Ann. Inst. Statist. Math.
Vol. 51, No. 3, 585–602 (1999)

A METHOD FOR TESTING NESTED POINT NULL HYPOTHESES USING MULTIPLE BAYES FACTOR

HEA-JUNG KIM

Department of Statistics, Dongguk University, Seoul 100-715, Korea

(Received January 17, 1997; revised February 27, 1998)

Abstract. As a flexible Bayesian test criterion for nested point null hypotheses, asymmetric and multiple Bayes factors are introduced in the form of a modified Savage-Dickey density ratio. This leads to a simple method for obtaining pairwise comparisons of hypotheses in a statistical experiment with a partition on the parameter space. The method is derived from the fact that in general, the asymmetric Bayes factor can be written as the product of the Savage-Dickey ratio and a correction factor where both terms are easily estimated by means of posterior simulation. Analyses of a censored data problem and a serial correlation problem are illustrated for the method. For these cases, the method is straightforward for specifying distributionally and to implement computationally, with output readily adapted for required tests.

Key words and phrases: Asymmetric and multiple Bayes factors, Savage-Dickey density ratio, Gibbs sampler, point null hypothesis, censored data, serial correlation.

1. Introduction

Consider an experiment with a parameter vector (or suitably reparameterized vector) $\theta = (\omega, \psi)' \in \Theta(= \Omega \times \Psi; \Omega \subseteq R \text{ and } \Psi \subseteq R)$, where we are interested in testing a nested point null hypothesis $H_0 : \omega = \omega_0$ (ω_0 given) against a general alternative $H_A : \omega \neq \omega_0$, where $\omega \in \Omega$ is an unknown parameter of the statistical model. For Bayesian inference, Bayes factor is usually used as measure of evidence in favor of H_0 versus H_A, which is complement of H_0 (cf. Kass and Raftery (1995); Verdinelli and Wasserman (1996)). Let $\pi \in (0, 1)$ be the prior probability of H_0, and let p_0 and p be the densities conditional on H_0 and H_A (here p_0 is also a density of θ, but the conditioning on $H_0 : \omega = \omega_0$ eliminates ω from the argument). Thus the overall prior distribution function of $\theta = (\omega, \psi)'$ is the mixture

$$F(\theta) = \pi I_{[\omega_0,\infty)}(\omega) \int_{-\infty}^{\psi} p_0(t_2)dt_2 + (1 - \pi) \int_{-\infty}^{\omega} \int_{-\infty}^{\psi} p(t_1, t_2)dt_2dt_1.$$

The Bayes factor, B, for testing H_0 is

$$B = \frac{\int_\Psi L(\omega_0, \psi) p_0(\psi) d\psi}{\int_\Omega \int_\Psi L(\omega, \psi) p(\omega, \psi) d\omega d\psi}, \tag{1.1}$$

where $L(\omega, \psi)$ is the likelihood function (cf. Jeffreys (1961)). In constructing the Bayes factor for the point null hypothesis, it has been usually thought sensible to take the prior distribution of the parameter ω, conditional on H_A, to be symmetrical with respect to ω_0 (so as to have an equal weight function).

In an effective Bayesian analysis concerning a real parameter ω, however, it often occurs that prior attitudes with respect to $\omega < \omega_0$ and $\omega > \omega_0$ are different. For example, it is unusual that the effect of a policy will be given a priori an equal chance of being positive or negative. In the particular case where there is no nuisance parameter (ψ) involved in the null hypothesis, Zellner (1987) and Bertolino *et al.* (1995) suggested a flexible Bayes factor which deals simultaneously, and without any constraints of symmetry, with any number of hypotheses. Given a partition ($\Omega_0 = \{\omega_0\}$, $\Omega_1 = \{\omega < \omega_0\}$, $\Omega_3 = \{\omega > \omega_0\}$) of parameter space Ω, they introduced a multiple Bayes factor as a vector with two components, the asymmetric Bayes factors of Ω_0 versus Ω_1, Ω_2 respectively (they called asymmetric Bayes factor as partial Bayes factor in their works). Therefore, this approach makes a separate analysis of suitable subsets of Ω possible so that, from a classical perspective, the approach may be seen as a multiple decision problem (see, for example, Ferguson (1967)). It is seen that the approach is appropriate in practical problems and useful in robust perspectives (cf. Bertolino *et al.* (1995)). However, the approach is not directly applicable for the case of a nested null hypothesis where ω is a subcomponent of the whole parameter θ. This indicates that the application of the concepts of multiple Bayes factor to the comparison of hypotheses in the presence of a nuisance parameter can be interesting; a modification to the multiple Bayes factor is needed for encompassing a nested point null hypothesis.

