The Annals of Statistics
1984, Vol. 12, No. 3, 1136–1144

CONSTRAINED SIMULTANEOUS CONFIDENCE INTERVALS FOR MULTIPLE COMPARISONS WITH THE BEST[1]

BY JASON C. HSU

The Ohio State University

For comparing k treatment effects $\theta_1, \theta_2, \cdots \theta_k$, often the parameters of primary interest are $\theta_i - \max_{j \neq i} \theta_j$, $i = 1, \cdots, k$. In this article, we develop constrained $100P^*\%$ two-sided simultaneous confidence intervals for $\theta_i - \max_{j \neq i} \theta_j$ which we refer to as (constrained) MCB intervals. It turns out that the *lower* bounds of the intervals imply Indifference Zone selection inference, and the *upper* bounds of these intervals imply Subset Selection inference, each given at the same confidence level $100P^*\%$ as MCB. We also extend our method to give $100P^*\%$ simultaneous confidence intervals for $\theta_i - \theta^{(i)}_{(k-t)}$, $i = 1, \cdots, k$, where $\theta^{(i)}_{(k-t)}$ is the tth largest among the θ's excluding θ_i.

1. Motivation and background. Let $\pi_1, \pi_2, \cdots, \pi_k$ be k treatments and $\theta_1, \theta_2, \cdots, \theta_k$ their respective treatment effects. Suppose π_i is better than π_j if $\theta_i > \theta_j$. Then simultaneous *two-sided* confidence intervals for $\theta_i - \max_{j \neq i} \theta_j$, $i = 1, \cdots, k$, give the most direct inference for deciding which treatments to use or not to use. For example, if the upper bounds for $\theta_i - \max_{j \neq i} \theta_j$, $i = 1, 2$, are ≤ 0, then π_1 and π_2 are at most second best and can be rejected. On the other hand, if the lower bounds for $\theta_i - \max_{j \neq i} \theta_j$, $i = 3, 4$, are $\geq -\delta^*$ where δ^* is close to zero, then both π_3 and π_4 may be acceptable.

In Section 3, we derive constrained $100P^*\%$ simultaneous *two-sided* confidence intervals for $\theta_i - \max_{j \neq i} \theta_j$, $i = 1, \cdots, k$, which we refer to as (constrained) *multiple comparisons with the best* (MCB) interval. In the parametric case, these intervals have the simple form

$$[-(Y_i - \max_{j \neq i} Y_j - d)^-, (Y_i - \max_{j \neq i} Y_j + d)^+]$$

$i = 1, \cdots, k$. (In this article, $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$.) It turns out that the *upper* bounds of these intervals imply the Subset Selection inference of Gupta (1956, 1965), and the *lower* bounds imply the Indifference Zone Selection inference of Bechhofer (1954), each given at the same confidence level $100P^*\%$ as the MCB intervals.

Note that previous MCB results were presented in terms of the parameters $\theta_i - \max_{1 \leq j \leq k} \theta_j$. Hsu (1981) gave simultaneous lower confidence bounds on $\theta_i - \max_{1 \leq j \leq k} \theta_j$. They were later generalized by Edwards and Hsu (1983) to two-sided simultaneous confidence intervals for $\theta_i - \max_{1 \leq j \leq k} \theta_j$ which reduce to the lower bounds of Hsu (1981) when the upper bounds are set equal to $+\infty$. It turns out that the result of Hsu (1981) can be significantly strengthened. Specifically, the

Received October 1982; revised November 1983.

[1] Research supported in part by Grant 1 R01 GM31287-01, awarded by the National Institute of General Medical Sciences, NIH.

AMS 1980 *subject classification.* Primary 62F07; secondary 62J15.

Key words and phrases. Multiple comparisons, simultaneous confidence intervals, best treatment.

lower bounds of this article and those in Hsu (1981) are the same numerically. They are the same inferentially as well because lower bounds on $\theta_i - \max_{1\le j\le k}\theta_j$ are also lower bounds on $\theta_i - \max_{j\neq i}\theta_j$, and *vice versa*, if lower bounds on the latter are constrained to be nonpositive, which they are in this article. Thus, this article shows that upper bounds on $\theta_i - \max_{j\neq i}\theta_j$ can be added to the lower bounds of Hsu (1981) at no cost (without decreasing the confidence level).

