ON A CLASS OF NONPARAMETRIC TESTS FOR INTERACTIONS IN FACTORIAL EXPERIMENTS

by

Pranab Kumar Sen
University of North Carolina

Institute of Statistics Mimeo Series No. 525

May 1967

Work supported by the National Institute of Health, Public Health Service, Grant GM-12868.

DEPARTMENT OF BIOSTATISTICS

UNIVERSITY OF NORTH CAROLINA

Chapel Hill, N. C.

ON A CLASS OF NONPARAMETRIC TESTS FOR INTERACTIONS IN FACTORIAL EXPERIMENTS*

By

Pranab Kumar Sen
University of North Carolina at Chapel Hill

1. Summary and introduction. This paper deals with a class of permutationally distribution-free aligned rank order tests for interactions in factorial experiments replicated in complete blocks. The asymptotic power-efficiencies of the proposed tests with respect to the classical analysis of variance test are also studied.

Nonparametric analysis of variance tests, available in the literature, mostly relate to one way or two way (without interaction) layouts. Though the approach of Lehmann (1964) (see also Puri and Sen (1966)) can be adapted to construct tests for interactions in factorial experiments, the necessity of avoiding incompatibility of the unadjusted estimates as well as of estimating some functional of the parent distribution (appearing in the expression for the dispersion matrix of the estimators) makes such tests only asymptotically distribution-free and somewhat tedious to apply. In the present paper, the theory of aligned rank order tests based on Chernoff-Savage (1958) type of statistics, developed in Sen (1966b), is further extended to provide suitable tests for interactions in factorial layouts with equal number of observations per cell. Under certain permutational invariance arguments the nonparametric structure of the proposed tests is established. These tests are also free from the other two difficulties mentioned earlier. Further, using a generalization of Chernoff-Savage (1958) theorem on the asymptotic normality of rank order statistics to aligned observations, the asymptotic power-efficiencies of the proposed tests (along with certain bounds are studied.

2. Preliminary notions. We shall consider in detail only the case of replicated two factor experiments with one observation per call and indicate briefly the theory

* Work supported by the National Institute of Health, Public Health Service, Grant GM-12868.

for the case of several factors and/or observations per cell. The chance variable Y_{ijk} associated with the yield of the plot placed in the ith replicate and receiving the combination of the jth variety (or level of the first factor) and the k-th treatment (or level of the second factor), is expressed, in accordance with the usual fixed-effects model, as

$$(2.1) \qquad Y_{ijk} = \mu_i + \nu_j + \tau_k + \eta_{jk} + U_{ijk}, \; i=1,\ldots,n; \; j=1,\ldots,p; \; k=1,\ldots,q,$$

where $\mu_1,\ldots,\mu_n$ stand for the replication-effects, $\nu_1,\ldots,\nu_p$ for the variety-effects, $\tau_1,\ldots,\tau_q$ for the treatment-effects, $\eta_{11},\ldots,\eta_{pq}$ for the variety x treatment interactions, and U_{ijk}'s are the residual error components. In (2.1), we may put

$$(2.2) \qquad \sum_{j=1}^{p} \nu_j=0, \; \sum_{k=1}^{q} \tau_k=0, \; \sum_{k=1}^{q} \eta_{jk}=0, \; j=1,\ldots,p, \text{ and } \sum_{j=1}^{p} \eta_{jk}=0, \; k=1,\ldots,q.$$

It is assumed that $(U_{i11},\ldots,U_{ipq})$, $i=1,\ldots,n$ are independent and identically distributed stochastic vectors having a common continuous (joint) cumulative distribution function (cdf) $F(x_{11},\ldots,x_{pq})$ which is symmetric in its pq arguments; this includes the conventional assumption of independence and identity of distributions of all the npq error components as a particular case. We want to test

$$(2.3) \qquad H_o: \; \underset{\sim}{T} = (\eta_{jk}) = \underset{\sim}{0}^{pxq},$$

against the set of alternatives that $\underset{\sim}{T}$ is non-null. By means of the following intra-block transformations, we eliminate the nuisance parameters ν_j's and τ_k's. Let us consider the p x q matrices

$$(2.4) \qquad \underset{\sim}{Y}_i = (Y_{ijk}), \; \underset{\sim}{U}_i = (U_{ijk}), \; \underset{\sim}{Z}_i = (Z_{ijk}) \text{ and } \underset{\sim}{E}_i = (E_{ijk}), \; i=1,\ldots,n,$$

where we define

(2.5) $$\underset{\sim}{Z}_i = (\underset{\sim}{I}_p - \frac{1}{p}\underset{\sim}{\ell}'_p\underset{\sim}{\ell}_p)\underset{\sim}{Y}_i(\underset{\sim}{I}_q - \frac{1}{q}\underset{\sim}{\ell}'_q\underset{\sim}{\ell}_q),\ i=1,\ldots,n;$$

(2.6) $$\underset{\sim}{E}_i = (\underset{\sim}{I}_p - \frac{1}{p}\underset{\sim}{\ell}'_p\underset{\sim}{\ell}_p)\underset{\sim}{U}_i(\underset{\sim}{I}_q - \frac{1}{q}\underset{\sim}{\ell}'_q\underset{\sim}{\ell}_q),\ i=1,\ldots,n,$$

