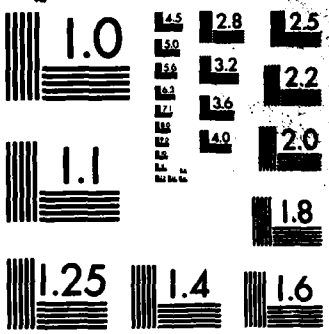
AD-A166 138 ON THE SELECTION OF THE BEST GAMMA POPULATION DETERMINATION OF MINIMAX SA (U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS N DAILAMI ET AL 1/1

UNCLASSIFIED NOV 85 RR-85-42 AFOSR-TR-86-0066 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 86-0066

ON THE SELECTION OF BEST GAMMA POPULATION: DETERMINATION OF MINIMAX SAMPLE SIZES*

by

N. Dailami
M. Bhaskara Rao
Department of Probability & Statistics

and

K. Subramanyam
Department of Mathematics & Statistics
University of Pittsburgh

AD-A166 138

Center for Multivariate Analysis

University of Pittsburgh

DTIC FILE COPY

DTIC
ELECTE
APR 1 1986
S D
A

Approved for public release;
distribution unlimited.

86 4 1 011

ON THE SELECTION OF BEST GAMMA POPULATION: DETERMINATION OF MINIMAX SAMPLE SIZES*

by

N. Dailami
M. Bhaskara Rao
Department of Probability & Statistics

and

K. Subramanyam
Department of Mathematics & Statistics
University of Pittsburgh

November 1985

Technical Report 85-42

DTIC
ELECTE
APR 1 1986
S D
A

Center for Multivariate Analysis
Fifth Floor, Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12.
Distribution is unlimited.
MATTHEW J. KERPER
Chief, Technical Information Division

*The work of K. Subramanyam and part of the work of M. Bhaskara Rao was done at the Center for Multivariate Analysis, University of Pittsburgh. It was supported partially by Contract F49620-85-C-008 of the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

(A)

ON THE SELECTION OF BEST GAMMA POPULATION: DETERMINATION OF MINIMAX SAMPLE SIZES

by

N. Dailami
M. Bhaskara Rao
Department of Probability & Statistics
The University, Sheffield

and

K. Subramanyam
Department of Mathematics & Statistics
University of Pittsburgh

ABSTRACT

Selecting the best Gamma population from a given set of Gamma populations is treated from a decision theoretic point of view. Cost of sampling and penalties for wrong decision play a role in the determination of optimum common sample sizes. Minimax sample sizes are determined under two different penalty functions.

AMS 1980 subject classification: primary: 62F07, secondary: 62C20.

Key words and phrases. Selection problems, Gamma distribution, cost of sampling, penalty functions, minimax criterion, minimax sample size.

ON THE SELECTION OF BEST GAMMA POPULATION: DETERMINATION OF MINIMAX SAMPLE SIZES

N. Dailami
M. Bhaskara Rao

Department of Probability & Statistics
The University
Sheffield S3 7RH
UK

and

K. Subramanyam

Department of Mathematics & Statistics
University of Pittsburgh
Pittsburgh, PA 15260

Availability Codes
Avail and/or
Dist Special
A-1

1. Introduction

Let $\pi_1, \pi_2, \ldots, \pi_k$ be k Gamma populations with unknown scale parameters $\theta_1, \theta_2, \ldots, \theta_k$ respectively and a common known shape parameters $c > 0$. The probability density function of the i-th population π_i $(i = 1,2,\ldots,k)$ is given by

$$f_i(x) = \theta_i^{-c}(\Gamma(c))^{-1} \exp(-\frac{x}{\theta_i})x^{c-1}, \quad 0 < x < \infty.$$

The main objective of this article is to describe a method of selecting that Gamma population with the least θ-value. This problem has considerable bearing on the problem of selecting the best of several normal populations in the sense of selecting that normal population with least variance. Basically, any selection problem consists of two components.

1. Draw a random sample of size n_i from population π_i, $i = 1,2,\ldots,k$.
2. Suggest a statistical procedure R which once the data are given clearly spells out the best population.

In many selection problems, a natural statistical procedure R manifests itself. The real problem is the selection of the sample sizes. The sample sizes are determined following certain optimality criterion. Following the lead given by Somerville (1954) and Ofosu (1972), we determine the sample sizes taking into account the cost of sampling and penalties imposed when wrong decisions are taken. The nature of the determination of the sample sizes is decision-theoretic in character and we adopt the minimax criterion as the optimality criterion. Ofosu (1972) has studied this problem under a particular penalty function and determined the optimal sample sizes using minimax principle. But his arguments have a gap and it is debatable whether the sample sizes he has given are really optimal sample sizes.

In this paper, we consider two types of penalty functions one of which is slightly less general than the one introduced by Ofosu (1972). Under these two penalty functions, we minimize the resultant loss functions over the entire parameter space for every fixed common sample size n. The solution to this minimization problem works out quite explicitly overcoming the gap present in the paper by Ofosu (1972).

For a general introduction to ranking and selection problems, the reader may refer to Gibbons, Olkin and Sobel (1977) or Gupta and Panchapakesan (1979).

2. Preliminaries

If $X_1, X_2, \ldots, X_n$ is a random sample of size n from a Gamma population with scale parameter θ and shape parameter c, then $\sum_{i=1}^{n} X_i$ is a sufficient statistic for θ for known c. Let $T = T(X_1, X_2, \ldots, X_n) = n^{-1} \sum_{i=1}^{n} X_i$. Then $ET = c\theta$ and T has a Gamma distribution with scale parameter θn and shape parameter cn. For the problem mentioned in the introduction, suppose that we wish to draw random samples of same size from each population. Let X_{ij}, $j = 1, 2, \ldots, n$ be a random sample of size n

from population π_i, $i=1,2,\ldots,k$. Let $T_i = T(X_{i1},X_{i2},\ldots,X_{in})$, $i=1,2,\ldots,k$. Now, a natural statistical procedure manifests itself for selecting the best Gamma population.

Statistical Procedure R

Declare population π_i $(i=1,2,\ldots,k)$ to be the best if $T_i < T_j$ for all $j \neq i$.

