PHILOSOPHICAL TRANSACTIONS
OF THE
ROYAL SOCIETY OF LONDON.
SERIES A.
VOL. 192, pp. 257–330.

9136

MATHEMATICAL CONTRIBUTIONS TO THE THEORY OF EVOLUTION.

—GENETIC (REPRODUCTIVE) SELECTION: INHERITANCE OF FERTILITY IN MAN, AND OF FECUNDITY IN THOROUGHBRED RACEHORSES.

BY
KARL PEARSON, F.R.S., ALICE LEE, B.A., B.Sc.,
AND
LESLIE BRAMLEY-MOORE.

LONDON:
PUBLISHED FOR THE ROYAL SOCIETY
BY DULAU AND CO., 37, SOHO SQUARE, W.
CONTINENTAL AGENTS; MESSRS. FRIEDLÄNDER AND SON, BERLIN.
1899.

32.

29.3.99.

Price Three Shillings and Sixpence.

VI. *Mathematical Contributions to the Theory of Evolution.*—VI. *Genetic* (*Reproductive*[*]) *Selection: Inheritance of Fertility in Man, and of Fecundity in Thoroughbred Racehorses.*

By KARL PEARSON, *F.R.S.*, ALICE LEE, *B.A.*, *B.Sc.*, and LESLIE BRAMLEY-MOORE.

Received November 14,—Read December 8, 1898.

CONTENTS.

Part I.—*Theory of Genetic* (*Reproductive*) *Selection.*

Part II.—*On the Inheritance of Fertility in Mankind.*

Part III.—*On the Inheritance of Fecundity in Thoroughbred Brood-mares.*

---

[*] The name Reproductive Selection is retained here, although objection has been taken to it, because it has been used in other memoirs of this series. I owe Genetic Selection to Mr. F. GALTON.

Introductory.

I UNDERSTAND by a *factor of evolution* any source of progressive change in the constants—mean values, variabilities, correlations—which suffice to define an organ or character, or the interrelations of a group of organs or characters, at any stage in any form of life. To demonstrate the existence of such a factor we require to show more than the plausibility of its effectiveness, we need that a numerical measure of the changes in the organic constants shall be obtained from actual statistical data. These data must be of sufficient extent to render the numerical determinations large as compared with their probable errors.

In a "Note on Reproductive Selection," published in the 'Roy. Soc. Proc.,' vol. 59, p. 301, I have pointed out that if fertility be inherited or if it be correlated with any inherited character—those who are thoroughly conversant with the theory of correlation will recognise that these two things are not the same—then we have a source of progressive change, a *vera causa* of evolution. I then termed this factor of evolution *Reproductive Selection.* As the term has been objected to, I have adopted *Genetic Selection* as an alternative. I mean by this term the influence of different grades of reproductivity in producing change in the predominant type.

If there be two organs A and B both correlated with fertility, but not necessarily correlated with each other,* then genetic or reproductive selection may ultimately cause the predominance in the population of two groups, in which the organs A and B are widely different from their primitive types—'widely different,' because reproductive selection is a source of *progressive* change. Thus this form of selection can be a source, not only of change, but of differential change. As this differentiation is progressive, it may amount in time to that degree of divergence at which crossing between the two groups begins to be difficult or distasteful. We then reach in genetic or reproductive selection a source of the origin of species.

When I assert that genetic (reproductive) selection is a factor of evolution, I do not intend at present to dogmatise as to the amount it is playing or has played in evolution. I intend to isolate it so far as possible from all other factors, and then measure its intensity numerically. If this be sensible, then the demonstration that it is a factor is complete. How far it may be held in check by other factors—*e.g.*, natural or sexual selection—is a matter for further inquiry. If three forces, F_1, F_2, F_3 hold a system sensibly in equilibrium, then F_1 cannot be asserted to be non-effective because no progressive change is visible; its absence would soon bring to light its effectiveness.

The manner in which genetic (reproductive) selection is to some extent held in check will be clearer when my memoir on the influence of directed selection on

---

* If r_{ab} be the correlation of two organic characters A and B, and C be a third character, there is a considerable range of values of r_{ac} and r_{bc}, for which r_{ab} may be zero (*see* YULE, 'Roy. Soc. Proc.' vol. 60, p. 486).

variation and correlation is published. Meanwhile Mr. FILON and I have shown that even a random selection of one organ alters the whole system of correlated organs.* Hence genetic (reproductive) selection indirectly modifies not only organs A and B, but all correlated organs. These modifications must be consistent with the maintenance of stamina, physique and fitness to the environment, if the change is not to be counteracted by natural selection.

So far as man is concerned, I have shown† that in the case of civilised man, the selective death-rate—*i.e.*, natural selection—does not appear to counteract reproductive selection. A small element of the population produces the larger part of the following generation. I thus concluded that *if* fertility were inherited, reproductive selection was not only a factor of evolution, but in civilised man a very sensible factor, *i.e.*, an apparently incompletely balanced factor.

In the three years which have intervened since writing the essay just referred to, members of the Department of Applied Mathematics in University College, as well as other friends, have occupied their spare time in the collection of data as to fertility and fecundity in the cases of man and of the thoroughbred racehorse. About 16,000 extracts were made in the case of man, and more than 7000 in the case of thoroughbred racehorses. In the course of the work, which proved far more laborious than we had anticipated, many difficulties and pitfalls appeared. But as a general conclusion it seems certain that: *Both fertility and fecundity are inherited, and probably in the manner prescribed by the Law of Ancestral Heredity.*‡

The object of this memoir is to set forth the theory and data by aid of which this conclusion was reached. It will be seen that it completes the establishment of genetic or reproductive selection as a factor of evolution by determining the much disputed point as to whether fertility is or is not inherited.

I. *Theory of Genetic or Reproductive Selection.* By KARL PEARSON, F.R.S.

(1.) While the physical result of fertility in an individual is measurable, the quality of fertility or fecundity in an individual differs from other physical characters in that it does not allow of direct measurements except when the potentiality is exerted and the effects recorded. At present we are not able to measure any series of organs or characters in individuals and so ascertain their fertility or fecundity. At the same time there is little doubt that these characters are functions of the physical and measurable organs and characters of the body. Such organs and characters we have good ground for supposing to be inherited according to the Law

---

* "Contributions to the Theory of Evolution.—IV. On the Influence of Random Selection on Variation and Correlation," 'Phil. Trans.,' A, vol. 191, p. 234 *et seq.*

† "The Chances of Death and other Studies in Evolution. Reproductive Selection," vol. 1, p. 63.

‡ *See* 'Roy. Soc. Proc.,' vol. 62, p. 386.

of Ancestral Heredity. It seems therefore worth while to prove the following proposition :

Proposition I.—Any character not itself directly measurable, but a function of physically measurable characters and organs inherited according to the Law of Ancestral Heredity, will itself be inherited according to that law.

Thus if we assume intellectual and emotional characters to be ultimately a result of physical conformation, we may be fairly certain that although we know neither the organs of which they are a function, nor the nature of that function, still they will be inherited according to the same law as that which holds for physically measurable organs.

Let y be the character in a parent, and let it be an unknown function f of the unknown physical organs $x_1, x_2, x_3, \ldots x_m$, or let :

$$y = f(x_1, x_2, x_3, \ldots x_m) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i.)}.$$

Let Δy denote the deviation from the mean value of the character y in some special individual, and Δx the deviation from the mean of any x organ in the same individual. Then if these deviations be small compared with the mean values of the organs considered, we have from (i.) above :

$$\Delta y = a_1 \Delta x_1 + a_2 \Delta x_2 + a_3 \Delta x_3 + \ldots \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii.)},$$

where $a_1, a_2 \ldots$ are constants independent of the individual variations.

Let σ denote a standard deviation, ρ a coefficient of interorganic correlation, S a summation with regard to all individuals with character y dealt with, and let them be n in number. Then :

$$\begin{aligned} n\sigma_y^2 &= \mathrm{S}(\Delta y)^2 = \mathrm{S}(a_1 \Delta x_1 + a_2 \Delta x_2 + a_3 \Delta x_3 + \ldots)^2 \\ &= n(a_1^2 \sigma_{x_1}^2 + a_2^2 \sigma_{x_2}^2 + a_3 \sigma_{x_3}^2 + \ldots + 2a_1 a_2 \sigma_{x_1} \sigma_{x_2} \rho_{x_1 x_2} + 2a_1 a_3 \sigma_{x_1} \sigma_{x_3} \rho_{x_1 x_3} + \ldots); \end{aligned}$$

or

$$\sigma_y^2 = \Sigma(a_1^2 \sigma_{x_1}^2) + 2\Sigma(a_1 a_2 \sigma_{x_1} \sigma_{x_2} \rho_{x_1 x_2}) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iii.)},$$

where Σ denotes a summation through the group of m organs.

Let y' denote the character in an individual who is the offspring of the individual of character y, and $x'_1, x'_2, x'_3 \ldots$ the corresponding organs. Then, if we do not suppose the nature of the function f to have changed in a single generation, we have :

$$y' = f(x'_1, x'_2, x'_3 \ldots x'_m),$$

and

$$\Delta y' = a_1 \Delta x'_1 + a_2 \Delta x'_2 + a_3 \Delta x'_3 + \ldots \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv.)},$$

$$\sigma_{y'}^2 = \Sigma(a_1^2 \sigma_{x'_1}^2) + 2\Sigma(a_1 a_2 \sigma_{x'_1} \sigma_{x'_2} \rho_{x'_1 x'_2}) \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(v.)}.$$

Let r be a coefficient of direct heredity expressing the correlation between parent and offspring, and according to the Law of Ancestral Heredity the same for all organs. Then multiplying (ii.) and (iv.) together and summing we have :

$$n\sigma_y\sigma_{y'}\mathrm{R} = \mathrm{S}(\Delta y\Delta y') = \Sigma(a_1^2\mathrm{S}(\Delta x_1\Delta x'_1)) + \Sigma(a_1a_2\mathrm{S}(\Delta x_1\Delta x'_2 + \Delta x_2\Delta x'_1)),$$

where R is the coefficient of correlation between the characters y and y' in parent and offspring. Now :

$$\mathrm{S}(\Delta x_1\Delta x'_1) = n\sigma_{x_1}\sigma_{x'_1}r$$

$$\mathrm{S}(\Delta x_1\Delta x'_2 + \Delta x_2\Delta x'_1) = n\sigma_{x_1}\sigma_{x'_2}r_{x_1x'_2} + n\sigma_{x'_1}\sigma_{x_2}r_{x'_1x_2},$$

where $r_{x_1x'_2}$ and $r_{x'_1x_2}$ are what I have elsewhere termed coefficients of cross-heredity. Now if the race be stable or sensibly stable for two generations we shall have for all organs $\sigma_{x'} = \sigma_x$. Hence :

$$\mathrm{S}(\Delta x_1\Delta x'_1) = n\sigma_{x_1}^2 \times r$$

$$\mathrm{S}(\Delta x_1\Delta x'_2 + \Delta x_2\Delta x'_1) = n\sigma_{x_1}\sigma_{x_2}(r_{x_1x'_2} + r_{x'_1x_2}) = n\sigma_{x_1}\sigma_{x_2} \times 2r\rho_{x_1x_2},$$

for it is shown in my memoir on the Law of Ancestral Heredity* that on a probable hypothesis :

$$\tfrac{1}{2}(r_{x_1x'_2} + r_{x'_1x_2}) = r \times \rho_{x_1x_2}.$$

Thus we find on substitution :

$$\sigma_y\sigma_{y'}\mathrm{R} = r(\Sigma(a_1^2\sigma_{x_1}^2) + 2\Sigma(a_1a_2\sigma_{x_1}\sigma_{x_2}\rho_{x_1x_2})).$$

But (iii.) and (iv.) show us that $\sigma_y = \sigma_{y'}$, if there be no sensible changes in a generation. Hence :

$$\sigma_y\sigma_{y'} = (\Sigma(a_1^2\sigma_{x_1}^2) + 2\Sigma(a_1a_2\sigma_{x_1}\sigma_{x_2}\rho_{x_1x_2})),$$

and

$$\mathrm{R} = r.$$

Thus the character which is a function of physical organs is inherited at the same rate as those organs themselves.

As we may not unreasonably consider fertility and fecundity to be functions of physically measurable organs, even if we cannot specify which organs, we may, *à priori*, expect fertility and fecundity to be inherited characters.

(2.) *Proposition II.—To determine the numerical values of the changes in mean variation and correlation if fertility be inherited.*

Let us first define two terms which will be frequently used in the sequel.

(*a*.) The *fertility* of an individual shall be defined as the total number of actual offspring.

---

* 'Roy. Soc. Proc.,' vol. 62, p. 411. The hypothesis yet awaits an experimental verification. The need to use it prevents Proposition I. being self-evident.

(*b.*) The *fecundity* of an individual shall be defined as the ratio of the total number of actual offspring to the total number of offspring which might have come into existence under the circumstances.

These definitions are not intended to give precise statistical measures at this stage of our investigations. They are merely meant to convey a general sense of the words, which will be more precisely limited when they are applied to any given species. Fertility and fecundity, as we have thus defined them, leave out of account individual conditions and definite conditions of period, age and environment, which must be fully stated before numerical measures can be made in any special case. When the words are used in this theoretical section the reader must suppose the phrase, "under definite individual and environmental conditions," to be always inserted.

Let M_1 = the mean fertility of parents of one sex; M'_1 = the mean fertility of parents of one sex weighted with their fertility;* N_1 the number of parents considered in the first case, N'_1 the apparent number dealt with in the second case; let σ_1 and σ'_1 be the standard deviations in the two cases, and let x represent the fertility of an individual parent and z its frequency among N_1 parents. Let S denote summation for N_1 parents. Then, without any assumption as to the type of frequency, $N'_1 = S(\lambda xz) = \lambda M_1 N_1$, where λ is a constant such that λx is the weight of a parent of fertility x. This follows at once, since:

$$N_1 = S(z), \qquad M_1 = S(xz)/S(z).$$

Further,

$$M'_1 = S(\lambda x \times xz)/N'_1 = \frac{S(x^2 z)}{M_1 N_1}$$

$$= \frac{S\{(x - M_1)^2 z + 2M_1(xz) - M_1^2 z\}}{M_1 N_1},$$

$$= \frac{N_1\sigma_1^2 + 2M_1^2 N_1 - M_1^2 N_1}{M_1 N_1},$$

by the definition of standard-deviation. Hence, finally:

$$M'_1 = \frac{\sigma_1^2}{M_1} + M_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{i.}).$$

Further:

$$\sigma'^2_1 = \frac{S\{\lambda x(x - M'_1)^2 z\}}{N'_1} = \frac{S\{(x - M_1 + M_1)(x - M_1 + M_1 - M'_1)^2 z\}}{M_1 N_1}.$$

Hence, multiplying out, we find after some reductions:

$$\sigma'^2_1 = \sigma_1^2\left(1 - \frac{\sigma_1^2}{M_1^2}\right) + \frac{S\{(x - M_1)^3 z\}}{M_1 N_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{ii.}).$$

At first sight it might seem a comparatively easy matter to avoid weighting parents with their fertility, but practically it is almost impossible. For example, if records

---

* *i.e.*, if f be the fertility of a parent, each parent is repeated λf times, where λ is a constant.

are sought of the fertility of mothers in mankind, the women will appear under their husbands' names, and the labour of ascertaining whether two sisters have been included is enormous, when large numbers are dealt with. But if two or more sisters have been included, their mother has been weighted with her fertility, and when we seek the correlation between mother and daughter, it will be between mothers and daughters when weighted with fertility. But a still more serious difficulty arises from the fact that all records are themselves weighted records; the same number are not married from each family, hence we are more likely to find a member of a large family included than a member of a small. The large families, when we seek a record of two generations, are more likely to appear than small families. Precisely the same difficulty occurs when we are dealing with thoroughbred horses; a mare with large fertility is less likely to have all her offspring colts, or all her progeny sold abroad, some one or more will probably ultimately come to the stud, and thus mares of large fertility are, *à priori*, more likely to contribute to our fecundity correlation cards. We do not get over this difficulty by taking the mother and only one of her offspring. The record is still weighted with fertility. The practical verification of this lies in the experience that the fertility of mothers will always be found to be greater than that of daughters, although the fertility of the community may really be increasing; the weighting, of course, excludes sterility in the generation of mothers, but the mere exclusion of the sterile is far from accounting for the whole difference.

What we actually find from our records are M'_1 and σ'_1, but what we want for the problem of heredity are M_1 and σ_1. Equations (i.) and (ii.) do not suffice to determine these, because we cannot evaluate the third moment $S\{(x - M_1)^3 z\}$. We can hardly, even for a first approximation, assume it zero, for the standard-deviation, and therefore the individual variation is large as compared with the mean in the case of fertility, *i.e.*, the distribution is markedly skew.

Turning to offspring of the same sex as the parents, say: let M_2 be the mean fertility of offspring taking *one* only to one parent for the number N_1 of parents, supposing the parents not weighted with their fertility; let M'_2 be the mean in the same case when the parents are weighted with their fertility; and let M''_2 be the mean of *all* recorded offspring of the second generation. Let σ_2, σ'_2, σ''_2 be the standard deviations in the fertility of the offspring for the same three cases, and r, r', r'' be the corresponding coefficients of correlation between fertility in parent and in offspring. It seems to me that r is the coefficient which actually measures the real inheritance of fertility, but that in any correlation table that we can form we shall get r' or r''.

Let y be the fertility of any individual among the offspring, and x the fertility of the corresponding parent; let λx as before be the weighting of the parent, and $\lambda' x$ the number of offspring included in the record, λ' being supposed a constant.*

---

* I have been unable so far to find any sensible correlation between size of family and number married in man, but the point is worth a more elaborate investigation.

We have at once the following results for the total numbers dealt with in each case:

$$N_1 = S(z), \qquad N'_1 = S(\lambda x z) = \lambda M_1 N_1,$$

$$N''_1 = S(\lambda x \lambda' x z) = \lambda\lambda' S(x^2 z) = \lambda\lambda'(\sigma_1^2 + M_1^2) N_1 \quad . \quad . \quad . \quad . \quad . \quad \text{(iii.)}.$$

Turning to the means:

$$M_2 = S(yz)/N_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv.)}.$$

$$M'_2 = S(\lambda x y z)/N'_1 = [S\{(x - M_1)(y - M_2) z\} + M_1 M_2 S(z)]/M_1 N_1 = M_2 + r\frac{\sigma_1\sigma_2}{M_1} \quad \text{(v.)}.$$

$$M''_2 = S(\lambda x \lambda' x y z)/N''_1 = M_2 + \frac{2M_1\sigma_1\sigma_2 r}{\sigma_1^2 + M_1^2} + \frac{S\{(x - M_1)^2 (y - M_2) z\}}{N_1(\sigma_1^2 + M_1^2)}$$

$$= M_2 + r\frac{\sigma_1\sigma_2}{M_1} + r\frac{\sigma_1\sigma_2}{M_1}\frac{M_1^2 - \sigma_1^2}{\sigma_1^2 + M_1^2} + \frac{S\{(x - M_1)^2 (y - M_2) z\}}{N_1(\sigma_1^2 + M_1^2)}$$

after some reductions. Now make use of (ii.) and we have:

$$M''_2 = M_2 + r\frac{\sigma_1\sigma_2}{M_1} + r\frac{\sigma_1\sigma_2}{M_1}\frac{\sigma'^2_1/\sigma_1^2}{1 + \sigma_1^2/M_1^2} + \frac{S\left\{(x - M_1)^2\left((y - M_2) - r\frac{\sigma_2}{\sigma_1}(x - M_1)\right) z\right\}}{(1 + \sigma_1^2/M_1^2) M_1^2 N_1} \quad \text{(vi.)}.$$

But for normal correlation the equation to the straight line of regression is:

$$y - M_2 = r\frac{\sigma_2}{\sigma_1}(x - M_1).$$

Hence for such correlation the mean value of $y - M_2$ for parents $x - M_1$ is equal to $r\frac{\sigma_2}{\sigma_1}(x - M_1)$ and the summation term would vanish. For skew correlation, Mr. YULE has shown that the line just given is the line of closest fit to the curve of regression. Hence even in the case of fertility, where the correlation is certainly skew, the summation term must be extremely small, or even zero. It follows, therefore, that we may write:

$$M''_2 = M_2 + r\frac{\sigma_1\sigma_2}{M_1}\left(1 + \frac{\sigma'^2_1/\sigma_1^2}{1 + \sigma_1^2/M_1^2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(vii.)}.$$

There is still another mean which ought to be found, namely, that of parents, M''_1, when all their recorded offspring have been entered on the correlation table. We have:

$$M''_1 = S(\lambda x \lambda' x x z)/N''_1 = S(x^3 z)/\{N_1(\sigma_1^2 + M_1^2)\},$$

or, after some reductions:

$$M''_1 = M_1 + \frac{\sigma_1^2}{M_1}\left(1 + \frac{\sigma'^2_1/\sigma_1^2}{1 + \sigma_1^2/M_1^2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(viii.)}.$$

I now proceed to the standard deviations for the three cases, and the additional case (σ''_1) for parents.

$$\sigma_2^2 = \mathrm{S}\{(y - \mathrm{M}_2)^2 z\}/\mathrm{N}'_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(ix.)}.$$

$$\begin{aligned}\sigma'^2_2 &= \mathrm{S}\{\lambda x (y - \mathrm{M}'_2)^2 z\}/\mathrm{N}'_1 \\ &= \mathrm{S}\{x (y - \mathrm{M}_2 + \mathrm{M}_2 - \mathrm{M}'_2)^2 z\}/\mathrm{M}_1\mathrm{N}_1 \\ &= \frac{\mathrm{S}\{x (y - \mathrm{M}_2)^2 z\} + 2\mathrm{S}\{x (y - \mathrm{M}_2) z\} (\mathrm{M}_2 - \mathrm{M}'_2) + \mathrm{M}_1 (\mathrm{M}_2 - \mathrm{M}'_2)^2 \mathrm{N}_1}{\mathrm{M}_1\mathrm{N}_1}\end{aligned}$$

Whence, after some reductions, we find :

$$\sigma'^2_2 = \sigma_2^2 \left\{ 1 + r^2 \left(\frac{\sigma'^2_1}{\sigma_1^2} - 1 \right) \right\} + \frac{\mathrm{S}\left\{ (x - \mathrm{M}_1) \left((y - \mathrm{M}_2)^2 - r^2 \frac{\sigma_2^2}{\sigma_1^2} (x - \mathrm{M}_1)^2 \right) z \right\}}{\mathrm{M}_1\mathrm{N}_1} \quad . \qquad \text{(x.)}.$$

Now for a nearly straight line of regression :

$$y - \mathrm{M}_2 = r \frac{\sigma_2}{\sigma_1} (x - \mathrm{M}_1) + \eta$$

where η is uncorrelated with $x - \mathrm{M}_1$. It follows accordingly that $\mathrm{S}\{(x - \mathrm{M}_1)^2 \eta z\}$ and $\mathrm{S}\{(x - \mathrm{M}_1)\eta^2 z\}$ will both vanish, since $\mathrm{S}(\eta)$ for an array and $\mathrm{S}(x - \mathrm{M}_1)$ for the whole correlation surface will be zero. Hence the summation term in (x.) is either absolutely zero or extremely small. We have accordingly :

$$\sigma'^2_2 = \sigma_2^2 \left\{ 1 + r^2 \left(\frac{\sigma'^2_1}{\sigma_1^2} - 1 \right) \right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xi.)}.$$

Before we proceed to determine σ''_2 and σ''_1 it seems simplest to find the coefficients of correlation r_1 r' and r''. We have :

$$r = \mathrm{S}\{ (x - \mathrm{M}_1) (y - \mathrm{M}_2) z\}/(\mathrm{N}_1\sigma_1\sigma_2) \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xii.)}.$$

To find r' we have :

$$r' = \mathrm{S}\{\lambda x z (x - \mathrm{M}'_1) (y - \mathrm{M}'_2)\}/(\mathrm{N}'_1\sigma'_1\sigma'_2).$$

Now

$$y - \mathrm{M}_2 = \frac{r\sigma_2}{\sigma_1} (x - \mathrm{M}_1) + \eta,$$

where η is sensibly un-correlated with $x - \mathrm{M}_1$. Hence :

$$\mathrm{N}'_1\sigma'_1\sigma'_2 r' = \mathrm{S}\left\{ \lambda x z (x - \mathrm{M}'_1) \left(r \frac{\sigma_2}{\sigma_1} (x - \mathrm{M}_1) + \mathrm{M}_2 - \mathrm{M}'_2 + \eta \right) \right\}.$$

Expanding, the summations with η vanish, and

$$M_2 - M'_2 - r\frac{\sigma_2}{\sigma_1}M_1 = -\frac{r\sigma_2}{\sigma_1}\left(\frac{\sigma_1^2}{M_1} + M_1\right) \text{ by (v.)}$$

$$= -r\frac{\sigma_2}{\sigma_1}M'_1 \text{ by (i.)}$$

But

$$\sigma'^2_1 = S\{\lambda xz(x - M'_1)^2\}/N'_1,$$

thus :

$$N'_1\sigma'_1\sigma'_2 r' = S\left\{\lambda xz(x - M'_1)\frac{r\sigma_2}{\sigma_1}(x - M'_1)\right\}$$

$$= \frac{r\sigma_2}{\sigma_1}S\{\lambda xz(x - M'_1)^2\}$$

$$= \frac{r\sigma_2}{\sigma_1}N'_1\sigma'^2_1.$$

Thus we deduce :

$$r' = r\frac{\sigma_2}{\sigma_1}\frac{\sigma'_1}{\sigma'_2}$$

or :

$$r'\sigma'_2/\sigma'_1 = r\sigma_2/\sigma_1 \quad . \; . \; . \; . \; . \; . \; . \; . \; . \quad \text{(xiii.)}.$$

This result has the simple interpretation *that while the coefficient of correlation is changed, the coefficient of regression is unchanged by weighting fertility, or by reproductive selection.*

This important conclusion is only an illustration of a very interesting theorem, which has been referred to in another memoir* and will be proved generally in a memoir on directed selection, written but not yet published, *i.e.*, that in a wide range of cases selection, whether random or directed (natural and artificial) changes correlation but not regression.

Before proceeding further a general remark will enable us to considerably simplify the otherwise lengthy algebra. Namely, the relation of M''_1, M''_2, σ''_1, σ''_2, r'' to M'_1, M'_2, σ'_1, σ'_2, r' is precisely the same as that of M'_1, M'_2, σ'_1, σ'_2, r' themselves to M_1, M_2, σ_1, σ_2, r. Consequently an interchange of symbols in results already found will lead us to the remaining formulæ needful.

As an illustration of this, let us verify the result we have found for M''_2. By an interchange in (v.) :

$$M''_1 = M'_2 + r'\frac{\sigma'_1\sigma'_2}{M'_1},$$

hence using (v.), (i.) and (xiii.), we find :

$$M''_2 = M_2 + \frac{r\sigma_1\sigma_2}{M_1} + r\sigma_2\frac{\sigma'^2_1}{\sigma_1}\frac{1}{M_1 + \dfrac{\sigma_1^2}{M_1}}$$

$$= M_2 + \frac{r\sigma_1\sigma_2}{M_1}\left(1 + \frac{\sigma'^2_1/\sigma_1^2}{1 + \sigma_1^2/M_1^2}\right),$$

---

* "Contributions to the Theory of Evolution.—V. On the Reconstruction of Stature," 'Phil. Trans.,' A, vol. 192, p. 177.

exactly the result reached by a longer process in (vii.). Similarly (viii.) may be deduced from (i.). Applying this to find r'' we have from (xiii.):

$$r''\sigma''_2/\sigma''_1 = r'\sigma'_2/\sigma'_1, \text{ and therefore } = r\sigma_2/\sigma_1 \quad . \quad . \quad . \quad . \quad \text{(xiv.)},$$

a result which again extends the constancy of the regression coefficient under the action of reproductive selection.

Next from (xi.):

$$\begin{aligned}\sigma''^2_2 &= \sigma'^2_2 \left\{1 + r'^2\left(\frac{\sigma''^2_1}{\sigma'^2_1} - 1\right)\right\} \\ &= \sigma^2_2\left\{1 + r^2\left(\frac{\sigma'^2_1}{\sigma^2_1} - 1\right)\right\} + r'^2\sigma'^2_2\left(\frac{\sigma''^2_1}{\sigma'^2_1} - 1\right),\end{aligned}$$

or using (xiii.) and rearranging:

$$\sigma''^2_2 = \sigma^2_2\left\{1 + r^2\left(\frac{\sigma''^2_1}{\sigma^2_1} - 1\right)\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xv.)}.$$

Again by interchanges in (ii.):

$$\sigma''^2_1 = \sigma'^2_1\left(1 - \frac{\sigma'^2_1}{\text{M}'^2_1}\right) + \frac{\text{S}\{(x - \text{M}'_1)^3 z'\}}{\text{M}'_1\text{N}'_1} \quad . \quad . \quad . \quad . \quad . \quad \text{(xvi.)}.$$

Here z' stands for λxz, and we should obtain a fourth moment of the original system of unweighted parents by substitution. But it is practically impossible to obtain a correlation table for such a system. Thus it is better to allow the summation term to stand as it is, where it represents the third moment of a system of parents, weighted for fertility owing to the nature of the record, but not weighted with all their recorded offspring. (xvi.) is then a relation between the standard-deviations of parents weighted solely by forming a record and weighted both by this and by their offspring.

Equations (i.) to (xvi.) contain the chief theoretical relations of our subject,[*] and I shall consider some points with regard to them in the following section.

(3.) (*a.*) If we wish to ascertain whether fertility is inherited, we have to discover whether r is or is not zero. Now by (xiv.) r vanishes with both r' and r'', and accordingly either of these will suffice to answer the problem. Still better, we may ascertain the coefficient of regression, and then whether our statistics weight for progeny or not we shall obtain the same value. If there be no secular change taking place in the population, due to something else than reproductive selection, we should expect, provided the Law of Ancestral Heredity holds for fertility, that the regression will be near ·3 for parent and offspring.[†]

---

[*] Two of these formulæ, (v.) and (xi.), were given, but in a less precisely defined manner, in my "Note on Reproductive Selection" of 1896, 'Roy. Soc. Proc.,' vol. 59, p. 303.

[†] See "Law of Ancestral Heredity," 'Roy. Soc. Proc.,' vol. 62, p. 397.

(*b.*) If no reproductive selection exists, *i.e.*, if fertility be not inherited, then $r = 0$, and

$$\sigma''_2 = \sigma'_2 = \sigma_2, \quad M_2 = M'_2 = M''_2,$$

or, however we form a record of offspring, the mean value and variability of their fertility ought not to be changed. We shall see later that this is very far from the truth, and that these values are in whole or part sensibly affected by the manner in which the record is formed.

(*c.*) Although there be no reproductive selection, M_1, M'_1, and M''_1 will not all be equal, it is impossible that they should be. Further, σ_1, σ'_1 and σ''_1 need not be equal; their degree of sensible divergence will depend on the nature of the primitive frequency distribution for parents.

(*d.*) If fertility be inherited, or reproductive selection be an actual factor of evolution, then we see, by comparing (v.) with (i.) and (vii.) with (viii.), that the mean fertility of mothers will always be *apparently* greater than the mean fertility of daughters. This follows, since r is always less than unity, and if the race be not subjected to secular evolution, other than that due to reproductive selection, σ_2 cannot differ very widely from σ_1.*

(*e.*) An argument from means, as to whether fertility is inherited or not, is very likely to be misleading. We may choose two groups from the record for comparison, neglecting the fact that their frequency in the record is not necessarily that of their frequency in the general population. Thus, if one person, say, in four were married, a marriage record of the community might exhibit the proper frequency of families of four, but it would not do so of families of one. The sort of fallacious arguments we have to be prepared for are, for example:

(i.) That the fertility of the community is diminishing, because M'_2 is less than M'_1.

(ii.) That the fertility of the community is increasing, because M'_2 might be $> M_1$ or M''_2 be $> M'_1$.

(iii.) That fertility is not inherited, because, owing to natural selection, or other factor of evolution, one or other of these means for offspring is sensibly equal to one or other of these means for parents.

Owing to the extreme difficulty of insuring that the method of extracting the record really gives us definitely M'_2, say, and not M''_2 (or M''_2 in part), I have discarded all use of the mean values in attempting to ascertain whether fertility is inherited. The following result, however, is tempting, and might possibly be made

---

* A difference between σ_1 and σ_2 would mark natural selection, sexual selection, or some other factor of secular evolution at work; of *secular*, not periodic, evolution, as parents and offspring must have reached the same adult stage to have had their fertility measured.

use of in direct experiments on breeding insects, where a record could be kept *ad hoc.* It follows at once from (i.), (v.), (vii.) and (viii.):

$$\frac{M'_2 - M_2}{M'_1 - M_1} = \frac{M''_2 - M'_2}{M''_1 - M'_1} = r\frac{\sigma_2}{\sigma_1} = \text{coefficient of regression} \quad . \quad . \qquad \text{(xvii.)}.$$

It is the second ratio which, I think, might with profit be experimentally evaluated.

(*f.*) Since the mean fertility of daughters loaded with the fertility of their mothers is the fertility of the next generation, and we see that this is always greater than M_2, if r be not zero, it follows that the inheritance of fertility marks a progressive change. The only means of counteracting its influence would be the reduction of M_2 to or below M_1 by the action of other equally potent factors of evolution. For the existence of such factors in man I shall later give evidence.

(4.) *Proposition III.—To extend the results obtained for fertility to the problem of fecundity.*

While the fecundity of an individual can often, at any rate approximately, be measured, the fertility is not ascertainable. Thus we can ascertain the number of occasions on which a brood mare has gone to the stallion and the number of foals she has produced, but her fertility, the produce she might have had, if she had throughout her whole career had every facility for breeding, is unknown to us. But if we proceed to form tables for the inheritance of fecundity, we are met by precisely the same difficulties as in the case of fertility. The more fertile individuals are *à priori* more likely to appear in the record, and will be likely to be weighted again with their fertility when we come to deal with their offspring.*

Now it is certain that fertility must be correlated with fecundity; or, if x now represents the fecundity and f the fertility, we shall have for the *mean* fertility for a given fecundity x an expression of the form $\lambda_0 + \lambda_1 x$, always supposing the regression to be sensibly linear. But the fertility must vanish with the fecundity, hence $\lambda_0 = 0$, and λ_1 is really the ratio of mean fertility to mean fecundity. Thus we may write for the fertility f

$$f = \lambda_1 x + \zeta,$$

where ζ may vary widely, but it is not correlated with x.

If now all the symbols we have used with regard to fertility in Section (2) be interpreted as referring to fecundity, we must weight with a factor λf instead of a factor λx, or with a factor $\lambda\lambda_1 x + \lambda\zeta$. So long as this factor is linear, absolutely no change can be made in the results, for, ζ being uncorrelated with x, all summations including

---

* In the case of sires especially, if we are dealing with thoroughbred horses, their comparative fewness at each period renders it quite impossible to deal with one offspring of each parent only.

$S(\zeta)$ vanish. Thus all the values given for M'_1, M'_2, σ'_1, and σ'_2 remain the same, if their results be interpreted in the sense of fecundity and not fertility. If ρ be the correlation between fecundity and fertility, and σ_1, σ_3 the standard deviations of these quantities, then $\lambda_1 = \rho\sigma_3/\sigma_1$; but we have seen that it is also the ratio of mean fertility to mean fecundity. It follows accordingly that ρ is the ratio of the coefficient of variation in fecundity to the coefficient of variation in fertility. If we may judge by the cases of man and horse, so far as I know the only cases in which fertility and fecundity have yet been examined, a coefficient of variation in fecundity amounts to about 30 per cent., while one in fertility is something like 50 per cent. Thus the correlation of fertility with fecundity would be about ·6. We should expect it to have a high value, perhaps even a higher value than this. In the case of thoroughbred horses, ρ will be the correlation between fecundity and *apparent* fertility. By direct investigation in the case of 1000 brood mares I find its value to be ·5152.

Passing now to the correlations r, r', r'', I observe that the proof given for fertility is valid with but few modifications, if these be fecundity correlations (see p. 266), for the proof involves no expansion of the factor $(\lambda_1 x + \zeta)^2$. Hence we conclude that the regression coefficient for the inheritance of fecundity will not be modified by the nature of the record or the weighting of individuals with their fertility.

When we come to the last series of constants, M''_1, M''_2, σ''_1, σ''_2, we find that these will be modified, owing to the presence of the square factor $(\lambda_1 x + \zeta)^2$, although ζ is not correlated with x. The term ζ^2 now comes in, and $S(\zeta^2)$ will give the standard-deviation of an array of fertilities corresponding to a given fecundity, *i.e.*, $S(\zeta^2) = \sigma_3^2(1 - \rho^2) \times$ number in the array.*

I find after some reductions that M''_2 and M''_1 are given by

$$M''_2 = M_2 + r\frac{\sigma_1\sigma_2}{M_1}\left(1 + \frac{\sigma'^2_1/\sigma^2_1}{1 + \sigma^2_1/(\rho^2 M^2_1)}\right) \quad . \quad . \quad . \quad . \qquad \text{(xviii.)},$$

$$M''_1 = M_1 + \frac{\sigma^2_1}{M_1}\left(1 + \frac{\sigma'^2_1/\sigma^2_1}{1 + \sigma^2_1/(\rho^2 M^2_1)}\right) \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xix.)},$$

the correlation of fertility and fecundity being now introduced into the results.

Clearly the result (xvii.)

$$\frac{M''_2 - M'_2}{M''_1 - M'_1} = \text{coefficient of regression} \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xx.)}$$

still remains true.

For the remaining two constants σ''_2 and σ''_1, I find, after some rather long analysis in the second case, which it seems unnecessary to reproduce,†

---

* Should the regression not be linear, $\sigma_3\sqrt{(1 - \rho^2)}$ is the mean of the standard-deviations of the arrays.

† In the course of the work the squared standard-deviation of a fertility array is assumed to be the same for all arrays $= \sigma_3^2(1 - \rho^2)$, and λ_1 is given its value $\rho\sigma_3/\sigma_1$. See, however, the previous footnote.

$$\sigma''^2_2 = \sigma^2_2 \left\{1 + r^2 \left(\frac{\sigma''^2_1}{\sigma^2_1} - 1\right)\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xxi.)},$$

$$\sigma''^2_1 = \gamma\sigma'^2_1 \left(1 - \frac{\gamma\sigma'^2_1}{M'^2_1}\right) + \gamma \frac{1 - \rho^2}{\rho^2} \frac{\sigma^4_1}{M^2_1} + \frac{S\{(x - M'_1)^3 Z'\}}{N'_1 M'_1} \quad . \quad . \quad . \qquad \text{(xxii.)},$$

and γ is the factor $\frac{M^2_1 + \sigma^2_1}{M^2_1 + \sigma^2_1/\rho^2}$, or as we can write it

$$\gamma = M'_1 \Big/ \left(M'_1 + \frac{\sigma^2_1}{M_1} \frac{1 - \rho^2}{\rho^2}\right).$$

If ρ be unity or near unity, *i.e.*, fecundity very closely correlated with fertility, $\gamma = 1$, the second term vanishes and (xxii.) becomes identical with the corresponding fertility formula (xvi.), just as (xxi.) is already identical with (xv.).

Thus we see that the whole series of fecundity relations are strikingly like those for fertility, except that in certain of them—those for M''_1, M''_2, σ''_1 and σ''_2—the correlation ρ of fertility and fecundity is introduced. If ρ be considerable, all the remarks we have made on the fertility formulæ may, *mutatis mutandis*, be applied to the measurement of fecundity.

(5.) *Proposition IV.—To deduce formulæ for finding the correlation between any grades of kindred from the means of arrays into which the kindred may be grouped.*

This problem is of very great practical importance. In the case of Man, families are so small that there is comparatively small difficulty in forming all the possible pairs of brethren, say, for any family; but when we come to animals or insects where the fertility may be extremely large, it is practically impossible to form a correlation table involving 50,000 to 100,000 entries.* One thoroughbred sire may have 50 to 80 daughters, and thus give us roughly 1200 to 3200 pairs of sisters to be entered in a correlation table. Still higher results occur in the case of aunts and nieces. It may be asked why we do not content ourselves with one or two pairs from each parent; the answer is simple: we have not (*e.g.*, in the case of thoroughbred animals, pedigree moths, &c.) a great number of sires, and the sire with 50 offspring cannot, for accuracy of result, be put on the same footing as the sire with only 2 to 4. Our process is really an indirect weighting of our results.

(A.) *To find the coefficient of correlation between brethren from the means of the arrays.*

Let x be the measure of any character or organ in one brother (sister), and x' that of a second brother (sister): let m be the mean of one set of brothers, and m' of the

---

* Even with the reduction in labour, introduced by this proposition and by the use of mechanical calculators, Mr. LESLIE BRAMLEY-MOORE and I took practically a week, of eight-hour days, to deduce two coefficients of correlation, *after* the means of the arrays had already been found.

second set. Let n be the number of brothers in an array, and therefore $\frac{1}{2}n(n-1)$ the number of pairs of brothers in the array. Let σ and σ' be the standard deviations of the two sets of brothers, and r the coefficient of correlation between brothers for the organ in question. Let S denote a summation with regard to all pairs of brothers in the community, and Σ with regard to all brothers in an array. Let N be the total number of brothers in the community. Then if we selected our pairs of brothers for tabulation at random (*e.g.*, not by seniority or other character), we should find $m' = m$ and $\sigma' = \sigma$. Further, by definition of correlation

$$\mathrm{N}r\sigma\sigma' = \mathrm{S}\,(x-m)(x'-m') = \mathrm{S}\Sigma\,(x-\mathrm{M}+\mathrm{M}-m)(x'-\mathrm{M}'+\mathrm{M}'-m'),$$

where M and M′ are the means of the two sets of brothers in any array and are clearly equal.

Further, $\Sigma\,(x-\mathrm{M}) = \Sigma\,(x'-\mathrm{M}') = 0$, when summed for an array, and $\Sigma\,(x-\mathrm{M})(x'-\mathrm{M}') = 0$, for there is no correlation within the array when the deviations are measured from the mean of the array. Hence :

$$\mathrm{N}r\sigma\sigma' = \mathrm{S}\,\{\tfrac{1}{2}n(n-1)(\mathrm{M}-m)(\mathrm{M}'-m')\},$$

or

$$\mathrm{N}r\sigma^2 = \mathrm{S}\,\{\tfrac{1}{2}n(n-1)\,\mathrm{M}^2\} - 2m\mathrm{S}\,\{\tfrac{1}{2}n(n-1)\,\mathrm{M}\} + m^2\mathrm{N}\,;$$

but

$$\mathrm{S}\,\{\tfrac{1}{2}n(n-1)\,\mathrm{M}\} = \mathrm{N}m.$$

Thus, finally,

$$r = \frac{\mathrm{S}\,\{\frac{1}{2}n(n-1)\,\mathrm{M}^2\}/\mathrm{N} - m^2}{\sigma^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xxiii.)}.$$

This can be written

$$r = \sigma_a^2/\sigma^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xxiv.)}$$

where σ_a is the standard deviation of the arrays concentrated into their means and loaded with their sizes ; σ is the standard deviation of all brethren loaded with the number of times they are counted as brethren ; m is the mean of all the offspring loaded with the number of times they are counted as brethren.

Let σ_0 be the standard deviation of offspring, and ρ the correlation between parent and offspring, then the standard deviation of an array of offspring, if correlation be sensibly linear,[*] will be $\sigma_0\sqrt{(1-\rho^2)}$. We have, further,

$$m = \mathrm{S}\,(x) = \mathrm{S}\Sigma\,(x-\mathrm{M}+\mathrm{M}) = \mathrm{S}\,\{\tfrac{1}{2}n(n-1)\,\mathrm{M}\},$$

$$\mathrm{N}\sigma^2 = \mathrm{S}\,(x-m)^2 = \mathrm{S}\Sigma\,(x-\mathrm{M}+\mathrm{M}-m)^2 = \mathrm{S}\,\{\Sigma\,(x-\mathrm{M})^2 + \tfrac{1}{2}n(n-1)(\mathrm{M}-m)^2\}.$$

But

$$\Sigma\,(x-\mathrm{M})^2 = \tfrac{1}{2}n(n-1)\,\sigma_0^2(1-\rho^2).$$

---

[*] See, however, the first footnote p. 270.

Thus :

$$N\sigma^2 = N\sigma_0^2(1-\rho^2) + N\sigma_a^2,$$

or :

$$\sigma^2 = \sigma_0^2(1-\rho^2) + \sigma_a^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xxv.}),$$

and r may be written :

$$r = \frac{\sigma_a^2}{\sigma_0^2(1-\rho^2) + \sigma_a^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xxvi.}).$$

Here σ_a can be found from the arrays, and σ_0 and ρ will in many cases have been previously ascertained.

(B.) *To find the correlation between "uncles" and "nephews" ("aunts" and "nieces") from the means of the corresponding arrays.*

Let n_1 be the number of uncles in an array, n_2 be the number of nephews in the associated array, so that n_1n_2 is the number of pairs of uncles and nephews provided by the associated arrays. Let $N = S(n_1n_2)$ be the total number of pairs of uncles and nephews in the community under consideration. Let x be the measure of the organ or character in the uncle, x' in the nephew. Let M and M′ be the means of two associated arrays of uncles and nephews respectively. Let m and m' be the means of all uncles weighted with their nephews and all nephews weighted with their uncles respectively, and let $\bar{\sigma}$, $\bar{\sigma}'$ be the corresponding standard deviations under the same circumstances ; r' the correlation of uncle and nephew. Then :

$$Nr'\sigma\sigma' = S(x-m)(x'-m') = S\Sigma(x - M + M - m)(x' - M' + M' - m').$$

Now $\Sigma(x-M) = \Sigma(x'-M') = 0$, and within the arrays there is no association of individual uncles with individual nephews, *i.e.*, $\Sigma(x-M)(x'-M') = 0$. Thus :

$$Nr'\sigma\sigma' = S\{n_1n_2(M-m)(M'-m')\} = S(n_1n_2MM') - Nmm',$$

since

$$m = S(n_1n_2M)/N, \qquad m' = S(n_1n_2M')/N.$$

Thus :

$$r' = \frac{S(n_1n_2MM')/N - mm'}{\bar{\sigma}\bar{\sigma}'} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xxvii.}).$$

If $\bar{\sigma}_a$ and $\bar{\sigma}'_a$ be the standard deviations of the means of the arrays of uncles and nephews and R the correlation of these means, the numerator is clearly $R\bar{\sigma}_a\bar{\sigma}'_a$. Thus :

$$r' = R\frac{\bar{\sigma}_a\bar{\sigma}'_a}{\bar{\sigma}\bar{\sigma}'} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xxviii.}).$$

Here the numerator as a whole or in parts is easily found from the means of the arrays. If $\bar{\sigma}_0$ and $\bar{\sigma}'_0$ be the means of unloaded uncles and nephews, we note that they are arrays owing to common parentage, and hence their array standard deviations* will be $\bar{\sigma}_0\sqrt{1-\rho^2}$ and $\bar{\sigma}'_0\sqrt{1-\rho^2}$, ρ being the standard deviation of parent and offspring. As before we find:

$$\left.\begin{aligned}\bar{\sigma}^2 &= \bar{\sigma}_a^2 + \bar{\sigma}_0^2(1-\rho^2) \quad . \quad . \quad . \quad . \quad . \quad . \\ \bar{\sigma}'^2 &= \bar{\sigma}'^2_a + \bar{\sigma}'^2_0(1-\rho^2) \quad . \quad . \quad . \quad . \quad . \quad .\end{aligned}\right\} \text{(xxviii.)}.$$

If, as will probably be the case, there be no secular change between uncles and nephews, then $\bar{\sigma} = \bar{\sigma}'$, $\bar{\sigma}_a = \bar{\sigma}'_a$, $\bar{\sigma}_0 = \bar{\sigma}'_0$, and accordingly $r' = \mathrm{R}\bar{\sigma}_a^2/\bar{\sigma}^2$; whence, using (xxiv.), we have:

$$r' = r \times \mathrm{R} \times \frac{\bar{\sigma}_a^2}{\sigma_a^2}\,\frac{\sigma^2}{\bar{\sigma}^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxix.)}.$$

If we could assume $\sigma_a = \bar{\sigma}_a$ and $\sigma = \bar{\sigma}$, this result would reduce to the very simple form:

$$r' = r \times \mathrm{R}.$$

Now the assumption $\bar{\sigma}_0 = \sigma_0$ is, I think, legitimate, for the distribution for an unloaded array of nephews or uncles should be sensibly that of an array of brethren. But the equality of σ_a and $\bar{\sigma}_a$, which would now involve that of σ and $\bar{\sigma}$, is a much more doubtful point. σ_a and $\bar{\sigma}_a$ mark indeed quite different systems of loading. Both, it is true, are of the form

$$\mathrm{S}(nn'\mathrm{M}^2)/\mathrm{N} - \{\mathrm{S}(nn'\mathrm{M})/\mathrm{N}\}^2,$$

but in the case of brethren $n' = \frac{1}{2}(n-1)$ or n' has perfect correlation with n, while in the case of uncles and nephews n' is only imperfectly correlated with n. The intensity of this correlation depends upon the correlation between the sizes of arrays of uncles and nephews, a quantity which may be very small, or not, according to the nature of the record. Hence it appears necessary in applying the method to make some attempt to appreciate the value of $\bar{\sigma}_a$ as well as σ_a. If this be done R can be found from (xxix.), if not directly. This value of R is not without importance for the inheritance of characters latent in one or other sex.

We have thus reduced the correlations of individuals to a calculation of the correlation of arrays.

(6.) *Proposition V.—To find a measure of the effect of mingling uncorrelated material with correlated material.*

The importance of this investigation lies in the fact that death, restraint, or other

---

* Or, again, the means of the standard-deviations of the arrays.

circumstances, completely screen, in a certain number of cases, both the potential fertility and the real fecundity of man. Precisely similar circumstances, which will be considered more at length later, hinder our obtaining in horses a true measure of fecundity for all cases. We are thus really dealing with a mixture of correlated and apparently uncorrelated material. In what manner does the influence of this mixture effect our results?

Let a group N consist of $n_1 + n_2 + n_3 + n_4$ pairs of individuals. Of these, in the case of n_1 pairs, both individuals have the true value of the character under investigation recorded; in the case of n_2 pairs, neither have the true value recorded; in the case of n_3 pairs, it is the first individual of the pair which has a true recorded value, and the second an apparent or fictitious value; lastly, in n_4 cases, let the fictitious value be in the first and the real value in the second individual of the pair. Then there will be no correlation between individuals in the groups n_2, n_3, n_4. Let r be the correlation in the group n_1 and R that observed in the whole group of $\mathrm{N} = n_1 + n_2 + n_3 + n_4$. Let x be the measure of a character in the first, x' in the second individual. Let M and M′ be the means of the total groups of the two individuals and Σ, Σ' their standard deviations. In group n_1 let the corresponding quantities be m_1, m'_1, σ_1, σ'_1, and a similar notation hold for the other sub-groups. Then $m_1 = m_3$ and $\sigma_1 = \sigma_3$; $m_2 = m_4$ and $\sigma_2 = \sigma_4$; while $m'_1 = m'_4$ and $\sigma'_1 = \sigma'_4$; $m'_2 = m'_3$, $\sigma'_2 = \sigma'_3$.

We have at once:

$$\mathrm{M} = \frac{n_1 m_1 + n_2 m_2 + n_3 m_3 + n_4 m_4}{m_1 + m_2 + n_3 + n_4} = \frac{(n_1 + n_3)\, m_1 + (n_2 + n_4)\, m_2}{n_1 + n_2 + n_3 + n_4},$$

while

$$\mathrm{M}' = \frac{(n_1 + n_4)\, m'_1 + (n_2 + n_3)\, m'_2}{n_1 + n_2 + n_3 + n_4}.$$

Further:

$$(n_1 + n_2 + n_3 + n_4)\, \Sigma\Sigma'\mathrm{R} = \mathrm{S}\,(x - \mathrm{M})\,(x' - \mathrm{M}'),$$

by the usual properties of product moments

$$\begin{aligned} &= n_1\sigma_1\sigma'_1 r + n_1 (m_1 - \mathrm{M})(m'_1 - \mathrm{M}') + n_2 (m_2 - \mathrm{M})(m'_2 - \mathrm{M}') \\ &\qquad + n_3 (m_3 - \mathrm{M})(m'_3 - \mathrm{M}') + n_4 (m_4 - \mathrm{M})(m'_4 - \mathrm{M}') \\ &= n_1\sigma_1\sigma'_1 r + n_1 m_1 m'_1 + n_2 m_2 m'_2 + n_3 m_3 m'_3 + n_4 m_4 m'_4 \\ &\qquad - \mathrm{M}\,(n_1 m'_1 + n_2 m'_2 + n_3 m'_3 + n_4 m'_4) - \mathrm{M}'\,(n_1 m_1 + n_2 m_2 + n_3 m_3 + n_4 m_4) \\ &\qquad + \mathrm{MM}'\,(n_1 + n_2 + n_3 + n_4) \\ &= n_1\sigma_1\sigma'_1 r + n_1 m_1 m'_1 + n_2 m_2 m'_2 + n_3 m_3 m'_3 + n_4 m_4 m'_4 - \mathrm{MM}'\,(n_1 + n_2 + n_3 + n_4). \end{aligned}$$

Substituting the values of M and M′ and using the relations between the m's, we find after some reductions:

$$\mathrm{N}\Sigma\Sigma'\mathrm{R} = n_1\sigma_1\sigma'_1 r + \frac{n_3 n_4 - n_1 n_2}{n_1 + n_2 + n_3 + n_4}(m_1 - m_2)(m'_1 - m'_2) \quad . \quad . \quad \text{(xxx.)}$$

Let $\frac{1}{p}$ of the N first individuals and $\frac{1}{q}$ of the N second individuals have fictitious values, then $\frac{p-1}{p}\mathrm{N}$ and $\frac{q-1}{q}\mathrm{N}$ will have their true values. If, now, there is no correlation between the fictitious values in the two cases, we have at once:

$$n_1 = \frac{(p-1)(q-1)}{pq}\mathrm{N}, \qquad n_2 = \frac{1}{pq}\mathrm{N}, \qquad n_3 = \frac{p-1}{pq}\mathrm{N}, \qquad n_4 = \frac{q-1}{pq}\mathrm{N}.$$

From this it follows at once that

$$n_3 n_4 = n_1 n_2,$$

or the second term in (xxx.) vanishes. Thus:

$$\mathrm{R} = \frac{n_1}{\mathrm{N}}\,\frac{\sigma_1\sigma'_1 r}{\Sigma\Sigma'} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxx.) } bis.$$

Thus R vanishes with r, and no spurious correlation could arise from the existence of fictitious values distributed at random through the correlation table. This result might, indeed, (as it often is tacitly) be assumed by some, but it seems very desirable to have a definite proof.

It remains to consider Σ and Σ'. We have:

$$\begin{aligned}\mathrm{N}\Sigma^2 &= n_1\sigma_1^2 + n_2\sigma'^2_2 + n_3\sigma_3^2 + n_4\sigma_4^2 \\ &\qquad + n_1(m_1 - \mathrm{M})^2 + n_2(m_2 - \mathrm{M})^2 + n_3(m_3 - \mathrm{M})^2 + n_4(m_4 - \mathrm{M})^2 \\ &= (n_1 + n_3)\sigma_1^2 + (n_2 + n_4)\sigma_2^2 \\ &\qquad + (n_1 + n_3)m_1^2 + (n_2 + n_4)m_2^2 - (n_1 + n_2 + n_3 + n_4)\mathrm{M}^2,\end{aligned}$$

or

$$\begin{aligned}\Sigma^2 &= \frac{n_1 + n_3}{\mathrm{N}}\sigma_1^2 + \frac{n_2 + n_4}{\mathrm{N}}\sigma_2^2 + \frac{n_1 + n_3}{\mathrm{N}}\,\frac{n_2 + n_4}{\mathrm{N}}(m_1 - m_2)^2 \\ &= \left(1 - \frac{1}{p}\right)\sigma_1^2 + \frac{1}{p}\sigma_2^2 + \left(1 - \frac{1}{p}\right)\frac{1}{p}(m_1 - m_4)^2 \\ &= \sigma_1^2 + \frac{1}{p}(\sigma_2^2 - \sigma_1^2) + \left(1 - \frac{1}{p}\right)\frac{1}{p}(m_1 - m_4)^2 \quad . \quad . \quad . \quad . \quad \text{(xxxi.)}.\end{aligned}$$

Similarly:

$$\Sigma'^2 = \sigma'^2_1 + \frac{1}{q}(\sigma'^2_2 - \sigma'^2_1) + \left(1 - \frac{1}{q}\right)\frac{1}{q}(m'_1 - m'_2)^2 \quad . \quad . \quad \text{(xxxii.)}.$$

Now if the introduction of the fictitious values consisted of anything of the nature of a wrong pairing of certain individuals, we should simply have $\sigma_1 = \sigma_2$, $\sigma'_1 = \sigma'_2$, $m_1 = m_2$, $m'_1 = m'_2$ and, accordingly, $\Sigma = \sigma_1$, and $\Sigma' = \sigma'_1$.

In any case, if the percentage of fictitious values be not large, the second and third terms are of the second order of small quantities, since $\frac{1}{p}$ and $\frac{1}{q}$ are small. The maximum value of the third term cannot be greater than $\frac{1}{4}(m_1 - m_2)^2$, and this will be relatively small in the cases to which we shall apply it.

For example, no great changes are made in σ, when we vary the amount of fictitious cases introduced into our fertility tables. m_1 and m_2 do, however, change. Thus $\sigma_1 = \sigma_2 = 3$ approximately, and the range $m_1 - m_2 = 1{\cdot}2$. Hence:

$$\Sigma^2 = 9 + \tfrac{1}{4}(1{\cdot}2)^2, \text{ at a maximum, } = 9{\cdot}36,$$

or,

$$\Sigma = 3{\cdot}06.$$

Thus in this *extreme* case there is only 2 per cent. change in the value of Σ. In such cases accordingly we may take for rough approximations $\Sigma = \sigma$ and $\Sigma = \sigma'$. This leads us to:

$$R \doteq \frac{n_1}{N} r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xxxii.}).$$

Or, *the reduction of correlation, due to the introduction of fictitious values, is obtained by using as a factor the ratio of actual correlated pairs of individuals to the total number of pairs tabulated.*

This result will be of considerable service when we come to deal with the fecundity of thoroughbred racehorses.

(7.) *Proposition VI.—To obtain a measure of the spurious correlation apparently existing between two organs, when a mixture is made of heterogeneous materials.*

Let x and x' be measures of the two organs, and let there be N pairs of organs formed by i heterogeneous groups containing $n_1, n_2, n_3 \ldots$ pairs with means $m_1, m'_1, m_2, m'_2, m_3, m'_3 \ldots$, &c., standard deviations $\sigma_1, \sigma'_1, \sigma_2, \sigma'_2, \sigma_3, \sigma'_3 \ldots$, &c., and correlations $r_1, r_2, r_3 \ldots$, &c. Let M_1 M′ be the means of the whole heterogeneous community, Σ, Σ' the standard deviations, and R the correlation. Then:

$$R\Sigma\Sigma' N = S(n\sigma\sigma' r) + S\{n(m - M)(m' - M')\},$$

where S denotes a summation with regard to all i groups. Now if there were no correlation at all between the organs in any one of the i groups, R for the heterogeneous mixture would still not be zero so long as the second summation did not vanish. This, then, is a measure of the spurious correlation produced by making a mixture of uncorrelated materials.

Now $S\{n(m - M)(m' - M')\}$, remembering the values of M and M′ may be written:

$$S\left\{\frac{n_p n_q}{N}(m_p - m_q)(m'_p - m'_q)\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xxxiv.)}$$

where the summation S now refers to every possible pair p and q of the r groups.

Now it is very unlikely, unless i be very large and the numbers $n_1, n_2, n_3 \ldots$ be taken at random, that this expression will vanish. Suppose even that the means of our heterogeneous groups were uncorrelated, *i.e.*, $S(m - M)(m' - M') = 0$, it is unlikely that $S\{n(m - M)(m' - M')\}$ will also be zero, when n is taken at random. With a comparatively few groups, with numbers taken at random, it is extremely improbable that the principal axes of the i points loaded with $n_1, n_2, n_3 \ldots$ will exactly coincide with the directions of the axes of x and x'.

We are thus forced to the conclusion that a mixture of heterogeneous groups, each of which exhibits in itself no organic correlation, will exhibit a greater or less amount of correlation. This correlation may properly be called spurious, yet as it is almost impossible to guarantee the absolute homogeneity of any community, our results for correlation are always liable to an error, the amount of which cannot be foretold. To those who persist in looking upon all correlation as cause and effect, the fact that correlation can be produced between two quite uncorrelated characters A and B by taking an artificial mixture of two closely allied races, must come rather as a shock.*

The better to illustrate this, I take some data recently deduced by Miss C. D. FAWCETT. She finds for 806 male skulls, from the Paris Catacombs, the correlation for length and breadth ·0869 ± ·0236, and for 340 female skulls, from the same locality, − ·0424 ± ·0365. The existence of the negative sign and the comparative smallness of the correlation, as compared with the probable errors, might lead us to assert the correlation between the length and breadth of French skulls to be sensibly zero.

If now the two sexes be mixed, the heterogeneous group has for correlation ·1968 ± ·0192, a value which cannot possibly be considered zero. Thus the mixture exhibits a large spurious correlation.

Whether any given mixture increases or reduces the correlation will depend entirely on the signs of the differences of the means of the sub-groups. But the danger of heterogeneity for the problem of correlation will have been made manifest. If the value of R for any mixture, whose components are known, is to be calculated, then we have only to note that:

$$\Sigma^2 = \frac{S(n\sigma^2)}{N} + \frac{S(n_p n_q (m_p - m_q)^2)}{N^2}, \quad \Sigma'^2 = \frac{S(n\sigma'^2)}{N} + \frac{S(n_p n_q (m'_p - m'_q)^2)}{N^2} \qquad \text{(xxxv.)}.$$

---

* Thus the mere fact of breeding from *two or three* individuals selected at random can easily produce a correlation between organs in the offspring, which has no existence in the species at large.

II. *On the Inheritance of Fertility in Mankind.* By KARL PEARSON, F.R.S., and ALICE LEE, B.A., B.Sc.

(8.) In commencing an investigation of this kind where the results to be expected were quite unknown to us, but where we had reason to believe that the *apparent* strength of inheritance must be very small, we considered that the first thing to be done was to investigate the largest possible amount of material. Thus the probable errors of our results would be very small and any, however small, correlation between fertility in parent and offspring would be brought to light. Attempts might then be made to strengthen any correlation discovered by removing so far as possible one after another the various factors tending to screen the full effect of the inheritance of fertility.

Such factors are for example:

(*a*.) The age of both husband and wife at the time of marriage. The real fertility may be screened by late marriages of one or both parents. The relation of fertility to age at marriage has been dealt with by several writers, notably by DUNCAN and ANSELL.*

(*b*.) The duration of marriage. The data may be taken from a marriage not yet complete, both parents being still alive. Or from a marriage which is complete one or both parents being dead. In the former or the latter case the marriage may be complete so far as fertility is concerned, *i.e.*, details of offspring may be available till the wife has reached the age of 50 years, which for statistical purposes may be taken as an upper limit to fecundity.

(*c*.) Restriction of fertility during marriage. It has been shown in a paper on Reproductive Selection† that there is evidence of the sensible influence of this factor in man. It tends to give fictitious values to the fertility of the younger, rather than the elder generation, and so obscures the correlation.

We have accordingly two problems before us:

(i.) Supposing these and other factors tending to screen the effects of reproductive selection to exist, can we show that it still produces sensible effects in the case of man, and thus demonstrate that fertility is really inherited?

(ii.) Can we by eliminating these factors so far as possible obtain a lower limit to the coefficient of heredity in the case of fertility, and ascertain whether it approximates in value to what we might expect from the Law of Ancestral Heredity?

The first impression of the reader may be that it is only needful to select the

---

* J. MATHEWS DUNCAN, 'Fecundity, Fertility, Sterility and Allied Topics,' second edition, Edinburgh, 1871. CHARLES ANSELL, Junr., 'Statistics of Families in the Upper and Middle Classes,' London, 1874.

† 'The Chances of Death and other Studies in Evolution,' vol. I, pp. 77, 89.

fertility of marriages, which were formed with husband and wife between 20 and 28[*] say, and which have lasted till the wife is over 50. But these conditions must be true in *two* successive generations, and, had we adopted them, we may safely say that without immense labour it would have been impossible to collect even a thousand cases. From the whole of the peerage, the baronetage, the landed gentry, a variety of family histories, of private pedigrees, and a collection of data formed of families at first hand, it was not possible to extract more than about 4000 cases for the inheritance of fertility in the female line, when the limitations were far less stringent, being applied only to *one* generation, and consisting in our taking marriages entered into at any time of life for either husband or wife, and lasting till the death of one member or for at least fifteen years. Even in this case the pedigree of the wife had to be sought for from one record to another and often in vain. It is the male pedigree with which the recorder in nearly all cases occupies himself.

Only those who have attempted the labour of extracting, as has been done in this case, some 16,000 separate returns, will fully grasp the difficulty of making the limitations of selection more and more complex ; the quantity to be obtained becomes dangerously small and the labour immensely increases. Even could with time and patience a sufficient selection of ideal cases have been made, it does not follow that the result would be satisfactory ; for, we should have made a narrow *selection*, and this very fact might indicate that possibly we have been selecting one grade or class of fertility. It is possible that the less fertile are the weaker, and so more liable to die early ; or again it may be the more fertile women who are subjected to the more frequent risk of childbed, and thus are less likely to appear in the selection of long marriages. Even greater or less risk at birth may be an inherited character in women, and may not unfairly be looked upon in itself as a factor limiting fertility naturally.

Taking these points into consideration, it seemed that if we were to have enough material to draw conclusions from we must entirely drop all attempt to classify by age of parents at marriage. We might make some limitations but they must not be very stringent ; they must leave room for an increase of stringency in different directions, so that we could roughly appreciate the influence of the screening factors. Accordingly our plan has been to show that correlation actually does exist between parent and offspring with regard to fertility, and that when we make the conditions more stringent the correlation increases towards the value indicated by the law of ancestral heredity.

(9.) *On the Inheritance of Fertility in Woman.*—(i.) Table I. gives the result for 4418 cases of the fertility of a mother and of her daughter. These were extracted from FOSTER's 'Peerage and Baronetage,' BURKE's 'Landed Gentry,' some family

---

[*] As DUNCAN points out, an early marriage on the average means an earlier cessation of fecundity ; a somewhat later one does not necessarily connote less fertility.

histories and a collection of family data drawn from private pedigrees and other sources. In the case of the daughter, no marriage was taken which had not lasted at least 15 years, or until the death of husband or wife. In the case of the mother no limitation whatever was made, the number of brothers and sisters of the daughter, including herself, being counted. Weight was given to the fertility of the mother, for every possible case that could be got from the records under the above conditions was extracted. It is quite possible that a certain proportion of offspring dying in early infancy have not been entered in the records.

If M_m, M_d be the mean fertilities of mother and daughter, σ_m, σ_d their standard deviations, and r_{md} their correlation, we found :

$$M_d = 3{\cdot}494, \qquad M_m = 6{\cdot}225,$$
$$\sigma_d = 2{\cdot}975, \qquad \sigma_m = 3{\cdot}052,$$
$$r_{md} = {\cdot}0418.$$

Clearly owing to the near equality of σ_d and σ_m the regression of daughter's on mother's fertility is sensibly equal to the correlation.

The probable error of r_{md} is determined by the formula given by PEARSON and FILON[*] to be ·0101, or r_{md} is four times its probable error.

We thus conclude :

(i.) That fertility is inherited in the female line.
(ii.) That its effects are very largely screened by the factors to which we have previously referred.

Had we started with no limitation as to the daughter's family, it is highly probable that r_{md} would scarcely have been sensible relatively to its probable error, and, therefore, small series without due regard to screening causes may easily lead the recorder to suppose that fertility is not inherited.

Supposing we exclude from the daughters the 775 barren marriages, we find the mean for 3643 cases of fertile marriages to be 5·237. Comparing this fertility with the observed fertility 6·225 of mothers, a superficial inquirer might at once consider that a diminution of fertility has taken place. The fact is that neither of the results, M_m or M_d gives the actual fertility of the mothers or daughters. These are the means M''_1 and M''_2 of formulæ (viii.) and (vii.) of the theoretical investigation.

Let us apply the theory developed to our statistics. In the first place we note that r is small ; hence r^2 is still smaller, and thus by (xv.) σ''_2 will not differ much from σ_2. Since σ''_1 will be generally less than σ_1 by (xvi.), it follows that σ''_2 will probably be less than σ_2. Approximately, we can take $\sigma_2 = 3$. Turning to (vii.) we see that M''_2 cannot, since r is small, differ widely from M_2. If there be no secular

---

[*] "Contributions to Theory of Evolution.—IV." 'Phil. Trans.,' A, vol. 191, p. 242.

evolution in the real fertility sensible in the one generation, then M_1 would equal M_2. Hence to a first approximation we should have :

$$M_1 = M_2 = M''_2 = 3{\cdot}494.$$

To obtain a second approximation we may substitute this in the small terms of (vii.). Here σ'^2_1 must be found from (ii.) ; neglecting the cubic term we have :

$$\sigma'^2_1/\sigma^2_1 = 1 - \sigma^2_1/M^2_1 = {\cdot}2628.$$

Hence :

$$\begin{aligned} M_1 = M_2 &= 3{\cdot}494 - r\frac{\sigma^2_1}{M_1}\left(1 + \frac{{\cdot}2628}{1{\cdot}7372}\right) \\ &= 3{\cdot}494 - {\cdot}0418 \times 2{\cdot}5759 \times 1{\cdot}1513 \\ &= 3{\cdot}494 - {\cdot}124 = 3{\cdot}370. \end{aligned}$$

We can now substitute this value of M_1 in (viii.), and we find :

$$M''_1 = 3{\cdot}370 + 2{\cdot}980 = 6{\cdot}350.$$

This differs comparatively little from the actually observed value, 6·225, and is satisfactory evidence of the validity of our theory. The fact that the elder generation was in no way limited like the younger, and that we have neglected the third moment—although fertility distributions are never normal—as well as made other approximations, is quite sufficient to account for the difference observed.

We may take it that 3·4 is practically the fertility of the elder generation, and that this is raised to about 3·5 by reproductive selection in the younger generation. The result 6·2 for the elder generation is thus purely a result of weighting due to the nature of the record.

(ii.) Table II. gives the result of 1000 cases taken from the Peerage. Here the conditions of extraction were as follows :—

One member only was taken out of each family, or no weight was given to the fertility of mothers.

The daughters' marriages had all been completed by the death of one parent or had lasted at least 15 years.

There was no limitation with regard to the parents' marriages. We found :

$$M_d = 3{\cdot}923, \qquad M_m = 5{\cdot}856,$$
$$\sigma_d = 2{\cdot}758, \qquad \sigma_m = 2{\cdot}751,$$
$$r_{dm} = {\cdot}2096.$$

The coefficient of regression is sensibly equal to that of correlation. The probable error of $r_{dm} = {\cdot}0204$, or not a tenth of the value of r_{dm} itself. Again we conclude that fertility is certainly inherited in the female line. By selecting fairly homogeneous material with a more definite and complete record than exists for the heterogeneous material of the previous case, we have carried up the correlation to five times its previous value, and within a reasonable distance of the value ·3 which would be required by the law of ancestral heredity. The homogeneity of our material is evidenced by the reduction in both standard deviations; the greater completeness of the record by the rise in the fertility of daughters; and the non-weighting of the fertility of mothers by the fall in their mean fertility.

If the reader will turn back to the theory of the influence of heterogeneity on correlation in section (7), he will notice that the expression in (xxxiv.) will be negative, and therefore the apparent correlation less than the real, if we form a mixture of two groups in which $m_p > m_q$ and $m'_p < m'_q$. Now the entries of women in the Landed Gentry and other records are very often entries of "heiresses," while the entries of women in the Peerage are entries because of class. An "heiress" naturally has fewer brothers and sisters than another woman on an average, or we may expect $m'_q > m'_p$. On the other hand an "heiress" need not have fewer children than other women, unless her heritage is the result of her coming from an infertile stock, and is not a result of the incompleteness of her parents' marriage. If she belongs to a somewhat lower social grade, she may possibly be more fertile than the average of a higher social grade. In this case m_p will be $> m_q$, and when we come to mix records of the Peerage with those of the Landed Gentry and Family Histories, we need not be surprised to find the correlation of fertility much weakened, as it undoubtedly is (as shown by (i.) and (ii.) above) by the mixture.

Let us next apply our theory to the above results. We are now dealing with M'_1, M'_2, σ'_1, σ'_2. Assuming that there is no secular change $\sigma_1 = \sigma_2$, and accordingly since $\sigma'_1 = \sigma'_2$ sensibly, formula (xi.) shows us that both $= \sigma_1$.

Further, if $M_1 = M_2$, formula (v.) is a quadratic equation to find M_1; substituting for M'_1, σ_1^2, and r, we have, on solving and taking the only admissible root, $M_1 = 3{\cdot}4625$. Then, applying formula (i.) to find M'_1, we have:

$$M'_1 = 5{\cdot}660.$$

This is not quite as high as the observed value 5·856, but it suffices to show that our theory expresses the main facts. In all probability we have not entirely freed our results from weighting with fertility; because, although every endeavour was made to take only one from each family, it is possible that pairs of sisters have occasionally crept into the record.

(iii.) Table III. gives the result of 1000 cases taken from the Landed Gentry. As we have already noted, the women recorded are largely "heiresses," and we believed this might be one of the chief sources of the heterogeneity of the material in Table I. The conditions of selection were made somewhat more stringent, and were as follows:—Only one daughter was taken from each family, and her marriage must have lasted at least 15 years. No limitation was placed on the duration of the parents' marriage. We found:

$$M_d = 4{\cdot}232, \qquad M_m = 5{\cdot}403,$$
$$\sigma_d = 3{\cdot}292, \qquad \sigma_m = 3{\cdot}241,$$
$$r_{md} = {\cdot}1045.$$

The probable error of $r_{md} = {\cdot}0211$, and again we see that fertility is certainly inherited. The correlation has, however, sunk; probably, as the great increase of variation indicates, because we are dealing with much more heterogeneous material than in the case of the Peerage. While the selection of "heiresses" has largely reduced the number of brothers and sisters, *i.e.*, the fertility of mothers, the limitation to marriages of at least 15 years has increased the apparent fertility of daughters; nor is this increase at all balanced by the fact that heiresses come from small families, and may, therefore, be supposed to be the children of rather sterile mothers. The average number of children of heiresses is sensibly as large as the average number of children of women who are not in the bulk heiresses, and who have, as in the following case, been selected with the same condition as to duration of marriage. The fact is that heiresses are not on the whole the children of sterile mothers; their high fertility and *their small correlation with their mothers* shows us that heiresses in the bulk are rather the daughters of mothers whose apparent fertility is fictitious. They have, owing to the sterility or early death of their husband, to their own marriage late in life, or to some physical disability, or other restraint, never reached their true fertility. If this conclusion be correct, and a comparison of the values of M_d and r_{md} in this and the following cases thrusts it almost irresistibly upon us, then we see that the argument against the inheritance of fertility based upon the fertility of heiresses and non-heiresses is of no validity.* It could not be valid as against the values of the correlation we have found, but the present investigation shows by the value of r_{md} exactly wherein the error lies: the heiress is not infertile, but is the daughter of a fictitiously infertile mother.

Applying our theory to this case, we find from formula (xi.), putting $\sigma_1 = \sigma_2$:

$$\sigma_1^2 = (\sigma'^2_2 - r^2\sigma'^2_1)/(1 - r^2),$$

whence we find $\sigma_1 = 3{\cdot}293$, a result sensibly identical with σ_d. Solving the quadratic (v.) with $M_1 = M_2$ to find M_1 we find:

$$M_1 = 3{\cdot}952.$$

Hence by (i.) we have:

$$M'_1 = 6{\cdot}838,$$

the actually observed value being 5·403. Thus the theory completely fails to give the fertility of the heiresses' mothers; for such a fertility as we find in the daughters,

---

* See, for example, a recent letter of Mr. HOWARD COLLINS in 'Nature,' November 3, 1898.

the mothers' fertility is far too low. This again emphasises the point we have already referred to. The peculiar character of the selection, which leads to the female record in the Landed Gentry, is not one such as we have considered in our theory, where the record of any family is likely to appear in proportion to its size. Such a distribution is a *chance* distribution, but a selection of women inheriting land has not this character, and a woman who is the mother of co-heiresses is hardly doubly as likely to appear as the woman who is mother of one. A marriage in either case is likely to be arranged, and if we take only one daughter from each family the record will not already have weighted—at any rate to the full extent—every mother with her fertility. If the reader will compare the variation columns for both daughters and mothers in Table III. with the corresponding columns in Table II. or Table IV., he will at once see how anomalous is the selection of women given in the Landed Gentry.

(iv.) Table IV. gives the results for 1000 cases taken from the Peerage and Baronetage under the following limitations: one daughter only was taken for each mother, and in the case of both mother and daughter the marriage must have lasted at least 15 years. We found:

$$M_d = 4{\cdot}335, \qquad M_m = 5{\cdot}898,$$
$$\sigma_d = 2{\cdot}967, \qquad \sigma_m' = 2{\cdot}830,$$
$$r_{md} = {\cdot}2130.$$

The probable error of $r_{md} = {\cdot}0204$. Thus, as it is now hardly necessary to repeat, fertility is certainly and markedly inherited. The regression coefficient is now as high as ·2233, the closest limit we have yet reached to the theoretical ·3 of the law of ancestral heredity.

Owing to the limitation to marriages of 15 or more years, the means of the fertilities of both mothers and daughters have risen, in the latter case more, however, than the former. It might have been expected that the fertility of mothers would have risen more, but it must be remembered that M_m is the apparent and not the *real* fertility of mothers; and further, since the record largely weights the more fertile women, the bulk of the mothers are already those with large families, *i.e.*, those whose marriages have lasted at least 15 years.

Assuming that there is no sensible secular change in unweighted fertility, *i.e.*, $\sigma_1 = \sigma_2$, we have from the formula on p. 284:

$$\sigma_1 = 2{\cdot}973.$$

From (v.) with $M_1 = M_2$ we find:

$$M_1 = 3{\cdot}845$$

for the real fertility of mothers. This is a sensible increase on the value 3·463 given in Case (ii.), in which there was no minimum duration to the length of the mother's marriage.

Applying formula (i.) we find:

$$M'_1 = 6{\cdot}144,$$

which is somewhat more than the observed value 5·898. The reason for this lies, we think, in the difficulty already referred to on p. 263. If we start extracting mothers, it is often difficult to follow the daughter's history; starting with the daughter it is much easier, although still laborious, to trace back her ancestry, and find the number of her brothers and sisters. Even in this case the search may be lengthy. But as daughters when married change their name, it requires great care in extracting large quantities to be sure that a mother is not repeated, *i.e.*, some approach made to weighting her with her fertility. Every care was made in extracting the records, but we cannot hope to have always avoided weighting to some extent a mother, and if this be done we shall have a transition from formulæ (xi.), (v.), and (i.) towards formulæ (xv.), (vii.), and (viii.), which would well account for the difference found between theory and observation.

If we sum up for inheritance of fertility in the female line on the basis of these four cases, we draw from each one of them the unquestionable result that fertility in woman is an inherited character. Further, the more we remove causes of fictitious values for the fertility in either generation, the closer does the value approach that required by the law of ancestral heredity. The two chief disturbing factors which we have not been able to eliminate are (*a.*) the age at which marriage is entered upon, (*b.*) restraint giving a fictitious value to the fertility. Both these causes must give a lessened value to the correlation of fertility between mother and daughter, and the first, judging from the great influence of age at marriage on fertility, cannot fail to give a serious diminution. Hence if we find the regression coefficient as high as ·2233, when we neglect these factors, it is no stretching of facts to conclude that it would in all probability rise to ·3 could we take them into account.

Our conclusion, therefore, is that fertility in woman is certainly inherited through the female line, and most probably according to the law of ancestral heredity. Reproductive selection is actually a *vera causa* of progressive change, but its influence is largely, if not entirely screened by the numerous factors tending to make the apparent fertility of women differ from their real or potential fertility.

(10.) *On the Inheritance of Fertility in Man.*

(i.) While many of the difficulties involved in the extraction of data for women still exist for man, a new and important feature tending to screen the full influence of the law of ancestral heredity arises in his case. The full fertility of the husband is not in the average case at all approached in the case of monogamic marriage. Hence, in considering the size of a man's family as a measure of his fertility we are measuring a character which differs largely from the character of fertility in woman. It is only in the case of sterile or even very sterile men that there is likely to be a correlation shown between the sizes of the families of fathers and sons.

The intensity or duration of fecundity in the husband must, one or other, be less than that of the wife,—and this will hardly be so in the great run of cases—if his family is to be in any way a measure of his fertility, or, as it might be better to call it in this case, his sterility. We are seeking to find a correlation between two characters, one in father and one in the son, neither of which we can measure unless they fall short of a certain limit. The result is that our correlated material is weakened down by the admixture of a mass of uncorrelated material in the manner indicated in Proposition V. of the theoretical part of this investigation. Within the family we cannot hope to get a correlation which will approach that indicated by the law of ancestral heredity. We may still, however, hope to ascertain whether fertility, respectively sterility, is an inherited character in man as well as woman.

(ii.) Our first attempt was to collect as much material as possible, so that our limitations were few. The Peerage, Baronetage, Landed Gentry, Family Histories, private pedigrees, and collected data provided the 6,070 cases arranged in Table V. Here large families were weighted because several, where available, were taken from one family. The son's marriage must either have lasted till the death of one partner or at least 15 years; there was no condition as to the duration of the father's marriage.

We have spoken of the correlation between fertility of father and son, but since only a single marriage of the father is taken, it may be equally well termed a correlation between the fertility of the mother and son, which may, perhaps, to some extent explain the relatively high values reached.

Let M_s, σ_s be the mean and standard-deviation of the son, M_p, σ_p of the parent, and r_{sp} the correlation; then we found:

$$M_s = 3{\cdot}871, \qquad M_p = 5{\cdot}831,$$
$$\sigma_s = 3{\cdot}003, \qquad \sigma_p = 3{\cdot}190,$$
$$r_{sp} = {\cdot}0514.$$

The probable error of $r_{sp} = {\cdot}0087$. Thus the correlation is nearly six times the probable error, or fertility in man is certainly inherited.

(iii.) Table VI. contains the result of extracting 1,000 cases from the Peerage, only one son being taken from each family, and his marriage having lasted at least 15 years. No attention was paid to the length of parents' marriage.

We found:

$$M_s = 5{\cdot}070, \qquad M_p = 5{\cdot}827,$$
$$\sigma_s = 2{\cdot}910, \qquad \sigma_p = 3{\cdot}142,$$
$$r_{sp} = {\cdot}0656.$$

The probable error of $r_{sp} = {\cdot}0212$. This case closely confirms the previous case; M_p and σ_p remain sensibly the same, M_s has risen owing to the longer period of duration of the son's marriage, and since there is a longer period for the possible exhaustion of the male fertility, we find r_{sp} is slightly larger. Although the numbers are smaller than in Case (i.), the probable error is not so large but that we can still assert an inheritance of fertility in man.

(iv.) Lastly, to compare with Case (iii.) for women, 1000 cases were extracted from the Landed Gentry, and are given in Table VII. Here no marriage of the son or parents was taken under a minimum of 15 years' duration, and only one son taken from each family. We found :

$$M_s = 5{\cdot}304, \qquad M_p = 6{\cdot}272,$$
$$\sigma_s = 2{\cdot}951, \qquad \sigma_p = 2{\cdot}911,$$
$$r_{sp} = {\cdot}1161.$$

Thus the longer duration of the marriage, which gives a greater chance for the exhaustion of the fertility of a partially sterile father, leads to an increased correlation. The probable error here is ·0210, and the correlation is thus unquestionable.

It would be idle to apply the theory before developed to these male cases, for the simple reason that we must certainly look upon them as containing a large proportion of uncorrelated material. But they suffice to show that male fertility is an inherited character, and although the results are widely different from those indicated by the law of ancestral heredity, they are large when we consider how little male fertility appears measurable by the results of monogamic marriage. Were an approximately close measure of male fertility available, there is certainly in the above results no reason to induce us to believe that it would not be found to obey the law of ancestral heredity.

(11.) *On the Inheritance of Fertility in Woman through the Male Line.*

Although we are not able to measure the potential fertility of the male, we are able to determine whether he transfers fertility from his mother to his daughter. This may be simply done by correlating the fertility of a woman and that of her paternal grandmother. This problem belongs to an important class—namely, questions as to the extent to which a sexual character is inherited through the opposite sex. DARWIN has touched upon this "transmission without development" in Chapter viii. of the 'Descent of Man,'* and we shall find his views amply verified.

The problem before us is : Does a woman have as close correlation with her paternal as with her maternal grandmother in the matter of fertility ?

To solve this problem 1000 cases were taken out of the Peerage for the fertility of a woman and of her paternal grandmother. The marriages of the woman and of her grandmother were both taken with a minimum duration of fifteen years. Every care was taken that no weight should be given to fertile families by taking only one out

---

* Second Edition, p. 227, 'Laws of Inheritance.'

of each family, but, of course, the difficulty of avoiding this is increased when a pedigree must be traced through three instead of two generations.

If d denotes granddaughter, g grandmother, the following results were obtained (Table VIII.) :—

$$M_d = 4{\cdot}411, \qquad M_g = 5{\cdot}657,$$
$$\sigma_d = 2{\cdot}897, \qquad \sigma_g = 3{\cdot}056,$$
$$r_{dg} = {\cdot}1123.$$

The coefficient of regression of daughter's fertility on grandmother's fertility $= {\cdot}1065$.

The probable error of $r_{dg} = {\cdot}0211$.

According to the law of ancestral heredity[*] we should expect the grandparental correlation and regression to be half the parental and equal to ·15. Comparing the present result with Case (iv.), we see that ·1123 and ·1065 have to be compared with $\frac{1}{2}({\cdot}2130)$ and $\frac{1}{2}({\cdot}2233)$, or with ·1065 and ·1116. These are differences well within the probable error of our results, or we may conclude that the correlation of a woman with her paternal grandmother is exactly what from Case (iv.) of Section (9) we should expect to find for her correlation with her maternal grandmother. The reduction from ·15 to ·1123 is just what we might have predicted after the maternal reduction from ·3 to ·2130. We, therefore, conclude that the fertility of woman is inherited through the male line with the same intensity as through the female, and this intensity is most probably that which would be indicated by the law of ancestral heredity.

(12.) We do not stay to consider many points which flow from our tables, such, for example, as the amount of restraint indicated by the hump at the start of our various frequency distributions for size of families, partly because such consideration would lead us beyond our present scope, the inheritance of fertility, and partly because this point has been already dealt with by one of us in a paper on 'Reproductive Selection.' We consider that we have shown fertility in mankind to be an inherited character in both lines, and probably obeying the law of ancestral heredity.[†] By aid of our theoretical investigations it is clear that the average size of a family (M_1), as deduced from our record data (M'_1 or M''_1), is about 3·5 children, if the marriage lasts till the death of one partner, or at least till 15 years; it is about 3·9 to 4 children if the duration of the marriage is at least 15 years. Reproductive selection would increase this average by about ·5 child per generation were its influence not counteracted

---

[*] "Mathematical Contributions to the Theory of Heredity, on the Law of Ancestral Heredity," 'Roy. Soc. Proc.,' vol. 62, p. 397.

[†] In the paper on "The Law of Ancestral Heredity" ('Roy. Soc. Proc.,' vol. 62, p. 412) it is stated that fertility is probably inherited, but the amount falls below that which would be indicated by the law of ancestral heredity. At that time only Case (i.) of Section (9) and Case (i.) of Section (10) had been worked out in detail. It is the rise of correlation with more stringent limitation of opposing influences, which suggests that after all that law is true for fertility as for other characters.

by a variety of other factors of evolution. These factors are so active that the influence is reduced to ·12 of a child per generation if we take Case (i.) of Section (9), and, we have little doubt, would be practically insensible did we take all marriages without any limitation whatever. Reproductive selection must, therefore, be looked upon as always tending to increase the fertility of a race; races are not only ever tending to increase, but tending to increase the rate at which they increase—a feature not recognised by MALTHUS, but which strengthens certain of his arguments. So soon, therefore, as environment, or other circumstance, relieves the pressure of opposing factors, a race will not only increase in numbers, but also in fertility. It is this inherited character of fertility, and its constant tendency to change unless held in check by natural selection or other factor of evolution, which seems to us the source of the immense diversity in fertility to be observed not only in different species, but in local races of the same species.

III. *On the Inheritance of Fecundity in Thoroughbred Racehorses.* By KARL PEARSON, F.R.S., with the assistance of LESLIE BRAMLEY-MOORE.*

(13.) The data provided for the fertility of thoroughbred racehorses by the *stud-books*, are of a kind which cannot be hoped for except in the cases of pedigree animals kept for breeding purposes, and of specially-arranged experiments on insects, &c. We have a practically complete record of the stud-life of every brood-mare. The sire by whom she has been covered in each year is stated, and the result, barren, dead foal, living filly or colt, twins, &c., can be ascertained. It is also possible to find out whether the foal dies young, say as a yearling. By examining the whole series of stud-books the complete pedigree of any mare or sire can nearly always be found, and the correlation theoretically worked out for almost any degree of relationship.

In starting an investigation of this kind on such a great mass of raw material, it is necessary to draw up certain rules for the extraction and arrangement of data. These rules must be prepared without any definite knowledge of the character of the material in *bulk*, for this can only be found after, perhaps, some 1000 cases have been extracted and worked out. Hence the rules originally adopted are often not such as an investigator would have arranged had he known beforehand the general character of the conclusions he would reach. But the statistician cannot, like the experimental physicist, modify without immense labour his methods and repeat his experiment. The collection of his data has frequently been far too laborious a task for repetition. His raw material has been prepared in a certain manner; he may

---

* During the three years in which this investigation has been in progress, a considerable number of friends have given me substantial aid in the arithmetical work, or in the preparation of the 6,000 pedigree cards on which the results are based. Mr. BRAMLEY-MOORE has latterly been my chief helper, but I am also much indebted to Miss ALICE LEE and Mr. G. U. YULE. Miss MARGARET SHAEN and Miss LINA ECKENSTEIN have also contributed to the labour of extracting the raw data from the stud-books.

sort and rearrange his data cards in a variety of ways, but to prepare new cards on a different system is practically beyond his powers.

These remarks are made in order to meet criticism of the method in which my data cards were prepared. I could *now* possibly extract more convenient data, but that is only because of the knowledge gained in the process of examining the fecundity of several thousand horses. I did not even know, *ab initio*, the extent of variability in equine fertility; I did not even know the immense preponderance which would have to be given to certain sires, at any rate I had no numerical estimate of it. Nor had I any percentage of the number of cases in which a pedigree might end abruptly with an alternative sire.*

I saw at once that the apparent fertility of racehorses was even less close to their potential fertility (which I presume to be the inherited character) than in the case of man. Mares go at different ages to the stud, they remain—for reasons not stated—uncovered for occasional years, or periods of years; they return to the training stable for a time; they are sold abroad; they are converted into hunters, put into harness, or, as is occasionally recorded, sold to cab proprietors. This by no means invariably denotes that their fertility is exhausted; their offspring may be bad racers, or their stock unfashionable. Very frequently also we find the mare put to a cart-horse stallion for a year, a few years, or for the remainder of her career, and then no record at all is given of the result. Thus the total fertility recorded can have but small correlation with the potential fertility, and I was compelled to deal with fecundity. The insufficiency of the *apparent* fertilities, as recorded in my mare index, to solve the problem, may be illustrated in the following manner: 1100 cases of the apparent fertilities of mares and dams having had at least four coverings were tabulated (Table IX.). The following results were calculated from this table, the subscript m referring to mare and d to dam :—

$$M_d = 7\cdot6655, \qquad M_m = 6\cdot1391,$$
$$\sigma_d = 3\cdot3652, \qquad \sigma_m = 3\cdot1617,$$
$$r_{md} = -\,\cdot0868.$$

The probable error of $r_{md} = \cdot0202$, and thus we might argue that a fertile dam has, on the average, infertile offspring. But an examination of the above numbers shows us that the dams are more variable than the mares,† and yet the dams have been theoretically subjected to the greater selection, for they must all be granddams, or the fertility of the mares could not have been recorded. We are forced to conclude that the mares have been in some manner selected, and the form of the selection is fairly obvious on examining the table. There appears a great defect of

---

* Even the pedigree of such a famous racehorse as Gladiateur is soon checked by the occurrence of alternative sires. His sire, Monarque, was the son of either The Baron, or Sting, or the Emperor.

† The variability of mares, as a whole, not separated into mares and dams, is (see Art. 16) 3·2775.

mares in the third quadrant,* *i.e.*, of mares and dams of large fertility, the frequency is cut off abruptly in this quadrant. The reason for this is fairly clear. We have dealt with a limited number of years, about 30, of horse-breeding ; hence, when the dam has a long record, her later offspring at any rate cannot possibly have a long one ; when she has a short one, it is possible for them to have a long one. Accordingly, there has been a process of unconscious selection, which has led to a negative correlation of the apparent fertilities.

To illustrate the point further, two more correlation coefficients were obtained. In Table X. are given the apparent fertilities of mares and their dams with a minimum of eight coverings. We find :

$$M_d = 8\cdot6191, \qquad M_m = 7\cdot6309,$$
$$\sigma_d = 3\cdot1656, \qquad \sigma_m = 2\cdot8149,$$
$$r_{md} = -\cdot0876.$$

The probable error is again about ·0202. While the mares now form a group with their mean fertility almost equal to that of the dams in the previous result, their variability is markedly less. Relatively to the dams its reduction is even greater. The correlation is sensibly the same. It would thus seem that the anomalous selection of mares which thus reduces their variability so markedly below that of the dams is not in the *low* fertilities.

I now removed from the Table IX. all parts of it concerning mares with a fertility greater than 8 ; 867 mares and dams remained with a minimum limit of four coverings, the mares not having a greater fertility than 8 offspring. I found :

$$M_d = 7\cdot7636, \qquad M_m = 4\cdot8558,$$
$$\sigma_d = 3\cdot3983, \qquad \sigma_m = 1\cdot9887,$$
$$r = -\cdot0190.$$

The probable error of $r = \cdot0229$. Now the line of regression for dams on mares ought to be the same, whether we obtain it from this result or from the first results in which mares with more than 8 offspring are included. Yet, in this case, there is no sensible correlation at all. In other words, if we exclude the data for large fertilities, we should have to conclude that there was no correlation between the apparent fertilities recorded for mares and their dams. We are thus forced to conclude that apparent fertility is a character depending on the manner in which the record is formed, and must be useless for the investigation of inheritance. This investigation strengthens my *à priori* reasons for selecting fecundity, not apparent fertility, as the character to be investigated. I took the fecundity of a brood-mare to be the number of her living offspring divided by the potential number of her offspring under the given circumstances. Of both numerator and denominator of this ratio I must say a few words.

---

* The portion of the table cut off by vertical and horizontal lines through the means of dams and mares.

In considering the inheritance of fertility I had two different problems in my mind: (i.) Is fertility pure and simple inherited? *i.e.*, Does a very fertile mare have offspring more fertile than the average? And (ii.) What effect does reproductive selection actually have on the population? *i.e.*, To what extent is it screened by other factors of evolution; does the very fertile mare actually have more offspring than the less fertile? Is, for example, her stock weedy and likely to die early? In the case of mankind, the fertility of a woman is, as a rule, effectively brought to its limit with the end of her marriage, and accordingly I started with completed marriages. In the case of a brood-mare her effective fertility depends not on the offspring she has but on the number of these which survive foaldom. It would doubtless have been better to have treated these two problems of fertility separately, but being fairly confident from Proposition I., p. 260, that fertility must be inherited, I was more interested to test the actual effect of reproductive selection. Accordingly I selected as the numerator of my fecundity ratio, not the number of foals born, but those who survived to the yearling sales. The difference is not very great, but quite sensible. For example, the mean fecundity of 3909 brood-mares, measured in my way, = ·6343, *i.e.*, 63 surviving offspring on the average of 100 coverings.

The following table gives the result of reckoning merely barren mares and those slipping foals or giving birth to dead foals in a twenty-year period:—

AVERAGE Fecundity of Brood-mares.

Year.	Average fecundity.	Year.	Average fecundity.
1873	·712	1883	·693
1874	·703	1884	·678
1875	·707	1885	·702
1876	·697	1886	·700
1877	·692	1887	·682
1878	·680	1888	·695
1879	·683	1889	·685
1880	·666	1890	·686
1881	·680	1891	·679
1882	·667	1892	·675

The averages of five-year periods are:

·702, ·675, ·691, ·684,

and of the whole period, ·688.

There does not appear to be sufficient evidence for any secular change here, and we may take ·688 to represent the average fecundity of the brood-mare, reckoning viable offspring to the number of coverings. The difference of ·688 and ·634 gives a death-rate of 5·4 foals in 68·8, or a death-rate of 7·85 per cent. of foals before maturity. If a considerable part of this death-rate be differential, we have room for natural selection influencing the drift of reproductive selection. The standard deviation in the fecundity is, however, about ·191, or about 19 foals in the 63, or about 30 per cent.—a very great variation, so that if fecundity be inherited, a differential death-rate of the immature will hardly suffice to check it.

So much then of the numerator of my ratio. I have spoken immediately above of the denominator as if it were the number of times the mare had been covered. It is generally this, but in the relatively few cases where the mare has given birth to twins, I have counted that covering *twice*. Had this not been done the fecundity might have been greater than unity, for example even in some exceptional cases have risen to two. On the other hand, a loss of twins would have been marked by no greater change in fecundity than a loss of one foal, or the survival of *one* twin would not have been different in its effect on fecundity to the birth of a foal. In order, therefore, to avoid these difficulties—especially that of isolated individuals lying far beyond the fecundity range of 0 to 1—when twins were born the potentiality of the covering was reckoned in the denominator as two. The relative infrequency of twins causes, however, this modification of the denominator to have small influence on the result.

My next step was to form some estimate of the extent to which fecundity thus measured was the same for different periods in a mare's breeding career. I expected fecundity to diminish with age as in the case of mankind, but taking out a fairly large test number of mares, I found that their fecundity for the periods covered by two successive stud-books was in the majority of cases closely the same. With larger experience I should now lay more weight on the decrease of fecundity with age; and I also think fecundity is smaller when the mare first goes to the stud. But even thus much of the reduced fecundity of old mares seems to arise from breeders sending famous mares to the sire long after their breeding days are passed. I have several records of old mares being covered seven or eight times without offspring. This custom of breeders was much more rife in the early days of breeding than it appears to be now, when some breeders discard or sell a fairly old mare, even if she is barren two or three successive years. Clearly the custom gives the mare a fictitious fecundity, far below her real value, and probably accounts for granddams having a somewhat less fecundity than their granddaughters.

The next problem to be answered was the effect the method of forming my fecundity ratios might have on the relative numbers which would be found in different element-groups. For example, supposing the element of fecundity to be 1/10, or the element-groups 0–1/20, 1/20–3/20, 3/20–5/20, . . . 17/20–19/20, 19/20–1, would the fact that the fecundity ratio is a ratio of *whole* numbers cause, *à priori*, a greater probability of frequency in one of these element-groups than another?

To begin with, all estimation of fecundity based on less than *four* coverings was discarded. Three coverings give too rough an appreciation of a mare's fecundity, it can only fall into one of the values 0, 1/3, 2/3, and 1. The question then arises, if all the fecundities:

0/4, 1/4, 2/4, 3/4, 4/4,
0/5, 1/5, 2/5, 3/5, 4/5, 5/5,
0/6, 1/6, 2/6, 3/6, 4/6, 5/6, 6/6,
. . . .
0/26, 1/26, 2/26, ... 25/26, 26/26,

were equally likely, how would the frequency depend on the grouping ?*

Taking 26 coverings as the probable maximum—it actually occurs—we have for the total number of fecundities given above : $5 + 6 + 7 + \ldots + 27 = 368$ separate fecundities. Let us see how they would be divided in one or two cases.

Case (i.) Let the elements be based on 1/8, or be 0–1/16, 1/16–3/16, 3/16–5/16, 5/16–7/16, 7/16–9/16, 9/16–11/16, 11/16–13/16, 13/16–15/16, 15/16–1.

The half-groups at the ends are taken so that zero and perfect fecundity should really be plotted at the middle of a 1/8 element. We find, adding up the numbers of the above fecundities which fall into the nine groups, the following frequencies :—

33·5, 42, 43·5, 44, 42, 44, 43·5, 42, 33·5.

There is thus a somewhat deficient frequency in the terminal groups, and this would probably to some extent bias the distribution.

Case (ii.) Let the elements be based on 1/15, or be

0–1/30, 1/30–3/30, 3/30–5/30, . . . 25/30–27/30, 27/30–29/30, 29/30–1.

We have the following distribution :

23, 22, 23, 23·5, 22·5, 23·5, 23·5, 22, 22, 23·5, 23·5, 22·5, 23·5, 23, 22, 23.

The bias here is only slight and the distribution is on the whole very satisfactory.

Case (iii.) Let the elements be based on 1/20, or be

0–1/40, 1/40–3/40, 3/40–5/40, . . . 35/40–37/40, 37/40–39/40, 39/40–1.

We find for the groups :

23, 13, 17·5, 17·5, 17, 18, 17, 18, 17·5, 17, 18, 17, 17·5, 18, 17, 18, 17, 17·5, 17·5, 13, 23.

Here the terminal groups have too great a frequency, and the adjacent groups too little. It is clear that the division into 1/15 elements is better than those of 1/8 of 1/20, so far as these results go. But unfortunately the different coverings do not occur in anything like the same proportions. Their exact frequencies could only be found *à posteriori*, and I was desirous of having some idea of grouping before start-

---

* Such problems are really not infrequent in statistical investigations, and seem to be of some interest for the theory of fractional numbers. Mr. FILON worked out for me the details of the cases given below.

ing the labour of extraction I therefore weighted the different coverings on the basis of a small preliminary investigation as follows :

Case (iv.) Number	of coverings,	4 to 5	inclusive,	loaded with	2.
,,	,,	6 to 9	,,	,,	3.
,,	,,	10 to 15	,,	,,	4.
,,	,,	16 to 18	,,	,,	2.
,,	,,	19 to 26	,,	,,	1.

The resulting system of frequencies was :

54, 42·5, 45·5, 47, 45·5, 46·5, 46, 45, 45, 46, 46·5, 45·5, 47, 45·5, 42·5, 54.

This system is not so uniform as in Case (ii.). I had hoped that the 744 frequencies would have been fairly closely the double system of Case (ii.). The main irregularity occurs at the terminal groups, or those having fecundities nearly zero and nearly perfect. These I considered would be relatively infrequent, when we started with as many as four coverings, and had an average failure of about 37 in 100. The sequel showed that the assumption was legitimate, so far as regards zero fecundity, but that perfect fecundity was sufficiently frequent to cause a hump in the frequency curve for fecundity, corresponding to the group-element 29/30 to 1. The frequency of this group is greater than that of the group 27/30 to 29/30, when we start from at least four coverings. This hump entirely disappears, however, if we start with at least eight coverings. Thus I take the hump to be purely "spurious," *i.e.*, a result of the arithmetical processes employed, and not an organic character in fecundity. It depends upon our definition of fecundity, which is not a truly continuous quantity.

As the theory of correlation applied is not in any way dependent on the form of the correlation surface, beyond the assumption of nearly linear regression, the hump cannot, I think, sensibly affect our conclusions. Had I known, however, *à priori*, what the frequency of different coverings and the nature of the fecundity frequency curve would be, I should have attempted to choose such a group-element, that, with proper weighting of the coverings, there would have been no *arithmetical* bias to the terminal groups. As it was, it seemed to me that the group-element of 1/15 gave fairly little arithmetical bias—at any rate where the bulk of the frequency would occur—and it was accordingly adopted as a basis for classifying fecundities.

The dfficulty illustrates the point I have referred to, namely, that in statistical investigations the best classification can only be found *à posteriori*, but the classification adopted has usually to be selected *à priori*.

The 1/15 element being selected, the letters *a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*, *i*, *j*, *k*, *l*, *m*, *n*, *p*, *q* were given to the 16 groups of fecundities from 0 to 1, as cited under Case (ii.).[*]

---

[*] A table was formed of the 368 actually-occurring fecundities, from which it was possible to at once read off the group (or it might be two groups, *e.g.*, ·5 falls half into *h* and half into *i*) into which they each fell.

Thus the fecundity of a mare was described by one of these 16 letters. Here the centre of the *j* group, for example, is ·6, and it covers all fecundities from ·56 to ·63. Thus midway between *j* and *k* we are at about the mean fecundity.

The more recent Stud-Books, vols. 12 to 17, were taken as containing more complete details and, what is more important, less in-and-in breeding, although as we shall see, this is still an important factor. These volumes cover 30 and more years of English* stud life. From these 30 years' records upwards of 5000 mares, who had been covered upwards of four times, had their fecundity ascertained. The process was a very laborious one, as each mare had generally to be sought for in several volumes, and the records in each volume are not continuous, but overlap by quite arbitrary numbers of years. Further, great care had to be taken to identify each mare properly, as the same name is very frequently repeated, and the like difficulty occurs, though to a lesser extent, in the case of sires. A card was then written, giving the name of the mare and those of her sire, her dam, and her dam's sire. Upon this card the letter indicating her fecundity was placed. A card alphabet of mares was thus formed, consisting, in the first place, of about 3000 entries. This alphabet was again gone through and the fecundity of the dams of the mares inserted on the cards till there were about 2500 cases known of mare and dam. The dams were partly found from the existing series, but it was also largely necessary to work out fresh cases. Lastly, the cards were gone through and the fecundity of the grand-dams entered in upwards of 1000 cases. This forms the first series of cards.

In the next place a card index was formed of all the sires serving during these 30 years. This contained upwards of 1000 cards. On these cards the sire's sire was entered, and the fecundity of all the mares contained in the first or mare alphabet was now taken off and placed on the card of the mare's sire. Thus the card of each sire had the letters *a*, *b*, *c*, *d*, *e*, &c., upon it, and a frequency distribution was formed on the card of each sire for the fecundity of his daughters.

The same thing was done for the sires' sires; only here recourse had again to be had to the stud-books to obtain the fecundity of the daughters of the more ancient sires. Finally, a sire-alphabet was obtained which gave the average fertility of the daughters of a sire and of the daughters of his sire, or his half-sisters. On these cards was also placed the number of mares upon which each average was based.

These two card-alphabets, the mare and sire alphabets, form the "dressed" material upon which all the subsequent calculations were based.

(14.) At this point it seems desirable to insist somewhat on the many causes which tend to make the fecundity of mares, as thus determined, to a considerable extent fictitious. Many of these were only apparent to me as I became more and more familiar with the material.

---

* Irish mares were excluded except where, for pedigree purposes, it was necessary to deal with them. Many Irish mares were further included when it came to the valuation of the fertility of mares due to a given sire.

(*a.*) Mares appear to be less fecund at the beginning and end of their breeding career. Hence, when the fecundity is based on a part only of their career, as it often must be, we do not really get a fair appreciation.

(*b.*) A more fertile mare is likely to have more daughters go to the stud than a less fertile one, and hence we get a better appreciation of the fertility of the offspring of the former than of the latter.

(*c.*) Fashion among breeders interferes largely with the exhibition of the natural fecundity of a mare. She may be a famous mare and is sent to a famous sire, even though produce is not so likely as if she were put to a sire of a different class. This appears to be practically recognised when apparently barren mares are sent in one season to two, or even three sires, or again to half-bred horses or cart-horses.

(*d.*) Brood-mares which have produced performers are kept much longer at the stud, and we have the fecundity lowered by coverings after the mare is sensibly sterile. Less important mares are removed sooner from the stud.

(*e.*) Good racing mares are often put late to the stud.

(*f.*) In a certain number of cases we are simply told that the mare had no produce for a period of years, but whether she was covered or not is unrecorded.

(*g.*) Second-rate mares, or mares thought to be near the end of their fecundity, are often sold abroad. In the latter case the fecundity is fictitiously increased; in the former we have only a short period to base it on.

(*h.*) There is no record kept of the half-bred foals, which for our purpose are as important as the thoroughbred foals. "Put to a hunter" is a not uncommon record, with no statement of the result.

(*i.*) Comparatively infertile mares, unless of very valuable stock or famous as racers, are not kept long enough at the stud to get a reliable measure of their fecundity.

(*j.*) The smaller breeders will often put mares to inferior sires, already nearly worn out, either because they own them, or because their fee is low; and thus again a full chance is not given to the fecundity of the mare to exhibit itself.

(*k.*) We have excluded in our determination of the fecundity foals dying young. This is often due to the fault of the mare, but is often again due to the environment.

(*l.*) Lastly, thoroughbred mares are highly artificial creatures, and many must suffer from their environment,* either in the matter of barrenness or slipping foal, in a manner from which the wild horse or a more robust domesticated animal would be entirely free.

These considerations may suffice to show that our values of the fecundity will only roughly represent what may be termed the natural fecundity, and we ought not for

---

* I am told that there are like difficulties with cows. Cows are very liable to slip their calves, and one cow doing so, several others in the herd will or may follow her example. There is a strong folk-belief in Wiltshire—I give it merely as evidence of what a slight change in the environment is supposed to achieve—that the habitual presence of a donkey with the herd in some way soothes the cows, and renders them less ready to slip their calves.

a moment to expect inheritance in the full intensity of the Galtonian law to be exhibited by such material.

(15.) But there is another point of very considerable importance for the weakening of correlation, namely the effect of in-and-in breeding. To get correlation we must have a diversity of parents producing a diversity of offspring, but when the parents become more and more identical, we get larger and larger arrays between which and the parents the correlation is weakened. For example : suppose the correlation found between *all* parents and offspring in the general population, and now select only all the brothers in a large array and find the correlation between them and their offspring, we shall find that the correlation is lower than in the previous case.* It would be impossible to apply theory to the present case, however, because we can only roughly appreciate the extent of such in-and-in breeding. That it is great the following statistics will show.

Of the more than 1000 sires in my sire alphabet, only 760 were sires of mares which had been covered at least four times. These 760 sires had upwards of 5000 offspring, of whom I had the fecundity recorded, but when mares with alternative sires were excluded, there remained only 4677 available mares.† These mares were distributed as follows :—

Daughters	1	2	3	4	5	6	7	8	9	10	11	12
Sires	280	113	78	43	29	22	20	21	22	14	10	10
Daughters	13	14	15	16	17	18	19	20	Above 20			
Sires	11	11	8	6	2	6	4	4	46			

Here the second line gives the number of sires having the number of daughters in the first line in the 4677 cases, which I take to be a fair sample.

Thus over a third of the sires had only one mare. Two-thirds of the sires had together only one-fifth of the mares. *Seventy-six* of the sires were fathers of about half the mares, and 46 sires alone produced 1801 mares, almost as many as 642 sires did. We are here dealing with the fairly long period of 30 years, but even making due allowance for young stallions commencing and old stallions concluding their stud career, it will be manifest that our sample shows that the great bulk of mares for the period in question were the offspring of comparatively few sires.

But let us look at the problem from the standpoint of the sires. My 760 sires

---

* The theory of such cases is fully developed in a memoir on the influence of selection on correlation not yet published.

† Some other cases were also excluded for diverse reasons.

were *all* fathered among themselves except in 49 cases. In other words, they were the sons or grandsons of only 49 sires. Of these 49 sires, there were 12 whose pedigree I could not trace,[*] but they were very probably sons of sires already on my list or among the remaining 37. In the majority of cases they appeared only as the sire of one stallion. The remaining 37, whose pedigree I could trace, were descended at once or in very few generations from 9 sires.[†] Thus both from the standpoint of the mares and of the sires we are dealing with a closely in-bred stock, and this is one and probably a very important factor in the weakening of the fecundity correlation.

Having regard to these difficulties, if we can succeed in showing that fecundity in thoroughbred racehorses is inherited, we can be fairly confident that we have only reached a lower limit of the correlation coefficient.

(16.) *On the Inheritance of Fecundity in the Female Line.*

(i.) A preliminary investigation must here be made, in order to determine the ρ of the formulæ given in Proposition III. (p. 269) we want the correlation of fecundity with fertility. If ϕ be the fecundity, f the apparent fertility, and c the number of coverings, twins counting as a double covering, we have:

$$\phi = f/c,$$

whence if we determine the correlation between ϕ and f, numerous constants will follow. Table XI. gives the correlation between fertility and fecundity for 1000 brood-mares. We found:

$$M_\phi = \cdot 6375, \qquad M_f = 6\cdot 515,$$
$$\sigma_\phi = \cdot 1810, \qquad \sigma_f = 3\cdot 2775,$$
$$\rho = r_{\phi f} = \cdot 5152,$$
$$v_\phi = 100\sigma_\phi/M_\phi = 28\cdot 39, \qquad v_f = 100\sigma_f/M_f = 50\cdot 31,$$

where v_ϕ and v_f are the "coefficients of variation."[‡] Here by YULE's Theorem[§] $r_{\phi f}\sigma_\phi/\sigma_f$ is the slope of the line which most closely fits the curve of regression for fecundity on fertility. If we supposed this curve to be straight, then the line must coincide with it. Now since fecundity vanishes with fertility, the curve passes through the origin, and hence, if the regression be linear, the line must also pass through the origin. In this case, as is shown on p. 270, $r_{\phi f} = v_\phi/v = \cdot 5644$. The difference between $\cdot 5644$ and $\cdot 5152$ may be taken, as it is several times the probable error, to indicate that the regression curve between fecundity and fertility is only approximately linear.

The variations in both fertility and fecundity are here large. Accordingly we

---

[*] Stockmar, Sovereign, Andover, Phaeton, Prince Caradoc, Robert Houdin, Pylades, King of Kent, Garry Owen, Calaban, Homily and Taurus.

[†] Tramp, Sir Peter Teazle, Catton, Buzzard, Orville, Diomed, Sorcerer, Dr. Syntax, Marske.

[‡] 'Phil. Trans.,' A, vol. 187, p. 276.

[§] 'Roy. Soc. Proc.,' vol. 60, p. 477.

must use the formula (i.) for the mean value of an index given in my memoir on spurious correlation.* We shall then obtain an approximate value to the mean number of coverings of each mare. Formulæ (iii.) of the same paper will then give the standard deviation for the number of coverings. In our present notation :

$$c = f/\phi,$$

and therefore :

$$M_c = \frac{M_f}{M_\phi}\left(1 + \left(\frac{v_\phi}{100}\right)^2 - r_{f\phi}\left(\frac{v_f v_\phi}{10,000}\right)\right),$$

$$\sigma_c = M_c \sqrt{\left(\frac{v_f}{100}\right)^2 + \left(\frac{v_\phi}{100}\right)^2 - 2r_{f\phi}\frac{v_f v_\phi}{10,000}}.$$

We find :

$$M_f/M_\phi = 10{\cdot}2196,$$

and :

$$M_c = 10{\cdot}2196 \times 1{\cdot}007 = 10{\cdot}2911,$$

$$\sigma_c = 4{\cdot}4455,$$

$$v_c = 43{\cdot}20.$$

To the same degree of approximation we can further ascertain the correlations between the number of coverings and the apparent fertility and fecundity, *i.e.*, r_{cf}, and $r_{c\phi}$. A short investigation similar to those in the memoir on spurious correlation just cited shows us that :

$$r_{cf} = (v_f - r_{f\phi}v_\phi)/v_c,$$

$$r_{c\phi} = (r_{f\phi}v_f - v_\phi)/v_c.$$

These lead to the numerical results :

$$r_{cf} = {\cdot}8259, \qquad r_{c\phi} = -{\cdot}0572.$$

The conclusions to be drawn from these results are all of some interest. In the first place we may ask : How does M_ϕ agree with its value found from other and more complete series? For 4677 mares—my complete series without mares with alternative sires—the average fecundity was ·6373. A better agreement could not have been hoped for. In a group of 1509 mares dealt with for variation only and entered as "daughters" on the cards—so that they had not been selected by the fact that their daughters must have recorded offspring, as is the case with "dam" entries—I found the following results :—

VARIATION in Fecundity of 1509 Brood-mares (Four Coverings).

Fecundity.	*a.*	*b.*	*c.*	*d.*	*e.*	*f.*	*g.*	*h.*	*i.*	*j.*	*k.*	*l.*	*m.*	*n.*	*p.*	*q.*
Frequency	9	3	11	26	46	43·5	85	122·5	154·5	232·5	194	223	146	100	23	90

Total 1509. $M_\phi = {\cdot}6345.$ $\sigma_\phi = {\cdot}1965.$

---

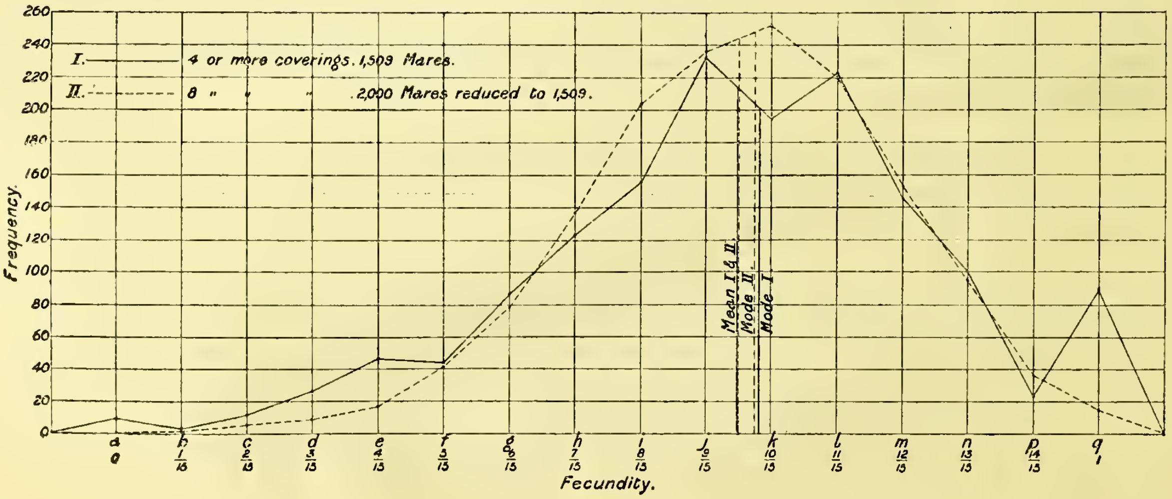
[*] 'Roy. Soc. Proc.,' vol. 60, p. 492.

Now this is precisely what we might expect; the mares belonged to a class, of which we are not certain whether their daughters have or have not recorded fecundity. The mean fecundity is therefore decreased and the variability increased. Add to this group 2400 mares, all of which had had their daughters' fertility recorded, and we find for 3909 mares, $M_\phi = \cdot 6345$ and $\sigma_\phi = \cdot 1910$, *i.e.*, the mean fecundity ascends and the variability falls. Illustration of this law will be found in the following two groups:—

	M_ϕ.	σ_ϕ.
1200 mares	·6337	·1888
1200 dams	·6525	·1643

Thus we send up the mean fertility and lower the variability by separating into two groups the pedigree of one which has a longer record. This is precisely in accordance with the theory already developed. Our mean fecundity and variability for brood-mares may be considered as constant characters, and variations in their values beyond their probable errors due to conscious or unconscious selection in the record itself, or in our extracting from it.

The reader will notice at once, if he turns to the diagram of the above frequency, (i.) that there is a small hump at (a) of no practical importance, and a larger one at (q), perfect fertility being fairly frequent with only four coverings, and there being from the arithmetical processes involved a bias towards (q) as compared with (p). (ii.) The distribution of frequency, although somewhat ragged, is quite clearly not normal, but of the character which in other papers I have called *skew*. Were there any occasion, it would be easy to fit it with one of my skew curves. To mark how (i.) will disappear and (ii.) become still more apparent, I have placed on the diagram the frequency distribution for 2000 mares reduced to the same scale.



VARIATION in Fecundity of 2000 Brood-mares (Eight Coverings).

Fecundity	*a.*	*b.*	*c.*	*d.*	*e.*	*f.*	*g.*	*h.*	*i.*	*j.*	*k.*	*l.*	*m.*	*n.*	*p.*	*q.*
Frequency	0	2	7·5	11·5	21·5	55	104·5	182	271·5	315	337	293·5	204	127	49	19
Total, 2000. $M_\phi = \cdot 6330$.. $\sigma_\phi = \cdot 1568$.																

Thus, making the minimum number of coverings 8 instead of 4, has removed the terminal humps, zero fecundity is now unknown, and perfect fecundity very rare. We have reached a smooth skew frequency distribution ; we see fecundity as a continuous character obeying the usual laws of variation.* The mean fecundity in the two cases is sensibly the same, ·633, but owing to the fact that we have made a selection of a limited group in the second case, the variability is considerably decreased.

The average apparent fertility of brood-mares, 6·515, must not be confused with their average real fertility, for, as we have seen, we have in many cases not a complete record of their stud-life, or such a full record has not been used (*e.g.*, in case of mares still at the stud, but having been already covered four or more times). Its 50 per cent. variation shows that an apparent fertility of 9 to 12 is not infrequent. The average number of coverings being 10 and more, it will be seen that the records of between 50,000 and 60,000 coverings have been dealt with to form our mare and sire alphabets. The large variability in the number of coverings shows that 15 to 20 coverings will not be infrequent, and cases of 26 actually occurred. Lastly, we have the correlation between fertility and the number of coverings, high as might be supposed, for a high apparent fertility could only be exhibited by many coverings. Although a low apparent fertility might correspond to any number of coverings, still, in practice a sterile mare will not be sent indefinitely to the sire. The correlation between the number of coverings and the fecundity is small and *negative* (— ·0572). This follows from the principle that, fertility being the same, a high number of coverings reduces the fecundity, and this factor is more potent than the high correlation of fertility and the number of coverings.

(ii.) Table XII. exhibits the correlation of 1200 mares and their dams with regard to fecundity. Here the more fertile dams are weighted with their fertility, and at least four coverings were required of each mare. If the subscript m refers to mare, and d to dam, we find :

---

* The actual equation to the curve referred to the mode ·6531 as origin, the axis of x being positive towards perfect fecundity, and the unit of x being 1/15 is:

$$y = 342 \cdot 187 \, (1 + x/47 \cdot 1358)^{82 \cdot 6261} \, (1 - x/12 \cdot 1106)^{21 \cdot 2291}.$$

The fit will be found to be very satisfactory.

$$M_m = \cdot 6337, \qquad M_d = \cdot 6525,$$
$$\sigma_m = \cdot 1888, \qquad \sigma_d = \cdot 1643,$$
$$r_{md} = \cdot 0831.$$

The coefficient of regression = ·0945.

The probable error of the correlation is ·0193 and of the regression* ·0195. Thus these quantities are four to five times their probable errors, and we conclude that fecundity is certainly inherited.

The intensity is far below that suggested by the law of ancestral heredity, but it nevertheless exists. Its lowness is probably due to the fictitious character of the fecundity owing to the causes indicated on pp. 298–9. An attempt must now be made to eliminate some of the factors disguising the fecundity, but to do so is by no means so easy as in the case of fertility in man.

(iii.) My first idea was that by taking a higher limit to the number of coverings a closer approach might be obtained to the true, *i.e.*, the inherited fecundity. Accordingly Table XIII. was formed for the correlation of 1000 mares and their dams, when the minimum number of coverings was eight. But I did not recognise that this would give far greater weight in the Table to the older mares, and that accordingly causes (*d*) and (*i*) of p. 298 would now play a much larger part in disguising the true fecundity than before. There appears to be no limit to the number of times a famous old mare may go to the stallion when there is very small hope of any offspring.

Table XIII. gives us the following results :

$$M_m = \cdot 6300, \qquad M_d = \cdot 6360,$$
$$\sigma_m = \cdot 1633, \qquad \sigma_d = \cdot 1500,$$
$$r_{md} = \cdot 0652.$$

The coefficient of regression = ·0708.

The probable error of the correlation is ·0212, and of the regression ·0213, both less than a third of the observed values. We should again conclude from this result that fecundity is inherited, although it offers less strong evidence than the previous case. The influence of selection† is at once apparent in the great reduction of the variabilities. The fact that we are throwing the determination of fecundity more on to the old age period of life appears from the reduced mean fecundities. I attribute the reduction in the fecundity-correlation to this source, *i.e.*, the very diverse treatment which old mares receive at the hands of different breeders.

(iv.) I made another attempt to remove screening causes by taking 1200 more

---

* Pearson and Filon : 'Phil. Trans.,' A, vol. 191, p. 214.

† The effect of such a selection as the above in reducing correlation is dealt with in my paper on the influence of selection on correlation.

mares, not identical with the series in (ii.)* and working out their dams' records most carefully, rejecting any cases in which the breeder was clearly sending the mare to the stallion long after it was obvious (*post facto*) that she was sterile. In this case four coverings were retained as a minimum, and the results are given in Table XIV. We find :

$$M_m = \cdot 6369, \qquad M_d = \cdot 6616,$$
$$\sigma_m = \cdot 1885, \qquad \sigma_d = \cdot 1604,$$
$$r_{md} = \cdot 0995.$$

The coefficient of regression = ·1169.

The probable error of the correlation is ·0193, and of the regression ·0194 ; the correlation is accordingly more than five, and the regression more than six times its probable error. We conclude that fecundity is most certainly inherited. The regression found is, however, only about two-fifths of what is required by the law of ancestral heredity,

(v.) It has been suggested that fertility or fecundity might alternate in *two* generations ; when the offspring are numerous their offspring might have less fertile or fecund offspring. I do not see how this would be possible without its exercising an influence on the correlation of two generations, for we must come to one fertile followed by an infertile generation. But I had made preparations in my alphabet of mares for testing the correlation between mares and their granddams, and I went on to the construction of a table, although the results for mares and their dams showed me that whatever result might be reached, it would be within the probable error of the observations. I reached this conclusion in the following manner : If we go back one generation we introduce, owing to the nature of the record, so much fictitious correlation and so much in-and-in breeding that the coefficient of inheritance is reduced to two-fifths or less of what its value should be according to the law of ancestral heredity. In going back two generations we come to fewer mares, to more in-and-in breeding, and to just the type of famous old mare, whose breeder kept her at the stud long after she was sterile. I expected accordingly a great and artificial fall in the fecundity of granddams and a double drop, something like $\frac{2}{5} \times \frac{2}{5}$, in the value of the regression as indicated by the law of ancestral heredity. This would reduce the apparent regression to about $\frac{2}{5} \times \frac{2}{5}$ of ·15, or to about ·025, say, a value about equal to the probable error of the table. The results actually reached are given in Table XV., and we find, if the subscript g refer to granddam :

$$M_m = \cdot 6345, \qquad M_g = \cdot 6232,$$
$$\sigma_m = \cdot 2040, \qquad \sigma_g = \cdot 1687,$$
$$r_{mg} = \cdot 0169.$$

The coefficient of regression = ·0204.

---

* In the first series the mares' names run from A to G; in the second from G to M, with 300 additions made to the A to G series, while I was completing my alphabet.

The probable error of the correlation is ·0213, and of the regression ·0213. Thus these results are not significant in themselves, but they are exactly what we might expect on the above hypothesis. Taken with the other five tables which we have worked out for the inheritance of fecundity, they are significant, for every one of them gives a *positive* correlation, however small it be, and thus adds to the accumulated evidence that fecundity is a heritable character.

(vi.) It remains to test our results by the theory developed on pp. 269 *et. seq.* But a difficulty comes in here. Turning to (xviii.) and (xix.) on p. 268, we cannot feel justified in putting $M_1 = M_2$, for there is a secular difference in the fecundity of mares and dams, owing to the fecundity of the older brood-mares being based on a longer period and liable to the disturbing causes so markedly manifest in the correlation of mares and granddams (see my remarks, p. 305). If we combine (xviii.) and (xix.) we find

$$M''_2 - M'_2 = r\frac{\sigma_2}{\sigma_1}(M''_1 - M'_1).$$

Now r is small, and it will accordingly be legitimate to put $M'_1 = M'_2$ and $\sigma_1 = \sigma_2$ on the right, we have then

$$(M''_2 - rM''_1)/(1 - r) = M'_2.$$

From this we deduce for the results in (ii.) on p. 304

$$M'_2 = \cdot 6321.$$

Turning now to (xix.), it may be written

$$M''_1 = M'_1\left\{1 + \frac{\sigma_1^2}{M_1^2}\left(\frac{\sigma'^2_1/\sigma_1^2}{1 + \sigma_1^2/(\rho^2 M_1^2)}\right)\right\}.$$

The second term in the curled brackets is small, and in it we may put to a first approximation $\sigma'_1 = \sigma_1 = \sigma_2$ and $M_1 = M'_2$. We then have

$$M''_1 = M'_1\left\{1 + \left(\frac{\cdot 1888}{\cdot 6321}\right)^2\left(\frac{1}{1 + \left(\frac{\cdot 1888}{\cdot 6321}\right)^2 \frac{1}{(\cdot 5132)^2}}\right)\right\},$$

or,

$$M''_1 = M'_1 \times 1\cdot 0666.$$

Substituting the value of M''_1 we find

$$M'_1 = \cdot 6118.$$

We thus see a difference in the fecundities of the unweighted dams and unweighted mares of ·6118 and ·6321, or about 2 foals more in the hundred appear to survive in the later generation. This is very probably due to the causes already indicated as affecting the apparent fecundity of the older mares (see p. 298). The influence of reproductive selection changes these quantities to ·6337 in the case of the daughters, and to the *apparent* high fecundity of ·6525 in the case of the dams.

We can now find σ_2 to a second approximation by aid of (xxi.). In the small term multiplied by r, we put $\sigma_1 = \sigma_2 = \sigma''_2$. Hence we find

$$\sigma_2^2 = \sigma''^2_2 + r^2 (\sigma''^2_2 - \sigma''^2_1),$$

and deduce, on substituting the numerical values,

$$\sigma_2 = \cdot 1896,$$

or is scarcely different from σ''_2. We accordingly conclude that we may quite reasonably assume the variability of the mares to represent the variability of the mares without reproductive selection, but the effect of weighting the dams with their fertility is to reduce the variability of the dams from about ·1896, if there be no secular change, to an apparent value as low as ·1643.

