by calculating the value of $2\mu_2(3\mu_2^2 - \mu_4) + 3\mu_3^2$. If this quantity is greater than zero the range is limited at both ends; if equal to zero, it is limited at one end only; and if less than zero, it is unlimited.

Pearson has distinguished seven types of curves of frequency, the descriptions and formulæ of which are given in the preceding table. In order to determine to which type our curve belongs we must find the values of the following quantities:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3},\ \beta_2 = \frac{\mu_4}{\mu_2^2},\ \kappa_1 = 2\,\beta_2 - 3\,\beta_1 - 6,$$

$$\text{and } \kappa_2 = \frac{\beta_1\,(\beta_2 + 3)^2}{4\kappa_1(4\beta_2 - 3\beta_1)}.$$

We may call κ_1 and κ_2 the first and second *criteria.* Looking up these values in the table we find there the formula for the curve, from which the ordinates may be calculated. The second criterion serves to differentiate all of the curves except the normal and Type II., in both of which $\kappa_2 = 0$, but Type II. differs from the normal in that β_2 is not equal to 3. The ordinates of the normal curve may be calculated by the aid of the improved tables given by Sheppard (1903), but for the other curves there are no such tables. It would unduly expand this article if we should attempt to explain fully the formulæ for Types I. to VI. and give directions for calculating the values of y, and must, therefore, refer the reader for further details to Pearson's original articles (1896, 1901, and 1902). It should be noted, however, that our curve may be regarded as normal, even if it does not quite satisfy the conditions given in the table. According to Dunker (1899, p. 134), if the product of κ_1 multiplied by by μ_2^3 is within the limits of ± 1, and the value of $\dfrac{3\,\nu_2^2 - 2\,\nu_1^4\,(+\frac{1}{80})}{\nu_4}$ is between 0.8 and 1.25, unity being the value with a strictly normal curve, we may still regard our curve as of the normal type.

Skewness is measured by the deviation (d) of the mean from the mode expressed in terms of the standard deviation. Thus $\text{Sk} = \dfrac{d}{\sigma}$. When the value of the mean is greater than the mode, the skewness is said to be positive; when less, negative. The skewness is calculated by a different formula for each type of curve. But it is not difficult to estimate it by the rule given previously for estimating the mode, if we know the mean, median, and standard deviation.

In fact, there is considerable doubt as to how far mathematical analysis of the statistics of variation can be carried with profit. The types of curves grade into one another in such a way that it may happen that in some cases either one of two or more curves will fit the polygon equally well for practical purposes. Thus Miss Hefferan (1900) found a case in which dropping out one extreme individual in a series of four hundred changed the curve from Type IV. to Type I., and the two curves when plotted are practically identical. Pearson (1901, p. 453) gives a case in which two curves of Types V. and VI. plotted from the same data are so close as to be hardly distinguishable to the eye.

Correlation of Variations.—So far we have dealt with variations of single characters only. But every organism possesses many characters and organs, and these are related in many ways, both as to structure and function. Now there may or may not be some definite relation on the average between the variations of any given pair of associated characters.

In the language of Galton (1889): "Two variable organs are said to be correlated when the variation of the one is accompanied on the average by more or less variation of the other, and in the same direction." This is *positive correlation.* When the variation of the second organ is in the opposite direction from that of the first, we have *negative correlation.* Many correlations are matters of every-day observation. Thus the right leg is generally of very nearly the same length as the left leg. There is a similar but less close correlation between the lengths of arms and legs, a man with light hair usually has blue eyes, and so on.

We may find correlations also between similar characters in associated individuals, as fathers and sons (see *Heredity*), or husbands and wives; also between characters of organisms and factors in their environment or their previous treatment or condition, as between color of plumage and latitude of habitat, susceptibility to disease and age, immunity against smallpox and vaccination.

Such correlations are clearly of great theoretical importance for the study of evolution and heredity, and they have very practical bearings upon medical and social questions.

Therefore it is important that we should have some means of expressing the amount of correlation in quantitative terms. This means has been furnished by Galton (1889) who showed how to find the *coefficient of correlation*, or what is often called Galton's function.

