ON A SELECTION PROCEDURE FOR SELECTING
THE BEST LOGISTIC POPULATION COMPARED WITH
A CONTROL*

by

Shanti S. Gupta           Zhengyan Lin**
Purdue University         Hangzhou University

and

Xun Lin
Purdue University

Technical Report #99-04

Department of Statistics
Purdue University
West Lafayette, IN USA

April 1999

*
ON A SELECTION PROCEDURE FOR SELECTING THE BEST LOGISTIC POPULATION COMPARED WITH A CONTROL*

. Shanti S. Gupta and Zhengyan Lin
  Department of Statistics
  Purdue University
  Hangzhou University
  West Lafayette, IN 47907
  Hangzhou, China 310028

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 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 selection procedure. The results of the simulation study are provided in the paper.

AMS Classification: primary 62F07; secondary 62C12.

Keywords: Asymptotically optimal; empirical approach; 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.
ON A SELECTION PROCEDURE FOR SELECTING
THE BEST LOGISTIC POPULATION COMPARED WITH
A CONTROL*

by

Shanti S. Gupta  Zhengyan Lin**
Purdue University  Hangzhou University

and

Xun Lin
Purdue University

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 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 selection procedure. The results of the simulation study are provided in the paper.

AMS Classification: primary 62F07, secondary 62C12.

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

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

** Visiting Professor from Department of Mathematics, Hangzhou University, China, 310028.
1 Introduction

Logistic distributions have been widely used in studies that are related with growth processes. Berkson (1957) used the logistic distribution as a model to analyze quantal response. Plackett (1958) considered the use of the logistic distribution with life test data. The importance of the logistic distribution has resulted in numerous investigations involving the statistical aspects of the distribution. For example, Talacko (1956) showed that it could be a limiting distribution in various situations. Birnbaum and Dudman (1963), and Gupta and Shah (1965) studied its order statistics and their limiting properties. Gupta and Gnanadesikan (1966), and Gupta, Qureishi and Shah (1967) have considered the estimation of parameters of the logistic distribution. Gupta, Qureishi and Shah have constructed the best linear unbiased estimators of both location and scale parameters using order statistics.

It is now well recognized that the classical techniques for testing homogeneity hypotheses are inadequate to serve, in many practical situations, the experimenter’s real purpose, which is to rank several competing populations or to select the best among them. Such realistic goals and formulations set the stage for the development of the ranking and selection theory. An important part of this development is the study of ranking and selection problems for specific parametric families of distributions including, of course, logistic distributions. Gupta and Han (1991) proposed an elimination type procedure based on the estimated sample means for selecting the best logistic population. In addition, Gupta and Han (1992) proposed another selection rule for selecting the best logistic population using the indifference zone approach. A very nice paper on ranking and selection procedures for the logistic populations is Panchapakesan (1992) which is published in “the Handbook of the Logistic Distribution”, edited by Balakrishnan (1992). In this book one can find a good deal of recent developments related to the logistic distribution.

In this paper, we investigate the problem of selecting the best logistic population by using the observed sample medians. Assume that there are $k$ independent logistic populations whose location parameters follow a prior normal distribution and the parameters of the prior normal distribution are unknown. Motivated by the empirical methodology, we propose an empirical selection procedure that is based on the past observed data.

2 Formulation of the Selection Problem and the Selection Rule

Let $\Pi_1, \ldots, \Pi_k$ be $k$ independent logistic populations with unknown means $\theta_1, \ldots, \theta_k$. Let $\theta_{[1]} \leq \cdots \leq \theta_{[k]}$ 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 \( \Pi_i \) with \( \theta_i = \theta_{[k]} \) is considered as the best population. For a given fixed control \( \theta_0 \), population \( \Pi_i \) is defined to be good if the corresponding \( \theta_i > \theta_0 \), and bad otherwise. Our goal is to select the one which is the best among the \( k \) logistic populations and also good compared with the given standard \( \theta_0 \). If there is no such population, we select none.

Let \( \Omega = \{ \theta = (\theta_1, \ldots, \theta_k) \} \) be the parameter space and \( a = (a_0, \ldots, a_k) \) be an action, where \( a_i = 0, \) or \( 1, \) for \( i = 0, 1, \ldots, k, \) and \( \sum_{i=0}^{k} a_i = 1. \) For each \( i = 1, \ldots, k, \) \( a_i = 1 \) means population \( \Pi_i \) is selected as the best among the \( k \) candidates and also good compared with \( \theta_0, \) while \( a_i = 0 \) means population \( \Pi_i \) is not selected either because it is not the best among the \( k \) candidates or because it is bad compared with the control. \( a_0 = 1 \) means that all the \( k \) populations are excluded as bad and none of these \( k \) logistic populations is selected. The following loss function will be considered:

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

For each \( i = 1, \ldots, k, \) let \( X_{i1}, \ldots, X_{iM} \) be a sample of size \( M \) from the \( i \)-th logistic population \( \Pi_i = L(\theta_i, \sigma_i^2) \) which has the following conditional density distribution given \( \theta_i \) and \( \sigma_i^2 \)

\[
\frac{1}{\sigma_i} \frac{e^{-(x_i - \theta_i)/\sigma_i}}{(1 + e^{-(x_i - \theta_i)/\sigma_i})^{2s+2}}, \quad -\infty < x_i < \infty. \tag{1}
\]

