Journal of the KSQC
Vol. 13, No. 2, 1985

# An Elimination Type Two-Stage Selection Procedure for Gamma Populations

Lee, Seung-Ho*
Choi, Kook Lyeol*

ABSTRACT

The problem of selecting the gamma population with the largest mean out of $k$ gamma populations, each of which has the same shape parameter is considered. An elimination type two-stage procedure is proposed which guarantees the same probability requirement using the indifference-zone approach as does the single-stage procedure of Gibbons, Olkin and Sobel (1977). The two-stage procedure has the highly desirable property that the expected total number of observations required by the procedure is always less than that of the corresponding single-stage procedure regardless of the configuration of the population parameters.

## 1. Introduction

The gamma model is considered frequently in the area of reliability and life testing. This may be partly due to its relationship to the Poisson process, since the waiting time to the $r$-th occurrence of a Poisson process follows a gamma distribution. More importantly, however, it is a generalization of the exponential distribution, and it provides a rather flexible skewed density defined over the positive range. Its hazard function may be increasing or decreasing, but it approaches a constant as time approaches infinity.

Suppose that $\pi(\theta, r)$ denotes a gamma population with density

$$(1.1)\quad f(x;\theta,r)=\frac{1}{\Gamma(r)\theta^r}x^{r-1}e^{-x/\theta},\qquad x>0;r,\theta>0.$$

The mean is $E(X)=r\theta$. The parameters $r$ and $\theta$ are referred to as shape and scale parameter, respectively.

Wilk, Gnanadesikan and Huyette (1962) fit a gamma distribution to failure time observed for transisters in a accelerated life test. Note that in this experiment $\theta$ is proportional to the time to failure (when $r$ is considered as fixed). Thus a large value of $\theta$ would indicate a large time to failure or a

* Department of Mathematics, Ajou University.
** This work was supported in part by the Research Fund of the Research Institude of Ajou University, 1984.

long expected lifetime.

In this paper, we consider the problem of identifying the gamma population with the largest value of $\theta$ out of $k$ gamma populations, each of which has the same known value or $r$, but has possibly different values of $\theta$. Note that because $r$ is the same for each population, choosing the population with the largest $\theta$ value is equivalent to selecting the population with the largest mean, which would be regarded as the best population.

Gupta (1963) investigated the problem of choosing the best of several gamma populations under the framework of subset selection approach, and tabulated the design constants of his procedure. The problem of selecting the best under the framework of indifference-zone approach can be solved using the table of Gupta (1963) which is given in the book by Gibbons, Olkin and Sobel (1977).

In the difference-zone approach, the single-stage procedure of Gibbons, Olkin and Sobel (1977) does not utilize the information from the data as they are observed. The single-stage procedure would require, in general, a large amount of sample size relative to the procedures which could react to the data.

In Section 2, we propose a two-stage procedure which has the property that it screens out non-contending populations in the preliminary stage, and concentrates sampling only on contending populations in the second stage. And we introduce the unrestricted minimax criterion as the design criterion of the proposed procedure.

In Section 3, a lower bound of the probability of correct selection and a general expression of the expected total sample size of the proposed procedure are derived. And in Section 4, an optimization problem is solved to determine the design constants to implement the proposed procedure.

## 2. An Elimination Type Two-Stage Procedure

Let $\Pi_i(\theta_i, r)$, $1 \leqq i \leqq k$ denote $k$ gamma populations with unknown scale parameter $\theta_i$ and a common known shape parameter $r$. Let $\theta_{[1]} \leqq \cdots \leqq \theta_{[k]}$ denote the ordered values of $\theta_1, \cdots, \theta_k$, where the correct pairing between $\theta_i$ and $\theta_{[i]}$ are unknown.

