ON THE ESTIMATION OF CONTRASTS IN LINEAR MODELS

By Subha Bhuchongkul[1] and Madan L. Puri[2]

University of California, Berkeley and New York University

1. Summary. In linear models with several observations per cell, a class of estimates of all contrasts are defined in terms of rank test statistics such as the Wilcoxon or normal scores statistic, which extend the results of Hodges and Lehmann (1963) and Lehmann (1963). The asymptotic efficiency of these estimates relative to the standard least squares estimates, as the number of observations in each cell gets large, is shown to be the same as the Pitman efficiency of the rank tests on which they are based to the corresponding t-tests.

2. A Class of Estimates of Contrasts. Let the observable random variables be $X_{i\alpha}$, and suppose they are of the form

$$X_{i\alpha} = \xi_i + U_{i\alpha} \qquad (\alpha = 1, \cdots, m_i; i = 1, \cdots, c) \tag{2.1}$$

where the variables $U_{i\alpha}$ are independently distributed with common distribution F having density f, and the ξ's are unknown constants. Denote by X_i the vector $(X_{i1}, \cdots, X_{im_i})$ and suppose that the Hodges-Lehmann statistic h [(3.1) of [4]] is calculated for every pair of samples, there being $c(c-1)/2$ pairs in all. We shall write $h_{ij}(X_i, X_j)$ for the value obtained from the ith and jth samples $(i, j = 1, \cdots, c; i \neq j)$. Thus we have

$$h_{ij}(X_i, X_j) = \sum_{K=1}^{m_j} E_\Psi[V^{(S_K)}], \tag{2.2}$$

where $S_1 < \cdots < S_{m_j}$ denote the ranks of $X_{j1}, \cdots, X_{jm_j}$ in the combined ith and jth samples, and where $V^{(1)} < \cdots < V^{(m_i+m_j)}$ denote an ordered sample of size $(m_i + m_j)$ from a distribution Ψ. Let

$$\begin{aligned} \Delta_{ij}^* &= \sup\{\Delta_{ij}: h_{ij}(X_i, X_j - \Delta_{ij}) > \mu\}, \\ \Delta_{ij}^{**} &= \inf\{\Delta_{ij}: h_{ij}(X_i, X_j - \Delta_{ij}) < \mu\}, \end{aligned} \tag{2.3}$$

where μ is the point of symmetry of the distribution of $h_{ij}(X_i, X_j)$ when $\Delta_{ij} = 0$ i.e. when $\xi_i = \xi_j$. It was shown in [4] that the estimate $\hat{\Delta}_{ij} = (\Delta_{ij}^* + \Delta_{ij}^{**})/2$ of $\xi_i - \xi_j$ has more robust efficiency than the classical estimate $T_{ij} = X_{i\cdot} - X_{j\cdot}$, where $X_{i\cdot} = \sum_{\alpha=1}^{m_i} X_{i\alpha}/m_i$.

Since the estimates $\hat{\Delta}_{ij}$ are incompatible in the sense that they do not satisfy the linear relations satisfied by the differences they estimate [see Lehmann [5], [6]], Lehmann proposed the adjusted estimates Z_{ij} of the type

$$Z_{ij} = \hat{\Delta}_{i\cdot} - \hat{\Delta}_{j\cdot}. \tag{2.4}$$

Received 19 March 1964; revised 30 June 1964.

[1] Research sponsored by the Office of Naval Research, Contract Nonr-222(43).

[2] Research sponsored by Sloan Foundation grant for statistics and by U. S. Navy Contract Nonr-285(38).

where

$$\hat{\Delta}_{i\cdot} = \sum_{j=1}^{c} \hat{\Delta}_{ij}/c. \tag{2.5}$$

(For a short cut method of computing $\hat{\Delta}_{ij}$, the reader is referred to [4], p. 602.)

