A NEW TECHNIQUE FOR IMPROVED CONFIDENCE BOUNDS FOR THE PROBABILITY OF CORRECT SELECTION*

by

Shanti S. Gupta, Yuning Liao, Chunfu Qiu and Jin Wang
Purdue University

Technical Report #91-62

Department of Statistics
Purdue University

October 1991
Revised May 1992
Revised February 1994

*This research was supported in part by NSF Grant DMS-8923071, Chunfu Qiu was supported by David Ross Fellowship and Jin Wang was supported in part by NSF Grant DMS-8717799 at Purdue University.

A New Technique for Improved Confidence Bounds for the Probability of Correct Selection

Shanti S. Gupta, Yuning Liao, Chunfu Qiu, and Jin Wang

Purdue University, New Drug Services Inc., University of Illinois and Purdue University

Abstract

This paper deals with the problem of estimating the probability of a correct selection (PCS) in location parameter models. Practical lower confidence bounds for the PCS in location parameter models are presented with a user's choice of dimension q $(1 \leq q \leq k-1)$ for computation, where k is the number of populations. It is shown that the larger the q, the better the lower bound, but the more complicated the computation. The result when $q = 1$ coincides with Kim's (1986) result. A numerical example is presented to show that our lower bound with $q = 2$ improves Kim's result considerably. With an appropriate modification, our result can be applied to location-scale parameter models with the scale parameter unknown.

Key words and phrases: Ranking and selection, probability of a correct selection, confidence region, location parameter.

1 Introduction

In many practical situations, the goal of the experimenter is to compare two or more populations in order to make a decision in the form of ranking the populations. The best studied ranking goal concerns the best population (the most efficient treatment for an ailment, the worst car in fuel efficiency and so on). The classical tests of homogeneity were not designed to provide answers to such questions. Rejecting the null hypothesis is not the final solution to the experimenter's problem; so the methods of ranking and selection come into play. In this paper, we focus on these problems in the case of location parameter models.

Let X_{ij} $(1 \leq i \leq k; 1 \leq j \leq n)$ be independent observations from each of k populations with cdf's $F(x - \theta_i)$. Let $Y_i = Y(X_{i1}, \cdots, X_{in})$ be an appropriate statistic with cdf $G(y - \theta_i)$ and pdf

$g(y - \theta_i)$. We denote the ordered values of θ_i's and y_i's by $\theta_{[1]} \leq \cdots \leq \theta_{[k]}$ and $y_{(1)} \leq \cdots \leq y_{(k)}$, respectively. Let $Y_{[i]}$ and $\pi_{(i)}$ denote the statistic and the population associated with the ordered but unknown parameter $\theta_{[i]}$, respectively. It is assumed that there is no a priori knowledge of the corresponding pairing of the $Y_{(i)}$ and $\theta_{[j]}$ $(1 \leq j \leq k)$.

The experimenter wishes to select the population associated with the largest unknown parameter $\theta_{[k]}$, so the natural selection rule "select the population corresponding to the largest Y_i value as the best" is used. Then the probability of a correct selection (i.e. the selected population is indeed the best) (PCS) is

$$PCS = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} G(y + \theta_{[k]} - \theta_{[i]}) dG(y). \tag{1}$$

Now, the PCS depends on the parameter $\underset{\sim}{\theta} = (\theta_1, \cdots, \theta_k)$ only via the differences $\theta_{[k]} - \theta_{[i]}$, $1 \leq i \leq k-1$. So, if we can find some reasonable lower bounds on the differences $\theta_{[k]} - \theta_{[i]}$, based on the sample, then we can provide a reasonable lower bound for the PCS.

Bechhofer (1954) applied an indifference zone for the difference $\theta_{[k]} - \theta_{[k-1]}$ to get a lower bound for the PCS, which requires the experimenter first to specify a positive constant δ^* so that $\theta_{[k]} - \theta_{[k-1]} \geq \delta^*$. Considering this as a retrospective analysis problem, Olkin, Sobel and Tong (1976, 1982) and Gibbons, Olkin and Sobel (1977) have presented estimators of the PCS. Faltin and McCulloch (1983) have studied the small-sample properties of the Olkin-Sobel-Tong estimator of the PCS for the case when $k = 2$. Bofinger (1985) has discussed the non-existence of a consistent estimator of the PCS. Anderson, Bishop and Dudewicz (1977) have given a lower confidence bound on the PCS in the case of normal populations having a common variance which is either known or unknown. Kim (1986) and Gupta, Leu and Liang (1990) found a $100(1-\alpha)\%$ lower confidence bound for the difference between the best and the second best parameters $\theta_{[k]} - \theta_{[k-1]}$ for two different types of distributions, respectively, and then derived a conservative lower confidence bound for the PCS for the corresponding distributions. In addition, Gupta and Liang (1991) obtained a lower confidence bound for the PCS by deriving simultaneous lower confidence bounds on the $\theta_{[k]} - \theta_{[i]}$, $i \neq k$, where a range statistic was used.

In this paper, using a new approach, we derive a confidence region for the differences $\theta_{[k-i+1]} - \theta_{[k-i]}$, $i = 1, ..., k-1$, and then obtain a lower confidence bound for the PCS which is sharper than that of Kim(1986). The paper is organized as follows: In Section 2, we formulate the problem and state the main result. Section 3 describes a numerically feasible approximation to the main result and compares it with that in Kim (1986). Section 4 is concerned with the case of an unknown scale parameter and illustrates two practical problems. Section 5 gives some final remarks. The Appendix provides a proof for the key Lemma 2.1.

