SELECTION OF THE BEST WITH A PRELIMINARY TEST FOR TWO-PARAMETER EXPONENTIAL DISTRIBUTIONS

by

Lii-Yuh Leu
Graduate Institute of Statistics
National Central University
Chung-Li, Taiwan, R.O.C.

and

TaChen Liang
Department of Mathematics
Wayne State University
Detroit, Michigan 48202

Technical Report #89-14

PURDUE UNIVERSITY

DEPARTMENT OF STATISTICS

1

SELECTION OF THE BEST WITH A PRELIMINARY TEST FOR TWO-PARAMETER EXPONENTIAL DISTRIBUTIONS

by

Lii-Yuh Leu	and	TaChen Liang
Graduate Institute of Statistics		Department of Mathematics
National Central University		Wayne State University
Chung-Li, Taiwan, R.O.C.		Detroit, Michigan 48202

Technical Report #89-14

Department of Statistics
Purdue University

June 1989

SELECTION OF THE BEST WITH A PRELIMINARY TEST FOR TWO-PARAMETER EXPONENTIAL DISTRIBUTIONS

Lii-Yuh Leu
Graduate Institute of Statistics
National Central University
Chung-Li, Taiwan, R.O.C.

TaChen Liang
Department of Mathematics
Wayne State University
Detroit, Michigan 48202

Key Words and Phrases: *best population; correct selection; indifference zone; non-selection zone; preference zone; two-parameter exponential distribution; guaranteed life time; two-stage procedure.*

ABSTRACT

This paper deals with the problem of selecting the best population from among $k(\geq 2)$ two-parameter exponential populations. New selection procedures are proposed for selecting the unique best. The procedures include preliminary tests which allow the experimenter to have an option to not select if the statistical evidence is not significant. Two probabilities, the probability to make a selection and the probability of a correct selection, are controlled by these selection procedures. Comparisons between the proposed selection procedures and certain earlier existing procedures are also made. The results show the superiority of the proposed selection procedures in terms of the required sample size.

1. INTRODUCTION

The problem of selecting the best population from among $k(\geq 2)$ populations has been studied extensively. A lot of selection procedures have been derived for different selection goals by several authors. Among them, Bechhofer (1954) introduced the indifference zone approach for selecting the normal population with the largest mean. In his approach, the determination of the sample size is heavily dependent of the indifference zone assumption. Also, the probability of a correct selection depends on the unknown parameters and is analogous to the power of a test. However, a probability that is analogous to the probability of type-I error of a test was not taken into consideration. Bechhofer's procedure forces the experimenter to make a selection, and often that procedure is not used in practical application because of the lack of a statistical test for the homogeneity of the parameters (see Simon (1977)).

Based on the preceding reasoning, Bishop and Pirie (1979) introduced a selection procedure in which a test of homogeneity was conducted. The procedure allows the experimenter to have the option not to make a selection if the statistical evidence is not significant. Later, Chen (1985) proposed modified selection procedures for the problem of selecting the best normal population. He considered a preliminary test based on the sampled spacing between the largest and the second largest ordered statistics. If the statistical evidence of the preliminary test is not significant, the experimenter decides not to make a selection. Otherwise, he selects the population yielding the largest sample mean value as the best population. The sample size is determined to control both the probability of type-I error for the preliminary test and the probability of a correct selection. Analogous to Chen (1985), Chen and Mithongtae (1986) proposed selection procedures for two-parameter exponential distribution models. However, their procedures are conservative in the sense that the determined sample size is always larger than the required minimum sample size. Also, their procedures can not be applied for a case where the common scale parameter is unknown.

In this paper, along the line of Chen (1985) and Chen and Mithongtae (1986), we study the problem of selecting the best population from among $k(\geq 2)$ two-parameter exponential populations. Selection procedures are derived according to whether the common scale parameter is known or unknown. Exact sample sizes are determined to control both the probability of type-I error and the probability of a correct selection for the cases considered by Chen and Mithongtae (1986). When the scale parameter is unknown, a two-stage selection procedure is also proposed. The two-stage selection procedure covers the case where the procedures of Chen and Mithongtae (1986) can not be used. It should be pointed out that the proposed procedures can be implemented by using certain existing tables.