The goal of the experimenter is to select the population associated with $\theta_{[k]}$, which is the best. Following the indifference-zone approach of Bechhofer (1954), the experimenter, prior to the experiment, specifies two constants $\delta^*(>0)$ and $P^*(1/k<P^*<1)$ which are incorporated into a probability requirement.

$$(2.1) \quad P_{\theta}\{CS\} \geqq P^*, \quad \text{for all} \quad \underline{\theta}=(\theta_1, \cdots, \theta_k) \in \Omega(\delta^*)$$

where $CS$ denotes the event of correct selection and $\Omega(\delta^*)=\{\underline{\theta} \mid \theta_{[k]}/\theta_{[k-1]} \geqq \delta^*\}$, which is called the preference-zone.

The following elimination type two-stage selection procedure $R_2$ is proposed as a generalization of the single-stage procedure $R_1$ of Gibbons, Olkin and Sobel (1977).

**Stage 1**; Take $n$ independent observations $X_{i1}, \cdots, X_{in}$ from $\Pi_i (1 \leqq i \leqq k)$, compute $T_i^{(1)} = \sum_{j=1}^{n} X_{ij}$ and determine a subset $I$ of $\{1, 2, \cdots, k\}$, where

$$(2.2) \quad I=\{i \mid cT_i^{(1)} \geqq \max_{1 \leq j \leq k} T_j^{(1)}\}, \quad c>1.$$

(a) If $I$ has only one element, stop sampling and assert that the population associated with $\max_{1 \leq i \leq k} T_i^{(1)}$ is the best.

(b) If $I$ has more than one element, go to the second stage.

**Stage 2**; Take $m$ additional independent observations $X_{i\,n+1},\cdots,X_{i\,n+m}$ from each $\pi_i$ for $i\in I$, compute $T_i=T_i^{(1)}+T_i^{(2)}=\sum_{j=1}^{n}x_{ij}+\sum_{j=n+1}^{m}x_{ij}=\sum_{j=1}^{n+m}x_{ij}$ and assert that the population associated with $\max_{i\in I} T_i$ is the best.

Note that the statistics $T_i^{(1)}$ and $T_i$ have the gamma distributions with scale parameter $\theta_i$ and the shape parameter $nr$ and $(n+m)r$, respectively.

**Remark.** If $c=0$ (or $+\infty$) the two-stage procedure $R_2$ reduces to the single-stage procedure $R_1$ with sample size $n$(or $n+m$) per population.

In the above definition of the two-stage procedure $R_2$, the sample size $n, m$ and the constant $c$ will be chosen so that the procedure guarantees the basic probability requirement (2.1). To make the choices unique as well as to have the total sample size (TSS) small, we adopt the following unrestricted minimax criterion:

minimize

$$(2.3) \qquad \sup_{\theta\in\Omega} E_{\underline{\theta}}(TSS\,|\,R_2)$$

subject to

$$(2.4) \qquad \inf_{\theta\in\Omega(\delta^*)} P_{\underline{\theta}}\{CR\,|\,R_2\}\geq P^*.$$

## 3. A Lower Bound of P{CS | R₂} and Expected Total Sample Size

A central problem concerned with the construction of procedures using the indifference-zone approach is to find the infimum of the probability of correct selection over the preference-zone $\Omega(\delta^*)$. Any parameter configuration at which such an infimum is attained is called a least favorable configuration (LFC) of the parameters for the procedure under study. For the procedure $R_2$, a LFC is $\theta_{[1]}=\cdots=\theta_{[k-1]}=\theta_{[k]}-\delta^*=\delta^*(\theta)$ which is intuitively obvious.

