

*Reprinted from*

*The Proceedings of the Second Indian Statistical Conference, Lahore, 1939.*

DISCUSSION ON THE  
APPLICATION OF STATISTICAL METHODS  
IN ANTHROPOMETRY

Dr. D. N. MAJUMDAR

Prof. P. C. MAHALANOBIS

Prof. K. P. CHATTOPADHYAY

STATISTICAL PUBLISHING SOCIETY

CALCUTTA

*Agents: P. S. King & Sons Ltd., London.*

# ANTHROPOLOGY

Chairman : Dr. D. N. MAJUMDAR, M.A., Ph.D. (Cantab.),  
*President, Section of Anthropology, Ind. Sc. Congress, 1939.*

7th January, 1939.

[Joint Meeting with the Section of Anthropology of the Indian Science Congress]

---

## DISCUSSION ON THE APPLICATION OF STATISTICAL METHODS IN ANTHROPOMETRY

DR. D. N. MAJUMDAR : NEED OF STANDARDISATION OF ANTHROPOMETRIC TECHNIQUE

This is the first occasion when the Indian Statistical Conference is having a joint session with the Anthropology Section of the Indian Science Congress. In his Presidential Address to the Anthropology Section of the Science Congress at Benares in 1925, Prof. P. C. Mahalanobis gave an important statistical analysis of Race-mixture in Bengal based on measurements of Anglo-Indians (old style) by the late Dr. N. Annandale of the Zoological Survey of India and on Risley's data relating to castes and tribes of Bengal. Since then Prof. Mahalanobis has taken keen interest in anthropometrical research in India. This Joint Session is a proof of the interest the statisticians are taking in our subject, and I hope, with the help and co-operation of the statisticians, it would be possible for anthropologists to put Anthropometry on a more scientific basis.

As a field anthropologist, I have to take physical measurements of living people. When I find within a certain geographical area a number of cultural groups who live independently of one another, or develop the closest co-operation among themselves, speak the same language but differ in definite somatic traits, I am tempted to investigate into the racial composition of the groups and ascertain the degree of racial affinity or divergence between them. Recently my investigations into the effects of culture contacts in certain areas have shown that a comparison of racial traits of the various cultural groups is indispensable for a proper evaluation of the role of each cultural group in the formation of a significant culture-complex. I have taken a large number of anthropometric measurements of primitive tribes and have undertaken a systematic racial survey of the people of the United Provinces. I have followed with interest the known methods of statistical analysis. I find, however, that there is very little agreement among anthropologists regarding the technique followed, the number of characters considered and also the size of the samples. It is unfortunate that the few anthropologists we have in India who take interest in anthropometry and craniometry do not recognise one another's methods and are working on divergent lines, so much so that it is impossible to compare the data collected by two anthropologists on the same population or even on the same samples.

Apart from the fact that there is no agreed number of measurements to be taken on any given sample of population, there is no uniformity in the methods of analysis and we have today rule of thumb methods like that of average and highly statistical ones like the coefficient of racial likeness or the coefficient of caste distance, all of which are followed with perfect equanimity.

It has been long recognised that mere description of indefinite physical characters of different racial groups does not exhaust the possibilities of comparison of racial traits. There are certain characters to which instruments of precision may be applied and the measurements can be expressed in numerical terms, so that the data can be statistically treated. Measurements of the human body or its parts are taken today for various purposes. Anthropologists, however, undertake these measurements for purposes of racial classification and for ascertaining the racial distance between groups, as also the degree of race admixture.

Sir Herbert Risley was the first to undertake a systematic anthropometric survey in India. He followed the technique of Topinard and Sir William Flower. The measurements were published in a volume of *Ethnographic Appendices to the Census of India, 1901*. The data of Risley, which were collected by his assistants in accordance with his instructions recorded in "Anthropometric Instructions" (Calcutta, 1886), were recently reviewed by Prof. P. C. Mahalanobis who found out a number of discrepancies, particularly in the calculation of average values and individual measurements. Like older anthropologists, Risley's analysis of the numerical data was based on the method of averages without considering how far these averages could be taken as typical measurements of the group from which the original measurements were obtained.