The purpose in this paper is to suggest the multiple Bayes factor that enables us to deal with the nested point null hypothesis when the different prior attitudes with respect to the partition $(\Omega_0, \Omega_1, \ldots, \Omega_K)$, $K \geq 1$, of parameter space Ω is appropriate. Analytic methods(exact or approximate) for carrying out required multi-dimensional integrations in calculating the multiple Bayes factor will sometimes be well-nigh impossible (especially for analysis of constrained parameter and truncated data problem as studied in Gelfand *et al.* (1992)). A simple method for approximating the multiple Bayes factor is also proposed. It is shown that the Bayesian calculations can be implemented routinely for the analytically impossible cases by means of the method. The method is obtained by extending the method of Verdinelli and Wasserman (1995) to the case where there are several hypotheses.

In Section 2, we introduce the multiple Bayes factor for testing a nested point null hypothesis. Section 3 interprets it in terms of a modified Savage-Dickey density ratio. Then we introduce the simple method for estimating the density ratio and discuss its implementation. In Section 4, we apply the simple method to a censored data problem. As another illustration, in Section 5, the method is applied to the analysis of a serial correlation, and a summary discussion is provided in Section 6.

2. Asymmetric Bayes factors for nested point null hypotheses

Let X having a density $f(x \mid \theta)$ be observed with unknown $\theta = (\omega, \psi) \in \Theta(= \Omega \times \Psi)$, and it is desired to test the nested point null hypothesis $H_0 : \omega = \omega_0$ versus $H_A : \omega \neq \omega_0$. Introducing the partition $(\Omega_0, \Omega_1, \Omega_2)$, where $\Omega_0 = \{\omega_0\}$, $\Omega_1 = \{\omega : \omega < \omega_0\}$, and $\Omega_2 = \{\omega : \omega > \omega_0\}$, makes it easier to perform a Bayesian analysis, especially the subsets Ω_1 and Ω_2 play a priori different roles. Let us assign the probabilities $\pi_i(\sum \pi_i = 1)$ to the events $\theta \in \Omega_i \times \Psi$ $(i = 0, 1, 2)$, and let p_0, p_1, and p_2 be respective prior densities of θ conditional on $\theta \in \Omega_0 \times \Psi$, $\theta \in \Omega_1 \times \Psi$, and $\theta \in \Omega_2 \times \Psi$. Then the overall prior distribution function of θ is

$$F(\theta) = \pi_0 I_{[\omega_0,\infty)}(\omega) \int_{-\infty}^{\psi} p_0(t_2)dt_2 + \pi_1 \int_{-\infty}^{\omega} \int_{-\infty}^{\psi} g_1(t_1 \mid t_2)p(t_2)dt_2 dt_1 + \pi_2 \int_{-\infty}^{\omega} \int_{-\infty}^{\psi} g_2(t_1 \mid t_2)p(t_2)dt_2 dt_1, \tag{2.1}$$

where $g_j(\omega \mid \psi)p(\psi) = p_j(\theta)$, is the conditional prior density for $\theta \in \Omega_j \times \Psi$, $j = 1, 2$, and $p_0(\psi)$ is the prior for ψ under H_0. Here $I(\cdot)$ denotes the indicator function. Different prior beliefs about ω in Ω_1 and Ω_2 can be represented by a suitable choice of the conditional density g_1 and g_2 (or, for robust analyses, of classes G_1 and G_2 of possible densities). Extending the asymmetric Bayes factors, introduced by Bertolino *et al.* (1995), to the case of the nested null hypothesis, we set up the following definition.

DEFINITION 1. The asymmetric Bayes factors of Ω_0 versus Ω_1 and Ω_2 are defined as

$$B_1 = \frac{\int_\Psi L(\omega_0, \psi)p_0(\psi)d\psi}{m(g_1)} \quad \text{and} \quad B_2 = \frac{\int_\Psi L(\omega_0, \psi)p_0(\psi)d\psi}{m(g_2)}, \tag{2.2}$$

and the multiple Bayes factor is a vector with components B_1 and B_2; (B_1, B_2), where $L(\omega, \psi)$ denotes the likelihood function, and $m(g_j) = \int_{\Omega_j} \int_\Psi L(\omega, \psi) g_j(\omega \mid \psi)p(\psi)d\psi d\omega$, $j = 1, 2$.

B_1 and B_2 can be seen as asymmetric Bayes factors of Ω_0 versus Ω_j, $j = 1, 2$, conditional on $\theta \in (\Omega_0 \cup \Omega_1) \times \Psi$ and $\theta \in (\Omega_0 \cup \Omega_2) \times \Psi$, respectively. They do not depend on π_0, π_1 and π_2 but rather depend on the given partition; $\Omega = (\Omega_0, \Omega_1, \Omega_2)$.

Extension to point null hypothesis $(H_0 : \omega = \omega_0)$ with the partition $(\Omega_0, \Omega_1, \ldots, \Omega_K)$ of the parameter space Ω can be easily obtained.