The same formulæ applied to the slightly better results in (iv.) on p. 305 give us:

$$M'_1 = \cdot 6205, \text{ and } M'_2 = \cdot 6342.$$

If we pass back from M'_1 and M'_2 to M_1 and M_2 we find:

	First case.	Second case.
M_1	·5460	·5567
M_2 . . .	·6266	·6278

If these results be considered as valid, we notice a remarkable difference between the fecundity of the younger and elder generation. While the crude results on pp. 304 and 305 might lead us on first examination to suppose the elder generation more fecund than the younger, these results show us that it is distinctly less so. The greater part of the difference, however, is due, not to a secular change, but to the causes we have so often referred to as weakening the fecundity recorded for the older mares. At the same time the whole system of breeding is so artificial that we may well doubt whether our equations (i.) and (v.) can be legitimately applied. For the chance of a mare getting into the stud-book as a dam, *i.e.*, having daughters at the stud, depends less on her fertility than on the degree of fashion in her stock. Thus the record weighting with fertility is hardly a probable hypothesis, and the values just given for M_1 are, I suspect, much below what they should be. For the above reason I have not proceeded to consider the changes in variability connoted by (ii.) and (xxii.). As I have made no attempt to form a correlation table for mares and dams in which the dam would have only one daughter to her record, I cannot make any plausible guess at the real magnitude of the cubic summation term in (xxii.). Apart, however, from the numerical application of these variation formulæ to a somewhat doubtful case, we see in these formulæ the theoretical basis for the observed fact that the fecundity of mothers is far less variable than that of daughters. It is really only an apparent divergence, due to the fact that the mothers have been weighted with their fertility; this, while it increases the apparent mean of their fecundity, reduces its apparent variability.

(17.) *On the Inheritance of Fecundity in the Brood-mare through the Male Line.*

For the thoroughbred horse this problem is fairly easily answered by investigating whether mares related to the same stallion have any correlation between their fecundities. The two cases I have selected are: (i.) "Sisters," daughters of the same sire, but in general not of the same mare; and then (ii.) "Nieces" and "Aunts," or daughters of a sire and the daughters of his sire. As we have only 760 sires and nearly 5000 mares, the daughters or aunts fall into rather large arrays, and we are compelled to use the methods discussed in Proposition IV., A and B. Even so the arithmetical work for a correlation based on the index of sires was far more laborious than for one based on the index of mares.

(i.) *To find the Correlation between Half-Sisters, Daughters of the same Sire.*

Here we have to use formulæ (xxiii.), (xxv.), and (xxvi.) of pp. 272–273. In order to do this a table was formed of the mean fecundity M of the array of sisters due to each sire, and of $\frac{1}{2}n(n-1)$, the number of pairs of sisters in each array. Then the products $\frac{1}{2}n(n-1)\mathrm{M}$ and $\frac{1}{2}n(n-1)\mathrm{M}^2$ were formed, and the numerator of (xxiii.), or σ_a^2, calculated by adding up for all the 760 sires. The result gave:

$$\sigma_a^2 = \cdot 6655167,$$

where the unit is the fecundity group element of 1/15. The number of pairs of sisters dealt with was 54,305. The denominator $\sigma_0^2(1-\rho^2)+\sigma_a^2$ is not so easily ascertained. σ_0 is the standard deviation of all the series of mares who are sisters without weighting; $\sigma_0\sqrt{(1-\rho^2)}$ is the standard deviation of an array of sisters, or if the regression be not linear, the mean of such standard deviations for all arrays, or rather its square is the mean of the squares of such standard deviations; ρ is the correlation between a patent character in the daughter and a purely latent character in the sire, and cannot therefore be found directly.

In order to get an appreciation of the standard deviation of an array of sisters—it being practically impossible to work out these quantities for 760 arrays—I selected twenty sires having fairly large arrays of daughters, and reached the following results:

TABLE of Arrays of Mares, which are Half-Sisters.

Sire.	No. of mares.	Mean fertility.	S.D. of array.
Speculum	76	9·697	2·989
Sterling	52	10·750	2·545
Scottish Chief	67	9·201	3·176
Newminster	64	8·875	2·497
Parmesan	37	9·708	3·076
Macaroni	81	10·210	2·770
King Tom	53	9·689	2·748
Lord Clifden	41	9·878	2·086
Hermit	79	9·437	3·003
Blair Athol	87	9·057	2·752
Lord Lyon	32	9·125	3·314
The Duke	35	9·186	2·474
Doncaster	37	9·297	3·021
Adventurer	58	10·466	2·621
Cathedral	43	9·267	2·847
Rosicrucian	59	10·932	3·094
Stockwell	80	9·131	2·093
Rataplan	40	8·222	2·201
Y. Melbourne	55	9·064	2·301
Thormanby	41	9·951	1·651
Totals	1117	191·143	53·259
Mean	55·85	9·55715	2·66295
Ditto in actual units*	..	·6371	·1775

I next took the mean and standard deviation of the 1117 mares to obtain σ_0. The mean fecundity was now found to be 9·5685 and $\sigma_0 = 2·7824$, or in actual units ·6379 and ·1855. Clearly only about $\frac{1}{8}$ per cent. difference is made whether we take the mean fecundity of the 1117 mares, or the mean of the unweighted means of the twenty arrays. Knowing σ_0 and $\sigma_0\sqrt{1-\rho^2}$ we can now find ρ. We have almost at once

$$\rho = ·2900.$$

This is probably the first determination of a coefficient of inheritance between a latent character in one sex and a patent character in the other sex. We see that it has almost exactly the value required (·3) by the law of ancestral heredity, or we conclude, *mares inherit from their sires a fecundity governed closely by the law of ancestral heredity.*

If the reader asks why is not the intensity reduced in this case in the same manner that we find it reduced in the case of the inheritance from the dam, the reply is :

(i.) In the case of the dam and mare, *both* quantities to be correlated are liable to fictitious values. In the case of sire and mare, we deal with only one.

---

[*] A fecundity unit is taken to be 1/15, for this is the unit of grouping.

(ii.) The influence of fictitious values has been shown on pp. 276–277 to chiefly affect the coefficient of correlation and not the standard deviation.

Now the present result is based solely on the calculation of standard deviations, or on the variability of fecundity as a whole and in arrays. It is accordingly not influenced to nearly the same extent by the existence of fictitious values. Could we calculate the variability of the arrays of daughters due to individual mares, we should probably get a better result for inheritance in the female line.*

The above result is so satisfactory that I have little doubt that we have determined a very good value for $\sigma_0\sqrt{1-\rho^2}$. Substituting it we find for the correlation between half-sisters:

$$r = \frac{\cdot 66552}{7\cdot 09130 + \cdot 66552} = \cdot 0858.$$

The law of ancestral heredity gives for *half*-sisters $r = \cdot 2$, and $\frac{2}{5}$ of this $= \cdot 08$.

Thus we see that the collateral heredity between half-sisters, daughters of the same sire, is quite sensible, and is almost what we might have predicted would be the result, if we supposed correlation to be weakened, as in the previous cases, to $\frac{2}{5}$ of its value by fictitious records.

It is worth while to consider the amount of fictitious fecundity suggested by the reduction factor $\frac{2}{5}$. We have only to suppose the n_1/N of our p. 277 to be $\frac{2}{5}$. Now we may well assume the chance of a fictitious fecundity being recorded to be the same for either one of a pair of sisters; hence we shall have $p \doteq q$, and therefore, from the result on p. 276, we find $(p-1)^2/p^2 = \frac{2}{5}$. This gives us $(p-1)p = \sqrt{\cdot 4}$, and $(n_1 + n_3)/\mathrm{N}$ the fraction without fictitious values $= (p-1)/p = \cdot 6325$. Thus in order to introduce the reduction factor of $\frac{2}{5}$ by the occurrence of fictitious values of the fecundity, we should have to suppose about 37 per cent. of fictitious values to occur. This is, of course, a sort of average; many values will probably be only partially fictitious, *i.e.*, will to some extent approximate to their real values. Considering the very artificial character of the thoroughbred brood-mare, and the uncertainty of her treatment by breeders, this does not seem such an immense percentage that it would force us to the conclusion that the law of ancestral heredity cannot be true for the inheritance of fecundity.

(ii.) *To find the Correlation in Fecundity between the Sisters of a Sire and his Daughters.*

What we want is really the correlation between aunts and nieces, but they

---

* The standard deviations for the arrays of mares in Table XII. were indeed worked out for the twelve cases of dams from *e* to *q*. The mean of these cases was sensibly the same whether the simple mean, or the mean weighted with the numbers in the array was taken, and equalled 2·8091 or ·1823. This is $\sigma_0(1-r^2)^{1/2}$. But by p. 48, $\sigma_0 = \cdot 1888$, whence we deduce $r = \cdot 1375$, and the regression equals ·1581. Thus we have found a substantially larger value for r than that on p. 304 by dealing with variabilities, and not direct correlations. This gives additional evidence, if any were needed, of the inheritance of fecundity.

TABLE XVI.—Correlation Table for Weighted Mean Fecundities of Arrays of Stallion's Daughters and Stallion's Sire's Daughters. 138,424 cases of Aunts and Nieces. (See p. 311.)

Stallion's sire.		Stallion.																											Totals.
		c	d		e		f		g		h		i		j		k		l		m		n		p		q		
		3-3·5.	3·5-4.	4-4·5.	4·5-5.	5-5·5.	5·5-6.	6-6·5.	6·5-7.	7-7·5.	7·5-8.	8-8·5.	8·5-9.	9-9·5.	9·5-10.	10-10·5.	10·5-11.	11-11·5.	11·5-12.	12-12·5.	12·5-13.	13-13·5.	13·5-14.	14-14·5.	14·5-15.	15-15·5.	15·5-16.	16-16·5.	
e	4·5-5	..	..	..	..	..	..	..	..	..	..	·25	·25	..	..	..	..	..	..	..	..	..	·25	·25	..	..	..	..	1
	5-5·5	..	..	..	..	..	..	..	..	..	..	·25	·25	..	..	..	..	1	1	..	..	..	·25	·25	..	..	..	..	3
f	5·5-6	..	..	..	..	..	..	..	..	..	..	..	..	..	60·25	·25	..	1	1	..	..	..	..	..	..	..	..	..	62·5
	6-6·5	..	..	..	..	..	..	..	..	..	..	·75	·75	..	·25	3·25	..	..	1·5	1·5	..	..	..	..	..	..	..	..	8
g	6·5-7	..	..	..	..	..	·5	·5	..	..	..	·75	·75	..	..	3	..	..	1·5	1·5	..	..	..	..	..	..	..	..	8·5
	7-7·5	..	..	..	..	..	·5	·5	..	..	..	..	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	5
h	7·5-8	..	..	..	..	..	..	..	..	..	..	28	24	..	4·25	16·25	..	..	..	..	..	..	..	..	..	..	2	2	76·5
	8-8·5	..	..	..	..	..	..	..	4	4	44·25	79·25	31	47·5	8·25	15·75	110	..	..	1·5	1·5	·25	·25	..	..	..	2	2	351·5
i	8·5-9	..	..	..	..	..	..	..	22	22	57·25	315·25	328	126·5	241·75	2162·75	1·5	396	5	30·5	163·5	18·25	·25	11	11	..	..	..	3912·5
	9-9·5	9	9	..	20	49	37·5	8·5	8·5	145	18	15	332·5	276	217·75	770·25	3603	317·5	196·5	226·5	22	31·5	9·5	..	..	..	20	20	6362·5
j	9·5-10	3	3	..	66	66	80·5	64·5	1·5	137	179·5	235·25	874·5	1415·25	3782·5	9020	5316	5197·5	1702·5	274·5	139	459·5	4	3	2	..	32	32	29,090·5
	10-10·5	..	..	..	107	107	74·5	56	127	221·5	428·5	357·75	1903	3554·25	3945·5	24,628	7165	7030·5	6263·5	640·5	152	202	161·5	5·5	42	..	114	114	57,400·5
k	10·5-11	..	3	3	19·5	40	243	10·5	22	22	76	224	705·5	676	2482·5	8148	2205·25	1502·75	1383	2122·5	406·75	72·25	470·5	14·5	..	..	73·5	73·5	20,999·5
	11-11·5	29·5	29·5	..	6·75	35·75	29·75	·75	30	18	301	67	245·5	747	526	944	1862·25	3692·25	3283·5	214	145·75	173·25	164	65·5	·5	·5	5·5	5·5	12,623
l	11·5-12	3·5	3·5	..	·75	6·25	13·25	7·75	23·75	23·75	44	99	86·25	367·25	1476 25	1347·75	1113·75	1877·25	343·5	263	72·5	20·5	4	..	7·5	7·5	10	10	7232·5
	12-12·5	..	..	..	..	..	..	..	2·25	2·25	..	7·5	25·75	·75	2·75	7·25	106·75	3·75	..	1	..	..	..	..	..	..	1	1	162
m	12·5-13	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1·5	1·5	3	3	..	32	..	..	..	..	..	..	41
	13-13·5	..	..	..	..	..	..	..	1·5	1·5	..	7·5	1·75	1	2·5	1·75	..	..	22·75	1·75	..	..	..	..	..	..	3	3	48
n	13·5-14	..	..	..	·25	·25	..	..	..	..	..	..	·25	1	1	·25	..	..	27·25	·25	..	..	..	·75	·75	..	..	..	32
	14-14·5	..	..	..	·25	·25	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	·75	·75	..	..	..	2
p	14·5-15	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	15-15·5	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
q	15·5-16	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	·5	·5	..	..	..	..	..	..	..	1
	16-16·5	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	·5	·5	..	..	..	..	..	..	..	1
Totals ..	..	45	48	3	220·5	304·5	479·5	149	242·5	597	1148·5	1437·5	4562	7214·5	12,751·5	47,068·5	21,485	20,021	13,235·5	3783	1104	1009·5	814·5	101·5	64·5	8	263	263	138,424

are not "aunts" and "nieces" in the human sense, for the aunts are only half-sisters of the sire. By a process similar to that on pp. 408 and 409 of my paper on the "Law of Ancestral Heredity,"* I deduce that the correlation between a sire's sisters and daughters ought to be ·05, and not ·15 as in the case of Man. If this be weakened down to the $\frac{2}{5}$ of previous results, we should not expect a result differing much from ·02. As the variability of the elder generation is always less than that of the younger, we ought to expect a coefficient of regression of about this value. The theory used will be that of p. 273 of the theoretical part of this paper. The weighted mean fecundity found for the arrays of aunts and nieces was as follows :—

	Without grouping.	With grouping.
Arrays of aunts	·6195	·6199
Arrays of nieces	·6346	·6338

The grouping was done in fecundity units of $\frac{1}{2}$, *i.e.*, 1/30 change in fecundity. The agreement may accordingly be considered very good. The "aunts" are the daughters of the older sires, who owing to in-and-in breeding form a comparatively small group, and are the sires of mares belonging to the older period, whose fecundity is much weakened by causes already referred to. Their mean fecundity is slightly less than that of granddams, given on p. 305, while the mean fecundity of their nieces agrees well with that for the corresponding group of mares.

The method of grouping being adopted, a correlation table was formed for the mean fecundities of arrays of mares, daughters of a sire, and of arrays of mares, daughters of his sire. This is Table XVI. Here each mean is weighted with the number of pairs of aunts and nieces in the two arrays, *i.e.*, the extent of the data on which it is based. It represents accordingly 138,424 pairs of aunts and nieces. The following results were obtained, corresponding to 687 pairs of sires :—

Sire's Sire.	Sire.
$M_a = \cdot 6199$.	$M_{a'} = \cdot 6338$.
$\bar{\sigma}_a = \cdot 04344$.	$\bar{\sigma}_{a'} = \cdot 07609$.

$$R = \cdot 1174.$$

It will be at once noticed how much more variable are the array-means for the sire than for the sire's sire. The means of many of the sire's arrays are based upon small numbers, which would have been selected out, if we had gone to another generation as in the case of the sire's sire.

It will clearly not be legitimate in this case to put $\bar{\sigma}'_a = \bar{\sigma}_a$ as suggested on p. 274. There is probably no secular change of importance here, but the sire's sire requiring

---

* 'Roy. Soc. Proc.,' vol. 62.

three generations from the record is really more stringently selected than the sire with only two. We can now form $\bar{\sigma}$ and $\bar{\sigma}'$ by (xxviii.) and (xxix.), if we adopt suitable values of $\bar{\sigma}_0$ and $\bar{\sigma}'_0$, ρ, as we have seen, may with high probability be put equal to ·3 (p. 309). σ_0 for groups of daughters, on p. 309, is given as ·1855, but since this certainly included a fair number of what are now aunts, it must be somewhat too low for $\bar{\sigma}'_0$. We can well put $\bar{\sigma}'_0$ equal to the ·1888 of the mares on p. 304. $\bar{\sigma}_0$ for aunts cannot be as low as the standard-deviation of dams on that page, as many of the aunts may never appear in the record as granddams,* *i.e.*, they are less stringently selected. The mean of the two results for mares and dams may, perhaps, be taken as a close enough approximation for our present purpose, or $\bar{\sigma}_0 = ·1765$. We then deduce

$$\bar{\sigma} = ·1739, \qquad \bar{\sigma}' = ·1955.$$

If we compare the results now found with those for sisters cited on pp. 308 and 309, we find :—

	"Aunts."	"Sisters."	"Nieces."
M	·6199	·6371	·6338
σ_a	·0434	·0544	·0761
σ'	·1739	·1855	·1955

The accordances and divergences are much what we might expect, except in the case of σ_a. We should, *à priori*, have expected "sisters" to have approached nieces more nearly than aunts. The work has been gone carefully through, but I have not succeeded in finding any error. In the "nieces," of course, the weighting of an outlying fecundity-mean due to a sire with but few daughters, may still be large, if his sire have numerous daughters; this cannot occur in the case of "sisters," as the weighting depends only on the number in the array. The like heavy weighting cannot usually occur in the case of "aunts," for they are, as a rule (owing to selection to the third generation) daughters of old and famous sires, with plenty of material for basing averages upon. We do not get *many* "nieces" attached to "aunts," who are not daughters of famous sires. Such is probably the source of divergence in σ_a between nieces and sisters.

Using formula (xxviii.), on p. 274, we find

$$r' = ·0114,$$

and for the regression coefficient ·0128.

This value is much below the ·05 of the law of ancestral heredity, and below the reduced value ·02, which we might have expected to reach. Still, it again shows

---

* Every dam appears as a granddam, otherwise the fecundity of the daughter could not have been found.

positive correlation, and we may conclude that the patent character in the daughter is inherited latently through the male line.*

But there is another and far more significant method of looking at this result, namely, by considering the meaning of R on p. 274. We may treat the fecundity of daughters as really a character of the sire, and their mean fecundity as a measure of a latent character in him. R is then the correlation between a latent character in both a stallion and his sire, and we see that it is sensibly inherited for R = ·1174. To compare with the law of ancestral heredity, we must use the coefficient of regression, for the stallions are much more variable than their sires. We find

Regression of stallion on sire = ·2056,

which carries us a long way in the direction indicated by that law. Thus it is extremely probable that this law of inheritance applies not only to the inheritance of a patent character, or of a character latent in one sex and transmitted to a second, but also to the inheritance of a character latent both in the transmitter and receiver. The present method accordingly seems applicable to the inheritance of a character latent in two individuals, if we take the mean of the character, when patent in the offspring, as a measure of its strength in the individual in whom it is latent. If l_1 be the measure of a latent character in a parent, then the offspring will have a mean value $ql_1 + c_1$ of this character, where q is the coefficient of parental regression and c_1 a constant. If l_2 be the measure of the same latent character in a relative, then the offspring in this case will have $ql_2 + c_2$ of the character. But the correlation of l_1 and l_2 will be identical with that of $ql_1 + c_1$ and $ql_2 + c_2$, as I have shown elsewhere.† Thus the mean of the patent character in the offspring may be used to measure the correlation between latent characters in their parents.

To sum up our results for thoroughbred mares, we conclude that their fecundity, notwithstanding the imperfections and difficulties of the record, has been demonstrated to be inherited, and this, both through the male and female line, so far as we can judge, with an equal intensity. The *apparent* value of this intensity, except in the case of latent characters, is much below that required by the law of ancestral heredity, roughly, perhaps, 2/5 of that value; but there is considerable reason to think that this reduction may take place owing to the presence of fictitious values in the record arising from the peculiar circumstances under which thoroughbred horses are reared and bred. These fictitious values would hardly influence the means and variability of arrays like they must do the relationship between pairs of individuals. Hence, when we deal with such means and variabilities as in the cases on pp. 309 and 313, we find a much closer approach to the law of ancestral heredity. Fecundity is certainly inherited; that it is inherited according to the Galtonian law

---

* As a matter of fact, this conclusion is stronger than it appears here, for the correlation between nieces and aunts was worked out, without grouping, for fourteen distinct series, and in *thirteen* of them was found to be sensibly positive; in the fourteenth it was found to have an insignificant negative value.

† "On the Reconstruction of the Stature of Prehistoric Races," 'Phil. Trans.,' A, vol. 192, p. 183.

is not demonstrated, but may be treated as probable until the results of further investigations—preferably by breeding experiments instituted for this very purpose—are available.

(18.) *Conclusion.*—The investigations of this memoir have been to some extent obscure and difficult, but the general result is beyond question.

Fertility and fecundity, as shown by investigations on mankind and on the thoroughbred horse, are inherited characters.

The laws of inheritance of these characters are with considerable probability those already developed in my memoir on the Law of Ancestral Heredity for the inheritance of directly measurable organic characters.

In the course of the work it has been shown how a numerical measure may be obtained for the inheritance of a character by one sex from the other, when it is patent in the former and latent in the latter. Fertility and fecundity purely latent in the male (in the sense here used) are shown to be transferred by him from his mother to his daughter. Thus DARWIN's views with regard to the transmission through one sex of a character peculiar to the other are given a quantitative corroboration.*

When we turn from these points to their weight and importance for the theory of evolution, we are at once encountered by all the wide-reaching principles which flow from the demonstration that genetic (reproductive) selection is a true factor of development. Let us look at these a little more closely.

If natural selection were to be absolutely suspended, *i.e.*, if there were no differential death-rate at all, then development would not for a moment cease. Not only is fertility inherited, but there can be small doubt that it is closely correlated with all sorts of organic characters; thus the inheritance of fertility marks, the moment natural selection is suspended, a progressive change in a great variety of organic characters. Without a differential death-rate the most fertile will form in every generation a larger and larger percentage of the whole population. There are very few characters which may not be supposed to be more or less directly correlated with fertility, and in reproductive selection we see a cause of progressive change continuously at work.† There is, so to speak, in every species an innate tendency to progressive change, quantitatively measurable by determining the correlation coefficients between fertility and organic characters, and between fertility in the parents and in the offspring. This "innate tendency" is no mysterious "force" causing evolution to take place in a pre-ordained direction; it is simply a part of the physical organisation of the individual, which does not leave fertility independent of

---

* The method is perfectly general, and a value can always be found for the intensity of transmission of a sexual character through the opposite sex. We could obtain, for example, a numerical measure of the manner in which a bull transmits good milking qualities to its offspring.

† I have endeavoured to show ('Roy. Soc. Proc.,' vol. 59, p. 303), that fertility is correlated with stature in woman. I hope later to return to the correlation of fertility and physique.

physique and organic relationship, or leave these characters uncontrolled by the principle of heredity. It seems to me, therefore, that the results of this memoir force on us some modification of current views of evolution. The suspension of natural selection does not denote either the regression of a race to past types, as the supporters of panmixia suggest, or the permanence of the existing type, as others have believed. It really denotes full play to genetic or reproductive selection, which will progressively develop the race in a manner which can be quantitatively predicted when once we know the numerical constants which define the characters of a race and their relation to racial fertility. In other words, natural selection must not be looked upon as moulding an otherwise permanent or stable type; it is occupied with checking, guiding, and otherwise controlling a progressive tendency to change.

So soon as a species is placed under a novel environment, either artificially or naturally, the equilibrium is disturbed, and it will begin to progress in the manner indicated by genetic (reproductive) selection, until this progress is checked by the development of characters in a manner or to an extent which is inconsistent with fitness to survive in the new surroundings. Within a very few generations a novel environment, sympathetic so to speak to the progressive tendency indicated by reproductive selection, produces the suitable variations without the assistance of natural selection. It seems to me that this principle ought to be borne in mind when, in laboratory experiments or in artificial breeding, natural selection is wholly or largely suspended, or again is altered in type; the species dealt with is unlikely to remain constant for several generations, but will develop in the direction indicated by genetic selection. Further, when stable types of life like the English sparrow are taken to America, or the English rabbit to Australia, where initially they fill a more or less vacant field among living forms, and natural selection is in part suspended, we should expect in a few generations a considerable divergence in type.* The converse aspect of the problem is also of great importance; namely, the natural selection of physical characters must tend to indirectly modify fertility and fecundity, if these be correlated with those characters. Variations in the fertility of local races need not be looked upon as due directly to environment, but may arise from the selection of characters correlated with fertility, combined with the law that fertility is itself an inherited character.

Lastly, the inheritance of fertility involves the "acceleration" of fertility; a race, natural selection being suspended, tends not only to increase but to increase at an increasing rate. This principle is again full of meaning, not only for the study of the manner in which lower types of life rapidly expand under changed environment, but also for the problems set to those philosophers who may desire that the most social and not the most fertile type of citizen may predominate in our modern civilised communities, where the state and public opinion to a greater or less extent hinder natural selection from playing the great part it does in wild life.

---

* It would be interesting to know whether the size or frequency of the litter of the Australian rabbit is greater than that of the English.

Table I.—Correlation of Fertility of a Woman and of her Daughter. Marriages completed, or having lasted at least 15 years in Daughter's case only. 4418 cases. Fertility of Mother weighted. (See p. 281.)

		Mother's fertility.																			Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
Daughter's fertility.	0	28	42	70	98	109	90	71	81	67	52	29	22	3	3	4	3	1	..	2	775
	1	22	56	78	104	80	70	64	51	39	31	20	19	1	4	4	4	..	..	..	647
	2	21	39	63	65	78	73	47	47	40	29	20	11	2	2	..	2	..	..	2	541
	3	20	35	44	53	53	75	57	57	33	37	21	13	10	5	1	2	..	..	..	516
	4	17	27	58	52	69	62	53	46	41	24	16	12	11	6	1	..	..	..	..	495
	5	16	28	26	37	52	55	43	32	40	25	12	14	5	2	1	..	..	..	..	388
	6	13	20	26	42	41	39	30	38	28	19	5	8	3	2	1	..	..	..	1	316
	7	3	14	33	23	33	31	24	32	26	14	10	9	5	1	1	1	..	..	..	260
	8	6	9	15	16	25	16	22	15	20	11	6	6	6	..	1	3	..	..	..	177
	9	4	7	13	13	13	10	14	12	10	8	5	8	3	1	..	1	..	..	..	122
	10	..	13	2	8	10	7	13	7	8	6	1	3	1	1	..	..	..	..	..	80
	11	7	2	6	9	7	1	4	4	4	3	2	1	1	1	..	..	..	..	..	52
	12	..	1	2	1	4	4	..	2	1	..	2	1	2	..	..	..	..	..	..	20
	13	..	..	2	1	1	2	4	2	1	2	1	1	1	..	..	..	..	..	..	18
	14	..	1	..	1	..	2	1	..	..	1	..	..	..	..	..	..	..	..	..	6
	15	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	1
	16	1	2	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	4
Totals ..		158	296	439	523	575	537	447	426	358	263	150	128	54	28	14	16	1	0	5	4418

Table II.—Correlation between the Fertility of a Woman and that of her Daughter. No weight to the Fertility of the Mother. Marriage completed or having lasted at least 15 years in the Daughter's case only. 1000 cases from the Peerage. (See p. 282.)

		Mother's fertility.															Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Daughter's fertility.	0	3	6	18	15	15	12	11	5	6	3	1	1	..	..	1	97
	1	6	13	16	23	17	16	10	11	6	7	2	..	..	..	..	127
	2	6	7	17	22	17	15	9	14	7	4	2	..	..	..	..	120
	3	8	12	18	20	13	22	12	12	10	6	1	3	..	..	..	137
	4	5	6	16	15	21	19	22	8	12	5	3	4	1	..	..	137
	5	4	7	4	12	16	16	17	11	13	7	6	3	..	..	..	116
	6	1	9	8	12	13	9	13	9	6	6	2	4	1	..	..	93
	7	1	2	5	6	9	9	5	7	5	6	2	2	2	..	1	62
	8	..	1	6	2	8	6	5	5	5	3	2	..	2	..	..	45
	9	1	1	1	5	4	4	6	1	1	1	2	..	..	..	..	27
	10	..	3	1	2	2	2	..	4	3	2	..	1	..	1	..	21
	11	..	..	1	2	1	1	2	..	..	..	1	1	..	..	..	9
	12	..	..	..	..	2	..	..	..	..	..	..	..	2	..	..	4
	13	..	..	..	..	..	1	2	..	..	..	..	..	1	..	1	5
Totals ..		35	67	111	136	138	132	114	87	74	50	24	19	9	1	3	1000

TABLE III.—Correlation between the Fertility of a Woman and that of her Daughter. No weight to Fertility of the Mother. Marriage completed or having lasted at least 15 years in Daughter's case only. 1000 cases, largely "Heiresses" from the Landed Gentry. (See p. 283.)

		Mother's fertility.																					Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
Daughter's fertility.	0	20	12	9	6	18	16	14	15	7	8	3	5	1	..	..	..	..	..	..	.	..	134
	1	16	10	17	23	12	11	11	7	3	7	4	..	..	..	..	..	..	..	..	..	..	121
	2	6	11	11	9	18	12	18	9	2	2	1	2	..	1	..	..	..	..	..	..	..	102
	3	15	9	8	13	19	10	6	10	4	4	1	..	..	1	..	..	..	..	..	..	..	100
	4	23	8	7	17	11	9	9	11	7	3	2	3	2	..	..	..	..	..	..	..	..	112
	5	17	8	8	16	13	10	12	10	6	6	2	..	3	..	..	..	..	..	..	..	..	111
	6	12	3	2	15	9	8	8	7	3	4	5	1	..	2	1	..	..	..	..	..	..	80
	7	14	3	11	3	7	7	7	9	4	3	3	..	3	1	1	..	..	..	..	..	..	76
	8	9	5	4	6	6	6	5	2	3	2	2	2	..	1	..	2	..	..	..	..	..	55
	9	10	3	1	2	1	2	5	6	..	1	1	2	1	2	..	..	..	..	..	..	..	37
	10	2	..	3	3	4	2	5	6	1	2	..	2	..	..	1	..	..	..	..	..	..	31
	11	4	..	2	..	3	2	1	3	1	..	..	..	..	..	..	..	..	..	..	..	..	16
	12	1	1	1	1	2	1	..	..	2	1	1	1	1	..	..	..	..	..	..	..	..	13
	13	..	..	..	..	2	..	1	2	..	..	..	..	1	..	..	..	..	..	..	..	1	7
	14	..	..	..	..	1	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	2
	15	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	1
	16	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	17	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	18	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	19	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	20	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	21	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
	22	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
Totals ..		149	73	84	114	126	98	103	97	43	43	26	18	12	8	3	2	0	0	0	0	1	1000

TABLE IV.—Correlation between the Fertility of a Woman and that of her Mother, when the Marriages of both have lasted at least 15 years and no weight is given to the Fertility of the Mother. 1000 cases from the Peerage. (See p. 285.)

		Mother's fertility.																	Totals.
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Daughter's fertility.	0	..	5	9	11	18	21	15	8	9	6	3	2	3	..	..	..	..	110
	1	..	12	5	14	15	10	13	9	8	5	3	2	2	..	..	..	..	98
	2	..	9	9	10	15	18	15	9	3	2	4	2	..	..	..	..	1	97
	3	..	5	10	16	11	9	14	13	10	4	8	2	3	..	..	..	..	105
	4	..	5	5	19	17	21	15	18	10	14	2	1	5	1	..	..	..	133
	5	..	7	6	7	17	23	9	12	13	14	8	3	2	2	..	..	..	123
	6	..	4	5	8	11	15	12	15	14	7	5	3	3	1	..	..	..	103
	7	..	5	4	3	8	4	13	9	8	5	10	2	1	1	..	..	..	73
	8	..	1	2	4	12	9	9	8	5	12	3	4	1	2	1	..	..	73
	9	..	..	..	4	3	3	4	7	5	3	2	2	1	..	..	..	..	34
	10	..	..	..	1	2	1	3	4	6	3	2	..	1	..	1	..	..	24
	11	..	..	..	2	1	1	1	..	..	1	2	..	..	..	..	..	..	8
	12	..	..	2	1	2	3	..	1	1	..	..	1	..	1	..	1	..	13
	13	..	..	..	..	..	2	1	..	..	..	..	1	..	2	..	..	..	6
	14	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	15	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	16	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
Totals ..		0	53	57	100	132	140	124	113	92	76	52	25	22	10	2	1	1	1000

TABLE V.—Correlation of the Fertility of a Man and of his Son as shown in Marriage. The Marriage of the Son being completed or having lasted at least 15 years. Weight given to the Fertility of Father. 6070 cases. (See p. 287.)

	Number of father's children.																								Totals.
Number of son's children.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
0	..	36	67	79	78	97	86	70	72	36	35	19	8	8	5	10	2	..	..	1	..	..	1	..	710
1	..	62	89	114	104	97	77	81	69	50	25	24	11	7	1	6	4	..	1	..	..	..	..	1	823
2	..	58	75	87	106	103	85	75	68	41	33	23	16	15	2	2	5	..	..	..	..	1	..	..	795
3	..	39	84	89	98	87	87	83	65	54	40	28	10	8	7	6	1	1	1	2	..	..	..	..	790
4	..	38	66	84	90	89	76	73	68	51	27	22	15	14	5	2	3	1	..	1	..	..	..	..	725
5	..	37	53	70	81	68	58	66	61	41	31	29	17	5	3	4	1	..	..	..	..	..	..	..	625
6	..	22	43	55	59	67	41	41	49	40	17	22	9	5	1	..	3	1	..	..	..	1	..	..	476
7	..	18	34	35	30	34	51	39	42	33	24	15	5	5	1	2	..	..	..	1	..	..	..	..	369
8	..	18	20	26	36	36	28	30	25	17	12	12	7	3	1	2	2	..	..	1	..	..	..	..	276
9	..	10	12	22	24	27	21	24	14	11	9	4	5	1	2	..	1	..	..	2	..	..	1	..	190
10	..	4	11	17	13	18	14	10	12	14	6	4	2	2	..	..	..	..	..	..	..	..	..	..	127
11	..	4	7	6	7	18	7	3	3	3	6	3	3	..	1	1	..	..	..	1	..	..	..	..	73
12	..	4	..	3	5	6	2	2	3	1	5	2	1	..	1	..	..	..	..	..	..	..	..	..	35
13	..	..	..	3	1	4	2	1	3	2	..	3	..	..	..	..	..	..	..	..	..	..	..	..	19
14	..	..	1	..	2	2	2	4	..	..	1	1	..	..	..	..	..	1	..	..	..	..	..	..	14
15	..	..	..	..	2	..	1	..	1	..	..	2	..	..	2	..	..	..	..	..	..	..	..	..	8
16	..	..	1	..	1	..	..	1	1	..	..	..	1	..	.	..	..	..	..	..	..	..	..	..	5
17	..	..	..	..	..	1	1	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	3
18	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	1
19	..	..	..	..	..	..	..	1	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	2
20	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
21	..	..	..	..	1	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	2
22	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
23	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
24	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
Totals ..	0	351	563	690	739	754	640	604	557	394	272	214	110	73	32	35	22	4	2	9	0	2	2	1	6070

TABLE VI.—Correlation of the Fertility in Marriage of a Man and his Son; no weight being given to the Fertility of the Father, and the Marriage of the Son being completed or having lasted at least 15 years. 1000 cases from the Peerage. (See p. 287.)

		Fertility of the father.																	Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Fertility of sons.	0	5	8	7	14	18	2	2	3	8	3	4	4	..	..	..	..	..	78
	1	3	3	6	5	8	8	6	5	4	..	2	..	1	..	..	..	..	51
	2	7	5	6	13	12	12	12	6	5	4	2	1	1	..	..	..	..	86
	3	5	10	13	11	17	13	13	12	10	4	1	2	1	..	..	..	..	112
	4	4	16	18	24	23	5	18	10	7	8	1	5	2	1	..	..	1	143
	5	9	8	11	14	16	12	16	12	2	5	8	6	3	..	2	..	..	124
	6	3	4	10	16	13	11	11	10	10	1	2	2	1	..	..	..	..	94
	7	5	6	8	7	10	14	11	12	4	7	1	1	..	..	..	..	..	86
	8	3	5	4	15	19	7	10	8	4	2	2	1	..	..	1	..	..	81
	9	1	6	5	9	5	5	8	5	4	3	3	1	..	..	..	..	..	55
	10	2	3	9	3	2	5	4	6	4	3	1	1	1	..	..	..	..	44
	11	1	1	1	4	2	1	1	2	1	1	..	..	..	..	..	..	..	15
	12	..	..	2	2	2	..	1	1	1	2	..	1	..	..	..	..	..	13
	13	1	..	..	1	3	..	1	1	1	1	..	1	..	1	..	..	..	10
	14	..	1	..	..	..	1	..	..	..	..	1	..	..	..	..	..	..	3
	15	..	..	1	..	..	..	..	1	..	..	1	..	..	..	..	..	..	3
	16	..	..	..	..	..	..	..	..	1	1	..	..	..	..	..	..	..	2
Totals ..		49	76	101	138	150	96	114	94	66	45	29	26	10	2	3	0	1	1000

TABLE VII.—Correlation of the Fertility in Marriage of a Man and his Son ; no weight being given to the Fertility of the Father, and the Marriages of both Father and Son having lasted at least 15 years. 1000 cases from the Landed Gentry. (See p. 288.)

		Fertiliy of father.																Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Fertility of the son.	0	1	1	..	11	9	6	4	5	6	2	1	..	2	..	..	..	48
	1	..	5	7	11	5	8	7	2	7	1	..	2	..	1	..	..	56
	2	2	6	9	12	8	7	8	4	1	4	3	1	1	..	..	..	66
	3	5	8	15	14	11	10	10	10	7	5	3	4	..	..	..	..	103
	4	4	9	14	22	16	21	10	12	11	7	1	8	..	2	1	..	138
	5	5	7	12	19	30	12	17	13	12	10	3	3	2	..	..	..	145
	6	5	4	9	12	17	16	20	14	8	8	5	4	..	3	..	..	125
	7	..	4	5	23	13	12	16	8	9	6	3	2	..	2	1	..	104
	8	3	2	5	7	8	12	7	11	8	5	4	3	..	1	..	1	77
	9	1	..	6	8	3	7	5	6	3	1	2	1	..	..	..	1	44
	10	..	1	5	7	4	3	4	7	1	4	2	1	..	1	1	..	41
	11	..	1	1	3	3	4	5	2	4	2	2	..	..	..	..	..	27
	12	..	..	1	..	2	1	..	4	..	3	1	..	..	2	..	..	14
	13	..	..	2	1	1	1	1	..	..	..	1	..	..	1	..	..	8
	14	..	..	..	1	1	..	..	..	..	..	..	..	..	..	..	..	2
	15	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	16	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	1
	17	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	1
Totals ..		26	48	91	151	132	120	114	98	78	58	31	29	5	13	3	3	1000

TABLE VIII.—Correlation of the Fertility of a Woman and her Paternal Grandmother. No weight being given to the Fertility of the Grandmother, and the Marriages of both having lasted for at least 15 years. 1000 cases from the Peerage and Baronetage. (See p. 289.)

		Fertility of paternal grandmother.																					Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
Fertility of granddaughter.	0	6	12	14	11	15	15	8	5	6	2	2	2	1	1	..	..	..	..	..	..	1	101
	1	10	10	7	13	13	5	7	4	2	7	1	1	..	..	..	..	..	..	..	..	..	80
	2	6	7	20	14	11	10	11	6	6	4	2	1	..	..	..	1	..	..	..	..	..	99
	3	9	15	12	20	13	9	15	14	10	1	..	5	..	..	..	1	..	..	1	..	..	125
	4	7	13	16	19	19	11	11	13	8	2	5	4	..	..	..	1	..	..	..	..	..	129
	5	4	9	10	14	20	15	16	12	9	9	5	2	..	..	..	..	..	..	..	..	..	125
	6	5	10	17	9	16	10	8	8	6	5	5	1	2	1	..	..	..	..	..	..	..	103
	7	7	4	8	7	8	9	10	15	10	5	2	1	2	..	..	..	..	..	..	..	..	88
	8	6	7	5	5	3	6	9	12	6	1	2	1	..	..	..	1	..	..	..	..	..	64
	9	3	3	3	4	4	4	1	2	6	..	1	..	2	..	..	..	..	..	1	..	..	34
	10	2	1	4	2	3	1	1	6	1	5	1	..	..	..	..	..	..	..	..	..	..	27
	11	1	..	2	2	..	2	1	..	..	..	1	..	..	..	..	..	..	..	..	..	..	9
	12	..	..	..	2	..	5	1	1	1	..	1	1	..	..	..	..	..	..	..	..	..	12
	13	1	..	..	1	..	..	..	..	1	..	..	1	..	..	..	..	..	..	..	..	..	4
Totals ..		67	91	118	123	125	102	99	98	72	41	28	20	7	2	0	4	0	0	2	0	1	1000

TABLE IX.—Correlation between the Apparent Fertilities of Mares and Dams, with a minimum of at least four coverings. 1100 cases. (See p. 291.)

		Mare's foals.																			Totals.
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Dam's foals.	1	..	..	..	..	..	1	..	1	..	..	1	..	..	2	..	..	..	..	..	5
	2	..	1	8	7	4	4	6	5	3	2	1	2	3	..	..	..	..	..	..	46
	3	..	2	5	5	3	5	10	7	5	3	4	1	2	3	..	1	..	..	..	56
	4	..	1	5	7	12	15	13	9	13	8	7	4	4	4	2	3	1	1	..	109
	5	..	2	6	10	7	13	18	7	11	8	6	10	2	1	3	1	..	1	..	106
	6	..	7	10	17	16	14	6	9	8	10	7	2	5	1	4	..	1	..	..	117
	7	1	2	15	17	22	18	10	10	14	8	6	2	2	..	1	..	1	..	1	130
	8	1	1	7	17	20	19	11	14	16	6	4	9	2	2	..	..	..	..	..	129
	9	1	1	8	15	6	6	7	8	6	1	6	3	2	2	1	..	..	..	..	73
	10	..	2	4	9	15	16	12	9	4	5	6	2	2	3	1	..	..	..	..	90
	11	..	2	7	14	11	7	10	13	8	4	..	3	2	..	1	..	..	..	..	82
	12	..	2	6	7	6	11	8	8	5	..	1	2	..	1	..	..	..	..	..	57
	13	..	..	2	4	9	3	3	5	3	4	1	..	..	..	..	..	..	..	..	34
	14	..	1	4	3	6	5	5	5	2	2	4	..	..	..	..	..	..	..	..	37
	15	..	2	..	1	1	1	4	1	1	3	..	..	1	..	1	..	..	..	..	16
	16	..	..	..	..	..	..	1	1	3	..	..	1	..	..	1	..	..	..	..	7
	17	..	..	..	..	..	2	..	1	1	1	..	1	..	..	..	..	..	..	..	6
	18	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
Totals ..		3	26	27	133	138	140	124	113	103	65	54	42	27	19	15	5	3	2	1	1100

Table X.—Correlation between the Apparent Fertility of Mares and Dams, with a minimum of at least eight coverings. 1100 cases. (See p. 292.)

		Mare's foals.																		Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Dam's foals.	1	..	..	..	..	1	..	..	..	..	..	..	..	1	..	..	..	..	..	2
	2	..	..	1	1	2	..	2	1	1	..	..	..	..	..	..	..	..	..	8
	3	1	..	1	..	1	2	4	..	2	..	2	..	2	.	..	1	..	..	16
	4	..	1	1	1	7	7	6	10	5	3	8	2	5	..	2	1	1	..	60
	5	..	1	1	5	7	15	9	13	12	1	11	4	3	3	..	2	1	..	88
	6	..	1	3	9	17	16	14	15	12	16	5	5	4	5	..	1	..	..	123
	7	2	2	7	16	18	17	24	22	19	9	9	1	3	3	1	1	..	1	155
	8	2	1	6	6	18	21	25	26	15	7	15	5	6	2	1	1	..	..	157
	9	..	1	6	4	9	10	20	9	4	9	7	3	3	5	..	..	..	..	90
	10	..	..	3	8	14	17	16	8	9	9	7	2	4	1	1	..	..	..	99
	11	..	1	3	4	10	19	18	14	7	1	6	2	1	1	..	..	..	..	87
	12	..	1	2	6	11	11	14	12	6	6	4	1	1	1	..	..	..	..	76
	13	..	..	1	4	7	11	8	9	9	1	3	1	..	..	..	..	..	..	54
	14	..	2	2	1	4	7	6	3	3	4	2	..	..	..	..	..	..	..	34
	15	..	..	..	1	2	5	2	4	6	1	..	2	..	1	..	..	..	..	24
	16	..	..	..	1	..	3	2	6	..	..	1	..	..	1	..	..	..	..	14
	17	..	..	..	..	2	..	2	1	1	..	2	..	1	..	..	..	..	..	9
	18	..	..	..	..	..	2	..	1	..	..	..	..	..	1	..	..	..	..	4
Totals..	..	5	11	37	67	130	163	172	154	111	67	82	28	34	24	5	7	2	1	1100

TABLE XI.—Correlation between the Fecundity and the *Apparent* Fertility of Brood-mares. 1000 cases. (See p. 300.)

		Apparent fertility.																			Totals.
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Fecundity.	*a*	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	2
	b	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
	c	..	2·5	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	3·5
	d	..	5·5	2	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	8·5
	e	..	13	5	3·5	2	..	1	..	..	..	..	..	..	..	..	..	..	..	..	24·5
	f	.	..	10	4·5	9	4	3	..	..	..	..	..	..	..	..	..	..	..	..	30·5
	g	..	..	16	13	6	13	9	5	1	1	..	..	..	..	..	..	..	..	..	64
	h	..	..	17·5	5	13	12·5	11	10	6·5	5·5	1·5	..	..	..	..	..	..	..	..	82·5
	i	..	..	16·5	5	6	19·5	23	14	17·5	13·5	6·5	3	1	..	..	..	..	..	..	125·5
	j	..	..	..	21	21	21	11	17	18	4	21	6	2	3	..	..	..	..	..	145
	k	..	..	4	..	22	..	9	21·5	12	18	5	15	9	2	3·5	..	..	..	..	121
	l	..	1	1	45	..	13	19	10·5	11	12	6	5	8	9	3·5	2	2	..	..	148
	m	..	..	..	..	31	13	..	10	17	10	12·5	7	3	4	6	3·5	..	1	..	118
	n	..	..	..	..	..	13	9	10	10	1·5	3·5	2	3	..	1	2·5	2	1	·5	59
	p	..	..	..	..	..	..	..	..	..	1·5	4	5	2	2	1	..	1	1	·5	18
	q	..	..	..	..	14	9	5	10	4	2	..	..	2	2	1	..	..	..	..	49
Totals ..		2	23	73	97	125	118	100	108	97	69	60	43	30	22	16	8	5	3	1	1000

TABLE XII.—Correlation Table for 1200 Mares and their Dams. Fecundity with a minimum of four coverings. (See p. 303.)

		Mares.																Totals.
		a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*	*p*	*q*	
Dams.	*a*	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	b	..	..	..	..	..	..	..	..	..	..	..	1	...	..	..	..	1
	c	..	..	..	..	..	·5	..	..	1	·5	..	..	..	..	..	..	2
	d	..	..	..	..	..	·5	..	1	1	2·5	..	3	1	..	..	..	9
	e	..	..	..	..	2·5	1	1·5	1	3	2·5	..	2·5	1	·5	·5	·5	16·5
	f	..	..	1	..	·5	1	4·5	2·5	2·5	·5	2	3·5	3	1	1	1·5	24·5
	g	..	..	..	1	2	3	5	2·5	10·5	10	3	6·5	1	1	1	4	50·5
	h	·5	..	·5	·5	3	5·5	9	10·25	10·25	14·5	14·25	15·75	13·5	4	3	6·5	111
	i	·5	..	..	1	4·5	3	12	14·25	19·25	21·5	17·25	18·25	10	5·5	1·5	4·5	133
	j	1	..	..	5	3	3	10	11	18·5	24	34	26	21·5	8	3·5	12	180·5
	k	..	..	1	1·5	5·5	4	14	17	21	19	28·75	31·75	23	13·25	3·75	8·5	192
	l	1	1	1·5	3	6·5	7	11	13·5	17·5	32	25·75	27·75	15·5	11·25	2·25	9·5	186
	m	..	..	·5	3	3	7·5	5·5	9·25	13·75	19·5	18·5	24	14	8	3	10	139·5
	n	..	..	·5	..	1·5	4·5	5·5	7	6·5	10·5	9·5	13	11·5	9·5	1	7·5	88
	p	..	..	..	..	1	·5	..	1·25	2·25	6	2·5	5·5	4	..	..	2·5	25·5
	q	..	..	.	1	1	..	4	3·5	4·5	9	5	4	3·5	2·5	1	2	41
Totals ..		3	1	5	16	34	41	82	94	131 5	172	160·5	182·5	122·5	64·5	21·5	69	1200

TABLE XIII.—Correlation between Brood-mares and their Dams with regard to Fecundity with a minimum of eight coverings. 1000 cases. (See p. 304.)

	Fecundity of mares.																
Fecundity of dams.	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*	*p*	*q*	Totals.
a	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
b	..	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	1
c	..	..	..	..	..	·5	..	..	1	·5	..	..	..	..	..	..	2
d	..	..	..	..	..	·5	..	..	1	·5	..	1	..	..	..	..	3
e	..	..	..	..	..	1	1	1	..	1	·5	2	..	..	..	..	6·5
f	..	..	1	..	..	1	4	3·5	2·5	1	3·5	2	2	..	1	..	21·5
g	..	..	..	..	1·5	2·5	3	2	11	10	7·5	4·5	4	4	2	..	52
h	..	..	·5	..	2·5	5	8	10	9·5	17	18·25	10·25	10·5	5	3·5	..	100
i	..	..	1·5	2	2	2·5	9	12·5	22	22	24·75	16·75	15·5	7	2·5	2	142
j	..	..	..	1	1·5	4·5	9	8	17	25	34	22	20·5	6	4·5	4	157
k	..	..	·5	·5	2·5	3·5	5·5	14·25	22·25	19·5	31·5	31	19	9·5	6	2	167·5
l	..	1	·5	1·5	5	4	5	14·25	19·25	25·5	26	20	17	5·5	3	2	149·5
m	..	..	·5	3·5	..	2	3	9·25	13·25	13·5	12	17·5	8·5	5·25	3·75	1	93
n	..	..	1	..	..	6	5	6·5	8	13	8·5	12	11·5	3·75	2·25	2	79·5
p	..	..	..	..	..	·5	..	2·25	2·25	7·5	2	4	2·5	1·5	..	..	20·5
q	..	..	..	..	..	..	..	·5	·5	2	1	..	..	..	..	1	5
Totals ..	0	1	5·5	8·5	15	33·5	52·5	82	129·5	158	169·5	144	111	47·5	28·5	14	1000

TABLE XIV.—Correlation of the Fecundity of Brood-mares and their Dams with a minimum of four coverings. Second series. 1200 cases. (See p. 305.)

	Fecundity of mares.																Totals.
	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*	*p*	*q*	
Fecundity of dams. *a*	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
b	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
c	..	..	..	..	..	..	..	1·25	·25	..	1	..	..	..	..	..	2·5
d	..	..	..	..	..	1	..	·75	·75	..	..	..	..	..	..	..	2·5
e	..	..	..	..	..	1	1	·25	·25	..	·5	1·5	·5	1·5	..	..	6·5
f	..	..	..	..	1	..	5	3·25	2·25	5	2·5	1·5	1·5	·5	1	..	23·5
g	..	..	..	..	2·5	1·5	6	5	8	9	7	7·5	6	7	1	..	60·5
h	·5	..	1·25	1·25	4	3	6	8·25	9·75	19	11	9·5	13·25	7·5	·25	3	97·5
i	1·5	..	2·25	3·25	3	3	6	9·25	14·75	30	15	15·5	16·75	6	2·25	5	133·5
j	..	..	·5	3·5	3·5	4·5	16	12·5	20·5	27	27·5	28·5	28	11	2	6	191
k	2	..	..	·5	5	4	13	18·25	28·75	23·5	20·25	25·75	23	12·75	2·75	13·5	193
l	1	..	..	2·5	6	3	4	22·25	23·75	24	27·75	18·25	22·5	13·25	4·75	13·5	186·5
m	..	..	·5	2	4·5	5	10	7·75	10·25	18	15	31	12	9·5	2	5·5	133
n	1	..	1·5	1	2·5	5	7	5·5	7	14·5	11·5	10·5	9·25	4·75	1·5	8	90·5
p	..	..	..	1	1	..	..	·25	1·25	6	2·5	5·5	3·25	3·25	..	2·5	26·5
q	..	..	..	..	..	1	2	2·5	5·5	6	9	10	7·5	2	2·5	5	53
Totals ..	6	0	6	15	33	32	76	97	133	182	150·5	165	143·5	79	20	62	1200

TABLE XV.—Correlation of the Fecundity of Brood-mares and of their Granddams, with a minimum of four coverings. 1000 cases. (See p. 305.)

		Fecundity of mares.																Totals.
		a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*	*p*	*q*	
Fecundity of granddams.	*a*	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	b	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	1
	c	..	..	..	..	·5	..	·5	·75	1·25	·5	..	1	·5	..	..	..	5
	d	..	..	..	..	·5	..	·5	·25	·75	1·5	1	1	·5	..	..	..	6
	e	..	1	..	..	..	·5	..	·5	1·5	1·5	..	3	..	·5	..	..	8·5
	f	..	..	..	..	..	1·5	..	2·5	3·5	5·5	4·5	3·5	6·5	2	..	1	30·5
	g	1	..	..	4	3	4	4	5	7	7	7·5	15·5	12·5	3·5	..	6	80
	h	..	..	1	1·5	4·5	4	6	7·5	12·5	11	11·5	19	10·25	5·25	..	5·5	99·5
	i	..	..	2	2·5	7	5·5	6	13·5	14·5	18·5	15·5	29	13·75	8·75	1	8·5	146
	j	1	..	2	3	7	6	8	10·5	14·5	16	18	22	19	11·5	·5	17	156
	k	1	..	1	3	4	7	11·5	11·25	15·25	25	10·75	21·75	10·5	14	1	13	150
	l	..	..	..	2	2	2	7·5	8·75	12·75	25·5	10·75	20·75	16	5	1·5	9	123·5
	m	..	..	..	..	7	4	3·5	8	13	5	11	11	14	5·5	·5	6·5	89
	n	1	..	..	·5	2	3	4·5	3·5	6·5	8·5	4·5	6	4	2·5	..	3·5	50
	p	..	..	..	·5	..	..	1	1·5	1·5	1·5	·5	4	1·5	1	..	4	17
	q	1	..	1	3	..	..	4	1	3	4	1·5	8·5	4·5	4·5	1	1	38
Totals ..		5	1	7	20	37·5	37·5	57	74·5	107·5	131	98	166	113·5	64	5·5	75	1000

Sold by Harrison and Sons, St. Martin's Lane.
Demy 8vo, pp. 224. Price 7*s*. 6*d*.

THE RECORD OF THE ROYAL SOCIETY, 1897.

Containing an Account of the Foundation, History, &c., of the Society.

Demy 8vo, pp. 179. Price 5*s*.

THE YEAR-BOOK OF THE ROYAL SOCIETY, 1899.

Containing an Account of the Chief Events and Arrangements of the Year.

PHILOSOPHICAL TRANSACTIONS.

Series A. Containing Papers of a Mathematical or Physical Character.
Vol. 190 (1897), Price £1 8*s*. 6*d*., with 17 Plates.
Vol. 191 (1898), Price £1 16*s*., with 23 Plates.

Series B. Containing Papers of a Biological Character.
Vol. 189 (1897), Price 17*s*. 6*d*., with 22 Plates.
Vol. 190 (1898), Price £2 5*s*., with 24 Plates.

Published by Clay and Sons, Ave Maria Lane.

CATALOGUE OF SCIENTIFIC PAPERS,

COMPILED BY THE ROYAL SOCIETY.

The volumes on sale as follows:—Vols. 1–6 (1800–63), cloth (vol. 1 in half-morocco) £4 net; half-morocco, £5 5*s*. net. Vols. 7, 8 (1864–73), cloth, £1 11*s*. 6*d*. net; half-morocco, £2 5*s*. net. Single vols., cloth, 20*s*.; half-morocco, 28*s*. net. Vols. 9, 10 (1874–83), cloth, £1 5*s*. net; half-morocco, £1 12*s*. net. Vol. 11 (1874–83), cloth, £1 5*s*. net; half-morocco, £1 12*s*. net.

A reduction of price to Fellows of the Royal Society

Published by Kegan Paul, Trench, Trübner and Co.

A few copies at the reduced price of £1. 1*s*. (published at £3).

Royal 4to, pp. iv.—936 cloth.

A MONOGRAPH OF THE HORNY SPONGES.

By R. von LENDENFELD

With 51 Lithographic and Photographic Plates.

A further reduction of price to Fellows of the Royal Society.

Published by Kegan Paul, Trench, Trübner, and Co.

In 1 vol., 4to. Pp. 500. With 6 Chromolithographs of the remarkable Sunsets of 1883 and 40 Maps and Diagrams.

THE ERUPTION OF KRAKATOA AND SUBSEQUENT PHENOMENA.

Report of the Krakatoa Committee of the Royal Society.

Edited by G. J. SYMONS, F.R.S.

Price 30*s*. To Fellows of the Royal Society, 20*s*.

Published by Kegan Paul, Trench, Trübner, and Co.

Royal 4to, pp. xiv.—326, cloth. Price 21*s*.

OBSERVATIONS OF THE INTERNATIONAL POLAR EXPEDITIONS,

1882–1883.

FORT RAE.

With 32 Lithographic Folding Plates.

A reduction of price to Fellows of the Royal Society.

Continental Agents for the publications of the Royal Society, Messrs. Friedländer and Son, Berlin.

Just Published, and Sold by Dulau and Co., Soho Square.

SEPARATE PAPERS FROM THE 'PHILOSOPHICAL TRANSACTIONS.'

1898-99.

A. 221. A Compensated Interference Dilatometer. By A. E. Tutton, Assoc. R.C.S. Price 3*s.*

A. 222. The Electric Conductivity of Nitric Acid. By V. H. Veley, M.A., F.R.S., and J. J. Manley, Daubeny Curator, Magdalen College, Oxford. Price 2*s.* 6*d.*

A. 223. On the Thermal Conductivities of Single and Mixed Solids and Liquids and their Variation with Temperature. By Charles H. Lees, D.Sc., Assistant Lecturer in Physics in the Owens College. Price 2*s.* 6*d.*

A. 224. Experiments on Aneroid Barometers at Kew Observatory, and their Discussion. By C. Chree, Sc.D., LL.D., F.R.S., Superintendent. Price 2*s.* 6*d.*

A. 225. On the Heat Dissipated by a Platinum Surface at High Temperatures. By. J. E. Petavel. Price 2*s.*

A. 226. Title, Contents, Index, &c., Vol. 191 (1898). Price 6*d.*

A. 227. On the Connexion of Algebraic Functions with Automorphic Functions. By E. T. Whittaker, B.A., Fellow of Trinity College, Cambridge. Price 1*s.* 6*d.*

A. 228. Researches in Vortex Motion.—Part III. On Spiral or Gyrostatic Vortex Aggregates. By W. M. Hicks, D.Sc., F.R.S., Professor of Physics in University College, Sheffield. Price 3*s.* 6*d.*

A. 229. On the Application of the Theory of Error to Cases of Normal Distribution and Normal Correlation. By W. F. Sheppard, M.A., LL.M., formerly Fellow of Trinity College, Cambridge. Price 4*s.*

A. 230. Mathematical Contributions to the Theory of Evolution.—V. On the Reconstruction of the Stature of Prehistoric Races. By Karl Pearson, F.R.S., University College, London. Price 4*s.*

A. 231. An Experiment in Search of a Directive Action of one Quartz Crystal on Another. By J. H. Poynting, Sc.D., F.R.S., and P. L. Gray, B.Sc. Price 1*s.*

B. 157. An Experimental Research upon Cerebro-Cortical Afferent and Efferent Tracts. By David Ferrier, M.D., F.R.S., Professor of Neuropathology, King's College; and William Aldren Turner, M.D., F.R.C.P., Demonstrator of Neuropathology, King's College. Price 3*s.*

B. 158. Experiments in Examination of the Peripheral Distribution of the Fibres of the Posterior Roots of some Spinal Nerves.—Part II. By Charles S. Sherrington, M.A., M.D., F.R.S., Professor of Physiology at University College, Liverpool. Price 9*s.*

B. 159. On Apogamy and the Development of Sporangia upon Fern Prothalli. By William H. Lang, M.B., B.Sc., Lecturer in Botany, Queen Margaret College, and Clark Scholar, Glasgow University. Price 4*s.* 6*d.*

B. 160. The Pharmacology of Aconitine, Diacetyl-Aconitine, Benzaconine, and Aconine, considered in relation to their Chemical Constitution. By J. Theodore Cash, M.D., F.R.S., and Wyndham R. Dunstan, M.A., F.R.S. Price 10*s.*

B. 161. On the Cytological Features of Fertilization and Related Phenomena in *Pinus silvestris, L.* By Vernon H. Blackman, B.A., F.L.S., Hutchinson Student, St. John's College, Cambridge, and Assistant, Department of Botany, British Museum. Price 3*s.*

B. 162. On the Nature of the Coccospheres and Rhabdospheres. By George Murray, F.R.S., Keeper of Botany, British Museum, and Vernon H. Blackman, B.A., F.L.S., Hutchinson Student, St. John's College, Cambridge, and Assistant, Department of Botany, British Museum. Price 2*s.*

B. 163. Upon the Structure and Development of the Enamel of Elasmobranch Fishes. By Charles S. Tomes, M.A., F.R.S. Price 2*s.*

B. 164. The Relations between the Hybrid and Parent Forms of Echinoid Larvæ. By H. M. Vernon, M.A., M.B., Radcliffe Travelling Fellow of the University of Oxford. Price 3*s.* 6*d.*

B. 165. Observations on Stomata. By Francis Darwin, F.R.S. Price 4*s.*

B. 166. Contributions to our Knowledge of the Fucaceæ: their Life-History and Cytology. By J. Bretland Farmer, M.A., and J. Ll. Williams. Price 3*s.* 6*d.*

B. 167. Title, Contents, Index, &c., Vol. 190 (1898). Price 6*d.*

Cloth cases for binding the Philosophical Transactions can be obtained of Harrison and Sons, St. Martin's Lane. Price 1*s.* 6*d.*

PHILOSOPHICAL TRANSACTIONS

OF THE

ROYAL SOCIETY OF LONDON.

SERIES A.

VOL. 195, pp. 1–47.

MATHEMATICAL CONTRIBUTIONS TO THE THEORY OF EVOLUTION.

VII. ON THE CORRELATION OF CHARACTERS NOT QUANTITATIVELY MEASURABLE.

BY

KARL PEARSON, F.R.S.

LONDON:
PUBLISHED FOR THE ROYAL SOCIETY
BY DULAU AND CO., 37, SOHO SQUARE, W.
CONTINENTAL AGENTS, MESSRS. FRIEDLÄNDER AND SON, BERLIN.
1900.

A. 262.

16.8 1900.

Price Two Shillings and Sixpence.

INDEX SLIP.

PEARSON, Karl.—Mathematical Contributions to the Theory of Evolution.—VII.—On the Correlation of Characters not Quantitatively Measurable.
Phil. Trans., A, vol. 195, 1900, pp. 1-47.

Antitoxin Treatment in Diphtheria—Effectiveness Quantitatively Measured.
PEARSON, Karl. Phil. Trans., A, vol. 195, 1900, pp. 1-47.

Association of Attributes.
PEARSON, Karl. Phil. Trans., A, vol. 195, 1900, pp. 1-47.

Correlation, Determination of, when Variables are not Quantitatively Measurable.
PEARSON, Karl. Phil. Trans., A, vol. 195, 1900, pp. 1-47.

Errors, Normal Theory of—Extension to Cases of Variables not Capable of Quantitative Measurement.
PEARSON, Karl. Phil. Trans., A, vol. 195, 1900, pp. 1-47.

Heredity—Rate of Production of Exceptional Offspring by Exceptional Stock.
PEARSON, Karl. Phil. Trans., A, vol. 195, 1900, pp. 1-47.

Vaccination, Effectiveness of, Quantitatively Measured.
PEARSON, Karl. Phil. Trans., A, vol. 195, 1900, pp. 1-47.

PHILOSOPHICAL TRANSACTIONS.

I. *Mathematical Contributions to the Theory of Evolution.*—VII. *On the Correlation of Characters not Quantitatively Measurable.*

By Karl Pearson, *F.R.S.*

(*From the Department of Applied Mathematics, University College, London.*)

Received February 7,—Read March 1, 1900.

NOTE.

In August, 1899, I presented a memoir to the Royal Society on the inheritance of coat-colour in the horse and of eye-colour in man, which was read November, 1899, and ultimately ordered to be published in the 'Phil. Trans.' Before that memoir was printed, Mr. Yule's valuable memoir on Association was read, and, further, Mr. Leslie Bramley-Moore showed me that the theory of my memoir as given in § 6 of the present memoir led to somewhat divergent results according to the methods of proportioning adopted. We therefore undertook a new investigation of the theory of the whole subject, which is embodied in the present memoir. The data involved in the paper on coat-colour in horses and eye-colour in man have all been recalculated, and that paper is nearly ready for presentation.* But it seemed best to separate the purely theoretical considerations from their application to special cases of inheritance, and accordingly the old memoir now reappears in two sections. The theory discussed in this paper was, further, the basis of a paper on the Law of Reversion with special reference to the Inheritance of Coat-colour in Basset Hounds recently communicated to the Society, and about to appear in the 'Proceedings.'†

While I am responsible for the general outlines of the present paper, the rough draft of it was taken up and carried on in leisure moments by Mr. Leslie Bramley-Moore, Mr. L. N. G. Filon, M.A., and Miss Alice Lee, D.Sc. Mr. Bramley-Moore discovered the u-functions; Mr. Filon proved most of their general properties and the convergency of the series; I alone am responsible for sections 4, 5, and 6. Mr. Leslie Bramley-Moore sent me, without proof, on the eve of his departure for the Cape, the general expansion for z on p. 26. I am responsible for the present proof and its applications. To Dr. Alice Lee we owe most of the illustrations and the table on p. 17. Thus the work is essentially a joint memoir in which we have equal part, and the use of the first personal pronoun is due to the fact that the material had to be put together and thrown into form by one of our number.—K. P.

Contents.

---

* Since ordered to be printed in the 'Phil. Trans.'

† Read January 25, 1900. 'Roy. Soc. Proc.,' vol. 66, p. 140.

§ (1.) *On a General Theorem in Normal Correlation.*

Let the frequency surface

$$z = \frac{\mathrm{N}}{2\pi\sqrt{(1-r^2)}\sigma_1\sigma_2} e^{-\frac{1}{2}\frac{1}{1-r^2}\left(\frac{x^2}{\sigma_1^2}+\frac{y^2}{\sigma_2^2}-\frac{2rxy}{\sigma_1\sigma_2}\right)},$$

where N = total number of observations,

σ_1, σ_2 = standard deviations of organs x and y,

r = correlation of x and y,

be divided into four parts by two planes at right angles to the axes of x and y at distances h' and k' from the origin. The total volumes or frequencies in these parts will be represented by a, b, c, and d in the manner indicated in the accompanying plan:—

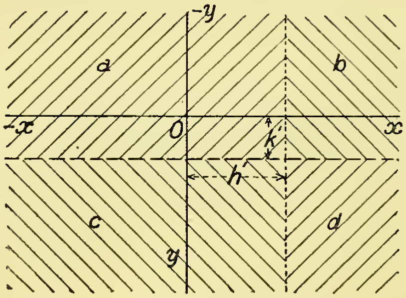


Table of Frequencies

a	b	a+b
c	d	c+d
a+c	b+d	N

Then clearly

$$d = \frac{\mathrm{N}}{2\pi\sqrt{(1-r^2)}\sigma_1\sigma_2}\int_{h'}^{\infty}\int_{k'}^{\infty} e^{-\frac{1}{2}\frac{1}{1-r^2}\left(\frac{x^2}{\sigma_1^2}+\frac{y^2}{\sigma_2^2}-\frac{2rxy}{\sigma_1\sigma_2}\right)}\,dx\,dy$$

$$= \frac{\mathrm{N}}{2\pi\sqrt{(1-r^2)}}\int_{h}^{\infty}\int_{k}^{\infty} e^{-\frac{1}{2}\frac{1}{1-r^2}(x^2+y^2-2rxy)}\,dx\,dy \quad . \; . \; . \; . \; . \; . \; . \; (\mathrm{i.}),$$

if $h = h'/\sigma_1$ and $k = k'/\sigma_2$.

Further,

$$b + d = \frac{\mathrm{N}}{\sqrt{2\pi}\,\sigma_1}\int_{h'}^{\infty} e^{-\frac{1}{2}\frac{x^2}{\sigma_1^2}}\,dx,$$

$$= \frac{\mathrm{N}}{\sqrt{2\pi}}\int_{h}^{\infty} e^{-\frac{1}{2}x^2}\,dx \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii.)},$$

and

$$c + d = \frac{\mathrm{N}}{\sqrt{2\pi}}\int_{k}^{\infty} e^{-\frac{1}{2}y^2}\,dy \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iii.)},$$

$$\frac{(a+c)-(b+d)}{\mathrm{N}} = \sqrt{\frac{2}{\pi}}\int_{0}^{h} e^{-\frac{1}{2}x^2}\,dx \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv.)},$$

$$\frac{(a+b)-(c+d)}{\mathrm{N}} = \sqrt{\frac{2}{\pi}}\int_{0}^{k} e^{-\frac{1}{2}y^2}\,dy \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(v.)}.$$

Thus, when a, b, c, and d are known, h and k can be found by the ordinary table of the probability integral, say that of Mr. SHEPPARD (' Phil. Trans.,' A, vol. 192, p. 167, Table VI.*). The limits accordingly of the integral for d in (i.) are known.

Now consider the expression

$$\frac{1}{\sqrt{1-r^2}}\,e^{-\frac{1}{2}\frac{1}{1-r^2}(x^2+y^2-2rxy)} = \mathrm{U}, \text{ say}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(vi.)},$$

and let us expand it in powers of r. Then, if the expansion be

$$\mathrm{U} = e^{-\frac{1}{2}(x^2+y^2)}\left(u_0 + \frac{u_1 r}{\underline{|1}} + \frac{u_2 r^2}{\underline{|2}} + \ldots + \frac{u_n r^n}{\underline{|n}} + \ldots\right) \quad . \quad . \quad . \quad . \quad \text{(vii.)},$$

we shall have

$$u_n = e^{\frac{1}{2}(x^2+y^2)}\left(\frac{d^n\mathrm{U}}{dr^n}\right)_{r=0} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(viii.)}.$$

Taking logarithmic differentials, we get at once

$$(1-r^2)^2\frac{d\mathrm{U}}{dr} = \{xy + r(1-x^2-y^2) + r^2xy - r^3\}\mathrm{U}.$$

Differentiating n times by LEIBNITZ'S theorem, and putting $r = 0$, we have, after some reductions

$$\begin{aligned} u_{n+1} = {} & n(2n-1-x^2-y^2)u_{n-1} \\ & - n(n-1)(n-2)^2 u_{n-3} \\ & + xy\{u_n + n(n-1)u_{n-2}\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ix.)}. \end{aligned}$$

Hence we find

$$\left.\begin{aligned} u_0 &= 1 \\ u_1 &= xy \\ u_2 &= (x^2-1)(y^2-1) \\ u_3 &= x(x^2-3)\,y(y^2-3) \\ u_4 &= (x^4-6x^2+3)(y^4-6y^2+3) \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad \text{(x.)}$$

---

* See, however, foot-note, p. 5.

Thus the following laws are indicated :—

$$u_n = v_n \times w_n \quad \text{. . . . . (xi.),}$$

where

$$v_n = xv_{n-1} - (n-1)v_{n-2} \quad \text{. . . . . (xii.),}$$

$$w_n = yw_{n-1} - (n-1)w_{n-2} \quad \text{. . . . . (xiii.).}$$

We shall now show that these laws hold good by induction. Assume To prove

$$u_{n+1} = v_{n+1}w_{n+1} = (xv_n - nv_{n-1})(yw_n - nw_{n-1}).$$

Or Thus

$$u_{n+1} = xyu_n + n^2u_{n-1} - n(yw_nv_{n-1} + xv_nw_{n-1}).$$

(xii)

B But by (ix.), substituting for u_{n-3} from (xi.) and (xiii.), putting n-1 for n,

$$\begin{aligned} u_{n+1} = xy\{v_nw_n + n(n-1)v_{n-2}w_{n-2}\} &+ n(2n-1-x^2-y^2)v_{n-1}w_{n-1} \\ &- n(n-1)v_{n-1}w_{n-1} - xyn(n-1)v_{n-2}w_{n-2} \\ &+ n(n-1)(yv_{n-1}w_{n-2} + xv_{n-2}w_{n-1}). \end{aligned}$$

$$\begin{aligned} &= xyv_nw_n + n^2v_{n-1}w_{n-1} - n(x^2+y^2)v_{n-1}w_{n-1} \\ &\qquad + n(n-1)(yv_{n-1}w_{n-2} + xv_{n-2}w_{n-1}) \end{aligned}$$

w_{n-1}

$$\begin{aligned} &= xyv_nw_n + n^2v_{n-1}w_{n-1} - n\{yv_{n-1}(yu_{n-1} - \overline{n-1}\,w_{n-2}) \\ &\qquad\qquad + xw_{n-1}(xv_{n-1} - \overline{n-1}\,v_{n-2}\} \end{aligned}$$

$$= xyv_nw_n + n^2v_{n-1}w_{n-1} - n(yv_{n-1}w_n + xw_{n-1}v_n) \text{, by xii and xiii,}$$

$$= v_{n+1}w_{n+1}, \text{ as we have seen above.}$$

Thus, if the theorem holds for u_n, it holds for u_{n+1}. Accordingly

$$U = e^{-\frac{1}{2}(x^2+y^2)}\left(1 + \frac{v_1w_1}{\underline{|1}}r + \frac{v_2w_2}{\underline{|2}}r^2 + \ldots + \frac{v_nw_n}{\underline{|n}}r^n + \ldots\right) \quad \text{. . . (xiv.),}$$

where the v's and w's are given by (x.), (xii.), and (xiii.).

It is thus clear that $\frac{1}{2\pi}\int_h^\infty\int_k^\infty U\,dx\,dy$ consists of a series of which the general term is

$$\frac{1}{\underline{|n}}V_nW_nr^n$$

where

$$V_n = \frac{1}{\sqrt{2\pi}}\int_h^\infty e^{-\frac{1}{2}x^2}v_n dx$$

$$W_n = \frac{1}{\sqrt{2\pi}}\int_k^\infty e^{-\frac{1}{2}y^2}w_n dy.$$

It remains to find these integrals.

The general form of v_n is given by

$$v_n = x^n - \frac{n(n-1)}{2\,\underline{|1}}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2^2\,\underline{|2}}x^{n-4} - \&\text{c.} \quad \text{. . (xv.).}$$

For this obviously gives (x.). Assume it true for v_{n-1} and v_{n-2}, then

$$xv_{n-1}-(n-1)v_{n-2}=x^n-\frac{(n-1)(n-2)}{2\,\underline{|1}}x^{n-2}+\frac{(n-1)(n-2)(n-3)(n-4)}{2^2\,\underline{|2}}x^{n-4}-\dots$$

$$-(n-1)x^{n-2}+\frac{(n-1)(n-2)(n-3)}{2\,\underline{|1}}x^{n-4}-\dots$$

$$=x^n-\frac{n(n-1)}{2\,\underline{|1}}x^{n-2}+\frac{n(n-1)(n-2)(n-3)}{2^2\,\underline{|2}}x^{n-4}-\dots$$

$$=v_n.$$

Thus the expression (xv.) is shown to hold by induction, the general terms being

$$(-1)^r\frac{(n-1)(n-2)\dots(n-2r+1)}{2^{r-1}\,\underline{|r-1}}\left(\frac{n-2r}{2r}+1\right)x^{n-2r}$$

$$=(-1)^r\frac{n(n-1)(n-2)\dots(n-2r+1)}{2^r\,\underline{|r}}x^{n-2r},$$

or the general term in u_n.

We notice at once that

$$\frac{dv_n}{dx}=nv_{n-1}\quad\dots\dots\dots\dots\quad\text{(xvi.)}.$$

Thus, by (xii.)

$$v_n=xv_{n-1}-\frac{dv_{n-1}}{dx}.$$

Multiply by $e^{-\frac{1}{2}x^2}$ and integrate

$$\int e^{-\frac{1}{2}x^2}v_n\,dx=\int xe^{-\frac{1}{2}x^2}v_{n-1}dx-\int e^{-\frac{1}{2}x^2}\frac{dv_{n-1}}{dx}dx.$$

Integrating the latter integral by parts, we have

$$\int v_ne^{-\frac{1}{2}x^2}dx=-e^{-\frac{1}{2}x^2}v_{n-1},$$

or

$$\mathrm{V}_n=\frac{1}{\sqrt{2\pi}}\int_h^\infty v_ne^{-\frac{1}{2}x^2}dx=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}h^2}(v_{n-1})_{x=h}.$$

Now $\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}h^2}$ can be found from any table of the ordinates of the normal curve, *e.g.*, Mr. SHEPPARD'S, 'Phil. Trans.,' A, vol. 192, p. 153, Table I.* We shall accordingly put

$$\mathrm{H}=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}h^2},\qquad \mathrm{K}=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}k^2}\quad\dots\dots\quad\text{(xvii.)},$$

and look upon H and K as known quantities.

---

* For our present purposes the differences of Mr. SHEPPARD'S tables are occasionally too large, but the following series give very close results :—

Let

$$\chi_1=\sqrt{\frac{\pi}{2}}\,\frac{(a+c)-(b+d)}{\mathrm{N}}=\int_0^h e^{-\frac{1}{2}x^2}dx\text{ by (iv.),}$$

$$\chi_2=\sqrt{\frac{\pi}{2}}\,\frac{(a+b)-(c+d)}{\mathrm{N}}=\int_0^k e^{-\frac{1}{2}y^2}dy\text{ by (v.).}$$

Further, let us write $(v_{n-1})_{x=h}$ as $\bar{v}_{n-1}$, and similarly $(w_{n-1})_{y=k}$ as $\bar{w}_{n-1}$. Thus

$$\mathrm{V}_n = \mathrm{H} \, . \, \bar{v}_{n-1}, \qquad \mathrm{W}_n = \mathrm{K} \, . \, \bar{w}_{n-1} \quad . \; . \; . \; . \; . \; . \; . \text{(xviii.)}.$$

We have then from (i.)

$$\frac{d}{\mathrm{N}} = \frac{1}{2\pi} \int_h^\infty \int_k^\infty \mathrm{U} \, dx \, dy$$

$$= \frac{1}{2\pi} \int_h^\infty \int_k^\infty e^{-\frac{1}{2}(x^2+y^2)} \, dx \, dy + \overset{\infty}{\underset{1}{\mathrm{S}}} \left(\frac{r^n}{\underline{|n}} \mathrm{HK} \, \bar{v}_{n-1} \bar{w}_{n-1} \right)$$

$$= \frac{(b+d)(c+d)}{\mathrm{N}^2} + \overset{\infty}{\underset{1}{\mathrm{S}}} \left(\frac{r^n}{\underline{|n}} \mathrm{HK} \, \bar{v}_{n-1} \bar{w}_{n-1} \right)$$

by (ii.) and (iii.).

Or, remembering that $\mathrm{N} = a + b + c + d$, we can write this

$$\frac{ad - bc}{\mathrm{N}^2 \mathrm{HK}} = \overset{\infty}{\underset{1}{\mathrm{S}}} \left(\frac{r^n}{\underline{|n}} \bar{v}_{n-1} \bar{w}_{n-1} \right)$$

$$= r + \frac{r^2}{2} hk + \frac{r^3}{6} (h^2 - 1)(k^2 - 1) + \frac{r^4}{24} h(h^2 - 3) k(k^2 - 3)$$

$$+ \frac{r^5}{120} (h^4 - 6h^2 + 3)(k^4 - 6k^2 + 3)$$

$$+ \frac{r^6}{720} h(h^4 - 10h^2 + 15) k(k^4 - 10k^2 + 15)$$

$$+ \frac{r^7}{5040} (h^6 - 15h^4 + 45h^2 - 15)(k^6 - 15k^4 + 45k^2 - 15)$$

$$+ \frac{r^8}{40320} h(h^6 - 21h^4 + 105h^2 - 105) k(k^6 - 21k^4 + 105k - 105) + \text{, \&c.} \quad . \; . \; . \; . \; \text{(xix.)}.$$

Then

$$h = \chi_1 + \frac{1}{\underline{|3}} \chi_1^3 + \frac{7}{\underline{|5}} \chi_1^5 + \frac{127}{\underline{|7}} \chi_1^7 + \ldots$$

$$\frac{1}{\mathrm{H}} = \sqrt{2\pi} \left(1 + \frac{1}{\underline{|2}} \chi_1^2 + \frac{7}{\underline{|4}} \chi_1^4 + \frac{127}{\underline{|6}} \chi_1^6 + \ldots \right),$$

and

$$k = \chi_2 + \frac{1}{\underline{|3}} \chi_2^3 + \frac{7}{\underline{|5}} \chi_2^5 + \frac{127}{\underline{|7}} \chi_2^7 + \ldots$$

$$\frac{1}{\mathrm{K}} = \sqrt{2\pi} \left(1 + \frac{1}{\underline{|2}} \chi_2^2 + \frac{7}{\underline{|4}} \chi_2^4 + \frac{127}{\underline{|6}} \chi_2^6 + \ldots \right):$$

These follow from the considerations that if

$$\chi_1 = \sqrt{2\pi} \phi_1, \qquad \chi_2 = \sqrt{2\pi} \phi_2,$$

$$\frac{d\phi_1}{dh} = \mathrm{H}, \qquad \frac{d\phi_2}{dk} = \mathrm{K},$$

$$\frac{d\mathrm{H}}{d\phi_1} = -h, \qquad \frac{d\mathrm{K}}{d\phi_2} = -k,$$

whence it is easy to find the successive differentials of h with regard to ϕ_1 and k with regard to ϕ_2, and then obtain the above results by MACLAURIN'S theorem. There is, of course, no difficulty in calculating H and K from (xvii.) directly. That method was adopted in the numerical illustrations.

Here the left-hand side is known, and since h and k are known, we can find the coefficients of any number of powers of r so soon as the first two have been found, from (xii.) and (xiii.).

Accordingly the correlation can be found if we have only made a grouping of our frequencies into the four divisions, a, b, c, and d.

If h and k be zero, we have from (xvii.) and (iv.)

$$\mathrm{H} = \mathrm{K} = \frac{1}{\sqrt{2\pi}}$$

$$a + c = b + d = \tfrac{1}{2}\mathrm{N}.$$

The right-hand side of (xix.) is now

$$r + \frac{1}{\underline{|3}} r^3 + \quad . \quad . \quad .$$

or equal to $\sin^{-1} r$.

Hence
$$r = \sin 2\pi \frac{(ad - bc)}{\mathrm{N}^2}$$

$$= \cos \pi \frac{b}{a + b} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\mathrm{xx.}),$$

which agrees with a result of Mr. SHEPPARD'S, 'Phil. Trans.,' A, vol. 192, p. 141. We have accordingly reached a generalised form of his result for *any* class-index whatever. Clearly, also, r being known, we can at once calculate the frequency of pairs of organs with deviations as great as or greater than h and k.

§ (2.) *Other Series for the Determination of r.*

For many purposes the series (xix.) is sufficiently convergent to give r for given h and k with but few approximations, but we will now turn to other developments.

We have by (vii.)

$$\int_0^r \mathrm{U}\, dr = e^{-\frac{1}{2}(x^2+y^2)} \left(u_0 r + \frac{u_1 r^2}{\underline{|2}} + \, . \, . \, . + u_n \frac{r^{n+1}}{\underline{|n + 1}} + \, . \, . \, . \right).$$

Put $x = h$, $y = k$, and write for brevity

$$\epsilon = \frac{ad - bc}{\mathrm{N}^2\mathrm{HK}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\mathrm{xxi.}).$$

It follows at once from (xix.) that

$$\epsilon = e^{\frac{1}{2}(h^2 + k^2)} \int_0^r \mathrm{U}\, dr$$

$$= e^{\frac{1}{2}(h^2+h^2)} \int_0^r \frac{1}{\sqrt{1 - r^2}} e^{-\frac{1}{2}\frac{1}{1-r^2}(h^2 + k^2 - 2rhk)}\, dr$$

$$\left.\begin{aligned} &= e^{\frac{1}{2}k^2}\int_0^\theta e^{-\frac{1}{2}(h\tan\theta\,-\,k\sec\theta)^2}\,d\theta \\ &= e^{\frac{1}{2}h^2}\int_0^\theta e^{-\frac{1}{2}(k\tan\theta\,-\,h\sec\theta)^2}\,d\theta \end{aligned}\right\} \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad \text{(xxii.)},$$

if $r = \sin\theta$.

Now either of the quantities under the sign of integration in (xxii.) can be expanded in powers of θ by MACLAURIN'S theorem. Thus let

$$\begin{aligned} \chi &= e^{-\frac{1}{2}(k\tan\theta\,-\,h\sec\theta)^2} \\ &= \chi_0 + \left(\frac{d\chi}{d\theta}\right)_0\theta + \left(\frac{d^2\chi}{d\theta^2}\right)_0\frac{\theta^2}{\underline{|2}} + \ldots + \left(\frac{d^n\chi}{d\theta^n}\right)_0\frac{\theta^n}{\underline{|n}} + \ldots \end{aligned}$$

Then

$$\epsilon = e^{\frac{1}{2}h^2}\left(\chi_0\theta + \left(\frac{d\chi}{d\theta}\right)_0\frac{\theta^2}{\underline{|2}} + \left(\frac{d^2\chi}{d\theta^2}\right)_0\frac{\theta^3}{\underline{|3}} + \ldots + \left(\frac{d^n\chi}{d\theta^n}\right)_0\frac{\theta^{n+1}}{\underline{|n+1}} + \ldots\right),$$

and it remains to find $\left(\dfrac{d^n\chi}{d\theta^n}\right)_0$.

Now
$$\log\chi = -\tfrac{1}{2}(k\tan\theta - h\sec\theta)^2.$$
Hence
$$\cos^3\theta\frac{d\chi}{d\theta} = -\chi[(h^2+k^2)\sin\theta - hk(\tfrac{3}{2} - \tfrac{1}{2}\cos 2\theta)].$$

Differentiating $\overline{n-1}$ times by LEIBNITZ'S theorem, and putting $\theta = 0$,

$$\begin{aligned} &4\left(\frac{d^n\chi}{d\theta^n}\right)_0 - 4hk\left(\frac{d^{n-1}\chi}{d\theta^{n-1}}\right) + \ldots \\ &\quad + \frac{(n-1)\ldots(n-r+1)}{\underline{|r-1}}\left[\cos\frac{r\pi}{2}\left\{\frac{n-r}{r}(3+3^r) - 4(h^2+k^2)\right\}\right. \\ &\quad \left. + \sin\frac{r\pi}{2}\,2^r hk\right]\left(\frac{d^{n-r}\chi}{d\theta^{n-r}}\right)_0 + \ldots = 0 \quad . \; . \; . \; . \; . \; . \; . \; . \quad \text{(xxiii.)}. \end{aligned}$$

Clearly $\chi_0 = e^{-\frac{1}{2}h^2}$, then we rapidly find

$$\begin{aligned} \left(\frac{d\chi}{d\theta}\right)_0 &= hk\,e^{-\frac{1}{2}h^2} \\ \left(\frac{d^2\chi}{d\theta^2}\right)_0 &= -(h^2+k^2-h^2k^2)e^{-\frac{1}{2}h^2} \\ \left(\frac{d^3\chi}{d\theta^3}\right)_0 &= hk\,\{h^2k^2 - 3(h^2+k^2) + 5\}. \end{aligned}$$

Or, finally

$$\epsilon = \theta + \tfrac{1}{2}hk\theta^2 - (h^2+k^2-h^2k^2)\frac{\theta^3}{6} + hk\{h^2k^2 - 3(h^2+k^2) + 5\}\frac{\theta^4}{24} + \ldots \quad \text{(xxiv.)},$$

where more terms if required can be found by (xxiii.). If θ be fairly small, θ^5 will be negligible. Or if h and k be small, the lowest term in the next factor will be $h^2 + k^2$, and this into $\theta^5/\underline{|5}$ is generally quite insensible. Very often two or three terms on the right-hand side of (xxiv.) give quite close enough values of θ, and accordingly of $r = \sin\theta$. (xxiv.) is clearly somewhat more convergent than (xix.) if h and k are, as usually happens, less than unity.

Returning now to (xix.), let us write it

$$\epsilon = f(r, h, k).$$

This is the equation that must be solved for r. Suppose r_0 a root of this when we retain only few terms on the right, say a root of the quadratic

$$\epsilon = r + \tfrac{1}{2}hkr^2.$$

Then if $r = r_0 + \rho$,

$$\epsilon = f(r_0, h, k) + \rho f'(r_0, h, k) + \frac{1}{\underline{|2}}\rho^2 f''(r_0, h, k) + \&c.$$

Hence $\rho = \dfrac{\epsilon - f(r_0hk)}{f'(r_0hk)}$ to a third approximation

$$= -\frac{\frac{1}{6}(h^2-1)(k^2-1)r_0^{\,3}}{\dfrac{1}{\sqrt{1-r_0^{\,2}}}\,e^{-\frac{1}{2}\frac{1}{1-r_0^{\,2}}(h^2+k^2-2r_0hk)}} \text{ nearly} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxv.)},$$

which gives us a value of ρ which, substituted in ρ^2 in the above equation, introduces only terms of the 6th order in r_0.

Another integral expression for ϵ of Equation (xxi.) may here be noticed :

$$\epsilon = e^{\frac{1}{2}(h^2+k^2)}\int_0^r \frac{dr}{\sqrt{1-r^2}}\,e^{-\frac{1}{2}\frac{1}{1-r^2}(h^2+k^2-2rhk)}.$$

Put $h = \dfrac{1}{\sqrt{2}}(\beta+\gamma)$, $k = \dfrac{1}{\sqrt{2}}(\beta-\gamma)$.

Hence

$$\epsilon = e^{\frac{1}{2}(\beta^2+\gamma^2)}\int_0^r \frac{dr}{\sqrt{1-r^2}}\,e^{-\frac{1}{2}\left(\frac{\beta^2}{1+r}+\frac{\gamma^2}{1-r}\right)}$$

$$= e^{\frac{1}{4}(\beta^2+\gamma^2)}\int_0^r \frac{dr}{\sqrt{1-r^2}}.\,e^{-\frac{1}{4}\left(\beta^2\frac{1-r}{1+r}+\gamma^2\frac{1+r}{1-r}\right)}.$$

Let $\tan 2\phi = \dfrac{1-r}{1+r}$, or, $r = \cos 2\,\phi$.

Therefore

$$\epsilon = 2e^{\frac{1}{4}(\beta^2+\gamma^2)}\int_\phi^{45^\circ} e^{-\frac{1}{4}(\beta^2\tan^2\phi+\gamma^2\cot^2\phi)}\,d\phi,$$

$$= 2e^{\frac{1}{4}(\beta^2+\gamma^2)}\int_1^v e^{-\frac{1}{4}\left(\frac{\beta^2}{v^2}+\gamma^2v^2\right)}\frac{dv}{1+v^2}$$

where $v = \cot\phi$ and is > 1.

It seems possible that interesting developments for ϵ might be deduced from this integral expression.

§ (3.) *To show that the Series for r is Convergent if $r < 1$, whatever be the Values of h and k.*

Write the series in the form of p. 6, *i.e.* :—

$$\epsilon = \overset{n}{\underset{1}{\mathrm{S}}} \frac{r^n}{\underline{|n}} \bar{v}_{n-1} \bar{w}_{n-1}.$$

Now

$$\left.\begin{aligned} \bar{v}_{n+1} &= h\bar{v}_n - n\bar{v}_{n-1} \\ \bar{w}_{n+1} &= k\bar{w}_n - n\bar{w}_{n-1} \end{aligned}\right\} \text{by (xii.) and (xiii.).}$$

From these we deduce

$$\bar{v}_{n+1} = \{h^2 - (2n-1)\}\,\bar{v}_{n-1} - (n-1)(n-2)\,\bar{v}_{n-3}$$
$$\bar{w}_{n+1} = \{k^2 - (2n-1)\}\,\bar{w}_{n-1} - (n-1)(n-2)\,\bar{w}_{n-3}$$

Now let $\quad s_n = \bar{v}_{n-1} r^{\frac{1}{2}n} \big/ \{\underline{|n}\}^{\frac{1}{2}}, \quad t_n = \bar{w}_{n-1} r^{\frac{1}{2}n} \big/ \{\underline{|n}\}^{\frac{1}{2}}.$

Then we find

$$s_{n+2} = \frac{h^2 - (2n-1)}{\sqrt{(n+1)(n+2)}}\, s_n r - \sqrt{\frac{(n-1)(n-2)^2}{n(n+1)(n+2)}}\, s_{n-2} r^2,$$

$$t_{n+2} = \frac{k^2 - (2n-1)}{\sqrt{(n+1)(n+2)}}\, t_n r - \sqrt{\frac{(n-1)(n-2)}{n(n+1)(n+2)}}\, t_{n-2} r^2.$$

Thus, when n is large, we find the ratio of successive terms s_{n+2}/s_n or t_{n+2}/t_n is given by ρ, where

$$\rho = -2r - r^2/\rho \text{ or, } \rho = -r.$$

The ultimate ratio of $s_{n+2}\, t_{n+2}$ to $s_n\, t_n$ is accordingly given by r^2, but this is the ratio of alternate terms of the original series. The original series thus breaks up into two series, one of odd and one of even powers of r. Both these series are absolutely convergent whatever h and k be, having an ultimate convergence ratio of r^2

§ (4.) *To find the Probable Error of the Correlation Coefficient as Determined by the Method of this Memoir.*

Given a division of the total frequency N into a, b, c, d groups, where $a + b + c + d = \mathrm{N}$, then the probable error of any one of them, say a, is $\cdot 67449\, \sigma_a$, where*

$$\sigma_a = \sqrt{\frac{a(\mathrm{N}-a)}{\mathrm{N}}} \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad \text{(xxvi.)}.$$

Let $b + d = n_1$, $c + d = n_2$, then

---

* The standard deviation of an event which happens np times and fails nq times in n trials is well known to be $\sqrt{npq}$. The probable errors here dealt with are throughout, of course, those arising from different samples of the same general population

$$\sigma_{n_1} = \sqrt{\frac{n_1(\mathrm{N} - n_1)}{\mathrm{N}}} \qquad \sigma_{n_2} = \sqrt{\frac{n_2(\mathrm{N} - n_2)}{\mathrm{N}}} \quad . \quad . \quad . \quad \text{(xxvii.)}.$$

To obtain r_{cd} we have, if $\delta\eta$ denotes an error in any quantity η,

$$\delta c + \delta d = \delta n_2,$$

$$\therefore \ \sigma_c^2 + \sigma_d^2 + 2\sigma_c\sigma_d r_{cd} = \sigma_{n_2}^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxviii.)},$$

by squaring, summing for all possible variations in c and d, and dividing by the total number of variations.

Hence, substituting the values of the standard deviations as found above, we deduce

$$\sigma_c\sigma_d r_{cd} = -cd/\mathrm{N} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxix.)}.$$

In a similar manner

$$\delta n_1\delta d = \delta b\,\delta d + (\delta d)^2,$$

$$\sigma_d\sigma_{n_1}r_{dn_1} = \sigma_b\sigma_d r_{bd} + \sigma_d^2$$

$$\sigma_d\sigma_{n_1}r_{dn_1} = d\,(a + c)/\mathrm{N} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxx.)}.$$

and

$$\sigma_d\sigma_{n_2}r_{dn_2} = d\,(a + b)/\mathrm{N} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxi.)},$$

Now

$$n_1 = \frac{\mathrm{N}}{\sqrt{2\pi}}\int_h^\infty e^{-\frac{1}{2}x^2}\,dx,$$

$$\therefore \ \delta n_1 = -\frac{\mathrm{N}}{\sqrt{2\pi}}\,e^{-\frac{1}{2}h^2}\delta h = -\mathrm{NH}\delta h.$$

Thus

$$\sigma_{n_1} = \mathrm{NH}\,\sigma_h \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxii.)},$$

and similarly

$$\sigma_{n_2} = \mathrm{NK}\,\sigma_k \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxiii.)}.$$

Hence the probable error of h

$$= \frac{\cdot 67449}{\mathrm{H}\sqrt{\mathrm{N}}}\sqrt{\frac{(b + d)(a + c)}{\mathrm{N}^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxiv.)},$$

and of k

$$= \frac{\cdot 67449}{\mathrm{K}\sqrt{\mathrm{N}}}\sqrt{\frac{(c + d)(a + b)}{\mathrm{N}^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxv.)}.$$

They can be found at once, therefore, when H and K have been found from an ordinate table of the exponential curve, and a, b, c, d are given. We have thus the probable error of the means as found from any double grouping of observations.

Next, noting that

$$\delta n_1\,\delta n_2 = \mathrm{N}^2\mathrm{HK}\,\delta h\,\delta k,$$

we have

$$\sigma_{n_1}\sigma_{n_2}r_{n_1n_2} = \mathrm{N}^2\mathrm{HK}\,\sigma_h\sigma_k r_{hk},$$

or

$$r_{n_1n_2} = r_{hk}.$$

But
$$\delta n_1 \delta n_2 = (\delta b + \delta d)(\delta c + \delta d),$$

$$\sigma_{n_1}\sigma_{n_2} r_{n_1 n_2} = \sigma_b \sigma_c r_{bc} + \sigma_b \sigma_d r_{bcd} + \sigma_c \sigma_d r_{cd} + \sigma_d^2,$$

$$= \frac{ad - bc}{\mathrm{N}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxvi.)},$$

therefore

$$\sigma_h \sigma_k r_{hk} = \frac{ad - bc}{\mathrm{NHK}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxvii.)}.$$

$$r_{hk} = \frac{ad - bc}{\sqrt{(b + d)(a + c)(c + d)(a + b)}} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxviii.)}.$$

This is an important result; it expresses the correlation between errors in the position of the means of the two characters under consideration. But if the probabilities were independent there could be no such correlation. Thus r_{hk} might be taken as a measure of divergence from independent variation. We shall return to this point later.

Since $\delta n_1 = -\mathrm{HN}\delta h$, we have $\delta n_1 \delta d = -\mathrm{HN}\delta d \delta h$, whence we easily deduce

$$r_{dn_1} = -r_{dh} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxix.)}.$$

Similarly
$$r_{dn_2} = -r_{dk} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xl.)}.$$

Now d is a function of r, h, and k. Hence if $d = f(r, h, k)$,

$$\delta d = \frac{df}{dr}\delta r + \frac{df}{dh}\delta h + \frac{df}{dk}\delta k$$

$$= \gamma_0 \delta r + \gamma_1 \delta h + \gamma_2 \delta k \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xli.)}.$$

Whence transposing, squaring, summing, and dividing by the total number of observations, we find

$$\gamma_0^2 \sigma_r^2 = \sigma_d^2 + \gamma_1^2 \sigma_h^2 + \gamma_2^2 \sigma_k^2 - 2\gamma_1 \sigma_d \sigma_h r_{dh} - 2\gamma_2 \sigma_d \sigma_k r_{dk} + 2\gamma_1 \gamma_2 \sigma_h \sigma_k r_{hk}$$

$$= \sigma_d^2 + \left(\frac{\gamma_1}{\mathrm{HN}}\right)^2 \sigma_{n_1}^2 + \left(\frac{\gamma_2}{\mathrm{KN}}\right)^2 \sigma_{n_2}^2 + 2\left(\frac{\gamma_1}{\mathrm{HN}}\right) \sigma_d \sigma_{n_1} r_{dn_1}$$

$$+ 2\left(\frac{\gamma_2}{\mathrm{HN}}\right) \sigma_d \sigma_{n_2} r_{dn_2} + \frac{2\gamma_1 \gamma_2}{\mathrm{N^2HK}} \sigma_{n_1} \sigma_{n_2} r_{n_1 n_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xlii.)}.$$

Substituting the values of the standard deviations and correlations as found above, we have

$$\sigma_r^2 = \frac{1}{\mathrm{N}\gamma_0^2}\left\{ d(a + b + c) + \left(\frac{\gamma_2}{\mathrm{HN}}\right)^2 (a + b)(d + c) + \left(\frac{\gamma_1}{\mathrm{KN}}\right)^2 (a + c)(d + b) \right.$$

$$\left. + \frac{2\gamma_1\gamma_2}{\mathrm{NHK}}(ad - bc) + \frac{2\gamma_2}{\mathrm{HN}} d(b + a) + \frac{2\gamma_1}{\mathrm{KN}} d(c + a) \right\} \quad . \quad . \quad . \quad \text{(xliii.)}.$$

It remains now to determine γ_0, γ_1, and γ_2.

By Equation (i.)

$$d = f(r, h, k) = \frac{\mathrm{N}}{2\pi\sqrt{1-r^2}} \int_h^\infty \int_k^\infty e^{-\frac{1}{2}\frac{x^2+y^2-2xy}{1-r^2}}\, dx\, dy,$$

$$\gamma_1 = \frac{df}{dh} = -\frac{\mathrm{N}}{2\pi\sqrt{1-r^2}} \int_k^\infty e^{-\frac{1}{2(1-r^2)}(h^2+y^2-2rhy)}\, dy$$

$$= -\frac{\mathrm{N}}{2\pi\sqrt{1-r^2}} e^{-\frac{1}{2}h^2} \int_k^\infty e^{-\frac{1}{2}\frac{(y-rh)^2}{1-r^2}}\, dy$$

$$= -\mathrm{H}\frac{\mathrm{N}}{\sqrt{2\pi}} \int_{\beta_2}^\infty e^{-\frac{1}{2}z^2}\, dz \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xliv.}),$$

where

$$\beta_2 = \frac{k - rh}{\sqrt{1-r^2}}.$$

Thus

$$\gamma_1/(\mathrm{NH}) = -\left(\frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-\frac{1}{2}z^2}\, dz - \frac{1}{\sqrt{2\pi}}\int_0^{\beta_2} e^{-\frac{1}{2}z^2}\, dz\right)$$

$$= \psi_2 - \tfrac{1}{2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xlv.}).$$

Similarly

$$\gamma_2/(\mathrm{NK}) = \psi_1 - \tfrac{1}{2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xlvi.}).$$

Here

$$\psi_1 = \frac{1}{\sqrt{2\pi}}\int_0^{\beta_1} e^{-\frac{1}{2}z^2}\, dz, \qquad \psi_2 = \frac{1}{\sqrt{2\pi}}\int_0^{\beta_2} e^{-\frac{1}{2}z^2}\, dz \quad . \quad . \quad . \quad . \quad . \quad (\text{xlvii.}),$$

where

$$\beta_1 = \frac{h - rk}{\sqrt{1-r^2}}, \qquad \beta_2 = \frac{k - rh}{\sqrt{1-r^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xlviii.}),$$

and thus ψ_1 and ψ_2 can be found at once from the tables when β_1 and β_2 are found from the known values of r, h, k.

Lastly, we have from Equation (xxi.)

$$\frac{ad - bc}{\mathrm{N}^2\mathrm{HK}} = e^{\frac{1}{2}(h^2+k^2)} \int_0^r \mathrm{U}\, dr,$$

or

$$\frac{d}{\mathrm{N}} = \frac{(d+b)(d+c)}{\mathrm{N}^2} + \frac{1}{2\pi}\int_0^r \mathrm{U}\, dr.$$

Thus*

$$\gamma_0 = df/dr = \frac{1}{2\pi}\mathrm{NU},$$

$$\gamma_0/\mathrm{N} = \chi_0,$$

where

$$\chi_0 = \frac{1}{2\pi}\frac{1}{\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}(h^2+k^2-2rhk)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xlix.})$$

a value which can again be found as soon as r, h, k are known. $\gamma_0 = \chi_0\mathrm{N}$ is clearly the ordinate of the frequency surface corresponding to $x = h$, $y = k$.

Substituting in Equation (xliii.) we have, after some reductions,

---

* By Equations (ii.) and (iii.), $d + b$ and $d + c$ are independent of r.

$$\text{Probable error of } r = {}^{\cdot}67449\sigma_r$$

$$= \frac{{}^{\cdot}67449}{\sqrt{N\chi_0}} \left\{ \frac{(a+d)(c+b)}{4N^2} + \psi_2{}^2 \frac{(a+c)(d+b)}{N^2} + \psi_1{}^2 \frac{(a+b)(d+c)}{N^2} \right.$$

$$\left. + 2\psi_1\psi_2 \frac{ad-bc}{N^2} - \psi_2 \frac{ab-cd}{N^2} - \psi_1 \frac{ac-bd}{N_2} \right\}^{\frac{1}{2}} \quad \text{. . . . . . (l.),}$$

where χ_0, ψ_1, and ψ_2 are readily found from Equations (xlix.), (xlvii.), and (xlviii.). Thus the probable error of r can be fairly readily found. It must be noted in using this formula, that a is the quadrant in which the mean falls, so that h and k are both *positive* (see fig., p. 2). In other words, we have supposed $a + c > b + d$ and $a + b > c + d$. Our lettering must always be arranged so as to suit this result before we apply the above formula.

§ (5.) *To Find a Physical Meaning for the Series in r, or for the* ϵ *of Equation* (xxi.).

Return to the original distribution $\dfrac{a \mid b}{c \mid d}$ of p. 2. If the probabilities of the two characters or organs were quite independent, we should expect the distribution

$$\begin{array}{c|c} N\dfrac{a+b}{N}\dfrac{a+c}{N} & N\dfrac{a+b}{N}\dfrac{b+d}{N} \\ \hline N\dfrac{c+d}{N}\dfrac{a+c}{N} & N\dfrac{c+d}{N}\dfrac{b+d}{N} \end{array}$$

Now re-arranging our actual data we may put it thus :

$$\frac{a \mid b}{c \mid d} = \begin{array}{c|c} N\dfrac{a+b}{N}\dfrac{a+c}{N} + \dfrac{ad-bc}{N} & N\dfrac{a+b}{N}\dfrac{b+d}{N} - \dfrac{ad-bc}{N} \\ \hline N\dfrac{c+d}{N}\dfrac{a+c}{N} - \dfrac{ad-bc}{N} & N\dfrac{c+d}{N}\dfrac{b+d}{N} + \dfrac{ad-bc}{N} \end{array}$$

Accordingly correlation denotes that $\dfrac{ad-bc}{N}$ has been transferred from each of the second and fourth compartments, and the same amount added to each of the first and third compartments. If $\eta = (ad - bc)/N^2$, then η is the *transfer per unit of the total frequency*. The magnitude of this transfer is clearly a measure of the divergence of the statistics from independent variation. It is physically quite as significant as the correlation coefficient itself, and of course much easier to determine. It must vanish with the correlation coefficient. We see from (xxi.) that

$$\eta = \epsilon \times HK,$$

or we have an interpretation for the series in r of (xix.).

Now, obviously any function of η, just like η itself, would serve as a measure of the divergence from perfectly independent variation. It is convenient to choose a function which shall lie arithmetically between 0 and 1.

Now consider what happens in the case of perfect correlation, *i.e.*, all the observations fall into a straight line. Hence if $ad > bc$, either b or c is zero, for a straight line cannot cut all four compartments, and a and d are obviously positive. Thus c and b can only be zero if $\eta = (c + d)(a + c)/\mathrm{N}^2$ or $(a + b)(b + d)/\mathrm{N}^2$. In order that b should be zero, it is needful that h and k, as given by (iv.) and (v.), should be positive or $a + c > b + d$, $a + b > c + d$, and the mean fall under the 45° line through the vertical and horizontal lines dividing the table into four compartments, *i.e.*, $h > k$. These conditions would be satisfied if $ad > bc$ and $a > d$, $c > b$. Now suppose our four-compartment table arranged so that

$$ad > bc, \quad a > d, \quad c > b,$$

and consider the function

$$\mathrm{Q}_1 = \sin\frac{\pi}{2}\frac{\eta}{(a + b)(b + d)/\mathrm{N}^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(li.)},$$

or

$$\mathrm{Q}_1 = \sin\frac{\pi}{2}\frac{ad - bc}{(a + b)(b + d)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(lii.)}.$$

This function vanishes if $\eta = 0$, and it further = unity if $b = 0$. Thus it agrees at the limits 0 and 1 with the value of the correlation coefficient. Again, when h and k are both zero, $a = d$, $b = c$, and $\mathrm{Q}_1 = \sin\frac{\pi}{2}\frac{a - b}{a + b}$, is thus r by (xx.). Hence we have found a function which vanishes with r and equals unity with r, while it is also equal to r if the divisions of the table be taken through the medians.

Now, I take it that these are very good conditions to make for any function of a, b, c, d which is to vanish with the "transfer," and to serve as a measure of the degree of dependent variability, or what Mr. YULE has termed the degree of "association." Mr. YULE has selected for his coefficient of association the expression

$$\mathrm{Q}_2 = \frac{ad - bc}{ad + bc} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(liii.)}.$$

This vanishes with the transfer, equals unity if b or c be zero, and minus unity if a or d be zero. The latter is, of course, unnecessary if we agree to arrange a, b, c, d so that ad is always greater than bc. Now it is clear that Q_2 possesses a great advantage over Q_1 in rapidity of calculation, but the coefficient of correlation is also a coefficient which measures the association, and it is a great advantage to select one which agrees to the closest extent with the correlation, for then it enables us to determine other important features of the system.

If we do not make all the above conditions, we easily obtain a number of coefficients which would vanish with the transfer. Thus for example the correlation r_{hk} of Equation (xxxviii.) is such an expression.* It has the advantage of a symmetrical form, and has a concise physical meaning. It does not, however, become unity when

---

* In fact (xxxvii.) gives us $\epsilon = \sigma_h \sigma_k r_{hk}$.

either, but not both, b and c vanish, nor does it, unless we multiply it by $\pi/2$ and take its sine, equal the coefficient of correlation when $a = d$ and $b = c$.

Again, we might deduce a fairly simple approximation to the coefficient of correlation from the Equation (xxiv.) for θ, using only its first few terms. Thus we find

$$\text{Sin } 2\pi \frac{ad - bc}{N^2\{1 - \frac{1}{2}(\chi_1{}^2 + \chi_2{}^2)\} + \pi\chi_1\chi_2(ad - bc)} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(liv.)},$$

where

$$\chi_1 = \sqrt{\frac{\pi}{2}} \frac{(a + c) - (b + d)}{N},$$

$$\chi_2 = \sqrt{\frac{\pi}{2}} \frac{(a + b) - (c + d)}{N},$$

as an expression which vanishes with the transfer, and will be fairly close to the coefficient of correlation. It is not, however, exactly unity when either b or c is zero. But without entering into a discussion of such expressions, we can write several down which fully satisfy the three conditions:—

(i.) Vanishing with the transfer.

(ii.) Being equal to unity if b or $c = 0$.

(iii.) Being equal to the correlation for median divisions.

Such are, for example:—

$$Q_3 = \sin \frac{\pi}{2} \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} - \sqrt{bc}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lv.)},$$

$$Q_4 = \sin \frac{\pi}{2} \frac{1}{1 + \frac{2bc \qquad N}{(ad - bc)(b + c)}}, \; ad > bc \quad . \quad . \quad . \quad \text{(lvi.)},$$

$$Q_5 = \sin \frac{\pi}{2} \frac{1}{\sqrt{1 + \kappa^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lvii.)},$$

where

$$\kappa^2 = \frac{4abcd \, N^2}{(ad - bc)^2 (a + d)(b + c)}.$$

Only by actual examination of the numerical results has it seemed possible to pick out the most efficient of these coefficients. Q_1 was found of little service. The following table gives the values of Q_2, Q_3, Q_4, and Q_5 in the case of fifteen series selected to cover a fairly wide range of values:—

No.	r.	h.	k.	Q_2.	Q_3.	Q_4.	Q_5.
1	·5939 ± ·0247	− ·0873	− ·4163	·7067	·6054	·6168	·6100
2	·5557 ± ·0261	− ·4189	− ·4163	·6688	·5657	·5405	·5570
3	·5529 ± ·0247	− ·0873	− ·0012	·6828	·5809	·5699	·5813
4	·5264 ± ·0264	+ ·2743	+ ·3537	·6345	·5331	·5200	·5283
5	·5213 ± ·0294	+ ·6413	+ ·6966	·6530	·5511	·4878	·5160
6	·5524 ± ·0307	+1·0234	+ ·3537	·7130	·6118	·6169	·6138
7	·5422 ± ·0288	+ ·6463	+ ·5828	·6693	·5673	·5136	·5452
8	·2222 ± ·0162	+ ·3190	+ ·3190	·2840	·2268	·2164	·2251
9	·3180 ± ·0361	+ ·1381	+ ·0696	·3959	·3185	·3176	·3183
10	·5954 ± ·0272	+1·5114	+ ·7414	·7860	·7100	·6099	·6803
11	·4708 ± ·0292	+ ·0865	− ·0054	·5692	·4712	·4720	·4715
12	·2335 ± ·0335	+ ·0405	+ ·0054	·2996	·2385	·2385	·2385
13	·2451 ± ·0205	+ ·2707	+ ·0873	·3103	·2473	·2456	·2470
14	·1002 ± ·0394	+ ·4557	+ ·1758	·1311	·1032	·0993	·1029
15	·6928 ± ·0164	+ ·5814	+ ·5814	·8032	·7108	·6699	·6897

Now an examination of this table shows that notwithstanding the extreme elegance and simplicity of Mr. YULE'S coefficient of association Q_2, the coefficients Q_3, Q_4, and Q_5, which satisfy also his requirements, are much nearer to the values assumed by the correlation. I take this to be such great gain that it more than counterbalances the somewhat greater labour of calculation. If we except cases (6) and (10), in which h or k take a large value exceeding unity, we find that Q_3, Q_4, and Q_5 in the fifteen cases hardly differ by as much as the probable error from the value of the correlation. If we take the mean percentage error of the difference between the correlation and these coefficients, we find

Mean difference of $Q_2 = 24{\cdot}38$ per cent.
,, ,, $Q_3 = 3{\cdot}95$,,
,, ,, $Q_4 = 2{\cdot}94$,,
,, ,, $Q_5 = 2{\cdot}72$,,

Thus although there is not much to choose between Q_4 and Q_5, we can take Q_5 as a good measure of the degree of independent variation.

The reader may ask: Why is it needful to seek for such a measure? Why cannot we always use the correlation as determined by the method of this paper? The answer is twofold. We want first to save the labour of calculating r for cases where the data are comparatively poor, and so reaching a fairly approximate result rapidly. But labour-saving is never a wholly satisfactory excuse for adopting an inferior method. The second and chief reason for seeking such a coefficient as Q lies in the fact that all our reasoning in this paper is based upon the normality of the frequency. We require to free ourselves from this assumption if possible, for the difficulty, as is exemplified in Illustration V. below, is to find material which actually obeys within the probable errors any such law. Now, by considering the coefficient of regression, $r\sigma_1/\sigma_2 = S(xy)/(N\sigma_1\sigma_2)$, as the slope of the line which best fits the series of points determined as the means of arrays of x for given values of y, we have once and for all freed ourselves from the difficulties attendant upon assuming normal frequency. We become indifferent to the deviations from that law, merely observing how closely or not our means of arrays fall on a line. When we are not given arrays but gross grouping under certain divisions, we have seen that the "transfer" is also a physical quantity of a significance independent of normality. We want accordingly to take a function which vanishes with the transfer, and does not diverge widely from the correlation in cases that we can test. Here the correlation is not taken as something peculiar to normal distributions, but something significant for all distributions whatever. Such a function of a suitable kind appears to be given by Q_5.

§ 6. *On the "Excess" and its Relation to Correlation and Relative Variability.*

There is another method of dealing with the correlation of characters for which we cannot directly discover a quantitative scale which deserves consideration. It is capable of fairly wide application, but, unlike the methods previously discussed, it requires the data to be collected in a special manner. It has the advantage of not applying only to the normal surface of frequency, but to any surface which can be converted into a surface of revolution by a slide and two stretches.

It is well known that not only the normal curve but the normal surface has a type form from which all others can be deduced by stretching or stretching and sliding. Thus in 1895 the Cambridge Instrument Company made for the instrument room at University College, London, a "biprojector," an instrument for giving arbitrary stretches in two directions at right angles to any curve. In this manner by the use of type-templates we were able to draw a variety of curves with arbitrary parameters, *e.g.*, all ellipses from one circle, parabolas from one parabola, normal curves from one normal curve template. Somewhat later Mr. G. U. YULE commenced a model of a normal frequency surface on the BRILL system of interlaced curves. This, by the variable amount of slide given to its two rectangular systems of normal curves, illustrated the changes from zero to perfect correlation. This model was exhibited at a College *soirée* in June, 1897. Geometrically this property has been taken by Mr. W. F. SHEPPARD as the basis of his valuable paper on correlation in the 'Phil. Trans.,' A, vol. 192, pp. 101–167. It is a slight addition to, and modification of, his results that I propose to consider in this section.

The equation to the normal frequency surface is, as we have seen in § 1,

$$z = \frac{\mathrm{N}}{2\pi\sigma_1\sigma_1\sqrt{1-r^2}} \text{ expt. } \left\{-\tfrac{1}{2}\left(\frac{x^2}{\sigma_1{}^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2{}^2}\right)\frac{1}{1-r^2}\right\}$$

Now write $x/(\sigma_1\sqrt{1-r^2}) = x'$, $y/\sigma_2 = y'$. This is merely giving the surface two uniform stretches (or squeezes) parallel to the coordinate axes. We have for the frequency of pairs lying between x, $x + \delta x$, and y, $\delta + \delta y$,

$$z\delta x\delta y = \frac{\mathrm{N}}{2\pi}\,\delta x'\delta y' \text{ expt.}\left\{-\tfrac{1}{2}\left(\left(x' - \frac{ry'}{\sqrt{1-r^2}}\right)^2 + y'^2\right)\right\}.$$

Now give the area a uniform slide parallel to the axis of x defined by $r/\sqrt{1-r^2}$ at unit distance from that axis. This will not change the basal unit of area $\delta\alpha = \delta x'\delta y'$, and analytically we may write

$$\mathrm{X} = x' - y'r/\sqrt{1-r^2}, \qquad \mathrm{Y} = y', \qquad \mathrm{R}^2 = \mathrm{X}^2 + \mathrm{Y}^2.$$

Whence we find

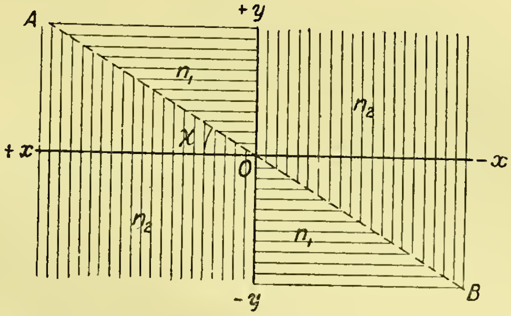
$$z\delta x\delta y = \frac{\mathrm{N}}{2\pi}\,\delta\alpha \text{ expt. } (-\tfrac{1}{2}\mathrm{R}^2).$$

This is the mechanical changing of the YULE-BRILL model analytically represented. The surface is now one of revolution, and the proof would have been precisely the same if we had written in the above results any function f, instead of the exponential.* It is easy to see that any volume cut off by two planes through the axis of the surface is to the whole volume as the angle between the two planes is to four right angles. Further the corresponding volumes of this surface and the original surface are to each other as unity to the product of the two stretches. Lastly, any plane through the z-axis of the original solid remains a plane through the z-axis after the two stretches and the slide. These points have all been dealt with by Mr. SHEPPARD (p. 101 *et seq.*, *loc. cit.*). I will here adopt his notation $r = \cos \mathrm{D}$, and term with him D the *divergence*. Thus cot D is (in the language of the theory of strain) the slide, and D is the angle between the strained positions of the original x and y directions. Now consider any plane which makes an angle χ with the plane of xz before strain. Then, since the contour lines of the correlation surface are ellipses, the volumes of the surface upon the like shaded opposite angles of the plan diagram below will be equal; and if they be n_1 and n_2, then $n_1 + n_2 = \frac{1}{2}\mathrm{N}$. If n_1' and n_2' be the volumes after strain, then by what precedes we shall have

$$n_1 = \sigma_1\sigma_2\sqrt{1-r^2} \times n_1', \qquad n_2 = \sigma_1\sigma_2\sqrt{1-r^2} \times n_2',$$

and

$$(n_2 - n_1)/(n_1 + n_2) = (n_2' - n_1')/(n_1' + n_2').$$



---

* The generalisation is not so great as might at first appear, for I have convinced myself that this property of conversion into a surface of revolution by stretches and slides does not hold for actual cases of markedly skew correlation.

Now n_1' and n_2' will be as the angles between the strained positions of the planes bounding n_1 and n_2. Ox does not change its direction. Oy is turned through an angle $\pi/2 - \mathrm{D}$ clockwise, and χ becomes χ'', say. Hence

$$n_1' : n_2' :: \frac{\pi}{2} - \chi'' + \frac{\pi}{2} - \mathrm{D} : \frac{\pi}{2} + \chi'' - \frac{\pi}{2} + \mathrm{D}.$$

or

$$(n_2' - n_1')/(n_2' + n_1') = \frac{2}{\pi}(\chi'' + \mathrm{D}) - 1.$$

Let us write $\mathrm{E}_1 = 2(n_2 - n_1)$ and term it the *excess* for the y-character for the line AB. Then we easily find:

$$\tan\left(\frac{\mathrm{E}_1}{\mathrm{N}}\frac{\pi}{2} + \frac{\pi}{2}\right) = \tan(\chi'' + \mathrm{D}) = \frac{\cot\chi'' + \cot \mathrm{D}}{\cot\chi''\cot \mathrm{D} - 1} \quad . \quad . \quad . \quad . \quad \text{(lviii.)}.$$

It remains to determine $\tan\chi''$ and substitute. The stretches alter $\tan\chi$ into $\tan\chi'$, such that

$$\tan\chi' = \frac{\sigma_1\sqrt{1 - r^2}}{\sigma_2}\tan\chi.$$

Further, by the slide

$$\cot\chi'' = \cot\chi' - \cot \mathrm{D} = \frac{\sigma_2}{\sigma_1\sqrt{1 - r^2}}\cot\chi - \cot \mathrm{D}.$$

Hence we have by (lviii.) above

$$-\cot\left(\frac{\mathrm{E}_1}{\mathrm{N}}\frac{\pi}{2}\right) = \frac{\sigma_2}{\sigma_1\sqrt{1 - r^2}}\cot\chi\Big/\left(\frac{\sigma_2}{\sigma_1\sqrt{1 - r^2}}\cot\chi\cot \mathrm{D} - \cot^2 \mathrm{D} - 1\right),$$

or,

$$-\tan\left(\frac{\mathrm{E}_1}{\mathrm{N}}\frac{\pi}{2}\right) = \cot \mathrm{D} - \frac{\sigma_1}{\sigma_2}\frac{\tan\chi}{\sin \mathrm{D}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lix.)}.$$

Now the excess E_1 is the difference of the frequencies in the sum of the strips of the volume made by planes parallel to the plane yz on the two sides of the plane ABz (defined by χ), taken without regard to sign. For on one side of the mean yy this is $n_2 - n_1$, and on the other $-(n_1 - n_2)$. Hence we have this definition of E_1, the *column excess* for any line through the mean of a correlation table: *Add up the frequencies above and below the line in each column and take their differences without regard to sign, and their sum is the column excess.*

If we are dealing with an actual correlation table and not with a method of collecting statistics, then care must be taken to properly proportion the frequencies in the column in which the mean occurs, and also in the groups which are crossed by the line. It is the difficulty of doing this satisfactorily, especially if the grouping, as in eye and coat colour, is large and somewhat rough, that hinders the effective use of the method, if the statistics have not been collected *ad hoc.*

Now let E_2 be the *row excess* for the line AB, defined in like manner, then we have in the same way

$$-\tan\left(\frac{E_2}{N}\frac{\pi}{2}\right) = \cot D - \frac{\sigma_2}{\sigma_1}\frac{\cot\chi}{\sin D} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{lix.}^{\text{bis}}).$$

Now eliminate σ_2/σ_1 between (lix.) and (lix.$^{\text{bis}}$); then

$$\left(\tan\left(\frac{E_1}{N}\frac{\pi}{2}\right) + \cot D\right)\left(\tan\left(\frac{E_2}{N}\frac{\pi}{2}\right) + \cot D\right) = \frac{1}{\sin^2 D}.$$

Whence we deduce

$$\cot D = \cot\frac{E_1 + E_2}{N}\frac{\pi}{2},$$

and, therefore,

$$r = \cos D = \cos\frac{E_1 + E_2}{N}\frac{\pi}{2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{lx.}).$$

Substituting for D in (lix.) we find further

$$\frac{\sigma_1}{\sigma_2} = \cot\chi\cos\left(\frac{E_2}{N}\frac{\pi}{2}\right)\Big/\cos\left(\frac{E_1}{N}\frac{\pi}{2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{lxi.}).$$

Thus Equations (lx.) and (lxi.) give the coefficient of correlation and the relative variability of the two characters. The latter is, I believe, quite new, the former novel in form.

If we call m_1 the frequency in the angle χ (AOx of the figure above), then it is easy to see that $E_1 = 2(n_2 - n_1) = N - 4n_1$, and similarly $E_2 = N - 4m_1$. Thus $(E_1 + E_2)/N = 2(N - 2(n_1 + m_1))/N$. But $n_1 + m_1$ is the frequency in the first quadrant. This Mr. SHEPPARD terms P, while that in the second he terms R. We have thus $(E_1 + E_2)/N = 2R/(R + P)$, or

$$r = \cos\frac{R}{R + P}\pi \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{lxii.}),$$

i.e., Mr. SHEPPARD's fundamental result* (' Phil. Trans.,' A, vol. 192, p. 141).

We can, of course, get Mr. SHEPPARD's result directly if we put $\chi = 0$, when we have at once $E_1 = 2(R - P)$, $E_2 = N = 2(R + P)$, and the result follows.

Equation (lxi.) may also be written in the form

$$\frac{\sigma_1}{\sigma_2} = \cot\chi\sin\left(\frac{m_1}{N}2\pi\right)\Big/\sin\left(\frac{n_1}{N}2\pi\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{lxiii.}).$$

If we put $\chi = 0$, then m_1 becomes zero, and the right-hand side of (lxiii.) is indeterminate. If we proceed, however, to the limit by evaluating the frequency in an indefinitely thin wedge of angle χ, we reach merely the identity $\sigma_1/\sigma_2 = \sigma_1/\sigma_2$. Hence there is no result corresponding to (lxi.) to be obtained by taking Mr. SHEPPARD's case of $\chi = 0$.

The following are the values of the probable errors of the quantities involved :—

---

* In the actual classification of data (lx.) and (lxii.) suggest quite different processes. We can apply (lx.) where (lxii.) is difficult or impossible, *e.g.*, correlation in shading of birds' eggs from the same clutch.

$$\text{Probable error of } E_1 = \cdot 67449\sqrt{N(1 - E_1^2/N^2)} \quad . \quad . \quad . \quad . \quad (\text{lxiv.}).$$

$$\text{,, \quad ,,} \quad E_2 = \cdot 67449\sqrt{N(1 - E_2^2/N^2)} \quad . \quad . \quad . \quad . \quad (\text{lxv.}).$$

$$\text{Correlation between errors in } E_1 \text{ and } E_2 = -\sqrt{\frac{(1 - E_1/N)(1 - E_2/N)}{(1 + E_1/N)(1 + E_2/N)}} \quad . \quad . \quad (\text{lxvi.}).$$

$$\text{Probable error in } r = \frac{\cdot 67449 \sin D\sqrt{D(\pi - D)}}{\sqrt{N}} \quad . \quad . \quad . \quad . \quad (\text{lxvii.}),$$

where $D = \frac{E_1 + E_2}{N}\frac{\pi}{2}$ (*cf.* SHEPPARD, *loc. cit.*, p. 148).

Probable error in ratio $\sigma_1/\sigma_2 =$

$$\frac{\cdot 67449}{\sqrt{N}}\frac{\sigma_1}{\sigma_2}\frac{\pi}{2}\left\{\left(1 - \frac{E_1^2}{N^2}\right)\tan^2\left(\frac{E_1}{N}\frac{\pi}{2}\right) + \left(1 - \frac{E_2^2}{N_2}\right)\tan^2\left(\frac{E_2}{N}\frac{\pi}{2}\right)\right.$$
$$\left. + 2\left(1 - \frac{E_1}{N}\right)\left(1 - \frac{E_2}{N}\right)\tan\left(\frac{E_1}{N}\frac{\pi}{2}\right)\tan\left(\frac{E_2}{N}\frac{\pi}{2}\right)\right\}^{\frac{1}{2}} \quad . \quad . \quad (\text{lxviii.}).$$

The application of the method here discussed to statistics without quantitative scale can now be indicated. If the characters we are dealing with have the same scale, although it be unknown, then, if the quantitative *order* be maintained, *i.e.*, individuals arranged in order of lightness or darkness of coat or eye-colour, the diagonal line on the table at 45° will remain unchanged, however we may suppose parts of the scale to be distorted, for the distortion will be the same at corresponding points of both axes. Further, if we suppose the mean of the two characters to be the same, this 45° line will pass through that mean, and will serve for the line AB of the above investigation. In this case we must take $\tan \chi = 1$, and consequently (lxi.) becomes

$$\sigma_1/\sigma_2 = \cos\left(\frac{E_2}{N}\frac{\pi}{2}\right)\Big/\cos\left(\frac{E_1}{N}\frac{\pi}{2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{lxix.}).$$

We can even, when the mean is a considerable way off the 45° line, get, in some cases, good results. Thus, the correlation in stature of husband and wife worked out by the ordinary product moment process is ·2872. But in this case $E_1 = 382{\cdot}062$ $E_2 = 806{\cdot}425$, and this gives the correlation ·2994. On the other hand, the actual ratio of variabilities is 1·12, while (lxix.) makes it 2·76! This arises from the fact that the errors in E_1 and E_2, due to the mean being off the 45° line, tend to cancel in $E_1 + E_2$, but tend in directly opposite directions in the ratio of the cosines. Similarly the correlation between father and son works out ·5666, which may be compared with the values given in Illustration V. below, ranging from ·5198 to ·5939. Again, correlation in eye-colour between husband and wife came out by the excess process ·0986, and by the process given earlier in the present Memoir ·1002. But all these are favourable examples, and many others gave much worse results. We ought really only to apply it to find σ_1/σ_2 when the means are on the 45° line, as in the correlation of the same character in brethren, and even in this case the statistics ought to be collected *ad hoc*, *i.e.*, we ought to make a very full quantitative order, and then notice for each individual case the number above and below the type. For example, suppose we had a diagram of some twenty-five to thirty eye tints in order (*e.g.*, like BERTRAND'S), then we take any individual, note his tint, and observe how many relatives of a particular class—brethren or cousins, say—have lighter and how many darker eyes; the difference of the two would be the excess for this individual. The same plan would be possible with horses' coat-colour and other characters. After trying the plan of the excesses on the data at my disposal for horses' coat-colour and human eye-colour (which were not collected *ad hoc*), I abandoned it for the earlier method of this Memoir; for, the classification being in large groups, the proportioning of the excess (as well as the differences in the means) introduced too great errors for such investigations.