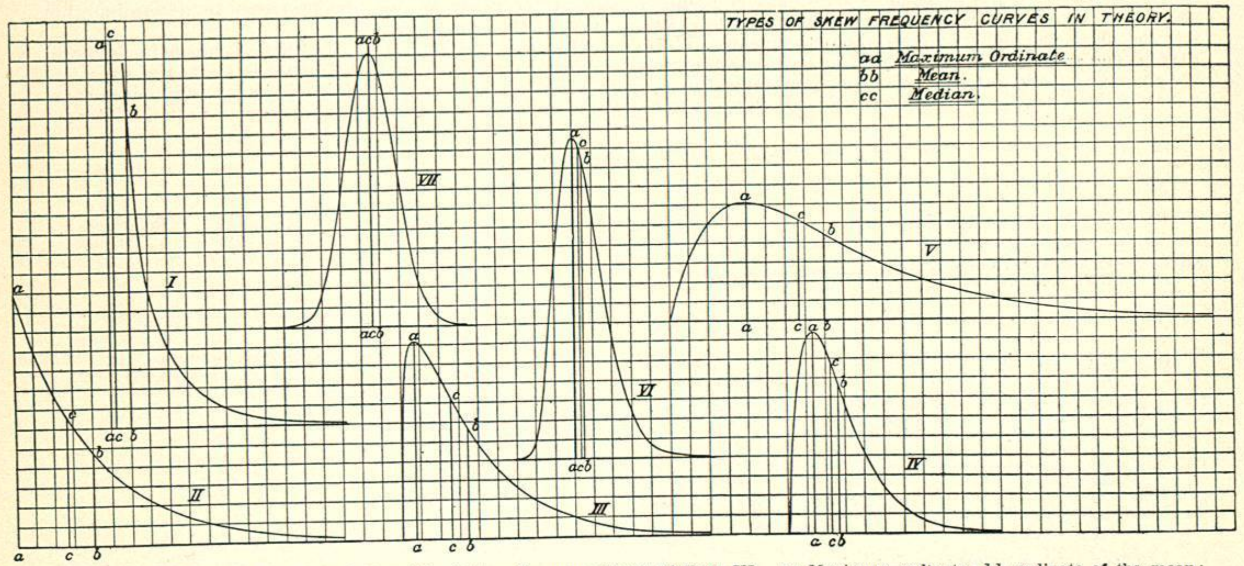

FIG. 4986.—The Various Sub-Types of Transitional Skew Frequency Curve of Type III. *aa*, Maximum ordinate; *bb*, ordinate of the mean; *cc*, ordinate of the median. (From Pearson.)

Suppose we have measurements of the length and breadth of the heads of one thousand men. The data may be arranged in a *correlation table*, like the one given below. This is constructed in the following manner: At the top is a horizontal scale of length, and we will call this the selected character, or "subject." Likewise at the left is a vertical scale of breadth, which is the associated character, or "relative." The entries in the body of the table are the frequencies of individuals in which this pair of characters have the values given in the corresponding positions in the horizontal and vertical scales, respectively. For example, there were 21 men whose head length was 7.1 inches. Of these 1 had a width of 5.7 inches, 4 were 5.8, 4 were 5.9, 6 were 6 inches broad, and so on. The columns and rows of frequencies are added up at the bottom and at the right. Each column forms an *array*, which shows the variations of breadth associated with the value of the length given at the head of the column. In the same way breadth may be taken as the subject, and then the rows of frequencies are the arrays, giving the variations of the associated length.

the axis of x; and the position of the mean of all the head breadths is similarly represented by the intersecting horizontal line, the axis of y. Now the horizontal distance of a circle from the axis of x will give the deviation (x) of the selected character, and the vertical distance of the same circle from the axis of y will give the mean deviation (y) of the associated array from the mean of all the arrays. Now if x and y are expressed in equivalent units, the ratio $\frac{y}{x}$ will give a measure of the correlation of the two characters in this particular array. If the ratio $\frac{y}{x}$ were the same in all the arrays, the zigzag line in Fig. 4987 would become a straight line. As it is, we see that it has a general trend which is represented by a straight line that may be drawn so that the sum of the distances of the small circles from it shall be as small as possible. This line will meet the axes of x and y at their point of intersection. Now the slope of this line may be measured in the same way that the grade of a

CORRELATION TABLE. HEAD LENGTH ON HEAD BREADTH, 1,000 CAMBRIDGE MEN. (From Macdonell.)

	Head Length in Inches.																			Totals.
Head Breadth (Inches).	6.8	6.9	7	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8	8.1	8.2	8.3	8.4	8.5	8.6	
5.5	..	..	..	..	1	..	2	..	..	..	..	..	..	..	..	..	..	..	..	3
5.6	..	..	..	..	2	3	1	4	2	..	..	..	..	..	..	..	..	..	..	12
5.7	1	1	..	1	2	6	5	7	5	8	3	4	..	..	..	..	..	..	..	43
5.8	..	..	2	4	6	7	15	12	12	11	7	3	1	..	..	..	..	..	..	80
5.9	..	..	1	4	6	16	16	24	20	28	7	5	4	..	..	..	..	..	..	131
6.0	..	1	..	6	14	12	24	40	45	40	28	9	13	4	..	..	..	..	..	236
6.1	..	..	1	3	4	6	11	27	42	40	22	15	10	2	2	..	..	..	..	185
6.2	..	..	1	2	6	3	14	18	22	25	13	14	17	2	4	1	..	..	..	142
6.3	..	..	1	..	1	4	9	8	16	15	15	16	6	7	..	..	..	..	1	99
6.4	..	..	..	..	1	..	2	4	6	6	3	4	7	2	1	1	..	..	..	37
6.5	..	..	..	..	..	1	1	..	3	3	..	2	3	1	1	..	..	..	..	15
6.6	..	..	..	1	..	..	..	1	2	3	..	1	..	2	1	1	..	..	..	12
6.7	..	..	..	..	..	..	..	..	..	1	..	..	..	2	..	..	..	..	..	3
6.8	..	..	..	..	..	..	..	..	..	..	..	2	..	..	..	..	..	..	..	2
Totals.	1	2	6	21	43	58	100	145	175	180	98	75	61	22	9	3	..	..	1	1,000

