Noninformative Priors for the Common Intraclass Correlation Coefficient

Dal Ho Kim[1,a]

[a]Department of Statistics, Kyungpook National University

Abstract

In this paper, we develop the noninformative priors for the common intraclass correlation coefficient when independent samples drawn from multivariate normal populations. We derive the first and second order matching priors. We reveal that the second order matching prior dose not match alternative coverage probabilities up to the second order and is not a HPD matching prior. It turns out that among all of the reference priors, one-at-a-time reference prior satisfies a second order matching criterion. Our simulation study indicates that one-at-a-time reference prior performs better than the other reference priors in terms of matching the target coverage probabilities in a frequentist sense.

Keywords: Frequentist coverage, intraclass correlation coefficient, matching priors, multivariate normal, reference priors.

1. Introduction

The intraclass correlation coefficient ρ is used to measure the degree of intrafamily resemblance with respect to characteristics such as blood pressure, cholesterol, weight, height, stature and lung capacity. Statistical inference concerning ρ for a single sample problem based on a normal distribution has been studied by many authors. However, its extension to multisample problems based on several multivariate normal distributions has received a little attention.

For testing the equality of two intraclass correlation coefficients based on two independent mutinormal samples, Donner and Bull (1983) discussed the likelihood ratio test. Konishi and Gupta (1989) proposed an approximate likelihood ratio test, and derived its asymptotic null and nonnull distributions for testing the equality of more than two intraclass correlation coefficients based on independent samples from several multinormal distributions. Huang and Sinha (1993) proposed several approximately optimum invariant tests for testing the equality of k (two or more) intraclass correlation coefficients based on mutinormal samples. In some other way, Young and Bhandary (1998) proposed several tests including the likelihood ratio test for the equality of two intraclass correlation coefficients under unequal family sizes based on two independent mutinormal samples.

The present paper focuses on the development of noninformative priors for the common intraclass correlation coefficient ρ when independent samples are drawn from multivariate normal populations. Although subjective Bayesian are often critical of such priors, these priors have clear pragmatic appeal especially when prior information is vague in nature.

The 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), Datta and Ghosh (1995b, 1996), Mukerjee and Ghosh (1997) and Mukerjee and Reid (1999).

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.

The outline of the remaining sections is as follows. In Section 2, we derive Fisher information matrix under the reparametrization. Then we develop the first order and second order probability matching priors. It turns out that the second order matching prior dose not match alternative coverage probabilities up to the second order and is not a HPD matching prior. Moreover, among all of the reference priors, one-at-a-time reference prior satisfies a second order matching criterion. In Section 3, we prove the propriety of the posterior distribution for a general class of prior distributions which includes the reference priors as well as the matching priors. In Section 4, simulated frequentist coverage probabilities under the proposed priors are investigated.

2. The Noninformative Priors

Let $X_i^{(l)} = (X_{i1}^{(l)}, \ldots, X_{ip}^{(l)})'$ be $p \times 1$ vector of observations from the multivariate normal distribution with mean vector $\boldsymbol{\mu}^{(l)}$ and covariance matrix $\boldsymbol{\Sigma}^{(l)}$, where $X_i^{(l)}$ is the ith observation vector in the l-population, $l = 1, \ldots, q$, $i = 1, \ldots, n_l$. The structure of the mean vector and the covariance matrix is given as

$$\boldsymbol{\mu}^{(l)} = \mu_l \mathbf{1}_p \quad \text{and} \quad \boldsymbol{\Sigma}^{(l)} = \sigma_l^2 \{(1 - \rho)\boldsymbol{I}_p + \rho \boldsymbol{J}_p\}, \tag{2.1}$$

where $\mathbf{1}_p$ is the $p \times 1$ vector of 1's, $\boldsymbol{I}_p$ is the $p \times p$ identity matrix and $\boldsymbol{J}_p$ is the $p \times p$ matrix containing only ones. Here ρ is a common intraclass correlation coefficient.

We consider an orthogonal transformation of the Helmert type applied to the $X_i^{(l)}$. Let $Y_i^{(l)} = (Y_{i1}^{(l)}, \ldots, Y_{ip}^{(l)})' = \boldsymbol{Q} X_i^{(l)}$, where $\boldsymbol{Q}$ is an orthogonal matrix. Under this orthogonal transformation, it is obvious that

$$Y_i^{(l)} \sim N_p\left(\boldsymbol{\mu}^{(l)*}, \boldsymbol{\Sigma}^{(l)*}\right), \quad l = 1, \ldots, q, \ i = 1, \ldots, n_l, \tag{2.2}$$

where

$$\boldsymbol{\mu}^{(l)*} = (\mu_l, 0, \ldots, 0)' \quad \text{and} \quad \boldsymbol{\Sigma}^{(l)*} = \sigma_l^2 \text{Diag}\left\{p^{-1}[1 + (p-1)\rho], (1 - \rho), \ldots, (1 - \rho)\right\}.$$

We begin with finding the probability matching priors for the common intraclass correlation ρ. For a prior π, let $\theta_1^{1-\alpha}(\pi; \boldsymbol{Y})$ denote the $(1 - \alpha)^{th}$ percentile of the posterior distribution of θ_1, that is,

$$P^\pi\left[\theta_1 \leq \theta_1^{1-\alpha}(\pi; \boldsymbol{Y}) \,\middle|\, \boldsymbol{Y}\right] = 1 - \alpha, \tag{2.3}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_t)^T$, $t = 2q = 1$ and $\theta_1 = \rho$ is the parameter of interest. We want to find priors π for which

$$P\left[\theta_1 \leq \theta_1^{1-\alpha}(\pi; \boldsymbol{Y}) \,\middle|\, \boldsymbol{\theta}\right] = 1 - \alpha + o\left(n^{-u}\right) \tag{2.4}$$

for some $u > 0$, as n goes to infinity. Priors π satisfying (2.4) 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, it is convenient to introduce orthogonal parametrization, (Cox and Reid, 1987; Tibshirani, 1989). To this end, let

$$\theta = \rho, \qquad \lambda_l = \sigma_l^2 \left[p^{-1}(1 + (p-1)\rho)(1-\rho)^{p-1} \right]^{\frac{1}{p}}, \qquad \delta_l = \mu_l, \quad l = 1, \ldots, q. \qquad (2.5)$$