§ 7. *On a Generalisation of the Fundamental Theorem of the Present Memoir.*

If we measure deviations in units of standard deviations, we may take for the equation to the correlation surface for n variables

$$z = \frac{\mathrm{N}}{(2\pi)^{\frac{1}{2}n}\sqrt{\mathrm{R}}} e^{-\frac{1}{2}\left\{\mathrm{S}_1\left(\frac{\mathrm{R}_{ss}}{\mathrm{R}} x_s{}^2\right) + 2\mathrm{S}_2\left(\frac{\mathrm{R}_{ss'} x_s x_{s'}}{\mathrm{R}}\right)\right\}} \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(lxx.)},$$

where

$$\mathrm{R} = \begin{vmatrix} 1 & r_{12} & r_{13} & \cdots & \cdots & r_{1n} \\ r_{21} & 1 & r_{23} & \cdots & \cdots & r_{2n} \\ r_{31} & r_{32} & 1 & \cdots & \cdots & r_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{n-1,1} & r_{n-1,2} & r_{n-1,3} & \cdots & 1 & r_{n-1,n} \\ r_{n,1} & r_{n,2} & r_{n,3} & \cdots & r_{n-1,n} & 1 \end{vmatrix}$$

and R_{pq} is the minor obtained by striking out the pth row and qth column. r_{pq} is, of course, the correlation between the pth and qth variables, and equals r_{qp}. S_1 denotes a summation for s from 1 to n, and S_2 a summation of every possible pair out of the n quantities 1 to n.

Now take the logarithmic differential of z with regard to r_{pq}. We find

$$\begin{aligned} \frac{1}{z}\frac{dz}{dr_{pq}} &= -\frac{1}{2\mathrm{R}}\frac{d\mathrm{R}}{dr_{pq}} - \tfrac{1}{2}\mathrm{S}_1\left\{\frac{d}{dr_{pq}}\left(\frac{\mathrm{R}_{ss}}{\mathrm{R}}\right)x_s{}^2\right\} - \mathrm{S}_2\left\{\frac{d}{dr_{pq}}\left(\frac{\mathrm{R}_{ss'}}{\mathrm{R}}\right)x_s x_{s'}\right\} \\ &= -\frac{\mathrm{R}_{pq}}{\mathrm{R}} + \mathrm{S}_1\left(\frac{\mathrm{R}_{ps}\,\mathrm{R}_{qs}}{\mathrm{R}^2}x_s{}^2\right) + \mathrm{S}_2\left(\frac{\mathrm{R}_{ps}\,\mathrm{R}_{qs'} + \mathrm{R}_{ps'}\,\mathrm{R}_{qs}}{\mathrm{R}^2}x_s x_{s'}\right) \end{aligned}$$

For

$$d\mathrm{R}/dr_{pq} = 2\mathrm{R}_{pq}$$

and, generally, whether s is or is not $= s'$, or these are or are not $= p$ and q, we have

$$\frac{d}{dr_{pq}}\left(\frac{R_{ss'}}{R}\right) = -\frac{R_{ps}\,R_{qs'} + R_{ps'}\,R_{qs}}{R^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxi.)}.$$

This follows thus:

$$\frac{d}{dr_{pq}}\left(\frac{R_{ss'}}{R}\right) = \frac{1}{R}\,\frac{dR_{ss'}}{dr_{pq}} - \frac{R_{ss'}}{R^2}\,\frac{dR}{dr_{pq}} = \frac{1}{R}\,\frac{dR_{ss'}}{dr_{pq}} - \frac{2R_{ss'}\,R_{pq}}{R^2},$$

or we have to show that

$$\begin{aligned}\frac{dR_{ss'}}{dr_{pq}} &= \frac{2R_{ss'}\,R_{pq} - R_{ps}\,R_{qs'} - R_{ps'}\,R_{qs}}{R} \\ &= \frac{R_{ss'}\,R_{pq} - R_{ps}\,R_{qs'}}{R} + \frac{R_{ss'}\,R_{pq} - R_{ps'}\,R_{qs}}{R} \\ &= {}_{pq}R_{ss'} + {}_{qp}R_{ss'}\end{aligned}$$

where ${}_{pq}R_{ss'}$ is the minor corresponding to the term r_{pq} in $R_{ss'}$, and ${}_{qp}R_{ss'}$ the minor corresponding to the term r_{qp}.* But this last result is obvious because $R_{ss'}$ only contains r_{pq} in two places, *i.e.*, as r_{pq} and r_{qp}.

Putting $s = s'$, we have the other identity required above, *i.e.*,

$$\frac{d}{dr_{pq}}\left(\frac{R_{ss}}{R}\right) = -\frac{2R_{ps}\,R_{qs}}{R^2}.$$

Returning now to the value for $\frac{1}{z}\,\frac{dz}{dr_{pq}}$ on the previous page, we see that the two sum terms may be expressed as a product, or we may put

$$\frac{1}{z}\,\frac{dz}{dr_{pq}} = -\frac{R_{pq}}{R} + S_1\left(\frac{R_{ps}}{R}\,x_s\right) \times S_1\left(\frac{R_{qs}}{R}\,x_s\right).$$

Now write

$$z = \frac{N}{(2\pi)^{\frac{1}{2}n}\sqrt{R}}\,e^{-\phi}.$$

Then $\quad \frac{d\phi}{dx_p} = S_1\left(\frac{R_{ps}}{R}\,x_s\right), \quad \frac{d\phi}{dx_q} = S_1\left(\frac{R_{qs}}{R}\,x_s\right)$ and $\frac{d^2\phi}{dx_p\,dx_q} = \frac{R_{pq}}{R}.$

Hence $\quad \frac{1}{z}\,\frac{dz}{dr_{pq}} = -\frac{d^2\phi}{dx_p\,dx_q} + \frac{d\phi}{dx_p}\,\frac{d\phi}{dx_q}.$

Now differentiate log z with regard to x_p. Then

$$\frac{dz}{dx_p} = -z\,\frac{d\phi}{dx_p}$$

---

* See also SCOTT, 'Theory of Determinants,' p. 59.

$$\frac{d^2z}{dx_pdx_q} = -z\frac{d^2\phi}{dx_pdx_q} - \frac{dz}{dx_q}\frac{d\phi}{dx_p}$$

$$\frac{1}{z}\frac{d^2z}{dx_pdx_q} = -\frac{d^2\phi}{dx_pdx_q} + \frac{d\phi}{dx_p}\frac{d\phi}{dx_q}$$

Thus finally

$$\frac{dz}{dr_{pq}} = \frac{d^2z}{dx_pdx_q} \quad . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \quad \text{(lxxii.)}.$$

In other words, the operator d/dr_{pq} acting on z can always be replaced by the operator d^2/dx_pdx_q.

Let $d/d\rho_{pq}$ denote the effect of applying the operator d/dr_{pq} to z, and putting r_{pq} zero after *all* differentiations have been performed, then the effect of this operator will be the same as if we used d^2/dx_pdx_q on z, putting r_{pq} zero *before* differentiation. Generally, let F be any series of operations like d/dr_{pq}, then we see that

$$\mathrm{F}\left(\frac{d}{dr_{pq}}, \ \frac{d}{dr_{p'q'}}, \ \frac{d}{dr_{p''q''}} \ . \ . \ . \ . \ . \ . \right) z$$

$$= \mathrm{F}\left(\frac{d^2}{dx_pdx_q}, \ \frac{d^2}{dx_{p'}dx_{q'}}, \ \frac{d^2}{dx_{p''}dx_{q''}}, \ . \ . \ . \ . \ . \ . \right) \frac{\mathrm{N}}{(2\pi)^{\frac{1}{2}n}} e^{-\frac{1}{2}\mathrm{S}_1(x_s^2)}.$$

Now let F be the function which gives the operation of expanding z by MACLAURIN'S theorem in powers of the correlation coefficients, *i.e.*,

$$\mathrm{F} = e^{\mathrm{S}_2\left(r_{ss'}\frac{d}{d\rho_{ss'}}\right)},$$

then

$$z = e^{\mathrm{S}_2\left(r_{ss'}\frac{d}{d\rho_{ss'}}\right)} z = \frac{\mathrm{N}}{(2\pi)^{\frac{1}{2}n}} e^{\mathrm{S}_2\left(r_{ss'}\frac{d}{dx_s dx_{s'}}\right)} e^{-\frac{1}{2}\mathrm{S}_1(x_s^2)}.$$

This is the generalised form of result (xiv.) reached above.

Now let

$$z_0 = \frac{\mathrm{N}}{(2\pi)^{\frac{1}{2}n}} e^{-\frac{1}{2}\mathrm{S}_1(x_s^2)},$$

then z_0 is the ordinate of a frequency surface of the nth order, in which the distribution of the n variables is absolutely independent. We have accordingly the extremely interesting geometrical interpretation that the operator

$$e^{\mathrm{S}_2\left(r_{ss'}\frac{d^2}{dx_sdx_{s'}}\right)},$$

applied to a surface of frequency for n independent variables converts it into a surface of frequency for n dependent variables, the correlation between the sth and s'th variables being $r_{ss'}$.*

---

* I should like to suggest to the pure mathematician the interest which a study of such operators would have, and in particular of the generalised form of projection in hyperspace indicated by them.

Expanding, we have

$$z = z_0 + \mathrm{S}_2\left(r_{rs'} \frac{d^2}{dx_s dx_{s'}}\right) z_0 + \frac{1}{\underline{2}} \left\{ \mathrm{S}_2 \left(r_{ss'} \frac{d^2}{dx_s dx_{s'}} \right) \right\}^2 z_0$$

$$+ \ldots + \frac{1}{\underline{|m}} \left\{ \mathrm{S}_2 \left(r_{ss'} \frac{d^2}{dx_s dx_{s'}} \right) \right\}^m z_0 + \ldots \ldots \quad \text{(lxxiii.)}.$$

Our next stage is to evaluate the operation

$$\mathrm{S}_2 \left(r_{ss'} \frac{d^2}{dx_s dx_{s'}} \right)^m z_0.$$

Let us put

$${}_sv_1 = x_s, \qquad {}_sv_2 = x_s^2 - 1, \qquad {}_sv_3 = x_s(x_s^2 - 3),$$

and ${}_sv_p$ = the pth function of x_s as defined by (xv.).

Let ϵ_s be a symbol such that ϵ_s^p represents ${}_sv_p$. Then we shall show that

$$\mathrm{S}_2 \left(r_{ss'} \frac{d^2}{dx_s dx_{s'}} \right)^m z_0 = z_0 \left\{ \mathrm{S}_r \left(r_{ss'} \epsilon_s \epsilon_{s'} \right) \right\}^m \quad \ldots \ldots \quad \text{(lxxiv.)}.$$

We shall prove this by induction.

By (xii.)

$${}_sv_{m+1} = x_s \, {}_sv_m - m \, {}_sv_{m-1},$$

or

$$\epsilon_s^{m+1} = x_s \, \epsilon_s^m - m \, \epsilon_s^{m-1},$$

and by (xvi.)

$$\frac{d_s v_m}{dx_s} = m \, {}_sv_{m-1}, \quad \text{or} \quad \frac{d\epsilon_s^m}{dx_s} = m \, \epsilon_s^{m-1}.$$

Now, let $\chi(\epsilon_s)$ be any function of ϵ_s

$$= \mathrm{S}(\mathrm{A}_q \epsilon_s^q),$$

if we suppose it can be expanded in powers of ϵ_s.

Then

$$\begin{aligned} \frac{d}{dx_s} \chi(\epsilon_s) &= \mathrm{S}\left(\mathrm{A}_q \frac{d}{dx_s} (\epsilon_s^q)\right) \\ &= \mathrm{S}(\mathrm{A}_q \, q\epsilon_s^{q-1}) \\ &= \mathrm{S}(\mathrm{A}_q(x_s \, \epsilon_s^q - \epsilon_s^{q+1})) \\ &= x_s \mathrm{S}(\mathrm{A}_q \, \epsilon_s^q) - \epsilon_s \, \mathrm{S}(\mathrm{A}_q \, \epsilon_s^q) \\ &= (x_s - \epsilon_s) \, \chi(\epsilon_s) \quad \ldots \ldots \ldots \quad \text{(lxxv.)}. \end{aligned}$$

Similarly

$$\frac{d^2}{dx_s dx_{s'}} \chi(\epsilon_s, \epsilon_{s'}) = (x_s - \epsilon_s)(x_{s'} - \epsilon_{s'}) \, \chi(\epsilon_s, \epsilon_{s'}) \quad \ldots \ldots \quad \text{(lxxvi.)}.$$

Now suppose that

$$\left\{ \mathrm{S}_2 \left(r_{ss'} \frac{d^2}{dx_s dx_{s'}} \right) \right\} z_0 = z_0 \{ \mathrm{S}_2 (r_{ss'} \epsilon_s \epsilon_{s'}) \}^m,$$

then

$$\left\{\mathrm{S}_2\left(r_{ss'}\frac{d^2}{dx_s dx_{s'}}\right)\right\}^{m+1} z_0 = \mathrm{S}_2\left(r_{ss'}\frac{d^2}{dx_s dx_{s'}}\right) z_0 \mathrm{U},$$

where U stands for $\{\mathrm{S}_2(r_{ss'}\epsilon_s\epsilon_{s'})\}^m$.

Hence, remembering that $dz_0/dx_s = -z_0 x_{s'}$,

$$\begin{aligned}\left\{\mathrm{S}_2\left(r_{ss'}\frac{d^2}{dx_s dx_{s'}}\right)\right\}^{m+1} z_0 &= z_0\mathrm{S}_2(r_{ss'}x_s x_{s'})\mathrm{U} + z_0\mathrm{S}_2\left(r_{ss'}\frac{d^2\mathrm{U}}{dx_s dx_{s'}}\right)\\ &\quad - z_0\mathrm{S}_2\left(r_{ss'}\left(x_s\frac{d\mathrm{U}}{dx_{s'}} + x_{s'}\frac{d\mathrm{U}}{dx_s}\right)\right)\\ &= z_0\mathrm{S}_2(r_{ss'}x_s x_{s'})\mathrm{U} + z_0\mathrm{S}_2\{r_{ss'}(x_s - \epsilon_s)(x_{s'} - \epsilon_{s'})\}\mathrm{U}\\ &\quad - z_0\mathrm{S}_2\{r_{ss'}\{x_s(x_{s'} - \epsilon_{s'}) + x_{s'}(x_s - \epsilon_s)\}\mathrm{U}\\ &= z_0\mathrm{S}_2(r_{ss'}\epsilon_s\epsilon_{s'})\mathrm{U}\\ &= z_0\{\mathrm{S}_2(r_{ss'}\epsilon_s\epsilon_{s'})\}^{m+1},\end{aligned}$$

which had to be proved.

But it is easy to show by simple differentiation that

$$\mathrm{S}_2\left(r_{ss'}\frac{d^2}{dx_s dx_{s'}}\right) z_0 = z_0\ \mathrm{S}_2(r_{ss'}\ {}_sv_1\ {}_{s'}v_1) = z_0\ \mathrm{S}_2(r_{ss'}\epsilon_s\epsilon_{s'})$$

$$\begin{aligned}\left\{\mathrm{S}_2\left(r_{ss'}\frac{d^2}{dx_s dx_{s'}}\right)\right\}^2 z_0 &= z_0\ \mathrm{S}_2(r^2_{ss'}\ {}_sv_2\ {}_{s'}v_2 + 2r_{pq}r_{ss'}\ {}_pv_1\ {}_qv_1\ {}_sv_1\ {}_{s'}v_1 + 2r_{ps}r_{ss'}\ {}_sv_2\ {}_pv_1\ {}_{s'}v_1)\\ &= z_0\ \{\mathrm{S}_2(r_{ss'}\epsilon_s\epsilon_{s'})\}^2 \qquad \text{(lxxvii.)}.\end{aligned}$$

Hence the theorem is generally true.

Thus we conclude that

$$\begin{aligned}z = z_0\Big[1 &+ \mathrm{S}_2(r_{ss'}\epsilon_s\epsilon_{s'}) + \frac{1}{\underline{|2}}\{\mathrm{S}_2(r_{ss'}\epsilon_s\epsilon_{s'})\}^2\\ &+ \ldots + \frac{1}{\underline{|m}}\{\mathrm{S}_2(r_{ss'}\epsilon_s\epsilon_{s'})\}^m + \ldots\Big] \qquad \text{(lxxviii.)}.\end{aligned}$$

It is quite straightforward, if laborious, to write down the expansion for any number of variables.

Now let Q be the total frequency of complices of variables with x_1 lying between h_1 and ∞, x_2 between h_2 and ∞, . . . x_s between h_s and ∞, . . . x_n between h_n and ∞; and let Q_0 be the frequency of such complices if there were no correlations. Then

$$\mathrm{Q} = \int_{h_1}^{\infty}\int_{h_2}^{\infty}\ldots\int_{h_3}^{\infty}\ldots\int_{h_n}^{\infty} z\, dx_1\, dx_2 \ldots dx_s \ldots dx_n$$

$$\mathrm{Q}_0 = \int_{h_1}^{\infty}\int_{h_2}^{\infty}\ldots\int_{h_s}^{\infty}\ldots\int_{h_n}^{\infty} z_0\, dx_1\, dx_2 \ldots dx_s \ldots dx_n$$

Now let

$$\beta_s \mathrm{H}_s = \frac{1}{\sqrt{2\pi}}\int_{h_s}^{\infty} e^{-\frac{1}{2}x_s^2} dx_s$$

where

$$\mathrm{H}_s = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}h_s^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxix.)}.$$

We have $\mathrm{Q}_0 = \mathrm{N}\beta_1\beta_2 \; . \; . \; \beta_s \; . \; . \; \beta_n \, \mathrm{H}_1 \, \mathrm{H}_2 \; . \; . \; \mathrm{H}_s \; . \; . \; \mathrm{H}_n.$

But by (xviii.)

$$\frac{1}{\sqrt{2\pi}} \int_{h_s}^{\infty} {}_s v_p \, e^{-\frac{1}{2}x_s^2} \, dx_s = \mathrm{H}_s \, {}_s\bar{v}_{p-1} = \mathrm{H}_s\beta_s \frac{{}_s\bar{v}_{p-1}}{\beta_s} \, ,$$

where

$${}_s\bar{v}_{p-1} = \left[{}_s v_{p-1}\right]_{x_s = h_s} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxx.)},$$

and as above,

$$\beta_s = \int_{h_s}^{\infty} e^{-\frac{1}{2}x_s^2} dx_s \Big/ e^{-\frac{1}{2}h_s^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxxi.)}.$$

Thus

$$\frac{1}{(\sqrt{2\pi})^n} \int_{h_1}^{\infty} \int_{h_2}^{\infty} \; . \; . \; . \int_{h_s}^{\infty} \; . \; . \; . \int_{h_n}^{\infty} {}_s v_p \, {}_{s'} v_{p'} \, {}_{s''} v_{p''} \; . \; . \; . \; e^{-\frac{1}{2}(x_1^2 + x_2^2 + \, . \, . \, . \, + x_s^2 + \, . \, . \, . \, + x_n^2)} \, dx_1 dx_2 \; . \; . \; . \; dx_s \; . \; . \; . \; dx_n$$

$$= \mathrm{H}_1\mathrm{H}_2 \; . \; . \; . \; \mathrm{H}_s \; . \; . \; . \; \mathrm{H}_n\beta_1\beta_2 \; . \; . \; . \; \beta_s \; . \; . \; . \; \beta_n \frac{{}_s\bar{v}_{p-1}}{\beta_s} \frac{{}_{s'}\bar{v}_{p'-1}}{\beta_{s'}} \frac{{}_{s''}\bar{v}_{p''-1}}{\beta_{s''}} \; . \; . \; . \, ,$$

or

$$\int_{h_1}^{\infty} \int_{h_2}^{\infty} \; . \; . \; . \int_{h_s}^{\infty} \; . \; . \; . \int_{h_n}^{\infty} z_0 \Pi({}_s v_p) \, dx_1 dx_2 \; . \; . \; . \; dx_s \; . \; . \; . \; dx_n = \mathrm{Q}_0 \Pi \left(\frac{{}_s\bar{v}_{p-1}}{\beta_s} \right)$$

$$. \quad . \quad . \quad . \quad . \quad \text{(lxxxii.)}.$$

where Π denotes a product of ${}_s v_p$ for any number of v's with any s and p. The rule, therefore, is very simple. We must expand the value of z in v's as given by (lxxviii.) above, then the multiple integral of this will be obtained by lowering every v's right-hand subscript by unity (remembering that ${}_s v_0 = 1$), and further dividing by the β of the left-hand subscript. The general expression up to terms of the fourth order has been written down; it involves thirty-four sums, each represented by a type term All these would only occur in the case of the correlation of eight organs, or when we have to deal with twenty-eight coefficients of correlation. Such a number seems beyond our present power of arithmetical manipulation, so that I have not printed the general expressions. At the same time, the theory of multiple correlation is of such great importance for problems of evolution, in which over and over again we have three or four correlated characters to deal with,* that it seems desirable to place on record the expansion for these cases. I give four variables up to the fourth and three variables up to the fifth order terms. Afterwards I will consider special cases.

---

* In my memoir on Prehistoric Stature I have dealt with five correlated organs, *i.e.*, ten coefficients. In some barometric investigations now in hand we propose to deal with at least fifteen coefficients, while Mr. BRAMLEY-MOORE, in the correlation of parts of the skeleton, has, in a memoir not yet published, dealt with between forty and fifty cases of four variables or six coefficients.

*Value of the Quadruple Integral in the Case of Four Variables.**

$$
\begin{aligned}
\frac{\mathrm{Q}-\mathrm{Q}_0}{\mathrm{Q}_0} = {} & \frac{r_{12}}{\beta_1\beta_2} + \frac{r_{13}}{\beta_1\beta_3} + \frac{r_{14}}{\beta_1\beta_4} + \frac{r_{23}}{\beta_2\beta_3} + \frac{r_{24}}{\beta_2\beta_4} + \frac{r_{34}}{\beta_3\beta_4} \\
& + \frac{1}{\underline{|2}}\left\{\frac{r_{12}^{\ 2}}{\beta_1\beta_2} v_1'v_1'' + \frac{r_{13}^{\ 2}}{\beta_1\beta_3} v_1'v_1''' + \frac{r_{14}^{\ 2}}{\beta_1\beta_4} v_1'v_1^{\text{iv}} + \frac{r_{23}^{\ 2}}{\beta_2\beta_3} v_1''v_1''' + \frac{r_{24}^{\ 2}}{\beta_2\beta_4} v_1''v_1^{\text{iv}}\right. \\
& + \frac{r_{34}^{\ 2}}{\beta_3\beta_4} v_1'''v_1^{\text{iv}} + \frac{2r_{12}r_{13}}{\beta_1\beta_2\beta_3} v_1' + \frac{2r_{12}r_{14}}{\beta_1\beta_2\beta_4} v_1' + \frac{2r_{13}r_{14}}{\beta_1\beta_3\beta_4} v_1' \\
& + \frac{2r_{12}r_{23}}{\beta_1\beta_2\beta_3} v_1'' + \frac{2r_{13}r_{23}}{\beta_1\beta_2\beta_3} v_1''' + \frac{2r_{14}r_{23}}{\beta_1\beta_2\beta_3\beta_4} + \frac{2r_{12}r_{34}}{\beta_1\beta_2\beta_3\beta_4} \\
& + \frac{2r_{13}r_{34}}{\beta_1\beta_3\beta_4} v_1''' + \frac{2r_{14}r_{34}}{\beta_1\beta_3\beta_4} v_1^{\text{iv}} + \frac{2r_{23}r_{34}}{\beta_2\beta_3\beta_4} v_1''' \\
& + \frac{2r_{24}r_{34}}{\beta_2\beta_3\beta_4} v_1^{\text{iv}} + \frac{2r_{12}r_{24}}{\beta_1\beta_2\beta_4} v_1'' + \frac{2r_{13}r_{24}}{\beta_1\beta_2\beta_3\beta_4} \\
& \left. + \frac{2r_{14}r_{24}}{\beta_1\beta_2\beta_4} v_1^{\text{iv}} + \frac{2r_{23}r_{24}}{\beta_2\beta_3\beta_4} v_1''\right\} \\
+ \frac{1}{\underline{|3}} & \left\{\frac{r_{12}^{\ 3}}{\beta_1\beta_2} v_2'v_2'' + \frac{r_{13}^{\ 3}}{\beta_1\beta_3} v_2'v_2''' + \frac{r_{14}^{\ 3}}{\beta_1\beta_4} v_2'v_2^{\text{iv}} + \frac{r_{23}^{\ 3}}{\beta_1\beta_3} v_2''v_2''' + \frac{r_{24}^{\ 3}}{\beta_2\beta_4} v_2''v_2^{\text{iv}}\right. \\
& + \frac{r_{34}^{\ 3}}{\beta_3\beta_4} v_2'''v_2^{\text{iv}} + \frac{3r_{12}^{\ 2}r_{13}}{\beta_1\beta_2\beta_3} v_2'v_1'' + \frac{3r_{12}r_{13}^{\ 2}}{\beta_1\beta_2\beta_3} v_2'v_1''' + \frac{3r_{12}^{\ 2}r_{14}}{\beta_1\beta_2\beta_4} v_2'v_1''' \\
& + \frac{3r_{12}r_{14}^{\ 2}}{\beta_2\beta_2\beta_4} v_2'v_1^{\text{iv}} + \frac{3r_{13}^{\ 2}r_{14}}{\beta_1\beta_2\beta_4} v_2'v_1''' + \frac{3r_{13}r_{14}^{\ 2}}{\beta_1\beta_3\beta_4} v_2'v_1^{\text{iv}} \\
& + \frac{3r_{12}^{\ 2}r_{23}}{\beta_1\beta_2\beta_3} v_2''v_1' + \frac{3r_{12}r_{23}^{\ 2}}{\beta_1\beta_2\beta_3} v_2''v_1''' + \frac{3r_{13}^{\ 2}r_{23}}{\beta_1\beta_2\beta_3} v_2'''v_1' \\
& + \frac{3r_{13}r_{23}^{\ 2}}{\beta_1\beta_2\beta_3} v_2'''v_1'' + \frac{3r_{14}^{\ 2}r_{23}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1^{\text{iv}} + \frac{3r_{14}r_{23}^{\ 2}}{\beta_1\beta_2\beta_3\beta_4} v_1''v_1''' \\
& + \frac{3r_{12}^{\ 2}r_{24}}{\beta_1\beta_2\beta_4} v_2''v_1' + \frac{3r_{24}^{\ 2}r_{12}}{\beta_1\beta_2\beta_4} v_2''v_1^{\text{iv}} + \frac{3r_{13}^{\ 2}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''' \\
& + \frac{3r_{13}r_{24}^{\ 2}}{\beta_1\beta_2\beta_3\beta_4} v_1''v_1^{\text{iv}} + \frac{3r_{14}^{\ 2}r_{24}}{\beta_1\beta_2\beta_4} v_2^{\text{iv}}v_1' + \frac{3r_{14}r_{24}^{\ 2}}{\beta_1\beta_2\beta_4} v_2^{\text{iv}}v_1'' \\
& + \frac{3r_{23}^{\ 2}r_{24}}{\beta_2\beta_3\beta_4} v_2''v_1''' + \frac{3r_{23}r_{24}^{\ 2}}{\beta_2\beta_3\beta_4} v_2''v_1^{\text{iv}} + \frac{3r_{34}r_{12}^{\ 2}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1'' \\
& + \frac{3r_{34}^{\ 2}r_{12}}{\beta_1\beta_2\beta_3\beta_4} v_1'''v_1^{\text{iv}} + \frac{3r_{13}^{\ 2}r_{34}}{\beta_1\beta_2\beta_4} v_2'''v_1' + \frac{3r_{13}r_{34}^{\ 2}}{\beta_1\beta_3\beta_4} v_2'''v_1^{\text{iv}} \\
& + \frac{3r_{14}^{\ 2}r_{34}}{\beta_1\beta_3\beta_4} v_2^{\text{iv}}v_1' + \frac{3r_{14}r_{34}^{\ 2}}{\beta_1\beta_3\beta_4} v_2^{\text{iv}}v_1''' + \frac{3r_{23}^{\ 2}r_{34}}{\beta_2\beta_3\beta_4} v_2'''v_1''
\end{aligned}
$$

* To simplify the notation, v_s', v_s'', v_s''' v_s^{iv} have been used for ${}_1\bar{v}_s$, ${}_2\bar{v}_s$, ${}_3\bar{v}_s$, ${}_4\bar{v}_s$.

$$+ \frac{3r_{23}r_{34}^{2}}{\beta_2\beta_3\beta_4} v_2'''v_1^{\text{iv}} + \frac{3r_{24}^{2}r_{34}}{\beta_2\beta_3\beta_4} v_2^{\text{iv}}v_1'' + \frac{3r_{24}r_{34}^{2}}{\beta_2\beta_3\beta_4} v_2^{\text{iv}}v_1'''$$

$$+ \frac{6r_{12}r_{13}r_{14}}{\beta_1\beta_2\beta_3\beta_4} v_2' + \frac{6r_{12}r_{13}r_{23}}{\beta_1\beta_2\beta_3} v_1'v_1''v_1''' + \frac{6r_{12}r_{14}r_{23}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''$$

$$+ \frac{6r_{13}r_{14}r_{23}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''' + \frac{6r_{12}r_{13}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1'' + \frac{6r_{12}r_{14}r_{24}}{\beta_1\beta_2\beta_4} v_1'v_1''v_1^{\text{iv}}$$

$$+ \frac{6r_{12}r_{14}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1^{\text{iv}} + \frac{6r_{12}r_{23}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_2'' + \frac{6r_{13}r_{23}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1''v_1'''$$

$$+ \frac{6r_{14}r_{23}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1''v_1^{\text{iv}} + \frac{6r_{12}r_{13}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''' + \frac{6r_{12}r_{14}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1^{\text{iv}}$$

$$+ \frac{6r_{12}r_{23}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''' + \frac{6r_{13}r_{23}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_2''' + \frac{6r_{14}r_{23}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'''v_1^{\text{iv}}$$

$$+ \frac{6r_{12}r_{24}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1^{\text{iv}} + \frac{6r_{13}r_{24}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'''v_1^{\text{iv}} + \frac{6r_{14}r_{24}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_2^{\text{iv}}$$

$$+ \frac{6r_{23}r_{24}r_{34}}{\beta_2\beta_3\beta_4} v_1''v_1'''v_1^{\text{iv}} + \frac{6r_{13}r_{14}r_{34}}{\beta_1\beta_3\beta_4} v_1\, v_1'''v_1^{\text{iv}} \Bigg\}$$

$$+\frac{1}{\underline{|4}}\Bigg\{\frac{r_{12}^{4}}{\beta_1\beta_2} v_3'v_3'' + \frac{r_{13}^{4}}{\beta_1\beta_3} v_3'v_3''' + \frac{r_{14}^{4}}{\beta_1\beta_4} v_3'v_3^{\text{iv}} + \frac{r_{23}^{4}}{\beta_2\beta_3} v_3''v_3''' + \frac{r_{24}^{4}}{\beta_2\beta_4} v_3''v_3^{\text{iv}} + \frac{r_{34}^{4}}{\beta_3\beta_4} v_3'''v_3^{\text{iv}}$$

$$+ \frac{4r_{12}^{3}r_{13}}{\beta_1\beta_2\beta_3} v_3\, v_2'' + \frac{4r_{12}r_{13}^{3}}{\beta_1\beta_2\beta_3} v_3'v_2''' + \frac{4r_{12}^{3}r_{14}}{\beta_1\beta_2\beta_4} v_3'v_2'' + \frac{4r_{12}r_{14}^{3}}{\beta_1\beta_2\beta_4} v_3'v_2^{\text{iv}}$$

$$+ \frac{4r_{13}^{3}r_{14}}{\beta_1\beta_3\beta_4} v_3'v_3''' + \frac{4r_{13}r_{14}^{3}}{\beta_1\beta_3\beta_4} v_3'v_2^{\text{iv}} + \frac{4r_{12}^{3}r_{23}}{\beta_1\beta_2\beta_3} v_3''v_2' + \frac{4r_{12}r_{23}^{3}}{\beta_1\beta_2\beta_3} v_3''v_2'''$$

$$+ \frac{4r_{13}^{3}r_{23}}{\beta_1\beta_2\beta_3} v_3'''v_2' + \frac{4r_{13}r_{23}^{3}}{\beta_1\beta_2\beta_3} v_3'''v_2'' + \frac{4r_{14}^{3}r_{23}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_2^{\text{iv}} + \frac{4r_{14}r_{23}^{3}}{\beta_1\beta_2\beta_3\beta_4} v_2''v_2'''$$

$$+ \frac{4r_{12}^{3}r_{24}}{\beta_1\beta_2\beta_4} v_2'v_3'' + \frac{4r_{12}r_{24}^{3}}{\beta_1\beta_2\beta_4} v_3''v_2^{\text{iv}} + \frac{4r_{13}^{3}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_2''' + \frac{4r_{13}r_{24}^{3}}{\beta_1\beta_2\beta_3\beta_4} v_2'\, v_2^{\text{iv}}$$

$$+ \frac{4r_{14}^{3}r_{24}}{\beta_1\beta_2\beta_4} v_2'v_3^{\text{iv}} + \frac{4r_{14}r_{24}^{3}}{\beta_1\beta_2\beta_4} v_2''v_3^{\text{iv}} + \frac{4r_{23}^{3}r_{24}}{\beta_2\beta_3\beta_4} v_3''v_2''' + \frac{4r_{23}r_{24}^{3}}{\beta_2\beta_3\beta_4} v_3''v_2^{\text{iv}}$$

$$+ \frac{4r_{12}^{3}r_{34}}{\beta_1\beta_3\beta_4} v_2'v_2'' + \frac{4r_{12}r_{34}^{3}}{\beta_1\beta_2\beta_3\beta_4} v_2'''v_2^{\text{iv}} + \frac{4r_{13}^{3}r_{34}}{\beta_1\beta_3\beta_4} v_3'''v_2' + \frac{4r_{13}r_{34}^{3}}{\beta_1\beta_3\beta_4} v_3'''v_2^{\text{iv}}$$

$$+ \frac{4r_{14}^{3}r_{34}}{\beta_1\beta_3\beta_4} v_2'v_3^{\text{iv}} + \frac{4r_{14}r_{34}^{3}}{\beta_1\beta_3\beta_4} v_2'''v_3^{\text{iv}} + \frac{4r_{23}^{3}r_{34}}{\beta_2\beta_3\beta_4} v_2''v_3'' + \frac{4r_{23}r_{34}^{3}}{\beta_2\beta_3\beta_4} v_3'''v_2^{\text{iv}}$$

$$+ \frac{4r_{24}^{3}r_{34}}{\beta_1\beta_3\beta_4} v_2''v_3^{\text{iv}} + \frac{4r_{24}r_{34}^{3}}{\beta_2\beta_3\beta_4} v_2'''v_3^{\text{iv}} + \frac{6r_{12}^{2}r_{13}^{2}}{\beta_1\beta_2\beta_3} v_3'v_1''v_1''' + \frac{6r_{12}^{2}r_{14}^{2}}{\beta_1\beta_2\beta_4} v_3'v_1''v_1^{\text{iv}}$$

$$+ \frac{6r_{13}^{2}r_{14}^{2}}{\beta_1\beta_3\beta_4} v_3'v_1'''v_1^{\text{iv}} + \frac{6r_{12}^{2}r_{23}^{2}}{\beta_1\beta_2\beta_3} v_1'v_3''v_1''' + \frac{6r_{13}^{2}r_{23}^{2}}{\beta_1\beta_2\beta_3} v_1'v_1''v_3''' + \frac{6r_{14}^{2}r_{23}^{2}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''v_1'''v_1^{\text{iv}}$$

$$+ \frac{6r_{12}^{2}r_{24}^{2}}{\beta_1\beta_2\beta_4} v_1'v_3''v_1^{\text{iv}} + \frac{6r_{13}^{2}r_{24}^{2}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''v_1'''v_1^{\text{iv}} + \frac{6r_{14}^{2}r_{24}^{2}}{\beta_1\beta_2\beta_4} v_1'v_1''v_3^{\text{iv}} + \frac{6r_{23}^{2}r_{24}^{2}}{\beta_2\beta_3\beta_4} v_3'\, v_1'''v_1^{\text{iv}}$$

$$+ \frac{6r_{12}^2 r_{34}^2}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''v_1'''v_1^{\mathrm{iv}} + \frac{6r_{13}^2 r_{34}^2}{\beta_1\beta_3\beta_4} v_1'v_3'''v_1^{\mathrm{iv}} + \frac{6r_{14}^2 r_{34}^2}{\beta_1\beta_3\beta_4} v_1'v_1'''v_3^{\mathrm{iv}} + \frac{6r_{23}^3 r_{34}^2}{\beta_2\beta_3\beta_4} v_1''v_3'''v_1^{\mathrm{iv}}$$

$$+ \frac{6r_{24}^2 r_{34}^2}{\beta_2\beta_3\beta_4} v_1''v_1'''v_3^{\mathrm{iv}} + \frac{12r_{12}^2 r_{13}r_{14}}{\beta_1\beta_2\beta_3\beta_4} v_3'v_1'' + \frac{12r_{12}r_{13}^2 r_{14}}{\beta_1\beta_2\beta_3\beta_4} v_3'v_1''' + \frac{12r_{12}r_{13}r_{14}^2}{\beta_1\beta_2\beta_3\beta_4} v_3'v_1^{\mathrm{iv}}$$

$$+ \frac{12r_{12}^2 r_{13}r_{23}}{\beta_1\beta_2\beta_3} v_2'v_2''v_1''' + \frac{12r_{12}r_{13}^2 r_{23}}{\beta_1\beta_2\beta_3} v_2'v_1''v_2''' + \frac{12r_{12}r_{13}r_{23}^2}{\beta_1\beta_2\beta_3} v_1'v_2''v_2'''$$

$$+ \frac{12r_{12}^2 r_{14}r_{23}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_2'' + \frac{12r_{12}r_{14}^2 r_{23}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_1''v_1^{\mathrm{iv}} + \frac{12r_{12}r_{14}r_{23}^2}{\beta^1\beta_2\beta_3\beta_4} v_1'v_2''v_1'$$

$$+ \frac{12r_{13}^2 r_{14}r_{23}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_2'' + \frac{12r_{13}r_{14}^2 r_{23}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_1'''v^{\mathrm{iv}} + \frac{12r_{13}r_{14}r_{23}^2}{\beta_1\beta_2\beta_3\beta_4} v_1'v_2'''v_1''$$

$$+ \frac{12r_{12}^2 r_{13}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_2'' + \frac{12r_{12}r_{13}^2 r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_1''v_1''' + \frac{12r_{12}r_{13}r_{24}^2}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''v_1^{\mathrm{iv}}$$

$$+ \frac{12r_{12}^2 r_{14}r_{24}}{\beta_1\beta_2\beta_4} v_2'v_2''v_1^{\mathrm{iv}} + \frac{12r_{12}r_{14}^2 r_{24}}{\beta_1\beta_2\beta_4} v_2'v_1''v_2^{\mathrm{iv}} + \frac{12r_{12}r_{14}r_{24}^2}{\beta_1\beta_2\beta_4} v_1'v_2''v_2^{\mathrm{iv}}$$

$$+ \frac{12r_{13}^2 r_{14}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_1'''v_1^{\mathrm{iv}} + \frac{12r_{13}r_{14}^2 r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_2^{\mathrm{iv}} + \frac{12r_{13}r_{14}r_{24}^2}{\beta_1\beta_2\beta_3\beta_4} v_1'v_2^{\mathrm{iv}}v_1''$$

$$+ \frac{12r_{12}^2 r_{23}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_3'' + \frac{12r_{12}r_{23}^2 r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_3''v_1''' + \frac{12r_{12}r_{23}r_{24}^2}{\beta_1\beta_2\beta_3\beta_4} v_3''v_1^{\mathrm{iv}}$$

$$+ \frac{12r_{13}^2 r_{23}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''v_2''' + \frac{12r_{13}r_{23}^2 r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_2''v_2''' + \frac{12r_{13}r_{23}r_{24}^2}{\beta_1\beta_2\beta_3\beta_4} v_2''v_1'''v_1^{\mathrm{iv}}$$

$$+ \frac{12r_{14}^2 r_{23}r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1''v_2^{\mathrm{iv}} + \frac{12r_{14}r_{23}^2 r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_2''v_1'''v_1^{\mathrm{iv}} + \frac{12r_{14}r_{23}r_{24}^2}{\beta_1\beta_2\beta_3\beta_4} v_2''v_2^{\mathrm{iv}}$$

$$+ \frac{12r_{12}^2 r_{13}r_{34}}{\beta_1\beta_2\beta_3\beta_3} v_2'v_1''v_2''' + \frac{12r_{12}r_{13}^2 r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_2''' + \frac{12r_{12}r_{13}r_{34}^2}{\beta_1\beta_2\beta_3\beta_4} v_1'v_2'''v_1^{\mathrm{iv}}$$

$$+ \frac{12r_{12}^2 r_{14}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_1''v_1^{\mathrm{iv}} + \frac{12r_{12}r_{14}^2 r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_2'v_2^{\mathrm{iv}} + \frac{12r_{12}r_{14}r_{34}^2}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1'''v_2^{\mathrm{iv}}$$

$$+ \frac{12r_{13}^2 r_{14}r_{34}}{\beta_1\beta_3\beta_4} v_2'v_2'''v_1^{\mathrm{iv}} + \frac{12r_{13}r_{14}^2 r_{34}}{\beta_1\beta_3\beta_4} v_2'v_1'''v_2^{\mathrm{iv}} + \frac{12r_{13}r_{14}r_{34}^2}{\beta_1\beta_3\beta_4} v_1'v_2'''v_2^{\mathrm{iv}}$$

$$+ \frac{12r_{12}^2 r_{23}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_2''v_1''' + \frac{12r_{12}r_{23}^2 r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_2''v_2''' + \frac{12r_{12}r_{23}r_{34}^2}{\beta_1\beta_2\beta_3\beta_4} v_1''v_2'''v_1^{\mathrm{iv}}$$

$$+ \frac{12r_{13}^2 r_{23}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_3^{\mathrm{iv}} + \frac{12r_{13}r_{23}^2 r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1''v_3''' + \frac{12r_{13}r_{23}r_{34}^2}{\beta_1\beta_2\beta_3\beta_4} v_3'''v_1^{\mathrm{iv}}$$

$$+ \frac{12r_{14}^2 r_{23}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_1'''v_2^{\mathrm{iv}} + \frac{12r_{14}r_{23}^2 r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1''v_2'''v_1^{\mathrm{iv}} + \frac{12r_{14}r_{23}r_{34}^2}{\beta_1\beta_2\beta_3\beta_4} v_2'''v_2^{\mathrm{iv}}$$

$$+ \frac{12r_{12}^2 r_{24}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_2''v_1^{\mathrm{iv}} + \frac{12r_{12}r_{24}^2 r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_2''v_2^{\mathrm{iv}} + \frac{12r_{12}r_{24}r_{34}^2}{\beta_1\beta_2\beta_3\beta_4} v_1''v_1'''v_2^{\mathrm{iv}}$$

$$+ \frac{12r_{13}^2 r_{24}r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'v_2'''v_1^{\mathrm{iv}} + \frac{12r_{13}r_{24}^2 r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1''v_1'''v_2^{\mathrm{iv}} + \frac{12r_{13}r_{24}r_{34}^2}{\beta_1\beta_2\beta_3\beta_4} v_2'''v_2^{\mathrm{iv}}$$

$$
\begin{aligned}
&+ \frac{12r_{14}^2 r_{24} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1' v_3^{\text{iv}} + \frac{12r_{14} r_{24}^2 r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'' v_3^{\text{iv}} + \frac{12r_{14} r_{24} r_{34}^2}{\beta_1\beta_2\beta_3\beta_4} v_1''' v_3^{\text{iv}} \\
&+ \frac{12r_{23}^2 r_{24} r_{34}}{\beta_2\beta_3\beta_4} v_2'' v_2''' v_1^{\text{iv}} + \frac{12r_{23} r_{24}^2 r_{34}}{\beta_2\beta_3\beta_4} v_2'' v_1''' v_2^{\text{iv}} + \frac{12r_{23} r_{24} r_{34}^2}{\beta_2\beta_3\beta_4} v_1'' v_2''' v_2^{\text{iv}} \\
&+ \frac{24r_{12} r_{13} r_{14} r_{23}}{\beta_1\beta_2\beta_3\beta_4} v_2' v_1'' v_1''' + \frac{24r_{12} r_{13} r_{14} r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_2' v_1'' v_1^{\text{iv}} + \frac{24r_{12} r_{13} r_{23} r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1' v_2'' v_1''' \\
&+ \frac{24r_{12} r_{14} r_{23} r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1' v_2'' v_1^{\text{iv}} + \frac{24r_{13} r_{14} r_{23} r_{24}}{\beta_1\beta_2\beta_3\beta_4} v_1' v_1'' v_1''' v_1^{\text{iv}} + \frac{24r_{12} r_{13} r_{14} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_2' v_1''' v_1^{\text{iv}} \\
&+ \frac{24r_{12} r_{13} r_{23} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1' v_1'' v_2''' + \frac{24r_{12} r_{14} r_{23} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1' v_1'' v_1''' v_1^{\text{iv}} + \frac{24r_{13} r_{14} r_{23} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1' v_2''' v_1^{\text{iv}} \\
&+ \frac{24r_{12} r_{13} r_{14} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_2' v_1''' v_1^{\text{iv}} + \frac{24r_{12} r_{14} r_{24} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1' v_1'' v_2^{\text{iv}} + \frac{24r_{13} r_{14} r_{24} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1' v_1''' v_2^{\text{iv}} \\
&+ \frac{24r_{12} r_{23} r_{24} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_2'' v_1''' v_1^{\text{iv}} + \frac{24r_{13} r_{23} r_{24} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'' v_2''' v_1^{\text{iv}} + \frac{24r_{14} r_{23} r_{24} r_{34}}{\beta_1\beta_2\beta_3\beta_4} v_1'' v_1''' v_2^{\text{iv}} \Bigg\} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{. . . . . (lxxxiii.).}
\end{aligned}
$$

In the case of three variables, we must cancel in the above expression all terms involving β_4. Thus we shall have 3 instead of 6 first order terms, 6 instead of 21 second order terms, 10 instead of 56 third order terms, and 15 instead of 126 fourth order terms—a much more manageable series.

I give below the extra term necessary for calculating the value of $(Q - Q_0)/Q_0$ as far as the fifth order terms in the case of three variables.

Fifth Order Terms for Three Variables.

$$
\begin{aligned}
\frac{1}{|\underline{5}} \Bigg\{ & \frac{r_{23}^5}{\beta_2\beta_3} v_4'' v_4''' + \frac{r_{31}^5}{\beta_3\beta_1} v_4''' v_4' + \frac{r_{12}^5}{\beta_1\beta_2} v_4' v_4'' + \frac{5r_{31}^4 r_{12}}{\beta_1\beta_2\beta_3} v_3''' v_4' \\
&+ \frac{5r_{31}^4 r_{23}}{\beta_1\beta_2\beta_3} v_3' v_4''' + \frac{5r_{12}^4 r_{23}}{\beta_1\beta_2\beta_3} v_3' v_4'' + \frac{5r_{12}^4 r_{31}}{\beta_1\beta_2\beta_3} v_4' v_3'' \\
&+ \frac{5r_{23}^4 r_{31}}{\beta_1\beta_2\beta_3} v_3'' v_4''' + \frac{5r_{23}^4 r_{12}}{\beta_1\beta_2\beta_3} v_4'' v_3''' + \frac{10r_{31}^3 r_{12}^2}{\beta_1\beta_2\beta_3} v_4' v_1'' v_2''' \\
&+ \frac{10r_{31}^3 r_{23}^2}{\beta_1\beta_2\beta_3} v_2' v_1' v_4''' + \frac{10r_{12}^3 r_{23}^2}{\beta_1\beta_3\beta_2} v_2' v_4'' v_1''' + \frac{10r_{12}^3 r_{31}^2}{\beta_1\beta_2\beta_3} v_4' v_2'' v_1''' \\
&+ \frac{10r_{23}^3 r_{31}^2}{\beta_1\beta_2\beta_3} v_1' v_2'' v_4''' + \frac{10r_{23}^3 r_{12}^2}{\beta_1\beta_2\beta_3} v_1' v_4'' v_2''' + \frac{20r_{23}^3 r_{31} r_{12}}{\beta_1\beta_2\beta_3} v_1' v_3'' v_3''' \\
&+ \frac{20r_{31}^3 r_{12} r_{23}}{\beta_1\beta_2\beta_3} v_3' v_1'' v_3''' + \frac{20r_{12}^3 r_{23} r_{31}}{\beta_1\beta_2\beta_3} v_3' v_3'' v_1''' + \frac{30r_{23} r_{31}^2 r_{12}^2}{\beta_1\beta_2\beta_3} v_3' v_3'' v_2''' \\
&+ \frac{30r_{31} r_{12}^2 r_{23}^2}{\beta_1\beta_2\beta_3} v_2' v_3'' v_2''' + \frac{30r_{12} r_{23}^2 r_{31}^2}{\beta_1\beta_2\beta_3} v_2' v_2'' v_3''' \Bigg\} . \quad \text{. . . . . . (lxxxiv.).}
\end{aligned}
$$

A numerical illustration of these formulæ will be given in the latter part of this Memoir. It will, however, be clear that what we want are tables of $\log\left(\frac{{}_s\bar{v}_p}{\beta_s}\right)$, including $\log\left(\frac{{}_s\bar{v}_0}{\beta_s}\right)$ or $\log\left(\frac{1}{\beta_s}\right)$ for a series of values of h. Such tables would render the computation of $\frac{Q - Q_0}{Q_0}$ fairly direct and rapid; they could be fairly easily calculated from existing tables for the ordinate and area of the normal curve, and I hope later to find some one willing to undertake them.

Meanwhile let us look at special cases. In the first place, suppose, in the case of three variables, that the division of the groups is taken at the mean, *i.e.*, $h_1 = h_2 = h_3 = 0$. Then we have

$$\beta_1 = \beta_2 = \beta_3 = \int_0^\infty e^{-\frac{1}{2}x^2}\,dx = \sqrt{\frac{\pi}{2}}.$$

$$v_1{}' = v_1{}'' = v_1{}''' = 0$$

$$v_2{}' = v_2{}'' = v_2{}''' = -1$$

$$v_3{}' = v_3{}'' = v_3{}''' = 0$$

$$v_4{}' = v_4{}'' = v_4{}''' = 3.$$

Hence we have

$$Q = \int_0^\infty\int_0^\infty\int_0^\infty z\,dx_1\,dx_2\,dx_3 = Q_0\left\{1 + \frac{2}{\pi}(r_{12} + r_{13} + r_{23})\right\}$$
$$+ \frac{1}{\underline{|3}}\left\{\frac{2}{\pi}(r_{12}{}^3 + r_{13}{}^3 + r_{23}{}^3)\right\} + \frac{1}{\underline{|5}}\left\{\frac{2}{\pi}9(r_{23}{}^5 + r_{31}{}^5 + r_{12}{}^5)\right\} + \ldots$$
$$= Q_0\left\{1 + \frac{2}{\pi}(\sin^{-1} r_{12} + \sin^{-1} r_{13} + \sin^{-1} r_{23})\right\} \quad \text{. . . . . . . (lxxxv.).}$$

Let $r_{12} = \cos D_{12}$, $r_{13} = \cos D_{13}$, $r_{23} = \cos D_{23}$, and let E be the spherical excess of the spherical triangle whose angles are the divergences D_{12}, D_{13}, D_{23}. Then we have

$$\frac{Q - Q_0}{Q_0}\frac{\pi}{2} = \frac{3\pi}{2} - D_{12} - D_{13} - D_{23} = \frac{\pi}{2} - E.$$

Or:

$$\sin\frac{Q - Q_0}{Q_0}\frac{\pi}{2} = \cos E \quad \text{. . . . . (lxxxvi.).}$$

Now take the case of four variables. Here we have

$$\beta_1 = \beta_2 = \beta_3 = \beta_4 = \sqrt{\frac{\pi}{2}}$$

$$v_2{}' = v_2{}'' = v_2{}''' = v_2{}^{\text{iv}} = 1$$

$$v_4{}' = v_4{}'' = v_4{}''' = v_4{}^{\text{iv}} = 3,$$

and all the odd v's zero. Hence

$$\begin{aligned}\frac{\mathrm{Q}-\mathrm{Q}_0}{\mathrm{Q}} = {} & \frac{2}{\pi}(r_{12}+r_{13}+r_{14}+r_{23}+r_{24}+r_{34}) + \left(\frac{2}{\pi}\right)^2(r_{14}r_{23}+r_{12}r_{34}+r_{13}r_{24}) \\ & + \frac{2}{\pi}\frac{1}{\underline{|3}}(r_{12}^3+r_{13}^3+r_{14}^3+r_{23}^3+r_{24}^3+r_{34}^3) + \left(\frac{2}{\pi}\right)^2(r_{12}r_{13}r_{14}+r_{12}r_{23}r_{24} \\ & + r_{13}r_{23}r_{34}+r_{14}r_{24}r_{34}) + \left(\frac{2}{\pi}\right)^2\frac{1}{\underline{|3}}(r_{14}^3r_{23}+r_{14}r_{23}^3+r_{13}r_{24}^3+r_{13}^3r_{24}+r_{12}^3r_{34} \\ & + r_{12}r_{34}^3) + \left(\frac{2}{\pi}\right)^2\frac{1}{\underline{|2}}(r_{12}^2r_{14}r_{23}+r_{13}^2r_{14}r_{23}+r_{12}^2r_{13}r_{24}+r_{13}r_{14}^2r_{24}+r_{13}r_{23}^2r_{24} \\ & + r_{14}r_{23}r_{24}^2+r_{12}r_{13}^2r_{34}+r_{12}r_{14}^2r_{34}+r_{12}r_{23}^2r_{34}+r_{14}r_{23}r_{34}^2+r_{12}r_{24}^2r_{34} \\ & + r_{13}r_{24}r_{34}^2) + \ldots\ldots\ldots \qquad \text{(lxxxvii.)}.\end{aligned}$$

This is the correct value including terms of the fourth order, but to this order of approximation we can throw it into a much simpler form. Let $r_{ss'} = \sin\delta_{ss'}$, then

$$\begin{aligned}\frac{\mathrm{Q}-\mathrm{Q}_0}{\mathrm{Q}_0}\frac{\pi}{2} = {} & \sin^{-1}r_{12}+\sin^{-1}r_{13}+\sin^{-1}r_{14}+\sin^{-1}r_{23}+\sin^{-1}r_{24}+\sin^{-1}r_{34} \\ & + \frac{2}{\pi}(\sin^{-1}r_{12}\sin^{-1}r_{13}\sin^{-1}r_{14}+\sin^{-1}r_{12}\sin^{-1}r_{23}\sin^{-1}r_{24} \\ & + \sin^{-1}r_{13}\sin^{-1}r_{23}\sin^{-1}r_{34}+\sin^{-1}r_{14}\sin^{-1}r_{24}\sin^{-1}r_{34}) \\ & + \frac{2}{\pi}[\sin^{-1}r_{14}\sin^{-1}r_{23}\{(1-r_{12}^2)(1-r_{13}^2)(1-r_{24}^2)(1-r_{34}^2)\}^{-\frac{1}{2}} \\ & \quad + \sin^{-1}r_{12}\sin^{-1}r_{34}\{(1-r_{13}^2)(1-r_{14}^2)(1-r_{23}^2)(1-r_{24}^2)\}^{-\frac{1}{2}} \\ & \quad + \sin^{-1}r_{13}\sin^{-1}r_{24}\{(1-r_{12}^2)(1-r_{14}^2)(1-r_{23}^2)(1-r_{34}^2)\}^{-\frac{1}{2}}] \\ = {} & \delta_{12}+\delta_{13}+\delta_{14}+\delta_{23}+\delta_{24}+\delta_{34} \\ & + \frac{2}{\pi}(\delta_{12}\delta_{13}\delta_{14}+\delta_{12}\delta_{23}\delta_{24}+\delta_{13}\delta_{23}\delta_{34}+\delta_{14}\delta_{24}\delta_{34}) \\ & + \frac{2}{\pi}\left(\frac{\delta_{14}\delta_{23}\cos\delta_{14}\cos\delta_{23}+\delta_{12}\delta_{34}\cos\delta_{12}\cos\delta_{34}+\delta_{13}\delta_{24}\cos\delta_{13}\cos\delta_{24}}{\cos\delta_{12}\cos\delta_{13}\cos\delta_{14}\cos\delta_{23}\cos\delta_{24}\cos\delta_{34}}\right) \\ & \qquad \ldots\ldots \quad \text{(lxxxviii.)}.\end{aligned}$$

We may write this
$$\sin\frac{\mathrm{Q}-\mathrm{Q}_0}{\mathrm{Q}_0}\frac{\pi}{2} = \cos\mathrm{E}' \quad \ldots\ldots\ldots \quad \text{(lxxxix.)}$$
where

$$\begin{aligned}\mathrm{E}' = {} & \frac{\pi}{2}-\delta_{12}-\delta_{13}-\delta_{14}-\delta_{23}-\delta_{24}-\delta_{34} \\ & - \frac{2}{\pi}(\delta_{12}\delta_{13}\delta_{14}+\delta_{12}\delta_{23}\delta_{24}+\delta_{13}\delta_{23}\delta_{24}+\delta_{14}\delta_{24}\delta_{34}) \\ & - \frac{2}{\pi}\left(\frac{\delta_{14}\delta_{23}\cos\delta_{14}\cos\delta_{23}+\delta_{12}\delta_{34}\cos\delta_{12}\cos\delta_{34}+\delta_{13}\delta_{24}\cos\delta_{13}\cos\delta_{24}}{\cos\delta_{12}\cos\delta_{13}\cos\delta_{14}\cos\delta_{23}\cos\delta_{24}\cos\delta_{34}}\right)\end{aligned}$$

The expressions E and E′ of (lxxxvi.) and (lxxxix.) are of considerable interest, for they enable us to express the area of a spherical triangle in three-dimensioned space, and (up to the above degree of approximation) the volume of a "tetrahedron" on a "sphere" in hyperspace of four dimensions. In fact, the whole theory of hyperspace "spherical trigonometry" needs investigation in relation to the properties of multiple correlation.

In our illustrations (viii.) and (ix.) will be found examples of the above formulæ applied to important cases in triple and quadruple correlation in the theory of heredity. I consider that the formulæ above given will cover numerous novel applications, for many of which greater simplicity will be introduced owing to the choice of special values for the h's or for the correlation coefficients.

(8.) *Illustrations of the New Methods.*

Illustration I. Inheritance of Coat-colour in Horses.—The following represents the distribution of sires and fillies in 1050 cases of thoroughbred racehorses, the grouping being made into all coat-colour classed as "bay and darker," "chesnut and lighter":—

	Colour.	Sires.		
		Bay and darker.	Chesnut and lighter.	
Fillies.	Bay and darker . . .	631	125	756
	Chesnut and lighter .	147	147	294
		778	272	1050

a	b	$a + b$
c	d	$c + d$
$a + c$	$b + d$	N

Then we require the correlation between sire and filly in the matter of coat-colour, and also the probable error of its determination.

We have from (iv.) and (v.)

$$\alpha_1 = \frac{(a + c) - (b + d)}{\text{N}} = \sqrt{\frac{2}{\pi}} \int_0^h e^{-\frac{1}{2}x^2} dx = \cdot 481,905,$$

$$\alpha_2 = \frac{(a + b) - (c + d)}{\text{N}} = \sqrt{\frac{2}{\pi}} \int_0^k e^{-\frac{1}{2}y^2} dy = \cdot 440,000.$$

Hence from the probability integral tables

$$h = \cdot 64630, \qquad k = \cdot 58284.$$

We have then: log HK $= 1 \cdot 037,3514$ by (xvii.), Thence $\epsilon = \dfrac{ad - bc}{\mathrm{N}^2\mathrm{HK}} = \cdot 619{,}068$ from (xxi.).

Calculating out the coefficients of the series in r in (xix.) we find

$$\cdot 619{,}068 = r + \cdot 188{,}345 r^2 + \cdot 064{,}0814 r^3 + \cdot 107{,}8220 r^4 + \cdot 005{,}9986 r^5 + \cdot 067{,}2682 r^6 + \&c.$$

Neglecting powers of r above the second, we find by solving the quadratic and taking the positive root

$$r = \cdot 5600.$$

Solving by two approximations the sextic we finally determine

$$r = \cdot 5422,$$

correct, I think, to four places of figures.

Turning now to the probable error as given by Equation (l.), I find

$$h^2 + k^2 - 2rhk = \cdot 348{,}924,$$

and from (xlix.)

$$\log \chi_0 = \bar{1}\cdot 170{,}0947.$$

Further: $\dfrac{k - rh}{\sqrt{1 - r^2}} = \cdot 275{,}642$, $\dfrac{h - rk}{\sqrt{1 - r^2}} = \cdot 393{,}078$.

Hence from (xlvii.) and (xlviii.) we find

$$\psi_1 = \frac{1}{\sqrt{2\pi}} \int_0^{\cdot 393{,}078} e^{-\frac{1}{2}z^2 dz}, \qquad \psi_2 = \frac{1}{\sqrt{2\pi}} \int_0^{\cdot 275{,}642} e^{-\frac{1}{2}z^2 dz},$$

and by means of the probability integral table

$$\psi_1 = \cdot 108{,}884, \qquad \psi_2 = \cdot 152{,}865.$$

By substituting in (l.), we find

$$\text{probable error of } r = \cdot 0288.$$

From (xxxiv.) and (xxxv.) we find

$$\text{p.e. of } h = \cdot 0282. \qquad \text{p.e. of } k = \cdot 0278.$$

Thus, finally, we may sum up our results

$$h = \cdot 6463 \pm \cdot 0282, \qquad k = \cdot 5828 \pm \cdot 0278,$$

$$r = \cdot 5422 \pm \cdot 0288.$$

The probable error of this r, if we had been able to find it from the product moment, would have been ·0147, or only about one-half its present value.

Illustration II.—Our analysis opens a large field suggested by the following problem:—*What is the chance that an exceptional man is born of an exceptional father?*

Of course much depends on how we define "exceptional," and any numerical measure of it must be quite arbitrary. As an illustration, let us take a man who possesses a character only possessed by one man in twenty as exceptional. For example, only one man in twenty is more than 6 feet 1·2 inches in height, and such a stature may be considered "exceptional." In a class of twenty students we generally find one of "exceptional" ability, and so on. Accordingly we have classed fathers and sons who possess characters only possessed by one man in twenty as exceptional. We first determine h and k, so that the tail of the frequency curve cut off is $\frac{1}{20}$ of its whole area. This gives us $h = k = 1·64485$.

Next we determine $\mathrm{HK} = \frac{1}{2\pi} e^{-\frac{1}{2}(h^2+k^2)}$, and find $\log \mathrm{HK} = \bar{2}·026{,}8228$.

Then we calculate the coefficients of the various powers of r in (xix.). We find

$$\log \tfrac{1}{2}hk = ·131{,}2225.$$

$$\log \tfrac{1}{6}(h^2-1)(k^2-1) = \bar{1}·685{,}5683.$$

$$\log \frac{hk}{24}(h^2-3)(k^2-3) = \bar{3}·990{,}1176.$$

$$\log \tfrac{1}{120}(h^4-6h^2+3)(k^4-6k^2+3) = \bar{1}·464{,}4772.$$

$$\log \frac{hk}{720}(h^4-10h^2+15)(k^4-10k^2+15) = \bar{2}·925{,}6367.$$

It remains to determine what value we shall give to r, the paternal correlation. It ranges from ·3 to ·5 for my own measurements as we turn from blended to exclusive inheritance. Taking these two extreme values we find

$$\frac{ad-bc}{\mathrm{N}^2} = ·0046344 \quad \text{or} \quad ·0096779.$$

But $\frac{ad-bc}{\mathrm{N}^2} = \frac{d}{\mathrm{N}} - \frac{(d+b)(d+c)}{\mathrm{N}^2}$, and the second term is the chance of exceptional fathers with exceptional sons, when variation is independent, *i.e.*, when there is no heredity, $= \frac{1}{20} \times \frac{1}{20} = ·0025$.

Thus $d/\mathrm{N} = ·007134$ or $·012178$;

accordingly $b/\mathrm{N} = ·042866$ or $·037822$.

Hence we conclude that of the 5 per cent. of exceptional men ·71 per cent. in the first case, and 1·22 per cent. in the second case, are born of exceptional fathers, and 4·29 per cent. in the first case and 3·78 per cent. in the second case of non-exceptional fathers. In other words, out of 1000 men of mark we may expect 142 in the first case, 244 in the second, to be born of exceptional parents, while 858 in the first and 756 in the second are born of undistinguished fathers. In the former case the odds are about 6 to 1, in the latter 3 to 1 against a distinguished son having a distinguished father. This result confirms what I have elsewhere stated, that we trust to the great mass of our population for the bulk of our distinguished men. On the other hand it does not invalidate what I have written on the importance of creating good stock, for a good stock means a bias largely above that due to an exceptional father alone.

In addition to this the $\frac{1}{20}$ of the population forming the exceptional fathers produce 142 or 244 exceptional sons to compare with the 858 or 756 exceptional sons produced by the $\frac{19}{20}$ of the population who are non-exceptional. That is to say that the *relative* production is as 142 to 45·2, or 244 to 39·8, *i.e.*, in the one case as more than 3 to 1, in the other case as more than 6 to 1. *In other words, exceptional fathers produce exceptional sons at a rate* 3 *to* 6 *times as great as non-exceptional fathers.* It is only because exceptional fathers are themselves so rare that we must trust for the bulk of our distinguished men to the non-exceptional class.

Illustration III. Heredity in Coat-colour of Hounds.—To find the correlation in coat-colour between Basset hounds which are half-brethren, say, offspring of the same dam.

Here the classification is simply into lemon and white (*lw*) and lemon, black and white or tricolour (*t*),

The following is the table for 4172 cases :—

Colour.	*t.*	*lw.*	Totals.
t.	1766	842	2608
lw.	842	722	1564
Totals	2608	1564	4172

Proceeding precisely in the same way as in the first illustration we find :

$$\alpha_1 = \alpha_2 = \cdot 25024$$
$$h = k = \cdot 318,957$$
$$\log \mathrm{KH} = \bar{1}\cdot 157,6378$$
$$\epsilon = \cdot 226,234.$$

It will be sufficient now to go to r^4. We have

$$\cdot 226,234 = r + \cdot 050,867\, r^2 + \cdot 134,480\, r^3 + \cdot 035,587\, r^4.$$

The quadratic gives $r = \cdot 2237$. Using the NEWTONIAN method of approximating to the root we find

$$r = \cdot 2222.$$

Summing up as before, after finding the probable errors, we have

$$h = k = \cdot 3190 \pm \cdot 0133,$$

$$r = \cdot 2222 \pm \cdot 0162.$$

Illustration IV. Inheritance of Eye-colour in Man.—To find the correlation in eye-colour between a maternal grandmother and her granddaughter. Here the classification is into eyes described as grey or lighter, and eyes described as dark grey or darker.*

	Tint.	Maternal grandmother.		Totals.
		Grey or lighter.	Dark grey or darker.	
Granddaughter.	Grey or lighter	254	136	390
	Dark grey or darker . .	156	193	349
	Totals	410	329	739

As before, we find

$$\alpha_1 = \cdot 109{,}607, \qquad \alpha_2 = \cdot 055{,}480,$$

$$h = \cdot 138{,}105, \qquad k = \cdot 069{,}593,$$

$$\log \mathrm{HK} = \bar{1}\cdot 196{,}6267,$$

$$\epsilon = \cdot 323{,}760.$$

Series for r up to r^4

$$\cdot 323{,}760 = r + \cdot 004{,}806 r^2 + \cdot 162\cdot 696 r^3 + \cdot 000{,}358 r^4.$$

The quadratic gives $r = \cdot 3233$, and the biquadratic

$$r = \cdot 3180,$$

the value of the term in r^4 being $\cdot 000{,}00366$, so that higher terms may be neglected.

Determining the probable errors as in Illustration I., we sum up :—

---

* According to Mr. GALTON'S classification, the first group contains eyes described as light blue, blue, dark blue, blue-green, grey; and the second eyes described as dark grey, hazel, light brown, brown, dark brown, verv dark brown, black.

$$h = \cdot1381 \pm \cdot0312,$$
$$k = \cdot0696 \pm \cdot0311,$$
$$r = \cdot3180 \pm \cdot0361.$$

Illustration V. Inheritance of Stature.—The following data have been found for the inheritance of stature between father and son from my Family Data cards, 1078 cases :—

Mean stature of father	67″·698
,, son	68″·661
Standard deviation of father . .	2″·7048
,, son . . .	2″·7321

Correlation $= \cdot5198 \pm \cdot0150$.

Now for purposes of comparison of methods the correlation has been determined for this material from various groupings of fathers and sons :—

(A.) *Fathers.*

	Class.	Below 67″·5.	Above 67″·5.	Totals.
Sons.	Below 67″·5 . .	269·25	95·75	365
	Above 67″·5 . .	232·25	480·75	713
	Totals . . .	501·5	576·5	1078

(B.) *Fathers.*

	Class.	Below 66″·5.	Above 66″·5.	Totals.
Sons.	Below 67″·5 . .	211·25	153·75	365
	Above 67″·5 . .	152·75	560·25	713
	Totals . . .	364	714	1078

(C.)

Fathers.

	Class.	Below 67″·5.	Above 67″·5.	Totals.
Sons.	Below 68·5″ . .	356·25	182·25	538·5
	Above 68·5″ . .	145·25	394·25	539·5
	Totals . . .	501·5	576·5	1078

(D.)

Fathers.

	Class.	Below 68″·5.	Above 68″·5.	Totals.
Sons.	Below 69″·5 . .	506	182	688
	Above 69″·5 . .	149·5	240·5	390
	Totals . . .	655·5	422·5	1078

(E.)

Fathers.

	Class.	Below 69″·5.	Above 69″·5.	Totals.
Sons.	Below 70″·5 . .	669	147	816
	Above 70″·5 . .	128	134	262
	Totals . . .	797	281	1078

(F.)

Fathers.

	Class.	Below 70″·5.	Above 70″·5.	Totals.
Sons.	Below 69″·5 . .	641·25	46·75	688
	Above 69″·5 . .	271·75	118·25	390
	Totals . . .	913	165	1078

TABLE of Results.

Classificatiom.	Correlation.	Mean of sons.	Mean of fathers.
		k.	*h*.
A	·5939 ± ·0247	68″·64 (– ·416,32)	67″·74 (– ·087,00)
B	·5557 ± ·0261	68″·64 (– ·416,32)	67″·63 (– ·418,86)
C	·5529 ± ·0247	68″·50 (– ·001,16)	67″·74 (– ·087,30)
D	·5264 ± ·0264	68″·53 (·353,71)	67″·77 (·274,30)
E	·5213 ± ·0294	68″·60 (·696,57)	67″·76 (·641,30)
F	·5524 ± ·0307	68″·53 (·353,71)	67″·73 (1·023,44)

Now these results are of quite peculiar interest. They show us :—

(i.) That the probable error of r, as found by the present method, increases with h and k. But the increase is not very rapid, so that the probable errors of the series range only between ·025 and ·031. Hence while it is an advantage, it is not a very great advantage, to take the divisions of the groups near the medians. It is an advantage which may be easily counterbalanced by some practical gain in the method of observation when the division is not close to the medians.

(ii.) While the probable error, as found from the present method of calculation, is 1·5 to 2 times the probable error as found from the product moment, it is by no means so large as to seriously weigh against the new process, if the old is unavailable. It is quite true that the results given by the present process for six arbitrary divisions differ very considerably among themselves. But a consideration of the probable errors shows that the differences are sensibly larger than the probable error of the differences, even in some case double ; hence it is not the method but the assumption of normal correlation for such distributions which is at fault. As we shall hardly get a better variable than stature to hypothesise normality for, we see the weakness of the position which assumes without qualification the generality of the GAUSSIAN law of frequency.

(iii.) We cannot assert that the smaller the probable error the more nearly will the correlation, as given by the present process, agree with its value as found by the product moment. If we did we should discard ·5213, a very accordant result, in favour of ·5529, or even ·5939. The fact is that the higher the correlation the lower, *ceteris paribus*, the probable error, and this fact may obscure the really best result. Judging by the smallness of h and k and of the probable error, we should be inclined to select C or the value ·5529. This only differs from ·5198 by slightly more than the probable error of the difference (·033 as compared with ·029) ; but since both are found from the *same* statistics, and not from different samplings of the same population, this forms sufficient evidence in itself of want of normality. The approximate character of all results based on the theory of normal frequency must be carefully borne in mind ; and all we ought to conclude from the present data for inheritance of stature from father to son would be that the correlation $= \cdot 55 \pm \cdot 015$, while the product moment method would tell us more definitely that its value was $\cdot 52 \pm \cdot 015$. There is no question that the latter method is the better, but this does not hinder the new method from being extremely serviceable; for many cases it is the only one available.

Illustration VI. Effectiveness of Vaccination.—To find the correlation between strength to resist small-pox and the degree of effective vaccination.

We have in the earlier illustrations chosen cases in which in all probability a scale of character might possibly, if with difficulty, be determined. In the present case, the relationship is a very important one, but a quantitative scale is hardly discoverable. Nevertheless, it is of great interest to consider what results flow from the application of our method. We may consider our two characters as strength to resist the ravages of small-pox and as degree of effective vaccination. No quantitative scales are here available; all the statistics provide are the number of recoveries and deaths from small-pox, and the absence or presence of a definite vaccination cicatrix. Taking the Metropolitan Asylums Board statistics for the epidemic of 1893, we have the table given below, where the cases of "no evidence" have been omitted. Proceeding in the usual manner we find

$$\alpha_1 = \cdot 86929 \qquad \alpha_2 = \cdot 54157$$

$$h = 1\cdot 51139 \qquad k = \cdot 74145$$

$$\epsilon = \cdot 782454.$$

Hence the equation for r is

$$\cdot 782{,}454 = r + \cdot 560{,}310 r^2 - \cdot 096{,}378 r^3 + \cdot 081{,}881 r^4 - \cdot 000{,}172 r^5 - \cdot 040{,}059 r^6$$

whence $r = \cdot 5954$.

Summing up we have, after calculating the probable errors,

$$h = 1\cdot 5114 \pm \cdot 0287,$$

$$k = \cdot 7414 \pm \cdot 0205,$$

$$r = \cdot 5954 \pm \cdot 0272.$$

Strength to resist Small-pox when incurred.

	Cicatrix.	Recoveries.	Deaths.	Total.
Degree of effective Vaccination.	Present	1562	42	1604
	Absent	383	94	477
	Total	1945	136	2081

We see accordingly that there is quite a large correlation between recovery and the presence of the cicatrix. The two things are about as closely related as a child to its "mid-parent." While the correlation is very substantial and indicates the protective character of vaccination, even after small-pox is incurred, it is, perhaps, smaller than some over-ardent supporters of vaccination would have led us to believe.

Illustration VII. Effectiveness of Antitoxin Treatment.—To measure quantitatively the effect of antitoxin in diphtheria cases.

In like manner we may find the correlation between recovery and the administration of antitoxin in diphtheria cases. The statistics here are, however, somewhat difficult to obtain in a form suited to our purpose. The treatment by antitoxin began in the Metropolitan Asylums Board hospitals in 1895, but the serum was then administered only in those cases which gave rise to anxiety. Hence we cannot correlate recovery and death with the cases treated or not treated in that year, for those who were likely to recover were not dosed. In the year 1896 the majority of the cases were, on the contrary, treated with antitoxin, and those not treated were the slight cases of very small risk; hence, again, we are in great difficulties in drawing up a table.* Further, if we compare an antitoxin year with a non-antitoxin year, we ought to compare the cases treated with antitoxin in the former year with those which would probably have been treated with it in the latter year. Lastly, the dosage, nature of cases treated, and time of treatment have been modified by the experience gained, so that it seems impossible to club a number of years together, and so obtain a satisfactorily wide range of statistics. In 1897, practically all the laryngeal cases were treated with antitoxin. Hence the best we can do is to compare the laryngeal cases in two years, one before and one after the introduction of antitoxin. The numbers available are thus rather few, but will help us to form some idea of the correlation. I take the following data from p. 8 of the Metropolitan Asylums Board 'Report upon the Use of Antitoxic Serum for 1896':—

Laryngeal cases.	Recoveries.	Deaths.	Totals.
With antitoxin, 1896	319	143	462
Without antitoxin, 1894 . . .	177	289	466
Totals	496	432	928

---

* When a new drug or process is introduced the medical profession are naturally anxious to give every patient the possible benefit of it, and patients of course rush to those who first adopt it. But if the real efficiency of the process or drug is to be measured this is very undesirable. No definite data by which to measure the effectiveness of the novelty are thus available.

Here I find $r = \cdot4708 \pm \cdot0292$.

A further table is of interest :—

Laryngeal cases.	Requiring tracheotomy.	Not requiring it.	Totals.
Without antitoxin, 1894	261	205	466
With antitoxin, 1896	188	274	462
Totals	449	479	928

In this case we have $r = \cdot2385 \pm \cdot0335$.

Lastly, I have drawn up a third table :—

Total Infantile Cases, Ages 0—5 years.

	Recovery.	Death.	Totals.
With antitoxin, 1896 . . .	912	434	1346
Without antitoxin, 1894 . .	615	556	1171
Totals	1527	990	2517

Here we have* $r = \cdot2451 \pm \cdot0205$.

The three coefficients are all sensible as compared with their probable errors, and that between the administration of antitoxin and recovery in laryngeal cases is substantial. But the relationship is by no means so great as in the case of vaccination, and if its magnitude justifies the use of antitoxin, even when balanced against other ills which may follow in its train, it does not justify the sweeping statements of its effectiveness which I have heard made by medical friends. It seems until wider statistics are forthcoming a case for cautiously feeling the way forward rather than for hasty generalisations.

Illustration VIII. Effect on Produce of Superior Stock.—To find the effect of superiority of stock on percentage goodness of produce.

To illustrate this and also the formula (lxxxiii.) for six correlation coefficients, we will investigate the effect of selecting sire, dam, and one grandsire on the produce when there

---

* The values of r for all the three cases of this Illustration were determined with great ease from Equation (xxiv.).

is selective pairing of dam and sire. We will suppose grandsire, dam, and sire to be above the average, and investigate what proportion of the produce will be above the average. As numbers very like those actually occurring in the case of dogs, horses, and even men, we may take

Correlation of grandsire and offspring . = ·25
,, sire or dam and offspring = ·5 in both cases
,, sire and grandsire . . . = ·5
Selective mating for sire and dam. . . = ·2

We will suppose zero correlation between paternal grandsire and dam, although with selective mating this may actually exist.* We have then the following system :—

$$r_{14} = \cdot 25, \quad r_{24} = \cdot 5, \quad r_{34} = \cdot 5, \quad r_{23} = \cdot 2, \quad r_{12} = \cdot 5, \quad r_{13} = 0.$$

Hence, substituting these values in (lxxxvii.), we find—after some arithmetic :

$$(\mathrm{Q} - \mathrm{Q}_0)/\mathrm{Q}_0 = 1\cdot 4851.$$

But Q_0 is the chance of produce above the average if there were no heredity between grandsire, sire, and dam, and no assortative mating.

Hence it equals $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\mathrm{N} = \frac{\mathrm{N}}{16} \quad \therefore \mathrm{Q} = \cdot 1553\,\mathrm{N}.$

Or, of the produce ·5 N above the average, ·1553 N instead of ·0625 N are born of the superior stock owing to inheritance, &c. In other words, out of the ·5 N above the average, ·1553 N are produced by the stock in sire, dam, and grandsire above the average, or by ·1827 of the total stock.† The remaining ·8173 only produce ·3447 N, or the superior stock produces produce above the average at over twice the rate of the inferior stock. Absolutely, the inferior stock being seven times as numerous produces about seven-tenths of the superior offspring.

Illustration IX. Effect of Exceptional Parentage.—Chance of an exceptional man being born of exceptional parents.

Let us enlarge the example in Illustration II., and seek the proportion of exceptional men, defined as one in twenty, born of exceptional parents in a community with assortative mating.

---

* A correlation, if there be substantial selective mating, may exist between a man and his mother-in-law. Its rumoured absence, if established scientifically, would not, however, prove the non-existence of selective mating, for A may be correlated with B and C, but these not correlated with each other.

† The proportion of pairs of parents associated with a grandsire above the average was found by putting ·5, ·2, and 0 for the three correlation coefficients in (lxxxv.). In comparing with Illustration II., the reader must remember we there dealt with an exceptional father, 1 in 20, here only with relatives above the average—a very less stringent selection.

Here we take for father and son $r_{12} = \cdot 5$, for mother and son $r_{13} = \cdot 5$, and for assortative mating, $r_{23} = \cdot 2$.

We have then to apply the general formulæ (lxxxiii.) and (lxxxiv.) for the case of three variables. We have

$$\begin{aligned} h_1 &= h_2 = h_3 = 1{\cdot}64485 \\ \beta_1 &= \beta_2 = \beta_3 = {\cdot}484{,}795 \\ v_1' &= v_1'' = v_1''' = 1{\cdot}644{,}850 \\ v_2' &= v_2'' = v_2''' = 1{\cdot}705{,}532 \\ v_3' &= v_3'' = v_3''' = -\ {\cdot}484{,}356 \\ v_4' &= v_4'' = v_4''' = -\ 5{\cdot}913{,}290 \end{aligned}$$

Whence, after some arithmetical reduction, we find

$$(Q - Q_0)/Q_0 = 20{\cdot}0389.$$

But $Q_0 = \frac{1}{20} \times \frac{1}{20} \times \frac{1}{20} N = \frac{1}{8000} N$. Hence $Q = {\cdot}00263\ N$.

We must now distinguish between the absolute and relative production of exceptional men by exceptional and non-exceptional parents. The exceptional pairs of parents are obtained by (xix.), whence we deduce, putting $r = \cdot 2$, $h = k = 1{\cdot}64485$,

$$\frac{ad - bc}{N^2} = \frac{d}{N} - \frac{(d + b)(d + c)}{N^2} = \frac{d}{N} - \frac{1}{400} = {\cdot}002745.$$

Whence the number of pairs of parents, both exceptional

$$= {\cdot}005245\ N.$$

Thus, ·005245 N pairs of exceptional parents produce ·00263 N exceptional sons, and ·994755 N pairs of parents, non-exceptional in character, produce ·04737 N exceptional sons, *i.e.*, the remainder of the $\frac{1}{20}$ N. The rates of production are thus as ·5014 to ·0476. Or: *Pairs of exceptional parents produce exceptional sons at a rate more than ten times as great as pairs of non-exceptional parents.* At the same time, eighteen times as many exceptional sons are born to non-exceptional as to exceptional parents, for the latter form only about ½ per cent. of the community.

The reader who will carefully investigate Illustrations II., VIII., and IX. will grasp fully why so many famous men are born of undistinguished parents, but will, at the same time, realise the overwhelming advantage of coming of a good stock.

SOLD BY HARRISON AND SONS, ST. MARTIN'S LANE.

Demy 8vo, pp. 224. Price 7*s*. 6d.

THE RECORD OF THE ROYAL SOCIETY, 1897.

Containing an Account of the Foundation, History, &c., of the Society.

Demy 8vo, pp. 209. Price 5*s*.

THE YEAR-BOOK OF THE ROYAL SOCIETY, 1900.

Containing an Account of the Chief Events and Arrangements of the Year.

PHILOSOPHICAL TRANSACTIONS.

SERIES A. Containing Papers of a Mathematical or Physical Character.
VOL. 193 (1900), Price £1 10*s*., with 28 Plates.
VOL. 194 (1900), Price £1 1*s*. 6*d*., with 5 Plates.

SERIES B. Containing Papers of a Biological Character.
VOL. 191 (1899), Price £1 12*s*., with 31 Plates.
VOL. 192 (1900), Price £1 7*s*. 6*d*., with 13 Plates.

PUBLISHED BY CLAY AND SONS, AVE MARIA LANE.

CATALOGUE OF SCIENTIFIC PAPERS,

COMPILED BY THE ROYAL SOCIETY.

The volumes on sale as follows:—Vols. 1–6 (1800–63), cloth (vol. 1 in half-morocco), £4 net; half-morocco, £5 5*s*. net. Vols. 7, 8 (1864–73), cloth, £1 11*s*. 6*d*. net; half-morocco, £2 5*s*. net. Single vols., cloth, 20*s*.; half-morocco, 28*s*. net. Vols. 9, 10 (1874–83), cloth, £1 5*s*. net; half-morocco, £1 12*s*. net. Vol. 11 (1874–83), cloth, £1 5*s*. net; half-morocco, £1 12*s*. net.

A reduction of price to Fellows of the Royal Society.

PUBLISHED BY KEGAN PAUL, TRENCH, TRÜBNER, AND CO.

A few copies at the reduced price of £1 1*s*. (published at £3).

Royal 4to, pp. iv.—936, cloth.

A MONOGRAPH OF THE HORNY SPONGES.

BY R. VON LENDENFELD.

With 51 Lithographic and Photographic Plates.

A further reduction of price to Fellows of the Royal Society.

PUBLISHED BY KEGAN PAUL, TRENCH, TRÜBNER, AND CO.

In 1 vol., 4to. Pp. 500. With 6 Chromolithographs of the remarkable Sunsets of 1883 and 40 Maps and Diagrams.

THE ERUPTION OF KRAKATOA AND SUBSEQUENT PHENOMENA.

Report of the Krakatoa Committee of the Royal Society.

EDITED BY G. J. SYMONS, F.R.S.

Price 30*s*. To Fellows of the Royal Society, 20*s*.

PUBLISHED BY KEGAN PAUL, TRENCH, TRÜBNER, AND CO.

Royal 4to, pp. xiv.—326, cloth. Price 21*s*.

OBSERVATIONS OF THE INTERNATIONAL POLAR EXPEDITIONS.

1882–1883.

FORT RAE.

With 32 Lithographic Folding Plates.

A reduction of price to Fellows of the Royal Society.

Continental Agents for the publications of the Royal Society, Messrs. FRIEDLÄNDER and SON, Berlin.

Just Published, and Sold by Dulau and Co., Soho Square.

SEPARATE PAPERS FROM THE 'PHILOSOPHICAL TRANSACTIONS.'
1899–1900.

A. 252. An Experimental Investigation of the Thermo-dynamical Properties of Superheated Steam.—On the Cooling of Saturated Steam by Free Expansion. By John H. Grindley, B.Sc., Wh.Sc. Exhibition (1851) Scholar, late Fellow of the Victoria University. Price 1*s.* 6*d.*

A. 253. A Comparison of Platinum and Gas Thermometers, including a Determination of the Boiling-Point of Sulphur on the Nitrogen Scale.—An Account of Experiments made in the Laboratory of the Bureau International des Poids et Mesures at Sèvres. By John Allen Harker, D.Sc., formerly Bishop Berkeley Fellow of Owens College, Manchester, Research Assistant at Kew Observatory, and Pierre Chappuis, Ph.D., Savant Attaché au Bureau International des Poids et Mesures, Sèvres. Price 5*s.*

A. 254. On the Propagation of Earthquake Motion to Great Distances. By R. D. Oldham, Geological Survey of India. Price 1*s.* 6*d.*

A. 255. Impact with a Liquid Surface studied by the Aid of Instantaneous Photography. Paper II. By A. M. Worthington, M.A., F.R.S., and R. S. Cole, M.A. Price 2*s.* 6*d.*

A. 256. Gold-aluminium Alloys. By C. T. Heycock, F.R.S., and F. H. Neville, F.R.S. Price 2*s.* 6*d.*

A. 257. Bakerian Lecture.—The Specific Heats of Metals, and the Relation of Specific Heat to Atomic Weight. By W. A. Tilden, D.Sc., F.R.S., Professor of Chemistry in the Royal College of Science, London. With an Appendix by Professor John Perry, F.R.S. Price 1*s.*

A. 258. On the Association of Attributes in Statistics: with Illustrations from the Material of the Childhood Society, &c. By G. Udny Yule, formerly Assistant Professor of Applied Mathematics, University College, London. Price 3*s.*

A. 259. The Ionization of Dilute Solutions at the Freezing Point. By W. C. D. Whetham, M.A., Fellow of Trinity College, Cambridge. Price 2*s.* 6*d.*

A. 260. Combinatorial Analysis.—The Foundations of a New Theory. By Major P. A. MacMahon, D.Sc., F.R.S. Price 1*s.* 6*d.*

A. 261. Title, Contents, Index, &c., Vol. 194. Price 6*d.*

B. 180. Agricultural, Botanical, and Chemical Results of Experiments on the Mixed Herbage of Permanent Grass-land, conducted for many years in succession on the same Land.—Part III. The Chemical Results—Section I. By Sir John Bennet Lawes. Bart., D.C.L., Sc.D., F.R.S., and Sir J. Henry Gilbert, LL.D., Sc.D., F.R.S. Price 3*s.* 6*d.*

B. 181. On Intestinal Absorption, especially on the Absorption of Serum, Peptone, and Glucose. By E. Waymouth Reid, F.R.S., Professor of Physiology in University College, Dundee, St. Andrew's University, N.B. Price 4*s.*

B. 182. The External Features in the Development of *Lepidosiren paradoxa*, Fitz. By J. Graham Kerr. Price 7*s.*

B. 183. On the Formation of the Pelvic Plexus, with especial reference to the Nervus Collector in the Genus *Mustelus* By R. C. Punnett, B.A., Scholar of Gonville and Caius College, Cambridge. Price 2*s.*

B. 184. Title, Contents, Index, &c., Vol. 192. Price 6*d.*

B. 185. On the Gastric Gland of Mollusca and Decapod Crustacea: its Structure and Functions. By C. A. MacMunn, M.A., M.D. Price 5*s.*

B. 186. Upon the Development of the Enamel in Certain Osseous Fish. By Charles S. Tomes, M.A., F.R.S. Price 1*s.*

B. 187. Extinct Mammalia from Madagascar.—I. *Megaladapis insignis*, sp. n. By C. I. Forsyth Major, M.D., F.Z.S. Price 1*s.*

B. 188. Further Observations on the Nature and Functions of the Nodules of Leguminous Plants By Maria Dawson, B.Sc. (London and Wales), 1851 Exhibition Science Research Scholar. Price 2*s.*

Cloth cases for binding the Philosophical Transactions can be obtained of Harrison and Sons, St. Martin's Lane. Price 1*s.* 6*d.*

PHILOSOPHICAL TRANSACTIONS

OF THE

ROYAL SOCIETY OF LONDON.

SERIES A.

Vol. 195, pp. 79–150.

MATHEMATICAL CONTRIBUTIONS TO THE THEORY OF EVOLUTION.—VIII. ON THE INHERITANCE OF CHARACTERS NOT CAPABLE OF EXACT QUANTITATIVE MEASUREMENT.

PART I.—INTRODUCTORY.
PART II.—ON THE INHERITANCE OF COAT-COLOUR IN HORSES.
PART III.—ON THE INHERITANCE OF EYE-COLOUR IN MAN.

BY

KARL PEARSON, F.R.S.,

WITH THE ASSISTANCE OF

ALICE LEE, D.Sc.,

UNIVERSITY COLLEGE, LONDON.

LONDON:

PUBLISHED FOR THE ROYAL SOCIETY
BY DULAU AND CO., 37, SOHO SQUARE, W.
CONTINENTAL AGENTS, MESSRS. FRIEDLÄNDER AND SON, BERLIN.

1900.

A. 264.

29.10.1900.

Price Three Shillings and Sixpence.

INDEX SLIP

Pearson (Karl), with assistance of Alice Lee—Mathematical Contributions to the Theory of Evolution. VIII.—On the Inheritance of Characters not capable of Exact Quantitative Measurement.
Phil. Trans., A, vol. 195, 1900, pp. 79-150.

Horse, Thoroughbred; Inheritance of Coat-Colour.
Pearson (Karl), with assistance of Alice Lee.
Phil. Trans., A, vol. 195, 1900, pp. 79-150.

Inheritance—Blended, exclusive and particulate; of characters not quantitatively measurable; collateral, direct, &c.
Pearson (Karl), with assistance of Alice Lee.
Phil. Trans., A, vol. 195, 1900, pp. 79–150.

Man; Inheritance of Eye-Colour.
Pearson (Karl), with assistance of Alice Lee.
Phil. Trans., A, vol. 195, 1900, pp. 79–150.

III. *Mathematical Contributions to the Theory of Evolution.*—VIII. *On the Inheritance of Characters not capable of Exact Quantitative Measurement.*—Part I. *Introductory.* Part II. *On the Inheritance of Coat-colour in Horses.* Part III. *On the Inheritance of Eye-colour in Man.*

By KARL PEARSON, *F.R.S.*, *with the assistance of* ALICE LEE, *D.Sc.*, *University College, London.*

Received August 5,—Read November 16, 1899; withdrawn, rewritten, and again received March 5, 1900.

CONTENTS.

PART I.—*Introductory.*

PART II.—*On Coat-colour Inheritance in the Thoroughbred Horse.*

PART III.—*On Eye-colour Inheritance in Man.*

NOTE.

This memoir was originally presented to the Society on August 5, 1899, and read on November 16, 1899. In working out by the same theory the coefficients of inheritance for Basset Hounds, Mr. LESLIE BRAMLEY-MOORE discovered that the method adopted was not exact enough in its process of proportioning. Accordingly, with the assistance of Mr. L. N. G. FILON, we immensely developed the theory, so that it was necessary to rewrite the theoretical part of the original memoir. This has been carried out in Part VII. of this series. The present memoir consists substantially of the portions of the original memoir relating to the inheritance of coat-colour in Horses and eye-colour in Man, with the numerical details and the resulting conclusions modified, so far as the extended theory rendered this necessary. In the very laborious work of reconstructing my original tables I have received the greatest possible assistance from Dr. ALICE LEE, and I now wish to associate her name with mine on the memoir.* The memoir was at my request returned to me for revision after it had been accepted for the 'Philosophical Transactions.'

PART I.—INTRODUCTORY.

(1.) A CERTAIN number of characters in living forms are capable of easy observation, and thus are in themselves suitable for observation, but they do not admit of an exact quantitative measurement, or only admit of this with very great labour. The object of the present paper is to illustrate a method by which the correlation of such characters may be effectively dealt with in a considerable number of cases. The conditions requisite are the following:—

(i.) The characters should admit of a quantitative order, although it may be impossible to give a numerical value to the character in any individual.

Thus it is impossible at present to give a quantitative value to a brown, a bay, or a roan horse, but it is not impossible to put them in order of relative darkness of shade. Or, again, we see that a blue eye is lighter than a hazel one, although we cannot *à priori* determine their relative positions numerically on a quantitative scale.

Even in the markings on the wings of butterflies or moths, where it might be indefinitely laborious to count the scales, some half dozen or dozen specimens may be taken to fix a quantitative order, and all other specimens may be grouped by inspection in the intervals so determined.

We can even go a stage further and group men or beasts into simply two categories—light and dark, tall and short, dolichocephalic and brachycephalic—and so we might ascertain by the method adopted whether there is, for example, correlation between complexion and stature, or stature and cephalic index.

(ii.) We assume that the characters are a function of some variable, which, if we

---

* I have further to thank Mr. LESLIE BRAMLEY-MOORE, Mr. L. N. G. FILON, M.A., Mr. W. R. MACDONELL, M.A., LL.D. and Miss C. D. FAWCETT, B.Sc., for much help in the arithmetic, often for laborious calculations by processes and on tables, which were none the less of service if they were afterwards discarded for others. To Mr. BRAMLEY-MOORE I owe the extraction and part of the arithmetical reduction of the horse-colour tables.

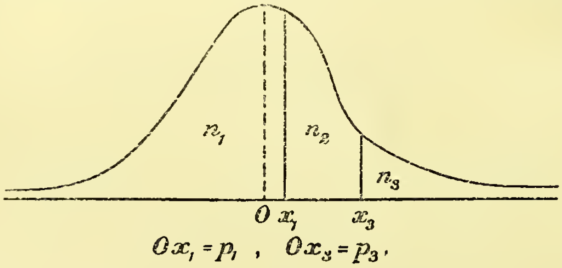
could determine a quantitative scale, would give a distribution obeying—at any rate to a first approximation—the normal law of frequency.

The whole of the theoretical investigations are given in a separate memoir, in which the method applied is illustrated by numerical examples taken from inheritance of eye-colour in man, of coat-colour in horses and dogs, and from other fields. We shall not therefore in this paper consider the processes involved, but we may make one or two remarks on the justification for their use. If we take a problem like that of coat-colour in horses, it is by no means difficult to construct an order of intensity of shade. The variable on which it depends may be the amount of a certain pigment in the hair, or the relative amounts of two pigments. Much the same applies to eye-colour. In both cases we may fail to obtain a true quantitative scale, but we may reasonably argue that, if we could find the quantity of pigment, we should be able to form a continuous curve of frequency. We make the assumption that this curve—to at any rate a first approximation—is a normal curve. Now if we take any line parallel to the axis of frequency and dividing the curve, we divide the total frequency into two classes, which, so long as there is a quantitative order of tint or colour, will have their relative frequency unchanged, however we, in our ignorance of the fundamental variable, distort its scale. For example, if we classify horses into bay and darker, chestnut and lighter, we have a division which is quite independent of the quantitative range we may give to black, brown, bay, chestnut, roan, grey, &c.

Precisely the same thing occurs with eye-colour; we classify into brown and darker, hazel and lighter, and the numbers in these classes will not change with the quantitative scale ultimately given to the various eye-tints. Our problem thus reduces to the following one: Given two classes of one variable, and two classes of a second variable correlated with it, deduce the value of the correlation. Classify sire and foal into bay and darker, chestnut and lighter; mother and daughter into brown and darker, hazel and lighter, and then find the correlation due to inheritance between the coat-colour or eye-colour of these pairs of relations. The method of doing this is given in Memoir VII. of this series. Its legitimacy depends on the assumptions (i.) and (ii.) made above, which may I think be looked upon as justifiable approximations to the truth.

Of course the probable error of the method is larger than we find it to be when correlation is determined from the product-moment. Its value varies with the inequality of the frequency in the two classes given by the arbitrary division. It will be least when we make that frequency as nearly equal as possible—a result which can often be approximately reached by a proper classification. In our present data the probable errors vary from about ·02 to ·04, values which by no means hinder us from drawing general conclusions, and which allow of quite satisfactory general results.

(2.) So far we have only spoken of the two classes, which are necessary if we merely want to determine the correlation. But if we wish to deal with relative variability we must have more than two classes. We have, in fact, in our tables preserved Mr. GALTON'S eight eye-colour classes and the seventeen classes under which the coat-colour of thoroughbred horses is classified in WETHERBY'S studbooks. Such a classification enables us at any rate approximately to ascertain relative variability, and, what is more, to reconstruct approximately the quantitative scale according to which the tints must be distributed in order that the frequency should be normal. For, in order to attain this result, we have to ascertain from a table of the areas of the normal curve the ratio of the length of the abscissa to the standard deviation which corresponds to any given increase of frequency. Let us suppose that three classes have been made—n_1, n_2, n_3, represented by the areas of the normal curve in the accompanying diagram so marked. Let p_1 and p_3 be the distances of the mean from the two boundaries of n_2. Here p_1 may be negative, or p_3 infinite, &c. Then if $h_1 = p_1/\sigma$, $h_3 = p_3/\sigma$, we find at once, if N = total frequency,



$$\frac{n_1 - n_2 - n_3}{\mathrm{N}} = \sqrt{\frac{2}{\pi}} \int_0^{h_1} e^{-\frac{1}{2}x^2} dx \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{i.}).$$

$$\frac{n_1 + n_2 - n_3}{\mathrm{N}} = \sqrt{\frac{2}{\pi}} \int_0^{h_3} e^{-\frac{1}{2}x^2} dx \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{ii.}).$$

Now the integrals on the right are tabulated, and thus, since the left-hand side is a known numerical quantity, it follows that p_1/σ and p_3/σ, and accordingly the range $(p_3 - p_1)/\sigma$ of the class in terms of the standard deviation, are fully determined.

Thus, if ϵ be the range on the scale of tint or colour of the group of which the observed frequency is n_2, we have $\epsilon = p_3 - p_1$, and thus $\epsilon/\sigma = q$ say, is known. For a second series $\epsilon/\sigma' = q'$. Hence $\sigma/\sigma' = q'/q$, and accordingly the ratio of the variabilities of the two series is determined.

Again, the ratio $p_1/(p_3 - p_1)$ enables us to find the position of the mean in terms of the range on the scale occupied by the tint corresponding to the frequency n_2. As a rule we shall take this tint to be that in which the mean actually lies, in which case we shall have $p_1/(p_3 + p_1)$ as determining the ratio in which the mean divides the true quantitative range of this particular tint.

(3.) Let

$$\eta = p_1(p_3 - p_1) = h_1/(h_3 - h_1) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{iii.}),$$

$$\zeta = \sigma/\sigma' = (h_3' - h_1')/(h_3 - h_1) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{iv.}).$$

It remains to find the probable errors of these quantities.

Suppose Σ_x to be the standard deviation for the errors in a quantity x, and R_{xy} the correlation coefficient for errors in two quantities x and y.

Further let

$$H = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}h^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{v.}),$$

where subscripts and dashes may be attached to H to correspond to like distinguishing marks attached to h.

Since

$$\frac{2n_1 - N}{2N} = \frac{1}{\sqrt{2\pi}} \int_0^{h_1} e^{-\frac{1}{2}x^2} dx \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{vi.}),$$

we have at once $\qquad \delta n_1 = NH_1 \delta h_1,$

and $\qquad \Sigma_{h_1} = \Sigma_{n_1}/(NH_1) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{vii.}).$

Similarly, $\delta n_3 = - NH_3 \delta h_3$, whence:

$$\Sigma_{h_3} = \Sigma_{n_3}/(NH_3) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{viii.}).$$

Further, we have $\qquad \Sigma_{h_1}\Sigma_{h_3}R_{h_1h_3} = - \Sigma_{n_1}\Sigma_{n_3}R_{n_1n_3}/(N^2H_1H_3) \quad . \quad . \quad . \quad . \quad . \quad (\text{ix.})$;

but, as is shewn in Part VII., § 4,

$$\Sigma_{n_1}^{\ 2} = \frac{n_1(N - n_1)}{N}, \qquad \Sigma_{n_3}^{\ 2} = \frac{n_3(N - n_3)}{N} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{x.}),$$

$$\Sigma_{n_1}\Sigma_{n_3}R_{n_1n_3} = - \frac{n_1n_3}{N} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xi.}).$$

Thus we find

$$\text{Probable error of } h_1 = \cdot 67449 \Sigma_{h_1}$$

$$= \frac{\cdot 67449}{\sqrt{N}} \frac{1}{H_1} \sqrt{\frac{n_1(N - n_1)}{N^2}} \quad . \quad . \quad . \quad . \quad . \quad (\text{xii.}).$$

$$\text{Probable error of } h_3 = \frac{\cdot 67449}{\sqrt{N}} \frac{1}{H_3} \sqrt{\frac{n_3(N - n_3)}{N^2}} \quad . \quad . \quad . \quad . \quad . \quad (\text{xiii.}).$$

Correlation in errors in h_1 and h_3, or $R_{h_1h_3}$, is given by

$$\Sigma_{h_1}\Sigma_{h_3}R_{h_1h_3} = \frac{n_1n_3}{N^3H_1H_3} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xiv.}).$$

Let $u = h_3 - h_1$, $u' = h_3' - h_1'$ be the ratio to the respective standard deviations of the ranges corresponding to the groups n_2 and n_2'. Then

$$\Sigma_u^{\ 2} = \Sigma_{h_1}^{\ 2} + \Sigma_{h_3}^{\ 2} - 2\Sigma_{h_1}\Sigma_{h_3}R_{h_1h_3}$$

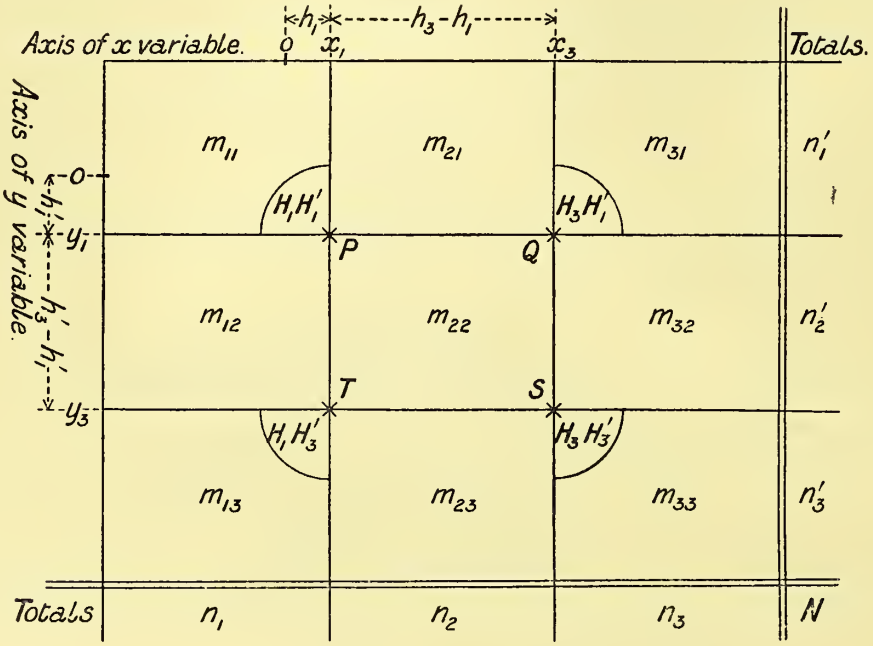
$$= \frac{1}{N^3}\left\{\frac{n_1(N - n_1)}{H_1^{\ 2}} + \frac{n_3(N - n_3)}{H_3^{\ 2}} - \frac{2n_1n_3}{H_1H_3}\right\};$$

whence, if ν be a proportional frequency $= n/N$, we readily find

$$\text{Probable error of } u = \frac{\cdot 67449}{\sqrt{N}} \left\{\frac{\nu_1}{H_1^{\ 2}} + \frac{\nu_3}{H_3^{\ 2}} - \left(\frac{\nu_1}{H_1} + \frac{\nu_3}{H_3}\right)^2\right\}^{\frac{1}{2}} \quad . \quad . \quad . \quad (\text{xv.}).$$

$$\text{Probable error of } u' = \frac{\cdot 67449}{\sqrt{N}} \left\{\frac{\nu_1'}{H_1'^2} + \frac{\nu_3'}{H_3'^2} - \left(\frac{\nu_1'}{H_1'} + \frac{\nu_3'}{H_3'}\right)^2\right\}^{\frac{1}{2}} . \quad . \quad . \quad (\text{xvi.}).$$

I now proceed to determine the correlation in the errors made in determining the ranges corresponding to any two classes of any two variables which are correlated. For this purpose let the frequency correlation table be dressed as follows, in the diagram below.



Here m_{ij} denotes the frequency of individuals common to the two classes n_i and n'_j. Let M_{ij} denote its "conjugate," or all the frequency which appears in neither n_i nor n'_j; then

$$N = M_{ij} + n_i + n'_j - m_{ij} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xvii.)}.$$

As before, we have

$$\Sigma_{n_i}{}^2 = \frac{n_i(N - n_i)}{N}, \qquad \Sigma_{n'_j}{}^2 = \frac{n'_j(N - n'_j)}{N} \quad . \quad . \quad . \quad . \quad \text{(xviii.)}.$$

$$\Sigma_{M_{ij}}{}^2 = \frac{M_{ij}(N - M_{ij})}{N}, \qquad \Sigma_{m_{ij}}{}^2 = \frac{m_{ij}(N - m_{ij})}{N} \quad . \quad . \quad . \quad . \quad \text{(xix.)}.$$

Further, since m_{ij} and M_{ij} are mutually exclusive, we have

$$\Sigma_{M_{ij}}\Sigma_{m_{ij}}R_{M_{ij}m_{ij}} = -\frac{M_{ij}m_{ij}}{N} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xx.)}.$$

From (xvii.) we have for small variations

$$\delta n_i + \delta n'_j = \delta m_{ij} - \delta M_{ij}.$$

Hence

$$2\Sigma_{n_i}\Sigma_{n'_j}R_{n_in'_j} = \Sigma_{m_{ij}}{}^2 + \Sigma_{M_{ij}}{}^2 - \Sigma_{n_i}{}^2 - \Sigma_{n_{ij}}{}^2 - 2\Sigma_{m_{ij}}\Sigma_{M_{ij}}R_{M_{ij}m_{ij}} \quad . \quad . \quad . \quad \text{(xxi)}.$$

Substituting the values given above we find, after some reductions,

$$\Sigma_{n_i}\Sigma_{n'_j}R_{n_i n'_j} = \frac{Nm_{ij} - n_i n'_j}{N} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxii.)}$$

This result, which is extremely simple in form, gives the correlation in errors made in determining the frequencies in any two classes whatever of any two correlated variables.

I next proceed to find the correlation between errors in u and u', the ratio of the ranges occupied by any two classes to their respective standard deviations.

We have

$$\begin{aligned} \delta n_2 + \delta n_3 &= -NH_1\delta h_1 ; \\ \delta n_3 &= -NH_3\delta h_3. \end{aligned}$$

Hence
$$\delta(h_3 - h_1) = \frac{\delta n_2}{NH_1} + \frac{\delta n_3}{N}\left(\frac{1}{H_1} - \frac{1}{H_3}\right).$$

Similarly
$$\delta(h_3' - h_1') = \frac{\delta n'_2}{NH_1'} + \frac{\delta n_3'}{N}\left(\frac{1}{H_1'} - \frac{1}{H_3'}\right).$$

Multiply the first by the second, and summing as usual for all possible errors, we have, by using (xxii.)

$$\Sigma_u\Sigma_{u'}R_{uu'} = \frac{1}{N}\left\{\frac{Nm_{22} - n_2n_2'}{N^2H_1H_1'} + \frac{Nm_{23} - n_2n_3'}{N^2H_1}\left(\frac{1}{H_1'} - \frac{1}{H_3'}\right) + \frac{Nm_{32} - n_2'n_3}{N^2H_1'}\left(\frac{1}{H_1} - \frac{1}{H_3}\right) \right.$$
$$\left. + \left(\frac{Nm_{33} - n_3n_3'}{N^2}\right)\left(\frac{1}{H_1} - \frac{1}{H_3}\right)\left(\frac{1}{H_1'} - \frac{1}{H_3'}\right)\right\}.$$

Collecting the like H's we find, after very considerable reductions,

$$\Sigma_u\Sigma_{u'}R_{uu'} = \frac{1}{N}\left\{\frac{Nm_{11} - n_1n_1'}{N^2H_1H_1'} + \frac{Nm_{33} - n_3n_3'}{N^2H_3H_3'} + \frac{Nm_{13} - n_1n_3'}{N^2H_1H_3'} + \frac{Nm_{31} - n_3n_1'}{N^2H_3H_1'}\right\}$$
$$. \quad . \quad . \quad . \quad \text{(xxiii.)},$$

or, $$\Sigma_u\Sigma_{u'}R_{uu'} = \frac{1}{N}\left\{\frac{\mu_{11} - \nu_1\nu_1'}{H_1H_1'} + \frac{\mu_{33} - \nu_3\nu_3'}{H_3H_3'} + \frac{\mu_{13} - \nu_1\nu_3'}{H_1H_3'} + \frac{\mu_{31} - \nu_3\nu_1'}{H_3H_1'}\right\} \quad . \quad . \; (\text{xxiii.}^{\text{bis}}).$$

where $\mu_{ij} = m_{ij}/N$ = proportional frequency.

A glance at our diagram on the previous page of the correlation table divided into nine classes, shows at once the symmetrical formation of this result. By writing at the points P, Q, S, and T, the ordinate there of the normal surface, on the supposition of no correlation and $N = 1$, the construction of the result is still more clearly brought out.

We are now in a position to determine the probable errors of η and ζ. We have

$$\delta\eta = \frac{h_3\delta h_1 - h_1\delta h_3}{u^2}.$$

Hence

$$\Sigma_\eta^2 = \frac{1}{u^4}\left(h_3{}^2\Sigma_{h_1}{}^2 + h_1{}^2\Sigma_{h_3}{}^2 - 2h_3h_1\Sigma_{h_1}\Sigma_{h_3}\mathrm{R}_{h_1h_3}\right)$$

$$= \frac{h_1{}^2h_3{}^2}{u^4\mathrm{N}}\left\{\frac{n_1(\mathrm{N}-n_1)}{(h_1\mathrm{H}_1)^2\,\mathrm{N}^2} + \frac{n_3(\mathrm{N}-n_3)}{(h_3\mathrm{H}_3)^2\mathrm{N}^2} - \frac{2n_1n_3}{(h_1\mathrm{H}_1)(h_2\mathrm{H}_3)\mathrm{N}^2}\right\}.$$

Or, Probable error of η

$$= \frac{\cdot 67449}{\sqrt{\mathrm{N}}}\,\frac{h_1h_3}{u^2}\left\{\frac{\nu_1}{(h_1\mathrm{H}_1)^2} + \frac{\nu_3}{(h_3\mathrm{H}_3)^2} - \left(\frac{\nu_1}{h_1\mathrm{H}_1} + \frac{\nu_3}{h_3\mathrm{H}_3}\right)^2\right\}^{\frac{1}{2}} \quad . \quad . \quad . \quad (\text{xxiv.}).$$

where u is the range $h_3 - h_1$, and ν_1 and ν_3 are the proportional frequencies, as before. Care must be taken, if the class n_2 cover, as it usually will in our present investigations, the mean, to put h_1 negative within the radical. In other words, for a class covering the mean we have

Probable error of η

$$= \frac{\cdot 67449}{\sqrt{\mathrm{N}}}\,\frac{h_1h_3}{(h_1+h_3)^2}\left\{\frac{\nu_1}{(h_1\mathrm{H}_1)^2} + \frac{\nu_3}{(h_3\mathrm{H}_3)^2} - \left(\frac{\nu_1}{h_1\mathrm{H}_1} - \frac{\nu_3}{h_3\mathrm{H}_3}\right)^2\right\}^{\frac{1}{2}} \quad . \quad . \quad (\text{xxv.}).$$

Lastly we have

$$\zeta = u'/u\,,$$

or,

$$\delta\zeta = \frac{u\delta u' - u'\delta u}{u^2} = \frac{u'}{u}\left(\frac{\delta u'}{u'} - \frac{\delta u}{u}\right).$$

Hence

$$\Sigma_\zeta{}^2 = \frac{u'^2}{u^2}\left\{\frac{\Sigma_{u'}{}^2}{u'^2} + \frac{\Sigma_u{}^2}{u^2} - \frac{2\Sigma_u\Sigma_u\mathrm{R}_{uu'}}{uu'}\right\}.$$

Thus: Probable error of ζ

$$= \cdot 67449\zeta\left\{\frac{\Sigma_u{}^2}{u^2} + \frac{\Sigma_{u'}{}^2}{u'^2} - \frac{2\Sigma_u\Sigma_{u'}\mathrm{R}_{uu'}}{uu'}\right\}^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad . \quad (\text{xxvi.}),$$

where we have by (xv.), (xvi.), and (xxiii.$^{\text{bis}}$)

$$\Sigma_u{}^2 = \frac{1}{\mathrm{N}}\left\{\frac{\nu_1}{\mathrm{H}_1{}^2} + \frac{\nu_3}{\mathrm{H}_3{}^2} - \left(\frac{\nu_1}{\mathrm{H}_1} + \frac{\nu_3}{\mathrm{H}_3}\right)^2\right\},$$

$$\Sigma_{u'}{}^2 = \frac{1}{\mathrm{N}}\left\{\frac{\nu_1'}{\mathrm{H}_1{}'^2} + \frac{\nu_3'}{\mathrm{H}_3{}'^2} - \left(\frac{\nu_1'}{\mathrm{H}_1'} + \frac{\nu_3'}{\mathrm{H}_3'}\right)^2\right\},$$

$$\Sigma_u\Sigma_{u'}\mathrm{R}_{u'u} = \frac{1}{\mathrm{N}}\left\{\frac{\mu_{11} - \nu_1\nu_1'}{\mathrm{H}_1\mathrm{H}_1'} + \frac{\mu_{33} - \nu_3\nu_3'}{\mathrm{H}_3\mathrm{H}_3'} + \frac{\mu_{13} - \nu_1\nu_3'}{\mathrm{H}_1\mathrm{H}_3'} + \frac{\mu_{31} - \nu_3\nu_1'}{\mathrm{H}_3\mathrm{H}_1'}\right\},$$

where, as before, μ's and ν's represent proportional frequencies.

In the following investigations on coat-colour and eye-colour inheritance I have not thought it needful to give in every one of the thirty-six relationships dealt with the probable errors of the means, ratio of variabilities, and the coefficients of inheritance (η, ζ, and r). The arithmetical labour would have been too great, for the expressions as given above are somewhat complex. It is, however, necessary to keep the approximate values of these probable errors in view, and, as our results classify themselves easily into groups for which our data, as well as the intensity of heredity, are approximately the same, one series of these errors has been found for each group.

(4.) If we have ground for our assumption that the variable at the basis of our tint classification can be so selected as to give a normal distribution, we may determine the relative lengths on the scale of that variable occupied by each tint or shade. Thus if σ_1 be the standard deviation of the variable for male eye-colour, σ_2 for female eye-colour, I measured the range on the scale in terms of σ_1 and σ_2 for Mr. GALTON'S eight eye-colour tints for 3000 cases of male and 3000 cases of female eye-colour. I found the spaces occupied on the unknown scale to be as follows :—

No.	Tint.	Range in terms of σ_2.	Range in terms of σ_1.
1	Light blue	∞	∞
2	Blue, dark blue . . .	1·39276	1·34918
3	Grey, blue-green . . .	·73468	·77596
4	Dark grey, hazel . . .	·40027	·41992
5	Light brown	·03893	·00856
6	Brown	·43679	·35895
7	Dark brown.	·84161	·64167
8	Very dark brown, black.	∞	∞

These results are not so regular as we might have hoped for, on the assumption that the ratio of σ_1/σ_2 would be the same from whatever part of the scale it be determined. The general conclusion, however, would be that σ_1 is slightly larger than σ_2, which is confirmed by other investigations. Actually a tint may be rather vaguely described, and where the data were obtained by untrained observers without the assistance of a plate of eye-colours, a good deal of rather rough classification is likely to have taken place. I do not think it would be safe to go further than stating that on the quantitative colour scale the tints as described must occupy spaces in about the following proportions :—

Light Blue.	Blue, Dark Blue.	Grey, Blue-Green.	Dark Grey, Hazel.	Light Brown.	Brown.	Dark Brown.	Very dark Brown, Black.
∞	1·37	·75	·41	·02	·40	·74	∞

Taking 2000 colts and 2000 fillies, the standard deviations being σ_1 and σ_2 respectively, I have worked out the coat-colour ranges in terms of σ_2 and σ_1 for each of the sixteen colours* occurring in the records. We have the following results :—

---

* See p. 92, below.

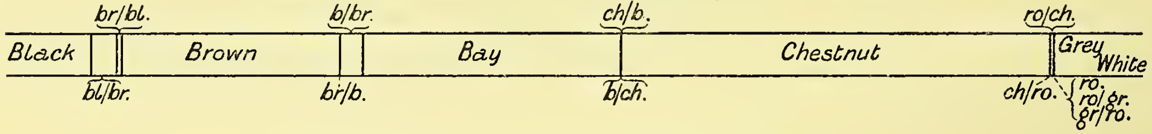
Tint.	Range in σ_2.	Range in σ_1.		Range in σ_2.	Range in σ_1.
1	∞	∞	9	·00000	·00000
2	·12683	·10768	10	1·96956	2·01658
3	·00000	·03313	11	·00000	·00000
4	·91747	1·11055	12	·02490	·00000
5	·00000	·00352	13	·00000	·00000
6	·11059	·10451	14	·00000	·00000
7	1·34684	1·27688	15	·00000	·00000
8	·00000	·00000	16	∞	∞

Here again it seems to me that the most we can safely do is to consider that on a suitable scale the relative lengths occupied by the classes of coat-colours recognised by thoroughbred horse breeders would be somewhat as follows :—

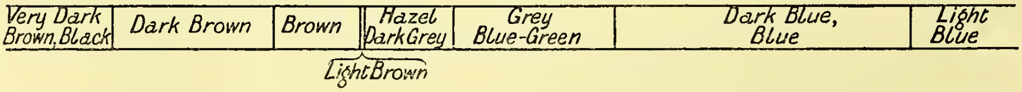
bl.	bl./br.	br./bl.	br.	br./b.	b./br.	b.	b./ch.	ch./b.	ch.	ch./ro.	ro./ch.	ro.	ro /gr.	gr./ro.	gr.
∞	·12	·02	1·01	·00	·11	1·31	·00	·00	1·99	·00	·01	·00	·00	·00	∞

The reader must carefully bear in mind that these represent scale-lengths occupied by the coat-colour and not the frequency of horses of these individual coat-colours. What we are to understand is this : that if eye-colour in man and coat-colour in horses were measured on such quantitative scales as we have given in skeleton, then the distribution of the frequency of the several colours would be very approximately normal. The actual skeleton scales are represented in the accompanying diagram, which puts them at once before the eye.

Normal Scale of Colour Ranges in Thoroughbred Horses.



Normal Scale of Eye Colour Ranges in Man.



(5.) It is necessary here to draw attention to a distinction of some importance in heredity, namely, that between *blended* and *exclusive* inheritance. In my treatment of the law of ancestral heredity,* it is assumed that we have to deal with a quantitatively measurable character, and that the ancestry contribute to the quantity of this character in certain proportions which on the average are fixed and follow certain definite numerical laws. Such an inheritance is *blended* inheritance. But another

---

* 'Roy. Soc. Proc.,' vol. 62, p. 386.

type of inheritance is possible. We may have one progenitor, prepotent over all others and absorbing all their shares, who hands down to the offspring not a proportion of his character, but the whole of it without blend. If this progenitor is a parent we have *exclusive* inheritance, if a higher ancestor a case of reversion. I have dealt at some length with this type of inheritance under the title of the Law of Reversion in another paper.* We must consider in outline the main features of such inheritance, for the cases of eye-colour in man and coat-colour in the horse approximate more closely to the numerical values required by it, than to those indicated by the law of ancestral heredity. The chief feature of exclusive inheritance is the absolute prepotency of one parent with regard to some organ or character. It need not always be the parent of the same sex, or the same parent throughout the same family. Some offspring may take absolutely after one, others after another parent for this or that organ or character only. I believe Mr. GALTON first drew attention, in his 'Natural Inheritance' (p. 139), to this exclusive or, as he terms it, *alternative* heritage in eye-colour. In going through his data again I have been extremely impressed by it; even those cases in which children might be described as a blend, rare as they are, are quite possibly the result of reversion rather than blending. If we suppose exclusive inheritance to be absolute, and there to be no blending or reversion, it is not hard to determine the laws of inheritance. Supposing the population stable, one-half the offspring of parentages with one parent of given eye-colour would be identical with that parent in eye-colour, the other half would regress to the general population mean, *i.e.*, the mean eye-colour of all parents. Hence, taken as a whole, the regression of children on the parent would be ·5. In the case of the grandparent the regression would be ·25; of a great grandparent ·125, and so on. With an uncle a quarter of the offspring of his brother will be identical in eye-colour with him, the other three-quarters will regress to the population mean, thus the regression will be ·25. If we have n brethren in a family, and take all possible pairs of fraternal relations out of it, there will be $\frac{1}{2}n(n - 1)$ such pairs; $\frac{1}{2}n$ brothers will have the same eye-colour that of one parent, the other $\frac{1}{2}n$ brother that of the other parent. Hence selecting any one brother, $\frac{1}{2}n - 1$ would have his eye-colour, and on the average $\frac{1}{2}n$ would have regressed to the mean of the general population. In other words, the coefficient of regression would be $(\frac{1}{2}n - 1)/(\frac{1}{2}n - 1 + \frac{1}{2}n) = (\frac{1}{2}n - 1)/(n - 1)$.

Accordingly $n = 3$ Regression $= \cdot 25$
$n = 4$,, $= \cdot 3333$
$n = 4\cdot 7$,, $= \cdot 3649$
$n = 5$,, $= \cdot 375$
$n = 5\cdot 3$,, $= \cdot 3833$
$n = 6$,, $= \cdot 4$
$n = \infty$,, $= \cdot 5$

---

* 'Roy. Soc. Proc.,' vol. 66, pp. 140 *et seq.*

It will at first appear, therefore, that the fraternal regression with the size of families actually occurring will vary from ·35 to ·4.

To some extent these values would be modified by assortative mating, which actually exists in the case of eye-colour. The correlations between parent and offspring and between brothers would both be slightly increased. Thus if ρ be the coefficient of assortative mating, R_f the fraternal correlation with and r_f without assortative mating, and r the coefficient for parent and offspring,[*]

$$R_f = \frac{r_f + 2\rho r^2}{1 - 2\rho r^2}.$$

If we put $r_f = ·36$, $r = ·5$, $\rho = ·1$, we find

$$R_f = ·39.$$

Thus we see that the regression or correlation for fraternal inheritance in the case of exclusive inheritance could not, with the average size of families, be very far from ·4 of blended inheritance.

A further source which can modify immensely, however, the fraternal correlation is the prepotency of one or other parent, not universally, but within the individual family. In the extreme case all the offspring might be alike in each individual family. Thus fraternal correlation might be perfect although parental correlation were no greater than ·5. Hence, where for small families we get a fraternal correlation greater than ·4 to ·5, it is highly probable that there exists either a sex prepotency (in this case, one of the parental correlations will be considerably greater than the other) or an individual prepotency (in which case the parental correlations based on the average may be equal). We shall see that fraternal correlations occur greater than ·5 in our present investigations. I have dealt with these points in my Memoir on the 'Law of Reversion,'[†] and also in the second edition of the 'Grammar of Science.'[‡]

Another point also deserves notice, namely, that with the series ·5, ·25, ·125, &c., for the ancestral coefficients in the direct line, the theorems proved in my Memoir on Regression, Heredity, and Panmixia[§] for the series of coefficients $r, r^2, r^3 \ldots$ exactly hold. That is to say, if we have absolutely exclusive inheritance, the partial regression coefficients for direct ancestry are all zero except in the case of the parents. This it will be observed is not in agreement with Mr. GALTON'S views as expressed in Chapter VIII. of the 'Natural Inheritance.' But I do not see how it is possible to treat exclusive inheritance on the hypothesis that the parental regression is about ·3.[||] Actual investigation shows that for this very character, *i.e.*, eye-colour, it is nearer ·5. If we take Table XIX. of Mr. GALTON'S appendix, and make the following groups, both

---

[*] This is shown in a paper not yet published on the influence of selection on correlation.

[†] 'Roy. Soc. Proc.,' vol. 66, pp. 140 *et seq.*.

[‡] "On Prepotency," p. 459; "On Exclusive Inheritance," p. 486.

[§] 'Phil. Trans.,' A, vol. 187, p. 302, etc.

[||] Mr. GALTON takes $\frac{1}{3}$ throughout his arithmetic.

parents light, one parent light and one medium, one light and one dark, we reach the following results :—

Parents' eye-colour.	Children, actual.		Light-eyed children, calculated.	
	Total.	Light-eyed.	Exclusive inheritance.	Ancestral law with knowledge of parents and grandparents.
Both light	355	334	355	321
Light and medium .	215	170	161	160
Light and dark . .	211	107	105	117

Here the exclusive inheritance leads us to misplace thirty-two and the ancestral law thirty-three children. The evidence, therefore, of the correctness of the latter is hardly greater than that of the former. Indeed, if the former were modified for reversion, it would very possibly give better results than the latter.

I am inclined accordingly to look upon eye-colour inheritance as an exclusive inheritance modified by reversion, and, to some extent, by assortative mating, rather than a mixture of exclusive inheritance with a slight amount of blending. In either case exclusive inheritance gives results like the above so closely in accord with the ancestral law that the latter might be supposed to hold. But, theoretically, I do not understand how the ancestral law is compatible with exclusive inheritance. Theoretically, we have values of parental, avuncular, and grand-parental correlation greater than the ancestral law would permit of, and these theoretical values are, on the whole, closer to observation, as we shall see in the sequel, than those given by the law of ancestral heredity. The following table gives the two systems :—

TABLE I.—Theoretical Values of the Regression Coefficients.

Relationship.	Blended inheritance, ancestral law.	Exclusive inheritance, absolute, no reversion.
Parent and offspring	·3	·5
Grandparent and offspring . . .	·15	·25
Great-grandparent and offspring .	·075	·125
Brethren	·4	·35 to ·5*
Uncle and nephew	·15	·25

Now, if exclusive inheritance be modified by reversion or assortative mating, or if blended inheritance be modified by "taxation,"† then we shall get values different

* This varies with the size of the family.

† 'Roy. Soc. Proc.,' vol. 62, p. 402.

from the above, and possibly filling up the numerical gap between them. To this point I shall return when dealing with the observed values for eye-colour in man.

Part II.—On Colour-Inheritance in Thoroughbred Racehorses.

