NONPARAMETRIC INFERENCE IN n REPLICATED 2^m FACTORIAL EXPERIMENTS*

PRANAB KUMAR SEN

(Received Nov. 24, 1969)

For experiments involving m factors $(A_1,\cdots,A_m)$, each at 2 levels $(1,2)$, and replicated in $n(\geqq 2)$ blocks, a class of nonparametric procedures for estimating and testing the various main effects and interactions are considered. The procedures are based on a simple alignment process and involve the use of some well known rank statistics. Their performance characteristics are compared with those of the standard (normal-theory) parametric procedures. Extensions to confounded or partially confounded designs are also considered.

1. Introduction

Let $\boldsymbol{j}=(j_1,\cdots,j_m)$ represent the combination of the levels $j_1,\cdots,j_m$ of the m factors, where $j_k=1,2$, for $k=1,\cdots,m(\geqq 2)$. We denote by $\boldsymbol{J}$ the set of all possible (i.e., 2^m) realizations of $\boldsymbol{j}$. The response $X_{i\boldsymbol{j}}$ of the plot in the ith block receiving the treatment $\boldsymbol{j}$ is expressed as

$$X_{i\boldsymbol{j}}=\beta_i+\frac{1}{2}[\Sigma_S(-1)^{\boldsymbol{j}\boldsymbol{r}'}\tau_{\boldsymbol{r}}]+e_{i\boldsymbol{j}}\,,\qquad \boldsymbol{j}\in\boldsymbol{J},\ \text{and}\ i=1,\cdots,n\,, \tag{1.1}$$

where $\beta_1,\cdots,\beta_n$ represent the block effects, $e_{i\boldsymbol{j}}$'s are the error variables,

$$\boldsymbol{r}=(r_1,\cdots,r_m)\qquad \text{where } r_j \text{ is either 0 or 1, } j=1,\cdots,m, \tag{1.2}$$

the summation S extends over all possible 2^m values of $\boldsymbol{r}$, and the treatment effects $\{\tau_{\boldsymbol{r}}\}$ are defined as follows. We let $\tau_{\boldsymbol{0}}=0$, and

$$\tau_{\boldsymbol{r}}=\tau_{A_1^{r_1}\cdots A_m^{r_m}}\qquad \text{for } \boldsymbol{r}\neq\boldsymbol{0},\ \text{where } A_j^0=0,\ j=1,\cdots,m. \tag{1.3}$$

Thus, $\tau_{A_1}=\tau_{(1,0,\cdots,0)},\cdots,\tau_{A_m}=\tau_{(0,\cdots,0,1)}$ represent the main effects of the m factors, $\tau_{A_1A_2}=\tau_{(1,1,0,\cdots,0)},\cdots,\tau_{A_{m-1}A_m}=\tau_{(0,\cdots,0,1,1)}$ represent the 2-factor (or first order) interactions, and so on; $\tau_{\boldsymbol{r}}$ is a k-factor interaction if $\boldsymbol{r}\boldsymbol{1}_m'=k$,

* Work supported by the Army Research Office, Durham, Grant DA-ARO-D-31-124-G-746, and presented at the 37th session of the International Statistical Institute, held at London, U.K., during September 3-11, 1969.

$1 \leqq k \leqq m$, where $\boldsymbol{l}_m=(1,\cdots,1)$. We denote by

$$\boldsymbol{R}=\{\boldsymbol{r}=(r_1,\cdots,r_m):\ \boldsymbol{r}\neq\boldsymbol{0}\}\,. \tag{1.4}$$

In the normal theory model (cf. Cochran and Cox ([2], Chapter 5) and Kempthorne ([3], Chapter 13)), the usual assumptions are the additivity of the block and treatment effects, and independence, homoscedasticity and normality of the errors. In the nonparametric model, we relax these assumptions considerably. First, we assume that the 2^m random variables $[e_{ij},\, \boldsymbol{j}\in \boldsymbol{J}]$ have jointly a continuous cumulative distribution function (cdf) G_i which is symmetric in its 2^m arguments, for $i=1,\cdots,n$, but $G_1,\cdots,G_n$ may otherwise be quite different from each other. If the errors within the same block are independent and identically distributed, then of course, their joint cdf is symmetric in the 2^m arguments, but the converse is not true. In fact, in many "mixed model" experiments, the symmetry of the joint distribution can be justified in the absense of the block vs treatment interactions, but the same can not be made of the independence (cf. Koch and Sen [4]). Thus, the normality of the distribution of the errors is disposed with, while the independence and homoscedasticity of all the $n2^m$ errors are replaced by the independence of the n sets of 2^m within block errors and the interchangeability of the errors within each block. Further, as $G_1,\cdots,G_n$ may be arbitrarily different, the condition of homoscedasticity is not imposed on errors belonging to different blocks. Second, we need not even assume that the block effects are non-stochastic or additive in nature. By virtue of the first assumption, we may absorb the β_i in the cdf G_i, and write (1.1) equivalently as

$$P[X_{ij}-\Sigma_S(-1)^{\boldsymbol{j}\boldsymbol{r}'}\tau_{\boldsymbol{r}}\leqq x_{\boldsymbol{j}},\ \boldsymbol{j}\in\boldsymbol{J}]=G_i(x_{\boldsymbol{j}},\ \boldsymbol{j}\in\boldsymbol{J})\,,\qquad i=1,\cdots,n\,, \tag{1.5}$$

where $G_i(x_1,\cdots,x_{2^m})$ is symmetric in its 2^m arguments. Thus, excepting the additivity of the treatment effects, all the other assumptions will be relaxed here.