2. FORMULATION OF THE PROBLEM

Let $\pi_1, \ldots, \pi_k$ denote $k(\geq 2)$ independent two-parameter exponential distributions with guaranteed life times $\mu_1, \ldots, \mu_k$, respectively, and a common (known or unknown) standard deviation θ. Let $\mu_{[1]} \leq \ldots \leq \mu_{[k]}$ denote the ordered values of $\mu_1, \ldots, \mu_k$ in the parameter space Ω where

$$\Omega = \{(\underline{\mu}, \theta) | \underline{\mu} = (\mu_1, \ldots, \mu_k), -\infty < \mu_i < \infty,\ \theta > 0\}.$$

We partition the parameter space into the following three subspaces:

the preference zone: $\Omega(PZ) = \{(\underline{\mu}, \theta) \in \Omega | \frac{\mu_{[k]} - \mu_{[k-1]}}{\theta} \geq \delta,\ \delta > 0\}$,

the nonselection zone: $\Omega(NZ) = \{(\underline{\mu}, \theta) \in \Omega | \mu_{[k-1]} = \mu_{[k]}\}$,

and

the indifference zone: $\Omega(IZ) = \Omega - \Omega(PZ) - \Omega(NZ)$,

where δ is a known positive constant.

Denote the event of a correct selection by CS and the event of selection by S. The goal is to develop selection procedures R to select a single best population with a minimum sample size from each population such that the following probability requirements are satisfied:

$$P_{(\underline{\mu}, \theta)}(S|R) \leq \alpha \qquad \text{for all } (\underline{\mu}, \theta) \in \Omega(NZ) \tag{2.1}$$

and

$$P_{(\underline{\mu}, \theta)}(CS|R) \geq P^* \text{ for all } (\underline{\mu}, \theta) \in \Omega(PZ) \tag{2.2}$$

where $\alpha \in (0,1)$ and $P^* \in (1/k, 1)$ are preassigned constants.

The selection procedure R depends on whether the common standard deviation is known or unknown.

3. SELECTION PROCEDURE FOR STANDARD DEVIATION KNOWN CASE

Let $X_{i1}, \ldots, X_{in}$ be a sample of size n arising from population π_i and let $Y_i = \min_{1 \le j \le n} X_{ij}$, $i = 1, \ldots, k$. Also, let $Y_{[1]} \le \ldots \le Y_{[k]}$ denote the ordered statistics of $Y_1, \ldots, Y_k$. When θ is known, we propose a selection procedure R_1 as follows:

R_1: select the population yielding $Y_{[k]}$ as the best population if $Y_{[k]} - Y_{[k-1]} > \lambda\theta/n$; otherwise, do not make a selection, where λ and n are chosen to satisfy the probability requirements (2.1) and (2.2).

For the rule R_1, we need to investigate the supremum of $P_{(\underline{\mu},\theta)}(S|R_1)$ for $(\underline{\mu}, \theta) \in \Omega(NZ)$ and the infimum of $P_{(\underline{\mu},\theta)}(CS|R_1)$ for $(\underline{\mu}, \theta) \in \Omega(PZ)$.

Firstly, let X_i, $i = 1, \ldots, k$ be independently distributed as two-parameter exponential distribution with guaranteed life time μ_i, $i = 1, \ldots, k$ and common standard deviation 1 respectively. Let $X_{[1]} \le \ldots \le X_{[k]}$ denote the ordered statistics of $X_1, \ldots, X_k$, $\mu_{[1]} \le \ldots \le \mu_{[k]}$ be the ordered guaranteed life times, and $X_{(i)}$ be the random variable associated with parameter $\mu_{[i]}$, $i = 1, \ldots, k$, respectively. Then we have the following Lemma:

Lemma 3.1. For any positive constant c, $P_{\underline{\mu}}\{X_{[k]} - X_{[k-1]} > c\}$ is nonincreasing in $\mu_{[i]}$, $i = 1, \ldots, k-2$.

Proof. Let F be the cdf and f be the pdf of the standard exponential random variable. By symmetry we may assume that $\mu_1 \le \mu_2 \le \ldots \le \mu_k$. For any $c > 0$, we have

$$\begin{aligned} &P_{\underline{\mu}}\{X_{[k]} - X_{[k-1]} > c\} \\ &= \sum_{i=1}^{k} \int_{d_i}^{\infty} \prod_{\substack{j=1 \\ j \ne i}}^{k} F(x + \mu_i - \mu_j - c) f(x)\,dx, \end{aligned}$$

where

$$d_i = \begin{cases} \mu_k - \mu_i + c, & i = 1, \ldots, k-1, \\ \max(0, \mu_{k-1} - \mu_k + c), & i = k. \end{cases}$$