**Theorem 3.1.** *For the procedure $R_2$, the following inequality holds.*

$$(3.1) \qquad \inf_{\theta\in\Omega(\delta^*)} P_{\underline{\theta}}\{CS\,|\,R_2\}\geq \inf_{\theta>0} E_\theta[H^{k-1}\{cT_k^{(1)},T_k\,|\,\delta^*(\theta)\}]$$

*where $H(\cdot,\cdot\,|\,\theta)$ is the joint c.d.f. of $T_j^{(1)}$ and $T_j$ with scale parameter $\theta$.*

**Proof.** Without loss of generality, we may assume that $\theta_1\leq\theta_2\leq\cdots\leq\theta_k$. Then for all $\underline{\theta}\in\Omega(\delta^*)$

$$P_{\underline{\theta}}\{CS|R_2\} = P_{\underline{\theta}}\{cT_k^{(1)}\geq \max_{1\leq j\leq k} T_j^{(1)},\ T_k=\max_{i\in I} T_i\}$$

$$\geqq P_{\underline{\theta}}\{cT_k^{(1)} \geqq T_j^{(1)},\ T_k \geqq T_j \quad \text{for all} \quad j=1,\cdots,k-1\}$$

$$= \int \prod_{j=1}^{k-1} H(cx, y \mid \theta_j) dH(x, y \mid \theta_k)$$

$$\geqq \int \prod_{j=1}^{k-1} H(cx, y \mid \delta^*(\theta)) dH(x, y \mid \theta)$$

$$= E_\theta[H^{k-1}\{cT_k^{(1)}, T_k \mid \delta^*(\theta)\}]$$

The last inequality comes from the fact that the joint cdf $H(x, y \mid \theta_j)$ of $(T_j^{(1)}, T_j)$ is non-increasing in $\theta_j$ and that $\theta_j \leqq \delta(\theta_k)$ for $j=1, \cdots, k-1$ whenever $\underline{\theta} \in \Omega(\delta^*)$.

The lower bound in (3.1) would be difficult to compute in practice due to the dependence between $T_k^{(1)}$ and $T_k$. Thus it seems reasonable to find a lower bound which is slightly less sharp but more easily computable. Such a lower bound can be provided by the following result.

**Lemma 3.2.**

$$E_\theta[H^{k-1}\{cT_k^{(1)}, T_k \mid \delta^*(\theta)\}]$$

$$\geqq E_\theta[F^{k-1}\{cT_k^{(1)} \mid \delta^*(\theta)\}]\, E_\theta[G^{k-1}\{T_k \mid \delta^*(\theta)\}]$$

*where $F(\cdot \mid \theta)$ and $G(\cdot \mid \theta)$ are the marginal cdf's of $T_j^{(1)}$ and $T_j$, respectively.*

**Proof.**

$$H\{cT_k^{(1)}, T_k \mid \delta^*(\theta)\} = P_\theta\{T_j^{(1)} \leqq cT_k^{(1)},\ T_j^{(1)} + T_j^{(2)} \leqq T_k\}$$

$$= E_\theta[P_\theta\{T_j^{(1)} \leqq cT_k^{(1)},\ T_j^{(1)} \leqq T_k - T_j^{(2)} \mid T_j^{(2)}\}]$$

$$\geqq E_\theta[P_\theta\{T_j^{(1)} \leqq cT_k^{(1)} \mid T_j^{(2)}\}\ P_\theta\{T_j^{(1)} \leqq T_k - T_j^{(2)} \mid T_j(2)\}]$$

$$= P_\theta\{T_j^{(1)} \leqq cT_k^{(1)}\}\ P_\theta\{T_j^{(1)} + T_j^{(2)} \leqq T_k\}$$

$$= F(cT_k^{(1)}) G(T_k)$$

**Corollary 3.3.**

$$\inf_{\underline{\theta} \in \Omega(\delta^*)} P_{\underline{\theta}}\{CS \mid R_2\}$$

(3.2)

$$\geqq \int_0^\infty F_{nr}^{k-1}(c\delta^* x) dF_{nr}(x) \int_0^\infty F_{(n+m)r}^{k-1}(\delta^* y) dF_{(n+m)r}(y)$$

*where $F_\nu(\cdot)$ denotes the cdf of the gamma distribution with the shape parameter $\nu$ and scale parameter unity.*

There are an infinite number of combinations of $(n,m,c)$ which for given $k$ and $(\delta^*, P^*)$ will guarantee the basic probability requirement (2.1). Hence we adopt the unrestricted minimax criterion (2.3) and (2.4). To solve (2.3) and (2.4) we find an analytical expression for $E_{\underline{\theta}}(TSS\,|\,R_2)$.

