AN ELIMINATION TYPE TWO-STAGE PROCEDURE FOR SELECTING THE POPULATION WITH THE LARGEST MEAN FROM K LOGISTIC POPULATIONS

by

Shanti S. Gupta
Purdue University

SangHyun Han
South Korean Military Academy
Seoul, South Korea

Technical Report #87-39

Department of Statistics
Purdue University

August 1987
Revised June 1988
Revised May 1991

*This research was supported by the Office of Naval Research Contracts N00014-84-C-0167, N00014-88-K-0170, and NSF Grant DMS-8606964 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

AN ELIMINATION TYPE TWO-STAGE PROCEDURE FOR SELECTING THE POPULATION WITH THE LARGEST MEAN FROM k LOGISTIC POPULATIONS

Shanti S. Gupta
Purdue University
West Lafayette, IN

and

SangHyun Han
South Korean Military Academy
Seoul, South Korea

SYNOPTIC ABSTRACT

A formula for an approximation to the distribution of the sample means of a logistic population is obtained by using the Edgeworth series expansions. Using this approximation, we consider an elimination type two-stage procedure based on the sample means for selecting the population with the largest mean from k logistic populations when their common variance is known. A short table of the constants needed to implement this procedure is provided and the efficiency of this procedure relative to the single-stage procedure is investigated.

Key Words and Phrases: Two-Stage Procedure; Selection Procedure; Largest Mean; Subset Selection; Logistic Populations.

1. INTRODUCTION

For the problem of selecting the population having the largest mean from normal populations with a common known variance σ^2, Cohen (1959), Alam (1970) and Tamhane and Bechhofer (1977, 1979) have studied two-stage elimination type procedures, in which they used Gupta's (1956, 1965) subset selection procedure in the first stage to screen out non-contending populations and Bechhofer's (1954) indifference zone approach to select the best from among the populations in the second stage.

Tamhane and Bechhofer (1977, 1979) studied in depth a two-stage elimination type procedure ($\mathcal{P}_2'$) for selecting the largest normal mean when the common variance is known. In order to determine a set of constants necessary to implement $\mathcal{P}_2'$, they proposed a criterion of minimizing the maximum over the entire parameter space of the expected total sample size required by $\mathcal{P}_2'$ subject to the procedure guaranteeing a specified probability of a correct selection. As a consequence, $\mathcal{P}_2'$ based on this unrestricted minimax design criterion possesses the highly desirable property that the expected total sample size required by $\mathcal{P}_2'$ is always less than or equal to the total sample size required by the best competing single-stage procedure of Bechhofer (1954), regardless of the true configuration of the population means.

The logistic distribution has been widely used by Berkson (1944,1951,1957) as a model for analyzing experiments involving quantal response. Pearl and Reed (1920) used this in studies connected with population growth. Plackett (1958,1959) has considered the use of this distribution with life test data. Gupta (1962) has studied this distribution as a model in life testing problems.

The importance of the logistic distribution in the modeling of stochastic phenomena has resulted in numerous other studies involving probabilistic and statistical aspects of the distribution. For example, Gumbel (1944), Gumbel and Keeney (1950) and Talacko (1956) show that it arises as a limiting distribution in various situations; Birnbaum and Dudman (1963), Gupta and Shah (1965) study its order statistics. Many other authors, for example, Antle, Klimko and Harkness (1970), Gupta and Gnanadesikan (1966) and Tarter and Clark (1965), investigate inference questions about its parameters.

In this paper we consider an elimination type two-stage procedure for selecting the logistic population with the largest population mean when the populations have a common known variance. Using an approximation to the distribution of the sample mean from a logistic population, we propose a two-stage elimination type procedure $\mathcal{P}_2$ and a non-linear optimization problem by using a minimax criterion to find a set of constants needed to implement $\mathcal{P}_2$. We derive lower bounds of the probability of a correct selection and the infimum over the preference zone of the lower bounds. We determine the supremum of the expected total sample size needed for $\mathcal{P}_2$ over the whole parameter space. We provide tables of constants to implement $\mathcal{P}_2$ and of the efficiency of $\mathcal{P}_2$ relative to the corresponding single-stage procedure $\mathcal{P}_1$ for the two special cases of equally spaced and slippage configurations.

2. DISTRIBUTION OF LOGISTIC SAMPLE MEANS

Because of the similarity between the logistic and the normal distributions, the sample mean and variance, the moment estimators of the logistic population parameters, are effective tools for statistical decisions involving the logistic distribution. Antle, Klimko and Harkness (1970) give a function of the sample mean as a confidence interval estimate of the population mean when the population variance is known. Schafer and Sheffield (1973) show that in terms of the mean squared error the moment estimators of the logistic population parameters are as good as their maximum likelihood estimators. The fact that the distribution of a sample mean has monotone likelihood ratio (MLR) with respect to the population mean when the variance is known is used by Goel (1975) to obtain a uniformly most accurate confidence interval for the population mean and a uniformly most powerful test for one-sided hypotheses involving the population mean. The sampling distribution of the mean is a primary requirement for these statistical purposes. The papers by Antle, Klimko and Harkness (1970), and Tarter and Clark (1965) used a Monte Carlo method for this distribution.

Goel (1975) obtained an expression for the distribution function of the sum of independent and identically distributed (*iid*) logistic variates by using the Laplace transform inverse method for convolutions of Pólya type functions, a technique developed by Schoenberg (1953) and Hirschman and Widder (1955). He provides a table of the cumulative distribution function (cdf) of the sum of *iid* logistic variates for sample size $n = 2(1)12$, $x = 0(0.01)3.99$, and $n = 13(1)15$, $x = 1.20(0.01)3.99$. George and Mudholkar (1983) obtained an expression for the distribution of a convolution of the *iid* logistic variables by directly inverting the characteristic function. However, since both formulas of Goel (1975) and George and Mudholkar (1983) contain terms $(1-e^x)^{-k}, k = 1, \ldots, n$, the problem of precision of the computation at values of x near zero arises when n is large. George and Mudholkar (1983) also show that a standardized Student's t distribution provides a very good approximation for the distribution of a convolution of *iid* logistic random variables compared to the standardized normal distribution and the Edgeworth series approximation terminating with the n^{-1} term.

In this section, we consider an approximation for the distribution of a standardized mean of samples from a logistic population by using Edgeworth series expansions including terms upto n^{-3}. Han (1987) has compared this with the Student's t distribution and found this approximation better (differing from the exact values of Goel by not more than 0.0001) for $n \geq 7$.

