ROBUST ESTIMATION IN INCOMPLETE BLOCKS DESIGNS[1]

By Vida L. Greenberg

University of California, Berkeley and San Francisco State College

1. Introduction and summary. Robust estimates of contrasts in treatment effects for experiments with one observation per cell were proposed by Lehmann [6] for complete (randomized) blocks designs. The model for the observations $X_{i\alpha}$ $(i = 1, \cdots, c; \alpha = 1, \cdots, n)$ is in this case

$$X_{i\alpha} = \nu + \xi_i + \mu_\alpha + U_{i\alpha} \qquad (\textstyle\sum \xi_i = \sum \mu_\alpha = 0) \tag{1}$$

where the ξ's are the treatment effects, the μ's are the block effects, and the U's are independent with a common continuous distribution. Here we shall generalize these estimates to experiments in which the block size is smaller than the number of treatments to be compared, and we shall obtain their asymptotic efficiencies relative to the classical estimates.

Since we are concerned with large sample theory, we shall be interested in designs in which the blocks are replicated a large number (at least 4) times. Such designs could be applied to situations in which only a few different treatment combinations are practicable but each could be replicated several times. For example, in an experiment to compare various diets for pigs, the natural block is the litter. One may wish to compare c diets and have available a number of litters of size $b < c$. An incomplete blocks design using some J litters could first be selected and then the whole design replicated several times using the remaining litters or (e.g. if some comparisons were of greater interest than others) some groups of b diets could be given to more litters than others. Thus the situation to be considered is that in which c treatments are to be compared and the blocks of experimental units are all of size $b < c$. An incomplete blocks design D consisting of J blocks of size b is selected, the number n_j of replications of the jth block is decided upon $(j = 1, \cdots, J; n_j = \rho_j n)$, $\sum n_j$ blocks of experimental units are selected and numbered, and $\sum n_j$ sets of b treatments are assigned to the selected blocks as specified by D and the n_j. The set of blocks receiving the same treatments will be called a replication set. After the assignment of treatments to blocks, the order of application within the blocks is randomized.

Assuming fixed effects and no interaction between treatment and block effects, the model for D is

$$X_{ij} = \nu' + \xi_i + \mu_j + U_{ij} \quad (j = 1, \cdots, J; i \in S_j) \tag{2}$$
$$\textstyle\sum_{i=1}^{c} \xi_i = \sum_{j=1}^{J} \mu_j = 0$$

where S_j consists of the numbers of the b treatments applied in the jth block, the

Received 6 December 1965.

[1] Prepared with the partial support of the National Science Foundation, Grant GP-2593 (at the Statistical Laboratory, University of California, Berkeley).

ξ's are the treatment effects, the μ's are the block effects, and the U's are inde-dependent and identically distributed according to a continuous distribution F with mean zero and variance σ^2 (not necessarily finite). Under the same assumptions, the model for the whole design is

$$
\begin{gathered}
X_{ija} = \nu + \xi_i + \mu_j + \beta_{j\alpha} + U_{ij\alpha}, \\
(j = 1, \cdots, J; i \varepsilon S_j; \alpha = 1, \cdots, n_j) \\
\textstyle\sum_{i=1}^{c} \xi_i = \sum_{j=1}^{J} \mu_j = 0; \sum_{\alpha=1}^{n_j} \beta_{j\alpha} = 0 \quad \text{for each } j
\end{gathered} \tag{3}
$$

where S_j is defined as before (the treatments applied to the jth block of D now being applied in the n_j blocks of the jth replication set), ξ_i is the effect of the ith treatment, μ_j is the effect of the jth replication set, $\beta_{j\alpha}$ is the effect of the αth block in the jth replication set, and the U's are distributed as in the model (2). (Although it might appear that the models (2) and (3) are valid only for a fixed order of application of the assigned treatments to the units within a block, the models remain valid under randomization. For, consider a model in which any of the b treatments may be assigned to each unit within the block. If the corresponding U's are independent, identically distributed random variables, then any selection according to specified probabilities will again be independent and identically distributed, which justifies the assumptions of the above models.)