LEMMA 1. *Assigning probabilities* $\pi_0, \pi_1, \ldots, \pi_K, \sum \pi_i = 1$, *to subsets* $\{\Omega_0 \times \Psi\}, \{\Omega_1 \times \Psi\}, \ldots, \{\Omega_K \times \Psi\}$ *and joint prior densities* $p_i(\theta)$ *for* $\theta = (\omega, \psi) \in \Omega_i \times \Psi$, $i = 1, \ldots, K$, *we have the asymmetric Bayes factor of* Ω_0 *versus* Ω_i *given by*

$$B_i = \frac{\int_\Psi L(\omega_0, \psi)p_0(\psi)d\psi}{m(g_i)}, \quad \text{for} \quad \omega \in \Omega_0 \cup \Omega_i, \quad i = 1, \ldots, K, \tag{2.3}$$

and the overall Bayes factor:

$$OB = \sum_{i=1}^{K} \pi_i / \sum_{i=1}^{K} \pi_i B_i^{-1}, \tag{2.4}$$

where $p_0(\psi)$ *denotes the prior for* ψ *under the space* $\Omega_0 \times \Psi$, $m(g_i) = \int_{\Omega_i} \int_{\Psi} L(\omega, \psi) g_i(\omega \mid \psi) p(\psi) d\psi d\omega$, *and* $p_i(\theta) = g_i(\omega \mid \psi) p(\psi)$.

PROOF. Given the partition of the parameter space, the overall distribution of $\theta = (\omega, \psi)$ is

$$\begin{aligned} F(\theta) = \pi_0 I_{[\omega_0, \infty)}(\omega) \int_{-\infty}^{\psi} p_0(t_2) dt_2 + \pi_1 \int_{-\infty}^{\omega} \int_{-\infty}^{\psi} g_1(t_1 \mid t_2) p(t_2) dt_2 dt_1 \\ + \cdots + \pi_K \int_{-\infty}^{\omega} \int_{-\infty}^{\psi} g_K(t_1 \mid t_2) p(t_2) dt_2 dt_1. \end{aligned} \tag{2.5}$$

Thus, from (1.1), we get the Bayes factor of Ω_0 versus Ω_i conditional on $\theta \in (\Omega_0 \cup \Omega_i) \times \Psi$ as B_i, $i = 1, \ldots, K$. Now the overall Bayes factor is

$$OB = \frac{\int_{\Psi} L(\omega_0, \psi) p_0(\psi) d\psi}{\int_{\Omega} \int_{\Psi} L(\omega, \psi) \tilde{g}(\omega \mid \psi) p(\psi) d\psi d\omega},$$

where

$$\tilde{g}(\omega \mid \psi) = \frac{\pi_1}{1 - \pi_0} g_1(\omega \mid \psi) + \cdots + \frac{\pi_K}{1 - \pi_0} g_K(\omega \mid \psi) \tag{2.6}$$

is the prior density of θ conditional on H_A. Expressing OB in terms of B_i, $i = 1, \ldots, K$, we have the equation (2.4).

Note that if $g_1(\omega \mid \psi) = \cdots = g_K(\omega \mid \psi)$ for $\Omega_1 \cup \cdots \cup \Omega_K$, then the overall Bayes factor becomes the same as the usual Bayes factor in (1.1). Extensions to a nested point null hypothesis with more than one nuisance parameter can be easily obtained by setting $\psi \in \Psi \subseteq R^p$, $p \geq 2$. For the general case, the final probabilities of Ω_0 and Ω_i turn out to be

$$\begin{aligned} P(\Omega_0 \mid \text{data}) &= \left\{ 1 + \pi_0^{-1} \sum_{i=1}^{K} \pi_i B_i^{-1} \right\}^{-1}, \\ P(\Omega_i \mid \text{data}) &= \left\{ 1 + \pi_0 / \pi_i B_i + \pi_i^{-1} B_i \sum_{j=1, j \neq i}^{K} \pi_j B_j^{-1} \right\}^{-1}, \\ & \qquad i = 1, \ldots, K. \end{aligned} \tag{2.7}$$

It is easily seen that if we are interested in the asymmetric Bayes factor B' for Ω_0 versus, say, $\Omega_{i_1} \cup \Omega_{i_2} \cup \cdots \cup \Omega_{i_h}$, where $\{i_1, i_2, \ldots, i_h\}$ is any subset of $\{1, 2, \ldots, K\}$, then we have directly B' as the weighted harmonic mean of

$B_{i_1}, B_{i_2}, \ldots, B_{i_h}$. Moreover, the asymmetric Bayes factors of Ω_i versus Ω_j are given simply by B_j/B_i.

