Asymptotic Expansion of the Quadratic Discriminant Function for Large Dimension and Samples

Chieko Matsumoto and Hirofumi Wakaki

Abstract. For the classification problem between two normal populations with different covariance matrix, we consider the quadratic discriminant function (QDF shortly) and the sample quadratic discriminant function (SQDF shortly). Our purpose of this paper is to derive an asymptotic expansion of misclassification probabilities for QDF and SQDF when both the sample sizes and the dimension are large.

1. Introduction

Consider the problem of classifying an observation vector $\boldsymbol{x}$ into one of two normal populations Π_1:$\mathrm{N}_p(\boldsymbol{\mu}_1, \Sigma_1)$ and Π_2:$\mathrm{N}_p(\boldsymbol{\mu}_2, \Sigma_2)$. If the mean vector $\boldsymbol{\mu}_i$ and the covariance matrix Σ_i are unknown, suppose that we have a training sample of $\boldsymbol{x}_{11}, \cdots, \boldsymbol{x}_{1N_1}$ from Π_1 and another independent training sample of $\boldsymbol{x}_{21}, \cdots, \boldsymbol{x}_{2N_2}$ from Π_2. Let $\bar{\boldsymbol{x}}_i$ and S_i $(i = 1, 2)$ be the sample mean and the sample covariance matrix given by

$$\bar{\boldsymbol{x}}_i = \frac{1}{N_i}\sum_{j=1}^{N_i}\boldsymbol{x}_{ij}, \qquad S_i = \frac{1}{n_i}\sum_{j=1}^{N_i}(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)',$$

respectively, where $n_i = N_i - 1$.

2000 *Mathematics Subject Classification.* primary 62E20, secondly 62H30.
Key words and phrases. Asymptotic expansion, Edgeworth expansion, Large dimension and sample sizes, Misclassification probability, Quadratic discriminant function, Sample quadratic discriminant function.

We define a discriminant function by

$$\begin{aligned} Q(\boldsymbol{x};\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\Sigma_1,\Sigma_2) &= (\boldsymbol{x}-\boldsymbol{\mu}_1)'\Sigma_1^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_1)-(\boldsymbol{x}-\boldsymbol{\mu}_2)'\Sigma_2^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_2) \\ &\quad +\log|\Sigma_1\Sigma_2^{-1}|, \end{aligned} \tag{1.1}$$

if the parameters $\boldsymbol{\mu}_i$ and Σ_i are known. When the parameters are unknown, we define a discriminant function by

$$\begin{aligned} Q(\boldsymbol{x};\bar{\boldsymbol{x}}_1,\bar{\boldsymbol{x}}_2,S_1,S_2) &= (\boldsymbol{x}-\bar{\boldsymbol{x}}_1)'S_1^{-1}(\boldsymbol{x}-\bar{\boldsymbol{x}}_1)-(\boldsymbol{x}-\bar{\boldsymbol{x}}_2)'S_2^{-1}(\boldsymbol{x}-\bar{\boldsymbol{x}}_2) \\ &\quad +\log|S_1S_2^{-1}|, \end{aligned} \tag{1.2}$$

using $\bar{\boldsymbol{x}}_i$ and S_i replaced by $\boldsymbol{\mu}_i$ and Σ_i in (1.1). We usually call (1.1) as the quadratic discriminant function (QDF shortly) and (1.2) as the sample quadratic discriminant function (SQDF shortly), respectively.

For the most part, the exact distributions of discriminant functions are too complicated to handle. Then, we often use asymptotic approximations instead of the exact distributions.

If we consider $\Sigma_1=\Sigma_2$, many results about asymptotic approximations were obtained. (For the discriminant rules when $\Sigma_1=\Sigma_2$, for example, W-rule and Z-rule, see Siotani, Hayakawa and Fujikoshi (1985), e.g.) When only the sample sizes are large, Okamoto (1963,1968) derived an asymptotic expansion formula for the W-rule. For the Z-rule, Memon and Okamoto (1971) derived an asymptotic expansion formula. It is known that the accuracy of these formulas depends on the dimension and the Mahalanobis distance between the populations. Therefore, when the dimension is large, the approximations by these expansion formulas are poor.

Recently Saranadasa (1993) obtained the limiting distribution function of the Z-rule when both the sample sizes and the dimension are large. Fujikoshi and Seo (1998) also derived a limiting distribution in the class of discriminant function included both the W-rule and the Z-rule, and their numerical experiments show that the asymptotic expansion formula gives a good approximation even for the small dimension. Moreover, Tonda and Wakaki (2003) derived an asymptotic expansion of misclassification probability for the W-rule. Matsumoto (2004) derived an asymptotic expansion of misclassifucation probability in the class included the the W-rule and the Z-rule.

In the case of proportional covariance matrices, Wakaki (1990) derived an asymptotic expansion of the SQDF when only the sample sizes are large. However, when there are no restrictions about the covariance matrices, we cannot find any asymptotic expansion formulas of the distribution function for the QDF and the SQDF. One of the reasons is the complexity of the limiting distribution of the SQDF, which is distributed as the weight sum of independent noncentral chi-square distribution. When both the sample sizes and the dimension are large, we can apply the central limiting theorem for the weight sum. Our purpose of this paper is to derive an asymptotic expansion of the distribution function of the SQDF when both the sample sizes and the dimension are large. In Section 2 we derive the asymptotic expansions of misclassification probability for the QDF and the SQDF. In Section 3, some numerical experiments are carried out to examine the performance of the derived asymptotic approximations and to compare with the previous results. And for the Appendix we show the parameters of derived result.

2. Asymptotic expansion

We may derive the asymptotic approximations of the misclassification probability of the QDF and the SQDF using the Edgeworth expansion for large dimension and sample sizes. Then we assume that

$$\liminf_{n\to\infty} \frac{p}{n_i} > 0, \quad n_i > p, \tag{2.1}$$

and define

$$\nu_i = \frac{p}{n_i}, \tag{2.2}$$

when the sample size N_i and the dimension p are sufficiently large.

2.1. Asymptotic expansion of the misclassification probability

The following sections, we assume that an observation vector $\boldsymbol{x}$ comes from Π_1.

We consider the formal asymptotic expansion of the distribution of the sum of the random variables which are independent but aren't distributed identically. Let $X_1, X_2, \cdots, X_n$ be mutually independent random variables having $\mathrm{E}(X_j) = 0$ and $\mathrm{E}(|X_j|^k) < \infty$ for each j, with integral $k \geq 3$, and let $\mathrm{E}(X_j^2) = \sigma_j^2$ and $\sum_{j=1}^n \sigma_j^2 > 0$. Let F_n be the distribution function of $\sum_{j=1}^n X_j / \sum_{j=1}^n \sigma_j^2$. Then, the approximation of $F_n(x)$ is given by

$$\begin{aligned} F_n(x) \sim\ & \Phi(x) \\ & -\phi(x)\left\{(x^2-1)\frac{\kappa_3}{6\sqrt{v_p}^3} + (x^3-3x)\frac{\kappa_4}{24v_p^2} + (x^5-10x^3+15x)\frac{\kappa_3^2}{72v_p^3}\right\} \\ & +o(n^{-1}), \end{aligned} \tag{2.3}$$

where v_n is the sum of the variance and κ_i is the sum of the i-th cumulant of X_j. As later mentioned, the QDF and the SQDF are shown as the sum of (conditional) independent random vectors. Hence, we use (2.3) to calculate the asymptotic expansions of the misclassification probability of the QDF and the SQDF.

For the validity of (2.3), we show the following theorem.

Theorem 2.1. (Bhattacharya and Rao(1976)). *Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n, \cdots$ be a sequence of independent random vectors in $\mathbf{R}^k$ with zero means and positive definite covariance matrices* $\mathrm{V}[\boldsymbol{x}_j]$. *Assume that*

- $\mathrm{E}[|\boldsymbol{x}_j|^s] < \infty$ *(for some $s > 3$);*
- $\liminf_{n\to\infty} \frac{1}{n} B_n^2 > 0 \quad B_n^2 = \sum_{j=1}^{n} \mathrm{V}[\boldsymbol{x}_j]$*;*
- $\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \int_{|\boldsymbol{x}_j| > \epsilon\sqrt{n}} |\boldsymbol{x}_j|^s d(P_{\boldsymbol{x}_j}) = 0$ *for any $\epsilon > 0$; and*
- *$g_n(t)$:the characteristic function of $\boldsymbol{x}_n$, $\overline{\lim}_{n\to\infty} \sup_{|t|>b} |g_n(t)| < 1$ for any $b > 0$.*

In this time, the distribution function of $B_n^{-1} \sum_{j=1}^{n} \boldsymbol{x}_j$ can product Edgeworth expansion under $O(n^{-(s-2)/2})$. (The formal expansion with cumulant has the validity.)

However, when the sample sizes and the dimension are large, we have to improve Theorem 2.1. Nevertheless, for a little improvement, we'll check the validity of (2.3) for the QDF and the SQDF. If (2.3) has the validity, the misclassification probability for the QDF and the SQDF is asymptotically given by

$$\begin{aligned}
\Pr(2|1) &= \mathrm{P}[Q(\boldsymbol{x}) > -2\log c|\Pi_1] \\
&\sim \Phi(-a) + \phi(a) \left\{ (a^2 - 1) \frac{\kappa_3[Q(\boldsymbol{x})]}{6\mathrm{Var}[Q(\boldsymbol{x})]^{3/2}} + (a^3 - 3a) \frac{\kappa_4[Q(\boldsymbol{x})]}{24\mathrm{Var}[Q(\boldsymbol{x})]^2} \right. \\
&\left. + (a^5 - 10x^3 + 15a) \frac{\kappa_3^2[Q(\boldsymbol{x})]}{72\mathrm{Var}[Q(\boldsymbol{x})]^3} \right\} + o(p^{-1}), \qquad (2.4)
\end{aligned}$$

using (2.3), where $Q(\boldsymbol{x})$ is shown as the QDF or the SQDF,

$$a = \frac{-2\log c - \mathrm{E}[Q(\boldsymbol{x})]}{\mathrm{Var}[Q(\boldsymbol{x})]^{1/2}},$$

and $\kappa_i[Q(\boldsymbol{x})]$ is the i-th cumulant of $Q(\boldsymbol{x})$. Therefore, we have to calculate the expectation, the variance, the 3rd and the 4th cumulant of the QDF and the SQDF. For applying these results to (2.4), we may obtain the asymptotic expansions of misclassification probability.