Let $\bar{\bar{I}}$ denote the cardinality of the set $I$ in stage 1 and let

$$S = \begin{cases} 0 & \text{if } \bar{\bar{I}} = 1 \\ \bar{\bar{I}} & \text{if } \bar{\bar{I}} \geqq 2. \end{cases}$$

Then the total sample size (TSS) required by $R_2(n, m, c)$ is

$$TSS = kn + Sn.$$

A general expression for $E_{\underline{\theta}}(TSS|R_2)$ is summarized in the following theorem.

**Theorem 3.4.** *For any $\underline{\theta} \in \Omega$ we have*

$$E_{\underline{\theta}}(TSS\,|\,R_2)$$

(3.3)

$$= kn + m \sum_{i=1}^{k} \{\int \prod_{j\neq i} F(cx\,|\,\theta_j)dF(x\,|\,\theta_i) - \int \prod_{j\neq i} F(x/c\,|\,\theta_j)dF(x\,|\,\theta_i)\}$$

**Proof.**

$$E_{\underline{\theta}}(S\,|\,R_2) = E_{\underline{\theta}}(\bar{\bar{I}}\,|\,R_2) - P_{\underline{\theta}}\{\bar{\bar{I}} = 1\,|\,R_2\}$$

$$= \sum_{i=1}^{k} [P_{\underline{\theta}}\{cT_i^{(1)} \geqq \max_{1\leq j\leq k} T_j^{(1)}\} - P_{\underline{\theta}}\{T_i^{(1)} \geqq \max_{j\neq i} (cT_j^{(1)}\}]$$

$$= \sum_{i=1}^{k} [\int \prod_{j\neq i} F(cx\,|\,\theta_j)dF(x\,|\,\theta_i) - \int \prod_{j\neq i} F(x/c\,|\,\theta_i)dF(x\,|\,\theta_i)].$$

Along the lines of Gupta (1965) it can be shown that the supremum of $E_{\underline{\theta}}(TSS\,|\,R_2)$ is attained when $\theta_1 = \theta_2 = \cdots = \theta_k$. Thus we have the following result.

**Collary 3.5.**

$$\sup_{\theta\in\Omega} E_{\underline{\theta}}(TSS\,|\,R_2)$$

(3.4)

$$= kn + km\{\int_0^\infty F_{nr}^{k-1}(cx)dF_{nr}(x) - \int_0^\infty F_{nr}^{k-1}(x/c)dF_{nr}(x)\}.$$

Instead of solving the optimization problem given by (2.3) and (2.4), we are to solve the optimization problem of minimizing (3.4) with the constraints (3.2), which are conservative since (3.2) is a lower bound of $\inf_{\theta \in \Omega(\delta^*)} P_\theta\{CS|R_2\}$.

## 4. Optimization Problems Yielding Conservative Solutions and the Performance of $R_2$ Relative to $R_1$.

The problem (3.2) and (3.4) are extremely complicated integer programing problems with nonlinear constraints and objective function. In solving the optimization problem, we have treated $n$ and $m$ as continuous variables, and have used SUMT algorithm of Fiacco and McCormick (1968). We shall denote by $(\hat{n},\hat{m},\hat{c})$ a solution to this continuous version of the optimization problem. This optimization problem has been solved numerically for $k$ = 2(1) 10, $P^*$ = 0.90, 0.95 and $\delta^*$ = 1.75, 2.0, 3.0. The results are given in Table 4.1.