2.1. Logistic Distribution and Edgeworth Series Expansion for the Distribution of the Sample Mean. A random variable X has the logistic distribution with mean μ and variance σ^2, denoted by $L(\mu, \sigma^2)$, if the probability density function (pdf) of X is given by

$$f(x) = (g/\sigma)[\exp\{-g(x-\mu)/\sigma\}][1+\exp\{-g(x-\mu)/\sigma\}]^{-2} \tag{1}$$

with the cdf of X given by

$$F(x) = [1+\exp\{-g(x-\mu)/\sigma\}]^{-1}, \tag{2}$$

where $-\infty < x < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$ and $g = \pi/\sqrt{3}$. This distribution is symmetrical about the mean μ.

Letting $Y = (X-\mu)g/\sigma$, the random variable Y has the logistic distribution with mean zero and variance $\pi^2/3$. The pdf and cdf of the random variable Y are given by

$$f(y) = [\exp\{-y\}][1+\exp\{-y\}]^{-2} \tag{3}$$

and

$$F(y) = [1 + \exp\{-y\}]^{-1} \tag{4}$$

respectively, where $-\infty < y < \infty$. (3) may be written in terms of $F(y)$ as

$$f(y) = F(y)(1 - F(y)). \tag{5}$$

The moment generating function (mgf) of Y is given by

$$\begin{aligned} M_Y(t) &= \Gamma(1+t)\Gamma(1-t) \\ &= \pi t / \sin \pi t, \ |t| < 1. \end{aligned} \tag{6}$$

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a logistic population $L(\mu, \sigma^2)$. Define a standardized mean Z of a sample of size n from $L(\mu, \sigma^2)$ as

$$\begin{aligned} Z &= \frac{1}{\sqrt{n}\sigma} \sum_{i=1}^{n} (X_i - \mu) \\ &= \frac{\sqrt{n}}{\sigma}(\overline{X} - \mu), \end{aligned} \tag{7}$$

where $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ is the sample mean.

Let $f_n(z)$ and $F_n(z)$ denote the pdf and cdf of the standardized mean of samples of size n from $L(\mu, \sigma^2)$. Then the Edgeworth series expansions of $f_n(z)$ and $F_n(z)$ are given symbolically as

$$f_n(z, \nu) = \phi(z) + \phi(z) \sum_{j=1}^{\nu} p_j(z) n^{-j/2} + O(n^{-(\nu+1)/2})$$

and

$$F_n(z, \nu) = \Phi(z) - \phi(z) \sum_{j=1}^{\nu} P_j(z) n^{-j/2} + O(n^{-(\nu+1)/2})$$

respectively, where $\phi(z)$ and $\Phi(z)$ are the standard normal pdf and cdf respectively and $p_j(z)$ and $P_j(z)$ are polynomials in z, which are obtained up to $\nu = 10$ in Draper and Tierney (1973).

Using $p_j(z)$ and $P_j(z)$ from Table II of Draper and Tierney (1973) and expressions for relative cumulants of X, the Edgeworth series expansions of $f_n(z)$ and

$F_n(z)$ correct to order n^{-3} can be obtained as

$$\begin{aligned} & f_n(z, \nu = 6) \\ & \quad = \phi(z)\{1 + [(\frac{1}{4!})(\frac{6}{5})H_4(z)]n^{-1} \\ & \qquad + [(\frac{1}{6!})(\frac{48}{7})H_6(z) + (\frac{35}{8!})(\frac{6}{5})^2 H_8(z)]n^{-2} \\ & \qquad + [(\frac{1}{8!})(\frac{432}{5})H_8(z) \\ & \qquad + (\frac{210}{10!})(\frac{48}{7})(\frac{6}{5})H_{10}(z) + (\frac{5775}{12!})(\frac{6}{5})^3 H_{12}(z)]n^{-3}\} + O(n^{-7/2}) \end{aligned} \tag{8}$$

and

$$\begin{aligned} & F_n(z, \nu = 6) \\ & \quad = \Phi(z) - \phi(z)\{[(\frac{1}{4!})(\frac{6}{5})H_3(z)]n^{-1} \\ & \qquad + [(\frac{1}{6!})(\frac{48}{7})H_5(z) + (\frac{35}{8!})(\frac{6}{5})^2 H_7(z)]n^{-2} \\ & \qquad + [(\frac{1}{8!})(\frac{432}{5})H_7(z) \\ & \qquad + (\frac{210}{10!})(\frac{48}{7})(\frac{6}{5})H_9(z) + (\frac{5775}{12!})(\frac{6}{5})^3 H_{11}(z)]n^{-3}\} + O(n^{-7/2}) \end{aligned} \tag{9}$$

where the $H_j(x)$ are the Hermite polynomials of degree j, defined by

$$(\frac{d}{dx})^j \exp(-x^2/2) = (-1)^j H_j(x) \exp(-x^2/2), \quad j = 0, 1, \ldots.$$

The Hermite polynomials follow the recurrence relation:

$$H_j(x) = xH_{j-1}(x) - (j-1)H_{j-2}(x), \quad j = 2, 3, \ldots,$$

and are given in Table III in Draper and Tierney (1973) for $j = 1(1)30$.

3. AN ELIMINATION TYPE TWO-STAGE PROCEDURE FOR SELECTING THE BEST POPULATION

3.1. Preliminaries. Let π_i, $i = 1, \ldots, k$, denote k logistic populations with unknown means μ_i and a common known variance σ^2, and let

$$\Omega = \{\vec{\mu} = (\mu_1, \ldots, \mu_k); -\infty < \mu_i < \infty, \ i = 1, \ldots, k\}$$

be the parameter space. Denote the ranked values of the μ_i by

$$\mu_{[1]} \leq \cdots \leq \mu_{[k]}$$

and let

$$\delta_{ij} = \mu_{[i]} - \mu_{[j]}.$$

We assume that the experimenter has no prior knowledge concerning the pairing of π_i with $\mu_{[j]}, i = 1, \ldots, k,\ j = 1, \ldots, k$. Let $\pi_{(j)}$ denote the population associated with $\mu_{[j]}$.

The goal of the experimenter is to select the 'best' population which is defined as the population with the largest mean. This event is referred to as a correct selection (CS). The experimenter restricts consideration to procedures ($\mathcal{P}$) which guarantee the probability requirement

$$P_{\vec{\mu}}[CS|\mathcal{P}] \geq P^*, \quad \forall\ \vec{\mu} \in \Omega(\delta), \tag{10}$$

where $\delta > 0$ and $1/k < P^* < 1$ are specified prior to the start of experimentation and

$$\Omega(\delta) = \{\vec{\mu} \in \Omega | (\mu_{[k]} - \mu_{[k-1]}) \geq \delta\}$$

which is defined as the preference zone for a correct selection.

