CLASSICAL AND BAYES-P* SUBSET SELECTION PROCEDURES
FOR DOUBLE EXPONENTIAL POPULATIONS*

by
Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, IN, USA

and
Yuning Liao
Department of Statistics
Purdue University
West Lafayette, IN, USA

Technical Report # 90–25C

Department of Statistics
Purdue University
May 1990

*Research supported in part by the Office of Naval Research Contract N00014–88–K–0170 and NSF Grant DMS–8702620 at Purdue University.
CLASSICAL and BAYES-P* SUBSET SELECTION PROCEDURES
for DOUBLE EXPONENTIAL POPULATIONS*

Shanti S. Gupta and Yuning Liao
Department of Statistics, Purdue University, West Lafayette, Indiana 47907, U.S.A.

Abstract

The exact distribution of the sample mean from a double exponential (Laplace) model is derived. A classical subset selection procedure based on the sample mean for selecting the population associated with the largest location parameter of \( k \) double exponential (Laplace) distributions is studied. For the case when a non-informative prior is introduced into the problem, the relation between the classical Maximum-Type Procedure Rule \( R^{\text{max}} \) and the so-called Bayes-P* subset selection procedure rule is studied. An improved bound for the guarantee probability of a correct selection for the classical subset selection rule \( R^{\text{max}} \) that relates the rule \( R^{\text{max}} \) to the selected subset size (notice that the subset selection rule \( R^{\text{max}} \) may select all the populations) is studied and some improved rules of the type \( R^{\text{max}} \) are provided.

1 Introduction

Suppose we have \( k \) double exponential populations \( \Pi_1, \Pi_2, \cdots, \Pi_k \), where each \( \Pi_i \) is characterized by the location parameter \( \theta_i, i = 1, 2, \cdots, k \). The parameters \( \theta_1, \theta_2, \cdots, \theta_k \) are assumed to be unknown. Let \( X_i \) be the observable random variable from \( \Pi_i \) with probability density function

\[
f(x; \theta_i, \sigma) = \frac{1}{2\sigma} \exp\left(\frac{-(|x - \theta_i|)}{\sigma}\right), \quad -\infty < x, \theta_i < \infty, \sigma > 0,
\]

*Research supported in part by the Office of Naval Research Contract N00014-88-K-0170 and NSF Grant DMS-8702620 at Purdue University.
where \( \sigma \), is a common known value for all \( i = 1, 2, \cdots, k \), so that without loss of generality, we can assume that \( \sigma = 1 \). The ranked parameters are denoted by \( \theta_{[1]} \leq \theta_{[2]} \leq \cdots \leq \theta_{[k]} \), and it is assumed that the correct pairing of the ordered \( \theta_{[i]} \)'s and the unordered \( \theta_i \)'s is unknown.

In this paper, we are mainly interested in the subset selection procedures. First, we assume that there is no prior information about the parameters. Then we study the case where \( \theta_i \)'s are independently distributed, and each \( \theta_i \) has a non-informative prior.

### 2 Distribution of the Sample Mean

In connection with the selection procedures based on the sample means, we first derive the distribution of the sample mean.

Let \( X_{ij} \) be a random sample from \( i \)th population \( i = 1, 2, \cdots, k, \ j = 1, 2, \cdots, n \), i.e.

\[
X_{ij} \sim f(x|\theta_i, 1) = \frac{1}{2} \exp\{-|x - \theta_i|\}.
\]

Hence

\[
U_{ij} = X_{ij} - \theta_i \sim f(x|0, 1) = \frac{1}{2} \exp\{-|x|\}. \tag{2}
\]

From the characteristic function of \( U_i = \sum U_{ij} \), we can derive the following lemma

**Lemma 2.1** (Weida (1935)) Suppose \( U_i = \sum_{j=1}^{n} U_{ij} \), where \( U_{ij} \) has density (2), then the density function of \( U_i \) is given by following formula

\[
p(u) = \frac{1}{2\pi} 2\pi i (-1)^{n-1} \frac{1}{(n-1)!} \frac{d^{n-1}}{ds^{n-1}} \left( \frac{e^{-itu}}{(1+it)^n} \right)|_{s=-1}, \tag{3}
\]

where \( u > 0 \) and

\[p(u) = p(-u) \quad \text{for} \quad u \leq 0. \]

Let \( s = -it \), then (3) becomes

\[
p(u) = \frac{1}{(n-1)!} \frac{d^{n-1}}{ds^{n-1}} \left( \frac{e^{su}}{(1-s)^n} \right)|_{s=-1}
= \frac{e^{-u}}{2^n(n-1)!} \sum_{j=1}^{n} c_{n,n-j} u^{n-j}
= e^{-u} \sum_{j=1}^{n} \frac{c_{n,n-j} u^{n-j}}{c_n}, \tag{4}
\]

\[2\]


where
\[ c_n = 2^n(n-1)!, \]
\[ c_{n,n-j} = \frac{(n+j-2)!}{(j-1)!(n-j)!2^{j-1}}, \quad j = 1, 2, \ldots, n. \]  

Therefore the density function of \( X_i = \sum_{j=1}^{n} X_{ij}/n \) is
\[ f_n(x) = e^{-n|x-\theta_i|} \sum_{j=1}^{n} \alpha_{n,n-j} x^{n-j}, \quad -\infty < x < \infty, \]
where \( \alpha_{n,n-j} = n^{n-j+1} c_{n,n-j}/c_n, \quad j = 1, 2, \ldots, n. \)

To obtain the coefficients \( \{c_{n,i}\} \), \( i = 0, 1, 2, \ldots, n-1, \) \( n = 2, 3, \ldots, \) it is helpful to rewrite the formula (5) as
\[ c_{n,i} = \frac{(2n-i-2)!}{(n-i-1)!i!2^{n-i-1}}. \]

Note that
\[ c_{n,n-1} = 1, \quad c_{n,i} = \frac{(2n-i-2)(2n-i-3)}{2(n-i-1)} c_{n-1,i}. \]

In particular
\[ c_{n,0} = c_{n,1}, \quad c_{n,1} = (2n-3)c_{n-1,1}. \]

In Table 1, we have provided the values of \( \{c_n\} \) and \( \{c_{n,i}\} \) for \( n = 2(1)10; \) \( i = 1(1)n-1. \)

To find the cdf of \( X_i \), let us first find the cdf of \( U_i \). Integrating the density function (4) of \( U_i \), we have
\[ P(u) = \int_{-\infty}^{u} p(t)dt \quad (u > 0) \]
\[ = 1 - e^u \sum_{j=1}^{n} \frac{a_{n,n-j} u^{n-j}}{c_n}, \]
where \( \{a_{n,n-j}\} \) satisfy:
\[ a_{n,n-j} = a_{n,n-j} + (n-j+1)a_{n,n-j+1} \quad j = 1, 2, \ldots, n, \quad a_{n,n} = 0. \]

Again we have
\[ a_{n,n-1} = 1, \quad a_{n,n-2} = (n-1)(n+2)/2. \]
Hence the cdf of $X_i$ is given by,

$$
F_n(x|\theta_i) = \begin{cases} 
    e^{-n|x-\theta_i|} \sum_{j=1}^{n} \frac{(n-n^*)^{n-j}}{c_n} |x-\theta_i|^{n-j}, & x < \theta_i \\
    1 - e^{-n|x-\theta_i|} \sum_{j=1}^{n} \frac{(n-n^*)^{n-j}}{c_n} |x-\theta_i|^{n-j}, & \text{otherwise.}
\end{cases} 
$$

(11)

In Table 2, we provide the values of $\{c_n\}$ and $\{a_{n,i}\}$ for $n = 2(1)10; i = 1(1)n - 1$.

