Hiroshima Statistical Research Group: Technical Report

Approximate interval estimation for EPMC for improved linear discriminant rule under high dimensional frame work

Masashi Hyodo, Tomohiro Mitani, Tetsuto Himeno and Takashi Seo

(Last Modified: March 9, 2015)

Abstract. An observation is to be classified into one of two multivariate normal populations with equal covariance matrix. In this paper, we consider the confidence intervals for expected probability of misclassification (EPMC) for improved linear discriminant rule in two types of data:namely, large sample data and high dimensional data. Our approximate confidence interval is based on the asymptotic normality of consistent estimator of EPMC. We obtain new results of stochastic expression for two bilinear forms and two quadratic forms which are important for our asymptotic evaluation of EPMC. We prove asymptotic normality under two different frameworks which could be convenient in different situations based on these results. Through simulation study, it is observed that our approximate confidence interval has a good performance not only in high dimensional and large sample settings, but also in large sample settings.

AMS 2010 *Mathematics Subject Classification.* 62H12, 62E30.

Key words and phrases. asymptotic approximations, expected probability of misclassification, high dimensional data, linear discriminant function.

§1. Introduction

We consider the problem of classifying a future observation vector into one of the two population groups Π_1 and Π_2. For each $i = 1, 2$, Π_i denotes a population from a multivariate normal distribution $\mathcal{N}_p(\boldsymbol{\mu}_i, \Sigma)$, and it is supposed that $\boldsymbol{x}_{ij}$, $j = 1, \ldots, N_i$, are observed from the population Π_i. Here, $\boldsymbol{\mu}_i$ $(i = 1, 2)$ and Σ are unknown parameters, and they are estimated by the sample mean vectors $\overline{\boldsymbol{x}}_i = N_i^{-1} \sum_{j=1}^{N_i} \boldsymbol{x}_{ij}$ $(i = 1, 2)$ and the pooled sample covariance matrix $S = n^{-1} \sum_{i=1}^{2} \sum_{j=1}^{N_i} (\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)'$ for $n = N_1 + N_2 - 2$.

The linear discriminant function is defined as

$$\tilde{T}(\boldsymbol{x}) = (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)' S^{-1}\{\boldsymbol{x} - \tfrac{1}{2}(\bar{\boldsymbol{x}}_1 + \bar{\boldsymbol{x}}_2)\}.$$

Observe however that the linear discriminant function $\tilde{T}(\boldsymbol{x})$ has a bias. In fact,

$$\mathrm{E}[\widetilde{T}(\boldsymbol{x})|\boldsymbol{x} \in \Pi_i] = \frac{n(-1)^{i-1}}{2(n-p-1)}\tilde{\Delta}^2 + \frac{n(N_1 - N_2)p}{2(n-p-1)N_1N_2}, \ i = 1, 2,$$

where $\tilde{\Delta}^2 = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)'\Sigma^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$. For this reason, we use the bias-corrected discriminant function defined as

$$T(\boldsymbol{x}) = (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)' S^{-1}\{\boldsymbol{x} - \tfrac{1}{2}(\bar{\boldsymbol{x}}_1 + \bar{\boldsymbol{x}}_2)\} - \frac{n(N_1 - N_2)p}{2(n-p-1)N_1N_2}, \tag{1.1}$$

where the subtraction of $n(N_1 - N_2)p/\{2(n-p-1)N_1N_2\}$ in (1.1) is to guarantee that $\mathrm{E}[T(\boldsymbol{x})|\boldsymbol{x} \in \Pi_i] = n/\{2(n-p-1)\}(-1)^{i-1}\tilde{\Delta}^2, \ i = 1, 2$. Now using $T(\boldsymbol{x})$, a new observation $\boldsymbol{x}$ is to be assigned to Π_1 if $T(\boldsymbol{x}) > 0$, and to Π_2 otherwise.

The performance of this discriminant rule is evaluated by its probabilities of misclassification. The probabilities of misclassification have been obtained with respect to the distribution of the linear discriminant function $\widetilde{T}(\boldsymbol{x})$. There are different types of misclassification probability associated with $\widetilde{T}(\boldsymbol{x})$. These are the conditional probabilities of misclassification (CPMC) and expected probabilities of misclassification (EPMC). The CPMC is defined by

$$L_1 = P[T(\boldsymbol{x}) \le 0|\boldsymbol{x} \in \Pi_1, X], \ L_2 = P[T(\boldsymbol{x}) > 0|\boldsymbol{x} \in \Pi_2, X], \tag{1.2}$$

where $X = (\boldsymbol{x}_{11}, \ldots, \boldsymbol{x}_{1N_1}, \boldsymbol{x}_{21}, \ldots, \boldsymbol{x}_{2N_2})$. We note that the CPMC is the conditional probability of misclassifying an observation $\boldsymbol{x}$ from Π_i into Π_j, $i, j = 1, 2, \ i \neq j$. On the other hand, the EPMC is defined by

$$R_1 = \mathrm{E}[L_1], \ R_2 = \mathrm{E}[L_2]. \tag{1.3}$$

We note that the EPMC is the unconditional probability of misclassifying an observation $\boldsymbol{x}$ from Π_i into Π_j, $i, j = 1, 2, \ i \neq j$. Since the exact expression for the EPMC is very complicated, there are much works for the approximation of EPMC. The asymptotic approximation of EPMC under a framework such that N_1 and N_2 are large with p is fixed has been studied. This approximation is called "large sample approximation". For a review of these results, see, e.g., Okamoto (1963, 1968) and Siotani (1982). Further, asymptotic approximation of EPMC under a framework that N_1, N_2 and p are all large have also been studied (see, e.g., Lachenbruch (1968) and Fujikoshi and Seo (1998)).

This approximation is called "high dimensional and large sample approximation". In addition, Fujikoshi (2000) gave an explicit formula of error bounds for a high dimensional and large sample approximation of EPMC proposed by Lachenbruch (1968). However, as their approximations are functions of unknown parameters, it must be estimated in practice. Based on the large sample approximation, Lachenbruch and Mickey (1968) proposed the asymptotic unbiased estimator of EPMC. On the other hand, Kubokawa, Hyodo and Srivastava (2013) proposed the second order asymptotic unbiased estimator of the EPMC in high dimensional and large sample framework.

In this paper, we consider the interval estimations for the EPMC. Since the exact interval estimations for the EPMC are very difficult problem, there are some works for the approximate confidence interval. McLachlan (1975) proposed an approximate confidence interval for the CPMC based on the large sample approximation. Recently, Chung and Han (2009) proposed the jackknife confidence interval and the bootstrap confidence interval for the CPMC. The problems with these methods are listed below.
(A) Since CPMC is conditional probability, it is more desirable to derive interval estimation of EPMC.
(B) Since these methods are based on large sample asymptotic results, these methods do not perform well in high dimensional settings.
For the problems (A) and (B), we derive the asymptotic distribution of the estimator of EPMC under the high dimensional and large sample frame works, and propose the approximate confidence interval for the EPMC. For that purpose, we derive explicit expression of stochastic for two bilinear forms and two quadratic forms. The method used in this paper is to express based on eight primitive random variables, namely four random variables having the standard normal distribution and four random variables having chi-square distributions. This approach not only makes it easier to derive asymptotic distribution of estimator of EPMC, but also enables us to show the asymptotic normality of CPMC. As a by product, we show asymptotic normality of CPMC.

The organization of this paper is as follows. In Section 2, we propose consistent estimator of EPMC. In Section 3, we propose new approximate confidence interval of EPMC and show the asymptotic normality of CPMC. In Section 4, we investigate the performances of our approximate confidence intervals through the numerical studies. The conclusion of our study is summarized in Section 5. Some preliminary results are given in Appendix.

§2. The consistent estimator of EPMC

In this section, we propose the consistent estimator of the EPMC. Since R_2 can be obtained from R_1 simply by interchanging N_1 and N_2, we only deal with R_1. Let $\tilde{c} = p/n,\ \tilde{\gamma}_1 = N_1/n,\ \tilde{\gamma}_2 = N_2/n$. We assume the following asymptotic frameworks, in order to derive limiting value of R_1.

(A1) $n,\ p \to \infty$ with $n(\tilde{c} - c) \to 0$ for some $c \in (0, 1)$,

(A2) $n,\ N_1,\ N_2 \to \infty$ with $n(\tilde{\gamma}_1 - \gamma_1) \to 0,\ n(\tilde{\gamma}_2 - \gamma_2) \to 0$ for some $\gamma_1, \gamma_2 \in (0, 1)$,

(A3) $n \to \infty$ with $n(\tilde{\Delta}^2 - \Delta^2) \to 0$ for some $\Delta^2 \in (0, \infty)$.

Suppose that $\boldsymbol{x} \in \Pi_1$. Under these conditions, a conditional distribution of $T(\boldsymbol{x})$ given $(\overline{\boldsymbol{x}}_1, \overline{\boldsymbol{x}}_2, S)$ is distributed as $\mathcal{N}(-U, V)$, where

$$
\begin{aligned}
U =&(\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2)' S^{-1} (\overline{\boldsymbol{x}}_1 - \boldsymbol{\mu}_1) - \frac{1}{2} (\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2)' S^{-1} (\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2) + \frac{n(N_1 - N_2)p}{2(n-p-1)N_1 N_2},\\
V =&(\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2)' S^{-1} \Sigma S^{-1} (\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2).
\end{aligned}
$$

Then, R_1 can be expressed as

$$
R_1 = \mathrm{E}\left[\Phi\left(U V^{-1/2}\right)\right],
$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of $\mathcal{N}(0, 1)$. We rewrite U and V by using

$$
\begin{aligned}
&\boldsymbol{\tau} = \sqrt{(N_1 N_2)/(n+2)} \Sigma^{-1/2} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2),\ \boldsymbol{u}_1 = \sqrt{\frac{N_1 N_2}{n+2}} \Sigma^{-1/2} (\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2),\\
&\boldsymbol{u}_2 = \frac{1}{\sqrt{n+2}} \Sigma^{-1/2} (N_1 \overline{\boldsymbol{x}}_1 + N_2 \overline{\boldsymbol{x}}_2 - N_1 \boldsymbol{\mu}_1 - N_2 \boldsymbol{\mu}_2),\ W = n \Sigma^{-1/2} S \Sigma^{-1/2}.
\end{aligned}
$$

It is seen that $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ and W are mutually independently and distributed as $\boldsymbol{u}_1 \sim \mathcal{N}_p(\boldsymbol{\tau}, I_p)$, $\boldsymbol{u}_2 \sim \mathcal{N}_p(\boldsymbol{0}, I_p)$ and $W \sim \mathcal{W}_p(n, I_p)$, respectively. Using these variables, we can rewrite U and V as

$$
\begin{aligned}
(2.1)\qquad U =& - \frac{(N_1 - N_2)n}{2N_1 N_2} \boldsymbol{u}_1' W^{-1} \boldsymbol{u}_1 + \frac{n}{\sqrt{N_1 N_2}} \boldsymbol{u}_1' W^{-1} \boldsymbol{u}_2 - \frac{n}{N_1} \boldsymbol{\tau}' W^{-1} \boldsymbol{u}_1\\
&+ \frac{n(N_1 - N_2)p}{2(n-p-1)N_1 N_2},\\
(2.2)\qquad V =& \frac{n^2(n+2)}{N_1 N_2} \boldsymbol{u}_1' W^{-2} \boldsymbol{u}_1.
\end{aligned}
$$

Applying Lemma A.1 to (2.1) and (2.2), we obtain the constants U_0 and V_0 as

$$
\begin{aligned}
U_0 &= \lim_{n,p\to\infty} \mathrm{E}[U] = -\frac{\Delta^2}{2(1-c)},\\
V_0 &= \lim_{n,p\to\infty} \mathrm{E}[V] = \frac{1}{(1-c)^3} \left(\Delta^2 + \frac{c}{\gamma_1 \gamma_2}\right).
\end{aligned}
$$

Also, the expectations $\mathrm{E}[(U-U_0)^2]$ and $\mathrm{E}[(V-V_0)^2]$ can be evaluated as

$$\begin{aligned}
\text{(2.3)} \quad \mathrm{E}\left[(U-U_0)^2\right] &= \frac{1}{2n(1-c)^3}\left\{\Delta^4+\frac{2}{\gamma_2}\left(\frac{c}{\gamma_1}+\Delta^2\right)\right. \\
&\quad \left.+\frac{c(\gamma_1-\gamma_2)^2}{\gamma_1^2\gamma_2^2}\right\}+o(n^{-1}), \\
\text{(2.4)} \quad \mathrm{E}\left[(V-V_0)^2\right] &= \frac{2}{n(1-c)^7}\left[(c+4)\Delta^4+\frac{2\left\{(c+1)^2+c\right\}}{\gamma_1\gamma_2}\Delta^2\right. \\
&\quad \left.+\frac{c\left\{(c+1)^2+c\right\}}{\gamma_1^2\gamma_2^2}\right]+o(n^{-1}).
\end{aligned}$$

under the asymptotic frameworks (A1)-(A3). (See details in Appendix B and C.) Thus, using (2.3), (2.4) and Chebyshev's inequality, we have that $U \xrightarrow{p} U_0$ and $V \xrightarrow{p} V_0$. Furthermore, using continuous mapping theorem, we obtain that

$$\left|\Phi\left(UV^{-1/2}\right)-\Phi\left(U_0V_0^{-1/2}\right)\right| \xrightarrow{p} 0 \tag{2.5}$$

under the asymptotic frameworks (A1)-(A3). On the other hand, it holds that

$$\left|\Phi\left(UV^{-1/2}\right)-\Phi\left(U_0V_0^{-1/2}\right)\right| < 1 \ a.s. \tag{2.6}$$

Combining (2.5), (2.6) and dominated convergence theorem, we obtain the following lemma.