2.2. Asymptotic expansion of misclassification probability for the quadratic discriminant function

First, we consider the asymptotic expansion of misclassification probability for the QDF (1.1) up to $p^{-1/2}$. In this case, we assume that the means and the covariance matrices are known. For making the discussion simply, we standardize $\boldsymbol{x}$ and $\boldsymbol{x}_{ij}$ as

$$\boldsymbol{y} = L(\boldsymbol{x} - \mu_1), \quad \boldsymbol{y}_{ij} = L(\boldsymbol{x}_{ij} - \mu_1) \ \ (i = 1, 2; j = 1, \cdots, N_i), \tag{2.5}$$

where L is the nonsingular matrix such that $L'\Sigma_1 L = I_p$ and $L'\Sigma_2 L = \Lambda_p = diag(\lambda_1, \cdots, \lambda_p)$, with λ_i's being the roots of $|\Sigma_2 - \lambda\Sigma_1| = 0$. Then, the quadratic discriminant function (1.1) is expressed as

$$\begin{aligned} Q(\boldsymbol{y}; \mathbf{0}, \eta, I_p, \Lambda_p) &= \boldsymbol{y}' I_p \boldsymbol{y} - (\boldsymbol{y} - \eta)'\Lambda_p^{-1}(\boldsymbol{y} - \eta) + \sum_{i=1}^{p} \log \lambda_i^{-1} \\ &= \sum_{j=1}^{p} \{y_j^2 - \lambda_j^{-1}(y_j - \eta_j)^2 + \log \lambda_i^{-1}\}, \end{aligned} \tag{2.6}$$

where

$$\boldsymbol{y} = (y_1, \cdots, y_p)' \sim \mathrm{N}_p(\mathbf{0}, I_p), \quad \eta = L(\mu_2 - \mu_1).$$

(2.6) is the form of the sum of the independent random vectors. Then, the expectation, variance and 3rd cumulant of statistic $T = \boldsymbol{y}'\boldsymbol{y} - (\boldsymbol{y} - \eta)'\Lambda_p^{-1}(\boldsymbol{y} - \eta)$ are calculated respectively as

$$\begin{aligned} &\mathrm{E}[T] = p - (\mathrm{tr}\Lambda_p^{-1} + \eta'\Lambda_p^{-1}\eta), \\ &\mathrm{Var}[T] = 2p - 4\mathrm{tr}\Lambda_p^{-1} + 2\mathrm{tr}\Lambda_p^{-2} + 4\eta'\Lambda_p^{-2}\eta, \\ &\kappa_3[T] = 8p + 24\eta'(\Lambda_p^{-2} - \Lambda_p^{-3})\eta - 24\mathrm{tr}\Lambda_p^{-1} + 24\mathrm{tr}\Lambda_p^{-2} - 8\mathrm{tr}\Lambda_p^{-3}. \end{aligned}$$

For the validity, we consider the following lemma. (About the proof, see Theorem 2.1.)

LEMMA 2.1. *Denote that* $e_0 := E[Q(\boldsymbol{y}; \mathbf{0}, \eta, I_p, \Lambda_p)]$, $v_0 := \mathrm{Var}[Q(\boldsymbol{y}; \mathbf{0}, \eta, I_p, \Lambda_p)]$ *and* $\kappa := \kappa_3[Q(\boldsymbol{y}; \mathbf{0}, \eta, I_p, \Lambda_p)]$. *If*

$$\underline{\lim}_{p\to\infty} \frac{1}{p}\left(p - 2\mathrm{tr}\Lambda_p^{-1} + \mathrm{tr}\Lambda_p^{-2} + 2\eta'\Lambda_p^{-2}\eta\right) > 0,$$

$$\frac{1}{p}\sum_{j=1}^{p} \lambda_j^3 |\eta_j|^k = O(1) \qquad k = 0, 1, \cdots, 6,$$

then uniformly for x,

$$\mathrm{P}\Big(\frac{Q(\boldsymbol{y}; \mathbf{0}, \eta, I_p, \Lambda_p) - e_0}{v_0^{1/2}} \le x\Big) = \Phi(x) + \phi(x)(x^2 - 1)\frac{\kappa_3}{6v_0^{3/2}} + o(p^{-1/2}).$$

Then, we obtain the following theorem.

THEOREM 2.2. *The misclassification probability of quadratic discriminant function* (1.1) *is asymptotically given as*

$$\mathrm{P}(2|1) \sim \Phi(-a) + \phi(a)(a^2 - 1)\frac{\kappa}{6v_0^{3/2}} + o(\frac{1}{\sqrt{p}}), \tag{2.7}$$

where

$$\begin{aligned} a &= \frac{-2\log c - e_0}{v_0^{1/2}}, \\ e_0 &= p - (\mathrm{tr}\Lambda_p^{-1} + \eta'\Lambda_p^{-1}\eta) + \sum_{i=1}^{p} \log \lambda_i^{-1}, \\ v_0 &= 2p - 4\mathrm{tr}\Lambda_p^{-1} + 2\mathrm{tr}\Lambda_p^{-2} + 4\eta'\Lambda_p^{-2}\eta, \\ \kappa &= 8p + 24\eta'(\Lambda_p^{-2} - \Lambda_p^{-3})\eta - 24\mathrm{tr}\Lambda_p^{-1} + 24\mathrm{tr}\Lambda_p^{-2} - 8\mathrm{tr}\Lambda_p^{-3}. \end{aligned}$$

When we assume that $a = O(1)$, i.e.

$$p - (\mathrm{tr}\Lambda_p^{-1} + \eta'\Lambda_p^{-1}\eta) = O(\sqrt{p}) \quad (p \to \infty), \tag{2.8}$$

the approximation will be good.

2.3. Asymptotic expansion of misclassification probability for the sample quadratic discriminant function

Next, we consider the asymptotic expansion of misclassification probability for the SQDF (1.2). In the same way for the QDF, we standardize $\mathbf{X}$ and $\mathbf{X}_{ij}$ using (2.5). Then, we obtain

$$\begin{aligned} Q(\boldsymbol{y};\bar{\boldsymbol{y}}_1,\bar{\boldsymbol{y}}_2,B_1,B_2) &= n_1(\boldsymbol{y}-\bar{\boldsymbol{y}}_1)'B_1^{-1}(\boldsymbol{y}-\bar{\boldsymbol{y}}_1) - n_2(\boldsymbol{y}-\bar{\boldsymbol{y}}_2)'B_2^{-1}(\boldsymbol{y}-\bar{\boldsymbol{y}}_2) \\ &\quad + \log|B_1/B_2| - p\log(n_1/n_2), \end{aligned} \tag{2.9}$$

where $n_iB_i = L'S_iL$ and the random variables $\boldsymbol{y}$, $\boldsymbol{y}_{1j}$, $\boldsymbol{y}_{2j}$, $\bar{\boldsymbol{y}}_1$, $\bar{\boldsymbol{y}}_2$, B_1 and B_2 are distributed as

$$\begin{aligned} &\boldsymbol{y},\boldsymbol{y}_{1j} \sim \mathrm{N}_p(\mathbf{0},I_p), \quad \boldsymbol{y}_{2j} \sim \mathrm{N}_p(\eta,\Lambda_p), \quad \eta = L(\mu_2-\mu_1), \\ &\bar{\boldsymbol{y}}_1 \sim \mathrm{N}_p(\mathbf{0},\frac{1}{N_1}I_p), \quad \bar{\boldsymbol{y}}_2 \sim \mathrm{N}_p(\eta,\frac{1}{N_2}\Lambda_p), \\ &B_1 \sim \mathrm{W}_p(n_1,I_p), \quad B_2 \sim \mathrm{W}_p(n_2,\Lambda_p), \end{aligned}$$

respectively. Besides, we need to show (2.9) as the independent random variables to calculate the asymptotic approximation of misclassification probability using (2.4). For using some theorems (see Muirhead (1982), e.g.), we obtain the following result.

Lemma 2.2. *The sample quadratic discriminant function* (2.9) *is represented as*

$$\begin{aligned} &Q(\boldsymbol{y};\bar{\boldsymbol{y}}_1,\bar{\boldsymbol{y}}_2,B_1,B_2) \\ &\quad = n_1V_{11}^{-1}(\boldsymbol{y}-\bar{\boldsymbol{y}}_1)'(\boldsymbol{y}-\bar{\boldsymbol{y}}_1) - n_2V_{21}^{-1}(\boldsymbol{y}-\bar{\boldsymbol{y}}_2)'\Lambda_p^{-1}(\boldsymbol{y}-\bar{\boldsymbol{y}}_2) \\ &\quad + \sum_{k=1}^{p}(\log(V_{1k}/n_1) - \log(V_{2k}/n_2) - \log\lambda_k) \\ &\quad = n_1V_{11}^{-1}\sum_{i=1}^{p}(y_i-\bar{y}_{1i})^2 - n_2V_{21}^{-1}\sum_{i=1}^{p}\lambda_i^{-1}(y_i-\bar{y}_{2i})^2 \\ &\quad + \sum_{k=1}^{p}(\log(V_{1k}/n_1) - \log(V_{2k}/n_2) - \log\lambda_k), \end{aligned} \tag{2.10}$$

where $\boldsymbol{y} = (y_1, \cdots, y_p)'$, $\bar{\boldsymbol{y}}_1 = (\bar{y}_{11}, \cdots, \bar{y}_{1p})'$, $\bar{\boldsymbol{y}}_2 = (\bar{y}_{21}, \cdots, \bar{y}_{2p})'$ *and* V_{ij} *are independent and these random vectors are distributed as*

$$\begin{aligned}
&\boldsymbol{y} \sim \mathrm{N}_p(\mathbf{0}, I_p), \quad \bar{\boldsymbol{y}}_1 \sim \mathrm{N}_p(\mathbf{0}, \frac{1}{N_1} I_p), \quad \bar{\boldsymbol{y}}_2 \sim \mathrm{N}_p(\eta, \frac{1}{N_2}\Lambda_p), \\
&V_{ik} \sim \chi^2_{n_i - p + k} (i = 1, 2; k = 1, \cdots, N_i),
\end{aligned}$$

respectively.

