ASYMPTOTICALLY NONPARAMETRIC INFERENCE IN SOME LINEAR MODELS WITH ONE OBSERVATION PER CELL[1]

By E. L. Lehmann

University of California, Berkeley

1. Introduction and summary. In [5] and [6] the author proposed point estimates, tests and confidence procedures for the parameters in a linear model, which have the same asymptotic efficiency (relative to the corresponding classical methods) as the Wilcoxon test has relative to the t-test. Here, "asymptotic" refers to the case that the numbers of observations per cell tend to infinity; in practice, they should presumably be at least equal to four.

In the present paper, we shall consider experiments with only one observation per cell. It then of course becomes necessary to impose some restrictions on the parameters of the model. We shall suppose that an experiment, concerned with various factors at several levels, is replicated at different levels of a nuisance factor (i.e. in different "blocks") and that this nuisance factor does not interact with the factors of interest. Nonparametric procedures for this situation were considered earlier by Friedman (1937) and in [3], where, however, only tests were proposed for the hypothesis of no effect of the factors of interest. In the present paper we shall be concerned primarily with estimating arbitrary contrasts in the factors of interest, and also with the problem of testing such contrasts.

Let us assume for the observations $X_{i\alpha}(i = 1, \cdots, c;\ \alpha = 1, \cdots, N)$ the model

$$X_{i\alpha} = \nu + \xi_i + \mu_\alpha + U_{i\alpha}\ (\sum \xi_i = \sum \mu_\alpha = 0) \tag{1.1}$$

where the ξ's are the parameters of interest, the μ's are the effects of the nuisance factor (or factors), and the U's are independently distributed according to a common continuous distribution F. Except for the assumption of equal sample sizes, this agrees with the model of [5] and [6] when the μ's are assumed to be zero. We shall be interested in inference methods concerning the ξ's, and their asymptotic properties as $N \to \infty$.

2. Estimation of a single contrast. Let $\theta = \sum c_i\xi_i(\sum c_i = 0)$ be any contrast in the ξ's, and let $V_\alpha = \sum c_i X_{i\alpha}$. Then $V_1, \cdots, V_N$ are independently, identically distributed; let their common distribution function be G_c. A possible

Received 23 August 1963.

[1] This work was begun while the author was a Professor of the Adolph C. and Mary Sprague Miller Institute for Basic Research in Science, University of California, Berkeley, and was completed with the partial support of the Office of Naval Research, Contract Nonr-222-(43). This paper in whole or in part may be reproduced for any purpose of the United States Government.

J. Rojo (ed.), *Selected Works of E. L. Lehmann*, Selected Works in Probability and Statistics, DOI 10.1007/978-1-4614-1412-4_49, © Springer Science+Business Media, LLC 2012

estimate of θ is that proposed in [4], namely

$$\text{med } [\tfrac{1}{2}(V_\alpha + V_\beta)] \tag{2.1}$$

where the median is taken over all $\alpha \leqq \beta$. Suppose that G_c is symmetric (as will of course in particular be the case when F is symmetric). It then follows from the results of [4] that this estimate is symmetric about θ, and that its asymptotic efficiency relative to the classical estimate $\sum c_i X_{i\cdot}$ (with $X_{i\cdot} = \sum X_{i\alpha}/N$) is $12\tau_c^2(\int g_c^2(x)\,dx)^2$ where g_c is the density and τ_c^2 the variance of G_c. This efficiency is therefore always $\geqq .864$ and is $3/\pi$ when F (and hence G_c) is normal.

As in [5], the estimates (2.1) for different contrasts are incompatible, and it may be desirable to replace them by a mutually compatible system. Such a system can, as in [5], be based on the estimates of the differences $\xi_i - \xi_j$ which by specialization of (2.1) are given by

$$Y_{ij} = \text{med } [\tfrac{1}{2}(X_{i\alpha} - X_{j\alpha} + X_{i\beta} - X_{j\beta})]. \tag{2.2}$$

We shall denote the common distribution of the difference $X_{i\alpha} - X_{j\beta}$ of any two X's by G, and note that G is symmetric without any assumption about F. As basis for a discussion of a compatible set of estimates based on the Y's we require the joint asymptotic distribution of the Y's, which is given by the following theorem.