Lemma 2.1. *Under the asymptotic frameworks* (A1)-(A3)*, it holds that*

$$R_1 \to \Phi\left(-\frac{(1-c)^{1/2}\Delta^2}{2\sqrt{\Delta^2+c/(\gamma_1\gamma_2)}}\right).$$

Since the limiting value of R_1 is a function of Δ^2, we begin by obtaining its consistent estimator.

Lemma 2.2. *The estimator of Δ^2 is defined by*

$$\widehat{\Delta}^2 = \frac{n-p-1}{n}(\overline{\boldsymbol{x}}_1-\overline{\boldsymbol{x}}_2)'S^{-1}(\overline{\boldsymbol{x}}_1-\overline{\boldsymbol{x}}_2)-\frac{(n+2)p}{N_1N_2}.$$

Under the asymptotic frameworks (A1)-(A3)*, it holds that* $\widehat{\Delta}^2 \xrightarrow{p} \Delta^2$.

(Proof) We can rewrite the estimator $\widehat{\Delta}^2$

$$\widehat{\Delta}^2 = \frac{(n-p-1)(n+2)}{N_1N_2}\boldsymbol{u}_1'W^{-1}\boldsymbol{u}_1-\frac{(n+2)p}{N_1N_2}. \tag{2.7}$$

Applying Lemma A.1 to (2.7), we have

$$\mathrm{E}[(\widehat{\Delta}^2 - \Delta^2)^2] = \frac{1}{n(1-c)} \left(2\Delta^4 + \frac{4\Delta^2}{\gamma_1\gamma_2} + \frac{2c}{\gamma_1^2\gamma_2^2} \right) + o(n^{-1}) \tag{2.8}$$

under the asymptotic frameworks (A1)-(A3). (See details in Appendix D.) Thus, using (2.8) and Chebyshev's inequality, we have $\widehat{\Delta}^2 \xrightarrow{p} \Delta^2$ under the asymptotic frameworks (A1)-(A3). □

Substituting the consistent estimator $\widehat{\Delta}^2$ into the limiting term $\Phi(U_0 V_0^{-1/2})$, the consistent estimator of R_1 is obtained by

$$\widehat{R}_1 = \Phi\left(\widehat{U}_0 \widehat{V}_0^{-\frac{1}{2}} \right),$$

where $\widehat{U}_0 = -2^{-1}(1-c)^{-1}\widehat{\Delta}^2$ and $\widehat{V}_0 = (1-c)^{-3}\{\widehat{\Delta}^2 + c/(\gamma_1\gamma_2)\}$. The following corollary is obtained from continuous mapping theorem and consistency of estimator $\widehat{\Delta}^2$.

Corollary 2.1. *Under the asymptotic frameworks* (A1)-(A3)*, it holds that* $\widehat{R}_1 \xrightarrow{p} R_1$.

§3. Approximate interval estimation for EPMC and asymptotic normality of CPMC

In Section 3.1, we show the asymptotic normality of the estimator of EPMC under two different frameworks, and propose the approximate confidence interval. In Section 3.2, we also show the asymptotic normality of CPMC.

3.1. The asymptotic normality of the estimator of EPMC

At first, we derive the asymptotic distribution of the studentized statistics under the high dimensional frameworks (A1)-(A3). We consider the following random variable

$$\sqrt{n}\left(\widehat{R}_1 - \Phi\left(U_0 V_0^{-\frac{1}{2}} \right) \right).$$

To show the asymptotic normality of the above random variable, we consider the stochastic expansions of $\widehat{U}$ and $\widehat{V}$. Since the statistics $\widehat{U}$ and $\widehat{V}$ are the functions of $\widehat{\Delta}^2$, it is essential to derive the stochastic expansion of $\widehat{\Delta}^2$. By using $\boldsymbol{u}_1$ and W, we rewrite $\widehat{\Delta}^2$ as

$$\widehat{\Delta}^2 = \frac{(n-p-1)(n+2)}{N_1 N_2} \boldsymbol{u}_1 W^{-1} \boldsymbol{u}_1 - \frac{(n+2)p}{N_1 N_2}.$$

Define the variables

$$v_1 = \frac{\tilde{v}_1 - (p-2)}{\sqrt{2(p-2)}},\ v_2 = \frac{\tilde{v}_2 - (n-p+1)}{\sqrt{2(n-p+1)}},$$

where

$$\tilde{v}_1 \sim \chi^2_{p-2},\ \tilde{v}_2 \sim \chi^2_{n-p+1}.$$

Here, χ^2_a $(a \in \mathbb{N})$ means chi-square distribution with a degrees of freedom. The estimator $\widehat{\Delta}^2$ is expanded as

$$\widehat{\Delta}^2 = \Delta^2 + \frac{D_1}{\sqrt{n}} + o_p(n^{-1/2}), \tag{3.1}$$

where $D_1 = g_1 v_1 + g_2 v_2 + g_3 u_1$. Here,

$$u_1 \sim \mathcal{N}(0,1),\ g_1 = \frac{\sqrt{2c}}{\gamma_1\gamma_2},\ g_2 = -\frac{\sqrt{2}\left(c + \Delta^2\gamma_1\gamma_2\right)}{\sqrt{1-c}\gamma_1\gamma_2},\ g_3 = \frac{2\Delta}{\sqrt{\gamma_1\gamma_2}}$$

and v_1, v_2 and u_1 are mutually independent. From (3.1), it is noted that

$$\widehat{U}_0 = U_0 + c_1\frac{D_1}{\sqrt{n}} + o_p(n^{-1/2}),\ \widehat{V}_0 = V_0 + c_2\frac{D_1}{\sqrt{n}} + o_p(n^{-1/2}), \tag{3.2}$$

for $c_1 = -\{2(1-c)\}^{-1}$ and $c_2 = (1-c)^{-3}$. Using (3.2) and Taylor series expansion, it follows that

$$\begin{aligned}\widehat{U}_0\widehat{V}_0^{-\frac{1}{2}} =& U_0V_0^{-\frac{1}{2}} + V_0^{-\frac{1}{2}}\left(c_1\frac{D_1}{\sqrt{n}} - \frac{U_0}{2V_0}c_2\frac{D_1}{\sqrt{n}}\right) + o_p(n^{-1/2})\\ =& U_0V_0^{-\frac{1}{2}} + \frac{1}{\sqrt{n}}Q_1 + o_p(n^{-1/2}),\end{aligned}$$

where

$$Q_1 = q_1v_1 + q_2v_2 + q_3u_1.$$

Here

$$\begin{aligned} q_1 &= -\frac{\sqrt{c(1-c)}\left(2c + \Delta^2\gamma_1\gamma_2\right)}{2\sqrt{2}\gamma_1^2\gamma_2^2\left\{c(\gamma_1\gamma_2)^{-1} + \Delta^2\right\}^{3/2}},\ q_2 = \frac{2c + \Delta^2\gamma_1\gamma_2}{2\sqrt{2}\gamma_1\gamma_2\sqrt{c(\gamma_1\gamma_2)^{-1} + \Delta^2}},\\ q_3 &= -\frac{\sqrt{1-c}\left(2c\Delta + \Delta^3\gamma_1\gamma_2\right)}{2\left(c + \Delta^2\gamma_1\gamma_2\right)^{3/2}}.\end{aligned}$$

From the stochastic expansion of $\widehat{U}_0\widehat{V}_0^{-\frac{1}{2}}$, we have

(3.3)

$$\widehat{R}_1 = \Phi\left(-\frac{(1-c)^{1/2}\Delta^2}{2\sqrt{\Delta^2 + c/(\gamma_1\gamma_2)}}\right) + \phi\left(-\frac{(1-c)^{1/2}\Delta^2}{2\sqrt{\Delta^2 + c/(\gamma_1\gamma_2)}}\right)\frac{Q_1}{\sqrt{n}} + o_p(n^{-1/2}),$$

where $\phi(\cdot)$ is the p.d.f. of the standard normal distribution. Note that u_1 is distributed as $\mathcal{N}(0,1)$, v_1 and v_2 are asymptotically distributed as $\mathcal{N}(0,1)$ under the asymptotic framework (A1), and these variables are mutually independent. Hence, under the asymptotic frameworks (A1)-(A3), it holds that

$$\frac{\sqrt{n}\left(\widehat{R}_1 - \Phi\left(-\frac{(1-c)^{1/2}\Delta^2}{2\sqrt{\Delta^2+c/(\gamma_1\gamma_2)}}\right)\right)}{\sigma_e(\Delta^2)} \xrightarrow{d} \mathcal{N}(0,1), \tag{3.4}$$

where

$$\begin{aligned}\sigma_e(\Delta^2) &= \phi\left(-\frac{(1-c)^{1/2}\Delta^2}{2\sqrt{\Delta^2+c/(\gamma_1\gamma_2)}}\right)\sqrt{q_1^2+q_2^2+q_3^2} \\ &= \phi\left(-\frac{(1-c)^{1/2}\Delta^2}{2\sqrt{\Delta^2+c/(\gamma_1\gamma_2)}}\right)\frac{\left(2c+\Delta^2\gamma_1\gamma_2\right)\sqrt{c+\Delta^2\gamma_1\gamma_2\left(\Delta^2\gamma_1\gamma_2+2\right)}}{2\sqrt{2\gamma_1\gamma_2}\left(c+\Delta^2\gamma_1\gamma_2\right)^{3/2}}.\end{aligned}$$

Now turn to evaluate the difference of the limiting value of R_1 and R_1. The remainder after using first term of the Taylor series of $\Phi(\cdot)$ at $UV^{-1/2} = U_0V_0^{-1/2}$ is given by

$$\frac{\Phi^{(2)}(d)}{2!}\left(\frac{U}{V^{1/2}} - \frac{U_0}{V_0^{1/2}}\right)^2$$

for some value d between $UV^{-1/2}$ and $U_0V_0^{-1/2}$, and $|\Phi^{(2)}(d)|$ is equal or smaller than $1/(\sqrt{2\pi e})$ uniformly in $d \in (-\infty,\infty)$. Here, $\Phi^{(2)}(\cdot)$ is second derivative function of $\Phi(\cdot)$. Hence, we have that

$$\begin{aligned}&\left|R_1 - \left(\Phi\left(U_0V_0^{-1/2}\right) + \phi\left(U_0V_0^{-1/2}\right)\mathrm{E}\left[\frac{U}{V^{1/2}} - \frac{U_0}{V_0^{1/2}}\right]\right)\right| \\ &\leq \frac{1}{2\sqrt{2\pi e}}\mathrm{E}\left[\left(\frac{U}{V^{1/2}} - \frac{U_0}{V_0^{1/2}}\right)^2\right].\end{aligned} \tag{3.5}$$

