EMPIRICAL BAYES TWO-STAGE PROCEDURES FOR SELECTING THE BEST NORMAL POPULATION COMPARED WITH A CONTROL

by

Shanti S. Gupta
Purdue University

TaChen Liang
Wayne State University

and

Re-Bin Rau
National Central University

Technical Report #93-21C

Department of Statistics
Purdue University

April 1993
Revised August 1993
Revised November 1993
Revised April 1994

EMPIRICAL BAYES TWO-STAGE PROCEDURES FOR SELECTING THE BEST NORMAL POPULATION COMPARED WITH A CONTROL [1]

Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, IN 47907–1399

TaChen Liang
Department of Mathematics
Wayne State University
Detroit, MI 48202

and

Re-Bin Rau
Department of Mathematics
National Central University
Chung-Li, Taiwan 32054
REPUBLIC OF CHINA

Abstract

The problem of selecting the population associated with the largest mean from among $k(\geq 2)$ independent normal populations is investigated. The population to be selected must be as good as or better than a control. It is assumed that past observations are available when the current selection is made. Accordingly, the empirical Bayes approach is employed. Combining useful information from the past data, an empirical Bayes two-stage selection procedure is developed. It is proved that the empirical Bayes two-stage selection procedure is asymptotically optimal, having a rate of convergence of order $O(\frac{(\ln n)^2}{n})$, where n is the number of past observations at hand. The result of a simulation study, which indicates that the possible obtainable rate of convergence is of order $O(n^{-1})$, is described.

AMS 1991 Subject Classification: Primary 62F07; secondary 62C12, 62C10
Keywords and Phrases: Asymptotic optimality; best population; Bayes rule; empirical Bayes; rate of convergence; two-stage selection procedure.

[1]This research was supported in part by NSF Grant DMS–8923071 at Purdue University.

1 Introduction

Consider k independent normal populations $\pi_1, \ldots, \pi_k$ with unknown means $\theta_1, \cdots, \theta_k$. Let $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ denote the ordered θ_i's. A population π_i with $\theta_i = \theta_{[k]}$ is called the best population. The problem of selecting the best population was studied in the pioneering works of Bechhofer and Gupta, by using the indifference zone approach and the subset selection approach, respectively. Gupta and Panchapakesan (1979,1985) provide a comprehensive survey of the development in this research area.

In a practical situation, one may not only be interested in the selection of the best population, but also require the selected population to be good enough. For example, in medical studies, the performance of any proposed new treatment must be better than a standard treatment before it can be accepted by medical practitioners. In the literature, Bechhofer and Turnbull (1978), Dunnett (1984) and Wilcox (1984) investigated procedures for selecting the best normal population compared with a control, respectively. Using the subset selection approach, Gupta and Sobel (1958) and Lehmann (1961) have made some contributions to this problem.

In this paper, we employ the empirical Bayes approach to select the best normal population provided it is at least as good as a specified standard. The empirical Bayes methodology was introduced by Robbins (1956, 1964). The empirical Bayes approach has been used in selection problems by several authors. Recently, Gupta and Liang (1989), and Gupta and Leu (1991) have investigated empirical Bayes procedures for several selection problems. Many such empirical Bayes selection procedures have been shown to be asymptotically optimal in the sense that the Bayes risk of the empirical Bayes selection procedure converges to the minimum Bayes risk.

This paper deals with a two-stage selection procedure for selecting the best normal population compared with a specified standard using the parametric empirical Bayes approach. Using a parametric empirical Bayes approach, Gupta, Liang and Rau (1994) studied a two-stage selection procedure for selecting the best Binomial population compared with a control. Also, this paper is an extension of a one-stage procedure for selecting the best normal population compared with a control studied by Gupta, Liang and Rau (1994). The formulation of the selection problem is described in Section 2. A Bayes two-stage selection procedure is derived in Section 3. We then construct an empirical Bayes two-stage selection procedure in Section 4. The asymptotic optimality of the empirical Bayes two-stage selection procedure is investigated in Section 5. It is proved that the empirical Bayes two-stage selection procedure has a rate of convergence of order $O(\frac{(\ln n)^2}{n})$, where n is the number of past observations at hand. The result of a simulation study which indicates that the possible obtainable rate of convergence is of order $O(n^{-1})$, is given in Section 6.

2 Formulation of the Selection Problem

Consider k independent normal populations, $\pi_1, \ldots, \pi_k$, with unknown means $\theta_1, \ldots, \theta_k$, respectively. Let $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ denote the ordered values of the parameters $\theta_1, \ldots, \theta_k$. It is assumed that there is no prior information about the true pairing between the ordered and the unordered parameters. Any population associated with $\theta_{[k]}$ is defined as the best population. For a given standard θ_0, π_i is said to be good if $\theta_i \geq \theta_0$, and bad otherwise. Our selection goal is to select a population which should be the best among the k competitors and good compared with the standard θ_0. If there is no such population, we select none.

A two-stage selection procedure is described as follows. First, we have M independent observations taken from each of the k normal populations. For each $i = 1, \ldots, k$, let $X_{i1}, \cdots, X_{iM}$ be a sample of size M from π_i. Based on $\underset{\sim}{X} = (X_1, \ldots, X_k)$, where $X_i = \bar{X}_i = \frac{1}{M}\sum_{j=1}^{M} X_{ij}$, one decides whether the selection should be made immediately or not. If one decides to make the selection immediately, then based on $\underset{\sim}{X}$ one may select a population from among the k populations or one may select none in which case the k populations are excluded as bad populations. If one decides not to make the selection immediately, then one (potential) population is chosen, say π_i, and a further sample of size m, $Y_{i1}, Y_{i2}, \ldots, Y_{im}$, is taken from π_i. Let $Y_i = \bar{Y}_i = \frac{1}{m}\sum_{j=1}^{m} Y_{ij}$. Then, based on $\underset{\sim}{X}$ and Y_i, one may decide to either select π_i as the best population and consider π_i to be good, or select none and exclude all the k populations as bad.

Let $\Omega = \{\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k) | \theta_i \in R, i = 1, \ldots, k\}$ be the parameter space. Let $\underset{\sim}{a} = (a_0, a_1, \ldots, a_k)$ be an action, where $a_i = 0, 1; i = 0, 1, \ldots, k$ and $\sum_{i=0}^{k} a_i = 1$. When $a_i = 1$ for some $i = 1, \ldots, k$, it means that π_i is selected as the best and considered to be good. When $a_0 = 1$, it means that all the k populations are excluded as bad populations. Also, let t denote a function associated with the termination action. When $t = 1$, it means that the selection is made immediately after $\underset{\sim}{X}$ is observed. When $t = 0$, it means that additional m observations from some of the k populations are needed in order to make the selection. When $t = 0$, let $\underset{\sim}{\Delta} = (\Delta_1, \ldots, \Delta_k)$ be the identity action, where $\Delta_i = 0, 1, i = 1, \ldots, k$, and $\sum_{i=1}^{k} \Delta_i = 1$. When $\Delta_i = 1$, it means that the additional m observations are taken from π_i. For the parameter $\underset{\sim}{\theta}$ and action $(\underset{\sim}{a}, t, \underset{\sim}{\Delta})$, the loss function $L(\underset{\sim}{\theta}, (\underset{\sim}{a}, t, \underset{\sim}{\Delta}))$ is defined to be:

$$\begin{aligned} L(\underset{\sim}{\theta}, (\underset{\sim}{a}, t, \underset{\sim}{\Delta})) &= \max(\theta_{[k]}, \theta_0) - t\sum_{i=0}^{k} a_i\theta_i + Mkc_1 \\ &\quad + (1-t)\left\{-\sum_{i=1}^{k} \Delta_i[a_i\theta_i + (1-a_i)\theta_0] + mc_2\right\}, \end{aligned} \tag{2.1}$$

where $c_1 > 0$ is the cost for each observation taken at the first stage, and $c_2 > 0$ is the cost for each observation taken at the second stage. Note that conditional on the parameter $\theta_i, X_i \sim N(\theta_i, \frac{\sigma_i^2}{M}), Y_i \sim N(\theta_i, \frac{\sigma_i^2}{m})$ and X_i and Y_i are conditionally independent. We let $f_i(x|\theta_i)$ denote the conditional probability function of X_i.

It is assumed that for each $i = 1, \ldots, k, \theta_i$ is a realization of a random variable Θ_i which has a $N(\mu_i, \tau_i^2)$ prior distribution with probability density function $h_i(\theta)$, where both μ_i and τ_i^2 are unknown. The random variables $\Theta_1, \ldots, \Theta_k$ are assumed to be mutually independent.

Let $\mathcal{X}$ be the sample space generated by $\underset{\sim}{X}$ and let $\mathcal{Y}$ be the sample space generated by $\underset{\sim}{Y} = (Y_1, \ldots, Y_k)$. A two-stage selection procedure, in general, consists of the following rules:

(a) Stopping rule τ: For each $\underset{\sim}{x} \in \mathcal{X}, \tau(\underset{\sim}{x})$ is the probability of terminating the sampling after observing $\underset{\sim}{x}$ and making a selection immediately based on $\underset{\sim}{x}$.

(b) Identity rule $\underset{\sim}{\delta} = (\delta_1, \ldots, \delta_k)$: For each $\underset{\sim}{x} \in \mathcal{X}, \delta_i(\underset{\sim}{x})$ is the probability of taking the additional m observations from π_i when the decision of going to the second-stage is made. Note that $\underset{\sim}{\delta}$ should satisfy that $\sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) = 1$ for all $\underset{\sim}{x} \in \mathcal{X}$.

(c) First-stage selection rule $\underset{\sim}{d}_1 = (d_{10}, d_{11}, \ldots, d_{1k})$: For each $\underset{\sim}{x} \in \mathcal{X}$, $d_{1i}(\underset{\sim}{x}), i = 1, \ldots, k$, is the probability of selecting π_i as the best and good, and $d_{10}(\underset{\sim}{x})$ is the probability of excluding all the k populations as bad and selecting none. Also, $\sum_{i=0}^{k} d_{1i}(\underset{\sim}{x}) = 1$ for all $\underset{\sim}{x} \in \mathcal{X}$.

(d) Second-stage selection rule $\underset{\sim}{d}_2 = (d_{20}, d_{21}, \ldots, d_{2k})$: For each $\underset{\sim}{x} \in \mathcal{X}, \underset{\sim}{y} \in \mathcal{Y}$, when the decision of going to the second-stage sampling from π_i is made, $d_{2i}(\underset{\sim}{x}, \underset{\sim}{y})$ is the probability of selecting π_i as the best and good, $i = 1, \ldots, k$. It should be noted that $d_{2i}(\underset{\sim}{x}, \underset{\sim}{y})$ depends on $\underset{\sim}{y}$ only through y_i since there are no observations from other populations $\pi_j, j \neq i$. Also, $d_{20}(\underset{\sim}{x}, \underset{\sim}{y}) = \sum_{i=1}^{k} \delta_i(\underset{\sim}{x})[1 - d_{2i}(\underset{\sim}{x}, \underset{\sim}{y})]$ is the probability of selecting none based on $\underset{\sim}{x}$ and the second-stage observation y_i for some $i = 1, \ldots, k$. For notational convenience, in the sequel, we may use either $d_{2i}(\underset{\sim}{x}, \underset{\sim}{y})$ or $d_{2i}(\underset{\sim}{x}, y_i)$.

We denote the Bayes risk of the two-stage selection procedure $(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2)$ by $R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2)$. Then a straightforward computation yields the following:

$$
\begin{aligned}
& R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2) \\
= \;& C - \int_{\mathcal{X}} \tau(\underset{\sim}{x}) \left[\sum_{i=0}^{k} d_{1i}(\underset{\sim}{x}) \varphi_i(x_i) \right] f(\underset{\sim}{x}) d\underset{\sim}{x} && (2.2) \\
+ \;& \int_{\mathcal{X}} [1 - \tau(\underset{\sim}{x})] \left\{ mc_2 - \theta_0 + \sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) \left[\int_R d_{2i}(\underset{\sim}{x}, y_i)[\theta_0 - \psi_i(x_i, y_i)] f_i(y_i|x_i) dy_i \right] \right\} f(\underset{\sim}{x}) d\underset{\sim}{x} \\
= \;& C + \int_{\mathcal{X}} \tau(\underset{\sim}{x}) \left\{ \begin{array}{l} \sum_{i=0}^{k} d_{1i}(\underset{\sim}{x})[\theta_0 - \varphi_i(x_i)] - mc_2 \\ - \sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) \left[\int_R d_{2i}(\underset{\sim}{x}, y_i)[\theta_0 - \psi_i(x_i, y_i)] f_i(y_i|x_i) dy_i \right] \end{array} \right\} f(\underset{\sim}{x}) d\underset{\sim}{x} \\
+ \;& \int_{\mathcal{X}} \left\{ mc_2 - \theta_0 + \sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) \left[\int_R d_{2i}(\underset{\sim}{x}, y_i)[\theta_0 - \psi_i(x_i, y_i)] f_i(y_i|x_i) dy_i \right] \right\} f(\underset{\sim}{x}) d\underset{\sim}{x} \\
& && (2.3)
\end{aligned}
$$

where

$$
\left\{ \begin{array}{l}
\varphi_i(x_i) = E(\Theta_i|x_i) = (x_i \tau_i^2 + \frac{\sigma_i^2}{M} \mu_i)(\tau_i^2 + \frac{\sigma_i^2}{M})^{-1}, i = 1, \ldots, k; \;\; \varphi_0(X_0) \equiv \theta_0, \\
\psi_i(x_i, y_i) = E(\Theta_i|x_i, y_i) = (\frac{Mx_i + my_i}{M+m} \tau_i^2 + \frac{\sigma_i^2}{M+m} \mu_i)(\tau_i^2 + \frac{\sigma_i^2}{M+m})^{-1}, i \neq 0, \\
f(\underset{\sim}{x}) = \prod_{i=1}^{k} f_i(x_i), \;\; f_i(x_i) = \int_R f_i(x_i|\theta_i) h_i(\theta_i) d\theta_i : \;\; \text{the marginal density of } X_i, \\
f_i(y_i|x_i) : \;\; \text{the marginal conditional density of } Y_i \text{ given } X_i = x_i \;, \\
Y_i|X_i = x_i \sim N(\varphi_i(x_i), \frac{\sigma_i^2}{m} \frac{(M+m)\tau_i^2 + \sigma_i^2}{M\tau_i^2 + \sigma_i^2}), C = \int_{\Omega} \max(\theta_{[k]}, \theta_0) dH(\underset{\sim}{\theta}) + Mkc_1, \\
H(\underset{\sim}{\theta}) : \;\; \text{the joint distribution of } \underset{\sim}{\Theta} = (\Theta_1, \ldots, \Theta_k).
\end{array} \right. \qquad (2.4)
$$

3 Derivation of a Bayes Two-Stage Selection Procedure

In the sequel, we assume that $\sigma_1^2 = \ldots = \sigma_k^2 = \sigma_0^2$ and $\tau_1^2 = \ldots = \tau_k^2 = \tau_0^2$, where σ_0^2 and τ_0^2 are unknown. Also, μ_i is unknown for $i = 1, \ldots, k$. In order to develop an empirical Bayes two-stage selection procedure, as a first step, we derive a Bayes two-stage selection procedure for the selection problem under consideration.

