ARO 32900.18-MA

SF298 last page

PURDUE UNIVERSITY

CENTER FOR STATISTICAL DECISION SCIENCES AND DEPARTMENT OF STATISTICS

EMPIRICAL BAYES SELECTION PROCEDURES FOR SELECTING THE BEST LOGISTIC POPULATION COMPARED WITH A CONTROL

by

Shanti S. Gupta, Zhengyan Lin and Xun Lin
Purdue University

Technical Report #98-03C

Department of Statistics
Purdue University
West Lafayette, IN USA

February 1998

EMPIRICAL BAYES SELECTION PROCEDURES FOR SELECTING THE BEST LOGISTIC POPULATION COMPARED WITH A CONTROL*

Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, IN 47907

and

Zhengyan Lin**
Department of Statistics
Purdue University
West Lafayette, IN 47907

Xun Lin
Purdue University
Department of Statistics
West Lafayette, IN 47907

Abstract

In this paper we investigate the problem of selecting the best logistic population from $k(\geq 2)$ possible candidates. The selected population must also be better than a given control. We employ the empirical Bayes approach and develop a selection procedure. The performance (rate of convergence) of the proposed selection rule is also analyzed. We also carry out a simulation study to investigate the rate of convergence of the proposed empirical Bayes selection procedure. The results of simulation are provided in the paper.

AMS Classification: primary 62F07; secondary 62C12.

Keywords: Asymptotically optimal; empirical Bayes; selection procedure; logistic population; rate of convergence.

---

*This research was supported in part by US Army Research Office, Grant DAAH04-95-1-0165 at Purdue University.

** Visiting professor from Department of Mathematics, Hangzhou University, China, 310028.

1 Introduction

Consider k independent logistic populations $\Pi_1, \ldots, \Pi_k$ with unknown means $\theta_1, \ldots, \theta_k$. Let $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ denote the ordered values of the parameters $\theta_1, \ldots, \theta_k$. It is assumed that the exact pairing between the ordered and the unordered parameters is unknown. A population Π_i with $\theta_i = \theta_{[k]}$ is called the best among the k underlying populations. In many pratical situations, we may not only be interested in the selection of the best population, but also require the selected population to be good enough compared with a given control. The problem of selecting the best population has been studied by many researchers. Gupta and Panchapakesan (1996) provided a comprehensive review of the development in this area. It should be pointed out that the logistic distribution serves as a statistical model in many practical situations, see, for example, Balakrishnan (1992). The statistical selection problem for logistic populations has been studied in Gupta and Han (1991, 1992), among others.

In this paper, we employ the empirical Bayes approach to select the best logistic population provided it is also as good as a given control. We describe the formulation of the selection problem and derive a Bayes selection procedure in Section 2. In Section 3, we construct an empirical Bayes selection procedure. Then we investigate the asymptotic optimality of the proposed empirical Bayes selection procedure in Section 4. A simulation study is carried out to investigate the performance of the proposed selection procedure in Section 5.

2 Formulation of the Selection Problem

Let $\Pi_1, \ldots, \Pi_k$ be k independent logistic populations with unknown means $\theta_1, \ldots, \theta_k$. Let $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ denote the ordered values of the parameters $\theta_1, \ldots, \theta_k$. It is assumed that the exact pairing between the ordered and the unordered parameters is unknown. A population π_i with $\theta_i = \theta_{[k]}$ is considered as the best population. For a given control θ_0, population π_i is defined to be good if the corresponding $\theta_i > \theta_0$, and bad otherwise. Our goal is to select a population which is the best among the k populations and also good compared with the standard θ_0. If there is no such treatment, we select none.

Let $\Omega = \{\underline{\theta} = (\theta_1, \ldots, \theta_k)\}$ be the parameter space. Let $\underline{a} = (a_0, \ldots, a_k)$ be an action, where $a_i = 0, 1; i = 0, 1, \ldots, k$ and $\sum_{i=0}^{k} a_i = 1$. For each $i = 1, \ldots, k, a_i = 1$ means that population π_i is selected as the best and also considered to be good compared with θ_0. $a_0 = 1$ means that all the k populations are excluded as bad and none is selected. We consider the loss function

$$L(\underset{\sim}{\theta}, \underset{\sim}{a}) = \max(\theta_{[k]}, \theta_0) - \sum_{i=0}^{k} a_i\theta_i.$$

It is the absolute error loss.

For each $i = 1, \ldots, k$, let $X_{i1}, \ldots, X_{iM}$ be a sample of size M from the logistic population $\Pi_i = L(\theta_i, \sigma_i^2)$ which has unknown mean θ_i and unknown variance $(\pi^2\sigma_i^2)/3$, that is, the conditional density distribution of X_{ij} given θ_i and σ_i^2 is

$$\frac{1}{\sigma_i}\frac{e^{-(x_i-\theta_i)/\sigma_i}}{(1+e^{-(x_i-\theta_i)/\sigma_i})^2}, \qquad -\infty < x_i < \infty. \tag{1}$$

Since logistic distribution is symmetric about its mean, the mean and the median of a logisitc population distribution are identical. For convenience, suppose M is an odd number, and we denote $M = 2s + 1$. We also assume that the unknown population median (and also the mean) θ_i has a normal $N(\mu_i, \tau_i^2)$ prior distribution with unknown parameters (μ_i, τ_i^2). The random variables $\theta_1, \ldots, \theta_k$ are assumed to be mutually independent. Define X_i to be the median of $\{X_{i1}, \ldots, X_{iM}\}$, $i = 1, \ldots, k$. Let $f_i(x_i|\theta_i, \sigma_i^2)$ and $h_i(\theta_i|\mu_i, \tau_i^2)$ be the conditional distributions of X_i given (θ_i, σ_i^2) and θ_i given (μ_i, τ_i^2), respectively. We have, for $i = 1, \ldots, k$,

$$f_i(x_i|\theta_i, \sigma_i^2) = \frac{(2s+1)!}{(s!)^2}\frac{1}{\sigma_i}\frac{(e^{-(x_i-\theta_i)/\sigma_i})^{s+1}}{(1+e^{-(x_i-\theta_i)/\sigma_i})^{2s+2}}, \qquad -\infty < x_i < \infty. \tag{2}$$

From (2) we see that the density function $f_i(x_i|\theta_i, \sigma_i^2)$ is symmetric about θ_i given θ_i, therefore,

$$EX_i = E(E(X_i|\theta_i)) = E\theta_i = \mu_i. \tag{3}$$

The posterior density of θ_i given $X_i = x_i$ is proportional to

$$\frac{(e^{-(x_i-\theta_i)/\sigma_i})^{s+1}}{(1+e^{-(x_i-\theta_i)/\sigma_i})^{2s+2}} \cdot e^{-\frac{(\theta_i-\mu_i)^2}{2\tau_i^2}}, \qquad -\infty < \theta_i < \infty. \tag{4}$$

Let $\underset{\sim}{X}=(X_1, \ldots, X_k)$ and $\mathcal{X}$ be the sample space generated by $\underset{\sim}{X}$. A selection procedure $\underset{\sim}{d}=(d_0, \ldots, d_k)$ is a mapping defined on the sample space $\mathcal{X}$. For every $\underset{\sim}{x} \in \mathcal{X}$, $d_i(\underset{\sim}{x}), i = 1, \ldots, k$, is the probability of selecting population Π_i as the best among the k

populations and also good compared with the given control θ_0, $d_0(\underline{x})$ is the probability of excluding all k populations as bad and selecting none. Also, $\sum_{i=0}^{k} d_i(\underline{x}) = 1$, for all $\underline{x} \in \mathcal{X}$.

Under the absolute error loss, the posterior median is the Bayes estimator of θ_i. We denote $\varphi_i(x_i)$ to be the posterior median of θ_i given $X_i = x_i$, $i = 1, \ldots, k$.