If V_{11} and V_{21} are given, (2.10) are rated as the sum of independent random variables. We divide (2.10) into two parts, $Q(\boldsymbol{y}; \bar{\boldsymbol{y}}_1, \bar{\boldsymbol{y}}_2, B_1, B_2) = Q_1 + Q_2$, where

$$\begin{aligned}
Q_1 &= n_1 V_{11}^{-1} (\boldsymbol{y} - \bar{\boldsymbol{y}}_1)'(\boldsymbol{y} - \bar{\boldsymbol{y}}_1) - n_2 V_{21}^{-1} (\boldsymbol{y} - \bar{\boldsymbol{y}}_2)' \Lambda_p^{-1} (\boldsymbol{y} - \bar{\boldsymbol{y}}_2), \\
Q_2 &= \sum_{k=1}^{p} (\log(V_{1k}/n_1) - \log(V_{2k}/n_2) - \log \lambda_k).
\end{aligned}$$

For Q_1, it is easy to calculate the conditional cumulant, because $\boldsymbol{y}$, $\bar{\boldsymbol{y}}_1$ and $\bar{\boldsymbol{y}}_2$ is distributed as normal distribution. Then, conditional expectation, variance, the 3rd and the 4th cumulant of Q_1 are shown as

$$\mathrm{E}[Q_1|V_1, V_2] = p\left(a_{11}\frac{n_1}{V_1} + a_{12}\frac{n_2}{V_2}\right) + \left(a_{21}\frac{n_1}{V_1} + a_{22}\frac{n_2}{V_2}\right), \tag{2.11}$$

$$\begin{aligned}
\mathrm{Var}[Q_1|V_1, V_2] &= 2p\left(b_{11}\frac{n_1^2}{V_1^2} + b_{12}\frac{n_1 n_2}{V_1 V_2} + b_{13}\frac{n_2^2}{V_2^2}\right) \\
&\quad + \left(b_{21}\frac{n_1^2}{V_1^2} + b_{22}\frac{n_2^2}{V_2^2}\right) + O(p^{-1}),
\end{aligned} \tag{2.12}$$

$$\begin{aligned}
\kappa_3[Q_1|V_1, V_2] &= 8p\left(c_{11}\frac{n_1^3}{V_1^3} + c_{12}\frac{n_1^2 n_2}{V_1^2 V_2} + c_{13}\frac{n_1 n_2^2}{V_1 V_2^2} + c_{14}\frac{n_2^3}{V_2^3}\right) \\
&\quad + O(1),
\end{aligned} \tag{2.13}$$

$$\begin{aligned}
\kappa_4[Q_1|V_1, V_2] &= p\left(d_{11}\frac{n_1^4}{V_1^4} + d_{12}\frac{n_1^3 n_2}{V_1^3 V_2} + d_{13}\frac{n_1^2 n_2^2}{V_1^2 V_2^2} + d_{14}\frac{n_1 n_2^3}{V_1 V_2^3} + d_{15}\frac{n_2^4}{V_2^4}\right) \\
&\quad + O(1),
\end{aligned} \tag{2.14}$$

where

$$a_{11} = 1, \quad a_{12} = -\frac{1}{p}(\mathrm{tr}\Lambda^{-1} + \eta'\Lambda^{-1}\eta), \quad a_{21} = \frac{p}{N_1}, \quad a_{22} = -\frac{p}{N_2},$$

$$
\begin{aligned}
&b_{11} = 1, \quad b_{12} = -2p^{-1}\mathrm{tr}\Lambda^{-1}, \quad b_{13} = p^{-1}(\mathrm{tr}\Lambda^{-2} + 2\eta'\Lambda^{-2}\eta),\\
&b_{21} = 4\frac{p}{N_1}, \quad b_{22} = 4\frac{1}{N_2}(\mathrm{tr}\Lambda^{-1} + \eta'\Lambda^{-1}\eta),\\
&c_{11} = 1, \quad c_{12} = -3p^{-1}\mathrm{tr}\Lambda^{-1},\\
&c_{13} = 3p^{-1}(\mathrm{tr}\Lambda^{-2} + \eta'\Lambda^{-2}\eta), \quad c_{14} = -p^{-1}(\mathrm{tr}\Lambda^{-3} + \eta'\Lambda^{-3}\eta),\\
&d_{11} = 48, \quad d_{12} = -192p^{-1}\mathrm{tr}\Lambda^{-1}, \quad d_{13} = 96p^{-1}(3\mathrm{tr}\Lambda^{-2} + 2\eta'\Lambda^{-2}\eta),\\
&d_{14} = -192p^{-1}(\mathrm{tr}\Lambda^{-3} + 2\eta'\Lambda^{-3}\eta), \quad d_{15} = 48p^{-1}(\mathrm{tr}\Lambda^{-4} + 4\eta'\Lambda^{-4}\eta).
\end{aligned}
$$

For Q_2, it is difficult to calculate the cumulant of $\log(V_{ik}/n_i)$. Then, we standardize V_{ik} as

$$
U_{ik} = \frac{V_{ik} - (n_i - p + k)}{\sqrt{2(n_i - p + k)}}. \tag{2.15}
$$

U_{ik} is asymptotically distributed as standard normal distribution. Then the expectation, variance, the 3rd and 4th cumulant of Q_2 are shown as

$$
\begin{aligned}
&\mathrm{E}(Q_2) = pa_{13} + a_{23} + O(p^{-1}), \quad \mathrm{Var}(Q_2) = b_{23} + O(p^{-1}),\\
&\kappa_3(Q_2) \sim O(p^{-1}), \quad \kappa_4(Q_2) \sim O(p^{-2}),
\end{aligned} \tag{2.16}
$$

where

$$
\begin{aligned}
a_{13} &= -\frac{n_1 - p}{p}\log(\frac{n_1 - p}{n_1}) + \frac{n_2 - p}{p}\log(\frac{n_2 - p}{n_2}) - p^{-1}\sum_{k=1}^{p}\log\lambda_k,\\
a_{23} &= -\frac{1}{2}\log(\frac{n_1 - p}{n_1}) + \frac{1}{2}\log(\frac{n_2 - p}{n_2}) - \sum_{k=1}^{p}\frac{1}{n_1 - p + k} + \sum_{k=1}^{p}\frac{1}{n_2 - p + k},\\
b_{23} &= \sum_{k=1}^{p}\frac{2}{n_1 - p + k} + \sum_{k=1}^{p}\frac{2}{n_2 - p + k}.
\end{aligned}
$$

(We may need to consider the *conditional independence*.)

For the validity, we obtain the following lemma. (About the proof, see Theorem 2.1.)

LEMMA 2.3. *Let* $E_p = E[Q(\boldsymbol{y}; \bar{\boldsymbol{y}}_1, \bar{\boldsymbol{y}}_2, B_1, B_2)]$,
$v_p = V[Q(\boldsymbol{y}; \bar{\boldsymbol{y}}_1, \bar{\boldsymbol{y}}_2, B_1, B_2)]$, $\kappa_3 = \kappa_3[Q(\boldsymbol{y}; \bar{\boldsymbol{y}}_1, \bar{\boldsymbol{y}}_2, B_1, B_2)]$ *and*

$\kappa_4 = \kappa_4[Q(\boldsymbol{y};\bar{\boldsymbol{y}}_1,\bar{\boldsymbol{y}}_2,B_1,B_2)]$. *If*

$$\liminf_{p\to\infty}\frac{1}{p}\left[\mathrm{tr}\{(I_p-\Lambda_p^{-1})^2\}+2\eta'\Lambda^{-2}\eta\right]=O(1),$$
$$\frac{1}{p}\sum_{j=1}^{p}\lambda_j^{-k}|\eta_j|^{2l}=O(1) \qquad (k=1,\cdots,3,\ l=0,1,\cdots,3).$$

Then uniformly for x,

$$\Pr\left(\frac{Q(\boldsymbol{y};\bar{\boldsymbol{y}}_1,\bar{\boldsymbol{y}}_2,B_1,B_2))-E_p}{v_p^{1/2}}\le x\right)$$
$$=\Phi(x)+\phi(x)\left\{(x^2-1)\frac{\kappa_3}{6v_p^{3/2}}+(x^3-3x)\frac{\kappa_4}{24v_p^2}+(x^5-10x^3+15x)\frac{\kappa_3^2}{72v_p^3}\right\}$$
$$+o(p^{-1}).$$

Then, we obtain the following result.

Theorem 2.3. *The misclassification probability of the sample quadratic discriminant function is asymptotically given by the expectation form for* V_{11} *and* V_{21}, *as*

$$\mathrm{P}(2|1) \sim \mathrm{E}_{V_{11},V_{21}}\left[\Phi(-a)+\phi(a)\left\{(a^2-1)\frac{\kappa_3}{6v_p^{3/2}}\right.\right.$$
$$\left.\left.+(a^3-3a)\frac{\kappa_4}{24v_p^2}+(a^5-10a^3+15a)\frac{\kappa_3^2}{72v_p^3}\right\}\right]+o(p^{-1}), \qquad (2.17)$$

where

$$a=-\frac{2\log c+E_p}{v_p^{1/2}},$$
$$E_p\sim p\left(a_{11}\frac{n_1}{V_1}+a_{12}\frac{n_2}{V_2}+a_{13}\right)+\left(a_{21}\frac{n_1}{V_1}+a_{22}\frac{n_2}{V_2}+a_{23}\right)+O(p^{-1}), \qquad (2.18)$$
$$v_p\sim 2p\left(b_{11}\frac{n_1^2}{V_1^2}+b_{12}\frac{n_1n_2}{V_1V_2}+b_{13}\frac{n_2^2}{V_2^2}\right) \qquad (2.19)$$
$$+\left(b_{21}\frac{n_1^2}{V_1^2}+b_{22}\frac{n_2^2}{V_2^2}+b_{23}\right)+O(p^{-1}), \qquad (2.20)$$

$$\kappa_3 \sim 8p\left(c_{11}\frac{n_1^3}{V_1^3}+c_{12}\frac{n_1^2n_2}{V_1^2V_2}+c_{13}\frac{n_1n_2^2}{V_1V_2^2}+c_{14}\frac{n_2^3}{V_2^3}\right)+O(1), \tag{2.21}$$

$$\kappa_4 \sim p\left(d_{11}\frac{n_1^4}{V_1^4}+d_{12}\frac{n_1^3n_2}{V_1^3V_2}+d_{13}\frac{n_1^2n_2^2}{V_1^2V_2^2}+d_{14}\frac{n_1n_2^3}{V_1V_2^3}+d_{15}\frac{n_2^4}{V_2^4}\right)+O(1). \tag{2.22}$$

To calculate the expectation of (2.17) for V_{11} and V_{21}, we obtain the asymptotic expansion of the misclassification probability of the quadratic discriminant function. We use the standardization of V_{i1},

$$U_i = \frac{V_{i1}-(n_i-p+1)}{\sqrt{2(n_i-p+1)}}. \tag{2.23}$$

for the calculation. Then we obtain the following result.

Theorem 2.4. *The misclassification probability of the sample quadratic discriminant function* (1.2) *is asymptotically given by*

$$P(2|1) = f_1 + \frac{1}{\sqrt{p}}f_2 + \frac{1}{p}f_3 + o(p^{-1}), \tag{2.24}$$

where f_1, f_2 and f_3 are shown in Appendix.

3. Simulation

We carry out some numerical experiments for three purposes. The first purpose is to examine the accuracy of the asymptotic expansion formula given by Theorem2.2, at the case of parameter known. The second purpose is to examine the accuracy of the asymptotic expansion formula given by Theorem2.4, at the case of parameter unknown. And the third purpose is to compare the result in Wakaki (1990), at the case of only large samples.

