SOME DISTRIBUTION-FREE STATISTICS AND THEIR APPLICATION TO THE SELECTION PROBLEM*

NOEL S. BARTLETT AND ZAKKULA GOVINDARAJULU

(Received Feb. 3, 1967)

0. Summary

The definition of the randomized rank-sum statistics introduced by Bell and Doksum [2] is extended to the case of k populations, and properties of the statistics are developed which generalize some of the results of Bell and Doksum [2]. Both the randomized rank-sum and the non-randomized rank-sum (heretofore called 'scores') statistics are employed in Gupta's [9] subset approach to the k population selection problem. A useful small-sample property of the randomized rank-sum procedure is demonstrated. The asymptotic relative efficiency (ARE) of the procedures based on the two types of rank-sum statistics is shown to be equal to unity, and the ARE of either relative to the procedure based on sample means is shown to be the same as that of the 'scores' procedure relative to the means procedure in the k-sample testing problem. The three procedures are shown to have the same asymptotic performance characteristic when sample size ratios equal the asymptotic relative efficiencies. It is further shown that an 'indifference zone' procedure based on the randomized rank-sum statistics and the 'indifference zone' 'scores' procedure are asymptotically equally efficient.

1. Introduction

Let $X_{i,j}$ $(i=1, 2, \cdots, k;\ j=1, 2, \cdots, n_i)$ be independent random samples drawn from populations with continuous cumulative distribution functions $F(x-\theta_i)$. For the problem of selecting a subset of the populations containing the one with the largest (smallest) θ-value, a natural procedure proposed by Gupta [9] is the means procedure which retains in the selected subset those populations whose sample means $\bar{X}_i$ are

* This research was supported by the Standard Oil Company (Ohio). It was also partially supported by the National Science Foundation Grant NSF-Gp 569, and by the Air Force Office of Scientific Research under Grant AF-AFOSR 741-765.

sufficiently close to the largest (smallest) sample mean. Likewise, for selecting the population with the largest (smallest) θ-value, the natural procedure investigated by Bachhofer [1] is the means procedure which selects the population associated with the largest (smallest) sample mean. The asymptotic relative efficiency (ARE) of two subset selection procedures in the case $n_1=n_2=\cdots=n_k$ is defined as the limiting ratio of the sample sizes required to achieve the same upper bound on the expected size of the selected subset for a given parameter configuration. The asymptotic relative efficiency of two 'indifference zone' selection procedures proposed by Bechhofer [1] is defined as the limiting ratio of the sample sizes required to achieve the same minimum probability of selecting a 'good' population. Lehmann [14] applied the 'scores' statistics to Bechhofer's [1] 'indifference zone' approach to the selection problem and showed that the asymptotic efficiency of the 'scores' procedure relative to the means procedure is the same as for the associated tests in the k-sample problem. Hence it is of interest and importance to consider, as alternatives to the means procedures for the selection problem, procedures based on randomized rank-sum statistics of the type introduced by Bell and Doksum [2] and procedures based on the 'scores' statistics already studied for the k-sample problem by numerous authors (for a list of which see [2]). Of concern are the asymptotic efficiencies of these alternative procedures relative to the means procedures.

2. Some results concerning distribution-free statistics based on ranks

Let $X_{i,j}$ $(i=1,\cdots,k;\ j=1,\cdots,n_i)$ be independent samples from populations Π_i having continuous cumulative distribution functions (cdf's) of the form $P(X_{i,j}\leqq x)=F(x-\theta_i)$. Let us pool the $\sum n_i=N$ observations on the k populations and rank them, denoting the rank of $X_{i,j}$ by $R_{i,j}$. Let H be any cdf. Let $Z_1,\cdots,Z_N$ be a random sample with cdf H, and let $Z(j)$ denote the jth smallest order statistic of $Z_1,\cdots,Z_N$. Then one defines the rank-sum statistics

$$S'_{N,i}(H)=n_i^{-1}\sum_{j=1}^{n_i}\mathrm{E}(Z(R(X_{i,j}))\,|\,H)\,,\qquad T'_{N,i}(H)=n_i^{-1}\sum_{j=1}^{n_i}Z(R(X_{i,j}))\,,\qquad i=1,\cdots,k \tag{2.1}$$

$$S_{N,i}(H)=S'_{N,i}(H)-S'_{N,k}(H)\,,\qquad T_{N,i}(H)=T'_{N,i}(H)-T'_{N,k}(H)\,,\qquad i=1,\cdots,k-1\,.$$

When $k=2$, the statistic $S_{N,1}(H)=S_N(H)$ is, for various H, equivalent to statistics considered by Fisher and Yates, Terry, Hoeffding, Wilcoxon, Hodges and Lehmann, Savage, Chernoff and Savage, Lehmann, Capon, and others for the two-sample problem, and the statistic $T_{N,1}(H)=T_N(H)$ has been proposed by Bell and Doksum [2] for the two-sample problem. Randomized statistics have also been considered by Durbin [5] and Ehrenberg [7]. The important properties of $T_{N,i}(H)$ follow, primarily, from the lemma of Bell and Doksum [2].

LEMMA 2.1.1. (Bell and Doksum [2], see p. 203). *Let F be a continuous cdf and let H be any cdf. If $W_1, \cdots, W_N$ and $Z_1, \cdots, Z_N$ are independent random samples with cdf's F and H, respectively, if $R(W_i)$ denotes the rank of W_i among $W_1, \cdots, W_N$ and if $Z(i)$ is the ith order statistic of $Z_1, \cdots, Z_N$, then $Z(R(W_1)), \cdots, Z(R(W_N))$ have the same joint distribution as the random sample $Z_1, \cdots, Z_N$.*

As an immediate consequence of this lemma, one obtains the following theorem, which generalizes, to the case $k>2$, a result of Bell and Doksum [2].

THEOREM 2.1. *When $\theta_1=\cdots=\theta_k$ (that is, $F_1=\cdots=F_k$), $T_{N,1}(H), \cdots, T_{N,k-1}(H)$ are jointly distributed as the differences of means of independent samples of sizes $n_1, \cdots, n_k$ from populations with cdf's H.*

PROOF. If $\theta_1=\cdots=\theta_k$, then $X_{1,1}, \cdots, X_{1,n_1}, \cdots, X_{k,1}, \cdots, X_{k,n_k}$ constitute a random sample of size N from a population with cdf $F=F_1=\cdots=F_k$. It follows from Lemma 2.1 that $Z(R(X_{1,1})), \cdots, Z(R(X_{k,n_k}))$ are independent samples with cdf H, consequently the $T_{N,i}(H)$ are means of independent samples of sizes n_i from a population with cdf H, and the theorem follows.