Under the preceding statistical model, the Bayes risk of the selection procedure $\underline{d}$ is denoted by $R(\underline{d})$. We have

$$R(\underline{d}) = -\int_{\mathcal{X}} [\sum_{i=0}^{k} d_i(\underline{x})\varphi_i(x_i)] f(\underline{x}) d(\underline{x}) + C, \tag{5}$$

where

$C = \int_\Omega \max(\theta_{[k]}, \theta_0) dH(\underline{\theta})$,

$H(\underline{\theta})$: the joint distribution of $\underline{\theta} = (\theta_1, \ldots, \theta_k)$,

$f_i(x_i) = \int_R f_i(x_i|\theta_i, \sigma_i^2) h_i(\theta_i|\mu_i, \tau_i^2) d\theta_i$,

$f(\underline{x}) = \Pi_{i=1}^{k} f_i(x_i)$,

$\varphi_0(x_0) = \theta_0$.

For each $\underline{x} \in \mathcal{X}$, let $I(\underline{x}) = \{i | \varphi_i(x_i) = \max_{0 \le j \le k} \varphi_j(x_j), i = 0, 1, \ldots, k\}$, and $i^* = \min\{i | i \in I(\underline{x})\}$. Then a Bayes selection procedure $d^B(\underline{x}) = (d_0^B(\underline{x}), \ldots, d_k^B(\underline{x}))$ is given as follows:

$$\begin{cases} d_{i^*}^B(\underline{x}) = 1, & \\ d_j^B(\underline{x}) = 0, & \text{for} \quad j \neq i^*. \end{cases} \tag{6}$$

3 The empirical Bayes Framework

The Bayes selection procedure $d^B(\underline{x})$ defined in Section 2 depends on the unknown parameters (μ_i, τ_i^2), $i = 1, \ldots, k$ and the specific form of $\varphi_i(x_i)$. Since the parameters and the specific form of $\varphi_i(x_i)$ are both unknown, it is impossible to implement the Bayes selection procedure for the selection problem in practice. In the empirical Bayes framework, it is generally assumed that there are some past observations when the present selection is to be made. At time $l = 1, \ldots, n$, let X_{ijl} be the j-th observation from Π_i, that is, for each $i = 1, \ldots, k$, let

$$\theta_{il} \sim N(\mu_i, \tau_i^2), \qquad l = 1, \ldots, n, \tag{7}$$

and

$$X_{ijl} \sim L(\theta_{il}, \sigma_i^2), \qquad j = 1, \ldots, M. \tag{8}$$

For $l = 1, \ldots, n$, denote $X_{i,l}$ to be the median of $(X_{i1l}, \ldots, X_{iMl})$, and

$$X_i(n) = \frac{1}{n} \sum_{l=1}^{n} X_{i,l}, \tag{9}$$

$$S_i^2(n) = \frac{1}{n-1} \sum_{l=1}^{n} (X_{i,l} - X_i(n))^2. \tag{10}$$

Then,

$$E(X_{i,l}) = E(E(X_{i,l}|\theta_{il})) = E(\theta_{il}) = \mu_i, \tag{11}$$

and

$$\begin{aligned} Var(X_{i,l}) &= Var(E(X_{i,l}|\theta_{il})) + E(Var(X_{i,l}|\theta_{il})) \\ &= Var(\theta_{il}) + E(Var(X_{i,l}|\theta_{il})) \\ &= \tau_i^2 + E(Var(X_{i,l}|\theta_{il})) \\ &< \infty. \end{aligned} \tag{12}$$

Denote $\nu_i^2 = Var(X_{i,l})$. Since $(X_{i1}, \ldots, X_{in})$ are i.i.d., by the strong law of large numbers, we know that as $n \to \infty$,

$$\begin{cases} X_i(n) \longrightarrow \mu_i, & a.s. \\ S_i^2(n) \longrightarrow \nu_i^2, & a.s. \end{cases} \tag{13}$$

To derive the empirical Bayes selection procedure, we first consider the following lemmas. The following lemma is from Serfling (1980).

Lemma 3.1 Let $\{Y_i, 1 \leq i \leq m\}$ be m i.i.d. random observations from continuous distribution function F; also let $\hat{\xi}$ and ξ be the medians of $\{Y_i, 1 \leq i \leq m\}$ and F, respectively. Then, for any $\epsilon > 0$,

$$P\{|\hat{\xi} - \xi| > \epsilon\} \leq 2e^{-2m\delta_\epsilon^2}, \tag{14}$$

where $\delta_\epsilon = \min\{F(\xi + \epsilon) - \frac{1}{2}, \frac{1}{2} - F(\xi - \epsilon)\}$.

Put $\sigma' = \min_{1 \leq i \leq k} \sigma_i$, $\sigma^* = \max_{1 \leq i \leq k} \sigma_i$. $X_{i1}, \ldots, X_{iM}$ are i.i.d. from $L(\theta_i, \sigma_i^2)$, which has the following cumulative distribution function

$$F(t_i) = \frac{1}{1 + e^{-(t_i - \theta_i)/\sigma_i}} \qquad -\infty < t_i < \infty, \tag{15}$$

and for $0 < \epsilon \leq \sigma'$,

$$F(\theta_i + \epsilon) - \frac{1}{2} = \frac{1}{2} - F(\theta_i - \epsilon) = \frac{e^{\epsilon/\sigma_i} - 1}{2(e^{\epsilon/\sigma_i} + 1)} \geq \frac{\epsilon}{2(e+1)\sigma^*}. \tag{16}$$

Given θ_i, θ_i and X_i are the population median and sample median respectively, we have, from Lemma 3.1,

$$P\{|X_i - \theta_i| > \epsilon\} \leq 2e^{\frac{-(2s+1)\epsilon^2}{2(e+1)^2\sigma^{*2}}}. \tag{17}$$

For any $0 < \epsilon \leq \sigma'$, denote $S_i = \{\underset{\sim}{x} \in \mathcal{X} : \ |x_i - \theta_i| \leq \epsilon\}$. We show that the conditional density of X_i given θ_i and σ_i^2 is approximately $N(\theta_i, \frac{2}{s+1}\sigma_i^2)$ as $s \to \infty$.

From (2), the conditional density of X_i given θ_i and σ_i^2 is

$$\begin{aligned} f_i(x_i|\theta_i, \sigma_i^2) &= \frac{(2s+1)!}{(s!)^2} \frac{1}{\sigma_i} \frac{(e^{-(x_i-\theta_i)/\sigma_i})^{s+1}}{(1 + e^{-(x_i-\theta_i)/\sigma_i})^{2s+2}} \\ &= \frac{(2s+1)!}{(s!)^2} \frac{1}{\sigma_i} \frac{1}{(2 + e^{-(x_i-\theta_i)/\sigma_i} + e^{(x_i-\theta_i)/\sigma_i})^{s+1}}. \end{aligned} \tag{18}$$

By Stirling's formula, when s is large enough,

$$\frac{(2s+1)!}{(s!)^2} \approx \frac{2^{(2s+\frac{3}{2})}}{\sqrt{2\pi}}\sqrt{s+1}. \tag{19}$$

Also choosing $\epsilon = \epsilon_s \downarrow 0$ to be a sequence of fixed numbers which tend to 0 as $s \to \infty$, by Taylor's polynomial expansion, we have

$$\log(2 + e^{-(x_i-\theta_i)/\sigma_i} + e^{(x_i-\theta_i)/\sigma_i}) \approx \log 4 + \frac{1}{4}\frac{(x_i-\theta_i)^2}{\sigma_i^2} \tag{20}$$

on S_i. When $s \to \infty$, from (17),

$$P\{\underset{\sim}{X} \notin S_i\} \leq 2e^{\frac{-(2s+1)\epsilon^2}{2(\epsilon+1)^2\sigma^{*2}}} \longrightarrow 0. \tag{21}$$

Therefore, we see that as $s \to \infty$,

$$f_i(x_i|\theta_i, \sigma_i^2) \approx \frac{1}{\sqrt{2\pi}\sqrt{2/s+1}}\frac{1}{\sigma_i}e^{-\frac{s+1}{4}\frac{(x_i-\theta_i)^2}{\sigma_i^2}}, \tag{22}$$

that is, $f_i(x_i|\theta_i, \sigma_i^2)$ is approximately $N(\theta_i, \frac{2}{s+1}\sigma_i^2)$.

