On Second Order Probability Matching Criterion in the One-Way Random Effect Model

Dal Ho Kim¹⁾, Sang Gil Kang²⁾ and Woo Dong Lee³⁾

Abstract

In this paper, we consider the second order probability matching criterion for the ratio of the variance components under the one-way random effect model. It turns out that among all of the reference priors given in Ye(1994), the only one reference prior satisfies the second order matching criterion. Similar results are also obtained for the intraclass correlation as well.

Keywords: Intraclass Correlation; Matching Priors; Reference Priors; Variance Components.

1. Introduction

Consider the following one-way random effect model

$$Y_{ii} = \mu + \alpha_i + \varepsilon_{ii}, \quad i = 1, \dots, I; j = 1, \dots, J$$

$$(1.1)$$

where μ is an unknown constant, and the α_i and ϵ_{ij} are independent normal variables with 0 means and variances σ^2_{α} and σ^2 , respectively. Let $\phi = J\sigma^2_{\alpha}/\sigma^2$ be our parameter of interest.

The present paper focuses on noninformative priors for ϕ . We consider Bayesian priors such that the resulting credible intervals have coverage probabilities equivalent to their frequentist counterparts. Although this matching 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), Datta and Ghosh(1995a,b, 1996), Mukerjee and Ghosh(1997).

Assistant Professor, Kyungpook National University, Taegu 702-701, Korea. E-mail: dalkim@bh.kyungpook.ac.kr

²⁾ Lecturer, Department of Statistics, Kyungpook National University, Taegu 702-701, Korea.

³⁾ Assistant Professor, Faculty of Information Science, Kyungsan University, Kyungsan 712-240, Korea. E-mail: wdlee@kyungsan.ac.kr

Mukerjee and Dey(1993), Datta and Ghosh(1995a,b, 1996), Mukerjee and Ghosh(1997).

On the other hand, 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. Quite often reference priors satisfy the matching criterion described earlier.

The ratio of variance components in the random effect model has been of interest for a long time, especially in animal sciences, where 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 model (1.1) from the sampling theory point of view is the possible negative estimates for σ^2_{α} as well as for ϕ . Thus a Bayesian analysis for this model is desirable, not only because of its intrinsic merit, but also because it can resolve this problem. Also the commenly used intraclass correlation $\rho = \sigma^2_{\alpha}/(\sigma^2_{\alpha} + \sigma^2)$ is a one-to-one transformation of ϕ . Therefore, when a good prior distribution for ϕ is obtained, its transformed prior distribution for ρ can be used as well. Moreover, the probability matching priors are invariant for the transformation between ρ and ϕ . (cf. Datta and Ghosh, 1996).

The problem of estimating variance components in the one-way random effect model has been investigated by many authors from a Bayesian point of view. We may refer to Hill(1965), Box and Tiao(1973), and Palmer and Broemeling(1990), among others. Recently, Ye(1994) developed reference priors for ϕ , examined frequentist coverage probabilities of posterior quantiles for various ϕ and compared risk functions of the Bayes estimators for reference priors. Also Chung and Dey(1998) derived reference priors and first order probability matching priors for ρ and examined the frequentist coverage probabilities of posterior quantiles for various ρ .

The outline of the remaining sections is as follows. In Section 2, we derive Fisher information matrix under the reparametrization. Then we provide a class of second order probability matching priors of which particularly simple one is recommended. In Section 3, we investigate which one satisfies the second order matching criterion among reference priors given in Ye(1994) and Chung and Dey(1998) respectively.