One notes that the randomized statistics $T_{N,i}(H)$ have the advantage over the non-randomized rank-sum statistics $S_{N,i}(H)$ of having a known and well-tabulated joint distribution when $F_1=\cdots=F_k$, for proper choices of H. In particular,

(i) If $H=\Phi$, the standard normal cdf, then $\{[n_in_k/(n_i+n_k)]^{1/2}T_{N,i}(\Phi)\}=\{[n_in_k/(n_i+n_k)]^{1/2}[n_i^{-1}\sum Z(R(X_{ij}))-n_k^{-1}\sum Z(R(X_{kj}))]\}$ has a $(k-1)$-variate normal cdf with zero mean vector, variances equal to unity and covariances $\sigma_{i,j}=\{n_in_j/(n_i+n_k)(n_j+n_k)\}^{1/2}$;

(ii) if $H=U$, the standard uniform cdf, then the cdf of

$$\left\{\frac{1}{2}+\frac{1}{2}T_{N,i}(U)\right\}=\left\{\frac{1}{2}+\frac{1}{2}[n_i^{-1}\sum Z(R(X_{i,j}))-n_k^{-1}\sum Z(R(X_{k,j}))]\right\}$$

is the same as the joint cdf of $k-1$ random variables, which are the sums of n_i+n_k standard uniform variables, and the marginal density of $\frac{1}{2}+\frac{1}{2}T_{N,i}(U)$ is $(m)^m[(m-1)!]^{-1}\sum(-1)^r\binom{m}{r}(x-r/m)^{m-1}$, where the summation is over $r\leqq mx$, and $m=n_i+n_k$;

(iii) if $H=K$, the standard exponential cdf on $[0, \infty)$, and if $n_1=\cdots=n_k=n$, then the joint cdf of $\{2nT_{N,i}(K)\}=\{2[\sum Z(R(X_{i,j}))-\sum Z(R(X_{k,j}))]\}$ is that of $k-1$ random variables, each of which is the difference of two independent chi-square variables, each with $2n$ degrees of freedom.

As one might expect from the construction of the $T_{N,i}(H)$, these statistics are in a certain sense asymptotically equivalent to their rank sum counterparts $S_{N,i}(H)$. We will note this from the theorems that follow, which are generalizations to the case $k>2$ of some of the results of Bell and Doksum [2].

LEMMA 2.2. *Let $a_{N,i,j}=1$ if the jth observation in the ordered combined sample of $N=\sum n_i$ observations is from Π_i, $a_{N,i,j}=0$ otherwise. Also let $\lambda_i=n_i/N$ $(i=1, 2, \cdots, k)$. Then one can write*

$$T_{N,i}(H)=(\lambda_i\lambda_k N)^{-1}\sum_{j=1}^{N} Z(j)(\lambda_k a_{N,i,j}-\lambda_i a_{N,k,j})$$

$$S_{N,i}(H)=(\lambda_i\lambda_k N)^{-1}\sum_{j=1}^{N} \mathrm{E}(Z(j)\mid H)(\lambda_k a_{N,i,j}-\lambda_i a_{N,k,j}) ,$$

where the Z's and the $a_{N,i,j}$'s are independent.

PROOF. Follows from straightforward calculations. For $k=2$ this reduces to Lemma 2.3 of Bell and Doksum [2].

We now generalize, to the case $k>2$, theorem 2.5 of Bell and Doksum [2].

THEOREM 2.2. *If H is any cdf, then*

(i) $\mathrm{E}(T_{N,i}(H)\mid F_1, \cdots, F_k)=\mathrm{E}(S_{N,i}(H)\mid F_1, \cdots, F_k)$ *for each* $i=1, \cdots, k-1$. *If H has second moments and if λ_i is bounded away from 0 and 1 for all i, then*

(ii) $\mathrm{Var}(N^{1/2}[T_{N,i}(H)-S_{N,i}(H)]\mid F_1=F_2=\cdots=F_k=F)\to 0$ *as* $N\to\infty$ *for each* $i=1, \cdots, k-1$ *and*

(iii) $\mathrm{Var}(N^{1/2}[T_{N,i}(H)-S_{N,i}(H)]\mid F_1, \cdots, F_k)\to 0$ *as* $N\to\infty$ *whenever one of the following is true*

$$\text{(a)}\quad A_{i,N}=o(N) \quad \text{as} \quad N\to\infty \qquad \text{or} \qquad \text{(b)}\quad A_{i,N}\leqq 0 ,$$

where

$$A_{i,N}=\sum_{s<t}\mathrm{Cov}\{Z(s), Z(t)\}\,\mathrm{E}[(\lambda_k a_{N,i,s}-\lambda_i a_{N,k,s})(\lambda_k a_{N,i,t}-\lambda_1 a_{N,k,t})\mid F_1, \cdots, F_k] .$$

PROOF. (i) follows immediately from Lemma 2.2, which also may be employed to show that

$$\begin{aligned}&\operatorname{Var}(N^{1/2}[T_{N,i}(H)-S_{N,i}(H)\mid F_1,\cdots,F_k)\\&=(\lambda_i\lambda_k)^{-2}N^{-1}\sum_{s=1}^{N}\operatorname{Var}(Z(s)\mid H)\,\mathrm{E}[(\lambda_k a_{N,i,s}-\lambda_i a_{N,k,s})^2\mid F_1,\cdots,F_k]\\&\quad+2(\lambda_i\lambda_k)^{-2}N^{-1}A_{i,N}\,.\end{aligned}$$

(iii) now follows immediately from Lemma 2.4 of Bell and Doksum [2]. (ii) then follows from (iii) (b) upon observing that, when $F_1=\cdots=F_k=F$,

$$\mathrm{E}[(\lambda_k a_{N,i,s}-\lambda_i a_{N,k,s})(\lambda_k a_{N,i,t}-\lambda_i a_{N,k,t})]=-\lambda_i\lambda_k(\lambda_i+\lambda_k)/(N-1)\,,\qquad\text{for}\quad s\neq t\,.$$

Since the $T_{N,i}(H)$ are asymptotically jointly normally distributed when $F_1=\cdots=F_k$ by virtue of Theorem 2.1, one obtains, using (i) and (ii) of Theorem 2.2, the following corollary.

COROLLARY 2.2.1. *If H has second moments, $F_1=\cdots=F_k$, and λ_i is bounded away from 0 and 1 for all i, then the distribution of $\{[n_in_k/(n_i+n_k)]^{1/2}\sigma_H^{-1}S_{N,i}(H)\}$ tends to a $(k-1)$-variate normal distribution with zero mean vector, variances equal to unity and covariances $\sigma_{i,j}=[n_in_j/(n_i+n_k)(n_j+n_k)]^{1/2}$ as $N\to\infty$.*

Similar results were obtained for $k=2$ by Dwass [6] using essentially the same approach. These results have also been obtained for $k=2$ by Chernoff and Savage [4] by means of a different approach, in which they invoke the regularity condition

$$\left|\frac{d^j}{du^j}J(u)\right|\leqq K[u(1-u)]^{-j-1/2+\delta}\,, \tag{2.2}$$

for $j=0,1,2$, for some $\delta>0$ and for $0<u<1$ where $J=H^{-1}$ and K is a constant. Puri [16] has extended Chernoff-Savage results for $k>2$. These results have been generalized for $k\geqq2$ by Govindarajulu et al. [8] under the weaker conditions that J be absolutely continuous and that (2.2) holds for $j=1$ and some $\delta>0$. The results of Govindarajulu et al. [8] hold when the F_i's are not all equal. Thus one has

COROLLARY 2.2.2. *If either* (a) *or* (b) *of Theorem* 2.2 *holds, J is absolutely continuous and* (2.2) *is satisfied for $j=1$ and some $\delta>0$, and λ_i is bounded away from* 0 *and* 1 *for all $i=1,\cdots,k-1$, then as $N\to\infty$ the cdf of the $\{(T_{N,i}(H)-\mu_{N,i})/\sigma_{N,i}\}$ tends to a $(k-1)$-variate normal distribution with zero mean vector and with covariance matrix $\sum$, where $\mu_{N,i}$ and $\sigma_{N,i}^2$ are the means and variances of the $S_{N,i}(H)$ and $\sum$ is the covariance matrix of the limiting joint normal distribution of the standardized $S_{N,i}(H)$.*

