LINEAR TRANSFORMATIONS AND THE PRODUCT-MOMENT MATRIX

By K. Nagabhushanam

Institute of Mathematical Statistics, Stockholm

Using linear transformations G. Rasch has deduced Wishart's distribution in his paper on "A Functional Equation for Wishart's Distribution" (*Annals of Math. Stat.*, Vol. 19 (1948), pp. 262–266). This note is of the nature of some observations on the Jacobian of the transformation induced by a linear transformation of coordinates with constant coefficients in the distinct elements of the product-moment matrix of a sample of n vectors, each of k components, drawn from a universe of a normal k-variate distribution with zero means. If the r-th vector of the drawn sample has components $(x_i^{(r)})$, $r = 1, 2, \cdots, n$, $i = 1, 2, \cdots, k$, the sum of the products of the i-th and j-th components of each of the n vectors is denoted by

$$M_{ij} = \sum_{r=1}^{n} x_i^{(r)} x_j^{(r)}.$$

Let the variables of the vector variate be $x_1, x_2, \cdots, x_k$, or shortly (x), and let a nonsingular (i.e., reversible) linear transformation of the variables with constant coefficients be made from (x) to (y), viz.,

$$x_r = \sum_{i=1}^{k} a_{ri}\, y_i \qquad (r = 1, 2, \cdots, k).$$

The distinct elements $M_{11}, M_{12}, \cdots, M_{1k}, M_{22}, M_{23}, \cdots, M_{kk}$ undergo a consequential or induced transformation which is also linear in terms of the corresponding elements of the product-moment matrix $\| M'_{ij} \|$ of the same n-vector sample in the coordinates (y).

Let the matrix of the coefficients of the induced transformation which is also the matrix of partial derivatives in this case be denoted by $\| J \|$, and let its determinant which is the Jacobian of the transformation be denoted by J. The elements of $\| J \|$ are functions of the elements of $\| a_{ri} \|$. When $\| a_{ri} \|$ is in the diagonal form, so is also $\| J \|$ with elements $a_{11}a_{11}, a_{11}a_{22}, \cdots, a_{11}a_{kk}$, $a_{22}a_{22}, a_{22}a_{33}, \cdots, a_{kk}a_{kk}$. In this case we have $J = A^{k+1}$, where A stands for the determinant of $\| a_{ri} \|$ which is nonzero on account of the nonsingularity of the transformation considered. It is then natural to ask the question whether it can be asserted that the same relationship holds even when the matrix $\| a_{ri} \|$ is not in the diagonal form or reducible to it. The answer is in the affirmative, the result being a particular case of the following theorem of Escherich (see C. C. Macduffee, *Theory of Matrices*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Julius Springer, Berlin, 1933, p. 86, theorem 44.3): *The determinant of the m-th power-matrix*[1] *with a nonvanishing determinant A is* $A^{(m+k-1)C_k}$, *where*

[1] It may be noted that the m-th power-matrix mentioned in Escherich's theorem is not the matrix multiplied by itself by the ordinary matrix multiplication. For its definition see Macduffee, loc. cit., p. 85.

k is the order of the square matrix whose determinant is A. Here $\| J \|$ is the second power-matrix, and so we have in all cases $J = A^{k+1}$.

Due to the importance of this in connection with Wishart's distribution, the following observations are of special interest.

(1) When the latent roots of the matrix $\| a_{ri} \|$ are all distinct, it is reducible to the diagonal form, and it has been shown by Rasch that in such a case $J = A^{k+1}$, and the method of analytic continuation has been suggested as a means of establishing the same result when not all the latent roots are distinct. We shall show this here by a consideration of the limit of a polynomial function.

Let us suppose that some of the latent roots are repeated. Let now a suitable number of the elements of the matrix be replaced by neighbouring values, i.e., by $a_{ri} + \epsilon\eta_{ri}$, so that the latent roots of the altered matrix, $\| a'_{ri} \|$ say, are all distinct. (This is always possible as there are an adequate or more than adequate number of elements for this purpose.) We shall now consider a linear transformation from the variables (x) to the variables (y) with $\| a'_{ri} \|$ as the transformation matrix. Using primes to denote the corresponding quantities in relation to the new transformation of the variables, we have $J' = (A')^{k+1}$, since $\| a'_{ri} \|$ is reducible to the diagonal form. Further J', expressed as an expansion in its elements, is a polynomial function whose continuity properties yield by proceeding to the limit

$$\begin{aligned} J = \lim_{\epsilon \to 0} J' &= \lim_{\epsilon \to 0} (A')^{k+1} \\ &= (\lim_{\epsilon \to 0} A')^{k+1} \\ &= A^{k+1}. \end{aligned}$$