But in his instructions to his assistants, Risley asked them to reject persons of very black complexion and with very broad and depressed noses when measuring the higher castes, and also to exclude men of very fair complexion and high caste type of features while measuring the lower castes. This search for the standard type has naturally led to a selection of subjects and it is difficult to say how far these instructions were carried out by his assistants, but it appears that the selection has had significant effect on the calculated values as the standard deviations of Risley's measurements tend in most cases to be lower than values for corresponding characters in groups measured by other anthropologists. It is no wonder, therefore, that many of Risley's conclusions have been challenged by later workers.

For a long time anthropologists in India were content with following the method of averages as detailed by Risley. As there was no agreed formula about the number of characters to be measured, even the 10 or 12 characters recorded by Risley were not regarded as necessary by later workers; and very often we find anthropologists visualising the broad outlines of Indian racial history on the basis of 3 or 4 absolute physical characters.

Some anthropologists in India accepted the fundamental types of Prof. Ronald Dixon and began to look for these types in the population of India. Instead of averaging the several measurements and indices, as was commonly the practice, Dixon found it convenient to analyse them into a number of groups based upon the combinations of the different indices. For example, he analysed the cephalic and nasal indices of 25 Khasis of Assam and found that almost half were mesocephalic and at the same time mesorhine; about one sixth (16 per cent) were dolichocephalic and at the same time mesorhine, etc. Assuming that a mesocephalic group is, in the ultimate analysis, the result of a blending of brachycephalic and dolichocephalic elements, and that, similarly, a mesorhine group owes its origin to the mixture of platyrhine and leptorhine elements, the 48 per cent of mesocephalic and mesorhines may be regarded as derived from the mixture of equal parts of dolichocephalic platyrhine and brachycephalic leptorhine; the 16 per cent dolichocephalic mesorhines from a blending of dolichocephalic platyrhine and dolichocephalic leptorhine, etc. On the basis of such analysis, the Khasi people are the result of the mixture of two main types, brachycephalic leptorhine and dolichocephalic platyrhine, but also include small factors

of the other two possible types, brachycephalic platyrrhine and dolichocephalic leptorrhine. Dixon, however, admitted that 25 was a very small number and conclusions derived from them were extremely tentative, but he believed that intermediate forms are the result of the blending of extreme forms, *e.g.*, that the mixture of a people characterised by dolichocephaly with one of brachycephaly type would in the end lead to a group characterised in the main by mesocephaly. As no definite proof exists of the inheritance of metrical factors in accordance with Mendelian laws, Dixon argued that he was justified in regarding these factors as on the whole characterised by blended inheritance and continuous variation. Dixon's methods of analysis have not been accepted by most of his colleagues, as even the most homogeneous group on analysis provides such classification as that among the Khasis.

An investigation into the nature of the material dealt with by anthropologists shows that not only is there a diversity of methods employed by them but the number of subjects on which such methods have been employed varies from 5 to 20 or more. As regards the technique of measurements, although an increasing number of workers are following the Monaco Agreement, there are many who mix the various techniques and adopt only those which are convenient to them. The time has, therefore, come when anthropologists should decide on a standard technique which should be followed in all cases so that the data collected by various workers may be conveniently compared and conclusions drawn from them.

Regarding the number of characters also, it is desirable that some such standardisation should be made, so that the characters measured by one may be comparable with those recorded by another. There is an urgent need also of *weighting* the characters, so that the racial significance of the characters may be understood by the person taking measurements. There are some anthropologists who make a fetish of the number of characters and the size of the sample. For example, one author has recently recorded 22 measurements on the head, 30 measurements of the body proportions and 40 observations. All these were taken, we are told, on 258 subjects. But however satisfactory this method of investigation may be, it presents obvious difficulties. It is impossible to secure subjects for measurements who will submit to 92 tests; and even if it were possible in the laboratory, it is extremely unlikely that such data can be safely collected in the field.