2 Main Result

For convenience of computation, let $\delta_i = \theta_{[k-i+1]} - \theta_{[k-i]}$ and $Z_i = Y_{(k-i+1)} - Y_{(k-i)}$, $(1 \le i \le k-1)$. Then for an observed sample of $X_{ij} = x_{ij}$ $(1 \le i \le k; 1 \le j \le n)$, we have a corresponding sample value $Z_i = z_i$ $(1 \le i \le k-1)$. We first consider the distribution of Z_1. As one would expect, we have

Lemma 2.1 *(a) The distribution of Z_1 depends on $\underset{\sim}{\theta}$ only via $\underset{\sim}{\delta} = (\delta_1, \cdots, \delta_{k-1})^T$;*

(b) For any value of $\underset{\sim}{\delta} \ge 0$ and any constant $c > 0$,

$$P_{\underset{\sim}{\delta}}(Z_1 > c) = \sum_{i=1}^{k} \int_{-\infty}^{\infty} \prod_{j=1}^{i-1} G(y - c + \sum_{l=j}^{i-1} \delta_{k-l}) \prod_{j=i+1}^{k} G(y - c - \sum_{l=i}^{j-1} \delta_{k-l}) dG(y); \tag{2}$$

(c) If $g(y)$ is (strictly) log-concave, then the cdf of Z_1 is (strictly) decreasing in $\underset{\sim}{\delta}$ in the sense that $P_{\underset{\sim}{\delta}}(Z_1 \le c) \le P_{\underset{\sim}{\delta}^*}(Z_1 \le c)$ for any fixed c whenever $\underset{\sim}{\delta} \ge \underset{\sim}{\delta}^*$, where $\underset{\sim}{\delta} = (\delta_1, \cdots, \delta_{k-1})^T \ge \underset{\sim}{\delta}^* = (\delta_1^*, \cdots, \delta_{k-1}^*)^T$ if and only if $\delta_i \ge \delta_i^*$ for every i.

Proof: Proof is given in the Appendix. □

Let $K(\underset{\sim}{\delta})$ be the $(1-\alpha)$ quantile of Z_1 given $\underset{\sim}{\delta}$. That is, $K(\underset{\sim}{\delta}) = c$ if and only if $P_{\underset{\sim}{\delta}}\{Z_1 \le c\} = 1 - \alpha$. Then $K(\cdot)$ is a well-defined mapping from $(R^+)^{k-1}$ to R^+. Also, for the function $K(\underset{\sim}{\delta})$, we have

$$P_{\underset{\sim}{\delta}}(K(\underset{\sim}{\delta}) \ge Z_1) = 1 - \alpha, \tag{3}$$

for all $\underset{\sim}{\delta}(\ge 0)$. Let $C(z_1) = \{\underset{\sim}{\delta} : K(\underset{\sim}{\delta}) \ge z_1\}$, then

Theorem 2.1 *$C(z_1)$ is a $100(1-\alpha)\%$ confidence region for the parameter $\underset{\sim}{\delta}$, based on the sample $\underset{\sim}{Z} = \underset{\sim}{z}$. Consequently, a $100(1-\alpha)\%$ lower confidence bound for the PCS is*

$$\widehat{PCS} = \inf_{\underset{\sim}{\delta} \in C(z_1)} \left\{ \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} G(y + \sum_{j=1}^{i} \delta_j) dG(y) \right\}. \tag{4}$$

That is, $P_{\underset{\sim}{\delta}}\{PCS \ge \widehat{PCS}\} \ge 1 - \alpha$ for all $\underset{\sim}{\delta}$ (≥ 0).

Proof: It is a straightforward result from (3). □

In the following we study the case when the density $g(y-\theta_i)$ of the distribution is log-concave. We need the following lemma which is an immediate consequence of Lemma 2.1(c).

Lemma 2.2 : *If $g(y)$ is (strictly) log-concave, then $K(\underset{\sim}{\delta})$ is (strictly) increasing in $\underset{\sim}{\delta}$.* □

Notice that, under the log-concave condition, $C(z_1)$ is not empty for any sample value z_1 and consists of all the values of $\underset{\sim}{\delta}(\geq 0)$ if $z_1 \leq x_{\alpha,k}$, where $x_{\alpha,k}$ is the solution of the following equation

$$k\int_{-\infty}^{\infty} G^{k-1}(y-c)dG(y) = \alpha. \tag{5}$$

The left hand side of (5) is obtained from (2) by setting $\underset{\sim}{\delta} = \underset{\sim}{0}$. Hence $\widehat{PCS} = 1/k$ if $z_1 \leq x_{\alpha,k}$, which is a trivial lower bound for the PCS. See also Kim (1986) for the explanation of a similar situation.

For $z_1 > x_{\alpha,k}$, let $C^*(z_1) = \{\underset{\sim}{\delta} : K(\underset{\sim}{\delta}) = z_1\}$, then

$$C^*(z_1) = \{\underset{\sim}{\delta} : P_{\underset{\sim}{\delta}}(Z_1 \leq z_1) = 1-\alpha\}, \tag{6}$$

in which $\underset{\sim}{\delta}$ satisfies the equation

$$\sum_{i=1}^{k}\int_{-\infty}^{\infty}\prod_{j=1}^{i-1} G(y-z_1+\sum_{l=j}^{i-1}\delta_{k-l})\prod_{j=i+1}^{k} G(y-z_1-\sum_{l=i}^{j-1}\delta_{k-l})dG(y) = \alpha. \tag{7}$$

Since both $K(\underset{\sim}{\delta})$ and the integrand in (4) are non-decreasing in $\underset{\sim}{\delta}$, it follows that the infimum of (4) is achieved on $C^*(z_1)$. Hence we have

Corollary 2.1 *If $z_1 > x_{\alpha,k}$, then*

$$\widehat{PCS} = \inf_{\underset{\sim}{\delta}\in C^*(z_1)}\left\{\int_{-\infty}^{\infty}\prod_{i=1}^{k-1} G(y+\sum_{j=1}^{i}\delta_j)dG(y)\right\}, \tag{8}$$

where $C^*(z_1)$ is determined by (7). Otherwise, $\widehat{PCS} = 1/k$. □

So, to evaluate $\widehat{PCS}$ for a given sample value $z_1 > x_{\alpha,k}$, first we figure out $C^*(z_1)$ numerically at some grid points, then find out the infimum of (8) on those points.

