

APPENDIX I. HISTORICAL NOTE ON THE D^2 -STATISTIC.

By P. C. MAHALANOBIS

Karl Pearson had devised the Coefficient of Racial Likeness (C.R.L.) for comparing the resemblance of racial groups and this coefficient was first used in 1921 by Miss M.L. Tildesley in her paper on "A First Study of the Burmese Skull" (*Biom.* Vol. 13, 1921 247-251). At about the same time I had started the statistical analysis of anthropological measurements of Anglo-Indians taken by the late Dr. N. Annandale, and had been considering the problem of classification. I soon realised that Karl Pearson's C^2 was properly a test of divergence between two samples rather than a measure of the actual magnitude of the divergence. I, however, wanted an actual measure of the group distance, and adopting the notion of a generalized distance I formulated an early expression and used it in my paper on the "Analysis of Race Mixture in Bengal" in 1925.

The point to be emphasized is that the Pearsonian C.R.L. is, properly speaking, a test and not a measure of group divergence. The magnitude of C^2 determines the degree of certainty with which differentiation between the samples under consideration can be asserted, but does not supply any information regarding the magnitude or extent of such divergence. So long as samples are drawn from the same population, C^2 would be approximately equal to 0 (within the margin of errors of sampling), whatever the size of the samples. When the two samples are however drawn from two different populations, the magnitude of C^2 would depend on the size of the samples. The value of D^2 , however, would remain sensibly constant (within the margin of errors of sampling) when samples are drawn from the same two different populations, whatever might be the size of the respective samples. In other words, D^2 supplies a *measure* of the actual magnitude of divergence between the two groups under comparison. This was my chief reason for using D^2 in preference to C^2 .

In June 1927, when I was working in the University College, London, during a period of leave abroad, I showed my work on D^2 to Karl Pearson and discussed with him on several occasions the difficulties in using C.R.L. arising from the fluctuating size of samples. Karl Pearson was unable to accept my views, and pointed out that I had obtained only an approximate expression for the standard error of D^2 as I had not retained deviations of an order higher than the second. A second line of argument used by him was that the C.R.L. had actually proved to be of great value in the classification of races (based on skull measurements), a subject which he had discussed a short time ago in considerable detail in a paper "On the Coefficient of Racial Likeness."

After my return to India in 1927, I made further calculations retaining statistical deviations up to much higher orders, and succeeded in getting the exact distribution of D^2 for the null-hypothesis, that is, for the case when the two samples were drawn from the same population. I also obtained expressions for moment-coefficients up to the fourth order (which later on turned out to be exact) of the D^2 -statistic in the non-null case when the samples were drawn from different populations. These results were communicated to the Indian Science Congress towards the end of 1928.

As already stated, Pearson had shown in a review of about 750 computed values of C^2 that in actual practice the C.R.L. had been found to be an extremely useful tool in craniometric researches (*Biom.* 1926, 105-177). For purposes of comparison, I obtained by

direct calculation values of C^2 reviewed by Pearson. I found that a very large number of C^2 -coefficients (reviewed by Pearson) referred to closely associated groups for which both C^2 and D^2 would have low values. Further, owing to paucity of material, the number of skulls in each sample was also usually small, so that the size of samples did not fluctuate very widely from sample to sample. In other words, in craniometric work reviewed by Pearson, values of C^2 and D^2 gave more or less concordant results in a large number of cases. I, however, came across a certain number of comparisons for which C^2 and D^2 gave widely different results. In most of these critical cases, values of D^2 (rather than of C^2) were more in accordance with known anthropological facts showing the superiority of D^2 for purposes of classifications.

I also used the D^2 -statistic for the analysis of extensive measurements given by H. Lundborg and F. J. Linders in their great publication on "Racial Characters of the Swedish Nation" (Swedish Institute for Race Biology, Upsala, 1926). I prepared a memoir containing (a) theoretical work on D^2 -statistics, (b) its application to the Swedish material, and (c) the comparison of about 750 values of C^2 and D^2 to which I have already referred, and sent it to Karl Pearson in 1929 for *Biometrika*, but the paper was not accepted for publication. I, therefore, sent a second copy to R. A. Fisher requesting that he might communicate to some other journal for publication. In the mean time Karl Pearson, without sending any information to me, had published the anthropological portion of the work on the Swedish material under my name in a paper on "A Statistical Study of Certain Anthropometric measurements from Sweden" (*Biometrika*, Vol. 22, 1930, 94-108). This naturally prevented the publication of the full paper in England. The theoretical portion of the work was subsequently published in the form of a paper "On Tests and Measures of Group Divergence", in the *Journal of the Asiatic Society of Bengal*, New Series Vol. 26, 1930, No. 4. The second portion on the Swedish material, as already noted, was published in *Biometrika*. I had sent to the Asiatic Society of Bengal the third portion of the original memoir (which dealt with the comparison of about 750 values of C^2 and D^2). It appeared, however, that papers dealing only with Asiatic matters could be published in the *Journal of the Asiatic Society of Bengal*; and as the C^2 and D^2 coefficients related mostly to non-Asiatic races it was held that this portion of the paper could not be published in this journal. The work on the comparison of C^2 and D^2 values was thus never published.