In connection with the Census of India, 1931, Dr. B. S. Guha of the Zoological Survey of India has published the results of analysis of measurements of 34 racial groups. Dr. Guha has used the C. R. L. method in his analysis of the anthropometric data and has shown the different racial strains in the population of India. As we were settling down to this method, we were warned by Prof. Fisher that this method of C. R. L. should not be made to tell more than it was ever meant to. He says that as a test of significance it is not a reliable one as it does not take into consideration the correlation among different measurements but treats them as though they were statistically independent. The effect of this is to cause very high or very low values of the coefficient to occur more frequently by chance than they should. This effect, Prof. Fisher adds, increases rapidly both for statistical and for anatomical reasons as the number of different measurements used is increased and may perhaps account for some of the great contrasts in the values of the coefficients which have been found.

Dr. G. M. Morant, who was one of the earliest to use the coefficient of racial likeness in anthropometry, points out the limitations of the C. R. L. method in the following words: "It is not a true measure of absolute divergence and must not for a moment be considered as such. When it is said that a low coefficient between two races A and B indicates a closer relationship than a high coefficient between A and C, what is meant always is that

it is more probable that A and B are random samples from the same population than that A and C are."

The points raised in the above paragraphs indicate the need for standardisation in technique, in the size of samples and the number of characters. While we recognise these as urgent and vital to our science, we also appreciate the difficulties of anthropologists faced with so many methods of statistical analysis, for they find it increasingly difficult to adopt any one as the standard technique. When the statisticians claim that they forge the tools which are indispensable if we are to interpret our anthropometric data, is it too much to expect of them that their tools should be not only workable but also durable? The present plight of the science of Anthropometry in India and elsewhere is not the working of the anthropologists alone. The statisticians have a rightful share in it; and in order that the science may be placed on a sounder footing, the statistician and the anthropologist should recognise each other's limitations and work out a practical scheme which will raise the prestige of the science and increase its usefulness.

PROF. P. C. MAHALANOBIS: THE APPLICATION OF STATISTICAL METHODS IN  
PHYSICAL ANTHROPOMETRY

For many years I have been pleading for a closer co-operation between physical anthropologists and statistical workers. If physical measurements are to be used in anthropometry it is essential that such measurements should be both collected and interpreted in such a way as to enable valid conclusions being drawn on an objective basis.

Several things are needed for this purpose. For valid comparisons it is desirable that the same set of characters should be used in surveying different castes, tribes, or other groups of human individuals. Secondly, the same standard definition and the same technique of measurement should be used by all observers for each character. Thirdly, there should not be any conscious selection or bias in choosing the individuals, that is, the individuals measured should form what is known as a "random sample" of the group or population under survey. This question, however, does not require any special consideration in the present connexion, as the procedure to be adopted is practically similar in every problem in which a sample method has to be used. Finally, the same standard methods should be used in the statistical analysis of the material so that results for different groups can be compared without ambiguity.

I discussed the question of a uniform technique in a paper on the "Need of Standardization of Measurements in the Living" in the *Biometrika*<sup>1</sup> more than ten years ago. Although some progress has been made since then a great deal of laborious work is still necessary before the aim of standardization is adequately attained. One promising line of advance will be to investigate carefully the question of personal equation or individual variations between the different observers in taking measurements of the same character. This work cannot be undertaken single handed; active co-operation of a number of observers is essential for this purpose. But provided two or more physical anthropologists are willing to participate in the work, the procedure will be straightforward. Broadly speaking all that is needed is to have measurements of a selected number of characters taken on the same set of subjects (who need not belong to any particular caste, or tribe, or homogeneous group), by different observers on different occasions so as to enable exact and valid comparisons being made between the different sets of measurements. Once this is done it will be possible to study on an objective basis how far and to what

1. Mahalanobis, P. C.: On the Need for Standardization in Measurements on the Living. *Biometrika*, Vol. XX., July, 1928, 1-31.