$\underset{\sim}{I}_t$ being the identity matrix of order t and $\underset{\sim}{\ell}_t$ the (row) t-vector having all the t elements equal to unity, $t \geq 1$. Then from (2.1) through (2.6), we obtain that

(2.7) $$\underset{\sim}{Z}_i = \underset{\sim}{\Upsilon} + \underset{\sim}{E}_i,\ i=1,\ldots,n.$$

In the sequel, we shall work with the nuisance parameter-free model (2.7). Also, we will only consider the case when $p,\ q \geq 3$. If either of them is less than 3, the situation simplifies as follows. Suppose $q=2$, $p > 2$, then from (2.1) and (2.2) we have $\eta_{j1} = -\eta_{j2} = \eta_j$ (say), $j=1,\ldots,p$; thus (2.3) reduces to H_o^*: $\eta_1=\ldots=\eta_p = 0$. Again from (2.5) and (2.6), we have

(2.8) $$Z_{ij1} = -Z_{ij2} = Z_{ij} \text{ (say)},\ e_{ij1} = -e_{ij2} = e_{ij} \text{ (say) for all } i=1,\ldots,n.$$

It follows from lemma 3.1 [to be proved in section 3] that $(e_{i1},\ldots,e_{ip})$ are symmetric dependent or interchangeable random variables for all $i=1,\ldots,n$. Consequently, based on the set of observations $\{Z_{ij},\ j=1,\ldots,p;\ i=1,\ldots,n\}$, the problem of testing H_o in (2.3) reduces to that of testing the interchangeability of $(Z_{i1},\ldots,Z_{ip})$ (for all $i=1,\ldots,n$), against shift alternatives. As such, the results of Sen (1966b) will directly apply, and the details are omitted. If $p=q=2$, we have $\eta_{11} = -\eta_{12} = -\eta_{21} = \eta_{22} = \eta$ (say), and

(2.9) $$Z_{i11} = -Z_{i12} = -Z_{i21} = Z_i \text{ (say)},\ i=1,\ldots,n,$$

and it is also easily seen that the distribution of Z_i is symmetric about η. So the problem of testing H_o in (2.3) reduces to that of testing the symmetry (about zero) of the distribution of Z_i; this is the well known one sample location problem, and hence, is not discussed.

3. The basic permutation principle. We define $\underset{\sim}{U}_i$ and $\underset{\sim}{E}_i$ (i=1,...,n) as in (2.4) and (2.6), and let $F^*(\underset{\sim}{E})$ be the joint cdf of E_i for i=1,...,n. Let $\underset{\sim}{j} = (j_1,\ldots,j_p)$ be any permutation of (1,...,p) and $\underset{\sim}{J}$, the set of all possible (p!) permutations, so that $\underset{\sim}{j} \in \underset{\sim}{J}$. Also let

(3.1) $$\underset{\sim}{I}_p(\underset{\sim}{j}) = (\delta_{ij_k})_{i,k=1,\ldots,p},$$

where δ_{rs} is the usual Kronecker delta. Now for any $\underset{\sim}{j} \in \underset{\sim}{J}$, $\underset{\sim}{I}_p(\underset{\sim}{j})$ is non-singular and has a unique reciprocal $\underset{\sim}{I}_p(\underset{\sim}{j}^*)$ (say), which also belongs to the set $\{\underset{\sim}{I}_p(\underset{\sim}{j}): \underset{\sim}{j} \in \underset{\sim}{J}\}$. Further, it is easily seen that if $\underset{\sim}{I}_p(\underset{\sim}{j}_1)$ and $\underset{\sim}{I}_p(\underset{\sim}{j}_2)$ be defined as in (3.1) for $\underset{\sim}{j}_1 \in \underset{\sim}{J}$, $\underset{\sim}{j}_2 \in \underset{\sim}{J}$, then $\underset{\sim}{I}_p(\underset{\sim}{j}_1)\underset{\sim}{I}_p(\underset{\sim}{j}_2)$ also belongs to $\{\underset{\sim}{I}_p(\underset{\sim}{j}): \underset{\sim}{j} \in \underset{\sim}{J}\}$. Thus the set $\{\underset{\sim}{I}_p(\underset{\sim}{j}): \underset{\sim}{j} \in \underset{\sim}{J}\}$ forms a finite group of elementary transformations. It can also be verified that

(3.2) $$\underset{\sim}{I}_p(\underset{\sim}{j})(\underset{\sim}{I}_p - \frac{1}{p}\underset{\sim}{\ell}'_p\underset{\sim}{\ell}_p)\underset{\sim}{I}_p(\underset{\sim}{j}^*) = \underset{\sim}{I}_p - \frac{1}{p}\underset{\sim}{\ell}'_p\underset{\sim}{\ell}_p \quad \text{if} \quad \underset{\sim}{I}_p(\underset{\sim}{j})\underset{\sim}{I}_p(\underset{\sim}{j}^*) = \underset{\sim}{I}_p.$$

Similarly, let $\underset{\sim}{k} = (k_1,\ldots,k_q)$ be any permutation of (1,...,q); the set of all possible (q!) permutations is denoted by $\underset{\sim}{K}$. A second group of elementary transformation matrices is then defined by