Example: If we want to obtain the density and the cumulative distribution function of the sample mean of size $n=4$ from a double exponential model, checking the column $n = 4$ from both Table 1 and Table 2, we can easily see that

$$
f_4(|x-\theta_i|) = \frac{1}{96} e^{-4|x-\theta_i|} (4^4|x-\theta_i|^3 + 6 \times 4^3|x-\theta_i|^2 + 15 \times 4^2|x-\theta_i| + 15 \times 4),
$$

and

$$
F_4(x|\theta_i) = \begin{cases} 
    \frac{1}{96} e^{-4|x-\theta_i|} (4^3|x-\theta_i|^3 + 9 \times 4^2|x-\theta_i|^2 + 33 \times 4|x-\theta_i|^1 + 48), & x < \theta_i \\
    1 - \frac{1}{96} e^{-4|x-\theta_i|} (4^3|x-\theta_i|^3 + 9 \times 4^2|x-\theta_i|^2 + 33 \times 4|x-\theta_i|^1 + 48), & \text{otherwise.}
\end{cases}
$$

To compare the percentage points of the sample mean and the sample median, let

$$
Z_n = \frac{\bar{X}_n - \theta}{\sigma} \quad \text{and} \quad Z^*_n = \frac{X_{(\frac{n}{2})} - \theta}{\sigma}.
$$

Since the cdf of $Z^*_n$ for odd number $n$ is much easier to derive (see Gupta and Leong 1979), we will only provide the comparison of the percentage points of $Z_n$ and $Z^*_n$ for $n = 3, 5, \ldots, 21$(Table 3). The percentage points for the distribution of the sample mean $Z_n$ when $n = 2, 4, \ldots, 20$ are provided in a separate table(Table 4).

3 Using the Sample Mean to Select the Largest Location Parameter

If we assume that no prior information about the parameter $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$ is available, then we usually will use either the classical subset selection approach or the indifference zone formulation in our ranking and selection problem. In the following, we only study the subset selection approach.

(A) Formulation of the Problem: The classical Maximum-Type Approach for any location type problem have been well studied, so we would not give too many details, but simply state some interesting results without any proof.
For selecting the population associated with the largest location parameter with a correct selec-
tion (CS) probability at least \( P^*(1/k < P^* < 1) \) from \( k \) double exponential populations, where we
have a sample mean \( \bar{X}_i \) of size \( n \) from each \( \Pi_i, \ i = 1, 2, \ldots, k \), the Classical Maximum-Type Subset
Selection Rule \( (R^{\text{max}}) \) proposed by Gupta (1956) is defined as follows:

\[
R^{\text{max}}: \text{ Select } \Pi_i, \text{ iff: } \bar{X}_i \geq \max_{1 \leq j \leq k} \bar{X}_j - d/\sqrt{n} \text{ for some } d(>0),
\]

where \( d(>0) \) is the smallest value satisfying:

\[
\int_{-\infty}^{\infty} F_{n}^{k-1}(u + d/\sqrt{n}) f_n(u) du \geq P^*.
\]

The usual condition of \( P(CS|R^{\text{max}}) \geq P^* \) is guaranteed by the following theorem:

**Theorem 3.1**

\[
\inf_{\theta \in \Omega} P_{\theta}(CS|R^{\text{max}}) = \inf_{\theta \in \Omega_0} P_{\theta}(CS|R^{\text{max}}) = \int_{-\infty}^{\infty} F_n^{k-1}(u + d/\sqrt{n}) f_n(u) du,
\]

where \( \Omega \supset \Omega_0 = \{ \theta : \theta_1 = \theta_2 = \cdots = \theta_k, -\infty < \theta_i < \infty, i = 1, 2, \ldots, k \} \).

\( \text{(B) Table of Necessary Constants For } R^{\text{max}}: \) for given \( k, n, \) and some particular values of
\( P^* \), the constants \( d/\sqrt{n} = d(k, n, P^*) \) which satisfy

\[
P^* = \int_{-\infty}^{\infty} F_n^{k-1}(u + d/\sqrt{n}) f_n(u) du,
\]

are given in Table 5.

\( \text{(C) Asymptotic Results for the Procedure } R^{\text{max}}: \) For large \( n \), we can certainly use the normal
distribution to approximate the infimum of \( P_{\theta}(CS|R^{\text{max}}) \). Since

\[
\inf_{\theta \in \Omega} P_{\theta}(CS|R^{\text{max}}) = \inf_{\theta \in \Omega_0} P_{\theta}(CS|R^{\text{max}}),
\]

it suffices to consider the case where \( \theta \in \Omega_0 \), now we have

\[
\frac{\bar{X}_n - \theta}{\sigma_n} \rightarrow N(0,1),
\]

where \( \sigma^2_n = 2/n \), so the probability of the following event

\[
\bar{X}_k \geq \max_{1 \leq j \leq k} \bar{X}_j - d/\sqrt{n},
\]
is, asymptotically, the same as that of

\[ Z_k \geq \max_{1 \leq j \leq k} Z_j - d/\sqrt{2}, \]

where \( Z_j, j = 1, 2, \cdots, k \) are i.i.d. standard normal variables, thus

\[ \inf_{g \in \U} P_g(CS|R_{\max}) \approx P_g(Z_k \geq \max_{1 \leq j \leq k} Z_j - d/\sqrt{2}) = \int_{-\infty}^{\infty} \Phi^{k-1}(u + d/\sqrt{2}) d\Phi(u). \]  \hfill (12)

On the other hand, if we use the sample median in the selection procedure, we will have, asymptotically,

\[ \inf_{g \in \U} P_g(CS|R_{\max}) \approx P_g(Z_k \geq \max_{1 \leq j \leq k} Z_j - d_{\text{median}}) = \int_{-\infty}^{\infty} \Phi^{k-1}(u + d_{\text{median}}) d\Phi(u). \]

Thus, in order to have the same probability of a correct selection for both selection rules based on the different statistics, we must have, for large \( n \),

\[ d \approx \sqrt{2} d_{\text{median}}. \]  \hfill (13)

(D) Sensitivity of the Assumption of Double Exponential: Suppose we have \( k \) populations \( \Pi_1, \Pi_2, \cdots, \Pi_k \), where \( \Pi_i \) is characterized by a location parameter \( \theta_i \). If we do not know whether these \( k \) populations have normal, logistic, or double exponential distributions, then selecting the population associated with the largest location parameter becomes a problem, because the real distribution of the populations is unknown. We will show that the double exponential distribution model provides a safeguard as explained below.