We note that

$$\begin{aligned}\frac{U}{V^{1/2}} - \frac{U_0}{V_0^{1/2}} &= \frac{1}{\sqrt{V_0}}(U-U_0) + \frac{U_0}{2V_0^{3/2}}(V_0-V) \\ &+\frac{U_0}{V_0^{3/2}}\left(\frac{1}{2\left(\sqrt{V_0/V}+1\right)} + \frac{\sqrt{V_0}}{2\sqrt{V}\left(\sqrt{V_0/V}+1\right)^2}\right)\frac{(V_0-V)^2}{V} \\ &+\frac{1}{\sqrt{V_0}+V_0/\sqrt{V}}\frac{(U-U_0)(V_0-V)}{V}.\end{aligned} \tag{3.6}$$

From (A.8) and (A.11)

$$\mathrm{E}\left[\frac{1}{\sqrt{V_0}}(U-U_0)\right] = -\frac{\Delta^2}{2\sqrt{V_0}(1-c)^2 n} + o(n^{-1}) \tag{3.7}$$

$$\begin{aligned}\mathrm{E}\left[\frac{U_0}{2V_0^{3/2}}(V_0-V)\right] &= -\frac{U_0}{2n(1-c)^3 V_0^{3/2}}\left\{\left(\frac{4}{1-c}-1\right)\Delta^2\right.\\ &\left.+\frac{c}{\gamma_1\gamma_2}\left(\frac{4}{1-c}+1\right)\right\}+o(n^{-1}).\end{aligned} \tag{3.8}$$

Since $\sqrt{V_0/V}+1>1$ and $\sqrt{V}\left(\sqrt{V_0/V}+1\right)^2>2\sqrt{V_0}$,

$$\begin{aligned}&\left|\mathrm{E}\left[\frac{U_0}{V_0^{3/2}}\left(\frac{1}{2\left(\sqrt{V_0/V}+1\right)}+\frac{\sqrt{V_0}}{2\sqrt{V}\left(\sqrt{V_0/V}+1\right)^2}\right)\frac{(V_0-V)^2}{V}\right]\right|\\ &<\frac{3|U_0|}{4V_0^{3/2}}\mathrm{E}\left[\frac{(V_0-V)^2}{V}\right]\end{aligned} \tag{3.9}$$

By using Lemma A.1, we obtain that

$$\mathrm{E}\left[\frac{(V-V_0)^2}{V}\right]=O(n^{-1}) \tag{3.10}$$

under the asymptotic frameworks (A1)-(A3). (See details in Appendix E.) From (3.9) and (3.10),

$$\mathrm{E}\left[\frac{U_0}{V_0^{3/2}}\left(\frac{1}{2\left(\sqrt{V_0/V}+1\right)}+\frac{\sqrt{V_0}}{2\sqrt{V}\left(\sqrt{V_0/V}+1\right)^2}\right)\frac{(V_0-V)^2}{V}\right]=O(n^{-1}). \tag{3.11}$$

By using $\sqrt{V_0}+V_0/\sqrt{V}>\sqrt{V_0}>0$ and Cauchy Schwarz inequality,

$$\begin{aligned}&\left|\mathrm{E}\left[\frac{1}{\sqrt{V_0}+V_0/\sqrt{V}}\frac{(U-U_0)(V_0-V)}{V}\right]\right|\\ &<\mathrm{E}\left[\frac{1}{\sqrt{V_0}}\frac{|U-U_0||V_0-V|}{V}\right]\le\frac{1}{\sqrt{V_0}}\sqrt{\mathrm{E}\left[\frac{|U-U_0|^2}{V}\right]}\sqrt{\mathrm{E}\left[\frac{|V_0-V|^2}{V}\right]}.\end{aligned} \tag{3.12}$$

By using Lemma A.1, we obtain that

$$\mathrm{E}\left[\frac{(U-U_0)^2}{V}\right]=O(n^{-1}) \tag{3.13}$$

under the asymptotic frameworks (A1)-(A3). (See details in Appendix F.) From (3.12) and (3.13),

$$\mathrm{E}\left[\frac{1}{\sqrt{V_0}+V_0/\sqrt{V}}\frac{(U-U_0)(V_0-V)}{V}\right]=O(n^{-1}). \tag{3.14}$$

Combining (3.7),(3.8),(3.11) and (3.14), under the asymptotic frameworks (A1)-(A3), it holds that

$$\mathrm{E}\left[\frac{U}{V^{1/2}}-\frac{U_0}{V_0^{1/2}}\right]=O(n^{-1}). \tag{3.15}$$

Since $\sqrt{V_0V}+V_0\geq V_0>0$,

$$\begin{aligned}\left(\frac{U}{V^{1/2}}-\frac{U_0}{V_0^{1/2}}\right)^2 &= \left(\frac{U-U_0}{\sqrt{V}}+\frac{U_0}{\sqrt{V_0V}+V_0}\frac{V_0-V}{\sqrt{V}}\right)^2\\ &= \frac{(U-U_0)^2}{V}+\frac{U_0^2}{(\sqrt{V_0V}+V_0)^2}\frac{(V_0-V)^2}{V}\\ &\quad +2\frac{U_0}{\sqrt{V_0V}+V_0}\frac{(U-U_0)(V_0-V)}{V}\\ &\leq \frac{(U-U_0)^2}{V}+\frac{U_0^2}{V_0^2}\frac{(V-V_0)^2}{V}+\frac{2U_0}{V_0}\frac{|U-U_0||V-V_0|}{V}.\end{aligned}$$

By using Cauchy Schwarz inequality, we obtain that

$$\begin{aligned}\mathrm{E}\left[\left(\frac{U}{V^{1/2}}-\frac{U_0}{V_0^{1/2}}\right)^2\right] &\leq \mathrm{E}\left[\frac{(U-U_0)^2}{V}\right]+\frac{U_0^2}{V_0^2}\mathrm{E}\left[\frac{(V-V_0)^2}{V}\right]\\ &\quad +\frac{2|U_0|}{V_0}\left(\mathrm{E}\left[\frac{(U-U_0)^2}{V}\right]\right)^{1/2}\left(\mathrm{E}\left[\frac{(V-V_0)^2}{V}\right]\right)^{1/2}.\end{aligned} \tag{3.16}$$

From (3.10),(3.13) and (3.16), we obtain that

$$\mathrm{E}\left[\left(\frac{U}{V^{1/2}}-\frac{U_0}{V_0^{1/2}}\right)^2\right]=O(n^{-1}) \tag{3.17}$$

under the asymptotic frameworks (A1)-(A3). Combining (3.5),(3.15) and (3.17), under the asymptotic frameworks (A1)-(A3), it holds that

$$\left|R_1-\Phi\left(-\frac{(1-c)^{1/2}\Delta^2}{2\sqrt{\Delta^2+c/(\gamma_1\gamma_2)}}\right)\right|=O(n^{-1}). \tag{3.18}$$

By using (3.4) and (3.18), we obtain the following theorem.

Theorem 3.1. *Under the asymptotic frameworks (A1)-(A3), it holds that*

$$T_e = \frac{\sqrt{n}\left(\widehat{R}_1 - R_1\right)}{\sigma_e(\Delta^2)} \xrightarrow{d} \mathcal{N}(0,1).$$

To propose the interval estimation of the EPMC, we need to estimate $\sigma_e(\Delta^2)$. We use truncated estimator

$$\hat{\Delta}_*^2 = \max(\widehat{\Delta}^2, 0),$$

so that the estimator of $\sigma_e(\Delta^2)$ may be negative. Then it holds that

$$|\max(\widehat{\Delta}^2, 0) - \Delta^2| \le |\widehat{\Delta}^2 - \Delta^2| \ a.s. \tag{3.19}$$

By using Markov's inequality, (2.8) and (3.19), we obtain $\hat{\Delta}_*^2 \xrightarrow{p} \Delta^2$ under the asymptotic frameworks (A1)-(A3). Hence, $\hat{\Delta}_*^2$ is a consistent estimator of Δ^2. Assigning the truncated estimator Δ_*^2 to the portion of $\sigma_e(\Delta^2)$ which may be negative, we propose

$$\sigma_e(\hat{\Delta}_*^2) = \phi\left(-\frac{(1-c)^{1/2}\hat{\Delta}_*^2}{2\sqrt{\hat{\Delta}_*^2 + c/(\gamma_1\gamma_2)}}\right)\frac{\left(2c + \hat{\Delta}_*^2\gamma_1\gamma_2\right)\sqrt{c + \hat{\Delta}_*^2\gamma_1\gamma_2\left(\hat{\Delta}_*^2\gamma_1\gamma_2 + 2\right)}}{2\sqrt{2\gamma_1\gamma_2}\left(c + \hat{\Delta}_*^2\gamma_1\gamma_2\right)^{3/2}}.$$

By using the consistent estimator $\sigma_e(\hat{\Delta}_*^2)$, we obtain the following statistics of T_e

$$T_e^* = \frac{\sqrt{n}\left(\widehat{R}_1 - R_1\right)}{\sigma_e(\hat{\Delta}_*^2)}.$$

Therefore we can obtain the following corollary.

Corollary 3.1. *Under the asymptotic frameworks (A1)-(A3), it holds that*

$$T_e^* \xrightarrow{d} \mathcal{N}(0,1).$$

Next, we show that asymptotic normality of T_e^* is also established under the large sample framework

$$(\mathrm{A}'1): p \text{ is fixed and } n \to \infty$$

or

$$(\mathrm{A}''1): n,\ p \to \infty \text{ with } p/\sqrt{n} \to 0.$$

Under the frameworks (A′1) and (A2) or the frameworks (A″1), (A2) and (A3), it holds that

$$R_1 = \Phi\left(-\frac{\Delta}{2}\right) + o(n^{-1/2}),\ \Phi\left(\frac{U_0}{\sqrt{V_0}}\right) = \Phi\left(-\frac{\Delta}{2}\right) + o(n^{-1/2}), \tag{3.20}$$

$$\sigma_e(\Delta^2) = \phi\left(-\frac{\Delta}{2}\right)\frac{\sqrt{\Delta^2 + 2/\gamma_1\gamma_2}}{2\sqrt{2}} + o(1), \tag{3.21}$$

$$\sigma_e(\hat{\Delta}_*^2) \xrightarrow{p} \phi\left(-\frac{\Delta}{2}\right)\frac{\sqrt{\Delta^2 + 2/\gamma_1\gamma_2}}{2\sqrt{2}},$$

$$T_e = \frac{\phi\left(-\frac{\Delta}{2}\right)}{\sigma_e(\Delta^2)}\left(\frac{\Delta}{2\sqrt{2}}v_2 - \frac{1}{2(\gamma_1\gamma_2)^{1/2}}u_1\right) + o_p(1). \tag{3.22}$$

From (3.20)-(3.22), we have that

$$T_e^* = \frac{1}{\sqrt{\frac{\Delta^2}{8} + \frac{1}{4\gamma_1\gamma_2}}}\left(\frac{\Delta}{2\sqrt{2}}v_2 - \frac{1}{2(\gamma_1\gamma_2)^{1/2}}u_1\right) + o_p(1).$$

Therefore we can obtain the following corollary.

Corollary 3.2. *Assume the conditions* (A′1) *and* (A2) *or the conditions* (A″1), (A2) *and* (A3). *Then, it holds that*

$$T_e^* \xrightarrow{d} \mathcal{N}(0, 1).$$

Remark 3.1. *From Corollary 3.1 and 3.2, T_e^* has a asymptotic normality not only under high dimensional and large sample frame work, but also under the large sample framework.*

Based on Corollary 3.1 and 3.2, we propose an approximate $100(1-\alpha)$ percentile confidence interval for EPMC as following:

$$\mathcal{C}_{T_1} = \left[\widehat{R}_1 + \frac{\sigma_e(\widehat{\Delta}_*^2)}{\sqrt{n}}y_{1-\frac{\alpha}{2}},\ \widehat{R}_1 + \frac{\sigma_e(\widehat{\Delta}_*^2)}{\sqrt{n}}y_{\frac{\alpha}{2}}\right], \tag{3.23}$$

where y_α denotes upper 100α percentile of standard normal distribution.