With this parametrization, the likelihood function of parameters $(\theta, \lambda_1, \ldots, \lambda_q, \delta_1, \ldots, \delta_q)$ for model (2.2) is given by

$$L(\theta, \boldsymbol{\lambda}, \boldsymbol{\delta}) \propto \left[\prod_{l=1}^{q} \lambda_l^{-\frac{n_l p}{2}} \right] \exp \left\{ - \sum_{l=1}^{q} \frac{1}{2\lambda_l} \left[p^{-1}(1 + (p-1)\theta)(1-\theta)^{p-1} \right]^{\frac{1}{p}} \right.$$
$$\left. \times \left[\frac{p \sum_{i=1}^{n_l} \left(y_{i1}^{(l)} - \delta_l \right)^2}{1 + (p-1)\theta} + \sum_{i=1}^{n_l} \sum_{j=2}^{p} \frac{\left(y_{ij}^{(l)} \right)^2}{1 - \theta} \right] \right\}, \qquad (2.6)$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_q)$ and $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_q)$. Based on (2.6), the Fisher information matrix is given by

$$I = \begin{pmatrix} I_{11} & 0 & 0 \\ 0 & \boldsymbol{I}_{22} & 0 \\ 0 & 0 & \boldsymbol{I}_{33} \end{pmatrix},$$

where

$$I_{11} = \frac{p(p-1)}{2} [1 + (p-1)\theta]^{-2} (1-\theta)^{-2} \sum_{l=1}^{q} n_l,$$

$$\boldsymbol{I}_{22} = \mathrm{Diag} \left\{ \frac{pn_1}{2\lambda_1^2}, \ldots, \frac{pn_q}{2\lambda_q^2} \right\},$$

$$\boldsymbol{I}_{33} = \left[p^{-1}(1 + (p-1)\theta) \right]^{\frac{1-p}{p}} (1-\theta)^{\frac{p-1}{p}} \mathrm{Diag} \left\{ \frac{n_1}{\lambda_1}, \ldots, \frac{n_q}{\lambda_q} \right\}.$$

Thus θ is orthogonal to $\boldsymbol{\lambda}$ and $\boldsymbol{\delta}$ in the sense of Cox and Reid (1987). Following Tibshirani (1989), the class of first order probability matching priors is characterized by

$$\pi_m^{(1)}(\theta, \boldsymbol{\lambda}, \boldsymbol{\delta}) \propto [1 + (p-1)\theta]^{-1} (1-\theta)^{-1} d(\boldsymbol{\lambda}, \boldsymbol{\delta}), \qquad (2.7)$$

where $d(\boldsymbol{\lambda}, \boldsymbol{\delta})$ is an arbitrary function differentiable in its arguments.

Clearly the class of prior given in (2.7), is quite large, and it is important to narrow down this class of priors. To this end, we consider the class of second order probability matching priors as given in Mukerjee and Ghosh (1997).

Theorem 1. *The second order probability matching prior is given by*

$$\pi_m^{(2)}(\theta, \boldsymbol{\lambda}, \boldsymbol{\delta}) = [1 + (p-1)\theta]^{-1} (1-\theta)^{-1} \lambda_1^{-1} \cdots \lambda_q^{-1} h(\boldsymbol{\delta}), \qquad (2.8)$$

where $h(\boldsymbol{\delta})$ is any smooth function of $\boldsymbol{\delta}$.

Proof: A second order probability matching prior is of the form (2.7), and also d must satisfy an additional differential equation (*cf* (2.10)) of Mukerjee and Ghosh (1997), namely

$$\frac{1}{6} d(\boldsymbol{\lambda}, \boldsymbol{\delta}) \frac{\partial}{\partial \theta} \left\{ I_{11}^{-\frac{3}{2}} L_{1,1,1} \right\} + \sum_{v=1}^{q} \frac{\partial}{\partial \lambda_v} \left\{ I_{11}^{-\frac{1}{2}} L_{11v} I^{vv} d(\boldsymbol{\lambda}, \boldsymbol{\delta}) \right\} + \sum_{v=1}^{q} \frac{\partial}{\partial \mu_v} \left\{ I_{11}^{-\frac{1}{2}} L_{11(q+v)} I^{(q+v)(q+v)} d(\boldsymbol{\lambda}, \boldsymbol{\delta}) \right\} = 0. \quad (2.9)$$

Here

$$L_{1,1,1} = E\left[\left(\frac{\partial \log L}{\partial \theta}\right)^3\right] = c[1 + (p-1)\theta]^{-3}(1-\theta)^{-3}, \quad c = \text{a constant}$$

$$L_{11v} = E\left[\frac{\partial^3 \log L}{\partial \theta^2 \partial \lambda_v}\right] = \frac{n_v}{2\lambda_v} p(p-1)[1 + (p-1)\theta]^{-2}(1-\theta)^{-2}, \quad v = 1, \ldots, q$$

$$L_{11(q+v)} = E\left[\frac{\partial^3 \log L}{\partial \theta^2 \partial \delta_v}\right] = 0, \quad v = 1, \ldots, q$$

and

$$I^{-1} = \begin{pmatrix} I^{11} & 0 & 0 \\ 0 & I^{22} & 0 \\ 0 & 0 & I^{33} \end{pmatrix},$$

where

$$I^{11} = \frac{2}{p(p-1)[\sum_{l=1}^q n_l]}[1 + (p-1)\theta]^2(1-\theta)^2,$$

$$I^{22} = \text{Diag}\left\{\frac{2\lambda_1^2}{pn_1}, \ldots, \frac{2\lambda_q^2}{pn_q}\right\},$$

$$I^{33} = \left[p^{-1}(1 + (p-1)\theta)\right]^{\frac{-1+p}{p}} (1-\theta)^{\frac{-p+1}{p}} \text{Diag}\left\{\frac{\lambda_1}{n_1}, \ldots, \frac{\lambda_q}{n_q}\right\}.$$