From above, we can see that for sufficiently large s, the conditional density of $X_{i,l}$ is approximately $N(\theta_i, \frac{2}{s+1}\sigma_i^2)$, given θ_i and σ_i. Since the prior distribution of θ_i is $N(\mu_i, \tau_i^2)$, the unconditional density of $X_{i,l}$ is approximately $N(\mu_i, \tau_i^2 + \frac{2}{s+1}\sigma_i^2)$.

For each population Π_i, let $W_i^2(n)$ be the measure of the overall sample variation for the past observations. That is,

$$\begin{cases} \bar{X}_{il} = \frac{1}{M}\sum_{j=1}^{M} X_{ijl}, \\ W_i^2(n) = \frac{1}{(M-1)n}\sum_{j=1}^{M}\sum_{l=1}^{n}(X_{ijl} - \bar{X}_{il})^2. \end{cases} \tag{23}$$

Then we define, for $i = 1, \ldots, k$,

$$\begin{cases} \hat{\mu}_i = X_i(n), \\ \hat{\sigma}_i^2 = \frac{3}{\pi^2}W_i(n)^2, \\ \hat{\nu}_i^2 = S_i^2(n), \\ \hat{\tau}_i^2 = \max(\hat{\nu}_i^2 - \frac{2}{s+1}\hat{\sigma}_i^2, 0). \end{cases} \tag{24}$$

and

$$\begin{cases} \hat{\varphi}_i(x_i) = \begin{cases} (x_i\hat{\tau}_i^2 + \frac{2\hat{\sigma}_i^2}{s+1}\hat{\mu}_i)/\hat{\nu}_i^2, & \text{if } \hat{\nu}_i^2 - \frac{2}{s+1}\hat{\sigma}_i^2 > 0, \\ \hat{\mu}_i, & \text{if } \hat{\nu}_i^2 - \frac{2}{s+1}\hat{\sigma}_i^2 \leq 0, \end{cases} \\ \hat{\varphi}_0(x_0) = \theta_0. \end{cases} \tag{25}$$

Then for each $\underline{x} \in \mathcal{X}$, let $\hat{I}(\underline{x}) = \{i|\hat{\varphi}_i(x_i) = \max_{0\leq j\leq k} \hat{\varphi}_j(x_j), i = 0, 1, \ldots, k\}$, and $\hat{i}^* = \min\{i|i \in \hat{I}(\underline{x})\}$. We propose the following empirical Bayes selection procedure $d^{(n,s)}(\underline{x}) = (d_0^{(n,s)}(\underline{x}), \ldots, d_k^{(n,s)}(\underline{x}))$ as follows:

$$\begin{cases} d_{\hat{i}^*}^{(n,s)} = 1, \\ d_j^{(n,s)} = 0, \quad \text{for} \quad j \neq \hat{i}^*. \end{cases} \tag{26}$$

4 Performance of the proposed selection procedure

Consider the empirical Bayes selection procedure $d^{(n,s)}(\underline{x})$ constructed in Section 3. Let $R(d^{(n,s)}(\underline{x}))$ be the conditional Bayes risk given the past observations $\{X_{ijl}, i = 1, \ldots, k; j = 1, \ldots, M;$ and $l = 1, \ldots, n\}$ and $ER(d^{(n,s)}(\underline{x}))$ the Bayes risk of the empirical Bayes selection procedure respectively, where E is the expectation taken with respect to the past observations $\{X_{ijl}\}$. From (5),

$$R(\underline{d}^{(n,s)}(\underline{x})) = -\int_{\mathcal{X}} [\sum_{i=0}^{k} d_i^{(n,s)}(\underline{x})\hat{\varphi}_i(x_i)] f(\underline{x}) d(\underline{x}) + C. \tag{27}$$

Note that $R(d^{(n,s)}(\underline{x})) - R(d^B(\underline{x})) \geq 0$, since $d^B(\underline{x})$ is the Bayes selection procedure. Therefore, $E(R(d^{(n,s)}(\underline{x})) - R(d^B(\underline{x}))) \geq 0$. We use the nonnegative difference regret risk $E(R(d^{(n,s)}(\underline{x})) - R(d^B(\underline{x}))) \geq 0$ as a measure of the performance of the selection procedure $d^{(n,s)}(\underline{x})$.

We first state some facts about $\varphi_i(x_i)$, the posterior median of θ_i given $X_i = x_i$ and μ_i. From the definition of $\varphi_i(x_i)$, we can see that $\varphi_i(x_i)$ is between x_i and μ_i. Besides,

Lemma 4.1 When s is large enough, for $1 \leq i \leq k$,

$$|\varphi_i(x_i) - x_i| \leq 2\sigma_i\sqrt{\frac{\log s}{s}} \tag{28}$$

Proof. We only prove $\varphi_i(x_i) \leq x_i + 2\sigma_i\sqrt{\frac{\log s}{s}}$ here. The proof of $\varphi_i(x_i) \geq x_i - 2\sigma_i\sqrt{\frac{\log s}{s}}$ is similar. To prove $\varphi_i(x_i) \leq x_i + 2\sigma_i\sqrt{\frac{\log s}{s}}$, it suffices to show that

$$\begin{aligned}
&\int_{x_i+2\sigma_i\sqrt{\frac{\log s}{s}}}^{\infty} f_i(x_i|\theta_i,\sigma_i^2)h_i(\theta_i|\mu_i,\tau_i^2)d\theta_i \\
= &\int_{x_i+2\sigma_i\sqrt{\frac{\log s}{s}}}^{\infty} \frac{(2s+1)!}{(s!)^2}\frac{1}{\sigma_i}\frac{1}{\sqrt{2\pi}\tau_i}\frac{(e^{(\theta_i-x_i)/\sigma_i})^{s+1}}{(1+e^{(\theta_i-x_i)/\sigma_i})^{2s+2}}\cdot e^{-\frac{(\theta_i-\mu_i)^2}{2\tau_i^2}}d\theta_i \\
= &\int_{2\sqrt{\frac{\log s}{s}}}^{\infty} \frac{1}{\sqrt{2\pi}\tau_i}\frac{(2s+1)!}{(s!)^2}\left(\frac{e^\theta}{(1+e^\theta)^2}\right)^{s+1}\cdot e^{-\frac{(\sigma_i\theta+x_i-\mu_i)^2}{2\tau_i^2}}d\theta \longrightarrow 0,
\end{aligned} \tag{29}$$

as $s \to \infty$. We first show

$$t(\theta,s) := \frac{(2s+1)!}{(s!)^2}\left(\frac{e^\theta}{(1+e^\theta)^2}\right)^{s+1} \longrightarrow 0 \qquad as\ s \to \infty, \tag{30}$$

uniformly for $\theta \geq 2\sqrt{\frac{\log s}{s}}$. Obviously it is enough to consider the case of $\theta = 2\sqrt{\frac{\log s}{s}}$ since $t(\theta,s)$ is decreasing on $\theta > 0$. When $\theta = 2\sqrt{\frac{\log s}{s}}$ and s is large enough, by Taylor's formula,

$$\log(1+e^\theta) = \log 2 + \frac{1}{2}\theta + \frac{1}{8}\theta^2 + o(\theta^2), \tag{31}$$

and by (19), when s is large enough,

$$\log\frac{(2s+1)!}{(s!)^2} \leq 2(s+1)\log 2 + \frac{1}{2}\log(s+1). \tag{32}$$

From (31) and (32), we obtain that

$$\begin{aligned}
&\log t(\theta,s) \\
= &(s+1)[\theta - 2\log(1+e^\theta)] + \log\frac{(2s+1)!}{(s!)^2} \\
\leq &-2(s+1)\log 2 - \frac{s+1}{4}\theta^2 + 2(s+1)\log 2 + \frac{1}{2}\log s + o(s\theta^2) \\
= &-(\frac{s+1}{s} - \frac{1}{2})\log s + o(\log s) \longrightarrow -\infty,
\end{aligned} \tag{33}$$

as $s \to \infty$. Therefore, (30) is proved, from which we can immediately see that (29) holds true. It completes the proof of Lemma 4.1.