If the sample size \( n \) is large, we know that the infimum of \( P_g(CS|R_{\max}) \) for the double exponential populations is approximately given by (12). On the other hand, for the normal means problem, we have

\[ \inf_{g \in \U} P_g(CS|R_{\max}^{N}) = \int_{-\infty}^{\infty} \Phi^{k-1}(u + d_N) d\Phi(u), \]

because

\[ Z_k \geq \max_{1 \leq j \leq k} Z_j - d_N/\sqrt{n} \iff \sqrt{n}(Z_k - \theta) \geq \max_{1 \leq j \leq k} \sqrt{n}(Z_j - \theta) - d_N, \]
and $\sqrt{n}(\bar{Z}_j - \theta) \sim N(0,1)$. Similarly, for the logistic distribution model, we have

$$\inf_{\text{gen}} P_\theta(\text{CS} | R^\text{max}_L) \simeq \int_{-\infty}^{\infty} \Phi^{k-1}(u + d_L) d\Phi(u),$$

therefore,

$$d \simeq \sqrt{2}d_N \simeq \sqrt{2}d_L.$$

It is clear from the above that the $d$-values for the double exponential provide conservative bounds for the other two models, if $n$ is large.

When $n$ is small, for instance, for $n = 10$, $k = 10$, we have the following:

<table>
<thead>
<tr>
<th></th>
<th>$P^*$-value</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.75</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td>$d_L$</td>
<td>2.2639</td>
<td>2.9925</td>
<td>3.4390</td>
</tr>
<tr>
<td>$d_N$</td>
<td>2.2637</td>
<td>2.9829</td>
<td>3.4182</td>
</tr>
</tbody>
</table>

$d_N$-value excerpted from Bechhofer(1954)

$d_L$-value excerpted from Han(1987 Ph.D. Thesis)

From this we again see that the $d$-values for the double exponential provide conservative bounds for the normal and logistic models for the problem of selecting the unknown location parameter.

### 4 Selection Using a Non-informative Prior

In the Classical Maximum-Type Subset Selection Procedure, it is easy to notice that the selected subset size $|s|$ is a random variable which is not fixed in advance.

In general, for any location or scale parameter situation, Gupta(1965) proved that:

1. The procedure of the above type is monotone, and
2. If the distribution $F(x, \theta)$ possesses a density $f(x, \theta)$ having a monotone likelihood ratio (MLR) in $x$, then $E(|s|)$ is maximized when $\theta_1 = \theta_2 = \cdots = \theta_k$ and the maximum is $kP^*$.

So, in the worst case, the expected proportion in the selected subset is equal to $P^*$. Furthermore, it may select populations such that, depending on the unknown parameter $\theta$, we may get an actual $P(CS)$ much larger than $P^*$. 


In this section, we will regard the likelihood function of $\theta_i$ as the distribution of $\theta_i$ given $X$. It is the same as saying that based on the distribution of a statistic (in our case it is the sample mean and the sample median), we assume that, independently, each $\Theta_i$ has a non-informative prior, $i = 1, 2, \ldots, k$.

4.1 Bayes Selection Procedure

In the following, we will consider a more general case, we assume

$$X_i \sim f(|x - \theta_i|),$$

i.e. the density of $X_i$ given $\Theta_i = \theta_i$ is symmetric about $\theta_i$ (for the case where $f(.)$ is not symmetrical, we have obtained some results which will be available later), and

$$\Theta_i \sim \Pi(\theta) = 1, \quad i = 1, 2, \ldots, k.$$

Now, we will make decisions based on the posterior distributions of $\Theta|X$.

From a Bayes perspective, in order to select the population associated with the largest parameter $\theta_{[k]}$ with a guaranteed posterior probability of a correct selection to be at least $P^*(1/k < P^* < 1)$ (the so-called $PP^*$-condition, see Gupta and Yang(1985)), we should consider the following events

$$A_i = \{i \text{ is the largest } |X = x\}, \quad i = 1, 2, \ldots, k.$$

Now, using the non-informative prior, we have

$$\Theta_i|X = x \sim f(|x_i - \theta_i|), \quad i = 1, 2, \ldots, k.$$

Let $p_i(x)$ be the probability of event $A_i$, then

$$p_i(x) = P(\theta_i \text{ is the largest } |x) = P(\theta_i > \theta_j, \forall j, j \neq i |x) = P(\theta_i - x_i > \theta_j - x_j - (x_i - x_j), \forall j, j \neq i |x) = \int_{-\infty}^{\infty} \prod_{j \neq i} F(u + (x_i - x_j))f(u)du,$$

where $F(.)$ is the cdf of $f(.)$. 

8
Lemma 4.1: (1) The posterior probability $p_i(z)$ depends only on the differences $x_i - x_j, \ i, j = 1, 2, \ldots, k, \ j \neq i$.

(2) $p_i(z)$ is non-increasing in $x_j, \ j \neq i$, keeping other components of $z$ fixed and it is non-decreasing in $x_i$, keeping other components of $z$ fixed.

(3) $p_i(z) \geq p_j(z)$ if and only if $x_i \geq x_j$.

Proof: The proof is straightforward and hence omitted. □

Theorem 4.1: For any subset $S$ of the whole populations $\Pi_1, \Pi_2, \ldots, \Pi_k$, let $PP(CS|S, z)$ denote the posterior probability of a correct selection for the subset $S$ (i.e. the subset $S$ contains the best population) based on a random sample $z$, then

(1) $PP(CS|S, z)$ is non-increasing in $x_j, \ j \notin S$, keeping other components of $z$ fixed, and

(2) $PP(CS|S, z)$ is non-decreasing in $x_i, \ i \in S$, keeping other components of $z$ fixed.

Proof: Since

$$PP(CS|S, z) = \sum_{i \in S} p_i(z) = 1 - \sum_{i \notin S} p_i(z).$$

Now, $p_i(z)$ is non-increasing in $x_j, \ j \notin S$ for all $i \in S$, so $PP(CS|S, z)$ is non-increasing in $x_j, \ j \notin S$ by first part of equation (14).

