TESTING HOMOGENEITY AGAINST ORDERED ALTERNATIVES

By V. J. CHACKO[1]

Forest Research Institute, Dehra Dun, India

0. Introduction and summary. In a one-way analysis of variance situation in which the populations are ordered under the alternative hypothesis, one desires a test that, unlike the usual normal theory F test, concentrates its power on the ordered alternatives, not on any alternatives. In this paper, two contributions are made. First, under the usual normal assumptions, work by Bartholomew [4], [7] on the likelihood ratio test, when the ordering is complete under the alternative hypothesis, is extended. By suitable characterization of the partition which the likelihood ratio induces on the sample space, the likelihood ratio test is shown to depend on incomplete Beta functions and certain probabilities of the above partitions of the sample space. The major contribution in this paper is for the case of equal sample sizes, where explicit expressions for these probabilities are obtained by indicating their relationship to Sparre Andersen's [1], [2] results. Second, under the analogous nonparametric assumptions and for equal sample sizes, a parallel test based on ranks is proposed and discussed for stochastic ordering of the populations. The asymptotic Pitman efficiency of the nonparametric test relative to the test in the normal case is derived.

1. Statement of the first problem. Consider k independent normal variates $x_1, x_2, \cdots, x_k$ with unknown means $\theta_1, \theta_2, \cdots, \theta_k$ respectively and a common but unknown variance σ^2. Let x_{ij} ($i = 1, 2, \cdots, k$; $j = 1, 2, \cdots, n_i$) be independent observations on the k variables, where x_{ij} is the jth observation from the ith variable. Let $\bar{x}_i = \sum_{j=1}^{n_i} x_{ij}/n_i$ and $s_i^2 = \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2/n_i$ denote the sample mean and standard deviation for the ith variable. We are interested in testing the hypothesis $H_0: \theta_1 = \theta_2 = \cdots = \theta_k$ against $H_1: \theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_k$ (with at least one inequality strong), where σ^2 is unspecified for both hypotheses. Denote by

$$(1) \quad \bar{x}_{[t,s]} = (n_t\bar{x}_t + n_{t+1}\bar{x}_{t+1} + \cdots + n_s\bar{x}_s)/(n_t + n_{t+1} + \cdots + n_s)$$

the pooled sample mean of $\bar{x}_t, \bar{x}_{t+1}, \cdots, \bar{x}_s$ where s and t are positive integers with $1 \leqq t < s \leqq k$.

2. Estimation of parameters. The MLE's (maximum likelihood estimates) for the θ's under H_0 are well known and are $\hat{\theta}_1^0 = \hat{\theta}_2^0 = \cdots = \hat{\theta}_k^0 = \bar{x}_{[1,k]}$. Brunk

Received July 6, 1959; revised March 20, 1963.

[1] This work was carried out while the author was at the University of California, Berkeley with the partial support of the Office of Naval Research (Nonr-222-43). This paper in whole or in part may be reproduced for any purpose of the United States Government.

[8], [9] and Constance van Eeden [14] proved that the MLE's under the alternative H_1 are unique and can be formally represented as

$$\hat{\theta}_i = \max_{1 \leqq r \leqq i} \min_{i \leqq s \leqq k} \bar{x}_{[r,s]} . \tag{2}$$

A convenient procedure indicated by Brunk [9] for obtaining the MLE's under H_1 may be described in the following way. If $\bar{x}_1 \leqq \bar{x}_2 \leqq \cdots \leqq \bar{x}_k$, then $\hat{\theta}_i = \bar{x}_i$, $i = 1, 2, \cdots, k$. If $\bar{x}_i > \bar{x}_{i+1}$ for some i ($i = 1, 2, \cdots, k - 1$) then $\hat{\theta}_i = \hat{\theta}_{i+1}$ and the means $\bar{x}_i$ and $\bar{x}_{i+1}$ are replaced in the sequence $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_k$ by the pooled weighted mean $\bar{x}_{[i,i+1]}$, obtaining an ordered set of only $k - 1$ quantities ($k - 2$ of which are sample means and one is a weighted mean of two sample means). The value of the weighted mean is then compared with $\bar{x}_{i-1}$ and $\bar{x}_{i+2}$. Then

(i) if $\bar{x}_{[i,i+1]} > \bar{x}_{i+2}$, the three values $\bar{x}_i$, $\bar{x}_{i+1}$ and $\bar{x}_{i+2}$ are replaced by $\bar{x}_{[i,i+2]}$ the pooled weighted mean of the three.

(ii) if $\bar{x}_{[i,i+1]} < \bar{x}_{i-1}$, the three values $\bar{x}_{i-1}$, $\bar{x}_i$ and $\bar{x}_{i+1}$ are replaced by the pooled mean $\bar{x}_{[i-1,i+1]}$.

(iii) if $\bar{x}_{i+2} < \bar{x}_{[i,i+1]} < \bar{x}_{i-1}$ then the procedure in either (i) or (ii) above is carried out.

(iv) otherwise the values are left unaltered.