The next lemma is well known and can be found in Baum and Katz (1965).

Lemma 4.2 Let $X_1, \ldots, X_n$ be i.i.d. random variables with mean 0. Suppose for $\alpha > 1$, $E|X_i|^\alpha < \infty$, for $i = 1, \ldots, n$, then for any $\epsilon > 0$,

$$P\{|\sum_{i=1}^{n} X_i/n| \geq \epsilon\} = o(n^{-(\alpha-1)}). \tag{34}$$

As a consequence of Lemma 4.2, we have

Lemma 4.3 Let $X_1, \ldots, X_n$ be independent random variables, with mean $EX_i = \mu$ and variance $\mathrm{Var}X_i = \sigma^2$, for $i = 1, \ldots, n$. Also let $\bar{X} = \frac{1}{n}\sum X_i$ and $S_n^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$. Suppose for $i = 1, \ldots, n$ and a fixed number $\alpha > 2$, $E|X_i|^\alpha < \infty$, then for any $\epsilon > 0$,

$$P\{|S_n^2 - \sigma^2| \geq \epsilon\} = o(n^{-(\alpha/2-1)}). \tag{35}$$

The proof of Lemma 4.3 can be found in Gupta and Lin (1997).

Since $EX_{i.l}^4 < \infty$, for any $\epsilon > 0$, by Lemma 4.2,

$$P\{|\hat{\mu}_i - \mu_i| \geq \epsilon\} = o(n^{-3}), \tag{36}$$

also by Lemma 4.3,

$$P\{|\hat{\nu}_i^2 - \nu_i^2| \geq \epsilon\} = o(n^{-1}). \tag{37}$$

Similarly, we have for any $\epsilon > 0$,

$$P\{|\hat{\sigma}_i^2 - \sigma_i^2| \geq \epsilon\} = o(n^{-1}). \tag{38}$$

When s is large enough, $\nu_i^2 - \frac{2}{s+1}\sigma_i^2 > 0$. Therefore, from (37) and (38), when s is sufficiently large,

$$P\{\hat{\nu}_i^2 - \frac{2}{s+1}\hat{\sigma}_i^2 \leq 0\} = o(n^{-1}). \tag{39}$$

Besides, $\tau_i^2 = \nu_i^2 - E(Var(X_{i,l}|\theta_i))$ by (12) and

$$\begin{aligned}
E(Var(X_{i,l}|\theta_i)) &= \int_{-\infty}^{\infty} (x_{il}-\theta_i)^2 \frac{(2s+1)!}{(s!)^2} \frac{1}{\sigma_i} \frac{(e^{-(x_{il}-\theta_i)/\sigma_i})^{s+1}}{(1+e^{-(x_{il}-\theta_i)/\sigma_i})^{2s+2}} dx_{il} \\
&= \sigma_i \int_{-\infty}^{\infty} x^2 \frac{(2s+1)!}{(s!)^2} \Big(\frac{e^x}{(1+e^x)^2}\Big)^{s+1} dx. \qquad (40)
\end{aligned}$$

We have

Lemma 4.4

$$\int_{-\infty}^{\infty} x^2 \frac{(2s+1)!}{(s!)^2} \Big(\frac{e^x}{(1+e^x)^2}\Big)^{s+1} dx = o\Big(\sqrt{\frac{\log s}{s}}\Big). \qquad (41)$$

Proof.

$$\begin{aligned}
&\int_{-\infty}^{\infty} x^2 \frac{(2s+1)!}{(s!)^2} \Big(\frac{e^x}{(1+e^x)^2}\Big)^{s+1} dx \\
= \;& 2\int_0^{\infty} x^2 \frac{(2s+1)!}{(s!)^2} \Big(\frac{e^x}{(1+e^x)^2}\Big)^{s+1} dx \\
= \;& 2\Big(\int_0^{\sqrt{8\frac{\log s}{s}}} + \int_{\sqrt{8\frac{\log s}{s}}}^{3} + \int_3^{\infty}\Big) x^2 \frac{(2s+1)!}{(s!)^2} \Big(\frac{e^x}{(1+e^x)^2}\Big)^{s+1} dx \\
:= \;& T_1 + T_2 + T_3. \qquad (42)
\end{aligned}$$

By Stirling's formula, when s is large enough,

$$\begin{aligned}
T_1 &= 2\frac{(2s+1)!}{(s!)^2} \int_0^{\sqrt{8\frac{\log s}{s}}} x^2 \Big(\frac{e^x}{(1+e^x)^2}\Big)^{s+1} dx \\
&\leq 2 \cdot 2^{2(s+1)} \sqrt{s+1} \cdot 2^{-2(s+1)} \int_0^{\sqrt{8\frac{\log s}{s}}} x^2 dx \\
&\leq \sqrt{s+1} \Big(8\frac{\log s}{s}\Big)^{3/2} \\
&= o\Big(\sqrt{\frac{\log s}{s}}\Big). \qquad (43)
\end{aligned}$$

Using the same approach as in the proof of Lemma 4.1, we have

$$T_2 = 2\frac{(2s+1)!}{(s!)^2} \int_{\sqrt{8\frac{\log s}{s}}}^{3} x^2 \Big(\frac{e^x}{(1+e^x)^2}\Big)^{s+1} dx$$

$$
\begin{aligned}
&\leq 2\frac{(2s+1)!}{(s!)^2}\Big(\frac{e^{\sqrt{8\frac{\log s}{s}}}}{(1+e^{\sqrt{8\frac{\log s}{s}}})^2}\Big)^{s+1}\int_{\sqrt{8\frac{\log s}{s}}}^{3} x^2 dx \\
&= o\Big(\sqrt{\frac{\log s}{s}}\Big). \qquad (44)
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
T_3 &= 2\frac{(2s+1)!}{(s!)^2}\int_3^\infty x^2\Big(\frac{e^x}{(1+e^x)^2}\Big)^{s+1}dx \\
&\leq 2\frac{(2s+1)!}{(s!)^2}\int_3^\infty x^2 e^{-(s+1)x}dx \\
&= o(\sqrt{\frac{\log s}{s}}). \qquad (45)
\end{aligned}
$$

This completes the proof of Lemma 4.4.