On the other hand, the second part of equation (14) and the fact that $p_j(z)$ is non-increasing in $x_i, \ i \in S$ for all $j \notin S$ imply that $PP(CS|S, z)$ is non-decreasing in $x_i, \ i \in S$. □

From the Bayesian analysis, we know that the Bayes Decision Rule ($R^B$) will select the $t$ populations which associated with the $t$ largest values of $p_i(z)$ values (i.e. the Bayes set $s^B = \{\Pi[k], \ldots, \Pi[k-t+1]\}$), where the integer $t(\geq 1)$ satisfies

$$\sum_{m=k-t+1}^{k} p_{[m]}(z) \geq P^*,$$

and

$$\sum_{m=k-t+2}^{k} p_{[m]}(z) < P^*,$$

where $p_{[1]}(z) \leq p_{[2]}(z) \leq \cdots \leq p_{[k]}(z)$ are the ordered values of $p_i(z)$'s, and $s^B$ is the subset selected by the Bayes selection rule $R^B$. 

9
4.2 A Lower Bound on the \(PP(CS)\) for the Subset Selection Rule \(R^{\text{max}}\)

Under the Maximum-Type Subset Selection Rule \(R^{\text{max}}\) defined in the previous section, we know that the larger the value \(z_i\) is, the larger the chance that the corresponding population \(\Pi_i\) will be selected.

Under the rule \(R^{\text{max}}\), we will pick the population \(\Pi_i\) if its \(z_i\) value is larger than \(z_{[k]} - d\), and reject \(\Pi_i\) if \(z_i < z_{[k]} - d\). Thus the following observations:

**Observations:** For the Maximum-Type Subset Selection Rule \(R^{\text{max}}\), we know at least the following two facts

1. \(R^{\text{max}}\) will always pick population \(\Pi_{[k]}\), i.e. the population associated with the largest value \(z_{[k]}\).
2. All \(\Pi_i\) not being selected by \(R^{\text{max}}\) must has its \(z_i\) value less than \(z_{[k]} - d\).

**Theorem 4.2:** If the subset selection rule \(R^{\text{max}}\) selects \(i\) populations (i.e. select population \(\Pi_{[k]}, \cdots, \Pi_{[k-i+1]}\), where \(\Pi_{[j]}\) is the population associated with the \(j\)th largest value \(z_{[j]}\)), under the classical selection procedure, then

\[
PP(CS|R^{\text{max}}, z) \geq PP(CS|R^{\text{max}}, z \in \mathcal{X}_0) = P^* + \frac{i-1}{k-1}(1 - P^*),
\]

where \(\mathcal{X}_0 = \{z : z_{[k]} - d = z_{[k-1]} = \cdots = z_{[1]}\}\).

**Remark 4.2:** A similar result for the normal model has been given in Gupta and Yang (1985).

Here, we will give a probabilistic proof of the above theorem.

**Proof:** The first part of the inequality [i.e. \(PP(CS|R^{\text{max}}, z) \geq PP(CS|R^{\text{max}}, z \in \mathcal{X}_0)\)] follows from the above observations and Theorem 4.1.

When \(z \in \mathcal{X}_0\), we know that \(p_{[k]}(z) > p_{[k-1]}(z) = \cdots = p_{[1]}(z)\), and

\[
p_{[k]}(z) = \int_{-\infty}^{\infty} \prod_{j \neq k} F(u + (z_{[k]} - z_{[j]}))f(u)du \\
= \int_{-\infty}^{\infty} \prod_{j \neq k} F(u + d)f(u)du \\
= \int_{-\infty}^{\infty} F^{k-1}(u + d)f(u)du = P^*,
\]

since \(\sum p_i(z) = 1\), so \(p_{[1]}(z) = p_{[2]}(z) = \cdots = p_{[k-1]}(z) = \frac{1}{k-1}(1 - P^*)\) and \(|s_{R^{\text{max}}}| = i\), hence the result follows. \(\square\)
Since $PP(CS|R_{\text{max}}, \mathcal{X}) \geq P^*$ and it is strictly larger than $P^*$, once we pick more than one population, we certainly can find a better subset selection rule by simply utilizing the lower bound on $PP(CS|R_{\text{max}}, \mathcal{X})$.

4.3 Some New Selection Procedures

First, let us consider the following selection procedure:

Let $\Delta x_{[i]} = x_{[k]} - x_{[k-i]}$ for $i = 1, 2, \ldots, k - 1$, where $x_{[1]} \leq x_{[2]} \leq \cdots \leq x_{[k]}$ are the ordered values of $x_{i}$'s. then, we compute the following $k - 1$ numbers:

$$P^*_{(1)} = \int_{-\infty}^{+\infty} F_{k-1}(u + \Delta x_{[i]})dF(u).$$

(16)

Since $0 \leq \Delta x_{[1]} \leq \cdots \leq \Delta x_{[k-1]}$, therefore

$$0 \leq P^*_{(1)} \leq P^*_{(2)} \leq \cdots \leq P^*_{(k-1)}(< 1).$$

Next, we compute:

$$Q^*_{(i)} = P^*_{(i)} + \frac{i-1}{k-1}(1 - P^*_{(i)}).$$

(17)

Lemma 4.2: For values of $\Delta x_{[i]}$, where $0 \leq \Delta x_{[1]} \leq \cdots \leq \Delta x_{[k-1]}$, we have

$$0 \leq Q^*_{(1)} \leq \cdots \leq Q^*_{(k-1)}(< 1).$$

(18)

Proof: Actually, we have

$$Q^*_{(i)} = 1 - \frac{k-i}{k-1}(1 - P^*_{(i)}),$$

so $Q^*_{(i)}$ is increasing in $i$, because $k-i$ is decreasing in $i$ and $1 - P^*_{(i)}$ is decreasing in $\Delta x_{[i]}$ (thus in $i$); hence the result.

Now, we propose the following subset selection rule $R_1$:

For any preassigned guarantee probability $P^*(1/k < P^* < 1)$, if there exists the smallest $Q^*_{(i_0)}$ which satisfies $Q^*_{(i_0)} \geq P^*$, then the subset selection rule $R_1$ is

$$R_1: \text{Select } \Pi(j) \text{ iff } j > i_0.$$  

(19)

The subset selection rule $R_1$ will take $s = \{\Pi(k), \ldots, \Pi(k-i_0+1)\}$ as our selected subset.
otherwise, $R_1$ will select all populations.