3.1. Methods of the examination

We denote $\Pi_1 \sim N_p(\mathbf{0}, I_p)$ and $\Pi_2 \sim N_p(\eta, \Lambda_p)$ $(\Lambda_p = diag(\lambda_1, \cdots, \lambda_p))$, and consider the parameters,

Table 1: The results of sumulation when parameters are known.

		ζ	$p=10$ Sim.	$p=10$ (2.7)	$p=30$ Sim.	$p=30$ (2.7)	$p=30$ Sim.	$p=30$ (2.7)
$\alpha=1.0$		1.28	0.26046	0.261086	0.26087	0.261086	0.26114	0.204686
		1.65	0.20407	0.204686	0.20434	0.204686	0.20472	0.204686
		1.96	0.16347	0.163543	0.16356	0.163543	0.16374	0.163543
$\alpha=1.5$	λ_j $=\alpha^{(p-j)/p}$	1.28	0.23649	0.251002	0.22059	0.234978	0.20578	0.220524
		1.65	0.18537	0.204636	0.17717	0.193706	0.16889	0.183161
		1.96	0.15272	0.168724	0.14245	0.160875	0.13586	0.152876
	λ_j $=\alpha^{(p-2j)/p}$	1.28	0.27149	0.213971	0.25130	0.202815	0.23231	0.191302
		1.65	0.20983	0.147622	0.19930	0.142658	0.18811	0.136093
		1.96	0.17107	0.104062	0.15949	0.102030	0.14891	0.098083

Table 2: The results of sumulation when $p=10$ and parameters are unknown.

$p=10$		ζ	Simulation	(2.24) First Term	(2.24) Up to $p^{-1/2}$	(2.24) Up to p^{-1}
$\alpha=1.0$		1.28	0.50631	0.430105	0.538842	0.487638
		1.65	0.47727	0.387604	0.518028	0.482598
		1.96	0.44314	0.347019	0.498134	0.482363
$\alpha=1.5$	λ_j $=\alpha^{(p-j)/p}$	1.28	0.56486	0.473576	0.507339	0.488065
		1.65	0.52694	0.431938	0.503722	0.454312
		1.96	0.49797	0.391813	0.491756	0.429851
	λ_j $=\alpha^{(p-2j)/p}$	1.28	0.44754	0.397475	0.516215	0.508781
		1.65	0.41341	0.352555	0.490327	0.504309
		1.96	0.38052	0.310994	0.464419	0.506498

- $p = 10,\ 20,\ 30$;
- $N_i = 1.5p \quad (i = 1, 2)$;
- $c = 1$;
- $\lambda_j = \alpha^{(p-j)/p}$ or $\lambda_j = \alpha^{(p-2j)/p}$, where $\alpha = 1.0, 1.5$;
- $\eta = \delta(1, 2, \cdots, p)'$, where $\delta = \{\sum_{j=1}^{p} j^2(\frac{1}{2} + \frac{\lambda_j}{2})^{-1}\}^{-1/2} \times \zeta$ and $\zeta = 1.28, 1.65, 1.96$.

Then we make the samples from these parameters, and estimate the misclassification probability for (1.1) and (1.2) using Monte-Carlo methods. Besides, we calculate the asymptotic approximation by using (2.7) and (2.24). Then, we compare the result of (1.1) with that of (2.7), and the result of (1.2) with that of (2.24).

Table 3: The results of sumulation when $p = 20$ and parameters are unknown.

$p = 20$		ζ	Simulation	(2.24) First Term	Up to $p^{-1/2}$	Up to p^{-1}
$\alpha = 1.0$		1.28	0.49538	0.449489	0.516060	0.471124
		1.65	0.46938	0.417512	0.493675	0.453709
		1.96	0.44089	0.385872	0.472595	0.437448
$\alpha = 1.5$	$\lambda_j = \alpha^{(p-j)/p}$	1.28	0.59701	0.516046	0.495208	0.491296
		1.65	0.56655	0.483104	0.493039	0.488429
		1.96	0.53539	0.451499	0.480653	0.462304
	$\lambda_j = \alpha^{(p-2j)/p}$	1.28	0.41743	0.417274	0.489708	0.466100
		1.65	0.38696	0.382673	0.464844	0.446993
		1.96	0.35879	0.349117	0.441681	0.430393

Table 4: The results of sumulation when $p = 30$ and parameters are unknown.

$p = 30$		ζ	Simulation	(2.24) First Term	Up to $p^{-1/2}$	Up to p^{-1}
$\alpha = 1.0$		1.28	0.49564	0.458448	0.509690	0.470534
		1.65	0.46932	0.431766	0.488643	0.452221
		1.96	0.44503	0.405011	0.468404	0.434383
$\alpha = 1.5$	$\lambda_j = \alpha^{(p-j)/p}$	1.28	0.62762	0.544742	0.502985	0.471521
		1.65	0.60015	0.514772	0.501787	0.500689
		1.96	0.57320	0.487004	0.491157	0.489193
	$\lambda_j = \alpha^{(p-2j)/p}$	1.28	0.40353	0.424925	0.480545	0.456517
		1.65	0.37633	0.395834	0.457442	0.436433
		1.96	0.35051	0.367108	0.435737	0.417622

3.2 Results and comments

Table 1 shows the result to compare (2.7) with the numerical result using (1.1). Table 2, 3 and 4 show the results to compare (2.24) with the numerical result using (1.2) when $p = 10$, 20 and 30 respectively. Standard deviations of these numerical simulations are not over 0.01, everything.

In this simulation, it is shown that (2.7) has a good approximation. But (2.24) doesn't have a good approximation much. Its reasons are, we consider, that

- The sample quadratic discriminant function is the distribution of weight sum of independent noncentral chi-square random variables;

Table 5: Comparing the results with Wakaki(1990) (1).

p	N	μ	λ	Sim.	Wakaki's result		(2.24)		
					First Term	Up to N^{-1}	First Term	Up to $p^{-1/2}$	Up to p^{-1}
5	10	1.5	1.2	0.4610	0.7328	0.4384	0.3844	0.5632	0.3272
			1.6	0.4805	0.5850	0.4427	0.4337	0.5236	0.3757
		2.0	1.2	0.4025	0.5147	0.4287	0.3003	0.5665	0.4253
			1.6	0.4398	0.4345	0.4000	0.3631	0.5296	0.1904
		2.5	1.2	0.3426	0.2947	0.3183	0.2183	0.5423	0.7567
			1.6	0.3946	0.2821	0.3089	0.2875	0.5329	0.1758
	15	1.5	1.2	0.3705	0.7328	0.5366	0.3407	0.5158	0.3995
			1.6	0.3554	0.5850	0.4901	0.3794	0.4930	0.3085
		2.0	1.2	0.3017	0.5147	0.4574	0.2508	0.4626	0.5044
			1.6	0.3136	0.4345	0.4115	0.2992	0.4652	0.2642
		2.5	1.2	0.2385	0.2947	0.3104	0.1723	0.3836	0.6458
			1.6	0.2679	0.2821	0.2999	0.2203	0.4193	0.3327
	20	1.5	1.2	0.3313	0.7328	0.5856	0.3156	0.4774	0.4196
			1.6	0.2975	0.5850	0.5138	0.3455	0.4615	0.3073
		2.0	1.2	0.2622	0.5147	0.4717	0.2264	0.3971	0.4672
			1.6	0.2571	0.4345	0.4172	0.2633	0.4115	0.2985
		2.5	1.2	0.1989	0.2947	0.3065	0.1525	0.3038	0.5007
			1.6	0.2142	0.2821	0.2955	0.1870	0.3422	0.3418
	30	1.5	1.2	0.2872	0.7328	0.6347	0.2879	0.4246	0.4064
			1.6	0.2385	0.5850	0.5376	0.3055	0.4129	0.3126
		2.0	1.2	0.2207	0.5147	0.4861	0.2024	0.3229	0.3811
			1.6	0.2026	0.4345	0.4230	0.2248	0.3391	0.2984
		2.5	1.2	0.1626	0.2947	0.3026	0.1346	0.2264	0.3403
			1.6	0.1650	0.2821	0.2910	0.1546	0.2543	0.2897
	45	1.5	1.2	0.2594	0.7328	0.6674	0.2677	0.3792	0.3671
			1.6	0.2019	0.5850	0.5534	0.2745	0.3667	0.3043
		2.0	1.2	0.1954	0.5147	0.4956	0.1867	0.2693	0.3025
			1.6	0.1681	0.4345	0.4268	0.1978	0.2792	0.2661
		2.5	1.2	0.1410	0.2947	0.3000	0.1237	0.1771	0.2394
			1.6	0.1350	0.2821	0.2880	0.1338	0.1920	0.2261

- The dimension p of this simulation is too small; and
- As the condition for the parameters $\left(\sum_{j=1}^{p} \lambda_j^{-k} |\eta_j|^{2l}\right)/p$, its size has effect on the actual error.

3.3 Comparison

To compare the result of Wakaki (1990), we denote $\Pi_1 \sim N_p(\mathbf{0}, I_p)$ and $\Pi_2 \sim N_p(\boldsymbol{\mu}, I_p)$, where $N_1 = N_2 = N$ and $\boldsymbol{\mu} = (\mu, 0, \cdots, 0)^T$. Then we consider the parameters,

- $p = 5,\ 10,\ 20;$

Table 6: Comparing the results with Wakaki(1990) (2).