For $\ell = 1, 2, \ldots, k-2$,

$$\begin{aligned}
&\frac{\partial}{\partial \mu_\ell} P_{\underline{\mu}}\{X_{[k]} - X_{[k-1]} > c\} \\
&= \sum_{\substack{i=1 \\ i \neq \ell}}^{k} \int_{d_\ell}^{\infty} \prod_{\substack{j=1 \\ j \neq i,\ell}}^{k} F(x + \mu_\ell - \mu_j - c) f(x + \mu_\ell - \mu_i - c) f(x) dx \\
&- \sum_{\substack{i=1 \\ i \neq \ell}}^{k-1} \int_{d_i}^{\infty} \prod_{\substack{j=1 \\ j \neq i,\ell}}^{k} F(x + \mu_i - \mu_j - c) f(x + \mu_i - \mu_\ell - c) f(x) dx \\
&- \int_{d_k}^{\infty} \prod_{\substack{j=1 \\ j \neq \ell}}^{k-1} F(x + \mu_k - \mu_j - c) f(x + \mu_k - \mu_\ell - c) f(x) dx \\
&= I_1 + I_2 - I_3,
\end{aligned} \tag{3.1}$$

where

$$\begin{aligned}
I_1 = \sum_{\substack{i=1 \\ i \neq \ell}}^{k-1} \int_{d_\ell}^{\infty} \left(\prod_{\substack{j=1 \\ j \neq i,\ell}}^{k} F(x + \mu_\ell - \mu_j + c) \right) [f(x + \mu_\ell - \mu_i - c) f(x) \\
- f(x - c) f(x + \mu_\ell - \mu_i)] dx,
\end{aligned}$$

$$I_2 = \int_{d_\ell}^{\infty} \left(\prod_{\substack{j=1 \\ j \neq \ell}}^{k-1} F(x + \mu_\ell - \mu_j - c) \right) f(x + \mu_\ell - \mu_k - c) f(x) dx,$$

and

$$I_3 = \int_{d_k}^{\infty} \left(\prod_{\substack{j=1 \\ j \neq \ell}}^{k-1} F(x + \mu_k - \mu_j - c) \right) f(x + \mu_k - \mu_\ell - c) f(x) dx.$$

$I_1 = 0$, since $f(x + \mu_\ell - \mu_i - c) f(x) = f(x - c) f(x + \mu_\ell - \mu_i)$. For $I_2 - I_3$, we consider the following two cases:

<u>Case 1.</u> $\mu_{k-1} - \mu_k + c > 0$.

After changing variables,

$I_2 - I_3$

$$= \int_{\mu_k-\mu_\ell+c}^{\infty} \left(\prod_{\substack{j=1\\ j\neq\ell}}^{k-1} F(x+\mu_\ell-\mu_j-c) \right) [f(x+\mu_\ell-\mu_k-c)f(x) - f(x-c)$$

$$f(x+\mu_\ell-\mu_k)]dx$$

$$- \int_{\mu_{k-1}-\mu_\ell+c}^{\mu_k-\mu_\ell+c} \left(\prod_{\substack{j=1\\ j\neq\ell}}^{k-1} F(x+\mu_\ell-\mu_j-c) \right) f(x-c)f(x+\mu_\ell-\mu_k)dx. \qquad (3.2)$$

The expression in the bracket of the first term in (3.2) is zero and the second term is nonnegative, hence the derivative (3.1) is nonpositive.

Case 2. $\mu_{k-1}-\mu_k+c \leq 0$.

After changing variables, analogous to the argument of Case 1, we have

$$I_2 - I_3$$

$$= \int_{\mu_k-\mu_\ell+c}^{\infty} \left(\prod_{\substack{j=1\\ j\neq\ell}}^{k-1} F(x+\mu_\ell-\mu_j-c) \right) [f(x+\mu_\ell-\mu_k-c)f(x)$$

$$- f(x-c)f(x+\mu_\ell-\mu_k)]dx$$

$$- \int_{\mu_k-\mu_\ell}^{\mu_k-\mu_\ell+c} \left(\prod_{\substack{j=1\\ j\neq\ell}}^{k-1} F(x+\mu_\ell-\mu_j-c) \right) f(x-c)f(x+\mu_\ell-\mu_k)dx$$

$$\leq 0.$$

This completes the proof of Lemma. □

From Lemma 3.1., we have the following corollary:

Corollary 3.2. $P_{\underline{\mu}}\{X_{[k]}-X_{[k-1]} > c\} \leq P\{|Z_1-Z_2| > c\}$ for all $\underline{\mu} \in \Omega(NZ)$, where Z_1 and Z_2 are independent having standard exponential distribution. Furthermore,

$$\sup_{\Omega(NZ)} P_{\underline{\mu}}\{X_{[k]}-X_{[k-1]} > c\} = 2H(-c),$$

where $H(t)$ denotes the cdf of $Z_1 - Z_2$.