For convenience, suppose (for now) \( M \) is an odd number, and we denote \( M = 2s + 1. \) Since logistic distribution is symmetric about its mean, the population mean and median are identical. We assume that for each \( i = 1, \ldots, k, \) the population median (and also the mean) \( \theta_i \) is a realization of random variable \( \Theta_i \) which follows a normal \( N(\mu_i, \tau_i^2) \) prior distribution with parameters \( (\mu_i, \tau_i^2). \) The random variables \( \Theta_1, \ldots, \Theta_k \) are mutually independent. \( \sigma_i^2, \mu_i, \tau_i^2 \) are unknown but fixed. In other words, \( \sigma_i^2, \mu_i, \tau_i^2 \) are fixed nuisance parameters. Let \( X_i \) be the median of \( \{X_{i1}, \ldots, X_{iM}\}, i = 1, \ldots, k, \) then the conditional distribution of \( X_i \) given \( (\theta_i, \sigma_i^2) \) can be explicitly written out as follows:

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

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

\[
EX_i = E(E(X_i|\Theta_i = \theta_i)) = E\Theta_i = \mu_i. \tag{3}
\]

The posterior distribution density of \( \Theta_i \) given \( X_i = x_i \) is proportional to

\[
\frac{(e^{-(x_i - \theta_i)/\sigma_i})^{s+1}}{(1 + e^{-(x_i - \theta_i)/\sigma_i})^{2s+2}} \cdot e^{-\frac{(\theta_i - \mu_i)^2}{\tau_i^2}}, \quad -\infty < \theta_i < \infty. \tag{4}
\]
The selection procedure will be based on the sample medians $X_i$. An estimator of $\Theta_i$ given $X_i = x_i$ is the median of the posterior distribution of $\Theta_i$. For $i = 1, \ldots, k$, denote $\varphi_i(x_i)$ to be the median of the posterior distribution of $\Theta_i$ given $X_i = x_i$.

Let $X = (X_1, \ldots, X_k)$ and $\mathcal{X}$ be the sample space generated by $X$. A selection procedure $d = (d_0, \ldots, d_k)$ is a mapping defined on the sample space $\mathcal{X}$. For every $x \in \mathcal{X}$, $d_i(x), i = 1, \ldots, k$, is the probability of selecting population $\Pi_i$ as the best among the $k$ populations and also good compared with the given control $\theta_0$. $d_0(x)$ is the probability of excluding all $k$ populations as bad and selecting none. Also, $\sum_{i=0}^{k} d_i(x) = 1$, for all $x \in \mathcal{X}$.

We next derive a selection rule $d(x)$ based on the posterior median $\varphi_i(x_i)$, $i = 1, \ldots, k$. For each $x \in \mathcal{X}$, let $I(x) = \{i|\varphi_i(x_i) = \max_{0 \leq j \leq k} \varphi_j(x_j), i = 0, 1, \ldots, k\}$, and $i^* = \min\{i|i \in I(x)\}$. Then based on $\varphi_i(x_i)$, a selection procedure $d(x) = (d_0(x), \ldots, d_k(x))$ is constructed as follows:

\[
\left\{
\begin{array}{l}
d_{i^*}(x) = 1, \\
d_j(x) = 0, \quad \text{for } j \neq i^*.
\end{array}
\right.
\]

Under the preceding statistical model, the expected risk of the selection procedure $d(x)$ is denoted by $R(d(x))$. Denote $h_i(\theta_i|\mu_i, \tau^2_i)$ to be the prior density function of $\Theta_i$ given $(\mu_i, \tau^2_i)$, we have

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

where

- $C = \int_\Theta \max(\{\theta_i\}, \theta_0) dH(\theta)$,
- $H(\Theta)$: the joint distribution of $\Theta = (\theta_1, \ldots, \theta_k)$,
- $f_i(x_i) = \int_{\Theta_i} f_i(x_i|\theta_i, \sigma^2_i) h_i(\theta_i|\mu_i, \tau^2_i) d\theta_i$,
- $f(x) = \prod_{i=1}^{k} f_i(x_i)$,
- $\varphi_0(x_0) = \theta_0$.

Note that sample median $X_i$ is not a sufficient statistic for $\theta_i$ (the observation vector is a minimal sufficient statistic). So $d(x)$ may not be a Bayes rule. Also, the selection procedure $d(x)$ defined above depends on the unknown parameters $(\mu_i, \tau^2_i), i = 1, \ldots, k$ 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 this selection procedure for the selection problem in practice.

To derive a practical selection rule, we assume there are 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 $\Pi_i$, that is, for each $i = 1, \ldots, k$, let

\[
\Theta_{\ell l} \sim N(\mu_i, \tau^2_i), \quad l = 1, \ldots, n,
\]

and

\[
X_{ijl} \sim L(\theta_{ij}, \sigma^2_i), \quad j = 1, \ldots, M.
\]
For $l = 1, \ldots, n$, denote $X_{i,l}$ to be the median of $(X_{i1}, \ldots, X_{iM})$, and

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

(9)

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

(10)

Then,

$$E(X_{i,l}) = E(E(X_{i,l} | \theta_d = \theta_d)) = E(\Theta_d) = \mu_i,$$

(11)

and

$$Var(X_{i,l}) = Var(E(X_{i,l} | \theta_d)) + E(Var(X_{i,l} | \theta_d))$$

$$= Var(\Theta_d) + E(Var(X_{i,l} | \theta_d))$$

$$= \tau_i^2 + E(Var(X_{i,l} | \theta_d))$$

< \infty.  

(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 $n \to \infty$,

$$\left\{ \begin{array}{l}
X_i(n) \to \mu_i, \\
S_i^2(n) \to \nu_i^2,
\end{array} \right. \quad a.s.$$

(13)

To derive an empirical selection procedure, we first consider the following lemmas.