Then (2.9) simplifies to

$$\sum_{i=1}^q \frac{\partial}{\partial \lambda_i}\{\lambda_i d(\lambda, \delta)\} = 0. \tag{2.10}$$

Hence the set of solution of (2.10) is of the form

$$d(\lambda, \delta) = \lambda_1^{-1} \cdots \lambda_q^{-1} h(\delta), \tag{2.11}$$

where $h(\delta)$ is any smooth function of δ. Thus the resulting second order matching prior is

$$\pi_m^{(2)}(\theta, \lambda, \delta) = [1 + (p-1)\theta]^{-1} (1-\theta)^{-1} \lambda_1^{-1} \cdots \lambda_q^{-1} h(\delta).$$

This completes the proof.   □

Remark 1. There are alternative ways through which matching can be accomplished. One such approach (DiCiccio and Stern, 1994; Ghosh and Mukerjee, 1995) is matching through the HPD region. Datta *et al.* (2000) provided a theorem which establishes the equivalence of second order matching priors and HPD matching priors within the class of first order matching priors. The equivalence condition is that $I_{11}^{-3/2} L_{111}$ dose not depend on θ_1. Since

$$L_{111} = E\left[\frac{\partial^3 \log L}{\partial \theta^3}\right] = -p(p-1)\left[3(p-1)\theta - 2p + 4\right]\left[1 + (p-1)\theta\right]^{-3}(1-\theta)^{-3}\sum_{l=1}^q n_l$$

and

$$I_{11} = \frac{p(p-1)}{2} \left[1 + (p-1)\theta\right]^{-2} (1-\theta)^{-2} \sum_{l=1}^{q} n_l,$$

thus

$$I_{11}^{-\frac{3}{2}} L_{111} = -2^{\frac{3}{2}} \left[p(p-1)\right]^{-\frac{1}{2}} \left[\sum_{l=1}^{q} n_l\right]^{-\frac{1}{2}} \left[3(p-1)\theta - 2p + 4\right]$$

depends on θ. Therefore the second order probability matching prior for θ does not HPD matching prior. Mukerjee and Reid (1999) studied that a prior satisfying (2.4) matches $P[\theta_1 + \beta(I^{11}/n)^{1/2} \leq \theta_1^{1-\alpha}(\pi; Y)|\boldsymbol{\theta}]$ with the corresponding posterior probability, up to the same order and for each β and α, where the scalar β is free from n, $\boldsymbol{\theta}$ and Y. If a matching prior matches the alternative coverage probabilities then there is a stronger justification for calling it noninformative in so far as agreement with a frequentist is concerned. In general the second order matching prior may or may not match the alternative coverage probabilities up to the same order of approximation. Under orthogonal parametrization, Mukerjee and Reid (1999) gives the simple differential equations that the second order probability matching prior matches alternative coverage probabilities up to the second order. The differential Equation (3.20) of Mukerjee and Reid (1999) is $\partial/\partial\theta_1\{I_{11}^{-3/2}L_{111}\} = 0$. Since $\partial/\partial\theta\{I_{11}^{-3/2}L_{111}\} = -2^{3/2}3(p-1)[p(p-1)]^{-1/2}[\sum_{l=1}^{q} n_l]^{-1/2} \neq 0$, thus the second order probability matching prior for θ does not match the alternative coverage probabilities up to the second order.

Remark 2. In the original parameterization, the second order matching prior becomes

$$\pi_m^{(2)}\left(\rho, \sigma_1^2, \ldots, \sigma_q^2, \mu_1, \ldots, \mu_q\right) = \left[1 + (p-1)\rho\right]^{-1} (1-\rho)^{-1} \sigma_1^{-2} \cdots \sigma_q^{-2} h\left(\mu_1, \ldots, \mu_q\right). \qquad (2.12)$$

Reference priors introduced by Bernardo (1979), and extended further by Berger and Bernardo (1992) have become very popular over the years for the development of noninformative priors. We derive the reference priors for different groups of ordering of $(\theta, \lambda, \delta)$. Then due to the orthogonality of the parameters, we can follow the algorithm by Datta and Ghosh (1996). Choosing rectangular compacts for each θ, λ and δ, when θ is the parameter of interest, the reference priors are given by as follows.

Theorem 2. *If θ is the parameter of interest, then the reference prior distributions for different groups of ordering of $(\theta, \lambda, \delta)$ are:*

Group ordering	Reference prior
$\{(\theta, \delta), \lambda\}$	$\pi_1 \propto \lambda_1^{-1} \cdots \lambda_q^{-1} \left[1 + (p-1)\theta\right]^{\frac{-p(q+2)+q}{2p}} (1-\theta)^{\frac{p(q-2)-q}{2p}}$
$\{(\theta, \lambda, \delta)\}$	$\pi_2 \propto \lambda_1^{-\frac{3}{2}} \cdots \lambda_q^{-\frac{3}{2}} \left[1 + (p-1)\theta\right]^{\frac{-p(q+2)+q}{2p}} (1-\theta)^{\frac{p(q-2)-q}{2p}}$
$\{\theta, \lambda, \delta\}, \{\theta, \delta, \lambda\}, \{(\theta, \lambda), \delta\}$	$\pi_3 \propto \lambda_1^{-1} \cdots \lambda_q^{-1} \left[1 + (p-1)\theta\right]^{-1} (1-\theta)^{-1}$
$\{\theta, (\lambda, \delta)\}$	$\pi_4 \propto \lambda_1^{-\frac{3}{2}} \cdots \lambda_q^{-\frac{3}{2}} \left[1 + (p-1)\theta\right]^{-1} (1-\theta)^{-1}$

Remark 3. Note that one-at-a-time reference priors π_3 satisfies the second order matching criterion for θ. Here "one-at-a-time" means that we derive the reference priors by splitting the parameters into groups one at a time. The reference priors π_3 and π_4 do not depend on p and q.

In the Theorem 1, the second order probability matching prior is given by

$$\pi_m^{(2)}(\theta, \lambda, \delta) = [1 + (p-1)\theta]^{-1}(1-\theta)^{-1}\lambda_1^{-1}\cdots\lambda_q^{-1}h(\delta),$$

where $h(\delta)$ is any smooth function of δ. However every function is not permissible in the construction of priors. For instance, we consider any function of the form $(\delta_1\cdots\delta_q)^k$. If k is a positive integer, then the posterior distribution under functions of the form $(\delta_1\cdots\delta_q)^k$ is proper. But under the odd numbers, the marginal posterior distribution of (θ, λ) is zero. For the even numbers, the condition of propriety in this form strongly depend on the k. That is, as $k \to \infty$, the sample size $n \to \infty$ to satisfy the condition of propriety. Thus any function of the form θ_2^k excluded from consideration. Also $\delta_1, \cdots, \delta_q$ are the location parameters. Hence we consider a particular second order matching prior where h is a constant in the above matching prior. This prior is given by

$$\pi_m^{(2)}(\theta, \lambda, \delta) = \lambda_1^{-1}\cdots\lambda_q^{-1}[1 + (p-1)\theta]^{-1}(1-\theta)^{-1}. \tag{2.13}$$

3. Implementation of the Bayesian Procedure

We investigate the propriety of posteriors for a general class of priors which includes the reference priors and the second order matching prior (2.13). We consider the class of priors

$$\pi(\theta, \lambda, \delta) \propto (\lambda_1\cdots\lambda_q)^{-a}[1 + (p-1)\theta]^{-b}(1-\theta)^{-c}, \tag{3.1}$$

where $a > 0, b > 0$ and $-\infty < c < \infty$. The following general theorem can be proved.

Theorem 3. *The posterior distribution of $(\theta, \lambda, \delta)$ under the prior π, (3.1), is proper if $n_l p + 2a - 3 > 0$, $l = 1, \ldots, q$, $(p-1)(Np + 2qa - 2q) - 2bp \geq 0$ and $p(N - q - 2c) + 2q(a-1) \geq 0$, where $N = \sum_{l=1}^{q} n_l$.*

Proof: Note that the joint posterior for θ, λ and δ given **y** is

$$\pi(\theta, \lambda, \delta|y) \propto \left[\prod_{l=1}^{q} \lambda_l^{-\frac{n_l p}{2}-a}\right][1 + (p-1)\theta]^{-b}(1-\theta)^{-c} \tag{3.2}$$

$$\times \exp\left\{-\sum_{l=1}^{q}\frac{1}{2\lambda_l}\left[p^{-1}(1+(p-1)\theta)(1-\theta)^{p-1}\right]^{\frac{1}{p}}\left[\frac{p\sum_{i=1}^{n_l}\left(y_{i1}^{(l)}-\delta_l\right)^2}{1+(p-1)\theta}+\sum_{i=1}^{n_l}\sum_{j=2}^{p}\frac{\left(y_{ij}^{(l)}\right)^2}{1-\theta}\right]\right\},$$

where $y = (y_1^{(1)}, \ldots, y_{n_1}^{(1)}, \ldots, y_1^{(q)}, \ldots, y_{n_q}^{(q)})$. Firstly, we integrate with respect to $\delta_1, \ldots, \delta_q$ from (3.2). Then one gets

$$\pi(\theta, \lambda|y) \propto \left[\prod_{l=1}^{q}\lambda_l^{-\frac{n_l p+2a-1}{2}}\right][1+(p-1)\theta]^{\frac{-2bp-q+pq}{2p}}(1-\theta)^{\frac{-2cp-(p-1)q}{2p}} \tag{3.3}$$

$$\times \exp\left\{-\sum_{l=1}^{q}\frac{1}{2\lambda_l}\left[p^{-1}(1+(p-1)\theta)(1-\theta)^{p-1}\right]^{\frac{1}{p}}\left[\frac{p\sum_{i=1}^{n_l}\left(y_{i1}^{(l)}-\bar{y}_l\right)^2}{1+(p-1)\theta}+\sum_{i=1}^{n_l}\sum_{j=2}^{p}\frac{\left(y_{ij}^{(l)}\right)^2}{1-\theta}\right]\right\},$$

where $\bar{y}_l = \sum_{i=1}^{n_l}y_{i1}^{(l)}/n_l$, $l = 1, \ldots, q$. Next, integrating with respect to $\lambda_1, \ldots, \lambda_q$, it follows from (3.3) that

$$\pi(\theta|y) \propto [1+(p-1)\theta]^{\frac{(p-1)(Np+2qa-2q)-2bp}{2p}}(1-\theta)^{\frac{p(N-q-2c)+2q(a-1)}{2p}}$$

$$\times \prod_{l=1}^{q}\left\{p(1-\theta)\sum_{i=1}^{n_l}\left(y_{i1}^{(l)}-\bar{y}_l\right)^2+[1+(p-1)\theta]\sum_{i=1}^{n_l}\sum_{j=2}^{p}\left(y_{ij}^{(l)}\right)^2\right\}^{-\frac{n_l p+2a-3}{2}}, \tag{3.4}$$

where $N = \sum_{l=1}^{q} n_l$. Here $n_l p + 2a - 3 > 0$, $l = 1, \ldots, q$. Finally, we integrate with respect to θ.

$$\int_{-\frac{1}{p-1}}^{1} \pi(\theta|\mathbf{y})d\theta = \int_{-\frac{1}{p-1}}^{1} [1 + (p-1)\theta]^{\frac{(p-1)(Np+2qa-2q)-2bp}{2p}} (1-\theta)^{\frac{p(N-q-2c)+2q(a-1)}{2p}}$$

$$\times \prod_{l=1}^{q} \left\{ p(1-\theta) \sum_{i=1}^{n_l} \left(y_{i1}^{(l)} - \bar{y}_l \right)^2 + [1 + (p-1)\theta] \sum_{i=1}^{n_l} \sum_{j=2}^{p} \left(y_{ij}^{(l)} \right)^2 \right\}^{-\frac{n_l p + 2a - 3}{2}} d\theta. \quad (3.5)$$

Thus the last integral is finite if $(p-1)(Np + 2qa - 2q) - 2bp \geq 0$ and $p(N - q - 2c) + 2q(a-1) \geq 0$. This completes the proof. $\qquad\square$

Next, we provide the marginal density of θ under the above noninformative prior.

Theorem 4. *Under the prior π, (3.1), the marginal posterior density of ρ, is given by*

$$\pi(\theta|\mathbf{y}) \propto [1 + (p-1)\theta]^{\frac{(p-1)(Np+2qa-q)-2bp}{2p}} (1-\theta)^{\frac{p(N-2q-2c)+q(2a-1)}{2p}}$$

$$\times \prod_{l=1}^{q} \left\{ p(1-\theta) \sum_{i=1}^{n_l} \left(y_{i1}^{(l)} - \bar{y}_l \right)^2 + [1 + (p-1)\theta] \sum_{i=1}^{n_l} \sum_{j=2}^{p} \left(y_{ij}^{(l)} \right)^2 \right\}^{-\frac{n_l p + 2a - 3}{2}}, \quad (3.6)$$

where $\bar{y}_l = \sum_{i=1}^{n_l} y_{i1}^{(l)}/n_l$, $l = 1, \ldots, q$ and $N = \sum_{l=1}^{q} n_l$.

The normalizing constant for the marginal density of θ requires a one dimensional integration. Therefore we have the marginal posterior density of θ, and so it is easy to compute the marginal moment of θ. For π_1 reference prior, $a = 1$, $b = (p(q + 2) - q)/(2p)$ and $c = (q - (q - 2)p)/(2p)$, for π_2 reference prior, $a = 3/2$, $b = (p(q + 2) - q)/(2p)$ and $c = (q - (q - 2)p)/(2p)$, for π_3 reference prior, $a = 1$, $b = 1$ and $c = 1$ and for π_4 reference prior, $a = 3/2$, $b = 1$ and $c = 1$ in this marginal density. Note that except for π_3 and π_4, the other cases depend on p and q. In Section 4, we investigate the frequentist coverage probabilities for the π_1, π_2, π_3 and π_4.

4. Numerical Analysis

4.1. Simulation study

We evaluate the frequentist coverage probability by investigating the credible interval using the marginal posterior density of θ under the noninformative prior π given in (3.1) for several configurations. That is to say, the frequentist coverage of a αth posterior quantile should be close to α. This is done numerically. Table 1 and Table 2 give numerical values of the frequentist coverage probabilites 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 $(\rho, \sigma_1^2, \ldots, \sigma_q^2, \mu_1, \ldots, \mu_q)$ and any prespecified probability value α. Here α is 0.05 (0.95). Let $\theta^{\alpha}(\pi; \mathbf{Y})$ be the posterior α-quantile of θ given $\mathbf{Y}$. That is to say, $F(\theta^{\alpha}(\pi; \mathbf{Y})|\mathbf{Y}) = \alpha$, where $F(\cdot|\mathbf{Y})$ is the marginal posterior distribution of θ. Then the frequentist coverage probability of this one sided credible interval of θ is

$$P_{(\theta, \boldsymbol{\lambda}, \boldsymbol{\delta})}(\alpha; \theta) = P_{(\theta, \boldsymbol{\lambda}, \boldsymbol{\delta})}(0 < \theta \leq \theta^{\alpha}(\pi; \mathbf{Y})). \quad (4.1)$$

The estimated $P_{(\theta, \boldsymbol{\lambda}, \boldsymbol{\delta})}(\alpha; \theta)$ when $\alpha = 0.05(0.95)$ is shown in Table 1 and Table 2.

Table 1: Frequentist coverage probability of 0.05 (0.95) posterior quantiles of ρ when $q = 3$