This corollary yields the asymptotic equivalence of $\{S_{N,i}(H)\}$ and $\{T_{N,i}(H)\}$ for all parameter points $(\theta_1,\cdots,\theta_k)$. The condition for this result is that (a) or (b) of Theorem 2.2 holds, which is somewhat difficult to verify. However, if only local asymptotic properties are desired, a more easily verifiable condition under which the result of part (iii) of Theorem 2.2 holds locally (that is, for those parameter points in a neighborhood of $\theta_1 = \cdots = \theta_k$) is the contiguity condition of Hajék and LeCam (see Hajék [11]). Define

$$g_\theta(u) = \frac{\partial}{\partial u} G_\theta(u) = \frac{\partial}{\partial u} F[F^{-1}(u) - \theta] , \qquad 0 < u < 1 \tag{2.3}$$

where $g_\theta(u)$ is assumed to be continuous with respect to θ in some non-degenerate closed interval about $\theta = 0$ for almost all u in $(0, 1)$. Also, let U denote the random variable distributed uniformly on $[0, 1]$. Then, the contiguity condition of Hajék and LeCam [11] is given by

$$\lim_{\theta \to 0} \mathrm{E}\{[(g_\theta^{1/2}(U) - g_0^{1/2}(U))]/\theta g_0^{1/2}(U)] - J(u)/2\} = 0 \tag{2.4}$$

where

$$H^{-1}(u) = J(u) = \left. \frac{\partial(2ng_\theta(u))}{\partial \theta} \right|_{\theta=0} . \tag{2.5}$$

If H has second moments and (2.4) is satisfied, then Lemma 3.1 of Matthes and Traux [15] can be used to show that $\mathrm{Var}\,(N^{1/2}[T_{N,i}(H) - S_{N,i}(H)]) \to 0$ as $N \to \infty$ for sequences of parameter points of the form $\theta_{i,n} - \theta_{k,n} = c_i n^{-1/2}$. Also, it is known from Bell and Doksum [3] that the function H given by (2.5) is 'optimal' in the sense that randomized tests based on this H are locally most powerful for the translation alternatives. However, the contiguity condition (2.4) and the optimal choice of H presupposes the knowledge of F. Hence, the preceding results are not adequate for our purposes.

We consider next, the application of randomized and non-randomized rank-sum statistics to the subset formulation of the selection problem, and determine their efficiency relative to the sample mean procedure.

3. Application of the distribution-free statistics to the subset selection problem

Suppose that one wishes to select a non-empty subset of the k populations, which he will assert contains the population with the largest θ-value. Using the means procedure (M) of Gupta [9], one will include Π_i in the subset if and only if $\bar{X}_i \geqq \max_j \bar{X}_j - c_M$, where c_M is a number chosen to satisfy the requirement

$$\mathrm{P}\{CS;\, c_M,\, n\} \geqq P^* \qquad \text{for all} \quad \tilde{\theta} = \{\theta_1, \cdots, \theta_k\} , \tag{3.1}$$

where $\mathrm{P}\{CS\}$ denotes the probability of correct selection, i.e. the probability of including in the subset the population with the largest θ-value.

If more than one population has the largest θ-value, one considers that a single 'tagged' population from among these is best.

Assuming, without loss of generality, that $\theta_1 \leqq \cdots \leqq \theta_{k-1} < \theta_k$, the left-hand side of (3.1) can be written $\mathrm{P}\{CS;\, c_M, n\} = \mathrm{P}(\bar{X}_k \geqq \bar{X}_j - c_M$, for $j=1, \cdots, k-1)$. Clearly $\mathrm{P}\{CS;\, c_M, n\}$ is a decreasing function of θ_j for $j=1, \cdots, k-1$, and, for all F, is smallest when

$$(3.2) \qquad \theta_1 = \cdots = \theta_{k-1} = \theta_k \,,$$

so that one determines c_M to satisfy

$$(3.3) \qquad \mathrm{P}(\bar{X}_i \leqq \bar{X}_k + c_M, \text{ for all } i) = P^*$$

when $\tilde{\theta}$ satisfies (3.2). Having determined $c_M(n)$ from (3.3), one may determine the common sample size n by imposing the additional requirement that the expected size of the subset (expected number of populations retained), $\mathrm{E}(S)$ satisfy

$$(3.4) \qquad \mathrm{E}\{S;\, c_M(n)\} \leqq 1 + \varepsilon$$

for some $\varepsilon > 0$, whenever $\tilde{\theta}$ lies in a given proper subset of the parameter space, for example, the subset defined by

$$(3.5) \qquad \theta_1 = \cdots = \theta_{k-1} = \theta_k - d^* \,, \qquad d^* > 0 \,.$$

In similar fashion, one can define a subset selection procedure based on each of the distribution-free statistics. The randomized rank-sum procedure $T(H)$ is to include Π_i in the subset if and only if $T'_{N,i}(H) \geqq \max_j T'_{N,j}(H) - c_T$, where $T'_{N,i}(H)$ is defined by (2.1) with $n_1 = \cdots = n_k$ and c_T is a number chosen to satisfy $\mathrm{P}\{CS;\, c_T, n\} \geqq P^*$ for all $\tilde{\theta}$. Since with probability unity $Z(j) \geqq Z(i)$ for $j > i$, $\mathrm{P}\{CS;\, \tilde{\theta}\} = \mathrm{P}\{n^{-1} \sum_j Z(R(X_{i,j})) \leqq n^{-1} \sum_j Z(R(X_{k,j})) + c_T$, for all $i\}$ is a non-increasing function of θ_i for $i=1, \cdots, k-1$. Thus $\mathrm{P}\{CS;\, T(H)\}$ is also minimized when (3.2) holds, and one determines c_T from the requirement

$$(3.6) \qquad \mathrm{P}(T'_{N,i}(H) \leqq T'_{N,k}(H) + c_T, \text{ for all } i) = P^*$$

when $\tilde{\theta}$ satisfies (3.2). The common sample size n_T for the procedure $T(H)$ may then be determined from a requirement similar to (3.4), namely

$$(3.7) \qquad \mathrm{E}\{S;\, c_T(n_T)\} \leqq 1 + \varepsilon \,,$$

when $\tilde{\theta}$ satisfies (3.5).

Finally, one can define a subset selection procedure $S(H)$ based on the non-randomized rank-sum statistics such that it includes Π_i in the subset if and only if $S'_{N,i}(H) \geqq \max_j S'_{N,j}(H) - c_S$, where $S'_{N,i}(H)$ is defined by (2.1) with $n_1 = \cdots = n_k$ and c_S is a number chosen to satisfy $\mathrm{P}\{CS;\, c_S, n\} \geqq P^*$ for all $\tilde{\theta}$. Since $\mathrm{E}(Z(j) \mid H)$ is a non-decreasing function of j, $\mathrm{P}\{CS;\, \theta_1, \cdots, \theta_k\} = \mathrm{P}\{S'_{N,i}(H) \leqq S'_{N,k}(H) + c_S$, for all $i\}$ is a non-increasing function of θ_i for $i = 1, \cdots, k-1$. It follows that $\mathrm{P}\{CS;\, S(H)\}$ is minimized when (3.2) holds, and one determines c_S from the requirement

$$\mathrm{P}(S'_{N,i}(H) \leqq S'_{N,k}(H) + c_S, \text{ for all } i) = P^* \tag{3.8}$$

when $\tilde{\theta}$ satisfies (3.2). The common sample size n_S for the procedure $S(H)$ may then be determined from

$$\mathrm{E}\{S;\, c_S(n_S)\} \leqq 1 + \varepsilon\ , \tag{3.9}$$

when $\tilde{\theta}$ satisfies (3.5).

Before proceeding to a comparison of the three procedures based on their asymptotic efficiencies, it is well to note the following interesting property of the randomized procedure $T(H)$, which holds for all sample sizes.

