The Annals of Mathematical Statistics
1969, Vol. 40, No. 6, 1933–1949

CONTRIBUTIONS TO THE k-SAMPLE PROBLEM: A SYMMETRIC STATISTIC[1]

By M. Schulzer

Utrecht University

0. Summary. The k-sample problem is studied. Given a set of independent observations from k populations, the pooled sample is ordered, and certain predetermined order statistics are chosen to form the endpoints of random intervals. A random vector is formed, whose components represent the number of observations from each sample in each of these intervals. The exact and limiting distributions of this vector are derived, both under the null-hypothesis of a single underlying distribution and under the alternative of unequal distributions for the k populations. This leads to the definition of a test-statistic for the null-hypothesis, and its limiting null-distribution and limiting distribution under a suitable sequence of alternative hypotheses are obtained. Hence a consistent test of the null-hypothesis is determined and shown to be an extension of the Mood-Brown test. Asymptotic efficiencies are calculated for this test relative to a family of common tests for the k-sample problem. To this end, a method is suggested for comparing efficiencies of tests whose limiting distributions under an alternative sequence are non-central chi-square, with unequal degrees of freedom. Some consideration is given to the problem of design for high relative efficiency.

1. Introduction. The context of the k-sample problem is that of testing the hypothesis that k samples have been independently drawn from the same distribution. This arises as a generalization of the well-studied two-sample problem, for which numerous techniques based mostly on ranks or on spacings have been proposed, for example by Dixon [3], Mood [7], Wallis and Kruskal [19], Weiss [21], and Blumenthal [2].

Weiss [20] suggests an extension of his two-sample spacings technique to the k-sample problem, according to which one picks a particular sample and measures the number of observations from this sample in successive intervals whose endpoints are made up of the pooled order statistics from all other samples. The asymmetry involved in deciding which samples should form the random intervals and which should fall within them renders this extension to the k-sample problem somewhat arbitrary.

Mood [7] proposes a two-sample test based on the number of those observations from each sample which are smaller than the median of the pooled sample. Massey [6] indicates how the test might be extended to k-samples, while retaining its attractive feature of symmetry.

Received 19 August 1968; revised 16 April 1969.

[1] Supported by the National Research Council of Canada. This paper represents part of the author's dissertation at the University of Washington, Seattle.

In the present paper, such an extension is made. The choice of the pooled median as reference point is replaced by a set of random points, the selection of which is at the discretion of the experimenter. These reference points are based on the pooled order statistics from all k samples. It is shown that when these points are chosen appropriately, higher power and better asymptotic relative efficiency can be gained.

Thus, suppose $X_{i1}, X_{i2}, \cdots, X_{in_i}$, $i = 1, 2 \cdots, k$ is a set of independent random variables, with $\sum_{i=1}^{k} n_i = N$. Let $F^{(i)}$ be the probability distribution function of X_{ih}, $h = 1, 2, \cdots, n_i$, for each i. Since density functions are used heavily in the sequel, we suppose that each $F^{(i)}$ is an element of the class of absolutely continuous distribution functions. We wish to test the hypothesis

$$H_0{:}F^{(1)} = F^{(2)} = \cdots = F^{(k)}, \tag{1.1}$$

against the class of alternatives consisting of the sets $\{F^{(1)}, F^{(2)}, \cdots, F^{(k)}\}$ which violate (1.1). Let the X_{ih} be arranged in increasing order and relabelled, to give $Z_1 < Z_2 < \cdots < Z_N$. Suppose $Z_0 = -\infty$, $Z_{N+1} = +\infty$.

Let

$$0 = r_0 < r_1 < r_2 < \cdots < r_{p-1} < r_p = N + 1 \tag{1.2}$$

be a preassigned set of non-negative integers. Further, let

$$\begin{aligned} t_j &= r_j - r_{j-1}, \qquad j = 1, 2, \cdots, p-1, \\ t_p &= r_p - r_{p-1} - 1 = N - r_{p-1}. \end{aligned} \tag{1.3}$$

Consider the p random intervals $]Z_{r_{j-1}}, Z_{r_j}]$, $j = 1, 2, \cdots, p$, and define

$$M_{ij} = \sum_{r=1}^{t_j} W_{ij}(r), \qquad i = 1, 2, \cdots, k, \qquad j = 1, 2, \cdots, p, \tag{1.4}$$

where

$$\begin{aligned} W_{ij}(r) &= 1 \quad \text{if the } r\text{th observation in the } j\text{th interval comes from the } i\text{th population,} \\ &= 0 \quad \text{otherwise,} \end{aligned} \tag{1.5}$$

for $r = 1, 2, \cdots, t_j$.

Thus M_{ij} represents the number of X's from the ith population which fall in the random interval $]Z_{r_{j-1}}, Z_{r_j}]$.

Let $\mathbf{M}$ be the pk-dimensional random vector

$$\mathbf{M} = (M_{11}, M_{12}, \cdots, M_{1p}, M_{21}, M_{22}, \cdots, M_{2p}, \cdots, M_{k1}, M_{k2}, \cdots, M_{kp}).$$

2. Exact null and alternative distributions of M. Define the vector $\mathbf{m}$ by

$$\mathbf{m} = (m_{11}, \cdots, m_{1p}, m_{21}, \cdots, m_{2p}, \cdots, m_{k1}, \cdots, m_{kp}),$$

where the m_{ij} are non-negative integers with

$$\begin{aligned} \sum_{j=1}^{p} m_{ij} &= n_i, \qquad i = 1, 2, \cdots, k, \\ \sum_{i=1}^{k} m_{ij} &= t_j, \qquad j = 1, 2, \cdots, p. \end{aligned} \tag{2.1}$$

An expression for $\Pr[\mathbf{M} = \mathbf{m}]$ was given by Massey [6], who obtained first the joint distribution of the M_{ij} and Z_{r_j}, and then integrated the Z_{r_j} over their entire range. The distribution can also be derived as follows:

THEOREM 2.1. *Under the null-hypothesis, subject to restrictions* (2.1)

$$\Pr[\mathbf{M} = \mathbf{m}] = [\textstyle\prod_{j=1}^{p} t_j! \prod_{i=1}^{k} n_i!]/[N! \prod_{j=1}^{p} \prod_{i=1}^{k} m_{ij}!]. \tag{2.2}$$

PROOF. We note that, under H_0, all ordered arrangements of the X_{ih} are equally likely. It follows that, for any j,

$$\Pr[M_{1j} = m_{1j}, \cdots, M_{kj} = m_{kj} \mid \textstyle\sum_{l=1}^{j-1} M_{il}, i = 1, 2, \cdots, k]$$
$$= \binom{N - \sum_{l=1}^{j-1} t_l}{t_j}^{-1} \prod_{i=1}^{k} \binom{n_i - \sum_{l=1}^{j-1} m_{il}}{m_{ij}}.$$

The joint distribution over all p intervals is, therefore,

$$\Pr[\mathbf{M} = \mathbf{m}] = \prod_{j=1}^{p} \left[\binom{N - \sum_{l=1}^{j-1} t_l}{t_j}^{-1} \prod_{i=1}^{k} \binom{n_i - \sum_{l=1}^{j-1} m_{il}}{m_{ij}}\right],$$

which is just (2.2), thus completing the proof.