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

$$P\{|\xi - \xi| > \epsilon\} \leq 2e^{-2m\delta^2},$$

(14)

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

This lemma is from Serfling (1980) and the proof can be found in it.

Back to our selection problem. Put

$$\sigma' = \min_{1 \leq i \leq k} \sigma_i, \quad \sigma^* = \max_{1 \leq i \leq k} \sigma_i.$$

$X_{i1}, \ldots, X_{iM}$ are i.i.d. from $L(\theta_i, \sigma_i^2)$, which has the following cumulative distribution function

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

(15)
and for $0 < \epsilon \leq \sigma'$,

$$F(\theta_i + \epsilon) - \frac{1}{2} = \frac{1}{2} - F(\theta_i - \epsilon) = \frac{e^{x_i^{(s)}} - 1}{2(e^{x_i^{(s)}} + 1)} \geq \frac{\epsilon}{2(e + 1)\sigma^*}. \quad (16)$$

Given $\Theta_i = \theta_i$, $\theta_i$ and $X_i$ are the population median and sample median of $L(\theta_i, \sigma_i)$ respectively. We have, from Lemma 1,

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

For any $0 < \epsilon \leq \sigma'$, denote $A_i = \{x \in \mathcal{X} : |x_i - \theta_i| \leq \epsilon\}$. We show that the conditional density of $X_i$ given $\theta_i$ and $\sigma_i^2$ 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 $\theta_i$ and $\sigma_i^2$ is

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

$$= \frac{(2s+1)!}{(s!)^2} \frac{1}{\sigma_i} \frac{1}{(2 + e^{-{(x_i-\theta_i)/\sigma_i}} + e^{(x_i-\theta_i)/\sigma_i})^{s+1}}. \quad (18)$$

By Stirling’s formula, when $s$ is large enough,

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

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

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

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

$$P\{X \not\in A_i\} \leq 2e^{-\frac{(2s+1)\epsilon^2}{2(\epsilon^2 + 1)}} \to 0. \quad (21)$$

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

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

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 $\theta_i$ and $\sigma_i$. Since the prior distribution of $\theta_i$ is $N(\mu_i, r_i^2)$, the unconditional density of $X_{i,l}$ is approximately $N(\mu_i, r_i^2 + \frac{2}{s+1} \sigma_i^2)$. 

6
For each population $\Pi_i$, let $W_i^2(n)$ be the measure of the overall sample variation for the past observations. That is,
\[
\begin{align*}
&\bar{X}_{il} = \frac{1}{M} \sum_{j=1}^{M} X_{ijl}, \\
&W_i^2(n) = \frac{1}{(M-1)n} \sum_{j=1}^{M} \sum_{l=1}^{n} (X_{ijl} - \bar{X}_{il})^2.
\end{align*}
\]

(23)

Then we define, for $i = 1, \ldots, k$,
\[
\begin{align*}
\hat{\mu}_i &= X_i(n), \\
\hat{\sigma}_i^2 &= \frac{3}{n^2} W_i^2(n), \\
\hat{\nu}_i^2 &= S_i^2(n), \\
\hat{\tau}_i^2 &= \max(\hat{\nu}_i^2 - \frac{2}{s+1} \hat{\sigma}_i^2, 0).
\end{align*}
\]

(24)

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

(25)

Then for each $x \in \mathcal{X}$, let $\hat{I}(x) = \{i|\hat{\varphi}_i(x_i) = \max_{0 \leq j \leq k} \hat{\varphi}_j(x_j), i = 0, 1, \ldots, k\}$, and $\hat{i}^* = \min\{i|i \in \hat{I}(x)\}$. We propose the following selection procedure $d^{(n,s)}(x) = (d_0^{(n,s)}(x), \ldots, d_k^{(n,s)}(x))$ as follows:
\[
\begin{align*}
&d_{\hat{i}^*}^{(n,s)} = 1, \\
&d_{j}^{(n,s)} = 0, \quad \text{for } j \neq \hat{i}^*.
\end{align*}
\]

(26)

3 Asymptotic Optimality of the Proposed Selection Procedure

Consider the selection procedure $d^{(n,s)}(x)$ constructed in (26). $d^{(n,s)}(x)$ is similar to selection rule $d(x)$ except that normal approximation is used to estimate the unknown prior parameters and the specific form of $\varphi_i(x_i)$ for $d^{(n,s)}(x)$. A natural question to ask is: How good is the selection rule $d^{(n,s)}(x)$ compared with $d(x)$? Let $R(d^{(n,s)}(x))$ be the conditional expected risk given the past observations $\{X_{ijl}, i = 1, \ldots, k, j = 1, \ldots, M, \text{and } l = 1, \ldots, n\}$, then
\[
R(d^{(n,s)}(x)) = -\int_{\mathcal{X}} \left[ \sum_{i=0}^{k} d_i^{(n,s)}(x) \varphi_i(x_i) \right] f(x) d(x) + C.
\]

(27)
Since \( d^{(n,s)}(x) \) is mimicking \( d(x) \), \( R(d^{(n,s)}(x)) - R(d(x)) \) should be close to 0 if the empirical selection rule works well. Note that \( R(d^{(n,s)}(x)) - R(d(x)) \) can be negative because \( d(x) \) is not a Bayes rule. Therefore, we use the overall integrated risk \( E|R(d^{(n,s)}(x)) - R(d(x))| \geq 0 \) as a measure of the performance of the selection procedure \( d^{(n,s)}(x) \), where \( E \) is the expectation taken with respect to the past observations \( \{X_i\} \).