					Wakaki's result		(2.24)		
p	N	μ	λ	Sim.	First Term	Up to N^{-1}	First Term	Up to $p^{-1/2}$	Up to p^{-1}
10	15	1.5	1.2	0.5645	0.9599	1.0653	0.4636	0.5164	0.6898
			1.6	0.6320	0.8063	0.3795	0.5334	0.5013	0.3347
		2.0	1.2	0.5199	0.8805	0.5685	0.4071	0.5119	0.3655
			1.6	0.5989	0.7064	0.2815	0.4811	0.5033	0.5200
		2.5	1.2	0.4709	0.7225	0.2901	0.3415	0.5079	0.2025
			1.6	0.5610	0.5698	0.2639	0.4262	0.4924	0.3969
	20	1.5	1.2	0.4883	0.9599	1.0389	0.4328	0.5022	0.4784
			1.6	0.5150	0.8063	0.4862	0.4917	0.4998	0.5018
		2.0	1.2	0.4301	0.8805	0.6465	0.3603	0.4754	0.3242
			1.6	0.4757	0.7064	0.3877	0.4298	0.4794	0.4172
		2.5	1.2	0.3664	0.7225	0.3982	0.2821	0.4407	0.2940
			1.6	0.4301	0.5698	0.3404	0.3625	0.4506	0.2824
	30	1.5	1.2	0.4060	0.9599	1.0126	0.3894	0.4715	0.3921
			1.6	0.3704	0.8063	0.5929	0.4283	0.4658	0.4211
		2.0	1.2	0.3380	0.8805	0.7245	0.3042	0.4168	0.3284
			1.6	0.3311	0.7064	0.4940	0.3574	0.4255	0.3181
		2.5	1.2	0.2707	0.7225	0.5063	0.2213	0.3482	0.3198
			1.6	0.2890	0.5698	0.4169	0.2822	0.3762	0.2412
	45	1.5	1.2	0.3460	0.9599	0.9950	0.3499	0.4343	0.3660
			1.6	0.2730	0.8063	0.6640	0.3700	0.4189	0.3484
		2.0	1.2	0.2772	0.8805	0.7765	0.2603	0.3565	0.3179
			1.6	0.2389	0.7064	0.5648	0.2950	0.3638	0.2745
		2.5	1.2	0.2125	0.7225	0.5783	0.1805	0.2695	0.2814
			1.6	0.2009	0.5698	0.4679	0.2201	0.2983	0.2250

- $N = 10,\ 15,\ 20,\ 30,\ 45$, where $N \geq 1.5p$;
- $c = 1$;
- $\mu = 1.5\ ,2,\ 2.5$;
- $\lambda = 1.2,\ 1.6$.

For the same methods in Section 3.1, we obtain the numerical results at Table 5, 6 and 7. In this simulation, it is show that the approximation in (2.24) is better than Wakaki's works for large dimension and sample sizes. In Wakaki's results, the dimension p is used for the parameter of χ^2-distribution. Therefore, the approximation by his paper will be poor for large dimension.

Table 7: Comparing the results with Wakaki(1990) (3).

p	N	μ	λ	Sim.	Wakaki's result		(2.24)		
					First Term	Up to N^{-1}	First Term	Up to $p^{-1/2}$	Up to p^{-1}
20	30	1.5	1.2	0.5962	0.9994	1.2013	0.5064	0.5096	0.3557
			1.6	0.6998	0.9510	1.1591	0.6213	0.5496	0.1960
		2.0	1.2	0.5577	0.9968	1.3553	0.4620	0.4929	0.6478
			1.6	0.6719	0.9195	0.9376	0.5645	0.5329	0.3516
		2.5	1.2	0.5123	0.9845	1.3985	0.4094	0.4670	0.4570
			1.6	0.6378	0.8639	0.6370	0.5116	0.5094	0.4939
	45	1.5	1.2	0.5220	0.9994	1.1340	0.4705	0.4952	0.5639
			1.6	0.5191	0.9511	1.0897	0.5417	0.5208	0.5055
		2.0	1.2	0.4644	0.9968	1.2358	0.4067	0.4539	0.4360
			1.6	0.4837	0.9195	0.9316	0.4829	0.4888	0.4886
		2.5	1.2	0.3998	0.9845	1.2605	0.3344	0.4039	0.3239
			1.6	0.4421	0.8639	0.7126	0.4198	0.4463	0.4192

A. Appendix – Parameters

We may show the parameters of the derived results in Section 2.3.

Using U_i of (2.23), standardization of V_{i1}, (2.18), (2.20), (2.21) and (2.22) are shown as

$$
\begin{aligned}
E_p &= pr_0 + \sqrt{p}(r_{11}U_1 + r_{12}U_2) + (r_{21}U_1^2 + r_{22}U_2^2 + r_{23}) \\
&\quad + \frac{1}{\sqrt{p}}(r_{31}U_1^3 + r_{32}U_2^3 + r_{33}U_1 + r_{34}U_2) + O(p^{-1}), \\
v_p &= ps_0 + \sqrt{p}(s_{11}U_1 + s_{12}U_2) + (s_{21}U_1^2 + s_{22}U_1U_2 + s_{23}U_2^2 + s_{24}) \\
&\quad + \frac{1}{\sqrt{p}}(s_{31}U_1^3 + s_{32}U_1^2U_2 + s_{33}U_1U_2^2 + s_{34}U_2^3 + s_{35}U_1 + s_{36}U_2) + O(p^{-1}), \\
\kappa_3 &= pt_0 + \sqrt{p}(t_{11}U_1 + t_{12}U_2) + O(1), \\
\kappa_4 &= pw_0 + O(p^{1/2}),
\end{aligned}
$$

where

$$
\begin{aligned}
r_0 &= \frac{n_1}{n_1 - p + 1}a_{11} + \frac{n_2}{n_2 - p + 1}a_{12} + a_{13}, \\
r_{11} &= -\sqrt{2}\frac{p^{1/2}n_1}{(n_1 - p + 1)^{3/2}}a_{11}, \quad r_{12} = -\sqrt{2}\frac{p^{1/2}n_2}{(n_2 - p + 1)^{3/2}}a_{12}, \\
r_{21} &= 2\frac{pn_1}{(n_1 - p + 1)^2}a_{11}, \quad r_{22} = 2\frac{pn_2}{(n_2 - p + 1)^2}a_{12}, \\
r_{23} &= \frac{p}{n_1 - p + 1}a_{21} + \frac{p}{n_2 - p + 1}a_{22} + a_{23},
\end{aligned}
$$

$$r_{31} = -2\sqrt{2}\frac{p^{3/2}n_1}{(n_1-p+1)^{5/2}}a_{11}, \quad r_{32} = -2\sqrt{2}\frac{p^{3/2}n_2}{(n_2-p+1)^{5/2}}a_{12},$$

$$r_{33} = -\sqrt{2}\frac{p^{1/2}n_1}{(n_1-p+1)^{3/2}}a_{21}, \quad r_{34} = -\sqrt{2}\frac{p^{1/2}n_2}{(n_2-p+1)^{3/2}}a_{22},$$

$$s_0 = 2\left(\frac{n_1^2}{(n_1-p+1)^2}b_{11} + \frac{n_1n_2}{(n_1-p+1)(n_2-p+1)}b_{12} + \frac{n_2^2}{(n_2-p+1)^2}b_{13}\right),$$

$$s_{11} = -2\sqrt{2}\frac{p^{1/2}n_1}{(n_1-p+1)^{3/2}}\left(2\frac{n_1}{n_1-p+1}b_{11} + \frac{n_2}{n_2-p+1}b_{12}\right),$$

$$s_{12} = -2\sqrt{2}\frac{p^{1/2}n_2}{(n_2-p+1)^{3/2}}\left(2\frac{n_2}{n_2-p+1}b_{13} + \frac{n_1}{n_1-p+1}b_{12}\right),$$

$$s_{21} = 4\frac{pn_1}{(n_1-p+1)^2}\left(3\frac{n_1}{n_1-p+1}b_{11} + \frac{n_2}{n_2-p+1}b_{12}\right),$$

$$s_{22} = 4\frac{pn_1n_2}{(n_1-p+1)^{3/2}(n_2-p+1)^{3/2}}b_{12},$$

$$s_{23} = 4\frac{pn_2}{(n_2-p+1)^2}\left(3\frac{n_2}{n_2-p+1}b_{13} + \frac{n_1}{n_1-p+1}b_{12}\right),$$

$$s_{24} = \frac{n_1^2}{(n_1-p+1)^2}b_{21} + \frac{n_2^2}{(n_2-p+1)^2}b_{22} + b_{23},$$

$$s_{31} = -4\sqrt{2}\frac{p^{3/2}n_1}{(n_1-p+1)^{5/2}}\left(4\frac{n_1}{n_1-p+1}b_{11} + \frac{n_2}{n_2-p+1}b_{12}\right),$$

$$s_{32} = -4\sqrt{2}\frac{p^{3/2}n_1n_2}{(n_1-p+1)^2(n_2-p+1)^{3/2}}b_{12},$$

$$s_{33} = -4\sqrt{2}\frac{p^{3/2}n_1n_2}{(n_1-p+1)^{3/2}(n_2-p+1)^2}b_{12},$$

$$s_{34} = -4\sqrt{2}\frac{p^{3/2}n_2}{(n_2-p+1)^{5/2}}\left(4\frac{n_2}{n_2-p+1}b_{13} + \frac{n_1}{n_1-p+1}b_{12}\right),$$

$$s_{35} = -8\sqrt{2}\frac{p^{1/2}n_1^2}{(n_1-p+1)^{5/2}}b_{21}, \quad s_{36} = -8\sqrt{2}\frac{p^{1/2}n_2^2}{(n_2-p+1)^{5/2}}b_{22},$$

$$t_0 = 8\left(\frac{n_1^3}{(n_1-p+1)^3}c_{11} + \frac{n_1^2n_2}{(n_1-p+1)^2(n_2-p+1)}c_{12}\right.$$

$$\left.+\frac{n_1n_2^2}{(n_1-p+1)(n_2-p+1)^2}c_{13} + \frac{n_2^3}{(n_2-p+1)^3}c_{14}\right),$$

$$t_{11} = -8\sqrt{2}\frac{p^{1/2}n_1}{(n_1-p+1)^{3/2}}\left(3\frac{p^{1/2}n_1^2}{(n_1-p+1)^2}c_{11}\right.$$

$$+2\frac{p^{1/2}n_1n_2}{(n_1-p+1)(n_2-p+1)}c_{12}+\frac{p^{1/2}n_2^2}{(n_2-p+1)^2}c_{13}\Bigg),$$

$$t_{12}=-8\sqrt{2}\frac{p^{1/2}n_2}{(n_2-p+1)^{3/2}}\Bigg(3\frac{p^{1/2}n_2^2}{(n_2-p+1)^2}c_{14}$$

$$+2\frac{p^{1/2}n_1n_2}{(n_1-p+1)(n_2-p+1)}c_{13}+\frac{p^{1/2}n_1^2}{(n_1-p+1)^2}c_{12}\Bigg),$$

$$w_0=\Bigg(\frac{n_1^4}{(n_1-p+1)^4}d_{11}+\frac{n_1^3n_2}{(n_1-p+1)^3(n_2-p+1)}d_{12}$$

$$+\frac{n_1^2n_2^2}{(n_1-p+1)^2(n_2-p+1)^2}d_{13}+\frac{n_1n_2^3}{(n_1-p+1)(n_1-p+1)^3}d_{14}$$

$$+\frac{n_2^4}{(n_2-p+1)^4}d_{15}\Bigg).$$

For the condition of the good approximation and the expansion, if we assume that $r_0\times\sqrt{p}=O(1)$ and $s_0=O(1)$, a, $\kappa_3/6v_p^{3/2}$,$\kappa_4/24v_p^2$ and $\kappa_3^2/72v_p^3$ of (2.17) are shown as