We would like to point out the difference between the method here and the method in Gupta-Liang (1991). First, we briefly describe the Gupta-Liang method. The notations are the same as defined in Section 1. We know that $Y_i - \theta_i \sim G(y)$, independent of θ_i, $i = 1, \cdots, k$. For a given α, $0 < \alpha < 1$, let $c(k, \alpha)$ be the value such that

$$P\{\max_{1 \le i \le k}(Y_i - \theta_i) - \min_{1 \le i \le k}(Y_i - \theta_i) \le c(k, \alpha)\} = 1 - \alpha.$$

Define

$$\hat{\delta}_{L,i} = (Y_{(k)} - Y_{(i)} - c(k, \alpha))^+, \ i = 1, \cdots, k-1;$$

and

$$\hat{P}_L = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} G(y + \hat{\delta}_{L,i}) dG(y);$$

where $(x)^+ = \max(0, x)$, and $\hat{P}_L$ is obtained by replacing $\theta_{[k]} - \theta_{[i]}$ by $\hat{\delta}_{L,i}$ in (1). Gupta and Liang (1991) showed that

$$P_{\underset{\sim}{\theta}}\{\hat{\delta}_{L,i} \le \theta_{[k]} - \theta_{[i]}, \ \ i = 1, \cdots, k-1\} \ge 1 - \alpha.$$

Then $\hat{P}_L$ is a lower confidence bound for the PCS.

The difference between our method and the Gupta-Liang method lies in using data differently to construct lower confidence bounds for $\theta_{[k]} - \theta_{[i]}$, $i = 1, \cdots, k-1$. We employ $Y_{(k)} - Y_{(k-1)}$ whereas Gupta and Liang use all data $Y_{(k)} - Y_{(i)}$, $i = 1, \ldots, k-1$. However, one method does not dominate the other completely in the sense of providing larger PCS values (see Examples 4.2 and 4.3 in applications). Usually our method does better when $Y_{(k)} - Y_{(k-1)}$ is not too small. The Gupta-Liang method may perform better when $Y_{(k)} - Y_{(k-1)}$ is small and, at the same time, $Y_{(k)} - Y_{(k-i)}$, $i = 1, \cdots, k-2$, are large. Generally speaking, for moderate k, say $k \ge 5$, the Gupta-Liang method tends to underestimate the PCS since $c(k, \alpha)$ is quite large, and the estimate for $\theta_{[k]} - \theta_{[k-1]}$ would be close to 0, and then $\hat{P}_L$ would be only around 0.5. The method in this paper obtains the best lower confidence bound for $\theta_{[k]} - \theta_{[k-1]}$ based only on data $Y_{(k)} - Y_{(k-1)}$. It is better not to underestimate $\theta_{[k]} - \theta_{[k-1]}$ since $\theta_{[k]} - \theta_{[i]} \ge \theta_{[k]} - \theta_{[k-1]}$, for all i.

3 Some Practical Lower Bounds

There is not much difficulty in evaluating $\widehat{PCS}$ for small k. However, this computation becomes very difficult for moderate or large k (≥ 5). In the following, we propose a method to reduce the dimensionality of $\underset{\sim}{\delta}$ involved in (7). However, this may make the value of $\widehat{PCS}$ more conservative.

For $1 \le q \le k-1$, denote $q(\underset{\sim}{\delta}) = (\delta_1, ..., \delta_q)^T$, and define

$$K(q(\underset{\sim}{\delta})) = \lim_{\substack{\delta_i \to \infty \\ i=q+1,\dots,k}} K(\underset{\sim}{\delta}) = K((\delta_1, \cdots, \delta_q, \infty, \cdots, \infty)^T),$$

and

$$C_q(z_1) = \{\underset{\sim}{\delta} : \; K(q(\underset{\sim}{\delta})) \ge z_1 \}.$$

Denote $x_{\alpha,q+1}$ as the solution of the following equation in c (setting $\delta_i = 0, \; i = 1, \cdots, q$, and $\delta_j = \infty, \; j = q+1, \cdots, k-1$ in (2)):

$$(q+1) \int_{-\infty}^{\infty} G^q(y-c) dG(y) = \alpha. \tag{9}$$

For $z_1 > x_{\alpha,q+1}$, define $C_q^*(z_1) = \{q(\underset{\sim}{\delta}) : \; K(q(\underset{\sim}{\delta})) = z_1 \}$, in which $(\delta_1, \cdots, \delta_q)^T$ is determined by the equation

$$\sum_{i=k-q}^{k} \int_{-\infty}^{\infty} \prod_{j=k-q}^{i-1} G(y - z_1 + \sum_{l=j}^{i-1} \delta_{k-l}) \prod_{j=i+1}^{k} G(y - z_1 + \sum_{l=i}^{j-1} \delta_{k-l}) dG(y) = \alpha. \tag{10}$$

Note that $C_q(z_1)$ is in $(R^+)^{k-1}$, $C_q^*(z_1)$ is in $(R^+)^q$, but both are determined only by $(\delta_1, ..., \delta_q)^T$.