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

**Lemma 2** When \( s \) is large enough, for \( 1 \leq i \leq k \),

\[
|\varphi_i(x_i) - x_i| \leq 2\sigma_i \sqrt{\log s \over s} \tag{28}
\]

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

\[
\int_{x_i + 2\sigma_i \sqrt{\log s \over s}}^{\infty} f_i(x_i|\theta_i, \sigma_i^2) h_i(\theta_i|\mu_i, \tau_i^2) \, d\theta_i
\]

\[
= \int_{x_i + 2\sigma_i \sqrt{\log s \over s}}^{\infty} \frac{(2s + 1)!}{(s!)^2} \frac{1}{\sigma_i \sqrt{2\pi \tau_i}} \frac{1}{(1 + e^{(\theta_i-x_i)/\sigma_i})^{2s+2}} \cdot e^{-\frac{(\theta_i-\mu_i)^2}{2\tau_i}} \, d\theta_i
\]

\[
= \int_{2\sqrt{\log s \over s} \sqrt{2\pi \tau_i}}^{\infty} \frac{1}{\sqrt{2\pi \tau_i}} \frac{1}{(s!)^2} \frac{1}{(1 + e^\theta)^2} \cdot e^{-\frac{e^\theta + x_i - \mu_i^2}{2\tau_i}} \, d\theta \rightarrow 0, \tag{29}
\]

as \( s \rightarrow \infty \). We first show

\[
t(\theta, s) := \frac{(2s + 1)!}{(s!)^2} \left( \frac{e^\theta}{1 + e^\theta} \right)^{s+1} \rightarrow 0 \quad \text{as} \quad s \rightarrow \infty, \tag{30}
\]

uniformly for \( \theta \geq 2\sqrt{\log s \over s} \). Obviously it is enough to consider the case of \( \theta = 2\sqrt{\log s \over s} \) since \( t(\theta, s) \) is decreasing on \( \theta > 0 \). When \( \theta = 2\sqrt{\log s \over s} \) 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

\[
\log t(\theta, s) = (s + 1)[\theta - 2 \log(1 + e^\theta)] + \log \frac{(2s + 1)!}{(s!)^2}
\]

8
\begin{equation}
\begin{aligned}
&\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) \\
&= -\left(\frac{s+1}{s} - \frac{1}{2}\right) \log s + o(\log s) \rightarrow -\infty,
\end{aligned}
\end{equation}

as \( s \to \infty \). Therefore, (30) is proved, from which we can immediately see that (29) holds true. It completes the proof of Lemma 2.

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

**Lemma 3** 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 \( \epsilon > 0 \),

\begin{equation}
P\{|\sum_{i=1}^{n} X_i/n| \geq \epsilon\} = o(n^{-(\alpha-1)}).
\end{equation}

As a consequence of Lemma 3, we have

**Lemma 4** Let \( X_1, \ldots, X_n \) be independent random variables, with mean \( EX_i = \mu \) and variance \( \text{Var}X_i = \sigma^2 \), 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 \( \epsilon > 0 \),

\begin{equation}
P\{|S_n^2 - \sigma^2| \geq \epsilon\} = o(n^{-(\alpha/2-1)}).
\end{equation}

Since \( EX_i^4 < \infty \), for any \( \epsilon > 0 \), by Lemma 3,

\begin{equation}
P\{|ar{\mu}_i - \mu_i| \geq \epsilon\} = o(n^{-3}),
\end{equation}

also by Lemma 4,

\begin{equation}
P\{\hat{\nu}_i^2 - \nu_i^2 \geq \epsilon\} = o(n^{-1}).
\end{equation}

Similarly, we have for any \( \epsilon > 0 \),

\begin{equation}
P\{|\hat{\sigma}_i^2 - \sigma^2_i| \geq \epsilon\} = o(n^{-1}).
\end{equation}

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

\begin{equation}
P\{\hat{\nu}_i^2 - \frac{2}{s+1} \hat{\sigma}_i^2 \leq 0\} = o(n^{-1}).
\end{equation}

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

\begin{equation}
E(\text{Var}(X_{i,l}|	heta_i)) = \int_{-\infty}^{\infty} (x_i - \theta_i)^2 (2s+1)! \frac{1}{(s!)^2} \frac{(1 + e^{-(x_i - \theta_i)/\sigma_i})^{s+1}}{\sigma_i} \frac{e^{x_i^2}}{(1 + e^{x_i})^2} dx_i
\end{equation}

\begin{equation}
= \sigma_i^2 \int_{-\infty}^{\infty} x_i^2 (2s+1)! \frac{1}{(s!)^2} \frac{e^{x_i^2}}{(1 + e^{x_i})^2} dx_i.
\end{equation}
We have

Lemma 5

\[
\int_{-\infty}^{\infty} x^2 \frac{(2s+1)!}{(s!)^2} \left( \frac{e^x}{1+e^x} \right)^{s+1} dx = o\left( \sqrt{\frac{\log s}{s}} \right). \tag{41}
\]

Proof.

\[
\begin{align*}
\int_{-\infty}^{\infty} x^2 \frac{(2s+1)!}{(s!)^2} \left( \frac{e^x}{1+e^x} \right)^{s+1} dx \\
= 2 \int_{0}^{\infty} x^2 \frac{(2s+1)!}{(s!)^2} \left( \frac{e^x}{1+e^x} \right)^{s+1} dx \\
= 2 \left( \int_{0}^{\sqrt{\frac{\log s}{s}}} + \int_{\sqrt{\frac{\log s}{s}}}^{3} + \int_{3}^{\infty} \right) x^2 \frac{(2s+1)!}{(s!)^2} \left( \frac{e^x}{1+e^x} \right)^{s+1} dx \\
:= T_1 + T_2 + T_3. \tag{42}
\end{align*}
\]

