Noninformative Priors for the Intraclass Coefficient of a Symmetric Normal Distribution

In Hong Chang[1] and Byung Hwee Kim[2]

ABSTRACT

In this paper, we develop the Jeffreys prior, reference priors and the probability matching priors for the intraclass correlation coefficient of a symmetric normal distribution. We next verify propriety of posterior distributions under those noninformative priors. We examine whether reference priors satisfy the probability matching criterion.

1. INTRODUCTION

A multivariate normal observation $X=(X_1,\cdots,X_p)^T$ is said to have a Symmetric normal distribution if

$$EX=\begin{pmatrix}\mu\\ \vdots\\ \mu\end{pmatrix}=\mu\mathbf{1} \quad \text{and} \quad CovX=\sigma^2\begin{pmatrix}1 & \rho & \cdots & \rho\\ \rho & 1 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \rho\\ \rho & \cdots & \rho & 1\end{pmatrix}, \tag{1.1}$$

where $\mu\in(-\infty,\infty)$ and $-(p-1)^{-1}<\rho<1$. The above distribution is extensively studied in Rao(1973) and can be used for many real-life applications. One such example is discussed in Rao(1973), where a symmetric model distribution can be used to estimate the correlation coefficient between heights of brothers on the basis of measurement taken on ρ brothers in each of the n families.

The intraclass correlation coefficient ρ is frequently used to measure the degree of intrafamily resemblance with respect to characteristics such as blood pressure, cholesterol, weight, height, stature, lung capacity, and so forth. Statistical inference concerning ρ for a single-sample problem based on a normal distribution has been studied by several authors. Among others, we may refer to Rao(1973), Rosner, Donner, and Hennekens(1977), Donner and Bull(1983), Srivastava and Carter(1983), Srivastava and Katapa(1986), Gokhale and SenGupta(1986), Velu and Rao(1990).

There is a considerable study of a statistical inference for intraclass correlation coefficient from familial data by several authors. In this paper, we consider estimating problem for intraclass correlation coefficient ρ using noninformative priors that applying in Bayesian estimation.

The most frequently used noninformative prior is Jeffreys' prior, which is proportional to

1) Full-time Instructor, Department of Computer Science and Statistics, Chosun University, 501-759, Gwangju, Korea
2) Professor, Department of Mathematics, Hanyang University, 133-791, Seoul, Korea

the positive square root of the determinant of the Fisher information matrix. But, in spite of its success in one-parameter problems, Jeffreys' prior frequently runs into serious difficulties in the presence of nuisance parameters. To overcome these deficiencies of Jeffreys' prior, Berger and Bernardo(1989,1992) extended the reference prior approach of Bernardo(1979) for deriving noninformative priors in multiparameter situations by dividing the parameters into parameters of interest and nuisance parameters. This approach is very successful in various practical problems. As an alternative, we use the method of Peers(1965) to find priors which require the frequentist coverage probability of the posterior region of a real-valued parametric function to match the nominal level with a remainder of $o(n^{-1/2})$. These priors, as usually referred to as the first order matching priors. Mukerjee and Dey (1993) extended this result with a remainder of $o(n^{-1})$ for which the priors are called the second order matching priors.

In this paper, we consider the problem of inferencing ρ using noninformative priors. In Section 2 we derives the reference priors. In Section 3, we show that posterior distributions are proper and provide marginal posterior distributions under reference priors. In Section 4, we examine whether reference priors satisfy the probability matching criterion.

2. NONINFORMATIVE PRIORS

Suppose we have a sample of measurements from k families, and let $X_1, \cdots, X_k$ represent measurements from the i-th family, where $X_i = (X_{i1}, \cdots, X_{ip})'$ is a $p \times 1$ vector of observations from the i-th family, $i = 1, 2, \cdots, k$. The structure of the mean vector and the covariance matrix for the familial data is given by (Rao, 1973) as

$$\mu = \mu\, \mathbf{1}_p \quad \text{and} \quad \underset{p\times p}{\Sigma} = \sigma^2(1-\rho)\, I_p + \rho\, J_p \tag{2.1}$$

where $\mathbf{1}_p$ is the $p \times 1$ vector of 1's, I_p is the $p \times p$ identity matrix and J_p is the $p \times p$ matrix containing only ones, $\mu(-\infty < \mu < \infty)$ is the common mean, $\sigma^2(\sigma^2 > 0)$ is the common variance of members of family, and ρ, called the intraclass correlation coefficient, is the coefficient of correlation among the members of family and $-1/(p-1) < \rho < 1$.

It is assumed that $X_i \sim N(\mu, \Sigma)$, $i = 1, 2, \cdots k$, where represents a p-variate normal distribution and μ and Σ are defined (2.1).