$$\begin{aligned}
a &= \alpha_0+\alpha_{11}U_1+\alpha_{12}U_2+\frac{1}{\sqrt{p}}(\alpha_{21}U_1^2+\alpha_{22}U_1U_2+\alpha_{23}U_2^2+\alpha_{24})\\
&\quad+\frac{1}{p}(\alpha_{31}U_1^3+\alpha_{32}U_1^2U_2+\alpha_{33}U_1U_2^2+\alpha_{34}U_2^3+\alpha_{35}U_1+\alpha_{36}U_2)+o(p^{-1}),\\
\frac{\kappa_3}{6v_p^{3/2}} &= \frac{1}{\sqrt{p}}\beta_0+\frac{1}{p}(\beta_{11}U_1+\beta_{12}U_2)+o(p^{-1}),\\
\frac{\kappa_4}{24v_p^2} &= \frac{1}{p}\gamma_0+o(p^{-1}),\quad \frac{\kappa_3^2}{72v_p^3}=\frac{1}{p}\delta_0+o(p^{-1}),
\end{aligned}$$

where

$$\alpha_0=-\sqrt{p}\frac{r_0}{\sqrt{s_0}},\quad \alpha_{11}=\frac{r_0s_{11}}{2\sqrt{s_0}^3}-\frac{r_{11}}{\sqrt{s_0}},\quad \alpha_{12}=\frac{r_0s_{12}}{2\sqrt{s_0}^3}-\frac{r_{12}}{\sqrt{s_0}},$$

$$\alpha_{21}=-\frac{3r_0s_{11}^2}{4\sqrt{s_0}^5}+\frac{2r_0s_{21}+s_{11}r_{11}}{2\sqrt{s_0}^3}-\frac{r_{21}}{\sqrt{s_0}},$$

$$\alpha_{22}=-\frac{3r_0s_{11}s_{12}}{4\sqrt{s_0}^5}+\frac{r_0s_{22}+s_{12}r_{11}+r_{12}s_{11}}{2\sqrt{s_0}^3},$$

$$\alpha_{23}=-\frac{3r_0s_{12}^2}{4\sqrt{s_0}^5}+\frac{2r_0s_{23}+s_{12}r_{12}}{2\sqrt{s_0}^3}-\frac{r_{22}}{\sqrt{s_0}},\quad \alpha_{24}=\frac{r_0s_{24}}{2\sqrt{s_0}^3}-\frac{r_{23}+2\log c}{\sqrt{s_0}},$$

$$\alpha_{31} = \frac{15r_0s_{11}^3}{8\sqrt{s_0}^7} - \frac{18r_0s_{11}s_{21} + 3r_{11}s_{11}^2}{4\sqrt{s_0}^5} + \frac{6r_0s_{31} + 2s_{21}r_{11} + r_{21}s_{11}}{2\sqrt{s_0}^3} - \frac{r_{31}}{\sqrt{s_0}},$$

$$\alpha_{32} = \frac{15r_0s_{11}^2s_{12}}{8\sqrt{s_0}^7} - \frac{6r_0(s_{12}s_{21} + s_{11}s_{22}) + 3r_{12}s_{11}^2 + 3r_{11}s_{11}s_{12}}{4\sqrt{s_0}^5}$$

$$+\frac{2r_0s_{32} + 2s_{21}r_{12} + r_{11}s_{22} + r_{21}s_{12}}{2\sqrt{s_0}^3},$$

$$\alpha_{33} = \frac{15r_0s_{11}s_{12}^2}{8\sqrt{s_0}^7} - \frac{6r_0(s_{12}s_{22} + s_{11}s_{23}) + 3r_{11}s_{12}^2 + 3r_{12}s_{11}s_{12}}{4\sqrt{s_0}^5}$$

$$+\frac{2r_0s_{33} + 2s_{23}r_{11} + r_{12}s_{22} + r_{22}s_{11}}{2\sqrt{s_0}^3},$$

$$\alpha_{34} = \frac{15r_0s_{12}^3}{8\sqrt{s_0}^7} - \frac{18r_0s_{12}s_{23} + 3r_{12}s_{12}^2}{4\sqrt{s_0}^5} + \frac{6r_0s_{34} + 2s_{23}r_{12} + r_{22}s_{12}}{2\sqrt{s_0}^3} - \frac{r_{32}}{\sqrt{s_0}},$$

$$\alpha_{35} = -\frac{3r_0s_{11}s_{24}}{4\sqrt{s_0}^5} + \frac{r_0s_{35} + s_{24}r_{11} + s_{11}r_{23}}{2\sqrt{s_0}^3} - \frac{r_{33}}{\sqrt{s_0}},$$

$$\alpha_{36} = -\frac{3r_0s_{12}s_{24}}{4\sqrt{s_0}^5} + \frac{r_0s_{36} + s_{24}r_{12} + s_{12}r_{23}}{2\sqrt{s_0}^3} - \frac{r_{34}}{\sqrt{s_0}},$$

$$\beta_0 = \frac{t_0}{6\sqrt{p}^3}, \quad \beta_{11} = -\frac{s_{11}t_0}{4\sqrt{s_0}^5} + \frac{t_{11}}{6\sqrt{s_0}^3}, \quad \beta_{12} = -\frac{s_{12}t_0}{4\sqrt{s_0}^5} + \frac{t_{12}}{6\sqrt{s_0}^3},$$

$$\gamma_0 = \frac{w_0}{24s_0^2}, \quad \delta_0 = \frac{t_0^2}{72s_0^3}.$$

Therefore, we can calculate the right side of (2.17) as the expectation of U_i. We use the asymptotic distribution of U_i,

$$f_{U_i}(x) := \phi(x)\{1 + \frac{1}{\sqrt{p}}\tau_{1i}(x^3 - 3x) + \frac{1}{p}\tau_{2i}(2x^6 - 21x^4 + 36x^2 - 3)\} + o(p^{-1}), \tag{A.1}$$

where

$$\tau_{1i} = \frac{\sqrt{2}\sqrt{p}}{3\sqrt{n_i - p + 1}}, \quad \tau_{2i} = \frac{p}{18(n_i - p + 1)},$$

and some property about standard normal distribution function. Then, we obtain the parameters f_1, f_2 and f_3 of Theorem 2.4,

$$f_1 = \Phi\left(-\frac{\alpha_0}{\sqrt{\alpha_{11}^2 + \alpha_{12}^2 + 1}}\right),$$

$$f_2 = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}\frac{\alpha_0^2}{\alpha_{11}^2+\alpha_{12}^2+1})\frac{1}{\sqrt{\alpha_{11}^2+\alpha_{12}^2+1}}$$

$$\times\{\zeta_0 - \frac{\alpha_{11}\alpha_0}{\alpha_{11}^2+\alpha_{12}^2+1}\zeta_{11} - \frac{\alpha_{12}\alpha_0}{\alpha_{11}^2+\alpha_{12}^2+1}\zeta_{12}$$

$$+\frac{(\alpha_{12}^2+1)^2+\alpha_{11}^2(1+\alpha_0^2+\alpha_{12}^2)}{(\alpha_{11}^2+\alpha_{12}^2+1)^2}\zeta_{21} - \frac{\alpha_{11}\alpha_{12}(1-\alpha_0^2+\alpha_{11}^2+\alpha_{12}^2)}{(\alpha_{11}^2+\alpha_{12}^2+1)^2}\zeta_{22}$$

$$+\frac{(\alpha_{11}^2+1)^2+\alpha_{12}^2(1+\alpha_0^2+\alpha_{11}^2)}{(\alpha_{11}^2+\alpha_{12}^2+1)^2}\zeta_{23}\},$$

$$f_3 = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}\frac{\alpha_0^2}{\alpha_{11}^2+\alpha_{12}^2+1})\frac{1}{\sqrt{\alpha_{11}^2+\alpha_{12}^2+1}}$$

$$\times\{\theta_0 - \frac{\alpha_{11}\alpha_0}{\alpha_{11}^2+\alpha_{12}^2+1}\theta_{11} - \frac{\alpha_{12}\alpha_0}{\alpha_{11}^2+\alpha_{12}^2+1}\theta_{12}$$

$$+\frac{(\alpha_{12}^2+1)^2+\alpha_{11}^2(1+\alpha_0^2+\alpha_{12}^2)}{(\alpha_{11}^2+\alpha_{12}^2+1)^2}\theta_{21}$$

$$-\frac{\alpha_{11}\alpha_{12}(1-\alpha_0^2+\alpha_{11}^2+\alpha_{12}^2)}{(\alpha_{11}^2+\alpha_{12}^2+1)^2}\theta_{22}$$

$$+\frac{(\alpha_{11}^2+1)^2+\alpha_{12}^2(1+\alpha_0^2+\alpha_{11}^2)}{(\alpha_{11}^2+\alpha_{12}^2+1)^2}\theta_{23}$$

$$-\frac{\alpha_0\alpha_{11}\{3(1+\alpha_{12}^2)^2+\alpha_{11}^2(3+\alpha_0^2+3\alpha_{12}^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^3}\theta_{31}$$

$$-\frac{\alpha_0\alpha_{12}\{(1+\alpha_{12}^2)^2+\alpha_{11}^2(-1+\alpha_0^2-\alpha_{12}^2)-2\alpha_{11}^4\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^3}\theta_{32}$$

$$-\frac{\alpha_0\alpha_{11}\{(1+\alpha_{11}^2)^2+\alpha_{12}^2(-1+\alpha_0^2-\alpha_{11}^2)-2\alpha_{12}^4\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^3}\theta_{33}$$

$$-\frac{\alpha_0\alpha_{12}\{3(1+\alpha_{11}^2)^2+\alpha_{12}^2(3+\alpha_0^2+3\alpha_{11}^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^3}\theta_{34}$$

$$+\Bigg[\frac{3(1+\alpha_{12}^2)^4+6\alpha_{11}^2(1+\alpha_{12}^2)^2(1+\alpha_0^2+\alpha_{12}^2)}{(\alpha_{11}^2+\alpha_{12}^2+1)^4}$$

$$+\frac{\alpha_{11}^4\{\alpha_0^4+6\alpha_0^2(1+\alpha_{12}^2)+3(1+\alpha_{12}^2)^2\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^4}\Bigg]\theta_{41}$$

$$+\frac{\alpha_{11}\alpha_{12}\{\alpha_0^4\alpha_{11}^2-3(1+\alpha_{12}^2)(1+\alpha_{11}^2+\alpha_{12}^2)^2-3\alpha_0^2(\alpha_{11}^4-(1+\alpha_{12}^2)^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^4}\theta_{42}$$

$$+\Bigg[\frac{\alpha_{11}^6(1+\alpha_0^2+3\alpha_{12}^2)+(1+\alpha_{12}^2)^2\{1+(1+\alpha_0^2)\alpha_{12}^2\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^4}$$