A First-Stage Selection Rule: For each $\underset{\sim}{x} \in \mathcal{X}$, let

$$\begin{cases} I(\underset{\sim}{x}) = \{i \mid \varphi_i(x_i) = \max_{0 \le j \le k} \varphi_j(x_j), i = 0, \ldots, k\}, \\ i^* \equiv i^*(\underset{\sim}{x}) = \begin{cases} 0 & \text{if } I(\underset{\sim}{x}) = \{0\}; \\ \min\{i \mid i \in I(\underset{\sim}{x}),\ i \neq 0\} & \text{otherwise.} \end{cases} \end{cases} \tag{3.1}$$

Then a first-stage selection rule $\underset{\sim}{d}_1^B = (d_{10}^B, \ldots, d_{1k}^B)$ is given as follows:

$$d_{1i^*}^B(\underset{\sim}{x}) = 1, \quad \text{and} \quad d_{1j}^B(\underset{\sim}{x}) = 0 \text{ for } j \neq i^*. \tag{3.2}$$

A Second-Stage Selection Rule: We define a second-stage selection rule $\underset{\sim}{d}_2^B = (d_{20}^B, \ldots, d_{2k}^B)$ as follows: For each $\underset{\sim}{x} \in \mathcal{X}, \underset{\sim}{y} \in \mathcal{Y}$, and $i = 1, \ldots, k$, define

$$d_{2i}^B(\underset{\sim}{x}, \underset{\sim}{y}) = \begin{cases} 1 & \text{if } \psi_i(x_i, y_i) \geq \theta_0; \\ 0 & \text{otherwise;} \end{cases} \quad \text{and } d_{20}^B(\underset{\sim}{x}, \underset{\sim}{y}) = \sum_{i=1}^{k} \delta_i^B(\underset{\sim}{x})[1 - d_{2i}^B(\underset{\sim}{x}, \underset{\sim}{y})], \tag{3.3}$$

where $\underset{\sim}{\delta}^B = (\delta_1^B, \ldots, \delta_k^B)$ is the identity rule defined below. Note that $d_{2i}^B(\underset{\sim}{x}, \underset{\sim}{y})$ depends on $(\underset{\sim}{x}, \underset{\sim}{y})$ only through (x_i, y_i), $i \neq 0$, see (3.3).

An Identity Rule: For each $i = 1, \ldots, k$, and $\underset{\sim}{x} \in \mathcal{X}$, define

$$T_i(\underset{\sim}{x}) \equiv \int_R d_{2i}^B(\underset{\sim}{x}, y_i)[\theta_0 - \psi_i(x_i, y_i)] f_i(y_i \mid x_i) dy_i. \tag{3.4}$$

By the definition of $d_{2i}^B(\underset{\sim}{x}, y_i)$ (see (3.3)), a straightforward computation yields

$$T_i(\underset{\sim}{x}) = (\theta_0 - \varphi_i(x_i))\Phi[q_0(\varphi_i(x_i) - \theta_0)] - \frac{1}{q_0}\phi[q_0(\varphi_i(x_i) - \theta_0)] \equiv T(\varphi_i(x_i)),$$

where $T(z) = (\theta_0 - z)\Phi[q_0(z - \theta_0)] - \frac{1}{q_0}\phi[q_0(z - \theta_0)]$, $q_0 = \frac{\sqrt{(M+m)M\tau_0^4 + (2M+m)\tau_0^2\sigma_0^2 + \sigma_0^4}}{\tau_0^2\sqrt{m\sigma_0^2}}$ and Φ and ϕ are cdf and pdf of the $N(0,1)$ distribution, respectively. Since $\frac{dT(z)}{dz} = -\Phi[q_0(z - \theta_0)] < 0$, $T(\varphi_i(x_i))$ is a decreasing function of $\varphi_i(x_i)$.

For each $\underset{\sim}{x} \in \mathcal{X}$, let

$$\begin{cases} J(\underset{\sim}{x}) = \{j \mid \varphi_j(x_j) = \max_{1 \le i \le k} \varphi_i(x_i), j = 1, \ldots, k\}, \\ j^* \equiv j^*(\underset{\sim}{x}) = \min\{j \mid j \in J(\underset{\sim}{x})\}. \end{cases} \tag{3.5}$$

We then define an identity rule $\underset{\sim}{\delta}^B = (\delta_1^B, \ldots, \delta_k^B)$ as follows:

$$\delta_j^B(\underset{\sim}{x}) = 1(0) \text{ if } j = (\neq) j^*. \tag{3.6}$$

A Stopping Rule: Consider the function $L : R^2 \to R$,

$$L(q,z) = z\Phi(qz) + \frac{1}{q}\phi(qz) - mc_2. \tag{3.7}$$

For any fixed $q > 0$, let $L_q(z) = L(q,z)$. Since $\frac{dL_q(z)}{dz} = \Phi(qz) > 0, \lim_{z\to-\infty} L_q(z) = -mc_2$ and $\lim_{z\to+\infty} L_q(z) = +\infty$, $L_{q_0}(z)$ is strictly increasing and has a unique zero, say z_0.

For each $\underset{\sim}{x} \in \mathcal{X}$, let

$$\begin{aligned} Q(\underset{\sim}{x}) &\equiv \sum_{i=0}^{k} d_{1i}^B(\underset{\sim}{x})[\theta_0 - \varphi_i(x_i)] - mc_2 - \sum_{j=1}^{k} \delta_j^B(\underset{\sim}{x}) T_j(\underset{\sim}{x}) \\ &= \sum_{j=1}^{k} \mathrm{I}_{\{i^*=0, j^*=j\}} L(q_0, \varphi_j(x_j) - \theta_0) + \sum_{i=1}^{k} \mathrm{I}_{\{i^*=j^*=i\}} L(q_0, \theta_0 - \varphi_i(x_i)), \end{aligned} \tag{3.8}$$

where the second equality is obtained by noting that if $i^* \neq 0$, then $i^* = j^*$. We then define a stopping rule τ^B as follows:

$$\tau^B(\underset{\sim}{x}) = 1(0) \text{ if } Q(\underset{\sim}{x}) \leq (>) 0. \tag{3.9}$$

Remark 3.1: Observe that

$$Q(\underset{\sim}{x}) \leq 0 \Leftrightarrow \begin{cases} \text{if } i^* = 0, j^* = j \text{ then } \varphi_j(x_j) \leq \theta_0 + z_0; \\ \text{if } i^* = j^* = i \text{ then } \varphi_i(x_i) \geq \theta_0 - z_0. \end{cases}$$

Hence, if $z_0 \geq 0$ then $\tau^B(\underset{\sim}{x}) = 1$ for all $\underset{\sim}{x}$. Note that z_0 is implicitly defined as a function of $t = mc_2$ by $z_0\Phi(q_0 z_0) + \frac{1}{q_0}\phi(q_0 z_0) - t = 0$. Taking derivatives on both side of the preceding equation with respect to t, we then obtain that $\frac{dz_0}{dt} = \frac{1}{\Phi(q_0 z_0)} > 0$. Therefore $z_0(t)$ is strictly increasing in $t = mc_2$. In the following, we assume that c_2 is small enough such that $z_0 < 0$.

Then, we have the following result:

Theorem 3.1. The two-stage selection procedure $(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$ defined through $(3.1) - (3.3), (3.5), (3.6), (3.8)$ and (3.9) is a Bayes two-stage selection procedure.

Proof: Let $(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2)$ be any two-stage selection procedure. We only need to prove that $R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \geq 0$. Now, $R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) = I + II + III$, where $I = R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2) - R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$, $II = R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) - R(\tau, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$, and $III = R(\tau, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$.

From (2.2),

$$\begin{aligned} I = &\int_{\mathcal{X}} \tau(\underset{\sim}{x}) \left\{ \sum_{i=0}^{k} [d_{1i}^B(\underset{\sim}{x}) - d_{1i}(\underset{\sim}{x})] \varphi_i(x_i) \right\} f(\underset{\sim}{x}) d\underset{\sim}{x} \\ + &\int_{\mathcal{X}} [1 - \tau(\underset{\sim}{x})] \left\{ \sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) \left[\int_R [d_{2i}(\underset{\sim}{x}, y_i) - d_{2i}^B(\underset{\sim}{x}, y_i)][\theta_0 - \psi_i(x_i, y_i)] f_i(y_i | x_i) dy_i \right] \right\} f(\underset{\sim}{x}) d\underset{\sim}{x}. \end{aligned} \tag{3.10}$$

By the definition of $\underset{\sim}{d}_1^B$ and $\underset{\sim}{d}_1$, $\sum_{i=0}^{k}[d_{1i}^B(\underset{\sim}{x}) - d_{1i}(\underset{\sim}{x})]\varphi_i(x_i) \geq 0$. Also, by the definition of $\underset{\sim}{d}_2^B$, for each $i = 1, \ldots, k$, $[d_{2i}(\underset{\sim}{x}, y_i) - d_{2i}^B(\underset{\sim}{x}, y_i)][\theta_0 - \psi_i(x_i, y_i)] \geq 0$. Hence, $I \geq 0$ since all other terms in (3.10) are nonnegative.

From (2.2) again,

$$II = \int_{\mathcal{X}} [1 - \tau(\underset{\sim}{x})] \left\{ \sum_{i=1}^{k} [\delta_i(\underset{\sim}{x}) - \delta_i^B(\underset{\sim}{x})] T_i(\underset{\sim}{x}) \right\} f(\underset{\sim}{x}) d\underset{\sim}{x} \geq 0, \tag{3.11}$$

since, by the decreasing property of T and the definition of $\underset{\sim}{\delta}^B$,

$$\sum_{i=1}^{k} [\delta_i(\underset{\sim}{x}) - \delta_i^B(\underset{\sim}{x})] T_i(\underset{\sim}{x}) = \sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) T(\varphi_i(x_i)) - T(\max_{1 \leq i \leq k} \varphi_i(x_i)) \geq 0.$$

Now, from (2.3)

$$III = \int_{\mathcal{X}} [\tau(\underset{\sim}{x}) - \tau^B(\underset{\sim}{x})] Q(\underset{\sim}{x}) f(\underset{\sim}{x}) d\underset{\sim}{x} \geq 0 \tag{3.12}$$

which holds by the definition of τ^B.

The proof of the above theorem is then completed by combining (3.10) – (3.12). □

4 The Empirical Bayes Two-Stage Selection Procedure

Since the parameters σ_0^2, τ_0^2 and μ_i, $i = 1, \ldots, k$, are unknown, it is not possible to apply the Bayes two-stage selection procedure $(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$ for the selection problem at hand. In the empirical Bayes framework, it is assumed that certain past data are available when the present selection is made. Let $X_{ijl}, j = 1, \ldots, M$, denote a sample of size M from π_i at time $l, l = 1, \ldots, n$. It is assumed that conditional on $(\theta_{il}, \sigma_0^2)$, $X_{ijl}, j = 1, \ldots, M$, follow a $N(\theta_{il}, \sigma_0^2)$ distribution and θ_{il} is a realization of a random variable Θ_{il} which has a $N(\mu_i, \tau_0^2)$ prior distribution. It is also assumed that Θ_{il}, $i = 1, \ldots, k$, $l = 1, 2, \ldots$, are mutually independent. For ease of notation, we denote the current random observations $X_{ij\,n+1}$ by X_{ij}, $j = 1, \ldots, M$, $i = 1, \ldots, k$. For each i, let

$$\begin{cases} X_{i.l} & = \frac{1}{M}\sum_{j=1}^{M} X_{ijl}, X_i(n) = \frac{1}{n}\sum_{l=1}^{n} X_{i.l}, \\ S_i^2(n) & = \frac{1}{n-1}\sum_{l=1}^{n}(X_{i.l} - X_i(n))^2, S^2(n) = \frac{1}{k}\sum_{i=1}^{k} S_i^2(n). \end{cases} \tag{4.1}$$

Then, $X_{i.1}, X_{i.2}, \ldots, X_{i.n}$ are marginally independent with a $N(\mu_i, v_0^2)$ distribution, where $v_0^2 = \tau_0^2 + \frac{\sigma_0^2}{M}$. Hence, $X_i(n)$ has a $N(\mu_i, \frac{v_0^2}{n})$ distribution and $\frac{k(n-1)}{v_0^2} S^2(n)$ has a $\chi^2(k(n-1))$ distribution.