The asymmetric Bayes factors based on asymmetry of the conditional prior densities also reveals the conflict between p-value (observed significance level) and $P(\Omega_0 \mid \text{data})$. This can be seen by using the example in Berger and Sellke (1987): Suppose $X = (x_1, \ldots, x_n)$ is a vector of independently $N(\omega, \psi)$ random variables, and that ψ is known, and suppose it is desired to test the point null hypothesis $H_0 : \omega = \omega_0$ versus $H_A : \omega \neq \omega_0$ using the asymmetric Bayes factors B_1 and B_2. Under the partition $(\Omega_0, \Omega_1, \Omega_2)$ of the parameter space, where $\Omega_0 = \{\omega_0\}$, $\Omega_1 = \{\omega : \omega < \omega_0\}$ and $\Omega_2 = \{\omega : \omega > \omega_0\}$, we assume the conditional prior density $g_i(\omega)$ is $N_{\Omega_i}(\omega_0, \tau_i\sigma^2)$, $j = 1, 2$. Since a sufficient statistic for ω is $\bar{X} \sim N(\omega, \sigma^2/n)$, we have that $m(g_i)$ is an $N(\omega_0, (\tau_i + n^{-1})\sigma^2)$, $i = 1, 2$. Thus

$$B_i = (1 + n\tau_i)^{1/2} \exp\left\{-\frac{t^2}{2(1 + (n\tau_i)^{-1})}\right\}, \quad i = 1, 2, \tag{2.8}$$

and, from (2.7),

$$P(H_0 \mid x) = \left[1 + \sum_{i=1}^{2} \frac{\pi_i}{\pi_0}(1 + n\tau_i)^{-1/2} \exp\left\{-\frac{t^2}{2(1 + (n\tau_i)^{-1})}\right\}\right]^{-1}, \tag{2.9}$$

where $t^2 = n(\bar{x} - \omega_0)^2/\sigma^2$.

Therefore, the Lindley paradox is apparent from this expression: For arbitrary π_0, π_1 and π_2, if t is fixed, corresponding to a fixed p-value, but $n \to \infty$, then $P(H_0 \mid x) \to 1$ no matter how small the p-value.

3. The simple method

As seen in (2.3), it is often difficult to compute B_i due to required multi-dimensional integrations (especially for $\Psi \subseteq R^p$). One possibility is to approximate B_i analytically by approximating the integrals in B_i by Laplace's method (cf. Tierney *et al.* (1989); Hsiao (1997)). Recent advances in statistical computing make it feasible to use posterior simulation output to estimate the Bayes factor, or integrated likelihood ratio, for a large class of problems (cf. Gelfand and Dey (1994); Newton and Raftery (1994); Chib (1995); Lewis and Raftery (1997)). In this section we consider yet another simple method (an extension of the method by Verdinelli and Wasserman (1995)) for estimating B_i which exploits Gibbs sampling algorithm to avoid or improve on the analytic approximation.

3.1 *Modified Savage-Dickey density ratio*

Dickey (1971) derived an alternative expression for the Bayes factor defined in (1.1). The expression can be directly applied to the asymmetric Bayes factors: If the prior in Lemma 1 has a relation that

$$p_i(\psi \mid \omega_0) = p_0(\psi) \quad \text{for} \quad (\omega, \psi) \in (\Omega_0 \cup \Omega_i) \times \Psi, \quad i = 1, \ldots K, \tag{3.1}$$

then $B_i = p_i(\omega_0 \mid x)/p_i(\omega_0)$, where $p_i(\omega \mid x) = \int p_i(\omega, \psi \mid x)d\psi$, the marginal posterior of $\omega \in (\Omega_0 \cup \Omega_i)$. Dickey attributed this formula to Savage. Thus the alternative expression for B_i is called Savage-Dickey density ratio. We see that when the special form (3.1) holds, computing B_i reduces to the problem of estimating the marginal posterior density $p_i(\omega \mid x)$ at the point ω_0. But there are cases where (3.1) is not appropriate. The next section presents an example of this.

When (3.1) is not appropriate for the prior specification, the following theorem gives an alternative expression of the asymmetric Bayes factor suited for a simple approximation method.

THEOREM 1. *Assume that $0 < p_i(\omega_0 \mid x) < \infty$ and that $0 < p_i(\omega_0, \psi) < \infty$ for almost all ψ, $i = 1, \ldots, K$. Then the asymmetric Bayes factor of Ω_0 versus Ω_i defined in (2.3) is*

$$B_i = p_i(\omega_0 \mid x)E\left[\frac{p_0(\psi)}{p_i(\omega_0, \psi)}\right] = \frac{p_i(\omega_0 \mid x)}{p_i(\omega_0)}E\left[\frac{p_0(\psi)}{p_i(\psi \mid \omega_0)}\right], \tag{3.2}$$

assuming that the expectation is finite, where the expectation is with respect to $p_i(\psi \mid \omega_0, x)$. Here $p_i(\omega \mid x) = m(g_i)^{-1} \int_\Psi L(\omega, \psi)p_i(\omega, \psi)d\psi I_{\Omega_0 \cup \Omega_1}(\omega)$ is the truncated marginal posterior density, and $p_i(\omega, \psi)$ is the conditional prior density of $(\omega, \psi) \in \Omega_i \times \Psi$.

