

BAYESIAN TEST FOR THE EQUALITY OF THE MEANS AND VARIANCES OF THE TWO NORMAL POPULATIONS WITH VARIANCES RELATED TO THE MEANS USING NONINFORMATIVE PRIORS

DAL HO KIM¹, SANG GIL KANG² AND WOO DONG LEE³

ABSTRACT

In this paper, when the variance of the normal distribution is related to the mean, we develop noninformative priors such as matching priors and reference priors. We prove that the second order matching prior matches alternative coverage probabilities up to the same order and also it is a HPD matching prior. It turns out that one-at-a-time reference prior satisfies a second order matching criterion. Then using these noninformative priors, we develop a Bayesian test procedure for the equality of the means and variances of two independent normal distributions using fractional Bayes factor. Some simulation study is performed, and a real data example is also provided.

AMS 2000 subject classifications. Primary 62F15; Secondary 62G10.

Keywords. Matching priors, alternative coverage probabilities, HPD matching, reference priors, fractional Bayes factor, variance and mean relationships.

1. INTRODUCTION

Suppose that the observations are normally distributed with mean μ and variance $(c + \mu)^k \sigma^2$. That is,

$$f(x) = \frac{1}{\sqrt{2\pi(c + \mu)^k \sigma^2}} \exp \left\{ -\frac{1}{2(c + \mu)^k \sigma^2} (x - \mu)^2 \right\}, \quad (1.1)$$

Received July 2002; accepted May 2003.

¹Department of Statistics, Kyungpook National University, Taegu 702-701, Korea (e-mail : dalkim@knu.ac.kr)

²Department of Applied Statistics, Sangji University, Wonju 220-702, Korea

³Faculty of Information Science, Kyungsan University, Kyungsan 712-240, Korea

where $c + \mu > 0$. Here the constant c (commonly 1 or 0) and the exponent k are known. Let $\theta_1 = c + \mu$ be the parameter of interest.

The present paper focuses on noninformative priors for θ_1 and a Bayesian test procedure for the equality of the means and variances of two normal populations. We consider Bayesian priors such that the resulting credible intervals for θ_1 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), DiCiccio and Stern (1994), Datta and Ghosh (1995), Datta and Ghosh (1995, 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.

Variance proportional to mean situations are indicated whenever coefficients of variation are approximately constant, as referenced for human stature, for example, in Snedecor and Cochran (1980) and apparent in the pheasant weight data in Steel and Torrie (1980). A nonbiological example is the experimentally determined result (Cox and Roseberry, 1966), that the variance of the average number in sequential probability ratio tests increases with the square of the average sample number. When the observations are normally distributed with variances that are related to the means, Cox (1985) gave the procedure for estimating the ratio of the means of two populations. The other estimating procedures for the ratio of the means, for example Fieller (1944) and Elston (1969), are not used relationship of the variance and the mean. Cox (1985) showed that his procedure produced the shorter confidence length than the other classical procedures. But his method is required to determine circumstances giving real roots for $k > 2$ in model (1.1).

Almost all the work mentioned above is the analysis based on the frequentist point of view, but we consider this problem from the viewpoint of Bayesian framework. There seems to be a necessity to develop objective Bayesian priors for dealing this problem. And we want to develop the Bayesian test procedure for

testing the equality of two independent normal populations using Bayes factor. We calculate the posterior probabilities of the hypotheses using the fractional Bayes factor of O'Hagan (1995) based on noninformative priors.

The outline of the remaining sections is as follows. In Section 2, we develop first order and second order probability matching priors for θ_1 . We reveal that the second order matching prior matches the alternative coverage probabilities up to the same order and is also a HPD matching prior up to the same order. Also we derive the reference priors for the parameters. It turns out that the one-at-a-time reference prior satisfies a second order matching criterion. We provide the propriety of the posterior distribution for a general class of prior distributions which includes the reference priors as well as second order matching priors. In Section 3, using the developed prior, we provide the Bayesian test procedure based on the fractional Bayes factor for the testing ratio of the means of two populations. In Section 4, simulated frequentist coverage probabilities under the proposed priors are investigated. A real example is also given.

2. THE NONINFORMATIVE PRIORS

2.1. The probability matching priors

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

$$P^\pi \{ \theta_1 \leq \theta_1^{1-\alpha}(\pi; \mathbf{X}) | \mathbf{X} \} = 1 - \alpha, \quad (2.1)$$

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

$$P \{ \theta_1 \leq \theta_1^{1-\alpha}(\pi; \mathbf{X}) | \boldsymbol{\theta} \} = 1 - \alpha + o(n^{-u}), \quad (2.2)$$

for some $u > 0$, as 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

$$\theta_1 = c + \mu, \quad \theta_2 = \sigma^2(c + \mu)^k.$$

With this parametrization, the likelihood function of parameters (θ_1, θ_2) for the

model (1.1) is given by

$$L(\theta_1, \theta_2) \propto \theta_2^{-1/2} \exp \left\{ -\frac{(x - \theta_1 + c)^2}{2\theta_2} \right\}. \quad (2.3)$$

Based on (2.3), the Fisher information matrix is given by

$$\mathbf{I} = (I_{ij}) = \begin{pmatrix} \theta_2^{-1} & 0 \\ 0 & \frac{1}{2}\theta_2^{-2} \end{pmatrix}.$$

For notational convenience, we denote $\mathbf{I}^{-1} = (I^{ij})$.

