VII. *Mathematical Contributions to the Theory of Evolution.—IV. On the Probable Errors of Frequency Constants and on the Influence of Random Selection on Variation and Correlation.*

By KARL PEARSON, *F.R.S.*, *and* L. N. G. FILON, *B.A.*, *University College, London.*

Received October 18,—Read November 25, 1897.

CONTENTS.

I. INTRODUCTORY.—GENERAL THEOREM.

(1) In earlier memoirs by one of the present authors, methods have been discussed for the calculation of the constants (a) of variation, normal or skew,* (b) of correla-

* "Mathematical Contributions to the Theory of Evolution.—II. On Skew Variation," 'Phil. Trans.,' A, vol. 186, pp. 343-414.

12.7.98.

tion, when normal.* The subject of skew correlation would now naturally present itself, but although several important conclusions with regard to skew correlation have been worked out, there are still difficulties which impede the completion of the memoir on that topic. Meanwhile Mr. G. U. YULE has shown that the constants of normal correlation are significant, if not completely descriptive, even in the case of skew correlation.† It seems desirable to take, somewhat out of its natural order, the subject of the present memoir, partly because the formulæ involved have been once or twice cited and several times used in memoirs by one of the present writers, and partly because the need of such formulæ seems to have been disregarded by various authors in somewhat too readily drawing conclusions from statistical data. Differences in the constants of variation or of correlation have been not infrequently asserted to be significant or non-significant of class or of type, or of race differences, without a due investigation of whether those differences are, from the standpoint of mathematical statistics, greater or less than the probable errors of the differences. Notwithstanding that every artificial or even random selection of a group out of a community changes not only the amount of variation, but the amount of correlation of the organs of its members as compared with those of the primitive group,‡ it has been supposed that correlation might be a racial constant, and the approximate constancy of coefficients of correlation of the same organs in allied species has been used as a valid argument. In the like manner differences in variation have been used as an argument for the activity of natural selection without a discussion of the probable errors of those differences.

In dealing with variation and correlation we find the distribution described by certain curves or surfaces fully determined when certain constants are known. These are the so-called constants of variation and correlation, the number of which may run up from two to a very considerable figure in the case of a complex of organs. If we deal with a complex of organs in two groups containing, say, n and n' individuals, we can only ascertain whether there is a significant or insignificant difference between those groups by measuring the extent to which the differences of corresponding constants exceed the probable errors of those differences. The probable error of a difference can at once be found by taking the square root of the sum of the squares of the probable errors of the quantities forming the difference. Hence the first step towards determining the significance of a group difference—*i.e.*, towards ascertaining whether it is really a class, race, or type difference—is to calculate the probable errors of the constants of variation and correlation of the individual groups. This will be the object of our first general theorem.

* "Mathematical Contributions to the Theory of Evolution.—III. Regression, Heredity, and Panmixia," 'Phil. Trans.,' A, vol. 187, pp. 253–318.

† "On the Significance of BRAVAIS' Formulæ for Skew Correlation," 'Roy. Soc. Proc.,' vol. 60, pp. 477–489.

‡ This will be sufficiently indicated in the latter part of the present memoir, but has been more fully dealt with in a paper on the "Influence of Selection on Correlation," written, but not yet published.

II. On the Probable Errors of a System of Frequency Constants.—General Theorem.

(2) Let there be a group of n individuals, for each of whom a complex of m organs is measured, and let $z\,\delta x_1\,\delta x_2\ldots\delta x_m$ be the frequency with which individuals having a complex of organs lying between x_1, $x_1+\delta x_1$; x_2, $x_2+\delta x_2$; $\ldots$ x_m, $x_m+\delta x_m$, occur in the total group of n. Here x shall measure the deviation of any organ from the mean of all like organs in the group. Let $h_1, h_2, h_3, \ldots h_m$, be the mean measurements on the organs, so that $h_1+x_1, h_2+x_2, \ldots h_m+x_m$, is the system giving the actual measurements on any individual. Then $h_1, h_2, h_3, \ldots h_m$, are the first set of constants of the frequency; they determine the "origin" of the frequency surface.

Let this frequency surface be given by

$$z = f(x_1, x_2, x_3, \ldots x_m;\ c_1, c_2, c_3, \ldots c_p),$$

where $c_1, c_2, c_3, \ldots c_p$, are p frequency constants, which define the form as distinguished from the position of the frequency surface, and which will be functions of standard deviations, moments, skewnesses, coefficients of correlation, &c., &c., of individual organs, and of pairs of organs in the complex.

The problem before us is to find the probable errors of the h's and the c's, which constants fully determine the position and shape of the frequency surface. Let $\sigma_{h_1}, \sigma_{h_2}, \sigma_{h_3}, \ldots \sigma_{h_m}, \sigma_{c_1}, \sigma_{c_2}, \ldots \sigma_{c_m}$, be the standard deviations of the quantities $h_1, h_2, \ldots h_m$, and $c_1, c_2, \ldots c_p$. Then a knowledge of these standard deviations will give us at once the probable errors of the frequency constants, for we have only to multiply the former by the numerical factor ·6745 to obtain the latter.

Let us now suppose that the value of the frequency constants had been $h_1+\Delta h_1$, $h_2+\Delta h_2, \ldots h_m+\Delta h_m$, $c_1+\Delta c_1$, $c_2+\Delta c_2, \ldots c_p+\Delta c_p$, instead of the observed values.

Then the frequency of any observed individual would have varied as

$$f(x_1+\Delta h_1, x_2+\Delta h_2, \ldots x_m+\Delta h_m;\ c_1+\Delta c_1, c_2+\Delta c_2, \ldots c_p+\Delta c_p)$$

instead of as

$$f(x_1, x_2, x_3, \ldots x_m;\ c_1, c_2, c_3, \ldots c_p).$$

Hence on this hypothesis the probability, P_Δ, of the set of individuals observed in the group actually occurring is to the probability, P_0, of the set occurring when the constants are $h_1, h_2, \ldots h_m$; $c_1, c_2, \ldots c_p$, in the ratio of the product of all quantities like $f(x_1+\Delta h_1, x_2+\Delta h_2, \ldots x_m+\Delta h_m;\ c_1+\Delta c_1, c_2+\Delta c_2, \ldots c_p+\Delta c_p)$ for all values of $x_1, x_2, \ldots x_m$, to the like product of all quantities like $f(x_1, x_2, \ldots x_m;$ $c_1, c_2, \ldots c_p)$, or

$$\frac{P_\Delta}{P_0} = \frac{\Pi f(x_1+\Delta h_1, x_2+\Delta h_2, \ldots x_m+\Delta h_m;\ c_1+\Delta c_1, c_2+\Delta c_2, \ldots c_p+\Delta c_p)}{\Pi f(x_1, x_2, \ldots x_m;\ c_1, c_2, \ldots c_p)}.$$

Taking logarithms, the products Π become sums S, or

$$\log(P_\Delta/P_0) = S\log f(x_1+\Delta h_1, x_2+\Delta h_2, \ldots x_m+\Delta h_m;\ c_1+\Delta c_1, c_2+\Delta c_2, \ldots c_p+\Delta c_p) \\ - S\log f(x_1, x_2, \ldots x_m;\ c_1, c_2, \ldots c_p).$$

Let the first summation now be expanded by TAYLOR's theorem, and typical terms up to the second order be written down. Then we have

$$\log(P_\Delta/P_0) = \Delta h_r\, S\frac{d}{dx_r}(\log f) + \tfrac{1}{2}(\Delta h_r)^2\, S\frac{d^2}{dx_r^2}(\log f) + \Delta h_r\,\Delta h_{r'}\, S\frac{d^2}{dx_r\,dx_s}(\log f) \\ + \Delta c_s\, S\frac{d}{dc_s}(\log f) + \tfrac{1}{2}(\Delta c_s)^2\, S\frac{d^2}{dc_s^2}(\log f) + \Delta c_s\,\Delta c_{s'}\, S\frac{d^2}{dc_s\,dc_{s'}}(\log f) \\ + \Delta h_r\,\Delta c_s\, S\frac{d^2}{dh_r dc_s}(\log f) + \ldots + \text{cubic terms in } \Delta h \text{ and } \Delta c + \&c.,$$

where f stands for $f(x_1, x_2, x_3, \ldots x_m, c_1, c_2, \ldots c_p)$.

Here r is to be given every value from 1 to m, and s to be given every value from 1 to p, but r' and s' in the third and sixth sums are only to be given values from 1 to m and 1 to p other than r and s respectively. In the above formula we may replace the sums by integrals, if we remember that the frequency of the system $x_1, x_2, \ldots x_m$, is simply $f\,\delta x_1\,\delta x_2 \ldots \delta x_n$.

Writing

$$\log(P_\Delta/P_0) = A_r\,\Delta h_r - \tfrac{1}{2}B_r(\Delta h_r)^2 + C_{rr'}\,\Delta h_r\,\Delta h_{r'} + D_s\,\Delta c_s - \tfrac{1}{2}E_s(\Delta c_s)^2 + F_{ss'}\,\Delta c_s\,\Delta c_{s'} \\ + G_{rs}\,\Delta h_r\,\Delta c_s + \&c. \ldots,$$

we will investigate the values of the constants separately.

First,

$$A_r = \iiint \ldots f\frac{d\log f}{dx_r}\,dx_1\,dx_2 \ldots dx_m, \\ = \iiint \ldots \frac{df}{dx_r}\,dx_1\,dx_2 \ldots dx_m, \\ = \iiint \ldots [f]\,dx_1\,dx_2 \ldots dx_{r-1}\,dx_{r+1} \ldots dx_m,$$

the integrals now not including one with regard to dx_r and $[f]$ denoting that f is to be taken between the extreme limits of f for x_r. Now in most cases of frequency the frequencies for extreme values of any organ are zero.* Hence $[f]$ equals nothing. Thus we have $A_r = 0$.

* In most cases, but not *invariably*, as, for example, in the case of some florets and petals. In the cases, however, in which A_r does not vanish, the conclusions finally reached will be the same, as A_r only marks a change of origin for the constant frequency distribution.

Secondly,

$$\begin{aligned} D_s &= \iiint \ldots f \frac{d \log f}{dc_s} dx_1\, dx_2 \ldots dx_m, \\ &= \iiint \ldots \frac{df}{dc_s} dx_1\, dx_2 \ldots dx_m, \\ &= \frac{d}{dc_s} \iiint \ldots f\, dx_1\, dx_2\, dx_m \\ &= dn/dc_s, \end{aligned}$$

where n is the total number of individuals measured, which is independent of c_s. Therefore $D_s = 0$.

Thirdly,

$$C_{rr'} = \iiint \ldots f \frac{d^2 \log f}{dx_r\, dx_s} dx_1\, dx_2 \ldots dx_m \quad \ldots \ldots \quad \text{(i.)}.$$

This will not as a rule vanish. If the frequency be normal, it will still not as a rule vanish. It will vanish either if the frequency be symmetrical about $x_r = 0$, and $x_s = 0$, or if there be no correlation between the x_r and x_s organs, *i.e.*, if f be of the form $f_1(x_r^2) \times f_2(x_s^2)$.

Fourthly,

$$G_{rs} = \iiint \ldots f \frac{d^2 \log f}{dx_r\, dc_s} dx_1\, dx_2 \ldots dx_m \quad \ldots \ldots \quad \text{(ii.)}.$$

This will not as a rule vanish. It will vanish, however, if the frequency of x_r be symmetrical about its mean.

Fifthly,

$$B_r = -\iiint \ldots f \frac{d^2 \log f}{d^2 x_r} dx_1\, dx_2 \ldots dx_m \quad \ldots \ldots \quad \text{(iii.)},$$

$$E_s = -\iiint \ldots f \frac{d^2 \log f}{d^2 c^s} dx_1\, dx_2 \ldots dx_m \quad \ldots \ldots \quad \text{(iv.)},$$

$$F_{ss'} = \iiint \ldots f \frac{d^2 \log f}{dc_s\, dc_{s'}} dx_1\, dx_2 \ldots dx_m \quad \ldots \ldots \quad \text{(v.)},$$

all of which will generally be finite, but admit, like $C_{rr'}$ and G_{rs}, of calculation when the form of the frequency f is given. Hence

$$\begin{aligned} P_\Delta = P_0 \text{ expt.} - \tfrac{1}{2} \{ & B_r (\Delta h_r)^2 - 2C_{rr'} \Delta h_r \Delta h_{r'} - 2G_{rs} \Delta h_r \Delta c_s \\ & + E_s (\Delta c_s)^2 - 2F_{ss'} \Delta c_s \Delta c_{s'} + \&\text{c.} \ldots \}, \\ & \times \text{ expt. (terms in cubic and higher orders of the } \Delta\text{'s)} \quad \ldots \quad \text{(vi.)}. \end{aligned}$$

This represents the probability of the observed unit, *i.e.*, the individuals (x_1, x_2, ... x_m, for all sets), occurring, on the assumption that errors Δh_1, Δh_2, ...

Δh_m, Δc_1, Δc_2, ... Δc_p, have been made in the determination of the frequency constants. In other words, we have here the frequency distribution for errors in the values of the frequency constants.

(3) *Conclusions to be drawn from the form of* (vi.).

(α) The distribution of the errors of frequency constants, if treated exactly, will generally be skew, for the cubic and higher terms in the Δ's do not vanish. If, however, the cubic terms are small as compared with the square terms, the frequency distribution of errors will approximate closely to a normal correlation surface.

It would be impossible to evaluate the remainder after the second power terms in TAYLOR's series for any general expression $f(x_1, x_2, \ldots x_m; c_1, c_2, \ldots c_p)$ for frequency. In special cases we have found that terms of the third order amount in the most unfavourable circumstances to 4 per cent. of the terms of the second order, generally to a good deal less.* Probably the series in most cases converges with considerable rapidity. The fact, however, that we are dealing with the first terms of a series should be borne in mind. It does not seem to have been sufficiently emphasised when the probable error of the standard deviation is taken to be $67{\cdot}449/\sqrt{2n}$ per cent. of the standard deviation. The usual proof of this result, however, involves the same assumption as to the smallness of the cubic terms.

(β) Supposing the errors so small that we may neglect the cubic terms, we conclude that the errors made in calculating the constants of any frequency distribution are—

(i.) Themselves distributed according to the normal law of errors,

(ii.) Correlated among themselves.

Both these conclusions are of the utmost importance. The first enables us to obtain the probable errors of the frequency constants; the second depends upon the fact that $C_{rr'}$, G_{rs}, and $F_{ss'}$ are in general not zero. The standard deviations of, and the correlations between, the frequency constant errors can now be calculated by the ordinary theory of normal correlation.

Before, however, proceeding to these calculations, we may draw one or two other conclusions of considerable generality and wide significance.

(γ) Consider a race fully defined by the variations σ_1, σ_2, σ_3, ..., &c. of the organs of its members and their correlations r_{12}, r_{23}, r_{13}, ... Now let a *random* small selection be made of this race, defined by

$$\sigma_1 + \delta\sigma_1,\ \sigma_2 + \delta\sigma_2,\ \sigma_3 + \delta\sigma_3,\ \ldots\ r_{12} + \delta r_{12},\ r_{23} + \delta r_{23},\ r_{13} + \delta r_{13},\ \ldots,$$

where the magnitudes of $\delta\sigma_1$, $\delta\sigma_2$, $\delta\sigma_3$, ... δr_{12}, δr_{23}, δr_{13}, ..., are quantities depending

* See, for the relative order of two terms of the second and third orders, "Regression, Heredity, and Panmixia," 'Phil. Trans.,' A, vol. 187, p. 266.

on the magnitude of the probable errors of the variation constants, and therefore on the size of the selection. Then the system $\delta\sigma_1$, $\delta\sigma_2$, $\delta\sigma_3$, . . . δr_{12}, δr_{23}, δr_{13}, . . ., is not a system of independent variations, but the changes in variation and correlation are correlated together, since the terms $F_{ss'}$ do not generally vanish. We therefore conclude that if a random selection be made out of a population with regard to one organ σ_1, there will be tendencies for the variation of all other organs and the correlation between all organs to change also in certain directions, which can be definitely indicated so soon as the general population has been measured, and the effect of the random selection on one organ has been ascertained. What is proved here of random selection will be shown to be true still more intensively for artificial selection,* *i.e.*, every selection of one organ modifies in a correlated manner the variation and correlation of all other organs. It is impossible to alter one organ without altering all other organs and their relation to each other.

(δ) The remarks in (γ) are not only true for a random selection of variation, but equally well apply to selection of size. This follows because the terms $C_{rr'}$ and G_{rs} do not as a rule vanish. If a random selection of 100 or 1,000 individuals be made out of a general population, then the mean size of any organ in this sub-group will probably differ from that of the general population. The result is that the size of all other organs, their amounts of variation and correlation, will probably have values differing in definite directions from those of the general population.

(ϵ) Take two random selections out of a general population; the probability is that they will have different means for any one organ, and a result will be that they will have correlated systems of changes in the sizes, variations, and correlations of all other organs. In other words, random selection produces a differentiation of all characters, which differentiation will be the more marked the smaller the random selection.†

(ζ) This principle—that a random selection gives a system of correlated changes in the deviations of all the characters of a species—seems, to some extent, capable of explaining the small but systematic differences to be found occasionally between closely allied species. It is not necessary to suppose them due to a long process of natural selection acting on a variety of organs; a small random selection, or possibly a natural selection of one organ, might suffice to produce the systematic differences of character in all organs.

How far a succession of random selections would give an evolution biassed by the first random selection requires further consideration; but it seems impossible for the characters of a race to remain fixed under the influence of a *heavy* but non-selective death-rate. They will vary from year to year, although this systematic change of

* "Memoir on the Influence of Selection on Correlation." Selection not only modifies correlation, but the selection of one organ can create correlation between organs previously uncorrelated.

† Continuous artificial selection of an organ produces a still more marked differentiation of all other characters, but this is treated of in another memoir.

characters be not always in the same direction. Systematic change of characters produced by random selection may be spoken of as *random evolution*. Random evolution is theoretically a possible cause of systematic change; experiment only can determine how great is its effectiveness in differentiating local races.

(η) In the case of a normal distribution of variation defined by the mean h and the standard deviation σ, it has been usual to suppose that the error made in the mean is independent of the error made in the variation. In other words, it has been assumed that G_{rs} vanishes, although no proof has been given, or possibly it has not been realised that a proof was necessary. In this case f is of the form $\frac{n}{\sqrt{(2\pi)}\,\sigma}$ expt. $-\{x^2/2\sigma^2\}$ and $\frac{d^2(\log f)}{d\sigma\, dx} = \frac{2x}{\sigma^3}$, whence

$$G = 2\int_{-\infty}^{+\infty} \frac{x}{\sigma^3} \text{ expt. } - \{x^2/2\sigma^2\} \frac{n}{\sqrt{(2\pi)}\,\sigma}\, dx = 0.$$

Thus there is no correlation between error in the mean and error in the standard deviation. This assumes that we stop at the square terms in (vi.). If, however, we include the cubic terms, &c., product terms in Δh and $\Delta\sigma$ do arise, and we cannot state straight off that no correlation exists, although it may be very small. In the case of all skew variation, such as is so frequent among plants and animals, a correlation will always be found between deviation or error in the mean and the like in the standard deviation. In other words, to alter the mean by selection (artificial or random) is to alter the variation of an organ.

With the exception of the statements in this paragraph (η), the whole of our general conclusions in this section are independent of any particular law of frequency.

(4) *On the Determination of the Probable Errors and the Error Correlations of the Frequency Constants.*

Let $\eta_1, \eta_2, \eta_3, \ldots$, be the frequency constants, whether they be the means, standard deviations, or correlations of a complex of organs. Then if we neglect cubic and higher terms in the deviations $\Delta\eta_1, \Delta\eta_2, \Delta\eta_3, \ldots$, the frequency surface giving the distribution of the variations in the deviations is

$$P_\Delta = P_0 \text{ expt. } -\tfrac{1}{2}\left[S\{a_{rr}(\Delta\eta_r)^2\} - 2S\{a_{rs}\,\Delta\eta_r\,\Delta\eta_s\}\right],$$

where

$$a_r = -\iiint \ldots f \frac{d^2(\log f)}{d\eta_{rr}^2}\, dx_1\, dx_2 \ldots dx_m,$$

$$a_{rs} = \iiint \ldots f \frac{d^2(\log f)}{d\eta_r\, d\eta_s}\, dx_1\, dx^2 \ldots dx_m \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(vii.)}.$$

It is required to find Σ_r, the standard deviation of $\Delta\eta_r$, and R_{rs}, the coefficient of correlation between $\Delta\eta_r$ and $\Delta\eta_s$.

Let Δ be the discriminant :

$$\begin{vmatrix} a_{11}, & -a_{12}, & -a_{13}, & \ldots \\ -a_{21}, & a_{22}, & -a_{23}, & \ldots \\ -a_{31}, & -a_{32}, & a_{33}, & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{vmatrix}$$

where $a_{rs} = a_{sr}$. Further, let A_{rs} be the minor of the r^{th} row and s^{th} column ; then

$$\Sigma_r^2 = A_{rr}/\Delta \quad \text{and} \quad R_{rs} = A_{rs}/(\Delta\Sigma_r\Sigma_s) \; . \quad . \quad . \quad . \quad . \quad . \quad \text{(viii.)}$$

give the required values of Σ_r and R_{rs}.*

Further, the standard deviation of $\Delta\eta_r$ for *selected* values of all the other $\Delta\eta$'s is $1/\sqrt{a_{rr}}$.

This value can often be of service. Thus suppose a considerable number of skeletons found, but that only in a comparatively few cases is it possible to pair together the femur and tibia of the same individuals. Then the variations of femur and tibia in the race will be known with great exactness, but the probable error of the correlation between femur and tibia will be given with close approximation by a form like $\cdot 67449 \dfrac{1}{\sqrt{a_{rr}}}$. In fact, whenever we have obtained from a large number of observations the values of the frequency constants of a race with great exactness, then, using these values to obtain an additional variation or correlation constant from a few observations, the probable error will be of the form just indicated, and not of the form $\cdot 67449\ A_{rr}/\Delta$. It is needless, perhaps, to remark that the former is far easier to calculate than the latter.

III. On the Probable Errors and the Coefficients of Correlation between Errors made in the Determination of the Constants of a Normal Frequency Distribution.

(5) In order to exhibit more clearly the method of investigation, it is desirable that a simple case be first taken. Accordingly we will start with the following problem :—

To find the Probable Errors and Error Correlations of the Constants of a Normal Frequency Distribution for Two Organs.

Let h_1, h_2, be the means, σ_1, σ_2, the standard deviations, r_{12} the coefficient of correlation, n the total number of pairs of observations, x_1, x_2, any pair of corresponding deviations of the organs. Then the frequency $z\,dx_1\,dx_2$ is given by the surface

$$z = \frac{n}{2\pi\sigma_1\sigma_2\sqrt{(1-r_{12}^2)}} \text{ expt.} - \tfrac{1}{2}\left\{\frac{x_1^2}{\sigma_1^2(1-r_{12}^2)} - \frac{2x_1x_2r_{12}}{\sigma_1\sigma_2(1-r_{12}^2)} + \frac{x_2^2}{\sigma_2^2(1-r_{12}^2)}\right\}. \quad \text{(ix.)}.$$

* See "Regression, Panmixia, and Heredity," 'Phil. Trans.,' A, vol. 187, p. 301 (*e*).

We require to find

Σ_{σ_1}, the standard deviation of errors in σ_1,
Σ_{σ_2}, ,, ,, ,, σ_2,
$\Sigma_{r_{12}}$, ,, ,, ,, r_{12},
Σ_{ρ_1}, ,, ,, ,, $\rho_1 = r_{12}\sigma_1/\sigma_2$,
which is the coefficient of regression of x_1,*
$R_{\sigma_1\sigma_2}$, the coefficient of correlation between errors in σ_1 and σ_2,
$R_{\sigma_1 r_{12}}$, ,, ,, ,, ,, σ_1 ,, r_{12},
$R_{\sigma_2 r_{12}}$, ,, ,, ,, ,, σ_2 ,, r_{12},
Σ_{h_1}, the standard deviation of errors in h_1,
Σ_{h_2}, ,, ,, ,, h_2,
$R_{h_1 h_2}$, the coefficient of correlation between errors in h_1 and h_2.

It follows that $R_{h_1\sigma_1}$, $R_{h_2\sigma_2}$, $R_{h_1\sigma_2}$, $R_{h_2\sigma_1}$, $R_{h_1 r_{12}}$, $R_{h_2 r_{12}}$, are all zero, since, by (ii.) of p. 233, G_{rs} will vanish when the distributions of x_1 and x_2 are symmetrical. These correlations would not, however, vanish for the skew frequency distributions, which are of most frequent occurrence in problems of heredity and fertility in man, &c.

The first stage in the investigation is to write down the second differentials of the logarithm of z for all quantities occurring in it.

We find

$$\frac{d^2(\log z)}{d\sigma_1^2} = \frac{1}{\sigma_1^2}\left\{1 - \frac{3x_1^2}{\sigma_1^2(1-r_{12}^2)} + \frac{2x_1x_2r_{12}}{\sigma_1\sigma_2(1-r_{12}^2)}\right\},$$

$$\frac{d^2(\log z)}{d\sigma_2^2} = \frac{1}{\sigma_2^2}\left\{1 - \frac{3x_2^2}{\sigma_2^2(1-r_{12}^2)} + \frac{2x_1x_2r_{12}}{\sigma_1\sigma_2(1-r_{12}^2)}\right\},$$

$$\frac{d^2(\log z)}{d\sigma_1\,d\sigma_2} = \frac{x_1x_2r_{12}}{\sigma_1^2\sigma_2^2(1-r_{12}^2)},$$

$$\frac{d^2(\log z)}{d\sigma_1\,dr_{12}} = \frac{1}{\sigma_1}\left\{\frac{2r_{12}x_1^2}{\sigma_1^2(1-r_{12}^2)^2} - \frac{x_1x_2(1+r_{12}^2)}{\sigma_1\sigma_2(1-r_{12}^2)^2}\right\},$$

$$\frac{d^2(\log z)}{d\sigma_2\,dr_{12}} = \frac{1}{\sigma_2}\left\{\frac{2r_{12}x_2^2}{\sigma_2^2(1-r_{12}^2)^2} - \frac{x_1x_2(1+r_{12}^2)}{\sigma_1\sigma_2(1-r_{12}^2)^2}\right\},$$

$$\frac{d^2(\log z)}{dr_{12}^2} = \frac{1+r_{12}^2}{(1-r_{12}^2)^2} - \frac{1+3r_{12}^2}{(1-r_{12}^2)^3}\left(\frac{x_1^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2}\right) + \frac{6r_{12}+2r_{12}^3}{(1-r_{12}^2)^3}\frac{x_1x_2}{\sigma_1\sigma_2},$$

$$\frac{d^2(\log z)}{dx_1^2} = -\frac{1}{\sigma_1^2(1-r_{12}^2)},$$

$$\frac{d^2(\log z)}{dx_2^2} = -\frac{1}{\sigma_2^2(1-r_{12}^2)},$$

$$\frac{d^2(\log z)}{dx_1\,dx_2} = \frac{r_{12}}{\sigma_1\sigma_2(1-r_{12}^2)}.$$

* See the memoir on "Heredity, Regression, and Panmixia," p. 268.

Now these must be multiplied by z and integrated for x_1 and x_2 from $-\infty$ to $+\infty$. These integrations follow at once, if we remember that

$$n\sigma_1^2 = \iint zx_1^2\,dx_1\,dx_2, \qquad n\sigma_2^2 = \iint zx_2^2\,dx_1\,dx_2, \qquad n\sigma_1\sigma_2 r_{12} = \iint zx_1x_2\,dx_1\,dx_2,$$

by definition of σ_1, σ_2, and r_{12}.

Thus we obtain the following system :

$$\begin{aligned}
a_{11} &= \frac{n}{\sigma_1^2}\,\frac{2 - r_{12}^2}{1 - r_{12}^2}, & a_{22} &= \frac{n}{\sigma_2^2}\,\frac{2 - r_{12}^2}{1 - r_{12}^2}, \\
a_{12} &= \frac{nr_{12}^2}{\sigma_1\sigma_2(1 - r_{12}^2)}, & a_{13} &= \frac{nr_{12}}{\sigma_1(1 - r_{12}^2)}, \\
a_{23} &= \frac{nr_{12}}{\sigma_2(1 - r_{12}^2)}, & a_{33} &= \frac{n(1 + r_{12}^2)}{(1 - r_{12}^2)}, \\
a_{44} &= \frac{n}{\sigma_1^2(1 - r_{12}^2)}, & a_{55} &= \frac{n}{\sigma_2^2(1 - r_{12}^2)}, & a_{45} &= \frac{nr_{12}}{\sigma_1\sigma_2(1 - r_{12}^2)} \quad . \quad . \quad \text{(x.)},
\end{aligned}$$

where the a's are those of Eqn. (vii.), p. 236, obtained by taking the above differentials of the logarithms of z in order.

We can now write down the correlation surface, giving the frequencies of errors in the constants:

$$\begin{aligned}
\mathrm{P}_\Delta = \mathrm{P}_0 \text{ expt.} &- \tfrac{1}{2}\left\{\frac{n(\Delta h_1)^2}{\sigma_1^2(1 - r_{12}^2)} - \frac{2nr_{12}\Delta h_1 \Delta h_2}{\sigma_1\sigma_2(1 - r_{12}^2)} + \frac{n(\Delta h_2)^2}{\sigma_2^2(1 - r_{12}^2)}\right\} \\
\times \text{ expt.} &- \tfrac{1}{2}\left\{\frac{n}{\sigma_1^2}\,\frac{(2 - r_{12}^2)}{1 - r_{12}^2}(\Delta\sigma_1)^2 + \frac{n}{\sigma_2^2}\,\frac{2 - r_{12}^2}{1 - r_{12}^2}(\Delta\sigma_2)^2 + \frac{n(1 + r_{12}^2)}{(1 - r_{12}^2)^2}(\Delta r_{12})^2 \right. \\
&\left. - \frac{2nr_{12}}{\sigma_2(1 - r_{12}^2)}\Delta\sigma_2\,\Delta r_{12} - \frac{2nr_{12}}{\sigma_1(1 - r_{12}^2)}\Delta\sigma_1\,\Delta r_{12} - \frac{2nr_{12}^2}{\sigma_1\sigma_2(1 - r_{12}^2)}\Delta\sigma_1\,\Delta\sigma_2\right\}. \quad \text{(xi.)}.
\end{aligned}$$

Now several important conclusions follow at once from this result :

(α) For the case of normal frequency (but in general only for this case) the errors in the means are uncorrelated with the errors in the variations and correlations. The error correlation surface breaks up into two parts, of which the first part we have written down involves only the means, and would coincide exactly with the correlation surface (ix.), with which we started, if we write in (ix.) Δh_1 for x_1, Δh_2 for x_2, and $\sigma_1/\sqrt{n}$, $\sigma_2/\sqrt{n}$, for σ_1 and σ_2 respectively.

It follows accordingly that the standard deviations of the errors in the means are

$$\Sigma_{h_1} = \sigma_1/\sqrt{n}, \qquad \Sigma_{h_2} = \sigma_2/\sqrt{n} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xii.)},$$

and the correlation of the errors made in the means is

$$\mathrm{R}_{h_1h_2} = r_{12} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xiii.)}.$$

Further, the standard deviations for errors in h_1 when the error in h_2 is known, or for errors in h_2 when the error in h_1 is known, are respectively

$$\frac{\sigma_1}{\sqrt{n}}(1 - r_{12}^2), \qquad \frac{\sigma_2}{\sqrt{n}}(1 - r_{12}^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xiv.)}.$$

The results (xii.) are, of course, well known; the results (xiii.) and (xiv.) are, we believe, novel and important. An illustration may be of service.

Suppose the stature and arm-length of a population to be under consideration. Let us suppose the mean stature of the population known from a great number of observations. Now let the arm-lengths be determined for a random selection of the population; then, if the stature of these individuals so selected differs Δh_1 in excess from the mean stature of the whole population, the arm-length of the random selection will most probably differ from that of the whole population by Δh_2, a quantity fixed by (xiii.), or rather by the coefficient of regression

$$R_{\sigma_1\sigma_2}\Sigma_{h_2}/\Sigma_{h_1}, \text{ i.e., } r_{12}\sigma_2/\sigma_1.$$

Thus

$$\Delta h_2 = r_{12}\sigma_2/\sigma_1 . \Delta h_1,$$

with a probable error of $\cdot 67449 \times \frac{\sigma_2}{\sqrt{n}}(1 - r_{12}^2)$.

In other words, if the arm-length is to be found from a selection of the general population only, and the stature of this selection differs from that of the general population, it is most reasonable to take the arm-length of the general population to be the mean arm-length of the selected population less the quantity $r_{12}\sigma_2/\sigma_1 . \Delta h_1$.

