Journal of the Korean Statistical Society (2000), **29**: *4, pp 443–454*

Intrinsic Priors for Testing Two Normal Means with the Default Bayes Factors

Jongsig Bae [1], Hyunsoo Kim [2] and Seong W. Kim [3]

ABSTRACT

In Bayesian model selection or testing problems of different dimensions, the conventional Bayes factors with improper noninformative priors are not well defined. The intrinsic Bayes factor and the fractional Bayes factor are used to overcome such problems by using a data-splitting idea and fraction, respectively. This article addresses a Bayesian testing for the comparison of two normal means with unknown variance. We derive proper intrinsic priors, whose Bayes factors are asymptotically equivalent to the corresponding fractional Bayes factors. We demonstrate our results with two examples.

Keywords: Default Bayes factor; Intrinsic Bayes factor; Fractional Bayes factor; Intrinsic prior; Noninformative prior.

1. INTRODUCTION

Bayesian model selection and testing problems have established progressive development of default Bayes factors that can be used in the lack of subjective prior information. Two very general such default Bayes factors are the fractional Bayes factor (FBF) of O'Hagan (1995) and the intrinsic Bayes factors (IBF) of Berger and Pericchi (1996). These methodologies have been applied in a wide variety of settings, and have undergone progressive study in situations involving testing of nested hypotheses or models.

Bayes factors under proper priors or informative priors have been successful in testing or model selection problems. However, limited information often require the use of noninformative priors such as Jeffreys's priors (Jeffreys, 1961)

[1]Department of Mathematics, SungKyunKwan University, Suwon, 440-746, Korea.
[2]Department of Mathematics, SungKyunKwan University, Suwon, 440-746, Korea.
[3]Department of Statistics, Seoul National University, Seoul, 151-742, Korea.

and reference priors (Berger and Bernado, 1992). In this article, we use the reference prior to test the equality of two independent normal means with unknown variance and we compare the numerical results of the arithmetic IBF and the FBF.

Suppose that the data $\mathbf{x}$ has a parametric distribution with density $f(\mathbf{x}|\xi_i)$, where ξ_i, $i = 1, 2$, is a vector of unknown parameters which belongs to the the parameter space Ξ_i, $i = 1, 2$, respectively. Let $\pi_i^N(\xi_i)$, $i = 1, 2$, be the improper prior density. The Bayes factor B_{21}^N of a model M_2 to a model M_1 is

$$B_{21}^N = \frac{m_2^N(\mathbf{x})}{m_1^N(\mathbf{x})} = \frac{\int_{\Xi_2} f(\mathbf{x}|\xi_2)\pi_2^N(\xi_2)d\xi_2}{\int_{\Xi_1} f(\mathbf{x}|\xi_1)\pi_1^N(\xi_1)d\xi_1}, \tag{1.1}$$

where $m_1^N(\mathbf{x})$ and $m_2^N(\mathbf{x})$ are the marginal densities under M_1 and M_2 respectively. Since both $\pi_1^N(\xi_1)$ and $\pi_2^N(\xi_2)$ are improper, they are defined only up to arbitrary constants, say c_1 and c_2 respectively. Thus, B_{21}^N is defined only up to (c_2/c_1), which is also arbitrary. So, the resulting Bayes factor is not well defined. This issue has been initially addressed by several authors including Geisser and Eddy (1979), Spiegelhalter and Smith (1982), and San Martini and Spezzaferri (1984).

The IBF criterion is based on a data-splitting method, which removes the arbitrariness of improper priors. There have been several articles written which use the IBF. Varshavsky (1996) made use of the IBF for a stationary autoregressive process. Lingham and Sivaganesan (1997) performed a test for the shape parameter of the power law process. Kim (2000) analyzed comparisons of two exponential means. Kim and Ibrahim (2000) derived an explicit form of the IBF for generalized linear models.

The difficulty of an IBF approach is due to its considerable computational expense; for large sample sizes it is fairly time-consuming. Another useful criterion is the FBF which is based on a similar intuition to that of the IBF. It is computed by exponentiating the likelihood to a power δ, where $0 \leq \delta \leq 1$. It is well defined as in the IBF method. Moreover, it does not require a heavy computation. The FBF version is thus much more computationally effective. So, in this paper, we concentrate on the FBF method with a specific choice of a fraction δ. We refer the reader to Kass and Raftery (1995) for further discussions about Bayes factors.