The means and standard deviations of the subject and relative may be calculated from the total frequencies given at the bottom and right side of the table, respectively, by the methods previously described. In the same way the mean may be determined for each array. Now transferring our horizontal and vertical scales to plotting paper, the positions of the means of the arrays may be plotted as in Fig. 4987, which is made from a correlation table of head breadth and stature in three thousand criminals. In this diagram the position of the mean of an array, is indicated by a small circle, and the

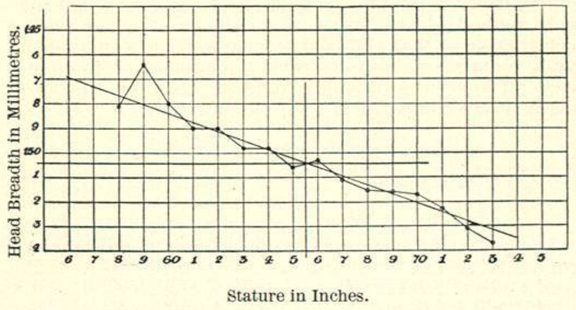

FIG. 4987.—Diagram of Correlation Table. Stature, subject; head breadth, relative. Three thousand criminals, mean height = 65.5 inches, mean head breadth = 150.4 mm. Coefficient of regression = 0.3612. (From Macdonell.)

circles are connected by a zigzag line. The position of the mean of all the statures is represented by a vertical line drawn through the middle of the diagram: call this

railway track is measured, by the rise or fall in a certain unit of horizontal distance, as ten feet in a mile, or in our case by the height y_a of any point a at the distance x_a from m_1 (Fig. 4988). The slope of the line may be

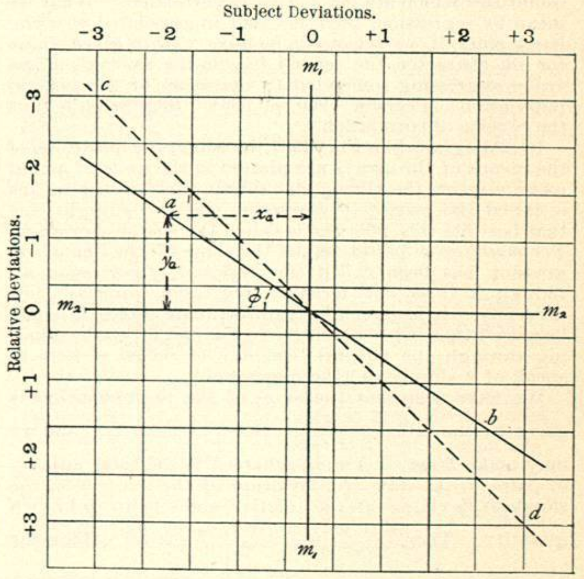

FIG. 4988.—Diagram of a Correlation Table with Deviations Plotted in Terms of Standard Deviation. m_1, Position of mean of subject; m_2, mean of relative; *a b*, regression line.

measured also by the angle ϕ which it makes with the line m_2, and the ratio $\frac{y_a}{x_a}$ is what mathematicians call the tangent of that angle, $\frac{y_a}{x_a} = \tan\phi$. It may be shown by the principle of least squares that, taking all the observations into consideration, the most probable value of $\tan\phi$ will be $\tan\phi = \frac{\Sigma\,(x\ y\ z)}{\Sigma\,(x^2\ z)}$, where x and y are the deviations of each pair of observations from the means of the subject and relative, respectively, z is the number of pairs, or frequency of each class in the arrays, and Σ is the sign for summation. But in a diagram like Fig. 4987 the ratio $\frac{\Sigma\,(x\ y\ z)}{\Sigma\,(x^2\ z)}$ will not give a constant measure of the correlation between the two characters, because, in the first place, x is expressed in inches while y is in millimetres. If we changed these units, we would change the ratio. Moreover, stature and head breadth have different standard deviations, so that the diagram is longer in one direction than in the other, and the slope of the diagonal would differ according as to whether head breadth or height were taken as the subject.