Or, again, if a selection from a general population show a mean organ Δh_1 in excess of that of the general population, the whole system of correlated organs will exhibit changes of which the magnitudes are most probably given by the type $\frac{r_{12}\sigma_{12}}{\sigma_1}\Delta h_1$.

The bearing of this on what we have termed random evolution will be obvious.

(β) Turning to the second part of the error correlation surface, we note at once that, *if two organs be correlated, random selection will give a system of correlated deviations in their variations and their correlation.*

Random selection (and *a fortiori* it may be added, artificial or natural selection), which alters an organ's variability, alters the variability and correlation of all other organs. In fact, when it is once realised how two random selections from a general population will as a rule have organs of different means, of different variabilities, and of different correlations, the means among themselves and the variabilities and correlations among themselves forming *systematic* groups, it becomes obvious how any assumption of the coefficient of correlation as a constant for local races runs wide of the mark; and this, whether natural or random evolution, is to be looked upon as the source of the observed differences in character. What chiefly concerns the biologist in this matter at present is this, that *even* a random selection of one organ

will produce changes in all other organic characters, which, if small, are still sensible and capable of quantitative expression.*

(6) Returning now to the algebra of our investigation, we have to discuss the second part of (xi.) by aid of the formulæ given in (viii.), p. 237.

We require, in the first place, to evaluate the determinant Δ.

Now

$$\Delta = \begin{vmatrix} \dfrac{n}{\sigma_1^2}\dfrac{2-r_{12}^2}{1-r_{12}^2}, & -\dfrac{r_{12}^2}{\sigma_1\sigma_2(1-r_{12}^2)}, & -\dfrac{nr_{12}}{\sigma_1(1-r_{12}^2)} \\ -\dfrac{nr_{12}^2}{\sigma_1\sigma_2(1-r_{12}^2)}, & \dfrac{n}{\sigma_2^2}\dfrac{2-r_{12}^2}{1-r_{12}^2}, & -\dfrac{nr_{12}}{\sigma_2(1-r_{12}^2)} \\ -\dfrac{nr_{12}}{\sigma_1(1-r_{12}^2)}, & -\dfrac{nr_{12}}{\sigma_2(1-r_{12}^2)}, & \dfrac{n(1+r_{12}^2)}{(1-r_{12}^2)^2} \end{vmatrix}.$$

Divide the first row by $\dfrac{n}{\sigma_1(1-r_{12}^2)}$, the second row by $\dfrac{n}{\sigma_2(1-r_{12}^2)}$, the third row by $\dfrac{n}{1-r_{12}^2}$, the first column by σ_1, and the second by σ_2. Hence

$$\Delta = \frac{n^3}{\sigma_1^2\sigma_2^2(1-r_{12}^2)^3}\begin{vmatrix} 2-r_{12}^2 & -r_{12}^2 & -r_{12} \\ -r_{12}^2 & 2-r_{12}^2 & -r_{12} \\ -r_{12} & -r_{12} & \dfrac{1+r_{12}^2}{1-r_{12}^2} \end{vmatrix}.$$

Subtract the second column from the first, and then add the first row to the second; we have

$$\Delta = \frac{n^3}{\sigma_1^2\sigma_2^2(1-r_{12}^2)^3}\begin{vmatrix} 2 & -r_{12}^2 & -r_{12} \\ 0 & 2(1-r_{12}^2) & -2r_{12} \\ 0 & -r_{12} & \dfrac{1+r_{12}^2}{1-r_1^2} \end{vmatrix},$$

$$= \frac{4n^3}{\sigma_1^2\sigma_2^2(1-r_{12}^2)^3}.$$

The minors are now easily found to be

$$A_{11} = \frac{2n^2}{\sigma_2^2(1-r_{12}^2)^3}, \qquad A_{22} = \frac{2n^2}{\sigma_1^2(1-r_{12}^2)^3},$$

$$A_{33} = \frac{4n}{\sigma_1^2\sigma_2^2(1-r_{12}^2)}, \qquad A_{12} = \frac{2n^2r_{12}^2}{\sigma_1\sigma_2(1-r_{12}^2)^3},$$

$$A_{13} = \frac{2n^2r_{12}}{\sigma_1\sigma_2^2(1-r_{12}^2)^2}, \qquad A_{23} = \frac{2n^2r_{12}}{\sigma_2\sigma_1^2(1-r_{12}^2)^2}.$$

* Take (xvii.) below, for example; it expresses for the first time quantitatively the important biological principle that, if a group be selected at random from the general population, and it has more variability in one character, it will be more variable than the general population in all other characters.

From these expressions the values of the standard deviations and error correlations are at once found by (viii.).

We have

$$\Sigma_{\sigma_1} = \frac{\sigma_1}{\sqrt{2n}}, \quad \Sigma_{\sigma_2} = \frac{\sigma_2}{\sqrt{2n}} \quad \text{(xv.)},$$

$$\Sigma_{r_{12}} = \frac{1 - r_{12}^2}{\sqrt{n}} \quad \text{(xvi.)},$$

$$R_{\sigma_1\sigma_2} = r_{12}^2 \quad \text{(xvii.)},$$

$$R_{\sigma_1 r_{12}} = \frac{r_{12}}{\sqrt{2}}, \quad R_{\sigma_2 r_{12}} = \frac{r_{12}}{\sqrt{2}} \quad \text{(xviii.)}.$$

The result (xv.) is, of course, old; the results (xvi.) to (xviii.) are novel, and lead to interesting conclusions, which are considered in the following paragraphs.

(α) The probable error of a coefficient of correlation r_{12} is $\cdot67449\,(1 - r_{12}^2)/\sqrt{n}$. If, therefore, the correlation between two organs is less than once to twice $\cdot67449/\sqrt{n}$, they cannot be safely assumed to be correlated at all.

(β) If we know definitely the errors made in σ_1 and σ_2—if, for example, we know the variations accurately—then the probable error of r_{12} is that of an array of r_{12}'s for definite $\Delta\sigma_1$ and $\Delta\sigma_2$. It is given at once by the coefficient of $(\Delta r_{12})^2$ in (xi.) as

$$\cdot67449\,(1 - r_{12}^2)/\sqrt{\{n\,(1 + r_{12}^2)\}} \quad \text{(xix.)}.$$

This is the value given for the probable error of r_{12} by one of the present authors in a former paper.* At that time he had not fully realised the importance of the principle of the correlation of errors made in determining the magnitude of frequency constants.

The following table will enable the reader to appreciate the difference in magnitude between the "absolute" and "partial" probable errors of r_{12}:

r_{12}	$1 - r_{12}^2$	$(1 - r_{12}^2)/\sqrt{(1 + r_{12}^2)}$
0	1	1
·1	·99	·985
·2	·96	·94
·3	·91	·87
·4	·84	·78
·5	·75	·67
·6	·64	·55
·7	·51	·42
·8	·36	·28
·9	·19	·14
1	0	0

* "Heredity, Regression, and Panmixia," 'Phil. Trans.,' A, vol. 187, p. 266.

It will thus be seen that when r_{12} is small the absolute and partial probable errors are both large and nearly equal; that when r_{12} is large the absolute and partial probable errors differ more widely, but, as both are small in this case, their difference is not of much importance. Bearing these facts in mind, it will be found that the reasoning based on the partial error in previous memoirs remains valid, even if the partial error be replaced, as it generally should be, by the somewhat larger absolute error.*

(γ) If we know definitely the variability of any organ, and we take a definite group of the general population to find the variability of a second correlated organ, then there will be correlation between the deviation of the variation of this group with regard to the first organ from the variation of the first organ in the general population and the deviation of the variation of the group with regard to its second organ from that of the general population's variation for the second organ. This correlation is measured by r_{12}^2, and thus, if the organs be slightly correlated, it is small; but if the organs be closely correlated, it is large. Suppose, for example, we know the variability of the tibia, and require to find that of the ulna from a comparatively few specimens. Let $\sigma_1 + \Delta\sigma_1$ be the variability of the tibia in the specimens for which the ulna can be measured, and r_{12} the correlation observed between the ulna and tibia in these specimens; then, the variability of the ulna being observed as $\sigma_2' = \sigma_2 + \Delta\sigma_2$, the most probable variability of the ulna in the general population is

$$\sigma_2' - r_{12}^2 \Sigma_{\sigma_2} \Delta\sigma_1 / \Sigma_{\sigma_1}$$

or

$$\sigma_2' - \frac{r_{12}^2 \sigma_2}{\sigma_1} \Delta\sigma_1,$$

or, since the second term is small, we may write σ_2' for σ_2, and the above expression equals

$$= \sigma_2' \left(1 - r_{12}^2 \frac{\Delta\sigma_1}{\sigma_1}\right).$$

For the long bones, $r_{12} = \cdot 9$ roughly, and therefore we have the ratio of variability of the ulna in the general population to the variability observed in the group $= 1 - \cdot 8\, \Delta\sigma_1/\sigma_1$.

It is clear that this expression also measures the change in the variability of the ulna due to a random selection of tibia.

(δ) Although the correlation between deviations in the variability of two organs from their mean variabilities only varies as the square of their correlation, the correlation between the deviations in the variability of an organ and in its correlation with a second organ varies as the first power of the correlation of the two organs. In other words,

* See, for example, the reasoning as to the non-constancy of the correlation coefficient for local races in 'Phil. Trans.,' A, vol. 187, pp. 267 and 278.

while a selection of variability may produce only a small or moderate change on the variability of correlated organs, a selection of correlation or a selection of variability is likely to produce considerable changes on variability and correlation respectively. Let σ_1, σ_2, r_{12}, be the mean values of the standard deviations and the coefficient of correlation for any three organs; let $\sigma_1 + \Delta\sigma_1$, $\sigma_2 + \Delta\sigma_2$, $r_{12} + \Delta r_{12}$, be the like quantities for a group selected at random. Then the principle of regression tells us that most probably

$$\Delta\sigma_1 = \mathrm{R}_{\sigma_1 r_{12}} \frac{\Sigma_{\sigma_1}}{\Sigma_{r_{12}}} \Delta r_{12},$$

$$\Delta r_{12} = \mathrm{R}_{\sigma_1 r_{12}} \frac{\Sigma_{r_{12}}}{\Sigma_{\sigma_1}} \Delta\sigma_1.$$

Substituting the values given by (xv.) to (xviii.), we find

$$\left.\begin{aligned} \Delta\sigma_1 &= \tfrac{1}{2}\,\sigma_1 \frac{r_{12}}{1 - r_{12}^2} \Delta r_{12} \\ \Delta r_{12} &= r_{12}(1 - r_{12}^2)\frac{\Delta\sigma_1}{\sigma_1} \end{aligned}\right\} \quad \text{(xx.).}$$

Now these equations lead us to some important conclusions. In the first place, if the correlation be very small or very large, then a random selection of variability ($\Delta\sigma_1$) makes only a small change in correlation (Δr_{12}). The change in correlation for a selection of variability is greatest when $r_{12} = 1/\sqrt{3}$, and then is approximately $\cdot 385\, \Delta\sigma_1/\sigma_1$, or over 6 per cent., if $\Delta\sigma_1/\sigma_1$ were as high as $\frac{1}{10}$. On the other hand, the change in variability ($\Delta\sigma_1$) due to a selection (Δr_{12}) of correlation is small if the correlation be small, but increases rapidly if the correlation become nearly perfect. Of course, for perfect correlation the probable error of r_{12} is zero, and accordingly it is infinitely improbable that a selection can be made with Δr_{12} differing from zero. But if r_{12} be not unity, then a selection in which Δr_{12} is large, however improbable, will give very large changes in the variability, if r_{12} be very large. Our conclusion is accordingly that considerable changes in variability are likely to be produced whenever there is a correlation selection among highly correlated organs.

(7) *To find the Probable Error of the Regression Coefficient for Two Organs.*

The regression coefficient ρ_1 is given by

$$\rho_1 = r_{12}\sigma_1/\sigma_2.$$

Its standard deviation Σ_{ρ_1} is given by the summation equation

$$(\Sigma_{\rho_1})^2 = \mathrm{S}\,(\Delta\rho_1)^2/n.$$

To find its value we adopt a method, which we give on this first occasion at length, as it will be frequently used in the sequel.

Take logarithmic differentials

$$\frac{\Delta\rho_1}{\rho_1} = \frac{\Delta r_{12}}{r_{12}} + \frac{\Delta\sigma_1}{\sigma_1} - \frac{\Delta\sigma_2}{\sigma_2}.$$

Square and divide by n after summing

$$\frac{S(\Delta\rho_1)^2}{n\rho_1^2} = \frac{S(\Delta r_{12})^2}{nr_{12}^2} + \frac{S(\Delta\sigma_1)^2}{n\sigma_1^2} + \frac{S(\Delta\sigma_2)^2}{n\sigma_2^2} + \frac{2S(\Delta r_{12}\,\Delta\sigma_1)}{nr_{12}\sigma_1} - \frac{2S(\Delta r_{12}\,\Delta\sigma_2)}{nr_{12}\sigma_2} - \frac{2S(\Delta\sigma_1\,\Delta\sigma_2)}{n\sigma_1\sigma_2}.$$

Now, remembering the definitions of standard deviations and coefficients of correlation, this may be written

$$\frac{(\Sigma_{\rho_1})^2}{\rho_1^2} = \frac{(\Sigma_{r_{12}})^2}{r_{12}^2} + \frac{(\Sigma_{\sigma_1})^2}{\sigma_1^2} + \frac{(\Sigma_{\sigma_1})^2}{\sigma_2^2} - \frac{2\Sigma_{\sigma_1}\Sigma_{\sigma_2}R_{\sigma_1\sigma_2}}{\sigma_1\sigma_2} + \frac{2\Sigma_{r_{12}}\Sigma_{\sigma_1}R_{r_{12}\sigma_1}}{r_{12}\sigma_1} - \frac{2\Sigma_{r_{12}}\Sigma_{\sigma_2}R_{r_{12}\sigma_2}}{r_{12}\sigma_2}.$$

Now all the quantities on the right have already been found in equations (xv.) to (xviii.). Hence, substituting, we have

$$\frac{(\Sigma_{\sigma_1})^2}{\rho_1^2} = \frac{1 - r_{12}^2}{nr_{12}^2}.$$

Hence

$$\Sigma_{\rho_1} = \frac{\sigma_1}{\sigma_2}\sqrt{\left(\frac{1 - r_{12}^2}{n}\right)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xxi.}).$$

Thus the probable error of a regression coefficient

$$= \cdot 67449\,\frac{\sigma_1}{\sigma_2}\sqrt{\left(\frac{1 - r_{12}^2}{n}\right)}.$$

This is of fundamental importance for testing the significance of results obtained by applying the theory of regression to problems in heredity, panmixia, &c.

The probable percentage error in a regression coefficient $= \dfrac{67\cdot 449}{\sqrt{n}}\,\dfrac{\sqrt{(1 - r_{12}^2)}}{r_{12}}$, and hence is small if the correlation be close, and increases rapidly if the correlation be small. This again illustrates the point to which reference has been made in another memoir,* namely, that when only a few individuals can be measured, the most reliable results for the purposes of the quantitative theory of evolution are to be found from the measurements of the most highly correlated organs.

* PEARSON and LEE: "Correlation in Civilised and Uncivilised Races," 'Roy. Soc. Proc.,' vol. 61, p. 345.

Attention should be drawn to the fact that we have replaced errors by differentials. This is only legitimate so long as product terms in the errors are negligible as compared with linear terms. This is the assumption almost universally made by writers on the theory of errors.* It will not lead us astray, so long as we take care in any practical applications to verify the smallness of Δr_{12}, $\Delta\sigma_1$, $\Delta\sigma_2$, as compared with r_{12}, σ_1, and σ_2 respectively.

(8) *To find the Probable Errors and Error Correlations of the Constants of a Normal Frequency Distribution for Three Organs or more.*

It will scarcely have escaped the attentive reader that our investigation hitherto, only involving two organs, has left several important problems untouched. For example, it has dealt only with the *direct* effect of random selection. But we may ask such a question as this: What is the change in the correlation of two organs when the variability of a third is randomly selected? Or again: What is the change in the correlation of two organs when the correlation between one of these and a third, or between a third and a fourth, is randomly selected? All these are important problems in the theory of evolution.

The general equation to a normal frequency-surface for m organs is :—

$$z = \frac{n}{(2\pi)^{\frac{1}{2}m}\sigma_1\sigma_2\ldots\sigma_m\sqrt{R}}\,\text{expt.} -\tfrac{1}{2}\frac{1}{R}\left\{R_{11}\frac{x_1^2}{\sigma_1^2} + R_{22}\frac{x_2^2}{\sigma_2^2} + \ldots + 2R_{12}\frac{x_1x_2}{\sigma_1\sigma_2} + \ldots\right\}$$

where R is the determinant

$$\begin{vmatrix} 1 & r_{12} & r_{13} & \ldots & r_{1m} \\ r_{21} & 1 & r_{23} & \ldots & r_{2m} \\ r_{31} & r_{32} & 1 & \ldots & r_{3m} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{m1} & r_{m2} & \ldots & \ldots & 1 \end{vmatrix}$$

and $R_{ss'}$ is the minor of the term in the s^{th} row and s'^{th} column.

We require first to find the quantities like (vii.) of our Art. (4).

$$\log z = \log\left(n/(2\pi)^{\frac{1}{2}m}\right) - \tfrac{1}{2}\log R - S_s(\log\sigma_s) - \tfrac{1}{2}S_s\left(\frac{R_{ss}}{R}\,\frac{x_s^2}{\sigma_s^2}\right) - S_{ss'}\left(\frac{R_{ss'}}{R}\,\frac{x_sx_{s'}}{\sigma_s\sigma_{s'}}\right),$$

$$\frac{d(\log z)}{d\sigma_1} = -\frac{1}{\sigma_1}\left\{1 - \frac{R_{11}x_1^2}{R\sigma_1^2} - S_s\left(\frac{R_{1s}}{R}\,\frac{x_1x_s}{\sigma_1\sigma_s}\right)\right\} \quad \ldots\ldots \quad \text{(xxii.)},$$

* GAUSS, admittedly, 'Theoria Combinationis Observationum' .... p. 53, Problema; LAPLACE and POISSON, actually but obscurely; see 'Théorie analytique des Probabilités, Liv. II., chap. IV., and 'Recherches sur la Probabilité des Jugements,' chap. IV.; more clearly in TODHUNTER'S account, 'History of Theory of Probability,' Art. 1,002 *et seq.* Further, CROFTON, Article 'Probability,' § 48, for a like assumption.

$$\frac{d(\log z)}{dr_{12}} = -\tfrac{1}{2}\frac{d(\log \mathrm{R})}{dr_{12}} - \tfrac{1}{2}\frac{d}{dr_{12}}\mathrm{S}_s\left(\frac{\mathrm{R}_{1s}x_s^2}{\mathrm{R}\sigma_s^2}\right) - \frac{d}{dr_{12}}\mathrm{S}_{ss'}\left(\frac{\mathrm{R}_{ss'}x_s x_{s'}}{\mathrm{R}\sigma_s\sigma_{s'}}\right) \quad . \quad \text{(xxiii.)}.$$

Differentiating the first of these again with regard to σ_1, and summing for all possible values of x's, we find

$$-a_{11} = \iiint \ldots z\frac{d^2(\log z)}{d\sigma_1^2}\,dx_1\,dx_2\ldots dx_m = \frac{n}{\sigma_1^2}\left(1 - \frac{3\mathrm{R}_{11}}{\mathrm{R}} - \frac{2\mathrm{S}_s(\mathrm{R}_{1s}r_{1s})}{\mathrm{R}}\right).$$

But

$$\mathrm{R} = \mathrm{R}_{11} + \mathrm{S}_s(\mathrm{R}_{1s}r_{1s}).$$

Hence

$$a_{11} = \frac{n}{\sigma_1^2}\left(\frac{\mathrm{R}+\mathrm{R}_{11}}{\mathrm{R}}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxiv.)}.$$

Differentiating (xxii.) with regard to σ_2 and summing, we have at once

$$\iiint \ldots z\frac{d^2(\log z)}{d\sigma_1\,d\sigma_2}\,dx_1\,dx_2\ldots dx_m = a_{12} = -\frac{n}{\sigma_1\sigma_2}\frac{r_{12}\mathrm{R}_{12}}{\mathrm{R}} \quad . \quad . \quad . \quad \text{(xxv.)}.$$

Differentiating (xxii.) with regard to r_{12} and summing, we have

$$\iiint \ldots z\frac{d^2(\log z)}{d\sigma_1\,dr_{12}}\,dx_1\,dx_2\ldots dx_m = {}_1c_{12} = \frac{n}{\sigma_1}\left\{\frac{d}{dr_{12}}\left(\frac{\mathrm{R}_{11}}{\mathrm{R}}\right) + \mathrm{S}_s r_{1s}\frac{d}{dr_{12}}\left(\frac{\mathrm{R}_{1s}}{\mathrm{R}}\right)\right\},$$

$$= \frac{n}{\sigma_1}\left\{\frac{d}{dr_{12}}\left(\frac{\mathrm{R}_{11}+\mathrm{S}_s(r_{1s}\mathrm{R}_{1s})}{\mathrm{R}}\right) - \frac{\mathrm{R}_{12}}{\mathrm{R}}\right\},$$

or

$${}_1c_{12} = -n\mathrm{R}_{12}/\sigma_1\mathrm{R} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxvi.)}.$$

Differentiating (xxii.) with regard to r_{23} and summing, we find

$${}_1c_{23} = \frac{n}{\sigma_1}\frac{d}{dr_{23}}\frac{\mathrm{R}_{11}}{\mathrm{R}} + \frac{n}{\sigma_1}\frac{d}{dr_{23}}\mathrm{S}_s\left(\frac{\mathrm{R}_{1s}r_{1s}}{\mathrm{R}}\right),$$

$$= \frac{n}{\sigma_1}\frac{d}{dr_{23}}\left(\frac{\mathrm{R}_{11}+\mathrm{S}_s(\mathrm{R}_{1s}r_{1s})}{\mathrm{R}}\right).$$

or

$${}_1c_{23} = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxvii.)}.$$

Our next step is to differentiate (xxiii.) with regard to r_{12} and sum. We have

$$\iiint \ldots z \frac{d^2(\log z)}{dr_{12}^2} dx_1 \ldots dx_m = - {}_{12}b_{12},$$

$$= -\frac{n}{2}\frac{d^2(\log \mathrm{R})}{dr_{12}^2} - \frac{n}{2}\frac{d^2}{dr_{12}^2}\mathrm{S}_{ss}\left(\frac{\mathrm{R}_{ss}}{\mathrm{R}}\right) - n\mathrm{S}_{ss'}\left(r_{ss'}\frac{d^2}{dr_{12}^2}\left(\frac{\mathrm{R}_{ss'}}{\mathrm{R}}\right)\right),$$

$$= -\frac{n}{2}\frac{d^2(\log \mathrm{R})}{dr_{12}^2} - \frac{n}{2}\frac{d^2}{dr_{12}^2}\mathrm{S}_{ss}\left(\frac{\mathrm{R}_{ss}}{\mathrm{R}}\right) + 2n\frac{d}{dr_{12}}\left(\frac{\mathrm{R}_{12}}{\mathrm{R}}\right) - n\mathrm{S}_{ss'}\left(\frac{d^2}{dr_{12}^2}\left(\frac{r_{ss'}\mathrm{R}_{ss'}}{\mathrm{R}}\right)\right).$$

Now, since $\mathrm{R} = \mathrm{R}_{11} + \mathrm{S}_s(\mathrm{R}_{1s}r_{1s})$,

$$m\mathrm{R} = \mathrm{S}_{ss}(\mathrm{R}_{ss}) + 2\mathrm{S}_{ss'}(\mathrm{R}_{ss'}r_{ss'}),$$

and therefore

$$\mathrm{S}_{ss'}(\mathrm{R}_{ss'}r_{ss'}/\mathrm{R}) = \tfrac{1}{2}m - \tfrac{1}{2}\mathrm{S}_{ss}(\mathrm{R}_{ss}/\mathrm{R}).$$

Substituting, we find

$$- {}_{12}b_{12} = -\frac{n}{2}\frac{d^2(\log \mathrm{R})}{dr_{12}^2} + 2n\frac{d}{dr_{12}}\left(\frac{\mathrm{R}_{12}}{\mathrm{R}}\right).$$

But $\dfrac{d}{dr_{12}}(\log \mathrm{R}) = \dfrac{1}{\mathrm{R}}\dfrac{d\mathrm{R}}{dr_{12}}$ and $2\mathrm{R}_{12} = \dfrac{d\mathrm{R}}{dr_{12}}$; hence

$$\frac{d^2(\log \mathrm{R})}{dr_{12}^2} = 2\frac{d(\mathrm{R}_{12}/\mathrm{R})}{dr_{12}}.$$

Thus finally

$${}_{12}b_{12} = -n\frac{d}{dr_{12}}\left(\frac{\mathrm{R}_{12}}{\mathrm{R}}\right) = n\frac{2\mathrm{R}_{12}^2 - \mathrm{R}\mathrm{R}_{12,12}}{\mathrm{R}^2} \quad \ldots\ldots\ldots \quad \text{(xxviii.)},$$

where $\mathrm{R}_{12,12}$ is the second minor, found by striking out from R the first and second rows and columns.

In the next place let us differentiate (xxiii.) with regard to r_{13} and sum. We find in precisely similar manner

$$\iiint \ldots z\frac{d^2 \log z}{dr_{12}d_{13}} dx_1 dx_2 \ldots dx_m = {}_{12}b_{13}$$

$$= n\frac{d}{dr_{12}}\left(\frac{\mathrm{R}_{13}}{\mathrm{R}}\right), \text{ or, } = n\frac{d}{dr_{13}}\left(\frac{\mathrm{R}_{12}}{\mathrm{R}}\right),$$

or

$${}_{12}b_{13} = -n\frac{(2\mathrm{R}_{12}\mathrm{R}_{13} - \mathrm{R}\mathrm{R}_{12,13})}{\mathrm{R}^2} \quad \ldots\ldots\ldots \quad \text{(xxix.)}.$$

Similarly, we deduce

$${}_{12}b_{34} = n\frac{d}{dr_{12}}\left(\frac{\mathrm{R}_{34}}{\mathrm{R}}\right) = -n\frac{(2\mathrm{R}_{12}\mathrm{R}_{34} - 2\mathrm{R}\mathrm{R}_{12,34})}{\mathrm{R}^2} \quad \ldots\ldots \quad \text{(xxx.)}.$$

This completes all the possible types.

We are now in a position to write down from (vii.) of Art. (4) the complete frequency surface for errors in the constants of a normal frequency surface for m organs. It will suffice to write a type of each term. We have

$$P_\Delta = P_0 \times \text{exponent} - \frac{n}{2}\left\{\frac{R + R_{11}}{R}\frac{(\Delta\sigma_1)^2}{\sigma_1^2} + \frac{2R_{12}r_{12}}{R}\frac{\Delta\sigma_1\,\Delta\sigma_2}{\sigma_1\sigma_2}\right.$$

$$+ \ldots + \frac{2R_{12}^2 - RR_{12,12}}{R^2}(\Delta r_{12})^2 + \ldots + \frac{2R_{12}}{R}\frac{\Delta\sigma_1}{\sigma_1}\Delta r_{12} + \ldots$$

$$\left. + 2\,\frac{2R_{12}R_{13} - RR_{12,13}}{R^2}\Delta r_{12}\,\Delta r_{13} + \ldots + 2\,\frac{2R_{12}R_{34} - 2RR_{12,34}}{R^2}\Delta r_{12}\,\Delta r_{34} + \ldots\right\} \quad \text{(xxxi.)}.$$

Now this result again seems at once to give conclusions of considerable importance. Thus:—

(α) Since there are no terms of the type $\Delta\sigma_1\,\Delta r_{23}$, we infer (i.) that the random selection of variation in one organ will most likely only vary the correlation between two other organs by terms of the second order; and that (ii.) the random selection of correlation between two organs will in all probability only change the variability of a third organ by terms of the second order.

(β) The selection of correlation between any two organs will most probable vary the correlation between a second pair, *i.e.*, terms exist in $\Delta r_{12}\,\Delta_{34}$, &c.

(γ) The selection of the variation for any organ varies the correlation between that organ and a third organ, and *vice versâ* the selection of correlation between two organs changes the variability of both organs. And lastly

(δ) The selection of correlation between two organs varies the correlation between either organ and any other organs.

We may exhibit these results more clearly by taking four special organs, say, femur, tibia, humerus, and radius. Then a group having the variability of its femur different from that of the general population, will also have, in all probability, the variability in its tibia, humerus, and radius different; the correlations femur-tibia, femur-humerus, and femur-radius different; but those of tibia-humerus, humerus-radius, and radius-tibia only slightly different. Further, a group having the correlation of its femur-tibia different from that of the general population, will also have all the other correlations, humerus-radius, femur-humerus, femur-radius, tibia-humerus, tibia-radius, different from the values for the general population. Further, the variability in femur and tibia will be changed; but in all likelihood the variability in humerus and radius only slightly changed.

These general conclusions, which seem to cast considerable light on the manner in which selection influences the variability and correlation of organs, must now be reduced to quantitative expression.

(9) *Evaluation of the Constants* a_{11}, a_{12}, ${}_1c_{12}$, ${}_1c_{23}$, ${}_{12}b_{12}$, ${}_{12}b_{13}$, ${}_{12}b_{34}$, &c., *of the last Article for the cases of Three or Four Correlated Organs.*

We will proceed first to calculate a general form for the determinant Δ of our Article (4).