(3.3) $$\{\underset{\sim}{I}_q(\underset{\sim}{k}): \underset{\sim}{k} \in \underset{\sim}{K}\} \quad \text{where } \underset{\sim}{I}_q(\underset{\sim}{k}) = (\delta_{ik_\ell})_{i,\ell=1,\ldots,q}.$$

Let us then define a finite group $\mathcal{G}_i$ of transformations $\{g_i(\underset{\sim}{j},\underset{\sim}{k}): \underset{\sim}{j} \in \underset{\sim}{J}, \underset{\sim}{k} \in \underset{\sim}{K}\}$ by

(3.4) $$\begin{aligned} \underset{\sim}{E}_i(\underset{\sim}{j},\underset{\sim}{k}) &= g_i(\underset{\sim}{j},\underset{\sim}{k})\underset{\sim}{E}_i \\ &= \underset{\sim}{I}_p(\underset{\sim}{j})\underset{\sim}{E}_i\underset{\sim}{I}'_q(\underset{\sim}{k}), \quad \underset{\sim}{j} \in \underset{\sim}{J}, \ \underset{\sim}{k} \in \underset{\sim}{K}, \text{ for } i=1,\ldots,n. \end{aligned}$$

Finally, the group of all the $(p!q!)^n$ transformations in (3.4) is denoted by $\mathcal{G}_n^*$ i.e.,

$$\mathcal{G}_n^* = (\mathcal{G}_1, \ldots, \mathcal{G}_n). \tag{3.5}$$

As before, we denote the cdf's of $\underset{\sim}{U}_i$ and $\underset{\sim}{E}_i$ by F and F*, respectively. Let now $\mathcal{F}$ be the class of all pq-variate continuous cdf's for which the pq variates are interchangeable. By definition (in section 2), $F \in \mathcal{F}$.

LEMMA 3.1 If $F \in \mathcal{F}$, F* is $\mathcal{G}$-invariant.

PROOF. On defining $\underset{\sim}{U}_i$, $\underset{\sim}{I}_p(\underset{\sim}{j})$ and $\underset{\sim}{I}_q(\underset{\sim}{k})$ as in (2.4), (3.1) and (3.3), we let

$$\underset{\sim}{U}_i(\underset{\sim}{j},\underset{\sim}{k}) = \underset{\sim}{I}_p(\underset{\sim}{j})\underset{\sim}{U}_i\underset{\sim}{I}_q'(\underset{\sim}{k}), \quad \underset{\sim}{j} \in \underset{\sim}{J},\ \underset{\sim}{k} \in \underset{\sim}{K}. \tag{3.6}$$

Since $F \in \mathcal{F}$, it remains invariant under row (or column) permutations. Hence, $\underset{\sim}{U}_i(\underset{\sim}{j},\underset{\sim}{k})$ has the same cdf F for all $\underset{\sim}{j} \in \underset{\sim}{J}$, $\underset{\sim}{k} \in \underset{\sim}{K}$. Now, from (2.6), (3.2) and (3.4), we obtain that

$$\begin{aligned}
\underset{\sim}{E}_i(\underset{\sim}{j},\underset{\sim}{k}) &= \underset{\sim}{I}_p(\underset{\sim}{j})\underset{\sim}{E}_i\underset{\sim}{I}_q'(\underset{\sim}{k}) \\
&= \underset{\sim}{I}_p(\underset{\sim}{j})(\underset{\sim}{I}_p - \tfrac{1}{p}\underset{\sim}{\ell}_p'\underset{\sim}{\ell}_p)\underset{\sim}{U}_i(\underset{\sim}{I}_q - \tfrac{1}{q}\underset{\sim}{\ell}_q'\underset{\sim}{\ell}_q)\underset{\sim}{I}_q'(\underset{\sim}{k}) \\
&= \underset{\sim}{I}_p(\underset{\sim}{j})(I_p - \tfrac{1}{p}\underset{\sim}{\ell}_p'\underset{\sim}{\ell}_p)\underset{\sim}{I}_p(\underset{\sim}{j}^*)\underset{\sim}{I}_p(\underset{\sim}{j})\underset{\sim}{U}_i\underset{\sim}{I}_q'(\underset{\sim}{k})\underset{\sim}{I}_q'(\underset{\sim}{k}^*)(\underset{\sim}{I}_q - \tfrac{1}{q}\underset{\sim}{\ell}_q'\underset{\sim}{\ell}_q)\underset{\sim}{I}_q'(\underset{\sim}{k}) \\
&= (\underset{\sim}{I}_p - \tfrac{1}{p}\underset{\sim}{\ell}_p'\underset{\sim}{\ell}_p)\underset{\sim}{U}_i(\underset{\sim}{j},\underset{\sim}{k})(\underset{\sim}{I}_q - \tfrac{1}{q}\underset{\sim}{\ell}_q'\underset{\sim}{\ell}_q)
\end{aligned} \tag{3.7}$$

Thus, the invariance of the cdf of $\underset{\sim}{U}_i(\underset{\sim}{j},\underset{\sim}{k})$ (under $\mathcal{G}_i$) implies the invariance of the cdf $\underset{\sim}{E}_i(\underset{\sim}{j},\underset{\sim}{k})$ under $\mathcal{G}_i$. Hence the lemma.

