2002 Proceedings of the Spring Conference, Korean Statistical Society

A Bayesian Comparison of Two Multivariate Normal Genralized Variances

Hea-Jung Kim*

Abstract

In this paper we develop a method for constructing a Bayesian HPD (highest probability density) interval of a ratio of two multivariate normal generalized variances. The method gives a way of comparing two multivariate populations in terms of their dispersion or spread, because the generalized variance is a scalar measure of the overall multivariate scatter. Fully parametric frequentist approaches for the interval is intractable and thus a Bayesian HPD (highest probability densith) interval is pursued using a variant of weighted Monte Carlo (WMC) sampling based approach introduced by Chen and Shao (1999). Necessary theory involved in the method and computation is provided.

Key Wards : HPD interval; multivarite normal population; ratio of two generalized variances; vague prior; weighted Monte Carlo approach

1 Introduction

In applications where variability of the multivariate population is of great practical importance, the generalized variance can be used to rank distinct groups and populations in order of their dispersion or spread. For example, a certain product, such as semiconductor, produced by a number of companies is characterized by a vector of p measurements. Although the same product is produced on the average, the companies can be distinguished on the basis of their associated generalized variances. The usage of the generalized variance has been widely accepted by statisticians (see Grizzle and Allen 1969, Press 1982, and Rencher 1995). However, due to complex sampling distribution involved in inferencing the generalized variance, the analysis of it is yet to be seen in applied settings.

*Department of Statistics, Dongguk University, Seoul 100-715, Korea, kim3hj@dongguk.edu

This paper, therefore, consider how inferences might be made about a possible difference in the variability of two multivariate normal populations. In comparing generalized variances of two multivariate normal populations, $N_p(\theta_1, \Sigma_1)$ and $N_p(\theta_1, \Sigma_1)$, there are many measures that could be used besides $|\Sigma_1|/|\Sigma_2|$, for example, the reciprocal ratio $|\Sigma_2|/|\Sigma_1|$, $|\Sigma_1|^{1/2}/|\Sigma_2|^{1/2}$and $\log(|\Sigma_1|) - \log(|\Sigma_2|)$. Unlike a Bayesian credible interval, however, the HPD interval is not invariant under a nonlinear transformation (see Box and Tiao 1992). Therefore, among them we focuses on a Bayesian method to coming at the HPD interval of $|\Sigma_1|/|\Sigma_2|$. This is done by deriving the joint posterior distribution of the eigen values of $(V_1^{1/2}\Lambda_1 V_1^{1/2})^{-1/2}(V_2^{1/2}\Lambda_2 V_2^{1/2})(V_1^{1/2}\Lambda_1 V_1^{1/2})^{-1/2}$, where $\Lambda_i = \Sigma_i^{-1}$, $i = 1, 2$, precision matrices, and V_1 and V_2 are the mean corrected sums of squares and cross product matrices obtained from samples of two populations, respectively. Under the posterior distribution, we are driven to do multidimensional integration to evaluate the HPD interval of $|\Sigma_1|/|\Sigma_2|$. Since $|\Sigma_1|/|\Sigma_2|$ is nonlinear function of the eigen values, we can't expect analytical and exact numerical evaluation of the interval. As an alternative solution, we develop a weighted Monte Carlo (WMC), a variant of Chen and Shao (1999) method, to compute the interval. Finally, an illustrative example is given to demonstrate utility of the proposed model.

2 Preliminary

In multivariate normal theory, Bayesian posterior distribution of a precision matrix comes out as an Wishart distribution. So that many Bayesian inferences for the multivariate normal covariance matrices (inverse of the precision matrices) involve a function of Wishart matrices. Same case applies to our HPD inteval. To derive our main results (The HPD interval of $|\Sigma_1|/|\Sigma_2|$), we shall need to make use of the following two lemmas of some functions of Wishart matrices.

Lemma 1. (Olkin and Rubin 1964) Let S_1 and S_2 be independently distributed, $S_i \sim W_p(I_p, n_i), i = 1, 2$, $S_1 = (S_1^{1/2})^2$. The distribution of $Y = S_1^{-1/2} S_2 S_1^{-1/2} \sim B_{II}(p; n_1/2, n_2/2)$, a multivariate Beta II distribution, with density given by

$$f_{II}(Y) = \{B_p(n_1/2, n_2/2)\}^{-1} |Y|^{(n_1-p-1)/2} |I_p + Y|^{-(n_1+n_2)/2}, \quad Y > 0, \qquad (1)$$

where $B(a,b) = \Gamma_p(a)\Gamma_p(b)/\Gamma_p(a+b)$ and $\Gamma_p(\lambda) = \pi^{p(p-1)/2} \prod_{j=1}^{p} \Gamma(\lambda - (j-1)/2)$.

Lemma 2. If $Y \sim B_{II}(p; n_1/2, n_2/2)$, then $W = (I_p + Y)^{-1} Y \sim B_I(p; n_1/2, n_2/2)$, a multivariate Beta I distribution, with its density

$$f_I(W) = \{B_p(n_1/2, n_2/2)\}^{-1} |W|^{(n_1-p-1)/2} |I_p - W|^{(n_2-p-1)/2}, \quad 0 < W < I_p. \qquad (2)$$

So that in univariate case, if $W \sim Beta(a,b)$, and if $Y = W(1-W)^{-1}$, then $Y \sim Beta^{-1}(a,b)$, an inverted Beta distribution (see Raiffa and Schlaifer, 1961, p. 221), whose density is given by

$$f(y) = \{B(a,b)\}^{-1} y^{(a-1)}(1+y)^{-(a+b)}, \;\; y > 0. \tag{3}$$

Above distributions lead to multivariate analogs of the Beta or F distribution. Some of these distributions arise naturally in various multivariate problems, e.g., multivariate analysis of variance tests and multivariate slippage problems (see Box and Tiao 1992).

3 Posterior Distrbution of the Eigen Values

Suppose $X_1(1), \ldots, X_{N_1}(1)$ are independent p-vectors from $N_p(\theta_1, \Lambda_1^{-1})$, $X_1(2), \ldots, X_{N_2}(2)$ are independent p-vectors from $N_p(\theta_2, \Lambda_2^{-1})$ and are independent of the $X_j(1)$'s, where $\Lambda_i = \Sigma_i$, the percision matrix. Let

$$\bar{X}(1) = \sum_{j=1}^{N_1} X_j(1), \;\; \bar{X}(2) = \sum_{j=1}^{N_2} X_j(2), \;\; V_1 = \sum_{j=1}^{N_1}(X_j(1) - \bar{X}(1))(X_j(1) - \bar{X}(2))',$$

and $V_2 = \sum_{j=1}^{N_2}(X_j(2) - \bar{X}(2))(X_j(2) - \bar{X}(2))'$. Then the joint p.d.f. of $\bar{X}(1), \bar{X}(2), V_1$, and V_2 is proportional to

$$\prod_{i=1}^{2} |V_i|^{(N_i-p-2)/2} |\Lambda_i|^{N_i/2} \mathrm{etr}\left\{-\frac{1}{2}\Lambda_i[V_i + N_i(\bar{X}(i) - \theta_i)(\bar{X}(i) - \theta_i)']\right\}, \tag{4}$$

where $\mathrm{etr}\{A\} = \exp\{\mathrm{tr}A\}$. To assure very little information is contirbuted to the analysis by a subjective prior density, we assume diffuse prior

$$p(\theta_1, \theta_2, \Lambda_1, \Lambda_2) \propto \prod_{i=1}^{2} |\Lambda_i|^{-(p+1)/2}. \tag{5}$$

The joint posterior density of the parameters is proportional to

$$\prod_{i=1}^{2} |\Lambda_i|^{(N_i-p-1)/2} \mathrm{etr}\left\{-\frac{1}{2}\Lambda_i[V_i + N_i(\bar{X}(i) - \theta_i)(\bar{X}(i) - \theta_i)']\right\}. \tag{6}$$

Integrating (6) with respect θ_i's, we have the marginal posterior distributions of Λ_i, $i = 1, 2$,

$$\Lambda_i | \bar{X}(i), V_i \sim W_p(V_i^{-1}, N_i - 1), \;\; i = 1, 2, \tag{7}$$

for $N_i \geq p+1$, a Wishart distribution with scale parameter V_i^{-1} and $N_i - 1$ degrees of freedom.

Theorem 1. The posterior distribution of

$$\Omega = (V_1^{1/2}\Lambda_1 V_1^{1/2})^{-1/2}(V_2^{1/2}\Lambda_2 V_2^{1/2})(V_1^{1/2}\Lambda_1 V_1^{1/2})^{-1/2}$$

is $B_{II}(p; n_1/2, n_2/2)$, a multivariate Beta II distribution, with density given by

$$p(\Omega|Data) = \{B_p(n_1/2, n_2/2)\}^{-1}|\Omega|^{(n_1-p-1)/2}|I_p + \Omega|^{-(n_1+n_2)/2}, \quad \Omega > 0, \qquad (8)$$

where $n_i = N_i - 1$.