Cost of sampling

We assume that the cost $C(n)$ of drawing a random sample of size n from all populations to be of the following form

$$C(n) = c_o + c_1 n^d, \quad n=1,2,\ldots,$$

where c_o, c_1 and d are non-negative constants. c_o and c_1 are measured in the same units, and c_o represents fixed administrative costs involved in setting up a sampling plan. If $d=1$, the cost of taking additional samples rises linearly with n. If $d<1$, the rise in the cost does not increase relatively with increasing sample sizes. If $d>1$, it will become more and more expensive to take additional samples.

Wrong Decisions

Let $\theta_1,\theta_2,\ldots,\theta_k$ be a configuration of the parameters of the k populations $\pi_1,\pi_2,\ldots,\pi_k$ respectively. Let $\theta_{(1)} \leq \theta_{(2)} \leq \cdots \leq \theta_{(k)}$ be the ordered arrangement of $\theta_1,\theta_2,\ldots,\theta_k$. Let $\pi_{i_1},\pi_{i_2},\ldots,\pi_{i_k}$ correspond to the parameters $\theta_{(1)},\theta_{(2)},\ldots,\theta_{(k)}$ respectively. We follow the usual convention that if two or more θ-values coincide the corresponding i_j's are taken to be in increasing order of magnitude. Let T_{i_j} correspond to π_{i_j} for $j=1,2,\ldots,k$. According to the configuration $\theta_1,\theta_2,\ldots,\theta_k$, π_{i_1} is the best population. There are k-1 different ways of going wrong. These are listed below.

Type of wrong decision	Description	Event
2	π_{i_2} is declared to be the best.	$\{T_{i_2} < T_{i_j},\ j \neq 2\}$
3	π_{i_3} is declared to be the best.	$\{T_{i_3} < T_{i_j},\ j \neq 3\}$
...	... ...	... .
k	π_{i_k} is declared to be the best.	$\{T_{i_k} < T_{i_j},\ j \neq k\}$

Correct decision

Description: π_{i_1} is declared to be the best.

Event: $\{T_{i_1} < T_{i_j}$ for all $j = 2,3,\ldots,k\}$.

Calculation of probabilities of wrong and correct decisions.

Let $b = cn$, $g_b(\cdot)$ the probability density function of a Gamma random variable with scale parameter equal to unity and shape parameter equal to b and $G_b(\cdot)$ the distribution function associated with $g_b(\cdot)$. Note that $nT_{i_j}/\theta_{(j)}$ has Gamma distribution with scale parameter equal to unity and shape parameter equal to b. Let

$$p_r(\theta_1,\theta_2,\ldots,\theta_k;\ n)$$

$$= Pr_{\theta_1,\theta_2,\ldots,\theta_k}\ (\text{committing wrong decision of type } r)$$

$$= Pr_{\theta_1,\theta_2,\ldots,\theta_k}\ (T_{i_r} < T_{i_j} \quad \text{for all } j \neq r)$$

$$= Pr_{\theta_1,\theta_2,\ldots,\theta_k}(nT_{i_r}/\theta_{(r)} < (nT_{i_j}/\theta_{(j)})(\theta_{(j)}/\theta_{(r)}),\ j \neq r)$$

$$= \int_0^\infty Pr_{\theta_1,\theta_2,\ldots,\theta_k}\ (nT_{i_j}/\theta_{(j)} > (\theta_{(r)}/\theta_{(j)})x \quad \text{for all } j \neq r) g_b(x)\,dx$$

since T_{i_r} is independent of T_{i_j}'s, $j \neq r$.

$$= \int_0^\infty \prod_{\substack{j=1 \\ j\neq r}}^{k} (1 - G_b((\theta_{(r)}/\theta_{(j)})x))\ g_b(x)\ dx, \tag{2.1}$$

since T_i's are independent,

for $r = 2,3,\ldots,k$.

Let

$$p_1(\theta_1,\theta_2,\ldots,\theta_k;\ n)$$

$$= Pr_{\theta_1,\theta_2,\ldots,\theta_k}(\text{R takes the correct decision})$$

$$= Pr_{\theta_1,\theta_2,\ldots,\theta_k}(T_{i_1} < T_{i_j} \quad \text{for all } j = 2,3,\ldots,k)$$

$$= \int_0^\infty \prod_{j=2}^{k} (1 - G_b((\theta_{(1)}/\theta_{(j)})x))\ g_b(x)\ dx. \tag{2.2}$$

<u>Penalty functions</u>

We consider two types of penalty functions.

<u>Penalty functions of type 1</u>

Let $\Theta = \{(\theta_1,\theta_2,\ldots,\theta_k);\theta_i > 0 \text{ for all } i\}$.

For every $(\theta_1,\theta_2,\ldots,\theta_k)$ in Θ and $j = 2,3,\ldots,k$, let

$$W_j(\theta_1,\theta_2,\ldots,\theta_k)$$

= Penalty for taking wrong decision of type j when the configuration of the parameters is $(\theta_1,\theta_2,\ldots,\theta_k)$

$$= c_2 \log(\theta_{(2)}/\theta_{(1)})$$

for some $c_2 > 0$ which is measured in the same units as those of c_o and c_1.

The rationale behind these penalty functions is explained as follows. A critical point in the analysis of selection problems is the ability to discriminate between the best population π_{i_1} with the parameter value $\theta_{(1)}$ and the second best population π_{i_2} with the parameter value $\theta_{(2)}$. If we accept the second best population as the best, the penalty is $c_2 \log(\theta_{(2)}/\theta_{(1)})$. The penalty for other types of wrong decisions should be at least the one as the above, and for mathematical expediency, we take this penalty to be exactly the same as above. The loss function now works out to be

$$L(\theta_1,\theta_2,\ldots,\theta_k;\ n)$$

$$= \text{Cost of sampling} + \text{expected penalty}$$

$$= C(n) + c_2 \sum_{j=2}^{k} (\log(\theta_{(2)}/\theta_{(1)}) p_j(\theta_1,\theta_2,\ldots,\theta_k;\ n)$$

for all $(\theta_1,\theta_2,\ldots,\theta_k)$ in Θ and n in N, where N is the set of all natural numbers.

(Ofosu (1972) has taken $W_j(\theta_1,\theta_2,\ldots,\theta_k)$ to be equal to $c_2 \log(\theta_{(j)}/\theta_{(1)})$, $j = 2,3,\ldots$ This is more general than the one we considered above. Mathematically, with this choice of penalty functions, the loss function seems to be intractable to carry out optimization.)