For any $\underset{\sim}{\delta}(\ge 0)$, we have $K(q(\underset{\sim}{\delta})) \ge K(\underset{\sim}{\delta})$ by Lemma 2.2. Therefore

$$P_{\underset{\sim}{\delta}}\{K(q(\underset{\sim}{\delta})) \ge Z_1\} \ge P_{\underset{\sim}{\delta}}\{K(\underset{\sim}{\delta}) \ge Z_1\} = 1 - \alpha.$$

That is, $P_{\underset{\sim}{\delta}}\{\underset{\sim}{\delta} \in C_q(Z_1)\} \ge 1 - \alpha$. Hence, we have the following theorem:

Theorem 3.1 *For any given sample $\underset{\sim}{Z} = \underset{\sim}{z}$, $C_q(z_1)$ is a $100(1-\alpha)\%$ confidence region for $\underset{\sim}{\delta}$. Consequently, a $100(1-\alpha)\%$ lower confidence bound for the PCS is*

$$\begin{aligned} \widehat{PCS}_q &= \inf_{\underset{\sim}{\delta} \in C_q(z_1)} \left\{ \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} G(y + \sum_{j=1}^{i} \delta_j) dG(y) \right\} \\ &= \inf_{q(\underset{\sim}{\delta}) \in C_q^*(z_1)} \left\{ \int_{-\infty}^{\infty} \prod_{i=1}^{q-1} G(y + \sum_{j=1}^{i} \delta_j) [G(y + \sum_{j=1}^{q} \delta_j)]^{k-q} dG(y) \right\}, \end{aligned}$$

if $z_1 > x_{\alpha,q+1}$, and $1/k$ otherwise. □

Notice that $C_q(z_1) \supseteq C_{q+1}(z_1)$ if $g(y)$ is log-concave; so we have:

Theorem 3.2 *If the density function $g(y)$ is log-concave, then $\widehat{PCS}_q \le \widehat{PCS}_{q+1}$ and $x_{\alpha,q+1} \ge x_{\alpha,q+2}$, for $q = 1, ..., k-2$, with strict inequalities if strict concavity is assumed.* □

Theorem 3.2 implies that as q increases, the accuracy of the lower bound improves and the critical value $x_{\alpha,q+1}$ decreases. However, it should be noted that the price of choosing a larger q to get a more accurate lower confidence bound for the PCS is more complicated computation.

Specifically, for $q = 2$, if $z_1 > x_{\alpha,3}$, the set $C_2^*(z_1)$ is a collection of $(\delta_1, \delta_2)^T$ satisfying following equation

$$\int_{-\infty}^{\infty} \{G(y - z_1 - \delta_1 - \delta_2)G(y - z_1 - \delta_2) + G(y - z_1 - \delta_1)G(y - z_1 + \delta_2) \\ + G(y - z_1 + \delta_1 + \delta_2)G(y - z_1 + \delta_1)\} \, dG(y) = \alpha, \tag{11}$$

which is a curve in the (δ_1, δ_2)-plane. The corresponding lower bound is given by

$$\widehat{PCS}_2 = \inf_{(\delta_1,\delta_2)\in C_2^*(z_1)} \int_{-\infty}^{\infty} G(y + \delta_1)[G(y + \delta_1 + \delta_2)]^{k-2} dG(y). \tag{12}$$

The case $q = 1$ corresponds to Kim's(1986) result, in which case, the set $C_1^*(z_1)$ consists of only one value of δ_1 determined by

$$\int_{-\infty}^{\infty} [G(y - z_1 - \delta_1) + G(y - z_1 + \delta_1)] \, dG(y) = \alpha, \tag{13}$$

and

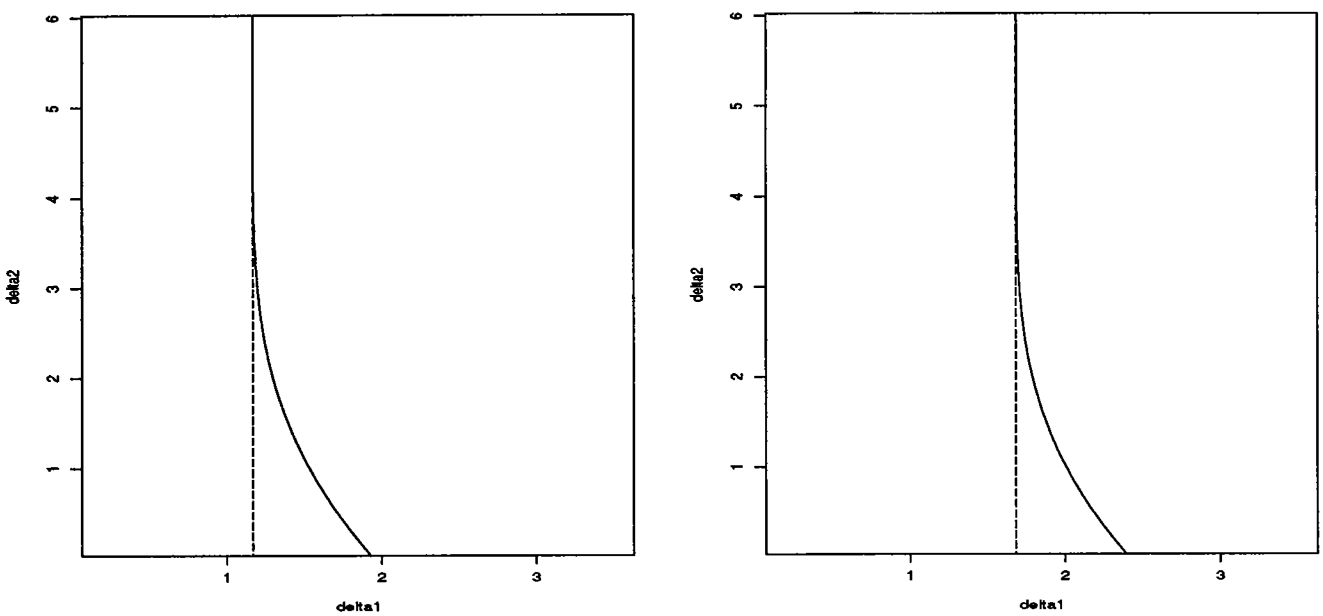
$$\widehat{PCS}_1 = \int_{-\infty}^{\infty} [G(y + \delta_1)]^{k-1} dG(y). \tag{14}$$