From the above Fisher information matrix \mathbf{I} , θ_1 is orthogonal to θ_2 in the sense of Cox and Reid (1987). Following Tibshirani (1989), the class of first order probability matching prior for θ_1 is characterized by

$$\pi_m^{(1)}(\theta_1, \theta_2) \propto \theta_2^{-1/2} d(\theta_2), \quad (2.4)$$

where $d(\theta_2)$ is an arbitrary positive 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 for θ_1 are given by*

$$\pi_m^{(2)}(\theta_1, \theta_2) = \theta_2^{-1}. \quad (2.5)$$

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

$$\frac{1}{6}d(\theta_2) \frac{\partial}{\partial \theta_1} \left\{ I_{11}^{-3/2} L_{1,1,1} \right\} + \frac{\partial}{\partial \theta_2} \left\{ I_{11}^{-1/2} L_{112} I^{22} d(\theta_2) \right\} = 0, \quad (2.6)$$

where

$$\begin{aligned} I_{11} &= \theta_2^{-1}, \quad I^{22} = 2\theta_2^2, \quad L_{1,1,1} = E \left\{ \frac{\partial \log L}{\partial \theta_1} \right\}^3 = 0, \\ L_{112} &= E \left\{ \frac{\partial^3 \log L}{\partial \theta_1^2 \partial \theta_2} \right\} = \theta_2^{-2}. \end{aligned}$$

Then (2.6) simplifies to

$$\frac{\partial}{\partial \theta_2} \left\{ 2\theta_2^{1/2} d(\theta_2) \right\} = 0. \quad (2.7)$$

Hence the set of solution of (2.7) is of the form $d(\theta_2) = \theta_2^{-1/2}$. Thus the resulting second order probability matching prior is

$$\pi_m^{(2)}(\theta_1, \theta_2) = \theta_2^{-1}.$$

This completes the proof. \square

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

Mukerjee and Reid (1999) studied that a prior satisfying (2.2) matches $P\{\theta_1 + \beta(I^{11}/n)^{1/2} \leq \theta_1^{1-\alpha}(\pi; \mathbf{X})|\boldsymbol{\theta}\}$ 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{X} . 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) gives 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

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

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

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

where

$$L_{11j} = E \left\{ \frac{\partial^3 \log L}{\partial \theta_1^2 \partial \theta_j} \right\},$$

$$L_{j,11} = E \left\{ \frac{\partial \log L}{\partial \theta_j} \cdot \frac{\partial^2 \log L}{\partial \theta_1^2} \right\}, \quad j = 1, \dots, t, \quad \boldsymbol{\theta} = (\theta_1, \dots, \theta_t)^T$$

and θ_1 is the parameter of interest.

THEOREM 2.2. *The second order probability matching prior for θ_1 ,*

$$\pi_m^{(2)}(\theta_1, \theta_2) = \theta_2^{-1}, \quad (2.11)$$

matches the alternative coverage probabilities up to the second order.

PROOF. Due to the orthogonality of θ_1 with θ_2 , the differential equations are given by

$$\begin{aligned}\frac{\partial}{\partial \theta_2} \left\{ L_{112} I^{22} I_{11}^{-1/2} d(\theta_2) \right\} &= 0, \\ \frac{\partial}{\partial \theta_2} \left\{ L_{2,11} I^{22} I_{11}^{-1/2} d(\theta_2) \right\} &= 0, \\ \frac{\partial}{\partial \theta_1} \left\{ I_{11}^{-3/2} L_{111} \right\} &= 0, \quad \frac{\partial}{\partial \theta_1} \left\{ I_{11}^{-3/2} L_{1,11} \right\} = 0.\end{aligned}$$

Since

$$\begin{aligned}d(\theta_2) &= \theta_2^{-1/2}, \quad I_{11} = \theta_2^{-1}, \quad I^{22} = 2\theta_2^2, \\ L_{111} &= 0, \quad L_{112} = \theta_2^{-2}, \quad L_{1,11} = 0, \quad L_{2,11} = 0,\end{aligned}$$

we obtain

$$\begin{aligned}\frac{\partial}{\partial \theta_2} \left\{ \theta_2^{-2} \cdot 2\theta_2^2 \cdot \theta_2^{1/2} \cdot \theta_2^{-1/2} \right\} &= 0, \\ \frac{\partial}{\partial \theta_2} \left\{ 0 \cdot 2\theta_2^2 \cdot \theta_2^{1/2} \cdot \theta_2^{-1/2} \right\} &= 0, \\ \frac{\partial}{\partial \theta_1} \left\{ \theta_2^{3/2} \cdot 0 \right\} &= 0, \quad \frac{\partial}{\partial \theta_1} \left\{ \theta_2^{3/2} \cdot 0 \right\} = 0.\end{aligned}$$

