PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Mathematical Contributions to the Theory of Evolution. XI. On the Influence of Natural Selection on the Variability and Correlation of Organs

Karl Pearson

Phil. Trans. R. Soc. Lond. A 1903 **200**, 1-66
doi: 10.1098/rsta.1903.0001

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:

$$z = z_0 \text{ expt.} - \tfrac{1}{2}\left(\mathrm{S}_1\left(c_{pp}x_{pp}^2\right) + 2\mathrm{S}_2\left(c_{pq}x_p x_q\right)\right) \quad . \quad . \quad . \quad . \quad . \quad \text{(i.)}.$$

Here z_0, c_{pp}, c_{pq} . . . 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, $\bar{x}_3, \bar{x}_4 \ldots \bar{x}_n$, will be given by the equations:

$$\begin{aligned}
&c_{13}x_1 + c_{23}x_2 + c_{33}\bar{x}_3 + c_{43}\bar{x}_4 + \ldots + c_{n3}\bar{x}_n = 0.\\
&c_{14}x_1 + c_{24}x_2 + c_{34}\bar{x}_3 + c_{44}\bar{x}_4 + \ldots + c_{n4}\bar{x}_n = 0.\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&c_{1n}x_1 + c_{2n}x_2 + c_{3n}\bar{x}_3 + c_{4n}\bar{x}_4 + \ldots + c_{nn}\bar{x}_n = 0. \quad \ldots \quad \text{(ii.)}.
\end{aligned}$$

The following expressions will not be zero, but will be written α and β:

$$\begin{aligned}
&c_{11}x_1 + c_{21}x_2 + c_{31}\bar{x}_3 + c_{41}\bar{x}_4 + \ldots + c_{n1}\bar{x}_n = \alpha.\\
&c_{12}x_1 + c_{22}x_2 + c_{32}\bar{x}_3 + c_{42}\bar{x}_4 + \ldots + c_{n2}\bar{x}_n = \beta. \quad \ldots \quad \text{(iii.)}.
\end{aligned}$$

Now, if Δ be the discriminant,

$$\begin{vmatrix}
c_{11}, & c_{12}, & c_{13}, & \ldots & c_{1n}\\
c_{21}, & c_{22}, & c_{23}, & \ldots & c_{2n}\\
c_{31}, & c_{32}, & c_{33}, & \ldots & c_{3n}\\
\ldots & \ldots & \ldots & \ldots & \ldots\\
c_{n1}, & c_{n2}, & c_{n3}, & \ldots & c_{nn}
\end{vmatrix} \quad \ldots\ldots\ldots \quad \text{(iv.)},$$

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:

$$\alpha = \frac{\Delta(x_1C_{22} - x_2C_{12})}{C_{11}C_{22} - C_{12}^2}, \qquad \beta = \frac{\Delta(x_2C_{11} - x_1C_{12})}{C_{11}C_{22} - C_{12}^2} \quad \ldots\ldots \quad \text{(v.)}.$$

Generally also:

$$\begin{aligned}
\bar{x}_q &= \frac{\alpha C_{1q} + \beta C_{2q}}{\Delta}\\
&= \frac{(C_{1q}C_{22} - C_{2q}C_{12})x_1 + (C_{11}C_{2q} - C_{1q}C_{12})x_2}{C_{11}C_{22} - C_{12}^2}\\
&= D_{1q}x_1 + D_{2q}x_2, \quad \text{say}, \quad \ldots\ldots\ldots\ldots \quad \text{(vi.)}.
\end{aligned}$$

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

$$\begin{aligned}
\mathrm{Q} &= \mathrm{S}_{q=1}^{q=n} x_q (c_{1q}x_1 + c_{2q}x_2 + c_{3q}x_3 \quad + \ldots + c_{nq}x_n) \\
&= \mathrm{S}_{q=3}^{q=n} (\bar{x}_q + x'_q) \begin{pmatrix} c_{1q}x_1 + c_{2q}x_2 + c_{3q}\bar{x}_3 + \ldots + c_{nq}\bar{x}_n \\ + c_{3q}x'_3 + \ldots + c_{nq}x'_n \end{pmatrix} \\
&\quad + x_1 \begin{pmatrix} c_{11}x_1 + c_{21}x_2 + c_{31}\bar{x}_3 + \ldots + c_{n1}\bar{x}_n \\ + c_{31}x'_3 + \ldots + c_{n1}x'_n \end{pmatrix} \\
&\quad + x_2 \begin{pmatrix} c_{12}x_1 + c_{22}x_2 + c_{32}\bar{x}_3 + \ldots + c_{n2}\bar{x}_n \\ + c_{32}x'_3 + \ldots + c_{n2}x'_n \end{pmatrix}.
\end{aligned}$$

Making use of the n linear equations (ii.) and (iii.):

$$\begin{aligned}
\mathrm{Q} &= \mathrm{S}_{q=3}^{q=n} (\bar{x}_q + x'_q)(c_{3q}x'_3 + c_{4q}x'_4 + \ldots + c_{nq}x'_n) \\
&\quad + x_1 (\alpha + c_{31}x'_3 + c_{41}x'_4 + \ldots + c_{n1}x'_n) \\
&\quad + x_2 (\beta + c_{32}x'_3 + c_{42}x'_4 + \ldots + c_{n2}x'_n) \\
&= \mathrm{S}_{q=3}^{q=n} x'_q (c_{3q}x'_3 + c_{4q}x'_4 \quad + \ldots + c_{nq}x'_n) \\
&\quad + \alpha x_1 + \beta x_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(vii.)}.
\end{aligned}$$

For arranging vertical columns in rows, the remaining terms are

$$\begin{aligned}
&x'_3 (c_{31}x_1 + c_{32}x_2 + \mathrm{S}_{q=3}^{q=n} (c_{3q}\bar{x}_q)) \\
&+ x'_4 (c_{41}x_1 + c_{42}x_2 + \mathrm{S}_{q=3}^{q=n} (c_{4q}\bar{x}_q)) \\
&\quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\
&+ x'_n (c_{n1}x_1 + c_{n2}x_2 + \mathrm{S}_{q=3}^{q=n} (c_{nq}\bar{x}_q)),
\end{aligned}$$

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

$$z = \zeta_0 \text{ expt.} - \tfrac{1}{2}\Delta \frac{x_1^2 C_{22} - 2C_{12}x_1x_2 + C_{11}x_2^2}{C_{11}C_{22} - C_{12}^2} \quad \ldots \ldots \quad \text{(viii.)}.$$

Comparing this with the known form of the correlation surface for two variables,*

$$z = \frac{N}{2\pi\sigma_1\sigma_2\sqrt{1-r_{12}^2}} \text{ expt.} - \tfrac{1}{2}\frac{1}{1-r_{12}^2}\left(\frac{x_1^2}{\sigma_1^2} - \frac{2r_{12}x_1x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right) \quad \ldots \quad \text{(ix.)},$$

we have at once

$$\sigma_1^2(1-r_{12}^2) = \frac{C_{11}C_{22} - C_{12}^2}{\Delta C_{22}}, \qquad \sigma_2^2(1-r_{12}^2) = \frac{C_{11}C_{22} - C_{12}^2}{\Delta C_{11}},$$

$$r_{12}\sigma_1\sigma_2 = C_{12}/\Delta \quad \ldots \ldots \quad \text{(x.)}.$$

Whence

$$r_{12} = C_{12}/\sqrt{C_{11}C_{22}} \quad \ldots \ldots \quad \text{(xi.)},$$

$$\sigma_1^2 = C_{11}/\Delta, \qquad \sigma_2^2 = C_{22}/\Delta \quad \ldots \ldots \quad \text{(xii.)}.$$

Or, generally:

$$r_{pq} = C_{pq}/\sqrt{C_{pp}C_{qq}} \quad \ldots \ldots \quad \text{(xiii.)},$$

$$\sigma_p = \sqrt{C_{pp}/\Delta} \quad \ldots \ldots \quad \text{(xiv.)}.$$

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:

$$\begin{array}{l}
c_{11}C_{11} + c_{12}C_{12} + c_{13}C_{13} + \quad \ldots \quad + c_{1n}C_{1n} = \Delta. \\
c_{11}C_{21} + c_{12}C_{22} + c_{13}C_{23} + \quad \ldots \quad + c_{1n}C_{2n} = 0. \\
c_{11}C_{31} + c_{12}C_{32} + c_{13}C_{33} + \quad \ldots \quad + c_{1n}C_{3n} = 0. \\
\ldots\ldots\ldots\ldots\ldots\ldots \\
\ldots\ldots\ldots\ldots\ldots\ldots \\
c_{11}C_{n1} + c_{12}C_{n2} + c_{13}C_{n3} + \quad \ldots \quad + c_{1n}C_{nn} = 0.
\end{array}$$

Or,

$$\begin{array}{l}
c_{11}\sigma_1^2 + c_{12}r_{12}\sigma_1\sigma_2 + c_{13}r_{13}\sigma_1\sigma_3 + \quad \ldots \quad + c_{1n}r_{1n}\sigma_1\sigma_n = 1, \\
c_{11}r_{n1}\sigma_1\sigma_2 + c_{12}\sigma_2^2 + c_{13}r_{23}\sigma_2\sigma_3 + \quad \ldots \quad + c_{1n}r_{2n}\sigma_2\sigma_n = 0, \\
c_{11}r_{31}\sigma_1\sigma_3 + c_{12}r_{32}\sigma_2\sigma_3 + c_{13}\sigma_3^2 + \quad \ldots \quad + c_{1n}r_{3n}\sigma_3\sigma_n = 0. \\
\ldots\ldots\ldots\ldots\ldots\ldots \\
\ldots\ldots\ldots\ldots\ldots\ldots \\
c_{11}r_{n1}\sigma_1\sigma_n + c_{12}r_{n2}\sigma_2\sigma_n + c_{13}r_{n3}\sigma_3\sigma_n + \quad \ldots \quad + c_{1n}\sigma_n^2 = 0.
\end{array}$$

Hence solving

$$c_{11} = S_{11}/S, \qquad c_{1q} = S_{1q}/S \quad \ldots \ldots \quad \text{(xv.)},$$

* PEARSON, 'Phil. Trans.,' A, vol. 187, p. 264.

where $$S = \begin{vmatrix} \sigma_1^2, & r_{12}\sigma_1\sigma_2, & r_{13}\sigma_1\sigma_3, & \ldots & r_{1n}\sigma_1\sigma_n \\ r_{21}\sigma_1\sigma_2, & \sigma_2^2, & r_{23}\sigma_2\sigma_3, & \ldots & r_{2n}\sigma_2\sigma_n \\ r_{31}\sigma_1\sigma_3, & r_{32}\sigma_2\sigma_3, & \sigma_3^2, & \ldots & r_{3n}\sigma_3\sigma_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{n1}\sigma_1\sigma_n, & r_{n2}\sigma_2\sigma_n, & r_{n3}\sigma_3\sigma_n, & \ldots & \sigma_n^2 \end{vmatrix}$$

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,

$$\begin{vmatrix} 1, & r_{12}, & r_{13}, & \ldots & r_{1n} \\ r_{21}, & 1, & r_{23}, & \ldots & r_{2n} \\ r_{31}, & r_{32}, & 1, & \ldots & r_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{n1}, & r_{n2}, & r_{n3}, & \ldots & 1 \end{vmatrix} \quad \ldots\ldots\ldots \text{(xvi.)},$$

and R_{pq} is the minor corresponding to the constituent r_{pq}. Thus we have

$$c_{11} = \frac{1}{\sigma_1^2} R_{11}/R, \qquad c_{1q} = \frac{1}{\sigma_1\sigma_q} R_{1q}/R \quad \ldots\ldots \text{(xvii.)},$$

or, generally,

$$c_{pp} = \frac{1}{\sigma_p^2}\frac{R_{pp}}{R}, \qquad c_{pq} = \frac{1}{\sigma_p\sigma_q} R_{pq}/R \quad \ldots\ldots \text{(xviii.)}.$$

Thus z may be written*

$$z = z_0 \text{ expt.} - \tfrac{1}{2}\left\{S_1\left(\frac{R_{pp}}{R}\frac{x_p^2}{\sigma_p^2}\right) + 2S_2\left(\frac{R_{pq}}{R}\frac{x_p x_q}{\sigma_p\sigma_q}\right)\right\} \quad \ldots \text{(xix.)}.$$

It remains to determine z_0 from the fact that the volume of the surface $= N$.