ρ	μ_1, μ_2, μ_3	$\sigma_1, \sigma_2, \sigma_3$	n_1, n_2, n_3	π_1	π_2	π_3	π_4
0.1	0, 0, 0	1, 1, 1	3, 3, 3	0.010(0.789)	0.012(0.772)	0.051(0.951)	0.051(0.937)
			3, 3, 5	0.011(0.825)	0.012(0.812)	0.049(0.953)	0.049(0.941)
			5, 5, 5	0.015(0.859)	0.017(0.852)	0.049(0.951)	0.050(0.944)
			5, 5, 7	0.016(0.870)	0.017(0.863)	0.047(0.951)	0.048(0.946)
			10, 10, 10	0.024(0.897)	0.024(0.894)	0.051(0.949)	0.052(0.946)
		1, 2, 3	3, 3, 3	0.009(0.786)	0.011(0.769)	0.050(0.946)	0.050(0.931)
			3, 3, 5	0.012(0.826)	0.014(0.814)	0.050(0.954)	0.051(0.944)
			5, 5, 5	0.016(0.856)	0.017(0.848)	0.049(0.948)	0.050(0.943)
			5, 5, 7	0.017(0.863)	0.018(0.856)	0.047(0.949)	0.049(0.944)
			10, 10, 10	0.023(0.894)	0.024(0.891)	0.050(0.947)	0.050(0.944)
	1, 2, 3	1, 1, 1	3, 3, 3	0.009(0.800)	0.010(0.782)	0.053(0.952)	0.053(0.936)
			3, 3, 5	0.013(0.827)	0.014(0.813)	0.050(0.952)	0.050(0.942)
			5, 5, 5	0.015(0.860)	0.016(0.852)	0.051(0.952)	0.052(0.944)
			5, 5, 7	0.016(0.870)	0.017(0.864)	0.047(0.952)	0.048(0.945)
			10, 10, 10	0.024(0.895)	0.024(0.891)	0.050(0.948)	0.051(0.945)
		1, 2, 3	3, 3, 3	0.007(0.791)	0.009(0.774)	0.050(0.948)	0.050(0.933)
			3, 3, 5	0.014(0.831)	0.015(0.816)	0.052(0.954)	0.053(0.945)
			5, 5, 5	0.017(0.853)	0.018(0.844)	0.052(0.950)	0.053(0.943)
			5, 5, 7	0.015(0.865)	0.016(0.858)	0.049(0.948)	0.049(0.941)
			10, 10, 10	0.024(0.896)	0.025(0.893)	0.052(0.950)	0.053(0.947)
0.5	0, 0, 0	1, 1, 1	3, 3, 3	0.008(0.792)	0.009(0.774)	0.049(0.950)	0.049(0.935)
			3, 3, 5	0.012(0.822)	0.014(0.810)	0.046(0.946)	0.046(0.936)
			5, 5, 5	0.016(0.854)	0.017(0.845)	0.050(0.951)	0.051(0.943)
			5, 5, 7	0.016(0.866)	0.018(0.859)	0.047(0.948)	0.047(0.941)
			10, 10, 10	0.022(0.899)	0.023(0.896)	0.051(0.953)	0.052(0.949)
		1, 2, 3	3, 3, 3	0.009(0.790)	0.010(0.771)	0.049(0.950)	0.049(0.934)
			3, 3, 5	0.010(0.824)	0.011(0.810)	0.050(0.949)	0.050(0.939)
			5, 5, 5	0.015(0.851)	0.017(0.842)	0.047(0.948)	0.048(0.941)
			5, 5, 7	0.017(0.864)	0.019(0.858)	0.052(0.949)	0.053(0.944)
			10, 10, 10	0.024(0.896)	0.025(0.893)	0.050(0.950)	0.051(0.947)
	1, 2, 3	1, 1, 1	3, 3, 3	0.010(0.796)	0.012(0.779)	0.053(0.951)	0.053(0.936)
			3, 3, 5	0.013(0.829)	0.015(0.815)	0.051(0.952)	0.051(0.941)
			5, 5, 5	0.017(0.858)	0.018(0.849)	0.052(0.949)	0.053(0.942)
			5, 5, 7	0.019(0.866)	0.020(0.859)	0.048(0.949)	0.049(0.944)
			10, 10, 10	0.023(0.895)	0.024(0.892)	0.050(0.948)	0.051(0.946)
		1, 2, 3	3, 3, 3	0.010(0.794)	0.012(0.778)	0.051(0.953)	0.051(0.937)
			3, 3, 5	0.012(0.826)	0.014(0.815)	0.049(0.950)	0.049(0.939)
			5, 5, 5	0.019(0.856)	0.020(0.848)	0.053(0.948)	0.054(0.941)
			5, 5, 7	0.017(0.864)	0.017(0.857)	0.050(0.948)	0.050(0.943)
			10, 10, 10	0.024(0.897)	0.025(0.894)	0.051(0.951)	0.053(0.948)
0.9	0, 0, 0	1, 1, 1	3, 3, 3	0.009(0.795)	0.009(0.777)	0.049(0.950)	0.049(0.935)
			3, 3, 5	0.014(0.821)	0.015(0.806)	0.050(0.947)	0.050(0.936)
			5, 5, 5	0.018(0.854)	0.019(0.846)	0.052(0.951)	0.053(0.944)
			5, 5, 7	0.018(0.870)	0.019(0.862)	0.051(0.950)	0.052(0.944)
			10, 10, 10	0.021(0.895)	0.022(0.892)	0.048(0.949)	0.049(0.944)
		1, 2, 3	3, 3, 3	0.010(0.796)	0.012(0.779)	0.056(0.948)	0.055(0.934)
			3, 3, 5	0.012(0.821)	0.013(0.808)	0.053(0.951)	0.054(0.941)
			5, 5, 5	0.015(0.855)	0.016(0.846)	0.052(0.948)	0.052(0.940)
			5, 5, 7	0.017(0.861)	0.017(0.856)	0.048(0.947)	0.048(0.940)
			10, 10, 10	0.021(0.900)	0.022(0.896)	0.046(0.952)	0.046(0.950)
	1, 2, 3	1, 1, 1	3, 3, 3	0.009(0.796)	0.011(0.778)	0.049(0.955)	0.049(0.942)
			3, 3, 5	0.013(0.826)	0.014(0.812)	0.047(0.950)	0.047(0.941)
			5, 5, 5	0.015(0.856)	0.017(0.847)	0.050(0.951)	0.051(0.945)
			5, 5, 7	0.016(0.868)	0.017(0.861)	0.050(0.951)	0.050(0.945)
			10, 10, 10	0.024(0.896)	0.025(0.893)	0.053(0.948)	0.053(0.946)
		1, 2, 3	3, 3, 3	0.010(0.794)	0.012(0.779)	0.050(0.950)	0.050(0.935)
			3, 3, 5	0.012(0.822)	0.014(0.810)	0.050(0.952)	0.050(0.939)
			5, 5, 5	0.017(0.858)	0.018(0.850)	0.050(0.952)	0.051(0.944)
			5, 5, 7	0.018(0.861)	0.019(0.853)	0.054(0.946)	0.055(0.939)
			10, 10, 10	0.022(0.892)	0.023(0.889)	0.050(0.946)	0.052(0.943)

In particular, for fixed $(\rho, \sigma_1^2, \ldots, \sigma_q^2, \mu_1, \ldots, \mu_q)$, p, q and $(n_1, \ldots, n_q)$, we take 10,000 independent random samples of Y from the model (2.2). Note that under the prior π, for fixed Y, $\theta \leq \theta^\alpha(\pi; Y)$ if and only if $F(\theta|Y) \leq \alpha$. Under the prior π, $P_{(\theta, \lambda, \delta)}(\alpha; \rho)$ can be estimated by the relative frequency of

Table 2: Frequentist coverage probability of 0.05 (0.95) posterior quantiles of ρ when $q = 5$