THEOREM 3.1. *For any continuous cdf F, and any cdf H, $c_T(k, n, P^*)$, the solution of* (3.6) *for given H, can be obtained as the solution of*

$$\mathrm{P}(\bar{Y}_i \leqq \bar{Y}_k + c_T, \text{ for } i = 1, \cdots, k-1 \mid H) = P^* \tag{3.10}$$

when $\tilde{\theta}$ satisfies (3.2), *that is, $c_T(k, n, P^*) = c_M(k, n, P^*)$, where $c_M(k, n, P^*)$ is the constant required to apply the means procedure for selecting a subset containing the largest θ-value when the cdf of Π_i is $H(x - \theta_i)$, for $i = 1, \cdots, k$.*

PROOF. As a consequence of Lemma 2.1, when $F_1 = \cdots = F_k$, the random variables $T'_{N,1}(H), \cdots, T'_{N,k}(H)$ are independent and identically distributed as means $\bar{Y}_1, \cdots, \bar{Y}_k$ of random samples of size n from populations with cdf's H. Thus (3.6) is equivalent to (3.10). Suppose that one has k populations $\Pi'_1, \cdots, \Pi'_k$ with cdf's of the form $\mathrm{P}(Y_i \leqq y) = H(y - \theta_i)$, and that one uses the means procedure for subset selection. The constant c_M required to apply the means procedure is to be determined from (3.3), where $\bar{X}_1, \cdots, \bar{X}_k$ are independent and identically distributed as means of random samples of size n from populations with cdf's H. Thus with $c_T = c_M$, (3.10) is equivalent to (3.3), and the theorem follows.

This means that for the subset approach to the selection problem, the procedure $T(H)$ has the advantage over the procedure $S(H)$ that one can calculate, or determine from already existing tables, the constant necessary to carry out the former procedure, for any given sample size. For instance, one has

Example 3.1. (i) If $H=\Phi$, the standard normal cdf, then, for any continuous cdf F, $c_T(k, n, P^*)$ for the randomized rank-sum procedure for selecting a subset containing the best population when Π_i has cdf $F(x-\theta_i)$ is the same as $c_M(k, n, P^*)$ for the means procedure for selecting a subset containing the normal population with the largest mean when the variances are known (and equal to unity). Gupta [10] provides values of $(n/2)^{1/2}c_M(n, k, P^*)$ for $k=2(1)51$ and $P^*=.75$, .90, .95, .975 and .99.

If, in particular, F is a normal distribution with unknown variance, then the problem is to pick a subset containing the normal population with the largest mean when the populations have common unknown variance, and by using the randomized rank-sum procedure with $H=\Phi$ one can evaluate the constant c_T as though he were dealing with normal populations with known variances.

Note that it is not possible to find n_T to satisfy (3.7) for the randomized rank-sum procedure, since the joint distribution of the $T'_{N,i}(H)$, $i=1, \cdots, k$ is not known when the F_i are not all the same, that is, when the θ_i are not all equal.

It will be convenient in what follows to replace (3.5), when the sample size is n, by

$$\theta_1=\cdots=\theta_{k-1}=\theta_k-d^{(n)} . \tag{3.11}$$

The sample sizes n, n_S, and n_T for the means, non-randomized rank-sum, and randomized rank-sum procedures will be determined from (3.4), (3.7) and (3.9), respectively when $\tilde{\theta}$ satisfies (3.11). We are not changing our definition of goodness of a procedure with sample size, since we will later associate $d^{(n)}$ with d^*. This temporary device emphasizes that we are first going to consider d as a function of n, rather than n as a function of d.

LEMMA 3.1. *For fixed P^*, if $c_M(n)$ is determined so that* (3.3) *holds, then as $n\to\infty$,*

$$c_M(n)=n^{-1/2}c\sigma+o(n^{-1/2}) , \tag{3.12}$$

where σ^2 is the variance of F and c is the solution of

$$Q(2^{-1/2}c, \cdots, 2^{-1/2}c)=P^* , \tag{3.13}$$

where Q is the cdf of a normally distributed vector $(W_1, \cdots, W_{k-1})$ with

$$\text{(3.14)}\quad \mathrm{E}(W_i)=0\ ,\quad \operatorname{Var}(W_i)=1\ ,\quad \operatorname{Cov}(W_i, W_j)=\frac{1}{2}\quad \text{for all}\quad i,\ j\ .$$

PROOF. The condition $\bar{X}_k \geqq \bar{X}_{\max}-c_M(n)$ is equivalent to $Y_i \leqq (2\sigma^2/n)^{-1/2}c_M(n)$, where $Y_i=(2\sigma^2/n)^{-1/2}(\bar{X}_i-\bar{X}_k)$, for each $i=1, \cdots, k-1$. Since $\theta_1=\cdots=\theta_k$, the Y_i certainly satisfy (3.14). By the central limit theorem the cdf of the random vector $(Y_1, \cdots, Y_{k-1})$ converges, uniformly in its arguments, to the cdf of the normally distributed vector $(W_1, \cdots, W_{k-1})$. Consequently requirement (3.3) is equivalent to $\lim_{n\to\infty} \mathrm{P}(W_i \leqq (2\sigma^2/n)^{-1/2}c_M(n)$, for all $i)=P^*$. If c is defined by (3.13), this equation will be satisfied if and only if (3.12) holds.

LEMMA 3.2. *For given* ε, *with* $c_M(n)$ *given by* (3.12) *and if* n *is determined from* (3.4) *when* $\tilde{\theta}$ *satisfies* (3.11), *then as* $n \to \infty$

$$\text{(3.15)}\qquad d^{(n)}=n^{-1/2}\, d\sigma+o(n^{-1/2})\ ,$$

where d *is the solution of*

$$\text{(3.16)}\qquad Q(2^{-1/2}(c+d), \cdots, 2^{-1/2}(c+d)) \\ +(k-1)\, Q(2^{-1/2}c, \cdots, 2^{-1/2}c, 2^{-1/2}(c-d))=1+\varepsilon$$

where Q *is the cdf of a normally distributed vector satisfying* (3.14).

PROOF. One can easily show that $\mathrm{E}\{S\}=\sum_{j=1}^{k} \mathrm{P}\{\Pi_j$ is included in the subset$\}$, hence, $\mathrm{E}\{S\}=\sum_j \mathrm{P}\{\bar{X}_j \geqq \bar{X}_i-c_M(n)$, for all $i\neq j\}$. Let $\tilde{\theta}_n=(\theta_{1,n}, \cdots, \theta_{k,n})$ be a sequence of parameter points satisfying (3.11), and let $Y_{i,j}=(2\sigma^2/n)^{-1/2}(\bar{X}_i-\bar{X}_j-\theta_{i,n}+\theta_{j,n})$, for $i, j=1, \cdots, k$. Then for each j and n, the random vector $\{Y_{i,j}, i\neq j\}$ satisfies (3.14). Furthermore, expected value of S is equal to $\sum_j \mathrm{P}\{Y_{i,j} \leqq (2\sigma^2/n)^{-1/2}(\theta_{j,n}-\theta_{i,n}+c_M(n))$, for all $i\neq j\}$. By the central limit theorem each of the cdf's in the last expression tends to a normal cdf uniformly in its arguments. Thus (3.4) is equivalent to

$$\lim_{n\to\infty} [\mathrm{P}\{Y_{i,k} \leqq (2\sigma^2/n)^{-1/2}(c_M(n)+d^{(n)}), i<k\}+(k-1)\mathrm{P}\{Y_{i,1} \leqq (2\sigma^2/n)^{-1/2}c_M(n), \\ 1<i<k,\ Y_{k,1} \leqq (2\sigma^2/n)^{-1/2}(c_M(n)-d^{(n)})\}]=1+\varepsilon\ .$$

If c is defined by (3.13), $c_M(n)$ by (3.12), and d by (3.16), this equation will be satisfied if and only if (3.15) holds. Note that there is a unique solution d of (3.16) for given c_M and ε since the left-hand side of (3.16) is a strictly decreasing function of d.