Let now Z_n^* be the npq-dimensional (Euclidean) space of the sample point $\underset{\sim}{Z}_n^* = (\underset{\sim}{Z}_1,\ldots,\underset{\sim}{Z}_n)$. Then the finite group $(\mathcal{G}_n^*)$ of transformations in (3.4) and (3.5) maps the sample space onto itself, and under H_0 in (2.3), the distribution of $\underset{\sim}{Z}_n^*$ is $\mathcal{G}_n^*$-invariant. Thus, proceeding as in Hoeffding (1952, pp. 169-170), we may prove the existence of similar size α $(0 < \alpha < 1)$ tests for H_0 in (2.3), valid for all $F \in \mathfrak{F}$. These tests are essentially conditional tests based on the consideration of equiprobable all possible row and column permutations of the matrices $\underset{\sim}{Z}_1, \ldots, \underset{\sim}{Z}_n$. Such a conditional test is termed a permutation test. In this paper, we shall study permutation tests based on a celebrated class of rank order statistics due to Chernoff and Savage (1958).

4. Formulation of the rank order tests. Let $c(u)$ be 1, $\frac{1}{2}$ or 0 according as u is $>$, $=$ or < 0, and let

$$(4.1) \qquad R_{ijk} = \tfrac{1}{2} + \sum_{r=1}^{n} \sum_{s=1}^{p} \sum_{t=1}^{q} c(Z_{ijk} - Z_{rst}), \quad i=1,\ldots,n;\ j=1,\ldots,p;\ k=1,\ldots,q;$$

by virtue of the assumed continuity of F, ties among Z_{ijk}'s may be ignored, in probability. We define a sequence of real numbers $\underset{\sim}{S}_N = (S_{N1}, \ldots, S_{NN})$, where

$$(4.2) \qquad S_{N\alpha} = J_N(\alpha/(N+1)), \quad 1 \leq \alpha \leq N;$$

the function J_N is defined as in Chernoff and Savage (1958) and is assumed to satisfy the regularity conditions of theorem 1 of Chernoff and Savage (1958). Let

$$(4.3) \qquad \underset{\sim}{T}_N = (T_{N,jk}); \quad T_{N,jk} = \frac{1}{n} \sum_{i=1}^{n} S_{NR_{ijk}} \quad \text{for } j=1,\ldots,p,\ k=1,\ldots,q.$$

Also, let

$$(4.4) \qquad \underset{\sim}{T}_N^* = (\underset{\sim}{I}_p - \frac{1}{p} \underset{\sim}{\ell}_p' \underset{\sim}{\ell}_p) \underset{\sim}{T}_N (\underset{\sim}{I}_q - \frac{1}{q} \underset{\sim}{\ell}_q' \underset{\sim}{\ell}_q) = (T_{N,jk}^*),$$

$$(4.5)\qquad S^*_{NR_{ijk}} = S_{NR_{ijk}} - \frac{1}{p}\sum_{j=1}^{p} S_{NR_{ijk}} - \frac{1}{q}\sum_{k=1}^{q} S_{NR_{ijk}} + \frac{1}{pq}\sum_{j=1}^{p}\sum_{k=1}^{q} S_{NR_{ijk}},$$

for $i=1,\ldots,n$; $j=1,\ldots,p$, $k=1,\ldots,q$. Then, from (4.3), (4.4) and (4.5), we obtain

$$(4.6)\qquad T^*_{N.\,jk} = \frac{1}{n}\sum_{i=1}^{n} S^*_{NR_{ijk}},\quad j=1,\ldots,p,\ k=1,\ldots,q.$$

We denote by $\mathcal{P}_N$, the permutational (conditional) probability measure induced by the $(p!q!)^n$ equally likely transformations in $\mathcal{G}^*_n$, defined by (3.4) and (3.5). Then, by simple arguments it follows that

$$(4.7)\qquad E\{\underset{\sim}{T}^*_N \mid \mathcal{P}_N\} = \underset{\sim}{0}^{p\times q}\ .$$

Also, let

$$(4.8)\qquad \sigma^2(\mathcal{P}_N) = \frac{1}{n(p-1)(q-1)}\sum_{i=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{q}(S^*_{NR_{ijk}})^2,$$

and let $\otimes$ stand for the symbol for Kronecker product of two square matrices. Then by routine computations, we obtain that

$$(4.9)\qquad n\,E\{\underset{\sim}{T}^*_N \otimes \underset{\sim}{T}^*_N \mid \mathcal{P}_N\} = (\underset{\sim}{I}_p - \frac{1}{p}\underset{\sim}{\ell}'_p\underset{\sim}{\ell}_p) \otimes (\underset{\sim}{I}_q - \frac{1}{q}\underset{\sim}{\ell}'_q\underset{\sim}{\ell}_q)\sigma^2(\mathcal{P}_N).$$

Thus considering the generalized inverse of the pq x pq matrix in (4.9) and employing it to construct a quadratic form in the elements of $\underset{\sim}{T}^*_N$, we derive the following test statistic

$$(4.10)\qquad \mathcal{L}_N = [n/\sigma^2(\mathcal{P}_N)]\sum_{j=1}^{p}\sum_{k=1}^{q}\{T^*_{N,\,jk}\}^2,$$

which is analogous to the classical parametric test based on the variance ratio criterion [cf. Scheffé (1959)].