In Section 4, we extend our method to give $100P^*\%$ simultaneous *two-sided* confidence intervals for $\theta_i - \theta^{(i)}_{(k-t)}$, $i = 1, \cdots, k$, where $\theta^{(i)}_{(k-t)}$ is the tth largest among the θ's excluding θ_i.

2. Notations and assumptions. Suppose independent random samples of size n are taken from the treatments $\pi_1, \cdots, \pi_k$. Let

$$\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \cdots, Y_{in})$$

denote the random sample from π_i. Suppose the distributions under the k treatments differ in location only, so that the joint distribution of $(\mathbf{Y}_1, \cdots, \mathbf{Y}_k)$ is $\prod_{i=1}^{k}\prod_{\alpha=1}^{n} F(y_{i\alpha} - \theta_i)$ for some absolutely continuous F. Let (1), (2), $\cdots$, (k) be the unknown *indices* such that

$$\theta_{(1)} \le \theta_{(2)} \le \cdots \le \theta_{(k)}. \tag{2.1}$$

In case of ties in the θ-values, the indices can be chosen in any manner so long as the order relation (2.1) is satisfied. Note that $\theta_{(i)}$ is the ith smallest θ-value. For each i, let

$$\theta^{(i)}_{(1)} \le \theta^{(i)}_{(2)} \le \cdots \le \theta^{(i)}_{(k-1)}$$

denote the ordered $\{\theta_j, j \neq i\}$, so for example $\theta_i - \max_{j\neq i}\theta_j = \theta_i - \theta^{(i)}_{(k-1)}$.

REMARK. Consistent throughout the article, ( ) in the *subscript* denotes ordering according to the *parameter* values under consideration. (i) in the *superscript* means the treatment π_i is being excluded from consideration. [ ] in the subscript signifies ordering according to the *observed* values under consideration.

Let T: $\mathbb{R}^{2n} \to \mathbb{R}$ be a statistic such that

$$T(\mathbf{Y}_i + \delta_1\mathbf{1}, \mathbf{Y}_j + \delta_2\mathbf{1}) = T(\mathbf{Y}_i, \mathbf{Y}_j) + (\delta_1 - \delta_2).$$

In the sequel, we use the abbreviation $T_{ij} \equiv T(\mathbf{Y}_i, \mathbf{Y}_j)$.

Below we describe two classes of T which give rise to parametric and nonparametric simultaneous confidence intervals respectively.

Parametric statistics. Let $T_{ij} \equiv T(\mathbf{Y}_i, \mathbf{Y}_j) = U(\mathbf{Y}_i) - U(\mathbf{Y}_j)$ where U: $\mathbb{R}^n \to \mathbb{R}$ is a translation equivariant estimator (e.g. sample mean, median, etc.). We use the notation $Y_i \equiv U(\mathbf{Y}_i)$ so $T_{ij} = Y_i - Y_j$.

Nonparametric statistics. Here T_{ij} will be a nonparametric bound on $\theta_i - \theta_j$.

Let $R_{ij\alpha}(\delta)$ denote the rank of $Y_{i\alpha} - \delta$ in the combined sample

$$Y_{i1} - \delta, \cdots, Y_{in} - \delta, Y_{j1}, \cdots, Y_{jn}.$$

For $i \neq j$, let

$$R_{ij}(\delta) = \sum_{\alpha=1}^{n} a_{2n}(R_{ij\alpha}(\delta)) - n\bar{a}_{2n},$$

where $a_n(\cdot)$ is some nondecreasing score function converging in quadratic mean to a nonconstant square integrable function $\phi(\cdot)$:

$$\lim_{n\to\infty} \int_0^1 \{a_n(1 + [un]) - \phi(u)\}^2 \, du = 0.$$

Here

$$\bar{a}_n = n^{-1} \sum_{\alpha=1}^{n} a_n(\alpha) \to \int_0^1 \phi(u) \, du$$

and $[un]$ denotes the largest integer contained in un. With the constant r to be determined according to the desired confidence level P^*, define

$$T_{ij} = \sup\{\delta: R_{ij}(\delta) \geq - r\}.$$

Two cases of $a_n(\cdot)$ are of particular interest.

