Hindawi Publishing Corporation
Journal of Probability and Statistics
Volume 2011, Article ID 765058, 11 pages
doi:10.1155/2011/765058

Research Article

Lower Confidence Bounds for the Probabilities of Correct Selection

Radhey S. Singh[1] and Narinder Kumar[2]

[1] *Department of Mathematics and Statistics, University of Guelph, ON, Canada N1G 2W1*
[2] *Department of Statistics, Panjab University, Chandigarh 160014, India*

Correspondence should be addressed to Radhey S. Singh, rssingh@uoguelph.ca

Received 6 October 2010; Accepted 12 January 2011

Academic Editor: A. Thavaneswaran

We extend the results of Gupta and Liang (1998), derived for location parameters, to obtain lower confidence bounds for the probability of correctly selecting the t best populations (PCS_t) simultaneously for all $t = 1, \ldots, k-1$ for the general scale parameter models, where k is the number of populations involved in the selection problem. The application of the results to the exponential and normal probability models is discussed. The implementation of the simultaneous lower confidence bounds for PCS_t is illustrated through real-life datasets.

1. Introduction

The population Π_i is characterized by an unknown scale parameter $\theta_i(> 0)$, $i = 1, \ldots, k$. Let T_i be an appropriate statistic for θ_i, based on a random sample of size n from population Π_i, having the probability density function (pdf) $f_{\theta_i}(x) = (1/\theta_i) f(x/\theta_i)$ with the corresponding cumulative distribution function (cdf) $F_{\theta_i}(x) = F(x/\theta_i)$, $x > 0$, $\theta_i > 0$, $i = 1, \ldots, k$. $F(\cdot)$ is an arbitrary continuous cdf with pdf $f(\cdot)$. Let the ordered values of T_i's and θ_i's be denoted by $T_{[1]}, \ldots, T_{[k]}$ and $\theta_{[1]}, \ldots, \theta_{[k]}$, respectively. Let $T_{(i)}$ be the statistic having a scale parameter $\theta_{[i]}$. Let $\Pi_{(i)}$ denote the population associated with $\theta_{[i]}$, the ith smallest of θ_i's. Any other population or sample quantity associated with $\Pi_{(i)}$ will be denoted by the subscript (i) attached to it. Throughout, we assume that there is no prior knowledge about which of $\Pi_1, \ldots, \Pi_k$ is $\Pi_{(i)}$, $i = 1, \ldots, k$ and that $\theta_1, \ldots, \theta_k$ are unknown. Call the populations $\Pi_{(k)}, \Pi_{(k-1)}, \ldots, \Pi_{(k-t+1)}$ as the t best populations.

In practice, the interest is to select the populations $\Pi_{(k)}, \Pi_{(k-1)}, \ldots, \Pi_{(k-t+1)}$, that is, the populations associated with the largest unknown parameters $\theta_{[k]}, \theta_{[k-1]}, \ldots, \theta_{[k-t+1]}$. For this, the natural selection rule "select the populations corresponding to t largest T_i's, that is, $T_{[k]}, T_{[k-1]}, \ldots, T_{[k-t+1]}$ as the t best populations" is used. However, it is possible that selected populations according to the natural selection rule may not be the best. Therefore, a question which naturally arises is: what kind of confidence statement can be made about these selection results? Motivated by this, we make an effort to answer this question.

Let CS_t (a correct selection of the t best populations) denote the event that t best populations are actually selected. Then, the probability of correct selection of the t best populations (PCS_t) is:

$$\begin{aligned}
\mathrm{PCS}_t(\theta) &= P\left\{\max_{1\leq i\leq k-t} T_{(i)} < \min_{k-t+1\leq j\leq k} T_{(j)}\right\} \\
&= \int \prod_{i=1}^{k-t} F\left(\frac{y}{\theta_{[i]}}\right) d\left\{1-\prod_{j=k-t+1}^{k} \bar{F}\left(\frac{y}{\theta_{[j]}}\right)\right\} \qquad (1.1a)\\
&= \int \prod_{j=k-t+1}^{k} \overline{F}\left(\frac{y}{\theta_{[j]}}\right) d\prod_{i=i}^{k-t} F\left(\frac{y}{\theta_{[i]}}\right), \qquad (1.1b)
\end{aligned}$$

where $\overline{F}(\cdot) = 1 - F(\cdot)$ and $\theta = (\theta_1, \ldots, \theta_k)$.