Galton overcame this difficulty by expressing all the deviations of each character in terms of its probable deviation. Then his correlation table took the form represented in Fig. 4988, when $\frac{y_a}{x_a}$ gives the coefficient of correlation no matter which character is taken as subject. If there is perfect correlation the line *ab* will coincide with *cd* and $\frac{y_a}{x_a} = 1$. But if there is no correlation *ab* coincides with m_2 and $\frac{y_a}{x_a} = 0$. So the coefficient of correlation is a fraction which may have any value between 0 and 1, in Fig. 4988 its value is $\frac{2}{3}$, and, if A be the selected character and B the relative, then this coefficient expresses the probable proportion of the mean deviation of B to the corresponding selected deviation of A. In general the deviations of B will be found to be less than the corresponding deviations of A. Galton first noticed this phenomenon in his early studies of heredity and he called it *regression.*

But regression is not peculiar to heredity. It is found in all statistics concerned with two associated variable quantities which are not perfectly correlated. What we mean by regression, then, is "that in associated or correlated pairs, if we select one member with a given value for its character, the second has on the average a less value, regressing somewhat to the mean of the general population" (Pearson, 1900, p. 394). Regression is thus the reverse of correlation.

In a diagram like Fig. 4987, in which the positions of the means of the arrays are plotted in the units of actual measurement, the zigzag line drawn through the means is called the *polygon of regression*, and the straight line that best fits this polygon is called the *line of regression*. For want of a better name, the slope of the line of regression has been called the *coefficient of regression*, although it does not measure regression directly, but inversely. It is a peculiarly unfortunate choice of name, because it has led to much confusion and misunderstanding through the natural tendency of recent writers to speak of it simply as "the regression."

We have seen that the slope of the regression line is given by the ratio $\frac{\Sigma\,(x\ y\ z)}{\Sigma\,(x^2\ z)}$. But $\Sigma\,(x^2 z) = N\sigma_1^2$ and we may make $\Sigma(xyz) = Nr\sigma_1\sigma_2$, where N is the total number of pairs, σ_1 the standard deviation of the subject, σ_2 the standard deviation of the relative and r is an unknown quantity. Then $\frac{\Sigma\,(x\ y\ z)}{\Sigma\,(x^2\ z)} = \frac{Nr\,\sigma_1\,\sigma_2}{N\sigma_1^2} = r\frac{\sigma_2}{\sigma_1}$. Thus the formula for the *coefficient of regression* becomes $r\frac{\sigma_2}{\sigma_1}$, and then it may be shown that $r =$ the *coefficient of correlation.*

They both measure the same phenomenon. The former gives the probable ratio of the deviations as they are actually observed. The latter gives a corrected value of this quantity after due allowance has been made for differences in units of measurement and in standard deviation. In cases in which $\sigma_1 = \sigma_2$, then $\tan\phi = r\frac{\sigma_2}{\sigma_1} = r$.

There is one more property of the correlation surface, as it is called, that should be noted. If x and y are plotted in terms of σ_1 and σ_2, as in Fig. 4988, and the distribution of frequencies is normal, then when $r = 0$, a series of contour lines drawn through the classes of equal frequency will be perfect circles. When the regression line, *ab*, moves from m_2 toward *cd*, the contour lines become ellipses with *cd* as their major axis. As *ab* approaches *cd* the ellipses become more and more narrow until when *ab* coincides with *cd*, that is, with perfect correlation, the frequency of each array would be concentrated along the line *cd*. Now the range of each array is measured by its standard deviation, which with perfectly normal distribution would be given by the formula $\sigma_3 = \sigma\sqrt{1 - r^2}$. It is evident that as r approaches the value of 1, $1 - r^2$ will become smaller and smaller until finally $\sigma_3 = 0$, that is, until the distribution of frequencies of the arrays is concentrated along the major axis of the ellipses. This shows that the accuracy with which one can use the value of a character to predict the value of its associated character will depend to a great extent upon the amount of correlation existing between them.

There are two methods at present in use for calculating the coefficient of correlation. The first method depends on the formula $r = \frac{\Sigma\,(x\ y\ z)}{N\sigma_1\,\sigma_2}$, which has been shown by Pearson to give the best value for r. The formula indicates the method of procedure, which is simple but laborious. Having constructed a correlation table and found the means and standard deviations of subject and relative, the deviations x and y for each array class are found, and then multiplied together and by the frequency of the class. All these products are added and the sum is divided by the product of the whole number of pairs multiplied by the standard deviations of subject and relative. The result is the coefficient of correlation.

Dunker (1899, p. 154) has devised a modification of this method, which is intended to lessen the labor, and is fully described by Davenport (1899, p. 33).

The second method is more complex in theory, but is shorter in practice. Theoretically it is not so good as the first, but it has the advantage of saving a great deal of time and labor, and it is the only method that can be used with characters that are not quantitatively measurable. It was for this purpose that the method was invented by Pearson (1901). Macdonell (1902) has made a careful comparison of the two methods, and concludes that the new method gives a value of r with an accuracy fairly comparable to that obtained from an ordinary frequency table containing about one-third the number of frequencies. The results obtained in three cases by the two methods are compared in the following table:

COMPARISON OF THE TWO METHODS.

	Coefficient of Correlation.	
	First method.	Second method.
Head length and head breadth	.4016 ± .0103	.3977 ± .0176
Head breadth and stature	.1831 ± .0119	.1811 ± .0210
Left middle finger and stature	.6608 ± .0069	.6633 ± .0142

The theory of this second method is explained in untechnical language by Pearson in the "Grammar of Science" (1900, pp. 432–436). We will proceed at once with a description of the manner in which the theory is applied in the calculation of r.