2. Main Results

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

$$P^{\pi}[\theta_1 \le \theta_1^{1-\alpha}(\pi; Y) | Y] = 1-\alpha,$$
 (2.1)

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)^T$ and θ_1 is the parameter of interest and θ_2 and θ_3 are nuisance parameters. We want to find priors π for which

$$P[\theta_1 \le \theta_1^{1-\alpha}(\pi; Y) | \theta] = 1 - \alpha + o(n^{-u}). \tag{2.2}$$

for some u > 0, as n goes to infinity. Priors π satisfying (2.2) are called probability 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 , it is convenient to introduce orthogonal parametrization (Cox and Reid, 1987; Tibshirani, 1989). To this end, let

$$w_1 = J \ \sigma_a^2 / \sigma^2, \quad w_2 = \sigma^2 \left(\frac{\sigma^2 + J \sigma_a^2}{\sigma^2} \right)^{1/J}, \quad w_3 = \mu.$$
 (2.3)

With this parametrization, the likelihood function of parameters (w_1, w_2, w_3) for the model (1.1) is given by

$$L(w_1, w_2, w_3) \propto w_2^{-IJ/2} \exp\left\{-\frac{1}{2(1+w_1)^{-1/J}w_2} \left[S_2 + \frac{S_1 + IJ(\overline{Y} - w_3)^2}{1+w_1}\right]\right\}. \quad (2.4)$$

Based on (2.4), the Fisher information matrix is given by

$$I = \begin{pmatrix} \frac{I(J-1)(1+w_1)^{-2}}{2J} & 0 & 0\\ 0 & \frac{IJ}{2w_2^2} & 0\\ 0 & 0 & \frac{IJ(1+w_1)^{-1/J-1}}{w_2} \end{pmatrix}.$$

Thus w_1 is orthogonal to w_2 and w_3 in the sense of Cox and Reid(1987). This orthogonality was suggested by Solomon and Taylor(1999). Following Tibshirani(1989), the class of first order probability matching prior is characterized by

$$\pi_m^{(1)}(w_1, w_2, w_3) \propto (1 + w_1)^{-1} d(w_2, w_3),$$
 (2.5)

where $d(w_2, w_3)$ is an arbitrary function differentiable in its arguments.

Clearly the class of prior given in (2.5) is quite large, and it is important to narrow down this class of priors. To this end, we derive the class of second order probability matching priors for one-way random effect models following Mukerjee and Ghosh(1997).

Theorem 1. The second order probability matching priors are given by

$$\pi_m^{(2)}(w_1, w_2, w_3) \propto (1 + w_1)^{-1} w_2^{-1} h(w_3),$$
 (2.6)

where $h(w_3)$ is any smooth function of w_3 .

Proof. A second order probability matching prior is of the form (2.5), and also d must satisfy an additional differential equation (cf (2.10)) of Mukerjee and Ghosh(1997), namely

$$\frac{1}{6} d(w_{2}, w_{3}) \frac{\partial}{\partial w_{1}} (I_{11}^{-3/2} L_{1,1,1}) + \frac{\partial}{\partial w_{2}} \{I_{11}^{-1/2} L_{112} I^{22} d(w_{2}, w_{3})\}
+ \frac{\partial}{\partial w_{3}} \{I_{11}^{-1/2} L_{113} I^{33} d(w_{2}, w_{3})\} = 0,$$
(2.7)

where

$$\begin{split} L_{1.1.1} &= E[(\frac{\partial \log L}{\partial w_1})^3] = \frac{I(J-1)(J-2)}{J^2} (1+w_1)^{-3}, \\ L_{112} &= E[\frac{\partial^3 \log L}{\partial w_1^2 \partial w_2}] = \frac{I(J-1)}{2J} (1+w_1)^{-2} w_2^{-1}, \\ L_{113} &= E[\frac{\partial^3 \log L}{\partial w_1^2 \partial w_3}] = 0, \end{split}$$

and

$$\begin{pmatrix} I^{11} & I^{12} & I^{13} \\ I^{21} & I^{22} & I^{23} \\ I^{31} & I^{32} & I^{33} \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{2J}{I(J-1)(1+w_1)^2} & 0 & 0 \\ 0 & \frac{2w_2^2}{IJ} & 0 \\ 0 & 0 & \frac{w_2}{IJ(1+w_1)^{1/J-1}} \end{pmatrix}.$$