The new quantities are then compared with the adjacent ones and so on. The above procedure is continued until an ordered set of monotone nondecreasing quantities (sample means or pooled weighted means) are obtained. Thus for each i the MLE $\hat{\theta}_i$ of θ_i is equal to that one of the final quantities to which the original mean $\bar{x}_i$ contributed. Though the order of combining the means is not uniquely determined, the above procedure for obtaining the MLE's gives a final set of estimates which is unique.

If there are m distinct estimates obtained by pooling, respectively, the first t_1 means, the next t_2 means, $\cdots$, and the last t_m means, $t_j > 0$, $t_1 + t_2 + \cdots + t_m = k$, and if we set $\tau_0 = 0$

$$\begin{aligned} \tau_i &= t_1 + t_2 + \cdots t_i \; (i = 1, 2, \cdots, m) \\ \tau_m &= k, \end{aligned} \tag{3}$$

then

$$\hat{\theta}_{\tau_i+1} = \hat{\theta}_{\tau_i+2} = \cdots = \hat{\theta}_{\tau_{i+1}} = \bar{x}_{[\tau_i+1,\tau_{i+1}]} \quad (i = 0, 1, 2, \cdots, m - 1). \tag{4}$$

For convenience in notation, denote these m distinct estimates by $\bar{x}_{[t_j]}$, ($j = 1, 2, \cdots, m$), where $\bar{x}_{[t_j]} = \bar{x}_{[\tau_{j-1}+1,\tau_j]}$. Let the sum of the sample sizes for the t_j means in $\bar{x}_{[t_j]}$ be denoted by $N_{[t_j]}$.

3. Criterion for equality of all MLE. To characterize the region in the sample space where all the k sample means are pooled (i.e., $\hat{\theta}_i = \bar{x}_{[1,k]}$ for $i = 1, 2, \cdots, k$) we shall prove the following theorem.

THEOREM 1. *A necessary and sufficient condition that all the k consecutive sample means $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_k$ are pooled in forming the* MLE *is*

$$\bar{x}_{[1,j]} > \bar{x}_{[1,k]} \qquad \textit{for} \quad j = 1, 2, \cdots, k - 1. \tag{5}$$

PROOF. Since the normal distribution is continuous we shall exclude from the following argument the possibility that $\bar{x}_{[1,j]} = \bar{x}_{[1,k]}$ for any $j = 1, 2, \cdots, k-1$.

NECESSITY. If all the k means are pooled, then for every i, $\max_{1 \leqq r \leqq i} \min_{i \leqq s \leqq k} \bar{x}_{[r,s]} = \bar{x}_{[1,k]}$. In particular, for $i = 1$, $\min_{1 \leqq s \leqq k} \bar{x}_{[1,s]} = \bar{x}_{[1,k]}$. This shows that $\bar{x}_{[1,j]} > \bar{x}_{[1,k]}$ for $j = 1, 2, \cdots, k-1$.

SUFFICIENCY. Assume that Condition (5) holds. We have to prove that $\hat{\theta}_i = \bar{x}_{[1,k]}$ for $i = 1, 2, \cdots, k$. Notice that by Assumption (5), $\min_{i \leqq s \leqq k} \bar{x}_{[1,s]} = \bar{x}_{[1,k]}$. We shall show that for $r = 2, 3, \cdots, i$, $\min_{i \leqq s \leqq k} \bar{x}_{[r,s]} < \bar{x}_{[1,k]}$. The idea of the proof is to show that at least one of the terms on the left hand side is less than $\bar{x}_{[1,k]}$. Consider $\bar{x}_{[r,k]}$, where $2 \leqq r \leqq i$. For $j = r-1$ in (5) we have $\bar{x}_{[1,r-1]} > \bar{x}_{[1,k]}$, i.e., $\sum_{i=1}^{r-1} n_i\bar{x}_i > \bar{x}_{[1,k]}\sum_{i=1}^{r-1} n_i$. Now

$$\sum_{i=r}^{k} n_i\bar{x}_i = \bar{x}_{[1,k]}\sum_{i=1}^{k} n_i - \sum_{i=1}^{r-1} n_i\bar{x}_i < \bar{x}_{[1,k]}\sum_{i=1}^{k} n_i - \bar{x}_{[1,k]}\sum_{i=1}^{r-1} n_i = \bar{x}_{[1,k]}\sum_{i=r}^{k} n_i .$$

Thus $\bar{x}_{[r,k]} < \bar{x}_{[1,k]}$ and $\min_{i \leqq s \leqq k} \bar{x}_{[r,s]} < \bar{x}_{[1,k]}$ for every $r = 2, 3, \cdots, i$. But for $r = 1$ and $s = k$ the value is $\bar{x}_{[1,k]}$, which proves the theorem.

COROLLARY 1. *A necessary and sufficient condition for pooling q consecutive means $\bar{x}_{p+1}, \bar{x}_{p+2}, \cdots, \bar{x}_{p+q}$ is $\bar{x}_{[p+1,p+i]} > \bar{x}_{[p+1,p+q]}$ for $i = 1, 2, \cdots, q-1$.* The proof is similar to that of Theorem 1.