Using the values (xxiv. to xxx.) we find for four organs

$$\Delta = \begin{vmatrix}
\frac{n(\mathrm{R}+\mathrm{R}_{11})}{\sigma_1^2\mathrm{R}}, & \frac{nr_{12}\mathrm{R}_{12}}{\sigma_1\sigma_2\mathrm{R}}, & \frac{nr_{13}\mathrm{R}_{13}}{\sigma_1\sigma_2\mathrm{R}}, & \frac{nr_{14}\mathrm{R}_{14}}{\sigma_1\sigma_4\mathrm{R}}, & \frac{n\mathrm{R}_{12}}{\sigma_1\mathrm{R}}, & \frac{n\mathrm{R}_{13}}{\sigma_1\mathrm{R}}, & \frac{n\mathrm{R}_{14}}{\sigma_1\mathrm{R}}, & 0, & 0, & 0, \\
\frac{nr_{12}\mathrm{R}_{12}}{\sigma_1\sigma_2\mathrm{R}}, & \frac{n(\mathrm{R}+\mathrm{R}_{22})}{\sigma_2^2\mathrm{R}}, & \frac{nr_{23}\mathrm{R}_{23}}{\sigma_2\sigma_3\mathrm{R}}, & \frac{nr_{24}\mathrm{R}_{24}}{\sigma_2\sigma_4\mathrm{R}}, & \frac{n\mathrm{R}_{12}}{\sigma_2\mathrm{R}}, & 0, & 0, & \frac{n\mathrm{R}_{23}}{\sigma_2\mathrm{R}}, & \frac{n\mathrm{R}_{24}}{\sigma_2\mathrm{R}}, & 0, \\
\frac{nr_{13}\mathrm{R}_{13}}{\sigma_1\sigma_3\mathrm{R}}, & \frac{nr_{23}\mathrm{R}_{23}}{\sigma_3^2\mathrm{R}}, & \frac{n(\mathrm{R}+\mathrm{R}_{33})}{\sigma_3^2\mathrm{R}}, & \frac{nr_{34}\mathrm{R}_{34}}{\sigma_3\sigma_4\mathrm{R}}, & 0, & \frac{n\mathrm{R}_{13}}{\sigma_3\mathrm{R}}, & 0, & \frac{n\mathrm{R}_{23}}{\sigma_3\mathrm{R}}, & 0, & \frac{n\mathrm{R}_{34}}{\sigma_3\mathrm{R}}, \\
\frac{nr_{14}\mathrm{R}_{14}}{\sigma_1\sigma_4\mathrm{R}}, & \frac{nr_{24}\mathrm{R}_{24}}{\sigma_2\sigma_4\mathrm{R}}, & \frac{n\mathrm{R}_{34}}{\sigma_3\sigma_4\mathrm{R}}, & \frac{n(\mathrm{R}+\mathrm{R}_{44})}{\sigma_4^2\mathrm{R}}, & 0, & 0, & \frac{n\mathrm{R}_{14}}{\sigma_4\mathrm{R}}, & 0, & \frac{n\mathrm{R}_{24}}{\sigma_4\mathrm{R}}, & \frac{n\mathrm{R}_{34}}{\sigma_4\mathrm{R}}, \\
\frac{n\mathrm{R}_{12}}{\sigma_1\mathrm{R}}, & \frac{n\mathrm{R}_{12}}{\sigma_2\mathrm{R}}, & 0, & 0, & -\frac{nd}{dr_{12}}\left(\frac{\mathrm{R}_{12}}{\mathrm{R}}\right), & -\frac{nd}{dr_{12}}\left(\frac{\mathrm{R}_{13}}{\mathrm{R}}\right), & -\frac{nd}{dr_{12}}\left(\frac{\mathrm{R}_{14}}{\mathrm{R}}\right), & -\frac{nd}{dr_{12}}\left(\frac{\mathrm{R}_{23}}{\mathrm{R}}\right), & -\frac{nd}{dr_{12}}\left(\frac{\mathrm{R}_{24}}{\mathrm{R}}\right), & -\frac{nd}{dr_{12}}\left(\frac{\mathrm{R}_{34}}{\mathrm{R}}\right), \\
\frac{n\mathrm{R}_{13}}{\sigma_1\mathrm{R}}, & 0, & \frac{n\mathrm{R}_{13}}{\sigma_3\mathrm{R}}, & 0, & -\frac{nd}{dr_{13}}\left(\frac{\mathrm{R}_{12}}{\mathrm{R}}\right), & -\frac{nd}{dr_{13}}\left(\frac{\mathrm{R}_{13}}{\mathrm{R}}\right), & -\frac{nd}{dr_{13}}\left(\frac{\mathrm{R}_{14}}{\mathrm{R}}\right), & -\frac{nd}{dr_{13}}\left(\frac{\mathrm{R}_{23}}{\mathrm{R}}\right), & -\frac{nd}{dr_{13}}\left(\frac{\mathrm{R}_{24}}{\mathrm{R}}\right), & -\frac{nd}{dr_{13}}\left(\frac{\mathrm{R}_{34}}{\mathrm{R}}\right), \\
\frac{n\mathrm{R}_{14}}{\sigma_1\mathrm{R}}, & 0, & 0, & \frac{n\mathrm{R}_{14}}{\sigma_4\mathrm{R}}, & -\frac{nd}{dr_{14}}\left(\frac{\mathrm{R}_{12}}{\mathrm{R}}\right), & -\frac{nd}{dr_{14}}\left(\frac{\mathrm{R}_{13}}{\mathrm{R}}\right), & -\frac{nd}{dr_{14}}\left(\frac{\mathrm{R}_{14}}{\mathrm{R}}\right), & -\frac{nd}{dr_{14}}\left(\frac{\mathrm{R}_{23}}{\mathrm{R}}\right), & -\frac{nd}{dr_{14}}\left(\frac{\mathrm{R}_{24}}{\mathrm{R}}\right), & -\frac{nd}{dr_{14}}\left(\frac{\mathrm{R}_{34}}{\mathrm{R}}\right), \\
0, & \frac{n\mathrm{R}_{23}}{\sigma_2\mathrm{R}}, & \frac{n\mathrm{R}_{23}}{\sigma_3\mathrm{R}}, & 0, & -\frac{nd}{dr_{23}}\left(\frac{\mathrm{R}_{12}}{\mathrm{R}}\right), & -\frac{nd}{dr_{23}}\left(\frac{\mathrm{R}_{13}}{\mathrm{R}}\right), & -\frac{nd}{dr_{23}}\left(\frac{\mathrm{R}_{14}}{\mathrm{R}}\right), & -\frac{nd}{dr_{23}}\left(\frac{\mathrm{R}_{23}}{\mathrm{R}}\right), & -\frac{nd}{dr_{23}}\left(\frac{\mathrm{R}_{24}}{\mathrm{R}}\right), & -\frac{nd}{dr_{23}}\left(\frac{\mathrm{R}_{34}}{\mathrm{R}}\right), \\
0, & \frac{n\mathrm{R}_{24}}{\sigma_2\mathrm{R}}, & 0, & \frac{n\mathrm{R}_{24}}{\sigma_4\mathrm{R}}, & -\frac{nd}{dr_{24}}\left(\frac{\mathrm{R}_{12}}{\mathrm{R}}\right), & -\frac{nd}{dr_{24}}\left(\frac{\mathrm{R}_{13}}{\mathrm{R}}\right), & -\frac{nd}{dr_{24}}\left(\frac{\mathrm{R}_{14}}{\mathrm{R}}\right), & -\frac{nd}{dr_{24}}\left(\frac{\mathrm{R}_{23}}{\mathrm{R}}\right), & -\frac{nd}{dr_{24}}\left(\frac{\mathrm{R}_{24}}{\mathrm{R}}\right), & -\frac{nd}{dr_{24}}\left(\frac{\mathrm{R}_{34}}{\mathrm{R}}\right), \\
0, & 0, & \frac{n\mathrm{R}_{34}}{\sigma_3\mathrm{R}}, & \frac{n\mathrm{R}_{34}}{\sigma_4\mathrm{R}}, & -\frac{nd}{dr_{34}}\left(\frac{\mathrm{R}_{12}}{\mathrm{R}}\right), & -\frac{nd}{dr_{34}}\left(\frac{\mathrm{R}_{13}}{\mathrm{R}}\right), & -\frac{nd}{dr_{34}}\left(\frac{\mathrm{R}_{14}}{\mathrm{R}}\right), & -\frac{nd}{dr_{34}}\left(\frac{\mathrm{R}_{23}}{\mathrm{R}}\right), & -\frac{nd}{dr_{34}}\left(\frac{\mathrm{R}_{24}}{\mathrm{R}}\right), & -\frac{nd}{dr_{34}}\left(\frac{\mathrm{R}_{34}}{\mathrm{R}}\right),
\end{vmatrix} \qquad \text{(xxxii.)}$$

Divide each row by n, the first, second, third, and fourth rows by σ_1, σ_2, σ_3, σ_4 respectively, and the first, second, third, and fourth columns by σ_1, σ_2, σ_3, σ_4 respectively. Multiply the fifth column by r_{12} and subtract it from the first and second; the sixth column by r_{13} and subtract it from the first and third; the seventh column by r_{14} and subtract it from the first and fourth; the eighth column by r_{23} and subtract it from the second and third; the ninth column by r_{24} and subtract it from the second and fourth; the tenth column by r_{34} and subtract it from the third and fourth. Remembering that

$$R = R_{11} + r_{12}R_{12} + r_{13}R_{13} + r_{14}R_{14},$$
$$R = R_{22} + r_{21}R_{12} + r_{23}R_{23} + r_{24}R_{24}, \text{ \&c.},$$

and that

$$\frac{R_{12}}{R} + r_{12}\frac{d}{dr_{12}}\left(\frac{R_{12}}{R}\right) + r_{13}\frac{d}{dr_{12}}\left(\frac{R_{13}}{R}\right) + r_{14}\frac{d}{dr_{14}}\left(\frac{R_{14}}{R}\right)$$
$$= \frac{d}{dr_{12}}\left(\frac{r_{12}R_{12} + r_{13}R_{13} + r_{14}R_{14}}{R}\right) = -\frac{d}{dr_{12}}\left(\frac{R_{11}}{R}\right),$$

$$r_{13}\frac{d}{dr_{12}}\left(\frac{R_{13}}{R}\right) + r_{23}\frac{d}{dr_{12}}\left(\frac{R_{23}}{R}\right) + r_{34}\frac{d}{dr_{12}}\left(\frac{R_{34}}{R}\right) = -\frac{d}{dr_{12}}\left(\frac{R_{33}}{R}\right), \text{ \&c.},$$

we find, writing d_{12} for d/dr_{12}, &c., for brevity,

$$\Delta = \frac{n^{10}}{\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4^2}$$

$$\times \left|\begin{matrix}
2R_{11}/R, & 0, & 0, & 0, & R_{12}/R, & R_{13}/R, & R_{14}/R, & 0, & 0, & 0, \\
0, & 2R_{22}/R, & 0, & 0, & R_{12}/R, & 0, & 0, & R_{23}/R, & R_{24}/R, & 0, \\
0, & 0, & 2R_{33}/R, & 0, & 0, & R_{13}/R, & 0, & R_{23}/R, & 0, & R_{34}/R, \\
0, & 0, & 0, & 2R_{44}/R, & 0, & 0, & R_{14}/R, & 0, & R_{24}/R, & R_{34}/R, \\
d_{12}\left(\frac{R_{11}}{R}\right), & d_{12}\left(\frac{R_{22}}{R}\right), & d_{12}\left(\frac{R_{33}}{R}\right), & d_{12}\left(\frac{R_{44}}{R}\right), & d_{12}\left(\frac{R_{12}}{R}\right), & d_{12}\left(\frac{R_{13}}{R}\right), & d_{12}\left(\frac{R_{14}}{R}\right), & d_{12}\left(\frac{R_{23}}{R}\right), & d_{12}\left(\frac{R_{24}}{R}\right), & d_{12}\left(\frac{R_{34}}{R}\right), \\
d_{13}\left(\frac{R_{11}}{R}\right), & d_{13}\left(\frac{R_{22}}{R}\right), & d_{13}\left(\frac{R_{33}}{R}\right), & d_{13}\left(\frac{R_{44}}{R}\right), & d_{13}\left(\frac{R_{12}}{R}\right), & d_{13}\left(\frac{R_{13}}{R}\right), & d_{13}\left(\frac{R_{14}}{R}\right), & d_{13}\left(\frac{R_{23}}{R}\right), & d_{13}\left(\frac{R_{24}}{R}\right), & d_{13}\left(\frac{R_{34}}{R}\right), \\
d_{14}\left(\frac{R_{11}}{R}\right), & d_{14}\left(\frac{R_{22}}{R}\right), & d_{14}\left(\frac{R_{33}}{R}\right), & d_{14}\left(\frac{R_{44}}{R}\right), & d_{14}\left(\frac{R_{12}}{R}\right), & d_{14}\left(\frac{R_{13}}{R}\right), & d_{14}\left(\frac{R_{14}}{R}\right), & d_{14}\left(\frac{R_{23}}{R}\right), & d_{14}\left(\frac{R_{24}}{R}\right), & d_{14}\left(\frac{R_{34}}{R}\right), \\
d_{23}\left(\frac{R_{11}}{R}\right), & d_{23}\left(\frac{R_{22}}{R}\right), & d_{23}\left(\frac{R_{33}}{R}\right), & d_{23}\left(\frac{R_{44}}{R}\right), & d_{23}\left(\frac{R_{12}}{R}\right), & d_{23}\left(\frac{R_{13}}{R}\right), & d_{23}\left(\frac{R_{14}}{R}\right), & d_{23}\left(\frac{R_{23}}{R}\right), & d_{23}\left(\frac{R_{24}}{R}\right), & d_{23}\left(\frac{R_{34}}{R}\right), \\
d_{24}\left(\frac{R_{11}}{R}\right), & d_{24}\left(\frac{R_{22}}{R}\right), & d_{24}\left(\frac{R_{33}}{R}\right), & d_{24}\left(\frac{R_{44}}{R}\right), & d_{24}\left(\frac{R_{12}}{R}\right), & d_{24}\left(\frac{R_{13}}{R}\right), & d_{24}\left(\frac{R_{14}}{R}\right), & d_{24}\left(\frac{R_{23}}{R}\right), & d_{24}\left(\frac{R_{24}}{R}\right), & d_{24}\left(\frac{R_{34}}{R}\right), \\
d_{34}\left(\frac{R_{11}}{R}\right), & d_{34}\left(\frac{R_{22}}{R}\right), & d_{34}\left(\frac{R_{33}}{R}\right), & d_{34}\left(\frac{R_{44}}{R}\right), & d_{34}\left(\frac{R_{12}}{R}\right), & d_{34}\left(\frac{R_{13}}{R}\right), & d_{34}\left(\frac{R_{14}}{R}\right), & d_{34}\left(\frac{R_{23}}{R}\right), & d_{34}\left(\frac{R_{24}}{R}\right), & d_{34}\left(\frac{R_{34}}{R}\right),
\end{matrix}\right|$$

Here the signs of the terms in the last six rows have been changed from minus to plus.* Now multiply the first four rows by R_{12}/R and add them to the fifth row, the first four rows by R_{13}/R and add them to the sixth row, the first four rows by R_{14}/R and add them to the seventh row, and so on. Then, remembering that

$$d_{12}\left(\frac{R_{11}}{R}\right) = -\frac{2R_{11}R_{12}}{R^2} + \frac{1}{R}\frac{dR_{11}}{dr_{12}},$$

$$d_{12}\left(\frac{R_{22}}{R}\right) = -\frac{2R_{22}R_{12}}{R^2} + \frac{1}{R}\frac{dR_{22}}{dr_{12}},$$

$$d_{12}\left(\frac{R_{34}}{R}\right) = -\frac{2R_{34}R_{12}}{R_2} + \frac{1}{R}\frac{dR_{34}}{dr_{12}}, \text{ \&c., \&c.,}$$

and, further, that R_{pp} does not contain r_{ps}, so that $d(R_{pp})/dr_{ps} = 0$, we have, taking a factor $1/R$ out of each row,

$$\Delta = \frac{(-1)^6 n^{10}}{\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4^2}\frac{1}{R^{10}} \begin{vmatrix} 2R_{11}, & 0, & 0, & 0, & R_{12}, & R_{13}, & R_{14}, & 0, & 0, & 0, \\ 0, & 2R_{22}, & 0, & 0, & R_{12}, & 0, & 0, & R_{23}, & R_{24}, & 0, \\ 0, & 0, & 2R_{33}, & 0, & 0, & R_{13}, & 0, & R_{23}, & 0, & R_{34}, \\ 0, & 0, & 0, & 2R_{44}, & 0, & 0, & R_{14}, & 0, & R_{24}, & R_{34}, \\ 0, & 0, & \frac{dR_{33}}{dr_{12}}, & \frac{dR_{44}}{dr_{12}}, & \frac{dR_{12}}{dr_{12}}, & \frac{dR_{13}}{dr_{12}}, & \frac{dR_{14}}{dr_{12}}, & \frac{dR_{23}}{dr_{12}}, & \frac{dR_{24}}{dr_{12}}, & \frac{dR_{34}}{dr_{12}}, \\ 0, & \frac{dR_{22}}{dr_{13}}, & 0, & \frac{dR_{44}}{dr_{13}}, & \frac{dR_{12}}{dr_{13}}, & \frac{dR_{13}}{dr_{13}}, & \frac{dR_{14}}{dr_{13}}, & \frac{dR_{23}}{dr_{13}}, & \frac{dR_{24}}{dr_{13}}, & \frac{dR_{34}}{dr_{13}}, \\ 0, & \frac{dR_{22}}{dr_{14}}, & \frac{dR_{33}}{dr_{14}}, & 0, & \frac{dR_{12}}{dr_{14}}, & \frac{dR_{13}}{dr_{14}}, & \frac{dR_{14}}{dr_{14}}, & \frac{dR_{23}}{dr_{14}}, & \frac{dR_{24}}{dr_{14}}, & \frac{dR_{34}}{dr_{14}}, \\ \frac{dR_{11}}{dr_{23}}, & 0, & 0, & \frac{dR_{44}}{dr_{23}}, & \frac{dR_{12}}{dr_{23}}, & \frac{dR_{13}}{dr_{23}}, & \frac{dR_{14}}{dr_{23}}, & \frac{dR_{23}}{dr_{23}}, & \frac{dR_{24}}{dr_{23}}, & \frac{dR_{34}}{dr_{23}}, \\ \frac{dR_{11}}{dr_{24}}, & 0, & \frac{dR_{33}}{dr_{24}}, & 0, & \frac{dR_{12}}{dr_{24}}, & \frac{dR_{13}}{dr_{24}}, & \frac{dR_{14}}{dr_{24}}, & \frac{dR_{23}}{dr_{24}}, & \frac{dR_{24}}{dr_{24}}, & \frac{dR_{34}}{dr_{24}}, \\ \frac{dR_{11}}{dr_{34}}, & \frac{dR_{22}}{dr_{34}}, & 0, & 0, & \frac{dR_{12}}{dr_{34}}, & \frac{dR_{13}}{dr_{34}}, & \frac{dR_{14}}{dr_{34}}, & \frac{dR_{23}}{dr_{34}}, & \frac{dR_{24}}{dr_{34}}, & \frac{dR_{34}}{dr_{34}}, \end{vmatrix} \quad \text{(xxxiii.).}$$

The form of Δ is now clear for the case of any number of correlated organs.†

(10) *Case (i).* Let us evaluate Δ and its minors for the case of three correlated organs.

* The factor $(-1)^{\frac{1}{2}p(p-1)}$ must be introduced if we deal with p organs.

† We may reduce this determinant as follows to one of the 6th order. Divide the 1st, 2nd, 3rd, 4th, 5th, 6th 7th, 8th, 9th, 10th columns by $2R_{11}$, $2R_{22}$, $2R_{33}$, $2R_{44}$, R_{12}, R_{13}, R_{14}, R_{23}, R_{24}, R_{34} respectively. After this division, subtract the first column from the 5th, 6th, and 7th. The determinant reduces to one of the 9th order; subtract the first column of this new determinant from its 4th, 7th, and 8th

In this case $R = 1 - r_{23}^2 - r_{31}^2 - r_{12}^2 + 2r_{23}r_{31}r_{12}$, and we have

$$R_{11} = 1 - r_{23}^2, \qquad R_{22} = 1 - r_{13}^2, \qquad R_{33} = 1 - r_{12}^2,$$
$$R_{23} = r_{31}r_{12} - r_{23}, \qquad R_{31} = r_{12}r_{23} - r_{13}, \qquad R_{12} = r_{23}r_{13} - r_{12},$$

whence we have

$$\Delta = \frac{(-1)^3 n^6}{\sigma_1^2\sigma_2^2\sigma_3^2}\frac{1}{R^6}\left|\begin{array}{cccccc} 2(1-r_{23}^2), & 0, & 0, & r_{23}r_{13}-r_{12}, & r_{12}r_{23}-r_{13}, & 0, \\ 0, & 2(1-r_{13}^2), & 0, & r_{23}r_{13}-r_{12}, & 0, & r_{31}r_{12}-r_{23}, \\ 0, & 0, & 2(1-r_{12}^2) & 0, & r_{12}r_{23}-r_{13}, & r_{31}r_{12}-r_{23}, \\ 0, & 0, & -2r_{12} & -1, & r_{23}, & r_{31}, \\ 0, & -2r_{13}, & 0, & r_{23}, & -1, & r_{12}, \\ -2r_{23}, & 0, & 0, & r_{31}, & r_{12}, & -1. \end{array}\right|$$

Divide the first three columns by 2; multiply the last row by r_{23} and subtract from the first row; the fifth row by r_{13} and subtract from the second; the fourth row by r_{12} and subtract from the third; we find

column, it again reduces by one degree. Repeat the process twice more, and after a slight rearrangement, the fact that $d(R_{pp})/dr_{ps} = 0$ being remembered, we have, if $\xi_{uv} = R_{uv}/\sqrt{R_{uu}/R_{vv}}$,

$$\Delta = \frac{n^{10}}{\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4^2} \cdot \frac{2^4 R_{11}R_{22}R_{33}R_{44}}{R^{10}} \cdot R_{12}R_{13}R_{14}R_{23}R_{24}R_{34}$$

$$\times \left|\begin{array}{cccccccc} \frac{d}{dr_{12}}\log\xi_{12}, & \frac{d}{dr_{12}}\log\xi_{13}, & \frac{d}{dr_{12}}\log\xi_{14}, & \frac{d}{dr_{12}}\log\xi_{23}, \ldots & \ldots & \ldots & \ldots \\ \frac{d}{dr_{13}}\log\xi_{12}, & \frac{d}{dr_{13}}\log\xi_{13}, & \frac{d}{dr_{13}}\log\xi_{14}, & \ldots & \ldots & \ldots & \ldots & \ldots \\ \frac{d}{dr_{14}}\log\xi_{12}, & \frac{d}{dr_{14}}\log\xi_{13}, & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \frac{d}{dr_{23}}\log\xi_{12}, & \frac{d}{dr_{23}}\log\xi_{13}, & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \frac{d}{dr_{24}}\log\xi_{12}, & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \frac{d}{dr_{34}}\log\xi_{12}, & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{array}\right|$$

the general run of terms being obvious.

In precisely similar manner the value of Δ for p organs can be written down, its degree being p less than the form given in (xxxiii.). We have not succeeded in reducing Δ for the general case [since writing this, Mr. ARTHUR BERRY, of King's College, Cambridge, has succeeded in reducing the determinant for $p = 4$, and also in showing its relation to elliptic space], but we feel fairly confident that its value will be found to be $\dfrac{2^p n^{\frac{1}{2}p(p+1)}}{\sigma_1^2\sigma_2^2 \ldots \sigma_p^2 R^{p+1}}$.

$$\Delta = \frac{8n^6(-1)^3}{\sigma_1^2\sigma_2^2\sigma_3^2}\frac{1}{R^6}\left|\begin{array}{cccccc} 1, & 0, & 0, & -r_{12}, & -r_{13}, & r_{23}, \\ 0, & 1, & 0, & -r_{12}, & r_{13}, & -r_{23}, \\ 0, & 0, & 1, & r_{12}, & -r_{13}, & -r_{23}, \\ 0, & 0, & -r_{12}, & -1, & r_{23}, & r_{31}, \\ 0, & -r_{13}, & 0, & r_{23}, & -1, & r_{12}, \\ -r_{23}, & 0, & 0, & r_{31}, & r_{12}, & -1. \end{array}\right|$$

Add the first column multiplied by r_{12} to the fourth; add the first multiplied by r_{13} to the fifth; and subtract the first multiplied by r_{23} from the sixth, the determinant will reduce to the minor of the first row and column. Continuing this process twice more, we ultimately deduce

$$\Delta = \frac{8n^6(-1)^3}{\sigma_1^2\sigma_2^2\sigma_3^2}\frac{1}{R^6}\left|\begin{array}{ccc} -R_{11}, & -R_{12}, & -R_{13} \\ -R_{12}, & -R_{22}, & -R_{23} \\ -R_{13}, & -R_{23}, & -R_{23} \end{array}\right| = \frac{8n^6}{\sigma_1^2\sigma_2^2\sigma_3^2}\frac{1}{R^4} \quad . \quad . \text{ (xxxiv.)}.$$

We will now proceed to calculate such of the minors of Δ as will give us results beyond those obtained for two correlated organs.*

We require the correlation of σ_1 and r_{23}, and of r_{12} and r_{13}.

Taking σ_1 and r_{23} we must strike out in (xxxii.), as we have taken only three organs, the 4th, 7th, 9th, and 10th rows and columns straight off, and for the required minor the 1st row and the 8th column. We have then for the minor M (σ_1, r_{23}),

$$M(\sigma_1, r_{23}) = -\frac{n^5}{\sigma_1\sigma_2^2\sigma_3^2}\left|\begin{array}{ccccc} \frac{r_{12}R_{12}}{R}, & \frac{R+R_{22}}{R}, & \frac{r_{23}R_{23}}{R}, & \frac{R_{12}}{R}, & 0, \\ \frac{r_{13}R_{13}}{R}, & \frac{r_{23}R_{23}}{R}, & \frac{R+R_{33}}{R}, & 0, & \frac{R_{13}}{R}, \\ \frac{R_{12}}{R}, & \frac{R_{12}}{R}, & 0, & -\frac{d}{dr_{12}}\left(\frac{R_{12}}{R}\right), & -\frac{d}{dr_{12}}\left(\frac{R_{13}}{R}\right), \\ \frac{R_{13}}{R}, & 0, & \frac{R_{13}}{R}, & -\frac{d}{dr_{13}}\left(\frac{R_{12}}{R}\right), & -\frac{d}{dr_{13}}\left(\frac{R_{13}}{R}\right), \\ 0, & \frac{R_{23}}{R}, & \frac{R_{23}}{R}, & -\frac{d}{dr_{23}}\left(\frac{R_{12}}{R}\right), & -\frac{d}{dr_{23}}\left(\frac{R_{13}}{R}\right). \end{array}\right|$$

To reduce this expression take the third row multiplied by r_{12} from the first, and the fifth row multiplied by r_{23} from the first. Then take the fourth row multiplied by r_{13} from the second, and the fifth row multiplied by r_{23} from the second, then remembering that

* As a matter of fact all the minors were worked out and the results of (xv.) to (xviii.) thus verified.

$$\mathrm{R} = \mathrm{R}_{22} + r_{12}\mathrm{R}_{12} + r_{23}\mathrm{R}_{23},$$
$$\mathrm{R} = \mathrm{R}_{33} + r_{13}\mathrm{R}_{13} + r_{23}\mathrm{R}_{23},$$

and that generally,

$$\frac{d\,(\mathrm{R}_{s'p'}/\mathrm{R})}{dr_{sp}} = \frac{d}{dr_{s'p'}}\left(\mathrm{R}_{sp}/\mathrm{R}\right).$$

We find :

$$\mathrm{M}\,(\sigma_1, r_{23}) = -\frac{n^5}{\sigma_1\sigma_2^2\sigma_3^2}\begin{vmatrix} 0, & \dfrac{2\mathrm{R}_{22}}{\mathrm{R}}, & 0, & -\dfrac{d}{dr_{12}}\left(\dfrac{\mathrm{R}_{22}}{\mathrm{R}}\right), & -\dfrac{d}{dr_{13}}\left(\dfrac{\mathrm{R}_{22}}{\mathrm{R}}\right), \\ 0, & 0, & \dfrac{2\mathrm{R}_{33}}{\mathrm{R}}, & -\dfrac{d}{dr_{12}}\left(\dfrac{\mathrm{R}_{33}}{\mathrm{R}}\right), & -\dfrac{d}{dr_{13}}\left(\dfrac{\mathrm{R}_{33}}{\mathrm{R}}\right), \\ \dfrac{\mathrm{R}_{12}}{\mathrm{R}}, & \dfrac{\mathrm{R}_{12}}{\mathrm{R}}, & 0, & -\dfrac{d}{dr_{12}}\left(\dfrac{\mathrm{R}_{12}}{\mathrm{R}}\right), & -\dfrac{d}{dr_{12}}\left(\dfrac{\mathrm{R}_{13}}{\mathrm{R}}\right), \\ \dfrac{\mathrm{R}_{13}}{\mathrm{R}}, & 0, & \dfrac{\mathrm{R}_{13}}{\mathrm{R}}, & -\dfrac{d}{dr_{13}}\left(\dfrac{\mathrm{R}_{12}}{\mathrm{R}}\right), & -\dfrac{d}{dr_{13}}\left(\dfrac{\mathrm{R}_{13}}{\mathrm{R}}\right), \\ 0, & \dfrac{\mathrm{R}_{23}}{\mathrm{R}}, & \dfrac{\mathrm{R}_{23}}{\mathrm{R}}, & -\dfrac{d}{dr_{23}}\left(\dfrac{\mathrm{R}_{12}}{\mathrm{R}}\right), & -\dfrac{d}{dr_{23}}\left(\dfrac{\mathrm{R}_{13}}{\mathrm{R}}\right). \end{vmatrix}$$

Multiply the first two columns by $\mathrm{R}_{12}/\mathrm{R}$ and subtract their sum from the fourth; multiply by the second two columns by $\mathrm{R}_{13}/\mathrm{R}$ and subtract their sum from the fifth; divide out by the factor $1/\mathrm{R}^5$. We obtain

$$\mathrm{M}\,(\sigma_1, r_{23}) = -\frac{n^5}{\sigma_1\sigma_2^2\sigma_3^2}\,\frac{1}{\mathrm{R}^5}\begin{vmatrix} 0, & 2\mathrm{R}_{22}, & 0, & \dfrac{d\mathrm{R}_{22}}{dr_{12}}, & \dfrac{d\mathrm{R}_{22}}{dr_{13}}, \\ 0, & 0, & 2\mathrm{R}_{33}, & \dfrac{d\mathrm{R}_{33}}{dr_{12}}, & \dfrac{d\mathrm{R}_{33}}{dr_{13}}, \\ \mathrm{R}_{12}, & \mathrm{R}_{12}, & 0, & \dfrac{d\mathrm{R}_{12}}{dr_{12}}, & \dfrac{d\mathrm{R}_{13}}{dr_{12}}, \\ \mathrm{R}_{13}, & 0, & \mathrm{R}_{13}, & \dfrac{d\mathrm{R}_{12}}{dr_{13}}, & \dfrac{d\mathrm{R}_{13}}{dr_{13}}, \\ 0, & \mathrm{R}_{23}, & \mathrm{R}_{23}, & \dfrac{d\mathrm{R}_{12}}{dr_{23}}, & \dfrac{d\mathrm{R}_{13}}{dr_{23}}. \end{vmatrix}$$

Now remembering that

$$\mathrm{R} = 1 - r_{23}^2 - r_{31}^2 - r_{12}^2 + 2r_{23}r_{31}r_{12},$$

substitute the various terms, and we find :

$$\mathrm{M}\,(\sigma_1, r_{23}) = -\frac{4n^5}{\sigma_1\sigma_2^2\sigma_3^2}\,\frac{1}{\mathrm{R}^5}\begin{vmatrix} 0, & 1 - r_{13}^2, & 0, & 0, & -r_{13}, \\ 0, & 0, & 1 - r_{12}^2, & -r_{12}, & 0, \\ -r_{12} + r_{23}r_{13}, & -r_{12} + r_{23}r_{13}, & 0, & -1, & r_{23}, \\ -r_{13} + r_{23}r_{12}, & 0, & -r_{13} + r_{12}r_{23}, & r_{23}, & -1, \\ 0, & -r_{23} + r_{12}r_{13}, & -r_{23} + r_{12}r_{13}, & r_{13}, & r_{12}. \end{vmatrix}$$

Subtract r_{13} times the fifth from the second, and r_{12} times the fourth from the third column :

$$M(\sigma_1, r_{23}) = -\frac{4n^5}{\sigma_1\sigma_2^2\sigma_3^2}\frac{1}{R^5}\begin{vmatrix} 0, & 1, & 0, & 0, & -r_{13} \\ 0, & 0, & 1, & -r_{12}, & 0 \\ -r_{12}+r_{23}r_{13}, & -r_{12}, & r_{12}, & -1, & r_{23} \\ -r_{13}+r_{23}r_{12}, & r_{13}, & -r_{13}, & r_{23}, & -1 \\ 0, & -r_{23}, & -r_{23}, & r_{13}, & r_{12} \end{vmatrix}.$$

Add r_{13} times the second column to the fifth and r_{12} times the third to the fourth, we have

$$M(\sigma_1, r_{23}) = -\frac{4n^5}{\sigma_1\sigma_2^2\sigma_3^2}\frac{1}{R^5}\begin{vmatrix} 0, & 1, & 0, & 0, & 0 \\ 0, & 0, & 1, & 0, & 0 \\ R_{12}, & -r_{12}, & r_{12}, & -R_{33}, & -R_{23} \\ R_{13}, & r_{13}, & -r_{13}, & -R_{23}, & -R_{22} \\ 0, & -r_{23}, & -r_{23}, & -R_{13}, & -R_{12} \end{vmatrix}$$

$$= -\frac{4n^5}{\sigma_1\sigma_2^2\sigma_3^2}\frac{1}{R^5}\begin{vmatrix} R_{33}, & R_{23}, & R_{21} \\ R_{23}, & R_{22}, & R_{13} \\ R_{13}, & R_{12}, & 0 \end{vmatrix}$$

$$= -\frac{4n^5}{\sigma_1\sigma_2^2\sigma_3^2}\frac{1}{R^5}\{R_{13}(R_{23}R_{13} - R_{21}R_{33}) + R_{12}(R_{21}R_{23} - R_{13}R_{22})\}$$

$$= -\frac{4n^5}{\sigma_1\sigma_2^2\sigma_3^2}\frac{1}{R^5}\{R_{13}r_{12}R + R_{12}r_{13}R\}$$

$$= \frac{4n^5}{\sigma_1\sigma_2^2\sigma_3^2}\frac{1}{R^4}\{r_{12}(r_{13} - r_{12}r_{23}) + r_{13}(r_{12} - r_{13}r_{23})\} \quad . \quad . \quad . \quad . \quad . \quad (\text{xxxv.}).$$

Now we have seen that

$$R_{\sigma_1 r_{23}} = \frac{M_{(\sigma_1, r_{23})}}{\Delta\Sigma_{\sigma_1}\Sigma_{r_{23}}} \quad \text{by (viii.) of Art. (4),}$$

and further Σ_{σ_1} and $\Sigma_{r_{23}}$ are given by (xv.) and (xvi.) of Art. (6). Hence

$$R_{\sigma_1 r_{23}} = \frac{r_{12}(r_{13} - r_{12}r_{23}) + r_{13}(r_{12} - r_{13}r_{23})}{\sqrt{2}\,.\,(1 - r_{23}^2)} \quad . \quad . \quad . \quad . \quad . \quad (\text{xxxvi.}).$$

