ON ROBUST ESTIMATION IN INCOMPLETE BLOCK DESIGNS[1]

By Madan Lal Puri and Pranab Kumar Sen

Courant Institute of Mathematical Sciences, New York University and University of North Carolina at Chapel Hill

1. Introduction and summary. The object of the present investigation is to generalize the results of Greenberg (1966) to a wider class of robust estimators which includes her estimator as a special case.

As in Greenberg (1966), we consider an incomplete block design D consisting of J blocks of (constant) size b to which $c(>b)$ treatments are applied, there being n_j replications of the jth block, for $j = 1, \cdots, J$. Let $n = \sum_{j=1}^{J} n_j$ and let S_j consist of the numbers of the b treatments applied in the jth block, for $j = 1, \cdots, J$. The observable random variables are then

$$X_{ij\alpha} = \nu + \xi_i + \mu_j + \beta_{j\alpha} + U_{ij\alpha}, \qquad \alpha = 1, \cdots, n_j; \quad i \,\varepsilon\, S_j; \quad j = 1, \cdots, J, \tag{1.1}$$

where ξ_i is the ith treatment-effect, μ_j the jth replication effect, $\beta_{j\alpha}$ the effect of the αth block in the jth replication set and $U_{ij\alpha}$'s are independent and identically distributed residual error components with a common distribution $F(u)$. We may set (without any loss of generality) that

$$\sum_{i=1}^{c} \xi_i = 0, \qquad \sum_{j=1}^{J} \mu_j = 0; \qquad \sum_{\alpha=1}^{n_j} \beta_{j\alpha} = 0 \quad \text{for all} \quad j = 1, \cdots, J. \tag{1.2}$$

Our intention is to provide some robust estimators of contrasts among ξ_i's and to study their various properties.

2. A class of rank order estimates. Define

$$Y_{(i,t)j\alpha} = X_{ij\alpha} - X_{tj\alpha}, \qquad \Delta_{it} = \xi_i - \xi_t, \qquad U_{(i,t)j\alpha} = U_{ij\alpha} - U_{tj\alpha} \tag{2.1}$$

for all $\alpha = 1, \cdots, n_j$, $(i \,\varepsilon\, S_j, t \,\varepsilon\, S_j)$, $j = 1, \cdots, J$, and we denote the common cumulative distribution function (cdf) of $U_{(i,t)j\alpha}$ by $G(u)$. By definition $G(u)$ is symmetric about $u = 0$ and if it has a finite variance that will be equal to $\sigma_G^2 = 2\sigma^2$, where σ^2 is the variance of $U_{ij\alpha}$. From (1.1) and (2.1), the cdf of $Y_{(i,t)j\alpha}$ is $G(u - \Delta_{it})$. Let

$$\mathbf{Y}_{(i,t)j} = (Y_{(i,t)j1}, \cdots, Y_{(i,t)jn_j}), \qquad (i, t) \,\varepsilon\, S_j, \quad j = 1, \cdots, J. \tag{2.2}$$

We consider the rank order statistic

$$h_{n_j}(\mathbf{Y}_{(i,t)j}) = \sum_{\alpha=1}^{n_j} E_{n_j,\alpha} Z_{n_j,\alpha}^{(i,t)} / n_j, \tag{2.3}$$

Received 8 March 1967.

[1] This work was partially supported by the Office of Naval Research, Nonr-285(38) and the National Institutes of Health, Public Health Service, Grant GM-12868. Reproduction in whole or in part is permitted for any purpose of the United States Government.

where $E_{n_j,\alpha}$ is the expected value of the αth order statistic of a sample of size n_j drawn from a distribution

$$\begin{aligned}\Psi^*(x) &= \Psi(x) - \Psi(-x-), \quad x \geqq 0,\\ &= 0, \qquad\qquad\qquad x < 0,\\ \Psi(x) &= 1 - \Psi(-x-), -\infty < x < \infty;\end{aligned} \tag{2.4}$$

and $Z_{n_j,\alpha}^{(i,t)} = 1$ if the αth smallest observation among $|Y_{(i,t)j\alpha}|$, $\alpha = 1, \cdots, n_j$, is from a positive $Y_{(i,t)j\alpha}$ and $Z_{n_j,\alpha}^{(i,t)} = 0$, otherwise. In passing, we remark that the cdf $\Psi(x)$ in (2.4) is assumed to satisfy the conditions of Theorem 1 of Chernoff and Savage (1958).

