GENERALIZED DISTANCE IN FAMILIAL STUDIES OF ANTHROPOMETRICAL CHARACTERS

C. Susanne
Laboratory of Human Genetics, Free University of Brussels (Belgium)

INTRODUCTION

In the analysis of multiple continuous variables in man and in the study of the heritable component of variation, methods have been published to partition the variation and covariation among a group of characters into its environmental and genetical components (Martin et al., 1977, Eaves et al., 1977). This methodology has been influenced by the fundamental works on analysis of covariances of Jöreskog (1973). This method has the advantage to determine by statistical test the adequacy of the models.

We tried to use a descriptive methodology which cannot estimate or distinct additive genetical variation from common specific environmental effects, we hope however that this method will help less mathematical minded biologists to describe in a first approach familial material in terms of genetic, environmental, and assortative mating factors.

We used a methodology based on the concept of generalized distances and described by Defrise (1967, 1968). Subjects of different age and sex can be pooled, we are thus able to compare the distances observed between different relatives. The differences between groups of relatives (twins, sibs, parent offspring, spouses) are expressed quantitatively in number of standard deviations between observed and expected mean distances. The method does not take into account sampling effects but, on large samples, such as we used in this study, these effects are negligible.

MATERIAL AND METHODS

For p normally distributed characters and a covariance matrix \( \Sigma \), the generalized distance (GD) between two subjects A and B
characterized respectively by the vectors \(a = (a_1, \ldots, a_p)\) and \(b = (b_1, \ldots, b_p)\) is defined by Defrise (1967):
\[
\Delta(A,B) = \left( (a-b)'\Sigma^{-1}(a-b) \right)^{\frac{1}{2}}
\]
It was shown (Defrise, 1955) that these GD have common properties with the classical GD of Mahalanobis.
The p variables of a subject can be interpreted as the coordinates on p orthogonal axes of a hyperspace. These measurements are transformed, by a translation of the origin of the axes to the center of the hyperellipsoids (hypersurfaces of equally probable points), then by a rotation of the axes with the effect that the original axes coincide with the axes of the hyperellipsoids and finally by a reduction of the units making the variances of the new variables equal to 1. In this new hyperspace the Euclidean distance between the new points \(A'\) and \(B'\) representing the subjects is equal to the GD between \(A\) and \(B\)
\[
A'B' = \Delta(A,B)
\]
If the individuals \(A\) and \(B\) are chosen at random in a normal \(p\) variate population,
\[
\Delta^2(A,B)/2
\]
is distributed as \(\chi^2\) with \(p\) degrees of freedom. This property can be applied to \(n\) pairs of subjects belonging to a normal \(p\) variate population. The half sum of the \(\Delta^2\) between \(n\) pairs chosen at random is distributed as \(\chi^2\) with \(np\) degrees of freedom:
\[
\sum \Delta^2(A,B)/2 = \chi^2_{np}.
\]
Moreover, in these conditions, when \(np \geq 100\),
\[
\sqrt{\sum \Delta^2(A,B)}
\]
is normally distributed about mean \(\sqrt{2np-1}\) with variance 1.
It is also worthwhile to take into account the distance of each subject to the centre of the population \(M\): indeed, if one of the members of a pair, say \(A\), lies far from the center \(M\), \(\Delta^2(A,B)\) has a greater chance of being large than when \(A\) lies nearer the center. M. Defrise (1968) showed that when \(A\) is given and \(B\) chosen at random, the distribution of \(\Delta^2(A,B)\) is a non-central \(\chi^2\) with \(p\) degrees of freedom (number of variables) and \(\Delta^2(A,M)\) as non-centrality parameter.
The probability \(P\) that the distance in the new hyperspace between a random subject and given \(A\) is smaller than the distance between \(A\) and his related partner (e.g. cotwin) \(B\) can then be calculated by a non-central \(\chi^2\) test. A small value of \(P\) indicates that the