Hence the second order matching prior matches the alternative coverage probabilities up to the second order. This completes the proof. \square

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 $k_\alpha \equiv k_\alpha(\pi; \mathbf{X})$ is such that

$$P^\pi \{ \tilde{\pi}(\theta_1 | \mathbf{X}) \geq k_\alpha | \mathbf{X} \} = 1 - \alpha + o(n^{-u}), \quad (2.12)$$

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

$$H_\alpha(\pi; \mathbf{X}) = \{ \theta_1 | \tilde{\pi}(\theta_1 | \mathbf{X}) \geq k_\alpha \}. \quad (2.13)$$

DiCiccio and Stern (1994) and Ghosh and Mukerjee (1995) characterized priors π for which

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

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 θ_1 with θ_2 , from equation (33) of DiCiccio 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

$$\frac{\partial^2}{\partial \theta_1^2} \{I^{11} \pi\} - \frac{\partial}{\partial \theta_1} \{L_{111} (I^{11})^2 \pi\} - \frac{\partial}{\partial \theta_2} \{L_{112} I^{22} I^{11} \pi\} = 0. \quad (2.15)$$

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 $L_{111} = 0$, the second order probability matching prior for θ_1 ,

$$\pi_m^{(2)}(\theta_1, \theta_2) = \theta_2^{-1}, \quad (2.16)$$

is a HPD matching prior up to the same order.

2.4. The reference priors

In this section, we derive the reference priors for different groups of ordering of (θ_1, θ_2) . Due to the orthogonality of the parameters, by following Datta and Ghosh (1995) and choosing rectangular compacts for each θ_1 and θ_2 the reference priors are obtained as follows.

THEOREM 2.3. *If θ_1 or θ_2 is the parameter of interest, then the reference prior distributions for different groups of ordering of (θ_1, θ_2) are:*

Group ordering	Reference prior
$\{(\theta_1, \theta_2)\}$	$\pi_1^R(\theta_1, \theta_2) \propto \theta_2^{-3/2}$
$\{\theta_1, \theta_2\}, \{\theta_2, \theta_1\}$	$\pi_2^R(\theta_1, \theta_2) \propto \theta_2^{-1}.$

REMARK 1. One-at-a-time reference prior π_2^R is a second order matching prior as well as a HPD matching prior, and matches the alternative coverage probabilities up to the second order.

2.5. 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. Consider the class of priors

$$\pi(\theta_1, \theta_2) \propto \theta_2^{-a}, \quad (2.17)$$

where $a > 0$. Using the general priors (2.17), we will prove the propriety of posterior distribution in the following theorem.

THEOREM 2.4. *The posterior distribution of (θ_1, θ_2) under the prior (2.17) is proper if $n + 2a - 3 > 0$.*

PROOF. Note that the joint posterior for θ_1 and θ_2 given \mathbf{x} is

$$\pi(\theta_1, \theta_2 | \mathbf{x}) \propto \theta_2^{-n/2-a} \exp \left\{ -\frac{1}{2\theta_2} \sum_{i=1}^n (x_i - \theta_1 + c)^2 \right\}. \quad (2.18)$$

First, we integrate with respect to θ_2 the right hand side of (2.18). If $n+2a-2 > 0$, then

$$\pi(\theta_1 | \mathbf{x}) \propto \left\{ \sum_{i=1}^n (x_i - \theta_1 + c)^2 \right\}^{-(n+2a-2)/2}.$$

Thus if $n + 2a - 3 > 0$, then

$$\begin{aligned} & \int_0^\infty \pi(\theta_1 | \mathbf{x}) d\theta_1 \\ &= \int_0^\infty \left\{ \sum_{i=1}^n (x_i - \theta_1 + c)^2 \right\}^{-(n+2a-2)/2} d\theta_1 \\ &< \left\{ \sum_{i=1}^n (x_i - \bar{x})^2 \right\}^{-(n+2a-2)/2} \int_{-\infty}^\infty \left\{ 1 + \frac{(w - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2 / n} \right\}^{-(n+2a-2)/2} dw \\ &< \infty. \end{aligned}$$

This completes the proof. □

THEOREM 2.5. *Under the prior (2.17), the marginal posterior density of θ_1 is given by*

$$\pi(\theta_1 | \mathbf{x}) \propto \left\{ \sum_{i=1}^n (x_i - \theta_1 + c)^2 \right\}^{-(n+2a-2)/2}. \quad (2.19)$$

The normalizing constant for the marginal density of θ_1 requires a one dimensional integration. Therefore we have the marginal posterior density of θ_1 , and so it is easy to compute the marginal moment of θ_1 . For reference prior, π_1 , $a = 3/2$ in the above marginal density (2.19). For π_2 , $a = 1$. In Section 4, we investigate the frequentist coverage probabilities for the π_1 and π_2 respectively.