Now if one is given a value d^*, and wishes to find the sample size

n for which (3.4) is satisfied when $\tilde{\theta}$ satisfies (3.5), then he sets $d^* = d^{(n)} = n^{-1/2}dc$, or

$$n = (d\sigma/d^*)^2 . \tag{3.17}$$

This defines n as a function of k, P^*, d^*, and ε.

Consider now the rank-sum procedures. From Theorem 3.1 and Lemma 3.1, one obtains the following lemma for the randomized rank-sum procedure $T(H)$.

LEMMA 3.3. *For fixed P^*, let $c_T(n_T)$ be determined so that* (3.6) *holds, and suppose that H has second moments. Then as $n_T \to \infty$,*

$$c_T(n_T) = n_T^{-1/2} cA + o(n_T^{-1/2}) , \tag{3.18}$$

where c is defined by (3.13) *and A^2 is defined by*

$$A^2 = \int_0^1 J^2(u)\,du - \left(\int_0^1 J(u)\,du\right)^2 ,$$

where $J = H^{-1}$ and thus $\sigma_H^2 = A^2$.

One expects the same result to be true for the non-randomized procedure, and in fact one obtains

LEMMA 3.4. *For fixed P^*, let $c_S(n_S)$ be determined so that* (3.8) *holds and suppose that either* (i) *H has second moments, or* (ii) *$J = H^{-1}$ is absolutely continuous and satisfies* (2.2) *for $j=1$ and some $\delta > 0$. Then as $n_S \to \infty$,*

$$c_S(n_S) = n_S^{-1/2} cA + o(n_S^{-1/2}) , \tag{3.19}$$

where c is defined by (3.13).

PROOF. A correct selection occurs if, for each $i = 1, \cdots, k-1$, $S_{N,i}(H) \leqq c_S(n_S)$. By Corollary 2.2.1 and the comments immediately following it, if (i) or (ii) is satisfied then the vector $\{(2A^2/n_S)^{-1/2} S_{N,1}(H), \cdots, (2A^2/n_S)^{-1/2} S_{N,k-1}(H)\}$ tends to a normally distributed vector $(W_1, \cdots, W_{k-1})$ which satisfies (3.14). The requirement (3.8) is therefore equivalent to

$$\lim_{n_S \to \infty} \mathrm{P}\{ W_i \leqq (2A^2/n_S)^{-1/2} c_S(n_S), \text{ for all } i\} = P^* ,$$

so that with c defined by (3.13), this equation will hold if and only if (3.19) is satisfied.

The following theorem, which is analogous to Lemma 3.2, enables one to determine the sample size required for the non-randomized rank-sum procedure and to establish the ARE of $S(H)$ relative to the means procedure. One can then use Corollary 2.2.2 to prove a corollary to the theorem, which states the same results for the randomized rank-sum procedure.

THEOREM 3.2. *With $c_S(n_S)$ given by* (3.19), *let n_S be determined so that* (3.7) *holds when $\tilde{\theta}$ satisfies* (3.5). *If* (ii) *of Lemma* 3.4 *holds, then as $n_S \to \infty$,*

$$d^{(n_S)} = n_S^{-1/2} dA\left[\int_{-\infty}^{\infty} \frac{d}{dx}\{J[F(x)]\}\, dF(x)\right]^{-1} + o(n_S^{-1/2}) , \tag{3.20}$$

where d is defined by (3.16).

PROOF. Let $\theta_{n_S} = (\theta_{1,n_S}, \cdots, \theta_{k,n_S})$ be a sequence of parameter points satisfying (3.11). Then one sees that

$$\begin{aligned} E(S) &= \sum_{j=1}^{k} P\{S'_{N,j}(H) \geqq S'_{N,i}(H) - c_S(n_S), \text{ for all } i \neq j\} \\ &= \sum_{j=1}^{k} P\{(2A^2/n_S)^{-1/2}[S'_{N,i}(H) - S'_{N,j}(H) - \mu_i(\theta_{n_S}) + \mu_j(\theta_{n_S})] \\ &\qquad \leqq [\mu_j(\theta_{n_S}) - \mu_i(\theta_{n_S}) + c_S(n_S)](2A^2/n_S)^{-1/2},\ i \neq j\} . \end{aligned}$$

The vector $\{n_S^{1/2}[S'_{N,1}(H) - \mu_1(\theta_{n_S})], \cdots, n_S^{1/2}[S'_{N,k}(H) - \mu_k(\theta_{n_S})]\}$, is asymptotically normal with the limiting covariance matrix $\sigma_{i,j} = (\delta_{i,j} - k^{-1})A^2$, where $\delta_{i,j}$ are Kronecker deltas. (See Puri [16] or Govindarajulu, et al. [8].) Therefore by virtue of Lemma 3 of Lehmann [14], the random vector $\{U_{i,j}, i \neq j\}$ has for each $j = 1, \cdots, k$, an asymptotic normal distribution which satisfies (3.14), where $U_{i,j} = (2A^2/n_S)^{-1/2}[S'_{N,i}(H) - S'_{N,j}(H) - \mu_i(\theta_{n_S}) + \mu_j(\theta_{n_S})]$. We temporarily defer the proof that

$$\lim_{n_S \to \infty} [\mu_i(\theta_{n_S}) - \mu_i(0)] = (\delta_{i,k} - k^{-1})\, d^{(n_S)} \int_{-\infty}^{\infty} \frac{d}{dx}\{J[F(x)]\}\, dF(x) , \quad \text{for } i = 1, \cdots, k , \tag{3.21}$$

from which it immediately follows that

$$\lim_{n_S \to \infty} [\mu_j(\theta_{n_S}) - \mu_i(\theta_{n_S})] = (\delta_{j,k} - \delta_{i,k})\, d^{(n_S)} \int_{-\infty}^{\infty} \frac{d}{dx}\{J[F(x)]\}\, dF(x) , \quad \text{for } i, j = 1, \cdots, k .$$

Using these results, one notes that the equation for determing n_S becomes

$$\lim_{n_S \to \infty} E(S) = P\left\{ U_{i,k} \leqq (2A^2/n_S)^{-1/2}\left[d^{(n_S)} \int_{-\infty}^{\infty} \frac{d}{dx}\{J[F(x)]\}\, dF(x) + c_S(n_S) \right], \quad 1 \leqq i < k \right\}$$

$$+(k-1)\,\mathrm{P}\Big\{U_{i,1}\leqq(2A^2/n_S)^{-1/2}\Big[c_S(n_S)$$

$$-\delta_{i,k}\,d^{(n_S)}\int_{-\infty}^{\infty}\frac{d}{dx}\{J[F(x)]\}\,dF(x)\Big],\ 1<i\leqq k\Big\}$$

$$=1+\varepsilon\ .$$

Thus, if d is defined by (3.16), the preceding equation will be satisfied if and only if $d^{(n_S)}$ satisfies (3.20).