Here we propose an elimination type two-stage procedure $\mathcal{P}_2 = \mathcal{P}_2(n_1, n_2, h)$ which depends on non-negative integers n_1, n_2 and a real constant $h > 0$ which are determined prior to the start of experimentation. The constants (n_1, n_2, h) depend on k, δ and P^* and they are chosen so that $\mathcal{P}_2$ guarantees the probability requirement (10) and possesses a certain minimax property.

Procedure $\mathcal{P}_2$;

Stage 1: Take n_1 independent observations

$$X_{ij}^{(1)},\ j = 1, \ldots, n_1,$$

from $\pi_i,\ i = 1, \ldots k$, and compute the k sample means

$$\overline{X}_i^{(1)} = \frac{1}{n_1} \sum_{j=1}^{n_1} X_{ij}^{(1)},\ i = 1, \ldots, k.$$

Let $\overline{X}_{[k]}^{(1)} = \max_{1 \le j \le k} \overline{X}_j^{(1)}$. Determine the subset $\mathbf{I}$ of $\{1, \ldots, k\}$ where

$$\mathbf{I} = \{i | \overline{X}_i^{(1)} \ge \overline{X}_{[k]}^{(1)} - h\sigma/\sqrt{n_1}\},$$

and let π_I denote the associated subset of $\{\pi_1, \ldots, \pi_k\}$.

1. If π_I consists of only one population, stop sampling and assert that the population associated with $\overline{X}_{[k]}^{(1)}$ is best.
2. If π_I consists of more than one population, proceed to the second stage.

Stage 2: Take n_2 additional independent observations $X_{ij}^{(2)}$, $j = 1, \ldots, n_2$, from each population in π_I, and compute the cumulative sample means

$$\begin{aligned} \overline{X}_i &= \frac{1}{n_1 + n_2}\left(\sum_{j=1}^{n_1} X_{ij}^{(1)} + \sum_{j=1}^{n_2} X_{ij}^{(2)}\right) \\ &= \frac{1}{n_1 + n_2}\left(n_1 \overline{X}_i^{(1)} + n_2 \overline{X}_i^{(2)}\right) \end{aligned}$$

for $i \in \mathbf{I}$, where

$$\overline{X}_i^{(2)} = \frac{1}{n_2} \sum_{j=1}^{n_2} X_{ij}^{(2)}.$$

Assert that the population associated with $\max_{i \in \mathbf{I}} \overline{X}_i$ is the best.

There are an infinite number of combinations of (n_1, n_2, h) for given k, δ and P^*, which will exactly guarantee the probability requirement given by (10), and different design criteria lead to different choices. We will consider one of these criteria.

Let S' denote the cardinality of the set $\mathbf{I}$ in stage one and let

$$S = \begin{cases} 0; & \text{if } S' = 1 \\ S'; & \text{if } S' > 1. \end{cases} \tag{11}$$

Then the total sample size required by $\mathcal{P}_2$, TSS say, is given by

$$TSS = kn_1 + Sn_2.$$

Let $E_{\vec{\mu}}[TSS|\mathcal{P}_2]$ denote the expected total sample size for $\mathcal{P}_2$ under $\vec{\mu}$.

We adopt the following unrestricted minimax criterion to make a choice of (n_1, n_2, h) as well as to have the total sample size TSS small. For given k and specified δ and P^*, choose (n_1, n_2, h) to

$$\begin{aligned} &\textit{minimize} \sup_{\vec{\mu} \in \Omega} E_{\vec{\mu}}[TSS|\mathcal{P}_2] \\ &\textit{subject to} \inf_{\vec{\mu} \in \Omega(\delta)} P_{\vec{\mu}}[CS|\mathcal{P}_2] \ge P^*, \end{aligned} \tag{12}$$

where (n_1, n_2) are non-negative integers and $h \geq 0$.

For any population whose sample mean has the MLR property, Bhandari and Chaudhuri (1987) proved that the least favorable configuration (LFC) of the two-stage population means problem is a slippage configuration. However, the problem of evaluating the exact probability of a correct selection in the LFC associated with $\mathcal{P}_2$ is complicated and still remains to be solved. Here we will consider lower bounds for $P_{\vec{\mu}}[CS|\mathcal{P}_2]$ and construct conservative two-stage procedures.

3.2. Lower Bounds for the Probability of A Correct Selection for $\mathcal{P}_2$. In this section we derive lower bounds for $P_{\vec{\mu}}[CS|\mathcal{P}_2]$. These lower bounds will prove to be particularly useful since they achieve their infimum over $\Omega(\delta)$ at $\vec{\mu}(\delta) = (\mu, \ldots, \mu, \mu + \delta)$. This result will permit us to construct a conservative two-stage procedure which guarantees the probability requirement (10).

The next theorem gives one of these lower bounds for $P_{\vec{\mu}}[CS|\mathcal{P}_2]$.

Theorem 13: For all $\vec{\mu} \in \Omega(\delta)$ we have

$$\inf_{\vec{\mu} \in \Omega(\delta)} P_{\vec{\mu}}[CS|\mathcal{P}_2] \geq$$
$$\int_{-\infty}^{\infty} \{F_{n_1}(x + \delta\sqrt{n_1}/\sigma + h)\}^{k-1} dF_{n_1}(x)$$
$$+ \int_{-\infty}^{\infty} \{F_{n_1+n_2}(x + \delta\sqrt{n_1+n_2}/\sigma)\}^{k-1} dF_{n_1+n_2}(x) - 1, \tag{14}$$

where $F_n(x)$ is the cdf the standardized mean of a sample of size n from $L(\mu, \sigma^2)$.

Proof: The proof is omitted; a sharper bound is obtained in Theorem 17. □

Remark 15: The distribution of the mean of samples from logistic population has the monotone likelihood ratio (MLR) property with respect to the location parameter (Goel (1975)) and hence the distributions of $\overline{X}_i^{(1)}$ and $\overline{X}_i^{(2)}$ are stochastically increasing (SI) families in μ_i, $i = 1, \ldots, k$.

Remark 16: The cumulative sample means

$$\overline{X}_i = \frac{n_1}{n_1 + n_2}\overline{X}_i^{(1)} + \frac{n_2}{n_1 + n_2}\overline{X}_i^{(2)}$$

are strictly increasing in each $\overline{X}_i^{(j)}$, $j = 1, 2$, $i = 1, \ldots, k$.

We now obtain another lower bound to $P_{\vec{\mu}}[CS|\mathcal{P}_2]$ which can be shown to be uniformly superior to the one given in Theorem 3.1. It is also straightforward to determine the LFC of the population means relative to this new lower bound.