It may be noted that, for any $i = 1, 2, \cdots, k$, and $j = 1, 2, \cdots, p$, M_{ij} has a marginal hypergeometric distribution under H_0, with $E(M_{ij}) = N^{-1}t_j n_i$, and $\operatorname{Var}(M_{ij}) = [N^2(N-1)]^{-1} t_j n_i (N - t_j)(N - n_i)$. Moreover, from (1.4) and (1.5) it follows that

$$\operatorname{Cov}(M_{i_1 j}, M_{i_2 j}) = -[N^2(N-1)]^{-1} t_j n_{i_1}(N - t_j) n_{i_2},$$
$$\operatorname{Cov}(M_{i j_1}, M_{i j_2}) = -[N^2(N-1)]^{-1} t_{j_1} n_i t_{j_2}(N - n_i),$$

and

$$\operatorname{Cov}(M_{i_1 j_1}, M_{i_2 j_2}) = [N^2(N-1)]^{-1} t_{j_1} n_{i_1} t_{j_2} n_{i_2},$$

where $i_1 \neq i_2$, $j_1 \neq j_2$.

In order to find the exact distribution of $\mathbf{M}$ under the alternative, we suppose $F^{(i)}$ is absolutely continuous, for $i = 1, 2, \cdots, k$, with density $f^{(i)}$. Let H be the mixed distribution function

$$H = \textstyle\sum_{i=1}^{k} N^{-1} n_i F^{(i)} \tag{2.3}$$

of the N random variables $X_{11}, \cdots, X_{1n_1}, \cdots, X_{k1}, \cdots, X_{kn_k}$, and let h be the corresponding density function. Both H and h depend on N.

In analogy with Hoeffding [5], we consider the Euclidean N-space of points

$$x = (x_{11}, x_{12}, \cdots, x_{1n_1}, \cdots, x_{k1}, \cdots, x_{kn_k}).$$

Since all distribution functions are continuous, we lose no generality by assuming all components of x to be distinct. Let $z_x = (z_1, z_2, \cdots, z_N)$ be the vector which corresponds to x with its coordinates arranged in increasing order. Let

$$R = (r_{11}, r_{12}, \cdots, r_{1n_1}, \cdots, r_{k1}, \cdots, r_{kn_k})$$

be an N-vector representing the permutation which must be applied to x to obtain z_x. Thus,

$$x_R = z_x, \qquad z_{r_{ih}} = x_{ih}, \qquad i = 1, 2, \cdots, k, \qquad h = 1, 2, \cdots, n_i.$$

Define

$$S(R) = \{x : x_R = z_x\}. \tag{2.4}$$

Thus $S(R)$ is the set of all N-tuples with distinct entries, whose components x_{ih} have ranks r_{ih} when they are ordered.

Let σ be a permutation defined on R, which in terms of the coordinates of R, is characterized by $\sigma(r_{ih_1}) = r_{ih_2}$, for any i, h_1, h_2, with $i = 1, 2, \cdots, k$, $h_1 = 1, 2, \cdots, n_i$, $h_2 = 1, 2, \cdots, n_i$.

Let τ be a permutation on R, such that, for any i, h, ($i = 1, 2, \cdots, k$, $h = 1, 2, \cdots, n_i$), $r_{j-1} < r_{ih} \leqq r_j$ implies $r_{j-1} < \tau(r_{ih}) \leqq r_j$, and for no value of i from 1 to k is $\tau(r_{ih_1})$ equal to r_{ih_2}, where $h_1 = 1, 2, \cdots, n_i$, $h_2 = 1, 2, \cdots, n_i$, with $h_1 \neq h_2$. Let $(\tau\sigma)'(i, h)$ denote the first subscript of $(\tau\sigma)(r_{ih})$.

Let the symbol $\sum_{\tau(\sigma)}$ denote the sum over all possible composed permutations $\tau(\sigma)$ on R, thus giving rise to $(\prod_{i=1}^k n_i! \prod_{j=1}^p t_j!)/(\prod_{j=1}^p \prod_{i=1}^k m_{ij}!)$ terms. We can now prove

THEOREM 2.2. *Under the alternative hypothesis,*

$$\Pr[\mathbf{M} = \mathbf{m}] = \sum_{\tau(\sigma)} \int_{z_N=-\infty}^{\infty} \cdots \int_{z_2=-\infty}^{z_3} \int_{z_1=-\infty}^{z_2} \prod_{i=1}^{k} \prod_{h=1}^{n_i} f^{(\tau\sigma)'(i,h)}(z_{\tau\sigma r_{ih}})\, dz_1 \cdots dz_N \tag{2.5}$$

whenever $z_1 < z_2 < \cdots < z_N$, *and* 0 *otherwise.*

Alternatively,

$$\Pr[\mathbf{M} = \mathbf{m}] = \sum_{\tau(\sigma)} \frac{1}{N!} E\left[\prod_{i=1}^{k} \prod_{h=1}^{n_i} \frac{f^{(\tau\sigma)'(i,h)}(Z_{\tau\sigma r_{ih}})}{h(Z_{r_{ih}})}\right]$$

where the Z's *are* N *order statistics from* h.

PROOF. With $S(R)$ defined in (2.4), we have for $z_1 < z_2 < \cdots < z_N$,

$$\begin{aligned} P[S(R)] &= \int_{z_N=-\infty}^{\infty} \cdots \int_{z_2=-\infty}^{z_3} \int_{z_1=-\infty}^{z_2} \prod_{i=1}^{k} \prod_{h=1}^{n_i} f^{(i)}(z_{r_{ih}})\, dz_1 \cdots dz_N \\ &= \frac{1}{N!} \int_{z_N=-\infty}^{\infty} \cdots \int_{z_2=-\infty}^{z_3} \int_{z_1=-\infty}^{z_2} \\ &\quad \cdot \prod_{i=1}^{k} \prod_{h=1}^{n_i} \frac{f^{(i)}(z_{r_{ih}})}{h(z_{r_{ih}})} \cdot h(z_{r_{ih}}) \cdot N!\, dz_1 \cdots dz_N \\ &= \frac{1}{N!} E\left[\prod_{i=1}^{k} \prod_{h=1}^{n_i} \frac{f^{(i)}(Z_{r_{ih}})}{h(Z_{r_{ih}})}\right], \end{aligned} \tag{2.6}$$

where $f^{(i)}(z)/h(z)$ is defined to be zero whenever $h(z)$ (and so also $f^{(i)}(z)$) is zero.

Now, to each point x in $S(R)$ there corresponds the same vector

$$\mathbf{m} = (m_{11}, m_{12}, \cdots, m_{1p}, \cdots, m_{k1}, \cdots, m_{kp}).$$

Moreover, for any σ and τ described above, the subsets of N-space $S(\sigma(R))$, $S(\tau(R))$, $S(\tau\sigma(R))$ still correspond to the same vector $\mathbf{m}$. Thus, to obtain $\Pr[\mathbf{M} = \mathbf{m}]$, we must sum together volumes of the form (2.6) over all permutations $\tau(\sigma)$. Hence,

$$\Pr[\mathbf{M} = \mathbf{m}] = \sum\nolimits_{\tau(\sigma)} \frac{1}{N!} E\left[\prod_{i=1}^{k} \prod_{h=1}^{n_i} \frac{f^{(\tau\sigma)'(i,h)}(Z_{\tau\sigma r_{ih}})}{h(Z_{r_{ih}})}\right]. \tag{2.7}$$

REMARK. Introducing the integral transformation H to the pooled data, we can express (2.7) in slightly different form. Define H^{-1} to be left-continuous. Then (2.7) becomes,

$$\Pr[\mathbf{M} = \mathbf{m}] = \sum\nolimits_{\tau(\sigma)} \frac{1}{N!} E\left[\prod_{i=1}^{k} \prod_{h=1}^{n_i} \frac{f^{(\tau\sigma)'(i,h)}(H^{-1}(U_{\tau\sigma r_{ih}}))}{h(H^{-1}(U_{r_{ih}}))}\right],$$

where the U's are uniform variables on [0, 1], and

$$U_{r_{ih}} = H(Z_{r_{ih}}) = H(X_{ih}).$$

3. Limiting normality of M under the alternative. We assume, for the remainder of this paper, that as $N \to \infty$, one has for $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, p$, with t_j as in (1.3),

$$N^{-1}n_i - \pi_i = O(N^{-\frac{1}{2}}), \qquad \lim_{N\to\infty} N^{-1}t_j = \alpha_j, \qquad \pi_i > 0,\ \alpha_j > 0, \tag{3.1}$$
$$\sum\nolimits_{i=1}^{k} \pi_i = \sum\nolimits_{j=1}^{p} \alpha_j = 1.$$

Define

$$\beta_0 = 0, \qquad \beta_j = \sum\nolimits_{l=1}^{j} \alpha_l, \qquad j = 1, 2, \cdots, p. \tag{3.2}$$

Thus

$$\beta_0 = \lim_{N\to\infty} N^{-1}r_0, \qquad \beta_j = \lim_{N\to\infty} N^{-1}r_j, \quad j = 1, 2, \cdots, p, \tag{3.3}$$

with r_j as in (1.2). Let $F_{n_i}^{(i)}$ denote the empirical distribution function for the ith sample and H_N the empirical distribution function for the pooled sample. Let H be as in (2.3), and define

$$H_\pi = \sum\nolimits_{i=1}^{k} \pi_i F^{(i)}, \qquad K_\pi = F^{(i)}H_\pi^{-1}, \tag{3.4}$$

where inverse functions are taken to be left-continuous. Assume K_π has a continuous derivative a_π in a neighborhood of each β_j, $j = 1, 2, \cdots, p - 1$, and that a_π has a left limit at $\beta_p = 1$. We then have,

THEOREM 3.1. *Under the alternative hypothesis and the above assumptions, the vector $N^{-\frac{1}{2}}\mathbf{M}$ has a limiting multivariate normal distribution as $N \to \infty$.*

PROOF. Let $\{W_{n_i}^{(i)}(t): 0 \leqq t \leqq 1\}$ be the empirical process of the ith sample.

Thus,

$$W_{n_i}^{(i)}(t) = n_i^{\frac{1}{2}}(F_{n_i}^{(i)}[F^{(i)^{-1}}(t)] - t).$$

It is well known that this process converges weakly to a separable, tied-down Wiener process $\{W_0{}^{(i)}(t):0 \leqq t \leqq 1\}$, with

$$E(W_0{}^{(i)}(t)) = 0, \qquad E(W_0{}^{(i)}(s)W_0{}^{(i)}(t)) = s(1-t)$$

for $0 \leqq s \leqq t \leqq 1$. (Donsker [4], Prokhorov [10].)

For purposes of this proof we consider especially constructed processes, defined on a single probability space, which converge almost surely in the uniform metric ρ defined on the sample function space (Skorokhod, [18]). A direct construction is given by Root [14]. Thus,

$$\rho(W_{n_i}^{(i)}, W_0{}^{(i)}) \to_{\text{a.s.}} 0. \tag{3.5}$$

Now, from Lemma 2.3 in Pyke and Shorack [13] we have

$$\rho(F^{(i)}H_N^{-1}, F^{(i)}H^{-1}) \to_{\text{a.s.}} 0. \tag{3.6}$$

Then, (3.3) and (3.6) give

$$F^{(i)}[H_N^{-1}(r_j/N)] \to_{\text{a.s.}} F^{(i)}[H_\pi^{-1}(\beta_j)]. \tag{3.7}$$

Hence, by (3.5), we have

$$W_{n_i}^{(i)}(F^{(i)}[H_N^{-1}(r_j/N)]) \to_L W_0{}^{(i)}(F^{(i)}[H_\pi^{-1}(\beta_j)]), \tag{3.8}$$

where the symbol $\to_L$ denotes convergence in law, and where

$$W_{n_i}^{(i)}(F^{(i)}[H_N^{-1}(N^{-1}r_j)]) = n_i^{\frac{1}{2}}(F_{n_i}^{(i)}[H_N^{-1}(N^{-1}r_j)] - F^{(i)}[H_N^{-1}(N^{-1}r_j)]).$$

Then, for any i, j,

$$\begin{aligned}
&n_i^{\frac{1}{2}}\left(F_{n_i}^{(i)}\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right] - F^{(i)}\left[H^{-1}\left(\frac{r_j}{N}\right)\right]\right) \\
&\quad = n_i^{\frac{1}{2}}\left(F_{n_i}^{(i)}\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right] - F^{(i)}\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right]\right) \\
&\qquad + \left(\frac{n_i}{N}\right)^{\frac{1}{2}} \frac{K_\pi\left(H_\pi\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right]\right) - K_\pi\left(H_\pi\left[H^{-1}\left(\frac{r_j}{N}\right)\right]\right)}{H_\pi\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right] - H_\pi\left[H^{-1}\left(\frac{r_j}{N}\right)\right]} \\
&\qquad \cdot \frac{H_\pi\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right] - H_\pi\left[H^{-1}\left(\frac{r_j}{N}\right)\right]}{H\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right] - H\left[H^{-1}\left(\frac{r_j}{N}\right)\right]} \\
&\qquad \cdot N^{\frac{1}{2}}\left(H\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right] - H_N\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right]\right. \\
&\qquad \left. + H_N\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right] - H\left[H^{-1}\left(\frac{r_j}{N}\right)\right]\right),
\end{aligned} \tag{3.9}$$

where H_π, K_π are defined by (3.4).

From Lemma 5.2 in [13], the differentiability assumption on K_π, (3.3) and the fact that all convergences are in the uniform metric, we conclude that

$$(3.10)\qquad \frac{K_\pi\left(H_\pi\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right]-K_\pi\left(H_\pi\left[H^{-1}\left(\frac{r_j}{N}\right)\right]\right)\right.}{H_\pi\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right]-H_\pi\left[H^{-1}\left(\frac{r_j}{N}\right)\right]}\rightarrow_{\text{a.s.}}\alpha_\pi(\beta_j),$$

and

$$(3.11)\qquad \frac{H_\pi\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right]-H_\pi\left[H^{-1}\left(\frac{r_j}{N}\right)\right]}{H\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right]-H\left[H^{-1}\left(\frac{r_j}{N}\right)\right]}\rightarrow_{\text{a.s.}} 1.$$

Further, the process

$$W_N(t) = N^{\frac{1}{2}}(H_N[H^{-1}(t)] - t)$$

converges in law to the Wiener process $W_0(t)$, and, as in (3.8),

$$(3.12)\qquad W_N(H[H_N^{-1}(N^{-1}r_j)]) \rightarrow_L W_0(\beta_j).$$

Moreover, since

$$(3.13)\qquad 0 \leqq H_N[H_N^{-1}(N^{-1}r_j)] - H[H^{-1}(N^{-1}r_j)] \leqq N^{-1}$$

it follows from (3.9)–(3.13) that

$$(3.14)\qquad n_i^{\frac{1}{2}}\left(F_{n_i}^{(i)}\left[H_N^{-1}\left(\frac{r_j}{N}\right)\right]-F^{(i)}\left[H^{-1}\left(\frac{r_j}{N}\right)\right]\right)$$
$$\rightarrow_L W_0^{(i)}(F^{(i)}[H_\pi^{-1}(\beta_j)]) - \pi_i^{\frac{1}{2}}\frac{f^{(i)}[H_\pi^{-1}(\beta_j)]}{h[H_\pi^{-1}(\beta_j)]}W_0(\beta_j),$$

where $f^{(i)}/h$ is defined to be zero whenever $h = f^{(i)} = 0$.

Since

$$N^{-\frac{1}{2}}M_{ij} = N^{-\frac{1}{2}}(n_iF_{n_i}^{(i)}[H_N^{-1}(N^{-1}r_j)] - n_iF_{n_i}^{(i)}[H_N^{-1}(N^{-1}r_{j-1})]),$$

it follows from (3.14) that $N^{-\frac{1}{2}}\mathbf{M}$ is asymptotically normally distributed. This completes the proof.

Theorem 3.1 implies, in particular, the asymptotic normality of the vector $N^{-\frac{1}{2}}\mathbf{M}$ under the null-hypothesis. The limiting distribution is singular, due to restrictions (2.1). However, for the subvector

$$(3.15)\qquad \mathbf{M}_1 = (M_{11}, M_{12}, \cdots, M_{1p-1}, \cdots, M_{k-11}, M_{k-12}, \cdots, M_{k-1p-1})$$

we have

COROLLARY 3.1. *Under H_0 and assumptions* (3.1), $N^{-\frac{1}{2}}(\mathbf{M}_1 - \boldsymbol{\mu}_1)$ *has a limiting* $N(0, \Lambda)$ *distribution, where*

$$(3.16)\qquad \boldsymbol{\mu}_1 = N(\pi_1\alpha_1, \cdots, \pi_1\alpha_{p-1}, \cdots, \pi_{k-1}\alpha_1, \cdots, \pi_{k-1}\alpha_{p-1}),$$

and Λ is the non-singular $(p - 1)(k - 1) \times (p - 1)(k - 1)$ matrix

$$\Lambda = \begin{bmatrix} \alpha_1 \pi_1(1 - \alpha_1)(1 - \pi_1), - \alpha_1 \pi_1 \alpha_2(1 - \pi_1), \cdots, - \alpha_1 \pi_1(1 - \alpha_1)\pi_2, \alpha_1 \pi_1 \alpha_2 \pi_2, \cdots \\ - \alpha_2 \pi_1 \alpha_1(1 - \pi_1), \alpha_2 \pi_1(1 - \alpha_2)(1 - \pi_1), \cdots \\ \vdots \\ \alpha_{p-1} \pi_{k-1} \alpha_1 \pi_1, \alpha_{p-1} \pi_{k-1} \alpha_2 \pi_1, \cdots, \alpha_{p-1} \pi_{k-1}(1 - \alpha_{p-1})(1 - \pi_{k-1}) \end{bmatrix}$$

This follows from Theorem 3.1, or may be proved directly, [17], by an argument similar to de Moivre-Laplace's multinomial approximation model.

4. The C-test. Define the random variable C_N by

$$C_N = \sum_{j=1}^{p} \sum_{i=1}^{k} \frac{(M_{ij} - N\pi_i \alpha_j)^2}{N\pi_i \alpha_j}. \tag{4.1}$$

Suppose that $\{K_N\}$ is a sequence of alternative hypotheses, and let the subscripts N, 0, denote integration under K_N and H_0 respectively. Assume that, for $i = 1, 2, \cdots, k, j = 1, 2, \cdots, p$, the following two conditions are fulfilled:

$$E_N(M_{ij}/N) - \lim_{N\to\infty} E_0(M_{ij}/N) = A_{ij}N^{-\frac{1}{2}} + o(N^{-\frac{1}{2}}), \tag{4.2}$$

$$\sigma_N(M_{i_1j_1}N^{-\frac{1}{2}}, M_{i_2j_2}N^{-\frac{1}{2}}) - \lim_{N\to\infty} \sigma_0(M_{i_1j_1}N^{-\frac{1}{2}}, M_{i_2j_2}N^{-\frac{1}{2}}) = o(1) \tag{4.3}$$

where the A_{ij}'s are constants, $i_1 = 1, 2, \cdots, k, i_2 = 1, 2, \cdots, k, j_1 = 1, 2, \cdots, p$, $j_2 = 1, 2, \cdots, p$. Let $\mathbf{A}_1 = (A_{11}, \cdots, A_{1p-1}, \cdots, A_{k-11}, \cdots, A_{k-1p-1})$.

Suppose $\mathbf{M}_{N,1}$ represents the $(p - 1)(k - 1)$ dimensional vector $\mathbf{M}_1$, (3.15), based on sample size N. Let $\boldsymbol{\mu}_{N,1}$ be the mean vector for $\mathbf{M}_{N,1}$ and Λ_N the dispersion matrix of $N^{-\frac{1}{2}}\mathbf{M}_{N,1}$ under K_N. Finally, let $\boldsymbol{\mu}_{0,1} = \boldsymbol{\mu}_1$, where $\boldsymbol{\mu}_1$ is defined in (3.16). Then we have,

THEOREM 4.1. *Under conditions* (4.2) *and* (4.3), *assuming K_N to be true for each N, the limiting distribution of C_N as $N \to \infty$ is non-central χ^2, with $(p - 1)(k - 1)$ degrees of freedom, and non-centrality parameter*

$$\lambda^C = \mathbf{A}_1\Lambda^{-1}\mathbf{A}_1'.$$

The proof follows at once from Theorem 3.1 and conditions (4.2) and (4.3).

Since Theorem 4.1 implies in particular that under H_0, C_N has a limiting central χ^2 distribution with $(p - 1)(k - 1)$ degrees of freedom, we can define, for large N, a size α "C-test" of the null-hypothesis (1.1), by

$$\text{"Reject } H_0 \text{ if } C_N > c_\alpha\text{"} \tag{4.4}$$

where c_α is the $100(1 - \alpha)$th percentile point of the χ^2 distribution with $(p - 1)(k - 1)$ degrees of freedom. This test may be seen to be consistent against any alternative satisfying the condition that, for some $i = 1, 2, \cdots, k$, and some $j = 1, 2, \cdots, p$, one has $\beta_j \neq F^{(i)}H_\pi^{-1}(\beta_j)$.

COROLLARY 4.1. *Suppose that, for each N,*

$$K_N: F^{(i)}(x) = F(x + \theta_i N^{-\frac{1}{2}}), \quad \theta_i \, real, i = 1, 2, \cdots, k, \tag{4.5}$$

where F has a finite second derivative, and F^{-1} has a finite derivative. Then C_N has a *limiting non-central chi-square distribution, with* $(p - 1)(k - 1)$ *degrees of freedom, and non-centrality parameter*

$$(4.6)\qquad \lambda^C = \sum_{i=1}^k \pi_i(\theta_i - \bar{\theta})^2 \sum_{j=1}^p \alpha_j^{-1}(f[F^{-1}(\beta_j)] - f[F^{-1}(\beta_{j-1})])^2,$$

where $\bar{\theta} = \sum_{l=1}^k \pi_l\theta_l$.

The corollary is proved by verifying that conditions (4.2) and (4.3) hold, so that Theorem 4.1 applies. The verification may be carried out through a number of Taylor series expansions [17]. A similar argument yields

COROLLARY 4.2. *Suppose that, for each* N,

$$(4.7)\qquad K_N:F^{(i)}(x) = F([1 + \theta_i N^{-\frac{1}{2}}]x), \quad \theta_i \text{ real}, i = 1, 2, \cdots, k,$$

where F *satisfies the assumptions of Corollary* 4.1. *Then, as* $N \to \infty$, C_N *has a limiting non-central chi-square distribution, with* $(p - 1)(k - 1)$ *degrees of freedom, and non-centrality parameter*

$$(4.8)\qquad \lambda^C = \sum_{i=1}^k \pi_i(\theta_i - \bar{\theta})^2 \sum_{j=1}^p \alpha_j^{-1}(F^{-1}(\beta_j)f[F^{-1}(\beta_j)] - F^{-1}(\beta_{j-1})f[F^{-1}(\beta_{j-1})])^2.$$

5. Asymptotic relative efficiency of the C-test when two random intervals are used ($p = 2$). Since most common rank tests for the k-sample problem are consistent, a useful criterion for comparison of these tests is the asymptotic relative efficiency (A.R.E.) in Pitman's sense [8].

Let M denote the Mood-Brown statistic, [7], given by

$$(5.1)\qquad M = N(N - 1)[b(N - b)]^{-1} \sum_{i=1}^k n_i^{-1}(m_i - N^{-1}bn_i)^2,$$

where $b = \frac{1}{2}(N - 1)$ for N odd, $b = \frac{1}{2}N$ for N even, and m_i denotes the number of observations in the ith sample which are less than the median of all observations.

Let H be the Kruskal-Wallis statistic [19],

$$H = 12[N(N + 1)]^{-1} \sum_{i=1}^k n_i(\bar{R}_i - \tfrac{1}{2}(N + 1))^2,$$

where $\bar{R}_i$ is the average rank of the ith sample in the pooled ordering of all N observations.

Let $\mathfrak{F}$ represent the classical analysis of variance statistic. Finally, let $\mathfrak{L}$ denote the test-statistic proposed by Puri, [11],

$$\mathfrak{L} = \sum_{i=1}^k n_i[(T_{N,i} - \mu_{N,i})/A_N]^2,$$

where $\mu_{N,i}$ and A_N are normalizing constants and

$$(5.2)\qquad T_{N,i} = n_i^{-1} \sum_{l=1}^N E_{N,l}Z_{N,l}^{(i)},$$

with $E_{N,l}$ constants, and $Z_{N,l}^{(i)} = 1$ if the lth smallest observation in the combined sample comes from the ith sample, while $Z_{N,l}^{(i)} = 0$ otherwise.

It is shown by Andrews [1] and Puri [11] that, under certain regularity conditions on the distribution F, and under the translation sequence (4.5), the above statistics have a limiting non-central χ^2 distribution, with $(k - 1)$ degrees of freedom, and non-centrality parameters which are, respectively,

$$\lambda^M = 4[f[F^{-1}(\tfrac{1}{2})]]^2 \textstyle\sum_{i=1}^k \pi_i(\theta_i - \bar{\theta})^2, \tag{5.3}$$

$$\lambda^H = 12[\textstyle\int_{-\infty}^{\infty} f(x)\, dF(x)]^2 \sum_{i=1}^k \pi_i(\theta_i - \bar{\theta})^2, \tag{5.4}$$

$$\lambda^{\mathfrak{F}} = \sigma_F^{-2} \textstyle\sum_{i=1}^k \pi_i(\theta_i - \bar{\theta})^2, \tag{5.5}$$

$$\lambda^{\mathfrak{L}} = \sum_{i=1}^k \pi_i(\theta_i - \bar{\theta})^2 \left[\int_{-\infty}^{\infty} \frac{d}{dx} J[F(x)]\, dF(x)\right]^2 \Big/ A^2; \tag{5.6}$$

where $f = F'$, $\sigma_F^2 = \int_{-\infty}^{\infty} x^2\, dF(x) - [\int_{-\infty}^{\infty} x\, dF(x)]^2$; $J(x)$ on $[0, 1]$ is defined by $J(H) = \lim_{N\to\infty} J_N(H)$, for $H = \sum_{i=1}^k N^{-1}n_iF^{(i)}$, with $J_N(0) = J_N(0+)$ and $J_N(t) = E_{N,l}$ for $l/N < t \leqq (l + 1)/N$; $E_{N,l}$ is given in (5.2); and $A^2 = \int_0^1 J^2(x)\, dx - [\int_0^1 J(x)\, dx]^2$.

Now suppose that C_N is calculated on the basis of two random intervals (i.e. $p = 2$). It is easy to verify that, in this event, the choice in (1.2) of $r_0 = 0$, $r_1 = \frac{1}{2}N$, $r_2 = N + 1$ (for N even), yields

$$M = C_N(1 - 1/N),$$

where M is given in (5.1.) (A similar argument holds for N odd.) Thus the Mood-Brown test becomes a special case of the C-test, for large N; indeed, it can be seen at once that, under the sequence of alternatives (4.5), both C_N and M have the same limiting non-central χ^2 distribution.

Remark. More generally, under the conditions of Corollary 4.1, assuming that $\int_{-\infty}^{\infty} x^2\, dF(x) - [\int_{-\infty}^{\infty} x\, dF(x)]^2 = \sigma_F^2$ exists, we have, for $p = 2$ and for $T = M$, H, $\mathfrak{F}$ or $\mathfrak{L}$, that $e_{C,T} = \lambda^C/\lambda^T$, where $e_{C,T}$ denotes the asymptotic efficiency of the C-test relative to the T-test, λ^C is given in (4.6) and λ^T in (5.3)–(5.6). This follows from Theorem 5.1 in Andrews [1].

Since λ^C, and thus $e_{C,T}$, depend on the choice of r_1 in (1.2), optimal values of the A.R.E. may be attained through appropriate selection of $\beta_1 = \lim_{N\to\infty} N^{-1}r_1$. Table I gives the optimal values of $e_{C,M}$, $e_{C,\mathfrak{F}}$, and $e_{C,H}$ for different densities, under β_1 values suggested in Table III, section 7.

To evaluate $e_{C,\mathfrak{L}}$, it is necessary to assume a specific form for $J(x)$ in (5.6).

It is shown by Puri [11] that, if $J(x) = x$ for $0 \leqq x \leqq 1$, $\mathfrak{L}$ becomes the Kruskal-

TABLE I

$(p = 2)$ (*Location*)

Density	$\sup_{0\leqq\beta_1\leqq 1} e_{C,M}$	$\sup_{0\leqq\beta_1\leqq 1} e_{C,\mathfrak{F}}$	$\sup_{0\leqq\beta_1\leqq 1} e_{C,H}$
$(2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}x^2)$, $-\infty < x < \infty$	1	$2/\pi$	2/3
$\exp(-x)\,[1 + \exp(-x)]^{-2}$, $-\infty < x < \infty$	1	$\pi^2/12$	3/4
$\frac{1}{2}\exp(-\lvert x\rvert)$, $-\infty < x < \infty$	1	2	4/3
$\frac{1}{4}x \exp(-\frac{1}{2}x)$, $x \geqq 0$, (χ^2 with 4 d.f.)	5.10	4	8/3
$\exp(-x)$, $x \geqq 0$	$+\infty$	$+\infty$	$+\infty$

Wallis H statistic, and thus

$$e_{C,\mathfrak{L}} = e_{C,H} = e_{C,M} \cdot e_{M,H} = e_{C,\mathfrak{F}} \cdot e_{\mathfrak{F},H}.$$

Again, Puri proves that, when $J = \Phi^{-1}$, where Φ is the standard normal distribution function, with density ϕ, then

$$e_{\mathfrak{L},\mathfrak{F}} = \sigma_F^2[\int_{-\infty}^{\infty} f^2(x)/\phi\{\Phi^{-1}[F(x)]\} \, dx]^2.$$

Then, since $e_{C,\mathfrak{L}} = e_{C,\mathfrak{F}} \cdot e_{\mathfrak{L},\mathfrak{F}}^{-1}$, $e_{C,\mathfrak{L}}$ can be computed. In particular, when $F = \Phi$, $e_{\mathfrak{L},\mathfrak{F}} = 1$ so that $e_{C,\mathfrak{L}} = e_{C,\mathfrak{F}} \leqq 2/\pi$, as in Table I.

The C-test may also be used for the k-sample scale problem under the alternative sequence (4.7). Puri [12] proves that, under this sequence, the k-sample generalizations of the Ansari-Bradley, Mood and normal scores statistics have limiting non-central chi-square distributions with $k - 1$ degrees of freedom. In particular, the Ansari-Bradley test is based on the statistic

$$\mathfrak{L}(B) = 48 \sum_{i=1}^{k} n_i (B_{N,i} - \mathfrak{B}_{N,i})^2,$$

where

$$m_i B_{N,i} = \sum_{l=1}^{N} [\tfrac{1}{2} + \tfrac{1}{2}N^{-1} - |\tfrac{1}{2} + \tfrac{1}{2}N^{-1} - lN^{-1}|] Z_{N,l}^{(i)},$$

$Z_{N,l}^{(i)}$ are defined as in (5.2) and $\mathfrak{B}_{N,i}$ are normalizing constants. The limiting non-centrality parameter for this statistic is then given by

$$\lambda^B = 48(\int_{-\infty}^{0} xf^2(x) \, dx - \int_0^{\infty} xf^2(x) \, dx)^2 \sum_{i=1}^{k} \pi_i(\theta_i - \bar{\theta})^2.$$

Optimal A.R.E. values of the C-test (with $p = 2$) against $\mathfrak{L}(B)$ are given in Table II for different densities and for suitably chosen β_1 values. Efficiencies relative to the other scale tests may be evaluated from the corresponding non-centrality parameters. (Puri [12].)

6. Asymptotic relative efficiency when more than two intervals are used **($p > 2$).** When the C-test is based on three or more random intervals, the problem of measuring its asymptotic efficiency under the location sequence (4.5) in relation to the tests above becomes more complicated. This is due to the discrep-

TABLE II
($p = 2$) (*Scale*)

Density	$\text{Sup}_{0 \leqq \beta_1 \leqq 1} e_{C,B}$	Optimal β_1
$(2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}x^2)$, $-\infty < x < \infty$	0.500	.9463 (or .0537)
$\exp(-x) [1 + \exp(-x)]^{-2}$, $-\infty < x < \infty$	0.419	.9168 (or .0832)
$\frac{1}{2} \exp(-\lvert x \rvert)$, $-\infty < x < \infty$	0.389	.9206 (or .0794)
$\frac{1}{4} x \exp(-\frac{1}{2}x)$, $x \geqq 0$	0.190	.7000
$\exp(-x)$, $x \geqq 0$	0.216	.7961

ancy in the degrees of freedom of the limiting distributions of the corresponding statistics, which renders the method described in Section 5 invalid. On the other hand, we shall see in Section 7 that increasing the numbers of intervals results in increasing λ^C, the non-centrality parameter, thus possibly improving the efficiency of the C-test. It is therefore of particular interest to have some means of calculating the efficiency for larger values of p.

THEOREM 6.1. *Let the assumptions of the Remark in Section 5 hold, with $p > 2$ and k large, and let $T = M, H, \mathfrak{F}$ or $\mathfrak{L}$. Then, in testing the null-hypothesis against the sequence of alternatives* (4.5), *we have,*

$$(6.1) \qquad e_{C,T} \cong (\lambda^T)^{-1}[z_\alpha(2\nu_T)^{\frac{1}{2}} + 2R_C^2 \pm 2R_C(R_C^2 + z_\alpha(2\nu_T)^{\frac{1}{2}} + \tfrac{1}{2}\nu_T)^{\frac{1}{2}}]$$

where λ^T is given by (5.3)–(5.6), *z_α is the* $100(1 - \alpha)$th *percentile point of the $N(0, 1)$ distribution, $\nu_T = k - 1$, R_C is a function of p, k and λ^C, defined by*

$$R_C = (z_\alpha(2\nu_C)^{\frac{1}{2}} - \lambda^C)(2[\nu_C + 2\lambda^C],)^{-\frac{1}{2}},$$

with $\nu_C = (p - 1)(k - 1)$, and the sign to be chosen in (6.1) *is that which satisfies*

$$R_C = (z_\alpha(2\nu_T)^{\frac{1}{2}} - \lambda^T e_{C,T})(2[\nu_T + 2\lambda^T e_{C,T}])^{-\frac{1}{2}}.$$

PROOF. Since k is large, both ν_T and ν_C are large, so that $(C_N - \nu_C)(2\nu_C)^{-\frac{1}{2}}$ and $(T - \nu_T)(2\nu_T)^{-\frac{1}{2}}$ have, approximately, a $N(0, 1)$ distribution under the null-hypothesis. We require both tests to have size α. Thus, under H_0,

$$P_0[(C_N - \nu_C)(2\nu_C)^{-\frac{1}{2}} > z_\alpha] = \alpha = P_0[(T - \nu_T)(2\nu_T)^{-\frac{1}{2}} > z_\alpha].$$

Now, suppose that when C_N is computed on the basis of N observations, the power β is achieved against the alternative K_N. Thus, under K_N,

$$(6.2) \quad P_N[(C_N - \nu_C)(2\nu_C)^{-\frac{1}{2}} > z_\alpha] = \beta = P_N[(C_N - \nu_C - \lambda^C)(2[\nu_C + 2\lambda^C])^{-\frac{1}{2}} > (z_\alpha(2\nu_C)^{\frac{1}{2}} - \lambda^C)(2[\nu_C + 2\lambda^C])^{-\frac{1}{2}}].$$

Since k is assumed to be large, we have from Corollary 4.1 and the normal approximation to the non-central χ^2 distribution that

$$(6.3) \qquad \int_{R_C}^{\infty} (2\pi)^{-\frac{1}{2}} \exp(-t^2/2)\, dt = \beta,$$

where

$$R_C = (z_\alpha(2\nu_C)^{\frac{1}{2}} - \lambda^C)(2[\nu_C + 2\lambda^C])^{-\frac{1}{2}}.$$

Suppose the test based on T requires N_T observations to attain power β (at size α) against the alternative K_{N_T}. Applying once again the normal approximation,

$$(6.4) \qquad \int_{R_T}^{\infty} (2\pi)^{-\frac{1}{2}} \exp(-t^2/2)\, dt = \beta,$$

where

$$R_T = (z_\alpha(2\nu_T)^{\frac{1}{2}} - \lambda^T)(2[\nu_C + 2\lambda^C])^{-\frac{1}{2}}.$$

From (6.3) and (6.4) it follows that

$$R_C = R_T. \tag{6.5}$$

When the alternatives K_N and K_{N_T} are identified, we have from (6.5),

$$(z_\alpha(2\nu_C)^{\frac{1}{2}} - \lambda^C)(2[\nu_C + 2\lambda^C])^{-\frac{1}{2}} = (z_\alpha(2\nu_T)^{\frac{1}{2}} - \lambda^T e_{C,T})(2[\nu_T + 2\lambda^T e_{C,T}])^{-\frac{1}{2}}, \tag{6.6}$$

and (6.1) is obtained by solving (6.6) as a quadratic in $e_{C,T}$. Thus the theorem is established.

THEOREM 6.2. *Under the assumptions of Theorem* 6.1 *but with k small, $p > 2$, one has, for large values of ν_C or λ^C, and λ^T, that*

$$e_{C,T} \cong (\lambda^T)^{-1}[c_{\alpha,T} + 2R_C^2 - \nu_T \pm 2R_C(R_C^2 - \tfrac{1}{2}\nu_T + c_{\alpha,T})^{\frac{1}{2}}], \tag{6.7}$$

where $c_{\alpha,T}$ is the $100(1 - \alpha)$th *percentile point of the central χ^2 distribution with $\nu_T = k - 1$ degrees of freedom, R_C is defined by*

$$R_C = (c_{\alpha,C} - [\nu_C + \lambda^C])(2[\nu_C + \lambda^C])^{-\frac{1}{2}},$$

with $\nu_C = (p - 1)(k - 1)$ and $c_{\alpha,C}$ the $100(1 - \alpha)$th *percentile point of the central χ^2 distribution with ν_C degrees of freedom, and the sign to be chosen in* (6.7) *is that which satisfies*

$$R_C = (c_{\alpha,T} - [\nu_T + \lambda^T e_{C,T}])(2[\nu_T + 2\lambda^T e_{C,T}])^{-\frac{1}{2}}.$$

The proof is analogous to that of Theorem 6.1, except that, due to the smaller number of degrees of freedom, the tests are made to have size α on the basis of the central χ^2 distribution, with parameters ν_C and ν_T respectively. (Recall Theorem 4.1.)

THEOREM 6.3. *Suppose the assumptions of Theorem* 6.1 *hold, with k small, $p > 2$, and ν_C or λ^C large. Then,*

$$e_{T,C} = e_{C,T}^{-1} \cong (\lambda^C)^{-1}[c_{\alpha,C} + 2R_T^2 - \nu_C \pm 2R_T(R_T^2 - \tfrac{1}{2}\nu_C + c_{\alpha,C})^{\frac{1}{2}}], \tag{6.8}$$

where R_T is defined by

$$R_T = \left(\frac{2c_{\alpha,T}(\nu_T + \lambda^T)}{\nu_T + 2\lambda^T}\right)^{\frac{1}{2}} - \left(\frac{2(\nu_T + \lambda^T)^2}{\nu_T + 2\lambda^T} - 1\right)^{\frac{1}{2}},$$

with $c_{\alpha,C}$, $c_{\alpha,T}$, ν_C, ν_T, λ^C, λ^T as in Theorem 6.2. *The sign to be picked in* (6.8) *is that which satisfies*

$$R_T = (c_{\alpha,C} - [\nu_C + \lambda^C e_{T,C}])(2[\nu_C + 2\lambda^C e_{T,C}])^{-\frac{1}{2}}.$$

PROOF. Since k is small, and λ^T is not assumed sufficiently large to warrant the normal approximation to $(T - [\nu_T + \lambda^T])(2[\nu_T + 2\lambda^T])^{-\frac{1}{2}}$, a better normal approximation under the alternative sequence is to be desired. Patnaik [9] proposes a faster normal approximation to the non-central χ^2, (denoted by χ'^2), with parameters ν and λ, by showing that the random variable $[2\chi'^2(\nu + \lambda)/(\nu + 2\lambda)]^{\frac{1}{2}}$ converges in law quite rapidly to a $N([2(\nu + \lambda)^2/(\nu + 2\lambda) - 1]^{\frac{1}{2}}, 1)$ variable. Our proof incorporates this approximation to T under the alternative sequence, with the standard approximation to C_N. In other respects, the proof is analogous to that of Theorem 6.2.

REMARK. The efficiency can be obtained from Theorem 6.1 when p and k are both large, or when at least k is large, since this guarantees that $\nu_T = k - 1$ and $\nu_C = (p-1)(k-1)$ are sufficiently large for the approximations introduced under H_0 and under K_N. When k is smaller, but p is large, one may choose to apply Theorem 6.3, unless the non-centrality parameter λ^T is large, in which case one may compute instead the simpler form of $e_{C,T}$ given by Theorem 6.2. Again, if p is smaller, with k small, one would apply Theorem 6.2 when both non-centrality parameters are large, or Theorem 6.3 when λ^T is smaller. Finally, if p and k are fairly small, and λ^T is large while λ^C is smaller, Theorem 6.3 can be used with T and C interchanged.

EXAMPLE. Suppose $k = 4, p = 4$, and let $F = \Phi$, the standard normal distribution. Further, let

$$\theta_1 = -\theta_2 = 5, \quad \theta_3 = -\theta_4 = 10, \quad \pi_1 = \pi_2 = \tfrac{1}{8}, \quad \pi_3 = \pi_4 = \tfrac{3}{8},$$

and pick

$$\alpha_1 = .16, \quad \alpha_2 = .34, \quad \alpha_3 = .34, \quad \alpha_4 = .16,$$

in accordance with the optimal choice, to be specified later (Table III). Thus

$$F^{-1}(\beta_1) = -.994, \quad F^{-1}(\beta_2) = 0, \quad F^{-1}(\beta_3) = .994,$$

so that

$$r_0 = 0, \quad r_1 = .16N, \quad r_2 = .5N, \quad r_3 = .84N, \quad r_4 = N + 1.$$

To compute $e_{C,M}$ we use Theorem 6.2, since k is small while λ^C and λ^M are both large ($\lambda^C = 71.687, \lambda^M = 51.713$). Hence we have, $e_{C,M} = 1.445 > 1 = \sup_{\beta_1} e_{C,M}$, where the right side of the inequality was obtained for $p = 2$ (Table I).

Similarly, $e_{C,\mathfrak{F}} = e_{C,M} e_{M,\mathfrak{F}} = 0.920$.

7. Design for high relative efficiency. The asymptotic relative efficiency $e_{C,T}$ increases with λ^C. This follows at once from Remark (Section 5) when $p = 2$, and from Theorems 6.1–6.3 and the identity $e_{C,T} = e_{T,C}^{-1}$, when $p > 2$.

We recall (4.6) that, under the location sequence (4.5),

$$\lambda^C = K(\boldsymbol{\theta}) \cdot I_p,$$

where

$$K(\boldsymbol{\theta}) = \textstyle\sum_{i=1}^k \pi_i(\theta_i - \bar{\theta})^2, \qquad \bar{\theta} = \textstyle\sum_{l=1}^k \pi_l \theta_l,$$

and

$$I_p = \sum_{j=1}^{p} \frac{(f[F^{-1}(\beta_j)] - f[F^{-1}(\beta_{j-1})])^2}{\beta_j - \beta_{j-1}}.$$

To maximize λ^C for a fixed λ^T ($T = M, H, \mathfrak{F}, \mathfrak{L}$), we hold $K(\boldsymbol{\theta})$ fixed and attempt to maximize I_p in the design. To this end, we show first that I_p is an increasing function of p.

THEOREM 7.1. *For* $a < x < b$, $F(a) < F(x) < F(b)$, *one has*

$$\frac{[f(b) - f(a)]^2}{F(b) - F(a)} \leqq \frac{[f(b) - f(x)]^2}{F(b) - F(x)} + \frac{[f(x) - f(a)]^2}{F(x) - F(a)},$$

where $f = F'$.

PROOF. Let

$$f(b) - f(x) = u, \quad f(x) - f(a) = v, \quad \frac{F(b) - F(x)}{F(b) - F(a)} = \beta,$$

and note that

$$u^2(1 - \beta) + v^2\beta - (u + v)^2\beta(1 - \beta) = [u(1 - \beta) - v\beta]^2 \geqq 0.$$

This completes the proof.

Consider next the problem of finding an optimal set $\beta_1, \beta_2, \cdots, \beta_p$ which maximizes I_p for a fixed p and for a given distribution F. Table III lists a few such optimal designs for different values of p.

Conditions for the existence of optimal designs for each p, as well as other examples, are given in Sacks and Ylvisaker [15], Särndal [16], and others, in the unrelated context of regression problems. In more difficult cases, an "approximately-optimum" design method is put forward in [15], whereby, under suitable regularity conditions on the function $g(s) = f[F^{-1}(s)]$, $0 \leqq s \leqq 1$ the β_j are chosen

TABLE III
(*Location*)

Density		Optimal design	I_p
$(2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}x^2)$,	$p = 2$	$\beta_1 = \frac{1}{2}$	$2/\pi$
$-\infty < x < \infty$	$p = 3$	$\beta_1 = .2709, \quad \beta_2 = .7291$	.8098
	$p = 4$	$\beta_1 = .16, \quad \beta_2 = .50, \quad \beta_3 = .84$	.8823
$\exp(-x)[1 + \exp(-x)]^{-2}$,	$p = 2$	$\beta_1 = \frac{1}{2}$	1/4
$-\infty < x < \infty$	$p = 3$	$\beta_1 = \frac{1}{3}, \quad \beta_2 = \frac{2}{3}$	8/27
	$p = 4$	$\beta_1 = \frac{1}{4}, \quad \beta_2 = \frac{1}{2}, \quad \beta_3 = \frac{3}{4}$	5/16
$\frac{1}{2}\exp(-\lvert x\rvert), \quad -\infty < x < \infty$	$p = 2$	$\beta_1 = \frac{1}{2}$	1
	$p = 3$	no improvement	
	$p = 4$		
$\frac{1}{4} x \exp(-\frac{1}{2}x), \quad x \geqq 0$	$p = 2$	$\beta_1 = 0+$	1/2
	$p = 3$	$F^{-1}(\beta_1) = 6.1 \times 10^{-5}$	.9900
		$F^{-1}(\beta_2) = 1.36 \times 10^{-2}$	
	$p = 4$	$F^{-1}(\beta_1) = 1.5 \times 10^{-5}$	1.3919
		$F^{-1}(\beta_2) = 6.8 \times 10^{-3}$	
		$F^{-1}(\beta_3) = 1.46 \times 10^{-1}$	
$\exp(-x), \quad x \geqq 0$	$p = 2$	$\beta_1 = 0+$	∞
	$p = 3$	no improvement	
	$p = 4$		

to satisfy

$$\int_0^{\beta_j} [g''(s)]^{\frac{2}{3}}\, ds = p^{-1} j \int_0^1 [g''(s)]^{\frac{2}{3}}\, ds,$$

i.e., the β's determine equal areas under the curve $[g''(s)]^{\frac{2}{3}}$. This method is asymptotically optimal in the following sense: if the optimal design for size p gives the value I_p^*, and if $\lim_{p\to\infty} I_p^* = I^*$ exists, and I_p is the value attained by the "approximate" design for size p, then

$$\lim_{p\to\infty} \frac{I_p - I^*}{I_p^* - I^*} = 1.$$

REMARK. We can write

$$I_p = \sum_{j=1}^{p} \left[\left(\frac{F'(x_j) - F'(x_{j-1})}{x_j - x_{j-1}} \right)^2 \div \frac{F(x_j) - F(x_{j-1})}{x_j - x_{j-1}} \right] (x_j - x_{j-1}).$$

Thus, if F has two continuous derivatives, it follows that, as $p \to \infty$ with $\max_{1 \leq j \leq p} (x_j - x_{j-1}) \to 0$, then

$$I_p \to E_F \left[\frac{d \ln F'(X)}{dX} \right]^2 = I,$$

the well-known "information integral". It is of interest to note that in the regression context of maximizing I_p the quantity I_p^{-1} is a measure of the variance of the regression coefficient estimate. Thus, the best estimate will have variance at least as large as I^{-1}, the Cramér-Rao lower bound.

EXAMPLES. (i) For the standard normal distribution, $\lim_{p\to\infty} I_p = 1$.

We note that the optimal designs picked in Table III for $p = 2, 3, 4$, gave $I_2 = 2/\pi = .6365$, $I_3 = .8098$, $I_4 = .8823$, approaching this limit.

(ii) For the logistic distribution, $\lim_{p\to\infty} I_p = 1/3$, and the optimal values were $I_2 = 1/4$, $I_3 = 8/27$, $I_4 = 5/16$, approaching 1/3.

(iii) For the bilateral exponential, $\lim_{p\to\infty} I_p = 1$, and this limit can be attained with $p = 2$.

(iv) For χ^2 with 4 d.f., $\lim_{p\to\infty} I_p = \infty$. Here the pth interval may contribute at most 1/2 to I_{p-1}, as $\beta_p \to 0+$. Table III shows that, for the $F^{-1}(\beta)$ values given, $I_3 = .9900$, $I_4 = 1.3919$. Computations suggest that larger $F^{-1}(\beta)$ values still give favorable results. Thus, with $F^{-1}(\beta_1) = 1.56 \times 10^{-2}$, $F^{-1}(\beta_2) = 2.25 \times 10^{-1}$, one has $I_3 = .8657$. Adding the point $F^{-1}(\beta_3) = 9.48 \times 10^{-1}$ gives $I_4 = 1.0105$.

(v) For the exponential distribution, I_2 may be made arbitrarily large by letting $\beta_1 \to 0+$, thus leaving the choice of other β's arbitrary.

8. Acknowledgments. The author wishes to express his deep appreciation to Professor Douglas G. Chapman for providing continual assistance and advice during the present work. A debt of gratitude is due to Professor Ronald Pyke for the valuable discussions and suggestions which he offered. The author would also like to thank Professor Donald N. Ylvisaker and Professor Galen R. Shorack for their useful comments and criticisms. Finally, sincere appreciation is due to Prof. Dr. G. J. Leppink and his Mathematical Statistics staff at the University of Utrecht, Netherlands, for their ready assistance with many of the computations.

REFERENCES

[1] ANDREWS, F. C. (1954). Asymptotic behavior of some rank tests for analysis of variance. *Ann. Math. Statist.* **25** 724–736.

[2] BLUMENTHAL, S. (1967). Limit theorems for functions of shortest 2-sample spacings, and a related test. *Ann. Math. Statist.* **38** 108–116.

[3] DIXON, W. J. (1940). A criterion for testing the hypothesis that two samples are from the same population. *Ann. Math. Statist.* **11** 199–204.

[4] DONSKER, M. L. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **23** 277–281.

[5] HOEFFDING, W. (1951). 'Optimum' nonparametric tests. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 83–92. Univ. of California Press.

[6] MASSEY, F. J., JR. (1951). A note on a two sample test. *Ann. Math. Statist.* **22** 304–306.

[7] MOOD, A. M. (1950). *Introduction to the Theory of Statistics.* McGraw-Hill, New York.

[8] NOETHER, G. E. (1955). On a theorem of Pitman. *Ann. Math. Statist.* **26** 64–68.

[9] PATNAIK, P. B. (1949). The non-central χ^2 and F distributions and their applications. *Biometrika* **36** 202–232.

[10] PROKHOROV, YU., V. (1956). Convergence of random processes and limit theorems in probability theory. *Theor. Probability Appl.* **1** 157–214.

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

[12] PURI, M. L. (1965). On some tests of homogeneity of variances. *Ann. Inst. Statist. Math.* **17** 323–330.

[13] PYKE, R., and SHORACK, G. R. (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage Theorems. *Ann. Math. Statist.* **39** 755–771.

[14] ROOT, D. (1967). Construction of almost surely convergent random processes. Ph.D. thesis, Univ. of Washington.

[15] SACKS, J., and YLVISAKER, D. (1966). Designs for regression problems with correlated errors. *Ann. Math. Statist.* **37** 66–89.

[16] SÄRNDAL, CARL-ERIK. (1962). On a maximizing problem with several applications in statistical theory. *Ark. Mat.* **4** 385–392.

[17] SCHULZER, M. (1967). Contributions to the k-sample problem: a symmetric spacings statistic. Ph.D. thesis, Mathematics Department, Univ. of Washington.

[18] SKOROKHOD, A. V. (1956). Limit theorems for stochastic processes. *Theor. Probability Appl.* **1** 261–290.

[19] WALLIS, W. A., and KRUSKAL, W. H. (1952). Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Assoc.* **47** 583–621.

[20] WEISS, L. (1958). Spacings generated by mixed samples. *Ann. Math. Statist.* **29** 316–318.

[21] WEISS, L. (1961). Tests of fit based on the number of observations falling in the shortest sample spacings determined by earlier observations. *Ann. Math. Statist.* **32** 838–845.