3.2. Asymptotic normality of CPMC

In this section, we show asymptotic normality of CPMC. The CPMC can be expressed as

$$L_1 = \Phi\left(UV^{-\frac{1}{2}}\right).$$

Applying Lemma A.1 to (2.1) and (2.2), we obtain

$$\begin{aligned} U &= -\frac{\tilde{\Delta}^2 n}{2\tilde{v}_2} + \left\{ \frac{np(N_1 - N_2)}{2N_1N_2(n-p-1)} - \frac{n(N_1 - N_2)}{2N_1N_2\tilde{v}_2}\left(u_1^2 + u_2^2 + \tilde{v}_1\right) \right\} \\ &\quad + \frac{nu_3}{\tilde{v}_2\sqrt{N_1N_2}}\sqrt{\left(\tilde{\Delta}\sqrt{\frac{N_1N_2}{n+2}} + u_1\right)^2 + u_2^2 + \tilde{v}_1} \\ &\quad - \frac{nu_4}{\tilde{v}_2\sqrt{N_1N_2}}\sqrt{\frac{\tilde{v}_3}{\tilde{v}_4}}\sqrt{\left(\tilde{\Delta}\sqrt{\frac{N_1N_2}{n+2}} + u_1\right)^2 + u_2^2 + \tilde{v}_1} \\ &\quad - \frac{\tilde{\Delta} n}{\tilde{v}_2\sqrt{(n+2)N_1N_2}}\left(N_1u_1 + N_2u_2\sqrt{\frac{\tilde{v}_3}{\tilde{v}_4}}\right), \\ V &= \left\{ \frac{\tilde{\Delta}^2 n^2}{\tilde{v}_2^2}\left(1 + \frac{\tilde{v}_3}{\tilde{v}_4}\right) + \frac{n^2(n+2)}{N_1N_2\tilde{v}_2^2}\left(1 + \frac{\tilde{v}_3}{\tilde{v}_4}\right)\left(u_1^2 + u_2^2 + \tilde{v}_1\right) \right\} \\ &\quad + \frac{2\tilde{\Delta} n^2\sqrt{n+2}}{\tilde{v}_2^2\sqrt{N_1N_2}}\left(1 + \frac{\tilde{v}_3}{\tilde{v}_4}\right)u_1, \end{aligned}$$

where

$$u_i \sim \mathcal{N}(0,1)\ (i = 1,2,3,4),\ \tilde{v}_1 \sim \chi^2_{p-2},\ \tilde{v}_2 \sim \chi^2_{n-p+1},\ \tilde{v}_3 \sim \chi^2_{p-1},\ \tilde{v}_4 \sim \chi^2_{n-p+2},$$

and these variables are mutually independent. Define the variables

$$v_1 = \frac{\tilde{v}_1 - (p-2)}{\sqrt{2(p-2)}},\ v_2 = \frac{\tilde{v}_2 - (n-p+1)}{\sqrt{2(n-p+1)}},\ v_3 = \frac{\tilde{v}_3 - (p-1)}{\sqrt{2(p-1)}},\ v_4 = \frac{\tilde{v}_2 - (n-p+2)}{\sqrt{2(n-p+2)}}.$$

Note that

$$\begin{aligned} \tilde{v}_1 &= (p-2) + \sqrt{2(p-2)}v_1, \\ \tilde{v}_2 &= (n-p+1) + \sqrt{2(n-p+1)}v_2, \\ \tilde{v}_3 &= (p-1) + \sqrt{2(p-1)}v_3, \\ \tilde{v}_4 &= (n-p+2) + \sqrt{2(n-p+2)}v_4, \end{aligned}$$

and v_1, v_2, v_3 and v_4 are asymptotically distributed as $\mathcal{N}(0,1)$ under the asymptotic framework (A1). By using Taylor series expansion based on these variables, we can expand U stochastically,

$$U = U_0 + \frac{1}{\sqrt{n}}U_1 + o_p(n^{-1/2}), \tag{3.24}$$

where

$$\begin{aligned}
U_0 &= -\frac{1}{2(1-c)}\Delta^2,\\
U_1 &= \frac{\sqrt{c}(\gamma_2-\gamma_1)}{\sqrt{2}(1-c)\gamma_1\gamma_2}v_1 + \frac{\big(c(\gamma_1-\gamma_2)+\Delta^2\gamma_1\gamma_2\big)}{\sqrt{2}(1-c)^{3/2}\gamma_1\gamma_2}v_2 - \frac{\sqrt{\gamma_1}\Delta}{(1-c)\sqrt{\gamma_2}}u_1\\
&\quad -\frac{\sqrt{c\gamma_2}\Delta}{(1-c)^{3/2}\sqrt{\gamma_1}}u_2 + \frac{\sqrt{c+\Delta^2\gamma_1\gamma_2}}{(1-c)\sqrt{\gamma_1\gamma_2}}u_3 - \frac{\sqrt{c}\sqrt{c+\Delta^2\gamma_1\gamma_2}}{(1-c)^{3/2}\sqrt{\gamma_1\gamma_2}}u_4.
\end{aligned}$$

Using similar arguments, we can expand V stochastically,

$$V = V_0 + \frac{V_1}{\sqrt{n}} + o_p(n^{-1/2}), \tag{3.25}$$

where

$$\begin{aligned}
V_0 &= \frac{1}{(1-c)^3}\left(\frac{c}{\gamma_1\gamma_2}+\Delta^2\right),\\
V_1 &= \frac{\sqrt{2c}}{(1-c)^3\gamma_1\gamma_2}v_1 - \frac{2\sqrt{2}\left(c+\Delta^2\gamma_1\gamma_2\right)}{(1-c)^{7/2}\gamma_1\gamma_2}v_2 + \frac{\sqrt{2c}\left(c+\Delta^2\gamma_1\gamma_2\right)}{(1-c)^3\gamma_1\gamma_2}v_3\\
&\quad -\frac{\sqrt{2}c\left(c+\Delta^2\gamma_1\gamma_2\right)}{(1-c)^{7/2}\gamma_1\gamma_2}v_4 + \frac{2\Delta}{(1-c)^3\sqrt{\gamma_1\gamma_2}}u_1.
\end{aligned}$$

By using (3.24), (3.25) and Taylor series expansion, it follows that

$$\begin{aligned}
UV^{-\frac{1}{2}} &= U_0V_0^{-\frac{1}{2}} + \frac{1}{\sqrt{n}V_0^{1/2}}\left\{U_1 - \frac{U_0}{2V_0}V_1\right\} + o_p(n^{-1/2})\\
&= U_0V_0^{-\frac{1}{2}} + \frac{W_1}{\sqrt{n}} + o_p(n^{-1/2}),
\end{aligned}$$

where

$$W_1 = w_1v_1 + w_2v_2 + w_3v_3 + w_4v_4 + w_5u_1 + w_6u_2 + w_7u_3 + w_8u_4.$$

Here,

$$\begin{aligned}
w_1 &= \frac{\sqrt{c(1-c)}\Delta^2}{2\sqrt{2}\gamma_1\gamma_2\left\{c(\gamma_1\gamma_2)^{-1}+\Delta^2\right\}^{3/2}} + \frac{\sqrt{c(1-c)}(\gamma_2-\gamma_1)}{\sqrt{2}\gamma_1\gamma_2\sqrt{c(\gamma_1\gamma_2)^{-1}+\Delta^2}},\\
w_2 &= \frac{(1-2\gamma_2)c}{\sqrt{2}\gamma_1\gamma_2\sqrt{c(\gamma_1\gamma_2)^{-1}+\Delta^2}},\\
w_3 &= \frac{\sqrt{c(1-c)}\Delta^2}{2\sqrt{2}\sqrt{c(\gamma_1\gamma_2)^{-1}+\Delta^2}},\ w_4 = -\frac{c\Delta^2}{2\sqrt{2}\sqrt{c(\gamma_1\gamma_2)^{-1}+\Delta^2}},\\
w_5 &= \frac{\sqrt{1-c}\Delta^3}{2\sqrt{\gamma_1\gamma_2}\left\{c(\gamma_1\gamma_2)^{-1}+\Delta^2\right\}^{3/2}} - \frac{\sqrt{1-c}\Delta\gamma_1}{\sqrt{\gamma_1\gamma_2}\sqrt{c(\gamma_1\gamma_2)^{-1}+\Delta^2}},\\
w_6 &= -\frac{\sqrt{c}\Delta\gamma_2}{\sqrt{\gamma_1\gamma_2}\sqrt{c(\gamma_1\gamma_2)^{-1}+\Delta^2}},\ w_7 = \sqrt{1-c},\ w_8 = -\sqrt{c}.
\end{aligned}$$

Using the Taylor series expansion, L_1 is expressed as

$$L_1 = \Phi(U_0 V_0^{-\frac{1}{2}}) + \phi(U_0 V_0^{-\frac{1}{2}}) \frac{W_1}{\sqrt{n}} + o_p(n^{-1/2}).$$

Since the random variables $v_1, v_2, v_3, v_4, u_1, u_2, u_3$ and u_4 in W_1 are mutually independent and asymptotically (or exactly) distributed as $\mathcal{N}(0,1)$, we obtain the following theorem.

Theorem 3.2. *Under the asymptotic frameworks* (A1)-(A3)*, it holds that*

$$\sqrt{n}(L_1 - R_1) \xrightarrow{d} \mathcal{N}(0, \sigma_2(\Delta^2)),$$

where $\sigma_2(\Delta^2)^2 = \{\phi(U_0 V_0^{-\frac{1}{2}})\}^2 \sum_{i=1}^{8} w_i^2$.

Next, we evaluate asymptotic property of L_1 under the large sample framework. We assume the conditions (A$'$1) and (A2) or the conditions (A$''$1), (A2) and (A3). Then it holds that

$$L_1 \;=\; \Phi\left(-\frac{\Delta}{2}\right) + \phi\left(-\frac{\Delta}{2}\right) \frac{1}{\sqrt{n}} \left(\frac{\gamma_2 - \gamma_1}{2\sqrt{\gamma_1 \gamma_2}} u_1 + u_3 \right) + o_p(n^{-1/2}).$$

Thus, we obtain the following corollary.

Corollary 3.3. *Assume the conditions* (A$'$1) *and* (A2) *or the conditions* (A$''$1)*,* (A2) *and* (A3)*. Then, it holds that*

$$\sqrt{n}\left(L_1 - \Phi\left(-\frac{\Delta}{2}\right)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{4\gamma_1\gamma_2}\phi^2\left(-\frac{\Delta}{2}\right)\right).$$

Remark 3.2. *We consider the relation between the optimal rule*

$$T_{opt}(\boldsymbol{x}) > (\text{resp.} \le) 0 \;\Rightarrow\; \boldsymbol{x} \in \Pi_1 (\text{resp.} \Pi_2), \tag{3.26}$$

and our suggested rule

$$\widetilde{T}(\boldsymbol{x}) > (\text{resp.} \le) 0 \;\Rightarrow\; \boldsymbol{x} \in \Pi_1 (\text{resp.} \Pi_2), \tag{3.27}$$

where

$$\begin{aligned} T_{opt}(\boldsymbol{x}) &= (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)' \Sigma^{-1} \{\boldsymbol{x} - \tfrac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)\}, \\ \widetilde{T}(\boldsymbol{x}) &= (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)' S^{-1} \{\boldsymbol{x} - \tfrac{1}{2}(\bar{\boldsymbol{x}}_1 + \bar{\boldsymbol{x}}_2)\}. \end{aligned}$$

From Corollary 3.3, we note that the distribution of the CPMC of the rule (3.27) under the condition (A$'$1) *or* (A$''$1) *approaches a normal distribution with standard deviation shrinking in proportion to* $1/\sqrt{n}$ *around the error rate of the optimal rule (3.26).*

§4. Simulation study

In this section, we investigate the performance of proposed approximate confidence intervals (3.23). In order to evaluate coverage probabilities of the approximate confidence intervals and the expected lengths, a Monte Carlo study is conducted. Without loss of generality, multivariate normal random samples are generated from $\Pi_1 : \mathcal{N}_p(\mathbf{0}, I_p)$ and $\Pi_2 : \mathcal{N}_p((\sqrt{5}, \mathbf{0}'_{p-1})', I_p)$. The values of N_1, N_2 and p are chosen as follows:

$$\begin{aligned}
&(\text{CaseA}) \quad p = 100, 200,\ \frac{n+2}{p} = 2, 3, 4,\ (N_1 : N_2) = (1:1), (3:1), (1:3), \\
&(\text{CaseB}) \quad p = 5,\ n+2 = 100, 300, 500,\ (N_1 : N_2) = (1:1), (3:1), (1:3).
\end{aligned}$$

In above configuration, we calculate the following coverage probabilities

$$CP = \frac{\sharp\{(\widehat{R}_1, \widehat{\Delta}^2_*) | R_1 \in [\widehat{R}_1 + n^{-1/2}\sigma_e(\widehat{\Delta}^2_*)y_{1-\alpha/2},\ \widehat{R}_1 + n^{-1/2}\sigma_e(\widehat{\Delta}^2_*)y_{\alpha/2}]\}}{sim},$$

and the following expected lengths of approximate confidence interval

$$EL = \mathrm{E}[n^{-1/2}\sigma_e(\widehat{\Delta}^2_*)(y_{\alpha/2} - y_{1-\alpha/2})],$$

where $\sharp\{\cdot\}$ denotes number of element of set $\{\cdot\}$, sim denotes replication number of simulation. We also estimate the exact expected length by using Monte Carlo simulation as follows:

$$EEL = \widehat{R}_{1(\alpha/2 \times sim)} - \widehat{R}_{1((1-\alpha/2)\times sim)},$$

where $\widehat{R}_{1(i)}$ denotes i-th largest value among the sim values. Tables 1-3 give the coverage probabilities when $p = 100, 200$ and 5, respectively. Tables 4-6 give the expected lengths of approximate confidence interval and exact expected length when $p = 100, 200$ and 5, respectively. As can be seen from the Tables 1-3, when the sample size or dimension is increased, probability for approximate confidence interval is close to confidence level. In addition, we observe that our approximations have a high level of accuracy in different situations: large sample settings (Table 3), high dimensional and large sample settings (Table 1-2). From Tables 4-6, when the sample sizes increase, the expected lengths become narrower for each case. Through these simulation results, we can see that our approximate confidence interval has a good performance not only in high dimensional and large sample settings, but also in large sample settings.

