The Korean Communications
in Statistics Vol. 8, No. 1, 2001
pp. 97-107

A Bayesian Criterion for a Multiple test of Two Multivariate Normal Populations

Hea Jung Kim[1] and Young Sook Son[2]

Abstract

A simultaneous test criterion for multiple hypotheses concerning comparison of two multivariate normal populations is considered by using the so called Bayes factor method. Fully parametric frequentist approach for the test is not available and thus Bayesian criterion is pursued using a Bayes factor that eliminates its arbitrariness problem induced by improper priors. Specifically, the fractional Bayes factor (FBF) by O'Hagan (1995) is used to derive the criterion. Necessary theories involved in the derivation and computation of the criterion are provided. Finally, an illustrative simulation study is given to show the properties of the criterion.

Keywords : multivariate normal populations, Jeffreys prior, noninformative improper prior, fractional Bayes factor(FBF), intrinsic Bayes factor(IBF), posterior probability.

1. Introduction

Let Π_1 and Π_2 be two independent p-variate normal populations, where for $k=1,2$, $\Pi_k \sim N_p(\mu_k, \Sigma_k)$ with a $p \times 1$ unknown mean vector μ_k and a $p \times p$ unknown covariance matrix Σ_k. Suppose that we wish to do a multiple test composed of four models,

$$\begin{cases} M_0 : \mu_1 \neq \mu_2 & \text{and} \quad \Sigma_1 \neq \Sigma_2, \\ M_1 : \mu_1 = \mu_2 & \text{and} \quad \Sigma_1 = \Sigma_2, \\ M_2 : \mu_1 \neq \mu_2 & \text{and} \quad \Sigma_1 = \Sigma_2, \\ M_3 : \mu_1 = \mu_2 & \text{and} \quad \Sigma_1 \neq \Sigma_2. \end{cases} \tag{1.1}$$

In the frequentist approach, test problem of two multivariate normal means and/or

1) Professor, Department of Statistics, Dongguk University, Seoul, 100-715, Korea.
E-mail : kim3hj@dgu.ac.kr.
2) Professor, Department of Statistics, Chonnam National University, Kwangju, 500-757, Korea
E-mail : ysson@chonnam.chonnam.ac.kr.

covariance matrices has been studied by many authors and various approximate procedures have been proposed, including Siotani (1987) and Yao (1965) for the multivariate Behrens-Fisher problem, Nagao (1973) and Perlman (1980) for the test of equality of covariance matrices, and Anderson (1984) for the equality of two multivariate populations. We refer Krzanowski and Marriott (1994) for the other frequentist methods. But, a test (say, numltiple test) that simultaneously tests all the models, M_0 through M_3, has not been seen yet. The Bayesian approach, however, to the multiple test can be seen in Kim and Kim (2001). They used the arithmetic intrinsic Bayes factor (AIBF) by Berger and Pericchi (1996) for deriving the multiple test criterion. Even though the criterion performs well in the multiple test, it is impractical when the dimension of multivariate and the number of observations get larger, i.e. the number of minimal training samples required for calculating the AIBF increases.

To circumvent the problem, in this paper, we propose an alternative Bayesian criterion for the multiple test. It is obtained by use of the fractional Bayes factor approach introduced by O'Hagan (1995). Then we develop a numerical technique to calculate the criterion.

In the next section, the FBF, the IBF, and the posterior probability of hypothesis are introduced. In section 3, we compute the FBF, and in section 4, a Bayesian criterion proposed in this paper is applied to some simulated data.

2. The Fractional Bayes Factor

Suppose that we wish to test q models,

$$M_i \,:\, \boldsymbol{X} \sim f_i(X|\theta_i), \quad \theta_i \in \Theta_i,$$

for $i = 1, 2, \ldots, q$, with a random sample $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ of size n, where $f_i(X|\theta_i)$ is a probability density function, and θ_i and Θ_i are a parameter vector and a parameter space under the model M_i, respectively. We define the function of a random sample $\boldsymbol{X}$ and a constant b as follows, for $i = 1, 2, \ldots, q$, $j \neq i$,

$$B_{ji}(\boldsymbol{X}|b) = \frac{m_j(\boldsymbol{X}|b)}{m_i(\boldsymbol{X}|b)},$$

where

$$m_i(\boldsymbol{X}|b) = \int_{\Theta_i} \pi_i(\theta_i) L_i^b(\theta_i|\boldsymbol{X}) d\theta_i, \tag{2.1}$$

$\pi_i(\theta_i)$ is a prior distribution of θ_i, $L_i(\theta_i|\boldsymbol{X}) = \Pi_{k=1}^n f_i(\boldsymbol{X}|\theta_i)$ is a likelihood function, and b is a constant such that $0 < b \leq 1$.