Theorem 17: For any $\vec{\mu} \in \Omega$ we have

$$\inf_{\vec{\mu} \in \Omega(\delta)} P_{\vec{\mu}}[CS|\mathcal{P}_2] \geq \int_{-\infty}^{\infty} \{F_{n_1}(x + \delta\sqrt{n_1}/\sigma + h)\}^{k-1} dF_{n_1}(x) \cdot \int_{-\infty}^{\infty} \{F_{n_1+n_2}(x + \delta\sqrt{n_1+n_2}/\sigma)\}^{k-1} dF_{n_1+n_2}(x), \quad (18)$$

where $F_n(x)$ is the cdf of the standardized mean of a sample of size n from $L(\mu, \sigma^2)$.

Proof: Let $F(.|\mu_i)$ and $G(.|\mu_i)$ denote the cdf's of $\overline{X}_i^{(1)}$ and $\overline{X}_i$ respectively and let $H(.,.|\mu_i)$ denote the joint cdf of $\overline{X}_i^{(1)}$ and $\overline{X}_i$. Then $F(.|\mu_i)$, $G(.|\mu_i)$ and $H(.,.|\mu_i)$ are non-increasing in μ_i, $i = 1, \ldots, k$, from Remarks 15 and 16. Without loss of generality we assume that $\mu_1 \leq \cdots \leq \mu_k$. Then for all $\vec{\mu} \in \Omega(\delta)$,

$$\begin{aligned}
&P_{\vec{\mu}}[CS|\mathcal{P}_2] \\
&= P_{\vec{\mu}}[\overline{X}_{(k)}^{(1)} \geq \max_{1 \leq j \leq k} \overline{X}_{(j)}^{(1)} - h\sigma/\sqrt{n_1}, \overline{X}_{(k)} = \max_{j \in \mathbf{I}} \overline{X}_{(j)}] \\
&\geq P_{\vec{\mu}}[\overline{X}_{(k)}^{(1)} \geq \overline{X}_{(j)}^{(1)} - h\sigma/\sqrt{n_1}, \overline{X}_{(k)} \geq \overline{X}_{(j)}, \forall j = 1, \ldots, k-1] \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} H(x + h\sigma/\sqrt{n_1}, y|\mu_i) dH(x, y|\mu_k) \\
&\geq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} H(x + h\sigma/\sqrt{n_1}, y|\mu_k - \delta) dH(x, y|\mu_k) \\
&= E_{\mu_k}[H^{k-1}\{\overline{X}_{(k)}^{(1)} + h\sigma/\sqrt{n_1}, \overline{X}_{(k)}|\mu_k - \delta\}],
\end{aligned}$$

where the expectation is with respect to the joint distribution of $\overline{X}_{(k)}^{(1)}$ and $\overline{X}_{(k)}$. Hence

$$\inf_{\vec{\mu} \in \Omega(\delta)} P_{\vec{\mu}}[CS|\mathcal{P}_2] \geq \inf_{\vec{\mu} \in \Omega(\delta)} E_{\mu_k}[H^{k-1}\{\overline{X}_{(k)}^{(1)} + h\sigma/\sqrt{n_1}, \overline{X}_{(k)}|\mu_k - \delta\}]$$

and it is enough to show that for all $\vec{\mu} \in \Omega(\delta)$,

$$\begin{aligned}
&E_{\mu_k}[H^{k-1}\{\overline{X}_{(k)}^{(1)} + h\sigma/\sqrt{n_1}, \overline{X}_{(k)}|\mu_k - \delta\}] \\
&\geq E_{\mu_k}[F^{k-1}(\overline{X}_{(k)}^{(1)} + h\sigma/\sqrt{n_1}|\mu_k - \delta)] E_{\mu_k}[G^{k-1}(\overline{X}_{(k)}|\mu_k - \delta)].
\end{aligned}$$

By Remark 16, for all a, b and μ,

$$\begin{aligned}
&P_\mu\{\overline{X}^{(1)}_{(j)} \le a, \overline{X}_{(j)} \le b\} \\
&= P_\mu\{\overline{X}^{(1)}_{(j)} \le a, \overline{X}^{(1)}_{(j)} \le \frac{n_1+n_2}{n_1}(b - \frac{n_2}{n_1+n_2}\overline{X}^{(2)}_{(j)})\} \\
&= E_\mu[P_\mu\{\overline{X}^{(1)}_{(j)} \le a, \overline{X}^{(1)}_{(j)} \le \frac{n_1+n_2}{n_1}(b - \frac{n_2}{n_1+n_2}\overline{X}^{(2)}_{(j)})|\overline{X}^{(2)}_{(j)}\}] \\
&\ge E_\mu[P_\mu\{\overline{X}^{(1)}_{(j)} \le a|\overline{X}^{(2)}_{(j)}\} \\
&\quad \cdot P_\mu\{\overline{X}^{(1)}_{(j)} \le \frac{n_1+n_2}{n_1}(b - \frac{n_2}{n_1+n_2}\overline{X}^{(2)}_{(j)})|\overline{X}^{(2)}_{(j)}\}] \\
&= P_\mu\{\overline{X}^{(1)}_{(j)} \le a\}P_\mu\{\overline{X}_{(j)} \le b\}.
\end{aligned}$$

Hence

$$\begin{aligned}
&E_{\mu_k}[H^{k-1}\{\overline{X}^{(1)}_{(k)} + h\sigma/\sqrt{n_1}, \overline{X}_{(k)}|\mu_k - \delta\}] \\
&\ge E_{\mu_k}[F^{k-1}\{\overline{X}^{(1)}_{(k)} + h\sigma/\sqrt{n_1}|\mu_k - \delta\} \\
&\quad \cdot G^{k-1}\{\overline{X}_{(k)}|\mu_k - \delta\}] \\
&\ge E_{\mu_k}[F^{k-1}\{\overline{X}^{(1)}_{(k)} + h\sigma/\sqrt{n_1}|\mu_k - \delta\}] \\
&\quad \cdot E_{\mu_k}[G^{k-1}\{\overline{X}_{(k)}|\mu_k - \delta\}]
\end{aligned}$$

by Chebyshev's inequality (Hardy, Littlewood and Pólya (1934)). □

<u>Remark 19</u>: If we let

$$a = \int_{-\infty}^{\infty} \{F_{n_1}(x + \delta\sqrt{n_1}/\sigma + h)\}^{k-1} dF_{n_1}(x)$$

and

$$b = \int_{-\infty}^{\infty} \{F_{n_1+n_2}(x + \delta\sqrt{n_1+n_2}/\sigma)\}^{k-1} dF_{n_1+n_2}(x),$$

then (14) states that

$$\inf_{\vec{\mu}\in\Omega(\delta)} P_{\vec{\mu}}[CS|\mathcal{P}_2] \ge a + b - 1$$

and (18) states that

$$\inf_{\vec{\mu}\in\Omega(\delta)} P_{\vec{\mu}}[CS|\mathcal{P}_2] \ge ab.$$

Since $a + b - 1 < ab$ for all $a, b \in (0, 1)$, the lower bound (18) is sharper and we use this henceforth.