Or, $$N = z_0 \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \text{expt.}\,(-\tfrac{1}{2}Q)\, dx_1\, dx_2 \ldots dx_n$$

$$= z_0\sigma_1\sigma_2 \ldots \sigma_n \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \text{expt.} - \tfrac{1}{2}\left\{S_1\left(\frac{R_{pp}}{R}x'^2_p\right) + 2S_2\left(\frac{R_{pq}}{R}x'_p x'_q\right)\right\}$$
$$dx'_1\, dx'_2 \ldots dx'_n \quad \ldots\ldots \text{(xx.)},$$

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.

$$S_1\left(\frac{R_{pp}}{R}x'^2_p\right)+2S_2\left(\frac{R_{pq}}{R}x'_px'_q\right)$$
$$=\frac{R_{11}}{R}\left(x'_1+\frac{R_{12}}{R_{11}}x'_2+\frac{R_{13}}{R_{11}}x'_3+\ldots+\frac{R_{1n}}{R_{11}}x'_n\right)^2$$
$$+\Sigma_1\left(\frac{R_{pp}R_{11}-R_{1p}^2}{RR_{11}}x'^2_p\right)+2\Sigma_2\left(\frac{R_{pq}R_{11}-R_{1p}R_{1q}}{RR_{11}}x'_px'_q\right),$$

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,

$$=\frac{R_{11}}{R}(x'_1+H'_1)^2+\Sigma_1\left(\frac{R'_{pp}}{R'}x'^2_p\right)+2\Sigma_2\left(\frac{R'_{pq}}{R'}x'_px'_q\right)\text{ where }H'_1=\Sigma_1\left(\frac{R_{1p}}{R_{11}}x'_p\right).$$

For if $R'=R_{11}$, 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 $(.R_{11})_{pq}$ of R is given by

$$R'_{pq}=(R_{pq}R_{11}-R_{1p}R_{1q})/R \quad\ldots\ldots\quad\text{(xxi.)},$$

and for $p=q$,

$$R'_{pp}=(R_{pp}R_{11}-R_{1p}^2)/R \quad\ldots\ldots\quad\text{(xxii.)}.$$

Hence, integrating x'_1 between the limits $+\infty$ and $-\infty$, we have

$$N=z_0\sigma_1\sigma_2\ldots\sigma_n.\sqrt{2\pi}\sqrt{\frac{R}{R'}}$$
$$\times\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty}\text{expt.}-\tfrac{1}{2}\left\{\Sigma_1\left(\frac{R'_{pp}}{R'}x'^2_p\right)+2\Sigma_2\left(\frac{R'_{pq}}{R'}x'_px'_q\right)\right\}dx'_2\,dx'_3\ldots dx'_n.$$

This is of precisely the same form as before, except that we have the factor $\sqrt{2\pi}\sqrt{R/R'}$, 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 $\sqrt{2\pi}\sqrt{R'/R''}$, where R″ is the minor of R′ not involving correlations of the second variable. Thus, by repeating the process, we have ultimately

$$N=z_0\sigma_1\sigma_2\ldots\sigma_n(\sqrt{2\pi})^n\sqrt{R/R'}\sqrt{R'/R''}\sqrt{R''/R'''}\ldots\sqrt{\frac{R^{n-1}}{1}},$$

or,

$$z_0=\frac{N}{(2\pi)^{\frac{1}{2}n}\sigma_1\sigma_2\ldots\sigma_n\sqrt{R}} \quad\ldots\ldots\quad\text{(xxiii.)},$$

which gives the constant of the surface.

The preceding investigation enables us also to deal with two further points.

(α) 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 $x'_2x'_3 \ldots x'_n$ 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 $\sigma_1, \sigma_2, \ldots \sigma_n$, we have a distribution about the point given by

$$\left.\begin{aligned} \bar{x}_1 &= -\sigma_1\left(\frac{R_{12}}{R_{11}}\frac{x_2}{\sigma_2} + \frac{R_{13}}{R_{11}}\frac{x_3}{\sigma_3} + \ldots + \frac{R_{1n}}{R_{11}}\frac{x_n}{\sigma_n}\right), \\ \text{with standard deviation} \qquad \bar{\sigma}'_1 &= \sigma_1\sqrt{R/R_{11}} \end{aligned}\right\} \quad \ldots \quad \text{(xxiv.)}.$$

(*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 $m_3 + h_3$, $m_4 + h_4 \ldots m_n + h_n$ to the remaining $(n-2)$ variables?

In this case we have from (i.)

$$\begin{aligned} z = \text{expt.} - \tfrac{1}{2}\{&c_{11}x_1^2 + 2c_{12}x_1x_2 + c_{22}x_2^2 \\ &+ 2(c_{13}h_3 + c_{14}h_{14} + \ldots + c_{1n}h_n)x_1 \\ &+ 2(c_{23}h_3 + c_{24}h_4 + \ldots + c_{2n}h_n)x_2 \\ &+ \text{terms not involving } x_1 \text{ and } x_2\} \quad \ldots \quad \text{(xxv.)}. \end{aligned}$$

Writing K_1 for the coefficient of x_1, and K_2 for that of x_2, we have for the centre

$$\bar{x}_1 = -\frac{(K_1c_{22} - K_2c_{12})}{c_{11}c_{22} - c_{12}^2}, \qquad \bar{x}_2 = -\frac{(-K_1c_{12} + c_{11}K_2)}{c_{11}c_{22} - c_{12}^2},$$

$$\bar{x}_1 = -\Sigma_1\left(\frac{c_{1p}c_{22} - c_{2p}c_{12}}{c_{11}c_{22} - c_{12}^2}h_p\right), \qquad \bar{x}_2 = -\Sigma_1\left(\frac{c_{2p}c_{11} - c_{1p}c_{12}}{c_{11}c_{22} - c_{12}^2}h_p\right),$$

$$\bar{x}_1 = -\sigma_1\Sigma_1\left(\frac{R_{1p}R_{22} - R_{2p}R_{12}}{R_{11}R_{22} - R_{12}^2}\frac{h_p}{\sigma_p}\right), \qquad \bar{x}_2 = -\sigma_2\Sigma_1\left(\frac{R_{2p}R_{11} - R_{1p}R_{12}}{R_{11}R_{22} - R_{12}^2}\frac{h_p}{\sigma_p}\right),$$

by transferring to the minors of R and the σ's. Or,

$$\bar{x}_1 = -\sigma_1\Sigma_1\left(\frac{\rho''_{1p}}{R''}\frac{h_p}{\sigma_p}\right), \qquad \bar{x}_2 = -\sigma_2\Sigma_2\left(\frac{\rho'_{2p}}{R''}\frac{h_p}{\sigma_p}\right) \quad \ldots \quad \text{(xxvi.)}.$$

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

$$\bar{r}_{12} = -\frac{c_{12}}{\sqrt{c_{11}c_{22}}} = -\mathrm{R}_{12}/\sqrt{\mathrm{R}_{11}\mathrm{R}_{22}} \quad . \quad . \quad . \quad . \quad . \quad \text{(xxvii.)}.$$

This is the *partial* correlation of the 1st and 2nd organs for the remaining $n - 2$ organs with constant values.

Again,

$$c_{11} = \frac{1}{\bar{\sigma}''_1{}^2(1 - \bar{r}_{12}{}^2)}, \quad c_{22} = \frac{1}{\bar{\sigma}''_2{}^2(1 - \bar{r}_{12}{}^2)}.$$

Whence we easily find from (xviii.) and (xxvii.)

$$\bar{\sigma}''_1{}^2 = \sigma_1{}^2 \frac{\mathrm{R}\,\mathrm{R}_{22}}{\mathrm{R}_{11}\mathrm{R}_{22} - \mathrm{R}_{12}{}^2}, \quad \bar{\sigma}''_2{}^2 = \sigma_2{}^2 \frac{\mathrm{R}\,\mathrm{R}_{11}}{\mathrm{R}_{11}\mathrm{R}_{22} - \mathrm{R}_{12}{}^2},$$

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.),

$$z = \text{constant} \times \text{expt.} - \tfrac{1}{2}\left\{S_1\left(\frac{R_{pp}}{R}\frac{x_p^2}{\sigma_p^2}\right) + 2S'_2\left(\frac{R_{pq}}{R}\frac{x_p x_q}{\sigma_p\sigma_q}\right)\right\} \quad . \quad . \quad \text{(xxix.)}.$$

For brevity, we can also write this in the form

$$z = \text{constant} \times \text{expt.} - \tfrac{1}{2}\{S_1(c_{pp}x_p^2) + 2S_2(c_{pq}x_p x_q)\} \quad . \quad . \quad \text{(xxix.)}^{\text{bis}}.$$

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 $R(u)$, if u be $> q$, for the determinant :

$$R(u) = \begin{vmatrix} 1, & r_{12}, & r_{13}, & \ldots & r_{1q}, & r_{1u} \\ r_{21}, & 1, & r_{23}, & \ldots & r_{2q}, & r_{2u} \\ r_{31}, & r_{32}, & 1, & \ldots & r_{3q}, & r_{3u} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{q1}, & r_{q2}, & r_{q3}, & \ldots & 1, & r_{qu} \\ r_{u1}, & r_{u2}, & r_{u3}, & \ldots & r_{uq}, & 1 \end{vmatrix} \quad . \quad . \quad . \quad . \quad \text{(xxx.)}.$$

Then if $R(u)_{p'p''}$ be the minor corresponding to the constituent $r_{p'p''}$, and if $b_{p'p''} = \dfrac{R(u)_{p'p''}}{R(u)}\dfrac{1}{\sigma_{p'}\sigma_p}$, the distribution of the $q+1$ organs will be given by the frequency distribution

$$z' = \text{constant} \times \text{expt.} - \tfrac{1}{2}\{S_1(b_{p'p'}x_{p'}^2) + 2S_2(b_{p'p''}x_{p'}x_{p''})\} \quad . \quad . \quad \text{(xxxi.)}.$$

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

$$\begin{aligned} x'_u &= -\left(\frac{b_{1u}}{b_{uu}}h_1 + \frac{b_{2u}}{b_{uu}}h_2 + \ldots + \frac{b_{qu}}{b_{uu}}h_q\right), \\ &= -\left(\frac{R(u)_{1u}}{R(u)_{uu}}\frac{\sigma_u}{\sigma_1}h_1 + \frac{R(u)_{2u}}{R(u)_{uu}}\frac{\sigma_u}{\sigma_2}h_2 + \ldots \frac{R(u)_{qu}}{R(u)_{uu}}\frac{\sigma_u}{\sigma_q}h_q\right) \quad . \quad . \quad \text{(xxxii.)}. \end{aligned}$$

Now these coefficients can be found at once if q be known.

For example :

$$q = 1, \quad -\frac{R(u)_{qu}}{R(u)_{uu}} = r_{1u},$$

$$q = 2, \quad -\frac{R(u)_{1u}}{R(u)_{uu}} = \frac{r_{1u} - r_{2u}r_{12}}{1 - r_{12}^2}, \qquad -\frac{R(u)_{2u}}{R(u)_{uu}} = \frac{r_{2u} - r_{1u}r_{12}}{1 - r_{12}^2},$$

$$q = 3, \quad -\frac{R(u)_{1u}}{R(u)_{uu}} = \frac{r_{1u}(1 - r_{23}^2) - r_{12}r_{2u} - r_{13}r_{3u} + r_{23}(r_{12}r_{3u} + r_{13}r_{2u})}{1 - r_{12}^2 - r_{23}^2 - r_{13}^2 + 2r_{12}r_{23}r_{13}}$$

and $-\dfrac{R(u)_{2u}}{R(u)_{uu}}$ and $-\dfrac{R(u)_{3u}}{R(u)_{uu}}$ can be written down by symmetry . (xxxiii.).