Let us denote the region in the sample space leading to m distinct estimates $\bar{x}_{[t_1]}, \bar{x}_{[t_2]}, \cdots, \bar{x}_{[t_m]}$ by

$$\mathfrak{X}_{(t_1,t_2,\cdots,t_m)} . \tag{6}$$

Let $\mathfrak{X}^*_{(m)} = \bigcup \mathfrak{X}_{(t_1,t_2,\cdots,t_m)}$ where the union is taken over all regions in the sample space leading to m distinct estimates. Obviously it is a union of $\binom{k-1}{m-1}$ disjoint sets. Let

$$p_{m,k} = P[\mathfrak{X}^*_{(m)}], \tag{7}$$

the probability of the set $\mathfrak{X}^*_{(m)}$ under the hypothesis H_0.

THEOREM 2. *Using the above nomenclature,*

$$p_{m,k} = \sum P(E_m)P(B_{t_1})P(B_{t_2}) \cdots P(B_{t_m}), \tag{8}$$

where

$$E_m = [\bar{x}_{[t_1]} < \bar{x}_{[t_2]} < \cdots < \bar{x}_{[t_m]}]$$

$$B_{t_i} = \bigcap_{j=\tau_{i-1}+1}^{\tau_i} [\bar{x}_{[\tau_{i-1}+1,j]} > \bar{x}_{[\tau_{i-1}+1,\tau_j]}] \quad (i = 1, 2, \cdots, m)$$

and the summation is over the $\binom{k-1}{m-1}$ values corresponding to the disjoint sets in $\mathfrak{X}^*_{(m)}$.

PROOF. We notice that the region $\mathfrak{X}_{(t_1,t_2,\cdots,t_m)}$ is the intersection of the sets $E_m, B_{t_1}, B_{t_2}, \cdots, B_{t_m}$, and that B_{t_i} is the region in the sample space where t_i sample means are pooled. By definition B_{t_i} depends on differences between the means $\bar{x}_{\tau_{i-1}+1}, \bar{x}_{\tau_{i-1}+2}, \cdots, \bar{x}_{\tau_i}$. The event E_m depends only on the differences between the sums of these means for $i = 1, 2, \cdots, m$. Since the variables are assumed to be normally distributed, the $m + 1$ events $E_m, B_{t_1}, B_{t_2}, \cdots, B_{t_m}$ are mutually independent. Thus $P[\mathfrak{X}_{(t_1, t_2, \cdots t_m)}] = P(E_m)P(B_{t_1})P(B_{t_2}) \cdots P(B_{t_m})$, and, since $\mathfrak{X}^*_{(m)}$ is the union of disjoint sets, $p_{m,k} = \sum P[\mathfrak{X}_{(t_1, t_2, \cdots, t_m)}]$, where the summation is over the $\binom{k-1}{m-1}$ values corresponding to the disjoint sets forming $\mathfrak{X}^*_{(m)}$.

4. Likelihood ratio test. It has been shown by Bartholomew [4], [7] and Chacko [10] that the likelihood ratio test at level α consists in observing the $\bar{x}_i$'s, computing the MLE's $\hat{\theta}_i$ $(i = 1, 2, \cdots, k)$ as in Section 2, T_k as in (10) and rejecting the hypothesis H_0 when $T_k \geqq C$, where

$$s_0^2 = \sum_{i=1}^{k} n_i[\bar{x}_i - \bar{x}_{[1,k]}]^2 + \sum_{i=1}^{k} n_i s_i^2, \tag{9}$$

$$T_k = \sum_{i=1}^{k} n_i[\hat{\theta}_i - \bar{x}_{[1,k]}]^2/s_0^2 = \sum_{j=1}^{m} N_{[t_j]}[\bar{x}_{[t_j]} - \bar{x}_{[1,k]}]^2/s_0^2 \tag{10}$$

and C is determined by

$$\alpha = \sum_{m=2}^{k} p_{m,k} P[\beta_{[(m-1)/2,(N-m)/2]} \geqq C]. \tag{11}$$

Here $\beta_{[(m-1)/2,(N-m)/2]}$ is a random variable having the Beta distribution with parameters $(m - 1)/2$ and $(N - m)/2$.

When σ^2 is known, the likelihood ratio test at level α is obtained by replacing s_0^2 by σ^2 in T_k and the Beta distribution by a χ^2 distribution with $m - 1$ degrees of freedom.

REMARKS. Bartholomew [6], [7] obtained some results on the power of the test based on T_k, when σ is known. The exact power function was obtained when $k = 3$ and $k = 4$. Two extreme cases were considered. First, for equal spacing of θ's and second, for the case when all but one of the θ's are equal. It was shown that the power in the latter case is lower than the power in the former case. Approximate results were also obtained in special cases when $k > 4$. The values of the power functions were compared with the ordinary χ^2 test, which assumes no prior information regarding the θ's, and the one-tail regression test, which assumes that the θ's are equally spaced. It was shown that the gain in power compared to the χ^2 test is substantial and that the relative gain increases with k. The power of the regression test is higher when the θ's are equally spaced. But this advantage is counterbalanced by lower power at the other extreme when all but one of the θ's are equal and the gap widens with larger values of k. The results suggested that the likelihood ratio test, for which the rejection level is constant, could not be substantially improved upon.

