Journal of the Korean
Data & Information Science Society
2006, Vol. 17, No. 2, pp. 559 ~ 568

Bayesian Analysis for the Ratio of Variance Components[1)]

Sang Gil Kang[2)]

Abstract

In this paper, we develop the noninformative priors for the linear mixed models when the parameter of interest is the ratio of variance components. We developed the first and second order matching priors. We reveal that the one-at-a-time reference prior satisfies the second order matching criterion. It turns out that the two group reference prior satisfies a first order matching criterion, but Jeffreys′ prior is not first order matching prior. Some simulation study is performed.

Keywords : Matching Prior, Ratio of Variance Components, Reference Prior

1. Inroduction

Consider the linear mixed model:

$$y_{ij} = x^T_{ij}\beta + \alpha_i + \varepsilon_{ij},\ i = 1, \cdots, k,\ j = 1, \cdots, n, \tag{1}$$

where x_{ij} is the $p\times 1$ vector of known covariates, β is the $p\times 1$ vector of unknown regression parameters. Also $\varepsilon_{ij} \sim N(0, \sigma^2)$ and $\alpha_i \sim N(0, \sigma^2_\alpha)$ are independently distributed. Let $\theta_1 = \sigma^2_\alpha / \sigma^2$ be our of interest.

The present paper focuses on noninformative priors for θ_1. We consider Bayesian priors such that the resulting credible intervals for θ_1 have coverage probabilities equivalent to their frequentist counterparts. Although this matching

1) This research was supported by Sangji University Research Fund, 2004.
2) Assistant Professor, Department of Applied Statistics, Sangji University, Wonju, 220-702, Korea.
E-mail : sangkg@mail.sangji.ac.kr

can be justified only asymptotically, our simulation results indicate that this is indeed achieved for small or moderate sample sizes as well.

This matching idea goes back to Welch and Peers (1963). Interest in such priors revived with the work of Stein (1985) and Tibshirani (1989). Among others, we may cite the work of Mukerjee and Dey (1993), DiCiccio and Stern (1994), Datta and Ghosh (1995a,b, 1996), Mukerjee and Ghosh (1997), Kang, Kim and Lee (2005).

On the other hand, Ghosh and Mukerjee (1992), and Berger and Bernardo (1989,1992) extended Bernardo's (1979) reference prior approach, giving a general algorithm to derive a reference prior by splitting the parameters into several groups according to their order of inferential importance. This approach is very successful in various practical problems. Quite often reference priors satisfy the matching criterion described earlier.

The ratio of variance components in random effect models has been of interest for a long time. Especially in animal science, this ratio is usually used to estimate the genetic heritability of a certain trait of livestock breeders (Graybill et al., 1956). One difficult part of the analysis of the random effect models from the sampling theory of view is the possible negative estimates for σ^2_a as well as for θ_1. Thus a Bayesian analysis for this model is desirable, not only because of its intrinsic merit, but also because it can resolve this problem.

The problem of estimating variance components in the one-way random effect model has been investigated by many authors from the Bayesian point of view. We may refer to Hill (1965), Box and Tiao (1973), Palmer and Broemeling (1990), among others. For noninformative priors, Ye (1994) developed the reference priors for θ_1, examined frequentist coverage probabilities for various θ_1 and compared risk functions of the Bayes estimators for reference priors. Kim, Kang and Lee (2001) provided a class of second order probability matching priors for θ_1. It is shown that among all of the reference priors, the only one-at-a-time reference prior satisfies a second order matching criterion.

The considered model is a general class model that strictly contains the balanced one-way random effect model (Ye, 1994). The outline of the remaining sections is as follows. In Section 2, we develop first order and second order probability matching priors for θ_1. Also we derive the reference priors for the parameters. It turns out that the two group and the one-at-a-time reference priors satisfies a first order matching criterion, but Jeffreys' prior is not a first order matching prior. Also the one-at-a-time reference prior is a second order matching prior. We provide that the propriety of the posterior distribution for the reference priors as well as the second order matching prior. In Section 4, simulated frequentist coverage probabilities under the proposed priors are given.

