NOTES

This section is devoted to brief research and expository articles and other short items.

A FUNCTIONAL EQUATION FOR WISHART'S DISTRIBUTION

By G. Rasch

State Serum Institute and University of Copenhagen

1. Introduction. The sampling distribution of the moment matrix for observations from a multivariate normal distribution was given by Wishart in 1928 [1]. This proof involved rather advanced multidimensional geometry but since then two analytical proofs have been given: one by Wishart and Bartlett in cooperation with Ingham by the use of the characteristic function [2] and a second by Hsu by induction with regard to the dimension of the observations, [3],

In the following section is given a new derivation of the form of Wishart's distribution in which a fundamental property of the multivariate normal distribution is utilized, *viz.* the invariance of the distribution type against a linear transformation. In section 3 the same principle is used for evaluation of the constant and determination of the moment matrix in the multidimensional normal distribution.

2. Derivation of Wishart's distribution. Let[1]

$$\mathbf{x} = (x_1, \cdots, x_k), \tag{1}$$

denote a k-dimensional normal variate with the mean vector 0 and the distribution matrix

$$\mathbf{\Phi} = (\varphi_{ij}), \tag{2}$$

viz.

$$p\{\mathbf{x}\} = \frac{\sqrt{\Delta(\mathbf{\Phi})}}{(\sqrt{2\pi})^k} \cdot e^{-\frac{1}{2}\mathbf{x}\mathbf{\Phi}\mathbf{x}^*}. \tag{3}$$

$\mathbf{\Phi}$ is symmetrical and positive definite.

Now consider n observations of $\mathbf{x}$: $\mathbf{x}_1, \cdots, \mathbf{x}_n$, which are stochastically independent. Their joint distribution is

$$p\{\mathbf{x}_1, \cdots \mathbf{x}_n\} = \left(\frac{\sqrt{\Delta(\mathbf{\Phi})}}{(\sqrt{2\pi})^k}\right)^n \cdot e^{-\frac{1}{2}\Sigma\mathbf{x}_\nu\mathbf{\Phi}\mathbf{x}_\nu^*}. \tag{4}$$

The estimation of $\mathbf{\Phi}$ is based upon the moment sums

$$m_{ij} = \Sigma x_{\nu i} x_{\nu j},$$

[1] Notations: Lower case latin and greek letters are scalars; boldface capital latin and greek letters denote matrices, and boldface lower case letters row vectors. * means transposition. Δ (**A**) stands for the determinant of the square matrix **A**.

which form the symmetrical, positive definite matrix

(5) $$\mathbf{M} = (m_{ij}) = \Sigma \mathbf{x}_\nu^* \mathbf{x}_\nu .$$

In order to derive the distribution of $\mathbf{M}$ the straightforward procedure seems to be to transform the distribution of the sample $(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ to a distribution of $\mathbf{M}$ and some other variables which then should be integrated away. As such, the transformation,

(6) $$\mathbf{x}_\nu = \mathbf{u}_\nu \mathbf{M}^{\frac{1}{2}}, \qquad \mathbf{M}\Sigma \mathbf{u}_\nu^* \mathbf{u}_\nu = 1,$$

might serve. The matrix

(7) $$\mathbf{U} = \begin{pmatrix} \mathbf{u}_1 \\ \cdots \\ \mathbf{u}_n \end{pmatrix}$$

contains nk elements linked together with $\frac{(k+1)k}{2}$ relations; $(\mathbf{U})$ symbolizes $\left(n - \frac{k+1}{2}\right) k$ of the elements taken as independent variables.

For the purpose of introducing $\mathbf{M}$ in the exponential term in (4) we shall define the "double dot multiplication" of two matrices:

(8) $$\mathbf{A} \cdot\cdot \mathbf{B} = (a_{ij}) \cdot\cdot (b_{ij}) = \sum_{(i)} \sum_{(j)} a_{ij} b_{ij},$$

for which we notice the rule

(9) $$\mathbf{A} \cdot\cdot (\mathbf{BCD}) = \mathbf{C} \cdot\cdot (\mathbf{B}^* \mathbf{A} \mathbf{D}^*).$$

As obviously

$$\mathbf{x}\Phi\mathbf{x}^* = \Sigma \varphi_{ij} x_i x_j = \Phi \cdot\cdot (\mathbf{x}^*\mathbf{x}),$$

we have

(10) $$\Sigma \mathbf{x}_\nu \Phi \mathbf{x}_\nu^* = \Phi \cdot\cdot \mathbf{M},$$

and accordingly

(11) $$p\{\mathbf{M}, (\mathbf{U})\} = \left(\frac{\sqrt{\Delta(\Phi)}}{(\sqrt{2\pi})^k}\right)^n \cdot e^{-\frac{1}{2}\Phi\cdot\cdot\mathbf{M}} \left| \frac{\partial(\mathbf{x}_1, \cdots, \mathbf{x}_n)}{\partial(\mathbf{M}, (\mathbf{U}))} \right|,$$

where $\frac{\partial(\ )}{\partial(\ )}$ denotes the jacobian of the transformation. On integrating with respect to $(\mathbf{U})$ we obtain

(12) $$p\{\mathbf{M}\} = (\sqrt{\Delta(\Phi)})^n \cdot e^{-\frac{1}{2}\Phi\cdot\cdot\mathbf{M}} \cdot \varphi(\mathbf{M}),$$

where $\varphi(\mathbf{M})$ is independent of Φ. From this it follows that $p\{\mathbf{x}_1, \cdots, \mathbf{x}_n \mid \mathbf{M}\}$ is independent of Φ, i.e. $\mathbf{M}$ is a sufficient statistic for Φ.