Then for any contrast $\sum c_i\xi_i$ with $\sum c_i = 0$, which can also be written in the form

$$\theta = \sum_{i=1}^{c} \sum_{j=1}^{c} d_{ij}(\xi_i - \xi_j) \tag{2.6}$$

the estimate

$$\hat{\theta} = \sum_{i=1}^{c} \sum_{j=1}^{c} d_{ij}Z_{ij} = \sum_{i=1}^{c} \sum_{j=1}^{c} d_{ij}(\hat{\Delta}_{i\cdot} - \hat{\Delta}_{j\cdot}) \tag{2.7}$$

is proposed.

3. Asymptotic distribution and efficiency. The asymptotic distribution of the adjusted estimates Z_{ij} is given by the following theorem, where the sample sizes m_i are assumed to tend to infinity in such a way that $m_i = \rho_i \cdot N$, $N \to \infty$ and $i = 1, \cdots, c$.

THEOREM 3.1.

(i) *The joint distribution of* $(V_1, \cdots, V_{c-1})$ *where*

$$V_i = N^{\frac{1}{2}}[\hat{\Delta}_{ic} - (\xi_i - \xi_c)] \tag{3.1}$$

is asymptotically normal with zero mean and covariance matrix

$$\begin{aligned} \operatorname{Var}(V_i) &= (1/\rho_i + 1/\rho_c)A^2/B^2, \\ \operatorname{Cov}(V_i, V_j) &= A^2/(\rho_c \cdot B^2) \end{aligned} \tag{3.2}$$

where

$$A^2 = \int_0^1 J^2(x)\,dx - \left(\int_0^1 J(x)\,dx\right)^2, \qquad J = \Psi^{-1} \tag{3.3}$$

$$B = \int J'[F(x)]f^2(x)\,dx. \tag{3.4}$$

Here the density f of F is assumed to satisfy the regularity conditions of Lemma 7.2 *of* [8].

(ii) *For any i and j,*

$$N^{\frac{1}{2}}\hat{\Delta}_{ij} \sim N^{\frac{1}{2}}(\hat{\Delta}_{ic} - \hat{\Delta}_{jc}), \tag{3.5}$$

where $\sim$ indicates that the difference of the two sides tend to zero in probability.

(iii) *The difference $N^{\frac{1}{2}}(Z_{ij} - \hat{\Delta}_{ij})$ tends to zero in probability for all i, j.*

The proof of (i) rests on the following lemma.

LEMMA 1. *Suppose that the variables $X_{i\alpha}$ have the distribution specified in connection with* (2.1) *with fixed F but a sequence of means*

$$(\xi_1, \cdots, \xi_c) = (\xi_1^{(N)}, \cdots, \xi_c^{(N)})$$

satisfying

$$\xi_i^{(N)} - \xi_c^{(N)} = -a_i/N^{\frac{1}{2}}. \tag{3.6}$$

Let $h_{ij}(X_i, X_j)$ be defined as in (2.2) with Ψ satisfying the assumptions of Theorem 1 of [1], *then the variables $(W_1, \cdots, W_{c-1})$ given by*

$$W_i = N^{\frac{1}{2}}[h_{ic}/m_c - \mu_{ic}] \qquad i = 1, \cdots, c-1 \tag{3.7}$$

have a joint asymptotic normal distribution as $N \to \infty$, with zero mean and covariance matrix

$$\begin{aligned} \text{Var}(W_i) &= A^2\rho_i/(\rho_i + \rho_c)\rho_c, \\ \text{Cov}(W_i, W_j) &= A^2\rho_i\rho_j/\rho_c(\rho_i + \rho_c)(\rho_j + \rho_c) \end{aligned} \tag{3.8}$$

and

$$\mu_{ic} = \int J\left[\frac{m_c}{m_i + m_c} F(x) + \frac{m_i}{m_i + m_c} F(x + a_i/N^{\frac{1}{2}})\right] dF(x).$$

The proof of this lemma is given in the appendix.