By Stirling's formula, when \( s \) is large enough,

\[
T_1 = 2 \frac{(2s+1)!}{(s!)^2} \int_{0}^{\sqrt{\frac{\log s}{s}}} x^2 \left( \frac{e^x}{1+e^x} \right)^{s+1} dx \\
\leq 2 \cdot 2^{2(s+1)} \sqrt{s+1} \cdot 2^{-2(s+1)} \int_{0}^{\sqrt{\frac{\log s}{s}}} x^2 dx \\
\leq \sqrt{s+1} \left( \frac{\log s}{s} \right)^{3/2} \\
= o\left( \sqrt{\frac{\log s}{s}} \right). \tag{43}
\]

Using the same approach as in the proof of Lemma 2, we have

\[
T_2 = 2 \frac{(2s+1)!}{(s!)^2} \int_{\sqrt{\frac{\log s}{s}}}^{3} x^2 \left( \frac{e^x}{1+e^x} \right)^{s+1} dx \\
\leq 2 \frac{(2s+1)!}{(s!)^2} \left( \frac{e^{\sqrt{\frac{\log s}{s}}}}{1+e^{\sqrt{\frac{\log s}{s}}}} \right)^{s+1} \int_{\sqrt{\frac{\log s}{s}}}^{3} x^2 dx \\
= o\left( \sqrt{\frac{\log s}{s}} \right). \tag{44}
\]

Moreover,

\[
T_3 = 2 \frac{(2s+1)!}{(s!)^2} \int_{3}^{\infty} x^2 \left( \frac{e^x}{1+e^x} \right)^{s+1} dx \\
\leq 2 \frac{(2s+1)!}{(s!)^2} \int_{3}^{\infty} x^2 e^{-(s+1)x} dx \\
= o\left( \sqrt{\frac{\log s}{s}} \right). \tag{45}
\]
This completes the proof of Lemma 5.

From Lemma 5, we observe that when $s$ is sufficiently large,

$$E(Var(X_i;|i|)) = o\left(\sqrt{\frac{\log s}{s}}\right), \quad (46)$$

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

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

and furthermore,

$$P\{\frac{\hat{c}^2_i - \hat{c}^2_i}{\hat{c}^2_i/(2\hat{c}^2_i)} = o(n^{-1})\} \quad (48)$$

Next we investigate the overall integrated risk $E[R(d^{(n,s)}(x)) - R(d(x))]$. Let $P_{n,s}$ be the probability measure generated by the past observations $X_{ij}, i = 1, \ldots k, j = 1, \ldots M$ and $l = 1, \ldots, n$.

$$E[R(d^{(n,s)}(x)) - R(d(x))] \leq \sum_{i=0}^{k} \sum_{j=0}^{k} \int_X P_{n,s}\{i^* = i, j^* = j\}|\varphi_i(x) - \varphi_j(x)|f(x)dx$$

$$= \sum_{i=1}^{k} \int_X P_{n,s}\{i^* = i, j^* = 0\}|\varphi_i(x) - \theta_0|f(x)dx$$

$$+ \sum_{j=1}^{k} \int_X P_{n,s}\{i^* = 0, j^* = j\}|\theta_0 - \varphi_j(x)|f(x)dx$$

$$+ \sum_{i=1}^{k} \sum_{j=1}^{k} \int_X P_{n,s}\{i^* = i, j^* = j\}|\varphi_i(x) - \varphi_j(x)|f(x)dx$$

$$\leq 2 \sum_{i=1}^{k} \int_R P_{n,s}\{|\varphi_i(x) - \varphi_i(x_i)| > |\varphi_i(x) - \theta_0|\}|\varphi_i(x_i) - \theta_0|f_i(x_i)dx_i$$

$$+ 2 \sum_{i=1}^{k} \sum_{j=1}^{k} \int_R P_{n,s}\{|\varphi_i(x) - \varphi_i(x_i)| > |\varphi_i(x_i) - \varphi_j(x_j)|\}|\varphi_i(x_i) - \varphi_j(x_j)|$$

$$\times f_i(x_i)f_j(x_j)dx_i dx_j$$

$$:= I_1 + I_2. \quad (49)$$

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

$$X_i = \{x_i : |\varphi_i(x_i) - \theta_0| \leq \epsilon\},$$

$$X_{ij} = \{(x_i, x_j) : |\varphi_i(x_i) - \varphi_j(x_j)| \leq \epsilon\}. \quad (50)$$

Then we have

$$I_1 = 2 \sum_{i=1}^{k} \int_{X_i} P_{n,s}\{|\varphi_i(x) - \varphi_i(x_i)| > |\varphi_i(x) - \theta_0|\}|\varphi_i(x) - \theta_0|f_i(x_i)dx_i$$
\[ +2 \sum_{i=1}^{k} \int_{R_{-X_i}} P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > |\varphi_i(x_i) - \theta_0| \} |\varphi_i(x_i) - \theta_0| f_i(x_i) dx_i \]

\[ \leq 2 \sum_{i=1}^{k} \int_{X_i} \epsilon f_i(x_i) dx_i \]

\[ +2 \sum_{i=1}^{k} \int_{R} P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \epsilon \} |\varphi_i(x_i) - \theta_0| f_i(x_i) dx_i. \quad (51) \]