Let

$$Y_i = (Y_{i1}, \cdots, Y_{ip})' = QX_i, \tag{2.2}$$

where Q is an orthogonal matrix. Under the orthogonal transformation (2.2), it is obvious that

$$Y_i \sim N(\mu^*, \Sigma^*),\ i = 1, 2, \cdots, k, \tag{2.3}$$

where

$$\mu^* = (\sqrt{p}\mu, 0, 0, \cdots, 0)' \quad \text{and} \quad \Sigma^* = \sigma^2 diag\{1 + (p-1)\rho, (1-\rho), \cdots, (1-\rho)\}.$$

The likelihood function of (ρ, σ^2, μ) for model (2.3) is

$$l(\rho, \sigma^2, \mu | \mathbf{y}) = (2\pi)^{-\frac{pk}{2}} (\sigma^2)^{-\frac{pk}{2}} [1+(p-\rho)]^{-\frac{k}{2}} (1-\rho)^{-\frac{(p-1)k}{2}} \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{k} [\frac{(y_{i1}-\sqrt{p}\mu)^2}{1+(p-1)\rho} + \frac{1}{1-\rho} \sum_{j=2}^{p} y_{ij}^2]} \quad (2.4)$$

From (2.4), Fisher information matrix is given by

$$I(\rho, \sigma^2, \mu) = \begin{pmatrix} \frac{(p-1)^2 k}{2(1+(1-p)\rho)^2} + \frac{(p-1)k}{2(1-\rho)^2} & -\frac{p(p-1)\rho k}{2\sigma^2} \frac{1}{(1+(p-1)\rho)} \frac{1}{1-\rho} & 0 \\ -\frac{p(p-1)\rho k}{2\sigma^2} \frac{1}{(1+(p-1)\rho)} \frac{1}{1-\rho} & \frac{pk}{2\sigma^4} & 0 \\ 0 & 0 & \frac{1}{\sigma^2} \frac{pk}{1+(p-1)\rho} \end{pmatrix} \quad (2.5)$$

We have the following theorem of the reference prior of the ordering group (ρ, σ^2, μ). using an algorithm by Berger and Bernard(1989) to (2.3).

Theorem 2.1 For model (2.3), if ρ is the parameter of interest, then the reference prior distributions for different groups of ordering of (ρ, σ^2, μ) are ;

Group ordering	Reference prior
$\{\rho, (\sigma^2, \mu)\}$	$\pi_1 \propto (\sigma^2)^{-\frac{3}{2}} [1+(p-1)\rho]^{-1} (1-\rho)^{-1}$
$\{\rho, \sigma^2, \mu\}$, $\{\rho, \mu, \sigma^2\}$, $\{(\rho, \sigma^2), \mu\}$	$\pi_2 \propto (\sigma^2)^{-1} [1+(p-1)\rho]^{-1} (1-\rho)^{-1}$
$\{(\rho, \mu), \sigma^2\}$	$\pi_3 \propto (\sigma^2)^{-1} [1+(p-1)\rho]^{-\frac{3}{2}} (1-\rho)^{-1}$
$\{(\rho, \sigma^2, \mu)\}$	$\pi_4 \propto (\sigma^2)^{-\frac{3}{2}} [1+(p-1)\rho]^{-\frac{3}{2}} (1-\rho)^{-1}$

All the reference prior distributions are proportional to a negative powers of ρ, $[1+(p-1)\rho]$ and $(1-\rho)$. Therefore , a general form of the prior can be written as

$$\pi \propto (\sigma^2)^{\alpha} [1+(p-1)\rho]^{\beta} (1-\rho)^{\gamma}. \quad (2.6)$$

Also, Jeffreys' prior which is the square root of the determinant of the expected Fisher information matrix is given by

$$\pi_J \propto (\sigma^2)^{-\frac{3}{2}} [1+(p-1)\rho]^{-\frac{3}{2}} (1-\rho)^{-1}, \quad (2.7)$$

Therefore, the Jeffreys' prior is the same as the reference prior π_4 for (ρ, σ^2, μ).

3. POSTERIOR DISTRIBUTIONS

According to Bayes theorem, the posterior distribution of (ρ, σ^2, μ) with respect to the priors in (2.6) is given by

$$\pi(\rho, \sigma^2, \mu | \mathbf{y}) = (2\pi)^{-\frac{pk}{2}} (\sigma^2)^{\alpha - \frac{pk}{2}} [1+(p-\rho)]^{\beta - \frac{k}{2}} (1-\rho)^{\gamma - \frac{(p-1)k}{2}} h(\mu) \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{k} [\frac{(y_{i1}-\sqrt{p}\mu)^2}{1+(p-1)\rho} + \frac{1}{1-\rho} \sum_{j=2}^{p} y_{ij}^2]}. \quad (3.1)$$

Theorem 3.1 The posterior distribution (3.1) is proper if $\alpha < \frac{pk-3}{2}$, $\beta - \alpha + \frac{(p-1)k}{2} > 0$, $\gamma - \alpha + \frac{k}{2} - \frac{1}{2} > 0$.

4. PROBABILITY MATCHING PRIORS

We have the first order matching priors for ρ, the parameter of interest, as follows:

Theorem 4.1. The first order probability matching priors are given by

$$\pi_M^{(1)}(\rho, \sigma^2, \mu) \propto [1+(p-1)\rho]^{\frac{1}{p}-1}(1-\rho)^{\frac{p-1}{p}-1} \cdot g(\sigma^2(1-\rho)^{1-\frac{1}{p}}[1+(p-1)\rho]^{\frac{1}{p}}, \mu), \tag{4.1}$$

where g is a positive arbitrary differentiable function.