The asymptotic normality obtained by Corollary 3.3 is also demonstrated. Let

$$B_{N_1,N_2} = \frac{2\sqrt{N_1N_2/n}(L_1 - \Phi(-\tilde{\Delta}/2))}{\phi(-\tilde{\Delta}/2)},\ H_{p,N_1,N_2} = \frac{\sqrt{n}(L_1 - R_1)}{\sigma_2(\tilde{\Delta}^2)}.$$

Then Corollary 3.3 (Theorem 3.2) show that B_{N_1,N_2} (H_{p,N_1,N_2}) converges in distribution to standard normal distribution as $n \to \infty$ ($n, p \to \infty$). To check for asymptotic normality make B_{N_1,N_2} (H_{p,N_1,N_2}) vs standard normal Q-Q plot in Case A. The straight line $y = x$ represents where asymptotic normality holds. Figure 1 display the Q-Q plots of B_{N_1,N_2} in Case B, and Figure 2, 3 display the Q-Q plots of H_{p,N_1,N_2} in Case A. From figures, it is confirmed that CPMC has normality when sample size is large enough compared with the dimension.

§5. Conclusion

The performance of classification procedure is evaluated by its error probability which usually depends on unknown parameters. In practice, we considered the interval estimation for EPMC of improved linear discriminant rule. To derive an approximate confidence interval, we obtained the explicit expression of stochastic for two bilinear forms and two quadratic forms, and derived the asymptotic distribution for the studentized statistics of estimator of EPMC under the high dimensional and large sample frame work. Our approximate confidence interval not only has been established in high dimensional and large sample settings, but also has been established in large sample settings. Also, we confirmed that the superiority of our approximate confidence intervals have been verified in the sense of the coverage probability and expected length by using Monte Carlo simulation.

Appendix

A. Stochastic expression quadratic form

We present here the preliminary results about the quadratic form.

Lemma A. 1. *Let* $\boldsymbol{z} \sim \mathcal{N}_p(\boldsymbol{\nu}, I_p)$, $\boldsymbol{g} \sim \mathcal{N}_p(\mathbf{0}, I_p)$, $W \sim \mathcal{W}_p(n, I_p)$ *and* $\nu = \sqrt{\boldsymbol{\nu}'\boldsymbol{\nu}}$. *Assume that* $n - p + 1 > 0$ *and* $p > 2$. *Then, it holds that*

$$\text{(i) } \boldsymbol{z}'W^{-1}\boldsymbol{z} = \frac{(u_1+\nu)^2 + u_2^2 + \tilde{v}_1}{\tilde{v}_2},$$

$$\text{(ii) } \boldsymbol{z}'W^{-2}\boldsymbol{z} = \frac{(u_1+\nu)^2 + u_2^2 + \tilde{v}_1}{\tilde{v}_2^2}\left(1 + \frac{\tilde{v}_3}{\tilde{v}_4}\right),$$

$$\text{(iii) } \boldsymbol{\nu}'W^{-1}\boldsymbol{z} = \frac{\nu}{\tilde{v}_2}\left\{\nu + u_1 + u_2\left(\frac{\tilde{v}_3}{\tilde{v}_4}\right)^{\frac{1}{2}}\right\},$$

$$\text{(iv) } \boldsymbol{z}'W^{-1}\boldsymbol{g} = \frac{\sqrt{(u_1+\nu)^2 + u_2^2 + \tilde{v}_1}}{\tilde{v}_2}\left\{u_3 - u_4\left(\frac{\tilde{v}_3}{\tilde{v}_4}\right)^{\frac{1}{2}}\right\},$$

where

$$u_i \sim \mathcal{N}(0,1)\ (i=1,2,3,4),\ \tilde{v}_1 \sim \chi^2_{p-2},\ \tilde{v}_2 \sim \chi^2_{n-p+1},\ \tilde{v}_3 \sim \chi^2_{p-1},\ \tilde{v}_4 \sim \chi^2_{n-p+2},$$

and these variables are mutually independent. Here, χ^2_p means chi-square distribution with p degrees of freedom.

(Proof) The proof of assertions (i)-(iv) follows directly by applying the technique derived in Lemma 1, in Yamada et al. (2015).

B. Derivation of (2.3)

By using Lemma A.1, U can be rewritten as

(A. 1)

$$\begin{aligned} U &= -\frac{\tilde{\Delta}^2 n}{2\tilde{v}_2} + \left\{ \frac{np(N_1-N_2)}{2N_1N_2(n-p-1)} - \frac{n(N_1-N_2)}{2N_1N_2\tilde{v}_2}\left(u_1^2+u_2^2+\tilde{v}_1\right) \right\} \\ &\quad + \frac{nu_3}{\tilde{v}_2\sqrt{N_1N_2}}\sqrt{\left(\tilde{\Delta}\sqrt{\frac{N_1N_2}{n+2}}+u_1\right)^2+u_2^2+\tilde{v}_1} \\ &\quad - \frac{nu_4}{\tilde{v}_2\sqrt{N_1N_2}}\sqrt{\frac{\tilde{v}_3}{\tilde{v}_4}}\sqrt{\left(\tilde{\Delta}\sqrt{\frac{N_1N_2}{n+2}}+u_1\right)^2+u_2^2+\tilde{v}_1} \\ &\quad - \frac{\tilde{\Delta} n}{\tilde{v}_2\sqrt{(n+2)N_1N_2}}\left(N_1u_1+N_2u_2\sqrt{\frac{\tilde{v}_3}{\tilde{v}_4}}\right). \end{aligned}$$

By using above expression, we calculate the expectation of U as

(A. 2)

$$\begin{aligned} \mathrm{E}[U] &= -\frac{n}{2(n-p-1)}\tilde{\Delta}^2 = -\frac{\tilde{\Delta}^2}{2(1-\tilde{c})}\left(1-\frac{1}{n(1-\tilde{c})}\right)^{-1} \\ &= -\frac{\Delta^2}{2(1-c)}\left(1+\frac{1}{n(1-c)}\right)+o(n^{-1}). \end{aligned}$$

The expectation of U^2 is obtained by calculating the second moment of each term in (A.1). The second moment of each term in (A.1) is calculated as follows:

$$\begin{aligned}
&\mathrm{E}\left[\left\{-\frac{\tilde{\Delta}^2 n}{2\tilde{v}_2}+\frac{np(N_1-N_2)}{2N_1N_2(n-p-1)}-\frac{n(N_1-N_2)}{2N_1N_2\tilde{v}_2}\left(u_1^2+u_2^2+\tilde{v}_1\right)\right\}^2\right]\\
&=\frac{n^2}{4(n-p-3)(n-p-1)}\tilde{\Delta}^4+\frac{n^2p(N_1-N_2)}{N_1N_2(n-p-3)(n-p-1)^2}\tilde{\Delta}^2\\
&\quad+\frac{(n-1)n^2p(N_1-N_2)^2}{2N_1^2N_2^2(n-p-3)(n-p-1)^2}\\
&=\frac{\Delta^4}{4(1-c)^2}\left(1+\frac{4}{n(1-c)}\right)+\frac{c\Delta^2(\gamma_1-\gamma_2)}{n(1-c)^3\gamma_1\gamma_2}+\frac{c(\gamma_1-\gamma_2)^2}{2n(1-c)^3\gamma_1^2\gamma_2^2}+o(n^{-1}),
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\left[\left\{\frac{nu_3}{\tilde{v}_2\sqrt{N_1N_2}}\sqrt{\left(\tilde{\Delta}\sqrt{\frac{N_1N_2}{n+2}}+u_1\right)^2+u_2^2+\tilde{v}_1}\right\}^2\right]\\
&=\frac{n^2}{(n+2)(n-p-3)(n-p-1)}\tilde{\Delta}^2+\frac{n^2p}{N_1N_2(n-p-3)(n-p-1)}\\
&=\frac{1}{n(1-c)^2}\Delta^2+\frac{c}{n(1-c)^2\gamma_1\gamma_2}+o(n^{-1}),
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\left[\left\{-\frac{nu_4}{\tilde{v}_2\sqrt{N_1N_2}}\sqrt{\frac{\tilde{v}_3}{\tilde{v}_4}}\sqrt{\left(\tilde{\Delta}\sqrt{\frac{N_1N_2}{n+2}}+u_1\right)^2+u_2^2+\tilde{v}_1}\right\}^2\right]\\
&=\frac{n^2(p-1)}{(n+2)(n-p-3)(n-p-1)(n-p)}\tilde{\Delta}^2\\
&\quad+\frac{n^2(p-1)p}{N_1N_2(n-p-3)(n-p-1)(n-p)}\\
&=\frac{c}{n(1-c)^3}\Delta^2+\frac{c^2}{n(1-c)^3\gamma_1\gamma_2}+o(n^{-1}),
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\left[\left\{-\frac{\tilde{\Delta}n}{\tilde{v}_2\sqrt{(n+2)N_1N_2}}\left(N_1u_1+N_2u_2\sqrt{\frac{\tilde{v}_3}{\tilde{v}_4}}\right)\right\}^2\right]\\
&=\frac{n^2\left\{N_1^2(n-p)+N_2^2(p-1)\right\}}{(n+2)N_1N_2(n-p-3)(n-p-1)(n-p)}\tilde{\Delta}^2\\
&=\frac{\left(\gamma_1^2-\gamma_1^2c+\gamma_2^2c\right)}{n\gamma_1\gamma_2(1-c)^3}\Delta^2+o(n^{-1}).
\end{aligned}$$

Summarizing these results, we obtain that

$$(\text{A. }3)\qquad \mathrm{E}\left[U^2\right] = \frac{\Delta^4}{4(1-c)^2}\left(1+\frac{4}{n(1-c)}\right)+\frac{\Delta^2}{n(1-c)^3\gamma_2} + \frac{c\,(\gamma_1-\gamma_2)^2}{2n(1-c)^3\gamma_1^2\gamma_2^2}+\frac{c}{n(1-c)^3\gamma_1\gamma_2}+o(n^{-1}).$$

From (A.2) and (A.3), we obtain that

$$\begin{aligned}\mathrm{E}\left[(U-U_0)^2\right] &= \mathrm{E}[U^2]-2U_0\mathrm{E}[U]+U_0^2\\ &= \frac{1}{2n(1-c)^3}\left\{\Delta^4+\frac{2}{\gamma_2}\left(\frac{c}{\gamma_1}+\Delta^2\right)+\frac{c(\gamma_1-\gamma_2)^2}{\gamma_1^2\gamma_2^2}\right\}+o(n^{-1}).\end{aligned}$$