<u>3.3. Expected Total Sample Size for $\mathcal{P}_2$</u>. In order to solve the optimization problem (12) we first obtain an expression for $E_{\vec{\mu}}[TSS|\mathcal{P}_2]$ and then determine sets of μ_i-values at which its supremum occurs.

Theorem 19: For any $\vec{\mu} \in \Omega$ we have

$$E_{\vec{\mu}}[TSS|\mathcal{P}_2] = kn_1 + n_2 \sum_{i=1}^{k} \int_{-\infty}^{\infty} \{\prod_{\substack{j=1\\ j\neq i}}^{k} F_{n_1}(x + \delta_{ij}\sqrt{n_1}/\sigma + h) - \prod_{\substack{j=1\\ j\neq i}}^{k} F_{n_1}(x + \delta_{ij}\sqrt{n_1}/\sigma - h)\} dF_{n_1}(x), \tag{20}$$

where $F_n(x)$ is the cdf of the standardized mean of a sample size n from $L(\mu, \sigma^2)$.

Proof: For any $\vec{\mu} \in \Omega$ we have

$$E_{\vec{\mu}}[TSS|\mathcal{P}_2] = kn_1 + n_2 E_{\vec{\mu}}[S|\mathcal{P}_2],$$

where S is defined in (11). Now

$$\begin{aligned} E_{\vec{\mu}}[S|\mathcal{P}_2] &= E_{\vec{\mu}}[S'|\mathcal{P}_2] - P_{\vec{\mu}}[S' = 1|\mathcal{P}_2] \\ &= \sum_{i=1}^{k} P_{\vec{\mu}}[\overline{X}_{(i)}^{(1)} \geq \overline{X}_{(j)}^{(1)} - h\sigma/\sqrt{n_1},\ \forall j \neq i] \\ &\quad - \sum_{i=1}^{k} P_{\vec{\mu}}[\overline{X}_{(i)}^{(1)} \geq \overline{X}_{(j)}^{(1)} + h\sigma/\sqrt{n_1},\ \forall j \neq i] \end{aligned} \tag{21}$$

and hence Theorem 19 follows immediately. □

The following theorem summarizes the result concerning the supremum of $E_{\vec{\mu}}[TSS|\mathcal{P}_2]$ for $\vec{\mu} \in \Omega$.

Theorem 22: For any $\vec{\mu} \in \Omega$, fixed k and (n_1, n_2, h) we have

$$\begin{aligned} &\sup_{\vec{\mu}\in\Omega} E_{\vec{\mu}}[TSS|\mathcal{P}_2] \\ &\quad = kn_1 + n_2 \int_{-\infty}^{\infty} [\{F_{n_1}(x+h)\}^{k-1} - \{F_{n_1}(x-h)\}^{k-1}] dF_{n_1}(x) \end{aligned} \tag{23}$$

which occurs when $\mu_{[1]} = \cdots = \mu_{[k]}$, where $F_n(x)$ is the cdf of the standardized mean of a sample of size n from $L(\mu, \sigma^2)$.

Proof: Using Remarks 15 and 16, and applying the results of Gupta (1965), it can be shown that both the supremum of $E_{\vec{\mu}}[S'|\mathcal{P}_2]$ and the infimum of $P_{\vec{\mu}}[S' = 1|\mathcal{P}_2]$ are achieved when $\mu_{[1]} = \cdots = \mu_{[k]}$. Hence the result follows immediately from Theorem 19. □

3.4. Optimization Problem Yielding Conservative Solutions. In this section we consider the optimization problem (12) which one must solve in order to determine the constants (n_1, n_2, h) which are necessary to implement $\mathcal{P}_2$. As we noted earlier, the problem of evaluating the exact probability of a correct selection in the LFC associated with $\mathcal{P}_2$ is very complicated. Thus we replace $\inf_{\vec{\mu} \in \Omega(\delta)} P_{\vec{\mu}}[CS|\mathcal{P}_2]$ by the conservative lower bound in (14) and consider the following optimization problem.

Table 1: Constants to implement the two-stage procedure $\mathcal{P}_2$ for selecting the logistic population with the largest mean: $P^* = 0.90$.

k	δ/σ	$\hat{n}_1$	$\hat{n}_2$	$\hat{h}$	$ETSS$
2	0.10	$0.1668e+03$	$0.1728e+03$	$0.2446e+01$	$0.650194e+03$
	0.50	$0.7013e+01$	$0.6404e+01$	$0.2591e+01$	$0.259726e+02$
	1.00	$0.1932e+01$	$0.1311e+01$	$0.3369e+01$	$0.643201e+01$
	2.00	$0.4011e+00$	$0.3724e+00$	$0.5331e+01$	$0.154620e+01$
	4.00	$0.1044e+00$	$0.8907e-01$	$0.5026e+01$	$0.386564e+00$
3	0.10	$0.2745e+03$	$0.2513e+03$	$0.2017e+01$	$0.146152e+04$
	0.50	$0.1126e+02$	$0.9634e+01$	$0.2071e+01$	$0.585665e+02$
	1.00	$0.2971e+01$	$0.2135e+01$	$0.2332e+01$	$0.146860e+02$
	2.00	$0.6894e+00$	$0.5189e+00$	$0.5004e+01$	$0.362197e+01$
	4.00	$0.1693e+00$	$0.1310e+00$	$0.4955e+01$	$0.900049e+00$
4	0.10	$0.3298e+03$	$0.3318e+03$	$0.1713e+01$	$0.229940e+04$
	0.50	$0.1340e+02$	$0.1300e+02$	$0.1728e+01$	$0.922982e+02$
	1.00	$0.3489e+01$	$0.3048e+01$	$0.1796e+01$	$0.232917e+02$
	2.00	$0.8374e+00$	$0.7008e+00$	$0.2643e+01$	$0.592662e+01$
	4.00	$0.2090e+00$	$0.1704e+00$	$0.2831e+01$	$0.147462e+01$
5	0.10	$0.3664e+03$	$0.4034e+03$	$0.1556e+01$	$0.315013e+04$
	0.50	$0.1488e+02$	$0.1595e+02$	$0.1553e+01$	$0.126542e+03$
	1.00	$0.3863e+01$	$0.3858e+01$	$0.1559e+01$	$0.320185e+02$
	2.00	$0.9610e+00$	$0.9217e+00$	$0.1867e+01$	$0.829954e+01$
	4.00	$0.2403e+00$	$0.2184e+00$	$0.2071e+01$	$0.208230e+01$
10	0.10	$0.4549e+03$	$0.6465e+03$	$0.1367e+01$	$0.750100e+04$
	0.50	$0.1844e+02$	$0.2588e+02$	$0.1357e+01$	$0.301614e+03$
	1.00	$0.4784e+01$	$0.6497e+01$	$0.1328e+01$	$0.765662e+02$
	2.00	$0.1328e+01$	$0.1644e+01$	$0.1257e+01$	$0.201395e+02$
	4.00	$0.3335e+00$	$0.4262e+00$	$0.1362e+01$	$0.528481e+01$
15	0.10	$0.4934e+03$	$0.7911e+03$	$0.1366e+01$	$0.119540e+05$
	0.50	$0.1999e+02$	$0.3177e+02$	$0.1358e+01$	$0.480822e+03$
	1.00	$0.5180e+01$	$0.8022e+01$	$0.1335e+01$	$0.122187e+03$
	2.00	$0.1433e+01$	$0.2074e+01$	$0.1280e+01$	$0.322460e+02$
	4.00	$0.3751e+00$	$0.5593e+00$	$0.1328e+01$	$0.865274e+01$