In order to estimate σ_0^2, it is assumed that $M \geq 2$. For each $i = 1, \cdots, k$, let

$$\begin{cases} W_{i.l}^2 & = \frac{1}{M-1}\sum_{j=1}^{M}(X_{ijl} - X_{i.l})^2, \\ W_i^2(n) & = \frac{1}{n}\sum_{l=1}^{n} W_{i.l}^2, \quad W^2(n) = \frac{1}{k}\sum_{i=1}^{k} W_i^2(n). \end{cases} \tag{4.2}$$

Then, $\frac{kn(M-1)}{\sigma_0^2} W^2(n)$ has a $\chi^2(kn(M-1))$ distribution. By the strong law of large numbers, $X_i(n) \longrightarrow \mu_i$ a.s. , $W^2(n) \longrightarrow \sigma_0^2$ a.s., $S^2(n) \longrightarrow v_0^2$ a.s., $S^2(n) - \frac{W^2(n)}{M} \longrightarrow$

τ_0^2 a.s. Also $E(X_i(n)) = \mu_i$, $E(S^2(n)) = v_0^2$, $E(W^2(n)) = \sigma_0^2$, $E(S^2(n) - \frac{W^2(n)}{M}) = v_0^2 - \frac{\sigma_0^2}{M} = \tau_0^2$. Since it is possible that $S^2(n) - \frac{W^2(n)}{M} < 0$, we define $\mu_{in}, \sigma_{0n}^2, v_{0n}^2, \tau_{0n}^2$ and q_{0n} as estimators of $\mu_i, \sigma_0^2, v_0^2, \tau_0^2$ and q_0, respectively, by the following:

$$\begin{cases} \mu_{in} &= X_i(n), \sigma_{0n}^2 = W^2(n), v_{0n}^2 = S^2(n), \tau_{0n}^2 = \max(v_{0n}^2 - \frac{\sigma_{0n}^2}{M},\ 0), \\ q_{0n} &= \frac{\sqrt{(M+m)M\tau_{0n}^4 + (2M+m)\tau_{0n}^2\sigma_{0n}^2 + \sigma_{0n}^4}}{\tau_{0n}^2\sqrt{m\sigma_{0n}^2}} \quad \text{if } \tau_{0n}^2 > 0. \end{cases} \tag{4.3}$$

Without loss of generality, we assume $\sigma_{0n}^2 > 0$, since $\sigma_{0n}^2 = 0$ with probability 0.

For $i = 1, 2, \cdots, k$, we define

$$\begin{cases} \varphi_{in}(x_i) &= \frac{x_i\tau_{0n}^2 + \frac{\sigma_{0n}^2}{M}\mu_{in}}{v_{0n}^2}, \varphi_{0n}(x_0) \equiv \theta_0, \\ \psi_{in}(x_i, y_i) &= \frac{\frac{Mx_i + my_i}{M+m}\tau_{0n}^2 + \frac{\sigma_{0n}^2}{M+m}\mu_{in}}{\tau_{0n}^2 + \frac{\sigma_{0n}^2}{M+m}}. \end{cases} \tag{4.4}$$

and use $\varphi_{in}(x_i)$ and $\psi_{in}(x_i, y_i)$ as estimators of $\varphi_i(x_i)$ and $\psi_i(x_i, y_i)$, respectively.

Recall that, for $i = 1, \ldots, k$, $f_i(y_i|x_i)$ is the density of a $N(\varphi_i(x_i), \rho_0^2)$ distribution, where $\rho_0^2 = \frac{\sigma_0^2}{m}\frac{(M+m)\tau_0^2 + \sigma_0^2}{M\tau_0^2 + \sigma_0^2}$. Hence, we let $f_{in}(y_i|x_i)$ be the density of a $N(\varphi_{in}(x_i), \rho_{0n}^2)$ distribution and use $f_{in}(y_i|x_i)$ as an estimator of $f_i(y_i|x_i)$, where $\rho_{0n}^2 = \frac{\sigma_{0n}^2}{m}\frac{(M+m)\tau_{0n}^2 + \sigma_{0n}^2}{M\tau_{0n}^2 + \sigma_{0n}^2}$.

Now, an empirical Bayes two-stage selection procedure $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$ is proposed as follows:

Empirical Bayes First-Stage Selection Rule $\underset{\sim}{d}_1^{*n} = (d_{10}^{*n}, \ldots, d_{1k}^{*n})$

For each $\underset{\sim}{x} \in \mathcal{X}$, let

$$\begin{cases} I_n(\underset{\sim}{x}) = \{i | \varphi_{in}(x_i) = \max_{0 \le j \le k} \varphi_{jn}(x_j), i = 0, \ldots, k\}, \\ i_n^* \equiv i_n^*(\underset{\sim}{x}) = \begin{cases} 0 & \text{if } I_n(\underset{\sim}{x}) = \{0\}, \\ \min\{i | i \in I_n(\underset{\sim}{x}),\ i \neq 0\} & \text{otherwise.} \end{cases} \end{cases} \tag{4.5}$$

We then obtain an empirical Bayes first-stage selection rule $\underset{\sim}{d}_1^{*n} = (d_{10}^{*n}, \ldots, d_{1k}^{*n})$ as follows:

$$d_{ij}^{*n}(\underset{\sim}{x}) = 1(0) \text{ if } j = (\neq) i_n^*. \tag{4.6}$$

Empirical Bayes Second-Stage Selection Rule $\underset{\sim}{d}_2^{*n} = (d_{20}^{*n}, \ldots, d_{2k}^{*n})$

We define $\underset{\sim}{d}_2^{*n}$ as follows: For each $\underset{\sim}{x} \in \mathcal{X}$ and $\underset{\sim}{y} \in \mathcal{Y}$, $i = 1, \ldots, k$,

$$d_{2i}^{*n}(\underset{\sim}{x}, \underset{\sim}{y}) = \begin{cases} 1 & \text{if } \psi_{in}(x_i, y_i) \geq \theta_0, \\ 0 & \text{otherwise;} \end{cases}, \text{ and } d_{20}^{*n}(\underset{\sim}{x}, \underset{\sim}{y}) = \sum_{i=1}^{k} \delta_i^{*n}(\underset{\sim}{x})[1 - d_{2i}^{*n}(\underset{\sim}{x}, \underset{\sim}{y})] \tag{4.7}$$

where $\underset{\sim}{\delta}^{*n} = (\delta_1^{*n}, \ldots, \delta_k^{*n})$ is the empirical Bayes identity rule defined below.

Empirical Bayes Identity Rule $\underset{\sim}{\delta}^{*n} = (\delta_1^{*n}, \ldots, \delta_k^{*n})$

For each $i = 1, \ldots, k$, and $\underset{\sim}{x} \in \mathcal{X}$, let

$$\begin{aligned} T_{in}(\underset{\sim}{x}) &\equiv \int_R d_{2i}^{*n}(\underset{\sim}{x}, y_i)[\theta_0 - \psi_{in}(x_i, y_i)] f_{in}(y_i|x_i) dy_i \\ &= \begin{cases} T_n(\varphi_{in}(x_i)) & \text{if } \tau_{0n}^2 > 0, \\ (\theta_0 - \varphi_{in}(x_i)) \, \mathrm{I}_{\{\varphi_{in}(x_i) > \theta_0\}} & \text{otherwise,} \end{cases} \end{aligned} \tag{4.8}$$

where $T_n(z) = (\theta_0 - z)\Phi[q_{0n}(z - \theta_0)] - \frac{1}{q_{0n}}\phi[q_{0n}(z - \theta_0)]$ and the second equality in (4.8) is obtained by the definition of d_{2i}^{*n} accordingly. Note that $\frac{dT_n(z)}{dz} = -\Phi[q_{0n}(z - \theta_0)] < 0$. Hence $T_n(z)$ is a decreasing function of z. Let

$$\begin{cases} J_n(\underset{\sim}{x}) = \{j | \varphi_{jn}(x_j) = \max_{1 \le i \le k} \varphi_{in}(x_i), j = 1, \ldots, k\}, \\ j_n^* \equiv j_n^*(\underset{\sim}{x}) = \min\{j | j \in J_n(\underset{\sim}{x})\}. \end{cases} \tag{4.9}$$

Then, the empirical Bayes identity rule $\underset{\sim}{\delta}^{*n} = (\delta_1^{*n}, \ldots, \delta_k^{*n})$ is defined as:

$$\delta_j^{*n}(\underset{\sim}{x}) = 1(0) \text{ if } j = (\neq) j_n^*. \tag{4.10}$$

Note that, by the definition of i^{*n} and j^{*n}, if $i^{*n} \neq 0$ then $i^{*n} = j^{*n}$.

Empirical Bayes Stopping Rule τ^{*n}

For each $\underset{\sim}{x} \in \mathcal{X}$, let

$$\begin{aligned} Q_n(\underset{\sim}{x}) &\equiv \sum_{i=0}^{k} d_{1i}^{*n}(\underset{\sim}{x})[\theta_0 - \varphi_{in}(x_i)] - mc_2 - \sum_{j=1}^{k} \delta_j^{*n}(\underset{\sim}{x}) T_{jn}(\underset{\sim}{x}) \\ &= \begin{cases} \begin{bmatrix} \sum_{j=1}^{k} \mathrm{I}_{\{i_n^*=0, j_n^*=j\}} L(q_{0n}, \varphi_{jn}(x_j) - \theta_0) \\ + \sum_{i=1}^{k} \mathrm{I}_{\{i_n^*=j_n^*=i\}} L(q_{0n}, \theta_0 - \varphi_{in}(x_i)) \end{bmatrix} & \text{if } \tau_{0n}^2 > 0, \\ -mc_2 & \text{otherwise.} \end{cases} \end{aligned} \tag{4.11}$$

We may use $Q_n(\underset{\sim}{x})$ to estimate $Q(\underset{\sim}{x})$ and propose an empirical Bayes stopping rule τ^{*n} accordingly. That is, for each $\underset{\sim}{x} \in \mathcal{X}$, define

$$\tau^{*n}(\underset{\sim}{x}) = 1(0) \text{ if } Q_n(\underset{\sim}{x}) \le (>) 0. \tag{4.12}$$

When $\tau_{0n}^2 > 0$, we let z_{0n} be the zero of $L(q_{0n}, z)$. Then,

$$Q_n(\underset{\sim}{x}) \le 0 \Leftrightarrow \begin{cases} \text{if } i_n^* = 0, j_n^* = j \ \text{ then } \varphi_{jn}(x_j) \le \theta_0 + z_{0n}; \\ \text{if } i_n^* = j_n^* = i \ \text{ then } \varphi_{in}(x_i) \ge \theta_0 - z_{0n}. \end{cases}$$

5 Asymptotic Optimality of $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$

Consider an empirical Bayes two-stage selection procedure $(\tau^n, \underset{\sim}{\delta}^n, \underset{\sim}{d}_1^n, \underset{\sim}{d}_2^n)$. Let $R(\tau^n, \underset{\sim}{\delta}^n, \underset{\sim}{d}_1^n, \underset{\sim}{d}_2^n)$ be the associated conditional Bayes risk (conditioning on the past observation X_{ijl}, $i = 1, \ldots, k$, $j = 1, \ldots, M$ and $l = 1, \ldots, n$) and let $E_n R(\tau^n, \underset{\sim}{\delta}^n, \underset{\sim}{d}_1^n, \underset{\sim}{d}_2^n)$ be the corresponding overall Bayes risk where the expectation E_n is taken with respect to the probability measure P_n generated by (X_{ijl}, $i = 1, \ldots, k$, $j = 1, \ldots, M$ and $l = 1, \ldots, n$). Since $(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$ is a Bayes two-stage selection procedure, $R(\tau^n, \underset{\sim}{\delta}^n, \underset{\sim}{d}_1^n, \underset{\sim}{d}_2^n) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \geq 0$ and $E_n R(\tau^n, \underset{\sim}{\delta}^n, \underset{\sim}{d}_1^n, \underset{\sim}{d}_2^n) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \geq 0$ for all n. The nonnegative regret risk $E_n R(\tau^n, \underset{\sim}{\delta}^n, \underset{\sim}{d}_1^n, \underset{\sim}{d}_2^n) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$ is generally used as a measure of performance of the empirical Bayes two-stage selection procedure $(\tau^n, \underset{\sim}{\delta}^n, \underset{\sim}{d}_1^n, \underset{\sim}{d}_2^n)$.