3. BAYESIAN TEST USING FRACTIONAL BAYES FACTOR

3.1. Preliminaries

Models (or Hypotheses), H_1, H_2, \dots, H_q are under consideration, with the data $\mathbf{x} = (x_1, x_2, \dots, x_n)$ having probability density function $f_i(\mathbf{x}|\boldsymbol{\theta}_i)$ under model H_i , $i = 1, 2, \dots, q$. The parameter vectors $\boldsymbol{\theta}_i$ are unknown. Let $\pi_i(\boldsymbol{\theta}_i)$ be the prior distribution of model H_i , and let p_i be the prior probabilities of model H_i , $i = 1, 2, \dots, q$. Then the posterior probability that the model H_i is true is

$$P(H_i|\mathbf{x}) = \left\{ \sum_{j=1}^q \frac{p_j}{p_i} \cdot B_{ji} \right\}^{-1}, \quad (3.1)$$

where B_{ji} is the Bayes factor of model H_j to model H_i defined by

$$B_{ji} = \frac{m_j(\mathbf{x})}{m_i(\mathbf{x})} = \frac{\int f_j(\mathbf{x}|\boldsymbol{\theta}_j)\pi_j(\boldsymbol{\theta}_j)d\boldsymbol{\theta}_j}{\int f_i(\mathbf{x}|\boldsymbol{\theta}_i)\pi_i(\boldsymbol{\theta}_i)d\boldsymbol{\theta}_i}. \quad (3.2)$$

The B_{ji} is interpreted as the comparative support of the data for the model j to i . The computation of B_{ji} needs specification of the prior distribution $\pi_i(\boldsymbol{\theta}_i)$ and $\pi_j(\boldsymbol{\theta}_j)$. Usually, one can use the noninformative prior, often improper, such as uniform, Jeffreys, reference or probability matching priors. Denote it as π_i^N . The use of improper priors $\pi_i^N(\cdot)$ in (3.2) causes the B_{ji} to contain unspecified constants. To solve this problem, O'Hagan (1995) proposed the fractional Bayes factor for Bayesian testing and model selection problem as follow.

When the $\pi_i^N(\boldsymbol{\theta}_i)$ is noninformative prior under H_i , equation (3.2) becomes

$$B_{ji}^N(\mathbf{x}) = \frac{\int f_j(\mathbf{x}|\boldsymbol{\theta}_j)\pi_j^N(\boldsymbol{\theta}_j)d\boldsymbol{\theta}_j}{\int f_i(\mathbf{x}|\boldsymbol{\theta}_i)\pi_i^N(\boldsymbol{\theta}_i)d\boldsymbol{\theta}_i}. \quad (3.3)$$

Then the fraction Bayes factor (FBF) of model H_j vs. model H_i is

$$B_{ji}^F = \frac{q_j(b, \mathbf{x})}{q_i(b, \mathbf{x})}, \quad (3.4)$$

where

$$q_i(b, \mathbf{x}) = \frac{\int f_i(\mathbf{x}|\boldsymbol{\theta}_i) \pi_i^N(\boldsymbol{\theta}_i) d\boldsymbol{\theta}_i}{\int f_i^b(\mathbf{x}|\boldsymbol{\theta}_i) \pi_i^N(\boldsymbol{\theta}_i) d\boldsymbol{\theta}_i},$$

and $f_i(\mathbf{x}|\boldsymbol{\theta}_i)$ is the likelihood function and b specifies a fraction of the likelihood which is to be used as a prior density. He proposed three ways for the choice of the fraction b . One frequently suggested choice is $b = m/n$, where m is the size of the minimal training sample, assuming this is well defined (see O'Hagan, 1995 and the discussion by Berger and Mortera of O'Hagan, 1995).

3.2. Bayesian test

Suppose that $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ is a set of independent random sample, where $\mathbf{X}_i = (X_{i1}, \dots, X_{in_i})$ is a sample from a normal distribution with mean μ_i and variance $(c + \mu_i)^k \sigma^2$, $c + \mu_i > 0$, $i = 1, 2$. Here the constant c and the exponent k are known. We want to test the hypotheses

$$H_1 : \frac{c + \mu_1}{c + \mu_2} = 1 \quad \text{vs.} \quad H_2 : \frac{c + \mu_1}{c + \mu_2} \neq 1.$$

The hypothesis H_1 indicates the equal mean. Our interest is to develop a Bayesian test for H_1 vs. H_2 which is an alternative to the classical tests.

Under the hypothesis H_1 , one-at-a-time reference prior for $\mu (\equiv \mu_1 = \mu_2)$ and σ^2 is

$$\pi_1(\mu, \sigma^2) = \frac{1}{\sigma^2}, \quad \sigma^2 > 0.$$

And the likelihood function is

$$L(\mu, \sigma^2 | \mathbf{x}) = \left\{ \frac{1}{2\pi(c + \mu)^k \sigma^2} \right\}^{(n_1 + n_2)/2} \exp \left\{ -\frac{1}{2(c + \mu)^k \sigma^2} \sum_{i=1}^2 \sum_{j=1}^{n_i} (x_{ij} - \mu)^2 \right\}.$$