In the next place we will determine $R_{r_{12}r_{13}}$. The minor $M(r_{12}, r_{13})$ is given by

$$M(r_{12}, r_{13}) = -\frac{n^5}{\sigma_1^2\sigma_2^2\sigma_3^2} \begin{vmatrix} \frac{R+R_{11}}{R}, & \frac{r_{12}R_{12}}{R}, & \frac{r_{13}R_{13}}{R}, & \frac{R_{13}}{R}, & 0, \\ \frac{r_{12}R_{12}}{R}, & \frac{R+R_{22}}{R}, & \frac{r_{23}R_{23}}{R}, & 0, & \frac{R_{23}}{R}, \\ \frac{r_{13}R_{13}}{R}, & \frac{r_{23}R_{23}}{R}, & \frac{R+R_{33}}{R}, & \frac{R_{13}}{R}, & \frac{R_{23}}{R}, \\ \frac{R_{13}}{R}, & \frac{R_{12}}{R}, & 0, & -\frac{d}{dr_{12}}\left(\frac{R_{13}}{R}\right), & -\frac{d}{dr_{12}}\left(\frac{R_{23}}{R}\right), \\ 0, & \frac{R_{23}}{R}, & \frac{R_{23}}{R}, & -\frac{d}{dr_{23}}\left(\frac{R_{13}}{R}\right), & -\frac{d}{dr_{23}}\left(\frac{R_{23}}{R}\right), \end{vmatrix}$$

$$= -\frac{n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^5} \begin{vmatrix} R+R_{11}, & r_{12}R_{12}, & r_{13}R_{13}, & R_{13}, & 0, \\ r_{12}R_{12}, & R+R_{22}, & r_{23}R_{23}, & 0, & R_{23}, \\ r_{13}R_{13}, & r_{23}R_{23}, & R+R_{33}, & R_{13}, & R_{23}, \\ R_{12}, & R_{12}, & 0, & \frac{2R_{13}R_{12}}{R} - r_{23}, & \frac{2R_{23}R_{12}}{R} - r_{13}, \\ 0, & R_{23}, & R_{23}, & \frac{2R_{13}R_{23}}{R} - r_{12}, & \frac{2R_{23}^2}{R} + 1. \end{vmatrix}$$

Add the first three rows together, multiply by R_{12}/R and subtract from the fourth, and by R_{23}/R and subtract from the fifth, we find

$$M(r_{12}, r_{13}) = -\frac{n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^5} \begin{vmatrix} R+R_{11}, & r_{12}R_{12}, & r_{13}R_{13}, & R_{13}, & 0, \\ r_{12}R_{12}, & R+R_{22}, & r_{23}R_{23}, & 0, & R_{23}, \\ r_{13}R_{13}, & r_{23}R_{23}, & R+R_{33}, & R_{13}, & R_{23}, \\ -R_{12}, & -R_{12}, & -2R_{12}, & -r_{23}, & -r_{13}, \\ -2R_{23}, & -R_{23}, & -R_{23}, & -r_{12}, & 1. \end{vmatrix}$$

Multiply the fourth column by r_{13}, and the fifth by r_{23}, and subtract from the third column

$$M(r_{12}, r_{13}) = -\frac{n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^5} \begin{vmatrix} R+R_{11}, & r_{12}R_{12}, & 0, & R_{13}, & 0 \\ r_{12}R_{12}, & R+R_{22}, & 0, & 0, & R_{23} \\ r_{13}R_{13}, & r_{23}R_{23}, & 2R_{33}, & R_{13}, & R_{23} \\ -R_{12}, & -R_{12}, & 2r_{12}, & -r_{23}, & -r_{13} \\ -2R_{23}, & -R_{23}, & 0, & -r_{12}, & 1 \end{vmatrix}.$$

Multiply the third row by r_{12}, and the fourth by $1 - r_{12}^2$ or R_{33}, and subtract the latter; the determinant now reduces to one of the fourth order, and we find :—

$$M(r_{12}, r_{13}) = -\frac{2n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^5}\begin{vmatrix} R + R_{11}, & r_{12}R_{12}, & R_{13}, & 0 \\ r_{12}R_{12}, & R + R_{22}, & 0, & R_{23} \\ r_{12}R_{33} + r_{13}R_{23}, & r_{12}R_{33} + r_{23}R_{13}, & R_{23}, & R_{13} \\ -2R_{23}, & -R_{23}, & -r_{12}, & 1 \end{vmatrix}.$$

Add the third column, multiplied by r_{13} to, and subtract the fourth multiplied by r_{23} from, the second, and then subtract the first

$$M(r_{12}, r_{13}) = -\frac{2n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^5}\begin{vmatrix} R + R_{11}, & -2R_{11}, & R_{13}, & 0 \\ r_{12}R_{12}, & 2R_{22}, & 0, & R_{23} \\ r_{12}R_{33} + r_{13}R_{23}, & 0, & R_{23}, & R_{13} \\ -2R_{23}, & -2r_{23}, & -r_{12}, & 1 \end{vmatrix}.$$

Divide out the 2 in the second column, add it to the first, and subtract the third column multiplied by r_{13}

$$M(r_{12}, r_{13}) = -\frac{4n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^5}\begin{vmatrix} R - r_{13}R_{13}, & -R_{11}, & R_{13}, & 0 \\ R - r_{23}R_{23}, & R_{22}, & 0, & R_{23} \\ r_{12}R_{33}, & 0, & R_{23}, & R_{13} \\ -R_{23}, & -r_{23}, & -r_{12}, & 1 \end{vmatrix}.$$

Add r_{23} times the last row to the first

$$M(r_{12}, r_{13}) = -\frac{4n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^5}\begin{vmatrix} R_{33}, & -1, & -r_{13}, & r_{23} \\ R - r_{23}R_{23}, & R_{22}, & 0, & R_{23} \\ r_{12}R_{33}, & 0, & R_{23}, & R_{13} \\ -R_{23}, & -r_{23}, & -r_{12}, & 1 \end{vmatrix}.$$

Subtract r_{23} times the first row from the last, and remember that $-R_{23} - r_{23}R_{33} = r_{12}R_{13}$,

$$M(r_{12}, r_{13}) = -\frac{4n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^5}\begin{vmatrix} R_{33}, & -1, & -r_{13}, & r_{23} \\ R - r_{23}R_{23}, & R_{22}, & 0, & R_{23} \\ r_{12}R_{33}, & 0, & R_{23}, & R_{13} \\ r_{12}R_{13}, & 0, & R_{12}, & R_{11} \end{vmatrix}.$$

Multiply the first row by R_{22} and add it to the second; the determinant reduces to the third order,

$$M(r_{12}, r_{13}) = -\frac{4n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^5}\begin{vmatrix} R - r_{23}R_{23} + R_{22}R_{33}, & -r_{13}R_{22}, & -r_{13}R_{12} \\ r_{12}R_{33}, & R_{23}, & R_{13} \\ r_{12}R_{13}, & R_{12}, & R_{11} \end{vmatrix}.$$

Expanding this determinant from its first column, remembering that $R_{23}R_{11}-R_{13}R_{12} = -r_{23}R$,

$$M(r_{12}, r_{13}) = -\frac{4n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^5}\{-r_{23}R(R - r_{23}R_{23} + R_{22}R_{33}) + r_{12}r_{13}(R_{22}R_{11} - R_{12}^2)R_{33} - r_{12}r_{13}(R_{22}R_{13} - R_{12}R_{23})R_{13}\}.$$

Or, since $R_{22}R_{11} - R_{12}^2 = R$, $R_{12}R_{23} - R_{22}R_{13} = r_{13}R$, and $R - r_{23}R_{23} + R_{22}R_{33} = 2R_{22}R_{33} - r_{12}r_{13}R_{23}$,

$$M(r_{12}, r_{13}) = \frac{4n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^4}\{2r_{23}R_{22}R_{33} - r_{12}r_{13}(R_{33} + r_{23}R_{23} + r_{13}R_{13})\},$$

$$= \frac{8n^5}{\sigma_1^2\sigma_2^2\sigma_3^2 R^4}\left\{r_{23}R_{22}R_{33} - \frac{r_{12}r_{13}}{2}R\right\}.$$

Hence, since

$$R_{r_{12}r_{13}} = \frac{M_{(r_{12}, r_{13})}}{\Delta\Sigma_{r_{12}}\Sigma_{r_{13}}},$$

we have by (xvi.) of Art. (6) and (xxxiv.) of Art. (9)

$$R_{r_{12}r_{13}} = r_{23} - \tfrac{1}{2}r_{12}r_{13}\frac{R}{R_{22}R_{33}}$$

$$= r_{23} - \frac{r_{12}r_{13}(1 - r_{23}^2 - r_{13}^2 - r_{12}^2 + 2r_{23}r_{12}r_{13})}{2(1 - r_{13}^2)(1 - r_{12}^2)} \quad . \quad . \quad . \quad . \quad \text{(xxxvii.)}.$$

To complete the theory for the errors made in an investigation of the constants for a system of three correlated organs, we require to determine the probable error of a regression coefficient for a partial regression of a first organ on a second, the third organ being constant. This coefficient is given by

$${}_3\rho_{12} = \frac{r_{12} - r_{23}r_{13}}{1 - r_{23}^2}\,\frac{\sigma_1}{\sigma_2}.$$

Take logarithmic differentials

$$\frac{\delta_3\rho_{12}}{{}_3\rho_{12}} = \frac{\delta r_{12}}{r_{12} - r_{23}r_{13}} - \frac{r_{23}\,\delta r_{13}}{r_{12} - r_{23}r_{13}} + \delta r_{23}\left\{\frac{-r_{13}}{r_{12} - r_{23}r_{13}} + \frac{2r_{23}}{1 - r_{23}^2}\right\} + \frac{\delta\sigma_1}{\sigma_1} - \frac{\delta\sigma_2}{\sigma_2}.$$

Let this be squared and divided by n, and then the values found above for the standard deviations of the errors in r_{12}, r_{13}, r_{23}, σ_1 and σ_2, and for the correlations of errors in these quantities be substituted. After some lengthy algebraic reductions, which it seems unnecessary to reproduce, there results

$$\frac{1}{({}_3\rho_{12})^2}(\Sigma_{{}_3\rho_{12}})^2 = \frac{1 - r_{23}^2 - r_{13}^2 - r_{12}^2 + 2r_{23}r_{13}r_{12}}{n(r_{12} - r_{13}r_{23})^2},$$

or

$$\Sigma_{_3\rho_{12}} = \frac{1}{\sqrt{n}} \frac{\sigma_1}{\sigma_2} \frac{\sqrt{(1 - r_{23}^2 - r_{13}^2 - r_{12}^2 + 2r_{23}r_{13}r_{12})}}{1 - r_{23}^2} \quad . \quad . \quad . \quad . \quad (\text{xxxviii.}).$$

The percentage probable error in a partial coefficient of regression is accordingly

$$67 \cdot 449 \sqrt{\mathrm{R}} \,/\, \sqrt{n} \, (- \mathrm{R}_{12}).$$

Before discussing the significance of these quantitative results for three organs, it seems desirable to complete the general case by investigating the correlation between the errors made in the correlation coefficients of a first pair of organs and a second different pair of organs.

(11.) Case (ii.). *Case of Four or more Organs.*—In the case of four or more organs the only new probable error will be that of a partial regression coefficient, but this can theoretically always be found by the method of the preceding paragraph, provided we know all the error correlations. The only novel correlation among the errors will be that of $r_{12}r_{34}$ and this we shall now proceed to investigate. The discovery of an error correlation coefficient of this type completes the theory of the errors of normal frequency constants.

Instead of evaluating Δ of (xxxiii.), which in the case of four organs appears to be very laborious, we may proceed as follows :—

If Δ be written in the form

$$\begin{vmatrix}
a_{\sigma_1\sigma_1}, & a_{\sigma_2\sigma_1}, & \cdot\cdot & a_{\sigma_4\sigma_1}, & a_{\sigma_1 r_{12}}, & a_{\sigma_1 r_{13}}, & \cdot\cdot & a_{\sigma_1 r_{34}}, \\
a_{\sigma_1\sigma_2}, & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
\cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
a_{\sigma_1\sigma_4}, & a_{\sigma_2\sigma_4}, & \cdot\cdot & a_{\sigma_4\sigma_4}, & a_{\sigma_4 r_{12}}, & a_{\sigma_4 r_{13}}, & \cdot\cdot & a_{\sigma_4 r_{34}}, \\
a_{\sigma_1 r_{12}}, & a_{\sigma_2 r_{12}}, & \cdot\cdot & a_{\sigma_4 r_{12}}, & a_{\sigma_2 r_{12}}, & a_{r_{12} r_{13}}, & \cdot\cdot & a_{r_{12} r_{34}}, \\
\cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
\cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
a_{\sigma_1 r_{34}}, & a_{\sigma_2 r_{34}}, & \cdot\cdot & a_{\sigma_4 r_{34}}, & a_{r_{12} r_{34}}, & a_{r_{12} r_{13}}, & \cdot\cdot & a_{r_{34} r_{34}},
\end{vmatrix}$$

and M denote the minor of the corresponding a in Δ, we have by a well known property of the determinant,

$$a_{\sigma_4\sigma_1}\mathrm{M}_{(r_{12},\, \sigma_1)} + a_{\sigma_4\sigma_2}\mathrm{M}_{(r_{12},\, \sigma_2)} + a_{\sigma_4\sigma_3}\mathrm{M}_{(r_{12},\, \sigma_3)} + \ldots + a_{\sigma_4 r_{34}}\mathrm{M}_{(r_{12},\, r_{34})} = 0.$$

Divide by Δ,

$$a_{\sigma_4\sigma_1}\frac{\mathrm{M}_{(r_{12},\, \sigma_1)}}{\Delta} + a_{\sigma_4\sigma_2}\frac{\mathrm{M}_{(r_{12},\, \sigma_2)}}{\Delta} + a_{\sigma_4\sigma_3}\frac{\mathrm{M}_{(r_{12},\, \sigma_3)}}{\Delta} + \ldots + a_{\sigma_4 r_{34}}\frac{\mathrm{M}_{(r_{12},\, r_{34})}}{\Delta} = 0.$$

Now $\frac{M_{(r_{12},\,\sigma_1)}}{\Delta\Sigma_{r_{12}}\Sigma_{\sigma_1}}$, $\frac{M_{(r_{12},\,\sigma_2)}}{\Delta\Sigma_{r_{12}}\Sigma_{\sigma_2}}$, $\frac{M_{(r_{12},\,\sigma_3)}}{\Delta\Sigma_{r_{12}}\Sigma_3}$, . . . $\frac{M_{(r_{12},\,r_{34})}}{\Delta\Sigma_{r_{12}}\Sigma_{r_{34}}}$, are the correlations between the errors made in the various quantities σ_1, σ_2, σ_3, σ_4, r_{12}, r_{13}, r_{14}, r_{23}, r_{24}, r_{34}, every one of which is known by the previous investigations except that of r_{12} and r_{34} or $M_{(r_{12},\,r_{34})}/\Delta\Sigma_{r_{12}}\Sigma_{r_{34}}$. Hence the above equation will suffice to find the latter quantity, since $\Sigma_{r_{12}}$ and $\Sigma_{r_{34}}$ are known. We have

$$\left.\begin{array}{ll} a_{\sigma_4\sigma_1} = \dfrac{n}{\sigma_4\sigma_1}\dfrac{R_{14}r_{14}}{R}, & a_{\sigma_4\sigma_2} = \dfrac{n}{\sigma_4\sigma_2}\dfrac{R_{24}r_{24}}{R}, \\ a_{\sigma_4\sigma_3} = \dfrac{n}{\sigma_4\sigma_3}\dfrac{R_{34}r_{34}}{R}, & a_{\sigma_4\sigma_4} = \dfrac{n}{\sigma_4^2}\left(1 + \dfrac{R_{44}}{R}\right), \end{array}\right\} \text{ by (xxiv.) and (xxv.),}$$

$$a_{\sigma_4 r_{12}} = 0, \qquad a_{\sigma_4 r_{13}} = 0, \qquad a_{\sigma_4 r_{23}} = 0, \quad \text{by (xxvii.),}$$

$$a_{\sigma_4 r_{14}} = \frac{n}{\sigma_4}\frac{R_{14}}{R}, \qquad a_{\sigma_4 r_{24}} = \frac{n}{\sigma_4}\frac{R_{24}}{R}, \qquad a_{\sigma_4 r_{34}} = \frac{n}{\sigma_4}\frac{R_{34}}{R}, \quad \text{by (xxvi.).}$$

Further,

$$\frac{M_{(r_{12},\,\sigma_1)}}{\Delta} = R_{r_{12}\sigma_1}\Sigma_{r_{12}}\Sigma_{\sigma_1} = \frac{\sigma_1}{2n}\,r_{12}\,(1 - r_{12}^2),$$

$$\frac{M_{(r_{12},\,\sigma_2)}}{\Delta} = \frac{\sigma_2}{2n}\,r_{12}\,(1 - r_{12}^2), \quad \text{by (xvi.) and (xviii.),}$$

$$\left.\begin{array}{l} \dfrac{M_{(r_{12},\,\sigma_3)}}{\Delta} = \dfrac{\sigma_3}{2n}\,\{r_{23}\,(r_{13} - r_{12}r_{23}) + r_{13}\,(r_{23} - r_{12}r_{13})\} \\ \dfrac{M_{(r_{12},\,\sigma_4)}}{\Delta} = \dfrac{\sigma_4}{2n}\,\{r_{24}\,(r_{14} - r_{24}r_{12}) + r_{14}\,(r_{24} - r_{12}r_{14})\}, \end{array}\right\} \text{ by (xxxvi.),}$$

$$\left.\begin{array}{l} \dfrac{M_{(r_{12},\,r_{14})}}{\Delta} = \dfrac{1}{n}\,\{r_{24}(1 - r_{12}^2)(1 - r_{14}^2) - \tfrac{1}{2}\,r_{12}r_{14}R_{33}\}, \\ \dfrac{M_{(r_{12},\,r_{24})}}{\Delta} = \dfrac{1}{n}\,\{r_{14}(1 - r_{12}^2)(1 - r_{24}^2) - \tfrac{1}{2}\,r_{12}r_{24}R_{33}\}, \end{array}\right\} \text{ by (xxxvii.),}$$

$$\frac{M_{(r_{12},\,r_{34})}}{\Delta} = \frac{1}{n}\,R_{r_{12}r_{34}}\,(1 - r_{12}^2)\,(1 - r_{34}^2), \quad \text{by (xvi.).}$$

Now substitute and divide out by common factors, and we find

$$\begin{aligned} &(r_{14}R_{14} + r_{24}R_{24})\,r_{12}\,(1 - r_{12}^2) + r_{34}R_{34}\,\{r_{23}\,(r_{13} - r_{12}r_{23}) + r_{13}\,(r_{23} - r_{12}r_{13})\} \\ &\quad + (R + R_{44})\,\{r_{24}\,(r_{14} - r_{24}r_{12}) + r_{14}\,(r_{24} - r_{12}r_{14})\} \\ &\quad + 2R_{14}\,\{r_{24}(1 - r_{12}^2)(1 - r_{14}^2) - \tfrac{1}{2}\,r_{12}r_{14}R_{33}\} \\ &\quad + 2R_{24}\,\{r_{14}(1 - r_{12}^2)(1 - r_{24}^2) - \tfrac{1}{2}\,r_{12}r_{24}R_{33}\} \\ &\quad + 2R_{34}R_{r_{12}r_{34}}(1 - r_{12}^2)(1 - r_{34}^2) = 0 \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad \text{(xxxix.).} \end{aligned}$$

Now

$$\begin{aligned} R &= R_{44} + r_{14}R_{14} + r_{24}R_{24} + r_{34}R_{34}, \\ 0 &= r_{14}R_{44} + R_{14} + r_{12}R_{24} + r_{31}R_{34}, \\ 0 &= r_{24}R_{44} + r_{12}R_{14} + R_{24} + r_{32}R_{34}. \end{aligned}$$

Multiply the third of these by r_{14} and the second by r_{24} and add, we have

$$0 = 2r_{14}r_{24}R_{44} + r_{24}R_{14} + r_{14}R_{24} + r_{12}(r_{14}R_{14} + r_{24}R_{24}) + (r_{13}r_{24} + r_{23}r_{14})R_{34}.$$

Hence, by the first,

$$r_{24}R_{14} + r_{14}R_{24} = R_{44}(r_{12} - 2r_{14}r_{24}) - Rr_{12} + R_{34}(r_{12}r_{34} - r_{13}r_{24} - r_{23}r_{14}),$$

while

$$r_{14}R_{14} + r_{24}R_{24} = R - R_{44} - r_{34}R_{34}.$$

By means of these relations, let us get rid of the terms in R_{14} and R_{24} in (xxxix.) above.

Re-arranging we have, after some reductions,

$$\begin{aligned} &- 2r_{12}(1 - r_{12}^2)R + 2r_{12}R_{33}R_{44} + R_{34}\{r_{34}r_{23}(r_{13} - r_{12}r_{23}) + r_{34}r_{13}(r_{23} - r_{12}r_{13}) \\ &\quad + 2(1 - r_{12}^2)(r_{12}r_{34} - r_{13}r_{24} - r_{23}r_{14}) + 2r_{14}r_{24}r_{34} - r_{12}r_{34}(r_{24}^2 + r_{14}^2)\} \\ &\quad + 2R_{34}R_{r_{12}r_{34}}(1 - r_{12}^2)(1 - r_{34}^2) = 0. \end{aligned}$$

But

$$(1 - r_{12}^2)R = R_{33}R_{44} - R_{34}^2.$$

Hence we can divide out by R_{34}, and accordingly,

$$\begin{aligned} &2R_{r_{12}r_{34}}(1 - r_{12}^2)(1 - r_{34}^2) \\ &\quad = - \{r_{34}r_{23}(r_{13} - r_{12}r_{23}) + r_{34}r_{13}(r_{23} - r_{12}r_{13}) + 2(1 - r_{12}^2)(r_{12}r_{34} - r_{13}r_{24} - r_{23}r_{41}) \\ &\qquad + 2r_{14}r_{24}r_{34} - r_{12}r_{34}(r_{24}^2 + r_{12}^2) + 2r_{12}R_{34}\}. \end{aligned}$$

Noting that

$$R_{34} = -r_{34}(1 - r_{12}^2) + r_{31}r_{41} + r_{32}r_{42} - r_{12}(r_{31}r_{42} + r_{32}r_{41}),$$

we have, after substitution and rearranging,

$$\begin{aligned} 2R_{r_{12}r_{34}}(1 - r_{12}^2)(1 - r_{34}^2) &= (r_{13} - r_{12}r_{23})(r_{24} - r_{23}r_{34}) + (r_{14} - r_{34}r_{13})(r_{23} - r_{12}r_{13}) \\ &\quad + (r_{13} - r_{14}r_{34})(r_{24} - r_{12}r_{14}) + (r_{14} - r_{12}r_{24})(r_{23} - r_{24}r_{34}). \end{aligned}$$

Or,

$$R_{r_{12}r_{34}} = \frac{\left\{\begin{array}{l} \quad (r_{13} - r_{12}r_{23})(r_{24} - r_{23}r_{34}) + (r_{14} - r_{34}r_{13})(r_{23} - r_{12}r_{13}) \\ + (r_{13} - r_{14}r_{34})(r_{24} - r_{12}r_{14}) + (r_{14} - r_{12}r_{24})(r_{23} - r_{24}r_{34}) \end{array}\right\}}{2(1 - r_{12}^2)(1 - r_{34}^2)} \quad . \quad . \quad (\text{xl.}).$$

If we put $4 = 1$ in this result and remember that $r_{11} = 1$, we find, after some reductions,

$$R_{r_{12}r_{31}} = r_{23} - \tfrac{1}{2} r_{12}r_{13} \frac{1 - r_{12}^2 - r_{13}^2 - r_{23}^2 + 2r_{12}r_{23}r_{31}}{(1 - r_{12}^2)(1 - r_{13}^2)},$$

which agrees with (xxxvii.), and may be taken as a verification of this result.*

(12.) We may draw several conclusions from the results (xxxvi.), (xxxvii.), and (xl.).

(α.) While errors in the correlations of a first organ with a second and a third have a correlation themselves of the first order, errors in the variation of a first organ and the correlation of two others; or in the correlation of two organs and in the correlation of a second two, have only correlation of the second order. Thus a selection of the correlation between two organs modifies the variation of all organs correlated with one or other or both of the first, but only in the second degree. Again, a selection of the correlation between two organs modifies the correlation of every other pair of organs, one or both of which are correlated with one or both of the first pair; but this is only in the second degree.

(β.) If two organs be entirely uncorrelated a random selection of the variation of a third organ correlated with both of them will tend to generate correlation between the hitherto uncorrelated organs, *i.e.*, put $r_{23} = 0$ in (xxxvi.), and we have

$$R_{\sigma_1 r_{23}} = \sqrt{2} \,.\, r_{12}r_{13}.$$

If a variation $\Delta\sigma_1$ be made in σ_1 the probable value of r_{23} is

$$\Delta r_{23} = R_{\sigma_1 r_{23}} \frac{\Sigma_{r_{23}}}{\Sigma_{\sigma_1}} \Delta\sigma_1 = 2r_{12}r_{13} \frac{\Delta\sigma_1}{\sigma_1},$$

which may clearly be of sensible magnitude. Thus correlation may be generated by selection of variation, and *vice versâ*.

(γ.) If two organs be each entirely uncorrelated with a third, yet a random selection, which produces a correlation between one of these organs and the third, will produce a correlation of the first order between the other of these organs and the third, *i.e.*, put $r_{12} = r_{13} = 0$ in (xxxvii.), we have

$$R_{r_{12}r_{13}} = r_{23},$$

a correlation of the first order between the probable changes.

(δ.) Consider four organs of which the first is alone correlated with the third and the second with the fourth, the third and fourth being themselves uncorrelated. Then any random selection which produces a correlation between the first and

* The probable error of a partial regression coefficient for p organs has not been worked owing to the labour involved, but judging by the cases on pp. 245 and 260, it may safely be taken as $67{\cdot}449 \sqrt{R}/\sqrt{n}\,(-R_{12})$, where R is now the determinant of the p^{th} degree.

second will tend to produce a correlation between the third and fourth, *i.e.*, if $r_{12} = r_{14} = r_{23} = r_{34} = 0$, we still have from (xl.)

$$R_{r_{12}r_{34}} = r_{13}r_{24}.$$

(ϵ.) We may further illustrate these principles by one or two hypothetical examples drawn from actual organs.

Let the actual organs be (1) physique of father, (2) artistic sense of mother, (3) physique of offspring, (4) artistic sense of offspring. Suppose in the general population there is no correlation between physique of father and artistic sense of mother, or between physique or artistic sense of parent, and artistic sense and physique respectively of offspring. Then $r_{12} = 0$, $r_{14} = 0$, $r_{23} = 0$, and, presumably, $r_{34} = 0$. Hence

$$R_{r_{12}r_{34}} = r_{13}r_{24}$$

is the product of the two coefficients for inheritance of physique from father to child, and for inheritance of artistic sense from mother to child.

Now let a random selection be made out of the general population in which assortative mating between physique in the male and artistic sense in the female presents itself, *i.e.*, let Δr_{12} be sensible; then we have, most probably,

$$\Delta r_{34} = R_{r_{12}r_{34}} \frac{\Sigma_{r_{34}}}{\Sigma_{r_{12}}} \Delta r_{12} = r_{13}r_{24} \Delta r_{12},$$

or, a correlation between physique and artistic sense in the offspring will tend to be developed. Generally, when r_{34} and r_{12} do not start from zero, we have,

$$\Delta r_{34} = \frac{r_{13}r_{24} + \frac{1}{2}r_{12}r_{34}(r_{13}^2 + r_{24}^2)}{(1 - r_{12}^2)^2} \Delta r_{12},$$

or, any increase of sexual selection in a group tends to emphasise the correlation of the selected qualities in the offspring.

Let the three characteristics be artistic sense (1) in a man, (2) in his mother, (3) in his wife. Then

$$R_{r_{12}r_{13}} = -\tfrac{1}{2}r_{12}r_{13}\frac{1 - r_{12}^2 - r_{13}^2}{(1 - r_{12}^2)(1 - r_{13}^2)},$$

if we suppose r_{23} to be zero.

Hence any selected group with a higher coefficient of maternal inheritance of heredity will have a less coefficient of sexual selection than the general population, and *vice versâ*. The *tendency* is, of course, independent of the magnitude of r, and really of the particular character. Supposing likeness of faculty or character to be a rough measure of "sympathy," we might conclude for any population with inheritance and sexual selection, that on the average a selected sub-group of men having greater sympathy with mothers than the general population will have less sympathy with wives, and *vice versâ*.

(13.) Many like propositions may be stated with regard to the action of selection on the correlation of characters. They require but little modification to state them for artificial or natural selection, as they are here stated for what we have termed random selection. The above will, however, suffice to indicate how every form of selection of variability or correlation influences in a manner capable of quantitative expression the variability and correlation of all other directly and indirectly correlated organs. Selection cannot be of service in altering *one* organ only, it alters at the same time the whole inter-relationship of a complex of organs. Evolution by natural selection can never be the change of one organ to suit a particular environment; it is the balance of advantage and disadvantage produced by the change of all organs involved in the attempt to select one of them. The moment the intimate correlation of organs in animal or plant life has been fully realised—and this realisation owing to recent statistical investigations has become fairly easy—then the conception of natural selection as moulding any single organ to what may be fittest to its surroundings must be discarded. The selection of the "fittest" in one organ would probably mean the selection of the unfit in other organs, and a general balance of fitness in the complex of organ is all that is possible.*

IV. On the Probable Errors and the Coefficients of Correlation between Errors made in the Determination of the Constants in the case of Skew Variation.

(14.) The case of Skew Variation has been dealt with at length by one of the present authors in the second paper of this series. He has shown that in a great variety of cases it can be dealt with by a series of curves having three principal algebraical types, each defined by a certain number of constants. The probable errors of the determination of these constants were not then investigated, but it is clearly of great importance for the practical use of these curves to know how far these constants can, for any given number of observations, be depended upon to give an accurate measure of the skewness and its special features. At the same time an investigation of the probable errors of these constants leads us to a number of novel properties which are connected with the theory of evolution in the frequent case of skew variation.

* Take, for example, result (xxxvi.); as far as terms of the second order are concerned $R_{\sigma_1 r_{23}} = \sqrt{2} \,.\, r_{12} r_{13}$. Hence, with positive correlation between three organs, the effect of trying to get a group very stable in one organ, *i.e.*, with a negative $\Delta\sigma_1$, is to reduce the correlation between every other pair of organs! In other words, we have to reduce variation at the expense of correlation, increased stability of one organ is gained at the expense of decreased stability in the inter-relationship of other organs. This may possibly be illustrated by the long bones of the French, where the lesser variability of the male relative to the female connotes also a lesser correlation. See Lee and Pearson: "On the Relative Variation and Correlation in Civilised and Uncivilised Races," 'R. S. Proc.,' vol. 61, pp. 354–356.

We shall deal first with the skew curve of Type III. ('Phil. Trans.' A, vol. 186, p. 373), because its treatment is less complex and leads at once to some general principles which must be borne in mind, whenever natural selection acts upon an organ exhibiting skew variation.

(15.) *Probable Errors and Error Correlations of the Constants of the Generalized Probability Curve of Type* $y = y_1\left(1 + \frac{\gamma x}{p+1}\right)^p e^{-\gamma x}$.

This is the equation of the curve referred to its mean as origin, where

$$y_1 = n\,\frac{\gamma e^{-(p+1)}(p+1)^p}{\Gamma(p+1)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xli.}).$$

Further, the moments about the centroid vertical are given by,

$$\mu_2 = \frac{p+1}{\gamma^2}, \quad \mu_3 = \frac{2(p+1)}{\gamma^3}, \quad \mu_4 = \frac{3(p+1)(p+3)}{\gamma^4} \quad . \quad . \quad . \quad (\text{xlii.}).$$

or,

$$\gamma = 2\mu_2/\mu_3, \; p = 4\mu_2^3/\mu_3^2 - 1 \; . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xliii.}).$$

The criterion for the application of this curve to any frequency distribution is

$$2\mu_2(3\mu_2^2 - \mu^4) + 3\mu_3^2 = 0,$$

or, if we write $\beta_2 = \mu_4/\mu_2^2$, $\beta_1 = \mu_3^2/\mu_2^3$,

$$6 - 2\beta_2 + 3\beta_1 = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xliv.}).$$

Lastly,

$$Sk. = \text{the skewness} = \tfrac{1}{2}\mu_3/(\mu_2)^{3/2} = \frac{1}{\sqrt{(p+1)}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xlv.}),$$

and the modal frequency,

$$y_0 = \frac{n\gamma p^p}{e^p\Gamma(p+1)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xlvi.}).*$$

We require to know the probable errors of p, γ, y_1, y_0, $\sigma = \sqrt{\mu_2}$, μ_2, μ_3, μ_4, and the skewness. We must discover the best physical constants to describe such skew frequencies and we shall at the same time succeed in deducing certain—we believe—novel properties of normal frequency distributions as limiting cases of this skew type of distribution.

* See 'Phil. Trans.,' A, vol. 186, pp. 373–4.

