Journal of the Korean Statistical Society (2004), **33**: 2, pp 203–218

NONINFORMATIVE PRIORS FOR LINEAR COMBINATION OF THE INDEPENDENT NORMAL MEANS

SANG GIL KANG[1], DAL HO KIM[2] AND WOO DONG LEE[3]

ABSTRACT

In this paper, we develop the matching priors and the reference priors for linear combination of the means under the normal populations with equal variances. We prove that the matching priors are actually the second order matching priors and reveal that the second order matching priors match alternative coverage probabilities up to the second order (Mukerjee and Reid, 1999) and also, are HPD matching priors. 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. We compute Bayesian credible intervals for linear combination of the means based on the reference priors.

AMS 2000 subject classifications. Primary 62F15; Secondary 62F25.
Keywords. Probability matching priors, alternative coverage probabilities, HPD matching priors, reference priors, linear combination of the means.

1. INTRODUCTION

Consider observations Y_{ij} from I normal populations specified by the following model:

$$Y_{ij} = \mu_i + \epsilon_{ij}, \text{ for } i = 1, \ldots, I \text{ and } j = 1, \ldots, J, \tag{1.1}$$

Received March 2003; accepted February 2004.
[1]Department of Applied Statistics, Sangji University, Wonju 220-702, Korea (e-mail : sangkg@sangji.ac.kr)
[2]Department of Statistics, Kyungpook National University, Taegu 702-701, Korea (e-mail : dalkim@knu.ac.kr)
[3]Department of Asset Managemental Science, Daegu Haany University, Kyungsan 712-240, Korea (e-mail : wdlee@dhu.ac.kr)

where $\mu_i, i = 1, \ldots, I$ are an unknown constants and ϵ_{ij} are independent normal variables with means 0 and variances σ^2. Let $\gamma = \sum_{i=1}^{I} \gamma_i \mu_i$, linear combination of the means, be our parameter of interest. The present paper focuses on noninformative priors for γ.

The problem of estimating linear combination of the means has received little attention from the Bayesian perspective. Among the little works, recently, Li and Stern (1997) studied the Bayesian intervals for linear combination of the means in a balanced nested design based on Jeffreys' prior. Kim *et al.* (2004) derived the noninformative priors for linear combination of the means in a balanced nested design.

A problem of making inference about the linear combination of normal means occurs when calculating the difference of two populations or evaluating the confidence interval for a treatment contrast under the ANOVA model. We consider Bayesian priors such that the resulting credible intervals for γ 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.

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, b, 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 described earlier.

The objective of this paper is to develop some noninformative priors on γ. We obtain the second order probability matching priors and these matching priors match the alternative coverage probabilities up to second order. We show that these priors are also HPD matching priors. And we show that among all of the reference priors, only one-at-a-time reference prior satisfies a second order matching criterion. Through simulation study, these priors are compared in the light of how accurately the coverage probabilities of Bayesian credible intervals match the corresponding frequentist coverage probabilities. A real data example is given.

The outline of the remaining sections is as follows. In Section 2, we derive Fisher information matrix under the reparametrization. Then we develop first order and second order probability matching priors for γ. We reveal that the second order matching prior matches the alternative coverage probabilities up to the same order. And we prove that the second order matching prior is a HPD matching prior. Also we develop the reference priors for different groups of ordering for the parameters. It turns out that among all of the reference priors, only one reference prior satisfies a second order matching criterion. In Section 3, we provide that the propriety of the posterior distribution for a general class of prior distributions which include the reference priors as well as second order matching priors. We prove a result which establishes the symmetry and unimodality of the posterior distribution of γ. In Section 4, simulated frequentist coverage probabilities under the proposed priors are given. We provide a real example for computation of the Bayes estimates and credible intervals using real data sets under the reference priors.

2. THE NONINFORMATIVE PRIORS

2.1. The probability matching priors

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

$$P^{\pi}\left\{\theta_1 \le \theta_1^{1-\alpha}(\pi;\mathbf{Y})\,|\mathbf{Y}\right\} = 1-\alpha, \tag{2.1}$$

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

$$P\left\{\theta_1 \le \theta_1^{1-\alpha}(\pi;\mathbf{Y})\,|\boldsymbol{\theta}\right\} = 1-\alpha + o(n^{-u}). \tag{2.2}$$

for some $u > 0$, as the sample size n goes to infinity. Priors π satisfying (2.2) 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

$$\gamma = \sum_{i=1}^{I}\gamma_i\mu_i,\quad \delta_i = \frac{\gamma_i(\sum_{j=1}^{I}\gamma_j\mu_j) - \mu_i(\sum_{j=1}^{I}\gamma_j^2)}{\gamma_1},\quad i = 2,\ldots,I,\quad \sigma^2 = \sigma^2.$$