The loss function introduced above is defined over the cartesian product space $\Theta \times N$. In order to find the minimax sample size n, we maximize $L(\theta_1,\theta_2,\ldots,\theta_k;\ n)$ over all $(\theta_1,\theta_2,\ldots,\theta_k)$ in Θ for every fixed n in N. Then the minimax sample size is obtained by minimizing

$$\max_{\underline{\theta}\in\Theta} L(\underline{\theta};n)$$

over all n in N. We will take up this work in Section 3.

Penalty functions of type 2

Let $0 < \delta < 1$ and $a > 0$ be prescribed. The constant a is measured in the same units as those of c_o, c_1 and c_2. We consider the following penalty functions. For $(\theta_1, \theta_2, \ldots, \theta_k)$ in Θ and $j = 2,3,\ldots,k$, let

$$W_j(\theta_1, \theta_2, \ldots, \theta_k) = a \quad \text{if } \theta_{(1)}/\theta_{(j)} \leq \delta,$$

$$= 0 \quad \text{if } \theta_{(1)}/\theta_{(j)} > \delta.$$

The rationale behind these penalty functions is as follows. Let $\theta_1, \theta_2, \ldots, \theta_k$ be a configuration of the parameters of the populations $\pi_1, \pi_2, \ldots, \pi_k$ respectively. Suppose the statistical procedure R declares π_{i_j} to be the best population for some $j = 2,3,\ldots,k$. The parameter values associated with the best population π_{i_1} and the population π_{i_j} are $\theta_{(1)}$ and $\theta_{(j)}$ respectively. If the ratio $\theta_{(1)}/\theta_{(j)}$ (≤ 1) of $\theta_{(1)}$ and $\theta_{(j)}$ is close to unity, we would not like to be penalized for taking the wrong decision of accepting π_{i_j} as the best population. On the other hand, if the ratio $\theta_{(1)}/\theta_{(j)}$ is small, we certainly wish to be penalized for accepting π_{i_j} as the best. A line has to be drawn somewhere between the statements that the ratio $\theta_{(1)}/\theta_{(j)}$ being close to unity and that it is being small. The number δ distinguishes these two statements and the choice of δ is subjective. We assume that the penalty for accepting the population π_{i_j} $(j = 2,3,\ldots,k)$ as the best when the ratio $\theta_{(1)}/\theta_{(j)}$ is small, i.e., $\leq \delta$, is the same constant a. The loss function then works out to be

$$L(\theta_1, \theta_2, \ldots, \theta_k; n) = C(n) + \sum_{j=2}^{k} W_j(\theta_1, \theta_2, \ldots, \theta_k) p_j(\theta_1, \theta_2, \ldots, \theta_k; n)$$

for $(\theta_1, \theta_2, \ldots, \theta_k)$ in Θ and n in N.

For a given configuration $(\theta_1,\theta_2,\ldots,\theta_k)$ in Θ, one could explicitly calculate the loss. In order to find the minimax sample size, we need to maximize the loss function over the parameter space Θ for every fixed n in N. We will take up this problem in the next section.

3. Minimax Sample Size

In this section, we maximize the loss function $L(\cdot: n)$ over the parameter space Θ for every fixed n in N under each of the penalty functions of types 1 and 2. First, we take up the case of penalty functions of type 1. The following subset of Θ plays a crucial role in the maximization problem of the loss function.

$$\Theta_o = \{(\theta_1,\theta_2,\ldots,\theta_k) \in \Theta;\ \theta_1 = r^{-1}\theta_2,\ \theta_2 = \theta_3 = \ldots = \theta_k \text{ for some } r \geq 1\}.$$

(In the literature, Θ_o is the so-called set of least favourable choices.)

The following chain of results helps to evaluate the maximum of the loss function over Θ explicitly.

Lemma 3.1 For each $(\theta_1,\theta_2,\ldots,\theta_k) \in \Theta_o$ with $\theta_1 = r^{-1}\theta_2$,

$$p_1(\theta_1,\theta_2,\ldots,\theta_k;\ n) = \int_0^\infty (1 - G_b(x/r))^{k-1} g_b(x)dx,$$

and

$$p_j(\theta_1,\theta_2,\ldots,\theta_k;\ n) = \int_0^\infty (1 - G_b(xr))(1 - G_b(x))^{k-2} g_b(x)dx$$

for every $j = 2,3,\ldots,k$.

Further,

$$\max_{\underline{\theta} \in \Theta_o} p_1(\underline{\theta};n) = \max_{r \geq 1} \int_0^\infty (1 - G_b(x/r))^{k-1} g_b(x)dx.$$

Proof. By direct verification. (Use (2.1) and (2.2).)

<u>Lemma 3.2</u> Let $(\theta_1,\theta_2,\ldots,\theta_k)\,\varepsilon\,\Theta$. Write $\theta_{(1)} = r^{-1}\theta_{(2)}$. (Obviously, $r \geq 1$.) Let $\theta_1^* = \theta_{(1)}$ and $\theta_2^* = \theta_3^* = \ldots = \theta_k^* = \theta_{(2)}$. Then $(\theta_1^*,\theta_2^*,\ldots,\theta_k^*)\,\varepsilon\,\Theta_o$ and

$$p_1(\theta_1,\theta_2,\ldots,\theta_k;\ n) \geq p_1(\theta_1^*,\theta_2^*,\ldots,\theta_k^*;\ n).$$

<u>Proof</u>. From (2.2)

$$p_1(\theta_1,\theta_2,\ldots,\theta_k;\ n) = \int_0^\infty \prod_{j=2}^{k} (1 - G_b((\theta_{(1)}/\theta_{(j)})x))g_b(x)dx$$

$$\geq \int_0^\infty (1 - G_b(x/r))^{k-1} g_b(x)dx,$$

since $1/r = \theta_{(1)}/\theta_{(2)} \geq \theta_{(1)}/\theta_{(3)} \geq \cdots \geq \theta_{(1)}/\theta_{(k)}$.

It is obvious that $(\theta_1^*,\theta_2^*,\ldots,\theta_k^*)\,\varepsilon\,\Theta_o$. By Lemma 3.1,

$$p_1(\theta_1^*,\theta_2^*,\ldots,\theta_k^*;\ n) = \int_0^\infty (1 - G_b(x/r))^{k-1} g_b(x)dx.$$

This completes the proof.