(6.) All the data were extracted from Weatherby's stud-books, the colours being those of the horses as yearlings. My first step was to form an order, not a quantitative scale, of horse-colours. With this end in view, the recorded colours were examined, and, including the arabs, the following seventeen colours were at first found :—

1. Black (bl.).
2. Black or brown (bl./br.).
3. Brown or black (br./bl.).
4. Brown (br.).
5. Brown or bay (br./b.).
6. Bay or brown (b./br.).
7. Bay (b.).
8. Bay or chestnut (b./ch.).
9. Chestnut or bay (ch./b.).
10. Chestnut (ch.).
11. Chestnut or roan (ch./ro.).
12. Roan or chestnut (ro./ch.).
13. Roan (ro.).
14. Roan or grey (ro./g.).
15. Grey or roan (g./ro.).
16. Grey (g.).
17. White (w.).

Now, if we take the alternative colours to mean that the first alternative is the prominent element, we see that these colours in use among breeders admit of *only one arrangement from black to white.* That is to say, that a continuous shade-change is actually in use for the colour-nomenclature of thoroughbred horses.* Thus without any hypothesis as to the quantitative relative values of bay or roan, we have an order which serves for all our present purposes. Following this order, Appendix I., Tables I.—XII., for the colour correlation of related pairs of horses was compiled by Mr. Leslie Bramley-Moore from the stud-books. When dealing with relationship in the ♀ line only, no weight has been given to fertility, as each mare has had only one foal attributed to it, or two in the case of fraternal correlation. In the case of the ♂ line, the colours of the older sires were harder to ascertain, and we did not obtain altogether more than 600 sire-colours. Thus one, two, or, in a few cases, three or four colts or fillies were taken from each sire.

I shall now discuss the results which may be drawn from these tables for the theory of heredity, first placing in a single table all the numerical constants calculated from the data in Tables I. to XII. of Appendix I.

---

* Among the 6000–8000 horses dealt with only *four* were found with colours not entered in this scale, but these entries of bl./ch., br./ch., b./ro., in no way contradict the *order* of the scale, but merely show a rougher appreciation on the part of the nomenclator, or possibly printers' or editor's errors.

TABLE II.—Coat-colour Inheritance in Thoroughbred Horses.

Pair of relatives.		Division of bay range by the mean		Ratio of variabilities.		Coefficients of correlation.	Coefficients of regression.		Number of cases.
x.	y.	η_x.	η_y.	$\zeta = \sigma_x/\sigma_y$.	$\zeta = \sigma_y/\sigma_x$.	r_{xy}.	R_{xy}.	R_{yx}.	N.
Sire.	Colt. .	·6111	·5713	·8712	1·1478	·4913	·4280	·5639	1300
Sire.	Filly .	·6061	·5719	·8298	1·2051	·5422	·4499	·6534	1050
Dam	Colt. .	·5359	·6027	·9500	1·0526	·4862	·4619	·5118	1000
Dam	Filly .	·5565	·6051	·9036	1·1067	·5668	·5122	·6273	1000
Maternal grandsire	Colt. .	·6583	·5867	·7030	1·4225	·3590	·2524	·5107	1000
Maternal grandsire	Filly .	·6359	·6042	·7678	1·3024	·3116	·2392	·4058	1000
Colt . . . . . . (Half	Colt. . siblings)	·5908	·5908	1	1	·3551	·3551	·3551	2000
Colt . . . . . . (Whole	Colt. . siblings)	·5620	·5620	1	1	·6232	·6232	·6232	2000
Filly . . . . . . (Half	Filly . siblings)	·5665	·5665	1	1	·4265	·4265	·4265	2000
Filly . . . . . . (Whole	Filly . siblings)	·5684	·5684	1	1	·6928	·6928	·6928	2000
Filly . . . . . . (Half	Colt. . siblings)	·5633	·5865	·9607	1·0409	·2834	·2723	·2950	1000
Filly . . . . . . (Whole	Colt. . siblings)	·5410	·5711	·9555	1·0466	·5827	·5568	1·0466	1000

In this table $R_{xy} = r_{xy}\sigma_x/\sigma_y$, $R_{yx} = r_{xy}\sigma_y/\sigma_x$. Half-siblings* are those having the same dam, but different sires. Further, η is measured from the brown end of the bay range up to the mean.

(7.) *On the Mean Coat-Colour of Horses.*—If our theory be correct, this colour will not differ much from the median colour, and we notice at once a secular change going on. We have the following order:—

Maternal grandsire of colt	$\eta =$ ·6583
Maternal grandsire of filly	= ·6359
Sire of colt	= ·6111
Sire of filly	= ·6061
Colt (mean value of seven series) . . .	= ·5816
Dam of colts	$\eta =$ ·5359
Dam of fillies	= ·5565
Fillies (mean value of seven series) . .	= ·5753

---

[*] I have introduced this expression in my paper on "The Law of Reversion," 'Roy. Soc. Proc.,' vol. 66, as a convenient expression for a pair of offspring from same parents whatever be their sex.

Now the colours of all the horses are returned when they are foals, so that there is no question of any variation of colour with age, yet we notice that—

(i.) The horse is lighter in colour than the mare.

(ii.) If we go back two generations (grandsire) the horse is lighter than if we only go back one generation (sire), and the sires are again lighter than their colts. In other words, there seems a progressive change towards a darker coat.

(iii.) On the other hand, the mares one generation back appear to be darker than their daughters.

(iv.) The average sire of colts is lighter than the average sire of fillies; the average dam of colts is darker than the average dam of fillies.

Now these conclusions seem to indicate that the older horse was lighter in coat, and the older mare darker in coat than either the colt or filly of to-day, and that there is a tendency in the thoroughbred racehorse of to-day to approach to an equality of colour in the two sexes, an equality which is a blend of the sensibly divergent sex-colour of the older generation.

Whether this secular change is due to the "breeding out" of the influence of light Arabian sires, or to a tendency in the past to select light-coloured sires and dark-coloured mares for breeding, or to the fact that such coloured sires and mares are the most fertile, *i.e.*, to an indirect effect of reproductive selection, is not so easy to determine. But what does appear certain is that the average thoroughbred is approaching to a blend between its male and female ancestry, which were sensibly divergent.*

(8.) *On the Relative Variability of Sex and Class in Horses.*—The following table gives the length of the bay range in terms of the standard deviation for each class. If ϵ represent this range, then in terms of the previous notation $\epsilon = u \times \sigma = u' \times \sigma'$, and from these values of u and u' the ratio, $\zeta = \sigma/\sigma'$ of Table II. was calculated.

---

* Mean of dams and sires of colts = ·5735, *i.e.*, $\frac{1}{2}$(·6111 + ·5359).
Mean of dams and sires of fillies = ·5813, *i.e.*, $\frac{1}{2}$(·6061 + ·5565).

These are curiously enough almost exactly equal to the mean values ·5753 and ·5816 obtained for fillies and colts. This *inverse* relationship is too close to the probable errors of the quantities under investigation for real stress to be laid on it, but it may still turn out to be suggestive.

III.—Table of Bay Ranges.

Relative Pair.		Bay Range.		Probable Error of Median.	
x.	y.	$u \times \sigma_x$.	$u' \times \sigma_y$.	x.	y.
Sire	Colt	1·46943σ_x	1·28019σ_y	±·0160	±·0183
Sire	Filly	1·64075σ_x	1·36149σ_y	±·0159	±·0192
Dam	Colt	1·36645σ_x	1·29819σ_y	±·0196	±·0206
Dam	Filly	1·38165σ_x	1·24845σ_y	±·0193	±·0214
Maternal grandsire .	Colt	1·69694σ_x	1·19293σ_y	±·0158	±·0224
Maternal grandsire .	Filly	1·65021σ_x	1·26702σ_y	±·0162	±·0211
Colt . . . . (Half	Colt . . . . siblings)	1·23953σ_x	1·23953σ_y	±·0153	±·0153
Colt . . . . (Whole	Colt . . . . siblings)	1·27688σ_x	1·27688σ_y	±·0148	±·0148
Filly . . . . (Half	Filly . . siblings)	1·39619σ_x	1·39619σ_y	±·0135	±·0135
Filly . . . . (Whole	Filly . . . . siblings)	1·34684σ_x	1·34684σ_y	±·0140	±·0140
Filly . . . . (Half	Colt . . . . siblings)	1·33479σ_x	1·28229σ_y	±·0202	±·0208
Filly . . . . (Whole	Colt . . . . siblings)	1·41501σ_x	1·35207σ_y	±·0189	±·0198

To explain the last double column I note that Mr. SHEPPARD has shown (' Phil. Trans.,' A, vol. 192, p. 134) that the probable error of the median equals

$$\cdot 84535\ \sigma/\sqrt{\mathrm{N}}.$$

Hence in terms of the bay range we have

$$\frac{\text{probable error of median}}{\text{length of bay range}} = \cdot 84535/(u\sqrt{\mathrm{N}}).$$

I have found that this simple result gives a value close to but slightly larger than the probable error of the quantity η (p. 82), from which I have determined the position of the mean in the bay range. It is much easier to calculate, but of course not so exact, as we take no account of possible errors in the bay range itself or their correlation with errors in the median. I have accordingly tabulated its values in the last double column as a rough guide to the errors made in the numbers upon which the statements in the previous section depend. I shall return to the consideration of the probable errors below. Turning to columns 3 and 4 of our Table II., we can draw the following conclusions as to the variability of sex and class:—

(*a*.) *The Younger Generation is more Variable than the Old.*—Thus, foals are more variable than their sires, and, looking at the expressions in Table III. for the bay range, sires than grandsires. This is a rule I have now found true in a very great number of cases of inheritance. Parents are a fairly closely selected body of the general population, and so less variable than that population at large. This might appear pretty obvious in the case of thoroughbred horses when we are dealing with sires and grandsires. I have already pointed out that it was impossible to take 1000 to 1300 colts or fillies with as many independent sires, the fashion in sires is too marked; and of course the number of independent grandsires was still fewer.* But even in the case of dams, where we have taken as many independent dams as fillies, we see this reduction in variability in the older generation. As it also occurs with stature, &c., in man as well as with coat-colour in horses—in which latter case we expect artificial selection—it deserves special consideration. Without weighting with fertility, there exists a selection of the individuals destined to be parents in each generation. We have to ask whether the change in mean and variability from parent to offspring in each generation is secular or periodic, or to what extent it is partly one and partly the other. The importance of settling this point is very great; it concerns the stability of races. So far as my fairly numerous series of measurements yet go, I cannot say that a "stable population" has definitely shown itself; the characters of each race when measured for two generations seem to vary both in mean and standard deviation. Palæontologists tell us of species that have remained stable for thousands of years, but this is a judgment hitherto based on a *qualitative* appreciation. A quantitative comparison of the means, variabilities, and correlations of some living species in its present and its fossil representatives would be of the greatest interest and value. For myself, I must confess that my numerical investigations so far tend to impress me with the unstable character of most populations.

(*b*.) *There is fairly good evidence that the horse is more variable than the mare in coat-colour.* It would be idle to argue from the first four results of Table III. that the mare is more variable than the horse, in that these results show the dam to be more variable than the sire. For, as we have shown, the process of breeding and our method of extracting the data tend to produce a much more intense selection of sires than of dams. But if we compare the mean bay range in terms of the standard deviation of colts for our seven series of colts with that for the seven series of fillies in Table III., we find for the first $1\cdot27458\,\sigma_c$ and for the second $1\cdot33854\,\sigma_f$. Hence we are justified in concluding that σ_c is greater than σ_f. In fact in only *one* case out of the seven does the series of fillies give a less variability than the corresponding series of colts, *i.e.*, colts corresponding to dams are somewhat less variable than fillies corresponding to dams. It must, however, be remembered that this conclusion is based upon the coat-colour of the animals recorded as yearling foals.† Thus, the coat-

---

* For some account of the extent of in and in breeding in the thoroughbred horse, see my memoir on "Reproductive Selection," 'Phil. Trans.,' A, vol. 192, p. 257 *et seq.*

† The reader must always bear in mind that when we speak of the variability of colour in sire or dam, &c., it means the variability of this class when they were yearlings.

colour may change both in intensity and variability with age, much as variability in stature changes with children from birth to adult life.

(*c.*) As a more or less natural result of (*b*) it follows *that any group, male or female, having male relatives is more variable than the same group with female relatives.* Thus sires of colts are more variable than sires of fillies; fillies half-sisters to colts are more variable than fillies half-sisters to fillies, &c. But out of the nine cases provided by our data there are three exceptions to the rule, and perhaps not much stress can be laid on it, at any rate in the above form. It would appear that males, relatives of males, are sensibly more variable than males relatives of females. The bay ranges are 1·3926 σ' and 1·4447 σ respectively, which indicates that the average σ' is larger than σ. But if we treat the groups of females alone, we find for females with male relatives the bay range = 1·3694 σ, and for females with female relatives 1·3433 σ', which indicates that the latter are more variable. The difference is, however, not very sensible. Possibly the rule is simply that extremes tend to produce their own sex, but our data are not sufficient for really definite conclusions on the point.

In order that we may have a fair appreciation of the probable errors of the quantities involved and the weight that is to be laid upon their differences, I place here a table* of the probable errors of η, of $\zeta = \sigma_x/\sigma_y$ and of r_{xy} for typical cases.

IV.—Table of Probable Errors.

Relations.	η_x	η_y	ζ	u	u'	r_{xy}
Sire and Filly . .	·0143	·0170	·0243	·0363	·0330	·0288
Grandsire and Colt	·0143	·0199	·0237	·0385	·0319	·0333
Colt and Colt . . (Whole siblings)	·0186	·0186	—	·0328	·0328	·0259
Filly and Colt . . (Half siblings)	·0179	·0185	·0315	·0335	·0328	·0363

It will be seen from this table that the probable error in η is about 3 per cent., in ζ about 2 to 4 per cent., in u about 2 to 2·5 per cent., and the values of r about ·03, growing somewhat larger as r grows smaller. The probable errors are thus somewhat larger than those which we obtain by the old processes when the characters are capable of quantitative measurement, but they are not so large as to seriously affect the use of the new processes in biological investigations. As we have already indicated, the probable errors of the η's may be roughly judged by Mr. SHEPPARD's formula for the median (p. 95).

It will be seen that the differences in the η's and ζ's of Table II., or the u's of

---

* I have to thank Mr. W. R. MACDONELL for friendly aid in the somewhat laborious arithmetic involved in calculating these probable errors.

Table III., are as a rule larger than the probable errors of the differences, sometimes several times larger. Yet in some cases they are not such large multiples of the probable errors of the differences that we can afford to lay great stress on the divergence of η or ζ or u in a pair of special cases. We must lay weight rather on the general tendency of the tables when all the series are taken together. Thus, while we may have small doubt about the correctness of (i.) of § 7 or (*b*) of § 8, we should look upon (iv.) of § 7 as an important suggestion which deserves serious consideration rather than a demonstrated law. The same again holds good for (*c*) of § 8. It is because of their suggestiveness that they are here included. That a differential fertility or an individualisation in the sex of offspring should be correlated with colour, would, if proved, be a fact of very considerable interest. It would again emphasise the important part which genetic selection plays in the modification of characters.* *À priori* it is not more unreasonable to expect coat-colour in horses than to suppose hair-colour in men to be correlated with fertility. But the fertility of man does seem to vary from the light to the dark races. The special feature of interest here, however, is that a different colour in the two sexes appears to influence the preponderance of one or other sex in the offspring. It would be an interesting inquiry to determine whether the sex-ratio in the offspring of "mixed marriages" varies when the races of the two parents are interchanged.

(9.) *On the Inheritance of Coat-colour in Thoroughbred Horses.*—(*a.*) *Direct Line. First Degree.*—Having regard to the probable errors—about ·03—in the values of the correlation coefficient r_{xy}, it seems quite reasonable to suppose that the mean parental correlation, actually ·5216, is practically ·5. It is quite impossible to imagine it the ·3 of Mr. GALTON's view of the Law of Ancestral Heredity. If we adopt the view of that law given in my paper on the Law of Ancestral Heredity,† and take the coefficient γ to be different from unity, then it is shown in my paper on the Law of Reversion‡ that we cannot on the theory of blended inheritance get parental correlation as high as ·5 without values of the fraternal correlation which are much higher than those hitherto observed, certainly much higher than, as we shall see later, we find in the case of coat-colour in horses. Coat-colour in horses does not thus appear to be at all in accord with Mr. GALTON's view of ancestral inheritance, or even with my generalised form of his theory. It does accord very well with what we might expect on the theory of exclusive inheritance as developed above, p. 91, on the assumption that there is no reversion.

Looking at the matter from the relative standpoint, we see that not much stress can be laid on the respective influences of the sire and dam on the colt, or of the sire and dam on the filly; but, on the other hand, the filly does appear to inherit more from

---

* See the concluding remarks in the memoir on "Genetic (Reproductive) Selection," 'Phil. Trans.,' A, vol. 192, pp. 257—330.

† 'Roy. Soc. Proc.,' vol. 62, p. 386 *et seq.*

‡ 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.*

both parents than the colt does. There is certainly (judged from coat-colour) no preponderance of the sire's influence over the dam's such as breeders appear occasionally to imagine. The average influence of the dam on the offspring indeed appears to be slightly greater than that of the sire, but the difference is of the order of the probable error, and not of the overwhelming character exhibited in the case of Basset Hounds. There is indeed in the case of thoroughbred horses not the same chance of carelessness produced by a misalliance afterwards screened by the defaulter. There exists, however, a far greater premium—considering the great value of yearlings from fashionable sires—set upon dishonesty. Again it is possible that when stallions receive too many public or private mares, or are still used in their old age, that they may, without losing the power of fertilising, lose some of the power of transmitting their characters. The divergences, so far as the probable errors are concerned, are not such that we are forced out of our way to explain them. With the single exception of sire and colt we see that our table shows the universal prevalence of the rule that:

Relatives of the same sex are more closely correlated than relatives of the same grades of the opposite sex. Thus :—

A colt is more like his sire than his dam.
A filly is more like her dam than her sire.
A dam is more like her filly than her colt.
A grandsire is more like his grand-colt than his grand-filly.
A colt is more like his brother colt than his sister filly.
A filly is more like her sister filly than her brother colt.

the latter two cases being true for both whole and half siblings.

The solitary exception is that a sire is more like his filly than his colt.

If we were to assume it a rule that a filly in the matter of coat-colour has stronger inheritance all round than a colt, we should find it agree with our results for parental inheritance, and receive considerable support for the much stronger correlation of fillies than of colts, when either whole or half siblings. But it would not be in accordance with our results for grandparents, for which, however, we have only details for two out of the eight possible cases. On the whole, I think we must content ourselves with the statements that parental correlation is certainly about ·5, and that with high probability each sex is more closely correlated with its own sex of the same grade of relationship.

(*b.*) *Direct Line, Second Degree.*—My data here are unfortunately only for two cases out of the possible eight. I hope some day to finish the series, but the labour of ascertaining from the studbooks the coat-colour of 700 or 800 separate sires is very great. Indeed it is not easy to follow up the pedigree through the male line when the sire is not one of the famous few. On the other hand, it is much easier through the female line. For this reason the *maternal* grandsire was taken. Even in this case we had to seek back for each sire—the year of whose birth was unknown—until we found the record of his coat-colour given under the heading of his dam in the year of his birth.

The average of our two cases gives a coefficient of correlation = ·3353, the colt having a greater degree of resemblance to the grandsire than the filly. This value is substantially greater than the ·25 we might expect for exclusive inheritance, and more than double the value ·15, to be expected for the grandparental correlation with Mr. GALTON's unmodified law for blended inheritance. Of course the ·25 is to be expected as the mean of the *eight* grandparental series, and, as we shall see for eye-colour in man, these may vary very much in magnitude. But allowing for this, it seems quite impossible that the average value could be reduced to ·15. I take it therefore that the grandparental, like the parental, data point to a law of inheritance other than that which has been described in my paper on the Law of Ancestral Heredity. This peculiar strengthening of the grandparental heritage has already been noted by me in my paper on the Law of Reversion,* and the difficulties of dealing with it on the principle of reversion therein discussed. There may be some opinion among breeders as to the desirability of emphasising the dam's strain in the choice of a sire which leads to this result, but if so it is unknown to me, nor do I see how it would work without also emphasising the correlation of the dam and foal. The mean value of the correlation for the maternal grandfather and grandchildren for eye-colour in man is ·3343—a result in capital agreement with that for coat-colour in horses. In that case the average of the eight series, as we shall see later, is considerably above ·25, and we must, I think, suspend our judgment as to whether this could possibly in the case of horses be the true mean value. As to the value ·15 it seems quite out of the question.

As already remarked, the influence of the maternal grandsire (unlike that of the sire) is substantially greater on the colt than on the filly.

(*c*.) *Collateral Heredity, First Degree.*—Here we have more ample data to go upon, namely, a complete set of six tables of both whole and half siblings of both sexes.

We notice one or two remarkable features straight off. In the first place, in the case of both fillies and colts, the whole siblings of the *same* sex have not a correlation the double of that of the half siblings, but have a correlation very considerably less than this. *À priori* we might very reasonably expect the one to be the double of the other. This is what would happen in the case of blended inheritance on the hypothesis of equipotency of the parents. As the half siblings are on the dam's side, we might assert a considerable prepotency of the dam over the sire. This cannot indeed be the explanation of the divergence in the case of Basset Hounds, where the half siblings have a correlation considerably *less* than half that of whole siblings,†

---

* 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.*

† 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.*

and yet the prepotency of the dam in coat-colour is very marked. But in the present case there is on the average only a slight, if indeed it be a real, prepotency of the dam. Further, if we turn to the correlation, no longer of siblings of the same sex, but of opposite sexes, we find the correlation of the whole siblings is approximately double that of the half siblings, as we should *à priori* have expected.

Taking averages on the assumption that the correlation for whole siblings should be double that for half siblings, we have the following results :—

Correlation between colts based on results for whole and half siblings	·6667
Correlation between fillies based on results for whole and half siblings	·7729
Correlation between filly and colt based on results for whole and half siblings	·5747
Mean correlation of siblings based upon all results for whole siblings	·6329
Mean correlation of siblings based upon all results for half siblings	·7100
Mean correlation of siblings based upon results for both whole and half siblings	·6714

We can draw the following conclusions :—

(i.) In whatever manner we deduce the fraternal correlation it is very much larger than the ·4 for whole brethren, or the ·2 for half brethren, required by the unmodified Galtonian law. Such values, as we see above, could be deduced from the modified Galtonian law by taking γ greater than unity,* but this would involve values for the parental correlation sensibly less than those given by theory. We are again compelled to assert that the modified as well as the unmodified theory of blended inheritance, based on the Galtonian law, does not fit the facts. The above values, however, are quite compatible with the theory of exclusive inheritance on the assumption that there is an individual (not a sexual) prepotency from one pairing to another.

(ii.) In whatever way we consider it, it would appear that the average value of the fraternal correlation, as deduced from siblings with the same dam only, is greater than that deduced from siblings with both the same dam and the same sire.

I am not able to explain this in any way, for I cannot assert a very substantial prepotency of the dam. All I can say from the data at present available is that for horses and dogs there appears to be no simple numerical relation between the correlation of whole and half brethren.

(iii.) Offspring of the same sex are more alike than offspring of opposite sexes.

This appears to be generally true, so far as our data at present extend, and will be fairly manifest from the table below.

---

* 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.*

TABLE V.—Collateral Heredity.

Pair.	Man.			Dog.	Horse.	
					Coat-colour.	
	Stature.[*]	Cephalic Index.[†]	Eye-colour.[‡]	Coat-colour.[§]	Whole Siblings.	Half Siblings.
Brother-Brother .	·3913	·3790	·5169	·5257	·6232	·3551
Sister-Sister . .	·4436	·4890	·4463		·6928	·4265
Brother-Sister .	·3754	·3400	·4615		·5827	·2834

It will be noted that, with the single exception of eye-colour in man, sister and sister are more alike than brother and brother.

The mean value of the fraternal correlation for stature is ·4034, and for cephalic index ·4027. These results are in excellent accordance with the ·4 required by the Galtonian theory of blended inheritance. The mean values for eye-colour in man, coat-colour in dogs, and coat-colour in horses are: ·4749, ·5257, and ·6329. These are quite incompatible with that theory. I venture accordingly to suggest that we have here cases of an inheritance which does not blend, and that it is an inheritance which is far more closely described by the numbers we have obtained on the theory before developed of exclusive inheritance than by the law of ancestral heredity.

Taking in conjunction with these results for collateral heredity, those for parental and grandparental inheritance, we see that coat-colour in horses diverges widely from the laws which have been found sufficient in the cases of stature and cephalic index in man. The latter characters are really based on a complex system of parts, while the determination of coat-colour may depend on a simple or even single factor in the plasmic mechanism. Here Mr. GALTON'S suggestion of an exclusive inheritance of separate parts ('Natural Inheritance,' p. 139) may enable us to understand why stature and cephalic index differ so widely in their laws of inheritance from coat- and eye-colours.

PART III.—ON THE INHERITANCE OF EYE-COLOUR IN MAN.

(10.) *On the Extraction and Reduction of the Data.*—The eye-colour data used in this memoir were most generously placed at my disposal by Mr. FRANCIS GALTON. They are contained in a manuscript wherein, by a simple notation, we can see at a

---

[*] PEARSON, 'Phil. Trans.,' A, vol. 187, p. 253 *et seq.* See Note I. at the end of this paper.

[†] FAWCETT and PEARSON, 'Roy. Soc. Proc.,' vol. 62, p. 413 *et seq.*

[‡] Present memoir, p. 113 *et seq.*

[§] PEARSON, 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.*

glance the distribution in eye-colour of a whole family in its numerous male and female lines. Such complete details of the various direct and collateral relationships I have not hitherto come across, and from them I was able to form, in the course of some months of work, the twenty-four tables of correlation which are given in Appendix II. These tables, for the first time, give the whole eight series of grand-parental and the whole eight series of avuncular relationships, besides such as we have deduced for other characters previously, *i.e.*, the four parental, the three fraternal relationships, and the table for assortative mating. The very great importance of this material will at once be obvious, and I cannot sufficiently express my gratitude to Mr. GALTON for allowing me to make free use of his valuable data.

At the same time we must pay due regard to the limitations of this material, which it is needful to enumerate, so that too great stress may not be laid on the irregularities and divergences which arise when we attempt to reduce the results to laws. These limitations are as follows :—

(*a.*) While the data of about 780 marriages are given in the record, they belong to less than 150 separate families. All our relationships, therefore, contain pairs weighted with the fertility of the individual families. Thus it was necessary to enter every child of a mother, every nephew of an uncle, and so forth. In the horse data we could take 1000 distinct mares and give to each one foal only. That is not possible in the present case.

(*b.*) The colour of eyes alters considerably with age. It is not clear that some of the eye-colours are not given for infants under twelve months, and certainly the eye-colours in the case of grandparents and others must have been taken in old, perhaps extreme old, age. In a large number of other cases of great grandfather, great great-grandfather, &c., great uncles, and so forth, the eye-colours must have been given from memory or taken from portraits—in neither alternative very trustworthy sources. Thus while the horse colour is always given for the yearling foal by the breeder, the eye-colour is given at very different ages, and comes through a variety of channels.

(*c.*) The personal equation in the statement of eye-colour, when the scale contains only a list of tint-names is, I think, very considerable. The issue for the collection of data of a plate of eye-colours like that of BERTRAND would be helpful, but we can hardly hope for a collection of family eye-colours so comprehensive as Mr. GALTON'S to be again made for a long time to come.

These causes seem to me to account for a good deal of the irregularity which appears in the numerical reduction of the results, but they are not, I hold, sufficient to largely impair the very great value of Mr. GALTON'S material.

In tabulating the data, I have followed the scale of tints adopted by Mr. GALTON, and I have used the entire material available in the cases of the grandparental, avuncular, and marital relations. I nearly exhausted the data for the parental relationships, but in these tables, which were first prepared, I stopped short at 1000 for the sake of whole numbers. I found, however, that it did not make the arithmetic sensibly shorter, and I afterwards dropped this limitation. In the case of brethren I took 1500 of each case—I daresay I could have got 2000 out of the records. As the light-eyed brethren are entered *first* in Mr. GALTON'S MS., the First Brother in my unsymmetrical tables is always lighter-eyed than the Second Brother, hence the tables had to be rendered symmetrical by interchanging and adding rows and columns before we could reduce them. Thus the symmetrical tables have an apparent entry of 3000 pairs. Of course 1500 is the number used in finding the probable error of the correlation coefficient. The like difficulty does not occur in the brother-sister table, where indeed the difference of mean eye-colour for the two sexes would not allow of our making the table symmetrical. A comparison of the symmetrical with unsymmetrical tables for colts-colts and fillies-fillies, will show how little need there is for rendering the tables symmetrical when pairs are taken out of any similar class and tabulated without regard to the relative magnitude of the character in the two individuals of the pair, *i.e.*, WEATHERBY'S record places the individuals simply in order of birth and not of darkness or lightness of coat-colour.

Table VII. gives the value of the chief numerical constants deduced from the twenty-four eye-colour tables in Appendix II.*

(11.) *On the Mean Eye-colour having regard to Sex and Generation.*—In order to test the degree of weight to be given to our conclusions, I have drawn up a table of probable errors for four typical cases—cases by no means selected to give the lowest possible values. Further, in Table VIII. I have given the probable error in the position of the median as determined in terms of the grey, blue-green range by the modification of Mr. SHEPPARD'S formula (see p. 95). The grey, blue-green range of eye-colour is about one-fifth of the total observed range, so that the probable error in the position of the median varies from about ·4 to 1 per cent. of that range. This is not a large error, but, relative to the small variations of value with generation and sex, it is sensible, and we must not draw too fine conclusions on the basis of single inequalities.

TABLE VI.—Table of Probable Errors in Eye-colour Data.

Relations.	η_x	η_y	ζ	u	u'	r_{xy}
Mother and Son . .	·0253	·0188	·0431	·0267	·0256	·0283
Maternal Grandmother and Granddaughter .	·0348	·0350	·0767	·0276	·0314	·0361
Sister and Sister . .	·0244	·0244	—	·0216	·0216	·0234
Maternal Aunt and Nephew	·0230	·0186	·0414	·0255	·0250	·0302

---

* The theoretical formulæ by aid of which these constants were determined, have been indicated in the earlier part of this memoir, and in Part VII. of the present series on Evolution. The actual work of reduction has been most laborious, but I trust that our results are free from serious error.

If we examine this table we see that the error in η amounts to from ·02 to ·025 when we have upwards of 1000 tabulated cases, but can amount to ·035 when we have as few as 700 to 750 tabulated cases. An examination of the values of η in Table VII. shows us that most of our differences with probable errors taken on this scale are very sensible. A comparison with Table VIII. shows us that the probable error of the median is always greater than the probable error of η, and accordingly the former, being much easier of calculation, may be taken as a maximum limit. The probable errors of ζ, *i.e.*, the ratio of σ_x to σ_y, are more considerable, amounting to about ·04 for our longer series, and even to ·077 in the case of the short series of grandmother and granddaughter, but in this case ζ actually takes its maximum value of 1·291, so that the error is under 6 per cent. ; in the longer series it is under 5 per cent. Again, we see that in most cases our differences in the ratio of variabilities are quite sensible. It must be admitted, however, that the ratio of variabilities as based on the grey blue-green range of eye-colour is not as satisfactory as that based on the bay range of coat-colour in horses. In the latter case, one-half of the horses fall into the bay range, but only about a quarter of mankind fall into the grey blue-green range of eye-colour, and, further, the appreciation of eye-colour seems to me by no means so satisfactory as that of coat-colour in horses.

I have tried a further series of values for the ratio of the variabilities by measuring the ranges occupied not only by the tints grey blue-green, but by the whole range of tints 3, 4, 5, and 6 of Mr. GALTON'S classification (see p. 87). Lastly, I have taken a third method of appreciating the relative variabilities, namely, by using the method of column and row excesses, E_1 and E_2, discussed in Part VII. of this series. While this method has the advantage of using all and not part of the observations to determine the ratio of σ_x to σ_y, and so naturally agrees better with the results based on the four than the one tint ranges, it suffers from the evil that these excesses can only be found by interpolation methods, which are not very satisfactory when our classes are, as in this case, so few and so disproportionate. On the whole, this investigation of relative variability is the least satisfactory part of our eye-colour inquiry, and I attribute this to two sources :—

(i.) The vagueness in appreciation of eye-colour when no colour scale accompanies the directions for observation (*cf.* p. 103, (*c*)).

(ii.) A possible deviation from true normality in the factor upon which eye-colour really depends (*cf.* p. 80, (ii) 80).

Lastly, we may note that the probable error in the correlation amounts in most cases to less than ·03, rising only somewhat above this value for grandparental inheritance, where our series are somewhat short—650 to 750 instead of 1000. Here again most of the divergences are quite sensible.

Allowing accordingly for the comparative largeness of our probable errors, we shall do best to base conclusions on the general average of series ; to insist on general inequalities rather than on exact quantitative differences, and to emphasise the

TABLE VII.—Eye-colour Inheritance in Man.

Pair of relatives.		Division of grey blue-green range by mean.		Ratio of variabilities.		Coefficient of correlation.	Coefficients of regression.		Number of cases.
x.	y.	η_x.	η_y.	σ_x/σ_y.	σ_y/σ_x.	r_{xy}.	R_{xy}.	R_{yx}.	N.
Father	Son	·5418	·5922	1·0667	·9375	·5503	·5870	·5159	1000
Father	Daughter . . .	·5064	·7322	1·1256	·8885	·4370	·4919	·3883	1000
Mother	Son	·8085	·5937	·9363	1·0680	·4817	·4510	·5145	1000
Mother	Daughter . . .	·8495	·7727	1·0188	·9815	·5096	·5192	·5002	1000
Brother	Brother	·6378	·6378	1	1	·5169	·5169	·5169	1500
Sister	Sister	·8827	·8827	1	1	·4463	·4463	·4463	1500
Brother	Sister	·6643	·8051	·9130	1·0953	·4615	·4213	·5055	1500
Husband	Wife	·4346	·7632	·9209	1·0859	·1002	·9227	1·0881	774
Paternal Grandfather .	Grandson . . .	·4665	·7135	1·0098	·9903	·4213	·4254	·4172	765
Paternal Grandfather .	Granddaughter .	·4698	·8134	1·0384	·9629	·3802	·3948	·3661	681
Maternal Grandfather .	Grandson . . .	·2154	·7129	·9122	1·0963	·3717	·3391	·4075	771
Maternal Grandfather .	Granddaughter .	·3115	·8652	·9675	1·0336	·2969	·2872	·3069	687
Paternal Grandmother .	Grandson . . .	·8491	·6733	·8659	1·1549	·2722	·2357	·3144	741
Paternal Grandmother .	Granddaughter .	·9493	·8156	·9259	1·0800	·2205	·2042	·2381	717
Maternal Grandmother .	Grandson . . .	·9371	·7159	1·1887	·8412	·2523	·2999	·2122	756
Maternal Grandmother .	Granddaughter .	·7673	·9090	1·2909	·7746	·3180	·4105	·2463	739
Paternal Uncle . . .	Nephew	·7387	·6402	·8815	1·1345	·3204	·2824	·3635	1290
Paternal Uncle . . .	Niece	·8087	·9063	·9874	1·0128	·2844	·2808	·2880	1128
Maternal Uncle . . .	Nephew	·4850	·7728	1·0192	·9812	·3706	·3777	·3636	1242
Maternal Uncle . . .	Niece	·4556	·8816	1·0520	·9506	·1938	·2039	·1842	1134
Paternal Aunt . . .	Nephew	·6982	·6563	·9012	1·1096	·2837	·2557	·3148	1186
Paternal Aunt . . .	Niece	·7988	·7539	1·1078	·9027	·1992	·2207	·1798	1149
Maternal Aunt . . .	Nephew	·8179	·6839	·9812	1·0192	·2576	·2528	·2625	1145
Maternal Aunt . . .	Niece	·7210	·7982	·9117	1·0968	·2101	·1915	·2304	1058

TABLE VIII.—On the relative Variability in Eye-colour.

Pair of relatives.		Ratio of variabilities from "excess."		Range of tint 3 in terms of σ. Range $= u\sigma$.		Range of tints 3, 4, 5, 6 in terms of σ. Range $= v\sigma$.		Probable error of median.	
x.	y.	σ_x/σ_y	σ_y/σ_x	u_x.	u_y.	v_x.	v_y.	x.	y.
Father	Son	·9981	1·0019	·6746	·7196	1·4950	1·5110	·0396	·0371
Father	Daughter . . .	1·0532	·9495	·6189	·6966	1·3413	1·6178	·0432	·0384
Mother	Son	·9959	1·0041	·7617	·7132	1·6653	1·4904	·0351	·0375
Mother	Daughter . . .	1·0424	·9593	·7178	·7313	1·5961	1·6500	·0372	·0366
Brother	Brother. . . .	1	1	·7759	·7759	1·5634	1·5634	·0199	·0199
Sister	Sister	1	1	·7347	·7347	1·6107	1·6107	·0210	·0210
Brother	Sister	·9917	1·0084	·8010	·7313	1·6492	1·5916	·0273	·0298
Husband	Wife	·9383	1·0657	·8061	·7423	1·7113	1·6203	·0377	·0409
Paternal Grandfather .	Grandson . . .	·9536	1·0487	·7939	·8017	1·3992	1·6008	·0385	·0381
Paternal Grandfather .	Granddaughter .	·9923	1·0077	·7144	·7418	1·3730	1·6793	·0453	·0437
Maternal Grandfather .	Grandson . . .	·9667	1·0344	·8120	·7407	1·6221	1·5528	·0366	·0411
Maternal Grandfather .	Granddaughter .	1·0775	·9281	·7990	·7730	1·4199	1·7836	·0404	·0417
Paternal Grandmother .	Grandson . . .	1·0242	·9764	·8882	·7691	1·6016	1·5846	·0350	·0404
Paternal Grandmother .	Granddaughter .	1·1019	·9075	·7925	·7338	1·5615	1·7600	·0398	·0430
Maternal Grandmother	Grandson . .	1·0197	·9807	·6332	·7527	1·6379	1·5891	·0486	·0408
Maternal Grandmother	Granddaughter .	1·1519	·8681	·5922	·7645	1·4447	1·7965	·0525	·0407
Paternal Uncle . . .	Nephew . . .	1·0179	·9825	·8242	·7265	1·4636	1·5268	·0286	·0324
Paternal Uncle . . .	Niece	1·1728	·8527	·7695	·7598	1·4253	1·8304	·0327	·0331
Maternal Uncle . . .	Nephew . . .	·9984	1·0016	·7977	·8130	1·5620	1·5763	·0301	·0295
Maternal Uncle . . .	Niece	1·0154	·9849	·7468	·7856	1·4396	1·8208	·0336	·0320
Paternal Aunt . . .	Nephew . . .	·9964	1·0036	·8208	·7397	1·6311	1·5929	·0299	·0332
Paternal Aunt . . .	Niece	1·1804	·8472	·7126	·7894	1·4974	1·8097	·0350	·0316
Maternal Aunt . . .	Nephew . . .	·9332	1·0716	·8022	·7871	1·7337	1·5371	·0311	·0317
Maternal Aunt . . .	Niece	1·0730	·9320	·8925	·8137	1·7202	1·7873	·0291	·0319

general tendency of a series rather than pick out single differences for special consideration. If we do this we shall still find that remarkable results flow from our Tables VII. and VIII., most of which seem hitherto to have escaped attention.

I return now to the special topic of the present section, the mean eye-colour, after this lengthy—if needful—digression on the probable error of the data given in our tables.

We may, I think, safely draw the following conclusions :—

(*a.*) *Man has a mean eye-colour very substantially lighter than that of woman.*

If we compare the mean eye-colour of father with mother, of son with daughter, of brother with sister, of grandfather with mother, of uncle with aunt, of grandson with granddaughter, of nephew with niece, we have the same result—man is distinctly lighter eyed than woman.

(*b.*) *There appears to be a secular change taking place in eye-colour, but this is more marked and definite in the man than in the woman.*

Thus we have the following mean values for η in classes, which must roughly represent successive generations :—

Grandfather	·3658	·4449	Grandmother	·8757	·8523
Father	·5241		Mother	·8290	
Son	·5929	·6484	Daughter	·7524	·8016
Grandson	·7039		Granddaughter	·8508	

Another comparison may be made by noting that grandsons are darker than grandfathers, sons than fathers, nephews than uncles. Similarly, granddaughters are lighter than grandmothers, daughters than mothers, but nieces are not lighter than aunts, as we might have expected. Thus, while the records show a definite darkening of the eyes of men, there appears to be a certain but less sensible lightening of the eyes of women. Again, the younger generations are much closer in eye-colour than the older generations. The difference in eye-colour between grandsons and granddaughters, sons and daughters, nephews and nieces is only about 15 per cent. of the grey blue-green range, but for fathers and mothers it is 30 per cent., and for grandfathers and grandmothers 50 per cent.

When we realise that difference in eye-colour appears to be a sexual character, the true explanation of this secular change in eye-colour becomes still more obscure.

If the lighter eye-colour of the older generation be due to an effect of old age, why is it conspicuous only in the male and not in the female ? Why is the mother sensibly darker than the daughter, but the father sensibly lighter than the son ?

Further, supposing light eyes much commoner among our grandfathers than among their grandsons, and dark eyes among our grandmothers than among their granddaughters, we cannot attribute the great approach in eye-colour to a blending of the parental characters, for, as we shall see later, eye-colour does not seem to blend, it is rather an exclusive character. We should, therefore, be thrown back on prepotency of the mother—a conclusion possibly warranted by our results in the case of daughters, but not in the case of sons. Again, why was there such a marked difference in eye-colour between the men and women of three or four generations back?[*] And if it was a sexual character, why is it disappearing? Was it not, perhaps, a racial difference? Light and dark eyes are not unusually associated with distinct races, and it is just possible that the change in eye-colour is a product of reproductive selection; the old blue-eyed element of the population may be dwindling owing to the greater fertility of the women of dark-eyed race, and thus without any obvious struggle for existence and survival of the fitter, the blue-eyed race may be disappearing from England, as the Langobard type has so largely gone from Italy and the Frank from France.[†] It will not do to be dogmatic about these matters, but the more one measures characters in different generations, the less stable do races appear to be. We speak of the national characters of the Englishman or the Frenchman based upon our experience of how these races have acted in past history, but although there has been no great racial invasion nor struggle, can we safely assert the physical characters of the Englishman to-day do not differ substantially from those of the Englishman of the Commonwealth? It seems to me that the possibly continuous change of characters in a race, not subjected to very apparent internal or external struggle, is a problem of the highest interest to the anthropologist and ultimately to the statesman.

Whatever be the explanation of this secular change in eye-colour, it appears to correspond singularly enough to the secular change we have noted in the coat-colour of thoroughbred horses—in the older generation the sexes differ more widely than in the younger.

(c.) *The maternal male relative (grandfather and uncle) is substantially lighter-eyed than the paternal male relative (grandfather and uncle).*—I see no explanation for this curious result, but it seems worth while to specially note it, for there are curious anomalies in the inheritance through the various male and female lines which may find their complete explanation some day when more and possibly more trustworthy characters have been investigated.

(12.) *On the Variability of Eye-colour with reference to Sex and Class.*—The determination of the relative variability of not exactly measurable characters is, as we have already seen (p. 105), a somewhat delicate problem. It is more so in the case of eye-colour in man than of coat-colour in horses, for there is greater difference in the means, and accordingly the ratio of σ_x/σ_y, as found from the ratio of the "excesses" (p. 105), will be even less reliable.[‡] The class indices corresponding to the

---

[*] Mr. GALTON'S records went back to great-grandfathers, many of whom accordingly appear in our data for grandfathers.

[†] See Note II. at the end of this paper.

[‡] The relative variability of all classes was worked out for thorough-bred horses by the "excess" method, and in only one case—that of dam and colt—did it differ from the bay range method in its determination of the class with the greater variability.

grey blue-green range are also not entirely satisfactory in their results, nor those taken for a still larger range covering tints 3, 4, 5, and 6, or blue-green, grey, hazel, light brown, and brown, which cover roughly about 1·5 to 1·6 times the standard deviation. We shall now consider the results of three methods of considering the relative variability, (α) from the excesses given in columns 1 and 2 of Table VIII.; (β) from the grey blue-green range given in columns 3 and 4 of Table VIII.; and (γ) from the range of tints 3 to 6 inclusive given in columns 5 and 6 of Table VIII. As we have already indicated, these methods are not likely to give the same relative magnitude numerically for the variabilities; we must content ourselves if they agree in making the ratio of σ_x to σ_y greater or less than unity. Now, in the twenty-two cases

α and β disagree in 10 cases.
β and γ disagree in 7 cases.
α and γ disagree in 5 cases.

Further, for the five cases in which α and γ disagree, those for father and son, paternal grandfather and grandson, maternal uncle and nephew, show so little difference of variability in the two sexes that both methods give sensibly the same results, *i.e.*, equality of variability. In the cases of the paternal grandfather and grandchildren, the two methods diverge rather markedly.

It will be of interest accordingly to work out the probable errors as given by the excess method for one, say the first of these cases. The theory is given in Part VII. of the present series. Here $E_1 = 275·165$, $E_2 = 309·013$, whence we find probable error of $E_1 = 17·273$, probable error of $E_2 = 16·925$, correlation between errors in E_1 and $E_2 = -·4424$, probable error in $\sigma_1/\sigma_2 = ·0394$.

Thus the probable error in the ratio of the variabilities is about 4 per cent., and of the order of the quantities by which we are distinguishing the relative size of σ_1 and σ_2.

Further, there is another source of error in evaluating E_1 and E_2 due to the method of interpolation used, and this would still further increase the probable error in σ_1/σ_2. We cannot therefore lay any very great stress on the manner in which the ratios of variabilities for the paternal grandfather and grandchildren have swung round from (α) to (γ).

A further examination shows us that in all five cases wherein γ differs from α it is in accord with β. I shall accordingly take γ as the standard criterion, but in those cases in which it has agreement with α, its conclusions must be given greater weight.

(a.) *On the Relative Variability of Sex in Eye-colour.*—The following male groups are more variable than the corresponding female groups:—

Sons of fathers than daughters of fathers.
Sons of mothers than daughters of mothers.
Brothers of brothers than sisters of sisters.

Grandsons (in four series) than granddaughters (in four same series).
Nephews (in four series) than nieces (in four same series).
Fathers (in two series) than mothers (in two series).
Grandfathers (in four series) than grandmothers (in four like series).
Uncles (in four series) than aunts (in four like series).

The following female groups are more variable than the corresponding male groups :—

Sisters of brothers than brothers of sisters.
Wives than husbands.

We have thus 21 series with male preponderance against only two with female preponderance of variability.

Again, the mean range of tints 3, 4, 5, 6 in 22 male series equals 1·5424 σ_x, and in 22 female series equals 1·6740 σ_y, or we have enough evidence to show that the ratio of male to female variability is about 1·08.*

This greater variability of the male in eye-colour is of considerable interest. It does not appear to be a result of sexual selection, for so far as our comparatively small series weighs, husbands are less variable than wives. That mothers are, however, less variable than fathers seems to indicate that dark-eyed women are more fertile† than light-eyed, for we must bear in mind that mothers have on the average a darker eye-colour than wives. We have thus again reached the same conclusion as before, namely, that a dark-eyed element in the population with a prepotent fertility is replacing the blue-eyed element.

The other female exception to the general rule of greater variability in the eye-colour of the male is that in mixed families the sisters appear to be more variable than their brothers, notwithstanding that brothers of brothers are more variable than sisters of sisters. In other words, so far as eye-colour is concerned an exceptional man is more likely to have brothers than sisters, but an exceptional woman also is more likely to have brothers than sisters. The inference is not very strong, as the excess method (α) makes brothers of sisters and sisters of brothers of sensibly equal variability; it rests therefore on (β) and (γ) only. Still it deserves fuller investigation.

(*b*.) Let A and B be two grades of relationship, of which A refers to the older generation, and A and B refer to either sex. Then the variability of all the A's

---

* It is worth noting that the ratio of male to female variability in the coat-colour of horses is 1·05 (see p. 96). In both cases the female is darker, *i.e.*, has less of "colour"; thus if we could take a coefficient of variation ratio instead of standard deviation ratio as the test, we should find the difference of variability less, possibly even zero.

† For if mothers are to be less variable than wives, their distribution must be more compressed round the mean than that of wives; this denotes that fertility is correlated with eye-colour, and the darker eye-colour goes with the greater fertility. [See Note II. at end of memoir, however.]

who have female B's is invariably greater than the variability of all the A's who have male B's.

The law appears to be universal, at least it is absolutely true for all the 10 cases to which we can apply it. Thus the father of sons is less variable than the father of daughters, the maternal grandmother of grandsons less variable than the maternal grandmother of granddaughters, or the paternal uncle of nephews less variable than the paternal uncle of nieces. In other words, although women appear, in eye-colour, to be less variable than men, they spring from more variable stocks.

This law is a remarkable one, but in face of the evidence for it, it seems difficult to doubt its validity. Should it be true for more characters in man than eye-colour,* the conclusions to be drawn from it will be somewhat far-reaching, however difficult it may be to interpret its physiological significance.

(c.) *On the Relative Variability of Different Generations.*—We have already had occasion to refer to the general rule that the older generation will be found less variable than the younger, for it is in itself a selection, namely, of those able to survive and reproduce themselves. But this rule is obscured in the present case by several extraneous factors, thus :—

(i.) The male is sensibly more variable than the female, consequently it is quite possible that an elder male generation should appear more variable than a younger female generation.

(ii.) There appears to be a secular change in eye-colour going on. Thus while the grandparental population is a selection from the general population, the general population, at a given time, is a selection from that of an earlier period.

Thus, taking means in the cases of the grandparental and avuncular relationships, we have from (γ) the following results :—

The father is more variable than son and than daughter.
The mother is less variable than son and more than daughter.
The grandfather is more variable than grandson and than granddaughter.
The grandmother is less variable than grandson and more than granddaughter.
The uncle is more variable than nephew and more than niece.
The aunt is less variable than nephew and more than niece.

In other words, the older generation is always more variable than the younger, except when rule (α), that the male is more variable than the female, comes in to overturn this law. If we confine ourselves to comparisons of the same sex the rule is seen to be universal.

We are thus forced again to ask for an explanation of the decreasing variability of eye-colour, and can only seek it in that secular change we have several times had

---

* Fathers of daughters are more variable in stature than fathers of sons ('Phil. Trans.,' A, vol. 187, p. 274). I propose to reinvestigate the question with regard to mothers from the material of my family measurement cards, which is far more extensive than the material I had at my disposal in 1895.

occasion to refer to. Mean and standard deviation of eye-colour appear to have changed sensibly during the few generations covered by Mr. GALTON's eye data.

It is difficult to understand how any obscurity about the recording of eye-colours could lead to anything but chaos in the numerical results. It does not seem to me possible that such results as we have reached under (*a*), (*b*), and (*c*), namely, greater variability in the male, greater variability in the stock of the female, and secular change in variability, can be due to any process of recording. I am forced to the conclusion that they are peculiar to the character under investigation, and are not due to the manner of taking the record or of dealing with it arithmetically. I have purposely avoided drawing attention to small differences and forming any conclusions which did not depend on whole series of groups and substantial averages.

(13.) *On the Inheritance of Eye-colour.* (*a.*) *Assortative Mating.*—Before we enter on the problem of inheritance, it is as well to look at the substantial correlation obtained between the eye-colour in husband and wife. When in 1895 I reached the value ·0931 ± ·0473 for stature, I wrote, "we are justified in considering that there is a definite amount of assortative mating with regard to height going on in the middle classes."* Since then we have worked out the coefficients of correlation in stature, forearm, and span for 1000 husbands and wives (instead of 200) from my family data† cards. The results, which are very substantial, will be dealt with in another paper, and amply confirm my view that assortative mating is very real in the case of mankind. The value (·0931) cited above is in close agreement with the result now reached (·1002 ± ·0378) for eye-colour in the *same* material. The correlation between husband and wife for two very divergent characters is thus shown to be about ·1, or is 25 *per cent. greater than is required between first cousins‡ by the law of ancestral heredity.*

This remarkable degree of likeness between husband and wife—the scientific demonstration that like seeks like—cannot be overlooked. It shows that sexual selection, at least as far as assortative mating is concerned, is a real factor of evolution, and that we must follow DARWIN rather than WALLACE in this matter.§

(*b.*) *Collateral Heredity. First Degree.*—I deal first with this form of heredity, as it offers least points for discussion. The values of the correlation ·5169 for brothers, and ·4463 for sisters and sisters are considerably less than what we have found for coat-colour in horses, but, like the value ·4615 for brothers and sisters, are substantially greater than ·4 to be expected from the unmodified Galtonian law. They could be reached by making γ greater than unity in my statement of the law of ancestral heredity.‖ They could also be given by the law of exclusive inheritance

---

* 'Phil. Trans.,' A, vol. 187, p. 273.

† See also 'Grammar of Science,' second edition, pp. 429–437.

‡ 'Roy. Soc. Proc.,' vol. 62, p. 410.

§ 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.*

‖ I have considered possible explanations of this apparently large assortative mating (i.) in both stature (see p. 90) with a certain degree of prepotency in the individual pairing. As we have already noted, collateral inheritance of the first degree alone considered will not enable us to discriminate between blended and exclusive inheritance.

We note that the male in collateral inheritance predominates over the female, brothers being more alike than sisters in eye-colour, and brother and sister more alike than sister and sister. The mean value for inheritance in the same sex is, however, greater than the value for inheritance between opposite sexes (*cf.* p. 102).

(*c.*) *Collateral Heredity. Second Degree.*—A very cursory inspection of the coefficients of correlation for the eight series of avuncular relationships shows us that it is quite impossible that the mean value should be ·15 as required by the Galtonian Law. The average value of the avuncular correlation is ·2650, and of the regression of nephew and niece on uncle or aunt is ·2733. The probable error of the former result will not be more than ·02, and of the latter something greater, as the ratio of the variabilities is open to larger error. This mean value is accordingly, within the limits of errors of investigation, identical with the ·25 to be expected on the theory of exclusive inheritance. It is a value which appears to be quite impossible on the theory of blended inheritance even with my generalised form of the ancestral law.

We may draw several other important conclusions from our table of avuncular correlations:—

(i.) Nephews are more closely related to both uncles and aunts than nieces are. This is true in each individual case, whether it be judged by correlation or regression. The mean correlations for uncles and for aunts are as ·3081 to ·2219 respectively.

(ii.) Uncles are more closely related to nephews and nieces than aunts are. This is true for three out of the four individual cases; in the fourth case the difference is of the order of the probable error of the difference. The mean correlations ot nephews and nieces are as ·2923 and ·2377 respectively.

(iii.) Paternal uncles and aunts are more closely correlated with their nieces and nephews than maternal uncles and aunts. The mean values are as ·2719 to ·2580.

(iv.) Resemblance between individuals of the same sex is closer than between individuals of opposite sex. The mean values for the avuncular correlation in the same sex and in the opposite sex are respectively ·2751 and ·2549.

(v.) Uncles are more closely related to nephews than aunts to nieces (mean correlations as ·3455 to ·2046). In fact, generally, we see a very considerable preponderance of heredity in the male line so far as these avuncular relations for

and eye-colour, being characters of local races, or even families, and the husband seeking his wife in his own locality or kin; (ii.) in a possible correlation of homogamy and fertility. See 'Roy. Soc. Proc.,' vol. 66, p. 28. Neither seem very satisfactory. Consciously or unconsciously, man and woman appear to select their own type in eye-colour and stature, until they are apparently more alike than such close blood relations as first cousins! Until we know how far this correlation extends to other characters, it would, perhaps, be idle to draw conclusions as to its bearing on widely current views as to first cousin marriages.

eye-colour extend. It is noteworthy that while the two highest correlations are reached for nephew with paternal and with maternal uncles, nearly the two lowest are found for niece with paternal and with maternal aunts. Without laying special stress on each small difference, it must be admitted that the avuncular correlations vary in a remarkable manner with sex, and differ very widely from the practical equality of resemblance which we might *à priori* have expected to exist in this relationship.

(*d.*) *Direct Heredity. First Degree.*—Here we have a mean value of the paternal correlation = ·4947. This is in excellent agreement with the ·5 to be expected by our theory of exclusive inheritance; it is thus in practical agreement with the value of the parental correlation obtained for the inheritance of coat-colour in horses. It would not be inconsistent with a high value for γ in the theory of blended inheritance, but such a value of γ is rendered impossible by the values we have obtained for collateral heredity (see 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.*).

We may draw the following special conclusions:—(i.) The son inherits more strongly from his parents than the daughter, the mean correlations are as ·5160 to ·4733; (ii.) The son inherits more strongly from his father than his mother, and the daughter more strongly from her mother than her father.

This is part of the general principle which we have seen to hold, namely: that change of sex weakens the intensity of heredity.

The correlation of father and daughter appears to be abnormally below the other three, but something of the same kind has been noted in certain stature data; as it is, the high correlation of father and son renders the mean paternal correlation with offspring (·4936) sensibly equal to the mean maternal correlation (·4956).

(*e.*) *Direct Heredity. Second Degree.*—If we take the mean value of the eight grandparental correlations, we find it equals ·3164, while the mean value of the regression of offspring on their grandparents is ·3136. These results are absolutely incompatible with the ·15 required by Mr. GALTON's unmodified theory, and they in fact put the theory of blended inheritance entirely out of court. At the same time, unlike the cases of parental, avuncular, and fraternal inheritance, they cannot be said to be in good agreement with the value ·25 required by the theory of exclusive inheritance. We have to admit that our grandparental data are shorter series than in the other cases, and that guesses as to grandparents' eye-colour, based on memory, miniatures, &c., were more likely to be made. Further, such guesses might easily be biased by a knowledge of the eye-colour of more recent members of the family. Still a reduction from ·32 to ·25 is a very large reduction, and we have to remember that for long series in the case of the thoroughbred horses, with no such guessing at colour as may occur with ancestors' eyes, we found ·3353 for the maternal grandsires, a result in excellent agreement with the ·3343 found for the maternal grandfathers in the present case. Thus while the theory of exclusive inheritance *without reversion* suffices to describe the quantitative values we have found for the parental, the avuncular and the fraternal correlation in the cases of both horse and man, it is yet in both these cases unsatisfactory so far as the grandparental inheritance is concerned. It may be imagined that if we allowed for reversion, we might emphasise the grandparental correlation beyond the value ·25 suggested by theory. But I have shown in my memoir on the "Law of Reversion," that with the parental correlation as high as ·5, we cannot hope to have the grandparental correlation even with reversion higher than ·25. (See 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.*) Clearly the values obtained for grandparental correlation in this paper—the first I believe hitherto investigated—seem to present anomalies which our theory of blended inheritance totally fails to account for, and which may require some modification of our views on reversion before we can meet them on our theory of exclusive inheritance.

I note the following general results deduced from our values of the grandparental correlations:—

(i.) Grandsons are more closely correlated with both grandparents than granddaughters are. This is true for three out of the four cases; the exception, maternal grandmother, is covered by another rule (iv.). The mean correlation for grandparents and grandsons is ·3294, and for grandparents and granddaughters ·3039.

(ii.) Grandfathers are more closely correlated with grandchildren than grandmothers are. This is true in three out of the four cases, the fourth being again subject to rule (iv.). The mean correlations for grandfathers and grandmothers are ·3675 and ·2658 respectively.

(iii.) Paternal grandparents appear to be more closely correlated with their grandchildren than maternal grandparents, the average values of the two correlations being ·3236 and ·3097 respectively.

(iv.) Resemblance between individuals of the same sex is closer than between individuals of the opposite sex. The mean values for the grandparental and grandchild correlation in the same and the opposite sexes are ·3329 and ·3004 respectively.

(v.) Grandfathers are more closely related to grandsons than grandmothers to granddaughters, the mean correlations being as ·3965 and ·2693 respectively. It will be noted at once that these five rules are identical with those we have obtained for the avuncular correlations. So that there is small doubt that they are general rules relating to all grades of relationship for this character.

It seems to me probable that the correct form of (iii.) is: Paternal grandfathers are more highly correlated with grandchildren (·4006) than maternal grandfathers (·3343), and paternal grandmothers (·2468) less highly correlated than maternal grandmothers (·2851). I have not stated the rule in this form, because it is not confirmed by the corresponding results for uncles and aunts. Paternal uncles (·3024) are more closely correlated with nephews and nieces than maternal uncles (·2722), but paternal aunts (·2414) are slightly more instead of less correlated with nephews and nieces than maternal aunts (·2338). I consider, however, that the correlation of paternal aunt and nephew (·2837) in our series is abnormally high.

Now it will, I believe, be seen that the investigation of the eight avuncular and the eight grandparental relationships, here made for the first time,* enables us to draw far wider conclusions than when, as hitherto, only parental and fraternal correlations are dealt with. In making the subjoined general statements, however, I must emphasise the following limitations :—

(α.) The rules are deduced only from data for one character in one type of life.

(β.) This character appears to be undergoing a secular change, a change very possibly due to a correlation between eye-colour and fertility in woman. Thus such a change might not unlikely differentiate the male and female influences in heredity.

My conclusions, definitely true for eye-colour in man, and at the very least suggestive for investigations on other characters in other types of life, are :—

(i.) That the younger generation takes, as a whole, more after its male than its female ascendants and collaterals.

(ii.) That the younger generation is more highly correlated with an ascendant or collateral of the same than of the opposite sex.

(iii.) That the younger generation is more highly correlated with an ascendant or higher collateral reached by a line passing through one sex only than if the line changes sex.

Thus correlation is greater with a paternal uncle than with a maternal uncle, or with a maternal grandmother than a paternal grandmother.

(iv.) Males are more highly correlated with their ascendants and collaterals than females are.

The above rules apply to the *averages;* individual exceptions will be generally found to arise from a conflict of rules. Thus (ii.) and (iii.) may in special cases come into conflict with (i.). When we have more data for a greater variety of characters, we shall see better the relative weight of these rules in cases where they conflict.

(*f.*) *Exclusive Inheritance in Eye-Colour.*—A cursory examination of the eye-colour records shows at once how rare is a blend of the parental tints. Even when such is recorded, it is by no means clear that we have not to deal with a medium tint which is really a case of reversion to a medium tinted ancestor. The failure of eye-colour to blend is, I think, well illustrated by what Mr. GALTON has termed cases of "particulate" inheritance. In the thousands of eye-colours I have been through, I noticed some half-dozen cases only in which the two eyes of the same individual were of different tint, or the iris of one pupil had streaks of a second tint upon it.†

---

* I anticipate equally valuable results when characters are first correlated for the nine possible cousin series.

† In the same manner the occurrence of particulate inheritance in coat-colour in horses may be really an argument against the existence of blends. In the many volumes of the studbooks I have examined, the recorded instances of piebalds are vanishingly few in number.

If we allow that it is from the theory of exclusive inheritance that we must seek results in the present cases, we see that for parental, collateral, and avuncular relationships we get quite excellent results, but that the grandparental relationship is somewhat anomalous. *À priori* it might appear that reversion would aid us in increasing the correlation between offspring and remote ascendants. But, as I have shown elsewhere,[*] this superficial view of reversion forgets that the parents as well as the offspring revert, and if we increase the grandparental correlation above ·25, we at once reach difficulties in the values of the parental correlation, provided we adopt what appear to be reasonable assumptions as to reversion being a continuous and decreasing factor from stage to stage of ancestry. I am inclined accordingly to suspend judgment on the grandparental relationships, thinking that the smallness of the number of families dealt with in Mr. GALTON'S data (200) may have something to do with my peculiar results. Meanwhile I shall endeavour to get the remaining six grandparental tables for thoroughbred horses worked out, and see whether they confirm the high values already found for the two maternal grandsires and offspring, or give an average value much nearer ·25.

That the reader may see at a glance the general results hitherto obtained in this and other papers, I append the following table of inheritance :—

---

[*] See my paper on "The Law of Reversion," 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.* Also 'The Grammar of Science,' second edition, 1900, pp. 486–96, "On Exclusive Inheritance."

TABLE IX.—Theoretical and Actual Results for Inheritance.

Relationship.	Theory.		Man.			Horse.	Hound.	Daphnia.
	Blended inheritance.[1]	Exclusive inheritance.[2]	Stature[3].	Head index.[4]	Eye-colour.[5]	Coat-colour.[6]	Coat-colour.[7]	Spine.[8]
Parental	·3000	·5000	·3355	·3348	·4947	·5216	·3507	[·3295]
Mid-parental . . .	·4242	—	·4745	·4735	—	—	—	·4660
Grandparental . .	·1500	·2500	—	—	·3166	·3353	·1340	[·1360]
G. Grandparental .	·0750	·1250	—	—	—	—	·0404	—
Avuncular . . .	·1500	·2500	—	—	·2650	—	—	—
Whole sibling . .	·4000	·4 to 1·0	·4034	·4025	·4749	·6329	·5170	·6934
Half sibling . . .	·2000	·2 to ·5	—	—	—	·3550	·1646	—

[1] Mr. GALTON'S unmodified hypothesis. See "Law of Ancestral Heredity," 'Roy. Soc. Proc.,' vol. 62, p. 397.

[2] Without any reversion. See "Law of Reversion," 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.* The values for the fraternal correlation depend on the degree of prepotency of either parent within the union.

[3] See 'Phil. Trans.,' A, vol. 187, p. 270.

[4] See 'Roy. Soc. Proc.,' vol. 62, p. 413. The paternal correlations, for reasons stated in the paper, are excluded from the result.

[5] See p. 113 *et seq.* of the present memoir.

[6] See p. 98 *et seq.* of the present memoir. The grandparental correlation is based on two cases only.

[7] See 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.*

[8] See 'Roy. Soc. Proc.,' vol. 65, p. 154. I have deduced the value for parents and grandparents from Dr. WARREN'S results for midparent and midgrandparent. The value for whole siblings I obtained from Dr. WARREN'S measurements, which he with great kindness placed at my disposal.

(14.) *Conclusions.*—The course of this investigation has not been without difficulties, and I am fully prepared to admit that more obscurity and greater probable errors are likely to arise when we deal with the inheritance of a character not directly measurable, than when we take that of a character to which we can at once apply a quantitative scale. But I contend that many of the characters, the inheritance of which it is most important to investigate, do not at present, and perhaps never will, admit of a quantitative measurement. We can arrange in order, we can classify, we can say more or less intense, but we cannot read off value on a scale. It is just such characters also, which the not highly trained observer can most easily appreciate and record. Hence we have been compelled to devise some method of dealing with them, and the present paper illustrates how the method invented can be applied to reach results of considerable interest and of substantial validity.

In order to illustrate the method, I chose two characters, coat-colour in horses and eye-colour in man, which seemed sufficiently diverse both as to origin and species.* The new method enabled me to reach results for half-brethren, grandparents and uncles and aunts, which had not yet been independently considered. The conclusions arrived at for eye-colour in man at no point conflict with those for coat-colour in horses, and both in the main accord with the theory of exclusive inheritance without reversion herein developed. We find—

(i.) No approach to a single value for the coefficient of inheritance for each grade of relationship; it varies widely with the sex, and the line through which the relationship is traced.

(ii.) No approach in average values to those which would be indicated by Mr. GALTON'S Law.

Nor does the modification of Mr. GALTON'S Law, which I have termed the Law of Ancestral Heredity, give better results. For, if we cause it to give the parental values, it then renders results inconsistent with the fraternal values.

(iii.) There is agreement with the theory of exclusive inheritance without reversion for the parental, avuncular and fraternal series; but there is some anomaly in the case of grandparental inheritance. This requires further investigation, and possibly a modification of our views on the nature of reversion.

We want a list formed of characters in various types of life, which are supposed to be exclusively inherited, and then experiments ought to be made and statistics collected with regard to these characters. It is in this field of exclusive inheritance that we must look for real light on the problem of reversion.

If we consider the three known forms of inheritance, the blended, the exclusive, and the particulate (which may possibly be combined in one individual, if we deal with different organs); if we consider further that these forms may possibly have to be supplemented by others not yet recognised (*e.g.*, reversional theories depending, say, on heterogamous unions), then it would appear that the time is hardly ripe even for provisional *mechanical* theories of heredity. What we require to know first is, the class of organs and the types of life for which one or other form of inheritance predominates. As variation in no wise depends on the existence of *two* germ-plasms, so biparental heredity can by no means be treated as the result of their simple quantitative mixture; the component parts of these germ-plasms corresponding to special characters and organs, must be able to act upon each other in a variety of qualitatively different ways. To adopt for a moment the language of DARWIN'S theory of pangenesis, the multiplying gemmules from an organ in the father must (i.) cross with gemmules from that organ in the mother, and the hybrid gemmules give rise to blended inheritance, (ii.) must without crossing multiply alongside the gemmules of the mother, and give rise to particulate inheritance, (iii.) must alone survive, or alone

---

* Since supplemented by my investigations, based on Mr. GALTON'S data, for coat-colour in hounds, 'Roy. Soc. Proc.,' vol. 66, p. 140 *et seq.*

be destroyed in a struggle for existence with those of the mother, and give rise to exclusive inheritance. And all these three processes may be going on within the same germ-plasm mixture at the same time! Even without using the language of gemmules, processes analogous to the above must be supposed to take place. Thus a quantitative "mixture of germ-plasms" becomes a mere name, screening a whole range of mechanical processes; and very possibly a new one could be found for each new form of heredity as it occurs. Such processes like the old ones would still remain without demonstrable reality under the veil of "mixture of germ-plasms."

What I venture to think we require at present is not a hypothetical plasmic mechanics, but careful classifications of inheritance for the several grades of relationship, for a great variety of characters, and for many types of life. This will require not only the formation of records and extensive breeding experiments, but ultimately statistics and most laborious arithmetic. Till we know what class of characters blend, and what class of characters is mutually exclusive, we have not within our cognizance the veriest outlines of the phenomena which the inventors of plasmic mechanisms are in such haste to account for. Such inventors are like planetary theorists rushing to prescribe a law of attraction for planets, the very orbital forms of which they have not first ascertained and described. Without the observations of TYCHO BRAHÉ, followed by the arithmetic of KEPLER, no NEWTON had been possible. The numerical laws for the intensity of inheritance must first be discovered from wide observation before plasmic mechanics can be anything but the purest hypothetical speculation.

APPENDIX I.

Tables of Colour Inheritance in Thoroughbred Racehorses, extracted by Mr. LESLIE BRAMLEY-MOORE from WEATHERBY'S Studbooks.

TABLE OF COLOURS.

1 = black (bl.)
2 = black or brown (bl./br.).
3 = brown or black (br./bl.).
4 = brown (br.).
5 = brown or bay (br./b.).
6 = bay or brown (b./br.).
7 = bay (b.).
8 = bay or chestnut (b./ch.).
9 = chestnut or bay (ch./b.).
10 = chestnut (ch.).
11 = chestnut or roan (ch./ro.).
12 = roan or chestnut (ro./ch.).
13 = roan (ro.).
14 = roan or grey (ro./gr.).
15 = grey or roan (gr./ro.)
16 = grey (gr.).

TABLE I.—Direct Inheritance. First Generation.

Sires and Colts.

Sires.

Colts.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	3	0	0	7	0	0	12	0	0	7	0	0	0	0	0	0	29
2	bl./br.	0	0	0	2	0	1	4	0	0	0	0	0	0	0	0	0	7
3	br./bl.	0	0	0	1	0	0	3	0	0	0	0	0	0	0	0	0	4
4	br.	2	0	0	64	0	4	105	0	0	38	1	0	0	0	0	1	215
5	br./b.	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	3
6	b./br.	0	0	0	12	0	0	23	0	0	8	0	0	0	0	0	1	44
7	b.	1	2	1	89	0	14	389	0	0	122	0	0	0	0	0	1	619
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
10	ch.	1	2	0	34	0	0	147	0	0	185	0	0	0	0	0	3	372
11	ch./ro.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	gr.	0	0	0	0	0	0	3	0	0	2	0	0	0	0	0	0	5
Totals...	...	7	4	1	209	0	19	691	0	0	362	1	0	0	0	0	6	1300

TABLE II.—Direct Inheritance. First Generation.

Sires and Fillies.

Sires.

Fillies.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	0	0	0	3	0	0	10	0	0	3	0	0	0	0	0	0	16
2	bl./br.	0	0	0	8	0	0	3	0	0	1	0	0	0	0	0	0	12
3	br./bl.	0	0	0	1	0	0	6	0	0	0	0	0	0	0	0	0	7
4	br.	2	1	1	42	0	6	76	0	0	23	0	0	0	0	0	0	151
5	br./b.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
6	b./br.	0	0	0	6	0	0	29	0	0	7	0	0	0	0	0	0	42
7	b.	2	1	0	57	0	12	364	0	0	91	0	0	0	0	0	0	527
8	b./ch.	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	2
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	2	1	0	21	0	1	118	0	0	133	2	0	0	0	0	1	279
11	ch./ro.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	3
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
16	gr.	0	0	0	1	0	0	0	0	0	4	0	0	0	0	0	3	8
Totals...	...	6	3	1	139	0	19	610	0	0	264	2	0	0	0	0	6	1050

TABLE III.—Direct Inheritance. First Generation.

Dams and Colts.

	Colts.	1 bl.	2 bl./br.	3 br./bl.	4 br.	5 br./b.	6 b./br.	7 b.	8 b./ch.	9 ch./b.	10 ch.	11 ch./ro.	12 ro./ch.	13 ro	14 ro./gr.	15 gr./ro.	16 gr.	Totals.
		Dams.																
1	bl.	5	1	0	3	0	0	3	0	0	4	0	0	0	0	0	0	16
2	bl./br.	1	0	0	2	0	0	4	0	0	0	0	0	0	0	0	0	7
3	br./bl.	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	br.	6	1	0	59	0	5	60	0	0	18	0	0	0	0	0	0	149
5	br./b.	0	0	0	3	0	0	1	0	0	1	0	0	0	0	0	0	5
6	b./br.	1	0	0	14	0	3	15	0	0	6	0	0	0	0	0	0	39
7	b.	9	1	0	76	0	6	295	0	1	86	0	1	1	0	0	4	480
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	6	0	0	27	0	2	127	0	0	133	0	0	2	0	0	3	300
11	ch./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
16	gr.	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	2
Totals...	...	28	3	1	184	0	16	505	0	1	249	1	1	4	0	0	7	1000

TABLE IV.—Direct Inheritance. First Generation.

Dams and Fillies.

Dams.

Fillies.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	1	1	0	1	0	1	4	0	0	2	0	0	0	0	0	0	10
2	bl./br.	1	0	0	4	0	0	8	0	0	2	0	0	0	0	0	0	15
3	br./bl.	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
4	br.	3	0	0	52	0	2	67	0	0	18	0	0	1	0	0	0	143
5	br./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	b./br.	3	0	0	15	0	1	31	0	0	6	0	0	0	0	0	0	56
7	b.	13	1	0	74	0	6	288	1	0	79	0	0	1	0	0	1	464
8	b./ch.	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	4	1	0	27	0	7	109	0	1	146	0	1	0	0	0	3	299
11	ch./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
15	gr./ro.	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	2
16	gr.	0	0	0	1	0	0	1	0	0	2	0	0	0	0	0	3	7
Totals ...	...	25	3	0	175	0	18	509	1	1	256	0	1	4	0	0	7	1000

TABLE V.—Direct Inheritance. Second Generation.

Maternal Grandsires and Colts.

Grandsires.

Colts.

		1 bl.	2 bl./br.	3 br./bl.	4 br.	5 br./b.	6 b./br.	7 b.	8 b./ch.	9 ch./b.	10 ch.	11 ch./ro.	12 ro./ch.	13 ro.	14 ro./gr.	15 gr./ro.	16 gr.	Totals.
1	bl.	0	0	0	1	0	1	12	0	0	4	0	0	0	0	0	0	18
2	bl./br.	0	0	0	2	0	0	11	0	0	0	0	0	0	0	0	0	13
3	br./bl.	0	0	0	1	0	0	2	0	0	1	0	0	0	0	0	0	4
4	br.	1	0	0	29	0	2	94	0	0	29	0	0	0	0	0	1	156
5	br./b.	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	3
6	b./br.	2	0	0	7	0	3	24	0	0	12	0	0	0	0	0	0	48
7	b.	0	1	0	51	0	8	286	0	0	100	0	0	0	0	0	1	447
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	0	2	0	17	0	1	153	0	0	128	2	0	0	0	0	1	304
11	ch./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
14	ro./gr.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
15	gr./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	gr.	0	0	0	0	0	1	2	0	0	1	0	0	0	0	0	1	5
Totals...	...	3	3	0	110	0	16	587	0	0	275	2	0	0	0	0	4	1000

TABLE VI.—Direct Inheritance. Second Generation.

Maternal Grandsires and Fillies.

Grandsires.

Fillies.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	0	0	0	5	0	0	16	0	0	8	0	0	0	0	0	1	30
2	bl./br.	0	0	0	2	0	0	6	0	0	1	0	0	0	0	0	0	9
3	br./bl.	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	3
4	br.	0	2	0	22	0	4	72	0	0	31	0	0	0	0	0	0	131
5	br./b.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
6	b./br.	0	1	0	6	0	3	31	0	0	7	0	0	0	0	0	0	48
7	b.	1	3	0	55	0	5	305	0	0	101	0	0	0	0	0	0	470
8	b./ch.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	0	1	0	34	0	3	141	0	0	116	1	0	0	0	0	1	297
11	ch./ro.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	2
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
16	gr.	0	0	0	0	0	0	2	0	0	2	0	0	0	0	0	2	6
Totals...	...	1	7	0	124	0	15	579	0	0	269	1	0	0	0	0	4	1000

TABLE VII.—Collateral Inheritance. First Degree. (Half-Brothers.)

Colts and Colts. Same Mare but different Sires.

First Colt.

Second Colt.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	2	1	0	6	0	1	8	0	0	5	0	0	0	0	0	0	23
2	bl./br.	0	0	0	2	0	0	5	0	0	6	0	0	0	0	0	0	13
3	br./bl.	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	2
4	br.	4	3	0	44	0	10	72	0	0	34	0	0	0	0	1	0	168
5	br./b.	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
6	b./br.	1	0	0	4	0	10	19	0	0	8	0	0	0	0	0	0	42
7	b.	5	3	2	62	1	19	242	0	0	119	0	0	0	0	0	1	454
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
10	ch.	2	2	0	24	0	5	119	0	0	134	0	0	0	0	0	0	286
11	ch./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	2
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	gr.	0	0	0	1	0	0	2	0	0	3	0	0	0	0	0	2	8
Totals...	...	14	9	2	143	1	46	470	0	0	311	0	0	0	0	1	3	1000

TABLE VII$^{\text{bis}}$.—Collateral Heredity. First Degree. (Half-Brothers.)
Colts and Colts. Same Mare but different Sires.

Symmetrical Distribution.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	4	1	0	10	0	2	13	0	0	7	0	0	0	0	0	0	37
2	bl./br.	1	0	0	5	0	0	8	0	0	8	0	0	0	0	0	0	22
3	br./bl.	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	4
4	br.	10	5	0	88	0	14	134	0	0	58	0	0	0	0	1	1	311
5	br./b.	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	2
6	b./br.	2	0	1	14	0	20	38	0	0	13	0	0	0	0	0	0	88
7	b.	13	8	3	134	1	38	484	0	1	238	0	0	1	0	0	3	924
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
10	ch.	7	8	0	58	1	13	238	0	0	268	0	0	1	0	0	3	597
11	ch./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	2
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
16	gr.	0	0	0	1	0	0	3	0	0	3	0	0	0	0	0	4	11
Totals...	...	37	22	4	311	2	88	924	0	1	597	0	0	2	0	1	11	2000

TABLE VIII.—Collateral Heredity. First Degree. (Whole Brothers.)

Colts and Colts. Same Mare and Sire.

First Colt.

Second Colt.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	3	0	0	7	0	0	8	0	0	0	0	0	0	0	0	0	18
2	bl/br.	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	4
3	br./bl.	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2
4	br.	5	1	0	62	0	2	67	0	0	30	0	0	1	0	0	0	168
5	br./b.	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
6	b./br.	0	0	0	14	0	2	17	0	0	3	0	0	0	0	0	0	36
7	b.	9	4	1	74	0	18	290	0	0	81	0	0	0	0	0	1	478
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	2	2	0	27	0	4	89	0	0	164	0	0	1	0	0	0	289
11	ch./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
14	ro./gr.	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
15	gr./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
Totals...	...	20	7	2	188	1	26	473	0	0	278	0	0	2	0	0	3	1000

TABLE $\mathrm{VIII}^{\mathrm{bis}}$.—Collateral Heredity. First Degree. (Whole Brothers.)

Colts and Colts. Same Mare and Sire.

Symmetrical Distribution.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	6	0	1	12	0	0	17	0	0	2	0	0	0	0	0	0	38
2	bl./br.	0	0	1	4	0	0	4	0	0	2	0	0	0	0	0	0	11
3	br./bl.	1	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	4
4	br.	12	4	0	124	0	16	141	0	0	57	0	0	1	1	0	0	356
5	br./b.	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	2
6	b./br.	0	0	0	16	0	4	35	0	0	7	0	0	0	0	0	0	62
7	b.	17	4	2	141	0	35	580	0	0	170	0	0	1	0	0	1	951
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	2	2	0	57	0	7	170	0	0	328	0	0	1	0	0	0	567
11	ch./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	3
14	ro./gr.	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
15	gr./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	gr.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	4	5
Totals...	...	38	11	4	356	2	62	951	0	0	567	0	0	3	1	0	5	2000

TABLE IX.—Collateral Heredity. (Half Sisters.)

Fillies and Fillies. Same Mare but different Sire

First Filly.

Second Filly.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	2	0	0	4	0	1	13	0	0	4	0	0	0	0	0	0	24
2	bl./br.	0	0	0	1	0	0	3	0	0	3	0	0	0	0	0	1	8
3	br./bl.	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
4	br.	4	1	0	42	0	5	64	0	0	25	0	0	0	0	0	0	141
5	br./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	b./br.	0	1	0	3	0	6	23	0	0	7	0	0	0	0	0	0	40
7	b.	9	3	1	61	0	21	313	0	0	103	0	0	0	0	0	2	513
8	b./ch.	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	5	5	0	26	0	10	91	0	0	119	0	0	0	0	1	1	258
11	ch./ro.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	2	0	0	2	0	0	0	0	0	0	4
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	2
16	gr.	0	0	0	3	0	0	2	0	0	1	0	0	0	0	0	1	7
Totals...	...	20	10	1	140	0	44	513	0	0	266	0	0	0	0	1	5	1000

TABLE IX^{bis}.—Collateral Heredity. First Degree. (Half-Sisters.)

Fillies and Fillies. Same Mare but different Sires.

Symmetrical Distribution.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	4	0	0	8	0	1	22	0	0	9	0	0	0	0	0	0	44
2	bl./br.	0	0	0	2	0	1	6	0	0	8	0	0	0	0	0	1	18
3	br./bl.	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	2
4	br.	8	2	0	84	0	8	125	0	0	51	0	0	0	0	0	3	281
5	br./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	b./br.	1	1	0	8	0	12	44	0	0	17	0	0	0	0	1	0	84
7	b.	22	6	1	125	0	44	626	0	0	194	1	0	2	0	1	4	1026
8	b./ch.	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	9	8	1	51	0	17	194	1	0	238	0	0	2	0	1	2	524
11	ch./ro.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	2	0	0	2	0	0	0	0	0	0	4
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	3
16	gr.	0	1	0	3	0	0	4	0	0	2	0	0	0	0	0	2	12
Totals...	...	44	18	2	281	0	84	1026	1	0	524	1	0	4	0	3	12	2000

TABLE X.—Collateral Heredity. First Degree. (Whole Sisters.)

Fillies and Fillies. Same Mare and same Sire.

First Filly.

Second Filly.		1. bl.	2. bl./br.	3. br./bl.	4 br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	6	1	0	6	0	1	11	0	0	4	0	0	0	0	0	0	29
2	bl./br.	1	0	0	2	0	2	3	0	0	0	0	0	0	0	0	0	8
3	br./bl.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	br.	6	5	0	64	0	4	53	0	0	21	0	0	0	0	0	0	153
5	br./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	b./br.	1	1	0	5	0	4	18	0	0	3	0	0	0	0	0	0	32
7	b.	8	2	0	55	0	17	336	0	0	77	0	0	0	0	0	1	496
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	4	1	0	23	0	3	77	0	0	168	0	1	0	0	0	0	277
11	ch./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	gr.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	4	5
Totals...	...	26	10	0	155	0	31	499	0	0	273	0	1	0	0	0	5	1000

TABLE X$^{\text{bis}}$.—Collateral Heredity. First Degree. (Whole Sisters).

Fillies and Fillies. Same Mare and same Sire.

Symmetrical Distribution.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	12	2	0	12	0	2	19	0	0	8	0	0	0	0	0	0	55
2	bl./br.	2	0	0	7	0	3	5	0	0	1	0	0	0	0	0	0	18
3	br./bl.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	br.	12	7	0	128	0	9	108	0	0	44	0	0	0	0	0	0	308
5	br./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	b./br.	2	3	0	9	0	8	35	0	0	6	0	0	0	0	0	0	63
7	b.	19	5	0	108	0	35	672	0	0	154	0	0	0	0	0	2	995
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	8	1	0	44	0	6	154	0	0	336	0	1	0	0	0	0	550
11	ch./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ro./ch.	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
13	ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	gr.	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	8	10
Totals...	...	55	18	0	308	0	63	995	0	0	550	0	1	0	0	0	10	2000

Table XI.—Collateral Heredity. First Degree. (Half Sister and Half Brother.)

Fillies and Colts. Same Mare but different Sires.

Fillies.

Colts.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	0	0	0	4	0	0	7	0	0	7	0	0	0	0	0	0	18
2	bl./br.	1	0	0	2	0	0	2	0	0	0	0	0	0	0	0	0	5
3	br./bl.	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
4	br.	3	0	0	34	0	8	71	0	0	32	0	0	1	0	1	0	150
5	br./b.	0	1	0	2	0	1	0	0	0	1	0	0	0	0	0	0	5
6	b./br.	0	0	0	13	0	10	18	0	0	5	0	0	0	0	0	1	47
7	b.	5	8	0	61	0	22	263	0	0	112	0	0	3	0	0	2	476
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	4	2	0	28	0	15	125	0	0	116	0	0	1	0	1	0	292
11	ch./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	2
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
16	gr.	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	3
Totals...	...	13	11	0	145	0	57	489	0	0	274	0	0	5	0	2	4	1000

TABLE XII.—Collateral Heredity, First Degree. (Whole Brother and Whole Sister.)

Fillies and Colts. Same Mare and Sire.

Fillies.

Colts.

		1. bl.	2. bl./br.	3. br./bl.	4. br.	5. br./b.	6. b./br.	7. b.	8. b./ch.	9. ch./b.	10. ch.	11. ch./ro.	12. ro./ch.	13. ro.	14. ro./gr.	15. gr./ro.	16. gr.	Totals.
1	bl.	6	0	0	5	0	0	6	0	0	3	0	0	0	0	0	0	20
2	bl./br.	0	2	0	0	0	0	1	0	0	1	0	0	0	0	0	0	4
3	br./bl.	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	br.	3	1	0	63	0	9	69	0	0	22	0	0	0	0	1	0	168
5	br./b.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
6	br./b.	1	0	0	5	0	3	11	0	0	6	0	0	0	0	0	0	26
7	b.	8	5	0	63	0	12	332	0	0	77	0	0	0	0	0	2	499
8	b./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ch./b.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ch.	1	1	0	26	0	4	97	0	0	141	0	0	1	0	0	0	271
11	ch./ro.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
12	ro./ch.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	ro./gr.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	gr./ro.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	gr.	0	1	0	0	0	2	2	0	0	2	0	0	0	0	0	2	9
Totals ...	...	20	10	0	162	0	30	520	0	0	252	0	0	1	0	1	4	1000

APPENDIX II.

Tables of Eye-colour Inheritance in Man, extracted by KARL PEARSON from Mr. FRANCIS GALTON'S Family Records.

TABLE OF TINTS.

1 = light blue.
2 = blue, dark blue.
3 = blue-green, grey.
4 = dark grey, hazel.
5 = light brown.
6 = brown.
7 = dark brown.
8 = very dark brown, black.

This grouping is not quite in keeping with more recent divisions of eye-colour, but being that adopted by Mr. GALTON in his original collection of data, it could not be modified in accordance with present practice.

TABLES for the Direct Inheritance of Eye-colour. First Generation.

I.—Fathers and Sons. 1000 Cases.

Fathers.

	Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Sons.	1	9	12	5	5	0	1	2	0	34
	2	10	163	65	36	1	7	15	4	301
	3	10	73	124	41	1	12	18	5	284
	4	4	21	34	55	0	11	11	1	137
	5	0	0	0	0	2	2	1	0	5
	6	1	26	12	19	1	19	16	6	100
	7	1	23	16	14	0	11	31	2	98
	8	1	4	8	10	0	1	7	10	41
	Totals	36	322	264	180	5	64	101	28	1000

II.—Fathers and Daughters. 1000 Cases.

Fathers.

	Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Daughters.	1	4	9	11	4	0	1	2	1	32
	2	11	139	57	31	0	6	24	5	273
	3	9	73	111	38	1	15	19	3	269
	4	5	43	34	54	2	10	14	3	165
	5	0	1	3	0	0	3	1	0	8
	6	1	45	13	19	0	23	15	3	119
	7	2	27	10	12	0	7	41	6	105
	8	0	8	4	0	0	2	11	4	29
	Totals	32	345	243	158	3	67	127	25	1000

III.—Mothers and Sons. 1000 Cases.

Mothers.

	Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Sons.	1	5	14	6	3	0	1	0	6	35
	2	12	119	83	29	8	20	21	9	301
	3	13	54	113	35	4	37	14	8	278
	4	3	21	26	54	1	17	6	6	134
	5	0	1	1	0	0	3	0	0	5
	6	1	9	26	10	1	30	24	3	104
	7	1	9	19	16	0	18	31	7	101
	8	0	7	15	4	1	3	5	7	42
	Totals	35	234	289	151	15	129	101	46	1000

IV.—Mothers and Daughters. 1000 Cases.

Mothers.

	Tints.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Daughters.	1	5	15	3	2	2	2	2	0	31
	2	7	99	67	29	2	15	23	13	255
	3	7	77	111	38	1	26	14	6	280
	4	5	22	34	46	2	27	21	7	164
	5	0	2	2	0	3	1	2	1	11
	6	0	13	27	20	1	35	17	7	120
	7	0	13	21	16	1	19	26	9	105
	8	1	5	7	2	1	4	12	2	34
	Totals.	25	246	272	153	13	129	117	45	1000

TABLES for the Collateral Inheritance of Eye-colour.

V^a.—Brothers and Brothers. 1500 Cases.

First Brother.

	First.	1.	2.	3.	4.	5.	6.	7	8.	Totals.
Second Brother.	1	8	2	3	4	0	0	0	0	19
	2	36	202	23	17	0	6	4	3	291
	3	16	182	209	26	0	4	2	2	441
	4	6	36	71	84	0	7	2	0	206
	5	0	3	2	1	0	1	0	0	7
	6	3	56	50	39	0	34	5	6	193
	7	6	37	76	48	1	36	36	2	242
	8	4	24	26	18	0	8	6	15	101
	Totals	79	542	460	237	1	96	55	30	1500

VI^a.—Sisters and Sisters. 1500 Cases.

First Sister.

Second Sister.

First.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
1	10	2	1	1	0	0	0	0	14
2	17	147	29	6	0	10	6	2	217
3	10	136	186	24	0	9	5	3	373
4	3	75	94	66	0	1	10	0	249
5	0	3	5	2	2	0	1	0	13
6	2	57	69	55	5	52	9	2	251
7	4	56	61	52	10	59	49	0	291
8	2	20	10	7	6	13	26	8	92
Totals	48	496	455	213	23	144	106	15	1500

V^b.—Brothers and Brothers. Symmetrical System.

Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
1	16	38	19	10	0	3	6	6	98
2	38	404	205	53	3	62	41	27	833
3	19	205	418	97	2	54	78	28	901
4	10	53	97	168	1	46	50	18	443
5	0	3	2	1	0	1	1	0	8
6	3	62	54	46	1	68	41	14	289
7	6	41	78	50	1	41	72	8	297
8	6	27	28	18	0	14	8	30	131
Totals	98	833	901	443	8	289	297	131	3000

VI^b.—Sisters and Sisters. Symmetrical System.

Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
1	20	19	11	4	0	2	4	2	62
2	19	294	165	81	3	67	62	22	713
3	11	165	372	118	5	78	66	13	828
4	4	81	118	132	2	56	62	7	462
5	0	3	5	2	4	5	11	6	36
6	2	67	78	56	5	104	68	15	395
7	4	62	66	62	11	68	98	26	397
8	2	22	13	7	6	15	26	16	107
Totals	62	713	828	462	36	395	397	107	3000

VII.—Brothers and Sisters. 1500 Cases.

Brothers.

Sisters.

Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
1	5	9	18	4	0	0	1	0	37
2	20	163	101	36	0	28	19	13	380
3	9	98	193	50	0	37	17	14	418
4	5	36	49	67	3	28	13	16	217
5	0	2	5	1	1	2	2	3	16
6	3	47	41	27	4	42	17	14	195
7	4	34	49	22	3	27	30	19	178
8	1	10	7	1	0	10	8	22	59
Totals	47	399	463	208	11	174	107	91	1500

TABLE for Assortative Mating in Eye-colour.

VIII.—Husbands and Wives. 774 Cases.

Husbands.

Wives.

Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
1	2	13	4	3	0	1	2	0	25
2	6	87	42	26	0	16	13	6	196
3	6	56	93	31	1	16	11	6	220
4	4	32	35	18	1	15	6	1	112
5	0	0	5	1	0	0	1	0	7
6	2	38	27	10	1	12	10	1	101
7	5	20	28	7	1	6	12	4	83
8	2	8	8	2	0	2	4	4	30
Totals	27	254	242	98	4	68	59	22	774

TABLES for the Direct Inheritance of Eye-colour. Second Generation.

IX.—Paternal Grandfather and Grandson. 765 Cases.

Paternal Grandfather.

Grandson.

First.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
1	4	10	3	0	0	1	3	0	21
2	7	115	31	20	1	6	13	3	196
3	5	64	109	21	0	10	22	4	235
4	2	25	40	21	0	9	13	6	116
5	0	0	0	1	0	1	0	0	2
6	0	14	32	11	0	15	5	2	79
7	4	16	16	9	1	11	21	2	80
8	0	6	5	0	0	3	16	6	36
Totals	22	250	236	83	2	56	93	23	765

X.—Paternal Grandfather and Granddaughter. 681 Cases.

Paternal Grandfather.

Granddaughter.

First.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
1	3	6	4	5	0	1	1	0	20
2	2	94	32	10	2	6	16	4	166
3	5	67	71	17	0	9	20	3	192
4	4	35	33	26	1	10	9	3	121
5	0	3	4	1	0	1	2	1	12
6	1	16	21	11	0	15	6	4	74
7	3	10	20	11	1	8	15	3	71
8	0	2	5	1	1	1	10	5	25
Totals	18	233	190	82	5	51	79	23	681

TABLE XI.—Maternal Grandfather and Grandson. 771 Cases.

Maternal Grandfather.

Grandson.

Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
1	3	11	3	1	—	—	1	—	19
2	8	113	46	22	1	13	5	3	211
3	10	87	89	12	—	11	8	3	220
4	4	33	35	22	—	15	6	2	117
5	—	1	1	—	—	—	—	—	2
6	2	25	25	7	—	14	7	4	84
7	1	22	26	6	2	9	10	4	80
8	—	12	12	6	—	4	3	1	38
Totals	28	304	237	76	3	66	40	17	771

TABLE XII.—Maternal Grandfather and Granddaughter. 687 Cases.

Maternal Grandfather.

Granddaughter.

Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
1	1	3	7	2	—	—	1	—	14
2	8	84	35	11	—	13	6	2	159
3	11	67	76	18	—	7	15	5	199
4	7	41	40	14	—	15	11	5	133
5	—	5	2	—	—	2	1	—	10
6	4	21	32	1	—	16	5	2	81
7	4	15	14	8	1	7	19	2	70
8	—	5	5	1	—	1	5	4	21
Totals	35	241	211	55	1	61	63	20	687

TABLE XIII.—Paternal Grandfather and Grandson. 741 Cases.

Paternal Grandmother.

	Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Grandson.	1	1	2	7	1	1	2	3	—	17
	2	6	62	69	22	4	15	25	4	207
	3	4	31	95	22	1	25	33	9	220
	4	3	18	36	20	4	15	16	4	116
	5	—	—	—	1	—	—	1	1	3
	6	1	10	23	6	2	16	10	4	72
	7	3	15	10	4	1	14	13	11	71
	8	1	10	3	3	—	5	5	8	35
	Totals	19	148	243	79	13	92	106	41	741

TABLE XIV.—Paternal Grandmother and Granddaughter. 717 Cases.

Paternal Grandmother.

	Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Granddaughter.	1	—	3	3	2	—	—	6	2	16
	2	7	53	56	14	2	13	28	8	181
	3	8	35	65	22	6	29	28	7	200
	4	4	29	36	20	3	16	23	8	139
	5	1	—	3	1	—	5	1	—	11
	6	—	9	29	2	4	27	8	3	82
	7	—	10	15	12	2	10	12	7	68
	8	1	2	1	2	1	4	4	5	20
	Totals	21	141	208	75	18	104	110	40	717

TABLE XV.—Maternal Grandmother and Grandson. 756 Cases.

Maternal Grandmother.

	Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Grandson.	1	1	10	1	3	—	—	3	1	19
	2	10	68	53	23	—	24	13	13	204
	3	9	39	67	38	—	32	23	11	219
	4	3	34	19	30	—	19	8	4	117
	5	—	—	1	1	—	—	—	—	2
	6	—	20	11	10	—	24	18	1	84
	7	2	9	23	11	1	17	17	—	80
	8	—	4	6	5	—	6	3	7	31
	Totals	25	184	181	121	1	122	85	37	756

TABLE XVI.—Maternal Grandmother and Granddaughter. 739 Cases.

Maternal Grandmother.

Granddaughter.

Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
1	2	16	—	—	—	—	—	—	18
2	7	66	34	13	—	21	15	6	162
3	12	62	55	25	1	27	23	5	210
4	6	32	36	25	—	23	15	7	144
5	—	1	3	2	—	3	3	—	12
6	1	14	21	11	—	27	17	2	93
7	—	19	17	7	—	16	17	3	76
8	—	5	4	2	—	7	3	3	24
Totals	28	212	170	85	1	124	93	26	739

TABLES for the Collateral Inheritance of Eye-colour. Second Degree.

XVII.—Paternal Uncle and Nephew. 1290 Cases.

Paternal Uncle.

Nephew.

Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Total.
1	4	10	11	6	1	4	5	2	43
2	11	136	98	40	0	26	48	12	371
3	8	84	157	26	1	27	54	7	364
4	0	29	69	36	1	19	27	12	193
5	0	2	1	0	0	2	1	0	6
6	1	31	35	7	1	30	19	3	127
7	2	21	27	24	1	13	34	11	133
8	0	11	7	6	0	10	8	11	53
Total	26	324	405	145	5	131	196	58	1290

XVIII.—Paternal Uncle and Niece. 1128 Cases.

Paternal Uncle.

Niece.

Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Total.
1	2	10	6	2	0	1	6	2	29
2	7	85	61	27	0	29	26	13	248
3	5	82	126	29	1	26	43	7	319
4	2	47	73	40	1	29	40	5	237
5	1	8	1	4	0	1	5	4	24
6	1	26	35	12	1	8	42	3	128
7	1	20	25	19	0	22	26	7	120
8	0	4	2	3	0	5	3	6	23
Total	26	282	329	136	3	121	191	47	1128

TABLE XIX.—Maternal Uncle and Nephew. 1242 Cases.

Maternal Uncle.

	Tint.	1.	2.	3.	4.	5.	6.	7	8.	Totals.
Nephew.	1	1	8	13	3	0	3	4	1	33
	2	17	137	71	29	0	19	14	9	296
	3	10	128	153	26	0	29	34	3	383
	4	2	50	62	28	0	22	14	1	179
	5	0	1	0	0	0	0	1	0	2
	6	4	33	29	12	0	35	20	3	136
	7	1	33	40	11	0	26	27	2	140
	8	0	9	17	23	0	8	3	13	73
	Totals	35	399	385	132	0	142	117	32	1242

TABLE XX.—Maternal Uncle and Niece. 1434 Cases.

Maternal Uncle.

	Tint.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Niece.	1	2	15	9	2	0	2	1	0	31
	2	6	99	76	23	0	18	13	11	246
	3	18	100	108	23	0	37	36	10	332
	4	4	72	64	28	0	16	21	9	214
	5	0	14	2	3	0	8	5	0	32
	6	5	38	41	10	0	23	11	4	132
	7	1	27	25	7	0	19	14	3	96
	8	0	15	5	6	0	9	11	5	51
	Totals	36	380	330	102	0	132	112	42	1134

TABLE XXI.—Paternal Aunt and Nephew. 1186 Cases.

Paternal Aunt.

	Tints.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Nephew.	1	6	13	5	3	1	4	6	4	42
	2	19	113	83	45	—	36	29	5	330
	3	10	81	147	30	—	29	35	8	340
	4	8	28	66	38	—	18	22	11	191
	5	—	—	—	—	—	—	—	—	0
	6	3	23	35	12	1	35	10	5	124
	7	5	22	28	19	—	18	15	5	112
	8	1	4	9	8	—	6	4	13	47
	Totals	52	284	373	155	2	148	121	51	1186

TABLE XXII.—Paternal Aunt and Niece. 1149 Cases.

Paternal Aunt.

	Tints.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Niece.	1	2	11	2	3	—	2	11	2	33
	2	15	89	62	37	2	25	40	14	284
	3	12	93	119	40	3	41	26	12	346
	4	10	36	62	43	5	25	21	11	213
	5	—	5	7	—	—	1	3	—	16
	6	1	24	33	16	1	29	19	5	128
	7	2	20	28	12	2	10	22	4	100
	8	—	7	—	4	—	5	9	4	29
	Totals	42	285	313	155	13	138	151	52	1149

TABLE XXIII.—Maternal Aunt and Nephew. 1145 Cases.

Maternal Aunt.

	First.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Nephew.	1	4	8	7	3	—	2	3	1	28
	2	5	117	81	29	—	43	29	6	310
	3	1	73	132	38	—	57	43	3	347
	4	1	20	54	27	1	21	11	—	135
	5	—	—	3	2	—	1	—	—	6
	6	—	24	35	22	—	30	23	3	137
	7	—	26	29	20	1	26	25	8	135
	8	—	14	5	10	—	12	4	2	47
	Totals	11	282	346	151	2	192	138	23	1145

TABLE XXIV.—Maternal Aunt and Niece. 1058 Cases.

Maternal Aunt.

	First.	1.	2.	3.	4.	5.	6.	7.	8.	Totals.
Niece.	1	2	3	10	—	—	—	—	—	15
	1	5	87	86	31	—	23	14	12	258
	3	3	71	125	32	1	49	41	3	325
	4	—	39	51	31	1	33	19	6	180
	5	—	4	6	1	—	8	5	3	27
	6	1	25	47	10	—	24	9	2	118
	7	—	30	29	11	—	14	18	10	112
	8	—	5	4	5	—	2	4	3	23
	Totals	11	264	358	121	2	153	110	39	1058

[Notes added July 3, 1900.

Note I. *Inheritance of Temper and Artistic Instinct.*—In additional to the fraternal correlations given on p. 102, I have dealt with Mr. GALTON'S statistics for the inheritance of good and bad temper given in his 'Natural Inheritance' (p. 235). The following gives the distribution of good and bad temper among 1,294 brethren, as deduced by Mr. YULE.

		First Brother.		
		Good Temper.	Bad Temper.	Totals.
Second Brother.	Good temper.	330	255	585
	Bad temper.	255	454	709
	Totals	585	709	1294

The correlation is ·3167.

A like table is that for artistic instinct in the direct line :—

		Parentage.		
		Artistic.	Non-artistic.	Totals.
Child.	Artistic.	296	173	469
	Non-artistic.	372	666	1038
	Totals	668	839	1507

In this case the correlation is ·4039.

The fraternal correlation is somewhat low. The exact significance of the parental correlation is also somewhat vague, as the parentage is classified as artistic when *one* or *both* parents are artistic. But the two tables are very suggestive, they indicate how the new method will enable us to deal quantitatively even with characters like temper and artistic instinct to which it is impossible to apply directly a quantitative scale. With the introduction of a third or medium class, I believe it will be possible to obtain excellent results for heredity from very simple observations, and I have in hand at the present time a large series of observations on collateral heredity based upon such simple classifications. The reader should further consult

Mr. G. U. Yule's remarks on the association of temper and of artistic instinct in his memoir on "Association," 'Phil. Trans.,' A, vol. 194, p. 290, 1900.

Note II. *On the Correlation of Fertility and Eye-Colour.*—In the course of the present paper I have frequently referred to a probable influence of reproductive selection as the source of the progressive change in eye-colour, *i.e.*, to a possibility that eye-colour is correlated with fertility. I saw from Mr. Galton's tables that in many cases the whole family had not been recorded, probably the eye-colour of the dead or of absentees being unknown. It appeared to me accordingly that it would be impossible to deal directly with the problem of fertility. However, it has since occurred to me that there is nothing likely to give the missing members of families a bias towards one rather than another eye-colour, and that we may simply treat them as a purely random subtraction from the total results. Assuming this, Mr. L. N. Filon, M.A., has prepared for me tables of father's and mother's eye-colour and of the recorded number of their children. From these* I take first the following results, premising (i.) that I call "light eyed," persons with eye-colours 1, 2 and 3, and "dark eyed," persons with eye-colours 4, 5, 6, 7, 8, *i.e.*, drawing the line between light and dark grey; (ii.) that I take as small families those with 0, 1, 2, or 3, recorded children and as large those with 4 or more recorded children.

Father.

		Light Eyed.	Dark Eyed.	Totals.
Family.	Small.	313	141	454
	Large.	264	139	403
Totals		577	280	857

Mother

		Light Eyed.	Dark Eyed.	Totals.
Family.	Small.	253	202	455
	Large.	225	169	394
Totals		478	371	849

---

* Correlation tables were prepared of the size of families 0 to 15, and of the eye-colours 1 to 8, but it does not seem needful to print them *in extenso*.

We have, accordingly, by the method of the present memoir :—

Correlation of size of family with darkness of eye-colour

$$= \cdot 0595, \text{ for fathers.}$$
$$= - \cdot 0239, \text{ for mothers.}$$

The former is just sensible, the latter hardly sensible relative to the probable error. So far as they can be relied upon, they would denote that fathers have more children the darker eyed they are, and mothers more children the lighter eyed they are. This is in accordance with the result given in the memoir, that the modern generation is darker than its male and lighter than its female ancestry, but it is not the explanation given in the text, although it is probably the true one. If it be the true one, dark fathers and light mothers ought to present the most fertile unions, and it seemed desirable to test this directly. We have already seen that there exists an assortative mating in eye-colour, like tending to mate with like, the co-efficient of correlation being about ·1 ; hence if we were to correlate the eye-colour of mothers and fathers, *i.e.*, husbands and wives weighted with their fertility, we ought to find this result substantially reduced. The following is the table :—

		Fathers.		
		Light Eyed.	Dark Eyed.	Totals.
Mothers.	Light Eyed.	1183	612	1795
	Dark Eyed.	826	455	1281
	Totals	2009	1067	3076

We find $r = \cdot 0239$, or the correlation has been reduced to a fifth of its previous value, and is now of the order of its probable error. To mark still further this increased fertility of heterogamous unions, I add two further tables, giving the mean number of recorded offspring for various classifications of parental eye-colour.

		Fathers.		
		Colours 1–3.	Colours 4–8.	Average of totals.
Mothers.	Colours 1–3.	3·83	4·57	4·07
	Colours 4–8.	3·82	3·73	3·79
	Average of totals	3·83	4·17	3·94

		Fathers.		
		Colours 1–2.	Colours 3–8.	Average of totals
Mothers.	Colours 1–2.	3·19	4·52	3·86
	Colours 3–8.	3·98	3·96	3·97
	Average of totals	3·68	4·09	3·94

The first table entirely confirms all the conclusions reached,—dark fathers and light mothers are most fertile absolutely and in union. The second table shows that it is the blue-green and grey rather than pure blue-eyed mothers who are most fertile. This supplementary investigation accordingly seems to support the view of the text of the memoir, namely that reproductive selection is the source of the secular change in eye-colour noted, only the prepotent fertility which is replacing the blue-eyed element is in the first place that of the dark-eyed male, and only in the second place due to mothers having eye-colours dark or light other than true blue. We seem accordingly in eye-colour to have reproductive selection working through heterogamy rather than through homogamy as in the case of stature.[*] The effect, however, is like,—the progressive elimination of one type of character.]

---

[*] See 'Roy. Soc. Proc.,' vol. 66, p. 30, and vol. 66, p. 316 *et seq.*

Sold by Harrison and Sons, St. Martin's Lane.

Demy 8vo, pp. 224. Price 7*s*. 6*d*.

THE RECORD OF THE ROYAL SOCIETY, 1897.

Containing an Account of the Foundation, History, &c., of the Society.

Demy 8vo, pp. 209. Price 5*s*.

THE YEAR-BOOK OF THE ROYAL SOCIETY, 1900.

Containing an Account of the Chief Events and Arrangements of the Year.

PHILOSOPHICAL TRANSACTIONS.

Series A. Containing Papers of a Mathematical or Physical Character.

Vol. 193 (1900), Price £1 10*s*., with 28 Plates.

Vol. 194 (1900), Price £1 1*s*. 6*d*., with 5 Plates.

Series B. Containing Papers of a Biological Character.

Vol. 190 (1898), Price £2 5*s*., with 24 Plates.

Vol. 191 (1899), Price £1 12*s*., with 31 Plates.

Published by Clay and Sons, Ave Maria Lane.

CATALOGUE OF SCIENTIFIC PAPERS,

COMPILED BY THE ROYAL SOCIETY.

The volumes on sale as follows:—Vols. 1–6 (1800–63), cloth (vol. 1 in half-morocco) £4 net; half-morocco, £5 5*s*. net. Vols. 7, 8 (1864–73), cloth, £1 11*s*. 6*d*. net; half-morocco, £2 5*s*. net. Single vols., cloth, 20*s*.; half-morocco, 28*s*. net. Vols. 9, 10 (1874–83), cloth, £1 5*s*. net; half-morocco, £1 12*s*. net. Vol. 11 (1874–83), cloth, £1 5*s*. net; half-morocco, £1 12*s*. net.

A reduction of price to Fellows of the Royal Society.

Published by Kegan Paul, Trench, Trübner and Co.

A few copies at the reduced price of £1. 1*s*. (published at £3).

Royal 4to, pp. iv.—936 cloth.

A MONOGRAPH OF THE HORNY SPONGES.

By R. von LENDENFELD.

With 51 Lithographic and Photographic Plates.

A further reduction of price to Fellows of the Royal Society.

Published by Kegan Paul, Trench, Trübner, and Co.

In 1 vol., 4to. Pp. 500. With 6 Chromolithographs of the remarkable Sunsets of 1883 and 40 Maps and Diagrams.

THE ERUPTION OF KRAKATOA AND SUBSEQUENT PHENOMENA.

Report of the Krakatoa Committee of the Royal Society.

Edited by G. J. SYMONS, F.R.S.

Price 30*s*. To Fellows of the Royal Society, 20*s*.

Published by Kegan Paul, Trench, Trübner, and Co.

Royal 4to, pp. xiv.—326, cloth. Price 21*s*.

OBSERVATIONS OF THE INTERNATIONAL POLAR EXPEDITIONS.

1882–1883.

FORT RAE.

With 32 Lithographic Folding Plates.

A reduction of price to Fellows of the Royal Society.

Continental Agents for the publications of the Royal Society, Messrs. Friedländer and Son, Berlin.

Just Published, and Sold by Dulau and Co., Soho Square.

SEPARATE PAPERS FROM THE 'PHILOSOPHICAL TRANSACTIONS.'

1899–1900.

A. 253. A Comparison of Platinum and Gas Thermometers, including a Determination of the Boiling-point of Sulphur on the Nitrogen Scale.—An Account of Experiments made in the Laboratory of the Bureau International des Poids et Mesures at Sèvres. By John Allen Harker, D.Sc., formerly Bishop Berkeley Fellow of Owens College, Manchester, Research Assistant at Kew Observatory, and Pierre Chappuis, Ph.D., Savant Attaché au Bureau International des Poids et Mesures, Sèvres. Price 5*s*.

A. 254. On the Propagation of Earthquake Motion to Great Distances. By R. D. Oldham, Geological Survey of India. Price 1*s*. 6*d*.

A. 255. Impact with a Liquid Surface studied by the Aid of Instantaneous Photography. Paper II. By A. M. Worthington, M.A., F.R.S., and R. S. Cole, M.A. Price 2*s*. 6*d*.

A. 256. Gold-aluminium Alloys. By C. T. Heycock, F.R.S., and F. H. Neville, F.R.S. Price 2*s*. 6*d*.

A. 257. Bakerian Lecture.—The Specific Heats of Metals, and the Relation of Specific Heat to Atomic Weight. By W. A. Tilden, D.Sc., F.R.S., Professor of Chemistry in the Royal College of Science, London. With an Appendix by Professor John Perry, F.R.S. Price 1*s*.

A. 258. On the Association of Attributes in Statistics: with Illustrations from the Material of the Childhood Society, &c. By G. Udny Yule, formerly Assistant Professor of Applied Mathematics, University College, London. Price 3*s*.

A. 259. The Ionization of Dilute Solutions at the Freezing Point. By W. C. D. Whetham, M.A., Fellow of Trinity College, Cambridge. Price 2*s*. 6*d*.

A. 260. Combinatorial Analysis.—The Foundations of a New Theory. By Major P. A. MacMahon, D.Sc., F R.S. Price 1*s*. 6*d*.

A. 261. Title, Contents, Index, &c., Vol. 194. Price 6*d*.

A. 262. Mathematical Contributions to the Theory of Evolution.—VII. On the Correlation of Characters not Quantitatively Measurable. By Karl Pearson, F.R.S. Price 2*s*. 6*d*.

A. 263. Electrical Conductivity in Gases traversed by Cathode Rays. By J. C. McLennan, Demonstrator in Physics, University of Toronto. Price 1*s*. (*In the press.*)

B. 180. Agricultural, Botanical, and Chemical Results of Experiments on the Mixed Herbage of Permanent Grass-land, conducted for many years in succession on the same Land.—Part III. The Chemical Results—Section I. By Sir John Bennet Lawes, Bart., D.C.L., Sc.D., F.R.S., and Sir J. Henry Gilbert, LL.D., Sc.D., F.R.S. Price 3*s*. 6*d*.

B. 181. On Intestinal Absorption, especially on the Absorption of Serum, Peptone, and Glucose. By E. Waymouth Reid, F.R.S., Professor of Physiology in University College, Dundee, St. Andrews University, N.B. Price 4*s*.

B. 182. The External Features in the Development of *Lepidosiren paradoxa*, Fitz. By J. Graham Kerr. Price 7*s*.

B. 183. On the Formation of the Pelvic Plexus, with especial reference to the Nervus Collector in the Genus *Mustelus*. By R. C. Punnett, B.A., Scholar of Gonville and Caius College, Cambridge. Price 2*s*.

B. 184. Title, Contents, Index, &c., Vol. 192. Price 6*d*.

B. 185. On the Gastric Gland of Mollusca and Decapod Crustacea: its Structure and Functions. By C. A. MacMunn, M.A., M.D. Price 5*s*.

B. 186. Upon the Development of the Enamel in Certain Osseous Fish. By Charles S. Tomes, M.A., F.R.S. Price 1*s*.

B. 187. Extinct Mammalia from Madagascar.—I. *Megaladapis insignis*, sp. n. By C. I. Forsyth Major, M.D., F.Z.S. Price 1*s*.

B. 188. Further Observations on the Nature and Functions of the Nodules of Leguminous Plants. By Maria Dawson, B.Sc. (London and Wales), 1851 Exhibition Science Research Scholar. Price 2*s*.

Cloth cases for binding the 'Philosophical Transactions' can be obtained of Harrison and Sons, St. Martin's Lane. Price 1*s*. 6*d*.

PHILOSOPHICAL TRANSACTIONS

OF THE

ROYAL SOCIETY OF LONDON.

SERIES A.

Vol. 197, pp. 285–379.

MATHEMATICAL CONTRIBUTIONS TO THE THEORY OF EVOLUTION.—IX. ON THE PRINCIPLE OF HOMOTYPOSIS AND ITS RELATION TO HEREDITY, TO THE VARIABILITY OF THE INDIVIDUAL, AND TO THAT OF THE RACE.

PART I. HOMOTYPOSIS IN THE VEGETABLE KINGDOM.

BY

KARL PEARSON, F.R.S.,

WITH THE ASSISTANCE OF

ALICE LEE, D.Sc., ERNEST WARREN, D.Sc., AGNES FRY,
CICELY D. FAWCETT, B.Sc., and others.

LONDON:
PUBLISHED FOR THE ROYAL SOCIETY
BY DULAU AND CO., 37 SOHO SQUARE, W.
CONTINENTAL AGENTS, MESSRS. FRIEDLÄNDER AND SON, BERLIN.
1901.

A. 294.

Price Five Shillings and Sixpence.

12.11.01.

INDEX SLIP.

PEARSON, Karl (and others).—Mathematical Contributions to the Theory of Evolution.—IX. On the Principle of Homotyposis and its Relation to Heredity, to the Variability of the Individual, and to that of the Race. Part I.—Homotyposis in the Vegetable Kingdom.
Phil. Trans., A, vol. 197, 1901, pp. 285–379.

Heredity—Relation to Homotyposis, to Variability of Species, to Evolutionary Descent.
PEARSON, Karl (and others).
Phil. Trans., A, vol. 197, 1901, pp. 285–379.

Homotyposis—Principle of, Relation to Heredity and Variability.
PEARSON, Karl (and others).
Phil. Trans., A, vol. 197, 1901, pp. 285–379.

Variability—in the Individual and in the Race—Relation to Homotyposis and Heredity.
PEARSON, Karl (and others).
Phil. Trans., A, vol. 197, 1901, pp. 285–379.

Vegetable Kingdom—Homotyposis in.
PEARSON, Karl (and others).
Phil. Trans., A, vol. 197, 1901, pp. 285–379.

VIII. *Mathematical Contributions to the Theory of Evolution.—IX. On the Principle of Homotyposis and its Relation to Heredity, to the Variability of the Individual, and to that of the Race. Part* I.—*Homotyposis in the Vegetable Kingdom.*

By Karl Pearson, *F.R.S.*,

With the assistance of Alice Lee, *D.Sc.*, Ernest Warren, *D.Sc.*, Agnes Fry, Cicely D. Fawcett, *B.Sc.*, *and others.*

Received October 6,—Read November 15, 1900.

Contents.

Page

I. Introductory.

II. Theory.

III. Observation. Measurement of Variability and Correlation of undifferentiated like Organs in the Individual.

IV. Actual Data.

12.11.1901.

I. Introductory.

(1.) The present paper endeavours to deal with a problem upon which I have long been occupied, adopting the widest basis compatible with the time and means at my disposal. In the first place, I have often been impressed with the small reduction in variability which can be produced by selection. The offspring of a single parent while diverging in character, possibly very widely from the average character of the race, will still have a variability in that character only slightly reduced, say at most 10 per cent. below the racial variability. Even if we select the ancestry for an indefinite number of generations, the offspring will have a variability upwards of 89 per cent. of that of the original race.* Now this capacity in the parent for producing variable offspring must be in some manner related to the degree of resemblance in those offspring. We have thus the two fundamental divisions of our subject:

(i.) What is the ratio of individual to racial variability?

(ii.) How is the variability in the individual related to inheritance within the race?

I must endeavour to explain my meaning a little more fully and clearly. The individual puts forth a number of like organs, corpuscles in the blood, petals of the flower, leaves of the trees, scales on the wing. These may or may not be divided up into differentiated groups. Special forms of leaves occur in the neighbourhood of the fruit; florets may be differentiated according to their position on the flower, scales

---

* "On the Law of Ancestral Heredity," 'Roy. Soc. Proc.,' vol. 62, p. 400.

according to their position on the wing; there may be two or more classes of blood-corpuscles. But if we take organs which are like, and so far as we can judge undifferentiated, there is still a difference between them. The individual in putting forth what I may perhaps term "undifferentiated like organs,"* does so with a certain measure of variability. I am not concerned at present with the source of this variability. It may be due to the individual environment, to the physique of the individual, or to some inherent or bathmic tendency in the individual due to its ancestral history. All I am concerned with at present is, that the undifferentiated like organs of an individual possess a certain variability, and that this variability is somewhat less than that of all like organs in the race. If, however, the variability were entirely due to *external* circumstances attending growth, we should hardly expect to find leaves gathered from different branches and on different sides of a tree more alike to each other than to leaves of other trees of the same race. But this is indeed the case; there is a considerable correlation among the undifferentiated like organs of an individual notwithstanding the small reduction on racial variability. Among the trees and plants considered in this paper we have data enough to enable us to determine whether, say, a hundred leaves placed before us were gathered from a single tree, or collected at random from a number of trees. Here, of course, I mean by to "determine" to state the odds for or against these two assumptions.

Just as we can find by the methods already discussed in earlier memoirs of this series, the degree of correlation between brothers and the variability of an array of brothers due to the same parentage, so we can determine the correlation, *i.e.*, the degree of resemblance between the undifferentiated like organs of the individual and the degree of variability within the individual. This determination is the answer to our first fundamental problem, that of the ratio of individual to racial variability. But turning to the process of reproduction, the offspring depend upon the parental germs, and it would thus seem that the degree of resemblance between offspring

---

* [I am fully conscious of *verbal* difficulty in the phrase "undifferentiated like organs," but I believe that the distinction between the differentiated and the undifferentiated is quite clear either from the standpoint of observation or from that of the frequency distribution itself. Differentiation, whether due to function, position on the individual, season of production, &c., is usually connected with one or two well-marked dominating factors; it is statistically discoverable by testing the frequency distribution for heterogeneity. On the other hand, variability in "undifferentiated like organs" is not to be associated with any one or two dominating factors which can be isolated; it is due to that combination of many small causes, inherent and environmental, which leads to what is familiar in both theory and observation as a homogeneous *chance* distribution. A diversity due to differentiation and a variability due to chance are quite distinct things. The one is the result of dominating factors which can be isolated and described; the other of a great number of small factors, varying from organ to organ, and incapable of being defined or specified. Indeed, upon each dominating factor of differentiation is superposed such a chance variability. Of course all things which differ even by chance variation are in a certain sense differentiated. But the term differentiation is throughout this paper reserved for the differences which arise, not from a multiplicity of small causes, but from dominating and usually easily recognisable single influences.—*July*, 1901.]

must depend on the variability of the sperm cells and the ova which may each be fairly considered as "undifferentiated like organs." Here again we are not compelled to assert that much or little is due to environment and little or much is due to inherent ancestral influence. All we assume is that such causes as produce the likeness between leaves of the same tree, or florets on the same flower, produce the likeness between spermatozoa or ova of the same individual, and that on this likeness the ultimate resemblance of offspring from the same parent depends. We have then to investigate how the quantitative resemblance between offspring of the same parents is related to the quantitative resemblance between the undifferentiated like organs in the individual; and then we must test on as wide a range of data as possible this theoretical relationship.

Now the reader will perceive at once that if we can throw back the resemblance of offspring of the same parents upon the resemblance between the undifferentiated like organs of the individual, we shall have largely simplified the whole problem of inheritance. Inheritance will not be a peculiar feature of the reproductive cells. One frog, let us say, differs from another in that it produces blood corpuscles more or less alike and unlike those of another frog. In the simplest forms of reproduction, budding and parthenogenesis, the offspring will not be absolutely alike, for buds and ova are undifferentiated like organs, and such organs have only a limited degree of resemblance. If this view be correct, variability is not a peculiarity of sexual reproduction, it is something peculiar to the production of undifferentiated like organs in the individual, and the problems of heredity must largely turn on how the resemblance between such organs is modified, if modified at all, by the conditions of nurture, growth, and environment generally. Our discussion of the subject will naturally break up into the following divisions:

(*a*.) On the theoretical relationship between the correlation of offspring and the correlation of undifferentiated like organs in the individual.

(*b*.) A determination of the correlation of undifferentiated like organs for as wide a range of life as possible.

The present paper deals only with variation and correlation in the vegetable kingdom, but the data for variation and correlation in the animal kingdom are being collected at the same time.

(*c*.) A comparison of the degree to which the results obtained from (*b*) satisfies the observed values of fraternal correlation already found when we use the relation determined in (*a*).

II. THEORETICAL DISCUSSION OF THE RELATION OF FRATERNAL CORRELATION TO THE CORRELATION OF UNDIFFERENTIATED LIKE ORGANS.

(2.) Let z be the deviation from the mean of the general population of any individual with regard to any character. Let us suppose z to depend upon certain characters in the spermatozoon and certain characters in the ovum from which the individual has developed. These characters cannot of course be determined, still less measured, but we have no reason to doubt their existence. In the particular spermatozoon from which the individual has developed, let them have deviations $x_1, x_2, x_3 \ldots$ from their mean values for all the spermatozoa of the race, and let $y_1, y_2, y_3 \ldots$ be the corresponding deviations for the ovum characters. Then

$$z = f(x_1, x_2, x_3 \ldots y_1, y_2, y_3 \ldots)$$

where f is a quite unknown function.

The mean of the z-character will, however, correspond to the mean values of the spermatozoon and ovum characters, and if we suppose the variation of these characters small as compared with their mean value, we assume as usual for such deviations:

$$z = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots + \beta_1 y_1 + \beta_2 y_2 + \beta_3 y_3 + \ldots \quad \text{(i.)},$$

where the α's and β's are independent of the x's and y's, and define the male and female inheritance.

Now let σ be the standard deviation of the character z in the population; σ_q the standard deviation of x_q, σ'_q of y_q. Let r_{pq} be the correlation of x_p and x_q, r'_{pq} of y_p and y_q. Then we will suppose that there is no selection of particular ova by particular spermatozoa, or that x_p and y_q are not correlated. Then if n = number of individuals in the population:

$$\sigma^2 = \frac{S(z^2)}{n} = \Sigma\alpha_p^2 \frac{S(x_p^2)}{n} + 2\Sigma(\alpha_p\alpha_q)\frac{S(x_p x_q)}{n} + \Sigma\beta_p^2 \frac{S(y_p^2)}{n} + 2\Sigma(\beta_p\beta_q)\frac{S(y_p y_q)}{n},$$

where S is the sum for all individuals of any x or y for constant subscript, and Σ is the sum of α and β for every possible subscript. This follows by simple squaring and remembering that $S(x_p y_q) = 0$. We thus reach:

$$\sigma^2 = \Sigma(\alpha_p^2\sigma_p^2) + \Sigma(\beta_p^2\sigma'^2_p) + 2\Sigma(\alpha_p\alpha_q\sigma_p\sigma_q r_{pq}) + 2\Sigma(\beta_p\beta_q\sigma'_p\sigma'_q r'_{pq}). \quad \text{(ii.)}.$$

Now let us consider the correlation of two individuals due to the spermatozoa and ova put forth by the same two individuals. Let z_1 and z_2 be the values of their characters, and x', x'', y', y'' represent the fundamental characters in the two spermatozoa, and two ova on which they depend.

Then we have

$$z_1 = \Sigma(\alpha_p x'_p) + \Sigma(\beta_p y'_p)$$
$$z_2 = \Sigma(\alpha_p x''_p) + \Sigma(\beta_p y''_p).$$

Now let us multiply z_1 by z_2 and sum for every fraternal pair; then if R be the fraternal correlation we have, since one set of brethren will have the same standard deviation as the other, if they be taken of the same sex,

$$S(z_1 \times z_2) = n \times \sigma \times \sigma \times R,$$

where n is the total number of pairs of brethren.

Hence

$$n \times \sigma^2 \times R = \Sigma(\alpha_p^2 (x'_p x''_p)) + \Sigma(\alpha_p \alpha_q S(x'_p x''_q + x''_p x'_q)) + \Sigma(\beta_p^2 S(y'_p y''_p)) + \Sigma(\beta_p \beta_q S(y'_p y''_q + y''_p y'_q))$$

with the same notation as before and the same assumption as to no selective process between the female and male reproductive cells.

Now x'_p, x''_p are undifferentiated like organs put forth by an individual. If ρ_{pp} be the correlation of such :

$$S(x'_p x''_p) = n\sigma_p^2 \rho_{pp},$$

Similarly

$$S(y'_p y''_p) = n\sigma'^2_p \rho'_{pp}.$$

In the next place we should have *all* pairs of brothers :

$$S(x'_p x''_q) = S(x''_p x'_q) = \sigma_p \sigma_q \rho_{pq},$$

$$S(y'_p y''_q) = S(y''_p y'_q) = \sigma'_p \sigma'_q \rho'_{pq}.$$

We thus deduce

$$\sigma^2 \times R = \Sigma(\alpha_p^2 \sigma_p^2 \rho_{pp}) + \Sigma(\beta_p^2 \sigma'^2_p \rho'_{pp}) + 2\Sigma(\alpha_p \alpha_q \sigma_p \sigma_q \rho_{pq}) + 2\Sigma(\beta_p \beta_q \sigma'_p \sigma'_q \rho'_{pq}) \quad . \quad . \quad . \quad . \quad . \quad \text{(iii.)}.$$

Now if the degree of resemblance between undifferentiated like organs vary largely from organ to organ, we could proceed no further with this investigation. We should have shown that there was a relation, namely (iii.) between the R and the ρ's, but as the characters in the spermatozoa and ova are unknown and unmeasurable, we could not proceed further. On the other hand, if the correlation between undifferentiated like organs is approximately constant, then we may replace correlation coefficients like ρ_{pp} and ρ'_{pq} by their *mean* value ρ, say.

We have next to consider such correlations as ρ_{pq} or ρ'_{pq}. These are the two correlations between two different characters in two spermatozoa or ova from the same individual. Such correlations must vanish (i.) if there be no organic correlation between characters of the same spermatozoon or ovum, (ii.) if there be no correlation between the reproductive cells put forth by the same individual. In other words, ρ_{pq} must contain ρ and r_{pq} as factors, and since if these two correlations are perfect ρ_{pq} must be perfect, it is a reasonable hypothesis to assume

$$\rho_{pq} = \rho \times r_{pq} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv.)}.$$

This, it will be seen, is identical in form with the result I have supposed to hold for cross-heredity,[*] and which appears to be approximately true for that case. We shall consider later statistics bearing on this result.