PROOF. For $\omega \in \Omega_0 \cup \Omega_i$, we have that

$$\begin{aligned}
B_i &= \frac{\int_\Psi L(\omega_0, \psi)p_0(\psi)d\psi}{m(g_i)} \\
&= p_i(\omega_0 \mid x)\frac{\int_\Psi L(\omega_0, \psi)p_0(\psi)d\psi}{p_i(\omega_0 \mid x)m(g_i)} \\
&= p_i(\omega_0 \mid x)\int_\Psi \frac{L(\omega_0, \psi)p_0(\psi)p_i(\psi \mid \omega_0, x)}{p_i(\omega_0, \psi \mid x)m(g_i)}d\psi \\
&= p_i(\omega_0 \mid x)\int_\Psi \frac{p_0(\psi)}{p_i(\omega_0, \psi)}p_i(\psi \mid \omega_0, x)d\psi,
\end{aligned}$$

because $p_i(\omega_0, \psi \mid x) = L(\omega_0, \psi)p_i(\omega_0, \psi)/m(g_i)$.

Note that, when (3.1) fails, B_i in (3.2) is equal to the Savage-Dickey density ratio times a correction factor, and hence (3.2) can be seen as a modified expression of the Savage-Dickey density ratio for the asymmetric Bayes factors.

3.2 *Simple approximation method*

Theorem above shows that computing B_i is reduced to the problem of estimating the truncated marginal posterior density $p_i(\omega \mid x)$ at the point ω_0 and the factor $C_i = E[\frac{p_0(\psi)}{p_i(\omega_0, \psi)}]$. An adaptive Monte Carlo integration technique known as the Gibbs sampler is proposed as a mechanism for implementing a conceptually and computationally simple solution in such a problem. We refer to Gelfand and Smith (1990) for a discussion of this technique and its properties.

Assume that samples $\{(\omega_j(i), \psi_j(i)), j = 1, \ldots, m\}$ and $\{\tilde{\psi}_j(i), j = 1, \ldots, m\}$, $i = 1, \ldots K$, are independent Markov chain samples from the posterior using the implementations of a Markov chain Monte Carlo. If the conditional posterior $p_i(\omega \mid \psi, x)$ defined for $\omega \in \Omega_0 \cup \Omega_i$ is in a closed form, then, following Gelfand and Smith (1990), we can estimate $p_i(\omega_0 \mid x)$ by the finite mixture density

$$\hat{p}_i(\omega_0 \mid x) = \frac{1}{m} \sum_{j=1}^{m} p_i(\omega_0 \mid \psi_j, x). \tag{3.3}$$

When it is not in a closed form, one could use a method by Chen (1992) or standard density estimation techniques such as a kernel density estimate and a smoothed histogram.

Using the sample $(\tilde{\psi}_1, \ldots, \tilde{\psi}_m)$ from $p_i(\psi \mid \omega_0, x)$, we estimate C_i by

$$\hat{C}_i = \frac{1}{m} \sum_{j=1}^{m} \frac{p_0(\tilde{\psi}_j)}{p_i(\omega_0, \tilde{\psi}_j)}. \tag{3.4}$$

These estimates lead to final estimate of B_i as $\hat{B}_i = \hat{p}_i(\omega_0 \mid x)\hat{C}_i$. $\hat{p}_i(\omega_0 \mid x)$ is consistent and a central limit theorem is available for the estimate (cf. Tierney (1994)). The consistency of $\hat{C}_i$ follows from the ergodicity of the Markov chain. Moreover, if the chain is uniformly ergodic and the expectation of $[\frac{p_0(\psi)}{p_i(\omega_0,\psi)}]^2$ under $p_i(\psi \mid \omega_0, x)$ is finite, a central limit theorem applies to $\hat{C}_i$ (cf. Tierney (1994)). Since $\hat{p}_i(\omega_0 \mid x)$ and $\hat{C}_i$ are based on two independent samples, Taylor series expansion yields a first-order approximate variance for $\hat{B}_i$ as

$$V(\hat{B}_i) = s_1^2 \hat{C}_i^2 + s_2^2 \hat{p}_i^2(\omega_0 \mid x), \tag{3.5}$$

where s_1 and s_2 are respective standard errors of $\hat{p}_i(\omega_0 \mid x)$ and $\hat{C}_i$. We refer Ripley ((1987), p. 155), Gelman and Rubin (1992), and Geyer (1992) for the methods of estimating the standard errors.