## STATISTICAL METHODS IN ANTHROPOMETRY

extent the definition and technique of measurements of each of the characters surveyed is adequate in the sense that measurements taken by different observers are in satisfactory agreement. Such controlled experiments may show for example that the definition of certain characters are more satisfactory than others in which case it would be necessary to modify the definitions of the unsatisfactory characters until they become adequately standardized. On the other hand we may find that in certain cases satisfactory standardization is not possible in which case the characters in question will have to be eliminated from our work ; or other characters may have to be substituted in their place. It is, therefore, quite likely that it will be necessary to explore the possibilities of new characters in the course of this work.

I am not suggesting that the work of standardization of definitions will be easy ; obviously a great deal of painstaking experimentation is needed in this connexion. Carefully designed experiments on the lines explained above are, however, bound to throw a flood of light on the subject and to offer many clues for a satisfactory solution of this fundamental question.

The second thing necessary is the setting up of a standard list or possibly a series of standard lists which may be used under appropriate conditions. At present different anthropologists use different lists which often makes it difficult or impossible to make valid comparisons between different sets of measurements.

In framing the standard lists of characters it will be necessary, of course, to keep several things in mind. The characters selected must, of course, be such as are capable of being measured with objective reliability by different observers. Secondly, the characters should be statistically independent as far as practicable, as this may reduce the computational work very considerably. The selected characters should also have high discrimination power for distinguishing between different racial groups. Finally, the number of characters selected must be as small as possible and preferably of the order of seven or eight as otherwise the computational labour would become unmanageable.

In preparing the standard lists it is quite possible, of course, to find that appropriate modifications are necessary to suit the special needs of particular geographical areas or particular racial groups. In the case of India, for example, we may find it convenient to use a basic list for India as a whole supplemented by lists of special characters for different provinces or groups of castes and tribes.

The question of differences in discriminating power can be considered only in relation to methods of statistical analysis. Karl Pearson's "Coefficient of Racial Likeness"<sup>2</sup> was developed for this purpose and was used, I believe, for the first time in 1921.<sup>3</sup> A little later, on a careful examination of the question I found that this particular method was defective in many respects from the point of view of statistical theory. In 1925 in my presidential address to the Anthropological Section of the Indian Science Congress the concept of a caste distance was used in discussing the question of race mixture in Bengal ,

- 
2. Pearson, Karl : On the Coefficient of Racial Likeness. *Biometrika*, Vol. XVIII, Pp. 105—117.
  3. Tildesley, M. L. : A First Study of the Burmese Skull. *Biometrika*, Vol. XIII, 1921, 247-251.
  4. Mahalanobis, P. C. : Analysis of Race-Mixture in Bengal, *Presidential Address, The Anthropological Section, Indian Science Congress, 1925. Jour. Asiat. Soc., Bengal, 23, No. 3, 301—333, 1927.*

and the method was further developed in a number of other papers.<sup>5</sup> Prof. R. A. Fisher in England and Prof. H. Hotelling in the United States independently worked on the same problem; and a comprehensive review of the whole subject has been given in a recent paper on the "Utilization of Multiple Measurements" by Prof. R. A. Fisher himself<sup>6</sup> in which it is gratifying to find that he has supported the use of what he called "Mahalanobis's generalized distance" for comparative work in anthropometry. Prof. Fisher has also pointed out the need of using certain directional vectors on which work is still proceeding. There are still outstanding difficulties, such as differences in variabilities and organic correlations between different characters among different racial groups, which are being studied carefully by statisticians. Prof. Fisher himself is working on this problem in England, and Mr. S. N. Roy of the Statistical Laboratory, Calcutta, has recently made significant progress in certain directions.<sup>7</sup>

The statistical theory of the subject cannot therefore be considered to have been completed, yet sufficient progress has been made to enable anthropologists using certain statistical methods with confidence and with the assurance that these methods are generally accepted by all responsible statistical workers.