Now suppose in the expression (xxix.)^{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

$$\begin{aligned}
&c_{1,q+1}h_1 + c_{2,q+1}h_2 + \ldots + c_{q,q+1}h_q + c_{q+1,q+1}x'_{q+1} + c_{q+2,q+1}x'_{q+2} + \ldots + c_{n,q+1}x'_n = 0.\\
&c_{1,q+2}h_1 + c_{2,q+2}h_2 + \ldots + c_{q,q+2}h_q + c_{q+1,q+2}x'_{q+1} + c_{q+2,q+2}x'_{q+2} + \ldots + c_{n,q+2}x'_n = 0.\\
&\quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&c_{1n}h_1 + c_{2n}h_2 + \ldots + c_{q,n}h_q + c_{q+1,n}x'_{q+1} + c_{q+2,n}x'_{q+2} + \ldots + c_{nn}x'_n = 0 \quad . \quad \text{(xxxiv.)},
\end{aligned}$$

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.)^{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:

$$\begin{aligned}
Z = \text{constant} \times \text{expt.} - \tfrac{1}{2}\{&\bar{c}_{11}x_1^2 + \bar{c}_{22}x_2^2 + \ldots + \bar{c}_{qq}x_q^2\\
&+ c_{q+1,q+2}x_{q+1}^2 + \ldots + c_{uu}x_u^2\}\\
&+ 2\bar{c}_{v'v''}x_{v'}x_{v''} + \ldots\ (v' \text{ and } v'' < q+1)\\
&+ 2c_{vu}x_v x_u + \ldots\ (v < q+1 \text{ and } u > q)\\
&+ 2c_{u'u''}x_{u'}x_{u''} + \ldots\ (u' \text{ and } u'' > q)\\
&+ \text{linear terms in } x_{v'}, x_{v''}\,(v' \text{ and } v'' < q+1) \quad . \quad \text{(xxxv.)},
\end{aligned}$$

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

$$x'_{q+1} = -\left\{\frac{\mathrm{R}(q+1)_{1,q+1}}{\mathrm{R}(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_1}h_1 + \frac{\mathrm{R}(q+1)_{2,q+1}}{\mathrm{R}(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_2}h_2 + \ldots + \frac{\mathrm{R}(q+1)_{q,q+1}}{\mathrm{R}(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_q}h_q\right\},$$

$$x'_{q+2} = -\left\{\frac{\mathrm{R}(q+2)_{1,q+2}}{\mathrm{R}(q+2)_{q+2,q+1}}\frac{\sigma_{q+2}}{\sigma_1}h_1 + \frac{\mathrm{R}(q+2)_{2,q+2}}{\mathrm{R}(q+2)_{q+2,q+2}}\frac{\sigma_{q+2}}{\sigma_2}h_2 + \ldots + \frac{\mathrm{R}(q+2)_{q,q+2}}{\mathrm{R}(q+2)_{q+2,q+2}}\frac{\sigma_{q+2}}{\sigma_q}h_q\right\},$$

. . . . . .

$$x'_n = -\left\{\frac{\mathrm{R}(n)_{1,n}}{\mathrm{R}(n)_{nn}}\frac{\sigma_n}{\sigma_1}h_1 + \frac{\mathrm{R}(n)_{2,n}}{\mathrm{R}(n)_{nn}}\frac{\sigma_n}{\sigma_2}h_2 + \ldots + \frac{\mathrm{R}(n)_{q,n}}{\mathrm{R}(n)_{nn}}\frac{\sigma_n}{\sigma_q}h_q\right\} \quad \ldots\ldots \quad \text{(xxxvi.)}.$$

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

$$\Delta = \begin{vmatrix} \bar{c}_{1,1}, & \bar{c}_{1,2}, & \bar{c}_{1,3} & \ldots & \bar{c}_{1,q}, & c_{1,q+1} & \ldots & c_{1,n} \\ \bar{c}_{2,1}, & \bar{c}_{2,2}, & \bar{c}_{2,3} & \ldots & \bar{c}_{2,q}, & c_{2,q+1} & \ldots & c_{2,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \bar{c}_{q1}, & \bar{c}_{q2}, & \bar{c}_{q,3} & \ldots & \bar{c}_{qq}, & c_{qq+1} & \ldots & c_{q,n} \\ c_{q+1,1}, & c_{q+1,2}, & c_{q+1,3} & \ldots & c_{q+1,q}, & c_{q+1,q+1} & \ldots & c_{q+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ c_{n,1}, & c_{n,2}, & c_{n,3} & \ldots & c_{n,q}, & c_{n,q+1} & \ldots & c_{nn} \end{vmatrix} \quad \text{(xxxvii.)}$$

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 r_{uv} be the correlation

coefficient of the u^{th} and v^{th} organs after selection, then if u and v be selected organs, $\mathfrak{r}_{uv} = \rho_{uv}$; 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.)

$$\Sigma_u^{\,2} = \frac{\mathrm{D}_{uu}}{\Delta} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxviii.)},$$

and

$$\Sigma_u \Sigma_v \mathfrak{r}_{uv} = \frac{\mathrm{D}_{uv}}{\Delta} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxix.)}.$$

Now let us write for brevity $\alpha_{uu} = \mathrm{D}_{uu}/\Delta$ and $\alpha_{uv} = \mathrm{D}_{uv}/\Delta$. 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 $s_u s_r \rho_{uv}$.

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

$$\alpha_{1,q+1} = -\left\{\frac{\mathrm{R}(q+1)_{1,q+1}}{\mathrm{R}(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_1}\alpha_{11} + \frac{\mathrm{R}(q+1)_{2,q+1}}{\mathrm{R}(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_2}\alpha_{12} + \ldots + \frac{\mathrm{R}(q+1)_{q,q+1}}{\mathrm{R}(q+1)_{q+1,q+1}}\frac{\sigma_{q+1}}{\sigma_q}\alpha_{1q}\right\}$$

$$\alpha_{1,q+2} = -\left\{\frac{\mathrm{R}(q+2)_{1,q+2}}{\mathrm{R}(q+2)_{q+2,q+2}}\frac{\sigma_{q+2}}{\sigma_1}\alpha_{11} + \frac{\mathrm{R}(q+2)_{2,q+2}}{\mathrm{R}(q+2)_{q+2,q+2}}\frac{\sigma_{q+2}}{\sigma_2}\alpha_{12} + \ldots + \frac{\mathrm{R}(q+2)_{q,q+2}}{\mathrm{R}(q+2)_{q+2,q+2}}\frac{\sigma_{q+2}}{\sigma_q}\alpha_{1q}\right\}$$

. . . . . .

$$\alpha_{1,n} = -\left\{\frac{\mathrm{R}(n)_{1,n}}{\mathrm{R}(n)_{n,n}}\frac{\sigma_n}{\sigma_1}\alpha_{11} + \frac{\mathrm{R}(n)_{2,n}}{\mathrm{R}(n)_{n,n}}\frac{\sigma_n}{\sigma_2}\alpha_{12} + \ldots + \frac{\mathrm{R}(n)_{q,n}}{\mathrm{R}(n)_{n,n}}\frac{\sigma_n}{\sigma_q}\alpha_{1q}\right\}$$

Provided v be $< q+1$ we might equally well have used $\alpha_{r1}, \alpha_{r2}, \ldots \alpha_{rq}$. Hence we conclude that if v be $< q+1$ and u be $> q$,

$$\alpha_{vu} = -\left\{\frac{\mathrm{R}(u)_{1u}}{\mathrm{R}(u)_{uu}}\frac{\sigma_u}{\sigma_1}\alpha_{v1} + \frac{\mathrm{R}(u)_{2u}}{\mathrm{R}(u)_{uu}}\frac{\sigma_u}{\sigma_2}\alpha_{r2} + \ldots + \frac{\mathrm{R}(u)_{qu}}{\mathrm{R}(u)_{uu}}\frac{\sigma_u}{\sigma_q}\alpha_{rq}\right\} \quad . \quad . \quad \text{(xli.)}.$$

Or, substituting the known values of $\alpha_{v1}, \alpha_{v2}, \ldots \alpha_{vq}$, we have for $v < q+1$ and $u > q$

$$\Sigma_v\Sigma_u\mathfrak{r}_{vu} = -s_v\sigma_u\left\{\frac{\mathrm{R}(u)_{1v}}{\mathrm{R}(u)_{uu}}\rho_{1v}\frac{s_1}{\sigma_1} + \frac{\mathrm{R}(u)_{2v}}{\mathrm{R}(u)_{uu}}\rho_{2v}\frac{s_2}{\sigma_2} + \ldots + \frac{\mathrm{R}(u)_{vu}}{\mathrm{R}(u)_{uu}}\frac{s_v}{\sigma_v} + \ldots + \frac{\mathrm{R}(u)_{qu}}{\mathrm{R}(u)_{uu}}\rho_{qv}\frac{s_q}{\sigma_q}\right\} \quad . \quad . \quad . \quad . \quad \text{(xli.)}^{\text{bis}}.$$

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}, α_{u2}, . . . α_{uq} . . . α_{uu} in the equations formed from the last $p-q$ lines of the determinant (xxxvii.).

We obtain the following system:

$$c_{1,q+1}\alpha_{u1} + c_{2,q+1}\alpha_{u2} + c_{3,q+1}\alpha_{u3} + \ldots + c_{q,q+1}\alpha_{uq} + c_{q+1,q+1}\alpha_{u,q+1} + \ldots + c_{u,q+1}\alpha_{uu} = 0$$

$$c_{1,q+2}\alpha_{u1} + c_{2,q+2}\alpha_{u2} + c_{3,q+2}\alpha_{u3} + \ldots + c_{q,q+2}\alpha_{uq} + c_{q+1,q+2}\alpha_{u,q+1} + \ldots + c_{u,q+2}\alpha_{uu} = 0$$

. . . . . .

. . . . . .

$$c_{1,u}\alpha_{u1} + c_{2,u}\alpha_{u2} + c_{3,u}\alpha_{u3} + \ldots + c_{q,u}\alpha_{uq} + c_{q+1,u}\alpha_{u,q+1} + \ldots + c_{u,u}\alpha_{uu} = 1$$

. . . . . .

. . . . . .

$$c_{1,u}\alpha_{u1} + c_{2u}\alpha_{u2} + c_{3u}\alpha_{u3} + \ldots + c_{q,u}\alpha_{uq} + c_{q+1,u}\alpha_{u,q+1} + \ldots + c_{u,u}\alpha_{uu} = 0.$$

These are identical with equations (xl.), except that the equation with the coefficients c_{1u}, c_{2u}, . . . c_{uu} has unity instead of zero on the right-hand side. Hence we see that $\alpha_{uv}\,(v > q)$ will be the same function of α_{u1}, α_{u2}, . . . α_{uq} that x'_v is of h_1, h_2, . . . 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

$$\begin{array}{llll} c_{q,q+1}, & c_{q+1,q+1}, & \ldots & c_{u,q+1} \\ c_{q,q+1}, & c_{q+1,q+2}, & \ldots & c_{u,q+2} \\ \ldots & \ldots & \ldots & \ldots \\ c_{q\,u}, & c_{q+1,u}, & \ldots & c_{u,u}. \end{array}$$

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, . . . s_q, the selected means h_1, h_2, . . . h_q, and the selected correlation coefficients ρ_{12}, ρ_{23}, . . . $\rho_{q-1,q}$. It depends only on the characters *before* selection.

We thus have

$$\alpha_{uv} = \gamma_{uv} - \left\{\frac{\mathrm{R}(v)_{1v}}{\mathrm{R}(v)_{vv}}\frac{\sigma_v}{\sigma_1}\alpha_{u1} + \frac{\mathrm{R}(v)_{2v}}{\mathrm{R}(v)_{vv}}\frac{\sigma_v}{\sigma_2}\alpha_{u2} + \ldots + \frac{\mathrm{R}(v)_{qv}}{\mathrm{R}(v)_{vv}}\alpha_{uq}\right\}.$$

Now the system α_{u1}, α_{u2}, . . . α_{uq} can be found from (xli.), since $\alpha_{u'u''} = \alpha_{u''u}$ whatever u' and u'' be.

Hence we have

$$\alpha_{uv} = \gamma_{uv} + \sigma_v\sigma_u\left[S_1\left\{\frac{R(v)_{pv}\,R(u)_{pv}}{R(v)_{vv}\,R(u)_{uu}}\left(\frac{s_p}{\sigma_p}\right)^2\right\} + S_2\left\{\left(\frac{R(v)_{p'v}\,R(u)_{p''u}}{R(v)_{vv}\,R(u)_{uu}} + \frac{R(v)_{p''v}\,R(u)_{p'u}}{R(v)_{vv}\,R(u)_{uu}}\right)\rho_{p'p''}\frac{s_{p'}s_{p''}}{\sigma_{p'}\sigma_{p''}}\right\}\right].$$

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, . . . q.

When $u = v$ we have simply

$$\alpha_{vv} = \gamma_{vv} + \sigma_v^2\left[S_1\left\{\left(\frac{R(v)_{pv}}{R(v)_{vv}}\right)^2\left(\frac{s_p}{\sigma_p}\right)^2 + 2S_2\left(\frac{R(v)_{p'v}\,R(v)_{p''v}}{R(v)_{vv}^2}\rho_{p'p''}\frac{s_{p'}s_{p''}}{\sigma_{p'}\sigma_{p''}}\right)\right\}\right].$$

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:

$$\left.\begin{aligned}\Sigma_u\Sigma_v\mathfrak{r}_{vu} = \gamma_{uv} &= -\sigma_u\sigma_v R(uv)_{uv}/R(q),\\ \Sigma_v^2 = \gamma_v^2 &= \sigma_v^2 R(uv)_{uu}/R(q),\\ \Sigma_u^2 = \gamma_u^2 &= \sigma_u^2 R(uv)_{vv}/R(q),\end{aligned}\right\} \quad \ldots\ldots \text{(xlii.)}.$$

The notation of that page has been changed so that $R(uv)$ now stands for the determinant

$$\left|\begin{array}{ccccccccc} 1, & r_{12}, & r_{13}, & \ldots & r_{1q}, & r_{1u}, & r_{1v}\\ r_{21}, & 1, & r_{23}, & \ldots & r_{2q}, & r_{2u}, & r_{2v}\\ r_{31}, & r_{32}, & 1, & \ldots & r_{3q}, & r_{3u}, & r_{3v}\\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\\ r_{q1}, & r_{q2}, & r_{q3}, & \ldots & 1, & r_{qu}, & r_{qv}\\ r_{u1}, & r_{u2}, & r_{u3}, & \ldots & r_{uq}, & 1, & r_{uv}\\ r_{v1}, & r_{v2}, & r_{v3}, & \ldots & r_{vq}, & r_{vu}, & 1 \end{array}\right| \quad \ldots\ldots \text{(xliii.)}.$$

$R(uv)_{uv}$ is the minor corresponding to the constituent r_{uv}; $R(uv)_{uu}$, the minor corresponding to the constituent at the meet of the u^{th} column and u^{th} row; and $R(q)$ the determinant with the last two rows and two columns struck out.