The following theorem simplifies the problem of maximization mentioned earlier.

<u>Theorem 3.3</u> For every n in N,

$$\max_{\underline{\theta}\varepsilon\Theta} L(\underline{\theta};\ n) = \max_{\underline{\theta}\varepsilon\Theta_o} L(\underline{\theta};\ n)$$

$$= C(n) + c_2 \max_{r\geq 1} (\log r)\left(1 - \int_0^\infty (1 - G_b(x/r))^{k-1} g_b(x)dx\right)$$

$$= C(n) + c_2(k-1) \max_{r\geq 1} (\log r) \int_0^\infty (1 - G_b(xr))(1 - G_b(x))^{k-2} g_b(x)dx.$$

<u>Proof</u> Let

$$p = \max_{\underline{\theta}\varepsilon\Theta} \sum_{j=2}^{k} \left(\log \frac{\theta_{(2)}}{\theta_{(1)}}\right) p_j(\theta_1,\theta_2,\ldots,\theta_k;n)$$

and

$$q = \max_{\underline{\theta}^*\varepsilon\Theta_o} \sum_{j=2}^{k} \left(\log \frac{\theta^*_{(2)}}{\theta^*_{(1)}}\right) p_j(\theta_1^*,\theta_2^*,\ldots,\theta_k^*;n).$$

To prove the first part of the theorem, it suffices to show that $p = q$. Since $\Theta_0 \subset \Theta$, $q \leq p$. We now prove that $p \leq q$. For this, it suffices to show the following. Given $(\theta_1, \theta_2, \ldots, \theta_k) \in \Theta$, there exists $(\theta_1^*, \theta_2^*, \ldots, \theta_k^*) \in \Theta_o$ such that

$$\sum_{j=2}^{k} (\log \frac{\theta_{(2)}}{\theta_{(1)}}) p_j(\theta_1, \theta_2, \ldots, \theta_k; n)$$
$$\leq \sum_{j=2}^{k} (\log \frac{\theta^*_{(2)}}{\theta^*_{(1)}}) p_j(\theta_1^*, \theta_2^*, \ldots, \theta_k^*; n).$$

Let $\theta_1^* = \theta_{(1)}$, $\theta_2^* = \theta_3^* = \ldots = \theta_k^* = \theta_{(2)}$. Obviously, $\theta^*_{(2)} = r\ \theta^*_{(1)}$ for some $r \geq 1$ and $(\theta_1^*, \theta_2^*, \ldots, \theta_k^*) \in \Theta_o$. Also,

$$\sum_{j=2}^{k} \log(\frac{\theta_{(2)}}{\theta_{(1)}}) p_j(\theta_1, \theta_2, \ldots, \theta_k; n)$$

$$= \sum_{j=2}^{k} (\log r)\, p_j(\theta_1, \theta_2, \ldots, \theta_k; n)$$

$$= (\log r)\,(1 - p_1(\theta_1, \theta_2, \ldots, \theta_k; n)$$

$$\leq (\log r)\ \ (1 - p_1(\theta_1^*, \theta_2^*, \ldots, \theta_k^*; n),$$

by Lemma 3.2

$$\leq (\log r) \sum_{j=2}^{k} p_j(\theta_1^*, \theta_2^*, \ldots, \theta_k^*; n)$$

$$= \sum_{j=2}^{k} \log(\frac{\theta^*_{(2)}}{\theta^*_{(1)}}) p_j(\theta_1^*, \theta_2^*, \ldots, \theta_k^*; n)$$

This shows that $p = q$ and hence $\max_{\underline{\theta} \in \Theta} L(\underline{\theta}; n) = \max_{\underline{\theta} \in \Theta_o} L(\underline{\theta}; n)$. Lemma 3.1 provides the exact expressions for $\max_{\underline{\theta} \in \Theta_o} L(\underline{\theta}; n)$.

In view of the above theorem, the critical function in the determination of the minimax sample size is the following one:

$$f(r) = (\log r)\,(1 - \int_0^\infty (1 - G_b(x/r))^{k-1} g_b(x)dx, \quad r \geq 1,$$

where $b = cn$.

We undertake the study of the above function in Section 4. Now, we work out the maximum of the loss function $L(\underline{\theta};n)$ under the penalty functions of type 2.

Let $0 < \delta < 1$ and $a > 0$ be fixed. We want to maximize $L(\underline{\theta};n)$ over all $\underline{\theta} \in \Theta$ for each sample size n under the penalty functions of type 2. See Section 2. For this purpose, we introduce the following subset of Θ.

$$\Theta_1 = \{(\theta_1,\theta_2,\ldots,\theta_k) \in \Theta \; ; \; \frac{1}{\delta}\theta_1 = \theta_2 = \ldots = \theta_k\}.$$

The following theorem works out explicitly the desired maximum of the loss function.

<u>Theorem 3.4</u> The following statements are true.

(1) If $(\theta_1,\theta_2,\ldots,\theta_k) \in \Theta_1$, then

$$p_1(\theta_1,\theta_2,\ldots,\theta_k;n) = \int_0^\infty (1 - G_b(\delta x))^{k-1} g_b(x)dx.$$

(2) If $(\theta_1,\theta_2,\ldots,\theta_k) \in \Theta_1$, then

$$\begin{aligned} &L(\theta_1,\theta_2,\ldots,\theta_k;n) \\ &\quad = C(n) + a\sum_{j=2}^{k} p_j(\theta_1,\theta_2,\ldots,\theta_k;n) \\ &\quad = C(n) + a(1 - p_1(\theta_1,\theta_2,\ldots,\theta_k;n)). \end{aligned}$$

(3)
$$\begin{aligned} &\max_{(\theta_1,\theta_2,\ldots,\theta_k)\in\Theta} L(\theta_1,\theta_2,\ldots,\theta_k;n) \\ &\quad = \max_{(\theta_1,\theta_2,\ldots,\theta_k)\in\Theta_1} L(\theta_1,\theta_2,\ldots,\theta_k;n) \\ &\quad = C(n) + a(1 - \int_0^\infty (1 - G_b(\delta x))^{k-1} g_b(x)dx). \end{aligned}$$

(Recall that $b = cn$.)