Note that

$$H(t) = \begin{cases} 1 - \frac{1}{2}e^{-t}, & \text{for } t \geq 0 \\ \frac{1}{2}e^{t}, & \text{for } t < 0. \end{cases} \qquad (3.3)$$

Since $n(Y_i - \mu_i)/\theta$ is distributed as a standard exponential, we have the following result:

Theorem 3.3. $\sup_{\Omega(NZ)} P_{(\underline{\mu},\theta)}(S|R_1) = e^{-\lambda}$.

Proof. For $(\underline{\mu}, \theta) \in \Omega(NZ)$,

$$\begin{aligned} &P_{(\underline{\mu},\theta)}(S|R_1) \\ &= P_{(\underline{\mu},\theta)}\{Y_{[k]} - Y_{[k-1]} > \lambda\theta/n\} \\ &= P_{(\underline{\mu},\theta)}\{\frac{nY_{[k]}}{\theta} - \frac{nY_{[k-1]}}{\theta} > \lambda\} \\ &\leq P\{|Z_1 - Z_2| > \lambda\}, \text{ by Corollary 3.2.} \\ &= 2H(-\lambda) = e^{-\lambda}. \end{aligned}$$

□

Remark 3.1. In order to satisfy the probability requirement (2.1), we may let $e^{-\lambda} = \alpha$. Thus

$$\lambda = -\ell n\ \alpha. \tag{3.4}$$

We next evaluate the infimum of $P_{(\underline{\mu},\theta)}(CS|R_1)$ over $\Omega(PZ)$.

Theorem 3.4. The infimum of $P_{(\underline{\mu},\theta)}(CS|R_1)$ over $\Omega(PZ)$ occurs at the configuration $\mu_{[1]} = \ldots = \mu_{[k-1]} = \mu_{[k]} - \delta\theta$ and

$$\inf_{\Omega(PZ)} P_{(\underline{\mu},\theta)}(CS|R_1) = \int_{\ell}^{\infty} (1 - e^{-(x+\tau-\lambda)})^{k-1} e^{-x} dx, \tag{3.5}$$

where $\ell = \max(0, \lambda - \tau)$ and $\tau = n\delta$. Further, if $\tau > \lambda$, then

$$\inf_{\Omega(PZ)} P_{(\underline{\mu},\theta)}(CS|R_1) = e^{\tau-\lambda}(1 - (1 - e^{-(\tau-\lambda)})^k)/k. \tag{3.6}$$

Proof. See Chen and Mithongtae (1986). □

Remark 3.2.

(a) If $\tau \leq \lambda$, then the right hand side of (3.5) is $\frac{1}{k}e^{\tau-\lambda}$ which is less than or equal to $1/k$.

(b) Since $\tau = n\delta$, for n large enough, we have $\tau > \lambda$. In order to satisfy the probability requirement (2.2), from (3.6), we may let

$$e^{\tau-\lambda}(1-(1-e^{-(\tau-\lambda)})^k)/k = P^*. \tag{3.7}$$

(c) Let $v = e^{-(\tau-\lambda)}$, then (3.7) can be rewritten as

$$v^{-1}(1-(1-v)^k) = kP^*. \tag{3.8}$$

Tables for the v values when $P^* = 0.90$, 0.95, 0.975 and 0.99; $k = 2(1)\ 15$ are available from Raghavachari and Starr (1970).

(d) For $\alpha \in (0,1)$, $P^* \in (1/k, 1)$ and k, let $\lambda = -\ell n\ \alpha$. Also let v be the solution of the equation (3.8). Then $\tau = -\ell n(\alpha v)$. Let $< x >$ denote the smallest integer not less than x, then

$$n =< -\ell n(\alpha v)/\delta > \tag{3.9}$$

is the smallest sample size satisfying the probability requirements (2.1) and (2.2).

(e) In order to compare our procedure with Chen and Mithongtae's procedure, we let n^*, λ^*, τ^* and δ^* denote the corresponding notations of Chen and Mithongtae's procedure.