From Lemma 4.4, we observe that when s is sufficiently large,

$$
E(Var(X_{i,l}|\theta_i)) = o\Big(\sqrt{\frac{\log s}{s}}\Big), \qquad (46)
$$

and therefore, by (37), (39) and the definition of $\hat{\tau}_i^2$, for $\epsilon \geq c\sqrt{\frac{\log s}{s}}$, where $c > 0$,

$$
P\{|\hat{\tau}_i^2 - \tau_i^2| \geq \epsilon\} = o(n^{-1}), \qquad (47)
$$

and furthermore,

$$
P\{\hat{\nu}_i^2/\hat{\tau}_i^2 \leq \nu_i^2/(2\tau_i^2)\} = o(n^{-1}). \qquad (48)
$$

Next we investigate the rate of convergence of $E(R(d^{(n,s)}(\underline{x})) - R(d^B(\underline{x})))$. Let $P_{n,s}$ be the probability measure generated by the past observations $X_{ijl}, i = 1, \dots k; j = 1, \dots M$ and $l = 1, \dots, n$.

$$
\begin{aligned}
&E(R(d^{(n,s)}(\underline{x})) - R(d^B(\underline{x}))) \\
= &\sum_{i=0}^{k}\sum_{j=0}^{k}\int_{\mathcal{X}} P_{n,s}\{i^* = i, \hat{i}^* = j\}(\varphi_i(x_i) - \varphi_j(x_j))f(\underline{x})d\underline{x}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{i=1}^{k} \int_{\mathcal{X}} P_{n,s}\{i^* = i, \hat{i}^* = 0\}(\varphi_i(x_i) - \theta_0) f(\underline{x}) d\underline{x} \\
&\quad + \sum_{j=1}^{k} \int_{\mathcal{X}} P_{n,s}\{i^* = 0, \hat{i}^* = j\}(\theta_0 - \varphi_j(x_j)) f(\underline{x}) d\underline{x} \\
&\quad + \sum_{i=1}^{k}\sum_{j=1}^{k} \int_{\mathcal{X}} P_{n,s}\{i^* = i, \hat{i}^* = j\}(\varphi_i(x_i) - \varphi_j(x_j)) f(\underline{x}) d\underline{x} \\
&\le 2\sum_{i=1}^{k} \int_{R} P_{n,s}\{|\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > |\varphi_i(x_i) - \theta_0|\}|\varphi_i(x_i) - \theta_0| f_i(x_i) dx_i \\
&\quad + 2\sum_{i=1}^{k}\sum_{j=1}^{k} \int_{R^2} P_{n,s}\{|\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{|\varphi_i(x_i) - \varphi_j(x_j)|}{2}\}|\varphi_i(x_i) - \varphi_j(x_j)| \\
&\qquad\qquad \times f_i(x_i) f_j(x_j) dx_i dx_j \\
&:= I_1 + I_2. \qquad (49)
\end{aligned}
$$

For any $\epsilon > 0$, and $i, j = 1, \ldots, k$, let

$$
\begin{cases}
\mathcal{X}_i = \{x_i : |\varphi_i(x_i) - \theta_0| \le \epsilon\}, \\
\mathcal{X}_{ij} = \{(x_i, x_j) : |\varphi_i(x_i) - \varphi_j(x_j)| \le \epsilon\}.
\end{cases} \qquad (50)
$$

Then we have

$$
\begin{aligned}
I_1 &= 2\sum_{i=1}^{k} \int_{\mathcal{X}_i} P_{n,s}\{|\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > |\varphi_i(x_i) - \theta_0|\}|\varphi_i(x_i) - \theta_0| f_i(x_i) dx_i \\
&\quad + 2\sum_{i=1}^{k} \int_{R-\mathcal{X}_i} P_{n,s}\{|\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > |\varphi_i(x_i) - \theta_0|\}|\varphi_i(x_i) - \theta_0| f_i(x_i) dx_i \\
&\le 2\sum_{i=1}^{k} \int_{\mathcal{X}_i} \epsilon f_i(x_i) dx_i \\
&\quad + 2\sum_{i=1}^{k} \int_{R} P_{n,s}\{|\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \epsilon\}|\varphi_i(x_i) - \theta_0| f_i(x_i) dx_i. \qquad (51)
\end{aligned}
$$

By Lemma 4.1, when s is large enough, $|\varphi_i(x_i) - x_i| \le 2\sigma_i\sqrt{\frac{\log s}{s}}$. From now on, we always set $\epsilon = 16\sigma^*\sqrt{\frac{\log s}{s}}$. Therefore, for sufficiently large s,

$$
|x_i - \theta_0| \le |\varphi_i(x_i) - x_i| + |\varphi_i(x_i) - \theta_0| \le 2\epsilon \qquad (52)
$$

on $\mathcal{X}_i$ and

$$
\begin{aligned}
\int_{\mathcal{X}_i} f_i(x_i)dx_i &\leq \int_{\{|x_i-\theta_0|\leq 2\epsilon\}} f_i(x_i)dx_i \\
&\leq \int_{\{|x_i-\theta_0|\leq 2\epsilon\}} \frac{1}{\sqrt{2\pi}\tau_i}dx_i \\
&= \frac{4\epsilon}{\sqrt{2\pi}\tau_i}. \qquad (53)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
I_1 &\leq \frac{8k}{\sqrt{2\pi}\tau_i}\epsilon^2 \\
&\quad +2\sum_{i=1}^{k}\int_R P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)|>\epsilon\}[|\varphi_i(x_i)-\mu_i|+|\mu_i-\theta_0|]f_i(x_i)dx_i. \qquad (54)
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
I_2 &= 2\sum_{i=1}^{k}\sum_{j=1}^{k}\int_{\mathcal{X}_{ij}} P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)|>\frac{|\varphi_i(x_i)-\varphi_j(x_j)|}{2}\}|\varphi_i(x_i)-\varphi_j(x_j)| \\
&\qquad \times f_i(x_i)f_j(x_j)dx_idx_j \\
&\quad +2\sum_{i=1}^{k}\sum_{j=1}^{k}\int_{R^2-\mathcal{X}_{ij}} P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)|>\frac{|\varphi_i(x_i)-\varphi_j(x_j)|}{2}\}|\varphi_i(x_i)-\varphi_j(x_j)| \\
&\qquad \times f_i(x_i)f_j(x_j)dx_idx_j \\
&\leq 2\epsilon\sum_{i=1}^{k}\sum_{j=1}^{k}\int_{\mathcal{X}_{ij}} f_i(x_i)f_j(x_j)dx_idx_j \\
&\quad +2\sum_{i=1}^{k}\sum_{j=1}^{k}\int_{R^2} P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)|>\frac{\epsilon}{2}\}|\varphi_i(x_i)-\varphi_j(x_j)| \\
&\qquad \times f_i(x_i)f_j(x_j)dx_idx_j. \qquad (55)
\end{aligned}
$$

From (28), when s is large enough, $|\varphi_i(x_i)-x_i|\leq\epsilon$ and $|\varphi_j(x_j)-x_j|\leq\epsilon$. Therefore, when s is sufficiently large,

$$
\{(x_i,x_j):|\varphi_i(x_i)-\varphi_j(x_j)|\leq\epsilon\}\subset\{(x_i,x_j):|x_i-x_j|\leq 3\epsilon\}. \qquad (56)
$$

Thus, similar to (53),

$$\int_{\mathcal{X}_{ij}} f_i(x_i) f_j(x_j) dx_i dx_j \leq \frac{6\epsilon}{\sqrt{2\pi}\min(\tau_i, \tau_j)}. \tag{57}$$

We observe that

$$\begin{aligned} I_2 \leq & \sum_{i=1}^{k}\sum_{j=1}^{k} \frac{12\epsilon^2}{\sqrt{2\pi}\min(\tau_i,\tau_j)} \\ & +2\sum_{i=1}^{k}\sum_{j=1}^{k}\int_{R^2} P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)| > \frac{\epsilon}{2}\}[|\varphi_i(x_i)-\mu_i| + |\varphi_j(x_j)-\mu_j| \\ & \qquad + |\mu_i - \mu_j|] f_i(x_i) f_j(x_j) dx_i dx_j. \end{aligned} \tag{58}$$

From (54) and (58), it suffices to analyze the limiting behaviors of

$$\begin{aligned} & \textstyle\int_R P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)| > \frac{\epsilon}{2}\} f_i(x_i) dx_i, \\ & \textstyle\int_R P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)| > \frac{\epsilon}{2}\} |\varphi_i(x_i)-\mu_i| f_i(x_i) dx_i. \end{aligned} \tag{59}$$

