ANTHROPOLOGICAL OBSERVATIONS ON THE ANGLO-INDIANS OF CALCUTTA.

PART I. ANALYSIS OF MALE STATURE.

By PRASANTA CHANDRA MAHALANOBIS, *B.Sc., B.A. (Cantab), Indian Educational Service, Professor of Physics, Presidency College, Calcutta.*

(Plates I—IV.)

INTRODUCTORY NOTE.

The people with whom these papers will deal are those officially called "Anglo-Indians" in India. They are not, however, the Anglo-Indian of English literature and common parlance in which the term is applied to persons of English, or rather British, birth who have spent a considerable part of their lives in India. Some years ago the Government of India, seeking to avoid the associations that had grown up round the name Eurasian, decided that persons of mixed Indian and European blood should be known henceforth as Anglo-Indians.[1] The word Eurasians had itself been invented to avoid a coarser and more descriptive term. That even the more recent designation was inaccurate in point of fact was pointed out at the time of its introduction in a letter published in a Calcutta newspaper and signed "Franco-Burman." The term Indian, indeed, had been stretched to include all native denizens of the Indian Empire—Burmese, Baluchis, etc., as well as Indians properly so-called; while it had been forgotten that any other European nation but the English had ever had a part in India.

The observations on which Professor Mahalanobis' analyses are based had their origin as follows. Ever since I began to take a serious interest in anthropometry, I have had doubts as to the value of bodily measurements taken on the living person. So long ago as 1903,[2] I pointed out that my own measurements of the faces of the people of the Faroe Islands were completely at variance with those of a previous observer, and attributed the different results mainly to slight difference in technique. The working out of the measurements of the various tribes of the Malay Peninsula obtained in 1901-1902[3] by Mr. H. C. Robinson and myself increased my doubts, and further made me suspicious

---

[1] I understand, however, that as early as 1830 the term Anglo-Indians had already been applied to persons of mixed descent.

[2] Annandale, *Proc. Roy. Soc. Edinburgh* XXV, pp. 2–24 (1903).

[3] Annandale and Robinson, *Fascicule Malayenses*, Anthropology (1903—1904).

that there was some inherent falacy in the whole method. These measurements were taken with special care, each individual being measured three times over and most by two observers. Although they showed the gross differences in head-measurements between the civilized and the uncivilized tribes, they failed completely to demonstrate differences between the heads of the Negrito and of the Indonesian jungle tribes.

Having in 1916 an opportunity of examining a number of Anglo-Indians anthropometrically, I determined to see whether my doubts were further justified by the investigation of a race known to be of recent mixed origin. Before discussing the methods adopted, I must say a few words about my subjects. They were with very few exceptions, young men between the ages of 18 and 40, and with few exceptions belonged to what I may call the middle class of so-called Anglo-Indians, mostly employed as clerks, mechanical engineers, overseers and so forth, or else fresh from school and about to take up employment of the kind. The fact is of importance, for social distinctions are somewhat rigidly maintained in this community. I am indebted to Mr. H. A. Stark, late Principal, Dacca Training College, now Principal Armenian College, Calcutta, for valuable information on the po
Among the Anglo-Indian community of Calcutta some families claim descent from Mahommedan ladies of noble and even princely birth, who in the old days entered into alliances of a perfectly regular kind from a Mahommedan point of view with Englishmen of good birth. These families are, however, comparatively few. At the other end of the social scale are the "Kintalis",[1] whose origin is thus described by Mr. Stark in a lecture on "Calcutta in Slavery Days" read before the Calcutta Social Study Society on March 13th, 1916.

"The liberated slaves [who, as Mr. Stark had previously explained, were mainly Indians but included not a few Negros] unbeknown to themselves that they had been doing what the Manumitted Roman slaves had done centuries before, in gratitude assumed the surnames of their late masters. Their descendants, for the most part, survive in the "Kintal" population of the city."

If this were a full statement of the case, it might be doubted whether the Kintalis have any real claim to be of mixed race, unless there is some slight admixture of Negro blood; but, as in all cities, there is a tendency for certain individuals of the more respectable classes to sink down to the slums and become a part of the submerged population, which is represented in Calcutta, so far as the Christian communities are concerned, by the Kintalis.

Be this as it may, few or no Kintalis are among the persons I measured, and probably none of very old family. So far as possible, moreover, we have eliminated from the measurements

1 The name is derived from the lodging-houses (*Kintal*) in which many of these people live or lived. The word Kintal, however, now means little more than a slum inhabited by low-class Christians.

analysed those of persons known to have recent Negro or Mongoloid blood, *i.e.* persons one of whose parents or grandparents was a Negro or belonged to a Mongoloid stock. This has been a necessary precaution, because the number of individuals in which the further complexity was introduced was large enough to affect the results without being sufficiently numerous to afford a sound basis for mathematical treatment. So far as recent Negro blood was concerned I was fairly confident in accepting the statements of those who offered themselves for measurement, as certain, not by any means all, Negro traits were present. I refer particularly to woolly hair, dark complexion, negroid nose and prognathism. The long lower limb and slender shin of the Negro, which is of a different type from that of the Indian, were not perpetuated in a single individual.[1] As to old Negro blood, no definite information was obtained.

To eliminate the recent Mongoloid element from our investigations was, however, a much less easy task and I am by no means sure that this has been done successfully. Here again I had to trust to the statements of individuals measured, but Mongoloid traits are often reproduced in a much more subtle manner than Negroid, and the Mongoloid element in the population of Calcutta is much larger than the Negroid. Indeed, I have observed that many of the most intelligent Anglo-Indians with whom I have had dealings have had distinctly Mongoloid features. This is not surprising, for the offspring of women of the various Mongoloid tribes of the Himalayas, Assam and Burma, who are not generally averse to unions of a more or less permanent nature with educated Europeans settled in their districts, are not only of respectable parentage in both lines but often receive a good education, and Calcutta is the natural goal of such people. So far as I could discover, it is unusual for an Anglo-Indian to know much of his family for more than two or three generations back and at the present time, in Calcutta at any rate, most of the community are the result of marriages of persons of mixed blood.[2]

The subjects of my investigations were, therefore, mainly of mixed Indo-European blood, probably in many individuals with some Mongoloid admixture, but not affiliated with the higher Hindu castes.

The measurements were taken in the zoological laboratory of the Indian Museum in the years 1916—1919. I had the help of

---

[1] As only about half a dozen Anglo-Indian-Negros were examined, I have refrained from giving details and merely cite the results for what they are worth. Recent Negro settlers in Calcutta are mostly West Indians. They and their families occupy a street practically by themselves.

[2] I may here note that further complexity is now being introduced into the Anglo-Indian community by the marriage of Anglo-Indian women to Canton Chinese, who are now numerous as cabinet-makers and bootmakers in Calcutta. These men keep themselves entirely apart from the Indian communities and frequently marry Anglo-Indians, though the custom of bringing their wives from China is becoming much common than it was a few years ago.

several assistants, among whom I may mention in particular my late laboratory assistant Mr. J Caunter, to whom I was indebted for obtaining many of my subjects. Dr. F. H. Gravely and Dr. K. S. Roy devoted much time and labour to helping me. The investigations were conducted in a less systematic manner than I would have wished, partly because they were in themselves of the nature of an experiment and I was perpetually attempting to discover more satisfactory methods, and partly because they had to be carried out at odd times, chiefly on Sundays and holidays, when subjects were available. The measurements that have been utilised by Prof. Mahalanobis were, however, made on one system and with the same instruments. The system was that recommended in the British Association's hand-book on anthropology and the instruments were the "Anthropometer" (112) and "Instrumentantascher" (203) supplied by Hermann of Zurich.

Prof. Mahalanobis has, in my opinion wisely, decided to treat the measurements as accurate only within 2 mm. He notes a tendency on my part to favour even numbers. Of this I was barely conscious at the time, but on attempting to reconstruct the process in my mind I seem to recollect that when I was not quite sure of a measurement within a millimetre, I had a prejudice in favour of even numbers. I never thought it possible to measure to within less than a millimetre. It is curious, however, that this prejudice seems to have communicated itself to my assistants, by several of whom the measurements were occasionally taken while I noted them down. That it has done so is evidence at any rate of uniformity of method.

The measurements, discussed without knowledge of mathematics, seemed to me so unsatisfactory that I had practically decided to reject them altogether, until I was so fortunate as to get into touch with Prof. Mahalanobis at the Nagpur meeting of the Indian Science Congress and he offered to analyse them statistically. The results he has already obtained seem to justify their publication, and to emphasize the value of co-operation and co-ordination of different branches of scientific work in anthropology, without which, in my opinion, further progress in most branches of biology has become impossible.

The special importance of investigations conducted on the Anglo-Indians lies in the fact that although we may not be able to trace out the history of any one family, we know that the whole race, if such it may be called, has arisen practically within the last 200 years by the admixture of other pre-existing races. After Prof. Mahalanobis has discussed my measurements on mathematic lines, I hope to have an opportunity of considering other aspects of the somatology of this interesting community We hope thus to throw some light on the question of the origin of human races by fusion.

N. ANNANDALE,

Director, Zoological Survey of India, Calcutta.

PART I.

ANALYSIS OF MALE STATURE.

CONTENTS.

SECTION I. GENERAL REMARKS.

In the present paper I have attempted a statistical examination of Anglo-Indian Stature based on Dr. Annandale's records. The measurements were all taken by Dr. Annandale or in a few cases under his direct supervision. Thus the present material may be considered free from large fluctuating errors due to different personal bias of different observers.

NATURE OF THE MATERIAL.

Dr. Annandale has explained in his introductory note the special character of the present material. After excluding "Negro," "West Indies" "Chinese," "Burmese" and "Bhutia" ancestry and omitting certain incomplete and doubtful records a series of 200 was obtained for Stature, Head Length, Head Breadth, Nasal Length, Nasal Breadth, Zygomatic Breadth and Upper Face Length.[1]

The great importance of the present material from a biometrical standpoint will be easily appreciated. So far as I am aware this is the first time that a true biologically mixed population is being studied by statistical methods.

From the statistical standpoint the coefficient of variability is considered to be a very important test of homogeneity.[2] Hitherto all attempts to fix the upper limit of homogeneous variability were necessarily confined to the study of artificially made up mixtures.[3] The Anglo-Indian data furnish us with a "natural mixture." A careful study may be expected to throw considerable light on this vexed question. Incidentally, it will be of great interest to compare the variability of such a "mixed" population with those of "purer" races.[4]

The Anglo-Indian population may really represent a new "race" in the making, and we hope to discuss in the sequel what indications may be afforded by a study of the present material as regards the mechanism of race formation.

It should be noted however that the word "race" is here used in its statistical sense. Pearson[5] says, "Any race may originally have arisen from a mixture of races, but such a mixed race is wholly different from a mixture of races, which have not interbred."

---

[1] Arithmetical work on these characters is nearly finished and I hope to publish the results at an early date.

[2] This is true of course for uni-modal data only, or more generally for distributions which cannot be dissected into component frequency groups. For a fuller discussion of this point see pp. 34, 93-94.

[3] C. S. Myer's—Man, February, 1903, pp. 28–32. Also see Karl Pearson's discussions on this point in *Biometrika* Vol.2, 1903, pp. 345–347, Myers' Reply and Pearson's Remarks on this Reply in *Biometrika* Vol. 2, 1903, pp. 504-508.

[4] "Purer" in a statistical sense, i.e. more homogeneous.

[5] *Biometrika* Vol. 2, 1903, p. 506.

The special significance of the present material is that it does represent a mixed race which has interbred and whose component races are still in a pure form.

Plan and Scope of the Paper.

Dr. Annandale took a very large number of measurements extending to forty different characters. But the records are not complete in each case. As I have already mentioned a series of 200 has been obtained for seven[1] metric characters. A second group[2] consists of about 120 to 180 and a third[3] of 50 to 100 complete records. In addition eye and skin colour were recorded, as also observations on hairyness in all cases.

In the present paper the frequency distribution and variability of stature has been discussed at some length. Certain points have been considered in great detail, much of which it will not be necessary to repeat in subsequent parts.

The second part (material for which is nearly ready) will contain a study of the frequency distribution and variability of individual organs included in the first group. Correlation between the organs of the first group will be next discussed and after that the study of the second and the third group will be taken up. Finally I hope to discuss the distribution and correlation of eye, hair and skin-colour in a separate paper.

I should make my position quite clear; I frankly confess that I know very little of anatomy. My work on the data supplied has been purely statistical.

Some of the results may appear to be thoroughly unconventional or sometimes perhaps even startling in character. With such a short series, it is of course impossible to lay emphasis on the numerical value of any particular constant. But I would like to draw the attention of Anthropologists to statistically significant magnitudes as not unworthy of careful study I have contented myself with pointing out statistical results and have refrained from drawing Anthropological conclusions.

The chief object of the present study is to invite the attention of Physical Anthropologists of India to the importance of the application of accurate statistical methods to their "crude" measurements. As some of the technical terms may be unfamiliar

---

Stature, Head Length, Head Breadth, Nasal Length, Nasal Breadth, Zygomatic Breadth, Upper Face Length.

[2] (i) Gonial breadth 181. (ii) Frontal breadth 142. (iii) Shoulder breadth 171. (iv) Thigh breadth 171. (v) Height of knee-joint, inside 174. (vi) Height of knee-joint, outside 120. (vii) Height of middlle finger 132. (viii) Styloid height 167. (ix) Trochanter height 180. (x) Iliac height 175. (xi) Upper radius height 118. (xii) Suprasternal height 119. (xiii) Acromion height 181. (xiv) Leg length 174. (xv) Chest, extended 137.

[3] (i) Total face length 93. (ii) External orbital breadth 93. (iii) Ocular breadth 91. (iv) Distance between eyes 87. (v) Chest, depressed 88. (vi) Kneeling height 87. (vii) Sitting height 93. (viii) Earhole height 87, (ix) Span of arms 93. (x) Cubit 87. (xi) Hand length 76. (xii) Humerus length 48. (xiii) Radius length 48. (xiv) Foot length 78. (xv) Foot breadth 78.

to Anthropologists, I have thought it advisable to include short explanatory notes, which would have been unnecessary in a purely Biometrical paper.

I must also offer my apologies to the trained statistician. Much of the work will no doubt appear to him to be quite superfluous. I would remind him that one of our objects has been to persuade the Anthropologists to adopt statistical methods. This has necessitated detailed consideration of certain points which may appear obvious to a trained statistician.

For example, a very full discussion of the effect of grouping has been given. All frequency constants were calculated several times over with very different units of grouping. It is then shown that the effect of grouping is quite negligible within very wide limits—a result which is of course quite familiar to all statisticians.[1] But as I found very wide-spread popular misapprehension regarding this point I have considered it desirable to give an actual empirical demonstration of the above fact. The discussion of various "correction" for grouping will have its own interest to the statistician.

Another consideration has guided me in this introductory paper. Any extension of a scientific method to new material requires caution. Our Anglo-Indian data cannot be assumed to be homogeneous in character, hence I have thought it desirable to justify *empirically* the application of statistical methods to such mixed data as the present material. The assumption of "normality" (i.e. of approximately Gaussian distribution) thoroughly permeates many important statistical methods. It was therefore necessary to investigate the question of frequency distribution in great detail.

The arithmetical labour has been very great specially as I did not have any modern calculating machine to help me. This want of mechanical accuracy may have introduced some uncertainty in the arithmetical results and this is why I have quoted the arithmetic very fully in order to facilitate checking by others. In the case of important "moments," I have checked them absolutely by working with different start points (i.e. different base numbers).

This is my first venture into the province of Biometry and it is not unlikely that I have made mistakes. I have included full details of the statistical work in the hope that competent Biometricians will kindly help me by pointing out errors. I have retained six places of decimal in the arithmetic, not in the vain hope of reaching an impossible degree of accuracy, but for convenience of checking. It is difficult to attain agreement to the second place in the final results unless about six figures are retained in the intermediate calculations in this type of work.

---

[1] K. Pearson, "Errors of Judgment &c." *Phil. Trans. Roy. Soc.* Vol. 198A (1902) "Assortative Mating in Man." *Biometrika* Vol. 2, 1903, p. 485. The authors note that "the system of grouping adopted is within wide limits immaterial."

I have intentionally made the present analysis very elaborate. A total of only 200 observations did not perhaps merit such close scrutiny. As there was no early prospect of increasing this total considerably, I thought it better to complete even a provisional investigation thoroughly rather than wait indefinitely for a larger sample. But the chief reason which prompted me to make an *intensive* study of the small available amount of material is this, so far as I am aware no work in this line has been done in India, no Anthropologist in India has ever made any use of the modern statistical calculus associated specially with the name of Karl Pearson and the Biometric School. The present study is intended to illustrate the urgent necessity of the application of statistical methods to Anthropology. The conclusions based on only 200 observations cannot of course claim any degree of finality. But these serve to show the kind of results which can be reached by statistical methods and also show the great scope and huge possibilities of statistical methods.

REMARKS ON THE APPLICATION OF STATISTICAL METHODS.

Before proceeding to the more systematic part of the work I wish to make a few general observations on the application of Statistical methods. I cannot do better than begin by quoting some remarks of Charles Goring in this connection.[1]

"Statistical enquiry, all scientific enquiry, is observational in character: that is to say, it is based upon the observation of individual facts. But these facts, in themselves, do not constitute knowledge. Knowledge consists in the discovery of relationships revealed by the systematic study, and by the legitimatised weighing of facts."

"No series of biological or social observations constitutes knowledge *in itself*. Knowledge lies potential in the facts, but ineffectual for use until their associations with each other have been accurately weighed. It is the weighing of observations which demands for the present enquiry, the employment of statistical methods: such methods being merely a regulated mechanism by which the relation between certain order of facts can be precisely determined."

"There is not, as is sometimes imagined, any *special* theory or hypothesis involved in conclusions revealed by statistics. The science of statistics provides only for the systematised study and legitimatised interpretation of observed facts: such interpretation consisting mainly in one and the same process—the associating or dissociating one set of facts with and from another. Before any association can be legitimately postulated, certain conditions must be fulfilled; evidence must be produced to show that the relation, affirmed to exist, is not a chance or accidental, but a natural asso-

[1] Charles Goring, The English Convict, pp. 19-20 (H.M.S.O. 1913)

ciation; that it is not one resulting from coincidence, but that it represents an inseparable connection between natural phenomena."

"The attributes and conditions of living things are so widely variable, are so delicately graduated in different individuals that their correlation can seldom be legitimately postulated, and can never be precisely estimated, without aid from a correlation calculus: that is to say, social science almost entirely, and biological and medical sciences to great extent, can only be built up after preliminary mathematical analysis of large series of carefully collected data" This is the reason why we assert that statistical methods are indispensable for our present enquiry.

We have got Anthropometric measurements of 200 Anglo-Indians as our material in the present case. We know that this constitutes only a very *small sample* of the whole Anglo-Indian population. We wish to investigate the Anthropometric characteristics of the whole population but we are constrained to do so from a study of the sample alone. If the sample exhibits certain typical features. we shall be justified in *inferring* the presence of these typical features in the general population. Thus our first statistical task is to find out the typical features of our sample. In order to do so, it is necessary to describe the given sample by means of a suitable typical curve, that is, to graduate the given sample suitably.

This very process of graduation itself will "smooth out" the irregularities peculiar to the particular sample considered. Hence when a typical formula is once obtained we get rid of the special individual peculiarities of the given sample and can replace the given sample by our graduated curve in all subsequent discussions. This graduated curve is, by logical induction, assumed to be typical of the whole population.

This typical frequency curve is defined by certain statistical constants [1] calculated from the measurements actually given in the sample. The reliability of each constant is determined by the internal consistency or uniformity of the particular set of measurements from which it is derived (and the total number of measurements). The reliability (measured by the probable error) can be precisely calculated with the help of the statistical calculus based on the theory of probabilities.

Thus in any statistical enquiry the first part of the work consists in the determining of the appropriate frequency constants and their probable errors. This is done in section II of the present paper, which also contains an elaborate technical discussion of the effect of grouping.

The next part of our work consists in constructing a type which is assumed to be true for the general population, within the limits of the probable error of the type. This is the problem discussed in section IV

---

[1] I have given a short account of some of these constants in non-technical language in Appendix I. pp. 90—94.

Once the typical curve is built up we can proceed to comparison with other general populations as represented by their own typical formulae. Goring observes "no valid comparison between two series of statistics is possible until the constants of each series have been determined." [1]

But even then, no conclusion can be safely asserted from the comparison, until a certain condition has been fulfilled. "Before drawing conclusions from the comparison of statistics, we must be certain that we are dealing with *strictly random samples of the same homogeneous material*" (italics mine).

This introduces the second part of our work. For valid comparison we must investigate the homogeneity (or otherwise) of our material. I have discussed the statistical tests of homogeneity in section III, and the application of these tests in section V

We then pass on to the question of comparison with other data. In section VI, I have considered the nature of the material for comparison and in the next section (section VII) I have investigated the question of comparative homogeneity in great detail.

In section VIII, I have added a preliminary note on the variation of stature with age. I shall discuss the question of age correlation and growth in a later paper.

---

[1] *Cf.* Goring, p. 33. "In order that complex groups such as two series of measurements, may be compared, these have to be reduced to a simple form, to the genius, as it were, of the series, i.e. certain values, called constants (the mean, mode, standard deviation, etc.), have to be extracted; and the groups compared through the medium of their constants. These values, however, are only themselves comparable in certain conditions. First, we must know that the statistics they represent are not chaotic in their distribution that the sequence of their frequencies have been determined by law. And, secondly, we must know the range of error to be discounted before any actual differences between the constants compared may be regarded as significant. Before we can assert that one series of measurements inherently differs from another, we must predict and allow for a certain amount of difference or arithmetical inexactness, which, according to the law of probability, is bound to appear in limited samples of the same homogeneous material. This predicted amount of insignificant difference is called, as we have already said, the probable error of the constants under consideration."

"Briefly resumed the matter stands thus: we must compare, not this or that particular measurement, but the whole series of measurements obtained from a random sample of (one population) with a similar whole series obtained from a random sample of (another) population. In order to make this comparison two things will be necessary: we must extract from each series its statistical constants, the mean, the standard deviation, etc., of the series: and by the theory of probability, we must determine for each constant obtained, its probable error. These constants, with their probable errors, will be the representatives of the series, which, through their medium, become comparable with each other. If the differences between the results compared are not greater than the probable errors of these results, such differences may be regarded as insignificant: if the difference is not greater than twice the probable error, it may be regarded as probably insignificat; and if it is not greater than three times the probable error, it may be regarded as possibly insignificant. On the other hand, if any difference found is greater than three times the probable error, it is reasonable to assume that the difference is due to some definite influence over and above those causes which are inherent in the sampling process."

The raw material in the form of the actual measurements, has been included in Appendix II.

"Tables," throughout the present paper, have reference to the indispensable volume edited by Karl Pearson, "Tables for Statisticians and Biometricians" (Cambridge University Press, 1914).

NOTE ON "BIAS" IN RECORDING MEASUREMENTS.

It is well known that different observers are affected with different 'personal bias' in taking measurements. In the present case the crude data showed an overwhelming preponderance of "even" readings as against 'odd" measurements.

In the case of Stature, we find no less than 193 "even" reading as against only 7 "odd." We have no reason to believe that Nature has any special preference for "even" number of millimeters, hence, apart from personal bias and fluctuations due to random sampling we should have had 100 "even" and "odd" readings each. Instead of this, we actually get 193 and 7.

The presence of "bias" is obvious, but I have calculated the "Contingency" [1] for the whole group of the above seven measurements.

TABLE I.

Contingency for "bias."

Organ.	Theoretical value.	Observed value.	$m-m'$.	$\left(\frac{m-m'}{m}\right)^2$
Stature	100	193	93	86·49
Head Length	100	174	74	54·76
Head Breadth	100	181	81	65·61
Nasal Length	100	111	11	1·21
Nasal Breadth	100	93	7	0·49
Zyg. Breadth ..	100	156	56	31·36
Upper Face Length	100	105	5	0·25
$n'=7$				$\chi^2=240·17$

The probability that "random sampling" would lead to as large or larger deviation between theory and observation is given by

$$P = e^{-\frac{1}{2}x^2}\left\{1+\frac{x^2}{2}+\frac{x^4}{2\cdot4}\right\}$$

$$\log P = -\tfrac{1}{2}x^2 \log_{10}e + \log\left\{1+\frac{x^2}{2}+\frac{x^2}{2\cdot4}\right\}$$

[1] Karl Pearson: Phil. Mag. Vol. L, pp. 157-175, 1900.

$$\log P = -\frac{240{\cdot}17}{2}\log_{10}e + \log_{10}\left\{1 + \frac{240{\cdot}17}{2} + \frac{576{\cdot}96}{2{\cdot}4}\right\}$$

$$\log P = -52{\cdot}15225289 + 3{\cdot}868282$$

$$= \bar{47}{\cdot}713029$$

Thus $P = 5{\cdot}17 \times 10^{-47}$, or the chances are 2×10^{46} to 1 against there being no bias.

In the case of Stature the unit of grouping is greater than 10 mm. and hence this preponderance of even values of millimetres is not a matter of great consequence.

SECTION II. EFFECT OF GROUPING ON THE FREQUENCY CONSTANTS.

FREQUENCY CONSTANTS AND PROBABLE ERRORS.

The object of the enquiry contained in this section may be best explained in Karl Pearson's words.[1]

"It is well known that if the distribution of errors follows the normal law, the "best" method of finding the mean is to add up all the errors and divide by their number, the "best" method of finding the square of the standard deviation is to form the squares of the deviations from the mean and divide by their number. . . These "best" methods become far too laborious in practice when the deviations run into hundreds or even thousands. The deviations are then grouped together, each group containing all deviations falling within a certain small range of quantity, and the means, standard deviations, and correlations are deduced from these grouped observations. If the means, standard deviations, and correlations be calculated from the grouped frequencies as if these frequencies were actually the frequency of deviations coinciding with the midpoints of the small ranges which serve for the basis of the grouping, we do not obtain the same values as in the cases of the ungrouped observations. It becomes of some importance what corrective terms ought to be applied to make the grouped and ungrouped results accord. This point has been considered by Mr. W F Sheppard (who has proposed certain corrections). Thus corrected the values of the constants of the distribution as found from the ungrouped and grouped deviations will nearly, but not of course absolutely, coincide."

In this section I have calculated both ungrouped and grouped constants with widely differing units of grouping. The constants as corrected by Sheppard's formulae have also been calculated in each case. By a comparison of the different constants we find that within very wide limits the effect of grouping is negligible.

The Stature list was classified into groups of 50 mm. The base number is taken to be 1655 mm. and the moment coefficients were calculated as shown below.[2]

We get the following table for "raw" moments about 1655:—

---

[1] Karl Pearson: "On the Mathematical Theory of Errors of Judgment and on the Personal Equation," *Phil. Trans. Roy. Soc.*, Vol. 198A, 1902, pp. 249, 250.

[2] For details, see K. Pearson: "On the Systematic Fitting of Curves, etc." Part I, *Biometrika*, Vol. I, 1902, pp. 265—303 and Vol. II, 1902, pp. 1—24. Also W. Palin Elderton "Frequency Curves and Correlation," pp. 13—19. (C. and E. Layton, 1917) and G. Udney Yule: "Theory of Statistics" (Charles Griffin & Co.)