Assuming accordingly that (iv.) holds, we find from (iii.)

$$\begin{aligned}\sigma^2 \times \mathrm{R} &= \rho\{\Sigma(\alpha_p{}^2\sigma_p{}^2) + \Sigma(\beta_p{}^2\sigma'_p{}^2) + 2\Sigma(\alpha_p\alpha_q\sigma_p\sigma_q r_{pq}) + 2\Sigma(\beta_p\beta_q\sigma'_p\sigma'_q r'_{pq})\} \\ &= \rho \times \sigma^2 \text{ by (ii.).}\end{aligned}$$

Hence we conclude that

$$\mathrm{R} = \rho \quad \text{. . . . . (v.).}$$

Or: *The correlation between brothers will be equal to the mean correlation between the undifferentiated like organs put forth by an individual.*

Now, if this result be true, it is very remarkable and very fundamental. We should hardly expect it to be absolutely true, for it is very unlikely that the coefficient of correlation between undifferentiated like organs is the same whatever the organs may be. This equality may rather be spoken of as belonging to an ideal theoretical vital state approaching the actual state, perhaps, as BOYLE'S Law or the perfect fluid approach phenomena observed in physical nature. What we should expect would be a general approximation between the values of R and ρ, and a tendency to equality when large series are averaged. This is the point which we shall investigate in the sixth section of this paper, after placing before the reader in the fourth section a fairly wide[†] range of actual statistics.

It will be seen at once that if (v.) represents an approximate or average truth of living nature, then the problem of inheritance is to a large extent the same as the problem of variability in the individual. We have not answered, of course, even in part, the problem of why two brothers resemble each other, but we see that it is part of a much more general problem having nothing to do with sexual reproduction. It is one with the problem of the likeness in leaves of the same tree, or the likeness in scales on the same spot of a moth's wing. It brings the problem of heredity into closer touch with the problem of variability. When we ascertain the sources of variation in the individual, then we shall have light on the problem of fraternal resemblance.

III. ON THE VARIABILITY AND CORRELATION OF UNDIFFERENTIATED LIKE ORGANS IN THE INDIVIDUAL.

(3.) I must frankly admit that I have collected my material from the standpoint of the mathematical statistician and not of the trained field naturalist. I have sought things which were easy to count and measure, and endeavoured to avoid "differentia-

---

[*] "The Law of Ancestral Heredity," 'Roy. Soc. Proc.,' vol. 62, p. 411.

[†] Not wide when we regard the natural range of living forms, but from the standpoint of the labour that has been spent on the collection.

tion" only by the roughest tests and by largely untrained powers of observation. I have examined my material to see if it gave obvious signs of heterogeneity in the tabulated bulk. I have examined my individuals and their organs for obvious outward signs of differentiation. But I have not studied the morphological evolution of the organs considered, or questioned whether the parts counted were all due to the same source. Hence it is quite possible that the botanist may reject at once some of my series. What I have endeavoured to do is to take as wide a range of as different organs as possible in different types of life and trust to the bulk of my statistics to give me a substantially accurate average value of ρ to compare with the values of R we have determined on other occasions. At the same time the material here presented does not by any means exhibit all the trials made; we often enough at the first attempt did not get a suitable character to measure or count, or again the individuals collected were occasionally insufficient. Nothing, however, has been omitted which was unfavourable to the conclusions ultimately drawn. Thus the *Nigella Hispanica* statistics have been retained, although there is not the slightest doubt of a very considerable differentiation in the flowers growing at different parts of the plant.

When such a differentiation takes place the result will generally be a great reduction in the correlation; for "like" organs on the same individual, say A and *a* if differentiated, will really be less closely related to each other than to B and *b* respectively, the corresponding organs in a second individual. On the other hand a heterogeneity of material, say a mixture of two different local races, will tend as a rule to raise the correlation,[*] for it generally amounts to compounding two very like correlation surfaces with the mean of one approximately shifted along the regression line of the second. If we bear these two opposing influences in mind, differentiation tending to reduce, heterogeneity to increase the actual degree of correlation of undifferentiated like organs, it will not seem incompatible with an actual approximate constancy of such correlation to find a fairly wide range of values in our statistics. We set them forth as the first rough attempt to appreciate the resemblance of like parts within the individual. To the specialist in the future must be left the work of selecting, with special knowledge, truly homogeneous material and absolutely undifferentiated characters, and thus obtaining the required correlation to a much higher degree of accuracy.

(4.) It remains to explain the process by which the correlation was deduced. Let us take as an illustration beech-leaves. One hundred trees fairly of the same age and belonging to the same district, were selected, and twenty-six leaves specified by the letters of the alphabet were gathered from each of these. The leaves were gathered so far as possible all round the outside of the tree, roughly about the same height from the ground, and scattered over different parts of the individual boughs. Thus each tree was supposed to be individualised by twenty-six leaves. The veins on these leaves were then counted, and varied for beech-leaves in general between ten and twenty-two.

---

[*] See 'Phil. Trans.,' A, vol. 192, p. 277.

All the possible pairs were now taken, *i.e.*, $\frac{1}{2}(26 \times 25) = 325$ in number, and entered on a correlation table in the usual manner, the two variables being the number of veins in the first leaf and the number of veins in the second leaf. But as either member of the pair might be a "first" leaf, the table so formed was rendered symmetrical by starting with either leaf in the pair as first or second. Thus a single tree led to 650 entries in the correlation table, or with 100 trees there were 65,000 entries. This large number must not lead the reader to overweight the importance of the constants calculated upon it. There were only 100 trees leading to 32,500 pairs of leaves, each pair coming from the same tree. I should have much preferred a thousand trees, but the great labour of collecting, counting, and calculating precluded any such number. In many cases also it would have been practically impossible to have obtained 100 individuals growing under fairly like environment. For example, I only succeeded in approximating to 100 Spanish chestnuts from one district and near one age. When I took a second hundred, gathered from several districts, partly old trees and partly others of a pollard growth of fifteen to twenty years, I found the heterogeneity at once increased the correlation (see below, p. 301). In many cases, of course, it was impossible to obtain twenty-six undifferentiated like organs from the same individual. In these cases the pairs were formed in the same manner, but in some series the total number due to each individual varied very considerably, and accordingly the work of verifying the tables was much increased. When but few pairs could be obtained from each individual, we have sometimes increased the number of individuals dealt with up to a couple of hundred. But the labour of dealing even with a hundred individuals is often—for example, in the case of mushrooms and onions—very serious. Had it not been for the generous help of a number not only of willing but of competent collectors and calculators, the material here dealt with would have taken not eighteen months but years of my own unaided efforts.

The calculation of the means, standard deviations, probable errors, and correlations of each table was carried out in the manner sufficiently discussed in earlier papers of this series.* The symmetry of the tables leads to slight simplifications in calculating the product moment which will readily suggest themselves to the reader, and of course only a single mean and standard deviation is required for each table. Tests for the accuracy of the last two constants are at once provided in the case in which the same number of organs are taken from each individual, for their values must be identical with those obtained for the whole series of organs entered only once and not for each possible pair.

(5.) The quantitative measurement of the degree of resemblance between undifferentiated-like organs being, so far as I am aware, a quite novel branch of investigation, I venture, with some hesitation, to introduce certain terms to describe oft-

---

* See especially Memoir III., 'Phil. Trans.,' A, vol. 187, pp. 253–318, and Memoir IV., 'Phil. Trans.,' A, vol. 191, pp. 229–311.

recurring ideas. I shall call undifferentiated like organs *homotypes*, indicating that they are types produced by the same mould or individual. Thus two leaves of the same tree, or two blood-corpuscles from the same frog are *homotypes*; their resemblance will be homotypic, and the character by which their resemblance is quantitatively measured will be the *homotypic* character. I shall speak of *homotypic* correlation and distinguish it from *organic* correlation—although, of course, the former is in a certain sense organic. By organic correlation I refer to the correlation between two different characters in the same organ; by homotypic correlation to the correlation between the same or different characters in a pair of homotypes. If the same character, then the correlation is *direct homotypic correlation*; if different characters, then the correlation is *cross homotypic correlation*. Lastly, the principle that homotypes are correlated, *i.e.*, that variation within the individual is less than that of the race, or that undifferentiated like organs have a certain degree of resemblance, I shall speak of as *homotyposis*. Thus homotyposis denotes not only likeness of the homotypes, but that this likeness has probably definite quantitative limits. If my view be correct, heredity is only a special case of homotyposis, any multiplication of cells is homotypical, and denotes a given degree of variation and a given degree of likeness. This does not "explain" heredity, but shows it merely as a phase of a much wider natural process.

IV. ACTUAL DATA.

SECTION I.—*Leaves of Trees.*

(6.) The leaves of trees provide material for fairly easy computation without measurement. Thus we may count the veins on the leaf or the leaflets on the compound leaf, and ascertain the degree of resemblance between leaves of the same tree. The variation is, however, in some cases very considerable, and the labour of forming the tables involving thousands of entries very great. Still the leaf-series are some of my best, and considered as a whole, perhaps the most satisfactory.

(7.) A. *Ash* (Fraxinus excelsior). *Number of* Pinnæ *on Leaf.*—My first series (i.) consists of 26 leaves taken from each of 109 trees by Dr. ALICE LEE or myself. The trees were nearly all large old trees, growing on the commons of Great and Little Hampden, Buckinghamshire. The leaves were taken as far as possible from a variety of branches all round the tree, and from different points on these branches. The labour of tabling and deducing the constants is due to Dr. LEE. The number of pairs = 70,850.

My second series (ii.) consist of 26 leaves taken from each of 120 trees by Miss C D. FAWCETT, B.Sc. These trees, many of which were young trees growing in the hedgerows, were taken in the neighbourhood of Lyme Regis in Dorsetshire. The same rules as to gathering were observed. Dr. LEE again undertook the labour of tabling and the deduction of the constants. The number of pairs = 78,000.

My third series consists of 26 leaves from each of 100 trees in Monmouthshire, young trees being included, but the same rules as to gathering being observed. The whole of the labour of gathering, counting and deducing the constants is due to Miss MILDRED E. BARWELL, of Girton College.*

The general numerical results are given in the table below :—

Resemblance of Ash Leaves from same Tree.

Series.	Number.			Mean No. of leaflets.	S. D. of leaflets.	Correlation.
	Trees.	Leaves.	Pairs.			
Buckinghamshire	109	2834	70850	10·1295 ± ·0214	1·6891 ± ·0151	·3743 [± ·0109]†
Dorsetshire . .	120	3120	78000	9·7260 ± ·0239	1·9759 ± ·0169	·3964 [± ·0102]
Monmouthshire .	100	2600	65000	9·8766 ± ·0265	2·0058 ± ·0188	·4047 [± ·0111]
Mean of series .	—	—	—	9·9107	1·8903	·3918

Thus a total of 329 trees, giving 8,554 compound leaves counted and yielding 213,850 pairs on the symmetrical tables, were dealt with in all.

It will be seen at once that the mean number of *pinnæ* to the leaf, and their variability differs considerably with the locality, the age and the environment of the tree, but the correlation of leaves from the same tree is fairly alike for the three series, and almost identical with the theoretical value ·4000 given by the Law of Ancestral Heredity for the degree of resemblance between pairs of brothers.

The following is the actual distribution of the frequency of the leaflets :—

Number of Pinnæ on Leaves.

Series.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	Total.
Buckingh'shire	3	0	16	21	201	67	879	156	1140	68	257	9	17	0	2834
Dorsetshire. .	4	0	84	30	396	115	959	228	911	72	280	11	29	1	3120
Monmouthshire	1	5	42	24	279	55	836	143	896	83	216	6	13	1	2600
Total . .	8	5	142	75	876	237	2674	527	2947	223	753	26	59	2	8554

---

* This was a remarkably satisfactory piece of work, carried through from the collecting of the raw material down to the calculation of the constants, the statistical methods having to be mastered and applied during the course of the investigation.

† In determining the probable errors of the correlation coefficients, there is some question as to whether we should use the number of individuals, which seems too small, the number of pairs, which seems too large, or the number of individual organs dealt with. I have used the latter, but enclose the probable errors in brackets to mark the doubt.

It will be remarked at once that the leaves with an even number of leaflets are relatively infrequent; there is usually a single leaflet at the end of the compound leaf. I was able at Hampden, however, to obtain leaves illustrating almost the whole evolution of this single leaflet. Sometimes this leaflet had the slightest division at its point; this division was found in most stages of double tongued leaflet down to a complete double leaf, which would be reckoned as two *pinnæ*. In the same way the centre leaflet was found slightly trisected, three tongued and finally a triple leaflet scarcely distinguishable from the usual single leaf at the end and the two nearest side leaflets. Very occasionally the double leaflet seemed to give one side leaflet and the final leaflet. Of course, all these anomalies were rare and had to be sought for, but they would form in themselves a suggestive study.

The Dorsetshire series gives 9 *pinnæ* instead of 11 as the mode, and 7 are more frequent than 13. This latter result is also true for the Monmouthshire series, which again has nothing like the same preponderance on its modal 11 as the Buckinghamshire series. I take it that the greater variability of both Dorsetshire and Monmouthshire series, together with their greater relative proportion of compound leaves with 9 and 7 *pinnæ*, is due to those series containing a much larger proportion than the Buckinghamshire series of small trees.

For ashes in general, I think, we may safely take 10 as the average (but *not the modal*) number of *pinnæ*, the standard deviation is 1·9 *pinnæ*, and the leaf resemblance is measured by a correlation of ·4.

If we compare racial and individual variabilities we have:

Series.	S. D.	S. D. of array.	Coefficient of variation.	Percentage variability.
Buckinghamshire . .	1·6891	1·5663	15·46	92·73
Dorsetshire	1·9759	1·8140	18·65	91·81
Monmouthshire . .	2·0058	1·8342	18·57	91·44
Mean . . .	1·8903	1·7382	17·56	91·99

Thus 92 per cent. of the total racial variability of this character of the ash is to be found in the individual tree, and here, as in all our other series, it is impossible to assert that variation is a result of sexual reproduction; it is clearly an essential factor of the individual growth. I conclude with the actual tables of data for the three series.

TABLE I.—(i.) Buckinghamshire Ashes.

Number of Pinnæ on First Leaf.

Number of Pinnæ on Second Leaf.

	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	Totals.
3	6	0	6	6	15	0	6	0	33	3	—	—	—	—	75
4	0	0	0	0	0	0	0	0	0	0	—	—	—	—	0
5	6	0	18	16	74	14	155	18	94	3	2	—	—	—	400
6	6	0	16	24	96	25	213	29	113	3	0	—	—	—	525
7	15	0	74	96	716	211	2168	235	1385	34	87	0	4	—	5025
8	0	0	14	25	211	50	732	82	485	17	56	0	3	—	1675
9	6	0	155	213	2168	732	8970	1305	7361	278	756	11	20	—	21975
10	0	0	18	29	235	82	1305	278	1648	69	213	6	17	—	3900
11	33	0	94	113	1385	485	7361	1648	13558	838	2825	57	103	—	28500
12	3	0	3	3	34	17	278	69	838	90	326	19	20	—	1700
13	—	—	2	0	87	56	756	213	2825	326	1896	88	176	—	6425
14	—	—	—	—	0	0	11	6	57	19	88	16	28	—	225
15	—	—	—	—	4	3	20	17	103	20	176	28	54	—	425
16	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
Totals	75	0	400	525	5025	1675	21975	3900	28500	1700	6425	225	425	0	70850

TABLE II.—(ii.) Dorsetshire Ashes.

Number of Pinnæ on First Leaf.

Number of Pinnæ on Second Leaf.

	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	Totals.
3	2	0	2	0	23	1	37	2	24	5	4	—	—	—	100
4	0	0	0	0	0	0	0	0	0	0	0	—	—	—	0
5	2	0	170	36	506	129	786	108	324	6	28	3	2	—	2100
6	0	0	36	22	114	39	245	62	186	11	31	0	4	—	750
7	23	0	506	114	2298	465	3726	641	1816	83	210	6	12	—	9900
8	1	0	129	39	465	454	782	338	536	31	88	2	10	—	2875
9	37	0	786	245	3726	782	9428	1662	6172	308	793	13	23	—	23975
10	2	0	108	62	641	338	1662	820	1554	103	360	9	40	1	5700
11	24	0	324	186	1816	536	6172	1554	8806	676	2421	62	192	6	22775
12	5	0	6	11	83	31	308	103	676	114	417	16	28	2	1800
13	4	0	28	31	210	88	793	360	2421	417	2276	120	238	14	7000
14	—	—	3	0	6	2	13	9	62	16	120	10	34	0	275
15	—	—	2	4	12	10	23	40	192	28	238	34	140	2	725
16	—	—	—	—	—	—	—	1	6	2	14	0	2	0	25
Totals...	100	0	2100	750	9900	2875	23975	5700	22775	1800	7000	275	725	25	78000

TABLE III.—(iii.) Monmouthshire Ashes.

Number of Pinnæ on First Leaf.

Number of Pinnæ on Second Leaf.

	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	Totals.
3	—	—	4	2	9	1	8	0	1	—	—	—	—	—	25
4	—	2	0	5	19	3	66	8	21	1	—	—	—	—	125
5	4	0	34	18	208	36	366	47	269	24	38	6	—	—	1050
6	2	5	18	10	106	21	251	32	138	2	15	0	—	—	600
7	9	19	208	106	1450	280	2594	283	1593	132	277	9	14	1	6975
8	1	3	36	21	280	60	464	86	365	19	38	1	1	0	1375
9	8	66	366	251	2594	464	8022	1227	6468	430	950	24	28	2	20900
10	0	8	47	32	283	86	1227	222	1344	94	216	2	13	1	3575
11	1	21	269	138	1593	365	6468	1344	9292	841	1924	46	93	5	22400
12	—	1	24	2	132	19	430	94	841	138	366	12	16	0	2075
13	—	—	38	15	277	38	950	216	1924	366	1388	45	131	12	5400
14	—	—	6	0	9	1	24	2	46	12	45	2	3	0	150
15	—	—	—	—	14	1	28	13	93	16	131	3	22	4	325
16	—	—	—	—	1	0	2	1	5	0	12	0	4	0	25
Totals...	25	125	1050	600	6975	1375	20900	3575	22400	2075	5400	150	325	25	65000

(8.) B. *Spanish Chestnut* (Castanea vulgaris). *Veins in the Leaf.*—A character which at first sight appears easy to determine is the number of veins running from the main rib to the edge of a leaf. The leaves of many trees, however, will on examination be found quite unsuitable, the veining being far too complex, too much indeed of a network. Among fairly practicable leaves are those of the beech and the Spanish chestnut. Here two fundamental difficulties occur : (i.) near the tip of the leaf the number of veins becomes rather difficult to register. In the case of the beech by holding the leaf in various lights or even using a lens, it was possible to be fairly sure of the separate veins. In the case of the chestnut, the total number of veins is so great that the addition or omission of a single vein is hardly likely on the average to lead to any serious error. (ii.) The occasional branching or bifurcation of the vein formed a difficulty, especially if it occurred near the edge of the leaf. If the bifurcation occurred close up to the main rib of the leaf, the vein system was counted as two ; if close to the edge of the leaf as one ; intermediate bifurcations had to be left to the judgment of the recorder, having regard to the general appearance of the leaf and its system of subsidiary veining. The doubtful cases were not very many, and occurred more frequently with the chestnut than with the beech leaves. In neither case was the main rib or vein included in the total number for the leaf, the record was of the side veins flowing from this. The numbers right and left in the beech leaf were found to be far more nearly equal than in the case of the chestnut, which thus possessed considerably greater asymmetry. In these cases as in those of the ash, the leaves were gathered at random round the tree, and at random on the branches, care being, however, taken in the case of the chestnut to avoid leaves in the neighbourhood of the fruit.*

Series (i.). *Hampden Spanish Chestnuts.*

I had hoped to get records for 26 leaves from each of 100 trees from the splendid chestnut avenue at Hampden House, but it provided records for some 60 trees only. These I owe to Miss LINA ECKENSTEIN. I was able to add by diligent search about 25 additional trees from the same neighbourhood. All these were large forest trees of splendid growth. In order to make the number up to 100, some 15 additional trees were added, these being all large old trees from other neighbourhoods. I look upon this series as a fairly homogeneous one, representing the degree of resemblance between leaves on fully developed forest trees. The counting of the veins on this series was undertaken by Miss C. D. FAWCETT, B.Sc., and the tabulating and numerical reduction by Dr. LEE. Both these processes involved an immense amount of labour, in fact days and days of careful and fatiguing work.

---

* The leaves were all gathered in the autumn, towards the fall.

[To face p. 300.

Table IV.—Hampden Chestnut Trees.

Number of Veins on First Leaf.

Number of Veins on Second Leaf.

	20.	21.	22.	23.	24.	25.	26.	27.	28.	29.	30.	31.	32.	33.	34.	35.	36.	37.	38.	39.	40.	41.	42.	43.	44.	45.	46.	47.	48.	49.	50.	51.	52.	53.	54.	55.	56.	57.	58.	59.	Totals.
20	—	—	—	—	1	0	1	0	0	1	2	7	1	1	0	2	0	4	0	2	0	0	1	0	0	1	0	1	—	—	—	—	—	—	—	—	—	—	—	—	25
21	—	—	—	—	0	1	2	1	2	1	2	2	0	2	6	5	1	3	3	3	7	1	1	2	1	1	3	0	—	—	—	—	—	—	—	—	—	—	—	—	50
22	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	—	—	—	0
23	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	—	—	—	0
24	1	0	0	0	0	1	2	1	1	2	3	9	10	5	5	12	7	7	11	9	6	6	5	5	2	7	3	4	1	—	—	—	—	—	—	—	—	—	—	—	125
25	0	1	0	0	1	0	2	1	2	1	6	4	4	0	6	10	7	8	8	9	8	6	9	0	4	1	1	0	0	0	0	0	1	—	—	—	—	—	—	—	100
26	1	2	0	0	2	2	2	3	4	6	9	9	5	6	9	17	4	10	9	11	12	7	6	3	0	6	1	4	0	0	0	0	0	—	—	—	—	—	—	—	150
27	0	1	0	0	1	1	3	8	4	11	16	17	21	29	37	26	24	20	18	21	31	9	20	11	5	7	2	4	1	0	0	0	0	0	2	—	—	—	—	—	350
28	0	2	0	0	1	2	4	4	4	11	25	12	7	26	25	32	13	27	27	26	31	13	19	6	7	9	6	5	1	2	0	0	1	0	2	—	—	—	—	—	350
29	1	1	0	0	2	1	6	11	11	12	36	23	19	29	31	42	23	32	45	32	44	19	23	8	8	7	3	4	0	2	0	0	0	0	0	—	—	—	—	—	475
30	2	2	0	0	3	6	9	16	25	36	38	53	43	66	68	100	61	73	88	69	86	44	53	27	30	18	11	9	5	4	0	0	3	0	2	—	—	—	—	—	1050
31	7	2	0	0	9	4	9	17	12	23	53	56	51	56	67	93	61	89	77	79	70	44	46	29	26	27	16	16	9	9	8	2	5	1	2	—	—	—	—	—	1075
32	1	0	0	0	10	4	5	21	7	19	43	51	66	59	76	99	73	74	107	74	98	75	52	56	34	38	33	13	13	9	7	4	2	0	2	—	—	—	—	—	1225
33	1	2	0	0	5	0	6	29	26	29	66	56	59	92	95	133	104	101	125	111	124	51	87	42	34	34	33	12	17	5	3	2	2	1	13	—	—	—	—	—	1500
34	0	6	0	0	5	6	9	37	25	31	68	67	76	95	124	156	158	158	171	155	183	113	131	95	87	67	52	43	48	15	16	6	14	3	5	—	—	—	—	—	2225
35	2	5	0	0	12	10	17	26	32	42	100	93	99	133	156	174	174	180	221	208	197	137	146	87	74	60	56	35	32	13	7	7	6	2	7	—	—	—	—	—	2550
36	0	1	0	0	7	7	4	24	13	23	61	61	73	104	158	174	222	233	252	244	251	189	240	140	140	85	84	58	63	15	25	16	22	3	8	—	—	—	—	—	3000
37	4	3	0	0	7	8	10	20	27	32	73	89	74	101	158	180	233	242	267	250	256	196	249	136	134	105	74	44	62	28	27	11	16	0	9	—	—	—	—	—	3125
38	0	3	0	0	11	8	9	18	27	45	88	77	107	125	171	221	252	267	292	328	307	240	288	181	172	129	103	65	72	34	40	15	20	5	5	—	—	—	—	—	3725
39	2	3	0	0	9	9	11	21	26	32	69	79	74	111	155	208	244	250	328	346	344	315	343	233	181	189	144	118	117	36	50	30	24	10	14	—	—	—	—	—	4125
40	0	7	0	0	6	8	12	31	31	44	86	70	98	124	183	197	251	256	307	344	396	367	393	311	277	280	219	162	163	67	77	49	75	17	26	5	4	4	0	3	4950
41	0	1	0	0	6	6	7	9	13	19	44	44	75	51	113	137	189	196	240	315	367	312	326	252	246	264	175	175	132	43	46	29	26	9	8	0	0	0	0	0	3875
42	1	1	0	0	5	9	6	20	19	23	53	46	52	87	131	146	240	249	288	343	393	326	378	300	299	256	207	174	145	64	74	34	57	19	16	4	0	4	0	6	4475
43	0	2	0	0	5	0	3	11	6	8	27	29	56	42	95	87	140	136	181	233	311	252	300	276	262	282	255	215	176	103	89	71	72	26	26	14	7	12	0	15	3825
44	0	1	0	0	2	4	0	5	7	8	30	26	34	34	87	74	140	134	172	181	277	246	299	262	226	277	237	233	172	96	101	56	64	23	27	4	1	4	0	6	3550
45	1	1	0	0	7	1	6	7	9	7	18	27	38	34	67	60	85	105	129	189	280	264	256	282	277	284	309	251	200	149	146	56	87	32	35	7	1	7	0	11	3725
46	0	3	0	0	3	1	1	2	6	3	11	16	33	33	52	56	84	74	103	144	219	175	207	255	237	309	342	256	258	158	146	99	120	38	47	10	11	7	0	6	3525
47	1	0	0	0	4	0	4	4	5	4	9	16	13	12	43	35	58	44	65	118	162	175	174	215	233	251	256	224	203	104	96	81	75	32	36	7	8	5	0	3	2775
48	—	—	—	—	1	0	0	1	1	0	5	9	13	17	48	32	63	62	72	117	163	132	145	176	172	200	258	203	230	128	131	106	114	25	33	5	9	3	0	1	2675
49	—	—	—	—	—	0	0	0	2	2	4	9	9	5	15	13	15	28	34	36	67	43	64	103	96	149	158	104	128	80	96	61	85	25	23	5	6	4	0	6	1475
50	—	—	—	—	—	0	0	0	0	0	0	8	7	3	16	7	25	27	40	50	77	46	74	89	101	146	146	96	131	96	122	67	92	34	39	2	3	2	0	4	1550
51	—	—	—	—	—	0	0	0	0	0	0	2	4	2	6	7	16	11	15	30	49	29	34	71	56	56	99	81	106	61	67	42	67	24	20	6	11	3	0	0	975
52	—	—	—	—	—	1	0	0	1	0	3	5	2	2	14	6	22	16	20	24	75	26	57	72	64	87	120	75	114	85	92	67	80	24	18	7	11	5	0	5	1200
53	—	—	—	—	—	—	—	0	0	0	0	1	0	1	3	2	3	0	5	10	17	9	19	26	23	32	38	32	25	25	34	24	24	4	11	2	4	1	0	0	375
54	—	—	—	—	—	—	—	2	2	0	2	2	2	13	5	7	8	9	5	14	26	8	16	26	27	35	47	36	33	23	39	20	18	11	18	7	9	4	0	1	475
55	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	5	0	4	14	4	7	10	7	5	5	2	6	7	2	7	2	6	4	0	3	100
56	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	4	0	0	7	1	1	11	8	9	6	3	11	11	4	9	6	6	3	0	0	100
57	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	4	0	4	12	4	7	7	5	3	4	2	3	5	1	4	4	3	0	0	3	75
58	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
59	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3	0	6	15	6	11	6	3	1	6	4	0	5	0	1	3	0	3	0	2	75
Totals	25	50	0	0	125	100	150	350	350	475	1050	1075	1225	1500	2225	2550	3000	3125	3725	4125	4950	3875	4475	3825	3550	3725	3525	2775	2675	1475	1550	975	1200	375	475	100	100	75	0	75	65000

Series (ii.). TABLE V.—Mixed Set Spanish Chestnut Leaves.

Number of Veins on First Leaf.

Number of Veins on Second Leaf.

	19.	20.	21.	22.	23.	24.	25.	26.	27.	28.	29.	30.	31.	32.	33.	34.	35.	36.	37.	38.	39.	40.	41.	42.	43.	44.	45.	46.	47.	48.	49.	50.	51.	52.	53.	54.	55.	56.	57.	58.	59.	60.	61.	62.	Totals.
19	—	—	—	1	1	2	1	2	4	2	5	3	1	0	0	0	1	0	0	0	1	0	0	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	25
20	—	—	—	0	0	0	0	1	0	0	1	1	1	3	1	8	0	1	1	3	1	3	0	0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	25
21	—	—	—	0	1	0	0	0	1	0	3	0	1	5	2	3	4	2	2	1	0	0	0	0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	25
22	1	0	0	0	1	2	2	3	5	3	10	8	4	1	5	1	2	0	0	0	1	0	0	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	50
23	1	0	1	1	0	2	2	3	9	2	9	9	6	13	8	6	7	7	5	1	3	3	0	1	0	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	100
24	2	0	0	2	2	4	4	8	16	14	14	13	9	5	7	10	8	5	2	1	8	6	1	4	2	2	0	0	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	150
25	1	0	0	2	2	4	4	5	10	11	20	20	21	12	20	19	11	10	6	7	7	6	1	5	11	6	2	2	0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	225
26	2	1	0	3	3	8	5	4	20	18	26	19	15	11	17	19	12	10	8	8	12	13	2	4	2	3	2	1	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	250
27	4	0	1	5	9	16	10	20	36	30	55	47	40	39	41	47	36	33	17	23	26	18	5	11	3	2	0	0	1	0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	575
28	2	0	0	3	2	14	11	18	30	22	41	28	36	22	38	36	39	20	18	16	27	20	6	9	9	4	0	1	1	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	475
29	5	1	3	10	9	14	20	26	55	41	70	65	65	55	75	78	79	51	41	42	44	40	26	27	26	13	6	5	4	4	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000
30	3	1	0	8	9	13	20	19	47	28	65	54	63	56	76	77	57	64	40	48	56	55	22	33	30	27	14	18	8	6	0	2	0	2	3	1	—	—	—	—	—	—	—	—	1025
31	1	1	1	4	6	9	21	15	40	36	65	63	50	62	89	96	82	73	59	61	70	69	42	29	53	29	14	13	3	9	3	2	1	2	1	1	—	—	—	—	—	—	—	—	1175
32	0	3	5	1	13	5	12	11	39	22	55	56	62	88	103	111	100	92	84	88	72	85	51	47	45	23	11	14	17	10	9	9	0	1	3	2	0	0	0	1	—	—	—	—	1350
33	0	1	2	5	8	7	20	17	41	38	75	76	89	103	94	121	119	113	98	102	113	117	61	55	70	52	24	21	6	19	11	9	1	3	3	6	0	0	0	0	—	—	—	—	1700
34	0	8	3	1	6	10	19	19	47	36	78	77	96	111	121	194	160	164	137	171	132	166	98	77	86	60	27	29	21	15	7	3	1	5	5	1	1	1	4	3	—	—	—	—	2200
35	1	0	4	2	7	8	11	12	36	39	79	57	82	100	119	160	158	171	175	154	178	190	123	104	124	81	42	46	23	20	14	9	5	7	3	2	0	4	0	0	—	—	—	—	2350
36	0	1	2	0	7	5	10	10	33	20	51	64	73	92	113	164	171	206	196	240	220	252	183	164	136	135	80	87	44	44	27	7	8	15	12	6	2	6	8	6	—	—	—	—	2900
37	0	1	2	0	5	2	6	8	17	18	41	40	59	84	98	137	175	196	166	229	213	255	205	149	181	132	95	84	41	44	20	15	6	14	10	6	3	4	8	6	—	—	—	—	2775
38	0	3	1	0	1	1	7	8	23	16	42	48	61	88	102	171	154	240	229	266	268	331	258	260	250	206	154	138	75	87	61	41	17	19	22	12	6	6	2	1	—	—	—	—	3675
39	1	1	0	1	3	8	7	12	26	27	44	56	70	72	113	132	178	220	213	268	266	354	291	261	261	255	153	154	89	81	58	20	12	20	8	2	2	7	1	2	1	—	—	—	3750
40	0	3	0	0	3	6	6	13	18	20	40	55	69	85	117	166	190	252	255	331	354	382	353	294	329	293	202	171	106	88	80	39	19	24	18	8	3	6	0	0	2	—	—	—	4400
41	0	0	0	0	0	1	1	2	5	6	26	22	42	51	61	98	123	183	205	258	291	353	382	310	351	276	214	207	142	107	96	43	34	26	23	21	8	10	10	12	0	—	—	—	4000
42	1	0	0	1	1	4	5	4	11	9	27	33	29	47	55	77	104	164	149	260	261	294	310	304	348	335	271	252	191	146	104	76	37	42	31	21	13	10	11	10	1	0	0	1	4050
43	—	—	—	—	0	2	11	2	3	9	26	30	53	45	70	86	124	136	181	250	261	329	351	348	354	334	260	232	183	143	122	85	56	53	29	22	9	9	5	12	0	0	0	0	4225
44	—	—	—	—	1	2	6	3	2	4	13	27	29	23	52	60	81	135	132	206	255	293	276	335	334	416	311	339	255	219	171	131	54	53	43	49	23	16	1	11	8	0	0	6	4375
45	—	—	—	—	—	0	2	2	0	0	6	14	14	11	24	27	42	80	95	154	153	202	214	271	260	311	274	271	220	213	166	121	52	62	35	24	22	23	14	14	2	0	0	5	3400
46	—	—	—	—	—	0	2	1	0	1	5	18	13	14	21	29	46	87	84	138	154	171	207	252	232	339	271	290	246	210	170	135	70	76	50	45	23	10	5	7	2	0	0	1	3425
47	—	—	—	—	—	1	0	1	1	1	4	8	3	17	6	21	23	44	41	75	89	106	142	191	183	255	220	246	238	194	173	140	79	117	66	45	33	35	20	22	3	0	0	7	2850
48	—	—	—	—	—	—	—	1	0	2	4	6	9	10	19	15	20	44	44	87	81	88	107	146	143	219	213	210	194	164	190	150	56	87	28	37	13	16	6	13	1	0	0	2	2425
49	—	—	—	—	—	—	—	—	—	—	—	0	3	9	11	7	14	27	20	61	58	80	96	104	122	171	166	170	173	190	162	172	95	113	61	57	26	26	17	13	1	0	0	0	2225
50	—	—	—	—	—	—	—	—	—	—	—	2	2	9	9	3	9	7	15	41	20	39	43	76	85	131	121	135	140	150	172	128	79	120	52	45	18	23	9	15	1	0	0	1	1700
51	—	—	—	—	—	—	—	—	—	—	—	0	1	0	1	1	5	8	6	17	12	19	34	37	56	54	52	70	79	56	95	79	42	72	49	44	28	16	12	5	0	0	0	0	950
52	—	—	—	—	—	—	—	—	—	—	—	2	2	1	3	5	7	15	14	19	20	24	26	42	53	53	62	76	117	87	113	120	72	92	58	38	20	25	16	18	0	0	0	0	1200
53	—	—	—	—	—	—	—	—	—	—	—	3	1	3	3	5	3	12	10	22	8	18	23	31	29	43	35	50	66	28	61	52	49	58	50	50	35	15	20	16	1	0	0	0	800
54	—	—	—	—	—	—	—	—	—	—	—	1	1	2	6	1	2	6	6	12	2	8	21	21	22	49	24	45	45	37	57	45	44	38	50	24	26	8	12	9	1	0	0	0	625
55	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	1	0	2	3	6	2	3	8	13	9	23	22	23	33	13	26	18	28	20	35	26	10	7	12	5	1	0	0	1	350
56	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	1	4	6	4	6	7	6	10	10	9	16	23	10	35	16	26	23	16	25	15	8	7	4	6	6	0	0	0	1	300
57	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	4	0	8	8	2	1	0	10	11	5	1	14	5	20	6	17	9	12	16	20	12	12	6	14	12	—	—	—	—	225
58	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0	3	0	6	6	1	2	0	12	10	12	11	14	7	22	13	13	15	5	18	16	9	5	6	12	6	—	—	—	—	225
59	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	2	0	1	0	8	2	2	3	1	1	1	0	0	1	1	1	0	—	—	—	—	—	—	25
60	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—	—	0
61	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—	—	0
62	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0	6	5	1	7	2	0	1	0	0	0	0	1	1	—	—	—	—	—	—	25
Totals	25	25	25	50	100	150	225	250	575	475	1000	1025	1175	1350	1700	2200	2350	2900	2775	3675	3750	4400	4000	4050	4225	4375	3400	3425	2850	2425	2225	1700	950	1200	800	625	350	300	225	225	25	0	0	25	67600

Series (ii.). *Mixed Spanish Chestnuts.*

My second series consists of 26 leaves from each of 104 trees. These were gathered in a variety of neighbourhoods. Many of these were young trees. In the case of nearly 50 sets, the trees were pollards from old stumps with about 15 years' growth; some of the others were from fine old trees, and as a result we have many districts and many ages combined. I have to thank Miss C. D. Fawcett for sets from the neighbourhood of Hampstead, Miss Grace Lee for sets from that of St. Albans, Miss Charles for sets from the neighbourhood of Dorking, while about half the series was gathered by Mr. Robert J. Parker from his chestnut woods at Haslemere. The counting of the veins was undertaken by Miss M. Notcutt, while the tabulation and numerical reduction are the work of Dr. W. R. Macdonell. The series being mixed, the variability is higher than that of the first series and thus the labour spent on it was even greater.

The following table gives the general results:—

Resemblance of Spanish Chestnut Leaves from same Tree.

Series.	Number.			Mean number of veins.	S. D. of veins.	Correlation.
	Trees.	Leaves.	Pairs.			
Hampden .	100	2600	65,000	41·0735 ± ·0775	5·8780 ± ·0550	·4655 [± ·0104]
Mixed . .	104	2704	67,600	41·2145 ± ·0840	6·4780 ± ·0594	·5913 [± ·0084]
Mean . .	—	—	—	41·1440	6·1780	·5284

We note at once that the mean number of veins for the forest trees and for the mixed group is sensibly the same. On the other hand, the variability and the correlation differ very considerably. The latter is sensibly higher than the ·4 obtained for the leaflets of the ash. I should be inclined to consider that the heterogeneity of the mixed series is at the bottom of the difference between the two series, but there cannot, I think, be any doubt of the greater resemblance of chestnut leaves for this character than of ash leaves in the number of *pinnæ.*

The actual distribution of veins in the 5304 leaves examined was as follows:—

No. of veins.	19.	20.	21.	22.	23.	24.	25.	26.	27.	28.	29.	30.	31.	32.	33.
1st series	—	1	2	0	0	5	4	6	14	14	19	42	43	49	60
2nd series	1	1	1	2	4	6	9	10	23	19	40	41	47	54	68
Total	1	2	3	2	4	11	13	16	37	33	59	83	90	103	128

No. of veins.	34.	35.	36.	37.	38.	39.	40.	41.	42.	43.	44.	45.	46.	47.	48.
1st series.	89	102	120	125	149	165	198	155	179	153	142	149	141	111	107
2nd series	88	94	116	111	147	150	176	160	162	169	175	136	137	114	97
Total	177	196	236	236	296	315	374	315	341	322	317	285	278	225	204

No. of veins.	49.	50.	51.	52.	53.	54.	55.	56.	57.	58.	59.	60.	61.	62.	Total.
1st series.	59	62	39	48	15	19	4	4	3	0	3	—	—	—	2600
2nd series	89	68	38	48	32	25	14	12	9	9	1	0	0	1	2704
Total	148	130	77	96	47	44	18	16	12	9	4	0	0	1	5304

The total gives a fairly smooth series with the mode at 40, and a secondary mode at 42. In the forest trees we have the same result as in the total, but in the mixed group the secondary mode is at 44, and almost equally important with the primary. The influence of symmetry makes itself felt in the preponderance of the even numbers, although not in such a marked degree as in the case of the ash, where the counting of the final leaflet gave, of course, preponderance to the odd numbers.

For the Spanish chestnut we may accordingly take 41 veins as the average, 40 as the modal number to the leaf; 6·2 as the S.D. of the veins; and ·5 (slightly weighting the homogeneous series) as the degree of resemblance between the leaves ot the same tree.

It remains to consider the comparative individual and racial variabilities :—

Series.	S. D.	S. D. of array.	Coefficient of variation.	Percentage variability.
Hampden	5·8780	5·2024	14·31	88·51
Mixed	6·4780	5·2242	15·72	80·65
Mean	6·1780	5·2133	15·01	84·58

We see that the variability of the chestnut leaf as measured by the coefficient of variation is somewhat less than that of the ash. The individual as compared with the racial variability is also somewhat less than in the case of the ash, but the individual exhibits still some 85 per cent. of the racial variability. Tables IV. and V. give the data for the chestnut series.

(9.) C. *Beech* (Fagus sylvatica). *Number of Veins on the Leaf.*—I have here only one series of 26 leaves from each of 100 trees. These were gathered and counted by Mrs. KARL PEARSON and myself in the neighbourhood of Great Hampden. The trees were in part fine isolated specimens growing on the commons, but as there were not sufficient of these we were compelled to gather from the somewhat smaller trees abounding in the beechwoods of this part of Buckinghamshire. The leaves were gathered about 4 to 5 feet from the ground on the outside of the tree.* The tabulation of the results and the calculation of the constants are again the work of Dr. LEE.

Hampden Beech Leaves.

Number.			Mean number of veins.	S.D. of veins.	Correlation of pairs.
Trees.	Leaves.	Pairs.			
100	2600	65,000	16·1062 ± ·0230	1·7351 ± ·0162	·5699 [± ·0087]

Distribution of Number of Veins.

Number of veins...	10	11	12	13	14	15	16	17	18	19	20	21	22	Total.
Frequency	1	7	34	110	318	479	595	516	307	181	36	15	1	2600

Individual and Racial Variation.

S.D.	S.D. of array.	Coefficient of variation.	Percentage variability.
1·7351	1·4258	10·77	82·17

Accordingly, we see that the modal value of the beech leaf veins is 16, which is very near the mean, 16·11. The standard deviation about 1·7, and the degree of resemblance among its leaves ·57 Measured by its coefficient of variation, it is the least variable of the trees so far dealt with, and the leaves of the same tree more closely resemble each other. Notwithstanding its relatively small variability, the individual on an average exhibits about 82 per cent. of the racial variability.

The table of reduced data is given below.

---

* I did not observe among the leaves gathered anything of the differentiation indicated by Professor MACLEOD between "licht- en schadernbladen." The material, so far as this character (number of veins) goes, seemed singularly homogeneous and of small variability.

TABLE VI.—Hampden Beech Trees.

Number of Veins on First Leaf.

Number of Veins on Second Leaf.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	Totals.
10	0	2	7	7	6	3	—	—	—	—	—	—	—	25
11	2	8	35	45	59	23	3	—	—	—	—	—	—	175
12	7	35	92	145	261	156	92	47	15	—	—	—	—	850
13	7	45	145	394	836	689	387	199	46	1	1	—	—	2750
14	6	59	261	836	2124	2021	1616	723	243	54	4	3	—	7950
15	3	23	156	689	2021	3256	3122	1856	605	221	20	3	—	11975
16	—	3	92	387	1616	3122	4576	3205	1380	437	51	6	—	14875
17	—	—	47	199	723	1856	3205	3914	1928	860	120	46	2	12900
18	—	—	15	46	243	605	1380	1928	1882	1278	213	80	5	7675
19	—	—	—	1	54	221	437	860	1278	1248	282	137	7	4525
20	—	—	—	1	4	20	51	120	213	282	138	64	7	900
21	—	—	—	—	3	3	6	46	80	137	64	32	4	375
22	—	—	—	—	—	—	—	2	5	7	7	4	0	25
Totals	25	175	850	2750	7950	11975	14875	12900	7675	4525	900	375	25	65000

(10.) D. *Holly* (Ilex Aquifolium). *Number of Prickles on the Leaf.*—This character seemed an easy one so far as mere counting goes, although the leaves were troublesome to gather. The leaves were gathered, so far as possible, all round the tree or bush, about the height at which cattle would browse, and from the outside of the bush. Internal and high branches (often differentiated by being almost prickleless) were to be avoided. Flowering trees, and, above all, leaves near the flowering buds or fruit were to be rejected. These conditions could only be approximately followed in practice when a large number of trees had to be dealt with from one neighbourhood, and these series, I fancy, allow more to the personal equation of the gatherer than is wholly satisfactory.

Series (i.). *Somersetshire Hollies.*

I owe this series of 26 leaves from each of 100 trees to Miss AGNES FRY. Of these trees, 56 were hedge-row hollies, and 44 wood trees or bushes growing in moist places.* The counting of the prickles was in some part undertaken by Mr. MACLEOD,

---

* Miss FRY writes to me: "The texture of the leaves seems rather different,—the wood hollies are much less stout and harsh and more often have few prickles. In fact, I rather wonder if *shade* does not affect the number. In the few cases where hedge-row hollies have very few prickles the leaves come from the inside of the bush. I have carefully avoided the neighbourhood of berries,—and what is more difficult of flower buds. But the holly has two flowers, one with abortive pistil and one with abortive stamens. The latter I suppose would drop away and leave no trace of their presence. But, on the whole, I have avoided flower and berry-bearing bushes—it is easy to find bushes without either. And, so far as my observation goes, I am not convinced that the neighbourhood of flowers or the top of the tree causes modification . . . . It is very difficult to keep to a standard of counting for the tip, but I have tried to do so. Doubtless the hollies too are of rather different ages, but I have avoided patriarchal trees, though doubtful if this makes much difference in this case." It will be seen that in this point, as in others, much had to be left to the judgment of a careful observer.

but for more than three-quarters of the trees by Miss Fry herself. The labour of tabulation and the deduction of the constants was undertaken by Dr. Lee.

Series (ii.). *Dorsetshire Hollies.*

This series of 26 leaves from each of 100 trees was gathered and counted by Miss C. D. Fawcett, B.Sc. It differs in several important respects from Miss Fry's series. As it came from the neighbourhood of the sea, by Lyme Regis, we might expect a considerable difference of environment; the trees were chiefly hedgerow trees. They show, however, a *lower* mean number of prickles than the Somersetshire series, which embraced a large element of wood trees. I think the series is less homogeneous than the Somersetshire series. In particular, Tree 91 contributes the almost abnormal series of 24 leaves with 1 and 2 leaves with 2 prickles. The lump at the end of the frequency distribution is practically due to this one tree. I do not feel able to reject it entirely, but I give the numerical constants as determined with and without it. I take it that of the two observers the personal equation of one tended rather to emphasise the need for a mathematically "random selection" of trees, and the personal equation of the other to emphasise the importance of a botanically "homogeneous group" of trees. The result has been the increase of correlation and variability in the first, and their decrease in the second case. Probably their mean is not far from the truth. It is noteworthy that the ash trees from Dorsetshire and Monmouthshire, although exhibiting a like substantial difference of environment, were far more alike.

The tabulation and reduction of this series is also due to Dr. Lee. I now summarise the results as before:—

Resemblance of Holly Leaves from the same Tree.

Series.	Number.			Mean number of prickles.	S. D. of prickles.	Correlation.
	Trees.	Leaves.	Pairs.			
Somersetshire (i.)	100	2600	65,000	15·2935 ± ·0379	2·8655 ± ·0268	·3648 [± ·0115]
Dorsetshire (ii.)*a*.	100	2600	65,000	13·4373 ± ·0497	3·7595 ± ·0352	·6423 [± ·0078]
" (ii.)*b*.	99	2574	64,350	13·5622 ± ·0474	3·5661 ± ·0335	·5985 [± ·0085]
Mean (i.) and (ii.)*a*	—	—	—	14·3654	3·3125	·5035
Mean (i.) and (ii.)*b*	—	—	—	14·4278	3·2158	·4816

(ii.)*b* is the Dorsetshire series, excluding tree (91).

In round numbers accordingly the number of prickles on the holly leaf is 14·4, the standard deviation 3·3 prickles, and the degree of resemblance ·5. The holly thus falls between our results for the ash and the chestnut, one series approaching our Buckinghamshire ashes, the other the group of mixed Spanish chestnuts, being somewhat in excess of the Buckinghamshire beeches.

The actual frequency of prickles is given by :—

Series.	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
Somersetshire (i.) . . .	1	2	2	3	6	6	10	8	10	24	43	74	135	276
Dorsetshire (ii.)*a* . . .	0	40	21	14	18	27	31	45	62	84	86	152	219	336
,, (ii.)*b* . . .	0	16	19	14	18	27	31	45	62	84	86	152	219	336
Total (i.) and (ii.)*a* .	1	42	23	17	24	33	41	53	72	108	129	226	354	612
Total (i.) and (ii.)*b* .	1	18	21	17	24	33	41	53	72	108	129	226	354	612

Series.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	26.	Total.
Somersetshire (i.) . . .	335	417	407	328	235	122	79	44	22	6	3	1	1	2600
Dorsetshire (ii.)*a* . . .	350	337	311	217	128	70	31	12	8	1	0	0	0	2600
,, (ii.)*b* . . .	350	337	311	217	128	70	31	12	8	1	0	0	0	2574
Total (i.) and (ii.)*a* .	685	754	718	545	363	192	110	56	30	7	3	1	1	5200
Total (i.) and (ii.)*b* .	685	754	718	545	363	192	110	56	30	7	3	1	1	5174

The series is thus seen to be fairly uniform, the mode corresponding in the total result as well as the Somersetshire series to 15 prickles, while that of the Dorsetshire series is at 14, an even number of prickles. The holly leaf has clearly not the marked symmetry of ash or chestnut leaves. Thus, 16 prickles in the Somersetshire series are almost as frequent as 15, or since there is almost always a prickle at the tip, asymmetrical leaves must be very nearly as frequent as symmetrical.

Lastly, turning to the comparison of racial and individual variabilities, we have the table :—

Series.	S. D.	S. D. of array.	Coefficient of variation.	Percentage variability.
Somersetshire (i.) . .	2·8655	2·6680	18·74	93·12
Dorsetshire (ii.)*a* . .	3·7595	2·8815	27·98	76·65
,, (ii.)*b* . .	3·5661	2·8282	26·29	80·11
Mean (i.) and (ii.)*a*	3·3125	2·7747	23·36	84·88
Mean (i.) and (ii.)*b*.	3·2158	2·7482	22·52	86·61

Here again we see how great is the variability within the individual. Looking upon Tree (91) as so largely biasing the complete Dorsetshire series that we must exclude it, we see that the individual possesses 80 to 90 per cent. of the variability of the race. The variability of the offspring of the individual will, of course, be larger still, even if we had merely uniparental reproduction. Thus we have still further evidence, if more were necessary, that variability is not a factor dependent upon sexual reproduction. For this character of the prickles it will be seen that the holly leaf is more variable than those of the ash, chestnut, or beech. The following are the fundamental tables of data :—

TABLE VII.—Somersetshire Hollies.

Number of Prickles on First Leaf.

Number of Prickles on Second Leaf.

	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	26.	Totals.
0	—	1	0	0	1	0	3	0	1	2	2	0	1	3	2	3	4	0	2	—	—	—	—	—	—	—	—	25
1	1	0	1	2	1	0	3	0	1	2	4	0	2	7	4	5	8	3	3	2	1	—	—	—	—	—	—	50
2	0	1	0	2	0	1	0	1	2	1	2	0	3	10	5	5	5	7	2	2	1	—	—	—	—	—	—	50
3	0	2	2	2	1	0	1	0	0	3	6	3	7	10	6	8	10	6	2	4	2	—	—	—	—	—	—	75
4	1	1	0	1	0	3	5	2	6	8	14	18	17	17	12	17	12	5	7	2	1	1	—	—	—	—	—	150
5	0	0	1	0	3	8	4	1	8	7	8	22	12	28	18	18	3	8	1	0	0	0	—	—	—	—	—	150
6	3	3	0	1	5	4	8	1	4	17	14	13	19	29	26	39	28	16	15	3	2	0	—	—	—	—	—	250
7	0	0	1	0	2	1	1	2	4	6	12	10	17	27	22	22	24	17	15	10	6	1	—	—	—	—	—	200
8	1	1	2	0	6	8	4	4	6	10	21	21	23	36	31	34	16	16	8	2	0	0	—	—	—	—	—	250
9	2	2	1	3	8	7	17	6	10	22	47	45	78	65	80	76	53	38	21	11	4	3	0	0	0	0	1	600
10	2	4	2	6	14	8	14	12	21	47	70	95	139	166	146	120	101	43	28	22	8	4	3	0	0	0	0	1075
11	0	0	0	3	18	22	13	10	21	45	95	162	237	327	286	212	167	125	52	32	11	8	2	0	1	0	1	1850
12	1	2	3	7	17	12	19	17	23	78	139	237	410	533	532	498	327	237	160	58	34	16	6	1	4	0	4	3375
13	3	7	10	10	17	28	29	27	36	65	166	327	533	1184	1199	1181	947	562	280	141	76	44	20	4	0	0	4	6900
14	2	4	5	6	12	18	26	22	31	80	146	286	532	1199	1360	1471	1325	885	491	237	112	78	36	4	2	0	5	8375
15	3	5	5	8	17	18	39	22	34	76	120	212	498	1181	1471	1970	1811	1343	738	415	225	135	65	8	4	2	0	10425
16	4	8	5	10	12	3	28	24	16	53	101	167	327	947	1325	1811	1832	1459	950	517	294	169	80	20	5	5	3	10175
17	0	3	7	6	5	8	16	17	16	38	43	125	237	562	885	1343	1459	1334	1026	512	272	175	71	18	12	7	3	8200
18	2	3	2	2	7	1	15	15	8	21	28	52	160	280	491	738	950	1026	944	484	324	185	76	33	20	6	2	5875
19	—	2	2	4	2	0	3	10	2	11	22	32	58	141	237	415	517	512	484	232	202	98	40	17	5	1	1	3050
20	—	1	1	2	1	0	2	6	0	4	8	11	34	76	112	225	294	272	324	202	186	99	78	24	11	1	1	1975
21	—	—	—	—	1	0	0	1	0	3	4	8	16	44	78	135	169	175	185	98	99	32	34	12	5	1	—	1100
22	—	—	—	—	—	—	—	—	—	0	3	2	6	20	36	65	80	71	76	40	78	34	24	8	5	2	—	550
23	—	—	—	—	—	—	—	—	—	0	0	0	1	4	4	8	20	18	33	17	24	12	8	0	1	—	—	150
24	—	—	—	—	—	—	—	—	—	0	0	1	4	0	2	4	5	12	20	5	11	5	5	1	0	—	—	75
25	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	2	5	7	6	1	1	1	2	—	—	—	—	25
26	—	—	—	—	—	—	—	—	—	1	0	1	4	4	5	0	3	3	2	1	1	—	—	—	—	—	—	25
Totals	25	50	50	75	150	150	250	200	250	600	1075	1850	3375	6900	8375	10425	10175	8200	5875	3050	1975	1100	550	150	75	25	25	65000

TABLE VIII.—Dorsetshire Hollies.

Number of Prickles on First Leaf.

Number of Prickles on Second Leaf.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	Totals.
1	590 [38]	105 [57]	36	19	40	35	8	41	15	19	27	43	14	4	3	1	—	—	—	—	—	—	—	1000 [400]
2	105 [57]	74 [72]	49	31	51	37	24	55	29	13	13	24	11	5	3	1	—	—	—	—	—	—	—	525 [475]
3	36	49	20	17	31	26	14	34	15	13	19	33	15	8	10	5	2	2	1	—	—	—	—	350
4	19	31	17	16	46	36	60	53	43	23	21	24	18	16	11	4	5	3	3	0	1	—	—	450
5	40	51	31	46	68	54	91	70	47	30	25	59	32	13	18	0	0	0	0	0	0	—	—	675
6	35	37	26	36	54	40	75	93	75	72	52	70	36	28	18	10	8	3	2	2	2	1	—	775
7	8	24	14	60	91	75	142	144	129	71	69	71	67	53	41	29	18	6	3	5	4	1	—	1125
8	41	55	34	53	70	93	141	188	203	132	156	129	87	64	37	38	16	2	3	3	2	0	—	1550
9	15	29	15	43	47	75	129	203	220	176	224	185	199	173	148	115	59	24	15	4	2	0	—	2100
10	19	13	13	23	30	72	71	132	176	180	268	268	296	205	156	113	54	27	21	8	4	1	—	2150
11	27	13	19	21	25	52	69	156	224	268	404	526	581	453	382	285	127	90	45	18	9	6	—	3800
12	43	24	33	24	59	70	71	129	185	268	526	784	1039	755	625	456	199	104	48	19	8	6	—	5475
13	14	11	15	18	32	36	67	87	199	296	581	1039	1822	1497	1234	725	391	196	80	40	11	6	3	8400
14	4	5	8	16	13	28	53	64	173	205	453	755	1497	1594	1393	1129	685	385	163	74	21	24	8	8750
15	3	3	10	11	18	18	41	37	148	156	382	625	1234	1393	1462	1223	814	480	223	90	32	17	5	8425
16	1	1	5	4	0	10	29	38	115	113	285	456	725	1129	1223	1432	1025	582	348	147	52	51	4	7775
17	—	—	2	5	0	8	18	16	59	54	127	199	391	685	814	1025	910	574	307	157	56	17	1	5425
18	—	—	2	3	0	3	6	2	24	27	90	104	196	385	480	582	574	340	222	99	35	24	2	3200
19	—	—	1	3	0	2	3	3	15	21	45	48	80	163	223	348	307	222	136	70	33	26	1	1750
20	—	—	—	0	0	2	5	3	4	8	18	19	40	74	90	147	157	99	70	22	13	4	0	775
21	—	—	—	1	0	2	4	2	2	4	9	8	11	21	32	52	56	35	33	13	6	9	0	300
22	—	—	—	—	—	1	1	0	0	1	6	6	6	24	17	51	17	24	26	4	9	6	1	200
23	—	—	—	—	—	—	—	—	—	—	—	—	3	8	5	4	1	2	1	0	0	1	—	25
Totals	1000 [400]	525 [475]	350	450	675	775	1125	1550	2100	2150	3800	5475	8400	8750	8425	7775	5425	3200	1750	775	300	200	25	65000 [64350]

The lower numbers in brackets give the results when Tree 91 is excluded.

Section II.—*Seed Vessels of Flowers.* A. *Poppies.*

(11.) Having dealt with several characters in a variety of trees, I thought it desirable to deal with a single character in a variety of races in a very different branch of the vegetable kingdom. After consultation with my colleague, Professor F. O. Oliver, the stigmatic bands on the seed capsules of poppies were selected, and they have proved an easily recorded character, which could be considered for a variety of races. Here again there was some diversity in the method of collection. The rules given being (i.) that the bands were to be counted at the periphery of the capsule and not at the centre; thus  was counted as 7, not 6, bands; and (ii.) every possible bud on the plant was to be counted. The work went on from spring to autumn of the year 1899; it was impossible to collect, dry, and preserve for counting all the wild poppies gathered as soon as they had ceased flowering. Thus some poppies were taken when they had completely done flowering, others with flowers on them, and others with many buds. Few, I think, would be included which had not formed all their buds. In my own series, I counted the bands on every capsule, in every flower, and in nearly every bud, opening the bud and using a microscope when necessary. The labour of this was very great, and was not carried out by all my helpers. I doubt, indeed, whether any advantage was derived from it, except in so far as there may be correlation between the number of stigmatic bands and the early or late development of the seed-vessel.

(12.) The following series were obtained :—

Series (i.). *Buckinghamshire Wild Poppy* (Papaver Rhœas).—I gathered 200 specimens from as wide a range as possible—roadside, field-hedgerow, and cornfield—avoiding *Papaver dubium*, and counting every set of stigmatic bands right down to the smallest bud I could manage. These specimens were gathered on the very top of the Chilterns; the soil and climate is not a very favourable one, and my mean number of buds was a small one. I myself tabled the data and calculated the constants.

Series (ii.). *Buckinghamshire Wild Poppy* (Papaver Rhœas).—Miss C. D. Fawcett gathered 236 poppy plants on the southern slope of the Chilterns, towards their foot. They came nearly all from a single cornfield, and thus their environment was much more uniform, and the climate more favourable than in Series (i.). It is quite possible that these poppies were largely the offspring of a comparatively few poppies, as they were gathered from a small area. Dr. Lee tabled the data and calculated the constants.

Series (iii.). *Somersetshire Wild Poppy* (Papaver Rhœas).—The bulk of this series were gathered by Miss Agnes Fry in the Quantocks, but being rather a short series, 86 plants, I added to it several other short series, 42 plants in all, gathered and counted by various less active helpers. Thus the series is really an omnibus series, although mainly from the Quantocks. Dr. LEE tabulated the data and calculated the constants.

My other two series are of the Shirley Poppy, a well-known and beautiful garden variety.

Series (iv.). *Chelsea Shirley Poppy.*—These plants were grown by my colleague, Professor F. O. OLIVER, in his garden at Chelsea. The plants had scarcely been thinned, and so, being very crowded, grew tall, and with few buds on the plant. The plants were taken up and forwarded to me by rail, but being delayed in transit suffered a great deal, so that a good many capsules were broken off, or could not be counted. These circumstances account for the smallness of the number of pairs obtained from each plant. Professor OLIVER kindly came and assisted me in the counting. The data were tabulated and reduced by myself.

Series (v.). *Buckinghamshire Shirley Poppy.*—These plants were grown at Hampden Farm House, and covered a large border. They were extremely healthy, large plants, and I have counted as many as 120 capsules on a single plant! The plants were taken up singly as they had done flowering and the bands counted. The labour of counting was severe, as there were 4443 capsules on 176 plants as compared with 1020 capsules on 325 Chelsea poppies. The tabulation involved the entry of 197,478 pairs, perhaps the largest number we have yet dealt with in a single correlation table; this and the determination of the constants is the work of Dr. LEE.

The reader will notice that the investigations on poppies differ largely from those on the leaves of trees. In the latter case, 26 leaves were taken at random from the trees and the character on each measured. In the former case, every available capsule was dealt with. These, of course, are not *all* the capsules put forth by the plant, and the proportion of these two sets would vary considerably from one series to another, approaching fairly close to equality in my wild poppies from the top of the Chilterns, and differing very considerably in the Shirley poppies from Chelsea. In the leaves of the trees no tree was weighted by its size or vigour; in the case of a vigorous poppy, however, it may contribute ten times as many pairs to our table as a feeble plant. There is a good deal to be said for both methods of dealing with the problem, and even for a third method to which I shall refer immediately. If we pick up two leaves at random and wish to consider the probability that they belong to one tree, we ought certainly to allow for the larger number of pairs of leaves on the larger trees; if we wish to discover to what extent two randomly chosen men have characters like to those of brothers, we ought certainly to allow for the large families, even if we do not deal with every possible pair of brothers in those families. These two different methods of dealing with resemblance due either to the production of like organs by the individual or to heredity must be carefully borne in mind. In dealing with Mr. GALTON's data for stature* I have taken every possible pair of

---

* 'Phil. Trans.,' A, vol. 187, pp. 253–318.

brothers; in my own measurements of 1000 families I have taken only two brothers from each family. We cannot *a priori* expect the results to come out exactly the same. Indeed, we might *a priori* expect the latter method to give rather higher results than the former, for it introduces more individual families, and if we tabulated all the brothers of one family only or all the like organs of one organism, we should find nothing but a spurious correlation growing indefinitely small as the number of brothers or organs was indefinitely increased.* Thus I anticipated that the Hampden Shirleys would give a lower correlation than the Chelsea Shirleys, the wild poppies from the top of the Chilterns a higher value than those from the bottom, &c.

As our series of poppy plants were in no case very large, it was impossible to directly test this result by taking only a single pair from each plant, even if it had been possible after the counting and marking to select a pair of capsules at random. I accordingly adopted a third or intermediate method. Miss FAWCETT had recorded her capsules under the letters *a*, *b*, *c*, &c., and not grouped them directly under the number of the stigmatic bands. Accordingly, it was possible to take *a* with *b*, *c* with *d*, *e* with *f*, &c., and thus repeat no single capsules, but get a number of pairs sensibly taken at random. Thus large and vigorous plants will give more pairs than feeble ones, and yet much fewer than if we take every possible pair. This series will be referred to as Series (ii.)^{bis}. Its general result confirms the anticipation referred to above. It will be noticed that nothing of this kind can be done in the case of leaves of trees, where we cannot gather *every* leaf. One great advantage of the 26 leaves from every tree is the means it provides of testing the accuracy of the lengthy process of tabulation and calculation. Each leaf is used 25 times, and thus the mean and standard-deviation of the total number of pairs ought to be the same as those of the total series of leaves (say, 2600 for 100 trees). This valuable method of verification is lost when we do not take the same number of like organs from each individual. It then becomes of considerable interest to ascertain how far the mean and variability of the total number of organs observed are identical with the mean and variability of the total number of pairs of organs observed. This comparison is made in Series (v.)^{bis}, the mean and variability of the 4443 capsules being compared with those of the corresponding 197,478 pairs. The whole of the tabulation and arithmetic involved in Series (ii.)^{bis} is mine, but Series (v.)^{bis} is again due to Dr. LEE.

I now proceed to sum up the results obtained from the above seven series.

---

* I worked out the theory of this some time ago, and may publish it on another occasion.

Resemblance of Poppy Capsules from the same Plant.

Series.	Number of			Mean number of bands.	S. D. of bands.	Correlation.
	Plants.	Capsules.	Pairs.*			
Wild poppy (i.) . (top of Chilterns)	200	1305	11,026	10·0443 ± ·0331	1·7704 ± ·0234	·5620 [± ·0128]
Wild poppy (ii.) . (southern slope of Chilterns)	236	2343	25,160	9·9055 ± ·0210	1·5087 ± ·0149	·3997 [± ·0117]
Wild poppy (ii.)$^{\text{bis}}$ (no capsule repeated)	236	2268	2,268	9·8444 ± ·0211	1·4878 ± ·0149	·4452 [± ·0114]
Wild poppy (iii.) . (Quantocks)	127	1324	19,790	8·7733 ± ·0270	1·4553 ± ·0191	·5333 [± ·0133]
Shirley poppy (iv.) (Chelsea)	325	1020	2,756	12·3679 ± ·0355	1·6803 ± ·0251	·6149 [± ·0131]
Shirley poppy (v.) (Hampden)	176	4443	197,478	12·6103 ± ·0191	1·8853 ± ·0135	·5238 [± ·0073]
Mean (i.), (ii.), (iii.)	563	4972	55,976	9·5744	1·5781	·4983
Mean (iv.), (v.) .	501	5463	200,234	12·4891	1·7828	·5693
Mean (i.) to (v.) .	1064	10,435	256,210	[10·6524]	[1·6600]	·5267

Now this table contains a good deal of interest. Upwards of a thousand poppy plants were dealt with, involving the counting of the bands on more than ten thousand capsules, and the entry of more than a quarter of a million of pairs in the correlation tables. We see at once that the mean and variability of the Shirley poppy differ so widely from those of the wild poppy, that it is idle to group these species for these characters together. It is interesting to note, however, that both in the number of bands and in the variability the Hampden poppies, whether wild or Shirley, stand at the top of their respective series. With regard to the wild poppy, I probably gathered from a wider area with a greater variety of environment than was the case in Series (ii.) or (iii.). That mine were gathered from much poorer soil is evidenced by my plants having on an average only 6 capsules, while the Quantocks and southern Chiltern poppies had about 10 capsules. The difference is possibly even greater than these numbers indicate, for I counted every bud I possibly could. The two Chiltern series do not differ much in the mean number of bands, but the difference of variability is very significant. I should take it to be a very fair measure of the relative range of environment in the two cases. When we turn to the Shirley series the difference is almost too great for any comparison. While the Hampden series had some twenty-five capsules, on an average, to the plant, the Chelsea series numbered

---

* The reader must bear in mind that this column gives the number of pairs of the symmetrical tables below, or *twice* the number of independent pairs.

only about three! In fact, besides being extremely crowded a process of selection had gone on in the transit. From examination of the uninjured plants, I think that five capsules a plant would be more than the average of these poppies before packing, and that at least one to two capsules per plant were destroyed on the journey. Accordingly, I do not consider that much weight should be given to the results of this series, but I felt bound to include it as I have included all the series dealt with, even when I hold them to be really unsuited to the present investigation. If we consider the two methods of dealing with Series (ii.), *i.e.*, that of including all possible pairs, and that of including pairs in which no capsule was repeated, we find that no very sensible change is made in the mean or variation, but a rather greater increase in the correlation than might have been anticipated. I think this is largely due to unconscious selection in the choice of pairs, *i.e.*, we took the records for (*a*) and (*b*), (*c*) and (*d*), &c., as our pairs; it is quite possible that in recording the capsules successive letters would correspond to capsules gathered from the same part of the plant, and so the members of the pair would be subject to the same or somewhat similar conditions of nutriment, light and warmth during development. Valuable investigations have yet to be made on the relative degree of resemblance of leaves or seed-vessels coming at random from the same tree or plant, and coming from the same branch or shoot. The effects of position on the individual undoubtedly produces some differentiation, and all we can hope to do at present is to get round numbers approximately describing the degree of resemblance of these *theoretically* undifferentiated like organs.

If we exclude the Chelsea series, we have for the

Wild poppy	·4983
Shirley poppy	·5238
Mean	·5110

For the four series of trees we have

Ash	·3918
Chestnut	·5284
Holly	·5035
Beech	·5699
Mean	·4984

So far as the seed-vessels of poppies and the leaves of trees are concerned we have a degree of resemblance in undifferentiated like organs which approximates to the value ·5.

It will be as well to put together the frequency distributions of the capsules as we have done in the case of the trees :—

Wild Poppies (*Papaver Rhœas*). Frequency of Stigmatic Bands.

Series.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	Totals.
Chilterns (i.) . . .	2	17	86	182	255	272	260	130	67	25	9	—	1305
Chilterns (ii.) . . .	1	13	103	304	571	634	428	198	56	26	6	3	2343
Quantocks (iii.) . .	17	61	189	358	316	227	96	39	20	1	—	—	1324
Total . . .	20	91	378	844	1142	1133	784	367	143	52	15	—	4972

Shirley Poppies. Frequency of Stigmatic Bands.

Series.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	Totals.
Chelsea (i.). .	0	1	1	4	40	95	167	245	202	145	86	26	7	1	0	1020
Hampden (ii.)	1	11	32	56	148	363	628	925	954	709	397	155	51	12	1	4443
Total .	1	12	33	60	188	458	795	1170	1156	854	483	181	58	13	1	5463

These distributions appear to be uniform although asymmetrical, and they give no obvious sign of a mixture. The mode is 10 for the Chilterns series and 8 for the Quantocks series, while the Shirley poppies give 12 for Chelsea and 13 for Hampden.

It remains to tabulate the several results obtained for variation. These are given in the following table. We see at once that the general effect of weighting the capsules with the number of pairs on the plant has the effect of increasing the mean and reducing the variability ; in other words, the plants with most capsules on and leading to most pairs, are those with the greatest number of stigmatic bands. We have some evidence accordingly that the most vigorous plants exhibit most bands on the capsules :—

Individual and Racial Variability.

Series.	Stigmatic bands.		Coefficient of variation.	S. D. of array.	Percentage variability.
	Mean.	S. D.			
Wild Poppy.					
(i.) from pairs	10·0443	1·7704	17·6259	1·4644	82·71
,, ,, capsules	9·9318	1·7684	17·8052	—	—
(ii.) from pairs	9·9055	1·5087	15·2309	1·3829	91·66
,, ,, capsules	9·8288	1·5007	15·2677	—	—
,, pairs without repetition of capsules	9·8444	1·4878	15·1132	1·3322	89·54
(iii.) from pairs	8·7733	1·4553	16·5878	1·2311	84·59
,, ,, capsules	8·7236	1·5402	17·6556	—	—
Mean from pairs	9·5744	1·5781	16·4809	1·3661	86·32
Mean from capsules . . .	9·4947	1·6031	16·9095	—	—
Shirley Poppy.					
(iv.) from pairs	12·3679	1·6803	13·5860	1·3251	78·86
,, ,, capsules	12·3873	1·7326	13·9869	—	—
(v.) from pairs	12·6103	1·8853	14·9505	1·6060	85·18
,, ,, capsules	12·5091	1·8977	15·1705	—	—
Mean from pairs	12·4891	1·7828	14·2682	1·4655	82·02
Mean from capsules. . .	12·4482	1·8151	14·5787	—	—
Total mean from pairs . .	—	—	15·5962	1·4019	84·60
Total mean from capsules .	—	—	15·9772	—	—

But we might well have expected a good deal more evidence of this. The differences of the means calculated from capsules and pairs is not very great, and in two out of three series of the wild poppies the variability as deduced from pairs is very slightly greater than (sensibly equal to) that deduced from capsules. The mean values of the coefficient of variation, as calculated from pairs and from capsules, are in close agreement. The only sensible deviation is that of the Quantocks series, and this arises from the large difference in the standard deviations of the two methods. The work has been carefully revised, but no error discovered.

Generally for the coefficient of variation we have—

Ash	17·56	Wild poppy	16·91
Spanish chestnut	15·05	Shirley poppy	14·58
Beech	10·77		
Holly	23·36	Mean . .	15·98
Mean . .	16·67		

So that the variability of both may be expressed in round numbers as 16. It is clearly not a constant for the whole vegetable kingdom, but 16 enables us to form some rough appreciation of the plasticity of vegetable forms, and may be compared with values obtained from other series later.

Comparing again the ratio of the racial and individual variabilities, we see that the latter reaches, on an average, over 85 per cent. of the former, the minimum value being 79 and the maximum 91. It is difficult to test this point directly, for the very simple reason that a single plant will, as a rule, have very few flowers; and upon small numbers it is impossible to test satisfactorily the full variability of the plant. The probable error becomes enormous if we have only four or five capsules to determine the standard deviation from. The method adopted has been that which arises in the treatment of parental inheritance; all the offspring of parents of a given stature are clubbed together and form an array. This array has a certain variability, and this variability is taken to be that of a single family of offspring having parents of the given stature. It would clearly be impossible to find the variability inside the family from numbers such as occur in a single human family. To some extent we can verify the assumptions thus made. Thus I took the wild poppy with the largest number of capsules I had come across, and determined their variability; there were 36 such capsules, and the standard deviation was 1·4519. I then took two of the most prolific Shirley poppies which had 126 and 85 capsules on them, and their variabilities were given by standard deviations of 1·3290 and 1·5889, *i.e.*, a mean of 1·4589. These are actually in excess of the standard deviations of the arrays, namely, 1·3661 and 1·4019 of the two series of poppies! I think we may accordingly take the latter as a reasonable measure of the possible variation of the individual.

TABLE IX.—Wild Poppy (Chilterns' Top) Series (i.).

Number of Stigmatic Bands on First Capsule.

Number of Stigmatic Bands on Second Capsule.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	Totals.
5	0	8	4	—	—	—	—	—	—	—	—	12
6	8	38	46	23	11	5	2	—	—	—	—	133
7	4	46	184	163	146	78	31	8	1	—	—	661
8	—	23	163	390	398	279	111	75	32	4	—	1475
9	—	11	146	398	554	415	250	161	68	9	5	2017
10	—	5	78	279	415	514	520	240	112	27	10	2200
11	—	2	31	111	250	520	770	366	160	52	9	2271
12	—	—	8	75	161	240	366	252	178	35	12	1327
13	—	—	1	32	68	112	160	178	92	23	11	677
14	—	—	—	4	9	27	52	35	23	28	12	190
15	—	—	—	—	5	10	9	12	11	12	4	63
Totals	12	133	661	1475	2017	2200	2271	1327	677	190	63	11026

TABLE X.—Wild Poppy (Chilterns' Base) Series (ii.).

Number of Stigmatic Bands on First Capsule.

Number of Stigmatic Bands on Second Capsule.

	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	Totals.
5	0	1	2	1	2	1	4	—	—	—	—	—	11
6	1	4	40	25	19	10	9	1	1	—	—	—	110
7	2	40	184	289	228	145	48	20	4	1	—	—	961
8	1	25	289	822	964	563	283	105	20	8	—	—	3080
9	2	19	228	964	1868	1633	855	353	78	32	5	4	6041
10	1	10	145	563	1633	2106	1469	682	151	60	17	8	6845
11	4	9	48	283	855	1469	1220	585	154	63	20	13	4723
12	—	1	20	105	353	682	585	336	134	54	14	14	2298
13	—	1	4	20	78	151	154	134	52	44	17	7	662
14	—	—	1	8	32	60	63	54	44	26	8	1	297
15	—	—	—	—	5	17	20	14	17	8	2	1	84
16	—	—	—	—	4	8	13	14	7	1	1	0	48
Totals	11	110	961	3080	6041	6845	4723	2298	662	297	84	48	25160

TABLE XI.—Wild Poppies (Chilterns' Base) (ii.)$^{\text{bis}}$.

(No repetition of any individual capsule.)

Number of Stigmatic Bands on First Capsule.

Number of Stigmatic Bands on Second Capsule.

	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	Totals.
5	—	—	—	—	1	—	—	—	—	—	—	—	1
6	—	2	5	1	1	0	3	—	—	—	—	—	12
7	—	5	24	28	19	13	1	1	—	—	—	—	91
8	—	1	28	80	79	73	24	8	2	—	—	—	295
9	1	1	19	79	202	140	74	25	5	4	—	—	550
10	—	0	13	73	140	198	127	50	10	7	1	—	619
11	—	3	1	24	74	127	112	59	10	6	0	2	418
12	—	—	1	8	25	50	59	32	13	5	2	0	195
13	—	—	—	2	5	10	10	13	10	3	1	0	54
14	—	—	—	—	4	7	6	5	3	0	0	0	25
15	—	—	—	—	—	1	0	2	1	0	0	1	5
16	—	—	—	—	—	—	2	0	0	0	1	0	3
Totals	1	12	91	295	550	619	418	195	54	25	5	3	2268

TABLE XII.—Wild Poppy (Quantocks) Series (iii.).

Number of Stigmatic Bands on First Capsule.

Number of Stigmatic Bands on Second Capsule.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	Totals.
5	14	41	31	22	16	7	4	0	1	0	136
6	41	170	299	202	93	40	9	0	9	0	863
7	31	299	922	895	375	113	22	2	7	0	2666
8	22	202	895	1716	1312	687	178	22	11	0	5045
9	16	93	375	1312	1590	1275	337	43	19	0	5060
10	7	40	113	687	1275	1234	406	92	52	0	3906
11	4	9	22	178	337	406	308	138	51	1	1454
12	0	0	2	22	43	92	138	102	37	1	437
13	1	9	7	11	19	52	51	37	26	4	217
14	0	0	0	0	0	0	1	1	4	0	6
Totals...	136	863	2666	5045	5060	3906	1454	437	217	6	19790

TABLE XIII.—Shirley Poppies (Chelsea) Series (iv.).

Number of Stigmatic Bands on First Capsule.

Number of Stigmatic Bands on Second Capsule.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	Totals.
6	—	—	1	—	—	—	—	—	—	—	—	—	—	1
7	—	—	0	0	1	—	—	—	—	—	—	—	—	1
8	1	0	0	0	2	0	0	3	—	—	—	—	—	6
9	—	0	0	20	37	29	15	9	—	—	—	—	—	110
10	—	1	2	37	88	64	39	19	4	1	—	—	—	255
11	—	—	0	29	64	140	138	70	19	4	1	—	—	465
12	—	—	0	15	39	138	196	154	68	25	2	—	—	637
13	—	—	3	9	19	70	154	174	105	55	8	—	—	597
14	—	—	—	—	4	19	68	105	112	63	17	2	1	391
15	—	—	—	—	1	4	25	55	63	52	16	4	0	220
16	—	—	—	—	—	1	2	8	17	16	4	4	1	53
17	—	—	—	—	—	—	—	—	2	4	4	6	1	17
18	—	—	—	—	—	—	—	—	1	0	1	1	0	3
Totals	1	1	6	110	255	465	637	597	391	220	53	17	3	2756

TABLE XIV.—Shirley Poppies (Hampden) Series (v.).

Number of Stigmatic Bands on First Capsule.

Number of Stigmatic Bands on Second Capsule.

	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	Totals.
5	0	10	19	10	4	3	8	4	5	—	—	—	—	—	—	63
6	10	90	190	101	51	42	80	40	50	—	—	—	—	—	—	654
7	19	190	354	208	122	140	308	194	145	7	—	—	—	—	—	1687
8	10	101	208	114	181	251	365	335	238	112	44	11	—	—	—	1970
9	4	51	122	181	762	1144	1085	721	390	119	50	16	—	—	—	4645
10	3	42	140	251	1144	2858	3943	3357	1948	698	249	86	15	7	—	14741
11	8	80	308	365	1085	3943	7444	6401	5327	2305	758	200	54	2	—	28280
12	4	40	194	335	721	3357	6401	10574	9095	5129	2234	715	179	49	—	39027
13	5	50	145	238	390	1948	5327	9095	11198	8494	4008	1321	303	57	—	42579
14	—	—	7	112	119	698	2305	5129	8494	9388	5569	2151	594	144	2	34712
15	—	—	—	44	50	249	758	2234	4008	5569	3664	1717	513	115	2	18923
16	—	—	—	11	16	86	200	715	1321	2151	1717	792	325	97	1	7432
17	—	—	—	—	—	15	54	179	303	594	513	325	160	58	4	2205
18	—	—	—	—	—	7	2	49	57	144	115	97	58	20	1	550
19	—	—	—	—	—	—	—	—	—	2	2	1	4	1	0	10
Totals ...	63	654	1687	1970	4645	14741	28280	39027	42579	34712	18923	7432	2205	550	10	197478

(13.) C. *Nigella Hispanica.*—Besides the poppy capsules, Professor OLIVER had suggested to me that it would be possible to count the segmentation of the seed vessel in the cases of Nigella and Mallow. I accordingly procured from Messrs. VEITCH, of Slough, upwards of a hundred plants of *Nigella Hispanica.* Meanwhile, Professor OLIVER sent me a list of some he had counted himself, and pointed out that there was a substantial differentiation between the seed vessels on the main stem and those on the side shoots. It was quite clear on examination that these seed vessels were not undifferentiated like organs, the number of segments in a seed vessel on the side shoots being much more like the number on a similar vessel on a separate plant than like those of a main-shoot capsule on the same plant. Further, although I do not know the history of these special plants, they were probably from highly selected seed and in a very unstable condition. Whereas in the examination of many thousand poppy capsules I only found one remarkable abnormal case,* abnormalities in the case of *Nigella Hispanica* abounded and rendered the counting of the segmentation very difficult. Double and treble capsules were found in every state of union up to complete fusion, and with every variety of deformity, such as the tips of one capsule growing through the sides of a second capsule. Thus the material was by no means suited to my present purpose. However, as I had the material it seemed worth while testing the influence of differentiation in reducing correlation. Accordingly I made no attempt to separate top from side capsules,† and formed the following table in the manner sufficiently indicated in the preceding pages :—

---

[*] A double capsule on one stem. It contained only a few seeds. These Professor F. O. OLIVER kindly sowed for me this year; there resulted two plants with a very few poor flowers, but the abnormality was not perpetuated, and the capsules rotted off without giving seed.

[†] If the top capsules only had been taken I should often have had only one on a plant.

TABLE XV.—Segmentation of *Nigella Hispanica*.

Number of Segments in First Seed Vessel.

Number of Segments in Second Seed Vessel.

	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	Totals.
2	—	—	—	5	2	3	—	—	—	—	—	—	—	—	—	—	—	—	—	10
3	—	—	—	1	1	0	4	0	0	0	0	1	—	—	—	—	—	—	—	7
4	—	—	—	2	5	4	6	1	1	0	0	0	0	0	1	—	—	—	—	20
5	5	1	2	64	58	55	101	13	1	0	2	1	0	0	0	—	—	—	—	303
6	2	1	5	58	60	76	168	25	9	2	6	0	0	0	0	—	—	—	—	412
7	3	0	4	55	76	88	240	41	7	10	9	0	0	0	1	—	—	—	—	534
8	—	4	6	101	168	240	866	107	28	15	11	4	0	0	2	—	—	—	—	1552
9	—	0	1	13	25	41	107	14	9	3	9	0	0	0	0	0	0	0	1	223
10	—	0	1	1	9	7	28	9	0	1	1	0	0	0	1	0	0	0	1	59
11	—	0	0	0	2	10	15	3	1	0	3	0	0	0	1	—	—	—	—	35
12	—	0	0	2	6	9	11	9	1	3	2	—	—	—	—	—	—	—	—	43
13	—	1	0	1	0	0	4	0	0	0	—	—	—	—	—	—	—	—	—	6
14	—	—	0	0	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	0
15	—	—	0	0	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	0
16	—	—	1	0	0	1	2	0	1	1	—	—	—	—	—	—	—	—	—	6
17	—	—	—	—	—	—	—	0	0	—	—	—	—	—	—	—	—	—	—	0
18	—	—	—	—	—	—	—	0	0	—	—	—	—	—	—	—	—	—	—	0
19	—	—	—	—	—	—	—	0	0	—	—	—	—	—	—	—	—	—	—	0
20	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	2
Totals...	10	7	20	303	412	534	1552	223	59	35	43	6	0	0	6	0	0	0	2	3212

While the modal value is conspicuously 8, the mean value is 7·4639. This is due to the second modal value of 5 on the side shoots. The standard deviation is 1·4893, and the correlation ·1899. It will be seen at once how the result fully confirmed my expectation, that the degree of resemblance between capsules on the same plant would be largely reduced, owing to the existence of differentiation. The actual distribution of capsules in 210 plants was as follows :—

Number of Segments.

	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	Total.
Frequency. . .	1	1	5	75	108	133	420	61	19	10	8	2	0	0	1	0	0	0	1	845

Thus on an average there were four seed vessels to a plant. The mean and standard deviation found directly from the capsules are 7·5101 and 1·4559 respectively. These are in substantial agreement with the above results found for pairs. They do not confirm the conclusion based upon the data for poppies, namely that if we work with pairs we shall raise the mean and lower the variability. The fact is that vigour seems to work in opposite directions with this plant and the poppy ; with the latter it means many capsules with a high number of stigmatic bands, but with the former it means many side shoots with the differentiated capsules with few segmentations. Such capsules of course decrease the mean and raise the variability.

The coefficient of variation is 19·385, sensibly greater than that of the wild poppy, but below that of the holly leaf. The standard deviation of an array is 1·4622, or the variability of the individual 98·18 per cent. of that of the race. The plant with the greatest number of seed vessels on had only 11, and these gave a standard deviation of 1·6188, greater than that of the whole race ! But of course this is subject to a very large probable error ; it suffices to emphasise, however, how the source of variation is to be sought within the individual, and is not a mere result of sexual reproduction.

The following table sums up the results obtained for *Nigella Hispanica* :—

	Number of			Mean.	S. D.	Coefficient of variation.	S. D. of array.	Percentage variation.	Correlation.
	Plants.	Capsules.	Pairs.						
Pairs. .	210	845	3212	7·4639	1·4893	—	1·4622	98·18	·1899
Capsules				7·5101	1·4559	19·385	—	—	—

(14.) D. *Mallow* (Malva Rotundifolia).—I had given up the idea of the ordinary wild mallow (*Malva sylvestris*) owing to the few plants to be found in the neighbourhood of Hampden, as well as on account of the immense number of seed vessels on the plants to be counted. But one day an inspection of my potato-patch revealed the whole ground covered with low plants of *Malva Rotundifolia.*[*] On inquiry I found that a cottage had recently stood on the site, and that the cottagers kept in their herb-bed a small clump of this plant for the preparation of a local ointment of high repute. When the cottage was removed the mallow had spread over the whole patch, I believe very little by seed, but largely by stolons. This seemed to provide excellent material ready to hand, and from it we can draw some conclusions as to the influence of such a division on the degree of resemblance of like parts. It will be remembered that DARWIN,[†] in experimenting on *Origanum vulgare*, found that cross-fertilisation of flowers on plants which had separate roots, but had spread by stolons from the same individual, presented no advantage over self-fertilisation, and he concluded that this arose from the fact that the plants were really one and the same individual so far as such processes were concerned. If this were so in the case of my mallow, we ought to find a great reduction in the relationship of undifferentiated like organs. Two points struck me at the outset—the small variability in the segmentation of the seed vessels and the low mean value (13·7) compared with that (15) given in botanical works. I was able to collect 127 plants; of these, 76 had more than 26 capsules, some many more, and of these I took 26 only; 51 plants had fewer than 26 capsules, and although I took buds in every stage of development I could not make up my number to this, often only obtaining 10, or even fewer. As the potatoes were being removed the mallows had to be taken up before the seeds were ripe, and the seed-vessels artificially dried. The counting of the segments proved a very laborious task. To Miss M. NOTCUTT I owe a considerable portion of the counting. The buds and small seed vessels requiring a lens or small microscope were counted by Dr. LEE, Mr. RADFORD SHARPE, and myself. I am responsible for the tabulation of the data and for all the numerical calculations.

The 127 plants gave rise to 57,740 pairs distributed thus:—

TABLE XVI.—*Malva Rotundifolia.*

Number of Segments in the First Seed-vessel.

Number of Segments in the Second Seed-vessel.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	Totals.
10	10	24	60	78	89	36	5	—	—	—	302
11	24	64	202	270	185	57	7	—	—	—	809
12	60	202	850	1802	1905	637	95	1	—	—	5552
13	78	270	1802	5036	6195	2409	324	8	0	5	16127
14	89	185	1905	6195	9760	4293	653	13	0	10	23103
15	36	57	637	2409	4293	2428	386	5	0	8	10259
16	5	7	95	324	653	386	60	2	0	2	1534
17	—	—	1	8	13	5	2	—	—	—	29
18	—	—	—	0	0	0	0	—	—	—	0
19	—	—	—	5	10	8	2	—	—	—	25
Totals	302	809	5552	16127	23103	10259	1534	29	0	25	57740

---

[*] Misnamed "marsh mallow" throughout this district of Buckinghamshire.

[†] 'Cross and Self-fertilisation in Plants,' 3rd edition, p. 94

The following table gives the chief numerical results :—

Malva Rotundifolia.

	Number of			Mean.	S. D.	Coefficient of variation.	S. D. of array.	Per-centage variation.	Correla-tion.
	Plants.	Capsules.	Pairs.						
Pairs. .	127	2568	57740	13·7000 ± ·0139	1·0407 ± ·0098	—	1·0232	98·32	·1827 ± [·0129]
Capsules				13·6632	1·0652	7·7961	—	—	—

The actual distribution of frequency of the capsules was :—

No. of segments.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	Total.
Frequency . .	18	41	263	736	994	445	68	2	0	1	2568

These results amply bear out my *a priori* conclusions.

The degree of resemblance of like parts has been very largely reduced. The variability (7·7961) is smaller than any we have yet come across, and the variability of the individual differs under 2 per cent. from that of the race. All these conclusions are compatible with the decreased individuality which may be expected to exist among the members of a group to a large extent propagated by stolons and not from seed.

Although in the majority of cases (76 out of 127) I took the same number of seed capsules (26) from the plant, still the influence of the plants with fewer capsules is quite marked; we see that weighting with the number of pairs lowers the variation and raises the mean, in other words, the large vigorous plants have more segmentation and less variability. The distribution of segmental frequency is fairly uniform, and shows a marked mode at 14 segments, not the 15 of the usual statement.

SECTION III.—*Members of Whorls.*

(15.) *Woodruff* (Asperula odorata).—I must confess that there is divergence between the standpoints of the botanist and of the statistician. To the latter a character is good or bad according as it affords facilities for fairly easy measurement or enumeration. He has first to seek such characters as a *sine quâ non,* and then inquire how far they occur in undifferentiated like organs. On the other hand, the biologist will readily provide a list of such characters in plant or animal life, and not one of them be available for measurement or counting. I had long fixed my eye on the members of the whorls of the woodruff. They were so fascinatingly easy to collect and count; they exhibited such a moderate amount of variation, and one felt sure the collecting, counting, tabulating, and calculating could all be done in the inside of a week; and the impulse to such a light task after the elaborate work on the earlier series was very great. Professor F. O. Oliver and others warned me that the members of the whorls were differentiated in their origin* and also in their position on the stem, and that such a series was unsuited to illustrate the degree of resemblance between undifferentiated like organs. However, the task seemed, comparatively speaking, so easy that I felt I would undertake it, if only to compare with the *Nigella Hispanica*, and ascertain in another case how differentiation does weaken the degree of resemblance of like parts. I accordingly collected 201 single sprays of woodruff† well distributed along about a mile of lane on the bank of a hedgerow facing south. These were gathered at Great Hampden, Buckinghamshire. In counting the members on the whorls I soon found evidences of differentiation in position, the whorls towards the top of the spray having, as a rule, fewer members than those lower down. The following is the table of frequency of different pairs of whorls:—

Table XVII.—Woodruff. Great Hampden.

Number of Members on First Whorl.

Number of Members on Second Whorl.	4.	5.	6.	7.	8.	9.	10.	Totals.
4	8	8	41	7	1	—	—	65
5	8	32	273	126	91	4	—	534
6	41	273	2376	1363	1150	41	4	5248
7	7	126	1363	1216	953	34	3	3702
8	1	91	1150	953	1284	54	2	3535
9	—	4	41	34	54	4	—	137
10	—	—	4	3	2	—	—	9
Totals ...	65	534	5248	3702	3535	137	9	13230

I now give the numerical statement of the constants deduced from the above table and the frequency distribution of the whorls:—

---

[*] Some of the members of the whorls are true leaves and others modified stipules.

[†] They were broken off as close to the ground as possible; they included all parts branching off above ground, but the forked sprays were only few in number.

Woodruff.

	Number of			Mean.	S. D.	Coefficient of variation.	S.D. of array.	Percentage variation.	Correlation.
	Sprays.	Whorls.	Pairs.						
Pairs ...	201	1465	13230	6·7978 ± ·0164	·9318 ± ·0116	—	·9177	98·49	·1733 [± ·0171]
Whorls				6·9010	·9291	13·4631	—	—	—

The frequency distribution of the 1465 whorls was :—

Number of	4.	5.	6.	7.	8.	9.	10.	Total.
Frequency . . .	6	40	537	410	455	16	1	1465

The double modes 6 and 8 are brought out by the above distribution, and this is evidence either of differentiation or of a tendency to an even number of members. Contrary to the usual rule, the mean number of members is reduced when we take pairs, indicating that when we take a spray with many whorls—such as occasionally arises from forking—the number of members in the whorls are reduced, and since the variation is also very slightly increased (as in *Nigella*), they are spread out over a greater range.

The great reduction in the correlation, the value of which is only ·1733, shows how unsuited the material was for the purposes of the present investigation. At the same time the ratio of the individual to the racial variability reaches the very high value of 98·5 ! We have selected a character to test individuality on, which exhibits differences which are largely racial, and not peculiar to the individual plant.

SECTION IV.—*Ferns.*

(16.) In choosing ferns for considering the resemblance between undifferentiated like organs I had in view not only the comparative ease of counting (as compared, for example, with the segmentation of mallow), but the fact which students of the fern strongly impressed upon me that it was peculiarly subject to its environment. One great authority went so far as to assert that the presence or absence of *sori* on an individual hartstongue depended *solely* on the environment and not at all on the individuality of the plant. While hardly prepared to accept to the full such a statement (for I had already learned in a variety of types of the existence of individuality in a marked manner apart from the influence of environment), I anticipated in accordance with it a fairly high correlation of undifferentiated like organs, and this has, indeed, turned out to be the case.

(17.) *Somersetshire Hartstongue* (Scolopendrium vulgare).—I owe the counting of the *sori* on 8 to 12 fronds of each of 101 hartstongue ferns to Miss AGNES FRY. It was not always possible to obtain the full suggested number of 10 fronds. In a few cases Miss FRY also counted more than 10. But the ferns with fewer than 10 are not many, and the distribution of the fronds is sensibly identical with that of the pairs of fronds. The number of *sori* was often very large, running up to 180; the grouping of the leaves was accordingly arranged in units of 10 *sori*, and, to save fractionising in the table, these groups were taken 0—9·5, 9·5—19·5, 19·5—29·5, &c. The grouping on the basis of Miss FRY's record was carried out by Mr. LESLIE BRAMLEY-MOORE. With regard to the counting of the *sori*, all the sterile as well as the fertile markings were included. Fronds with no markings were included; one plant, the only one found, with absolutely no markings on any of the fronds was excluded, and to this I shall recur immediately. It was difficult to procure ten perfect fronds; many of those gathered were found to be more or less torn; in many cases the loss of *sori* was easily ascertained; in others an estimate had to be formed. Such estimates were made, however, in only forty-four cases out of more than 1,000 fronds, and in the majority of these cases it meant the addition of very few *sori*, and a possible error in the estimate of still fewer. Sometimes one or two *sori* may have been missed at the tip, where the small markings are indistinct, but the errors due to this source as well as to the occasional error of a unit or two in counting such a large number are well under the unit of grouping of ten *sori* finally selected. The environment of the series was not quite uniform, some thirty-three coming from a stream bed and the remainder from a park glade. The plants varied considerably in age, but the necessity of procuring eight to twelve fronds was probably effective in excluding any quite young ferns. The absence or presence of *sori* on the frond was not directly associated with its age or size. Miss FRY has recorded a number of large and medium fronds with no *sori*, and some of the fronds recorded as young have occasionally a very considerable number of *sori* (as many as ninety-five on one at least). Still, a glance at the table on p. 330 shows that the fronds with zero *sori* form in same way an anomalous group; they give a hump in the frequency distribution in a manner somewhat similar to that due to barren wives in fertility-frequency curves.* Dr. LEE having tabulated the *sori*, and calculated the constants on the basis of including all fronds with zero number of *sori*, I went through Miss FRY's data again and excluded, not all the entries in the first row and column of Miss LEE's table, but only such of them as arose from the fronds with no *sori*. The result was a great smoothing of the table, a raising of the mean, and a lowering of the variability; but *no appreciable effect whatever was made on the degree of resemblance of like organs*, although I

---

* See 'The Chances of Death and other Studies in Evolution,' vol. 1, essay on "Reproductive Selection," and diagram.

threw out upwards of 1212 pairs! This was hardly to be anticipated, but may certainly give us confidence in the substantial validity of the numerical result reached. This result is clearly independent of any extraneous source hindering the production of any *sori* at all on the frond. About a third of the ferns in the series (31 in 101) had one or more fronds in ten without *sori* at all, the total number of fronds without *sori* being seventy-six. The following table gives the chief numerical constants:—

Hartstongue Fern.

	Number of			Mean number of *sori*.	S. D.	Coefficient of variation.	S. D. of array.	Percentage variation.	Correlation.
	Plants.	Fronds.	Pairs.						
Fronds without *sori* included .	101	1024	9384	73·7398 ± ·7419	35·1967 ± ·5246	47·7310	27·3136	77·60	·6307 [± ·0127]
Fronds without *sori* excluded .	101	948	8172	74·3948 ± ·5246	31·2176 ± ·4836	41·9640	24·2358	77·64	·6303 [± ·0132]

The following is the actual distribution of pairs, which in this case is very nearly proportional to the distribution of fronds:—

Number of *sori*.	0–9·5.	9·5–19·5.	19·5–29·5.	29·5–39·5.	39·5–49·5.	49·5–59·5.	59·5–69·5.	69·5–79·5.	79·5–89·5.	89·5–99·5.	99·5–109·5.	109·5–119·5.	119·5–129·5.	129·5–139·5.	139·5–149·5.	149·5–159·5.	159·5–169·5.	169·5–179·5.	179·5–189·5.	Totals.
Fronds without *sori* included .	838	254	352	727	783	886	979	1040	887	908	717	446	277	191	27	18	27	18	9	9384
Fronds without *sori* excluded	103	211	294	634	716	829	925	1004	856	893	709	439	276	184	27	18	27	18	9	8172

We see from the last table how the "hump" disappears from the frequency distribution when we exclude the fronds without *sori*. The double mode indicates at least considerable flatness at the top of the distribution, and may possibly mark some heterogeneity of material. I am inclined, however, to think that the great variability (about forty-two even when the fronds without *sori* are excluded) is fully accounted for by the sensitiveness of the fern in the matter of *sori* to very slight differences of environment. I look upon the high resemblance of like parts here as having been intensified by this cause; upon an inherent individuality we have superposed an individuality due partially, perhaps, to age, but largely to small differences of immediate environment. Although the variation in the race is so large, and the degree of individuality so great, the ratio of individual to racial variability is still 78 per cent. Thus while the variability of the hartstongue (as far as *sori* is concerned) is double that of the very variable holly, and almost thrice that of the poppy, still the percentage variability of the two latter species is to that of the former only as about 85 to 78. Thus in the most variable and most individual species we have yet come across, we still find the variation within the individual is more than three-fourths of the entire variation of the race. In view of facts like this, it seems impossible to maintain the position taken up by Mr. ADAM SEDGWICK, that variation is the outcome of bi-sexual reproduction. The source of variation exists within the individual and is extensively active without the occurrence of any form of mating whatever.

I place here the table for the distribution of *sori* in pairs of fronds. The numbers in brackets are those which must replace the unbracketed numbers, if fronds without *sori* be excluded.

TABLE XVIII.—Hartstongue

Number of Sori *on First Frond.*

Number fo Sori *on Second Frond of Pair.*

	0–9·5.	9·5–19·5.	19·5–29·5.	29·5–39·5.	39·5–49·5.	49·5–59·5.	59·5–69·5.	69·5–79·5.	79·5–89·5.	89·5–99·5.	99·5–109·5.	109·5–119·5.	119·5–129·5.	129·5–139·5.	139·5–149·5.	149·5–159·5.	159·5–169·5.	169·5–179·5.	179·5–189·5.	Totals.
0 – 9·5	272 (14)	47 (4)	70 (12)	111 (18)	88 (21)	65 (8)	65 (11)	37 (1)	35 (4)	18 (3)	10 (2)	7 (0)	1 (0)	12 (5)	—	—	—	—	—	838 (103)
9·5– 19·5	47 (4)	20	40	44	30	27	17	6	12	7	2	0	2	0	—	—	—	—	—	254 (211)
19·5– 29·5	70 (12)	40	32	61	36	38	28	11	18	13	5	0	0	0	—	—	—	—	—	352 (294)
29·5– 39·5	111 (18)	44	61	134	134	74	68	46	29	17	5	0	1	3	—	—	—	—	—	727 (634)
39·5– 49·5	88 (21)	30	36	134	146	112	79	68	39	27	11	7	5	0	0	1	0	—	—	783 (716)
49·5– 59·5	65 (8)	27	38	74	112	116	136	127	76	49	41	10	8	7	0	0	0	—	—	886 (829)
59·5– 69·5	65 (11)	17	28	68	79	136	152	173	109	89	31	12	3	6	1	4	3	2	1	979 (925)
69·5– 79·5	37 (1)	6	11	46	68	127	173	206	145	124	64	16	9	3	1	4	0	0	0	1040 (1004)
79·5– 89·5	35 (4)	12	18	29	39	76	109	145	122	133	113	29	13	10	3	1	0	0	0	887 (856)
89·5– 99·5	18 (3)	7	13	17	27	49	89	124	133	190	130	69	32	9	1	0	0	0	0	908 (893)
99·5–109·5	10 (2)	2	5	5	11	41	31	64	113	130	168	79	40	11	7	0	0	0	0	717 (709)
109·5–119·5	7 (0)	0	0	0	7	10	12	16	29	69	79	90	81	34	5	1	3	2	1	446 (439)
119·5–129·5	1 (0)	2	0	1	5	8	3	9	13	32	40	81	38	39	5	0	0	0	0	277 (276)
129·5–139·5	12 (5)	0	0	3	0	7	6	3	10	9	11	34	39	46	4	1	3	2	1	191 (184)
139·5–149·5	—	—	—	—	0	0	1	1	3	1	7	5	5	4	0	0	0	0	0	27 (27)
149·5–159·5	—	—	—	—	1	0	4	4	1	0	0	1	0	1	0	0	3	2	1	18 (18)
159·5–169·5	—	—	—	—	0	0	3	0	0	0	0	3	0	3	0	3	6	6	3	27 (27)
169·5–179·5	—	—	—	—	—	—	2	0	0	0	0	2	0	2	0	2	6	2	2	18 (18)
179·5–189·5	—	—	—	—	—	—	1	0	0	0	0	1	0	1	0	1	3	2	0	9 (9)
Totals. .	838 (103)	254 (211)	352 (294)	727 (634)	783 (716)	886 (829)	979 (925)	1040 (1004)	887 (856)	908 (893)	717 (709)	446 (439)	277 (276)	191 (184)	27 (27)	18 (18)	27 (27)	18 (18)	9 (9)	9384 (8172)

(18.) B. *Somersetshire Ceterach* (Ceterach officinarum).—Some time after I had tried in vain to count any characters in bracken, I had a letter from Miss Agnes Fry saying that she thought it just possible that the lobes on the fronds of ceterach might be counted. The chief difficulty, of course, was the indefinite character of the lobes near the tip—a difficulty which had rendered the lobes in bracken impossible. The rather indefinite tip is here of greater consequence than in the hartstongue or the chestnut leaf, for the total number of lobes is comparatively small. In some samples Miss Fry sent me, however, I agreed fairly closely with her estimates, and although the tip must form a difficulty,* we settled that an attempt should be made to include ceterach in the present series. Miss Fry accordingly counted 9 to 11 fronds on each of 99 plants. It was necessary to take plants of very different ages, and even if ceterach be not so sensible to its environment as hartstongue, it is quite possible that part of the correlation observed is due to similarity of age. The absolute agreement of the result obtained with that for hartstongue is one of the most striking things in the whole collection of data. Considering the difficulty of the tip in ceterach and the sensibility of the *sori* of hartstongue to environment, it may be a chance agreement, but it is certainly one that gives ground for pause, and suggests further investigations of the degree of resemblance between like organs in ferns. The tabulation of the data and the calculation of the constants are again due to Dr. Lee.

Ceterach.

Number of			Mean No. of lobes.	S. D.	Coefficient of variation	S. D. of array.	Percentage variation.	Correlation.
Plants.	Fronds.	Pairs.						
99	999	9098	23·1677 ± ·0902	4·2278 ± ·0638	18·2485	3·2795	77·57	·6311[± ·0128]

The distribution of frequency of the fronds is almost the same as that of the pairs, since the attempt was made to take 10 fronds; actually 80 plants had 10, 14 had 11, and 5 only 9 fronds.

Frequency from Pairs.

No. of lobes .	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.
Frequency .	55	73	146	229	375	418	584	575	811	726	800	805

No. of lobes .	25.	26.	27.	28.	29.	30.	31.	32.	33.	34.	35.	36.	37.	38.	39.	Total.
Frequency .	863	679	668	386	288	215	184	91	37	45	27	0	9	0	9	9098

---

* Miss Fry writes: "It is very difficult to keep to a standard of counting for the tip, but I have tried to do so."

Table XIX.—Ceterach, Somersetshire.

Number of Lobes in First Frond.

Number of Lobes in Second Frond.

	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	26.	27.	28.	29.	30.	31.	32.	33.	34.	35.	36.	37.	38.	39.	Totals.
13	2	4	5	7	5	8	7	5	7	4	0	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	55
14	4	2	8	10	7	11	11	10	5	3	2	0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	73
15	5	8	10	9	15	14	21	19	17	9	6	7	2	0	2	1	1	—	—	—	—	—	—	—	—	—	—	146
16	7	10	9	10	21	25	31	28	23	26	21	7	7	2	2	0	0	—	—	—	—	—	—	—	—	—	—	229
17	5	7	15	21	22	45	59	47	48	32	27	18	11	8	6	2	2	—	—	—	—	—	—	—	—	—	—	375
18	8	11	14	25	45	42	65	53	55	35	25	13	8	12	3	1	1	2	—	—	—	—	—	—	—	—	—	418
19	7	11	21	31	59	65	80	77	65	61	47	18	27	6	7	1	1	0	—	—	—	—	—	—	—	—	—	584
20	5	10	19	28	47	53	77	54	62	54	41	26	28	25	27	8	2	4	2	—	—	—	—	—	—	—	—	575
21	7	5	17	23	48	55	65	62	90	103	87	63	63	38	44	19	10	5	3	3	1	—	—	—	—	—	—	811
22	4	3	9	26	32	35	61	54	103	84	83	71	64	44	29	12	7	2	1	1	1	—	—	—	—	—	—	726
23	0	2	6	21	27	25	47	44	87	83	82	96	87	69	65	27	17	10	4	1	0	—	—	—	—	—	—	800
24	1	0	7	7	18	13	18	26	63	71	96	136	104	95	73	30	15	11	14	4	2	1	—	—	—	—	—	805
25	—	—	2	7	11	8	27	28	63	64	87	104	132	96	92	50	43	16	20	6	4	3	—	—	—	—	—	863
26	—	—	0	2	8	12	6	25	38	44	69	95	96	82	95	43	28	15	14	4	2	1	—	—	—	—	—	679
27	—	—	2	2	6	3	7	27	44	29	65	73	92	95	76	53	27	28	25	10	1	2	0	0	1	—	—	668
28	—	—	1	0	2	1	1	8	19	12	27	30	50	43	53	28	37	38	22	5	7	2	0	0	0	—	—	386
29	—	—	1	0	2	1	1	2	10	7	17	15	43	28	27	37	36	17	22	13	4	3	0	0	2	—	—	288
30	—	—	—	—	—	2	0	4	5	2	10	11	16	15	28	38	17	26	19	6	9	6	0	0	1	—	—	215
31	—	—	—	—	—	—	—	2	3	1	4	14	20	14	25	22	22	19	8	11	3	7	6	0	1	0	2	184
32	—	—	—	—	—	—	—	—	3	1	1	4	6	4	10	5	13	6	11	8	1	7	6	0	3	0	2	91
33	—	—	—	—	—	—	—	—	1	1	0	2	4	2	1	7	4	9	3	1	0	2	0	0	0	0	0	37
34	—	—	—	—	—	—	—	—	—	—	—	1	3	1	2	2	3	6	7	7	2	2	6	0	1	0	2	45
35	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	6	6	0	6	6	0	0	0	3	27
36	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	—	—	—	—	0
37	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0	2	1	1	3	0	1	0	—	—	—	—	9
38	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	—	—	—	—	0
39	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	2	0	2	3	—	—	—	—	9
Totals	55	73	146	229	375	418	584	575	811	726	800	805	863	679	668	386	288	215	184	91	37	45	27	0	9	0	9	9098

It will be seen at once that while in the case of ceterach the variability has fallen to less than half that of the hartstongue—being close to that of the *Nigella Hispanica* and less than that of holly—still the measure of the individuality, *i.e.*, the degree of sameness in like organs, is to the limit of accuracy defined by its probable error identical with that of hartstongue. Considering the difficulties arising from environment, difference of age, and from standards of reckoning, this agreement is remarkable. It seems to me that the basis of the identity, the bulk of the agreement, must have other sources than a balance between the resemblances produced in undifferentiated like organs by such diverse causes as environment on the one hand and age on the other acting upon the individual plants. I hope that further investigations upon ferns may be forthcoming. Table XIX. contains the classified data for ceterach.

SECTION V.—*Onions* (Allium cepa), *Great Hampden.*

(19.) The veins in the successive tunics of onions occurred to me as a possible character for computation, no differentiation of the number of veins with the position of the tunic being discoverable. The number of tunics on which the veins were countable was about six, but the process of counting itself was for more than one reason a most trying one. As there were so few "like organs" to each individual, it was needful to deal with at least 200. I accordingly purchased about 250 English onions; these were in three groups, and came from as many cottage gardens. Their average horizontal diameter was about 2 inches, and they appeared very uniform in character. It is quite possible, however, that there may be some heterogeneity introduced into the series by this mixture of onions from different sources, but I did not at the time see my way to obtaining the whole series from one homogeneous sowing.* The entire work of counting and calculating on 200 of these onions was most valiantly undertaken by Dr. LEE.† Each onion was cut through horizontally; the tunics then came easily apart, and these were held up to the light, and the number of veins, ranging from 18 to 54, then counted and recorded. The following table contains the chief numerical constants:—

---

* The onions were purchased at an auction following a harvest festival; the great bulk of them had formed a large chancel crucifix constructed of two different species from one cottage garden. The English onions from this crucifix formed a most homogeneous group, having been already selected for uniformity of size and appearance. I supplemented these by the purchase of two other smaller "offerings," closely resembling them.

† If any of my readers wish to appreciate the purely *physical* difficulties of the task, they should try the effect of slicing and counting four or five onions, and they will, I think, be inclined, as I soon was, to give up the task in despair.

English Onion.

Number of			Mean No. of veins.	S. D.	Coefficient of variation.	S. D. of array.	Percentage variation.	Correlation.
Plants.	Tunics.	Pairs.						
200	1085	4924	29·8725 ± ·1066	5·2060 ± ·0754	17·4274	4·1222	79·18	·6108 [± ·0128]

The actual number of tunics with veins countable was made up as follows: 121 onions with 6, 44 with 5, 34 with 4, and 1 with only 3 available tunics. It will be seen that the onion has a degree of resemblance in its undifferentiated-like organs closely resembling that of the Chelsea Shirley poppies, of the Dorsetshire hollies (without No. 91), or the mixed Spanish chestnuts, and somewhat less than that found for the two series of ferns. Its variability approaches closely that of the wild poppy. On the whole it fits well into our series of results, the somewhat high value of the correlation approaching those of several sets, the homogeneity of which is not entirely above suspicion. The accompanying table gives the classified data from which the constants were deduced.

SECTION VI.—*Leguminous Plants.*

(20.) I had at an early stage considered that the counting of seeds in the pods of leguminous plants would be a very easy manner of testing the intensity of homotypic correlation. But as the number of seeds fully developed would depend on the extent to which the flower had been fertilised, it seemed to me, after further thinking about the problem, that unless I counted all the abortive as well as the fully developed seeds, I should find the factor of homotyposis weakened by the external chances upon which I thought cross-fertilisation must depend, and which I feared would have nothing in the main to do with the individuality of the plant. I made some attempts to count all abortive as well as fully developed seeds in certain species, but I found the task not only very laborious, but my estimates doubtful. Examining for another purpose DARWIN's 'Cross- and Self-Fertilisation of Plants.' I was struck by the fact that he did not hesitate for the purpose of comparing their vigour to count the seeds of plants living in *the open and all alike subject to free visitation from insects and bees.* He distinctly states (3rd Edn., p. 115) that "the difference in the number of the contained seeds* must depend upon the constitution of the plants." This view of the matter impressed me without entirely removing my earlier doubts. Given a number of plants living under the like conditions, and with a superabundant visitation from insects, there would be ample and equal chance of fertilisation for all

---

* DARWIN in all cases appears to have calculated the good or fully developed seed.

Table XX.—English Onion. Great Hampden.

[*To face p.* 334.

Number of Veins in First Tunic.

Number of Veins in Second Tunic.

	18.	19.	20.	21.	22.	23.	24.	25.	26.	27.	28.	29.	30.	31.	32.	33.	34.	35.	36.	37.	38.	39.	40.	41.	42.	43.	44.	45.	46.	47.	48.	49.	50.	51.	52.	53.	54.	Totals.
18	—	—	2	1	0	2	2	2	0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	9
19	—	—	0	0	0	1	1	2	3	1	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	10
20	2	0	4	4	3	3	3	6	5	0	0	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	31
21	1	0	4	6	2	5	9	14	7	7	0	1	5	0	2	1	0	2	0	0	0	1	0	0	0	1	—	—	—	—	—	—	—	—	—	—	—	68
22	0	0	3	2	2	6	17	16	23	15	15	10	6	4	2	1	0	1	1	0	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	—	—	124
23	2	1	3	5	6	8	18	27	27	16	13	14	10	3	3	0	3	2	0	1	1	1	0	0	0	0	—	—	—	—	—	—	—	—	—	—	—	164
24	2	1	3	9	17	18	14	34	29	23	20	15	10	5	2	3	0	2	1	0	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	—	—	208
25	2	2	6	14	16	27	34	32	63	30	23	23	17	12	9	2	6	7	0	1	1	1	0	0	0	1	—	—	—	—	—	—	—	—	—	—	—	329
26	—	3	5	7	23	27	29	63	64	62	48	51	31	21	15	16	9	10	2	3	1	1	0	0	0	0	—	—	—	—	—	—	—	—	—	—	—	491
27	—	1	0	7	15	16	23	30	62	52	37	45	30	33	18	15	18	16	10	8	2	3	0	0	3	1	—	—	—	—	—	—	—	—	—	—	—	445
28	—	2	0	0	15	13	20	23	48	37	52	40	33	20	17	12	10	10	11	8	6	0	1	1	0	0	—	—	—	—	—	—	—	—	—	—	—	379
29	—	—	1	1	10	14	15	23	51	45	40	86	28	33	15	13	13	9	11	8	2	2	0	2	2	0	—	—	—	—	—	—	—	—	—	—	—	424
30	—	—	—	5	6	10	10	17	31	30	33	28	40	23	27	17	20	19	13	12	4	3	4	0	4	0	—	—	—	—	—	—	—	—	—	—	—	356
31	—	—	—	0	4	3	5	12	21	33	20	33	23	20	13	16	11	8	12	6	4	3	3	1	0	0	—	—	—	—	—	—	—	—	—	—	—	251
32	—	—	—	2	2	3	2	9	15	18	17	15	27	13	14	4	24	23	15	6	7	6	3	1	1	0	—	—	—	—	—	—	—	—	—	—	—	227
33	—	—	—	1	1	0	3	2	16	15	12	13	17	16	4	22	11	10	17	10	15	10	5	2	2	2	1	—	—	—	—	—	—	—	—	—	—	207
34	—	—	—	0	0	3	0	6	9	18	10	13	20	11	24	11	20	15	8	12	8	4	3	2	3	1	2	2	0	0	0	0	0	0	0	0	2	207
35	—	—	—	2	1	2	2	7	10	16	10	9	19	8	23	10	15	12	14	12	13	4	2	0	2	2	0	0	0	0	0	0	0	0	0	0	0	195
36	—	—	—	0	1	0	1	0	2	10	11	11	13	12	15	17	8	14	10	12	15	9	2	8	9	4	1	0	1	0	0	0	1	0	0	0	0	187
37	—	—	—	0	0	1	0	1	3	8	8	8	12	6	6	10	12	12	12	14	14	4	8	3	8	2	0	0	0	0	0	0	0	0	0	0	0	152
38	—	—	—	0	0	1	0	1	1	2	6	2	4	4	7	15	8	13	15	14	14	13	7	6	7	7	3	1	0	0	0	0	0	0	0	0	1	152
39	—	—	—	1	0	1	0	1	1	3	0	2	3	3	6	10	4	4	9	4	13	4	3	2	9	3	1	0	0	0	0	0	0	0	0	0	0	87
40	—	—	—	0	0	0	0	0	0	0	1	0	4	3	3	5	3	2	2	8	7	3	2	2	2	0	1	0	0	0	0	0	0	0	0	0	0	48
41	—	—	—	0	0	0	0	0	0	0	1	2	0	1	1	2	2	0	8	3	6	2	2	4	2	2	3	0	2	0	0	0	2	0	0	0	0	45
42	—	—	—	0	0	0	0	0	0	3	0	2	4	0	1	2	3	2	9	8	7	9	2	2	4	1	1	0	0	0	0	0	0	0	0	0	0	60
43	—	—	—	1	0	0	0	1	0	1	0	0	0	0	0	2	1	2	4	2	7	3	0	2	1	2	1	0	0	0	0	0	0	0	0	0	0	30
44	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	2	0	1	0	3	1	1	3	1	1	0	1	1	0	0	0	1	0	0	0	1	18
45	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	5
46	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	1	0	0	0	0	2	0	0	1	0	0	0	0	0	1	0	0	0	0	5
47	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	1	0	0	0	0	2	0	0	1	0	1	0	0	0	0	0	0	0	0	5
51	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
52	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
53	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
54	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	5
Totals	9	10	31	68	124	164	208	329	491	445	379	424	356	251	227	207	207	195	187	152	152	87	48	45	60	30	18	5	5	0	0	0	5	0	0	0	5	4924

of them, and excess or defect from the average number of fertile seeds would depend only on the constitution of the individual plant. It would seem accordingly that in the case of cross-fertilisation ample visitation was a *sine quâ non*, and this led me to select broom in the first place.

Broom (Cytisus Scoparius), *Danby Dale.*—We collected 10 pods from each of 120 broom bushes. These were growing within some quarter of a mile of each other, on the roadside at Botton, towards the head of Danby Dale, Yorkshire.

I roughly estimated that if abortive seeds were to be included, 16 would be about the modal number of ovules; the actual average of fully developed seeds in 1200 pods was 9·6425, with a modal value at 9, so that some 6 to 7 seeds failed of fertilisation on the average in each pod. The following is the frequency distribution of the seeds in the pods:—

No. of seeds.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	Total.
Frequency .	8	15	29	47	54	77	98	126	136	134	103	100	98	67	48	33	17	5	4	1	1200

I came across no pods with an entire absence of seeds, although such may exist. I refer to this because the reader might think that such were discarded. The distribution is fairly regular, but it is clear that the variation is very large, the distribution being very flat topped. As in nearly all the cases that I have dealt with in this memoir, it is markedly skew. I reserve, however, for the present the full consideration of variation in the vegetable world, as my data extend far beyond the material considered in this paper, which is limited to the cases in which the homotypic correlation has also been worked out—a far more laborious enquiry. I would only remark that among plants and trees I know of nothing approximating even to the "normal law," and that in many cases we appear to have mixtures of local races hardly yet differentiated by the botanist.

The following table gives the constants for broom:—

Broom. Seeds in Pods.

Number of			Mean No. of seeds.	S. D. of seeds.	Coefficient of variation.	S. D. of array.	Percentage variation.	Correlation.
Plants.	Pods.	Pairs.						
120	1200	10,800	9·6425 ± ·0691	3·54655 ± ·0488	36·7804	3·22595	90·96	·4155 [± ·0161]

The actual distribution of pairs is given in the table on the following page. It will be seen that some of the arrays are rather irregular, but in its results I look upon broom to be as satisfactory as any material I have dealt with. In the first place, the table fully bears out the conclusion drawn from the simple frequency

TABLE XXI.—Broom. Danby Dale.

Number of Seeds in First Pod.

Number of Seeds in Second Pod.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	Totals.
1	8	10	9	4	2	7	9	9	5	5	3	1	—	—	—	—	—	—	—	—	72
2	10	8	20	24	11	17	13	6	7	8	5	2	1	2	0	1	—	—	—	—	135
3	9	20	34	49	25	34	16	22	18	15	9	6	4	0	0	0	—	—	—	—	261
4	4	24	49	52	39	52	37	33	40	39	19	11	11	11	1	1	—	—	—	—	423
5	2	11	25	39	40	53	60	41	61	53	37	25	16	13	4	5	0	1	—	—	486
6	7	17	34	52	53	64	63	77	89	76	50	42	33	21	7	7	1	0	—	—	693
7	9	13	16	37	60	63	88	97	116	103	84	63	52	35	25	14	4	3	—	—	882
8	9	6	22	33	41	77	97	144	133	141	106	97	106	64	36	14	8	0	—	—	1134
9	5	7	18	40	61	89	116	133	146	136	125	97	86	67	47	27	18	4	2	—	1224
10	5	8	15	39	53	76	103	141	136	148	122	107	96	68	38	28	17	3	2	1	1206
11	3	5	9	19	37	50	84	106	125	122	88	77	91	51	32	20	7	0	0	1	927
12	1	2	6	11	25	42	63	97	97	107	77	88	107	72	47	31	14	8	4	1	900
13	—	1	4	11	16	33	52	106	86	96	91	107	82	76	61	37	16	4	3	0	882
14	—	2	0	11	13	21	35	64	67	68	51	72	76	32	39	35	13	2	1	1	603
15	—	0	0	1	4	7	25	36	47	38	32	47	61	39	40	28	17	4	4	2	432
16	—	1	0	1	5	7	14	14	27	28	20	31	37	35	28	20	16	5	6	2	297
17	—	—	—	—	0	1	4	8	18	17	7	14	16	13	17	16	10	4	7	1	153
18	—	—	—	—	1	0	3	0	4	3	0	8	4	2	4	5	4	2	5	0	45
19	—	—	—	—	—	—	—	—	2	2	0	4	3	1	4	6	7	5	2	—	36
20	—	—	—	—	—	—	—	—	—	1	1	1	0	1	2	2	1	0	—	—	9
Totals...	72	135	261	423	486	693	882	1134	1224	1206	927	900	882	603	432	297	153	45	36	9	10800

distribution of the high degree of variability in the broom pod. We have so far reached nothing but the *sori* on the Hartstongue fern with anything like such a high coefficient of variation. We might expect, if Mr. ADAM SEDGEWICK'S view were correct and great variation meant small intensity of heredity, that homotyposis, of which heredity is only a special case, would also be small if variation were large. But we find nothing of the kind. The resemblance of undifferentiated like organs reaches the value ·4155, fully equal to that of the *pinnæ* of the ash, which have less than half the variability of broom pods. This value is close also to the ·4000 given by the Ancestral Law of Heredity for the degree of resemblance between brothers. In the next place this normal behaviour as to the degree of resemblance of like organs is associated with a variability in the individual which amounts to upwards of 90 per cent. of the racial variability. It is impossible to form standard deviations for groups of 10, but if the reader will reduce any array of the above table to the number 10, rejecting all fractions of unity, he will obtain quite fair samples of what I actually found in the case of individual pods. I think we may feel fairly confident that the variability of a race is not in inverse proportion to either its heredity or its homotyposis.

It will be seen that broom, so far as it goes, supports the view held by DARWIN that the number of ripe seeds is a measure of individual constitution. At the same time the question of self-fertilisation arises. Would completely self-fertilised plants exhibit full homotyposis? May not outside influences—wind, shaking due to visits of insects or other causes of a random character—be also needful for the ripening of the seeds even in the case of self-fertilisation? Again, can we always suppose that a plentiful visitation will take place in the case of all cross-fertilised plants? I must confess that the value for the homotyposis found in the case of broom did not suffice to remove all my doubts as to any character depending on fertilisation being a suitable one for the determination of the intensity of homotyposis. The subject, however, is one of such great interest that it deserves an independent and fuller treatment than can be provided here.*

V. CROSS-HOMOTYPIC CORRELATION.

(21.) *Mushroom* (Agaricus campestris).—A more complete study was now made of two species—the mushroom and the ivy. We found some difficulty in discovering two easily measurable or countable characters in one organ, neither of which was largely influenced by the growth of the organ or the age of the organism. In selecting an organ in the mushroom, I was guided by the desire to take a simple organism, and an organ upon which fertilisation had no influence. The gill of the mushroom seemed to satisfy to some extent these conditions. Had we been able to grow our own mushrooms we might have succeeded in taking them all at the same stage of

---

* This has been to some extent provided in an Appendix to this memoir, added since its completion.

development, and our results would in this case have been more homogeneous and thus more satisfactory. But simply gathering them in the fields we could only make a very rough approximation to uniformity in age or development. We aimed at a condition in which the underskin had disappeared and the gills had lost all signs of a crease or notch. Dr. LEE and I collected more than 100 mushrooms in the Great and Little Hampdens. With the assistance of my wife we cut 10 gills out of each of these, distributed, as far as the condition of the mushroom would allow, uniformly round the axis. These gills were then placed on ruled paper, the horizontal lines of which served to mark the maximum length of the gill. This was marked by two fine needle-pricks, one at either end. Two further needle-pricks were now made at the points farthest removed from this length line, one on the upper and one on the lower contour of the gill. Thus by drawing through the four needle points lines parallel and perpendicular to the paper-ruling we obtained a rectangle approximately circumscribing the gill. The length and breadth of this rectangle are what I shall term the length and breadth of the gill. At first I endeavoured to take these measurements by the use of a micrometer microscope, but this method was very laborious, and gave an apparent exactness wholly out of keeping with the somewhat rough and ready manipulation of the gill. Even when the gill itself was placed under the microscope the gill length could only be placed parallel to the spider wires by a somewhat vague appreciation, and the upper boundary of the gill was under the lens rather less definite than without it. Accordingly we resorted to the needle-pricks and the ruling by fine lines of the above-mentioned rectangle. I determined, after various trials with engraved glass scales, &c., the length and breadth of these rectangles by the use of a pair of proportional compasses, set to multiply by ten, and an ordinary millimetre scale. I was thus able to get readings to ·01 of a millimetre, which agreed well on repetition. As the millimetre itself was afterwards adopted as the unit of grouping, I do not think the errors of manipulation will seriously affect our results. The needle-prick method enabled us to deal with the mushrooms quite fresh and reserve the measuring till another season.* Dr. LEE found what we may term the *gill-index*, *i.e.*, the ratio of length to breadth for the 1070 gills, by aid of a FULLER'S slide rule. We had thus three characters to deal with—length, breadth, and index. Of these we considered that the first two would undoubtedly be largely influenced by the stage of development of the mushroom, but we hoped that the index (like, for example, the cephalic index of children over two years of age) would not be much influenced by the growth factor. On the other hand, the correlation of lengths of gills or the correlation of breadths of gills from the same mushroom would of course be influenced by growth as well as

---

* In all such systems of mass-measurement, the method adopted has to be chosen with due regard to time and eyesight. The actual measurement of the gills occupied me about ten full days, the microscopic method would have taken about four times the time, even if my eyes would allow of such continuous working and the mushrooms could have been preserved fresh.

by individuality. Thus we should find the degree of resemblance between two gills of the same mushroom, when measured by length or breadth, to be influenced not only by the individuality of the mushroom, but by the stage of its growth, and accordingly likely to exceed, and possibly considerably exceed, the value about ·5, which we have determined in several other cases.*

In order to make a more complete study of the influence of stages of growth on the degree of resemblance of undifferentiated like organs, Dr. Lee with her characteristic energy undertook the five needful tables.

They are the following :—

(i.) Organic correlation of the length and breadth of the same gill.
(ii.) Homotypic correlation of the lengths of two gills of the same mushroom.
(iii.) Homotypic correlation of the breadths of two gills of the same mushroom.
(iv.) Cross homotypic correlation of the length of one, and the breadth of a second gill of the same mushroom.
(v.) Homotypic correlation of the indices of two gills of the same mushroom.