Then the element of FBF under H_1 is given by

$$\begin{aligned} & \int_0^\infty \int_{-c}^\infty L^b(\mu, \sigma^2 | \mathbf{x}) \pi_1(\mu, \sigma^2) d\mu d\sigma^2 \\ &= (\pi b)^{-b(n_1 + n_2)/2} \Gamma\left(\frac{b(n_1 + n_2)}{2}\right) \int_0^\infty \left\{ \sum_{i=1}^2 \sum_{j=1}^{n_i} (x_{ij} + c - \theta_1)^2 \right\}^{-b(n_1 + n_2)/2} d\theta_1, \end{aligned}$$

where $\theta_1 = c + \mu$ and $\theta_2 = (c + \mu)^k \sigma^2$. Let

$$S(\mathbf{x}) = \int_0^\infty \left\{ \sum_{i=1}^2 \sum_{j=1}^{n_i} (x_{ij} + c - \theta_1)^2 \right\}^{-(n_1+n_2)/2} d\theta_1,$$

$$S^b(\mathbf{x}) = \int_0^\infty \left\{ \sum_{i=1}^2 \sum_{j=1}^{n_i} (x_{ij} + c - \theta_1)^2 \right\}^{-b(n_1+n_2)/2} d\theta_1.$$

Then

$$q_1(b, \mathbf{x}) = \left\{ \pi^{-(n_1+n_2)/2} \Gamma\left(\frac{n_1+n_2}{2}\right) S(x) \right\} \\ \times \left\{ \pi^{-b(n_1+n_2)/2} b^{-b(n_1+n_2)/2} \Gamma\left(\frac{b(n_1+n_2)}{2}\right) S^b(x) \right\}^{-1}. \quad (3.5)$$

For the H_2 , the independent one-at-a-time reference prior for (μ_1, μ_2, σ^2) for the two-sample case is

$$\pi_2(\mu_1, \mu_2, \sigma^2) = \left(\frac{1}{\sigma^2}\right)^2, \quad \sigma^2 > 0.$$

The likelihood function is

$$L(\mu_1, \mu_2, \sigma^2 | \mathbf{x}) \\ = \{2\pi(c + \mu_1)^k \sigma^2\}^{-n_1/2} \exp\left\{-\frac{1}{2(c + \mu_1)^k \sigma^2} \sum_{j=1}^{n_1} (x_{1j} - \mu_1)^2\right\} \\ \times \{2\pi(c + \mu_2)^k \sigma^2\}^{-n_2/2} \exp\left\{-\frac{1}{2(c + \mu_2)^k \sigma^2} \sum_{j=1}^{n_2} (x_{2j} - \mu_2)^2\right\}.$$

Let $\nu_1 = c + \mu_1$ and $\nu_2 = c + \mu_2$. Then the element of FBF under H_2 gives as follows:

$$\int_0^\infty \int_{-c}^\infty \int_{-c}^\infty L^b(\mu_1, \mu_2, \sigma^2 | \mathbf{x}) \pi_2(\mu_1, \mu_2, \sigma^2) d\mu_1 d\mu_2 d\sigma^2 \\ = 2\pi^{-b(n_1+n_2)/2} b^{-1-b(n_1+n_2)/2} \Gamma\left(\frac{b(n_1+n_2)+2}{2}\right) \int_0^\infty \int_0^\infty \nu_1^{-bkn_1/2} \nu_2^{-bkn_2/2} \\ \times \left\{ \frac{\sum_{j=1}^{n_1} (x_{1j} + c - \nu_1)^2}{\nu_1^k} + \frac{\sum_{j=1}^{n_2} (x_{2j} + c - \nu_2)^2}{\nu_2^k} \right\}^{-1-b(n_1+n_2)/2} d\nu_1 d\nu_2.$$

Put

$$T(\mathbf{x}) = \int_0^\infty \int_0^\infty \nu_1^{-kn_1/2} \nu_2^{-kn_2/2} \\ \times \left\{ \frac{\sum_{j=1}^{n_1} (x_{1j} + c - \nu_1)^2}{\nu_1^k} + \frac{\sum_{j=1}^{n_2} (x_{2j} + c - \nu_2)^2}{\nu_2^k} \right\}^{-(n_1+n_2+2)/2} d\nu_1 d\nu_2,$$

$$T^b(\mathbf{x}) = \int_0^\infty \int_0^\infty \nu_1^{-bkn_1/2} \nu_2^{-bkn_2/2} \left\{ \frac{\sum_{j=1}^{n_1} (x_{1j} + c - \nu_1)^2}{\nu_1^k} + \frac{\sum_{j=1}^{n_2} (x_{2j} + c - \nu_2)^2}{\nu_2^k} \right\}^{-1-b(n_1+n_2)/2} d\nu_1 d\nu_2.$$

Then

$$q_2(b, \mathbf{x}) = \left\{ \pi^{-(n_1+n_2)/2} \Gamma\left(\frac{n_1+n_2+2}{2}\right) T(x) \right\} \times \left\{ \pi^{-b(n_1+n_2)/2} b^{-1-b(n_1+n_2)/2} \Gamma\left(\frac{b(n_1+n_2)+2}{2}\right) T^b(x) \right\}^{-1}. \quad (3.6)$$

Therefore the FBF of H_2 vs. H_1 is given by

$$B_{21}^F(\mathbf{x}) = \frac{T(\mathbf{x})S^b(\mathbf{x})}{T^b(\mathbf{x})S(\mathbf{x})}. \quad (3.7)$$

4. NUMERICAL STUDIES

4.1. Simulation study

We evaluate the frequentist coverage probability by investigating the credible interval of the marginal posterior density of θ_1 under the reference priors given in Theorem 2.3 for several configurations n , c , k , σ^2 , μ_1 and μ_2 .