PROOF OF THEOREM 3.1. (i) By 9.1 of [4],

$$\lim P\{N^{\frac{1}{2}}[\hat{\Delta}_{ic} - (\xi_i - \xi_c)] \leqq a_i \text{ for all } i\} = \lim P_N\{N^{\frac{1}{2}}[(1/m_c)h_{ic} - \alpha] \leqq 0 \text{ for all } i\}$$

where $\alpha = \int J[F(x)]\, dF(x)$ and P_N indicates that the probability is computed for a sequence of means satisfying (3.6). Furthermore since by Lemma 7.2 of [8] $N^{\frac{1}{2}}(\mu_{ic} - \alpha) \to -a_iB\rho_i/(\rho_i + \rho_c)$ as $N \to \infty$, it follows that

$$\lim P\{N^{\frac{1}{2}}[\hat{\Delta}_{ic} - (\xi_i - \xi_c)] \leqq a_i \text{ for all } i\} = \lim P_N\{N^{\frac{1}{2}}[(1/m_c)h_{ic} - \mu_{ic}] \leqq a_iB\rho_i/(\rho_i + \rho_c) \text{ for all } i\}.$$

By Lemma 1, this is equal to $Q(a_1, \cdots, a_{c-1})$ where Q is the $(c-1)$ dimensional multivariate normal distribution with zero mean and covariance matrix (3.2).

Parts (ii) and (iii) of the theorem follow by Lemma 2 of Lehmann (1963).

The proof of the following theorem exactly parallels Lehmann's argument, see for example Theorem 3 of [6], and is therefore omitted.

THEOREM 3.2. *The asymptotic efficiency of the estimate $\hat{\theta} = \sum_{i=1}^{c}\sum_{j=1}^{c} d_{ij}Z_{ij}$ of $\theta = \sum_{i=1}^{c}\sum_{j=1}^{c} d_{ij}(\xi_i - \xi_j)$ relative to the estimate $\sum_{i=1}^{c}\sum_{j=1}^{c} d_{ij}(X_{i\cdot} - X_{j\cdot})$ is*

$$e = \sigma^2B^2/A^2, \tag{3.9}$$

where $\sigma^2 = \text{Var}(X_{i\alpha})$, and where A^2 and B^2 are given by (3.3) and (3.4) respectively.

In particular when $J = \Phi^{-1}$, where Φ is the standard normal cumulative distribution function having the density ϕ then (3.9) is the same as the Pitman efficiency of the normal scores test relative to the t-test [cf. 1].

4. Appendix.

Proof of Lemma 1. Let $F_{m_i}(x)$ be the cdf (cumulative distribution function) of m_i observations $X_{i1}, \cdots, X_{im_i}$ of which the population cdf is $F_i(x) = F(x - \xi_i)$. Denote $m_{ic} = m_i + m_c$ and $\lambda_{ic} = m_c/m_{ic}$; $i = 1, \cdots, c-1$. Define $H_{m_i c}(x) = \lambda_{ic} F_{m_c}(x) + (1 - \lambda_{ic}) F_{m_i}(x)$ and $H_{ic}(x) = \lambda_{ic} F_c(x) + (1 - \lambda_{ic}) F_i(x)$. Then [cf. Chernoff-Savage (1958)] we can write

$$T_{ic} = h_{ic}/m_c = A^{(ic)} + B_{1N}^{(ic)} + B_{2N}^{(ic)} + \textstyle\sum_{K=1}^{6} C_{KN}^{(ic)}, \tag{4.1}$$

where

$$A^{(ic)} = \int J[H_{ic}(x)]\, dF_c(x), \tag{4.2}$$

$$B_{1N}^{(ic)} = \int J[H_{ic}(x)]\, d[F_{m_c}(x) - F_c(x)], \tag{4.3}$$

$$B_{2N}^{(ic)} = \int [H_{m_i c}(x) - H_{ic}(x)] J'[H_{ic}(x)]\, dF_c(x) \tag{4.4}$$

and the C-terms are all $o_p N^{-\frac{1}{2}}$.