Let now $\boldsymbol{P}$ be a subset of $\boldsymbol{R}$. Then, our first problem is to test the null hypothesis

$$H_{o,\boldsymbol{P}}:\ \tau_{\boldsymbol{r}}=0\,,\qquad \text{for all } \boldsymbol{r}\in\boldsymbol{P}\ (\subset\boldsymbol{R})\,. \tag{1.6}$$

For later convenience, we shall say that $\boldsymbol{P}$ is a monoatomic, diatomic or multiatomic set according as it contains a single, two or more than two elements. Thus, if we want to test for any single $\tau_{\boldsymbol{r}}$, $\boldsymbol{P}$ will be monoatomic, and if we are interested in more than one $\tau_{\boldsymbol{r}}$, $\boldsymbol{P}$ may be diatomic or multiatomic. Our second problem is to provide robust (point as well as interval) estimates of $\tau_{\boldsymbol{r}}$, $\boldsymbol{r}\in R$. Finally, we like to extend these results to the situations where not all the treatment combinations are applied in all the blocks, i.e., we are given some confounded or partially confounded design.

2. Fundamental properties of the aligned observations

Since the model (1.1) involves the block effects as nuisance parameters (or spurious random variables), by means of the following intra-block transformations, we obtain the aligned observations. These aligned observations provide both the least squares and the proposed estimates of τ_r. Let

$$t_{i,r}=2^{-(m-1)}\Sigma_J(-1)^{jr'}X_{ij}\,, \qquad r\in R, \text{ and } i=1,\cdots,n\,, \tag{2.1}$$

where the summation J extends over all $j\in J$. The proposed nonparametric procedures are based on the aligned observations, defined by (2.1). Incidentally, the parametric procedures are also based on these aligned observations ((cf. Cochran and Cox ([2], Chapter 5)). The following lemmas relate to some fundamental properties of these aligned observations and are used in the later sections.

LEMMA 2.1. *If G_i in (1.5) is symmetric in its 2^m arguments, then for each r ($\in R$) the (marginal) distribution of $t_{i,r}$ is symmetric about τ_r.*

PROOF. Let us write

$$g_{i,r}=2^{-(m-1)}\Sigma_J(-1)^{jr'}e_{ij}\,, \qquad r\in R, \text{ and } i=1,\cdots,n\,. \tag{2.2}$$

Then, it follows that

$$t_{i,r}=\tau_r+g_{i,r} \qquad \text{for all } r\in R, \text{ and } i=1,\cdots,n\,. \tag{2.3}$$

Thus, it suffices to show that the marginal distribution of each $g_{i,r}$ is symmetric about 0. Now for any $r\in R$, let J be decomposed into 2 disjoint subsets

$$J_r^{(1)}=\{r\colon jr'=\text{odd}\}\,, \qquad \text{and} \qquad J_r^{(2)}=\{r\colon jr'=\text{even}\}\,. \tag{2.4}$$

Then, we can easily show that $J_r^{(1)}$ and $J_r^{(2)}$ both contain 2^{m-1} elements, and hence,

$$g_{i,r}=\Sigma_{J_r^{(2)}}e_{ij}-\Sigma_{J_r^{(1)}}e_{ij}=U_{i,r}^{(2)}-U_{i,r}^{(1)}\,, \tag{2.5}$$

where $U_{i,r}^{(k)}$ is a linear function of $[e_{ij},\ j\in J]$ with 2^{m-1} of the coefficients equal to 1 and the rest all 0, for $k=1,2$. Now, by the hypothesis of the lemma, the cdf G_i is symmetric in its 2^m arguments, and this implies that the joint cdf of $(U_{i,r}^{(1)}, U_{i,r}^{(2)})$ is also symmetric in its 2 arguments. Hence, the distribution of $U_{i,r}^{(2)}-U_{i,r}^{(1)}$ is symmetric about 0. Q.E.D.

Now, by definition, $t_{i,r}$, $r \in \boldsymbol{R}$ are mutually uncorrelated random variables. But, these are not necessarily independent, unless we impose the normality on G_i. We shall next prove the uncorrelation of skew-symmetric functions of these random variables. Let $h(x)$ be a real valued skew-symmetric function of x, i.e.,

$$h(x)+h(-x)=0\,, \qquad \text{for all } x\,, \tag{2.6}$$

and suppose that

$$E[h^2(g_{i,r})]<\infty\,, \qquad \text{for all } r \in \boldsymbol{R}, \text{ and } i=1,\cdots,n\,. \tag{2.7}$$

LEMMA 2.2. *Under* (1.5), (2.6) *and* (2.7), $E[h(g_{i,r})]=0$ *for all* $r \in \boldsymbol{R}$, *and*

$$E[h(g_{i,r})h(g_{i,s})]=0\,, \qquad \textit{for all } r \neq s \in \boldsymbol{R}\,. \tag{2.8}$$

PROOF. By virtue of Lemma 2.1, the distribution of $g_{i,r}$ is symmetric about 0, and by (2.6), $h(x)$ is skew-symmetric. Hence, it follows readily that $E[h(g_{i,r})]=0$, where the existence of the expectation is insured by (2.7). Now, we say that $\boldsymbol{j}\boldsymbol{r}' \pmod 2 = k$, if $\boldsymbol{j}\boldsymbol{r}'=2l+k$, $k=0, 1$, where l is any non-negative integer. To prove (2.8), we partition $\boldsymbol{J}$ into 4 disjoint subsets

$$J_{rs}^{(k,q)}=\{\boldsymbol{j}\colon \boldsymbol{j}\boldsymbol{r}' \pmod 2 = k,\ \boldsymbol{j}\boldsymbol{s}' \pmod 2 = q\}\,; \qquad k, q=0, 1\,. \tag{2.9}$$