Wilcoxon statistics. Suppose $a_{2n}(x) = x/(2n + 1)$. If we let $W_{ij[1]} \leq \cdots \leq W_{ij[n^2]}$ denote the n^2 ordered differences $Y_{i\alpha} - Y_{j\beta}$, $1 \leq \alpha, \beta \leq n$, then $T_{ij} = W_{ij[w_0+w]}$ where $w_0 = (n^2 + 1)/2$ and $w = (2n + 1)r + 1/2$.

Median statistics. Suppose $a_{2n}(x) = 1, 0$, or -1 as $x >, =$, or $< (2n + 1)/2$. Let $Y_{i[1]} \leq \cdots \leq Y_{i[n]}$, $Y_{j[1]} \leq \cdots \leq Y_{j[n]}$ denote the ordered observations from π_i and π_j. Then $T_{ij} = Y_{i[m_0+m]} - Y_{j[m_0-m]}$ where $m_0 = (n + 1)/2$ and $m = (r + 1)/2$.

3. Multiple comparisons with the best (MCB). In this section, we derive simultaneous *two-sided* confidence intervals for $\theta_i - \max_{j\neq i}\theta_j$, $i = 1, \cdots, k$, by pivoting the event

$$E = \{(T_{(k)(i)} - (\theta_{(k)} - \theta_{(i)}), i = 1, \cdots, k - 1) \in A\}.$$

Here $A \subseteq \mathbb{R}^{k-1}$ is chosen to be *monotone, permutationally invariant,* (defined below) and making $P[E] \geq P^* \geq 1/k$.

DEFINITION 3.1. $A \subseteq \mathbb{R}^{\ell}$ is *monotone* if $(x_1, \cdots, x_\ell) \in A$ and $x_i \leq x_i'$ for all i imply $(x_1', \cdots, x_\ell') \in A$.

DEFINITION 3.2. $A \subseteq \mathbb{R}^{\ell}$ is *permutationally invariant* if $(x_1, \cdots, x_\ell) \in A$ and $(\tau_1, \cdots, \tau_\ell)$ is a permutation of $(1, \cdots, \ell)$ imply $(x_{\tau_1}, \cdots, x_{\tau_\ell}) \in A$.

Define, for $i = 1, \cdots, k$,

$$D_i^{**} = (\sup\{\delta: (T_{ij} - \delta, j \neq i) \in A\})^+.$$

We will show that, on the event E, $\theta_i - \max_{j\neq i}\theta_j \leq D_i^{**}$ for all i.

REMARK. D_i^{**} is the nonnegative part of the maximum amount by which the observations from π_i can be reduced (while keeping all other observations fixed) for π_i to still remain in the (generalized) Gupta's (1956, 1965) subset containing the best treatment.

For $i \neq j$, define

$$D_{ji} = \sup\{\delta: \delta > 0, (T_{ji} - \delta, T_{j\ell}, \ell \neq i \text{ or } j) \in A\},$$

$$D_i^* = -max_{j\neq i} D_{ji}, \ i = 1, \cdots, k,$$

where $\sup(\emptyset) = 0$. We will show that, on the event E, $\theta_i - \max_{j\neq i}\theta_j \geq D_i^*$ for all i.

THEOREM 3.1. *For all* $\boldsymbol{\theta}$,

$$P[D_i^* \leq \theta_i - \max_{j\neq i}\theta_j \leq D_i^{**} \text{ for all } i] \geq P^*.$$

PROOF.

$$\begin{aligned} E &= \{(T_{(k)(i)} - (\theta_{(k)} - \theta_{(i)}), i = 1, \cdots, k-1) \in A\} \\ &\subseteq \{(T_{(k)(i)} - (\theta_{(k)} - \theta_{(k-1)}), i = 1, \cdots, k-1) \in A\} \\ &= \{(T_{ji} - (\theta_{(k)} - \theta_{(k-1)}), i \neq j) \in A \text{ for } j = (k) \\ &\qquad \text{and } \theta_j - \theta_{(k-1)} \leq 0 \text{ for } j \neq (k)\} \\ &\subseteq \{\theta_j - \theta_{(k-1)} \leq D_j^{**} \text{ for all } j\} \\ &= \{\theta_i - \max_{j\neq i}\theta_j \leq D_i^{**} \text{ for all } i\} \equiv E_1 \quad \text{(say)}. \end{aligned}$$

