HIROSHIMA MATH. J.
23 (1993), 365–416

Admissibility of some tests, multiple decision procedures and classification procedures in multivariate analysis

Nobuo NISHIDA

(Received January 18, 1992)

1. Introduction

In this paper we study the admissibility of some tests, multiple decision procedures and classification procedures. In general, two methods are mainly used in multivariate analysis to show the admissibility of various procedures. One is to use Bayes procedures. The other is to use the structure of the exponential family. The former method has been used in Kiefer and Schwartz [16], Nishida [19], [20], [21], [22], [24]. The latter method has been seen in Ghosh [13], Birnbaum [9], Stein [33], Schwartz [29], Anderson and Takemura [5], etc. In this paper we use the former method. All the problems are studied for 0–1 loss function.

In Section 3, we consider testing problems related to a given structure of means (For testing a given structure of means, see, e.g., Rao [27], Mardia et al. [18], Siotani et al. [31]). Nishida [24] obtained a class of admissible tests for the combined problem of a given structure of means and $\Sigma = \Sigma_0$. In this section two testing problems are considered. One is to test the combined hypothesis of a given structure of means and the sphericity covariance structure. The other is to test a given structure of means under the sphericity covariance structure. The admissibility of the likelihood ratio test (LRT) is shown for each problem.

Testing problems for covariance matrices are studied in Section 4. As for testing independence of sets of variates, Kiefer and Schwartz [16] derived a class of admissible tests. They also treated the problem of testing equality of covariance matrices for k samples case. One sample case (that is, the problem that $\Sigma = \Sigma_0$) was studied by Nishida [19]. Each work obtained a class of admissible tests which contains the LRT. In this section we consider one sided tests for one and two samples cases. Linear structure for the inverse matrix of a covariance matrix is also considered. A class of admissible tests is obtained for each problem.

In Section 5, the admissibility of multiple decision procedures for covariance matrices is studied. Multiple decision problems or ranking problems on means of normal populations have been studied in many literature. For example, see Bechhofer [6], Bechhofer et al. [7], Dunnett [11], Paulson [26] and Seal [30]. The same problems for variances of normal populations have been studied by Bechhofer and Sobel [8], Eaton [12], Hall [14] and other authors. However, these problems are not sufficiently studied for multivariate case. In this section, first we consider a multiple decision problem for covariance matrices in two samples case. Secondly, the three samples case is studied. Various types of problems are considered in the three samples case. A theorem which gives a class of admissible procedures and its corollaries are derived for each problem.

The classification problem with unequal covariance matrices is studied in Section 6. Admissible classification rules were given by Kiefer and Schwartz [16], Nishida [19], [21] and [22] under various situations. In this section we derive three maximum likelihood (ML) classification rules for the unequal covariance matrices case. The rules are extensions of the ones in Kanazawa [15]. It is shown that two of them are admissible. Further, the limiting distributions of the three rules are also studied. As a result, it becomes clear that they have the same limiting distribution whose expectation is related to the Kullback-Leibler information (cf. Theorem 6.4). A class of admissible classification rules which have the same limiting distribution as stated above is also given. Finally, numerical simulations are carried out to examine some properties of the three rules.

2. Notations and preliminary lemmas

It is known (Kiefer and Schwartz [16]) that an admissible Bayes critical region (for 0–1 loss function) is of the form

$$\left\{X : \int f(X;\theta)\Pi_1(d\theta) \Big/ \int f(X;\theta)\Pi_0\, d\theta) \geq c\right\}, \tag{2.1}$$

for some positive constant c, where X is the matrix of total random sample, θ is the vector of parameters, $f(X;\theta)$ is the *p.d.f.* of X given θ, and Π_0 and Π_1 are the probability measures over the null parameter space H_0 and the alternative parameter space H_1, respectively. Here, it is assumed that the distribution of X is continuous type. We identify the hypothesis and the corresponding parameter space. Since c is arbitrary in (2.1) we only require for Π_0 and Π_1 to be finite instead of $\Pi(\Omega) = 1$, where $\Pi = \Pi_0 + \Pi_1$ and $\Omega = H_0 + H_1$. By the same reason, we often omit constant multiples in calculating Bayes rules (see [16]). The density of variables is always described by f, even if variables and/or parameters are changed. For example, we use the notations $f(X;\theta)$, $f(Y, Z; \mu, \Sigma)$ and so on.

For a multiple decision problem with three or more decisions, it is easily seen that the Bayes rule is given as follows: Let the total parameter space Ω be devided to a disjoint union of $H_1, H_2, \ldots, H_k$ and Π_i denote a finite measure over H_i $(i = 1, 2, \ldots, k)$. Then the Bayes rule is given by

$$\text{(2.2)} \qquad \text{choose } H_j \text{ if } c_j \int f(X;\theta)\Pi_j(d\theta) = \max_i c_i \int f(X;\theta) f(X;\theta)\Pi_i(d\theta),$$

where c_i's are any constants and maximum is taken for $i = 1, 2, \cdots k$.

Next, we state a lemma which is given in [16] and is useful for obtaining Bayes rules. Under H_1, let $X = (Y, U)$ be a random matrix whose columns are independently distributed as $N_p(\cdot, \Sigma)$. Also assume Σ is unknown and $EU = v(p \times 1)$ (unknown). Let θ^* be the parameter of Y and θ that of X, *i.e.*, $\theta = (\theta^*, v)$. Let H_i^* be the domain of θ^* under H_i, and consider the case where the domain of v is E^p and

$$\text{(2.3)} \qquad H_i = H_i^* \times E^p \quad (i = 0, 1).$$

Of course, (2.3) means that

$$\text{(2.4)} \qquad \theta \in H_i \text{ if and only if } \theta^* \in H_i^*.$$

Let H_i^{**} be a subset of H_i^* for which Σ can be written as $\Sigma = (C_0 + D_i)^{-1}$, where C_0 is a given positive definite matrix and D_i is a positive semidefinite matrix. Further, consider a finite measure Π_i^* on H_i^* which assigns a whole measure to H_i^{**}. Then, the following lemma due to Kiefer and Schwartz [16] holds:

LEMMA 1. *There exist finite measures Π_0 and Π_1 over H_0 and H_1, respectively, which satisfy*

(2.5)
$$\int f(X;\theta)\Pi_1(d\theta) \Big/ \int f(X;\theta)\Pi_0(d\theta) = \int f(Y;\theta^*)\Pi_1^*(d\theta^*) \Big/ \int f(Y;\theta^*)\Pi_0^*(d\theta^*).$$

Using this lemma, we can treat the problem without U and hence v. Since this lemma is proved by showing that it is possible to construct Π_i from Π_i^* which satisfy (for some positive constant d_i)

$$\text{(2.6)} \qquad \int f(X;\theta)\Pi_i(d\theta) = d_i \cdot \operatorname{etr}\{-C_0UU'/2\} \cdot \int f(Y;\theta^*)\Pi_i^*(d\theta) \quad (i = 1, 2),$$

we can generalize the lemma for the procedure (2.2). So, it is possible to eliminate U and v in such a case.

The following two lemmas which are useful for the integrability of prior densities were also given in [16].

LEMMA 2. *Let η be a $p \times q$ matrix. Then*

$$\int_{E^{pq}} |I_p + \eta\eta'|^{-h/2} d\eta < \infty \tag{2.7}$$

if and only if $h > p + q - 1$.

LEMMA 3. *Let η be a $p \times q$ matrix with $q \geq p$. If $p - 1 < q + t < h - p + 1$ and $t > -1$, then*

$$\int_{E^{pq}} |\eta\eta'|^{t/2} |I_p + \eta\eta'|^{-h/2} d\eta < \infty. \tag{2.8}$$

The following lemma is also used for the integrability of densities, especially in multiple decision problems. The result was given in [23], but it is necessary to correct its proof. So, we give a complete proof.

LEMMA 4. *Let $\eta = (\eta_1, \dots, \eta_k)$, where η_i is a $p \times q_i$ matrix with $q_i \geq p$ $(i = 1, \dots, k)$. If*

$$p - 1 < q_1 + t_i < h_i (i = 1, \dots, k-1),\ p - 1 < q_k + t_k < h_k - p + 1 \tag{2.9}$$
$$\text{and } t_i > -1 \ (i = 1, \dots, k),$$

then

$$\int_{E^{pq}} \prod_{i=1}^k (|\eta_i\eta_i'|^{t_i/2} |I_p + \textstyle\sum_{j=1}^i \eta_j\eta_j'|^{-h_j/2}) d\eta < \infty, \tag{2.10}$$

where $q = \sum_{i=1}^k q_i$.

PROOF. The proof is given only for $k = 2$, but it is easy to extend the proof to the case $k \geq 3$.

Using the transformation $\eta_2^* = (I_p + \eta_1\eta_1')^{-1/2}\eta_2$, the integral in (2.10) can be calculated as follows.

$$\begin{aligned} &\int |\eta_1\eta_1'|^{t_1/2} |\eta_2\eta_2'|^{t_2/2} |I_p + \eta_1\eta_1'|^{-h_1/2} |I_p + \eta_1\eta_1' + \eta_2\eta_2'|^{-h_2/2} d\eta \\ &= \int |I_p + \eta_1\eta_1'|^{-(h_1+h_2)/2} |I_p + (I_p + \eta_1\eta_1')^{-1}\eta_2\eta_2'|^{-h_2/2} \\ &\qquad \cdot |\eta_1\eta_1'|^{t_1/2} |\eta_2\eta_2'|^{t_2/2} d\eta \\ &= \int |I_p + \eta_1\eta_1'|^{-(h_1+h_2-t_2-q_2)/2} |\eta_1\eta_1'|^{t_1/2} d\eta_1 \\ &\qquad \cdot \int |I_p + \eta_2^*\eta_2^{*\prime}|^{-h_2/2} |\eta_2^*\eta_2^{*\prime}|^{t_2/2} d\eta_2^*. \end{aligned} \tag{2.11}$$

Here, integrals are carried out over Euclidean spaces. The integrals in the last line of (2.11) are integrable under the condition that

$$p-1<q_1+t_1<h_1+h_2-t_2-q_2-p+1, \tag{2.12}$$
$$p-1<q_2+t_2<h_2-p+1,\quad t_1>-1 \text{ and } t_2>-1.$$

Since $h_2-t_2-q_2-p+1>0$ from the second inequality, the first inequality in (2.12) is valid when

$$p-1<q_1+t_1<h_1. \tag{2.13}$$

In Lemma 4 with $k=2$, (2.10) holds if (2.12) is satisfied. However, (2.9) is rather convenient for use, because the condition for q_1+t_1 does not contain q_2 and t_2. The same argument holds for $k \geq 3$.

In calculating Bayes rules, integrations are usually carried out over Euclidean spaces or the space of whole positive definite matrices, and it is always neglected to state the spaces explicitly.

3. Admissibility of the LRT for a given structure of means

3.1. A given structure of means with the sphericity covariance structure

Let $X_1, X_2, \ldots, X_N$ be a random sample from $N_p(\mu, \Sigma)$. Consider the two problems of testing

$$H_0\colon H\mu=\xi_0 \text{ and } \Sigma=\sigma^2\Sigma_0 \text{ against } H_1\colon \text{not } H_0, \tag{3.1}$$

and

$$H_0\colon H\mu=\xi_0 \text{ against } H_1\colon \text{not } H_0 \tag{3.2}$$

under the assumption $\Sigma=\sigma^2\Sigma_0$. Here, $H(q\times p)$ and $\xi_0(q\times 1)$ are prespecified matrix and vector, respectively. It is assumed that rank $(H)=q\leq p$, $\Sigma_0(p\times p)$ is a given positive definite matrix, and σ^2 is an unknown positive number. It is also assumed that $p\geq 2$ and $N-1\geq p$.

These problems are regarded as the ones obtained by combining a given structure problem with the sphericity problem. For the sphericity problem, see, e.g., Anderson [3]. The problems considered in this section are slightly different from the one which was treated in Nishida [24]. The problem dealt there was

$$H_0\colon H\mu=\xi_0 \text{ and } \Sigma=\Sigma_0 \text{ against } H_1\colon \text{not } H_0, \tag{3.3}$$

and a class of admissible tests which include the LRT was obtained.

Admissibility of the LRT's is derived for (3.1) and (3.2) in the following subsections.

Now, let us summarize the maximum likelihood estimators (MLE's) under H_0 in (3.1) or (3.2). Letting $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ and $S = \sum_{i=1}^{N} (X_i - \bar{X})(X_i - \bar{X})'$, then

$$(3.4) \qquad \log L(\mu, \Sigma) = -\frac{N}{2}\log|2\pi\Sigma| - \frac{1}{2}\operatorname{tr}\Sigma^{-1}S - \frac{N}{2}\operatorname{tr}\Sigma^{-1}(\bar{X}-\mu)(\bar{X}-\mu)',$$

where $L(\mu, \Sigma)$ is the likelihood function. Under H_0, $\log L(\mu, \Sigma)$ is maximized by

$$(3.5) \qquad \mu^* = \bar{X} - \Sigma H'(H\Sigma H')^{-1}(H\bar{X} - \xi_0)$$

for any fixed Σ (This is shown by the calculations similar to the ones in p. 106–107 of [18]). Hence

$$(3.6) \qquad \hat{\mu} = \bar{X} - \Sigma_0 H'(H\Sigma_0 H')^{-1}(H\bar{X} - \xi_0)$$

is the MLE of μ under H_0, which coincides with the MLE under the hypothesis $H\mu = \xi_0$ and $\Sigma = \Sigma_0$. Since

(3.7)
$$\max_{\mu \in H_0} \log L(\mu, \Sigma) = -\frac{N}{2}\Big(p \cdot \log\sigma^2 + \log|2\pi\Sigma_0| + \frac{1}{\sigma^2}\Big\{\frac{1}{N}\operatorname{tr}\Sigma_0^{-1}S + (H\bar{X} - \xi_0)'(H\Sigma_0 H')^{-1}(H\bar{X} - \xi_0)\Big\}\Big),$$

the MLE of σ^2 is given by

$$(3.8) \qquad \hat{\sigma}^2 = \frac{1}{pN}\{\operatorname{tr}\Sigma_0^{-1}S + N(H\bar{X} - \xi_0)'(H\Sigma_0 H')^{-1}(H\bar{X} - \xi_0)\}.$$

3.2. Admissibility of the LRT for (3.1)

The problem (3.1) can be reduced to the following form: Let $Z_1, Z_2, \dots, Z_N$ be a random sample from $N_p(\nu, \Phi)$. Our problem is to test

$$(3.9) \qquad H_0\colon \nu_{r+1} = \cdots = \nu_p = 0 \text{ and } \Phi = \sigma^2 I \text{ against } H_1\colon \text{not } H_0,$$

where $\nu' = (\nu_1, \dots, \nu_p)$ and $r = p - q$. Let

$$(3.10) \qquad \bar{Z} = \frac{1}{N}\sum_{i=1}^{N} Z_i = (\bar{Z}_1, \dots, \bar{Z}_p)'$$

and

(3.11) $$A = \sum_{i=1}^{N} (Z_i - \bar{Z})(Z_i - \bar{Z})',$$

then the following theorem holds for (3.1).

THEOREM 3.1. *If $n = N - 1 > p$, a test with the critical region*

(3.12) $$\{\operatorname{tr} A + N \textstyle\sum_{i=r+1}^{p} \bar{Z}_i^2\}^{p/2} / |A|^{1/2} \geq c$$

is admissible Bayes for any c, and is the LRT.