(2) The following situation is sometimes met with. Consequent on the linearity of the transformation of the variables with constant coefficients we have

$$dx_r = \sum_{i=1}^{k} a_{ri}\, dy_i \qquad (r = 1, 2, \cdots k)$$

which gives

$$d(x_i^{(r)} x_j^{(r)}) = \sum_{l=1}^{k} \sum_{m=1}^{k} a_{il}\, a_{jm}\, d(y_l^{(r)} y_m^{(r)}).$$

Summing over the n values of r, we have dM_{ij} transformed in terms of (dM'_{lm}) with the same coefficients as those that appear in the transformation of $dx_i\, dx_j$ in terms of $(dy_l\, dy_m)$. Also we know that

$$\begin{aligned} &dM_{11}\, dM_{12} \cdots dM_{1k}\, dM_{22}\, dM_{23} \cdots dM_{kk} \\ &\qquad = J \cdot dM'_{11}\, dM'_{12} \cdots dM'_{1k}\, dM'_{22}\, dM'_{23} \cdots dM'_{kk}. \end{aligned}$$

By the sameness of coefficients of transformation we can write

$$\begin{aligned} &dx_1\, dx_1\, dx_1\, dx_2 \cdots dx_1\, dx_k\, dx_2\, dx_2\, dx_2\, dx_3 \cdots dx_k\, dx_k \\ &\qquad = J \cdot dy_1\, dy_1\, dy_1\, dy_2 \cdots dy_1\, dy_k\, dy_2\, dy_2\, dy_2\, dy_3 \cdots dy_k\, dy_k\,; \end{aligned}$$

that is,

$$(dx_1 dx_2 \cdots dx_k)^{k+1} = J \cdot (dy_1 dy_2 \cdots dy_k)^{k+1}.$$

But $dx_1 dx_2 \cdots dx_k = A \cdot dy_1 dy_2 \cdots dy_k$, so that $J = A^{k+1}$. These formal operations show that in this case the differentials when multiplied in the usual way work like the determinants they signify.

(3) After obtaining that $J = A^{k+1}$ Rasch's functional equation is seen to hold good.

(4) When the constant in Wishart's distribution is evaluated in Rasch's notation using H. Cramér's method (see *Mathematical Methods of Statistics*, Princeton University Press, 1946, pp. 390–393), it will be found that a power of n is missing in the numerator. This is due to the fact that we have not considered the estimate $1/n \, || M_{ij} ||$ but worked with $|| M_{ij} ||$.

I thank Professor H. Cramér and Mr. G. Blom for their interest in this work.

A NOTE ON A TWO SAMPLE TEST[1]

By Frank J. Massey, Jr.

University of Oregon

1. Summary. Mood ([1], p. 394) discusses a test for the hypothesis that two samples come from populations having the same continuous cumulative frequency distributions. It consists of arranging the observations from the two samples in a single group in order of size and then comparing the numbers in the two samples above the median. This technique is extended to using several order statistics from the combined samples, and to the case of several samples. The test is non-parametric and might be a good substitute for the single variable of classification analysis of variance in cases of doubtful normality. The application of the test would be the same as in Mood ([1], p. 398) except that there would be more than two rows in the table.

2. The distribution function. Suppose we have p populations all having the same continuous cumulative distribution function $F(x)$. Let X_{ij} $(i = 1, 2 \cdots, p; j = 1, 2 \cdots, n_i)$ be the jth observation in a sample of size n_i from the ith population. Let $\sum_{i=1}^{p} n_i = N$.

Arrange these N observations in a single series according to size and rename them $z_1 \leq z_2 \leq \cdots \leq z_N$. We choose $k - 1$ of the z values, for example, $z_{\alpha_1}, z_{\alpha_2}, \cdots, z_{\alpha_{k-1}}$ (the α_i are integers and $1 \leq \alpha_1 < \alpha_2 < \cdots \leq N$). Denote by m_{ij} the number of observations X_{ih} such that $z_{\alpha_{j-1}} < X_{ih} \leq z_{\alpha_j}$ for $j = 2, 3, \cdots, k-1$, by m_{i0} the number of $X_{ih} \leq z_{\alpha_1}$, and by m_{ik} the number of $X_{ih} > z_{\alpha_{k-1}}$. These can be illustrated by the following table.

[1] This paper sponsored in part by the Office of Naval Research, Contract N6-onr-218/IV, Project NR 042 063.