It is straightforward to see that, setting the joint priors $p_1(\omega, \psi) = \cdots = p_K(\omega, \psi) = p(\omega, \psi)$ for the parameter space $(\Omega_1 \cup \cdots \cup \Omega_K) \times \Psi$, we can readily apply the equations (3.2), (3.3), (3.4), and (3.5) for estimating the usual Bayes factor in (1.1). This yields the estimated Bayes factor as

$$\begin{aligned} \hat{B} &= \hat{p}(\omega_0 \mid x)\hat{C} \\ &= \left\{ \frac{1}{m} \sum_{j=1}^{m} p(\omega_0 \mid \psi_j, x) \right\} \left\{ \frac{1}{m} \sum_{j=1}^{m} \frac{p_0(\tilde{\psi}_j)}{p(\omega_0, \tilde{\psi}_j)} \right\}, \end{aligned} \tag{3.6}$$

where $p(\omega \mid x)$ is the marginal posterior density of $\omega \in \Omega$ obtained from using the common prior $p(\omega, \psi)$.

4. Application with censored data

4.1 *The simple method*

Consider the testing of normal mean with censored data. Assume we have a series of observations $x_h \sim N(\mu, \sigma^2)$, $h = 1, \ldots, n$, and we observe only that $x_h \in A_h$ for some set A_h. Our problem of interest is to test $H_0 : \mu = \mu_0$ with different prior attitudes with respect to a sensible choice of the partition $\Omega_0 = \{\mu = \mu_0\}$, $\Omega_1 = \{\mu < \mu_0\}$, and $\Omega_2 = \{\mu > \mu_0\}$. Let overall prior distribution function of (μ, σ) be

$$(4.1) \quad F(\mu,\sigma) = \pi_0 I_{[\mu_0,\infty)}(\mu) \int_0^\sigma p_0(t_2)dt_2 + \pi_1 \int_{-\infty}^{\mu} \int_0^\sigma g_1(t_1 \mid t_2)p(t_2)dt_2dt_1 + \pi_2 \int_{-\infty}^{\mu} \int_{-\infty}^{\sigma} g_2(t_1 \mid t_2)p(t_2)dt_2dt_1,$$

where $g_i(\mu \mid \sigma)$ is the pdf of $N_{\Omega_i}(\mu_0, \tau_i\sigma^2)$, denoting a normal distribution $N(\mu_0, \tau_i\sigma^2)$, truncated to the set Ω_i and $p(\sigma) = p_0(\sigma) \propto \sigma^{-1}$. Note that, in this case, $\tilde{g}$ defined in (2.6) with $K = 2$ is a unimodal density having a discontinuity point in the mode $\mu = \mu_0$, and see Lee ((1988), p. 140) for the dependent prior setting.

Standard Bayes calculations yield the following complete truncated conditional distributions involved in (3.2): Given the partition and the prior $p_i(\mu, \sigma)$, $i = 1, 2$,

$$\mu \mid x_1, \ldots, x_n, \sigma \sim N_{\Omega_i}\left(\frac{n\bar{x} + \mu_0\tau_i^{-1}}{n + \tau_i^{-1}}, \frac{\sigma^2}{n + \tau_i^{-1}}\right), \qquad \text{for} \quad \mu \in \Omega_i,$$

$$\sigma \mid x_1, \ldots, x_n, \mu \in \Omega_i \sim \frac{S_i^{1/2}}{\chi_{(n+1)}},$$

where $S_i = \sum_{h=1}^{n}(x_h - \mu)^2 + \tau_i^{-1}(\mu - \mu_0)^2$, $\bar{x} = \sum_{h=1}^{n} x_h/n$. Treating the true unobserved values of the censored x_h's as parameters, we also obtain the conditional distributions of x_h's:

$$x_h \mid \mu, \sigma \sim N_{A_h}(\mu, \sigma^2), \qquad h = 1, \ldots, n.$$

We may implement the Gibbs sampler using the complete conditional distributions given above. Run the Gibbs sampler for each of the Ω_i separately with m times to produce a sample whose empirical distribution approximates marginal posterior of μ (see, for example, Geyer (1992), for choosing m). Then we estimate $p_i(\mu_0 \mid x_1 \in A_1, \ldots, x_n \in A_n)$, $i = 1, 2$, by

$$(4.2) \quad \hat{p}_i(\mu_0 \mid x_1 \in A_1, \ldots, x_n \in A_n) = \frac{1}{m}\sum_{k=1}^{m} \frac{\phi\left(\mu_0; \frac{n\bar{x}_k + \mu_0\tau_i^{-1}}{n + \tau_i^{-1}}, \frac{\sigma_k^2}{n + \tau_i^{-1}}\right)}{\int_{\Omega_i} \phi\left(\mu; \frac{n\bar{x}_k + \mu_0\tau_i^{-1}}{n + \tau_i^{-1}}, \frac{\sigma_k^2}{n + \tau_i^{-1}}\right) d\mu},$$