We note that, by choosing $q = 2$, we obtain a better lower confidence bound for the PCS than that of Kim (1986). To see how much improvement can be achieved for this choice of $q = 2$, notice that $\delta_1' < \delta_1''$ for all $\delta_1' \in C_1^*(z_1)$ and $(\delta_1'', \delta_2'')^T \in C_2^*(z_1)$ by the increasing property of $K(\cdot)$ in $\underset{\sim}{\delta}$. Comparing (12) with (14), we see that there is considerable improvement. Even if the infimum of (12) occurs at $\delta_2'' \to \infty$, we have $\delta_1'' = \delta_1'$, but $\widehat{PCS}_2 = \int_{-\infty}^{\infty} G(y + \delta_1'')dG(y)$ and $\widehat{PCS}_1 = \int_{-\infty}^{\infty}[G(y + \delta_1')]^{k-1}dG(y)$, which again results in the same conclusion as above. Another advantage in choosing $q = 2$ is that $x_{\alpha,3} < x_{\alpha,2}$, which implies that $\widehat{PCS}_1$ becomes a trivial bound while $\widehat{PCS}_2$ can still be pretty good. The following is an example to show the difference numerically.

Example 3.1 Let Y_i be independently distributed as $N(\theta_i, 1)$, $i = 1, 2, ..., k$. And let the differences of ordered statistics and parameters be denoted as $Z_i = Y_{(k-i+1)} - Y_{(k-i)}$, and $\delta_i = \theta_{[k-i+1]} - \theta_{[k-i]}$, respectively, where $\theta_{[i]}$ and $Y_{(i)}$ are not directly related. For a given sample $\underset{\sim}{Z} = \underset{\sim}{z}$ and confidence $100(1-\alpha)\%$, $C_2^*(z_1)$ represents the curve in the $(\delta_1, \delta_2)^T$-plane, determined by

$$\int_{-\infty}^{\infty} \{\Phi(y-z_1-\delta_1-\delta_2)\Phi(y-z_1-\delta_2)+\Phi(y-z_1-\delta_1)\Phi(y-z_1+\delta_2) + \Phi(y-z_1+\delta_1+\delta_2)\Phi(y-z_1+\delta_1)\}\,\varphi(y)dy=\alpha, \tag{15}$$

and $\widehat{PCS}_2$ is the infimum of $\int_{-\infty}^{\infty}\Phi(y+\delta_1)[\Phi(y+\delta_1+\delta_2)]^k\varphi(y)dy$ on the curve. For $\alpha=0.05$ and 0.10, respectively, and $z_1=3.5$, the curves are shown as follows:



The vertical lines correspond to $\delta_1=\delta$, the value of δ in Kim (1986).

Figure 1. Plots of $C_2^*(z_1)$ for $z_1=3.50$, $\alpha=0.05$(left) and $\alpha=0.10$(right).

The lower bounds $\widehat{PCS}_2$(ours) and $\widehat{PCS}_1$(Kim's) are computed for $k=3,4,\cdots,12$ and are listed in the following table:

Table 1. The lower bounds $\widehat{PCS}_2$(top entry) and $\widehat{PCS}_1$(bottom entry) for $z_1 = 3.5$

	k									
α	3	4	5	6	7	8	9	10	11	12
0.05	0.7954	0.7954	0.7745	0.7438	0.7176	0.6947	0.6744	0.6562	0.6398	0.6248
	0.6809	0.6043	0.5483	0.5049	0.4699	0.4410	0.4163	0.3954	0.3771	0.3608
0.10	0.8835	0.8835	0.8703	0.8498	0.8317	0.8155	0.8007	0.7873	0.7749	0.7634
	0.8072	0.7511	0.7071	0.6711	0.6409	0.6149	0.5923	0.5722	0.5543	0.5381

4 Unknown Scale Parameter Case

In this section, we consider the normal distributions to demonstrate how to construct a lower confidence bound for the PCS when each population has the same scale parameter. The result can be easily extended to the general distribution of location-scale parameter models with a similar setup as that in Gupta-Leu-Liang (1990).

The case when the scale parameter is known can be easily solved as discussed in the previous sections. So we only focus on the case of unknown scale parameter. The derivation is similar to those in Kim (1986) and Gupta-Leu-Liang (1990). Hence we simply state the result.

Suppose that the samples X_{ij} $(1 \leq i \leq k; 1 \leq j \leq n)$ are from $N(\theta_i, \sigma^2)$, where the common variance $\sigma^2 > 0$ is unknown. The best population is the one associated with $\theta_{[k]}$, and we select the population corresponding to the largest $\bar{X}_i$ value as the best.

Let S^2 denote the pooled sample variance. Note that vS^2/σ^2 has a χ^2 distribution with $v = k(n-1)$ df. Denote the cdf of $\sqrt{\chi^2/v}$ as $Q_v(u)$. Let $Z_1 = \sqrt{n}(\bar{X}_{(k)} - \bar{X}_{(k-1)})/S$, $\delta_i = \sqrt{n}(\theta_{k-i+1} - \theta_{k-i})/\sigma$, $i = 1, ..., k-1$, and $\underset{\sim}{\delta} = (\delta_1, ..., \delta_{k-1})^T$. We note that the distribution of Z_1 depends only on $\underset{\sim}{\delta}$. For a fixed integer q $(1 \leq q \leq k-1)$, denote $x_{\alpha,q+1}$ as the solution of the following equation in c

$$(q+1)\int_{-\infty}^{\infty} H_v^q(y-c)dH_v(y) = \alpha, \tag{16}$$

where $H_v(\cdot)$ is the cdf of the t distribution of df v. For a given sample value $Z_1 = z_1 > x_{\alpha,q+1}$, define

$$C_q^*(z_1) = \left\{ (\delta_1, ..., \delta_q)^T : \lim_{\substack{\delta_i \to \infty \\ i=q+1,...,k}} P_{\underset{\sim}{\delta}}(Z_1 > z_1) = \alpha \right\}. \tag{17}$$