PROOF. Under H_0, set

(3.13) $$\begin{aligned} (\sigma^2)^{-1} &= 1 + \tau^2 \\ \nu_i &= \tau\gamma_i/(1 + \tau^2) \qquad (i = 1, \ldots, r), \end{aligned}$$

and consider

(3.14) $$d\Pi_0/d\theta = \frac{|\tau|^{p-1}}{(1 + \tau^2)^{pN/2}} \cdot \exp\left(-\frac{N}{2}\left\{1 - \frac{\tau^2}{1 + \tau^2}\right\} \textstyle\sum_{i=1}^{r} \gamma_i^2\right)$$

as a prior distribution on H_0. By the reason stated in preliminaries, constant multiples are usually omitted hereafter in calculating Bayes rules (except Subsection 6.1). Denote the sample matrix by Z. Then

(3.15) $$\begin{aligned} \int f(Z; \nu, \Phi) d\Pi_0 &= \int \frac{1}{|\Phi|^{N/2}} \cdot \operatorname{etr}\left(-\frac{1}{2}\Phi^{-1}\{A + N(\bar{Z} - \nu)(\bar{Z} - \nu)'\}\right) d\Pi_0 \\ &= \int |\tau|^{p-1} \operatorname{etr}\left(-\frac{1}{2}(1 + \tau^2)\{A + N(\bar{Z} - \nu)(\bar{Z} - \nu)'\}\right) \\ &\quad \cdot \exp\left(-\frac{N}{2}\left\{1 - \frac{\tau^2}{1 + \tau^2}\right\} \textstyle\sum_{i=1}^{r} \gamma_i^2\right) d\gamma_1 \cdots d\gamma_r \, d\tau \\ &= \operatorname{etr}\left\{-\frac{1}{2}(A + N\bar{Z}\bar{Z}')\right\} \int |\tau|^{p-1} \operatorname{etr}\left\{-\frac{1}{2}\tau^2(A + N\bar{Z}\bar{Z}')\right\} \\ &\quad \cdot \exp\left(-\frac{N}{2} \textstyle\sum_{i=1}^{r} \left\{-2\tau\gamma_i\bar{Z}_i + \dfrac{\tau^2\gamma_i^2}{1 + \tau^2}\right\}\right) \\ &\quad \cdot \exp\left(-\frac{N}{2}\left\{1 - \frac{\tau^2}{1 + \tau^2}\right\} \textstyle\sum_{i=1}^{r} \gamma_i^2\right) d\gamma_1 \cdots d\gamma_r \, d\tau. \end{aligned}$$

Since

(3.16) $$\begin{aligned} &\int \exp\left(-\frac{N}{2}\left\{-2\tau\gamma_i\bar{Z}_i + \frac{\tau^2\gamma_i^2}{1 + \tau^2}\right\}\right) \exp\left(-\frac{N}{2}\left\{1 - \frac{\tau^2}{1 + \tau^2}\right\}\gamma_i^2\right) d\gamma_i \\ &= \exp\left(\frac{N}{2}\tau^2 \bar{Z}_i^2\right) \end{aligned}$$

for any i, (3.15) reduces to

$$\text{(3.17)}\quad \begin{aligned} &\operatorname{etr}\{-(A+N\bar{Z}\bar{Z}')/2\}\int|\tau|^{p-1}\exp(-\{\operatorname{tr}A+N\textstyle\sum_{i=r+1}^{p}\bar{Z}_i^2\}\tau^2/2)\,d\tau \\ &= \operatorname{etr}\left\{-\frac{1}{2}(A+N\bar{Z}\bar{Z}')\right\}\cdot\{\operatorname{tr}A+N\textstyle\sum_{i=r+1}^{p}\bar{Z}_i^2\}^{-p/2}. \end{aligned}$$

Under H_1, we transform $Z=(Z_1,\dots,Z_N)$ by a suitable orthogonal matrix $T(N\times N)$ so that $V_{n+1}=\sqrt{N}\bar{Z}$ in $ZT=V=(V_1,\dots,V_n,V_{n+1})$. Then, the columns of V are independently distributed as $N_p(\cdot,\Phi)$ with $\mathrm{E}V_i=0$ for $i=1,\dots,n$. Since the domain of $\mathrm{E}V_{n+1}$ is E^p, we apply Lemma 1 in Nishida [20] for V_{n+1} with $v_0=0$ in calculating the Bayes rule. Set

$$\text{(3.18)}\qquad \Phi^{-1}=I_p+\eta\eta' \quad\text{with}\quad \eta(p\times 1)$$

and

$$\text{(3.19)}\qquad d\Pi_1^*/d\eta=|I_p+\eta\eta'|^{-n/2}.$$

Then we have

$$\text{(3.20)}\quad \begin{aligned} \int f(Z;v,\Phi)\,d\Pi_1 &= \operatorname{etr}\left(-\frac{1}{2}N\bar{Z}\bar{Z}'\right)\cdot\int f(V^*;\Phi)\,d\Pi_1^* \\ &= \operatorname{etr}\left(-\frac{1}{2}N\bar{Z}\bar{Z}'\right)\cdot\int\operatorname{etr}\left\{-\frac{1}{2}(I_p+\eta\eta')A\right\}d\eta \\ &= \operatorname{etr}\left\{-\frac{1}{2}(A+N\bar{Z}\bar{Z}')\right\}\cdot|A|^{-1/2}, \end{aligned}$$

where $V^*=(V_1,\dots,V_n)$. Therefore, the procedure

$$\text{(3.21)}\qquad \frac{\int f(Z;v,\Phi)\,d\Pi_1}{\int f(Z;v,\Phi)\,d\Pi_0}=\{\operatorname{tr}A+N\textstyle\sum_{i=r+1}^{p}\bar{Z}_i^2\}^{p/2}/|A|^{1/2}\geq c$$

is admissible Bayes. The prior density (3.19) is shown to be integrable by Lemma 2. Further, it is easily seen that the procedure is the LRT.

We can write (3.12) in the terms of original variables as follows:

$$\text{(3.22)}\qquad \{\operatorname{tr}\Sigma_0^{-1}S+N(H\bar{X}-\xi_0)'(H\Sigma_0H')^{-1}(H\bar{X}-\xi_0)\}^{p/2}/|\Sigma_0^{-1}S|^{1/2}\geq c.$$

REMARK 3.1. In the special case $q=p\,(r=0)$, (3.1) becomes to

$$\text{(3.23)}\qquad H_0\colon \mu=\mu_0 \text{ and } \Sigma=\sigma^2 I \text{ against } H_1\colon \text{not } H_0.$$

This is considered as the combined problem of $\mu = \mu_0$ and the sphericity structure.

3.3. Admissibility of the LRT for (3.2)

By the same notations as in the previous subsection, the testing problem (3.2) can be reduced as follows: Let $Z_1, Z_2, \dots, Z_N$ be a random sample from $N_p(\nu, \sigma^2 I)$. Then we want to test

$$H_0: \nu_{r+1} = \cdots = \nu_p = 0 \text{ against } H_1: \text{not } H_0. \tag{3.24}$$

Since $\Sigma = \sigma^2 I_p$ under both H_0 and H_1, this problem can be considered as the one of testing a linear hypotesis in univariate linear model. Therefore the admissibility of the LRT is already known. However, we give another derivation based on Bayes approach.

THEOREM 3.2. *If* $0 < \alpha < pn - p + 1$, *a test with the critical region*

$$\{\text{tr } A + N\textstyle\sum_{i=r+1}^{p} \bar{Z}_i^2\}^p/(\text{tr } A)^\alpha \geq c \tag{3.25}$$

is admissible Bayes for any c.

PROOF. Under H_0, we use the same prior distribution as used in the proof of Theorem 3.1. Under H_1, we consider the same transformation as in the previous subsection and apply Lemma 1 in [20] for $\sqrt{N}\bar{Z}$ with setting $\nu_0 = 0$. Further, set

$$(\sigma^2)^{-1} = 1 + \tau^2, \tag{3.26}$$

$$d\Pi_1^*/d\tau = |\tau|^{\alpha-1}(1 + \tau^2)^{-pn/2}. \tag{3.27}$$

This density is integrable. For this prior distribution, we have

$$\begin{aligned} \int f(Z; \nu, \sigma^2) d\Pi_1 &= \text{etr}\left(-\frac{1}{2}N\bar{Z}\bar{Z}'\right) \cdot \int f(V^*; \sigma^2) d\Pi_1^* \\ &= \text{etr}\left\{-\frac{1}{2}(A + N\bar{Z}\bar{Z}')\right\} \cdot \int |\tau|^{\alpha-1} \exp\left\{-\frac{1}{2}(\text{tr } A)\tau^2\right\} d\tau \\ &= \text{etr}\left\{-\frac{1}{2}(A + N\bar{Z}\bar{Z}')\right\} \cdot (\text{tr } A)^{-\alpha/2}. \end{aligned} \tag{3.28}$$

Combining this with (3.17), we obtain the Bayes critical region

$$\{\text{tr } A + N\textstyle\sum_{i=r+1}^{p} \bar{Z}_i^2\}^{p/2}/(\text{tr } A)^{\alpha/2} \geq c_1 \tag{3.29}$$

which is equivalent to (3.25).

If we set $\alpha = p$ in (3.25), the following corollary is obtained.

COROLLARY 3.2.1. *For any c, a test with the critical region*

$$\{\mathrm{tr}\, A + N\textstyle\sum_{i=r+1}^{p}\bar{Z}_i^2\}/(\mathrm{tr}\, A) \geq c \tag{3.30}$$

is admissible Bayes, and is the LRT.

In the terms of original variables, (3.30) can be expressed as

$$\{\mathrm{tr}\,\Sigma_0^{-1}S + N(H\bar{X} - \xi_0)'(H\Sigma_0 H')^{-1}(H\bar{X} - \xi_0)\}/\mathrm{tr}\,\Sigma_0^{-1}S \geq c. \tag{3.31}$$

REMARK 3.2. In the case $q = p$, (3.24) is considered as the problem of testing $\mu = \mu_0$ under $\Sigma = \sigma^2 I$, that is

$$H_0: \mu = \mu_0 \text{ against } H_1: \text{not } H_0 \text{ with the assumption } \Sigma = \sigma^2 I. \tag{3.32}$$

REMARK 3.3. Relating to (3.23) and (3.32), we recollect the problem

$$H_0: \Sigma = \sigma^2 I \text{ against } H_1: \text{not } H_0 \text{ with the assumption } \mu = \mu_0, \tag{3.33}$$

which is the sphericity problem in the known mean vector case. We note that the admissibility of the LRT for (3.33) is easily obtained by the following prior distribution.

Under H_0, set

$$(\sigma^2)^{-1} = 1 + \tau^2 \quad \text{and} \quad d\Pi_0/d\tau = |\tau|^{p-1}(1 + \tau^2)^{-pN/2}. \tag{3.34}$$

Under H_1, set

$$\Sigma^{-1} = I_p + \eta\eta' \quad \text{and} \quad d\Pi_1/d\eta = |I_p + \eta\eta'|^{-N/2}, \tag{3.35}$$

where η is a $p \times 1$ vector. The admissibility of the LRT for the sphericity problem with unknown mean vector was obtained in Kiefer and Schwartz [16].

4. Tests for covariance matrices

4.1. One sided test for one sample case

Suppose $X(p \times N) = (X_1, \ldots, X_N)$ be a random sample from $N_p(\mu, \Sigma)$. We consider the problem of testing

$$H_0: \Sigma = \Sigma_0 \quad \text{against} \quad H_1: \Sigma < \Sigma_0. \tag{4.1}$$

Here, Σ_0 is a given positive definite matrix and $\Sigma < \Sigma_0$ means that $\Sigma_0 - \Sigma$ is a positive definite matrix.

For the problem with the two sided alternative $\Sigma \neq \Sigma_0$, Nishida [19] obtained a class of admissible tests which includes the LRT and the modified LRT.

Let

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i, \quad S = \sum_{i=1}^{N}(X_i - \bar{X})(X_i - \bar{X})'. \tag{4.2}$$

Then we have the following theorem for (4.1).

THEOREM 4.1. *If* $p - 1 < r < n - p + 1$ *and* B *is a given positive definite matrix which satisfies* $B \geq \Sigma_0^{-1}$, *then a test with the critical region*

$$\operatorname{etr}\{(B - \Sigma_0^{-1})S\} \cdot |S|^r \leq c \tag{4.3}$$

is admissible Bayes for any c. *Here,* $B \geq \Sigma_0^{-1}$ *means that* $B - \Sigma_0^{-1}$ *is positive semidefinite.*

PROOF. At first, let us transfom X by an orthogonal matrix $T(N \times N)$ such that $XT = (Y_1, \ldots, Y_n, \sqrt{N}\bar{X})$. Here, $n = N - 1$ and the columns of XT have the same covariance matrix Σ. Further, we have $\mathrm{E}Y_i = 0$ for $i = 1, \ldots, n$. Putting the whole mass to Σ in (4.4) as a prior distribution for Σ, we apply Lemma 1. It is taken (e.g.) $C_0 = \frac{1}{2}\Sigma_0^{-1}$ in (2.6).

Choose t and integer q such that $q \geq p$, $t > -1$ and $r = q + t$. Then there exist such q and t if $p - 1 < r$. Under H_1, set

$$\Sigma^{-1} = B + \eta\eta', \tag{4.4}$$

where $\eta(p \times q)$ and

$$d\Pi_1^*(\eta)/d\eta = |\eta\eta'|^{t/2}|B + \eta\eta'|^{-n/2}. \tag{4.5}$$

By Lemma 3, this density is integrable under the conditions of the theorem. We have

$$\begin{aligned}
&\int |\Sigma|^{-n/2} \operatorname{etr}\left\{-\frac{1}{2}\Sigma^{-1}S\right\} d\Pi_1^*(\eta) \\
&= \operatorname{etr}\left\{-\frac{1}{2}BS\right\} \cdot \int |\eta\eta'|^{t/2} \operatorname{etr}\left\{-\frac{1}{2}\eta\eta' S\right\} d\eta \\
&= \operatorname{etr}\left\{-\frac{1}{2}BS\right\} \cdot |S|^{-(t+q)/2} \cdot \int |\eta^*\eta^{*\prime}|^{t/2} \operatorname{etr}\left\{-\frac{1}{2}\eta^*\eta^{*\prime}\right\} d\eta^*,
\end{aligned} \tag{4.6}$$

where $\eta^* = S^{1/2}\eta$. Since the integral of the last line of (4.6) is constant, we obtain the statistic

$$\operatorname{etr}\left\{-\frac{1}{2}BS\right\} \cdot |S|^{-(t+q)/2}. \tag{4.7}$$

Consequently

$$
(4.8)\qquad \begin{aligned}
&\int f(X;\theta)d\Pi_1(\theta)\Big/\int f(X;\theta)d\Pi_0(\theta)\\
&=\int f(Y;\Sigma)d\Pi_1^*(\eta)/\mathrm{etr}\left\{-\frac{1}{2}\Sigma_0^{-1}S\right\}\\
&=\mathrm{etr}\left\{-\frac{1}{2}(B-\Sigma_0^{-1})S\right\}\cdot|S|^{-(t+q)/2}\geq c'
\end{aligned}
$$

is admissible Bayes critical region, which is identical to (4.3).

Setting $B=(1+u)\Sigma_0^{-1}(u\geq 0)$ and $\alpha=u/r$ in the above theorem, we obtain the following corollary.

COROLLARY 4.1.1. *If* $n>2(p-1)$ *and* $\alpha\geq 0$, *then a test with the critical region*

$$
(4.9)\qquad \{\mathrm{etr}\,\Sigma_0^{-1}S\}^{\alpha}|S|\leq c
$$

is admissible Bayes for any c.

Anderson and Gupta [4] considered to test H_0 against the altenatives defined by

$$
(4.10)\qquad H_1^*:\gamma_p\geq 1\quad\text{and}\quad \textstyle\sum_{i=1}^{p}\gamma_i>p,
$$

where γ_i's ($\gamma_1\geq\cdots\geq\gamma_p$) are the characteristic roots of $\Sigma_0^{-1}\Sigma$. The alternatives mean that Σ is larger than Σ_0 in a sense. By using their result (p. 1063), it can be shown that (4.9) is an acceptance region for their problem which has monotonicity property.

REMARK 4.1. In the case that μ is a known vector, it is easily shown that the corresponding theorem and corollary hold. We have only to exchange n and S by N and $S^*=\sum_{i=1}^{N}(X_i-\mu)(X_i-\mu)'$, respectively. For example, if $N>2(p-1)$ and $\alpha\geq 0$, then a test with the critical region

$$
(4.11)\qquad \{\mathrm{etr}\,\Sigma_0^{-1}S^*\}^{\alpha}|S^*|\leq c
$$

is admissible Bayes for any c.

4.2. One sided test for two samples case

Suppose that $X(p\times N_1)=(X_1,\dots,X_{N_1})$ and $Y(p\times N_2)=(Y_1,\dots,Y_{N_2})$ are random samples from $N_p(\mu_1,\Sigma_1)$ and $N_p(\mu_2,\Sigma_2)$, respectively. We consider the problem of testing

$$
(4.12)\qquad H_0:\Sigma_1=\Sigma_2\quad\text{against}\quad H_1:\Sigma_1<\Sigma_2.
$$

Define $\bar{X}$ and S_1 by

$$\bar{X} = \frac{1}{N_1}\sum_{i=1}^{N_1} X_i, \quad S_1 = \sum_{i=1}^{N_1}(X_i - \bar{X})(X_i - \bar{X})'. \tag{4.13}$$

Further, define $\bar{Y}$ and S_2 similarly for the second sample. Let $n_i = N_i - 1$ as usual. Then we have the following theorem.

THEOREM 4.2. *If $p-1 < r < n_1 + n_2 - p + 1$, $p-1 < r_1 < n_2$ and $p-1 < r_2 < n_1 - p + 1$, then a test with the critical region*

$$|S_1 + S_2|^{r-r_1}/|S_1|^{r_2} \geq c \tag{4.14}$$

is admissible Bayes for any c.