In the limiting case, as $N \to \infty$, k fixed, the power of the test based on T_k is equivalent to the power of the test for known σ because σ^2 is then estimated with infinitely many degrees of freedom. Thus the results mentioned above for the case when σ was known may be considered as large sample results for the test based on T_k.

5. Evaluation of $p_{m,k}$. Attempts to obtain explicit expressions for $p_{m,k}$ in the general case that are valid for unequal sample sizes have so far been unsuccessful. For $k \leqq 5$, however, explicit expressions can be derived. See Bartholomew [4] and Chacko [10].

Evaluation of $p_{m,k}$ when the sample sizes are equal to n. Denote the sample means by $\bar{x}_i$ $(i = 1, 2, \cdots, k)$. Consider consecutive sums of $\bar{x}_i$ and let $S_0 = 0$, $S_i = \bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_i$.

LEMMA 1. *If H_0 is true, the probability of pooling any specific q consecutive means is $1/q$.*

PROOF. There is no loss of generality in assuming that the q means pooled are the first q means $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_q$. By Theorem 1, we find that the necessary and sufficient condition for pooling these q consecutive means is $\bar{x}_{[1,i]} > \bar{x}_{1,q}$ for $i = 1, 2, \cdots, q - 1$. Equivalently, all the points (j, S_j), $(j = 1, 2, \cdots, q - 1)$ lie above the straight line $(0, 0)$ to (q, S_q).

The required probability can thus be obtained from a theorem of Sparre Andersen [1]. The statement of the Theorem is as follows. "If $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_q$ be symmetrically dependent random variables then the number t of points (j, S_j) above the line $(0, 0)$ to (q, S_q) has a uniform distribution for $t = 0, 1, \cdots, q - 1$ if and only if $P[S_i/i = S_q/q] = 0$, $i = 1, 2, \cdots, q - 1$". The conditions of the theorem are valid in our case and taking $t = q - 1$ we obtain the result $p_{1,q} = 1/q$.

To evaluate $p_{m,k}$ in general for $m > 1$ we shall introduce the following definitions.

(i) T_i $(i = 0, 1, \cdots, k)$ form a convex sequence if and only if the sequence of differences $T_{i+1} - T_i$ $(i = 0, 1, \cdots, k - 1)$ is nondecreasing.

(ii) The largest convex minorant path of the set of points (j, S_j) $(j = 0, 1, \cdots, k)$ is the convex polygonal path from $(0, 0)$ to (k, S_k) with only points (j, S_j) as vertices and such that none of the points (j, S_j) lie below the polygonal path.

(iii) The largest convex minorant sequence of the sequence S_j $(j = 0, 1, \cdots, k)$ is that of the values T_j of the polygonal path in definition (ii) at the points $j = 0, 1, \cdots, k$.

THEOREM 3. *For given sample means $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_k$,*

(a) *The largest convex minorant sequence of the sequence S_i $(i = 0, 1, \cdots, k)$ is $T_0 = 0$ and $T_i = \hat{\theta}_1 + \hat{\theta}_2 + \cdots + \hat{\theta}_i$ $(i = 1, 2, \cdots, k)$ where $\hat{\theta}_i$ $(i = 1, 2, \cdots, k)$ are the* MLE *in the domain $[\theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_k]$,*

(b) *The number of distinct estimates m is the number of equalities $S_i = T_i$ $(i = 1, 2, \cdots, k)$.*

PROOF. From Theorem 1, it is clear that a necessary and sufficient condition for all the estimates to be identical is that all the points (j, S_j) $(j = 1, 2, \cdots,$

$k - 1$) lie above the line joining $(0, 0)$ to (k, S_k), i.e., $T_j = j\bar{x}_{[1,k]}$. Taking successive differences of T_j, we obtain $\hat{\theta}_i = \bar{x}_{[1,k]} (i = 1, 2, \cdots, k)$. The number of equalities $S_i = T_i$ $(i = 1, 2, \cdots, k)$ is only one, namely, for the suffix k.

Consider the case with m distinct estimates $\bar{x}_{[t_1]}, \bar{x}_{[t_2]}, \cdots, \bar{x}_{[t_m]}$. The largest convex minorant sequence is obviously the polygonal path joining the $m + 1$ points $(0, 0)$, (τ_1, S_{τ_1}), $\cdots$, (τ_m, S_{τ_m}) with the remaining points lying above this polygonal path. Forming the sequence T_i and taking successive differences we obtain the MLE. The number of equalities $S_i = T_i$ corresponds to the values $i = \tau_1, \tau_2, \cdots, \tau_m$ and the value is m, which proves the Theorem.