Table 2: Constants to implement the two-stage procedure $\mathcal{P}_2$ for selecting the logistic population with the largest mean: $P^* = 0.95$.

k	δ/σ	$\hat{n}_1$	$\hat{n}_2$	$\hat{h}$	$ETSS$
	0.10	0.3008e + 03	0.2827e + 03	0.1781e + 01	0.104953e + 04
	0.50	0.1227e + 02	0.1098e + 02	0.1810e + 01	0.421247e + 02
2	1.00	0.3215e + 01	0.2504e + 01	0.1958e + 01	0.106222e + 02
	2.00	0.7631e + 00	0.5883e + 00	0.3556e + 01	0.268233e + 01
	4.00	0.1899e + 00	0.1457e + 00	0.3785e + 01	0.667647e + 00
	0.10	0.4362e + 03	0.3657e + 03	0.1574e + 01	0.211419e + 04
	0.50	0.1768e + 02	0.1436e + 02	0.1589e + 01	0.849214e + 02
3	1.00	0.4579e + 01	0.3388e + 01	0.1654e + 01	0.214801e + 02
	2.00	0.1223e + 01	0.6952e + 00	0.2269e + 01	0.553339e + 01
	4.00	0.2858e + 00	0.1853e + 00	0.3237e + 01	0.139794e + 01
	0.10	0.4991e + 03	0.4519e + 03	0.1452e + 01	0.318364e + 04
	0.50	0.2023e + 02	0.1787e + 02	0.1453e + 01	0.127954e + 03
4	1.00	0.5232e + 01	0.4325e + 01	0.1464e + 01	0.324315e + 02
	2.00	0.1420e + 01	0.9417e + 00	0.1675e + 01	0.846044e + 01
	4.00	0.3393e + 00	0.2423e + 00	0.2163e + 01	0.218343e + 01
	0.10	0.5381e + 03	0.5259e + 03	0.1392e + 01	0.426098e + 04
	0.50	0.2182e + 02	0.2086e + 02	0.1388e + 01	0.171314e + 03
5	1.00	0.5649e + 01	0.5112e + 01	0.1379e + 01	0.434710e + 02
	2.00	0.1546e + 01	0.1182e + 01	0.1430e + 01	0.114045e + 02
	4.00	0.3809e + 00	0.3002e + 00	0.1751e + 01	0.299628e + 01
	0.10	0.6279e + 03	0.7682e + 03	0.1349e + 01	0.973702e + 04
	0.50	0.2544e + 02	0.3070e + 02	0.1342e + 01	0.391770e + 03
10	1.00	0.6592e + 01	0.7667e + 01	0.1321e + 01	0.996400e + 02
	2.00	0.1827e + 01	0.1923e + 01	0.1269e + 01	0.263641e + 02
	4.00	0.4897e + 00	0.5216e + 00	0.1344e + 01	0.724983e + 01
	0.10	0.6674e + 03	0.9126e + 03	0.1377e + 01	0.153152e + 05
	0.50	0.2703e + 02	0.3659e + 02	0.1370e + 01	0.616396e + 03
15	1.00	0.7002e + 01	0.9178e + 01	0.1354e + 01	0.156917e + 03
	2.00	0.1942e + 01	0.2339e + 01	0.1310e + 01	0.416523e + 02
	4.00	0.5293e + 00	0.6784e + 00	0.1300e + 01	0.115109e + 02

For the given k, δ and P^*, choose the constants (n_1, n_2, h) to

$$\text{minimize } kn_1 + n_2 \int_{-\infty}^{\infty} [\{F_{n_1}(x+h)\}^{k-1} - \{F_{n_1}(x-h)\}^{k-1}]dF_{n_1}(x)$$

$$\text{subject to } \int_{-\infty}^{\infty} \{F_{n_1}(x + \delta\sqrt{n_1}/\sigma + h)\}^{k-1}$$

$$\cdot \int_{-\infty}^{\infty} \{F_{n_1+n_2}(x + \delta\sqrt{n_1+n_2}/\sigma)\}^{k-1} dF_{n_1+n_2}(x) \geq P^*, \qquad (24)$$

where n_1 and n_2 are non-negative integers and $h \geq 0$.

Let $(\hat{n}_1, \hat{n}_2, \hat{h})$ denote the solution to the optimization problem (24) treating

n_1 and n_2 as continuous positive real variables. Then we can use the approximate design constants:

$$n_1 = [\hat{n}_1 + 1]^*, \; n_2 = [\hat{n}_2 + 1]^*, h = \hat{h},$$

where $[z]^*$ denotes the greatest integer which is less than z, to implement $\mathcal{P}_2$.

Tables 1 and 2 provide the constants $(\hat{n}_1, \hat{n}_2, \hat{h})$ and the values of the expected total sample size (ETSS) for $k = 2, 3, 4, 5, 10, 15$, $P^* = 0.90, 0.95$ and $\delta/\sigma = 0.1, 0.5, 1.0, 2.0, 4.0$. All computations were carried out in double-precision arithmetic on a Vax-11/780. The SUMT (Sequential Unconstrained Minimization Techniques: Fiacco and McCormick (1968)) algorithm is used to solve the nonlinear optimization problem. A source program in Fortran for the SUMT algorithm is given by Kuester and Mize (1973).