(16.) The first stage in the investigation is to apply the general proposition of our Art. 2, to

$$\log y = \log n + \log \gamma - (p+1) + p \log (p+1) - \log \Gamma (p+1)$$
$$+ p \log \left(1 + \frac{\gamma x}{p+1}\right) - \gamma x.$$

We find :

$$\frac{d^2 (\log y)}{dx^2} = - \frac{p}{a^2} \frac{1}{(1 + x/a)^2}, \text{ if } a = (p+1)/\gamma,$$

$$\frac{d^2 (\log y)}{dx\, dp} = \frac{1}{a} \left(\frac{1}{1 + x/a} - \frac{p}{p+1} \frac{1}{(1 + x/a)^2}\right),$$

$$\frac{d^2 (\log y)}{dx\, d\gamma} = \frac{p}{p+1} \frac{1}{(1 + x/a)^2} - 1,$$

$$\frac{d^2 (\log y)}{dp^2} = - \frac{d^2}{dp^2} \log \Gamma (p+1) + \frac{2}{p+1} \frac{1}{1 + x/a} - \frac{p}{(p+1)^2} \frac{1}{(1 + x/a)^2},$$

$$\frac{d^2 (\log y)}{d\gamma^2} = - \frac{1}{\gamma^2} \left(p + 1 - 2p \frac{1}{1 + x/a} + \frac{p}{(1 + x/a)^2}\right),$$

$$\frac{d^2 (\log y)}{dp\, d\gamma} = \frac{1}{\gamma} \left(1 - \frac{2p+1}{p+1} \frac{1}{1 + x/a} + \frac{p}{p+1} \frac{1}{(1 + x/a)^2}\right).$$

Let $I_p = y_1 \int_{-a}^{\infty} \left(1 + \frac{x}{a}\right)^p e^{-\gamma x} dx$, then we easily find $I_{p-1} = \frac{p+1}{p} I_p$, and $I_{p-2} = \frac{(p+1)^2}{p(p-1)} I_p$.

By aid of these we can at once write down the integrals of the above expressions multiplied by y, since $n = I_p$. We find with the notation of p. 243,

$$a_{11} = -\int_{-a}^{\infty} y \frac{d^2 (\log y)}{dx^2} dx = \frac{n\gamma^2}{p-1},$$

$$a_{12} = \int_{-a}^{\infty} y \frac{d^2 (\log y)}{dx\, dp} dx = - \frac{n\gamma}{p(p-1)},$$

$$a_{13} = \int_{-a}^{\infty} y \frac{d^2 (\log y)}{dx\, d\gamma} dx = \frac{2n}{p-1},$$

$$a_{22} = -\int_{-a}^{\infty} y \frac{d^2 (\log y)}{dp^2} dx = n \left(\frac{d^2}{dp^2} \log \Gamma (p+1) - \frac{p-2}{p(p-1)}\right),$$

$$a_{33} = -\int_{-a}^{\infty} y \frac{d^2 (\log y)}{d\gamma^2} dx = \frac{2n(p+1)}{\gamma^2 (p-1)},$$

$$a_{23} = \int_{-a}^{\infty} y \frac{d^2 (\log y)}{d\gamma\, dp} dx = \frac{n(p+1)}{\gamma p (p-1)}.$$

Before we consider the determinant and its minors, we may note that

$$\log \Gamma(p+1) = \log \sqrt{(2\pi)} + (p+\tfrac{1}{2}) \log p - p + \frac{B_1}{1.2.p} - \frac{B_3}{3.4.p^3} + \frac{B^5}{5.6.p^5} - \ldots$$

where the B's are the BERNOULLI numbers. Hence

$$\frac{d^2}{dp^2}\{\log \Gamma(p+1)\} = \frac{2p-1}{2p^2} + \frac{B_1}{p^3} - \frac{B_3}{p^5} + \frac{B_5}{p^7} - ,$$

and we have the convenient form,

$$a_{22} = n\left(\frac{d^2}{dp^2}\log \Gamma(p+1) - \frac{p-2}{p(p-1)}\right) = \frac{n}{p^2}\left\{\frac{1}{2}\frac{p+1}{p-1} + S\right\},$$

where S is the semi-convergent series $B_1/p - B_3/p^3 + B_5/p^5 -$, &c.

Now we have at once,

$$\Delta = \begin{vmatrix} a_{11}, & -a_{12}, & -a_{13}, \\ -a_{12}, & a_{22}, & -a_{23}, \\ -a_{13}, & -a_{23}, & a_{33}. \end{vmatrix} = n^3 \begin{vmatrix} \dfrac{\gamma}{p-1}, & \dfrac{\gamma}{p(p-1)}, & -\dfrac{2}{p-1}, \\ \dfrac{\gamma}{p(p-1)}, & \dfrac{1}{p^2}\left(\dfrac{1}{2}\dfrac{p+1}{p-1} + S\right), & -\dfrac{p+1}{\gamma(p-1)p}, \\ -\dfrac{2}{p-1}, & -\dfrac{p+1}{\gamma(p-1)p}, & \dfrac{2(p+1)}{\gamma^2(p-1)}. \end{vmatrix}$$

Divide all three columns by $1/(p-1)$; the first row and first column by γ, and then the last row and last column by $1/\gamma$; we find,

$$\Delta = \frac{n^3}{(p-1)^3} \begin{vmatrix} 1, & \dfrac{1}{p}, & -2, \\ \dfrac{1}{p}, & \dfrac{p-1}{p^2}\left(\dfrac{1}{2}\dfrac{p+1}{p-1} + S\right), & -\dfrac{p+1}{p}, \\ -2, & -\dfrac{p+1}{p}, & 2(p+1). \end{vmatrix}$$

Divide the second row and column by $1/p$; add half the last row to the second, and we find,

$$\Delta = \frac{n^3}{p^2(p-1)^3} \begin{vmatrix} 1, & 1, & -2 \\ 0, & (p-1)S, & 0 \\ -2, & -(p+1), & 2(p+1) \end{vmatrix}.$$

Subtract the first column from the second, and add the double of it to the last, we deduce

$$\Delta = \frac{n^3}{p^2(p-1)^3} \begin{vmatrix} 1, & 0, & 0 \\ 0, & (p-1)\,\mathrm{S}, & 0 \\ -2, & -(p-1), & 2(p-1) \end{vmatrix},$$

or,

$$\Delta = \frac{2n^3\mathrm{S}}{p^2(p-1)} \quad \ldots \ldots \ldots \quad \text{(xlvii.)}.$$

The minors can now be written down at once.

$$\mathrm{M}_{11} = \frac{2(p+1)n^2\mathrm{S}}{\gamma^2 p^2 (p-1)},$$

whence

$$\Sigma_h^2 = \mathrm{M}_{11}/\Delta = \frac{p+1}{n\gamma^2} = \mu_2 = \sigma^2/n,$$

or

$$\Sigma_h = \sigma/\sqrt{n} \quad \ldots \ldots \ldots \quad \text{(xlviii.)}.$$

$$\mathrm{M}_{22} = \frac{2n^2}{p-1},$$

whence

$$\Sigma_p^2 = \mathrm{M}_{22}/\Delta = p^2/n\mathrm{S},$$

or

$$\Sigma_p = \frac{p}{\sqrt{n}} \sqrt{\frac{1}{\mathrm{S}}} \quad \ldots \ldots \ldots \quad \text{(xlix.)}.$$

$$\mathrm{M}_{33} = \frac{n^2\gamma^2}{p^2(p-1)} (\tfrac{1}{2} + \mathrm{S}),$$

whence

$$\Sigma_\gamma^2 = \mathrm{M}_{33}/\Delta = \frac{\gamma^2}{2n}\left(1 + \frac{1}{2\mathrm{S}}\right),$$

or

$$\Sigma_\gamma = \frac{\gamma}{\sqrt{(2n)}} \sqrt{\left(1 + \frac{1}{2\mathrm{S}}\right)} \quad \ldots \ldots \ldots \quad \text{(l.)}.$$

$$\mathrm{M}_{23} = \frac{\gamma n^2}{p(p-1)},$$

whence

$$\mathrm{R}_{p\gamma} = \mathrm{M}_{23}/(\Delta\Sigma_p\Sigma_\gamma) = \sqrt{\{\tfrac{1}{2}/(\tfrac{1}{2} + \mathrm{S})\}} \quad \ldots \ldots \quad \text{(li.)}.$$

$$\mathrm{M}_{12} = 0,$$

whence

$$\mathrm{R}_{hp} = 0 \quad \ldots \ldots \ldots \quad \text{(lii.)}.$$

Lastly,

$$M_{13} = \frac{2n^2 S}{p^2(p-1)},$$

whence

$$R_{h\gamma} = M_{13}/(\Delta\Sigma_h\Sigma_\gamma) = \sqrt{\left\{\frac{2}{p+1} S/(\tfrac{1}{2} + S)\right\}} \quad . \quad . \quad . \quad . \quad \text{(liii.)}.$$

This completes the direct series of probable errors and error correlations. By aid of the above correlations and standard deviations we can now find a further series.

From (xlii.) we have for the standard deviation σ (about the mean), $\sigma^2 = \mu_2 = \frac{p+1}{\gamma^2}$, or $\sigma = \frac{\sqrt{(p+1)}}{\gamma}$. Hence

$$\frac{\Delta\sigma}{\sigma} = \frac{1}{2}\frac{\Delta p}{p+1} - \frac{\Delta\gamma}{\gamma} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(liv.)}.$$

Square both sides of this, divide by n and sum, we have at once from the definition of a coefficient of correlation

$$\left(\frac{\Sigma_\sigma}{\sigma}\right)^2 = \frac{1}{4}\frac{\Sigma_p^2}{(p+1)^2} + \frac{\Sigma_\gamma^2}{\gamma^2} - \frac{R_{p\gamma}\Sigma_p\Sigma_\gamma}{\gamma(p+1)}.$$

Hence, using (xlix.), (l.), and (li.), we find, after reductions,

$$\Sigma_\sigma = \frac{\sigma}{\sqrt{(2n)}}\left(1 + \frac{1}{2}\frac{1}{(p+1)^2 S}\right)^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lv.)}.$$

Multiply (liv.) by Δh, sum and divide by n, we have

$$\frac{\Sigma_\sigma\Sigma_h R_{h\sigma}}{\sigma} = \frac{1}{2}\frac{\Sigma_p\Sigma_h R_{hp}}{p+1} - \frac{\Sigma_\gamma\Sigma_h R_{h\gamma}}{\gamma}.$$

Whence, by (lii.),

$$R_{h\sigma} = -\frac{\sigma}{\gamma}\frac{\Sigma_\gamma}{\Sigma_\sigma}R_{h\gamma},$$

or, reducing by (l.), (liii.), and (lv.),

$$R_{h\sigma} = -\sqrt{\left(\frac{2}{p+1}\right)}\frac{1}{\sqrt{\left(1 + \frac{1}{2}\frac{1}{(p+1)^2 S}\right)}} \quad . \quad . \quad . \quad . \quad . \quad \text{(lvi.)}.$$

Next, if S_k be the skewness, we have from (xlv.)

$$\Delta S_k = -\frac{1}{2}\frac{\Delta p}{(p+1)^{3/2}},$$

or

$$\Sigma_{S_k} = \tfrac{1}{2}\frac{p}{p+1}\frac{1}{\sqrt{n}}\frac{1}{\sqrt{\{(p+1)S\}}} \quad \ldots \ldots \quad \text{(lvii.)}.^*$$

Similarly

$$R_{S_k h} = \frac{S(\Delta S_k \Delta_h)}{\Sigma_{S_k}\Sigma_h} = 0, \text{ since } R_{hp} = 0,$$

or

$$R_{hS_k} = 0 \quad \ldots \ldots \ldots \quad \text{(lviii.)}.$$

We easily obtain, by multiplying (liv.) by Δp, the result

$$R_{p\sigma} = -\frac{1}{\{1 + 2(p+1)^2 S\}^{\frac{1}{2}}} \quad \ldots \ldots \quad \text{(lix.)}.$$

We can now obtain $R_{\sigma S_k}$, for every Δ_{S_k} is negatively proportional to the corresponding Δp. Hence

$$R_{\sigma S_k} = \frac{1}{\{1 + 2(p+1)^2 S\}^{\frac{1}{2}}} \quad \ldots \ldots \quad \text{(lx.)}.$$

We next pass to the mean and modal frequencies as given by (xli.) and (xlvi.). We have, by taking logarithmic differentials,

$$\frac{\Delta y_1}{y_1} = \frac{\Delta\gamma}{\gamma} - J\Delta p,$$

where

$$J = \frac{1}{p+1} + \frac{d}{dp}\{\log\Gamma(p+1)\} - \log(p+1).$$

* If $S' = \frac{B_1}{p+1} - \frac{B_3}{(p+1)^3} + \frac{B_5}{(p+1)^5} - \&c.$, then it is easy to show that $\frac{1}{2} + p^2 S' = (p+1)^2 S$. Remembering that, if $S_k = q$, $q = \frac{1}{\sqrt{(p+1)}}$, we easily deduce

$$\Sigma_{S_k} = \frac{1}{2\sqrt{n}}\frac{1}{\left\{B_1 - B_3 q^4 + B_5 q^8 + \frac{1}{2}\frac{q^2}{(1-q^2)^2}\right\}^{\frac{1}{2}}}$$

$$= \frac{1}{2\sqrt{n}}\frac{1}{\{B_1 + \frac{1}{2}q^2 + (1-B_3)q^4 + \frac{3}{2}q^6 + (B_5+2)q^8 + \ldots\}^{\frac{1}{2}}}$$

$$= \sqrt{\left(\frac{3}{2n}\right)}\frac{1}{\sqrt{\left\{1 + 3(\text{Sk.})^2 + \frac{29}{5}(\text{Sk.})^4 + 9(\text{Sk.})^6\right\}}},$$

as far the 7th power of the skewness inclusive.

Very generally the probable error of the skewness may be taken as equal to

$$\sqrt{\left(\frac{3}{2n}\right)}\frac{1}{\sqrt{\{1 + 3(\text{Sk.})^2\}}},$$

and it is always less than $\sqrt{(3/2n)}$, its value in the case of a normal frequency.

Squaring, introducing the standard deviations, and rearranging, we find

$$\frac{\Sigma_{y_1}^2}{y_1^2} = \frac{\Sigma_\gamma^2}{\gamma^2}(1 - R_{\gamma p}^2) + \left(J - \frac{\Sigma_\gamma}{\Sigma_p}\frac{R_{\gamma p}}{\gamma}\right)^2 \Sigma_p^2 = \frac{1}{2n}\left\{1 + \frac{2}{S}(Jp - \tfrac{1}{2})^2\right\}.$$

We must now evaluate $Jp - \frac{1}{2}$. This is easily shown from the BERNOULLI number expansion for $\log \Gamma(p+1)$ to be given by

$$Jp - \tfrac{1}{2} = \frac{p}{p+1} - p \log \frac{p+1}{p} - T,$$

where

$$T = \frac{B_1}{2p} - \frac{B_3}{4p^3} + \frac{B_5}{6p^5} - \ldots$$

Thus we determine

$$\Sigma_{y_1} = \frac{y_1}{\sqrt{(2n)}}\left\{1 + \frac{2}{S}\left(p \log \frac{p+1}{p} - \frac{p}{p+1} + T\right)^2\right\}^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad \text{(lxi.)}.$$

Expanding the expression in brackets in inverse powers of p we find

$$\Sigma_{y_1} = \frac{y_1}{\sqrt{(2n)}}\left\{1 + \frac{49}{12p} - \frac{28}{3p^2} + \frac{248}{15p^3} - \right\}^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxii.)}.$$

Result (lxi.), however, with S and T calculated to $1/p^3$, gives a better value than (lxii.).

To find the modal frequency error we must take the logarithmic differential of (xlvi.) and proceed in the same way. We find almost at once

$$\frac{\Delta y_0}{y_0} = \frac{\Delta\gamma}{\gamma} - \frac{\Delta p}{p}(\tfrac{1}{2} - T).$$

Whence on squaring and completing the square of the factor of Σp^2, we find

$$\frac{\Sigma_{y_0}^2}{y_0^2} = \frac{1}{2n}\left\{1 + \frac{2T^2}{S}\right\},$$

and

$$\Sigma_{y_0} = \frac{y_0}{\sqrt{(2n)}}\left\{1 + \frac{2T^2}{S}\right\}^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxiii.}$$

Expanding as far as powers of $1/p^5$ *exclusive* we obtain

$$\Sigma_{y_0} = \frac{y_0}{\sqrt{(2n)}}\left\{1 + \frac{1}{12p}\right\}^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxiv.)},$$

a very simple expression for the probable error of the modal frequency, y_0.

In like manner we shall now determine the probable errors of the moments and their error correlations.

Take the logarithmic differentials of (lxii.)

$$\frac{\Delta\mu_2}{\mu_2} = \frac{\Delta p}{p+1} - \frac{2\Delta\gamma}{\gamma},$$

$$\frac{\Delta\mu_3}{\mu_3} = \frac{\Delta p}{p+1} - \frac{3\Delta\gamma}{\gamma},$$

$$\frac{\Delta\mu_4}{\mu_4} = \frac{2(p+2)\Delta p}{(p+1)(p+3)} - \frac{4\Delta\gamma}{\gamma}.$$

Squaring each of these in succession and using the known values of Σ_p, Σ_γ, $R_{p\gamma}$ we find

$$\Sigma_{\mu_2} = \mu_2 \frac{2}{\sqrt{(2n)}} \sqrt{\left\{1 + \frac{1}{2\,S(p+1)^2}\right\}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{lxv.}),$$

$$\Sigma_{\mu_3} = \mu_3 \frac{3}{\sqrt{(2n)}} \sqrt{\left\{1 + \frac{(p+3)^2}{18\,S(p+1)^2}\right\}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{lxvi.}),$$

$$\Sigma_{\mu_4} = \mu_4 \frac{4}{\sqrt{(2n)}} \sqrt{\left\{1 + \frac{(2p+3)^2}{2\,S(p+1)^2(p+3)^2}\right\}} \quad . \quad . \quad . \quad (\text{lxvii.}),$$

Now multiply $\Delta\mu_2/\mu_2$ by $\Delta\mu_3/\mu_3$ and we find, after some reductions,

$$R_{\mu_2\mu_3} = \frac{1 + \dfrac{p+3}{6\,S(p+1)^2}}{\sqrt{\left\{\left(1 + \dfrac{1}{2\,S(p+1)^2}\right)\left(1 + \dfrac{(p+3)^2}{18\,S(p+1)^2}\right)\right\}}} \quad . \quad . \quad . \quad (\text{lxviii.}).$$

Next multiply $\Delta\mu_2/\mu_2$ by $\Delta\mu_4/\mu_4$, and we ultimately have

$$R_{\mu_2\mu_4} = \frac{1 + \dfrac{2p+3}{2\,S(p+1)^2(p+3)}}{\sqrt{\left\{\left(1 + \dfrac{1}{2\,S(p+1)^2}\right)\left(1 + \dfrac{(2p+3)^2}{2\,S(p+1)^2(p+3)^2}\right)\right\}}} \quad . \quad (\text{lxix.}).$$

Lastly, multiplying $\Delta\mu_3/\mu_3$ by $\Delta\mu_4/\mu_4$, we deduce, after some reductions,

$$R_{\mu_3\mu_4} = \frac{1 + \dfrac{2p+3}{6\,S(p+1)^2}}{\sqrt{\left\{\left(1 + \dfrac{(p+3)^2}{18\,S(p+1)^2}\right)\right\}}\sqrt{\left\{\left(1 + \dfrac{(2p+3)^2}{2\,S(p+1)^2(p+3)^2}\right)\right\}}} \quad . \quad (\text{lxx.}).$$

We may add to these results the values of Σ_{β_1} and Σ_{β_2}, where β_1 and β_2 are given by (xliv.); we find

$$\Sigma_{\beta_1} = \frac{4}{\sqrt{n}} \frac{p}{(p+1)^2 \sqrt{S}}, \qquad \Sigma_{\beta_2} = \frac{6}{\sqrt{n}} \frac{p}{(p+1)^2 \sqrt{S}} \quad . \quad . \quad . \quad \text{(lxxi.)}.$$

The distances from the mode to the mean, d, and from the mean to the end of the range, a, are given by

$$d = 1/\gamma \quad \text{and} \quad a = \frac{p+1}{\gamma}.$$

Hence

$$\Sigma_d = \frac{d}{\sqrt{(2n)}} \sqrt{\left(1 + \frac{1}{2S}\right)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxii.)},$$

$$\Sigma_a = \frac{a}{\sqrt{(2n)}} \sqrt{\left\{1 + \frac{(p-1)^2}{2S(p+1)^2}\right\}} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxiii.)},$$

and further

$$\left.\begin{aligned} R_{dh} &= -\sqrt{\left(\frac{\frac{2}{p+1}S}{\frac{1}{2}+S}\right)} \\ R_{ah} &= -\sqrt{\left(\frac{\frac{2}{p+1}S}{\frac{1}{2}\left(\frac{p-1}{p+1}\right)^2 + S}\right)} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(lxxiv.)}.$$

The results (xlvii.) to (lxxiv.) must be now considered at length.

(17.) (α.) The frequency curve of the type considered is fully described by the three constants, the mean, γ, and p. But, since *any* three constants would do equally well—for example, what may be termed the three physical constants: mean, standard deviation (or variation), and skewness—it becomes of some importance to inquire which constants have the least percentage of probable error.

Now (xlviii.) shows us that the probable error in the mean is precisely the same as in the case of the normal curve and

$$= \cdot 67449 \sigma/\sqrt{n}.$$

Thus, the percentage error in the mean

$$= \cdot 67449 \frac{100\sigma}{h} \frac{1}{\sqrt{n}}$$

$$= \frac{\cdot 67449}{\sqrt{n}} \times \text{coefficient of variation},$$

and will certainly be small whenever the coefficient of variation is small. Its value is quite independent of the order of p.

On the other hand, the percentage probable errors of p and γ are from (xlix.) and (l.)

$$\frac{67 \cdot 449}{\sqrt{n}} \frac{1}{\sqrt{S}} \text{ and } \frac{67 \cdot 449}{\sqrt{(2n)}} \sqrt{\left(1 + \frac{1}{2S}\right)} \text{ respectively.}$$

Here S is equal to the series $B_1/p - B_3/p^3 + B_5/p^5 - \ldots$ which tends to zero as p increases.

The errors in p and γ thus tend to increase indefinitely as p increases. It may then be asked how the form of the curve can be determined with any degree of accuracy. The answer is simple: Equation (li.) shows us that the correlation between errors in p and γ tends, as p increases, to become "perfect," *i.e.*, unity. But as p increases indefinitely, it has been shown that the frequency curve of this type passes over into the normal form.* It is the high correlation between errors in p and γ which renders the curve, when plotted to observations, such an excellent fit. If the errors in p and γ were independent, this would not be so. At the same time it renders p and γ unsuitable for tabulation as physical or biological constants of the frequency.

Turning to (lv.) and (lvii.) we see that the standard deviation, σ, and the skewness, Sk., are suitable constants for tabulation. Their probable errors do not tend to increase indefinitely with p, and will always be small, if n be large.

Hence a frequency distribution of this type is best defined by its mean h, its standard-deviation σ, and its skewness Sk. These are constants characteristic of the group, for they are given with small probable errors. If it be desired to draw the form of the frequency-curve, then its algebraic constants, p and γ, may be found from

$$p = \frac{1}{(\text{Sk.})^2} - 1, \qquad \gamma = \frac{1}{\sigma \times \text{Sk.}},$$

and the possibly considerable errors in p and γ will not vary largely its actual shape.

(β.) The nature of the probable errors of the other allied constants may now be considered. The mean and modal frequencies per unit variation of organ, or y_1 and y_0 are seen by (lxii.) and (lxiv.) to have small percentage probable errors, and are, therefore, good for use as characteristic physical or biological constants. But it should be noted that the modal frequency is considerably more exact for moderate values of p than the mean frequency. For example it would be somewhat better to tabulate the modal than the mean frequency of the barometer as a physical characteristic of climate.

The probable errors of the distances from the mean to the mode and from the mean to the terminal of the range are given by (lxxii.) and (lxxiii.). Since $d = 1/\gamma = \sigma/\sqrt{(1 + p)}$, we may write the first

$$\Sigma_d = \frac{\sigma}{\sqrt{(2n)}} \sqrt{\left(\frac{1}{1 + p} + \frac{1}{2(1 + p)S}\right)}.$$

* 'Phil. Trans.,' A, vol. 186, p. 374.

This remains finite, even if p be indefinitely great. On the other hand, the probable error of a, and even its percentage probable error, becomes indefinitely great with p. It is to be noted that a in this case becomes infinite.

(γ.) Results (lxv.) to (lxvii.) give the probable errors of the second, third, and fourth moments. It will be seen that roughly, for a large p, the percentage error of the fourth moment is about double that of the second. It might thus appear, at first sight, safer to work with the second than with the fourth, but this is by no means necessarily the case, for to deduce any quantity from one or the other they *must be reduced to the same order.* For example, the square root of μ_2 must be compared with the fourth root of μ_4, and the probable errors of $\sqrt{\mu_2}$ and $(\mu_4)^{1/4}$ will be sensibly of the same order.

Remembering that $\mu_3 = \dfrac{2(p+1)}{\gamma^3} = \dfrac{2\sigma^3}{\sqrt{(p+1)}}$, we may write

$$\Sigma_{\mu_3} = \frac{6\sigma^3}{\sqrt{(2n)}} \sqrt{\left(\frac{1}{p+1} + \frac{(p+3)^2}{18\,\mathrm{S}(p+1)^3}\right)}.$$

This tends to a finite limit as p increases indefinitely, and we conclude that the probable error of μ_3 is always finite, and will in general be a small fraction of the cube of the standard deviation. The above remarks are a justification for the use of higher moments in frequency calculations.

Equations (lxviii.) to (lxx.) give the error-correlations between the first three moments. They show that an error in the value of one of these moments will most probably lead to an error in the other two. We see that for p fairly large $\mathrm{R}_{\mu_2\mu_4}$ is a large correlation, while $\mathrm{R}_{\mu_2\mu_3}$ and $\mathrm{R}_{\mu_3\mu_4}$ are small. In other words a random selection of an even moment makes a far larger correlated change in another even moment than in an odd moment. If p increase indefinitely we find the ratio $\mathrm{R}_{\mu_2\mu_3} / \mathrm{R}_{\mu_3\mu_4}$ approaches the value 2/3; in other words, μ_3 is more closely correlated to the higher moment μ_4 than to the lower moment μ_2.

Formulæ (lxxi.) give the probable errors of the useful constants $\beta_1 = \mu_3^2/\mu_2^3$ and $\beta_2 = \mu_4/\mu_2^2$. We see that they are small and approach the value zero as p is indefinitely increased.

(δ.) Let us restate the formulæ for p indefinitely great, *i.e.*, for the normal curve of frequency

$$y = \frac{n}{2\pi\sqrt{\sigma}}\, e^{-x^2/(2\sigma^2)}$$

In this case we have $\mu_2 = \sigma^2$, $\mu_3 = 0$, $\mu_4 = 3\sigma^4$, $\beta_1 = 0$, $\beta_2 = 3$, skewness $= 0$, mean and mode coincide. Several of these zero quantities, however, tend to have definite probable errors.

We have

$$\Sigma_p = \sigma/\sqrt{n},$$
$$\Sigma_\sigma = \sigma/\sqrt{(2n)},$$
$$\Sigma_{sk.} = \sqrt{\left(\frac{3}{2n}\right)},$$
$$\Sigma_{y_1} = \frac{y_1}{\sqrt{(2n)}},$$
$$\Sigma_{\mu_2} = \frac{2\sigma^2}{\sqrt{(2n)}},$$
$$\Sigma_{\mu_3} = 2\sqrt{\ }\sigma^3 = 2\sigma^3\,\Sigma_{sk},$$
$$\Sigma_{\mu_4} = \frac{12\sigma^4}{\sqrt{(2n)}},$$
$$R_{\mu_2\mu_3} = 0,$$
$$R_{\mu_2\mu_4} = 1,$$
$$R_{\mu_3\mu_3} = 0,$$
$$\Sigma_d = \sqrt{\left(\frac{3}{2n}\right)}\,\sigma = \sigma\Sigma_{sk}.$$

The first, second, fourth and fifth of these results are old; the rest appear to be novel and of some importance.

In the first place we notice that given a population which is really normal, we should not expect a random selection to exhibit all the signs of normality. Its skewness will differ from zero with a probable error of $\cdot67449\sqrt{\left(\frac{3}{2n}\right)}$. For example, in a random selection of 600 from a normal population, the skewness will be as likely to exceed as to fall short of ·034. Hence an exhibition of skewness of less than once to twice $\cdot67449\sqrt{\left(\frac{3}{2n}\right)}$ must not in itself merely be taken to indicate an absence of normality in a general population.

Again, in a random selection from a general population, the mode will differ from the mean, even if the population be normal, with a probable error of $\cdot67449\sqrt{\left(\frac{3}{2n}\right)}\,\sigma$. Thus, in a population of 600, a difference between the mean and the mode of $\cdot034\sigma$ should not be taken to indicate want of normality. Generally, the divergence between mean and mode in a population must at least exceed once to twice $\cdot67449\sqrt{\left(\frac{3}{2n}\right)}\,\sigma$, for us to be able to argue on this ground alone that the population has not a normal distribution.

Again, the third moment not being zero, but having a value of once or twice $\cdot67449 \times 2\sqrt{\left(\frac{3}{2n}\right)}\,\sigma^3$, is not in itself an argument for skew frequency.

The above statements are an important addition to the second memoir of this series; they give us the criterion, there wanting, to distinguish between a skewness which is characteristic of a population and one which might arise by the random selection of a population of the given size out of a larger, but really normal, population.

(ϵ.) We may now note the exceedingly interesting conclusions which these results have for the theory of evolution.

Suppose an organ to have, as so many do, skew variation, then we notice

(i.) Any selection of the organ by size tends to alter its variability, but not its skewness; this follows from (lvi.) and (lviii.). Further, if, as we have supposed, the range be limited on the side of dwarf organs, then any increase of size means a decrease of variability, and *vice versâ*.

(ii.) Any selection of variability is a selection of skewness; this follows from (lx.). If a selection be made from a general population, which has less variability, then it will tend to greater normality. In other words, it would appear that stringent selection tends to generate normal distribution. Thus, if out of a skewly distributed population we make a number of random selections, that with the least variability will be most normal. Select at random again out of this latter selection, and the least variable group will again be the most normal, and so on.

Now take a problem of this kind involving *group*, and not *individual*, selection. Let a large general population break itself up at random into groups, and let us suppose these groups, not individuals among them, to carry on a struggle for existence—an inter-group, not an intra-group, struggle. Then, if it be an advantage to a group that its members shall be among themselves close to a type, *i.e.*, less variable, then the more normal groups will survive, for variability is positively correlated with skewness. Now suppose each group to be periodically subdivided at random into new groups—the mathematical description of some process of group reproduction—then we see how normal distribution may be a result of a stringent inter-group selection of groups whose individuals have the closest resemblance to each other—intra-group resemblance.

(iii.) Any selection of the size of an organ produces by (lxxiv.) an alteration in the distances between the mean and the mode, and between the mean and the end of the range.

A random selection which has its mean larger than that of the general population, will, if the mode be on the dwarf side of the mean, tend to have its mode and mean nearer together than are the mode and mean of the general population, while on the other hand, to raise the mean is to raise the dwarf limit to the range.

A considerable number of like results might be stated, but the above will be sufficient to emphasize the general principle that a random and *à fortiori* an artificial selection of the size of an organ, does, whenever its distribution is skew, influence in a definite manner the variability of the organ. It is quite safe to assert that it will also influence the correlation of organs. When we notice how wide-spread is skew variation in nature, we may assert that the general rule is that no modification can be made in any of the features—mean sizes, variabilities and correlations—of a group of organs without at the same time modifying all the others.*

* A paper has recently been published by Messrs. DAVENPORT and BULLARD in the 'Proceedings of the American Academy of Science' (see Illustration II. below) on "The Variation and Correlation of the Glands in the Legs of Swine." Unfortunately the authors have overlooked the markedly skew character

As a result of Articles (15) and (16), it is possible to use the frequency curve of type $y = y_1 (1 + x/a)^p e^{-\gamma_1 x}$ with as much certainty as to the nature and magnitude of the errors made in the constants as has hitherto been possible in the case of the normal distribution $y = \frac{n}{2\pi\sqrt{\sigma}} e^{-x^2/(2\sigma^2)}$. The method has been exemplified numerically in twenty-three cases in a memoir on the "Variation of Barometric Frequency" (see 'Phil. Trans.,' A, vol. 190, p. 423). It may not, however, be amiss to illustrate it further in a special case having closer bearings on the theory of evolution.

