Sankhyā : The Indian Journal of Statistics
1981, Volume 43, Series B, Pt. 1, pp. 104–111

MONTE CARLO STUDIES FOR THE POWER OF A ONE SIDE TEST FOR HOMOGENEITY OF MEANS IN AN UNBALANCED ONE WAY MODEL

By R. PINCUS*

G.D.R. Academy of Science, Berlin

and

D. MAJUMDAR

Indian Statistical Institute

SUMMARY. For testing equality of treatment effects in an unbalanced one-way model against a restricted alternative explicit expressions for the likelihood ratio test and its null distribution are given. The gain of power compared with the F-test is demonstrated in a simultation experiment.

1. Introduction

Let us consider an unbalanced one way classification model, i.e. we have k normal populations with means $\xi_1, \xi_2, \ldots, \xi_k$, respectively and a common variance σ^2; x_{ij} denotes the j-th observation of the i-th population, $j = 1, 2, \ldots, n_i$, $n_1+n_2+\ldots+n_k = N$.

The problem is to test the hypothesis

$$H : \xi_1 = \xi_2 = \ldots = \xi_k$$

against the partially ordered alternative

$$K : \text{not } H,\ \xi_1 \leqslant \ldots \leqslant \xi_r\ (r \leqslant k). \qquad \ldots (1)$$

For such problems Bartholomew (1961) (see also Barlow *et al.*, 1972), developed the likelihood ratio criterion based on the statistic

$$\bar{E}^2 = \frac{\Sigma n_i(x_i^*-\bar{x})^2}{\Sigma\, n_i(\bar{x}_i-\bar{x})^2+(N-k)s^2}.$$

The x_i^* denote the 'amalgamized' means, or the projections of the $\bar{x}_i$'s on the cone defined by the restricted alternative (1). The null distribution of $\bar{E}^2$ is a mixture of beta distributions. However, except for balanced models the weights cannot be given in an analytical form when $r > 4$. An approximation for 'nearly balanced' models may be found in Siskind (1976).

* The paper was written during the stay of this author at the Indian Statistical Institute, New Delhi under the Indo-GDR agreement for scientific cooperation.

Evading these difficulties Pincus (1975) evaluated a modified version E_0^2 of the likelihood ratio test for $r = k$, based upon embedding the alternative region in a suitable circular cone, and gave its null distribution. In Section 2 those results are extended to the perhaps more realistic case $r \leqslant k$. Since the results are given for a cannonical form of the linear model, they are useful also in situations other than the one way models.

For the case $r = k$ there also exist other test statistics, e.g. Abelson and Tukey's (1963) 'Minimax linear contrast' or Schaafsma's (1966) 'Most stringent somewhere most powerful' test, both being t-distributed under the hypothesis, and with a fairly good power in the alternative region. If $r < k$ they are not applicable and a decomposition of the test problem so that these statistics could be used for testing the equality of the first r means, leads to complicated distributional problems.

In Section 3 we derive the explicit form of the E_0^2-statistic in the unbalanced one way model. Section 4 finally presents the outcome of Monte-Carlo studies for the behaviour of the power function of the E_0^2-test and compares it with the usual F-test.

2. The distribution of E_0^2

We assume a linear hypothesis to be given in a cannonical form. If (w, x, y, z) is a random variate with

$$\begin{aligned} w &\sim N_l(\eta, \sigma^2 I) \\ x &\sim N_p(\mu, \sigma^2 I) \\ y &\sim N_m(\nu, \sigma^2 I) \\ z &\sim N_n(0, \sigma^2 I) \end{aligned} \qquad \ldots \quad (2)$$

where w, x, y and z are independently distributed, η, μ, ν and σ^2 are unknown, we want to test the hypothesis

$$H\colon \mu = \nu = 0.$$

Further, we assume μ to lie in a circular cone $\mathcal{C}$ determined by its given axis c and angle ψ, $c \in R^p$, $0 \leqslant \psi \leqslant \frac{\pi}{2}$; i.e. the alternative reads

$$K : \text{not } H, \ \frac{c'\mu}{\|c\| \, \|\mu\|} \geqslant \cos\psi.$$

The likelihood ratio statistic

$$\lambda^{2/N} = \frac{\max\ (2\pi\sigma^2)^{-1} \exp\left\{ -\frac{1}{N\sigma^2}\ (\|w-\eta\|^2+\|x\|^2+\|y\|^2+\|z\|^2) \right\}}{\max\limits_{\mu\in\mathcal{C}}\ (2\pi\sigma^2)^{-1} \exp\left\{ -\frac{1}{N\sigma^2}\ (\|w-\eta\|^2+\|x-\mu\|^2+\|y-\nu\|^2+\|z\|^2) \right\}}$$

with $N = l+p+m+n$ equals

$$\lambda^{2/N} = \frac{(x-\mu_x)'(x-\mu_x)+z'z}{x'x+y'y+z'z}$$

where μ_x denotes the orthogonal projection of x onto $\mathcal{C}$, i.e.