In the following, we evaluate the asymptotic optimality of the proposed empirical Bayes two-stage selection procedure $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$. First, we have:

$$R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) = I_n + II_n + III_n, \tag{5.1}$$

where

$$\begin{aligned} 0 \leq I_n &\equiv R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \\ &= \int_{\mathcal{X}} \tau^{*n}(\underset{\sim}{x})[d_{1i^*}^B(\underset{\sim}{x})\varphi_{i^*}(x_{i^*}) - d_{1i_n^*}^{*n}(\underset{\sim}{x})\varphi_{i_n^*}(x_{i_n^*})]f(\underset{\sim}{x})d\underset{\sim}{x} \\ &\quad + \int_{\mathcal{X}} [1 - \tau^{*n}(\underset{\sim}{x})]\{\sum_{i=1}^{k} \delta_i^{*n}(\underset{\sim}{x})[\int_R [d_{2i}^{*n}(\underset{\sim}{x}, y_i) - d_{2i}^B(\underset{\sim}{x}, y_i)][\theta_0 - \psi_i(x_i, y_i)] \\ &\qquad \times f_i(y_i|x_i)dy_i]\}f(\underset{\sim}{x})d\underset{\sim}{x} \qquad (5.2) \\ &= I_{n,1} + I_{n,2}; \end{aligned}$$

$$\begin{aligned} 0 \leq II_n &\equiv R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) - R(\tau^{*n}, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \\ &= \int_{\mathcal{X}} [1 - \tau^{*n}(\underset{\sim}{x})][\delta_{j_n^*}^{*n}(\underset{\sim}{x})T_{j_n^*}(\underset{\sim}{x}) - \delta_{j^*}^B(\underset{\sim}{x})T_{j^*}(\underset{\sim}{x})]f(\underset{\sim}{x})d\underset{\sim}{x}; \end{aligned} \tag{5.3}$$

and

$$\begin{aligned} 0 \leq III_n &\equiv R(\tau^{*n}, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \\ &= \int_{\mathcal{X}} [\tau^{*n}(\underset{\sim}{x}) - \tau^B(\underset{\sim}{x})]Q(\underset{\sim}{x})f(\underset{\sim}{x})d\underset{\sim}{x}. \end{aligned} \tag{5.4}$$

To investigate the convergence rate of $E_n[I_{n,1}]$, we state some facts:

Fact 1

1.1 If $i^* = 0,\ \varphi_l(x_l) < \theta_0$ for all $l = 1, \ldots, k$. Then, if $i_n^* = j \neq 0$,

$$\begin{aligned} & P_n\{i^* = 0,\ i_n^* = j\} = P_n\{\varphi_l(x_l) < \theta_0\ \forall\, l \neq 0,\ \varphi_{jn}(x_j) \geq \varphi_{ln}(x_l)\ \forall\, l \neq j\} \\ \leq\ & P_n\{\varphi_{jn}(x_j) - \varphi_j(x_j) \geq \theta_0 - \varphi_j(x_j)\}. \end{aligned}$$

1.2 If $i_n^* = 0,\ \varphi_{ln}(x_l) < \theta_0$ for all $l = 1, \ldots, k$. Then, if $i^* = i \neq 0$,

$$\begin{aligned} & P_n\{i^* = i,\ i_n^* = 0\} = P_n\{\varphi_i(x_i) \geq \varphi_l(x_l)\ \forall\, l \neq i,\ \varphi_{ln}(x_l) < \theta_0\ \forall\, l \neq 0\} \\ \leq\ & P_n\{\varphi_{in}(x_i) - \varphi_i(x_i) < -(\varphi_i(x_i) - \theta_0)\}. \end{aligned}$$

1.3 If $i^* = i \neq 0,\ i_n^* = j \neq 0$ and $i \neq j$, then

$$\begin{aligned} & P_n\{i^* = i,\ i_n^* = j\} = P_n\{\varphi_i(x_i) \geq \varphi_l(x_l)\ \forall\, l \neq i,\ \varphi_{jn}(x_j) \geq \varphi_{ln}(x_l)\ \forall\, l \neq j\} \\ \leq\ & P_n\{|\varphi_{jn}(x_j) - \varphi_j(x_j)| \geq \frac{\varphi_i(x_i) - \varphi_j(x_j)}{2}\} + P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| \geq \frac{\varphi_i(x_i) - \varphi_j(x_j)}{2}\}. \end{aligned}$$

Recall that $\varphi_i(x_i) = (x_i\tau_0^2 + \frac{\sigma_0^2}{M}\mu_i)/(\tau_0^2 + \frac{\sigma_0^2}{M})$ and X_i is marginally $N(\mu_i, v_0^2)$ distributed. Therefore, $\varphi_i(X_i)$ is $N(\mu_i, \frac{\tau_0^4}{v_0^2})$ distributed. Moreover, $\varphi_i(X_i) - \varphi_j(X_j)$ has a $N(\mu_i - \mu_j, \frac{\tau_0^4}{v_0^2} + \frac{\tau_0^4}{v_0^2})$ distribution. For $\varepsilon_n > 0,\ i, j = 1, \ldots, k,\ i \neq j$, let

$$\left\{ \begin{array}{lcl} \mathcal{X}_i & = & \{x_i |\ |\varphi_i(x_i) - \theta_0| \leq \varepsilon_n\}, \\ \mathcal{X}_{ij} & = & \{(x_i, x_j) |\ |\varphi_i(x_i) - \varphi_j(x_j)| \leq \varepsilon_n\}. \end{array} \right. \tag{5.5}$$

From Facts 1.1 − 1.3, (5.2) and (5.5), we get

$$\begin{aligned} & E_n[I_{n,1}] \\ = \ & \sum_{i=1}^{k} E_n \int_{\mathcal{X}} \tau^{*n}(\underset{\sim}{x})\ \mathrm{I}_{\{i^*=i, i_n^*=0\}}[\varphi_i(x_i) - \theta_0] f(\underset{\sim}{x}) d\underset{\sim}{x} \\ & + \sum_{j=1}^{k} E_n \int_{\mathcal{X}} \tau^{*n}(\underset{\sim}{x})\ \mathrm{I}_{\{i^*=0, i_n^*=j\}}[\theta_0 - \varphi_j(x_j)] f(\underset{\sim}{x}) d\underset{\sim}{x} \\ & + \sum_{i=1}^{k}\sum_{j=1}^{k} E_n \int_{\mathcal{X}} \tau^{*n}(\underset{\sim}{x})\ \mathrm{I}_{\{i^*=i, i_n^*=j\}}[\varphi_i(x_i) - \varphi_j(x_j)] f(\underset{\sim}{x}) d\underset{\sim}{x} \\ \leq \ & \sum_{i=1}^{k} \int_{R} P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| \geq |\varphi_i(x_i) - \theta_0|\} |\varphi_i(x_i) - \theta_0| f_i(x_i) dx_i \\ & + \sum_{i=1}^{k}\sum_{j=1}^{k} \int_{R^2} \Bigg[P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| \geq \frac{|\varphi_i(x_i) - \varphi_j(x_j)|}{2}\} \\ & \qquad + P_n\{|\varphi_{jn}(x_j) - \varphi_j(x_j)| \geq \frac{|\varphi_i(x_i) - \varphi_j(x_j)|}{2}\} \Bigg] \end{aligned}$$

$$\times|\varphi_i(x_i) - \varphi_j(x_j)| f_i(x_i) f_j(x_j) dx_i dx_j$$

$$\begin{aligned}
\leq\ & \sum_{i=1}^{k} \int_{\mathcal{X}_i} \varepsilon_n f_i(x_i) dx_i \\
& + \sum_{i=1}^{k} \int_{R} P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| > \varepsilon_n\} |\varphi_i(x_i) - \theta_0| f_i(x_i) dx_i \\
& + \sum_{i=1}^{k} \sum_{j=1, j\neq i}^{k} \int_{\mathcal{X}_{ij}} 2\varepsilon_n f_i(x_i) f_j(x_j) dx_i dx_j \\
& + \sum_{i=1}^{k} \sum_{j=1, j\neq i}^{k} \int_{R^2} \left[P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| > \frac{\varepsilon_n}{2}\} + P_n\{|\varphi_{jn}(x_j) - \varphi_j(x_j)| > \frac{\varepsilon_n}{2}\} \right] \\
& \qquad \times|\varphi_i(x_i) - \varphi_j(x_j)| f_i(x_i) f_j(x_j) dx_i dx_j \\
\leq\ & O(\varepsilon_n^2) \qquad (5.6) \\
& + \sum_{i=1}^{k} \int_{R} P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| > \frac{\varepsilon_n}{2}\} [|\varphi_i(x_i) - \mu_i| + |\mu_i - \theta_0|] f_i(x_i) dx_i, \\
& + \sum_{i=1}^{k} \sum_{j=1, j\neq i}^{k} \int_{R^2} \left[P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| > \frac{\varepsilon_n}{2}\} + P_n\{|\varphi_{jn}(x_j) - \varphi_j(x_j)| > \frac{\varepsilon_n}{2}\} \right] \\
& \qquad \times [|\varphi_i(x_i) - \mu_i| + |\varphi_j(x_j) - \mu_j| + |\mu_i - \mu_j|]\, f_i(x_i) f_j(x_j) dx_i dx_j,
\end{aligned}$$

where $\sum_{i=1}^{k} \int_{\mathcal{X}_i} \varepsilon_n f_i(x_i) dx_i = O(\varepsilon_n^2)$ and $\sum_{i=1}^{k} \sum_{j=1, j\neq i}^{k} \int_{\mathcal{X}_{ij}} 2\varepsilon_n f_i(x_i) f_j(x_j) dx_i dx_j = O(\varepsilon_n^2)$, since

$$\begin{cases} \int_{\{x_i |\ |\varphi_i(x_i) - \theta_0| \leq \varepsilon_n\}} f_i(x_i) dx_i \leq \frac{2v_0}{\sqrt{2\pi\tau_0^2}} \varepsilon_n, \quad i = 1, \ldots, k, \\ \int_{\{(x_i, x_j) |\ |\varphi_i(x_i) - \varphi_j(x_j)| \leq \varepsilon_n\}} f_i(x_i) f_j(x_j) dx_i dx_j \leq \frac{v_0}{\sqrt{\pi\tau_0^2}} \varepsilon_n, \quad i \neq j, \quad i, j = 1, \ldots, k. \end{cases} \qquad (5.7)$$

For $I_{n,2}$, we state the following facts:

Fact 2

$$\begin{array}{lll}
2.1 & d_{2i}^{*n}(\underset{\sim}{x}, y_i) = 1,\ d_{2i}^{B}(\underset{\sim}{x}, y_i) = 0 \ \Rightarrow & \psi_{in}(x_i, y_i) - \psi_i(x_i, y_i) \geq \theta_0 - \psi_i(x_i, y_i) > 0. \\
2.2 & d_{2i}^{*n}(\underset{\sim}{x}, y_i) = 0,\ d_{2i}^{B}(\underset{\sim}{x}, y_i) = 1 \ \Rightarrow & \psi_{in}(x_i, y_i) - \psi_i(x_i, y_i) < -(\psi_i(x_i, y_i) - \theta_0) \leq 0.
\end{array}$$

Let $U_i = \{(x_i, y_i) |\ |\psi_i(x_i, y_i) - \theta_0| \leq \varepsilon_n\}$. Observe that $\psi_i(X_i, Y_i)$ is $N(\mu_i, \frac{\tau_0^4}{w_0^2})$, where $w_0^2 = \tau_0^2 + \frac{\sigma_0^2}{M+m}$. Then, by Facts 2.1, 2.2 and (5.2),

$$\begin{aligned}
& E_n[I_{n,2}] \\
=\ & E_n \sum_{i=1}^{k} \int_{R^{k-1}} \{ \int_{R^2} [1 - \tau^{*n}(\underset{\sim}{x})] \delta_i^{*n}(\underset{\sim}{x}) [d_{2i}^{*n}(\underset{\sim}{x}, y_i) - d_{2i}^{B}(\underset{\sim}{x}, y_i)] \\
& \qquad \times [\theta_0 - \psi_i(x_i, y_i)] f_i(y_i|x_i) f_i(x_i) dy_i dx_i \} \prod_{l=1, l\neq i}^{k} [f_l(x_l) dx_l]
\end{aligned}$$

$$\leq \sum_{i=1}^{k} \int_{U_i} \varepsilon_n f_i(y_i|x_i) f_i(x_i) dy_i dx_i \tag{5.8}$$

$$+\sum_{i=1}^{k} \int_{R^2-U_i} P_n\{|\psi_{in}(x_i,y_i) - \psi_i(x_i,y_i)| > |\psi_i(x_i,y_i) - \theta_0|\} f_i(y_i|x_i) f_i(x_i) dy_i dx_i$$

$$\leq O(\varepsilon_n^2)$$

$$+\sum_{i=1}^{k} \int_{R^2} P_n\{|\psi_{in}(x_i,y_i) - \psi_i(x_i,y_i)| > \varepsilon_n\} f_i(y_i|x_i) f_i(x_i) dy_i dx_i$$

where $\quad \sum_{i=1}^{k} \int_{U_i} \varepsilon_n f_i(y_i|x_i) f_i(x_i) dy_i dx_i = O(\varepsilon_n^2)$

since

$$\int_{\{(x_i,y_i)||\psi_i(x_i,y_i)-\theta_0|\leq\varepsilon_n\}} f_i(y_i|x_i) f_i(x_i) dy_i dx_i \leq \frac{2w_0}{\sqrt{2\pi\tau_0^2}}\varepsilon_n, \quad i = 1,\ldots,k. \tag{5.9}$$

Now, we consider

$$II_n = \sum_{i=1}^{k}\sum_{j=1}^{k} \int_{\mathcal{X}} [1 - \tau^{*n}(\underset{\sim}{x})] \, \mathrm{I}_{\{j^*=i, j_n^*=j\}} [T_j(\underset{\sim}{x}) - T_i(\underset{\sim}{x})] f(\underset{\sim}{x}) d\underset{\sim}{x}. \tag{5.10}$$

By the mean value theorem, we have

$$\begin{aligned} |T_j(\underset{\sim}{x}) - T_i(\underset{\sim}{x})| &= |T(\varphi_j(x_j)) - T(\varphi_i(x_i))| \leq \sup_{z\in R} |\frac{dT(z)}{dz}| |\varphi_j(x_j) - \varphi_i(x_i)| \\ &\leq |\varphi_j(x_j) - \varphi_i(x_i)| \end{aligned} \tag{5.11}$$

since $|\frac{dT(z)}{dz}| = |-\Phi(q(z-\theta_0))| \leq 1$. From the definitions of i^*, i_n^*, j^* and j_n^*, if $i^* \neq 0$ and $i_n^* \neq 0$ then $i^* = j^*$ and $i_n^* = j_n^*$, respectively. Therefore, by Fact 1.3,(4.9), (4.10), (5.5), (5.7), (5.10) and (5.11)

$$E_n[II_n]$$

$$\leq \sum_{i=1}^{k} \sum_{j=1, j\neq i}^{k} \int_{\mathcal{X}_{ij}} \varepsilon_n f_i(x_i) f_j(x_j) dx_i dx_j$$

$$+\sum_{i=1}^{k} \sum_{j=1, j\neq i}^{k} \int_{R^2-\mathcal{X}_{ij}} \Big[P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| > \frac{|\varphi_i(x_i) - \varphi_j(x_j)|}{2}\}$$

$$+P_n\{|\varphi_{jn}(x_j) - \varphi_j(x_j)| > \frac{|\varphi_i(x_i) - \varphi_j(x_j)|}{2}\}\Big]$$

$$\times |\varphi_i(x_i) - \varphi_j(x_j)| f_i(x_i) f_j(x_j) dx_i dx_j$$

$$\leq O(\varepsilon_n^2) \tag{5.12}$$

$$+\sum_{i=1}^{k} \sum_{j=1, j\neq i}^{k} \int_{R^2} \Big[P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| > \frac{\varepsilon_n}{2}\} + P_n\{|\varphi_{jn}(x_j) - \varphi_j(x_j)| > \frac{\varepsilon_n}{2}\}\Big]$$

$$\times \left[|\varphi_i(x_i) - \mu_i| + |\varphi_j(x_j) - \mu_j| + |\mu_i - \mu_j|\right] f_i(x_i) f_j(x_j) dx_i dx_j.$$