GROUP (mm.)	Mid-Ordinate.	x	y=Frequency.	xy	$x^2 \cdot y$	$x^3 \cdot y$	$x^4 \cdot y$	$x^5 \cdot y$	$x^6 \cdot y$
1430—1480 .	1455	−4	3	−12	48	−192	7 68	−30 72	1 22 88
1480—1530 .	1505	−3	5	−15	45	−135	4 05	−12 15	36 45
1530—1580 .	1555	−2	14	−28	56	−112	2 24	−24 48	8 96
1580—1630 .	1605	−1	45	−45	45	−45	45	−45	45
1630—1680 .	1655	0	60	−100		−484		−47 80	
1680—1730 .	1705	+1	48	48	48	48	48	48	48
1730—1780 .	1755	+2	20	40	80	160	3 20	6 40	12 80
1780—1830 .	1805	+3	3	9	27	81	2 43	7 29	21 87
1830—1880 .	1855	+4	2	8	32	128	5 12	20 48	81 92
				+105		+417		+34 65	
TOTAL .			200	+5	381	−67	25 65	−13 15	2 85 81

Dividing by the total, 200, we get for the "raw" moments, S denoting a summation for all groups.

$$\nu_1' = S\frac{(xy)}{N} = + \quad \cdot 025$$

$$\nu_2' = S\frac{(x^2y)}{N} = + \quad 1 \cdot 905$$

$$\nu_3' = S\frac{(x^3y)}{N} = - \quad 0 \cdot 335$$

$$\nu_4' = S\frac{(x^4y)}{N} = + \quad 12 \cdot 825$$

$$\nu_5' = S\frac{(x^5y)}{N} = - \quad 6 \cdot 575$$

$$\nu_6' = S\frac{(x^6y)}{N} = + 142 \cdot 905$$

The true Mean is given by

$$1655 + (\cdot 025 \times 50) = 1656 \cdot 15 \text{ mm.}$$

Transferring[1] to the true Mean with the help of ;—

$$\mu_2 = \nu_2' - \nu_1'^2$$
$$\mu_3 = \nu_3' - 3\nu_1'\nu_2' + 2\nu_1'^3$$
$$\mu_4 = \nu_4' - 4\nu_1'\nu_3' + 6\nu_1'^2\nu_2' - 3\nu_1'^4$$
$$\mu_5 = \nu_5' - 5\nu_1'\nu_4' + 10\nu_1'^2\nu_3' - 10\nu_1'^3\nu_2' + 4\nu_1'^5.$$

[1] Karl Pearson: "Contributions to the Mathematical Theory of Evolution—On the Dissection of Asymmetrical Frequency-curves," *Phil. Trans. Roy. Soc.*, Vol. 185 A (1894), p. 71.

we get moments about Mean (without correction)

$$\mu_2 = 1{\cdot}90\ 43\ 75$$
$$= -\ {\cdot}47\ 78\ 43\ 77$$
$$\mu_4 = 12{\cdot}86\ 56\ 42\ 58$$
$$\mu_5 = -8{\cdot}18\ 04\ 93\ 98.$$

The moments were checked by calculating the "raw" moments about 143·0 cm. (end of range) as base unit The "raw" moments were

$\nu_1' = -4{\cdot}52\ 5$, $\nu_2' = -22{\cdot}38$, $\nu_3' = -118{\cdot}02\ 625$, $\nu_4' = -657{\cdot}42\ 75$, $\nu_5' = -3846{\cdot}6203125$,

but after transferring to the Mean, the same values as before were obtained.

The Standard Deviation [1] (S.D.) is given by $\sigma = \sqrt{\mu_2}$

Thus $\sigma = +\ 1{\cdot}38$ in working units

$= +69{\cdot}00$ mm.

The Coefficient of Variation [2] (V) is defined by $\frac{100\sigma}{M}$ and we get

$$V = 4{\cdot}1660.$$

We must now proceed to find the other frequency constants [3]

$\beta_1 = \mu_3^2/\mu_2^3$ $\quad \beta_1 = {\cdot}033204$

$\beta_2 = \mu_4/\mu_2^2$ $\quad \beta_2 = 3{\cdot}547534$

Skewness = Sk. = ·069858

where skewness $= \dfrac{\sqrt{\beta_1}(\beta_2 + 3)}{2(5\beta_2 - 6\beta_1 - 9)}$.

Distance between Mode and Mean $= d = \sigma \times$ skewness.

It is now necessary to find the Probable Errors.[4]

---

[1] Also See Appendix I.

[2] Karl Pearson: "Regression, Heredity and Pan-mixia," *Phil. Trans., Roy. Soc.* Vol. 187A (1896), p. 203. See footnote on p. 34.

[3] (i) Karl Pearson:—"Skew Variation in Homogeneous Material," *Phil. Trans., Roy. Soc.* Vol. 186A (1895), pp. 343—414, Supplement, Vol. 197A (1901), pp. 443—459.

(ii) Karl Pearson: "On the Mathematical Theory of Errors of Judgment," *Phil. Trans., Roy. Soc.* Vol. 198A (1902), pp. 274—279 and p. 277.

(iii) "Skew Frequency Curves," *Biometrika*, Vol. 4 (1905), pp. 169—212; *Biometrika* Vol. 5 (1906), pp. 168—171 and pp. 172—175.

(iv) W Palin Elderton:—"Frequency Curves and Correlation" (Charles and Edwin Layton, London) with Addendum and Errata, 1917.

[4] The fundamental memoirs are Karl Pearson and L. N. G. Filon: (*a*) "On the Probable Errors of Frequency Constants and on the Influence of Random Selection on Variation and Correlation," *Phil. Trans. Roy. Soc.*, Vol. 191A (1898), pp. 229—311.

(*b*) W F. Sheppard: "On the application of the Theory of Error to cases of Normal Distribution and Normal Correlation," *Phil. Trans. Roy. Soc.*, Vol. 192A (1899), pp. 101—167.

The Probable Error of Mean [1]

$$= \frac{\cdot 6744898}{\sqrt{n}} \sigma = \chi_1 \sigma.$$

Probable Error of Standard Deviation

$$= \frac{\cdot 6744898}{\sqrt{2n}} \cdot \sigma = \chi_2 \cdot \sigma.$$

Probable Error of Coefficient of Variation

$$= \frac{\cdot 6744898}{\sqrt{2n}} V \left\{ 1 + 2 \left(\frac{V}{100} \right)^2 \right\}^{\frac{1}{2}}$$

We find, Probable Error of Mean = 0·32906 cm.

Probable Error of S.D. = 0·32267 cm.

Probable Error of V = 0·14166.

The Probable Error of S.D. requires correction for skewness.[2] The P.E. of S.D.

$$= \frac{\cdot 6744898 \ \sigma}{\sqrt{2n}} \sqrt{\{1 + \tfrac{1}{2}(\beta_2 - 3)\}}$$

which reduces to the usual expression involving $\frac{\sigma}{\sqrt{2n}}$ for normal curve, since $\beta_2 - 3 = 0$ approximately in this case. Making this correction we get P.E. of S.D. = 0·3643 cm. This correction has been made in all subsequent work, but the difference made is not considerable in any case.

The probable errors of β_1 and β_2, skewness and d were found from Table XXXVII, XXXVIII, XI and XII pp. 68–77 of Tables for Biometricians.[3]

Probable Errors of β_1. Table XXXVII p. 68.

$\beta_1 = \cdot 0332$

$\beta_2 = 3 \cdot 5$ $\qquad \sqrt{N}\, \Sigma_{\beta_1} = 0 + \frac{332}{500}(1 \cdot 37) = 0 \cdot 9069$

(*c*) "On the Probable Errors of Frequency Constants," *Biometrika* Vol. 2 (1903), pp. 272.

(*d*) Karl Pearson: "On the Mathematical Theory of Errors of Judgment," *Phil. Trans. Roy. Soc.*, Vol. 198A (1902), pp. 274—279.

(*e*) "Probable Errors of Frequency Constants," Part II, *Biometrika*, Vol. 9 (1913), pp.

1 Tables were published by W Gibson and Raymond Pearl (*Biometrika* Vol. pp. 385—393) to facilitate the calculation of probable errors. These have been now reprinted as Tables V and VI in "Tables for Satisticians and Biometricians" (Cambridge University Press, 1914).

2 Karl Pearson, Editorial Note on a paper by Raymond Pearl: "On Certain Points concerning the Probable Error of the Standard Deviation," *Biometrika* Vol. 6 (1909), p. 117.

3 These tables were originally published by A. Rhind in *Biometrika* Vol. 7 (1910), pp. 126–147 and pp. 386–397. Rhind gives an excellent summary of the whole subject.

$\beta_2 = 3{\cdot}6$ $\quad 0 + \frac{332}{500}(1{\cdot}50) = 0{\cdot}9930$

$\beta_2 = 3{\cdot}54\ 75$, $\sqrt{N}\,\Sigma_{\beta_1} = 0{\cdot}9069 + \frac{475}{1000}(\cdot0861)$

$= \cdot9478$

Multiplying by $\chi_1 = \cdot67449/\sqrt{n}$ $= \cdot04769$

we get, P.E. of $\beta_1 = \cdot045201$.

Then from Table XXXVIII, p. 71.

$\beta_1 = \cdot03\ 32$

$\beta_2 = 3{\cdot}5 \quad \sqrt{N}\Sigma_{\beta_2} = 10{\cdot}85 + \frac{332}{500}(0{\cdot}9) = 11{\cdot}4458$

$3{\cdot}6 \quad = 12{\cdot}67 + \frac{332}{500}(1{\cdot}07) = 13{\cdot}3783$

For $\beta_2 = 3{\cdot}54\ 74$, $\quad 11{\cdot}44\ 58 + \frac{475}{1000}(1{\cdot}9325)$

$\sqrt{N}\Sigma_{\beta_2} = 12{\cdot}3637$

hence P.E. of $\beta_2 = 5{\cdot}89625$

From Table XLI, p. 76.

$\beta_2 = 3{\cdot}5 \quad \sqrt{N}\Sigma_{sk} = 1{\cdot}31 - \frac{332}{500}(\cdot02) = 1{\cdot}30\ 87$

$3{\cdot}6 \quad 1{\cdot}32 - \frac{332}{500} \times \cdot02 = 1{\cdot}31\ 87$

$\beta_2 = 3{\cdot}54\ 75$, $\sqrt{N}\Sigma_{sk} = 1{\cdot}30\ 87 + \frac{475}{1000}(\cdot01) = 1{\cdot}3134$

P.E. of Skewness = $\cdot06\ 26\ 36$

We thus find

Mean, $M = 1656{\cdot}25 \pm 3{\cdot}2906$ mm.

S.D. σ $69{\cdot}00 \pm 2{\cdot}6431$ mm.

Coeff. of V, $V = 4{\cdot}1660 \pm \cdot1407$

The other constants are :—

$\beta_1 = \cdot03\ 32\ 04 \pm \cdot04\ 52\ 01$

$\beta_2 = 3{\cdot}54\ 75\ 34 \pm \cdot58\ 96\ 25$

Skewness $= sk = \cdot06\ 98\ 58 \pm \cdot06\ 26\ 36$

We thus find that the skewness is not significant. Hence we are justified in assuming normal distribution, at least to a first approximation.

On this assumption we can find the P.E. of the moments quite easily.

The S.D. of any moment μ_q in a sample of size n is given by [1]

$$n.\Sigma\mu_q^2 = \mu_{2q} - \mu_q^2 - 2q\mu_{q+1}\mu_{q-1} + q^2\sigma^2\mu^2_{q-1}$$

$$\text{P.E. of } \mu_2 = \cdot 67449\sqrt{\frac{\overline{2}}{n}} \cdot \mu_2 = \cdot 06\ 74\ 49\cdot\mu_2$$

$$\text{P.E. of } \mu_3 = \cdot 67449\sqrt{\frac{\overline{6}}{n}} \cdot \sigma^3 = \cdot 11\ 68\ 16\cdot\sigma^3$$

$$\text{P.E. of } \mu_4 = \cdot 67449\sqrt{\frac{\overline{96}}{n}}\ \mu_2^2 = 27\ 96\ 56\cdot\mu_2^2$$

For $q=5$, we must find μ_{10}.

But Sheppard [2] has shown that for the normal curve (in our present notation)

$$\mu_{2s+1} = 0$$
$$\mu_{2s} = (2s-1)(2s-3) \quad 1.\mu_2^s$$

Hence we get $$\Sigma^2\mu_5 = \frac{720}{n}\ \mu_2^5$$

Substituting in the above formula, we get

$$\text{Thus P.E. of } \mu_5 = \cdot 67\ 44\ 9\sqrt{\frac{\overline{720}}{n}} \cdot \sigma^5 = 1\cdot 27\ 96\ 56\cdot\sigma^5$$

We thus get:—

$$\mu_2 = 1\cdot 90\ 43\ 75 \pm 0\cdot 12\ 84\ 31$$
$$\mu_3 = -\ \cdot 47\ 78\ 43\ 77 \pm \cdot 30\ 69\ 85$$
$$\mu_4 = 12\cdot 86\ 56\ 42\ 58 \pm 1\cdot 46\ 74\ 85$$
$$\mu_5 = -8\cdot 18\ 04\ 93\ 98 \pm 6\cdot 40\ 45\ 50$$

SHEPPARD'S CORRECTION.

I shall now consider the question of corrections for grouping. The theoretical work in this subject now consists of a good deal of literature. I shall discuss this question from a purely practical point of view. The fundamental memoir is W F Sheppard [3]: "On the Calculation of the most Probable Values of Frequency

---

[1] "On Probable Errors of Frequency Constants," *Biometrika* Vol. 2 (1903), p. 276.

[2] W F. Sheppard: *Phil. Trans. Roy. Soc.*, 192A.

[3] (*a*) A summary of Sheppard's memoir (with some new results) is given in an Editorial Note: "On an Elemementary Proof of Sheppard's Formulae for correcting Raw Moments and on Other Allied Points" in *Biom.* Vol. 3, pp. 308—310.

(*b*) In Pearson's paper: "On Systematic Fitting of Curves, etc." *Biom.* Vols. 1 and 2, this question has been discussed from a different standpoint.

(*c*) Sheppard himself has given a simplified method of obtaining certain corrections in a later paper: "The Calculation of Moments of a Frequency-Distribution," *Biom.* Vol. 5 (07), pp. 450—459.

(*d*) Eleanor Pairman and Karl Pearson have published a memoir: "On Corrections for the Moment-coefficients of Limited Range Frequency Distributions etc." in *Biom.* Vol. 12 (1919), pp. 231—338, which I shall have occasion to discuss later on in greater detail.

Constants, for Data arranged according to Equidistant Divisions of Scale," *Proc. Lond. Math. Soc.*, Vol. 29, pp. 353—380.

In our notation the above correction (which is known as Sheppard's correction) is given by the following set of equations:—

$$\mu_1' = \nu_1'$$
$$\mu_2' = \nu_2' - \tfrac{1}{12} h^2$$
$$\mu_3' = \nu_3' - \tfrac{1}{4} h^2 \nu'$$
$$\mu_4' = \nu_4' - \tfrac{1}{2} h^2 \nu_2' + \tfrac{7}{240} h^4$$
$$\mu_5' = \nu_5' - \tfrac{5}{6} h^2 \nu_3' + \tfrac{7}{48} h^4 \nu_1'$$
$$\mu_6' = \nu_6' - \tfrac{5}{4} h^2 \nu_4' + \tfrac{7}{16} h^4 \nu_2' - \tfrac{31}{1344} h^6$$

h is the length of the base unit, it is usually = 1 for working units.

50 mm. unit of grouping.

Making these corrections we find adjusted moments about 1655 to be

$\mu_1' =$ ·025, $\mu_2' =$ 1·82 16 67, $\mu_3' = -$0·34 12 50,
$\mu_4' =$ 11·90 16 67, $\mu_5' = -$6·29 21 87, $\mu_6' =$ 147·33 97 83,

Now transferring to Mean we get

$\mu_2 =$ 1·82 10 42 $\mu_3 = -$ ·47 78
$\mu_4 =$ 11·94 16 22 $\mu_5 = -$7·78 23

Hence we finally get "corrected" constants:

Mean =	1656·25	± 3·21 7	mm.
S.D. =	67·47 3	± 2·61 62	,,
Coeff. of V, V =	4·07 38	± ·13 76	
β_1 =	·03 78 10	± ·05 41 33	
β_2 =	3·60 10	± ·71 20 69	
sk =	·07 31 10	± ·06 22 32	
d =	4·93 29 50	± 4·22 30 20	mm.

Note.—Starting with 1430 as our base unit, we reach the same results, thus the arithmetic is absolutely checked in this case.

The Frequency Constants were next calculated (both with and without Sheppard's correction) for widely different units of grouping. We have 1 mm., 20 mm., 30 mm., 50 mm. and finally 100 mm. as our unit of grouping. It will be observed that the unit of grouping is thus successively made the same, 10 times, 20 times, 50 times and finally 100 times the unit of measurement.

With "ungrouped" (i.e, 1 mm.) measurements, the arithmetical labour is tremendous. In this case the maximum value of x is −210, which involves calculating $(210)^4$ for the fourth moment. Hence it was not possible to go beyond the fourth moment. As it is, the actual sum of fourth-products, i.e., $S(x^4y)$ runs into 11 figures. I quote actual results

$S(xy)$ =	1 58
$S(x^2y)$ =	90 82 72
$S(x^3y)$ =	−6 76 88 78
$S(x^4y)$ =	144 04 28 60 67

which gives us (dividing by 200) :—

$$\nu_1' = \quad \cdot 79$$
$$\nu_2' = \quad 45\ 41\ \cdot 36$$
$$\nu_3' = \quad 3\ 48\ 44\ \cdot 39$$
$$\nu_4' = 72\ 02\ 14\ 30\ \cdot 38$$

For purposes of comparison it is necessary to reduce all moments to the same unit. 50 mm. was chosen as the standard unit.[1]

Let μ_n be any moment in units of grouping h, let M_n be the corresponding moment in standard units h_o, let $\rho = \frac{h}{h_o}$

Then $M_n = \rho^n \mu_n$, is the formula of reduction to standard unit.

For $h_o = 50$ mm., $\rho = \frac{1}{50}, \frac{2}{5}, \frac{3}{5}$ and 2 successively for units of 1 mm., 20 mm., 30 mm. and 100 mm. respectively.

The annexed table gives the Frequency Constants for the different units of grouping. I have added the probable errors in each case.

For the purpose of studying the effect of grouping it is natural to take the "ungrouped" constants as our standard. We have accordingly assumed that the 1 mm. constants are the "true" constants.

Different Values of Mean Stature.

Unit of Grouping.	
1 mm.	16 56·79 ± 3·23 mm.
20 "	16 56·85 ± 3·23 "
30 "	16 56·35 ± 3·23 "
50 "	16 56·25 ± 3·22 "
100 "	16 59·50 ± 3·09 "

When the unit of grouping is so large as 100 mm. (and the total record is divided only into 5 groups), there is considerable difference in the Mean. But this difference of 2·71 mm. is less than the probable error of over 3 mm. Thus even with 100 mm. grouping, the Mean is stable within the limits of its own probable error.

The agreement is almost perfect when we omit the 100 mm. group. The maximum "error" due to grouping amounts to only ·54 mm., which is considerably less than the unit of measurement itself and is about $\frac{1}{6}$ of the probable error.

Let us consider a very large sample of 7,500 individuals. It is not likely that the Standard Deviation will exceed 70 mm. The P.E. of Mean will be about ·55 mm. *The maximum observed difference in the present case, due to grouping, is thus of the same order as the random P.E. of the Mean in a sample of 7,500. We conclude therefore that for samples of 200, the effect of grouping on the Mean up to 50 mm. is quite negligible.*

---

[1] For reasons explained on pp. 39–40.

Standard Deviation.

Let us first consider the results without Sheppard's correction.

1 mm.	67·385 ± 2·557	mm.
20 ,,	67·894 ± 2·547	,,
30 ,,	68·365 ± 2·444	,,
50 ,,	69·00 ± 2·679	,,
100 ,,	70·922 ± 2·662	,,

With 100 mm. the difference is quite large. It is 3·537 mm. which is considerably greater than the prob. error. Omitting 100 mm. we find the maximum difference to be 1·615 mm., which is considerable, but is still less than the P.E. Such a P.E. will be obtained with samples of 400. *Thus the agreement without Sheppard's correction is not very good.*

With *Sheppard's correction*

1 mm.	67·385 ± 2·557	mm.
20 ,,	67·648 ± 2·539	,,
30 ,,	67·812 ± 2·426	,,
50 ,,	67·473 ± 2·619	,,
100 ,,	64·77 ± 2·432	,,

100 mm. is again discrepant. The difference is 2·615 mm. which is of just the same order as the P.E. Evidently 100 mm. grouping is too broad and the error due to grouping is no longer negligible. This is also obvious from the fact that Sheppard's correction makes the S.D. actually less than its true value, while the uncorrected value is considerably greater.

Omitting 100 mm. the agreement is excellent. The maximum difference (which is now in the 30 mm. group) is only ·427 mm., a value about a sixth of the probable error. It will require a sample of 6000 to produce a random error of the same amount.

Thus with Sheppard's correction, the effect of grouping is quite negligible up to 50 mm. These corrections are so easily applied that there can be no excuse for omitting them. We have thus empirically verified the great importance of Sheppard's correction in giving better values of the Frequency Constants. Henceforth it will not be necessary to compare the values obtained without Sheppard's correction.

Coefficient of variation: $V = \frac{100\sigma}{M}$

1 mm.	4·06 72 ± ·13 74	
20 ,,	4·08 29 ± ·13 79	
30 ,,	4·09 41 ± ·13 83	
50 ,,	4·07 38 ± ·13 76	
100 ,,	3·90 29 ± ·13 18	

100 mm. is obviously incorrect, we may omit this group from further consideration. The difference ·1643 is greater than the

P.E. Omitting 100 mm. the maximum difference is ·0269, which will be the P.E. in a random sample of 5,000 (with coeff. of variation equal to 4). *Thus the effect of grouping is of the same order as the effect of sampling in a group of 5,000.* Hence we conclude that different units of grouping do not introduce any appreciable errors in the Coefficient of Variation.

From the Anthropological standpoint, the Mean, the S.D., and the Coeff. of Variation are the most important constants. For stature, with samples of 200 *with Sheppard's correction the effect of even such a large unit of grouping as 50 times the unit of measurement is in all these cases absolutely inappreciable.*

We shall however consider the other statistical constants before concluding this portion of our work.

Values of μ_2.

With Sheppard's correction:—

1 mm.	1·81 62 61	± ·12 25 05
20 ,,	1·83 04 98	± ·12 34 77
30 ,,	1·83 94 72	± ·12 40 83
50 ,,	1·82 10 42	± .12 28 40
100 ,,	1·67 87	± ·11 32 39

100 mm. makes a difference of 1376, which is just about the same as the P.E. Otherwise the maximum difference is ·0232 which is only a sixth of the P.E. A random error of the same amount will be produced in samples of 2800.

Let us now compare the values obtained *without Sheppard's correction:*

1 mm.	1·81 62 94	± ·12 25 07
20 ,,	1·84 38 31	± ·12 43 09
30 ,,	1·86 94 72	± ·12 61 06
50 ,,	1·90 43 75	± ·12 84 60
100 ,,	2·01 20	± ·13 57 07

100 mm. introduces an error of ·1958 which is considerably greater than the P.E.

The effect of grouping has now become quite obvious, 20 mm., 30 mm. and 50 mm. now introduce steadily increasing error. With 50 mm. the error has now amounted to ·0881 which is only $\frac{2}{3}$rds of the P.E.

We thus see that Sheppard's correction is absolutely indispensable here. With Sheppard's correction the effect is quite negligible up to 50 mm.

Values of μ_3.

With Sheppard's correction:—

1 mm.	= − ·64 36 06	± ·28 59
20 ,,	= − ·30 87 16	± ·29 27
30 ,,	= − ·46 86 97	± ·29 86
50 ,,	= − ·47 78 44	± ·30 70
100 ,,	= − ·48 35 78	± ·33 33

100 mm. is not at all worse than others. The maximum error (which now occurs in the 20 mm. group) ·3349 just exceeds the P.E.

Without Sheppard's correction :—

20 mm.	=	·30 87	± ·28 93
30 ,,	=	.46 55	± ·29 14
50 ,,	=	.47 78	± ·28 71
100 ,,	=	·66 40	± ·25 39

Evidently Sheppard's correction does not produce substantial improvements. In this case the gross P.E. of μ_3 is of the same order as μ_3 itself and hence there is wide fluctuation in the result.

In view of the large P.E. we cannot say that grouping makes any significant difference. The asymmetry is very slight and very nearly zero, thus the fluctuations though large are not statistically significant. These wide fluctuations indicate the critical approach to the Gaussian curve.

Values of μ_4.

With Sheppard's correction :—

1 mm.	11·56 10 21	± 1·54 15
20 ,,	11·56 54 26	± 1·56 59
30 ,,	10·97 11 78	± 1·58 08
50 ,,	11·94 16 22	± 1·54 97
100 ,,	10·35 96	± 1·31 58

100 mm. makes a difference of 1·2014 which nearly equals the P.E. Otherwise the agreement is good. The maximum error is ·59 (in the 30 mm. group) which is much less than ½ the P.E. Random error of the same amount will require samples of 1300 individuals.

Without Sheppard's correction the agreement is much worse. We have

1 mm.	11·56 10 21	± 1·54 15
20 ,,	11·70 20	± 1·58 87
30 ,,	11·30 37	± 1·63 31
50 ,,	12·86 56 42	± 1·69 46
100 ,,	13·98 12 48	± 1·89 13

100 mm. has become too " rough " and 50 mm. itself introduces an error of about the same order as the P.E. Thus Sheppard's corrections make substantial improvement in the results. The percentage probable error of μ_4 for normal curves is given by $\frac{1}{3}\chi_1\sqrt{96} = 15{\cdot}7\%$ in our case. In view of this large percentage variation, observed agreement with different groupings is quite satisfactory.

Values of μ_5.

With Sheppard's correction:—[1]

30 mm.	−11·92 55 72	± 5·87 11
50 ,,	− 7·78 23	± 5·72 74
100 ,,	−11·09 49 34	± 4·66 78

Without Sheppard's correction:—

50 mm.	− 8·18 04 94	± 6·40 46
100 ,,	−14·56	± 7·34 65

The gross prob. error is again of the same order as μ_5 itself. Hence there is very wide fluctuation in its value and Sheppard's correction is not important. It should be noted however that even now the maximum difference (*inter se*) is less than the P.E.

Values of μ_6.

30 mm.	121·83 05	± 29·93 43
50 ,,	154·73 13	± 29·03 98

The percentage P.E. for normal curve is $\frac{1}{15}\sqrt{4\cdot80\ 95} = 32\ 63\%$ *With such large percentage variation it is quite idle to calculate the higher moments directly.*

Pearson says in this connection[2] "Constants based on high moments will be practically idle. They may enable us to describe closely an individual random sample, but no safe argument can be drawn from this individual sample as to the general population at large, at any rate so far as the argument is based on the constants depending upon these high moments."

Values of β_1.

1 mm.	·06 87 56	± ·07 97 81
20 ,,	·01 55 38	± ·01 93 24
30 ,,	·03 53	± ·03 57 68
50 ,,	·03 78 10	± ·06 55 55
100 ,,	·04 94 90	± ·06 31

Remembering that $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$, we are quite prepared for such wide fluctuations. It will be seen that β_1 differs from zero by just about the same amount as its own P.E. (calculated separately for each) which of course implies that there is a tendency towards β_1 differing *slightly* form zero, but that with a small sample of 200 this tendency has not become quite significant. The unit of grouping does not make any difference so far as this tendency is

---

[1] On account of the great Arithmetical labour, it has not been found possible to calculate μ_5 and μ_6 with lower units of grouping.

[2] Draper's Company Research Memoirs: "On the General Theory of Skew Correlation and Non-Linear Regression," p. 9.

concerned. With 50 mm. without correction, β_1 is = ·03 32 04 ± ·04 52 01. Thus Sheppard's correction is not important.

Values of β_2.