In order to get the insight into the savings in total sample size by the two-stage procedure $R_2$, we consider the ratios, which are termed relative efficiency (RE) as defined by

$$RE = \frac{\sup E_\theta(TSS|R_2)}{kn_0} \times 100(\%)$$

where $n_0$ is the sample size needed by the single-stage procedure $R_1$ to satisfy the same probability requirement (2.1).

The values of RE as well as the supremum of the expected total sample size are given in Table 4.2.

It can be observed from Table 4.2 that the RE is decreasing in $k$ although this has not been established analytically. Thus the effectiveness of $R_2$ appears to be increasing in $k$, which is in accordance with one's intuition that the screening process would be helpful when $k$ is large.

**Table 4.1. Design constants $(\hat{n}, \hat{m}, \hat{c})$ of the procedure $R_2$**

(a) $P^*$ = 0.90

| | $\delta^*$ = 1.75 | | | $\delta^*$ = 2.0 | | | $\delta^*$ = 3.0 | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $\hat{n}\ r$ | $\hat{m}\ r$ | $\hat{c}$ | $\hat{n}\ r$ | $\hat{m}\ r$ | $\hat{c}$ | $\hat{n}\ r$ | $\hat{m}\ r$ | $\hat{c}$ |
| 2 | 5.97 | 5.10 | 3.32 | 3.30 | 3.91 | 4.83 | 2.16 | 1.35 | 3.60 |
| 3 | 7.82 | 8.60 | 2.53 | 6.04 | 4.69 | 2.93 | 2.79 | 2.06 | 3.06 |
| 4 | 9.44 | 10.84 | 1.99 | 6.61 | 6.67 | 2.27 | 2.79 | 2.06 | 2.61 |
| 5 | 11.51 | 12.03 | 1.67 | 7.62 | 7.92 | 1.86 | 3.25 | 2.70 | 2.52 |
| 6 | 13.20 | 13.15 | 1.51 | 7.73 | 9.58 | 1.80 | 3.71 | 3.80 | 2.44 |
| 7 | 13.90 | 14.69 | 1.47 | 8.99 | 9.87 | 1.61 | 3.76 | 4.05 | 2.08 |
| 8 | 14.66 | 15.66 | 1.43 | 9.03 | 11.32 | 1.58 | 4.00 | 4.21 | 1.98 |
| 9 | 14.78 | 17.62 | 1.42 | 9.20 | 12.29 | 1.57 | 4.26 | 4.51 | 1.83 |
| 10 | 15.89 | 17.79 | 1.38 | 9.50 | 12.90 | 1.54 | 4.40 | 4.82 | 1.76 |

(b) $P^* = 0.95$

| $k$ | $\delta^* = 1.75$ $\hat{n}r$ | $\hat{m}r$ | $\hat{c}$ | $\delta^* = 2.0$ $\hat{n}r$ | $\hat{m}r$ | $\hat{c}$ | $\delta^* = 3.0$ $\hat{n}r$ | $\hat{m}r$ | $\hat{c}$ |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 9.05 | 9.65 | 2.05 | 6.43 | 5.73 | 2.52 | 3.12 | 2.10 | 3.48 |
| 3 | 12.56 | 12.66 | 1.71 | 8.88 | 7.64 | 1.89 | 3.90 | 3.23 | 2.37 |
| 4 | 15.12 | 14.84 | 1.52 | 10.14 | 9.48 | 1.67 | 4.38 | 3.91 | 2.11 |
| 5 | 16.37 | 16.54 | 1.48 | 11.71 | 10.21 | 1.52 | 4.88 | 4.63 | 1.81 |
| 6 | 17.69 | 17.96 | 1.42 | 12.41 | 11.29 | 1.47 | 4.96 | 4.59 | 1.99 |
| 7 | 18.71 | 19.25 | 1.39 | 12.79 | 12.49 | 1.45 | 5.42 | 4.74 | 1.82 |
| 8 | 19.54 | 20.30 | 1.37 | 13.10 | 13.56 | 1.43 | 5.51 | 5.18 | 1.79 |
| 9 | 20.29 | 21.38 | 1.35 | 13.37 | 14.49 | 1.42 | 5.59 | 5.47 | 1.79 |
| 10 | 21.04 | 22.48 | 1.32 | 13.54 | 15.46 | 1.41 | 5.96 | 5.70 | 1.66 |