For small values of n, p and q, the exact permutation distribution of $\mathcal{L}_N$ can be obtained by considering the $(p!q!)^n$ (conditionally) equally likely row and column permutations of the matrices $\underset{\sim}{S}_N^{(i)} = (S_{NR_{ijk}})$, $i=1,\ldots,n$. This procedure becomes prohibitively laborious for large values of n, p or q. For this reason, we consider the following large sample approach.

Let us denote the marginal cdf of Z_{ijk} by $F^*_{[jk]}(x)$ and the joint cdf of $(Z_{ijk}, Z_{ij'k'})$ by $F^*_{[jk;j'k']}(x,y)$ for all $j,j'=1,\ldots,p$; $k,k'=1,\ldots,q$. Let then

$$H(x) = \frac{1}{pq} \sum_{j=1}^{p} \sum_{k=1}^{q} F^*_{[jk]}(x); \tag{4.11}$$

$$H_{10}(x,y) = \frac{1}{qp(p-1)} \sum_{k=1}^{q} \sum_{j'\neq j=1}^{p} F^*_{[jk;j'k]}(x,y); \tag{4.12}$$

$$H_{01}(x,y) = \frac{1}{pq(q-1)} \sum_{j=1}^{p} \sum_{k\neq k'=1}^{q} F^*_{[jk;jk']}(x,y); \tag{4.13}$$

$$H_{11}(x,y) = \frac{1}{p(p-1)q(q-1)} \sum_{j\neq j'=1}^{p} \sum_{k\neq k'=1}^{q} F^*_{[jk;j'k']}(x,y). \tag{4.14}$$

We denote by $J(u) = \lim_{N\to\infty} J_N(u)$: $0 < u < 1$, and define

$$\delta^2 = \int_0^1 J^2(u)\,du - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J[H(x)]J[H(y)]\,d[H_{10}(x,y) + H_{01}(x,y) - H_{11}(x,y)]. \tag{4.15}$$

Then, proceeding as in the proof of theorem 4.2 of Puri and Sen (1966) and omitting the details, we obtain the following theorem.

THEOREM 4.1 Under the conditions of theorem 1 of Chernoff and Savage (1958), $[\sigma^2(\mathcal{P}_N) - \delta^2]$ converges in probability to zero.

Now, if we assume that

$$(4.16) \qquad P\{[J[H(Z_{jk})]-J[H(Z_{j'k})]-J[H(Z_{jk'})]+J[H(Z_{j'k'})]] = \text{Constant}\} < 1,$$

for at least one pair of $j \neq j'$ and $k \neq k'$, then as in theorem 4.1 of Sen (1966b), it can be shown that δ^2, defined by (4.15), is strictly positive. (4.16) will be termed, in the sequel, as the non-degeneracy condition of the cdf F*. The main theorem of this section is the following.

THEOREM 4.2 Under the conditions of theorem 1 of Chernoff and Savage (1958), the permutation distribution of $\mathcal{L}_N$ converges to a chi-square distribution with (p-1)(q-1) degrees of freedom (d.f.).

PROOF. By virtue of (4.7), (4.9) and (4.10), it suffices to show that for any non-null $\underset{\sim}{A} = (a_{jk})$, $Y_n = n^{\frac{1}{2}} \sum_{j=1}^{p} \sum_{k=1}^{q} a_{jk} T^*_{N,jk}$ converges in law (under $\mathcal{P}_N$) to a normal distribution as $n \rightarrow \infty$. Now, using (4.4), (4.5), (4.6) and the first two conditions of theorem 1 of Chernoff and Savage (1958), we write

$$(4.17) \qquad Y_n = n^{-\frac{1}{2}} \sum_{i=1}^{n} Y_{ni} + o_p(1); \; Y_{ni} = \sum_{j=1}^{p} \sum_{k=1}^{q} a^*_{jk} \; J(\frac{R_{ijk}}{N+1}), \; i=1,\ldots,n,$$

where a^*_{jk}'s are linear functions of a_{jk}'s and they satisfy the constraints that $\sum_{k=1}^{q} a^*_{jk} = o$, $j=1,\ldots,p$ and $\sum_{j=1}^{p} a^*_{jk} = o$, $k=1,\ldots,q$. We note that under $\mathcal{P}_N$, Y_{ni} can have only p!q! (conditionally) equally likely values obtained by permuting the rows and columns of the matrix $\underset{\sim}{R}_i = (R_{ijk})_{j=1,\ldots,p,\; k=1,\ldots,q}$, and $Y_{n1},\ldots,Y_{nn}$ are all stochastically independent (under $\mathcal{P}_N$). Thus, it readily follows that $E(Y_{n_i} \mid \mathcal{P}_N) = o$, and

$$E(Y_{ni}^2 \mid \mathcal{P}_N) = \sum_{j=1}^{p} \sum_{k=1}^{q} (a_{jk}^*)^2 \cdot \frac{1}{(p-1)(q-1)} \left\{ \sum_{j=1}^{p} \sum_{k=1}^{q} J^2\left(\frac{R_{ijk}}{N+1}\right) - \right.$$