<u>Proof</u>.

(1) If $(\theta_1,\theta_2,\ldots,\theta_k)\in\Theta_1$, then $\frac{1}{\delta}\theta_1=\theta_2=\theta_3=\ldots=\theta_k$. So, $\theta_{(1)}=\theta_1$, $\theta_{(2)}=\theta_2,\ldots,\theta_{(k)}=\theta_k$ according to our convention. Therefore, from (2.2),

$$P_1(\theta_1,\theta_2,\ldots,\theta_k;n)=\int_0^\infty \prod_{j=2}^{k}(1-G_b(\frac{\theta_{(1)}}{\theta_{(j)}}x))g_b(x)dx$$

$$=\int_0^\infty (1-G_b(\delta x))^{k-1}g_b(x)dx.$$

(2) If $(\theta_1,\theta_2,\ldots,\theta_k)\in\Theta_1$, then $W_j(\theta_1,\theta_2,\ldots,\theta_k)=a$ for all $j=2,3,\ldots,k$. Consequently,

$$L(\theta_1,\theta_2,\ldots,\theta_k;n)=C(n)+\sum_{j=2}^{k}W_j(\theta_1,\theta_2,\ldots,\theta_k)P_j(\theta_1,\theta_2,\ldots,\theta_k;n)$$

$$=C(n)+a\sum_{j=2}^{k}P_j(\theta_1,\theta_2,\ldots,\theta_k;n)$$

$$=C(n)+a(1-P_1(\theta_1,\theta_2,\ldots,\theta_k;n))$$

$$=C(n)+a(1-\int_0^\infty (1-G_b(\delta x))^{k-1}g_b(x)dx),$$

by (1) above.

(3) Since $\Theta_1\subset\Theta$,

$$p=\max_{(\theta_1,\theta_2,\ldots,\theta_k)\in\Theta}L(\theta_1,\theta_2,\ldots,\theta_k;n)$$

$$\geq\max_{(\theta_1,\theta_2,\ldots,\theta_k)\in\Theta_1}L(\theta_1,\theta_2,\ldots,\theta_k;n)=q,\text{ say.}$$

From (2) above, $q=C(n)+a(1-\int_0^\infty (1-G_b(\delta x))^{k-1}g_b(x)dx)$. To complete the proof, it is enough to show that $p\leq q$. For this, we show that for any $(\theta_1,\theta_2,\ldots,\theta_k)$ in Θ,

$$L(\theta_1,\theta_2,\ldots,\theta_k;n)\leq q.$$

Let $(\theta_1,\theta_2,\ldots,\theta_k) \in \Theta$. Let $\theta_{(1)} \leq \theta_{(2)} \leq \cdots \leq \theta_{(k)}$ be the ordered arrangement of $\theta_1,\theta_2,\ldots,\theta_k$ in increasing order of magnitude.

<u>Case (i)</u>. $\frac{\theta_{(1)}}{\theta_{(j)}} > \delta$ for all $j = 2,3,\ldots,k$. Then

$$L(\theta_1,\theta_2,\ldots,\theta_k;n) = C(n) + 0 \leq q, \quad \text{obviously.}$$

<u>Case (ii)</u>. There exists j in $\{2,3,\ldots,k\}$ such that $\frac{\theta_{(1)}}{\theta_{(j)}} \leq \delta$. Let j^* be the smallest integer in $\{2,3,\ldots,k\}$ such that $\frac{\theta_{(1)}}{\theta_{(j^*)}} \leq \delta$. This means that

$$\frac{\theta_{(1)}}{\theta_{(i)}} > \delta \qquad \text{for all} \quad i = 1,2,\ldots,j^*-1$$

and

$$\frac{\theta_{(1)}}{\theta_{(i)}} \leq \delta \qquad \text{for all} \quad i = j^*, j^*+1,\ldots,k.$$

Then

$$\begin{aligned} L(\theta_1,\theta_2,\ldots,\theta_k;n) &= C(n) + a \sum_{j=j^*}^{k} P_j(\theta_1,\theta_2,\ldots,\theta_k;n) \\ &\leq C(n) + a \sum_{j=2}^{k} P_j(\theta_1,\theta_2,\ldots,\theta_k;n) \\ &\leq C(n) + a(1 - P_1(\theta_1,\theta_2,\ldots,\theta_k;n)). \end{aligned}$$

Let us calculate $P_1(\theta_1,\theta_2,\ldots,\theta_k;n)$. Note that

$$\frac{\theta_{(1)}}{\theta_{(j)}} \leq 1 \quad \text{for} \quad j = 2,3,\ldots,j^*-1$$

and

$$\frac{\theta_{(1)}}{\theta_{(j)}} \leq \delta \quad \text{for } j = j^*,j^*+1,\ldots,k.$$

Therefore,

$$\frac{\theta_{(1)}}{\theta_{(j)}} x \leq x \quad \text{for} \quad j = 2,3,\ldots,j^*-1$$

and

$$\frac{\theta_{(1)}}{\theta_{(j)}} x \leq \delta x \quad \text{for} \quad j = j^*, j^*+1, \ldots, k$$

for all positive x.

Consequently,

$$\begin{aligned}
&P_1(\theta_1,\theta_2,\ldots,\theta_k;n) \\
&\quad = \int_0^\infty \prod_{j=2}^{k} (1 - G_b(\frac{\theta_{(1)}}{\theta_{(j)}} x)) g_b(x)dx \\
&\quad = \int_0^\infty \prod_{j=2}^{j^*-1} (1 - G_b(\frac{\theta_{(1)}}{\theta_{(j)}} x)) \prod_{j=j^*}^{k} (1 - G_b(\frac{\theta_{(1)}}{\theta_{(j)}} x)) g_b(x)dx \\
&\quad \geq \int_0^\infty (1 - G_b(x))^{j^*-2} (1 - G_b(\delta x))^{k-j^*+1} g_b(x)dx.
\end{aligned}$$

From the above inequality, we obtain

$$\begin{aligned}
&L(\theta_1,\theta_2,\ldots,\theta_k;n) \\
&\quad \leq C(n) + a(1 - P_1(\theta_1,\theta_2,\ldots,\theta_k;n)) \\
&\quad \leq C(n) + a(1 - \int_0^\infty (1 - G_b(x))^{j^*-2} (1-G_b(\delta x))^{k-j^*+1} g_b(x)dx).
\end{aligned}$$