For the k populations differing in their location parameters $\mu_1, \ldots, \mu_k$, Gupta and Liang [1] provided a novel idea to construct simultaneous lower confidence bounds for the PCS_t for all $t = 1, \ldots, k-1$. Their result was applied to the selection of the t best means of normal populations. For other references under location set up, one may refer to the papers cited therein.

For other relevant references, one may refer to Gupta et al. [2], Gupta and Panchpakesan [3], Mukhopadhyay and Solanky [4], and the review papers by Gupta and Panchapakesan [5, 6], Khamnei and Kumar [7], and the references cited therein.

In this article, we use the methodology and results of Gupta and Liang [1] to derive simultaneous lower confidence bounds for the PCS_t for all $t = 1, \ldots, k-1$ under the general scale parameter models. Section 2 deals with obtaining such intervals. The application of the results to the exponential and normal probability models is discussed in Section 3. In the case of an exponential distribution, Type-II censored data is also considered. In Section 4, we have given some numerical examples, based on real life data sets, to illustrate the procedure of finding out simultaneous lower confidence bounds for the probability of correctly selecting the t best populations (PCS_t).

2. Simultaneous Lower Confidence Bounds for PCS_t

Most of the results in this Section are as a simple consequence of the results obtained by Gupta and Liang [1].

From (1.1a), the $\mathrm{PCS}_t(\theta)$ can be expressed as

$$\mathrm{PCS}_t(\theta) = \sum_{j=k-t+1}^{k} P_{tj}(\theta), \qquad (2.1)$$

where for each $j = k - t + 1, \ldots, k$,

$$P_{tj}(\theta) = \int \prod_{i=1}^{k-t} F\left(y\Delta_{tji}(1)\right) \prod_{m=k-t+1}^{j-1} \overline{F}\left(y\Delta_{tjm}(2)\right) \prod_{l=j+1}^{k} \overline{F}\left(y\Delta_{tjl}(3)\right) dF(y), \tag{2.2}$$

where $\Delta_{tji}(1) = \theta_{[j]}/\theta_{[i]} \geq 1$ for $1 \leq i \leq k - t < j$; $\Delta_{tjm}(2) = \theta_{[j]}/\theta_{[m]} \geq 1$ for $k - t + 1 \leq m < j$ and $\Delta_{tjl}(3) = \theta_{[j]}/\theta_{[l]} \leq 1$ for $k - t + 1 \leq j < l \leq k$. Here, $\prod_s^t \equiv 1$ if $t < s$. Note that for each $j(k - t + 1 \leq j \leq k)$, $P_{tj}(\theta)$ is increasing in $\Delta_{tji}(1)$, and decreasing in $\Delta_{tjm}(2)$ and $\Delta_{tjl}(3)$, respectively. Thus, if we develop simultaneous lower confidence bounds for $\Delta_{tji}(1)$, $1 \leq i \leq k - t$ and upper confidence bounds for $\Delta_{tjm}(2)$ and $\Delta_{tjl}(3)$, $k - t + 1 \leq m \leq j \leq l \leq k$, $m \neq j$, $l \neq j$ for all $t = 1, \ldots, k - 1$, then, simultaneous lower confidence bounds for $\mathrm{PCS}_t(\theta)$ for all $t = 1, \ldots, k - 1$ can be established.

Also, from (1.1b), the $\mathrm{PCS}_t(\theta)$ can be expressed as

$$\mathrm{PCS}_t(\theta) = \sum_{i=1}^{k-t} Q_{ti}(\theta), \tag{2.3}$$

where for each $i = 1, \ldots, k - t$,

$$Q_{ti}(\theta) = \int \prod_{m=1}^{i-1} F(z\delta_{tim}(1)) \prod_{l=i+1}^{k-t} F(z\delta_{til}(2)) \prod_{j=k-t+1}^{k} \overline{F}\left(z\delta_{tij}(3)\right) dF(z) \tag{2.4}$$

