Journal of the Korean Data & Information Science Society
2012, **23**(3), 605–616
http://dx.doi.org/10.7465/jkdi.2012.23.3.605
한국데이터정보과학회지

Default Bayesian testing on the common mean of several normal distributions

Sang Gil Kang[1] · Dal Ho Kim[2] · Woo Dong Lee[3]

[1]Department of Computer and Data Information, Sangji University
[2]Department of Statistics, Kyungpook National University
[3]Department of Asset Management, Daegu Haany University
Received 13 April 2012, revised 12 May 2012, accepted 17 May 2012

Abstract

This article deals with the problem of testing on the common mean of several normal populations. We propose Bayesian hypothesis testing procedures for the common normal mean under the noninformative prior. The noninformative prior is usually improper and yields a calibration problem that makes the Bayes factor to be defined up to a multiplicative constant. So we propose the default Bayesian hypothesis testing procedures based on the fractional Bayes factor and the intrinsic Bayes factors under the reference priors. Simulation study and an example are provided.

Keywords: Common normal mean, fractional Bayes factor, intrinsic Bayes factor, reference prior.

1. Introduction

The inference on a common mean of several normal distributions with unequal variances has attracted the attention of many researches. This problem is quite natural in balanced incomplete block design with uncorrelated random block effects and fixed treatments effects (Montgomery, 1991, pp. 184-186). In this set up, the intra-block estimator and the interblock estimator of a treatment contrast are independent normal with a common mean, but their variances are unknown and unequal. A second example relates to meta analysis where for examples, several clinics and social and behavioral sciences provide estimates of a common parameter of interest, and the problem is how to combine these estimates meaningfully into a single one.

Point estimation of the common mean has been addressed both from the classical and decision theoretic points of view. Among others, we may refer to Graybill and Deal (1959), Zacks (1966,1970), Khatri and Shah (1974), Cohen and Sackrowitz (1974), Brown and Cohen (1974), Shinozaki (1978), Bhattacharya (1980), Sinha and Mouqadem (1982), Sinha (1985) and Kubokawa (1990).

[1] Associate professor, Department of Computer and Data Information, Sangji University, Wonju 220-702, Korea.
[2] Professor, Department of Statistics, Kyungpook National University, Daegu 702-701, korea.
[3] Corresponding author: Professor, Department of Asset Management, Daegu Haany University, Kyungsan 712-715, Korea. E-mail: wdlee@dhu.ac.kr

For interval estimation of the common mean, approximate confidence intervals are found in Meier (1953), Brown and Cohen (1974), Sinha (1985) and Eberhardt *et al.* (1989). Exact intervals are proposed in Fairweather (1972), Cohen and Sackrowitz (1974) and Jordan and Krishnamoorthy (1996). Yu *et al.* (1999) and Hartung *et al.* (2008) have these intervals and others exact intervals compared by their lengths when the confidence coefficients are the same. It should be noted that all the methods (except Fairweather's) considered in Yu *et al.* (1999) do not always produce nonempty confidence intervals. They required satisfying some conditions in order to yield nonempty intervals. Krishnamoorthy and Lu (2003) and Lin and Lee (2005) proposed a procedure based on the generalized confidence limits, but with different pivotal quantities.

In contrast, much less attention has been paid to the hypothesis testing problem, presumably due to the complicated sampling distributions of the test statistics involved. Cohen and Sackrowitz (1989) proposed a test combining individual tests by weighting with respect to their sample variances. This idea was extended by Zhou and Mathew (1993) who proposed two tests and compared their power functions with that of Fisher's (1932) test. Krishnamoorthy and Lu (2003) proposed a test based on the generalized p-value approach, and showed that the power of the generalized test is much higher than those of the other tests considered when the population size is five or more regardless of the sample sizes. Also Lin and Lee (2005) used the same the generalized p-value approach with a different pivot, and showed that the proposed method is better than the existing methods in the senses of having the highest powers by simulation study. However it is not clear how closely the size of the generalized test of Lin and Lee (2005) follow the nominal level (Chang and Pal, 2008). Chang and Pal (2008) proposed three tests based on the Graybill-Deal estimator as well as the maximum likelihood estimator, and showed that the three tests exhibit good size and power behavior by simulation study.

In Bayesian model selection or testing problem, the Bayes factor under proper priors or informative priors have been very successful. However, limited information and time constraints often require the use of noninformative priors. Since noninformative priors such as Jeffreys' prior or reference prior (Berger and Bernardo, 1989, 1992) are typically improper so that such priors are only defined up to arbitrary constants which affects the values of Bayes factors. Spiegelhalter and Smith (1982), O'Hagan (1995) and Berger and Pericchi (1996) have made efforts to compensate for that arbitrariness.

Spiegelhalter and Smith (1982) used the device of imaginary training sample in the context of linear model comparisons to choose the arbitrary constants. But the choice of imaginary training sample depends on the models under comparison, and so there is no guarantee that the Bayes factor of Spiegelhalter and Smith (1982) is coherent for multiple model comparisons. Berger and Pericchi (1996) introduced the intrinsic Bayes factor using a data-splitting idea, which would eliminate the arbitrariness of improper prior. O'Hagan (1995) proposed the fractional Bayes factor. For removing the arbitrariness he used to a portion of the likelihood with a so-called the fraction b. These approaches have shown to be quite useful in many statistical areas (Kang *et al.*, 2008, 2011; Lee and Kang, 2008). An excellent exposition of the objective Bayesian method to model selection is Berger and Pericchi (2001).

In this paper, we propose the objective Bayesian hypothesis testing procedures for the common mean of several normal distributions based on the Bayes factors. The outline of the remaining sections is as follows. In Section 2, we introduce the Bayesian hypothesis testing based on the Bayes factors. In Section 3, under the reference prior, we provide the Bayesian hypothesis testing procedures based on the fractional Bayes factor and the intrinsic Bayes factors. In Section 4, simulation study and an example are given.