Of these tables the cross homotypic correlation between length and breadth is a very laborious business, offering no means of readily testing the accuracy of the entries until the whole table is completed and its columns and rows added up. Ultimately the frequency distributions (as given by row or column marked totals) for the lengths and breadths in tables (ii.), (iii.), and (iv.), ought to be $n - 1$ times the frequency distributions of lengths and breadths as given by the like row and column in the organic correlation table (i.), n being the number of organs measured in each individual; in our case this is ten.

I begin by considering the results obtained for the degree of resemblance of gill-indices, *i.e.*, Table (v.) referred to above. The following were the results obtained :—

Mushroom. Gill Indices.

Number of			Mean index.	S. D.	Coefficient of variation.	S. D. of array.	Percentage variation.	Correlation.
Mushrooms.	Gills.	Pairs.						
107	1070	9630	2·8695 ± ·0108	·5246 ± ·0076	18·2819	·4385	83·58	·5490 [± ·0144]

The frequency distribution of the gills was the following :—

---

* The absurdity of disregarding the growth-factor may be realised if we consider how much the degree of resemblance in brethren would be increased if we took pairs of brothers, differing, say, by less than two years, and these pairs were scattered at all ages from four to twenty!

Gill Index.

	1·4995–1·5995.	1·5995–1·6995.	1·6995–1·7995.	1·7995–1·8995.	1·8995–1·9995.	1·9995–2·0995.	2·0995–2·1995.	2·1995–2·2995.	2·2995–2·3995.	2·3995–2·4995.	2·4995–2·5995.	2·5995–2·6995.	2·6995–2·7995.	2·7995–2·8995.	2·8995–2·9995.
Frequency .	1	1	4	7	13	18	30	39	55	78	101	90	100	93	73

	2·9995–3·0995.	3·0995–3·1995.	3·1995–3·2995.	3·2995–3·3995.	3·3995–3·4995.	3·4995–3·5995.	3·5995–3·6995.	3·6995–3·7995.	3·7995–3·8995.	3·8995–3·9995.	3·9995–4·0995.	4·0995–4·1995.	4·1995–4·2995.	4·2995–4·3995.	4·3995–4·4995.
Frequency .	66	62	39	43	31	43	15	20	10	9	1	8	3	3	4

	4·4995–4·5995.	4·5995–4·6995.	4·6995–4·7995.	4·7995–4·8995.	4·8995–4·9995.	4·9995–5·0995.	5·0995–5·1995.	5·1995–5·2995.	5·2995–5·3995.	5·3995–5·4995.	5·4995–5·5995.	5·5995–5·6995.	Total.
Frequency .	0	3	1	2	0	1	0	1	0	1	0	1	1070

This distribution is rather irregular and may suggest heterogeneity; the tail is somewhat inordinately extended. But the numerical constants given in the table above are quite comparable with the best of our other series, and are rather better than I had anticipated *a priori*, for I expected the diverse stages of growth would influence the index as well as the absolute measures. It would be of great interest to repeat the series on an artificially grown bed of mushrooms, taking each at the same stage.

It will be seen that the mushroom, with degree of resemblance of like organs about that of the Spanish chestnut, has a variability comparable with that of the ash. The variation of the individual is even in such a comparatively simple organism more than 80 p.c. of that of the race. It will thus be clear that the mushroom takes, so far as our present investigation is concerned, no special place apart in the vegetable world, but quantitatively is comparable with highly developed organisms like trees. The accompanying table, Table XXII., gives the data for the measurement of the gills.

I now turn to the relationships of the absolute measurements, which, as I have noted, we have good grounds for believing are much affected by growth.

The numerical constants are summed up in the following table, the dimensions being given in millimetres :—

Table XXII.—Mushroom Gill-indices.

[*To face p.* 340.

Index of First Gill.

Index of Second Gill.	1·5495.	1·6495.	1·7495.	1·8495.	1·9495.	2·0495.	2·1495.	2·2495.	2·3495.	2·4495.	2·5495.	2·6495.	2·7495.	2·8495.	2·9495.	3·0495.	3·1495.	3·2495.	3·3495.	3·4495.	3·5495.	3·6495.	3·7495.	3·8495.	3·9495.	4·0495.	4·1495.	4·2495.	4·3495.	4·4495.	4·5495.	4·6495.	4·7495.	4·8495.	4·9495.	5·0495.	5·1495.	5·2495.	5·3495.	5·4495.	5·5495.	5·6495.	Totals.
1·5495	—	—	—	—	1	1	0	2	1	1	0	2	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	9
1·6495	—	—	1	2	1	2	2	1	0	0	0	0	0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	9
1·7495	—	1	2	2	7	6	10	5	1	2	0	0	0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	36
1·8495	—	2	2	8	5	12	9	6	5	4	5	2	2	0	0	0	0	0	0	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	63
1·9495	1	1	7	5	14	18	16	11	20	7	5	4	3	0	1	3	0	0	1	0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	117
2·0495	1	2	6	12	18	10	26	25	21	9	12	3	6	1	1	2	0	1	0	3	1	0	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	162
2·1495	0	2	10	9	16	26	20	35	35	26	29	8	16	12	7	5	2	3	2	4	2	0	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	270
2·2495	2	1	5	6	11	25	35	40	42	28	44	28	32	15	8	6	7	6	1	3	3	0	3	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	351
2·3495	1	0	1	5	20	21	35	42	48	68	72	47	44	29	15	15	6	5	7	4	3	1	2	1	1	0	0	0	1	1	—	—	—	—	—	—	—	—	—	—	—	—	495
2·4495	1	0	2	4	7	9	26	28	68	68	102	75	62	79	38	46	27	13	13	7	12	3	6	0	3	0	2	0	1	0	—	—	—	—	—	—	—	—	—	—	—	—	702
2·5495	0	0	0	5	5	12	29	44	72	102	124	124	104	90	47	50	32	17	15	11	9	5	3	1	4	0	3	0	0	0	0	0	0	0	0	0	0	0	0	1	—	—	909
2·6495	2	0	0	2	4	3	8	28	47	75	124	138	100	84	50	42	36	17	10	4	11	1	2	4	7	0	2	0	2	4	0	1	0	0	0	0	0	0	0	2	—	—	810
2·7495	1	0	0	2	3	6	16	32	44	62	104	100	112	96	78	54	69	29	23	18	24	3	13	1	3	0	1	0	1	2	0	1	0	1	0	0	0	0	0	1	—	—	900
2·8495	—	—	—	0	0	1	12	15	29	79	90	84	96	98	74	63	60	26	34	16	23	9	9	3	3	0	2	2	5	2	0	0	0	1	0	0	0	0	0	1	—	—	837
2·9495	—	—	—	0	1	1	7	8	15	38	47	50	78	74	58	56	50	27	43	22	44	10	13	5	2	0	2	0	0	1	0	3	0	1	0	0	0	0	0	1	—	—	657
3·0495	—	—	—	0	3	2	5	6	15	46	50	42	54	63	56	44	55	32	25	20	29	16	10	5	7	2	4	0	0	3	0	0	0	0	0	0	0	0	0	0	—	—	594
3·1495	—	—	—	0	0	0	2	7	6	27	32	36	69	60	50	55	48	31	40	26	22	12	8	5	7	0	1	0	6	0	0	2	1	1	0	1	0	1	0	1	0	1	558
3·2495	—	—	—	0	0	1	3	6	5	13	17	17	29	26	27	32	31	18	20	23	20	13	21	10	4	1	2	4	2	1	0	3	0	1	0	0	0	0	0	0	0	1	351
3·3495	—	—	—	0	1	0	2	1	7	13	15	10	23	34	43	25	40	20	28	22	42	13	12	9	6	0	4	0	4	2	0	3	1	3	0	1	0	1	0	1	0	1	387
3·4495	—	—	—	1	0	3	4	3	4	7	11	4	18	16	22	20	26	23	22	22	16	7	12	10	5	2	6	3	0	4	0	2	1	2	0	1	0	1	0	1	0	0	279
3·5495	—	—	—	—	—	1	2	3	3	12	9	11	24	23	44	29	22	20	42	16	58	15	23	7	2	2	11	2	0	2	0	3	0	1	0	0	0	0	0	0	0	0	387
3·6495	—	—	—	—	—	0	0	0	1	3	5	1	3	9	10	16	12	13	13	7	15	4	10	4	2	0	1	0	3	2	0	0	0	1	0	0	0	0	0	0	0	0	135
3·7495	—	—	—	—	—	2	1	3	2	6	3	2	13	9	13	10	8	21	12	12	23	10	10	7	1	0	2	6	0	2	0	1	0	1	0	0	0	0	0	0	0	0	180
3·8495	—	—	—	—	—	—	—	—	1	0	1	4	1	3	5	5	5	10	9	10	7	4	7	4	2	0	6	3	0	1	0	1	0	1	0	0	0	0	0	0	0	0	90
3·9495	—	—	—	—	—	—	—	—	1	3	4	7	3	3	2	7	7	4	6	5	2	2	1	2	2	0	5	1	0	3	0	2	2	1	0	2	0	2	0	0	0	2	81
4·0495	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	2	0	1	0	2	2	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	9
4·1495	—	—	—	—	—	—	—	—	0	2	3	2	1	2	2	4	1	2	4	6	11	1	2	6	5	1	12	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	72
4·2495	—	—	—	—	—	—	—	—	0	0	0	0	0	2	0	0	0	4	0	3	2	0	6	3	1	0	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27
4·3495	—	—	—	—	—	—	—	—	1	1	0	2	1	5	0	0	6	2	4	0	0	3	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	27
4·4495	—	—	—	—	—	—	—	—	1	0	0	4	2	2	1	3	0	1	2	4	2	2	2	1	3	1	1	0	0	2	0	0	0	2	0	0	0	0	0	0	0	0	36
4·5495	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4·6495	—	—	—	—	—	—	—	—	—	—	0	1	1	0	3	0	2	3	3	2	3	0	1	1	2	0	0	0	0	0	0	0	1	1	0	1	0	1	0	0	0	1	27
4·7495	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	1	0	1	1	0	0	0	0	2	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	1	9
4·8495	—	—	—	—	—	—	—	—	—	—	0	0	1	1	1	0	1	1	3	2	1	1	1	1	1	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	18
4·9495	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5·0495	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	1	0	1	1	0	0	0	0	2	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	1	9
5·1495	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5·2495	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	1	0	1	1	0	0	0	0	2	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	1	9
5·3495	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	0
5·4495	—	—	—	—	—	—	—	—	—	—	1	2	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	9
5·5495	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	0
5·6495	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	0	0	0	0	0	2	0	0	0	0	0	0	1	1	0	0	1	0	1	—	—	—	—	9
Totals...	9	9	36	63	117	162	270	351	495	702	909	810	900	837	657	594	558	351	387	279	387	135	180	90	81	9	72	27	27	36	0	27	9	18	0	9	0	9	0	9	0	9	9630

N.B.—1·5495 denotes the group formed of all indices from 1·4995 to 1·5995, or since the ratio, length/breadth, was formed to three decimal places, all values of the index from 1·500 to 1·599 inclusive.

Absolute Dimensions of Mushroom Gills.

Character.	Mean.	S. D.	Coefficient of variation.	S. D. of array.	Percentage variation.	Direct homotypic correlation.
Length. .	19·1847 ± ·0831	4·0282 ± ·0587	20·9969	2·0510	50·92	·8607 [± ·0053]
Breadth .	6·8529 ± ·0331	1·6047 ± ·0234	23·4164	1·0859	67·67	·7363 [± ·0094]

Cross-correlations.

Length and breadth, organic . .	·7000 ± ·0105
Length and breadth, homotypic .	·6275[± ·0125]

The results here given seem to be of considerable suggestiveness. While the two means and standard deviations differ very considerably, the coefficients of variation for the length and breadth of the gill are approximately equal, and in addition very high. This is precisely what we might expect when dealing with a series of organisms in various stages of growth. Further, the variability of the individual as compared with that of the race is reduced immensely below the 80 to 90 per cent. with which our previous investigations have made us familiar. This might again be anticipated as a result of heterogeneity in the stage of growth. The organic correlation of length and breadth, as well as the homotypic correlations of the same two characters, both direct and cross, are all very high, and the latter are much beyond what we might legitimately put to the credit of the pure homotyposis factor. It seems, therefore, somewhat hazardous to consider how far they satisfy the relationship suggested in the introductory part of this paper. The direct homotypic correlations for length and breadth are not approximately equal as we might expect; this suggests that growth may be a somewhat more marked factor in the length than the breadth of the gill. If we take the product of direct homotypic correlation with the organic correlation, this ought, on our hypothesis, to give us roughly the cross-homotypic correlation. We have the following results :—

Product of direct homotypic length correlation with organic length-breadth correlation . . . . = ·6025.
Product of direct homotypic breadth correlation with organic length-breadth correlation . . . . = ·5154.
Product of mean of direct homotypic breadth and length correlation with organic length-breadth correlation . . . . = ·5590.

These numbers have to be compared with the cross homotypic length-breadth correlation, *i.e.*, with ·6275.

TABLE XXIII.—Mushroom. Length of Gill.

Length of First Gill.

Length of Second Gill.

Hundredths of millimetres.	899·5–999·5.	999·5–1099·5.	1099·5–1199·5.	1199·5–1299·5.	1299·5–1399·5.	1399·5–1499·5.	1499·5–1599·5.	1599·5–1699·5.	1699·5–1799·5.	1799·5–1899·5.	1899·5–1999·5.	1999·5–2099·5.	2099·5–2199·5.	2199·5–2299·5.	2299·5–2399·5.	2399·5–2499·5.	2499·5–2599·5.	2599·5–2699·5.	2699·5–2799·5.	2799·5–2899·5.	2899·5–2999·5.	2999·5–3099·5.	3099·5–3199·5.	3199·5–3299·5.	3299·5–3399·5.	3399·5–3499·5.	Totals.
899·5–999·5	—	1	8	2	6	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	18
999·5–1099·5	1	2	6	4	6	4	4	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	27
1099·5–1199·6	8	6	68	30	37	29	8	3	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	189
1199·5–1299·5	2	4	30	22	32	41	23	15	7	4	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	180
1299·5–1399·5	6	6	37	32	78	93	52	26	10	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	342
1399·5–1499·5	1	4	29	41	93	172	141	76	43	21	15	7	3	2	—	—	—	—	—	—	—	—	—	—	—	—	648
1499·5–1599·5	—	4	8	23	52	141	146	183	91	36	11	5	6	4	1	—	—	—	—	—	—	—	—	—	—	—	711
1599·5–1699·5	—	—	3	15	26	76	183	236	180	94	59	17	15	4	1	—	—	—	—	—	—	—	—	—	—	—	909
1699·5–1799·5	—	—	—	7	10	43	91	180	250	214	140	43	41	11	9	0	2	1	0	0	2	—	—	—	—	—	1044
1799·5–1899·5	—	—	—	4	2	21	36	94	214	178	191	67	82	43	16	1	2	1	0	0	2	—	—	—	—	—	954
1899·5–1999·5	—	—	—	—	—	15	11	59	140	191	172	110	97	53	15	6	3	0	0	0	1	—	—	—	—	—	873
1999·5–2099·5	—	—	—	—	—	7	5	17	43	67	110	228	150	85	14	6	3	15	4	0	2	—	—	—	—	—	756
2099·5–2199·5	—	—	—	—	—	3	6	15	41	82	97	150	204	134	39	27	14	4	1	0	2	—	—	—	—	—	819
2199·5–2299·5	—	—	—	—	—	2	4	4	11	43	53	85	134	120	61	46	21	20	16	0	1	—	—	—	—	—	621
2299·5–2399·5	—	—	—	—	—	—	1	1	9	16	15	14	39	61	42	55	33	25	10	8	4	—	—	—	—	—	333
2399·5–2499·5	—	—	—	—	—	—	—	—	0	1	6	6	27	46	55	92	62	35	16	11	3	—	—	—	—	—	360
2499·5–2599·5	—	—	—	—	—	—	—	—	2	2	3	3	14	21	33	62	36	34	11	13	9	—	—	—	—	—	243
2599·5–2699·5	—	—	—	—	—	—	—	—	1	1	0	15	4	20	25	35	34	54	24	8	12	6	0	4	—	—	243
2699·5–2799·5	—	—	—	—	—	—	—	—	0	0	0	4	1	16	10	16	11	24	20	5	3	1	2	3	0	1	117
2799·5–2899·5	—	—	—	—	—	—	—	—	0	0	0	0	0	0	8	11	13	8	5	8	1	0	0	0	0	0	54
2899·5–2999·5	—	—	—	—	—	—	—	—	2	2	1	2	2	1	4	3	9	12	3	1	10	11	4	12	0	2	81
2999·5–3099·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	6	1	0	11	6	2	9	0	1	36
3099·5–3199·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	2	0	4	2	2	6	0	2	18
3199·5–3299·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	4	3	0	12	9	6	8	0	3	45
3299·5–3399·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	—	—	0
3399·5–3499·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0	2	1	2	3	—	—	9
Totals . . .	18	27	189	180	342	648	711	909	1044	954	873	756	819	621	333	360	243	243	117	54	81	36	18	45	0	9	9630

TABLE XXIV.—Mushroom. Breadth of Gill.

Breadth of First Gill.

Breadth of Second Gill.

Hundreths of millimetres.	299·5–349·5.	349·5–399·5.	399·5–449·5	449·5–499·5	499·5–549·5.	549·5–599·5.	599·5–649·5.	649·5–699·5.	699·5–749·5.	749·5–799·5.	799·5–849·5.	849·5–899·5.	899·5–949·5.	949·5–999·5.	999·5–1049·5.	1049·5–1099·5.	1099·5–1149·5.	1149·5–1199·5.	1199·5–1249·5.	1249·5–1299·5.	Totals.
299·5– 349·5	14	22	4	15	11	3	2	1	—	—	—	—	—	—	—	—	—	—	—	—	72
349·5– 399·5	22	32	22	32	25	15	8	3	2	1	—	—	—	—	—	—	—	—	—	—	162
399·5– 449·5	4	22	44	56	69	40	39	14	14	1	1	1	1	—	—	—	—	—	—	—	306
449·5– 499·5	15	32	56	70	88	67	60	24	29	7	5	5	2	2	4	0	2	—	—	—	468
499·5– 549·5	11	25	69	88	250	206	181	128	55	24	5	2	0	0	0	0	0	—	—	—	1044
549·5– 599 5	3	15	40	67	206	158	192	167	75	34	10	4	1	0	0	0	0	—	—	—	972
599·5– 649·5	2	8	39	60	181	192	252	243	129	56	23	11	6	5	4	2	2	—	—	—	1215
649·5– 699·5	1	3	14	24	128	167	243	286	197	92	68	26	10	6	0	4	0	—	—	—	1269
699·5– 749·5	—	2	14	29	55	75	129	197	164	155	108	49	18	11	2	0	0	—	—	—	1008
749·5– 799·5	—	1	1	7	24	34	56	92	155	176	157	59	33	14	12	5	1	0	0	1	828
799·5– 849·5	—	—	1	5	5	10	23	68	108	157	170	116	74	45	19	9	0	0	0	0	810
849·5– 899·5	—	—	1	5	2	4	11	26	49	59	116	82	55	46	30	16	1	0	0	1	504
899·5– 949·5	—	—	1	2	0	1	6	10	18	33	74	55	62	57	26	18	1	1	3	1	369
949·5– 999·5	—	—	—	2	0	0	5	6	11	14	45	46	57	28	22	19	2	1	3	0	261
999·5–1049·5	—	—	—	4	0	0	4	0	2	12	19	30	26	22	24	13	2	0	0	4	162
1049·5–1099 5	—	—	—	0	0	0	2	4	0	5	9	16	18	19	13	14	3	3	9	2	117
1099·5–1149·5	—	—	—	2	0	0	2	0	0	1	0	1	1	2	2	3	0	1	3	—	18
1149·5–1199·5	—	—	—	—	—	—	—	—	—	0	0	0	1	1	0	3	1	0	3	—	9
1199·5–1249·5	—	—	—	—	—	—	—	—	—	0	0	0	3	3	0	9	3	3	6	—	27
1249·5–1299·5	—	—	—	—	—	—	—	—	—	1	0	1	1	0	4	2	—	—	—	—	9
Totals . . .	72	162	306	468	1044	972	1215	1269	1008	828	810	504	369	261	162	117	18	9	27	9	9630

TABLE XXV.—Mushroom. Length and Breadth of Gill, Organic Correlation.

Length of Gill.

Breadth of Gill.

Hundredths of millimetres.	899·5–999·5.	999·5–1099·5.	1099·5–1199·5.	1199·5–1299·5.	1299·5–1399·5.	1399·5–1499·5.	1499·5–1599·5.	1599·5–1699·5.	1699·5–1799·5.	1799·5–1899·5.	1899·5–1999·5.	1999·5–2099·5.	2099·5–2199·5.	2199·5–2299·5.	2299·5–2399·5.	2399·5–2499·5.	2499·5–2599·5.	2599·5–2699·5.	2699·5–2799·5.	2799·5–2899·5.	2899·5–2999·5.	2999·5–3099·5.	3099·5–3199·5.	3199·5–3299·5.	3299·5–3399·5.	3399·5–3499·5.	Totals.
299·5– 349·5	1	0	4	0	1	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	8
349·5– 399·5	0	2	0	3	7	1	1	0	1	1	1	0	0	1	—	—	—	—	—	—	—	—	—	—	—	—	18
399·5– 449·5	0	1	2	3	5	5	7	3	1	3	1	0	2	1	—	—	—	—	—	—	—	—	—	—	—	—	34
449·5– 499·5	1	0	7	2	5	7	9	6	8	4	1	0	1	1	—	—	—	—	—	—	—	—	—	—	—	—	52
499·5– 549·5	—	—	6	5	2	16	16	22	17	17	10	2	2	1	—	—	—	—	—	—	—	—	—	—	—	—	116
549·5– 599·5	—	—	0	2	6	14	12	21	15	18	5	8	3	1	1	0	2	—	—	—	—	—	—	—	—	—	108
599·5– 649·5	—	—	1	2	6	11	13	20	25	18	16	4	7	2	2	2	1	3	1	0	1	—	—	—	—	—	135
649·5– 699·5	—	—	1	1	3	12	6	16	28	16	15	20	9	4	4	1	1	2	1	0	1	—	—	—	—	—	141
699·5– 749·5	—	—	—	1	3	3	6	5	10	21	17	14	10	9	3	5	3	1	1	0	0	—	—	—	—	—	112
749·5– 799·5	—	—	—	1	0	1	2	4	5	3	14	14	22	10	2	5	2	5	1	1	0	—	—	—	—	—	92
799·5– 849·5	—	—	—	—	—	1	5	2	3	2	11	13	16	11	9	8	3	2	1	2	0	0	0	1	—	—	90
849·5– 899·5	—	—	—	—	—	—	1	1	1	2	2	5	14	10	4	7	1	3	0	1	3	1	0	0	—	—	56
899·5– 949·5	—	—	—	—	—	—	—	0	1	1	2	2	3	5	8	5	5	5	1	1	1	0	0	1	—	—	41
949·5– 999·5	—	—	—	—	—	—	—	1	0	0	2	0	2	6	2	4	5	2	3	0	0	1	0	1	—	—	29
999·5–1049·5	—	—	—	—	—	—	—	—	1	0	0	0	0	4	1	2	3	2	2	0	0	0	2	0	0	1	18
1049·5–1099·5	—	—	—	—	—	—	—	—	—	—	—	2	0	3	0	1	1	1	1	1	2	0	0	1	—	—	13
1099·5–1149·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0	0	0	0	0	0	1	0	0	—	—	2
1149·5–1199·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0	0	0	0	0	0	—	—	1
1199·5–1249·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	1	1	0	1	—	—	3
1249·5–1299·5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	1
Totals	2	3	21	20	38	72	79	101	116	106	97	84	91	69	37	40	27	27	13	6	9	4	2	5	0	1	1070

TABLE XXVI.—Mushroom. Length and Breadth of Gill, Homotypic Correlation.

Length of First Gill.

Breadth of Second Gill.

Hundredths of millimetres.	899·5–999·5.	999·5–1099·5.	1099·5–1199·5.	1199·5–1299·5.	1299·5–1399·5.	1399·5–1499·5.	1499·5–1599·5.	1599·5–1699·5.	1699·5–1799·5.	1799·5–1899·5.	1899·5–1999·5.	1999·5–2099·5.	2099·5–2199·5.	2199·5–2299·5.	2299·5–2399·5.	2399·5–2499·5.	2499·5–2599·5.	2599·5–2699·5.	2699·5–2799·5.	2799·5–2899·5.	2899·5–2999·5.	2999·5–3099·5.	3099·5–3199·5.	3199·5–3299·5.	3299·5–3399·5.	3399·5–3499·5.	Totals.
299·5– 349·5	3	4	22	9	10	8	7	6	3	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	72
349·5– 399·5	4	7	24	10	31	24	17	10	7	8	4	2	6	5	3	—	—	—	—	—	—	—	—	—	—	—	162
399·5– 449·5	0	5	14	12	18	45	60	49	29	31	13	3	12	12	3	—	—	—	—	—	—	—	—	—	—	—	306
449·5– 499·5	3	5	41	20	64	55	72	55	40	38	24	8	21	12	10	—	—	—	—	—	—	—	—	—	—	—	468
499·5– 549·5	2	4	44	29	39	129	112	158	158	141	107	51	36	16	5	2	2	7	0	0	2	—	—	—	—	—	1044
549·5– 599·5	2	2	11	24	57	107	112	159	150	123	87	47	39	25	6	2	3	11	1	1	3	—	—	—	—	—	972
599·5– 649·5	2	0	19	34	53	107	127	179	197	142	115	78	52	25	15	15	8	35	4	4	4	—	—	—	—	—	1215
649·5– 699·5	1	0	10	24	41	86	81	149	195	175	136	140	104	60	17	8	12	14	4	4	8	—	—	—	—	—	1269
699·5– 749·5	1	0	4	17	27	53	55	64	108	124	120	104	127	82	23	41	21	16	6	10	5	—	—	—	—	—	1008
719·5– 799·5	—	—	—	1	2	10	27	39	71	56	88	136	162	92	32	42	23	14	16	7	3	1	2	3	0	1	828
799·5– 849·5	—	—	—	—	—	24	30	25	39	53	86	114	115	81	51	70	53	32	16	10	5	1	2	2	0	1	810
849·5– 899·5	—	—	—	—	—	—	7	11	22	30	47	33	69	70	51	60	43	22	14	5	7	1	4	6	0	2	504
899·5– 949·5	—	—	—	—	—	—	0	0	9	12	19	22	25	53	57	55	35	32	18	8	6	5	4	7	0	2	369
949·5– 999·5	—	—	—	—	—	—	2	2	8	12	13	14	26	42	35	35	19	22	12	2	7	3	2	4	0	1	261
999·5–1049·5	—	—	—	—	—	—	2	3	6	7	4	1	14	24	16	17	13	12	15	2	8	3	4	9	0	2	162
1049·5–1099·5	—	—	—	—	—	—	—	—	0	0	9	3	10	18	7	12	10	15	9	1	9	9	0	5	—	—	117
1099·5–1149·5	—	—	—	—	—	—	—	—	2	2	1	0	1	1	2	0	0	2	0	0	3	2	0	2	—	—	18
1149·5–1199·5	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	1	0	0	3	3	0	2	—	—	9
1199·5–1249·5	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	6	0	0	8	8	0	5	—	—	27
1249·5–1299·5	—	—	—	—	—	—	—	—	—	—	—	—	—	3	0	1	1	2	2	—	—	—	—	—	—	—	9
Totals . . .	18	27	189	180	312	648	711	909	1044	954	873	756	819	621	333	360	243	243	117	54	81	36	18	45	0	9	9630

It will be seen at once that they are sensibly too small (say, ·56 as compared with ·63). But having recognised the influence of heterogeneity in the growth stages of our material, the divergence is, I take it, of a magnitude rather to confirm than confute the hypothesis as to the relationship of direct and cross homotypic correlations. It must be remembered that we are dealing with *average* results. Our main proposition is that heredity is not a factor of life peculiar to sexual reproduction, but merely a phase of the larger factor, which we have termed homotyposis, or the tendency of the individual to put forth undifferentiated like organs with a certain degree of resemblance. Our subsidiary proposition is: that if the direct homotypic correlation oscillates about a certain mean value, and the cross homotypic correlation be the product of organic and direct homotypic correlations, then we should expect to find the average degree of resemblance of brethren equal to the average degree of resemblance of undifferentiated like organs in the individual.

All then I think we can safely say for the gills of mushrooms is that the relation of the direct and cross homotypic correlations is by no means such that it condemns our hypothesis, or enforces us to reject our subsidiary proposition. The difficulty lies in finding *adult* organisms with undifferentiated like organs with two characters in sufficient quantity, easily counted or measured, upon which the hypothesis can be tested. The investigation of the length and breadth of ivy leaves, to which we shall soon turn, suffers to some extent from the same defects as that for the gills of mushrooms.

Tables XXIII.–XXVI. contain the length-breadth data for the mushrooms. The frequency distributions for length and breadth are given in the last row and column of Table XXV. above, and exhibit in their irregularity something of the heterogeneity of growth to which I have referred. Of the two distributions, I consider that for the breadth as the more irregular and consequently the less satisfactory. It is, of course, harder to determine a definite breadth for the gill than a definite length, and I should be well content to compare the product of the homotypic length correlation and the organic length-breadth correlation, *i.e.*, ·6025, with the cross homotypic length-breadth correlation, *i.e.*, ·6275, as the best basis for our subsidiary proposition available from these measurements on mushrooms.

(22.) *Wild Ivy* (Hedera Helix).—This series was originally undertaken by Dr. E. WARREN, his object being to measure the degree of resemblance between ivy leaves on the same plant, using as his character the index, or ratio of length to maximum breadth. It was hoped that in this manner, the growth factor might be fairly well eliminated.

The figures (page 348) indicate how the breadth and length were determined by Dr. WARREN in leaves of somewhat different shape.

It will be seen that, as in the case of the gill of the mushroom, the breadth is taken as the maximum breadth between tangents to the contour parallel to the length, and these tangents do not necessarily go through points like *a*, *a*. Twenty-five leaves were taken from each of 42 plants by Dr. WARREN, in the neighbourhood of Canterbury,

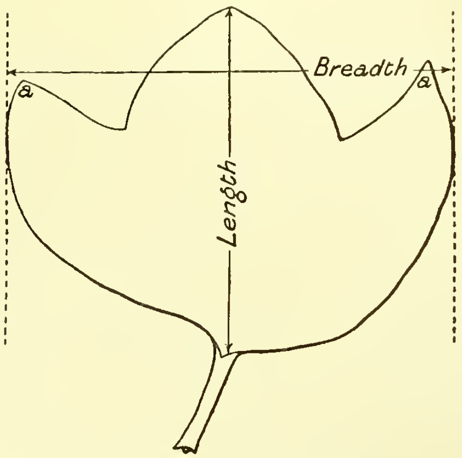
TABLE XXVII.—Ivy. Leaf Index.

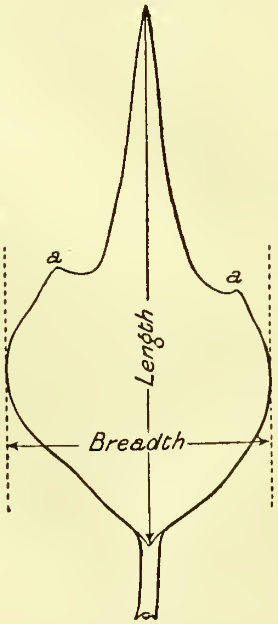
Index of First Leaf.

Index of Second Leaf.

Index.	·295–·395.	·395–·495.	·495–·595.	·595–·695.	·695–·795.	·795–·895.	·895–·995.	·995–1·095.	1·095–1·195.	1·195–1·295.	1·295–1·395.	1·395–1·495.	1·495–1·595.	1·595–1·695.	Totals.
·295– ·395	—	3	5	6	11	10	6	4	2	0	1	—	—	—	48
·395– ·495	3	8	14	31	55	71	36	12	5	3	0	2	—	—	240
·495– ·595	5	14	36	132	294	267	98	40	23	2	1	0	—	—	912
·595– ·695	6	31	132	1048	2186	1556	633	243	92	30	14	4	1	—	5976
·695– ·795	11	55	294	2186	5512	4728	2043	878	314	89	41	16	8	1	16176
·795– ·895	10	71	267	1556	4728	5536	3193	1420	578	195	102	40	12	4	17712
·895– ·995	6	36	98	633	2043	3193	2466	1182	459	190	101	41	12	4	10464
·995–1·095	4	12	40	243	878	1420	1182	628	303	130	68	39	15	6	4968
1·095–1·195	2	5	23	92	314	578	459	303	142	70	33	26	13	4	2064
1·195–1·295	0	3	2	30	89	195	190	130	70	36	17	2	4	0	768
1·295–1·395	1	0	1	14	41	102	101	68	33	17	4	1	1	0	384
1·395–1·495	—	2	0	4	16	40	41	39	26	2	1	12	5	4	192
1·495–1·595	—	—	—	1	8	12	12	15	13	4	1	5	0	1	72
1·595–1·695	—	—	—	—	1	4	4	6	4	0	0	4	1	0	24
Totals . . .	48	240	912	5976	16176	17712	10464	4968	2064	768	384	192	72	24	60000

and 25 leaves from each of 58 plants by me, in the neighbourhood of Great Hampden in Buckinghamshire. Care was taken to avoid the reproductive shoots with their larger uncut leaves. The Hampden plants were, I feel certain, a fairly homogeneous group, the plants were young runners up the trees, only a few feet in length. The Canterbury plants appear to have had a somewhat larger leaf, although in themselves an apparently homogeneous group.





It would undoubtedly have been better to have had the leaves collected by one person from one district; but until the leaves had all been measured and their indices found, I did not know of this difference between the two groups. It was very little noticeable in the table of index distribution, but became manifest to me when finding the homotypic correlation table of the absolute lengths, which I did in two sections, one for either group. The very considerable labour involved in measuring and determining the indices of 2500 leaves precluded a repetition of the work on more homogeneous material. The whole of this labour was undertaken by Dr. Warren, who further formed the table and determined the correlation coefficient.*

I give in the following scheme the frequency distribution of the indices to show its uniformity.

Index.	·295–·395.	·395–·495.	·495–·595.	·595–·695.	·695–·795.	·795–·895.	·895–·995.	·995–1·095.	1·095–1·195.	1·195–1·295.	1·295–1·395.	1·395–1·495.	1·495–1·595.	1·595–1·695.	Total.
Frequency	2	10	38	249	674	738	436	207	86	32	16	8	3	1	2500

---

[*] After measuring the gills of only 1070 mushrooms, I very fully appreciate Dr. Warren's task on the 2500 ivy leaves!

Of course the groupings here are large, but the distribution notwithstanding the mixture of material seems more regular than in some of our other series.

The following scheme gives the constants in the usual manner :

Leaf Index of Wild Ivy.

Number of			Mean index.	S. D.	Coefficient of variation.	S. D. of array.	Percentage variation.	Correlation.
Plants.	Leaves.	Pairs.						
100	2500	60,000	·8473 ± ·0020	·1506 ± ·0014	17·7735	·1449	96·21	·2726 [± ·0125]

It will be observed that the correlation is lower than we might have expected, although the mean of the index correlations for mushroom gills and ivy leaves, *i.e.*, $\frac{1}{2}$ (·5490 + ·2726) = ·4108, is almost identical with the value ·4 given for brothers by the law of ancestral heredity. The low value of the correlation gives of course a high value to the percentage variation, the variation of the individual plant being within 4 per cent. of the racial variation. The variability of the ivy leaf, as judged by this character, has not the intensity which is popularly associated with it ; it is sensibly less than that of the mushroom gill, and about equal to that of the tunics of the onion.

I now pass to the absolute measurements of the leaf.

Dr. WARREN being much pressed with other work, kindly placed all his measurements at my disposal, and I proceeded to draw up the same four tables as in the case of the mushroom gills. The work here was, however, much more laborious as I had to deal with 25 leaves instead of 10 gills, and this involved 300 pairs for each plant, instead of only 45 ! In the course of a fortnight's work I had completed the length-breadth organic correlation, the length-length homotypic correlation, and about a third of the length-breadth homotypic correlation. At this point Dr. LEE took the work off my hands and finished the last table and the breadth-breadth homotypic correlation. So that the results are again the product of co-operation.* I give below the data arranged as in the case of the mushroom gills. The dimensions are given in eighths of inches.

---

* We have now systematised the working of these long tables, involving 10,000 to 60,000 entries providing appropriate checks for accuracy up to each stage of construction. It seems unnecessary to describe these here, but we shall be glad to put our experience at the service of any one working at similar problems. At the same time the collection, measurement and formation of a table for the cross homotypic correlation of two characters in 25 undifferentiated like organs of 100 individuals will cost a *single* worker at least three weeks to a month's fairly continuous labour.

Absolute Dimensions of Wild Ivy Leaves.

Character.	Mean.	S. D.	Coefficient of variation.	S. D. of array.	Percentage variation.	Correlation.
Length. .	10·9504 ± ·0457	3·3885 ± ·0612	30·9442	2·8033	82·73	·5618 [± ·0092]
Breadth .	13·2148 ± ·0323	4·5384 ± ·0433	34·3430	3·8376	84·56	·5332 [± ·0096]

Cross-Correlations.

Length and breadth, organic . .	·8718 ± ·0032
Length and breadth, homotypic .	·5157 [± ·0099]

These results differ very considerably from those for the mushroom gills. The coefficients of variation are even higher than those for the gills, and I think this is a result of a certain amount of heterogeneity, as well as of the leaves on individual runners not being all quite at the same stage of development. In the case of the deciduous trees the leaves were gathered in the fall of the year, and no further development of veins was possible ; but in the ivy leaves, taken indeed at the same time, the equal development of all the leaves taken from the runner could hardly have been reached when they were gathered. On the other hand the correlations are much lower than in the case of the gills,—much more within the range of the results obtained for other characters. I am inclined to think, therefore, that there has been some balancing of opposing factors here, heterogeneity due to locus of collection and to stages of growth being to some extent counteracted by a differentiation due to position of the leaf on the runner.* I have not included these results for lengths and breadths of ivy leaves in my homotypic series because the values, although well within the range of the other determinations, appear to me to be somewhat fictitiously so. The disturbing factors referred to above seem to me to have also given the cross homotypic correlation a higher value than it ought to have. It will be seen that the correlation of length and breadth of pairs of leaves from the same runner is almost as high as that of the breadths of pairs of leaves from the same runner. I should expect a considerably greater inequality. I regret that the great labour of these cross-correlation investigations has hindered their being carried further than these two cases in the vegetable kingdom, but in the zoological data, which we have at present in hand, I hope to get material less open to criticism than in the cases of mushroom gills and ivy leaves. These two series were originally undertaken with the view of ascertaining how far the use of indices would cancel the influence of the factor of growth. As we have

---

* The reader must of course bear in mind that the leaves were taken from runners which had not reached the tops of their walls or trees, *i.e.*, they were not from reproductive shoots.

TABLE XXVIII.—Wild Ivy. Homotypic Correlation of Lengths of Leaves.

Length of First Leaf.

Length of Second Leaf.

Eighths of inches.	2·95–3·95.	3·95–4·95	4·95–5·95.	5·95–6·95.	6·95–7·95.	7·95–8·95.	8·95–9·95.	9·95–10·95.	10·95–11·95.	11·95–12·95.	12·95–13·95.	13·95–14·95.	14·95–15·95.	15·95–16·95.	16·95–17·95.	17·95–18·95.	18·95–19·95.	19·95–20·95.	20·95–21·95.	21·95–22·95.	22·95–23·95.	23·95–24·95.	24·95–25·95.	Totals.
2·95– 3·95	2	13	23	18	8	9	6	11	1	3	0	0	2	—	—	—	—	—	—	—	—	—	—	96
3·95– 4·95	13	112	218	198	112	97	52	26	5	7	0	0	0	—	—	—	—	—	—	—	—	—	—	840
4·95– 5·95	23	218	504	520	330	322	188	100	39	42	10	4	4	—	—	—	—	—	—	—	—	—	—	2304
5·95– 6·95	18	198	520	760	627	510	429	306	152	103	53	33	29	4	1	1	—	—	—	—	—	—	—	3744
6·95– 7·95	8	112	330	627	736	655	689	466	259	273	136	72	42	17	12	1	3	0	1	0	1	—	—	4440
7·95– 8·95	9	97	322	510	655	866	1034	803	558	464	255	184	55	56	21	10	5	0	0	0	0	—	—	5904
8·95– 9·95	6	52	188	429	689	1034	1254	1108	847	722	432	278	110	98	59	33	18	5	3	0	3	—	—	7368
9·95–10·95	11	26	100	306	466	803	1108	1200	1067	934	576	440	200	170	94	40	32	15	7	3	7	3	—	7608
10·95–11·95	1	5	39	152	259	558	847	1067	1040	989	653	514	239	222	105	80	66	13	9	11	14	5	—	6888
11·95–12·95	3	7	42	103	273	464	722	934	989	958	674	546	268	200	138	57	43	12	13	13	13	8	—	6480
12·95–13·95	0	0	10	53	136	255	432	576	653	674	516	436	228	187	136	69	29	21	15	14	11	12	1	4464
13·95–14·95	0	0	4	33	72	184	278	440	514	546	436	336	202	143	104	52	41	28	29	11	12	14	1	3480
14·95–15·95	2	0	4	29	42	55	110	200	239	268	228	202	100	104	75	55	43	16	16	12	9	11	4	1824
15·95–16·95	—	—	—	4	17	56	98	170	222	200	187	143	104	100	55	70	67	10	22	13	11	5	6	1560
16·95–17·95	—	—	—	1	12	21	59	94	105	138	136	104	75	55	38	36	30	22	20	14	10	13	1	984
17·95–18·95	—	—	—	1	1	10	33	40	80	57	69	52	55	70	36	40	41	17	17	13	3	8	5	648
18·95–19·95	—	—	—	—	3	5	18	32	66	43	29	41	43	67	30	41	44	17	21	8	8	8	4	528
19·95–20·95	—	—	—	—	0	0	5	15	13	12	21	28	16	10	22	17	17	12	14	4	1	9	0	216
20·95–21·95	—	—	—	—	1	0	3	7	9	13	15	29	16	22	20	17	21	14	8	7	5	8	1	216
21·95–22·95	—	—	—	—	0	0	0	3	11	13	14	11	12	13	14	13	8	4	7	6	6	8	1	144
22·95–23·95	—	—	—	—	1	0	3	7	14	13	11	12	9	11	10	3	8	1	5	6	2	4	—	120
23·95–24·95	—	—	—	—	—	—	—	3	5	8	12	14	11	5	13	8	8	9	8	8	4	4	—	120
24·95–25·95	—	—	—	—	—	—	—	—	—	—	1	1	4	6	1	5	4	0	1	1	—	—	—	24
Totals . .	96	840	2304	3744	4440	5904	7368	7608	6888	6480	4464	3480	1824	1560	984	648	528	216	216	144	120	120	24	60,000

TABLE XXIX.—Wild Ivy. Homotypic Correlation of Breadths of Leaves.

Breadth of First Leaf.

Breadth of Second Leaf.

Eighths of inches.	2·95–3·95.	3·95–4·95.	4·95–5·95.	5·95–6·95.	6·95–7·95.	7·95–8·95.	8·95–9·95.	9·95–10·95.	10·95–11·95.	11·95–12·95.	12·95–13·95.	13·95–14·95.	14·95–15·95.	15·95–16·95.	16·95–17·95.	17·95–18·95.	18·95–19·95.	19·95–20·95.	20·95–21·95.	21·95–22·95.	22·95–23·95.	23·95–24·95.	24·95–25·95.	25·95–26·95.	26·95–27·95.	27·95–28·95.	28·95–29·95.	29·95–30·95.	30·95–31·95.	31·95–32·95.	32·95–33·95.	33·95–34·95.	34·95–35·95.	35·95–36·95.	Totals.
2·95– 3·95	0	3	8	10	7	7	4	7	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	48
3·95– 4·95	3	42	132	60	71	54	27	17	4	10	5	3	1	0	1	0	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	432
4·95– 5·95	8	132	308	214	252	198	123	89	42	43	23	17	6	1	5	1	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1464
5·95– 6·95	10	60	214	358	346	269	188	215	94	70	41	23	15	10	14	2	15	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1944
6·95– 7·95	7	71	252	346	380	379	354	331	203	148	97	66	68	47	17	16	20	4	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2808
7·95– 8·95	7	54	198	269	379	328	439	428	289	214	173	129	106	73	50	33	19	15	7	4	2	—	—	—	—	—	—	—	—	—	—	—	—	—	3216
8·95– 9·95	4	27	123	188	354	439	572	634	527	389	372	209	241	160	88	73	53	17	12	5	0	0	1	—	—	—	—	—	—	—	—	—	—	—	4488
9·95–10·95	7	17	89	215	331	428	634	766	634	509	501	367	287	218	145	132	79	36	35	15	7	5	7	4	0	2	1	0	1	—	—	—	—	—	5472
10·95–11·95	2	4	42	94	203	289	527	634	680	574	602	436	412	227	182	156	106	55	46	18	12	1	14	5	2	1	3	0	1	—	—	—	—	—	5328
11·95–12·95	—	10	43	70	148	214	389	509	574	616	658	473	403	306	215	198	134	65	60	39	21	14	19	8	9	3	5	0	5	—	—	—	—	—	5208
12·95–13·95	—	5	23	41	97	173	372	501	602	658	748	625	537	393	285	239	179	92	83	50	28	5	34	12	6	7	5	3	2	0	1	1	0	1	5808
13·95–14·95	—	3	17	23	66	129	209	367	436	473	625	574	473	386	323	271	222	134	87	60	34	14	59	27	16	14	8	5	6	0	1	1	0	1	5064
14·95–15·95	—	1	6	15	68	106	241	287	412	403	537	473	396	377	255	222	165	87	77	54	31	13	42	19	19	14	12	5	4	0	1	1	0	1	4344
15·95–16·95	—	0	1	10	47	73	160	218	227	306	393	386	377	272	237	208	155	72	64	63	36	18	39	18	18	14	11	3	6	0	0	0	0	0	3432
16·95–17·95	—	1	5	14	17	50	88	145	182	215	285	323	255	237	194	211	133	71	61	39	33	13	45	23	11	12	7	4	8	0	2	2	0	2	2688
17·95–18·95	—	0	1	2	16	33	73	132	156	198	239	271	222	208	211	174	135	69	78	50	23	12	45	23	6	13	11	5	9	0	3	3	0	3	2424
18·95–19·95	—	2	2	15	20	19	53	79	106	134	179	222	165	155	133	135	100	56	38	39	21	10	27	22	3	15	10	4	6	0	2	2	0	2	1776
19·95–20·95	—	—	—	—	4	15	17	36	55	65	92	134	87	72	71	69	56	20	32	16	13	5	18	13	4	6	5	1	3	0	1	1	0	1	812
20·95–21·95	—	—	—	—	1	7	12	35	46	60	83	87	77	64	61	78	38	32	22	22	14	12	16	10	5	10	7	3	5	0	3	3	0	3	816
21·95–22·95	—	—	—	—	1	4	5	15	18	39	50	60	54	63	39	50	39	16	22	14	8	7	15	7	5	5	8	1	4	0	1	1	0	1	552
22·95–23·95	—	—	—	—	—	2	0	7	12	21	28	34	31	36	33	23	21	13	14	8	8	8	10	10	3	9	4	1	0	0	0	0	0	0	336
23·95–24·95	—	—	—	—	—	—	0	5	1	14	5	14	13	18	13	12	10	5	12	7	8	2	5	5	4	6	4	1	1	0	1	1	0	1	168
24·95–25·95	—	—	—	—	—	—	1	7	14	19	34	59	42	39	45	45	27	18	16	15	10	5	18	8	1	4	2	1	2	0	0	0	0	0	432
25·95–26·95	—	—	—	—	—	—	—	4	5	8	12	27	19	18	23	23	22	13	10	7	10	5	8	4	0	10	5	2	2	0	1	1	0	1	240
26·95–27·95	—	—	—	—	—	—	—	0	2	9	6	16	19	18	11	6	3	4	5	5	3	4	1	0	6	0	2	0	0	0	0	0	0	0	120
27·95–28·95	—	—	—	—	—	—	—	2	1	3	7	14	14	14	12	13	15	6	10	5	9	6	4	10	0	6	5	4	2	0	2	2	0	2	168
28·95–29·95	—	—	—	—	—	—	—	1	3	5	5	8	12	11	7	11	10	5	7	8	4	4	2	5	2	5	0	1	1	0	1	1	0	1	120
29·95–30·95	—	—	—	—	—	—	—	0	0	0	3	5	5	3	4	5	4	1	3	1	1	1	1	2	0	4	1	0	1	0	1	1	0	1	48
30·95–31·95	—	—	—	—	—	—	—	1	1	5	2	6	4	6	8	9	6	3	5	4	0	1	2	2	0	2	1	1	0	0	1	1	0	1	72
31·95–32·95	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32·95–33·95	—	—	—	—	—	—	—	—	—	—	1	1	1	0	2	3	2	1	3	1	0	1	0	1	0	2	1	1	1	0	0	1	0	1	24
33·95–34·95	—	—	—	—	—	—	—	—	—	—	1	1	1	0	2	3	2	1	3	1	0	1	0	1	0	2	1	1	1	0	1	0	0	1	24
34·95–35·95	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	0
35·95–36·95	—	—	—	—	—	—	—	—	—	—	1	1	1	0	2	3	2	1	3	1	0	1	0	1	0	2	1	1	1	0	1	1	—	—	24
Totals . .	48	432	1464	1944	2808	3216	4488	5472	5328	5208	5808	5064	4344	3432	2688	2424	1776	912	816	552	336	168	432	240	120	168	120	48	72	0	24	24	0	24	60,000

TABLE XXX.—Wild Ivy. Organic Correlation of Length and Breadth of Leaves.

Length of Leaf.

Breadth of Leaf.

Eighths of inches.	2·95–3·95.	3·95–4·95.	4·95–5·95.	5·95–6·95.	6·95–7·95.	7·95–8·95.	8·95–9·95.	9·95–10·95.	10·95–11·95.	11·95–12·95.	12·95–13·95.	13·95–14·95.	14·95–15·95.	15·95–16·95.	16·95–17·95.	17·95–18·95.	18·95–19·95.	19·95–20·95.	20·95–21·95.	21·95–22·95.	22·95–23·95.	23·95–24·95.	24·95–25·95.	Totals.
2·95– 3·95	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2
3·95– 4·95	1	5	8	4	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	18
4·95– 5·95	1	16	23	17	3	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	61
5·95– 6·95	0	9	22	35	9	5	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	81
6·95– 7·95	0	2	21	36	33	10	8	5	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	117
7·95– 8·95	1	0	10	34	28	35	19	5	0	2	—	—	—	—	—	—	—	—	—	—	—	—	—	134
8·95– 9·95	0	2	5	19	46	51	34	21	4	3	2	—	—	—	—	—	—	—	—	—	—	—	—	187
9·95–10·95	1	0	3	6	33	60	73	24	20	4	2	0	2	—	—	—	—	—	—	—	—	—	—	228
10·95–11·95	—	—	2	2	17	45	55	61	24	10	6	0	0	—	—	—	—	—	—	—	—	—	—	222
11·95–12·95	—	—	1	2	5	21	51	60	40	24	8	2	1	1	0	0	1	—	—	—	—	—	—	217
12·95–13·95	—	—	—	1	6	14	40	49	60	42	22	6	2	0	0	0	0	—	—	—	—	—	—	242
13·95–14·95	—	—	—	—	3	3	13	35	63	51	21	17	3	2	0	0	0	—	—	—	—	—	—	211
14·95–15·95	—	—	—	—	0	0	7	36	32	45	37	12	7	2	2	0	1	—	—	—	—	—	—	181
15·95–16·95	—	—	—	—	0	0	2	10	23	36	25	30	8	5	3	1	0	—	—	—	—	—	—	143
16·95–17·95	—	—	—	—	1	1	1	8	9	24	21	20	10	9	4	3	1	—	—	—	—	—	—	112
17·95–18·95	—	—	—	—	1	0	0	2	5	12	21	27	12	18	1	1	1	—	—	—	—	—	—	101
18·95–19·95	—	—	—	—	—	—	2	1	3	11	13	11	14	10	5	1	0	3	—	—	—	—	—	74
19·95–20·95	—	—	—	—	—	—	1	0	0	3	7	12	5	2	4	1	1	1	1	—	—	—	—	38
20·95–21·95	—	—	—	—	—	—	—	—	0	2	1	7	6	3	7	5	1	0	1	0	0	1	—	34
21·95–22·95	—	—	—	—	—	—	—	—	2	0	0	1	6	6	2	3	2	1	0	0	0	0	—	23
22·95–23·95	—	—	—	—	—	—	—	—	—	1	0	0	0	2	1	4	3	1	1	0	0	0	1	14
23·95–24·95	—	—	—	—	—	—	—	—	—	—	—	—	—	1	4	0	2	0	0	0	0	0	—	7
24·95–25·95	—	—	—	—	—	—	—	—	—	—	—	—	—	4	5	2	2	0	2	2	1	0	—	18
25·95–26·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	2	4	1	1	0	0	0	—	10
26·95–27·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	1	0	1	2	0	1	0	—	5
27·95–28·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	2	0	1	1	1	1	0	—	7
28·95–29·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	0	0	0	1	2	—	5
29·95–30·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0	0	1	0	0	—	2
30·95–31·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0	0	2	0	0	—	3
31·95–32·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0	—	0
32·95–33·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	0	—	1
33·95–34·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1
34·95–35·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	—	0
35·95–36·95	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1
Totals . .	4	35	96	156	185	246	307	317	287	270	186	145	76	65	41	27	22	9	9	6	5	5	1	2500

TABLE XXXI.—Wild Ivy. Homotypic Cross Correlation of Length and Breadth of Leaves.

Length of First Leaf.

Breadth of Second Leaf.

Eighths of inches.	2·95–3·95.	3·95–4·95.	4·95–5·95.	5·95–6·95.	6·95–7·95.	7·95–8·95.	8·95–9·95.	9·95–10·95.	10·95–11·95.	11·95–12·95.	12·95–13·95.	13·95–14·95.	14·95–15·95.	15·95–16·95.	16·95–17·95.	17·95–18·95.	18·95–19·95.	19·95–20·95.	20·95–21·95.	21·95–22·95.	22·95–23·95.	23·95–24·95.	24·95–25·95.	Totals.
2·95– 3·95	—	9	13	9	7	6	3	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	48
3·95– 4·95	6	44	144	97	49	48	26	9	1	6	0	0	2	—	—	—	—	—	—	—	—	—	—	432
4·95– 5·95	18	126	398	313	188	190	123	54	14	29	5	4	2	—	—	—	—	—	—	—	—	—	—	1464
5·95– 6·95	16	160	314	489	300	263	176	95	45	33	19	16	16	2	—	—	—	—	—	—	—	—	—	1944
6·95– 7·95	12	164	385	486	424	420	358	208	137	94	60	37	14	8	0	1	—	—	—	—	—	—	—	2808
7·95– 8·95	4	114	328	469	505	466	454	290	217	156	100	67	20	15	6	4	0	1	—	—	—	—	—	3216
8·95– 9·95	8	77	211	420	609	655	761	557	411	332	192	136	55	30	22	9	0	3	—	—	—	—	—	4488
9·95–10·95	3	72	200	434	521	776	919	722	544	499	301	230	93	66	49	20	9	6	4	1	1	2	—	5472
10·95–11·95	1	22	78	290	460	657	900	802	668	551	355	231	109	89	54	36	12	9	1	1	2	0	—	5328
11·95–12·95	9	25	84	214	348	541	764	869	669	597	402	274	118	127	67	40	26	13	10	3	4	3	1	5208
12·95–13·95	5	9	41	146	251	520	781	948	813	818	523	409	185	151	99	48	27	13	8	3	7	3	0	5808
13·95–14·95	3	10	33	102	201	384	556	723	758	741	476	402	201	177	106	65	61	16	21	8	13	5	2	5064
14·95–15·95	2	2	25	99	206	313	489	637	630	615	428	347	188	137	96	36	44	14	23	5	10	7	1	4344
15·95–16·95	3	3	18	53	128	242	307	476	467	505	374	313	176	117	94	40	46	26	21	5	6	11	1	3432
16·95–17·95	2	1	12	57	88	125	240	341	359	367	311	241	134	147	84	57	45	16	13	15	11	8	4	2688
17·95–18·95	1	1	11	19	55	109	180	295	375	358	272	219	145	125	72	74	42	12	17	18	10	10	4	2424
18·95–19·95	3	1	7	33	59	86	116	186	247	278	206	171	98	88	51	45	40	13	14	13	9	10	2	1776
19·95–20·95	—	—	2	8	22	41	71	125	145	132	95	74	45	47	25	25	23	7	7	7	6	4	1	912
20·95–21·95	—	—	—	4	7	28	56	98	103	108	93	71	49	52	31	34	24	10	13	16	7	9	3	816
21·95–22·95	—	—	—	1	6	14	24	54	73	75	67	58	44	38	19	25	23	5	7	7	4	6	2	552
22·95–23·95	—	—	—	1	0	4	12	22	36	37	41	35	20	26	22	20	22	14	12	4	2	6	0	336
23·95–24·95	—	—	—	—	0	3	8	11	8	9	15	19	14	12	9	12	11	10	10	6	3	7	1	168
24·95–25·95	—	—	—	—	5	10	29	40	64	39	43	37	38	35	19	23	27	6	7	2	3	3	2	432
25·95–26·95	—	—	—	—	1	0	5	15	34	32	20	22	14	25	13	10	17	7	8	6	5	6	—	240
26·95–27·95	—	—	—	—	—	1	3	13	8	11	17	24	10	10	7	2	4	3	4	0	2	1	—	120
27·95–28·95	—	—	—	—	—	0	2	3	23	20	12	15	9	14	13	8	16	7	7	7	4	8	—	168
28·95–29·95	—	—	—	—	—	1	3	7	12	12	13	12	11	10	7	6	8	5	4	4	2	3	—	120
29·95–30·95	—	—	—	—	—	0	1	0	9	8	5	4	3	4	4	2	1	0	1	2	2	2	—	48
30·95–31·95	—	—	—	—	—	1	1	7	12	9	10	6	5	5	6	3	0	0	1	2	2	2	—	72
31·95–32·95	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	0
32·95–33·95	—	—	—	—	—	—	—	—	2	3	3	2	2	1	3	1	0	0	1	3	1	2	—	24
33·95–34·95	—	—	—	—	—	—	—	—	2	3	3	2	2	1	3	1	0	0	1	3	2	1	—	24
34·95–35·95	—	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	0
35·95–36·95	—	—	—	—	—	—	—	—	2	3	3	2	2	1	3	1	0	0	1	3	2	1	—	24
Totals . .	96	840	2304	3744	4440	5904	7368	7608	6888	6480	4464	3480	1824	1560	984	648	528	216	216	144	120	120	24	60,000

seen the mean of the index-correlations for mushroom gills and ivy leaves = ·4108, a result by no means bad, or widely divergent from the mean result of all our data.

If we test the hypothesis as to the value of the cross-correlation we have the following results arranged as in the case of the mushrooms :

Product of direct homotypic length correlation with organic length-breadth correlation = ·4897

Product of direct homotypic breadth correlation and organic length-breadth correlation = ·4648

Product of mean of direct homotypic length and breadth correlations with the organic length-breadth correlations = ·4773

These numbers have to be compared with the cross length-breadth correlation, *i.e.*, with ·5157.

We see that, as in the case of the mushroom, they are somewhat too small, 48 say as compared with 52. But the difference is considerably less here, and allowing for the action of disturbing factors, I think we may say that the two quantities under investigation are at least of the same order of magnitude. There I think we must leave the hypothesis until my zoological measurements are reduced.

I give the four tables of classified data for the absolute lengths and breadths of the ivy leaves. In the first or organic correlation table the reader will be able, in the distribution at least of breadths, to find something of the irregularity to which I have already referred.

VI. Summary of Results.

(23.) In summing up my results and comparing them with those obtained for fraternal correlation by my co-workers and myself, I felt some difficulty. If I made a selection of what I considered the best homotypic correlation series, and the best fraternal correlation, I might well lay myself open to the charge of selecting statistics with a view to the demonstration of a theoretical law laid down beforehand. Accordingly, I determined to include all my homotypic results, except those for the absolute dimensions of mushroom gills and ivy leaves, where it was pretty evident that we had to a greater or less degree an influence exerted by the growth factor. I thus drew up Table XXXII., containing a summary of all my results. I am quite sure that heterogeneity due to one or another cause, exerts an influence of one kind at the top, and differentiation an influence of another kind at the bottom of this table. The amount of these influences and of other disturbing causes, one cannot measure and allow for. I can only hope that having taken a fairly wide range of races and characters, the influences tending to obscure the homotypic correlation, on the one hand by raising it, and on the other by reducing it, will about balance, and the average of the whole series be a fairly close approximation to the true mean value of homotypic correlation. The result is ·4570. Turning now to our results for fraternal correlation, I put into a second table every single result that we have

TABLE XXXII.—General Results for Homotypic Correlation.

Race.	Character.	Percentage variation.	Correlation.	Remarks.
Mushroom, Hampden	Lengths of gills	50·92	·8607	All these results introduce a correlation due to stages of growth and accordingly are not included in the determination of means.
,, ,,	Breadths of gills	67·67	·7363	
,, ,,	Lengths and breadths of gills	—	·6275	
Wild Ivy, mixed localities	Lengths of leaves	82·73	·5618	
,, ,,	Breadths of leaves	84·56	·5332	
,, ,,	Lengths and breadths of leaves	—	·5157	
(i.) Ceterach, Somersetshire	Lobes on fronds	77·57	·6311	Said to be largely affected by growth and environment.
(ii.) Hartstongue, Somersetshire	Sori on fronds	77·64	·6303	
(iii.) Shirley Poppy, Chelsea	Stigmatic bands	78·86	·6149	Much selected in transit.
(iv.) English Onion, Hampden	Veins in tunics	79·18	·6108	Possibly slightly heterogeneous.
(v.) Holly, Dorsetshire	Prickles on leaves	80·11	·5985	
(vi.) Spanish Chestnut, mixed	Veins in leaves	80·65	·5913	Heterogeneous.
(vii.) Beech, Buckinghamshire	Veins in leaves	82·17	·5699	
(viii.) *Papaver Rhœas*, Hampden	Stigmatic bands	82·71	·5620	
(ix.) Mushroom, Hampden	Gill indices	83·58	·5490	Possibly influenced by individual growth.
(x.) *Papaver Rhœas*, Quantocks	Stigmatic bands	84·59	·5333	
(xi.) Shirley Poppy, Hampden	Stigmatic bands	85·18	·5238	
(xii.) Spanish Chestnut, Buckinghamshire	Veins in leaves	88·51	·4655	
(xiii.) Broom, Yorkshire	Seeds in pods	90·96	·4155	
(xiv.) Ash, Monmouthshire	Leaflets on leaves	91·44	·4047	
(xv.) *Papaver Rhœas*, Lower Chilterns	Stigmatic bands	91·66	·3997	All from one field.
(xvi.) Ash, Dorsetshire	Leaflets on leaves	91·81	·3964	
(xvii.) Ash, Buckinghamshire	Leaflets on leaves	92·73	·3743	
(xviii.) Holly, Somersetshire	Prickles on leaves	93·12	·3548	
(xix.) Wild Ivy, mixed localities	Leaf indices	96·21	·2726	From two localities and possibly slightly influenced by differentiation.
(xx.) *Nigella Hispanica*, Slough	Segments of seed-capsules	98·18	·1899	Differentiation of organs due to position on stem.
(xxi.) *Malva Rotundifolia*, Hampden	Segments of seed-vessels	98·32	·1827	Principally spread from one clump by stolons.
(xxii.) Woodruff, Buckinghamshire	Members of whorls	98·49	·1733	Members really different in morphological origin.
Mean of 22 cases	—	87·44	·4570	—

worked out up to the date of writing this memoir. I felt this was the only safe method, although I might be justified in cutting out several values from the top and from the bottom of this table. I am sure the value for Daphnia is much too high,

TABLE XXXIII.—General Results for Fraternal Correlation.*

Race.	Sex.	Character.	Source of material.	No. of cases.	Reduced by	Correlation.	Remarks.
(i.) Daphnia	♀ & ♀	Length of spine	ERNEST WARREN . . .	330	K. PEARSON	·6934	Probably much too high, owing to heterogeneity introduced by the selection of a few mothers only.
(ii.) Horse .	♀ & ♀	Coat-colour. .	WEATHERBY'S Studbooks .	1000	K. PEARSON, L. BRAMLEY-MOORE, and A. LEE	·6928	Probably much too high, owing to heterogeneity introduced by the use of comparatively few sires.
(iii.) ,,	♂ & ♂	,,	,, ,,	1000		·6232	
(iv.) ,,	♂ & ♀	,,	,, ,,	1000		·5827	
(v.) Man . .	♀ & ♀	Forearm. . .	PEARSON, family data . .	441	A. LEE	·5424	One pair only from each family.
(vi.) Hound (Basset)	mixed	Coat-colour. .	GALTON, from studbook .	—	K. PEARSON and A. LEE	·5257	All members of litter without regard to sex.
(vii.) Man .	♂ & ♂	Eye-colour . .	GALTON, family data . .	1500	K. PEARSON	·5169	All possibly pairs in family taken.
(viii.) ,,	♀ & ♀	Cephalic index	FRANZ BOAS, N. A. Indians	—	C. FAWCETT	·4890	Paternity doubtful.
(ix.) ,,	♂ & ♀	Eye-colour . .	GALTON, family data . .	1500	K. PEARSON	·4615	See remark to (vii.).
(x.) ,,	♀ & ♀	,,	,, ,,	1500	,,	·4463	,, ,,
(xi.) ,,	♀ & ♀	Stature . . .	,, ,,	595	,,	·4436	,, ,,
(xii.) ,,	♂ & ♂	,,	,, ,,	605	,,	·3913	,, ,,
(xiii.) ,,	♂ & ♂	Cephalic index	FRANZ BOAS, N. A. Indians	—	C. FAWCETT	·3790	See remark to (viii.).
(xiv.) ,,	♂ & ♀	Stature . . .	GALTON, family data . .	1181	K. PEARSON	·3754	See remark to (vii.).
(xv.) ,,	♂ & ♀	Cephalic index	FRANZ BOAS, N. A. Indians	—	C. FAWCETT	·3400	See remark to (viii.).
(xvi.) ,,	♀ & ♀	Longevity . .	Quaker records	1050	M. BEETON	·3323	Reduced below true value by non-selective deaths.
(xvii.) ,,	mixed	Temper . . .	GALTON, family data . .	1294	K. PEARSON	·3167	Character very indefinite, and difficult to estimate.
(xviii.) ,,	♂ & ♂	Longevity . .	Peerage records	1000	M. BEETON	·2602	See remark to (xvi.).
(xix.) ,,	♂ & ♀	,,	Quaker records	1947	,,	·1973	,, ,,
—	—	—	—	—	Mean of 19 series . .	·4479	—

---

* [Since the above memoir was written I have deduced another exceedingly interesting value for fraternal correlation from the measurements of Prof. C. B. DAVENPORT on the statoblasts of the Bryozoa (*Pectinella Magnifica*, LEIDY). See 'The American Naturalist,' vol. 34, p. 964, 1900. DAVENPORT gives for the standard-deviation of the number of hooks in all statoblasts the value 1·326, and for the average standard deviation of colonies of statoblasts 1·197. If we represent the former by σ, the latter will be $\sigma\sqrt{1-r^2}$, whence I find for the "fraternal correlation" $r = \cdot 4302$, a result in excellent agreement with the mean values we have just found.—July, 1901.]

and the value for longevity in man much below the true fraternal correlation. In the former case, the mothers were few in number, in the latter the non-selective death-rate reduces very considerably the intensity of collateral inheritance. Both longevity and temper are included in this table for the same reason as *Nigella Hispanica* and woodruff in the first table. I would not run the risk of any apparent selection to reduce either homotypic or fraternal correlation to a closer range of values. The mean of this second table gives the value ·4479 for fraternal correlation. Now I do not propose to lay great stress on what at first sight might look like a most conclusive equality between the mean values of homotypic and fraternal correlations,—within the limits of the probable errors ·4479 and ·4570 are indeed equal. I am quite aware that a few further series added to either the homotypic or fraternal results might modify to some extent this equality. But what I would ask the reader to do is to examine the two tables side by side, to note how the first and last several results of both may fairly be held to be subject to quite definite modifying factors, and then to consider whether there is not very substantial evidence gathered from a fairly wide range of characters in nearly as wide a range of species to show that both homotypic and fraternal correlation fluctuate about a mean value between ·4 and ·5.

I will not venture to assert that either are absolutely constant, but I do realise that it is extremely difficult with the complex system of factors influencing living forms to reduce our conditions to that theoretically perfect state in which we shall measure solely the factor we are investigating. If the intensity of homotyposis were exactly ·45, I should be inclined to distrust any long series of results, one and all of which gave the answer ·45 exactly. There are so many other disturbing factors which only those who have endeavoured to collect series of this kind will fully appreciate. In the first place, the theoretical conception of undifferentiated like organs is very hard to realise practically; position of the organ on the branch or of the branch on a plant, however careful be the collector, may really have introduced differentiation, and so weakened the apparent homotyposis. Secondly, the environmental factor comes into play. It is difficult to obtain a hundred individuals with like environment; soil, position with regard to other growths, sunlight, insect life, &c., may differ in a manner that the collector cannot appreciate. Unlike environment may produce a fictitious likeness in the organs of the same individual when we pass from one individual to a second. The fact that half our series grew in one field, the other half in another, that part came from one side of a road, part from another, may introduce an unperceived heterogeneity which increases the apparent homotyposis. Thirdly, the difficulty of ensuring that all individuals are of the same age or in the same stage of development, is very great. The leaves of an old tree may have a fictitious likeness when compared with those of a young tree; we may gather organs from one individual when it is in a stage of development, which would only be reached in another individual some days or weeks later. These and other factors may perhaps be to some extent eliminated—far better of course by the trained botanist than by the mere mathematician—but I very much doubt the possibility of their complete elimination. If homotyposis had a practically constant value throughout nature, I should only expect this value to be ascertained as a result of the average of many series in which the opposing factors of differentiation, environment, age, stage of growth, &c., may more or less counteract each other. In this manner we may approach to a fair appreciation of the bathmic influence of individuality in the production of undifferentiated like organs. What I should accordingly deduce as legitimate from the above general results would be this, the intensity of pure homotyposis throughout the vegetable kingdom probably lies between ·4 and ·5 ; this is also the mean value found up to the present for fraternal correlation. We may accordingly conclude that heredity is really only a phase of the wider factor of homotyposis. At bottom it is only part of the principle that when an individual puts forth undifferentiated like organs these are not exactly the same, but with a definite intensity of variation have a definite degree of likeness. When we associate heredity with sexual reproduction, we are only considering the result of homotyposis (variation and likeness) between individual spermatozoa and between individual ova. Such homotyposis leads to a likeness of the individuals resulting from the zygotes, which has probably the same mean value as homotypic correlation itself. Thus if the continuity of the germ cells between parent and offspring be realised, we face no longer the problem of heredity, but that of homotyposis, and this again might possibly be reduced to the simplest problem of budding or cell multiplication. Why does the fundamental life-unit on self-multiplication produce homotypes with a definite degree of likeness and a definite degree of variation ? I shall hope for further light on this problem when my data for homotyposis in the animal kingdom, already being collected, are somewhat more complete ; but only the biologist, not the mathematician, can solve it.

(24.) Now let us turn to another point : homotyposis involves, as we have seen, not only a certain degree of likeness in the group of homotypes but a certain degree of variation. Our series does not include any groups of more than twenty-six homotypes, except in the case of the Hampden Shirley poppies. Hence it is not really possible to calculate directly the variability of the individual. But from theoretical considerations, as well as from the support of individual instances, we have seen that the standard deviation of the array is a reasonable measure of the variability of the individual.* Of course more elaborate direct investigations on this point would be of great interest. But I consider that the present series indicate that on an average the variation in the individual is some 87 to 88 per cent. of that of the race. If the reader will examine the column headed *Percentage Variation* in Table XXXII., he will notice that excluding the lengths and breadths of mushroom gills—cases in which the stage of growth is all important—no percentage variation falls below 77. Now

---

* It is quite easy to find isolated individuals with a greater degree of variability than this, and even a greater variability than that of the race, it is the *average* individual variability which is represented by the S. D. of the array.

this seems to me very instructive when we consider the statements made by some writers who theorise about variation rather than actually measure it: *The undifferentiated like organs put forth by the individual have in round numbers* 80 *to* 90 *per cent. of the variation of such organs in the race.* Does not this completely refute the views frequently expressed that variation is the result of sexual reproduction, and that it is quite insignificant in the case of budding? Sexual reproduction may produce a type which is not that of either parent, but this does not *à fortiori* alter the variability of the race with regard to any organ. With continuous variation such type would have previously existed as far as any special organ or character is concerned, and its repetition contributes nothing to the racial variability. We have to meet the fact that the individual produces undifferentiated like organs with a remarkable degree of variety, and if the investigations of the present memoir be valid the practical result of the homotyposis factor having a value of ·4 to ·5 would denote that the individual variability is 91·65 to 86·60 per cent. of the racial.

(25.) Lastly, we may consider a third point of very great interest, which is involved in Table XXXIV. Here we have the coefficients of variation given for some twenty-two series. Now this coefficient seems to me the only satisfactory comparative measure we can find at present of variability. A variation of two in the petals of a buttercup is far more significant than one of two in the florets of an ox-eyed daisy; the measures of the *absolute* variations as given by the standard deviations seem to me of no use when we are comparing different characters in different species. In default of the suggestion of any better standard, all we can do is to get rid of absolute size or number by using the percentage variation of the character as indicated in the coefficients of variation. Now our table gives a fairly continuous series from 7·80 up to 41·96.* The mean value of 19 is in very good accord with the results I have obtained for variation in a much wider series for the vegetable kingdom. Now I think it will be admitted:

(*a*.) that this variation is based on results for a wide number of species;
(*b*.) that this variation covers a considerable variety of characters;
(*c*.) that it is roughly continuous in value between 8 and 40.

But if we examine the fourth column in this table, which gives the place of the corresponding homotypic correlation, we find absolutely no relationship between the intensity of the homotyposis and of the variability. The homotypic order will not compare in any way with the variation order. The mean variability of the first eleven series is 22·95, and of the last eleven series 14·28, very sensible deviations from the mean 18·62 of the whole twenty-two series. But the mean homotypic correlation of the first eleven series is ·4559, and of the last eleven series ·4581, neither of which

---

* If we excluded the hartstongue as largely influenced by environment, we should find broom heading the list, one of my most satisfactory series, both as to similarity of individual environment and as to smoothness of frequency. We should still conclude that variation might reach to nearly 40.

exhibits any sensible difference from ·4570 the mean of all twenty-two series. We are compelled therefore to conclude that there is no relationship between the variability of a species with regard to any organ, and the homotypic correlation of such organs. It would be quite impossible to assert that as the variability of a species decreased, its homotyposis increased. But according to the results reached in this paper heredity is a result of homotyposis. Hence there seems, so far as our researches go, no ground for asserting that increased intensity of heredity means decreased intensity of variation and *vice versâ*. The general impression formed upon my mind by our measurements is the approximate constancy of the *pure* homotypic factor, or at any rate its variation within a fairly small range, and consequently the approximate constancy within a like small range of the hereditary factor for all species and all characters.

TABLE XXXIV.—General Results for Variability.

Race.	Character.	Coefficient of variation.	Position in correlation table.
Hartstongue, Somersetshire	Sori on fronds	41·96	(ii.)
Broom, Yorkshire	Seeds in pods	36·78	(xiii.)
Holly, Dorsetshire.	Prickles on leaves. . . .	26·29	(v.)
Nigella Hispanica, Slough	Segments in capsules. . .	19·39	(xx.)
Holly, Somersetshire.	Prickles on leaves. . . .	18·74	(xviii.)
Ash, Dorsetshire	Leaflets on leaf.	18·65	(xvi.)
Ash, Monmouthshire.	,, ,, ...	18·57	(xiv.)
Mushroom, Buckinghamshire	Gill-indices	18·28	(ix.)
Ceterach, Somersetshire.	Lobes on fronds	18·25	(i.)
Papaver Rhœas, top of Chilterns . . .	Stigmatic bands	17·81	(viii.)
Wild Ivy, mixed	Leaf-indices.	17·77	(xix.)
Papaver Rhœas, Quantocks	Stigmatic bands	17·66	(x.)
English Onion, Hampden	Veins in tunics.	17·43	(iv.)
Spanish Chestnut, mixed	Veins in leaves.	15·72	(vi.)
Ash, Buckinghamshire	,, ,, ...	15·46	(xvii.)
Papaver Rhœas, Lower Chilterns . . .	Stigmatic bands	15·27	(xv.)
Shirley Poppy, Hampden	,, ,, ...	15·17	(xi.)
Spanish Chestnut, Buckinghamshire. .	Veins in leaves.	14·31	(xii.)
Shirley Poppy, Chelsea	Stigmatic bands	13·99	(iii.)
Woodruff, Buckinghamshire	Members in whorls . . .	13·46	(xxii.)
Beech, Buckinghamshire	Veins in leaves.	10·77	(vii.)
Malva Rotundifolia, Hampden	Segments in seed-vessels .	7·80	(xxi.)
Mushroom, Buckinghamshire	Lengths of gills	21·00	—
,, ,, ...	Breadth of gills	23·42	—
Wild Ivy, mixed	Lengths of leaves	30·94	—
,, ,, ...	Breadths of leaves . . .	34·34	—
Mean of first 22 cases . .	—	18·62	—
Mean of all 26 cases . . .	—	19·97	—

On the other hand, great diversity in the variability of different species. I do not wish to be dogmatic on this point, however, but I would definitely assert that so far no relationship between variability and homotyposis is discoverable. Further, it does not seem to me that looking down the order of variability we can assert that our table in any way corresponds to the order of evolutionary descent. It does not seem to me that the species towards the top of it are the more primitive and those towards the bottom the more complex. Indeed two series of the same species may stand at very different points in the table. I am accordingly forced to the conclusion that my table gives no support whatever to the view that variability in the vegetable kingdom is greatest for the more primitive and simpler organisms. Nor, again examining Table XXXII., can we conclude that homotyposis is least intense for such organisms, for the second half of our table cannot be said to contain more of such organisms than the first. Shortly, I see no relation between the position of an organism in the evolutionary descent and the intensity of either its variability or its homotyposis. Hence it follows that, if homotyposis be the source of heredity, the intensity of the latter cannot be related to either variability or position in evolutionary descent.

Mr. ADAM SEDGWICK may be right in his view* that in the early stages of evolution variability is large and inheritance small, and that with development this relation is changed, variability decreasing with an increasing intensity of heredity. My statistics only cover a small portion of one kingdom. But so far as I am aware they represent the only material so far published on which a definite judgment can be founded—where by definite judgment, I mean one based on quantitative facts, and not drawn from a general estimate of nature which makes no statement whatever of the intensity of either variability or heredity for any single character in any one species of living things.

The data I have given seem entirely opposed to Mr. SEDGWICK'S view, but I very fully realise that they are far from exhaustive. They indicate, however, what I hold to be the only valid method of approaching any problem in variation or heredity, *i.e.*, laborious statistical collection of actual facts. My statistics may be insufficient, I would heartily welcome additions to them; or they may be dealt with fallaciously. Here is the material; let others add to and, if need be, modify the conclusions. Only let us follow the method so clearly indicated by DARWIN himself in his 'Cross and Self-fertilisation of Plants;' let us cease to propound hypotheses illustrating them by isolated facts or vague generalities; there are innumerable species at hand in Nature ready for us to measure and count. *Sine numero nihil demonstrandum est*, should now-a-days be the motto of every naturalist who desires to propound novel hypotheses with regard to variation or heredity. It is equally valid for those who merely seek either to establish more firmly or to modify the old.

(26.) I would sum up the results of this memoir in the following conclusions:

---

* Presidential Address to Section D of the British Association, 1899.

(i.) When an individual of any species produces a number of undifferentiated like organs (homotypes) these are not exactly alike when measured with regard to any character. They exhibit a certain degree of variation combined with a certain degree of likeness (homotypic correlation).

(ii.) The homotypic correlation as tested for twenty-two series of homotypes in a variety of species in the vegetable kingdom, has a mean value between ·4 and ·5. Its exact determination is rendered difficult by a number of conflicting factors; but it seems very possible that the actual value of pure homotypic correlation, *i.e.*, the correlation due to the individuality of the plant as apart from the influence of environment, &c., differs but little from the mean value stated above.

(iii.) If a character occurs in a series of homotypes, it will be found to have in the individual a variability 80 to 90 per cent. of the variability of this character in the species at large. It is thus seen that variability is a primary factor of living forms, and is not in any way dependent on sexual reproduction.

(iv.) With a certain hypothesis as to cross homotypic correlation, *i.e.*, the correlation of two different characters in homotypes, namely, that it is the product of the direct homotypic and of the organic correlations, it would follow that the mean fraternal correlation would equal the mean homotypic correlation. The mean of twenty-two homotypic series is found to be sensibly identical with the mean of nineteen fraternal series. A direct investigation of the value of the cross homotypic correlation is only made for two cases, that of the mushroom gill and that of the ivy leaf. The cross homotypic correlation is not found to differ very widely from the product of the direct homotypic and the organic correlations in these cases, but the results are not close enough to be conclusive.

(v.) It would seem, if (iv.) be true, that heredity is only a phase of the principle of homotyposis (ii.), and that the numerical value of its constants may be found from that principle.

(vi.) No relation whatever could be found between the intensity of variability and that of homotyposis in the twenty-two series dealt with. Nor was any relation to be observed between the relative simplicity of the organism and the intensity either of its variability or its homotypic correlation. Regarding heredity as a case of homotyposis, there seems no reason to suppose, as it has been suggested, that variability has decreased and heredity increased in the course of evolution. On the contrary there seems some ground for supposing that homotyposis (and therefore heredity) is a primary factor of living forms, a condition for the evolution of life by natural selection, and not a product of such selection. If the mushroom, the poppy, and the beech show approximately equal homotyposis, it seems well nigh impossible to consider it as a factor of life, increasing with advancing evolution.

APPENDIX.

On the Homotyposis of the Pods of Leguminous Plants.

I have already indicated that I had initially grave doubts as to whether a true measurement of homotyposis could be obtained by counting the ripe seeds in the pods of leguminous plants; but that a passage in a work of DARWIN'S led me to modify my opinion, and when finishing my memoir, I found time to incorporate in it the broom series from Danby Dale which seemed to confirm DARWIN'S views. Broom is cross-fertilised and much visited by insects, and accordingly the statement made on p. 334 of this memoir appears to have full validity. As the matter seemed of some interest, I suggested to Miss C. D. FAWCETT a fuller study of leguminous plants from this standpoint. Her results have reached me in time to be reduced by Dr. LEE and to be incorporated in this Appendix. I have myself worked series for the common bean and the tare vetch, so that if these be included we have now a fairly extensive range of material from which to draw conclusions.

I will first give the data and state the manner in which it has been reduced, and then consider the results which flow from its consideration. The following is a list of the species dealt with; I have remarked on some of the difficulties arising in the course of our work.

(i.) *Cytisus Scoparius.* Broom from Danby Dale (see p. 334).—It was found difficult to count the total number of ovules, and only the ripened seeds were dealt with. The plant is cross-fertilised. The pods were fairly free from insects.

(ii.) *Lotus Corniculatus.* Bird's-foot trefoil from the neighbourhood of Lyme Regis, Dorsetshire.—The aborted seeds could not be counted, and the pods were largely frequented by grubs, so that it was occasionally doubtful how far ripe seeds had been destroyed. The number of pods on each plant in this first series were too few too allow those with grubs in them to be neglected.*

(iii.) *Lotus Corniculatus.*—A second series from the neighbourhood of Lyme Regis. This series consisted of more vigorous plants from a slightly different environment, so that fourteen to twenty pods could be gathered from each, and thus difficulties as to the possible destruction of ripe seeds by grubs avoided. It would appear that the effect of rejecting such pods has raised the mean number of seeds, but at the same time artificially reduced the homotyposis. The plant is cross-fertilised. (DARWIN: 'Cross and Self-Fertilisation of Plants,' p. 361.)

(iv.) *Lathyrus Odoratus.* Sweet Pea.—This was obtained from a row in a nursery garden at Lyme Regis. Both the fertilised and aborted seeds were counted, so that the total number of ovules could be found. The plant is said to be cross-fertilised in

---

* As an illustration of this sort of difficulty I may note that I found something like 90 per cent. of the pods of gorse in the Danby district so full of insect life that nothing could be done in regard to counting their seeds.

countries where it is indigenous, but DARWIN ('Cross and Self-Fertilisation of Plants,' 3rd edition, p. 155, *et seq.*) strongly believes it to be wholly self-fertilised in England.

(v.) *Lathyrus Odoratus.*—A shorter series obtained from a second nursery garden at Lyme Regis.

(vi.) *Lathyrus Sylvestris.* Everlasting pea from the sea coast not very far from Lyme Regis.—Both the ripe and the aborted seeds were counted. The plant appears to be cross-fertilised. Professor F. O. OLIVER tells me that he has watched bees effectively visiting the cultivated everlasting pea.

(vii.) *Vicia Faba.* Common Bean.—I took 100 plants, each having at least ten pods, from a field in Danby Dale, in which a mixture of oats, the common pea, the common bean, and three or four vetches* (described by the owner as tares), was growing as food for cattle. The perfect beans only were counted, but the plants were a poor crop, many pods being stunted in their growth, and it was not always easy to determine whether the seeds had not been fertilised or there had been failure owing to want of nutrition. The common bean is both cross and self-fertilised. I have to thank Miss J. SHARPE and Miss E. CYRIAX for aid in the work on these beans.

(viii.) *Vicia Hirsuta.*—I found twenty-eight plants of the tare vetch growing on a strip of uncut grass along a cut cornfield at Botton, Danby Dale. I could find no more plants in the neighbourhood. There was plenty of the common vetch (*Vicia Sativa*) on the same strip, but the plants had rarely more than two to three pods on them. I was not able to get ten pods from each tare vetch, thirteen plants had fewer, and seven plants more. The series being short, the probable errors are high, but it seemed worth while to include the data. I endeavoured to count both ripe and aborted seeds, but here, as in one or two of the previous cases, I much doubt whether we have succeeded in counting all the ovules. In some of the green pods the number of ovules seemed to be considerably larger than in the fully ripe pods, and I think it possible that the non-fertilised seeds shrink till they are quite unnoticeable even by a cautious observer. Again, it is by no means certain here, as in other cases, that all the seeds reckoned as aborted are really non-fertilised. It is quite possible that in some cases fertilised seeds have dwindled for want of nutriment till they appear aborted.

Vicia Hirsuta is either cross or self-fertilised (DARWIN: 'Cross and Self-Fertilisation of Plants,' p. 367). Whether this vetch, or indeed the common bean, were in our case wholly cross-fertilised or in part self-fertilised, I do not see that we have any means of settling.

Table L. gives the frequency distributions of the various series. Now in these series we must bear in mind that we are only definitely certain of one thing, the number

---

* Neither the pea nor the vetches offered enough pods per plant for an investigation of their homotyposis to be of value.

TABLE L.—Variation of the Seeds in the Pods of Leguminous Plants per mille.

Number of Seeds in Pod.

Species.	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	26.	27.	Totals.
Cytisus Scoparius (ripe) .	—	7	12	24	39	45	61	82	105	113	112	86	83	82	56	40	28	14	4	3	1	—	—	—	—	—	—	—	1000*
Lotus Corniculatus (i.) (ripe)	9	6	42	50	73	85	108	101	88	72	80	56	63	39	23	25	25	15	8	10	5	8	4	4	0	1	—	—	1000
Ditto (ii.) (ripe) . . .	3	0	11	29	37	77	82	96	110	79	76	78	67	43	57	33	25	23	21	10	13	9	5	8	1	4	1	2	1000
Lathyrus Odoratus (i.) (ovules)	—	—	—	—	—	2	14	95	303	325	209	47	5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000
Ditto (ripe)	1	6	11	31	103	208	239	190	134	53	20	3	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000
Ditto (aborted)	74	167	253	220	178	67	31	8	1	0	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000
Lathyrus Odoratus (ii.) (ovules)	—	—	—	—	—	9	36	119	345	294	177	19	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000*
Ditto (ripe)	8	16	45	74	113	194	186	184	120	45	13	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000*
Ditto (aborted)	92	168	225	195	150	74	52	23	15	5	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000*
Lathyrus Sylvestris (ovules)	—	—	—	—	—	—	—	—	1	0	2	6	37	204	476	230	39	5	—	—	—	—	—	—	—	—	—	—	1000
Ditto (ripe)	1	7	20	74	183	227	206	132	79	43	17	9	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000
Ditto (aborted)	—	—	3	10	18	42	84	156	187	203	155	95	33	12	0	1	1	—	—	—	—	—	—	—	—	—	—	—	1000
Vicia Faba (ripe) . . .	35	43	153	408	352	9	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000
Vicia Hirsuta (ovules) .	—	37	169	408	282	91	13	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000*
Ditto (ripe)	2	191	438	308	52	9	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000*
Ditto (aborted)	226	554	167	38	15	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1000*

* These series were reduced to per milles.

TABLE LI.—The following table sums up our results.

Species.	Fertilisation.	Number of			Nature of seeds.	Mean No. of seeds.	S. D. of seeds.	Coefficient of variation.	S. D. of array.	Percentage variation.	Homotypic correlation.
		Plants.	Pods.	Pairs.							
Cytisus Scoparius . .	C.	120	1200	10,800	Ripe	9·6425	3·54655	36·78	3·22595	90·96	·4155 [± ·0161]
Lotus Corniculatus (i.) .	C.	100	1000	9000	Ripe	8·5020	4·4326	52·14	4·3080	97·19	·2354 [± ·0201]
Ditto (ii.)	C.	100	1000	9000	Ripe	9·9783	4·6276	46·38	4·5447	98·21	·1884 [± ·0206]
Lathyrus Odoratus (i.) .	S.	100	1000	9000	Ovules	8·7750	1·1101	12·65	1·0834	97·59	·2182 [± ·0203]
Ditto (i.)	S.	100	1000	9000	Ripe	6·1310	1·6964	27·67	1·6906	99·66	·0830 [± ·0212]
Ditto (i.)	S.	100	1000	9000	Aborted	2·6400	1·5442	58·49	1·4878	96·34	·2679 [± ·0198]
Lathyrus Odoratus (ii.)	S.	80	800	7200	Ovules	8·4925	1·1369	13·39	1·0581	93·07	·3658 [± ·0207]
Ditto (ii.)	S.	80	800	7200	Ripe	5·6725	1·9831	34·96	1·9393	97·79	·2091 [± ·0228]
Ditto (ii.)	S.	80	800	7200	Aborted	2·8212	1·8945	67·15	1·8650	98·44	·1759 [± ·0231]
Lathyrus Sylvestris . .	C.	100	1000	9000	Ovules	14·0130	·9512	6·79	·9374	98·55	·1695 [± ·0207]
Ditto	C.	100	1000	9000	Ripe	5·6080	1·8456	32·91	1·8015	97·59	·2184 [± ·0203]
Ditto	C.	100	1000	9000	Aborted	8·4150	1·9906	23·66	1·9717	99·05	·1376 [± ·0209]
Vicia Faba	C. or S.	100	1000	9000	Ripe	3·0260	1·0136	33·50	·9956	98·22	·1877 [± ·0206]
Vicia Hirsuta . . .	C. or S.	28	255	2328	Ovules	3·2612	1·0071	30·88	·9920	98·50	·1724 [± ·0410]
Ditto	C. or S.	28	255	2328	Ripe	2·2436	·8598	38·32	·8531	99·22	·1243 [± ·0416]
Ditto	C. or S.	28	255	2328	Aborted	1·0636	·8250	77·57	·8157	98·88	·1493 [± ·0413]
Mean	Four series				Ovules	—	—	13·43	—	96·93	·2315
Mean	Eight series				Ripe	—	—	37·83	—	97·355	·2077
Mean	Four series				Aborted	—	—	56·72	—	98·18	·1827
Mean	Sixteen series				All series	—	—	37·08	—	97·45	·2073

of fully ripened seeds—and this only when the pods were not infested with grubs. We did the best we could to count the aborted seeds, but we cannot be certain that all those counted as aborted were non-fertilised, or that we succeeded in counting all those which had shrunk to microscopic proportions. The number of ovules is simply the sum of the ripe and the aborted seeds counted in each pod, and this again must be somewhat doubtful. The reader will bear in mind that we proceeded with care, but that we think it right, if anything, to rather over- than under-emphasise possible sources of error. Yet allowing for such sources of error we cannot, on examining the results given in Table LI., allow that Broom is in the least representative of the degree of homotyposis to be found in the pods of leguminous plants. Whether we consider the ovules, the ripe, or the aborted seeds, our results are sensibly below that for broom, and the mean of the whole sixteen series gives us a value about *one-half* that of the homotypic correlation based upon characters not depending on fertilisation. If we deal with averages, it would certainly seem that in the results flowing from fertilisation, we have reduced the intensity of the individuality to about half its previous value.

To this extent only does the individual constitution appear influential in the number of seeds in the pod, the remainder of the homotypic intensity seems to have disappeared under random influences having nothing to do with the individuality of the plant. This is perhaps what we might expect in the case of *ripe* seeds in cross-fertilised plants, where the fertilisation may depend on the chance or not of insect visitation and the effectiveness or not of the pollen brought on such occasions. We should have to assert that the bird's-foot trefoil and the everlasting pea, whose average for homotypic correlation in the case of ripe seeds is about ·21 to ·22, lose half the intensity of their individuality through the random nature of the chances of the cross-fertilisation. But although this might be fairly satisfactory for these cases, what are we to say for the species which are self-fertilised absolutely or self-fertilised in default of cross-fertilisation ? We might have expected a high degree of homotyposis in the field bean or the tare vetch, where failing cross-fertilisation we are told there will be self-fertilisation. We find on the contrary, however, in these results some of the lowest homotypic correlations of the whole series. The sweet pea also in its two series presents some very remarkable results. If the sweet pea be entirely self-fertilised then we should expect the homotyposis of both ovules and ripe seeds to be fully up to the average. In neither case is such a result reached, although in one series we have a value for the ovules higher than that obtained for anything except the broom. The noteworthy fact, however, is that the results for the two series of sweet peas differ so widely in character ! In the first series it is the abortion which is most individual in character; in the second series it is the ovules. In the first series the ripe seeds have far less individuality than the aborted seeds, in fact, the smallest homotyposis I have yet observed; in the second series the ripe seeds have a value rising to ·2, which is higher than that of the aborted seeds. Ash trees from Monmouthshire, Buckinghamshire, and Dorsetshire, gave results in good agreement; but sweet peas from two different nursery gardens in the same district give strikingly divergent homotypic correlations!

According to DARWIN ('The Effects of Cross and Self-Fertilisation,' 3rd edition, p. 153) the sweet pea "in this country seems invariably to fertilise itself." He bases this statement on (*a*) the difficulty of access to bees and other insects, (*b*) the fact that the varieties are habitually grown by seed growers close together, and yet the colours do not blend, (*c*) the experience that when the varieties are artificially crossed the colours do change. Professor DELPHINO, of Florence, in a letter to DARWIN, writes that it is the fixed opinion of the Italian gardeners that the varieties do cross; but in Italy other insects may of course be available. Now it is most remarkable that in Series I., whether the sweet pea be crossed or self-fertilised, the number of ripe seeds should be hardly individual at all. It would indicate that, with the particular environment of this series, the chance even ot self-fertilisation depends upon extraneous causes. The comparatively high individuality in the tendency to abort may mark some peculiarity of this variety; there may be a distinct inheritance of sterile tendencies to be kept separate from an inheritance of fertility. But it is probably idle to guess at explanations of such discrepancies. I have gone carefully into the differences of environment in the two series which are interesting, although they do not obviously provide any key to the mystery. The first series were thickly planted in a long row, and the seed pods were gathered unripe. It was more difficult to separate each plant and to be quite certain that some of the seeds would not have had to be reckoned as aborted in the fully ripened pods. The second series consisted of withered plants, the pods being almost all completely ripe; the peas were planted in separate groups a few feet apart. There was no difficulty in ascertaining the individuality of the plant nor as to the number of ripe or aborted seeds. About forty hives of bees were kept in the immediate neighbourhood of this second series;* the first series were at the other end of Lyme Regis, and removed from hives. Both series had had their flowers freely cut. I might have been prepared to attribute the low value of the homotyposis in the sweet pea to this latter cause, but then the common bean and the tare vetch had not been subjected to any similar process, and they give on the whole much lower values. The ovules in either self-fertilised or cross-fertilised plants ought indeed to give a fair measure of homotyposis, but, for reasons already stated, it is not certain that our procedure in counting the perfect and the aborted seeds in the ripened pods has led to a just estimate of them. I am inclined on the whole to attempt no explanations for the anomalies observed in these seed investigations. The influences of self- and cross-

---

* I mention this, as I noticed in Yorkshire fairly frequent but apparently ineffectual visits of the hive bees to sweet peas. It is possible that their attempts may assist the self-fertilisation, just as shaking or wind has been observed to do in other species.

fertilisation on homotyposis deserve still fuller and more direct investigation.* I feel we know little as to the influence of external causes even on the completeness or incompleteness of self-fertilisation. So far as homotyposis in the pods of leguminous plants is concerned, I would draw conclusions based solely on averages, and state that :—

(i.) The homotyposis in the case of either ripe or aborted seeds in the pods oi plants seems weakened to one-half the average value it has in the case of characters not depending on fertilisation. Thus DARWIN'S view, that differences in the number of ripe seeds depend upon the constitution of the plants, seems to be only partially true. Extraneous causes about which we are not very clear appear to be generally influential.

(ii.) The extraneous causes which act in a random manner on the homotypic correlation seem to affect both self-fertilised and cross-fertilised plants, and this both with regard to ripe and to aborted seeds.

(iii.) The order of intensity of homotypic correlation is ovules, ripened seeds, and aborted seeds.

The reduction of homotypic correlation in the case of the fertility of pods may be profitably compared with the like reduction which we find in the case of the coefficients of inheritance of fertility and fecundity in man and the horse.†

Other points worth noting, I think, in our results are those of Table LII., which gives the correlation between ovules and ripe and aborted seeds. As we might anticipate, the more ripe the fewer aborted seeds, and *vice versâ*. The correlation is negative, high for the everlasting pea, remarkably low for the tare vetch. It is interesting to see that the correlation between the ovules and either ripe or aborted seeds is not very high. A large number of ovules not necessarily connoting either a very large number of ripe or of aborted seeds. In fact, in the case of the everlasting pea, the number of ovules has very small influence indeed on the number of seeds which ripen. In the tare vetch only is the relationship more marked. Professor F. O. OLIVER tells me that in certain cases evolution appears to be tending in the direction of the pod containing one ripe seed only. It seems, therefore, that such relations as are indicated in Table LII. (and others of a like kind, for further observations ought certainly to be made) may be useful in indicating the degree of fixity between the number of ovules and the number of ripe or aborted seed which are ultimately to be found in the seed vessel.

While the ovules have the least and the aborted seeds the greatest variability, as measured by the coefficient of variation, the ovules have the most and the aborted seeds the least homotypic correlation. This might at first sight appear to be opposed to the view expressed on p. 363, that there is no relationship between the intensities

---

* It seems to me that investigations of this kind ought to be carried out by those who have, what I unfortunately have not, the needful land for experimental investigations.

† 'Phil. Trans.,' A, vol. 192, p. 277 *et seq.*

of variation and of homotyposis. But it is really not so, for if we proceed as on that page to divide up our series into two groups in order of their variability, we find that the first eight give an average coefficient of variation of 22·68, and have an average homotypic correlation of ·1941; the last eight have an average coefficient of variation of 51·47—double and more that of the first eight—while their homotypic correlation is ·2207. Thus, so far from the homotyposis being reduced by greater variability, it actually appears to be slightly increased.

Table LII.—Relationship between the Number of Ovules and those of Ripe and of Aborted Seeds.

Species.	Correlation.		
	Ripe and aborted.	Ripe and ovules.	Aborted and ovules.
Lathyrus Odoratus (i.) . . .	− ·7692	·4581	·2155
" " (ii.) . .	− ·8290	·3628	·2203
Lathyrus Sylvestris	− ·8797	·0992	·3660
Vicia Hirsuta	− ·2859	·6195	·5751
Means . . .	− ·69095	·3849	·3442

Lastly, we may notice (Table LI.) that the variation in the individual amounts to upwards of 97 per cent. of that in the race, or we have another strong fact to aid in demolishing the theory that variability is a result of sexual reproduction. Note in particular the 97 per cent. in the case of the ovules whose number is determined before fertilisation!

Tables XXXV.–XLIX. give the data from which the constants of Tables L.–LII. have been calculated.

TABLE XXXV.—*Lotus Corniculatus.* Dorsetshire. First Series.

Number of Perfect Seeds in First Pod.

Number of Perfect Seeds in Second Pod.

	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	Totals.
0	8	1	8	6	13	9	9	3	2	8	1	2	5	2	2	0	0	2	—	—	—	—	—	—	—	—	81
1	1	0	5	1	4	5	9	5	5	2	6	3	2	2	1	1	1	1	—	—	—	—	—	—	—	—	54
2	8	5	24	25	48	42	60	39	37	21	18	10	16	8	4	4	4	3	1	1	—	—	—	—	—	—	378
3	6	1	25	36	45	53	53	59	33	22	37	24	11	12	5	6	7	4	3	2	1	4	1	—	—	—	450
4	13	4	48	45	58	73	95	58	65	28	42	28	34	18	14	5	7	7	3	3	2	4	2	1	—	—	657
5	9	5	42	53	73	88	87	67	61	53	60	45	24	30	10	14	14	7	6	4	3	8	1	1	—	—	765
6	9	9	60	53	95	87	140	57	109	56	90	43	52	31	23	17	17	7	5	8	2	1	1	0	—	—	972
7	3	5	39	59	58	67	57	136	82	75	63	54	62	39	14	26	24	14	5	3	7	9	1	5	0	2	909
8	2	5	37	33	65	61	109	82	54	57	77	48	54	23	14	15	16	9	3	10	4	5	6	3	0	0	792
9	8	2	21	22	28	53	56	75	57	76	38	43	43	28	16	22	16	13	8	7	5	4	3	4	0	0	648
10	1	6	18	37	42	60	90	63	77	38	96	31	47	25	10	18	23	14	5	8	2	4	5	0	0	0	720
11	2	3	10	24	28	45	43	54	48	43	31	38	38	24	13	12	13	6	5	7	4	4	4	4	0	1	504
12	5	2	16	11	34	24	52	62	54	43	47	38	38	36	23	17	21	10	6	8	3	8	5	3	0	1	567
13	2	2	8	12	18	30	31	39	23	28	25	24	36	20	13	12	8	5	2	5	1	3	1	2	0	1	351
14	2	1	4	5	14	10	23	14	14	16	10	13	23	13	2	14	9	3	3	6	1	4	0	2	0	1	207
15	0	1	4	6	5	14	17	26	15	22	18	12	17	12	14	12	10	5	5	1	1	4	0	1	0	3	225
16	0	1	4	7	7	14	17	24	16	16	23	13	21	8	9	10	10	11	2	2	1	2	3	4	—	—	225
17	2	1	3	4	7	7	7	14	9	13	14	6	10	5	3	5	11	0	4	4	3	1	0	2	—	—	135
18	—	—	1	3	3	6	5	5	3	8	5	5	6	2	3	5	2	4	2	1	1	2	0	0	—	—	72
19	—	—	1	2	3	4	8	3	10	7	8	7	8	5	6	1	2	4	1	6	0	0	0	4	—	—	90
20	—	—	—	1	2	3	2	7	4	5	2	4	3	1	1	1	1	3	1	0	0	3	1	—	—	—	45
21	—	—	—	4	4	8	1	9	5	4	4	4	8	3	4	4	2	1	2	0	3	0	2	—	—	—	72
22	—	—	—	1	2	1	1	1	6	3	5	4	5	1	0	0	3	0	0	0	1	2	—	—	—	—	36
23	—	—	—	—	1	1	0	5	3	4	0	4	3	2	2	1	4	2	0	4	—	—	—	—	—	—	36
24	—	—	—	—	—	—	—	0	0	0	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	—	0
25	—	—	—	—	—	—	—	2	0	0	0	1	1	1	1	3	—	—	—	—	—	—	—	—	—	—	9
Totals	81	54	378	450	657	765	972	909	792	648	720	504	567	351	207	225	225	135	72	90	45	72	36	36	0	9	9000

TABLE XXXVI.—*Lotus Corniculatus.* Dorsetshire. Second Series.

Number of Perfect Seeds in First Pod.

Number of Perfect Seeds in Second Pods.

	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	26.	27.	Totals.
0	6	0	3	0	3	0	0	6	0	3	0	0	3	0	0	0	0	3	—	—	—	—	—	—	—	—	—	—	27
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—	—	—	—	—	—	0
2	3	0	0	0	6	19	6	14	12	5	7	6	6	4	3	3	2	1	1	1	—	—	—	—	—	—	—	—	99
3	0	0	0	24	28	37	35	16	20	22	16	12	7	9	8	3	2	5	6	2	1	2	0	0	0	6	—	—	261
4	3	0	6	28	24	49	36	36	32	20	20	10	12	10	9	10	7	1	4	1	5	2	0	1	1	6	—	—	333
5	0	0	19	37	49	86	69	97	66	40	52	28	38	21	28	21	9	2	13	4	7	1	1	2	1	1	0	1	693
6	0	0	6	35	36	69	80	72	75	69	64	49	41	28	28	25	19	6	13	3	5	6		4	0	1	0	2	738
7	6	0	14	16	36	97	72	84	87	79	69	55	49	34	43	31	20	13	16	9	9	6	1	9	1	5	0	3	864
8	0	0	12	20	32	66	75	87	124	86	64	86	92	48	60	33	26	21	16	8	11	5	5	7	0	3	3	0	990
9	3	0	5	22	20	40	69	79	86	68	53	68	47	26	32	18	13	21	10	11	5	6	0	1	0	5	0	3	711
10	0	0	7	16	20	52	64	69	64	53	46	65	48	32	43	22	21	15	12	8	12	5	2	5	1	1	0	1	684
11	0	0	6	12	10	28	49	55	86	68	65	74	46	36	59	17	15	20	14	6	8	9	5	10	1	0	0	3	702
12	3	0	6	7	12	38	41	49	92	47	48	46	46	33	35	23	16	12	14	5	11	8	5	2	0	3	1	0	603
13	0	0	4	9	10	21	28	34	48	26	32	36	33	18	29	12	13	7	9	4	3	3	4	3	0	1	0	0	387
14	0	0	3	8	9	28	28	43	60	32	43	59	35	29	34	20	7	27	10	5	9	5	3	11	1	0	1	3	513
15	0	0	3	3	10	21	25	31	33	18	22	17	23	12	20	14	6	11	10	3	4	3	2	5	1	0	0	0	297
16	0	0	2	2	7	9	19	20	26	13	21	15	16	13	7	6	4	7	8	7	8	5	5	3	0	0	1	1	225
17	3	0	1	5	1	2	6	13	21	21	15	20	12	7	27	11	7	12	6	4	2	2	3	2	0	1	2	1	207
18	—	—	1	6	4	13	13	16	16	10	12	14	14	9	10	10	8	6	10	3	3	2	3	4	1	1	0	—	189
19	—	—	1	2	1	4	3	9	8	11	8	6	5	4	5	3	7	4	3	0	3	2	1	0	0	0	0	—	90
20	—	—	—	1	5	7	5	9	11	5	12	8	11	3	9	4	8	2	3	3	4	5	0	0	1	0	1	—	117
21	—	—	—	2	2	1	6	6	5	6	5	9	8	3	5	3	5	2	2	2	5	2	1	1	0	0	—	—	81
22	—	—	—	0	0	1	2	1	5	0	2	5	5	4	3	2	5	3	3	1	0	1	0	2	0	0	—	—	45
23	—	—	—	0	1	2	4	9	7	1	5	10	2	3	11	5	3	2	4	0	0	1	2	0	0	0	—	—	72
24	—	—	—	0	1	1	0	1	0	0	1	1	0	0	1	1	0	0	1	0	1	0	0	0	0	0	—	—	9
25	—	—	—	6	6	1	1	5	3	5	1	0	3	1	0	0	0	1	1	0	0	0	0	0	0	2	—	—	36
26	—	—	—	—	—	0	0	0	3	0	0	0	1	0	1	0	1	2	0	0	1	—	—	—	—	—	—	—	9
27	—	—	—	—	—	1	2	3	0	3	1	3	0	0	3	0	1	1	—	—	—	—	—	—	—	—	—	—	18
Totals	27	0	99	261	333	693	738	864	990	711	684	702	603	387	513	297	225	207	189	90	117	81	45	72	9	36	9	18	9000

TABLE XXXVII.—*Lathyrus Odoratus.* First Series.

Number of Ovules in First Pod.

Number of Ovules in Second Pod.	5.	6.	7.	8.	9.	10.	11.	12.	Totals.
5	—	—	—	7	7	3	1	—	18
6	—	22	24	23	35	20	2	—	126
7	—	24	172	282	205	134	36	2	855
8	7	23	282	1168	786	371	81	9	2727
9	7	35	205	786	1114	625	132	21	2925
10	3	20	134	371	625	592	127	9	1881
11	1	2	36	81	132	127	42	2	423
12	—	—	2	9	21	9	2	2	45
Totals	18	126	855	2727	2925	1881	423	45	9000

TABLE XXXVIII.—*Lathyrus Odoratus.* First Series.

Number of Ripe Seeds in First Pod.

Number of Ripe Seeds in Second Pod.	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	Totals.
0	—	—	—	—	—	2	1	3	2	1	—	—	—	9
1	—	—	1	2	14	13	8	10	2	3	1	—	—	54
2	—	1	2	1	9	34	23	20	6	3	0	—	—	99
3	—	2	1	6	45	61	59	41	33	22	8	1	—	279
4	—	14	9	45	120	193	201	164	124	41	14	2	—	927
5	2	13	34	61	193	426	461	351	208	86	29	6	2	1872
6	1	8	23	59	201	461	564	418	251	111	48	3	3	2151
7	3	10	20	41	164	351	418	338	238	87	29	7	4	1710
8	2	2	6	33	124	208	251	238	234	76	29	3	—	1206
9	1	3	3	22	41	86	111	87	76	30	15	2	—	477
10	—	1	0	8	14	29	48	29	29	15	4	3	—	180
11	—	—	—	1	2	6	3	7	3	2	3	—	—	27
12	—	—	—	—	—	2	3	4	—	—	—	—	—	9
Totals	9	54	99	279	927	1872	2151	1710	1206	477	180	27	9	9000

TABLE XXXIX.—*Lathyrus Odoratus.* First Series.

Number of Aborted Seeds in First Pod.

Number of Aborted Seeds in Second Pod.	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	Totals.
0	78	151	167	129	92	35	12	2	—	—	—	666
1	151	286	388	296	225	104	37	14	1	0	1	1503
2	167	388	636	468	375	161	61	17	1	0	3	2277
3	129	296	468	510	355	127	72	18	2	0	3	1980
4	92	225	375	355	362	117	61	10	3	0	2	1602
5	35	104	161	127	117	32	16	10	1	—	—	603
6	12	37	61	72	61	16	18	1	1	—	—	279
7	2	14	17	18	10	10	1	—	—	—	—	72
8	—	1	1	2	3	1	1	—	—	—	—	9
9	—	0	0	0	0	—	—	—	—	—	—	0
10	—	1	3	3	2	—	—	—	—	—	—	9
Totals	666	1503	2277	1980	1602	603	279	72	9	0	9	9000

TABLE XL.—*Lathyrus Odoratus.* Second Series.

Number of Ovules in First Pod.

Number of Ovules in Second Pod.	5.	6.	7.	8.	9.	10.	11.	12.	Totals.
5	2	17	24	14	6	—	—	—	63
6	17	62	95	59	26	2	—	—	261
7	24	95	196	300	150	82	8	—	855
8	14	59	300	1066	689	321	34	1	2484
9	6	26	150	689	754	439	48	3	2115
10	—	2	82	321	439	394	37	3	1278
11	—	—	8	34	48	37	6	2	135
12	—	—	—	1	3	3	2	—	9
Totals	63	261	855	2484	2115	1278	135	9	7200

TABLE XLI.—*Lathyrus Odoratus.* Second Series.

Number of Ripe Seeds in First Pod.

Number of Ripe Seeds in Second Pod.

	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	Totals.
0	2	3	2	10	13	11	3	6	3	1	—	—	54
1	3	2	8	13	19	31	11	12	15	1	1	1	117
2	2	8	22	50	44	68	45	38	29	16	2	0	324
3	10	13	50	62	85	105	67	70	46	17	4	2	531
4	13	19	44	85	108	178	149	129	65	20	5	4	819
5	11	31	68	105	178	300	257	251	146	42	6	0	1395
6	3	11	45	67	149	257	278	284	174	50	20	3	1341
7	6	12	38	70	129	251	284	258	183	71	18	3	1323
8	3	15	29	46	65	146	174	183	118	59	23	3	864
9	1	1	16	17	20	42	50	71	59	36	10	1	324
10	—	1	2	4	5	6	20	18	23	10	0	1	90
11	—	1	0	2	4	0	3	3	3	1	1	—	18
Totals	54	117	324	531	819	1395	1341	1323	864	324	90	18	7200

TABLE XLII.—*Lathyrus Odoratus.* Second Series.

Number of Aborted Seeds in First Pod.

Number of Aborted Seeds in Second Pod.

	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	Totals.
0	112	166	138	116	57	27	25	15	6	3	1	666
1	166	266	280	230	131	51	39	26	13	4	0	1206
2	138	280	396	312	249	115	71	29	18	10	2	1620
3	116	230	312	278	237	106	75	18	23	6	3	1404
4	57	131	249	237	156	113	82	30	19	5	1	1080
5	27	51	115	106	113	48	36	18	12	4	1	531
6	25	39	71	75	82	36	30	12	4	4	0	378
7	15	26	29	18	30	18	12	8	6	0	0	162
8	6	13	18	23	19	12	4	6	6	0	1	108
9	3	4	10	6	5	4	4	0	0	—	—	36
10	1	0	2	3	1	1	0	0	1	—	—	9
Totals	666	1206	1620	1404	1080	531	378	162	108	36	9	7200

TABLE XLIII.—*Lathyrus Sylvestris.* Sea Coast, Dorsetshire.

Number of Ovules in First Pod.

Number of Ovules in Second Pod.

	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	Totals.
8	—	—	—	—	—	2	3	3	0	1	9
9	—	—	—	—	—	0	0	0	0	0	0
10	—	—	—	—	2	9	7	0	0	0	18
11	—	—	—	—	5	10	28	10	1	0	54
12	—	—	2	5	28	90	167	38	2	1	333
13	2	0	9	10	90	486	837	348	51	3	1836
14	3	0	7	28	167	837	2162	920	143	17	4284
15	3	0	0	10	38	348	920	610	121	20	2070
16	0	0	0	1	2	51	143	121	30	3	351
17	1	0	0	0	1	3	17	20	3	0	45
Totals	9	0	18	54	333	1836	4284	2070	351	45	9000

TABLE XLIV.—*Lathyrus Sylvestris.* Sea Coast, Dorsetshire.

Number of Ripe Seeds in First Pod.

Number of Ripe Seeds in Second Pod.

	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	Totals.
0	—	—	—	1	1	1	0	1	1	4	—	—	—	9
1	—	—	4	4	16	12	11	7	7	1	1	—	—	63
2	—	4	16	10	40	33	37	29	5	6	0	—	—	180
3	1	4	10	42	146	170	135	65	53	20	12	6	2	666
4	1	16	40	146	412	378	299	197	86	42	16	13	1	1647
5	1	12	33	170	378	476	426	267	145	77	33	16	9	2043
6	0	11	37	135	299	426	402	251	163	84	31	13	2	1854
7	1	7	29	65	197	267	251	168	107	59	19	15	3	1188
8	1	7	5	53	86	145	163	107	70	39	24	10	1	711
9	4	1	6	20	42	77	84	59	39	38	12	5	—	387
10	—	1	0	12	16	33	31	19	24	12	2	3	—	153
11	—	—	—	6	13	16	13	15	10	5	3	—	—	81
12	—	—	—	2	1	9	2	3	1	—	—	—	—	18
Totals	9	63	180	666	1647	2043	1854	1188	711	387	153	81	18	9000

TABLE XLV.—*Lathyrus Sylvestris.* Sea Coast, Dorsetshire.

Number of Aborted Seeds in First Pod.

Number of Aborted Seeds in Second Pod.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	Totals.
2	—	—	1	6	4	2	1	7	5	1	—	—	—	—	—	27
3	—	—	3	4	11	21	18	13	10	8	2	—	—	—	—	90
4	1	3	2	12	18	23	24	27	29	14	5	3	0	0	1	162
5	6	4	12	28	46	61	73	66	45	25	7	5	0	0	0	378
6	4	11	18	46	84	155	141	123	107	42	15	6	0	4	0	756
7	2	21	23	61	155	250	286	262	164	123	41	14	0	2	0	1404
8	1	18	24	73	141	286	344	354	241	143	41	16	0	0	1	1683
9	7	13	27	66	123	262	354	396	305	195	55	22	0	0	2	1827
10	5	10	29	45	107	164	241	305	254	148	65	19	0	1	2	1395
11	1	8	14	25	42	123	143	195	148	100	39	15	0	1	1	855
12	—	2	5	7	15	41	41	55	65	39	18	7	0	1	1	297
13	—	—	3	5	6	14	16	22	19	15	7	0	0	0	1	108
14	—	—	0	0	0	0	0	0	0	0	0	0	—	—	—	0
15	—	—	0	0	4	2	0	0	1	1	1	0	—	—	—	9
16	—	—	1	0	0	0	1	2	2	1	1	1	—	—	—	9
Totals	27	90	162	378	756	1404	1683	1827	1395	855	297	108	0	9	9	9000

TABLE XLVI.—*Vicia Faba.* Danby Dale, Yorkshire.

Number of Ripe Seeds in First Pod.

Number of Ripe Seeds in Second Pod.	0.	1.	2.	3.	4.	5.	Totals.
0	46	38	35	120	76	—	315
1	38	46	70	142	89	2	387
2	35	70	274	608	386	4	1377
3	120	142	608	1542	1237	23	3672
4	76	89	386	1237	1332	48	3168
5	—	2	4	23	48	4	81
Totals	315	387	1377	3672	3168	81	9000

TABLE XLVII.—*Vicia Hirsuta.* Danby Dale, Yorkshire.

Number of Ovules in First Pod.

Number of Ovules in Second Pod.	1.	2.	3.	4.	5.	6.	Totals.
1	6	22	33	21	4	—	86
2	22	98	169	81	21	2	393
3	33	169	360	292	86	10	950
4	21	81	292	194	57	11	656
5	4	21	86	57	38	6	212
6	—	2	10	11	6	2	31
Totals	86	393	950	656	212	31	2328

TABLE XLVIII.—*Vicia Hirsuta.* Danby Dale, Yorkshire.

Number of Ripe Seeds in First Pod.

Number of Ripe Seeds in Second Pod.	0.	1.	2.	3.	4.	5.	Totals.
0	—	—	3	1	1	—	5
1	—	160	165	89	31	—	445
2	3	165	492	304	45	11	1020
3	1	89	304	278	38	6	716
4	1	31	45	38	10	—	120
5	—	—	11	6	—	—	22
Totals	5	445	1020	716	120	22	2328

TABLE XLIX.—*Vicia Hirsuta.* Danby Dale, Yorkshire.

Number of Aborted Seeds in First Pod.

Number of Aborted Seeds in Second Pod.	0.	1.	2.	3.	4.	Totals.
0	160	263	88	9	6	526
1	263	788	176	52	10	1289
2	88	176	92	16	16	388
3	9	52	16	10	2	89
4	6	10	16	2	2	36
Totals	526	1289	388	89	36	2328

f

SOLD BY HARRISON AND SONS, ST. MARTIN'S LANE.

Demy 8vo, pp. 224. Price 7*s.* 6*d.*

THE RECORD OF THE ROYAL SOCIETY, 1897.

Containing an Account of the Foundation, History, &c., of the Society.

Demy 8vo, pp. 243. Price 5*s.*

THE YEAR-BOOK OF THE ROYAL SOCIETY, 1901.

Containing an Account of the Chief Events and Arrangements of the Year.

PHILOSOPHICAL TRANSACTIONS.

SERIES A. Containing Papers of a Mathematical or Physical Character.
VOL. 195 (1901), Price £1 7*s.*, with 25 Plates.
VOL. 196 (1901), Price £1 11*s.*, with 4 Plates.

SERIES B. Containing Papers of a Biological Character.
VOL. 192 (1900), Price £1 7*s.* 6*d.*, with 13 Plates.
VOL. 193 (1900), Price £1 5*s.*, with 11 Plates.

PUBLISHED BY CLAY AND SONS, AVE MARIA LANE.

CATALOGUE OF SCIENTIFIC PAPERS,

COMPILED BY THE ROYAL SOCIETY.

The volumes on sale as follows:—Vols. 1–6 (1800–63), cloth (vol. 1 in half-morocco) £4 net; half-morocco, £5 5*s.* net. Vols. 7, 8 (1864–73), cloth, £1 11*s.* 6*d.* net; half-morocco, £2 5*s.* net. Single vols., cloth, 20*s.*; half-morocco, 28*s.* net. Vols. 9, 10 (1874–83), cloth, £1 5*s.* net; half-morocco, £1 12*s.* net. Vol. 11 (1874–83), cloth, £1 5*s.* net; half-morocco, £1 12*s.* net.

A reduction of price to Fellows of the Royal Society.

PUBLISHED BY KEGAN PAUL, TRENCH, TRÜBNER AND CO.

A few copies at the reduced price of £1. 1*s.* (published at £3).

Royal 4to, pp. iv.—936, cloth. With 51 Lithographic and Photographic Plates.

A MONOGRAPH OF THE HORNY SPONGES.

BY R. VON LENDENFELD.

A further reduction of price to Fellows of the Royal Society.

PUBLISHED BY KEGAN PAUL, TRENCH, TRÜBNER, AND CO.

In 1 vol., 4to. Pp. 500. With 6 Chromolithographs of the remarkable Sunsets of 1883 and 40 Maps and Diagrams.

THE ERUPTION OF KRAKATOA AND SUBSEQUENT PHENOMENA.

Report of the Krakatoa Committee of the Royal Society.

EDITED BY G. J. SYMONS, F.R.S.

Price 30*s.* To Fellows of the Royal Society, 20*s.*

PUBLISHED BY KEGAN PAUL, TRENCH, TRÜBNER, AND CO.

Royal 4to, pp. xiv.—326, cloth. Price 21*s.*

OBSERVATIONS OF THE INTERNATIONAL POLAR EXPEDITIONS.

1882–1883.

FORT RAE.

With 32 Lithographic Folding Plates.

A reduction of price to Fellows of the Royal Society.

Continental Agents for the publications of the Royal Society, Messrs. FRIEDLÄNDER and SON, Berlin.

JUST PUBLISHED, AND SOLD BY DULAU AND CO., SOHO SQUARE.

SEPARATE PAPERS FROM THE 'PHILOSOPHICAL TRANSACTIONS 1900–01.

Series A.—Mathematical and Physical.

A. 284. On the Conductivity of Gases under the Becquerel Rays. By Hon. R. J. STRUTT, Fellow of Trinity College, Cambridge. Price 1*s*. 6*d*.

A. 285. The Thermo-chemistry of the Alloys of Copper and Zinc. By T. J. BAKER, B.Sc. Price 1*s*.

A. 286. Title, Contents, Index, &c., Vol. 196. Price 6*d*.

A. 287. The Integration of the Equations of Propagation of Electric Waves. By A. E. H. LOVE, F.R.S., Sedleian Professor of Natural Philosophy in the University of Oxford. Price 2*s*.

A. 288. Argon and its Companions. By WILLIAM RAMSAY, F.R.S., and MORRIS W. TRAVERS. Price 3*s*.

A. 289. On the Theory of Consistence of Logical Class-frequencies, and its Geometrical Representation. By G. UDNY YULE, formerly Assistant Professor of Applied Mathematics in University College, London. Price 3*s*.

A. 290. The Action of Magnetised Electrodes upon Electrical Discharge Phenomena in Rarefied Gases. By C. E. S. PHILLIPS. Price 1*s*.

A. 291. Total Eclipse of the Sun, January 22, 1898.—Observations at Viziadrug. By Sir NORMAN LOCKYER, K.C.B., F.R.S., Captain CHISHOLM-BATTEN, R.N., and Professor A. PEDLER, F.R.S. Price 6*s*. 6*d*.

A. 292. On the Heat Dissipated by a Platinum Surface at High Temperatures.—Part IV. Thermal Emissivity in High-pressure Gases. By J. E. PETAVEL, A.M. Inst. C.E., A.M. Inst. E.E., Past 1851 Exhibition Scholar, John Harling Fellow of the Owens College, Manchester. Price 2*s*.

A. 293. A Comparative Crystallographical Study of the Double Selenates of the Series $R_2M(SeO_4)_2,6H_2O$. Part II. Salts in which M is Magnesium. By A. E. TUTTON, B.Sc., F.R.S. Price 2*s*.

Series B.—Biological.

B. 195. The Histology of the Cell Wall with Special Reference to the Mode of Connection of Cells. By WALTER GARDINER, M.A., F.R.S., and ARTHUR W. HILL, B.A. Part I.—The Distribution and Character of "Connecting Threads" in the Tissues of *Pinus sylvestris* and other Allied Species. By ARTHUR W. HILL. Price 5*s*.

B. 196. On the Restoration of Co-ordinated Movements after Nerve-crossing, with Interchange of Function of the Cerebral Cortical Centres. By ROBERT KENNEDY, M.A., D.Sc., M.D., Assistant Surgeon to the Western Infirmary, Glasgow. Price 2*s*.

B. 197. Investigations on the Abnormal Outgrowths or Intumescences on *Hibiscus vitifolius*, Linn. A Study in Experimental Plant Pathology. By ELIZABETH DALE, Pfeiffer Student, Girton College, Cambridge. Price 1*s*.

B. 198. On the "Blaze Currents" of the Frog's Eyeball. By AUGUSTUS D. WALLER, M.D., F.R.S. Price 4*s*.

B. 199. A Chemical Study of the Phosphoric Acid and Potash Contents of the Wheat Soils of Broadbalk Field, Rothamsted. By BERNARD DYER, D.Sc., F.I.C. Price 2*s*. 6*d*.

B. 200. On the Structure and Affinities of Fossil Plants from the Palæozoic Rocks.—IV. The Seed-like Fructification of *Lepidocarpon*, a Genus of Lycopodiaceous Cones from the Carboniferous Formation. By D. H. SCOTT, M.A., Ph.D., F.R.S., Hon. Keeper of the Jodrell Laboratory, Royal Gardens, Kew. (*Shortly.*)

B. 201. The Distribution of Vertebrate Animals in India, Ceylon, and Burma. By W. T. BLANFORD, F.R.S. Price 6*s*.

Cloth cases for binding the 'Philosophical Transactions' can be obtained of HARRISON and SONS, St. Martin's Lane. Price 1*s*. 6*d*.

PHILOSOPHICAL TRANSACTIONS

OF THE

ROYAL SOCIETY OF LONDON.

SERIES A.

Vol. 197, pp. 443–459.

MATHEMATICAL CONTRIBUTIONS TO THE THEORY OF EVOLUTION.—X. SUPPLEMENT TO A MEMOIR ON SKEW VARIATION.

BY

KARL PEARSON, F.R.S.,

UNIVERSITY COLLEGE, LONDON.

LONDON:
PUBLISHED FOR THE ROYAL SOCIETY
BY DULAU AND CO., 37, SOHO SQUARE, W.
CONTINENTAL AGENTS, MESSRS. FRIEDLÄNDER AND SON, BERLIN.
1901.

A. 297.

Price One Shilling.

29.11.01.

XI. *Mathematical Contributions to the Theory of Evolution.—X. Supplement to a Memoir on Skew Variation.**

By KARL PEARSON, *F.R.S., University College, London.*

Received May 22,—Read, June 20, 1901.

(1.) IN a memoir on Skew Variation published in the 'Phil. Trans.,' A, vol. 186, 1895, a series of frequency curves are discussed which are integrals of the differential equation

$$\frac{1}{y}\frac{dy}{dx} = \frac{-x}{c_1 + c_2x + c_3x^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i.)}.$$

(See p. 381 of the memoir.)

The discussion of four main types is given in detail, and a brief reference is made to various sub-types which may occur. The types considered in that memoir covered at the time all the frequency series, and they were fairly numerous, that I had had occasion to deal with. In the course of the last few years, however, I have been somewhat puzzled by frequency distributions for which the criterion $2\beta_2 - 3\beta_1 - 6$ (see p. 378) was positive, and therefore *à priori* a curve of the type

$$y = y_0 \frac{1}{\left\{1 + \left(\frac{x}{a}\right)^2\right\}^m} e^{-\nu \tan^{-1}(x/a)}$$

was to be expected, but which on calculation gave ν imaginary. The frequency distributions in question arose† occasionally in sociological statistics, but also in

---

* 'Phil. Trans.,' A, vol. 186, p. 343.

† Some other frequency distributions, which on first investigation fell under Types V. and VI. of the present paper, were found with improved values for the moments to fall under types already discussed. Mr. W. F. SHEPPARD'S values for the moments ('Lond. Math. Soc.,' vol. 29, p. 369, formula 30) should certainly be used in preference to those given by me ('Phil. Trans.,' A, vol. 186, p. 350) whenever we are calculating the moments of a curve from areas and not from true ordinates. I hope shortly to publish a paper on this point, which is one really of quadrature formulæ. Meanwhile for every true frequency curve *with high contact at both terminals* we ought to use

$$\mu_2 = c^2(\nu_2' - \nu_1'^2 - \tfrac{1}{12})$$
$$\mu_4 = c^4(\nu_4' - 4\nu_1'\nu_3' + 6\nu_1'^2\nu_2' - 3\nu_1'^4 - \tfrac{1}{2}(\nu_2' - \nu_1'^2) + \tfrac{7}{240}),$$

instead of the values given on p. 350, μ_3 remaining unchanged.

biological investigations. It seemed, therefore, desirable to enter a little more fully into the analysis of the cases in which the criterion was positive but ν imaginary, and discover what types of frequency curves had escaped my attention.*

The key to the solution lies in the fact noted on p. 369 of the memoir, namely, that even if the criterion be positive, there will still be a solution akin to Type I. and not to Type IV. if ϵ be negative. No frequency series satisfying these conditions had at that time come under my notice, and later, when collecting data of floral variability, my own remark as to ϵ had slipped from my memory. It is the object of this supplement to obtain an improved criterion of type, to discuss the nature of the curves which fill the gap observed, and to illustrate by one or two examples the fitting of such curves to actual statistics.