Then (2.7) simplifies to

$$\left[\frac{I(J-1)}{2J}\right]^{1/2} \frac{2}{IJ} (1+w_1)^{-1} \frac{\partial}{\partial w_2} \left\{ w_2 d(w_2, w_3) \right\} = 0. \tag{2.8}$$

Hence the set of solution of (2.8) is of the form

$$d(w_2, w_3) = w_2^{-1}h(w_3),$$

where $h(w_3)$ is any smooth function of w_3 . Thus the resulting second order probability matching prior is given by

$$\pi_m^{(2)}(w_1, w_2, w_3) \propto (1 + w_1)^{-1} w_2^{-1} h(w_3).$$

This completes the proof.

Also we are interest in the intraclass correlation, $\rho = \sigma_a^2/(\sigma_a^2 + \sigma^2)$. Especially, $\rho = w_1/(w_1 + f)$ is a one-to-one function of w_1 . Thus by using invariance property of probability matching priors under one-to-one transformation of the parameter vector (Datta and Ghosh, 1996; Mukerjee and Ghosh, 1997), we obtain the second order probability matching prior for $\pi_m^{(2)}(\rho, w_2, w_3)$ from the matching priors (2.6) as follows.

$$\pi_m^{(2)}(\rho, w_2, w_3) \propto [1 + (J-1)\rho]^{-1}(1-\rho)^{-1}w_2^{-1}h(w_3).$$
 (2.9)

We consider a particular second order matching prior where h is a constant in (2.6). This choice is very natural since w_3 is the location parameter. Thus we have

$$\pi_m^{(2)}(w_1, w_2, w_3) \propto (1 + w_1)^{-1} w_2^{-1}.$$
 (2.10)

Again by using invariance of probability matching priors under one-to-one transformation of the parameter vector, we obtain the second order probability matching prior $\pi_m^{(2)}(\phi, \sigma^2, \mu)$ and $\pi_m^{(2)}(\rho, \sigma^2, \mu)$ from the matching prior (2.10), respectively. The results are as follows.

Corollary 1. For the random effect model (1.1), (i) if $\phi = J\sigma_a^2/\sigma^2$ is the parameter of interest, then the second order probability matching prior for (ϕ, σ^2, μ) is given by

$$\pi_m^{(2)}(\phi, \sigma^2, \mu) \propto \sigma^{-2}(1+\phi)^{-1}$$
. (2.11)

(ii) If $\rho = \sigma_a^2/(\sigma_\alpha^2 + \sigma^2)$ is the parameter of interest, then the second order probability matching prior for (ρ, σ^2, μ) is given by

$$\pi_m^{(2)}(\rho, \sigma^2, \mu) \propto \sigma^{-2}[1 + (J-1)\rho]^{-1}(1-\rho)^{-1}.$$
 (2.12)

Using the results of Corollary 1, we can examine which one satisfies the second order matching criterion among reference priors given in Ye(1994) and Chung and Dey(1998) respectively.

3. Simulation Study and Discussion

Ye(1994) derived four reference priors for different groups of ordering of (ϕ, μ, σ^2) under the model (1.1). In specific, if $\phi = J\sigma_a^2/\sigma^2$ is the parameter of interest, then the reference prior distributions for different groups of ordering of (μ, σ^2, ϕ) are as follows.

Based on both the asymptotic frequentist coverage property and the decision theory points of view, Ye(1994) revealed that π_3^Y , the one-at-a-time grouping reference prior, is the best among all of these reference priors when ϕ is the parameter of interest. Table 1 (Ye, 1994) shows that π_3^Y seems to be the best in the sense of the frequentist coverage probabilities of 0.05(0.95) posterior quantiles for ϕ . Those numerical findings are consistent with our

theoretical results. That is, from (i) of the Corollary 1 in Section 2, only π_3^Y satisfies a second order probability matching criterion.

