ON THE ADMISSIBILITY OF SOME TESTS OF MANOVA

By M. N. Ghosh

Institute of Agricultural Research Statistics, New Delhi

1. Summary. Using Stein's [7] generalisation of a method of Lehmann and A. Birnbaum (A. Birnbaum [2]), we show that two of the tests commonly used in multivariate analysis of variance are admissible.

2. Introduction. In the canonical form of the multivariate linear hypothesis we observe independently normally distributed p-dimensional row vectors $z^{(1)}, \cdots z^{(Q+m)}, y^{(1)}, \cdots y^{(n)}$ $(n \geqq p)$ with common non-singular covariance matrix Σ_1 and means

$$E(z^{(r)}) = b^{(r)}, \qquad E(y^{(s)}) = 0 \qquad r = 1, \cdots Q + m, s = 1, \cdots, n \tag{2.1}$$

where $b^{(r)}$ are unknown p-vectors and we want to test the hypothesis

$$H : b^{(r)} = 0 \quad \text{for} \quad r = 1, \cdots Q. \tag{2.2}$$

Several tests have been suggested for this hypothesis depending on the characteristic roots of the matrix ZY^{-1} where $Z = \sum_{q=1}^{Q} z^{(q)\prime} z^{(q)}$ and $Y = \sum_{u=1}^{n} y^{(u)\prime} y^{(u)}$, i.e. roots $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_{\min(Q,p)}$ of the equation

$$|Z - \lambda Y| = \left| \sum_{q=1}^{Q} z^{(q)\prime} z^{(q)} - \lambda \sum_{u=1}^{n} y^{(u)\prime} y^{(u)} \right| = 0. \tag{2.3}$$

The following statistics which reject the hypothesis H for large values are commonly used for the test of H,

(i) the sum of roots, $\sum \lambda_i = \operatorname{Tr}(ZY^{-1})$ suggested by Hotelling [3] and Lawley [4],

(ii) the largest root, suggested by Roy [5] and

(iii) $\prod(1 + \lambda_i) = |Y + Z|/|Y|$ suggested by Wilks [8] and Bartlett [1].

When $Q = 1$, all these criteria reduce to Hotelling's T^2-test which Stein [7] proved admissible against unrestricted alternatives. His proof is based on a generalisation of a result of Lehmann and A. Birnbaum [2]. We shall use Stein's method to prove the admissibility of (i) and (ii) against unrestricted alternatives. The proof of admissibility given here has the same limitations as the corresponding proof of the admissibility of Hotelling's T^2 given by Stein [7], e.g., the superiority of the above tests is established for some large values of the parameters in the alternative hypotheses and would be useless if one restricts the domain of alternatives.

Stein considers random variables $(x_1, \cdots x_n)$ which belong to the n-dimensional linear space $\mathfrak{X}$ and defines an exponential family of distributions.

$$P_\xi(x) = \psi(\xi)^{-1} e^{\xi x} d\mu(x) \tag{2.4}$$

Received 17 June 1963; revised 1 November 1963.

where μ is a measure on the σ-algebra B of all ordinary Borel sets of $\mathfrak{X}$, and ξ is an element of the adjoint space $\mathfrak{X}'$ of $\mathfrak{X}$ corresponding to different values of the parameters. Let $\tilde{\Theta}$ denote the set of $\xi \varepsilon \mathfrak{X}'$ for which the integral of $P_\xi(x)$ is defined and is equal to unity and let $\theta \varepsilon \Theta \subset \tilde{\Theta}$ be the parameter space in the problem under consideration. The hypothesis to be tested is given by the points of $\Theta_0 \subset \Theta$ and the alternatives by points of $\Theta - \Theta_0$. Stein proves the following theorem.