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

\[ |x_i - \theta_0| \leq |\varphi_i(x_i) - x_i| + |\varphi_i(x_i) - \theta_0| \leq 2\epsilon \quad (52) \]
on \( X_i \) and

\[ \int_{X_i} f_i(x_i) dx_i \leq \int_{\{|x_i - \theta_0| \leq 2\epsilon\}} f_i(x_i) dx_i \]

\[ \leq \int_{\{|x_i - \theta_0| \leq 2\epsilon\}} \frac{1}{\sqrt{2\pi \tau_i}} dx_i \]

\[ = \frac{4\epsilon}{\sqrt{2\pi \tau_i}}. \quad (53) \]

Thus,

\[ I_1 \leq \frac{8k}{\sqrt{2\pi \tau_i}} \epsilon^2 \]

\[ +2 \sum_{i=1}^{k} \int_{R} P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \epsilon \} |\varphi_i(x_i) - \mu_i| + |\mu_i - \theta_0| f_i(x_i) dx_i. \quad (54) \]

Moreover,

\[ I_2 = 2 \sum_{i=1}^{k} \sum_{j=1}^{k} \int_{X_{ij}} P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{|\varphi_i(x_i) - \varphi_j(x_j)|}{2} \} |\varphi_i(x_i) - \varphi_j(x_j)| \]

\[ \times f_i(x_i) f_j(x_j) dx_i dx_j \]

\[ +2 \sum_{i=1}^{k} \sum_{j=1}^{k} \int_{R_{-X_i}} P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{|\varphi_i(x_i) - \varphi_j(x_j)|}{2} \} |\varphi_i(x_i) - \varphi_j(x_j)| \]

\[ \times f_i(x_i) f_j(x_j) dx_i dx_j \]

\[ \leq 2\epsilon \sum_{i=1}^{k} \sum_{j=1}^{k} \int_{X_{ij}} f_i(x_i) f_j(x_j) dx_i dx_j \]

\[ +2 \sum_{i=1}^{k} \sum_{j=1}^{k} \int_{R^2} P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{\epsilon}{2} \} |\varphi_i(x_i) - \varphi_j(x_j)| \]

\[ \times f_i(x_i) f_j(x_j) dx_i dx_j. \quad (55) \]
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,

\[
\{(x_i, x_j) : |\varphi_i(x_i) - \varphi_j(x_j)| \leq \epsilon\} \subset \{(x_i, x_j) : |x_i - x_j| \leq 3\epsilon\}. \tag{56}
\]

Thus, similar to (53),

\[
\int_{X_{ij}} f_i(x_i) f_j(x_j) dx_i dx_j \leq \frac{6\epsilon}{\sqrt{2\pi} \min(\tau_i, \tau_j)}. \tag{57}
\]

We observe that

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

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

\[
\int_R P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{\epsilon}{2} \} f_i(x_i) dx_i,
\]

\[
\int_R P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{\epsilon}{2} \} |\varphi_i(x_i) - \mu_i| f_i(x_i) dx_i. \tag{59}
\]

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

\[
\mathcal{Y}_i = \{ x_i : |\varphi_i(x_i) - \theta_i| \leq \frac{\epsilon}{8} \},
\]

\[
\mathcal{Z}_i = \{ x_i : |x_i - \theta_i| \leq \frac{\epsilon}{8} \}. \tag{60}
\]