and $\delta_{tim}(1) = \theta_{[i]}/\theta_{[m]} \geq 1$ for $1 \leq m < i \leq k - t$; $\delta_{til}(2) = \theta_{[i]}/\theta_{[l]} \leq 1$ for $1 \leq i < l \leq k - t$; and $\delta_{tij}(3) = \theta_{[i]}/\theta_{[j]} \leq 1$ for $i \leq k - t < j \leq k$. Note that for each $i = 1, \ldots, k - t$, $Q_{ti}(\theta)$ is increasing in $\delta_{tim}(1)$, $\delta_{til}(2)$, and decreasing in $\delta_{tij}(3)$, respectively. Thus, if simultaneous lower confidence bounds for $\delta_{tim}(1)$ and $\delta_{til}(2)$, $1 \leq m \leq i \leq l \leq k - t$, $m \neq i$, $l \neq i$ and upper confidence bounds for $\delta_{til}(3)$, $i \leq k - t < j \leq k$ can be obtained, and, thereafter, by using (2.3) and (2.4), we can obtain simultaneous lower confidence bounds for the $\mathrm{PCS}_t(\theta)$ for all $t = 1, \ldots, k - 1$.

We use the results of Gupta and Liang [1] to construct simultaneous lower confidence bounds for all $\Delta_{tji}(1)$, $\delta_{tim}(1)$, $\delta_{til}(2)$, and upper confidence bounds for all $\Delta_{tjm}(2)$, $\Delta_{tjl}(3)$, and $\delta_{til}(3)$ for all $t = 1, \ldots, k - 1$.

For each $P^*(0 < P^* < 1)$, let $c(k, n, P^*)$ be the value such that

$$P_{\underset{\sim}{\theta}}\left\{\left[\frac{\max_{1\leq i\leq k}(T_i/\theta_i)}{\min_{1\leq j\leq k}(T_j/\theta_j)}\right] \leq c(k, n, P^*)\right\} = P^*. \tag{2.5}$$

Note that since T_i has a distribution function $F(y/\theta_i)$, $i = 1, \ldots, k$, the value of $c = c(k, n, P^*)$ is independent of the parameter θ. Let

$$
\begin{aligned}
E &= \left\{ \frac{\max_{1\le i\le k}(T_i/\theta_i)}{\min_{1\le j\le k}(T_j/\theta_j)} \le c \right\},\\
E_1 &= \left\{ \left(\frac{T_{[i]}}{cT_{[j]}} \right)^{+} \le \frac{\theta_{[i]}}{\theta_{[j]}} \le \left(\frac{cT_{[i]}}{T_{[j]}} \right),\ \forall 1 \le j < i \le k \right\},\\
E_2 &= \left\{ \left(\frac{T_{[i]}}{cT_{[j]}} \right) \le \frac{\theta_{[i]}}{\theta_{[j]}} \le \left(\frac{cT_{[i]}}{T_{[j]}} \right)^{-},\ \forall 1 \le i < j \le k \right\},
\end{aligned}
\tag{2.6}
$$

where $y^+ = \max(1, y)$ and $y^- = \min(1, y)$.

Lemma 2.1. (a) *$E \subset E_1 \cap E_2$ and, therefore,*
(b) *$P_\theta\{E_1 \cap E_2\} \ge P_\theta\{E\} = P^*$ for all θ.*

Proof. Part (a) follows on the lines of Lemma 2.1 of Gupta and Liang [1] by noting that $\theta_{[i]}/\theta_{[j]} \ge 1$ as $j < i$ and $\theta_{[i]}/\theta_{[j]} \le 1$ for $i < j$, we have $E \subset E_1$ and $E \subset E_2$. Therefore, $E \subset E_1 \cap E_2$.

Part (b) follows immediately from part (a) and (2.5).