Table 1 : Frequentist Coverage Probabilities of 0.05 (0.95) Posterior Quantiles for ϕ

ϕ	(I,J)	π_1^Y	π_2^Y	π_3^Y	π_4^Y
1	(6,5)	0.03(1.00)	0.03(1.00)	0.05(1.00)	0.06(1.00)
	(7,5)	0.03(1.00)	0.03(1.00)	0.05(1.00)	0.06(1.00)
	(5,10)	0.03(1.00)	0.03(1.00)	0.05(1.00)	0.06(1.00)
	(10,5)	0.03(1.00)	0.03(1.00)	0.05(1.00)	0.06(1.00)
	(10,10)	0.03(1.00)	0.03(1.00)	0.05(1.00)	0.05(1.00)
	(5,20)	0.02(1.00)	0.02(1.00)	0.05(1.00)	0.05(1.00)
	(20,5)	0.03(0.97)	0.04(0.97)	0.05(0.98)	0.05(0.98)
5	(6,5)	0.02(0.91)	0.03(0.93)	0.05(0.97)	0.06(0.97)
	(7,5)	0.02(0.91)	0.03(0.92)	0.05(0.96)	0.06(0.96)
	(5,10)	0.02(0.92)	0.03(0.93)	0.05(0.99)	0.05(0.99)
	(10,5)	0.03(0.91)	0.03(0.92)	0.05(0.95)	0.06(0.95)
	(10,10)	0.03(0.91)	0.03(0.92)	0.05(0.95)	0.05(0.95)
	(5,20)	0.02(0.92)	0.03(0.92)	0.05(0.99)	0.05(0.99)
	(20,5)	0.03(0.93)	0.03(0.93)	0.05(0.95)	0.05(0.95)
10	(6,5)	0.02(0.89)	0.03(0.90)	0.05(0.95)	0.06(0.96)
	(7,5)	0.03(0.91)	0.03(0.91)	0.05(0.95)	0.06(0.96)
	(5,10)	0.03(0.89)	0.03(0.89)	0.05(0.96)	0.05(0.96)
	(10,5)	0.03(0.91)	0.03(0.92)	0.05(0.95)	0.06(0.95)
	(10,10)	0.03(0.91)	0.03(0.91)	0.05(0.95)	0.06(0.95)
	(5,20)	0.03(0.89)	0.03(0.90)	0.05(0.96)	0.05(0.96)
	(20,5)	0.03(0.93)	0.03(0.93)	0.05(0.95)	0.05(0.95)
50	(6,5)	0.02(0.90)	0.03(0.90)	0.05(0.95)	0.06(0.95)
	(7,5)	0.03(0.90)	0.03(0.90)	0.05(0.95)	0.05(0.95)
	(5,10)	0.02(0.88)	0.03(0.89)	0.05(0.95)	0.06(0.95)
	(10,5)	0.03(0.91)	0.04(0.92)	0.05(0.95)	0.06(0.95)
	(10,10)	0.03(0.92)	0.03(0.92)	0.05(0.95)	0.05(0.95)
	(5,20)	0.02(0.89)	0.03(0.89)	0.05(0.95)	0.05(0.95)
	(20,5)	0.03(0.93)	0.03(0.93)	0.05(0.95)	0.05(0.95)

Also when ρ is the parameter of interest, Chung and Dey(1998) derived four reference priors for different groups of ordering for (ρ, μ, σ^2) and the first order probability matching priors for ρ under the model (1.1). Specifically, if $\rho = \sigma_a^2/(\sigma_a^2 + \sigma^2)$ is the parameter of interest, then the reference prior distributions for different groups of ordering of (μ, σ^2, ρ) are as follows.