With this parametrization, the likelihood function of parameters $(\gamma, \delta_2, \ldots, \delta_I, \sigma^2)$ for the model (1.1) is given by

$$L(\gamma, \delta_2, \cdots, \delta_I, \sigma^2) \propto (\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2}\left(S_1^2 + S_2^2\right)\right\}, \tag{2.3}$$

where

$$Y_{i\cdot} = \sum_{j=1}^{J} \frac{Y_{ij}}{J}, \quad S_1^2 = \sum_{i=1}^{I}\sum_{j=1}^{I}(Y_{ij} - Y_{i\cdot})^2,$$

$$S_2^2 = \left\{ J\left(Y_{1\cdot} - \frac{\gamma_1\gamma + \sum_{j=2}^{I}\gamma_j\delta_j}{\sum_{j=1}^{I}\gamma_j^2}\right)^2 + J\sum_{i=2}^{I}\left(Y_{i\cdot} - \frac{\gamma_i\gamma - \gamma_1\delta_i}{\sum_{j=1}^{I}\gamma_j^2}\right)^2\right\},$$

and $N = IJ$. Based on (2.3), the Fisher information matrix is given by

$$\mathbf{I}_F = \begin{pmatrix} I_{11} & \mathbf{0}^T & 0 \\ \mathbf{0} & \mathbf{I}_{22} & \mathbf{0}^T \\ 0 & \mathbf{0} & I_{(I+1)(I+1)} \end{pmatrix},$$

where

$$I_{11} = \frac{J}{\sum_{j=1}^{I}\gamma_j^2}(\sigma^2)^{-1}, \quad I_{(I+1)(I+1)} = \frac{N}{2}(\sigma^2)^{-2},$$

$$\mathbf{I}_{22} = \begin{pmatrix} a_{22} & a_{23} & \cdots & a_{2I} \\ a_{32} & a_{33} & \cdots & a_{3I} \\ \vdots & \vdots & & \vdots \\ a_{I2} & a_{I3} & \cdots & a_{II} \end{pmatrix}$$

with

$$a_{ii} = \frac{J(\gamma_i^2 + \gamma_1^2)}{(\sum_{j=1}^{I}\gamma_j^2)^2}(\sigma^2)^{-1}, \quad i = 2, \ldots, I,$$

$$a_{ij} = \frac{J\gamma_i\gamma_j}{(\sum_{j=1}^{I}\gamma_j^2)^2}(\sigma^2)^{-1}, \quad i, j = 2, \ldots, I.$$

Let $\boldsymbol{\delta} = (\delta_2, \ldots, \delta_I)^T$. From the above Fisher information matrix $\mathbf{I}_F$, γ is orthogonal to $\boldsymbol{\delta}$ and σ^2 in the sense of Cox and Reid (1987). Following Tibshirani (1989), the class of first order probability matching prior is characterized by

$$\pi_m^{(1)}(\gamma, \boldsymbol{\delta}, \sigma^2) \propto (\sigma^2)^{-1/2} d(\boldsymbol{\delta}, \sigma^2). \tag{2.4}$$

where $d(\boldsymbol{\delta},\sigma^2)$ is an arbitrary function differentiable in its arguments.

The class of prior given in (2.4) can be narrowed down to the second order probability matching priors as given in Mukerjee and Ghosh (1997).

THEOREM 2.1. *The second order probability matching priors are given by*

$$\pi_m^{(2)}(\gamma,\boldsymbol{\delta},\sigma^2)=(\sigma^2)^{-1}h(\boldsymbol{\delta}), \tag{2.5}$$

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

PROOF. A second order probability matching prior is of the form (2.4), and also d must satisfy an additional differential equation (*cf.* (2.10)) of Mukerjee and Ghosh (1997), namely

$$\frac{1}{6}d(\boldsymbol{\delta},\sigma^2)\frac{\partial}{\partial\gamma}\left(I_{11}^{-3/2}L_{1,1,1}\right)+\sum_{v=2}^{I}\sum_{s=2}^{I}\frac{\partial}{\partial\delta_v}\left\{I_{11}^{-1/2}L_{11s}I^{sv}d(\boldsymbol{\delta},\sigma^2)\right\}$$
$$+\frac{\partial}{\partial\phi}\left\{I_{11}^{-1/2}L_{11(I+1)}I^{(I+1)(I+1)}d(\boldsymbol{\delta},\sigma^2)\right\}=0, \tag{2.6}$$

where

$$L_{1,1,1}=E\left\{\left(\frac{\partial\log L}{\partial\gamma}\right)^3\right\}=0,\quad L_{11s}=E\left(\frac{\partial^3\log L}{\partial\gamma^2\partial\delta_s}\right)=0,\quad s=2,\ldots,I,$$
$$L_{11(I+1)}=E\left(\frac{\partial^3\log L}{\partial\gamma^2\partial\sigma^2}\right)=\frac{J}{\sum_{j=1}^{I}\gamma_j^2}(\sigma^2)^{-2},\quad I^{(I+1)(I+1)}=\frac{2}{N}(\sigma^2)^2$$

and

$$\mathbf{I}^{22}=\mathbf{I}_{22}^{-1}=\begin{pmatrix} I^{22} & I^{23} & \ldots & I^{2I}\\ I^{32} & I^{33} & \ldots & I^{3I}\\ \vdots & \vdots & & \vdots\\ I^{I2} & I^{I3} & \ldots & I^{II}\end{pmatrix}.$$

Then (2.6) simplifies to

$$\frac{\partial}{\partial\sigma^2}\left\{c(\sigma^2)^{1/2}d(\boldsymbol{\delta},\sigma^2)\right\}=0, \tag{2.7}$$

where c is a constant. Hence the set of solution of (2.7) is of the form

$$d(\boldsymbol{\delta},\sigma^2)=(\sigma^2)^{-1/2}h(\boldsymbol{\delta}),$$

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

$$\pi_m^{(2)}(\gamma,\boldsymbol{\delta},\sigma^2)=(\sigma^2)^{-1}h(\boldsymbol{\delta}).$$