The usual factor B_{ji} as a Bayesian tool to test the model M_j to the model M_i is defined by

$$B_{ji}(X|b=1) = \frac{m_j(X|b=1)}{m_i(X|b=1)},$$

where $m_i(X|b=1)$ is usually called a marginal or a predictive density of the model M_i.

The first step in a Bayesian inference is to choose the prior distributions of all the parameters in hypotheses or models. Default priors, most of which are typically noninformative improper, are objective priors that need not any subjective consideration. But the Bayes factor $B_{ji}(X|b=1)$ cannot be used because of arbitrary constants incorporated into the Bayes factor if priors are noninformative improper priors, $\pi_i^N(\theta_i)$ and $\pi_j^N(\theta_j)$, where throughout this paper the notation of superscript N implies the noninformative improper prior or its use, and have different dimensions in parameters. The fractional Bayes factor(FBF) of O'Hagan(1995) and the intrinsic Bayes factor(IBF) of Berger and Pericchi(1996, 1998) to overcome the problem due to the arbitrariness of noninformative improper priors are automatic and objective.

The idea of IBF is to use minimal training samples $\{X_m(l), l=1,2,\ldots,L\}$ to convert the improper prior to the proper posterior density. The minimal training sample implies the part of full sample with the minimal sample size to guarantee $0 < m_i^N(X|b=1) < \infty$ for all i. The IBF, $B_{ji}^I(l)$, given a minimal training sample, $X_m(l)$, for some l is defined by

$$B_{ji}^I(l) = B_{ji}^N(X|b=1) \cdot B_{ij}^N(X_m(l)|b=1).$$

But, practically to prevent the IBF from depending on only one minimal training sample is used an arithmetic IBF(AIBF) as substituting an arithmetic mean of $B_{ij}^N(X_m(l)|b=1)$, $l=1,2,\ldots,L$, for $B_{ij}^N(X_m(l)|b=1)$ for some l, a geometric IBF(GIBF) as a geometric mean, or a median IBF(MIBF) as a median. The FBF uses a fraction b of each likelihood function to change noninformative improper priors into proper priors. The FBF is defined by

$$B_{ji}^F = B_{ji}^N(X|b=1) \cdot B_{ij}^N(X|b). \tag{2.2}$$

O'Hagan proposed three ways to set b, where m is the minimal training sample size: (a) $b=m/n$, when robustness is no concern, (b) $b=n^{-1}\max\{m,\sqrt{n}\}$, when robustness is a serious concern, and (c) $b=n^{-1}\max\{m, \log n\}$, as an intermediate option. Generally, for a Bayesian multiple test the posterior probabilities of hypotheses via the Bayes factors are useful. Under the assumption of prior probability p_i of the model M_i being true the posterior probability of M_i via the FBF is given by

$$P(M_i|\boldsymbol{X}) = \left\{ \sum_{j=1}^{q}(p_j/p_i) \cdot B_{ji}^{F} \right\}^{-1}, \quad i=1,2,\dots,q. \tag{2.3}$$