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 probabilities of 0.05 (0.95) posterior quantiles for the our prior. The computation of these numerical values is based on the following algorithm for any fixed true (θ_1, θ_2) and any prespecified probability value α . Here α is 0.05 (0.95). Let $\theta_1^\alpha(\pi|\mathbf{X})$ be the posterior α -quantile of θ_1 given \mathbf{X} . That is, $F(\theta_1^\alpha(\pi|\mathbf{X})|\mathbf{X}) = \alpha$, where $F(\cdot|\mathbf{X})$ is the marginal posterior distribution of θ_1 . Then the frequentist coverage probability of this one sided credible interval of θ_1 is

$$P_{(\theta_1, \theta_2)}(\alpha; \theta_1) = P_{(\theta_1, \theta_2)}\{0 < \theta_1 \leq \theta_1^\alpha(\pi|\mathbf{X})\}. \quad (4.1)$$

The estimated $P_{(\theta_1, \theta_2)}(\alpha; \theta_1)$ when $\alpha = 0.05$ (0.95) is shown in Table 4.1. In particular, for fixed n and (θ_1, θ_2) , we take 10,000 independent random samples of \mathbf{X} from the normal population. In our simulation, we take $\sigma^2 = 1$ without loss of generality. Note that under the prior π , for fixed \mathbf{X} , $\theta_1 \leq \theta_1^\alpha(\pi|\mathbf{X})$ if and only

TABLE 4.1 *Frequentist coverage probability of 0.05 (0.95) posterior quantiles of θ_1*

n	θ_1	k	c	π_1^R	π_2^R
3	1	1	0	0.1078(0.9269)	0.0658(0.9699)
			1	0.1002(0.9340)	0.0616(0.9744)
		2	0	0.0995(0.9327)	0.0594(0.9710)
			1	0.1124(0.9290)	0.0698(0.9727)
		3	1	0.1023(0.9061)	0.0584(0.9558)
			1	0.0976(0.9043)	0.0547(0.9526)
		2	0	0.1049(0.9238)	0.0663(0.9680)
			1	0.1036(0.9308)	0.0599(0.9715)
	5	1	0	0.0992(0.9007)	0.0569(0.9484)
			1	0.1114(0.9067)	0.0629(0.9511)
		2	0	0.1039(0.9285)	0.0653(0.9716)
			1	0.1087(0.9278)	0.0680(0.9681)
	10	1	0	0.0987(0.9091)	0.0517(0.9550)
			1	0.0980(0.8994)	0.0550(0.9505)
		2	0	0.1053(0.9320)	0.0625(0.9690)
			1	0.1087(0.9307)	0.0657(0.9704)
5	1	1	0	0.0734(0.9395)	0.0504(0.9617)
			1	0.0769(0.9401)	0.0548(0.9627)
		2	0	0.0770(0.9413)	0.0553(0.9648)
			1	0.0771(0.9378)	0.0543(0.9611)
		3	1	0.0734(0.9246)	0.0523(0.9495)
			1	0.0706(0.9301)	0.0486(0.9532)
		2	0	0.0705(0.9394)	0.0489(0.9615)
			1	0.0742(0.9395)	0.0523(0.9613)
	5	1	0	0.0751(0.9261)	0.0491(0.9495)
			1	0.0703(0.9280)	0.0482(0.9506)
		2	0	0.0717(0.9402)	0.0516(0.9620)
			1	0.0780(0.9409)	0.0560(0.9633)
	10	1	0	0.0772(0.9245)	0.0516(0.9491)
			1	0.0697(0.9263)	0.0486(0.9478)
		2	0	0.0704(0.9359)	0.0500(0.9605)
			1	0.0760(0.9428)	0.0555(0.9643)
10	1	1	0	0.0597(0.9452)	0.0501(0.9543)
			1	0.0620(0.9459)	0.0515(0.9550)
		2	0	0.0584(0.9447)	0.0486(0.9563)
			1	0.0604(0.9436)	0.0499(0.9543)
		3	1	0.0634(0.9427)	0.0532(0.9526)
			1	0.0613(0.9440)	0.0502(0.9531)
		2	0	0.0598(0.9409)	0.0501(0.9518)
			1	0.0595(0.9446)	0.0504(0.9550)

(continued)

TABLE 4.1 *Frequentist coverage probability of 0.05 (0.95) posterior quantiles of θ_1 (continued)*

n	θ_1	k	c	π_1^R	π_2^R
10	5	1	0	0.0630(0.9391)	0.0513(0.9491)
			1	0.0600(0.9440)	0.0514(0.9533)
		2	0	0.0626(0.9415)	0.0524(0.9518)
			1	0.0602(0.9431)	0.0506(0.9528)
	10	1	0	0.0593(0.9393)	0.0488(0.9494)
			1	0.0601(0.9350)	0.0507(0.9468)
		2	0	0.0576(0.9411)	0.0479(0.9521)
			1	0.0601(0.9460)	0.0500(0.9557)

if $F(\theta_1^\alpha(\pi|\mathbf{X})|\mathbf{X}) \leq \alpha$. Under the prior π , $P_{(\theta_1, \theta_2)}(\alpha; \theta_1)$ can be estimated by the relative frequency of $F(\theta_1^\alpha|\mathbf{X}) \leq \alpha$.