This completes the proof. □

2.2. The probability matching priors : Matching the alternative coverage probabilities

Mukerjee and Reid (1999) studied that a prior satisfying (2.2) matches $P\{\theta_1 + \beta(I^{11}/n)^{1/2} \le \theta_1^{1-\alpha}(\pi; \mathbf{Y})|\boldsymbol{\theta}\}$ 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 $\mathbf{Y}$. 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, a 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) give the simple differential equations that a second order probability matching prior matches alternative coverage probabilities up to the second order. The differential equations are given by

$$\sum_{i=2}^{t}\sum_{j=2}^{t} \frac{\partial}{\partial \theta_i}\left\{L_{11j} I^{ij} I_{11}^{-1/2} d(\theta_2, \ldots, \theta_t)\right\} = 0, \tag{2.8}$$

$$\sum_{i=2}^{t}\sum_{j=2}^{t} \frac{\partial}{\partial \theta_i}\left\{L_{j,11} I^{ij} I_{11}^{-1/2} d(\theta_2, \ldots, \theta_t)\right\} = 0, \tag{2.9}$$

$$\frac{\partial}{\partial \theta_1}\left(I_{11}^{-3/2} L_{111}\right) = 0, \quad \frac{\partial}{\partial \theta_1}\left(I_{11}^{-3/2} L_{1,11}\right) = 0, \tag{2.10}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_t)^T$, θ_1 is the parameter of interest and

$$L_{11j} = E\left(\frac{\partial^3 \log L}{\partial \theta_1^2 \partial \theta_j}\right),$$
$$L_{j,11} = E\left(\frac{\partial \log L}{\partial \theta_j} \cdot \frac{\partial^2 \log L}{\partial \theta_1^2}\right), \quad i = 1, \ldots, t.$$

THEOREM 2.2. *The second order probability matching prior,*

$$\pi_m^{(2)}(\gamma, \boldsymbol{\delta}, \sigma^2) = (\sigma^2)^{-1} h(\boldsymbol{\delta}), \tag{2.11}$$

where $h(\boldsymbol{\delta})$ is any smooth function of $\boldsymbol{\delta}$, matches the alternative coverage probabilities up to the second order.

PROOF. Due to the orthogonality of γ with $\boldsymbol{\delta}$ and σ^2, the differential equations are given by

$$\sum_{i,j=2}^{I} \frac{\partial}{\partial \delta_i}\left\{L_{11j} I^{ij} I_{11}^{-1/2} d(\boldsymbol{\delta},\sigma^2)\right\} + \frac{\partial}{\partial \sigma^2}\left\{L_{11(I+1)} I^{(I+1)(I+1)} I_{11}^{-1/2} d(\boldsymbol{\delta},\sigma^2)\right\} = 0,$$

$$\sum_{i,j=2}^{I} \frac{\partial}{\partial \delta_i}\left\{L_{j,11} I^{ij} I_{11}^{-1/2} d(\boldsymbol{\delta},\sigma^2)\right\} + \frac{\partial}{\partial \sigma^2}\left\{L_{(I+1),11} I^{(I+1)(I+1)} I_{11}^{-1/2} d(\boldsymbol{\delta},\sigma^2)\right\} = 0,$$

$$\frac{\partial}{\partial \gamma}\left(I_{11}^{-3/2} L_{111}\right) = 0, \quad \frac{\partial}{\partial \gamma}\left(I_{11}^{-3/2} L_{1,11}\right) = 0.$$

Since

$$d(\boldsymbol{\delta},\sigma^2) = (\sigma^2)^{-1/2} h(\boldsymbol{\delta}), \quad L_{11j} = 0,\ j = 1,\ldots,I, \quad L_{11(I+1)} = \frac{J}{\sum_{j=1}^I \gamma_j^2}(\sigma^2)^{-2},$$

$$L_{j,11} = 0, \quad j = 1,\ldots,I, \quad L_{(I+1),11} = c(\sigma^2)^{-2}, \quad c = \text{a constant},$$

$$I_{11} = \frac{J}{\sum_{j=1}^I \gamma_j^2}(\sigma^2)^{-1}, \quad I^{(I+1)(I+1)} = \frac{2}{N}(\sigma^2)^2,$$

thus

$$\sum_{i=2}^{I}\sum_{j=2}^{I} \frac{\partial}{\partial \delta_i}\left\{0 \cdot I^{ij} I_{11}^{-1/2} d(\boldsymbol{\delta},\sigma^2)\right\} + \frac{\partial}{\partial \sigma^2}\left\{\frac{2}{N}\left(\frac{J}{\sum_{j=1}^I \gamma_j^2}\right)^{1/2} h(\boldsymbol{\delta})\right\} = 0,$$

$$\sum_{i=2}^{I}\sum_{j=2}^{I} \frac{\partial}{\partial \delta_i}\left\{0 \cdot I^{ij} I_{11}^{-1/2} d(\boldsymbol{\delta},\sigma^2)\right\} + \frac{\partial}{\partial \sigma^2}\left\{\frac{2c}{N}\left(\frac{J}{\sum_{j=1}^I \gamma_j^2}\right)^{-1/2} h(\boldsymbol{\delta})\right\} = 0,$$

$$\frac{\partial}{\partial \gamma}\left(I_{11}^{-3/2} \cdot 0\right) = 0, \quad \frac{\partial}{\partial \gamma}\left(I_{11}^{-3/2} \cdot 0\right) = 0.$$