Note we have *equality* for the last line because $D_i^{**} \geq 0$.

$$\begin{aligned} E &= \{(T_{(k)(i)} - (\theta_{(k)} - \theta_{(i)}), i = 1, \cdots, k-1) \in A\} \\ &\subseteq \{\theta_{(k)} - \theta_{(i)} \leq D_{(k)(i)} \text{ for } i = 1, \cdots, k-1\} \\ &= \{-D_{(k)i} \leq \theta_i - \theta_{(k)} \text{ for all } i \neq (k) \text{ and} \\ &\qquad 0 = \theta_i - \theta_{(k)} \text{ for } i = (k)\} \\ &\subseteq \{-\max_{j\neq i} D_{ji} \leq \theta_i - \theta_{(k)} \text{ for all } i\} \\ &= \{D_i^* \leq \theta_i - \max_{j\neq i}\theta_j \text{ for all } i\} = E_2 \quad \text{(say)}. \end{aligned}$$

Note we have *equality* for the last line since $D_i^* \leq 0$. Thus $E \subseteq (E_1 \cap E_2)$. Noting $P[E] \geq P^*$ completes the proof.

Parametric confidence intervals. In the parametric case $T_{ij} = Y_i - Y_j$, and we can let

$$E = \{Y_{(k)} - Y_{(i)} - (\theta_{(k)} - \theta_{(i)}) > -d \text{ for } i = 1, \cdots, k-1\}$$

so $A = (-d, \infty) \times \cdots \times (-d, \infty) \subseteq \mathbb{R}^{k-1}$ is monotone and permutationally invariant. It can be verified that

$$D_i^* = -(Y_i - \max_{j\neq i} Y_j - d)^-, \quad D_i^{**} = (Y_i - \max_{j\neq i} Y_j + d)^+$$

so the confidence statement can be written as

$$(\theta_i - \max_{j\neq i}\theta_j) \in \pm\,(Y_i - \max_{j\neq i} Y_j \pm d)^{\pm} \quad \text{for} \quad i = 1, \cdots, k.$$

Note the striking simplicity and symmetry.

If the common distribution G of $Y_i - \theta_i$ is completely known, then to satisfy $P[E] = P^*$, we take d to be the solution of

$$\int_{-\infty}^{+\infty} [G(z+d)]^{k-1}\, dG(z) = P^*.$$

For the usual normal distribution model, however, we have $Y_i = \bar{Y}_i. = \sum_{a=1}^n Y_{ia}/n$, $G(y_i - \theta_i) = \Phi((y_i - \theta_i)/(n^{-1/2}\sigma))$ where Φ is the standard normal distribution and the scale factor σ is *unknown*. To satisfy $P[E] = P^*$, we take $d = d's/n^{1/2}$ where s is the usual pooled estimator of σ, and d' is the solution of

$$\int_0^{\infty}\int_{-\infty}^{+\infty} [\Phi(z + d's)]^{k-1}\, d\Phi(z)\, dQ(s) = P^*.$$

Here Q is the distribution of s/σ. The point is, d is always exactly the number which enters the Subset Selection procedure of Gupta (1956, 1965) and the Indifference Zone Selection procedure of Bechhofer (1954), each at the same confidence level as MCB.

Nonparametric confidence intervals. In the nonparametric case we can let

$$E = \{T_{(k)(i)} - (\theta_{(k)} - \theta_{(i)}) > 0 \text{ for } i = 1, \cdots, k-1\}$$

so $A = (0, \infty) \times \cdots \times (0, \infty) \subseteq \mathbb{R}^{k-1}$ is monotone and permutationally invariant. Now

$$E = \{R_{(k)(i)}(\theta_{(k)} - \theta_{(i)}) \ge -r \text{ for } i = 1, \cdots, k-1\} \quad \text{a.s.},$$

so to satisfy $P[E] \ge P^*$, we set r to be the smallest nonnegative number such that

$$P_0[R_{ki}(0) \ge -r \text{ for } i = 1, \cdots, k-1] \ge P^*.$$

Here P_0 indicates that the probability is computed under $\theta_1 = \cdots = \theta_k$.