For III_n, we proceed as follows: Recall that $L(q_0, z_0) = 0$ and consider the equation $L(q, z) = 0$. By implicit differentiation,

$$\frac{dq}{dz} = \frac{q^2\Phi(qz)}{\phi(qz)} > 0. \tag{5.13}$$

So, under the condition $L(q, z) = 0$, q is an increasing function of z, and $q(z_0) = q_0 > 0$. Hence, by the continuity property of $q(z)$, there exists $\delta_0 > 0$ such that $0 < q(z) < \infty$, for all $z \in [z_0 - \delta_0, z_0 + \delta_0]$. Let q_{0n} be the estimator of q_0 given in (4.3) and when $\tau_{0n}^2 > 0$, let z_{0n} be the zero of $L(q_{0n}, z)$. That is, $L(q_{0n}, z_{0n}) = 0$. Moreover, we define a function $u : R^2 \to R$ by $u(s,t) = (t\sqrt{ms})^{-1}\sqrt{(M+m)Mt^2 + (2M+m)ts + s^2}$. Observe that $u(\sigma_0^2, \tau_0^2) = q_0$ and

$$\begin{cases} \frac{\partial u}{\partial s} = [s^2 - (M+m)Mt^2]\left(2st\sqrt{ms}\sqrt{(M+m)Mt^2 + (2M+m)ts + s^2}\right)^{-1}, \\ \frac{\partial u}{\partial t} = -s[(2M+m)t + 2s]\left(2t^2\sqrt{ms}\sqrt{(M+m)Mt^2 + (2M+m)ts + s^2}\right)^{-1}. \end{cases} \tag{5.14}$$

Recall that $z_0 < 0$ (see Remark 3.1) and $\frac{\partial L(q,z)}{\partial z} = \Phi(qz)$. For $\varepsilon_n > 0$, $i, j = 1, \ldots, k$, define

$$\begin{cases} H_j &= \{x_j |\ |(\varphi_j(x_j) - \theta_0) - z_0| \le \varepsilon_n\}, \\ D_i &= \{x_i |\ |(\theta_0 - \varphi_i(x_i)) - z_0| \le \varepsilon_n\}, \\ B &= \{(\sigma^2, \tau^2) |\ \sqrt{(\sigma^2 - \sigma_0^2)^2 + (\tau^2 - \tau_0^2)^2} \le \frac{1}{2}\min(\sigma_0^2, \tau_0^2)\}. \end{cases} \tag{5.15}$$

Note that

$$\begin{cases} \int_{H_j} f_j(x_j)dx_j &\le \frac{2v_0}{\sqrt{2\pi\tau_0^2}}\varepsilon_n, \quad j = 1, \ldots, k, \\ \int_{D_i} f_i(x_i)dx_i &\le \frac{2v_0}{\sqrt{2\pi\tau_0^2}}\varepsilon_n, \quad i = 1, \ldots, k. \end{cases} \tag{5.16}$$

In the following, we assume $\varepsilon_n \le \min(\delta_0, \frac{-z_0}{2}, b_0\frac{c_u}{c_q})$ where $b_0 = \frac{1}{2}\min(\sigma_0^2, \tau_0^2)$, $c_u = \max_{(\sigma^2,\tau^2)\in B}\sqrt{(\frac{\partial u(\sigma^2,\tau^2)}{\partial \sigma^2})^2 + (\frac{\partial u(\sigma^2,\tau^2)}{\partial \tau^2})^2} < \infty$ and $c_q = \frac{1}{2}\min_{z_0-\delta_0 \le z \le z_0+\delta_0}|q'(z)| > 0$. Note that c_u is finite by (5.14) and (5.15) and c_q is positive by (5.13) and the definition of δ_0. We state some facts:

Fact 3

3.1 $(\sigma^2, \tau^2) \in R^2 - B \Rightarrow \quad |\sigma^2 - \sigma_0^2| > \frac{b_0}{\sqrt{2}}$ or $|\tau^2 - \tau_0^2| > \frac{b_0}{\sqrt{2}}$.

3.2 If $|z - z_0| > \frac{\varepsilon_n}{2}$ and $(\sigma^2, \tau^2) \in B$, then
$\Rightarrow \quad |q(z) - q(z_0)| > \min\{q(z_0 + \frac{\varepsilon_n}{2}) - q(z_0), q(z_0) - q(z_0 - \frac{\varepsilon_n}{2})\} \ge c_q\varepsilon_n$
$\Rightarrow \quad c_q\varepsilon_n < |q(z) - q(z_0)| = |u(\sigma^2, \tau^2) - u(\sigma_0^2, \tau_0^2)| \le c_u\sqrt{(\sigma^2 - \sigma_0^2)^2 + (\tau^2 - \tau_0^2)^2}$
$\Rightarrow \quad |\sigma^2 - \sigma_0^2| > \frac{c_q}{\sqrt{2}c_u}\varepsilon_n$ or $|\tau^2 - \tau_0^2| > \frac{c_q}{\sqrt{2}c_u}\varepsilon_n$.

3.3 If $i^* = 0, j^* = j$, and $x_j \in H_j$, then

$$\begin{aligned} L(q_0, \varphi_j(x_j) - \theta_0) &= L(q_0, \varphi_j(x_j) - \theta_0) - L(q_0, z_0) \\ &\le \sup_{z \in R} |\frac{\partial L(q_0, z)}{\partial z}|\varepsilon_n \\ &= \varepsilon_n. \end{aligned}$$

3.4 If $i^* = 0,\ j^* = j,\ x_j \in R - H_j$ and $|z_{0n} - z_0| \le \frac{\varepsilon_n}{2}$, then

(a) $\tau^B(\underset{\sim}{x}) = 1,\ \tau^{*n}(\underset{\sim}{x}) = 0$
$\Rightarrow\ \varphi_j(x_j) < \theta_0 - (-z_0) - \varepsilon_n$ and $\theta_0 - (-z_{0n}) < \varphi_{j_n^* n}(x_{j_n^*}) < \theta_0 + (-z_{0n})$
$\Rightarrow\ \varphi_{j_n^* n}(x_{j_n^*}) - \varphi_{j_n^*}(x_{j_n^*}) > \frac{\varepsilon_n}{2}$.

(b) $\tau^B(\underset{\sim}{x}) = 0,\ \tau^{*n}(\underset{\sim}{x}) = 1$
$\Rightarrow\ \theta_0 - (-z_0) + \varepsilon_n < \varphi_j(x_j) < \theta_0$
and ($\varphi_{j_n^* n}(x_{j_n^*}) \le \theta_0 - (-z_{0n})$ or $\varphi_{j_n^* n}(x_{j_n^*}) \ge \theta_0 + (-z_{0n})$)
$\Rightarrow\ \varphi_{jn}(x_j) - \varphi_j(x_j) < -\frac{\varepsilon_n}{2}$ or $\varphi_{j_n^* n}(x_{j_n^*}) - \varphi_{j_n^*}(x_{j_n^*}) > \frac{\varepsilon_n}{2}$.

3.5 If $i^* = j^* = i \neq 0$ and $x_i \in D_i$, then

$$\begin{aligned} L(q_0, \theta_0 - \varphi_i(x_i)) &= L(q_0, \theta_0 - \varphi_i(x_i)) - L(q_0, z_0) \\ &\le \sup_{z \in R} \left|\frac{\partial L(q_0, z)}{\partial z}\right| \varepsilon_n \\ &= \varepsilon_n. \end{aligned}$$

3.6 If $i^* = j^* = i \neq 0,\ x_i \in R - D_i$ and $|z_{0n} - z_0| \le \frac{\varepsilon_n}{2}$, then

(a) $\tau^B(\underset{\sim}{x}) = 1,\ \tau^{*n}(\underset{\sim}{x}) = 0$
$\Rightarrow\ \varphi_i(x_i) > \theta_0 + (-z_0) + \varepsilon_n$ and $\theta_0 - (-z_{0n}) < \varphi_{j_n^* n}(x_{j_n^*}) < \theta_0 + (-z_{0n})$
$\Rightarrow\ \varphi_{in}(x_i) - \varphi_i(x_i) < -\frac{\varepsilon_n}{2}$.

(b) $\tau^B(\underset{\sim}{x}) = 0,\ \tau^{*n}(\underset{\sim}{x}) = 1$
$\Rightarrow\ \theta_0 \le \varphi_i(x_i) < \theta_0 + (-z_0) - \varepsilon_n$
and ($\varphi_{j_n^* n}(x_{j_n^*}) \le \theta_0 - (-z_{0n})$ or $\varphi_{j_n^* n}(x_{j_n^*}) \ge \theta_0 + (-z_{0n})$)
$\Rightarrow\ \varphi_{in}(x_i) - \varphi_i(x_i) < -\frac{\varepsilon_n}{2}$ or $\varphi_{j_n^* n}(x_{j_n^*}) - \varphi_{j_n^*}(x_{j_n^*}) > \frac{\varepsilon_n}{2}$.

From (3.8) and (5.4), we can write III_n as follows:

$$\begin{aligned} III_n &= \sum_{j=1}^{k} \int_{\mathcal{X}} [\tau^{*n}(\underset{\sim}{x}) - \tau^B(\underset{\sim}{x})]\ \mathrm{I}_{\{i^*=0, j^*=j\}} L(q_0, \varphi_j(x_j) - \theta_0) f(\underset{\sim}{x}) d\underset{\sim}{x} \\ &\quad + \sum_{i=1}^{k} \int_{\mathcal{X}} [\tau^{*n}(\underset{\sim}{x}) - \tau^B(\underset{\sim}{x})]\ \mathrm{I}_{\{i^*=j^*=i\}} L(q_0, \theta_0 - \varphi_i(x_i)) f(\underset{\sim}{x}) d\underset{\sim}{x} \\ &= III_{n,1} + III_{n,2}. \end{aligned} \tag{5.17}$$

Now,

$$\begin{aligned} & E_n[III_{n,1}] \\ \le\ & \sum_{j=1}^{k} \int_{H_j} \varepsilon_n f_j(x_j) dx_j \qquad \text{(by Fact 3.3)} \end{aligned}$$

$$
\begin{aligned}
&+\sum_{j=1}^{k} E_n \int_{\mathcal{X}} [\tau^{*n}(\underset{\sim}{x}) - \tau^B(\underset{\sim}{x})] \ \mathrm{I}_{\{i^*=0, j^*=j\}} \quad \mathrm{I}_{\{x_j \in R - H_j\}} \ \mathrm{I}_{\{|z_{0n} - z_0| > \frac{\varepsilon_n}{2}\}} \\
&\qquad\qquad \times \left[\ \mathrm{I}_{\{(\sigma_{0n}^2, \tau_{0n}^2) \in R^2 - B\}} + \ \mathrm{I}_{\{(\sigma_{0n}^2, \tau_{0n}^2) \in B\}} \right] L(q_0, \varphi_j(x_j) - \theta_0) f(\underset{\sim}{x}) d\underset{\sim}{x} \\
&+\sum_{j=1}^{k} E_n \int_{\mathcal{X}} [\tau^{*n}(\underset{\sim}{x}) - \tau^B(\underset{\sim}{x})] \ \mathrm{I}_{\{i^*=0, j^*=j\}} \quad \mathrm{I}_{\{x_j \in R - H_j\}} \ \mathrm{I}_{\{|z_{0n} - z_0| \le \frac{\varepsilon_n}{2}\}} \\
&\qquad\qquad \times L(q_0, \varphi_j(x_j) - \theta_0) f(\underset{\sim}{x}) d\underset{\sim}{x} \\
\le\ & O(\varepsilon_n^2) \qquad\qquad \text{(by (5.16))} \qquad\qquad (5.18) \\
&+ \left[P_n\{|\sigma_{0n}^2 - \sigma_0^2| > \frac{c_q}{\sqrt{2} c_u} \varepsilon_n\} + P_n\{|\tau_{0n}^2 - \tau_0^2| > \frac{c_q}{\sqrt{2} c_u} \varepsilon_n\} \right] \\
&\quad \times \sum_{j=1}^{k} (E[|\varphi_j(X_j) - \mu_j|] + |\mu_j - \theta_0| + \frac{1}{q_0\sqrt{2\pi}} + mc_2) \quad \text{(by Facts 3.1, 3.2 and (3.7))} \\
&+ \sum_{j=1}^{k} \sum_{l=1}^{k} \int_{\mathcal{X}} P_n\{|\varphi_{ln}(x_l) - \varphi_l(x_l)| > \frac{\varepsilon_n}{2}\} \\
&\qquad \times (|\varphi_j(x_j) - \mu_j| + |\mu_j - \theta_0| + \frac{1}{q_0\sqrt{2\pi}} + mc_2) f(\underset{\sim}{x}) d\underset{\sim}{x}. \quad \text{(by Fact 3.4 and (3.7))}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&E_n[III_{n,2}] \\
\le\ & O(\varepsilon_n^2) \qquad\qquad \text{(by Fact 3.5, (5.16))} \qquad\qquad (5.19) \\
&+ \left[P_n\{|\sigma_{0n}^2 - \sigma_0^2| > \frac{c_q}{\sqrt{2} c_u} \varepsilon_n\} + P_n\{|\tau_{0n}^2 - \tau_0^2| > \frac{c_q}{\sqrt{2} c_u} \varepsilon_n\} \right] \\
&\quad \times \sum_{i=1}^{k} (E[|\varphi_i(X_i) - \mu_i|] + |\mu_i - \theta_0| + \frac{1}{q_0\sqrt{2\pi}} + mc_2) \qquad \text{(by Facts 3.1, 3.2)} \\
&+ \sum_{i=1}^{k} \sum_{l=1}^{k} \int_{\mathcal{X}} P_n\{|\varphi_{ln}(x_l) - \varphi_l(x_l)| > \frac{\varepsilon_n}{2}\} \\
&\qquad \times (|\varphi_i(x_i) - \mu_i| + |\mu_i - \theta_0| + \frac{1}{q_0\sqrt{2\pi}} + mc_2) f(\underset{\sim}{x}) d\underset{\sim}{x}. \quad \text{(by Fact 3.6 and (3.7))}
\end{aligned}
$$