We first analyze $\int_R P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)| > \frac{\epsilon}{2}\} f_i(x_i) dx_i$. Denote

$$\begin{aligned} \mathcal{Y}_i &= \{x_i : |\varphi_i(x_i) - \theta_i| \leq \tfrac{\epsilon}{4}\}, \\ \mathcal{Z}_i &= \{x_i : |x_i - \theta_i| \leq \tfrac{\epsilon}{8}\}. \end{aligned} \tag{60}$$

By Lemma 4.1, we know that when s is large enough, $|\varphi_i(x_i) - x_i| \leq \frac{\epsilon}{8}$. Therefore, for sufficiently large s, we have

$$R - \mathcal{Y}_i \subset R - \mathcal{Z}_i \tag{61}$$

and

$$\begin{aligned} & \int_R P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)| > \frac{\epsilon}{2}\} f_i(x_i) dx_i \\ \leq & \int_R \left(\int_{R-\mathcal{Z}_i} P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)| > \frac{\epsilon}{2}\} f_i(x_i|\theta_i,\sigma_i^2) h_i(\theta_i|\mu_i,\tau_i^2) dx_i\right) d\theta_i \\ & + \int_R \left(\int_{\mathcal{Y}_i} P_{n,s}\{|\hat{\varphi}_i(x_i)-\varphi_i(x_i)| > \frac{\epsilon}{2}\} f_i(x_i|\theta_i,\sigma_i^2) h_i(\theta_i|\mu_i,\tau_i^2) dx_i\right) d\theta_i \end{aligned}$$

$$
\begin{aligned}
&\le \int_R \Big(\int_{R-\mathcal{Z}_i} f_i(x_i|\theta_i,\sigma_i^2) h_i(\theta_i|\mu_i,\tau_i^2) dx_i \Big) d\theta_i \\
&\quad + \int_R \Big(\int_R P_{n,s}\{|\hat{\varphi}_i(x_i) - \theta_i| \ge \frac{\epsilon}{4}\} f_i(x_i|\theta_i,\sigma_i^2) h_i(\theta_i|\mu_i,\tau_i^2) dx_i \Big) d\theta_i \\
&\le \int_R \Big(\int_{|x_i-\theta_i|>\frac{\epsilon}{8}} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + \int_R \Big(\int_R P_{n,s}\{|\hat{\varphi}_i(x_i) - \theta_i| \ge \frac{\epsilon}{4}, \hat{\nu}_i^2 - 2\hat{\sigma}_i^2/(s+1) > 0\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) \\
&\qquad \times h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + \int_R \Big(\int_R P_{n,s}\{\hat{\nu}_i^2 - 2\hat{\sigma}_i^2/(s+1) \le 0\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\le 2\int_R e^{-\frac{(2s+1)\epsilon^2}{128(e+1)^2\sigma^{*2}}} h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + \int_R \Big(\int_R P_{n,s}\{|(x_i\hat{\tau}_i^2 + \frac{2\hat{\sigma}_i^2}{s+1}\hat{\mu}_i)/\hat{\nu}_i^2 - \theta_i| \ge \frac{\epsilon}{4}\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + o(n^{-1}) \\
&\le 2e^{-\frac{(2s+1)\epsilon^2}{128(e+1)^2\sigma^{*2}}} \\
&\quad + \int_R \Big(\int_R P_{n,s}\{|x_i - \theta_i| \ge \frac{\hat{\nu}_i^2}{\hat{\tau}_i^2}\frac{\epsilon}{8}\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + \int_R \Big(\int_R P_{n,s}\{|\hat{\mu}_i - \theta_i| \ge \frac{(s+1)\hat{\nu}_i^2\epsilon}{16\hat{\sigma}_i^2}\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + o(n^{-1}) \\
&\le O(s^{-1}) \\
&\quad + \int_R \Big(\int_R P_{n,s}\{|x_i - \theta_i| \ge \frac{\hat{\nu}_i^2}{\hat{\tau}_i^2}\frac{\epsilon}{8}, \frac{\hat{\nu}_i^2}{\hat{\tau}_i^2} \ge \nu_i^2/(2\tau_i^2)\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + \int_R \Big(\int_R P_{n,s}\{\frac{\hat{\nu}_i^2}{\hat{\tau}_i^2} \le \nu_i^2/(2\tau_i^2)\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + \int_R \Big(\int_R P_{n,s}\{|\hat{\mu}_i - \theta_i| \ge \frac{(s+1)\hat{\nu}_i^2\epsilon}{16\hat{\sigma}_i^2}, \hat{\nu}_i^2/\hat{\sigma}_i^2 \ge \nu_i^2/(2\sigma_i^2)\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) \\
&\qquad \times h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + \int_R \Big(\int_R P_{n,s}\{\hat{\nu}_i^2/\hat{\sigma}_i^2 \le \nu_i^2/(2\sigma_i^2)\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + o(n^{-1}) \\
&\le O(s^{-1}) + \int_R \Big(\int_{|x_i-\theta_i|\ge\frac{\nu_i^2}{2\tau_i^2}\frac{\epsilon}{16}} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i + o(n^{-1}) \\
&\quad + \int_R \Big(\int_R P_{n,s}\{|\theta_i - \mu_i| \ge \frac{(s+1)\nu_i^2\epsilon}{64\sigma_i^2}\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + \int_R \Big(\int_R P_{n,s}\{|\hat{\mu}_i - \mu_i| \ge \frac{(s+1)\nu_i^2\epsilon}{64\sigma_i^2}\} f_i(x_i|\theta_i,\sigma_i^2) dx_i \Big) h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + o(n^{-1}) \\
&\le O(s^{-1}) + 2e^{-\frac{(2s+1)\nu_i^4\epsilon^2}{2\times 32^2(e+1)^2\sigma^{*2}\tau_i^4}} + \int_{|\theta_i-\mu_i|\ge\frac{(s+1)\nu_i^2\epsilon}{64\sigma_i^2}} h_i(\theta_i|\mu_i,\tau_i^2) d\theta_i \\
&\quad + o(n^{-3}) + o(n^{-1})
\end{aligned}
$$

$$
\begin{aligned}
&\leq O(s^{-1}) + e^{-\frac{(s+1)^2 \nu_i^4 \epsilon^2}{2\times 64^2 \sigma_i^4 \tau_i^2}} + o(n^{-1}) \\
&= O(s^{-1}) + o(n^{-1}).
\end{aligned} \tag{62}
$$

Similarly, we can obtain

$$
\int_R P_{n,s}\{|\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{\epsilon}{2}\}|\varphi_i(x_i) - \mu_i| f_i(x_i) dx_i = o(\frac{1}{n}) + O(\frac{1}{s}). \tag{63}
$$

Combining (49), (54), (58), (59), (62) and (63), we finally obtain the rate of convergence of the proposed selection procedure.

Theorem 1. The selection procedure $d(\underset{\sim}{x})$ defined in (26) is asymptotically optimal with convergence rate of order $o(\frac{1}{n}) + O(\frac{\log s}{s})$. That is,

$$
E(R(d^{(n,s)}(\underset{\sim}{x})) - R(d^B(\underset{\sim}{x}))) = o(\frac{1}{n}) + O(\frac{\log s}{s}). \tag{64}
$$

5 Simulations

We carried out a simulation study to investigate the preformance of the selection procedure $d^{(n,s)}(\underset{\sim}{x})$. The expected risk $E(R(d^{(n,s)}(\underset{\sim}{x})) - R(d^B(\underset{\sim}{x})))$ is used as measure of the performance of the selection rule.

We consider the following case in which $k = 3$, that is, we have 3 logistic populations Π_1, Π_2 and Π_3 and we would like to use the proposed selection procedure to select the best population compared with a control.