**Remark 4.3:** To implement the procedure $R_1$, we examine the posterior probabilities at following $k - 1$ stages:

Stage 1. pull all $k - 2$ values of $x_{[i]}$, $i = 1, 2, \cdots, k - 2$ to the point $x_{[k-1]}$, and check if

$$\frac{k - 1}{k - 1}(1 - P_{(1)}^*) \leq 1 - P^*,$$

if the above holds, we select $s = \{\Pi_{[k]}\}$ and terminate the process, if not, we go to

Stage 2. pull all $k - 2$ values of $x_{[i]}$, $i \neq k - 2$, $i = 1, 2, \cdots, k - 1$ to the point $x_{[k-2]}$, check if

$$\frac{k - 2}{k - 1}(1 - P_{(2)}^*) \leq 1 - P^*,$$

if it holds, we select $s = \{\Pi_{[k]}, \Pi_{[k-1]}\}$ and terminate the process, if not, we go to Stage 3, and so on, until we can find an $i$ such that

$$\frac{k - i}{k - 1}(1 - P_{(i)}^*) \leq 1 - P^*,$$

and then we select $s = \{\Pi_{[k]}, \Pi_{[k-1]}, \cdots, \Pi_{[k-i+1]}\}$; If there does not exist such an $i$, we select all populations.

For other subset selection rules $R_2, \cdots, R_{k-1}$, we give the following remark:

**Remark 4.4:** Note that in the process of deriving the subset selection rule $R_1$, we divided the data into two groups, and put only one value (i.e. $x_{[k]}$) into the first group. Now we can develop it in two directions.

(a) By putting more $x_{[i]}$’s into the first group, we can actually replace $Q_{(i)}^*$ by $Q_{(i)}^{**}$ as follows:

$$Q_{(i)}^{**} = \max_{0 \leq m \leq i-1} \left(1 - \frac{k - i}{k - m - 1} P_{(i,m)}^{**}\right),$$

where

$$P_{(i,m)}^{**} = \int_{-\infty}^{+\infty} F_{m+1}^*(u - (x_{[k-m]} - x_{[k-1]}))dF_{k-m-1}^*(u), \quad m = 0, 1, \cdots, i - 1,$$

is the posterior probability of $p_{[1]}(z) = \cdots = p_{[k-m-1]}(z)$, when we pull $x_{[k]}, \cdots, x_{[k-m+1]}$ to $x_{[k-m]}$ and $x_{[k-m-1]}, \cdots, x_{[1]}$ to $x_{[k-i]}$.

When $m = 0$, we have

$$P_{(i,0)}^{**} = 1 - P_{(i)}^*,$$
which is the value we used in the rule $R_1$.

(b) We can also divide the data into $3, 4, \ldots, k$ groups. Let $R_2$ be the rule for the case of 3 groups, $\cdots$, and $R_{k-1}$ be the rule for the case of $k$ groups. Then in the case of $k$ groups, the subset selection rule and the previous rule $R^B$ are identical. Later, it will be shown that $R_2$ can be as good as $R^B$ and it is easier to implement from the computational viewpoint.

4.4 Properties of Subset Selection Rule

We can easily prove the following:

Proposition 4.1: The subset selection rule $R_1$ is better than rule $R^\text{max}$, in the sense that

(a) $PP(CS|R_1) \geq P^*$, because $PP(CS|R_1) \geq Q^*_{(i_0)}$, and

(b) $s_1 \subseteq s_{R^\text{max}}$, because $P^*_{(i_0)} \leq P^*$.

Proposition 4.2: (a) The subset selection rule $R_1$ and $R^B$ will take the same action, if $x_{[1]} = \cdots = x_{[k-1]} < x_{[k]}$, or when the subset selection rule $R_1$ selects all populations in its selected subset.

(b) The subset selection rule $R_1$, $R^B$ and $R^\text{max}$ will take the same action, if the subset selection rule $R_1$ selects only one population.

Proposition 4.3: The subset selection rule $R_1$ possesses the advantage of the rule $R^\text{max}$, because the forms of the involved integration for $P^*_{(i)}$ and $P^*$ are identical.

Remark 4.5: The selection rule $R_1$ is like a modified rule of $R^B$, where, it like that the population associated with the largest statistic possesses the probability $P^*$ of a correct selection, and the remaining $|s| - 1$ populations in $s$ have the $P(CS)$ at least equal to $\frac{|s| - 1}{k-1}(1 - P^*)$.

5 An Example for Comparison of the Several Subset Selection Rules

A data set of exponential random numbers generated by a statistical package G6-RVP designed by H.Rubin and C.Hinkle at Purdue University was given in Gupta and Leong's paper (1979), where 9 observations for each of 5 sets of double exponential random numbers with location parameters $\theta_i$ equal to 0, 2.5, 3.4, -2.0, -0.65 were taken.
To see how each subset selection rule performs, let

\[ x_i = \text{the sample mean of } \Pi_i \quad \text{and} \quad y_i = \text{sample median of } \Pi_i, \]

then

\[ \mathbf{z} = (x_1, \ldots, x_5)' = (0.0980, 2.5980, 3.4980, -1.9020, -0.5520)', \]

\[ \mathbf{y} = (y_1, \ldots, y_5)' = (-0.1761, 2.3239, 3.2239, -2.1761, -0.8261)' \]

Hence the difference of \( x_i \)'s and \( y_i \)'s are \( \Delta x_{32} = \Delta y_{32} = 0.90, \Delta x_{31} = \Delta y_{31} = 3.40, \Delta x_{35} = \Delta y_{35} = 4.05, \Delta x_{34} = \Delta y_{34} = 5.40. \)

(a) Now, we have the following:

\[ PP(\mathcal{CS}|R, \mathbf{z}) \text{ for } R = R^B, R_1, R_i(i \geq 2) \]

when one population is picked

<table>
<thead>
<tr>
<th>( R^B ) or ( R_i(i \geq 2) )</th>
<th>using mean</th>
<th>using median</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathcal{R}^{\text{max}} ) or ( R_1 )</td>
<td>0.9131</td>
<td>0.9380</td>
</tr>
</tbody>
</table>

where, in the case of the sample mean, the integration for \( R^B \) is

\[ P^* = \int_{-\infty}^{\infty} F_9(u+0.9) \times F_9(u+3.4) \times F_9(u+4.05) \times F_9(u+5.4) dF_9(u). \]

The integration for \( R_2 \) is

\[ P^* = \int_{-\infty}^{\infty} F_9(u+0.9) \times F_9^2(u+3.4) dF_9(u). \]
Also, the integration for $R_1$ or $R_{\text{max}}^*$ is

$$P^* = \int_{-\infty}^{\infty} F_9^*(u + 0.9)dF_9(u),$$

where $F_9(.)$ is the cdf of the sample mean of size 9.

The same applies to the case of the sample median. Note that the rule $R_2$ is as good as $R_B$.