THEOREM 1. *If the density g of G satisfies the regularity conditions of Lemma* 3(a) *of* [2], *the joint limiting distribution of the random variables Y_{ij} is the $\binom{c}{2}$-variate normal distribution with zero mean and covariance matrix $\Sigma^* = (\sigma^*_{ij,kl})$ where the variances are given by*

$$\sigma_{ij,ij} = 1/12(\int g^2(x)\,dx)^2 \qquad \text{for all } i, j \tag{2.3}$$

and the covariances by

$$\begin{aligned} \sigma_{ij,kl} &= 0, && \text{if } i, j, k, l \text{ are distinct,} \\ &= [\lambda(F) - \tfrac{1}{4}]/(\int g^2(x)\,dx)^2, && \text{if } i = k \text{ or } j = l, \\ &= [\tfrac{1}{4} - \lambda(F)]/(\int g^2(x)\,dx)^2, && \text{if } i = l \text{ or } j = k, \end{aligned} \tag{2.4}$$

and where

$$\lambda(F) = P(X_1 < X_2 + X_3 - X_4 \text{ and } X_1 < X_5 + X_6 - X_7), \tag{2.5}$$

$X_1, \cdots, X_7$ being independent random variables with distribution F.

PROOF. Let $\binom{N}{2}U^{(ij)}$ be equal to the number of pairs (α, β) with $\alpha < \beta$ and such that

$$X_{j\alpha} - X_{i\alpha} + X_{j\beta} - X_{i\beta} > 2a_{ij}/N^{\frac{1}{2}},$$

and let Q^* denote the $\binom{c}{2}$-variate normal distribution with zero mean and covariance matrix Σ^*. Then the argument leading to (10.3) of [4] and used in proving Theorem 4 of [5], shows that

$$\begin{aligned}(2.6)\qquad &\lim P\{N^{\frac{1}{2}}[Y_{ij} - (\xi_j - \xi_j)] \leqq a_{ij} \quad \text{for all} \quad 1 \leqq i < j \leqq c\} \\ &= \lim P_0\{N^{\frac{1}{2}}[U^{(ij)} - \tfrac{1}{2}] \leqq 0 \quad \text{for all} \quad 1 \leqq i < j \leqq c\},\end{aligned}$$

where P_0 indicates that the probability is being computed under the assumption $\xi_1 = \cdots = \xi_c$.

To evaluate the right hand side of (2.6) we shall use the fact that the theorem on U-statistics stated in the appendix of [5] holds if in Part (iii) the distributions $F, G, H, \cdots$ are independent of N but instead the functions $\phi^{(i)}$ depend on N, provided only these functions are uniformly bounded.

Let $X_\alpha = (X_{1\alpha}, \cdots, X_{c\alpha})$ and define

$$\begin{aligned}\phi_N^{(ij)}(x_\alpha, x_\beta) &= 1 \qquad \text{if} \quad x_{j\alpha} - x_{i\alpha} + x_{j\beta} - x_{i\beta} > 2a_{ij}/N^{\frac{1}{2}} \\ &= 0 \qquad \text{otherwise.}\end{aligned}$$

Then the quantities ζ defined in the appendix of [5] are given by

$$\zeta_N^{(ij)} = \operatorname{Cov}[\phi_N^{(ij)}(X_1, X_2), \phi_N^{(ij)}(X_1, X_3)]$$

which tends to $\zeta^{(ij)} = \frac{1}{12}$, while $\zeta_N^{(ij,kl)}$ tends to

$$\begin{aligned}\zeta^{(ij,kl)} &= 0 \qquad \text{if} \quad i, j, k, l \quad \text{are distinct} \\ &= \lambda(F) - \tfrac{1}{4} \qquad \text{if} \quad i = k \quad \text{or} \quad j = l \\ &= \tfrac{1}{4} - \lambda(F) \qquad \text{if} \quad i = l \quad \text{or} \quad j = k.\end{aligned}$$