$$\|x-\mu_x\|^2 = \min_{\mu\in\mathcal{C}}\ \|x-\mu\|^2.$$

The L. R. criterion rejects the hypothesis if λ is too small or equivalently, considering $\mu_x'(x-\mu_x) = 0$ etc. (c.f. Barlow *et. al*, 1972, p. 121) and setting $s^2 = \frac{1}{n} z'z$, if

$$E_0^2 = \frac{\mu_x'\mu_x+y'y}{x'x+y'y+ns^2} > v_\alpha;\ 0 \leqslant v_\alpha \leqslant 1, \qquad \ldots \quad (3),$$

where v_α is chosen so that the level of significance of the test (3) equals a given α.

We notice, c.f. Pincus (1975), that

$$\mu_x'\mu_x = \begin{cases} 0 & \text{if} \quad \dfrac{c'x}{\|c\|\ \|x\|} < -\sin\psi \\ x'x & \text{if} \quad \dfrac{c'x}{\|c\|\ \|x\|} > \cos\psi \\ \left[\dfrac{c'x}{\|c\|} \cos\psi + \left(x'x - \dfrac{(c'x)^2}{\|c\|^2} \right)^{1/2} \sin\psi \right]^2 & \text{elsewhere.} \end{cases} \qquad \ldots \quad (4)$$

The distribution of E_0^2 under the hypothesis is evaluated for $m = 0$ in Pincus (1975). Proceeding on the same lines one gets the null distribution for arbitrary m,

$$P(E_0^2 \leqslant v) = \sum_{j=0}^{p} q_j\, B_{(j+m)/2,\,(n+p-j)/2}(v), \quad 0 \leqslant v \leqslant 1 \qquad \ldots \quad (5)$$

Here $B_{a,b}(\cdot)$ stands for the distribution function of a Beta variate with parameters a and b. The q_j's are defined by

$$q_0 = \frac{1}{2} B_{(p-1)/2,\frac{1}{2}}(\cos^2 \psi),$$

$$q_j = \frac{1}{2}\binom{p-2}{j-1} \frac{\beta\left(\frac{j}{2}, \frac{p-j}{2}\right)}{\beta\left(\frac{1}{2}, \frac{p-1}{2}\right)} \sin^{j-1}\psi \cos^{p-j-1}\psi,\ j = 1 \,..,p-1,$$

$$q_p = \frac{1}{2} B_{(p-1)/2,\frac{1}{2}}(\sin^2 \psi).$$

The distribution of E_0^2 under the hypothesis therefore depends on the parameters m, p, n and ψ. Formula (5) may be used either for evaluation of significance points, or by substituting the v in (5) by the observed value, to compare the right hand side of (5) with $1-\alpha$, (c.f. example in Section 4).

2. Evaluation of the E_0^2 statistic

The crucial point in evaluating the E_0^2-statistic, if the alternative region is given by some inequalities like (1), is the determination of the axis c and the angle ψ which determine the smallest circular cone containing the admissible parameter region, and its representation in the original coordinates, since testing problems usually are not given in a cannonical form.

Consider now the unbalanced one way model of Section 1. Introduce a new orthogonal basis in the N dimensional space so that the first basis vector is $(N^{-1/2}, \ldots, N^{-1/2})'$, the next $(r-1)$ lie in the subspace spanned by $(1, 1, \ldots, 1, 0, \ldots, 0, 0, \ldots, 0')$ (n_1 units). ..., $(0, \ldots, 0, 1, \ldots, 1, 0, \ldots, 0)'$ (n_r units), and the following $k-r$ vectors lie in the span of $(1, \ldots, 1, 0, \ldots, 0)'$ (n_1 units), ..., $(0, \ldots, 0, 1, \ldots, 1)'$ (n_k units); the model gets the cannonical form (2) with

$$l = 1,\ p = r-1,\ m = k-r \text{ and } n = N-k. \tag{6}$$

Following Schaafsma (1966) who proved it for $k = r$, Petrowiak (1978) showed that the inequalities in (1) induce a cone in the x-space of the cannonical form, and that the axis c and the angle ψ of the smallest circular cone in that subspace which contains the alternative region are determined by

$$c = (c_1, \ldots, c_1, \ldots, c_r, \ldots, c_r, 0, \ldots, 0)'$$

and

$$\cos^2 \psi = \left(\sum_{i=1}^{r} n_i c_i^2 \right)^{-1}. \qquad \dots \quad (7)$$