This paper is organized as follows. In Section 2, we review the two default Bayes factors. In Section 3, we consider the testing problem for comparing the two normal means with unknown variance. In particular, the study of intrinsic priors is restricted to the FBF and we derive proper intrinsic priors, whose Bayes factors are asymptotically equivalent to the corresponding FBFs. A real dataset and a simulated dataset are analyzed in Section 4. A brief conclusion is given in Section 5.

2. PRELIMINARIES

We have known that the Bayes factor B_{21}^N in (1.1) involves arbitrary constants. The methods for removing this arbitrariness are to use a subset of data, a training sample, and a portion of the likelihood, a fraction δ.

First, we discuss the method for removing this arbitrariness by a training sample. Let $\mathbf{x}(l)$ be a training sample and let $\mathbf{x}(-l)$ be the remainder of the data. First, compute the posterior $\pi_i^N(\xi|\mathbf{x}(l))$, then compute the Bayes factors with the $\mathbf{x}(-l)$ as data, using $\pi_i^N(\xi|\mathbf{x}(l))$ as the prior. Consequently, the Bayes factor of model M_2 to model M_1 is given as follows :

$$B_{21}(l) = B_{21}^N \cdot B_{12}^N(\mathbf{x}(l)), \tag{2.1}$$

where $B_{12}^N(\mathbf{x}(l)) = m_1^N(\mathbf{x}(l))/m_2^N(\mathbf{x}(l))$ is the Bayes factor for the training sample.

In practice, $\mathbf{x}(l)$ is chosen to be minimal in the sense that the marginal $m_i^N(\mathbf{x}(l))$ is finite for $i = 1, 2$, and no subset of $\mathbf{x}(l)$ gives finite marginals. Note that in (2.1), $B_{21}(l)$ does not depend on arbitrary constants, and thus is well defined. Furthermore, the Bayes factor defined by (2.1), depends on the choice of the minimal training sample. To avoid this dependence, Berger and Pericchi (1996) suggested to take the average of $B_{21}(l)$ over all $\mathbf{x}(l)$.

Definition 3.1 The arithmetic intrinsic (AI) Bayes factor of M_2 to M_1 is given by

$$B_{21}^{AI} = \frac{1}{L}\sum_{l=1}^{L} B_{21}(l) = B_{21}^N \cdot CFA_{12}, \tag{2.2}$$

where L is the number of all possible minimal training samples, and the correction factor is given by

$$CFA_{12} = \frac{1}{L}\sum_{l=1}^{L} B_{12}^N(\mathbf{x}(l)). \tag{2.3}$$

Second, we introduce the method for removing the arbitrariness in (1.1) by a portion of the likelihood with the fraction δ. O'Hagan (1995) proposed the fractional Bayes factor (FBF) as a default Bayes factor. The FBF of model M_2 to model M_1 is

$$B_{21}^F = B_{21}^N \cdot CFR_{12}(\delta), \tag{2.4}$$

where the correction factor is defined by

$$CFR_{12}(\delta) = \frac{\int_{\Xi_1} [L_1(\xi_1)]^\delta \, \pi_1^N(\xi_1) d\xi_1}{\int_{\Xi_2} [L_2(\xi_2)]^\delta \, \pi_2^N(\xi_2) d\xi_2}.$$

Here, $L_i(\xi_i)$ is the likelihood function under model M_i, $i = 1, 2$ and δ, known as a fraction of the likelihood function, is a number between 0 and 1. A commonly suggested choice is $\delta = m/n$, where m is the size of the minimal training sample proposed by Berger and Pericchi (1996) and n is the size of the whole sample. We will use this choice in our problems. However, the choice of δ may vary to specify for obtaining a stable Bayes factor. See O'Hagan (1995), Berger and Pericchi (1998), and among others.

It is of considerable interest to see the asymptotic behavior of default Bayes factors to real Bayes factors which can be computed with prior distributions, often called *intrinsic priors*. This can detect systematic biases of default Bayes factors towards one of the hypotheses. Further, intrinsic priors can be directly used to compute Bayes factors especially for small sample sizes. This issue was established by Berger and Pericchi (1996), and several intrinsic priors were derived in various settings. See Dmochowski (1996), Lingham and Sivaganesan (1997), Berger and Mortera (1999), and Kim (2000) for related work.