3.5. Performance of the Two-stage Procedure Relative to A Single-stage Procedure. We consider the single-stage procedure $\mathcal{P}_1$ that selects the population which yielded the largest sample mean based on samples of common size n. For $\mathcal{P}_1$ and $\mathcal{P}_2$ satisfying the same basic probability requirement (10), the relative efficiency (RE) of $\mathcal{P}_2$ w.r.t. $\mathcal{P}_1$ is defined by the ratio $E_{\vec{\mu}}[TSS|\mathcal{P}_2]/kn_s$ where n_s is the minimum sample size needed in $\mathcal{P}_1$. Clearly, RE depends on $\vec{\mu}$, δ and P^*. Values of the RE less than unity favor $\mathcal{P}_2$ over $\mathcal{P}_1$.

Now RE is approximately given by

$$\widetilde{RE} = \frac{1}{k\hat{n}_s}[k\hat{n}_1 + \hat{n}_2 \sum_{i=1}^{k} \int_{-\infty}^{\infty} \{\prod_{\substack{j=1 \\ j \neq i}}^{k} F_{\hat{n}_1}(t + \delta_{ij}\sqrt{\hat{n}_1}/\sigma + \hat{h}) - \prod_{\substack{j=1 \\ j \neq i}}^{k} F_{\hat{n}_1}(t + \delta_{ij}\sqrt{\hat{n}_1}/\sigma - \hat{h})\} dF_{\hat{n}_1}(t)]$$

where $\hat{n}_s$ is the smallest integer n_s for which

$$\int_{-\infty}^{\infty} \{F_{n_s}(t + (\sqrt{n_s}/\sigma)\delta)\}^{k-1} dF_{n_s}(t) \geq P^*.$$

We consider the relative efficiency for two special cases, namely, the equally spaced and the slippage configurations. First, for the equally spaced configuration, we assume that the unknown means of $\pi_1, \pi_2, \ldots, \pi_k$ are μ, $\mu + \delta, \ldots,$ $\mu + (k-1)\delta$, respectively. Let RE_{eq} denote the relative efficiency with respect to the above configuration. Then, since $\delta_{ij} = \mu_{[i]} - \mu_{[j]} = (i-j)\delta$,

$$\widetilde{RE}_{eq} = \frac{1}{k\hat{n}_s}[k\hat{n}_1 + \hat{n}_2 \sum_{i=1}^{k} \int_{-\infty}^{\infty} \{\prod_{\substack{j=1 \\ j\neq i}}^{k} F_{\hat{n}_1}(t + \sqrt{\hat{n}_1}(i-j)\delta/\sigma + \hat{h})$$

$$- \prod_{\substack{j=1 \\ j\neq i}}^{k} F_{\hat{n}_1}(t + \sqrt{\hat{n}_1}(i-j)\delta/\sigma - \hat{h})\} dF_{\hat{n}_1}(t)].$$

Table 3: Relative efficiency of the two-stage procedure $\mathcal{P}_2$ w.r.t. $\mathcal{P}_1$.

Slippage Configuration						
P^*	k	δ/σ				
		0.1	0.5	1.0	2.0	4.0
0.900	2	0.922	0.935	0.975	0.998	0.998
	3	0.796	0.809	0.852	0.994	0.993
	4	0.698	0.706	0.730	0.854	0.877
	5	0.636	0.642	0.658	0.711	0.745
	10	0.527	0.532	0.546	0.590	0.583
	15	0.494	0.499	0.513	0.551	0.551
0.950	2	0.820	0.826	0.847	0.965	0.974
	3	0.709	0.715	0.734	0.818	0.908
	4	0.651	0.656	0.671	0.722	0.754
	5	0.616	0.621	0.636	0.678	0.685
	10	0.545	0.550	0.564	0.606	0.612
	15	0.517	0.522	0.535	0.574	0.575
Equally Spaced Configuration						
0.900	2	0.922	0.935	0.975	0.998	0.998
	3	0.782	0.792	0.824	0.981	0.981
	4	0.689	0.695	0.713	0.762	0.776
	5	0.642	0.648	0.663	0.677	0.687
	10	0.554	0.559	0.573	0.612	0.603
	15	0.518	0.523	0.535	0.570	0.569
0.950	2	0.820	0.826	0.847	0.965	0.974
	3	0.716	0.721	0.736	0.795	0.852
	4	0.664	0.669	0.683	0.721	0.718
	5	0.634	0.639	0.653	0.690	0.678
	10	0.564	0.569	0.582	0.620	0.625
	15	0.533	0.537	0.550	0.586	0.585

Next, for the slippage configuration, we assume that the unknown means of the k populations are $\mu_{[j]} = \mu,\ j = 1, \ldots, k-1$, and $\mu_{[k]} = \mu + \delta,\ \delta \geq 0$. Then the relative efficiency with respect to the above configuration, RE_{sp}, is approximately given by

$$\widetilde{RE}_{sp} = \frac{1}{k\hat{n}_s}[k\hat{n}_1 + \hat{n}_2\{(k-1)\int_{-\infty}^{\infty}(F_{\hat{n}_1}(t+\hat{h}) - F_{\hat{n}_1}(t-\hat{h}))^{k-2} \cdot (F_{\hat{n}_1}(t-\sqrt{\hat{n}_1}\delta/\sigma+\hat{h}) - F_{\hat{n}_1}(t-\sqrt{\hat{n}_1}\delta/\sigma-\hat{h}))dF_{\hat{n}_1}(t) + \int_{-\infty}^{\infty}(F_{\hat{n}_1}(t+\sqrt{\hat{n}_1}\delta/\sigma+\hat{h}) - F_{\hat{n}_1}(t+\sqrt{\hat{n}_1}\delta/\sigma-\hat{h}))^{k-1}dF_{\hat{n}_1}(t)\}].$$

Table 3 gives the values of $\widetilde{RE}_{eq}$ and $\widetilde{RE}_{sp}$ for given values of $P^* = 0.90, 0.95$, $k = 2, 3, 4, 5, 10, 15$ and $\delta/\sigma = 0.1, 0.5, 1.0, 2.0, 4.0$.

For any values of P^*, k and δ, $\widetilde{RE}_{eq} \leq 1$ and $\widetilde{RE}_{sp} \leq 1$ and hence the two-stage procedure is more efficient than the single-stage procedure in terms of the expected total sample sizes. Furthermore, the effectiveness of $\mathcal{P}_2$ appears to be increasing with k.