(18.) *Numerical Illustration—Incidence of Enteric Fever.*

In a memoir in the 'Phil. Trans.,' A, vol. 186, p. 391, it is shown that the curve

$$y = 1894{\cdot}57 \left(1 + \frac{x}{3{\cdot}428094}\right)^{3{\cdot}673,042} e^{-1{\cdot}071453x}$$

closely represents the distribution with age of 8,689 cases of enteric fever received into the Metropolitan Asylums Board Fever Hospitals. The unit of x is five years, and the origin is the mode at 14·3025 years. The criterion is not very nearly zero, although small, but the curve is graphically a good fit (see Plate 12, fig. 9).

The following are the numerical values of all the constants :—

Mean = 18·9691 years. d = mean-mode.
Mode = 14·3025 years. = ·933,313 unit.
Sk = skewness = ·462,594.
γ = 1·071,453. p = 3·673,042.
a = 3·428,094.
y_0 = modal frequency = 1894·57.
y_1 = mean frequency = 1687·80.
σ = standard deviation = 2·01756 units = 10·0878 years.

From these the numerical values of the probable errors and of the correlations between the errors of the constants were found by the processes indicated and the formulæ given above.

We found

$$T = {\cdot}022{,}525, \qquad S = {\cdot}044{,}735,$$

whence

of the distribution. It is, however, clear from their tables and plate that no selection could be made of the absolute number of glands without altering the variability of the gland distribution and the correlation between different systems of glands.

	Probable Error.	Percentage Probable Error.
p	·125659	3·4211
γ	·019130	1·7854
Correlation of errors in p and γ =		·9581
a	·061202	1·7853
y_0	9·8029	·5174
y_1	23·5465	1·3951
mean =	·014600 = ·073 year.	
mode =	·024126 = ·121 year.	

These are the constants which determine the position and algebraical equation to the frequency curve, and we see at once that they are all determined with a close degree of accuracy. The largest percentage probable error is in p, but this is under 3·5 per cent., and, owing to the high correlation between p and γ a much larger error would produce no sensible change in the shape of the curve.

Two important facts may also be drawn from these results, which indeed follow from the general formulæ, namely :

(i.) The position of the mean is sensibly more exactly determined than the position of the mode. Here about 1·7 times as accurately.

(ii.) The modal frequency, on the other hand, is sensibly more accurate than the mean frequency. Here about 2·8 times as accurate.

Hence the advantage of using the mean as origin of measurement for the curve is accompanied by the counterbalancing, and here relatively greater, disadvantage of the increased inaccuracy of determination of the mean frequency.

Passing to the "physical" constants of the curve, we have

	Probable Error.	Percentage Probable Error.
σ	·012693	·6291
Sk.	·022845	1·3445
d	·016663	1·7854

These fully determine the non-symmetrical nature and spread of the curve, and since the errors in the skewness and in the distance between the mean and mode are less than 1·4 and 1·8 per cent. of the respective values of these quantities, we conclude that skewness and divergence between mode and mean are characteristic features of enteric fever distribution, and not mere anomalies due to a random selection of cases. They are significant constants peculiar to each type of fever distribution and no description of such a distribution is sufficient unless their values are stated.

Before giving a table of the correlations between what we have termed the "physical" constants, it may be well to write down some of the correlations between the errors in the physical and algebraical constants, which arise in the course of their calculation. We find

$$R_{p\gamma} = \cdot9581, \qquad R_{pm} = 0,$$
$$R_{m\gamma} = \cdot1875, \qquad R_{d\gamma} = -1.$$

By aid of these we find the following table of error correlations :—

	Mean.	y_0.	σ.	d.	Sk.
Mean.	1	·6469	−·5321	−·1875	0
y_0.	·6469	1	−·8908	−·4260	−·1489
σ.	−·5321	−·8908	1	·7905	·1584
d.	−·1875	−·4260	·7905	1	·9592
Sk.	0	−·1489	·1584	·9592	1

Now this table enables us to draw some remarkable conclusions with regard to enteric fever. We see at once that no random selection of a group of individuals, which has any single characteristic differing from that of the general population will, except in the case of mean age of incidence and skewness, leave the other characteristics unmodified. Thus the most probable result of any selection which alters the nature of the distribution of enteric fever can be predicted. The reader will possibly appreciate this better, if we replace the above table by another giving the absolute progressions in years, number of cases per thousand, &c.

PROGRESSION TABLE.

Corresponds to a probable change in the same units of	Unit change of				
	One year in mean age of incidence.	One case per cent. in modal year of frequency.*	One year in standard deviation or "spread."	One year in number of years between modal and mean incidence.	A unit of 1/10 in the skewness.
Mean age of incidence	1	·0909	− ·6120	− ·1643	0
Modal frequency . .	4·5852	1	−7·2626	−2·6456	− ·3372
"Spread"	− ·4626	− ·1093	1	·6022	·0440
Interval between mode and mean	− ·2140	− ·0686	1·0377	1	·3498
Skewness	0	− ·0657	·5702	2·6301	1

* The frequency of incidence in the modal year $= y_0 \times \frac{1}{5}$ since the unit is five years $= 1894\cdot57 \times \frac{1}{5}$ To make this 1000 we must multiply by $\dfrac{1000 \times 5}{1894\cdot57}$. Similarly $\Delta y_0 \times \frac{1}{5} =$ error in incidence of modal year. Thus we have to replace Δy_0 by $\dfrac{1894\cdot57}{1000}\left(\dfrac{\frac{1}{5}\Delta y_0}{\frac{1}{5}y_0}1000\right) = \dfrac{1894\cdot57}{1000}$ (error per thousand in modal year of incidence).

We see at once from the above table that if the mean age of incidence of enteric fever in any group were raised, the disease would be concentrated in fewer years, the modal and mean incidence would be brought closer together, and the incidence in the modal year of frequency would be heavier. The changes here are very sensible. Thus, if we raised the mean age of attack to that of phthisis, or about nine years, the modal frequency would be increased about 41 per cent., the concentration of the incidence of the fever increased about 40 per cent., while the distance between mode and mean would be reduced to nearly 2/5 of its original value. The skewness would not be changed. Much less marked effects would arise from a selection of modal frequency. Any increase of modal frequency tends to slightly raise the mean age of attack, to increase slightly the concentration, to draw the mode towards the mean and reduce the skewness.

The changes produced by closer concentration of the attacks of the disease, *i.e.*, the limitation of its incidence to fewer years, would be of a more marked character, they would raise the mean age of attack and the modal frequency, they would decrease the interval between mode and mean, and reduce the skewness. Concentration of the disease would thus tend to render its distribution more normal.

To increase the interval between mean and mode lowers the mean age of attack, reduces the modal frequency, increases the period of liability to incidence, and much increases the skewness.

Finally, increase of skewness decreases the modal frequency, increases the period of liability and the interval between mean and mode.

These statements with regard to the manner in which enteric fever would affect different groups selected at random from the general population seem of considerable interest, for there is reason to believe that what is thus stated for enteric fever in different groups may be applied to different fevers in one and the same group. For example, the lower the mean age of attack of any fever, the greater its concentration; the less the concentration, the more nearly normal is its distribution, &c., &c.

(19.) *Probable Errors and Error Correlations of the Constants of the Generalised Probability Curve of the Type* $y = y_1\left(1 + \frac{x}{a_1}\right)^{m_1}\left(1 - \frac{x}{a_2}\right)^{m_2}$.

Transfer the origin to one end of the range, and the equation to the curve becomes

$$y = \frac{n}{b}\frac{\Gamma(m_1 + m_2 + 2)}{\Gamma(m_1 + 1)\,\Gamma(m_2 + 1)}\left(\frac{x}{b}\right)^{m_1}\left(1 - \frac{x}{b}\right)^{m_2} \quad . \quad . \quad . \quad . \quad \text{(lxxv.)},$$

where n is the number of observations and b is the range.

The following values are given in 'Phil. Trans.,' A, vol. 186, pp. 368–9, where $r = m_1 + m_2 + 2$:

$$\text{Mean } \bar{x} = \frac{b(m_1+1)}{m_1+m_2+2} \quad \ldots \ldots \ldots \quad \text{(lxxvi.)},$$

$$\text{Mode } x_1 = \frac{m_1 b}{m_1+m_2} \quad \ldots \ldots \ldots \quad \text{(lxxvii.)},$$

$$d = \bar{x} - x_1 = \frac{b(m_1-m_2)}{(m_1+m_2)(m_1+m_2+2)} \quad \ldots \ldots \quad \text{(lxxviii.)},$$

$$\sigma = \frac{b}{r}\sqrt{\left\{\frac{(m_1+1)(m_2+1)}{r+1}\right\}} \quad \ldots \ldots \quad \text{(lxxix.)},$$

$$\text{Skewness} = S_k = \frac{m_1-m_2}{m_1+m_2}\sqrt{\left\{\frac{m_1+m_2+3}{(m_1+1)(m_2+1)}\right\}} \quad \ldots \quad \text{(lxxx.)}.$$

For the moments we have

$$\mu_2 = \frac{b^2(m_1+1)(m_2+1)}{r^2(r+1)} \quad \ldots \ldots \ldots \quad \text{(lxxxi.)},$$

$$\mu_3 = \frac{2b^3(m_1+1)(m_2+1)(m_2-m_1)}{r^3(r+1)(r+2)} \quad \ldots \ldots \ldots \quad \text{(lxxxii.)},$$

$$\mu_4 = \frac{3b^4(m_1+1)(m_2+1)\{(m_1+1)(m_2+1)(r-6)+2r^2\}}{r^4(r+1)(r+2)(r+3)} \quad \ldots \quad \text{(lxxxiii.)}.$$

Lastly, for the mean and modal frequencies of $y_1\,\delta x$ and $y_2\,\delta x$, we have

$$y_1 = \frac{n}{b}\frac{(m_1+m_2+2)^{3/2}}{\sqrt{\{2\pi(m_1+1)(m_2+1)\}}}\text{expt.}\{\bar{S}(m_1+m_2+2)-\bar{S}(m_1+1)-\bar{S}(m_2+1)\} \quad \text{(lxxxiv.)},$$

$$y_2 = \frac{n}{b}\frac{(m_1+m_2+1)\sqrt{(m_1+m_2)}}{\sqrt{(2\pi m_1 m_2)}}\text{expt.}\{\bar{S}(m_1+m_2)-\bar{S}(m_1)-\bar{S}(m_2)\} \quad \ldots \quad \text{(lxxxv.)}.$$

where

$$\bar{S}(p) = \frac{B_1}{1.2p} - \frac{B_3}{3.4p^3} + \frac{B_5}{5.6.p^5} - \text{etc.} \quad \ldots \ldots \quad \text{(lxxxvi.)}.$$

Let

$$I(m_1, m_2) = y_0\int_0^b\left(\frac{x}{b}\right)^{m_1}\left(1-\frac{x}{b}\right)^{m_2}dx$$

where

$$y_0 = \frac{n}{b}\frac{\Gamma(m_1+m_2+2)}{\Gamma(m_1+1)\Gamma(m_2+1)},$$

then $I(m_1, m_2) = n$, and we easily find, by the fundamental property of Γ functions, that

$$\left.\begin{aligned} \mathrm{I}\,(m_1 - 1,\, m_2) &= \frac{m_1 + m_2 + 1}{m_1}\, n, \\ \mathrm{I}\,(m_1,\, m_2 - 1) &= \frac{m_1 + m_2 + 1}{m_2}\, n, \\ \mathrm{I}\,(m_1 - 2,\, m_2) &= \frac{(m_1 + m_2)\,(m_1 + m_2 + 1)}{m_1\,(m_1 - 1)}\, n, \\ \mathrm{I}\,(m_1,\, m_2 - 2) &= \frac{(m_1 + m_2)\,(m_1 + m_2 + 1)}{m_2\,(m_2 - 1)}\, n, \end{aligned}\right\} \quad \text{(lxxxvii.).}$$

From (i.) we have

$$\log y = \log n - \log b + \log \chi + m_1 \log \frac{x}{b} + m_2 \log \left(1 - \frac{x}{b}\right) \quad \text{(lxxxviii.),}$$

where

$$\chi = \frac{\Gamma\,(m_1 + m_2 + 2)}{\Gamma\,(m_1 + 1)\,\Gamma\,(m_2 + 1)}.$$

It will be needful to find $\frac{d^2\,(\log \chi)}{dm_1^2}$, $\frac{d^2\,(\log \chi)}{dm_2^2}$, and $\frac{d^2\,(\log \chi)}{dm_1\,dm_2}$.

$$\begin{aligned} \frac{d^2\,(\log \chi)}{dm_1^2} &= \frac{d^2\,\Gamma\,(m_1 + m_2 + 2)}{dm_1^2} - \frac{d^2\,\Gamma\,(m_1 + 1)}{dm_1^2} \\ &= \frac{d^2\,\Gamma\,(m_1 + m_2 + 2)}{d\,(m_1 + m_2 + 1)^2} - \frac{d^2\,\Gamma\,(m_1 + 1)}{dm_1^2} = \epsilon_3 - \epsilon_1 \quad \text{(lxxxix.).} \end{aligned}$$

Similarly

$$d^2\,(\log \chi)/dm_2^2 = \epsilon_3 - \epsilon_2 \quad \text{(xc.),}$$

$$d^2\,(\log \chi)/dm_1\,dm_2 = \epsilon_3 \quad \text{(xci.),}$$

where

$$\left.\begin{aligned} \epsilon_1 = d^2\,\Gamma\,(m_1 + 1)/dm_1^2, \qquad \epsilon_2 = d^2\,\Gamma\,(m_2 + 1)/dm_2^2, \\ \epsilon_3 = d^2\,\Gamma\,(m_1 + m_2 + 2)/d\,(m_1 + m_2 + 1)^2 \end{aligned}\right\} \quad \text{(xcii.).}$$

ϵ_1, ϵ_2, and ϵ_3 can now be readily expressed in semi-convergent series admitting of calculation.

$$\left.\begin{aligned} \epsilon_1 &= \frac{2m_1 - 1}{2m_1^2} + \frac{\mathrm{S}\,(m_1)}{m_1^2} \\ \epsilon_2 &= \frac{2m_2 - 1}{2m_2^2} + \frac{\mathrm{S}\,(m_2)}{m_2^2} \\ \epsilon_3 &= \frac{2\,(m_1 + m_2 + 1) - 1}{2\,(m_1 + m_2 + 1)^2} + \frac{\mathrm{S}\,(m_1 + m_2 + 1)}{(m_1 + m_2 + 1)^2} \end{aligned}\right\} \quad \text{(xciii.),}$$

where

$$\mathrm{S}\,(p) = \mathrm{B}_1/p - \mathrm{B}_3/p^3 + \mathrm{B}_5/p^5 -$$

and

$$\frac{\mathrm{S}\,(p)}{p^2} = \frac{d^2}{dp^2}\,\{\overline{\mathrm{S}}\,(p)\} \quad \text{(xciv.).}$$

It is clear that if m_1 and m_2 be at all large, which they frequently will be, we may omit the series S, or even reduce ϵ_1, ϵ_2, ϵ_3 to $1/m_1$, $1/m_2$, and $1/(m_1 + m_2 + 1)$, respectively.

Making use of (lxxxvii.) and (lxxxviii.) we easily find

$$a_{11} = nb_{11} = -\int_0^b y \frac{d^2(\log y)}{dx^2} dx = \frac{n}{b^2}(m_1 + m_2)(m_1 + m_2 + 1)\left(\frac{1}{m_1 - 1} + \frac{1}{m_2 - 1}\right) \quad \text{(xcv.)}.$$

$$a_{12} = nb_{12} = -\int_0^b y \frac{d^2(\log y)}{dx\, db} dx = -\frac{n}{b^2}\frac{(m_1 + m_2)(m_1 + m_2 + 1)}{m_2 - 1} \quad \text{(xcvi.)}.$$

$$a_{13} = nb_{13} = -\int_0^b y \frac{d^2(\log y)}{dx\, dm_1} dx = -\frac{n}{b}\frac{m_1 + m_2 + 1}{m_1} \quad \text{(xcvii.)}.$$

$$a_{14} = nb_{14} = -\int_0^b y \frac{d^2(\log y)}{dx\, dm_2} dx = \frac{n}{b}\frac{(m_1 + m_2 + 1)}{m_2} \quad \text{(xcviii.)}.$$

$$a_{22} = nb_{22} = -\int_0^b y \frac{d^2(\log y)}{db^2} dx = \frac{n}{b^2}(m_1 + m_2 + 1)\frac{m_1 + 1}{m_2 - 1} \quad \text{(xcix.)}.$$

$$a_{23} = nb_{23} = -\int_0^b y \frac{d^2(\log y)}{db\, dm_1} dx = \frac{n}{b} \quad \text{(c.)}.$$

$$a_{24} = nb_{24} = -\int_0^b y \frac{d^2(\log y)}{db\, dm_2} dx = -\frac{n}{b}\frac{m_1 + 1}{m_2} \quad \text{(ci.)}.$$

$$a_{33} = nb_{33} = -\int_0^b y \frac{d^2(\log y)}{dm_1^2} dx = n(\epsilon_1 - \epsilon_3) \quad \text{(cii.)}.$$

$$a_{34} = nb_{34} = -\int_0^b y \frac{d^2(\log y)}{dm_1\, dm_2} dx = -n\epsilon_3 \quad \text{(ciii.)}.$$

$$a_{44} = nb_{44} = -\int_0^b y \frac{d^2(\log y)}{dm_2^2} = n(\epsilon_2 - \epsilon_3) \quad \text{(civ.)}.$$

The next stage is to calculate the determinant

$$\Delta' = \begin{vmatrix} b_{11}, & b_{12}, & b_{13}, & b_{14} \\ b_{12}, & b_{22}, & b_{23}, & b_{24} \\ b_{13}, & b_{23}, & b_{33}, & b_{34} \\ b_{14}, & b_{24}, & b_{34}, & b_{44} \end{vmatrix}$$

and the minors $B_{ss'}$, &c., of $b_{ss'}$, &c. We shall then determine

$$\Sigma_h^2 = \frac{1}{n}\frac{B_{11}}{\Delta'}, \quad \Sigma_b^2 = \frac{1}{n}\frac{B_{22}}{\Delta'}, \quad \Sigma_{m_1}^2 = \frac{1}{n}\frac{B_{33}}{\Delta'}, \quad \Sigma_{m_2}^2 = \frac{1}{n}\frac{B_{44}}{\Delta'} \quad \text{(cv.)},$$

and the correlations

$$\left.\begin{aligned} R_{hb} &= \frac{B_{12}}{\sqrt{(B_{11}B_{22})}}, \quad R_{hm_1} = \frac{B_{13}}{\sqrt{(B_{11}B_{33})}}, \quad R_{hm_2} = \frac{B_{14}}{\sqrt{(B_{11}B_{44})}} \\ R_{bm_1} &= \frac{B_{23}}{\sqrt{(B_{22}B_{33})}}, \quad R_{bm_2} = \frac{B_{24}}{\sqrt{(B_{22}B_{44})}}, \quad R_{m_1m_2} = \frac{B_{34}}{\sqrt{(B_{33}B_{44})}} \end{aligned}\right\} \quad . \quad . \quad . \quad \text{(cvi.)}.$$

The algebraic expressions for the expanded determinant and its minors are very lengthy, and it will be found easiest in any numerical case to calculate the numerical values of the b_{11}, b_{12}, b_{13}, b_{14}, b_{22} . . . , and then find the values of the determinant and its minors numerically. So soon as the above four standard deviations and six error correlations have been calculated, the determination of the probable errors in the fundamental constants of the frequency distribution becomes easy.

We have for the mean organ M_e, if h now defines the origin of coordinates,

$$\begin{aligned} \Delta M_e &= \Delta h + \Delta x \\ &= \Delta h + \frac{m_1 + 1}{m_1 + m_2 + 2}\Delta b + \frac{b(m_2+1)\Delta m_1}{(m_1+m_2+2)^2} - \frac{b(m_1+1)\Delta m_2}{(m_1+m_2+2)^2} \quad . \quad . \quad \text{(cvii.)}. \end{aligned}$$

Whence

$$\begin{aligned} \Sigma^2_{M_e} = \Sigma^2_h &+ \left(\frac{m_1+1}{r}\right)^2 \Sigma^2_b + \frac{b^2(m_2+1)^2}{r^4}\Sigma^2_{m_1} + \frac{b^2(m_1+1)^2}{r^4}\Sigma^2_{m_2} \\ &+ 2\,\frac{m_1+1}{r}\,\Sigma_h\Sigma_b R_{hb} + \frac{2b(m_2+1)}{r^2}\Sigma_h\Sigma_{m_1}R_{hm_1} - \frac{2b(m_1+1)}{r^2}\Sigma_h\Sigma_{m_2}R_{hm_2} \\ &+ \frac{2b(m_1+1)(m_2+1)}{r^3}\Sigma_b\Sigma_{m_1}R_{bm_1} - \frac{2b(m_1+1)}{r^3}\Sigma_b\Sigma_{m_2}R_{bm_2} \\ &- \frac{2b^2(m_1+1)(m_2+1)}{r^4}\Sigma_{m_1}\Sigma_{m_2}R_{m_1m_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(cviii.)}. \end{aligned}$$

Similarly, the modal value of the organ can be found from*

$$\Delta M_0 = \Delta h + \frac{m_1}{m_1+m_2}\Delta b + \frac{bm_2}{(m_1+m_2)_2}\Delta m_1 - \frac{bm_1}{(m_1+m_2)^2}\Delta m_2 \quad . \quad \text{(cix.)}.$$

* The easiest numerical method in this, as in the previous case, is to proceed as follows:—Write

$$\Sigma_{M_0}\chi_{M_0} = \Sigma_h\chi_h + \frac{m_1}{m_1+m_2}\Sigma_b\chi_b + \frac{bm_2}{(m_1+m_2)^2}\Sigma_{m_1}\chi_{m_1} - \frac{bm_1}{(m_1+m_2)^2}\Sigma_{m_2}\chi_{m_2},$$

where the χ's are umbral symbols, and let N_1, N_2, N_3, N_4 be the *numerical* values of the coefficients on the right. Then put

$$\Sigma_{M_0}\chi_{M_0} = N_1\chi_h + N_2\chi_b + N_3\chi_{m_1} - N_4\chi_{m_2}.$$

Now square this equation, and, whenever a product, $\chi_q\chi_{q'}$, occurs, multiply it by the corresponding error correlation, $R_{qq'}$, already calculated, putting $R_{qq'} = 1$ if $q = q'$. Then, actually,

$$\begin{aligned} \Sigma^2_{M_0} = N_1^2 + N_2^2 + N_3^2 + N_4^2 + 2N_1N_2R_{hb} + 2N_1N_3R_{hm_1} - 2N_1N_4R_{hm_2} + 2N_2N_3R_{bm_1} \\ - 2N_1N_4R_{bm_2} - 2N_3N_3R_{m_1m_2}. \end{aligned}$$

BARLOW's Tables rapidly give the squares and CRELLE's Tables, or a Brunsviga, the products.

In like manner from (lxxviii.)

$$\frac{\Delta d}{d} = \frac{\Delta b}{b} + \Delta m_1 \left\{ \frac{1}{m_1 - m_2} - \frac{1}{m_1 + m_2} - \frac{1}{m_1 + m_2 + 2} \right\}$$
$$+ \Delta m_2 \left\{ - \frac{1}{m_1 - m_2} - \frac{1}{m_1 + m_2} - \frac{1}{m_1 + m_2 + 2} \right\}$$
$$= \Delta b/b + c_1 \Delta m_1 + c_2 \Delta m_2,$$

say, where the numerical values of c_1 and c_2 can easily be found in any actual case. Hence

$$\Sigma^2_d/d^2 = \Sigma^2_b/b^2 + c_1^2\Sigma^2_{m_1} + c_2^2\Sigma^2_{m_2} + 2c_1c_2\Sigma_{m_1}\Sigma_{m_2}\mathrm{R}_{m_1m_2} + \frac{2c_1}{b}\Sigma_b\Sigma_{m_1}\mathrm{R}_{bm_1} + \frac{2c_2}{b}\Sigma_b\Sigma_{m_2}\mathrm{R}_{bm_2} \quad \text{(cx.)}.$$

Again, from (lxxix.)

$$\frac{\Delta\sigma}{\sigma} = \frac{\Delta b}{b} + \Delta m_1 \left(- \frac{1}{r} - \frac{1}{2(r+1)} + \frac{1}{2(m_1+1)} \right) + \Delta m_2 \left(- \frac{1}{r} - \frac{1}{2(r+1)} + \frac{1}{2(m_2+1)} \right).$$

Or $\quad \Delta\sigma/\sigma = \Delta b/b + e_1 \Delta m_1 + e_2 \Delta m_2,$

$$\frac{\Sigma^2_\sigma}{\sigma^2} = \frac{\Sigma^2_b}{b^2} + e_1^2\Sigma^2_{m_1} + e_2^2\Sigma^2_{m_2} + 2e_1e_2\Sigma_{m_1}\Sigma_{m_2}\mathrm{R}_{m_1m_2} + \frac{2e_1}{b}\Sigma_b\Sigma_{m_1}\mathrm{R}_{bm_1} + \frac{2e_2}{b}\Sigma_b\Sigma_{m_2}\mathrm{R}_{bm_2} \quad \text{(cxi.)}.$$

Further from (lxxx.)

$$\frac{\Delta \mathrm{S}_k}{\mathrm{S}_k} = \Delta m_1 \left\{ \frac{1}{m_1 - m_2} - \frac{1}{m_1 + m_2} + \frac{1}{2(m_1 + m_2 + 3)} - \frac{1}{2(m_1 + 1)} \right\}$$
$$+ \Delta m_2 \left\{ - \frac{1}{m_1 - m_2} - \frac{1}{m_1 + m_2} + \frac{1}{2(m_1 + m_2 + 3)} - \frac{1}{2(m_2 + 1)} \right\}$$
$$= f_1 \Delta m_1 + f_2 \Delta m_2, \text{ say.}$$

Hence

$$\Sigma^2_{\mathrm{S}_k} = \mathrm{S}_k^2 \{ f_1^2 \Sigma^2_{m_1} + f_2^2 \Sigma^2_{m_2} + 2f_1f_2 \Sigma_{m_1}\Sigma_{m_2}\mathrm{R}_{m_1m_2} \} \quad \text{. . . . (cxii.)}.$$

From the results given above we can deduce the effects on size, range, variability, or skewness of a selection at random of any one of these four.

Writing (cvii.)

$$\Delta\mathrm{M}_e = \Delta h + g_1 \Delta b + g_2 \Delta m_1 - g_3 \Delta m_2$$

we find

$$\mathrm{R}_{\mathrm{M}_e\sigma} = \frac{\sigma}{\Sigma_{\mathrm{M}_e}\Sigma_\sigma} \left\{ \frac{\Sigma_b\Sigma_h\mathrm{R}_{bh}}{b} + \frac{g_1}{b}\Sigma^2_b + \frac{g_2}{b}\Sigma_b\Sigma_{m_1}\mathrm{R}_{bm_1} - \frac{g_3}{b}\Sigma_b\Sigma_{m_2}\mathrm{R}_{bm_2} + e_1\Sigma_h\Sigma_{m_1}\mathrm{R}_{hm_1} \right.$$
$$+ e_1g_1\Sigma_{m_1}\Sigma_b\mathrm{R}_{bm_1} + e_1g_2\Sigma^2_{m_1} - e_1g_3\Sigma_{m_1}\Sigma_{m_2}\mathrm{R}_{m_1m_2} + e_2\Sigma_h\Sigma_{m_2}\mathrm{R}_{hm_2} + e_2g_1\Sigma_b\Sigma_{m_2}\mathrm{R}_{bm_2}$$
$$\left. + e_2g_2\Sigma_{m_1}\Sigma_{m_2}\mathrm{R}_{m_1m_2} - e_2g_3\Sigma^2_{m_2} \right\} \quad \text{. . . . . (cxiii.)}.$$

$$\mathrm{R}_{\mathrm{M}_e b} = \frac{1}{\Sigma_{\mathrm{M}_e}} \left\{ \Sigma_h \mathrm{R}_{hb} + g_1 \Sigma_b + g_2 \Sigma_{m_1} \mathrm{R}_{m_1 b} - g_3 \Sigma_{m_2} \mathrm{R}_{m_2 b} \right\} \quad \text{(cxiv.).}$$

$$\begin{aligned}\mathrm{R}_{\mathrm{M}_e \mathrm{S}_k} = \frac{\mathrm{S}_k}{\Sigma_{\mathrm{S}_k} \Sigma_{\mathrm{M}_e}} \Big\{ & f_1 \Sigma_h \Sigma_{m_1} \mathrm{R}_{hm_1} + g_1 f_1 \Sigma_{m_1} \Sigma_b \mathrm{R}_{bm_1} + f_1 g_2 \Sigma^2_{m_1} - f_1 g_3 \Sigma_{m_1} \Sigma_{m_2} \mathrm{R}_{m_1 m_2} + f_2 \Sigma_h \Sigma_{m_2} \mathrm{R}_{hm_2} \\ & + f_2 g_1 \Sigma_b \Sigma_{m_2} \mathrm{R}_{bm_2} + f_2 g_2 \Sigma_{m_1} \Sigma_{m_2} \mathrm{R}_{m_1 m_2} - f_2 g_3 \Sigma^2_{m_2} \Big\} \quad \text{(cxv.).}\end{aligned}$$

These results show us *that it is, in the general case of skew variation, impossible to select any one of the quantities—mean size, range, variability, or skewness of an organ, without at the same time in all probability modifying all the others.*

For example, the frequency of the incidence of certain types of diseases at different ages follows a distribution of this character. Hence, if any special class of the community had a mean age of incidence differing from that of the general population, we should expect correlated changes in such other characteristics of the disease as (i.) its first appearance; (ii.) its last appearance; (iii.) its tendence to heavier incidence above or below the mean age of incidence; (iv.) the concentration of its incidence about the mean age of incidence for this selected class.

Precisely similar series of changes would arise in the case of a random selection of individuals having the variability of a certain organ greater or less than that of the general population, there would be correlated changes in the size, range, and skewness of the distribution of this organ.

Turning to the mean and modal frequencies, we have

$$\begin{aligned}\frac{\Delta y_1}{y_1} &= -\frac{\Delta b}{b} + \left\{ \tfrac{3}{2} \frac{1}{m_1 + m_2 + 2} - \tfrac{1}{2} \frac{1}{m_1 + 1} + \frac{d\,\overline{\mathrm{S}}\,(m_1 + m_2 + 2)}{d\,(m_1 + m_2 + 2)} - \frac{d\,\overline{\mathrm{S}}\,(m_1 + 1)}{d\,(m_1 + 1)} \right\} \Delta m_1 \\ &\quad + \left\{ \tfrac{3}{2} \frac{1}{m_1 + m_2 + 2} - \tfrac{1}{2} \frac{1}{m_2 + 1} + \frac{d\,\overline{\mathrm{S}}\,(m_1 + m_2 + 2)}{d\,(m_1 + m_2 + 2)} - \frac{d\,\overline{\mathrm{S}}\,(m_2 + 1)}{d\,(m_2 + 1)} \right\} \Delta m_2 \\ &= -\Delta b/b + h_1 \Delta m_1 + h_2 \Delta m_2, \text{ say} \quad \text{(cxvi.).}\end{aligned}$$

Similarly,

$$\begin{aligned}\frac{\Delta y_2}{y_2} &= -\frac{\Delta b}{b} + \left\{ \frac{1}{m_1 + m_2 + 1} + \frac{1}{2\,(m_1 + m_2)} - \frac{1}{2m_1} + \frac{d\,\overline{\mathrm{S}}\,(m_1 + m_2)}{d\,(m_1 + m_2)} - \frac{d\overline{\mathrm{S}}(m_1)}{dm_1} \right\} \Delta m_1 \\ &\quad + \left\{ \frac{1}{m_1 + m_2 + 1} + \frac{1}{2\,(m_1 + m_2)} - \frac{1}{2m_2} + \frac{d\,\overline{\mathrm{S}}\,(m_1 + m_2)}{d\,(m_1 + m_2)} - \frac{d\overline{\mathrm{S}}(m_2)}{dm_2} \right\} \Delta m_2 \\ &= -\Delta b/b + k_1 \Delta m_1 + k_2 \Delta m_2, \text{ say} \quad \text{(cxvii.).}\end{aligned}$$

Here h_1, h_2, k_1, k_2 can be easily calculated, if we note that

$$\begin{aligned} d\,\overline{\mathrm{S}}\,(p)/dp &= -\mathrm{B}_1/2p^2 + \mathrm{B}_3/4p^4 - \mathrm{B}_5/6p^6 + \ldots \quad \text{(cxviii.).} \\ &= -\mathrm{T}/p, \text{ where T is the same as on p. 272.}\end{aligned}$$

Σ_{y_1} and Σ_{y_2} can be found in the usual manner by squaring after the insertion of the numerical values.

From (lxxxi.)–(lxxiii.) the probable errors and correlations of the moments can be found if required, the calculation being numerically somewhat laborious, but presenting nothing of novelty.