2. The estimates. The classical least squares estimate of a treatment difference $\xi_k - \xi_l$ has the form

$$C_{kl} = \textstyle\sum_{s,t,j} A_{stj}^{kl} C_{st}^{j}, \tag{4}$$

where the A's are constants and

$$C_{st}^{j} = X_{sj\cdot} - X_{tj\cdot} = n_j^{-1} \textstyle\sum_{\alpha=1}^{n_j} (X_{sj\alpha} - X_{tj\alpha}), \quad s, t \varepsilon S_j, j = 1, \cdots, J. \tag{5}$$

(This is seen by observing that the least squares estimates $\hat{\xi}_i$ are unchanged by the addition of constants to the block effects and hence are functions of intrablock differences only; also differences within the blocks of the same replication set are weighted equally.)

Estimates more robust than (4) will be defined as functions of the random variables

$$Y_{st}^{j} = \text{med}_{1 \leq \alpha \leq \beta \leq n_j} (X_{sj\alpha} - X_{tj\alpha} + X_{sj\beta} - X_{tj\beta})/2; \tag{6}$$

Y_{st}^{j} is the one-sample Hodges-Lehmann estimate of $\xi_s - \xi_t$ based on the differences of observations in the jth replication set.

The following definitions and notation will be needed: Let G be the common distribution of the random variables $U_{sj\alpha} - U_{tj\alpha}$; G is symmetric about zero with variance $\tau^2 = 2\sigma^2$. Let g be the density of G. Let

$$\lambda(F) = P[X_1 < X_2 + X_3 - X_4, X_1 < X_5 + X_6 - X_7],$$

$X_1, \cdots, X_7$ being independent random variables with common distribution F. It was shown in [7] that $\frac{1}{4} \leq \lambda(F) \leq \frac{7}{24}$ for all continuous distributions F.

LEMMA 1. *If the density g of G satisfies the regularity conditions of Lemma* 3(a) of [3], *then the random variables $n^{\frac{1}{2}}(Y_{st}^j - (\xi_s - \xi_t))$, $s, t \in S_j$, $j = 1, \cdots, J$, have a joint normal limiting distribution with mean zero and covariance matrix $\Sigma_{Y^j} = (\sigma_{jst,juv})$ given by*

$$\begin{aligned}
\sigma_{jst,j'uv} &= 0 && \text{if} \quad s, t, u, v \text{ distinct and } j = j' \text{ or if } j \neq j' \\
&= 1/12\rho_j(\textstyle\int g^2(x)\,dx)^2 && \text{if} \quad j = j', s = u \text{ and } t = v \qquad (7)\\
&= (\lambda(F) - \tfrac{1}{4})/\rho_j(\textstyle\int g^2(x)\,dx)^2 && \text{if} \quad j = j' \text{ and } s = u \text{ or } t = v \\
&= (\tfrac{1}{4} - \lambda(F))/\rho_j(\textstyle\int g^2(x)\,dx)^2 && \text{if} \quad j = j' \text{ and } s = v \text{ or } t = u.
\end{aligned}$$

PROOF. For $j \neq j'$, Y_{st}^j and $Y_{st}^{j'}$ are independent. For $j = j'$, the proof is a slight modification of the proof of Theorem 1 of [6] and hence will be omitted.

We shall now derive a suitable set of compatible estimates of the differences $\xi_k - \xi_l$, at first restricting attention to the case in which the overall design is balanced, i.e. all $n_j = n$, each treatment occurs the same number of times, and each pair of treatments occurs together in the same number of blocks. The design with model (3) must then consist of n replications of a balanced incomplete blocks design D.