where $\phi(\gamma;\alpha,\beta)$ denotes the pdf of $N(\alpha,\beta)$ with variable γ, and $\bar{x}_k$ and σ_k are the mean of $x_1,\ldots,x_n$ and the value of σ at the k-th replication, respectively. We proceed the same Gibbs sampling with μ fixed at μ_0 to draw a sample $(\tilde{\sigma}_1,\ldots,\tilde{\sigma}_m)$ from the posterior $p_i(\sigma \mid \mu_0, x_1 \in A_1,\ldots,x_n \in A_n)$. Since $p_0(\sigma)/p_i(\mu_0,\sigma) = 1/g_i(\mu_0 \mid \sigma) = 2^{-1/2}(\tau_i\pi)^{1/2}\sigma$, $i=1,2$, the factor defined in (3.2) is given by

$$\begin{aligned} C_i &= E\left[\frac{p_0(\sigma)}{p_i(\mu_0,\sigma)} \mid \mu_0, X_1 \in A_1,\ldots,X_n \in A_n\right] \\ &= 2^{-1/2}(\pi\tau_i)^{1/2}E[\sigma \mid \mu_0, X_1 \in A_1,\ldots,X_n \in A_n], \end{aligned}$$

which can be estimated by

$$\hat{C}_i = \frac{2^{-1/2}(\pi\tau_i)^{1/2}}{m}\sum_{k=1}^{m}\tilde{\sigma}_k. \tag{4.3}$$

Finally, using (2.8), we calculate posterior probability of each hypothesis based upon the estimated asymmetric Bayes factors given by

$$\hat{B}_i = \hat{C}_i\hat{p}_i(\mu_0 \mid x_1 \in A_1,\ldots,x_n \in A_n), \qquad i=1,2. \tag{4.4}$$

In the course of sampling from the truncated normal distributions, we may use a convenient "one-for-one" sampling method by Devroye (1986): To draw an observation from $N(\alpha,\beta)$ restricted to an interval (a,b), we generate U, a random uniform $(0,1)$ variate and calculate $X = \alpha + \beta^{1/2}\Phi^{-1}(\ell)$, where

$$\ell = \Phi\left(\frac{a-\alpha}{\beta^{1/2}}\right) + U\left\{\Phi\left(\frac{b-\alpha}{\beta^{1/2}}\right) - \Phi\left(\frac{a-\alpha}{\beta^{1/2}}\right)\right\},$$

with Φ denoting the standard normal cdf. It is straightforward to show that X has the desired distribution.

For estimating the asymmetric Bayes factors, several methods are also available. In particular, Newton and Raftery (1994) and Gelfand and Dey (1994) took advantage of posterior simulation to compute the Bayes factor. However, their methods need evaluation of the likelihood $L(\mu,\sigma)$ that might be a computational burden for this example: In the censored data problem where x_h is known only to lie in a set A_h, the likelihood evaluation involves integration over the set A_h, $h=1,\ldots,n$.

4.2 *Simulation results*

The aim of this simulation study is to see the performance of the simple method and when and how much the use of the multiple Bayes factor yields substantially different conclusions with respect to the use of the usual Bayes factor.

For the study, we considered two cases of censored data of size n from $N(\mu,\sigma^2)$:

Case 1. $\{\langle -.2(x_1), x_2, x_3, x_4,\ldots,x_{n-2},x_{n-1}\rangle.3(x_n)\}$, where the first observation is left censored, and the last observation is right censored.

Case 2. $\{\langle -.2(x_1)\langle -.5(x_2), x_3, x_4, \ldots, x_{n-2}\rangle .1(x_{n-1})\rangle .3(x_n)\}$, where the first two observations are left censored, and the last two observations are right censored.

Under the partition $\Omega_0 = \{\mu = 0\}$, $\Omega_1 = \{\mu < 0\}$, and $\Omega_2 = \{\mu > 0\}$, we estimated $B_i, i = 1, 2$, and then calculated the posterior probabilities of $H_0 : \mu = 0$, $H_1 : \mu < 0$ and $H_2 : \mu > 0$. For illustration, we specified conditional priors $g_i(\mu \mid \sigma)$ as $\phi(\mu; 0, \tau_i\sigma^2)I_{\Omega_i}(\mu)$. Thus $\{n, \mu, \sigma, \tau_1, \tau_2\}$ is the parameters set to be varied in this simulation.

Given a generated data, Gibbs sampling in the previous subsection was conducted. Running the Gibbs sampler for each of the $\hat{p}_i(\mu_0 = 0 \mid x_1 \in A_1, \ldots, x_n \in A_n)$ and the $\hat{C}_i$ separately with $m = 1000$, we monitored the collection $\{\mu_k^{(t)}, k = 1, \ldots, m\}$ and $\{\tilde{\sigma}_k^{(t)}, k = 1, \ldots, m\}$ over 5-unit t-increments ($t = 5, 10, \ldots$). Stabilization of each empirical quantiles of $\mu_k^{(t)}$ and $\tilde{\sigma}_k^{(t)}$, and successive density estimates obtained by (4.2) indicated convergence of the algorithm within $t = 30$ iterations. Under the same simulation, to estimate the usual Bayes factor for testing H_0 versus H_A ($H_A = H_1 \cup H_2$), we specified the conditional priors to be $g_1(\mu \mid \sigma) = g_2(\mu \mid \sigma) = \phi(\mu; 0, \tau\sigma^2)I_{\Omega_1\cup\Omega_2}(\omega)$.