3. Computation of The FBF

Let μ denote the common value of $\mu_1=\mu_2$, and Σ the common value of $\Sigma_1=\Sigma_2$. We use Jeffreys prior π_i^N, $i=0,1,2,3$, noninformative improper prior, for each model $M_i, i=0,1,2,3$, under the assumption of independence between a mean vector and a covariance matrix as follows

$$\begin{cases} \pi_0^N(\mu_1,\mu_2,\Sigma_1,\Sigma_2) = c_0 \Pi_{k=1}^{2} |\Sigma_k|^{-\frac{1}{2}(p+1)}, & \Sigma_1>0,\ \Sigma_2>0, \\ \pi_1^N(\mu,\Sigma) = c_1 |\Sigma|^{-\frac{1}{2}(p+1)}, & \Sigma>0, \\ \pi_2^N(\mu_1,\mu_2,\Sigma) = c_2 |\Sigma|^{-\frac{1}{2}(p+1)}, & \Sigma>0, \\ \pi_3^N(\mu,\Sigma_1,\Sigma_2) = c_3 \Pi_{k=1}^{2} |\Sigma_k|^{-\frac{1}{2}(p+1)}, & \Sigma_1>0,\ \Sigma_2>0, \end{cases} \tag{3.1}$$

where c_i, $i=0,1,2,3$, is an undefined normalizing constant.

Let $\boldsymbol{X}_k=\{X_{k1},X_{k2},\dots,X_{kn_k}\}$ be a p-variate random sample of size n_k from Π_k with a distribution $N_p(\mu_k,\Sigma_k)$, $k=1,2$. We use the following notation throughout this paper,

$$\begin{cases} n = n_1+n_2, \\ \boldsymbol{X} = \{\boldsymbol{X}_1,\boldsymbol{X}_2\}, \\ \overline{X}_k = \dfrac{\sum_{j=1}^{n_k} X_{kj}}{n_k}, \\ \overline{X} = \dfrac{\sum_{k=1}^{2} n_k \overline{X}_k}{n}, \\ V_k = \sum_{j=1}^{n_k}(X_{kj}-\overline{X}_k)(X_{kj}-\overline{X}_k)', \\ V = \sum_{k=1}^{2}\sum_{j=1}^{n_k}(X_{kj}-\overline{X})(X_{kj}-\overline{X})'. \end{cases}$$

The likelihood function $L_i(\cdot)$, $i=0,1,2,3$, under each hypothesis is given by

$$\begin{cases} L_0(\mu_1,\mu_2,\Sigma_1,\Sigma_2) = \Pi_{k=1}^{2}(2\pi)^{-\frac{n_k p}{2}}|\Sigma_k|^{-\frac{n_k}{2}}\exp\left\{-\frac{1}{2}tr[\Sigma_k^{-1}\Omega_k^*]\right\}, \\ L_1(\mu,\Sigma) = (2\pi)^{-\frac{np}{2}}|\Sigma|^{-\frac{n}{2}}\exp\left\{-\frac{1}{2}tr[\Sigma^{-1}\Omega]\right\}, \\ L_2(\mu_1,\mu_2,\Sigma) = (2\pi)^{-\frac{np}{2}}|\Sigma|^{-\frac{n}{2}}\Pi_{k=1}^{2}\exp\left\{-\frac{1}{2}tr[\Sigma^{-1}\Omega_k^*]\right\}, \\ L_3(\mu,\Sigma_1,\Sigma_2) = \Pi_{k=1}^{2}(2\pi)^{-\frac{n_k p}{2}}|\Sigma_k|^{-\frac{n_k}{2}}\exp\left\{-\frac{1}{2}tr[\Sigma_k^{-1}\Omega_k]\right\}, \end{cases} \tag{3.2}$$

where $\Omega = V + n(\mu-\overline{X})(\mu-\overline{X})'$, $\Omega_k = V_k + n_k(\mu-\overline{X}_k)(\mu-\overline{X}_k)'$, and $\Omega_k^* = V_k + n_k(\mu_k - \overline{X}_k)(\mu_k-\overline{X}_k)'$.

After using the kernel of multivariate normal density for the integration over a mean vector and the kernel of the inverted Wishart density for the integration over a covariance matrix the computation result of the function (2.2) with (3.1) and (3.2) is as follows

$$\begin{cases} m_0^N(\boldsymbol{X}|b) = c_0(2\pi)^{-\frac{(bn-1)p}{2}}\Pi_{k=1}^{2}\Delta_k(0)|V_k|^{-\frac{bn_k-1}{2}}n_k^{-\frac{p}{2}}, \\ m_1^N(\boldsymbol{X}|b) = c_1(2\pi)^{-\frac{(bn-1)p}{2}}\Delta(1)|V|^{-\frac{bn-1}{2}}n^{-\frac{p}{2}}, \\ m_2^N(\boldsymbol{X}|b) = c_2(2\pi)^{-\frac{(bn-2)p}{2}}\Delta(2)|V_1+V_2|^{-\frac{bn_k-2}{2}}\Pi_{k=1}^{2}n_k^{-\frac{p}{2}}, \\ m_3^N(\boldsymbol{X}|b) = c_3(2\pi)^{-\frac{bnp}{2}}\left\{\Pi_{k=1}^{2}\Delta_k(3)|V_k|^{-\frac{bn_k}{2}}\right\}\int_{-\infty}^{\infty}\Pi_{k=1}^{2}\{1+W_k(\mu)\}^{-\frac{bn_k}{2}}d\mu, \end{cases} \tag{3.3}$$

where

$$\begin{cases} \Delta_k(0) = 2^{\frac{p(bn_k-1)}{2}}b^{-\frac{bn_k p}{2}}\Gamma_p\left\{\frac{1}{2}(bn_k-1)\right\}, \\ \Delta(1) = 2^{\frac{p(bn-1)}{2}}b^{-\frac{bnp}{2}}\Gamma_p\left\{\frac{1}{2}(bn-1)\right\}, \\ \Delta(2) = 2^{\frac{p(bn-2)}{2}}b^{-\frac{bnp}{2}}\Gamma_p\left\{\frac{1}{2}(bn-2)\right\}, \\ \Delta_k(3) = 2^{\frac{bn_k p}{2}}b^{-\frac{bn_k p}{2}}\Gamma_p\left\{\frac{1}{2}bn_k\right\}, \\ W_k(\mu) = (\mu-\overline{X}_k)'S_k^{-1}(\mu-\overline{X}_k), \quad S_k = V_k/N_k, \\ \Gamma_p(t) = \pi^{\frac{p(p-1)}{4}}\Pi_{i=1}^{p}\Gamma(t-\frac{i-1}{2}). \end{cases}$$

The integration in $m_3^N(\boldsymbol{X}|b)$ of (3.3) is not analytically solved. This integration can be performed by the numerical integration procedure or Monte Carlo integration. In a simulation study of section 4, we estimate the integral function using the Monte Carlo integration method through the importance sampling. In $m_3^N(\boldsymbol{X}|b)$, we need to compute the following integral

$$M=\int_{-\infty}^{\infty} g(\mu)d\mu = \int_{-\infty}^{\infty} \frac{g(\mu)}{f_\mu(\mu)} \cdot f_\mu(\mu)\,d\mu = E_{f_\mu(\mu)}\left[\frac{g(\mu)}{f_\mu(\mu)}\right],$$

where $g(\mu)=\Pi_{k=1}^{2}\{1+W_k(\mu)\}^{-\frac{bn_k}{2}}$, and $f_\mu(\mu)$ is an importance sampling density function. The Monte Carlo estimate of M is $\hat{M}=\frac{1}{G}\sum_{j=1}^{G}\frac{g(\mu_j)}{f_\mu(\mu_j)}$, where μ_j, $j=1,2,\ldots,G$, is generated from the importance sampling density $f_\mu(\mu)$. It is well known that $Var(\hat{M})$ is small when $f_\mu(\mu) \propto |g(\mu)|$. The function $g(\mu)$ can be rewritten as

$$\begin{aligned} g(\mu) &= \Pi_{k=1}^{2}\exp\left\{-\frac{1}{2}\,bn_k\cdot\ln(1+W_k(\mu))\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left(bn_1W_1(\mu)+bn_2W_2(\mu)\right)\right\} \\ &\propto \exp\left\{-\frac{1}{2}(\mu-\mu_0)'J(\mu-\mu_0)\right\}, \end{aligned}$$

where the first proportional term is obtained using only the first term of Taylor series on $W_k(\mu)=0$ of $\ln\{1+W_k(\mu)\}$. Thus the importance sampling density function of μ is $N_p(\mu_0, J^{-1})$, where $\mu_0=K\overline{X}_1+(I_p-K)\overline{X}_2$, $K=bn_1(S_1J)^{-1}$, $J=b(n_1S_1^{-1}+n_2S_2^{-1})$, and I_p is a $p\times p$ identity matrix throughout this paper.

The size m of a minimal training sample equals to the condition that the marginal density $m_0^N(\boldsymbol{X}|b=1)$ of encompassing model M_0 is to be finite. If $n_k\le p$, then $rank(V_k)\le n_k-1\le p-1$. But $|V_k|=0$, since a matrix V_k is a $p\times p$ matrix. Hence $n_k\ge p+1$ for $k=1,2$. Thus the size of a minimal training sample is $m=2(p+1)$. The conventional selection of a fraction, b, of likelihood function in the computation of FBF is $b=m/n=2(p+1)/n$. For $k=1,2$ the sample size n_k must be restricted to $n_k\ge[p/b]+1$, where $[\cdot]$ is a Gauss symbol, in order that the arguments of gamma functions in $m_i^N(\boldsymbol{X}|b)$, $i=0,1,2,3$, are to be positive.

Finally, the computation of the FBF, and the posterior probabilities of hypotheses via the FBF are straightforward from (2.2) and (2.3).

4. A Simulation Study

We directly follow a simulation study of Kim and Kim(2001) originally based on the simulation scheme in Marks and Dunn(1974). All the experiments are performed for two independent p-variate $(p=2,4)$ normal samples with sample size $n_1=n_2=30$, the importance samples of size 500, and 200 replications assuming equal prior model probabilities. Let 0_p and 1_p be p-variate row vectors of zeroes and ones, respectively. Now we set $\mu_1 = 0_p$, $\mu_2 = [\tau(1+\sqrt{\lambda}), 0_{p-1}]'$, $\Sigma_1 = I_p$, and $\Sigma_2 = \Lambda$, where Λ is a diagonal matrix with a vector, $[\lambda \cdot 1_{p/2}, 1_{p/2}]'$ of diagonal elements. Here $\tau(\lambda)$ is the measure of degree of separation between means(covariances) of two populations. If $\tau = 0$ then $\mu_1 = \mu_2$ else $\mu_1 \neq \mu_2$. If $\lambda = 1$ then $\Sigma_1 = \Sigma_2$ else $\Sigma_1 \neq \Sigma_2$. Data with different choices, $\tau = 0, 2$ and $\lambda = 1, 4, 8$ of τ and λ are generated. For each data the FBF's are computed with the common use of fraction b, $b = m/n = 2(p+1)/n$.

Table 4.1 shows the results of the averages and the standard deviations in parentheses of posterior probabilities for each model based on 200 replications. Figure 4.1-4.6 are frequency plots on 200 replications of posterior probabilities for each model when $p = 4$. Though the plots of $p = 2$ are not presented here because of the limit of space, their behaviors are similar to $p = 4$.

Table 4.1 : The averages and the standard deviations in parentheses of posterior probabilities on 200 replications.

p	τ	λ	$P(M_0 \mid X)$	$P(M_1 \mid X)$	$P(M_2 \mid X)$	$P(M_3 \mid X)$
2	0	1	0.0044 (0.0084)	0.7872 (0.1448)	0.1256 (0.1061)	0.0829 (0.1141)
		4	0.0425 (0.0453)	0.1355 (0.2085)	0.0272 (0.0894)	0.7948 (0.2398)
		8	0.0721 (0.1213)	0.0111 (0.0626)	0.0012 (0.0070)	0.9155 (0.1353)
	2	1	0.0411 (0.0689)	0.0000 (0.0000)	0.9589 (0.0689)	0.0000 (0.0000)
		4	0.7425 (0.3064)	0.0000 (0.0000)	0.2575 (0.3064)	0.0000 (0.0000)
		8	0.9867 (0.0535)	0.0000 (0.0000)	0.0133 (0.0535)	0.0000 (0.0000)
4	0	1	0.0001 (0.0007)	0.8685 (0.1999)	0.1126 (0.1824)	0.0188 (0.0997)
		4	0.0141 (0.0627)	0.1691 (0.2574)	0.0276 (0.1060)	0.7891 (0.2963)
		8	0.0271 (0.0942)	0.0005 (0.0035)	0.0003 (0.0033)	0.9721 (0.0953)
	2	1	0.0021 (0.0158)	0.0000 (0.0000)	0.9979 (0.0158)	0.0000 (0.0000)
		4	0.6470 (0.3788)	0.0000 (0.0000)	0.3530 (0.3788)	0.0000 (0.0000)
		8	0.9931 (0.0572)	0.0000 (0.0000)	0.0069 (0.0572)	0.0000 (0.0000)

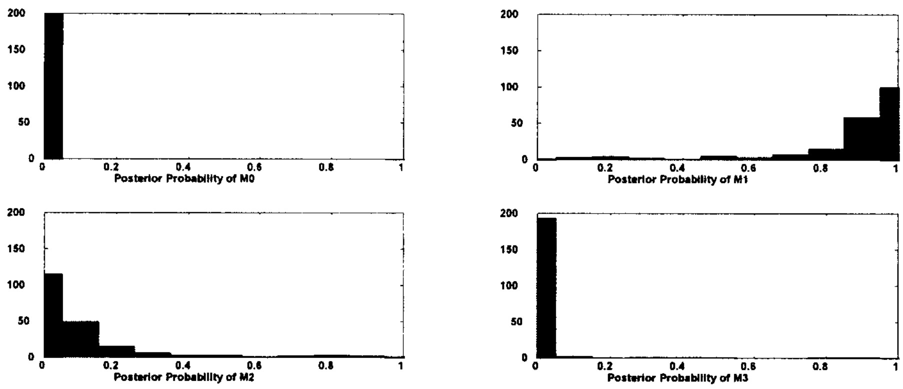

Figure 4.1 : The frequency plot when $(p, \tau, \lambda) = (4, 0, 1)$.

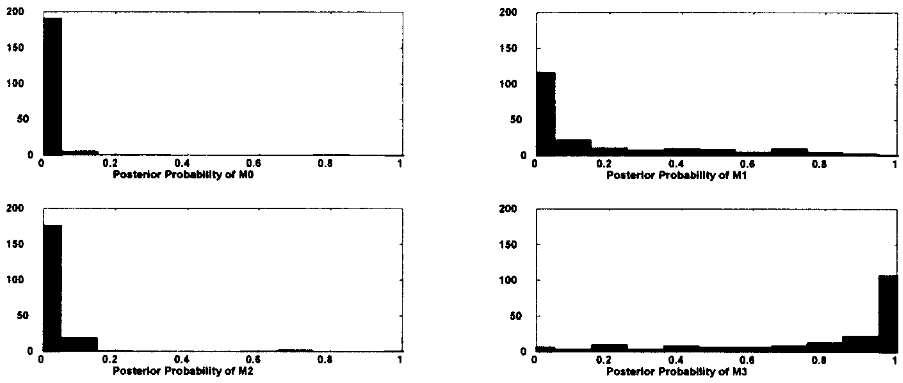

Figure 4.2 : The frequency plot when $(p, \tau, \lambda) = (4, 0, 4)$.

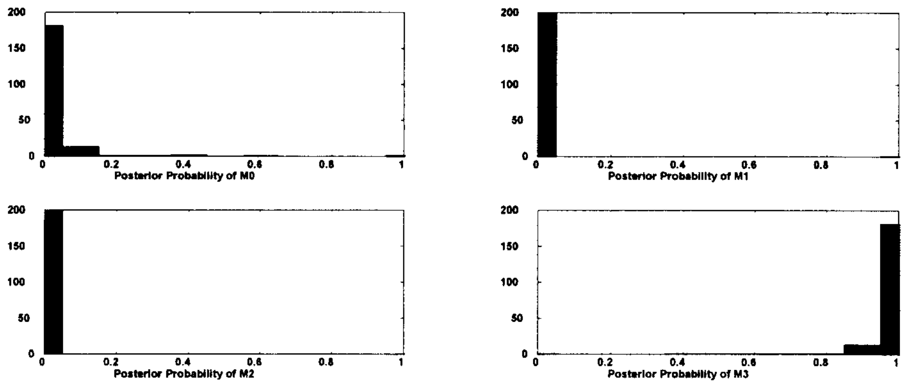

Figure 4.3 : The frequency plot when $(p, \tau, \lambda) = (4, 0, 8)$.

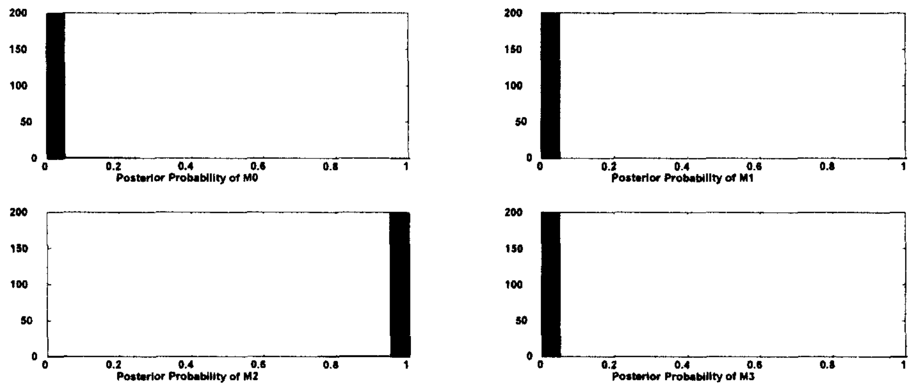

Figure 4.4 : The frequency plot when $(p, \tau, \lambda) = (4, 2, 1)$.

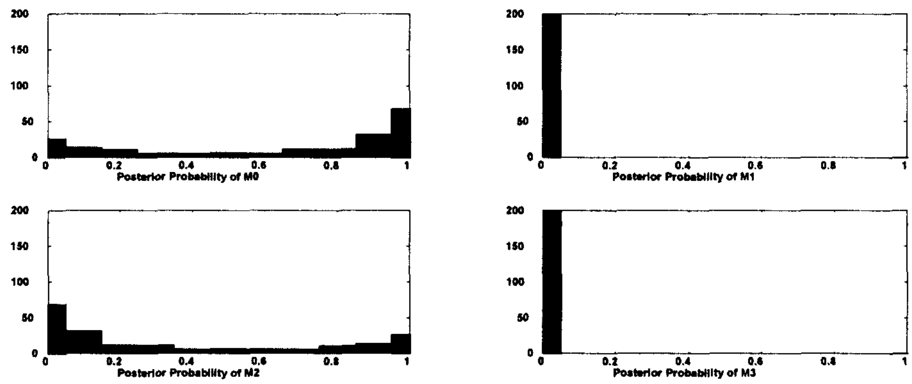

Figure 4.5 : The frequency plot when $(p, \tau, \lambda) = (4, 2, 4)$.

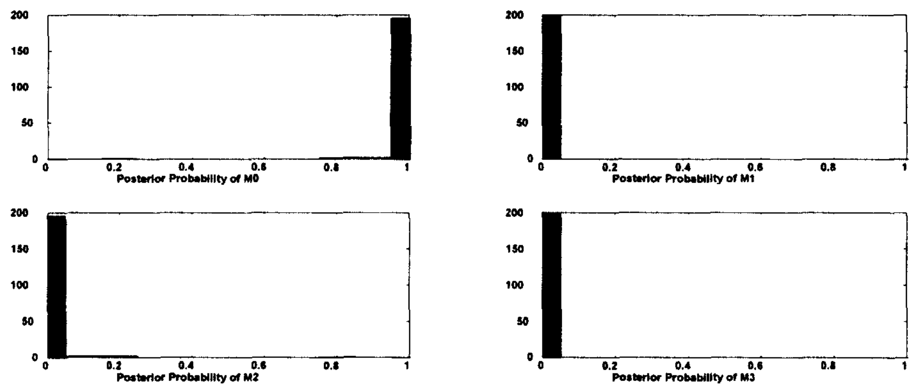

Figure 4.6 : The frequency plot when $(p, \tau, \lambda) = (4, 2, 8)$.