**Table 4.2. Relative efficiency (RE) of the procedure $R_2$**

(a) $P^* = 0.90$

| $k$ | $\delta^* = 1.75$ E(TSS $\mid R_2$) | RE(%) | $\delta^* = 2.0$ E(TSS $\mid R_2$) | RE(%) | $\delta^* = 3.0$ E(TSS $\mid R_2$) | RE(%) |
|---|---|---|---|---|---|---|
| 2 | 21.6 | 98.2 | 14.4 | 98.0 | 6.4 | 80.0 |
| 3 | 47.5 | 98.9 | 31.3 | 94.8 | 13.3 | 88.6 |
| 4 | 74.3 | 97.8 | 48.9 | 94.0 | 20.5 | 85.4 |
| 5 | 101.3 | 92.1 | 66.7 | 93.2 | 28.1 | 90.6 |
| 6 | 128.7 | 93.3 | 84.9 | 94.3 | 35.0 | 83.3 |
| 7 | 156.3 | 89.3 | 102.7 | 91.7 | 42.6 | 86.9 |
| 8 | 184.0 | 85.2 | 120.9 | 88.9 | 50.1 | 89.4 |
| 9 | 211.7 | 84.0 | 139.2 | 85.9 | 57.5 | 91.2 |
| 10 | 240.0 | 82.8 | 157.5 | 87.5 | 65.0 | 81.2 |

(b) $P^* = 0.95$

| $k$ | $\delta^* = 1.75$ E(TSS $\mid R_2$) | RE(%) | $\delta^* = 2.0$ E(TSS $\mid R_2$) | RE(%) | $\delta^* = 3.0$ E(TSS $\mid R_2$) | RE(%) |
|---|---|---|---|---|---|---|
| 2 | 34.8 | 96.7 | 23.0 | 95.8 | 9.8 | 98.2 |
| 3 | 69.1 | 95.9 | 45.5 | 94.8 | 19.2 | 91.4 |
| 4 | 103.5 | 92.4 | 68.2 | 94.7 | 28.5 | 89.0 |
| 5 | 138.2 | 89.2 | 90.8 | 90.8 | 38.0 | 94.9 |
| 6 | 172.9 | 87.3 | 113.6 | 90.2 | 47.3 | 87.5 |
| 7 | 207.8 | 84.8 | 136.6 | 88.7 | 56.6 | 89.8 |
| 8 | 242.9 | 84.3 | 159.7 | 86.8 | 66.1 | 82.6 |
| 9 | 278.1 | 83.5 | 182.8 | 84.6 | 75.7 | 84.1 |
| 10 | 313.5 | 82.5 | 206.2 | 85.9 | 85.2 | 85.2 |

## REFERENCES


1. Bechhofer R.E. (1954), "A single-sample multiple decision procedure for ranking means of normal populations with known variances", *Ann. Math. Statist.*, 25, 16-39.
2. Fiacco, A.V. and McCormick, G.P. (1968). *Nonlinear Sequential Unconstrained Minimization Techniques,* John Wiley and Sons, Inc., New York.
3. Gibbons, J.D., Olkin, I. and Sobel, M. (1977)., *Selecting and Ordering Populations: A New Statistical Methodology,* John Wiley and Sons, Inc., New York.
4. Gupta, S.S. (1963). "On a selection and ranking procedure for gamma populations", *Ann. Inst. Statist. Math.*, 7, 225-245.
5. Wilk, M.B., Gnanadesikan, R. and Huyette, M.T. (1962). "Estimation of parameters of the gamma distribution using order statistics", *Biometrika,* 49, 525-546.