Group ordering

Reference prior

$$\{\rho, \mu, \sigma^2\}, \{\rho, \sigma^2, \mu\}, \{(\rho, \sigma^2), \mu\} \qquad \pi_1^C \propto \sigma^{-2} (1 + (J-1)\rho)^{-1} (1-\rho)^{-1} \\ \{\rho, (\mu, \sigma^2)\} \qquad \qquad \pi_2^C \propto \sigma^{-3} (1 + (J-1)\rho)^{-1} (1-\rho)^{-1} \\ \{(\rho, \mu), \sigma^2\} \qquad \qquad \pi_3^C \propto \sigma^{-2} (1 + (J-1)\rho)^{-3/2} (1-\rho)^{-1/2} \\ \{(\rho, \sigma^2, \mu)\} \qquad \qquad \pi_4^C \propto \sigma^{-3} (1 + (J-1)\rho)^{-3/2} (1-\rho)^{-1/2}$$

Also the first order matching priors are of the form

$$\pi_m^{(1)}(\rho, \mu, \sigma^2) \propto (k_1 (\ln \sigma^2 + \rho) + k_2) \sigma^{-2} (1 + (J-1)\rho)^{-1} (1-\rho)^{-1}$$
,

where k_1 and k_2 are arbitrary constants.

Tale 2 : Frequentist Coverage Probabilities of 0.05(0.95) Posterior Quantiles

ρ	(I,J)	π_1^Y	π_2^Y	π_3^Y	π_4^Y
3/4	(6,5)	0.05(0.95)	0.06(0.95)	0.02(0.89)	0.03(0.90)
1	(7,5)	0.05(0.95)	0.06(0.95)	0.03(0.90)	0.03(0.91)
	(5,10)	0.05(0.95)	0.05(0.95)	0.02(0.89)	0.03(0.89)
2/3	(6,5)	0.05(0.95)	0.06(0.96)	0.03(0.89)	0.03(0.90)
	(7,5)	0.05(0.95)	0.06(0.96)	0.03(0.90)	0.03(0.91)
	(5,10)	0.05(0.95)	0.05(0.95)	0.02(0.88)	0.03(0.89)
1/2	(6,5)	0.05(0.97)	0.06(0.97)	0.02(0.92)	0.03(0.92)
	(7,5)	0.05(0.97)	0.06(0.97)	0.03(0.91)	0.03(0.92)
	(5,10)	0.05(0.96)	0.05(0.96)	0.02(0.89)	0.03(0.89)
1/3	(6,5)	0.05(1.00)	0.06(1.00)	0.03(0.98)	0.03(0.98)
	(7,5)	0.05(0.99)	0.06(0.99)	0.03(0.96)	0.03(0.97)
	(5,10)	0.05(0.98)	0.05(0.99)	0.02(0.92)	0.02(0.92)
1/5	(6,5)	0.05(1.00)	0.06(1.00)	0.03(1.00)	0.04(1.00)
	(7,5)	0.05(1.00)	0.06(1.00)	0.02(1.00)	0.03(1.00)
	(5,10)	0.05(1.00)	0.05(1.00)	0.03(1.00)	0.03(1.00)

Chung and Dey(1998) revealed that the π_1^C and π_2^C are better than π_3^C and π_4^C in the sense of the asymptotic frequentist coverage probability, where only π_1^C is the first order matching prior. But this conclusion for reference priors for ρ somewhat differs with the results of Ye(1994). Since the $\rho = \phi/(\phi + f)$ is a one-to-one function of ϕ , the results for ϕ in Ye(1994) are expected to be consistent with the results for ρ in Chung and Dey(1998). Our theoretical results in Section 2 show that the only reference prior π_1^C satisfies the second order matching criterion from the (ii) of Corollary 1. Thus we reinvestigate the credible interval for ρ under four reference priors for comparison using the design situation of Chung and Dey(1998). 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 2 provides frequentist coverage probabilities of 0.05(0.95) posterior quantiles for ρ under four reference priors. The computation of these numerical values is based on the following algorithm for any fixed true ρ and any prespecified probability value α . Here α is 0.05(0.95). Let $\rho^{\pi}(\alpha \mid Y)$ be the posterior α -quantile of ρ given Y. That is to say, $F(\rho^{\pi}(\alpha \mid Y) \mid Y) = \alpha$, where $F(\cdot \mid Y)$ is the marginal posterior distribution of ρ . Then the frequentist coverage probability of this one sided credible interval of ρ is