Now, we show that

$$\begin{aligned}
&\int_0^\infty (1 - G_b(x))^{j^*-2} (1 - G_b(\delta x))^{k-j^*+1} g_b(x)dx \\
&\geq \int_0^\infty (1 - G_b(\delta x))^{k-1} g_b(x)dx.
\end{aligned}$$

Since $0 < \delta < 1$, $G_b(\delta x) < G_b(x)$ for all positive x from which it follows that $(1 - G_b(x))^{j^*-2} < (1 - G_b(\delta x))^{j^*-2}$ for all positive x. This inequaltiy yields

$$L(\theta_1,\theta_2,\ldots,\theta_k;n)$$

$$\leq C(n) + a(1 - \int_0^\infty (1 - G_b(x))^{j^*-2}(1 - G_b(\delta x))^{k-j^*+1} g_b(x)dx$$

$$\leq C(n) + a(1 - \int_0^\infty (1 - G_b(\delta x))^{k-1} g_b(x)dx) = q.$$

This completes the proof.

In the case of penalty functions of type 2, the maximum of the loss function $L(\underline{\theta};n)$ over Θ works out quite explicitly. In order to find the minimax sample size, we need to minimize this maximum with respect to n. We take up this problem in Section 4.

4. Numerical Illustrations

The selection of the best Gamma population has considerable bearing in the selection of the population which is least variable. Let $\pi_1,\pi_2,\ldots,\pi_k$ be k normal populations with i-th population having mean C_i and variance $\theta_i > 0$. The best population among these is the one with the smallest θ_i. If X_{ij}, $j = 1,2,\ldots,n$ is a ramdom sample of size n from π_i, then we could use the statistics $T_i = \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2$, $i = 1,2,\ldots,k$ to discriminate the populations $\pi_1,\pi_2,\ldots,\pi_k$. $T_i/2\theta_i$ has Gamma distribution with shape parameter equal to $(n-1)/2$ and scale parameter equal to unity. In this case, in order to find the minimax sample size, one has to maximize

$$L(r;n) = C(n) + C_2(\log r)\,(1 - \int_0^\infty (1 - G_b(x/r))^{k-1} g_b(x)dx)$$

over $r \geq 1$ for every fixed $n = _,3,\ldots$, and then this maximum is to be minimized over the set $\{2,3,4,\ldots\}$, where $b = (n-1)/2$. In this section, we address ourselves to the problem of determining minimax sample sizes when $b = n/2$.

The critical function in the maximization of $L(r;n)$ over $r \geq 1$ is the function

$$f_{n,k}(r) = \log r(1 - \int_0^\infty (1 - G_b(x/r))^{k-1} g_b(x)dx), \quad r \geq 1,$$

where $b = n/2$. Determination of the maximum of this function analytically does not seem to be feasible. It is obvious that $f_{n,k}(1) = 0$ and one can prove that $\lim_{r\to\infty} f_{n,k}(r) = 0$. An extensive tabulation of this function over the domain of definition $[1,\infty)$ of r for various values of n and k suggests that this function is unimodal. These calculations yielded the following tables. (Let $r_{n,k}$ denote the number at which $f_{n,k}(r_{n,k})$ is maximum.)

TABLE 1

No. of populations k	2	3	4	5	6
sample size n	$r_{n,k}$	$r_{n,k}$	$r_{n,k}$	$r_{n,k}$	$r_{n,k}$
2	4.06	5.21	6.26	7.24	8.18
3	3.09	3.77	4.34	4.85	5.32
4	2.69	3.20	3.60	3.96	4.27
5	2.47	2.88	3.21	3.48	3.72
6	2.32	2.68	2.96	3.19	3.39
7	2.22	2.54	2.79	2.99	3.16
8	2.15	2.44	2.66	2.84	3.00
9	2.09	2.36	2.57	2.73	2.87
10	2.04	2.30	2.49	2.65	2.77
11	2.00	2.25	2.43	2.57	2.69
12	1.97	2.21	2.38	2.52	2.63
13	1.95	2.18	2.34	2.46	2.57
14	1.92	2.14	2.30	2.42	2.52
15	1.90	2.12	2.27	2.39	2.48
16	1.89	2.10	2.24	2.35	2.45
17	1.87	2.07	2.22	2.33	2.41
18	1.86	2.06	2.20	2.30	2.39
19	1.85	2.04	2.18	2.28	2.36
20	1.84	2.03	2.16	2.26	2.34

The following pattern emerges among $r_{n,k}$'s from the above table.

(i) $r_{m,k} \le r_{n,k}$ if $m \ge n$.

(ii) $r_{n,k} \le r_{n,s}$ if $k \le s$.

Using the property (i) mentioned above, we promulgate the following strategy to determine minimax sample sizes.

<u>Strategy</u>. k is fixed. Find $r_{2,k}$, the point at which $f_{2,k}(r)$ is maximum. Evaluate $L(r;n)$ for $r = 1.0, 1.1, 1.2, \ldots, r_{2,k}$ and $n = 2,3,\ldots,20$. These values are tabulated

in a two-way grid,rows corresponding to r and columns to n. For each column in the two-way grid, locate the maximum. Under each column representing a particular sample size n, we know that the maximum of L(r;n) occurs at some value of r between 1.0 and $r_{2,k}$. See property (i) above. Then the column for which this maximum is minimum is located. The corresponding sample size is the required minimax sample size.

If the column maximums are increasing right upto r = 20, then the minimax sample size is ≥ 20. In such a case, we need to include some more sample sizes n = 21,22,... etc. For largevalues of n, evaluation of L(r,n) may be beyond the reach of many computers. Later, in this section, we give asymptotic formulas for L(r;n).

The above method of determining minimax sample sizes is a modification of a procedure suggested by Ofosu (1972). This modification works faster than the methods outlined by Ofosu (1972).

By way of illustration of the working of the above strategy, we include the following examples.