5. Concluding Remarks

We have proposed a Bayesian criterion for a multiple test of two independent multivariate normal populations. The test is performed by comparing with posterior probabilities of models via the FBF under the assumption of Jeffreys priors, noninformative improper priors. This multiple test doesn't require the prior knowledge or test on the equality or the unequalness of two means or two covariances, while the classical test requires that. Also, a Bayesian multiple test suggested in this paper can be flexibly applied to the classical tests of two independent multivariate normal populations. For example, the test of M_0' : $\Sigma_1=\Sigma_2$ *versus* M_1' : $\Sigma_1 \neq \Sigma_2$ is to reject M_0' if $P(M_0|X)+P(M_3|X) > 0.5$. Then the Beherens-Fisher problem, the test of M_3 versus M_0, can be solved by comparing $P(M_0|X)$ with $P(M_3|X)$.

Concerning with the use of the IBF, the number of minimal training samples possible over the full sample is $L = \binom{n_1}{p+1} \cdot \binom{n_2}{p+1}$. Under the work of Varshavksy(1995) based on the theory of U-statistics the number of minimal training samples is only mn with such an accuracy as all the possible minimal training samples. For a example of $n_1=n_2=30$, $mn=240$, $L=189,225$ when $p=1$, $mn=360$, $L=16,483,600$ when $p=2$, $mn=480$, $L=145,422,675$, when $p=3$, and $mn=600$, $L=20,308,000,000\ 675$, when $p=4$. Now, the IBF can be calculated by random sampling mn minimal training samples from total L minimal training samples. Of course, a sensitivity analysis for several sets of mn minimal training samples should be followed to check the stability of the IBF. But even though mn of the $B_{ij}^N(X_m(l)|b=1)$ are computed, a burden of computation remains since importance sampling for each minimal training sample must be performed. So, the computation of posterior probabilities of hypotheses via the AIBF, GIBF, or MIBF which additionally need times for sorting is a job requiring much more computation times than the FBF, while the FBF is very simple to use without the need of sampling minimal training samples. Also, we can see that the results in this paper via the FBF confirm to our theoretical expectation for the test.

References

[1] Anderson, T. W. (1984), *An Introduction to Multivariate Statistical Analysis*, second edition, John Wiley and Sons, Inc.

[2] Berger, J. O. and Pericchi, L. R.(1996), The Intrinsic Bayes Factor for Model Selection and Prediction, *Journal of the American Statistical Association*, Vol. 91, No.433, 109-122.

[3] Berger, J. O. and Pericchi, L. R.(1998), Accurate and Stable Bayesian Model Selection: The Median Intrinsic Bayes Factor, *Sankhya*, B, Vol.60, 1-18.

[4] Jeffreys, H.(1961), *Theory of Probability*, London: Oxford University Press.

[5] Kim, H. J. and Kim, S. W.(2001), Two-Smple Bayesian Test using Intrinsic Bayes Factors for Multivariate Normal Observations, *Communications in Statistics : Simulation and Computation*.

[6] Krzanowski, W. J. and Marriott, F. H. C. (1994), *Multivariate Analysis; Part 1 Distributions, Ordination and Inference*, Edward Arnold.

[7] Marks, S. & Dunn, O.J.(1974) Discriminant Funtions When Covariance Matrices are Unequal, *Journal of the American Statistical Association*, 69, pp. 555-559.

[8] Nagao, H. (1973), On Some Test Criteria for Covariance Matrix, *Annals of Statistics*, Vol. 1, 700-709.

[9] O'Hagan, A.(1995), Fractional Bayes Factors for Model Comparsion, *Journal of the Royal Statistical Society*, B, Vol.57, No.1, 99-138.

[10] Perlman, M. D. (1980), Unbiasedness of the Likelihood Ratio Tests for Equality of Several Covariance Matrices and Equality of Several Multivariate Normal Populations, *Annals of Statistics*, Vol. 8, 247-263.

[11] Siotani, M. (1987), *Advances in Multivariate Analysis*, A. K. Gupta (ed.), D. Reidel Publishing Co.

[12] Varshavsky, J(1995), On the Development of Intrinsic Bayes Factors, Ph. D. Thesis, Purdue University.

[13] Yao, Y. (1965), An Approximate Degrees of Freedom Solution to the Multivariate Behrens-Fisher problem, *Biometrika*, Vol. 52, 139-147.