Let R be the class of all linear functions of the random variables Y_{st}^j, $s, t \in S_j$, $j = 1, \cdots, J$, which are unbiased estimates of a difference $\xi_k - \xi_l$. Let $Y_{s\cdot}^j = b^{-1}\sum_{t \in S_j} Y_{st}^j$ and $Z_{st}^j = Y_{s\cdot}^j - Y_{t\cdot}^j$; Z_{st}^j is the Hodges-Lehmann adjusted estimate of $\xi_s - \xi_t$ based on the differences of observations in the jth replication set.

THEOREM 1. *In the class R an asymptotically minimum variance unbiased estimate is obtained by substituting Z_{st}^j for C_{st}^j, $s, t \in S_j$, $j = 1, \cdots, J$ in the classical least squares estimate. If $\lambda(F) < \frac{7}{24}$, this is the unique asymptotically minimum variance estimate in R.*

PROOF. Though $\lambda(F)$ is unknown, it will be considered as known throughout the proof; it turns out that the minimum variance estimate in R does not depend on $\lambda(F)$.

If $\lambda(F) < \frac{7}{24}$, the covariance matrix Σ defined by (7) with $\rho_j = 1$ is non-singular, and the Gauss-Markov theorem applies, yielding a unique minimum variance unbiased estimate in the class R. This least squares estimate may be found by a straightforward computation using Lagrange multipliers to minimize the asymptotic variance of an arbitrary linear function in the class R.

If $\lambda(F) = \frac{7}{24}$ (which may not be possible), it is seen from (7) that the covariance matrix Σ_{Y^j} is proportioned to Σ_{C^j}, the covariance matrix of the joint normal limiting distribution of the random variables $n^{\frac{1}{2}}(C_{st}^j - (\xi_s - \xi_t))$ where $C_{st}^j = X_{sj\cdot} - X_{tj\cdot}$, $s, t \in S_j$, $j = 1, \cdots, J$. Since the Y's and the C's each have a joint normal limiting distribution and their covariance matrices are proportional, any linear function of the Y's has asymptotic variance proportional to the asymptotic variance of the same linear function of the C's, with the same proportionality factor. Hence, the same linear function of the Y's as the classical least squares estimate is of the C's has minimum asymptotic variance among unbiased estimates which are linear functions of the Y's. From (7) with $\lambda = \frac{7}{24}$, it is seen that the asymptotic variance of $n^{\frac{1}{2}}(Y^j_{st} + Y^j_{tu} + Y^j_{us})$ is zero, so that $n^{\frac{1}{2}}(Y^j_{st} + Y^j_{tu} + Y^j_{us}) \to 0$ in probability, and the same proof as in [4] shows that $n^{\frac{1}{2}}(Y^j_{st} - Z^j_{st}) \to 0$ in probability. Thus the classical estimate with Z's substituted for C's, while not the same function as that obtained by substituting Y's for C's, has the same asymptotic distribution and hence the same minimum asymptotic variance in the class R.

The above result also throws some new light on the estimates proposed in [6] for the case of complete blocks designs. There the observations are given by (1), and the incompatible estimates

$$Y_{ij} = \operatorname{med}_{1 \leqq \alpha \leqq \beta \leqq n} [(X_{i\alpha} - X_{j\alpha} + X_{i\beta} - X_{j\beta})/2] \tag{8}$$

are adjusted by minimizing the sum of squares $\sum_{i \neq j} [Y_{ij} - (\xi_i - \xi_j)]^2$ to obtain compatible estimates

$$Z_{ij} = Y_{i\cdot} - Y_{j\cdot}\,. \tag{9}$$

We now see that the estimates Z_{ij} have the following optimum property:

Corollary 1. *Let* (1) *be the model for the observations in a complete blocks design and let* Y_{ij} *and* Z_{ij} *be defined by* (8) *and* (9), $1 \leqq i, j \leqq c$ [6]. *In the class of all linear functions of the random variables* Y_{ij} *which are unbiased estimates of* $\xi_k - \xi_l$, Z_{kl} *uniformly minimizes the asymptotic variance, and, if* $\lambda(F) < \frac{7}{24}$, *is the unique estimate in this class which does so.*

Proof. In the preceding theorem let $b = c$, $J = 1$.

3. Asymptotic distribution and efficiency of the estimates. Theorem 1 shows that, in the case of a balanced design, among all linear functions of the Y's which are unbiased estimates of a treatment difference the largest asymptotic efficiency (in the sense of the reciprocal of the ratios of the variances) relative to the least squares estimate is achieved by putting Z^j_{st} for C^j_{st} in the classical estimate. Although this optimum property has not been proved in the unbalanced case, this suggests the corresponding estimate

$$Z_{kl} = \sum_{s,t,j} A^{kl}_{stj} Z^j_{st}\,, \tag{10}$$

where the A's are given by (4), for the general (not necessarily balanced) case. The efficiency of the results justifies the choice of the estimate (10). We shall now find the efficiency of (10) in the general case.

Lemma 2. *Under the assumptions about* g *of Lemma* 1, *the random variables* $n^{\frac{1}{2}}(Z^j_{st} - (\xi_s - \xi_t))$, $s, t \in S_j$, $j = 1, \cdots, J$, *are asymptotically joint normal with mean zero and covariance matrix* $(\sigma_{jst,j'uv})$ *given by*

$$\begin{aligned} \sigma_{jst,j'uv} &= 0 && \text{if} \quad j \neq j' \text{ or if } j = j' \text{ and } s, t, u, v \text{ distinct} \\ &= S^2/2\rho_j && \text{if} \quad j = j' \text{ and } s = u \text{ or } t = v \\ &= -S^2/2\rho_j && \text{if} \quad j = j' \text{ and } s = v \text{ or } t = u \end{aligned} \tag{11}$$

$$= S^2/\rho_j \qquad \textit{if} \qquad j = j', s = u, t = v$$

where $S^2 = (2/b)[\frac{1}{12} + (b - 2)(\lambda(F) - \frac{1}{4})]/(\int g^2(x)\,dx)^2$.

PROOF. The joint asymptotic normality follows from Lemma 1, since the Z^j_{st} are linear functions of the Y^j_{st}, $s, t \varepsilon S_j$, $j = 1, \cdots, J$. The variances and covariances are found by direct computation from those of the Y's, exactly as in Lemma 8 of [6].

Let Σ_Z be the covariance matrix (11) and let Σ_C be the covariance matrix of the random variables $n^{\frac{1}{2}}(C^j_{st} - (\xi_s - \xi_t))$, $s, t \varepsilon S_j$, $j = 1, \cdots, J$. Then $\Sigma_C = (\tau^2/S^2)\Sigma_C$. Hence we have from Lemma 2

LEMMA 3. *The two sets of random variables* $\{(\tau/S)n^{\frac{1}{2}}(Z^j_{st} - (\xi_s - \xi_t)), s, t \varepsilon S_j, j = 1, \cdots, J\}$ *and* $\{n^{\frac{1}{2}}(C^j_{st} - (\xi_s - \xi_t)), s, t \varepsilon S_j, j = 1, \cdots, J\}$ *have the same limiting distribution.*

Lemma 3 now yields the desired efficiency result:

THEOREM 2. *If* C_{kl}, *as defined by* (4), *is the least squares estimate of* $\xi_k - \xi_l$, *then* (10) *is an unbiased estimate of* $\xi_k - \xi_l$, *and the asymptotic efficiency of* (10) *relative to* (4) *is*

$$(12) \qquad e = \tau^2/S^2 = 12\tau^2(\int g^2(x)\,dx)^2/(24/b)[\tfrac{1}{12} + (b - 2)(\lambda(F) - \tfrac{1}{4})].$$

In general the efficiency is seen to depend on the block size b and to be independent of the number c of treatments. In the case that $b = c$ (hence $J = 1$), the design reduces to a complete blocks design [6] and our results agree with those of [6]. The efficiency e is the same function of b as the efficiency in the case of complete blocks is of c; a table giving values of the efficiency in the complete blocks case for various values of c in the case of F normal is given in [6]. For fixed $\lambda(F) < \frac{7}{24}$, e is an increasing function of b, with a minimum at $b = 2$ of $12\tau^2(\int g^2(x)\,dx)^2$. For fixed $b > 2$, e is a decreasing function of $\lambda(F)$ bounded below by its value $12\tau^2(\int g^2(x)\,dx)^2$ at $\lambda = \frac{7}{24}$. Thus, for all b, $e \geqq 12\tau^2(\int g^2(x)\,dx)^2$, which is $\geqq .864$ for all distributions F and $= 3/\pi$ if F is normal [2].

LEMMA 4. *The random variables* Z_{kl}, $1 \leqq k, l \leqq c$, *form a compatible set of estimates, in the sense that* $Z_{kl} + Z_{lm} = Z_{km}$.

PROOF. Considering each replication set as a complete blocks design, we have from [6] that for each j the random variables $\{Z^j_{kl}, k, l \varepsilon S_j\}$ form a compatible set. That is, the $\{Z^j_{kl}\}$ satisfy the same linear restrictions as do the $C^j_{kl} = X_{kj\cdot} - X_{lj\cdot}$ for each j. Hence linear combinations of the Z^j_{kl} satisfy the same linear relations as do the same linear combinations of the C^j_{kl}, and thus $C_{kl} + C_{lm} = C_{km}$ implies $Z_{kl} + Z_{lm} = Z_{km}$, i.e. the Z's are compatible.

The efficiency (12) holds not only for differences but extends to the estimation of any contrast. Let $\theta = \sum c_k\xi_k$, $\sum c_k = 0$, be any contrast; θ is a function only of differences of the ξ's. The representation $\theta = \sum\sum d_{kl}(\xi_k - \xi_l)$ is not unique, but the estimate $\theta^* = \sum\sum d_{kl}Z_{kl}$ is independent of the representation. This follows from that compatibility of the Z's and the fact that the classical estimate of any contrast is, by the Gauss-Markov theorem ([7] p. 14) independent of the representation of that contrast. That is, if $\sum\sum d_{kl}(\xi_k - \xi_l) = \sum\sum e_{kl}(\xi_k - \xi_l)$, then $\sum\sum d_{kl}(\hat{\xi}_k - \hat{\xi}_l) = \sum\sum e_{kl}(\hat{\xi}_k - \hat{\xi}_l)$, i.e. $\sum\sum d_{kl}C_{kl} = \sum\sum e_{kl}C_{kl}$. But the last equality together with the fact that the Z's satisfy the same linear relations as the C's yields the desired equality $\sum\sum d_{kl}Z_{kl} = \sum\sum e_{kl}Z_{kl}$. In order to extend the efficiency result of Theorem 2 to the estimation of any contrast, it suffices to show that the asymptotic covariance matrices Σ_Z and Σ_C are proportional, with the same proportionality factor as that relating the asymptotic variances of $n^{\frac{1}{2}}Z_{kl}$ and $n^{\frac{1}{2}}C_{kl}$.

The following lemma enables us to obtain from the fact that the estimates Z_{kl} and C_{kl}, $1 \leqq k, l \leqq c$ have proportional asymptotic variances that they have proportional asymptotic covariance matrices.

LEMMA 5. *Let V_{ij} and W_{ij}, $1 \leqq i, j \leqq c$ be two sets of random variables with covariance matrices Σ_V and Σ_W and such that each set forms a compatible system of estimates of a set of differences $\xi_i - \xi_j$. If, for all i, j,* Var $(V_{ij}) = K$ Var (W_{ij}) *for some constant K, then $\Sigma_V = K\Sigma_W$.*

PROOF. We shall show that for all i, j, k, l

$$(13) \quad \operatorname{Cov}(W_{ij}, W_{kl}) = \tfrac{1}{2}[\operatorname{Var}(W_{il}) - \operatorname{Var}(W_{ik}) + \operatorname{Var}(W_{jk}) - \operatorname{Var}(W_{jl})].$$

The assertion of the lemma will follow from (13). Since $W_{ij} + W_{jk} = W_{ik}$, $\operatorname{Var}(W_{ij}) + \operatorname{Var}(W_{jk}) + 2\operatorname{Cov}(W_{ij}, W_{jk}) = \operatorname{Var}(W_{ik})$ and

$$(14) \quad \operatorname{Cov}(W_{ij}, W_{jk}) = \tfrac{1}{2}[\operatorname{Var}(W_{ik}) - \operatorname{Var}(W_{ij}) - \operatorname{Var}(W_{jk})].$$

Now (13) is obtained by a similar computation from the relation $\operatorname{Var}(W_{il}) = \operatorname{Var}(W_{ij} + W_{jk} + W_{kl})$ with the use of (14).

The joint asymptotic distribution of the random variables $n^{\frac{1}{2}}(Z_{kl} - (\xi_k - \xi_l))$, $1 \leqq k, l \leqq c$, is now given by the following lemma.

LEMMA 6. *The random variables $n^{\frac{1}{2}}(Z_{kl} - (\xi_k - \xi_l))$, $1 \leqq k, l \leqq c$, have a joint normal limiting distribution with zero mean and covariance matrix $\Sigma_Z = (S^2/\tau^2)\Sigma_C$, where Σ_C is the covariance matrix of the random variables $n^{\frac{1}{2}}(C_{kl} - (\xi_k - \xi_l))$.*

PROOF. The joint asymptotic normality follows from Lemma 2, since the Z_{kl} are linear functions of the Z_{st}^j. From Theorem 2, the ratio of the asymptotic variances of $n^{\frac{1}{2}}Z_{kl}$ and $n^{\frac{1}{2}}C_{kl}$ is S^2/τ^2, hence by Lemma 5, $\Sigma_Z = (S^2/\tau^2)\sum_C$.

COROLLARY 2. *Let $\theta = \sum c_i\xi_i = \sum\sum d_{kl}(\xi_k - \xi_l)$ be any contrast. The asymptotic efficiency of the estimate $\sum\sum d_{kl}Z_{kl}$ relative to the classical estimate $\sum\sum d_{kl}C_{kl}$ is $e = \tau^2/S^2$.*

4. Extensions and acknowledgment. Tests of the hypothesis that all the treatment effects are zero can be obtained from the estimates Z_{kl} of $\xi_k - \xi_l$ using the two approaches suggested in [5]. Also, it is easy to extend the above estimates and tests to certain Model II designs (one-way layout and nested designs). In this extension the distribution of the errors may be any unspecified continuous distribution, but the factors of interest are assumed normally distributed. Such an assumption may not be unreasonable in cases where the errors are more likely to be subject to gross errors than are the factors of interest Details of these extensions are given in [1].

I would like to express my deepest gratitude to Professor E. L. Lehmann, whose guidance and encouragement made this work possible.

REFERENCES

[1] GREENBERG, V. L. (1964). Robust inference in some experimental designs. Unpublished dissertation, University of California, Berkeley.

[2] HODGES, J. L., JR. and LEHMANN, E. L. (1956). The efficiency of some nonparametric competitors of the t-test. *Ann. Math. Statist.* **27** 324–335.

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

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

[5] LEHMANN, E. L. (1963). Asymptotically nonparametric inference: an alternative approach to linear models. *Ann. Math. Statist.* **34** 1494–1506.

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

[7] SCHEFFÉ, HENRY (1959). *The Analysis of Variance.* Wiley, New York.