I have already mentioned the question of limiting the number of characters. This is a problem of great practical importance. It appears probable that the use of say  $p$  different characters will involve the reduction of determinants with  $p$  rows and columns; so that computational labour is likely to increase very rapidly as the number of characters is increased. For example, to increase the number of characters from eight to nine would increase the computational labour roughly by nine times. This is why I have suggested restricting the number of characters to eight if possible.

I should, however, mention in this connexion that in making quantitative or metrical comparisons the restriction of the number of characters involves certain fundamental logical questions which require careful consideration. The situation is much simpler so long as we restrict ourselves to qualitative test of difference. For example, let us consider two samples  $S_1$  and  $S_2$  consisting of measurements on say  $p$  characters. Using appropriate statistical tests we may, on any desired level of significance, reach either of the two conclusions:

(A) The two samples  $S_1$  and  $S_2$  may be considered (at the assigned level of significance) to have been drawn from the same population, that is, they both belong to the same anthropological group; or

(B) the two samples  $S_1$  and  $S_2$  may be considered (at the assigned level of significance) to have been drawn from two different populations or two different anthropological groups.

5(a). Mahalanobis, P. C.: Tests and Measures of Divergence. *Jour. Asiat. Soc., Bengal*, 26, No. 4, 541-588, 1930.

5(b). Mahalanobis, P. C.: On the Generalised Distance in Statistics. *Proc. Nat. Inst. of Science*, Vol. II (1), 1936, 44-55.

5(c). Roy, S. N.: A Note on the Distribution of the Studentised  $D^2$ -Statistic. *Sankhyā*, Vol. 4 (3), 1939.

5(d). Bose, R. C. and Roy, S. N.: The Distribution of the Studentised  $D^2$ -Statistic. *Sankhyā*, Vol. 4 (1), 1939, 19-38.

6. Fisher, R. A.: The Statistical Utilization of Multiple Measurements. *Annals of Eugenics*, Vol. VIII (IV), Pp. 376-386.

7. Roy, S. N.: On the K-statistic. *Science and Culture*, Vol. V (2), 1939, 131-132.

In this situation if we include measurements on additional characters (besides the measurements on the  $p$  old characters) judgment (A) may get altered. That is, two samples which do not differ significantly in the first set of  $p$  characters may show significant difference when other characters are included. This possibility may always remain open. But in the case of judgment (B) we are on more sure grounds; if two samples are considered to have come from two different populations on the basis of measurements of say  $p$  characters this judgment is likely to remain unchanged even if we take fresh characters into consideration.

The position, however, is far more complicated when we are making quantitative comparisons. For example, let us consider three samples  $S_1$ ,  $S_2$  and  $S_3$ . In using the "generalized distance" we try to estimate the distance, say  $D_{12}$  between sample  $S_1$  and  $S_2$ ; distance  $D_{13}$  between  $S_1$  and  $S_3$ ; and  $D_{23}$  between  $S_2$  and  $S_3$ ; and we try to compare these distances to find out for example whether  $D_{12}$  is greater than  $D_{13}$ , that is, whether  $S_1$  and  $S_2$  are more widely separated than  $S_1$  and  $S_3$ . In this situation it is quite conceivable that results based on  $p$  characters will become different when additional characters are taken into consideration or a new set of characters is used.

I have discussed this question in an earlier paper<sup>8</sup> in which I pointed out that it appeared necessary to make certain assumptions regarding the actual physical or anthropological situation in order to reach stable judgments. For example, it may be necessary to assume that the total number of statistically independent characters is finite; or if infinite that they may be arranged in a sequence in such a way that the contribution of successive characters to the "generalized distance" becomes smaller and smaller so that the "generalized distance" itself approaches a limit as the number of characters is increased indefinitely.