The difference $N^{\frac{1}{2}}(T_{ic} - A^{(ic)}) - N^{\frac{1}{2}}(B_{1N}^{(ic)} + B_{2N}^{(ic)})$ tends to zero in probability and so, by a well-known theorem ([2], p. 299) the vectors $(W_1, \cdots, W_{c-1})$ and $(Z_i, \cdots, Z_{c-1})$ where $Z_i = N^{\frac{1}{2}}(B_{1N}^{(ic)} + B_{2N}^{(ic)})$ possess the same limiting distribution. Thus to prove the lemma it suffices to show that for any real δ_i; $i = 1, \cdots, c-1$, not all zero, $\sum_{i=1}^{c-1} \delta_i Z_i$ has normal distribution in the limit. Now proceeding as in [1] or [8], we find

$$\begin{aligned}\textstyle\sum_{i=1}^{c-1} \delta_i Z_i = & -\textstyle\sum_{i=1}^{c-1} [\delta_i \cdot ((1 - \lambda_{ic})/m_i) \sum_{\alpha=1}^{m_i} \{B_{ic}^{*}(X_{i\alpha}) - EB_{ic}^{*}(X_i)\}] \\ & + \textstyle\sum_{i=1}^{c-1} \delta_i (1 - \lambda_{ic}) \{m_c^{-1} \sum_{\alpha=1}^{m_c} B_{ic}(X_{c\alpha}) - EB_{ic}(X_c)\},\end{aligned} \tag{4.5}$$

where

$$B_{ic}(x) = \textstyle\int_{x_{i0}}^{x} J'[H_{ic}(y)]\, dF_i(y), \tag{4.6}$$

$$B_{ic}^{*}(x) = \textstyle\int_{x_{i0}}^{x} J'[H_{ic}(y)]\, dF_c(y) \tag{4.7}$$

and x_{i0} is such that $H_{ic}(x_0) = \frac{1}{2}$.

The above summations involve independent samples of identically distributed random variables having finite first two moments. Hence $\sum_{i=1}^{c-1} \delta_i Z_i$ when properly normalized has normal distribution in the limit. The proof follows.

The covariance matrix (3.8) is obtained by taking limits of N Var $(B_{1N}^{(ic)} + B_{2N}^{(ic)})$ and N Cov $(B_{1N}^{(ic)} + B_{2N}^{(ic)}, B_{1N}^{(jc)} + B_{2N}^{(jc)})$ as $N \to \infty$.

REFERENCES

[1] Chernoff, Herman and Savage, I. Richard (1958). Asymptotic normality and efficiency of certain non-parametric tests. *Ann. Math. Statist.* **29** 972–994.

[2] Cramér, Harald (1946). *Mathematical Methods of Statistics.* Princeton University Press.

[3] Hodges, J. L., Jr. and Lehmann, E. L. (1961). Comparison of the normal scores and Wilcoxon Tests. *Proc. Fourth Berkeley Sym. Math. Statist. Prob.* **1** 307–318.

[4] Hodges, J. L., Jr. and Lehmann, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611.

[5] Lehmann, E. L. (1957). A theory of some multiple decision problems, I. *Ann. Math. Statist.* **28** 1–25.

[6] LEHMANN, E. L. (1963). Robust estimation in analysis of variance. *Ann. Math. Statist.* **34** 957–966.

[7] LEHMANN, E. L. (1963). Asymptotically non-parametric inference: an alternative approach to linear models. *Ann. Math. Statist.* **34** 1494–1505.

[8] PURI, MADAN L. (1964). Asymptotic efficiency of a class of c-sample tests. *Ann. Math. Statist.* **35** 102–121.

[9] PURI, MADAN L. (1964). Some distribution-free k-sample rank tests of homogeneity against ordered alternatives. *Ann. Math. Statist.* (abstract) **35** 461.