For each $t = 1, \ldots, k-1$ and $j = k-t+1, \ldots, k$, let

$$
\begin{aligned}
\widehat{\Delta}_{tji}(1) &= \left(\frac{T_{[j]}}{cT_{[i]}} \right)^{+} \quad \text{for } 1 \le i \le k-t;\\
\widehat{\Delta}_{tjm}(2) &= \left(\frac{cT_{[j]}}{T_{[m]}} \right) \quad \text{for } k-t+1 \le m < j;\\
\widehat{\Delta}_{tjl}(3) &= \left(\frac{cT_{[j]}}{T_{[l]}} \right)^{-} \quad \text{for } j < l \le k.
\end{aligned}
\tag{2.7}
$$

Also, for each $t = 1, \ldots, k-1$ and $i = 1, \ldots, k-t$, let

$$
\begin{aligned}
\widehat{\delta}_{tim}(1) &= \left(\frac{T_{[i]}}{cT_{[m]}} \right)^{+} \quad \text{for } 1 \le m \le i-1;\\
\widehat{\delta}_{til}(2) &= \left(\frac{T_{[i]}}{cT_{[l]}} \right) \quad \text{for } i+1 \le l \le k-t;\\
\widehat{\delta}_{tij}(3) &= \left(\frac{cT_{[i]}}{T_{[j]}} \right)^{-} \quad \text{for } k-t+1 \le j \le k.
\end{aligned}
\tag{2.8}
$$

□

The following Lemma is a direct result of Lemma 2.1.

Lemma 2.2. *With probability at least P^*, the following* (A1) *and* (A2) *hold simultaneously.*

(A1) *For each $t = 1, \ldots, k-1$ and each $j = k-t+1, \ldots, k$,*

$$\begin{aligned}
\Delta_{tji}(1) &\geq \widehat{\Delta}_{tji}(1), \quad \forall i = 1, \ldots, k-t; \\
\Delta_{tjm}(2) &\leq \widehat{\Delta}_{tjm}(2), \quad \forall k-t+1 \leq m < j; \\
\Delta_{tjl}(3) &\leq \widehat{\Delta}_{tjl}(3), \quad \forall j < l \leq k.
\end{aligned} \tag{2.9}$$

(A2) *For each $t = 1, \ldots, k-1$ and each $i = 1, \ldots, k-t$,*

$$\begin{aligned}
\delta_{tim}(1) &\geq \widehat{\delta}_{tim}(1), \quad \forall 1 \leq m \leq i-1; \\
\delta_{til}(2) &\geq \widehat{\delta}_{til}(2), \quad \forall i+1 \leq l \leq k-t; \\
\delta_{tij}(3) &\leq \widehat{\delta}_{tij}(3), \quad \forall k-t+1 \leq j \leq k.
\end{aligned} \tag{2.10}$$

Now, for each $t = 1, \ldots, k-1$ and each $j = k-t+1, \ldots, k$, define

$$\widehat{P}_{tj} = \int \prod_{i=1}^{k-t} F\left(y\widehat{\Delta}_{tji}(1)\right) \prod_{m=k-t+1}^{j-1} \overline{F}\left(y\widehat{\Delta}_{tjm}(2)\right) \prod_{l=j+1}^{k} \overline{F}\left(y\widehat{\Delta}_{tjl}(3)\right) dF(y), \tag{2.11}$$

and for each $t = 1, \ldots, k-1$, define

$$\widehat{P}_t = \sum_{j=k-t+1}^{k} \widehat{P}_{tj}. \tag{2.12}$$

Also, for each $t = 1, \ldots, k-1$ and each $i = 1, \ldots, k-t$, define

$$\widehat{Q}_{ti} = \int \prod_{m=1}^{i-1} F\left(z\widehat{\delta}_{tim}(1)\right) \prod_{l=i+1}^{k-t} F\left(z\widehat{\delta}_{til}(2)\right) \prod_{j=k-t+1}^{k} \overline{F}\left(z\widehat{\delta}_{tij}(3)\right) dF(z), \tag{2.13}$$

$$\widehat{Q}_t = \sum_{i=1}^{k-t} \widehat{Q}_{ti}. \tag{2.14}$$

Define

$$P_{tL} = \max\left(\widehat{P}_t, \widehat{Q}_t\right). \tag{2.15}$$

The authors propose $P_{tL} = \max(\widehat{P}_t, \widehat{Q}_t)$ as an estimator of a lower confidence bound of the $PCS_t(\theta)$ for each $t = 1, \ldots, k-1$. The authors have the following theorem.