The probable error of the criterion

$$\kappa = 3\beta_1 - 2\beta_2 + 6,$$

where $\beta_1 = \mu_3^2/\mu_2^3$ and $\beta_2 = \mu_4/\mu_2^2$, may be found as follows: Put $\epsilon = (m_1 + 1)(m_2 + 1)$, $r = m_1 + m_2 + 2$; then we find

$$\kappa = \frac{12r^2(r + 1 + \epsilon)}{(r + 2)^2(r + 3)\epsilon} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxix.}),$$

and accordingly

$$\begin{aligned} \frac{\Delta\kappa}{\kappa} &= \left(\frac{2}{r} - \frac{2}{r+2} - \frac{1}{r+3} + \frac{1}{r+1+\epsilon}\right)\Delta r + \left(\frac{1}{r+1+\epsilon} - \frac{1}{\epsilon}\right)\Delta\epsilon \\ &= \left(\frac{2}{r} - \frac{2}{r+2} - \frac{1}{r+3} + \frac{1}{r+1+\epsilon} - \frac{r+1}{(r+1+\epsilon)(m_1+1)}\right)\Delta m_1 \\ &\quad + \left(\frac{2}{r} - \frac{2}{r+2} - \frac{1}{r+3} + \frac{1}{r+1+\epsilon} - \frac{r+1}{(r+1+\epsilon)(m_2+1)}\right)\Delta m_2 \\ &= i_1\Delta m_1 + i_2\Delta m_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxx.}), \end{aligned}$$

where i_1 and i_2 admit of easy calculation. Hence

$$\Sigma_\kappa^2/\kappa^2 = i_1^2\Sigma_{m_1}^2 + i_2^2\Sigma_{m_2}^2 + 2i_1i_2\Sigma_{m_1}\Sigma_{m_2}R_{m_1m_2} \quad . \quad . \quad . \quad . \quad . \quad (\text{cxxi.}).$$

The value of Σ_κ can thus be found, and the steadiness of the curve to its type ascertained.

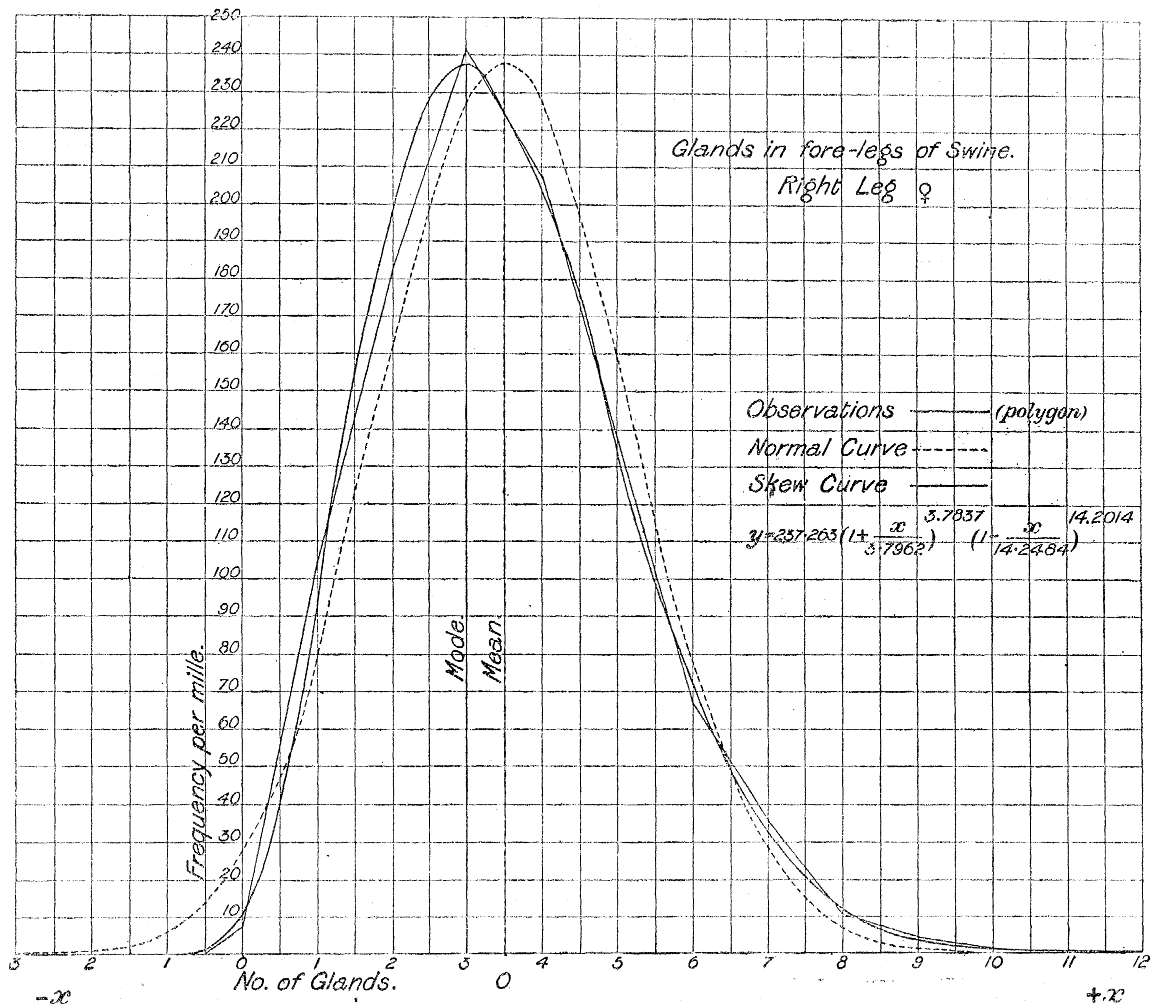
Illustration.—Glands of the Fore-legs of Swine.

In the 'Proceedings of the American Academy of Arts and Sciences,' vol. 32, p. 87, 1896, is a memoir by C. B. DAVENPORT and C. BULLARD, on the variation in number of the Müllerian glands in the fore-legs of 4,000 swine. The paper especially attracted our attention, because the authors are content to describe the frequency distribution of these glands by means of a normal curve. They write, after discussing the plotting of the normal curve on their diagram (pp. 90–91):—

"These and other characters of the 'probability' curve are indicated in that shown in dotted line* in the accompanying diagram. The diagram also shows the curve of

* The authors actually represent the normal curve by an 18-sided polygon.

distribution of the various numbers of glands occurring on a leg from 1 to 10*. This curve is drawn from the right female leg only; the curve for the other legs would be very similar. We shall speak in a moment of the method of construction of these curves; but we want now to call attention to the fairly close similarity of the two curves—that gained by observation and the theoretical one—a similarity so close that we are justified in concluding that the law of distribution of the variants in the leg glands of swine is the same as that of accidental errors."



Now, in our opinion, the curve was markedly skew, and it seemed to us that most interesting properties bearing on the action of selection on the Müllerian glands in swine actually depended on this skewness. We have taken the distribution of glands for 2,000 ♀ swine.

To illustrate the difficulty of applying the normal curve we may remark that it gives about 6 swine per mille with − 1 gland, and about 1·5 with − 2 glands, while

* The authors have forgotten that there is a sensible percentage of zero-glands.

it gives 30 per mille instead of 10 per mille with no glands. These difficulties are entirely met by the skew curve, which gives no frequency whatever of negative glands (see figure).

Taking the number of glands in the right fore-leg of female swine, we have the frequency series:

No. of glands .	0	1	2	3	4	5	6	7	8	9	10
Frequency . .	15	209	365	482	414	277	134	72	22	8	2
Per mille . . .	7·5	104·5	182·5	241	207	138·5	67	36	11	4	1

We have worked with the frequency per mille for convenience of reduction, although the actual number of observed cases, 2,000, is used, of course, in the determination of the probable errors.

Using the method of the paper in the 'Phil. Trans.,' A, vol. 186, p. 367, we found

$$\text{mean} = 3{\cdot}501 \text{ glands.} \qquad \sigma = 1{\cdot}680,774$$
$$\mu_2 = 2{\cdot}824,999 \qquad \beta_1 = 0{\cdot}259,1825$$
$$\mu_3 = 2{\cdot}417,278 \qquad \beta_2 = 3{\cdot}110,8211$$
$$\mu_4 = 24{\cdot}826,297 \qquad 6 + 3\beta_1 - 2\beta_2 = 0{\cdot}555,905$$

Thus the criterion is greater than zero, or the frequency distribution is of Type I., or has a limited range.

Proceeding we found

$$r = 19{\cdot}985119 \qquad \epsilon = 72{\cdot}71918$$
$$m_1 = 3{\cdot}783718 \qquad m_2 = 14{\cdot}201402$$
$$a_1 = 3{\cdot}79623 \qquad a_2 = 14{\cdot}24837$$
$$b = 18{\cdot}0446 \qquad y_0 = 237{\cdot}263$$
$$d = 0{\cdot}522996 \qquad \text{Sk.} = 0{\cdot}311164$$
$$\text{Mode} = 2{\cdot}978 \qquad \text{Start of curve} = -0{\cdot}818 \text{ gland.}$$

Thus it would appear that both the distance (d) from mean to mode and the skewness are very sensible, and that, unless their probable errors be very large, it is quite impossible to represent the results by a normal curve.

We may note that the range starts from $-0{\cdot}818$ gland and runs to $17{\cdot}227$ glands, so since it gives zero at -1 gland, we see that it sensibly confines the possible number of glands between 0 and 17, but we should have to examine considerably more than 2,000 swine to have a probability of more than 10 glands occurring. The

total range given is thus both in magnitude and position extremely satisfactory, and supposing only the frequency, not the actually measured quantity, *i.e.*, number of glands, to be known, the theory would have given a very accurate determination of the limits of possibility, especially the start of possibility with the whole number of glands.

In order to work out easily the determinant, Δ, and its minors, we found it desirable to bring out certain factors and reduce the formulæ given above to slightly different forms, which, as they are likely to be of general service, are here repeated.

Let $\alpha = m_1\epsilon_1 - (m_1 + m_2)\epsilon_3$, $\beta = m_2\epsilon_2 - (m_1 + m_2)\epsilon_3$, where ϵ_1, ϵ_2, ϵ_3 have the values given on p. 284, then we found

$$\Delta = \frac{n^4(m_1+m_2+1)}{b^4 m_1^2 m_2^2 (m_1-1)(m_2-1)} \begin{vmatrix} m_1 + m_2 & 0, & 1, & 0, \\ -(m_1-1), & m_1\alpha + m_2\beta, & -m_1\alpha, & -(m_1+m_2-2) \\ m_2(m_1-1), & m_2\beta, & m_1 m_2 \epsilon_3, & m_1 - 2m_2 + 1 \\ m_2 + 1, & -1, & 0, & 2(m_1+m_2+1) \end{vmatrix},$$

$$A_{11} = \frac{n^3}{b^2 m_1^2 m_2^2 (m_2-1)} \begin{vmatrix} m_1\alpha + m_2\beta, & -m_1\alpha, & -1 \\ m_2\beta, & m_1 m_2 \epsilon_3, & -(m_1+1) \\ -(m_2-1), & -m_1(m_2-1), & (m_1+1)(m_1+m_2+1) \end{vmatrix},$$

$$A_{22} = \frac{n^3(m_1+m_2+1)}{b^4 m_2^2 (m_1-1)(m_2-1)} \begin{vmatrix} m_1 + m_2, & 1, & 0, \\ m_2(m_1-1), & m_2^2(\epsilon_2 - \epsilon_3), & m_1 - 2m_2 + 1 \\ m_2 + 1, & -1, & 2(m_1+m_2+1) \end{vmatrix},$$

$$A_{33} = \frac{n^3(m_1+m_2+1)}{b^4 m_1^2 (m_1-1)(m_2-1)} \begin{vmatrix} m_1 + m_2, & -1, & 0 \\ -(m_2+1)(m_1-1), & m_1^2(\epsilon_1 - \epsilon_3), & m_2 - 2m_1 + 1 \\ m_2 + 1, & 0, & 2(m_1+m_2+1) \end{vmatrix},$$

$$A_{44} = \frac{n^3(m_1+m_2+1)}{b^2 m_1^2 m_2^2 (m_1-1)(m_2-1)} \begin{vmatrix} (m_1+m_2)(m_1+m_2-2), & 0, & (m_1-1)(m_2-1) \\ 0, & m_1\alpha + m_2\beta, & -m_1\alpha \\ m_1 + m_2 + 1, & m_2\beta, & m_1 m_2 \epsilon_3 \end{vmatrix},$$

$$A_{12} = -\frac{n^3(m_1+m_2+1)}{b^3 m_1 m_2^2 (m_2-1)} \begin{vmatrix} -1, & -m_1 m_2 \epsilon_3, & 0 \\ 0, & m_2\beta, & -1 \\ -(m_1+m_2), & -(m_1+1)(m_2-1), & m_2 + 2m_1 + 1 \end{vmatrix},$$

$$A_{13} = \frac{n^3(m_1+m_2+1)}{b^3 m_1 m_2 (m_2-1)} \begin{vmatrix} 0, & \alpha, & -1 \\ 1, & -m_2\epsilon_3, & 0 \\ -(m_1+m_2), & (m_2-1), & (m_1+1) \end{vmatrix},$$

$$A_{14} = -\frac{n^3(m_1+m_2+1)}{b^2 m_1^2 m_2^2 (m_2-1)} \begin{vmatrix} 0, & m_1\alpha + m_2\beta, & -m_1\alpha \\ m_2 - 1, & m_2\beta, & m_1 m_2 \epsilon_3 \\ -(m_1+m_2), & -1, & -m_1 \end{vmatrix},$$

$$A_{23} = -\frac{n^3(m_1+m_2+1)}{b^4 m_1 m_2 (m_1-1)(m_2-1)} \begin{vmatrix} m_1+m_2, & -1, & 0, \\ m_2(m_1-1), & -m_1 m_2 \epsilon_3, & m_1 - 2m_2 + 1 \\ m_2+1, & 0, & 2(m_1+m_2+1) \end{vmatrix},$$

$$A_{24} = \frac{n^3(m_1+m_2+1)}{b^3 m_1 m_2^2 (m_1-1)(m_2-1)} \begin{vmatrix} m_1+m_2, & -(m_2+1)(m_1-1), & -(m_1-1) \\ 1, & -m_1 m_2 \epsilon_3, & m_2\beta \\ 0, & m_2 - 2m_1 + 1, & -(m_1+m_2-2) \end{vmatrix},$$

$$A_{34} = -\frac{n^3(m_1+m_2+1)}{b^3 m_1^2 m_2 (m_1-1)(m_2-1)} \begin{vmatrix} m_1+m_2, & -(m_1-1), & m_2(m_1-1) \\ -1, & m_1\alpha, & -m_1 m_2 \epsilon_3 \\ 0, & -(m_1+m_2-2), & m_1 - 2m_2 + 1 \end{vmatrix}.$$

In our particular case we found

$$\epsilon_1 = \cdot 232{,}4012, \qquad \epsilon_2 = \cdot 067{,}9945, \qquad \epsilon_3 = \cdot 051{,}3099,$$
$$m_1\alpha = -\cdot 164{,}4934, \qquad m_2\beta = \cdot 607{,}8513.$$

With the aid of the values for m_1 and m_2 given above, the determinantal parts of the A's were then calculated. If these be δ, α_{11}, α_{22}, α_{33}, α_{44}, α_{12}, α_{13}, α_{14}, α_{23}, α_{24}, α_{34}, we have

$$\begin{aligned} \delta &= \cdot 153{,}7969, & \alpha_{12} &= \cdot 274{,}9969, \\ \alpha_{11} &= \cdot 348{,}3713, & \alpha_{13} &= \cdot 111{,}8280, \\ \alpha_{22} &= 13\cdot 018{,}7332, & \alpha_{14} &= \cdot 211{,}9650, \\ \alpha_{33} &= 47\cdot 671{,}9443, & \alpha_{23} &= 22\cdot 706{,}5156, \\ \alpha_{44} &= 13\cdot 357{,}1309, & \alpha_{24} &= 11\cdot 507{,}0153, \\ & & \alpha_{34} &= 25\cdot 088{,}6121. \end{aligned}$$

From these the standard deviations and correlations of errors in the algebraical constants are at once found. We have

$$\begin{aligned} \Sigma_h &= \cdot 2325, & R_{hm_1} &= \cdot 9387, \\ \Sigma_{m_1} &= \cdot 7784, & R_{hm_2} &= \cdot 7548, \\ \Sigma_{m_2} &= 5\cdot 5908, & R_{hb} &= \cdot 7143, \\ \Sigma_b &= 3\cdot 7602, & R_{m_1 m_2} &= \cdot 91145, \\ & & R_{m_1 b} &= \cdot 8726, \\ & & R_{m_2 b} &= \cdot 9942, \end{aligned}$$

Then as a step towards the determination of other probable errors, the standard deviation and umbral equation* for $\gamma = m_1 + m_2$ were found

$$\Sigma_\gamma = 6{\cdot}3085,$$
$$\chi_\gamma = \text{Antl. } \bar{1}{\cdot}091{,}2932\chi_{m_1} + \text{Antl. } \bar{1}{\cdot}947{,}5528\chi_{m_2}.$$

This led to

$$R_{\gamma m_1} = {\cdot}9312, \qquad R_{\gamma b} = {\cdot}9888, \qquad R_{\gamma h} = {\cdot}7848.$$

By aid of these auxiliary results the probable errors of all the algebraical and "physical" constants were determined.

PROBABLE Error Table.

	Constant.	Probable error.	Percentage probable error.
Algebraic constants.	m_1	0·5250	13·8762
	m_2	3·7709	26·5533
	a_1	0·1748	4·6056
	a_2	2·4351	17·0903
"Physical" constants.	Range	2·5362	14·0554
	Start of range	0·1568	—
	Mode	0·0398	—
	Mean	0·0253	—
	Standard deviation	0·0183	1·0911
	Mean to mode	0·0294	5·6308
	Skewness	0·0158	5·0655
	Modal frequency	3·2455	1·3679

Now it will be clear from an examination of these results that all the "physical" constants are determined with great accuracy.† The mean is subject to less probable error than the mode, the modal frequency has a slightly less probable error than the mean, and as it is less than 1·4 per cent. in the former case, either are closely known. The skewness and distance from mean to mode are known respectively with less than 5·1 and with 5·6 probable errors. Thus they are both significant constants. In other words, the curve differs significantly from a normal curve, and it is erroneous to represent the frequency by such a normal curve. The range which ought to be such that there is no frequency at -1 gland, gives no frequency at

* See footnote, p. 286, and later, p. 305. It may be as well to remind the reader that here, as in the other illustrations, logarithms of the full, not the cited values, were used in the calculations.

† The probable percentage errors m_1, m_2, a_1, a_2 are high, but this, as we have several times pointed out, is of small importance, as, owing to their high correlation, the actual shape of the curve is not changed sensibly by large changes in m_1 and m_2.

$-\cdot 818$ gland with a probable error of $\pm \cdot 157$. It is, therefore, clear that our method gives the start of the range with very considerable accuracy. The whole length of the range runs to 17·227 glands, with a probable error of 2·536. We may, accordingly, conclude that the maximum possible number of glands is hardly likely to be less than 16 or more than 20. We consider that this example is a good illustration of the accuracy with which the principal "physical" characteristics of a distribution may be obtained by aid of skew curves, and how they provide much information which is not given by the use of the normal curve.

The next point is the determination of the umbral equations giving the error correlations of the "physical" constants. They are, if Antl. stands for antilogarithm :

$$\chi_{\text{mean}} = \text{Antl.}\ \cdot 792{,}4156\chi_h - \text{Antl.}\ 1\cdot 380{,}2040\chi_b - \text{Antl.}\ 1\cdot 153{,}9620$$
$$+ \text{Antl.}\ 1\cdot 508{,}1033\chi_{m_2},$$

$$\chi_{\text{range}} = \chi_b,$$
$$\chi_{y_0} = -\text{Antl.}\ 1\cdot 011{,}7885\chi_b - \text{Antl.}\ \cdot 248{,}2856\chi_{m_1} + \text{Antl.}\ 1\cdot 097{,}6534\chi_{m_2},$$
$$\chi_\sigma = \text{Antl.}\ 1\cdot 109{,}9660\chi_b + \text{Antl.}\ \cdot 168{,}8507\chi_{m_1} - \text{Antl.}\ 1\cdot 151{,}0582\chi_{m_2},$$
$$\chi_d = \text{Antl.}\ \cdot 397{,}2701\chi_b - \text{Antl.}\ \cdot 274{,}1702\chi_{m_1} - \text{Antl.}\ \bar{1}\cdot 810{,}3180\chi_{m_2},$$
$$\chi_{sk.} = -\text{Antl.}\ \cdot 381{,}5919\chi_{m_1} + \text{Antl.}\ \cdot 367{,}7012\chi_{m_2}.$$

Multiplying these out pair and pair, we found

ERROR Correlation Table.

	Mean.	Range.	Modal frequency.	Standard deviation.	Mean to mode.	Skewness.
Mean	1	·0232	·3400	−·3493	·0500	·1309
Range	·0232	1	·6284	·0906	·2132	·2175
Modal frequency . . .	·3400	·6284	1	−·6944	−·1473	−·0141
Standard deviation . .	−·3493	·0906	−·6944	1	·5891	·4394
Mean to mode	·0500	·2132	−·1473	·5891	1	·9847
Skewness	·1309	·2175	−·0141	·4394	·9847	1

Hence, proceeding to multiply rows and divide columns by the corresponding standard deviations, we have, after altering the units, the following

PROGRESSION Table.

Corresponds to probable change in same units of	Unit change of					
	One gland in the mean.	One gland in the range.	One per cent. in the modal frequency.	One gland in the standard deviation.	One gland in the interval from mean to mode.	$\frac{1}{10}$ in the skewness.
Mean	1	·0002	·0063	−·4818	·0430	·0210
Range	2·3231	1	1·1651	12·5279	18·3604	3·4992
Modal frequency	18·3876	·3389	1	−51·7973	−6·8411	−·1225
Standard deviation . . .	−·2533	·0007	−·0093	1	·3668	·0511
Interval from mean to mode	·0583	·0025	−·0032	·9459	1	·1840
Skewness	·8156	·0135	−·0016	3·7761	5·2706	1

An examination of this table brings out several interesting features of the frequency distribution of Müllerian glands in the fore-legs of swine. If a group of swine were isolated, and found to have a higher mean number of glands, then this group would most probably have an increased possible range, but at the same time a decreased variability and a marked increase of skewness. This increase of the possible range with a decreased variability is especially notable, since the rough-and-ready class of statistician is very apt to treat the range observed as a measure of variability; we have here a case in which the same cause, raising of the mean, produces opposite effects on range and variability. Increase of range, it will next be observed, produces very little effect on any of the physical constants, but such effect as there is, is an increase of them all. To increase the modal frequency is to increase the range and to reduce both the variability and the skewness. Thus the more mediocre swine there exist in any group, the more nearly their distribution will be normal. Change in the variability is the cause which on the whole produces most effect. Increased variability means lowered mean and less mediocrity, but much increased skewness. Finally increased skewness denotes probable increase of range, variability, and mean.

As we have suggested in a previous illustration the principles of multiple correlation easily enable us to predict the probable change in a random selection in which two or more of the characters differ from those of the general population.

(20.) *Probable Errors and Error-Correlations of the Constants of the Generalised Probability Curve of Type*

$$y = y_0 \frac{1}{\{1 + (x/a)^2\}^m} e^{-\nu \tan^{-1}(x/a)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxxii.}).$$

This curve is discussed at length, 'Phil. Trans.,' A, vol. 186, pp. 376–80. The chief constants are given as follows, if $m = \frac{1}{2}(r + 2)$, $z = \nu^2 + r^2$, and h denote the origin :—

Moments—

$$\mu_2 = \frac{a^2}{r^2(r-1)}(r^2 + \nu^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxxiii.}),$$

$$\mu_3 = -\frac{4a^3\nu(r^2 + \nu^2)}{r^3(r-1)(r-2)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxxiv.}),$$

$$\mu_4 = \frac{3a^4(r^2 + \nu^2)\{(r+6)(r^2 + \nu^2) - 8r^2\}}{r^4(r-1)(r-2)(r-3)} \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxxv.}).$$

$$\text{Distance of centroid from origin} = -a\nu/r \quad . \quad . \quad . \quad . \quad (\text{cxxvi.}),$$

$$\text{Size of mean organ} = h - \frac{a\nu}{r} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxxvii.}),$$

$$\text{Size of modal organ} = h - \frac{a\nu}{r+2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxxviii.}),$$

$$\text{Distance from mean to mode} = d = \frac{2\nu a}{r(r+2)} \quad . \quad . \quad . \quad (\text{cxxix.}),$$

$$\text{Skewness} = \frac{2\nu}{r+2}\sqrt{\left(\frac{r-1}{r^2 + \nu^2}\right)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxxx.}),$$

$$\text{Standard deviation} = \sigma = \frac{a}{r\sqrt{(r-1)}}\sqrt{(r^2 + \nu^2)} \quad . \quad . \quad (\text{cxxxi.}),$$

$$y_0 = \frac{n}{a} e^{\frac{1}{2}\nu\pi}/G(r, \nu) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxxxii.}),$$

where

$$G(r, \nu) = \int_0^\pi \sin^r\theta e^{\nu\theta} \, d\theta$$

and

$$G(r, \nu) = \frac{r(r-1)}{r^2 + \nu^2} G(r-2, \nu) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{cxxxiii.}),$$

is the formula of reduction.

Further, we have the following BERNOULLI number series for $G(r, \nu)$, where $\tan\phi = \nu/r$:—

$$\log\{e^{-\frac{1}{2}\pi\nu} G(r, \nu)\} = \log\sqrt{(2\pi/r)} + (r+1)\log\cos\phi + \nu\phi$$
$$+ \overset{\infty}{\underset{0}{S}}\left\{\frac{B_{2s+1}(-1)^s}{(2s+1)(2s+2)\,r^{2s+1}}\left(1 - 2^{2s+2}\cos^{2s+1}\phi\cos\overline{2s+1}\,\phi\right)\right\} \quad . \quad \text{(cxxxiv.)}.$$

To find $y_1\,\delta x$, the mean frequency, we have only to put $x = -a\nu/r$ in (i.), and we have

$$\begin{aligned}\log y_1 &= \log y_0 + (r+2)\log\cos\phi + \nu\phi \\ &= \log n - \log a - \log\{e^{-\frac{1}{2}\pi\nu} G(r, \nu)\} + (r+2)\log\cos\phi + \nu\phi \\ &= \log n - \log a - \log\sqrt{(2\pi/r)} + \log\cos\phi - \chi,\end{aligned}$$

where χ stands for the summation in (cxxxiv.). Hence

$$y_1 = \frac{n}{a}\sqrt{\left(\frac{r}{2\pi}\right)}\,e^{-\chi}\cos\phi \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(cxxxv.)},$$

or,

$$= \frac{n}{\sqrt{(2\pi)}\,\sigma}\sqrt{\left(\frac{r}{r-1}\right)}\,e^{-\chi} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(cxxxvi.)}.$$

As typical constants we require the probable errors of the mean, the standard deviation, the skewness, and the mean frequency. It is clear that these will require us first to find Σ_h, Σ_r, Σ_ν, Σ_a, and R_{ha}, R_{hr}, $R_{h\nu}$, R_{ha}, $R_{r\nu}$, R_{ra}, $R_{\nu a}$.

We shall only indicate briefly the steps towards finding the integrals of the second differentials of $\log y$.

$$\log y = \log y_0 - m\log\{1 + (x/a)^2\} - \nu\tan^{-1}(x/a),$$

$$\frac{d(\log y)}{dx} = -\frac{\nu}{a}\,\frac{1}{\{1+(x/a)^2\}} - \frac{(2mx)/a^2}{1+(x/a)^2},$$

$$\frac{d^2(\log y)}{dx^2} = -\frac{2}{a^2}\left\{-\frac{\nu x/a}{\{1+(x/a)^2\}^2} - \frac{m}{1+(x/a)^2} + \frac{2m}{(1+x^2/a^2)^2}\right\}$$

$$= -(2/a^2)\{-\nu\sin\theta\cos^3\theta - m\cos^2\theta + 2m\cos^4\theta\}$$

$$\int_{-\infty}^{+\infty} y\,\frac{d^2\log y}{dx^2}\,dx$$
$$= -\frac{2n}{a^2\,G(r,\nu)}\left\{-e^{\frac{1}{2}\nu\pi}\nu\int_{-\pi/2}^{\pi/2}\sin\theta\cos^{r+3}\theta\,e^{-\nu\theta}\,d\theta - mG(r+2,\nu) + 2mG(r+4,\nu)\right\},$$

whence, remembering that $2m = r + 2$, and integrating the first integral by parts, we find

$$\int_{-\infty}^{+\infty} y \frac{d^2 \log y}{dx^2} dx = -\frac{2n}{a^2}\left\{\left(\frac{\nu^2}{r+4} + (r+2)\right)\frac{G(r+4,\nu)}{G(r,\nu)} - \tfrac{1}{2}(r+2)\frac{G(r+2,\nu)}{G(r,\nu)}\right\}$$

$$= -\frac{n}{a^2}\frac{(r+4)\{\nu^2+(r+2)^2\}}{\nu^2+(r+4)^2}\frac{G(r+2,\nu)}{G(r,\nu)}$$

or,

$$-a_{11} = -\frac{n}{a^2}\frac{(r+1)(r+2)(r+4)}{\nu^2+(r+4)^2} \quad \ldots \ldots \quad \text{(cxxxvii.)}.$$

Precisely similar reductions lead us to

$$-a_{12} = \int_{-\infty}^{+\infty} y \frac{d^2(\log y)}{dx\,da} dx = -\frac{n\nu(r+1)(r+2)}{a^2\{\nu^2+(r+4)^2\}} \quad \ldots \quad \text{(cxxxviii.)},$$

$$-a_{13} = \int_{-\infty}^{+\infty} y \frac{d^2(\log y)}{dx\,dr} dx = \frac{n\nu(r+1)}{a\{\nu^2+(r+2)^2\}} \quad \ldots \quad \text{(cxxxix.)},$$

$$-a_{14} = \int_{-\infty}^{+\infty} y \frac{d^2(\log y)}{dx\,d\nu} dx = -\frac{n}{a}\frac{(r+1)(r+2)}{\nu^2+(r+2)^2} \quad \ldots \quad \text{(cxl.)},$$

$$-a_{22} = \int_{-\infty}^{+\infty} y \frac{d^2(\log y)}{da^2} dx = -\frac{n}{a^2}(r+1)\frac{\nu^2+2(r+4)}{\nu^2+(r+4)^2} \quad \ldots \quad \text{(cxli.)},$$

$$-a_{23} = \int_{-\infty}^{+\infty} y \frac{d^2(\log y)}{da\,dr} dx = \frac{n}{a}\frac{\nu^2+(r+2)}{\nu^2+(r+2)^2} \quad \ldots \quad \text{(cxlii.)},$$

$$-a_{24} = \int_{-\infty}^{+\infty} y \frac{d^2(\log y)}{da\,d\nu} dx = -\frac{n}{a}\frac{\nu(r+1)}{\nu^2+(r+2)^2} \quad \ldots \quad \text{(cxliii.)},$$

$$-a_{33} = \int_{-\infty}^{+\infty} y \frac{d^2(\log y)}{dr^2} dx = -n\frac{d^2}{dr^2}\{\log G(r,\nu)\} \quad \ldots \quad \text{(cxliv.)},$$

$$-a_{34} = \int_{-\infty}^{+\infty} y \frac{d^2(\log y)}{dr\,d\nu} dx = -n\frac{d^2}{dr\,d\nu}\{\log G(r,\nu)\} \quad \ldots \quad \text{(cxlv.)},$$

$$-a_{44} = \int_{-\infty}^{+\infty} y \frac{d^2(\log y)}{d\nu^2} = -n\frac{d^2}{d\nu^2}\{\log G(r,\nu)\} \quad \ldots \quad \text{(cxlvi.)}.$$

It will now be needful to find easily calculable series for the second differentials of $\log G(r, \nu)$. These can be obtained from (cxxxiv.). We find

$$\frac{d}{dr}\{\log G(r,\nu)\} = -\frac{1}{2r} + \log\cos\phi + \frac{\sin^2\phi}{r}$$
$$-\overset{\infty}{\underset{0}{S}}\frac{B_{2s+1}(-1)^s}{(2s+2)r^{2s+2}}\{1 - 2^{2s+2}\cos^{2s+2}\phi\cos(2s+2)\phi\} \quad \ldots \quad \text{(cxlvii.)},$$

$$\frac{d}{d\nu}\{\log G(r,\nu)\} = \tfrac{1}{2}\pi - \frac{\sin\phi\cos\phi}{r} + \phi$$
$$+\overset{\infty}{\underset{0}{S}}\frac{(-1)^s B_{2s+1}}{(2s+2)r^{2s+2}}2^{2s+2}\cos^{2s+2}\phi\sin(2s+2)\phi \quad \ldots \quad \text{(cxlviii.)}.$$

Hence

$$\frac{d^2}{dr^2}\{\log G(r, \nu)\} = \frac{1}{r^2}\Big\{\cdot 5 + \sin^2\phi\,(r - 1 - 2\cos^2\phi) + \overset{\infty}{\underset{0}{S}} \frac{(-1)^s B_{2s+1}}{r^{2s+1}}\{1 - 2^{2s+2}\cos^{2s+3}\phi\cos(2s+3)\phi\}\Big\} \quad \text{(cxlix.)},$$

$$\frac{d^2}{d\nu^2}\{\log G(r, \nu)\} = \frac{1}{r^2}\Big\{\cos^2\phi\,(2 - 1 + 2\sin^2\phi) + \overset{\infty}{\underset{0}{S}} \frac{(-1)^s B_{2s+1}}{r^{2s+1}} 2^{2s+2}\cos^{2s+3}\phi\cos(2s+3)\phi\Big\} \quad . \quad . \quad . \quad \text{(cl.)},$$

$$\frac{d^2}{dr\,d\nu}\{\log G(r, \nu)\} = \frac{1}{r^2}\Big\{\sin\phi\cos\phi\,(2\cos^2\phi - r) - \overset{\infty}{\underset{0}{S}} \frac{(-1)^s B_{2s+1}}{r^{2s+1}} 2^{2s+2}\cos^{2s+3}\phi\sin(2s+3)\phi\Big\} \quad . \quad . \quad . \quad \text{(cli.)}.$$

These allow of the fairly rapid calculation of a_{33}, a_{34}, a_{44}. The values of the standard deviations of the errors, and of the error correlations, can then all be calculated from the determinant

$$\Delta = \begin{vmatrix} a_{11}, & a_{12}, & a_{13}, & a_{14}, \\ a_{12}, & a_{22}, & a_{23}, & a_{24}, \\ a_{13}, & a_{23}, & a_{33}, & a_{34}, \\ a_{14}, & a_{24}, & a_{34}, & a_{44}, \end{vmatrix}$$

and its minors.*

Let $b_{pp'} = \frac{1}{n}a_{pp'}$, and let $\Delta' = \frac{1}{n^4}\Delta$, then if $B_{pp'}$ be the minor corresponding to $b_{pp'}$ in Δ', we must work out for any special numerical case

$$\Sigma_h^2 = \frac{1}{n}\frac{B_{11}}{\Delta'}, \qquad \Sigma_a^2 = \frac{1}{n}\frac{B_{22}}{\Delta'},$$

$$\Sigma_r^2 = \frac{1}{n}\frac{B_{33}}{\Delta'}, \qquad \Sigma_\nu^2 = \frac{1}{n}\frac{B_{44}}{\Delta'},$$

$$R_{ah} = \frac{B_{12}}{\sqrt{(B_{11}B_{22})}}, \qquad R_{rh} = \frac{B_{13}}{\sqrt{(B_{11}B_{33})}},$$

$$R_{\nu h} = \frac{B_{14}}{\sqrt{(B_{11}B_{22})}}, \qquad R_{ar} = \frac{B_{23}}{\sqrt{(B_{22}B_{33})}},$$

$$R_{a\nu} = \frac{B_{24}}{\sqrt{(B_{22}B_{44})}}, \qquad R_{r\nu} = \frac{B_{34}}{\sqrt{(B_{33}B_{44})}}.$$

* As in the former case, these were all developed, but the extreme length of the resulting formulæ gives them no advantage over working in any special case with the numerical determinants.