The material is arranged in a fourfold correlation table like the following:

CORRELATION OF STATURE AND BREADTH OF HEAD. 3,000 CRIMINALS. (Macdonell.)

Head breadth (centimetres).	Stature in Feet and Inches.		Totals.
	5′ 4$\frac{9}{16}$″ and under.	Over 5′ 4$\frac{9}{16}$″.	
14.8 and under	455	622	1,077
Over 14.8..........	599	1,324	1,923
Totals...........	1,054	1,946	3,000

CORRELATION OF PREVIOUS TREATMENT AND RESULT OF SMALLPOX. HOMERTON AND FULHAM HOSPITALS (DOUBTFUL CASES EXCLUDED). (Macdonell.)

	Recoveries.	Deaths.	Totals.
Vaccinated..............	8,207	692	8,899
Unvaccinated	1,424	1,103	2,527
Totals.................	9,631	1,795	11,426

Let the correlation table be represented by the following diagrams:

TABLE OF FREQUENCIES.

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

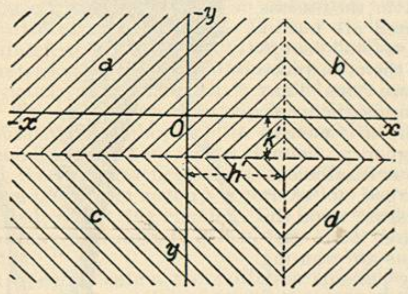

FIG. 4989.—Diagram Showing the Theoretical Relation of the Table of Frequencies to a Normal Correlation Surface. The shaded rectangles a, b, c, d represent the divisions of the table of frequencies, h' and k' the deviations of the lines of division from the means of subject and relative represented by the axes of y and x. (From Pearson.)

Calculate the values of a_1 and a_2 by the formulæ

$$a_1 = \frac{(a + c) - (b + d)}{N}.$$

$$a_2 = \frac{(a + b) - (c + d)}{N}.$$

Now we need a table of the "probability integral," which gives the "sums from the beginning" of the normal curve of frequency. In "Biometrika," vol. ii., pp. 182–190, is a series of tables of this kind prepared especially for our purpose. Turning to Tables III. and IV., pp. 189–190, we find in the a column our value of a_1 and in the x column read off the corresponding tabular entry, interpolating if necessary, and the value thus obtained is called h. Continuing to the z column we find the tabular entry and call that H. In the same way for a_2 we find the corresponding entries under x and z and call them k and K.

Then find ε by the formula $\varepsilon = \frac{ad - bc}{N^2\,H\,K}$, and r, the coefficient of correlation, may be found from the equation,

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

(Pearson, 1901*a*, p. 6).

Or, if we wish to be a little more exact, we may find r by the following equations:

$$\varepsilon = \theta + \tfrac{1}{2}hk\theta^2 - (h^2 + k^2 - h^2k^2)\frac{\theta^3}{6} + hk\left\{h^2k^2 - 3(h^2 + k^2) + 5\right\}\frac{\theta^4}{24} + \text{etc.}$$

$r = \sin\theta$ (Fig. 4990). (Pearson, *loc. cit.*, p. 8.)

The equations for finding r and θ may be solved by Newton's method of approximation, or better by Horner's method (see Wells, "College Algebra," p. 520); or Burnside and Panton, 1892, p. 212). In the second equation the value of the angle θ is given in radians, ratio of arc to radius. As $\frac{1}{2}\pi = 90°$, the value of θ in seconds is given by the equation $A = 206{,}265\,\theta$ (De Morgan, 1902, p. 276). Reducing A to degrees, minutes, and seconds, $r = \sin\theta$ is found easily in any table of natural sines (*e.g.*, Jones, "Logarithmic Tables," pp. 60 *et seq.*). A still easier method of finding r from θ is to take a table like Jones's Table II. (*loc. cit.*, pp. 15–19) where the angles are given both in degrees and in radians, and $\log r = \log\sin\theta$ can be read off directly, and r, found from a table of logarithms.

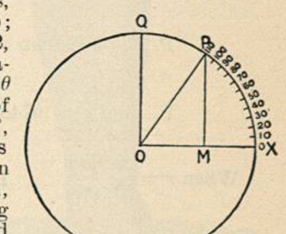

FIG. 4990. — Illustrating the Meaning of $r = \sin\theta$. If the angle $POX = \theta$ is measured by the arc PX, then $r = \sin\theta = PM$, OP being taken as unity. (From Pearson.)

Pearson has given (1902 *b*, pp. 13–14) a graphical method for fitting a straight line to a set of observations that seems applicable in the plotting of regression lines and the calculation of coefficients of correlation. But it is doubtful if this method would have any advantage over the one just described.