Theorem 2.3. *$P_{\underline{\theta}}\{PCS_t(\theta) \geq P_{tL}$ for all $t = 1, \ldots, k-1\} \geq P^*$ for all θ.*

Proof. Note that $P_{tj}(\theta)$ is increasing in $\Delta_{tji}(1)$ and decreasing in $\Delta_{tjm}(2)$ and$\Delta_{tjl}(3)$. Also, $Q_{ti}(\theta)$ is increasing in $\delta_{tim}(1)$, $\delta_{til}(2)$ and decreasing in $\delta_{tij}(3)$. Then, by using (2.2), (2.4), (2.11), (2.13), and Lemma 2.2, we have

$$P_\theta\left\{P_{tj}(\theta) \geq \widehat{P}_{tj},\ \forall j = k-t+1, ..., k,\ \text{and}\ Q_{ti}(\theta) \geq \widehat{Q}_{ti},\ \forall i = 1, \ldots, k-t,\ t = 1, \ldots, k-1\right\} \geq P^*. \tag{2.16}$$

Then, by (2.1), (2.3), (2.12), (2.14), and (2.16), we have

$$\begin{aligned} P^* &\leq P\left\{\mathrm{PCS}_t(\theta) \geq \widehat{P}_t, \mathrm{PCS}_t(\theta) \geq \widehat{Q}_t,\ \forall t = 1, \ldots, k-1\right\} \\ &= P_\theta(\mathrm{PCS}_t(\theta) \geq P_{tL}\ \forall t = 1, \ldots, k-1\}. \end{aligned} \tag{2.17}$$

This proves the theorem. □

3. Applications to Exponential and Normal Distributions

3.1. Exponential Distribution

(i) Complete Data

Let X_{ij}, $j = 1, \ldots, n$ denote a random sample of size n from the two-parameter exponential population Π_i having pdf $f(x) = (1/\theta_i)\exp\{-(x-\mu_i)/\theta_i\}$, $i = 1, \ldots, k$. Let $M_i = \min_{1\leq j\leq n} X_{ij}$ and $Y_i = \sum_{j=1}^n (X_{ij} - M_i)$. Here, (M_i, Y_i) is a sufficient statistic for (μ_i, θ_i), $i = 1, \ldots, k$. Y_i/θ_i has a standardized gamma distribution with shape parameter $\theta = n-1$, $i = 1, \ldots, k$. Then, based on statistics $Y_1, \ldots, Y_k$ by applying the natural selection rule for each $t = 1, \ldots, k-1$, the associated PCS_t is

$$\begin{aligned} \mathrm{PCS}_t(\theta) &= \sum_{j=k-t+1}^{k} P_{tj}(\theta) \\ &= \sum_{i=1}^{k-t} Q_{ti}(\theta), \end{aligned} \tag{3.1}$$

where

$$\begin{aligned} P_{tj}(\theta) &= \int \prod_{i=1}^{k-t} F\left(y\Delta_{tji}(1)\right) \prod_{m=k-t+1}^{j-1} \overline{F}\left(y\Delta_{tjm}(2)\right) \prod_{l=j+1}^{k} \overline{F}\left(y\Delta_{tjl}(3)\right) dF(y), \\ Q_{ti}(\theta) &= \int \prod_{m=1}^{i-1} F(z\delta_{tim}(1)) \prod_{l=i+1}^{k-t} F(z\delta_{til}(2)) \prod_{j=k-t+1}^{k} \overline{F}\left(z\delta_{tij}(3)\right) dF(z), \end{aligned} \tag{3.2}$$

and $F(\cdot)$ is the distribution function of the standardized gamma distribution with shape parameter $\theta = n-1$.

For each P^* $(0 < P^* < 1)$, let $c = c(k, P^*, n)$ be the P^* quantile of the distribution of the random variable Z defined as $Z = \{\max_{1\le i\le k}(Y_i/\theta_i)\}/\{\min_{1\le i\le k}(Y_i/\theta_i)\}$, the extreme quotient of independent and identically distributed random variables Y_i.

Given k, n, P^* the value of c can be obtained from the tables of Hartley's ratio Z with $2(n-1)$ degrees of freedom refer to Pearson and Hartley [8].