Hence, if we put $\theta_N^{(ij)} = EU^{(ij)}$, it follows that the variables $N^{\frac{1}{2}}(U^{(ij)} - \theta_N^{(ij)})$ have a joint normal limiting distribution with zero mean and covariance matrix $(\sigma_{ij,kl}) = 4(\zeta^{(ij,kl)})$. Now

$$\theta_N^{(ij)} = P\{X_{j1} - X_{i1} + X_{j2} - X_{i2} > 2a_{ij}/N^{\frac{1}{2}}\}$$

and a standard argument shows that

$$N^{\frac{1}{2}}(\tfrac{1}{2} - \theta_N^{(ij)}) \to 2a_{ij} \int g^2(x)\, dx.$$

This proves that the right hand side or (2.6) tends to

$$Q(2a_{12} \int g^2(x)\, dx, 2a_{13} \int g^2(x)\, dx, \cdots, 2a_{c-1c} \int g^2(x)\, dx)$$

where Q is the $\binom{c}{2}$-variate normal distribution with zero mean and covariance matrix $(\sigma_{ij,kl})$, and the assertion of the theorem follows.

In the above argument it was tacitly assumed that the covariance matrix Σ^* is non-singular as is typically the case. Should there exist cases in which Σ^* is singular, the argument could be modified along the lines of the proof of Theorem 2 of [5].

For all that follows it is essential to have some knowledge regarding the quantity $\lambda(F)$ defined by (2.5). We therefore prove next the following inequalities.

THEOREM 2. *For all distributions F we have*

$$(2.7)\qquad \tfrac{1}{4} \leqq \lambda(F) \leqq \tfrac{7}{24}.$$

PROOF.

(i) The left inequality is a consequence of the following more general one. Let X_1, X_2, Y_1, Y_2 be independently distributed, the X's with distribution F_X and the Y's with distribution F_Y, then

$$P(X_1 < Y_1 \text{ and } X_1 < Y_2) \geqq P(X_1 < Y_1 \text{ and } X_2 < Y_2). \tag{2.8}$$

This follows immediately from the fact that the inequality (2.8) is equivalent to

$$\int [1 - F_Y(x)]^2 \, dF_X(x) \geqq \{\int [1 - F_Y(x)] \, dF_X(x)\}^2$$

which is a consequence of the Schwarz' inequality. Identifying the left hand side of (2.8) with that of (2.5) with $Y_1 = X_2 + X_3 - X_4$ and $Y_2 = X_5 + X_6 - X_7$, we find that the left hand side of (2.5) is $\geqq [P(X_1 + X_4 < X_2 + X_3)]^2 = \frac{1}{4}$, as was to be proved.

(ii) To prove the right inequality, consider the covariance matrix defined by (2.3) and (2.4) for the random variables Y_{12}, Y_{13}, Y_{23}. Putting $\gamma = 3[4\lambda(F) - 1]$, the determinant of this matrix is seen to be proportional to $(1 + \gamma)^2(1 - 2\gamma)$. This can be nonnegative only if either $\gamma = -1$ and hence $\lambda = \frac{1}{6}$, which is ruled out by (i), or if $\gamma \leqq \frac{1}{2}$ and hence $\lambda \leqq \frac{7}{24}$, as was to be proved.

Three values of $\lambda(F)$ are

F	Normal	Rectangular	Cauchy
$\lambda(F)$	.2902	.2909	.2879

(2.9)

These values are all only slightly below the upper bound $\frac{7}{24} = .2917$.

3. A compatible set of estimates for all contrasts. We have seen above (in the proof of (ii) of Theorem 2) that for any distribution F with $\lambda(F) \neq \frac{7}{24}$, the joint limiting distribution of the three random variables

$$N^{\frac{1}{2}}[Y_{ij} - (\xi_i - \xi_j)], \qquad N^{\frac{1}{2}}[Y_{jk} - (\xi_j - \xi_k)], \qquad N^{\frac{1}{2}}[Y_{ki} - (\xi_k - \xi_i)] \tag{3.1}$$

is nondegenerate. In particular, the variables therefore do not satisfy the condition $N^{\frac{1}{2}}(Y_{ij} + Y_{jk} + Y_{ki}) \to 0$ in probability. They are thus not only incompatible but even asymptotically incompatible in the sense of [5]. As in [5], we may wish to replace the incompatible estimates Y_{ij} of $\xi_i - \xi_j$ by the compatible estimates

$$Z_{ij} = Y_{i\cdot} - Y_{j\cdot} \tag{3.2}$$

where the dot as usual denotes averaging with respect to the indicated subscript. It follows from the asymptotic incompatibility of the Y's that in contrast to the corresponding result of [5], the Z's and the Y's are not asymptotically equivalent. In the present section, we shall investigate the asymptotic behaviour of the Z's and in particular shall prove that asymptotically the Z's are at least as efficient as the Y's for all distributions F.