THEOREM. *Let* $(\mathfrak{X}, \mu, \Theta, P)$ *be an exponential family and* Θ_0 *a nonempty proper subset of* Θ. *Let* A *be a closed convex subset of* $\mathfrak{X}$ *such that for every* $\xi \varepsilon \mathfrak{X}'$ *and a real* c *for which*

$$\{x : \xi x > c\} \cap A = 0 \qquad (\textit{the null set}) \tag{2.5}$$

there exists $\theta_1 \varepsilon \tilde{\Theta}$ *such that there exist arbitrarily large* λ *for which* $\theta_1 + \lambda\xi \varepsilon \Theta - \Theta_0$, *then the test* ϕ_0 *given by*

$$\begin{aligned} \phi_0(x) &= 0 \quad \text{if} \quad x \varepsilon A \\ &= 1 \quad \text{if} \quad x \not\varepsilon A \end{aligned} \tag{2.6}$$

is admissible for testing the hypothesis that a random element x *of* $\mathfrak{X}$ *is distributed according to some* P_θ *with* $\theta \varepsilon \Theta_0$ *against the alternatives* $\theta \varepsilon \Theta - \Theta_0$.

The distribution of the variables $z^{(r)}, y^{(u)}$ can be expressed as

$$\begin{aligned} \frac{\exp\{-\frac{1}{2}\sum b^{(r)}\Sigma_1^{-1}b^{(r)\prime}\}}{(2\pi)^{\frac{1}{2}(Q+m+n)p}|\Sigma_1|^{\frac{1}{2}(Q+m+n)}} \times \exp\left\{-\frac{1}{2}\operatorname{Tr}(\Sigma_1^{-1}S) + \sum_{r=1}^{Q+m} b^{(r)}\Sigma_1^{-1}z^{(r)\prime}\right\} \\ = \psi(\xi)^{-1}e^{\xi x}\,d\mu(x) \end{aligned} \tag{2.7}$$

where x is a vector variable with $\frac{1}{2}p(p+1) + (Q+m)p$, components corresponding to the $\frac{1}{2}p(p+1)$ components of the matrix $S = \sum_{r=1}^{Q+m} z^{(r)\prime}z^{(r)} + \sum_{u=1}^{n} y^{(u)\prime}y^{(u)}$ and the components of the vectors $z^{(1)}, \cdots z^{(Q+m)}$ and ξ is a vector with $\frac{1}{2}p(p+1) + (Q+m)p$ components corresponding to the $\frac{1}{2}p(p+1)$ components of Σ_1^{-1} and $(Q+m)p$ components of $b^{(1)}\Sigma_1^{-1}, \cdots b^{(Q+m)}\Sigma_1^{-1}$, i.e.

$$x = (S, z^{(1)}, \cdots z^{(Q+m)}), \qquad \xi = (\Sigma_1^{-1}, b^{(1)}\Sigma_1^{-1}, \cdots b^{(Q+m)}\Sigma_1^{-1}) \tag{2.8}$$

and $\xi x = -\frac{1}{2}\operatorname{Tr}(\Sigma_1^{-1}S) + \sum_{r=1}^{Q+m} b^{(r)}\Sigma_1^{-1}z^{(r)\prime}$ and $\mu(x)$ is obtained from the joint distribution of S and $z^{(1)}, \cdots z^{(Q+m)}$, and is absolutely continuous with respect to the Lebesgue measure on $\mathfrak{X}$. All vectors considered below are rowvectors unless specified otherwise.

3. We shall use the following equivalence relation due to Roy [5] which is a consequence of the simultaneous diagonalisability of two quadratic forms in a suitable basis.

Let S_1 and S_2 be $p \times p$ symmetric positive definite matrices whose elements are functions of x and δ any row vector, then

$$\left\{x : \frac{\delta S_1 \delta'}{\delta S_2 \delta'} \leqq c \text{ for all } \delta\right\} \equiv \{x : \text{largest ch root of } S_1 S_2^{-1} \leqq c\}. \tag{3.1}$$

Here "≡" denotes that the two regions are identical. In the special case of this when the matrix $S_1 = y'y$, y' being a p-vector, $S_1S_2^{-1} = y'yS_2^{-1}$ has only one characteristic root $yS_2^{-1}y' = \text{Tr}(y'yS_2^{-1})$ so that (3.1) becomes

$$\{x : (y\delta')^2/\delta S_2\delta' \leqq c \text{ for all } \delta\} \equiv \{x : yS_2^{-1}y' \leqq c\} \tag{3.2}$$

which was given by Scheffé [6].