$$
\begin{aligned}
&+\frac{\alpha_{11}^4\{3+9\alpha_{12}^2+6\alpha_{12}^4+\alpha_0^2(2-3\alpha_{12}^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^4} \\
&\left.+\frac{\alpha_{11}^2\{\alpha_0^4\alpha_{12}^2+3(1+\alpha_{12}^2)^3+\alpha_0^2(1-2\alpha_{12}^2-3\alpha_{12}^4)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^4}\right]\theta_{43} \\
&+\frac{\alpha_{11}\alpha_{12}\{\alpha_0^4\alpha_{12}^2-3(1+\alpha_{11}^2)(1+\alpha_{11}^2+\alpha_{12}^2)^2-3\alpha_0^2(\alpha_{12}^4-(1+\alpha_{11}^2)^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^4}\theta_{44} \\
&+\left[\frac{3(1+\alpha_{11}^2)^4+6\alpha_{12}^2(1+\alpha_{11}^2)^2(1+\alpha_0^2+\alpha_{11}^2)}{(\alpha_{11}^2+\alpha_{12}^2+1)^4}\right. \\
&\left.+\frac{\alpha_{12}^4\{\alpha_0^4+6\alpha_0^2(1+\alpha_{11}^2)+3(1+\alpha_{11}^2)^2\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^4}\right]\theta_{45} \\
&-\left[\frac{\alpha_0\alpha_{11}\{15(1+\alpha_{12}^2)^4+10\alpha_{11}^2(1+\alpha_{11}^2)^2(3+\alpha_0^2+3\alpha_{12}^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\right. \\
&\left.+\frac{\alpha_0\alpha_{11}^5(\alpha_0^4+10\alpha_0^2(1+\alpha_{12}^2)+15(1+\alpha_{12}^2)^2)}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\right]\theta_{51} \\
&-\left[\frac{\alpha_0\alpha_{12}\{6\alpha_{11}^2(-1+\alpha_0^2-\alpha_{12}^2)(1+\alpha_{12}^2)^2}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\right. \\
&+\frac{\alpha_0\alpha_{12}\{3(1+\alpha_{12}^2)^4-4\alpha_{11}^6(3+\alpha_0^2+3\alpha_{12}^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5} \\
&\left.+\frac{\alpha_0\alpha_{12}\{\alpha_{11}^4(\alpha^4+2\alpha^2(1+\alpha_{12}^2)-21(1+\alpha_{12}^2)^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\right]\theta_{52} \\
&-\left[\frac{\alpha_0\alpha_{11}\{\alpha_{11}^6(3+\alpha_0^2+9\alpha_{12}^2)-3(1+\alpha_{12}^2)^2(-1-(-1+\alpha_0^2)\alpha_{12}^2+2\alpha_{12}^4)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\right. \\
&+\frac{\alpha_0\alpha_{11}\{\alpha_{11}^2(\alpha_0^4\alpha_{12}^2-3(-3+\alpha_{12}^2)(1+\alpha_{12}^2)^2+\alpha_0^2(1-2\alpha_{12}^2-3\alpha_{12}^4))\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5} \\
&\left.+\frac{\alpha_0\alpha_{11}^5\{(\alpha_0^2(2-5\alpha_{12}^2)+3(3+7\alpha_{12}^2+4\alpha_{12}^4)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\right]\theta_{53} \\
&-\left[\frac{\alpha_0\alpha_{12}\{\alpha_{12}^6(3+\alpha_0^2+9\alpha_{11}^2)-3(1+\alpha_{11}^2)^2(-1-(-1+\alpha_0^2)\alpha_{11}^2+2\alpha_{11}^4)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\right. \\
&+\frac{\alpha_0\alpha_{12}\{\alpha_{12}^2(\alpha_0^4\alpha_{11}^2-3(-3+\alpha_{11}^2)(1+\alpha_{11}^2)^2+\alpha_0^2(1-2\alpha_{11}^2-3\alpha_{11}^4))\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5} \\
&\left.+\frac{\alpha_0\alpha_{12}^5\{(\alpha_0^2(2-5\alpha_{11}^2)+3(3+7\alpha_{11}^2+4\alpha_{11}^4)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\right]\theta_{54} \\
&-\left[\frac{\alpha_0\alpha_{11}\{6\alpha_{12}^2(-1+\alpha_0^2-\alpha_{11}^2)(1+\alpha_{11}^2)^2}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\right.
\end{aligned}
$$

$$+\frac{3(1+\alpha_{11}^2)^4-4\alpha_{12}^6(3+\alpha_0^2+3\alpha_{11}^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}$$
$$+\frac{\alpha_0\alpha_{11}\{\alpha_{12}^4(\alpha_0^4+2\alpha_0^2(1+\alpha_{11}^2)-21(1+\alpha_{11}^2)^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\Bigg]\theta_{55}$$
$$-\Bigg[\frac{\alpha_0\alpha_{12}\{15(1+\alpha_{11}^2)^4+10\alpha_{12}^2(1+\alpha_{12}^2)^2(3+\alpha^2+3\alpha_{11}^2)\}}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}$$
$$+\frac{\alpha_0\alpha_{12}^5(\alpha_0^4+10\alpha_0^2(1+\alpha_{11}^2)+15(1+\alpha_{11}^2)^2)}{(\alpha_{11}^2+\alpha_{12}^2+1)^5}\Bigg]\theta_{56}\},$$

where

$$\zeta_0=-\alpha_{24}+\beta_0-\alpha_0^2\beta_0+\alpha_{11}\tau_{11}+\alpha_{12}\tau_{12},\quad \zeta_{11}=-2\alpha_0\beta_0\alpha_{11},$$
$$\zeta_{12}=-2\alpha_0\beta_0\alpha_{12},\quad \zeta_{21}=-\alpha_{21}-\alpha_{11}^2\beta_0-\alpha_{11}\tau_{11},$$
$$\zeta_{22}=-\alpha_{22}-2\alpha_{11}\alpha_{12}\beta_0,\quad \zeta_{23}=-\alpha_{23}-\alpha_{12}^2\beta_0-\alpha_{12}\tau_{12},$$
$$\theta_0=\frac{\alpha_0\alpha_{24}^2}{2}-3\alpha_0\alpha_{24}\beta_0+\alpha_0^3\alpha_{24}\beta_0$$
$$-15\alpha_0\delta_0+10\alpha_0^3\delta_0-\alpha_0^5\delta_0+3\alpha_0\gamma_0-\alpha_0^3\gamma_0,$$
$$\theta_{11}=\frac{\alpha_{11}\alpha_{24}^2}{2}-\alpha_{35}-3\alpha_{11}\alpha_{24}\beta_0+3\alpha_0^2\alpha_{11}\alpha_{24}\beta_0+\beta_{11}$$
$$-\alpha_0^2\beta_{11}-15\alpha_{11}\delta_0+30\alpha_0^2\alpha_{11}\delta_0-5\alpha_0^4\alpha_{11}\delta_0+3\alpha_{11}\gamma_0-3\alpha_0^2\alpha_{11}\gamma_0$$
$$+3\alpha_{24}\tau_{11}-3\beta_0\tau_{11}+3\alpha_0^2\beta_0\tau_{11}-3\alpha_{11}\tau_{21},$$
$$\theta_{12}=\frac{\alpha_{12}\alpha_{24}^2}{2}-\alpha_{36}-3\alpha_{12}\alpha_{24}\beta_0+3\alpha_0^2\alpha_{12}\alpha_{24}\beta_0+\beta_{12}$$
$$-\alpha_0^2\beta_{12}-15\alpha_{12}\delta_0+30\alpha_0^2\alpha_{12}\delta_0-5\alpha_0^4\alpha_{12}\delta_0+3\alpha_{12}\gamma_0-3\alpha_0^2\alpha_{12}\gamma_0$$
$$+3\alpha_{24}\tau_{12}-3\beta_0\tau_{12}+3\alpha_0^2\beta_0\tau_{12}-3\alpha_{12}\tau_{22},$$
$$\theta_{21}=\alpha_0\alpha_{21}\alpha_{24}-3\alpha_0\alpha_{21}\beta_0+\alpha_0^3\alpha_{21}\beta_0+3\alpha_0\alpha_{11}^2\alpha_{24}\beta_0$$
$$-2\alpha_0\alpha_{11}\beta_{11}+30\alpha_0\alpha_{11}^2\delta_0-10\alpha_0^3\alpha_{11}^2\delta_0-3\alpha_0\alpha_{11}^2\gamma_0+6\alpha_0\alpha_{11}\beta_0\tau_{11},$$
$$\theta_{22}=\alpha_0\alpha_{22}\alpha_{24}-3\alpha_0\alpha_{22}\beta_0+\alpha_0^3\alpha_{22}\beta_0+6\alpha_0\alpha_{11}\alpha_{12}\alpha_{24}\beta_0$$
$$-2\alpha_0\alpha_{12}\beta_{11}-2\alpha_0\alpha_{11}\beta_{12}+60\alpha_0\alpha_{11}\alpha_{12}\delta_0-20\alpha_0^3\alpha_{11}\alpha_{12}\delta_0$$
$$-6\alpha_0\alpha_{11}\alpha_{12}\gamma_0+6\alpha_0\alpha_{12}\beta_0\tau_{11}+6\alpha_0\alpha_{11}\beta_0\tau_{12},$$
$$\theta_{23}=\alpha_0\alpha_{23}\alpha_{24}-3\alpha_0\alpha_{23}\beta_0+\alpha_0^3\alpha_{23}\beta_0+3\alpha_0\alpha_{12}^2\alpha_{24}\beta_0$$
$$-2\alpha_0\alpha_1\beta_{12}+30\alpha_0\alpha_{12}^2\delta_0-10\alpha_0^3\alpha_{12}^2\delta_0-3\alpha_0\alpha_{12}^2\gamma_0+6\alpha_0\alpha_{12}\beta_0\tau_{12},$$