LEMMA 1. *If the covariance matrix of the variables $N^{\frac{1}{2}}Y_{ij}$ is given by* (2.3) *and* (2.4), *and if we put*

$$\text{(3.3)} \qquad \operatorname{Var}(N^{\frac{1}{2}}Y_{i\cdot}) = \sigma_0^2, \qquad \operatorname{Cov}(N^{\frac{1}{2}}Y_{i\cdot}, N^{\frac{1}{2}}Y_{j\cdot}) = \rho\sigma_0^2,$$

then

$$\text{(3.4)} \qquad \sigma_0^2 = [(c-1)/c^2]\{\tfrac{1}{12} + (c-2)[\lambda(F) - \tfrac{1}{4}]\}/[\int g^2(x)\,dx]^2$$

and

$$\text{(3.5)} \qquad \rho = -1/(c-1).$$

PROOF. An easy calculation shows that

$$\operatorname{Var}(N^{\frac{1}{2}}Y_{i\cdot}) = (c-1)c^{-2}[\operatorname{Var}(N^{\frac{1}{2}}Y_{ij}) + (c-2)\operatorname{Cov}(N^{\frac{1}{2}}Y_{ij}, N^{\frac{1}{2}}Y_{ik})]$$

and (3.4) follows by substituting from (2.3) and (2.4). Equation (3.5) is a consequence of the fact that $\sum Y_i = 0$ and the joint distribution of $(Y_{1\cdot}, \cdots, Y_{c\cdot})$ is symmetric.

THEOREM 3. *Suppose that the density g of G satisfies the regularity conditions of Lemma* 3a *of* [2]. *Then the set of random variables*

$$\text{(3.6)} \qquad N^{\frac{1}{2}}[Z_{ic} - (\xi_i - \xi_c)], \qquad i = 1, \cdots, c-1$$

is asymptotically (as $N \to \infty$) normally distributed with mean zero and covariance matrix $\Sigma = (\sigma_{ij})$,

$$\text{(3.7)} \qquad \sigma_{ii} = 2\sigma_0^2(1-\rho), \qquad \sigma_{ij} = \sigma_0^2(1-\rho)$$

where σ_0^2 and ρ are given by (3.4) *and* (3.5).

PROOF. Since the Z_{ic} are linear functions of the Y's, joint asymptotic normality follows from Theorem 1. That the covariance of the limiting distribution is given by (3.7), is an immediate consequence of Lemma 1.

We are now in position to compute the asymptotic efficiency e^* of Z_{ij} relative to Y_{ij} (in the sense of reciprocal of the ratio of the asymptotic variances).

THEOREM 4. *The efficiency e^* is given by*

$$\text{(3.8)} \qquad e_\lambda^*(c) = c/24\{\tfrac{1}{12} + (c-2)[\lambda(F) - \tfrac{1}{4}]\}$$

and is $\geqq 1$ (i.e. Z_{ij} is at least as efficient as Y_{ij}) for all F.

PROOF. Formula (3.8) follows directly from (2.3) and (3.7). For $c = 2$, the quantity (3.8) is equal to 1, as it should be since in this case $Y_{ij} = Z_{ij}$. For $c > 2$, (3.8) is seen to be $\geqq 1$ since $\lambda(F) \leqq \frac{7}{24}$.