When $S_2 = I$, (3.2) reduces to

$$\{x : (y\delta')^2 \leqq c\delta\delta' \text{ for all } \delta\} \equiv \{x : \textstyle\sum y_i^2 \leqq c\}. \tag{3.2$'$}$$

For the case of $Q = 1$, Stein has shown that the region

$$\left\{(S, z^{(1)}, \cdots z^{(m+1)}) : S - \sum_{r=2}^{m+1} z^{(r)'}z^{(r)} \text{ positive definite} \right. \\ \left. \text{and } z^{(1)}\left(S - \sum_{r=2}^{m+1} z^{(r)'}z^{(r)}\right)^{-1} z^{(1)'} \leqq c\right\} \tag{3.3}$$

is μ-equivalent (differ in a set of zero μ-measure, i.e. Lebesgue measure) to the region contained in the intersection of all half-spaces of the form

$$\left\{(S, z^{(1)}, \cdots z^{(m+1)}) : lz^{(1)'} + \sum_{r=2}^{m+1} k^{(r)}lz^{(r)'} \right. \\ \left. - \frac{1}{2}\text{Tr}\,(l'lS) \leqq \frac{1}{2}\left[c + \sum_{r=2}^{m+1}(k^{(r)})^2\right]\right\} \\ \left\{(S, z^{(1)}, \cdots z^{(m+1)}) : \sum_{r=2}^{m+1} k^{(r)}lz^{(r)'} - \frac{1}{2}\text{Tr}\,(l'lS) \leqq \frac{1}{2}\sum_{r=2}^{m+1}(k^{(r)})^2\right\} \tag{3.4}$$

where l ranges over all p-vectors different from zero and $k^{(2)}, \cdots k^{(m+1)}$ range over the real line. We shall extend this result to regions in the $\mathfrak{X}$-space defined by

$$\{x : \text{Tr}(ZY^{-1}) \leqq c\} \quad \text{and} \quad \{x : \lambda_1 \leqq c\}.$$

Consider the pQ-vectors

$$\tilde{z} = (z_1^{(1)}, \cdots z_p^{(1)}, \cdots z_1^{(Q)}, \cdots z_p^{(Q)}) \equiv (z^{(1)}, \cdots z^{(Q)})$$
$$l = (l_1^{(1)}, \cdots l_p^{(1)}, \cdots l_1^{(Q)}, \cdots l_p^{(Q)}) \equiv (l^{(1)}, \cdots l^{(Q)})$$

and $\tilde{U} = I_Q \times U$, where I_Q is the unit matrix in Q dimensions, U any symmetric and positive definite $p \times p$ matrix and $I_Q \times U$ denotes the Kronecker product of I_Q with U. Consider also the regions

$$A_1(\tilde{U}) = \{x : (\tilde{z}l')^2/l\tilde{U}l' \leqq c \text{ for all } l\}, \\ A_2(\tilde{U}) = \{x : (\tilde{z}l')^2/l\tilde{U}l' \leqq c \text{ for all } l \text{ with } l_i^{(q)} = l_q\delta_i\}. \tag{3.5}$$

We shall denote the pQ-dimensional vector space $\{l\}$ by V_1 and the vector space formed by vectors of the type $l_i^{(q)} = l_q\delta_i$ by V_2. We shall show that the regions A_1 and A_2 can be generated by half-spaces similar to (3.4). Since $(cl\tilde{U}l')^{\frac{1}{2}} \leqq \frac{1}{2}(c + l\tilde{U}l')$ for all $l \in V_1$

$$A_i(\tilde{U}) \subset \{x : |\bar{z}l'| \leqq \tfrac{1}{2}(c + l\tilde{U}l') \text{ for all } l \,\varepsilon\, V_i\} = B_i(\tilde{U}).$$

Since for any $l \,\varepsilon\, V_i$, $-l$ also belongs to V_i, we need not take the absolute value of $\bar{z}l'$ in defining $B_i(\tilde{U})$. For any given l and $\tilde{U}$, if $\bar{z}_1 \,\varepsilon\, B_i(\tilde{U})$, then $|\bar{z}_1 l'/\lambda| \leqq \tfrac{1}{2}\{c + l\tilde{U}l'/\lambda^2\}$ holds for all λ. Putting $\lambda = (l\tilde{U}l'/c)^{\frac{1}{2}}$ we get

$$(\bar{z}_1 l'/\lambda)^2 \leqq c^2 = cl\tilde{U}l'/\lambda^2,$$

i.e. $(\bar{z}_1 l')^2 \leqq cl\tilde{U}l'$ for all $l \,\varepsilon\, V_i$. Thus $\bar{z}_1 \,\varepsilon\, A_i(\tilde{U})$ for given $\tilde{U}$, so that generally

$$A_i(\tilde{U}) \equiv B_i(\tilde{U}). \tag{3.6}$$

Let $S_1 = S - \sum_{r=1}^{Q+m} z^{(r)\prime} z^{(r)} = Y$ and $\tilde{S}_1 = I_Q \times Y$. Then from (3.6)

$$\begin{aligned} A_i(\tilde{S}_1) &= \{x : (\bar{z}l')^2/l\tilde{S}_1 l' \leqq c \text{ for all } l \,\varepsilon\, V_i\} \\ &\equiv \left\{x : \frac{1}{2}\, l\left(I_Q \times \sum_{r=1}^{Q+m} z^{(r)\prime} z^{(r)}\right) l' \leqq \frac{1}{2}\, c - \bar{z}l' \right. \\ &\qquad\qquad \left. + \frac{1}{2}\, l(I_Q \times S)l' \text{ for all } \varepsilon V_i \right\}. \end{aligned} \tag{3.7}$$

Now

$$\begin{aligned} & l\left(I_Q \times \sum_{r=1}^{Q+m} z^{(r)\prime} z^{(r)}\right) l' \\ &\quad = l\left(\sum_{r=1}^{Q+m} (I_Q \times z^{(r)\prime})(I_Q \times z^{(r)\prime})'\right) l' \\ &\quad = \sum_{r=1}^{Q+m} (l\cdot I_Q \times z^{(r)\prime})(l\cdot I_Q \times z^{(r)\prime})' \\ &\quad = \sum_{r=1}^{Q+m} \sum_{q=1}^{Q} (l^{(q)} z^{(r)\prime})^2 \quad \text{since} \quad l\cdot I_Q \times z^{(r)\prime} = (l^{(1)} z^{(r)\prime}, \cdots l^{(Q)}\cdot z^{(r)\prime}). \end{aligned} \tag{3.8}$$

From $(3.2)'$ we get

$$\begin{aligned} & \left\{x : \sum_{r=1}^{Q+m} \sum_{q=1}^{Q} (l^{(q)}\cdot z^{(r)\prime})^2 \leqq 2\alpha \right\} \\ &\quad \equiv \left\{x : \left[\sum_{r=1}^{Q+m} \sum_{q=1}^{Q} k_q^{(r)} l^{(q)} z^{(r)\prime}\right]^2 \leqq 2\alpha \sum_{r=1}^{Q+m} \sum_{q=1}^{Q} (k_q^{(r)})^2 \text{ for all } k_q^{(r)} \right\}. \end{aligned} \tag{3.9}$$

From (3.6), (3.8) and (3.9)

$$\begin{aligned} & \left\{x : l\left(I_Q \times \sum_{r=1}^{Q+m} z^{(r)\prime} z^{(r)}\right) l' \leqq 2\alpha \right\} \\ &\quad \equiv \left\{x : \sum_{r=1}^{Q+m} \sum_{q=1}^{Q} k_q^{(r)} l^{(q)} z^{(r)\prime} - \tfrac{1}{2} \sum_{r=1}^{Q+m} \sum_{q=1}^{Q} (k_q^{(r)})^2 \leqq \alpha \text{ for all } k_q^{(r)} \right\}. \end{aligned} \tag{3.10}$$

From (3.7) and (3.10) we get

$$(3.11)\quad \begin{aligned} A_i(\tilde{S}_1) \equiv \Big\{x : \tilde{z}l' + \sum_{r=1}^{Q+m} \sum_{q=1}^{Q} k_q^{(r)} l^{(q)} z^{(r)} & - \tfrac{1}{2} \operatorname{Tr}(LS) \\ & \leqq \left[c + \sum_{r=1}^{Q+m} \sum_{q=1}^{Q} (k_q^{(r)})^2 \right] \text{for all } l \in V_i \text{ and } k_q^{(r)} \Big\} \end{aligned}$$

where we have used the relation

$$l(I_Q \times S)l' = \operatorname{Tr}\left[\sum_{q=1}^{Q} l^{(q)\prime} l^{(q)} S\right] = \operatorname{Tr}(LS) \quad \text{where} \quad L = \sum_{q=1}^{Q} l^{(q)\prime} l^{(q)}.$$

We now introduce the additional restriction that $S - \sum_{r=1}^{Q+m} z^{(r)\prime} z^{(r)}$ should be positive definite. This is equivalent to $\delta(S - \sum_{r=1}^{Q+m} z^{(r)\prime} z^{(r)})\delta' > 0$ for all non-null p-vectors δ, i.e. the smallest characteristic root is positive.

Thus we get the condition

$$(3.12)\qquad \delta\left(\sum_{r=1}^{Q+m} z^{(r)\prime} z^{(r)}\right)\delta' < \delta S \delta'.$$

Analogous to (3.10) this is μ-equivalent to

$$(3.13)\qquad \left\{x : \sum_{r=1}^{Q+m} k^{(r)} \delta z^{(r)\prime} - \tfrac{1}{2} \operatorname{Tr}(\delta' \delta S) \leqq \tfrac{1}{2} \sum_{r=1}^{Q+m} (k^{(r)})^2 \right\}.$$

Thus the regions $A_i(\tilde{S}_1)$, $(i = 1, 2)$, with the additional constraint that S_1 is positive definite are μ-equivalent to the region contained in the intersection of half-spaces (3.11) and (3.13).

4. We shall now show that the regions A_1, A_2 respectively are the regions $\{x : \operatorname{Tr}(ZY^{-1}) \leqq c\}$ and $\{x : \lambda_1 \leqq c\}$. In fact

$$\tilde{z}\tilde{S}_1^{-1}\tilde{z}' = \sum_{q=1}^{Q} z^{(q)} y^{-1} z^{(q)\prime} = \operatorname{Tr}\left[\left(\sum_{q=1}^{Q} z^{(q)\prime} z^{(q)}\right) Y^{-1}\right] = \operatorname{Tr}(ZY^{-1})$$

and from (3.2)

$$(4.1)\quad \begin{aligned} A_1(\tilde{S}_1) = \{x : (\tilde{z}l')^2 / l\tilde{S}_1 l' \leqq c \text{ for } l \in V_1\} & \equiv \{x : \tilde{z}\tilde{S}_1^{-1}\tilde{z} \leqq c\} \\ & \equiv \{x : \operatorname{Tr}(ZY^{-1}) \leqq c\}. \end{aligned}$$

When $l \in V_2$

$$(4.2)\quad \begin{aligned} A_2(\tilde{S}_1) & = \left\{x : \frac{\left[\sum_{q=1}^{Q} l_q (z^{(q)} \delta')\right]^2}{\sum l_q^2 \, \delta Y \delta'} \leqq c \text{ for all } l \text{ and } \delta \right\} \\ & \equiv \left\{x : \sum_{q=1}^{Q} [z^{(q)} \cdot \delta']^2 \leqq c \delta Y \delta' \text{ for all } \delta \right\} \qquad \text{from } (3.2)' \\ & \equiv \left\{x : \delta\left(\sum_{q=1}^{Q} z^{(q)\prime} z^{(q)}\right)\delta' \leqq c\delta Y \delta' \text{ for all } \delta\right\} \\ & \equiv \{x : \lambda_1 \leqq c\} \qquad \text{from } (3.1). \end{aligned}$$