Therefore the second order matching prior matches the alternative coverage probabilities up to the second order. This completes the proof. □

2.3. HPD matching priors

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. Specifically, if $\tilde{\pi}$ denotes the posterior distribution of θ_1 under a prior π, and $k_\alpha \equiv k_\alpha(\pi, \mathbf{Y})$ is such that

$$P^{\pi}\left\{\tilde{\pi}(\theta_1|\mathbf{Y}) \geq k_\alpha|\mathbf{Y}\right\} = 1 - \alpha + o(n^{-u}). \tag{2.12}$$

then the HPD region for θ_1 with posterior coverage probability $1-\alpha+o(n^{-u})$ is given by

$$H_\alpha(\pi;\mathbf{Y}) = \{\theta_1 : \tilde{\pi}(\theta_1|\mathbf{Y}) \geq k_\alpha\}. \tag{2.13}$$

DiCicco and Stern (1994) and Ghosh and Murkerjee (1995) characterized priors π for which

$$P\{\theta_1 \in H_\alpha(\pi;\mathbf{Y})|\boldsymbol{\theta}\} = 1-\alpha+o(n^{-u}), \tag{2.14}$$

for all $\boldsymbol{\theta}$ and all $\alpha \in (0,1)$. They found necessary and sufficient conditions under which π satisfies (2.14). Due to the orthogonality of γ with $\boldsymbol{\delta}$ and σ^2, from equation (33) of DiCicco and Stern (1994) or equation (4.1) of Ghosh and Mukerjee (1995), a prior π is a HPD matching prior if and only if it satisfies

$$\begin{aligned}&\frac{\partial^2}{\partial\gamma^2}\left(I^{11}\pi\right) - \frac{\partial}{\partial\gamma}\left\{L_{111}(I^{11})^2\pi\right\} - \sum_{v=2}^{I}\frac{\partial}{\partial\delta_v}\left\{\left(\sum_{s=2}^{I}L_{11s}I^{vs}\right)I^{11}\pi\right\}\\ &-\frac{\partial}{\partial\sigma^2}\left\{L_{11(I+1)}I^{(I+1)(I+1)}I^{11}\pi\right\} = 0.\end{aligned} \tag{2.15}$$

Recently, 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 γ. Since $L_{111} = E\left\{(\partial^3\log L)/(\partial\gamma^3)\right\} = 0$, the second order probability matching prior for γ,

$$\pi_m^{(2)}(\gamma,\boldsymbol{\delta},\sigma^2) = (\sigma^2)^{-1}h(\boldsymbol{\delta}), \tag{2.16}$$

where $h(\boldsymbol{\delta})$ is any smooth function of $\boldsymbol{\delta}$, is a HPD matching prior up to the same order.

2.4. 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 $(\gamma,\boldsymbol{\delta},\sigma^2)$.

For the group ordering, the notation $\{\gamma,(\boldsymbol{\delta},\sigma^2)\}$ will be used to represent the case where there are two groups, with γ being the most important and $\boldsymbol{\delta}$ and σ^2 being of equal importance. Similarly, $\{\gamma,\boldsymbol{\delta},\sigma^2\}$ means there are three groups, with γ being most important and σ^2 being least important.

Then due to the orthogonality of the parameters, following Datta and Ghosh (1995), choosing rectangular compacts for each γ, $\boldsymbol{\delta}$ and σ^2 when γ is the parameter of interest, the reference priors are given by as follows.

THEOREM 2.3. *For the one-way model (1.1), if $\gamma = \sum_{i=1}^{I} \gamma_i \mu_i$ is the parameter of interest, then the reference prior distributions for different groups of ordering of $(\gamma, \boldsymbol{\delta}, \sigma^2)$ are:*

Group ordering	*Reference prior*
$\{\ (\gamma, \boldsymbol{\delta}, \sigma^2)\ \}$	$\pi_1 \propto (\sigma^2)^{-(I+2)/2}$
$\{\ \gamma, (\boldsymbol{\delta}, \sigma^2)\ \}$	$\pi_2 \propto (\sigma^2)^{-(I+1)/2}$
$\{\gamma, \boldsymbol{\delta}, \sigma^2\ \}$, $\{\gamma, \sigma^2, \boldsymbol{\delta}\ \}$, $\{(\gamma, \boldsymbol{\delta}), \sigma^2\ \}$	$\pi_3 \propto (\sigma^2)^{-1}$
$\{\ (\gamma, \boldsymbol{\delta}), \sigma^2\}$	$\pi_4 \propto (\sigma^2)^{-3/2}$.

REMARK 1. The reference prior,π_3, satisfies a second order matching criterion. Thus the π_3 is a HPD matching prior and matches the alternative coverage probabilities up to the second order. And it does not depend on I.