For any fixed $\lambda < \frac{7}{24}$, the efficiency $e_\lambda^*(c)$ is an increasing function of c, and for $\lambda < \lambda'$, we have $e_\lambda^*(c) > e_{\lambda'}^*(c)$ for all $c > 2$. For the values of λ found in (2.9), the variation in e^* with changing c is seen to be slight. In the normal case, for example, e^* increases from $e^*(3) = 1.012$ to $e^*(\infty) = 1.036$. For the asymptotic efficiency e' of Z_{ij} relative to the classical estimate $X_{i\cdot} - X_{j\cdot}$, given by

$$e' = 12r^2(\int g^2(x)\,dx)^2 c/24\{\tfrac{1}{12} + (c-2)[\lambda(F) - \tfrac{1}{4}]\},$$

we have of course a corresponding increase. In particular, in the normal case, we find for e' the values

c	3	5	10	∞
e'	.9664	.9755	.9826	.9893

The efficiency result stated in Theorem 4 holds not only for the estimation of the difference $\xi_i - \xi_j$ but immediately generalizes to the estimation of any contrast $\theta = \sum c_i\xi_i = \sum\sum d_{ij}(\xi_i - \xi_j)$.

COROLLARY 1. *Let $\hat{\theta}_1$ be the estimate of θ defined by* (2.1) *and $\hat{\theta}_2$ the estimate $\sum\sum d_{ij}(Y_{i\cdot} - Y_{j\cdot})$. Then the asymptotic efficiency e^* of $\hat{\theta}_2$ relative to $\hat{\theta}_1$ is given by* (3.8)

PROOF. Let e' and e denote the asymptotic efficiency of $\hat{\theta}_2$ relative to the classical estimate $\hat{\theta}_0 = \sum\sum d_{ij}(X_{i\cdot} - X_{j\cdot})$ and of $\hat{\theta}_1$ relative to $\hat{\theta}_0$ respectively. Then it was seen in Section 2 that $e(F) = 12\tau^2\,(\int g^2(x)\,dx)^2$, independent of the coefficient vector $\mathbf{c} = (c_1, c_2, \cdots)$. On the other hand, since the covariance matrix of the asymptotic distribution of the Z's is proportional to that of the classical estimates $X_{i\cdot} - X_{j\cdot}$, e' is independent of $\mathbf{c}$. Thus $e^* = e'/e$ is also independent of $\mathbf{c}$ as was to be proved.

4. Tests and confidence procedures. So far, we have discussed only the problem of estimating contrasts of the ξ's. The corresponding testing and confidence problems can now be treated quite similarly to their treatment in [6]. Analogously to (3.9) of [6], the classical test of H: $(\xi_1, \cdots, \xi_c)\ \varepsilon\ \pi_w$ concerning contrasts of the ξ's, is based on the statistic $N\sum(X_{i\cdot} - \bar{\xi}_i)^2/\sigma^2$ where π_w is a $(c - r)$-dimensional subspace of c-space and where $(\bar{\xi}_1, \cdots, \bar{\xi}_c)$ is the projection of $(X_{1\cdot}, \cdots, X_{c\cdot})$ on π_w. This statistic can be written in the form $N\sum[\sum a_{ij}(X_{j\cdot} - X_{c\cdot})]^2/\sigma^2$, and under H has a limiting χ^2 distribution with r degrees of freedom. It then follows as in [6] that under H

$$N\sum(\sum a_{ij}Z_{ij})^2/\sigma_0^2 = 12c^2\,(\int g^2(x)\,dx)^2 N\sum(\sum a_{ij}Z_{ij})^2/(c-1)\{1 + 12(c-2)[\lambda(F) - \tfrac{1}{4}]\} \tag{4.1}$$

also has a limiting χ^2-distribution with r degrees of freedom.

We can now apply either of the approaches (i) or (ii) of Section 3 of [6].

(i) Since the variables $(X_{i\alpha} - X_{j\alpha}) - (\xi_i - \xi_j)$ constitute a sample from the symmetric distribution G, the estimate (15) of [7] is a consistent estimate of $1/\int g^2(x)\,dx$. If we denote this estimate by T_{ij}, we can replace the unknown quantity $\int g^2(x)\,dx$ in (4.1) by the average of the T_{ij}'s without changing the asymptotic distribution of the statistic. Clearly, it is necessary also to replace the unknown $\lambda(F)$ by a consistent estimate. Such an estimate, which incidentally is unbiased, is obtained by considering for each sextuple $(i, j, j'; \alpha, \beta, \beta')$ whether or not the simultaneous inequalities

$$X_{i\alpha} < X_{j\alpha} + X_{i\beta} - X_{j\beta}, \quad X_{i\alpha} < X_{j'\alpha} + X_{i\beta'} - X_{j'\beta'}$$

are satisfied. The proportion of cases in which these inequalities hold is the desired estimate. The total number of sextuples will typically be very large, and one may be satisfied with basing this estimate on a subset.