C. Derivation of (2.4)

By using Lemma A.1, V can be rewritten as

$$(\text{A. }4)\quad V = \left\{\frac{\tilde{\Delta}^2n^2}{\tilde{v}_2^2}\left(1+\frac{\tilde{v}_3}{\tilde{v}_4}\right)+\frac{n^2(n+2)}{N_1N_2\tilde{v}_2^2}\left(1+\frac{\tilde{v}_3}{\tilde{v}_4}\right)\left(u_1^2+u_2^2+\tilde{v}_1\right)\right\} +\frac{2\tilde{\Delta}n^2\sqrt{n+2}}{\tilde{v}_2^2\sqrt{N_1N_2}}\left(1+\frac{\tilde{v}_3}{\tilde{v}_4}\right)u_1.$$

By using above expression, we calculate the expectation of V as

(A. 5)

$$\begin{aligned}\mathrm{E}[V] &= \mathrm{E}\left[\frac{\tilde{\Delta}^2n^2}{\tilde{v}_2^2}\left(1+\frac{\tilde{v}_3}{\tilde{v}_4}\right)+\frac{n^2(n+2)}{N_1N_2\tilde{v}_2^2}\left(1+\frac{\tilde{v}_3}{\tilde{v}_4}\right)\left(u_1^2+u_2^2+\tilde{v}_1\right)\right]\\ &= \frac{(n-1)n^2}{(n-p-3)(n-p-1)(n-p)}\tilde{\Delta}^2\\ &\quad+\frac{(n-1)(n+2)n^2p}{N_1N_2(n-p-3)(n-p-1)(n-p)}\\ &= \frac{\Delta^2}{(1-c)^3}\left\{1+\frac{1}{n}\left(\frac{4}{1-c}-1\right)\right\}\\ &\quad+\frac{c}{(1-c)^3\gamma_1\gamma_2}\left\{1+\frac{1}{n}\left(\frac{4}{1-c}+1\right)\right\}+o(n^{-1}).\end{aligned}$$

The expectation of V^2 is obtained by calculating the second moment of each term in (A.4). The second moment of each term in (A.4) is calculated as follows:

(A. 6)
$$\mathrm{E}\left[\left\{\frac{\tilde{\Delta}^2 n^2}{\tilde{v}_2^2}\left(1+\frac{\tilde{v}_3}{\tilde{v}_4}\right)+\frac{n^2(n+2)}{N_1 N_2 \tilde{v}_2^2}\left(1+\frac{\tilde{v}_3}{\tilde{v}_4}\right)\left(u_1^2+u_2^2+\tilde{v}_1\right)\right\}^2\right]$$
$$=\frac{n^4\left\{\left(\tilde{\Delta}^2 N_1 N_2+(n+2)p\right)^2+2(n+2)^2 p\right\}}{N_1^2 N_2^2}\mathrm{E}\left[\frac{p^2-1}{\tilde{v}_2^4\tilde{v}_4^2}+\frac{2p-2}{\tilde{v}_2^4\tilde{v}_4}+\frac{1}{\tilde{v}_2^4}\right],$$

(A. 7)
$$\mathrm{E}\left[\left\{\frac{2\tilde{\Delta} n^2\sqrt{n+2}}{\tilde{v}_2^2\sqrt{N_1 N_2}}\left(1+\frac{\tilde{v}_3}{\tilde{v}_4}\right)u_1\right\}^2\right]$$
$$=\frac{4\tilde{\Delta}^2 n^4(n+2)}{N_1 N_2}\mathrm{E}\left[\frac{p^2-1}{\tilde{v}_2^4\tilde{v}_4^2}+\frac{2p-2}{\tilde{v}_2^4\tilde{v}_4}+\frac{1}{\tilde{v}_2^4}\right].$$

We note that

$$\mathrm{E}\left[\frac{1}{\tilde{v}_2^4}\right]=\frac{1}{(1-c)^4 n^4}+\frac{16}{(1-c)^5 n^5}+o(n^{-5}),$$
$$\mathrm{E}\left[\frac{1}{\tilde{v}_4^2}\right]=\frac{1}{(1-c)^2 n^2}+\frac{2}{(1-c)^3 n^3}+o(n^{-3}),$$
$$\mathrm{E}\left[\frac{1}{\tilde{v}_4}\right]=\frac{1}{(1-\tilde{c})n}.$$

Thus we obtain that

$$\text{(A. 8)}\quad \mathrm{E}\left[\frac{p^2-1}{\tilde{v}_2^4\tilde{v}_4^2}+\frac{2p-2}{\tilde{v}_2^4\tilde{v}_4}+\frac{1}{\tilde{v}_2^4}\right]=\frac{1}{(1-c)^6 n^4}+\frac{2(2c+7)}{(1-c)^7 n^5}+o(n^{-5}).$$

Substitute (A.8) into (A.6) and (A.7), we obtain that

$$\text{(A. 9)}\quad \mathrm{E}\left[V^2\right] = \frac{1}{(1-c)^6}\Delta^4+\frac{2c}{(1-c)^6\gamma_1\gamma_2}\Delta^2+\frac{c^2}{(1-c)^6\gamma_1^2\gamma_2^2}$$
$$+\frac{1}{n}\left(\frac{2(2c+7)\Delta^4}{(1-c)^7}+\frac{4\{c(c+7)+1\}\Delta^2}{(1-c)^7\gamma_1\gamma_2}\right.$$
$$\left.+\frac{2c(8c+1)}{(1-c)^7\gamma_1^2\gamma_2^2}\right)+o(n^{-1}).$$

From (A.5) and (A.9), we obtain that

$$\begin{aligned}
\mathrm{E}\left[(V-V_0)^2\right] &= \mathrm{E}[V^2]-2V_0\mathrm{E}[V]+V_0^2\\
&= \frac{2}{n(1-c)^7}\left[(c+4)\Delta^4+\frac{2\left\{(c+1)^2+c\right\}}{\gamma_1\gamma_2}\Delta^2\right.\\
&\quad \left.+\frac{c\left\{(c+1)^2+c\right\}}{\gamma_1^2\gamma_2^2}\right]+o(n^{-1}).
\end{aligned}$$

D. Derivation of (2.8)

By using Lemma A.1, $\hat{\Delta}^2$ can be rewritten as

$$\hat{\Delta}^2=\frac{(n-p-1)(n+2)}{N_1N_2}\frac{(u_1+\tau)^2+u_2^2+\tilde{v}_1}{\tilde{v}_2}-\frac{(n+2)p}{N_1N_2} \tag{A. 10}$$

By using above expression, we calculate the expectation of $\hat{\Delta}^2$ as

(A. 11)

$$\begin{aligned}
\mathrm{E}[\hat{\Delta}^2] &= \frac{(n-p-1)(n+2)}{N_1N_2}\mathrm{E}\left[\frac{(u_1+\tau)^2+u_2^2+\tilde{v}_1}{\tilde{v}_2}\right]-\frac{(n+2)p}{N_1N_2}\\
&= \frac{(n-p-1)(n+2)}{N_1N_2}\frac{N_1N_2\Delta^2+(n+2)p}{(n+2)(n-p-1)}-\frac{(n+2)p}{N_1N_2}\\
&= \tilde{\Delta}^2.
\end{aligned}$$

Also, we calculate the second moment of $\hat{\Delta}^2$ as

(A. 12)

$$\begin{aligned}
\mathrm{E}[\hat{\Delta}^4] &= \frac{(n-p-1)^2(n+2)^2}{N_1^2N_2^2}\mathrm{E}\left[\left(\frac{(u_1+\tau)^2+u_2^2+\tilde{v}_1}{\tilde{v}_2}\right)^2\right]\\
&\quad -\frac{2(n+2)^2(n-p-1)p}{N_1^2N_2^2}\mathrm{E}\left[\frac{(u_1+\tau)^2+u_2^2+\tilde{v}_1}{\tilde{v}_2}\right]+\frac{(n+2)^2p^2}{N_1^2N_2^2}.
\end{aligned}$$

The expected term in (A.12) can be calculated as

(A. 13)

$$\mathrm{E}\left[\frac{(u_1+\tau)^2+u_2^2+\tilde{v}_1}{\tilde{v}_2}\right] = \frac{N_1N_2\tilde{\Delta}^2+(n+2)p}{(n+2)(n-p-1)},$$

(A. 14)

$$\mathrm{E}\left[\left(\frac{(u_1+\tau)^2+u_2^2+\tilde{v}_1}{\tilde{v}_2}\right)^2\right] = \frac{1}{(n-p-3)(n-p-1)(n+2)^2} \{N_1^2N_2^2\tilde{\Delta}^4+2(n+2)(p+2)N_1N_2\tilde{\Delta}^2 +(n+2)^2p(p+2)\}.$$

Substitute (A.13) and (A.14) into (A.12), we obtain that

(A. 15)

$$\mathrm{E}[\hat{\Delta}^4] = \left(1+\frac{2}{n-p-3}\right)\tilde{\Delta}^4+\frac{4(n-1)(n+2)}{(n-p-3)N_1N_2}\tilde{\Delta}^2+\frac{2(n+2)^2p(n-1)}{N_1^2N_2^2(n-p-3)}.$$

From (A.11) and (A.15), we obtain that

$$\begin{aligned}\mathrm{E}[(\widehat{\Delta}^2-\Delta^2)^2] &= \mathrm{E}[\widehat{\Delta}^4]-2\Delta^2\mathrm{E}[\widehat{\Delta}^2]+\Delta^4\\ &= \frac{1}{n(1-c)}\left(2\Delta^4+\frac{4\Delta^2}{\gamma_1\gamma_2}+\frac{2c}{\gamma_1^2\gamma_2^2}\right)+o(n^{-1}).\end{aligned}$$

E. Derivation of (3.10)

From Lemma A.1, we note that

$$0<\frac{(V-V_0)^2}{V}<\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}(V-V_0)^2 \ a.s.$$

So, we consider to evaluate

(A. 16)

$$\mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}(V-V_0)^2\right] = \mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}V^2\right]-2V_0\mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}V\right] +V_0^2\mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}\right].$$

The each term on right hand side in (A.16) is evaluated as

$$
\begin{aligned}
\text{(A. 17)}\quad & \mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}V^2\right] \\
&= \mathrm{E}\left[\left(\frac{n(\tilde{v}_3+\tilde{v}_4)\left\{\tilde{\Delta}^2N_1N_2+(n+2)\left(u_1^2+u_2^2+\tilde{v}_1\right)\right\}}{\sqrt{(n+2)N_1N_2}\sqrt{\tilde{v}_1}\tilde{v}_2\tilde{v}_4}\right.\right. \\
&\quad \left.\left.+\frac{2n\tilde{\Delta}u_1(\tilde{v}_3+\tilde{v}_4)}{\sqrt{\tilde{v}_1}\tilde{v}_2\tilde{v}_4}\right)^2\right] \\
&= \frac{(n-3)(n-1)n^2N_1N_2\tilde{\Delta}^4}{(n+2)(n-p-3)(n-p-2)(n-p-1)(n-p)(p-4)} \\
&\quad +\frac{2(n-3)(n-1)n^2p\tilde{\Delta}^2}{(n-p-3)(n-p-2)(n-p-1)(n-p)(p-4)} \\
&\quad +\frac{(n-3)(n-1)n^2(n+2)(p-2)p}{N_1N_2(n-p-3)(n-p-2)(n-p-1)(n-p)(p-4)} \\
&= \left\{\frac{\gamma_1\gamma_2}{(1-c)^4c}+\frac{2(3c^2-2c+2)\gamma_1\gamma_2}{(1-c)^5c^2n}\right\}\Delta^4 \\
&\quad +\left\{\frac{2}{(1-c)^4}+\frac{4(2c^2-c+2)}{(1-c)^5cn}\right\}\Delta^2 \\
&\quad +\frac{c}{(1-c)^4\gamma_1\gamma_2}+\frac{2\left(c^2+c+1\right)}{(1-c)^5n\gamma_1\gamma_2}+o(n^{-1}),
\end{aligned}
$$

and

$$
\begin{aligned}
\text{(A. 18)}\quad & \mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}V\right] \\
&= \mathrm{E}\left[\frac{(\tilde{v}_3+\tilde{v}_4)\left\{\tilde{\Delta}^2N_1N_2+(n+2)\left(u_1^2+u_2^2+\tilde{v}_1\right)\right\}}{(n+2)\tilde{v}_1\tilde{v}_4}\right. \\
&\quad \left.+\frac{2\sqrt{N_1N_2}\tilde{\Delta}u_1(\tilde{v}_3+\tilde{v}_4)}{\sqrt{n+2}\tilde{v}_1\tilde{v}_4}\right] \\
&= \frac{(n-1)N_1N_2}{(n+2)(n-p)(p-4)}\tilde{\Delta}^2+\frac{(n-1)(p-2)}{(n-p)(p-4)} \\
&= \left(\frac{\gamma_1\gamma_2}{(1-c)c}+\frac{(4-3c)\gamma_1\gamma_2}{(1-c)c^2n}\right)\Delta^2+\left(\frac{1}{1-c}+\frac{2-c}{(1-c)cn}\right)+o(n^{-1}).
\end{aligned}
$$