In the Wilcoxon case

$$D_i^* = \min_{j\neq i}(W_{ij[w_0-w]} I_{\{D_j^{**}>0\}}), \quad D_i^{**} = (\min_{j\neq i} W_{ij[w_0+w]})^+$$

where $w_0 = (n^2+1)/2$ and $w = (2n+1)r + ½$.

In the median statistics case

$$D_i^* = -(Y_{i[m_0-m]} - \max_{j\neq i} Y_{j[m_0+m]})^-, \quad D_i^{**} = (Y_{i[m_0+m]} - \max_{j\neq i} Y_{j[m_0-m]})^+$$

where $m_0 = (n+1)/2$ and $m = (r+1)/2$.

Asymptotic relative efficiencies. If we measure the relative efficiencies of confidence intervals by their abilities to exclude false parameter values, then the same results as for the lower bounds of Hsu (1981) can be obtained by using the arguments therein.

3.1 *Implications for ranking and selection.*

Subset Selection. The parametric Subset Selection procedure of Gupta (1956, 1965) selects π_i if $(Y_i - \max_{j\neq i} Y_j + d) > 0$ where d is the same as in our confidence intervals. Thus the traditional Subset Selection confidence statement is

$$P[\pi_{(k)} \in \{\pi_i: D_i^{**} > 0\}] \geq P^*. \tag{3.1}$$

The confidence statement given by the *upper* bounds of our intervals is

$$P[\theta_i - \max_{j\neq i}\theta_j \leq D_i^{**} \text{ for all } i] \geq P^* \tag{3.2}$$

which, since $D_i^{**} \geq 0$, implies

$$P[\theta_i - \theta_{(k-1)} \leq 0 \text{ for all } i \text{ such that } D_i^{**} \leq 0] \geq P^*.$$

Note now the Subset Selection inference

$$[\pi_{(k)} \in \{\pi_i: D_i^{**} > 0\}] = A \quad \text{(say)}$$

is essentially identical to the inference

$$[\theta_i - \theta_{(k-1)} \leq 0 \text{ for all } i \text{ such that } D_i^{**} \leq 0] = B \quad \text{(say)}.$$

A is B if $\theta_{(k-1)} < \theta_{(k)}$. If $\theta_{(k-1)} = \theta_{(k)}$, then A is B with the additional restriction that B does not include the inference $\theta_i - \theta_{(k-1)} \leq 0$ where π_i is the treatment tied for the best that has been *arbitrarily* designated as the best treatment $\pi_{(k)}$. This additional restriction is not useful, we feel, since in fact $\theta_i - \theta_{(k-1)} = 0$ in this case. Actually (3.2) is easier to understand than (3.1) and avoids completely the technicality inherent in (3.1) of having the validity of the inference dependent on an *arbitrary* assignment of $\pi_{(k)}$ in case of ties.

Indifference Zone Selection. Since the lower confidence bounds D_i^* for $\theta_i - \max_{j\neq i}\theta_j$ are constrained to be nonpositive, they are also bounds for $\theta_i - \theta_{(k)}$. In that form they were given in Hsu (1981), and shown to imply the *union* of the Indifference Zone Selection inferences of Bechhofer (1954), Fabian (1962), and Desu (1970), each given at the same confidence level as MCB.

4. Multiple comparisons with the *t*th best (MC*t*B). In this section we indicate how the method can be extended to give simultaneous *two-sided* confidence intervals for $(\theta_i - \theta_{(k-t)}^{(i)})$, $i = 1, \cdots, k$, the difference between each treatment and the tth best ($t < k - 1$) among the rest of the treatments. This is achieved by pivoting the event

$$E = \cap_{j=k-t+1}^{k} \{(T_{(j)(i)} - (\theta_{(j)} - \theta_{(i)}),\ i = 1, \cdots, k - t) \in A\}.$$

Here we choose $A \subseteq \mathbb{R}^{k-t}$ to be *monotone, Schur concave*, and making $P[E] \geq P^* \geq \binom{k}{t}^{-1}$.