The simulation scheme is described as follows:

(1) For each n, s and for each $i = 1, 2, 3$, generate independent random variables $X_{i1}, \ldots,$ X_{iM} as follows:

$$
\begin{cases}
\text{for } \; l = 1, \ldots, n, \\
\text{(a) first generate } \theta_{il} \text{ from normal distribution with density } N(\mu_i, \tau_i^2) \\
\text{(b) then generate } X_{ijl} \text{ from logistic distribution } L(\theta_{il}, \sigma_i)
\end{cases} \tag{65}
$$

(2) Based on the past observations X_{ijl}, and the present observations $\underset{\sim}{X} = (X_1, \ldots, X_k)$, we construct the empirical Bayes selection procedure $d^{(n,s)}(\underset{\sim}{x})$ and compute the conditional difference

$$D = R((d^{(n,s)}(\underline{x}) - R(d^B(\underline{x}))). \tag{66}$$

(3) Repeat steps (1) and (2) 400 times. The average of the conditional differences on the 400 repetitions which is denoted by $\overline{D}(n,s)$, is used as an estimator of the differences $ER((d^{(n,s)}(\underline{x}) - R(d^B(\underline{x})))$.

Tables (1) gives the simulation results on the performance of the proposed empirical Bayes selection procedures. We choose $\theta_0 = 0.5$, $\mu_1 = 0.4$, $\mu_2 = 0.5$, and $\mu_3 = 0.6$, $\tau_1 = \tau_2 = \tau_3 = 1$.

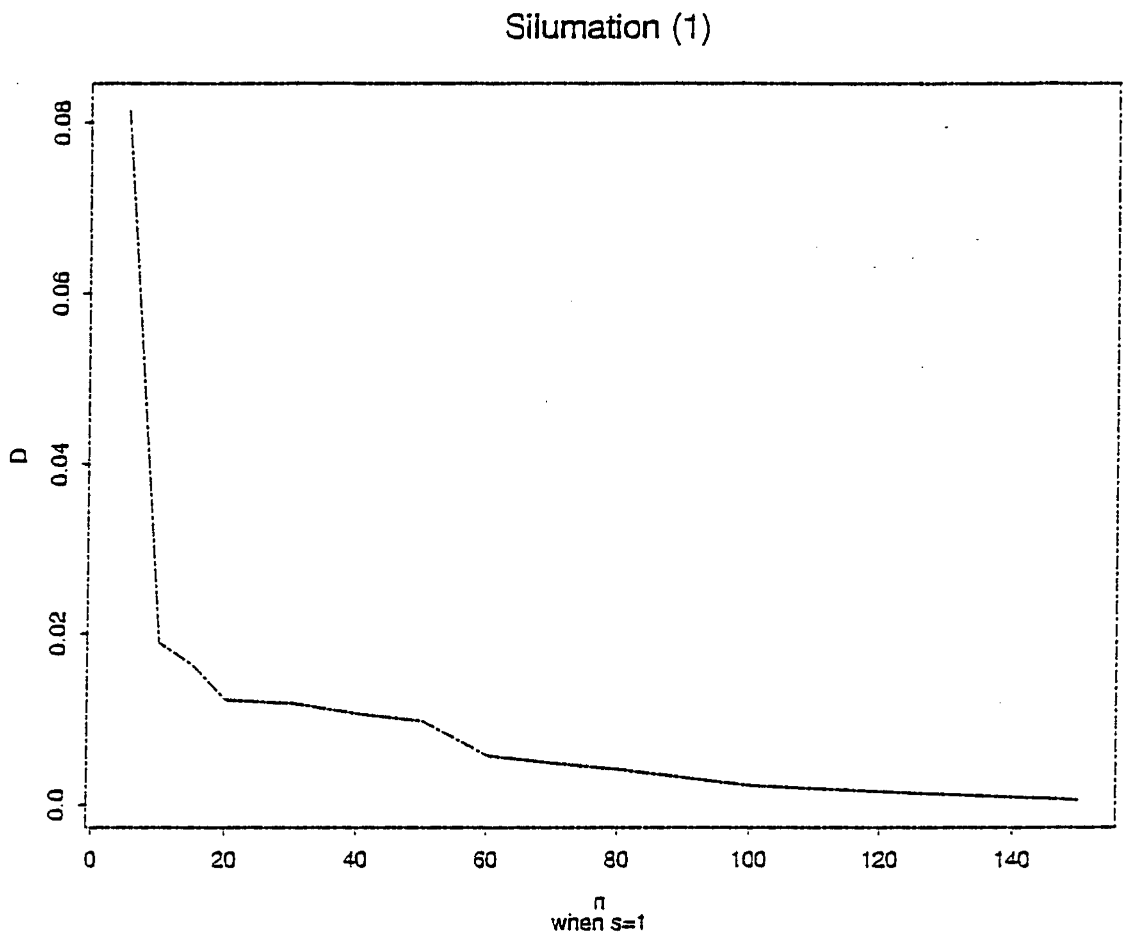
From these results, we see that $\overline{D}(n,s)$ decreases to zero very rapidly. It supports Theorem 1 that the convergence rate is $o(\frac{1}{n}) + O(\frac{\log s}{s})$.