(2.) *The Two Criteria.*

Throughout this supplement the notation of the previous memoir will be assumed to be familiar to the reader.

Turning to p. 378 of that memoir, we note that since β_1 and $r - 1$ are necessarily positive, z if positive must be $> r^2$. Hence ν can only become imaginary if z be negative, or

$$\frac{\beta_1 (r-2)^2}{16 (r-1)} > 1.$$

Substitute in this the value of r and it becomes

$$\frac{\beta_1 (\beta_2 + 3)^2}{4 (4\beta_2 - 3\beta_1)(2\beta_2 - 3\beta_1 - 6)} > 1 \quad . \; . \; . \; . \; . \; . \; . \; . \quad \text{(ii.)}.$$

Hence the complete condition that a curve of Type IV. shall give the distribution of frequency is not only

$$\kappa_1 = 2\beta_2 - 3\beta_1 - 6 > 0,$$

but also

$$\kappa_2 = \frac{\beta_1 (\beta_2 + 3)^2}{4 (4\beta_2 - 3\beta_1)(2\beta_2 - 3\beta_1 - 6)} < 1.$$

Turning back to p. 369, we see that ϵ being positive the complete conditions for a curve of Type I. giving the distribution of frequency are

$$\kappa_1 = 2\beta_2 - 3\beta_1 - 6 < 0,$$

---

* I was very loath to adopt Professor EDGEWORTH's method of inventing new frequency curves by putting $x = f(x')$ in a normal frequency distribution, $y = y_0 e^{-cx^2}$. Besides strong theoretical objections to this process, I had found Equation (i.) so sufficient for a great variety of cases that I felt confident it must cover the newly discovered outstanding cases, and this confidence seems justified by the result.

and

$$\kappa_2 = \frac{\beta_1 (\beta_2 + 3)^2}{4 (4\beta_2 - 3\beta_1) (2\beta_2 - 3\beta_1 - 6)} < 0.$$

The latter condition will be always satisfied since β_1 and $4\beta_2 - 3\beta_1$ are positive for any distribution whatever, and $2\beta_2 - 3\beta_1 - 6$ is negative by hypothesis.

Further, in the previous case κ_2 is seen to be essentially positive.

Hence the criteria written down cover all possible cases but those for which

$$\kappa_2 > 1.$$

Sub-cases which arise from transition curves just at the limits will, however, be likely to be of interest, What happens when $\kappa_2 = \infty$ and when $\kappa_2 = 1$? The only possibility for $\kappa_2 = \infty$ is $2\beta_2 - 3\beta_1 - 6$, or $\kappa_1 = 0$. But this curve has been fully treated under Type III. in the memoir.

We shall see later that $\kappa_2 = 1$ leads us up to a novel transition curve of considerable interest.

To ascertain something about the general case in which $\kappa_2 > 1$, let us return to the memoir again and examine the value of ϵ on p. 369. It can only be negative if

$$4 + \tfrac{1}{4}\beta_1 (r + 2)^2/(r + 1) \text{ be } < 0,$$

where r is here

$$= \frac{6 (\beta_2 - \beta_1 - 1)}{3\beta_1 - 2\beta_2 + 6}.$$

Substituting, we find at once

$$\kappa_2 > 1,$$

which in itself involves $\kappa_1 > 0$.

Hence the missing gap corresponds to those cases in which ϵ is negative.

It will be clear that κ_2, although in form giving a more complex criterion than κ_1, is really more effective, as covering all the possible cases. We have then the following scheme :—

Criterion κ_2.	Corresponding frequency curve.
$\kappa_2 = \infty$	Transition curve, Type III. (Memoir, p. 373).
$\kappa_2 > 1$ & $< \infty$	Type VI. (see p. 448 below).
$\kappa_2 = 1$	Transition curve, Type V. (see p. 446 below).
$\kappa_2 > 0$ & < 1.	Type IV. (Memoir, p. 376).
$\kappa_2 = 0, \beta_1 = 0, \beta_2 = 3$.	Normal curve.
$\kappa_2 = 0, \beta_1 = 0, \beta_2$ not $= 3$	Type II. (Memoir, p. 372).
$\kappa_2 < 0$	Type I. (Memoir, p 367).

The object of this supplement is to discuss the calculation of curves of Type V., and to consider those of Type VI. somewhat more at length, they being only briefly referred to on p. 369 of the memoir. It will be seen that Type I. of the memoir has now broken up into two divisions. One portion is the old Type I. passing into the normal curve on one side and Type III. on the other. This Type III. separates the second portion, Type VI., of the old Type I. from the first portion. Type VI. passes from Type III. to the new transition curve Type V., which, like Type III., will be found to have a range limited in one direction only. Finally this new Type V. is the transition to the old Type IV. bounded on the other side by the sub-curve, the old Type II., and beyond that the normal curve. Thus we see that Types I. and IV. do not pass directly into each other through Type III., as might be supposed by the criterion $\kappa_1 >$ or < 0, but that there are a series of intervening curves, two of which, Types V. and VI., require further consideration, if we are to complete the whole round of frequency distributions embraced under the differential equation (i.).

(3.) *On the Frequency Curve of Type V.*

Returning to the fundamental differential equation (i.), let us consider what transformation takes place when the denominator on the right has *equal* roots.* We may then write it in the form

$$\frac{1}{y}\frac{dy}{dx} = \frac{-x}{c_0(c_1+x)^2} = \frac{c_1}{c_0}\frac{1}{(c_1+x)^2} - \frac{1}{c_0(c_1+x)}.$$

Hence
$$\log y = -\frac{c_1}{c_0}\frac{1}{(c_1+x)} - \frac{1}{c_0}\log(c_1+x) + \text{const.}$$

Thus
$$y = y_0\, e^{-\frac{\gamma}{c_1+x}}(c_1+x)^{-p},$$

where, y_0 is a constant, $\gamma = c_1/c_0$ and $p = 1/c_0$. Thus changing the origin we may write the curve :

$$y = y_0\, x^{-p}\, e^{-\gamma/x} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iii.)},$$

where $x_{m0} = \gamma/p$ gives the distance of the mode from the new origin.

To find the moments about this origin, we notice that, p and γ being positive, $y = 0$ when $x = 0$ and when $x = \infty$. Thus as in the curve of Type III. we have a range limited at one end only.

To find the moments we have, if a be the area,

$$a\,\mu'_n = \int_0^\infty y_0\, x^{-p+n}\, e^{-\gamma/x}\, dx \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv.)}.$$

---

* I owe to Miss Agnes Kelly, Ph.D., the suggestion that this type of frequency curve deserved fuller treatment.

Put $\gamma/x = z$, and we find

$$\alpha\,\mu'_n = y_0\,\gamma^{n-p+1}\int_0^\infty e^{-z}\,z^{p-n-2}\,dz$$

$$= y_0\,\gamma^{n-p+1}\,\Gamma\,(p-n-1) \quad \text{. . . . . (v.).}$$

Thus:

$$\alpha = y_0\,\gamma^{1-p}\,\Gamma\,(p-1) \quad \text{. . . . . (vi.).}$$

$$\left.\begin{aligned} \mu_1' &= \gamma/(p-2) \\ \mu_2' &= \gamma^2/(p-2)(p-3) \\ \mu_3' &= \gamma^3/(p-2)(p-3)(p-4) \\ \mu_4' &= \gamma^4/(p-2)(p-3)(p-4)(p-5) \end{aligned}\right\} \quad \text{. . . . . . (vii.).}$$

Transferring to the centroid we find

$$\left.\begin{aligned} \mu_2 &= \frac{\gamma^2}{(p-2)^2(p-3)} \\ \mu_3 &= \frac{4\gamma^3}{(p-2)^3(p-3)(p-4)} \\ \mu_4 &= \frac{3(p+4)\gamma^4}{(p-2)^4(p-3)(p-4)(p-5)} \end{aligned}\right\} \quad \text{. . . . . . . (viii.).}$$

$$\beta_1 = \mu_3^2/\mu_2^3 = \frac{16\,(p-3)}{(p-4)^2} \quad \text{. . . . . (ix.).}$$

$$\beta_2 = \mu_4/\mu_2^2 = \frac{3\,(p+4)\,(p-3)}{(p-5)\,(p-4)} \quad \text{. . . . . (x.).}$$

Eliminating p between β_1 and β_2 we find after some reductions:

$$\beta_1\,(\beta_2+3)^2 = 4\,(2\beta_2 - 3\beta_1 - 6)\,(4\beta_2 - 3\beta_1)$$

or,

$$\kappa_2 = 1 \quad \text{. . . . . (xi.).}$$

Clearly, since this is the condition for Type V., that transition curve is none other than the curve obtained by making the denominator of the right-hand side of the differential equation have equal roots. The curve is clearly of considerable interest, and its existence had not been noticed in the previous series of frequency curves.

The manner of fitting it is now easily described.

Equation (ix.) gives us a quadratic to find $p-4$:

$$(p-4)^2 - \frac{16}{\beta_1}\,(p-4) - \frac{16}{\beta_1} = 0 \quad \text{. . . . . (xii.).}$$

The positive root of this is the required solution.

γ is then found from the first of equations (viii.), or if σ be the standard deviation $= \sqrt{\mu_2}$, then :*

$$\gamma = \sigma (p - 2) \,.\, \sqrt{(p - 3)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xiii.}).$$

Then (vi.) gives :

$$y_0 = \frac{\alpha \gamma^{p-1}}{\Gamma(p - 1)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xiv.}),$$

which determines y_0, the remaining constant for the shape of the curve.

For the position of the curve, we have for the distance from origin to mean, from the first of equations (vii.) :

$$\mu_1' = \gamma/(p - 2) = \sigma \sqrt{(p - 3)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xv.}).$$

If d be the distance from mode to mean we have :

$$d = \mu_1' - \gamma/p$$

$$= \frac{\gamma}{p - 2} - \frac{\gamma}{p} = \frac{2\gamma}{p(p - 2)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xvi.}).$$

Further, the skewness :

$$\text{Sk.} = d/\sigma = \frac{2\sqrt{(p - 3)}}{p} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xvii.}).$$

Thus the solution is completed.

(4.) *On the Frequency Curve of Type VI.*

Type VI., as we have seen, corresponds to the case in which Type I. of the memoir has its ϵ negative. Hence either m_1' or m_2' is negative and the curve transferring the origin takes the form

$$y = y_0 (x - a)^{m_1}/x^{m_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xviii.}).$$

Now it is possible that this curve falls under the limited range type of a frequency from $x = 0$ to $x = a$, but as we see that the criterion places Type VI. between two curves of range limited in one direction only, we expect Type VI. also to be of that character, and a complete solution is obtained by taking the range from $x = a$ to $x = \infty$; this indeed fills up the gap for $\kappa_2 > 1$ and $< \infty$, and (xviii.) with this range is seen to pass into one or other of the two transition curves

$$y = y_0 \, x^p e^{-\gamma x},$$

or

$$y = y_0 \, x^{-p} e^{-\gamma/x},$$

according as we allow the first or second factor to approach a limit.†

---

* The sign of μ_3 will determine the sign of γ, or, what may be taken as the same thing, the direction of the axis of x.

† Write : $y = \text{const.} \times x^{-m_2} (1 - x/a)^{m_1}$ and make $m_1 = \infty$, and $a = \infty$ but m_1/a finite.

Or, $y = \text{const.} \times (1 - a/x)^{m_1}/x^{m_2 - m_1}$, and make $a = 0$, $m_1 = \infty$, and $a \times m_1$ together with $m_2 - m_1$ finite.

Accordingly we shall write Type VI. in the form

$$y = y_0 (x - a)^{q_2}/x^{q_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xix.}),$$

and take the range from a to ∞.

Differentiating to find the position of the mode we have

$$x_{mo} = \frac{a q_1}{q_1 - q_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xx.}).$$

For the moments about the origin:

$$a\mu'_n = \int_a^\infty y_0 \frac{x^n (x-a)^{q_2}}{x^{q_1}}\, dx.$$

Put $a/x = z$, hence

$$\begin{aligned} a\mu'_n &= \int_0^1 \frac{y_0}{a^{q_1 - q_2 - n - 1}} z^{q_1 - q_2 - n - 2} (1 - z)^{q_2}\, dz \\ &= \frac{y_0}{a^{q_1 - q_2 - n - 1}} \mathrm{B}(q_1 - q_2 - n - 1, q_2 + 1) \\ &= \frac{y_0}{a^{q_1 - q_2 - n - 1}} \frac{\Gamma(q_1 - q_2 - n - 1)\, \Gamma(q_2 + 1)}{\Gamma(q_1 - n)}. \end{aligned}$$

Hence we deduce

$$a = \frac{y_0}{a^{q_1 - q_2 - 1}} \frac{\Gamma(q_1 - q_2 - 1)\, \Gamma(q_2 + 1)}{\Gamma(q_1)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xxi.}).$$

$$\left.\begin{aligned} \mu_1' &= \frac{a(q_1 - 1)}{q_1 - q_2 - 2} \\ \mu_2' &= \frac{a^2 (q_1 - 1)(q_1 - 2)}{(q_1 - q_2 - 2)(q_1 - q_2 - 3)} \\ \mu_3' &= \frac{a^3 (q_1 - 1)(q_1 - 2)(q_1 - 3)}{(q_1 - q_2 - 2)(q_1 - q_2 - 3)(q_1 - q_2 - 4)} \\ \mu_4' &= \frac{a^4 (q_1 - 1)(q_1 - 2)(q_1 - 3)(q_1 - 4)}{(q_1 - q_2 - 2)(q_1 - q_2 - 3)(q_1 - q_2 - 4)(q_1 - q_2 - 5)} \end{aligned}\right\} \quad . \quad . \quad . \quad (\text{xxii.}).$$

Now if we compare these results with those on p. 368 of the earlier memoir we see that the one set can be at once deduced from the other by writing $m_1 = -q_1$, $m_2 = q_2$. Thus with this interchange the whole of that solution holds, if we bear in mind that the range is now from $x = a$ to ∞.

We easily find:

$$r = -q_1 + q_2 + 2 \qquad \epsilon = 1 - q_1 + q_2 - q_1 q_2$$

and $1 - q_1$ and $q_2 + 1$ are the roots of

$$z^2 - rz + \epsilon = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxiii.)},$$

where r and ϵ are to be determined as in that memoir, pp. 368-369.

We have :

$$\mu_2 = \frac{a^2 (1 - q_1)(1 + q_2)}{r^2 (r + 1)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxiv.)},$$

where $1 - q_1$ and r are both negative. This gives a.* Thus q_1, q_2 and a are known, and from Equation (xxi.)

$$y_0 = \frac{\alpha\, a^{q_1 - q_2 - 1}\, \Gamma(q_1)}{\Gamma(q_1 - q_2 - 1)\, \Gamma(q_2 + 1)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxv.)}$$

we find the remaining unknown constant for the shape of the curve, y_0. As before, various approximations may be used to the values of the Γ functions when either q_1 or q_2 or both are large.†

We easily obtain for the distance between mode and mean

$$d = \frac{a(q_1 + q_2)}{(q_1 - q_2)(q_1 - q_2 - 2)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxvi.)},$$

and for the skewness :

$$\text{Sk.} = \frac{(q_1 + q_2)\sqrt{(q_1 - q_2 - 3)}}{(q_1 - q_2)\sqrt{\{(q_1 - 1)(q_2 + 1)\}}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxvii.)}.$$

(5.) A special case of some interest arises when the start of the curve is *à priori* known. Suppose its distance from the mean to be c and let (using moments about centroid)

$$\mu_2/c^2 = \gamma_2, \quad \mu_3/(2\mu_2 c) = \gamma_3 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxviii.)}.$$

Then we easily find :

$$\gamma_2 = \frac{1 - q_1}{(1 + q_2)(-q_1 + q_2 + 3)}, \qquad \gamma_3 = \frac{q_1 + q_2}{(1 + q_2)(q_1 - q_2 - 4)}.$$

---

[*] $1 - q_1$ being negative, ϵ is negative, and accordingly by what goes before κ_2 lies between 1 and ∞.

[†] The value of y_0 for curves of Type I., if m_1 be small but m_2 large ('Phil. Trans.,' A, vol. 186, p. 369, foot-note), is

$$y_0 = \frac{\alpha}{b}(m_1 + m_2 + 1)\sqrt{\frac{m_1 + m_2}{m_2}}\, e^{\frac{1}{12}\left(\frac{1}{m_1 + m_2} - \frac{1}{m_2}\right)} \frac{m_1{}^{m_1} e^{-m_1}}{\Gamma(m_1 + 1)},$$

and this can be easily modified to suit (xxv.) above. A very convenient and exact formula for $\Gamma(n + 1)$, if n be large, is that given by FORSYTH ('B.A. Report,' 1883, p. 47) :

$$\Gamma(n + 1) = \sqrt{2\pi}\left(\frac{\sqrt{n^2 + n + \frac{1}{6}}}{e}\right)^{n + \frac{1}{2}},$$

the error being less than $\dfrac{1}{240n^3}$ of the whole.

Whence we deduce to determine q_1 and q_2:

$$q_2 - q_1 = \frac{1 - 3\gamma_2 + 4\gamma_3}{\gamma_2 - \gamma_3} \qquad q_1 + q_2 = \frac{\gamma_3(1 + \gamma_2)(\gamma_2 - 1 - 2\gamma_3)}{(2\gamma_2 - \gamma_3 + \gamma_2\gamma_3)(\gamma_2 - \gamma_3)} \quad . \quad . \quad \text{(xxix.)},$$

and the solution proceeds as before.

(6.) *Illustrations.*—I propose to note a few distributions of frequency in which I have come across Types V. and VI.

(A.) *Statistics of Age of Bride at Marriage, the Bridegroom's Age being between 24 and 25 years.*[*]

The observations given in the table, p. 454, are taken from PEROZZO's memoir: "Nuove Applicazioni del Calcolo delle Probabilità . . . ," 'Reale Accademia dei Lincei,' Anno CCLXXIX., 1881–2, Tavola I.

The total number of recorded marriages is 28,454. The moments were calculated by using SHEPPARD's corrections ('London Math. Soc. Proc.,' vol. 29, p. 369), and are as follows:—

Mean age of bride = 22·1877.

$$\mu_2 = 13{\cdot}3346$$
$$\mu_3 = 67{\cdot}8145$$
$$\mu_4 = 1224{\cdot}6342$$

Whence:

$$\beta_1 = 1{\cdot}9396$$
$$\beta_2 = 6{\cdot}8873$$
$$\kappa_1 = 1{\cdot}9558$$
$$\kappa_2 = 1{\cdot}1094$$

Thus by p. 445 we see that Type VI. is the frequency curve to be selected, but as κ_2 does not differ widely from unity, we shall probably get a good fit from Type V. as well.

Taking Type VI. first, we find:

$$r = -12{\cdot}11075, \qquad \epsilon = -317{\cdot}84987.$$

The quadratic (xxiii.) is accordingly:

$$z^2 + 12{\cdot}11075\,z - 317{\cdot}84987 = 0.$$

---

[*] I selected this example at random, as one out of several leading to the curve types it was my object to illustrate. There is so much tampering with statistics, however, whenever they refer to the ages of women, that it would probably have been better to have used the men.

Thus: $q_1 = 25\cdot 88401, \qquad q_2 = 11\cdot 77326.$

Hence by (xxiv.) $a = 8\cdot 268,405,$

and by (xxv.) $\log y_0 = 24\cdot 275,3032.$

We have accordingly for the equation to the curve:

$$y = 10^{24} \times 1\cdot 884,965 \frac{(x - 8\cdot 268,405)^{11\cdot 77326}}{x^{25\cdot 88401}}.$$

The distance from the origin to the mean is given by the first equation of (xxii.):

$$\mu_1' = 16\cdot 98913,$$

or, the theoretical range starts with brides of $5\cdot 198,570 + 8\cdot 268,405 = 13\cdot 466,975$ years. This is an excellent underlimit to the age of women marrying men of 24 to 25 in a country like Italy. Our first group is at 15·5, and the above start is just two base units before this initial group.

The skewness = ·498,953, and the distance from mode to mean = 1·822,004, or the mode is at 20·3657 years.

Turning now to Type V. we have the following results:—

$$16/\beta_1 = 8\cdot 249,262.$$

Hence Equation (xii.) is:

$$(p - 4)^2 - 8\cdot 249,262\,(p - 4) - 8\cdot 249,262 = 0.$$

Thus the positive value of p is:

$$p = 13\cdot 150,747.$$

Equation (xiii.) gives:

$$\gamma = 129\cdot 73081.$$

Then (xiv.) gives:

$$\log y_0 = 22\cdot 367,6952.$$

Thus the equation to the curve is:

$$y = 10^{22} \times 2\cdot 331,821\, x^{-13\cdot 150,747}\, e^{-129\cdot 73081/x}.$$

To find the position of its start we have by (xv.):

$$\mu_1' = 11\cdot 6343,$$

or, since the mean age of brides is 22·1877, the youngest possible theoretical bride is 10·5534 years. This is probably a worse determination of the underlimit than in the case of Type VI. At the same time I notice that out of about 180,000 women, 101 were married between 14 and 15 years of age, and all the curves begin with a sensibly finite ordinate at 14·5; it is accordingly possible that a somewhat lower age than 13·5 actually occurs in Italy.

Equation (xvi.) gives us for the distance from mode to mean:

$$d = 1·7694,$$

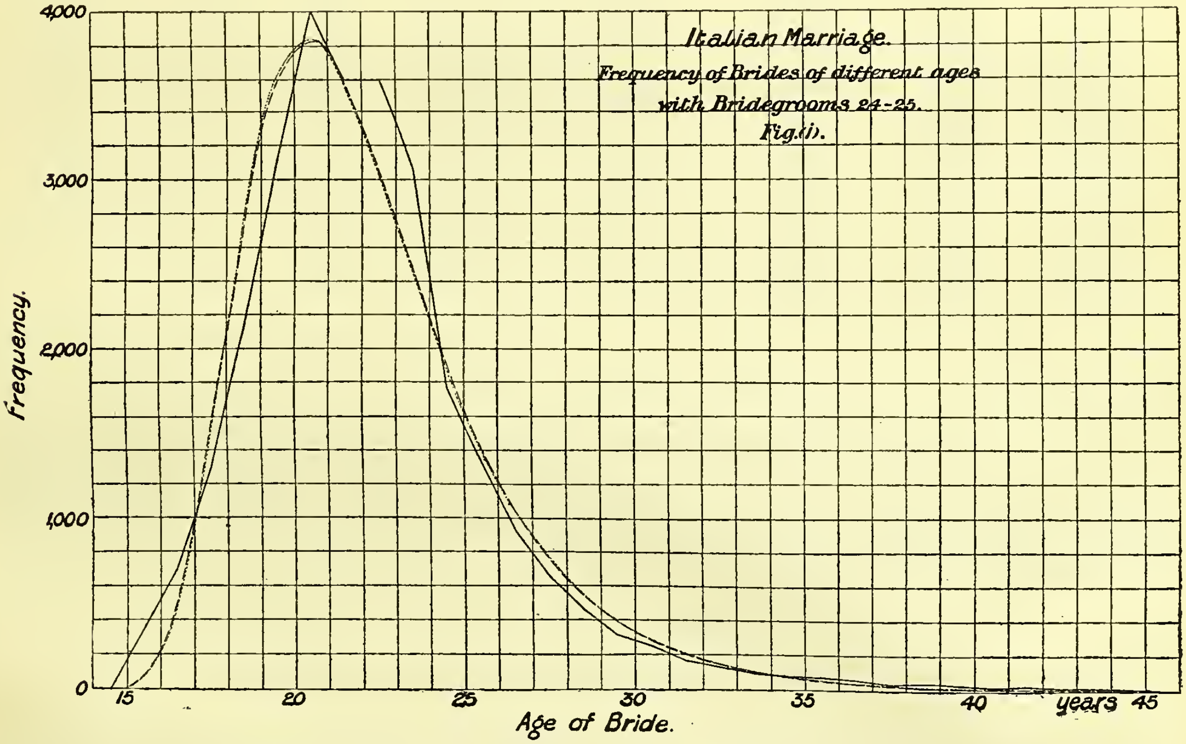
or the modal age at marriage is 20·4183 years. This is only about ·053 of a year or about 19 days different from the modal age as given by Type VI., a most satisfactory agreement.

For the skewness we have from Equation (xvii.):

$$\text{Sk.} = ·4845,$$

or, it differs by less than 3 per cent. from the skewness as given by Type (VI.).

The diagram (fig. i.) shows the two curves, and the table compares the results obtained from either with the observations.*



Type V. - - - - - , Type VI. . . . .

It is clear that for all practical purposes the curve of Type V. is as good as that of Type VI. Indeed, there is practically no difference between them except for the

---

* The observation data are really areas, while to save lengthy calculations we have compared both in diagram and table the ordinates of the theoretical curves. This is in general legitimate, if, as in this case, the number of groups is very large.

ages 15 to 17. The fit is, however, not a very good one, and although it is indefinitely better than a normal curve, and we see why in the absence of these types the statistics could not be fitted with any of the first series of skew curves, yet we are compelled to consider that there are causes other than chance at work very definitely affecting the frequency of the *recorded* ages. Thus the bridegrooms being 24 to 25, the desire of the bride to be recorded as younger than her husband probably fully accounts for the bulk of the preponderance of observation over theory in the frequency of the brides of 22 to 24. The defect of brides between 17 and 20 may be again due to the tendency to state the age as over 21, and so free the woman from the need for parental sanction.* These causes, giving a false displacement of age frequency, are probably in themselves sufficient to account for the theoretical defect in brides of 15 to 17.

TABLE of Observed and Calculated Frequencies.

Age.	Observed frequency.	Calculated frequency.		Age.	Observed frequency.	Calculated frequency.	
		Type V.	Type VI.			Type V.	Type VI.
15–16	367	70	49	30–31	256	281	282
16–17	717	514	489	31–32	164	201	198
17–18	1294	1538	1560	32–33	134	148	146
18–19	2121	2751	2800	33–34	94	104	105
19–20	3156	3591	3622	34–35	77	75	76
20–21	4009	3830	3831	35–36	68	55	55
21–22	3593	3577	3560	36–37	59	40	40
22–23	3604	3055	3034	37–38	33	29	29
23–24	3060	2456	2439	38–39	40	21	22
24–25	1774	1894	1884	39–40	27	16	16
25–26	1353	1419	1415	40–41	18	12	12
26–27	936	1044	1043	41–42	21	9	9
27–28	663	758	760	42–43	11	7	7
28–29	468	546	549	43–44	14	5	5
29–30	319	392	395	44–45	4	4	4

(7.) (B.) *On the Variation in the Number of Lips of the Medusa* P. Pentata.

My data are the following, taken from a paper by ALFRED GOLDSBOROUGH MAYER: "The Variations of a Newly Arisen Species of Medusa," 'Science Bulletin of the Museum of the Brooklyn Institute,' vol. 1, p. 1, 1901.

---

[*] I have found in England the statement of the bride's age in the marriage licence is for the same reason occasionally not in accordance with the year of birth as shown by the parish register.

	Frequency.	No. of lips
	2	1
	5	2
	18	3
	123	4
	798	5
	49	6
	1	7
Total	996	

Mr. MAYER (p. 12) notes the failure of my curve of Type IV. I find for the constants:

$$\text{Mean} = 4{\cdot}8685 \text{ lips.}$$

$$\mu_2 = {\cdot}309{,}006, \quad \sigma = {\cdot}55588$$

$$\mu_3 = -\ {\cdot}350{,}697$$

$$\mu_4 = 1{\cdot}181{,}718$$

$$\beta_1 = 4{\cdot}16834$$

$$\beta_2 = 12{\cdot}37598$$

$$\kappa_1 = 6{\cdot}24694$$

$$\kappa_2 = 1{\cdot}06594$$

Since κ_2 is so nearly unity we may use Type V.
Hence I find:

$$p = 8{\cdot}66184 \qquad \gamma = -\ 8{\cdot}811634$$

(γ must be negative since μ_3 is negative)

$$\mu_1' = 1{\cdot}32270.$$

Thus the curve starts at 6·19118 lips, or the one medusa with seven lips is theoretically excluded. Here I have worked with the uncorrected moments because the lips are discontinuous variants. Working with SHEPPARD'S corrective terms the limit is about six lips, and with the corrective terms suggested in my memoir on skew variation the limit is 7·65. Further we have:

$$\log y_0 = 6{\cdot}829{,}3633,$$

$$\text{distance from mean to mode} = {\cdot}30541,$$

$$\text{Sk.} = {\cdot}54941.$$

The mode is thus at 5·17389, in good agreement with observation.

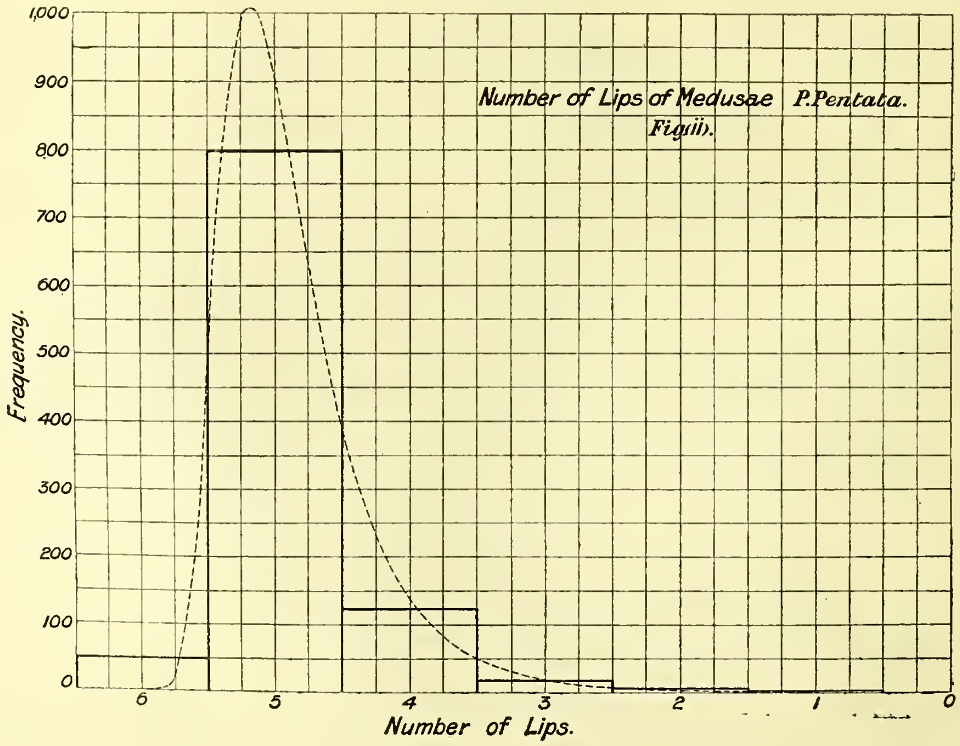
The equation to the curve is, taking x *positive* from 6·19118 lips towards lesser values:

$$\log y = 6{\cdot}829{,}3633 - 8{\cdot}66184 \log x - \frac{3{\cdot}826{,}8435}{x}.$$

This curve was drawn on a large scale and its areas read off with an integrator. The following theoretical frequencies were obtained :

No. of lips.	Observation.	Calculation.
6 and over	50	47
5	798	762
4	123	160·5
3	18	20
2	5	5
1	2	1·5

There would not be any serious divergence here, were it not for the group with four lips, which observation shows to be much under-represented. But it must be remembered that we have only seven groups, and that such a number is very insufficient for a good determination of the moments of a curve. Further, the variation is not really continuous, as indicated by the curve, but *discrete*. We have at present no clear statement as to how the moments of a discrete system of variation should be modified or corrected so as to give the best results for the moments of the continuous curve which is to theoretically represent the series. I am doubtful whether SHEPPARD'S corrections—the best for continuous variation—are equally appropriate in this case. Above I have used merely the rough moments, but I



have found by considerable experience that in the case of discrete variables, to treat the system as a polygon and correct, as in my memoir on Skew Variation ('Phil. Trans.,' A, vol. 186, p. 350), appears to give the best results when the areas are compared with the discrete groups. The point wants further investigation; when we have a large number of groups it is of little importance, but it makes a considerable difference in these excessively skew distributions of discrete variables when the number of groups are small.*

Above all, the diagram (fig. ii.) shows how all important it is to compare *areas* and not merely the *ordinates* of the frequency curve with the blocks representing the discrete frequencies in such a case as this. The wide-spread custom among foreign investigators of comparing merely the ordinates of the theoretical frequency curve with the observed frequencies leads in such cases to most fallacious results.

(8.) (C.) *On the Distribution of Incidence of Scarlet Fever Cases with Age.*

It seems desirable to give an illustration of the method of dealing with a distribution which falls under the class dealt with in Section (5) of this paper. Dr. MACDONELL, in dealing with the intensity of incidence of different diseases at various ages, has come across in scarlet fever a good illustration of curves of the types now under consideration. The whole of the arithmetical work on the present example is due to him, and I have to thank him very heartily for allowing me to use it here.

The statistics are taken from the 'Report of the Metropolitan Asylums Board' (Statistical Part, 1899). They involve 39,253 *male* cases, distributed as follows :—

Year of life.	Frequency.	Year of life.	Frequency.
Under 1	443	20–25	926
1–2	1456	25–30	420
2–3	2631	30–35	215
3–4	3599	35–40	91
4–5	3862	40–45	45
5–10	15791	45–50	26
10–15	7359	50–55	17
15–20	2366	55–60	5
		60–65	1

The data being grouped partly in one and partly in five-year periods the moments had to be calculated with caution, separating the material into two pieces. Taking five years as the unit, Dr. MACDONELL found for the uncorrected moments :

---

* *E.g.*, petals of buttercups, teeth on the carapace of prawns, lips of medusæ, as compared with veins on chestnut leaves, florets on ox-eyed daisy, &c.

Mean age of incidence, 8·60975 years.

$$\mu_2 = 1·369,345$$
$$\mu_3 = 3·233,194$$
$$\mu_4 = 19·143,575.$$

The moments were not modified by SHEPPARD'S corrections, for these suppose contact of a high order at both terminals of the curve, and it was quite apparent that the curve must rise at a finite angle on the birth side. The following additional constants were then determined :—

$$\beta_1 = 4·071,222, \qquad \beta_2 = 10·209,333,$$
$$\kappa_1 = 2·205,000, \qquad \kappa_2 = 2·813,783.$$

Thus κ_2 is >1 and $<\infty$ and the distribution is of Type VI. Now let us suppose the incidence of scarlet fever to start with birth, although there might, as in the case of enteric fever, be really some antenatal cases.*

Turning to Section (5) we have :

$$c = \text{distance from birth to mean} = 8·60975 \text{ years} = 1·72195 \text{ units.}$$

Hence we deduce

$$\gamma_2 = ·461,819, \qquad \gamma_3 = ·685,596.$$

And so from (xxix.)

$$q_2 - q_1 = -10·532,485, \qquad q_2 + q_1 = 15·417,281 \,;$$

or, $$q_1 = 12·974,883, \qquad q_2 = 2·442,398.$$

Then from

$$c = \mu_1' - a = a\,(q_2+1)/(q_1 - q_2 - 2)$$

we find

$$a = 4·268,104,$$

and, finally, after determining y_0 from (xxi.),

$$\log y_0 = 13·652,5078.$$

Thus the values of the frequency are given by

$$\log y = 13·652,5078 + 2·442,398 \log (x - 4·268,104) - 12·974,883 \log x.$$

The origin of the curve is thus 4·268,104 before birth. The mode is given by

$$x_{\text{mo}} = aq_1/(q_1 - q_2) = 5·257,842.$$

Thus : $x_{\text{mo}} - a = ·989,738 = 4·94869$ yrs.

---

* See 'Phil. Trans.,' A, vol. 186, p. 390. The remarkably sharp rise of the scarlet-fever distribution as compared with the enteric is, however, much against this.

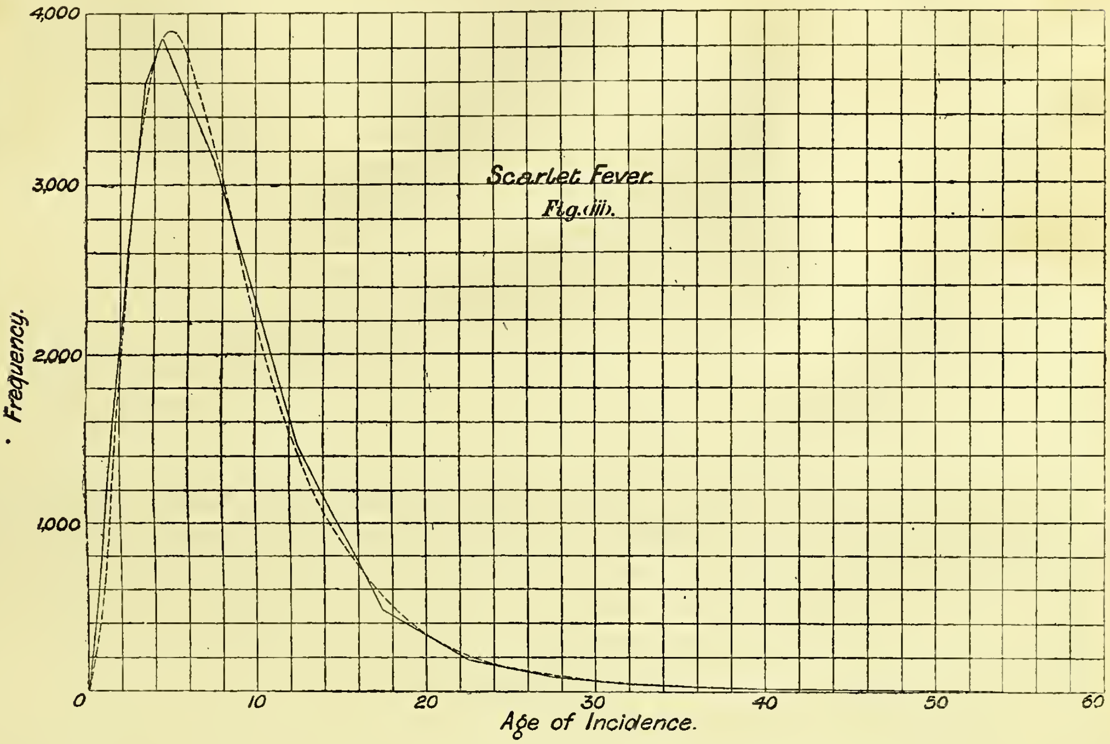
This gives for y_{mode} the value 3892.

Distance between mode and mean = 3·66106 yrs.

Whence we find for skewness the value

$$\text{Sk.} = \cdot 5347.$$

The diagram (fig. iii.) shows that the fit may be considered a good one.



(9.) The conclusions of this paper are, I think, of some interest from the general standpoint of scientific investigation. A certain number of frequency distributions had been found, not only by my co-workers and myself here, but by biologists in America, not to fit into the general system of skew distributions dealt with by me in my original memoir. The first conclusion was that however wide-reaching that system appeared to be, it was a failure for a few remarkably skew distributions. But on more careful investigation of the differential equation it appeared that two types of solution had been left out of consideration, and that these were precisely those needed in the recorded cases of failure.

I owe some apology to authors like Professor DAVENPORT and Dr. DUNCKER, who have recently issued text-books on the application of statistical methods to biological variation, because although we have known and used these curves for some years past, no account has hitherto been published of them, and, consequently, biological investigators* using their *résumés* of my methods have been, and I fear still may be, occasionally puzzled.

---

* *E.g.*, Mr. A. G. MAYER in the paper on Medusæ referred to above.

SOLD BY HARRISON AND SONS, ST. MARTIN'S LANE.

Demy 8vo, pp. 224. Price 7*s.* 6*d.*

THE RECORD OF THE ROYAL SOCIETY, 1897.

Containing an Account of the Foundation, History, &c., of the Society.

Demy 8vo, pp. 243. Price 5*s.*

THE YEAR-BOOK OF THE ROYAL SOCIETY, 1901.

Containing an Account of the Chief Events and Arrangements of the Year.

PHILOSOPHICAL TRANSACTIONS.

SERIES A. Containing Papers of a Mathematical or Physical Character.
VOL. 195 (1901), Price £1 7*s.*, with 25 Plates.
VOL. 196 (1901), Price £1 11*s.*, with 4 Plates.

SERIES B. Containing Papers of a Biological Character.
VOL. 192 (1900), Price £1 7*s.* 6*d.*, with 13 Plates.
VOL. 193 (1900), Price £1 5*s.*, with 11 Plates.

PUBLISHED BY CLAY AND SONS, AVE MARIA LANE.

CATALOGUE OF SCIENTIFIC PAPERS,

COMPILED BY THE ROYAL SOCIETY.

The volumes on sale as follows:—Vols. 1–6 (1800–63), cloth (vol. 1 in half-morocco) £4 net; half-morocco, £5 5*s.* net. Vols. 7, 8 (1864–73), cloth, £1 11*s.* 6*d.* net; half-morocco, £2 5*s.* net. Single vols., cloth, 20*s.*; half-morocco, 28*s.* net. Vols. 9, 10 (1874–83), cloth, £1 5*s.* net; half-morocco, £1 12*s.* net. Vol. 11 (1874–83), cloth, £1 5*s.* net; half-morocco, £1 12*s.* net.

A reduction of price to Fellows of the Royal Society.

PUBLISHED BY KEGAN PAUL, TRENCH, TRÜBNER AND CO.

A few copies at the reduced price of £1. 1*s.* (published at £3).

Royal 4to, pp. iv.—936, cloth. With 51 Lithographic and Photographic Plates.

A MONOGRAPH OF THE HORNY SPONGES.

BY R. VON LENDENFELD.

A further reduction of price to Fellows of the Royal Society.

PUBLISHED BY KEGAN PAUL, TRENCH, TRÜBNER, AND CO.

In 1 vol., 4to. Pp. 500. With 6 Chromolithographs of the remarkable Sunsets of 1883 and 40 Maps and Diagrams.

THE ERUPTION OF KRAKATOA AND SUBSEQUENT PHENOMENA.

Report of the Krakatoa Committee of the Royal Society.

EDITED BY G. J. SYMONS, F.R.S.

Price 30*s.* To Fellows of the Royal Society, 20*s.*

PUBLISHED BY KEGAN PAUL, TRENCH, TRÜBNER, AND CO.

Royal 4to, pp. xiv.—326, cloth. Price 21*s.*

OBSERVATIONS OF THE INTERNATIONAL POLAR EXPEDITIONS.

1882–1883.

FORT RAE.

With 32 Lithographic Folding Plates.

A reduction of price to Fellows of the Royal Society.

Continental Agents for the publications of the Royal Society, Messrs. FRIEDLÄNDER and SON, Berlin.

JUST PUBLISHED, AND SOLD BY DULAU AND CO., SOHO SQUARE.

SEPARATE PAPERS FROM THE 'PHILOSOPHICAL TRANSACTIONS,' 1901.

Series A.—Mathematical and Physical.

A. 289. On the Theory of Consistence of Logical Class-frequencies, and its Geometrical Representation. By G. UDNY YULE, formerly Assistant Professor of Applied Mathematics in University College, London. Price 3*s*.

A. 290. The Action of Magnetised Electrodes upon Electrical Discharge Phenomena in Rarefied Gases. By C. E. S. PHILLIPS. Price 1*s*.

A. 291. Total Eclipse of the Sun, January 22, 1898.—Observations at Viziadrug. By Sir NORMAN LOCKYER, K.C.B., F.R.S., Captain CHISHOLM-BATTEN, R.N., and Professor A. PEDLER, F.R.S. Price 6*s*. 6*d*.

A. 292. On the Heat Dissipated by a Platinum Surface at High Temperatures.—Part IV. Thermal Emissivity in High-pressure Gases. By J. E. PETAVEL, A.M. Inst. C.E., A.M. Inst. E.E., Past 1851 Exhibition Scholar, John Harling Fellow of the Owens College, Manchester. Price 2*s*.

A. 293. A Comparative Crystallographical Study of the Double Selenates of the Series $R_2M(SeO_4)_2,6H_2O$. Part II. Salts in which M is Magnesium. By A. E. TUTTON, B.Sc., F.R.S. Price 2*s*.

A. 294. Mathematical Contributions to the Theory of Evolution.—IX. On the Principle of Homotyposis and its Relation to Heredity, to the Variability of the Individual, and to that of the Race. Part I. Homotyposis in the Vegetable Kingdom. By KARL PEARSON, F.R.S., with the assistance of ALICE LEE, D.Sc., ERNEST WARREN, D.Sc., AGNES FRY, CICELY D. FAWCETT, B.Sc., and others. Price 5*s*. 6*d*.

A. 295. Wave-length Determinations and General Results obtained from a Detailed Examination of Spectra photographed at the Solar Eclipse of January 22, 1898. By JOHN EVERSHED, F.R.A.S. Price 2*s*. 6*d*.

A. 296. On the Electrical Conductivity of Air and Salt Vapours. By HAROLD A. WILSON, D.Sc., M.Sc., B.A., Allen Scholar, Cavendish Laboratory, Cambridge. Price 1*s*. 6*d*.

Series B.—Biological.

B. 197. Investigations on the Abnormal Outgrowths or Intumescences on *Hibiscus vitifolius*, Linn. A Study in Experimental Plant Pathology. By ELIZABETH DALE, Pfeiffer Student, Girton College, Cambridge. Price 1*s*.

B. 198. On the "Blaze Currents" of the Frog's Eyeball. By AUGUSTUS D. WALLER, M.D., F.R.S. Price 4*s*.

B. 199. A Chemical Study of the Phosphoric Acid and Potash Contents of the Wheat Soils of Broadbalk Field, Rothamsted. By BERNARD DYER, D.Sc., F.I.C. Price 2*s*. 6*d*.

B. 200. On the Structure and Affinities of Fossil Plants from the Palæozoic Rocks.—IV. The Seed-like Fructification of *Lepidocarpon*, a Genus of Lycopodiaceous Cones from the Carboniferous Formation. By D. H. SCOTT, M.A., Ph.D., F.R.S., Hon. Keeper of the Jodrell Laboratory, Royal Gardens, Kew. (*Shortly*.)

B. 201. The Distribution of Vertebrate Animals in India, Ceylon, and Burma. By W. T. BLANFORD, F.R.S. Price 6*s*.

B. 202. The Chemistry of Nerve-degeneration. By F. W. MOTT, M.D., F.R.S., and W. D. HALLIBURTON, M.D., F.R.S. Price 3*s*.

B. 203. A Conjugating "Yeast." By B. T. P. BARKER, B.A., Gonville and Caius College, Cambridge. Price 1*s*. 6*d*.

Cloth cases for binding the 'Philosophical Transactions' can be obtained of HARRISON and SONS, St. Martin's Lane. Price 1*s*. 6*d*.

INDEX SLIP.

Pearson, Karl.—Mathematical Contributions to the Theory of Evolution. XI—On the Influence of Natural Selection on the Variability and Correlation of Organs.
Phil. Trans., A, vol. 200, 1902, pp. 1-66.

Correlation of Organs, Relation to Natural Selection.
Pearson, Karl. Phil. Trans., A, vol. 200, 1902, pp. 1-66.

Evolution, Mathematical Theory of.
Pearson, Karl. Phil. Trans., A, vol. 200, 1902, pp. 1-66.

Selection—Its Relation to Variability and Correlation of Organs, and to Heredity.
Pearson, Karl. Phil. Trans., A, vol. 200, 1902, pp. 1-66.

Variability of Organs, Relation to Natural Selection.
Pearson, Karl. Phil. Trans., A, vol. 200, 1902, pp. 1-66.

PHILOSOPHICAL TRANSACTIONS.

I. *Mathematical Contributions to the Theory of Evolution.*—XI. *On the Influence of Natural Selection on the Variability and Correlation of Organs.*

By KARL PEARSON, *F.R.S.*, *University College, London.*

Received December 20, 1901,—Read January 23, 1902.

CONTENTS.

(1.) *Introductory. On the Influence of Selection upon Correlation.*

At an earlier stage in the development of the statistical theory of evolution it was suggested that the coefficient of correlation (Galton's function) might be found constant for all races of the same species—in fact, it was considered possible that this coefficient might be the long-sought-for criterion of identity in species. Professor Weldon, following up this suggestion of Mr. Galton's, then made the elaborate series of measurements on crabs with which his name will always be closely associated. To a first approximation these researches seemed to confirm the possibility of Galton's function being a true criterion of species. When, however, a finer mathematical test was applied to Professor Weldon's observations as well as to other statistical series for organs in man,* it became clear that the coefficient of correlation varied from local race to local race, and could not be used as a criterion of species. A slight investigation undertaken in the summer of 1896 convinced me that the coefficient of correlation between any two organs, is just as much peculiar and characteristic of a local race as the means and variations of those organs. In fact, if local races be the outcome of natural selection, then their coefficients of correlation must in general differ. The object of the present paper is to show, not only that natural selection must determine the amount of correlation, but that it is probably the chief factor in the production of correlation. If selection, natural or artificial, be capable of producing correlation, then it seems impossible to regard all correlation as evidence of a causal nexus,† although the converse proposition that all causal nexus denotes correlation, is undoubtedly the most philosophical method of regarding causality.

In dealing with the influence of selection on correlation, I shall suppose the distribution of complex groups of organs to follow the normal correlation surface—the generalised Gaussian law of frequency. I shall further assume the selection surfaces to be normal in character. Neither of these assumptions is absolutely true, but the Gaussian law in a good many cases describes the frequency sufficiently closely to enable us to obtain fair numerical results by its application. Probably in all cases it will enable us to reach *qualitative* if not accurate quantitative theoretical deductions. I have the less hesitation in asserting this, as Mr. G. U. Yule has recently succeeded in deducing the chief formulæ for correlation and regression as given by the Gaussian law from general principles, which make no appeal to a special law of frequency.‡

---

* "Mathematical Contributions to the Theory of Evolution.—III.," 'Phil. Trans.,' A, vol. 189, pp. 266 and 280. See also E. Warren, "Variation in *Portunus depurator*," 'Roy. Soc. Proc.,' vol. 60, pp. 233–4.

† See a series of letters in 'Nature,' vol. 54, 1896, arising from a discussion upon a paper by A. R. Wallace.

‡ 'Roy. Soc. Proc.,' vol. 60, p. 477.

In particular, if a selected group be not given with very great accuracy by a normal frequency surface, still we may, I think, consider ourselves justified in supposing that the effects of the actual selection, and those of a normal selection with the same means, the same amounts of variation, and with correlations of the same intensity will be at least qualitatively alike in character, if they be not indeed exactly the same quantitatively within the limits fixed by the probable errors of the constants.

My plan in this memoir will be as follows:—I shall first state the fundamental theorem in multiple correlation with a new proof, so that the formulæ required may be once for all collected for reference.* I shall then give the algebraic investigation of the new formulæ for selection. I shall afterwards consider special simple cases, and illustrate them by examples. Finally I shall draw attention to the nature of the selective death-rate as indicated in cases of this kind, and consider at length its algebraic theory. Throughout I shall endeavour to illustrate the somewhat complex algebra by arithmetical examples.

(2.) *On the Fundamental Theorem in Multiple Correlation.*

I have shown in my memoir on "Regression, Heredity, and Panmixia" ('Phil. Trans.,' A, vol. 187, p. 261) that if the n variables of a complex be functions of m $(m > n)$ independent variables with frequency distributions following the normal law, and such that the principle of superposition holds for the deviations from the means supposed small; then the frequency of the complex with deviations from the means of the n variables lying between x_1 and $x_1 + \delta x_1$, x_2 and $x_2 + \delta x_2 \ldots x_n$ and $x_n + \delta x_n$ will be $z\, \delta x_1\, \delta x_2 \ldots \delta x_n$ where z:

$$z = z_0 \text{ expt.} - \tfrac{1}{2}\left(\mathrm{S}_1\left(c_{pp}x_{pp}{}^2\right) + 2\mathrm{S}_2\left(c_{pq}x_px_q\right)\right) \quad . \quad . \quad . \quad . \quad . \quad \text{(i.)}.$$

Here z_0, c_{pp}, $c_{pq} \ldots$ are constants, and S_1 denotes a summation for every value of p, and S_2 for every pair of values of p and q in the series from 1 to n.

In the same memoir (p. 302) I have determined the values of z_0, c_{pp}, c_{pq} in terms of the correlations r_{pq} and the standard deviations σ_p and σ_q of the n variables. This point had already been considered by Professor EDGEWORTH ('Phil. Mag.,' vol. 34, p. 201, 1892), and some further results by Mr. A. BLACK, reached before his death in 1893, were published in the 'Camb. Phil. Trans.' (vol. 16, p. 219, 1897). The present investigation is, I think, novel, and adds to results already reached others required in the present memoir, so that it thus places together with a fairly simple proof all the fundamental results to which I shall have occasion to appeal later.

---

* We have used these formulæ for several years, but they do not appear to have been hitherto published in a collected form.

Let us consider the quadric of the n^{th} order

$$Q = S_1 (c_{pp}x_p^2) + 2S_2 (c_{pq}x_px_q) = \text{constant},$$

and fix our attention on two of the variables, say the first two, x_1 and x_2. If these be considered constants, the quadric of the remaining $n - 2$ variables will not now be referred to its "centre." But its centre, $\bar{x}_3, \bar{x}_4 \dots \bar{x}_n$, will be given by the equations:

$$\begin{array}{l}
c_{13}x_1 + c_{23}x_2 + c_{33}\bar{x}_3 + c_{43}\bar{x}_4 + \quad \dots \quad + c_{n3}\bar{x}_n = 0. \\
c_{14}x_1 + c_{24}x_2 + c_{34}\bar{x}_3 + c_{44}\bar{x}_4 + \quad \dots \quad + c_{n4}\bar{x}_n = 0. \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
c_{1n}x_1 + c_{2n}x_2 + c_{3n}\bar{x}_3 + c_{4n}\bar{x}_4 + \quad \dots \quad + c_{nn}\bar{x}_n = 0. \qquad \dots \quad \text{(ii.)}.
\end{array}$$

The following expressions will not be zero, but will be written α and β:

$$\begin{array}{l}
c_{11}x_1 + c_{21}x_2 + c_{31}\bar{x}_3 + c_{41}\bar{x}_4 + \quad \dots \quad + c_{n1}\bar{x}_n = \alpha. \\
c_{12}x_1 + c_{22}x_2 + c_{32}\bar{x}_3 + c_{42}\bar{x}_4 + \quad \dots \quad + c_{n2}\bar{x}_n = \beta. \qquad \dots \quad \text{(iii.)}.
\end{array}$$

Now, if Δ be the discriminant,

$$\left|\begin{array}{ccccc}
c_{11}, & c_{12}, & c_{13}, & \dots & c_{1n} \\
c_{21}, & c_{22}, & c_{23}, & \dots & c_{2n} \\
c_{31}, & c_{32}, & c_{33}, & \dots & c_{3n} \\
\dots & \dots & \dots & \dots & \dots \\
c_{n1}, & c_{n2}, & c_{n3}, & \dots & c_{nn}
\end{array}\right| \qquad \dots\dots \quad \text{(iv.)},$$

and C_{pq} the minor corresponding to c_{pq}, we have by solving the above n linear equations,

$$x_1 = \frac{\alpha C_{11} + \beta C_{12}}{\Delta}, \qquad x_2 = \frac{\alpha C_{12} + \beta C_{22}}{\Delta},$$

whence:

$$\alpha = \frac{\Delta (x_1C_{22} - x_2C_{12})}{C_{11}C_{22} - C_{12}^2}, \qquad \beta = \frac{\Delta (x_2C_{11} - x_1C_{12})}{C_{11}C_{22} - C_{12}^2} \quad \dots\dots \quad \text{(v.)}.$$

Generally also:

$$\begin{aligned}
\bar{x}_q &= \frac{\alpha C_{1q} + \beta C_{2q}}{\Delta} \\
&= \frac{(C_{1q}C_{22} - C_{2q}C_{12})x_1 + (C_{11}C_{2q} - C_{1q}C_{12})x_2}{C_{11}C_{22} - C_{12}^2} \\
&= D_{1q}x_1 + D_{2q}x_2, \quad \text{say}, \dots\dots\dots \quad \text{(vi.)}.
\end{aligned}$$

This determines the central co-ordinate for any variable x_q for a given value of x_1 and x_2.

Now let us transfer the quadric to $\bar{x}_3, \bar{x}_4 \ldots \bar{x}_n$ as origin. It may be written

$$\begin{aligned} \mathrm{Q} &= \mathrm{S}_{q=1}^{q=n}\, x_q\,(c_{1q}x_1 + c_{2q}x_2 + c_{3q}x_3 + \ldots + c_{nq}x_n) \\ &= \mathrm{S}_{q=3}^{q=n}\,(\bar{x}_q + x'_q)\begin{pmatrix} c_{1q}x_1 + c_{2q}x_2 + c_{3q}\bar{x}_3 + \ldots + c_{nq}\bar{x}_n \\ + c_{3q}x'_3 + \ldots + c_{nq}x'_n \end{pmatrix} \\ &\quad + x_1\begin{pmatrix} c_{11}x_1 + c_{21}x_2 + c_{31}\bar{x}_3 + \ldots + c_{n1}\bar{x}_n \\ + c_{31}x'_3 + \ldots + c_{n1}x'_n \end{pmatrix} \\ &\quad + x_2\begin{pmatrix} c_{12}x_1 + c_{22}x_2 + c_{32}\bar{x}_3 + \ldots + c_{n2}\bar{x}_n \\ + c_{32}x'_3 + \ldots + c_{n2}x'_n \end{pmatrix}. \end{aligned}$$

Making use of the n linear equations (ii.) and (iii.):

$$\begin{aligned} \mathrm{Q} &= \mathrm{S}_{q=3}^{q=n}\,(\bar{x}_q + x'_q)\,(c_{3q}x'_3 + c_{4q}x'_4 + \ldots + c_{nq}x'_n) \\ &\quad + x_1\,(\alpha + c_{31}x'_3 + c_{41}x'_4 + \ldots + c_{n1}x'_n) \\ &\quad + x_2\,(\beta + c_{32}x'_3 + c_{42}x'_4 + \ldots + c_{n2}x'_n) \\ &= \mathrm{S}_{q=3}^{q=n}\, x'_q\,(c_{3q}x'_3 + c_{4q}x'_4 + \ldots + c_{nq}x'_n) \\ &\quad + \alpha x_1 + \beta x_2 \quad \ldots\ldots\ldots\ldots \quad \text{(vii.)}. \end{aligned}$$

For arranging vertical columns in rows, the remaining terms are

$$\begin{aligned} & x'_3\,(c_{31}x_1 + c_{32}x_2 + \mathrm{S}_{q=3}^{q=n}\,(c_{3q}\bar{x}_q)) \\ + & x'_4\,(c_{41}x_1 + c_{42}x_2 + \mathrm{S}_{q=3}^{q=n}\,(c_{4q}\bar{x}_q)) \\ & \ldots\ldots\ldots\ldots\ldots \\ + & x'_n\,(c_{n1}x_1 + c_{n2}x_2 + \mathrm{S}_{q=3}^{q=n}\,(c_{nq}\bar{x}_q)), \end{aligned}$$

each line of which vanishes by the equations (ii.) for the centre.

Accordingly:

$$\mathrm{Q} = \mathrm{Q}' + \alpha x_1 + \beta x_2,$$

where Q′ is a quadratic function of $x'_3, x'_4, \ldots x'_n$, not involving x_1 and x_2 at all.

Hence: $z = z_0\, e^{-\frac{1}{2}(\mathrm{Q}' + \alpha x_1 + \beta x_2)}$.

Now integrate z with respect to all the variables $x'_3, x'_4, \ldots x'_n$ from $-\infty$ to $+\infty$, keeping x_1 and x_2 constant.

Then, although the origin is a function of x_1 and x_2,

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\!\!\ldots\int_{-\infty}^{+\infty} z_0 e^{-\frac{1}{2}\mathrm{Q}'}\, dx'_3\, dx'_4 \ldots dx'_n$$

cannot involve x_1 and x_2 but only $c_{3q}, c_{4q} \ldots c_{nq}$, &c.; let the result be ζ_0. Then:

$$z = \zeta_0 e^{-\frac{1}{2}(\alpha x_1 + \beta x_2)}$$

must be the correlation surface for x_1 and x_2, for all values of $x_3, x_4, \ldots x_n$. We may write it in the form

$$z = \zeta_0 \text{ expt.} - \tfrac{1}{2}\Delta \frac{x_1^2 C_{22} - 2C_{12}x_1x_2 + C_{11}x_2^2}{C_{11}C_{22} - C_{12}^2} \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(viii.)}.$$

Comparing this with the known form of the correlation surface for two variables,*

$$z = \frac{N}{2\pi\sigma_1\sigma_2\sqrt{1 - r_{12}^2}} \text{ expt.} - \tfrac{1}{2}\frac{1}{1 - r_{12}^2}\left(\frac{x_1^2}{\sigma_1^2} - \frac{2r_{12}x_1x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right) \quad . \quad . \quad . \qquad \text{(ix.)},$$

we have at once

$$\sigma_1^2(1 - r_{12}^2) = \frac{C_{11}C_{22} - C_{12}^2}{\Delta C_{22}}, \qquad \sigma_2^2(1 - r_{12}^2) = \frac{C_{11}C_{22} - C_{12}^2}{\Delta C_{11}},$$

$$r_{12}\sigma_1\sigma_2 = C_{12}/\Delta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(x.)}.$$

Whence

$$r_{12} = C_{12}/\sqrt{C_{11}C_{22}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xi.)},$$

$$\sigma_1^2 = C_{11}/\Delta, \qquad \sigma_2^2 = C_{22}/\Delta. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xii.)}.$$

Or, generally :

$$r_{pq} = C_{pq}/\sqrt{C_{pp}C_{qq}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xiii.)},$$

$$\sigma_p = \sqrt{C_{pp}/\Delta} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xiv.)}.$$

Thus correlations and variations are fully determined in terms of the discriminant and its minor for the constants $c_{11}, c_{22}, c_{12} \ldots c_{pp}, c_{qq}, c_{pq}, \ldots c_{nn}$.

We have next the inverse proposition to find the c's in terms of the r's and σ's.

We have, by well-known propositions in the theory of determinants :

$$\begin{array}{l}
c_{11}C_{11} + c_{12}C_{12} + c_{13}C_{13} + \quad . \quad . \quad . \quad + c_{1n}C_{1n} = \Delta. \\
c_{11}C_{21} + c_{12}C_{22} + c_{13}C_{23} + \quad . \quad . \quad . \quad + c_{1n}C_{2n} = 0. \\
c_{11}C_{31} + c_{12}C_{32} + c_{13}C_{33} + \quad . \quad . \quad . \quad + c_{1n}C_{3n} = 0. \\
. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\
. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\
c_{11}C_{n1} + c_{12}C_{n2} + c_{13}C_{n3} + \quad . \quad . \quad . \quad + c_{1n}C_{nn} = 0.
\end{array}$$

Or,

$$\begin{array}{l}
c_{11}\sigma_1^2 + c_{12}r_{12}\sigma_1\sigma_2 + c_{13}r_{13}\sigma_1\sigma_3 + \quad . \quad . \quad . + c_{1n}r_{1n}\sigma_1\sigma_n = 1, \\
c_{11}r_{n1}\sigma_1\sigma_2 + c_{12}\sigma_2^2 + c_{13}r_{23}\sigma_2\sigma_3 + \quad . \quad . \quad . + c_{1n}r_{2n}\sigma_2\sigma_n = 0, \\
c_{11}r_{31}\sigma_1\sigma_3 + c_{12}r_{32}\sigma_2\sigma_3 + c_{13}\sigma_3^2 + \quad . \quad . \quad . + c_{1n}r_{3n}\sigma_3\sigma_n = 0. \\
. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\
. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\
c_{11}r_{n1}\sigma_1\sigma_n + c_{12}r_{n2}\sigma_2\sigma_n + c_{13}r_{n3}\sigma_3\sigma_n + \quad . \quad . \quad + c_{1n}\sigma_n^2 = 0.
\end{array}$$

Hence solving

$$c_{11} = S_{11}/S, \qquad c_{1q} = S_{1q}/S \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xv.)},$$

---

* PEARSON, 'Phil. Trans.,' A, vol. 187, p. 264.

where
$$S = \begin{vmatrix} \sigma_1^2, & r_{12}\sigma_1\sigma_2, & r_{13}\sigma_1\sigma_3, & \cdots & r_{1n}\sigma_1\sigma_n \\ r_{21}\sigma_1\sigma_2, & \sigma_2^2, & r_{23}\sigma_2\sigma_3, & \cdots & r_{2n}\sigma_2\sigma_n \\ r_{31}\sigma_1\sigma_3, & r_{32}\sigma_2\sigma_3, & \sigma_3^2, & \cdots & r_{3n}\sigma_3\sigma_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{n1}\sigma_1\sigma_n, & r_{n2}\sigma_2\sigma_n, & r_{n3}\sigma_3\sigma_n, & \cdots & \sigma_n^2 \end{vmatrix}$$

and S_{pq} is the minor corresponding to the constituent $r_{pq}\sigma_p\sigma_q$.

But
$$S = \sigma_1^2\sigma_2^2\sigma_3^2 \ldots \sigma_n^2 R,$$
$$S_{pq} = \sigma_1^2\sigma_2^2\sigma_3^2 \ldots \sigma_p\sigma_q \ldots \sigma_n^2 R_{pq},$$

where R is the determinant,

$$\begin{vmatrix} 1, & r_{12}, & r_{13}, & \cdots & r_{1n} \\ r_{21}, & 1, & r_{23}, & \cdots & r_{2n} \\ r_{31}, & r_{32}, & 1, & \cdots & r_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{n1}, & r_{n2}, & r_{n3}, & \cdots & 1 \end{vmatrix} \quad \ldots\ldots\ldots \quad \text{(xvi.)},$$

and R_{pq} is the minor corresponding to the constituent r_{pq}. Thus we have

$$c_{11} = \frac{1}{\sigma_1^2} R_{11}/R, \qquad c_{1q} = \frac{1}{\sigma_1\sigma_q} R_{1q}/R \quad \ldots\ldots \quad \text{(xvii.)},$$

or, generally,

$$c_{pp} = \frac{1}{\sigma_p^2}\frac{R_{pp}}{R}, \qquad c_{pq} = \frac{1}{\sigma_p\sigma_q} R_{pq}/R \quad \ldots\ldots\ldots \quad \text{(xviii.)}.$$

Thus z may be written*

$$z = z_0 \text{ expt.} - \tfrac{1}{2}\left\{S_1\left(\frac{R_{pp}}{R}\frac{x_p^2}{\sigma_p^2}\right) + 2S_2\left(\frac{R_{pq}}{R}\frac{x_px_q}{\sigma_p\sigma_q}\right)\right\} \quad \ldots \quad \text{(xix.)}.$$

It remains to determine z_0 from the fact that the volume of the surface $= N$.

Or,
$$N = z_0 \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \text{expt.} \left(-\tfrac{1}{2}Q\right) dx_1\, dx_2 \ldots dx_n$$
$$= z_0\sigma_1\sigma_2 \ldots \sigma_n \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \text{expt.} - \tfrac{1}{2}\left\{S_1\left(\frac{R_{pp}}{R}x'^2_p\right) + 2S_2\left(\frac{R_{pq}}{R}x'_px'_q\right)\right\}$$
$$dx'_1\, dx'_2 \ldots dx'_n \quad \ldots\ldots\ldots \quad \text{(xx.)},$$

by writing $x_p/\sigma_p = x'_p$, &c.

Now, integrate first with regard to x'_1 writing first

---

* 'Phil. Trans.,' A, vol. 187, p. 302.

$$\mathrm{S}_1\left(\frac{\mathrm{R}_{pp}}{\mathrm{R}}\,x'^{\,2}_p\right)+2\mathrm{S}_2\left(\frac{\mathrm{R}_{pq}}{\mathrm{R}}\,x'_px'_q\right)$$

$$=\frac{\mathrm{R}_{11}}{\mathrm{R}}\left(x'_1+\frac{\mathrm{R}_{12}}{\mathrm{R}_{11}}\,x'_2+\frac{\mathrm{R}_{13}}{\mathrm{R}_{11}}\,x'_3+\ldots+\frac{\mathrm{R}_{1n}}{\mathrm{R}_{11}}\,x'_n\right)^2$$

$$+\Sigma_1\left(\frac{\mathrm{R}_{pp}\mathrm{R}_{11}-\mathrm{R}_{1p}^{\,2}}{\mathrm{R}\mathrm{R}_{11}}\,x'^{\,2}_p\right)+2\Sigma_2\left(\frac{\mathrm{R}_{pq}\mathrm{R}_{11}-\mathrm{R}_{1p}\mathrm{R}_{1q}}{\mathrm{R}\mathrm{R}_{11}}\,x'_px'_q\right),$$

where Σ_1 is the sum for all values of p, and Σ_2 for all pairs of values of p and q, from 2 to n inclusive,

$$=\frac{\mathrm{R}_{11}}{\mathrm{R}}(x'_1+\mathrm{H}'_1)^2+\Sigma_1\left(\frac{\mathrm{R}'_{pp}}{\mathrm{R}'}\,x'^{\,2}_p\right)+2\Sigma_2\left(\frac{\mathrm{R}'_{pq}}{\mathrm{R}'}\,x'_px'_q\right)\text{ where }\mathrm{H}'_1=\Sigma_1\left(\frac{\mathrm{R}_{1p}}{\mathrm{R}_{11}}\,x'_p\right).$$

For if $\mathrm{R}'=\mathrm{R}_{11}$, or the determinant of the correlation coefficients, omitting all involving the first variable, *i.e.*, the first row and column of R, the determinant R'_{pq} corresponding to the minor of the constituent r_{pq} in R′ or to the second minor $(.\,\mathrm{R}_{11})_{pq}$ of R is given by

$$\mathrm{R}'_{pq}=(\mathrm{R}_{pq}\mathrm{R}_{11}-\mathrm{R}_{1p}\mathrm{R}_{1q})/\mathrm{R}\quad\ldots\ldots\ldots\quad\text{(xxi.)},$$

and for $p=q$,

$$\mathrm{R}'_{pp}=(\mathrm{R}_{pp}\mathrm{R}_{11}-\mathrm{R}_{1p}^{\,2})/\mathrm{R}\quad\ldots\ldots\ldots\quad\text{(xxii.)}.$$

Hence, integrating x'_1 between the limits $+\infty$ and $-\infty$, we have

$$\mathrm{N}=z_0\sigma_1\sigma_2\ldots\sigma_n\,.\,\sqrt{2\pi}\sqrt{\frac{\mathrm{R}}{\mathrm{R}'}}$$

$$\times\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty}\text{expt.}-\tfrac{1}{2}\left\{\Sigma_1\left(\frac{\mathrm{R}'_{pp}}{\mathrm{R}'}\,x'^{\,2}_p\right)+2\Sigma_2\left(\frac{\mathrm{R}'_{pq}}{\mathrm{R}'}\,x'_px'_q\right)\right\}dx'_2\,dx'_3\ldots dx'_n.$$

This is of precisely the same form as before, except that we have the factor $\sqrt{2\pi}\,\sqrt{\mathrm{R}/\mathrm{R}'}$, and the multiple integral is reduced by one integration and by the disappearance of all correlations involving the first variable. Now, integrate with regard to x'_2. The sole effect will be to multiply by a factor $\sqrt{2\pi}\,\sqrt{\mathrm{R}'/\mathrm{R}''}$, where R″ is the minor of R′ not involving correlations of the second variable. Thus, by repeating the process, we have ultimately

$$\mathrm{N}=z_0\sigma_1\sigma_2\ldots\sigma_n\,(\sqrt{2\pi})^n\sqrt{\mathrm{R}/\mathrm{R}'}\,\sqrt{\mathrm{R}'/\mathrm{R}''}\,\sqrt{\mathrm{R}''/\mathrm{R}'''}\ldots\sqrt{\frac{\mathrm{R}^{n-1}}{1}},$$

or,

$$z_0=\frac{\mathrm{N}}{(2\pi)^{\frac{1}{2}n}\,\sigma_1\sigma_2\ldots\sigma_n\sqrt{\mathrm{R}}}\quad\ldots\ldots\ldots\quad\text{(xxiii.)},$$

which gives the constant of the surface.

The preceding investigation enables us also to deal with two further points.

(*a*) Given n variables, what is the mean value $m_1+\overline{x_1}$ of the first variable and its variability $\overline{\sigma'_1}$ for definite values m_2+h_2, $m_3+h_3\ldots m_n+h_n$ of the other $(n-1)$ variables?

Clearly, when $x'_2 x'_3 \dots x'_n$ are constants, the distribution of x'_1 is of the form

$$\text{constant} \times \text{expt.} - \tfrac{1}{2} \frac{R_{11}}{R} \left\{ x'_1 + \Sigma_1 \left(\frac{R_{1p}}{R_{11}} x'_p \right) \right\}^2.$$

Or, re-introducing the $\sigma_1, \sigma_2, \dots \sigma_n$, we have a distribution about the point given by

$$\left. \begin{aligned} \bar{x}_1 &= -\sigma_1 \left(\frac{R_{12}}{R_{11}} \frac{x_2}{\sigma_2} + \frac{R_{13}}{R_{11}} \frac{x_3}{\sigma_3} + \dots + \frac{R_{1n}}{R_{11}} \frac{x_n}{\sigma_n} \right), \\ \text{with standard deviation} \qquad \bar{\sigma}'_1 &= \sigma_1 \sqrt{R/R_{11}} \end{aligned} \right\} \quad . \quad . \quad . \quad . \quad \text{(xxiv.)}.$$

(*b*) Given n variables, what are the mean values $m_1 + \bar{x}_1$, $m_2 + \bar{x}_2$, the standard deviations $\bar{\sigma}''_1$, $\bar{\sigma}''_2$, and the correlation $\bar{r}_{12}$ of two of them, when we give definite values $m_3 + h_3$, $m_4 + h_4 \dots m_n + h_n$ to the remaining $(n - 2)$ variables?

In this case we have from (i.)

$$\begin{aligned} z = \text{expt.} - \tfrac{1}{2} \{ c_{11} x_1^2 &+ 2 c_{12} x_1 x_2 + c_{22} x_2^2 \\ &+ 2 (c_{13} h_3 + c_{14} h_{14} + \dots + c_{1n} h_n) x_1 \\ &+ 2 (c_{23} h_3 + c_{24} h_4 \ + \dots + c_{2n} h_n) x_2 \\ &+ \text{terms not involving } x_1 \text{ and } x_2 \} \quad . \quad . \quad . \quad . \quad \text{(xxv.)}. \end{aligned}$$

Writing K_1 for the coefficient of x_1, and K_2 for that of x_2, we have for the centre

$$\bar{x}_1 = -\frac{(K_1 c_{22} - K_2 c_{12})}{c_{11} c_{22} - c_{12}^2}, \qquad \bar{x}_2 = -\frac{(-K_1 c_{12} + c_{11} K_2)}{c_{11} c_{22} - c_{12}^2},$$

$$\bar{x}_1 = -\Sigma_1 \left(\frac{c_{1p} c_{22} - c_{2p} c_{12}}{c_{11} c_{22} - c_{12}^2} h_p \right), \qquad \bar{x}_2 = -\Sigma_1 \left(\frac{c_{2p} c_{11} - c_{1p} c_{12}}{c_{11} c_{22} - c_{12}^2} h_p \right),$$

$$\bar{x}_1 = -\sigma_1 \Sigma_1 \left(\frac{R_{1p} R_{22} - R_{2p} R_{12}}{R_{11} R_{22} - R_{12}^2} \frac{h_p}{\sigma_p} \right), \qquad \bar{x}_2 = -\sigma_2 \Sigma_1 \left(\frac{R_{2p} R_{11} - R_{1p} R_{12}}{R_{11} R_{22} - R_{12}^2} \frac{h_p}{\sigma_p} \right),$$

by transferring to the minors of R and the σ's. Or,

$$\bar{x}_1 = -\sigma_1 \Sigma_1 \left(\frac{\rho''_{1p}}{R''} \frac{h_p}{\sigma_p} \right), \qquad \bar{x}_2 = -\sigma_2 \Sigma_2 \left(\frac{\rho'_{2p}}{R''} \frac{h_p}{\sigma_p} \right) \quad . \quad . \quad . \quad \text{(xxvi.)}.$$

Here R'' is the determinant formed by striking out the first two rows and columns of R; ρ''_{1p} is the minor obtained by striking out the second row and column from R, and then the first row and p^{th} column; ρ'_{2p} the minor obtained by striking out the first row and column, and then the second row and p^{th} column. But a comparison with (xxiv.) shows us that these values for $\bar{x}_1$ and $\bar{x}_2$ are precisely what we should have obtained for the regression equations of the 1st and 2nd variables respectively *alone* on the other $n - 2$ variables. Thus the existence and the correlations of x_2 have no effect on the value of $\bar{x}_1$, nor those of x_1 on the value of $\bar{x}_2$.

Returning to (xxv.), we remark that the terms of the second order in x_1 and x_2, on which the correlation and variations depend, are not altered by a transfer to the centre $\bar{x}_1$ and $\bar{x}_2$ of the array.

Hence by (ix.) and (xvii.) we have

$$\bar{r}_{12} = -\frac{c_{12}}{\sqrt{c_{11}c_{22}}} = -\mathrm{R}_{12}/\sqrt{\mathrm{R}_{11}\mathrm{R}_{22}} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxvii.)}.$$

This is the *partial* correlation of the 1st and 2nd organs for the remaining $n - 2$ organs with constant values.

Again,

$$c_{11} = \frac{1}{\bar{\sigma}''^2_1(1 - \bar{r}_{12}^2)}, \quad c_{22} = \frac{1}{\bar{\sigma}''^2_2(1 - \bar{r}_{12}^2)}.$$

Whence we easily find from (xviii.) and (xxvii.)

$$\bar{\sigma}''^2_1 = \sigma_1^2 \frac{\mathrm{R}\,\mathrm{R}_{22}}{\mathrm{R}_{11}\mathrm{R}_{22} - \mathrm{R}_{12}^2}, \quad \bar{\sigma}''^2_2 = \sigma_2^2 \frac{\mathrm{R}\,\mathrm{R}_{11}}{\mathrm{R}_{11}\mathrm{R}_{22} - \mathrm{R}_{12}^2},$$

or,

$$\bar{\sigma}''_1 = \sigma_1 \sqrt{\mathrm{R}_{22}/\mathrm{R}''}, \quad \bar{\sigma}''_2 = \sigma_2 \sqrt{\mathrm{R}_{11}/\mathrm{R}''} \quad . \quad . \quad . \quad . \quad \text{(xxviii.)},$$

where R″ is, as before, the determinant R without its first two rows and columns. These by (xxiv.), are what we should have reached by ignoring x_2 in finding $\bar{\sigma}'_1$, and x_1 in finding $\bar{\sigma}'_2$.

(3.) *General Theorem in Selection.*

To find the selected means, the selected variations and selected correlations, when q organs are selected, naturally or artificially, out of a complex of n organs.

Let the selected group of q organs have their means raised $h_1, h_2, h_3, \ldots h_q$ (some of these quantities may be negative); their standard deviations changed from $\sigma_1, \sigma_2, \ldots \sigma_q$ to $s_1, s_2, s_3, \ldots s_q$, and their mutual correlations from $r_{12}, r_{13}, \ldots r_{1q}, r_{23}, r_{24}, \ldots r_{2q}, \ldots r_{q-1,q}$ to $\rho_{12}, \rho_{13}, \ldots \rho_{1q}, \rho_{23}, \rho_{24}, \ldots \rho_{2q}, \ldots \rho_{q-1,q}$.

The whole system of n organs before selection will be defined by the means as origin of measurement for each organ, by the standard deviations $\sigma_1, \sigma_2, \sigma_3, \ldots \sigma_n$, and by the coefficients of correlation $r_{12}, r_{13}, \ldots r_{1n}, r_{23}, r_{24}, \ldots r_{2n}, \ldots r_{n-1n}$. Let R be the determinant

$$\begin{vmatrix} 1, & r_{12}, & r_{13}, & . & . & . & r_{1n} \\ r_{21}, & 1, & r_{23}, & . & . & . & r_{2n} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ r_{n1}, & r_{n2}, & r_{n3}, & . & . & . & 1 \end{vmatrix}$$

and R_{uv} the minor corresponding to the constituent r_{uv}. Then the unselected population is given by the frequency surface of equation (xix.),

$$z = \text{constant} \times \text{expt.} - \tfrac{1}{2}\left\{\mathrm{S}_1\left(\frac{\mathrm{R}_{pp}}{\mathrm{R}}\frac{x_p^2}{\sigma_p^2}\right) + 2\mathrm{S}'_2\left(\frac{\mathrm{R}_{pq}}{\mathrm{R}}\frac{x_p x_q}{\sigma_p\sigma_q}\right)\right\} \quad . \quad . \quad \text{(xxix.)}.$$

For brevity, we can also write this in the form

$$z = \text{constant} \times \text{expt.} - \tfrac{1}{2}\{\mathrm{S}_1(c_{pp}x_p^2) + 2\mathrm{S}_2(c_{pq}x_px_q)\} \quad . \quad . \quad \text{(xxix.)}^{\text{bis}}.$$

Now consider for the time only $q + 1$ organs—namely, the first q organs and the m^{th} organ ($m > q$), and let us write $\mathrm{R}(u)$, if u be $> q$, for the determinant :

$$\mathrm{R}(u) = \begin{vmatrix} 1, & r_{12}, & r_{13}, & . & . & . & r_{1q}, & r_{1u} \\ r_{21}, & 1, & r_{23}, & . & . & . & r_{2q}, & r_{2u} \\ r_{31}, & r_{32}, & 1, & . & . & . & r_{3q}, & r_{3u} \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ r_{q1}, & r_{q2}, & r_{q3}, & . & . & . & 1, & r_{qu} \\ r_{u1}, & r_{u2}, & r_{u3}, & . & . & . & r_{uq}, & 1 \end{vmatrix} \quad . \quad . \quad . \quad . \quad . \quad \text{(xxx.)}.$$

Then if $\mathrm{R}(u)_{p'p''}$ be the minor corresponding to the constituent $r_{p'p''}$, and if $b_{p'p''} = \dfrac{\mathrm{R}(u)_{p'p''}}{\mathrm{R}(u)}\dfrac{1}{\sigma_{p'}\sigma_{p}}$, the distribution of the $q + 1$ organs will be given by the frequency distribution

$$z' = \text{constant} \times \text{expt.} - \tfrac{1}{2}\{\mathrm{S}_1(b_{p'p'}x_{p'}^2) + 2\mathrm{S}_2(b_{p'p''}x_{p'}x_{p''})\} \quad . \quad . \quad \text{(xxxi.)}.$$

S_1 being a sum for every value of p' throughout the $q + 1$ organs, and S_2 for every pair of values.

Now let the first q organs be given values $h_1, h_2, \ldots h_q$, then the mean value of x_u will be given by

$$\begin{aligned} x'_u &= -\left(\frac{b_{1u}}{b_{uu}}h_1 + \frac{b_{2u}}{b_{uu}}h_2 + \ldots + \frac{b_{qu}}{b_{uu}}h_q\right), \\ &= -\left(\frac{\mathrm{R}(u)_{1u}}{\mathrm{R}(u)_{uu}}\frac{\sigma_u}{\sigma_1}h_1 + \frac{\mathrm{R}(u)_{2u}}{\mathrm{R}(u)_{uu}}\frac{\sigma_u}{\sigma_2}h_2 + \ldots \frac{\mathrm{R}(u)_{qu}}{\mathrm{R}(u)_{uu}}\frac{\sigma_u}{\sigma_q}h_q\right) \quad . \quad . \quad \text{(xxxii.)}. \end{aligned}$$

Now these coefficients can be found at once if q be known.

For example :

$$q = 1, \qquad -\frac{\mathrm{R}(u)_{qu}}{\mathrm{R}(u)_{uu}} = r_{1u},$$

$$q = 2, \qquad -\frac{\mathrm{R}(u)_{1u}}{\mathrm{R}(u)_{uu}} = \frac{r_{1u} - r_{2u}r_{12}}{1 - r_{12}^2}, \qquad -\frac{\mathrm{R}(u)_{2u}}{\mathrm{R}(u)_{uu}} = \frac{r_{2u} - r_{1u}r_{12}}{1 - r_{12}^2},$$

$$q = 3, \qquad -\frac{\mathrm{R}(u)_{1u}}{\mathrm{R}(u)_{uu}} = \frac{r_{1u}(1 - r_{23}^2) - r_{12}r_{2u} - r_{13}r_{3u} + r_{23}(r_{12}r_{3u} + r_{13}r_{2u})}{1 - r_{12}^2 - r_{23}^2 - r_{13}^2 + 2r_{12}r_{23}r_{13}}$$

and $-\dfrac{\mathrm{R}(u)_{2u}}{\mathrm{R}(u)_{uu}}$ and $-\dfrac{\mathrm{R}(u)_{3u}}{\mathrm{R}(u)_{uu}}$ can be written down by symmetry . (xxxiii.).

Now suppose in the expression (xxix.)$^{\text{bis}}$ we were to put $x_1, x_2, \ldots x_q$ equal to $h_1, h_2, \ldots h_q$ respectively, then the system of equations to find the means of $x_{q+1}, x_{q+2}, \ldots x_{q+u}, \ldots x_n$ for this array would be the $n - q$ equations

$$
\begin{array}{l}
c_{1,q+1}\,h_1 + c_{2,q+1}\,h_2 + \ldots + c_{q,q+1}\,h_q + c_{q+1,q+1}\,x'_{q+1} + c_{q+2,q+1}\,x'_{q+2} + \ldots + c_{n,q+1}\,x'_n = 0. \\
c_{1,q+2}\,h_1 + c_{2,q+2}\,h_2 + \ldots + c_{q,q+2}\,h_q + c_{q+1,q+2}\,x'_{q+1} + c_{q+2,q+2}\,x'_{q+2} + \ldots + c_{n,q+2}\,x'_n = 0. \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
c_{1n}\,h_1 + c_{2n}\,h_2 + \ldots + c_{q,n}\,h_q + c_{q+1,n}\,x'_{q+1} + c_{q+2,n}\,x'_{q+2} + \ldots + c_{nn}\,x'_n = 0 \quad . \quad \text{(xxxiv.)},
\end{array}
$$

where $x'_{q+1}, x'_{q+2}, \ldots x'_n$ are the "co-ordinates of the centre" of the array of the $n - q$ organs.

If we were to solve these equations we ought to get precisely the solution for x'_u $(u > q < n + 1)$ that we have found for x'_u in (xxxii.) above, where none of the coefficients involve correlation-coefficients other than those of the first q organs among themselves and with the u^{th} organ. This result flows from the pretty obvious law that the mean of the u^{th} organ for an array determined by values of the first q organs cannot be in any way dependent on our considering the relation of this selection of q organs to any additional organs beside the u^{th}: see p. 9.

Thus the solution of (xxxiv.) is simply obtained by putting $u = q + 1$, $q + 2$, $q + 3, \ldots n$ successively in (xxxiii.).

Let us now select the first q organs not with absolute values, but varying about means $h_1, h_2, \ldots h_q$, with standard deviations $s_1, s_2 \ldots s_q$, and with mutual correlations $\rho_{12}, \rho_{13}, \ldots \rho_{1q}, \rho_{23}, \ldots \rho_{2q}, \ldots \rho_{q-1q}$. We have then to multiply ζ in (xxix.)$^{\text{bis}}$ by an exponential quadratic function of the $x_1 + h_1, x_2 + h_2, \ldots x_q + h_q$, *i.e.*, the selective correlation surface, and divide it by another exponential quadratic surface, *i.e.*, the primary correlation surface of the q organs $x_1, x_2 \ldots x_q$. This follows, since the frequency of each complex of n organs must be reduced in the ratio of the selected to the primary frequency of the complex of q selected organs. But it will be clear that such a reduction must give us a result of the following form for the final frequency surface of the n organs :

$$
\begin{array}{rl}
\text{Z} = \text{constant} \times \text{expt.} - \tfrac{1}{2} & \{\bar{c}_{11}x_1^2 + \bar{c}_{22}x_2^2 + \ldots + \bar{c}_{qq}x_q^2 \\
& + c_{q+1,q+2}x_{q+1}^2 + \ldots + c_{nn}x_n^2\} \\
& + 2\bar{c}_{v'v''}x_{v'}x_{v''} + \ldots \ (v' \text{ and } v'' < q + 1) \\
& + 2c_{vu}x_v x_u + \ldots \ (v < q + 1 \text{ and } u > q) \\
& + 2c_{u'u''}x_{u'}x_{u''} + \ldots \ (u' \text{ and } u'' > q) \\
& + \text{linear terms in } x_{v'}, x_{v''} \ (v' \text{ and } v'' < q + 1) \quad . \quad \text{(xxxv.)},
\end{array}
$$

where the $\bar{c}$'s denote the changed c's.

Now if we differentiate the quadric to find its "centre" we have n equations in $x_1 \ldots x_q \ldots x_n$, but the solutions of these, if $x_{q+1} \ldots x_n$ were eliminated, are known to be the "centre" $h_1, h_2 \ldots h_q$. Hence we require only $n - q$ equations involving $x_{q+1} \ldots x_n$ and we can put the h's for the remaining values $x_1 \ldots x_q$. Let us take the differentials of the quadric with regard to $x_{q+1} \ldots x_n$, then the resulting equations involve none of the $\bar{c}$'s, but only the c's. They reproduce in fact (xxxiv.). But the values of $x'_{q+1} \ldots x'_n$ found from (xxxiv.) are, we have seen, identical with the values of (xxxii.). Thus we have

$$x'_{q+1} = -\left\{\frac{\mathrm{R}(q+1)_{1,q+1}}{\mathrm{R}(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_1}h_1 + \frac{\mathrm{R}(q+1)_{2,q+1}}{\mathrm{R}(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_2}h_2 + \ldots + \frac{\mathrm{R}(q+1)_{q,q+1}}{\mathrm{R}(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_q}h_q\right\},$$

$$x'_{q+2} = -\left\{\frac{\mathrm{R}(q+2)_{1,q+2}}{\mathrm{R}(q+2)_{q+2,q+1}}\frac{\sigma_{q+2}}{\sigma_1}h_1 + \frac{\mathrm{R}(q+2)_{2,q+2}}{\mathrm{R}(q+2)_{q+2,q+2}}\frac{\sigma_{q+2}}{\sigma_2}h_2 + \ldots + \frac{\mathrm{R}(q+2)_{q,q+2}}{\mathrm{R}(q+2)_{q+2,q+2}}\frac{\sigma_{q+2}}{\sigma_q}h_q\right\},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$x'_n = -\left\{\frac{\mathrm{R}(n)_{1,n}}{\mathrm{R}(n)_{nn}}\frac{\sigma_n}{\sigma_1}h_1 + \frac{\mathrm{R}(n)_{2,n}}{\mathrm{R}(n)_{nn}}\frac{\sigma_n}{\sigma_2}h_2 + \ldots + \frac{\mathrm{R}(n)_{q,n}}{\mathrm{R}(n)_{nn}}\frac{\sigma_n}{\sigma_q}h_q\right\} \quad \ldots\ldots \quad \text{(xxxvi.)}.$$

These give the most general form of a theorem proved for a particular case in 'Phil. Trans.,' A, vol. 187, p. 300, c (ii.). *If systems of q organs be selected with any arbitrary variations and correlations out of complexes of n organs, then the* mean *sizes of the remaining $n - q$ organs have precisely the same values as if the selection of all the systems of q organs had been to one size and not varied about mean values. The arbitrary variations of the selected systems about these mean values, as well as the arbitrary correlations, have no influence on the mean changes of the $n - q$ organs.*

Returning to equation (xxxv.) we know that if the determinant

$$\Delta = \begin{vmatrix} \bar{c}_{1,1}, & \bar{c}_{1,2}, & \bar{c}_{1,3} & \ldots & \bar{c}_{1,q}, & c_{1,q+1} & \ldots & c_{1,n} \\ \bar{c}_{2,1}, & \bar{c}_{2,2}, & \bar{c}_{2,3} & \ldots & \bar{c}_{2,q}, & c_{2,q+1} & \ldots & c_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \bar{c}_{q1}, & \bar{c}_{q2}, & \bar{c}_{q,3} & \ldots & \bar{c}_{qq}, & c_{qq+1} & \ldots & c_{q,n} \\ c_{q+1,1}, & c_{q+1,2}, & c_{q+1,3} & \ldots & c_{q+1,q}, & c_{q+1,q+1} & \ldots & c_{q+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n,1}, & c_{n,2}, & c_{n,3} & \ldots & c_{n,q}, & c_{n,q+1} & \ldots & c_{nn} \end{vmatrix} \quad \text{(xxxvii)}$$

be formed, its constituents and not the linear terms in the exponential of (xxxv.) determine all the standard deviations and correlations. Let Σ_u be the variation after selection of the u^{th} organ; then if u be one of the selected organs $\Sigma_u = s_u$, if u be for one of the unselected organs Σ_u has still to be found. Let $\mathfrak{r}_{uv}$ be the correlation coefficient of the u^{th} and v^{th} organs after selection, then if u and v be selected organs, $\mathrm{r}_{uv} = \rho_{uv}$; if they be not both selected organs, then it has to be found. Let D_{uv} be the minor corresponding to the constituent c_{uv} in the above determinant, then by (x) and (xiv.)

$$\Sigma_u^{\,2} = \frac{D_{uu}}{\Delta} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xxxviii.}),$$

and

$$\Sigma_u \Sigma_v \mathrm{r}_{uv} = \frac{D_{uv}}{\Delta} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xxxix.}).$$

Now let us write for brevity $\alpha_{uu} = D_{uu}/\Delta$ and $\alpha_{uv} = D_{uv}/\Delta$. Clearly as long as u and v are less than $q + 1$, α_{uu} and x_{uv} will both be known, *i.e.*, are equal to $s_u^{\,2}$ and $s_u s_v \rho_{uv}$.

Now by a well-known property of determinants

$$c_{1,q+1}\alpha_{11} + c_{2,q+1}\alpha_{12} + c_{3,q+1}\alpha_{13} + \ldots + c_{q,q+1}\alpha_{1q} + c_{q+1,q+1}\alpha_{1,q+1} + \ldots + c_{n,q+1}\alpha_{1n} = 0$$

$$c_{1,q+2}\alpha_{11} + c_{2,q+2}\alpha_{12} + c_{3,q+2}\alpha_{13} + \ldots + c_{q,q+2}\alpha_{1q} + c_{q+1,q+2}\alpha_{1,q+2} + \ldots + c_{n,q+2}\alpha_{1n} = 0$$

. . . .

. . . .

$$c_{1,n}\alpha_{11} + c_{2,n}\alpha_{12} + c_{3,n}\alpha_{13} + \ldots + c_{qn}\alpha_{1q} + c_{q+1,n}\alpha_{1,q+2} + \ldots + c_{nn}\alpha_{1n} = 0 \quad . \quad . \quad (\text{xl.}).$$

Comparing these equations for the $n - q$ unknowns $\alpha_{1,q+1}$, $\alpha_{1,q+2} \ldots \alpha_{1n}$ with (xxxiv.) for finding x'_{q+1}, $x'_{q+2} \ldots x'_p$, we see that they are absolutely identical if we change $h_1, h_2 \ldots h_q$ in the latter into $\alpha_{11}, \alpha_{12} \ldots \alpha_{1q}$. Accordingly the solution is given by (xxxvi.), or we have

$$\alpha_{1,q+1} = -\left\{\frac{R(q+1)_{1,q+1}}{R(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_1}\alpha_{11} + \frac{R(q+1)_{2,q+1}}{R(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_2}\alpha_{12} + \ldots + \frac{R(q+1)_{q,q+1}}{R(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_q}\alpha_{1q}\right\}$$

$$\alpha_{1,q+2} = -\left\{\frac{R(q+2)_{1,q+2}}{R(q+2)_{q+2,q+2}}\frac{\sigma_{q+2}}{\sigma_1}\alpha_{11} + \frac{R(q+2)_{2,q+2}}{R(q+2)_{q+2,q+2}}\frac{\sigma_{q+2}}{\sigma_2}\alpha_{12} + \ldots + \frac{R(q+2)_{q,q+2}}{R(q+2)_{q+2,q+2}}\frac{\sigma_{q+2}}{\sigma_q}\alpha_{1q}\right\}$$

. . . .

$$\alpha_{1,n} = -\left\{\frac{R(n)_{1,n}}{R(n)_{n,n}}\frac{\sigma_n}{\sigma_1}\alpha_{11} + \frac{R(n)_{2,n}}{R(n)_{n,n}}\frac{\sigma_n}{\sigma_2}\alpha_{12} + \ldots + \frac{R(n)_{q,n}}{R(n)_{n,n}}\frac{\sigma_n}{\sigma_q}\alpha_{1q}\right\}$$

Provided v be $< q + 1$ we might equally well have used $\alpha_{v1}, \alpha_{v2}, \ldots \alpha_{vq}$. Hence we conclude that if v be $< q + 1$ and u be $> q$,

$$\alpha_{vu} = -\left\{\frac{R(u)_{1u}}{R(u)_{uu}}\frac{\sigma_u}{\sigma_1}\alpha_{v1} + \frac{R(u)_{2u}}{R(u)_{uu}}\frac{\sigma_u}{\sigma_2}\alpha_{v2} + \ldots + \frac{R(u)_{qu}}{R(u)_{uu}}\frac{\sigma_u}{\sigma_q}\alpha_{vq}\right\} \quad . \quad . \quad (\text{xli.}).$$

Or, substituting the known values of $\alpha_{v1}, \alpha_{v2}, \ldots \alpha_{vq}$, we have for $v < q + 1$ and $u > q$

$$\Sigma_v \Sigma_u r_{vu} = - s_v \sigma_u \left\{ \frac{\mathrm{R}\,(u)_{1u}}{\mathrm{R}\,(u)_{uu}} \rho_{1v} \frac{s_1}{\sigma_1} + \frac{\mathrm{R}\,(u)_{2u}}{\mathrm{R}\,(u)_{uu}} \rho_{2v} \frac{s_2}{\sigma_2} + \ldots + \frac{\mathrm{R}\,(u)_{vu}}{\mathrm{R}\,(u)_{uu}} \frac{s_v}{\sigma_v} + \ldots \right.$$
$$\left. + \frac{\mathrm{R}\,(u)_{qu}}{\mathrm{R}\,(u)_{uu}} \rho_{qv} \frac{s_q}{\sigma_q} \right\} \quad . \quad . \quad . \quad . \quad (\text{xli.})^{\text{bis}}.$$

The next stage in our work is to find α_{vu} when both v and u are $> q$, and also α_{uu} when $v = u$ and both are $> q$. This is done at once by substituting the minors α_{u1}, $\alpha_{u2}, \ldots \alpha_{uq} \ldots \alpha_{uu}$ in the equations formed from the last p—q lines of the determinant (xxxvii.).

We obtain the following system:

$$c_{1,q+1}\alpha_{u1} + c_{2,q+1}\alpha_{u2} + c_{3,q+1}\alpha_{u3} + \ldots + c_{q,q+1}\alpha_{uq} + c_{q+1,q+1}\alpha_{u,q+1} + \ldots + c_{n,q+1}\alpha_{un} = 0$$
$$c_{1,q+2}\alpha_{u1} + c_{2,q+2}\alpha_{u2} + c_{3,q+2}\alpha_{u3} + \ldots + c_{q,q+2}\alpha_{uq} + c_{q+1,q+2}\alpha_{u,q+1} + \ldots + c_{n,q+2}\alpha_{un} = 0$$
$$. \quad .$$
$$. \quad .$$
$$c_{1,u}\alpha_{u1} + c_{2,u}\alpha_{u2} + c_{3,u}\alpha_{u3} + \ldots + c_{q,u}\alpha_{uq} + c_{q+1,u}\alpha_{u,q+1} + \ldots + c_{n,u}\alpha_{un} = 1$$
$$. \quad .$$
$$. \quad .$$
$$c_{1,n}\alpha_{u1} + c_{2n}\alpha_{u2} + c_{3n}\alpha_{u3} + \ldots + c_{q,n}\alpha_{uq} + c_{q+1,n}\alpha_{u,q+1} + \ldots + c_{n,n}\alpha_{un} = 0.$$

These are identical with equations (xl.), except that the equation with the coefficients c_{1u}, $c_{2u}, \ldots c_{nu}$ has unity instead of zero on the right-hand side. Hence we see that α_{uv} $(v > q)$ will be the same function of α_{u1}, $\alpha_{u2}, \ldots \alpha_{uq}$ that x'_v is of h_1, $h_2, \ldots h_q$ in equation (xxxvi.), but it will add to this a function of the last p—q system of c's, *i.e.*, the c's

$$\begin{array}{llcl} c_{q,q+1}, & c_{q+1,q+1}, & . \quad . \quad . & c_{n,q+1}\,, \\ c_{q,q+1}, & c_{q+1,q+2}, & . \quad . \quad . & c_{n,q+2} \\ . \quad . \quad . & . \quad . \quad . & . \quad . \quad . & . \quad . \\ c_{q\ n}, & c_{q+1,n}, & . \quad . \quad . & c_{n,n}. \end{array}$$

Whatever this function may be we will represent it for the time by γ_{uv}; we notice that it is independent of the selected variations s_1, $s_2, \ldots s_q$, the selected means h_1, $h_2, \ldots h_q$, and the selected correlation coefficients ρ_{12}, $\rho_{23}, \ldots \rho_{q-1,q}$. It depends only on the characters *before* selection.

We thus have

$$\alpha_{uv} = \gamma_{uv} - \left\{ \frac{\mathrm{R}\,(v)_{1v}}{\mathrm{R}\,(v)_{vv}} \frac{\sigma_r}{\sigma_1} \alpha_{u1} + \frac{\mathrm{R}\,(v)_{2v}}{\mathrm{R}\,(v)_{vv}} \frac{\sigma_v}{\sigma_2} \alpha_{u2} + \ldots + \frac{\mathrm{R}\,(v)_{qv}}{\mathrm{R}\,(v)_{vv}} \alpha_{uq} \right\}.$$

Now the system α_{u1}, $\alpha_{u2}, \ldots \alpha_{uq}$ can be found from (xli.), since $\alpha_{u'u''} = \alpha_{u''u'}$ whatever u' and u'' be.

Hence we have

$$\alpha_{uv} = \gamma_{uv} + \sigma_v\sigma_u\left[\mathrm{S}_1\left\{\frac{\mathrm{R}(v)_{pv}\,\mathrm{R}(u)_{pv}}{\mathrm{R}(v)_{vv}\,\mathrm{R}(u)_{uu}}\left(\frac{s_p}{\sigma_p}\right)^2\right\}\right.$$
$$\left.+\,\mathrm{S}_2\left\{\left(\frac{\mathrm{R}(v)_{p'v}\,\mathrm{R}(u)_{p''u}}{\mathrm{R}(v)_{vv}\,\mathrm{R}(u)_{uu}} + \frac{\mathrm{R}(v)_{p''v}\,\mathrm{R}(u)_{p'u}}{\mathrm{R}(v)_{vv}\,\mathrm{R}(u)_{uu}}\right)\rho_{p'p''}\frac{s_{p'}s_{p''}}{\sigma_{p'}\sigma_{p''}}\right\}\right].$$

Here S_1 denotes a sum from $p = 1$ to $p = q$, and S_2 a sum for every pair of values of p' and p'' out of $1, 2, 3, \ldots q$.

When $u = v$ we have simply

$$\alpha_{vv} = \gamma_{vv} + \sigma_v^2\left[\mathrm{S}_1\left\{\left(\frac{\mathrm{R}(v)_{pv}}{\mathrm{R}(v)_{vv}}\right)^2\left(\frac{s_p}{\sigma_p}\right)^2 + 2\mathrm{S}_2\left(\frac{\mathrm{R}(v)_{p'v}\,\mathrm{R}(v)_{p''v}}{(\mathrm{R}(v)_{vv})^2}\,\rho_{p'p''}\frac{s_{p'}s_{p''}}{\sigma_{p'}\sigma_{p''}}\right)\right\}\right].$$

It only remains to determine γ_{uv} and γ_{vv}. This we can do by putting all the s's zero, or selecting our q organs of one size only. We see at once that γ_{uv} and γ_{vv} are the values of α_{uv} and α_{vv}, that is, of $\Sigma_u\Sigma_v\mathfrak{r}_{vu}$, and Σ_v^2, when we select q organs of definite values and seek the correlation and the variabilities of two others, the u^{th} and the v^{th}. These values have already been found on p. 10. Or:

$$\left.\begin{aligned}\Sigma_u\Sigma_v\mathfrak{r}_{vu} &= \gamma_{uv} = -\,\sigma_u\sigma_v\mathrm{R}(uv)_{uv}/\mathrm{R}(q),\\ \Sigma_v^2 &= \gamma_v^2 = \sigma_v^2\mathrm{R}(uv)_{uu}/\mathrm{R}(q),\\ \Sigma_u^2 &= \gamma_u^2 = \sigma_u^2\mathrm{R}(uv)_{vv}/\mathrm{R}(q),\end{aligned}\right\} \quad \ldots\ldots \quad \text{(xlii.)}$$

The notation of that page has been changed so that $\mathrm{R}(uv)$ now stands for the determinant

$$\begin{vmatrix} 1, & r_{12}, & r_{13}, & \ldots & r_{1q}, & r_{1u}, & r_{1v} \\ r_{21}, & 1, & r_{23}, & \ldots & r_{2q}, & r_{2u}, & r_{2v} \\ r_{31}, & r_{32}, & 1, & \ldots & r_{3q}, & r_{3u}, & r_{3v} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ r_{q1}, & r_{q2}, & r_{q3}, & \ldots & 1, & r_{qu}, & r_{qv} \\ r_{u1}, & r_{u2}, & r_{u3}, & \ldots & r_{uq}, & 1, & r_{uv} \\ r_{v1}, & r_{v2}, & r_{v3}, & \ldots & r_{vq}, & r_{vu}, & 1 \end{vmatrix} \quad \ldots\ldots \quad \text{(xliii.)}$$

$\mathrm{R}(uv)_{uv}$ is the minor corresponding to the constituent r_{uv}; $\mathrm{R}(uv)_{uu}$, the minor corresponding to the constituent at the meet of the u^{th} column and u^{th} row; and $\mathrm{R}(q)$ the determinant with the last two rows and two columns struck out.

For example:

$$q = 1: \quad \gamma_{vv} = \sigma_v^2 (1 - r_{1v}^2) \qquad \gamma_{uu} = \sigma_u^2 (1 - r_{1u}^2),$$

$$\gamma_{uv} = \sigma_v \sigma_u (r_{vu} - r_{1v} r_{1u}),$$

$$q = 2: \quad \gamma_{vv} = \sigma_v^2 (1 - r_{12}^2 - r_{1v}^2 - r_{2v}^2 + 2 r_{12} r_{1v} r_{2v}) / (1 - r_{12}^2),$$

$$\gamma_{uu} = \sigma_u^2 (1 - r_{12}^2 - r_{1u}^2 - r_{2u}^2 + 2 r_{12} r_{1u} r_{2u}) / (1 - r_{12}^2),$$

$$\gamma_{uv} = \sigma_u \sigma_v \{ r_{vu} (1 - r_{12}^2) - r_{1v} r_{1u} - r_{2v} r_{2u} + r_{12} (r_{1u} r_{2v} + r_{2u} r_{1v}) \} / (1 - r_{12}^2),$$

$$q = 3:^* \quad \gamma_{vv} = \sigma_v^2 \times \begin{vmatrix} 1, & r_{12}, & r_{13}, & r_{1v} \\ r_{21}, & 1, & r_{23}, & r_{2v} \\ r_{31}, & r_{32}, & 1, & r_{3v} \\ r_{v1}, & r_{v2}, & r_{v3}, & 1 \end{vmatrix} \div (1 - r_{12}^2 - r_{23}^2 - r_{31}^2 + 2 r_{12} r_{23} r_{31})$$

$$\gamma_{uu} = \sigma_u^2 \times \begin{vmatrix} 1, & r_{12}, & r_{13}, & r_{1u} \\ r_{21}, & 1, & r_{23}, & r_{2u} \\ r_{31}, & r_{32}, & 1, & r_{3u} \\ r_{u1}, & r_{u2}, & r_{u3}, & 1 \end{vmatrix} \div (1 - r_{12}^2 - r_{23}^2 - r_{31}^2 + 2 r_{12} r_{23} r_{31})$$

$$\gamma_{uv} = \sigma_v \sigma_u \begin{vmatrix} 1, & r_{12}, & r_{13}, & r_{1v} \\ r_{21}, & 1, & r_{23}, & r_{2v} \\ r_{31}, & r_{32}, & 1, & r_{3v} \\ r_{u1}, & r_{u2}, & r_{u3}, & r_{uv} \end{vmatrix} \div (1 - r_{12}^2 - r_{23}^2 - r_{31}^2 + 2 r_{12} r_{23} r_{31}) \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xliv.)}.$$

We can now collect our complete results.

The variability of a non-selected organ $v > q$ is after the selection of q organs given by

$$\Sigma_v^2 = \sigma_v^2 \left\{ \frac{\mathrm{R}(uv)_{uu}}{\mathrm{R}(q)} + \mathrm{S}_1 \left[\left(\frac{\mathrm{R}(v)_{pv}}{\mathrm{R}(v)_{vv}} \right)^2 \left(\frac{s_p}{\sigma_p} \right)^2 \right] + 2\mathrm{S}_2 \left[\frac{\mathrm{R}(v)_{p'v}\, \mathrm{R}(v)_{p''v}}{(\mathrm{R}(v)_{vv})^2}\, \rho_{p'p''} \frac{s_{p'} s_{p''}}{\sigma_{p'} \sigma_{p''}} \right] \right\}. \quad \text{(xlv.)}$$

The correlation of two non-selected organs v and u both $> q$ is after the selection of q organs given by

$$\Sigma_u \Sigma_v \mathrm{r}_{uv} = \sigma_u \sigma_v \left[- \frac{\mathrm{R}(uv)_{uv}}{\mathrm{R}(q)} + \mathrm{S}_1 \left\{ \frac{\mathrm{R}(v)_{pv}\, \mathrm{R}(u)_{pv}}{\mathrm{R}(v)_{vv}\, \mathrm{R}(u)_{uu}} \left(\frac{s_p}{\sigma_p} \right)^2 \right\} \right.$$

$$\left. + \mathrm{S}_2 \left\{ \left(\frac{\mathrm{R}(v)_{p'v}\, \mathrm{R}(u)_{p''u}}{\mathrm{R}(v)_{vv}\, \mathrm{R}(u)_{uu}} + \frac{\mathrm{R}(v)_{p''v}\, \mathrm{R}(u)_{p'u}}{\mathrm{R}(v)_{vv}\, \mathrm{R}(u)_{uu}} \right) \rho_{p'p''} \frac{s_{p'} s_{p''}}{\sigma_p \sigma_{p''}} \right\} \right] \quad . \quad . \quad \text{(xlvi.)}.$$

The correlation between a non-selected organ $u > q$ and a selected organ $v < q + 1$ is given by (xli.)$^{\text{bis}}$ or,

$$s_v \Sigma_u \mathrm{r}_{vu} = - s_v \sigma_u \left[\mathrm{S}_1 \left(\frac{\mathrm{R}(u)_{pu}}{\mathrm{R}(u)_{uu}}\, \rho_{pv} \frac{s_p}{\sigma_p} \right) \right] \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xlvii.)}.$$

---

[*] The expanded values of these determinants are given, 'Phil. Trans.,' A, vol. 187, p. 294.

Here in S_1 : p takes every value from 1 to q, and in S_2 : p' and p'' every possible pair of values from 1 to q. Equations (xlv.)–(xlvii.) fully determine all the required quantities and form the full solution of the problem of selection. Before we see the remarkably simple forms they take for simpler cases, we may draw some general conclusions of a most important character.

In the first place we must distinguish between directly selected and what we have termed non-selected organs. It would be better to term the latter *indirectly* selected organs. Suppose the recruiting sergeant were to pay attention only to stature and seek to form a regiment of men of *about* 5 feet 10 inches. He might have a real range of stature about 6 or 7 inches, but he would strive to get men of *about* this height from the population. We will suppose that he did not consider chest-breadth, head-length, foot-length, lungs or any other character. The distribution of these "non-selected" characters in the regiment would not be the same as in the general population. Their means would have changed by (xxxvi.) and their variabilities and correlations be given by (xlv.)–(xlvii.). In other words, an indirect selection would have taken place. A selection by stature would change foot-length and head-length and indeed every other correlated organ. Much the same result must occur in natural selection. If it be advantageous for a species to have a certain group of its organs of definite size, falling within a definite range, and related to each other in a definite manner, then these changes cannot take place without modifying not only the size, but the variability and correlation of all the other organs correlated with these, although these organs themselves be not directly selected. Practically this means all the other organs, for so far one can hardly say with certainty that we have come across any two characters in an organism which are uncorrelated. Many of those investigated are highly correlated, all appear to have some correlation, even if it be very small or negative.

We may therefore conclude as follows:—

(*a*) The selection of any complex of characters or organs in an organism changes all the other characters and organs not directly selected.

(*b*) If the change in the complex be continuous and progressive, the other characters will continue to be modified until the change in them is so considerable that selection begins to act directly upon them also.

(*c*) The changes noted here are not confined to the average value of a non-directly selected character and to its variability; the correlations between non-directly selected characters and the correlations between directly and non-directly selected characters are also both changed.

(*d*.) If local races have been produced by selection from a common stock, it will be impossible to look upon correlation as a criterion for species. Every selection will modify such correlation, and it has no greater fixity than either type value (mean) or variability (standard deviation).

The whole of these statements will become more manifest as we apply our general theorem to special cases, but we must note that if two organs were uncorrelated with each other, it still might be possible by selecting a third, or a third and fourth, to produce correlation between them. Further, by selection of one or more organs, two non-directly selected organs can have their existing correlation increased, lessened or even changed in sign.

(4.) A primary difficulty will of course arise in the case of natural selection. How are we to determine which are the directly and which are the indirectly selected organs? With artificial selection by man, we know which organs have been selected fairly well; attention has been paid to colour, size, proportion of parts, &c. Even in the case of the medical examination of the recruit, it is chest, lungs, heart, stature, &c., which form the basis of the acceptance or rejection. If the head or foot be not absolutely deformed, little if any attention is paid to them, so with hair-colour, probably eye-colour, and a mass of other details. No doubt the direct medical selection indirectly selects these, but we could roughly class the selected and non-selected organs or characters and investigate the changes in the correlations of the latter owing to the indirect selection. But how are we to form these classes in the case of natural selection?

The investigations may look difficult, and even from the standpoint of arithmetic appalling, but it seems to me that the differentiation of organs into directly and indirectly selected classes is the keynote to the problem of evolution by natural selection.

Let us look at a simple case and see whether it will throw any light on the problem of distinguishing between directly and indirectly selected organs. Suppose we have two organs only, with means m_1, m_2, standard deviations σ_1, σ_2, correlation r_{12}, and let the first be selected so as to have a mean value $m_1 + h_1$, and standard deviation s_1. Let Σ_2 be the standard deviation of the second organ and $\mathbf{r}_{12}$ the correlation of the two organs after selection, and $m_2 + x'_2$ the mean of the non-selected organ.

Then by (xxxii.):

$$x'_2 = r_{12}\frac{\sigma_2}{\sigma_1}h_1,$$

and it will be shown later (see p. 23) that

$$\Sigma_2^2 = \sigma_2^2\left(1 - \left(1 - \frac{s_1^2}{\sigma_1^2}\right)r_{12}^2\right),$$

and

$$\Sigma_2\mathbf{r}_{12} = \sigma_2\frac{s_1}{\sigma_1}r_{12}.$$

Hence we have:

$$\frac{\mathbf{r}_{12}\Sigma_2}{s_1} = \frac{r_{12}\sigma_2}{\sigma_1},$$

$$\frac{\mathbf{r}_{12}s_1}{\Sigma_2} = \frac{s_1^2}{\sigma_1\sigma_2}\frac{r_{12}}{1 - \left(1 - \frac{s_1^2}{\sigma_1^2}\right)r_{12}^2},$$

In other words the regression coefficient of the non-selected organ on the selected remains unchanged, while that of the selected organ on the non-selected will, as a rule, be widely modified.

Further, let X_2 be the mean value of the second organ before selection corresponding to a value H_1 of the first; let M_1 and M_2 be the means of the organs after selection, and Y_2 be the mean value of the second organ corresponding to a value K_1 of the first. Then the equation to the regression line before selection is

$$X_2 = r_{12}\frac{\sigma_2}{\sigma_1}H_1 + m_2 - r_1\frac{\sigma_2}{\sigma_1}m_1,$$

and after selection it is

$$\begin{aligned}
Y_2 &= \mathrm{r}_{12}\frac{\Sigma_2}{s_1}K_1 + M_2 - \mathrm{r}_{12}\frac{\Sigma_2}{s_1}M_1,\\
&= r_{12}\frac{\sigma_2}{\sigma_1}K_1 + m_2 + r_{12}\frac{\sigma_2}{\sigma_1}h_1 - r_{12}\frac{\sigma_2}{\sigma_1}(m_1 + h_1),\\
&= r_{12}\frac{\sigma_2}{\sigma_1}K_1 + m_2 - r_{12}\frac{\sigma_2}{\sigma_1}m_1.
\end{aligned}$$

But this is identically the same line as the regression line before selection. Hence not only the slope (regression coefficient) of the line, but its position is identical, and we have the following result:—

*If two local races have been evolved from a single stock by the selection in different ways of one organ only, then the regression lines for the two races of any non-directly selected organ on the directly selected organ will be the same in direction and position; but the regression lines of the selected organ on any non-selected organ will differ for the two races.**

Of course the means, standard deviations and correlations, not only of the selected organ but of all the non-selected organs also, will probably have changed. It is only certain of the regression lines which remain unchanged and serve as a criterion to enable us to distinguish between directly and non-directly selected organs.

Of course the problem in Nature will not be as simple as this, for differentiation of the two local races may have arisen from the selection of more than one organ, or may have arisen from the selection of two different organs, but the illustration will, I think, indicate the nature of the investigation we are proposing.

We can easily generalise our theorem by considering the form of the selection surface given on p. 12. Any result obtained from (xxxv.) which does not involve any of the $\bar{c}$'s will be a result unaffected by the selection that has gone on. Now to obtain a regression equation we put any number of the x's equal to constants, to h's

---

* The geometrical interpretation in this simple case that the regression line is unchanged is quite obvious, and, indeed, may serve as a proof.

say, and find the "centre" of the quadric of the remaining x's, the co-ordinates of this centre, expressed in terms of the h's, are the regression equations. Now it will be clear, that if we put all the selected x's equal to h's, the differentials of the quadric with regard to the remaining or non-selected x's can contain no $\bar{c}$'s or the coefficients of the regression equations thus found will not be modified by selection.

Further, we might have given not only the selected organs, but any number of the non-selected organs constant values, and the resulting regression equations would involve only the c's and not the $\bar{c}$'s.

Hence we have the following general theorems :—

(i.) *If an organ has been modified only by indirect selection, then its partial regression coefficients on any complex of other organs, however large or small, provided it includes all the directly selected organs, will remain unchanged by the selection.*

(ii.) *The same organ in two different local races which have been derived from a common stock by the selection of two complexes of organs, some of which may or may not be common, will, if it has not been directly selected in either case, give the same partial regression coefficients for any group of organs which includes the members of both complexes and any number of non-directly selected organs besides.*

If the partial regression equations have changed coefficients, then we cannot at once determine whether—

(a) We are dealing with a non-directly selected organ, and have not included *all* the directly selected organs in the group upon which we are calculating the regression ; or

(b) We are really dealing with a directly selected organ. In this case, we have also certainly not included at least one directly selected organ in the regression group.

Theoretically, however, (i.) and (ii.) suffice to find out which, if any, are the non-directly selected organs in the differentiation of local races. Practically, however, the number of organs and characters may be so great, and our ignorance of those probably selected so complete, that the arithmetic of determining so extensive a series of partial regression coefficients may be quite beyond our powers. Still, where the divergence between local races is not too great, and the source of the differentiation not too obscure, it is probable that the above theorems will lead to results of great interest.*

Without laying too great weight on these theorems, I would still venture to suggest that if the criterion of a species be the discovery of any numerical constant

---

* Mr. L. BRAMLEY-MOORE has been working with this end in view at the long-bones in man. But even here the direct selection of parts of the vertebral column—for which, at present, we have no correlation values either among themselves or with the long-bones—and of the hand and foot, which Dr. W. R. MACDONELL has just shown, are very highly correlated with the long-bones, may render nugatory all attempt to ascertain which, if any, long-bone has been only indirectly, or, at any rate, least directly selected.

or group of constants, which is the same for all local races, then these constants must not be sought in the values of mean characters, degrees of variability, or of correlation, but in a system of partial regression coefficients, and the discovery of these is therefore of first class biological importance; it is the classification of the characters into directly and non-directly selected groups, *i.e.*, it is the discovery of the *modus operandi* of the factors by which the differentiation has taken place. We are a long way from solution yet, but we may venture, perhaps, to admit a faint glimmer of light in the direction of what might seem the culminating problem of the mathematical method as applied to evolution—the piecing together by quantitative analysis of the stages of descent.

(5.) I will take now the application of the above results to simple cases; but for the benefit of those who cannot easily follow the main principles of our investigation through the stages of determinant analysis, I will prove directly the proposition that: *the selection of an organ* A *alters the mean and variability of a correlated organ* B, *and also the correlation between* A *and* B.

Let the correlation surface for the two organs be

$$z = \frac{N}{2\pi\sigma_1\sigma_2}\,\frac{1}{\sqrt{1-r_{12}^2}}\,e^{-\frac{1}{2}\left(\frac{x_1^2}{\sigma_1^2(1-r_{12}^2)} - \frac{2r_{12}x_1x_2}{\sigma_1\sigma_2(1-r_{12}^2)} + \frac{x_2^2}{\sigma_2^2\,(1-r_{12}^2)}\right)},$$

where N is the number of individuals in the general population before selection, and the subscripts 1 and 2 refer to the organs A and B respectively.

Let the distribution of the population after selection of the A organ be

$$z' = \frac{n}{\sqrt{2\pi}\,s_1}\,e^{-\frac{(x_1-h_1)^2}{2s_1^2}},$$

where N − n is the total destruction, h_1 the mean and s_1 the variability of the population with regard to A after selection. Before selection this distribution was

$$z_1 = \frac{N}{\sqrt{2\pi}\sigma_1}\,e^{-\frac{x_1^2}{2\sigma_1^2}}.$$

Hence, the selection being random with regard to the array of B's corresponding to any A, we have for the surface after selection

$$Z = z \times z'/z_1,$$

for each array must be altered in the ratio of the corresponding z' to z_1.

This gives for the surface in full

$$Z = \frac{n}{2\pi\sigma_2 s_1}\,\text{expt.} - \tfrac{1}{2}\left\{x_1^2\left(\frac{1}{s_1^2} - \frac{1}{\sigma_1^2} + \frac{1}{\sigma_1^2\,(1-r_{12}^2)}\right) - \frac{2r_{12}x_1x_2}{\sigma_1\sigma_2\,(1-r_{12}^2)} \right.$$
$$\left. + \, x_2^2\,\frac{1}{\sigma_2^2\,(1-r_{12}^2)} - \frac{2h_1x_1}{s_1^2}\right\} \quad . \quad . \quad . \quad \text{(xlviii.)}.$$

Integrate this for every possible value of x_1 and we shall have the distribution h_2 and s_2 of x_2 or B after the selection of A. After some reductions we find for the frequency $\zeta \delta x_2$

$$\zeta = \frac{n}{\sqrt{2\pi}\Sigma_2} \cdot e^{-\frac{1}{2}\frac{1}{\Sigma_2^2}\left(x_2 - \frac{h_1 r_{12}\sigma_2}{\sigma_1}\right)^2},$$

where

$$\Sigma_2^2 = \sigma_2^2\left(1 - \left(1 - \frac{s_1^2}{\sigma_1^2}\right) r_{12}^2\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xlix.)},$$

which gives the standard deviation of the indirectly selected organ, and the "centre" of this organ is given by

$$h_2 = \frac{r_{12}\sigma_2}{\sigma_1} h_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(l.)}.$$

These are special cases of our results (xlv.) and (xxxvi.) above respectively.

Further, returning to the correlation surface (xlviii.), the coefficient of correlation r_{12} is the coefficient of $x_1 x_2$ divided by the product of the coefficients of x_1^2 and x_2^2. Hence we find

$$\mathrm{r}_{12} = \frac{s_1}{\sigma_1} \frac{r_{12}}{\sqrt{1 - \left(1 - \frac{s_1^2}{\sigma_1^2}\right) r_{12}^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(li.)}.$$

Let $r_{12} = \sin\theta_{12}$, $\mathrm{r}_{12} = \sin\phi_{12}$, then we have

$$\tan\phi_{12} = \frac{s_1}{\sigma_1}\tan\theta_{12} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lii.)}.$$

This shows us that ϕ_{12} decreases with s_1, or that r_{12} decreases with s_1, that is to say, the more intense the selection the less is the correlation. This in broad terms demonstrates the general principle that intensity of selection connotes a lessening of correlation. It is this principle which very possibly accounts for the fact that the more civilized races of man appear to be not only more variable but more highly correlated than the less civilized, among whom the struggle for existence is more intense. It may, perhaps, also account for the skeletons of women of the civilized races having their parts more highly correlated than the parts of those of men.* Lastly, it may well throw some light on the markedly plastic character of races which have been stringently selected with regard to one or a few organs only.

As an illustration, suppose that the correlation between femur and tibia were ·7, and let us investigate what would be the effect of reducing the variability of the tibia by direct selection 50 per cent. We find at once on substituting $s_1/\sigma_1 = \cdot 5$ and $r_{12} = \cdot 7$ in (li.) above that $\mathrm{r}_{12} = \cdot 44$, or a reduction of about 37 per cent. This will, perhaps, be sufficient to indicate what immense changes must be made in the correlation of highly correlated organs whenever selection, artificial or natural, is stringent. It is important to notice that the change in the size of the organ in no

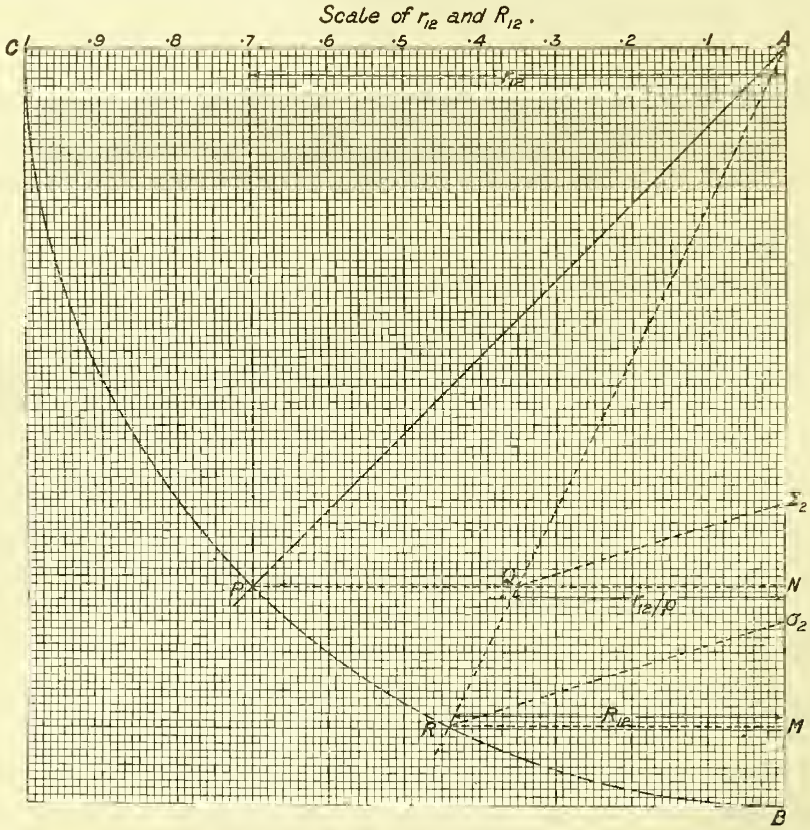
---

* See LEE and PEARSON, 'Roy. Soc. Proc.,' vol. 61, p. 354; and LEE, 'Phil. Trans.,' A, vol. 196, p. 231.

way influences the change in the correlation between organs, if the distribution be normal,* the change depends only on the stringency of the selection. Breeders who select by the size of an organ only are in that case very likely to reduce the variability of the organ in the selected group by far more than 50 per cent. Accordingly, it is not to be wondered at if they, to a great extent, destroy the correlation between the selected organ and other organs. This destruction would appear as a want of relationship between parts, possibly as a want of fixity in type.

By means of equation (lii.) r_{12} can easily be found from r_{12} and $p = s_1/\sigma_1$. All we need is a table of trigonometrical functions. We observe that r_{12} is always of the same sign and less than r_{12}. For many biological purposes the following graphical construction gives quite sufficiently accurate results. Let CAB be a quadrant, say of 10 centims. radius, and take the point P on this quadrant distant PN $= 10r_{12}$ from AB. Take QN $= \dfrac{1}{p}$ PN, and let AQ meet the quadrant in R, then RM the distance of R from AB $= 10\mathrm{r}_{12}$, and consequently determines r_{12}. If the figure be drawn on decimal paper the determination of r_{12} is peculiarly easy.

Graphical method of finding correlation between organs A and B after selection has acted on A.



In the above example r_{12} = PN = ·7 ; $s_1/\sigma = \dfrac{1}{p} = \frac{1}{2}$, and therefore QN $= r_{12}/p$ = ·35.

$\mathrm{r}_{12} = \mathrm{R}_{12}$ = RM = ·44, as before. See p. 23.

---

* It will do so if the distribution be skew, see 'Phil. Trans.,' A, vol. 191, p. 231.

Further, taking AP as our unit, $AN^2 = 1 - r_{12}{}^2$, and $QN^2 = r_{12}{}^2 \times \frac{s_1{}^2}{\sigma_1{}^2}$. Hence, from (lxix.):

$$\Sigma_2{}^2 = \sigma_2{}^2\,(AN^2 + QN^2) = \sigma_2{}^2 AQ^2.$$

Therefore if $A\sigma_2$ in the diagram be taken equal to σ_2 and $Q\Sigma_2$ be drawn parallel to $R\sigma_2$, we shall have $A\Sigma_2 = \Sigma_2$, or we can scale off the reduced variability.