We consider a particular second order matching prior where h is a constant in matching priors (2.5). This prior is given by

$$\pi_m^{(2)}(\gamma, \boldsymbol{\delta}, \sigma^2) = (\sigma^2)^{-1}. \tag{2.17}$$

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 (2.17). We consider the class of priors

$$\pi(\gamma, \boldsymbol{\delta}, \sigma^2) \propto (\sigma^2)^{-a}, \tag{3.1}$$

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

THEOREM 3.1. *The posterior distribution of $(\gamma, \boldsymbol{\delta}, \sigma^2)$ under the prior π, (3.1), is proper if $N + 2a - I - 2 > 0$, where $N = IJ$.*

PROOF. Note that the joint posterior for γ, $\boldsymbol{\delta}$ and σ^2 given $\mathbf{y}$ is

$$\pi(\gamma, \boldsymbol{\delta}, \sigma^2 | \mathbf{y}) \propto (\sigma^2)^{-N/2-a} \exp\left\{ -\frac{1}{2\sigma^2} \left(S_1^2 + S_2^2 \right) \right\}. \tag{3.2}$$

where

$$S_1^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - Y_{i\cdot})^2,$$

$$S_2^2 = \left\{ J\left(Y_{1\cdot} - \frac{\gamma_1\gamma + \sum_{j=2}^{I}\gamma_j\delta_j}{\sum_{j=1}^{I}\gamma_j^2}\right)^2 + J\sum_{i=2}^{I}\left(Y_{i\cdot} - \frac{\gamma_i\gamma - \gamma_1\delta_i}{\sum_{j=1}^{I}\gamma_j^2}\right)^2 \right\}.$$

Let

$$\mu_1 = \frac{\gamma_1\gamma + \sum_{j=2}^{I}\gamma_j\delta_j}{\sum_{j=1}^{I}\gamma_j^2}, \quad \mu_i = \frac{\gamma_i\gamma - \gamma_1\delta_i}{\sum_{j=1}^{I}\gamma_j^2}, \quad i = 2, \ldots, I.$$

Then

$$\int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \pi(\gamma, \boldsymbol{\delta}, \sigma^2|\mathbf{y}) d\gamma d\boldsymbol{\delta} d\sigma^2$$
$$= \int_0^\infty \int_{-\infty}^\infty c_1(\sigma^2)^{-N/2-a} \exp\left[-\frac{1}{2\sigma^2}\left\{S_1^2 + J\sum_{i=1}^{I}(Y_{i\cdot} - \mu_i)^2\right\}\right] d\boldsymbol{\mu} d\sigma^2$$
$$= \int_0^\infty c_2\left(\sigma^2\right)^{-(N-I)/2-a} \exp\left(-\frac{1}{2\sigma^2}S_1^2\right) d\sigma^2.$$

Here c_1 and c_2 are a constant and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_I)^T$. Thus the last integral is finite if $N + 2a - I - 2 > 0$. This completes the proof. □

THEOREM 3.2. *Under the prior (3.1), the marginal posterior density of γ is given by*

$$\pi(\gamma|\mathbf{y}) \propto \left[\left(\sum_{j=1}^{I}\gamma_j^2\right)^2 \frac{S_1^2}{J} + \sum_{i=1}^{I}\left\{\left(\sum_{j=1}^{I}\gamma_j^2\right) y_{i\cdot} - \gamma_i\gamma\right\}^2 - \sum_{i=2}^{I}\left(\gamma_i^2 f_i + \gamma_1^2\right)^{-1} g_i^2\right]^{-(N+2a-I-1)/2}, \tag{3.3}$$

where

$$f_i = f_{i-1} - f_{i-1}^2\gamma_{i-1}^2(f_{i-1}\gamma_{i-1}^2 + \gamma_1^2)^{-1}, \quad i \geq 3,$$

with $f_2 = 1$ and

$$g_i = \left(\sum_{j=1}^{I}\gamma_j^2\right)\gamma_i y_{1\cdot} - \left(\sum_{j=1}^{I}\gamma_j^2\right)\gamma_1 y_{i\cdot} - \sum_{j=2}^{i-1} f_j(\gamma_j^2 f_j + \gamma_1^2)^{-1} g_j\gamma_j\gamma_i, \quad i \geq 3,$$

with $g_2 = \left(\sum_{j=1}^{I}\gamma_j^2\right)\gamma_2 y_{1\cdot} - \left(\sum_{j=1}^{I}\gamma_j^2\right)\gamma_1 y_{2\cdot}$.

PROOF. Firstly, we integrate with respect to $\delta_2, \ldots, \delta_I$ from (3.2). Then one gets

$$\pi(\gamma, \sigma^2|\mathbf{y}) \propto (\sigma^2)^{-\{N+2a-(I-1)\}/2} \exp\left(-\frac{J}{2(\sum_{j=1}^{I}\gamma_j^2)^2\sigma^2}\left\{\left(\sum_{j=1}^{I}\gamma_j^2\right)^2\frac{S_1^2}{J} + \sum_{i=1}^{I}\left\{\left(\sum_{j=1}^{I}\gamma_j^2\right)y_{i\cdot} - \gamma_i\gamma\right\}^2 - \sum_{i=2}^{I}(\gamma_i^2 f_i + \gamma_1^2)^{-1}g_{ii}^2\right\}\right), \quad (3.4)$$

where

$$f_i = f_{i-1} - f_{i-1}^2\gamma_{i-1}^2(f_{i-1}\gamma_{i-1}^2 + \gamma_1^2)^{-1}, \quad i \geq 3$$

with $f_2 = 1$ and

$$g_i = \left(\sum_{j=1}^{I}\gamma_j^2\right)\gamma_i y_{1\cdot} - \left(\sum_{j=1}^{I}\gamma_j^2\right)\gamma_1 y_{i\cdot} - \sum_{j=2}^{i-1} f_j(\gamma_j^2 f_j + \gamma_1^2)^{-1}g_j\gamma_j\gamma_i, \quad i \geq 3,$$