Combining (A.16)-(A.18), we obtain that

$$\mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}(V-V_0)^2\right] = \frac{2}{n}\left(\frac{(c+4)\gamma_1\gamma_2}{(1-c)^5c}\Delta^4 + \frac{2\{c(c+3)+1\}}{(1-c)^5c}\Delta^2 + \frac{c(c+3)+1}{(1-c)^5\gamma_1\gamma_2}\right) + o(n^{-1}).$$

F. Derivation of (3.13)

From Lemma A.1, it holds that

$$0 < \frac{(U-U_0)^2}{V} < \frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}(U-U_0)^2 \ a.s.$$

So, we consider to evaluate

(A. 19)

$$\mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}(U-U_0)^2\right] = \mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}U^2\right] - 2U_0\mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}U\right] + U_0^2\mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}\right].$$

We evaluate the first term on right hand side of (A.19).
The random variable $\sqrt{N_1N_2\tilde{v}_2^2/\{n^2(n+2)\tilde{v}_1\}}U$ can be rewritten as

(A. 20)

$$\left(\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}\right)^{1/2}U = \left\{\frac{(N_1-N_2)\tilde{v}_2}{2\sqrt{(n+2)N_1N_2}\sqrt{\tilde{v}_1}}\left(\frac{p}{n-p-1} - \frac{u_1^2+u_2^2+\tilde{v}_1}{\tilde{v}_2}\right) - \frac{\sqrt{N_1N_2}}{2(n+2)^{1/2}\sqrt{\tilde{v}_1}}\tilde{\Delta}^2\right\} + \frac{u_3}{\sqrt{n+2}\sqrt{\tilde{v}_1}}\sqrt{\left(\tilde{\Delta}\sqrt{\frac{N_1N_2}{n+2}}+u_1\right)^2+u_2^2+\tilde{v}_1} - \frac{u_4}{\sqrt{n+2}}\sqrt{\frac{\tilde{v}_3}{\tilde{v}_1\tilde{v}_4}}\sqrt{\left(\tilde{\Delta}\sqrt{\frac{N_1N_2}{n+2}}+u_1\right)^2+u_2^2+\tilde{v}_1} - \frac{\tilde{\Delta}}{(n+2)\sqrt{\tilde{v}_1}}\left(N_1u_1+N_2u_2\sqrt{\frac{\tilde{v}_3}{\tilde{v}_4}}\right).$$

The expectation of $(N_1N_2\tilde{v}_2^2)/\{(n+2)\tilde{v}_1\}U^2$ is obtained by calculating the second moment of each term on right hand side of (A.20). These second moments can be calculated as follows:

$$\text{(A. 21)}\quad \mathrm{E}\left[\left(\frac{(N_1-N_2)\tilde{v}_2\left(\frac{p}{n-p-1}-\frac{u_1^2+u_2^2+\tilde{v}_1}{\tilde{v}_2}\right)}{2\sqrt{(n+2)N_1N_2}\sqrt{\tilde{v}_1}}-\frac{\sqrt{N_1N_2}}{2(n+2)^{1/2}\sqrt{\tilde{v}_1}}\tilde{\Delta}^2\right)^2\right]$$
$$=\frac{N_1N_2}{4(n+2)(p-4)}\tilde{\Delta}^4-\frac{(n-1)\left(N_1-N_2\right)}{(n+2)(p-4)(n-p-1)}\tilde{\Delta}^2$$
$$+\frac{(n-1)(n-p+3)(N_1-N_2)^2p}{2(n+2)(n-p-1)^2N_1N_2(p-4)},$$

$$\text{(A. 22)}\quad \mathrm{E}\left[\frac{u_3^2}{(n+2)\tilde{v}_1}\left\{\left(\tilde{\Delta}\sqrt{\frac{N_1N_2}{n+2}}+u_1\right)^2+u_2^2+\tilde{v}_1\right\}\right]$$
$$=\frac{N_1N_2}{(n+2)^2(p-4)}\tilde{\Delta}^2+\frac{p-2}{(n+2)(p-4)}$$
$$=\frac{(n+2)N_1-N_1^2}{(n+2)^2(p-4)}\tilde{\Delta}^2+\frac{p-2}{(n+2)(p-4)},$$

$$\text{(A. 23)}\quad \mathrm{E}\left[\frac{u_4^2\tilde{v}_3}{(n+2)\tilde{v}_1\tilde{v}_4}\left\{\left(\tilde{\Delta}\sqrt{\frac{N_1N_2}{n+2}}+u_1\right)^2+u_2^2+\tilde{v}_1\right\}\right]$$
$$=\frac{N_1N_2(p-1)}{(n+2)^2(p-4)(n-p)}\tilde{\Delta}^2+\frac{(p-1)(p-2)}{(n+2)(p-4)(n-p)}$$
$$=\frac{\{(n+2)N_2-N_2^2\}(p-1)}{(n+2)^2(p-4)(n-p)}\tilde{\Delta}^2+\frac{(p-1)(p-2)}{(n+2)(p-4)(n-p)},$$

$$\text{(A. 24)}\quad \mathrm{E}\left[\frac{\tilde{\Delta}^2}{(n+2)^2\tilde{v}_1}\left(N_1u_1+N_2u_2\sqrt{\frac{\tilde{v}_3}{\tilde{v}_4}}\right)^2\right]=\frac{N_1^2(n-p)+N_2^2(p-1)}{(n+2)^2(p-4)(n-p)}\tilde{\Delta}^2.$$

From (A.21)-(A.24), we can obtain that

$$\text{(A. 25)}\quad \mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}U^2\right]$$
$$=\frac{N_1N_2}{4(n+2)(p-4)}\tilde{\Delta}^4+\frac{N_1p(p-n)+N_2\left\{(n-1)^2-p^2+p\right\}}{(n+2)(p-4)(n-p-1)(n-p)}\tilde{\Delta}^2$$
$$+\frac{n-1}{2(n+2)(p-4)}\left\{\frac{p(n-p+3)(N_1-N_2)^2}{N_1N_2(n-p-1)^2}+\frac{2(p-2)}{n-p}\right\}$$
$$=\frac{\Delta^4\gamma_1\gamma_2}{4c}+\frac{1}{n}\left[\frac{(2-c)\Delta^4\gamma_1\gamma_2}{2c^2}-\frac{\Delta^2\{c\gamma_1-(c+1)\gamma_2\}}{(1-c)c}\right.$$
$$\left.+\frac{\gamma_1^2+\gamma_2^2}{2(1-c)\gamma_1\gamma_2}\right]+o(n^{-1}).$$

Also, we have that

(A. 26)
$$
\begin{aligned}
&\mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}U\right] \\
&= -\frac{N_1N_2(n-p+1)}{2n(n+2)(p-4)}\tilde{\Delta}^2 + \frac{(N_1-N_2)(n-p+1)(n+p-1)}{n(n+2)(n-p-1)(p-4)} \\
&= -\frac{(1-c)\gamma_1\gamma_2}{2c}\Delta^2 + \frac{1}{n}\left[\frac{\{(5-2c)c-4\}\Delta^2\gamma_1\gamma_2}{2c^2}\right. \\
&\left. +\frac{(c+1)(\gamma_1-\gamma_2)}{c}\right] + o(n^{-1}),
\end{aligned}
$$

and

(A. 27)
$$
\begin{aligned}
&\mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}\right] \\
&= \frac{N_1N_2(n-p+1)(n-p+3)}{n^2(n+2)(p-4)} \\
&= \frac{(1-c)^2\gamma_1\gamma_2}{c} + \frac{2\{2-(1-c)c\}(1-c)\gamma_1\gamma_2}{c^2n} + o(n^{-1}).
\end{aligned}
$$

Combining (A.25)-(A.27), we obtain that

$$
\begin{aligned}
\mathrm{E}\left[\frac{N_1N_2\tilde{v}_2^2}{n^2(n+2)\tilde{v}_1}(U-U_0)^2\right] &= \frac{1}{2n(1-c)c\gamma_1\gamma_2}\left\{\Delta^4\gamma_1^2\gamma_2^2 + 2\Delta^2\gamma_1^2\gamma_2\right. \\
&\left. +c(\gamma_1^2+\gamma_2^2)\right\} + o(n^{-1}).
\end{aligned}
$$

Acknowledgments

The authors would like to express their gratitude to Professor Yasunori Fujikoshi for many valuable comments and discussions.

References

[1] Chung, H.-C. and Han, C.-P. (2009). Conditional confidence intervals for classification error rate. *Computational Statistics and Data Analysis*, **53**, 4358-4369.

[2] Fujikoshi, Y. (2000). Error bounds for asymptotic approximations of the linear discriminant function when the sample sizes and dimensionality are large. *Journal of Multivariate Analysis*, **73**, 1-17.

[3] Fujikoshi, Y. and Seo, T. (1998). Asymptotic approximations for EPMC's of the linear and the quadratic discriminant functions when the sample sizes and the dimension are large. *Random Operators and Stochastic Equations*, **6**, 269-280.

[4] Kubokawa, T., Hyodo, M. and Srivastava, M. S. (2013). Asymptotic expansion and estimation of EPMC for linear classification rules in high dimension. *Journal of Multivariate Analysis*, **115**, 496-515.

[5] Lachenbruch, P. A. (1968). On expected probabilities of misclassification in discriminant analysis, necessary sample size, and a relation with the multiple correlation coefficient. *Biometrics.*, **24**, 823-834.

[6] Lachenbruch, P. A. and Mickey, M. R. (1968). Estimation of error rates in discriminant analysis. *Technometrics*, **10**, 1-11.

[7] McLachlan, G. J. (1975). Confidence intervals for the conditional probability of misallocation in discriminant analysis. *Biometrics*, **31**, 161-167.

[8] Okamoto, M. (1963). An asymptotic expansion for the distribution of the linear discriminant function. *Annals of Mathematical Statistics*, **34**, 1286-1301.

[9] Okamoto, M. (1968). Correction to "An asymptotic expansion for the distribution of the linear discriminant function". *Annals of Mathematical Statistics*, **39**, 1358-1359.

[10] Siotani, M. (1982). Large sample approximations and asymptotic expansions of classification statistic. *Handbook of Statistics* 2 (P. R. Krishnaiah and L. N. Kanal, Eds.), North-Holland Publishing Company, 61-100.