To complete the proof of Theorem 3.2, one needs to demonstrate the validity of (3.21). To this end let $d_i^{(n_S)}$ be defined by $d_i^{(n_S)}=(1-\delta_{i,k})\,d^{(n_S)}$, for $i=1,\cdots,k$. θ_{n_S} will denote the configuration $\theta_{i,n_S}=\theta_{k,n_S}-d_i^{(n_S)}$, for $i=1,\cdots,k-1$. Setting $F_j(x)=F(x+d_j^{(n_S)})$, and applying the definition of $\mu_i(\theta)$, one obtains

$$\mu_i(\theta_{n_S})-\mu_i(0)=\int_{-\infty}^{\infty}J\Big[k^{-1}\sum_{j=1}^{n}F_j(x)\Big]\,dF_i(x)-\int_{-\infty}^{\infty}J[F(x)]\,dF(x)$$

$$=\int_{-\infty}^{\infty}\Big\{J\Big[k^{-1}\sum_{j=1}^{k}F(x+d_j^{(n_S)}-d_i^{(n_S)})\Big]-J[F(x)]\Big\}\,dF(x)\ ;$$

The left-hand side of (3.21) is thus equivalent to $\int A_{n_S}(x)B_{n_S}(x)\,dF(x)$, where

$$A_{n_S}(x)=\Big\{J\Big[k^{-1}\sum_{j=1}^{k}F(x+d_j^{(n_S)}-d_i^{(n_S)})\Big]-J[F(x)]\Big\}\Big/B_{n_S}(x)$$

and

$$B_{n_S}(x)=k^{-1}\sum_{j=1}^{k}F(x+d_j^{(n_S)}-d_i^{(n_S)})-F(x)\ .$$

Now, $d_j^{(n_S)}\to 0$ for $j=1,\cdots,k$ as $n_S\to\infty$. Thus, under the assumed regularity conditions,

$$\lim_{n_S\to\infty}A_{n_S}(x)=\frac{d}{du}J(u)\Big|_{u=F(x)}\ ,\quad\text{and}\quad\lim_{n_S\to\infty}B_{n_S}(x)=k^{-1}\sum_{j=1}^{k}(d_j^{(n_S)}-d_i^{(n_S)})F'(x)\ .$$

It follows that

$$\lim_{n_S\to\infty}[\mu_i(\theta_{n_S})-\mu_i(0)]=k^{-1}\sum_{j=1}^{k}(d_j^{(n_S)}-d_i^{(n_S)})\int_{-\infty}^{\infty}\frac{d}{dx}\{J[F(x)]\}\,dF(x)\ ,$$

from which (3.21) follows. This completes the proof of Theorem 3.2.

COROLLARY 3.2.1. *With $c_T(n_T)$ given by* (3.18), *let n_T be determined so that* (3.7) *holds when $\tilde{\theta}$ satisfies* (3.11). *If either* (a) *or* (b) *of Theorem* 2.2 *holds and $J=H^{-1}$ is absolutely continuous and satisfies* (2.2) *for $j=1$ and some $\delta>0$, then as $n_T\to\infty$,*

$$d^{(n_T)} = n_T^{-1/2} dA \left[\int_{-\infty}^{\infty} \frac{d}{dx} \{J[F(x)]\} \, dF(x) \right]^{-1} + o(n_T^{-1/2}) \,, \tag{3.22}$$

where d is defined by (3.16).

Now, given d^*, one wishes, as with the means procedure, to find the common sample sizes n_S and n_T, for the rank-sum procedures, for which (3.9) and (3.7) are satisfied when $\tilde{\theta}$ satisfies (3.5). This is done by letting $d^* = d^{(n_S)} = d^{(n_T)}$, and from (3.20) and (3.22) one obtains

$$n_S = n_T = \left[dA \Big/ \left\{ d^* \int_{-\infty}^{\infty} \frac{d}{dx} \{J[F(x)]\} \, dF(x) \right\} \right]^2 . \tag{3.23}$$

From this result and (3.17) one concludes that

THEOREM 3.3. *The non-randomized and randomized rank-sum procedures are asymptotically equally efficient for the subset approach to the selection problem, and the ARE of either relative to the means procedure is*

$$A(T(H), M) = A(S(H), M) = \lim_{n \to \infty} \frac{n}{n_S} \left[\sigma \int_{-\infty}^{\infty} \frac{d}{dx} \{J[F(x)]\} \, dF(x)/A \right]^2 . \tag{3.24}$$

This ARE of the non-randomized rank-sum procedure to the means procedure is the same as that found by Chernoff and Savage [4] and Puri [16] for the two-and k-sample problems, and by Lehmann [14] for the indifference zone formulation of the selection problem. Therefore, as is by now well known, one has

Example 3.2. (i) If $H = \Phi$, or in fact if H is any normal distribution, then $A(S(\Phi), M) = A(T(\Phi), M) \geqq 1$ for all F, with equality if and only if F is normal. The non-randomized procedure when $H = \Phi$ is the well-known normal scores procedure, which is also known to be asymptotically equivalent to the Van der Waerden X-test [17].

(ii) If $H = U$, the standard uniform distribution, or any uniform distribution, then, as shown by Hodges and Lehmann [12], $A(S(U), M) = A(T(U), M) \geqq .864$ for all F; $A(S(U), M) = A(T(U), M) = 3/\pi \sim .955$ when F is normal; and $A(S(U), M) = A(T(U), M) > 1$ for many non-normal distributions. (For comparison of the efficiency of this procedure, often called the rank-sum procedure, relative to the normal scores procedure, see Hodges and Lehmann [13].)

We have studied three procedures for selecting a subset containing the best population, imposing the common requirement that the minimum probability of including the true best population is the same for all procedures, and have found that if the sample sizes are related as in (3.24), then the expected size of the subset for parameter points satisfying

(3.11) is approximately the same for all three procedures. One may ask whether, for parameter points not satisfying (3.11) or (3.2), the probability of correct selection is asymptotically the same for all procedures. The answer is contained in the following theorem.

THEOREM 3.4. *Let n and $n_S = n_T$ be related as in* (3.24), *that is, let the sample sizes be determined by identical requirements. Then $S(H)$, $T(H)$ and the means procedure have the same asymptotic probability of correct selection and the same expected subset size for any parameter configuration.*

PROOF. Let us investigate first the behavior of the means procedure, considering any sequence of parameter points satisfying

$$\theta_{k,n} - \theta_{i,n} = d_{i,n} = n^{-1/2} d_i \sigma + o(n^{-1/2}) , \tag{3.25}$$

for $i = 1, \cdots, k-1$. The joint limiting cdf of the random variables $Y_i = (2\sigma^2/n)^{-1/2}(\bar{X}_i - \bar{X}_k)$ is the same as that of a $(k-1)$-dimensional normal vector $(W_1, \cdots, W_{k-1})$ for which

$$\mathrm{E}(W_i) = -2^{-1/2} d_i , \qquad \mathrm{Var}(W_i) = 1 , \qquad \mathrm{Cov}(W_i, W_j) = \frac{1}{2} \quad \text{for all} \quad i, j . \tag{3.26}$$

Therefore, since $\mathrm{P}\{CS\} = \mathrm{P}(\bar{X}_k \geqq \bar{X}_{\max} - c_M(n))$,

$$\lim_{n\to\infty} \mathrm{P}\{CS;\, d_1, \cdots, d_{k-1}\} = Q(2^{-1/2}(c + d_1), \cdots, 2^{-1/2}(c + d_{k-1})) . \tag{3.27}$$