with $g_2 = (\sum_{j=1}^{I}\gamma_j^2)\gamma_2 y_{1\cdot} - (\sum_{j=1}^{I}\gamma_j^2)\gamma_1 y_{2\cdot}$. Next, integrating with respect to σ^2, it follows from (3.4) that

$$\pi(\gamma|\mathbf{y}) \propto \left[\left(\sum_{j=1}^{I}\gamma_j^2\right)^2\frac{S_1^2}{J} + \sum_{i=1}^{I}\left\{\left(\sum_{j=1}^{I}\gamma_j^2\right)y_{i\cdot} - \gamma_i\gamma\right\}^2 - \sum_{i=2}^{I}(\gamma_i^2 f_i + \gamma_1^2)^{-1}g_{ii}^2\right]^{-(N+2a-I-1)/2}$$

This completes the proof. □

Next we prove a result which establishes the symmetry and unimodality of the posterior distribution of γ, (3.3), about $\sum_{i=1}^{I}\gamma_i Y_{i\cdot}$.

THEOREM 3.3. *The posterior distribution of γ is symmetric and unimodal about $\sum_{i=1}^{I}\gamma_i Y_{i\cdot}$ under the noninformative prior (3.1).*

PROOF. Firstly, we prove that the posterior distribution of γ is symmetric about $\sum_{i=1}^{I}\gamma_i Y_{i\cdot}$. We will show that the value of posterior density at $\sum_{i=1}^{I}\gamma_i Y_{i\cdot}+c$

is equal to the value of posterior density at $\sum_{i=1}^I \gamma_i Y_{i\cdot} - c$, where c is arbitrary constant. Since from the posterior distribution of γ, (3.3),

$$\pi\left(\sum_{i=1}^I \gamma_i Y_{i\cdot} + c \middle| \mathbf{y}\right) = \left\{\left(\sum_{j=1}^I \gamma_j^2\right)^2 \frac{S_1^2}{J} + c^2 \sum_{j=1}^I \gamma_j^2\right\}^{-(N+2a-I-1)/2}$$

and

$$\pi\left(\sum_{i=1}^I \gamma_i Y_{i\cdot} - c \middle| \mathbf{y}\right) = \left\{\left(\sum_{j=1}^I \gamma_j^2\right)^2 \frac{S_1^2}{J} + c^2 \sum_{j=1}^I \gamma_j^2\right\}^{-(N+2a-I-1)/2},$$

thus

$$\pi\left(\sum_{i=1}^I \gamma_i Y_{i\cdot} + c \middle| \mathbf{y}\right) = \pi\left(\sum_{i=1}^I \gamma_i Y_{i\cdot} - c \middle| \mathbf{y}\right).$$

Therefore the posterior distribution of γ is symmetric about $\sum_{i=1}^I \gamma_i Y_{i\cdot}$ under the noninformative prior (3.1). Next we will prove unimodality. Now

$$\frac{d^2}{d\gamma^2}\pi(\gamma|\mathbf{y})\Big|_{\gamma=\sum_{i=1}^I \gamma_i Y_{i\cdot}} = -(N+2a-I-1)\left\{\left(\sum_{j=1}^I \gamma_j^2\right)^2 \frac{S_1^2}{J}\right\}^{-(N+2a-I+1)/2} < 0.$$

Thus the posterior distribution of γ is unimodal about $\sum_{i=1}^I \gamma_i Y_{i\cdot}$ under the noninformative prior (3.1). □

The normalizing constant for the marginal density of γ requires a one dimensional integration. Therefore we have the marginal posterior density of γ, and so it is easy to compute the marginal moment of γ. For reference prior, π_1, $a = (I+2)/2$ in the above marginal density (3.3). For π_2, $a = (I+1)/2$. For π_3, $a = 1$. For π_4, $a = 3/2$. In Section 4, we investigate the frequentist coverage probabilities for the π_1, π_2, π_3 and π_4, respectively. Also we compute the Bayesian credible intervals using real data.

4. NUMERICAL ANALYSIS

4.1. Simulation study

We evaluate the frequentist coverage probability by investigating the credible interval of the marginal posteriors density of γ under the noninformative

TABLE 4.1 *Frequentist coverage probability of 0.05 (0.95) posterior quantiles of* γ