Then

$$\lambda^* = \lambda + \ell n\ u, \quad \tau^* = n^*\delta, \quad \delta^* = \delta\theta,\ \tau = n\delta$$

and

$$v = e^{-(n\delta-\lambda)} = e^{-(n^*\delta-\lambda^*)},$$

where $u = \frac{1}{2} + \sum_{i=1}^{k-1} 1/(k-i+1)$.

Hence

$$n^* = n + (\ell n\ u)/\delta.$$

Thus, $n^* > n$ as $k \geq 3$.

(f) The Table 1 of Chen and Mithongtae (1986) can be used to implement our procedure. The minimum sample size required is $n =< -\ell n(\alpha v)/\delta >$, where $v = e^{-(\tau^* - \lambda^*)}$, and λ^* and τ^* can be found from Table 1 of Chen and Mithongtae (1986).

Example. For $k = 5$, $\theta = 2$, $\delta^* = 0.5$, $\alpha = 0.05$ and $P^* = 0.95$, then $n^* = 29$ and $n = 27$.

4. SELECTION PROCEDURE FOR STANDARD DEVIATION UNKNOWN CASE

Under the formulation of $\Omega(PZ)$ in Section 2, we have a single stage selection procedure when the standard deviation is unknown. We take a sample of size n, say X_{ij}, $j = 1, \ldots, n$ from each population π_i, $i = 1, \ldots, k$. Let Y_i be the first order statistic. The minimum variance unbiased estimator of θ is

$$\hat{\theta} = \sum_{i=1}^{k} \sum_{j=1}^{n} (X_{ij} - Y_i)/v$$

where $v = k(n-1)$. Let $Y_{[1]} \leq \cdots \leq Y_{[k]}$ denote the ordered values of Y_i's. We propose a selection procedure R_2 as follows:

R_2: select the population yielding $Y_{[k]}$ as the best population if $Y_{[k]} - Y_{[k-1]} > \gamma\hat{\theta}/n$; otherwise, do not make a selection, where γ and n are chosen to satisfy the probability requirements (2.1) and (2.2).

Note that $Z_i = n(Y_{(i)} - \mu_{[i]})/\theta$, $i = 1, \ldots, k$ and $\hat{\theta}$ are mutually independent. $Z_i \sim$ standard exponential and $Q = \hat{\theta}/\theta$ has a gamma distribution with parameters v and v. We use $G(x)$ to denote the cdf of Q. Then $g(x) = G'(x) = \frac{v^v x^{v-1}}{\Gamma(v)} e^{-vx}$. By applying Corollary 3.2, we have the following result:

Theorem 4.1.

$$\sup_{\Omega(NZ)} P_{(\underline{\mu},\theta)}(S|R_2) = (\frac{v}{\gamma+v})^v. \tag{4.1}$$

Proof.

$$P_{(\underline{\mu},\theta)}(S|R_2) = P\{Y_{[k]} - Y_{[k-1]} > \gamma\hat{\theta}/n\}$$
$$= E(P_{(\underline{\mu},\theta)}\{Y_{[k]} - Y_{[k-1]} > \gamma Q\theta/n|Q\}).$$

Now

$$P_{(\underline{\mu},\theta)}\{Y_{[k]} - Y_{[k-1]} > \gamma Q\theta/n|Q = x\}$$
$$= P_{(\underline{\mu},\theta)}\{Y_{[k]} - Y_{[k-1]} > \gamma x\theta/n\}$$
$$\leq 2H(-\gamma x), \text{ by Corollary 3.2.}$$
$$= e^{-\gamma x}.$$

Hence

$$\sup_{\Omega(NZ)} P_{(\underline{\mu},\theta)}(S|R_2)$$
$$= \int_0^\infty e^{-\gamma x} dG(x)$$
$$= \int_0^\infty \frac{v^v x^{v-1}}{\Gamma(v)} e^{-(\gamma+v)x} dx$$
$$= (\frac{v}{\gamma+v})^v.$$

□

By setting $\sup_{\Omega(NZ)} P_{(\underline{\mu},\theta)}(S|R_2)$ equal to α, we obtain

$$\gamma = v\{\alpha^{-1/v} - 1\}. \tag{4.2}$$

Now we evaluate the infimum of $P_{(\underline{\mu},\theta)}(CS|R_2)$ for $(\underline{\mu},\theta) \in \Omega(PZ)$.