It will be interesting to note at this point some of the results that have been obtained by the study of correlation. The following coefficients of correlation were obtained by Macdonell from three thousand criminals:

Correlations	Coefficient.
Stature and head length..............................	0.33993
Stature and head breadth.............................	.18308
Stature and face breadth	.34527
Stature and finger......................................	.66084
Stature and cubit	.79986
Stature and foot...	.73636

From the results of which these form a part he has arrived at important conclusions concerning the identification of criminals.

Some correlations of especial interest to physicians are:

DIPHTHERIA. Pearson (1901)—Correlation of antitoxin treatment and recovery.................$r = 0.2451 \pm 0.0205$

SMALLPOX. Macdonell (1902)—Effectiveness of vaccination and strength to resist the disease.

For six towns	$r = 0.6561 \pm 0.0092$	Doubtful cases included in the vaccinated.
For Sheffield	.7694 ± .0124	
For Leicester...............	.6112 ± .0728	
For Gloucester	.5897 ± .0198	
For Homerton and Fulham .	.5760 ± .0089	
For Homerton and Fulham.	.6615 ± .0083	Doubtful cases excluded.
For Glasgow................	.7783 ± .0365	
For London, 1901......	.5779 ± .0311	
For London, 1892-93........	.5954 ± .0272	

Precision of the Constants.—The constants of our curves of variation and correlation—mean, median, standard deviation, coefficient of correlation, and the like—are the predictions made from observations as to the probable condition of the given character in the whole population. Our data are like the balls drawn from the urn which we cannot see into nor upset. We can only guess at what is in the urn by what has been drawn out of it. We have seen that the probability of our guess being right depends upon the number of observations. This is expressed by the, so-called, probable errors of the constants, which, preceded by the ± sign, are written after the constants, as in the examples given in previous sections. We will now give the formulæ for calculating the probable errors (P.E.) of some of the most important constants. When distribution is normal:

$$\text{P. E. of mean} = .6745\frac{\sigma}{\sqrt{N}}.$$

$$\text{P. E. of median} = .8454\frac{\sigma}{\sqrt{N}}.$$

$$\text{P. E. of } \sigma = .6745\frac{\sigma}{\sqrt{2N}}.$$

$$\text{When } r = \frac{\Sigma(xyz)}{N\sigma_1\sigma_2}. \qquad \text{P. E. of } r = .6745\frac{1-r^2}{\sqrt{N}}.$$

When r is calculated by the second method, its P. E. has a very complex formula (Pearson, 1901, p. 14). But, according to Macdonell, it will be about the same as if r were calculated by the first method with one-third the number of pairs. Yule gives a table (1897, p. 854) of the P.E. of r by the first method for various values of N.

Biologists are indebted to Elderton for a useful table ("Biometrika," vol. i., part ii.), by which it is possible to determine whether the differences between a polygon of frequency and the corresponding theoretical curve are greater or less than the probability of random sampling should lead one to expect. If they are not greater, the curve chosen may be regarded as a satisfactory representation of the relative frequencies of the variations in nature.

We will now bring to a close this description of the methods employed in statistical biology, which we believe to be the first one that is both fairly complete and in a language that can be understood by the ordinary American biologist or physician; and we venture to express the hope that it may lead to a wider investigation in the fields that these new methods make available for study.

Robert Payne Bigelow.

Bibliographical References.