Proof. the first order probability matching priors are the solutions π to the differential equation

$$\frac{\partial}{\partial\rho}\left\{\left[\frac{2[1+(p-1)\rho]^2(1-\rho)^2}{p(p-1)k}\right]^{\frac{1}{2}}\pi\right\} + \frac{\partial}{\partial\sigma^2}\left\{\left[\frac{2[1+(p-1)\rho]^2(1-\rho)^2}{p(p-1)k}\right]^{\frac{1}{2}} \cdot \left[\frac{2\sigma^2[1+(p-1)\rho](1-\rho)\rho}{pk}\right]\pi\right\} = 0. \tag{4.2}$$

The solutions to (4.2) are given by

$$\pi_M^{(1)}(\rho, \sigma^2, \mu) \propto [1+(p-1)\rho]^{\frac{1}{p}-1}(1-\rho)^{\frac{p-1}{p}-1}\ g(\sigma^2(1-\rho)^{1-\frac{1}{p}}[1+(p-1)\rho]^{\frac{1}{p}}, \mu),$$

for any positive differentiable function g.

Clearly the class of priors given (4.1) is quite large, and it is important to narrow down this class of priors. We have the second order matching priors for ρ, the parameter of interest, as follows:

Theorem 4.2. The second order probability matching priors are given by

$$\pi_M^{(2)}(\rho, \sigma^2, \mu) \propto (\sigma^2)^{-1}[1+(p-1)\rho]^{-1}(1-\rho)^{-1}\ h(\mu), \tag{4.3}$$

where $h(\mu)$ is a positive arbitrary differentiable function.

Remark 4.1. There are infinitely many matching priors for ρ up to $o(n^{-1})$. Among the reference priors developed in Section 2 π_2 is the only second order probability matching prior with $h(\mu)=1$.

REFERENCES

Berger, J. O. and Bernardo, J. M.(1989). "Estimating a product of means : Bayesian analysis with reference priors." *J. Amer. statist. Assoc*, ***84***, 200-207.

Berger, J. O. and Bernardo, J. M.(1992). On the development of reference priors (with discussion). Bayesian Statistics 4 (Bernardo, J. M. et al., eds.). Oxford Univ.

Press, *Oxford*, 35-60.

Bernardo, J. M(1979). "Reference posterior distributions for Bayesian inference ." *J. Roy Statist. Soc. (Ser. B)*, ***41***, 113-147.

Cox, D. R. and Reid, N.(1987). "Orthogonal parameters and approximate conditional inference (with discussion)." *J. Roy. Statist. Soc. (Ser. B)*, ***49***, *1*-39.

Datta, G. S. and Ghosh, M.(1996). "On the invariance of noninformative priors." *Ann. Statist.*, ***24***, 141-159.

Donner, A. and Bul, S.(1983). "Inferences concerning a common intraclass correlation coefficient." *Biometrics*, ***39***, 771-775.

Ghosh, J. K. and Mukerjee, R.(1993). "On priors that match posterior and frequentist distribution functions." *Canad. J. Statist.*, ***21***, 89-96.

Ghokale, D. V. and SenGupta, A.(1986). "Optimal tests for the correlation coefficient in a symmetric multivariate normal population." *Jounal of Statistical Planning and Inference*, ***14***, 263-268.

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.

Mukerjee, R. and Ghosh, M.(1997). "Second order probability matching priors." *Biometrika*, ***84***, 970-975.

Peers, H. W.(1965). On confidence sets and Bayesian probability points in the case of several parameters. *J. Royal Statist. Soc. (Ser B)*, ***27***, 9-16.

Rao, C. R.(1973). *"Linear statistical inference and Its applications."*, New York: Wiley

Rosner, B., Donner, A., and Honnekens, C. H.(1977). "Estimation of intraclass correlations from familial data." *Applied Statistics*, ***26***, 179-187.

Rosner, B., Donner, A., and Honnekens, C. H.(1979). "Significance testing of intraclass correlations from familial data." *Biometrics*, ***35***, 461-471.

Srivastava, M. S.(1984). "Estimation of intraclass correlation in familial data." *Biometrica*, ***71***, 177-185.

Srivastava, M. S. and Carter, E. M.(1983). *An introduction to a applied multivariate statistics.* North Holland. New Yok

Srivastava, M. S. and Katapa, R. S.(1986). "Comparison of estimators of interclass and intraclass correlations from familial data ." *Canadian Journal of Statistics*, ***14***, 29-42.

Velu, R. and Rao, M. B.(1990). "Estimation of parent-offspring correlation." *Biometrika*, ***77***, 557-562.

Young, D. J. and Bhandary, M.(1998). "Test for equality of intraclass correlation coefficients under unequal family sizes." *Biometrics*, ***54***, 1363-1373.