ρ	$\mu_1,\ldots,\mu_5$	$\sigma_1,\ldots,\sigma_5$	$n_1,\ldots,n_5$	π_1	π_2	π_3	π_4
0.1	0, 0, 0, 0, 0	1, 1, 1, 1, 1	3, 3, 3, 3, 3	0.006(0.730)	0.007(0.708)	0.050(0.948)	0.049(0.932)
			3, 3, 3, 5, 5	0.008(0.779)	0.009(0.765)	0.049(0.948)	0.048(0.938)
			5, 5, 5, 5, 5	0.010(0.825)	0.011(0.818)	0.048(0.952)	0.048(0.945)
			5, 5, 5, 7, 7	0.013(0.837)	0.014(0.829)	0.052(0.950)	0.052(0.943)
			10, 10, 10, 10, 10	0.018(0.879)	0.019(0.875)	0.048(0.949)	0.048(0.945)
		1, 2, 3, 4, 5	3, 3, 3, 3, 3	0.005(0.728)	0.006(0.709)	0.050(0.950)	0.049(0.936)
			3, 3, 3, 4, 5	0.008(0.777)	0.010(0.760)	0.051(0.948)	0.049(0.937)
			5, 5, 5, 5, 5	0.010(0.817)	0.011(0.809)	0.048(0.952)	0.048(0.944)
			5, 5, 5, 7, 7	0.011(0.833)	0.012(0.826)	0.050(0.951)	0.051(0.943)
			10, 10, 10, 10, 10	0.019(0.869)	0.019(0.866)	0.050(0.947)	0.051(0.944)
	1, 2, 3, 4, 5	1, 1, 1, 1, 1	3, 3, 3, 3, 3	0.005(0.734)	0.006(0.714)	0.054(0.954)	0.053(0.937)
			3, 3, 3, 5, 5	0.008(0.777)	0.008(0.763)	0.050(0.949)	0.049(0.937)
			5, 5, 5, 5, 5	0.011(0.820)	0.012(0.810)	0.053(0.951)	0.053(0.943)
			5, 5, 5, 7, 7	0.011(0.832)	0.012(0.825)	0.046(0.950)	0.047(0.944)
			10, 10, 10, 10, 10	0.017(0.872)	0.017(0.870)	0.047(0.950)	0.048(0.946)
		1, 2, 3, 4, 5	3, 3, 3, 3, 3	0.006(0.732)	0.007(0.714)	0.048(0.950)	0.047(0.934)
			3, 3, 3, 5, 5	0.007(0.779)	0.008(0.766)	0.049(0.950)	0.047(0.940)
			5, 5, 5, 5, 5	0.010(0.817)	0.011(0.806)	0.050(0.950)	0.051(0.944)
			5, 5, 5, 7, 7	0.011(0.837)	0.012(0.828)	0.053(0.948)	0.053(0.941)
			10, 10, 10, 10, 10	0.016(0.876)	0.016(0.872)	0.048(0.947)	0.049(0.944)
0.5	0, 0, 0, 0, 0	1, 1, 1, 1, 1	3, 3, 3, 3, 3	0.006(0.728)	0.007(0.710)	0.049(0.951)	0.048(0.935)
			3, 3, 3, 5, 5	0.008(0.774)	0.009(0.758)	0.051(0.947)	0.051(0.937)
			5, 5, 5, 5, 5	0.009(0.814)	0.010(0.802)	0.048(0.945)	0.048(0.938)
			5, 5, 5, 7, 7	0.012(0.833)	0.012(0.824)	0.050(0.948)	0.051(0.941)
			10, 10, 10, 10, 10	0.019(0.876)	0.020(0.872)	0.050(0.947)	0.050(0.943)
		1, 2, 3, 4, 5	3, 3, 3, 3, 3	0.005(0.723)	0.005(0.703)	0.047(0.949)	0.045(0.931)
			3, 3, 3, 5, 5	0.006(0.774)	0.007(0.761)	0.050(0.953)	0.050(0.942)
			5, 5, 5, 5, 5	0.011(0.807)	0.012(0.798)	0.052(0.945)	0.052(0.937)
			5, 5, 5, 7, 7	0.011(0.836)	0.012(0.828)	0.051(0.949)	0.051(0.944)
			10, 10, 10, 10, 10	0.016(0.876)	0.017(0.873)	0.049(0.949)	0.050(0.947)
	1, 2, 3, 4, 5	1, 1, 1, 1, 1	3, 3, 3, 3, 3	0.005(0.736)	0.006(0.715)	0.048(0.949)	0.046(0.934)
			3, 3, 3, 5, 5	0.008(0.778)	0.008(0.762)	0.049(0.952)	0.048(0.942)
			5, 5, 5, 5, 5	0.011(0.819)	0.012(0.810)	0.049(0.950)	0.050(0.944)
			5, 5, 5, 7, 7	0.013(0.834)	0.014(0.828)	0.052(0.950)	0.053(0.944)
			10, 10, 10, 10, 10	0.018(0.877)	0.019(0.875)	0.048(0.948)	0.049(0.944)
		1, 2, 3, 4, 5	3, 3, 3, 3, 3	0.006(0.726)	0.006(0.706)	0.050(0.947)	0.048(0.932)
			3, 3, 3, 5, 5	0.007(0.774)	0.008(0.760)	0.048(0.949)	0.048(0.937)
			5, 5, 5, 5, 5	0.011(0.821)	0.012(0.813)	0.053(0.949)	0.053(0.943)
			5, 5, 5, 7, 7	0.014(0.843)	0.014(0.835)	0.056(0.953)	0.056(0.947)
			10, 10, 10, 10, 10	0.018(0.873)	0.019(0.869)	0.052(0.948)	0.052(0.945)
0.9	0, 0, 0, 0, 0	1, 1, 1, 1, 1	3, 3, 3, 3, 3	0.005(0.731)	0.007(0.711)	0.049(0.950)	0.047(0.933)
			3, 3, 3, 5, 5	0.007(0.783)	0.008(0.766)	0.048(0.949)	0.047(0.938)
			5, 5, 5, 5, 5	0.011(0.819)	0.012(0.811)	0.053(0.949)	0.053(0.943)
			5, 5, 5, 7, 7	0.013(0.839)	0.014(0.831)	0.053(0.950)	0.054(0.945)
			10, 10, 10, 10, 10	0.017(0.875)	0.017(0.872)	0.049(0.948)	0.050(0.946)
		1, 2, 3, 4, 5	3, 3, 3, 3, 3	0.006(0.724)	0.007(0.706)	0.051(0.947)	0.050(0.931)
			3, 3, 3, 4, 5	0.007(0.776)	0.008(0.762)	0.047(0.952)	0.045(0.941)
			5, 5, 5, 5, 5	0.010(0.813)	0.011(0.803)	0.051(0.949)	0.052(0.940)
			5, 5, 5, 7, 7	0.011(0.842)	0.011(0.835)	0.051(0.953)	0.052(0.946)
			10, 10, 10, 10, 10	0.015(0.881)	0.016(0.877)	0.049(0.953)	0.050(0.951)
	1, 2, 3, 4, 5	1, 1, 1, 1, 1	3, 3, 3, 3, 3	0.005(0.733)	0.005(0.714)	0.050(0.951)	0.048(0.935)
			3, 3, 3, 5, 5	0.010(0.785)	0.010(0.769)	0.054(0.953)	0.053(0.943)
			5, 5, 5, 5, 5	0.011(0.817)	0.012(0.806)	0.052(0.948)	0.052(0.941)
			5, 5, 5, 7, 7	0.013(0.834)	0.014(0.826)	0.052(0.949)	0.052(0.944)
			10, 10, 10, 10, 10	0.017(0.878)	0.018(0.874)	0.053(0.952)	0.053(0.949)
		1, 2, 3, 4, 5	3, 3, 3, 3, 3	0.006(0.725)	0.007(0.706)	0.049(0.948)	0.047(0.935)
			3, 3, 3, 5, 5	0.008(0.773)	0.009(0.759)	0.049(0.950)	0.048(0.941)
			5, 5, 5, 5, 5	0.011(0.819)	0.012(0.810)	0.050(0.948)	0.050(0.940)
			5, 5, 5, 7, 7	0.014(0.835)	0.015(0.827)	0.051(0.949)	0.052(0.941)
			10, 10, 10, 10, 10	0.018(0.874)	0.018(0.870)	0.048(0.948)	0.049(0.945)