2. The Noninformative Priors

2.1 The Probability Matching Priors

For a prior π, let $\theta_1^{1-\alpha}(\pi;Y)$ denote the $(1-\alpha)$th percentile of the posterior distribution of θ_1, that is,

$$P^{\pi}[\theta_1 \le \theta_1^{1-\alpha}(\pi;Y) \mid Y] = 1-\alpha, \tag{2}$$

where $\theta = (\theta_1, \cdots, \theta_t)^T$ and θ_1 is the parameter of interest. We want to find priors π for which

$$P[\theta_1 \le \theta_1^{1-\alpha}(\pi;Y) \mid \theta] = 1-\alpha + o(n^{-u}), \tag{3}$$

for some $u > 0$, as n goes to infinity. Priors π satisfying (3) are called matching priors. If $u = 1/2$, then π is referred to as a first order matching prior, while if $u = 1$, π is referred to as a second order matching prior.

In order to find such matching priors π, let

$$\theta_1 = \sigma_\alpha^2/\sigma^2, \quad \theta_2 = \sigma_\alpha^{-2}(1+n\sigma_\alpha^2/\sigma^2)^{-1/n} \text{ and } \theta_3 = \beta.$$

With this parametrization, the likelihood function of parameters $(\theta_1, \theta_2, \theta_3)$ for the model (1) is given by

$$L(\theta_1, \theta_2, \theta_3) \propto \theta^{\frac{N}{2}} \exp - \frac{\theta_2(1+n\theta_1)^{1/n}}{2} \sum_{i=1}^{k} (y_i - X_i\theta_3)^T \left(I - \frac{\theta_1}{1+n\theta_1} J\right)(y_i - X_i\theta_3), 2 \tag{4}$$

where $X_i = (x_{i1}, \cdots, x_{in})^T$ is the $n \times p$ design matrix, $y_i = (y_{i1}, \cdots, y_{in})^T$ is the $n \times 1$ vector, I is the identity matrix of order n, J is the $n \times n$ matrix with each element equal to 1 and $N = nk$. Based on (4), the Fisher information matrix is given by

$$I_F = Diag\{I_{11}, I_{22}, I_{33}\},$$

where $I_{11} = \frac{1}{2} kn(n-1)(1+n\theta_1)^{-2}$, $I_{22} = \frac{nk}{2}\theta_2^{-2}$ and $I_{33} = \theta_2(1+n\theta_1)^{1/n} \sum_{i=1}^{k}(X_i^T X_i - \frac{\theta_1}{1+n\theta_1} X_i^T J X_i)$. From the above Fisher information matrix I_F, θ_1

is orthogonal to (θ_2, θ_3) in the sense of Cox and Reid(1987). Following Tibshirani(1989), the class of first order probability matching prior is characterized by

$$\pi_m^{(1)}(\theta_1, \theta_2, \theta_3) \propto (1+n\theta_1)^{-1} d(\theta_2, \theta_3), \tag{5}$$

where $d(\theta_2, \theta_3) > 0$ is an arbitrary function differentiable in its arguments.

The class of prior given in (5) can be narrowed down to the second order probability matching priors as given in Mukerjee and Ghosh (1997). A second order probability matching prior is of the form (5), and also d must satisfy an additional differential equation (cf (2.10)) of Mukerjee and Ghosh (1997), namely

$$\begin{aligned} &\frac{1}{6} d(\theta_2, \theta_3) \frac{\partial}{\partial \theta_1} \{ I_{11}^{-\frac{3}{2}} L_{1,1,1} \} + \frac{\partial}{\partial \theta_2} \{ I_{11}^{-\frac{1}{2}} L_{112} I^{22} d(\theta_2, \theta_3) \} \\ &+ \sum_{v=3}^{p+2} \sum_{s=3}^{p+2} \frac{\partial}{\partial \theta_v} \{ I_{11}^{-\frac{1}{2}} L_{11s} I^{sv} d(\theta_2, \theta_3) \} = 0, \end{aligned} \tag{6}$$

where

$$L_{1,1,1} = E[(\frac{\partial \log L}{\partial \theta_1})^3] = c_1 (1+n\theta_1)^{-3}, \quad c_1 = \text{ a constant}$$

$$L_{112} = E[\frac{\partial^3 \log L}{\partial \theta_1^2 \partial \theta_2}] = c_2 \theta_2^{-1} (1+n\theta_1)^{-2}, \quad c_2 = \text{ a constant}$$

$$L_{11k} = E[\frac{\partial^3 \log L}{\partial \theta_1^2 \partial \theta_k}] = 0, \; k = 3, \cdots, p+2,$$

$$I_{11} = \frac{1}{2} kn(n-1)(1+n\theta_1)^{-2}, \; I^{22} = \frac{2}{nk} \theta_2^2.$$

Then (6) simplifies to

$$\frac{\partial}{\partial \theta_2} \{ \theta_2 d(\theta_2, \theta_3) \} = 0. \tag{7}$$

Hence the set of solution of (7) is of the form

$$d(\theta_2, \theta_3) = \theta_2^{-1} h(\theta_3),$$

where $h(\theta_3) > 0$ is an arbitrary function differentiable in its arguments. Thus the resulting second order probability matching prior is given by

$$\pi_m^{(2)}(\theta_1, \theta_2, \theta_3) \propto (1+n\theta_1)^{-1} \theta_2^{-1} h(\theta_3). \tag{8}$$

2.2 The Reference Priors

Reference priors introduced by Barnardo (1979), and extended further by Berger and Barnardo (1992) have become very popular over the years for the development of noninformative priors. In this section, we derive the reference priors for different groups of ordering of $(\theta_1, \theta_2, \theta_3)$. Then due to the orthogonality of the parameters, following Datta and Ghosh (1995b), choosing rectangular compacts for each θ_1, θ_2 and θ_3 when θ_1 is the parameter of interest, the reference priors are given by as follows.

For the linear mixed model (1), if θ_1 is the parameter of interest, then the reference prior distribution for group of ordering of $\{(\theta_1, \theta_2, \theta_3)\}$ is

$$\pi_1 \propto \theta_2^{-\frac{2-p}{2}}(1+n\theta_1)^{\frac{p-2n}{2n}} \left| \sum_{i=1}^{k}(X_i^T X_i - \frac{\theta_1}{1+n\theta_1} X_i^T J X_i)\right|^{\frac{1}{2}}.$$

For group of ordering of $\{\theta_1, (\theta_2, \theta_3)\}$, the reference prior is

$$\pi_2 \propto \theta_2^{-\frac{2-p}{2}}(1+n\theta_1)^{-1}.$$

And for group of ordering of $\{\theta_1, \theta_2, \theta_3\}$, the reference prior is

$$\pi_3 \propto \theta_2^{-1}(1+n\theta_1)^{-1}.$$

Remark 2. The reference priors π_2 and π_3 satisfy a first order matching criterion, but Jeffreys' prior π_1 is not a first order matching prior. The one-at-a-time reference prior satisfies the second order matching criterion.

Notice that the matching priors (8) include many different matching priors because of the arbitrary selection of the function h. The θ_3 is the location parameter vector. So we consider a particular second order matching prior where h is a constant in matching priors. Because the matching priors based on the selection of this function are exactly the one-at-a-time reference priors. This prior is given by

$$\pi_m^{(2)}(\theta_1, \theta_2, \theta_3) \propto (1+n\theta_1)^{-1}\theta_2^{-1}. \qquad (9)$$

3. Implementation of the Bayesian Procedure

We investigate the propriety of posteriors for a general class of priors which include the reference priors and the second order matching prior (9). We consider the class of priors

$$\pi_G(\theta_1, \theta_2, \theta_3) \propto (1+n\theta_1)^{-a}\theta_2^{-b}. \tag{10}$$

where $a>0$ and $b>0$. The following general theorem can be proved.

Theorem 1. The posterior distribution of θ_1, θ_2 and θ_3 under the prior (10), is proper if $N-2b-p+2>0$ and $N+2na-2b-2n+2>0$, where $N=nk$.

Proof. Under the prior (10), the joint posterior for θ_1, θ_2 and θ_3 given y is

$$\begin{aligned} &\pi(\theta_1, \theta_2, \theta_3 \mid y) \\ \propto\ & \theta_2^{\frac{N-2b}{2}}(1+n\theta_1)^{-a} \\ \times\ & \exp\{-\frac{\theta_2(1+n\theta_1)^{\frac{1}{n}}}{2}\sum_{i=1}^{k}(y_i-X_i\theta_3)^T(I-\frac{\theta_1}{1+n\theta_1}J)(y_i-X_i\theta_3)\}. \end{aligned} \tag{11}$$

Now for $i=1,\cdots,k$,

$$(y_i-X_i\theta_3)^T(I-\frac{\theta_1}{1+n\theta_1}J)(y_i-X_i\theta_3)\geq(y_i-X_i\theta_3)^T(I-\frac{1}{n}J)(y_i-X_i\theta_3).$$

Thus

$$\begin{aligned} \pi(\theta_1, \theta_2, \theta_3 \mid y) \leq\ & \theta_2^{\frac{N-2b}{2}}(1+n\theta_1)^{-a} \\ \times\ & \exp\{-\frac{\theta_2(1+n\theta_1)^{\frac{1}{n}}}{2}[\sum_{i=1}^{k}(y_i-X_i\theta_3)^T(I-\frac{1}{n}J)(y_i-X_i\theta_3)]\} \\ \equiv\ & \pi'(\theta_1, \theta_2, \theta_3 \mid y). \end{aligned} \tag{12}$$

Integrating with respect to θ_3 in (12), then

$$\pi'(\theta_1, \theta_2 \mid y) \propto \theta_2^{\frac{N-2b-p}{2}} (1+n\theta_1)^{-a-\frac{p}{2n}} \exp\{-\frac{\theta_2(1+n\theta_1)^{\frac{1}{n}}}{2} \times [\sum_{i=1}^{k}(y_i - X_i\widehat{\theta}_3)^T(I-\frac{1}{n}J)(y_i - X_i\widehat{\theta}_3)]\}, \quad (13)$$

where $\widehat{\theta}_3 = [\sum_{i=1}^{k} X_i^T(I-\frac{1}{n}J)X_i]^{-1}[\sum_{i=1}^{k} X_i^T(I-\frac{1}{n}J)y_i]$. If $N-2b-p+2>0$, and integrating with respect to θ_2 in (13), then

$$\pi'(\theta_1 \mid y) \propto (1+n\theta_1)^{-\frac{N+2na-2b+2}{2n}}. \quad (14)$$

Therefore the marginal posterior (14) is proper if $N+2na-2b+2-2n>0$. This completes the proof.

Remark 3. Using the similar method in Theorem 1, we can easily prove that the posterior distribution under Jeffreys' prior π_1 is proper.

Theorem 2. Under the general prior (10), the marginal posterior density of θ_1 is given by

$$\pi(\theta_1 \mid y) \propto |\sum_{i=1}^{k} X_i^T(I-\frac{\theta_1}{1+n\theta_1}J)X_i|^{-\frac{1}{2}}(1+n\theta_1)^{-\frac{N+2na-2b+2}{2n}} \times [\sum_{i=1}^{k}(y_i - X_i\widehat{\theta}_3(\theta_1))^T(I-\frac{\theta_1}{1+n\theta_1}J)(y_i - X_i\widehat{\theta}_3(\theta_1))]^{-\frac{N-2b-p+2}{2}}$$

and under Jeffrey's prior π_1, the marginal posterior density of θ_1 is given by

$$\pi(\theta_1 \mid y) \propto (1+n\theta_1)^{-\frac{k+2}{2}} \times [\sum_{i=1}^{k}(y_i - X_i\widehat{\theta}_3(\theta_1))^T(I-\frac{\theta_1}{1+n\theta_1}J)(y_i - X_i\widehat{\theta}_3(\theta_1))]^{-\frac{N}{2}},$$

where $\widehat{\theta}_3(\theta_1) = [\sum_{i=1}^{k} X_i^T(I-\frac{\theta_1}{1+n\theta_1}J)X_i]^{-1}[\sum_{i=1}^{k} X_i^T(I-\frac{\theta_1}{1+n\theta_1}J)y_i]$.

The normalizing constant for the marginal density of θ_1 requires an one dimensional integration. Therefore we have the marginal posterior density of θ_1, and so it is to compute the marginal moment of θ_1. In Section 4, we investigate the frequentist coverage probabilities for the reference priors π_2, π_3 and Jeffreys' prior π_1, respectively.

4. Numerical Studies and Discussion

We evaluate the frequentist coverage probability by investigating the credible interval of the marginal posteriors density of θ_1 under the noninformative prior π given in Section 3 for several configurations $(\beta, \sigma^2_\alpha, \sigma^2), k, n$ and $X_i, i=1,\cdots,k$. That is to say, the frequentist coverage of a $(1-\alpha)$th posterior quantile should be close to $(1-\alpha)$. This is done numerically. Table 1 gives numerical values of the frequentist coverage probabilities of 0.05 (0.95) posterior quantiles for the our prior. The computation of these numerical values is based on the following algorithm for any fixed true $(\beta, \sigma^2_\alpha, \sigma^2)$, the design matrix X_i and any prespecified probability value α. Here α is 0.05 (0.95). Let $\theta_1^\pi(\alpha \mid Y)$ be the posterior α-quantile of θ_1 given Y. That is to say, $F(\theta_1^\pi(\alpha \mid Y) \mid Y)=\alpha$, where $F(\cdot \mid Y)$ is the marginal posterior distribution of θ_1. Then the frequentist coverage probability of this one sided credible interval of θ_1 is

$$P_{(\beta,\sigma^2_\alpha,\sigma^2)}(\alpha;\theta_1)=P_{(\beta,\sigma^2_\alpha,\sigma^2)}(0<\theta_1\le\theta_1^\pi(\alpha \mid Y)).$$

The estimated $P_{(\beta,\sigma^2_\alpha,\sigma^2)}(\alpha;\theta_1)$ when $\alpha=0.05$ (0.95) is shown in Table 1 for the trivariate case.

In particular, for fixed $(\beta, \sigma^2_\alpha, \sigma^2)$ and X_i, we take 10,000 independent random samples of Y from the model (1). Our simulation, we take $\beta=(1,1,1)^T$ without loss of generality. In the results of Table 1, the design matrixes are given by

$$X_1=X_2=X_3=\begin{pmatrix}1&0&0\\1&1&0\\1&1&1\end{pmatrix} \text{ and } X_4=X_5=\begin{pmatrix}1&0&1\\0&1&0\\1&1&1\end{pmatrix}.$$

For the cases presented in Table 1, we see that reference prior π_3 satisfying the second order matching criterion meet very well the target coverage probabilities. But although the two group reference prior π_2 satisfies the first order matching criterion, the prior has not good coverage probabilities. Thus we recommend to use the one-at-a-time reference prior π_3.

<Table 1> Frequentist Coverage Probabilities of 0.05 (0.95) Posterior Quantiles for θ_1

σ^2	σ_α^2	π_1	π_2	π_3
1	0.5	0.080(1.000)	0.073(1.000)	0.052(1.000)
	1	0.089(0.995)	0.074(0.995)	0.052(1.000)
	10	0.123(0.951)	0.077(0.931)	0.048(0.950)
	20	0.131(0.955)	0.080(0.929)	0.048(0.945)
	50	0.149(0.965)	0.086(0.934)	0.051(0.947)

In the linear mixed model, we have found a prior which is a second order matching prior and reference prior for the variance ratio. It turns out that the one-at-a-time reference prior satisfies the second order matching criterion. As illustrated in our numerical study, the one-at-a-time reference prior seems to be the best appropriate results than the other reference priors in the sense of asymptotic frequentist coverage property.