(b) In the case where two populations are taken, we have the probability one for all selection rules, because

$$\int_{-\infty}^{\infty} F_9^*(u + 3.4)dF_9(u) \simeq \int_{-\infty}^{\infty} G_4^*(u + 3.4)dG_4(u) \simeq 1,$$

where $G_4(.)$ is the cdf of the sample median of size 9 (see Gupta and Leong (1979)).
Table 1: Table of \( \{c_n\} \) and \( \{c_{n,i}\} \) for \( n = 2, 3, \ldots, 10 \)

<table>
<thead>
<tr>
<th>( c_{n,i} )</th>
<th>( n )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_n )</td>
<td></td>
<td>4</td>
<td>16</td>
<td>96</td>
<td>768</td>
<td>7680</td>
<td>92160</td>
<td>1290240</td>
<td>20643840</td>
<td>371589120</td>
</tr>
<tr>
<td>( c_{n,0} )</td>
<td></td>
<td>1</td>
<td>3</td>
<td>15</td>
<td>105</td>
<td>945</td>
<td>10395</td>
<td>135135</td>
<td>2027025</td>
<td>34459425</td>
</tr>
<tr>
<td>( c_{n,1} )</td>
<td></td>
<td>1</td>
<td>3</td>
<td>15</td>
<td>105</td>
<td>945</td>
<td>10395</td>
<td>135135</td>
<td>2027025</td>
<td>34459425</td>
</tr>
<tr>
<td>( c_{n,2} )</td>
<td></td>
<td>1</td>
<td>6</td>
<td>45</td>
<td>420</td>
<td>4725</td>
<td>62370</td>
<td>945945</td>
<td>16216200</td>
<td></td>
</tr>
<tr>
<td>( c_{n,3} )</td>
<td></td>
<td>1</td>
<td>10</td>
<td>105</td>
<td>1260</td>
<td>17325</td>
<td>270270</td>
<td>4729725</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( c_{n,4} )</td>
<td></td>
<td>1</td>
<td>15</td>
<td>210</td>
<td>3150</td>
<td>3150</td>
<td>3150</td>
<td>51975</td>
<td>945945</td>
<td></td>
</tr>
<tr>
<td>( c_{n,5} )</td>
<td></td>
<td>1</td>
<td>21</td>
<td>378</td>
<td>6930</td>
<td>135135</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( c_{n,6} )</td>
<td></td>
<td></td>
<td>28</td>
<td>630</td>
<td>13860</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( c_{n,7} )</td>
<td></td>
<td>1</td>
<td>36</td>
<td>990</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( c_{n,8} )</td>
<td></td>
<td></td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( c_{n,9} )</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Table of \( \{c_n\} \) and \( \{a_{n,i}\} \) for \( n = 2,3,\ldots,10 \)

<table>
<thead>
<tr>
<th>( a_{n,i} )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_n )</td>
<td>4</td>
<td>16</td>
<td>96</td>
<td>768</td>
<td>7680</td>
<td>92160</td>
<td>1290240</td>
<td>20643840</td>
<td>371589120</td>
</tr>
<tr>
<td>( a_{n,0} )</td>
<td>2</td>
<td>8</td>
<td>48</td>
<td>384</td>
<td>3840</td>
<td>46080</td>
<td>645120</td>
<td>10321920</td>
<td>185794560</td>
</tr>
<tr>
<td>( a_{n,1} )</td>
<td>1</td>
<td>5</td>
<td>33</td>
<td>279</td>
<td>2895</td>
<td>35685</td>
<td>509985</td>
<td>8294895</td>
<td>151335135</td>
</tr>
<tr>
<td>( a_{n,2} )</td>
<td>1</td>
<td>9</td>
<td>87</td>
<td>975</td>
<td>12645</td>
<td>187425</td>
<td>3133935</td>
<td>58437855</td>
<td></td>
</tr>
<tr>
<td>( a_{n,3} )</td>
<td>1</td>
<td>14</td>
<td>185</td>
<td>2640</td>
<td>41685</td>
<td>6090</td>
<td>114765</td>
<td>2336040</td>
<td></td>
</tr>
<tr>
<td>( a_{n,4} )</td>
<td>1</td>
<td>20</td>
<td>345</td>
<td>6090</td>
<td>114765</td>
<td>12558</td>
<td>278019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( a_{n,5} )</td>
<td>1</td>
<td>27</td>
<td>588</td>
<td>12558</td>
<td>278019</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( a_{n,6} )</td>
<td>1</td>
<td>35</td>
<td>938</td>
<td>23814</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( a_{n,7} )</td>
<td>1</td>
<td>44</td>
<td>1422</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( a_{n,8} )</td>
<td>1</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( a_{n,9} )</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: 1) Upper 100(1 - \( \alpha \)) Percentage Points \( \xi_\alpha \) of \( Z_n \)(Top Entry); 2) \( \Delta_\alpha = \xi_\alpha^* - \xi_\alpha \), where \( \xi_\alpha^* \) is the Upper 100(1 - \( \alpha \)) Percentage Points of \( Z_n^* \)(Bottom Entry).

<table>
<thead>
<tr>
<th>( 1 - \alpha )</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>11</th>
<th>13</th>
<th>15</th>
<th>17</th>
<th>19</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.750</td>
<td>0.6050</td>
<td>0.6321</td>
<td>0.6440</td>
<td>0.6507</td>
<td>0.6550</td>
<td>0.6580</td>
<td>0.6602</td>
<td>0.6619</td>
<td>0.6632</td>
<td>0.6643</td>
</tr>
<tr>
<td>-.0825</td>
<td>-.1102</td>
<td>-.1249</td>
<td>-.1343</td>
<td>-.1409</td>
<td>-.1459</td>
<td>-.1499</td>
<td>-.1531</td>
<td>-.1557</td>
<td>-.1579</td>
<td></td>
</tr>
<tr>
<td>0.900</td>
<td>1.2221</td>
<td>1.2432</td>
<td>1.2532</td>
<td>1.2590</td>
<td>1.2629</td>
<td>1.2656</td>
<td>1.2676</td>
<td>1.2692</td>
<td>1.2704</td>
<td>1.2715</td>
</tr>
<tr>
<td>-.0739</td>
<td>-.1259</td>
<td>-.1591</td>
<td>-.1823</td>
<td>-.1995</td>
<td>-.2129</td>
<td>-.2237</td>
<td>-.2326</td>
<td>-.2401</td>
<td>-.2467</td>
<td></td>
</tr>
<tr>
<td>0.950</td>
<td>1.6372</td>
<td>1.6385</td>
<td>1.6395</td>
<td>1.6402</td>
<td>1.6408</td>
<td>1.6412</td>
<td>1.6416</td>
<td>1.6419</td>
<td>1.6422</td>
<td>1.6424</td>
</tr>
<tr>
<td>-.0369</td>
<td>-.1025</td>
<td>-.1484</td>
<td>-.1815</td>
<td>-.2066</td>
<td>-.2264</td>
<td>-.2426</td>
<td>-.2560</td>
<td>-.2675</td>
<td>-.2774</td>
<td></td>
</tr>
<tr>
<td>0.975</td>
<td>2.0284</td>
<td>2.0026</td>
<td>1.9905</td>
<td>1.9836</td>
<td>1.9794</td>
<td>1.9763</td>
<td>1.9740</td>
<td>1.9724</td>
<td>1.9711</td>
<td>1.9699</td>
</tr>
<tr>
<td>.0144</td>
<td>.0632</td>
<td>-.1210</td>
<td>-.1637</td>
<td>-.1967</td>
<td>-.2229</td>
<td>-.2443</td>
<td>-.2623</td>
<td>-.2778</td>
<td>-.2910</td>
<td></td>
</tr>
<tr>
<td>0.990</td>
<td>2.5214</td>
<td>2.4524</td>
<td>2.4194</td>
<td>2.3999</td>
<td>2.3871</td>
<td>2.3782</td>
<td>2.3715</td>
<td>2.3663</td>
<td>2.3621</td>
<td>2.3590</td>
</tr>
<tr>
<td>.0976</td>
<td>.0055</td>
<td>-.0683</td>
<td>-.1236</td>
<td>-.1666</td>
<td>-.2014</td>
<td>-.2298</td>
<td>-.2539</td>
<td>-.2741</td>
<td>-.2923</td>
<td></td>
</tr>
<tr>
<td>0.995</td>
<td>2.8821</td>
<td>2.7759</td>
<td>2.7246</td>
<td>2.6941</td>
<td>2.6746</td>
<td>2.6599</td>
<td>2.6495</td>
<td>2.6410</td>
<td>2.6343</td>
<td>2.6294</td>
</tr>
<tr>
<td>.1684</td>
<td>.0659</td>
<td>-.0195</td>
<td>-.0842</td>
<td>-.1355</td>
<td>-.1758</td>
<td>-.2099</td>
<td>-.2387</td>
<td>-.2625</td>
<td>-.2844</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Upper 100(1 - \( \alpha \)) Percentage Points \( \xi_\alpha \) of \( Z_n \) for even values of \( n \)