Theorem 4.2. The infimum of $P_{(\underline{\mu},\theta)}(CS|R_2)$ over $\Omega(PZ)$ occurs at the configuration $\mu_{[1]} = \ldots = \mu_{[k-1]} = \mu_{[k]} - \delta\theta$ and

$$\inf_{\Omega(PZ)} P_{(\underline{\mu},\theta)}(CS|R_2) = \int_0^{n\delta/\gamma} \frac{1}{k} e^{n\delta-\gamma x}(1-(1-e^{\gamma x - n\delta})^k) dG(x)$$
$$+ \int_{n\delta/\gamma}^\infty \frac{1}{k} e^{n\delta-\gamma x} dG(x). \tag{4.3}$$

Remark 4.1. Although Theorem 4.2. is similar to Theorem 4 of Chen and Mithongtae (1986), the concept is different. Their procedure should specify the value of δ^*/θ in advance. However, it seems hard or not possible since the value of the parameter θ is unknown.

In order to satisfy the probability requirement (2.2), we set

$$\int_0^{n\delta/\gamma} \frac{1}{k} e^{n\delta-\gamma x}(1-(1-e^{\gamma x-n\delta})^k)dG(x) + \int_{n\delta/\gamma}^{\infty} \frac{1}{k} e^{n\delta-\gamma x} dG(x) = P^*. \tag{4.4}$$

The numerical methods of Chen and Mithongtae (1986) can be applied for solving equation (4.4).

5. A TWO-STAGE SELECTION PROCEDURE WITH UNKNOWN STANDARD DEVIATION

In the following, we consider the case where the parameter space is partitioned by the same way as that of Chen and Mithongtae (1986). That is,

$$\tilde{\Omega}(PZ) = \{(\underline{\mu}, \theta) | \mu_{[k]} - \mu_{[k-1]} \geq \delta,\ \delta > 0\},$$

and

$$\tilde{\Omega}(NZ) = \{(\underline{\mu}, \theta) | \mu_{[k-1]} = \mu_{[k]}\}.$$

When θ is known, the situation is equivalent to the one discussed in Section 3. However, when θ is unknown, a single stage selection procedure proposed by Chen and Mithongtae (1986) can not be implemented since it seems not possible to specify the value of the ratio δ/θ in advance. Analogous to that of Desu, Narula and Villarreal (1977), a two-stage selection procedure R_3 is proposed as follows:

(i) Take an initial sample of size n_0 from each of the k populations, say $X_{i1}, \ldots, X_{in_0}$, $i = 1, \ldots, k$. Let $Y_i(n_0) = \min_{1 \leq j \leq n_0} X_{ij}$, $i = 1, \ldots, k$, and $\hat{\theta}_0 = \sum_{i=1}^{k} \sum_{j=1}^{n_0} (X_{ij} - Y_i(n_0))/v_0$, where $v_0 = k(n_0 - 1)$.

(ii) compute $\eta = h\hat{\theta}_0/\delta$, where $h > \gamma_0$ is determined so that

$$\int_0^{\infty} \int_0^{\infty} (1 - e^{-\zeta-(h-\gamma_0)x})^{k-1} e^{-\zeta} d\zeta g(x) dx = P^*, \tag{5.1}$$

$\gamma_0 = v_0\{\alpha^{-1/v_0} - 1\}$ and $g(x)$ is the pdf of $Q = \hat{\theta}_0/\theta$. Let $N = \max\{n_0, < \eta >\}$, where $< x >$ denotes the smallest integer not less than x.

(iii) If necessary, take $N - n_0$ additional observations from each of the k populations and compute

$$Y_i(N) = \min_{1 \le j \le N} X_{ij}, \; i = 1, \ldots, k.$$

(iv) Select the population yielding $Y_{[k]}(N)$ as the best population if $Y_{[k]}(N) - Y_{[k-1]}(N) > \gamma_0\hat{\theta}_0/N$; otherwise, do not make a selection.

Let $Z_i = N(Y_i(N) - \mu_i)/\theta$, $i = 1, \ldots, k$. It is well-known that, given $N = n$, Z_i, $i = 1, \ldots, k$, are i.i.d. having standard exponential distribution.

Now, we evaluate the supremum of $P_{(\underline{\mu},\theta)}(S|R_3)$ for $(\underline{\mu}, \theta) \in \tilde{\Omega}(NZ)$.

Theorem 5.1. For any $(\underline{\mu}, \theta) \in \tilde{\Omega}(NZ)$ and any positive constant γ_0, $P_{(\underline{\mu},\theta)}\{Y_{[k]}(N) - Y_{[k-1]}(N) > \gamma_0\hat{\theta}_0/N\} \le (\frac{v_0}{\gamma_0 + v_0})^{v_0}$, where $v_0 = k(n_0 - 1)$.