Then, in view of (5.1)–(5.4), (5.6), (5.8), (5.12) and (5.17)–(5.19), it suffices to investigate the following:

$$
\left\{
\begin{array}{l}
\int_R P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| > \frac{\varepsilon_n}{2}\} f_i(x_i) dx_i, \\
\int_R P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| > \frac{\varepsilon_n}{2}\} |\varphi_i(x_i) - \mu_i| f_i(x_i) dx_i, \\
\int_{R^2} P_n\{|\psi_{in}(x_i, y_i) - \psi_i(x_i, y_i)| > \varepsilon_n\} f_i(y_i|x_i) f_i(x_i) dy_i dx_i, \\
P_n\{|\sigma_{0n}^2 - \sigma_0^2| > \frac{c_q}{\sqrt{2} c_u} \varepsilon_n\}, \\
P_n\{|\tau_{0n}^2 - \tau_0^2| > \frac{c_q}{\sqrt{2} c_u} \varepsilon_n\}.
\end{array}
\right. \qquad (5.20)
$$

Lemma 5.1 Let S_n be a random variable having a $\chi^2(n)$ distribution. Then,

$$(a)\quad P\{S_n \le n(1-\eta)\} \le \exp(-\frac{n}{2}g_1(\eta)) \text{ for any } \eta,\ 0<\eta<1,$$
$$(b)\quad P\{S_n \ge n(1+\eta)\} \le \exp(-\frac{n}{2}g_2(\eta)) \text{ for any } \eta,\ \eta>0;$$

where
$$g_1(\eta) = -\eta - \ln(1-\eta) \text{ for any } \eta,\ 0<\eta<1,$$
$$g_2(\eta) = \eta - \ln(1+\eta) \text{ for any } \eta,\ \eta>0.$$

Proof: See Corollary 4.1 of Gupta, Liang and Rau (1994).

Remark 5.1 Observe that $g_1(0) = g_2(0) = 0$, $\frac{d}{d\eta}g_1(\eta) > 0$, for $0<\eta<1$, and $\frac{d}{d\eta}g_2(\eta) > 0$, for $\eta>0$. Thus, $g_1(\eta)$ and $g_2(\eta)$ are positive and strictly increasing functions for $0<\eta<1$ and $\eta>0$, respectively. Also, $\lim_{\eta\to 0}\frac{g_1(\eta)}{\eta^2} = \frac{1}{2}$ and $\lim_{\eta\to 0}\frac{g_2(\eta)}{\eta^2} = \frac{1}{2}$.

In the following three lemmas, since the proof is just a straightforward computation, we omit the details.

Lemma 5.2 Let $\mu_{in}, \sigma_{0n}^2, v_{0n}^2$ and τ_{0n}^2 be the estimators of μ_i, σ_0^2, v_0^2 and τ_0^2, respectively, as defined in (4.4). Then, for any $c>0, 0<c_\sigma<\sigma_0^2, 0<c_v<v_0^2$ and $0<c_\tau<\min(2v_0^2, \frac{2\sigma_0^2}{M})$, we have

$$(a)\quad P_n\{|\mu_{in}-\mu_i| \ge c\} \le \frac{2v_0}{\sqrt{2\pi}c}\frac{1}{\sqrt{n}}\exp(\frac{-c^2}{2v_0^2}n),$$
$$(b)\quad P_n\{|\sigma_{0n}^2-\sigma_0^2| \ge c_\sigma\} \le \exp(-\frac{kn(M-1)}{2}g_1(\frac{c_\sigma}{\sigma_0^2})) + \exp(-\frac{kn(M-1)}{2}g_2(\frac{c_\sigma}{\sigma_0^2})),$$
$$(c)\quad P_n\{|v_{0n}^2-v_0^2| \ge c_v\} \le \exp(-\frac{k(n-1)}{2}g_1(\frac{c_v}{v_0^2})) + \exp(-\frac{k(n-1)}{2}g_2(\frac{c_v}{v_0^2})),$$
$$(d)\quad P_n\{|\tau_{0n}^2-\tau_0^2| \ge c_\tau\} \le P_n\{|v_{0n}^2-v_0^2| \ge \frac{\tau_0^2}{2}\} + P_n\{|\sigma_{0n}^2-\sigma_0^2| \ge \frac{M\tau_0^2}{2}\}$$
$$+P_n\{|v_{0n}^2-v_0^2| \ge \frac{c_\tau}{2}\} + P_n\{|\sigma_{0n}^2-\sigma_0^2| \ge \frac{Mc_\tau}{2}\}.$$

Lemma 5.3 Let $\varphi_i(x_i)$ and $\varphi_{in}(x_i)$ be defined as in (2.4) and (4.3), respectively. Then, for any $\varepsilon>0$, any $\kappa>0$ and any $x_i \in R$, we have

$$\frac{1}{2}P_n\{|\varphi_{in}(x_i)-\varphi_i(x_i)| > \varepsilon\}$$
$$\le P_n\{|\mu_{in}-\mu_i| \ge \kappa\} + P_n\{|\mu_{in}-\mu_i| > \frac{Mv_0^2\varepsilon}{5\sigma_0^2}\} + P_n\{|\sigma_{0n}^2-\sigma_0^2| > \frac{M\tau_0^2}{2}\}$$
$$+P_n\{|\sigma_{0n}^2-\sigma_0^2| > \frac{Mv_0^2\varepsilon}{5\kappa}\} + 2P_n\{|\sigma_{0n}^2-\sigma_0^2| > \frac{\sigma_0^2}{5}\frac{\varepsilon v_0^2}{\frac{\sigma_0^2}{M}|x_i-\mu_i|+\varepsilon v_0^2}\}$$
$$+P_n\{|v_{0n}^2-v_0^2| > \frac{\tau_0^2}{2}\} + P_n\{|v_{0n}^2-v_0^2| > \frac{v_0^2}{5}\frac{\varepsilon v_0^2}{\frac{\sigma_0^2}{M}|x_i-\mu_i|+\varepsilon v_0^2}\},$$

Lemma 5.4 Let $\psi_i(x_i, y_i)$ and $\psi_{in}(x_i, y_i)$ be defined as in (2.4) and (4.4), respectively. Also, recall $w_0^2 = \tau_0^2 + \frac{\sigma_0^2}{M+m}$. Then, for any $\varepsilon > 0$, any $\kappa > 0$ and any $(x_i, y_i) \in R^2$, we have

$$\begin{aligned}
&\frac{1}{2} P_n\{|\psi_{in}(x_i, y_i) - \psi_i(x_i, y_i)| > \varepsilon\} \\
\leq\ & P_n\{|\mu_{in} - \mu_i| \geq \kappa\} + P_n\{|\mu_{in} - \mu_i| > \frac{(M+m)w_0^2\varepsilon}{5\sigma_0^2}\} \\
&+ P_n\{|\sigma_{0n}^2 - \sigma_0^2| > \frac{(M+m)w_0^2\varepsilon}{5\kappa}\} + P_n\{|\sigma_{0n}^2 - \sigma_0^2| > \frac{M\tau_0^2}{2}\} \\
&+ 4P_n\{|\sigma_{0n}^2 - \sigma_0^2| \geq \min(\frac{\sigma_0^2}{5}, \frac{Mw_0^2}{20}) \frac{\varepsilon w_0^2}{\frac{\sigma_0^2}{M+m}|\frac{Mx_i+my_i}{M+m} - \mu_i| + \varepsilon w_0^2}\} \\
&+ P_n\{|v_{0n}^2 - v_0^2| \geq \frac{\tau_0^2}{2}\} + P_n\{|v_{0n}^2 - v_0^2| \geq \frac{w_0^2}{20} \frac{\varepsilon w_0^2}{\frac{\sigma_0^2}{M+m}|\frac{Mx_i+my_i}{M+m} - \mu_i| + \varepsilon w_0^2}\},
\end{aligned}$$

Now, consider the terms in (5.20). By Lemma 5.3,

$$\begin{aligned}
&P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| > \frac{\varepsilon_n}{2}\} \\
\leq\ & 2\Big[P_n\{|\mu_{in} - \mu_i| \geq \kappa\} + P_n\{|\mu_{in} - \mu_i| > \frac{Mv_0^2\frac{\varepsilon_n}{2}}{5\sigma_0^2}\} + P_n\{|\sigma_{0n}^2 - \sigma_0^2| > \frac{M\tau_0^2}{2}\} \\
&+ P_n\{|\sigma_{0n}^2 - \sigma_0^2| > \frac{Mv_0^2\frac{\varepsilon_n}{2}}{5\kappa}\} + 2P_n\{|\sigma_{0n}^2 - \sigma_0^2| > \frac{\sigma_0^2}{5} \frac{\frac{\varepsilon_n}{2}v_0^2}{\frac{\sigma_0^2}{M}|x_i - \mu_i| + \frac{\varepsilon_n}{2}v_0^2}\} \\
&+ P_n\{|v_{0n}^2 - v_0^2| > \frac{\tau_0^2}{2}\} + P_n\{|v_{0n}^2 - v_0^2| > \frac{v_0^2}{5} \frac{\frac{\varepsilon_n}{2}v_0^2}{\frac{\sigma_0^2}{M}|x_i - \mu_i| + \frac{\varepsilon_n}{2}v_0^2}\}\Big].
\end{aligned} \tag{5.21}$$

Also, $|\varphi_i(x_i) - \mu_i| = \frac{\tau_0^2}{v_0^2}|x_i - \mu_i|$. By Lemma 5.4, we have

$$\begin{aligned}
&P_n\{|\psi_{in}(x_i, y_i) - \psi_i(x_i, y_i)| > \varepsilon_n\} \\
\leq\ & 2\Big[P_n\{|\mu_{in} - \mu_i| \geq \kappa\} + P_n\{|\mu_{in} - \mu_i| > \frac{(M+m)w_0^2\varepsilon_n}{5\sigma_0^2}\} \\
&+ P_n\{|\sigma_{0n}^2 - \sigma_0^2| > \frac{(M+m)w_0^2\varepsilon_n}{5\kappa}\} + P_n\{|\sigma_{0n}^2 - \sigma_0^2| > \frac{M\tau_0^2}{2}\} \\
&+ 4P_n\{|\sigma_{0n}^2 - \sigma_0^2| \geq \min(\frac{\sigma_0^2}{5}, \frac{Mw_0^2}{20}) \frac{\varepsilon_n w_0^2}{\frac{\sigma_0^2}{M+m}|\frac{Mx_i+my_i}{M+m} - \mu_i| + \varepsilon_n w_0^2}\} \\
&+ P_n\{|v_{0n}^2 - v_0^2| \geq \frac{\tau_0^2}{2}\} + P_n\{|v_{0n}^2 - v_0^2| \geq \frac{w_0^2}{20} \frac{\varepsilon_n w_0^2}{\frac{\sigma_0^2}{M+m}|\frac{Mx_i+my_i}{M+m} - \mu_i| + \varepsilon_n w_0^2}\}\Big].
\end{aligned} \tag{5.22}$$

Moreover, by Lemma 5.2(d),

$$P_n\{|\tau_{0n}^2 - \tau_0^2| > \frac{c_q}{\sqrt{2}c_u}\varepsilon_n\}$$

$$\leq \; P_n\{|v_{0n}^2 - v_0^2| \geq \frac{\tau_0^2}{2}\} + P_n\{|\sigma_{0n}^2 - \sigma_0^2| \geq \frac{M\tau_0^2}{2}\} \qquad (5.23)$$
$$+P_n\{|v_{0n}^2 - v_0^2| \geq \frac{c_q}{2\sqrt{2}c_u}\varepsilon_n\} + P_n\{|\sigma_{0n}^2 - \sigma_0^2| \geq \frac{Mc_q}{2\sqrt{2}c_u}\varepsilon_n\}.$$

Let

$$\begin{cases} \zeta_n &= (\frac{\varepsilon_n}{2}v_0^2)(\frac{\sigma_0^2}{M}|x_i - \mu_i| + \frac{\varepsilon_n}{2}v_0^2)^{-1}, \\ \lambda_n &= (\varepsilon_n w_0^2)(\frac{\sigma_0^2}{M+m}|\frac{Mx_i+my_i}{M+m} - \mu_i| + \varepsilon_n w_0^2)^{-1}. \end{cases} \qquad (5.24)$$