<table>
<thead>
<tr>
<th>( 1 - \alpha )</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
<th>14</th>
<th>16</th>
<th>18</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.750</td>
<td>0.5731</td>
<td>0.6218</td>
<td>0.6390</td>
<td>0.6478</td>
<td>0.6531</td>
<td>0.6566</td>
<td>0.6592</td>
<td>0.6611</td>
<td>0.6626</td>
<td>0.6638</td>
</tr>
<tr>
<td>0.900</td>
<td>1.1986</td>
<td>1.2350</td>
<td>1.2489</td>
<td>1.2564</td>
<td>1.2611</td>
<td>1.2643</td>
<td>1.2667</td>
<td>1.2685</td>
<td>1.2698</td>
<td>1.2710</td>
</tr>
<tr>
<td>0.950</td>
<td>1.6359</td>
<td>1.6379</td>
<td>1.6390</td>
<td>1.6399</td>
<td>1.6405</td>
<td>1.6411</td>
<td>1.6415</td>
<td>1.6418</td>
<td>1.6420</td>
<td>1.6423</td>
</tr>
<tr>
<td>0.975</td>
<td>2.0563</td>
<td>2.0125</td>
<td>1.9955</td>
<td>1.9867</td>
<td>1.9814</td>
<td>1.9777</td>
<td>1.9751</td>
<td>1.9733</td>
<td>1.9717</td>
<td>1.9705</td>
</tr>
<tr>
<td>0.990</td>
<td>2.5858</td>
<td>2.4792</td>
<td>2.4335</td>
<td>2.4084</td>
<td>2.3929</td>
<td>2.3822</td>
<td>2.3746</td>
<td>2.3688</td>
<td>2.3642</td>
<td>2.3605</td>
</tr>
<tr>
<td>0.995</td>
<td>2.9944</td>
<td>2.8174</td>
<td>2.7466</td>
<td>2.7075</td>
<td>2.6831</td>
<td>2.6666</td>
<td>2.6544</td>
<td>2.6453</td>
<td>2.6379</td>
<td>2.6318</td>
</tr>
</tbody>
</table>

18
Table 5: Values of $d/\sqrt{n} = d(n,k,P^*)$ for $n, k = 2, 3, \ldots, 10$

<table>
<thead>
<tr>
<th>n</th>
<th>$P^*$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.75</td>
<td>1.1462</td>
<td>1.7849</td>
<td>2.1575</td>
<td>2.4258</td>
<td>2.6365</td>
<td>2.8104</td>
<td>2.9584</td>
<td>3.0874</td>
<td>3.2615</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>2.3972</td>
<td>3.0504</td>
<td>3.4336</td>
<td>3.7083</td>
<td>3.9234</td>
<td>4.1002</td>
<td>4.2503</td>
<td>4.3809</td>
<td>4.4964</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>3.2716</td>
<td>3.9322</td>
<td>4.3197</td>
<td>4.5971</td>
<td>4.8138</td>
<td>4.9917</td>
<td>5.1427</td>
<td>5.2739</td>
<td>5.3899</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>0.9580</td>
<td>1.3575</td>
<td>1.6011</td>
<td>1.7756</td>
<td>1.9110</td>
<td>2.0214</td>
<td>2.1144</td>
<td>2.1947</td>
<td>2.2653</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>1.7893</td>
<td>2.1966</td>
<td>2.4393</td>
<td>2.6118</td>
<td>2.7452</td>
<td>2.8539</td>
<td>2.9454</td>
<td>3.0244</td>
<td>3.0939</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>2.3470</td>
<td>2.7550</td>
<td>2.9962</td>
<td>3.1670</td>
<td>3.2992</td>
<td>3.4069</td>
<td>3.4975</td>
<td>3.5758</td>
<td>3.6447</td>
</tr>
<tr>
<td>3</td>
<td>0.75</td>
<td>0.7379</td>
<td>1.1187</td>
<td>1.3251</td>
<td>1.4666</td>
<td>1.5740</td>
<td>1.6605</td>
<td>1.7328</td>
<td>1.7949</td>
<td>1.8493</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>1.4421</td>
<td>1.7960</td>
<td>1.9924</td>
<td>2.1281</td>
<td>2.2316</td>
<td>2.3152</td>
<td>2.3853</td>
<td>2.4455</td>
<td>2.4983</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>1.8926</td>
<td>2.2339</td>
<td>2.4250</td>
<td>2.5576</td>
<td>2.6591</td>
<td>2.7410</td>
<td>2.8098</td>
<td>2.8690</td>
<td>2.9209</td>
</tr>
<tr>
<td>4</td>
<td>0.75</td>
<td>0.6478</td>
<td>0.9796</td>
<td>1.1575</td>
<td>1.2784</td>
<td>1.3696</td>
<td>1.4428</td>
<td>1.5037</td>
<td>1.5558</td>
<td>1.6013</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>1.2564</td>
<td>1.5601</td>
<td>1.7268</td>
<td>1.8414</td>
<td>1.9283</td>
<td>1.9983</td>
<td>2.0567</td>
<td>2.1068</td>
<td>2.1507</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>1.6399</td>
<td>1.9298</td>
<td>2.0909</td>
<td>2.2020</td>
<td>2.2866</td>
<td>2.3548</td>
<td>2.4119</td>
<td>2.4608</td>
<td>2.5038</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>2.4086</td>
<td>2.6774</td>
<td>2.8290</td>
<td>2.9344</td>
<td>3.0150</td>
<td>3.0802</td>
<td>3.1351</td>
<td>3.1820</td>
<td>3.2234</td>
</tr>
<tr>
<td>5</td>
<td>0.75</td>
<td>0.5842</td>
<td>0.8821</td>
<td>1.0407</td>
<td>1.1480</td>
<td>1.2286</td>
<td>1.2931</td>
<td>1.3466</td>
<td>1.3923</td>
<td>1.4322</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>1.1280</td>
<td>1.3980</td>
<td>1.5454</td>
<td>1.6462</td>
<td>1.7224</td>
<td>1.7836</td>
<td>1.8346</td>
<td>1.8783</td>
<td>1.9165</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>1.4673</td>
<td>1.7236</td>
<td>1.8650</td>
<td>1.9623</td>
<td>2.0362</td>
<td>2.0956</td>
<td>2.1452</td>
<td>2.1877</td>
<td>2.2249</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>2.1401</td>
<td>2.3752</td>
<td>2.5071</td>
<td>2.5983</td>
<td>2.6682</td>
<td>2.7246</td>
<td>2.7719</td>
<td>2.8121</td>
<td>2.8477</td>
</tr>
</tbody>
</table>
### Table 5 (continued)