For example:

$$q=1: \quad \gamma_{vv}=\sigma_v^2(1-r_{1v}^2) \qquad \gamma_{uu}=\sigma_u^2(1-r_{1u}^2),$$
$$\gamma_{uv}=\sigma_v\sigma_u(r_{vu}-r_{1v}r_{1u}),$$

$$q=2: \quad \gamma_{vv}=\sigma_v^2(1-r_{12}^2-r_{1v}^2-r_{2v}^2+2r_{12}r_{1v}r_{2v})/(1-r_{12}^2),$$
$$\gamma_{uu}=\sigma_u^2(1-r_{12}^2-r_{1u}^2-r_{2u}^2+2r_{12}r_{1u}r_{2u})/(1-r_{12}^2),$$
$$\gamma_{uv}=\sigma_u\sigma_v\{r_{vu}(1-r_{12}^2)-r_{1v}r_{1u}-r_{2v}r_{2u}+r_{12}(r_{1v}r_{2v}+r_{2u}r_{1v})\}/(1-r_{12}^2),$$

$q=3$:* $\gamma_{vv}=\sigma_v^2 \times$

$$\begin{vmatrix} 1, & r_{12}, & r_{13}, & r_{1v} \\ r_{21}, & 1, & r_{23}, & r_{2v} \\ r_{31}, & r_{32}, & 1, & r_{3v} \\ r_{v1}, & r_{v2}, & r_{v3}, & 1 \end{vmatrix} \div (1-r_{12}^2-r_{23}^2-r_{31}^2+2r_{12}r_{23}r_{31})$$

$\gamma_{uu}=\sigma_u^2 \times$

$$\begin{vmatrix} 1, & r_{12}, & r_{13}, & r_{1u} \\ r_{21}, & 1, & r_{23}, & r_{2u} \\ r_{31}, & r_{32}, & 1, & r_{3u} \\ r_{u1}, & r_{u2}, & r_{u3}, & 1 \end{vmatrix} \div (1-r_{12}^2-r_{23}^2-r_{31}^2+2r_{12}r_{23}r_{31})$$

$\gamma_{uv}=\sigma_v\sigma_u$

$$\begin{vmatrix} 1, & r_{12}, & r_{13}, & r_{1v} \\ r_{21}, & 1, & r_{23}, & r_{2v} \\ r_{31}, & r_{32}, & 1, & r_{3v} \\ r_{u1}, & r_{u2}, & r_{u3}, & r_{uv} \end{vmatrix} \div (1-r_{12}^2-r_{23}^2-r_{31}^2+2r_{12}r_{23}r_{31}) \quad \ldots\ldots \text{(xliv.)}.$$

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

$$\Sigma_u\Sigma_v\mathfrak{r}_{uv}=\sigma_u\sigma_v\left[-\frac{\mathrm{R}(uv)_{uv}}{\mathrm{R}(q)}+\mathrm{S}_1\left\{\frac{\mathrm{R}(v)_{pv}\,\mathrm{R}(u)_{pu}}{\mathrm{R}(v)_{vv}\,\mathrm{R}(u)_{uu}}\left(\frac{s_p}{\sigma_p}\right)^2\right\}\right.$$
$$\left.+\mathrm{S}_2\left\{\left(\frac{\mathrm{R}(v)_{p'v}\,\mathrm{R}(u)_{p''u}}{\mathrm{R}(v)_{vv}\,\mathrm{R}(u)_{uu}}+\frac{\mathrm{R}(v)_{p''v}\,\mathrm{R}(u)_{p'u}}{\mathrm{R}(v)_{vv}\,\mathrm{R}(u)_{uu}}\right)\rho_{p'p''}\frac{s_{p'}s_{p''}}{\sigma_{p'}\sigma_{p''}}\right\}\right] \quad \ldots \text{(xlvi.)}.$$

The correlation between a non-selected organ $u>q$ and a selected organ $v<q+1$ is given by (xli.)$^{\text{bis}}$ or,

$$s_v\Sigma_u\mathfrak{r}_{vu}=-s_v\sigma_u\left[\mathrm{S}_1\left(\frac{\mathrm{R}(u)_{pu}}{\mathrm{R}(u)_{uu}}\rho_{pv}\frac{s_p}{\sigma_p}\right)\right] \quad \ldots\ldots \text{(xlvii.)}.$$

* 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 $\mathfrak{r}_{12}$ the correlation of the two organs after selection, and $m_2 + \mathfrak{x}'_2$ the mean of the non-selected organ.

Then by (xxxii.):

$$\mathfrak{x}'_2 = r_{12}\frac{\sigma_2}{\sigma_1}h_1,$$

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

$$\Sigma_2\mathfrak{r}_{12} = \sigma_2\frac{s_1}{\sigma_1}r_{12}.$$

Hence we have:

$$\frac{\mathfrak{r}_{12}\Sigma_2}{s_1} = \frac{r_{12}\sigma_2}{\sigma_1},$$

$$\frac{\mathfrak{r}_{12}s_1}{\Sigma_2} = \frac{s_1^2}{\sigma_1\sigma_2}\frac{r_{12}}{1 - \left(1 - \frac{s_1^2}{\sigma_1^2}\right)r_{12}^2},$$

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

$$\begin{aligned} Y_2 &= \mathfrak{r}_{12}\frac{\Sigma_2}{s_1}K_1 + M_2 - \mathfrak{r}_{12}\frac{\Sigma_2}{s_1}M_1, \\ &= r_{12}\frac{\sigma_2}{\sigma_1}K_1 + m_2 + r_{12}\frac{\sigma_2}{\sigma_1}h_1 - r_{12}\frac{\sigma_2}{\sigma_1}(m_1 + h_1), \\ &= r_{12}\frac{\sigma_2}{\sigma_1}K_1 + m_2 - r_{12}\frac{\sigma_2}{\sigma_1}m_1. \end{aligned}$$

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

$$z = \frac{\mathrm{N}}{2\pi\sigma_1\sigma_2}\frac{1}{\sqrt{1-r_{12}^2}}\,e^{-\frac{1}{2}\left(\frac{x_1^2}{\sigma_1^2(1-r_{12}^2)}-\frac{2r_{12}x_1x_2}{\sigma_1\sigma_2(1-r_{12}^2)}+\frac{x_2^2}{\sigma_2^2(1-r_{12}^2)}\right)},$$

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 $\mathrm{N}-n$ 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

$$z_1 = \frac{\mathrm{N}}{\sqrt{2\pi}\sigma_1}\,e^{-\frac{x_1^2}{2\sigma_1^2}}.$$

Hence, the selection being random with regard to the array of B's corresponding to any A, we have for the surface after selection

$$\mathrm{Z} = z \times z'/z_1,$$

for each array must be altered in the ratio of the corresponding z' to z_1.

This gives for the surface in full

$$\mathrm{Z} = \frac{n}{2\pi\sigma_2 s_1}\,\mathrm{expt.}\,-\frac{1}{2}\left\{x_1^2\left(\frac{1}{s_1^2}-\frac{1}{\sigma_1^2}+\frac{1}{\sigma_1^2(1-r_{12}^2)}\right)-\frac{2r_{12}x_1x_2}{\sigma_1\sigma_2(1-r_{12}^2)}\right.$$
$$\left.+\,x_2^2\,\frac{1}{\sigma_2^2(1-r_{12}^2)}-\frac{2h_1x_1}{s_1^2}\right\}\quad .\quad .\quad .\quad \text{(xlviii.)}.$$

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

$$\Sigma_2^2 = \sigma_2^2\left(1 - \left(1 - \frac{s_1^2}{\sigma_1^2}\right) r_{12}^2\right). \quad \ldots\ldots\ldots \quad \text{(xlix.)},$$

which gives the standard deviation of the indirectly selected organ, and the "centre" of this organ is given by

$$h_2 = \frac{r_{12}\sigma_2}{\sigma_1} h_1 . \quad \ldots\ldots\ldots \quad \text{(l.)}.$$

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

$$\mathrm{r}_{12} = \frac{s_1}{\sigma_1} \frac{r_{12}}{\sqrt{1 - \left(1 - \frac{s_1^2}{\sigma_1^2}\right) r_{12}^2}} \quad \ldots\ldots\ldots \quad \text{(li.)}.$$

Let $r_{12} = \sin\theta_{12}$, $\mathrm{r}_{12} = \sin\phi_{12}$, then we have

$$\tan\phi_{12} = \frac{s_1}{\sigma_1}\tan\theta_{12} \quad \ldots\ldots\ldots \quad \text{(lii.)}.$$

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.

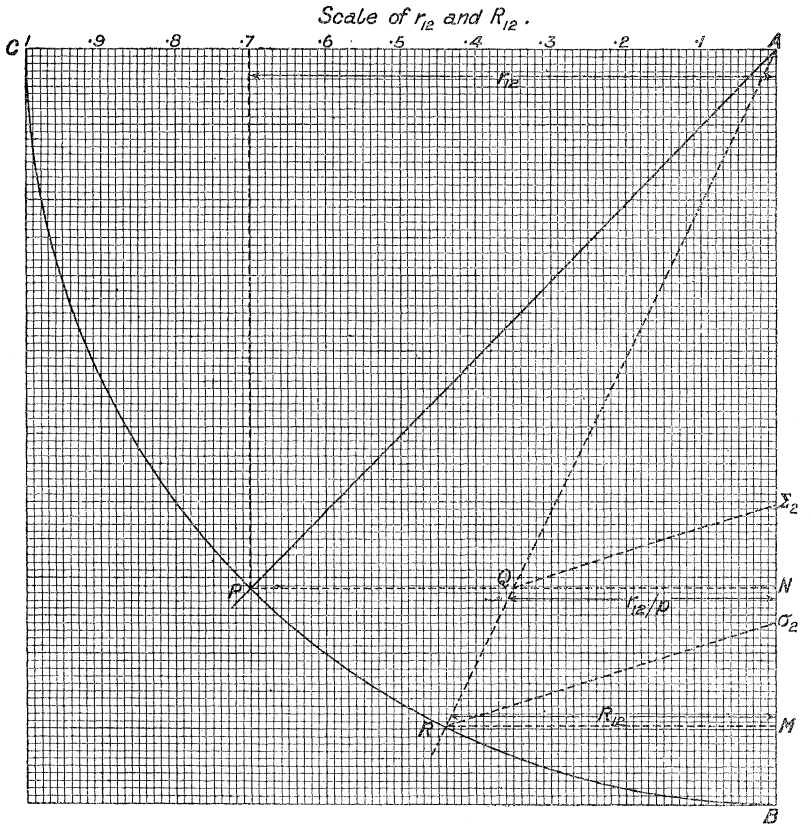
As an illustration, suppose that the correlation between femur and tibia were ·7, and let us investigate what would be the effect of reducing the variability of the tibia by direct selection 50 per cent. We find at once on substituting $s_1/\sigma_1 = \cdot 5$ and $r_{12} = \cdot 7$ in (li.) above that $\mathrm{r}_{12} = \cdot 44$, 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.) $\mathfrak{r}_{12}$ 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 $\mathfrak{r}_{12}$ is always of the same sign and less than $\mathfrak{r}_{12}$. 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 $QN = \frac{1}{p} PN$, and let AQ meet the quadrant in R, then RM the distance of R from $AB = 10\mathfrak{r}_{12}$, and consequently determines $\mathfrak{r}_{12}$. If the figure be drawn on decimal paper the determination of $\mathfrak{r}_{12}$ is peculiarly easy.

Graphical method of finding correlation between organs A and B after selection has acted on A.



In the above example $r_{12} = PN = \cdot 7$; $s_1/\sigma = \frac{1}{p} = \frac{1}{2}$, and therefore $QN = r_{12}/p = \cdot 35$.