$$\theta_{31} = \alpha_{11}\alpha_{21}\alpha_{24} - \alpha_{31} - 3\alpha_{11}\alpha_{21}\beta_0 + 3\alpha_0^2\alpha_{11}\alpha_{21}\beta_0 + \alpha_{11}^3\alpha_{24}\beta_0$$
$$-\alpha_{11}^2\beta_{11} + 10\alpha_{11}^3\delta_0 - 10\alpha_0^2\alpha_{11}^3\delta_0 - \alpha_{11}^3\gamma_0$$
$$+3\alpha_{21}\tau_{11} - \alpha_{24}\tau_{11} + \beta_0\tau_{11} - \alpha_0^2\beta_0\tau_{11} + 3\alpha_{11}^2\beta_0\tau_{11} + 11\alpha_{11}\tau_{21},$$
$$\theta_{32} = \alpha_{12}\alpha_{21}\alpha_{24} + \alpha_{11}\alpha_{22}\alpha_{24} - \alpha_{32}$$
$$-3\alpha_{12}\alpha_{21}\beta_0 + 3\alpha_0^2\alpha_{12}\alpha_{21}\beta_0 - 3\alpha_{11}\alpha_{22}\beta_0 + 3\alpha_0^2\alpha_{11}\alpha_{22}\beta_0 + 3\alpha_{11}^2\alpha_{12}\alpha_{24}\beta_0$$
$$-2\alpha_{11}\alpha_{12}\beta_{11} - \alpha_{11}^2\beta_{12} + 30\alpha_{11}^2\alpha_{12}\delta_0 - 30\alpha_0^2\alpha_{11}^2\alpha_{12}\delta_0 - 3\alpha_{11}^2\alpha_{12}\gamma_0$$
$$+3\alpha_{22}\tau_{11} + 6\alpha_{11}\alpha_{12}\beta_0\tau_{11} + 3\alpha_{21}\tau_{12} + 3\alpha_{11}^2\beta_0\tau_{12},$$
$$\theta_{33} = \alpha_{12}\alpha_{22}\alpha_{24} + \alpha_{11}\alpha_{23}\alpha_{24} - \alpha_{33}$$
$$-3\alpha_{12}\alpha_{22}\beta_0 + 3\alpha_0^2\alpha_{12}\alpha_{22}\beta_0 - 3\alpha_{11}\alpha_{23}\beta_0 + 3\alpha_0^2\alpha_{11}\alpha_{23}\beta_0 + 3\alpha_{11}\alpha_{12}^2\alpha_{24}\beta_0$$
$$-2\alpha_{11}\alpha_{12}\beta_{12} - \alpha_{12}^2\beta_{11} + 30\alpha_{11}\alpha_{12}^2\delta_0 - 30\alpha_0^2\alpha_{11}\alpha_{12}^2\delta_0 - 3\alpha_{11}\alpha_{12}^2\gamma_0$$
$$+3\alpha_{23}\tau_{11} + 6\alpha_{11}\alpha_{12}\beta_0\tau_{12} + 3\alpha_{22}\tau_{12} + 3\alpha_{12}^2\beta_0\tau_{11},$$
$$\theta_{34} = \alpha_{12}\alpha_{23}\alpha_{24} - \alpha_{34} - 3\alpha_{12}\alpha_{23}\beta_0 + 3\alpha_0^2\alpha_{12}\alpha_{23}\beta_0 + \alpha_{12}^3\alpha_{24}\beta_0$$
$$-\alpha_{12}^2\beta_{12} + 10\alpha_{12}^3\delta_0 - 10\alpha_0^2\alpha_{12}^3\delta_0 - \alpha_{12}^3\gamma_0$$
$$+3\alpha_{23}\tau_{12} - \alpha_{24}\tau_{12} + \beta_0\tau_{12} - \alpha_0^2\beta_0\tau_{12} + 3\alpha_{12}^2\beta_0\tau_{12} + 11\alpha_{12}\tau_2,$$
$$\theta_{41} = \frac{\alpha_0\alpha_{21}^2}{2} + 3\alpha_0\alpha_{11}^2\alpha_{21}\beta_0 - 5\alpha_0\alpha_{11}^4\delta_0 - 2\alpha_0\alpha_{11}\beta_0\tau_{11},$$
$$\theta_{42} = \alpha_0\alpha_{21}\alpha_{22} + 6\alpha_0\alpha_{11}\alpha_{12}\alpha_{21}\beta_0 + 3\alpha_0\alpha_{11}^2\alpha_{22}\beta_0$$
$$-20\alpha_0\alpha_{11}^3\alpha_{12}\delta_0 - 2\alpha_0\alpha_{12}\beta_0\tau_{11},$$
$$\theta_{43} = \frac{\alpha_0\alpha_{22}^2}{2} + \alpha_0\alpha_{21}\alpha_{23} + 3\alpha_0\alpha_{12}^2\alpha_{21}\beta_0$$
$$+6\alpha_0\alpha_{11}\alpha_{12}\alpha_{22}\beta_0 + 3\alpha_0\alpha_{11}^2\alpha_{23}\beta_0 - 30\alpha_0\alpha_{11}^2\alpha_{12}^2\delta_0,$$
$$\theta_{44} = \alpha_0\alpha_{22}\alpha_{23} + 6\alpha_0\alpha_{11}\alpha_{12}\alpha_{23}\beta_0 + 3\alpha_0\alpha_{12}^2\alpha_{22}\beta_0$$
$$-20\alpha_0\alpha_{11}\alpha_{12}^3\delta_0 - 2\alpha_0\alpha_{11}\beta_0\tau_{12},$$
$$\theta_{45} = \frac{\alpha_0\alpha_{23}^2}{2} + 3\alpha_0\alpha_{12}^2\alpha_{23}\beta_0 - 5\alpha_0\alpha_{12}^4\delta_0 - 2\alpha_0\alpha_{12}\beta_0\tau_{12},$$
$$\theta_{51} = \frac{\alpha_{11}\alpha_{21}^2}{2} + \alpha_{11}^3\alpha_{21}\beta_0 - \alpha_{11}^5\delta_0 - \alpha_{21}\tau_{11} - \alpha_{11}^2\beta_0\tau_{11} - 2\alpha_{11}\tau_{21},$$
$$\theta_{52} = \frac{\alpha_{12}\alpha_{21}^2}{2} + \alpha_{11}\alpha_{21}\alpha_{22} + 3\alpha_{11}^2\alpha_{12}\alpha_{21}\beta_0 + \alpha_{11}^3\alpha_{22}\beta_0$$

$$-5\alpha_{11}^4\alpha_{12}\delta_0 - \alpha_{22}\tau_{11} - 2\alpha_{11}\alpha_{12}\beta_0\tau_{11},$$

$$\theta_{53} = \frac{\alpha_{11}\alpha_{22}^2}{2} + \alpha_{12}\alpha_{21}\alpha_{22} + \alpha_{11}\alpha_{21}\alpha_{23} + 3\alpha_{11}\alpha_{12}^2\alpha_{21}\beta_0$$

$$+3\alpha_{11}^2\alpha_{12}\alpha_{22}\beta_0 + \alpha_{11}^3\alpha_{23}\beta_0 - 10\alpha_{11}^3\alpha_{12}^2\delta_0 - \alpha_{23}\tau_{11} - \alpha_{12}^2\beta_0\tau_{11},$$

$$\theta_{54} = \frac{\alpha_{12}\alpha_{22}^2}{2} + \alpha_{12}\alpha_{21}\alpha_{23} + \alpha_{11}\alpha_{22}\alpha_{23} + 3\alpha_{11}\alpha_{12}^2\alpha_{22}\beta_0$$

$$+3\alpha_{11}^2\alpha_{12}\alpha_{23}\beta_0 + \alpha_{12}^3\alpha_{21}\beta_0 - 10\alpha_{11}^2\alpha_{12}^3\delta_0 - \alpha_{21}\tau_{12} - \alpha_{11}\alpha_{11}\beta_0\tau_{12},$$

$$\theta_{55} = \frac{\alpha_{11}\alpha_{23}^2}{2} + \alpha_{12}\alpha_{22}\alpha_{23} + 3\alpha_{11}\alpha_{12}^2\alpha_{23}\beta_0 + \alpha_{12}^3\alpha_{22}\beta_0$$

$$-5\alpha_{11}\alpha_{12}^4\delta_0 - \alpha_{22}\tau_{12} - 2\alpha_{11}\alpha_{12}\beta_0\tau_{12},$$

$$\theta_{56} = \frac{\alpha_{12}\alpha_{23}^2}{2} + \alpha_{12}^3\alpha_{23}\beta_0 - \alpha_{12}^5\delta_0 - \alpha_{23}\tau_{12} - \alpha_{12}^2\beta_0\tau_{12} - 2\alpha_{12}\tau_{22}.$$

Acknowledgment

The authors would like to thank Professor Y. Fujikoshi, Hiroshima University, for several useful comments and suggestions.

Reference

[1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis,* Second Edition. New York, Wiley, 1984.

[2] Z. D. Bai and L. C. Zhao, Edgeworth expansions of distribution functions of independent random variables, *Sci.China Ser.A*, **29** (1986), 1-22.

[3] O. E. Barndorff-Nielsen and D. R. Cox, *Asymptotic Techniques for Use in Statistics,* Chapman and Hall, 1989.

[4] R. N. Bhattacharya and J. K. Ghosh, On the validity of the formal Edgeworth expansions, *Ann. Statist.*, **6** (1978), 434-451.

[5] R. N. Bhattacharya and R. Ranga Rao, *Normal Approximation and Asymptotic Expansions,* Wiley, New York. Reprint with corrections and Supplemental material (1986). Krieger, Malabar, Florida, 1976.

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

[7] C. Matsumoto, An optimal discriminant rule in the class of linear and quadratic discriminant functions for large dimension and samples, *Hiroshima Mathematical Journal*, **34** (2004) (in printing).

[8] A. Z. Memon and M. Okamoto, Asymptotic expansion of the distribution of the Z statistics in discriminani anarysis, *J. Multivariate Anal.* **1** (1971), 294-307.

[9] R. J. Muirhead, *Aspects of Multivariate Statistical Theory,* John Willey and Sons, Inc., New York, 1982.

[10] M. Okamoto, An asymptotic expansion for the distribution of the linear discriminant function, *Ann. Math. Statist.*, **34** (1963), 1286-1301. Correction. *Ann. Math. Statist.*, **39** (1968), 1358-1359.

[11] H. Saranadasa, Asymptotic expansion of the misclassification probabilities of D- and A-criteria for discrimination from two high dimensional populations using the theory of large dimensional random matrices, *J.Multivariate Anal.*, **46** (1993), 154-174.

[12] M. Siotani, T. Hayakawa and Y. Fujikoshi, *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook,* American Sciences Press, INC, 1985.

[13] T. Tonda and H. Wakaki, EPMC estimation in discriminant analysis when the dimension and sample sizes are large, TR No. 03-08, Statistical Research Groupe, Hiroshima University, 2003.

[14] H. Wakaki, Comparison of linear and quadratic discriminant functions, *Biometrika*, **77** (1990), 227-229.

Chieko Matsumoto
Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603, Japan
e-mail. shisato@ier.hit-u.ac.jp

Hirofumi Wakaki
Department of Mathematics
Graduate school of Science
Hiroshima University
Higashi-Hiroshima 739-8526, Japan
e-mail. wakaki@math.sci.hiroshima-u.ac.jp