Under the regularity conditions in Berger and Pericchi (1996), a set of intrinsic priors denoted by (π_1^I, π_2^I) is a solution of the following system of equations:

$$\begin{cases} \dfrac{\pi_2^I(\phi_2(\xi_1))\pi_1^N(\xi_1)}{\pi_2^N(\phi_2(\xi_1))\pi_1^I(\xi_1)} &= B_1^*(\xi_1), \\ \dfrac{\pi_2^I(\xi_2)\pi_1^N(\phi_1(\xi_2))}{\pi_2^N(\xi_2)\pi_1^I(\phi_1(\xi_2))} &= B_2^*(\xi_2), \end{cases} \tag{2.5}$$

where for $i \neq j$,

$$\phi_i(\xi_j) = \lim_{n\to\infty} E_{\xi_j}^{M_j}(\hat{\xi}_i) \text{ under } M_j.$$

Here, $\hat{\xi}_i$ is the MLE under M_i, and for $i = 1, 2$,

$$B_i^*(\xi_i) = \lim_{n\to\infty} CF \text{ under } M_i,$$

with CF being CFA_{12} or $CFR_{12}(\delta)$.

Remark 1 The noninformative priors $\pi_1^N(\xi_1)$ and $\pi_2^N(\xi_2)$ are called starting priors. We note that solutions are not necessarily unique nor proper. It is of interest to find proper intrinsic priors for given starting priors. Once we derive proper intrinsic priors, the fractional Bayes factor B_{21}^F can be replaced by the ordinary Bayes factors B_{21}^I computed with intrinsic priors at least asymptotically.

For the nested model comparison, i.e. when $\Xi_1 \subseteq \Xi_2$, from (2.5) the intrinsic prior $\pi_2^I(\xi_2)$ can be derived as

$$\pi_2^I(\xi_2) = \pi_2^N(\xi_2)\pi_1^I(\phi_1(\xi_2))B_2^*(\xi_2)/\pi_1^N(\phi_1(\xi_2)).$$

Obviously $\pi_2^I(\xi_2)$ depends on the choice of $\pi_1^I(\xi_1)$. So there could be a class of intrinsic priors. Kim (2000) found a class of intrinsic priors for testing two exponential means. Furthermore, it is of interest to see whether $\pi_2^I(\xi_2)$ is proper, since one cannot guarantee the propriety of $\pi_2^I(\xi_2)$. Thus, it is different from Theorem 1 of Berger and Pericchi (1996).

3. TESTING NORMAL MEANS

For the general location-scale model, closed expressions for the default Bayes factors are not typically obtainable. We thus consider only the normal model, where closed expressions can be found and allow an interesting comparison of the two default Bayes factors.

Suppose that we have independent observations $x_{ij} \sim N(\mu_i, \theta)$, $i = 1, 2$; $j = 1, 2 \ldots, n_i$ with unknown variance θ. We use the following notation throughout this paper. Let $\Theta_1 = \{(\mu, \theta) | -\infty < \mu < \infty, 0 < \theta < \infty\}$ and $\Theta_2 = \{(\mu_1, \mu_2, \theta) | -\infty < \mu_1, \mu_2 < \infty, 0 < \theta < \infty\}$. Let $N = n_1 + n_2$, and let $\bar{\mathbf{x}}_i = \sum_{j=1}^{n_i} x_{ij}/n_i$ and $\mathbf{s}_i^2 = \sum_{j=1}^{n_i}(x_{ij} - \bar{\mathbf{x}}_i)^2, i = 1, 2$. Assume that $n_1/N \to a$ as $N \to \infty$. Consider the following models,

$$M_1 : \mu_1 = \mu_2, \text{ and } M_2 : \mu_1 \neq \mu_2.$$

Let μ denote the common value of μ_i under M_1. We start with reference prior for each model, i.e., $\pi_1^N(\mu, \theta) = 1/\theta$ and $\pi_2^N(\mu_1, \mu_2, \theta) = 1/\theta$, respectively. The Bayes factor for the full sample is

$$B_{21}^N = \sqrt{\pi}\left(\frac{n_1 n_2}{N}\right)^{-1/2} \frac{\Gamma(N/2-1)}{\Gamma((N-1)/2)} \frac{\{\mathbf{s}_1^2 + \mathbf{s}_2^2 + n_1 n_2(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^2/N\}^{(N-1)/2}}{\{\mathbf{s}_1^2 + \mathbf{s}_2^2\}^{N/2-1}}. \quad (3.1)$$