$$P_{\rho}(\alpha;\rho) = P_{\rho}(0 \langle \rho \leq \rho^{\pi}(\alpha | Y)). \tag{3.1}$$

The estimated $P_{\rho}(\alpha; \rho)$ when $\alpha = 0.05(0.95)$ is shown in Table 2. Actually Table 2 was computed in the following way. For fixed ρ , we take 10,000 independent random samples of \mathbf{Y} from the model (1.1). Note that under the prior π , for fixed \mathbf{Y} , $\rho \leq \rho^{\pi}(\alpha \mid \mathbf{Y})$ if and only if $F(\rho^{\pi}(\alpha \mid \mathbf{Y}) \mid \mathbf{Y}) \leq \alpha$. Under the prior π , $P_{\rho}(\alpha; \rho)$ can be estimated by the relative frequency of $F(\rho^{\pi} \mid \mathbf{Y}) \leq \alpha$. From the Table 2, π_1^C is the most appealing reference prior distribution in the sense of the asymptotic frequentist coverage probability.

Consequently the second order matching criterion seems to give more appropriate choice among several reference priors for ϕ and ρ in the asymptotic frequentist coverage probability points of view.

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] Chung, Y. and Dey D.K.(1998). Bayesian approach to estimation of intraclass correlation using reference prior, Communications in Statistics - Theory and Methods, 27, 2241-2255.
- [6] Cox, D.R. and Reid, N.(1987). Orthogonal parameters and approximate conditional inference (with discussion), *Journal of Royal Statistical Society*, B, 49, 1-39.
- [7] Datta, G.S. and Ghosh, J.K.(1995a). On priors providing frequentist validity for Bayesian inference, *Biometrika*, 82, 37-45.
- [8] Datta, G.S. and Ghosh, M.(1995b). Some remarks on noninformative priors, Journal of the

- American Statistical Association, 90, 1357-1363.
- [9] Datta, G.S. and Ghosh, M.(1996). On the invariance of noninformative priors, *The Annal of Statistics*, 24, 141-159.
- [10] Graybill, F.A., Martin, F. and Godfrey, G.(1956). Confidence intervals for variance ratios specifying genetic heritability, *Biometrics*, 12, 99-109.
- [11] Hill, B.(1965). Inference about variance components in the one-way model, *Journal of the American Statistical Association*, 60, 806-825.
- [12] 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.
- [13] Mukerjee, R. and Ghosh, M.(1997). Second order probability matching priors, *Biometrika*, 84, 970-975.
- [14] Palmer, J.L. and Broemeling, L.D.(1990). A comparision of Bayes and maximum likelihood estimation of the intraclass correlation coefficient, *Communications in Statistics Theory and Methods*, 19, 953–975.
- [15] Stein, C.(1985). On the coverage probability of confidence sets based on a prior distribution, *Sequential Methods in Statistics*, 16, 485-514.
- [16] Solomon, P.J. and Taylor, J.M.G.(1999). Orthogonality and transformations in variance components models, *Biometrika*, 86, 289-300.
- [17] Tibshirani, R.(1989). Noninformative Priors for One Parameter of Many, *Biometrika*, 76, 604-608.
- [18] 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.
- [19] Welch, B.N. and Peers, B.(1963). On formulae for confidence points based on integrals of weighted likelihood, *Journal of Royal Statistical Society*, 35, 318–3298.