PROOF. After transforming X and Y by orthogonal matrices analogous to the previous subsection, we use Lemma 1 for $\sqrt{N_1}\bar{X}$ and $\sqrt{N_2}\bar{Y}$. Set

$$\Sigma_1^{-1} = \Sigma_2^{-1} = \Sigma^{-1} = I_p + \eta\eta' \quad \text{with} \quad \eta(p \times q) \tag{4.15}$$

and

$$d\Pi_0^*(\eta)/d\eta = |\eta\eta'|^{t/2}|I_p + \eta\eta'|^{-(n_1+n_2)/2} \tag{4.16}$$

under H_0. Here, t and integer q are chosen such that $q \geq p$, $t > -1$ and $t + q = r$. Then we have

$$\begin{aligned} &\int |\Sigma|^{-(n_1+n_2)/2} \operatorname{etr}\left\{-\frac{1}{2}\Sigma^{-1}(S_1 + S_2)\right\} d\Pi_0^*(\eta) \\ &= \operatorname{etr}\left\{-\frac{1}{2}(S_1 + S_2)\right\} \cdot \int |\eta\eta'|^{t/2} \operatorname{etr}\left\{-\frac{1}{2}\eta\eta'(S_1 + S_2)\right\} d\eta \\ &= \operatorname{etr}\left\{-\frac{1}{2}(S_1 + S_2)\right\} \cdot |S_1 + S_2|^{-(t+q)/2}. \end{aligned} \tag{4.17}$$

Under H_1, set

$$\Sigma_2^{-1} = I_p + \eta_1\eta_1' \quad \text{with} \quad \eta_1(p \times q_1), \tag{4.18}$$

$$\Sigma_1^{-1} = I_p + \eta_1\eta_1' + \eta_2\eta_2' \quad \text{with} \quad \eta_2(p \times q_2) \tag{4.19}$$

and

$$\begin{aligned} &d\Pi_1^*(\eta)/d\eta \\ &= |\eta_1\eta_1'|^{t_1/2}|\eta_2\eta_2'|^{t_2/2} \cdot |I_p + \eta_1\eta_1' + \eta_2\eta_2'|^{-n_1/2}|I_p + \eta_1\eta_1'|^{-n_2/2}. \end{aligned} \tag{4.20}$$

Here, t_i's and q_i's are chosen to satisfy $q_i \geq p$, $t_i > -1$ and $t_i + q_i = r_i$. Then we have

$$\begin{aligned}
&\int |\Sigma_1|^{-n_1/2}|\Sigma_2|^{-n_2/2}\,\mathrm{etr}\left\{-\frac{1}{2}(\Sigma_1^{-1}S_1+\Sigma_2^{-1}S_2)\right\}d\Pi_1^*(\eta)\\
(4.21)\quad &= \mathrm{etr}\left\{-\frac{1}{2}(S_1+S_2)\right\}\cdot\int|\eta_1\eta_1'|^{t_1/2}|\eta_2\eta_2'|^{t_2/2}\\
&\qquad\cdot\,\mathrm{etr}\left(-\frac{1}{2}\{\eta_1\eta_1'(S_1+S_2)+\eta_2\eta_2'S_1\}\right)d\eta_1 d\eta_2\\
&= \mathrm{etr}\left\{-\frac{1}{2}(S_1+S_2)\right\}\cdot|S_1+S_2|^{-(t_1+q_1)/2}|S_1|^{-(t_2+q_2)/2}.
\end{aligned}$$

So, we obtain

$$(4.22)\qquad \frac{\int f(X, Y;\theta)d\Pi_1(\theta)}{\int f(X, Y;\theta)\,d\Pi_0(\theta)} = |S_1+S_2|^{(r-r_1)/2}/|S_1|^{r_2/2},$$

which implies the theorem. The integrability of the densities (4.16) and (4.20) is assured by Lemmas 3 and 4.

COROLLARY 4.2.1. *If* $n_1 > 2(p-1)$ *and* $n_2 > p-1$, *then a test with the critical region*

$$(4.23)\qquad |S_1+S_2|/|S_1| = |I_p + S_1^{-1}S_2| \geq c$$

is admissible Bayes for any c.

PROOF. Set $r_1 = r_2 = d$ and $r = 2d$ where d is chosen as slightly larger than $p-1$, then we obtain (4.23). Further, these r_1, r_2 and r satisfy the integrability conditions.

Anderson and Gupta [4] also considered to test H_0 against the alternatives (4.10), where γ_i's ($\gamma_1 \geq \cdots \geq \gamma_p$) are the chracteristic roots of $\Sigma_1^{-1}\Sigma_2$. They obtained a class of tests which have the monotonicity property. Since $H_1 \subset H_1^*$, the above theorem and corollary hold for their problem (see Remark 5.1). It can be shown that the test (4.22) is contained in their class.

REMARK 4.2. Strictly speaking, we must determine how Σ is (or Σ's are) set in the prior distribution before applying Lemma 1, as we have done in the previous subsection. However, in order to make the argument simple we beforehand delete the variables (like $\sqrt{N_1}\bar{X}$ and $\sqrt{N_2}\bar{Y}$ in this subsection) hereafter, on the premise that Σ's will be set later as $\Sigma^{-1} = C_0 + D_i$ for some C_0 and D_i.

REMARK 4.3. When the mean vectors μ_1 and μ_2 are known, we can obtain the similar theorem and corollary by a slight modification of the above argument. Let

$$S_1^* = \sum_{i=1}^{N_1}(X_i - \mu_1)(X_i - \mu_1)' \quad \text{and} \quad S_2^* = \sum_{i=1}^{N_2}(Y_i - \mu_2)(Y_i - \mu_2)', \tag{4.24}$$

then in the propositions corresponding to Theorem 4.2 and Corollary 4.2.1, n_1, n_2, S_1, S_2 should be exchanged by N_1, N_2, S_1^*, S_2^*, respectively. For example, if $\min(N_1, N_2) > 2(p-1)$

$$|S_1^* + S_2^*|/|S_1^*| = |I_p + S_1^{*-1}S_2^*| \geq c \tag{4.25}$$

is an admissible Bayes critical region for any c. We can easily obtain similar modifications for propositions appeared in the following sections, so, we do not mention such modifications hereafter.

4.3. Linear structure of the inverse matrix of the covariance matrix

Let $X(p \times N) = (X_1, \ldots, X_N)$ be a random sample of size N from $N_p(\mu, \Sigma)$. We consider the problem of testing

$$H_0: \Sigma^{-1} = \sigma_0^2\Sigma_0 + \sigma_1^2 G_1 + \cdots + \sigma_k^2 G_k \text{ against } H_1: \text{not } H_0, \tag{4.26}$$

where Σ_0 is a given positive definite matrix and $G_1, \ldots, G_k$ are given positive semidefinite matrices. The multiples $\sigma_0^2, \sigma_1^2, \ldots, \sigma_k^2$ are unknown constants. This problem includes the sphericity problem and the intraclass correlation model as a special case. It is also regarded as a generalization of the problem which was considered in Kiefer and Schwartz [16].

The linear structure of the inverse matrix of the covariance matrix was considered by Anderson [1], [2]. The linear structure for the covariance matrix itself was considered in Bock and Bargmann [10], Srivastava [32] and Anderson [1], [2]. Krishnaiah and Lee [17] studied an extension of the problem. For a summary of these structures, see, e.g., Siotani et al. [31].

Let $\bar{X}$, S and n be the ones defined as usual (like in Subsection 4.1). Then, the following theorem holds.

THEOREM 4.3. *If $0 < r_0 < n$ and $p - 1 < r_1 < n - p + 1$, then a test with the critical region*

$$(\operatorname{tr}\{(a\Sigma_0 + b_1 G_1 + \cdots + b_k G_k)S\})^{r_0}/|S|^{r_1} \geq c \tag{4.27}$$

is admissible Bayes for any c, where a is any positive number and $b_1, \ldots, b_k$ are any nonnegative numbers.

PROOF. Let us transform X by an orthogonal matrix T such that $XT = (Z, \sqrt{N}\bar{X})$, where $\mathrm{E}Z = O(p \times n)$, and eliminate $\sqrt{N}\bar{X}$, μ by Lemma 1 with the following setting: Under H_0, set

$$\Sigma^{-1} = (1 + a\eta^2)\Sigma_0 + b_1\eta^2 G_1 + \cdots + b_k\eta^2 G_k \tag{4.28}$$

and

$$d\Pi_0^*(\eta)/d\eta = |(1 + a\eta^2)\Sigma_0 + b_1\eta^2 G_1 + \cdots + b_k\eta^2 G_k|^{-n/2}|\eta|^{r_0-1}, \tag{4.29}$$

where η is a scalar variable. Then we have

(4.30)

$$\begin{aligned}\int f(Z;\Sigma)d\Pi_0^*(\Sigma) &= \operatorname{etr}\left\{-\frac{1}{2}\Sigma_0^{-1}S\right\}\\ &\quad\cdot\int|\eta|^{r_0-1}\operatorname{etr}\left\{-\frac{1}{2}(a\Sigma_0 + b_1G_1 + \cdots + b_kG_k)S\eta^2\right\}d\eta\\ &= \operatorname{etr}\left\{-\frac{1}{2}\Sigma_0^{-1}S\right\}\cdot(\operatorname{tr}\{(a\Sigma_0 + b_1G_1 + \cdots + b_kG_k)S\})^{-r_0/2}.\end{aligned}$$

Under H_1, set

$$\Sigma^{-1} = \Sigma_0 + \eta_1\eta_1' \quad \text{with} \quad \eta_1(p \times q_1) \tag{4.31}$$

and

$$d\Pi_1^*(\eta)/d\eta = |\Sigma_0 + \eta_1\eta_1'|^{-n/2}|\eta_1\eta_1'|^{t_1/2}, \tag{4.32}$$

where $p \le q_1$, $t_1 > -1$ and $q_1 + t_1 = r_1$. Then we have

$$\begin{aligned}\int f(Z;\Sigma)d\Pi_1^*(\Sigma) &= \operatorname{etr}\left\{-\frac{1}{2}\Sigma_0^{-1}S\right\}\cdot\int|\eta_1\eta_1'|^{t_1/2}\operatorname{etr}\left\{-\frac{1}{2}\eta\eta' S\right\}d\eta_1\\ &= \operatorname{etr}\left\{-\frac{1}{2}\Sigma_0^{-1}S\right\}\cdot|S|^{-(t_1+q_1)/2}.\end{aligned} \tag{4.33}$$

Therefore

$$\begin{aligned}\frac{\int f(X;\theta)d\Pi_1(\theta)}{\int f(X;\theta)d\Pi_0(\theta)} &= \frac{\int f(Z;\Sigma)d\Pi_1^*(\Sigma)}{\int f(Z;\Sigma)d\Pi_0^*(\Sigma)}\\ &= (\operatorname{tr}\{(a\Sigma_0 + b_1G_1 + \cdots + b_kG_k)S\})^{r_0/2}/|S|^{r_1/2},\end{aligned} \tag{4.34}$$

which implies the theorem.

If $n > 2(p-1)$, it is possible to set $r_0 = r = d$ in the theorem, where d is a number which satisfies $p - 1 < d < n - p + 1$. Consequently, the following corollary holds.

COROLLARY 4.3.1. *If* $n > 2(p-1)$, *then a test with the critical region*

$$\mathrm{tr}\,\{(a\Sigma_0 + b_1 G_1 + \cdots + b_k G_k)S\}/|S| \geq c \tag{4.35}$$

is admissible Bayes for any positive constant c, *where* a *is any positive number and* $b_1, \dots, b_k$ *are any nonnegative numbers.*

5. Multiple decision problems for covariance matrices

5.1. Two samples case

Under the same situation (two normal populations and their random samples) as in Subsection 4.2, we consider the multiple decision problem of deciding whether of the following three hypotheses are true:

$$H_0 : \Sigma_1 = \Sigma_2, \quad H_1 : \Sigma_1 < \Sigma_2, \quad H_2 : \Sigma_1 > \Sigma_2. \tag{5.1}$$

Our interest is to obtain a class of admissible procedures. By using the same notations as in Subsection 4.2, the following theorem holds.

THEOREM 5.1. *If* $p-1 < r < n_1 + n_2 - p + 1$, $p - 1 < r_1 < n_2$, $p - 1 < r_2 < n_1 - p + 1$, $p - 1 < r_3 < n_2 - p + 1$ *and* $p - 1 < r_4 < n_1$, *then the procedure which selects* H_i *when* $T_i = \min_j T_j$ *is admissible Bayes, where*

$$T_0 = c_0 |S_1 + S_2|^r, \quad T_1 = c_1 |S_1 + S_2|^{r_1} |S_1|^{r_2}, \quad T_2 = c_2 |S_1 + S_2|^{r_4} |S_2|^{r_3} \tag{5.2}$$

and c_j*'s are any positive constants.*

PROOF. Consider the usual orthogonal transformations for X and Y, and use Lemma 1 like the way as that in Subsection 4.2. Under H_0, we use the same prior distribution as (4.16). So, we obtain the statistic

$$\int f(X, Y; \theta)\, d\Pi_0^*(\theta) = \mathrm{etr}\left\{-\frac{1}{2}(S_1 + S_2)\right\} \cdot |S_1 + S_2|^{-(t+q)/2}. \tag{5.3}$$

Under H_1, we also use the same prior distrbution as (4.20), which yields

(5.4)
$$\int f(X, Y; \theta) d\Pi_1^*(\theta) = \mathrm{etr}\left\{-\frac{1}{2}(S_1 + S_2)\right\} \cdot |S_1 + S_2|^{-(t_1+q_1)/2} |S_1|^{-(t_2+q_2)/2}.$$

Since the hypothesis H_2 is obtained by exchanging the suffixes in H_1, we use the prior distribution for H_2 which is obtained by exchanging the suffixes in that for H_1. Then we have

(5.5)

$$\int f(X, Y; \theta) d\Pi_2^*(\theta) = \operatorname{etr}\left\{-\frac{1}{2}(S_1 + S_2)\right\} \cdot |S_1 + S_2|^{-(t_4+q_4)/2} |S_2|^{-(t_3+q_3)/2}.$$

Therefore, the theorem follows from (2.2), by letting $r = q + t$ and $r_i = q_i + t_i$ $(i = 1, 2, 3, 4)$.

For this problem, Roy and Gnanadeskian [28] proposed a procedure based on the largest and smallest roots of $S_2^{-1}S_1$. However, their procedure is not contained in the class of Theorem 5.1 unless $p = 1$.

COROLLARY 5.1.1. *If* $\min(n_1, n_2) > 2(p-1)$, *then the procedure which selects* H_i *when* $U_i = \min_j U_j$ *is admissible Bayes, where*

(5.6) $$U_0 = c_0|S_1 + S_2|^{n_1+n_2}, \quad U_j = c_j|S_j|^{n_j}|S_1 + S_2|^{n_{3-j}} \qquad (j = 1, 2)$$

and c_j*'s are any positive constants.*

PROOF. Setting $r = c(n_1 + n_2)$, $r_1 = r_3 = cn_2$ and $r_2 = r_4 = cn_1$ in the theorem, where c is slightly larger than $(p-1)/\min(n_1, n_2)$, then we obtain the corollary.

COROLLARY 5.1.2. *If* $\min(n_1, n_2) > 2(p-1)$, *then the procedure which selects* H_i *when* $V_i = \max_j V_j$ *is admissible Bayes, where*

(5.7) $$V_0 = c_0, \quad V_1 = c_1|S_1 + S_2|/|S_1| = c_1|I_p + S_2S_1^{-1}|,$$
$$V_2 = c_2|S_1 + S_2|/|S_2| = c_2|I_p + S_1S_2^{-1}|$$

and c_j*'s are any positive constants.*

PROOF. Set $r_i = d$ $(i = 1, 2, 3, 4)$ and $r = 2d$ in the theorem, where d is slightly larger than $p - 1$. Then we obtain the rule which essentially coincides with that of the corollary.

COROLLARY 5.1.3. *If* $\min(n_1, n_2) > 2(p-1)$, *then the procedure which selects* H_i *when* $W_i = \min_j W_j$ *is admissible Bayes, where*

(5.8) $$W_0 = c_0, \quad W_1 = c_1|S_1|, \quad W_2 = c_2|S_2|$$

and c_j*'s are any positive constants.*

PROOF. The corollary is obtained by letting $r = r_1 = r_4 = d_1$ and $r_2 = r_3 = d_2$ in the theorem, where d_1 and d_2 are chosen to be slightly larger than $p - 1$.