Furthermore, if one defines $p_j = \mathrm{P}$ (include Π_j in subset), then $p_j = \mathrm{P}(\bar{X}_j \geqq \bar{X}_{\max} - c_M(n))$. The joint limiting distribution of the random variables $Y_{i,j} = (2\sigma^2/n)^{-1/2}(\bar{X}_i - \bar{X}_j)$, for $i \neq j$, is the same as that of a $(k-1)$-dimensional normal vector $(W_{1,j}, \cdots, W_{j-1,j}, W_{j+1,j}, \cdots, W_{k,j})$ for which

$$\mathrm{E}(W_{i,j}) = 2^{-1/2}(d_j - d_i) , \qquad \mathrm{Var}(W_{i,j}) = 1 , \qquad \mathrm{Cov}(W_{i,j}, W_{t,j}) = \frac{1}{2} , \quad \text{for all} \quad i, t \tag{3.28}$$

where $d_k = 0$ by definition. It follows that

$$\lim_{n\to\infty} p_j(d_1, \cdots, d_{k-1}) = \mathrm{P}\{W_{i,j} \leqq (2\sigma^2/n)^{-1/2} c_M(n) \text{ for all } i \neq j\} = Q(2^{-1/2}(c + d_1 - d_j), \cdots, 2^{-1/2}(c + d_k - d_j)) . \tag{3.29}$$

Let us now consider the non-randomized and randomized rank-sum procedures. Using the asymptotic joint normality of the random variables $n_S^{1/2}[S'_{N,i}(H) - \mu_i(\theta_{n_S})]$, Lehmann [14] has shown (Theorem 1) that the mean of $S_{N,i}(H)$ in the limiting distribution of $(S_{N,1}(H), \cdots, S_{N,k-1}(H))$ is equal to $-d_i/2^{1/2}$, that is, for $i=1, \cdots, k-1$,

$$\lim_{n\to\infty} (2A^2/n_S)^{-1/2}[\mu_i(\theta_n)-\mu_k(\theta_n)] = -d_i/2^{1/2} , \tag{3.30}$$

when n and n_S are related by (3.24). Consequently, the joint limiting distribution of the random variables $(2A^2/n_S)^{-1/2}S_{N,i}(H)$ is that of a $(k-1)$-dimensional normal vector satisfying (3.26). For the procedure $S(H)$, $\mathrm{P}\{CS; S(H)\} = \mathrm{P}(S_{N,i}(H) \leqq c_S(n_S))$; thus (3.27) must also hold for $S(H)$, and this means that $S(H)$ has the same asymptotic performance characteristic as the means procedure, where the characteristic is taken to be the probability of correct selection. The same must also be true of the procedure $T(H)$ under the conditions of Corollary 2.2.2.

Now for $S(H)$, $p_j = \mathrm{P}(S'_{N,j}(H) \geqq S'_{N,i}(H) - c_S(n_S)$, for $i \neq j$). It follows from (3.30) that the joint limiting distribution of the random variables $S_{N,i,j}(H) = (2A^2/n_S)^{-1/2}[S'_{N,i}(H) - S'_{N,j}(H)]$ for $i \neq j$ is the same as that of a $(k-1)$-dimensional normal vector $(W_{1,j}, \cdots, W_{j-1,j}, W_{j+1,j}, W_{k,j})$ satisfying (3.28), from which one concludes that $S(H)$, and therefore $T(H)$, satisfies (3.29). Since $\mathrm{E}(S) = \sum_{j=1}^{k} p_j$, this means that $S(H)$, $T(H)$ and the means procedure also have the same asymptotic performance characteristic when that characteristic is taken to be the expected size of the subset.

Let us observe that these remarks remain true for sequences of parameter points for which not all the differences $\theta_{k,n} - \theta_{i,n}$ tend to zero at the $n^{-1/2}$ rate assumed in (3.25). If any difference tends to zero more rapidly, we replace d_1 by 0, and if it tends to zero more slowly, or tends to a finite limit, then we replace d_i by ∞, and still obtain the same asymptotic behavior. This completes the proof of Theorem 3.4.

4. Application of the distribution-free statistics to the indifference zone formulation of the selection problem

Suppose now that one desires to select from among the k populations a single population which he will assert has the largest θ-value. Using the means procedure (M) of Bechhofer [1], one will select Π_i as best if $\bar{X}_i = \max_s \bar{X}_s$ where $\bar{X}_s = n^{-1} \sum_{j=1}^{n} X_{s,j}$. One determines the common sample size n from the requirement that $\mathrm{P}\{CS\}$, the probability of correct selection of the true best population, satisfy

$$\mathrm{P}\{CS; n\} \geqq P^* , \tag{4.1}$$

whenever

(4.2) $$\theta_i < \theta_k - d^* , \qquad i=1, \cdots, k-1 ,$$

where P^* and d^* are given constants, and where, without loss of generality, we suppose that $\theta_k = \theta_{\max}$.

When (4.2) holds and one is using the means procedure, $\mathrm{P}\{CS; n\} = \mathrm{P}(Y_j \leqq Y_k + \theta_k - \theta_j$, for $j=1, \cdots, k-1)$, where $Y_j = \bar{X}_j - \theta_j$ $(j=1, \cdots, k)$ are independent, identically distributed random variables with cdf $F(y)$. Clearly $\mathrm{P}\{CS; n\}$ is a decreasing function of θ_j for all $j<k$, so that, for all F, $\mathrm{P}\{CS; n\}$ takes on its minimum value when $\tilde{\theta}$ satisfies (3.5), and one determines the sample size from the condition

(4.3) $$\mathrm{P}\{CS; n\} = \mathrm{P}(\bar{X}_k = \bar{X}_{\max}) = P^*$$

when $\tilde{\theta}$ satisfies (3.5). As alternatives to the means procedure one defines the non-randomized rank-sum procedure $S(H)$, under which one selects Π_i as best if $S'_{N,i}(H) = \max_j S'_{N,j}(H)$, and the randomized rank-sum procedure $T(H)$, under which one selects Π_i as best if $T'_{N,i}(H) = \max_j T'_{N,j}(H)$, where $S'_{N,j}(H)$ and $T'_{N,j}(H)$ are defined by (2.1) $n_1 = n_2 = \cdots = n_k$. The procedure $S(H)$ just defined was suggested and its properties investigated by Lehmann [14].

The sample sizes n_S and n_T needed for $S(H)$ and $T(H)$, respectively, are determined from the requirement that (4.1) hold whenever $\tilde{\theta}$ satisfies (4.2), for given P^* and d^*. Because, as was mentioned in discussing the subset formulation, $\mathrm{E}(Z(j) \mid H)$ and $Z(j)$ are non-decreasing functions of j, $\mathrm{P}(S'_{N,k}(H) \geqq S'_{N,j}(H))$ and $\mathrm{P}(T'_{N,k}(H) \geqq T'_{N,j}(H))$ are non-increasing functions of θ_j for all $j<k$. Therefore, the infimum of $\mathrm{P}\{CS; S(H)\}$ and $\mathrm{P}\{CS; T(H)\}$ over all parameter points (4.2) occurs at the configuration (3.5), and one determines n_S and n_T from the conditions

(4.4) $$\mathrm{P}(S'_{N,k}(H) \geqq S'_{N,i}(H), \text{ for all } i) = P^*$$

when $\tilde{\theta}$ satisfies (3.5), and

(4.5) $$\mathrm{P}(T'_{N,k}(H) \geqq T'_{N,i}(H), \text{ for all } i) = P^*$$

when $\tilde{\theta}$ satisfies (3.5). Let us note first of all that for this 'indifference zone' formulation there is no result comparable to Theorem 3.1, that is, one cannot compute n_T explicitly for $T(H)$, since under the configuration (3.5) the joint distribution of the $T'_{N,i}(H)$ is not known. Thus $T(H)$ has no apparent advantage over $S(H)$ for finite sample sizes in the present formulation, as it did in the subset formulation.