Since the coefficient vectors of $g_{i,r}$ and $g_{i,s}$ are mutually orthogonal and each orthogonal to $(1,\cdots,1)$, it follows readily that each of these subsets contains 2^{m-2} elements ($\boldsymbol{j}$). Let then

$$U_{i,rs}^{(k,q)}=\Sigma_{J_{rs}^{(k,q)}} e_{ij}\,, \qquad k, q=0, 1,\ i=1,\cdots,n\,. \tag{2.10}$$

Then, we may write

$$\begin{aligned} g_{i,r}&=U_{i,rs}^{(1,1)}+U_{i,rs}^{(1,0)}-U_{i,rs}^{(0,1)}-U_{i,rs}^{(0,0)}\,,\\ g_{i,s}&=U_{i,rs}^{(1,1)}-U_{i,rs}^{(1,0)}+U_{i,rs}^{(0,1)}-U_{i,rs}^{(0,0)}\,. \end{aligned} \tag{2.11}$$

Now, by (1.5), the symmetry of the cdf G_i implies that the joint distribution of $[U_{i,rs}^{(k,q)},\ k, q=0, 1]$ is also symmetric in its four arguments. Hence,

$$E[h(g_{i,r})h(g_{i,s})]=E[h(Z_1+Z_2-Z_3-Z_4)h(Z_1-Z_2+Z_3-Z_4)]\,, \tag{2.12}$$

where Z_1, Z_2, Z_3 and Z_4 are interchangeable random variables. Let $Y=[Y_{(1)}\leqq Y_{(2)}\leqq Y_{(3)}\leqq Y_{(4)}]$ be the order statistics associated with $(Z_1,\cdots,Z_4)$. Then, from the symmetric dependence of the Z_i, it follows that the conditional distribution of $Z_1, \cdots, Z_4$, given Y, is uniform over the 24 permutations of the ordered variables among themselves. Thus,

$$E[h(g_{i,r})h(g_{i,s})\,|\,Y] \tag{2.13}$$
$$=\frac{1}{24}\Sigma^* h(Y_{(i_1)}+Y_{(i_2)}-Y_{(i_3)}-Y_{(i_4)})h(Y_{(i_1)}-Y_{(i_2)}+Y_{(i_3)}-Y_{(i_4)}),$$

where the summation Σ^* extends over all possible 24 permutations of $i_1, \cdots, i_4$ over $1, \cdots, 4$. Now, using (2.6), it can be easily verified that the right-hand side of (2.13) is equal to 0, for all Y. Hence, taking expectation over Y, the lemma follows.

Remark. The alignment procedure considered here is a natural extension of the same in Sen [8]. Also, Lemma 2.2 generalizes Lemma 2.1 of Sen [10] to a more complicated situation.

Suppose now that $X=(X_1, \cdots, X_p)$ be a stochastic vector following a continuous cdf $G(x)$, $x \in R^p$, the p-dimensional real space. We say that the cdf $G(x)$ is diagonally symmetric about $\mathbf{0}$, if both X and $(-1)X$ have the same distribution $G(x)$ (for details, see Sen and Puri [12]). It follows from (2.11) and the discussion following it that for any $r \neq s$, $(g_{i,r}, g_{i,s})$ and $(-g_{i,r}, -g_{i,s})$ have the same (bivariate) distribution. Hence, we have the following lemma.

LEMMA 2.3. *Under the condition of Lemma* 2.1, *all the* $\binom{m}{2}$ *bivariate (marginal) distributions of* $(g_{i,r}, g_{i,s})$, $r \neq s$ $(\in \boldsymbol{R})$ *are diagonally symmetric about* $(0, 0)$.

It is to be noted that (1.5) or even the independence and identity of the distributions of the errors within the same block is not sufficient to guarantee the diagonal symmetry of the joint distribution of more than two $g_{i,r}$. This, we show by means of the following simple counter example. Suppose that $m=2$ and the errors in the same block are independent and identically distributed with a characteristic function $\phi(\theta)$. Consider the three variables

$$g_1 = 2g_{i,10} = e_{i11} + e_{i10} - e_{i01} - e_{i00}$$
$$g_2 = 2g_{i,01} = e_{i11} - e_{i10} + e_{i01} - e_{i00}$$
$$g_3 = 2g_{i,11} = e_{i11} - e_{i10} - e_{i01} + e_{i00}\,.$$

Then, the joint characteristic function of (g_1, g_2, g_3) is

$$\phi(\theta_1, \theta_2, \theta_3) = \phi(\theta_1+\theta_2+\theta_3)\phi(\theta_1-\theta_2-\theta_3)\phi(-\theta_1+\theta_2-\theta_3) \cdot \phi(-\theta_1-\theta_2+\theta_3)\,. \tag{2.14}$$

The joint distribution of (g_1, g_2, g_3) will be diagonally symmetric about

(0, 0, 0), iff $\phi(\theta_1, \theta_2, \theta_3)=\phi(-\theta_1, -\theta_2, -\theta_3)$, for all $(\theta_1, \theta_2, \theta_3)$, and that, in general, it is not true can easily be verified by considering the $\phi(\theta)$, corresponding to some skew distributions, such as the Gamma or the exponential ones. It may be noted that the diagonal symmetry of any set $\{t_{i,r}, r \in \boldsymbol{P}(\subset \boldsymbol{R})\}$ can be established if we strengthen the assumption in (1.5) a little more as follows:

(2.15) the joint distribution of the errors $[e_{ij}, \boldsymbol{j} \in \boldsymbol{J}]$ is not only symmetric in the 2^m arguments but also is diagonally symmetric about **0**.