By Lemma 2, 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

\[
\int_R P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{\epsilon}{2} \} f_i(x_i) dx_i \\
\leq \int_R \left( \int_{R - \mathcal{Z}_i} P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{\epsilon}{2} \} f_i(x_i|\theta_i, \sigma_i^2) h_i(\theta_i|\mu_i, \tau_i^2) dx_i \right) d\theta_i \\
+ \int_R \left( \int_{\mathcal{Y}_i} P_{n,s} \{ |\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{\epsilon}{2} \} f_i(x_i|\theta_i, \sigma_i^2) h_i(\theta_i|\mu_i, \tau_i^2) dx_i \right) d\theta_i \\
\leq \int_R \left( \int_{R - \mathcal{Z}_i} f_i(x_i|\theta_i, \sigma_i^2) h_i(\theta_i|\mu_i, \tau_i^2) dx_i \right) d\theta_i \\
+ \int_R \left( \int_R P_{n,s} \{ |\varphi_i(x_i) - \theta_i| \geq \frac{\epsilon}{4} \} f_i(x_i|\theta_i, \sigma_i^2) h_i(\theta_i|\mu_i, \tau_i^2) dx_i \right) d\theta_i \\
\leq \int_R \left( \int_{|\varphi_i - \theta_i| > \frac{\epsilon}{8}} f_i(x_i|\theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i|\mu_i, \tau_i^2) d\theta_i.
\]
\begin{equation}
\begin{aligned}
+ \int_R \left( \int_R P_{n,s} \{ |i(x_i, \theta_i) - \mu_i| \geq \frac{\epsilon}{4}, \nu_i^2 - 2\sigma_i^2/(s+1) > 0 \} f_i(x_i | \theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i \\
+ \int_R \left( \int_R P_{n,s} \{ \nu_i^2 - 2\sigma_i^2/(s+1) \leq 0 \} f_i(x_i | \theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i \\
\leq 2 \int_R e^{-\frac{(2s+1)\epsilon}{128(s+1)2^{s+2}}} h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i \\
+ \int_R \left( \int_R P_{n,s} \{ |i(x_i, \theta_i) - \mu_i| \geq \frac{\epsilon}{4}, \nu_i^2 - 2\sigma_i^2/(s+1) \geq 0 \} f_i(x_i | \theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i \\
+ o(n^{-1})
\end{aligned}
\end{equation}
\begin{equation}
\begin{aligned}
\leq 2e^{-\frac{(2s+1)\epsilon}{128(s+1)^22^{s+2}}} \\
+ \int_R \left( \int_R P_{n,s} \{ |i(x_i, \theta_i) - \mu_i| \geq \frac{\epsilon}{4}, \nu_i^2 - 2\sigma_i^2/(s+1) \geq 0 \} f_i(x_i | \theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i \\
+ \int_R \left( \int_R P_{n,s} \{ |\hat{\mu}_i - \theta_i| \geq \frac{(s+1)\nu_i^2}{16\sigma_i^2} \} f_i(x_i | \theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i \\
+ o(n^{-1})
\end{aligned}
\end{equation}
\begin{equation}
\begin{aligned}
\leq O(s^{-1}) \\
+ \int_R \left( \int_R P_{n,s} \{ |i(x_i, \theta_i) - \mu_i| \geq \frac{\epsilon}{4}, \nu_i^2 - 2\sigma_i^2/(s+1) \geq 0 \} f_i(x_i | \theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i \\
+ \int_R \left( \int_R P_{n,s} \{ \nu_i^2 - 2\sigma_i^2/(s+1) \leq \frac{(s+1)\nu_i^2}{16\sigma_i^2} \} f_i(x_i | \theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i \\
+ \int_R \left( \int_R P_{n,s} \{ |\hat{\mu}_i - \theta_i| \geq \frac{(s+1)\nu_i^2}{16\sigma_i^2} \} f_i(x_i | \theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i \\
+ o(n^{-1})
\end{aligned}
\end{equation}
\begin{equation}
\begin{aligned}
\leq O(s^{-1}) + \int_R \left( \int_{|i(x_i, \theta_i) - \mu_i| \geq \frac{\epsilon}{4}, \nu_i^2 - 2\sigma_i^2/(s+1) \geq 0} f_i(x_i | \theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i + o(n^{-1})
\end{aligned}
\end{equation}
\begin{equation}
\begin{aligned}
+ \int_R \left( \int_R P_{n,s} \{ |\hat{\mu}_i - \theta_i| \geq \frac{(s+1)\nu_i^2}{16\sigma_i^2} \} f_i(x_i | \theta_i, \sigma_i^2) dx_i \right) h_i(\theta_i | \mu_i, \tau_i^2) d\theta_i \\
+ o(n^{-3}) + o(n^{-1})
\end{aligned}
\end{equation}
\begin{equation}
\leq O(s^{-1}) + o(n^{-1})
\end{equation}
\begin{equation}
= O(s^{-1}) + o(n^{-1}).
\end{equation}
Using similar approach, we can obtain

\[
\int_R P_{n,s}\{|\hat{\varphi}_i(x_i) - \varphi_i(x_i)| > \frac{c}{2}\}|\varphi_i(x_i) - \mu_i|f_i(x_i)dx_i = o\left(\frac{1}{n}\right) + O\left(\frac{1}{s}\right). \quad (63)
\]

At the beginning of this paper, \( M \) is assumed to be an odd number. However, from the proof we can see that this condition can be dropped. In other words, no matter \( M \) is even or odd, the asymptotic property will hold true. Combining (49), (54), (58), (59), (62) and (63), we finally obtain the asymptotic property of the derived selection procedure.

**Theorem 1** The selection procedure \( d^{(n,s)}(\bar{x}) \) defined in (26) is asymptotically optimal with a convergence rate of order \( o\left(\frac{1}{n}\right) + O\left(\frac{\log s}{s}\right) \). That is,

\[
E|R(d^{(n,s)}(\bar{x})) - R(d(\bar{x}))| = o\left(\frac{1}{n}\right) + O\left(\frac{\log s}{s}\right). \quad (64)
\]

Theorem 1 establishes the rate of convergence of \( E|R(d^{(n,s)}(\bar{x})) - R(d(\bar{x}))| \) as both \( n \) and \( s \) go to infinity in an additive form. This implies that \( E|R(d^{(n,s)}(\bar{x})) - R(d(\bar{x}))| \) will converge to 0 when both \( n \) and \( s \) go to infinity.

### 4 Simulations

We carried out a simulation study to investigate the performance of the selection procedure \( d^{(n,s)}(\bar{x}) \). The overall integrated risk \( E|R(d^{(n,s)}(\bar{x})) - R(d(\bar{x}))| \) 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 \( \Pi_1, \Pi_2 \) and \( \Pi_3 \) 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 \( i \), generate past observations as follows:

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

2. For each \( i \), generate current observations \( \Theta_i \) from \( N(\mu_i, \tau_i^2) \) and \( (X_{i1}, \ldots, X_{iM}) \) i.i.d. from \( L(\theta_i, \sigma_i) \).

3. Based on the past observations \( X_{ijl} \), and the present observations, we construct \( d(\bar{x}) \) and \( d^{(n,s)}(\bar{x}) \). Then compute the losses \( L(d(\bar{x})) \) and \( L(d^{(n,s)}(\bar{x})) \).

4. Repeat Steps (2) and (3) 1000 times. Calculate the averages of the conditional losses \( L(d(\bar{x})) \) and \( L(d^{(n,s)}(\bar{x})) \), respectively. Denote the averages to be \( \hat{R}(d(\bar{x})) \) and \( \hat{R}(d^{(n,s)}(\bar{x})) \). Then compute the absolute difference

\[
D = |\hat{R}(d^{(n,s)}(\bar{x})) - \hat{R}(d(\bar{x}))|. \quad (66)
\]
(5) Repeat steps (1), (2), (3) and (4) 5000 times. The average of the $D$s in (66), denoted by $D(n, s)$, is used as an estimator of the differences $E[R(d^{(n,s)}(x))] - R(d(x))]$.