For each $t = 1, \ldots, k-1$ and each $j = k-t+1, \ldots, k$, let

$$\widehat{P}_{tj} = \int \prod_{i=1}^{k-t} F\left(y\widehat{\Delta}_{tji}(1)\right) \prod_{m=k-t+1}^{j-1} \overline{F}\left(y\widehat{\Delta}_{tjm}(2)\right) \prod_{l=j+1}^{k} \bar{F}\left(y\widehat{\Delta}_{tjl}(3)\right) dF(y), \tag{3.3}$$

and for each $t = 1, \ldots, k-1$ and each $i = 1, \ldots, k-t$, let

$$\widehat{Q}_{ti} = \int \prod_{m=1}^{i-1} F\left(z\widehat{\delta}_{tim}(1)\right) \prod_{l=i+1}^{k-t} F\left(z\widehat{\delta}_{til}(2)\right) \prod_{j=k-t+1}^{k} \overline{F}\left(z\widehat{\delta}_{tij}(3)\right) dF(z), \tag{3.4}$$

where $\widehat{\Delta}_{tji}(1)$, $\widehat{\Delta}_{tjm}(2)$, and $\widehat{\Delta}_{tjl}(3)$ are defined as (2.7) and $\widehat{\delta}_{tim}(1)$, $\widehat{\delta}_{til}(2)$, and $\widehat{\delta}_{tij}(3)$ are defined in (2.8) with c chosen from Pearson and Hartley's tables.

For each $t = 1, \ldots, k-1$, let

$$\begin{aligned} \widehat{P}_t &= \sum_{j=k-t+1}^{k} \widehat{P}_{tj}, \\ \widehat{Q}_t &= \sum_{i=1}^{k-t} \widehat{Q}_{ti}. \end{aligned} \tag{3.5}$$

Then, by Theorem 2.3, we can conclude the following.

Theorem 3.1. *$P_\theta\{PCS_t(\theta) \ge \max(\widehat{P}_t, \widehat{Q}_t)$ for all $t = 1, \ldots, k-1\} \ge P^*$ for all θ.*

(ii) Type-II Censored Data

From each population Π_i, $i = 1, \ldots, k$, we take a sample of n items. Let $X_{i[1]}, \ldots, X_{i[n]}$ denote the order statistic representing the failure times of n items from population Π_i, $i = 1, \ldots, k$. Let r be a fixed integer such that $1 \le r \le n$. Under Type-II censoring, the first r failures from each population Π_i are to be observed. The observations from population Π_i cease after observing $X_{i[r]}$. The $(n-r)$ items whose failure times are not observable beyond $X_{i[r]}$ become the censored observations. Type-II censoring was investigated by Epstein and Sobel [9]. The sufficient statistic for θ_i, when location parameters are known, is

$$U_i = \sum_{j=1}^{r} X_{i[j]} + (n-r)X_{i[r]}, \quad i = 1, \ldots, k. \tag{3.6}$$

U_i is called the total time on test (TTOT) statistic. It is easy to verify that U_i/θ_i has standardized gamma distribution with shape parameter $r, i = 1, \ldots, k$. Again, the results of complete data can be applied simply by taking $\vartheta = r$.

Table 1

π_1:	999	112	242	991	111	1	587	389	38	25	357
π_2:	25	21	13	87	2	20	7	24	99	8	99
π_3:	24	18	31	51	90	52	73	8	36	48	7
π_4:	52	164	19	53	15	43	340	133	111	231	378

3.2. *Normal Distribution*

Let Π_i denote the normal population with mean μ_i and variance θ_i (both unknown), $i = 1, \ldots, k$. The sufficient statistic for θ_i based on a random sample $X_{i1}, \ldots, X_{in}$ of size n from Π_i is $Y_i^* = (1/(n-1)) \sum_{j=1}^{n} (X_{ij} - \overline{X}_i)^2$, where $\overline{X}_i = (1/n) \sum_{j=1}^{n} X_{ij}$, $i = 1, \ldots, k$. It can be verified that $\{(n-1)Y_i^*\}/(2\theta_i)$ is a standardized gamma variate with shape parameter $(n-1)/2$, $i = 1, \ldots, k$. Once again, the above results of exponential distribution can be used with $\vartheta = (n-1)/2$.

To illustrate the implementation of the simultaneous lower confidence bounds for the probability of correctly selecting the t best populations (PCS_t), we consider the following examples.

4. Examples

Example 4.1. Hill et al. [10] considered data on survival days of patients with inoperable lung cancer, who were subjected to a test chemotherapeutic agent. The patients are divided into the following four categories depending on the histological type of their tumor: squamous, small, adeno, and large denoted by π_1, π_2, π_3, and π_4, respectively, in this article. The data are a part of a larger data set collected by the Veterans Administrative Lung Cancer Study Group in the USA.

We consider a random sample of eleven survival times from each group, and they are given in Table 1.

Using the standard results of reliability (refer to Lawless [11]), one can check the validity of the two-parameter exponential model for Table 1. In this example, the populations with larger survival times (i.e., larger Y_i's) are desirable.

For Table 1 data set:

$$Y_1 = 3841, \qquad Y_2 = 383, \qquad Y_3 = 361, \qquad Y_4 = 1374. \tag{4.1}$$

Hence, according to natural selection rule, the populations π_1, π_2, and π_4 are selected as the t ($t = 1, 2, 3$) best populations, that is, for $t = 1$, population π_1 which has largest survival time is the best; for $t = 2$, populations π_1 and π_4 which have the two largest survival times are the best; and for $t = 3$, populations π_1, π_2, and π_4 which have the three largest survival times are the best. However, it i,s possible that selected populations according to the natural selection rule may not be the best. Therefore, we wish to find out a confidence statement that can be made about the probability of correctly selecting the t best populations (PCS_t) simultaneously for all $t = 1, 2, 3$.

Here, $k = 4$, $n = 11$, and, by taking $P^* = 0.95$, we get, from the tables of Pearson and Hartley [8], $c = c(k, n, P^*) = 3.29$.

Table 2

T	1	2	3
$\widehat{P}_t$	0.407125	0.143943	0.088946
$\widehat{Q}_t$	0.551725	0.33380	0.174162
Max$(\widehat{P}_t,\widehat{Q}_t)$	0.551725	0.33380	0.174162

Table 3

T	1	2
$\widehat{P}_t$	0.424471	0.164871
$\widehat{Q}_t$	0.163855	0.248274
max$(\widehat{P}_t,\widehat{Q}_t)$	0.424471	0.248274

Table 4

π_1:	1.54	0.66	1.70	1.82	2.75	0.66	0.55	0.18	10.6	10.63	0.71
π_2:	1.99	2.15	1.08	0.93	0.82	0.49	2.80	3.82	0.02	3.72	3.57
π_3:	3.17	0.80	1.13	1.08	2.12	1.56	1.34	2.10	7.21	3.83	5.13

Then, $\widehat{P}_t$ and $\widehat{Q}_t$ computed for the above data set using (3.5) are given in Table 2.

From Table 2, we have, with at least a 95% confidence coefficient, that simultaneously $\mathrm{PCS}_1(\theta) \geq 0.551725$, $\mathrm{PCS}_2(\theta) \geq 0.33380$, and $\mathrm{PCS}_3(\theta) \geq 0.174162$.

Example 4.2. Nelson [12] considered the data which represent times to breakdown in minutes of an insulating fluid subjected to high voltage stress. The times in their observed order were divided into three groups. After analyzing the data, it was shown to follow an exponential distribution. We consider the following data based on a random sample of size 11 each from the three groups and the observations are in Table 4.

For the above data set:

$$Y_1 = 20.82, \qquad Y_2 = 21.17, \qquad Y_3 = 20.67. \tag{4.2}$$

Hence, according to natural selection rule, the populations π_1, π_2 are selected as the t ($t = 1,2$) best populations, that is, for $t = 1$, population π_1 which has largest survival time is the best; and for $t = 2$, populations π_1 and π_2 which have the two largest survival times are the best. However, it is possible that selected populations according to the natural selection rule may not be the best. Therefore, we wish to find out a confidence statement that can be made about the probability of correctly selecting the t best populations (PCS_t) simultaneously for all $t = 1,2$.