Table 1

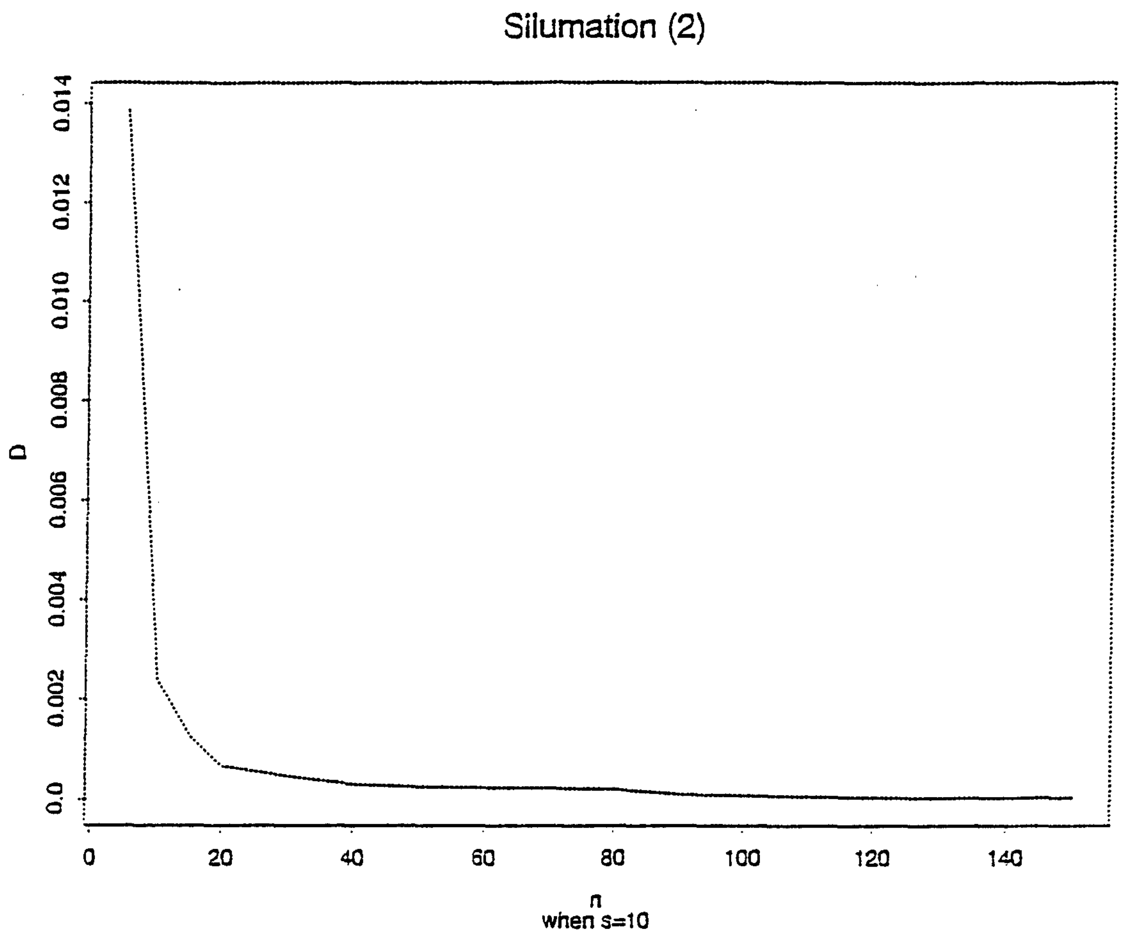
Performance of the selection rule

n	$\overline{D}(n, s=1)$	$\overline{D}(n, s=10)$	$\overline{D}(n, s=50)$
5	0.05132320	0.01647000	0.00560000
10	0.03145200	0.00653760	0.00218600
15	0.00636600	0.00367570	0.00079450
20	0.00389500	0.00293676	0.00010670
30	0.00278474	0.00089434	0.00008610
40	0.00283848	0.00008932	0.00004989
50	0.00019361	0.00023743	0.00003889
60	0.00056436	0.00010391	0.00004021
70	0.00023664	0.00009736	0.00002519
80	0.00035232	0.06272372	0.00001805
90	0.00636233	0.00211873	0.00001781
100	0.00036277	0.00012751	0.00001664
125	0.00326283	0.00032525	0.00001033
150	0.03272747	0.00003257	0.00000819

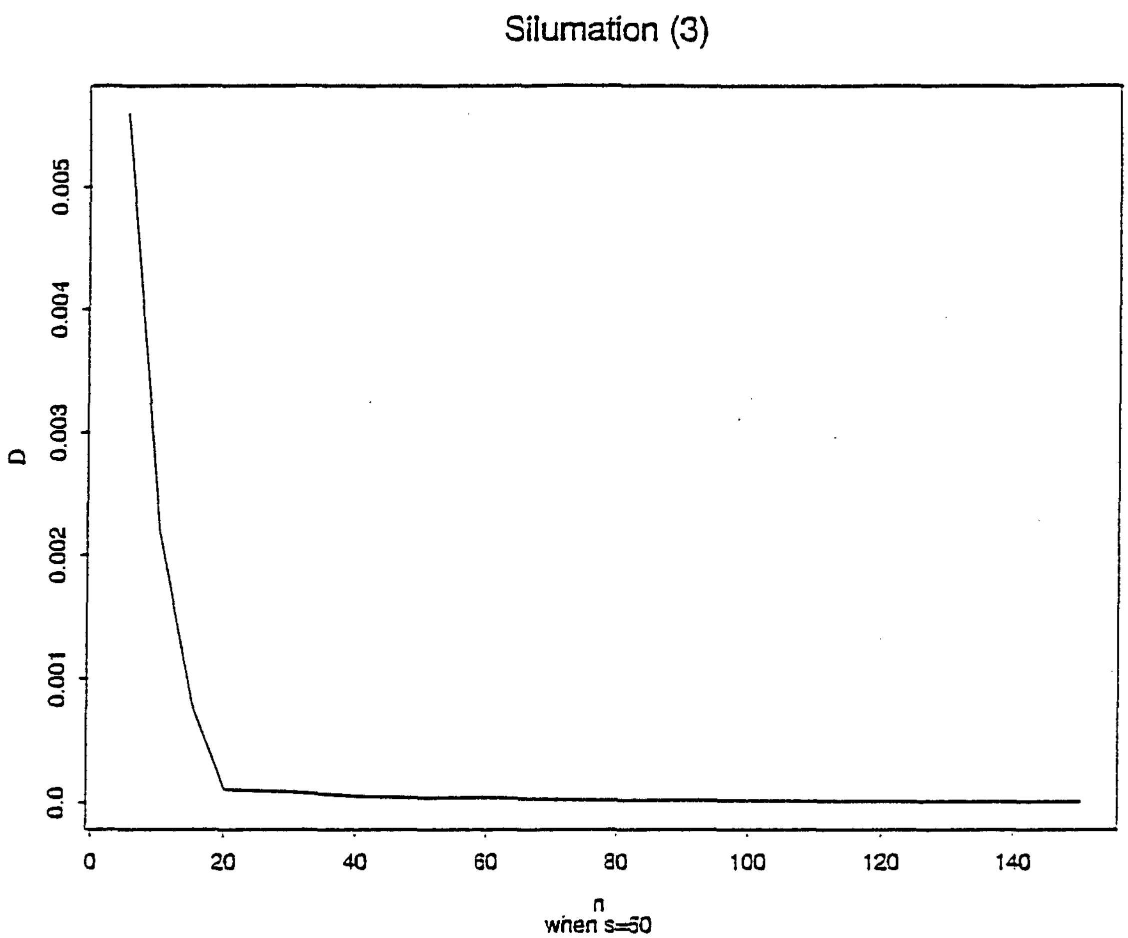
Graph for Table 1 (when $s = 1$)



Graph for Table 1 (when $s = 10$)



Graph for Table 1 (when $s = 50$)



References

Balakrishnan, N. (1992). *Handbook of the logistic distribution.* Marcel Dekker, Inc, New York.

Baum, L. E. and Katz, M. (1965). Convergence Rates in the Law of Large Numbers, Transactions of the Amer. Math. Society, Vol. 120, 1, 108-123.

Berger, J. O. (1985). Statistical Decision Theory and Bayesian Analysis, Springer-Verlag.

Gupta S. S. and Balakrishnan, N. (1992). Logistic order statistics and their properties. *Handbook of the logistic distribution.* (Balakrishnan, N. eds.), Marcel Dekker, Inc, New York, 17-48.

Gupta, S. S. and Han, S., (1991). An Elimination type two-stage procedure for selecting the population with the largest mean for k logistic populations. Amer. Journal Math. Mgmt. Sciences. Vol. 11, 351-370.

Gupta, S. S. and Han, S., (1992). Selection and ranking problems for logistic populations. *Order Statistics and Nonparametrics.* (Eds. I. A. Salama and P. K. Sen), 377-404, Amsterdam, Elsevier Science Publishers B. V.

Gupta S. S. and Liang T. (1996). On empirical Bayes simultaneous selection procedures for comparing normal populations with a standard. Department of Statistics, Purdue University, Technical Report #96-25C.

Gupta S. S. and Lin, X. (1997). A Selection Problem in Measurement Error Models. Department of Statistics, Purdue University, Technical Report #97-18C.

Gupta, S. S. and Panchapakesan S. (1996). Design of experiments with selection and ranking goals. *Handbook of Statistics.* (S. Ghosh and C. R. Rao, eds.), Elsevier Science B. V., Vol. 13, 555-584.

Serfling, J. R. (1980). Approximation Theorems of Mathematical Statistics, New York, Wiley.

SF 298 MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	2/11/98	Final February, 1998

4. TITLE AND SUBTITLE
Empirical Bayes Selection Procedures for Selecting the Best Logistic Population Compared with a Control

5. FUNDING NUMBERS
DAAH04-95-01-0165

6. AUTHOR(S)
Shanti S. Gupta, Zhengyan Lin and Xun Lin

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
Purdue University
West Lafayette, IN 47907-1399

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
ARO 32922.18-MA

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
In this paper we investigate the problem of selecting the best logistic population from $k(\geq 2)$ possible candidates. The selected population must also be better than a given control. We employ the empirical Bayes approach and develop a selection procedure. The performance (rate of convergence) of the proposed selection rule is also analyzed. We also carry out a simulation study to investigate the rate of convergence of the proposed empirical Bayes selection procedure. The results of simulation are provided in the paper.

14. SUBJECT TERMS
Asymptotically optimal; empirical Bayes; logistic population; rate of convergence; selection procedure

15. NUMBER IF PAGES
21

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Enclosure 1

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102