$F(\theta|Y) \leq \alpha$. For the cases presented in Table 1 and Table 2 we see that the noninformative prior π_3 and π_4 meet very well the target coverage probabilities. Also note that the results of tables are not much sensitive to the change of the values of $\rho, \sigma_1^2, \ldots, \sigma_5^2$ and $\mu_1, \ldots, \mu_5$, and one-at-a-time reference

Table 3: Number of finger ridges on both hands of individuals in 12 pairs of female identical twins

Population	Finger Ridges	Population	Finger Ridges
1	71 71	2	76 70
1	79 82	2	83 82
1	105 99	2	114 113
1	115 114	2	57 44
1	114 113	2	94 91
1	75 83	2	76 72

prior π_3 matches the target very well. Thus we recommend to use the one-at-a-time reference prior when using the matching criterion.

4.2. Example

We provide several estimators for the common intraclass correlation coefficients ρ using real data from Snedecor and Cochran (1980), which are randomly split into two populations as shown in Table 3. The data are numbers of finger ridges on both hands of individuals in 12 pairs of female identical twins.

Three estimators for the common intraclass correlation coefficients are the maximum likelihood estimator ρ_M (Donner and Bull, 1983), the approximate estimator ρ_{KG} (Konishi and Gupta, 1989) and the Bayes estimator using one-at-a-time reference priors ρ_B. They are given by $\rho_M = 0.9608$, $\rho_{KG} = 0.9595$ and $\rho_B = 0.9611$, respectively. Three estimates are very similar. Because we split the data into two populations from one population, we calculate the maximum likelihood estimate for the intraclass correlation coefficient in one population which is given by $\rho^M = 0.9626$. The Bayes estimate is closer to ρ^M.

References

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.

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.

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference (with discussion), *Journal of Royal Statistical Society, Series B*, **41**, 113–147.

Cox, D. R. and Reid, N. (1987). Orthogonal parameters and approximate conditional inference (with discussion), *Journal of Royal Statistical Society, Series B*, **49**, 1–39.

Datta, G. S. and Ghosh, J. K. (1995a). On priors providing frequentist validity for Bayesian inference, *Biometrika*, **82**, 37–45.

Datta, G. S. and Ghosh, M. (1995b). Some remarks on noninformative priors, *Journal of the American Statistical Association*, **90**, 1357–1363.

Datta, G. S. and Ghosh, M. (1996). On the invariance of noninformative priors, *The Annal of Statistics*, **24**, 141–159.

Datta, G. S., Ghosh, M. and Mukerjee, R. (2000). Some new results on probability matching priors, *Calcutta Statistical Association Bulletin*, **50**, 179–192.

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, Series B*, **56**, 397–408.

Donner, A. and Bull, S. (1983). Inference concerning a common intraclass correlation coefficient, *Biometrics*, **39**, 771–775.

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.

Ghosh, J. K. and Mukerjee, R. (1995). Frequentist validity of highest posterior density regions in the presence of nuisance parameters, *Statistics & Decisions*, **13**, 131–139.

Huang, W. and Sinha, B. K. (1993). On optimum invariant tests of equality of intraclass correlation coefficents, *Annals of the Institute of Statistical Mathematics*, **45**, 579–597.

Konishi, S. and Gupta, A. K. (1989). Testing the equality several intraclass correlation coefficients, *Journal of Statistical Planning and Inference*, **21**, 93–105.

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.

Mukerjee, R. and Reid, N. (1999). On a property of probability matching priors: Matching the alternative coverage probabilities, *Biometrika*, **86**, 333–340.

Snedecor, G. W. and Cochran, G. (1980) *Statistical Methods, 7th Ed.*, Iowa State University Press, Ames, Iowa.

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.

Tibshirani, R. (1989). Noninformative priors for one parameter of many, *Biometrika*, **76**, 604–608.

Welch, B. N. and Peers, B. (1963). On formulae for confidence points based on integrals of weighted likelihood, *Journal of Royal Statistical Society, Series B*, **25**, 318–329.

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