Tables 1, 2, and 3 give the results of simulation for the performance of the proposed empirical selection procedures. We choose $\theta_0 = 0.5$, $\mu_1 = 0.4$, $\mu_2 = 0.5$, $\mu_3 = 0.6$, $\tau_1 = \tau_2 = \tau_3 = 1$, and $\sigma_1 = \sigma_2 = \sigma_3 = 1$. The related figures are also attached.
References


Table 1

Performance of the selection rule

when $s = 5$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$D(n, s)$</th>
<th>$SE(D(n, s))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.08509841</td>
<td>0.027865659</td>
</tr>
<tr>
<td>10</td>
<td>0.04994583</td>
<td>0.017846225</td>
</tr>
<tr>
<td>15</td>
<td>0.02375784</td>
<td>0.009525936</td>
</tr>
<tr>
<td>20</td>
<td>0.01263961</td>
<td>0.006187404</td>
</tr>
<tr>
<td>30</td>
<td>0.00738123</td>
<td>0.003425718</td>
</tr>
<tr>
<td>40</td>
<td>0.00651246</td>
<td>0.003046143</td>
</tr>
<tr>
<td>50</td>
<td>0.00519491</td>
<td>0.002195022</td>
</tr>
<tr>
<td>60</td>
<td>0.00415648</td>
<td>0.001433615</td>
</tr>
<tr>
<td>70</td>
<td>0.00381372</td>
<td>0.001086059</td>
</tr>
<tr>
<td>80</td>
<td>0.00361584</td>
<td>0.000818471</td>
</tr>
<tr>
<td>90</td>
<td>0.00347570</td>
<td>0.000713619</td>
</tr>
<tr>
<td>100</td>
<td>0.00338605</td>
<td>0.000689365</td>
</tr>
<tr>
<td>125</td>
<td>0.00316974</td>
<td>0.000642861</td>
</tr>
<tr>
<td>150</td>
<td>0.00309849</td>
<td>0.000627380</td>
</tr>
</tbody>
</table>
Table 2

Performance of the selection rule

when $s = 10$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$D(n, s)$</th>
<th>$SE(D(n, s))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.07304765</td>
<td>0.019237152</td>
</tr>
<tr>
<td>10</td>
<td>0.03923180</td>
<td>0.012678379</td>
</tr>
<tr>
<td>15</td>
<td>0.01681043</td>
<td>0.008253716</td>
</tr>
<tr>
<td>20</td>
<td>0.01064427</td>
<td>0.004316348</td>
</tr>
<tr>
<td>30</td>
<td>0.00517343</td>
<td>0.002579812</td>
</tr>
<tr>
<td>40</td>
<td>0.00325936</td>
<td>0.001308793</td>
</tr>
<tr>
<td>50</td>
<td>0.00216874</td>
<td>0.000845426</td>
</tr>
<tr>
<td>60</td>
<td>0.00157431</td>
<td>0.000393872</td>
</tr>
<tr>
<td>70</td>
<td>0.00128317</td>
<td>0.000251801</td>
</tr>
<tr>
<td>80</td>
<td>0.00107369</td>
<td>0.000137618</td>
</tr>
<tr>
<td>90</td>
<td>0.00098735</td>
<td>0.000113625</td>
</tr>
<tr>
<td>100</td>
<td>0.00089170</td>
<td>0.000093612</td>
</tr>
<tr>
<td>125</td>
<td>0.00084682</td>
<td>0.000079186</td>
</tr>
<tr>
<td>150</td>
<td>0.00081364</td>
<td>0.000072564</td>
</tr>
</tbody>
</table>
Table 3

Performance of the selection rule

when \( s = 50 \).

<table>
<thead>
<tr>
<th>( n )</th>
<th>( D(n, s) )</th>
<th>( SE(D(n, s)) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.05812975</td>
<td>0.010374924</td>
</tr>
<tr>
<td>10</td>
<td>0.01450924</td>
<td>0.007316987</td>
</tr>
<tr>
<td>15</td>
<td>0.00810957</td>
<td>0.003264550</td>
</tr>
<tr>
<td>20</td>
<td>0.00496331</td>
<td>0.001684319</td>
</tr>
<tr>
<td>30</td>
<td>0.00156783</td>
<td>0.000873342</td>
</tr>
<tr>
<td>40</td>
<td>0.00084367</td>
<td>0.000439320</td>
</tr>
<tr>
<td>50</td>
<td>0.00058916</td>
<td>0.000231186</td>
</tr>
<tr>
<td>60</td>
<td>0.00033976</td>
<td>0.000124923</td>
</tr>
<tr>
<td>70</td>
<td>0.00026253</td>
<td>0.000095978</td>
</tr>
<tr>
<td>80</td>
<td>0.00023609</td>
<td>0.000075646</td>
</tr>
<tr>
<td>90</td>
<td>0.00019546</td>
<td>0.000057012</td>
</tr>
<tr>
<td>100</td>
<td>0.00017925</td>
<td>0.000043645</td>
</tr>
<tr>
<td>125</td>
<td>0.00013687</td>
<td>0.000037202</td>
</tr>
<tr>
<td>150</td>
<td>0.00012795</td>
<td>0.000030269</td>
</tr>
</tbody>
</table>
Figure 1: Graph for Table 1
Figure 2: Graph for Table 2
Figure 3: Graph for Table 3