Then $C_q^*(z_1)$ is the collection of $(\delta_1, ..., \delta_q)^T$ satisfying the following equation

$$\int_0^\infty \left[\sum_{i=k-q}^{k} \int_{-\infty}^{\infty} \prod_{j=k-q}^{i-1} \Phi(y - uz_1 + \sum_{l=j}^{i-1} \delta_{k-l}) \prod_{j=i+1}^{k} \Phi(y - uz_1 + \sum_{l=i}^{j-1} \delta_{k-l}) \varphi(y) dy \right] dQ_v(u) = \alpha. \tag{18}$$

And with $100(1-\alpha)\%$ confidence, we have

$$PCS \geq \inf_{q(\underset{\sim}{\delta}) \in C_q^*(z_1)} \left\{ \int_{-\infty}^{\infty} \prod_{i=1}^{q-1} \Phi(y + \sum_{j=1}^{i} \delta_j) [\Phi(y + \sum_{j=1}^{q} \delta_j)]^{k-q} d\Phi(y) \right\}. \tag{19}$$

If $z_1 \leq x_{\alpha,q+1}$, only the trivial lower bound $1/k$ can be achieved.

Specifically, for $q = 2$, the set $C_2^*(z_1)$ represents the curve in the (δ_1, δ_2)-plane determined by

$$\begin{aligned} \int_0^\infty \int_{-\infty}^{\infty} \{ & \Phi(y - uz_1 - \delta_1 - \delta_2)\Phi(y - uz_1 - \delta_2) + \Phi(y - uz_1 - \delta_1)\Phi(y - uz_1 + \delta_2) \\ & + \Phi(y - uz_1 + \delta_1 + \delta_2)\Phi(y - uz_1 + \delta_1) \} \varphi(y) dy dQ_v(u) = \alpha. \end{aligned} \tag{20}$$

And a $100(1-\alpha)\%$ lower confidence bound for the PCS can still be evaluated by (12) with $C_2^*(z_1)$ defined in (20). Again, the case $q = 1$ corresponds to Kim's (1986) result.

Example 4.1 *Kim's (1986) example.* Consider the example in Section 4 of Kim (1986), where $k = 5$, $v = 5(49) = 245$, $S = 228.26$, and $z_1 = \sqrt{2}(2.92)$. Choosing $q = 2$ as an illustration, we computed a 90% lower confidence bound for the PCS as 0.9465 by our method. The bound given by Kim (1986) for this case is 0.856.

Example 4.2 *Gupta-Liang (1991) example.* Consider the example in Section 4 of Gupta-Liang (1991). The data is taken from Example 3, page 506, of Gupta and Panchapakesan (1979), in which an experimenter wants to compare the glowing time of five different types of phosphorescent coatings of airplane instrument dials. Assume that the distributions of the glowing time for each type of phosphorescent coatings are normal with a common unknown variance. A two-stage selection rule as described in Gupta-Liang (1991) requires eight observations for each population. The data are summarized as $k = 5$, $n = 8$, $v = k(n-1) = 35$, $S = 5.06$, $\bar{x}_1 = 50.44$, $\bar{x}_2 = 50.83$, $\bar{x}_3 = 55.76$, $\bar{x}_4 = 57.56$, and $\bar{x}_5 = 64.88$. Then $z_1 = \sqrt{n}(\bar{x}_{(5)} - \bar{x}_{(4)})/S = 4.09$. For $\alpha = 0.1$, the Gupta-Liang bound is 0.518, and ours with $q = 2$ is 0.9295.

The reason that the method here gives a much better result is because the value of z_1 (4.09) from the data is relatively large and our method is good in using z_1, while the lower bound for $\theta_{[k]} - \theta_{[k-1]}$ given by the Gupta-Liang method is $\hat{\delta}_{L,4} = 0.4835$, which is very small.

Example 4.3 *The case when the Gupta-Liang method is better.* Let us look at an example showing that the Gupta-Liang method performs better than the method here. The data is taken from Problem 3.1 of Gibbons-Olkin-Sobel (1977), in which Black and Olson (1947) reported a study to compare dry shear strength of $k = 6$ different resin glues for bonding yellow birch plywood. They obtained $n = 10$ observations for each glue. Assume that the distributions for each glue are normal with common unknown variance. The data are summarized as $k = 6$, $n = 10$, $v = k(n-1) = 54$, $S = 25.63$, $\bar{x}_1 = 56.0$, $\bar{x}_2 = 78.8$, $\bar{x}_3 = 92.4$, $\bar{x}_4 = 128.8$, $\bar{x}_5 = 178.6$, and $\bar{x}_6 = 196.5$. Then $z_1 = \sqrt{n}(\bar{x}_{(6)} - \bar{x}_{(5)})/S = 2.2$. For $\alpha = 0.10$, only the trivial bound 1/6 can be obtained by using Kim's method because $2.2/\sqrt{2} < s_{\frac{\alpha}{2}} = 1.671$. Our lower bound with $q = 2$ is 0.4505, while the Gupta-Liang bound is very close to 0.5.

Examining the data carefully, we note that $\bar{x}_{(6)} - \bar{x}_{(5)} = 17.9$ is not large enough to claim that $\delta_1 > 0$ under any procedure, but $\bar{x}_{(6)} - \bar{x}_{(4)} = 67.7$ is so large that one of the first two populations can be claimed as the best at level $\alpha = 0.10$. The Gupta-Liang method uses the latter information to give a slightly better result. However, both bounds are low due to the nature of the data. Also see the following remark.