Let us denote by $\mathbf{I}_{n_j}$ the n_j-vector having all elements equal to 1 and define

$$\begin{aligned}\Delta_{it}^{*(j)} &= \sup\{a: h_{n_j}(\mathbf{Y}_{(i,t)j} - a\mathbf{I}_{n_j}) > \mu\},\\ \Delta_{it}^{**(j)} &= \inf\{a: h_{n_j}(\mathbf{Y}_{(i,t)j} - a\mathbf{I}_{n_j}) < \mu\},\end{aligned} \tag{2.5}$$

μ being the point of symmetry of the cdf of h_{n_j} when for the cdf $G(u - \Delta_{it})$ of $Y_{(i,t)j\alpha}$, $\Delta_{it} = 0$. It is well known (cf. [2], [5]) that

$$\hat{\Delta}_{it}^{(j)} = \tfrac{1}{2}[\Delta_{ij}^{*(j)} + \Delta_{it}^{**(j)}] \tag{2.6}$$

is a translation invariant estimator of Δ_{it} and its distribution is symmetric about 0. It may be noted that if we work with $\Psi(x) = (x + 1)/2$, $-1 \leqq x \leqq 1$, we obtain the Wilcoxon-type of estimator which has been studied in detail by Greenberg (1966). Another important estimator, termed the normal score estimator, may be obtained by using $\Psi(x)$ as the standardized normal cdf and will be shown to have some desirable properties.

Let us denote

$$\hat{\Delta}_{i.}^{(j)} = (1/b)\textstyle\sum_{t\varepsilon S_j}\hat{\Delta}_{it}^{(j)} \qquad (\text{where } \hat{\Delta}_{ii}^{(j)} = 0), \tag{2.7}$$

and we define the *compatible or adjusted estimator* of $\Delta_{it}^{(j)}$ as

$$Z_{it}^{(j)} = \hat{\Delta}_{i.}^{(j)} - \hat{\Delta}_{t.}^{(j)} \qquad \text{for all } i, t\,\varepsilon\, S_j,\ j = 1, \cdots, J. \tag{2.8}$$

For the study of the asymptotic distribution of the adjusted estimators in (2.8), we shall assume that

$$n_j = n\rho_j{:}0 < \rho_j < 1 \qquad \text{for all } j = 1, \cdots, J; \qquad n \to \infty. \tag{2.9}$$

Then, we have the following.

THEOREM 2.1. *If the density function $g(x) = G'(x)$ satisfies the regularity conditions of Lemma 3(a) of Hodges and Lehmann (1961) and the* cdf $\Psi(x)$ *in* (2.4) *satisfies the conditions of Theorem 1 of Chernoff and Savage (1958), then subject to* (2.9)

$$n^{\frac{1}{2}}(Z_{it}^{(j)} - \Delta_{it}) \qquad i, t\,\varepsilon\, S_j,\ j = 1, \cdots, J,$$

have asymptotically a joint normal distribution with means zero and a covariance *matrix having elements*

$$\begin{aligned} &\sigma_{jit,j'i't'} \\ &= 0, && \text{if } i, i', t, t' \text{ are distinct and } j = j' \text{ or if } j \neq j', \\ (2.10) \quad &= A^2/\rho_j B^2, && \text{if } j = j', \ i = i', \ t = t', \\ &= S^2/2\rho_j, && \text{if } j = j', \ i = i' \text{ or } t = t', \\ &= -S^2/2\rho_j, && \text{if } j = j', \ i = t' \text{ or } t = i', \end{aligned}$$

where

$$(2.11) \qquad A^2 = \int_0^1 J^2(u)\,du, \qquad B = \int_{-\infty}^{\infty} (d/dx)J[G(x)]\,dG(x),$$

$$(2.12) \qquad S^2 = (2/b)[A^2 + (b-2)\lambda_J(G)]/B^2,$$

$$(2.13) \qquad \lambda_J(F) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J[G(x)]J[G(y)]\,dG^*(x, y), \quad J = \Psi^{-1},$$

and $G^*(x, y)$ *is the joint* cdf *of* $U_{(i,t)j\alpha}$, $U_{(i,t')j\alpha}(t \neq t')$ *whose marginal* cdf's *are* $G(x)$ *and* $G(y)$, *respectively.*

Since the $Z_{it}^{(j)}$'s are linear functions of $\Delta_{it}^{(j)}$, $i, t \in S_j$, $j = 1, \cdots, J$, it is enough to prove the following:

LEMMA 2.2. *Under the assumptions of Theorem* 2.1, *the random variables* $n^{\frac{1}{2}}(\hat{\Delta}_{it}^{(j)} - \Delta_{it})$, $i, t \in S_j$, $j = 1, \cdots, J$, *have asymptotically a multi-normal distribution with null mean vector and covariance matrix having the elements*

$$\begin{aligned} &\sigma^*_{jit,j'i't'} \\ &= 0 && \text{if } i, i', t, t' \text{ are distinct and } j = j' \text{ or if } j \neq j' \\ (2.14) \quad &= A^2/\rho_j B^2 && \text{if } j = j', \ i = i' \text{ and } t = t' \\ &= \lambda_J(F)/\rho_j B^2 && \text{if } j = j', \ i = i' \text{ or } t = t' \\ &= -\lambda_J(F)/\rho_j B^2 && \text{if } j = j', \ i = t' \text{ or } t = i', \end{aligned}$$

where A^2, B^2 *and* $\lambda_J(F)$ *are defined by* (2.11) *and* (2.13).

The proof of this lemma follows from Theorem 3.1 of Puri and Sen (1966) as does Lemma 1 of Greenberg (1966) from Theorem 1 of Lehmann (1964). The computations of the covariance terms in (2.14) are straightforward and are therefore omitted.