<table>
<thead>
<tr>
<th></th>
<th>0.75</th>
<th>0.90</th>
<th>0.95</th>
<th>0.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.5361</td>
<td>0.8088</td>
<td>0.9532</td>
<td>1.0507</td>
</tr>
<tr>
<td></td>
<td>1.1237</td>
<td>1.1819</td>
<td>1.2301</td>
<td>1.2713</td>
</tr>
<tr>
<td></td>
<td>1.3072</td>
<td>1.0015</td>
<td>0.7512</td>
<td>0.7512</td>
</tr>
<tr>
<td>7</td>
<td>0.4983</td>
<td>0.7514</td>
<td>0.8850</td>
<td>0.9748</td>
</tr>
<tr>
<td></td>
<td>1.0420</td>
<td>1.0955</td>
<td>1.1397</td>
<td>1.1775</td>
</tr>
<tr>
<td></td>
<td>1.2103</td>
<td>0.90</td>
<td>0.9575</td>
<td>1.1843</td>
</tr>
<tr>
<td></td>
<td>1.3071</td>
<td>1.3906</td>
<td>1.4535</td>
<td>1.5038</td>
</tr>
<tr>
<td></td>
<td>1.5457</td>
<td>1.5814</td>
<td>1.6126</td>
<td>1.6615</td>
</tr>
<tr>
<td></td>
<td>1.7106</td>
<td>1.7450</td>
<td>1.7450</td>
<td>1.7450</td>
</tr>
<tr>
<td>8</td>
<td>0.4675</td>
<td>0.7045</td>
<td>0.8295</td>
<td>0.9132</td>
</tr>
<tr>
<td></td>
<td>0.9758</td>
<td>1.0255</td>
<td>1.0666</td>
<td>1.1017</td>
</tr>
<tr>
<td></td>
<td>1.1321</td>
<td>0.90</td>
<td>0.8969</td>
<td>1.1086</td>
</tr>
<tr>
<td></td>
<td>1.2230</td>
<td>1.3006</td>
<td>1.3590</td>
<td>1.4057</td>
</tr>
<tr>
<td></td>
<td>1.4444</td>
<td>1.4775</td>
<td>1.5064</td>
<td>1.5254</td>
</tr>
<tr>
<td></td>
<td>1.7140</td>
<td>1.7410</td>
<td>1.7410</td>
<td>1.7410</td>
</tr>
<tr>
<td>9</td>
<td>0.4419</td>
<td>0.6658</td>
<td>0.7835</td>
<td>0.8623</td>
</tr>
<tr>
<td></td>
<td>0.9210</td>
<td>0.9677</td>
<td>1.0063</td>
<td>1.0391</td>
</tr>
<tr>
<td></td>
<td>1.0676</td>
<td>0.90</td>
<td>0.8468</td>
<td>1.0461</td>
</tr>
<tr>
<td></td>
<td>1.1535</td>
<td>1.2263</td>
<td>1.2811</td>
<td>1.3248</td>
</tr>
<tr>
<td></td>
<td>1.3610</td>
<td>1.3920</td>
<td>1.4189</td>
<td>1.4456</td>
</tr>
<tr>
<td></td>
<td>1.6384</td>
<td>1.6123</td>
<td>1.6384</td>
<td>1.6384</td>
</tr>
<tr>
<td>10</td>
<td>0.0015</td>
<td>0.5712</td>
<td>0.7179</td>
<td>0.8051</td>
</tr>
<tr>
<td></td>
<td>0.8665</td>
<td>0.9135</td>
<td>0.9516</td>
<td>0.9835</td>
</tr>
<tr>
<td></td>
<td>1.0110</td>
<td>0.90</td>
<td>0.7512</td>
<td>0.9761</td>
</tr>
<tr>
<td></td>
<td>1.0869</td>
<td>1.1593</td>
<td>1.2126</td>
<td>1.2547</td>
</tr>
<tr>
<td></td>
<td>1.2893</td>
<td>1.3188</td>
<td>1.3443</td>
<td>1.3707</td>
</tr>
<tr>
<td></td>
<td>1.5515</td>
<td>1.5269</td>
<td>1.5515</td>
<td>1.5515</td>
</tr>
<tr>
<td></td>
<td>1.9596</td>
<td>1.9365</td>
<td>1.9596</td>
<td>1.9596</td>
</tr>
</tbody>
</table>

20
References


CLASSICAL AND BAYES-P* SUBSET SELECTION PROCEDURES FOR DOUBLE EXPONENTIAL POPULATIONS

Shanti S. Gupta and Yuning Liao

The exact distribution of the sample mean from a double exponential (Laplace) model is derived. A classical subset selection procedure based on the sample mean for selecting the population associated with the largest location parameter of \( k \) double exponential (Laplace) distributions is studied. For the case when a non-informative prior is introduced into the problem, the relation between the classical Maximum-Type Procedure Rule \( R_{\max} \) and the so-called Bayes-P* subset selection procedure rule is studied. An improved bound for the guarantee probability of a correct selection for the classical subset selection rule \( R_{\max} \) that relates the rule \( R_{\max} \) to the selected subset size (notice that the subset selection rule \( R_{\max} \) may select all the populations) is studied and some improved rules of the type \( R_{\max} \) are provided.