In general none of these correlations vanish, and their values must all be found before the errors and correlations of the chief characteristics of the frequency can be found.

The following results, easily obtained by aid of the relation $\tan\phi = \nu/r$, will be of service

$$\Sigma_\phi^2 = \frac{\cos^4\phi}{r^2}\{\Sigma_\nu^2 + \tan^2\phi\Sigma_r^2 - 2\tan\phi\Sigma_\nu\Sigma_r R_{\nu r}\} \quad . \quad . \quad . \quad . \quad \text{(clii.)},$$

$$R_{\phi\nu} = \frac{\cos^2\phi}{r\Sigma_\phi}\{\Sigma_\nu - \tan\phi\Sigma_r R_{r\nu}\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(cliii.)},$$

$$R_{\phi r} = \frac{\cos^2\phi}{r\Sigma_\phi}\{\Sigma_\nu R_{r\nu} - \tan\phi\Sigma_r\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(cliv.)},$$

$$R_{\phi a} = \frac{\cos^2\phi}{r\Sigma_\phi}\{\Sigma_\nu R_{a\nu} - \tan\phi\Sigma_r R_{ar}\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(clv.)},$$

$$R_{\phi h} = \frac{\cos^2\phi}{r\Sigma_\phi}\{\Sigma_\nu R_{h\nu} - \tan\phi\Sigma_r R_{hr}\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(clvi.)}.$$

By (cxxvii.) and (cxxxi.) if $\bar{x}$ be the mean size of organ,

$$\Delta\bar{x} = \Delta h - \tan\phi\Delta a - \frac{a}{\cos^2\phi}\Delta\phi,$$

$$\frac{\Delta\sigma}{\sigma} = \frac{\Delta a}{a} + \tan\phi\Delta\phi - \tfrac{1}{2}\frac{\Delta r}{r-1}.$$

Hence

$$\Sigma_{\bar{x}}^2 = \Sigma_h^2 + \tan^2\phi\Sigma_a^2 + \frac{a^2}{\cos^4\phi}\Sigma_\phi^2 - 2\tan\phi\Sigma_h\Sigma_a R_{ha} - \frac{2a}{\cos^2\phi}\Sigma_h\Sigma_\phi R_{\phi h}$$
$$+ 2\tan\phi\frac{a}{\cos^2\phi}\Sigma_a\Sigma_\phi R_{a\phi} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(clvii.)},$$

$$\frac{\Sigma_\sigma^2}{\sigma^2} = \frac{\Sigma_a^2}{a^2} + \tan^2\phi\Sigma_\phi^2 + \frac{1}{4(r-1)^2}\Sigma_r^2 + \frac{2}{a}\tan\phi\Sigma_a\Sigma_\phi R_{a\phi} - \frac{1}{a(r-1)}\Sigma_a\Sigma_r R_{ar}$$
$$- \tan\phi\frac{1}{r-1}\Sigma_\phi\Sigma_r R_{\phi r} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(clviii.)},$$

and

$$R_{\bar{x}\sigma} = \frac{\sigma}{\Sigma_\sigma\Sigma_{\bar{x}}}\left\{\frac{\Sigma_a\Sigma_h R_{ah}}{a} - \frac{\tan\phi}{a}\Sigma_a^2 - \frac{\Sigma_a\Sigma_\phi R_{a\phi}}{\cos^2\phi} + \tan\phi\Sigma_h\Sigma_\phi R_{h\phi} - \tan^2\phi\Sigma_a\Sigma_\phi R_{a\phi}\right.$$
$$\left. - \frac{a\sin\phi}{\cos^3\phi}\Sigma_\phi^2 - \frac{1}{2(r-1)}\Sigma_h\Sigma_r R_{hr} + \frac{\tan\phi}{2(r-1)}\Sigma_a\Sigma_r R_{ar} + \frac{a}{2\cos^2\phi(r-1)}\Sigma_r\Sigma_\phi R_{r\phi}\right. . \quad \text{(clix.)}.$$

From (cxxx.) we have, if S_k = skewness,

$$S_k = \frac{2\sin\phi}{r+2}\sqrt{(r-1)}.$$

Or, taking logarithmic differentials,

$$\Delta S_k/S_k = \cot\phi\,\Delta\phi - \left(\frac{1}{r+2} - \tfrac{1}{2}\frac{1}{r-1}\right)\Delta r,$$

whence,

$$\Sigma^2_{S_k}/S_k^2 = \cot^2\phi\,\Sigma^2_\phi + \frac{(r-4)^2}{4(r-1)^2(r+2)^2}\Sigma^2_r - \cot\phi\,\frac{r-4}{(r-1)(r+2)}\Sigma_\phi\Sigma_r R_{r\phi}\,. \quad \text{(clx.)}.$$

In a similar manner $R_{\bar{x}S_k}$ and $R_{\sigma S_k}$ can be found if desired. None of these quantities will, as a rule, vanish, and as very many measurements on animals give curves of the tangent type, we conclude that in general *all selection of the size of an organ alters its variability and the skewness of its distribution, and again all selection of variability connotes alteration of the size and skewness of the selected organ.*

The probable errors of μ_2, μ_3, and μ_4, as well as the error-correlations of these quantities, can all be found from the differentials of (cxxiii.), (cxxiv.), and (cxxv.); the calculation is laborious, but presents no novelty.

Lastly, the probable errors of the mean and modal frequencies may be deduced.

For the mean frequency we start from (cxxxv.) and use (cxxxiv.).

This requires us to know $\Delta\chi$, where

$$\chi = \mathop{S}_0^\infty \frac{(-1)^s B_{2s+1}}{(2s+1)(2s+2)r^{2s+1}}\left(1 - 2^{2s+2}\cos^{2s+1}\phi\cos(2s+1)\phi\right) \quad . \quad . \quad \text{(clxi.)}.$$

We have, as in (cxlvii.) and (cxlviii.),

$$\begin{aligned}\Delta\chi = &-\frac{\Delta r}{r}\mathop{S}_0^\infty \frac{(-1)^s B_{2s+1}}{(2s+2)r^{2s+1}}\left(1 - 2^{2s+2}\cos^{2s+2}\phi\cos(2s+2)\phi\right)\\ &+\frac{\Delta\nu}{r}\mathop{S}_0^\infty \frac{(-1)^s B_{2s+1}}{(2s+2)r^{2s+1}}\,2^{2s+2}\cos^{2s+2}\phi\sin(2s+2)\phi\\ = &-c_1\,\Delta r/r + c_2\,\Delta\nu/r,\end{aligned}$$

where c_1 and c_2 admit of fairly easy calculation.

Hence, by (cxxxv.), we find,

$$\Delta y_1/y_1 = -\Delta a/a + (\tfrac{1}{2} + \sin^2\phi + c_1)\Delta r/r - (c_2 + \cos\phi\sin\phi)\Delta\nu/r.$$

$$\begin{aligned}\Sigma^2_{y_1}/y_1^2 = &\ \Sigma^2_a/a^2 + (\tfrac{1}{2} + \sin^2\phi + c_1)^2\Sigma^2_r/r^2 + (c_2 + \cos\phi\sin\phi)^2\Sigma^2_\nu/r^2\\ &-\frac{2}{ra}(\tfrac{1}{2} + \sin^2\phi + c_1)\Sigma_a\Sigma_r R_{ar} + \frac{2(c_2 + \cos\phi\sin\phi)}{ar}\Sigma_a\Sigma_\nu R_{a\nu}\\ &-\frac{2(\tfrac{1}{2} + \sin^2\phi + c_1)(c_2 + \cos\phi\sin\phi)}{r^2}\Sigma_r\Sigma_\nu R_{\nu r}\ . \quad . \quad . \quad . \quad . \quad . \quad \text{(clxii.)}.\end{aligned}$$

If the problem be to find the modal frequency $y_2\,\delta x$, we easily deduce y_2 by putting

$x = -\dfrac{\nu}{a(r+2)}$ in the equation to the curve. Writing $\tan\phi' = \nu/(r+2)$ and χ' the same function of $r+2$ that χ is of r, we have

$$y_2 = \frac{n}{a}\sqrt{\left(\frac{r+2}{2\pi}\right)}\cos\phi'^{\frac{r+1}{r+2}}e^{-\chi'} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(clxiii.)}.$$

Since $\dfrac{r+1}{\sqrt{(r+2)}}$ is greater than $\sqrt{r}$, and ϕ' and χ' are less than ϕ and χ respectively, it follows that y_2 is greater than y_1, as it should be. Further we find

$$\frac{\Delta y_2}{y_2} = -\frac{\Delta a}{a} + \left(\tfrac{1}{2}\frac{r+3}{r+1} + \sin^2\phi' + c_1'\right)\frac{\Delta r}{r+2} - (c_2' + \cos\phi'\sin\phi')\frac{\Delta\nu}{r+2} \quad . \quad \text{(clxiv.)}.$$

Here c_1' and c_2' are the same functions of $r+2$ and ϕ' that c_1 and c_2 are of r and ϕ. The usual process of squaring and introducing the standard deviations into the square terms and the product of standard deviations and correlations into the product terms will give us $\Sigma^2_{y_2}$.

(21.) *Illustration.—Stature of Children.*

In order to illustrate the difficulties which may arise in determining the probable errors of the constants and the error correlations, we have selected for this illustration not a curve markedly skew, but one which is extremely nearly normal. The problem in this case is accordingly the following one: Are the values of the constants obtained for the distribution and distinguishing it from a normal distribution really significant? The difficulties which arise in the course of the arithmetical work depend upon the fact that, as the distribution is nearly normal, its constants approach the values at which the type of the skew curve passes over into the normal curve, and consequently not only will their probable errors be large, but, as in all cases of approach to limits, they will depend upon expressions tending to become indeterminate. Thus in the evaluation of the determinant Δ and its minors, we at once found our results depended on the ratio of the differences of very small quantities. We were accordingly in this case obliged to calculate our constituents to a degree of accuracy which will, in general, be quite unnecessary, and which was only possible and straightforward owing to the ready help of a large sized Brunsviga. That the method, even in a critical case of this kind, gives correct results is evidenced by the agreement of our values of the constants with those (probable errors of mean and standard deviation) which can be readily calculated by other processes.

The example we have selected is that given for the stature of 2192 St. Louis school girls of 8 years of age in 'Phil. Trans.,' A, vol. 186, p. 386.

The equation to the frequency curve is

$$x = 14{\cdot}9917 \tan\theta,$$
$$y = 235{\cdot}323 \cos^{32{\cdot}8023}\theta e^{-4{\cdot}56967\theta},$$

the axis of x being *positive towards dwarfs* and the origin 2·2241 on the positive side of the mean. The unit of x is 2 cms. of height, and all the constants except the mean height are given in two-centimetre units.

We have the following values of the constants:—

$$\begin{array}{llrl}
\mu_2 = & 7{\cdot}70739, & \text{Mean height} = & 118{\cdot}271 \text{ centims.},\\
\mu_3 = & -2{\cdot}38064, & \sigma = & 2{\cdot}77622,\\
\mu_4 = & 192{\cdot}17419, & y_1 = \text{modal frequency} = & 324{\cdot}18,\\
 & & y_2 = \text{mean frequency} = & 323{\cdot}76,\\
d = & {\cdot}135{,}606, & \text{Sk.} = \text{skewness} = & {\cdot}04885,\\
r = & 30{\cdot}8023, & m = & 16{\cdot}4011,\\
\nu = & 4{\cdot}56967, & a = & 14{\cdot}9917.
\end{array}$$

It will be seen at once that the skewness is small, that the mean and mode are close together, and the mean and modal frequencies are almost identical. Our problem is: Are these differences significant or not?

Let $n = 2192$, the total population; then the values of a, r, ν given above were assumed to be absolutely exact, and Δ calculated with its constituents to 9 places of figures, as it depends on the differences of very small quantities. We shall indicate one or two stages in the arithmetical work

$$\frac{\Delta}{n^4} = \begin{vmatrix}
{\cdot}131{,}108{,}064 & {\cdot}017{,}214{,}971 & -{\cdot}008{,}837{,}638 & {\cdot}063{,}438{,}906\\
{\cdot}017{,}214{,}971 & {\cdot}010{,}392{,}042 & -{\cdot}003{,}264{,}669 & {\cdot}008{,}837{,}638\\
-{\cdot}008{,}837{,}638 & -{\cdot}003{,}264{,}669 & {\cdot}001{,}144{,}775 & -{\cdot}004{,}422{,}657\\
{\cdot}063{,}438{,}906 & {\cdot}008{,}837{,}638 & -{\cdot}004{,}422{,}657 & {\cdot}030{,}799{,}043
\end{vmatrix}.$$

The evaluation of this determinant and its minors was then carried out by means of the Brunsviga, and we found

$$\frac{\Delta}{n^4} = \frac{{\cdot}104{,}824{,}472}{10^{12}}, \qquad \frac{A_{12}}{n^3} = -\frac{1200{\cdot}842{,}528}{10^{12}},$$

$$\frac{A_{11}}{n^3} = \frac{670{\cdot}195{,}695{,}496}{10^{12}}, \qquad \frac{A_{13}}{n^3} = -\frac{5059{\cdot}387{,}378}{(10)^{12}},$$

$$\frac{A_{22}}{n^3} = \frac{4606{\cdot}123{,}523}{10^{12}}, \qquad \frac{A_{14}}{n^3} = -\frac{1762{\cdot}570{,}609}{10^{12}},$$

$$\frac{A_{33}}{n^3} = \frac{76025{\cdot}131{,}845}{10^{12}}, \qquad \frac{A_{23}}{n^3} = \frac{18675{\cdot}261{,}289}{10^{12}},$$

$$\frac{A_{44}}{n^3} = \frac{4828{\cdot}382{,}384}{10^{12}}, \qquad \frac{A_{24}}{n^3} = \frac{3833{\cdot}460{,}555}{10^{1}},$$

$$\frac{A_{34}}{n^3} = \frac{15979{\cdot}332{,}581}{10^{1}}$$

We can now, remembering that

$$\Sigma_p = \sqrt{(A_{pp}/\Delta)}, \quad \text{and} \quad R_{pq} = A_{pq}/\sqrt{(A_{pp}A_{qq})},$$

write down the standard deviations and error correlations of the algebraical constants

$$\begin{array}{llll} \Sigma_h = & 1{\cdot}7078, & R_{ha} = & -\,{\cdot}6835, \\ \Sigma_a = & 4{\cdot}4773, & R_{hr} = & -\,{\cdot}7088, \\ \Sigma_r = & 18{\cdot}1898, & R_{h\nu} = & -\,{\cdot}9798, \\ \Sigma_\nu = & 4{\cdot}5840, & R_{ar} = & {\cdot}9980, \\ & & R_{a\nu} = & {\cdot}8129, \\ & & R_{r\nu} = & {\cdot}8340. \end{array}$$

Here h marks the position of the origin of the curve, and the numerical values are only retained to four places of figures, although, of course, in the further calculations the logarithms of the full values of the Σ's and R's have been used.

It will be noticed at once that though a, r, and ν have very considerable probable errors, the correlation between them is very high. In other words, as the curve approaches its limiting shape, a, r, and ν may vary very considerably, but owing to their close correlation this will not sensibly affect the geometrical shape of the curve.

The next stage was to determine the standard deviations and error correlations of certain subsidiary constants. Here, as in the determination later of the like quantities for the "physical" characters, we found the umbral notation of great service. It consists, as we have seen, in writing down a difference equation between any constants, and then replacing the differences δu by $\Sigma_u\chi_u$, δv by $\Sigma_v\chi_v$, &c., where χ_u, χ_v, &c., are quantities which obey the relations $\chi_u^2 = 1$, $\chi_v^2 = 1$, $\chi_u\chi_v = R_{uv}$. Thus, if $\tan\phi = \nu/r$, we find for the umbral equation

$$\Sigma_\phi\chi_\phi = \frac{\cos^2\phi\Sigma_\nu}{r}\chi_\nu - \frac{\sin\phi\cos\phi}{r}\Sigma_r\chi_r.$$

Whence, putting in the numerical values, we have

$$\Sigma_\phi\chi_\phi = \text{Antl. } \bar{1}{\cdot}163{,}2115\chi_\nu - \text{Antl. } \bar{2}{\cdot}933{,}0706\chi_r$$

where Antl. stands for antilogarithm, in which form we found it easiest to keep the umbral coefficients. The square of this result gave at once

$$\Sigma_\phi = {\cdot}087{,}926$$

and dividing out by its logarithm, we have the pure umbral equation

$$\chi_\phi = \text{Antl. } {\cdot}219{,}0937\chi_\nu - \text{Antl. } \bar{1}{\cdot}988{,}9728\chi_r.$$

Our object was then to find such pure umbral equations connecting all the "physical" constants with the algebraic constants. Their products will then give the error correlations of all the "physical" constants in terms of the correlations already known between the algebraical constants.

For example, multiplying the above equation for χ_ϕ by χ_h, χ_a, χ_r, χ_ν we have, since $\chi_h\chi_a = R_{ha}$, $\chi_a\chi_r = R_{ar}$, &c., are already known,

$$\begin{aligned} R_{h\phi} &= -\cdot 9317, \text{ actually } \log(-R_{h\phi}) = \bar{1}\cdot 969{,}2668, \\ R_{a\phi} &= \cdot 3733, \quad ,, \quad \log R_{a\phi} = \bar{1}\cdot 572{,}0115, \\ R_{r\phi} &= \cdot 4063, \quad ,, \quad \log R_{r\phi} = \bar{1}\cdot 608{,}8746, \\ R_{\nu\phi} &= \cdot 8430, \quad ,, \quad \log R_{\nu\phi} = \bar{1}\cdot 925{,}8379. \end{aligned}$$

It was these logarithms, of course, which were used in the further calculations.

Since h is measured negatively (*i.e.*, towards dwarfs, x is positive), we must write for transferring origin to the mean

$$x' = x + a \tan \phi,$$

where $a \tan \phi$ is the distance between the old origin and the mean, or if m be used to represent the mean we have

$$m = h + a \tan \phi.$$

Hence we find the umbral equation

$$\Sigma_m\chi_m = \text{Antl. } \cdot 232{,}4493\chi_h + \text{Antl. } \bar{1}\cdot 822{,}3179\chi_a + \text{Antl. } \cdot 129{,}4233\chi_\phi.$$

Hence we determine

$$\Sigma_m = \cdot 0549,$$

and the pure umbral equation

$$\chi_m = \text{Antl. } \bar{1}\cdot 492{,}5897\chi_h + \text{Antl. } \bar{1}\cdot 082{,}4583\chi_a + \text{Antl. } \bar{1}\cdot 389{,}5637\chi_\phi.$$

In precisely the same way all the other "physical" constants, *i.e.*, the standard deviation, σ, the mean frequency, y_2, the distance between mean and mode, d, and the skewness, Sk., were found, and the umbral equations investigated. It is only necessary here to give the results.

	Quantity.	Probable error.	Percentage probable error.
Algebraic constants.	h, position of origin	1·1519	—
	a	3·0199	20·1438
	r	12·2688	39·8309
	ν	3·0919	67·6612
"Physical" constants.	Position of mean	0·03705	—
	Position of mode	0·05950	—
	Mean frequency, y_2	4·4362	1·3703
	Standard deviation, σ	0·02984	1·0750
	Mean to mode, d	0·0497	36·6420
	Skewness, Sk.	0·02661	54·4690

Now it will be seen at once that the probable errors in the algebraic constants are large, but that the probable errors in the position of the mean, of the mode, and in the magnitudes of the mean frequency and standard deviation are small. The position of the mean is sensibly more correct than that of the mode. On the other hand, the distance of the mean from the mode and the skewness have large probable errors, not, however, so large but what these quantities are probably significant. The frequency distribution probably differs significantly from the normal distribution, but the difference is small and would require a very large number of observations to determine it with extreme accuracy. That there is a significant divergence from normality is also indicated by the sensible difference between the percentage errors in y_2 and σ, which would be equal for a normal distribution. Had we taken a normal distribution, the probable error of the mean would have been ·0400, and of the standard deviation, ·02831. In fact, the standard deviation of the standard deviation, if calculated for the normal curve = ·04197, if calculated by our present method = ·044246, and if calculated by a modified form of the fourth moment formula given by CZUBER* = ·044240. This shows that the arithmetic of our process has been substantially correct.

We now place together the umbral equations for the correlations of the errors in the "physical" constants. They are

$$\begin{aligned}
\chi_m &= \text{Antl. } 1{\cdot}492{,}5897\chi_h + \text{Antl. } 1{\cdot}082{,}4583\chi_a + \text{Antl. } 1{\cdot}389{,}5637\chi_\phi \\
\chi_\sigma &= \text{Antl. } 1{\cdot}272{,}7484\chi_a - \text{Antl. } 1{\cdot}282{,}1306\chi_r + \text{Antl. } \bar{1}{\cdot}913{,}0028\chi_\phi \\
\chi_{y_2} &= \text{Antl. } \bar{1}{\cdot}816{,}6966\chi_\phi - \text{Antl. } 1{\cdot}167{,}3480\chi_a + \text{Antl. } 1{\cdot}155{,}7688\chi_r \\
\chi_d &= \text{Antl. } {\cdot}266{,}3616\chi_\nu + \text{Antl. } \bar{1}{\cdot}740{,}1625\chi_a - \text{Antl. } {\cdot}323{,}8256\chi_\nu \\
\chi_{sk.} &= \text{Antl. } \bar{1}{\cdot}865{,}6420\chi_\phi - \text{Antl. } {\cdot}017{,}0466\chi_r.
\end{aligned}$$

From these results any correlation between pairs of errors, "physical" or algebraic, can be found at once. The following table gives the chief results:—

CORRELATION Coefficients between Errors in Constants.

	m.	σ.	y_2.	d.	sk.
m	1	·0772	−·0584	·0826	·0426
σ	·0772	1	−·7062	·1177	·1431
y_2	−·0584	−·7062	1	·1779	·4086
d	·0826	·1177	·1779	1	·6843
sk.	·0426	·1431	·4086	·6843	1

* 'Theorie der Beobachtungsfehler,' p. 133.

Now it is clear that although the curve is nearly normal, there is still sensible correlation between quantities—*e.g.*, mean and σ or d—which would have no correlation between them if the curve were absolutely normal. This will be clearer if, as in the previous illustration, we replace this table by a table of regression coefficients.

	m.	y_2.	σ.	d.	sk.
m	1	−·0005	·0959	·0616	·05935
y_2	−6·9883	1	−104·9745	15·8875	68·1271
σ	·0622	−·00475	1	·0707	·16045
d	·1108	·0020	·1960	1	1·2780
sk.	·0306	·00245	·12755	·3665	1

This table has now finally to be thrown into more suitable units and attention paid to the fact that *m increases towards dwarfs*. We have, after the proper changes, the following results :—

PROGRESSION Table.

Corresponds to probable changes in the same units of	Unit change of				
	One centim. in mean stature.	One child per hundred in frequency of mean stature.	One centim. in standard deviation.	One centim. in interval from mean to mode.	1/10 in the skewness.
Mean stature	1	·0032	−·0959	−·0616	−·0119
Mean frequency	1·0798	1	−16·1745	2·4536	2·1042
Standard deviation	−·0622	−·0308	1	·0707	·0321
Interval from mean to mode	−·1108	·0129	·1960	1	·2556
Skewness	−·1530	·0794	·6378	1·8333	1

This table is extremely suggestive. It shows us that a random selection of girls of eight which had an increase of stature would have a less standard deviation, less distance between the mode and mean and less skewness. In other words, a selection giving taller children would be less variable and more nearly normal. Now as children grow older their stature increases, is less variable, and is more normal in its distribution. Thus, a selection of taller children from among children of eight would broadly tend to reproduce the characters of the stature distribution of older

children. In the same manner a selection of shorter children is more variable and less normal than the distribution of the general population of eight years of age, *i.e.*, tends to reproduce the characteristics of a younger population. Generally, a random selection, which increases variability, very sensibly increases skewness and decreases stature. What, perhaps, would hardly be expected, is that increase of skewness as well as increase of interval from mean to mode, *i.e.*, greater divergence from normality, increases the frequency of the mean stature.

It will be clear that by aid of this table we are able to predict the probable changes in all the other physical characters of the distribution when any sub-class has been selected at random from the general population with a difference of one character. If two or more characters differ in the sub-class, the probable changes in the other characters can be found by the principles of multiple correlation from the correlation table on page 307.

(22.) *Conclusion.*

This study of the probable errors and error correlations shows us that these quantities can be determined for the most complex system of organs in the case of normal correlation, and in the case of either normal or skew variation with considerable ease. It is only in the case of skew variation that the arithmetic becomes at all laborious. But numerical examples suffice to show that the errors here made are of the same order as in the case of normal variation, if we confine our attention to the characteristic features of the frequency, *e.g.*, the mean or modal frequency, the standard deviation, the skewness, &c. Certain constants of the algebraic form of the frequency curves have large probable errors, but these errors are so highly correlated, that their existence does not suffice to substantially modify either the form of the curve, or the "physical" characteristics of the distribution calculated from such values.

For the theory of evolution certain very important principles flow, beyond the mere advantage of knowing the probable errors made in the measurement of racial or organic characters. Above all we note the importance of a random selection in altering in a systematic manner all racial constants. In most cases even size cannot be altered without alteration of the size, variation and correlation of all correlated organs. This principle is developed more at length in a memoir, nearly completed, on the influence of *directed* selection, which covers as a special case that of random selection.

Later, we hope to apply the general theorem from which our memoir starts to determine the probable errors in the constants of the components into which a heterogeneous frequency distribution may be resolved by the method of the first memoir of this series.* It applies equally to such an investigation.

* The importance of such a determination was emphasized by Professor GEORGE DARWIN in the discussion which took place at the reading of that memoir.

[NOTE.—Added May 25, 1898. One point ought to have been more fully dealt with in the above memoir, namely, the probable error of the criterion $\kappa = 6 + 3\beta_1 - 2\beta_2$, upon which the selection of the type of the frequency depends. Clearly, if the probable error of this criterion is as large as the criterion itself, there can be no stability of type, or the frequency may change over from one type to another.

On page 289 we have found the standard deviation of the criterion in terms of known quantities for the curve

$$y = y_1 (1 + x/a_1)^{m_1} (1 - x/a_2)^{m_2}.$$

It is in fact given by the umbral equation

$$\chi_\kappa \Sigma_\kappa / \kappa = i_1 \Sigma_{m_1} \chi_{m_1} + i_2 \Sigma_{m_2} \chi_{m_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(clxv.)}$$

where i_1 and i_2 are functions of m_1 and m_2 given in (cxx.) and $\chi_{m_1}\chi_{m_2} = R_{m_1 m_2}$ is known from (cvi.).

The standard deviation of the criterion for the curve of type

$$y = y_0 \, e^{-\nu \tan^{-1}(x/a)} / \{1 + (x/a)^2\}^m$$

may be found by taking differentials of

$$\kappa = 6 + 3\beta_1 - 2\beta_2 = -\frac{12}{r-3}\left\{4 \sin^2 \phi \frac{r-1}{(r-2)^2} + 1\right\},$$

a value readily obtainable from 'Phil. Trans.,' A, vol. 186, p. 377. We thus find the umbral equation

$$\chi_\kappa \Sigma_\kappa = \left\{96 \sin^2 \phi \frac{r^2 - 3r + 1}{(r-3)^2 (r-2)^3} + \frac{12}{(r-3)^2}\right\} \Sigma_r \chi_r - \frac{96 \sin \phi \cos \phi \, (r-1)}{(r-3)(r-2)^2} \Sigma_\phi \chi_\phi \qquad \text{(clxvi.)}$$

$$= i_1' \Sigma_r \chi_r - i_2' \Sigma_\phi \chi_\phi, \text{ say,}$$

where Σ_r, Σ_ϕ and $\chi_r \chi_\phi = R_{r\phi}$ are given by (clii.) and (cliv.).

Applying these results to the numerical examples, we find:—

(a.) For the glands of swine

$$\frac{\Sigma_\kappa}{\kappa} \chi_\kappa = - \cdot 070{,}5459 \, \Sigma_{m_1} \chi_{m_1} - \cdot 038{,}4629 \, \Sigma_{m_2} \chi_{m_2},$$

whence the probable error of $\kappa = \cdot 67449 \, \Sigma_\kappa = \cdot 1012$; or,

$$\kappa = \cdot 5559 \pm \cdot 1012.$$

(b.) For the stature of children

$$\Sigma_\kappa \chi_\kappa = \cdot 015{,}6206 \, \Sigma_r \chi_r - \cdot 018{,}0067 \, \Sigma_\phi \chi_\phi,$$

whence the probable error of $\kappa = \cdot 67449 \, \Sigma_\kappa = \cdot 1919$; or,

$$\kappa = - \cdot 4330 \pm \cdot 1919.$$

In both cases, therefore, we may consider that the sign of κ is beyond question, or that the type selected is really a significant character of the frequency.

With regard to the probable error made in estimating a criterion to be zero, and using a curve of type

$$y = y_1 (1 + x/a)^{\nu} e^{-\gamma x},$$

we must remark that, the criterion being assumed zero is equivalent to assuming that its probable error is zero. Accordingly the only satisfactory method of testing whether a curve really falls under this type is to work out the probable error of its criterion on the hypothesis that it belongs to one or other of the two types, with positive or negative criterion as the case may be. If the probable error of the criterion thus calculated is sensibly as large as the criterion itself, then we may assume that the frequency distribution is of the type

$$y = y_1 (1 + x/a)^{\nu} e^{-\gamma x}.]$$