Definition. In a particular m partition $(t_1, t_2, \cdots, t_m)$ of k corresponding to the estimates $\bar{x}_{[t_1]}, \bar{x}_{[t_2]}, \cdots, \bar{x}_{[t_m]}$ let β_i denote the number of estimates of type i, that is the number of estimates obtained by pooling i sample means.

Notice that in a particular m partition many of the β's will be zero, that $0 \leqq \beta_j \leqq [k/j]$, the largest integer in (k/j), and that

$$\sum_{i=1}^{k} \beta_i = m, \tag{12}$$

$$\sum_{i=1}^{k} i\beta_i = k. \tag{13}$$

Two m partitions will be defined to belong to the same class if the β's of one are a permutation of the β's of the other.

Theorem 4. *Using the above nomenclature,*

$$p_{m,k} = \sum \prod_{i=1}^{k} i^{-\beta_i} (\beta_i!)^{-1}, \tag{14}$$

where the summation is over all the different classes giving m distinct estimates.

Proof. The result is immediate from Theorem 3 and the following Theorem of Sparre Andersen [2], which we shall state in our notation.

Theorem A. (*Sparre Andersen*). *Let the random variables* $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_k$ *be symmetrically dependent and let the joint distribution satisfy*

$$P[S_i/i = S_j/j] = 0, \qquad 1 \leqq i < j \leqq k.$$

Let H_k *be the number of equalities* $S_i = T_i$ $(i = 1, 2, \cdots, k - 1)$. *Then for* $m = 0, 1, 2, \cdots, k - 1$

$$P[H_k = m] = \sum \prod_{i=1}^{k} i^{-\beta_i} (\beta_i!)^{-1},$$

where the summation is over those values of $\beta_1, \beta_2, \cdots, \beta_k$ *from the set of values* $0, 1, 2, \cdots, k$ *for which* $\beta_1 + \beta_2 + \cdots + \beta_k = m + 1$ *and* $\beta_1 + 2\beta_2 + \cdots + k\beta_k = k$.

The only point to note is that we are counting the number of equalities $S_i = T_i$ $(i = 1, 2, \cdots, k)$ with a view to obtaining m estimates. Thus $p_{m,k} = P[H_k = m - 1]$.

COROLLARY. *Let $x_1, x_2, \cdots, x_k$ be random variables satisfying the condition of Theorem* A; *then the probability of the region in the sample space defined similarly to $\mathfrak{X}^*_{(m)}$ is also given as in Theorem* 4.

The proof of Theorem A given by Sparre Andersen is rather lengthy and is based on fairly general assumptions. It may be of interest to give, under the assumption of normality, an alternate proof of Theorem 4 based on a simple combinatorial argument.

ALTERNATE PROOF OF THEOREM 4. Using Lemma 1 and Equation (8),

$$p_{m,k} = \sum \left[\prod_{i=1}^{m} \frac{1}{t_i} \right] P(E_m),$$

where the summation is over the $\binom{k-1}{m-1}$ values corresponding to disjoint sets giving m distinct estimates. For all the cases in a class with values $\beta_1, \beta_2, \cdots, \beta_k$, $\prod_{i=1}^{m} 1/t_i = \prod_{i=1}^{k} i^{-\beta_i}$. Hence the probability of the region determined by a class is

$$\left[\prod_{i=1}^{k} i^{-\beta_i} \right] \sum{}' P(E_m),$$

where the summation is over all the cases in the class. To evaluate this sum, we notice that this is the conditional probability for the class given that the pooling of the samples has occurred. We may examine the problem in the following manner. Consider a population of m things belonging to k groups such that β_i of them belong to the ith group (elements belonging to the same group are assumed to be indistinguishable), and such that $\beta_1 + \beta_2 + \cdots + \beta_k = m$. They could be arranged in $m!/\beta_1!\, \beta_2! \cdots \beta_k!$ distinguishable orders. But the total number of ordering is $m!$. Thus the probability of an ordering with $\beta_1, \beta_2, \cdots, \beta_k$ is $\prod_{i=1}^{k} (\beta_i!)^{-1}$. Thus

$$p_{m,k} = \sum \prod_{i=1}^{k} (\beta_i!)^{-1} i^{-\beta_i},$$

where the summation is now over all the classes.

REMARKS. It is well known that

$$\sum \prod_{i=1}^{k} (-1)^{\beta_i} i^{-\beta_i} (\beta_i!)^{-1} = 0$$

where the summation is over all the classes. In terms of $p_{m,k}$, the above equation will be $\sum_{m=1}^{k} (-1)^m p_{m,k} = 0$. Since $\sum_{m=1}^{k} p_{m,k} = 1$ we conclude that, under H_0, the probability of an even number of distinct estimates, and the probability of an odd number of them, are both $\frac{1}{2}$.

Numerical computation of $p_{m,k}$. To use Theorem 4 we have to enumerate all the classes and the corresponding β's for fixed k and m. We need a method of listing all the partitions of k into m parts.