The minimal training samples are size of 3 (see Remark 2), i.e. $\mathbf{x}_1(l) = (x_{1i}, x_{2j_1}, x_{2j_2})$ and $\mathbf{x}_2(l) = (x_{1i_1}, x_{1i_2}, x_{2j})$. Then the marginals for $\mathbf{x}_1(l)$ and $\mathbf{x}_2(l)$ are

$$\begin{cases} m_1^N(\mathbf{x}_1(l)) = \frac{\sqrt{3}}{\pi}\{(x_{1i} - x_{2j_1})^2 + (x_{2j_1} - x_{2j_2})^2 + (x_{2j_2} - x_{1i})^2\}^{-1}, \\ m_2^N(\mathbf{x}_1(l)) = |x_{2j_1} - x_{2j_2}|^{-1}, \end{cases}$$

and

$$\begin{cases} m_1^N(\mathbf{x}_2(l)) = \frac{\sqrt{3}}{\pi}\{(x_{1i_1} - x_{1i_2})^2 + (x_{1i_2} - x_{2j})^2 + (x_{2j} - x_{1i_1})^2\}^{-1}, \\ m_2^N(\mathbf{x}_2(l)) = |x_{1i_1} - x_{1i_2}|^{-1}. \end{cases}$$

Thus, the AI Bayes factor is

$$B_{21}^{AI} = B_{21}^N \cdot CFA_{12}, \tag{3.2}$$

where the correction factor is given by

$$CFA_{12} = \frac{1}{2}\left(\frac{1}{L_1}\sum_{l=1}^{L_1} B_{12}^N(\mathbf{x}_1(l)) + \frac{1}{L_2}\sum_{l=1}^{L_2} B_{12}^N(\mathbf{x}_2(l))\right),$$

with $B_{12}^N(\mathbf{x}_i(l))$ being the Bayes factor for the training sample $\mathbf{x}_i(l)$, and L_i being the number of all possible minimal training samples $\mathbf{x}_i(l)$ for $i = 1, 2$.

Remark 2 Note that if we take one observation from each population, the marginal density for M_1 is finite. However, for such a training sample, the marginal density for M_2 is not finite. So, we need one more observation. Now, let $\mathbf{x}_1(l) = (x_{1i}, x_{2j_1}, x_{2j_2})$ and $\mathbf{x}_2(l) = (x_{1i_1}, x_{1i_2}, x_{2j})$ be training sample for $i, i_k \in \{1, \cdots, n_1\}$ and $j, j_k \in \{1, \cdots, n_2\}$, $k = 1, 2$. Since the marginals for training samples $\mathbf{x}_1(l)$ and $\mathbf{x}_2(l)$ are finite and no subset gives finite marginal, the size of the minimal training sample is 3.

Meanwhile, the fractional Bayes factor B_{21}^F of M_2 to M_1 is

$$B_{21}^F = B_{21}^N \cdot CFR_{12}(\delta), \tag{3.3}$$

where B_{21}^N is in (3.1) and the correction factor at $\delta = 3/N$ is

$$CFR_{12}\left(\frac{3}{N}\right) = \frac{1}{\pi}\left(\frac{n_1}{N}\frac{n_2}{N}\right)^{1/2} \frac{\{(\mathbf{s}_1^2 + \mathbf{s}_2^2)/N - n_1 n_2(\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2)^2/N^2\}^{-1}}{\{(\mathbf{s}_1^2 + \mathbf{s}_2^2)/N\}^{-1/2}}.$$

We need to compute $B_2^*(\mu_1, \mu_2, \theta)$ to drive intrinsic priors.

Proposition 3.1 The quantity $B_2^*(\mu_1, \mu_2, \theta)$ is

$$B_2^*(\mu_1, \mu_2, \theta) = \frac{\sqrt{ab}}{\pi}\{\theta + ab(\mu_1 - \mu_2)^2\}^{-1}\theta^{1/2}, \tag{3.4}$$

where $b = 1 - a$.

Proof. The result immediately follows from the strong law of large numbers. □

After taking the limit, the set of intrinsic priors is

$$\begin{cases} \pi_1^I(\mu,\theta) = g(\mu,\theta), \ (\mu,\theta) \in \Theta_1, \\ \pi_2^I(\mu_1,\mu_2,\theta) = g(a\mu_1 + b\mu_2, \theta) \cdot B_2^*(\mu_1,\mu_2,\theta), \ (\mu_1,\mu_2,\theta) \in \Theta_2, \end{cases} \tag{3.5}$$

where B_2^* is given by (3.4) and $g(\mu, \theta)$ is any proper density for $(\mu, \theta) \in \Theta_1$.