Bateson, W.: Materials for the Study of Variation, London, 1894.
Burnside, W. S., and Panton, A. W.: Theory of Equations, ed. 3, Dublin, 1892, pp. 212–227.
Davenport, C. B.: Statistical Methods with special reference to Biological Variation, New York, 1899.—On the Variation of the Shell of *Pecten irradians* Lamarck from Long Island. Amer. Nat., vol. 34, 1900, pp. 863–877.
Darwin, C.: Variation of Animals and Plants under Domestication, London, 1868.
De Morgan, A.: On the Study and Difficulties of Mathematics, 2d reprint ed., Chicago, 1902.
Dunker, G.: Die Methode der Variationsstatistik. Arch. f. Entw. d. Org., vol. 8, 1899, pp. 112–183.
Elderton, W. P.: Tables for Testing the Goodness of Fit of Theory and Observation. Biometrika, vol. 1, 1902, pp. 155–163.
Galton, F.: Correlations and Their Measurement, Chiefly from Anthropometric Data. Proc. Royal Soc., vol. 45, 1889 (*a*), pp. 136–145.—Natural Inheritance, London, 1889 (*b*).
Goodman, J.: Mechanics Applied to Engineering, London, 1899, pp. 1–19 and 50–105 (Moments).
Hefferan, Mary: Variation in the Teeth of Nereis. Biol. Bull., vol. 2, 1900, pp. 124–143.
Laplace, P. S. de: Philosophical Essay on Probabilities. Trans. by Truscott and Emory, New York, 1902.
Ludwig, F.: Ueber Variationskurven und Variationsflächen der Pflanzen. Bot. Centralbl., vol. 64, 1895.
Macdonell, W. R.: On Criminal Anthropometry and the Identification of Criminals. Biometrika, vol. 1, 1902 (*a*), pp. 177–227.—On the Influence of Previous Vaccination in Cases of Smallpox, *ibid.*, 1902 (*b*), pp. 375–383.
Pearson, K.: Skew Variation in Homogeneous Material. Phil. Trans. A., vol. 186, 1896, pp. 343–414.—Grammar of Science, ed. 2, London, 1900.—On the Correlation of Characters Not Quantitatively Measurable. Phil. Trans. A., vol. 195, 1901 (*a*), pp. 1–48.—Supplement to a Memoir on Skew Variation. Phil. Trans. A., vol. 197, 1901 (*b*), pp. 443–459.—On the Modal Value of an Organ or Character. Biom., vol. 1, 1902 (*a*), pp. 260–251.—On the Systematic Fitting of Curves and Observations and Measurements. Biom., 1902 (*b*), vol. 1, pp. 265–303, and vol. 2, pp. 1–23.—On the Fundamental Conceptions of Biology. Biom., vol. 1, 1902 (*c*), pp. 320–344.
Quetelet, A.: Lettres . . . sur la théorie des probabilités, appliquée aux sciences morales et politiques, Bruxelles, 1846.—Letters on Probabilities. Trans. by Downes, London, 1849.
Sheppard, W. F.: New Tables of the Probability Integral. Biom., vol. 2, 1903, pp. 174–190.
Vernon, H. W.: Variation in Animals and Plants, London, 1903.
Vries, H. de: Ueber halbe Galtonkurven als Zeichen diskontinuirlichen Variation. Ber. Deut. bot. Ges., vol. 12, 1894, pp. 197–207.
Weldon, W. F. R.: Address to the Section of Zoology. Nature, vol. 58, 1898, pp. 499–506, also Report Brit. Assn. Adv. Sci., 68, 1899, pp. 887–902.
Yule, G. U.: On the Theory of Correlation. Jour. Royal Statistical Soc., vol. 60, 1897, pp. 812–854.

VARICELLA. See *Chickenpox*.

VARICOCELE. See *Sexual Organs, Male, Diseases of*.

VARICOSE VEINS.—Etiology.—Dilated and tortuous veins are spoken of as varices, or varicose veins. This abnormal condition is dependent in some way upon an interference with the flow of blood in the veins. Thus gravity plays its part, since the trouble is most common in the veins of the legs, and since it is more often seen in tall people than in short ones. Age and disease of the heart, producing imperfect valvular action, are also potent factors. As instances of local venous obstruction may be mentioned abdominal tumors, pregnancy, and the wearing of constricting clothing, such as circular garters. The left spermatic vein empties into the left renal vein at a right angle, being crossed by the sigmoid flexure. The pressure exerted upon it by an accumulation of fæces is thought to be one reason why varicocele is common on the left side.

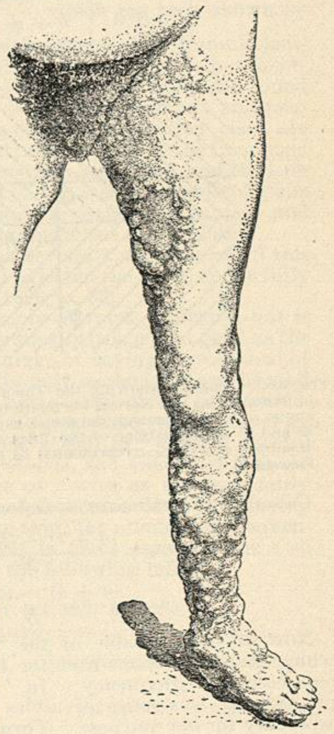

Fig. 4991.—Varicose Veins of the Lower Limbs and Genital Organs in the Case of a Woman soon to be Delivered of her Tenth Child. (Duplay et Reclus: "Traité de Chirurgie.")

However, the reasons just stated are not of themselves sufficient to explain the occurrence of varicose veins. Thus the veins begin to dilate in the early months of pregnancy before the uterus is sufficiently enlarged to exert any pressure upon the iliac veins, while varicose veins are wanting in many patients with large abdominal tumors. Nor is the change due to the atrophy of old age, as it is usually noticed before the fortieth year, and in a great many instances before the twenty-fifth year. It is equally difficult to explain why the veins of one leg should be much dilated and tortuous, while those of the other are scarcely affected; or why in certain persons the large veins are chiefly affected, and in others the smaller radicles in the skin. As muscular contraction and muscular relaxation constitute the chief aids to venous flow, it is not surprising that persons who are obliged to stand for hours at a time (washerwomen, car motormen, etc.) suffer more from this trouble than those who are constantly changing their position.

While any of the veins of the lower extremity may be affected, the trouble is most often found in some of the radicles of the internal saphenous vein. The tortuous and dilated vessels are easily recognized on inspection, and if they lie near the surface their bluish color is visible through the skin. If they are not or have not been inflamed, they collapse under moderate pressure and disappear almost entirely when the foot is elevated. In some places their channels through the skin feel like a break in its continuity; while here and there they lie so near the surface that they seem to be covered with little more than epithelium. The deep veins of the leg may be affected as well as the superficial ones. If the dilatation goes on to a still greater degree a distinct thin-walled sac filled with fluid blood is formed. Such a large varix is shown in Fig. 4992.