In order to prove the rate of convergence, in view of (5.20) – (5.24), we need to consider the following:

$$\begin{aligned}
&(1) \quad P_n\{|\mu_{in} - \mu_i| \geq \kappa\}, \\
&(2) \quad P_n\{|\mu_{in} - \mu_i| > b_1\varepsilon_n\}, \\
&(3) \quad P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_2\}, \\
&(4) \quad P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_3\varepsilon_n\}, \\
&(5) \quad P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_4\frac{\varepsilon_n}{\kappa}\}, \\
&(6) \quad \int_R P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_5\zeta_n\}f_i(x_i)dx_i, \\
&(7) \quad \int_{R^2} P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_6\lambda_n\}f_i(y_i|x_i)f_i(x_i)dy_idx_i, \qquad (5.25) \\
&(8) \quad P_n\{|v_{0n}^2 - v_0^2| > b_7\}, \\
&(9) \quad P_n\{|v_{0n}^2 - v_0^2| > b_8\varepsilon_n\}, \\
&(10) \quad \int_R P_n\{|v_{0n}^2 - v_0^2| > b_9\zeta_n\}f_i(x_i)dx_i, \\
&(11) \quad \int_{R^2} P_n\{|v_{0n}^2 - v_0^2| > b_{10}\lambda_n\}f_i(y_i|x_i)f_i(x_i)dy_idx_i, \\
&(12) \quad \int_R P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_5\zeta_n\}|x_i - \mu_i|f_i(x_i)dx_i, \\
&(13) \quad \int_R P_n\{|v_{0n}^2 - v_0^2| > b_9\zeta_n\}|x_i - \mu_i|f_i(x_i)dx_i.
\end{aligned}$$

(In the above, (5.25) represents (1)-(13).) where

$$\begin{cases} b_1 = \frac{Mv_0^2}{10\sigma_0^2}, & b_2 = \min(\frac{M\tau_0^2}{2}, \frac{\sigma_0^2}{2}), \quad b_3 = \min(\frac{c_q}{\sqrt{2}c_u}, \frac{\sigma_0^2}{2}), \\ b_4 = \min(\frac{Mv_0^2}{10}, \frac{(M+m)w_0^2}{5}, \frac{\sigma_0^2}{2}), & b_5 = \frac{\sigma_0^2}{5}, \quad b_6 = \min(\frac{\sigma_0^2}{5}, \frac{Mw_0^2}{20}), \\ b_7 = \frac{\tau_0^2}{2}, & b_8 = \min(\frac{c_q}{2\sqrt{2}c_u}, \frac{v_0^2}{2}), \quad b_9 = \frac{v_0^2}{5}, \\ b_{10} = \frac{w_0^2}{20}. \end{cases} \qquad (5.26)$$

In the sequel, we let

$$\begin{cases} \kappa \equiv \kappa_n = \sqrt{\ln n}, \quad \varepsilon_n = \frac{\ln n}{\sqrt{cn}}, \\ \mathcal{G}_i = \{x_i | \, |x_i - \mu_i| < \frac{Mv_0^2}{2\sigma_0^2}\varepsilon_n\sqrt{\frac{n}{c_s \ln n}}\}, \; i = 1, \ldots, k, \\ \mathcal{A}_i = \{(x_i, y_i) | \, |\frac{Mx_i + my_i}{M+m} - \mu_i| < \frac{(M+m)w_0^2}{\sigma_0^2}\varepsilon_n\sqrt{\frac{n}{c_s \ln n}}\}, \; i = 1, \ldots, k, \end{cases} \tag{5.27}$$

where c and c_s are positive constants which will be given later. Then, by the inequality for a $N(0,1)$ distributed random variable Z (see Pollard (1984) Appendix B),

$$P\{Z \geq \eta\} < \frac{1}{\eta}\frac{exp(\frac{-\eta^2}{2})}{\sqrt{2\pi}}, \quad \eta > 0, \tag{5.28}$$

we have, if $c \leq \frac{2v_0^2}{b_1^2}$, (in the sequel, we use (1), (2),..., (13) for the corresponding expressions in (5.25)),

$$(1) \; P_n\{|\mu_{in} - \mu_i| \geq \kappa\} \;\leq\; O(\frac{1}{\sqrt{n \ln n}} \exp(-\frac{1}{2v_0^2} n \ln n)), \tag{5.29}$$

$$(2) \; P_n\{|\mu_{in} - \mu_i| > b_1\varepsilon_n\} \;\leq\; O(\frac{1}{n \ln n}). \tag{5.30}$$

By Lemma 5.2(b), the fact $0 < \frac{b_2}{\sigma_0^2} < 1$ and Remark 5.1, we get $g_j(\frac{b_2}{\sigma_0^2}) > 0, \; j = 1, 2$ and

$$\begin{aligned} (3) \; P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_2\} \;&\leq\; \exp(-\frac{kn(M-1)}{2} g_1(\frac{b_2}{\sigma_0^2})) + \exp(-\frac{kn(M-1)}{2} g_2(\frac{b_2}{\sigma_0^2})) \\ &\leq\; \exp(-\frac{kn(M-1)}{2} \min(g_1(\frac{b_2}{\sigma_0^2}), g_2(\frac{b_2}{\sigma_0^2}))). \end{aligned} \tag{5.31}$$

Similarly, by Lemma 5.2(c), the fact $0 < \frac{b_7}{v_0^2} < 1$ and Remark 5.1, we have $g_j(\frac{b_7}{v_0^2}) > 0, \; j = 1, 2$ and

$$(8) \; P_n\{|v_{0n}^2 - v_0^2| > b_7\} \leq \exp(-\frac{k(n-1)}{2} \min(g_1(\frac{b_7}{v_0^2}), g_2(\frac{b_7}{v_0^2}))). \tag{5.32}$$

From Lemma 5.2(b) and Remark 5.1, if $c \leq \frac{k(M-1)b_3^2}{8\sigma_0^4}$, we get

$$\begin{aligned} &(4) \; P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_3\varepsilon_n\} \\ \leq\;& \exp(-\frac{kn(M-1)}{2} g_1(\frac{b_3\varepsilon_n}{\sigma_0^2})) + \exp(-\frac{kn(M-1)}{2} g_2(\frac{b_3\varepsilon_n}{\sigma_0^2})) \\ =\;& \exp(-\frac{kn(M-1)}{2} (\frac{b_3\varepsilon_n}{\sigma_0^2})^2 \frac{g_1(\frac{b_3\varepsilon_n}{\sigma_0^2})}{(\frac{b_3\varepsilon_n}{\sigma_0^2})^2}) + \exp(-\frac{kn(M-1)}{2} (\frac{b_3\varepsilon_n}{\sigma_0^2})^2 \frac{g_2(\frac{b_3\varepsilon_n}{\sigma_0^2})}{(\frac{b_3\varepsilon_n}{\sigma_0^2})^2}) \\ \leq\;& O(\frac{1}{n}). \end{aligned} \tag{5.33}$$

Similarly, if $c \leq \frac{k(M-1)b_4^2}{8\sigma_0^4}$,

$$(5)\ P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_4\frac{\varepsilon_n}{\kappa}\} \leq O(\frac{1}{n}). \tag{5.34}$$

Also, by Lemma 5.2(c) and Remark 5.1, if $c \leq \frac{kb_8^2}{8v_0^4}$, we have

$$(9)\ P_n\{|v_{0n}^2 - v_0^2| > b_8\varepsilon_n\} \leq O(\frac{1}{n}). \tag{5.35}$$

For (6), we proceed as follows. By (5.27), Remark 5.1 if $c_s \geq \frac{32\sigma_0^4}{k(M-1)b_5^2}$,

$$\begin{aligned}
& x_i \in \mathcal{G}_i \Rightarrow |x_i - \mu_i| < \frac{Mv_0^2}{2\sigma_0^2}\varepsilon_n\sqrt{\frac{n}{c_s \ln n}} \\
& \Rightarrow \zeta_n = \frac{\frac{\varepsilon_n}{2}v_0^2}{\frac{\sigma_0^2}{M}|x_i - \mu_i| + \frac{\varepsilon_n}{2}v_0^2} > \frac{1}{\sqrt{\frac{n}{c_s \ln n}} + 1} > \frac{1}{2}\sqrt{\frac{c_s \ln n}{n}} \\
& \Rightarrow g_j(\frac{b_5}{\sigma_0^2}\zeta_n) > g_j(\frac{b_5}{2\sigma_0^2}\sqrt{\frac{c_s \ln n}{n}}), \quad j = 1, 2 \\
& \Rightarrow \exp(-\frac{kn(M-1)}{2}g_j(\frac{b_5}{\sigma_0^2}\zeta_n)) \\
& \qquad < \exp(-\frac{kn(M-1)}{2}g_j(\frac{b_5}{2\sigma_0^2}\sqrt{\frac{c_s \ln n}{n}})) \\
& \qquad = \exp(-\frac{kn(M-1)}{2}(\frac{b_5}{2\sigma_0^2}\sqrt{\frac{c_s \ln n}{n}})^2 \frac{g_j(\frac{b_5}{2\sigma_0^2}\sqrt{\frac{c_s \ln n}{n}})}{(\frac{b_5}{2\sigma_0^2}\sqrt{\frac{c_s \ln n}{n}})^2}) \\
& \qquad = O(\frac{1}{n}).
\end{aligned} \tag{5.36}$$

Also, by (5.27), (5.28) and the fact that $\frac{|X_i - \mu_i|}{v_0}$ is $N(0,1)$ distributed, if $c \leq \frac{M^2 v_0^2}{8\sigma_0^4 c_s}$, we have

$$\begin{aligned}
EI_{\{R-\mathcal{G}_i\}} &= P\{\frac{|X_i - \mu_i|}{v_0} \geq \frac{Mv_0}{2\sigma_0^2}\varepsilon_n\sqrt{\frac{n}{c_s \ln n}}\} \\
&\leq 2\frac{1}{\frac{Mv_0}{2\sigma_0^2}\varepsilon_n\sqrt{\frac{n}{c_s \ln n}}}\frac{1}{\sqrt{2\pi}}\exp(-\frac{M^2v_0^2}{8\sigma_0^4}\frac{\varepsilon_n^2 n}{c_s \ln n}) \\
&= \frac{4\sigma_0^2}{Mv_0}\sqrt{\frac{c_s c}{\ln n}}\frac{1}{\sqrt{2\pi}}\exp(-\frac{M^2v_0^2 \ln n}{8\sigma_0^4}\frac{}{c_s c}) \\
&\leq O(\frac{1}{n\sqrt{\ln n}}).
\end{aligned} \tag{5.37}$$

Hence, by Lemma 5.2(*b*), (5.36), (5.37) and the fact $0 < \frac{b_5}{\sigma_0^2}\zeta_n < 1$ (which implies $g_j(\frac{b_5}{\sigma_0^2}\zeta_n) > 0,\ j = 1, 2$, by Remark 5.1), if $c_s \geq \frac{32\sigma_0^4}{k(M-1)b_5^2}$ and $c \leq \frac{M^2 v_0^2}{8\sigma_0^4 c_s}$, we obtain

$$
\begin{aligned}
& (6)\ \int_R P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_5\zeta_n\} f_i(x_i)dx_i \\
\leq\ & \int_R \exp(-\frac{kn(M-1)}{2}g_1(\frac{b_5}{\sigma_0^2}\zeta_n)) + \exp(-\frac{kn(M-1)}{2}g_2(\frac{b_5}{\sigma_0^2}\zeta_n)) f_i(x_i)dx_i \\
=\ & \int_{\mathcal{G}_i} \exp(-\frac{kn(M-1)}{2}g_1(\frac{b_5}{\sigma_0^2}\zeta_n)) + \exp(-\frac{kn(M-1)}{2}g_2(\frac{b_5}{\sigma_0^2}\zeta_n)) f_i(x_i)dx_i \qquad (5.38) \\
& + \int_{R-\mathcal{G}_i} \exp(-\frac{kn(M-1)}{2}g_1(\frac{b_5}{\sigma_0^2}\zeta_n)) + \exp(-\frac{kn(M-1)}{2}g_2(\frac{b_5}{\sigma_0^2}\zeta_n)) f_i(x_i)dx_i \\
\leq\ & O(\frac{1}{n}) + O(\frac{1}{n\sqrt{\ln n}}) \\
=\ & O(\frac{1}{n}).
\end{aligned}
$$

Similarly, by Lemma 5.2(*c*), (5.36), (5.37) and the fact $0 < \frac{b_9}{v_0^2}\zeta_n < 1$, if $c_s \geq \frac{32v_0^4}{kb_9^2}$ and $c \leq \frac{M^2 v_0^2}{8\sigma_0^4 c_s}$, we have

$$
(10)\ \int_R P_n\{|v_{0n}^2 - v_0^2| > b_9\zeta_n\} f_i(x_i)dx_i \leq O(\frac{1}{n}). \qquad (5.39)
$$

For (7), following an argument analogous to that of (6), if $c_s \geq \frac{32\sigma_0^4}{k(M-1)b_6^2}$, we have

$$
\begin{aligned}
(x_i, y_i) \in \mathcal{A}_i\ &\Rightarrow\ \lambda_n > \frac{1}{2}\sqrt{\frac{c_s \ln n}{n}} \qquad (5.40) \\
&\Rightarrow\ \exp(-\frac{kn(M-1)}{2}g_j(\frac{b_6}{\sigma_0^2}\lambda_n)) \leq O(\frac{1}{n}),\quad j = 1, 2.
\end{aligned}
$$

Also, if $c \leq \frac{(M+m)^2 w_0^2}{2\sigma_0^4 c_s}$, we get

$$
E\mathrm{I}_{\{R^2-\mathcal{A}_i\}} \leq O(\frac{1}{n\sqrt{\ln n}}). \qquad (5.41)
$$