1. No. of populations, k = 2,3,4,5,6.
2. $c_0 = 0$, $c_1 = 1$, $d = 1$ so that $C(n) = kn$.
3. $c_2 = 50(25)\ 500$.

The loss function then becomes

$$L(r;n) = kn + c_2(\log r)\left(1 - \int_0^\infty (1 - G_b(x/r))^{k-1} g_b(x)dx\right),$$

$$r \geq 1,\ n = 1,2,3,4,\ldots,\ \text{where } b = n/2.$$

The minimax sample sizes are given in the following Table.

Table 2

Minimax sample sizes (n) and minimax loss L(n) under penalty functions of type 1.

k	2		3		4		5		6	
c_2	n	L(n)	n	L(n)	n	L(n)	n	L(n)	n	L(n)
50	3	27.79	3	42.60	3	53.79	3	63.07	3	71.18
75	4	38.55	4	52.25	4	72.87	4	84.90	4	95.63
100	4	48.74	5	73.23	5	90.99	5	105.53	5	118.15
125	5	58.76	5	87.79	6	108.69	6	125.63	5	140.19
150	5	68.53	6	101.97	6	125.63	6	144.75	6	161.15
175	5	78.29	6	115.97	7	142.57	7	163.80	6	182.01
200	6	87.95	7	129.79	7	158.93	7	182.20	7	202.02
225	6	97.44	7	143.39	7	175.30	7	200.60	7	222.02
250	6	106.94	7	156.99	8	191.33	8	218.47	8	241.58
275	6	116.43	8	170.43	8	207.26	8	236.32	8	260.94
300	7	125.90	8	183.74	8	223.20	9	254.05	9	280.17
325	7	135.23	8	197.05	9	238.82	9	271.47	9	299.02
350	7	144.55	9	210.36	9	254.42	9	288.90	9	317.86
375	7	153.88	9	223.45	9	270.02	10	306.11	10	336.56
400	7	163.20	9	236.55	10	285.46	10	323.18	10	355.00
425	7	172.53	9	245.65	10	300.80	10	340.25	10	373.43
450	8	181.76	9	262.74	10	316.14	10	357.33	11	391.85
475	8	190.96	10	275.77	11	331.36	11	374.18	11	409.95
500	8	200.17	10	288.71	11	346.48	11	390.97	11	428.05

Now, we consider the case of penalty functions of type 2. From Theorem 3.4,

$$\max_{\underline{\theta}\in\Theta} L(\underline{\theta};n)$$

$$= C(n) + a(1 - \int_0^\infty (1 - G_b(\delta x))^{k-1} g_b(x)dx,$$

for $n = 1,2,3,\ldots$, $0 < \delta < 1$ and $b = cn$. In order to find the minimax sample size, we need to minimize the above function over $n = 1,2,3,\ldots$ for a given k, δ, c and $C(n) = kn$. As an illustration, we take $c = \frac{1}{2}$ and $a = 100(100)600$ and determine the minimax sample sizes for $\delta = 0.5$, 0.6, 0.7, 0.8, 0.9 and $k = 2,3,4,5$.

Table 3

Minimax sample sizes (n) and minimax loss function L(n) under penalty functions of type 2.

	k	2		3		4		5	
	a	n	L(n)	n	L(n)	n	L(n)	n	L(n)
δ=0.50	100	7	32.94	6	51.91	4	66.34	3	76.62
	200	12	48.42	13	77.74	12	103.05	10	124.94
	300	16	58.53	17	94.24	17	126.31	16	155.67
	400	18	66.50	20	106.85	20	143.55	20	177.86
	500	19	73.66	21	118.06	22	158.19	22	195.95
	600	20	80.40	22	128.63	23	171.90	23	212.43
	k	2		3		4		5	
	a	n	L(n)	n	L(n)	n	L(n)	n	L(n)
δ=0.60	100	5	39.51	4	59.71	3	72.64	3	82.13
	200	12	62.87	12	98.81	9	126.25	7	146.84
	300	18	79.27	19	126.09	16	165.13	13	197.58
	400	20	92.87	22	147.61	22	94.60	20	236.20
	500	21	105.62	24	167.39	24	220.04	24	267.55
	600	22	117.95	25	186.07	26	243.92	26	296.22
	k	2		3		4		5	
	a	n	L(n)	n	L(n)	n	L(n)	n	L(n)
δ=0.70	100	3	44.92	3	64.84	3	77.35	3	84.46
	200	9	78.27	7	117.35	4	142.40	3	157.92
	300	17	104.50	15	160.86	9	200.31	7	227.19
	400	21	126.52	22	196.51	17	249.92	11	289.45
	500	22	147.28	24	228.24	24	292.15	18	343.79
	600	23	167.53	26	258.89	26	330.52	25	391.25
	k	2		3		4		5	
	a	n	L(n)	n	L(n)	n	L(n)	n	L(n)
δ=0.80	100	2	48.64	2	68.73	2	78.72	2	86.38
	200	6	91.40	3	129.12	2	149.56	2	162.77
	300	9	129.59	7	186.78	3	219.63	2	239.15
	400	17	164.14	11	241.04	7	286.27	3	314.75
	500	21	195.64	17	291.00	9	350.36	7	387.85
	600	23	225.92	23	337.22	14	411.22	7	458.42

Table 3 (continued)

Minimax sample sizes (n) and minimax loss function L(n) under penalty functions of type 2.

	a \ k	2		3		4		5	
		n	L(n)	n	L(n)	n	L(n)	n	L(n)
	100	2	51.56	2	70.48	2	81.17	2	88.46
	200	2	99.13	2	134.96	2	154.33	2	166.91
$\delta=0.90$	300	3	146.30	2	199.44	2	227.50	2	245.37
	400	7	192.21	3	263.69	3	301.68	3	326.10
	500	7	236.76	4	327.34	3	374.10	2	402.27
	600	9	281.05	7	389.15	3	446.52	2	480.73

5. Asymptotic Minimax Value of n

In this section, we approximate the following integral

$$\int_0^\infty (1 - G_n(\tfrac{x}{r}))^{k-1} g_n(x)\,dx$$

for large values of n by an integral involving the standard normal distribution, where $g_n(\cdot)$ is the probability density function of a Gamma distribution with parameters n and 1, i.e.,

$$g_n(x) = \frac{1}{\Gamma(n)} e^{-x} x^{n-1}, \quad 0 < x < \infty,$$

and $G_n(\cdot)$ is the distribution function associated with $g_n(\cdot)$. This integral appears in the loss functions considered under penalty functions of types 1 and 2. This integral can be realized probabilistically as follows.