DEFINITION 4.1. $A \subseteq \mathbb{R}^{\ell}$ is *Schur concave* if $(x_1, \cdots, x_\ell) \in A$ and $(x_1, \cdots, x_\ell)$ majorizes $(x_1', \cdots, x_\ell')$ imply $(x_1', \cdots, x_\ell') \in A$. (See for example Tong, 1980, for a definition of majorization.)

The reason for the Schur concavity requirement is that, when used with the following lemma due to Hollander, Proschan, Sethuraman (1977), it allows us to replace the unobservable $T_{(j)(i)}$ in E by appropriate observable ordered T_{ji}.

LEMMA 4.1. *Suppose $\lambda_1 \leq \cdots \leq \lambda_\ell$ are fixed. Let $x_{[1]} \leq x_{[2]} \leq \cdots \leq x_{[\ell]}$ denote the ordered $x_1, \cdots, x_\ell$. If $A \subseteq \mathbb{R}^{\ell}$ is Schur concave, then $\{(x_1 - \lambda_1, \cdots, x_\ell - \lambda_\ell) \in A\} \subseteq \{(x_{[1]} - \lambda_1, \cdots, x_{[\ell]} - \lambda_\ell) \in A\}$.*

For each i, let

$$T_{i[1]} \leq T_{i[2]} \leq \cdots \leq T_{i[k-1]}$$

denote the ordered $\{T_{ij}, j \neq i\}$. Define, for $i = 1, \cdots, k$,

$$D_i^{**} = (\sup\{\delta: (T_{i[t]} - \delta, \cdots, T_{i[k-1]} - \delta) \in A\})^+.$$

It can be shown that, on the event E, $\theta_i - \theta_{(k-t)}^{(i)} \leq D_i^{**}$ for all i. The proof consists of noting

$$E = \cap_{j=k-t+1}^{k} \{(T_{(j)(i)} - (\theta_{(j)} - \theta_{(i)}), i = 1, \cdots, k - t) \in A\}$$

$$\subseteq \cap_{j=k-t+1}^{k} \{(T_{(j)[k-i]} - (\theta_{(j)} - \theta_{(i)}), i = 1, \cdots, k - t) \in A\}$$

by monotonicity and the lemma and then proceeding much the same as in Theorem 3.1.

For $i \neq j$, let

$$T_{j[1]}^{(i)} \leq \cdots \leq T_{j[k-2]}^{(i)}$$

be the ordered $\{T_{j\ell}, \ell \neq i \text{ or } j\}$. Now define

$$D_{ji} = \sup\{\delta: \delta > 0, (T_{ji} - \delta, T_{j[t]}^{(i)}, \cdots, T_{j[k-2]}^{(i)}) \in A\}.$$

Here $\sup(\varnothing) = 0$. For each i, let

$$D_{[1]i} \leq \cdots \leq D_{[k-1]i}$$

denote the ordered $\{D_{ji}, j \neq i\}$. Define

$$D_i^* = -D_{[k-t]i}, \quad i = 1, \cdots, k.$$

It can be shown that, on the event E, $\theta_i - \theta_{(k-t)}^{(i)} \geq D_i^*$ for all i. The proof proceeds much the same as in Theorem 3.1 with the additional intermediate step that

$$\{\theta_{(j)} - \theta_{(i)} \leq D_{(j)(i)}, j = k - t + 1, \cdots, k, i = 1, \cdots, k - t\}$$

$$\subseteq \{\theta_{(j)} - \theta_{(i)} \leq D_{[j-1](i)}, j = k - t + 1, \cdots, k, i = 1, \cdots, k - t\}$$

by the monotonicity and the lemma.

Since $P[E] \geq P^*$, combining the two sets of inequalities above gives

(4.1) $$P[D_i^* \leq \theta_i - \theta_{(k-t)}^{(i)} \leq D_i^{**} \text{ for } i = 1, \cdots, k] \geq P^*$$

for all $\boldsymbol{\theta}$.

For illustration, we give the exact expressions for D_i^* and D_i^{**} in the parametric case, the expressions in the nonparametric case being analogous.