$\mu_1,\cdots,\mu_5$	$\gamma_1,\cdots,\gamma_5$	$n_1,\cdots,n_5$	π_1	π_2	π_3	π_4
1,1,1,1,1	$\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5}$	3,3,3,3,3	0.0920(0.9142)	0.0842(0.9200)	0.0530(0.9532)	0.0632(0.9448)
		5,5,5,5,5	0.0666(0.9308)	0.0620(0.9366)	0.0438(0.9526)	0.0488(0.9488)
		7,7,7,7,7	0.0680(0.9418)	0.0646(0.9438)	0.0526(0.9540)	0.0552(0.9516)
		10,10,10,10,10	0.0602(0.9390)	0.0586(0.9410)	0.0510(0.9482)	0.0530(0.9462)
	1,1,1,1,-4	3,3,3,3,3	0.0852(0.9098)	0.0770(0.9178)	0.0476(0.9508)	0.0554(0.9408)
		5,5,5,5,5	0.0706(0.9356)	0.0668(0.9400)	0.0502(0.9560)	0.0554(0.9526)
		7,7,7,7,7	0.0670(0.9392)	0.0650(0.9422)	0.0538(0.9520)	0.0564(0.9506)
		10,10,10,10,10	0.0636(0.9444)	0.0614(0.9456)	0.0516(0.9528)	0.0538(0.9516)
	1,0,0,0,-1	3,3,3,3,3	0.0846(0.9098)	0.0766(0.9164)	0.0456(0.9500)	0.0542(0.9410)
		5,5,5,5,5	0.0716(0.9296)	0.0658(0.9338)	0.0486(0.9500)	0.0534(0.9454)
		7,7,7,7,7	0.0648(0.9378)	0.0620(0.9404)	0.0492(0.9486)	0.0520(0.9464)
		10,10,10,10,10	0.0596(0.9376)	0.0580(0.9412)	0.0484(0.9502)	0.0516(0.9480)
1,1,2,2,4	$\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5}$	3,3,3,3,3	0.0920(0.9142)	0.0842(0.9200)	0.0530(0.9532)	0.0632(0.9448)
		5,5,5,5,5	0.0666(0.9308)	0.0620(0.9366)	0.0438(0.9526)	0.0488(0.9488)
		7,7,7,7,7	0.0680(0.9418)	0.0646(0.9438)	0.0526(0.9540)	0.0552(0.9516)
		10,10,10,10,10	0.0602(0.9390)	0.0586(0.9410)	0.0510(0.9482)	0.0530(0.9462)
	2,2,1,1,-2	3,3,3,3,3	0.0870(0.9106)	0.0804(0.9180)	0.0476(0.9484)	0.0564(0.9394)
		5,5,5,5,5	0.0724(0.9300)	0.0678(0.9342)	0.0492(0.9532)	0.0540(0.9480)
		7,7,7,7,7	0.0610(0.9366)	0.0586(0.9404)	0.0496(0.9492)	0.0524(0.9470)
		10,10,10,10,10	0.0624(0.9420)	0.0590(0.9438)	0.0522(0.9502)	0.0540(0.9482)
	1,-1,0,0,0	3,3,3,3,3	0.0910(0.9116)	0.0834(0.9176)	0.0476(0.9526)	0.0574(0.9420)
		5,5,5,5,5	0.0760(0.9274)	0.0714(0.9312)	0.0516(0.9480)	0.0560(0.9436)
		7,7,7,7,7	0.0644(0.9330)	0.0620(0.9346)	0.0496(0.9458)	0.0530(0.9428)
		10,10,10,10,10	0.0620(0.9442)	0.0604(0.9470)	0.0496(0.9489)	0.0519(0.9471)

prior π given in (3.1) for several configurations $(\mu_1,\ldots,\mu_5,\sigma^2)$, $(\gamma_1,\ldots,\gamma_5)$ and $n_1,\ldots,n_5$ under 5 populations. 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 4.1 gives 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 $(\mu_1,\ldots,\mu_5,\sigma^2)$ and any prespecified probability value α. Here α is 0.05 (0.95). Let $\gamma^\pi(\alpha|\mathbf{Y_1},\ldots,\mathbf{Y_5})$ be the posterior α-quantile of γ given $(\mathbf{Y_1},\ldots,\mathbf{Y_5})$. That is to say, $F(\gamma^\pi(\alpha|\mathbf{Y_1},\ldots,\mathbf{Y_5})|\mathbf{Y_1},\ldots,\mathbf{Y_5})=\alpha$, where $F(\cdot|\mathbf{Y_1},\ldots,\mathbf{Y_5})$ is the marginal posterior distribution of γ. Then the frequentist coverage probability of this one sided credible interval of γ is

$$P_{(\mu_1,\ldots,\mu_5,\sigma^2)}(\alpha;\gamma)=P_{(\mu_1,\ldots,\mu_5,\sigma^2)}\left\{0<\theta_1\le\gamma^\pi(\alpha|\mathbf{Y_1},\ldots,\mathbf{Y_5})\right\}. \quad (4.1)$$

The estimated $P_{(\mu_1,\ldots,\mu_5,\sigma^2)}(\alpha;\gamma)$ when $\alpha=0.05(0.95)$ is shown in Table 4.1.

In particular, for fixed $(\mu_1,\ldots,\mu_5,\sigma^2)$, and $n_1\ldots,n_5$, we take 5,000 independent random samples of $(\mathbf{Y_1},\ldots,\mathbf{Y_5})$ from the model (1.1). Our simulation, we take $\sigma^2=1$ without loss of generality. Note that under the prior π, for fixed $(\mathbf{Y_1},\ldots,\mathbf{Y_5})$, $\gamma\le\gamma^\pi(\alpha|\mathbf{Y_1},\ldots,\mathbf{Y_5})$ if and only if $F\{\gamma^\pi(\alpha|\mathbf{Y_1},\ldots,\mathbf{Y_5})$ $|\mathbf{Y_1},\ldots,\mathbf{Y_5}\}\le\alpha$. Under the prior π, $P_{(\mu_1,\ldots,\mu_5,\sigma^2)}(\alpha;\gamma)$ can be estimated by the relative frequency of $F(\gamma^\pi|\mathbf{Y_1},\ldots,\mathbf{Y_5})\le\alpha$. For the cases presented in Table

TABLE 4.2 *Data (in lb/in^2) from the tensile strength experiment from Mongomery (1976)*

	Observation				
Cotton Weight Percent	1	2	3	4	5
15	7	7	15	11	9
20	12	17	12	18	18
25	14	18	18	19	19
30	19	25	22	19	23
35	7	10	11	15	11

4.1, we see that one-at-a-time reference prior satisfying the second order matching criterion meet very well the target coverage probabilities. Note that the results of tables are not much sensitive to the change of the values of $(\mu_1, \ldots, \mu_5)$ and $(\gamma_1, \cdots, \gamma_5)$. Thus we recommend to use the one-at-a-time reference prior when using the second order matching criterion.