(4.18)

$$\left. \frac{1}{p} \sum_{k=1}^{q} \left(\sum_{j=1}^{p} J\left(\frac{R_{ijk}}{N+1}\right)\right)^2 - \frac{1}{q} \sum_{j=1}^{p} \left(\sum_{k=1}^{q} J\left(\frac{R_{ijk}}{N+1}\right)\right)^2 + \frac{1}{pq} \left(\sum_{j=1}^{p} \sum_{k=1}^{q} J\left(\frac{R_{ijk}}{N+1}\right)\right)^2 \right\} .$$

Thus, by routine analysis, it follows as in theorem 4.1 that

(4.19) $$\frac{1}{n} \sum_{i=1}^{n} E(Y_{ni}^2 \mid \mathcal{P}_N) \to \delta^2 \sum_{j=1}^{p} \sum_{k=1}^{q} (a_{jk}^*)^2 > 0,$$

where δ^2 is defined by (4.15) and is positive by (4.16). Further, using the growth condition of theorem 1 of Chernoff and Savage (1958), it follows that

(4.20) $$\frac{1}{n} \sum_{i=1}^{n} E(|Y_{ni}|^{2+\delta} \mid \mathcal{P}_N) < \infty \text{ , for some } \delta > 0.$$

Consequently, using the Berry-Esseen theorem (cf. [4 , p. 288]), the asymptotic normality of Y_n follows from (4.19) and (4.20). Hence the theorem.

By virtue of theorem 4.2, an asymptotically size $\alpha (0 < \alpha < 1)$ test for the hypothesis of no interaction may be proposed as follows. If

(4.21) $$\mathcal{L}_N \begin{cases} \geq \chi^2_{(p-1)(q-1),\alpha}, & \text{reject } H_o \text{ in (2.3),} \\ < \chi^2_{(p-1)(q-1),\alpha}, & \text{accept } H_o, \end{cases}$$

where $\chi^2_{t,\alpha}$ is the upper $100\alpha\%$ point of a chi-square distribution having t d.f. .

5. Asymptotic efficiency of the test based on $\mathcal{L}_N$. It can be easily shown that the test in (4.19) is consistent for any non-null $\underset{\sim}{\eta} = (\eta_{jk})$. For the study of

the asymptotic efficiency of the test based on $\mathcal{S}_N$, we shall therefore consider the following sequence of Pitman-alternatives, specified by

$$(5.1) \qquad H_N: \quad \underset{\sim}{\Upsilon} = \underset{\sim}{\Upsilon}_N = N^{-\frac{1}{2}} \underset{\sim}{\Lambda}, \ \underset{\sim}{\Lambda} = (\lambda_{jk}),$$

where λ_{jk}'s are real and finite and they satisfy

$$(5.2) \qquad \sum_{j=1}^{p} \lambda_{jk} = 0, \ k=1,\ldots,q; \ \sum_{k=1}^{q} \lambda_{jk} = 0 \text{ for } j=1,\ldots,p.$$

Thus, under $\{H_N\}$, the cdf of $\underset{\sim}{Z}_i$ (defined by (2.5),) is specified by $F^*(\underset{\sim}{x} - N^{-\frac{1}{2}}\underset{\sim}{\Lambda})$, (where $\underset{\sim}{x}$ is a p x q matrix), and $F^*(\underset{\sim}{x})$ is invariant under the row or column permutations of $\underset{\sim}{x}$. Thus, the univariate marginal cdf $F^*_{[jk]}(x)$ (of $Z_{ijk}-N^{-\frac{1}{2}}\lambda_{jk}$) is independent of (j,k) and is denoted by H(x) [cf. (4.11)]. Similarly, the bivariate marginal cdf $F^*_{[jk,jk']}(x,y)$ will be independent of $(j,k\neq k')$ and is denoted by $H_{01}(x,y)$ [cf. (4.13)], $F^*_{[jk,j'k]}(x,y)$ will be independent of $(j\neq j',k)$ and is denoted by $H_{10}(x,y)$ [cf. (4.12)], and $F^*_{[jk;j'k']}(x,y)$ will be independent of $(j\neq j',k\neq k')$ and is denoted by $H_{11}(x,y)$ [cf. (4.14)].

Now, for arbitary $F^*(\underset{\sim}{x})$, the asymptotic normality of $N^{\frac{1}{2}}(\underset{\sim}{T}_N-\underset{\sim}{\mu}_N)$ (where $\underset{\sim}{\mu}_N = (\mu_{N,jk})$, $\mu_{N,jk} = \int_{-\infty}^{\infty} J[H(x)]dF^*_{[jk]}(x)$, $j=1,\ldots,p$, $k=1,\ldots,q$) can be proved along the same line as in the proof of theorem 5.1 of Sen (1966a) (with direct-extension to the matrix case). We shall specifically consider the case when $\{H_N\}$ holds. For this, we define

$$(5.3) \qquad B(H) = \int_{-\infty}^{\infty} \frac{d}{dx} J[H(x)]dH(x),$$

$$(5.4) \qquad A^2 = \int_0^1 J^2(u)du - [\int_0^1 J(u)du]^2,$$

$$(5.5)\qquad \rho_{ij} = \frac{1}{A^2}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J[H(x)]J[H(y)]dH_{ij}(x,y) - \{\int_0^1 J(u)du\}^2\right],$$

for $(i,j) = (0,1)$, $(1,0)$ and $(1,1)$, where H_{ij}'s are defined earlier. Then, by the same technique as in theorem 5.1 of Sen (1966a), we obtain that under $\{H_N\}$,

$$(5.6)\qquad n^{\frac{1}{2}}E\{\underset{\sim}{T}^*_N | H_N\} = [(pq)^{-\frac{1}{2}}B(H)]\underset{\sim}{\Lambda} + o(1)$$

$$(5.7)\qquad n\,E\{\underset{\sim}{T}^*_N \otimes \underset{\sim}{T}^*_N | H_N\} = (\underset{\sim}{I}_p - \frac{1}{p}\underset{\sim}{\ell}_p'\underset{\sim}{\ell}_p) \otimes (\underset{\sim}{I}_q - \frac{1}{q}\underset{\sim}{\ell}_q'\underset{\sim}{\ell}_q)A^2(1-\rho_{10}-\rho_{01}+\rho_{11})+o(1),$$

where $B(H)$, A^2, ρ_{ij}'s are defined in (5.3), (5.4) and (5.5). Again, using (4.15), theorem 4.1 and some routine computations, it follows that under $\{H_N\}$ in (5.1)

$$(5.8)\qquad \sigma^2(\mathcal{P}_N) \xrightarrow{p} A^2(1-\rho_{10}-\rho_{01}+\rho_{11}).$$

(5.6), (5.7), (5.8) and the asymptotic normality of $n^{\frac{1}{2}}\underset{\sim}{T}_{N,jk}$ lead to the following theorem.

<u>THEOREM 5.1</u> <u>If</u> (i) $\{H_N\}$ <u>in (5.1) holds,</u> (ii) <u>the conditions of theorem 1 of Chernoff and Savage (1958) hold and</u> (iii) <u>the conditions of lemma 7.2 of Puri (1964) hold,</u> $\mathcal{L}_N$, <u>defined by (4.10), has asymptotically a non-central chi-square distribution with (p-1)(q-1) d.f. and the non-centrality parameter</u>

$$(5.9)\qquad \Delta_{\mathcal{L}} = \left[\frac{1}{pq}\sum_{j=1}^{p}\sum_{k=1}^{q}\lambda_{jk}^2\right][B^2(H)/A^2(1-\rho_{10}-\rho_{01}+\rho_{11})].$$

Referring back to the model in (2.1), let σ_u^2 be the variance of U_{ijk} and ρ_u be the correlation between any two U_{ijk}'s belonging to the same block. Let $F_{(p-1)(q-1),(n-1)(pq-1)}$ be the classical analysis of variance ratio test statistic for testing H_o in (2.3) when the parent cdf is assumed to be normal. Then, it can be shown that under $\{H_N\}$ in (5.1), $Q_N = (p-1)(q-1)\ F_{(p-1)(q-1),(n-1)(pq-1)}$ has asymptotically a non-central chi-square distribution with (p-1)(q-1) d.f. and the non-centrality parameter

(5.10)
$$\Delta_Q = [\frac{1}{pq} \sum_{j=1}^{p} \sum_{k=1}^{q} \lambda_{jk}^2]/\sigma_u^2(1-\rho_u).$$

Let now σ_e^2 be the variance of e_{ijk}, defined by (2.6). Then, after some simplifications, we obtain from (2.6) that

(5.11)
$$\sigma_e^2 = [(p-1)(q-1)/pq]\sigma_u^2(1-\rho_u).$$

Consequently, from theorem 5.1, (5.10) and (5.11), we arrive at the following.

<u>THEOREM 5.2</u> <u>When the conditions of theorem 5.1 hold, the asymptotic relative efficiency (A.R.E.) of the $\mathcal{S}_N$-test with respect to the classical analysis of variance test is given by</u>

(5.12)
$$e_{(\{\mathcal{S}_N\},\{Q_N\})} = \frac{pq}{(p-1)(q-1)(1-\rho_{10}-\rho_{01}+\rho_{11})} \left[\sigma_e^2 B^2(H)/A^2\right].$$

We note that σ_e^2 is the variance of the cdf H, and hence the second factor on the right hand side of (5.12) resembles the usual efficiency factor for the well-known Chernoff-Savage (1958) type of test statistics. Also, it follows from lemmas 4.4 and 4.5 of Sen (1966b) that

(5.13)
$$\rho_{10} \geq -1/(p-1), \quad \rho_{01} \geq -1/(q-1),$$

where the equality sign holds iff $J = H^{-1}$ (apart from an additive constant). Thus, from (5.12) and (5.13), we obtain that

$$(5.14) \qquad e_{(\{\mathcal{L}_N\},\{Q_N\})} \geq \frac{1}{\{1 - \frac{1}{pq}[1-(p-1)(q-1)\rho_{11}]\}} \cdot \left[\sigma_e^2 B^2(H)/A^2\right] .$$

This leads to the following corollary.

COROLLARY 5.2.1 A sufficient condition for $e_{(\{\mathcal{L}_N\},\{Q_N\}}$ to be at least as large as $[\sigma_e^2 B^2(H)/A^2]$ is that $\rho_{11} < 1/(p-1)(q-1)$.

We shall now consider two special $\mathcal{L}_N$-statistics, namely, the Normal Scores and Wilcoxon scores statistics. In the first case, $J_N(\frac{\alpha}{N+1})$, defined by (4.2), is the expected value of the α-th smallest observation of a sample of size N from a standard normal distribution, for $\alpha=1,\ldots,N$. In this case, it is well-known [cf. Chernoff and Savage (1958)] that $\sigma_e^2 B^2(H)/A^2$ is greater than or equal to 1, where the equality sign holds only when H is also normal. Thus, the minimum A.R.E. of the normal scores test with respect to the classical analysis of variance test is equal to $1/\{1 - \frac{1}{pq}[1-(p-1)(q-1)\rho_{11}]\} \geq \frac{1}{2-(\frac{1}{p}+\frac{1}{q})} > \frac{1}{2}$. On the other hand, for the class of parent cdf's for which $\rho_{11} \leq 1/(p-1)(q-1)$, the normal scores test will be at least as efficient as the Q_N-test. In particular, if $H(x)$ is normal, $\rho_{11} = 1/(p-1)(q-1)$ and $\sigma_e^2 B^2(H)/A^2 = 1$, so that the normal scores test and the Q_N-test become asymptotically power equivalent. For Wilcoxon scores, $J_N(\frac{\alpha}{N+1}) = \frac{\alpha}{N+1} : 1 \leq \alpha \leq N$. In this case, $[\sigma_e^2 B^2(H)/A^2]$ is known to be greater than or equal to 0.864 for all H. Consequently, the A.R.E. of the Wilcoxon scores test with respect to the Q_N-test is bounded below by

$$(5.15) \qquad 0.864/\{1 - \frac{1}{pq}[1-(p-1)(q-1)\rho_{11}]\} \geq 0.432 .$$

For normal F, it is known that

$$(5.16) \qquad \rho_{10} = \frac{6}{\pi}\,\mathrm{Sin}^{-1}\left(\frac{-1}{2(p-1)}\right),\ \rho_{01} = \frac{6}{\pi}\,\mathrm{Sin}^{-1}\left(\frac{-1}{2(q-1)}\right),\ \rho_{11} = \frac{6}{\pi}\,\mathrm{Sin}^{-1}\left(\frac{1}{2(p-1)(q-1)}\right),$$

and hence from (5.12), we obtain that for normal distributions the A.R.E. of the Wilcoxon-scores test with respect to the Q_N-test is given by

$$(5.17) \qquad \frac{3pq}{\pi(p-1)(q-1)[1+\frac{6}{\pi}\{\mathrm{Sin}^{-1}\frac{1}{2(p-1)}+\mathrm{Sin}^{-1}\frac{1}{2(q-1)}+\mathrm{Sin}^{-1}\frac{1}{2(p-1)(q-1)}\}]} .$$

The following table illustrates the numerical values of (5.17).

q \ p	2	3	4	5	6	7	8	9	10	15	20	∞
2	.955	.966	.965	.963	.962	.961	.960	.960	.959	.958	.957	.955
3		.975	.974	.972	.971	.961	.970	.970	.969	.968	.968	.966
4			.972	.971	.970	.970	.969	.969	.968	.967	.966	.965
5				.970	.969	.968	.968	.967	.967	.966	.965	.963
6					.969	.967	.966	.966	.966	.964	.964	.962
7						.966	.966	.965	.965	.964	.963	.961
8							.965	.964	.964	.963	.962	.960
9								.964	.964	.962	.962	.960
10									.963	.962	.961	.959
15										.961	.960	.958
20											.959	.957
∞												.955

Thus, the efficiency is bounded below by $3/\pi$ and may be as high as 0.974.

If we have more than one observation per cell, we may still work with the aligned observations obtained by making adjustments for row, column and grand means. The permutation argument is essentially the same (with p and q replaced by pr and qr, respectively, r being the number of observations per cell). For more than two factors, summing over a subset of factors, we may arrive at the desired nuisance parameter-free model and proceed as in this paper. For brevity, the details are omitted.

REFERENCES

[1] CHERNOFF, H., and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric tests. Ann. Math. Statist. 29, 972-994.

[2] HOEFFDING, W. (1952). The large sample power of tests based on permutations of observations. Ann. Math. Statist. 23, 169-192.

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

[4] LOEVE, M. (1962). Probability Theory. D. Van Nostrand Co.

[5] PURI, M. L. (1964). Asymptotic efficiency of a class of c-sample test. Ann. Math. Statist. 35, 102-121.

[6] PURI, M. L., and SEN, P. K. (1966a). On a class of multivariate multi-sample rank order tests. Sankhyā, Ser A. 28, 353-376.

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

[8] SEN, P. K. (1966a). On some nonparametric generalizations of Wilks' tests for H_M, H_{VC} and H_{MVC}, I. Inst. Statist. Univ. North Carolina Mimeo Report No. 468, March 1966.

[9] SEN, P. K. (1966b). On a class of aligned rank order tests in two way layouts. Ann. Math. Statist. (Submitted).