The defining half-spaces for both the regions $\{x : \mathrm{Tr}(ZY^{-1}) \leqq c\}$ and $\{x : \lambda_1 \leqq c\}$ with appropriate restrictions on $l_i^{(q)}$ can be written as

$$\left(L, l^{(1)} + \sum_{q=1}^{Q} k_q^{(1)} l^{(q)}, \cdots, l^{(Q)} + \sum_{q=1}^{Q} k_q^{(Q)} l^{(q)}, \sum_{q=1}^{Q} k_q^{(Q+1)} l^{(q)}, \cdots \sum_{q=1}^{Q} k_q^{(Q+m)} l^{(q)}\right) .$$

$$\tag{4.3} \cdot (S, Z^{(1)}, \cdots Z^{(Q)}, Z^{(Q+1)}, \cdots Z^{(Q+m)}) \leqq \tfrac{1}{2}\left[c + \sum_{r=1}^{Q+m} \sum_{q=1}^{Q} (k_q^{(r)})^2\right]$$

$$(\delta'\delta, k^{(1)}\delta, \cdots k^{(Q+m)}\delta) \cdot (S, Z^{(1)}, \cdots Z^{(Q+m)}) \leqq \tfrac{1}{2} \sum_{r=1}^{Q+m} (k^{(r)})^2.$$

If ξ is any point in $\mathfrak{X}'$ say $\xi = (\Gamma, \eta_1, \cdots \eta_{Q+m})$ for which the region $\{x : \xi x > c\}$ has no common points with either of these regions, it may be obtained as a limit of a finite number of positive linear combinations of halfspaces (4.3), so that Γ is positive semidefinite at least. We now choose $\theta_1 \,\varepsilon\, \Theta - \Theta_0$. Then the first component of $\theta_1 + \lambda\xi$ for $\lambda > 0$ is a positive definite matrix. If $\eta = 0$, the second component is non zero for all λ while if $\eta \neq 0$, it is non-zero for large values of λ. Thus $\theta_1 + \lambda\xi \,\varepsilon\, \Theta - \Theta_0$ for large λ and the conditions of Stein's theorem are satisfied. Thus the acceptance regions $\{x : \mathrm{Tr}(ZY^{-1}) \leqq c\}$ and $\{x : \lambda_1 \leqq c\}$ are admissible for the test of the hypothesis $b^{(q)} = 0$ $(q = 1, 2, \cdots Q)$, when the alternatives are unrestricted.

REFERENCES

[1] Bartlett, M. S. (1947). Multivariate analysis. *J. Roy. Statist. Soc. Ser. B* **9** 176–197.

[2] Birnbaum, A. (1955). Characterizations of complete classes of tests of some multiparametric hypotheses, with applications to likelihood ratio tests. *Ann. Math. Statist.* **26** 21–36.

[3] Hotelling, H. (1951). A generalised T test and measure of multivariate dispersion. *Proc. Second Berkeley Symp. Prob. Statist.* 23–41.

[4] Lawley, D. N. (1938). A generalisation of Fisher's Z-test. *Biometrika* **30** 180–187.

[5] Roy, S. N. (1958). *Some Aspects of Multivariate Analysis.* Asia Publications, Bombay.

[6] Scheffé, H. (1960). *Analysis of Variance*, (1st ed.). Wiley, New York.

[7] Stein, C. (1956). The admissibility of Hotelling's T^2-test. *Ann. Math. Statist.* **27** 616–623.

[8] Wilks, S. S. (1932). Certain generalisations in the analysis of variance. *Biometrika* **24** 471–494.