(2.15) holds, in particular, when e_{ij}, $\boldsymbol{j} \in \boldsymbol{J}$, are all distributed independently and identically according to a distribution which is symmetric about 0. Of course, this is more restrictive than (1.5), but is less restrictive than the assumption of normality of the errors.

LEMMA 2.4. *Under* (1.5) *and* (2.15), *the joint distribution of any set* $\{g_{i,r}, r \in \boldsymbol{P} \subset \boldsymbol{R}\}$ *is diagonally symmetric about the origin.*

PROOF. By the hypothesis in (2.15), both the sets $[e_{ij}, \boldsymbol{j} \in \boldsymbol{J}]$ and $[-e_{ij}, \boldsymbol{j} \in \boldsymbol{J}]$ have a common distribution which is symmetric in its 2^m arguments. Now, if in (2.2), we replace the e_{ij} by the corresponding $-e_{ij}$, $\boldsymbol{j} \in \boldsymbol{J}$, we obtain $-g_{i,r}$ for all $\boldsymbol{r} \in \boldsymbol{R}$. Hence, the joint distribution of $[-g_{i,r}, \boldsymbol{r} \in \boldsymbol{R}]$ is the same as that of $[g_{i,r}, \boldsymbol{r} \in \boldsymbol{R}]$. Q.E.D.

LEMMA 2.5. $P[g_{i,r}>0,\ g_{i,s}>0]=1/4$, *for all* $\boldsymbol{r} \neq \boldsymbol{s}$ $(\in \boldsymbol{R})$, *and* $i=1, \cdots, n$.

PROOF. The proof follows from Lemma 2.2 by letting $h(x)$ to be 1, 0 or -1 according as x is $>$, $=$ or <0, and noting that by virtue of Lemma 2.1, the population median of $g_{i,r}$ is equal to 0, for all $\boldsymbol{r} \in \boldsymbol{R}$.

3. Nonparametric tests for τ_r, $\boldsymbol{r} \in \boldsymbol{P} \subset \boldsymbol{R}$

In the normal theory model, the total corrected sum of squares is partitioned into the various components due to each of the main effects, interactions and errors. Then, the test for any hypothesis is based on the variance-ratio ($\mathcal{F}$-) criterion comparing the mean square due to the hypothesis with the error mean square; the reader is referred to Cochran and Cox ([2], Chapter 5) for details.

In the nonparametric model, we are faced with the following situation. If the subset $\boldsymbol{P}$ is monoatomic, exact (i.e., small sample) as well as large sample tests for H_0: $\tau_r=0$ can be constructed. If $\boldsymbol{P}$ is diatomic, conditionally distribution-free tests can be constructed for small samples. If $\boldsymbol{P}$ is multi-atomic, such conditionally distribution-free tests can be constructed only under the assumption (2.15). However, if $\boldsymbol{P}$ is diatomic or multi-atomic, large sample tests can always be constructed under the assumption (1.5). As such, we shall consider the cases separately.

3.1. *Nonparametric tests for monoatomic* $\boldsymbol{P}$. The problem is to test the null hypothesis H_0: $\tau_r=0$, for some specified $\boldsymbol{r}\,(\in \boldsymbol{R})$. (Actually, we may test for H_0: $\tau_r=\tau_r^0$ (known), by working with the variables $t_{i,r}-\tau_r^0$, instead of the $t_{i,r}$.)

(1) *The sign test.* Under the null hypothesis, by (2.3) and Lemma 2.1, $t_{i,r}$, $i=1,\cdots,n$ are all distributed independently and symmetrically about 0. Let $n(r)$ be the number of positive observations among $t_{1,r},\cdots,t_{n,r}$. Then, we have

$$P[n(r)=k\,|\,H_0\colon \tau_r=0]=\binom{n}{k}2^{-n}, \qquad \text{for } k=0,1,\cdots,n\,. \tag{3.1}$$

Thus, we can use the simple binomial tables to construct a test based on the observed $n(r)$. For large values of n, we may use the asymptotic normality of the random variable $Z_r=2n^{-1/2}[n(r)-n/2]$. The test is known to be consistent and unbiased for any $\tau_r\neq 0$. The asymptotic relative efficiency (A.R.E.) of the test with respect to the standard parametric test is equal to the A.R.E. of the sample median with respect to the sample mean. Since, we are dealing with the situation where the distributions can differ from block to block, the classical results are not directly applicable. However, we may use the results in Sen ([9], Section 3) and claim the same robustness properties as are studied there.

(2) *The general scores test.* Let us arrange (in ascending order of magnitude) the observations $|t_{1,r}|,\cdots,|t_{n,r}|$, and let $R_{i,r}$ stand for the rank of $|t_{i,r}|$ in this set, $i=1,\cdots,n$. Also, let $S_{i,r}$ be equal to 1 or 0 according as $t_{i,r}$ is positive or not. Finally, let $a_n(i)$ be a single valued function of i $(=1,\cdots,n)$. Two notable forms of this function are as follows: (a) the Wilcoxon scores, where $a_n(i)=i/(n+1)$, $i=1,\cdots,n$, and (b) the normal scores, where $a_n(i)$ is the expected value of the ith smallest observation in a sample of size n from a chi-distribution with 1 degree of freedom (d.f.), $i=1,\cdots,n$. In general, we may work with suitable $[a_n(1),\cdots,a_n(n)]$ satisfying the regularity conditions of Chernoff and Savage, which have been studied in detail in a previous paper (Puri and Sen [6]). For simplicity of presentation, we shall consider only the cases of the Wilcoxon and the normal scores and refer to Puri and Sen [6], which may be used to extend the results to the case of the general scores. We now define the statistic

$$Q(r)=n^{-1}\Sigma_{i=1}^{n}a_n(R_{i,r})[S_{i,r}-1/2]\,, \qquad r\in \boldsymbol{R}\,. \tag{3.2}$$