5.2. Three samples case

Suppose that $X(p \times N_1) = (X_1, \ldots, X_{N_1})$, $Y(p \times N_2) = (Y_1, \ldots, Y_{N_2})$ and $Z(p \times N_3) = (Z_1, \ldots, Z_{N_3})$ are random samples from $N_p(\mu_1, \Sigma_1)$, $N_p(\mu_2, \Sigma_2)$ and $N_p(\mu_3, \Sigma_3)$, respectively. Let $\bar{X}$, $\bar{Y}$, $\bar{Z}$, S_1, S_2 and S_3 be the ones analogous to (4.13). Further, let $S = S_1 + S_2 + S_3$ and $n = n_1 + n_2 + n_3$ $(n_i = N_i - 1)$. We consider the following three multiple decision problems:

$$\begin{aligned} &H_0: \Sigma_1 = \Sigma_2 = \Sigma_3, \quad H_1: \Sigma_1 \neq \Sigma_2 = \Sigma_3, \quad H_2: \Sigma_2 \neq \Sigma_1 = \Sigma_3, \\ &H_3: \Sigma_3 \neq \Sigma_1 = \Sigma_2 \text{ and } H_4: \Sigma_i \neq \Sigma_j \ (i \neq j). \end{aligned} \tag{5.9}$$

$$\begin{aligned} &H_0: \Sigma_1 = \Sigma_2 = \Sigma_3, \quad H_1: \Sigma_1 < \Sigma_2 = \Sigma_3, \quad H_2: \Sigma_2 < \Sigma_1 = \Sigma_3 \\ &\text{and } H_3: \Sigma_3 < \Sigma_1 = \Sigma_2. \end{aligned} \tag{5.10}$$

$$\begin{aligned} &H_0: \Sigma_1 = \Sigma_2 = \Sigma_3, \quad H_1: \Sigma_1 > \Sigma_2 = \Sigma_3, \quad H_2: \Sigma_2 > \Sigma_1 = \Sigma_3 \\ &\text{and } H_3: \Sigma_3 > \Sigma_1 = \Sigma_2. \end{aligned} \tag{5.11}$$

At first, we transform samples by the usual orthogonal matrices and use Lemma 1. So, we can treat these problems without $\bar{X}$, $\bar{Y}$, $\bar{Z}$ and μ_i's.

THEOREM 5.2. *If* $p - 1 < r < n - p + 1$, $p - 1 < r_{2i-1} < n_i - p + 1$, $p - 1 < r_{2i} < n - n_i - p + 1$ $(i = 1, 2, 3)$ *and* $p - 1 < r_{6+j} < n_j - p + 1$ $(j = 1, 2, 3)$, *then the procedure which selects* H_i *when* $T_i^{(1)} = \min_j T_j^{(1)}$ *is admissible Bayes for* (5.9), *where*

$$\begin{aligned} &T_0^{(1)} = c_0 |S|^r, \quad T_j^{(1)} = c_j |S_j|^{r_{2j-1}} |S - S_j|^{r_{2j}} \qquad (j = 1, 2, 3), \\ &T_4^{(1)} = c_4 |S_1|^{r_7} \cdot |S_2|^{r_8} \cdot |S_3|^{r_9}, \end{aligned} \tag{5.12}$$

and c_j*'s are any positive constants.*

PROOF. Under H_0, set

$$\Sigma_1 = \Sigma_2 = \Sigma_3 = \Sigma = (I_p + \eta\eta')^{-1} \quad \text{with} \quad \eta(p \times q) \tag{5.13}$$

and

$$d\Pi_0^*(\eta)/d\eta = |I_p + \eta\eta'|^{-(n_1+n_2+n_3)/2} |\eta\eta'|^{t/2}, \tag{5.14}$$

where t and integer q are chosen such that $q \geq p$, $t > -1$ and $q + t = r$. Then we have

$$\begin{aligned} &\int |\Sigma|^{-(n_1+n_2+n_3)/2} \operatorname{etr}\left\{-\frac{1}{2}\Sigma^{-1}(S_1 + S_2 + S_3)\right\} d\Pi_0^*(\eta) \\ &= \operatorname{etr}\left\{-\frac{1}{2}S\right\} \cdot \int |\eta\eta'|^{t/2} \operatorname{etr}\left\{-\frac{1}{2}\eta\eta' S\right\} d\eta \end{aligned} \tag{5.15}$$

$$= \operatorname{etr}\left\{-\frac{1}{2}S\right\}\cdot|S|^{-(t+q)/2}.$$

Under H_1, set

$$\Sigma_1 = (I_p + \eta_1\eta_1')^{-1} \quad \text{with} \quad \eta(p\times q_1), \tag{5.16}$$

$$\Sigma_2 = \Sigma_3 = (I_p + \eta_2\eta_2')^{-1} \quad \text{with} \quad \eta(p\times q_2) \tag{5.17}$$

and

$$d\Pi_1^*(\eta)/d\eta = |I_p + \eta_1\eta_1'|^{-n_1/2}|I_p + \eta_2\eta_2'|^{-(n_2+n_3)/2}\cdot|\eta_1\eta_1'|^{t_1/2}|\eta_2\eta_2'|^{t_2/2}, \tag{5.18}$$

where t_i's and r_i's are required to satisfy the restrictions $q_i \geq p$, $t_i > -1$ and $r_i = t_i + q_i$ (these restrictions appeared several times before. We neglect hereafter to mention such restrictions explicitly. However, t_i's and q_i's can be choosen to satisfy these restrictions under the conditions in the theorems). Then we have

$$\int|\Sigma_1|^{-n_1/2}|\Sigma_2|^{-(n_2+n_3)/2}\operatorname{etr}\left(-\frac{1}{2}\{\Sigma_1^{-1}S_1 + \Sigma_2^{-1}(S_2+S_3)\}\right)d\Pi_1^*(\eta) \tag{5.19}$$

$$= \operatorname{etr}\left\{-\frac{1}{2}S\right\}\cdot|S_1|^{-(t_1+q_1)/2}|S_2+S_3|^{-(t_2+q_2)/2}.$$

Under H_2 and H_3, we consider prior distributions similar to the one under H_1. This gives the statistics

$$\operatorname{etr}\left\{-\frac{1}{2}S\right\}\cdot|S_j|^{-(t_{2j-1}+q_{2j-1})/2}|S-S_j|^{-(t_{2j}+q_{2j})/2} \qquad (j=2,3). \tag{5.20}$$

Under H_4, set

$$\Sigma_i = (I_p + \eta_i\eta_i')^{-1} \quad \text{with} \quad \eta(p\times q_{6+i}) \tag{5.21}$$

for $i = 1, 2, 3$ and

$$d\Pi_4^*(\eta)/d\eta = \textstyle\prod_{i=1}^3(|I_p + \eta_i\eta_i'|^{-n_i/2}|\eta_i\eta_i'|^{t_{6+i}/2}). \tag{5.22}$$

Then we have

$$\int\textstyle\prod_{i=1}^3\left(|\Sigma_i|^{-n_i/2}\operatorname{etr}\left\{-\frac{1}{2}\Sigma_i^{-1}S_i\right\}\right)d\Pi_4^*(\eta) \tag{5.23}$$

$$= \operatorname{etr}\left\{-\frac{1}{2}S\right\}\cdot\textstyle\prod_{i=1}^3(|S_i|^{-(t_{6+i}+q_{6+i})/2}).$$

These results imply the theorem.

COROLLARY 5.2.1. *If* $\min_j n_j > 2(p-1)$, *then the procedure which selects* H_i *when* $V_i^{(1)} = \min_j V_j^{(1)}$ *is admissible Bayes for* (5.9), *where*

$$(5.24)\qquad \begin{aligned} &V_0^{(1)} = c_0|S|^n, \quad V_j^{(1)} = c_j|S_j|^{n_j}|S - S_j|^{n-n_j} \qquad (j = 1, 2, 3), \\ &V_4^{(1)} = c_4|S_1|^{n_1} \cdot |S_2|^{n_2} \cdot |S_3|^{n_3}, \end{aligned}$$

and c_j*'s are any positive constants. This is the modified ML procedure.*

PROOF. The corollary is obtained by letting $r = dn$, $r_{2i-1} = dn_i$, $r_{2i} = d(n - n_i)$ $(i = 1, 2, 3)$ and $r_{6+j} = dn_j$ $(j = 1, 2, 3)$, where d is slightly larger than $(p-1)/\min n_i$.

If n_i's and n are exchanged by N_i's and $N (N = N_1 + N_2 + N_3)$ in the above proof, we obtain the admissibility of the ML procedure. Such modification can be done for the first corollary of each theorem in this section. The above corollary can be regarded as the one corresponding to Corollary 5.1.1. The propositions which correspond to Corollary 5.1.2 and 5.1.3 can be also proved for Theorem 5.2.

Now we treat the multiple decision problems which are slightly modified from (5.9) as follows:

$$(5.25)\qquad \begin{aligned} &H_0: \Sigma_1 = \Sigma_2 = \Sigma_3, \quad H_1: \Sigma_1 \neq \Sigma_2 = \Sigma_3, \quad H_2: \Sigma_2 \neq \Sigma_1 = \Sigma_3 \\ &\qquad \text{and } H_3: \Sigma_3 \neq \Sigma_1 = \Sigma_2, \end{aligned}$$

and

$$(5.26)\qquad \begin{aligned} &H_1: \Sigma_1 \neq \Sigma_2 = \Sigma_3, \quad H_2: \Sigma_2 \neq \Sigma_1 = \Sigma_3, \quad H_3: \Sigma_3 \neq \Sigma_1 = \Sigma_2 \\ &\qquad \text{and } H_4: \Sigma_i \neq \Sigma_j \qquad (i \neq j). \end{aligned}$$

Similar admissible procedures for these problems are immediately given from the above theorem.

COROLLARY 5.2.2. *If* $p - 1 < r < n - p + 1$, $p - 1 < r_{2i-1} < n_i - p + 1$ *and* $p - 1 < r_{2i} < n - n_i - p + 1$ $(i = 1, 2, 3)$, *then the procedure which selects* H_i *when* $T_i^{(1)} = \min_j T_j^{(1)}$ *is admissible Bayes for* (5.25), *where* $T_j^{(1)}$ *is given by* (5.12) $(j = 0, 1, 2, 3)$.

COROLLARY 5.2.3. *If* $p - 1 < r_{2i-1} < n_i - p + 1$, $p - 1 < r_{2i} < n - n_i - p + 1$ $(i = 1, 2, 3)$ *and* $p - 1 < r_{6+j} < n_j - p + 1$ $(j = 1, 2, 3)$, *then the procedure which selects* H_i *when* $T_i^{(1)} = \min_j T_j^{(1)}$ *is admissible Bayes for* (5.26), *where* $T_j^{(1)}$ *is given by* (5.12) $(j = 1, 2, 3, 4)$.

These corollaries are easily proved by putting probability 0 to H_4 or to H_0 in the proof of Theorem 5.2. The propositions which correspond to Corollary 5.2.1 (etc.) also hold. Thus, for a problem which is given by eliminating a hypothesis (or hypotheses) from a multiple decision problem, an admissible Bayes procedure can be immediately obtained from an admissible Bayes procedure for the original problem.

THEOREM 5.3. *If* $p-1<r<n-p+1$, $p-1<r_{2j-1}<n-n_j$, $p-1<r_{2j}<n_j-p+1$ $(j=1, 2, 3)$, *then the procedure which selects* H_i *when* $T_i^{(2)}=\min_j T_j^{(2)}$ *is admissible Bayes for* (5.10), *where*

$$T_0^{(2)}=c_0|S|^r, \quad T_j^{(2)}=c_j|S|^{r_{2j-1}}|S_j|^{r_{2j}} \qquad (j=1, 2, 3) \tag{5.27}$$

and c_j*'s are any positive constants.*

PROOF. Under H_0, we use the same prior distribution as the one defined by (5.13) and (5.14). So, we obtain the statistic

$$\mathrm{etr}\left\{-\frac{1}{2}S\right\}\cdot|S|^{-(t+q)/2}. \tag{5.28}$$

Under H_1, set

$$\Sigma_2^{-1}=\Sigma_3^{-1}=I_p+\eta_1\eta_1' \quad \text{with} \quad \eta_1(p\times q_1), \tag{5.29}$$

$$\Sigma_1^{-1}=I_p+\eta_1\eta_1'+\eta_2\eta_2' \quad \text{with} \quad \eta_2(p\times q_2) \tag{5.30}$$

and

$$\begin{aligned} d\Pi_1^*(\eta)/d\eta &= |\eta_1\eta_1'|^{t_1/2}|\eta_2\eta_2'|^{t_2/2} \\ &\quad\cdot|I_p+\eta_1\eta_1'+\eta_2\eta_2'|^{-n_1/2}|I_p+\eta_1\eta_1'|^{-(n_2+n_3)/2}. \end{aligned} \tag{5.31}$$

The integrability of (5.31) is assured by Lemma 4. Then

$$\begin{aligned} &\int|\Sigma_1|^{-n_1/2}|\Sigma_2|^{-(n_2+n_3)/2}\,\mathrm{etr}\left(-\frac{1}{2}\{\Sigma_1^{-1}S_1+\Sigma_2^{-1}(S_2+S_3)\}\right)d\Pi_1^*(\eta) \\ &=\mathrm{etr}\left\{-\frac{1}{2}(S_1+S_2+S_3)\right\}\cdot\int|\eta_1\eta_1'|^{t_1/2}|\eta_2\eta_2'|^{t_2/2} \\ &\qquad\cdot\mathrm{etr}\left(-\frac{1}{2}\{\eta_1\eta_1'(S_1+S_2+S_3)+\eta_2\eta_2'S_1\}\right)d\eta_1 d\eta_2 \\ &=\mathrm{etr}\left\{-\frac{1}{2}S\right\}\cdot|S|^{-(t_1+q_1)/2}|S_1|^{-(t_2+q_2)/2}. \end{aligned} \tag{5.32}$$

Under H_2 and H_3, considering the prior distributions similar to the one under

H_1, we obtain the theorem.

The following three corollaries are easily proved by the same technique as in deriving Corollary 5.1.1, 5.1.2 and 5.1.3.

COROLLARY 5.3.1. *If* $\min_j n_j > 2(p-1)$, *then the procedure which selects* H_i *when* $U_i^{(2)} = \min_j U_j^{(2)}$ *is admissible Bayes for* (5.10), *where*

$$U_0^{(2)} = c_0 |S|^n, \quad U_j^{(2)} = c_j |S_j|^{n_j} \cdot |S|^{n-n_j} \qquad (j = 1, 2, 3), \tag{5.33}$$

and c_j*'s are any positive constants.*

COROLLARY 5.3.2. *If* $\min_j n_j > 2(p-1)$, *then the procedure which selects* H_i *when* $V_i^{(2)} = \min_j V_j^{(2)}$ *is admissible Bayes for* (5.10), *where*

$$V_0^{(2)} = c_0 |S|, \quad V_j^{(2)} = c_j |S_j| \qquad (j = 1, 2, 3), \tag{5.34}$$

and c_j*'s are any positive constants.*

COROLLARY 5.3.3. *If* $\min_j n_j > 2(p-1)$, *then the procedure which selects* H_i *when* $W_i^{(2)} = \min_j W_j^{(2)}$ *is admissible Bayes for* (5.10), *where*

$$W_0^{(2)} = c_0, \quad W_j^{(2)} = c_j |S_j| \qquad (j = 1, 2, 3), \tag{5.35}$$

and c_j*'s are any positive constants.*

Similarly we can derive a class of admissible procedures for (5.11), which is given in the following theorem.

THEOREM 5.4. *If* $p-1 < r < n-p+1$, $p-1 < r_{2j-1} < n_j$, $p-1 < r_{2j} < n - n_j - p + 1$ $(j = 1, 2, 3)$, *then the procedure which selects* H_i *when* $T_i^{(3)} = \min_j T_j^{(3)}$ *is admissible Bayes for* (5.11), *where*

$$T_0^{(3)} = c_0 |S|^r, \quad T_j^{(3)} = c_j |S|^{r_{2j-1}} |S - S_j|^{r_{2j}} \qquad (j = 1, 2, 3) \tag{5.36}$$

and c_j*'s are any positive constants.*

PROOF. For this theorem, under H_0 we use the prior distribution which is defined by (5.13) and (5.14). Under H_1, set

$$\Sigma_1^{-1} = I_p + \eta_1 \eta_1' \quad \text{with} \quad \eta_1(p \times q_1), \tag{5.37}$$

$$\Sigma_2^{-1} = \Sigma_3^{-1} = I_p + \eta_1 \eta_1' + \eta_2 \eta_2' \quad \text{with} \quad \eta_2(p \times q_2) \tag{5.38}$$

and

$$d\Pi_1^*(\eta)/d\eta = |\eta_1\eta_1'|^{t_1/2}|\eta_2\eta_2'|^{t_2/2} \cdot |I_p + \eta_1\eta_1'|^{-n_1/2}|I_p + \eta_1\eta_1' + \eta_2\eta_2'|^{-(n_2+n_3)/2}. \tag{5.39}$$

Then

$$\int |\Sigma_1|^{-n_1/2}|\Sigma_2|^{-(n_2+n_3)/2} \operatorname{etr}\left(-\frac{1}{2}\{\Sigma_1^{-1}S_1 + \Sigma_2^{-1}(S_2+S_3)\}\right) d\Pi_1^*(\eta) = \operatorname{etr}\left\{-\frac{1}{2}S\right\} \cdot |S|^{-(t_1+q_1)/2}|S-S_1|^{-(t_2+q_2)/2}. \tag{5.40}$$

By the obivious exchange of the suffixes, we obtain the similar statistics for H_2 and H_3, which lead the theorem.

The following corollaries are also proved by a slight modification of the proofs of Corollary 5.1.1, 5.1.2 and 5.1.3.

COROLLARY 5.4.1. *If* $\min_j n_j > 2(p-1)$, *then the procedure which selects* H_i *when* $U_i^{(3)} = \min_j U_j^{(3)}$ *is admissible Bayes for* (5.11), *where*

$$U_0^{(3)} = c_0|S|^n, \quad U_j^{(3)} = c_j|S|^{n_j} \cdot |S-S_j|^{n-n_j} \qquad (j=1, 2, 3), \tag{5.41}$$

and c_j*'s are any positive constants.*

COROLLARY 5.4.2. *If* $\min_j n_j > 2(p-1)$, *then the procedure which selects* H_i *when* $V_i^{(3)} = \min_j V_j^{(3)}$ *is admissible Bayes for* (5.11), *where*

$$V_0^{(3)} = c_0|S|, \quad V_j^{(3)} = c_j|S-S_j| \qquad (j=1, 2, 3), \tag{5.42}$$

and c_j*'s are any positive constants.*

COROLLARY 5.4.3. *If* $\min_j n_j > 2(p-1)$, *then the procedure which selects* H_i *when* $W_i^{(3)} = \min_j W_j^{(3)}$ *is admissible Bayes for* (5.11), *where*

$$W_0^{(3)} = c_0, \quad W_j^{(3)} = c_j|S-S_j| \qquad (j=1, 2, 3), \tag{5.43}$$

and c_j*'s are any positive constants.*

There exist multiple decision problems which are more complicated (sometimes, consist of much more hypotheses) than the ones treated in this paper. From mathematical viewpoint, it is not difficult to derive admissible procedures for such problems by using similar prior distributions (e.g., see (5.46) $\sim$ (5.49)). Integrability of densities can be assured by Lemma 4 or its generalizations. However, it is doubtful to use 0–1 loss function for such problems with many hypotheses. It will be required to use a more general loss function, such as in Eaton [12]. So, we treat only one more problem, that is

$$H_{ijl}: \Sigma_i > \Sigma_j > \Sigma_l \qquad (i, j, l = 1, 2, 3; i \neq j \neq l \neq i). \tag{5.44}$$

Then the following theorem holds for (5.44).

THEOREM 5.5. *If* $p - 1 < r_{ijl}(1) < n_i$, $p - 1 < r_{ijl}(2) < n_j$ *and* $p - 1 < r_{ijl}(3) < n_l - p + 1$ $(i, j, l = 1, 2, 3; i \neq j \neq l \neq i)$, *then the procedure which selects* $H_{i'j'l'}$ *when* $T_{i'j'l'} = \min T_{ijl}$ *is admissible Bayes, where*

$$T_{ijl} = c_{ijl}|S_1 + S_2 + S_3|^{r_{ijl}(1)}|S_j + S_l|^{r_{ijl}(2)}|S_l|^{r_{ijl}(3)} \tag{5.45}$$

and c_{ijl}*'s are any positive constants.*

PROOF. Let us use Lemma 1 at first, then we may consider the proof after removing sample means and population means. By the symmetry of hypotheses, we consider only the prior distribution for H_{123}.

Under H_{123}, set

$$\Sigma_1^{-1} = I_p + \eta_1\eta_1' \quad \text{with} \quad \eta_1(p \times q_1), \tag{5.46}$$

$$\Sigma_2^{-1} = I_p + \eta_1\eta_1' + \eta_2\eta_2' \quad \text{with} \quad \eta_2(p \times q_2) \tag{5.47}$$

$$\Sigma_3^{-1} = I_p + \eta_1\eta_1' + \eta_2\eta_2' + \eta_3\eta_3' \quad \text{with} \quad \eta_3(p \times q_3) \tag{5.48}$$

and

$$\begin{aligned} d\Pi_{123}^*(\eta)/d\eta &= |\eta_1\eta_1'|^{t_1/2}|\eta_2\eta_2'|^{t_2/2}|\eta_3\eta_3'|^{t_3/2}|I_p + \eta_1\eta_1'|^{-n_1/2} \\ &\quad \cdot|I_p + \eta_1\eta_1' + \eta_2\eta_2'|^{-n_2/2}|I_p + \eta_1\eta_1' + \eta_2\eta_2' + \eta_3\eta_3'|^{-n_3/2}. \end{aligned} \tag{5.49}$$

By Lemma 4, the density (5.49) is integrable under the conditions of Theorem 5.5. Further, we have

(5.50)

$$\begin{aligned} &\int |\Sigma_1|^{-n_1/2}|\Sigma_2|^{-n_2/2}|\Sigma_3|^{-n_3/2}\,\mathrm{etr}\left(-\frac{1}{2}\{\Sigma_1^{-1}S_1 + \Sigma_2^{-1}S_2 + \Sigma_3^{-1}S_3\}\right)d\Pi_{123}^*(\eta) \\ &= \mathrm{etr}\left\{-\frac{1}{2}(S_1 + S_2 + S_3)\right\}\cdot\int |\eta_1\eta_1'|^{t_1/2}|\eta_2\eta_2'|^{t_2/2}|\eta_3\eta_3'|^{t_3/2} \\ &\quad \cdot\mathrm{etr}\left(-\frac{1}{2}\{\eta_1\eta_1'(S_1 + S_2 + S_3) + \eta_2\eta_2'(S_2 + S_3) + \eta_3\eta_3'S_3\}\right)d\eta \end{aligned}$$

$$= \operatorname{etr}\left\{-\frac{1}{2}S\right\} \cdot |S_1 + S_2 + S_3|^{-(t_1+q_1)/2} |S_2 + S_3|^{-(t_2+q_2)/2} |S_3|^{-(t_3+q_3)/2}.$$

Therefore, by putting $r_{123}(i) = t_i + q_i$ we obtain the statistic T_{123}, which leads the theorem.

The following two corollaries can be easily derived as special cases of the above theorem.

COROLLARY 5.5.1. *If* $\min n_i > 2(p-1)$, *then the procedure which selects* $H_{i'j'l'}$ *when* $U_{i'j'l'} = \min U_{ijl}$ *is admissible Bayes, where*

(5.51)
$$U_{ijl} = c_{ijl}|S_1 + S_2 + S_3|^{n_i}|S_j + S_i|^{n_j}|S_l|^{n_l} \qquad (i, j, l = 1, 2, 3;\ i \neq j \neq l \neq i)$$

and c_{ijl}*'s are any positive constants.*

COROLLARY 5.5.2. *If* $\min n_i > 2(p-1)$, *then the procedure which selects* $H_{i'j'l'}$ *when* $W_{i'j'l'} = \min W_{ijl}$ *is admissible Bayes, where*

(5.52)
$$W_{ijl} = c_{ijl}|S_j + S_l| \cdot |S_l| \qquad (i, j, l = 1, 2, 3;\ i \neq j \neq l \neq i)$$

and c_{ijl}*'s are any positive constants.*

REMARK 5.1. It is possible to treat the multiple decision problems whose hypotheses are described by the determinants of the covariance matrices (generalized variances). All propositions in this section also hold even if the covariance matrices are exchanged by their determinants in hypotheses. For example, the results of Theorem 5.1 and its corollaries also hold for the problem of deciding whether the following three hypotheses are true:

(5.53)
$$H_0^*: |\Sigma_1| = |\Sigma_2|, \quad H_1^*: |\Sigma_1| < |\Sigma_2|, \ H_2^*: |\Sigma_1| > |\Sigma_2|.$$

under same conditions. This can be easily shown by using the entirely same prior distributins as the one of Theorem 5.1. Because $H_i \subset H_i^*$, it is possible to consider the prior distribution on H_i^* with the whole mass for H_i. Here H_i's are given by (5.1). Of course, this argument holds for testing problems such as in Subsection 4.1 or 4.2.

6. Classification problem with unequal covariance matrices

6.1. Classification rules

Let us consider the classification problem with unequal covariance matrices. The p-variates normal population $N_p(\mu_i, \Sigma_i)$ is denoted by $\pi_i(i = 1, 2, 3)$. Suppose that $X(p \times N_1) = (X_1, \dots, X_{N_1})$, $Y(p \times N_2) = (Y_1, \dots,$

Y_{N_2}) and $Z(p \times N_3) = (Z_1, \ldots, Z_{N_3})$ are random samples from normal populations π_1, π_2, π_3, respectively. It is assumed that $\mu_1 \neq \mu_2$ and $\Sigma_1 \neq \Sigma_2$. Then we consider the problem of testing

$$(6.1) \qquad H_1: \mu_3 = \mu_1, \quad \Sigma_3 = \Sigma_1 \quad \text{against} \quad H_2: \mu_3 = \mu_2, \quad \Sigma_3 = \Sigma_2.$$

This problem is equivalent to classify a sample from π_3 as either π_1 or π_2. Kiefer and Schwartz [16] showed the admissibility of some procedure for the case $\Sigma_1 = \Sigma_2 = \Sigma_3$. Kanazawa [15] studied three classification rules for the unequal covariance matrices case with $N_3 = 1$, which are called classification rule-W, -Z and -B. We will extend these three procedures to the case $N_3 \geq 1$. First, some estimators are defined for describing the classification procedures.

$$(6.2) \qquad \begin{aligned} \bar{X} &= \frac{1}{N_1} \textstyle\sum_{i=1}^{N_1} X_i, \quad S_1 = \sum_{i=1}^{N_1} (X_i - \bar{X})(X_i - \bar{X})', \\ \bar{Y} &= \frac{1}{N_2} \textstyle\sum_{i=1}^{N_2} Y_i, \quad S_2 = \sum_{i=1}^{N_2} (Y_i - \bar{Y})(Y_i - \bar{Y})'. \end{aligned}$$

Under H_1, we may regard that (X, Z) is a random sample of size $N_1 + N_3$ from π_1. So, we can define the estimators of μ_1, Σ_1, μ_2 and Σ_2 under H_1 as

$$(6.3) \qquad \begin{aligned} &\bar{X}^{(1)} = \frac{1}{N_1 + N_3} \{\textstyle\sum_{i=1}^{N_1} X_i + \sum_{i=1}^{N_3} Z_i\}, \\ &S_1^{(1)} = \textstyle\sum_{i=1}^{N_1} (X_i - \bar{X}^{(1)})(X_i - \bar{X}^{(1)})' + \sum_{i=1}^{N_3} (Z_i - \bar{X}^{(1)})(Z_i - \bar{X}^{(1)})', \\ &\bar{Y}^{(1)} = \bar{Y}, \quad S_2^{(1)} = S_2. \end{aligned}$$

Under H_2, the estimators are also defined as

$$(6.4) \qquad \begin{aligned} &\bar{X}^{(2)} = \bar{X}, \quad S_1^{(2)} = S_1, \\ &\bar{Y}^{(2)} = \frac{1}{N_2 + N_3} \{\textstyle\sum_{i=1}^{N_2} Y_i + \sum_{i=1}^{N_3} Z_i\}, \\ &S_2^{(2)} = \textstyle\sum_{i=1}^{N_2} (Y_i - \bar{Y}^{(2)})(Y_i - \bar{Y}^{(2)})' + \sum_{i=1}^{N_3} (Z_i - \bar{Y}^{(2)})(Z_i - \bar{Y}^{(2)})'. \end{aligned}$$

Classification rule-W If the parameters μ_i's and Σ_i's are known ($i = 1, 2$), the ML classification rule is given by using the ratio of $f(Z; \mu_3, \Sigma_3)$ in H_1 and H_2, *i.e.*

$$(6.5) \qquad \begin{aligned} \lambda = \{|\Sigma_1| \cdot |\Sigma_2|^{-1}\}^{N_3/2} &\operatorname{etr}\left(\frac{1}{2} \Sigma_1^{-1} \{\textstyle\sum_{i=1}^{N_3} (Z_i - \mu_1)(Z_i - \mu_1)'\}\right) \\ &\cdot \operatorname{etr}\left(-\frac{1}{2} \Sigma_2^{-1} \{\textstyle\sum_{i=1}^{N_3} (Z_i - \mu_2)(Z_i - \mu_2)'\}\right). \end{aligned}$$

The rule may be also expressed by using

$$(6.6)\qquad -2\log\lambda = N_3\log|\Sigma_2| + \operatorname{tr}\Sigma_2^{-1}\{\textstyle\sum_{i=1}^{N_3}(Z_i-\mu_2)(Z_i-\mu_2)'\}$$
$$-N_3\log|\Sigma_1| - \operatorname{tr}\Sigma_1^{-1}\{\textstyle\sum_{i=1}^{N_3}(Z_i-\mu_1)(Z_i-\mu_1)'\}.$$

The classification rule-W is obtained by substituting the usual unbiased estimators for μ_1, μ_2, Σ_1 and Σ_2 into (6.6). Here, only the samples X and Y are used for the unbised estimators (*i.e.* (6.2) is used). Namely, the statistic

$$(6.7)\quad DW$$
$$= N_3\log|S_2/(N_2-1)| + \operatorname{tr}(\{S_2/(N_2-1)\}^{-1}\{\textstyle\sum_{i=1}^{N_3}(Z_i-\bar{Y})(Z_i-\bar{Y})'\})$$
$$-N_3\log|S_1/(N_1-1)| - \operatorname{tr}(\{S_1/(N_1-1)\}^{-1}\{\textstyle\sum_{i=1}^{N_3}(Z_i-\bar{X})(Z_i-\bar{X})'\})$$

is used as the sample ML classification rule. According to the value of this statistic, the rule is defined as

$$(6.8)\qquad \begin{array}{l}\text{choose } H_1 \text{ if } DW \geq 0, \text{ and}\\ \text{choose } H_2 \text{ if } DW < 0.\end{array}$$

We call this as classification rule-W. This is an extension of the sample ML rule for the case $N_3 = 1$ in [15].

Classification rule-Z The (exact) ML rule is obtained by substituting (6.3) under H_1 and (6.4) under H_2 to the parameters in the likelihood ratio function. Therefore, the ML rule is given by

$$(6.9)\qquad DZ = -2\log\lambda_z = N_1\log|S_1^{(2)}/N_1| + (N_2+N_3)\log|S_2^{(2)}/(N_2+N_3)|$$
$$-(N_1+N_3)\log|S_1^{(1)}/(N_1+N_3)| - N_2\log|S_2^{(1)}/N_2|,$$

where λ_Z is the likelihood ratio in this case. The rule is defined as

$$(6.10)\qquad \begin{array}{l}\text{choose } H_1 \text{ if } DZ \geq 0, \text{ and}\\ \text{choose } H_0 \text{ if } DZ < 0,\end{array}$$

which we call classification rule-Z.

In the special case $N_3 = 1$, we can express DZ as

$$(6.11)\quad DZ$$
$$= (N_2+1)\log\left\{1 + \frac{N_2}{N_2+1}(Z-\bar{Y})'S_2^{-1}(Z-\bar{Y})\right\} + \log|S_2|$$
$$-(N_1+1)\log\left\{1 + \frac{N_1}{N_1+1}(Z-\bar{X})'S_1^{-1}(Z-\bar{X})\right\} - \log|S_1|$$
$$+(N_1+1)p\log(N_1+1) + N_2 p\log N_2 - N_1 p\log N_1 - (N_2+1)p\log(N_2+1),$$

(cf. Anderson [3]).

Classification rule-B This rule is obtained along the Bayesian approach. Let $\theta = (\mu_1, \mu_2, \Sigma_1, \Sigma_2)$, and consider a prior density $\Pi(\theta)$ which is defined by

$$\Pi(\theta) \propto (|\Sigma_1| \cdot |\Sigma_2|)^{-(p+1)/2}. \tag{6.12}$$

The prior density $\Pi(\theta)$ is used in common for H_1 and H_2, but the parameters of π_3 are different for H_1 and for H_2. This prior distribution was adopted in Mardia et al. [18], and is not a finite measure. For this prior distribution, we calculate the improper Bayes procedure which we call classification rule-*B*.

We consider at first under H_1.

$$\begin{aligned} J_1 &= \int f(X; \mu_1, \Sigma_1) \cdot f(Y; \mu_2, \Sigma_2) \cdot f(Z; \mu_1, \Sigma_1) d\Pi(\theta) \\ &= c \int |\Sigma_1|^{-(N_1+N_3)/2} |\Sigma_2|^{-N_2/2} \operatorname{etr}\Big(-\frac{1}{2}\Sigma_1^{-1}\{\textstyle\sum_{i=1}^{N_1}(X_i - \mu_1)(X_i - \mu_1)' \\ &\quad + \textstyle\sum_{i=1}^{N_3}(Z_i - \mu_1)(Z_i - \mu_1)'\} - \frac{1}{2}\Sigma_2^{-1}\{\textstyle\sum_{i=1}^{N_2}(Y_i - \mu_2)(Y_i - \mu_2)'\}\Big) \\ &\quad \cdot (|\Sigma_1| \cdot |\Sigma_2|)^{-(p+1)/2} d\mu_1 d\mu_2 d\Sigma_1 d\Sigma_2 \\ &= c \int |\Sigma_1|^{-(N_1+N_3+p+1)/2} |\Sigma_2|^{-(N_2+p+1)/2} \\ &\quad \cdot \operatorname{etr}\Big(-\frac{1}{2}\Sigma_1^{-1}\{S_1^{(1)} + (N_1 + N_3)(\bar{X}^{(1)} - \mu_1)(\bar{X}^{(1)} - \mu_1)'\} \\ &\quad - \frac{1}{2}\Sigma_2^{-1}\{S_2^{(1)} + N_2(\bar{Y}^{(1)} - \mu_2)(\bar{Y}^{(1)} - \mu_2)'\}\Big) d\mu_1 d\mu_2 d\Sigma_1 d\Sigma_2, \end{aligned} \tag{6.13}$$

where $c = \{2\pi\}^{-p(N_1+N_2+N_3)/2}$. The part concerning the integration of μ_1 can be carried out as

$$\begin{aligned} &\int \operatorname{etr}\Big(-\frac{1}{2}\Sigma_1^{-1}\{(N_1 + N_3)(\bar{X}^{(1)} - \mu_1)(\bar{X}^{(1)} - \mu_1)'\}\Big) d\mu_1 \\ &\quad = (2\pi)^{p/2}(N_1 + N_3)^{-p/2}|\Sigma_1|^{1/2}. \end{aligned} \tag{6.14}$$

Analogously

$$\begin{aligned} &\int \operatorname{etr}\Big(-\frac{1}{2}\Sigma_2^{-1}\{N_2(\bar{Y}^{(1)} - \mu_2)(\bar{Y}^{(1)} - \mu_2)'\}\Big) d\mu_2 \\ &\quad = (2\pi)^{p/2} N_2^{-p/2}|\Sigma_2|^{1/2}. \end{aligned} \tag{6.15}$$

Substituting these results into J_1, we have

(6.16) $$J_1 = b(N_1 + N_3)^{-p/2} N_2^{-p/2} \int |\Sigma_1|^{-(N_1+N_3+p)/2} |\Sigma_2|^{-(N_2+p)/2} \cdot \operatorname{etr}\left(-\frac{1}{2}\{\Sigma_1^{-1} S_1^{(1)} + \Sigma_2^{-1} S_2^{(1)}\}\right) d\Sigma_1 d\Sigma_2,$$

where $b = \{2\pi\}^{-(N_1+N_2+N_3)/2+p}$. Let us consider the integration with respect to Σ_1. Transform Σ_1 to $\Sigma_1^* = \Sigma_1^{-1}$, then

(6.17) $$d\Sigma_1 = |\Sigma_1^*|^{-(p+1)} d\Sigma_1^*.$$

Hence

(6.18) $$\begin{aligned} &\int |\Sigma_1|^{-(N_1+N_3+p)/2} \operatorname{etr}\left\{-\frac{1}{2}\Sigma_1^{-1} S_1^{(1)}\right\} d\Sigma_1 \\ &= \int |\Sigma_1^*|^{\{(N_1+N_3-1)-p-1\}/2} \operatorname{etr}\left\{-\frac{1}{2}\Sigma_1^* S_1^{(1)}\right\} d\Sigma_1^* \\ &= 2^{(N_1+N_3-1)p/2} |S_1^{(1)}|^{-(N_1+N_3-1)/2} \Gamma_p\left(\frac{N_1+N_3-1}{2}\right), \end{aligned}$$

where Γ_p is the p-variates gamma function defined by

(6.19) $$\Gamma_p(t) = \pi^{p(p-1)/4} \cdot \Pi_{i=1}^p \Gamma\left(t - \frac{(i-1)}{2}\right).$$

It also holds by analogous calculation,

(6.20) $$\int |\Sigma_2|^{-(N_2+p)/2} \operatorname{etr}\left\{-\frac{1}{2}\Sigma_2^{-1} S_2^{(1)}\right\} d\Sigma_2 = 2^{(N_2-1)p/2} |S_2^{(1)}|^{-(N_2-1)/2} \Gamma_p\left(\frac{N_2-1}{2}\right).$$

Therefore

(6.21) $$J_1 = b \cdot 2^{(N_1+N_2+N_3-2)p/2} (N_1 + N_3)^{-p/2} N_2^{-p/2} \Gamma_p\left(\frac{N_1+N_3-1}{2}\right) \cdot \Gamma_p\left(\frac{N_2-1}{2}\right) \cdot |S_1^{(1)}|^{-(N_1+N_3-1)/2} |S_2^{(1)}|^{-(N_2-1)/2}.$$

Similarly it can be seen that

(6.22) $$J_2 = \int f(X; \mu_1, \Sigma_1) \cdot f(Y; \mu_2, \Sigma_2) \cdot f(Z; \mu_2, \Sigma_2) d\Pi(\theta)$$

$$= b \cdot 2^{(N_1+N_2+N_3-2)p/2} N_1^{-p/2} (N_2+N_3)^{-p/2} \Gamma_p\left(\frac{N_1-1}{2}\right)$$

$$\cdot \Gamma_p\left(\frac{N_2+N_3-1}{2}\right) \cdot |S_1^{(2)}|^{-(N_1-1)/2} |S_2^{(2)}|^{-(N_2+N_3-1)/2}.$$

Now we define DB as $DB = -2\log(J_2/J_1)$. Then

$$DB = 2\log c_2 + (N_1-1)\log|S_1^{(2)}| + (N_2+N_3-1)\log|S_2^{(2)}| - 2\log c_1 - (N_1+N_3-1)\log|S_1^{(1)}| - (N_2-1)\log|S_2^{(1)}|, \tag{6.23}$$

where

$$c_i = \left(\frac{N_i+N_3}{N_i}\right)^{p/2} \cdot \left\{\Gamma_p\left(\frac{N_i-1}{2}\right) \Big/ \Gamma_p\left(\frac{N_i+N_3-1}{2}\right)\right\} \tag{6.24}$$

for $i = 1, 2$. The improper Bayes procedure is given by

$$\begin{aligned} &\text{choose } H_1 \text{ if } DB \geq 0, \text{ and} \\ &\text{choose } H_2 \text{ if } DB < 0. \end{aligned} \tag{6.25}$$

For the special case $N_3 = 1$,

$$S_1^{(1)} = S_1 + \left\{\frac{N_1}{N_1+1}\right\}(Z-\bar{X})(Z-\bar{X})', \tag{6.26}$$

$$\Gamma_p\left(\frac{N_1-1}{2}\right) \Big/ \Gamma_p\left(\frac{N_1}{2}\right) = \Gamma\left(\frac{N_1-p}{2}\right) \Big/ \Gamma\left(\frac{N_1}{2}\right) \tag{6.27}$$

and

$$\begin{aligned} |S_1^{(2)}|^{-1}|S_1^{(1)}| &= \left| I_p + \frac{N_1}{N_1+1} S_1^{-1/2}(Z-\bar{X})(Z-\bar{X})' S_1^{-1/2} \right| \\ &= 1 + \frac{N_1}{N_1+1}(Z-\bar{X})' S_1^{-1}(Z-\bar{X}). \end{aligned} \tag{6.28}$$

Of course, the equalities which are obtained by exchanging the suffixes 1 and 2 also hold. Consequently

$$\begin{aligned} DB &= 2\log c_2 + \log|S_2| + N_2 \log\left(1 + \frac{N_2}{N_2+1}(Z-\bar{Y})' S_2^{-1}(Z-\bar{Y})\right) \\ &\quad - 2\log c_1 - \log|S_1| - N_1 \log\left(1 + \frac{N_1}{N_1+1}(Z-\bar{X})' S_1^{-1}(Z-\bar{X})\right) \end{aligned} \tag{6.29}$$

for $N_3 = 1$. In this case, c_i reduces to

$$(6.30) \qquad c_i = \left(\frac{N_i + 1}{N_i}\right)^{p/2} \cdot \left\{\Gamma\left(\frac{N_i - p}{2}\right) / \Gamma\left(\frac{N_i}{2}\right)\right\}$$

The expression (6.29) coincides with that of Kanazawa [15].

REMARK 6.1. To derive (6.29), Kanazawa used the function

$$(6.31) \qquad f_i(Z|X, Y) = \int f(Z; \mu_i, \Sigma_i) \cdot f(\theta|X, Y) d\mu_1 d\mu_2 d\Sigma_1 d\Sigma_2$$

which is the conditional likelihood of Z under H_i given (X, Y), $i = 1, 2$. The posterior density of θ given (X, Y) is defined by

$$(6.32) \qquad f(\theta|X, Y) = \frac{f(X, Y; \theta)\Pi(\theta)}{\int f(X, Y; \theta)\Pi(\theta) d\mu_1 d\mu_2 d\Sigma_1 d\Sigma_2}.$$

The Bayes classification procedure is determined by

$$(6.33) \qquad \text{choose } H_1 \text{ or } H_2 \text{ according as } f_1(Z|X, Y) > \text{ or } < f_2(Z|X, Y).$$

Of course, this approach is equivalent to that of this paper.

REMARK 6.2. We can treat k-samples case similarly. That is, $\pi_1, \pi_2, \ldots, \pi_k, \pi_{k+1}$ are p-variates normal populations and consider the hypotheses

$$(6.34) \qquad H_i: \mu_{k+1} = \mu_i, \Sigma_{k+1} = \Sigma_i \quad (i = 1, \ldots, k).$$

Then the rule-Z (for example) is described as follows, by using similar notations. We can say that

$$H_i \text{ is preferable than } H_j \text{ if and only if } DZ(i, j) > 0,$$

where

$$(6.35) \qquad DZ(i, j) = N_i \log |S_i^{(j)}/N_i| + (N_j + N_{k+1}) \log |S_j^{(j)}/(N_j + N_{k+1})| \\ - (N_i + N_{k+1}) \log |S_i^{(i)}/(N_i + N_{k+1})| - N_j \log |S_j^{(i)}/N_j|.$$

Then we select H_i when H_i is preferable than any other H_j. Rule-W and rule-B can be described analogously. For simplicity, in the following subsections we state admissibility or other results only for the case $k = 2$, however, the results hold quite similarly for the case $k \geq 3$.

6.2. Admissible classification rules

For the problem which is slightly general than that of the previous subsection, Nishida [19] obtained a class of admissible rules. For the problem of this section, the theorem in [19] can be described as follows:

THEOREM 6.1. *If* $p-1<r_1<(N_1+N_3-1)-p+1$, $p-1<r_2<(N_2-1)-p+1$, $p-1<r_3<(N_1-1)-p+1$ *and* $p-1<r_4<(N_2+N_3-1)-p+1$, *then the classification rule*:

choose H_1 *or* H_2 *according as*

$$|S_1^{(2)}|^{r_3}|S_2^{(2)}|^{r_4}/|S_1^{(1)}|^{r_1}|S_2^{(1)}|^{r_2} > or < c \tag{6.36}$$

is admissible Bayes for any c.

As a special case of this theorem, rule-Z is shown to be admissible if $\min(N_1-1, N_2-1) > 2(p-1)$. Further, using this theorem, we can also derive that rule-B is a dmissible.

COROLLARY 6.1.1. *If* $\min(N_1-1, N_2-1) > 2(p-1)$, *then the classification rule-B is admissible Bayes.*

PROOF. Choose a constant d as slightly larger than $(p-1)/\min(N_1-1, N_2-1)$. Then $d<1$. Setting

$$r_1 = d(N_1+N_3-1),\ r_2 = d(N_2-1),\ r_3 = d(N_1-1),\ r_4 = d(N_2+N_3-1), \tag{6.37}$$

we obtain the corollary. The conditions for r_i's in the theorem are satisfied if $\min(N_1-1, N_2-1) > 2(p-1)$. This is shown by using the fact that the length of the intervals for r_i's are longer than or equal to 1 and that $d<1$.

REMARK 6.3. The condition $\min(N_1-1, N_2-1) > 2(p-1)$ in the above corollary is usually regarded as that for N_i's. However, the condition may be regarded as that for p. That is, if N_i's are not large, then p should be taken a small value for the validity of admissibility.

6.3. The limiting distribution of the classification rules

To study the limiting distributions of the classification statistics DW, DZ and DB, we put

$$dw(1) = N_3 \log|S_1/(N_1-1)| + \sum_{i=1}^{N_3}(Z_i-\bar{X})'\{S_1/(N_1-1)\}^{-1}(Z_i-\bar{X}), \tag{6.38}$$

$$dw(2) = N_3 \log|S_2/(N_2-1)| + \sum_{i=1}^{N_3}(Z_i-\bar{Y})'\{S_2/(N_2-1)\}^{-1}(Z_i-\bar{Y}), \tag{6.39}$$

(6.40) $dz(j)$

$$= (N_j + N_3)\log|S_j^{(j)}/(N_j + N_3)| - N_j\log|S_j^{(3-j)}/N_j|$$

and

(6.41) $db(j)$

$$= 2\log c_j + (N_j + N_3 - 1)\log|S_j^{(j)}| - (N_j - 1)\log|S_j^{(3-j)}|$$

for $j = 1, 2$. Of course, it holds that

$$DW = dw(2) - dw(1),$$
$$DZ = dz(2) - dz(1), \tag{6.42}$$
$$DB = db(2) - db(1).$$

If N_1 increases to ∞, $\bar{X} \to \mu_1$ and $S_1/(N_1 - 1) \to \Sigma_1$ in probability, respectively. So, it follows that

$$dw(1) \longrightarrow N_3 \log|\Sigma_1| + \sum_{i=1}^{N_3}(Z_i - \mu_1)'\Sigma_1^{-1}(Z_i - \mu_1) \tag{6.43}$$

in probability as $N_1 \to \infty$. Let

$$D(Z, j) = \sum_{i=1}^{N_3}(Z_i - \mu_j)'\Sigma_j^{-1}(Z_i - \mu_j) \qquad (j = 1, 2). \tag{6.44}$$

Then it is well known that $D(Z, j)$ is distributed as a chi-squre distribution χ_f^2 with $f = pN_3$ under H_j. Similarly, it can be shown that

$$dw(2) \longrightarrow N_3 \log|\Sigma_2| + D(Z, 2) \tag{6.45}$$

in probability as $N_2 \to \infty$. Now

(6.46) $dz(1)$

$$= N_3 \log|S_1^{(1)}/(N_1 + N_3)| + N_1\{\log|S_1^{(1)}/(N_1 + N_3)| - \log|S_1^{(2)}/N_1|\}$$
$$= N_3 \log|S_1^{(1)}/(N_1 + N_3)| + N_1(\log|\{S_1^{(2)}\}^{-1}S_1^{(1)}| + p\log\{N_1/(N_1 + N_3)\}).$$

Since

(6.47)

$$S_1^{(1)} = S_1^{(2)} + \sum_{i=1}^{N_3}(Z_i - \bar{Z})(Z_i - \bar{Z})' + \{N_1N_3/(N_1 + N_3)\}(\bar{Z} - \bar{X})(\bar{Z} - \bar{X})'$$

and

$$|I_p + A/n| = 1 + \operatorname{tr}(A/n) + O(n^{-2}), \tag{6.48}$$

it follows that

(6.49) $dz(1)$

$$= N_3|S_1^{(1)}/(N_1 + N_3)| + N_1\left(\frac{1}{N_1}\operatorname{tr}\sum_{i=1}^{N_3}\{S_1^{(2)}/N_1\}^{-1}(Z_i - \bar{Z})(Z_i - \bar{Z})'\right.$$

$$+\frac{N_3}{N_1}\operatorname{tr}\{S_1^{(2)}/N_1\}^{-1}(\bar{Z}-\bar{X})(\bar{Z}-\bar{X})'+O(N_1^{-2})\Big)-pN_1\log(1+N_3/N_1),$$

where $\bar{Z}=\frac{1}{N_3}\sum_{i=1}^{N_3}Z_i$. Here, $S_1^{(1)}/(N_1+N_3)\to\Sigma_1$ as $N_1\to\infty$, not only under H_1 but also under H_2, because

$$\begin{aligned}(6.50)\qquad & S_1^{(1)}/(N_1+N_3)\\ &=\{S_1^{(2)}/N_1\}\cdot N_1(N_1+N_3)^{-1}+(N_1+N_3)^{-1}\textstyle\sum_{i=1}^{N_3}(Z_i-\bar{Z})(Z_i-\bar{Z})'\\ &\quad+N_1N_3(N_1+N_3)^{-2}(\bar{Z}-\bar{X})(\bar{Z}-\bar{X})'.\end{aligned}$$

So, it holds that

$$\begin{aligned}(6.51)\qquad dz(1)\longrightarrow\ & N_3\log|\Sigma_1|+\textstyle\sum_{i=1}^{N_3}\operatorname{tr}\{\Sigma_1^{-1}(Z_i-\bar{Z})(Z_i-\bar{Z})'\}\\ &+N_3\operatorname{tr}\Sigma_1^{-1}(\bar{Z}-\mu_1)(\bar{Z}-\mu_1)'-pN_3\\ &=N_3\log|\Sigma_1|-pN_3+D(Z,1)\end{aligned}$$

in probability as $N_1\to\infty$. Consequently, as $N_1\to\infty$

$$(6.52)\qquad dz(1)\longrightarrow N_3\log|\Sigma_1|-pN_3+\chi_f^2$$

under H_1. Obviously, similar arguments hold for $dz(2)$ by exchanging the suffixes.

By the Stiring's formula,

$$(6.53)\qquad \Gamma(t+\alpha)/\Gamma(t)\sim t^\alpha$$

for large t, we have

$$(6.54)\qquad \log c_1\sim-(pN_3/2)\log(N_1/2)$$

for large N_1. Hence

$$\begin{aligned}(6.55)\qquad db(1)&=2\log c_1+N_3\log|S_1^{(2)}|+(N_1+N_3-1)\log|\{S_1^{(2)}\}^{-1}S_1^{(1)}|\\ &=2\log c_1+pN_3\log N_1+N_3\log|S_1^{(2)}/N_1|\\ &\quad+(N_1+N_3-1)\log|\{S_1^{(2)}\}^{-1}S_1^{(1)}|\\ &\quad\longrightarrow pN_3\log 2+N_3\log|\Sigma_1|+D(Z,1)\end{aligned}$$

in probability as $N_1\to\infty$, by a slight modification of the calculation for $dz(1)$. Of course, the corresponding result holds for $db(2)$.

Summarizing the results, the following theorem holds.

THEOREM 6.2. *When H_j is true,*

$$
(6.56)\qquad
\begin{aligned}
dw(j) &\longrightarrow N_3 \log|\Sigma_j| + \chi_f^2\\
dz(j) &\longrightarrow -pN_3 + N_3 \log|\Sigma_j| + \chi_f^2\\
db(j) &\longrightarrow pN_3 \log 2 + N_3 \log|\Sigma_j| + \chi_f^2
\end{aligned}
$$

in probability as $N_j \to \infty$.

From this theorem, the limits of the expectations of the statistics under H_j are given as

$$
(6.57)\qquad
\begin{aligned}
\mathrm{E}[dw(j)] &\longrightarrow pN_3 + N_3 \log|\Sigma_j|\\
\mathrm{E}[dz(j)] &\longrightarrow N_3 \log|\Sigma_j|\\
\mathrm{E}[db(j)] &\longrightarrow pN_3 \log 2 + pN_3 + N_3 \log|\Sigma_j|,
\end{aligned}
$$

and the variances of them have the same limit $2pN_3$.

By the above results, it is clear that the limit of the statistics $dw(j)$, $dz(j)$ and $db(j)$ are described as a sum of N_3 independent variables which are identically distributed. Namely, for example,

$$
(6.58)\qquad dw(j) \longrightarrow \textstyle\sum_{i=1}^{N_3} \{\log|\Sigma_j| + (Z_i - \mu_j)'\Sigma_j^{-1}(Z_i - \mu_j)\}.
$$

Kanazawa [15] studied the distribution of a variable of the form which appears in the brace of (6.58). Therefore, from [15] we have the following theorem:

THEOREM 6.3. *When* H_{3-j} *is true, it holds that*

(6.59)

$$
\begin{aligned}
\mathrm{E}[dw(j)] &\longrightarrow N_3\{\log|\Sigma_j| + \mathrm{tr}\,\Sigma_j^{-1}\Sigma_{3-j} + (\mu_2 - \mu_1)'\Sigma_j^{-1}(\mu_2 - \mu_1)\},\\
\mathrm{E}[dz(j)] &\longrightarrow N_3\{\log|\Sigma_j| - p + \mathrm{tr}\,\Sigma_j^{-1}\Sigma_{3-j} + (\mu_2 - \mu_1)'\Sigma_j^{-1}(\mu_2 - \mu_1)\},\\
\mathrm{E}[db(j)] &\longrightarrow N_3\{\log|\Sigma_j| + p\log 2 + \mathrm{tr}\,\Sigma_j^{-1}\Sigma_{3-j} + (\mu_2 - \mu_1)'\Sigma_j^{-1}(\mu_2 - \mu_1)\}
\end{aligned}
$$

as $N_j \to \infty$. *Further, the variances of them have the same limit*

$$
(6.60)\qquad V(j) = N_3\{2\,\mathrm{tr}(\Sigma_j^{-1}\Sigma_{3-j})^2 + 4(\mu_2 - \mu_1)'\Sigma_j^{-1}\Sigma_{3-j}\Sigma_j^{-1}(\mu_2 - \mu_1)\}.
$$

Let $I(j, 3-j)$ be the Kullback-Leibler information (see, e.g., Zacks [34]) for classification in favour H_j against H_{3-j} $(j = 1, 2)$. Then

$$
(6.61)\qquad
\begin{aligned}
I(j, 3-j) &= \mathrm{E}_j(\log\{f(Z;\mu_j,\Sigma_j)/f(Z;\mu_{3-j},\Sigma_{3-j})\})\\
&= \int f(Z;\mu_j,\Sigma_j)\log\{f(Z;\mu_j,\Sigma_j)/f(Z;\mu_{3-j},\Sigma_{3-j})\}\,dZ\\
&= \frac{N_3}{2}(\log\{|\Sigma_{3-j}|/|\Sigma_j|\} + \mathrm{tr}\{\Sigma_j(\Sigma_{3-j}^{-1} - \Sigma_j^{-1})\}\\
&\quad + \mathrm{tr}\{\Sigma_{3-j}(\mu_2 - \mu_1)(\mu_2 - \mu_1)'\}).
\end{aligned}
$$

The limiting distributions of DW, DZ and DB can be expressd in the term of

(6.62) $$D(1, 2) = D(Z, 2) - D(Z, 1) + N_3 \log|\Sigma_2 \Sigma_1^{-1}|.$$

THEOREM 6.4. *When $N_1, N_2 \to \infty$, DW, DZ and DB have the same limiting distribution, which is given as the distribution of $D(1, 2)$. If H_1 is true, its mean and variance are given by*

$$2I(1, 2) \text{ and } V(1, 2),$$

respectively. If H_2 is true, its mean and variance are given by

$$-2I(2, 1) \text{ and } V(2, 1),$$

respectively. Here

(6.63) $$V(j, 3-j) = N_3(2 \operatorname{tr}\{(\Sigma_{3-j}^{-1} - \Sigma_j^{-1})\Sigma_j(\Sigma_{3-j}^{-1} - \Sigma_j^{-1})\Sigma_j\} + 4(\mu_2 - \mu_1)' \Sigma_{3-j}^{-1}\Sigma_j\Sigma_{3-j}^{-1}(\mu_2 - \mu_1)).$$

Now, it becomes clear that rule-Z and -B are admissible and that DW, DZ and DB have the same limiting distribution. These results can be stated in a combined form as follows.

Consider the statistic defined by

(6.64) $$\begin{aligned} D_\delta = {} & (N_1 + \delta)\log|S_1^{(2)}| + (N_2 + N_3 + \delta)\log|S_2^{(2)}| \\ & - (N_1 + N_3 + \delta)\log|S_1^{(1)}| - (N_2 + \delta)\log|S_2^{(1)}| \\ & + N_3 p(\log N_1 - \log N_2) + O\{(N_1^2 + N_2^2)^{-1/2}\}, \end{aligned}$$

where $\delta > -\min(N_1, N_2)$. Here, the part $O(\cdot)$ does not contain any sample variables. Such a statistic is obtained from (6.36) by putting

(6.65) $$\begin{aligned} & r_1 = d(N_1 + N_3 + \delta),\ r_2 = d(N_2 + \delta),\ r_3 = d(N_1 + \delta), \\ & r_4 = d(N_2 + N_3 + \delta). \end{aligned}$$

THEOREM 6.5. *If $\min(N_1, N_2) > 2(p-1)$, then the procedure;*

select H_1 or H_2 according as $D_\delta >$ or < 0

is admissible Bayes for $\delta > p - 1 - \min(N_1, N_2)$. The limiting distribution of D_δ as $N_1, N_2 \to \infty$ is given as the distribution of $D(1, 2)$ for any δ.

PROOF. Choose d as slightly larger than $(p-1)/\{\min(N_1, N_2) + \delta\}$, and consider r_j's in (6.65). If $d < 1$, these r_j's satisfy the conditions of Theorem 6.1. If $\delta > p - 1 - \min(N_1, N_2)$, then there exists such d. The limiting distribution of D_δ can be obtained by similar calculation as for DZ (e.g.) and coincides with that of the previous three statistics.

DZ and DB are obtained as (6.64) with $\delta = 0$ and -1, respectively. For DW, it is clear that DW has a form of D_δ if the part $O(\cdot)$ is permitted to contain the sample variables. That is, using

$$d_\delta(j) = (N_j + N_3 + \delta)\log|S_j^{(j)}| + (N_j + \delta)\log|S_j^{(3-j)}|, \tag{6.66}$$

it holds that (after calculations like ones for $dz(1)$)

$$dw(j) = d_\delta(j) + O(N_j^{-1}). \tag{6.67}$$

and hence

$$DW = D_\delta + O\{(N_1^2 + N_2^2)^{-1/2}\} \tag{6.68}$$

for any fixed δ. In this case, however, the part $O(\cdot)$ contains the sample variables.

6.4. Numerical comparison of rule-W, -Z and -B

In this subsection we compare the three rules -W, -Z and -B by using simulation. We consider 12 cases of populations π_1, π_2 and π_3. For each case we set N_i $(i = 1, 2)$ equal to 6, 10, 30. The 12 cases are defined as follows. These cases are chosen to examine how the change of experimental conditions influence the characteristics of the three rules. It is assumed that π_1 is always $N_p(\mathbf{0}, I_p)$.

CASE 1.

$$p = 3,\ N_3 = 1,\ \mu_2 = \mu_2^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},\ \Sigma_2 = \Sigma_2^{(1)} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}.$$

CASE 2.

$$p = 3,\ N_3 = 3,\ \mu_2 = \mu_2^{(1)},\ \Sigma_2 = \Sigma_2^{(1)}.$$

CASE 3.

$$p = 5,\ N_3 = 1,\ \mu_2 = \mu_2^{(2)} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},\ \Sigma_2 = \Sigma_2^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

CASE 4.

$$p = 5,\ N_3 = 3,\ \mu_2 = \mu_2^{(2)},\ \Sigma_2 = \Sigma_2^{(2)}.$$

CASE 5.

$$p = 3,\ N_3 = 1,\ \mu_2 = \mu_2^{(1)},\ \Sigma_2 = \Sigma_2^{(3)} = \begin{bmatrix} 1 & .5 & .5 \\ .5 & 1 & .5 \\ .5 & .5 & 1 \end{bmatrix}.$$

CASE 6.

$$p = 3,\ N_3 = 3,\ \mu_2 = \mu_2^{(1)},\ \Sigma_2 = \Sigma_2^{(3)}.$$

CASE 7.

$$p = 5,\ N_3 = 1,\ \mu_2 = \mu_2^{(2)},\ \Sigma_2 = \Sigma_2^{(4)} = \begin{bmatrix} 1 & .5 & .5 & .5 & .5 \\ .5 & 1 & .5 & .5 & .5 \\ .5 & .5 & 1 & .5 & .5 \\ .5 & .5 & .5 & 1 & .5 \\ .5 & .5 & .5 & .5 & 1 \end{bmatrix}.$$

CASE 8.

$$p = 5,\ N_3 = 3,\ \mu_2 = \mu_2^{(2)},\ \Sigma_2 = \Sigma_2^{(4)}.$$

CASE 9.

$$p = 3,\ N_3 = 1,\ \mu_2 = \mu_2^{(3)} = \begin{bmatrix} .5 \\ 1 \\ 1.5 \end{bmatrix},\ \Sigma_2 = \Sigma_2^{(5)} = \begin{bmatrix} 1 & .2 & .8 \\ .2 & 2 & 1.8 \\ .8 & 1.8 & 3 \end{bmatrix}.$$

CASE 10.

$$p = 5,\ N_3 = 1,\ \mu_2 = \mu_2^{(4)} = \begin{bmatrix} .5 \\ 1 \\ 1.5 \\ 1 \\ 1 \end{bmatrix},\ \Sigma_2 = \Sigma_2^{(6)} = \begin{bmatrix} 1 & .2 & .8 & .5 & .4 \\ .2 & 2 & 1.8 & .8 & .5 \\ .8 & 1.8 & 3 & .9 & .5 \\ .5 & .8 & .9 & 4 & .4 \\ .4 & .5 & .5 & .4 & 5 \end{bmatrix}.$$

CASE 11.

$$p = 5,\ N_3 = 1,\ \mu_2 = \mu_2^{(5)} = [1, 1, 2, 3, 1]',\ \Sigma_2 = \Sigma_2^{(2)}.$$

CASE 12.

$$p = 5,\ N_3 = 1,\ \mu_2 = \mu_2^{(6)} = [1, 1, 1, 3, 3]',\ \Sigma_2 = \Sigma_2^{(4)}.$$

The normal pseudorandom numbers were generated by the Box-Müller method based uniform pseudorandom numbers (generated by personal computors). For each case, 1000 observations are carried out under H_1 and H_2, respectively. For Cases $1 \sim 8$, three tables are given. The first tables record the rates of correct classifications of three rules among 1000 observations which were carried out under H_1. The second tables record the corresponding ones under H_2. The third tables record the rates among 2000 observations which are obtained by averaging the first ones and the second ones. We call the third type of tables the averaged tables. The rules -*W*, -*Z* and -*B* are denoted by *W*, *Z* and *B* in tables. The values in the last row and the last three columns of the tables present the averages of the rows and columns. For each of Cases $9 \sim 12$, only the averaged tables are given. In paticular, the averaged tables which are restricted to $N_1, N_2 = 6, 10$ are given for Cases 11 and 12. Each table number or its first number correspond to the case number. For example, Table 1.2 is the second one for Case 1 and Table 9 is the one for Case 9.

Table 1.1.

N_1 \ N_2	6			10			30			TOTAL		
	W	*Z*	*B*	*W*	*Z*	*B*	*W*	*Z*	*B*	*W*	*Z*	*B*
6	.67	.68	.68	.58	.67	.61	.55	.71	.58	.600	.683	.623
10	.76	.70	.77	.72	.72	.72	.67	.75	.67	.716	.722	.719
30	.83	.72	.84	.79	.75	.81	.81	.81	.81	.809	.758	.818
TOTAL	.753	.699	.763	.696	.711	.710	.675	.753	.686	.708	.721	.719

Table 1.2.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.69	.70	.71	.76	.72	.76	.82	.72	.85	.756	.712	.770
10	.60	.68	.63	.71	.72	.71	.79	.74	.80	.702	.712	.716
30	.53	.68	.56	.70	.77	.63	.76	.76	.75	.661	.734	.647
TOTAL	.606	.686	.631	.723	.733	.702	.790	.739	.801	.706	.719	.711

Table 1.3.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.68	.69	.69	.67	.69	.69	.69	.71	.71	.678	.698	.697
10	.68	.69	.70	.72	.72	.72	.73	.75	.74	.709	.717	.717
30	.68	.70	.70	.74	.76	.72	.78	.78	.78	.735	.746	.732
TOTAL	.679	.693	.697	.709	.722	.706	.733	.746	.743	.707	.720	.715

Table 2.1.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.71	.76	.77	.60	.76	.68	.49	.81	.67	.598	.780	.707
10	.87	.82	.88	.81	.83	.83	.74	.86	.80	.806	.837	.836
30	.96	.85	.95	.91	.86	.90	.89	.89	.89	.922	.871	.916
TOTAL	.847	.815	.866	.771	.818	.805	.707	.856	.789	.775	.829	.820

Table 2.2.

N_2 / N_1	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.74	.77	.77	.89	.84	.89	.95	.87	.95	.859	.826	.870
10	.63	.78	.72	.82	.83	.83	.93	.86	.92	.792	.826	.822
30	.49	.80	.68	.75	.87	.82	.92	.92	.92	.722	.861	.806
TOTAL	.620	.781	.723	.820	.847	.847	.933	.885	.928	.791	.838	.833

Table 2.3.

N_2 / N_1	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.72	.77	.77	.74	.80	.79	.72	.84	.81	.729	.803	.789
10	.75	.80	.80	.81	.83	.83	.83	.86	.86	.799	.832	.829
30	.73	.83	.81	.83	.87	.86	.91	.91	.91	.822	.866	.861
TOTAL	.734	.798	.794	.796	.833	.826	.820	.870	.858	.783	.834	.826

Table 3.1.

N_2 / N_1	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.61	.61	.61	.30	.48	.38	.24	.47	.31	.382	.519	.431
10	.91	.83	.90	.72	.72	.72	.59	.75	.66	.740	.766	.760
30	.95	.84	.95	.87	.78	.86	.83	.83	.83	.883	.817	.881
TOTAL	.822	.761	.819	.632	.658	.654	.551	.683	.600	.668	.701	.691

Table 3.2.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.60	.63	.63	.90	.81	.88	.96	.85	.96	.819	.764	.824
10	.34	.51	.42	.72	.72	.72	.87	.78	.87	.646	.669	.668
30	.24	.48	.34	.63	.78	.69	.82	.82	.82	.561	.694	.615
TOTAL	.394	.539	.462	.749	.772	.764	.883	.816	.881	.675	.709	.702

Table 3.3.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.61	.62	.62	.60	.64	.63	.60	.66	.64	.601	.641	.628
10	.63	.67	.66	.72	.72	.72	.73	.76	.76	.693	.718	.714
30	.59	.66	.65	.75	.78	.78	.82	.82	.82	.722	.756	.748
TOTAL	.608	.650	.641	.690	.715	.709	.717	.749	.740	.672	.705	.697

Table 4.1.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.57	.65	.66	.21	.50	.36	.08	.52	.27	.286	.556	.430
10	.95	.90	.97	.80	.83	.83	.55	.85	.75	.764	.858	.847
30	1.0	.93	.99	.98	.91	.97	.96	.95	.95	.976	.932	.969
TOTAL	.839	.829	.873	.660	.745	.719	.527	.772	.654	.675	.782	.749

Table 4.2.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.62	.69	.70	.96	.91	.96	1.0	.95	1.0	.858	.848	.884
10	.23	.52	.37	.79	.83	.84	.98	.91	.96	.668	.753	.724
30	.13	.57	.34	.61	.89	.78	.95	.96	.96	.564	.807	.692
TOTAL	.327	.595	.471	.787	.875	.857	.976	.937	.973	.697	.802	.767

Table 4.3.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.59	.67	.68	.59	.70	.66	.54	.73	.63	.572	.702	.657
10	.59	.71	.67	.79	.83	.83	.76	.88	.86	.716	.805	.786
30	.56	.75	.67	.79	.90	.87	.95	.95	.95	.770	.869	.831
TOTAL	.583	.712	.672	.723	.810	.788	.751	.854	.814	.686	.792	.758

Table 5.1.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.65	.64	.63	.57	.63	.55	.51	.67	.51	.576	.647	.563
10	.78	.70	.77	.70	.70	.69	.66	.73	.65	.713	.709	.704
30	.84	.71	.85	.79	.71	.78	.78	.78	.78	.803	.731	.803
TOTAL	.756	.681	.746	.687	.678	.676	.649	.727	.647	.697	.695	.690

Table 5.2.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.64	.67	.67	.74	.70	.76	.82	.74	.85	.735	.703	.761
10	.56	.65	.60	.67	.68	.69	.76	.73	.78	.667	.685	.692
30	.50	.65	.55	.65	.70	.66	.74	.74	.74	.629	.697	.648
TOTAL	.569	.654	.607	.688	.696	.703	.774	.735	.792	.677	.695	.701

Table 5.3.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.65	.65	.65	.66	.67	.66	.66	.71	.68	.656	.675	.662
10	.68	.67	.69	.69	.69	.69	.71	.73	.72	.690	.697	.698
30	.67	.68	.70	.72	.70	.72	.76	.76	.76	.716	.714	.725
TOTAL	.663	.668	.677	.688	.687	.689	.711	.731	.720	.687	.695	.695

Table 6.1.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.70	.74	.74	.55	.73	.63	.40	.73	.55	.552	.731	.638
10	.85	.78	.83	.76	.77	.76	.72	.86	.77	.777	.802	.788
30	.95	.84	.94	.91	.83	.91	.89	.89	.89	.916	.852	.911
TOTAL	.833	.786	.836	.743	.775	.765	.669	.824	.736	.748	.795	.779

Table 6.2.

N_1 \ N_2	6			10			30			TOTAL		
	W	*Z*	*B*	*W*	*Z*	*B*	*W*	*Z*	*B*	*W*	*Z*	*B*
6	.69	.73	.75	.88	.83	.89	.95	.85	.96	.840	.802	.865
10	.59	.75	.69	.80	.83	.83	.91	.87	.92	.767	.813	.814
30	.51	.79	.66	.75	.86	.81	.89	.89	.89	.716	.848	.788
TOTAL	.594	.756	.701	.811	.840	.844	.917	.867	.923	.774	.821	.823

Table 6.3.

N_1 \ N_2	6			10			30			TOTAL		
	W	*Z*	*B*	*W*	*Z*	*B*	*W*	*Z*	*B*	*W*	*Z*	*B*
6	.70	.74	.74	.72	.78	.76	.68	.79	.75	.696	.766	.752
10	.72	.77	.76	.78	.80	.80	.82	.86	.85	.772	.807	.801
30	.73	.81	.80	.83	.85	.86	.89	.89	.89	.816	.850	.850
TOTAL	.714	.771	.768	.777	.807	.805	.793	.846	.829	.761	.808	.801

Table 7.1.

N_1 \ N_2	6			10			30			TOTAL		
	W	*Z*	*B*	*W*	*Z*	*B*	*W*	*Z*	*B*	*W*	*Z*	*B*
6	.57	.56	.55	.27	.39	.27	.19	.40	.21	.343	.449	.344
10	.89	.77	.86	.66	.63	.63	.50	.65	.52	.685	.682	.666
30	.94	.80	.94	.85	.71	.85	.82	.81	.80	.869	.771	.862
TOTAL	.800	.707	.784	.592	.576	.581	.505	.619	.508	.632	.634	.624

Table 7.2.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.60	.63	.64	.94	.89	.95	1.0	.95	1.0	.846	.823	.862
10	.21	.52	.39	.74	.81	.82	.96	.89	.96	.638	.741	.725
30	.10	.51	.29	.58	.84	.74	.90	.91	.92	.527	.757	.650
TOTAL	.303	.555	.441	.755	.847	.840	.953	.918	.957	.670	.774	.746

Table 7.3.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.58	.60	.60	.61	.64	.61	.60	.68	.60	.594	.636	.603
10	.55	.64	.63	.70	.72	.72	.73	.77	.74	.662	.712	.696
30	.52	.66	.62	.71	.78	.80	.86	.86	.86	.698	.764	.756
TOTAL	.551	.631	.613	.674	.712	.711	.729	.769	.732	.651	.704	.685

Table 8.1.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.61	.63	.63	.16	.39	.24	.07	.42	.16	.279	.478	.339
10	.96	.85	.94	.74	.74	.73	.48	.78	.59	.724	.790	.755
30	1.0	.91	.99	.97	.85	.95	.89	.88	.88	.952	.880	.938
TOTAL	.854	.797	.852	.622	.658	.639	.478	.694	.542	.652	.716	.677

Table 8.2.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.58	.66	.68	.95	.88	.95	.99	.93	1.0	.846	.823	.862
10	.23	.52	.38	.76	.82	.83	.95	.88	.95	.644	.741	.721
30	.11	.55	.33	.58	.86	.76	.91	.92	.92	.536	.775	.670
TOTAL	.307	.575	.441	.762	.854	.846	.951	.909	.957	.673	.779	.755

Table 8.3.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.60	.64	.65	.56	.63	.60	.53	.68	.58	.560	.650	.607
10	.59	.69	.66	.75	.78	.78	.72	.83	.77	.684	.766	.738
30	.56	.73	.66	.78	.86	.85	.90	.90	.90	.744	.828	.804
TOTAL	.581	.686	.658	.692	.756	.742	.715	.801	.749	.663	.748	.716

Table 9.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.63	.65	.65	.65	.67	.66	.63	.68	.64	.637	.665	.650
10	.65	.66	.68	.69	.70	.70	.72	.71	.72	.688	.692	.699
30	.67	.68	.71	.70	.70	.71	.74	.74	.74	.701	.707	.720
TOTAL	.649	.664	.679	.682	.690	.690	.695	.710	.700	.675	.688	.690

Table 10.

N_1 \ N_2	6			10			30			TOTAL		
	W	Z	B	W	Z	B	W	Z	B	W	Z	B
6	.58	.59	.60	.60	.62	.62	.57	.63	.59	.586	.613	.600
10	.62	.65	.66	.70	.71	.71	.70	.74	.73	.675	.701	.700
30	.60	.68	.69	.78	.80	.81	.81	.82	.82	.729	.765	.773
TOTAL	.601	.640	.646	.693	.707	.714	.696	.732	.713	.663	.693	.691

Table 11.

N_1 \ N_2	6			10		
	W	Z	B	W	Z	B
6	.72	.74	.74	.71	.79	.77
10	.73	.80	.78	.89	.90	.90

Table 12.

N_1 \ N_2	6			10		
	W	Z	B	W	Z	B
6	.72	.73	.73	.70	.78	.75
10	.73	.80	.80	.89	.89	.89

Discussion At first, we investigate Case 1. When H_1 is true (the data to be classified are taken from π_1), if $N_2 = 30$ and $N_1 = 6, 10$, then the rule-Z is better than -W and -B. When $N_1 = 30$ and $N_2 = 6, 10$ in table 1.2 (H_2 is true), rule-Z is also better than -W and -B. These facts seem to be caused by the estimating method in rule-Z. Since DZ is obtained by using $N_i + N_3$ observations for estimatimation under H_i, if N_i is small, N_3 observations are effective for estimation. If so, the tendancy should appear more notably for $N_3 = 3$ than for $N_3 = 1$. Hence, let us examine Case 2. The tendancy stated above also appears in Tables 2.1 and 2.2, more clearly than Tables 1.1 and 1.2. The other hand, if $N_1 = 30$ and $N_2 = 6, 10$ under H_1 or $N_2 = 30$ and $N_1 = 6, 10$ under H_2, the rule-Z is worse than -W and -B. However, the inferiority of -Z from -W and -B is almost same for $N_3 = 1$ and $N_3 = 3$. These tendancy is seen not only for Cases 1 and 2 but also for Cases 3 and 4, etc.

As a criterion for the goodness of the three rules, it is reasonable to use the averages of the values (last three values of the last rows) in the averaged tables. Examining these values, rule-Z and -B are rather better than

-*W*. Regarding to -*Z* and -*B*, -*Z* is slightly better than -*B*. This tendancy hold for almost all of Cases $1 \sim 10$ and other cases which are not written in this paper. Kanazawa [15] studied the property of the three rules. She carried out numerical simulations in the case $p = 2$ and $N_3 = 1$. From the results there it is known that rule-*Z* and -*B* have nearly the same goodness and both are better than rule-*W*. Our conclusion coincides with her one. However, we give a further comparison between rule-*Z* and -*B* as well as the case $p = 3$ and $N_3 > 1$.

We note that Case 3 is obtained from Case 1 by adding two variables. From Tables 1.3 and 3.3, it becomes clear that the rates in Table 1.3 are better than the corresponding ones in Table 3.3 for the case N_1 and/or N_2 equal to 6 (that is, small samples cases). Consequently, it seems useless to add variables in this case. This property also find in Tables 2.3 and 4.3, or in other cases. On the other hand, it is possible to give examples which improve the rates of correct classification by adding some variables in small samples cases. Those are Cases 11 and 12, in which the rates are improved than the corresponding ones for $p = 3$ (Tables 1.3 and 5.3). Thus, it is important to examine the deviation of the two populations when we attempt to add or delete variables.

Acknowledgment

I wish to thank Prof. Yasunori Fujikoshi, Hiroshima University for his encouragement and for his careful reading which led to the improved form of the manuscript.

References

[1] T. W. Anderson, Statistical inference for covariance matrices with linear structure. Multivariate Analysis II (P. R. Krishnaiah, ed.), Academic Press, New York (1969), 55–66.

[2] T. W. Anderson, Estimation of covariance matrices which are linear combinations or whose inverse are linear combinations of given matrices, Essays Prob. Statist. (R. C. Bose and Others, eds), Univ. North Carolina Press, Chapel Hill (1970), 1–24.

[3] T. W. Anderson, An Introduction to Multivariate Statistical Amalysis (2nd Ed.), John Wiley & Sons, New York, 1984.

[4] T. W. Anderson and S. Das Gupta, A monotonicity property of the power functions of some tests of the equality of two covariance matrices, Ann. Math. Statist. **35** (1964), 1059–1063.

[5] T. W. Anderson and A. Takemura, A new proof of admissibility of tests in multivariate analysis, J. Multivariate Anal. **12** (1982), 457–468.

[6] R. E. Bechhofer, A single sample multiple decision procedure for ranking means of normal populations with known variances, Ann. Math. Statist. **25** (1954), 16–39.

[7] R. E. Bechhofer, C. W. Dunnett and M. Sobel, A two-sample multiple decision procedure

for ranking means of noral populations with a common unknown variance, Biometrika **41** (1954), 170–176.

[8] R. E. Bechhofer and M. Sobel, A single sample multiple decision procedure for ranking variances of normal populations, Ann. Math. Statist. **25** (1954), 273–289.

[9] A. Birnbaum, Charactrization of complete classes of tests of some multiparametric hypotheses, with applications to likelihood ratio tests, Ann. Math. Statist. **26** (1955), 21–36.

[10] R. D. Bock and R. E. Bargmann, Analysis of covariance structures, Psychometrika, **31** (1966), 507–534.

[11] C. W. Dunnett, On selecting the largest of k normal populations means, J. Roy. Statist. Soc. B **22** (1960), 1–40.

[12] M. L. Eaton, Some optimum properties of ranking procedures, Ann. Math. Statist. **38** (1967), 124–137.

[13] M. N. GHOSH, On the admissibility of some tests of MANOVA, Ann. Math. Statist. **35** (1964), 789–794.

[14] W. J. Hall, The most economical character of some Bechhofer and Sobel decision rules, Ann. Math. Statist. **30** (1959), 964–969.

[15] M. Kanazawa, ML classfication distances and rules in the normal populations when covariance matrices are unequal, J. Japan Statist. Soc. **16** (1986), 25–36.

[16] J. Kiefer and R. Schwartz, Admissible Bayes character of T^2-, R^2-, and other fully invariant tests for classical multivariate normal problems, Ann. Math. Statist. **36** (1965), 747–770.

[17] P. R. Krishnaiah and J. C. Lee, On covariance structures, Sankhyā A, **38** (1974), 357–371.

[18] K. V. Mardia, J. T. Kent and J. M. Bibby, Multivariate Analysis, Academic Press, 1979.

[19] N. Nishida, A note on the admissible tests and classifications in multivariate analysis, Hiroshima Math. J. **1** (1971), 427–434.

[20] N. Nishida, The admissibility of tests for the equality of mean vectors and covariance matrices, Hiroshima Math. J. **2** (1972), 215–220.

[21] N. Nishida, Admissibility of classification procedures in multivariate analysis I, Bull. Hiroshima Women's Univ. **12** (1976), 5–9.

[22] N. Nishida, Admissibility of classification procedures in multivariate analysis II, Bull. Hiroshima Women's Univ. **13** (1977), 33–37.

[23] N. Nishida, Admissible selection procedures in multivariate analysis, Bull. Hiroshima Women's Univ. **24** (1988), 69–74.

[24] N. Nishida, Admissibility of the likelihood ratio test for linear constraints, J. Japan Statist. Soc. **20** (1990), 43–49.

[25] I. Olkin and S. J. Press, Testing and estimation for a circular stationary model, Ann. Math. Statist. **40** (1969), 1358–1373.

[26] E. Paulson, A multiple decision procedure for certain problems in analysis of variance, Ann. Math. Statist. **20** (1949), 95–98.

[27] C. R. Rao, Some problems involving linear hypotheses in multivariate analysis, Biometrika, **46** (1959), 49–58.

[28] S. N. Roy and R. Gnanadeskian, Two-sample comparisons of dispersion matrices for alternatives of intermediate specificity, Ann. Math. Statist. **33** (1962), 432–438.

[29] R. Schwartz, Admissible tests in multivariate analysis of variance, Ann. Math. Statist. **38** (1967), 698–710.

[30] K. C. Seal, On a class of decision procedures for ranking means of normal populations, Ann. Math. Statist. **26** (1955), 387–398.

[31] M. Siotani, T. Hayakawa and Y. Fujikoshi, Modern Multivariate Statistical Analysis, American Sciences Press, 1985.

[32] J. N. Srivastava, On testing hypotheses regarding a class of covariance structure, Psychometrika, **31** (1966), 147–164.

[33] C. Stein, The admissibility of Hotelling's T^2-test, Ann. Math. Statist. **27** (1956), 616–623.

[34] S. Zacks, The Theory of Statistical Inference, John Wiley & Sons, New York, 1971.

Department of Life Science
Facculty of Home Economics
Hiroshima Women's University
Hiroshima 734, *Japan*