(ii) Suppose that all contrasts in a subspace of the ξ's—of dimension r', say—are zero. Let Q'' be the quadratic form $N\sum[\sum a'_{ij}(X_{j\cdot} - X_{c\cdot})]^2$ appropri-

ate in the classical analysis of variance for testing the hypothesis that these contrasts are indeed zero, and let Q' be obtained from Q'' by substituting Z_{ij} for $X_{i\cdot} - X_{j\cdot}$. Then

$$N \sum (\sum a_{ij} Z_{ij})^2 / r \div Q'/r' \tag{4.2}$$

has, under H, a limiting F-distribution with r and r' degrees of freedom.

As in Section 5 of [6], it is easily seen that the asymptotic efficiency of the test discussed above under (i) relative to the classical F-test is the efficiency e^* given by (3.8) and that the corresponding efficiency of the test based on (4.2) tends to e^* as r' tends to infinity.

The method suggested in [6] for getting additional degrees of freedom for the quadratic form Q' by introducing spurious further parameters and then testing that they are zero, leads to difficulties in the present situation since it changes the dimensionality c of the space of parameters of interest and hence the factor $1 + 12(c - 2)(\lambda - \frac{1}{4})$ in (4.1).

Let us now illustrate the test based on (4.2). Suppose that in each of N blocks, we have a two-way layout, with one observation per cell; that both factors of the layout are of interest but not the block effect; and that we may assume all three factors to be additive. Changing slightly the notation of (1.2), we may then write

$$X_{ij\alpha} = \gamma + \xi_i + \zeta_j + \mu_\alpha \qquad i = 1, \cdots, a; j = 1, \cdots, b; \ \alpha = 1, \cdots, N \tag{4.3}$$

where $\sum \xi_i = \sum \zeta_j = \sum \mu_\alpha = 0$. From (2.2) we get

$$Y_{ij,kl} = \text{med}\,[\tfrac{1}{2}(X_{ij\alpha} - X_{kl\alpha} + X_{ij\beta} - X_{kl\beta})] \tag{4.4}$$

and $Z_{ij,kl} = Y_{ij\cdot\cdot} - Y_{kl\cdot\cdot}$. Substituting the Y's for the X's in the formulae $\hat{\xi}_i = X_{i\cdot\cdot} - X_{\cdots}$, $\hat{\zeta}_j = X_{\cdot j\cdot} - X_{\cdots}$ giving the classical estimates, we find

$$\hat{\xi}_i = Y_{i\cdots} - Y_{\cdots\cdot}, \qquad \hat{\zeta}_j = Y_{\cdot j\cdot\cdot} - Y_{\cdots\cdot} \tag{4.5}$$

where actually $Y_{\cdots\cdot} = 0$. If the interactions γ_{ij} had not been assumed to be zero, their estimates would be $\hat{\gamma}_{ij} = Y_{ij\cdot\cdot} - Y_{i\cdots} - Y_{\cdot j\cdot\cdot} + Y_{\cdots\cdot}$.

For testing the hypothesis H: $\xi_1 = \cdots = \xi_a = 0$, we can then use the test statistic

$$bN \sum \hat{\alpha}_i^2/(a - 1) \div N \sum \sum \hat{\gamma}_{ij}^2/(a - 1)(b - 1)$$

which under H has a limiting F-distribution with $a - 1$ and $(a - 1)(b - 1)$ degrees of freedom.

The classical test of H: $(\xi_1, \cdots, \xi_c) \, \varepsilon \, \Pi_\omega$ is the same whether the block effects μ_α in (1.1) are known to be zero as was assumed in [6] or whether they are unknown. It follows that the quadratic form $N \sum (\sum a_{ij} Z_{ij})^2 = N \sum \sum a_{ij}(Y_{i\cdot} - Y_{j\cdot})^2$ is also formally the same in both cases. Actually, the two statistics are of course different since Y_{ij} is defined by (2.2) in the present case but as med $(X_{i\alpha} - X_{j\beta})$ or as $\text{med}_{\alpha \leqq \beta}\,[\frac{1}{2}(X_{i\alpha} + X_{i\beta})] - \text{med}_{\gamma \leqq \delta}\,[\frac{1}{2}(X_{j\gamma} + X_{j\delta})]$ in [6].

