SOCIETIES AND ACADEMIES

LONDON

Royal Society, December 20, 1888.—"Correlations ant their Measurement, chiefly from Anthropometric Data." By Francis Galton, F.R.S.

Two organs are said to be co-related or correlated, when variations in the one are generally accompanied by variations in the other, in the same direction, while the closeness of the relation differs in different pairs of organs. All variations being due to the aggregate effect of many causes, the correlation is a consequence of a part of those causes having a common influence over both of the variables, and the larger the proportion of the common influences the closer will be the correlation. The length of the cubit is correlated with the stature, because a long cubit generally implies a tall man. If the correlation between them were very close, a very long cubit would usually imply a very tall stature, but if it were not very close, a very long cubit would be on the average associated with only a tall stature, and not a very tall one ; while, if it were *nil*, a very long cubit would be associated with no especial stature, and therefore, on the average, with mediocrity. The relation between the cubit and the stature will serve as a specimen of other correlations. It is expressed in its simplest form when the relation is not measured between their actual length, but between (a) the deviation of the length of the cubit from the mean of the lengths of all the cubits under discussion, and (b) the deviation of the mean of the corresponding statures from the mean of all the statures under discussion. Moreover, these deviations should be expressed on the following method in terms of their respective variabilities. In the case of the cubit, all the measures of the left cubit in the group under discussion, and which were recorded in inches, were marshalled in their magnitude, and those of them were noted that occupied the first, second, and third quarterly divisions of the series. Calling these measures Q_1, M, and Q_3, the deviations were measured from M, in terms of inches divided by $\frac{1}{2}(Q_3 - Q_1)$, which divisor we will call Q. Similarly as regards the statures. [It will be noted that Q is practically the same as the probable error.] This having been done, it was found that, whatever the deviation, y, of the cubit might be, the mean value of the corresponding deviations of stature was $0 \cdot 8y$; and conversely, whatever the deviation, y', of the stature might be, the mean value of the corresponding deviations of the cubit was also $0 \cdot 8y'$. Therefore this factor of $0 \cdot 8$, which may be expressed by the symbol r, measures the closeness of the correlation, or of the reciprocal relation between the cubit and the stature. The M and Q values of these and other elements were found to be as follows:[1]

---

[1] The head length is here the maximum length measured from the notch below the brow. The cubit is measured with the hand prone, from the flexed elbow to the tip of the middle finger. The height of knee is taken from a stool, on which the foot rests with the knee flexed at right angles ; from this the measured thickness of the heel of the boot is subtracted. All measures had to be made in ordinary clothing. The smallness of the number of measures, viz. 350, is of little importance, as the results run with fair smoothness. Neither does the fact of most of the persons measured being hardly full grown affect the main results. It somewhat diminishes the values of M, and very slightly influences that of Q, but it cannot be expected to have any sensible effect on the value of r.

left cubit, $18 \cdot 05$ and $0 \cdot 56$; stature $67 \cdot 2$ and $1 \cdot 75$; head length, $7 \cdot 62$ and $0 \cdot 19$; head breadth, $6 \cdot 00$ and $0 \cdot 18$; left middle finger, $4 \cdot 54$ and $0 \cdot 15$; height of right knee, $20 \cdot 50$ and $0 \cdot 80$; all the measurements being in inches. The values of r in the following pairs of variables were found to be: head length and stature, $0 \cdot 35$; left middle finger and stature, $0 \cdot 70$; head breadth and head length, $0 \cdot 45$; height of knee and stature, $0 \cdot 9$; left cubit and height of right knee, $0 \cdot 8$. The comparison of the observed results with those calculated from the above data showed a very close agreement. The measures were of 350 male adults, containing a large proportion of students barely above twenty-one years of age, made at the laboratory at South Kensington, belonging to the author.

These results are identical in form with those already arrived at by the author in his memoir on hereditary stature (Proc. Roy. Soc., vol. xl, p. 42, 1886), when discussing the general law of kinship. In that memoir, and in the appendix to it by Mr. J. Hamilton Dickson, their *rationale* is fully discussed. In fact, the family resemblance of kinsmen is nothing more than a special case of correlation.

The general result of the inquiry was that, when two variables that are severally conformable to the law of frequency of error, are correlated together, the conditions and measure of their closeness of correlation admits of being easily expressed. Let x_1, x_2, x_3, &c., be the deviations in inches, or other absolute measure, of the several "relatives" of a large number of "subjects," each of which has a deviation, y, and let X be the mean of the values of x_1, x_2, x_3, &c. Then (1) $y = r$X, whatever may be the value of y. (2) If the deviations are measured, not in inches or other absolute standards, but in units, each equal to the Q (that is, to the probable error) of their respective systems, then r will be the same, whichever of the two correlated variables is taken for the subject. In other words, the relation between them becomes reciprocal ; it is strictly a correlation. (3) r is always less than 1. (4) r (which, in the memoir on hereditary stature, was called the ratio of regression) is a measure of the closeness of correlation. Other points were dwelt upon in the memoir, that are not mentioned here : Among these was as follows : (5) The probable error, or Q, of the distribution of x_1, x_2, x_3, &c., about X, is the same for all values of y, and is equal to $\sqrt{(1 - r^2)}$ when the conditions specified in (2) are observed.

It should be noted that the use of the Q unit enables the variations of the most diverse quantities to be compared with as much precision as those of the same quantity. Thus, variations in lung-capacity which are measured in volume can be compared with those of strength measured by weight lifted, or of swiftness measured in time and distance. It places all variables on a common footing.

[*Nature* **39** (1889 January 3), 238.]