Thus the diagram enables us to see at a glance the reduction in correlation and variability.

(6.) Let us now write down the results when an organ A is selected out of a group of three organs, A, B, C, whose constants are marked by the subscripts 1, 2, 3, respectively. Let $\mu_1 = s_1/\sigma_1$, and be represented, when required, by $\cos\chi_1$. Then we find from (xlv.) (xlvii.):

$$x'_2 = r_{12}\frac{\sigma_2}{\sigma_1}h_1, \qquad x'_3 = r_{13}\frac{\sigma_3}{\sigma_1}h_1 \quad . \; . \; . \; . \; . \; . \; . \; . \quad \text{(liii.)},$$

$$\left.\begin{aligned} \Sigma_2 &= \sigma_2\left\{1 - \left(1 - \left(\frac{s_1}{\sigma_1}\right)^2\right)r_{12}{}^2\right\}^{\frac{1}{2}} = \sigma_2\,\{1 - \sin^2\chi_1\cos^2\theta_{12}\}^{\frac{1}{2}} \\ \Sigma_3 &= \sigma_3\left\{1 - \left(1 - \left(\frac{s_1}{\sigma_1}\right)^2\right)r_{13}{}^2\right\}^{\frac{1}{2}} = \sigma_3\,\{1 - \sin^2\chi_1\cos^2\theta_{13}\}^{\frac{1}{2}} \end{aligned}\right\} \quad . \quad \text{(liv.)},$$

$$\left.\begin{aligned} \mathfrak{r}_{12} &= \frac{\mu_1 r_{12}}{\{1 - (1 - \mu_1{}^2)\,r_{12}{}^2\}^{\frac{1}{2}}} = \frac{\cos\chi_1\cos\theta_{12}}{\sqrt{1 - \sin^2\chi_1\cos^2\theta_{12}}} \\ \mathfrak{r}_{13} &= \frac{\mu_1 r_{13}}{\{1 - (1 - \mu_1{}^2)\,r_{13}{}^2\}^{\frac{1}{2}}} = \frac{\cos\chi_1\cos\theta_{13}}{\sqrt{1 - \sin^2\chi_1\cos^2\theta_{13}}} \end{aligned}\right\} \quad . \; . \; . \; . \; . \quad \text{(lv.)},$$

$$\left.\begin{aligned} \mathfrak{r}_{23} &= \frac{(r_{23} - r_{13}r_{12})\left(1 - \dfrac{s_1{}^2}{\sigma_1{}^2}\right) + \dfrac{s_1{}^2}{\sigma_1{}^2}\,r_{23}}{\sqrt{\left\{1 - \left(1 - \left(\dfrac{s_1}{\sigma_1}\right)^2\right)r_{12}{}^2\right\}\left\{1 - \left(1 - \left(\dfrac{s_1}{\sigma_1}\right)^2\right)r_{13}{}^2\right\}}} \\ &= \frac{\cos\theta_{23} + \sin^2\chi_1\cos\theta_{13}\cos\theta_{12}}{\sqrt{(1 - \sin^2\chi_1\cos^2\theta_{12})\,(1 - \sin^2\chi_1\cos^2\theta_{13})}} \end{aligned}\right\} \quad . \; . \quad \text{(lvi.)},$$

where, as before, we write $r_{pq} = \cos\theta_{pq}$. Let us also write $\mathfrak{r}_{pq} = \cos\Theta_{pq}$, and

$$\sin\chi_1\cos\theta_{12} = \cos a_{12}, \qquad \sin\chi_1\cos\theta_{13} = \cos a_{13}.$$

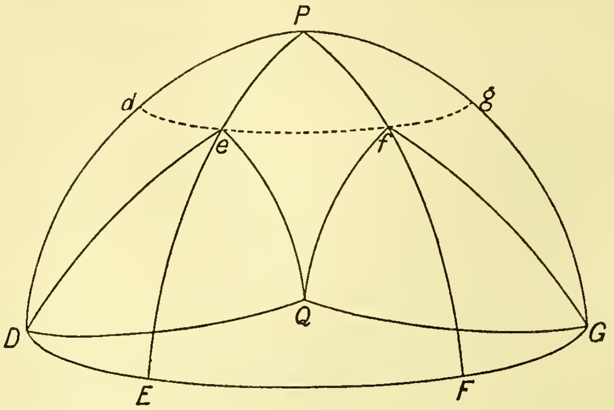
Then we can replace the above results by

$$\Sigma_2 = \sigma_2\sin a_{12}, \qquad \Sigma_3 = \sigma_3\sin a_{13},$$

$$\cos\Theta_{12} = \cot\chi_1\cot a_{12}, \qquad \cos\Theta_{13} = \cot\chi_1\cot a_{13},$$

$$\cos\Theta_{23} = \frac{\cos\theta_{23} - \cos a_{12}\cos a_{13}}{\sin a_{12}\sin a_{13}} \quad . \; . \; . \; . \; . \; . \; . \; . \; . \quad \text{(lvii.)}.$$

These equations admit of easy interpretation by spherical geometry.



Let P be the pole of the great circle DEFG. Take DG $= \theta_{23}$, DE $= \theta_{12}$, GF $= \theta_{13}$. Join P to E and F; let the small circle of radius χ_1 round P meet PD, PE, PF, and PG in d, e, f, g respectively. Draw the arcs De and Gf. Let the small circles with centres D and G and radii De and Gf respectively meet in Q. Join DQ and GQ. Then the quantities required are:

$$\Sigma_2/\sigma_2 = \sin \mathrm{DQ}, \qquad \Sigma_3/\sigma_3 = \sin \mathrm{GQ}.$$

$$\mathrm{r}_{12} = \cos \mathrm{D}e\mathrm{E}, \qquad \mathrm{r}_{13} = \cos \mathrm{F}f\mathrm{G}, \qquad \mathrm{r}_{23} = \cos \mathrm{DQG}.$$

For $$\mathrm{DE} = \theta_{12},\ \mathrm{E}e = \frac{\pi}{2} - \chi_1,\ \angle\ \mathrm{DE}e = \frac{\pi}{2};$$

hence: $$\cos \mathrm{D}e = \cos \theta_{12} \cos \mathrm{E}e = \sin \chi_1 \cos \theta_{12} = \cos \alpha_{12}, \text{ or } \alpha_{12} = \mathrm{D}e;$$

similarly $$\alpha_{13} = \mathrm{G}f.$$

Next, $\cos \mathrm{D}e\mathrm{E} = \cot \mathrm{D}e \tan e\mathrm{E} = \cot \alpha_{12} \cot \chi_1$, or $\angle\ \mathrm{D}e\mathrm{E} = \Theta_{12}$; similarly $< \mathrm{F}f\mathrm{G} = \Theta_{13}$. Lastly, from the triangle DQG: DQ $= \alpha_{12}$, QG $= \alpha_{13}$, and DG $= \theta_{23}$, but

$$\cos \mathrm{DG} = \cos \mathrm{DQ} \cos \mathrm{QG} + \sin \mathrm{DQ} \sin \mathrm{QG} \cos \mathrm{DQG},$$

or, $$\cos \mathrm{DQG} = \frac{\cos \theta_{23} - \cos \alpha_{12} \cos \alpha_{13}}{\sin \alpha_{12} \sin \alpha_{13}} = \cos \Theta_{23}; \text{ or } \angle\ \mathrm{DQG} = \Theta_{23}.$$

Thus all the relations can be expressed in terms of the sides and angles of a simple system of spherical triangles. For the degree of accuracy generally possible in biological and sociological investigations these triangles can be solved by a spherical trigonometer, such as that sold by KREIDL, of Prague.* The changes, however, which r_{23} undergoes for various values of r_{12}, r_{13}, r_{23} are, indeed, far more difficult to appreciate as a whole than those of r_{12} or r_{13}. In order that they may be followed easily, and in order to solve directly to a degree of approximation sufficient for many practical purposes problems in the influence of selection on correlation, my assistant, Dr. L. N. G. FILON, has kindly drawn up the tables which accompany this memoir.

---

* It will suffice fairly well for all but a few special values of r_{12}, r_{23}, r_{31}.

Of course to bring them within any reasonable compass we have had to limit the values taken. In the first place we have considered only eleven grades of selective stringency given by

$$s_1/\sigma_1 = 0,\ 1/10,\ 2/10,\ 3/10,\ 4/10,\ 5/10,\ 6/10,\ 7/10,\ 8/10,\ 9/10,\ 1,$$

the corresponding values of $\mathfrak{r}_{23}$ in the tables are entered as

$$R_0,\ R_1,\ R_2,\ R_3,\ R_4,\ R_5,\ R_6,\ R_7,\ R_8,\ R_9,\ R_{10}.$$

The tables are calculated for r_{23} both positive and negative, but r_{12} and r_{13} are always supposed *positive*. If r_{12} and r_{13} be both *negative*, then $\mathfrak{r}_{23}$ will be the same as if they were both positive. If r_{12} and r_{13} be of opposite signs, then all we have to do is to look out $\mathfrak{r}_{23}$ in the table in which r_{23} has a sign the reverse of its actual value, and having found the corresponding value of $\mathfrak{r}_{23}$, then change its sign to obtain the actual coefficient of correlation after selection. This follows, if r_{13} be the negative coefficient, by writing:

$$\mathfrak{r}_{23} = \frac{r_{23} + r_{13}r_{12}\sin^2\chi_1}{\sqrt{\{1 - r_{12}^2\sin^2\chi_1\}\{1 - r_{13}^2\sin^2\chi_1\}}} = -\frac{(-r_{23}) + (-r_{13})(r_{12})\sin^2\chi}{\sqrt{\{1 - r_{12}^2\sin^2\chi\}\{1 - (-r_{13})^2\sin^2\chi_1\}}}.$$

Lastly, it would clearly be very laborious to tabulate $\mathfrak{r}_{23}$ for a very great series of values of r_{12}, r_{13}, r_{23}. Accordingly a selection had to be made of these coefficients of correlation. They were given the values 0, ·25, ·5, ·75, and 1. These may be spoken of as zero, small, medium, large, and perfect correlations, and the ranges 0 to ·25, ·25 to ·5, ·5 to ·75, and ·75 to 1, as the ranges of little, moderate, considerable, and high correlation respectively. There would thus appear to be 15 combinations of values for r_{12}, r_{13}; these are given in the key to the tables as (*a*), (*b*), (*c*), (*d*)... (*m*), (*n*), (*p*), see p. 63. If these 15 values had to be combined with the 10 values (5 positive and 5 negative) of r_{23} and the 11 values of s_1/σ_1, we should have 1650 entries in our tables. But this number is much reduced by the consideration that the expression $1 - r_{23}^2 - r_{13}^2 - r_{12}^2 + 2r_{23}r_{13}r_{12}$ has for the real correlation of three characters to be always positive. $\mathfrak{r}_{23}$ can also never be greater than unity. Accordingly all values of r_{23}, r_{13}, r_{12}, which do not satisfy these conditions, have been excluded from the tables; they cannot arise in nature. A few impossible values of $\mathfrak{r}_{23}$ have been included in the tables, but these are placed there solely for the purpose of finding by interpolation values of $\mathfrak{r}_{23}$, which are less than unity. The following *purely hypothetical* illustrations of formula (lvi.) and the tables will serve to indicate their use.

Illustration I.—Suppose the correlation of tibia and femur with each other to be ·8, and of both with the stature to be ·6. How would their correlation be altered if the variation in stature were reduced by selection to half its present value?

Let $s_1/\sigma_1 = \mu_1$ as before, and suppose $\mathfrak{r}_{23} = R$; then let μ_1, r_{12}, r_{13}, r_{23} be the values of the constants next *below* the required values occurring in the tables, and giving

$r_{23} = R$; let R' be the true value of r_{23}, corresponding to the values $\mu_1 + \delta\mu_1$, $r_{12} + \delta r_{12}$, $r_{13} + \delta r_{13}$, $r_{23} + \delta r_{23}$. Thus we have, as far as first differences :

$$R' = R + 10\,(\Delta_{\mu_1}R)\,\delta\mu_1 + 4\,\{(\Delta_{r_{12}}R)\,\delta r_{12} + (\Delta_{r_{13}}R)\,\delta r_{13} + (\Delta_{r_{23}}R)\,\delta r_{23}\}.$$

In our case $\mu_1 = \frac{2}{10}$, $r_{12} = \cdot 5$, $r_{13} = \cdot 5$, $r_{23} = \cdot 75$, $\delta\mu_1 = \cdot 05$, $\delta r_{12} = \delta r_{13} = \cdot 1$, $\delta r_{23} = \cdot 05$. Further, we look up Table IV. (a), and the nearest case is (j) under R_2, which gives $R = \cdot 67105$. We then see that $(\Delta_{r_{12}}R) + (\Delta_{r_{13}}R)$ = difference between (j) and (m) cases $= - \cdot 21455$; $(\Delta_{r_{23}}R)$ = difference between result in IV. (a) and V. (a) $= \cdot 32895$; and lastly $(\Delta_{\mu_1}R)$ = difference between R_3 and R_2 columns of (j) row of Table IV. (a) $= \cdot 00535$. Thus we find :

$$R' = \cdot 671050 + \cdot 002675 - \cdot 04291 + \cdot 06578 = \cdot 6966.$$

The value by straightforward calculation of formula (lvi.) is ·6981, the two results giving substantially the same value ·7. Thus we see that such a selection would reduce the correlation of tibia and femur by 12·5 per cent.

Illustration (*II.*).—Suppose the correlation of humerus and femur to be ·5, and of those with stature to ·7 and ·8 respectively. How would the correlation of humerus and femur be modified by a selection of stature given by $s_1/\sigma_1 = \cdot 5$?

In this case, $\mu_1 = \cdot 5$, $r_{12} = \cdot 5$, $r_{13} = \cdot 75$, $r_{23} = \cdot 5$, $\delta\mu_1 = 0$, $\delta r_{12} = \cdot 2$, $\delta r_{13} = \cdot 05$, $\delta r_{23} = 0$. We turn to Table III. (a) and take out (k) under R_5, which gives us $R = \cdot 3192$. We have $\Delta_{r_{12}}R = - \cdot 1841$ and $\Delta_{r_{13}}R = - \cdot 04185$, whence we find $R' = \cdot 1636$, but the differences of the table are too great at this point for the result to be very trustworthy.[*] Suppose we take μ_1 and r_{23} as before, but $r_{12} = \cdot 75$, $r_{13} = \cdot 75$, and therefore R, to be found from (m), $= \cdot 1351$; then $\delta r_{12} = - \cdot 05$, $\delta r_{13} = \cdot 05$ and $\Delta_{r_{12}}R = - \cdot 1841$ as before, $\Delta_{r_{13}}R = - \cdot 2995$. Hence we deduce $R' = \cdot 1120$. The mean of these two values of R' is ·1377, and the true value calculated from (lvi.) is $R' = \cdot 1395$. Taking ·14 for the practical value, we see that the correlation of humerus and femur has been reduced by this comparatively moderate selection of stature upwards of 70 per cent. !

Illustration (*III.*).—Suppose a case in which humerus and femur were not correlated, but that both were correlated ·7 with stature. What would be the effect of selecting stature with the same intensity, *i.e.*, $s_1/\sigma_1 = \cdot 5$?

Our best results from the tables will be to take R_5 (m) from Table I. (a), which gives $R = - \cdot 7297$. We have then $r_{12} = r_{13} = \cdot 75$, hence $\delta r_{12} = - \cdot 05$, $\delta r_{13} = - \cdot 05$. $\Delta_{r_{12}}R$ is to be found from (m) and (h) and $= - \cdot 3193 = \Delta_{r_{13}}R$, and

$$\begin{aligned} R' &= - \cdot 7297 + (\Delta_{r_{12}}R)\,\delta r_{12} + (\Delta_{r_{13}}R)\,\delta r_{13}, \\ &= - \cdot 7297 + \cdot 1277 = - \cdot 6020. \end{aligned}$$

The actual value by formula is $- \cdot 5810$.

---

[*] Second differences ought to be used, and the process indicated is practically equivalent to using them.

Now this again is a remarkable result; by selecting an organ correlated with two others, neither of which are correlated with each other, we have produced a considerable correlation, and what is more, one of a *negative* sign.

In other words, if humerus and femur were unrelated to each other, but were related to stature, then a selection of stature would result in men of long femur having a short humerus, and *vice versâ.*

Illustration IV.—Suppose the correlation between greatest length and breadth of the skull to be ·25, and between both and the auricular height to be ·5. Now let a stringent selection, $\frac{s_1}{\sigma_1} = \frac{1}{10}$, of height take place. What modification will there be of the length and breadth correlation?

Table II. (a), R_1, case (j) gives us at once the result—

$$r_{23} = \cdot 0033.$$

In other words, the correlation between length and breadth would be sensibly destroyed by such a selection. Thus correlation can be created or destroyed or reversed by selection.

The above illustrations, hypothetical though they may be, will suffice to indicate how entirely dependent correlation is upon selection. We must look upon coefficients of correlation, in fact, as just as much the outcome of selection as coefficients of variation, standard-deviations, or even the mean size of organs. No selection can take place, in the sense in which it has usually been understood to take place—*i.e.*, by a change of mean and of variability, without at the same time the means, variabilities and the *correlations* of all correlated, but not directly selected, organs being varied. This is true whether the non-selected organs be initially correlated or not among themselves. We must always bear in mind this all-important fundamental conception, that natural or artificial selection, or even random sampling, are in themselves active factors in the modification (*i.e.*, creation, destruction, or reversal) of correlation. Thus not only is the impossibility of the constancy of correlation for local races obvious, but the primary importance of insuring that our samples are *representative*, and not accidentally selected samples, in all observations or experiments on heredity, homotyposis, or organic correlation becomes more and more manifest. We must not lay too much stress on two heredity constants—differing, for example, by more than the probable error of their difference—unless we are convinced, which practically it will be difficult to be, that all modification of correlation by unintentional and unmarked selection has really been avoided.

(7.) Let us now take the next most simple case. *If* A, B, C, D *be four mutually correlated organs* (*in either the same or different individuals*), *and a selection take place of* A *and* B, *to find the changes in the characters of the non-selected organs.*

Let subscripts 1, 2, 3, 4 mark the organs A, B, C, D respectively; let h_1, h_2, s_1, s_2, ρ_{12} be the constants which determine the selection of A and B; and let us apply (xlv.) to (xlvii.). Here the coefficients will be obtained from the set for $q = 2$ in (xxxiii.) and (xliv.). Hence we find:

$$\Sigma_3{}^2 = \sigma_3{}^2 \left\{ \frac{1 - r_{12}{}^2 - r_{13}{}^2 - r_{23}{}^2 + 2r_{12}r_{13}r_{23}}{1 - r_{12}{}^2} + \left(\frac{r_{13} - r_{23}r_{12}}{1 - r_{12}{}^2}\right)^2 \left(\frac{s_1}{\sigma_1}\right)^2 + \left(\frac{r_{23} - r_{13}r_{12}}{1 - r_{12}{}^2}\right)^2 \left(\frac{s_2}{\sigma_2}\right)^2 \right.$$
$$\left. + 2\rho_{12} \frac{r_{13} - r_{23}r_{12}}{1 - r_{12}{}^2} \frac{r_{23} - r_{13}r_{12}}{1 - r_{12}{}^2} \frac{s_1 s_2}{\sigma_1 \sigma_2} \right\} \quad \ldots \ldots \quad \text{(lviii.)}$$

$$\Sigma_4{}^2 = \sigma_4{}^2 \left\{ \frac{1 - r_{12}{}^2 - r_{14}{}^2 - r_{24}{}^2 + 2r_{12}r_{14}r_{24}}{1 - r_{12}{}^2} + \left(\frac{r_{14} - r_{24}r_{12}}{1 - r_{12}{}^2}\right)^2 \left(\frac{s_1}{\sigma_1}\right)^2 + \left(\frac{r_{24} - r_{14}r_{12}}{1 - r_{12}{}^2}\right)^2 \left(\frac{s_2}{\sigma_2}\right)^2 \right.$$
$$\left. + 2\rho_{12} \frac{r_{14} - r_{24}r_{12}}{1 - r_{12}{}^2} \frac{r_{24} - r_{14}r_{12}}{1 - r_{12}{}^2} \frac{s_1 s_2}{\sigma_1 \sigma_2} \right\} \quad \ldots \ldots \quad \text{(lix.)}$$

$$\Sigma_3 \Sigma_4 \mathbf{r}_{34} = \sigma_3 \sigma_4 \left\{ \frac{r_{34}(1 - r_{12}{}^2) - r_{13}r_{14} - r_{23}r_{24} + r_{12}(r_{13}r_{24} + r_{14}r_{23})}{1 - r_{12}{}^2} \right.$$
$$+ \frac{r_{13} - r_{23}r_{12}}{1 - r_{12}{}^2} \frac{r_{14} - r_{24}r_{12}}{1 - r_{12}{}^2} \left(\frac{s_1}{\sigma_1}\right)^2 + \frac{r_{23} - r_{13}r_{12}}{1 - r_{12}{}^2} \frac{r_{24} - r_{14}r_{12}}{1 - r_{12}{}^2} \left(\frac{s_2}{\sigma_2}\right)^2$$
$$\left. + \rho_{12} \left(\frac{r_{13} - r_{23}r_{12}}{1 - r_{12}{}^2} \frac{r_{24} - r_{14}r_{12}}{1 - r_{12}{}^2} + \frac{r_{14} - r_{24}r_{12}}{1 - r_{12}{}^2} \frac{r_{23} - r_{13}r_{12}}{1 - r_{12}{}^2} \right) \frac{s_1 s_2}{\sigma_1 \sigma_2} \right\} . \quad . \quad \text{(lx.)}$$

$$s_1 \Sigma_3 \mathbf{r}_{13} = s_1 \sigma_3 \left\{ \frac{r_{13} - r_{12}r_{23}}{1 - r_{12}{}^2} \frac{s_1}{\sigma_1} + \frac{r_{23} - r_{12}r_{13}}{1 - r_{12}{}^2} \rho_{12} \frac{s_2}{\sigma_2} \right\} . \quad \ldots \quad \text{(lxi.)}$$

$$s_1 \Sigma_4 \mathbf{r}_{14} = s_1 \sigma_4 \left\{ \frac{r_{14} - r_{12}r_{24}}{1 - r_{12}{}^2} \frac{s_1}{\sigma_1} + \frac{r_{24} - r_{12}r_{14}}{1 - r_{12}{}^2} \rho_{12} \frac{s_2}{\sigma_2} \right\} . \quad \ldots \quad \text{(lxii.)}$$

$$s_2 \Sigma_3 \mathbf{r}_{23} = s_2 \sigma_3 \left\{ \frac{r_{23} - r_{12}r_{13}}{1 - r_{12}{}^2} \frac{s_2}{\sigma_2} + \frac{r_{13} - r_{12}r_{23}}{1 - r_{12}{}^2} \rho_{12} \frac{s_1}{\sigma_1} \right\} . \quad \ldots \quad \text{(lxiii.)}$$

$$s_2 \Sigma_4 \mathbf{r}_{24} = s_2 \sigma_4 \left\{ \frac{r_{24} - r_{12}r_{14}}{1 - r_{12}{}^2} \frac{s_2}{\sigma_2} + \frac{r_{14} - r_{12}r_{24}}{1 - r_{12}{}^2} \rho_{12} \frac{s_1}{\sigma_1} \right\} . \quad \ldots \quad \text{(lxiv.)}$$

Finally, for the change of means of the non-directly selected organs, we have:

$$x'_3 = \frac{r_{13} - r_{12}r_{23}}{1 - r_{12}{}^2} \frac{\sigma_3}{\sigma_1} h_1 + \frac{r_{23} - r_{12}r_{13}}{1 - r_{12}{}^2} \frac{\sigma_3}{\sigma_2} h_2 \quad \ldots \ldots \quad \text{(lxv.)}$$

$$x'_4 = \frac{r_{14} - r_{12}r_{24}}{1 - r_{12}{}^2} \frac{\sigma_4}{\sigma_1} h_1 + \frac{r_{24} - r_{12}r_{14}}{1 - r_{12}{}^2} \frac{\sigma_4}{\sigma_2} h_2 \quad \ldots \ldots \quad \text{(lxvi.)}$$

If we write (lxv.) and (lxvi.) in the form—

$$\left.\begin{aligned} x'_3 &= \beta_{13}\frac{\sigma_3}{\sigma_1}h_1 + \beta_{23}\frac{\sigma_3}{\sigma_2}h_2 \\ x'_4 &= \beta_{14}\frac{\sigma_4}{\sigma_1}h_1 + \beta_{24}\frac{\sigma_4}{\sigma_2}h_2 \end{aligned}\right\} \quad . \; . \; . \; . \; . \; . \; . \quad \text{(lxvii.)},$$

the β'*s* are the partial regression coefficients, and the whole solution can be expressed in terms of them. Thus :

$$\Sigma_3^{\,2} = \sigma_3^{\,2}\left\{1 - \beta_{13}r_{13} - \beta_{23}r_{23} + \beta_{13}^{\,2}\left(\frac{s_1}{\sigma_1}\right)^2 + \beta_{23}^{\,2}\left(\frac{s_2}{\sigma_2}\right)^2 + 2\rho_{12}\beta_{13}\beta_{23}\frac{s_1s_2}{\sigma_1\sigma_2}\right\}. \quad \text{(lxviii.)}.$$

$$\begin{aligned} \Sigma_3\Sigma_4\mathfrak{r}_{34} = \sigma_3\sigma_4\Bigg\{r_{34} - \beta_{13}r_{14} - \beta_{23}r_{24} + \beta_{13}\beta_{14}\left(\frac{s_1}{\sigma_1}\right)^2 + \beta_{23}\beta_{24}\left(\frac{s_2}{\sigma_2}\right)^2 \\ + \rho_{12}(\beta_{13}\beta_{24} + \beta_{23}\beta_{14})\frac{s_1s_2}{\sigma_1\sigma_2}\Bigg\} \quad . \; . \; . \; . \; . \; . \; . \quad \text{(lxix.)}. \end{aligned}$$

$$s_1\Sigma_3\mathfrak{r}_{13} = s_1\sigma_3\left\{\beta_{13}\frac{s_1}{\sigma_1} + \rho_{12}\beta_{23}\frac{s_2}{\sigma_2}\right\} \quad . \; . \; . \; . \; . \; . \; . \quad \text{(lxx.)}.$$

Thus the whole series of results can be easily calculated, if the regression coefficients are first calculated.

I may make some remarks upon these results. A formula equivalent to (lxviii.) was first given by me in my memoir on "Heredity, Panmixia, and Regression" ('Phil. Trans.' A, vol. 187, p. 303), and used for certain problems of inheritance, and conclusions drawn from (lxix.) or (lxx.) have been cited or indicated in other memoirs.

Some interesting results follow at once. If the selection be very stringent, s_1/σ and $s_2/\sigma = 0$ sensibly, then all correlation between a selected and non-selected organ is destroyed. But

$$\mathfrak{r}_{34} = \frac{r_{34}(1 - r_{12}^{\,2}) - r_{13}r_{14} - r_{23}r_{24} + r_{12}(r_{13}r_{24} + r_{14}r_{23})}{\sqrt{(1 - r_{12}^{\,2} - r_{23}^{\,2} - r_{13}^{\,2} + 2r_{12}r_{13}r_{23})}\sqrt{(1 - r_{12}^{\,2} - r_{24}^{\,2} - r_{14}^{\,2} + 2r_{12}r_{14}r_{24})}} \;. \quad \text{(lxxi.)}.$$

This is what I have termed a *partial* correlation coefficient—*i.e.*, the correlation between C and D when fixed values are given to A and B. So far as I am aware, such coefficients were first directly used by Mr. G. U. YULE in certain economic problems.[*] They are of very considerable interest, but for natural or artificial selection are not quite so important as the generalised form (lxix.), for we generally select *about* a mean value, and not absolutely at it.

It will be noticed that the coefficient of correlation of two non-selected organs differs from the corresponding partial correlation coefficient by terms of the square order in s/σ, but the coefficient of correlation of a selected and non-selected organ

---

[*] 'Roy. Soc. Proc.,' vol. 60, pp. 485, 488; 'Economic Journal,' December, 1895, and December, 1896.

differs from zero by terms of the first order in s/σ. Hence, when selection is intense (s/σ small), we may neglect the former as compared with the latter, and we have thus the basis of a method of approximation very useful in some cases.

I have not yet succeeded in giving a geometrical interpretation to the above formulæ, but have little doubt that it would be quite easy if the "spherical" geometry of four dimensioned space were more familiar to me. It will suffice to observe that it is easy to find cases in which the correlation of a non-directly selected organ with a directly selected organ, or with another of its own class, can be reduced, destroyed, increased, or reversed. In fact, all our previous warnings as to the caution necessary in avoiding unintentional selection in collecting material for testing correlation remain in force, and, indeed, are emphasised.

The following illustrations will indicate the kind of problems which may be attacked by such formulæ as the above :—

(8). ILLUSTRATION I.—*Influence of a Selection of the Long Bones of the Leg on the Long Bones of the Arm, and on the Relation of the Leg to the Arm.*

A numerical example will throw light on the application of the above formulæ, and effectively illustrate the manner in which a selection influences size, variation, and correlation.

Consider the long bones femur, tibia, humerus, and radius, indicated by the subscripts 1, 2, 3, 4 respectively, and let m_1, m_2, m_3, m_4 be the mean values in centimetres. Then the following numerical values are given in a memoir by Miss ALICE LEE and myself :—*

	French ♂.	Aino ♂.
m_1	45·23	40·77
m_2	36·81	33·89
m_3	33·01	29·50
m_4	24·39	21·55
σ_1	2·37	1·90
σ_2	1·80	1·67
σ_3	1·54	1·34
σ_4	1·17	1·06
r_{12}	·806	·827
r_{13}	·842	·858
r_{14}	·744	·789
r_{23}	·860	·745
r_{24}	·780	·865
r_{34}	·845	·776

---

* 'Roy. Soc. Proc.,' vol. 61, p. 343 *et seq.* The correlations are not worked out for exactly the same lengths in the case of the two races, but the numbers will serve quite well for the purposes of illustration.

Now let us select from the French population a group having the same characteristics of the long bones of the leg as the Aino population, and then compare the characteristics of the arm bones of this selected group with those of the Aino population.

Our selection is given by :

$$h_1 = 40{\cdot}77 - 45{\cdot}23 = -4{\cdot}46,$$
$$h_2 = 33{\cdot}89 - 36{\cdot}81 = -2{\cdot}92,$$
$$s_1 = 1{\cdot}90 \quad \rho_{12} = {\cdot}827,$$
$$s_2 = 1{\cdot}67.$$

The following constants must now be determined arithmetically :

$$s_1/\sigma_1 = 1{\cdot}90/2{\cdot}37 = {\cdot}802,$$
$$s_2/\sigma_2 = 1{\cdot}67/1{\cdot}80 = {\cdot}928,$$

$$\frac{r_{13} - r_{12}r_{23}}{1 - r_{12}^2} = {\cdot}426, \qquad \frac{r_{23} - r_{12}r_{13}}{1 - r_{12}^2} = {\cdot}518,$$

$$\frac{r_{14} - r_{12}r_{24}}{1 - r_{12}^2} = {\cdot}329, \qquad \frac{r_{24} - r_{12}r_{14}}{1 - r_{12}^2} = {\cdot}515,$$

$$\frac{1 - r_{12}^2 - r_{23}^2 - r_{13}^2 + 2r_{12}r_{13}r_{23}}{1 - r_{12}^2} = {\cdot}196,$$

$$\frac{1 - r_{12}^2 - r_{24}^2 - r_{14}^2 + 2r_{12}r_{24}r_{14}}{1 - r_{12}^2} = {\cdot}354,$$

$$\frac{r_{34}(1 - r_{12}^2) - r_{13}r_{14} - r_{23}r_{24} + r_{12}(r_{13}r_{24} + r_{14}r_{23})}{1 - r_{12}^2} = {\cdot}125.$$

If x'_3 and x'_4 be the mean humerus and radius of a femur-tibia selection from the French population, we have from (lxv.)—

$$x'_3 = 33{\cdot}01 + {\cdot}277h_1 + {\cdot}443h_2,$$
$$x'_4 = 24{\cdot}39 + {\cdot}162h_1 + {\cdot}335h_2.$$

These would give the effect of selecting any femur and tibia defined by h_1 and h_2 from the mean values of the humerus and radius. For the particular selection indicated above :

$$x'_3 = 30{\cdot}48, \qquad x'_4 = 22{\cdot}69,$$

both of which are about a centimetre in excess of the Aino population. By selecting, therefore, from the French, a population with a mean leg like the Aino, we should still find the average arm of this population some two centimetres greater in length than the Aino. The variabilities Σ_3 and Σ_4 of humerus and radius for a population selected from the French by femur and tibia are obtained from (lxiii.) and (lix.).

We have :

$$\Sigma_3^2/\sigma_3^2 = {\cdot}196 + {\cdot}181\left(\frac{s_1}{\sigma_1}\right)^2 + {\cdot}268\left(\frac{s_2}{\sigma_2}\right)^2 + 2\,\frac{s_1}{\sigma_1}\,\frac{s_2}{\sigma_2}\,\rho_{12} \times {\cdot}221,$$

$$\Sigma_4^2/\sigma_4^2 = {\cdot}354 + {\cdot}108\left(\frac{s_1}{\sigma_1}\right)^2 + {\cdot}265\left(\frac{s_2}{\sigma_2}\right)^2 + 2\,\frac{s_1}{\sigma_1}\,\frac{s_2}{\sigma_2}\,\rho_{12} \times {\cdot}169.$$

These give for the particular case :

$$\Sigma_3 = 1{\cdot}39, \qquad \Sigma_4 = 1{\cdot}11.$$

Turning to the correlation of humerus and radius, we have by (lx.) :

$$r_{34} = \frac{\sigma_3}{\Sigma_3}\,\frac{\sigma_4}{\Sigma_4}\left\{{\cdot}125 + {\cdot}140\left(\frac{s_1}{\sigma_1}\right)^2 + {\cdot}267\left(\frac{s_2}{\sigma_2}\right)^2 + {\cdot}389\,\rho_{12}\,\frac{s_1}{\sigma_1}\,\frac{s_2}{\sigma_2}\right\},$$

giving for the particular case :

$$r_{34} = {\cdot}799.$$

It will thus be seen that if we selected from the French a group with the same variabilities and correlation of femur and tibia as the Aino, the variabilities and correlation of the humerus and radius of this group would not be very different from those of the Aino. On the other hand, the correlations between upper and lower members would be very significantly different.

Generally we have by (lxi.) for selection from the French :

$$r_{13} = \frac{\sigma_3}{\Sigma_3}\left\{{\cdot}426\,\frac{s_1}{\sigma_1} + {\cdot}518\,\rho\,\frac{s_2}{\sigma_2}\right\},$$

$$r_{14} = \frac{\sigma_4}{\Sigma_4}\left\{{\cdot}329\,\frac{s_1}{\sigma_1} + {\cdot}515\,\rho\,\frac{s_2}{\sigma_2}\right\},$$

$$r_{23} = \frac{\sigma_3}{\Sigma_3}\left\{{\cdot}518\,\frac{s_2}{\sigma_2} + {\cdot}426\,\rho\,\frac{s_1}{\sigma_1}\right\},$$

$$r_{24} = \frac{\sigma_4}{\Sigma_4}\left\{{\cdot}515\,\frac{s_2}{\sigma_2} + {\cdot}329\,\rho\,\frac{s_1}{\sigma_1}\right\}.$$

These yield for our particular case :

$$r_{13} = {\cdot}819, \qquad r_{14} = {\cdot}694,$$

$$r_{23} = {\cdot}845, \qquad r_{24} = {\cdot}768.$$

These are all smaller than the corresponding French values, the selection has reduced the correlation, but the Aino population has in all the cases but r_{23} a greater correlation than the French. We must accordingly conclude that by a leg selection from the French aimed at reproducing the proportions of the Aino leg, we should not obtain an arm equivalent to the Aino arm. The divergences are indicated in the accompanying table :—

Selection from the French.		Aino.	Unselected French.
Mean humerus	30·48	29·50	33·01
„ radius	22·69	21·55	24·39
Variability of humerus	1·39	1·34	1·54
„ radius	1·11	1·06	1·17
Correlation of humerus and radius	·799	·776	·845
„ „ femur and humerus	·819	·858	·842
„ „ femur and radius	·694	·789	·744
„ „ tibia and humerus	·845	·745	·860
„ „ tibia and radius	·768	·865	·780

There is, of course, no special reason for supposing that the French and Aino differ merely by an evolution which has acted by selection of femur and tibia. We might have obtained a race out of the French more nearly akin to the Aino by a selection of femur and humerus, but the process would numerically be exactly similar. The particular illustration here chosen is taken merely as an instance, to indicate how the methods developed in this memoir enable us to ascertain with quantitative certainty how far racial differences may be due to the more or less stringent selection of a limited number of organs in the one race.

If we consider that local races have been differentiated from a parent stock by the selection of the chief or more markedly divergent organs, then we have in processes such as that just illustrated a method of ascertaining, at least tentatively, whether two races are to be considered as merely local varieties, and further the particular organs through selection of which the differentiation has taken place.

ILLUSTRATION II.—*Influence of a Selection of Femur and Humerus in Modifying Stature.*

The following data have been calculated for me by Miss ALICE LEE from ROLLET'S measurements on the French :—*

---

* They have been undertaking, with the view of determining more scientifically than appears to me yet to have been done, the mean stature of a race from a measurement of the long bones found in burial mounds, &c. ROLLET'S measurements are given in 'De la Mensuration des Os longs des Membres,' Lyons, 1889. I hope shortly to publish a memoir on the subject. [The memoir in question was published in 'Phil. Trans.,' A, vol. 192, pp. 169–244, 1898.]

	♂	♀
Mean stature, m_1	166·26 cms.	154·02 cms.
Variation of stature, σ_1	5·50 ,,	5·45 ,,
Mean femur, m_2	45·23 ,,	41·57 ,,
Variation of femur, σ_2	2·37 ,,	2·26 ,,
Mean humerus m_3	33·01 ,,	29·77 ,,
Variation of humerus, σ_3	1·54 ,,	1·53 ,,
Correlation, femur and humerus, r_{23}	·842	·872
,, stature and humerus, r_{13}	·809	·771
,, stature and femur, r_{12}	·811	·805

Now let us select from among French males a group having the same variability, correlation, and mean size of humerus and femur as French females, and let us ask how this would alter the variability (Σ_1), mean size (M_1) of stature in French males, and also the correlation between stature and humerus ($\mathfrak{r}_{13}$) and stature and femur ($\mathfrak{r}_{12}$).

We have at once from the second column—

$$h_2 = -3{\cdot}66, \qquad h_3 = -3{\cdot}24,$$

$$s_2 = 2{\cdot}26, \qquad s_3 = 1{\cdot}53,$$

$$\rho_{23} = {\cdot}872,$$

whence we find,

$$\frac{r_{12} - r_{23}r_{13}}{1 - r_{23}^2} = {\cdot}447, \qquad \frac{r_{13} - r_{23}r_{12}}{1 - r_{23}^2} = {\cdot}433.$$

From (lxv.) we deduce,

$$\mathrm{M}_1 = 166{\cdot}26 + 1{\cdot}037h_2 + 1{\cdot}546h_3.$$

This formula gives the stature of any group of males selected from the French, and having their femur and humerus respectively h_2 and h_3 centims. longer than the average.

For the special selection referred to above, $h_2 = -3{\cdot}66$ and $h_3 = -3{\cdot}24$, hence

$$\mathrm{M}_1 = 166{\cdot}26 - 3{\cdot}79 - 5{\cdot}01 = 157{\cdot}46.$$

This example shows us that if we selected French men with the same femur and humerus as French women, it would be the selection of the humerus which would contribute mostly to the reduction of stature—a somewhat singular result. Further, such a selected group of French men would be still some $3\frac{1}{2}$ centims. taller than the average of French women (instead of about $12\frac{1}{4}$ centims.). Probably had we selected the tibia as well, the greater portion of this remaining advantage in height would have disappeared.

To find the variability in stature of the selected group we must use (lviii.). We deduce:

$$\Sigma_1^2 = \sigma_1^2 \left\{ \cdot 287 + \cdot 200 \left(\frac{s_2}{\sigma_2}\right)^2 + \cdot 187 \left(\frac{s_3}{\sigma_3}\right)^2 + \cdot 387 \frac{s_2}{\sigma_2} \frac{s_3}{\sigma_3} \rho \right\}.$$

In our particular case this gives:

$$\Sigma_1 = \sigma_1 \times \cdot 987 = 5 \cdot 43.$$

The actual variability in stature of French women is measured by a standard deviation of 5·45. Hence our selected group of men would be sensibly equally variable with French women, as far as absolute variation is concerned.

Lastly, from (lxi.):

$$\mathrm{r}_{12} = \frac{\sigma_1}{\Sigma_1} \left\{ \cdot 447 \frac{s_2}{\sigma_2} + \cdot 433 \frac{s_3}{\sigma_3} \rho_{23} \right\},$$

$$\mathrm{r}_{13} = \frac{\sigma_1}{\Sigma_1} \left\{ \cdot 433 \frac{s_3}{\sigma_3} + \cdot 447 \frac{s_2}{\sigma_2} \rho_{23} \right\},$$

which give in our particular case:

$$\mathrm{r}_{12} = \frac{\sigma_1}{\Sigma_1} \times \cdot 8011 = \cdot 811,$$

$$\mathrm{r}_{13} = \frac{\sigma_1}{\Sigma_1} \times \cdot 8017 = \cdot 812.$$

Such a selection, therefore, would accordingly only increase insensibly the correlation between stature and humerus, while leaving that between stature and femur the same. The sensible reduction of correlation between stature and humerus (·809 to ·771), which is found as we pass from male to female, does not arise when we select a group of males with their femur and humerus of the same length, variation, and correlation as those of the females.

ILLUSTRATION III.—*Influence of a Selection of Stature in Modifying Femur and Humerus.*

Let us select a group of French men having the same height and variability in height as French women, and calculate the changes which will arise in their femur and humerus.

Here the selection is given by:

$$h_1 = -12 \cdot 24, \qquad s_1 = 5 \cdot 45.$$

We now need only the earlier formulæ of this memoir. From (l.) we find

$$\mathrm{M}_2 = m_2 + \cdot 349 h_1, \qquad \mathrm{M}_3 = m_3 + \cdot 227 h_1.$$

These give for our case:

$$M_2 = 40{\cdot}95, \qquad M_3 = 30{\cdot}24.$$

Thus a group of males, selected to have the same stature as the females, would have a slightly shorter femur and a slightly longer humerus. A slightly longer femur in woman and a slightly longer humerus in man would thus appear to be sexual characters.

Turning to the variations, these are given by (xlix.). We find:

$$\Sigma_2 = 2{\cdot}36, \qquad \Sigma_3 = 1{\cdot}53.$$

This shows us that while the selection would give the same variability of humerus to the men that women have, it would fail to produce the reduction of variability in the femur, which is characteristic of the women.

From (li.) we deduce

$$r_{12} = {\cdot}808, \qquad r_{13} = {\cdot}806,$$

while from (lvi.) we have

$$r_{23} = {\cdot}840.$$

Thus we see that very small changes would be made in the correlations, stature and femur, stature and humerus, and femur and humerus, if we selected French men to have the same size and variability of stature as French women. The explanation of this lies in the nearly equal absolute variability of the two sexes with regard to stature, for, as we have seen, it is the selection of variability which modifies correlation. Looking at the table of values on p. 36, we see that the largest difference of variability in the two sexes lies in the femur, and accordingly it is from a selection of femur that we should expect the greatest differences in the variability and correlation of the two sexes to have arisen, but even this difference alone would not account for the observed sexual differences in the correlation. Indeed, it would be surprising if it did, for the selection of other organs, notably the pelvis, must have played a considerable part in the differentiation of sex.*

(9.) I shall now proceed to a series of problems, which will show the application of results, such as those obtained in this memoir, to questions which arise in dealing with inheritance and selection. If we suppose a general population to have statistical "constants," which remain constant at any rate for a moderate interval, we still want to know not only the error which may arise from a *random* sampling, but also the sort of effect which results from our sample being too much drawn from one kind of environment, from a rather limited class, or from any other practically necessary or unconsciously introduced limitation of the random character of our sample.

---

* "Primitive man and woman are more nearly equal in size, variability, and correlation than highly civilized man and woman" ('Roy. Soc. Proc.,' vol. 61, p. 354).

ILLUSTRATION I.—*To find the Influence on the Intensity of Parental Heredity of the Selection of Parents.*

Let the subscripts 1, 2, 3 refer respectively to father, mother, child. Let us first select one parent—say, the father—very stringently, *i.e.*, $s_1/\sigma_1 = \mu_1$ is small. Then we need only equations (xlix.) and (li.). These give us :

$$\mathrm{r}_{13} = \frac{\mu_1 r_{13}}{\sqrt{1 - (1 - \mu_1^2)\, r_{13}^2}},$$

and

$$\Sigma_2^2 = \sigma_2^2 (1 - (1 - \mu_1^2)\, r_{12}^2).$$

The first may be written

$$\mathrm{r}_{13} = \frac{r_{13}}{\sqrt{r_{13}^2 + \dfrac{1 - r_{13}^2}{\mu_1^2}}},$$

or, we see that r_{13} will decrease, as μ_1 decreases. Thus if $r_{13} = \cdot 4$ we have for

$\mu_1 = 1/2,$	$\mathrm{r}_{13} = \cdot 2132,$	$\Sigma_2/\sigma_2 = \cdot 8367,$
$\mu_1 = 1/4,$	$\mathrm{r}_{13} = \cdot 1085,$	$\Sigma_2/\sigma_2 = \cdot 7906,$
$\mu_1 = 1/8,$	$\mathrm{r}_{13} = \cdot 0545,$	$\Sigma_2/\sigma_2 = \cdot 7786.$

It is clear, therefore, that the correlation of parent and child will be much reduced by such a selection. On the other hand, the regression coefficient will not be altered, *i.e.*, $\Sigma_2 \mathrm{r}_{13}/s_1 = \sigma_2 r_{13}/\sigma_1$, as we have seen. Hence in problems of heredity, where we suspect a parent to have been highly selected, we should seek for the regression of son on parent rather than for the correlation. Thus in the case of Basset Hounds,* some if not all the reduction in correlation between sire and offspring may be due to selection of the sire. A test of whether the reduction in correlation is due to selection of a parent ought to be given by a comparison of s_1 and Σ_2. We cannot, I think, suppose, unless natural selection be very stringent, that σ_2 differs much from σ_1. Hence it follows that Σ_2/s_1 ought generally to be large, if there be selection of a parent. We can hardly test this point effectively in the case of the Basset Hounds, owing to the nature of the classification. In racehorses, although the sire appears to be far more selected than the dam, there is not a great reduction of the coefficient of correlation between sire and offspring,† s_1 appears to be less than Σ_2, but not so greatly and certainly less, that we can be surprised that the correlation of sire and offspring is not much less than we have found it for material in which selection of the father is certainly far less marked. We must accept the warning as to the reduction

---

* 'Roy. Soc. Proc.,' vol. 66, p. 157.

† See 'Phil. Trans.,' A, vol. 195, p. 93.

in correlation produced by the stringent selection of one parent, but we must remember the complexity of the factors—the variety of other influences at work in selecting and modifying selection—before we lay much stress on this source of alteration in parental correlation.

Now let us deal with the case of both parents selected, and suppose their selection given by $s_1/\sigma_1 = \mu_1$, $s_2/\sigma_2 = \mu_2$, and the change of their coefficient of assortative mating from r_{12} to ρ_{12}. We have from formulæ (lxviii.) and (lxx.) by a little rearranging

$$\Sigma_3^2 = \sigma_3^2 \{1 - \beta_{13}^2 (1 - \mu_1^2) - \beta_{23}^2 (1 - \mu_2^2) - 2 (r_{12} - \rho_{12}\mu_1\mu_2) \beta_{13}\beta_{23}\}$$

$$\Sigma_3\mathfrak{r}_{13} = \sigma_3 \{r_{13} - (1 - \mu_1) \beta_{13} - (r_{12} - \rho_{12}\mu_2) \beta_{23}\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxii.)},$$

where $\beta_{13} = (r_{13} - r_{23}r_{12}) / (1 - r_{12}^2)$, $\beta_{23} = (r_{23} - r_{13}r_{12}) / (1 - r_{12}^2)$. Now let us take special cases to bring out points. Let us suppose $r_{12} = 0$, or no assortative mating to exist, and let us enquire what change would be made in parental correlation if we selected parents who had assortatively mated, without altering their variability, *i.e.*, let us take $\mu_1 = 1$, $\mu_2 = 1$, we have at once

$$\Sigma_3^2 = \sigma_3^2 (1 + 2\rho_{12}r_{13}r_{23}) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxiii.)},$$

$$\Sigma_3\mathfrak{r}_{13} = \sigma_3 (r_{13} + \rho_{12}r_{23}) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxiv.)},$$

whence we find for $r_{13} = r_{23} = \cdot 4$ and

$$\begin{array}{lll}
\rho_{12} = \cdot 1, & \Sigma_3/\sigma_3 = 1{\cdot}0149, & \mathfrak{r}_{13} = \cdot 4335, \\
\rho_{12} = \cdot 2, & \qquad\;\; = 1{\cdot}0315, & \quad\;\; = \cdot 4653, \\
\rho_{12} = \cdot 3, & \qquad\;\; = 1{\cdot}0469, & \quad\;\; = \cdot 4967, \\
\rho_{12} = \cdot 4, & \qquad\;\; = 1{\cdot}0621, & \quad\;\; = \cdot 5271, \\
\rho_{12} = \cdot 5, & \qquad\;\; = 1{\cdot}0770, & \quad\;\; = \cdot 5571, \\
\rho_{12} = 1, & \qquad\;\; = 1{\cdot}1489, & \quad\;\; = \cdot 6963.
\end{array}$$

Thus the general effect of assortative pairing of parents is to increase the correlation between parent and offspring sensibly, but not to very rapidly increase the variability of the offspring. Thus marriages within a class would, if heredity statistics were collected for a class, tend to show increased parental correlation. Very high assortative mating no doubt occurs with some forms of breeding, and we may well find in such cases higher values of the parental heredity than we should obtain for a population of the same species with random mating. I think this may be an effective factor in the raising of the parental correlation in the case of coat-colour in thoroughbred horses.

Now let us see what happens if we select both parents moderately. As test cases, let us take $\mu_1 = \mu_2 = \cdot 8$ and $\cdot 5$, and for the extreme $= 0$.

We have at once

$$\Sigma_3 = \sigma_3 \sqrt{\cdot 68 + \cdot 32 \mu_1^2 (1 + \rho_{12})} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxv.)},$$

$$\Sigma_3 \mathfrak{r}_{13} = \sigma_3 \{\cdot 4 \mu_1 (1 + \rho_{12})\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxvi.)}.$$

Hence we deduce

ρ_{12}.	$\mu_1 = \mu_2 = \cdot 8$.		$\mu_1 = \mu_2 = \cdot 5$.		$\mu_1 = \mu_2 = 0$.	
	Σ_3/σ_3.	$\mathfrak{r}_{13}$.	Σ_3/σ_3.	$\mathfrak{r}_{13}$.	Σ_3/σ_3.	$\mathfrak{r}_{13}$.
·1	·9515	·3700	·8764	·2865	·8246	0
·2	·9622	·3991	·8809	·3093	·8246	0
·3	·9727	·4277	·8854	·3316	·8246	0
·4	·9832	·4556	·8899	·3535	·8246	0
·5	·9936	·4831	·8944	·3750	·8246	0
1	1·0438	·6131	·9165	·4762	·8246	0

This table is very instructive. It shows us that selection and assortative mating are factors of opposite influence; that if selection be only moderate, then with considerable assortative mating the coefficient of parental correlation may be increased, but if selection be stringent, then assortative mating cannot counteract, even if as high as ·5, its destructive influence on parental correlation.

For example, if we take parents remarkable for some intellectual or physical character, say with a variability only a very small fraction of that of the general population, then, however proportionately we might pair them, we should find their relationship to their children, as measured by the coefficient of correlation, very sensibly reduced below that of the general population. I think we have here the reason why Mr. GALTON'S Family Data, which were drawn from a rather narrow class, and had only a small coefficient of assortative mating, give so much smaller parental correlation than my own Family Data, which seem to me drawn from a wider class, and have a considerably higher assortative mating.*

It will be clear that with factors like assortative mating, natural selection, artificial selection of breeders, unconscious selection of material from one class or one environment, modifying our coefficients of heredity in one or another direction, we can hardly hope for more in practical statistics than an approximation to the strength of the pure inheritance factor by dealing with the average of as many races and characters as possible.

---

* The work for Mr. GALTON'S Family Data is given, 'Phil. Trans.,' A, vol. 187, p. 270. My own results are as yet unpublished. The average value is about ·45, as compared with Mr. GALTON'S ·34

Illustration II.—*To Find the Influence of Parental Selection on Modifying Fraternal Correlation.*

Let the subscripts 1, 2 represent the parents, and 3 and 4 two of their offspring. Let us first select one parent only, the selection being given as before by $s_1/\sigma_1 = \mu_1$. Our formulæ will now be (liv.) and (lvi.). So far as the change in variability is concerned, we have already discussed it under our first illustration, so we need only consider:

$$\mathbf{r}_{34} = \frac{(r_{34} - r_{13}r_{14})(1 - \mu_1^2) + \mu_1^2 r_{34}}{\sqrt{\{1 - (1 - \mu_1^2)r_{13}^2\}\{1 - (1 - \mu_1^2)r_{14}^2\}}} \quad . \quad . \quad . \quad . \quad \text{(lxxvii.)}.$$

Now $r_{13} = r_{14}$, if the offspring are of one sex; hence:

$$\mathbf{r}_{34} = \frac{r_{34} - r_{13}^2(1 - \mu_1^2)}{1 - r_{13}^2(1 - \mu_1^2)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxviii.)}.$$

If we take $r_{13} = {\cdot}4$ and $r_{34} = {\cdot}5$ as reasonable values, we have

$$\mathbf{r}_{34} = \frac{{\cdot}34 + {\cdot}16\mu_1^2}{{\cdot}84 + {\cdot}16\mu_1^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxix.)}.$$

Thus $\mathbf{r}_{34}$ will be greatest when μ_1 is greatest, *i.e.*, when there is no selection, and will decrease with μ_1 until it reaches ·4048, when $\mu_1 = 0$, or there is selection of fathers of one value of the character only.[*]

The selection of one parent only does not, therefore, immensely modify the correlation of brothers. Still, if we work sensibly with one class of the community—say, men of genius—we should expect to find their sons rather less like each other than if we worked with the general population of brothers.

Now let us select both parents. Here again the variability of the offspring has already been dealt with. We are concerned with equation (lxix.), and we shall put $r_{13} = r_{14} = r_{23} = r_{24} = r$, or make parental influence equipotent for the two sexes. Hence

$$\beta_{13} = \beta_{23} = \beta_{14} = \beta_{24} = \beta = \frac{r}{1 + r_{12}},$$

where r is the parental correlation, and r_{12} the coefficient of assortative mating. Hence we find

$$\mathbf{r}_{34} = \frac{r_{34} - \beta^2\{1 - \mu_1^2 + 1 - \mu_2^2 + 2(r_{12} - \rho_{12}\mu_1\mu_2)\}}{1 - \beta^2\{1 - \mu_1^2 + 1 - \mu_2^2 + 2(r_{12} - \rho_{12}\mu_1\mu_2)\}} \quad . \quad . \quad . \quad \text{(lxxx.)}.$$

To reduce to numbers, suppose $\mu_1 = \mu_2$, and $r_{12} = 0$ for the general population. We have

$$\mathbf{r}_{34} = \frac{r_{34} - 2r^2\{1 - \mu_1^2(1 + \rho_{12})\}}{1 - 2r^2\{1 - \mu_1^2(1 + \rho_{12})\}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxxi.)}$$

---

[*] In general the value of $\mathbf{r}_{34}$ ranges from r_{34} down to $\dfrac{r_{34} - r_{13}^2}{1 - r_{13}^2}$.

Hence if we put $r_{34} = \cdot 5$, and $r = \cdot 4$,

$$\mathrm{r}_{34} = \frac{\cdot 18 + \cdot 32\mu_1{}^2(1 + \rho_{12})}{\cdot 68 + \cdot 32\mu_1{}^2(1 + \rho_{12})} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxxii.)}.$$

The following table will suffice to indicate the changes which take place, when we give a series of values to μ_1 and ρ_{12}. Thus the first row gives the influence of selecting parents without any assortative mating. We see that with increasing stringency of selection the reduction of correlation is very considerable, and that with such selection the influence of assortative mating becomes less and less. Nevertheless, assortative mating can produce quite sensible results, if there be little or no selection. I am, indeed, inclined to think that a good deal of the high values found for the fraternal colour correlation in the thoroughbred foals* is due to much assortative colour mating in sire and dam. Of course it cannot be all due to this source.

VALUES of Fraternal Correlation with Parental Selection.

	$\mu_1 = 1.$	$\mu_1 = \cdot 8.$	$\mu_1 = \cdot 6.$	$\mu_1 = \cdot 4.$	$\mu_1 = \cdot 2.$	$\mu_1 = 0.$
$\rho_{12} = 0$	·5000	·4349	·3712	·3162	·2783	·2647
$\rho_{12} = \cdot 1$	·5155	·4477	·3802	·3209	·2796	·2647
$\rho_{12} = \cdot 2$	·5301	·4599	·3889	·3256	·2809	·2647
$\rho_{12} = \cdot 3$	·5438	·4716	·3974	·3303	·2823	·2647
$\rho_{12} = \cdot 5$	·5690	·4935	·4137	·3392	·2849	·2647
$\rho_{12} = 1$	·6212	·5411	·4508	·3609	·2914	·2647

On the whole, I think, we may conclude, so far as the relative influences of sexual selection in the form of assortative mating and natural selection go, that:

Both sexual and natural selection can sensibly modify the intensity of inheritance as measured by the coefficient of correlation, the former tends to raise, the latter to lower, its intensity. But the effect of the latter, if at all stringent, is to completely mask the effect of the former.

In fact, we may write

$$\mathrm{r}_{34} = 1 - \frac{1 - r_{34}}{1 - 2r^2 + 2r^2\mu_1{}^2(1 + \rho_{12})}.$$

Hence the smaller $\mu_1{}^2(1 + \rho_{12})$, the smaller will be fraternal correlation. This varies as the square of μ_1 and only as the linear power of $1 + \rho_{12}$. Thus we see at once why stringency of selection is far more potent than assortative mating.

ILLUSTRATION III.—*To find the influence of selecting two organs A and B in a parent, on the correlation of the like organs A′ and B′ in the offspring.*

Let the organs in the parent be denoted by 1 and 2, and in the offspring by 3 and 4. Suppose the organic correlation of the two organs in the general population to be r',

---

* 'Phil. Trans.,' A, vol. 195, p. 93.

so that $r_{12} = r_{34} = r'$ before any selection takes place. Let r, the correlation of the organs in the parent and offspring be supposed to be the same for both organs; then $r = r_{13} = r_{24}$. Finally we have the coefficients of cross-heredity, r_{14} and r_{23}. These must vanish if there be no heredity and no organic correlation, and should be perfect if both these are perfect. Hence we will take $r_{14} = r_{23} = rr'$ as a probable hypothesis.* With these values of the correlation coefficients we easily find

$$\beta_{13} = r = \beta_{24},\ \beta_{23} = \beta_{14} = 0.$$

Hence from (lxviii.) and (lxix.) we have:

$$\Sigma_3^2 = \sigma_3^2\{1 - r^2 + r^2\mu_1^2\}, \qquad \Sigma_4^2 = \{1 - r^2 + r^2\mu_2^2\} \quad . \quad . \quad \text{(lxxxiii.)},$$

$$\mathfrak{r}_{34} = \frac{r'(1 - r^2) + \rho_{12}r^2\mu_1\mu_2}{\sqrt{\{1 - r^2 + r^2\mu_1^2\}\{1 - r^2 + r^2\mu_2^2\}}} \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxxiv.)}.$$

For simplicity, suppose the stringency of the selection to be the same for both organs, then:

$$\mathfrak{r}_{34} = \frac{r' + \rho_{12}\mu_1^2\dfrac{r^2}{1 - r^2}}{1 + \mu_1^2\dfrac{r^2}{1 - r^2}}.$$

If $r = \cdot 4$, and $\rho_{12} = \gamma r'$,

$$\mathfrak{r}_{34} = r' \times \frac{1 + \frac{4}{21}\gamma\mu_1^2}{1 + \frac{4}{21}\mu_1^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxxv.)}.$$

The following table indicates the value of $\mathfrak{r}_{34}/r'$:—

	$\mu_1 = 1.$	$\mu_1 = \cdot 8.$	$\mu_1 = \cdot 6.$	$\mu_1 = \cdot 4.$	$\mu_1 = \cdot 2.$	$\mu_1 = 0.$
$\gamma = 1$	1	1	1	1	1	1
$\gamma = 0 \cdot 8$	·9680	·9783	·9872	·9941	·9985	1
$\gamma = 0 \cdot 6$	·9360	·9565	·9743	·9882	·9970	1
$\gamma = 0 \cdot 4$	·9040	·9348	·9615	·9823	·9955	1
$\gamma = 0 \cdot 2$	·8720	·9131	·9487	·9763	·9940	1
$\gamma = 0 \cdot 0$	·8400	·8913	·9358	·9704	·9924	1

It will be clear from this table that if the selection be at all stringent, no *reduction* of organic correlation in the parents will affect substantially the organic correlation in the offspring.

On the other hand, if γ be > 1, we can have considerable modifications in the value of the correlation, even if the selection be stringent.

---

* See 'Roy. Soc. Proc.,' vol. 62, p. 411. I have a good deal of data on the value of these cross-heredity correlations now reduced and soon to be published.

Thus we have the following values of $\mathfrak{r}_{31}/r'$, if :—

	$\mu_1 = 1.$	$\mu_1 = \cdot 8.$	$\mu_1 = \cdot 6.$	$\mu_1 = \cdot 4.$	$\mu_1 = \cdot 2.$	$\mu_1 = 0.$
$\gamma = 2$	1·1600	1·1087	1·0642	1·0296	1·0076	1
$= 10$	2·4400	1·9779	1·5775	1·2662	1·0681	1
$= 50$	8·8400	6·3243	4·1448	2·4492	1·3705	1

Lastly, if $r' = 0$:

$$\mathfrak{r}_{34} = \rho_{12} \frac{1}{1 + \frac{21}{4} \frac{1}{\mu_1{}^2}},$$

or, $\mathfrak{r}_{34}/\rho_{12}$ is given by :

	$\mu_1 = 1.$	$\mu_1 = \cdot 8.$	$\mu_1 = \cdot 6.$	$\mu_1 = \cdot 4.$	$\mu_1 = \cdot 2.$	$\mu_1 = 0.$
$\mathfrak{r}_{34}/\rho_{12}$	·1600	·1087	·0642	·0296	·0076	0

Thus, even if there were no correlation between the organs A and B in the general population, still a selection of parents in which such organs were correlated would lead to offspring with correlated organs A′ and B′. The amount of such correlation would only be ·1600, if the variability of the parent were not selected, and would diminish rapidly with stringent selection of variability. Still ·1600 is quite sensible, and would, if the selection continued for a few generations, continue to increase. Thus we see how selection of a pair of organs in a parent may increase or even create correlation between the like organs in the offspring.

The reader will find other interesting illustrations in tracing the influence of an absolute selection of one parent only on the correlation of the offspring, *e.g.*, relation between pairs of foals which all have a common sire, the influence of selecting an organ A in the sire and an organ B in the dam on the correlation of the organs A and B in the offspring, the influence on assortative mating of selecting parents of men of genius,* and in many other problems.

(10.) It is not without value to consider how ρ_{12} arises in the case of natural or artificial selection. Suppose we have two organs, A and B, then we shall endeavour (i.) to give these definite values, say $\bar{x}$ and $\bar{y}$, but we shall not be able to get all our individuals with such absolute values, we shall select with certain deviations from $\bar{x}$ and $\bar{y}$, given by $x = \bar{x} + x'$ and $y = \bar{y} + y'$, say. Further, we shall endeavour to

---

* This is a peculiarly interesting case, for if we select men of remarkable intellectual ability, we should expect to find both parents above the average of the general population, but with a *negative* correlation between them amounting at a maximum to − ·1905.

make y some function of x, say $y = f(x)$ or $\bar{y} + y' = f(\bar{x}) + x'f'(\bar{x}) + \frac{x'^2}{1\cdot2} f''(\bar{x}) + \dots$ by TAYLOR's theorem. But $\bar{y} = f(\bar{x})$ and x' is small, so that our attempted relation will be of the form :

$$y' = mx'.$$

Here m is the slope of a line to which we endeavour to confine the selected organs. But we shall not be quite able to exactly hit this relation off; when $x' = \epsilon$, we shall find that $y' = m\epsilon + \eta$, where η is a small unavoidable error in selection of y' uncorrelated with ϵ. Thus, if s_1 and s_2 be the selected variabilities, we shall have :

$$s_2^2 = \frac{1}{n} \mathrm{S}(y'^2) = \frac{1}{n} \mathrm{S}(m\epsilon + \eta)^2, \quad s_1^2 = \frac{1}{n} \mathrm{S}(\epsilon^2).$$

Therefore :
$$s_2^2 = m^2 s_1^2 + \frac{1}{n} \mathrm{S}(\eta^2).$$

Further :
$$\rho_{12} = \frac{\mathrm{S}(y'x')}{n s_1 s_2} = \frac{\mathrm{S}\{(m\epsilon + \eta)\epsilon\}}{n s_1 s_2} = \frac{m s_1^2}{s_1 s_2} = \frac{m s_1}{s_2},$$

and therefore :
$$\frac{1}{n} \mathrm{S}(\eta^2) = s_2^2 (1 - \rho_{12}^2).$$

Or, ρ_{12} is at once obtained from the slope of the line m, by which we endeavour to fix the relationship of the organs A and B. Or, again, we may look upon $s_2\sqrt{1 - \rho_{12}^2}$ as a quantity measuring the mean divergence of the B organ from that absolute fulfilment of the relationship between A and B which we are striving to attain. Thus ρ_{12} is a quantity which naturally arises in every attempt, whether artificial or natural, to select organs having a definite relationship to each other.

Much the same considerations arise when we select three or more organs. In each case the selected coefficients of correlation are constants which enable us to express (i.) to a first approximation the form of relationship we are aiming at, and (ii.) the average degree of divergence from absolute fulfilment of this relationship.

Thus, without regard to any particular distribution of frequency, the s's and the ρ's are the appropriate constants to express approximately the nature of any form of natural or artificial selection.

(11.) *On the Probability of Survival and the Surface of Survival Rates.*

In the course of the present paper I have assumed that when measurements are made on any population for a complex of n organs, the frequency surface may be taken as approximately normal. If this holds for the population before and after selection, and measurements made on many groups at different periods of life seem to indicate that it is approximately true, it follows that we can determine the form of the probability of survival as a function of the means, variations, and correlations of the selected and unselected populations.

Let the unselected population be given by

$$Z = Z_0 \text{ expt. } - \tfrac{1}{2}\{c_{11}x_1^2 + c_{22}x_2^2 + \ldots + c_{nn}x_n^2 + 2c_{12}x_1x_2 + \ldots + 2c_{n-1,n}x_{n-1}x_n\} \quad . \quad . \quad . \quad \text{(lxxxvi.)}.$$

Let the probability of survival be given by

$$p = p_0 f(x_1 - k_1, \quad x_2 - k_2, \quad x_3 - k_3, \ldots x_n - k_n) \quad . \quad . \quad \text{(lxxxvii.)},$$

where f is at present an unknown function, which is to be a maximum for

$$x_1 = k_1, \quad x_2 = k_2, \quad x_3 = k_3, \ldots x_n = k_n,$$

and, if the selection be at all stringent, to take rapidly decreasing values as

$$x_1 - k_1, \quad x_2 - k_2, \quad x_3 - k_3, \ldots x_n - k_n$$

take increasing large negative or positive values. It will be clear then that the individuals who are "fittest to survive," *i.e.*, have the smallest death-rate, are those whose organs are defined by :

$$x_1 = k_1, \quad x_2 = k_2, \ldots x_n = k_n,$$

and fitness generally will be measured by the closeness of the individual to these "fittest" individuals.

In order to find the surface of survivors, immediately after the selection if growth be taking place,* or at any later stage if growth have ceased, we have only to multiply Z by p, or :

$$z = Z \times p \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxxviii.)},$$

is what in the earlier part of this memoir I have termed the selection surface. Now if this selection surface be itself normal, it will be of the form :

$$\begin{aligned} z = z_0 \text{ expt. } - \tfrac{1}{2}\{b_{11}(x_1 - h_1)^2 + b_{22}(x_2 - h_2)^2 + \ldots \\ + b_{nn}(x_n - h_n)^2 + 2b_{12}(x_1 - b_1)(x_2 - b_2) \\ - \ldots + 2b_{n-1,n}(x_{n-1} - h_{n-1})(x_n - h_n) \quad . \quad . \quad \text{(lxxxix.)}. \end{aligned}$$

Here, as in the value of Z, all the constants $b_{11}, b_{22}, \ldots b_{nn}, b_{12} \ldots b_{n-1,n}$ are known in terms of the variations and correlations. If there be selection of q organs only out of the n, then $b_{q+1,q+1} \ldots b_{nn}, b_{1,q+1}, b_{2,q+1}, \ldots b_{n-1,n}$, will all be zero. Since by Equation (lxxxviii.) $p = z/Z$, it follows that the function f which defines the probability of survival must be of the normal exponential type, or

---

* I propose to deal in another memoir with the important problems of slow selection during rapid growth, and of secular selection during several generations.

$$f(x_1 - k_1, x_2 - k_2, \dots x_n - k_n) = \text{expt.} - \tfrac{1}{2}\{a_{11}(x_1 - k_1)^2 + a_{22}(x_2 - k_2)^2 + \dots$$
$$+ 2a_{12}(x_1 - k_1)(x_2 - k_2) - \dots + 2a_{n-1,n}(x_{n-1} - k_{n-1})(x_n - k_n)\} \quad . \quad \text{(xc.)}.$$

Thus, to determine the probability of survival, we require to know the values of the a's and k's in terms of the b's, h's, and c's. The shortest method of finding p_0 is to put $x_1 = k_1$, $x_2 = k_2$, $\dots$ $x_n = k_n$, and then note that:

$$p_0 = \frac{z(x_1 = k_1, x_2 = k_2, \dots x_n = k_n)}{Z(x_1 = k_1, x_2 = k_2, \dots x_n = k_n)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xci.)}.$$

Since $p = z/Z$, and $z = p/Z^{-1}$,

we see that the relations for p, given z and Z, and for z, given p and Z, are cyclicly interchangeable if at the same time we change c_{11} to $-c_{11}$, c_{22} to $-c_{22}$ $\dots$ c_{12} to $-c_{12}$ $\dots$ $c_{n-1,n}$ to $-c_{n-1,n}$. If $\sigma_1, \sigma_2, \dots \sigma_n$ be the standard deviations of the unselected population, this amounts to changing $\sigma_1, \sigma_2, \dots \sigma_n$ to $\sqrt{-1}\sigma_1$, $\sqrt{-1}\sigma_2$, $\dots$ $\sqrt{-1}\sigma_n$ respectively. Thus the results which give the probability of survival in terms of unselected and selected populations can always by an easy interchange be used to obtain the selected population from a knowledge of the unselected population and of the probability of survival.

Let the unselected population be defined by $m_1, m_2, \dots m_n$, $\sigma_1, \sigma_2, \dots \sigma_n$, and r_{12}, r_{13}, r_{23}, $\dots$ $r_{n-1,n}$.

Let the selected population be defined by $m_1 + \mathrm{H}_1$, $m_2 + \mathrm{H}_2$, $\dots$ $m_n + \mathrm{H}_n$, Σ_1, Σ_2, $\dots$ Σ_n, and r_{12}, r_{13}, r_{23}, $\dots$ $\mathrm{r}_{n-1,n}$.

Let the constants of the probability of survival function, or $a_{11}, a_{22}, \dots a_{nn}$, $a_{12} \dots a_{n-1,n}$, be expressed by $\bar{s}_1, \bar{s}_2, \dots \bar{s}_n$, $\bar{\rho}_{12}$, $\bar{\rho}_{13}$, $\dots$ $\bar{\rho}_{n-1,n}$ as if it were a normal correlation surface.*

Then the problem will be solved, if we know the k's, $\bar{s}$'s, and $\bar{\rho}$'s in terms of the σ's, r's, H's, Σ's, and r's.

Equating the squares, products, and linear terms in the x's in the equation $p = z/Z$, we have at once the system:

$$\left.\begin{aligned} a_{uu} &= b_{uu} - c_{uu}, \\ a_{uv} &= b_{uv} - c_{uv}, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xcii.)},$$

for all values of u and v from 1 to n.

$$-a_{v1}k_1 - a_{v2}k_2 - \dots - a_{vv}k_v - \dots - a_{vn}k_n$$
$$= -b_{v1}\mathrm{H}_1 - b_{v2}\mathrm{H}_2 - \dots - b_{vv}\mathrm{H}_v - \dots - b_{vn}\mathrm{H}_n \quad . \quad . \quad \text{(xciii.)},$$

for all values of v from 1 to n.

---

* These must not be confused with the $s_1, s_2, \dots s_q$, $\rho_{12}, \rho_{13}, \rho_{23}, \dots \rho_{q,q-1}$ constants of the q selected organs of the previous discussion. The new quantities may be in part *imaginary*.

If we now substitute from the first two equations for the a's we find:

$$\begin{aligned}(c_{r1} - b_{r1})\,k_1 + (c_{r2} - b_{r2})\,k_2 + \ldots + (c_{rr} - b_{rr})\,k_r + \ldots + (c_{rn} - b_{rn})\,k_n \\ = - b_{r1}\mathrm{H}_1 - b_{r2}\mathrm{H}_2 - \ldots - b_{rv}\mathrm{H}_v - \ldots - b_{rn}\mathrm{H}_n \quad . \quad . \quad \text{(xciv.)}.\end{aligned}$$

Now let Δ be the determinant

$$\begin{vmatrix} c_{11} - b_{11}, & c_{12} - b_{12}, & c_{13} - b_{13}, & \ldots & c_{1n} - b_{1n} \\ c_{21} - b_{21}, & c_{22} - b_{22}, & c_{23} - b_{23}, & \ldots & c_{2n} - b_{2n} \\ c_{31} - b_{31}, & c_{32} - b_{32}, & c_{33} - b_{33}, & \ldots & c_{3n} - b_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ c_{n1} - b_{n1}, & c_{n2} - b_{n2}, & c_{n3} - b_{n3}, & \ldots & c_{nn} - b_{nn} \end{vmatrix}$$

and Δ_{uv} the minor of its uvth constituent.

Then we have at once:

$$\begin{aligned}\Delta \times k_r = {} & \Delta_{r1}(b_{11}\mathrm{H}_1 + b_{12}\mathrm{H}_2 + \ldots + b_{1n}\mathrm{H}_n) \\ & + \Delta_{r2}(b_{21}\mathrm{H}_1 + b_{22}\mathrm{H}_2 + \ldots + b_{2n}\mathrm{H}_n) \\ & + \Delta_{r3}(b_{31}\mathrm{H}_1 + b_{32}\mathrm{H}_2 + b_{33}\mathrm{H}_3 \ldots + b_{3n}\mathrm{H}_n) \\ & \ldots\ldots\ldots\ldots\ldots\ldots \\ & + \Delta_{rn}(b_{n1}\mathrm{H}_1 + b_{n2}\mathrm{H}_2 + \ldots + b_{nn}\mathrm{H}_n).\end{aligned}$$

Or:

$$\begin{aligned}\Delta \times k_v = {} & \mathrm{H}_1(b_{11}\Delta_{r1} + b_{12}\Delta_{r2} + b_{13}\Delta_{r3} + \ldots + b_{1n}\Delta_{rn}) \\ & + \mathrm{H}_2(b_{21}\Delta_{r1} + b_{22}\Delta_{r2} + b_{23}\Delta_{r3} + \ldots + b_{2n}\Delta_{rn}) \\ & + \mathrm{H}_3(b_{31}\Delta_{r1} + b_{32}\Delta_{r2} + b_{33}\Delta_{r3} + \ldots + b_{3n}\Delta_{rn}) \\ & + \ldots\ldots\ldots\ldots\ldots\ldots \\ & + \mathrm{H}_n(b_{n1}\Delta_{r1} + b_{n2}\Delta_{r2} + b_{n3}\Delta_{r3} + \ldots + b_{nn}\Delta_{rn}) \quad . \quad . \quad . \quad \text{(xcv.)}.\end{aligned}$$

Thus k_v is determined in terms of $\mathrm{H}_1, \mathrm{H}_2, \ldots \mathrm{H}_n$, which define the maximum frequency of survival.

In a similar manner by making the proper changes indicated above we have:

$$\begin{aligned}\Delta' \times \mathrm{H}_v = {} & k_1(a_{11}\Delta'_{r1} + a_{12}\Delta'_{r2} + a_{13}\Delta'_{r3} + \ldots + a_{1n}\Delta'_{rn}) \\ & + k_2(a_{21}\Delta'_{r1} + a_{22}\Delta'_{r2} + a_{23}\Delta'_{r3} + \ldots + a_{2n}\Delta'_{rn}) \\ & + k_3(a_{31}\Delta'_{r1} + a_{32}\Delta'_{r3} + a_{33}\Delta'_{r3} + \ldots + a_{3n}\Delta'_{rn}) \\ & + \ldots\ldots\ldots\ldots\ldots\ldots \\ & + k_n(a_{n1}\Delta'_{r1} + a_{n2}\Delta'_{r3} + a_{n3}\Delta'_{r3} - \ldots + a_{nn}\Delta'_{rn}) \quad . \quad . \quad . \quad \text{(xcvi.)},\end{aligned}$$

where Δ' is the determinant

$$\begin{vmatrix} (c_{11}+a_{11}), & (c_{12}+a_{12}), & (c_{13}+a_{13}) & \ldots & (c_{1n}+a_{1n}) \\ (c_{21}+a_{21}), & (c_{22}+a_{22}), & (c_{23}+a_{23}) & \ldots & (c_{2n}+a_{2n}) \\ (c_{31}+a_{31}), & (c_{32}+a_{32}), & (c_{33}+a_{33}) & \ldots & (c_{3n}+a_{3n}) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ (c_{n1}+a_{n1}), & (c_{n2}+a_{n2}), & (c_{n3}+a_{n3}) & \ldots & (c_{nn}+a_{nn}) \end{vmatrix}$$

and Δ'_{uv} the minor of its uv^{th} constituent.

Now it will be clear from these results that as a general rule it is impossible for k_v to be equal to H_v. In other words: *The individual most frequently met with in any given selected community*, i.e., *the mediocre individual, is not the individual fittest to survive.* It is only in the limiting case of natural selection being so stringent that one type of individual alone is able to survive, that the fittest class has a numerical majority over any other class of the community. This seems to me an important, algebraically almost self-obvious truth, and yet one which is very much obscured by the use of such a phrase as the "survival of the fittest."

Of course, if there be continuous selection, or an environment so stable that the probability of survival remains constant for a long period, there will be a gradual approach, never theoretically an actual identification of the mediocre and the fittest. But in actual nature the environment, at any rate so far as it depends on climatological conditions, must have a long period as compared with the vital and reproductive periods of innumerable forms of life. A hard winter, a drought, a flood, a famine, a plague or epidemic of any kind, even if fairly stringent, will rarely, if ever, render the most frequently surviving individual identical with the individual who is fittest to survive.* Still less will this identity take place in the many processes of artificial selection, which are becoming and will more and more become valuable laboratory aids in our appreciation of the action of natural selection. The divergence between the most frequently surviving and the fittest individual is measured by the above formulæ for the k's in terms of the H's.†

To complete the solution, the a's must be found from the equations of type $a_{uv} = b_{uv} - c_{uv}$, and then from the a's the $\bar{s}$'s and $\bar{\rho}$'s follow by the well-known determinants for multiple correlation: see our Equations (xi.) and (xii.).

Throughout the earlier part of this memoir I have used only the surface of selection, but the above investigation will enable us whenever desired to replace it by the probability of survival. I will illustrate this by obtaining the formulæ suitable to the simpler cases.

---

* We badly want a name for the selection which acts for a short time and rapidly modifies the adult population. It is practically the type of selection considered in this paper. It is epidemic or catastrophic in character.

† The point is of considerable importance, for more than one influential writer has spoken of the result of natural selection as the preservation of the type the mortality of which is *least* under the given conditions.

(12.) CASE (i.).—*Selection of a Single Organ only.*

The original population is given by

$$Z = Z_0 e^{-x^2(2\sigma^2)},$$

and the curve of survivors by

$$z = z_0 e^{-(x-H)^2/2\Sigma^2}.$$

The probability of survival is

$$p = p_0 e^{-(x-k)^2/2\bar{s}^2},$$

where we easily find, if $\Sigma/\sigma = \lambda$,

$$k = H/(1-\lambda^2) \quad \text{. . . . . (xcvii.)},$$

$$\bar{s} = \Sigma/\sqrt{1-\lambda^2} \quad \text{. . . . . (xcviii.)},$$

and

$$p_0 = \frac{z_0}{Z_0} e^{\frac{H^2}{2\sigma^2(1-\lambda^2)}} \quad \text{. . . . . (xcx.)}.$$

As an illustration consider a selection from modern French peasants, which should reduce the mean and variability of their cephalic index to those of the Libyan race.

French peasants :—

$m = 79{\cdot}786, \qquad \sigma = 3{\cdot}841.$

Libyans :—

$m + H = 72{\cdot}938, \qquad \Sigma = 2{\cdot}885.$

Hence : $\quad H = -6{\cdot}848, \qquad \lambda = {\cdot}7511.$

These give : $\quad k = -15{\cdot}712, \qquad \bar{s} = 4{\cdot}370.$

Thus for such a change as 7 points in the cephalic index to take place by selection* we should have to make the "fittest to survive" of such a ridiculously low cephalic index as 64·074, and such a high variation as 4·370.

We find $\quad p_0 = 51{\cdot}0474\, n/N,$

and accordingly the probability of survival given by

$$p = 51{\cdot}0474 \frac{n}{N} e^{-(x+15{\cdot}712)^2/(38{\cdot}1939)},$$

where N are the number of Frenchmen converted into n Libyans so far as cephalic index is concerned.

I have purposely taken a somewhat extreme case of selection in order to illustrate how widely the most frequently surviving individual can diverge from the fittest.

In this case, if the chances of survival (i.) of the fittest, (ii.) of the individuals most frequent after selection, and (iii.) of the individuals most frequent before selection, be C_1, C_2, and C_3 respectively, we have :

---

[*] This is, of course, supposing the change to occur by catastrophic selection and not by a continuous secular selection, see footnote preceding page.

$$C_1 : C_2 : C_3 :: 1 : \cdot 12782 : \cdot 00156,$$

or, the chances of survival of an individual of the fittest type would be about eight times as great as those of an individual of the most frequent type after selection and about 700 times as great as those of an individual of the most frequent type before selection. If ν_1, ν_2, ν_3 be the numbers after selection in the three classes* of the fittest to survive, the most frequent after selection and the most frequent before selection, we find

$$\nu_1 : \nu_2 : \nu_3 :: \cdot 00892 : 1 : \cdot 05978.$$

In other words, the most numerous type before selection is still after selection about 6·7 times as numerous as the type with the least mortality, and this latter type is only about $\frac{1}{111}$ as numerous as the type to be most frequently met with after selection has taken place.

Thus, although there would have been a very great evolution in cephalic index, due to a fairly stringent selection, the fittest to survive would always have formed but a small fraction of the dominant type. Even if we were to replace the selection here considered by a gradual evolution spread over several generations, we should still reach in the main the same conclusion, *i.e.*, *that natural selection never proceeds by the survival of the fittest, or the survival of those with the least death-rate.* These will always remain a small fraction of the community—they are the goal, but often the very distant goal, to which selection tends to shift the population.

(13.) CASE (ii.)—*Selection of Two Organs.*

In this case let the surface of survivors be :

$$z = z_0 \text{ expt.} - \tfrac{1}{2}\left\{\frac{(x_1 - \mathrm{H}_1)^2}{\Sigma_1^2(1 - \mathfrak{r}_{12}^2)} - \frac{2\mathfrak{r}_{12}(x_1 - \mathrm{H}_1)(x_2 - \mathrm{H}_2)}{\Sigma_1\Sigma_2(1 - \mathfrak{r}_{12}^2)} + \frac{(x_2 - \mathrm{H}_2)^2}{\Sigma_2^2(1 - \mathfrak{r}_{12}^2)}\right\} \quad . \quad \text{(c.)},$$

the original population :

$$Z = Z_0 \text{ expt.} - \tfrac{1}{2}\left\{\frac{x_1^2}{\sigma_1^2(1 - r_{12}^2)} - \frac{2r_{12}x_1x_2}{\sigma_1\sigma_2(1 - r_{12}^2)} + \frac{x_2^2}{\sigma_2^2(1 - r_{12}^2)}\right\} \quad . \quad . \quad \text{(ci.)},$$

and the curve of probability of survival :

$$p = p_0 \text{ expt.} - \tfrac{1}{2}\left\{\frac{(x_1 - k_1)^2}{\bar{s}_1^2(1 - \bar{\rho}_{12}^2)} - \frac{2\bar{\rho}_{12}(x_1 - k_1)(x_2 - k_2)}{\bar{s}_1\bar{s}_2(1 - \bar{\rho}_{12}^2)} + \frac{(x_2 - k_2)^2}{\bar{s}_2^2(1 - \bar{\rho}_{12}^2)}\right\} . \quad \text{(cii.)}.$$

Since :

$$p = z/Z,$$

---

* By individuals of a type or class is meant here, as elsewhere in this section, all the group falling within some small definite range of variation lying round a particular value of the organ (*e.g.*, m, $m + \mathrm{H}$, or $m + k$), which defines the type or class.

we find at once

$$\frac{1}{\bar{s}_1^2(1-\bar{\rho}_{12}^2)}=\frac{1}{\Sigma_1^2(1-\mathfrak{r}_{12}^2)}-\frac{1}{\sigma_1^2(1-r_{12}^2)}\quad.\ .\ .\ .\ .\ .\quad\text{(ciii.)}.$$

$$\frac{1}{\bar{s}_2^2(1-\bar{\rho}_{12}^2)}=\frac{1}{\Sigma_2^2(1-\mathfrak{r}_{12}^2)}-\frac{1}{\sigma_2^2}\frac{1}{(1-r_{12}^2)}\quad.\ .\ .\ .\ .\ .\quad\text{(civ.)}.$$

$$\frac{1}{\bar{s}_1\bar{s}_2}\frac{\bar{\rho}_{12}}{1-\bar{\rho}_{12}^2}=\frac{\mathfrak{r}_{12}}{\Sigma_1\Sigma_2(1-\mathfrak{r}_{12}^2)}-\frac{r_{12}}{\sigma_1\sigma_2(1-r_{12}^2)}\quad.\ .\ .\ .\ .\ .\quad\text{(cv.)}.$$

$$\frac{k_1}{\bar{s}_1^2(1-\bar{\rho}_{12}^2)}-\frac{k_2\bar{\rho}_{12}}{\bar{s}_1\bar{s}_2(1-\bar{\rho}_{12}^2)}=\frac{\mathrm{H}_1}{\Sigma_1^2(1-\mathfrak{r}_{12}^2)}-\frac{\mathrm{H}_2\mathfrak{r}_{12}}{\Sigma_1\Sigma_2(1-\mathfrak{r}_{12}^2)}\quad.\ .\ .\quad\text{(cvi.)}.$$

$$-\frac{k_1\bar{\rho}_{12}}{\bar{s}_1\bar{s}_2(1-\bar{\rho}_{12}^2)}+\frac{k_2}{\bar{s}_2^2(1-\bar{\rho}_{12}^2)}=-\frac{r_{12}\mathrm{H}_1}{\Sigma_1\Sigma_2(1-\mathfrak{r}_{12}^2)}+\frac{\mathrm{H}_2}{\Sigma_2^2(1-\mathfrak{r}_{12}^2)}\quad.\ .\quad\text{(cvii.)}.$$

Let $\Sigma_1/\sigma_1=\lambda_1$, $\Sigma_2/\sigma_2=\lambda_2$ measure the stringency of the selection, and $\mu=\sqrt{\dfrac{1-\mathfrak{r}_{12}^2}{1-r_{12}^2}}$ measure the change in correlation.* Then solving the above equations we find :

$$\bar{\rho}_{12}=\frac{\mathfrak{r}_{12}-r_{12}\mu^2\lambda_1\lambda_2}{\sqrt{1-\mu^2\lambda_1^2}\sqrt{1-\mu^2\lambda_2^2}}\quad.\ .\ .\ .\ .\ .\ .\ .\quad\text{(cviii.)},$$

$$\bar{s}_1=\frac{\Sigma_1\sqrt{1-r_{12}^2-(1-\mathfrak{r}_{12}^2)\lambda_2^2}}{\sqrt{1-r_{12}^2-\lambda_1^2-\lambda_2^2+(1-\mathfrak{r}_{12}^2)\lambda_1^2\lambda_2^2+2\mathfrak{r}_{12}r_{12}\lambda_1\lambda_2}}\quad.\ .\ .\quad\text{(cix.)},$$

$$\bar{s}_2=\frac{\Sigma_2\sqrt{1-r_{12}^2-(1-\mathfrak{r}_{12}^2)\lambda_1^2}}{\sqrt{1-r_{12}^2-\lambda_1^2-\lambda_2^2+(1-\mathfrak{r}_{12}^2)\lambda_1^2\lambda_2^2+2\mathfrak{r}_{12}r_{12}\lambda_1\lambda_2}}\quad.\ .\ .\quad\text{(cx.)},$$

$$\frac{k_1}{\sigma_1}=\frac{\mathrm{H}_1}{\sigma_1}\frac{1-r_{12}^2-\lambda_2^2+\mathfrak{r}_{12}r_{12}\lambda_1\lambda_2}{\beta}+\frac{h_2}{\sigma_2}\frac{\mathfrak{r}_{12}\lambda_2\lambda_1-r_{12}\lambda_1^2}{\beta}\quad.\ .\ .\ .\quad\text{(cxi.)},$$

$$\frac{k_2}{\sigma_2}=\frac{h_1}{\sigma_1}\frac{\mathfrak{r}_{12}\lambda_1\lambda_2-r\lambda_2^2}{\beta}+\frac{h_2}{\sigma_2}\frac{1-r_{12}^2-\lambda_1^2+\mathfrak{r}_{12}r_{12}\lambda_1\lambda_2}{\beta}\quad.\ .\ .\ .\ .\quad\text{(cxii.)},$$

where

$$\beta=1-r_{12}^2-\lambda_1^2-\lambda_2^2+(1-\mathfrak{r}_{12}^2)\lambda_1^2\lambda_2^2+2\mathfrak{r}_{12}r_{12}\lambda_1\lambda_2.$$

Similarly, if the original population and the curve of probability of surviving or of survival rates be given, we have to find the selected population :

$$\mathfrak{r}_{12}=\frac{\bar{\rho}_{12}+r_{12}\nu^2\kappa_1\kappa_2}{\sqrt{1+\nu^2\kappa_1^2}\sqrt{1+\nu^2\kappa_2}}\quad.\ .\ .\ .\ .\ .\ .\ .\quad\text{(cxiii.)},$$

---

* If $r_{12}=\cos d$, $\mathfrak{r}_{12}=\cos \mathrm{D}$, $\mu=\sin \mathrm{D}/\sin d$. The quantity D has been conveniently termed the "divergence" by Mr. SHEPPARD. Hence μ is the ratio of the sines of the selected and unselected divergences. The above formula for ρ_{12} can be at once changed into one suitable for trigonometrical logarithmic calculation. Let $\sin\alpha_1=\mu\lambda_1$, $\sin\alpha_2=\mu\lambda_2$, and $\bar{\rho}_{12}=\cos\delta$; then, if Δ be the side of the spherical triangle, of which α_1, α_2 are the other sides and δ the included angle :

$$\sin\tfrac{1}{2}\delta=\sqrt{\left\{\frac{\sin\frac{1}{2}(\mathrm{D}-\Delta)\sin\frac{1}{2}(\mathrm{D}+\mathrm{A})}{\cos\alpha_1\cos\alpha_2}\right\}}.$$

where $$\nu = \sqrt{\frac{1-\bar{\rho}_{12}^{\,2}}{1-r_{12}^{\,2}}}, \quad \kappa_1 = s_1/\sigma_1 \quad \text{and} \quad \kappa_2 = s_2/\sigma_2,$$

$$\Sigma_1 = \frac{\bar{s}_1\sqrt{1-r_{12}^{\,2}+(1-\bar{\rho}_{12}^{\,2})\,\kappa_2^{\,2}}}{\sqrt{1-r_{12}^{\,2}+\kappa_1^{\,2}+\kappa_2^{\,2}+(1-\bar{\rho}_{12}^{\,2})\,\kappa_1^{\,2}\kappa_2^{\,2}-2\bar{\rho}_{12}r_{12}\kappa_1\kappa_2}} \quad . \quad . \quad . \qquad \text{(cxiv.)},$$

$$\Sigma_2 = \frac{\bar{s}_2\sqrt{1-r_{12}^{\,2}+(1-\bar{\rho}_{12}^{\,2})\,\kappa_1^{\,2}}}{\sqrt{1-r_{12}^{\,2}+\kappa_1^{\,2}+\kappa_2^{\,2}+(1-\bar{\rho}_{12}^{\,2})\,\kappa_1^{\,2}\kappa_2^{\,2}-2\bar{\rho}_{12}r_{12}\kappa_1\kappa_2}} \quad . \quad . \quad . \qquad \text{(cxv.)},$$

$$\frac{\mathrm{H}_1}{\sigma_1} = \frac{k_1}{\sigma_1}\,\frac{1-r_{12}^{\,2}+\kappa_2^{\,2}-\bar{\rho}_{12}r_{12}\kappa_1\kappa_2}{\gamma} + \frac{k_2}{\sigma_2}\,\frac{r\kappa_1^{\,2}-\rho_{12}\kappa_1\kappa_2}{\gamma}. \quad . \quad . \quad . \qquad \text{(cxvi.)},$$

$$\frac{\mathrm{H}_2}{\sigma_2} = \frac{k_1}{\sigma_1}\,\frac{r_{12}\kappa_2^{\,2}-\bar{\rho}_{12}\kappa_1\kappa_2}{\gamma} + \frac{k_2}{\sigma_2}\,\frac{1-r_{12}^{\,2}+\kappa_1^{\,2}-\bar{\rho}_{12}r_{12}\kappa_1\kappa_2}{\gamma}. \quad . \quad . \quad . \qquad \text{(cxvii.)},$$

where $$\gamma = 1 - r_{12}^{\,2} + \kappa_1^{\,2} + \kappa_2^{\,2} + (1-\bar{\rho}_{12}^{\,2})\,\kappa_1^{\,2}\kappa_2^{\,2} - 2\bar{\rho}_{12}r_{12}\kappa_1\kappa_2.$$

(14.) *Illustration.*—The following results are taken from the paper by Miss ALICE LEE and myself already cited:—

	French ♂.	Aino ♂.	
Femur	$m_1 =$ 45·228 centims.	40·770 centims.	$= m_1 + \mathrm{H}$.
	$\sigma_1 =$ 2·372 ,,	1·898 ,,	$= \Sigma_1$.
Humerus	$m_2 =$ 33·010 ,,	29·502 ,,	$= m_2 + \mathrm{H}_2$.
	$\sigma_2 =$ 1·538 ,,	1·343 ,,	$= \Sigma_2$.
Correlation of Femur and Humerus . .	$r_{12} =$ ·8421	·8584	$= \mathrm{r}_{12}$.

As indicated by the symbols above, let us select from the French a population having the same femur and humerus relations as the Aino.

We have at once:

$$\mathrm{H}_1 = -\,4{\cdot}458, \qquad \mathrm{H}_2 = -\,3{\cdot}508,$$
$$\lambda_1 = {\cdot}80017, \qquad \lambda_2 = {\cdot}87321,$$
$$\mu^2 = {\cdot}9047.$$

Whence we find:

$$\bar{\rho}_{12} = {\cdot}9027, \qquad \bar{s}_1 = 3{\cdot}4870, \qquad \bar{s}_2 = 2{\cdot}8736,$$
$$k_1 = -\,17{\cdot}8464, \qquad k_2 = -\,15{\cdot}8547,$$
$$p_0 = 208{,}425\,\frac{n}{\mathrm{N}},$$

where n are the number of Ainos which can be obtained from N Frenchmen.

We have accordingly the following form for the surface of fitness to survive:

$$p = 208{,}425 \frac{n}{\mathrm{N}} \text{ expt.} - \tfrac{1}{2}\{\cdot 222056\,(x + 17\cdot 8464)^2 + \cdot 326{,}960\,(y + 15\cdot 8547)^2 - \cdot 486{,}450\,(x + 17\cdot 8464)\,(y + 15\cdot 8547)\}.$$

Now it is clear that if we wanted by a "catastrophic" selection to convert the French into something resembling the Aino, we should have to give the least death-rate to those French with femur corresponding to $m_1 + k_1$ and humerus to $m_2 + k_2$, or to the dwarfs with femur $= 27\cdot 382$ centims. and humerus $= 17\cdot 155$ centims. ! By no other means could we shift the modal value of the French population down as low as the Aino modal value. The physical meaning of this is that we have been compelled to put on an excessive death-rate for the bigger Frenchmen.

An interesting point of our work is that

$$k_1 = 1\cdot 2514\,\mathrm{H}_1 + 3\cdot 4971\,\mathrm{H}_2,$$

$$k_2 = -1\cdot 0266\,\mathrm{H}_1 + 5\cdot 8242\,\mathrm{H}_2,$$

whence we see that while a selective reduction of humerus is far more effective in reducing both femur and humerus centres of survival than a reduction of femur, a selective reduction of femur occurring contemporaneously with that of the humerus actually tends to *raise* the centre of the humerus, *i.e.*, the coefficient of H_1 is *negative*.

Now let us consider the frequency of survivors per unit length, say centimetre of femur and humerus, at different points. The surface of survivors, *i.e.*, the Aino population, is

$$z = \frac{n}{2\pi\Sigma_1\Sigma_2\sqrt{(1-\mathfrak{r}_{12}{}^2)}}\,e^{-\frac{1}{2}\frac{1}{(1-\mathfrak{r}_{12}{}^2)}\left\{\frac{(x-\mathrm{H}_1)^2}{\Sigma_1{}^2} - \frac{2\mathfrak{r}_{12}(x-\mathrm{H}_1)(y-\mathrm{H}_2)}{\Sigma_1\Sigma_2} + \frac{(y-\mathrm{H}_2)^2}{\Sigma_2{}^2}\right\}}.$$

If we put $x = 0$, $y = 0$ we have the frequency after selection of the original population type ; if we put $x = \mathrm{H}_1$, $y = \mathrm{H}_2$ we have the frequency after selection of the new population type ; and if we put $x = k_1$, $y = k_2$, we shall have the frequency after selection of those best fitted to survive. If these frequencies be ν_3, ν_2, ν_1 respectively, we find on substituting the numerical values that

$$\nu_1 : \nu_2 : \nu_3 :: \cdot 117/10^{18} : 1 : \cdot 032289.$$

Thus the most frequent type of the new population is now about thirty times as frequent as the old most frequent type, while the type most fitted to survive has practically no existence at all. It probably lies outside the actual boundary of the French population.

Here really arises the question as to how we are, in any actual problem, to fix the ratio of n to N, or, what amounts to the same thing, to fix a practical boundary to a given population. Such a boundary must be *conventional*, but I think that for practical purposes we are quite safe if we assume that an individual who occurs only once per thousand can produce no effect on the physical evolution of the population as a whole.

Now the form of a correlation-surface for two organs, x and y, is

$$z = \frac{N}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\, e^{-\frac{1}{2}\frac{1}{1-r^2}\left\{\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right\}}.$$

Let us write $\kappa^2 = \frac{1}{1-r^2}\left\{\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right\}$; then $\kappa = a$ constant gives a series of similar ellipses which are the contour lines of the surface, or lines of equal frequency, *i.e.*, giving individuals with equal probability of occurrence. Let the equation to these contour lines referred to their principal axes be

$$\kappa^2 = \frac{X^2}{A^2} + \frac{Y^2}{B^2}.$$

Then we have at once:

$$\frac{1}{A^2} + \frac{1}{B^2} = \frac{1}{\sigma_1^2(1-r^2)} + \frac{1}{\sigma_2^2(1-r^2)} = \frac{1}{1-r^2}\,\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2},$$

$$\frac{1}{A^2B^2} = \frac{1}{\sigma_1^2\sigma_2^2(1-r^2)^2} - \frac{r^2}{\sigma_1^2\sigma_2^2(1-r^2)^2} = \frac{1}{(1-r^2)\,\sigma_1^2\sigma_2^2}$$

or,

$$AB = \sigma_1\sigma_2\sqrt{1-r^2}, \qquad A^2 + B^2 = \sigma_1^2 + \sigma_2^2.$$

Further, if ϕ be the angle the A principal axis makes with the axis of x, we have:*

$$\tan 2\phi = 2r\sigma_1\sigma_2/(\sigma_1^2 - \sigma_2^2).$$

These fully determine the principal axes of the frequency surface. Now consider the frequency between the elliptic cylinders corresponding to κ and $\kappa + \delta\kappa$; we have it

$$= z \times 2\pi AB\kappa d\kappa = z \times 2\pi\sigma_1\sigma_2\sqrt{1-r^2}\,\kappa d\kappa = Ne^{-\frac{1}{2}\kappa^2}\kappa d\kappa.$$

Hence, if N_κ be the frequency outside any contour κ,

$$N_\kappa = N\int_\kappa^\infty e^{-\frac{1}{2}\kappa^2}\kappa d\kappa = Ne^{-\frac{1}{2}\kappa^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(cxviii.)}.$$

For N_κ to be $\frac{1}{1000}$ of N we have simply

$$\kappa^2 = \frac{6}{\log e}, \quad \text{whence } \kappa = 3{\cdot}716{,}923.$$

---

* For easy calculation put $\gamma = \sqrt{\sigma_1^2 + \sigma_2^2}$, $\tan\psi = \sigma_2/\sigma_1$. Then we have at once if $r = \cos D$:

$$A = \gamma\cos\chi, \qquad B = \gamma\sin\chi,$$

where:

$$\sin 2\chi = \sin 2\psi \sin D, \qquad \tan 2\phi = \tan 2\psi \cos D.$$

This will enable us to determine our conventional boundary to effective population.

Now let us refer our non-selected and selected populations to their centres and principal axes.

We find for the contour curves :—

	Unselected Population (*i.e.*, French).	Selected Population (*i.e.*, Aino).	Surface of Survival (*i.e.*, Rate of Survival).
x centre (femur)	45·228	40·770	27·382
y centre (humerus) . . .	33·010	29·502	17·155
tan ϕ (slope to x)	·601,3775	·670,1454	·807,3371
Principal axes A . . .	2·7338	2·2514	4·4115
Principal axes B . . .	·7197	·5808	·9775
1 in 1000 limit κA . . .	10·1615	8·3682	—
1 in 1000 limit κB . . .	2·6750	2·1588	—

Referred to its principal axes, the rate of survival is now

$$p = 208{,}425 \, \frac{n}{\mathrm{N}} \, e^{-\frac{1}{2}\left\{\frac{\mathrm{X}^2}{(4\cdot4115)^2} + \frac{\mathrm{Y}^2}{(\cdot9775)^2}\right\}}.$$

Suppose we require to get at least 1000 Aino out of the French population, N, then $n = 1000$. Now suppose the Aino limiting ellipse drawn, then the French population must be sufficiently large to give the individuals inside this ellipse. Now p gets smaller as we go further from the centre of the survival surface. Hence the contour line of the survival surface corresponding to $p = 1$ must be touched *externally* by the limiting contour of the Aino population, in order that we may get at least 1000 Aino out of N Frenchmen. Now, by a graphical construction, I find the major axis of the elliptic contour line of the survival surface which touches the Aino limiting ellipse, is about 11·44. This gives for the parameter κ_1 of this ellipse, $\kappa_1 \times 4\cdot1149 = 11\cdot44$, or $\kappa_1 = 2\cdot5932$. Whence :

$$p_{\kappa_1} = 1 = 208{,}425 \, \frac{1000}{\mathrm{N}} \, e^{-\frac{1}{2}(2\cdot5932)^2}$$

gives the greatest possible value of p and the least possible value of N. Numerically this gives us N = 7,200,000 about, or we should want more than 7,000,000 of Frenchmen to obtain our 1000 Aino by a catastrophic selection. The actual bounding contour line of this least possible number of Frenchmen* has for its major axis 15·285 centims., and it touches the Aino limiting ellipse at the point where it is touched by the survival contour $p = 1$.

Now let us turn the problem round and ask what is the least population of Aino

---

* The least possible to reproduce the Aino, as far as femur and humerus are concerned, to 1 in a 1000 of the population.

from which we could produce 1000 Frenchmen by a catastrophic selection. In this case the surface of survivals is simply obtained by inverting p, and, if N' be the number of Aino, $n' = 1000$ = required number of Frenchmen, we have:

$$p' = \frac{1}{208{,}425}\,\frac{n'}{\mathrm{N}}\,e^{\frac{1}{2}\left\{\frac{x^2}{(4\cdot4115)^2}+\frac{y^2}{(\cdot9775)^2}\right\}}.$$

Here p' gets larger as we go away from the centre of the surface of survivals, and we must therefore make the French limiting ellipse just touched *internally* by the contour line of the surface of survivors for which $p' = 1$. The major axis of this contour line for $p' = 1$ was found by a graphical process to be about 33·91. This gives for κ_2

$$\kappa_2 \times (4\cdot4115) = 33\cdot91, \quad \text{or} \quad \kappa_2 = 7\cdot6867.$$

Thus:

$$p'_{\kappa_2} = 1 = \frac{1000}{208{,}425}\,\frac{1}{\mathrm{N}'}\,e^{\frac{1}{2}(7\cdot6867)^2},$$

leads to $\mathrm{N}' = 32{,}460{,}000{,}000$ about, or we should want upwards of 32,460,000,000 of Aino to produce the 1000 Frenchmen. The bounding contour line of this number of Aino has a major axis of 15·890 centims. about, and touches the French limiting contour in the point in which it is touched by the $p' = 1$ contour of the surface of survivals.

Now the difference between these two unselected populations is very great. We see that to get the Aino a very great number of Frenchmen would have to be exterminated, about 7000 for each Aino selected; but to get the Frenchmen from the Aino an appalling number of Aino would have to be destroyed, upwards of 32,000,000 for each Frenchman selected. Even if the selection were not catastrophic but spread over centuries and centuries, we must recognise what a large consumption of life there must be—individuals destroyed without progeny*—if we are to suppose any highly civilized race like the French produced by selection from an apparently primitive type like the Aino. Indeed, the return journey in this case seems much easier than the upward ascent. Beyond all this we have only made French and Aino alike for two organs, and only for one character of each of them! Allowing for our conventional limit to the population, allowing for the fact that our Aino data are drawn from a very limited population of remarkably small variability, it seems very improbable that the French have ever been produced by selection from a primitive race at all resembling the Aino. The fact that the Aino could be so much more easily obtained by selection from the French seems to indicate that they are rather

---

* Of course, with a secular selection spread over many generations, it is largely the potentiality and not the actuality of life which is destroyed. Still, while the gross number killed among a small primitive community may not be large, the death-rate must still be immense. I hope to return to these points when dealing with secular selection as distinguished from catastrophic selection.

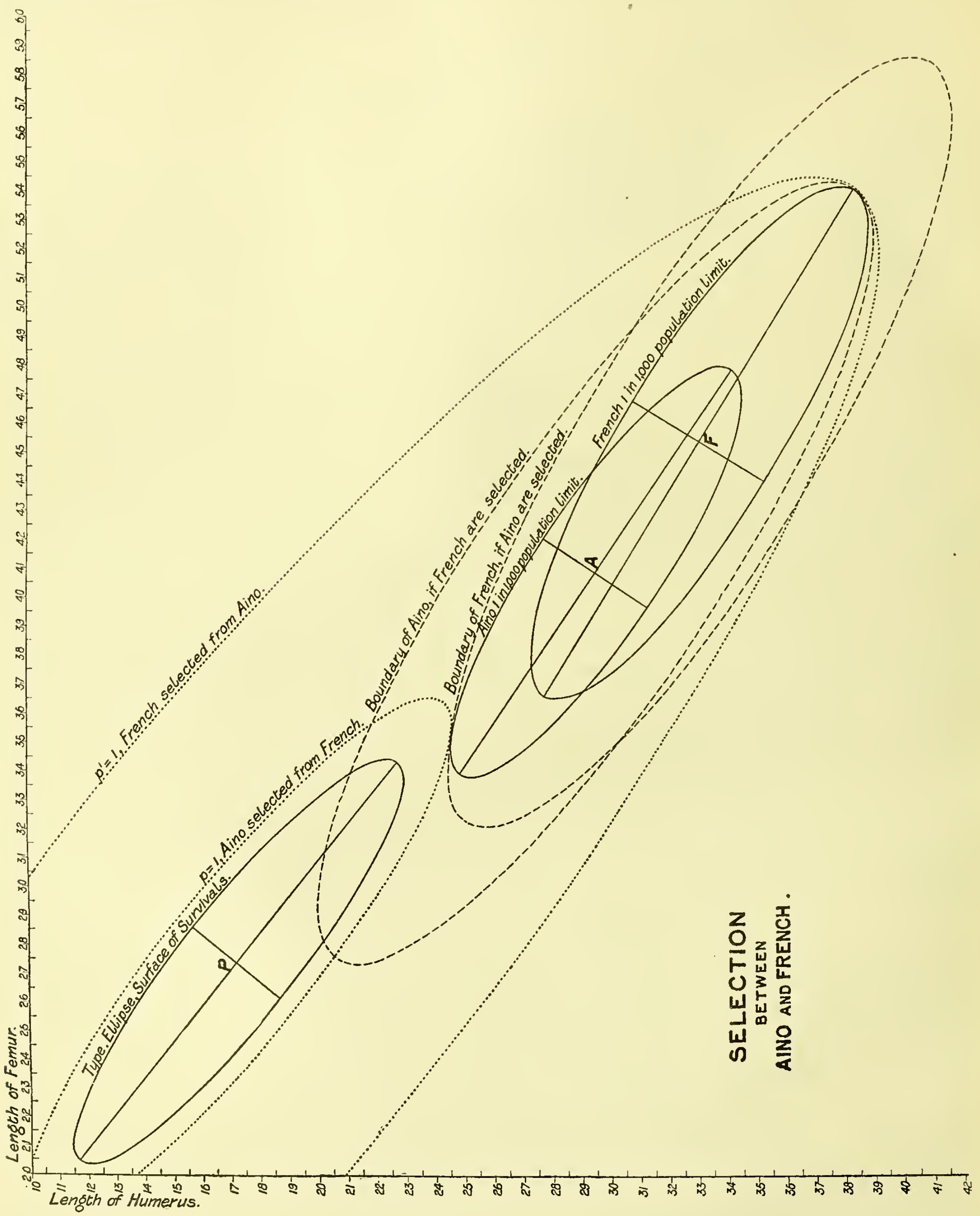
some degenerate offshoot of a race superior to themselves than a sample of the primitive people from which the Circassian races may be supposed to have sprung.

The whole of this discussion is, of course, very hypothetical; no stress whatever is to be laid upon it except as an illustration of method, and a rough appreciation of the vast amount of elimination which must be necessary to evolve one race from a second in the case of organs which we know by measurement to have continuity of variation, and only saltatory changes in pathological cases, which have, as far as we can judge, no influence on the mass-evolution which has produced the local races of man.

But given fair samples of material our method will enable us to determine whether a race A—for of course a limited number of characters—could with less destruction be deduced from a race B, than the race B from A. It will not therefore follow that the path of least selection is that which necessarily was used by Nature. Possibly both A and B have been reached by far less expenditure of material from C. Still it is something definite in the midst of our gropings after truth in problems of descent to have even a rough appreciation of the amount of selective destruction which would arise from alternative suggestions. That is why this special numerical illustration of the surface of survival has been given.

The reader will possibly find the matter rendered somewhat clearer by the diagram. The femur is measured along the horizontal and the humerus along the vertical. A is the type or mean femur-humerus of the Aino population. Within in the continuous ellipse round A the whole Aino population up to 1 in 1000 would fall. F is the type of the French population and the continuous ellipse round F gives the area within which up to 1 in 1000 of the French population fall. Since the diagram is drawn to centimetres of the bones, it will be seen how very small are the limits of variation within both populations. P is the centre of the surface of survivals; for the selection of Aino from French it makes the "fittest to survive." In the case of the selection of the French from the Aino, P is no longer the centre of fitness, but the "centre of unfitness"; the Aino are killed off with an intensity which increases the closer we approach to P. Now it seems to me that these two cases, which are quite distinct in theory, ought to manifest themselves in Nature and require distinguishing names. A race may be modified because a complex of organs with a certain system of values is good for it, or because it is bad for it. The race may be modified because a certain element of it is fittest or because it is unfittest to survive. In the former case we select for survival round the centre, in the latter case we select for destruction. I propose to call these cases *positive* and *negative* selection respectively. It may be said that if there be positive selection in one part of the population there will be negative in another. But the kernel of the matter is in either case the existence of a centre, a definite set of most fit or of most unfit organs, while in positive selection the less fit organs, and in negative selection the more fit organs are distributed over wide areas of the field, and do not reach a maximum of unfitness or a maximum of fitness respectively for any definite individual.

In the diagram we have also drawn the contour line to which the French population must extend if we are to get at least a representative population of 1000 Aino from it, and further the contour line to which the Aino population must extend if we are to get at least a representative population of 1000 French from it. A consideration of



the nature of the contour lines of the surface of survivals shows that the contour lines above referred to, and marked "boundary" in the diagram, must touch the Aino 1 in 1000 limit and the French 1 in 1000 limit respectively at the points in which they are touched by the contour lines $p = 1$ and $p' = 1$ of the corresponding surfaces of survivals. I have already indicated that the major axes of these boundaries are 15·285 for the French and 15·890 for the Aino. The corresponding values of the parameter κ are respectively given by

$$\kappa_F = \left(\frac{15·285}{2·7338}\right) = 5·5911, \qquad \kappa_A = \frac{15·890}{2·2514} = 7·0578.$$

Hence by (cxviii.) we can easily find the frequency of population outside the contours κ_F and κ_A; if these be ν_F and ν_A we have:

$$\nu_F = ·000{,}000{,}163, \qquad \nu_A = ·000{,}000{,}000{,}015.$$

Thus the French population would have to be extended to a boundary in which only about 1 in six millions was excluded, and the Aino population to a boundary excluding only 15 in the billion! The boundaries of what we may thus term the selection populations are far larger than our conventional boundaries of 1 in 1000 for representative populations. In fact, it would be impossible to select a representative Aino population from a conventional representative French population and *vice versa*—in either case the very exceptional members of French or Aino populations are required to complete the conventional representative populations of Aino or French by selection.

(15.) I have devoted most of my consideration of the surface of survivals to a particular case in which two organs have been selected, and we consider the nature of p which determines the fraction of each group of individuals which survives. I have done this partly because normal surfaces are at best only an approximate representation of our selectable and selected distributions, and partly because I have thought a concrete case would best bring out the general points of investigations of this kind.

But some little indication of the properties of the surface of survival-rates ought to be indicated here, or it may appear that they have been overlooked. While the contour lines of the correlation frequency surfaces for two organs must be ellipses, this does not follow in the case of the surface of survival-rates. In our illustration they were ellipses, but they may be also parabolas, hyperbolas, or even straight lines. We must not therefore expect to find always a "centre" of positive or negative selection. We may come across a "saddle-back system" of contours with the rate of survival constant along two intersecting lines, but rising in one pair of opposite angles and falling in the other pair. In this case we have fields of negative and positive selection separated by two independent relations between the two organs, which are linear and for which the survival-rate is the same, they may be termed the "critical lines." For one pair of angles the centre is now a "centre of fitness," for the other pair of angles a "centre of unfitness." It seems to me that these critical organic relations may possess considerable biological importance.

If the contour lines of the surface of survival-rates are parabolas, we have really only a limiting case of the centre at a very great distance. It is one in which the fittest (or most unfit) has no practical existence, but there is a direction towards which the rate of survival will be found to be always increasing or decreasing.

If the contour lines of the surface of survival-rates are parallel straight lines, then so long as the deviation in one organ has a certain definite relation to that in the other, the survival-rate will remain constant. In this case the survival-rate will fall uniformly in one direction and remain constant in the direction at right angles to it.

All the cases I have given here can occur just as easily as the elliptic contour system of our illustration and diagram. Each is marked by quite definite biological characteristics, and we may, perhaps, class them as elliptic, hyperbolic, parabolic, and linear selection. Even if the surface of survival-rates be not of the exponential quadric type discussed in this paper, yet to the neighbourhood of each part of it this classification of selection types will apply.

If we pass to more than two organs, then similar considerations will apply; we shall only be reproducing the geometry of quadric surfaces in space of three and higher dimensions. But before we allow ourselves excursions into the higher geometry of the surface of survival-rates, it seems desirable that we should obtain quantitative determinations of this surface by experiments in artificial selection. We shall then be better able to see what part of our geometry will really be of service for the problems of natural selection. The field is too large to be cultivated for merely theoretical interests. We must first determine what parts of it are likely to have practical application to life as we find it, but of death-rates in the case of any living form but man, we are at present sadly ignorant.

Key to Selective Correlation Tables.

Case.	Values of	
	r_{12}.	r_{13}.
(*a*)	0	0
(*b*)	0	·25
(*c*)	0	·5
(*d*)	0	·75
(*e*)	0	1
(*f*)	·25	·25
(*g*)	·25	·5
(*h*)	·25	·75
(*i*)	·25	1
(*j*)	·5	·5
(*k*)	·5	·75
(*l*)	·5	1
(*m*)	·75	·75
(*n*)	·75	1
(*p*)	1	1

Formula :

$$R' = R + 10\,(\Delta_{\mu_1}R)\,\delta\mu_1 + 4\,\{(\Delta_{r_{12}}R)\,\delta r_{12} + (\Delta_{r_{13}}R)\,\delta r_{13} + (\Delta_{r_{23}}R)\,\delta r_{23}\}.$$

Occasionally second differences must be used.

SELECTIVE CORRELATION TABLES.

Values of r_{23} positive.

Table I (A).—$r_{23} = 0$. The possible values are (*a*), (*b*), (*c*), (*d*), (*e*) [limiting] (*f*), (*g*), (*h*), (*j*), (*k*). (*i*), (*m*), (*l*) are needed for interpolation.

[In one case, viz. (*e*), R_0 is indeterminate].

	R_0	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}
(*a*), (*b*), (*c*), (*d*), (*e*)	0	0	0	0	0	0	0	0	0	0	0
(*f*)	–·0667	–·0660	–·0638	–·0603	–·0554	–·0492	–·0417	–·0329	–·0230	–·0120	0
(*g*)	–·1491	·1473	–·1420	–·1333	–·1214	–·1065	–·0891	–·0694	–·0477	–·0245	0
(*h*)	–·2928	–·2879	–·2737	–·2515	–·2228	–·1894	–·1531	–·1151	–·0765	–·0379	0
(*i*)	–∞	–2·5553	–1·2377	–·7809	–·5394	–·3841	–·2722	–·1851	–·1138	–·0531	0
(*j*)	–·3333	–·3289	–·3158	–·2945	–·2658	–·2308	–·1905	–·1461	–·0989	–·0499	0
(*k*)	–·6547	–·6429	–·6089	–·5557	–·4880	–·4104	–·3273	–·2425	–·1585	–·07725	0
(*l*)	–∞	–5·7063	–2·7530	–1·7256	–1·1813	–·83205	–·5819	–·3900	–·2359	–·1082	0
(*m*)	–1·2857	–1·2567	–1·1739	–1·0487	–·8957	–·7297	–·5625	–·4023	–·2539	–·1197	0

Table II (A).—r_{23} = ·25 the possible values are (*a*), (*b*), (*c*), (*d*), (*f*), (*g*), (*h*), (*i*) [limiting] (*j*), (*k*), (*m*). (*e*), (*l*), (*n*) are needed for interpolation.

	R_0	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}
(*a*)	·2500	·2500	·2500	·2500	·2500	·2500	·2500	·2500	·2500	·2500	·25
(*b*)	·2582	·2581	·25785	·2574	·2568	·2561	·2552	·2541	·2529	·2515	·25
(*c*)	·2887	·2882	·2868	·28445	·2813	·27735	·2728	·2676	·26205	·2562	·25
(*d*)	·3780	·37555	·3686	·3578	·3442	·3288	·3125	·2960	·28645	·2645	·25
(*e*)	+∞	2·5000	1·2500	·8333	·6250	·5000	·4167	·3571	·3125	·2778	·25
(*f*)	·2000	·2005	·2021	·2048	·20845	·2131	·21875	·2253	·2328	·2410	·25
(*g*)	·1491	·1503	·1538	·1596	·1676	·1776	·1893	·20265	·21735	·2332	·25
(*h*)	·0976	·09985	·10645	·1170	·1308	·14735	·1659	·1858	·2067	·2282	·25
(*i*)	0	·0258	·0516	·0772	·1027	·1280	·1531	·1779	·2023	·2263	·25
(*j*)	0	·0033	·01315	·0291	·0506	·0769	·10715	·1404	·1758	·2126	·25
(*k*)	–·2182	–·2100	–·18605	–·1486	–·10065	–·0456	·0136	·0745	·1350	·1938	·25
(*l*)	–∞	–2·8243	–1·3192	–·7775	–·4782	–·27735	–·1273	–·0076	·0917	·1765	·25
(*m*)	–·7143	–·6925	–·6304	–·5365	–·4217	–·2973	–·1719	–·0517	·0596	·1602	·25
(*n*)	–∞	–7·3985	–3·4649	–2·06345	–1·3080	–·8220	–·4792	–·2241	–·0280	·1264	·25

Selective Correlation Tables.

Positive values of r_{23}.

Table III. (A). $r_{23} = \cdot 5$. The possible values are (*a*), (*b*), (*c*), (*d*), (*f*), (*g*), (*h*), (*j*), (*k*), (*l*), (*m*) [limiting]. (*e*), (*i*), (*n*), are needed for interpolation.

	R_0.	R_1.	R_2.	R_3.	R_4.	R_5.	R_6.	R_7.	R_8.	R_9.	R_{10}.
(*a*)	·5000	·5000	·5000	·5000	·5000	·5000	·5000	·5000	·5000	·5000	·5000
(*b*)	·5164	·5162	·5157	·51485	·5137	·51215	·5103	·5082	·5057	·5030	·5000
(*c*)	·57735	·5764	·5735	·5689	·5625	·5547	·54555	·5353	·5241	·5123	·5000
(*d*)	·7559	·7511	·7372	·71565	·6884	·6576	·6250	·5924	·5599	·5291	·5000
(*e*)	+ ∞	5·0000	2·5000	1·6667	1·2500	1·0000	·8333	·7143	·6250	·5556	·5000
(*f*)	·4667	·4670	·4681	·46985	·4723	·4754	·4792	·4835	·4885	·4940	·5000
(*g*)	·4472	·4478	·4496	·4525	·45655	·46165	·4677	·4747	·4824	·49095	·5000
(*h*)	·48795	·4876	·4866	·48545	·48445	·4841	·4848	·4867	·4898	·4943	·5000
(*i*)	+ ∞	2·6069	1·34085	·9353	·7448	·6402	·57835	·5408	·5184	·5058	·5000
(*j*)	·3333	·33555	·3421	·35275	·3671	·3846	·4048	·4269	·45055	·4751	·5000
(*k*)	·2182	·2230	·2368	·2585	·2866	·3192	·3546	·3914	·4285	·46485	·5000
(*l*)	0	·0576	·1147	·1707	·2250	·27735	·3273	·3747	·4193	·4611	·5000
(*m*)	− ·14285	− ·12835	− ·08695	− ·0243	·0521	·1351	·2188	·2989	·37305	·44015	·5000
(*n*)	− ∞	− 3·6429	− 1·6219	− ·8707	− ·4475	− ·1644	·0417	·1988	·3219	·4203	·5000

Table IV. (A). $r_{23} = \cdot 75$. The possible values are (*a*), (*b*), (*c*), (*f*), (*g*), (*h*), (*j*), (*k*), (*m*), (*n*) [limiting]. (*d*), (*i*), (*l*), are needed for interpolation.

	R_0.	R_1.	R_2.	R_3.	R_4.	R_5.	R_6.	R_7.	R_8.	R_9.	R_{10}.
(*a*)	·7500	·7500	·7500	·7500	·7500	·7500	·7500	·7500	·7500	·7500	·7500
(*b*)	·7746	·7743	·7736	·7723	·7705	·7682	·7655	·76225	·7586	·7545	·7500
(*c*)	·8660	·8646	·8603	·8533	·8438	·83205	·8183	·8029	·7862	·7685	·7500
(*d*)	1·1339	1·1267	1·1058	1·0735	1·0327	·9864	·9375	·8881	·83985	·7936	·7500
(*f*)	·7333	·7335	·7340	·7349	·73615	·7377	·7396	·7418	·74425	·7470	·7500
(*g*)	·74536	·74536	·74537	·7454	·7455	·7457	·7461	·7467	·7476	·74865	·7500
(*h*)	·8783	·8753	·8668	·8539	·8381	·8209	·80375	·78755	·7730	·76045	·7500
(*i*)	+ ∞	5·1881	2·6301	1·7934	1·3869	1·1523	1·0036	·9038	·8345	·78525	·7500
(*j*)	·6667	·6678	·67105	·6764	·6835	·6923	·7024	·7135	·7253	·7376	·7500
(*k*)	·65465	·6559	·6596	·6656	·67385	·6839	·69555	·7084	·7220	·7359	·7500
(*l*)	+ ∞	2·9396	1·5486	1·1188	·9282	·83205	·78195	·75705	·7469	·7457	·7500
(*m*)	·4286	·4358	·4565	·4878	·5261	·5676	·6094	·6494	·6865	·7201	·7500
(*n*)	0	·1127	·2212	·32205	·4131	·4932	·5625	·6217	·6719	·71425	·7500

Table V. (A). $r_{23} = 1$. The only possible values are $r_{12} = r_{13}$, *i.e.*, (*a*), (*f*), (*j*), (*m*), (*p*). All these give $R = 1$ for all values of the selective intensity λ.

SELECTIVE CORRELATION TABLES.

Values of r_{23} negative.

Table I. (B). This is the same as Table I. (A). $r_{23} = 0$.

The possible values are (*a*), (*b*), (*c*), (*d*), (*e*) [limiting], (*f*), (*g*), (*h*), (*j*), (*k*). Values which are impossible, but which are tabulated for purposes of interpolation, are (*i*), (*m*), (*l*).

Table II. (B). $r_{23} = -\cdot 25$.

The possible values are (*a*), (*b*), (*c*), (*d*), (*f*), (*g*), (*h*), (*j*). Values which are impossible, but are needed for purposes of interpolation, are (*e*), (*i*), (*k*). (*a*), (*b*), (*c*), (*d*), (*e*) are the same as in Table II. (A) with the sign changed.

	R_0.	R_1.	R_2.	R_3.	R_4.	R_5.	R_6.	R_7.	R_8.	R_9.	R_{10}.
(*f*)	− ·3333	− ·3324	− ·3298	− ·3254	− ·3193	− ·3115	− ·3021	− ·2911	− ·2788	− ·2650	− ·2500
(*g*)	− ·4472	− ·4449	− ·4377	− ·4261	− ·4103	− ·3907	− ·3675	− ·3414	− ·3128	− ·2822	− ·2500
(*h*)	− ·6831	− ·67565	− ·6539	− ·61995	− ·5764	− ·5262	− ·4721	− ·4160	− ·3596	− ·3040	− ·2500
(*i*)	− ∞	− 5·1364	− 2·5270	− 1·6389	− 1·1814	− ·8962	− ·6974	− ·5481	− ·4299	− ·3325	− ·2500
(*j*)	− ·6667	− ·6611	− ·6447	− ·6181	− ·5822	− ·5384	− ·4881	− ·4327	− ·3736	− ·3123	− ·2500
(*k*)	− 1·0911	− 1·0758	− 1·0317	− ·9628	− ·8753	− ·7751	− ·6683	− ·5594	− ·4519	− ·3483	− ·2500

Table III. (B). $r_{23} = -\cdot 5$.

The possible values are (*a*), (*b*), (*c*), (*d*), (*f*), (*g*), (*j*) [limiting]. (*h*) and (*e*) are tabulated for purposes of interpolation. (*a*), (*b*), (*c*), (*d*), (*e*) are the same as in Table III. (A) with the sign changed.

	R_0.	R_1.	R_2.	R_3.	R_4.	R_5.	R_6.	R_7.	R_8.	R_9.	R_{10}.
(*f*)	− ·6000	− ·5989	− ·5958	− ·5905	− ·5831	− ·5738	− ·5625	− ·5493	− ·5345	− ·5180	− ·5
(*g*)	− ·7453	− ·7424	− ·7335	− ·7190	− ·6993	− ·6747	− ·6459	− ·6134	− ·5778	− ·5399	− ·5
(*h*)	− 1·0735	− 1·0634	− 1·0341	− ·9884	− ·9300	− ·8630	− ·7910	− ·7169	− ·6427	− ·5701	− ·5
(*j*)	− 1·0000	− ·9934	− ·9737	− ·9417	− ·8987	− ·8461	− ·7858	− ·7192	− ·6484	− ·5748	− ·5

Table IV. (B). $r_{23} = -\cdot 75$.

The possible values are (*a*), (*b*), (*c*), (*f*). (*d*) and (*g*) are needed for interpolation. (*a*), (*b*), (*c*), (*d*) are the same as in Table IV. (A) with the sign changed.

	R_0.	R_1.	R_2.	R_3	R_4.	R_5.	R_6.	R_7.	R_8.	R_9.	R_{10}.
(*f*)	− ·8667	− ·8654	− ·8617	− ·8555	− ·8470	− ·8361	− ·8229	− ·8076	− ·7903	− ·7710	− ·75
(*g*)	− 1·0435	− 1·0399	− 1·0293	− 1·0119	− ·9882	− ·9588	− ·9243	− ·8854	− ·8430	− ·7976	− ·75

Table V. (B) is entirely impossible except (*a*). $r_{12} = r_{13} = 0$. $R = -1$ for all values of selective intensity λ.