$\mathfrak{r}_{12} = R_{12} = RM = \cdot 44$, 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.):

$$x'_2 = r_{12} \frac{\sigma_2}{\sigma_1} h_1, \qquad x'_3 = r_{13} \frac{\sigma_3}{\sigma_1} h_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(liii.)},$$

$$\left.\begin{aligned} \Sigma_2 &= \sigma_2 \left\{1 - \left(1 - \left(\frac{s_1}{\sigma_1}\right)^2\right) r_{12}^2\right\}^{\frac{1}{2}} = \sigma_2 \{1 - \sin^2 \chi_1 \cos^2 \theta_{12}\}^{\frac{1}{2}} \\ \Sigma_3 &= \sigma_3 \left\{1 - \left(1 - \left(\frac{s_1}{\sigma_1}\right)^2\right) r_{13}^2\right\}^{\frac{1}{2}} = \sigma_3 \{1 - \sin^2 \chi_1 \cos^2 \theta_{13}\}^{\frac{1}{2}} \end{aligned}\right\} \quad . \quad \text{(liv.)},$$

$$\left.\begin{aligned} \mathfrak{r}_{12} &= \frac{\mu_1 r_{12}}{\{1 - (1 - \mu_1^2) r_{12}^2\}^{\frac{1}{2}}} = \frac{\cos \chi_1 \cos \theta_{12}}{\sqrt{1 - \sin^2 \chi_1 \cos^2 \theta_{12}}} \\ \mathfrak{r}_{13} &= \frac{\mu_1 r_{13}}{\{1 - (1 - \mu_1^2) r_{13}^2\}^{\frac{1}{2}}} = \frac{\cos \chi_1 \cos \theta_{13}}{\sqrt{1 - \sin^2 \chi_1 \cos^2 \theta_{13}}} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad \text{(lv.)},$$

$$\left.\begin{aligned} \mathfrak{r}_{23} &= \frac{(r_{23} - r_{13} r_{12}) \left(1 - \frac{s_1^2}{\sigma_1^2}\right) + \frac{s_1^2}{\sigma_1^2} r_{23}}{\sqrt{\left\{1 - \left(1 - \left(\frac{s_1}{\sigma_1}\right)^2\right) r_{12}^2\right\} \left\{1 - \left(1 - \left(\frac{s_1}{\sigma_1}\right)^2\right) r_{13}^2\right\}}} \\ &= \frac{\cos \theta_{23} + \sin^2 \chi_1 \cos \theta_{13} \cos \theta_{12}}{\sqrt{(1 - \sin^2 \chi_1 \cos^2 \theta_{12}) (1 - \sin^2 \chi_1 \cos^2 \theta_{13})}} \end{aligned}\right\} \quad . \quad . \quad \text{(lvi.)},$$

where, as before, we write $r_{pq} = \cos \theta_{pq}$. Let us also write $\mathfrak{r}_{pq} = \cos \Theta_{pq}$, and

$$\sin \chi_1 \cos \theta_{12} = \cos \alpha_{12}, \qquad \sin \chi_1 \cos \theta_{13} = \cos \alpha_{13}.$$

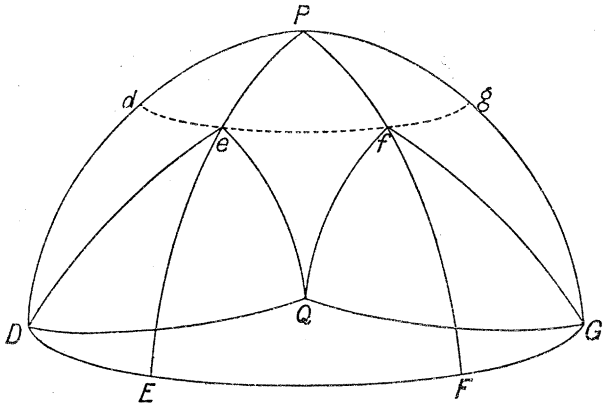
Then we can replace the above results by

$$\Sigma_2 = \sigma_2 \sin \alpha_{12}, \qquad \Sigma_3 = \sigma_3 \sin \alpha_{13},$$

$$\cos \Theta_{12} = \cot \chi_1 \cot \alpha_{12}, \qquad \cos \Theta_{13} = \cot \chi_1 \cot \alpha_{13},$$

$$\cos \Theta_{23} = \frac{\cos \theta_{23} - \cos \alpha_{12} \cos \alpha_{13}}{\sin \alpha_{12} \sin \alpha_{13}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lvii.)}.$$

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:

$$\Sigma_2/\sigma_2 = \sin DQ, \qquad \Sigma_3/\sigma_3 = \sin GQ.$$

$$r_{12} = \cos DeE, \qquad r_{13} = \cos EfG, \qquad r_{23} = \cos DQG.$$

For $$DE = \theta_{12},\ Ee = \frac{\pi}{2} - \chi_1,\ \angle\ DEe = \frac{\pi}{2};$$

hence: $$\cos De = \cos\theta_{12} \cos Ee = \sin\chi_1 \cos\theta_{12} = \cos\alpha_{12}, \text{ or } \alpha_{12} = De;$$

similarly $$\alpha_{13} = Gf.$$

Next, $\cos DeE = \cot De \tan eE = \cot\alpha_{12} \cot\chi_1$, or $\angle\ DeE = \Theta_{12}$; similarly $\angle EfG = \Theta_{13}$. Lastly, from the triangle DQG: $DQ = \alpha_{12}$, $QG = \alpha_{13}$, and $DG = \theta_{23}$, but

$$\cos DG = \cos DQ \cos QG + \sin DQ \sin QG \cos DQG,$$

or, $$\cos DQG = \frac{\cos\theta_{23} - \cos\alpha_{12}\cos\alpha_{13}}{\sin\alpha_{12}\sin\alpha_{13}} = \cos\Theta_{23};\ \text{or}\ \angle\ DQG = \Theta_{23}.$$

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), $(d)\ldots(m)$, (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 $\mu_1 = \frac{2}{10}$, $r_{12} = \cdot5$, $r_{13} = \cdot5$, $r_{23} = \cdot75$, $\delta\mu_1 = \cdot05$, $\delta r_{12} = \delta r_{13} = \cdot1$, $\delta r_{23} = \cdot05$. Further, we look up Table IV. (a), and the nearest case is (j) under R_2, which gives $R = \cdot67105$. We then see that $(\Delta_{r_{12}} R) + (\Delta_{r_{13}} R)$ = difference between (j) and (m) cases $= -\cdot21455$; $(\Delta_{r_{23}} R)$ = difference between result in IV. (a) and V. (a) $= \cdot32895$; and lastly $(\Delta_{\mu_1} R)$ = difference between R_3 and R_2 columns of (j) row of Table IV. (a) $= \cdot00535$. Thus we find:

$$R' = \cdot671050 + \cdot002675 - \cdot04291 + \cdot06578 = \cdot6966.$$

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 $s_1/\sigma_1 = \cdot5$?

In this case, $\mu_1 = \cdot5$, $r_{12} = \cdot5$, $r_{13} = \cdot75$, $r_{23} = \cdot5$, $\delta\mu_1 = 0$, $\delta r_{12} = \cdot2$, $\delta r_{13} = \cdot05$, $\delta r_{23} = 0$. We turn to Table III. (a) and take out (k) under R_5, which gives us $R = \cdot3192$. We have $\Delta_{r_{12}} R = -\cdot1841$ and $\Delta_{r_{13}} R = -\cdot04185$, whence we find $R' = \cdot1636$, 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 $r_{12} = \cdot75$, $r_{13} = \cdot75$, and therefore R, to be found from (m), $= \cdot1351$; then $\delta r_{12} = -\cdot05$, $\delta r_{13} = \cdot05$ and $\Delta_{r_{12}} R = -\cdot1841$ as before, $\Delta_{r_{13}} R = -\cdot2995$. Hence we deduce $R' = \cdot1120$. The mean of these two values of R' is ·1377, and the true value calculated from (lvi.) is $R' = \cdot1395$. 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.*, $s_1/\sigma_1 = \cdot5$?

Our best results from the tables will be to take R_5 (m) from Table I. (a), which gives $R = -\cdot7297$. We have then $r_{12} = r_{13} = \cdot75$, hence $\delta r_{12} = -\cdot05$, $\delta r_{13} = -\cdot05$. $\Delta_{r_{12}} R$ is to be found from (m) and (h) and $= -\cdot3193 = \Delta_{r_{13}} R$, and

$$R' = -\cdot7297 + (\Delta_{r_{12}} R)\,\delta r_{12} + (\Delta_{r_{13}} R)\,\delta r_{13},$$
$$= -\cdot7297 + \cdot1277 = -\cdot6020.$$

The actual value by formula is $-\cdot5810$.

* 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:

$$\Sigma_3^2 = \sigma_3^2 \left\{ \frac{1 - r_{12}^2 - r_{13}^2 - r_{23}^2 + 2r_{12}r_{13}r_{23}}{1 - r_{12}^2} + \left(\frac{r_{13} - r_{23}r_{12}}{1 - r_{12}^2}\right)^2 \left(\frac{s_1}{\sigma_1}\right)^2 + \left(\frac{r_{23} - r_{13}r_{12}}{1 - r_{12}^2}\right)^2 \left(\frac{s_2}{\sigma_2}\right)^2 + 2\rho_{12} \frac{r_{13} - r_{23}r_{12}}{1 - r_{12}^2} \frac{r_{23} - r_{13}r_{12}}{1 - r_{12}^2} \frac{s_1 s_2}{\sigma_1 \sigma_2} \right\} \quad \text{(lviii.)}.$$

$$\Sigma_4^2 = \sigma_4^2 \left\{ \frac{1 - r_{12}^2 - r_{14}^2 - r_{24}^2 + 2r_{12}r_{14}r_{24}}{1 - r_{12}^2} + \left(\frac{r_{14} - r_{24}r_{12}}{1 - r_{12}^2}\right)^2 \left(\frac{s_1}{\sigma_1}\right)^2 + \left(\frac{r_{24} - r_{14}r_{12}}{1 - r_{12}^2}\right)^2 \left(\frac{s_2}{\sigma_2}\right)^2 + 2\rho_{12} \frac{r_{14} - r_{24}r_{12}}{1 - r_{12}^2} \frac{r_{24} - r_{14}r_{12}}{1 - r_{12}^2} \frac{s_1 s_2}{\sigma_1 \sigma_2} \right\} \quad \text{(lix.)}.$$

$$\Sigma_3\Sigma_4\mathfrak{r}_{34} = \sigma_3\sigma_4 \left\{ \frac{r_{34}(1 - r_{12}^2) - r_{13}r_{14} - r_{23}r_{24} + r_{12}(r_{13}r_{24} + r_{14}r_{23})}{1 - r_{12}^2} + \frac{r_{13} - r_{23}r_{12}}{1 - r_{12}^2} \frac{r_{14} - r_{24}r_{12}}{1 - r_{12}^2} \left(\frac{s_1}{\sigma_1}\right)^2 + \frac{r_{23} - r_{13}r_{12}}{1 - r_{12}^2} \frac{r_{24} - r_{14}r_{12}}{1 - r_{12}^2} \left(\frac{s_2}{\sigma_2}\right)^2 + \rho_{12} \left(\frac{r_{13} - r_{23}r_{12}}{1 - r_{12}^2} \frac{r_{24} - r_{14}r_{12}}{1 - r_{12}^2} + \frac{r_{14} - r_{24}r_{12}}{1 - r_{12}^2} \frac{r_{23} - r_{13}r_{12}}{1 - r_{12}^2} \right) \frac{s_1 s_2}{\sigma_1 \sigma_2} \right\} \quad \text{(lx.)}.$$

$$s_1\Sigma_3\mathfrak{r}_{13} = s_1\sigma_3 \left\{ \frac{r_{13} - r_{12}r_{23}}{1 - r_{12}^2} \frac{s_1}{\sigma_1} + \frac{r_{23} - r_{12}r_{13}}{1 - r_{12}^2} \rho_{12} \frac{s_2}{\sigma_2} \right\} \quad \text{(lxi.)}.$$

$$s_1\Sigma_4\mathfrak{r}_{14} = s_1\sigma_4 \left\{ \frac{r_{14} - r_{12}r_{24}}{1 - r_{12}^2} \frac{s_1}{\sigma_1} + \frac{r_{24} - r_{12}r_{14}}{1 - r_{12}^2} \rho_{12} \frac{s_2}{\sigma_2} \right\} \quad \text{(lxii.)}.$$

$$s_2\Sigma_3\mathfrak{r}_{23} = s_2\sigma_3 \left\{ \frac{r_{23} - r_{12}r_{13}}{1 - r_{12}^2} \frac{s_2}{\sigma_2} + \frac{r_{13} - r_{12}r_{23}}{1 - r_{12}^2} \rho_{12} \frac{s_1}{\sigma_1} \right\} \quad \text{(lxiii.)}.$$

$$s_2\Sigma_4\mathfrak{r}_{24} = s_2\sigma_4 \left\{ \frac{r_{24} - r_{12}r_{14}}{1 - r_{12}^2} \frac{s_2}{\sigma_2} + \frac{r_{14} - r_{12}r_{24}}{1 - r_{12}^2} \rho_{12} \frac{s_1}{\sigma_1} \right\} \quad \text{(lxiv.)}.$$