In Table 4.1, we can observe that one-at-a-time reference prior π_2^R meets very well the target coverage probabilities. Also note that the results of tables are not much sensitive to the change of the values of (θ_1, θ_2) . Thus we recommend to use the one-at-a-time reference prior in our situation.

For the hypotheses

$$H_1 : \frac{c + \mu_1}{c + \mu_2} = 1 \quad \text{vs.} \quad H_2 : \frac{c + \mu_1}{c + \mu_2} \neq 1,$$

we want to compare the classical F -test (Cox, 1985) with Bayesian test using the one-at-a-time reference prior based on p -values and posterior probabilities of H_1 . We evaluate the p -values and posterior probabilities of H_1 for several configurations n , c , k , σ^2 , μ_1 and μ_2 . The p -values are computed based on F -test (Cox, 1985) with 1 and $n_1 + n_2$ degrees of freedom. The Bayes factors and the posterior probabilities of H_1 being true are computed assuming equal prior probabilities. The numerical values of P -values, Bayes factors and posterior probabilities are given in Table 4.2.

From Table 4.2, when $(\mu_1, \mu_2) = (1, 2)$ and $(\mu_1, \mu_2) = (1, 3)$, the Bayes factors select H_2 properly, but the p -values do not choose H_2 for some cases. Actually for this case, as the sample sizes become larger, the p -values will select H_2 . The both p -values and Bayes factors support H_1 for the case of $(\mu_1, \mu_2) = (1, 1)$. Thus the Bayesian testing procedure gives fairly reasonable answers.

TABLE 4.2 Bayes factors and posterior probabilities for testing $H_1 : (c + \mu_1)/(c + \mu_2) = 1$ vs.
 $H_2 : (c + \mu_1)/(c + \mu_2) \neq 1$

μ_1	μ_2	k	c	n_1	n_2	Bayes factor	Posterior probability	p-value
1	1	1	0	5	5	0.5742	0.6352	0.7392
				10	10	0.4529	0.6883	0.4856
				20	20	0.3163	0.7597	0.6970
				30	30	0.1727	0.8527	0.7871
		1	1	5	5	0.6220	0.6165	0.8843
				10	10	0.4850	0.6734	0.5824
				20	20	0.3443	0.7439	0.6565
				30	30	0.2669	0.7894	0.9257
	2	0	0	5	5	0.3082	0.7644	0.7057
				10	10	0.1656	0.8579	0.4675
				20	20	0.0757	0.9296	0.7908
				30	30	0.0606	0.9429	0.6698
		1	1	5	5	0.4756	0.6777	0.7705
				10	10	0.2731	0.7855	0.9946
				20	20	0.1729	0.8526	0.9549
				30	30	0.0809	0.9251	0.6563
1	2	1	0	5	5	2.5611	0.2808	0.3545
				10	10	5.4849	0.1542	0.1300
				20	20	32.3203	0.0300	0.0023
				30	30	326.1383	0.0031	0.0053
		1	1	5	5	0.9317	0.5177	0.4563
				10	10	2.9585	0.2526	0.1034
				20	20	11.3815	0.0808	0.0150
				30	30	151.5732	0.0066	0.0062
	2	0	0	5	5	0.7220	0.5807	0.4467
				10	10	3.1379	0.2417	0.1842
				20	20	15.0477	0.0623	0.0903
				30	30	262.5194	0.0038	0.0218
		1	1	5	5	0.7784	0.5623	0.3570
				10	10	0.8126	0.5517	0.2398
				20	20	1.8484	0.3511	0.2097
				30	30	2.8366	0.2606	0.0815
1	3	1	0	50	50	240.8619	0.0041	0.0392
				5	5	3.3030	0.2324	0.0294
				10	10	168.9921	0.0059	0.0009
				20	20	389600.7181	0.0000	0.0001
				30	30	915588.9037	0.0000	0.0000
		1	1	5	5	2.8754	0.2580	0.0398
				10	10	21.4005	0.0446	0.0029
				20	20	498.7589	0.0020	0.0002
				30	30	26629.4126	0.0000	0.0000

(continued)

TABLE 4.2 Bayes factors and posterior probabilities for testing $H_1 : (c + \mu_1)/(c + \mu_2) = 1$ vs. $H_2 : (c + \mu_1)/(c + \mu_2) \neq 1$ (continued)