Also, it has been shown by Puri and Sen (1966) that $\lambda_J(F) \leqq \frac{1}{2}A^2$ for all continuous distributions F. As such, upon considering the balanced incomplete block designs (for which $n_1 = \cdots = n_J$) and proceeding as in the proof of Theorem 1 of Greenberg (1966), we obtain the following theorem with the aid of our Lemma 2.2.

THEOREM 2.3. *In the class* R *of all linear functions of the random variables* $\hat{\Delta}_{it}^{(j)}$, $i, t \in S_j$, $j = 1, \cdots, J$, *which are unbiased estimators of* Δ_{it}, *an asymptotically minimum variance unbiased estimator is obtained by substituting* $Z_{it}^{(j)}$ *for*

C_{it}^j, $i, t \varepsilon S_j$, $j = 1, \cdots, J$, *in the classical least square estimate, where* C_{it}^j *is defined by* (4) *in Greenberg* (1966). *If* $\lambda_J(F) < \frac{1}{2}A^2$, *this is the unique asymptotically minimum variance unbiased estimator in* R.

3. Asymptotic efficiency of the estimates. On defining S^2 by (2.12), we may note that [cf. Lemma 3 of Greenberg (1966) and our Theorem 2.1] $\{(2^{\frac{1}{2}}\sigma/S)n^{\frac{1}{2}} \cdot (Z_{it}^{(j)} - \Delta_{it}), i, t \varepsilon S_j, j = 1, \cdots, J\}$ and $\{n^{\frac{1}{2}}(C_{it}^j - \Delta_{it}), i, t \varepsilon S_j, j = 1, \cdots, J\}$ having the same limiting normal distribution when $n_1 = \cdots = n_J$ (i.e., $n = Jn_1$). Hence, the asymptotic relative efficiency (ARE) of $Z_{it}^{(j)}$ with respect to the least square estimate C_{jt}^i is given by

$$(3.1) \qquad e(\Psi) = 2\sigma^2/S^2 = b\sigma^2B^2/[A^2 + (b-2)\lambda_J(F)],$$

where A^2, B and $\lambda_J(F)$ are defined by (2.11) and (2.13). It may be noted that (3.1) is independent of $i, t \varepsilon S_j$, $j = 1, \cdots, J$, but depends on b, the block size. On substituting $b = c$ (the number of treatments), (3.1) agrees with the expression for the efficiency in the complete block experiments [cf. Puri and Sen (1966), (3.11)]. Again, if we take $\Psi(x) = (x+1)/2: -1 \leqq x \leqq 1$, we obtain the results of Greenberg (1966) and of Lehmann (1964) (when $b = c$). Since, σ^2, B^2, A^2 and $\lambda_J(F)$ are all independent of b (and c) and $\lambda_J(F) \leqq \frac{1}{2}A^2$ [cf. Puri and Sen (1966)], it is easily seen that (3.1) is an increasing function of $b: 2 \leqq b \leqq c$; its minimum value being equal to $2\sigma^2B^2/A^2$. Thus,

$$(3.2) \qquad 2\sigma^2B^2/A^2 \leqq e(\Psi) \leqq c\sigma^2B^2/[A^2 + (c-2)\lambda_J(F)].$$

Now, if we use $\Psi(x)$ as the standardized normal cdf (i.e., (2.3) as the one-sample normal score test-statistic), then noting that the cdf $G(x)$ [of $U_{(i,t)j\alpha}$] has the variance $2\sigma^2$, we get from (2.11) and the well-known result of Chernoff and Savage (1958) that

$$(3.3) \qquad \inf_{G \varepsilon \mathcal{G}} 2\sigma^2B^2/A^2 = 1,$$

where $\mathcal{G}$ is the class of all continuous cdf's and the equality sign in (3.3) holds only when G is a normal cdf. Thus from (3.2) and (3.3), we get for the normal score estimator

$$(3.4) \qquad \inf_{G \varepsilon \mathcal{G}} e(\Psi) \geqq 1 \qquad \text{for all} \quad b = 2, \cdots, c,$$

where the equality sign holds only for normal G's.

Finally, it can be shown that the efficiency (2.7) holds not only for the differences $\xi_i - \xi_t$ but extends to the estimation of any contrast $\theta = \sum \sum d_{it}(\xi_i - \xi_t)$. The details are omitted for intended brevity.

REFERENCES

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

[2] Greenberg, V. L. (1962). Robust estimation in incomplete block designs. *Ann. Math. Statist.* **37** 1331–1337.

[3] HODGES JR., J. L. and LEHMANN, E. L. (1961). Comparison of the normal scores and Wilcoxon tests. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 307–318. Univ. of California Press.

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

[5] LEHMANN, E. L. (1964). Asymptotically nonparametric inference in some linear models with one observation per cell. *Ann. Math. Statist.* **35** 726–734.

[6] PURI, M. L. and SEN, P. K. (1966). Some optimum nonparametric procedures in two way layout. *Inst. Statist. Univ. of North Carolina, Mimeo Ser. Report No.* 485.