Under H_0: $\tau_r=0$, $Q(r)$ has 2^n equally likely realizations obtained by considering the 2^n equally likely realizations of $(S_{1,r}, \cdots, S_{n,r})$, where each $S_{i,r}$ can assume the values 0 and 1 with equal probability 1/2. Hence the distribution of $Q(r)$ over the 2^n equally likely realizations of $(S_{1,r}, \cdots, S_{n,r})$ does not depend on the parent distributions $G_1, \cdots, G_n$. Thus the test based on $Q(r)$ is a distribution-free test. For the particular case of the Wilcoxon scores, the usual critical values available in Owen ([5], pp. 325–329) (computed under the assumption that the $t_{i,r}$ have all a common distribution) can still be used in our case where the parent distributions are not necessarily all identical. For large values of n, it follows from the results of Sen ([7], [11]) that the large sample distribution of $W(r)=2n^{1/2}Q(r)/A_n$ (where $A_n^2=n^{-1}\Sigma_{i=1}^n a_n^2(i)$) has closely a standard normal distribution when the null hypothesis holds. The asymptotic relative efficiency of the test based on the Wilcoxon scores with respect to the standard parametric variance-ratio test is the same as that of the Wilcoxon signed-rank test with respect to the Student t-test, and as this has been studied in detail in Sen [7], we omit the discussion here. For the normal scores test, it follows from the results of Sen [11] that the A.R.E. is bounded below by 1 for all $G_1, \cdots, G_n$. This clearly explains the supremacy of the normal scores test over the standard parametric test.

3.2. *Tests for diatomic P.* Here we desire to test the null hypothesis that $\tau_r=\tau_s=0$, for some $r\neq s$ $(\in R)$.

(1) *The sign test* (Chatterjee [1]). We define $C_{rs}^{(1)}$ $(C_{rs}^{(2)})$ as the number of $(t_{i,r}, t_{i,s})$, $i=1, \cdots, n$, for which both the coordinates are positive (negative), and let $D_{rs}^{(1)}$ $(D_{rs}^{(2)})$ be number of observations for which the first (second) coordinate is positive and the other is negative. Then, $C_{rs}=C_{rs}^{(1)}+C_{rs}^{(2)}$ and $D_{rs}=n-C_{rs}=D_{rs}^{(1)}+D_{rs}^{(2)}$ are the number of concordant and discordant observations. If both C_{rs} and D_{rs} are positive, we define

$$T_{rs}=4\left[\left(C_{rs}^{(1)}-\frac{1}{2}C_{rs}\right)^2\Big/C_{rs}+\left(D_{rs}^{(1)}-\frac{1}{2}D_{rs}\right)^2\Big/D_{rs}\right]. \tag{3.3}$$

If C_{rs} is 0 or n, one of the terms in (3.3) is absent. Under the null hypothesis, it follows from the results of Chatterjee [1] that conditioned on C_{rs}, $C_{rs}^{(1)}$ has a binomial distribution with parameters $(C_{rs}, 1/2)$, $D_{rs}^{(1)}$ has a binomial distribution with parameters $(n-C_{rs}, 1/2)$, and $C_{rs}^{(1)}$ and $D_{rs}^{(1)}$ are stochastically independent. Thus, the exact (conditional) null distribution of T_{rs} can be readily traced with the aid of simple binomial tables. For large n, it follows from the results of Chatterjee [1] that the null distribution of T_{rs} can be approximated by a chi-squared distribution with 2 d.f.; omitting details we say that by virtue of our Lemma 2.5, the needed regularity conditions are all satisfied in our case. The test is shown by Chatterjee to be unbiased and consistent for any

$(\tau_r, \tau_s) \neq (0, 0)$. Again, using our Lemma 2.5, and proceeding as in Chatterjee [1], we may prove along the lines of Section 3 of Sen [9] that the A.R.E. of this test with respect to the parametric variance ratio test is the same as in the corresponding sign test for monoatomic $\boldsymbol{P}$.

(2) *The general scores tests.* The discussions of the Wilcoxon and the normal scores tests follow on the same line as in the multiatomic case considered in Section 3.3.

3.3. *Tests for multiatomic $\boldsymbol{P}$.* We want to test the null hypothesis that $\tau_r=0$ for all $\boldsymbol{r} \in \boldsymbol{P}$, where $\boldsymbol{P}$ contains 2 or more elements. Unless dealing with the general scores tests for the diatomic case, we assume that (2.15) holds for the parent distributions, so that the diagonal symmetry is taken to be granted.