Finally, for the change of means of the non-directly selected organs, we have:

$$x'_3 = \frac{r_{13} - r_{12}r_{23}}{1 - r_{12}^2} \frac{\sigma_3}{\sigma_1} h_1 + \frac{r_{23} - r_{12}r_{13}}{1 - r_{12}^2} \frac{\sigma_3}{\sigma_2} h_2 \quad \text{(lxv.)}.$$

$$x'_4 = \frac{r_{14} - r_{12}r_{24}}{1 - r_{12}^2} \frac{\sigma_4}{\sigma_1} h_1 + \frac{r_{24} - r_{12}r_{14}}{1 - r_{12}^2} \frac{\sigma_4}{\sigma_2} h_2 \quad \text{(lxvi.)}.$$

If we write (lxv.) and (lxvi.) in the form—

$$\left.\begin{aligned} x'_3 &= \beta_{13}\frac{\sigma_3}{\sigma_1}h_1 + \beta_{23}\frac{\sigma_3}{\sigma_2}h_2 \\ x'_4 &= \beta_{14}\frac{\sigma_4}{\sigma_1}h_1 + \beta_{24}\frac{\sigma_4}{\sigma_2}h_2 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxvii.)},$$

the β's are the partial regression coefficients, and the whole solution can be expressed in terms of them. Thus :

$$\Sigma_3^2 = \sigma_3^2\left\{1 - \beta_{13}r_{13} - \beta_{23}r_{23} + \beta_{13}^2\left(\frac{s_1}{\sigma_1}\right)^2 + \beta_{23}^2\left(\frac{s_2}{\sigma_2}\right)^2 + 2\rho_{12}\beta_{13}\beta_{23}\frac{s_1s_2}{\sigma_1\sigma_2}\right\}. \quad \text{(lxviii.)}.$$

$$\Sigma_3\Sigma_4\mathfrak{r}_{34} = \sigma_3\sigma_4\left\{r_{34} - \beta_{13}r_{14} - \beta_{23}r_{24} + \beta_{13}\beta_{14}\left(\frac{s_1}{\sigma_1}\right)^2 + \beta_{23}\beta_{24}\left(\frac{s_2}{\sigma_2}\right)^2 + \rho_{12}\left(\beta_{13}\beta_{24} + \beta_{23}\beta_{14}\right)\frac{s_1s_2}{\sigma_1\sigma_2}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxix.)}.$$

$$s_1\Sigma_3\mathfrak{r}_{13} = s_1\sigma_3\left\{\beta_{13}\frac{s_1}{\sigma_1} + \rho_{12}\beta_{23}\frac{s_2}{\sigma_2}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxx.)}.$$

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

$$\mathfrak{r}_{34} = \frac{r_{34}\left(1 - r_{12}^2\right) - r_{13}r_{14} - r_{23}r_{24} + r_{12}\left(r_{13}r_{24} + r_{14}r_{23}\right)}{\sqrt{\left(1 - r_{12}^2 - r_{23}^2 - r_{13}^2 + 2r_{12}r_{13}r_{23}\right)}\sqrt{\left(1 - r_{12}^2 - r_{24}^2 - r_{14}^2 + 2r_{12}r_{14}r_{24}\right)}} \quad . \quad \text{(lxxi.)}.$$

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 :

$$\Sigma_3^2/\sigma_3^2 = \cdot196 + \cdot181\left(\frac{s_1}{\sigma_1}\right)^2 + \cdot268\left(\frac{s_2}{\sigma_2}\right)^2 + 2\,\frac{s_1}{\sigma_1}\,\frac{s_2}{\sigma_2}\,\rho_{12} \times \cdot221,$$

$$\Sigma_4^2/\sigma_4^2 = \cdot354 + \cdot108\left(\frac{s_1}{\sigma_1}\right)^2 + \cdot265\left(\frac{s_2}{\sigma_2}\right)^2 + 2\,\frac{s_1}{\sigma_1}\,\frac{s_2}{\sigma_2}\,\rho_{12} \times \cdot169.$$

These give for the particular case :

$$\Sigma_3 = 1\cdot39, \qquad \Sigma_4 = 1\cdot11.$$

Turning to the correlation of humerus and radius, we have by (lx.) :

$$\mathrm{r}_{34} = \frac{\sigma_3}{\Sigma_3}\,\frac{\sigma_4}{\Sigma_4}\left\{\cdot125 + \cdot140\left(\frac{s_1}{\sigma_1}\right)^2 + \cdot267\left(\frac{s_2}{\sigma_2}\right)^2 + \cdot389\,\rho_{12}\,\frac{s_1}{\sigma_1}\,\frac{s_2}{\sigma_2}\right\},$$

giving for the particular case :

$$\mathrm{r}_{34} = \cdot799.$$

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 :

$$\mathrm{r}_{13} = \frac{\sigma_3}{\Sigma_3}\left\{\cdot426\,\frac{s_1}{\sigma_1} + \cdot518\,\rho\,\frac{s_2}{\sigma_2}\right\},$$

$$\mathrm{r}_{14} = \frac{\sigma_4}{\Sigma_4}\left\{\cdot329\,\frac{s_1}{\sigma_1} + \cdot515\,\rho\,\frac{s_2}{\sigma_2}\right\},$$

$$\mathrm{r}_{23} = \frac{\sigma_3}{\Sigma_3}\left\{\cdot518\,\frac{s_2}{\sigma_2} + \cdot426\,\rho\,\frac{s_1}{\sigma_1}\right\},$$

$$\mathrm{r}_{24} = \frac{\sigma_4}{\Sigma_4}\left\{\cdot515\,\frac{s_2}{\sigma_2} + \cdot329\,\rho\,\frac{s_1}{\sigma_1}\right\}.$$

These yield for our particular case :

$$\mathrm{r}_{13} = \cdot819, \qquad \mathrm{r}_{14} = \cdot694,$$
$$\mathrm{r}_{23} = \cdot845, \qquad \mathrm{r}_{24} = \cdot768.$$

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,

$$M_1 = 166{\cdot}26 + 1{\cdot}037h_2 + 1{\cdot}546h_3.$$

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

$$M_1 = 166{\cdot}26 - 3{\cdot}79 - 5{\cdot}01 = 157{\cdot}46.$$

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 :

$$\mathfrak{r}_{13} = \frac{\mu_1 r_{13}}{\sqrt{1 - (1 - \mu_1^2)\, r_{13}^2}},$$

and

$$\Sigma_2^2 = \sigma_2^2 (1 - (1 - \mu_1^2)\, r_{13}^2).$$

The first may be written

$$\mathfrak{r}_{13} = \frac{r_{13}}{\sqrt{r_{13}^2 + \dfrac{1 - r_{13}^2}{\mu_1^2}}},$$

or, we see that $\mathfrak{r}_{13}$ will decrease, as μ_1 decreases. Thus if $r_{13} = \cdot 4$ we have for

$\mu_1 = 1/2,$	$\mathfrak{r}_{13} = \cdot 2132,$	$\Sigma_2/\sigma_2 = \cdot 8367,$
$\mu_1 = 1/4,$	$\mathfrak{r}_{13} = \cdot 1085,$	$\Sigma_2/\sigma_2 = \cdot 7906,$
$\mu_1 = 1/8,$	$\mathfrak{r}_{13} = \cdot 0545,$	$\Sigma_2/\sigma_2 = \cdot 7786.$

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.*, $\Sigma_2 \mathfrak{r}_{13}/s_1 = \sigma_2 r_{13}/\sigma_1$, 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

$$\Sigma_3^2 = \sigma_3^2 (1 + 2\rho_{12}r_{13}r_{23}) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxiii.)},$$

$$\Sigma_3 \mathfrak{r}_{13} = \sigma_3 (r_{13} + \rho_{12}r_{23}) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxiv.)},$$

whence we find for $r_{13} = r_{23} = \cdot 4$ and

$\rho_{12} = \cdot 1$,	$\Sigma_3/\sigma_3 = 1 \cdot 0149$,	$\mathfrak{r}_{13} = \cdot 4335$,
$\rho_{12} = \cdot 2$,	$= 1 \cdot 0315$,	$= \cdot 4653$,
$\rho_{12} = \cdot 3$,	$= 1 \cdot 0469$,	$= \cdot 4967$,
$\rho_{12} = \cdot 4$,	$= 1 \cdot 0621$,	$= \cdot 5271$,
$\rho_{12} = \cdot 5$,	$= 1 \cdot 0770$,	$= \cdot 5571$,
$\rho_{12} = 1$,	$= 1 \cdot 1489$,	$= \cdot 6963$.

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:

$$\mathfrak{r}_{34} = \frac{(r_{34} - r_{13}r_{14})(1-\mu_1^2) + \mu_1^2 r_{34}}{\sqrt{\{1-(1-\mu_1^2)r_{13}^2\}\{1-(1-\mu_1^2)r_{14}^2\}}} \quad . \quad . \quad . \quad . \quad \text{(lxxvii.)}.$$

Now $r_{13} = r_{14}$, if the offspring are of one sex; hence:

$$\mathfrak{r}_{34} = \frac{r_{34} - r_{13}^2(1-\mu_1^2)}{1 - r_{13}^2(1-\mu_1^2)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxviii.)}.$$

If we take $r_{13} = \cdot 4$ and $r_{34} = \cdot 5$ as reasonable values, we have

$$\mathfrak{r}_{34} = \frac{\cdot 34 + \cdot 16\mu_1^2}{\cdot 84 + \cdot 16\mu_1^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxix.)}.$$

Thus $\mathfrak{r}_{34}$ will be greatest when μ_1 is greatest, *i.e.*, when there is no selection, and will decrease with μ_1 until it reaches $\cdot 4048$, 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

$$\mathfrak{r}_{34} = \frac{r_{34} - \beta^2\{1-\mu_1^2 + 1 - \mu_2^2 + 2(r_{12} - \rho_{12}\mu_1\mu_2)\}}{1 - \beta^2\{1-\mu_1^2 + 1 - \mu_2^2 + 2(r_{12} - \rho_{12}\mu_1\mu_2)\}} \quad . \quad . \quad . \quad \text{(lxxx.)}.$$

To reduce to numbers, suppose $\mu_1 = \mu_2$, and $r_{12} = 0$ for the general population. We have

$$\mathfrak{r}_{34} = \frac{r_{34} - 2r^2\{1-\mu_1^2(1+\rho_{12})\}}{1 - 2r^2\{1-\mu_1^2(1+\rho_{12})\}} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxxi.)}$$

* In general the value of $\mathfrak{r}_{34}$ ranges from r_{34} down to $\frac{r_{34} - r_{13}^2}{1 - r_{13}^2}$.

Hence if we put $r_{34} = \cdot 5$, and $r = \cdot 4$,

$$r_{34} = \frac{\cdot 18 + \cdot 32\mu_1^2(1 + \rho_{12})}{\cdot 68 + \cdot 32\mu_1^2(1 + \rho_{12})} \quad \ldots\ldots\ldots \quad \text{(lxxxii.)}.$$

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

$$r_{31} = 1 - \frac{1 - r_{34}}{1 - 2r^2 + 2r^2\mu_1^2(1 + \rho_{12})}.$$

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

$$\beta_{13} = r = \beta_{24}, \quad \beta_{23} = \beta_{14} = 0.$$

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 $\rho_{12} = \gamma r'$,

$$\mathfrak{r}_{34} = r' \times \frac{1 + \frac{4}{21}\gamma\mu_1^2}{1 + \frac{4}{21}\mu_1^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxxv.)}.$$

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 $\mathfrak{r}_{34}/r'$, 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 $\bar{y} + y' = f(\bar{x}) + x' f'(\bar{x}) + \frac{x^2}{1\cdot 2} f''(\bar{x}) + \ldots$ 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

$$Z = Z_0 \text{ expt. } -\tfrac{1}{2}\{c_{11}x_1^2 + c_{22}x_2^2 + \ldots + c_{nn}x_n^2 + 2c_{12}x_1x_2 + \ldots + 2c_{n-1,n}x_{n-1}x_n\} \quad \text{(lxxxvi.)}.$$

Let the probability of survival be given by

$$p = p_0 f(x_1 - k_1, \quad x_2 - k_2, \quad x_3 - k_3, \ldots x_n - k_n) \quad \text{(lxxxvii.)},$$

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 :

$$z = Z \times p \quad \text{(lxxxviii.)},$$

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 :

$$z = z_0 \text{ expt. } -\tfrac{1}{2}\{b_{11}(x_1 - h_1)^2 + b_{22}(x_2 - h_2)^2 + \ldots + b_{nn}(x_n - h_n)^2 + 2b_{12}(x_1 - b_1)(x_2 - b_2) - \ldots + 2b_{n-1,n}(x_{n-1} - h_{n-1})(x_n - h_n) \quad \text{(lxxxix.)}.$$

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.

$$f(x_1 - k_1, x_2 - k_2, \ldots x_n - k_n) = \text{expt.} - \tfrac{1}{2}\{a_{11}(x_1 - k_1)^2 + a_{22}(x_2 - k_2)^2 + \ldots + 2a_{12}(x_1 - k_1)(x_2 - k_2) - \ldots + 2a_{n-1,n}(x_{n-1} - k_{n-1})(x_n - k_n)\} \quad . \quad \text{(xc.)}.$$

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 $x_1 = k_1$, $x_2 = k_2$, . . . $x_n = k_n$, and then note that:

$$p_0 = \frac{z(x_1 = k_1, x_2 = k_2, \ldots x_n = k_n)}{Z(x_1 = k_1, x_2 = k_2, \ldots x_n = k_n)} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xci.)}.$$

Since $$p = z/Z, \quad \text{and} \quad z = p/Z^{-1},$$

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 $-c_{22}$ . . . c_{12} to $-c_{12}$ . . . $c_{n-1,n}$ to $-c_{n-1,n}$. If σ_1, σ_2, . . . σ_n be the standard deviations of the unselected population, this amounts to changing σ_1, σ_2, . . . σ_n to $\sqrt{-1}\sigma_1$, $\sqrt{-1}\sigma_2$, . . . $\sqrt{-1}\sigma_n$ 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, . . . m_n, σ_1, σ_2, . . . σ_n, and r_{12}, r_{13}, r_{23}, . . . $r_{n-1,n}$.