The justification for using a finite number of characters smaller than the total number requires a different kind of assumption of "dimensional convergence" which was mentioned in the earlier paper. For any given number of characters, say  $p$ , let the expectation of the generalized distance for all possible random selections of  $p$  characters be  $\Delta_p$ ; we may then assume that the values of  $\Delta_p$  can be arranged in a convergent sequence as  $p$  increases.

An alternative approach is also possible. Suppose we decide to work with  $p$  characters; we may enquire what particular  $p$  characters we should choose in order that the value of the generalized distance based on these particular  $p$  characters approaches most closely the true value of the generalized distance based on the complete set of characters. On this view of the matter it is necessary to find out what characters are most useful for this purpose. It is also necessary to find out how closely the generalized distance based on  $p$  characters approaches the true value; so that the actual numerical value of  $p$  may be settled in reference to the "standard of dimensional approximation".

The subject presents difficulties; and a completely satisfactory logical solution may not be attainable in practice; but this need not discourage us, for the lack of a logically perfect foundation has not prevented progress in the physical and biological sciences. The really important thing is to proceed with the work on the basis of careful experimental observations in accordance with approved statistical principles.

---

8. Mahalanobis, P. C., Bose, R. C., and Roy, S. N.: Normalisation of Statistical Variates and the Use of Rectangular Co-ordinates in the Theory of Sampling Distribution (Appendix). *Sankhyā*, Vol. 3 (1), 1937, 35-40.



It will be seen, therefore, that the standardization of measurements, the technique of the preparation of standard lists of characters, and the use of appropriate and standard statistical method for the analysis of the material are all intimately connected. In fact it is not possible to make any appreciable progress in any one of them independently of the other two. This, I think, is the strongest argument in favour of the thesis with which I started namely the urgent need of active, close, and sustained co-operation between statisticians and anthropologists.

PROF. K. P. CHATTOPADHYAY : STATISTICAL TREATMENT OF ANTHROPOMETRIC DATA

The speaker stated that a serious defect in measurements on the living was that inadequate attention was paid to various necessary corrections. No attempt was made to calibrate instruments, so that variations due to the use of different sets of instruments by the same or different observers were not allowed for even in osteometry. An even more important point was that no attempts had been made to determine how far measurements by different observers were comparable. Some work had indeed been done comparing the different results obtained for the same measurement with different types of instruments or using different techniques. But no one apparently had tried to find out the variation in measurements taken by the same observer at different times or by different observers, on the same series of subjects. Professor P. C. Mahalanobis had indeed drawn attention to the necessity of this work several years ago ; but nothing has so far been done. The speaker said he would like to add that he was getting some preliminary work done on the same subject. If he received the wholehearted co-operation of those interested in somatometry, this work would be carried forward.

## ABSTRACT

### "THE NASION IN THE LIVING"\*

PROFESSOR K. P. CHATTOPADHYAY, CALCUTTA UNIVERSITY

The author discussed the known methods of locating the nasion and criticised the suggestions of Ashley Montagu that the position of the nasion is sharply defined by the tangent to the superior palpebral sulci. Ashley Montagu's experiments had been conducted on cadavera. The author pointed out that if the relation between the tangent and the nasion was as reported by Ashley Montagu for cadavera, then it ought to hold also for the living. If the nasion is correctly located by palpation, as it can be done in a large proportion of subjects of any group, the relation of the tangent to the superior palpebral sulcus to the nasion can be checked with ease in those cases where the nasion is definitely located. This was done by the author who had selected 120 subjects out of 156, using only those cases where the location was definite. It was found that the distance of the tangent to the sulci from the nasion was not constant. The coefficient of variation was in fact extremely high.

The author pointed out that in those cases where the nasion cannot be clearly located but is determined only indirectly, it is desirable that the measurements with reference to this point should be shown separately. Otherwise measurements which are correct to different extents may be treated as having the same degree of accuracy and lead to inaccurate results and incorrect interpretations.

---

\* Paper read in the Conference and published in *Am. J. Phys. Anthropol.*, V. 25, No. 2, July-Sept., 1939.