As a result of the dilatation many of the valves become insufficient. This can be tested in the following manner: After the veins of the lower extremities have been emptied by elevation of the foot and stroking of the limb toward the body, the thumb is placed upon the main saphenous trunk and the patient is directed to stand. The varicose veins will fill slowly and only to a moderate degree. The moment the thumb is removed the column of venous blood falling into them from above instantly distends them to their fullest capacity if the valves are incompetent. Such are the conditions of simple varicosity. Sooner or later one or more complications are likely to arise, such as thrombosis, rupture, periphlebitis, œdema, eczema, and ulcer.

Complications.—Thrombosis may occur in dilated veins of the leg exactly as it may occur in the dilated vein of an external hemorrhoid. It is accompanied by a good deal of pain and tenderness, by slight redness and by œdema plainly limited to the immediate vicinity of the vein involved. Thus if a not very tortuous vein be affected for a distance of five or six inches its course can be accurately mapped out as an indurated strip about three-fourths of an inch wide. If the vein is tortuous, the indurated area will have an irregular outline.

The nutrition of the parts drained by varicose veins is often seriously affected, so that a wound may become infected. The result may be erysipelas, cellulitis, abscess, or suppurative thrombophlebitis, although the last-mentioned condition is by no means common.

When the vein lies near the surface it is easily ruptured by a blow from a sharp object, and as there is little elastic tissue about the opening, the hemorrhage is profuse and may be serious if it is not stopped by pressure or ligation.

A fourth complication, more often seen in older individuals, is an extensive œdema. At first this is of the usual type, revealing itself by pitting on pressure; but after it has existed for many months the production of fibrous tissue may be sufficient to prevent much indentation on pressure. This condition may be due to other causes than varicosity of the veins, and it greatly interferes with the nutrition of the parts, and especially with the repair of a chronic ulcer, whether varicose or not.

Eczema is another complication due to an imperfect nutrition which is apt to lead to ulcers starting in the small scratches made by the patient in the vain attempt to relieve himself from the intolerable itching.

Not every ulcer occurring in a patient whose veins are varicose is to be attributed to such varicosity. A long-standing ulcer of the leg of a non-malignant, non-syphilitic, non-tuberculous character is better spoken of as a chronic ulcer. It may be the direct or indirect result of varicose veins, but it may also be due to traumatism or eczema, or œdema, or anæmia. It is misleading to call all such ulcers varicose ulcers. They are all due to poor local nutrition, of which varicosity of the veins is simply one cause.

When varicose veins have existed for a considerable time there will often be noted a brown pigmentation of the skin, occurring more or less in patches, and due either to small subcutaneous ruptures of the venous radicles or to transudation of red blood cells through the dilated venous walls. In either case the blood pigment becomes permanently fixed in the fibrous tissue of the skin, giving it a characteristic yellow-brown color.

Symptoms.—Varicose veins often give rise to no symptoms whatever. Such is apt to be the case in young and

Fig. 4992.—Varicose Veins of Both Legs with a Large Varicose Vein below the Left Knee and a Varicose Ulcer just above the Right Ankle. (Von Bruns.)

healthy persons, and also when the veins are dilated throughout a small area—for example, over one-half of the vulva, or around the saphenous opening.

The symptoms in uncomplicated cases are: A sense of weight and more or less dull aching relieved by elevation of the affected extremity.

The symptoms of thrombosis are: Marked tenderness on pressure, acute local pain, which is considerably but not wholly relieved by a recumbent position, and a rise of temperature of one, two, or three degrees.

The symptoms of the inflammatory complications mentioned above are such as accompany these processes wherever they occur in the body.

The diagnosis of varicose veins and of the different complications above enumerated is easy for any one who is able to recognize the different forms of inflammation and ulceration. The inflamed strip of a varicose vein is broader than that of infection in a lymph vessel, and the overlying skin is not so red. A sharply localized varix can hardly be mistaken for any sort of a cystic growth, as it is collapsible on pressure, refills slowly, and has less tension than most cysts. It does not pulsate like an aneurism.

Treatment.—Palliative treatment of uncomplicated varicose veins consists in attention to the general health; in the avoidance of such occupations and such clothing as tend to interfere with the venous flow; in the elevation of the feet as much as possible when the patient is sitting down, and in the wearing, during the daytime, of an elastic bandage or stocking extending from the toes to the knee. Even though the varicose veins extend well into the thigh, firm compression of the leg will usually relieve the symptoms, and it is difficult to apply a stocking or bandage with comfort above the knee-joint.

A woven elastic, cotton or silk-web stocking costs from $2.50 to $8. Its advantages are the ease with which it may be applied and the firm pressure which it exerts when new. A thin white lisle thread stocking should be