Let the selected population be defined by $m_1 + \mathrm{H}_1$, $m_2 + \mathrm{H}_2$, . . . $m_n + \mathrm{H}_n$, Σ_1, Σ_2, . . . Σ_n, and $\mathfrak{r}_{12}$, $\mathfrak{r}_{13}$, $\mathfrak{r}_{23}$, . . . $\mathfrak{r}_{n-1,n}$.

Let the constants of the probability of survival function, or a_{11}, a_{22}, . . . a_{nn}, a_{12} . . . $a_{n-1,n}$, be expressed by $\bar{s}_1$, $\bar{s}_2$, . . . $\bar{s}_n$, $\bar{\rho}_{12}$, $\bar{\rho}_{13}$, . . . $\bar{\rho}_{n-1,n}$ 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 $\mathfrak{r}$'s.

Equating the squares, products, and linear terms in the x's in the equation $p = z/Z$, we have at once the system:

$$\left.\begin{aligned} a_{uu} &= b_{uu} - c_{uu}, \\ a_{uv} &= b_{uv} - c_{uv}, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xcii.)},$$

for all values of u and v from 1 to n.

$$-a_{v1}k_1 - a_{v2}k_2 - \ldots - a_{vv}k_v - \ldots - a_{vn}k_n = -b_{v1}\mathrm{H}_1 - b_{v2}\mathrm{H}_2 - \ldots - b_{vv}\mathrm{H}_v - \ldots - b_{vn}\mathrm{H}_n \quad . \quad . \quad \text{(xciii.)},$$

for all values of v from 1 to n.

* These must not be confused with the s_1, s_2, . . . s_q, ρ_{12}, ρ_{13}, ρ_{23}, . . . $\rho_{q,q-1}$ 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 α's we find:

$$(c_{r1} - b_{r1})\,k_1 + (c_{r2} - b_{r2})\,k_2 + \ldots + (c_{rv} - b_{rv})\,k_v + \ldots + (c_{rn} - b_{rn})\,k_n = -\,b_{r1}\mathrm{H}_1 - b_{r2}\mathrm{H}_2 - \ldots - b_{rv}\mathrm{H}_v - \ldots - b_{rn}\mathrm{H}_n \quad \text{(xciv.)}.$$

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 uv^{th} constituent.

Then we have at once:

$$\begin{aligned} \Delta \times k_v = {} & \Delta_{v1}(b_{11}\mathrm{H}_1 + b_{12}\mathrm{H}_2 + \ldots + b_{1n}\mathrm{H}_n) \\ & + \Delta_{v2}(b_{21}\mathrm{H}_1 + b_{22}\mathrm{H}_2 + \ldots + b_{2n}\mathrm{H}_n) \\ & + \Delta_{v3}(b_{31}\mathrm{H}_1 + b_{32}\mathrm{H}_2 + b_{33}\mathrm{H}_3 \ldots + b_{3n}\mathrm{H}_n) \\ & \ldots \\ & + \Delta_{vn}(b_{n1}\mathrm{H}_1 + b_{n2}\mathrm{H}_2 + \ldots + b_{nn}\mathrm{H}_n). \end{aligned}$$

Or:

$$\begin{aligned} \Delta \times k_v = {} & \mathrm{H}_1(b_{11}\Delta_{v1} + b_{12}\Delta_{v2} + b_{13}\Delta_{v3} + \ldots + b_{1n}\Delta_{vn}) \\ & + \mathrm{H}_2(b_{21}\Delta_{v1} + b_{22}\Delta_{v2} + b_{23}\Delta_{v3} + \ldots + b_{2n}\Delta_{vn}) \\ & + \mathrm{H}_3(b_{31}\Delta_{v1} + b_{32}\Delta_{v2} + b_{33}\Delta_{v3} + \ldots + b_{3n}\Delta_{vn}) \\ & + \ldots \\ & + \mathrm{H}_n(b_{n1}\Delta_{v1} + b_{n2}\Delta_{v2} + b_{n3}\Delta_{v3} + \ldots + b_{nn}\Delta_{vn}) \quad \ldots \quad \text{(xcv.)}. \end{aligned}$$

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:

$$\begin{aligned} \Delta' \times \mathrm{H}_v = {} & k_1(a_{11}\Delta'_{v1} + a_{12}\Delta'_{v2} + a_{13}\Delta'_{v3} + \ldots + a_{1n}\Delta'_{vn}) \\ & + k_2(a_{21}\Delta'_{v1} + a_{22}\Delta'_{v2} + a_{23}\Delta'_{v3} + \ldots + a_{2n}\Delta'_{vn}) \\ & + k_3(a_{31}\Delta'_{v1} + a_{32}\Delta'_{v3} + a_{33}\Delta'_{v3} + \ldots + a_{3n}\Delta'_{vn}) \\ & + \ldots \\ & + k_n(a_{n1}\Delta'_{v1} + a_{n2}\Delta'_{v3} + a_{n3}\Delta'_{v3} - \ldots + a_{nn}\Delta'_{vn}) \quad \ldots \quad \text{(xcvi.)}, \end{aligned}$$

where Δ' is the determinant

$$\begin{vmatrix} (c_{11}+a_{11}), & (c_{12}+a_{12}), & (c_{13}+a_{13}) \ldots (c_{1n}+a_{1n}) \\ (c_{21}+a_{21}), & (c_{22}+a_{22}), & (c_{23}+a_{23}) \ldots (c_{2n}+a_{2n}) \\ (c_{31}+a_{31}), & (c_{32}+a_{32}), & (c_{33}+a_{33}) \ldots (c_{3n}+a_{3n}) \\ \cdot \cdot \cdot \cdot & \cdot \cdot \cdot \cdot & \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\ \cdot \cdot \cdot \cdot & \cdot \cdot \cdot \cdot & \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\ (c_{n1}+a_{n1}), & (c_{n2}+a_{n2}), & (c_{n3}+a_{n3}) \ldots (c_{nn}+a_{nn}) \end{vmatrix}$$

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_r to be equal to H_r. 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

$$p_0 = \frac{z_0}{Z_0} e^{\frac{H^2}{2\sigma^2(1-\lambda^2)}} \quad \text{(xcix.)}.$$

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 :

$$H = -6{\cdot}848, \qquad \lambda = {\cdot}7511.$$

These give :

$$k = -15{\cdot}712, \qquad \bar{s} = 4{\cdot}370.$$

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

$$p_0 = 51{\cdot}0474\, n/N,$$

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

$$\frac{1}{\bar{s}_1^2(1-\bar{\rho}_{12}^2)} = \frac{1}{\Sigma_1^2(1-\mathfrak{r}_{12}^2)} - \frac{1}{\sigma_1^2(1-r_{12}^2)} \quad \text{(ciii.)},$$

$$\frac{1}{\bar{s}_2^2(1-\bar{\rho}_{12}^2)} = \frac{1}{\Sigma_2^2(1-\mathfrak{r}_{12}^2)} - \frac{1}{\sigma_2^2}\frac{1}{(1-r_{12}^2)} \quad \text{(civ.)},$$

$$\frac{1}{\bar{s}_1\bar{s}_2}\frac{\bar{\rho}_{12}}{1-\bar{\rho}_{12}^2} = \frac{\mathfrak{r}_{12}}{\Sigma_1\Sigma_2(1-\mathfrak{r}_{12}^2)} - \frac{r_{12}}{\sigma_1\sigma_2(1-r_{12}^2)} \quad \text{(cv.)},$$

$$\frac{k_1}{\bar{s}_1^2(1-\bar{\rho}_{12}^2)} - \frac{k_2\bar{\rho}_{12}}{\bar{s}_1\bar{s}_2(1-\bar{\rho}_{12}^2)} = \frac{\mathrm{H}_1}{\Sigma_1^2(1-\mathfrak{r}_{12}^2)} - \frac{\mathrm{H}_2\mathfrak{r}_{12}}{\Sigma_1\Sigma_2(1-\mathfrak{r}_{12}^2)} \quad \text{(cvi.)},$$

$$-\frac{k_1\bar{\rho}_{12}}{\bar{s}_1\bar{s}_2(1-\bar{\rho}_{12}^2)} + \frac{k_2}{\bar{s}_2^2(1-\bar{\rho}_{12}^2)} = -\frac{r_{12}\mathrm{H}_1}{\Sigma_1\Sigma_2(1-\mathfrak{r}_{12}^2)} + \frac{\mathrm{H}_2}{\Sigma_2^2(1-\mathfrak{r}_{12}^2)} \quad \text{(cvii.)}.$$

Let $\Sigma_1/\sigma_1 = \lambda_1$, $\Sigma_2/\sigma_2 = \lambda_2$ measure the stringency of the selection, and $\mu = \sqrt{\frac{1-\mathfrak{r}_{12}^2}{1-r_{12}^2}}$ measure the change in correlation.* Then solving the above equations we find :

$$\bar{\rho}_{12} = \frac{\mathfrak{r}_{12} - r_{12}\mu^2\lambda_1\lambda_2}{\sqrt{1-\mu^2\lambda_1^2}\sqrt{1-\mu^2\lambda_2^2}} \quad \text{(cviii.)},$$

$$\bar{s}_1 = \frac{\Sigma_1\sqrt{1-r_{12}^2-(1-\mathfrak{r}_{12}^2)\lambda_2^2}}{\sqrt{1-r_{12}^2-\lambda_1^2-\lambda_2^2+(1-\mathfrak{r}_{12}^2)\lambda_1^2\lambda_2^2+2\mathfrak{r}_{12}r_{12}\lambda_1\lambda_2}} \quad \text{(cix.)},$$

$$\bar{s}_2 = \frac{\Sigma_2\sqrt{1-r_{12}^2-(1-\mathfrak{r}_{12}^2)\lambda_1^2}}{\sqrt{1-r_{12}^2-\lambda_1^2-\lambda_2^2+(1-\mathfrak{r}_{12}^2)\lambda_1^2\lambda_2^2+2\mathfrak{r}_{12}r_{12}\lambda_1\lambda_2}} \quad \text{(cx.)},$$

$$\frac{k_1}{\sigma_1} = \frac{\mathrm{H}_1}{\sigma_1}\frac{1-r_{12}^2-\lambda_2^2+\mathfrak{r}_{12}r_{12}\lambda_1\lambda_2}{\beta} + \frac{k_2}{\sigma_2}\frac{\mathfrak{r}_{12}\lambda_2\lambda_1-r_{12}\lambda_1^2}{\beta} \quad \text{(cxi.)},$$

$$\frac{k_2}{\sigma_2} = \frac{k_1}{\sigma_1}\frac{\mathfrak{r}_{12}\lambda_1\lambda_2-r\lambda_2^2}{\beta} + \frac{k_2}{\sigma_2}\frac{1-r_{12}^2-\lambda_1^2+\mathfrak{r}_{12}r_{12}\lambda_1\lambda_2}{\beta} \quad \text{(cxii.)},$$

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 :

$$\mathfrak{r}_{12} = \frac{\bar{\rho}_{12}+r_{12}\nu^2\kappa_1\kappa_2}{\sqrt{1+\nu^2\kappa_1^2}\sqrt{1+\nu^2\kappa_2}} \quad \text{(cxiii.)},$$