Remark 4.1 If the bound given by Kim (1986) is trivial, then our bound (others should be the same) can not be larger than 0.50, because there is a $\underset{\sim}{\delta}$ with $\delta_1 = 0$ in the confidence region $C(z_1)$.

Note: *Details of Computation.* To compute the lower confidence bounds using the method here with $q = 2$ in the three examples above, we took δ_2 from 0 to 7.5 with increment 0.25 and found the corresponding δ_1 determined by (20) using the method of bisection. Then the corresponding integral in (12) was evaluated, and its minimum was recorded, which gave our bound. The double integration in (20) was carried out via the method of Monte Carlo with sample size of 10,000, and the integral in (12) was evaluated via IMSL's subroutine QDAGI.

5 Appendix: Proof of Lemma 2.1

Actually, it is easy to show that (Kim (1986))

$$P_{\underset{\sim}{\theta}}[Y_{(k)} - Y_{(k-1)} > c] = \sum_{i=1}^{k} \int_{-\infty}^{\infty} \prod_{j \neq i} G(y + \theta_{[i]} - \theta_{[j]} - c) dG(y). \tag{21}$$

Hence, (a) and (b) follow.

For (c), it is sufficient to show that, given $c \geq 0$, $P_{\underset{\sim}{\delta}}\{Z_1 > c\}$ is increasing in δ_{k-m}, keeping other δ_i, $i \neq k-m$, fixed, for $m = 1, 2, ..., k-1$. Now we write $P_{\underset{\sim}{\delta}}\{Z_1 > c\}$ as

$$\begin{aligned} &\sum_{i=1}^{m} \int_{-\infty}^{\infty} \prod_{j=1}^{i-1} G(y - c + \sum_{l=j}^{i-1} \delta_{k-l}) \prod_{j=i+1}^{m} G(y - c - \sum_{l=i}^{j-1} \delta_{k-l}) \prod_{j=m+1}^{k} G(y - c - \sum_{l=i}^{j-1} \delta_{k-l}) dG(y) \\ &+ \sum_{i=m+1}^{k} \int_{-\infty}^{\infty} \prod_{j=1}^{m} G(y - c + \sum_{l=j}^{i-1} \delta_{k-l}) \prod_{j=m+1}^{i-1} G(y - c + \sum_{l=j}^{i-1} \delta_{k-l}) \prod_{j=i+1}^{k} G(y - c - \sum_{l=i}^{j-1} \delta_{k-l}) dG(y), \end{aligned} \tag{22}$$

therefore

$$\begin{aligned} &\frac{\partial P}{\partial \delta_{k-m}} \\ &= \sum_{i=1}^{m} \sum_{t=m+1}^{k} \Bigg\{ \int_{-\infty}^{\infty} \Bigg[g(y - c + \sum_{l=i}^{t-1} \delta_{k-l}) g(y) \prod_{j=1}^{i-1} G(y - c + \sum_{l=j}^{i-1} \delta_{k-l} + \sum_{l=i}^{t-1} \delta_{k-l}) \\ &\quad \cdot \prod_{j=i+1}^{t-1} G(y - c - \sum_{l=i}^{j-1} \delta_{k-l} + \sum_{l=i}^{t-1} \delta_{k-l}) \prod_{j=t+1}^{k} G(y - c - \sum_{l=i}^{j-1} \delta_{k-l} + \sum_{l=i}^{t-1} \delta_{k-l}) \Bigg] dy \\ &\quad - \int_{-\infty}^{\infty} \Bigg[g(y - c - \sum_{l=i}^{t-1} \delta_{k-l}) g(y) \prod_{j=1}^{i-1} G(y - c + \sum_{l=j}^{i-1} \delta_{k-l}) \prod_{j=i+1}^{t-1} G(y - c - \sum_{l=i}^{j-1} \delta_{k-l}) \prod_{j=t+1}^{k} G(y - c - \sum_{l=i}^{j-1} \delta_{k-l}) \Bigg] dy \Bigg\} \\ &= \sum_{i=1}^{m} \sum_{t=m+1}^{k} \int_{-\infty}^{\infty} \Big\{ g(y - c) g(y - \sum_{l=i}^{t-1} \delta_{k-l}) - g(y - c - \sum_{l=i}^{t-1} \delta_{k-l}) g(y) \Big\} \\ &\quad \cdot \prod_{j=1}^{i-1} G(y - c + \sum_{l=j}^{i-1} \delta_{k-l}) \prod_{j=i+1}^{t-1} G(y - c - \sum_{l=i}^{j-1} \delta_{k-l}) \prod_{j=t+1}^{k} G(y - c - \sum_{l=i}^{j-1} \delta_{k-l}) dy. \end{aligned}$$

The first equality is obtained through differentiating the third product in the first term and the first product in the second term of (22). The last equality results from a variable transformation $y' = y + \sum_{l=i}^{t-1} \delta_{k-l}$ in the first term of the previous equation.

By the equivalence between the concavity of $\log g(y)$ and the monotone likelihood ratio property of $g(y - \theta)$ in y, we have

$$g(y - c) g(y - \sum_{l=i}^{t-1} \delta_{k-l}) - g(y - c - \sum_{l=i}^{t-1} \delta_{k-l}) g(y) \geq 0,$$

with strict inequality if the concavity is strict. Hence the result follows. □

Acknowledgements

The research of Shanti S. Gupta and Yuning Liao was supported in part by NSF Grant DMS-8923071. The research of Chunfu Qiu was supported by a David Ross Fellowship at Purdue University. Jin Wang was supported in part by NSF Grant DMS-8717799. The authors wish to acknowledge gratefully the helpful comments and suggestions of two anonymous referees.