In order to determine the mathematical form of $\varphi(\mathbf{M})$ we shall apply an arbitrary linear transformation to the original variates:

(13) $$\mathbf{x}_\nu = \mathbf{x}_\nu' \mathbf{A}.$$

The new variates $\mathbf{x}'_\nu$ are obviously normally distributed about 0 with the distribution matrix

$$\Phi' = \mathbf{A}\Phi\mathbf{A}^*. \tag{14}$$

Therefore the distribution function of the new moment matrix, given by

$$\mathbf{M} = \mathbf{A}^*\mathbf{M}'\mathbf{A}, \tag{15}$$

is

$$p\{\mathbf{M}'\} = (\sqrt{\Delta(\Phi')})^n \cdot e^{-\frac{1}{2}\Phi' \cdot\cdot \mathbf{M}'} \varphi(\mathbf{M}'). \tag{16}$$

On the other hand the transformation from $\mathbf{M}$ to $\mathbf{M}'$ is a linear one, the jacobian of which therefore is a constant depending on $\mathbf{A}$ only:

$$\frac{\partial(\mathbf{M})}{\partial(\mathbf{M}')} = \psi(\mathbf{A}), \quad \text{say.} \tag{17}$$

Consequently,

$$p\{\mathbf{M}'\} = \sqrt{\Delta(\Phi)} \cdot e^{-\frac{1}{2}\Phi \cdot\cdot \mathbf{M}} \cdot \varphi(\mathbf{M}) \cdot |\, \psi(\mathbf{A}) \,|. \tag{18}$$

The two expressions for $p\{\mathbf{M}'\}$ must be identical, and as

$$\Delta(\Phi') = \Delta(\Phi)\Delta^2(\mathbf{A}), \tag{19}$$

and

$$\Phi' \cdot\cdot \mathbf{M}' = (\mathbf{A}\Phi\mathbf{A}^*) \cdot\cdot \mathbf{M}' = (\mathbf{A}^*\mathbf{M}'\mathbf{A}) \cdot\cdot \Phi = \mathbf{M} \cdot\cdot \Phi, \tag{20}$$

it follows that $\varphi(\mathbf{M})$ satisfies the functional equation

$$|\, \Delta(\mathbf{A}) \,|^n \varphi(\mathbf{M}') = \varphi(\mathbf{M}) \cdot |\, \psi(\mathbf{A}) \,|. \tag{21}$$

Now, since the transformation $\mathbf{M} = (\mathbf{AB})^* \mathbf{M}'(\mathbf{AB})$ may be carried out in two steps, $\psi(\mathbf{A})$ also satisfies a functional equation

$$\psi(\mathbf{AB}) = \psi(\mathbf{A})\psi(\mathbf{B}). \tag{22}$$

Furthermore, if $\mathbf{A}$ is a diagonal matrix it is easily seen that

$$\psi(\mathbf{A}) = (\Delta(\mathbf{A}))^{k+1}, \tag{23}$$

and this relation holds generally. In fact, considering the case where the normal form of $\mathbf{A}$ is a diagonal matrix:

$$\mathbf{A} = \mathbf{TDT}^{-1}, \text{ say,}$$

we get

$$\begin{aligned} \psi(\mathbf{A}) &= \psi(\mathbf{T})\psi(\mathbf{D})\psi(\mathbf{T}^{-1}) \\ &= (\Delta(\mathbf{D}))^{k+1} \psi(\mathbf{TT}^{-1}) \\ &= (\Delta(\mathbf{A}))^{k+1}, \end{aligned}$$

and by analytical continuation this is seen to be true for any $\mathbf{A}$.

Now, inserting this result in the functional equation (21) and taking for **A** the real symmetrical square root of **M** so that[2] $\mathbf{M}' = 1$, we readily obtain the solution

$$\varphi(\mathbf{M}) = (\Delta(\mathbf{M}^{\frac{1}{2}}))^{n-k-1} \cdot \varphi(1). \tag{24}$$

It follows that

$$p\{\mathbf{M}\} = \gamma_k(n)(\Delta(\Phi))^{n/2} \cdot e^{-\frac{1}{2}\Phi \cdot \cdot \mathbf{M}} \cdot (\Delta(\mathbf{M}))^{(n-k-1)/2}, \tag{25}$$

where $\gamma_k(n) = \varphi(1)$ is a constant which may be determined in various ways (cf. for instance Cramér [4]).