Theorem 3.1 The intrinsic prior π_2^I in (3.5) is proper.

Proof. Let $s = a\mu_1 + b\mu_2$, $t = \mu_1 - \mu_2$. Then

$$\begin{aligned} \int\int\int_{\Theta_2} \pi_2^I(\mu_1,\mu_2,\theta)d\mu_2 d\mu_1 d\theta &= \int\int\int_{\Theta_2} \frac{\sqrt{ab}}{\pi}(\theta + ab(\mu_1-\mu_2)^2)^{-1}\theta^{1/2} \\ &\qquad g(a\mu_1 + b\mu_2,\theta)d\mu_1 d\mu_2 d\theta \\ &= \int\int\int_{\Theta_2} \frac{\sqrt{ab}}{\pi}(\theta + abt^2)^{-1}\theta^{1/2} g(s,\theta)dtdsd\theta \\ &= \int_0^\infty \int_{-\infty}^\infty g(s,\theta)dsd\theta \\ &= 1. \end{aligned}$$

□

Corollay 3.1 When $g(\mu,\theta)$ is given by $g(\mu,\theta) = g_1(\mu|\theta)g_2(\theta)$ with $\mu|\theta \sim N(0, \tau\theta)$ and $\theta \sim$ *Inverse Gamma*(λ, η), the set of intrinsic priors is

$$\begin{cases} \pi_1^I(\mu,\theta) = (2\pi\tau\theta)^{-1/2}\exp\{-\frac{\mu^2}{2\tau\theta}\} \cdot \frac{\eta^\lambda}{\Gamma(\lambda)\theta^{\lambda+1}}\exp\{-\frac{\eta}{\mu}\}, \ \text{under } \Theta_1, \\ \pi_2^I(\mu_1,\mu_2,\theta) = g(a\mu_1 + b\mu_2,\theta) \cdot B_2^*(\mu_1,\mu_2,\theta), \ \text{under } \Theta_2. \end{cases} \tag{3.6}$$

where

$$g(a\mu_1 + b\mu_2, \theta) = (2\pi\tau\theta)^{-1/2}\exp\{-\frac{(a\mu_1 + b\mu_2)^2}{2\tau\theta}\} \cdot \frac{\eta^\lambda}{\Gamma(\lambda)\theta^{\lambda+1}}\exp\{-\frac{\eta}{\theta}\}.$$

Remark 3 The joint prior $\pi_1^I(\mu,\theta)$ in (3.5) is from a conjugate family for the distribution of $(\mathbf{x_1}, \mathbf{x_2})$, where $\mathbf{x_1} = (x_{11}, \ldots, x_{1n_1})$ and $\mathbf{x_2} = (x_{21}, \ldots, x_{2n_2})$.

This prior is commonly used for which the data follow the normal distributions (cf. Berger (1980)).

Remark 4 We were not able to derive intrinsic priors for the IBF case, since the closed form of B_2^* is unobtainable. As O'Hagan (1995) suggested, there are two more different choices of δ. However, for those choices, it is impossible to derive intrinsic priors in this setting. If we restrict the parameter space, it could be possible to compute the Bayes factor by the intrinsic limiting procedure of Moreno et *al.* (1998).

4. NUMERICAL RESULTS

Example 1: The following data are given by Rohatgi (1976). During World War II bacterial polysaccharides were investigated as blood plasma extenders. Sixteen samples of hydrolixed polusaccharides supplied by various manufacturers in order to assess two chemical methods for determining the average molecular weight yielded the following results (data unit is 1000):

MethodA	62.7	29.1	44.4	47.8	36.3	40.0	43.4	35.8
	33.9	44.2	34.3	31.3	38.4	47.1	42.1	42.2
MethodB	56.4	27.5	42.2	46.8	33.3	37.1	37.3	36.2
	35.2	38.0	32.2	27.3	36.1	43.1	38.4	39.9