Since the asymptotic efficiency $e^* = e^*(F)$ of the present procedures has been shown to be uniformly higher than the efficiency $e = e(F)$ of the procedures proposed in [6], it is of course tempting to apply the procedures of the present paper also in the case where the μ_α's are zero. The only disadvantages of this approach appear to be (a) the restriction to equal sample sizes, (b) the difficulty in using possible additional degrees of freedom for Q'.

5. Several observations per cell. If instead of just one observation we had many observations in each cell, the ξ's and μ's could be estimated or tested by the methods of [5] and [6]. We shall here consider the case where there are only few, perhaps two or three, observations in each cell and where again we are concerned only with the ξ's. Let the observations be denoted by

$$X_{i\alpha r} = \nu + \xi_i + \mu_\alpha + U_{i\alpha r} \tag{5.1}$$

where $i = 1, \cdots, a$; $\alpha = 1, \cdots, N$; $r = 1, \cdots, m_\alpha$. The numbers of observations in different cells of the same block must therefore be equal but they may differ from block to block.

With the $X_{i\alpha r}$ defined by (5.1), the $N' = \sum m_\alpha$ differences $(X_{i\alpha r} - X_{j\alpha r}) - (\xi_i - \xi_j)$ constitute a sample from the distribution G. If we denote the N' differences $X_{i\alpha r} - X_{j\alpha r}$ by $D_\beta (\beta = 1, \cdots, N')$ and define $Y_{ij} = \text{med}_{\beta \leqq \beta'} [\frac{1}{2}(D_\beta + D_{\beta'})]$ and $Z_{ij} = Y_{i\cdot} - Y_{j\cdot}$, the asymptotic theory established in Sections 2 and 3 applies with N' in place of N.

This method has an unattractive feature, the random pairing of the observations in the $i\alpha$th and $j\alpha$th cell. Such pairing seems not to make the best possible use of the data since it does not utilize all possible differences between observations in the two cells. We shall however now show that the asymptotic efficiency of the proposed procedure relative to the classical procedure is as before given by e^*. This is in fact immediately seen by noting that both the present estimate Y_{ij} and the classical estimate $X_{i\cdot\cdot\cdot} - X_{j\cdot\cdot\cdot}$ of $\xi_i - \xi_j$ (and their distributions) are unchanged if the quantity μ_α in (5.1) is replaced by $\mu_{\alpha r}$. The earlier efficiency results therefore apply with N' instead of N, and since they do not involve the sample size, this establishes the desired result.

One might expect that the efficiency loss due to random pairing would become serious as the number of observations per cell increases. This seems however not to be the case. For consider the case of a single block with a large number of observations in each cell. If the observations in the ith cell are denoted by $X_{ir} (r = 1, \cdots, m)$, pairing leads to the differences $X_{ir} - X_{jr}$, and the estimate $Y_{ij} = \text{med}\, [\frac{1}{2}(X_{ir} - X_{jr} + X_{is} - X_{js})]$ of $\xi_i - \xi_j$. The asymptotic efficiency of this as $m \to \infty$ is $e = 12\tau^2 (\int g^2(x)\, dx)^2$.

It was assumed above that the numbers of observations in different cells of the same block must be equal. What should be done if, for example, in a given block it was intended to have two observations per cell, but the second observation in one of the cells is missing? A simple possibility, which presumably would not seriously affect the performance, consists in simply duplicating the single observation and proceeding as before.

REFERENCES

[1] FRIEDMAN, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Assoc.* **32** 675–701.

[2] 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** 307–317.

[3] HODGES, J. L., JR. and LEHMANN, E. L. (1962). Rank methods for combination of independent experiments in analysis of variance. *Ann. Math. Statist.* **33** 482–497.

[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. (1963). Robust estimation in analysis of variance. *Ann. Math. Statist.* **34** 957–966.

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

[7] LEHMANN, E. L. (1963). Nonparametric confidence intervals for a shift parameter. *Ann. Math. Statist.* **34** 1507–1512.