* 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 $$\nu = \sqrt{\frac{1-\bar{\rho}_{12}^2}{1-r_{12}^2}}, \quad \kappa_1 = s_1/\sigma_1 \quad \text{and} \quad \kappa_2 = s_2/\sigma_2,$$

$$\Sigma_1 = \frac{\bar{s}_1\sqrt{1-r_{12}^2+(1-\bar{\rho}_{12}^2)\,\kappa_2^2}}{\sqrt{1-r_{12}^2+\kappa_1^2+\kappa_2^2+(1-\bar{\rho}_{12}^2)\,\kappa_1^2\kappa_2^2-2\bar{\rho}_{12}r_{12}\kappa_1\kappa_2}} \quad \text{(cxiv.)},$$

$$\Sigma_2 = \frac{\bar{s}_2\sqrt{1-r_{12}^2+(1-\bar{\rho}_{12}^2)\,\kappa_1^2}}{\sqrt{1-r_{12}^2+\kappa_1^2+\kappa_2^2+(1-\bar{\rho}_{12}^2)\,\kappa_1^2\kappa_2^2-2\bar{\rho}_{12}r_{12}\kappa_1\kappa_2}} \quad \text{(cxv.)},$$

$$\frac{\mathrm{H}_1}{\sigma_1} = \frac{k_1}{\sigma_1}\frac{1-r_{12}^2+\kappa_2^2-\bar{\rho}_{12}r_{12}\kappa_1\kappa_2}{\gamma} + \frac{k_2}{\sigma_2}\frac{r\kappa_1^2-\bar{\rho}_{12}\kappa_1\kappa_2}{\gamma} \quad \text{(cxvi.)},$$

$$\frac{\mathrm{H}_2}{\sigma_2} = \frac{k_1}{\sigma_1}\frac{r_{12}\kappa_2^2-\bar{\rho}_{12}\kappa_1\kappa_2}{\gamma} + \frac{k_2}{\sigma_2}\frac{1-r_{12}^2+\kappa_1^2-\bar{\rho}_{12}r_{12}\kappa_1\kappa_2}{\gamma} \quad \text{(cxvii.)},$$

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.	$= m_1 + \mathrm{H}.$
	σ_1	2·372 ,,		1·898 ,,	$= \Sigma_1.$
Humerus	m_2	33·010 ,,		29·502 ,,	$= m_2 + \mathrm{H}_2.$
	σ_2	1·538 ,,		1·343 ,,	$= \Sigma_2.$
Correlation of Femur and Humerus	r_{12}	·8421		·8584	$= \mathfrak{r}_{12}.$

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 :

$$p = 208{,}425\,\frac{n}{N}\,\text{expt.} - \tfrac{1}{2}\{\cdot 222056\,(x + 17\cdot 8464)^2 + \cdot 326{,}960\,(y + 15\cdot 8547)^2 - \cdot 486{,}450\,(x + 17\cdot 8464)\,(y + 15\cdot 8547)\}.$$

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 = 27·382 centims. and humerus = 17·155 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

$$k_1 = \quad 1\cdot 2514\ \mathrm{H}_1 + 3\cdot 4971\ \mathrm{H}_2,$$
$$k_2 = -\,1\cdot 0266\ \mathrm{H}_1 + 5\cdot 8242\ \mathrm{H}_2,$$

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

$$z = \frac{n}{2\pi\Sigma_1\Sigma_2\sqrt{(1-\mathrm{r}_{12}^2)}}\,e^{-\frac{1}{2}\frac{1}{(1-\mathrm{r}_{12}^2)}\left\{\frac{(x-\mathrm{H}_1)^2}{\Sigma_1^2} - \frac{2\mathrm{r}_{12}(x-\mathrm{H}_1)(y-\mathrm{H}_2)}{\Sigma_1\Sigma_2} + \frac{(y-\mathrm{H}_2)^2}{\Sigma_2^2}\right\}}.$$

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 $\kappa^2 = \dfrac{1}{1-r^2}\left\{\dfrac{x^2}{\sigma_1^2}-\dfrac{2rxy}{\sigma_1\sigma_2}+\dfrac{y^2}{\sigma_2^2}\right\}$; 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,

$$N_\kappa = N\int_\kappa^\infty e^{-\frac{1}{2}\kappa^2}\kappa d\kappa = Ne^{-\frac{1}{2}\kappa^2} \quad \ldots\ldots\ldots \text{(cxviii.)}.$$

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: $\sin 2\chi = \sin 2\psi \sin D$, $\quad \tan 2\phi = \tan 2\psi \cos D$.

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

$$p = 208{,}425\ \frac{n}{\mathrm{N}}\ e^{-\frac{1}{2}\left\{\frac{X^2}{(4\cdot4115)^2}+\frac{Y^2}{(\cdot9775)^2}\right\}}.$$

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, $\kappa_1 \times 4{\cdot}1149 = 11{\cdot}44$, or $\kappa_1 = 2{\cdot}5932$. 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 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.

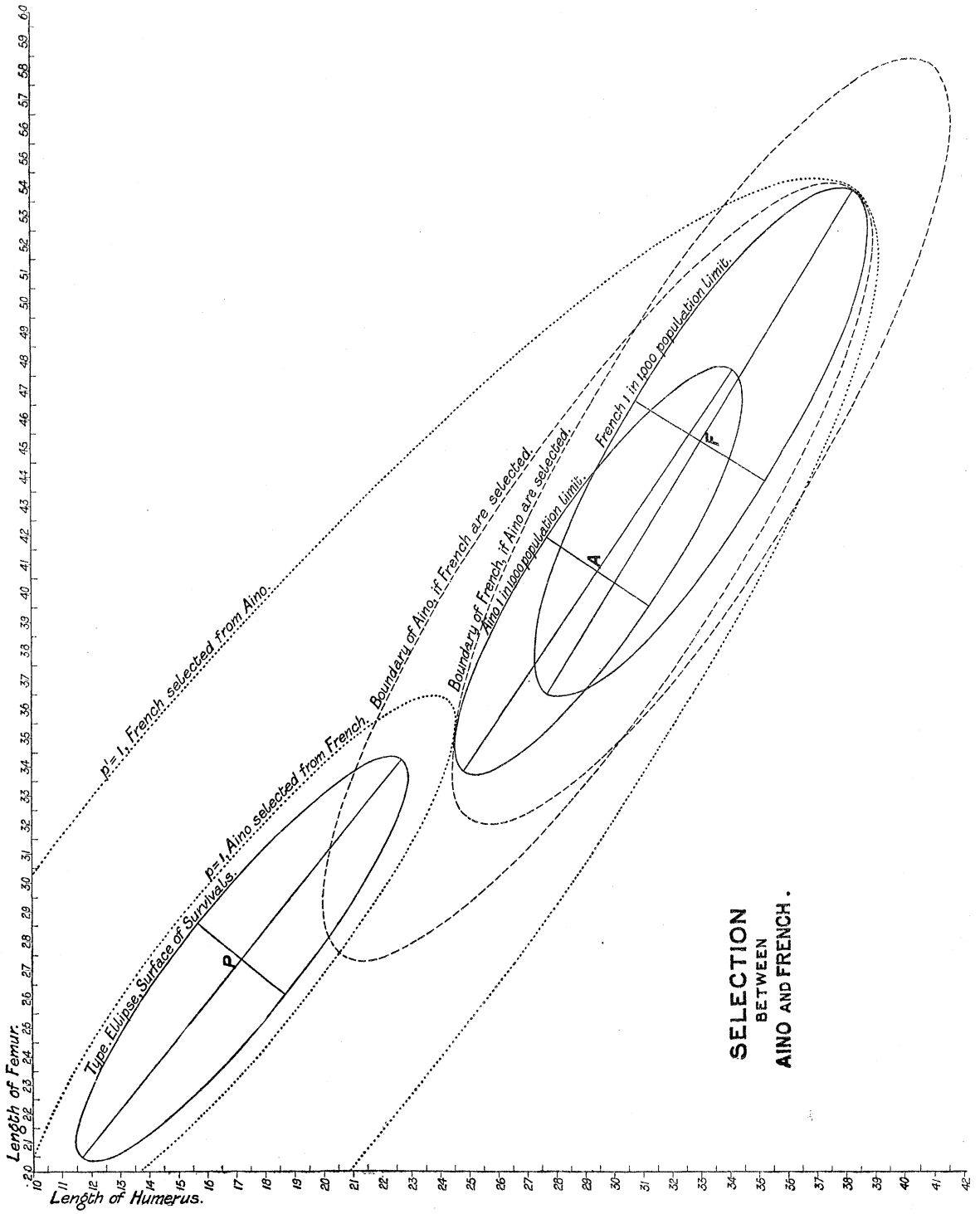
some degenerate offshoot of a race superior to themselves than a sample of the primitive people from which the Circassian races may be supposed to have sprung.

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

$$\kappa_F = \left(\frac{15{\cdot}285}{2{\cdot}7338}\right) = 5{\cdot}5911, \qquad \kappa_A = \frac{15{\cdot}890}{2{\cdot}2514} = 7{\cdot}0578.$$

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:

$$\nu_F = {\cdot}000{,}000{,}163, \qquad \nu_A = {\cdot}000{,}000{,}000{,}015.$$

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.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

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.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

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

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

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)	$+\infty$	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)	$+\infty$	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)	$-\infty$	−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)	$+\infty$	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)	$+\infty$	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.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

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)	$-\cdot 3333$	$-\cdot 3324$	$-\cdot 3298$	$-\cdot 3254$	$-\cdot 3193$	$-\cdot 3115$	$-\cdot 3021$	$-\cdot 2911$	$-\cdot 2788$	$-\cdot 2650$	$-\cdot 2500$
(g)	$-\cdot 4472$	$-\cdot 4449$	$-\cdot 4377$	$-\cdot 4261$	$-\cdot 4103$	$-\cdot 3907$	$-\cdot 3675$	$-\cdot 3414$	$-\cdot 3128$	$-\cdot 2822$	$-\cdot 2500$
(h)	$-\cdot 6831$	$-\cdot 67565$	$-\cdot 6539$	$-\cdot 61995$	$-\cdot 5764$	$-\cdot 5262$	$-\cdot 4721$	$-\cdot 4160$	$-\cdot 3596$	$-\cdot 3040$	$-\cdot 2500$
(i)	$-\infty$	$-5\cdot 1364$	$-2\cdot 5270$	$-1\cdot 6389$	$-1\cdot 1814$	$-\cdot 8962$	$-\cdot 6974$	$-\cdot 5481$	$-\cdot 4299$	$-\cdot 3325$	$-\cdot 2500$
(j)	$-\cdot 6667$	$-\cdot 6611$	$-\cdot 6447$	$-\cdot 6181$	$-\cdot 5822$	$-\cdot 5384$	$-\cdot 4881$	$-\cdot 4327$	$-\cdot 3736$	$-\cdot 3123$	$-\cdot 2500$
(k)	$-1\cdot 0911$	$-1\cdot 0758$	$-1\cdot 0317$	$-\cdot 9628$	$-\cdot 8753$	$-\cdot 7751$	$-\cdot 6683$	$-\cdot 5594$	$-\cdot 4519$	$-\cdot 3483$	$-\cdot 2500$

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)	$-\cdot 6000$	$-\cdot 5989$	$-\cdot 5958$	$-\cdot 5905$	$-\cdot 5831$	$-\cdot 5738$	$-\cdot 5625$	$-\cdot 5493$	$-\cdot 5345$	$-\cdot 5180$	$-\cdot 5$
(g)	$-\cdot 7453$	$-\cdot 7424$	$-\cdot 7335$	$-\cdot 7190$	$-\cdot 6993$	$-\cdot 6747$	$-\cdot 6459$	$-\cdot 6134$	$-\cdot 5778$	$-\cdot 5399$	$-\cdot 5$
(h)	$-1\cdot 0735$	$-1\cdot 0634$	$-1\cdot 0341$	$-\cdot 9884$	$-\cdot 9300$	$-\cdot 8630$	$-\cdot 7910$	$-\cdot 7169$	$-\cdot 6427$	$-\cdot 5701$	$-\cdot 5$
(j)	$-1\cdot 0000$	$-\cdot 9934$	$-\cdot 9737$	$-\cdot 9417$	$-\cdot 8987$	$-\cdot 8461$	$-\cdot 7858$	$-\cdot 7192$	$-\cdot 6484$	$-\cdot 5748$	$-\cdot 5$

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)	$-\cdot 8667$	$-\cdot 8654$	$-\cdot 8617$	$-\cdot 8555$	$-\cdot 8470$	$-\cdot 8361$	$-\cdot 8229$	$-\cdot 8076$	$-\cdot 7903$	$-\cdot 7710$	$-\cdot 75$
(g)	$-1\cdot 0435$	$-1\cdot 0399$	$-1\cdot 0293$	$-1\cdot 0119$	$-\cdot 9882$	$-\cdot 9588$	$-\cdot 9243$	$-\cdot 8854$	$-\cdot 8430$	$-\cdot 7976$	$-\cdot 75$

Table V. (B) is entirely impossible except (a). $r_{12} = r_{13} = 0$. $R = -1$ for all values of selective intensity λ.