In the parametric case $T_{ij} = Y_i - Y_j$ and we can let

$$E = \cap_{j=k-t+1}^{k} \{Y_{(j)} - Y_{(i)} - (\theta_{(j)} - \theta_{(i)}) > -d \text{ for } i = 1, \cdots, k-t\}$$

so $A = (-d, \infty) \times \cdots \times (-d, \infty) \subseteq \mathbb{R}^{k-t}$ is monotone and Schur concave. If, for each i,

$$Y_{[1]}^{(i)} \leq \cdots \leq Y_{[k-1]}^{(i)}$$

denote the ordered $\{Y_j, j \neq i\}$, then it can be shown that

$$D_i^* = -(Y_i - Y_{[k-t]}^{(i)} - d)^-, \qquad D_i^{**} = (Y_i - Y_{[k-t]}^{(i)} + d)^+.$$

Implications for ranking and selection. For Indifference Zone selection of the best t treatments without regard to order, the confidence statement of Bechhofer (1954) can be written as

(4.2) $$P_\theta[0 \leq \theta_i - \theta_{(k-t+1)} \text{ for all } i \text{ with } -\delta^* < D_i^*] \geq P^* \quad \text{if } -\delta^* \geq \theta_{(k-t)} - \theta_{(k-t+1)}.$$

A stronger confidence statement by Chiu (1974) is implied by

(4.3) $$P_\theta[-\delta^* < \theta_i - \theta_{(k-t+1)} \text{ for all } i \text{ with } -\delta^* < D_i^* \text{ and } \theta_i - \theta_{(k-t)} < \delta^* \text{ for all } i \text{ with } D_i^{**} < \delta^*] \geq P^*.$$

For selecting a subset of the treatments so that all those selected are good, the confidence statement of Carroll, Gupta, and Huang (1975) can be written as

(4.4) $$P_\theta[-\delta^* < \theta_i - \theta_{(k-t+1)} \text{ for all } i \text{ with } -\delta^* < D_i^*] \geq P^*.$$

For selecting a subset of the treatments to contain the best t treatments, the confidence statement of Carroll, Gupta, and Huang (1975) was essentially

(4.5) $$P_\theta[\theta_i - \theta_{(k-t)} \leq 0 \text{ for all } i \text{ with } D_i^{**} \leq 0] \geq P^*.$$

Clearly, the inference given by (4.1) contains the *union* of the inference given by (4.2)–(4.5), each given at the same confidence level $100P^*\%$ as (4.1).

5. Concluding note. A computer program for MCB in the normal distribution case has been written and is available for distribution from the author.

REFERENCES

BECHHOFER, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Ann. Math. Statist.* **25** 16–39.

CARROLL, R. J., GUPTA, S. S. and HUANG, D. Y. (1975). Selection procedures for the t best populations. *Comm. Statist.* **4** 987–1008.

CHIU, W. K. (1974). The ranking of means of normal populations for a generalized selection goal. *Biometrika* **61** 579–584.

DESU, M. M. (1970). A selection problem. *Ann. Math. Statist.* **41** 1596–1603.

EDWARDS, D. G. and HSU, J. C. (1983). Multiple comparisons with the best treatment. *J. Amer. Statist. Assoc.* **78** 965–971. Correction, (1984). *J. Amer. Statist. Assoc.* **79,** to appear.

FABIAN, V. (1962). On multiple decision methods for ranking population means. *Ann. Math. Statist.* **33** 248–254.

GUPTA, S. S. (1956). On a decision rule for a problem in ranking means. *Inst. Statist. Mimeo Series* 150, Univ. of North Carolina, Chapel Hill.

GUPTA, S. S. (1965). On some multiple decision (selection and ranking) rules. *Technometrics* **7** 225–245.

HOLLANDER, M., PROSCHAN, F. and SETHURAMAN, J. (1977). Functions decreasing in transposition and their applications in ranking problems. *Ann. Statist.* **5** 722–733.

HSU, J. C. (1981). Simultaneous confidence intervals for all distances from the "best". *Ann. Statist.* **9** 1026–1034.

TONG, Y. L. (1980). *Probability Inequalities in Multivariate Distributions.* Academic, New York.

DEPARTMENT OF STATISTICS
THE OHIO STATE UNIVERSITY
1958 NEIL AVE.
COLUMBUS, OHIO 43210