References

1. Berger, J.O. and Bernardo, J.M. (1989). Estimating a Product of Means : Bayesian Analysis with Reference Priors. *Journal of the American Statistical Association*, 84, 200-207.
2. Berger, J.O. and Bernardo, J.M. (1992). On the Development of Reference Priors (with discussion). *Bayesian Statistics IV*, J.M. Bernardo, et. al., Oxford University Press, Oxford, 35-60.
3. Bernardo, J.M. (1979). Reference Posterior Distributions for Bayesian Inference (with discussion). *Journal of Royal Statistical Society*, B, 41, 113-147.
4. Box, G. and Tiao, G. (1973). *Bayesian Inference in Statistical Analysis*. Addison-Wesly, Readings, Massachusetts.
5. Cox, D.R. and Reid, N. (1987). Orthogonal Parameters and Approximate Conditional Inference (with discussion). *Journal of Royal Statistical Society*, B, 49, 1-39.
6. Datta, G.S. and Ghosh, J.K. (1995a). On Priors Providing Frequentist Validity for Bayesian Inference. *Biometrika*, 82, 37-45.
7. Datta, G.S. and Ghosh, M. (1995b). Some Remarks on Noninformative Priors. *Journal of the American Statistical Association*, 90, 1357-1363.
8. Datta, G.S. and Ghosh, M. (1996). On the Invariance of Noninformative Priors. *The Annal of Statistics*, 24, 141-159.
9. DiCiccio, T.J. and Stern, S.E. (1994). Frequentist and Bayesian Bartlett

Correction of Test Statistics based on Adjusted Profile Likelihood. *Journal of Royal Statistical Society*, B, 56, 397-408.
10. Ghosh, J.K. and Mukerjee, R. (1992). Noninformative Priors (with discussion). *Bayesian Statistics IV*, J.M. Bernardo, et. al., Oxford University Press, Oxford, 195-210.
11. Graybill, F.A., Martin, F. and Godfrey, G. (1956). Confidence Intervals for Variance Ratios Specifying Genetic Heritability. *Biometrics*, 12, 99-109.
12. Hill, B. (1965). Inference about Variance Components in the One-Way Model. *Journal of the American Statistical Association*, 60, 806-825.
13. Kim, D.H., Kang, S.G. and Lee, W.D. (2001). On the Second Order Probability Matching Criterion in the One-Way Random Effect Model. *The Korean Communications in Statistics*, 8, 29-37.
14. Kang, S.G. Kim, D.H. and Lee, W.D. (2005). Bayesian Analysis for the Difference of Exponential Means. *Journal of Korean Data & Information Science Society*, 16, 1067-1078.
15. Mukerjee, R. and Dey, D.K. (1993). Frequentist validity of Posterior Quantiles in the Presence of a Nuisance Parameter: Higher Order Asymptotics. *Biometrika*, 80, 499-505.
16. Mukerjee, R. and Ghosh, M. (1997). Second Order Probability Matching Priors. *Biometrika*, 84, 970-975.
17. Palmer, J.L. and Broemeling, L.D. (1990). A Comparison of Bayes and Maximum Likelihood Estimation of the Intraclass Correlation Coefficient. *Communications in Statistics - Theory and Methods*, 1, 953-975.
18. Stein, C. (1985). On the Coverage Probability of Confidence Sets based on a Prior Distribution. *Sequential Methods in Statistics*, Banach Center Publications, 16, 485-514.
19. Tibshirani, R. (1989). Noninformative Priors for One Parameter of Many. *Biometrika*, 76, 604-608.
20. Welch, B.L. and Peers, H.W. (1963). On Formulae for Confidence Points based on Integrals of Weighted Likelihood. *Journal of Royal Statistical Society*, B, 25, 318-329.
21. Ye, K. (1994). Bayesian Reference Prior Analysis on the Ratio of Variances for the Balanced One-Way Random Effect Model. *Journal of Statistical Planning and Inference*, 41, 267-280.

[received date : Mar. 2006, accepted date : May. 2006]