

STATISTICAL ANALYSIS OF THE DATA

In the statistical treatment of the data the possible desirability of subsequent analytical comparisons other than merely those of averages and standard deviations has been kept in mind. The peculiarities of a population with respect to the attributes of its component individuals cannot adequately be appraised statistically without a knowledge of at least four of the parameters of the frequency distributions involved. This means, in terms of the method of moments of Pearson, the calculation of the first four moments of each frequency distribution. These yield respectively the *mean*, the *standard deviation*, β_1 and β_2 . The first and second of these statistical constants are familiar to most readers, being the usual quantitative expressions of the average size and of the normal variability of a character respectively. From the latter, the standard deviation, are derived the probable errors of the various constants. In order to provide a *relative* measure of variability, thus ruling out the influence of the differences in the units of absolute measurement, the coefficient of variation (V) is also calculated; this being merely the percentage ratio of the standard deviation to the mean.

The last two of the constants mentioned above, β_1 and β_2 , are derived respectively from the additional calculations of the third and fourth moments of the frequency distribution, and provide the bases of measures of *skewness* or asymmetry of the distribution on the one hand, and of a factor known as *kurtosis* or relative peakedness or flatness of the frequency curve at the mid-point on the other. Although these last two constants have not as yet been calculated for many human populations, nevertheless they are as essential to a complete description of human groups as they are for other biologic groups. The direction of skewness is determined by the sign of the third moment (μ_3) which is tabled immediately after the value of β_1 .

The criteria for the so-called normal distribution are $\beta_1=0$ and $\beta_2=3$. The significance or non-significance of observed deviations from these values is to be judged on the basis of the size of the derived measures of skewness or kurtosis (cf. formulae in accompanying table) as compared with the indicated probable errors of these constants for the normal curve.

In addition to these purely descriptive statistical constants, the coefficients of correlation between the various absolute measurements and the chief linear measure of total body size, namely the stature, have also been calculated and are included in the tabulations presented below.

Constants	Symbols	Formulae	Probable Errors .67449 x
Mean	M	$\frac{\sum fx}{N}$	$\frac{\sigma}{\sqrt{N}}$
Standard Deviation	σ	$\sqrt{\mu_2} = \sqrt{\frac{\sum fx^2}{N}}$	$\frac{\sigma}{\sqrt{2N}}$
Coefficient of Variation	V	$\frac{\sigma \times 100}{M}$	$\frac{V}{\sqrt{2N}} \left(1 + 2 \left(\frac{V}{100} \right)^2 \right)^{\frac{1}{2}}$
β_1	β_1	$\frac{\mu_{\frac{3}{2}}}{\mu_{\frac{2}{2}}}$	$\sigma \sqrt{\frac{2}{N}}$
β_2	β_2	$\frac{\mu_4}{\mu_{\frac{2}{2}}^2}$	$\sigma \sqrt{\frac{6}{N}}$
Skewness	X	$\frac{\sqrt{\beta_1}(\beta_2+3)}{2(5\beta_2-6\beta_1-9)}$	$\sqrt{\frac{3}{2N}}$ (for normal curve)
Kurtosis	η	β_2-3	$\sqrt{\frac{24}{N}}$ (for normal curve)
Coefficient of Correlation	r_{12}	$\frac{\Sigma(x_1x_2)}{N \sigma_1\sigma_2}$	$\frac{(1-r)}{\sqrt{N}}$

Σ stands for summation; f, frequency; x, deviation from mean ($x' - x$). Sheppard's corrections for grouping are included in the calculation of the various moments.

Although it may be assumed that the reader is either familiar with or is able to look up the various formulae involved in the calculation of the statistical constants mentioned nevertheless the working forms of these formulae are assembled in the above table for reference or information.

NON-METRIC SOMATOLOGICAL OBSERVATIONS

Data based on non-metric observations possess a slightly different order of importance than those derived from metric measurements. It is unfortunate that at the present time there exists no international agreement in regard to the selection and standards for such qualitative somatological observations. Within the rather wide limits of a generally accepted range of such observations, the selection of the particular observations to be made and the terminology to be used in their recording in any particular case are matters depending largely upon the personal interest and experience of the observer. In the present instance the following observations have been made. The accompanying glossary of descriptive terms used is added for convenience of reference, and the findings are tabled in the last table of the series (table 14).

NON-METRIC OBSERVATIONS AND GLOSSARY OF DESCRIPTIVE TERMS

CHARACTERS	DESCRIPTIVE TERMS
SKIN: COLOR FRECKLES SCRATCH TEST	WHITE—YELLOW—pale, tawny, dusky YELLOW-BROWN—BROWN PRESENT—ABSENT POSITIVE—NEGATIVE
HAIR: COLOR (1) Head CHARACTER (2) Body AXILLA TORSO PUBIS	BLACK—BROWN—BLACK—Mixed with grey STRAIGHT—WAVY NONE—SCANTY—MEDIUM—MARKED