Begin with an initial partition having unity for each of the first $m - 1$ elements and $k - m - 1$ as the last element. To obtain a new partition from a given one, pass over the elements of the latter from right to left, stopping at the first element h which is less by at least two units than the last element. Without altering any element to the left of h, write $h + 1$ in place of h and every element to the right of h except that the last element is taken so as to give the sum k. Continue this procedure until we reach a partition in which no part differs from the last element by more than one unit. For example, if $k = 6$ and $m = 3$, then the classes are

$$\begin{matrix} 1 & 1 & 4 \\ 1 & 2 & 3 \\ 2 & 2 & 2 \end{matrix}$$

$$p_{3,6} = \frac{1}{2!}\cdot\frac{1}{4} + \frac{1}{2}\cdot\frac{1}{3} + \frac{1}{3!}\cdot\frac{1}{8} = \frac{225}{6!}$$

Remarks. It may be stressed that the proof of Theorem 2 involves the assumption of normality. Although we have proved Theorem 3 under the assumption of normality, the same proof holds for any continuous distribution where the estimation procedure in Section 2 is valid. In particular, it holds for populations belonging to the exponential family, since, as Brunk [9] has shown, the estimation procedure is valid. The alternate proof of Theorem 4 is based on Theorem 2 and hence depends on the assumption of normality. But Theorem 4, as stated in the corollary to the Theorem, is valid under more general conditions.

6. The second problem. A rank analogue of the standard one-way analysis of variance test was proposed and discussed in 1952 by Kruskal and Wallis [12]. When the sample sizes are equal to n, following the method in [12] and [13], we propose and investigate in this section a test similar to the test proposed in Section 4 for normal samples, with ranks replacing the original observations. The rank test is proposed only for the case of equal sample sizes since the theory is based on Theorem 4 which is valid only for symmetrically dependent random variables.

Let k independent random samples of equal size n be drawn from k univariate populations with unknown cumulative distributions F_i $(i = 1, 2, \cdots, k)$ respectively. To avoid the problem of ties we assume that each F_i is continuous. It is desired to test the hypothesis $H_0: F_1 = F_2 = \cdots = F_k$ against the alternative that the populations are stochastically ordered, i.e., $F_1 \geqq F_2 \geqq \cdots \geqq F_k$ with at least one inequality strong.

The proposed test procedure.

Step 1. Replace each observation x_{ij} by R_{ij}, its rank in the overall sample. Let $\bar{R}_i = (1/n) \sum_{j=1}^{n} R_{ij}$, $N = nk$.

Step 2. Formally operate on the $\bar{R}$'s as if one were obtaining normal MLE's and replace the $\bar{R}$'s by a set of nondecreasing means or pooled means in exactly the same way as was done for the $\bar{x}$'s in Section 2. Notice that pooling is very much simplified for the $\bar{R}$'s since the sample sizes are equal. Following our previous notation let us denote the final distinct set of m quantities thus obtained by $\bar{R}_{[t_1]}, \bar{R}_{[t_2]}, \cdots, \bar{R}_{[t_m]}$. The test statistic proposed is

$$\bar{H}_k = \frac{12n}{N(N+1)} \sum_{j=1}^{m} t_j \left[\bar{R}_{[t_j]} - \frac{N+1}{2} \right]^2 \tag{15}$$

and the test rejects H_0 for large values of the statistic.

THEOREM 5. *Let $\bar{R}_1, \bar{R}_2, \cdots, \bar{R}_k$ be the mean ranks of k samples. The test defined with the critical region $\bar{H}_k \geqq C$, where C is determined by*

$$\alpha = \sum_{m=2}^{k} p_{m,k} P[\chi^2_{m-1} \geqq C] \tag{16}$$

and χ^2_{m-1} is a random variable having the χ^2 distribution with $m - 1$ degrees of freedom, is, for large n, approximately a level α test of H_0.

PROOF. We notice that $\bar{H}_k$ is the H of Kruskal and Wallis [12] for m samples with sample sizes nt_j $(j = 1, 2, \cdots, m)$. Kruskal [13] proved that if H_0 is true the random variables $n(12^{\frac{1}{2}})[\bar{R}_i - (N+1)/2]/N^{\frac{3}{2}}$, $i = 1, 2, \cdots, k$ have asymptotically a degenerate symmetric multivariate normal distribution and that the distribution of $\bar{H}_k$ is approximately a central χ^2 with $m - 1$ degrees of freedom. Now apply the corollary to Theorem 4 and the proof is immediate.

Asymptotic Pitman efficiency. Restricting ourselves to translation alternatives we shall study the asymptotic Pitman efficiency of the nonparametric test based on $\bar{H}_k$ relative to the test based on T_k in the normal case with equal sample sizes. The method of proof is mainly an adaptation of the results of Andrews [3]. We shall follow Andrews in introducing the following assumptions.

ASSUMPTION 1. Let H_n specify the hypothesis that for each $i = 1, 2, \cdots, k$, $F_i(x) = F(x - \Delta_i/n^{\frac{1}{2}})$, where F is an arbitrary continuous distribution and $\Delta_1 \geqq \Delta_2 \geqq \cdots \geqq \Delta_k \geqq 0$ with at least one strict inequality.