The results of Lehmann [14] for the means procedure and the procedure $S(H)$ will now be briefly mentioned for the sake of completeness. It will then be shown, using the results of Section 2, that the same results are true for the randomized procedure $T(H)$.

To this end, consider again a sequence of situations for increasing n, in which (4.3), (4.4), and (4.5) are required to hold when $\tilde{\theta}$ satisfies (3.11). For the means procedure, if n is determined by (4.3) when $\tilde{\theta}$ satisfies (3.11), then as $n \to \infty$ (3.15) holds, where d is the solution of (3.13) with d in place of c (for a proof, see Lemma 1 of [14]). For the procedure $S(H)$, if n_S is determined by (4.4) when $\tilde{\theta}$ satisfies (3.11) and if F and $J=H^{-1}$ satisfy certain regularity conditions (for instance, the sufficient condition of Govindarajulu, et al. [8] and the assumptions of Lemma 7.2 of Puri [16]), then as $n_S \to \infty$, (3.20) is satisfied, where d is the solution of (3.13) with d in place of c (for a proof see Lemma 2 of [10]). As a consequence of Corollary 2.2.2, one has the following theorem.

THEOREM 4.1. *If either* (a) *or* (b) *of Theorem* 2.2 *holds, if F and $J=H^{-1}$ satisfy the regularity condition of Govindarajulu, et al.* [5] *and the assumptions of Lemma* 7.2 *of Puri* [16] *and if, for fixed P^*, n_T is determined so that* (4.5) *holds when $\tilde{\theta}$ satisfies* (3.11), *then as $n_T \to \infty$,* (3.22) *is satisfied, where d is the solution of* (3.13) *with d replaced by c.*

PROOF. (4.5) may be written as $\mathrm{P}(T_{N,i}(H) \leqq 0,\ i=1, \cdots, k-1)=P^*$. Let θ_{n_S} be a sequence of parameter points satisfying (3.11). Then as shown in [14], the variables $(2A^2/n_S)^{-1/2}[S_{N,i}(H)-\mu_i(\theta_{n_S})+\mu_k(\theta_{n_S})]$ have a limiting joint normal distribution with means, variances and covariances satisfying (3.14). Therefore, by Corollary 2.2.2, the same is true of the variables $(2A^2/n_T)^{-1/2}[T_{N,i}(H)-\mu_i(\theta_{n_T})+\mu_k(\theta_{n_T})]$. Thus (4.5) is equivalent to $\lim_{n_T\to\infty} \mathrm{P}\{W_i \leqq (2A^2/n_T)^{-1/2}[\mu_k(\theta_{n_T})-\mu_i(\theta_{n_T})],\ \text{all } i\}=P^*$. Consequently, one requires that $\lim_{n_T\to\infty}(2A^2/n_T)^{-1/2}[\mu_k(\theta_{n_T})-\mu_i(\theta_{n_T})]=d/2^{1/2}$, which will be true if and only if (3.22) holds.

Remark 4.1.1. For a given d^*, the sample sizes n, n_S and n_T required for the three procedures to have the same minimum probability P^* of correct selection are to be determined from (4.3), (4.4) and (4.5) when $\tilde{\theta}$ satisfies (3.11). Setting $d^*=d^{(n)}=d^{(n_S)}=d^{(n_T)}$, it follows from (3.15), (3.20) and (3.22) that n must satisfy (3.17) and n_S and n_T must satisfy (3.23). Consequently Theorem 3.3 and Example 3.2 hold also for the indifference zone approach to the selection problem. (The expression for $A(S(H), M)$ in (3.24) was obtained by Lehmann [14] for the indifference zone approach.)

Remark 4.1.2. Let the ith population be considered good if $\theta_i \geqq \theta_{\max}-d^*$, where d^* is a given constant, and consider the probability that the selected distribution is good, for parameter points not satisfying (3.5). Lehmann [14] has shown that if the sample sizes are related by (3.24), then $S(H)$ and the means procedure have the same asymptotic probability of selecting a good population. The proof hinges on the fact that for any sequence of parameter points satisfying (3.25), the joint limiting distribution of the random variables $(2A^2/n_S)^{-1/2}S_{N,i}(H)$ is that of a $(k-1)$-dimensional normal vector satisfying (3.26), when n and n_S are related by (3.24). Since in the proof of Theorem 3.4 it was shown that the random variables $(2A^2/n_T)^{-1/2}T_{N,i}(H)$ have the same limiting joint distribution under the identical conditions, it follows that $T(H)$ has the same asymptotic probability of selecting a good population as $S(H)$ and the means procedure, when the sample sizes are related by (3.24).

STANDARD OIL COMPANY (OHIO)
CASE INSTITUTE OF TECHNOLOGY*

REFERENCES

[1] R. E. Bechhofer, "A single-sample multiple decision procedure for ranking means of of normal populations with known variances," *Ann. Math. Statist.*, 25 (1954), 16-39.
[2] C. B. Bell and K. A. Doksum, "Some new distribution-free statistics," *Ann. Math. Statist.*, 36 (1965), 203-214.
[3] C. B. Bell and K. A. Doksum, ""Optimal," one-sample distribution-free tests and their two-sample extensions," *Ann. Math. Statist.*, 37 (1966), 120-132.
[4] H. Chernoff and I. R. Savage, "Asymptotic normality and efficiency of certain nonparametric test statistics," *Ann. Math. Statist.*, 29 (1958), 972-994.
[5] J. Durbin, "Some methods of constructing exact tests," *Biometrika*, 48 (1961), 41-55.
[6] M. Dwass, "On the asymptotic normality of some statistics used in non-parametric tests," *Ann. Math. Statist.*, 26 (1953), 334-339.
[7] A. S. C. Ehrenberg, "Note on normal transformations of ranks," *Brit. Jour. Psych., Stat.*, Section 4 (1951), 133-134.
[8] Z. Govindarajulu, L. LeCam and M. Raghavachari, "Generalizations of theorems of Chernoff and Savage on the asymptotic normality of test statistics," *Proc. Fifth Berkeley Symp. Math. Stat. & Prob.*, Univ. of California Press, Berkeley, (1966), 609-638.
[9] S. S. Gupta, "On a decision rule for a problem in ranking means," Univ. of North Carolina, Institute of Statistics, Mimeo Series No. 150, 1956.
[10] S. S. Gupta, "Probability integrals of multivariate normal and multivariate *t*," *Ann. Math. Statist.*, 34 (1963), 792-828.
[11] J. Hajék, "Asymptotically most powerful rank-order tests," *Ann. Math. Statist.*, 33 (1962), 1124-1147.
[12] J. L. Hodges, Jr. and E. L. Lehmann, "The efficiency of some nonparametric competitors of the *t*-test," *Ann. Math. Statist.*, 27 (1956), 324-335.
[13] J. L. Hodges, Jr. and E. L. Lehmann, "Comparison of the normal scores and Wilcoxon tests," *Proc. Fourth Berkeley Symp. Math. Statist. & Prob.*, Univ. of California Press, Berkeley, 1 (1961), 307-318.
[14] E. L. Lehmann, "A class of selection procedures based on ranks," *Math. Ann.*, 150 (1963), 268-275.
[15] T. K. Matthes and D. R. Truax, "Optimal invariant rank tests for the *k*-sample problem," *Ann. Math. Statist.*, 36 (1965), 1207-1222.
[16] M. L. Puri, "Asymptotic efficiency of a class of *c*-sample tests," *Ann. Math. Statist.*, 35 (1964), 102-121.
[17] B. L. van der Waerden, "Ein neuer Test für das Problem der zwei Stichproben," *Math. Ann.*, 126 (1953), 93-107.

* Now at University of Kentucky.