4.2. Example

We provide the calculations on an example from Mongomery (1976). A product development engineer is interested in investigating the tensile strength of a new synthetic fiber that will be used to make cloth for men's shirts. The engineer knows from previous experience that the strength is affected by the weight percent of cotton used in the blend of materials for the fiber. He has run a completely randomized experiment with five levels of cotton weight percent and five replicates. The data is provided in Table 4.2.

The sample means based upon $n_i = 5$ observations for each of the five treatments are

$$Y_{1\cdot} = 9.8,\ Y_{2\cdot} = 15.4,\ Y_{3\cdot} = 17.6,\ Y_{4\cdot} = 21.6,\ Y_{5\cdot} = 10.8.$$

In this example, it is of interest to determine whether the orthogonal contrasts $\gamma^1 = \mu_4 - \mu_5$, $\gamma^2 = \mu_1 + \mu_3 - \mu_4 - \mu_5$, $\gamma^3 = \mu_1 - \mu_3$ and $\gamma^4 = \mu_1 - 4\mu_2 + \mu_3 + \mu_4 + \mu_5$ are zero. In Theorem 3.3, we showed that the posterior of γ is symmetric and unimodal at $\sum_{i=1}^{I} \gamma_i Y_{i\cdot}$. Thus the Bayes estimator for γ is $\sum_{i=1}^{I} \gamma_i Y_{i\cdot}$. Therefore the Bayes estimate for $\gamma^1, \gamma^2, \gamma^3$ and γ^4 are $10.8, -5.0, -7.8$ and -1.8, respectively. Next, the Bayesian credible intervals and the Scheffé confidence intervals for $\gamma^1, \gamma^2, \gamma^3$ and γ^4 are given in Table 4.3. We compute the 95% Bayesian credible interval from the posterior distribution (3.3) for the reference priors, respectively.

TABLE 4.3 *95% Bayesian credible intervals for the reference priors and the Scheffé 95% confidence interval*

γ	π_1	π_2	π_3	π_4	*Scheffé*
γ^1	(7.4924,14.1076)	(7.4171,14.1829)	(7.0545,14.5455)	(7.1559,14.4441)	(4.7163,16.8837)
γ^2	(-9.6776,-0.3224)	(-9.7842,-0.2158)	(-10.2969,0.2969)	(-10.1535,0.1535)	(-13.6037,3.6037)
γ^3	(-11.1076,-4.4924)	(-11.1829,-4.4171)	(-11.5455,-4.0545)	(-11.4441,-4.1559)	(-13.8837,-1.7163)
γ^4	(-12.2595,8.6595)	(-12.4978,8.8978)	(-13.6442,10.0442)	(-13.3235,9.7235)	(-21.0384,17.4384)

The inferences can be made that $\gamma^1 \neq 0$ and $\gamma^3 \neq 0$ since these credible intervals do not contain zero. This results agree with the results of Scheffé. Note that for γ^2, the Bayesian credible interval using the reference prior π_3 is more reasonable answer than Scheffé.

From the results of Bayesian credible interval in Table 4.3, there is a little difference in interval between the reference priors. The length of interval for the one-at-a-time reference prior π_3 is slightly longer than that of the other reference priors. This fact agree with our simulation study in Section 4.1. Also the length of the Bayesian credible intervals are less than the length of the Scheffé confidence interval, respectively.

Consequently the the one-at-a-time reference prior satisfying the second order matching criterion seems to give better results than the other reference priors for γ in the sense of asymptotic frequentist coverage property and length of confidence interval.

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)", In *Bayesian Statistics IV* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.), 35–60, Oxford University Press, Oxford.

BERNARDO, J. M. (1979). "Reference posterior distributions for Bayesian inference (with discussion)", *Journal of Royal Statistical Society*, **B41**, 113–147.

COX, D. R. AND REID, N. (1987). "Parameter orthogonality and approximate conditional inference (with discussion)", *Journal of Royal Statistical Society*, **B49**, 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 Annals 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 likelihoods", *Journal of the Royal Statistical Society*, **B56**, 397–408.

GHOSH, J. K. AND MUKERJEE, R. (1992). "Noninformative priors (with discussion)", In *Bayesian Statistics IV* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.), 195–210, Oxford University Press, Oxford.

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.

KIM, D. H., KANG, S. G. AND LEE, W. D. (2004). "Noninformative priors for the nested eesign", *Journal of Statistical Planning and Inference*, To Appear.

LI, H. AND STERN, H. S. (1997). "Bayesian inference for nested designs based on Jeffreys's Prior", *The American Statistician*, **51**, 219–224.

MONTGOMERY, D. C. (1976). *Design and Analysis of Experiments*, John Wiley and Sons, New York.

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.

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. L. AND PEERS, H. W. (1963). "On formulae for confidence points based on integrals of weighted likelihoods", *Journal of Royal Statistical Society*, **B25**, 318–329.