μ_1	μ_2	k	c	n_1	n_2	Bayes factor	Posterior probability	p-value
1	3	2	0	5	5	11.0401	0.0831	0.2905
				10	10	24.5567	0.0391	0.0228
				20	20	2183.8789	0.0005	0.0003
				30	30	76075329.1029	0.0000	0.0008
		1	5	5		1.0751	0.4819	0.3705
				10	10	11.1865	0.0821	0.0866
				20	20	97.8067	0.0101	0.0470
				30	30	291.9831	0.0034	0.0012

TABLE 4.3 Fine gravel in surface soils

Soil type	Fine gravel (%)								\bar{x}_i	S_i^2
Good (1)	5.9	3.8	6.5	18.3	18.2	16.1	7.6		10.9	40.13
Poor (2)	7.6	0.4	1.1	3.2	6.5	4.1	4.7		3.94	6.95

4.2. Example

This example was analyzed by Cox (1935). We analyze the same data using our proposed method. The data in Table 4.3 are used in Steel and Torrie (1980) to illustrate the Satterthwaite (1946) procedure for the unequal variance case. Since $\bar{x}_1/\bar{x}_2 = 2.77$ and $s_1/s_2 = 2.40$, this may be regarded as a $c = 0$, $k = 2$ or as $c = 1$, $k = 2$ because $(1 + \bar{x}_1)/(1 + \bar{x}_2) = 2.41$ is closer to s_1/s_2 by Cox (1985). The 95% confidence intervals for $(c + \mu_1)/(c + \mu_2)$ are given in Table 4.4. The procedures are those of Fieller (1944), Elston (1969), the conventional analysis using the log transformation and Cox (1985). The procedure given by Fieller (1945) used the assumption that the data are normally distributed with common variance. The Elston (1969) procedure predicated normality and unequal variances. In some other way, Cox (1985) gave the procedure based on the relation of variance and mean.

From Table 4.4, the classical procedures exclude unit consistently with the inference that $(c + \mu_1)/(c + \mu_2) \neq 1$. When the prior probabilities are assumed to be equal, Table 4.5 gives the Bayes factor and posterior probability for testing

$$H_1 : \frac{c + \mu_1}{c + \mu_2} = 1 \text{ vs. } H_2 : \frac{c + \mu_1}{c + \mu_2} \neq 1.$$

From Table 4.5, we see that there is very strong evidence for H_2 in terms of

TABLE 4.4 The 95% confidence intervals for $(c + \mu_1)/(c + \mu_2)$

Procedure	c	Interval limits
Cox (1985)	0	1.27, 6.68
Fieller (1944)	0	Improper
Elston (1969)	0	1.13, 7.85
Log Transformation	0	1.16, 8.90
Cox (1985)	1	1.23, 4.75
Log Transformation	1	1.21, 5.10

TABLE 4.5 The Bayes factor and posterior probability for testing $H_1 : (c + \mu_1)/(c + \mu_2) = 1$ vs. $H_2 : (c + \mu_1)/(c + \mu_2) \neq 1$

c	B_{21}^F	$P^F(H_2 x_1, x_2)$
0	23.5327	0.9592
1	28.2342	0.9658

posterior probability. The Bayesian test procedure together with the classical procedures gives fairly reasonable answers.

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 the Royal Statistical Society*, **B41**, 113–147.
- COX, C. P. (1985). "Interval estimates for the ratio of the means of two normal populations with variances related to the means", *Biometrics*, **41**, 261–265.
- COX, C. P. AND ROSEBERRY, T. D. (1966). "A note on the variance of the distribution of sample number in sequential probability ratio tests", *Technometrics*, **8**, 700–704.
- COX, D. R. AND REID, N. (1987). "Parameter orthogonality and approximate conditional inference (with discussion)", *Journal of the Royal Statistical Society*, **B49**, 1–39.
- DATTA, G. S. AND GHOSH, J. K. (1995). "On priors providing frequentist validity for Bayesian inference", *Biometrika*, **82**, 37–45.
- DATTA, G. S. AND GHOSH, M. (1995). "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.
- ELSTON, R. C. (1969). "An analogue to Fieller's theorem using Scheffé's solution to the Fisher-Behrens problem", *The American Statistician*, **23**, 26–28.
- FIELLER, E. C. (1944). "A fundamental formula in the statistics of biological assay, and some applications", *Quarterly Journal of Pharmacy and Pharmacology*, **17**, 117–123.
- 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.
- 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.
- O'HAGAN, A. (1995). "Fractional Bayes factors for model comparison (with discussion)", *Journal of the Royal Statistical Society*, **B57**, 99–138.
- SATTERTHWAITE, F. E. (1946). "An approximate distribution of estimates of variance components", *Biometrics Bulletin*, **2**, 110–114.
- SNEDECOR, G. W. AND COCHRAN, W. G. (1980). *Statistical Methods*, 7th ed., Iowa State University Press, Ames, Iowa.
- STEEL, R. G. D. AND TORRIE, J. H. (1980). *Principles and Procedures of Statistics*, 2nd ed., McGraw-Hill, New York.
- 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 likelihoods", *Journal of the Royal Statistical Society*, **25**, 318–329.