Therefore, by Lemma (5.2)(*b*), (5.40), (5.41) and the fact $0 < \frac{b_6}{\sigma_0^2}\lambda_n < 1$, if $c_s \geq \frac{32\sigma_0^4}{k(M-1)b_6^2}$ and $c \leq \frac{(M+m)^2 w_0^2}{2\sigma_0^4 c_s}$, we get

$$
\begin{aligned}
& (7)\ \int_{R^2} P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_6\lambda_n\} f_i(y_i|x_i) f_i(x_i)dy_i dx_i \\
\leq\ & \int_{\mathcal{A}_i} [\exp(-\frac{kn(M-1)}{2}g_1(\frac{b_6}{\sigma_0^2}\lambda_n)) + \exp(-\frac{kn(M-1)}{2}g_2(\frac{b_6}{\sigma_0^2}\lambda_n))] f_i(y_i|x_i) f_i(x_i)dy_i dx_i
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{R^2-\mathcal{A}_i}[\exp(-\frac{kn(M-1)}{2}g_1(\frac{b_6}{\sigma_0^2}\lambda_n))+\exp(-\frac{kn(M-1)}{2}g_2(\frac{b_6}{\sigma_0^2}\lambda_n))]f_i(y_i|x_i)f_i(x_i)dy_idx_i \\
\leq\ & O(\frac{1}{n})+O(\frac{1}{n\sqrt{\ln n}}) \qquad (5.42)\\
=\ & O(\frac{1}{n}).
\end{aligned}
$$

Similarly, if $c_s \geq \frac{32v_0^4}{kb_{10}^2}$ and $c \leq \frac{(M+m)^2 w_0^2}{2\sigma_0^4 c_s}$, we have

$$
(11)\ \int_{R^2} P_n\{|v_{0n}^2-v_0^2|>b_{10}\lambda_n\}f_i(y_i|x_i)f_i(x_i)dy_idx_i \leq O(\frac{1}{n}). \qquad (5.43)
$$

Now, we consider (12), by Lemma 5.2(b),

$$
\begin{aligned}
&(12)\ \int_R P_n\{|\sigma_{0n}^2-\sigma_0^2|>b_5\zeta_n\}|x_i-\mu_i|f_i(x_i)dx_i \\
\leq\ & \int_R[\exp(-\frac{kn(M-1)}{2}g_1(\frac{b_5}{\sigma_0^2}\zeta_n))+\exp(-\frac{kn(M-1)}{2}g_2(\frac{b_5}{\sigma_0^2}\zeta_n))]|x_i-\mu_i|f_i(x_i)dx_i \\
=\ & I_{12,1}+I_{12,2}, \qquad (5.44)
\end{aligned}
$$

where

$$
\begin{cases}
I_{12,1}=\int_{\mathcal{G}_i}[\exp(-\frac{kn(M-1)}{2}g_1(\frac{b_5}{\sigma_0^2}\zeta_n))+\exp(-\frac{kn(M-1)}{2}g_2(\frac{b_5}{\sigma_0^2}\zeta_n))]|x_i-\mu_i|f_i(x_i)dx_i, \\
I_{12,2}=\int_{R-\mathcal{G}_i}[\exp(-\frac{kn(M-1)}{2}g_1(\frac{b_5}{\sigma_0^2}\zeta_n))+\exp(-\frac{kn(M-1)}{2}g_2(\frac{b_5}{\sigma_0^2}\zeta_n))]|x_i-\mu_i|f_i(x_i)dx_i.
\end{cases}
$$

By (5.36) and the fact $E|X_i-\mu_i|<+\infty$, if $c_s \geq \frac{32\sigma_0^4}{k(M-1)b_5^2}$, we get

$$
I_{12,1} \leq O(\frac{1}{n}). \qquad (5.45)
$$

By (5.27), Remark 5.1 and the fact $0<\frac{b_5}{\sigma_0^2}\zeta_n<1$, if $c \leq \frac{M^2v_0^2}{8\sigma_0^4 c_s}$, it follows

$$
\begin{aligned}
I_{12,2}\ \leq\ & 2\int_{R-\mathcal{G}_i}|x_i-\mu_i|f_i(x_i)dx_i \\
=\ & 2\int_{\{|x_i-\mu_i|\geq \frac{Mv_0^2}{2\sigma_0^2}\varepsilon_n\sqrt{\frac{n}{c_s\ln n}}\}}|x_i-\mu_i|f_i(x_i)dx_i \\
\leq\ & 2v_0\int_{\{|z|\geq\frac{Mv_0}{2\sigma_0^2}\sqrt{\frac{\ln n}{c_s c}}\}}|z|d\Phi(z) \\
\leq\ & \frac{4v_0}{\sqrt{2\pi}}\exp(-\frac{M^2v_0^2}{8\sigma_0^4}\frac{\ln n}{c_s c}) \qquad (5.46)\\
\leq\ & O(\frac{1}{n}).
\end{aligned}
$$

Hence, by (5.44) – (5.46), if $c_s \geq \frac{32\sigma_0^4}{k(M-1)b_5^2}$ and $c \leq \frac{M^2 v_0^2}{8\sigma_0^4 c_s}$, we obtain

$$(12) \int_R P_n\{|\sigma_{0n}^2 - \sigma_0^2| > b_5\zeta_n\}|x_i - \mu_i| f_i(x_i)dx_i \leq O(\frac{1}{n}). \tag{5.47}$$

Similarly, by Lemma 5.2(c), if $c_s \geq \frac{32 v_0^4}{k b_9^2}$ and $c \leq \frac{M^2 v_0^2}{8\sigma_0^4 c_s}$, we get

$$(13) \int_R P_n\{|v_{0n}^2 - v_0^2| > b_9\zeta_n\}|x_i - \mu_i| f_i(x_i)dx_i \leq O(\frac{1}{n}). \tag{5.48}$$

Hence, by (5.20), (5.25) – (5.27), (5.29) – (5.35), (5.38), (5.39), (5.42), (5.43), (5.47) and (5.48), if $c_s \geq \frac{32}{k}\max(\frac{\sigma_0^4}{(M-1)b_6^2}, \frac{v_0^4}{b_{10}^2})$ and $c \leq \min(\frac{2v_0^2}{b_1^2}, \frac{k(M-1)b_3^2}{8\sigma_0^4}, \frac{k(M-1)b_4^2}{8\sigma_0^4}, \frac{kb_8^2}{8v_0^4}, \frac{M^2v_0^2}{8\sigma_0^4 c_s})$, we see that the expressions in (5.20) have the rate of convergence at least of order $O(\frac{1}{n})$. Now, by (5.1) – (5.4), (5.6), (5.8), (5.12), (5.17) – (5.20) and the above fact, we conclude the following theorem.

Theorem 5.1 Let $\{(\tau^{*n}, \utilde{\delta}^{*n}, \utilde{d}_1^{*n}, \utilde{d}_2^{*n})\}_{n=1}^{\infty}$ be the sequence of empirical Bayes two-stage selection procedures constructed in Section 4. Then,

$$E_n R(\tau^{*n}, \utilde{\delta}^{*n}, \utilde{d}_1^{*n}, \utilde{d}_2^{*n}) - R(\tau^B, \utilde{\delta}^B, \utilde{d}_1^B, \utilde{d}_2^B) \leq O(\varepsilon_n^2) = O(\frac{(\ln n)^2}{n}).$$

That is, the empirical Bayes two-stage selection procedure $(\tau^{*n}, \utilde{\delta}^{*n}, \utilde{d}_1^{*n}, \utilde{d}_2^{*n})$ is asymptotically optimal with convergence rate of order $O(\frac{(\ln n)^2}{n})$.

6 Small Sample Performance: Simulation Study

Let $E_{\mathcal{X}}$ and $E_{\mathcal{Y}|\mathcal{X}}$ be the expectations taken with respect to the probability measures generated by $\utilde{X} = (X_1, \ldots, X_k)$ and $(Y_i|X_i,\ i = 1, \ldots, k)$, respectively. For any first-stage observation $\utilde{X}$, second-stage observation $\utilde{Y}$ and past observations $X_{ijl}(i = 1, \ldots, k,\ j = 1, \ldots, M$ and let $l = 1, \ldots, n)$, let

$$\begin{aligned}
D_{I,1,n}(\utilde{X}) &= \tau^{*n}(\utilde{X})[d_{1i^*}^B(\utilde{X})\varphi_{i^*}(X_{i^*}) - d_{1i_n^*}^{*n}(\utilde{X})\varphi_{i_n^*}(X_{i_n^*})], \\
D_{I,2,n}(\utilde{X}, \utilde{Y}) &= [1 - \tau^{*n}(\utilde{X})]\textstyle\sum_{i=1}^k \delta_i^{*n}(\utilde{X})[d_{2i}^{*n}(\utilde{X}, Y_i) - d_{2i}^B(\utilde{X}, Y_i)][\theta_0 - \psi_i(X_i, Y_i)], \\
D_{II,n}(\utilde{X}) &= [1 - \tau^{*n}(\utilde{X})][\delta_{j_n^*}^{*n}(\utilde{X})T_{j_n^*}(\utilde{X}) - \delta_{j^*}^B(\utilde{X})T_{j^*}(\utilde{X})], \\
D_{III,n}(\utilde{X}) &= [\tau^{*n}(\utilde{X}) - \tau^B(\utilde{X})]Q(\utilde{X}), \\
D_n(\utilde{X}, \utilde{Y}) &= D_{I,1,n}(\utilde{X}) + D_{I,2,n}(\utilde{X}, \utilde{Y}) + D_{II,n}(\utilde{X}) + D_{III,n}(\utilde{X}).
\end{aligned}$$

Then, by (5.1) – (5.4) we have

$$E_n R(\tau^{*n}, \utilde{\delta}^{*n}, \utilde{d}_1^{*n}, \utilde{d}_2^{*n}) - R(\tau^B, \utilde{\delta}^B, \utilde{d}_1^B, \utilde{d}_2^B) = E_n E_{\mathcal{X}} E_{\mathcal{Y}|\mathcal{X}} D_n(\utilde{X}, \utilde{Y}).$$

Therefore, by the law of large numbers, the sample mean of $D_n(\utilde{X}, \utilde{Y})$, based on the observations of $\utilde{X}, \utilde{Y}$ and X_{ijl} $(i = 1, \ldots, k,\ j = 1, \ldots, M$ and $l = 1, \ldots, n)$, can be used as an estimator of $E_n R(\tau^{*n}, \utilde{\delta}^{*n}, \utilde{d}_1^{*n}, \utilde{d}_2^{*n}) - R(\tau^B, \utilde{\delta}^B, \utilde{d}_1^B, \utilde{d}_2^B)$.

We carried out a simulation study to investigate the performance of $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$ for small to moderate values of n based on $D_n(\underset{\sim}{X}, \underset{\sim}{Y})$ for $k = 3$ case. The average of $D_n(\underset{\sim}{X}, \underset{\sim}{Y})$ based on 4000 repetitions, which is denoted by $\bar{D}_n$, is used as an estimator of $E_n R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$. We find that the values of $\bar{D}_n$ decrease quite rapidly as n increases. Moreover, the values of $n\bar{D}_n$ decrease in general. This supports Theorem 5.1 that the rate of convergence is at least of order $O(\frac{(\ell n\ n)^2}{n})$. It also seems to indicate that the possible obtainable rate of convergence is of order $O(\frac{1}{n})$.

References

Bechhofer, R. E. and Turnbull, B. W. (1978). Two $(k+1)$-decision selection procedures for comparing k normal means with a specified standard. *J. Amer. Statist. Assoc.*, **73**, 385–392.

Dunnett, Charles W. (1984). Selection of the best treatment in comparison to a control with an application to a medical trial. *Design of Experiments: Ranking and Selection*, (Eds. A. C. Tamhane and T. Santner), Marcel Dekker, 47–66.

Gupta, S. S. and Leu, L. Y. (1991). On Bayes and empirical Bayes rules for selecting good populations. In *Proceedings of the 1990 Taipei Symposium in Statistics,* (Eds. M. T. Chao and P. E. Cheng), Institute of Statistical Science, Academia Sinica, Taipei, 167-182.

Gupta, S. S. and Liang, T. (1989). Selecting the best binomial population: parametric empirical Bayes approach. *J. Statist. Plann. Inference*, **23**, 21–31.

Gupta, S. S., Liang, T. and Rau, R. B. (1994). Empirical Bayes two-stage procedures for selecting the best Bernoulli population compared with a control. In *Statistical Decision Theory and Related Topics - V*, (Eds. S. S. Gupta and J. O. Berger), Springer-Verlag, Berlin-New York, 277–292.

Gupta, S. S., Liang, T. and Rau, R. B. (1994). Empirical Bayes rules for selecting the best normal population compared with a control. To appear in *Statistics & Decisions*.

Gupta, S. S. and Panchapakesan, S. (1979). *Multiple Decision Procedures*. New York: John Wiley.

Gupta, S. S. and Panchapakesan, S. (1985). Subset selection procedures: review and assessment. *American Journal of Mathematical and Management Sciences*, American Sciences Press, **Vol. 5**, Nos. 3 & 4, 235–311.

Gupta, S. S. and Sobel, M. (1958). On selecting a subset which contains all populations better than a standard. *Ann. Math. Statist.*, **29**, 235–244.

Lehmann, E. L. (1961). Some Model I problems of selection. *Ann. Math. Statist.*, **32**, 990–1012.

Pollard, David (1984). *Convergence of Stochastic Processes.* Springer-Verlag.

Robbins, H. (1956). An empirical Bayes approach to statistics. *Proc. Third Berkeley Symp. Math. Statist. Probab.*, **1**, University of California Press, 157–163.

Robbins, H. (1964). The empirical Bayes approach to statistical decision problems. *Ann. Math. Statist.*, **35**, 1–20.

Wilcox, R. R. (1984). Selecting the best population, provided it is better than a standard: the unequal variance case. *J. Amer. Statist. Assoc.*, **79**, 887–891.