(1) *The sign test.* We define C_{rs} as in Section 3.2, and let $h_{rs}=4C_{rs}-1$. Also, let $n(r)$ be defined as in Section 3.1. Let now $\boldsymbol{P}$ contain k ($\geqq 2$) elements $\boldsymbol{r}_j$, $j=1,\cdots,k$. Define then $\boldsymbol{n}(\boldsymbol{r})=(n(r_1),\cdots,n(r_k))'$, and let $\boldsymbol{H}$ be a $k\times k$ matrix whose diagonal elements are all equal to 1 and the off-diagonal elements are given by $h_{r_j r_{j'}}$, $j\neq j'=1,\cdots,k$. By Lemma 2.5, $\boldsymbol{H}$ is positive definite, in probability. In any case, we may work with the generalized inverse of $\boldsymbol{H}$, and denote it by $\boldsymbol{H}^*$. Then, the proposed test statistic is

$$Z_{(\boldsymbol{P})}=\frac{4}{n}\left[\boldsymbol{n}(\boldsymbol{r})-\frac{1}{2}n\boldsymbol{l}_k\right]'\boldsymbol{H}^*\left[\boldsymbol{n}(\boldsymbol{r})-\frac{1}{2}n\boldsymbol{l}_k\right]. \tag{3.4}$$

For $k=2$, (3.4) reduces to (3.3). We denote by $\boldsymbol{S}_{i,\boldsymbol{r}}=(S_{i,r_1},\cdots,S_{i,r_k})$, $i=1,\cdots,n$, where the $S_{i,r}$ are defined as in Section 3.1. Then, the conditional distribution of $Z_{(\boldsymbol{P})}$ over the 2^n (conditionally) equally likely realizations of $(\boldsymbol{S}_{1,\boldsymbol{r}},\cdots,\boldsymbol{S}_{n,\boldsymbol{r}})$ (where each $\boldsymbol{S}_{i,\boldsymbol{r}}$ can only assume the values $(-1)^j\boldsymbol{S}_{i,\boldsymbol{r}}$, $j=0,1$, with probability 1/2) generates the exact conditional distribution of $Z_{(\boldsymbol{P})}$. For large n, it again follows from our Lemma 2.5 and some standard manipulations that the null (conditional) distribution of $Z_{(\boldsymbol{P})}$ can be closely approximated by the chi-square distribution with k d.f. The null hypothesis is rejected when $Z_{(\boldsymbol{P})}$ is larger than its critical value at a specified level of significance. By using Lemma 2.5, it follows that H approaches the identity matrix of order k, in probability, as $n\to\infty$. Thus, some standard manipulations show that the A.R.E. of this test with respect to the parametric variance ratio test is the same as in the monoatomic case.

(2) *The general scores tests.* We define $a_n(i)$ as in Section 3.1, and let R_{i,r_j}, $i=1,\cdots,n$, be defined as in (2) of Section 3.1, for all $j=1,\cdots,k$. Define then

$$v_n(\boldsymbol{r}_j, \boldsymbol{r}_{j'})=(1/n)\Sigma_{i=1}^n a_n(R_{i,r_j})a_n(R_{i,r_{j'}})(S_{i,r_j}-1/2)(S_{i,r_{j'}}-1/2)\,, \tag{3.5}$$

for all $j, j'=1,\cdots,k$. Note that $v_n(\boldsymbol{r}_j, \boldsymbol{r}_j)=(1/4)A_n^2$ for all $j=1,\cdots,k$.

while for $j \neq j'$, the quantities depend on the sample rank matrix. We denote by $\boldsymbol{V}_n$ the $k \times k$ matrix whose elements are given by (3.5), and we denote its generalized inverse by $\boldsymbol{V}_n^*$. Then the proposed test statistic is

$$W_{(P)} = n\boldsymbol{Q}(\boldsymbol{r})'\boldsymbol{V}_n^*\boldsymbol{Q}(\boldsymbol{r}) , \tag{3.6}$$

where $\boldsymbol{Q}(\boldsymbol{r}) = (Q(\boldsymbol{r}_1), \cdots, Q(\boldsymbol{r}_k))'$ and the $Q(\boldsymbol{r}_j)$ are defined as in (3.2). Two special cases of (3.6) are the multivariate signed rank statistic and the multivariate normal scores statistic for which $Q(\boldsymbol{r}_j)$ are defined as in (2) of Section 3.1. Such statistics are studied in details by Sen and Puri [6]. We note that though they considered the permutation distribution theory of such rank statistics for the case of identical distributions, their theory readily extends to non-identical cdfs, as the sign-invariance has nothing to do with the identity of the parent cdfs. Let us denote the marginal cdf of $g_{i,r}$ by $F_i(x)$, $i=1,\cdots,n$, where we note that by definition the cdf F_i is the same for all $\boldsymbol{r} \in \boldsymbol{R}$. Let then $\bar{F}_n = n^{-1}\Sigma_{i=1}^n F_i$. Since, by Lemma 2.1., $F_1, \cdots, F_n$ are all symmetric about 0, so is $\bar{F}_n$. Thus, by definition $\bar{F}_n(x) - 1/2$ is skew-symmetric about 0. Also, let $F_i^*(x, y)$ be the bivariate cdf of $g_{i,r}$, $g_{i,s}$ for any $\boldsymbol{r} \neq \boldsymbol{s} \in R$, $i=1,\cdots,n$, and let $\bar{F}_n^* = n^{-1}\Sigma_{i=1}^n F_i^*$. Then, for the particular case of the Wilcoxon scores, it is easy to verify that for any $\boldsymbol{r}_j \neq \boldsymbol{r}_{j'}$, $v_n(\boldsymbol{r}_j, \boldsymbol{r}_{j'})$ converges in probability to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\bar{F}_n(x) - \frac{1}{2}\right]\left[\bar{F}_n(y) - \frac{1}{2}\right]d\bar{F}_n^*(x, y) , \tag{3.7}$$

which is equal to 0, by our Lemma 2.2. Hence, it follows that the matrix $\boldsymbol{V}_n$ converges in probability to a diagonal matrix whose elements are all given by 1/12. As such, it follows along the lines of Theorem 3.2 that the large sample permutation distribution of the Wilcoxon scores statistic can be approximated by the chi-square distribution with k d.f. Of course, as in (1), for small n, we can evaluate the exact conditional distribution over the 2^n conditionally equally likely realizations of $S_{1,r}, \cdots, S_{n,r}$. The A.R.E. of the Wilcoxon test with respect to the parametric variance ratio test again coincides with the monoatomic case. For the normal scores test, we let $\Phi(x)$ to be the cdf of the standard normal distribution. Then, let $h(x) = \Phi^{-1}[\bar{F}_n(x)]$. Since $\bar{F}_n$ has been shown to be symmetric about 0, and $\Phi(x)$ is skew-symmetric about 0, it follows that $h(x)$ is also skew-symmetric about 0. Then, for the normal scores, it can be shown that $v_n(\boldsymbol{r}_j, \boldsymbol{r}_{j'})$ converges, in probability, to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)h(y)d\bar{F}_n^*(x, y) , \qquad \text{for all } j \neq j' = 1, \cdots, k . \tag{3.8}$$

Again, by Lemma 2.2, (3.8) equals to 0, for all $j \neq j'$. Consequently, it follows that in this case, the matrix V_n converges, in probability, to a diagonal matrix all whose elements are equal to 1/4. As such, again using Theorem 3.2 of Sen and Puri [6], it follows that the permutation distribution of the normal scores statistic converges to a chi-square distribution with k d.f. For small values of n, we may again use the 2^n conditionally equally likely realizations of $S_{1,r},\cdots,S_{n,r}$ to evaluate the exact conditional distribution of $W_{(P)}$. The A.R.E. of the normal scores test with respect to the variance ratio test is bounded below by 1, uniformly in $G_1,\cdots,G_n$.

3.4. *Large sample tests for multiatomic* $\boldsymbol{P}$. The tests in Section 3.3 are conditional in nature and require the assumption (2.15), which is more restrictive than (1.5). We shall now show that for large values of n, we can construct unconditional tests which are valid even under (1.5) and are much simpler in nature. Here also, we want to test the null hypothesis that $\tau_r=0$, for all $\boldsymbol{r}\in\boldsymbol{P}$, where $\boldsymbol{P}$ contains k elements $\boldsymbol{r}_1,\cdots,\boldsymbol{r}_k$.

(1) *The sign test.* We define the vector $\boldsymbol{n}(\boldsymbol{r})$ as in (1) of Section 3.3. It follows by some standard arguments based on our Lemma 2.5 that under the null hypothesis, $n^{-1/2}2[\boldsymbol{n}(\boldsymbol{r})-n\boldsymbol{l}_k/2]$ has asymptotically a multinormal distribution with null mean vector and dispersion matrix $\boldsymbol{I}_k$, the identity matrix of order k. Thus, under the null hypothesis the statistic

$$Z^*_{(P)}=\frac{4}{n}\sum_{j=1}^{k}\left[n(\boldsymbol{r}_j)-\frac{n}{2}\right]^2 \tag{3.9}$$

has asymptotically a chi-square distribution with k d.f. Thus, an unconditional test for the null hypothesis can be based on $Z^*_{(P)}$, using the tail of the chi-square distribution (with k d.f.) as its critical region. This test has the A.R.E. as the other sign tests in Sections 3.1–3.3.

(2) *The general scores tests.* For the case of the Wilcoxon and the normal scores tests, we have shown in Section 3.3 that the corresponding permutation covariance matrix converges in probability to a diagonal matrix as $n\to\infty$. It follows from the results of Sections 4 and 5 of Sen and Puri [6] with extensions along the lines of Sen [7], [11] that even when $G_1,\cdots,G_n$ are not necessarily identical, the unconditional covariance matrices of the Wilcoxon scores or the normal scores statistics are diagonal, when the null hypothesis holds and the conditions of Lemma 2.2 hold. Thus, it can be easily shown along the lines of Sen [11] that the statistic

$$W^*_{(P)}=\frac{4n}{A_n^2}\sum_{j=1}^{k}[Q(\boldsymbol{r}_j)]^2 \tag{3.10}$$

has asymptotically a chi-square distribution with k d.f. when the null hypothesis holds. The A.R.E. of this test with respect to the variance ratio test is the same as that of the test based on $W_{(P)}$. Thus, for the normal scores test, the A.R.E. is bounded below by 1, for all $G_1, \cdots, G_n$.

4. Robust estimation of the main effects and interactions

The conventional estimate of the treatment effect τ_r is the sample average of $t_{i,r}$, $i=1, \cdots, n$, for all $\boldsymbol{r} \in \boldsymbol{R}$. This estimate is known to be sensitive to gross-errors and outlying observations. As in Puri and Sen [6], we may desire to provide robust estimators of τ_r based on suitable rank statistics. For this purpose, we use the rank statistics considered in Section 3. Specifically, we consider the following three estimates.

(1) *The median estimators.* We denote the ordered observations among $t_{1,r}, \cdots, t_{n,r}$ by $t_{(1)r}, \cdots, t_{(n)r}$. Then, if n is odd $(=2n_0+1)$, the sample median is $t_{(n_0+1)r}$, while if n is even $(=2n_0)$, we define it as $[t_{(n_0)r}+t_{(n_0)r}]/2$, for $\boldsymbol{r} \in \boldsymbol{R}$. Then, on using the sign statistics of Section 3.1, and proceeding as in Puri and Sen [6], we obtain the sample median of the $t_{i,r}$ as an estimate of τ_r, for all $\boldsymbol{r} \in \boldsymbol{R}$. The estimate is unbiased and consistent. Moreover, it is very insensitive to fluctuations of the sample extreme values and is robust for gross errors. The A.R.E. of this estimator with respect to the conventional estimator is the same as that of the corresponding tests considered in Section 3.1. Hence, referring to Section 3 of Sen [9], we omit the details here.