Let μ_1 and μ_2 denote the mean molecular weights for Method A and Method B respectively. Suppose that we want to test $M_1 : \mu_1 = \mu_2$ versus $M_2 : \mu_1 \neq \mu_2$. Here $(n_1, n_2, \hat{\mu}_1, \hat{\mu}_2) = (16, 16, 40.813, 37.938)$, where $\hat{\mu}_i$ is the MLE of μ_i for $i = 1, 2$. We computed the fractional Bayes factor and the Bayes factors using the set of intrinsic priors given by (3.6) with five choices of τ and three choices of (λ, η), respectively. They are $\tau = 1$, 10, 50, 100, 500 and (0.01, 0.01), (0.1, 0.1), and (1.0, 1.0). The computation requires a two-dimensional numerical integration. This can be done by the IMSL routines linked to Fortran 77. The numerical values are reported in Table 1. The Bayes factors with intrinsic priors are nearly free of hyperparameters (λ, η) but they get close to B_{21}^F as τ increases. We notice that starting priors for both M_1 and M_2 do not depend on location parameters. Thus, as τ increases, the effect of the conditional prior $\mu|\theta$ in Corollary 3.1 decreases. Since the Bayes factors are less than 1, one may conclude that the difference between the two chemical methods is fairly small. Furthermore, there is not much difference between each value of (λ, η). Therefore, the Bayes factors using intrinsic priors are quite robust in terms of the hyperparameters $\tau > 10$ and (λ, η).

Table 1: Bayes factors for testing $M_1 : \mu_1 = \mu_2$ versus $M_2 : \mu_1 \neq \mu_2$

				(λ, η)		
$(\hat{\mu}_1, \hat{\mu}_2)$	B_{21}^{AI}	B_{21}^{F}	τ	$(0.01, 0.01)$	$(0.1, 0.1)$	$(1.0, 1.0)$
			1	0.186	0.186	0.190
			10	0.231	0.231	0.238
$(40.81, 37.94)$	0.361	0.253	50	0.239	0.240	0.247
			100	0.240	0.241	0.249
			500	0.241	0.242	0.250

Example 2: We performed a simulation study for testing M_1 versus M_2. We examined the cases when $\mu_1 = \mu_2 = 0$ for some choices of n_1 and n_2. We computed the average of the relative differences between the FBF and the Bayes factors with intrinsic priors for choices of (λ, η) and τ given by (3.6). We used 200 replication to see the stability of numerical values. We also computed the standard deviations of relative differences based on 200 replication. The numerical values are reported in Table 2 and Table 3. The relative differences are quite small for each simulated dataset. Especially, as the sample size increases, the relative difference decreases. This is what we would expect from the theoretical results. We also note that the values are quite stable.

Table 2: Default Bayes factors and MLE's (Numerical values are averaged over 200 replication.)

(n_1, n_2)	$(\hat{\mu}_1, \hat{\mu}_2)$	B_{21}^{I}	B_{21}^{F}
$(5, 5)$	$(1.024, 1.011)$	0.721	0.629
$(10, 10)$	$(0.999, 0.971)$	0.611	0.459
$(20, 10)$	$(0.979, 1.010)$	0.585	0.342
$(20, 20)$	$(1.019, 0.980)$	0.488	0.386
$(30, 30)$	$(1.017, 0.976)$	0.320	0.249

Table 3: Relative difference $|B^I_{21} - B^F_{21}|/B^F_{21}$ for estimating the fractional Bayes factor. The relative difference (R.D.) is averaged over 200 replication. The numbers in parentheses are the standard deviations of the relative differences, based on 200 replication

	(λ, η)	(.01, .01)	(.1, .1)	(1.0, 1.0)
(n_1, n_2)	τ	$R.D(STE)$	$R.D(STE)$	$R.D(STE)$
(5, 5)	1	0.153(0.067)	0.154(0.065)	0.177(0.070)
(10, 10)		0.081(0.029)	0.081(0.028)	0.087(0.030)
(20, 10)		0.054(0.018)	0.054(0.018)	0.057(0.019)
(20, 20)		0.042(0.014)	0.042(0.014)	0.045(0.015)
(30, 30)		0.028(0.009)	0.028(0.009)	0.030(0.008)
	10	0.155(0.073)	0.154(0.067)	0.175(0.069)
		0.081(0.030)	0.081(0.029)	0.086(0.030)
		0.054(0.018)	0.054(0.018)	0.057(0.019)
		0.042(0.015)	0.042(0.015)	0.044(0.015)
		0.028(0.009)	0.028(0.009)	0.030(0.008)
	50	0.156(0.074)	0.154(0.068)	0.174(0.069)
		0.081(0.030)	0.081(0.029)	0.086(0.030)
		0.054(0.018)	0.054(0.018)	0.057(0.019)
		0.042(0.015)	0.042(0.015)	0.044(0.015)
		0.028(0.009)	0.028(0.009)	0.030(0.008)
	100	0.156(0.074)	0.154(0.068)	0.174(0.069)
		0.081(0.030)	0.081(0.029)	0.086(0.030)
		0.054(0.018)	0.054(0.018)	0.057(0.019)
		0.042(0.015)	0.042(0.015)	0.044(0.015)
		0.028(0.009)	0.028(0.009)	0.030(0.008)
	500	0.156(0.074)	0.154(0.068)	0.174(0.069)
		0.081(0.030)	0.081(0.029)	0.086(0.030)
		0.054(0.018)	0.018(0.054)	0.057(0.019)
		0.042(0.015)	0.042(0.015)	0.044(0.015)
		0.028(0.009)	0.028(0.009)	0.030(0.008)