This brings the researches in the problem up to 1930. Since then a good deal of further work has been done. In a brilliant piece of investigation, Rajchandra Bose succeeded in 1936 in obtaining the exact distribution of the D^2 -statistic in the classical form (that is, when the population values of the dispersion matrix were known). A little later, in 1938, the distribution of the D^2 -statistic in the studentised form was given by Rajchandra Bose and Samarendra Nath Roy. Suitable tests to judge the significance of differences in the dispersion matrices of two multivariate populations have been developed by Samarendra Nath Roy in a series of theoretical papers from 1939. The appropriate generalizations of the D^2 -statistic in the case of more than two populations and suitable statistics to test the equality of means of each character for a given number of populations were given by Fisher in 1939. The distribution of the ratio of D^2 's calculated on p and $(p+r)$ characters has been found by C. R. Rao in 1946. The test based on this is useful to decide whether the inclusion of r more characters to p basic characters results in an increase in the divergence measure between the populations.

U. P. ANTHROPOMETRIC SURVEY, 1941

H. Hotelling, working from an entirely different point of view, had also given in 1931, a generalized test (T) which was closely related to the D^2 -statistic. R. A. Fisher developed the use of the discriminant function in the problem of classification with multiple measurements in 1936, and later in 1939 obtained suitable generalizations of the analysis of variance to multiple measurements. He also gave a general review of the whole subject in his 1938 paper.

An account of some of the methods that were developed in the course of the analysis of the present material has been given in later appendices.

REFERENCES

1. Bose, R. C. (1936) : On the exact distribution and moment coefficients of the D^2 -statistic. *Sankhyā*, 2, 143-154.
2. Bose, R. C. and Roy, S. N. (1938) : The distribution of the studentized D^2 -statistic. *Sankhyā*, 4, 19-38.
3. Fisher, R. A. (1936) : The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7, 179-188.
4. Fisher, R. A. (1938) : The statistical utilization of multiple measurements. *Annals of Eugenics*, 8, 376-386.
5. Fisher, R. A. (1939) : The sampling distribution of some statistics obtained from non-linear regression. *Annals of Eugenics*, 9, 238-249.
6. Hotelling, H. (1931) : The generalisation of Student's ratio. *Ann. Math. Stat.* 2, 360-378.
7. Karl Pearson (1936) : On the Coefficient of Racial Likeness. *Biometrika*, 13, 105-117.
8. Karl Pearson (1928) : Note on standardization of method of using the Coefficient of Racial Likeness. *Biometrika*, 20B, 376-378.
9. Mahalanobis, P. C. (1922) : Anthropological observations on the Anglo-Indians of Calcutta. Part I, Analysis of male stature. *Rec. Indian Museum*, 23, 1-96.
10. Mahalanobis, P. C. (1925) : Analysis of race mixture in Bengal. (Presidential Address, Anthropology Section, Indian Science Congress, Benares, 1925). *Jour. Asiatic Soc. Bengal*, 23, 301-333.
11. Mahalanobis, P. C. (1928) : A statistical study of the Chinese head. *Man in India*, 8, 107-122.
12. Mahalanobis, P. C. (1930) : A statistical study of certain anthropometric measurements from Sweden. *Biometrika* 22, 94-108.
13. Mahalanobis, P. C. (1930) : On tests and measures of group divergence. *Jour. Asiatic Soc. Bengal*, 26, 541-588.
14. Mahalanobis, P. C. (1931) : Anthropological observations on the Anglo-Indians of Calcutta. Part II, Analysis of Anglo-Indian head length. *Rec. Indian Museum*, 23.
15. Mahalanobis, P. C. (1936) : On the generalized distance in statistics. *Proc. Nat. Inst. of Science*, 2, 49-55.
16. Mahalanobis, P. C. (1940) : Anthropological observations on the Anglo-Indians of Calcutta. Statistical analysis of measurements of seven characters. *Rec. Indian Museum*, 23, 151-187.
17. Mahalanobis, P. C. and Bose, Chameli (1943) : Correlation between anthropometric characters in some Bengal castes and tribes. *Sankhyā*, 5, 249-260.