3.6. An Example and Application of the Selection Procedure. We would like to illustrate the use of the two-stage selection procedure. Using the IMSL package we generated a set of logistic random deviates with a common variance $\sigma^2 = 1$ and location parameters $\theta_i = 0,\ 1.0,\ 2.5,\ 4.5,\ 5.5$. For $P^* = 0.90$, $\delta/\sigma = 1$, Table 1 gives $n_1 = 4$, $n_2 = 4$ and $h = 1.559$. At the first stage, we take four observations from each population.

Π_1	Π_2	Π_3	Π_4	Π_5
-0.375142	0.300190	2.62890	4.25264	5.84068
1.34968	-0.996658	2.22858	4.17833	4.98076
-0.0568658	1.57423	3.18402	4.00930	3.36522
0.00534297	1.99014	3.59607	4.71256	5.98789

Then $\overline{x}_1^{(1)} = 0.230753$, $\overline{x}_2^{(1)} = 0.716977$, $\overline{x}_3^{(1)} = 2.90939$, $\overline{x}_4^{(1)} = 4.28821$, $\overline{x}_5^{(1)} = 5.04364$, and $\max \overline{x}_j^{(1)} - h = 3.48464$. Hence we select populations Π_4 and Π_5 after the first stage. Then take additional 4 observations from each of the populations selected in the subset. These are:

Π_3	Π_4
5.60940	5.43388
2.84378	4.78641
2.35752	5.66428
2.93089	5.09631

Then $\overline{x}_4^{(2)} = 3.43540$, and $\overline{x}_5^{(2)} = 5.24522$. The cumulative sample means are given by $\overline{x}_4 = 3.86180$, and $\overline{x}_5 = 5.14443$. Hence the procedure selects Π_5 as the best population.

ACKNOWLEDGMENTS

This research was partially supported by the Office of Naval Research Contract N00014-88-K-0170 and NSF Grants DMS-8606964, DMS-8923071 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

REFERENCES

Alam, K. (1970). A two-sample procedure for selecting the population with the largest mean from k normal populations. Annals of the Institute of Statistical Mathematics, 22, 127–136.

Antle, C. E., Klimko, L., and Harkness, W. (1970). Confidence intervals for the parameters of the logistic distribution. Biometrika, **57**, 397–402.

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.

Berkson, J. (1944). Application of the Logistic Function to Bio-assay. Journal of the American Statistical Association, **39**, 357–365.

Berkson, J. (1951). Why I prefer logits to probits. Biometrics, **7**, 327–339.

Berkson, J. (1957). Tables for the maximum likelihood estimate of the logistic function. Biometrics, **13**, 28–34.

Birnbaum, A. and Dudman, J. (1963). Logistic Order Statistics. Annals of Mathematical Statistics, **34**, 658–663.

Bhandari, S. K. and Chaudhuri, A. R. (1987). On two conjectures about two-stage selection problem (to appear in Sankhyā Series B).

Cohen, D. S. (1959). A Two-Sample Decision Procedure for Ranking Means of Normal Populations with a Common Known Variance. M.S. Thesis, Department of Operations Research, Cornell University, Ithaca, New York.

Draper, N. R. and Tierney, D. E. (1973). Exact formulas for additional terms in some important series expansions. Communications in Statistics, **1**, 495-524.

Fiacco, A. V. and McCormick, G. P. (1968). Nonlinear Sequential Unconstrained Minimization Techniques. John Wiley and Sons, Inc., New York.

Goel, P. K. (1975). On the distribution of standardized mean of samples from the logistic population. Sankhyā, 2B, 165–172.

George, E. O. and Mudholkar, G. S. (1983). On the convolution of logistic random variables. Metrika, **30**, 1–13.

Gumbel, E. J. (1944). Ranges and midranges. Annals of Mathematical Statistics, **15**, 414–422.

Gumbel, E. J. and Keeney, R. D. (1950). The extremal quotient. Annals of Mathematical Statistics, **21**, 523–538.

Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Ph.D. Thesis (Mimeograph Series No. 150). Institute of Statistics, University of North Carolina, Chapel Hill, North Carolina.

Gupta, S. S. (1962). Life test sampling plans for normal and lognormal distributions. Technometrics, **4**, 151–175.

Gupta, S. S. (1965). On some multiple decision (selection and ranking) rules. Technometrics, **7**, 225–245.

Gupta, S. S. and Gnanadesikan, M. (1966). Estimation of the parameters of the logistic distribution. Biometrika, **53**, 565–570.

Gupta, S. S. and Shah, B. K. (1965). Exact moments and percentage points of the order statistics and the distribution of the range from the logistic distribution. Annals of Mathematical Statistics, **36**, 907–920.

Han, S-H. (1987). Contributions to selection and ranking theory with special reference to logistic populations. Ph.D. thesis (Technical Report # 87-38). Department of Statistics, Purdue University.

Hardy, G. H., Littlewood, J. E., and Pólya, G. (1934). Inequalities. Cambridge Univ. Press, Cambridge.

Hirschman, I. I. and Widder, D. V. (1955). The Convolution Transform. Princeton, N.J.

Kuester, J. L. and Mize, J. H. (1973). Optimization Techniques with Fortran.

McGraw-Hill, Inc., New York.

Pearl, R. and Reed, L. J. (1920). On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proceedings of National Academy of Science, **6**, 275–288.

Plackett, R. L. (1958). Linear estimation from censored data. Annals of Mathematical Statistics, **29**, 131–142.

Plackett, R. L. (1959). The analysis of life test data. Technometrics, **1**, 9–19.

Schafer, R. E. and Sheffield, T. S. (1973). Inferences on the parameters of the logistic distribution. Biometrics, **29**, 449–455.

Schoenberg, I. J. (1953). On Pólya frequency functions and their Laplace transformations. Journal d'Analyse Mathematique, **1**, 331–374.

Talacko, J. (1956). Perk's distributions and their role in the theory of Wiener's stochastic variates. Trabajos de Estadistica, **7**, 159–174.

Tamhane, A. C. and Bechhofer, R. E. (1977). A two-stage minimax procedure with screening for selecting the largest normal mean. Communications in Statistics–Theory and Methods, **A6**, 1003–1033.

Tamhane, A. C. and Bechhofer, R. E. (1979). A two-stage minimax procedure with screening for selecting the largest normal mean (II): an improved PCS lower bound and associated tables. Communications in Statistics–Theory and Methods, **A8**, 337–358.

Tarter, M. E. and Clark, V. A. (1965). Properties of the Median and other order statistics of logistic variates. Ann. Math. Statist. 36, 1779-1786.