The c_i are given, (c.f. Schaafsma, 1966, p. 79) by

$$c_i = s_r^{-1/2} n_i^{-1} \{ -s_i^{-1/2} (s_r - s_i)^{1/2} + s_{i-1}^{1/2} (s_r - s_{i-1})^{1/2} \}, \; i = 1, \dots, r, \qquad \dots \quad (8)$$

where

$$s_i = \sum_{j=1}^{i} n_j, \; s_0 = 0.$$

Denoting

$$\bar{x}_i = n_i^{-1} \sum_{j=1}^{n_i} x_{ij}, \; i = 1, \dots, k; \; \bar{x} = N^{-1} \sum_{i=1}^{k} n_i \bar{x}_i,$$

$$\bar{x}_{[1]} = s_r^{-1} \sum_{i=1}^{r} n_i \bar{x}_i ; \; \bar{x}_{[2]} = (N - s_r)^{-1} \sum_{i=r+1}^{k} n_i \bar{x}_i,$$

and using (3), (4), (7), the E_0^2 statistic may be given explicitly, (see also Petrowiak, 1978),

$$E_0^2 = \frac{T(x) + \sum_{j=r+1}^{k} n_j (\bar{x}_j - \bar{x}_{[2]})^2 + s_r (\bar{x}_{[1]} - \bar{x})^2 + (N - s_r)(\bar{x}_{[2]} - \bar{x})^2}{\sum_{i=1}^{k} n_i (\bar{x}_i - \bar{x})^2 + (N - k) s^2} \qquad \dots \quad (9)$$

where according to (4) and (7),

$$T(x) = \begin{cases} 0 \quad \text{if} \quad \dfrac{\sum_{i=1}^{r} n_i c_i \bar{x}_i}{\left\{ \sum_{i=1}^{r} n_i (\bar{x}_i - \bar{x}_{[1]})^2 \right\}^{1/2}} < -\left(\sum_{i=1}^{r} n_i c_i^2 - 1 \right)^{1/2} \\[2ex] \sum_{i=1}^{r} n_i (\bar{x}_i - \bar{x}_{[1]})^2 \quad \text{if} \quad \dots > 1 \\[2ex] \left[\dfrac{\sum_{i=1}^{r} n_i c_i \bar{x}_i}{\sum_{i=1}^{r} n_i c_i^2} + \left(\sum_{i=1}^{r} n_i \bar{x}_i - \bar{x}_{[1]})^2 - \dfrac{\left(\sum_{i=1}^{r} n_i c_i \bar{x}_i \right)^2}{\sum_{i=1}^{r} n_i c_i^2} \right)^{1/2} \left(1 - \dfrac{1}{\sum_{i=1}^{r} n_i c_i^2} \right)^{1/2} \right]^2 \text{elsewhere} \end{cases}$$

Obviously the valuation of E_0^2 statistic demands in most cases at least a desk calculator, but it is easily programmable. The distribution of E_0^2 is given by (5) with parameters determined by (6) and (7).

The individual terms in the numerator of (9) are useful for the interpretation of the E_0^2 statistic. If e.g. E_0^2 exceeds the $\alpha\%$-significance point, a large value of $T(x)$ indicates inhomogeneity among the first r groups, etc.

4. Monte Carlo studies of the power function

We will test the equality of six treatment effects in an experiment, where e.g., the first four treatments represent different dosages of the same drug, so that a priori we may suppose $\xi_1 \leqslant \xi_2 \leqslant \xi_3 \leqslant \xi_4$.

Let the sample sizes and the (fictitious) data be as in Table 1.

TABLE 1

	1	2	3	4	5	6
n	6	4	3	5	4	6
$\bar{x}_i$	3.7	4.1	2.1	3.5	3.2	4.0

$N = 28$

$s^2 = .44$

The usual ANOVA test gives an F-value 4·26 which is beyond the 1% significance point $F_{5,22,01} = 3{\cdot}99$. Evaluation of the E_0^2 statistic according to (9) gives ·1383, while by (5) we have $P(E_0^2 < {\cdot}1383) = {\cdot}3338$, so that using a 1%, 5% or even 10% significance level we had to accept the hypothesis.