Let $Y_1, Y_2, \ldots, Y_k$ be k independent identically distributed random variables with common probability density function $g_n(\cdot)$ given above. Then for any $r \geq 1$,

$$\begin{aligned} P_k(r;n) &= \Pr(Y_1 < rY_j, \quad j = 2,3,\ldots,k) \\ &= \int_0^\infty \prod_{j=2}^{k} \Pr(rY_j > x)\; g_n(x)\,dx \\ &= \int_0^\infty (1 - G_n(\tfrac{x}{r}))^{k-1} g_n(x)\,dx. \end{aligned}$$

We observe that $2Y_j \sim x^2_{2n}$ for every $j = 1,2,\ldots,k$. We want to apply Wilson-Hilferty's transformation to each $2Y_j$. See Kendall and Stuart (1977, p. 398-399). Let

$$Z_j = \frac{(\frac{2Y_j}{2n})^{1/3} - (1 - \frac{1}{9n})}{(\frac{1}{9n})^{1/2}}$$

for $j = 1,2,\ldots,k$.

$Z_1, Z_2, \ldots, Z_k$ are independent random variables and for large n, each Z_j has standard normal distribution. We also note that

$$2Y_1 < 2rY_j$$

if and only if

$$\frac{(\frac{2Y_1}{2n})^{1/3} - (1 - \frac{1}{9n})}{(\frac{1}{9n})^{1/2}} < \frac{r^{1/3}(\frac{2Y_j}{2n})^{1/3} - (1 - \frac{1}{9n})}{(\frac{1}{9n})^{1/2}}$$

$$= \frac{r^{1/3}[(\frac{2Y_j}{2n})^{1/3} - (1 - \frac{1}{9n})]}{(\frac{1}{9n})^{1/2}} + \frac{(1 - \frac{1}{9n})(r^{1/3} - 1)}{(\frac{1}{9n})^{1/2}}$$

if and only if

$$Z_1 < r^{1/3} Z_j + a_{n,r}$$

for any $j = 2,3,\ldots,k$, where

$$a_{n,r} = \frac{(1 - \frac{1}{9n})(r^{1/3} - 1)}{(\frac{1}{9n})^{1/2}} .$$

Now, we are in a position to approximate the integral in focus as follows. For large n,

$$\int_0^\infty (1 - G_n(\frac{x}{r}))^{k-1} g_n(x)dx = P_k(r;n)$$

$$= Pr(Y_1 < rY_j,\ j = 2,3,\ldots,k)$$

$$= Pr(2Y_1 < 2rY_j,\ j = 2,3,\ldots,k)$$

$$= Pr(Z_1 < r^{1/3}Z_j + a_{n,r},\ j = 2,3,\ldots,k)$$

$$= \int_{-\infty}^{\infty} (1 - \Phi(\frac{x - a_{n,r}}{r^{1/3}}))^{k-1} \frac{1}{(2\pi)^{1/2}} \exp(-\frac{x^2}{2})dx,$$

where $\Phi(\cdot)$ is the distribution function of the standard normal probability model.

The loss function under penalty functions of type 1 for alrge n, becomes

$$L(r;n) = C(n) + c_2(\log r)\ (1 - \int_{-\infty}^{\infty} (1 - \Phi(\frac{x - a_{b,r}}{r^{1/3}}))^{k-1}$$

$$\frac{1}{(2\pi)^{1/2}} \exp(-\frac{x^2}{x})dx)$$

for $r \geq 1$ and n in N, where $b = cn$.

To find the minimax sample size n, the above approximation can be used in the strategy outlined in Section 4.

The maximum of the loss function over the entire parameter space under penalty functions of type 2, for large n, becomes

$$C(n) + a(1 - \int_{-\infty}^{\infty} (1 - (\frac{x - b_{d,\delta}}{(\frac{1}{\delta})^{1/3}}))^{k-1} \frac{1}{(2\pi)^{1/2}} e^{-\frac{x^2}{2}} dx)$$

for n in N, where $d = cn$ and

$$b_{d,\delta} = \frac{(1 - \frac{1}{9d})((\frac{1}{\delta})^{1/3} - 1)}{(\frac{1}{9d})^{1/2}} .$$

Similar approximations can be provided for the problem of selecting the best normal population with least variance.

REFERENCES

1. Gibbons, J.D., Olkin, I., and Sobel, M. (1977). Selecting and ordering populations - A new statistical methodology, John Wiley and Sons, New York.

2. Gupta, S..S and Panchapakesan, S. (1979). Multiple decision procedures, theory and methodology of selecting and ranking populations. John Wiley and Sons, New York.

3. Kendall, M.G. and Stuart, A. (1977). The Advanced Theory of Statistics, Volume I, Charles Griffin and Company, Ltd. London, Fourth edition.

4. Ofosu, J.B. (1972). A minimax procedure for selecting the population with the largest (smallest) scale parameter, Calcutta Statist. Assoc. Bull., 21, 143-154. Errata 22 (1973), 129.

5. Somerville, P.N. (1954). Some problems of optimum sampling, Biometrika, 41, 420-429.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 86-0066
2. GOVT ACCESSION NO.: AD-A166 138
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): On the selection of best gamma population: Determination of minimax sample sizes.

5. TYPE OF REPORT & PERIOD COVERED: Technical- November 1985

6. PERFORMING ORG. REPORT NUMBER: 85-42

7. AUTHOR(s): N. Dailami, M. Bhaskara Rao and K. Subramanyam

8. CONTRACT OR GRANT NUMBER(s): F49620-85-C-0008

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Center for Multivariate Analysis, 515 Thackeray Hall, University of Pittsburgh, Pittsburgh, PA

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F / 2304/A5

11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research, Department of the Air Force, Bolling Air Force DC 20332

12. REPORT DATE: November 1985

13. NUMBER OF PAGES: 24

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Cost of sampling, Gamma distribution, minimax criterion, minimax sample size, penalty functions, selection problems.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Selecting the best Gamma population from a given set of Gamma populations is treated from a decision theoretic point of view. Cost of sampling and penalties for wrong decision play a role in the determination of optimum common sample sizes. Minimax sample sizes are determined under two different penalty functions.

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END

DTIC

5-86