(2) *The Wilcoxon scores estimators.* We consider the Wilcoxon signed rank statistics, considered in Section 3.1, and proceeding as in Puri and Sen [6], we obtain the following estimator of τ_r:

$$\tau_r = \underset{1 \leq i \leq i' \leq n}{\text{median}} \{[t_{i,r}+t_{i',r}]/2\}, \qquad \boldsymbol{r} \in \boldsymbol{R}. \tag{4.1}$$

The estimator is also unbiased and consistent. Further, like the median, it is insensitive to the fluctuations of the extreme values and to gross errors. Moreover, its A.R.E. with respect to the conventional estimator can never be less than 86%, while the same can be indefinitely large. This suggests itself to be a very strong competitor of the conventional estimator.

(3) *The normal scores estimators.* We define the normal scores statistics as in Section 3.1. If instead of the observations $t_{i,r}$, $i=1, \cdots, n$, we work with the observations $t_{i,r}-b$, $i=1, \cdots, n$ (where b is any real quantity), it can be shown (cf. Puri and Sen [6]) that there will be an half open interval in b (say, $b_r^{(1)} \leqq b < b_r^{(2)}$) for which the statistic is equal to 0 (or there will be a value of b, say b_r, such that for b less than b_r, the statistic assumes positive values, while for $b \geqq b_r$, it is negative). In the first case, we define $b_r = (b_r^{(1)}+b_r^{(2)})/2$. Then, our proposed estimator

of τ_r is b_r. This estimator is also unbiased and consistent. Further, it is robust for gross errors and insensitive to outliers. The A.R.E. of this estimator with respect to the conventional estimator is bounded below by 1, while it can also be indefinitely large. The only disadvantage of this estimator is that we have to use a trial and error method for the computation of b_r. However, if we use (4.1) as a trial solution, usually the iterative procedure converges quite rapidly. In fact, in Puri and Sen [6], it is shown that this computation is not very serious specially when n is not very large.

So far we have considered the problem of point estimation only. We may also consider the problem of confidence regions for τ_r, based on suitable rank statistics. As these follow on the same line as in Section 4 of Puri and Sen [6], for brevity, the details are omitted.

5. Extensions to confounded or partially confounded designs

Suppose now that the 2^m experiment is conducted in blocks of size $2^{m'}$, $m'=m-p$, $p\geqq 1$. Then, it is well known that within each replicate 2^p-1 of the treatment effects are confounded with the block effects and these are not estimable. Thus, from the n replicates, we obtain $n2^p$ of the intra-replicate $t_{i,r}$ confounded with the block effects. This situation poses no serious problem to our procedure. Suppose that we are able to obtain the unconfounded $t_{i,r}$ (for a fixed r) only from n'_r ($\leqq n$) replicates. Then, we can apply all the results in Section 3.1 and Section 4 with the simple change that n be replaced by n'_r. To apply the results of Sections 3.2 and 3.3 we require that for all $r\in \boldsymbol{P}$, the observations $t_{i,r}$, $r\in \boldsymbol{P}$, are unconfounded in an equal number of blocks. This can be justified in the case of balanced partially confounded designs. Finally, the results of Section 3.4 can also be extended to the confounded case, by replacing n in (3.9) or (3.10) by n'_r and shifting it within the summation.

UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL

REFERENCES

[1] Chatterjee, S. K. (1966). A bivariate sign test for location, *Ann. Math. Statist.*, **37**, 1771-1782.

[2] Cochran, W. G. and Cox, G. M. (1957). *Experimental Designs*, John Wiley & Sons, Inc., Second ed., New York.

[3] Kempthorne, O. (1952). *The Design and Analysis of Experiments*, John Wiley & Sons, Inc., New York.

[4] Koch, G. G. and Sen, P. K. (1968). Some aspects of the statistical analysis of the 'mixed model', *Biometrics*, **24**, 27-48.

[5] Owen, D. B. (1962). *Handbook of Statistical Tables*, Addison-Wesley, Reading, Mass.

[6] Puri, M. L. and Sen, P. K. (1967). On some optimum nonparametric procedures in two-way layouts, *Jour. Amer. Statist. Assoc.*, **62**, 1214-1229.

[7] Sen, P. K. (1968a). On a further robustness property of the test and estimator based on Wilcoxon signed rank statistic, *Ann. Math. Statist.*, **39**, 282-285.

[8] Sen, P. K. (1968b). On a class of aligned rank order tests in two-way layouts, *Ann. Math. Statist.*, **39**, 1115-1124.

[9] Sen, P. K. (1968c). Asymptotic normality of sample quantiles for m-dependent processes, *Ann. Math. Statist.*, **39**, 1724-1730.

[10] Sen, P. K. (1968d). Robustness of some nonparametric procedures in linear models, *Ann. Math. Statist.*, **39**, 1913-1922.

[11] Sen, P. K. (1969). On the distribution of the one-sample rank order statistics, *Proc. First Internat. Symp. on Nonparametric Tech. in Statist. Inf.*, Edited by M. L. Puri, Cambridge Univ. Press. (In press).

[12] Sen, P. K. and Puri, M. L. (1967). On the theory of rank order tests for location in the multivariate one sample problem, *Ann. Math. Statist.*, **38**, 1216-1228.