References

Anderson, P. O., Bishop, T. A. and Dudewicz, E. J. (1977). Indifference zone ranking and selection: confidence intervals for true achieved P(CD). *Comm. Statist. Theory Method* **A(6)**, 1121-1132.

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.

Bofinger, E. (1985). On the non-existence of consistent estimators for $P(CS)$. *Amer. J. Math. Manag. Sci.* **5**, 63-76.

Faltin, F. and McCulloch, C. E. (1983). On the small-sample properties of the Olkin-Sobel-Tong estimator of the probability of correct selection. *J. Amer. Statist. Assoc.* **78**, 464-467.

Gibbons, J. D., Olkin, I. and Sobel, M. (1977). *Selecting and Ordering Populations: A New Statistical Methodology.* John Wiley, New York.

Gupta, S. S. and Panchapakesan, S. (1979). *Multiple Decision Procedures.* John Wiley, New York.

Gupta, S. S., Leu L. Y., and Liang, T. (1990). On lower confidence bounds for PCS in truncated location parameter models.*Comm. Statist. Theory Method* **19(2)**, 527-546.

Gupta, S. S., and Liang, T. (1991). On a lower confidence bound for the probability of a correct selection: analytical and simulation studies. In *The Frontiers of Statistical Scientific Theory*

& Industrial Applications (Volume II of the Proceedings of ICOSCO-I, The First International Conference on Statistical Computing, March-April 1987.). American Sciences Press, Inc., Syracuse, New York, 77-95.

Kim, W. C. (1986). A lower confidence bound on the probability of a correct selection. *J. Amer. Statist. Assoc.* **81**, 1012-1017.

Olkin, I., Sobel, M. and Tong, Y. L. (1976). Estimating the true probability of a correct selection for location and scale parameter families. *Technical Report* **110**, Department of Statistics, Stanford University.

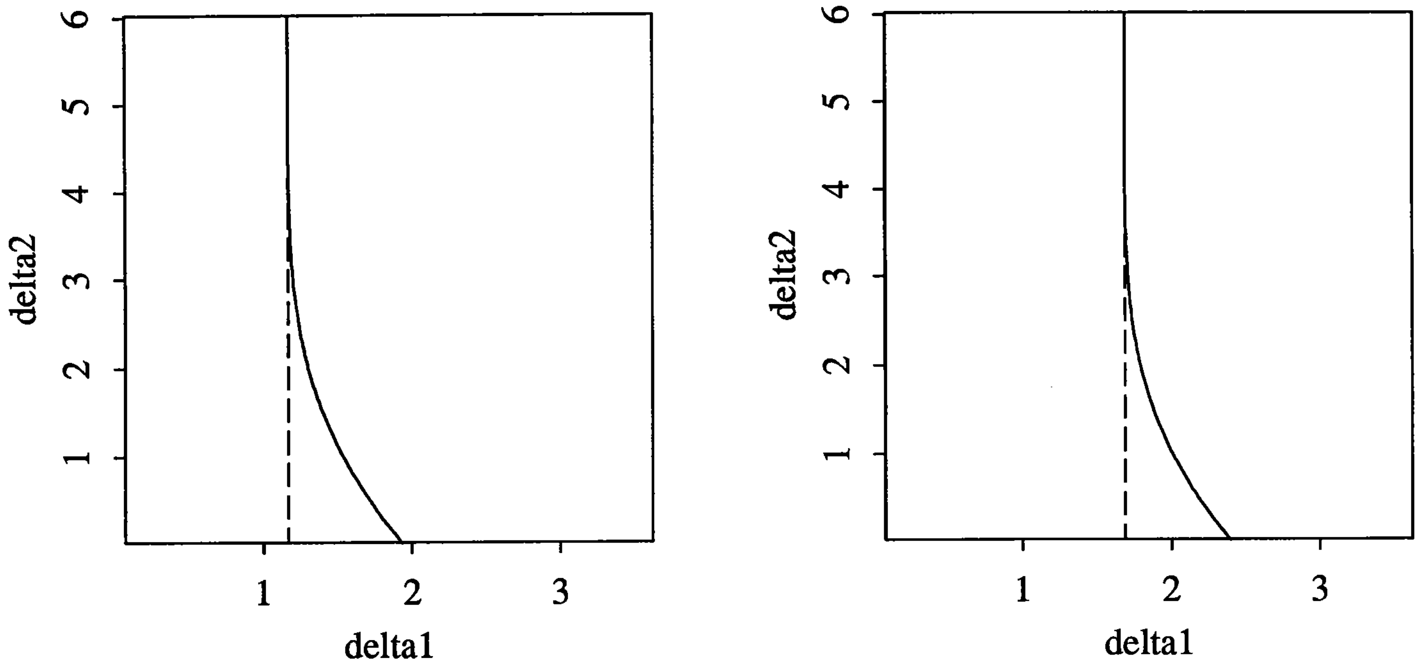
Olkin, I., Sobel, M. and Tong, Y. L. (1982). Bounds for a k-fold integral for location and scale parameter models with applications to statistical ranking and selection problem. *Statistical Decision Theory and Related Topics III* **Vol. 2.** Eds. S. S. Gupta and J. O. Berger. Academic Press, New York. 193-212.

Department of Statistics, Purdue University, West Lafayette, IN 47907, U.S.A.

New Drug Services Inc., 415 McFarlan Road, Suite 201, Kennette Square, PA 19348, U.S.A.

Department of Statistics, University of Illinois, Champaign, IL 61820, U.S.A.

School of Industrial Engineering, Purdue University, West Lafayette, IN 47907, U.S.A



The vertical lines correspond to $\delta_1 = \delta$, the value of δ in Kim (1986).

Figure 1. Plots of $C_2^*(z_1)$ for $z_1 = 3.50$, $\alpha = 0.05$(left) and $\alpha = 0.10$(right).