Interpretation of the different outcomes is easy-the deviation e.g., between $\bar{x}_3$ and $\bar{x}$ which alone contributes 2·83 to the F value is automatically adjusted by the E_0^2 statistic.

The example looks of course somewhat pathological and in any case, says nothing about the quality of the E_0^2 test with respect to its power function. Therefore we simulated the power of the E_0^2-test for this model. This was done in the following way.

28 standard normal variables were generated and the $\bar{x}_i$ were and s^2 evaluated. Subsequently the $\bar{x}_i$ were replaced by $\bar{x}_i + \xi_i$ for different combinations of the ξ_i and the value of E_0^2 was computed and compared with the 5% significance point. The whole procedure was repeated 10,000 times. Or

more exactly, we choose the first 9500 repetitions for which the E_0^2 statistic was in the acceptance region, and the first 500 for which it was in the rejection region, when all the random variables had zero means; thus making the empirical size 0·05.

Since the power of the E_0^2-test depends not only on the noncentrality parameter $\sum_{i=1}^{6} n_i(\bar{\xi}_i-\bar{\xi})^2$ like the F-test, but also on the configuration of $\xi_1, \ldots, \xi_6$, they were chosen so that they formed either multiples of $(c_1, \ldots, c_4, 0, 0) = (-\cdot 333, -\cdot 027, \cdot 069, \cdot 380, 0, 0)$ which guarantees maximum power for fixed noncentrality, or multiples of $(0, 0, 0, 0, -\cdot 6, \cdot 4)$ which gives minimum power, (c.f. Pincus, 1975). The factors were chosen so that

$$\phi = \sqrt{\sum_{i=1}^{6} n_i(\xi_i-\bar{\xi})^2/6}$$

equals ·3, ·5 etc. as indicated in Table 2, where the empirical power function (percentage of rejections) is compared with that of the F-test. The exact values for the F-test are taken from Tiku (1967).

TABLE 2

ϕ	0	.3	.5	1.0	1.2	1.4	1.6	1.8	2.0	2.6
exact F-power	.050	—	.109	.343	.487	.637	.770	.871	.937	.997
emp. F-power	.051	.069	.109	.345	.488	.640	.771	.873	.937	.997
emp. E-power (max. config.)	.050	.117	.206	.549	.693	.819	.907	.958	.984	.999
emp. E-power (min. config.)	.050	.074	.119	.405	.563	.717	.838	.921	.965	.999

The results suggest that the application of the E_0^2 test for partially ordered alternatives gives a considerable gain in power, even in situations where the inhomogeneity is due to means not involved in order restrictions. The conclusions are supported by power evaluations of the likelihood ratio test, for a related model with σ^2 known in Barlow *et al.* (1972, p. 159).

Acknowledgement. The authors are grateful to Dr. M. Hartwig, Berlin, for his help in providing the simulations for Table 2.

References

ABELSON, R. P. and TUKEY, J. W. (1963): Efficient utilization of non-numerical information in quantitative analysis; general theory and the case of simple order. *Ann. Math. Statist.*, **34**, 1347-1864.

BARLOW, R. E., BARTHOLOMEW, D. J., BRENNER, J. M. and BRUNK, H. D. (1972): *Statistical Inference Under Order Restrictions*, Wiley, London and New York.

BARTHOLOMEW, D. J. (1961): Ordered tests in the analysis of variance. *Biometrika*, **48**, 325-332.

PETROWIAK, C. (1978): Testing homogeneity of means in the unbalanced one way model (in German), Master-Thesis, Humboldt University, Berlin.

PINCUS, R. (1975): Testing linear hypothesis under restricted alternatives, *Math. Operation forsch. Statist.*, **6**, 733-751.

SCHAAFSMA, W. (1966): *Hypothesis Testing Problems with the Alternative Restricted by a Number of Inequalities*, P. Noordhoff, Gronigen.

SISKIND, V. (1976): Approximate probability integrals and critical values for Bartholomew's test for ordered means. *Biometrika*, **63**, 647-654.

TIKU, M. L. (1967): Tables for the power of the F-test. *J.A.S.A.*, **62**, 525-539.

Paper received : July, 1979.
Revised : June, 1980.