ASSUMPTION 2. Let F possess a continuous derivative $f(x)$ except, perhaps, on a set of F measure zero.

ASSUMPTION 3. Let

$$\int_{-\infty}^{\infty} x^2 \, dF(x) - \left[\int_{-\infty}^{\infty} x \, dF(x) \right]^2 = \sigma_F^2$$

exist.

LEMMA A (Andrews). *If Assumptions 1 and 2 hold and if for any real number b*

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} n^{\frac{1}{2}} \left[F\left(x + \frac{b}{n^{\frac{1}{2}}} \right) - F(x) \right] dF(x)$$

exists and is finite, then the Kruskal-Wallis H based on k samples has asymptotically a noncentral χ^2 distribution with the noncentrality parameter

$$\lambda^H = 12 \left[\int_{-\infty}^{\infty} f(x) \, dF(x) \right]^2 \sum_{i=1}^{k} (\Delta_i - \bar{\Delta})^2, \tag{17}$$

where $\bar{\Delta} = \sum_{i=1}^{k} \Delta_i/k$.

We shall use the result proved by Andrews that the limiting distribution of $\bar{R}_1, \bar{R}_2, \cdots, \bar{R}_k$ under H_n is multivariate normal with the same covariance matrix had the hypothesis H_0 been true. Consider

$$\bar{H}_k = [12/N(N+1)]\sum_{i=1}^{m} \{\bar{R}_{[t_i]} - [(N+1)/2]\}^2.$$

The limiting distribution of $\bar{H}_k$ is noncentral χ^2 with $m - 1$ degrees of freedom. To compute the noncentrality parameter, we should obtain expressions for the limiting means. Following the same procedure as Andrews, the noncentrality parameter is

$$\lambda^{\bar{H}} = 12\left[\int_{-\infty}^{\infty} f(x)\, dF(x)\right]^2 \sum_{j=1}^{m} N_{[t_j]}\,[\bar{\Delta}_{[t_j]} - \bar{\Delta}]^2, \tag{18}$$

where $\bar{\Delta}_{[t_j]}$ is defined for the Δ's analogously to (1) and $\bar{\Delta} = \sum_{i=1}^{k} \Delta_i/k$.

DEFINITION. Let $p'_{m,k}(\theta)$ represent the probability of the region $\mathfrak{X}_{(t_1,t_2,\cdots,t_m)}$ under the hypothesis H_n.

The limiting power of the test based on $\bar{H}_k$ can be represented as

$$\sum p'_{m,k}(\theta)P[\chi^2_{m-1}(\lambda^{\bar{H}}) \geqq C],$$

where the summation is for all the $\binom{k-1}{m-1}$ regions $\mathfrak{X}_{(t_1,t_2,\cdots,t_m)}$ and for $m = 2, 3, \cdots, k$.

Consider the test based on T_k. Applying the results of Andrews, it is easy to show that under Assumptions 1, 2 and 3 the statistic T_k, in the region $\mathfrak{X}_{(t_1,t_2,\cdots,t_m)}$, has asymptotically a noncentral χ^2 distribution with noncentrality parameter.

$$\lambda^{T} = \sum_{j=1}^{m} N_{[t_j]}[\bar{\Delta}_{(t_j)} - \bar{\Delta}]^2/\sigma_F^2 \tag{19}$$

THEOREM 6. *If the distribution function F satisfies Assumptions* 1, 2 *and* 3, *then the asymptotic Pitman efficiency of the test based on $\bar{H}_k$ relative to the test based on T_k for equal sample sizes, is*

$$12\sigma_F^2\left[\int_{-\infty}^{\infty} f(x)\, dF(x)\right]^2. \tag{20}$$

PROOF. Let us fix the level of significance at α, and the limiting power at β. From (18) and (19), to obtain the same limiting power we should have

$$\sum p'_{mk}(\theta)P[\chi^2_{m-1}(\lambda^{\bar{H}}) \geqq C] = \sum p'_{mk}(\theta)P[\chi^2_{m-1}(\lambda^{T}) \geqq C].$$

Thus $\lambda^{\bar{H}} = \lambda^T$ for every $t_1, t_2, \cdots, t_m$. To have the same alternative we should have $\Delta_i/n^{\frac{1}{2}} = \Delta_i^*/(n^*)^{\frac{1}{2}}$. Thus, for each $t_1, t_2, \cdots, t_m$,

$$\lim \frac{n}{n^*} = 12\sigma_F^2\left[\int_{-\infty}^{\infty} f(x)\, dF(x)\right]^2,$$

which is a consequence of the theorem of Andrews and which shows that the relative efficiency of the test based on Kruskal-Wallis H compared to the classical F test is independent of the number of samples considered.

REMARKS. If F is the normal distribution, then the value of the relative efficiency is $3/\pi$. If F is the uniform distribution, then the value is 1.