2. Intrinsic and fractional Bayes factors

Suppose that hypotheses H_1,H_2 ,$\cdots$,H_q are under consideration, with the data $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ having probability density function $f_i(\mathbf{x}|\theta_i)$ under hypothesis H_i. The parameter vector θ_i is unknown. Let $\pi_i(\theta_i)$ be the prior distributions of hypothesis H_i, and let p_i be the prior probability of hypothesis H_i,$i = 1, 2, \cdots, q$. Then the posterior probability that the hypothesis H_i is true is

$$P(H_i|\mathbf{x}) = \left(\sum_{j=1}^{q} \frac{p_j}{p_i} \cdot B_{ji}\right)^{-1}, \tag{2.1}$$

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

$$B_{ji} = \frac{\int f_j(\mathbf{x}|\theta_j)\pi_j(\theta_j)d\theta_j}{\int f_i(\mathbf{x}|\theta_i)\pi_i(\theta_i)d\theta_i} = \frac{m_j(\mathbf{x})}{m_i(\mathbf{x})}. \tag{2.2}$$

The B_{ji} interpreted as the comparative support of the data for H_j versus H_i. The computation of B_{ji} needs specification of the prior distribution $\pi_i(\theta_i)$ and $\pi_j(\theta_j)$. Often in Bayesian analysis, one can use noninformative priors π_i^N. Common choices are the uniform prior, Jeffreys' prior and the reference prior. The noninformative prior π_i^N is typically improper. Hence the use of noninformative prior π_i^N in (2.2) causes the B_{ji} to contain unspecified constants. To solve this problem, Berger and Pericchi (1996) proposed the intrinsic Bayes factor, and O'Hagan (1995) proposed the fractional Bayes factor.

One solution to this indeterminacy problem is to use part of the data as a training sample. Let $\mathbf{x}(l)$ denote the part of the data to be so used and let $\mathbf{x}(-l)$ be the remainder of the data, such that

$$0 < m_i^N(\mathbf{x}(l)) < \infty, i = 1, \cdots, q. \tag{2.3}$$

In view (2.3), the posteriors $\pi_i^N(\theta_i|\mathbf{x}(l))$ are well defined. Now, consider the Bayes factor $B_{ji}(l)$ with the remainder of the data $\mathbf{x}(-l)$ using $\pi_i^N(\theta_i|\mathbf{x}(l))$ as the priors:

$$B_{ji}(l) = \frac{\int f(\mathbf{x}(-l)|\theta_j, \mathbf{x}(l))\pi_j^N(\theta_j|\mathbf{x}(l))d\theta_j}{\int f(\mathbf{x}(-l)|\theta_i, \mathbf{x}(l))\pi_i^N(\theta_i|\mathbf{x}(l))d\theta_i} = B_{ji}^N \cdot B_{ij}^N(\mathbf{x}(l)) \tag{2.4}$$

where

$$B_{ji}^N = B_{ji}^N(\mathbf{x}) = \frac{m_j^N(\mathbf{x})}{m_i^N(\mathbf{x})}$$

and

$$B_{ij}^N(\mathbf{x}(l)) = \frac{m_i^N(\mathbf{x}(l))}{m_j^N(\mathbf{x}(l))}$$

are the Bayes factors that would be obtained for the full data $\mathbf{x}$ and training samples $\mathbf{x}(l)$, respectively.

Berger and Pericchi (1996) proposed the use of a minimal training sample to compute $B_{ij}^N(\mathbf{x}(l))$. Then, an average over all the possible minimal training samples contained in the sample is computed. Thus the arithmetic intrinsic Bayes factor (AIBF) of H_j to H_i is

$$B_{ji}^{AI} = B_{ji}^N \times \frac{1}{L}\sum_{l=1}^{L} B_{ij}^N(\mathbf{x}(l)), \tag{2.5}$$

where L is the number of all possible minimal training samples. Also the median intrinsic Bayes factor (MIBF) by Berger and Pericchi (1998) of H_j to H_i is

$$B_{ji}^{MI} = B_{ji}^N \times ME[B_{ij}^N(\mathbf{x}(l))], \tag{2.6}$$

where ME indicates the median for all the training sample Bayes factors.

Therefore we can also calculate the posterior probability of H_i using (2.1), where B_{ji} is replaced by B_{ji}^{AI} and B_{ji}^{MI} from (2.5) and (2.6), respectively.

The fractional Bayes factor (O'Hagan, 1995) is based on a similar intuition to that behind the intrinsic Bayes factor but, instead of using part of the data to turn noninformative priors into proper priors, it uses a fraction, b ,of each likelihood function, $L(\theta_i) = f_i(\mathbf{x}|\theta_i)$, with the remaining $1-b$ fraction of the likelihood used for model discrimination. Then the fractional Bayes factor (FBF) of hypothesis H_j versus hypothesis H_i is

$$B_{ji}^F = B_{ji}^N \cdot \frac{\int L^b(\mathbf{x}|\theta_i)\pi_i^N(\theta_i)d\theta_i}{\int L^b(\mathbf{x}|\theta_j)\pi_j^N(\theta_j)d\theta_j} = B_{ji}^N \cdot \frac{m_i^b(\mathbf{x})}{m_j^b(\mathbf{x})}. \tag{2.7}$$

O'Hagan (1995) proposed three ways for the choice of the fraction b. One common choice of b is $b = m/n$, where m is the size of the minimal training sample, assuming that this number is uniquely defined. See O'Hagan (1995, 1997) and the discussion by Berger and Mortera in O'Hagan (1995).

3. Bayesian hypothesis testing procedures

Let $X_{ij}, i = 1, \cdots, k, j = 1, \cdots, n_i$, denote observations from $N(\mu, \sigma_i^2)$. Then likelihood function is given by

$$f(\mathbf{x}|\mu, \sigma_1, \cdots, \sigma_k) = \left(\sqrt{2\pi}\right)^{-n}\left(\prod_{i=1}^{k}\sigma_i^{-n_i}\right)\exp\left\{-\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{(x_{ij}-\mu)^2}{2\sigma_i^2}\right\}, \tag{3.1}$$

where $\mathbf{x} = (\mathbf{x}_1, \cdots, \mathbf{x}_k)$, $\mathbf{x}_i = (x_{i1}, \cdots, x_{in_i}), i = 1, \cdots, k$ and $n = \sum_{i=1}^{k} n_i$. We are interested in testing the hypotheses $H_1 : \mu = \mu_0$ versus $H_2 : \mu \neq \mu_0$ based on the fractional Bayes factor and the intrinsic Bayes factors.

3.1. Bayesian hypothesis testing procedure based on the fractional Bayes factor

From (3.1) the likelihood function under the hypothesis $H_1 : \mu = \mu_0$ is

$$L_1(\sigma_1, \cdots, \sigma_k|\mathbf{x}) = \left(\sqrt{2\pi}\right)^{-n}\left(\prod_{i=1}^{k}\sigma_i^{-n_i}\right)\exp\left\{-\sum_{i=1}^{k}\frac{1}{2\sigma_i^2}\left[S_i^2 + n_i(\bar{x}_i-\mu_0)^2\right]\right\}, \tag{3.2}$$

where $S_i^2 = \sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2$ and $\bar{x}_i = \sum_{j=1}^{n_i} x_{ij}/n_i, i = 1, \cdots, k$. And under the hypothesis H_1, the reference prior for $(\sigma_1, \cdots, \sigma_k)$ is

$$\pi_1^N(\sigma_1, \cdots, \sigma_k) \propto \prod_{i=1}^k \sigma_i^{-1}. \tag{3.3}$$

Then from the likelihood (3.2) and the reference prior (3.3) , the element $m_1^b(\mathbf{x})$ of the FBF under H_1 is given by

$$\begin{aligned} m_1^b(\mathbf{x}) &= \int_0^\infty \cdots \int_0^\infty L_1^b(\sigma_1, \cdots, \sigma_k|\mathbf{x})\pi_1^N(\sigma_1, \cdots, \sigma_k) d\sigma_1 \cdots d\sigma_k \\ &= \left(\sqrt{2\pi}\right)^{-bn} 2^{-k} \prod_{i=1}^k \Gamma\left[\frac{bn_i}{2}\right] \left\{\frac{b[S_i^2 + n_i(\bar{x}_i - \mu_0)^2]}{2}\right\}^{-\frac{bn_i}{2}}. \end{aligned} \tag{3.4}$$

For the hypothesis H_2, the reference prior for $(\mu, \sigma_1, \cdots, \sigma_k)$ is

$$\pi^N(\mu, \sigma_1, \cdots, \sigma_k) \propto \prod_{i=1}^k \sigma_i^{-1}. \tag{3.5}$$

The likelihood function under the hypothesis H_2 is

$$L_2(\mu, \sigma_1, \cdots, \sigma_k|\mathbf{x}) = \left(\sqrt{2\pi}\right)^{-n} \left(\prod_{i=1}^k \sigma_i^{-n_i}\right) \exp\left\{-\sum_{i=1}^k \frac{1}{2\sigma_i^2}\left[S_i^2 + n_i(\bar{x}_i - \mu)^2\right]\right\}. \tag{3.6}$$

Thus from the likelihood (3.6) and the reference prior (3.5), the element $m_2^b(\mathbf{x})$ of FBF under H_2 is given as follows.

$$\begin{aligned} m_2^b(\mathbf{x}) &= \int_{-\infty}^\infty \int_0^\infty \cdots \int_0^\infty L_2^b(\mu, \sigma_1, \cdots, \sigma_k|\mathbf{x})\pi_2^N(\mu, \sigma_1, \cdots, \sigma_k) d\sigma_1 \cdots d\sigma_k d\mu \\ &= \left(\sqrt{2\pi}\right)^{-bn} 2^{-k} \prod_{i=1}^k \Gamma\left[\frac{bn_i}{2}\right] \int_{-\infty}^\infty \prod_{i=1}^k \left\{\frac{b[S_i^2 + n_i(\bar{x}_i - \mu)^2]}{2}\right\}^{-\frac{bn_i}{2}} d\mu. \end{aligned} \tag{3.7}$$

Therefore the element B_{21}^N of FBF is given by

$$B_{21}^N = \frac{S_2(\mathbf{x})}{S_1(\mathbf{x})}, \tag{3.8}$$

where

$$S_1(\mathbf{x}) = \prod_{i=1}^k \left\{S_i^2 + n_i(\bar{x}_i - \mu_0)^2\right\}^{-\frac{n_i}{2}}$$

and

$$S_2(\mathbf{x}) = \int_{-\infty}^\infty \prod_{i=1}^k \left\{S_i^2 + n_i(\bar{x}_i - \mu)^2\right\}^{-\frac{n_i}{2}} d\mu.$$

And the ratio of marginal densities with fraction b is

$$\frac{m_1^b(\mathbf{x})}{m_2^b(\mathbf{x})} = \frac{S_1(\mathbf{x};b)}{S_2(\mathbf{x};b)}, \tag{3.9}$$

where

$$S_1(\mathbf{x};b) = \prod_{i=1}^{k} \left\{S_i^2 + n_i(\bar{x}_i - \mu_0)^2\right\}^{-\frac{bn_i}{2}}$$

and

$$S_2(\mathbf{x};b) = \int_{-\infty}^{\infty} \prod_{i=1}^{k} \left\{S_i^2 + n_i(\bar{x}_i - \mu)^2\right\}^{-\frac{bn_i}{2}} d\mu.$$

Thus the FBF of H_2 versus H_1 is given by

$$B_{21}^F = \frac{S_2(\mathbf{x})}{S_1(\mathbf{x})} \cdot \frac{S_1(\mathbf{x};b)}{S_2(\mathbf{x};b)}. \tag{3.10}$$

Note that the calculations of the FBF of H_2 versus H_1 requires only one dimensional integration.

3.2. Bayesian hypothesis testing procedure based on the intrinsic Bayes factor

The element B_{21}^N of the intrinsic Bayes factor is computed in the fractional Bayes factor. So under minimal training sample, we only calculate the marginal densities for the hypotheses H_1 and H_2, respectively.

Let $\mathbf{X}_l = (X_{li}, X_{lj}), i, j(i < j) = 1, 2, \cdots, n_l$ be a random sample of size 2 from population $l, l = 1, 2, \cdots, k$. The marginal density of $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_k$ under the hypothesis H_1 with prior (3.3) is

$$\begin{aligned} & m_1^N(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k) \\ &= \int_0^\infty \cdots \int_0^\infty f(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k|\sigma_1, \cdots, \sigma_k)\pi_1^N(\sigma_1, \cdots, \sigma_k)d\sigma_1 \cdots d\sigma_k \\ &= \left(\sqrt{2\pi}\right)^{-2k} 2^k \prod_{l=1}^{k} \left[(x_{li} - x_{lj})^2 + (x_{li} + x_{lj} - 2\mu_0)^2\right]^{-1}, \end{aligned}$$

where $\mathbf{x}_l = (x_{li}, x_{lj}), i, j(i < j) = 1, 2, \cdots, n_l,\ l = 1, 2, \cdots, k$.

And the marginal density $m_2^N(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k)$ under H_2 is given by

$$\begin{aligned} & m_2^N(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k) \\ &= \int_\infty^{-\infty} \int_0^\infty \cdots \int_0^\infty f(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k|\mu, \sigma_1, \cdots, \sigma_k)\pi_2^N(\mu, \sigma_1, \cdots, \sigma_k)d\sigma_1 \cdots d\sigma_k d\mu \\ &= \left(\sqrt{2\pi}\right)^{-2k} 2^k \int_{-\infty}^{\infty} \prod_{l=1}^{k} \left[(x_{li} - x_{lj})^2 + (x_{li} + x_{lj} - 2\mu)^2\right]^{-1} d\mu. \end{aligned}$$

Since the marginal densities m_1^N and m_2^N are finite, the minimal training sample is $\mathbf{x}_l = (x_{li}, x_{lj}), i, j(i < j) = 1, 2, \cdots, n_l,\ l = 1, 2, \cdots, k$. Thus we can conclude that any training sample of size $2k$ is a minimal training sample.

Therefore the AIBF of H_2 versus H_1 is given by

$$B_{21}^{AI} = \frac{S_2(\mathbf{x})}{S_1(\mathbf{x})}\left[\frac{1}{L}\sum_{1i<1j}^{n_1}\sum_{2i<2j}^{n_2}\cdots\sum_{ki<kj}^{n_k}\frac{T_1(x_{1i}, x_{1j}, x_{2i}, x_{2j}, \cdots, x_{ki}, x_{kj})}{T_2(x_{1i}, x_{1j}, x_{2i}, x_{2j}, \cdots, x_{ki}, x_{kj})}\right], \tag{3.11}$$

where $L = \prod_{i=1}^k [n_i(n_i - 1)]/2$,

$$T_1(x_{1i}, x_{1j}, x_{2i}, x_{2j}, \cdots, x_{ki}, x_{kj}) = \prod_{l=1}^k \left[(x_{li} - x_{lj})^2 + (x_{li} + x_{lj} - 2\mu_0)^2\right]^{-1}$$

and

$$T_2(x_{1i}, x_{1j}, x_{2i}, x_{2j}, \cdots, x_{ki}, x_{kj}) = \int_{-\infty}^{\infty}\prod_{l=1}^k \left[(x_{li} - x_{lj})^2 + (x_{li} + x_{lj} - 2\mu)^2\right]^{-1} d\mu.$$

Also the MIBF of H_2 versus H_1 is given by

$$B_{21}^{MI} = \frac{S_2(\mathbf{x})}{S_1(\mathbf{x})} ME\left[\frac{T_1(x_{1i}, x_{1j}, x_{2i}, x_{2j}, \cdots, x_{ki}, x_{kj})}{T_2(x_{1i}, x_{1j}, x_{2i}, x_{2j}, \cdots, x_{ki}, x_{kj})}\right]. \tag{3.12}$$

Note that the calculations of the AIBF and the MIBF of H_2 versus H_1 require only one dimensional integration.

Remark 3.1 To compare the GDE1 and GDE2 tests of Chang and Pal (2008) and the Bayes factors in Section 4, we describe the GDE1 and GDE2 tests. The GDE1 test uses Sinha's (1985) first order unbiased variance estimator of $\hat{\mu}_{GDE}$, and the GDE2 test uses the corrected exact unbiased variance estimator of $\hat{\mu}_{GDE}$ obtained Chang and Pal (2008). The GDE1 test is

$$\text{Reject } H_1 \text{ if } \Delta_{GDE1} = (\hat{\mu}_{GDE} - \mu_0)^2/\widehat{V}_{(1)}(\hat{\mu}_{GDE}) > \chi^2_{1,\alpha},$$

where $s_i^2 = S_i^2/(n_i - 1)$,

$$\begin{aligned}
\hat{\mu}_{GDE} &= \sum_{i=1}^k (n_i/s_i^2)\bar{x}_i / \sum_{i=1}^k (n_i/s_i^2), \\
\widehat{V}_{(1)}(\hat{\mu}_{GDE}) &= \left(\sum_{i=1}^k n_i/s_i^2\right)^{-1}\left[1 + 4\sum_{i=1}^k (n_i+1)^{-1}(n_i/s_i^2)\right. \\
&\quad \times \left.\left\{\sum_{i=1}^k (n_i/s_i^2) - (n_i/s_i^2)^2/\left(\sum_{i=1}^k n_i/s_i^2\right)^2\right\}^{-1}\right].
\end{aligned}$$

And the GDE2 test is

$$\text{Reject } H_1 \text{ if } \Delta_{GDE2} = (\hat{\mu}_{GDE} - \mu_0)^2/\widehat{V}(\hat{\mu}_{GDE}) > (t_{l,(\alpha/2)})^2,$$

where

$$\begin{aligned}
\widehat{V}(\hat{\mu}_{GDE}) &= \left(\sum_{i=1}^{k} n_i/s_i^2\right)^{-2} \sum_{i=1}^{k}(n_i/s_i^2)_2F_1\left(1,2;(n_i+1)/2;1-(n_i/s_i^2)/\left(\sum_{i=1}^{k} n_i/s_i^2\right)\right),\\
t_{l,(\alpha/2)} &= \{t_{[l],(\alpha/2)}\} + \{(t_{[l]+1,(\alpha/2)} - t_{[l],(\alpha/2)})(l-[l])\},\\
l &\approx \left(\sum_{i=1}^{k}(s_i^2/n_i)\right)^2 / \left(\sum_{i=1}^{k}(s_i^2/n_i)^2/(n_i-1)\right),
\end{aligned}$$

and here ${}_2F_1(a,b;c;z)$ is the Gaussian hypergeometric function and $[l]$ is the largest integer smaller than l.

4. Numerical studies

In order to assess the Bayesian hypothesis testing procedures, we evaluate the posterior probability for several configurations of $(\mu,\sigma_1),\cdots,(\mu,\sigma_k)$ and $(n_1,\cdots,n_k)$. In particular, for fixed $(\mu,\sigma_1),\cdots,(\mu,\sigma_k)$, we take 1,000 independent random samples of $\mathbf{X}_i$ with sample size n_i from $N(\mu,\sigma_i^2), i=1,\cdots,k$. In our simulation, we put $\mu=0$ without loss of generality. We want to test the hypotheses $H_1:\mu=\mu_0$ versus $H_2:\mu\neq\mu_0$. The posterior probabilities of H_1 being true are computed assuming equal prior probabilities. Tables 4.1 and 4.2 show the results of the averages and the standard deviations in parentheses of posterior probabilities. For fixed σ_1,σ_2 and μ_0, Table 4.1 shows results with $k=2$, $n_1=5,10$ and $n_2=5,10,20$. Table 4.2 is designed under the condition that $k=3$, $n_1=5,10$, $n_2=5,10$ and $n_3=5,10,20$. In Tables 4.1 and 4.2, $P^F(\cdot)$,$P^{AI}(\cdot)$ and $P^{MI}(\cdot)$ are the posterior probabilities of the hypothesis H_1 being true based on FBF, AIBF and MIBF, respectively. From Tables 4.1 and 4.2, the FBF, the AIBF and the MIBF accept the hypothesis H_1 when the values of μ_0 are close to 0, whereas reject the hypothesis H_1 when the values of μ_0 are far from 0. Also the FBF and the AIBF give a similar behavior for all sample sizes. However the MIBF favors the hypothesis H_1 than the FBF and the AIBF.

Example 4.1 In an example given by Snedecor (1950) the data from four experiments are used to estimate the percentage of albumin in plasma protein of normal human subjects. This dataset is reported in Meier (1953) and is analyzed in Jordan and Krishnamoorthy (1996) and Krishnamoorthy and Lu (2003). The data appear in the Table 4.3.

We want to test the hypotheses $H_1:\mu=\mu_0$ versus $H_2:\mu\neq\mu_0$. The p-values of GDE1 and GDE2 tests (Chang and Pal, 2008), and the values of the fractional Bayes factor and the posterior probabilities of H_1 are given in Table 4.4. The AIBF and the MIBF are not mentioned in Table 4.3 because we do not have the full original dataset available in order to compute the the AIBF and the MIBF. The results of Table 4.4 indicate that for values of μ_0 that are close to 60.5, any criteria accept the H_1. However for values of μ_0 between 59.8648 and 60.1479, the GDE1 and GDE2 tests accept the H_1 whereas the fractional Bayes factor reject H_1. The GDE1 and GDE2 tests favor the H_1 more than the fractional Bayes factor.

Table 4.1 The averages and the standard deviations in parentheses of posterior probabilities

σ_1	σ_2	μ_0	(n_1, n_2)	$P^F(H_1\|\mathbf{x},\mathbf{y})$	$P^{AI}(H_1\|\mathbf{x},\mathbf{y})$	$P^{MI}(H_1\|\mathbf{x},\mathbf{y})$
1.0	1.0	0.0	5,5	0.547 (0.127)	0.583 (0.177)	0.648 (0.147)
			5,10	0.589 (0.139)	0.617 (0.181)	0.691 (0.149)
			10,10	0.622 (0.146)	0.633 (0.189)	0.720 (0.154)
			10,20	0.667 (0.141)	0.673 (0.178)	0.762 (0.139)
		0.2	5,5	0.517 (0.145)	0.545 (0.192)	0.614 (0.166)
			5,10	0.543 (0.174)	0.568 (0.209)	0.645 (0.188)
			10,10	0.556 (0.191)	0.571 (0.222)	0.657 (0.202)
			10,20	0.562 (0.214)	0.574 (0.240)	0.659 (0.223)
		0.4	5,5	0.443 (0.181)	0.456 (0.221)	0.532 (0.202)
			5,10	0.436 (0.217)	0.454 (0.251)	0.532 (0.242)
			10,10	0.393 (0.235)	0.403 (0.258)	0.483 (0.262)
			10,20	0.338 (0.251)	0.348 (0.270)	0.422 (0.284)
		0.6	5,5	0.349 (0.195)	0.354 (0.229)	0.434 (0.221)
			5,10	0.286 (0.215)	0.293 (0.239)	0.365 (0.248)
			10,10	0.227 (0.214)	0.230 (0.234)	0.293 (0.254)
			10,20	0.145 (0.186)	0.149 (0.199)	0.193 (0.226)
		0.8	5,5	0.264 (0.184)	0.258 (0.205)	0.343 (0.216)
			5,10	0.160 (0.169)	0.159 (0.184)	0.219 (0.207)
			10,10	0.101 (0.141)	0.099 (0.148)	0.139 (0.176)
			10,20	0.031 (0.071)	0.031 (0.073)	0.045 (0.096)
		1.0	5,5	0.176 (0.160)	0.172 (0.176)	0.248 (0.196)
			5,10	0.082 (0.121)	0.079 (0.128)	0.119 (0.151)
			10,10	0.032 (0.070)	0.030 (0.072)	0.049 (0.094)
			10,20	0.006 (0.027)	0.006 (0.028)	0.009 (0.037)
1.0	2.0	0.0	5,5	0.553 (0.126)	0.583 (0.174)	0.652 (0.140)
			5,10	0.589 (0.150)	0.601 (0.179)	0.667 (0.156)
			10,10	0.638 (0.141)	0.656 (0.178)	0.733 (0.144)
			10,20	0.666 (0.152)	0.660 (0.181)	0.739 (0.152)
		0.2	5,5	0.538 (0.139)	0.573 (0.179)	0.636 (0.158)
			5,10	0.568 (0.166)	0.575 (0.199)	0.645 (0.175)
			10,10	0.601 (0.170)	0.625 (0.199)	0.699 (0.178)
			10,20	0.630 (0.185)	0.623 (0.210)	0.704 (0.186)
		0.4	5,5	0.492 (0.163)	0.516 (0.204)	0.587 (0.183)
			5,10	0.513 (0.192)	0.521 (0.219)	0.597 (0.200)
			10,10	0.479 (0.227)	0.490 (0.252)	0.573 (0.243)
			10,20	0.476 (0.248)	0.477 (0.260)	0.552 (0.259)
		0.6	5,5	0.429 (0.184)	0.447 (0.218)	0.525 (0.203)
			5,10	0.414 (0.220)	0.421 (0.238)	0.495 (0.234)
			10,10	0.354 (0.237)	0.364 (0.259)	0.437 (0.267)
			10,20	0.322 (0.246)	0.327 (0.257)	0.395 (0.273)
		0.8	5,5	0.347 (0.190)	0.353 (0.216)	0.435 (0.212)
			5,10	0.325 (0.216)	0.334 (0.235)	0.409 (0.238)
			10,10	0.207 (0.192)	0.215 (0.211)	0.276 (0.230)
			10,20	0.155 (0.190)	0.158 (0.199)	0.206 (0.230)
		1.0	5,5	0.279 (0.184)	0.281 (0.206)	0.359 (0.213)
			5,10	0.235 (0.202)	0.246 (0.218)	0.315 (0.235)
			10,10	0.114 (0.147)	0.116 (0.159)	0.160 (0.186)
			10,20	0.075 (0.125)	0.077 (0.132)	0.108 (0.159)
1.0	3.0	0.0	5,5	0.560 (0.130)	0.589 (0.173)	0.651 (0.144)
			5,10	0.619 (0.137)	0.615 (0.164)	0.679 (0.138)
			10,10	0.646 (0.138)	0.661 (0.174)	0.738 (0.139)
			10,20	0.696 (0.137)	0.667 (0.171)	0.749 (0.133)
		0.2	5,5	0.550 (0.144)	0.578 (0.187)	0.645 (0.160)
			5,10	0.591 (0.164)	0.587 (0.188)	0.652 (0.166)
			10,10	0.607 (0.175)	0.622 (0.208)	0.698 (0.180)
			10,20	0.651 (0.178)	0.633 (0.198)	0.710 (0.174)
		0.4	5,5	0.499 (0.166)	0.523 (0.203)	0.591 (0.184)
			5,10	0.533 (0.190)	0.535 (0.202)	0.600 (0.187)
			10,10	0.506 (0.219)	0.520 (0.241)	0.599 (0.231)
			10,20	0.539 (0.232)	0.525 (0.242)	0.600 (0.233)
		0.6	5,5	0.455 (0.176)	0.472 (0.210)	0.546 (0.193)
			5,10	0.460 (0.213)	0.460 (0.222)	0.530 (0.210)
			10,10	0.373 (0.236)	0.385 (0.257)	0.460 (0.261)
			10,20	0.385 (0.254)	0.373 (0.259)	0.447 (0.268)
		0.8	5,5	0.361 (0.186)	0.370 (0.210)	0.450 (0.210)
			5,10	0.371 (0.220)	0.374 (0.226)	0.451 (0.229)
			10,10	0.244 (0.206)	0.251 (0.223)	0.317 (0.241)
			10,20	0.245 (0.225)	0.240 (0.224)	0.302 (0.246)
		1.0	5,5	0.319 (0.183)	0.330 (0.207)	0.408 (0.213)
			5,10	0.296 (0.216)	0.306 (0.221)	0.380 (0.233)
			10,10	0.149 (0.172)	0.150 (0.179)	0.202 (0.207)
			10,20	0.124 (0.163)	0.125 (0.165)	0.169 (0.194)

Table 4.2 The averages and the standard deviations in parentheses of posterior probabilities

σ_1	σ_2	σ_3	μ_0	(n_1, n_2, n_3)	$P^F(H_1\|\mathbf{x},\mathbf{y})$	$P^{AI}(H_1\|\mathbf{x},\mathbf{y})$	$P^{MI}(H_1\|\mathbf{x},\mathbf{y})$
1.0	1.0	1.0	0.0	5,5,5	0.528 (0.129)	0.574 (0.189)	0.658 (0.150)
				5,10,10	0.586 (0.138)	0.625 (0.187)	0.715 (0.149)
				10,10,10	0.607 (0.144)	0.636 (0.190)	0.735 (0.151)
				10,10,20	0.640 (0.143)	0.659 (0.189)	0.765 (0.142)
			0.2	5,5,5	0.488 (0.161)	0.528 (0.216)	0.611 (0.190)
				5,10,10	0.507 (0.198)	0.544 (0.243)	0.633 (0.222)
				10,10,10	0.521 (0.201)	0.551 (0.240)	0.649 (0.221)
				10,10,20	0.504 (0.224)	0.522 (0.255)	0.627 (0.243)
			0.4	5,5,5	0.391 (0.192)	0.416 (0.238)	0.508 (0.225)
				5,10,10	0.331 (0.220)	0.352 (0.256)	0.439 (0.265)
				10,10,10	0.323 (0.237)	0.336 (0.267)	0.423 (0.280)
				10,10,20	0.246 (0.224)	0.256 (0.248)	0.338 (0.276)
			0.6	5,5,5	0.284 (0.193)	0.290 (0.232)	0.386 (0.235)
				5,10,10	0.164 (0.183)	0.172 (0.205)	0.237 (0.237)
				10,10,10	0.123 (0.164)	0.124 (0.176)	0.181 (0.214)
				10,10,20	0.059 (0.109)	0.062 (0.122)	0.090 (0.154)
			0.8	5,5,5	0.173 (0.168)	0.168 (0.193)	0.255 (0.219)
				5,10,10	0.058 (0.103)	0.058 (0.111)	0.095 (0.146)
				10,10,10	0.035 (0.084)	0.038 (0.095)	0.059 (0.125)
				10,10,20	0.008 (0.029)	0.009 (0.033)	0.015 (0.048)
			1.0	5,5,5	0.096 (0.121)	0.086 (0.131)	0.154 (0.168)
				5,10,10	0.017 (0.045)	0.015 (0.048)	0.030 (0.072)
				10,10,10	0.007 (0.028)	0.007 (0.030)	0.012 (0.043)
				10,10,20	0.001 (0.017)	0.001 (0.014)	0.002 (0.017)
1.0	1.0	3.0	0.0	5,5,5	0.534 (0.127)	0.575 (0.186)	0.652 (0.152)
				5,10,10	0.596 (0.142)	0.627 (0.185)	0.707 (0.153)
				10,10,10	0.614 (0.139)	0.645 (0.179)	0.733 (0.144)
				10,10,20	0.651 (0.147)	0.649 (0.185)	0.738 (0.149)
			0.2	5,5,5	0.510 (0.146)	0.548 (0.197)	0.628 (0.167)
				5,10,10	0.546 (0.181)	0.565 (0.219)	0.653 (0.196)
				10,10,10	0.540 (0.198)	0.567 (0.233)	0.659 (0.213)
				10,10,20	0.578 (0.202)	0.576 (0.230)	0.670 (0.206)
			0.4	5,5,5	0.438 (0.180)	0.469 (0.228)	0.552 (0.210)
				5,10,10	0.422 (0.220)	0.443 (0.251)	0.528 (0.246)
				10,10,10	0.363 (0.230)	0.384 (0.261)	0.470 (0.266)
				10,10,20	0.403 (0.251)	0.406 (0.267)	0.495 (0.272)
			0.6	5,5,5	0.349 (0.195)	0.363 (0.234)	0.451 (0.231)
				5,10,10	0.268 (0.213)	0.280 (0.237)	0.358 (0.252)
				10,10,10	0.213 (0.210)	0.221 (0.230)	0.289 (0.256)
				10,10,20	0.203 (0.219)	0.211 (0.233)	0.271 (0.257)
			0.8	5,5,5	0.248 (0.185)	0.249 (0.217)	0.338 (0.231)
				5,10,10	0.146 (0.165)	0.150 (0.181)	0.213 (0.213)
				10,10,10	0.086 (0.134)	0.088 (0.144)	0.128 (0.175)
				10,10,20	0.070 (0.120)	0.072 (0.128)	0.107 (0.160)
			1.0	5,5,5	0.164 (0.157)	0.159 (0.174)	0.243 (0.204)
				5,10,10	0.061 (0.106)	0.062 (0.117)	0.098 (0.147)
				10,10,10	0.032 (0.079)	0.033 (0.087)	0.051 (0.109)
				10,10,20	0.024 (0.065)	0.024 (0.065)	0.040 (0.089)
1.0	3.0	5.0	0.0	5,5,5	0.559 (0.124)	0.601 (0.170)	0.667 (0.142)
				5,10,10	0.613 (0.143)	0.614 (0.170)	0.682 (0.145)
				10,10,10	0.644 (0.130)	0.677 (0.161)	0.755 (0.128)
				10,10,20	0.670 (0.151)	0.662 (0.186)	0.748 (0.148)
			0.2	5,5,5	0.548 (0.137)	0.589 (0.183)	0.658 (0.153)
				5,10,10	0.591 (0.162)	0.595 (0.190)	0.665 (0.161)
				10,10,10	0.602 (0.167)	0.632 (0.197)	0.713 (0.171)
				10,10,20	0.633 (0.185)	0.628 (0.208)	0.711 (0.185)
			0.4	5,5,5	0.499 (0.162)	0.533 (0.207)	0.609 (0.183)
				5,10,10	0.535 (0.196)	0.542 (0.210)	0.616 (0.190)
				10,10,10	0.494 (0.220)	0.518 (0.251)	0.604 (0.238)
				10,10,20	0.531 (0.222)	0.533 (0.237)	0.617 (0.226)
			0.6	5,5,5	0.438 (0.182)	0.466 (0.220)	0.549 (0.206)
				5,10,10	0.463 (0.215)	0.477 (0.229)	0.553 (0.220)
				10,10,10	0.369 (0.234)	0.393 (0.258)	0.474 (0.265)
				10,10,20	0.401 (0.247)	0.405 (0.260)	0.486 (0.264)
			0.8	5,5,5	0.373 (0.188)	0.393 (0.221)	0.478 (0.214)
				5,10,10	0.363 (0.228)	0.374 (0.236)	0.454 (0.238)
				10,10,10	0.239 (0.208)	0.255 (0.229)	0.327 (0.249)
				10,10,20	0.226 (0.214)	0.232 (0.224)	0.299 (0.248)
			1.0	5,5,5	0.306 (0.190)	0.323 (0.219)	0.409 (0.222)
				5,10,10	0.265 (0.213)	0.284 (0.226)	0.359 (0.240)
				10,10,10	0.129 (0.154)	0.138 (0.173)	0.189 (0.202)
				10,10,20	0.121 (0.160)	0.126 (0.167)	0.174 (0.198)

Table 4.3 Percentage of albumin in plasma protein

Experiment	n_i	Mean	Variance
A	12	62.3	12.986
B	15	60.3	7.840
C	7	59.5	33.433
D	16	61.5	18.513

Table 4.4 p-value, Bayes factor and posterior probability of $H_1 : \mu = \mu_0$

μ_0	p_{GDE1}-value	p_{GDE2}-value	B_{21}^F	$P^F(H_1\|\mathbf{x})$
59.2	0.002	0.007	28.558	0.034
59.5	0.010	0.021	8.005	0.111
59.7643	0.033	0.050	3.033	0.248
59.8648	0.050	0.069	2.188	0.314
60.0	0.084	0.105	1.466	0.405
60.1479	0.142	0.162	1.000	0.500
60.5	0.391	0.403	.513	0.661

5. Concluding remarks

In this paper, we developed the objective Bayesian hypothesis testing procedures based on the fractional Bayes factor and the intrinsic Bayes factors for the common mean of several normal distributions under the reference priors. From our numerical results, the developed hypothesis testing procedures give fairly reasonable answers for all parameter configurations. However the MIBF favors the hypothesis H_1 than the FBF and the AIBF. From our simulation and example, we recommend the use of the FBF than the AIBF and MIBF for practical application in view of its simplicity and ease of implementation.

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*, edited by J.M. Bernardo, et al., Oxford University Press, Oxford, 35-60.

Berger, J. O. and Pericchi, L. R. (1996). The intrinsic Bayes factor for model selection and prediction. *Journal of the American Statistical Association*, **91**, 109-122.

Berger, J. O. and Pericchi, L. R. (1998). Accurate and stable Bayesian model selection: The median intrinsic Bayes factor. *Sankya B*, **60**, 1-18.

Berger, J. O. and Pericchi, L. R. (2001). Objective Bayesian methods for model selection: Introduction and comparison (with discussion). In *Model Selection, Institute of Mathematical Statistics Lecture Notes-Monograph Series*, **38**, edited by P. Lahiri, Beachwood, Ohio, 135-207.

Bhattacharya, C. G. (1980). Point and confidence estimation of a common mean and recovery of inter-block information. *The Annals of Statistics*, **2**, 963-976.

Brown, L. D. and Cohen, A. (1974). Estimation of a common mean and recovery of interblock information. *The Annals of Statistics*, **8**, 205-211.

Chang, C. H. and Pal, N. (2008). Testing on the common mean of several normal distributions. *Computational Statistics and Data Analysis*, **53**, 321-333.

Cohen A. and Sackrowitz, H. B. (1974). On estimating of the common mean of two normal populations. *The Annals of Statistics*, **2**, 1274-1282.

Cohen A. and Sackrowitz, H. B. (1989). Exact tests that recover interblock information in balanced incomplete blocks designs. *Journal of the American Statistical Association*, **84**, 556-560.

Eberhardt, K. R., Reeve, C. P. and Spiegelman, C. H. (1989). A minimax approach to combining means, with practical examples. *Chemometrics and Intelligent Laboratory Systems*, **5**, 129-148.

Fairweather, W. R. (1972). A method of obtaining an exact confidence interval for the common mean of several normal populations. *Applied Statistics*, **21**, 229-233.

Fisher, R. A. (1932). *Statistical methods of research workers*, Oliver & Boyd, London.

Graybill, F. A. and Deal, R. B. (1959). Combining unbiased estimators. *Biometrics*, **15**, 543-550.

Hartung, J., Knapp, G. and Sinha, B. K. (2008). *Statistical meta-analysis with applications*, Wiley, New York.

Jordan, S. M. and Krishnamoorthy, K. (1996). Exact confidence intervals for the common mean of several normal populations. *Biometrics*, **52**, 77-86.

Kang, S. G., Kim, D. H. and Lee, W. D. (2008). Bayesian model selection for inverse Gaussian populations with heterogeneity. *Journal of the Korean Data & Information Science Society*, **19**, 621-634.

Kang, S. G., Kim, D. H. and Lee, W. D. (2011). Default Bayesian testing for the bivariate normal correlation coefficient. *Journal of the Korean Data & Information Science Society*, **22**, 1007-1016.

Kharti, C. G. and Shah, K. R. (1974). Estimation of location parameters from two linear models under normality. *Communications in Statistics: Theory and Methods*, **3**, 647-663.

Krishnamoorthy, K. and Lu, Y. (2003). Inference on the common mean of several normal populations based on the generalized variable method. *Biometrics*, **59**, 237-247.

Kubokawa, T. (1990). Minimax estimation of common coefficient of several regression models under quadratic loss. *Journal of Statistical Planning and Inference*, **24**, 337-345.

Lee, W. D. and Kang, S. G. (2008). Bayesian multiple hypotheses testing for Poisson mean. *Journal of the Korean Data & Information Science Society*, **19**, 331-341.

Lin, S. H. and Lee, J. C. (2005). Generalized inferences on the common mean of several normal populations. *Journal of Statistical Planning and Inference*, **134**, 568-582.

Meier, P. (1953). Variance of weighted mean. *Biometrics*, **9**, 59-73.

Montgomery, D. C. (1991). *Design and analysis of experiments*, Wiley, New York.

O'Hagan, A. (1995). Fractional Bayes factors for model comparison (with discussion). *Journal of Royal Statistical Society B*, **57**, 99-118.

O'Hagan, A. (1997). Properties of intrinsic and fractional Bayes factors. *Test*, **6**, 101-118.

Shinozaki, N. (1978). A note on estimating the common mean of k normal distributions and the Stein problem. *Communications in Statistics: Theory and Methods*, **7**, 1421-1432.

Sinha, B. K. (1985). Unbiased estimation of the variance of the Graybill-Deal estimator of the common mean of several normal populations. *Canadian Journal of Statistics*, **13**, 243-247.

Sinha, B. K. and Mouqadem, O. (1982). Estimation of the common mean of two univariate normal populations. *Communications in Statistics: Theory and Methods*, **11**, 1603-1614.

Snedecor, G. W. (1950). The statistical part of the scientific method. *Annals of the New York Academy of Science*, **52**, 742-749.

Spiegelhalter, D. J. and Smith, A. F. M. (1982). Bayes factors for linear and log-linear models with vague prior information. *Journal of Royal Statistical Society B*, **44**, 377-387.

Yu, P. L., Sun, Y. and Sinha, B. K. (1999). On exact confidence intervals for the common mean of several normal populations. *Journal of Statistical Planning and Inference*, **81**, 263-277.

Zacks, S. (1966). Unbiased estimation of the common mean. *Journal of the American Statistical Association*, **61**, 467-476.

Zacks, S. (1970). Bayes and fiducial equivariant estimators of the common means of two populations. *The Annals of Mathematical Statistics*, **41**, 59-69.

Zhou, L. and Mathew, T. (1993). Combining independent tests in linear models. *Journal of the American Statistical Association*, **88**, 650-655.