3. Other applications of the linear transformation. It may be noticed that the linear transformation also leads to simple derivations of two fundamental properties of the normal multivariate distribution itself, *viz.* determination of the constant and the relation between the moment matrix and the distribution matrix.

Let

$$p\{\mathbf{x}\} = \gamma(\Phi) \cdot e^{-\frac{1}{2}\mathbf{x}\Phi\mathbf{x}^*}, \tag{26}$$

and transform by

$$\mathbf{x} = \mathbf{x}'\mathbf{A}. \tag{27}$$

The new variable obviously has the distribution matrix (14) and the constant $\gamma(\Phi')$. But on the other hand direct transformation of (26) leads to

$$P\{\mathbf{x}'\} = \gamma(\Phi) \cdot e^{-\frac{1}{2}\mathbf{x}\Phi\mathbf{x}^*} \cdot \left| \frac{\partial(\mathbf{x})}{\partial(\mathbf{x}')} \right|$$
$$= \gamma(\Phi) \mid \Delta(\mathbf{A}) \mid e^{-\frac{1}{2}\mathbf{x}\Phi\mathbf{x}^*},$$

and therefore we must have

$$\gamma(\Phi') = \gamma(\Phi) \mid \Delta(\mathbf{A}) \mid .$$

For $\mathbf{A} = \Phi^{-\frac{1}{2}}$ we get $\Phi' = 1$ and consequently

$$\gamma(\Phi) = \sqrt{\Delta(\Phi)} \cdot \gamma(1),$$

where obviously

$$\gamma(1) = \frac{1}{(\sqrt{2\pi})^n}.$$

Considering

$$\mathbf{M}(\Phi) = \int \mathbf{x}^* \mathbf{x} p\{\mathbf{X}\}\, d\mathbf{x},$$

---

[2] Exists because M is positive definite: Let $\mathrm{M} = \mathrm{ODO}^*$ where O is orthogonal and **D** the diagonal form of M; then $\mathrm{M}^{\frac{1}{2}} = \mathrm{O}D^{\frac{1}{2}}\mathrm{O}^*$ is real and symmetrical.

the same transformation gives

$$\mathbf{M}(\Phi) = \int \mathbf{A}^* \mathbf{x}^* \mathbf{x} \mathbf{A} p\{\mathbf{x}'\}\, d\mathbf{x}',$$

$$= \mathbf{A}^* \mathbf{M}(\Phi')\mathbf{A}$$

which for $A = \Phi^{-\frac{1}{2}}$ leaves us with

$$\mathbf{M}(\Phi) = (\Phi')^{-1}$$

because $\mathbf{M}(1) = 1$.

REFERENCES

[1] J. Wishart, "The generalised product moment distribution in samples from a normal multivariate population", *Biometrika*, Vol. 20A (1928), p. 32.
[2] J. Wishart and M. S. Bartlett, "The generalised product moment distribution in a normal system", *Proc. Cambr. Phil. Soc.*, Vol. 29 (1933), p. 260.
[3] P. L. Hsu, "A new proof of the joint product moment distribution", *Proc. Cambr. Phil. Soc.*, Vol. 35 (1939), p. 336.
[4] H. Cramér, *Mathematical Methods in Statistics*, Princeton Univ. Press, 1946, pp. 392–93.

THE DISTRIBUTION OF A DEFINITE QUADRATIC FORM

By Herbert Robbins

University of North Carolina

Let $X_1, \cdots, X_n$ be independent normal variates with zero means and unit variances, let $a_1, \cdots, a_n$ be positive constants and define

$$U_n = \frac{a_1}{2} X_1^2 + \cdots + \frac{a_n}{2} X_n^2, \tag{1}$$

$$F_n(x) = \Pr[U_n \leq x], \qquad f_n(x) = F_n'(x). \tag{2}$$

Setting

$$a = (a_1 \cdots a_n)^{1/n} \tag{3}$$

and using the convolution formula we may show by induction that for $x > 0$,

$$f_n(x) = a^{-\frac{1}{2}n} x^{\frac{1}{2}n-1} \sum_{k=0}^{\infty} \frac{c_k(-x)^k}{\Gamma(\frac{1}{2}n + k)}, \tag{4}$$

$$F_n(x) = a^{-\frac{1}{2}n} x^{\frac{1}{2}n} \sum_{k=0}^{\infty} \frac{c_k(-x)^k}{\Gamma(\frac{1}{2}n + k + 1)}, \tag{5}$$

where for $k = 0, 1, \cdots$

$$c_k = \pi^{-\frac{1}{2}n} \sum_{i_1+\cdots+i_n=k} \frac{\Gamma(i_1 + \frac{1}{2}) \cdots \Gamma(i_n + \frac{1}{2})}{i_1! \cdots i_n! \, a_1^{i_1} \cdots a_n^{i_n}} > 0. \tag{6}$$