Corollary 1. Let components of $\lambda' = (\lambda_1, \ldots, \lambda_p)$, $\lambda_1 > \cdots > \lambda_p > 0$, are ordered eigen values of Ω, the joint postrior distribution of λ is

$$p(\lambda|Data) = c_0 \left\{\prod_{\ell=1}^{p} \lambda_\ell^{(n_1-p-1)/2}\right\}\left\{\prod_{\ell=1}^{p}(1+\lambda_\ell)^{-(n_1+n_2)/2}\right\} J(\lambda), \qquad (9)$$

over $\{\lambda : \lambda_1 > \cdots > \lambda_p > 0$, where

$$c_0^{-1} = \pi^{-p^2/2}\Gamma_p\{p/2\}\{B_p(n_1/2, n_2/2)\} \text{ and } J(\lambda) = \prod_{\ell<v}^{p}(\lambda_\ell - \lambda_v).$$

4 Estimation of the HPD interval for $|\Sigma_1|/|\Sigma_2|$

4.1 The HPD interval

The distribution (9) enables us to obtain various posterior quantity of the form $h(|\Omega|)$, where $|\Omega| = (|\Sigma_1|/|\Sigma_2|)(|V_2|/|V_1|) = \prod_{\ell=1}^{p}\lambda_\ell$. Especially the HPD interval of $|\Sigma_1|/|\Sigma_2|$ can be expressed as an interval-type posterior quantity: A $(1-\alpha)\times 100\%$ HPD interval of $|\Sigma_1|/|\Sigma_2|$ is given by

$$R(\pi_\alpha) = \left\{|V_1|/|V_2|\prod_{\ell=1}^{p}\lambda_\ell : \; p(\prod_{\ell=1}^{p}\lambda_\ell|Data) \geq \pi_\alpha\right\},$$

where π_α is the largest contant such that

$$\int_{R^p} \mathrm{I}\left\{|\mathrm{V}_1|/|\mathrm{V}_2|\prod_{\ell=1}^{\mathrm{p}}\lambda_\ell \in \mathrm{R}(\pi_\alpha)\right\} p(\lambda|\mathrm{Data})\, \mathrm{d}\lambda \geq 1-\alpha,$$

where $I\{\cdot\}$ is the indicator function.

The analytic evaluation of $R(\pi_\alpha)$ is not available for closed form of the posterior p.d.f. of $\prod_{\ell=1}^{p}\lambda_\ell$ is not known. Moreover, computing the HPD interval is difficult and challenging, because computation of $R(\pi_\alpha)$ requires knowing π_α and then calculating the content defined by $R(\pi_\alpha)$. Recently, however, Chen and Shao (1999) propose a simple Monte Carlo method whose variant is applicable for the computation of the HPD interval. The merit of a variant of Chen and Shao's method is that the method does not require a closed-form expression of the marginal posterior distribution of $\prod_{\ell=1}^{p} \lambda_\ell$. Instead we only need an important function of the joint p.d.f. of λ to estimate the HPD interval.

4.2 Weighted Monte Carlo Method

Assume that $\{\lambda^{(j)}, j = 1, \ldots, m\}$ is an MCMC sample from an appropriate importance function $g(\lambda)$. Then we can obtain an HPD interval for $\xi = \prod_{\ell=1}^{p} \lambda_\ell$ using the weighted Monte Carlo method. Procedure of the method as follows: Let $\{\xi_j, j = 1, \ldots, m\}$ be a MCMC sample calculated from $\{\lambda^{(j)}, j = 1, \ldots, m\}$. Also let $\xi_{(j)}$ denote the ordered values of $\xi_j, j = 1, \ldots, m$. Then γth quantile of the marginal posterior distribution of ξ can be estimated by

$$\hat{\xi}^{(\gamma)} = \begin{cases} \xi_{(1)} & \text{if} \quad \gamma = 0 \\ \xi_{(j)} & \text{if} \quad \sum_{k=1}^{j-1} w_{(k)} < \gamma \le \sum_{k=1}^{j} w_{(k)}, \end{cases} \tag{10}$$

where $w_{(k)}$ is thw weight function associated with the k th ordered value $\xi_{(k)}$. More specifically, we first compute

$$w_k = \frac{p(\lambda^{(k)}|Data)/g(\lambda^{(k)})}{\sum_{j=1}^{m} p(\lambda^{(j)}|Data)/g(\lambda^{(j)})}. \tag{11}$$

Then we rewrite $\{w_k, k = 1, \ldots, m\}$ as $\{w_{(k)}, k = 1, \ldots, m\}$ so that kth value of $w_{(k)}$ corresponse to the kth ordered value $\xi_{(k)}$. Using (10) we compute

$$R_k(m) = \left(\hat{\xi}^{(k/m)},\ \hat{\xi}^{(k+[(1-\alpha)/m])}\right), \tag{12}$$

for $k = 1, 2, \ldots, m - [(1-\alpha)m]$ and a $100(1-\alpha)\%$ HPD interval of ξ is $R_{k^*}(m)$ that has smallest interval width among all $R_k(m)$'s. Here $[x]$ denotes the largest integer part of x. Thus desired HPR inteval for $|\Sigma_1|/|\Sigma_2|$ can be obtained by multiplying $|V_1|/|V_2|$ on both bounds in $R_{k^*}(m)$, for the HPD is invariant with respect to linear transformation (see Box and Tiao 1992). From (11), it is easy to observe that it is requires to know the joint posterior density $p(\lambda|Data)$ only up to a normalizing constant, since this normalizing constant camcels out in the calculation of w_k.

Comparing (3) and (9), we see that the most natural candidate for the importance function $g(\lambda)$ is the joint density of ordered statistics from a random sample of inverted Beta variates, $\lambda_\ell \sim Beta^{-1}((n_1 - p + 1)/2, (n_2 + p - 1)/2)$, $\ell = 1, \ldots, p$. This similarity is shape is exploited in developing the impotance sampling procedure. Undered the importance function the Gibbs sampler is consisting of a sequence of truncated inverted Beta distributions, so that, for $\ell = 1, \ldots, p$, the full conditional distributions are

$$[\lambda_\ell^{(j)} | \lambda_1^{(j)}, \ldots, \lambda_{\ell-1}^{(j)}, \lambda_{\ell+1}^{(j-1)}, \ldots, \lambda_p^{(j-1)}] \sim Beta^{-1}(a, b) I \left(\lambda_{\ell-1}^{(j)} < \lambda_\ell^{(j)} < \lambda_\ell^{(j-1)}\right),$$

where $a = (n_1 - p + 1)/2$, $b = (n_2 + p - 1)/2$, and $\lambda^{(j)} = (\lambda_1^{(j)}, \ldots, \lambda_p^{(j)})'$. A simple accept/reject algorithm applies for generating the truncated inverted Beta variate $\lambda_\ell^{(j)}$. Furthermore the weight in (11) reduces to

$$w_k = J(\lambda^{(k)}) / \sum_{j=1}^{m} J(\lambda^{(j)}), \quad \text{where } J(\lambda^{(j)}) = \prod_{\ell < v}^{p} (\lambda_\ell^{(j)} - \lambda_v^{(j)}). \tag{13}$$

5 Concluding Remarks

In this paper we propose a weighted Monte Carlo method for estimating HPD interval of the ratio of two generalized variances. The ratio is a useful criterion for ranking the dispersion or spread of two multivariate normal populations in terms of scalar measure. The methodology proposed in this paper can easily extended to compare internal dispersions among K multivariate normal populations. An immediate method for multiple comparison is to construct the simultaneous HPD intervals of $|\Sigma_1|/|\Sigma_2|, \ldots, |\Sigma_{K-1}|/|\Sigma_K|$ using the Bonferroni method.

References

Box, G. E. P. and Tiao, G. C. (1992). *Bayesian Inference in Statistical Analysis.* New York: Wiley.

Chen, M.-H, and Shao, Q.-M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of the Computational and Graphical Statistics* 8, 69-92.

Grizzle, J. E. and Allen, D. M. (1969). Analysis of growth and does response curves. *Biometrics* 25, 357-382.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* New York: Wiely.

Olkin, I. and Tomsky, J. L. (1981). A new class of multivariate tests based on the union-intersection principle. *Annals of Statistics* 9, 792-802.

Press, S. J. (1982). *Applied Multivariate Analysis.* Florida: Krieger Publishing Co.

Raiffa, H. and Schlaifer, R. (1961). *Applied Statistical Decision Theory.* Boston: Harvard University Press.

Rencher, A. C. (1995).*Methods of Multivariate Analysis.* New York: Wiley.