1 mm.	3·50 46	± ·60 17
20 ,,	3·45 16 21	± ·49 72 97
30 ,,	3·24 24	± ·35 44 89
50 ,,	3·60 10 00	± ·71 20 69
100 ,,	3·45 36	± ·48 51
50 ,,	3·54 75 34	± ·58 96 25 (without correction).

Though β_2 does not seem to differ significantly from 3, there is slight tendency towards lepto-kurtosis.[1]

The P.E. of β_2 for a Gaussian distribution is $\chi_1 \sqrt{24}$ and is about ± ·23 in our case. The magnitude of P.E. again shows the want of significant divergence from meso-kurtosis.

The effect of grouping is evidently quite negligible. The above investigation has been most elaborate in character and is sufficient to justify the application of "grouped" statistical methods to our present material.

The foregoing analysis may be summarized thus:—

(1) With samples of 200, even such broad grouping as 100 mm. does not introduce errors greater than the random error of sampling.

(2) Up to 50 mm. the effect of grouping is absolutely negligible. In the case of the Mean, the S.D. and the Coeff. of Variation, "grouping error" is of the same order as "random error" in samples of several thousands of individuals.

(3) Sheppard's correction leads to a very substantial improvement in the S.D. and the even moments. The odd moments (being near a critical value) are not affected very much. Speaking generally, Sheppard's correction should never be omitted.

(4) The percentage variation in the higher moments is too large to make it worth while calculating them directly.

I speak with hesitation about another inference which may perhaps be drawn from the above investigation. Small errors of estimating stature—even up to perhaps a few mm. are not likely to affect the Mean value very considerably (provided these errors are random errors and not systematic).

"FULL CORRECTIONS" OF PAIRMAN AND PEARSON.

We shall now consider certain "full corrections" recently discussed by Pairman and Pearson.[2] The object of the above

---

[1] K. Pearson: "Skew variation, a Rejoinder" *Biom.* Vol. 4 (1906), p. 175 Also appendix II.

[2] Eleanor Pairman and K. Pearson: "On Corrections for the Moment-Coefficients of Limited Range Frequency Distributions when there are Finite or Infinite Ordinates and any Slopes at the Terminals of the Range." *Biom.* Vol. 12 (1919), pp. 231–258.

paper was to investigate the full corrections for curtailed blocks of frequency.

The general shape of our curve showed that there was no significant curtailing, still I thought it advisable to investigate this point more carefully.

We choose 50 mm. unit of grouping as our standard and find "raw" moments about one end of range, i.e. 1430 mm.

Stature in mm.	Frequency $=y$.
1430–1480	3
1530	5
1580	14
1630	45
1680	60
1730	48
1780	20
1830	3
1880	2
TOTAL .	200

Raw Moments are:—

$\nu_1' = 4{\cdot}5250$

$\nu_2' = 22{\cdot}38$

$\nu_3' = 118{\cdot}0262$

$\nu_4' = 657{\cdot}4275$

Note.—These lead to the same moments about Mean as obtained from raw moments about 1655. Hence there is an absolute check on the Arithmetic.

Instead of working with n_1', n_2' (the proportional frequencies), we can work with y_1, y_2, the actual frequencies, and then divide the whole by 200. Thus we get the following (slightly modified) formulae from p. 233 of the paper cited above.

$$a_1 = -\tfrac{1}{200}\ \tfrac{1}{60}\{163y_1 - 163y_2 + 137y_3 - 63y_4 + 12y_5\}$$
$$a_2 = +\tfrac{1}{200}\ \tfrac{1}{12}\{45y_1 - 109y_2 + 105y_3 - 51y_4 + 10y_5\}$$
$$a_3 = -\tfrac{1}{200}\ \tfrac{1}{4}\{17y_1 - 54y_2 + 64y_3 - 34y_4 + 7y_5\}$$
$$a_4 = +\tfrac{1}{200}\ \{3y_1 - 11y_2 + 15y_3 - 9y_4 + 2y_5\}$$
$$a_5 = -\tfrac{1}{200}\ \{y_1 - 4y_2 + 6y_3 - 4y_4 + y_5\}$$

and for b's

$$b_1 = +\tfrac{1}{200}\ \tfrac{1}{60}\{137y_p - 163y_{p-1} + 137y_{p-2} - 63y_{p-3} + 12y_{p-4}\}$$
$$b_2 = -\tfrac{1}{200}\ \tfrac{1}{12}\{45y_p - 109y_{p-1} + 105y_{p-2} - 51y_{p-3} + 10y_{p-4}\}$$
$$b_3 = +\tfrac{1}{200}\ \tfrac{1}{4}\{17y_p - 54y_{p-1} + 64y_{p-2} - 34y_{p-3} + 7y_{p-4}\}$$
$$b_4 = -\tfrac{1}{200}\ \{3y_p - 11y_{p-1} + 15y_{p-2} - 9y_{p-3} + 2y_{p-4}\}$$
$$b_5 = +\tfrac{1}{200}\ \{y_p - 4y_{p-1} + 6y_{p-2} - 4y_{p-3} + y_{p-4}\}$$

In our case

$$y_1 = 3,\ y_2 = 5,\ y_3 = 14,\ y_4 = 45,\ y_5 = 60$$
$$y_p = 2,\ y_{p-1} = 3,\ y_{p-2} = 20,\ y_{p-3} = 48,\ y_{p-4} = 60$$

Hence we obtain

$a_1 = +\cdot05\ 00\ 83$ $\quad b_1 = +\cdot01\ 86\ 67$
$a_2 = -\cdot26\ 45\ 83$ $\quad b_2 = -\cdot00\ 62\ 50$
$a_3 = +\cdot54\ 12\ 50$ $\quad b_3 = -\cdot07\ 50\ 00$
$a_4 = -\cdot60\ 50\ 00$ $\quad b_4 = +\cdot19\ 50\ 00$
$a_5 = +\cdot26\ 50\ 00$ $\quad b_5 = -\cdot10\ 50\ 00$

From these we obtain:—

$A_1 = a_1' - \frac{1}{6\,0}\,a_3' + \frac{1}{2\,5\,2\,0}\,a_5' = +\cdot04\ 11\ 60\ 16$
$B_1 = b_1' - \frac{1}{6\,0}\,b_3' + \frac{1}{2\,5\,2\,0}\,b_5' = +\cdot01\ 98\ 75\ 00$
$A_2 = a_2' - \frac{5}{1\,2\,6}\,a_4' = -\cdot24\ 05\ 75\ 40$
$B_2 = b_2' - \frac{5}{1\,2\,6}\,b_4' = -\cdot01\ 39\ 88\ 09$
$A_3 = a_1' - \frac{5}{6\,3}\,a_3' + \frac{1}{2\,4\,0}\,a_5' = +\cdot00\ 82\ 31\ 24$
$B_3 = b_1' - \frac{5}{6\,3}\,b_3' + \frac{1}{2\,4\,0}\,b_5' = +\cdot02\ 41\ 81\ 55$
$A_4 = a_2' - \frac{7}{8\,0}\,a_4' = -\cdot21\ 16\ 45\ 80$
$B_4 = b_2' - \frac{7}{8\,0}\,b_4' = -\cdot02\ 38\ 75\ 00$

From Equations (xxii) to (xxv) on p. 240, we get the fully corrected raw moments to be:—

$$\mu_1' = \nu_1' + \tfrac{1}{1\,2}\{A_1 + B_1\}$$
$$\mu_2' = \nu_2' - \tfrac{1}{1\,2} + \tfrac{1}{1\,2\,0}\{B_2 - A_2\}$$
$$\mu_3' = \nu_3' - \tfrac{1}{4}\,\nu_1' + \{-\tfrac{1}{4\,0}(A_3 + B_3) + \tfrac{1}{4\,0}p\ B_2 + \tfrac{1}{4}p^2\ B_1\}$$
$$\mu_4' = \nu_4' - \tfrac{1}{2}\,\nu_2' + \tfrac{7}{2\,4\,0} + \{\tfrac{1}{1\,2\,6}(A_4 - B_4) - \tfrac{1}{1\,0}pB_3 + \tfrac{1}{2\,0}p^2B_2 + \tfrac{1}{3}p^3\ B_1\}$$

In our case the range $p = 9$, and we get:—

$\mu_1' = \nu_1' + \{\cdot00\ 50\ 86\ 26\}$
$\mu_2' = \nu_2' - \frac{1}{1\,2} + \{\cdot03\ 17\ 00\ 69\}$
$\mu_3' = \nu_3' - \frac{1}{4}\ \nu_1' + \{\cdot39\ 12\ 18\ 23\}$
$\mu_4' = \nu_4' - \frac{1}{2}\ \nu_2' + \{\cdot45\ 67\ 19\ 58\} + \frac{7}{2\,4\,0}$

Where the curled brackets give the correction over and above Sheppard's correction.

Thus we get fully adjusted raw moments to be

$\mu_1' = 4\cdot53\ 00\ 86\ 26\ 30$
$\mu_2' = 22\cdot32\ 83\ 67\ 35\ 85$
$\mu_3' = 117\cdot28\ 62\ 18\ 23\ 12$
$\mu_4' = 646\cdot72\ 33\ 86\ 24\ 84$

Transferring to the Mean (which itself is now changed) we obtain the Moment-Coefficients about the Mean.

Moments after "full correction"

$\mu_2 = 1\cdot80\ 66\ 86$
$\mu_3 = -0\cdot23\ 20\ 97$
$\mu_4 = 7\cdot53\ 03\ 39$

and the Mean $= 1656\cdot5043$ mm.
with S.D. $= 67\cdot1950$ mm.

Comparing with our "standard" values we see evident signs of "over correction." With such small samples as 200, the P.E. in terminal frequencies are too great to allow the a's and b's to be calculated with any degree of accuracy. The transfer of one individual from one group to another would seriously affect the results.

In order to test this point, I next calculated the a's and b's with a shorter sub-range, i.e. 40 mm.

Thus $$P_h = P_o = \frac{h_o}{h} = \frac{50}{40} = 1{\cdot}25$$

Hence $$a_s' = (1{\cdot}25)^s . a_s$$
$$b_s' = (1{\cdot}25)^s . b^s$$

1430–1470	–1510	–1550	–1590	–1630	–1670	–1710	–1750	–1790	–1830	–1870 mm.
y_1	y_2	y_3	y_4	y_5		y_{p-4}	y_{p-3}	y_{p-2}	y_{p-1}	y_p
2	2·5	6·0	18·5	38·0	51·0	40·0	24·0	15·0	1·0	2·0

From these we get

$a_1 = +{\cdot}00\ 17\ 50$ $\quad$ $b_1 = +{\cdot}09\ 36\ 67$
$a_2 = -{\cdot}04\ 83\ 33$ $\quad$ $b_2 = -{\cdot}30\ 5$
$a_3 = +{\cdot}10$ $\quad$ $b_3 = +{\cdot}50\ 5$
$a_4 = -{\cdot}11$ $\quad$ $b_4 = -{\cdot}42$
$a_5 = +{\cdot}04$ $\quad$ $b_5 = +{\cdot}16$

leading to

$a_1' = +{\cdot}00\ 21\ 88$ $\quad$ $b_1' = +\ {\cdot}11\ 70\ 83$
$a_2' = -{\cdot}07\ 55\ 21$ $\quad$ $b_2' = -\ {\cdot}47\ 65\ 63$
$a_3' = +{\cdot}19\ 53\ 13$ $\quad$ $b_3' = +\ {\cdot}98\ 63\ 28$
$a_4' = -{\cdot}26{\cdot}85\ 55$ $\quad$ $b_4' = -1{\cdot}02\ 53\ 91$
$a_5' = +{\cdot}12\ 20\ 70$ $\quad$ $b_5' = +\ {\cdot}48\ 82\ 81$

These give

$$\mu_2 = 1{\cdot}89\ 48\ 76\ 86$$

giving S.D. = 68·82 5

and Mean = 1656·66 58 mm.

The values are again quite discrepant from those given above.

With subrange of 25 mm. still more widely divergent values were obtained.

Hence we are obliged to conclude that with small samples, the probable errors of the terminal frequencies are much too large to allow Pairman and Pearson's "full corrections" being calculated with accuracy.

The general conclusion of the above investigation is this.

There is no indication of appreciable "curtailing" of our material. Further, with small samples, the "abruptness coefficients" cannot be calculated with any reasonable degree of accuracy and these "full corrections" will necessarily have to be omitted. But we have already seen that Sheppard's correction can be safely applied and should never be omitted.

SECTION III. ON THE STATISTICAL TESTS OF HOMOGENEITY

One of the main objects of our present enquiry is to investigate the "homogeneity" of our material. For this purpose it is necessary to have some precise definition of "homogeneity." I fully realise the great difficulties underlying any attempt at such a definition, but in order to avoid confusion of thought I have found it impossible to forego at least a working definition. I shall approach the problem from a purely statistical point of view.

"Homogeneity" implies similarity and functional equivalency among the members of a group of any class of objects. When all the members are identical with respect to some definite property, homogeneity is perfect with reference to that particular property. This is the ideal limit of thought, but in practice it always remains a mere intellectual abstraction.

Thus in actual practice diversity is always present. But if the similarity attains a certain intensity we can speak of the group as being homogeneous. The actual amount of similarity considered necessary to attain this intensity is of course a matter of practical convenience. A group which is homogeneous for one purpose may be quite heterogeneous for another.[1]

"Homogeneity" thus ultimately depends on our standard of discrimination.[2] If the actual difference between any two members of a group is *less* than our unit of discrimination, we can never become aware of this difference and the group will appear to be homogeneous. On the other hand if the actual difference is greater, heterogeneity will become evident. If our unit of discrimination is made indefinitely small and yet no heterogeneity is detected, we gradually approach identity, which is the ideal limit of thought.

The concept of "homogeneity" is thus essentially relative and practical. We can never have any absolute logical criterion of homogeneity. We must set up separate standards of homogeneity in each case. To this extent the definition of homogeneity is necessarily arbitrary and conventional. But having once set up a standard we must rigidly adhere to it. We cannot give it up in the middle of a discussion on the plea of arbitrariness.

The discriminant may be either qualitative or quantitative, in either case it should be precise and definite.

We can now proceed to set up tests of homogeneity for our special purpose.

---

[1] Cf. K. Pearson Skew Variation," *Biom.* Vol. 4 (1906), p. 176, 192 and p. 185.

[2] e.g. In statistics, the probable error is the fundamental discriminant, in Experimental Psychology the least perceptible difference is the ultimate unit.

From the statistical standpoint our first necessity is suitable graduation of the given sample. This is necessary in order to draw legitimate inferences about the general population from a study of the given sample.[1] Our first condition is :—

I. *We should be able to graduate the given sample by a smooth curve. That is, the given frequency distribution must be homotypic*[2] *in character.*[3]

The goodness of fit can be tested by the Pearsonian Contingency Coefficient.[4]

Possibility of graduation by a smooth curve is thus a necessary condition of statistical homogeneity.

This is not however sufficient. All heterotypic curves are excluded, but a homotypic frequency curve need not necessarily be homogeneous. For example, it may well happen that a mixture of two different homogeneous samples is amenable to graduation by a homotypic curve. But even then if the given curve can be split up into simpler components we get direct evidence of heterogeneity.

II. Thus our second condition is *that the sampled frequency curve should not be capable of being analysed*[5] *into simpler real*[6] *components.*

Pearson[7] has furnished us with a technical method for dissection into two components. But failure in dissection may also imply that the curve is multi-complex in character, i.e. that it is built up of more than two simple components. This second condition (impossibility of analysis) again though necessary, is yet not sufficient.

The concept of functional equivalency provides us with another test. If we consider any sub-sample[1] it should be generally equivalent to another sub-sample, that is, it should not differ significantly from other sub-samples. Thus we get :—

III. *The frequency constants of different sub-samples should agree within the limits of their own probable error.*[8]

---

[1] We assume throughout that all samples are random samples, that is, we definitely exclude heterogeneity due to mere "bias" in sampling.

[2] Homotypic curves will ordinarily include the Gaussian and the different Pearsonian skew curves. Other smooth curves (Edgeworth, Charlier, Thiele, Kapteyn etc.) may also be included.

[3] The possibility of suitable graduation of the present material has been discussed in Section IV, pp. 35–40.

[4] The original memoir was given in *Phil. Mag.* 1900, pp. 157—175. For a discussion of its use in testing goodness of fit see L. Isserlis : "On the Representation of Statistical Data." *Biometrika*, Vol. XI (1917), pp. 418—425.

[5] The possibility of dissection of the present material has been investigated in Section VI.

[6] Negative and imaginary solutions are sometimes obtained; until we can give a consistent interpretation of these, it is perhaps safer to ignore such purely mathematical solutions.

[7] Memoir on Dissection of Curves, already cited *Phil. Trans. Roy. Soc.*, 185A (1894).

[8] Strictly speaking, the agreement of subsamples is only an indirect test of homogeneity. What it actually does serve to show is the representative character of the given sample.

This condition ensures that the sub-samples will not differ significantly from the general sample.[1]

The above three tests are purely formal and have no reference to the nature of the material. We can proceed further by taking into consideration our previous experience of similar material.

Let us take the case of stature as an example. In all known cases stature distribution is either approximately Gaussian or is of Type IV or Type I. Consider the frequency distribution of some unknown sample. If we find that the curve though homotypic is J or U shaped, we are naturally suspicious about the homogeneity of the material. The curve may be smooth, it may successfully resist dissection, its sub-samples may agree quite well, yet in view of our previous experience we would, in the absence of other evidence, hesitate to call it homogeneous.

IV Our fourth criterion is that the general nature of the sampled frequency should be the same as that of known homogeneous material.

This criterion is quite empirical in character and its practical utility depends upon what exact significance we can attach to the concept of "general nature of known frequency constants." Though somewhat vague this condition is by no means useless.

Let us suppose that the given sample is really heterogeneous in character. Consider a "random" subsample of the given sample. Now if this subsample is to be representative in character, it must include the same degree of heterogeneity as is present in the sample itself, that is, in order that it may be a "fair" as well as a "random" subsample, it is necessary that it should be sufficiently large. Samples which are large enough to be "fair" will obviously agree among themselves. Thus the agreement of large fair subsamples cannot reveal the want of homogeneity of the given sample.

Now consider a subsample which is again "random" but which is not sufficiently large to include the same degree of heterogeneity as is present in the sample. Not being representative in character, it will not be surprising if these fail to agree. Thus want of agreement on the part of subsamples on account of their smallness of size will not necessarily prove the existence of heterogeneity in the material. The lower limit of agreement of random subsamples may however be looked upon as a measure of homogeneity.

In any case however, agreement of random subsamples does show that these subsamples are large enough to be representative in character. The given sample, being larger than its own subsamples, will obviously be large enough to be representative in character. Thus the agreement of subsamples is a test of the representative character of the sample, rather than any evidence of the homogeneity of the material.

An example may help. Consider an ordinary black and white chess board. Let us look at this chessboard through a sighting hole. The size of this sighting hole determines the size of the sample. If this size is larger than the size of one of the squares then each sample will show a mixed patch. In this case subsamples would agree. On the other hand, if the size of the sighting hole is only a fraction of the size of a square, then some samples will show white, some black and others mixed patches. The lower limit, up to which samples agree is evidently a measure of the size of the discontinuities. Agreement of subsamples of 100 shows that 200 is large enough be representative in character in the present case.

This implication serves as the basis of Pearson's discussion of P.E. of sub-samples for comparison with the general sample. K. Pearson: "Note on the Significant or Non-significant Character of a sub-Sample drawn from a Sample." *Biometrika* Vol. 5 (1906), pp. 181—183.

We require some further precise quantitative test. This is supplied by the variability (both absolute, as measured by the Standard Deviation and relative, as measured by the Coefficient of Variation) of the distribution.[1]

V *The variability of the sample should not be significantly greater than the average variability of the same organ for known homogeneous material.*

The Coefficient of Variation, V (multiplication by 100 is merely for arithmetical convenience) is a straightforward measure of variability. It is of course possible to set up other standards by choosing some other function of the S.D. and Mean, $f\left(\frac{\sigma}{M}\right)$, but it is quite unnecessary to enter into such subtleties in the present stage of our knowledge.

It is quite easy to extend the above condition to the case of more than one organ. In that case we shall have to define variability by the generalised[2] or multiple probable error of the group of organs considered.[3]

We have thus got five different tests of "homogeneity." It should be remembered that we have all along discussed statistical homogeneity. Whether statistical homogeneity necessarily implies anthropological homogeneity and vice versa, is a very difficult question,[4] into which I do not propose to enter. I confine myself to a consideration of purely statistical homogeneity.

---

[1] For a full discussion see Pearson: *Chances of Death* "Variation in Man and Woman," pp. 255—377, specially pp. 272—286. Also Appendix I.

[2] K. Pearson and Alice Lee: "On the Generalised Probable Error in Multiple Normal Correlation," *Biom.* Vol. 6 (1908), pp. 59 - 68.

[3] Incidentally we may note that variability gives us a convenient method of defining a "normal" group (in a medical, psychological or social sense) of individuals. The normal group (with reference to some particular trait) consists of the individuals included between the Mean, M, and p times the S.D. σ, where p is an arbitrary number. Thus a "normal" individual is one who does not differ from the average type of his class by more than $p\ \sigma$. By a proper choice of p we can make our definition as elastic or as stringent as we please. We can also extend the definition to cover more than one single trait, with the help of the generalised or multiple probable error.

[4] K. Pearson: "Craniological Notes. Homogeneity and Heterogeneity in Collections of Crania," *Biom.* Vol. 2 (1903), pp. 345—347. Also see C. Myer's Reply to above and Pearson's Remarks on the Reply, *Biom.* Vol. 2 (1903), pp. 504—508, and Aurel Von Török's Note and Pearson's Reply. *Ibid.*, pp. 508—510.

SECTION IV. TYPE OF CURVE AND "GOODNESS OF FIT"

We shall now test the "goodness of fit" with our "normal" curve. K. Pearson[1] has shown how this may be done. He shows that[2] if

$$x^2 = S\left[\frac{(m_1' - m)^2}{m}\right],$$

where S denotes a summation, m' and m are observed and theoretical values in each sub-group, then the chances of a system of errors with as great or greater frequency than that denoted by x^2 is given

by
$$P = \frac{\left[\int\int\int e^{-\frac{1}{2}x^2} dx_1\ dx_2\ dx_3 \ldots \ldots dx_n\right]_o^\infty}{\left[\int\int\int \ldots e^{-\frac{1}{2}x^2} dx_1 . dx_2\ dx_3 \ldots dx_n\right]_0^\infty}$$

$$= \frac{\int_x^\infty e^{-\frac{1}{2}x^2} x^{n-1}\, dx}{\int_0^\infty e^{-\frac{1}{2}x^2} x^{n-1}\, dx}$$

which reduces to for n' odd

$$P = e^{-\frac{1}{2}x^2}\left\{1 + \frac{x^2}{2} + \frac{x^4}{2.4} + \ldots + \frac{x^{n'-3}}{2\ 4\ 6 \ldots (n'-3)}\right\}$$

and n' even

$$P = \sqrt{\frac{2}{\pi}}\int_\infty^\infty e^{-\frac{1}{2}x^2} dx + \sqrt{\frac{2}{\pi}}\, e^{-\frac{1}{2}x^2}\left\{\frac{x}{1} + \frac{x^3}{1\ 3} + \ldots + \frac{x^{n'-3}}{1\ 3\ 5 \ldots (n'-3)}\right\}$$

Tables[3] have been calculated to facilitate calculation of P when x^2 is known.

Pearson then shows[4] that if x^2 for the sample is so small as to warrant us in speaking of the frequency distribution as a random

---

[1] K. Pearson: "On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling." *Phil. Mag.* July 1900, p. 157.

[2] x^2 is thus quite easy to calculate; it is given by

$$x^2 = \text{Sum}\left(\frac{\text{square of difference of theoretical and observed values}}{\text{theoretical value of frequency}}\right)$$

[3] W. Palin Elderton: "Tables for Testing Goodness of Fit." *Biom.* Vol 1 (1902), pp. 155—163. Reprinted as Table XII on p. 26 of Tables for Statisticians, etc.

[4] Pearson, paragraph 5 and following of reference 1.

variation of the frequency distribution determined from itself, then we may also speak of it as a random sample from a general population whose theoretical distribution differs only by quantities of the order of the probable errors of the constants from the distribution deduced from the observed sample.

Thus *if a curve is a good fit to a sample, to the same fineness of grouping it may be used to describe other samples from the same population.* If a curve serves to any degree, it will serve for all rougher degrees, but it does not follow that it will suffice for still finer groupings. A good fit for a large sample would be a good fit for a smaller sample but not necessarily for a larger one.[1]

I shall test the Goodness of Fit for different groupings. I shall next compare the fit for the same grouping given by the slightly different values of the Standard Deviation calculated with different unit of grouping. This will test how the Goodness of Fit is affected by different units of grouping adopted in calculating the frequency constants.

NORMAL CURVE.

I have calculated the theoretical frequencies from the "raw" (i.e. uncorrected by Sheppard's adjustment) values of the S.D. in some cases. For "if the ordinates of a normal curve be calculated from the raw second moment value of the Standard Deviation, these ordinates will more closely represent the actual frequencies than do the ordinates of the true normal curve, which have to be corrected by the factor

$$1 + \frac{1}{24} h^2 \frac{x_\gamma^2 - \sigma^2}{\sigma^2}$$

to obtain the actual frequencies."

If therefore our sole object is to compare observed and calculated frequencies for definite series of groups, there are advantages in using the "raw" second moment in the equation to the curve. Such a curve has been termed by Sheppard a "spurious curve of frequency"[2].

---

[1] For a discussion of another test of Goodness of Fit proposed by Prof. Edgeworth see a Note by L. Isserliss: "On the Representation of Statistical Data" *Biometrika* 1917, pp. 418—425.

[2] Editorial Note: "On an Elementary Proof of Sheppard's Formulae for correcting Raw Moments and on other Allied Points," *Biom.* Vol. 3 (1904), p. 311.

TABLE 2.

Mean = 1656·2938 mm.

Unit = 20 mm.

S.D. = 67·3845 mm. $1/\sigma = \cdot 296802$.

Stature.	Observed Value m'.	Theoretical Value m.	$(m'-m)$	$\frac{(m'-m)^2}{m}$
Beyond 1475	3	1·10 22	1·897	3·265
1475—1495	1	1·37 34	0·373	·101
—1515	4	2·66 11	1·338	·673
—1535	2	4·72 29	2·722	1·569
—1555	4	7·68 67	3·686	1·768
—1575	10	11·45 71	1·457	·185
—1595	12	15·64 81	3·684	·850
—1615	25	19·58 25	5·418	1·499
—1635	32	22·45 35	9·546	4·058
—1655	21	23·58 87	2·588	·284
—1675	17	22·71 04	5·710	1·436
—1695	21·5	20·23 05	1·269	·796
—1715	18·5	16·18 89	2·311	·330
—1735	10	11·98 74	1·987	·329
—1755	5	8·13 40	3·134	1·209
—1775	10	6·20 48	3·795	2·321
—1795	2	1·73 25	0·267	·041
—1815	0	1·50 37	1·504	1·504
Beyond 1825	2	1·22 91	0.770	·482
	200	200·19 75		$x^2 = 22·699$

The above table gives observed and theoretical values for 20 mm. grouping. These have been plotted both in histogram and in mid-ordinate continuous curve form. (See Plate I).

The equation to the theoretical Gaussian is (in 20 mm. working units):—

$$Y = 23·682 \times \exp.\left\{-\frac{(1656·25 - X)^2}{36·3259}\right\}$$

where X = stature in mm.

Y = frequency.

Mean = 16 56·29 38 mm.

S.D. = 67·38 49 8 mm.

Unit of grouping = 20 mm.

In order to avoid fractions of individuals in theoretical values we stop at 1475 mm. and 1825 mm.

with $n' = 19$ $\chi^2 = 22·699$

From Table XII, p. 26 we find

for $\chi^2 = 22$ $P = \cdot 23\ 19\ 85$

$\chi^2 = 23$ $\cdot 19\ 05\ 90$

$\cdot 04\ 13\ 95$

for $\chi^2 = 22·699$ $P = \cdot 23\ 19\ 85 - \cdot 699 \times (\cdot 04\ 13\ 95)$

Thus $P = \cdot 2030$.

We can now find the probable error of P. Pearson[1] has shown that

$$\sigma_P = \tfrac{1}{2}\sigma_{\chi^2}\{P_q(\chi^2) - P_{q-\xi}(\chi^2)\}$$

and $$\sigma_{\chi_2}{}^2 = \{2(q-1) + q/H + q(q-1)/N\}$$

where q = number of cells and H = harmonic mean of expected frequency.

In the present case, $q = 19$, $N = 200$, $q/H = 4{\cdot}4137$.

Hence, $\sigma_{\chi^2}{}^2 = 42{\cdot}1237$. Giving $\tfrac{1}{2}\sigma_{\chi^2} = 3{\cdot}245$

also $P_{19} = {\cdot}2030$ and $P_{17} = {\cdot}1226$

thus $\sigma_p = 0{\cdot}2609$.

we get finally, $P = {\cdot}2030 \pm {\cdot}1760$.

The chances are 4 to 1 against its being a random sample. In other words about once in five trials we would get worse fits than this. The probable error of P is large. Still the fit is not very bad, for odds of 4 to 1 cannot be considered excessive.

We notice that the contributions of the terminal ranges to χ^2 is heavy, being 3·265, 1·504 and ·482. Combining the two terminal groups at each end we find $\chi^2 = 18{\cdot}482$, and n' 17. We get $P = {\cdot}2978$ which gives a decent fit. In three trials out of ten, random sampling would give us worse fits.

TABLE 3.

Mean = 1656·25 mm. Unit of grouping = 50 mm.

S.D. = 67·3849 mm.

Stature in mm.	Observed Value m'.	Theoretical Value m.	$(m'-m)$.	$\frac{(m'-m)^2}{m}$.
Beyond 1530	8	6·0993	1·9007	·5902
1530—1580	14	19·6830	5·6830	1·6408
—1630	45	43·9045	1·0955	·0231
—1680	60	57·8639	2·1361	·0788
—1730	48	45·0741	2·9259	·1800
—1780	20	20·7464	0·7464	·0268
Beyond 1780	5	6·6292	1·6292	·4002
	200	200·0004	$n'=7$	$x^2=2{\cdot}8399$

From Tables by interpolation, we get

$$P = 0{\cdot}82\ 65\ 83 \pm {\cdot}28\ 86\ 86$$

the probable error, is large, but a high value of P is not improbable.

The fit is now excellent. In 83 trials out of 100 the fit will be worse than this. *We conclude therefore that with 50 mm. grouping, the Gaussian curve is quite adequate for purposes of graduation. With this unit of grouping we may then safely investi-*

[1] *Phil. Mag.* Vol. LII, 1916, pp. 369–378.....

gate the statistical properties of the general population.[1] In subsequent analysis we have for this reason always adopted 50 mm. as our unit. With finer groupings we are likely to obtain mere individual peculiarities of our sample which may not have any connexion whatever with the properties of the general population.

We shall try the effect of other values of Mean and S.D. on "Goodness of Fit."

With 20 mm., $M = 16\ 56{\cdot}2938$, S.D. $= 67{\cdot}13\ 25$

$n' = 19$ $\chi^2 = 25{\cdot}59\ 42$, $P = {\cdot}10\ 98\ 81$

Only once in ten trials the fit will be worse. The end contributions being rather large, we again combine the terminal frequencies and obtain a much better fit.

$n' = 17$, $\chi^2 = 21{\cdot}20\ 72$, $P = {\cdot}1712$

That is once in six trials we will get a worse fit.

TABLE 4.

Summary of "Goodness of Fit."

Mean.	S.D.	n'.	χ^2.	P.
	Unit of grouping = 20 mm.			
1656·29 38 mm.	67·13 15	19	25·59 42	·10 98 81
		17	21·20 73	.17 12
1656·29 38	67·38 49 8	19	22·69 9	·20 30 09
		17	18·48 2	·29 72 74
	Unit of grouping = 50 mm.			
16 56·25 mm.	69·00	7	3·47	·75 67 24
16 56·51	67·21 95	7	2·93 82	·81 56 98
16 56·25	67 47 5	7	3·02 09	·80 65 85
16 56·25	67·38 49 8	7	2·77 88	·83 33 98

20 mm. gives a fit of about the same order in each case. Even with such fine grouping, we get an indication that Gaussian distribution is not impossible, but we cannot assert that the normal curve is fully adequate.

With 50 mm., the fit is excellent in every case. Even with the highest observed value of S.D., namely 69·00 mm., we

[1] This is the reason why 50 mm. is selected as our standard unit of grouping. For purposes of comparison. See page 21.

get P greater than ·75, *i.e.* in three cases out of four, a random fit will be worse. Thus we see that *the effect of different units of grouping (in calculating moment coefficients) on the Goodness of Fit is negligible.*

We must note however that the Goodness of Fit is a much more sensitive criterion than P.E. in judging the accuracy of a S.D. We notice that with S.D. = 67·385 (the value finally adopted) P is ·83, which is substantially better than P = ·75 with S.D. = 69·00 mm.

We conclude that with 50 mm. unit of grouping, a Gaussian curve is fully adequate in every way.[1]

NOTE ON THE LIMITS OF THE UNIT OF GROUPING.

In section I we saw that up to a certain unit of grouping which in our case was 50 mm., the effect of grouping on the frequency constants were negligible. Let this *upper* limit of grouping be h_m On the other hand, in the present section, we have seen that there is a *lower* limit of grouping for which the goodness of fit is satisfactory. Let this lower limit be h_l In our case, it is again 50 mm.

Evidently, the size of h_m and h_l, both depend on the size of the sample. If the distribution is truly Gaussian, then these should depend *only* on the size of the sample and the S.D. It will be extremely useful to obtain even a rough idea about h_m and h_l for any given size of sample.

We can study the problem empirically. We must remember Bernouilli's law which requires that accuracy should depend on the square root of the total number of measurements. As the simplest alternative we can try, if N is the total size of sample and A and B are constants,

$$h_m = A\sqrt{N} \quad \text{and} \quad h_l = B/\sqrt{N}$$

In our case we have, h_m = 50 mm. and h_l = 50 mm. Substituting, we get

$$A = 50/\sqrt{200} = \quad 3{\cdot}53\ 55$$

$$B = 50\ \sqrt{200} = 707{\cdot}10\ 68$$

I provisionally suggest that

(*a*) *In the case of Stature, in calculating frequency constants, the unit of grouping should be less than* $3{\cdot}5\sqrt{N}$.

(*b*) *In testing goodness of fit, the unit of grouping should be greater than* $700/\sqrt{N}$

I do not of course attach much value to the numerical magnitudes of A and B given here; study of a single example is obviously not sufficient. I give the above analysis as a suggestion.

---

[1] This result is well brought out in the 50 mm. graph, but it is quite impossible to judge the goodness of fit by merely looking at a curve.

Adopting the above values of A and B, we get the following table :—

N	h_m	h_l
10	11	222
20	16	157
50	25	100
100	35	70
200	50	50
500	80	30
1000	110	20

With small samples of 10, h_m is 11. Grouping for calculation of frequency constants is thus justified even in the case of small samples. On the other hand for $N=10$, h_l is over 200 mm. which shows the absolute impossibility of judging the adequacy of fit in the case of small samples. In fact with samples of less than 50 (for which $h_l=100$ mm.) it is practically impossible to test the goodness of fit and hence to judge the reliability of any inference about the general population. Even with $N=1000$, the lower limit is not reduced below 20 mm. Thus, discontinuities of less than 20 mm. may easily escape in samples of 1000.

It should be observed that so long as h_l is greater than h_m, we cannot hope to attain great accuracy in judging the significance of a fit so far as the *general* population is concerned. We see, however, that with samples of 200, $h_m=h_l=50$ mm. It then becomes only just possible to assert anything about the population sampled with any certainty. It seems as if 200 is the lower limit of safe sampling for anthropological purposes (at least so far as stature is concerned).

TYPE IV SKEWNESS, LEPTO-KURTOSIS.

For Anglo-Indian Stature, our fundamental constants are (in 50 mm. working units).

Mean =	16 56·79	±3·21 36 mm.
S.D. =	67·38 49 8	±2.55 85 mm.
V =	4·06 72	±
β_1 =	·06 87 56	± ·07 97 81
β_2 =	3·50 46	± ·60 17
Sk. =	+·10 53	± ·05 68
d =	7·09 63	±4·78 18 mm.
μ'_2 =	1·81 62 94	± ·12 24 91
μ'_3 =	−·64 18 53	± ·28 60 80
μ'_4 =	11·56 14 03	±1·81 91 17
β_5 =	·86 66 61	(1)

1 From Biometric Table XLII (*a*), p. 78.

The curve is not significantly skew. But there is distinct tendency towards lepto-kurtosis.

The curve belongs to Type IV of Pearson's Skew Curves.[1] The probable errors of β_1 and β_2 are quite large, and we may investigate whether the $\beta_1-\beta_2$ "probability ellipse" touches the Gaussian point G.[2]

In order to do this we must find Σ_1 and Σ_2, the semi-minor and semi-major axis of the "probability ellipse."

$\beta_1 = \cdot 06\ 87\ 56$

$\beta_2 = 3\cdot 5$ $\quad 1\cdot 177\sqrt{N}.\Sigma_1 = 1\cdot 4 + \cdot 37\ 912 \times [\cdot 4] = 1\cdot 551648$

$= 3\cdot 6$ $\quad 1\cdot 5 + \cdot 37\ 912 \times [\cdot 5] = \dfrac{1\cdot 6885}{\cdot 137912}$

$\beta_2 = 3\cdot 50\ 46,$ $\quad 1\cdot 177\sqrt{N}\Sigma_1 = 1\cdot 55\ 79\ 91$

Similarly $\quad 1\cdot 177\sqrt{N}\Sigma_2 = 13\cdot 51\ 71\ 2$

Multiplying by $\chi_1 = \cdot 04769$, we get

semi-minor axis $= \cdot 0743$

semi-major axis $= \cdot 6446$

Tracing a probability ellipse with these values and centering the ellipse at the point $\beta_1 = \cdot 07$ and $\beta_2 = 3\cdot 5$ approximately, on the diagram on p. 66 of Biometric Tables, we find that the Gaussian point G falls just within the ellipse. We also note that the ellipse covers a small area of the Type III region.

We conclude therefore that a Gaussian distribution itself is not unlikely and may be expected to give a good fit. Type III is not altogether impossible but as the major portion of the ellipse lies within the Type IV region, the lepto-kurtosis is probably just significant.[3]

COMPARATIVE DATA.

Our frequency curve is approximately Gaussian in type The asymmetry is very slight, skewness is small and positive (Mode is greater than the Mean) and the curve belongs to Type IV with lepto-kurtosis.

A. O. Powys[4] has discussed distribution of stature for different age groups of New South Wales criminals. The author says, "by looking at the curves, we see that the material is extremely homogeneous[5] . . . the stature distribution of these

---

[1] See Memoirs cited above in footnote on p. 16.

[2] A discussion of these points is given by A. Rhind: "Additional Tables and Diagram for the Determination of the Errors of Type of Frequency Distribution." *Biometrika* Vol. 7 (1910), p. 386—397.

[3] The asymmetry is very slight and the distance between the Mode and the Mean is also quite small. On the whole there is very little to choose between the "normal" and a Type IV curve. The latter may give slightly improved fit.

[4] A. O. Powys: "Anthropometric Data from Australia," *Biometrika* Vol. 1 (1902), p. 30.

[5] *Ibid.*, p. 38.

homogeneous groups is nearly normal, but what divergence there is lies in the direction of Type IV "[1] In the case of males, the skewness is always positive and the Mode is greater than the Mean.[2] Powys used very long series of measurements extending to several thousands in each age group. The distribution is lepto-kurtic in every case.

W R. Macdonell[3] finds in the case of 3,000 English convicts the stature curve to be of Type IV The skewness is small and negative and there is slight lepto-kurtosis. Mode is less than the Mean.

In the case of Verona statistics[4] the stature of 16,203 conscripts show significant lepto-kurtosis[5] and a Type IV distribution while 3,810 selected recruits show equally significant "platy-kurtosis."[6] Both have significant positive asymmetry and Mode is greater than the Mean.

J F. Tocher[7] finds lepto-kurtosis for the Scottish Insane, the curve belongs to Type IV, and there is small positive skewness, Mode being greater than the Mean[8]. For long series then, viz. New South Wales males, Italian conscripts, Italian recruits and Scottish Insane, there is agreement as to skewness—in all four cases it is significantly positive; in one case, the American recruits,[9] there is quite significant negative asymmetry" American recruits also differ in showing meso-kurtosis.[10]

Charles Goring[11] in the case of the English convict found the distribution approximately Gaussian in type for all crime-groups excepting one. In the only case in which the distribution is significantly different from the normal, the curve is of Type IV with significant lepto-kurtosis and marked positive skewness.

Orensteen[12] found in the case of Cairo-born Egyptians, that the distribution was nearly symmetrical. The criterion K however is less than 1, hence the curve really belongs to Type IV

---

[1] *Ibid.*, p. 39.

[2] *Ibid.*, p. 43. Powys mentions skewness as negative. This is probably a slip.

[3] W. R. Macdonell: "On Criminal Anthropometry and the Identification of Criminals," *Biom.* Vol. 1 (1902), pp. 177—227.

[4] Quoted in Miscellena, *Biom.* Vol. 4 (1906), p. 506 and referred to by J. F. Tocher (see below).

[5] Lepto-kurtic curve are more *sharp-topped* than the normal cu ve, the rise being sharper than the Gaussian.

[6] Platy-kurtic is "flat-topped" as compared to the Gaussian.

[7] J. F. Tocher: "Anthropometric characteristics of the Inmates of Asylums in Scotland," *Biom.* Vol. 5 (1917), pp. 301.

[8] *Ibid.*, p. 182. Tocher says that for long series asymmetry is negative. He evidently means μ_3. This however is slightly ambiguous and may give rise to confusion. I have thought it better to refer to *Skewness* in each case, which has its sign opposite to that of μ_3, so that Mode is greater or less than the Mean according as skewness in positive or negative (and μ_3 negative or positive).

[9] K. Pearson, *Phil. Trans. Roy. Soc.*, Vol. 186 A (1894), p. 385

[10] Meso-kurtosis signifies about the same degree of flatness as the Gaussian.

[11] Charles Goring: "The English Convict," p. 199.

[12] Myer M. Orensteen: "Correlation if Anthropometrical Measurements in Cairo-born Natives," *Biom.* Vol. XI (1915), p. 71.

Conclusion.

(1) The Gaussian curve is quite adequate for graduating a short series of 200 Anglo-Indian measurements. This confirms C. D. Fawcett's rule of normal distribution for short series[1] of anthropometic measurements.

(2) There is some tendency towards *Type IV*, with lepto-kurtosis. All long series, with the exception of American and Italian recruits seem to be definitely lepto-kurtic. It is therefore likely that stature distribution is in general slightly lepto-kurtic in character, but this small lepto-kurtosis does not become statistically significant in small samples.

(3) Skewness is small and positive (Mode being greater than the Mean) for New South Wales criminals, Italian conscripts, Italian recruits, Scottish insane and a short series of several offenders among English criminals. It is negative in the case of several short series of English criminals, and for one long series viz. American recruits. For a short series of Anglo-Indians it is positive but is so small that it cannot be called significant. Hence we conclude that the small skewness of our present sample is not incompatible with homogeneity.

(4) *We conclude therefore that the distribution of stature in the case of Anglo-Indians is of the same nature as in the case of other samples where the material is known to be "homogeneous." In other words, the nature of distribution of stature does not reveal any presence of heterogeneity in the Anglo-Indian population.*[2]

---

[1] *Biometrika* Vol. 1 (1902), p. 443.

[2] Type IV of course is absolutely no indication against homogeneity. For a detailed discussion of this point see K. Pearson: "Skew Variation, a Rejoinder," *Biometrika* Vol. 4 (1905), p. 181.

SECTION V DISSECTION INTO COMPONENT CURVES.

I shall next consider the possibility of statistical dissection of our frequency curve. It might be possible that the sample consisted of two (statistically) different strains. If this were so then it would be possible to break up the frequency distribution into two component normal distributions.

The fundamental memoir on this subject is K. Pearson: "On the Dissection of Asymmetrical Frequency Curves."[1] Pearson has discussed the application of the theory in several[2] actual cases and[3] has given the fundamental equations in a somewhat better form in a paper "On the Problem of Sexing Osteometric Material".[4] I have followed the notation of the fundamental memoir, excepting in one or two instances, where I have used a slightly modified notation.

But before proceeding to a full discussion of the subject it will be useful to apply some simpler tests of homogeneity.

AGREEMENT OF SUB-SAMPLES.

The whole group of two hundred cards were arbitrarily divided into two sub-groups of 100 cards each. The Frequency Constants were calculated for each of these two sub-groups and compared.

The unit of grouping adopted was 50 mm. in each case.

Mean:—

1st group of 100	=	16 58·75 ± 4·64 36 mm.
2nd group of 100	=	16 57·00 ± 4·94 14
Difference	=	1·75 ± 6·78 08[5]

Standard Deviation:—

2nd group	= 73.26	± 3·49 mm.
1st group	= 68·85	± 3·28
Difference	= 4·41	± 4·79

---

[1] *Phil. Trans.* Vol. 184A (1894), pp. 71—110.

[2] K. Pearson: "On the Applications of the Theory of Chance to Racial Differentiation," *Phil. Mag.* 1901, p. 110.

[3] K. Pearson: "On the Probability that two Independent Distributions of Frequency are really Samples of the Same Population, with Special Reference to Recent Work on the Identity of Trypanosome Strains." *Biometrika* Vol. 10 (1915), p. 123 ff.

[4] *Biometrika* Vol. 10 (1915), pp. 479—487.

[5] It is well known that the P.E. of a sum or a difference is given by square root of the sum of the squares of P. E. (see Yule Statistics, p. 211).

Coeff. of variation :—

2nd group	= 4·42 12	±·21 13
1st group	= 4·15 04	±·19 83
Difference	= ·27 08	±·28 97

The difference is in no case significant.

Passing on to the other constants we get :—

μ_2 :—

1st group	= 2·14 66 51	±·20 48
2nd group	= 1·89 60 75	±·18 09
Difference	= 0·25 05 76	±·27 33

μ_3 :—

1st group	– 0·28 75 92	±·51 96
2nd group	– 1·29 90 10	±·43 14
Difference	= 1·01 14 18	±·67 53

μ_4 :—

1st group	– 12·26 96 01	±3·04 55
2nd group	= 11·91 27 23	±2·32 19
Difference	= 0·35 68 78	± .38 30

β_1 :—

1st group	= ·08 357	±·10 99
2nd group	= ·24 890	±·12 04
Difference	= ·16 533	±·16 30

β_2 :—

1st group	= 3·32 56	±·65 80
2nd group	= 2·66 27	±·26 09
Difference	= ·66 29	±·70 78

We conclude that the first hundred measurements are not significantly differentiated from the second hundred in any way. Both represent "random" samples of the same general population.

It should be noted however that the difference between the two samples of hundred each, is of the same order as the probable error of the difference. In one case viz. μ_3, the difference is actually greater than its probable error. This shows that 100 is very nearly approaching the critical limit of " fair (i.e. representative) sampling." [See section III, footnote 8, pp. 32–33].

There is grave danger of samples of less than one hundred being not representative in character (at least so far as the stature of populations of the same order of variability as the Anglo-Indian is concerned). The discussion on p. 40 Section IV shows however that *two hundred* is about the lower limit for *safe* inferences about the general population.

TRIAL SOLUTIONS BY "TAIL" FUNCTIONS.

Consider a mixture of two homogeneous components. If the Means of these components are sufficiently wide apart, the "tail" (i.e. the terminal frequencies) on each side will represent an approximately homogeneous part of the component on that side. Or if the variability of one component is sufficiently greater than the other, the terminal frequencies on its own side will give a fairly homogeneous "tail," even though the Means are not widely different.

We can fit a normal (Gaussian) curve to the "tail," that is, to the terminal frequencies only, with the help of the "tail" functions. If the "tail" is significantly different from the whole sample, then the Gaussian which describes the "tail" satisfactorily may be quite different from the Gaussian which fits the whole sample. For example if we get two "tail" distributions which are each different from the whole distribution, and yet when added together reproduce the total distribution, then we are pretty certain that these "tails" each represent one component of the given sample. Even when we find only one "tail" which is different from the total distribution we can always find the other component by subtraction from the total curve.

This method belongs to the trial and error type. The "tail curves" obtained by considering different portions of the tail, may themselves differ. The uncertainty in the terminal frequencies must be considerable and as Dr. Lee observes, "the chief weakness of the method, besides the assumption of the Gaussian, often quite legitimate, is the absence as yet of the values of probable errors, which must be very considerable for slender material."[1]

For the purposes of "tail" functions, 50 mm. gives too broad groupings. Hence I have found it necessary to work with 20 mm. groupings.

Curtailing at 1585, we get the following:—

Group	1585 –1505 mm.	1505 –1545	1545 –1525	1525 –1505	1505 –1485	1485 –1465	1465 –1445	Total.
Frequency.	10	4	2	4	1	1	2	24

Taking origin at end of range 1585, we get raw moments

$\nu_1' = d = 2{\cdot}20\ 83\ 33$ and $\nu_2' = 8{\cdot}66\ 66\ 67$

$\mu_2 = \Sigma^2 = 3{\cdot}78\ 99\ 31$

[1] K. Pearson and Alice Lee: Generalised Probable Error in Multiple Normal Correlation. *Biometrika* Vol. 6 (1908), pp. 59-68. Alice Lee: Table of the Gaussian Tail Functions. *Biometrika* Vol. 10 (1914), pp. 208-214; Biometric Tables. p. xxvii.

Hence $$\psi_1 = \frac{\Sigma^2}{d^2} = \frac{\mu^2}{\nu_1'^2} = \cdot 69\ 28$$

From Biometric Tables XI, p. 25, we get

$$\left.\begin{aligned} \psi_1 &= 0\cdot 69\ 28 \\ h' &= 0\cdot 77\ 71\ 45 \\ \psi_2 &= 1\cdot 75\ 70\ 30 \end{aligned}\right\}$$

Thus $\sigma = \psi_2 \quad d = 1\cdot 757030 \times 2\cdot 208333$

$= 3\cdot 880107.$

Mean is at a distance $h = \sigma \cdot h' = 3\cdot 015417$ (in working units) from origin.

From Table II:— $n/N = \cdot 21\ 85\ 68\ 5$

Thus we get a normal curve of

N = 110 individuals

Mean = 1645·3 mm.

S.D. = 77·6 mm.

Curtailing at 1605 we get a fresh table:—

Group	1605 1585 mm.	−1565	−1545	−1525	−1505	−1485	−1465	−1445	Total.
Frequency	12	10	4	2	4	1	1	2	36

Calculating "raw" moments about end of stump (1605 mm.) we get

$\nu_1' = d = 2\cdot 30\ 55.56$ $\qquad$ $\nu_2' = 9\cdot 47\ 22\ 22$

giving corrected $\qquad$ $\mu_2 = 3\cdot 71\ 23\ 89$

Thus $$\psi_1 = \frac{\mu^2}{\nu_1'^2} = 9\cdot 64\ 45\ 31$$

From Biometric Tables XI, p. 25, we get by interpolation

$\psi_1 = 0\cdot 64\ 45\ 31$

$h' = 0\cdot 44\ 33$

$\psi_2 = 1\cdot 52\ 32\ 67$

Thus $\sigma = \psi_2 \cdot d = 3\cdot 511747$ (in working units)

or $\sigma = 70\cdot 2349$ mm.

Mean is at distance

$h'\sigma = \cdot 4433 \times 70\cdot 2349$ mm. from 1605 mm.

Thus Mean = 1635·14 mm.

and $n/N = \cdot 32\ 27\ 64\ 2$ (from Table II).

We finally get the following for the *shorter end of the frequency distribution*

N = 112
Mean = 1635·14 mm.
S.D. = 70·23 mm.

This gives a "shorter" group differing in the average stature but with about the *same variability* as the total sample.

Let us now turn to the *taller* end.

Curtailing at 1705, we get

Group	1705 –1725	–1745	–1765	–1785	–1805	–1825	–1845	–1865	Total.
Frequency	18·5	10·0	5·0	10·0	2·0	0	1·0	1·0	47·5

With origin at 1705, raw moments are

$\nu_1' = 2{\cdot}00$
$\nu_2' = 6{\cdot}73\ 42\ 11$, leading to $\mu_2 = 2{\cdot}73\ 42\ 11$
$\psi_1 = 0{\cdot}68\ 33$
Thus $h' = 0{\cdot}70\ 71$
$\psi_2 = 1{\cdot}68\ 18$

and we obtain

N = 198
Mean = 1659·02 mm.
S.D. = 67·27 mm.

which is practically identical with the whole sample.

Thus *the "taller" end seems to represent a homogeneous sample of the whole group, and starting from the taller end, we do not succeed in breaking up the given frequency distribution into two normal sub-groups.*

The "shorter" end gives a pseudo-component. I shall show later on, when we consider the question of age-differentiation that *the shorter tail represents approximately the smaller age groups.*

ASYMMETRICAL DISSECTION.

We have seen that our frequency curve is slightly *asymmetric.* As Pearson observes,[1] "the asymmetry may arise from the fact that the units grouped together in the measured material are not really homogeneous. It may happen that we have a mixture of 2, 3, .. n homogeneous groups, each of which deviates about its mean symmetrically and in a manner represented by the normal curve."

[1] Karl Pearson: "Contributions to the Mathematical Theory of Evolution I. On the Dissection of Asymmetrical Frequency Curves," *Phil. Trans. Roy. Soc.*, Vol. 185A, 1894, p. 72.

Thus an asymmetrical frequency curve may be really built up of normal curves having parallel but not necessarily coincident axes and different parameters. The object of the present section is to discuss the possibiiity of splitting up our asymmetrical frequency curve into *two* component normal curves.[1]

Pearson gave necessary mathematical formulae[2] for this purpose in his memoir of 1894. The solution depends on finding the roots of a numerical equation of the ninth degree, and the arithmetical calculations are extremely laborious. Pearson has discussed the application of the theory in several actual cases.[3]

Let μ_2, μ_3, μ_4 and μ_5 be the moment-coefficients, M the mean and N the total of the given frequency curve. Let m_1, m_2, be the means, σ_1, σ_2, the standard deviations and n_1, n_2 the totals of the component curves.

Then if h is the unit of grouping

$$m_1 = M + \gamma_1 \ h \quad \text{and} \quad m_2 = M + \gamma_2 \ h$$

Also, taking $h = 1$, we have

$$\sigma_1^2 = \mu_2 - \tfrac{1}{3}\mu_3/\gamma_2 - \tfrac{1}{3}p_1\gamma_1 + p_2$$

$$\sigma_2^2 = \mu_2 - \tfrac{1}{3}\mu_3/\gamma_1 - \tfrac{1}{3}p_1\gamma_2 + p_2$$

$$n_1 = -\frac{\gamma_2}{\gamma_1 - \gamma_2}$$

$$n_2 = +\frac{\gamma_1}{\gamma_1 - \gamma_2}$$

Let $p_1 = \gamma_1 + \gamma_2$, $p_2 = \gamma_1 \cdot \gamma_2$ and $p_3 = p_1 \cdot p_2$

Also $\lambda_4 = 9\mu_2^2 - 3\mu_4$, $\lambda_5 = 30\mu_2\mu_3 - 3\mu_5$

Then $$p_3 = \frac{2\mu_3^3 - 2\mu_3\lambda_4 p_2 - \lambda_5 p_2^2 - 8\mu_3 p_2^3}{4\mu_3^2 - \lambda_4 p_2 + 2p_2^3}$$

"Hence, so soon as p_2 is known, $p_1 = p_3/p_2$ can be found, and then γ_1 and γ_2 will be the roots of :—

$$\gamma^2 - p_1\gamma + p_2 = 0$$

The equation for finding p_2 is one of the ninth degree :—

$$24p_2^9 - 28\lambda_4 p_2^7 + 36\mu_3^2 p_2^6 - (24\mu_3\lambda_5 - 10\lambda_4^2)p_2^5 - (148\mu_3^2\lambda_4 - 2\lambda_5^2)p_2^4$$
$$+ (288\mu_3^4 - 12\lambda_4\lambda_5\mu_3 - \lambda_4^3)p_2^3 + (24\mu_3^3\lambda_5 - 7\mu_3^2\lambda_4^2)p_2^2 + 32\mu_3^4\lambda_4 p_2 - 24\mu_3^6 = 0$$

---

[1] *Ibid.*, p. 72. "There are reasons, indeed, why the resolution into two is of special importance. A family probably breaks up into two species, rather than three or more, owing to the pressure at a given time of some particular form of natural selection Even where the heterogeneity may be three-fold or more, the dissection into two is likely to give us, at any rate, an approximation to the chief groups."

[2] The fundamental formulae have been expressed in a slightly modified form in terms of the β-constants in a recent paper "On Sexing Osteometric Measurements." *Biometrika* Vol. 10, 1915, pp. 479—487.

[3] K. Pearson : "On the Applications of the Theory of Chance to Racial Differentiations," *Phil. Mag.* 1901, p. 110.

K. Pearson : "On the Probability that two Independent Distributions of Frequencies are really Samples of the Same Population, etc.," *Biometrika* Vol. 10, 1915, p. 123 *et seq.*

In our case we have, for 50 mm. unit of grouping,

$$\mu_2 = 1{\cdot}82\ 10\ 42$$
$$\mu_3 = -\ 0{\cdot}47\ 78\ 43\ 77$$
$$\mu_4 = 11{\cdot}94\ 16\ 22\ 08$$
$$\mu_5 = -\ 7{\cdot}78\ 23$$

Thus

$$\lambda_4 = -\ 5{\cdot}97\ 91\ 20\ 55$$
$$\lambda_5 = -\ 2{\cdot}75\ 83\ 07\ 24$$

After some laborious arithmetical calculations[1] we find the fundamental nonic :—

$$p_2^9 + 6{\cdot}97\ 56\ 40\ 64p_2^7 + 0{\cdot}34\ 24\ 99\ 30p_2^6 + 13{\cdot}57\ 76\ 59\ 77p_2^5 + 7{\cdot}82\ 66\ 28\ 24p_2^4 + 17{\cdot}63\ 90\ 36\ 20p_2^3 - 1{\cdot}17\ 49\ 41\ 13p_2^2 - 0{\cdot}40\ 73\ 08\ 98p_2 - 0{\cdot}01\ 19\ 04\ 62 = 0.$$

I next form the nine Sturm's auxiliary functions, retaining four figures in the decimal.

$$f_1(x) = 9p_2^8 + 48{\cdot}82\ 95p_2^6 + 2{\cdot}05\ 50p_2^5 + 67{\cdot}88\ 83p_2^4 + 31{\cdot}30\ 65p_2^3 + 52{\cdot}91\ 71p_2^2 - 2{\cdot}34\ 99p_2 - 0{\cdot}40\ 73$$

$$f_2(x) = -1{\cdot}55\ 12p_2^7 - 0{\cdot}11\ 42p_2^6 - 6{\cdot}03\ 46p_2^5 - 4{\cdot}34\ 81p_2^4 - 11{\cdot}75\ 93p_2^3 + 0{\cdot}91\ 38p_2^2 + 0{\cdot}36\ 20p_2 + 0{\cdot}01\ 19$$

$$f_3(x) = -13{\cdot}86\ 58p_2^6 + 21{\cdot}3152p_2^5 - 4{\cdot}68\ 37p_2^4 - 36{\cdot}21\ 80p_2^3 - 54{\cdot}86\ 28p_2^2 + 2{\cdot}28\ 60p_2 + 0{\cdot}40\ 73$$

$$f_4(x) = +\ 1{\cdot}66\ 93p_2^5 - 0{\cdot}54\ 78p_2^4 - 0{\cdot}90\ 53p_2^3 - 9{\cdot}22\ 89p_2^2 + 0{\cdot}09\ 56p_2 + 0{\cdot}06\ 15$$

$$f_5(x) = +\ 6{\cdot}70\ 18p_2^4 + 103{\cdot}78\ 45p_2^3 - 38{\cdot}61\ 84p_2^2 - 1{\cdot}83\ 67p_2 + 0{\cdot}21\ 04$$

$$f_6(x) = -417{\cdot}52\ 59p_2^3 + 160{\cdot}89\ 11p_2^2 + 7{\cdot}18\ 96p_2 - 0{\cdot}89\ 03$$

$$f_7(x) = -2{\cdot}48\ 49p_2^2 + 0{\cdot}01\ 94p_2 + 0{\cdot}01\ 64$$

$$f_8(x) = -5{\cdot}66\ 47p_2 - 0{\cdot}15$$

$$f_9(x) = -0{\cdot}01\ 41$$

We can now find the number of real roots from the changes of sign in the Sturm's functions.

	$+\infty$	0	$-\infty$
$f(x)$	+	−	−
$f_1(x)$	+	−	+
$f_2(x)$	−	+	+
$f_3(x)$	−	+	−
$f_4(x)$	+	+	−
$f_5(x)$	+	+	+
$f_6(x)$	−	−	+
$f_7(x)$	−	+	−
$f_8(x)$	−	−	+
$f_9(x)$	−	−	−

[1] My best thanks are due to Prof. J. M. Bose M.A., B.Sc. of the Mathematics Department of the Presidency College, Calcutta for his kind help in checking the arithmetic in many places.

There are 3 changes of sign with $x = +\infty$, 4 changes with $x = 0$ and 6 changes with $x = -\infty$ Hence there is $4-3=1$ real positive root and $6-4=2$ real negative roots.

By trial I locate the positive root between 0 and 1, and the two negative roots between 0 and -1.

I try the following successive approximations by Horner's method.

$f(+0{\cdot}2) = +{\cdot}01\ 77$ $\quad$ $f(+0{\cdot}15) = -{\cdot}03\ 79$

$f(+0{\cdot}18) = -{\cdot}01\ 17$ $\quad$ $f(+0{\cdot}187) = -{\cdot}00\ 09$

$f(+0{\cdot}188) = +{\cdot}00\ 02$ $\quad$ $f(+0{\cdot}1878) = -{\cdot}00\ 002$

Thus we can take the positive root, $p_2 = +0{\cdot}1878$

For the negative roots I try

$f(0) = -{\cdot}01\ 19$ $\quad$ $f(-{\cdot}5) = -2{\cdot}27\ 75$

$f(-{\cdot}25) = -{\cdot}44\ 88$ $\quad$ $f(-{\cdot}01) = -\ {\cdot}00\ 80$

$f(-{\cdot}1) = +{\cdot}00\ 01$

Root is near $-{\cdot}1$. I try higher approximations, now retaining eight decimal figures.

$$\begin{aligned} f(-{\cdot}1) &= +0{\cdot}00\ 00\ 84\ 36 \\ f(-{\cdot}101) &= -0{\cdot}00\ 02\ 54\ 15 \\ f(-{\cdot}1001) &= +\ {\cdot}00\ 00\ 51\ 06 \\ f(-{\cdot}1003) &= -\ {\cdot}00\ 00\ 44\ 78 \\ f(-{\cdot}1002) &= +\ {\cdot}00\ 00\ 14\ 65 \end{aligned}$$

Thus $p_2 = -{\cdot}1002$ is another root.

Again

$$\begin{aligned} f(-{\cdot}05) &= +{\cdot}00\ 34 \\ f(-{\cdot}01) &= -{\cdot}00\ 80 \\ f(-{\cdot}03) &= -{\cdot}00\ 12 \\ f(-{\cdot}04) &= +{\cdot}00\ 14 \\ f(-{\cdot}034) &= -.00\ 00\ 97\ 79 \\ f(-{\cdot}0343) &= -{\cdot}00\ 00\ 17\ 84 \\ f(-{\cdot}0344) &= +{\cdot}00\ 00\ 08\ 69 \end{aligned}$$

Thus $p_2 = -{\cdot}0344$ is the third root.

It should be observed that if the material is a real mixture of two true *normal* components, then the mathematical solution would be theoretically *unique*. In practice, however, a statistical curve may be the sum of two asymmetric curves, and hence we must not be surprised if more than one solution is given by the present method of dissection. Each root of the fundamental nonic gives one distinct mode of dissection.

Case 1.

$$p_2 = +\ 0{\cdot}18\ 78$$

Then, $$p_3 = -\ 5{\cdot}28\ 28\ 44$$

$$p_1 = p_3/p_2 = -28{\cdot}11\ 01\ 59$$

Hence γ_1 and γ_2 are roots of

$$\gamma^2 + 28{\cdot}13\ 017 + 0{\cdot}18\ 78 = 0$$

We get

$$\gamma_1 = -\ 0{\cdot}00\ 665$$
$$\gamma_2 = -28{\cdot}12\ 345$$

We obtain, finally, for the *first component*,

$$\sigma_1^2 = \quad 1{\cdot}94\ 08\ 23$$
$$\sigma_1 = \quad 1{\cdot}39\ 31$$
$$n_1 = \frac{28{\cdot}12\ 345}{28{\cdot}11\ 680} \cdot 200$$
$$= 200{\cdot}0473 = 200, \text{ to the nearest integer,}$$

and $$m_1 = 1655{\cdot}91\ 75 \text{ mm.}$$

The *second component* is given by

$$\sigma_2^2 = -285{\cdot}64\ 89\ 43$$
$$n_2 = -\ 0{\cdot}04\ 73$$
$$m_2 = \quad 250{\cdot}08 \text{ mm.}$$

The second component has σ^2 *negative*, and is thus *imaginary*. Hence dissection into two *real* components is impossible in this case. The first component, which is the only real component, gives practically the whole of the given sample. The total frequency of the second component is only $-{\cdot}04\ 73$ and is quite negligible.

Case 2.

$$p_2 = -\ 0{\cdot}10\ 02$$

We find $$p = +\ 1{\cdot}21\ 09\ 58\ 36$$

and $$p_1 = -12{\cdot}08\ 54\ 13$$

Thus $$\gamma^2 + 12{\cdot}08\ 54\ 137 - {\cdot}10\ 02 = 0$$

and

$$\gamma_1 = +\ \ {\cdot}00\ 82\ 85$$
$$\gamma_2 = -12{\cdot}10\ 19\ 90$$

We get for the *first component*,

$$n_1 = \quad 199{\cdot}86\ 31$$
$$\sigma_1^2 = \quad 1{\cdot}74\ 10\ 46$$
$$\sigma_1 = \quad 1{\cdot}31\ 94\ 87$$
$$m_1 = 1656{\cdot}66\ 42 \text{ mm.}$$

The *second component* is

$$n_2 = +\ 0{\cdot}13\ 69$$
$$\sigma^2 = -29{\cdot}03\ 96\ 23\ 71$$
$$m_2 = 1051{\cdot}98 \text{ mm.}$$

We again find that the first curve gives practically the whole of the given sample, while the second is *imaginary*.

Case 3.

$$p_2 = -0{\cdot}03\ 44$$

Whence $p_3 = +0{\cdot}17\ 10\ 76$

$$p_1 = -4{\cdot}97\ 31\ 40$$

Thus $\gamma^2 + 4{\cdot}97\ 31\ 407 - {\cdot}03\ 44 = 0$

$$\gamma_1 = +\ {\cdot}00\ 69\ 075$$

$$\gamma_2 = -4{\cdot}98\ 00\ 475.$$

First component

$$\text{Mean} = 1656{\cdot}59\ 54 \text{ mm.}$$

$$n_1 = 199{\cdot}72\ 3$$

$$\sigma_1^2 = 1{\cdot}76\ 61\ 09$$

$$\sigma_1 = 1{\cdot}32\ 89\ 50$$

Second component

$$\text{Mean} = 1407{\cdot}24\ 76 \text{ mm.}$$

$$n_2 = +\cdot\ {\cdot}27{\cdot}7$$

$$\sigma_2^2 = +16{\cdot}59\ 03\ 29$$

The second component is *real* this time, but its frequency being only ·277, it is again negligible. The first component gives practically the whole of the distribution.

It will be seen that first solution ($p_2 = {\cdot}1878$) gives the frequency curve as the *difference* of two normal curves. "The probability curve, with positive area, may possibly be looked upon as the birth population (unselectively diminished by death). The negative probability curve is a selective diminution of units about a certain mean; that mean may, perhaps be the average of the less fit."[1] In our present case, however, the negative component is imaginary. Hence we conclude that the real component is describing the general population with sufficient accuracy.

In the case of the second solution ($p_2 = -{\cdot}1002$) the second component, though now additive, is still imaginary. The mean is at 1051·98 mm. This component may be interpreted as representing a "tendency" towards the presence of a small proportion of dwarfs.

This tendency becomes more prominent in the third solution ($p_2 = -{\cdot}0344$). We find that the second component, which is additive and real, definitely represents a "dwarf" distribution with an average stature of 1407·24 mm. The proportion, however, is extremely small. It is only 0·14% and can be safely neglected in samples of 200. In larger samples of over a thousand, we should not be surprised to get a few dwarfs.

So far as the present analysis goes we must conclude therefore that it is not possible to break up our given curve into two real

[1] Pearson, *Phil. Trans. Roy. Soc.*, Vol. 185 A, 1894, p. 76.

significant component distributions. The only sign of differentiation perceived so far is a tendency towards the presence of a very small proportion of dwarfs.

SYMMETRICAL DISSECTION.

We have already seen that β_1 (which measures the deviation from symmetry) is not significantly different from zero in our present case. In other words, within the limits of probable errors it is quite possible to look upon our curve as a symmetrical one. "Another important case of the dissection of a frequency curve can arise, when the frequency curve, without being asymmetrical, still consists of the sum or difference of two components, i.e. when the means about which the components groups are distributed are identical. This case is all the more the interesting and important, as it is not unlikely to occur in statistical investigations, and the symmetry of the frequency-curve is then in itself likely to lead the statistician to believe that he is dealing with an example of the normal frequency-curve." [1]

Pearson also notes that "symmetry may arise in the case of compound frequency curves, even without identity of the means of the components. In this case, for two components, we should have for different means, equality of component group totals and their standard deviations. This equality seems less likely than equality of means and divergence of totals and standard deviations." [2]

Pearson then shows that for this second type of symmetrical dissection (i.e. divergent means) a necessary condition is that $3\mu_2^2$ should be *greater* than μ_4, that is β_2 should be *less* than 3, or the curve should be platy-kurtic. But we have seen that our curve is lepto-kurtic (i.e. $3\mu_2^2$ is *less* than μ_4), hence this type of dissection is impossible in the present case.

I shall now discuss the possibility of the first type of symmetric dissection. The fundamental equations are given in the *Memoir* cited, p. 90. I shall slightly modify these equations in order to express them in terms of the β-variables.

Let N, n_1, n_2, represent the totals and Σ, σ_1 and σ_2 the standard deviations of the compound and the two component curves respectively. Then, as Pearson has shown, the solution is given by

$$n_1 = \frac{\mu_2 - w_2}{w_1 - w_2} N \qquad n_2 = \frac{w_1 - \mu_2}{w_1 - w_2} . N$$

$$\sigma_1^2 = w_1 \qquad \sigma_2^2 = w_2 \quad \text{where} \quad \mu_2 = \Sigma^2$$

and w_1 and w_2 are the roots of

$$(\mu_4 - 3\mu_2^2)w^2 + (\mu_2\mu_4 - \tfrac{1}{5}\mu_6)w - (\tfrac{1}{3}\mu_4^2 - \tfrac{1}{5}\mu_2\mu_6) = 0$$

---

[1] Karl Pearson: "On the Dissection of Symmetrical Frequency Curves," *Phil. Trans. Roy Soc.*, Vol. 185A, 1894, p. 90.

[2] *Ibid.*, footnote on pp. 90-91

This equation involves μ_6. We can however transform this equation to the β-variables.

Dividing throughout by μ_2^4, we get

$$\left(\frac{\mu_4}{\mu_2^2}-3\right)\frac{w^2}{\mu_2^2}+\frac{1}{5}\left(\frac{5\mu_4}{\mu_2^2}-\frac{\mu_6}{\mu_2^2}\right)\frac{w}{\mu_2}-\frac{1}{15}\left(\frac{5\mu_4^2}{\mu_2^4}-3\frac{\mu_6}{\mu_2^8}\right)=0$$

But $\beta_2=\mu_4/\mu_2^2$ and $\beta_4=\mu_6/\mu_2^3$

Changing to the β-variables and putting $x=w/\mu_2$ we get

$$(\beta_2-3)x^2+\frac{5\beta_2-\beta_4}{5}x-\frac{5\beta_2^2-3\beta_4}{15}=0$$

Thus $$x=\frac{w}{\mu_2}=\frac{\frac{1}{5}(\beta_4-5\beta_2)\pm\sqrt{\frac{1}{25}(\beta_4-5\beta_2)^2+\frac{4}{15}(\beta_2-3)(5\beta_2^2-3\beta_4)}}{2(\beta_2-3)}$$

The condition for a real solution is that

$$\tfrac{1}{5}(\beta_4-5\beta_2)>\sqrt{\tfrac{1}{25}(\beta_4-5\beta_2)^2+\tfrac{4}{15}(\beta_2-3)(5\beta_2^2-3\beta_4)}$$

Squaring and substracting

$$0>\tfrac{4}{25}(\beta_2-3)(5\beta_2^2-3\beta_4)$$

Pearson has shown that it is necessary that w_1 and w_2 should be of the same sign.

The necessary condition for real solution becomes:—

For lepto-kurtic curves, $\beta_2-3>0$ or $\beta_2>3$, it is necessary that $3\beta_4$ should be greater than $5\beta_2^2$.

For platy-kurtic curves, $\beta_2-3<0$ i.e. $\beta_2<3$, the condition is that $5\beta_2^2$ must be greater than $3\beta_4$.

With ungrouped distribution it is almost impossible to find β_4 directly. We can however find β_4 in terms of β_1 and β_2, from Table XLII (*b*), p. 78 of Tables for Biometricians and Statisticians.[1]

We have, $\beta_1=\cdot06\ 87\ 56$

$\beta_2=3\cdot50\ 46$

For $\beta_2=3\cdot5$ $\beta_4=23\cdot72\ 89+\frac{68\ 756}{100\ 000}\times[2\cdot0142]=25\cdot11\ 37$

$4\cdot0$ $31\cdot00+\frac{68\ 756}{100\ 000}\times[10\cdot766]=28\cdot40\ 23$

$\beta_2=3\cdot50\ 46$ $\beta_4=25\cdot11\ 37+\frac{46}{5000}\times[13\cdot28\ 86]$

$=25\cdot23\ 60$

We have β_2 greater than 3, and $3\beta_4$ greater than $5\beta_2^2$ hence we shall obtain a real solution.

The quadratic is

$$\cdot50\ 46x^2-1\cdot54\ 26x+0\cdot95\ 31\ 27=0$$

---

[1] Cf. K. Pearson: "Skew Correlation and Non-Linear Regression", p. 8 (Draper's Company Research Memoirs).

The solution is given by

$$w_1 = 2{\cdot}19\ 75$$
$$w_1 = \cdot 66\ 89$$

Since $\mu_2 = 1{\cdot}81\ 62$

We get $\sigma_1 = 1{\cdot}48\ 24$ $\qquad$ $\sigma_2 = \cdot 81\ 78\ 63$

$= 74{\cdot}12$ mm. $\qquad$ $= 44{\cdot}89$ mm.

And $n_1 = \dfrac{1{\cdot}14\ 73}{1{\cdot}52\ 86}\ 200$ $\qquad$ $n_2 = \dfrac{\cdot 38\ 13}{1{\cdot}52\ 86} \times 200$

$= 150{\cdot}11$ $\qquad$ $= 49{\cdot}89$

It is thus possible to break up the curve into two normal curves with the same Means but widely different Standard Deviations. It will be observed that nearly three-fourths of the sample has got a greater variability, while about one-fourth seems to be a very stringently selected group. This particular solution may be only a peculiarity of the sample and may have no reference to actual fact so far as the general population is concerned. A calculation of the probable error of β_4 may throw some light on the question.

Pearson[1] gives the percentage variation of β_4 to be 23·3 in a sample of 500. Multiplying this by

$$\sqrt{500/200} = \sqrt{2{\cdot}5},$$

we get the percentage variation in a sample of 200 to be 36 84. Hence the probable error in the present case is so large as $\pm 9{\cdot}28$.

We thus have $\beta_4 = 25{\cdot}236 \pm 9{\cdot}28$

If we take our actual value of $\beta_2 = 3.5$, the necessary condition for a real solution is that β_4 must be greater than 20·42. If the value of β_4 for the general population is less than 20·42 (with a value of $\beta_2 = 3{\cdot}5$) then the present method of dissection will fail.

This limiting value is only 4·82 less than the value of β_4 in the sample, while the probable error is $\pm 9{\cdot}28$. It is therefore not at all unlikely that β_4 should be less than 20·42 in the general population. We conclude therefore that it is not unlikely that the possibility of this particular type of dissection is only a peculiar property of the sample and has no reference to actual fact in the case of the general population.

Hence we are not justified, on this evidence alone, in concluding that the sampled population is heterogeneous in character.

Note added on the 27th November, 1920.

In view of the great importance of the question of heterogeneity I thought it desirable to consider this question in greater

[1] K. Pearson: "Skew Correlation and Non-Linear Regression", p. 8 (Draper's Company Research Memoirs).

detail. I calculated the grouped moment-coefficients directly upto μ_6 with 50 mm. as the unit of grouping. I find

$$\mu_2 = 1{\cdot}82\ 10\ 42$$
$$\mu_3 = -\ 0{\cdot}47\ 78\ 43\ 77$$
$$\mu_4 = 11{\cdot}94\ 16\ 22\ 08$$
$$\mu_5 = -\ 7{\cdot}78\ 23$$
$$\mu_6 = 129{\cdot}74\ 38\ 42\ 48$$

Thus

$$\beta_2 = 3{\cdot}601$$

and

$$\beta_4 = 21{\cdot}474$$

Since β_2 is greater than 3, it is necessary that $3\beta_4$ should be greater than $5\beta_2$. Actually we find

$$3\beta_4 = 64{\cdot}42\ 20$$

while $5\beta_2^2 = 64{\cdot}83\ 60$, so that $5\beta_2^2$ is $> 3\beta_4$

Thus no real solution is possible in this case. But we must note that there is some tendency towards a solution of this type. I do not propose to draw any inference from this result. I have not yet analysed the other frequency curves and so I am not in a position to either confirm or refute this *tendency* towards a very special type of splitting up.[1]

Goodness of fit with Sum of Dissected Components.

First component :—

Mean stature = 16 56·79 mm.
S.D. $= \sigma_1$ = 74·12 mm.
Total $= n_1$ = 150

Second component :—

Mean stature = 16 56·79 mm.
S.D. $= \sigma_2$ = 40·89 32 mm.
Total $= n_2$ = 50

I First Component.	II Second Component.	I+II (Total) Theoretical m'.	Observed m.	$m \sim m'$.	$\frac{(m-m')^2}{m'}$.
6·53 48	0.04 71	6·58 18	8	1·41 82	·30 55
15·97 53	1·46 22	17·43 75	14	3·43 75	·67 77
31·31 52	11·29 66	42·61 18	45	2·38 82	·13 07
39·43 29	22·93 16	62·36 45	60	2·36 45	·08 96
32·23 82	12·42 61	44·66 43	48	3·33 57	·24 91
17·26 74	1·77 15	19·03 89	20	·96 11	·04 85
7·23 61	·06 47	7·30 08	5	2·30 08	·72 46
		$n'=7$,			$\chi^2 = 2{\cdot}22\ 57$

[1] Since going to press, I have obtained expressions for the Probable Errors of the Component Frequency Constants, which confirms the non-significant character of the dissection in the present case. I hope to publish these new formulae for Probable Errors at an early date.

Thus, $P = \cdot 89\ 46\ 80$

With the single normal curve, we had $P = \cdot 82\ 65\ 83$

Difference $= \cdot 06\ 80\ 97$

Thus there is an improvement of 8·2% in the fit. This is satisfactory. But, in view of the discussion of probable errors perhaps this is not sufficient to warrant us in asserting that the possibility of the present type of dissection is unmistakeable evidence of heterogeneity of the material.

SECTION VI. DATA FOR COMPARISON.

SOURCE OF THE MATERIAL.

I have collected material from many different sources. In 1897, K. Pearson [1] gave the coeff. of variation for 1000 English middleclass men, 390 Bavarian men, 284 French (from statistics given in "Memoires de la Societe d'Anthropologie de Paris," 1888) and also some data for American school children (from the years 6 to 10, taken from Porter's "Growth of Saint Louis Children"). I have retained his French and German data but have substituted corrected values for Englishmen given by Pearson in a later paper. I have omitted the children as being all under the age of 10.

Pearson also reduced statistics for U.S.A. recruits [2] and gave final figures for his family data [3] in Biometrika in 1903. His family data consists of 1078 records of middle class English fathers and sons.

Powys [4] gave the heights of 2862 male criminals from New South Wales, distributed into different age-groups. I have selected the total variability [5] of the whole group, for in our Anglo-Indian data men of all ages are present. Powys considers his data to be "extremely homogeneous."[6]

In 1901, W R. Macdonell [7] discussed the measurements for 3000 English criminals. He also calculated the coeff. of variation for 1000 Cambridge undergraduates.[8]

Raymond Pearl [9] has calculated variabilities of stature for 416 Swedes, 475 Hessians, 266 Bohemians, and 365 Bavarians.[10] The measurements were all taken on dead bodies and the coeff. of variation are $4{\cdot}009 \pm {\cdot}094$, $3{\cdot}954 \pm {\cdot}117$, $4{\cdot}323 \pm {\cdot}127$ and $3{\cdot}838 \pm {\cdot}096$ respectively.

Blakeman [11] has analysed a short series of 117 English males who died in hospitals. The coeff. of variation [12] for stature is

---

[1] K. Pearson: "Chances of Death," Vol. 1, pp. 294—296.

[2] *Phil. Trans. Roy. Soc.*, Vol. 184A, p. 386.

[3] *Biometrika* Vol. 2 (1903), p. 370; K. Pearson and Alice Lee: "On the Laws of Inheritance in Man," pp. 357—482.

[4] A. O. Powys: "Anthropometric Data from Australia," *Biometrika* Vol. 1 (1901), pp. 30–49.

[5] *Ibid.*, p. 44. [6] *Ibid.*, p. 38.

[7] W. R. Macdonell: "On Criminal Anthropometry and the Identification of Criminals," *Biometrika* Vol. 1 (1901), pp. 177—277.

[8] *Ibid.*, p. 189.

[9] Raymond Pearl: "Variation and Correlation for Brain Weight," *Biometrika* Vol. 4 (1905), pp. 13–104.

[10] *Ibid.*, p. 23.

[11] J. Blakeman: "A Study of the Biometric Constants of English Brain-Weights, and their Relationships to External Physical Measurements," *Biometrika* Vol. 4 (1905), pp. 124–160.

[12] *Ibid.*, p. 126.

4·55±·20. Blakeman believes[1] the "increased variability in stature to be due to the measurements being taken on the corpse and not on the living subject." He mentions further[2] that the average V for males in Pearl's data is 4·11.

I have thought it best to omit the above series of corpse data for purposes of comparision. It will be observed that the variability is in each case considerably higher than the average variability (which is about 3·6) obtained by omitting them. Thus the only effect of including the "corpse" data would be to still further increase our average variability. We may further note that in most of the above cases, the variability is even higher than the variability of our Anglo-Indian data, which is about 4·06. Thus omission of the corpse date cannot affect our general conclusion that the variability of the Anglo-Indian series is *not* significantly greater than the average variability of stature for homogeneous material.

Tocher[3] gave in 1906, a very large series of measurements on the Scottish Insane, numbering 4381 males.

Schuster[4] in 1910 gave V for different age-groups of Oxford undergraduates. For reasons already explained I have taken the average variability for the whole group of 959 individuals. In an editorial note to the above,[5] some results for 493 Scottish (Aberdeen) undergraduates are quoted. I have calculated the coeff. of variability in this latter case also. I may note in passing that the different age-groups of the Oxford data do not give lower values of variability, in fact give slightly greater values than the total in many cases.[6]

Craig[7] gave the results of a very large series of measurements of modern Egyptians. These were classified in accordance with the town or district of birth.[8] The total number in each group is fairly large and this series gives us a very good list of variabilities for purposes of comparison. I have retained the separate variability for Aswan, omitting the total variability as the material is not homogeneous.

Garett[9] has given a series of measurements of the natives of Borneo and Java. The majority were coolies in the employ of the author. Unfortunately the number in the case of each people is not very extensive, and I have been only able to retain the values

---

[1] *Ibid.*, p. 131. [2] *Ibid.*, p. 132.

[3] J. F. Tocher: "The Anthropometric Characteristics of the Inmates of Asylums in Scotland.," *Biometrika* Vol. 5 (1906), pp. 298—350.

[4] E. Schuster: "First Results from the Oxford Anthropometric Laboratory," *Biometrika* Vol. 8 (1911), pp. 40-51.

[5] *Ibid.*, p. 49.

[6] Thus the lumping together of all age-groups cannot again affect the general validity of our conclusions.

[7] J. I. Craig "Anthropometry of Modern Egyptians," *Biometrika* Vol. 8 (1911), pp. 69—77.

[8] *Ibid.*, p. 75.

[9] T R. H. Garett: "Natives of the Eastern Portion of Borneo and Java," *Jour. Roy. Anthrop. Inst.*, Vol. XLII, 1912, pp. 60—66.

for Javanese (17), Banjerese (33) and Sundanese (37), as no other series includes more than 7 individuals.

Joyce [1] has given figures for 25 different groups of people of Chinese Turkestan and the Pamirs. But again the total number is rather small in most cases, even the longest series including only 67 individuals.

Leys and Joyce [2] gave measurements for 38 different groups of people from East Africa. Some of these are foreigners. Numbers are moderately large in some cases, the longest series containing 384 individuals.

Seligmann [3] has given measurements for 7 groups of people of Anglo-Egyptian Sudan. The number in each group is moderately large, being on an average about 50. Dr. Bowley has analysed the Dinka group containing 116 individuals. The absolute S.D. (9·66 mm.) as well as the coeff. of variation (5·4311) is exceptionally high. Dr. Bowely [4] concludes from the goodness of fit that "there is no indication of the mixture of two distinct groups with widely differing averages." [5]

Frankly speaking, such a high value of V as 5·4311 ± ·24 for homogeneous material is extremely puzzling. We have of course obtained several high values of V, but in all such cases the numbers are quite small and the P.E. quite large. One would like to obtain independent evidence regarding the homogeneity of the Dinka people. In any case, a fresh series of measurements of the Dinka people is urgently needed.

Goring [6] has given extensive data for English criminals, to which we shall have to refer again.

Whiting [7] has discussed the case of 500 English convicts belonging to Dr. Goring's data.

Orensteen [8] gave results for 802 adult male Egyptians born in Cairo.

Addendum.

Dudley Buxton has recently published the Variabilities of 10 Mediterranean and 3 Jewish races.[9]

---

[1] T. A. Joyce: "Notes on the Physical Anthropology of Chinese Turkestan and the Pamir," *Jour. Roy. Anthrop. Inst.*, Vol. XLII, 1912, p. 450.

[2] Norman M. Leys and T. A. Joyce: "Note on a series of Physical Measurements from East Africa," *Jour. Roy. Anthrop. Inst.* Vol. XLIII, 1913, p. 195.

[3] C. G. Seligmann: "Some Aspects of Hamitic Problem in the Anglo-Egyptian Sudan," *Jour. Roy. Anthrop. Inst.* Vol. XLIII, 1913, pp. 592–705.

[4] *Ibid.*, p. 705.

[5] In the absence of any attempt at statistical dissection, mere homotyposis in graduation cannot be considered conclusive evidence of homogeneity.

[6] Charles Goring: "The English Convict," 1913.

[7] Madeline H. Whiting: "On the Association of Temperature, Pulse and Respiration with Physique and Intelligence in Criminals," *Biometrika* Vol. 11 (1915), pp. 1–37.

[8] Myers M. Orensteen: "Measurements of Cairo-born Egyptians," *Biometrika* Vol. 11 (1915), pp. 67-81.

[9] *Biometrika*, Vol. XII, 1920, pp. 92–112.

N.B.—I may note that in many cases, the Coeff. of Variation has been calculated by me.

Risley[1] published the crude measurements of 87 Indian castes and tribes, but he did not calculate a single frequency constant or a single probable error. The size of sample varies from 185 to 2, yet every average has been given equal weight on the strength of his authority. The averages published in his book were in many cases hopelessly wrong, in one instance the difference amounted to no less than 60 mm.

I have just finished calculating the frequency constants for the whole of Risley data for Stature. I hope to publish my results at an early date. Meanwhile I shall use my summary table for purposes of comparison in this paper.

It should be noted that the present section was already submitted to the press when the Mediterranean data reached me. The Risley data also had not then been reduced. Thus the earlier part of the present section does not include the above two series of data. I have retained a portion of the older work, but have gone over the whole ground again with the inclusion of the new data.

The Caste data of Risley is substantially differentiated from other samples in showing a significant lower Variability, hence the Anglo-Indian sample is found to be significantly more variable than the Indian Castes and Tribes. Otherwise the inclusion of the new data does not upset the earlier conclusion that the Anglo-Indian Variability, though higher than the general Variability of " homogeneous " races, is not significantly different. As a matter of fact Anglo-Indian Variability is just about the same as the Variability of European (in a geographical sense only) races.

NOTE ON THE RETENTION OF CRIMINAL DATA.

It may be objected that a criminal population being substantially differentiated from the general population, it is not legitimate to use criminal data for comparative purposes. We can only reply that if there is any fundamental anthropological differentiation this has not yet been proved to be the case. On the other hand the bulk of available statistical evidence goes to show that there is no such thing as a different criminal type. J J Craig[2] says of his Egyptan data, " it may be objected that criminality in itself is a determining factor of selection, but the objection does not hold in Egypt " and he proceeds to explain why In the case of New South Wales also the same is true. There is no significant differentiation of criminals from the general population.[3]

As regards the English convict, we need only refer to the great work on the subject by Dr. Charles Goring (already cited several times in this paper). Goring comes to the conclusion that the Lombrosian doctrine of criminal types is false. " Criminals as

---

[1] " *Indian Castes and Tribes*," 2 Vols. (1904?) (Superintendent of Government Printing, Calcutta).

[2] J. I. Craig: *loc. cit. Biom.* Vol. 8 (1911).

[3] Goring: " The English Convict," (1913), p. 198.

criminals are not a physically differentiated class of the general community. The physical and mental constitution of both criminals and law abiding persons of same age, stature, class and intelligence are identical. There is no such thing as an anthropological criminal type.[1]" In view of Goring's work we may safely include criminal data for purposes of comparison, at least until statistical evidence in support of the Lombrosian doctrine is forth-coming.

TABLE 5.

Mean Stature, S.D. and Coeff. of Variation of 100 *different races.*

Note.—(1) The number immediately after the name of the race gives the reference of the source from which material is collected (see end of table).

(2) Second column gives number of individuals on which the average is based.

(3) *Races italicised were selected as more reliable.* It will be noticed that the total number in each case is greater than 25, and the P.E. of Coeff. of Variation is less than ·32 or $\sqrt{\cdot 1}$.

Name of Race.	Col. 2 No. in Sample.	Mean (mm.) ± P.E. of Mean.	S.D. in mm. ± P.E. of S.D.	100 × (Coeff. of Var. ± P.E. of V.)
1 Segua (1) ..	12	1670· ± 5·7	29·46 ± 4·0	176·42 ± 24·28
2 Digo (1)	15	1629·4 ± 5·9	33·78 ± 4·2	207·32 ± 25·52
3 Nyika (1)	18	1658·1 ± 6·3	39·37 ± 4·4	237·43 ± 26·68
4 Comoro (1)	23	1662·9 ± 5·9	41·66 ± 4·1	250·49 ± 24·91
5 Kaseri (1) ..	12	1696·5 ± 8·6	43·94 ± 6·0	259·02 ± 35·66
6 Javanese (2)	17	1570·59 ± 6·71	43·3 ± 6·4	261· ± 34·
7 Kelpin (3)	15	1650·00 ± 9·8	44·6 ± 7·0	270·30 ± 33·28
8 *Sarikoli* (3)	40	1637·7 ± 6·0	44·3 ± 4·3	270·50 ± 20·39
9 Nandi (1)	14	1676·4 ± 8·3	45·9 ± 5·9	274·24 ± 34·95
10 *Lamu* (1) ..	26	1637·0 ± 5·9	44·96 ± 4·2	274·63 ± 25·68
11 Dolan (3) .	16	1641·1 ± 9·5	46·10 ± 6·7	280·89 ± 33·49
12 *Muscat Arab* (1)	31	1648·4 ± 5·8	47·8 ± 4·1	289·67 ± 24·81
13 Faizabad (1)	12	1669·2 ± 11·0	49·2 ± 7·8	294·75 ± 40·58
14 Shilluk (5) ..	14	1776·0 ± 9·6	53·0 ± 6·8	298·42 ± 38·04
15 *Baganda* (1)	44	1664·7 ± 5·1	50·3 ± 3·6	302·10 ± 21·72
16 Hami (3)	21	1630· ± 8·3	49·5 ± 5·9	303·68 ± 31·60
17 Yemen Arab (1) .	20	1647·7 ± 7·6	50·29 ± 5·4	305·22 ± 32·55
18 *Swahili* (1) ..	53	1646·7 ± 4·7	50·3 ± 3·3	305·41 ± 20·01
19 *Wanyamwezi* (1)	101	1764·9 ± 3·5	51·6 ± 2·4	307·85 ± 14·61
20 Nissa (3)	9	1602·2 ± 12·7	49·5 ± 9·0	308·95 ± 49·11
21 *Pakhpo* (3)	25	1604·0 ± 7·6	49·5 ± 5·4	308·60 ± 29·43
22 *Segeju* (1)	36	1631·1 ± 5·7	50·5 ± 4·0	309·82 ± 24·62
23 Chinese (3)	20	1667·0 ± 8·5	51·7 ± 6·0	310·97 ± 33·08
24 *Banjerese* (2)	33	1569·64 ± 5·71	48·61 ± 4·04	310· ± 26·
25 Niya (3)	18	1626·0 ± 9·0	50·4 ± 6·4	310·15 ± 34·86
26 Karnaghu-Tagh (3)	21	1660·5 ± 8·3	52·9 ± 5·9	318·57 ± 33·15
27 *Canal Egyptians* (4)	127	1658·7 ± 3·2	54·2 ± 2·3	326·00 ± 14·00
28 Kababish (5)	23	1709·0 ± 7·9	56·0 ± 5·6	327·67 ± 32·58
29 Cutch (1)	24	1633·0 ± 7·4	54·1 ± 5·3	331·31 ± 32·25
30 Nejmps (1)	11	1723·1 ± 11·7	57·4 ± 8·3	333·13 ± 47·96
31 *Khotan* (3)	67	1655·2 ± 4·6	55·5 ± 3·2	335·30 ± 19·53
32 *Punjabi* (1)	60	1683·8 ± 5·0	57·2 ± 3·5	339·41 ± 20·89
33 Bantu Kavirondo (1)	24	1692·6 ± 7·9	57·4 ± 5·6	339·13 ± 33·01
34 *Minia* (4)	491	1669·70 ± 1·7	56·6 ± 1·2	339·00 ± 7·06
35 *Sundanese* (2)	37	1591·30 ± 6·00	54·07 ± 4·24	340· ± 27·
36 *Kamba* (1)	128	1656·6 ± 3·4	56·6 ± 2·4	341·92 ± 14·41
37 *Turfan* (3)	72	1662·6 ± 4·5	57·0 ± 3·20	342·83 ± 19·27
38 *Beheira* (4)	525	1676·8 ± 1·7	57·4 ± 1·2	342·00 ± 07·00

[1] Goring: *Ibid.*, p. 370.

Name of Race.	Col. 2 No. in Sample.	Mean (mm.) ± P.E. of Mean.	S.D. in mm. ± P.E. of S.D.	100 × (Coeff. of Var. ± P.E. of V.)
39 Biloch (1)	15	1649·7 ± 9·9	56·6 ± 7·0	343·34 ± 42·27
40 *Duruma* (1)	67	1649·2 ± 4·8	57·7 ± 3·4	349·60 ± 20·37
41 *Arab and Swahilil* (32)	32	1644·6 ± 6·9	57·7 ± 4·9	350·57 ± 29·55
42 *Giza* (4)	326	1678·0 ± 2·2	58·8 ± 1·6	350· ± 9·
43 Chitrali (3)	22	1684·5 ± 8·1	59·3 ± 5·8	352·03 ± 35·79
44 *Qena* (4)	824	1678·0 ± 1·4	59·0 ± 1·0	352· ± 6·
45 *Beni Amer* (5)	51	1643· ± 5·	58· ± 4·	353·01 ± 23·57
46 *Girga* (4)	610	1677·7 ± 1·6	59·2 ± 1.1	353 ± 7·
47 *Fayum* (4)	413	1672·0 ± 2·0	59·2 ± 1·4	354 ± 8·
48 *Polu* (3)	31	1644 2 ± 7·9	58·3 ± 4·9	354·57 ± 30·37
49 *Beni Suef* (4)	384	1662·3 ± 2·0	59·1 ± 1·4	355· ± 9·
50 *Gharbia* (4)	1105	1673·3 ± 1·2	59·4 ± 0·9	355·00 ± 5·
51 *Masai* (1)	91	1700·0 ± 4·3	60·7 ± 3·0	357·08 ± 17·85
52 *Hadendoa and Amara* (5)	54	1676 ± 5·	60· ± 4	357·99 ± 23·23
53 Aksu (3)	13	1637·7 ± 10·6	58·5 ± 7·5	357·20 ± 47·25
54 Sheher (1)	82	1615·7 ± 4·4	57·9 ± 3·1	358·43 ± 18·87
55 *Alexandria* (4)	643	1666·2 ± 1·6	59·7 ± 1·1	359·0 ± 7·
56 *Kökyar* (3)	37	1629·2 ± 6·3	58·9 ± 4·4	361·52 ± 28·34
57 Giriama (1)	24	1629·7 ± 8·1	58·9 ± 5·7	361·59 ± 35·20
58 *Daqahlia* (4)	504	1660·6 ± 1·8	60·0 ± 1·3	361· ± 08
59 *Assiut* (4)	889	1668·9 ± 1·4	160·3 ± 1·0	362 ± 06
60 *Cairo* (14)	802	1682·9 ± 1·4	59·3 ± 1·0	364 ± 06·
61 Wakhi (3)	19	1680· ± 8·8	61·8 ± 6 2	367·84 ± 40·25
62 *Camb. Students* (11).	1000	1748·88 ± 1·4	64·6 ± 0·97	369·58 ± 05·58
63 Ajawa (1)	16	1652·2 ± 10·3	61·2 ± 7·3	370·48 ± 64·17
64 *Aswan North* (4)	115	1683·3 ± 3·9	62·3 ± 2·8	370·00 ± 16·00
65 *Menufia* (4)	718	1677·0 ± 1·6	62·5 ± 1·1	371· ± 7·
66 *Embu* (1)	110	1630·1 ± 3·9	61·2 ± 2 9	375·50 ± 17·07
67 Kafir (3)	18	1667·8 ± 9·0	63·3 ± 6·4	379·54 ± 42·66
68 *Manyema* (1)	42	1667·5 ± 6·6	63·2 ± 4·7	379·28 ± 27·91
69 *Kikuyu* (1)	384	1640· ± 2·2	62·5 ± 1·5	380·98 ± 9·27
70 *Qualiubia* (4)	295	1662·4 ± 2·5	63·1 ± 1·8	280· ± 10·
71 *Sharqia* (4)	516	1655·4 ± 1·9	63·3 ± 1·3	382· ± 8
72 *U.S.A. Recruits* (6)	25,898	1709·4 ± 0·27	65·6 ± 0·19	383·76 ± 1·15
73 *Nuer* (5)	39	1806· ± 8·0	70· ± 5	387·59 ± 29·60
74 *N.S.W. Criminals* (8)	2871	1698·8 ± 0·83	65·8 ± 0 58	387·33 ± 03·45
75 Nyasa (1)	21	1640·0 ± 9·4	63 7 ± 7·3	390·27 ± 40·61
76 Keriya (3)	21	1612·5 ± 9·3	62·9 ± 6·5	390·07 ± 40·59
77 Sukuma (1)	21	1717·0 ± 9·9	67·3 ± 7 0	392·01 ± 42·80
78 *Kirghiz* (3)	38	1640·8 ± 6 2	64·6 ± 4·4	393·71 ± 30·46
79 Somali (1)	27	1735·1 ± 7·6	68·6 ± 5·4	395·25 ± 36·27
80 Suk (1)	15	1677·9 ± 11·6	66·3 ± 8·2	395·2 ± 48·66
81 *Eng. Sons* (9)	1078	1744 0 ± 1·42	69·4 ± 1·0	395 ± 6
82 *Eng. Fathers* (9)	1078	1719·5 ± 1·39	68·7 ± 1·0	399 ± 6
83 *Germans* (10)	390	1659·3 ± 2·3	66·8 ± 1·6	402·37 ± 10·38
84 *Eng. Criminals* (11)	3000	1658·1 ± 1 6	68·07 ± 1·2	411 ± 9
85 Nilotik Kavirondo (1)	37	1729·0 ± 7·9	71 4 ± 5·6	412·81 ± 32·36
86 *Loplik* (3)	38	1695·0 ± 6·2	70·3 ± 4·4	414·74 ± 32·08
87 *Barabra* (5)	70	1680 ± 7·0	70 ± 3·7	416·66 ± 23·75
88 *Kachamega* (1)	100	1668·3 ± 4·7	69·8 ± 3·3	418·69 ± 19·96
89 Kamasia (1)	20	1719·8 ± 10·9	72·4 ± 7·7	420·91 ± 44·89
90 *Aswan South* (4)	95	1650·6 ± 4·8	69·4 ± 3·4	421 ± 21
91 Mastuji (3)	28	1666·1 ± 7·2	70·4 ± 5·1	422·54 ± 38·08
92 Korla (3)	14	1667·9 ± 10·2	70·6 ± 7·2	423·28 ± 53·95
93 Scot. Insane (7)	4381	1673·8 ± 0·73	72·1 ± 0·52	430·95 ± 3·10
94 Scot. Total (7)	4401	1668·8 ± 0·75	73·7 ± 0·53	441·40 ± 3·17
95 Bagh-jigda (3)	12	1647·5 ± 11·0	73·2 ± 7·8	446·30 ± 61·17
96 Charklik (3)	12	1678·3 ± 11·0	74·6 ± 7 8	446·28 ± 61·44
97 Chaga (1)	18	1641·6 ± 12·2	76·96 ± 8·7	468·82 ± 52·70
98 Rabai (1)	13	1626·1 ± 14·5	77·4 ± 0·2	476·41 ± 63·01
99 Turkana (1)	9	1694·4 ± 19·4	86·1 ± 13·7	508·16 ± 80·78
100 Dinka (5)	116	1786· ± 6·0	97·0 ± 4 4	543·11 ± 24·04

Supplementary List.

In this List actual Coefficients of Variability are given.

Name of Race.	Col 2 No. in Sample.	Mean (mm.) ± P.E. of Mean.	S.D. in mm. ± P.E. of S.D.	100 × (Coeff. of Var. ± P.E of V.)
101 *Crete, whole Island* (12)	318	1706·1 ±2·6	67·5 ±1·8	3·96 ± ·12
102 *Eparchies* (*Selinos, Sphakia*) (12)	50	1752·6 ±5·4	57·1 ±3·9	3·26 ± ·22
103 *Albanian* (12) ..	140	1693·2 ±3·7	65·7 ±2·6	3·88 ± ·18
104 *Cyprus* (whole Island) (12)	585	1687·7 ±1·7	61·6 ±1·2	3·64 ± ·07
105 *Cyprus* (*Nicosia*) (12)		1678·8 ±3·9	60·5 ±2·7	3·60 ± ·16
106 ,, (*Lapitho*) (12).	221	1680·0 ±2·5	54·7 ±1·8	3·25 ± ·10
107 ,, (*Ekomi*) (12)	167	1690·5 ±3·2	60·8 ±2·2	3·59 ± ·13
108 ,, (*Levkonika*) (12)	87	1689·8 ±4·6	63·7 ±3·3	3·77 ± ·19
109 *Cyprus* (*Leukas*) (12)	42	1668·0 ±6·7	64·3 ±4·7	3·86 ± ·33
110 *Lycian Gypsies* (12).	53	1660·2 ±4·4	47·8 ±3·1	2·88 ± ·20
111 *Persian Jews* (12)	57	1643·5 ±5·2		3·53 ± ·22
112 *Yemen Jews* (12) .	78	1594·0 ±2·9	..	3·76 ± ·20
113 *Samarkand Jews* (12)	100	1664·2 ±3·9		3·52 ± ·17
114 *Oxford students* (13)	959	1765	66·08 ±	3·7439
115 *Aberdeen students* (11) ..	493	1717·0 ±1·8	59·4 ±1·3	3·4595

(1) Leys and Joyce, *Jour. Roy. Anthrop. Inst.*, Vol. XLIII (1913) p. 216.
(2) Garett, *Jour. Roy, Anthrop. Inst.*, Vol. XLII (1912), pp. 60–66.
(3) Joyce, *Jour. Roy. Anthrop. Inst.*. Vol. XLII (1912), p. 473.
(4) Craig, *Biometrika*, Vol. 8 (1911). p. 75.
(5) Seligmann, *Jour. Roy. Anthrop. Inst.* Vol. XLIII (1913), pp. 700–702.
(6) Pearson, *Phil. Trans. Roy. Soc.* Vol. 184A, p. 386.
(7) Tocher, *Biometrika*, Vol. 5 (1906–7) p. 307.
(8) Powys, *Biometrika*, Vol. 1 (1901), p. 44.
(9) Pearson, *Biometrika*, Vol. 2, (1903), p. 370.
(10) Pearson, *Chances of Death*, Vol. 1, pp. 294–296.
(11) Macdonell, *Biometrika*, Vol. 1 (1901) pp. 191.
(12) Buxton, *Biometrika*, Vol. 13 (1920), p. 104 and p. 108.
(13) Schuster, *Biometrika* Vol. 8 (1911), p. 49.
(14) Orensteen, *Biometrika*, Vol. 11 (1915), pp. 67–81

TABLE OF VARIABILITIES.

There are several remarkable points about the Table of Variabilities. The material is supposed to be homogeneous in each case, yet we note the extreme range of variation of the coeff. of variability. We have 1·76 42 ± ·24 28 and 5·08 16 ± ·80 78 as our extreme values.

The mean variability is very near 3·6, and one very remarkable fact is this, that—

I. *The more highly civilised races have greater variabilities than the average.*

This confirms Pearson's result for Cephalic Index.[1] Pearson concludes for Cephalic Index that greater variability is a characteristic of the "races which have been successful in the struggle for existence, and at the present time are the dominant races of the

[1] Chances of Death, Vol. 1, p. 292.

earth. At the same time the greater variability of the more dominant and civilised peoples admit of being interpreted as a result of the lesser severity of the struggle for existence among them. Thus greater variability would be an effect not a cause of the higher state of civilisation."

Another fact which may be gathered from the above table is this. The more civilised races though more variable, do not in any case occupy the extreme ends of the table. Thus one would probably be justified in inferring that a higher state of civilisation is not associated with extreme degrees of variability.

We may look at the same question from a different point of view. The less civilised races occupy the extreme ends of the table more frequently than the more civilised races. The less civilised races though on the whole less variable, may thus be associated with extreme degees of variabilities.

II. *The greater variability of more highly civilised races seems to be only moderate in degree and is never excessive.*[1]

It seems as if *slightly greater* variability than the stable type of the species is accompanied by greater adaptability and hence with a higher state of progress.

INTERRACIAL VARIABILITY.

There is another point which deserves attention. By looking at our general list of variabilities, we find some association between average stature (M) and standard deviation σ.

The point which we are considering now is *interracial* correlation between M an σ for the different races.[2]

If $$\mu_{11} = S(xy)/N,$$

then the correlation coefficient as determined by the product moment method,[3] is given by

$$r = \mu_{11}/(\sigma_x \ \sigma_y)$$

where σ_x and σ_y are S.D. of the two variables.

I find, without grouping, with base numbers 1660 mm. and 60 mm. respectively for average stature and S.D. the raw moments to be:—

For Stature $\nu_1' = 5{\cdot}24$ $\nu_2' = 1389{\cdot}48$

---

[1] In the selected list (see below) this fact is not so apparent. It seems as if the extremely high variability of less civilised races is due to unreliability of data.

[2] This is quite distinct from the *intra-racial* (or within the race) correlation between errors in Mean and errors in S.D.

In *Biom.* Vol. 2 (1903), Problem IX, p. 279, is shown that

$$R_{M,\sigma} = \mu_3/(\sigma \ \sigma_{\mu_2} \ N)$$

In our case, μ_3 is negative, hence a taller subsample of Anglo-Indians will show less variability and vice versa. This is actually the case with the two subsamples we have already considered. The subsample with a higher average 1658·75 mm. has a S.D. of 68·85 mm. as against the other with Mean = 1657·00 mm. and S.D. = 73·26. μ_3 being small, correlation however, is very small.

[3] See Yule: "Theory of Statistics," p. 171.

Standard Deviation (base No. 60)

$$\nu_1' = \cdot 77 \qquad \nu_2' = 92 \cdot 29$$

and for $\nu_{11}' = 110 \cdot 04$.

Transferring to Mean, we get:—

For Stature:—

Mean value of Average stature = 1665·24 mm.
Standard Deviation = 36·90 52 mm.
Coeff. of *interracial* Variability = 2·21 62

For Standard Deviation:—

Mean value of Standard Deviation = 59·23 mm.
S.D. of Standard Deviation = 9·57 58 mm.
Coeff. of *interracial* V of S.D. = 16·17

We also have $\mu_{11} = +114 \cdot 07\ 38$

Thus $$R_{M,\sigma} = \frac{+114 \cdot 07\ 38}{36 \cdot 90\ 52 \times 9 \cdot 57\ 58}$$

$$= +\cdot 30\ 98.$$

The Prob. Error [1] of R is given by $\frac{\cdot 67449}{\sqrt{n}}(1 - R^2)$

From Abac in Biometric Tables p. 19, we find for $N = 100$, P.E. of $R = \cdot 062$.

Thus $R_{M,\sigma} = \cdot 3098 \pm \cdot 062$

We may now consider the correlation for our *selected list of 55 reliable samples.*

Stature:—

Mean value of Average Stature = 16. 63·94 54 mm.
Standard Deviation of Average Stature = 36·59 53 mm.
Coeff. of Variability (*interracial*) = 2·19 93

Standard Deviation—

Mean value of Standard Deviation = 59·34 53 mm.
S.D. of Standard Deviation = 6·46 mm.
Coeff. of Variation (*interracial*) = 10·89

$$\mu_{11} = +125 \cdot 889$$

Thus $R_{M,\sigma} = +\cdot 3283 \pm \cdot 082$

Selection of more reliable values does not make any substantial difference. We may therefore conclude that there is a positive interracial correlation of about +·3 between Average Stature and Standard Deviation.

---

[1] K. Pearson and L. N. G. Filon: "Probable Errors of Frequency Constants" etc., *Phil. Trans.* Vol. 191A, pp. 231—241.

Interracially, taller races are on the whole more variable than shorter.

It will be noticed that the average stature of all the races is 1665·24 and in the case of selected races, 1663·94 mm.

The Anglo-Indians are thus slightly shorter than the general average of all the races. But the difference is only about 7 mm.

In this connection it is interesting to compare the figures given by Tschepourkowsky.[1] He finds for 92 Russian races the mean value of average stature to be 1647·4 mm. and S.D. 33·3 mm. while for Deniker's 84 living races, the values are 1639·6 and 55·9 respectively.

His coefficient of interracial variation of stature is 2·02. In our series of 100 races it is slightly higher, being about 2·22 but is of the same order.

Thus our value of interracial variability agrees generally with a previous value found independently by another worker.[2]

We can now pass on to the question of interracial correlation between M and V

If v_1, v_2, v_3, v_4 are the variabilities of x_1, x_2, x_3, x_4 and r_{12}, r_{13} are the correlation between x_1 and x_2, x_2 and x_3, etc., then the correlation between x_1/x_3 and x_2/x_4 has been shown[3] to be

$$\rho = \frac{r_{12}v_1v_2 - r_{14}v_1v_4 - r_{23}v_2v_3 + r_{34}v_3v_4}{\sqrt{(v_1^2 + v_2^2 - 2r_{13}v_1v_2)(v_2^2 + v_4^2 - 2r_{24}v_2v_4)}}$$

We get correlation between M and $V = 100\sigma/M$ by putting

$$x_1 = x_4 = M,\ x_3 = 1 \text{ and } x_2 = \sigma$$

Then

$$v_1 = v_4, \quad v_3 = 0, \quad r_{13} = r_{23} = r_{34} = 0,$$
$$r_{14} = r_{41} = 1, \quad r_{12} = r_{42}.$$

Thus

$$\rho_{M,V} = \frac{r_{12}v_2 - v_1}{\sqrt{v_1^2 + v_2^2 - 2r_{12}v_1v_2}}.$$

For the Whole Series of 100 races:—

$$v_1 = 2{\cdot}21\ 62$$
$$v_2 = 16{\cdot}17$$
$$r_{12} = +{\cdot}30\ 98$$

Hence $\rho_{M,V} = +{\cdot}17\ 87 \pm {\cdot}065$

For the Selected Series of 55 races:—

$$v_1 = 2{\cdot}19\ 93$$
$$v_2 = 10{\cdot}89$$

---

[1] E. Tschepourkowsky: "Contributions to the Study of Interracial Correlation." *Biom.* Vol. IV (1905), pp. 286—312.

[2] We may note however that the interracial variability is higher in our case. This implies that our sample of races is more representative in character than Tschepourkowsky's.

[3] K. Pearson: "On Spurious Correlation," *Proc. Roy Soc.* Vol. LX.

$$r_{12} = +\cdot 32\ 83$$

$$\rho_{M,\,V} = +\cdot 1303 \pm \cdot 088$$

The correlation in the latter case is scarcely significant, but seems to be slightly positive.

Thus *there seems to be a small positive interracial correlation between the average stature and the coefficient of variation.*

Assuming recent races to be more variable, the positive interracial correlation between stature and variability may be explained on the hypothesis that tallness is a recent acquirement of the human species. The greater variability is not merely due to the greater absolute size of the taller races, since the coefficient of variability i.e. proportional variability itself is also positively correlated with stature.

SECTION VII. COMPARISON OF VARIABILITIES.

Standard Deviation of Stature.

(*a*) *The Whole Series.*

Let us consider the 100 different values of Standard Deviation of Stature, which I have collected for purposes of comparison. We notice the great range of variation of the S.D. Our extreme values are 29·5 mm. and 97·0 mm.

Grouping by units of 5 mm. we get the following distribution:—

Distribution of 100 *S.D. of Stature.*

Group	29 to 34	34 to 39	39 to 44	44 to 49	49 to 54	54 to 59	59 to 64	64 to 69	69 to 74	74 to 79	79 to 84	84 to 89
Frequency	2	0	4	6	13·5	22	23·5	10	15	3	0	1

We get

Mean Value of Standard Deviation = 59·45 mm.
S.D. of Standard Deviation = 9·52 37 mm.
P.E. of Mean Standard Deviation = 6·42 12

We can now compare our Anglo-Indian S.D. with this Mean Value:—

Anglo-Indian S.D. = 67·38 mm.
Mean value of S.D. = 59·45 mm.
Difference 7·93 ± 6·42 mm.

The difference 7·93 ± 6·42 mm. is not at all significant. We can find the probability of this difference,

$$x = {}^{D}/_{\sigma} = \frac{7{\cdot}935}{9{\cdot}524} = 0{\cdot}83 \text{ approximately}$$

From Tables II, p. 2 $\frac{1}{2}(1+a)$ = ·79 67 30 6
$\frac{1}{2}(1-a)$ = ·20 32 69 4.

If we assume that our sample of 100 standard deviations is a random or representative samples then 20·3% of all "homogeneous" races will have a S.D. greater than the Anglo-Indians, and 40·6% will differ more from the average value than Anglo-Indians.

For Stature, the absolute variability (Standard Deviation) of Anglo-Indians is thus not significantly greater than the average absolute variability of homogeneous races.

It will be noticed that the list contains many small samples. It will be better to omit all samples of less than 25. Doing this we find that extreme values have been mostly eliminated by this process of selection, showing that such extreme values were probably in most cases due to uncertainty of sampling rather than to any peculiarity of the population.

I have also thought it best to exclude Scottish Insane as well as the Dinka group. We have already seen that Anglo-Indian variability is not significantly greater than the average variability of the *whole series*. The inclusion of any variability greater than Anglo-Indian variability will strengthen this conclusion, rejection of greater variabilities will go against our conclusion. The Insane is manifestly abnormal and may be neglected for the present. Variability of the Dinka people is greater than that of Anglo-Indians, its rejection will thus make the test more rigid. Separate figures for Aswan is also omitted for similar reasons.[1]

For the *selected series of Standard Deviations*

Selected Mean Stand. Dev. = 59·8929 mm.
S.D. of Standard Dev. = 6·3504 mm.

We notice that the selected Mean is almost exactly the same as the Mean for the whole series. We conclude that 60 *mm. is about the true average absolute Variability of stature for human races.*

Due to selection the S.D. of Variability is considerably reduced because the extreme values of Variability have in most cases been eliminated.

Anglo-Indian S.D. = 67·385 mm.
Selected Mean S.D. = 59·893 mm.
Anglo Indian Difference = 7·492 mm.

We find the probability :—

$$x = D/\sigma = \frac{7{\cdot}492}{6{\cdot}350} = 1{\cdot}18$$

From Biometric Tables II, $\frac{1}{2}(1+\alpha) = {\cdot}88\ 09\ 99\ 9$
$\frac{1}{2}(1-\alpha) = {\cdot}11\ 90\ 00\ 1$

Thus 11·9% of all races will have greater variabilities than Anglo-Indians while 23% will differ more from the Selected Mean.

As judged by a reliable series of standard deviations, the Absolute Variability of Anglo-Indians is not significantly greater than the Average Variability of different "homogeneous" samples.

RELATIVE VARIABILITY OF STATURE.

We shall now compare the *Relative Variability* (as measured by the Coefficient of Variation) of our Anglo-Indian data with the variability of samples recognised to be homogeneous.

[1] J. I. Craig : *Biometrika* Vol. 8 (1911), p. 70.

(*a*) *Whole Series.*

Distribution of 100 *Coefficients of Variation of Stature.*

Group	1·80 to 2·20	2·20 to 2·60	2·60 to 3·00	3·00 to 3·40	3·40 to 3·80	3·80 to 4·20	4·20 to 4·60	Beyond 4·60	Total.
Frequency	2	3	9	20·5	34·0	20·5	8·0	3	100

Grouping in units of 4, we find moment-coefficients about the Mean,

$$\mu_2 = 1{\cdot}88\ 66$$
$$\mu_3 = -{\cdot}77\ 80$$
$$\mu_4 = 11{\cdot}61\ 74\ 74$$

giving $\beta_1 = {\cdot}09\ 46\ 73$

$$\beta_2 = 3{\cdot}37\ 26\ 20 \pm {\cdot}63\ 30$$

with sk. = ·13 44 8

Mean Coefficient of Variation = 3·5700

and S.D. of Coefficient of Variation = ·5450

Curve belongs to Type IV, but the Gaussian itself will be quite adequate.

"Goodness of Fit" of Coefficients of Variation.

Coeff. of V.	Observed m'.	Theoretical m.	$m-m'$	$\frac{(m-m')^2}{m}$.
Beyond 2·20	2	·742	1·258	2·1320
2·20—2·60	3	3·538	·538	·0818
2·60—3·00	9	11·512	2·512	·5481
3·00—3·40	20·5	22·912	2·412	·2531
3·40—3·80	34·0	27·934	6·066	1·3170
3·80—4·20	20·5	20·769	·239	·0028
4·20—4·60	8·0	9·459	1·459	·2250
Beyond 4·60	3·0	3·130	·130	·0054

$\chi^2 =$ 4·5660

$n' = 8$ $\chi^2 = 4{\cdot}566$

$P =$ ·71 21 63

Thus the Gaussian gives excellent fit. In seven cases out of ten, the fit will be worse.

We notice that one terminal frequency gives rather a large value i.e. 2·1320, combining the two end groups, we get,

$\chi^2 = 2{\cdot}555$

$n' = 8$, $P =$ ·85 45 87

The fit is now considerably improved. *I conclude that the Coefficient of Variation (for homogenous groups) can itself be gra-*

duated by the Gaussian curve. We can now safely apply the theory of Errors (which is based on the Gauss-Laplacian Probability Integral) to judge the likelihood of deviations from the Mean.

Anglo-Indians $V =$ 4·0672
Average $V =$ 3·5700
Anglo-Indian Difference = ·4972
Now the S.D of $V =$ ·5450
Thus, P.E. of $V = \pm$·3676
Anglo-Indian Difference = ·4972 ± ·3676

$$x = \frac{D}{\sigma} = \frac{\cdot 4972}{\cdot 5450} = \cdot 91$$

From Biometric Table II, $\frac{1}{2}(1+\alpha) =$ ·81 85 88
$\frac{1}{2}(1-\alpha) =$ ·18 14 12

Thus we find that no less than 18·14% of "homogeneous" races will have *larger* Coefficients of Variation than Anglo-Indians. *The Anglo-Indian Coefficient of Variation is not significantly greater than the average Coefficient of Variation of the whole series.*

(*b*) *Selected Series.*

We obtain the following distribution of the Coefficients of Variation for 55 *selected*[1] races (unit of grouping = ·2).

Distribution of 55 selected Coefficients of Variation.

Group	2·7 to 2·9	2·9 to 3·1	3·1 to 3·3	3·3 to 3·5	3·5 to 3·7	3·7 to 3·9	3·9 to 4·1	4·1 to 4·3	Total.
Frequency	3	5·5	1·5	9·0	17·5	9·5	4·0	5·0	55

We get,
Mean Coefficient of Variation = 3·571
Standard Deviation of Coefficient of Variation = ·3590
P.E. of Mean V = ·2421

The other constants are:—

$\mu_2 =$ 3·22 26 45
$\mu_3 =$ 1·59 62 01
$\mu_4 =$ 29·89 12 68
$\beta_1 =$ ·06 61 53
$\beta_2 =$ 2·97 93

[1] It will be noticed that the *extreme* values have been automatically excluded by our principle of *rejection of unreliable* values.

The stability of the Mean Value is remarkable. For the whole series it was 3·57, for the selected races it is 3·571. It therefore seems likely that 3·57 *is very near the true typical coefficient of variation (of stature) for homogeneous non-Indian* samples.

The S.D. is much reduced by selection. This is now ·3590 as against ·5450 for the whole series. We have selected the more reliable values, but this has also excluded almost all extreme values. Great divergence from the Mean value is thus probably due more to paucity of material than to actual peculiarities of distribution.

$$\begin{aligned} \text{Anglo-Indian } V &= 4{\cdot}0672 \\ \text{Selected Mean } V &= 3{\cdot}5710 \\ \hline \text{Anglo-Indian Difference} &= {\cdot}4962 \pm {\cdot}2421 \end{aligned}$$

The actual difference is again the same, but this is now nearly twice the Probable Error.

We have,

$$x = \frac{D}{\sigma} = \frac{49{\cdot}62}{35{\cdot}90} = 1{\cdot}38 \text{ approximately}$$

From Biometric Table II, $\frac{1}{2}(1+\alpha) = {\cdot}91\ 62\ 04\ 7$

$\frac{1}{2}(1-\alpha) = {\cdot}08\ 37\ 95\ 3$

Thus 8·38% of all reliable samples will actually be more variable than Anglo-Indians, while 16·55% will differ more from the Mean.

Anglo-Indian Variability of stature is not significantly higher than the average Variability of selected samples.

(c) *Selected and Weighted Series.*

Still another course is open to us. We can consider the "weighted Mean" [1] and "weighted" Standard Deviation of the Coefficient of Variation. For this purpose, we choose our weights to be proportional to $1/E^2$, where E is the probable error, i. e. give "weights" proportional to reliability.

We get,

$$\begin{aligned} \text{Weighted Mean } V &= 3{\cdot}7622 \\ \text{Weighted S.D. of Mean } V &= {\cdot}1846 \end{aligned}$$

We notice that the Mean is now considerably higher. This is due to the much greater reliability in the measurements of the more civilised races, who have invariably higher variabilities. This greater value is also due in a large measure to the weight of the U.S.A. recruits ($w = 7623$ against 10 for the lowest weight) which includes 25,898 individuals.

---

[1] See Yule: "Theory of Statistics" (Charles Griffin, 1919), p. 220.

Anglo-Indian $V =$ 4·0672
(Selected Weighted Mean $V =$ 3·7622
Anglo-Indian Difference = ·3000
P.E. of Difference = ± ·1245

$$x = \frac{D}{\sigma} = 1{\cdot}63 \text{ approximately}$$

From Biometric Table II, $\frac{1}{2}(1+a) =$ ·94 84 49 3
$\frac{1}{2}(1-a) =$ ·05 15 50 7

5·1% will be more variable, while 10·2% will differ more from the weighted average than Anglo-Indians.

Thus even when compared with the weighted Mean, Anglo-Indian Variability is not significantly greater than average Variability.

We have seen that U.S.A. recruits raise the weighted Mean very considerably. But it is not at all certain that the recruits of the U.S.A. Army are possessed of any great degree of homogeneity. One would surmise rather that they are heterogeneous in character Let us see the effect of leaving out U.S.A. recruits.

Omitting U.S.A. recruits we get

Weighted Mean $V =$ 3·6413
Weighted S.D. of $V =$ ·2509
Weighted P.E. of $V =$ ± ·1683
Anglo-Indian $V =$ 4·0672
Weighted Mean $V =$ 3·6413
Difference = ·4259 ± 16·83

$$x = \frac{D}{\sigma} = \frac{\cdot 4259}{\cdot 2509} = 1{\cdot}70$$

From Biometric Table II, $\frac{1}{2}(1+a) =$ ·95 54 34 5
$\frac{1}{2}(1-a) =$ ·04 45 65 5

4·5% will have greater Variabilities than the Anglo-Indian sample. As regards the Coefficient of Variation, this is the most stringent test we can apply with the non-Indian material at our disposal. We find

Anglo-Indian Variability is within the limits of probability of homogeneous Variation. Study of the Coefficient of Variation for Stature does not enable us to assert definitely that the present Anglo-Indian sample is heterogeneous in character.

I shall now consider the whole series of non-Caste samples including the Mediterranian samples. I have omitted the separate age-groups for the New South Wales Criminals and the Oxford student data. As all these have greater Variability than the Average, the stringency of our test will not be diminished by this rejection.[1] Another reason why I have omitted the different age-

[1] See discussion on p. 72.

groups is this. My purpose is to compare the Anglo-Indian Variability with the general average Variability of other races. If the Coefficient of Variation for the same race is given several times over under different age-groups, too much weight will obviously be given to this particular race. I also omit Dinka.[1]

Distribution of Coefficients of Variation of 107 non-Caste Samples.

Group	Beyond 1·80	1·80 to 1·90	1·90 to 2·00	2·00 to 2·10	2·10 to 2·20	2·20 to 2·30	2·30 to 2·40	2·40 to 2·50	2·50 to 2·60	2·60 to 2·70
Frequency	1	0	0	1	0	0	1	0	2	1

Group	−2·80	−2·90	−3·00	−3·10	−3·20	−3·30	−3·40	−3·50	−3·60	−3·70
Frequency	4	3	3	8	4	3	5	6	18	9

Group	−3·80	−3·90	−4·00	−4·10	−4·20	−4·30	−4·40	−4·60	−4·80	−5·00
Frequency	7	8	9	1	5	4	1	1	1	1

Grouping in units of ·1, I find

$$\mu_2 = 28{\cdot}99\ 60\ 64,$$
$$\mu_3 = -50{\cdot}77\ 00\ 11,$$
$$\mu_4 = 3252{\cdot}08\ 73\ 71$$

Thus $\beta_1 = {\cdot}10573$ and

$\beta_2 = 3{\cdot}868$

Curve is of Type IV, but to a first approximation we can apply the "normal" curve of errors.

Mean Coefficient of Variation (107 samples) $= 3\ 5353 \pm {\cdot}0348$

Standard Deviation of Coefficient of Variation $= {\cdot}5385 \pm {\cdot}0245$

The Mean Value is slightly lower than the one found earlier. This is due to the fact that I have omitted the Dinka group here. If we include the Dinka group, the Mean Value would be raised to 3·553 which compares favourably with the value 3·570, a difference of ·035 only.

Anglo-Indian Coefficient of Variation $= 4{\cdot}0672$

Mean Coefficient of Variation $= 3{\cdot}5353$

Anglo-Indian Difference $= {\cdot}5319.$

$$x = \frac{D}{\sigma} = \frac{{\cdot}5319}{{\cdot}5385} = {\cdot}988$$

From Biometric Table II, $\frac{1}{2}(1+a) = 83\ 84\ 217$

$\frac{1}{2}(1-a) = {\cdot}16\ 15\ 783$

[1] See discussion on p. 62.

Thus as before the Anglo-Indian sample does not seem to be significantly more variable than homogeneous samples. About 16% of homogeneous samples will have a greater Variability.

(*b*) *Selected Series.*

Let us now select samples greater than 25. We get a total (omitting different age-groups) of 67 samples distributed as follows:—

Distributions of 67 Selected Coefficients of Variations.

Group	Beyond 2·70	−2·80	−2·90	−3·00	−3·10	−3·20	−3·30	−3·40	−3·50	−3·60
Frequency	2	2	1	5	1	2	2	5	15	7

Group	−3·70	−3·80	−3 90	−4·00	−4·10	−4·20	−4·30	Total.
Frequency .	6	8	4	1	4	1	1	67

We get, $\mu_2 = 12{\cdot}97\ 78\ 82$

$\mu_3 = 17{\cdot}77\ 40\ 69$

$\mu_4 = 477{\cdot}70\ 38\ 55$

giving $\beta_1 = {\cdot}14\ 43\ 94 \pm {\cdot}12\ 47$

$\beta_2 = 2{\cdot}83\ 53\ 67 \pm {\cdot}38\ 28$

Graduation by the "normal" curve is thus possible and we are justified in using the "normal" Probability Integral.

Mean Value of Coefficient of Variation $= 3{\cdot}5843 \pm {\cdot}0297$

Standard Deviation of Co-efficient of Variation $= {\cdot}3602 \pm {\cdot}0210$

It will be noticed that the Mean Value 3·584 is sensibly the same as we had obtained without including this Mediterranean data e.c. 3·571. The difference is only ·013 while the probable error is certainly greater than ·03. Thus 3·58 may be safely taken as a standard value for the Coefficient of Variation for Stature of homogeneous non-Caste samples.

The mean value for the whole series 3·5353 is smaller than the mean value for selected samples, 3·5843, because in small samples the dispersion is more likely to be smaller.[1]

Let us now compare the Anglo-Indian Variability with the above Mean Variability.

Anglo-Indian Coeff. of Variation $= 4{\cdot}06\ 72$

Mean Selected Coeff. of Variation $= 3{\cdot}58\ 43$

Anglo-Indian Difference $= 0{\cdot}48\ 29$

[1] For a discussion of the dependence of Standard Deviation on the size of sample see *Biometrika* Vol. 10(1915) p. 522 and Vol. 11 (1916) p. 277.

Hence $$x = \frac{D}{\sigma} = \frac{0{\cdot}4829}{0{\cdot}3602} = 1{\cdot}34$$

From Table II, $\frac{1}{2}(1+\alpha) = {\cdot}90\ 98\ 773$

$\frac{1}{2}(1-\alpha) = {\cdot}09\ 01\ 227$

Thus nearly 9% of homogeneous samples will have a greater Variability. The inclusion of the new Mediterranean series does not affect our previous conclusion.

The Variability of the Anglo-Indian sample though higher than the Average is not excessively so and the difference is not statistically significant.

INDIAN CASTE VARIABILITY.

(*a*) *Whole Series.*

I shall now consider the Coefficient of Variation of the Indian Caste data of Risley. Omitting 3 tribes in which the sample consists of only 2 individuals I get a total of 84 Castes and Tribes.

Distribution of 84 Caste Coefficients of Variation.

Group	-2·1	-2·2	-2·3	-2·4	-2·5	-2·6	-2·7	-2·8	-2·9	-3·0
Frequency	2	1	2	0	1	0	1	3	8	7
Group	-3·1	-3·2	-3·3	-3·4	-3·5	-3·6	-3·7	-3·8	-3·9	-4·0
Frequency	5	12	13	7	7	6	3	0	5	1

Grouping by ·1, I get

Mean Value of Caste Coefficient of Variation = 3·2395

Standard Deviation of Coefficient of Variation = ·3943

Anglo-Indian Coefficient of Variation = 4·0672

Anglo-Indian Difference = ·8277

$$x = \frac{D}{\sigma} = \frac{{\cdot}8277}{{\cdot}3943} = 2{\cdot}099,$$

From Biometric Table II, $\frac{1}{2}(1+\alpha) = {\cdot}98\ 21\ 356$

$\frac{1}{2}(1-\alpha) = {\cdot}01\ 78\ 644$

Only about two per cent of Indian Caste samples will show greater variability. It seems therefore likely that the Anglo-Indian sample is really differentiated from the Indian Castes in showing a just significant degree of greater variability.

It should be noted that the Caste Variability is much lower than the non-Caste Variability.

We have

Non-Caste Variability = $3{\cdot}5700 \pm {\cdot}0368$

Caste Variability $3{\cdot}2395 \pm {\cdot}0290$

Caste Difference = $3305 \pm {\cdot}0422$

The difference is nearly eight times the probable error of the difference. Hence we conclude that Caste Variability is *significantly lower* than the Average Variability of other homogeneous samples.

It is interesting to find that while the Anglo-Indian sample is not significantly more variable than non-Caste samples, it does seem to be just significantly more variable than Caste samples.

The Anglo-Indian sample is "mixed" from a Caste standpoint but is not so from the standpoint of ordinary stable populations. We shall see later that the Anglo-Indians are about as variable as modern European samples.

(*b*) *Selected Indian Castes.*

I now select samples of 25 and more from the Caste data.

Distributions of 70 Selected Caste Coefficients of Variation.

Group	2·2	−2·3	−2·4	−2·5	−2·6	−2·7	−2·8	−2·9	−3·0	−3·1
Frequency	1	1	0	1	0	1	2	7	5	4

Group	−3·2	−3·3	−3·4	−3·5	−3·6	−3·7	−3·8	−3·9	Total.
Frequency	11	12	6	6	5	3	0	5	70

With ·1 as the grouping unit, I find

Mean Selected Coefficient of Variation $= 3{\cdot}3043 \pm {\cdot}0278$

Standard Deviation of Coeff. of Variation $= {\cdot}3458 \pm {\cdot}0197$

Mean non-Caste Coeff. of Variation $= 3{\cdot}5710 \pm {\cdot}0326$

Caste Difference $= 0{\cdot}16\,67 \pm {\cdot}0429$

In this case also the difference is nearly four times the probable error. We conclude that *the Indian Caste samples have got a substantially lower Variability than non-Caste samples.*

We shall now compare Anglo-Indian Variability with the selected Caste Variability.

Anglo-Indian Variability $= 4{\cdot}0672$

Selected Caste Variability $= 3.3043$

Anglo-Indian Difference $= {\cdot}7629$

Thus $$x = \frac{{\cdot}7629}{3458} = 2{\cdot}806$$

From Biometric Table II, $\frac{1}{2}(1+a) = {\cdot}98\ 64\ 474$

$\frac{1}{2}(1-a) = {\cdot}01\ 35\ 526$

The chance is only 13 in 1000 that the Variability of an Indian Caste will be greater than Anglo-Indian Variability. This is the lowest odds we have got up till now.

To sum up,

The Anglo-Indian Variability is significantly greater than Caste Variability but is not beyond the range of homogeneous Variability.

Other Comparisons.

I shall give a short summary of other comparisons, reserving a fuller discussion for a future paper on the Caste data

Pooling together the 84 Caste and the 109 other samples we get a total of 193 (all samples).

I find

Mean Value of Coefficient of Variation $= 3{\cdot}4231 \pm {\cdot}0240$

Anglo-Indian Co-efficient of Variation $= 4{\cdot}0672$

Anglo-Indian Difference $= {\cdot}6441$

Standard Deviation $= {\cdot}4949 \pm {\cdot}0169$

Thus $$x = \frac{.6441}{.4949} = 1{\cdot}301,$$

From Biometric Table $\frac{1}{2}(1 \pm a) = {\cdot}90\ 31\ 995$

$\frac{1}{2}(1 - a) = {\cdot}09\ 68\ 005.$

Anglo-Indian Variability would be exceeded by nearly 10% of total (Caste and non-Caste) samples.

Selecting samples greater than 25 we get a total of 137 fairly reliable samples.

Distribution of 137 Selected Coefficients of Variation.

Group ..	Beyond 2·2	-2·3	-2·4	-2·5	-2·6	-2·7	-2·8	-2·9	-3·0	-3 1
Frequency	1	1	0	1	0	3	4	8	10	5
Group	-3·2	-3·3	-3·4	-3·5	-3.6	-3·7	-3·8	-3·9	-4·0	-4·1
Frequency	13	14	11	21	12	9	8	9	1	4
Group	-4·2	-4·3	Total.							
Frequency	1	1	137							

Grouping by ·1 I find :—

$$\mu_2 = 14{\cdot}42\ 12\ 29$$

$$\mu_3 = -\ 17{\cdot}98\ 72\ 90$$

$$\mu_4 = 760{\cdot}82\ 26\ 19$$

Hence

$$\beta_1 = {\cdot}10796 \pm {\cdot}14439$$

$$\beta_2 = 3{\cdot}65892 \pm {\cdot}91379$$

Thus we are justified in applying the normal integral for calculating the chances for any deviation.

Anglo-Indian Variability = 4·0672
Mean Selected (137 samples) = 3·4412 ± ·0219
Anglo-Indian Difference = ·6260
Standard Deviation = ·3797 ± ·0155

$$x = \frac{\cdot 6260}{\cdot 3797} = 1\cdot 62,$$

From Biometric Table II, $\frac{1}{2}(1+a) = \cdot 94\ 73\ 839$
$\frac{1}{2}(1-a) = \cdot 05\ 26\ 161$

Thus over 5% will have greater Variability. The difference can scarcely be called significant.

Standard Deviations.

I shall merely give the final results. (The complete figures will be published in a supplement).

(*a*) *All Samples* (*Caste and others*) *total* = 190.
Mean Standard Dev. = 57·0684 ± ·42 71 mm.
Anglo-Indian Standard Dev. = 67·3850
Anglo-Indian Difference = 10·3167
S.D. of Standard Dev. = 8·7302 · ± 3020.

$$x = \frac{10\cdot 3167}{8\cdot 7302} = 1\cdot 181,$$

From Biometric Table II, $\frac{1}{2}(1+a) = \cdot 88\ 09\ 999.$
$\frac{1}{2}(1-a) = \cdot 17\ 90\ 001.$

Thus nearly 18% will have a greater Standard Deviation than the Anglo-Indian sample.

(*b*) *Selected Samples* (*Caste and others*) *greater than* 25, *total* = 134
Mean Standard Dev. = 56·7612 ± ·3987 mm.
Anglo-Indian Standard Dev. = 67·385
Anglo-Indian Difference = 10·6238
S.D. of Standard Deviation = 6·8424

$$x = \frac{10\cdot 6238}{6\cdot 8424} = 1\cdot 552,$$

From Biometric Table II, $\frac{1}{2}(1+a) = \cdot 93\ 94\ 292$
$\frac{1}{2}(1-a) = \cdot 06\ 05\ 708$

Six per cent will have a greater variability than the Anglo-Indians.

(*a*) *All Non-Caste Samples, total* = 106.
Mean Standard Dev. = 59·2830 ± ·6138 mm.
Anglo-Indian Standard Dev. = 67·385

Anglo-Indian Difference = 8·102

S.D. of Standard Deviation = 9·3688 ± ·4340

$$x = \frac{8\cdot102}{9\cdot3688} = 0\cdot864,$$

From Biometric Table II, $\frac{1}{2}(1+a) =$ ·80 51 055

$\frac{1}{2}(1-a) =$ ·19 48 945

Over 19% will have greater absolute variability.

(*b*) *Selected Non-Caste Samples greater than* 25, *total* = 64.

Mean Standard Dev. = 60·6563 ± ·5453 mm.

Anglo-Indian Difference = 6·7287

S.D. of Standard Deviation = 6·4676 ± ·3856

$$x = \frac{6\cdot7287}{6\cdot4676} = 1\cdot04$$

$\frac{1}{2}(1+a) =$ ·85083

$\frac{1}{2}(1-a) =$ ·14917

(*a*) *All Caste Samples total* = 84

Mean Caste S.D. = 53·0714 ± ·4693 mm.

Anglo-Indian S.D. = 67·385

Anglo-Indian Difference = 14·314

S.D. of Standard Deviation = 6·3785 ± ·3320

$$x = \frac{14\cdot314}{6\cdot3785} = 2\cdot244,$$

From Biometric Table II, $\frac{1}{2}(1+a) =$ ·98 74 545

$\frac{1}{2}(1-a) =$ ·01 25 455

Only 12 in 1000 castes will have a greater variability than the Anglo-Indian sample. Thus we may conclude that *the Absolute Variability of the Anglo-Indian sample is appreciably greater than Caste Variability.*

Also

Non-Caste Mean S.D. = 59·2830 ± ·6138 mm

Caste Mean S.D. = 53·0714 ± ·4693

Caste Difference 6·2116 ± ·7727

Thus *Absolute Variability of Caste samples is significantly greater than Non-Caste Variability.*

(*b*) *Selected Caste Samples greater than* 25, *total* = 70

Mean Selected Caste S.D. = 53·8 ± ·4429 mm.

Anglo-Indian S.D. = 67·385

Anglo-Indian Difference = 13·585 ± 2·471

S.D. of Standard Deviation = 5·4938 ± ·3131

$$x = \frac{13\cdot585}{5\cdot4938} = 2\cdot471,$$

From Biometric Table II, $\frac{1}{2}(1+a) =$ ·99 32 443
$\frac{1}{2}(1-a) =$ ·00 67 557

Thus only about 7 in 1000 will have greater variability.

Again,

Selected Non-Caste S.D. = 60·6563 ± 5453 mm.
Selected Caste S.D. = 53·8 ± ·4429

Anglo-Indian Difference = 6·8563 ± ·7025

Selected Caste Variability is thus significantly greater.

We conclude from our comparative study of variabilities that Anglo-Indian Variability though high is not sufficiently so to enable us to assert that the material is heterogeneous. The Anglo-Indian sample is however markedly more variable than the Rislay Samples of Indian Castes and Tribes.

I shall now consider a series of modern European races for which reliable data is available.

Modern European Races.

	S.D.
	mm.
Aberdeen students (493)	59·4
Cyprus (585)	61·6
Cambridge students (1000)	64·6
U.S.A. recruits (25,898)	65·6
Albanians (140)	65·7
N.S.W. Criminals (2871)	65·8
Oxford students (959)	66·1
Germans (390)	66·8
Crete (318)	67·5
Eng. Criminals (3000)	68·1
Eng. Fathers (1078)	68·7
Eng. Sons (1078)	69·4

Anglo-Indian S.D. = 67·385 mm.
Average European S.D. = 65·775

Anglo-Indian Difference = 1·61
S.D. of S D. = 2·75
Anglo-Indian excess in terms of S.D.
= 1·61/2·75 = 0·5855

Thus the Anglo-Indian variability is only 1·61 mm. greater than average variability of European races. We have however included no less than five different English samples. If we retain the largest English sample (3000 criminals) we get the Mean variability to be 65·375 mm. with a S.D. of 2·513 mm. The Anglo-Indian excess is 2·1 mm. or in terms of the S.D. is 0·79586.

We conclude that Anglo-Indian Variability is of the same order as modern European variability.

CONCLUSIONS.

I have proposed five distinct tests of "homogeneity."

I The frequency distribution should be homotypic.

II It should resist statistical dissection;

III Subsamples should not differ significantly,

IV The general nature of the distribution should be similar to homogeneous distribution.

V The Variability should not differ significantly from the average Variability of homogeneous races.

(1) I have shown that graduation by the Gaussian (possibly still better by a Type IV curve) is adequate. *Anglo-Indian frequency distribution is certainly homotypic.* Our first test thus fails to show any sign of heterogeneity in the material.

(2) Excepting for a very special type of dissection (which is probably a peculiar feature of the particular sample considered) *statistical analysis into component groups is not possible.* Our second test too fails to detect heterogeneity.

(3) We have seen that the difference between subsamples is statistically insignificant. *Subsamples seem to agree quite well,* thus confirming statistical homogeneity of the material.

(4) The *general nature of Anglo-Indian frequency distribution is also similar to other homogeneous distribution.* Anglo-Indian distribution is approximately Gaussian with some tendency towards type IV, lepto-kurtosis and small asymmetry. Other known cases of stature distribution show the same characteristics. The fourth test thus supports the view that the present material is homogeneous.

(5) I have compared the Variability of the Anglo-Indian with Variabilities of other races in many different ways.

Anglo-Indians are more variable than the Indian Castes and Tribes but the Variability of the Anglo-Indian sample is not significantly greater than the average Variability of homogeneous samples in general.

SECTION VIII. NOTE ON CORRELATION BETWEEN AGE AND STATURE.

I shall give a short summary of the values of the Coefficient of Correlation between Age and Stature, reserving a fuller discussion for a future part.

(*a*) *The whole series* (*all ages*), *total* = 191.

The age has been recorded in the case of 191 out of the total group of 200 which we have been considering so far. I have used the standard "product moment" method.[1]

I find for stature, with 50 mm. unit of grouping and 1660 mm. as base number,

$$\nu_1' = -\ \cdot14136, \quad \text{and}$$

Thus Mean Stature = 1656·14 mm.

S.D. = σ_x = 65·4923 mm.

For age, with one year unit of grouping and base number = 24 years,

$$\mu_1' = +\cdot27$$

and

$$\nu_2' = 44\cdot98.$$

Thus Mean Age = 24·27 years

S.D. = σ_y = 6·7022 years.

With the same units and base numbers we find the product moment to be + 40· 26 70.

Correcting for base number, we have

μ_{11} = Product moment = + 40·22

Thus
$$r = +\frac{40\cdot22}{6\cdot7022 \times 65\cdot4923}$$
$$= +\ \cdot1089$$

The Probable Error is given by $\cdot6745\ (1-r^2)/\sqrt{n}$

N = 191, hence P.E. is = ± ·049.

We have then

$$r = +\cdot1089 \pm \cdot049.$$

The correlation coefficient is slightly over twice its Probable Error, hence it is not definitely significant. In any case the correlation between age and stature seems to be small.

The low average age of the whole sample shows the presence of a considerable number of individuals in their early youth. I next separated the measurements of those above 25 years of age

---

[1] Yule, *Statistics* Chap. IX.
Karl Pearson: "Regression, Heredity and Panminia" *Phil. Trans. Roy. Soc.* Vol. 187A, 1896.

from the measurements of those below 25, and considered the correlation for the two different age groups separately.

(*b*) *Age group below* 25, *total* = 125.

I find for age

	Mean Age	= 20·52 years
	S.D. $= \sigma_y$	= 2·2449 years
For stature,	Mean Stature	= 1649·35 mm.
	S.D. $= \sigma_x$	= 61·35
Also	r_{11} = Product moment	= + 20·16

We notice that the average stature of the lower age group is only 7 mm. less than the general average. The S.D. is also less than the general average, showing that the lower age group is less variable than the general sample. I shall come back to this point later on.

We find the coefficient of correlation to be

$$r = +\cdot1464 \pm \cdot058$$

The correlation is positive but small. It is just on the verge of being significant. The positive character of the coefficient is of course expected, it merely indicates, or rather actually measures the average rate of growth with age. The material includes only a few cases of 16, the lowest age group, and so it is not possible to say very much about the actual variations in the rate of growth. The smallness of the coefficient (if not due to errors of sampling) seems to suggest that the greater part of the increase in stature is attained before the age of 16 or 17. Thus the Anglo-Indian seems to be, so far as stature is concerned, rather precocious in growth. I shall discuss this point after investigating the correlation between age and the other characters.

(*c*) *Age group above* 25, *total* = 66.

I find

Mean age	= 31·38	years
S.D. $= \sigma_y$	= 6·5765	years
Mean Stature	= 1688·1818	mm.
S.D. $= \sigma_x$	= 71·072	mm.
Product moment	= −55·4884	

Thus, $r = -\cdot1187 \pm \cdot08$.

The coefficient is now negative but is scarcely significant in view of its large probable error. A small negative correlation is to be expected in view of the shrinkage which sets in after 25 or 30.[1]

---

[1] Powys: "Anthropometric Data from Australia," *Biometrika* Vol, I (1902) p. 49.

The value of the Absolute Variability is for the

Lower age group $= 65{\cdot}4923 \pm 2{\cdot}7939$ mm.
Higher age group $= 71{\cdot}0720 \pm 4{\cdot}1726$ mm.
Difference $= 5{\cdot}5797 \pm 5{\cdot}02$ mm.

The variability of the younger group is thus considerably less, but the difference is scarcely significant. Even though we cannot definitely assert that the variability is being reduced with time, the above noticed decrease is certainly interesting as giving an indication that such a view is not altogether untenable.

If we turn to the Relative Variability, i.e. the Coefficient of Variation, we find

Higher age group $= 4{\cdot}2073 \pm {\cdot}2470$ mm.
Lower age group $= 3{\cdot}9545 \pm {\cdot}1687$
Difference $= 0{\cdot}2528 \pm {\cdot}2991$

The difference is less significant than the previous one. But the reduction in even the relative variability is distinctly suggestive.

Another point must be carefully noted. The variability of the Anglo-Indian sample is not significantly diminished by selection of age groups. Thus the high value of the variability (both absolute and relative) is not merely due to the mixing of the different age groups but represents a real degree of dispersion.

SECTION IX. SUMMARY OF CONCLUSIONS.

Statistical.

(1) For stature, with samples of the order of 200, a grouping unit of 50 mm. is fairly satisfactory. For calculating frequency constants the grouping unit should be less than $3{\cdot}5\sqrt{N}$ (for samples of size N).

(2) Sheppard's corrections lead to substantial improvement in the frequency constants and should never be omitted. With small samples finer corrections (e.g. Pairman and Pearson) are useless.

(3) The Gauss-Laplacian normal curve is adequate for 50 mm. grouping. For proper graduation, i.e. for testing goodness of fit the grouping should be broader than $700/\sqrt{N}$

(4) The actual frequency curve belongs to Type IV of Pearson's Skew family. There is small positive asymmetry with the Mode greater than the Mean, and a slight tendency towards leptokurtosis. The general nature of the distribution is similar to other homogeneous distributions.

(5) There is no definite evidence of statistical heterogeneity. The Anglo-Indian sample may be accepted as a *statistically* homogeneous sample.

Anthropological (*Stature*).

(1) The more highly civilised races have greater variabilities than the average.

(2) This greater variability of more highly civilised races seem to be only moderate in degree and is never excessive.

(3) Interracially, taller races seem to be more variable than the shorter (both as regards the absolute and the relative variability).

(4) Indian Castes and Tribes are significantly less variable than the average.

(5) Anglo-Indian variability is greater than Indian Caste variability but is of the same order as the variability of modern European races.

(6) The variability of the Anglo-Indian sample though greater than the average is not beyond the range of possibility of homogeneous variability

(7) The Anglo-Indians seem to be rather precocious in growth, and there is some indication of the arrest of growth occurring at an earlier age than in the case of European races.

(8) Variability of the smaller age-groups is distinctly less, showing a decrease of variability with time (or increasing homogeneity of the younger generation).

APPENDIX I. NOTE ON STATISTICAL TERMS.

In this appendix I have made an attempt to explain, in non-mathematical language, some of the more frequently occurring technical terms of statistical theory. Considerations of space have prevented me from giving concrete illustrations. I hope however that the following pages will serve some useful purpose in helping anthropologists who lack the requisite mathematical training, in taking an intelligent interest in the various technical discussions contained in this paper. I have only attempted to give a general idea of the different terms; the statistician will, I hope, forgive me for the consequent lack of precision in many places.

Let us consider our 200 measurements of Anglo-Indian stature. Almost all individual measurements are different from one another. The existence of variability is patent. The important fact is, however, that this variability of stature is not chaotic in its distribution, but that it is governed by definite laws.[1]

We can classify our material into different groups in accordance with size. We find, for example, that there are 2 individuals whose heights are less than 1465 mm. Between 1465 and 1485, there is only one. Between 1485 and 1505, there are 4, and so on. Thus with a 20 mm. unit of grouping, we get the following distribution of frequency in each group. (The number of individuals in any group is called the frequency of that group).

Frequency Groups in units of 20 mm.

Group	1445 mm. to 1465 mm.	1465 mm. to 1485	1485 to 1505	1505 to 1525	1525 to 1545	1545 to 1565	1565 to 1585	1585 to 1605	1605 to 1625	1625 to 1645
Number	2	1	1	4	2	4	10	12	25	32

Group	1645 to 1665	1665 to 1685	1685 to 1705	1705 to 1725	1725 to 1745	1745 to 1765	1765 to 1785	1785 to 1805	1805 to 1825	1825 to 1845	1845 to 1865	Total
Number	21	17	21·5	18·5	10	5	10	2	0	1	1	200

These frequency groups are shown graphically in Plate I.

Let the horizontal x-axis represent stature. Then, at the middle point of each group, we can erect vertical lines proportional to the frequency in that group. For example, at 1455, which is the middlepoint of the group 1445–1465, we erect a vertical line whose length is two units, to represent the frequency in that group. At 1475, the height of the vertical line is one unit and so on. If the extremeties of these vertical

[1] Cf. Goring: "The English Convict," p. 29.

lines are joined by straight lines, we get the corresponding *frequency polygon.* With 20 mm. unit of grouping, the polygon is broken and irregular in outline, because many intermediate measurements are missing in the sample.

If we gradually increase the size of our sample, more and more of these gaps will be filled up and the polygon will become more and more regular. On the other hand, with an indefinitely large sample, we can make the size of each group as small as we please, without incurring any risk of meeting with gaps in the measurements. Thus, with a very large sample, and when the size of each group is indefinitely diminished, the discontinuous broken polygon will gradually pass into a continuous smooth curve. This *frequency curve* will give us the distribution of stature of an indefinitely large population.

Such distributions are usually termed *Chance* distributions. But as Pearson observes,[1] "in the first place, we have to recognise that our conception of chance is now utterly different from that of yore. Where we cannot predict, where we do not find order and regularity, there we should now assert that something else than chance is at work. What we are to understand by a chance distribution is one in accordance with law and order, and one the nature of which can for all practical purposes be closely predicted. It is not theory, but actual statistical experience, which forces us to the conclusion that, however little we know of what will happen in the individual instance, yet the frequency of a large number of instances is distributed round the mode in a manner more and more smooth and uniform the greater the number of instances. Our conception of chance is one of law and order in large numbers; it is not that idea of chaotic incidence that vexed the mediaeval mind."

The *Gaussian* distribution (named after the great mathematician Gauss) is one important standard type. It has got the following characteristics:—

(*a*) The frequency is maximum for the average value of the organ measured.

(*b*) The distribution is symmetrical with regard to this maximum.

(*c*) The curve slopes down, gradually and in a characteristic way, to zero, so that extreme degrees of variation become increasingly rare.

(*d*) The curve ends tangentially to the x-axis, so that infinitely large degrees of variation are theoretically possible.

Variability.—We have not yet investigated the question of variability of the distribution. Two frequency distributions may be both Gaussian and yet their variabilities may differ widely. Anthropologists have often used the *range*, which is defined as the difference in size of the most extreme members, as a measure of variability. A little reflection will, however, show that the range

[1] Chances of Death, Vol. I, p. 11.

is not at all suitable for this purpose. The inclusion in the sample of a single abnormal "dwarf" or "giant" will completely upset the value of the range. A measure so radically affected by stray items at the extremes is practically useless for scientific purposes.[1]

In current statistical practice it is usual to measure variability by the *Standard Deviation.* The deviation of each measurement from the Mean (or Average) is squared. The sum of all such squares divided by their total number gives the second moment μ_2, which is thus the average squared-deviation of all the measurements. The square-root of μ_2 finally, gives the Standard Deviation. It is the average root-square deviation of all the measurements, and is a precise mathematical measure of the variability of the sample. One great advantage in using the Standard Deviation is this that it uniquely defines the corresponding Gaussian curve, so that the Gaussian can be found as soon as the Standard Deviation is determined. Standard Deviation (or S.D.) is usually represented by σ.

Probable Errors.—The Gaussian distribution is also known as the "normal curve of errors," since it is assumed that this curve gives the distribution of "errors" made in physical measurements.[2] The greater the diversity in any set of measurements the greater will be the Standard Deviation of the set. Accuracy or reliability depends on the uniformity of the set of measurements, that is, on the smallness of the Standard Deviation The "probable error," which measures the accuracy or reliability of any set of measurements, is hence suitably defined by a particular sub-multiple of the Standard Deviation.

If σ is adopted as the unit of measurements (that is, all measurements in terms of ordinary units are divided by σ), then the curve of errors becomes the standard curve of probability. The mathematical theory of probability then enables us to find the probability of any given deviation from the Mean occurring in the sample.

For example, a deviation half of the Standard Deviation will occur no less than 62 times in 100 samples. A deviation as great as the Standard Deviation will occur in 32% instances, while a deviation four times as great will not happen more than once in 17,000 instances. The Probable Error is defined to be such a deviation as will be exceeded by half the total deviations, or in other words, the chances are even that any deviation will be greater than or less than the Probable Error.

We must now come back to Anthropology. It is well known that almost all anthropometric measurements have an approximately Gaussian distribution. This was originally pointed out by Quatelet, and since then has been confirmed by many different

---

[1] For a simple non-technical account of the different measures of dispersion, see King: "Elements of Statistical Theory" (MacMillan, 1919), p. 141.

[2] This assumption itself is not always strictly true. See Pearson's memoir on. "Errors of Judgement, etc." *Phil. Trans. Roy. Soc.* 198A (1902).

observers.[1] But it must be remembered that the distribution is only *approximately* normal and is almost never exactly so. We are thus obliged to study other types of frequency distribution.

It is often found that the maximum frequency does not occur at the Mean value of the character concerned. In such cases, the most frequent size, that is, the position of the maximum ordinate, is called the *Mode.* In anthropometric measurements it is very usual to find the Mode different from the Mean. When this happens, the distribution is no longer symmetrical about the Mean. Such *asymmetrical* distributions are called *skew* distributions.[2]

The distance between the Mode and the Mean is one obvious measure of *skewness*, or better still (for purposes of comparison), this distance divided by the Standard Deviation. The mathematical measure of skewness depends on the third moment μ_3, obtained by cubing the deviations from the Mean and taking the average. The positive and negative deviations (from the Mean) must, by the very definition of the Mean, balance exactly; so that the sum of all deviations is zero. For a symmetrical curve this is also true of the cubes of deviations. But in the case of an asymmetrical curve, the sum of all the cubes of deviations is not zero. Hence the third moment, which is merely the average sum of the cubes of all deviations, is not equal to zero. Thus μ_3 or more conveniently $\beta_1 = \mu_3^2/\mu_2^3$ is a precise measure of the degree of asymmetry. If β_1 is significantly different from zero, then the curve must be considered skew.

Frequency distributions may differ from the normal curve in another particular. The change of slope of the normal curve is a characteristic feature of the curve. Now a frequency curve may differ from the normal as regards the manner in which its slope changes. For example, if a curve rises more abruptly, than the normal curve, it is then called a *lepto-kurtic* curve. While if it is more flat-topped than the normal, it is called a *platy-kurtic* curve. Curves with the same degree of abruptness as the normal are known as *meso-kurtic* curves. The kurtosis is measured by $\beta_2 - 3$. For meso-kurtic curves β_2 is equal to 3, and the kurtosis is zero. For lepto-kurtic β_2 is greater than 3, and for platy-kurtic it is less than 3. A frequency curve may also differ from the normal in having a definitely limited range. The curve may be limited in one or in both directions. With these curves there is a definite theoretical limit to the size of deviations.

The Coefficient of Variation.—Pearson[3] says, "In dealing with the comparative variation of men and women.. , we have constantly to bear in mind that relative size influences not only the means but the deviations from the means. When dealing with absolute measurements, it is, of course idle to compare the

---

[1] For references see pp. 42-44.

[2] For literature on the subject see references quoted on p. 16. Also J. C Kapteyn: "Skew Frequency Curves in Biology and Statistics."

[3] Karl Pearson: "Regression, Heredity and Panmixia," *Phil. Trans. Roy. Soc.*, Vol. 187A, 1896, pp. 276-277.

variation of the larger male organ directly with the variation of the smaller female organ. The same remark also applies to the comparison of large and small built races . We may take as a measure of variation the ratio of Standard Deviation to mean, or what is more convenient, this quantity multiplied by 100. We shall, accordingly, define V, the coefficient of variation, as the percentage variation in the mean, the Standard Deviation being treated as the total variation in the mean .. Of course, it does not follow because we have defined in this manner our "coefficient of variation," that this coefficient is really significant in the comparison of various races, it may be only a convenient mathematical expression, but I believe there is evidence to show that it is a more reliable test of "efficiency" in a race than absolute variation . . By "race efficiency," I would denote stability, combined with capacity to play a part in the history of civilisation."

APPENDIX II.

TABLE OF MEASUREMENTS.

Stature of Anglo-Indians measured in the Anthropological Laboratory of the Indian Museum, Calcutta.

Card No.	Age in years.	Stature in mm.
87	15	1446
166	16	1624
186	16	1588
147	16	1666
144	16	1726
250	17	1656
175	17	1588
145	17	1588
289	17	1544
76	17	1642
120	17	1610
143	17	1662
44	17	1706
253	18	1746
141	18	1768
132	18	1610
258	18	1602
86	18	1636
191	18	1660
160	18	1570
251	18	1574
94	18	1580
288	19	1638
277	19	1636
294	19	1634
66	19	1630
286	19	1626
298	19	1614
53	19	1604
75	19	1586
288	19	1458
226	19	1768
4	19	1768
151	19	1760
295	19	1744
174	19	1718
299	19	1705
6	19	1706
56	19	1780
146	19	1674
176	19	1686
73	19	1666
246	19	1550
26	19	1646
140	19	1644
91	19	1640
46	20	1716
110	20	1712
142	20	1710
235	20	1700
8	20	1680
175	20	1670
14B.	20	1673
168	20	1664
241	20	1638
156	20	1622
280	20	1622
248	20	1562
65	20	1500
275	20	1510
217	20	1514
152	20	1610
67	20	1650
219	20	1620
172	20	1658
107	21	1626
102	21	1636
111	21	1650
287	21	1654
234	21	1656
99	21	1708
133	21	1730
101	21	1768
51	21	1768
106	21	1704
10	21	1694
281	21	1696
28	21	1672
227	21	1678
267	21	1624
88	21	1628
9	22	1730
148	22	1726
74	22	1716
180	22	1700
149	22	1700
108	22	1684
103	22	1688
170	22	1677
72	22	1650
96	22	1568
43	22	1576
177	22	1608
25	22	1644
68	22	1644
7	22	1636
136	22	1636
134	22	1616
243	22	1654
62	22	1658
40	23	1692
11	23	1692
265	23	1680
12B.	23	1775
64	23	1472
61	23	1572
269	23	1624
42	23	1646
224	24	1592
45	24	1610
54	24	1620
297	24	1690
50	24	1670
1	24	1684
13	24	1696
230	24	1634
284	24	1596
276	24	1636
47	24	1644
57	25	1738
262	25	1513
54	25	1580
285	25	1619
48	25	1630
3	25	1634
293	26	1644
240	26	1638
282	26	1656
263	26	1730
2	26	1710
60	26	1604
58	26	1628
231	26	1616
63	27	1522
27	27	1700
119	27	1692
38	27	1770
39	27	1776
29	27	1796
137	27	1840
20	27	1656
14	27	1650
31	28	1610
232	28	1636
223	28	1754
32	28	1662
78	28	1662
271	29	1780
268	29	1584
36	29	1730
278	29	1722
247	29	1620
33	29	1608
279	29	1562
55	29	1578
35	30	1656

Card No	Age in years.	Stature in mm.
135	30	1712
22	30	1734
34	30	1760
260	30	1708
216	30	1684
232	30	1640
229	30	1628
77	30	1614
18	30	1672
266	30	1694
37	30	1698
283	31	1716
220	32	1640
97	32	1606
105	32	1592

Card No.	Age in years.	Stature in mm.
19	32	1714
201	32	1720
70	32	1734
159	33	1617
264	33	1624
93	33	1788
79	35	1704
155	35	1722
17	35	1670
71	38	1644
82	39	1610
92	39	1714
252	40	1848
131	41	1638

Card No	Age in years.	Stature in mm.
49	42	1704
52	42	1756
165	43	1540
15	44	1610
5	45	1598
95	48	1574
98	..	1554
16	..	1586
228	..	1632
30	..	1654
12	..	1670
150	..	1690
118	..	1694
112	..	1700
6B	..	1711