Proof. Let $A_n = \{N = n\}$, then $A_{n_0} = \{Q \le \frac{n_0\delta}{h\theta}\}$ and $A_n = \{\frac{(n-1)\delta}{h\theta} \le Q < \frac{n\delta}{h\theta}\}$, $n > n_0$.

$$\begin{aligned}
&P_{(\underline{\mu},\theta)}\{Y_{[k]}(N) - Y_{[k-1]}(N) > \gamma_0\hat{\theta}_0/N\} \\
&= P_{(\underline{\mu},\theta)}\left\{\frac{NY_{[k]}(N)}{\theta} - \frac{NY_{[k-1]}(N)}{\theta} > \gamma_0 Q\right\} \\
&= \sum_{n=n_0}^{\infty} \int_{A_n} P_{(\underline{\mu},\theta)}\left\{\frac{NY_{[k]}(N)}{\theta} - \frac{NY_{[k-1]}(N)}{\theta} > \gamma_0 x\right\} g(x)dx \\
&\le \sum_{n=n_0}^{\infty} \int_{A_n} e^{-\gamma_0 x} g(x)dx, \text{ by corollary 3.2.} \\
&= \int_0^{\infty} e^{-\gamma_0 x} g(x)dx \\
&= \left(\frac{v_0}{\gamma_0 + v_0}\right)^{v_0}.
\end{aligned}$$

□

Remark 5.1. Note that $\sup_{\hat{\Omega}(NZ)} P_{(\underline{\mu},\theta)}(S|R_3) = (\frac{v_0}{\gamma_0 + v_0})^{v_0}$ which is independent of N. In order to satisfy the probability requirement (2.1), we may set $(\frac{v_0}{\gamma_0 + v_0})^{v_0} = \alpha$ and hence

$$\gamma_0 = v_0\{\alpha^{-1/v_0} - 1\}. \tag{5.2}$$

Next, we evaluate the infimum of the probability of a correct selection for $(\underline{\mu}, \theta) \in \tilde{\Omega}(PZ)$.

Theorem 5.2. The infimum of $P_{(\underline{\mu},\theta)}(CS|R_3)$ over $\tilde{\Omega}(PZ)$ occurs at the configuration $\mu_{[1]} = \ldots = \mu_{[k-1]} = \mu_{[k]} - \delta$ and

$$\inf_{\tilde{\Omega}(PZ)} P_{(\underline{\mu},\theta)}(CS|R_3) \geq \int_0^\infty \int_0^\infty (1 - e^{-\zeta-(h-\gamma_0)x})^{k-1} e^{-\zeta} d\zeta g(x) dx. \tag{5.3}$$

Proof. $P_{(\underline{\mu},\theta)}(CS|R_3)$

$$\begin{aligned}
&= P_{(\underline{\mu},\theta)}\{Y_{[k]}(N) - Y_{[k-1]}(N) > \gamma_0\hat{\theta}_0/N,\ Y_{(k)}(N) = Y_{[k]}(N)\} \\
&= P_{(\underline{\mu},\theta)}\{Y_{(j)}(N) < Y_{(k)}(N) - \gamma_0\hat{\theta}_0/N,\ j = 1,\ldots,k-1\} \\
&= P_{(\underline{\mu},\theta)}\{\frac{N(Y_{(j)}(N) - \mu_{[j]})}{\theta} < \frac{N(Y_{(k)}(N) - \mu_{[k]})}{\theta} + \frac{N(\mu_{[k]} - \mu_{[j]})}{\theta} \\
&\qquad - \frac{\gamma_0\hat{\theta}_0}{\theta},\ j = 1,\ldots,k-1\} \\
&\geq P\left\{Z_j \leq Z_k + \frac{N\delta}{\theta} - \gamma_0 Q,\ j = 1,\ldots,k-1\right\} \\
&\geq P\{Z_j \leq Z_k + hQ - \gamma_0 Q,\ j = 1,\ldots,k-1\}, \text{ since } \frac{N\delta}{\theta} \geq hQ \\
&= \int_0^\infty \int_0^\infty (1 - e^{-\zeta-(h-\gamma_0)x})^{k-1} e^{-\zeta} d\zeta g(x) dx. \qquad \square
\end{aligned}$$

Remark 5.2. If $h \leq \delta_0$, then the right hand side of (5.3) is

$$\begin{aligned}
&\int_0^\infty \int_{(\gamma_0-h)x}^\infty \left(1-e^{-\zeta-(h-\gamma_0)x}\right)^{k-1} e^{-\zeta} d\zeta g(x)dx \\
&= \int_0^\infty \int_0^\infty \left(1-e^{-u}\right)^{k-1} e^{-u+(h-\gamma_0)x} du g(x)dx \\
&= \frac{1}{k}\int_0^\infty e^{(h-\gamma_0)x} \frac{v_0^{v_0}\ x^{v_0-1}}{\Gamma(v_0)}\ e^{-v_0 x} dx \\
&= \frac{1}{k}\left(\frac{v_0}{\gamma_0-h+v_0}\right)^{v_0} \\
&\leq \frac{1}{k}, \text{ since } \gamma_0-h+v_0 \geq v_0.
\end{aligned}$$

Hence we should choose h such that $h > \gamma_0$.

In order to satisfy the probability requirement (2.2), we may set

$$\int_0^\infty \int_0^\infty \left(1-e^{-\zeta-(h-\gamma_0)x}\right)^{k-1} e^{-\zeta} d\zeta g(x)dx = P^*. \tag{5.4}$$

Remark 5.3. Table of Desu, Narula and Villarreal (1977) can be used to implement the proposed two–stage procedure R_3. Let d be the value obtained from their table, then $h = \gamma_0 + 2v_0 d$.

ACKNOWLEDGMENTS

This research was supported in part by the Office of Naval Research Contracts N00014–88–K–0170 and NSF Grant DMS-8606964 at Purdue University. The research of Dr. Lii–Yuh Leu was also supported in part by the National Science Council, Republic of China.

BIBLIOGRAPHY

Bechhofer, R.E. (1954). A single–sample multiple decision procedure for ranking means of normal populations with known variances. *Annals of Mathematical Statistics*, **25**, 16–39.

Bishop, T.A. and Pirie, W.R. (1979). A class of selection rules with options to not select. *Technical Report No.* **196**, Department of Statistics, The Ohio State University, Columbus, Ohio.

Chen, H.J. (1985). A new selection procedure for selecting the best population with a preliminary test (with discussion). *The Frontiers of Modern Statistical Inference Procedures.* American Sciences Press, Inc., Columbus, Ohio, 99–117.

Chen, H.J. and Mithongtae, J. (1986). Selection of the best exponential distribution with a preliminary test. *American Journal of Mathematical and Management Sciences*, **Vol. 6**, 219–249.

Desu, M.M., Narula, S.C. and Villarreal, B. (1977). A two-stage procedure for selecting the best of k exponential distributions. *Commun. Statist.-Theor. Meth.*, **A6(12)**, 1223-1230.

Raghavachari, M. and Starr, N. (1970). Selection problems for some terminal distributions. *Metron*, **28**, 185–197.

Simon, R. (1977). Adaptive treatment assignment methods and clinical trials. *Biometrics*, **33**, 743–749.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release, distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
Technical Report #89-14

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
Purdue University

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION

6c. ADDRESS *(City, State, and ZIP Code)*
Department of Statistics
West Lafayette, IN 47907

7b. ADDRESS *(City, State, and ZIP Code)*

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
Office of Naval Research

8b. OFFICE SYMBOL *(If applicable)*

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
N00014-88-K-0170 and
NSF Grant DMS-8606964

8c. ADDRESS *(City, State, and ZIP Code)*
Arlington, VA 22217-5000

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION

11. TITLE *(Include Security Classification)*
SELECTION OF THE BEST WITH A PRELIMINARY TEST FOR TWO-PARAMETER EXPONENTIAL DISTRIBUTIONS (Unclassified)

12. PERSONAL AUTHOR(S)
Lii-Yuh Leu and TaChen Liang

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	June 1989	15

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Best population, correct selection, indifference zone, non-selection zone, preference zone, two-parameter exponential distribution, guaranteed life time, two-stage ced

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This paper deals with the problem of selecting the best population from among $k(\geq 2)$ two-parameter exponential populations. New selection procedures are proposed for selecting the unique best. The procedures include preliminary tests which allow the experimenter to have an option to not select if the statistical evidence is not significant. Two probabilities, the probability to make a selection and the probability of a correct selection, are controlled by these selection procedures. Comparisons between the proposed selection procedures and certain earlier existing procedures are also made. The results show the superiority of the proposed selection procedures in terms of the required sample size.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
TaChen Liang

22b. TELEPHONE *(Include Area Code)*
313-577-2493

22c. OFFICE SYMBOL