5. CONCLUSION

IBFs are the most difficult default Bayes factors to compute because the number of training samples might be enormous with large sample sizes. But

FBFs are easier to compute than IBFs. The weakness of the FBF was typically inadequate for very small sample sizes. So, the FBF approach requires a large region and the range of applicability of the FBF is more limited than the IBF. However, the FBF methodology provides fully authentic Bayes factor in the sense of dealing only with default standard noninformative priors. It is well defined and seems to be reasonably close to actual Bayes factors. The computational results show that the FBFs appear to correspond to ordinary Bayes factors using intrinsic priors, at least asymptotically.

References

Berger, J. O. (1980). *Statistical Decision Theory*, Springer-Verlag.

Berger, J. O. and Bernardo, J. (1992). On the development of the reference priors, in *Bayesian Statistics 4*, (Bernardo, J. M., Berger, J. O., Dawid, A. P. and Smith, A. F. M. eds), Oxford University Press: London, pp. 35-60.

Berger, J. O. and Mortera, J. (1999). Default Bayes factors for Non-Nested Hypothesis Testing. *Journal of the American Statistical Association,* **94**, pp.542-554.

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

Berger, J. O. and Pericchi, L. R. (1998). On criticisms and comparisons of default Bayes factors for model selection and hypothesis testing, *In Proceedings of the Workshop on Model Selection,* Rassegna di Metodi Statitici ed Applicazioni, Pitagora Editrice, Bologna.

Dmochowski, J. (1996). "Intrinsic priors via Kullback-Leibler geometry" in *Bayesian Statistics V*, eds. J. M., Bernardo et al., eds., London, Oxford University Press.

Geisser, S. and Eddy, W.F. (1979). A predictive approach to model selection, *Journal of the American Statistical Association,* **74**, pp. 153-160.

Jeffreys, H. (1961). *Theory of Probability.* London: Oxford University Press.

Kass, R. E. and Raftery, A. (1995) Bayes Factors, *Journal of the American Statistical Association,* **90**, pp. 773-795.

Kim, S. W. (2000). Intrinsic priors for testing exponential means, *Statistics and Probability Letters*, **46**, pp 195-201.

Kim, S. W. and Ibrahim, J. G. (2000). Default Bayes factors for generalized linear models, to appear *Journal of Statistical Planning and Inference.*

Lingham, R. T. and Sivaganesan, S. (1997). Testing hypotheses about the power law process under failure truncation using intrinsic Bayes factors. *Ann. Inst. Statist. Math.*, **49**, pp. 693-710.

Moreno, E., Bertolino, F., and Racugno, W. (1998). An intrinsic limiting procedure for model selection and hypothesis testing, *Journal of the American Statistical Association,* **93**, pp. 1451-1460.

O'Hagan, A. (1995). Fractional Bayes factors for Model Comparison, *Journal of Royal Statistical Society, ser. B*, **57**, pp. 99-138.

Rohatgi, V. (1976). *An Introduction to Probability Theory and Mathematical Statistics*, New York: Wiley.

San Martini, A. and Spezzaferri, S. F. (1984). A predictive model selection criterion, *Journal of Royal Statistical Society, ser. B*, **46**, pp. 296-303.

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, ser. B*, **44**, pp. 377-387.

Varshavsky, J. A. (1996). Intrinsic Bayes factors for model selection with autoregressive data, in Bayesian Analysis V, J. M. Bernado, et al., eds., pp. 757-763, Oxford University Press.