Here, $k = 3$, $n = 11$, and, by taking $P^* = 0.95$, we get, from the tables of Pearson and Hartley [8], $c = c(k,n,P^*) = 2.95$.

Then, $\widehat{P}_t$ and $\widehat{Q}_t$ computed for the above data set using (3.5) are given in Table 3.

From Table 3, we have, with at least a 95% confidence coefficient, that simultaneously $\mathrm{PCS}_1(\theta) \geq 0.424471$ and $\mathrm{PCS}_2(\theta) \geq 0.248274$.

Table 5

π_1:	413	100	169	447	201	118	67
π_2:	10	14	20	44	29	26	23
π_3:	11	4	80	54	63	18	24
π_4:	22	3	46	22	30	23	14

Example 4.3. Proschan [13] considered the data on intervals between failures (in hours) of the air-conditioning system of a fleet of 13 Boeing 720 jet air planes. After analyzing the data, he found that the failure distributions of the air-conditioning system for each of the planes was well approximated as exponential. We consider the following data based on four random samples of size seven each, and the observations in the samples are mentioned in Table 5.

For the above data set:

$$Y_1 = 1046, \qquad Y_2 = 96, \qquad Y_3 = 226, \qquad Y_4 = 139. \tag{4.3}$$

Hence, according to natural selection rule, the populations π_1, π_3, and π_4 are selected as the t $(t = 1, 2, 3)$ best populations.

Here, $k = 4$, $n = 7$ and, by taking $P^* = 0.99$, we get, from the tables of Pearson and Hartley [8], $c = c(k, n, P^*) = 6.90$.

Proceeding on the lines similar to Examples 4.1 and 4.2, we have, with at least a 99% confidence coefficient, that simultaneously $\mathrm{PCS}_1(\theta) \geq 0.360517$, $\mathrm{PCS}_2(\theta) \geq 0.217558$, and $\mathrm{PCS}_3(\theta) \geq 0.154598$.

Acknowledgments

The authors thank the editor, the associate editor, and an anonymous referee for their helpful comments which led to the improvement of this paper.

References

[1] S. S. Gupta and T. Liang, "Simultaneous lower confidence bounds for probabilities of correct selections," *Journal of Statistical Planning and Inference*, vol. 72, no. 1-2, pp. 279–290, 1998.

[2] S.S. Gupta, I. Olkin, and M. Sobel, *Selecting and Ordering Populations: A New Statistical Methodology*, John Wiley & Sons, New York, NY, USA, 1977.

[3] S. S. Gupta and S. Panchapakesan, *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations*, John Wiley & Sons, 1979.

[4] N. Mukhopadhyay and T. K. S. Solanky, *Multistage Selection and Ranking Procedures*, vol. 142, Marcel Dekker, New York, NY, USA, 1994.

[5] S. S. Gupta and S. Panchapakesan, "Subset selection procedures: review and assessment," *American Journal of Mathematical and Management Sciences*, vol. 5, no. 3-4, pp. 235–311, 1985.

[6] S. S. Gupta and S. Panchpakesan, "Design of experiments with selection and ranking goals," in *Handbook of Statistics*, vol. 13, Elsevier Science, Amsterdam, The Netherlands, 1995.

[7] H. J. Khamnei and N. Kumar, "Improved confidence bounds for the probability of correct selection: the scale parameter case," *Metron*, vol. 57, no. 3-4, pp. 147–167, 1999.

[8] E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians. Vol. I*, Cambridge University Press, London, UK, 1954.

[9] B. Epstein and M. Sobel, "Life testing," *Journal of the American Statistical Association*, vol. 48, pp. 486–502, 1953.

[10] N. J. Hill, A. R. Padmanabhan, and Madan L. Puri, “Adaptive nonparametric procedures and applications,” *Journal of the Royal Statistical Society. Series C*, vol. 37, no. 2, pp. 205–218, 1988.
[11] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, New York, NY, USA, 1982.
[12] W. Nelson, *Applied Life Data Analysis*, John Wiley & Sons, New York, NY, USA, 1982.
[13] F. Proschan, “Theoretical explanation of observed decreasing failure rate,” *Technometrics*, vol. 5, pp. 375–383, 1963.