Relationship to other tests. When $k = 2$ the test discussed in this section is the same as the one-tail test of Wilcoxon [16]. Jonckheere [11] suggested a distribution-free test analogous to the one-tail regression test in the normal case when the means increase by a constant amount, and hence, appropriate when such specific knowledge regarding the means is available. The results proved by Bartholomew [7] indicate that the test based on $\bar{H}_k$ should show larger power than Jonckheere's test when there is considerable variation in the differences between consecutive means. Whitney [15] proposes a test for $k = 3$ that is analogous to a normal-theory test one might consider when, under the alterna-

TABLE I

Table of $p_{m,k}$

m	k							
	3	4	5	6	7	8	9	10
1	.333333	.250000	.200000	.166667	.142857	.125000	.111111	.100000
2	.500000	.458333	.416667	.380556	.350000	.324107	.301984	.282897
3	.166667	.250000	.291667	.312500	.322222	.325694	.325519	.323165
4		.041667	.083333	.118055	.145833	.167882	.185417	.199427
5			.008333	.020833	.034722	.048611	.061863	.074219
6				.001389	.004167	.007986	.012500	.017436
7					.000198	.000694	.001505	.002604
8						.000025	.000099	.000240
9							.000003	.000012
10								.000000

TABLE II

Five per cent and one per cent values of T_k *for equal sample sizes* n

n	$k = 3$		$k = 4$		$k = 5$		$k = 6$	
	5%	1%	5%	1%	5%	1%	5%	1%
2	.687	.878	.590	.787	.518	.708	.461	.641
3	.455	.665	.392	.575	.345	.506	.308	.453
4	.337	.522	.292	.447	.258	.391	.231	.348
5	.267	.427	.233	.364	.206	.318	.184	.282
6	.221	.361	.193	.307	.170	.268	.153	.237
7	.189	.312	.165	.265	.146	.231	.131	.205
8	.164	.287	.144	.223	.128	.203	.115	.180
10	.131	.222	.113	.188	.102	.164	.092	.145
16	.081	.140	.071	.119	.064	.103	.057	.091

tive hypothesis, any two of the population means are equal. The test based on $\bar{H}_3$ is likely to be more efficient when the differences between consecutive means are nearly equal. There is a reduction in power through using the test based on $\bar{H}_3$ instead of the test based on T_k. But the former, although valid only for equal sample sizes, has the compensating advantage of being independent of the assumption of normality.

7. Tables of $p_{m,k}$ and tail distribution of T_k. Table I gives the values of $p_{m,k}$ for equal sample sizes and $k = 3, 4, \cdots, 10$. Table II gives the 5 percent and 1 percent values for the test statistic T_k.

Acknowledgments. I wish to acknowledge my indebtedness to Professor Erich L. Lehmann, under whose considerate guidance this research was done, and to Professor Joseph L. Hodges, Jr., for many valuable suggestions.

REFERENCES

[1] ANDERSEN, E. S. (1953). On the fluctuations of sums of random variables. *Math. Scand.* **1** 263–285.

[2] ANDERSEN, E. S. (1954). On the fluctuations of sums of random variables II. *Math. Scand.* **2** 195–223.

[3] ANDREWS, F. C. (1954). Asymptotic behaviour of some rank tests for analysis of variance. *Ann. Math. Statist.* **25** 724–735.

[4] BARTHOLOMEW, D. J. (1959). A test of homogeneity for ordered alternatives. *Biometrika* **46** 36–48.

[5] BARTHOLOMEW, D. J. (1959). A test of homogeneity for ordered alternatives. *Biometrika* **46** 328–335.

[6] BARTHOLOMEW, D. J. (1961). A test of homogeneity of means under restricted alternatives. *J. Roy. Statist. Soc. Ser. B* **23** 239–281.

[7] BARTHOLOMEW, D. J. (1961). Ordered tests in the analysis of variance. *Biometrika* **48** 325–332.

[8] BRUNK, H. D. (1955). Maximum likelihood estimates of monotone parameters. *Ann. Math. Statist.* **26** 607–616.

[9] BRUNK, H. D. (1958). On the estimation of parameters restricted by inequalities. *Ann. Math. Statist.* **29** 437–454.

[10] CHACKO, V. J. (1959). Testing homogeneity against ordered alternatives. Ph.D. thesis, Univ. of California, Berkeley.

[11] JONCKHEERE, A. R. (1954). A distribution free k sample test against ordered alternatives. *Biometrika* **41** 135–145.

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

[13] KRUSKAL, W. H. (1952). A nonparametric test for the several sample problem. *Ann. Math. Statist.* **23** 525–540.

[14] VAN EEDEN, C. (1957). Maximum likelihood estimation of partially or completely ordered parameters. *Indag. Math.* **19** 128–211.

[15] WHITNEY, D. R. (1951). A bivariate extension of the U statistic. *Ann. Math. Statist.* **22** 274–282.

[16] WILCOXON, F. (1945). Individual comparisons by ranking methods. *Biometrics Bull.* **1** 80–83.