[11] Yamada, T., Himeno, T. and Sakurai, T. (2015). Cut-off point of linear discriminant rule for large dimension. Technical Report, No.15-04, *Hiroshima statistical research group*, *Hiroshima University.*

Table 1. The coverage probabilities ($p = 100$)

$\alpha \setminus n+2$	$(N_1 : N_2)$	200	300	400
0.01	(1 : 1)	0.987	0.988	0.989
	(1 : 3)	0.987	0.987	0.988
	(3 : 1)	0.986	0.987	0.988
0.05	(1 : 1)	0.946	0.948	0.948
	(1 : 3)	0.947	0.947	0.947
	(3 : 1)	0.945	0.947	0.948
0.10	(1 : 1)	0.898	0.899	0.898
	(1 : 3)	0.899	0.897	0.898
	(3 : 1)	0.896	0.897	0.899

Table 2. The coverage probabilities ($p = 200$)

$\alpha \setminus n+2$	$(N_1 : N_2)$	400	600	800
0.01	(1 : 1)	0.988	0.989	0.989
	(1 : 3)	0.989	0.988	0.989
	(3 : 1)	0.988	0.989	0.989
0.05	(1 : 1)	0.948	0.949	0.949
	(1 : 3)	0.949	0.948	0.950
	(3 : 1)	0.949	0.949	0.949
0.10	(1 : 1)	0.899	0.899	0.900
	(1 : 3)	0.900	0.899	0.900
	(3 : 1)	0.900	0.899	0.900

Table 3. The coverage probabilities ($p = 5$)

$\alpha \setminus n+2$	$(N_1 : N_2)$	100	300	500
0.01	(1 : 1)	0.984	0.987	0.989
	(1 : 3)	0.983	0.987	0.989
	(3 : 1)	0.982	0.987	0.989
0.05	(1 : 1)	0.944	0.947	0.950
	(1 : 3)	0.943	0.947	0.948
	(3 : 1)	0.941	0.947	0.949
0.10	(1 : 1)	0.896	0.897	0.901
	(1 : 3)	0.893	0.897	0.900
	(3 : 1)	0.894	0.898	0.900

Table 4. The expected lengths ($p = 100$)

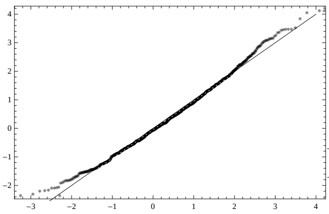
$n+2$		200		300		400	
α	$(N_1 : N_2)$	EL	EEL	EL	EEL	EL	EEL
0.01	$(1:1)$	0.170	0.173	0.122	0.124	0.097	0.098
	$(1:3)$	0.195	0.199	0.140	0.143	0.112	0.114
	$(3:1)$	0.195	0.198	0.140	0.142	0.111	0.114
0.05	$(1:1)$	0.129	0.131	0.093	0.094	0.074	0.075
	$(1:3)$	0.149	0.152	0.106	0.108	0.085	0.086
	$(3:1)$	0.149	0.151	0.106	0.107	0.085	0.085
0.10	$(1:1)$	0.109	0.110	0.078	0.078	0.062	0.062
	$(1:3)$	0.125	0.127	0.089	0.090	0.071	0.072
	$(3:1)$	0.125	0.127	0.089	0.089	0.071	0.071

Table 5. The expected lengths ($p = 200$)

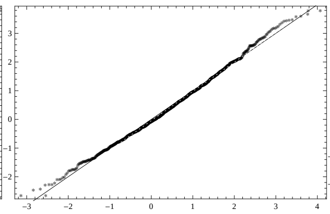
$n+2$		400		600		800	
α	$(N_1 : N_2)$	EL	EEL	EL	EEL	EL	EEL
0.01	$(1:1)$	0.120	0.122	0.086	0.086	0.069	0.069
	$(1:3)$	0.138	0.140	0.099	0.100	0.079	0.080
	$(3:1)$	0.139	0.141	0.099	0.101	0.079	0.079
0.05	$(1:1)$	0.092	0.092	0.065	0.066	0.052	0.053
	$(1:3)$	0.105	0.106	0.075	0.076	0.060	0.060
	$(3:1)$	0.105	0.107	0.075	0.076	0.060	0.060
0.10	$(1:1)$	0.077	0.077	0.055	0.055	0.044	0.044
	$(1:3)$	0.088	0.089	0.063	0.064	0.050	0.050
	$(3:1)$	0.088	0.089	0.063	0.064	0.050	0.051

Table 6. The expected lengths ($p = 5$)

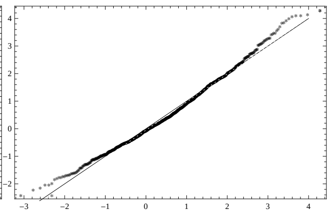
$n+2$		100		300		500	
α	$(N_1 : N_2)$	EL	EEL	EL	EEL	EL	EEL
0.01	$(1:1)$	0.151	0.163	0.103	0.107	0.083	0.085
	$(1:3)$	0.168	0.174	0.114	0.118	0.092	0.094
	$(3:1)$	0.169	0.174	0.114	0.119	0.092	0.094
0.05	$(1:1)$	0.115	0.119	0.078	0.080	0.063	0.064
	$(1:3)$	0.128	0.134	0.087	0.089	0.070	0.071
	$(3:1)$	0.128	0.134	0.087	0.089	0.070	0.071
0.10	$(1:1)$	0.097	0.099	0.066	0.067	0.053	0.054
	$(1:3)$	0.108	0.110	0.073	0.074	0.059	0.059
	$(3:1)$	0.108	0.110	0.073	0.074	0.059	0.059



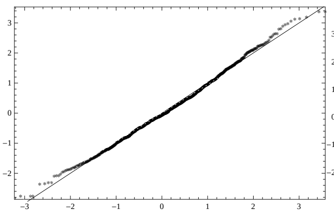
$(N_1, N_2) = (25, 75)$



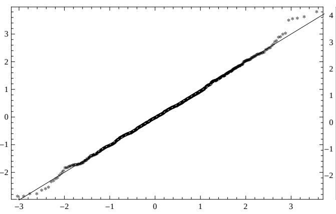
$(N_1, N_2) = (50, 50)$



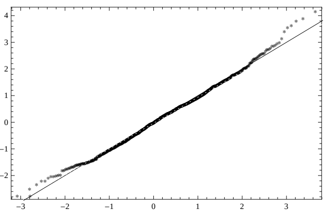
$(N_1, N_2) = (75, 25)$



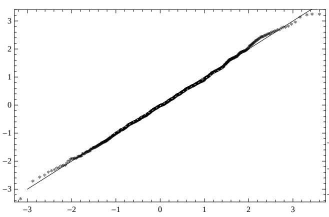
$(N_1, N_2) = (75, 225)$



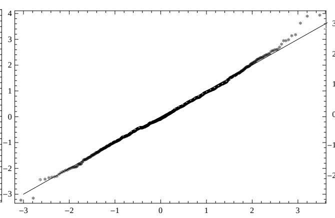
$(N_1, N_2) = (150, 150)$



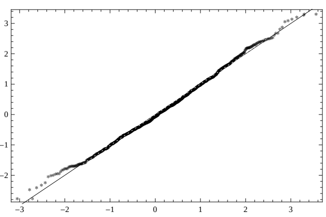
$(N_1, N_2) = (225, 75)$



$(N_1, N_2) = (125, 375)$

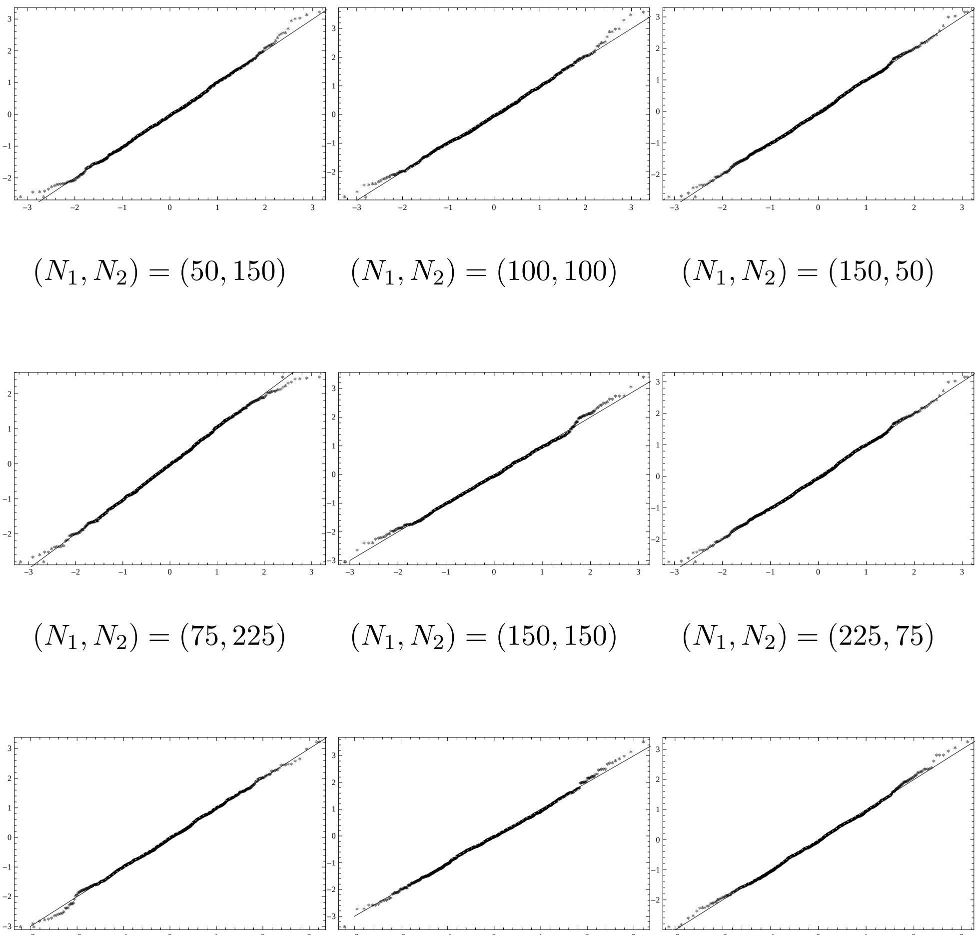


$(N_1, N_2) = (250, 250)$



$(N_1, N_2) = (375, 125)$

Figure 1. Q-Q plots of B_{N_1,N_2} for Case B

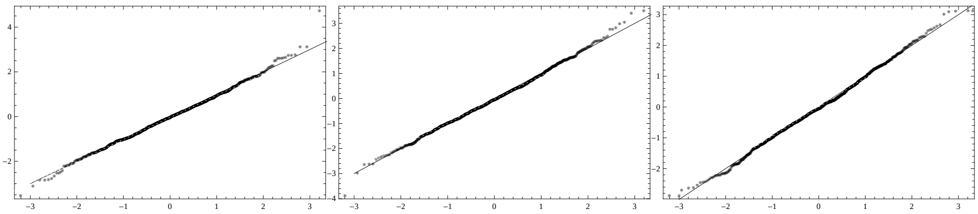


$(N_1, N_2) = (50, 150)$ $(N_1, N_2) = (100, 100)$ $(N_1, N_2) = (150, 50)$

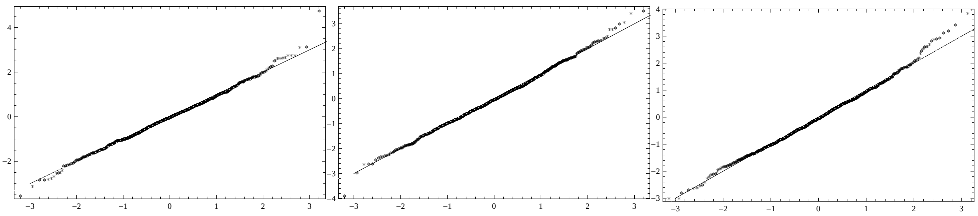
$(N_1, N_2) = (75, 225)$ $(N_1, N_2) = (150, 150)$ $(N_1, N_2) = (225, 75)$

$(N_1, N_2) = (100, 300)$ $(N_1, N_2) = (200, 200)$ $(N_1, N_2) = (300, 100)$

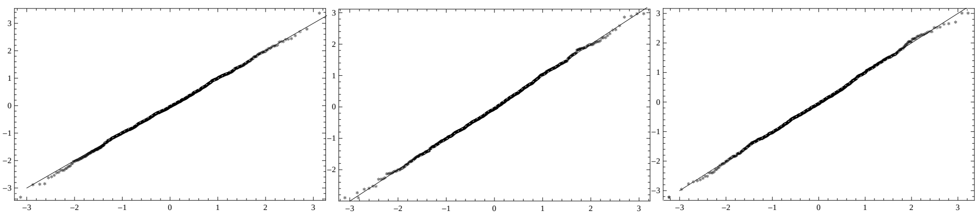
Figure 2. Q-Q plots of H_{p,N_1,N_2} in Case A ($p = 100$)



$(N_1, N_2) = (100, 300)$ $(N_1, N_2) = (200, 200)$ $(N_1, N_2) = (300, 100)$



$(N_1, N_2) = (150, 450)$ $(N_1, N_2) = (300, 300)$ $(N_1, N_2) = (450, 150)$



$(N_1, N_2) = (200, 600)$ $(N_1, N_2) = (400, 400)$ $(N_1, N_2) = (600, 200)$

Figure 3. Q-Q plots of H_{p,N_1,N_2} in Case A ($p = 200$)

Masashi Hyodo
Department of Mathematical Information Science, Tokyo University of Science
1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
E-mail: `caicmhy@gmail.com`

Tomohiro Mitani
Graduate School of Science, Tokyo University of Science
1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan

Tetsuto Himeno
Faculty of Science and Technology, Seikei University
3-3-1 Kichijoji-Kitamachi, Musashino-shi, Tokyo 180-8633, Japan

Takashi Seo
Department of Mathematical Information Science
Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan