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.

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 $H_1, H_2, \ldots, H_q$ are under consideration, with the data $x = (x_1, x_2, \ldots, x_n)$ having probability density function $f_i(x|\theta_i)$ under hypothesis $H_i$. The parameter vector $\theta_i$ 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 $H_i$, $i = 1, 2, \ldots, q$. Then the posterior probability that the hypothesis $H_i$ is true is

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

(2.1)

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

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

(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 $\pi_i^N$. Common choices are the uniform prior, Jeffreys’ prior and the reference prior. The noninformative prior $\pi_i^N$ is typically improper. Hence the use of noninformative prior $\pi_i^N$ 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 $x(l)$ denote the part of the data to be so used and let $x(-l)$ be the remainder of the data, such that

$$0 < m_i^N(x(l)) < \infty, i = 1, \ldots, q.$$  

(2.3)

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

$$B_{ji}(l) = \frac{\int f(x(-l)|\theta_i, x(l))\pi^N_i(\theta_i|x(l))d\theta_i}{\int f(x(-l)|\theta_i, x(l))\pi^N_i(\theta_i|x(l))d\theta_i} = B_{ji}^N \cdot B_{ji}^N(x(l))$$

(2.4)

where

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

and

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

are the Bayes factors that would be obtained for the full data $x$ and training samples $x(l)$, respectively.
Berger and Pericchi (1996) proposed the use of a minimal training sample to compute $B_{ji}^{l}(x(l))$. 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

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

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

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

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 $B_{ji}^{M}$ and $B_{ji}^{F}$ from (2.5) and (2.6), respectively.

The fractional Bayes factor (O’Hagan, 1995) is based on a similar intuition to that behind the fractional 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, $L(\theta_i) = f_i(x|\theta_i)$, with the remaining $1 - b$ fraction of the likelihood used for model discrimination. Then the fractional Bayes factor (FBF) of hypothesis $H_i$ versus hypothesis $H_j$ is

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

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

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

where $x = (x_1, \ldots, x_k), x_i = (x_{i1}, \ldots, x_{in_i}), i = 1, \ldots, k$ 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

$$L_1(\sigma_1, \ldots, \sigma_k|x) = (\sqrt{2\pi})^{-n}(\prod_{i=1}^{k} \sigma_i^{-n_i}) \exp \left\{-\sum_{i=1}^{k} \frac{1}{2\sigma_i^2} \left[S_i + n_i(\bar{x}_i - \mu_0)^2\right]\right\}, \quad (3.2)$$
where \( S_i^2 = \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 \) and \( \bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}, i = 1, \ldots, k. \) And under the hypothesis \( H_1, \) the reference prior for \((\sigma_1, \ldots, \sigma_k)\) is
\[
\pi_1^N(\sigma_1, \ldots, \sigma_k) \propto \prod_{i=1}^{k} \sigma_i^{-1}. \tag{3.3}
\]

Then from the likelihood (3.2) and the reference prior (3.3), the element \( m_1^h(x) \) of the FBF under \( H_1 \) is given by
\[
m_1^h(x) = \int_0^\infty \cdots \int_0^\infty L_1^h(\sigma_1, \ldots, \sigma_k|x)\pi_1^N(\sigma_1, \ldots, \sigma_k)d\sigma_1 \cdots d\sigma_k
= (\sqrt{2\pi})^{-bn} 2^{-k} \prod_{i=1}^{k} \Gamma \left[ \frac{bn_i}{2} \right] \left\{ \frac{b_i S_i^2 + n_i (\bar{x}_i - \mu_0)^2}{2} \right\}^{-bn}. \tag{3.4}
\]

For the hypothesis \( H_2, \) the reference prior for \((\mu, \sigma_1, \ldots, \sigma_k)\) is
\[
\pi_2^N(\mu, \sigma_1, \ldots, \sigma_k) \propto \prod_{i=1}^{k} \sigma_i^{-1}. \tag{3.5}
\]

The likelihood function under the hypothesis \( H_2 \) is
\[
L_2(\mu, \sigma_1, \ldots, \sigma_k|x) = \left( \sqrt{2\pi} \right)^{-n} \left\{ \prod_{i=1}^{k} \sigma_i^{-n_i} \right\} \exp \left\{ -\sum_{i=1}^{k} \frac{1}{2\sigma_i^2} \left[ S_i^2 + n_i (\bar{x}_i - \mu)^2 \right] \right\}. \tag{3.6}
\]

Thus from the likelihood (3.6) and the reference prior (3.5), the element \( m_2^h(x) \) of FBF under \( H_2 \) is given as follows.
\[
m_2^h(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L_2^h(\mu, \sigma_1, \ldots, \sigma_k|x)\pi_2^N(\mu, \sigma_1, \ldots, \sigma_k)d\sigma_1 \cdots d\sigma_k d\mu
= (\sqrt{2\pi})^{-bn} 2^{-k} \prod_{i=1}^{k} \Gamma \left[ \frac{bn_i}{2} \right] \int_{-\infty}^{\infty} \left\{ \frac{b_i S_i^2 + n_i (\bar{x}_i - \mu)^2}{2} \right\}^{-bn} d\mu. \tag{3.7}
\]

Therefore the element \( B_{21}^N \) of FBF is given by
\[
B_{21}^N = \frac{S_2(x)}{S_1(x)}, \tag{3.8}
\]

where
\[
S_1(x) = \prod_{i=1}^{k} \left\{ S_i^2 + n_i (\bar{x}_i - \mu_0)^2 \right\}^{-n_i/2}.
\]

and
\[
S_2(x) = \int_{-\infty}^{\infty} \prod_{i=1}^{k} \left\{ S_i^2 + n_i (\bar{x}_i - \mu)^2 \right\}^{-n_i/2} d\mu.
\]
And the ratio of marginal densities with fraction $b$ is
\[ \frac{m_1^b(x)}{m_2^b(x)} = \frac{S_1(x; b)}{S_2(x; b)} \] (3.9)
where
\[ S_1(x; b) = \prod_{i=1}^{k} \{ S_i^2 + n_i(\bar{x}_i - \mu_0)^2 \}^{-\frac{b_n}{2}} \]
and
\[ S_2(x; b) = \int_{-\infty}^{\infty} \prod_{i=1}^{k} \{ S_i^2 + n_i(\bar{x}_i - \mu)^2 \}^{-\frac{b_n}{2}} d\mu. \]
Thus the FBF of $H_2$ versus $H_1$ is given by
\[ B_{21}^F = \frac{S_2(x)}{S_1(x)} \frac{S_1(x; b)}{S_2(x; b)}. \] (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 $X_1 = (X_{i1}, X_{i2})$, $i, j (i < j) = 1, 2, \ldots, n_l$ be a random sample of size 2 from population $l, l = 1, 2, \ldots, k$. The marginal density of $X_1, X_2, \ldots, X_k$ under the hypothesis $H_1$ with prior (3.3) is
\[ m_1^N(x_1, x_2, \ldots, x_k) \]
\[ = \int_{0}^{\infty} \cdots \int_{0}^{\infty} f(x_1, x_2, \ldots, x_k | \sigma_1, \ldots, \sigma_k) \pi_1^N(\sigma_1, \ldots, \sigma_k) d\sigma_1 \cdots d\sigma_k \]
\[ = \left( \frac{\sqrt{2\pi}}{\sigma} \right)^{-2k} 2^k \prod_{i=1}^{k} \left[ (x_{i1} - x_{i2})^2 + (x_{i1} + x_{i2} - 2\mu_0)^2 \right]^{-1}, \]
where $x_i = (x_{i1}, x_{i2})$, $i, j (i < j) = 1, 2, \ldots, n_l, l = 1, 2, \ldots, k$.
And the marginal density $m_2^N(x_1, x_2, \ldots, x_k)$ under $H_2$ is given by
\[ m_2^N(x_1, x_2, \ldots, x_k) \]
\[ = \int_{-\infty}^{\infty} \cdots \int_{0}^{\infty} f(x_1, x_2, \ldots, x_k | \mu, \sigma_1, \ldots, \sigma_k) \pi_2^N(\mu, \sigma_1, \ldots, \sigma_k) d\sigma_1 \cdots d\sigma_k d\mu \]
\[ = \left( \frac{\sqrt{2\pi}}{\sigma} \right)^{-2k} 2^k \int_{-\infty}^{\infty} \prod_{i=1}^{k} \left[ (x_{i1} - x_{i2})^2 + (x_{i1} + x_{i2} - 2\mu)^2 \right]^{-1} d\mu. \]
Since the marginal densities \( m_1^N \) and \( m_2^N \) are finite, the minimal training sample is \( x_i = (x_{ij}, x_{lj}), i, j (i < j) = 1, 2, \ldots, n_i, l = 1, 2, \ldots, k \). 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

\[
B_{21}^{AI} = \frac{S_2(x)}{S_1(x)} \left[ \frac{1}{L} \sum_{i=1}^{n} \sum_{1<i<j}^{n_2} \sum_{k_i<k_j}^{n_k} \frac{T_1(x_{1i}, x_{1j}, x_{2i}, x_{2j}, \ldots, x_{ki}, x_{kj})}{T_2(x_{1i}, x_{1j}, x_{2i}, x_{2j}, \ldots, x_{ki}, x_{kj})} \right],
\]

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

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

and

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

Also the MIBF of \( H_2 \) versus \( H_1 \) is given by

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

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

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

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

\[
\hat{\mu}_{GDE} = \sum_{i=1}^{k} (n_i/s_i^2) \bar{x}_i / \sum_{i=1}^{k} (n_i/s_i^2),
\]

\[
\hat{\nu}_{(1)}(\hat{\mu}_{GDE}) = \left( \sum_{i=1}^{k} n_i/s_i^2 \right)^{-1} \left[ 1 + 4 \sum_{i=1}^{k} (n_i + 1)^{-1} (n_i/s_i^2) \right]
\]

\[
\times \left[ \sum_{i=1}^{k} (n_i/s_i^2) - (n_i/s_i^2)^2 / \left( \sum_{i=1}^{k} n_i/s_i^2 \right)^2 \right]^{-1}.
\]

And the GDE2 test is

 Reject \( H_1 \) if \( \Delta_{GDE2} = (\hat{\mu}_{GDE} - \mu_0)^2/\hat{\nu}(\hat{\mu}_{GDE}) > (t_{l, (\alpha/2)})^2 \),
where
\[
\tilde{V}(\hat{\mu}_{\text{GDE}}) = \left( \sum_{i=1}^{k} \frac{n_i}{s_i^2} \right)^{-2} \sum_{i=1}^{k} \frac{(n_i/s_i^2)_{2F1}}{(1,2;(n_i+1)/2;1-(n_i/s_i^2)/\left(\sum_{i=1}^{k} n_i/s_i^2\right))},
\]
\[
t_i(\alpha/2) = \{t_{i\lceil\alpha/2\rceil} - t_{i\lceil\alpha/2\rceil}(l-\lfloor l\rfloor)\},
\]
\[
l \approx \left( \sum_{i=1}^{k} \frac{s_i^2/n_i}{\sum_{i=1}^{k} (s_i^2/n_i)} \right)^2 + \left( \sum_{i=1}^{k} \frac{(s_i^2/n_i)^2}{(n_i-1)} \right),
\]
and here $2F1(a,b;c;z)$ is the Gaussian hypergeometric function and $\lfloor l\rfloor$ 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 $(\mu, \sigma_1, \ldots, \mu, \sigma_k)$ and $(n_1, \ldots, n_k)$. In particular, for fixed $(\mu, \sigma_1, \ldots, \mu, \sigma_k)$, we take 1,000 independent random samples of $X_i$ with sample size $n_i$ from $N(\mu, \sigma_i^2), i = 1, \ldots, k$. 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 $\sigma_1, \sigma_2$ and $\mu_0$, 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, $P^F(\cdot), P^Al(\cdot)$ and $P^M(\cdot)$ 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 $\mu_0$ are close to 0, whereas reject the hypothesis $H_1$ when the values of $\mu_0$ 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 $\mu_0$ that are close to 60.5, any criteria accept the $H_1$. However for values of $\mu_0$ 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.
| $n_1$ | $n_2$ | $n_0$ | $P(P_{x,y}|x,y)$ | $P(P_{x,y}^1|x,y)$ | $P(P_{x,y}^2|x,y)$ |
|-------|-------|-------|------------------|------------------|------------------|
| 613   |       |       |                  |                  |                  |
| 1.0   | 1.0   |       |                  |                  |                  |
| 1.0   | 2.0   |       |                  |                  |                  |
| 2.0   | 3.0   |       |                  |                  |                  |

Table 4.1 The averages and the standard deviations in parentheses of posterior probabilities.
<table>
<thead>
<tr>
<th>$\sigma_1$</th>
<th>$\sigma_2$</th>
<th>$\sigma_3$</th>
<th>$\mu_1$</th>
<th>$\mu_2$</th>
<th>$\mu_3$</th>
<th>$p_{11}(x, y)$</th>
<th>$p_{12}(x, y)$</th>
<th>$p_{13}(x, y)$</th>
<th>$p_{21}(x, y)$</th>
<th>$p_{22}(x, y)$</th>
<th>$p_{23}(x, y)$</th>
<th>$p_{31}(x, y)$</th>
<th>$p_{32}(x, y)$</th>
<th>$p_{33}(x, y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 1.0 1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>1.3</td>
<td>1.3</td>
<td>1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.6</td>
<td>1.6</td>
<td>1.6</td>
<td>1.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>1.7</td>
<td>1.7</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.9</td>
<td>1.9</td>
<td>1.9</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2 The averages and the standard deviations in parentheses of posterior probabilities.
Table 4.3 Percentage of albumin in plasma protein

<table>
<thead>
<tr>
<th>Experiment</th>
<th>n_i</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>12</td>
<td>62.3</td>
<td>12.986</td>
</tr>
<tr>
<td>B</td>
<td>15</td>
<td>60.3</td>
<td>7.840</td>
</tr>
<tr>
<td>C</td>
<td>7</td>
<td>59.5</td>
<td>33.433</td>
</tr>
<tr>
<td>D</td>
<td>16</td>
<td>61.5</td>
<td>18.513</td>
</tr>
</tbody>
</table>

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

| $\mu_0$ | $p_{GDE1}$-value | $p_{GDE2}$-value | $B_{12}$ | $P^* (H_1|x)$ |
|---------|------------------|------------------|----------|---------------|
| 59.2    | 0.002            | 0.007            | 28.538   | 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            | 0.513    | 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