For each parameters set with $\sigma = 1$, we carried out 200 runs of the posterior simulation. In the case of $\mu = 0$, Table 1 tabulates means and standard deviations (in parentheses) of the estimated asymmetric Bayes factors and the usual Bayes factor for a couple of sample sizes n and for different types of τ_1, τ_2 and τ. Posterior probabilities of H_0 (denoted by $\Pr(H_0 \mid x)$) calculated from the estimated Bayes factors are also tabulated in the tables. For the calculation of the probabilities from (2.7), we assumed that the prior probabilities of the three hypotheses are $\pi_0 = 1/2$ and $\pi_1 = \pi_2 = 1/4$.

The table notes several points: (i) The posterior probabilities for H_0 indicate that the simple method in Section 3 estimates of the asymmetric Bayes factors for testing the null hypothesis ($H_0 : \mu = 0$) properly. (ii) It is also noted from the tables that, regardless of prior distributions, the test using the asymmetric Bayes factors results in higher posterior probability of H_0 than that using the usual Bayes factor. This favorable performance of the test by the asymmetric Bayes factors is shown to be consistent with different settings of censored data (Case 1 and Case 2), and the parameters set. This suggests the idea that the asymmetric Bayes factors carry information about H_0 that could be remain hidden when considering the usual Bayes factor. (iii) Finally, the above implications become more apparent with increase of the sample size.

The advantage of using the asymmetric Bayes factors (B_1 and B_2) is that, when H_0 is not true, they enable us to make a correct decision between H_1 and H_2 incorporating the different prior attitudes with respect to the hypotheses. This is not available from the use of the usual Bayes factor. Table 2 highlights the merit of the asymmetric Bayes factors by comparing the posterior probabilities of $H_1 : \mu < 0$ and $H_2 : \mu > 0$ with those obtained from the symmetric priors. The probabilities noted in the table are calculated from (2.7) using $\pi_0 = 1/2$ and $\pi_1 = \pi_2 = 1/4$ for the Case 2. For the comparison, the conditional prior distributions ($\mu \mid \sigma \sim N_{\Omega_i}(\mu_0, \tau_i\sigma^2)$, $i = 1, 2$) to be used are as follows:

Table 1. Estimated B_1, B_2, B and corresponding $\Pr(H_0 \mid x)$ for the true mean $\mu = 0$ $(\alpha = .5)$.

		$(\tau_1 = \alpha, \tau_2 = 2\alpha)$			$(\tau_1 = 2\alpha, \tau_2 = \alpha)$			$(\tau = \alpha)$	
		Asymmetric Bayes factor			Asymmetric Bayes factor			Bayes factor	
Case	n	B_1	B_2	$\Pr(H_0 \mid x)$	B_1	B_2	$\Pr(H_0 \mid x)$	B	$\Pr(H_0 \mid x)$
Case 1	10	2.6420	3.7745	.7565	3.3351	2.7494	.7508	1.6730	.6258
		(1.5614)	(2.2865)		(2.1856)	(1.5545)		(.7225)	
	20	5.4130	2.8488	.7887	4.4913	3.8836	.8064	2.4341	.7088
		(3.1079)	(.9435)		(2.8225)	(2.1492)		(.8345)	
Case 2	10	2.1761	3.6104	.7308	3.2406	2.6507	.7446	1.7322	.6339
		(1.1373)	(1.9201)		(1.7865)	(1.3010)		(.0259)	
	20	3.1037	5.2813	.7963	4.5413	3.7954	.8052	2.3217	.6989

Table 2. $\Pr(H_1 \mid x)$ and $\Pr(H_2 \mid x)$ for the true mean μ.

		$\Pr(H_1 \mid x)$				$\Pr(H_2 \mid x)$			
$\lvert\mu\rvert$	n	Prior 1	Prior 2	Prior 3	Prior 4	Prior 1	Prior 2	Prior 3	Prior 4
−1.5	10	0.83072	0.78033	0.82888	0.78254	0.00480	0.00899	0.00700	0.00617
	20	0.99777	0.99675	0.99776	0.99677	0.00003	0.00007	0.00004	0.00005
−1.0	10	0.56186	0.46638	0.55736	0.47092	0.01798	0.03115	0.02584	0.02171
	20	0.92556	0.89528	0.92495	0.89610	0.00151	0.00302	0.00216	0.00210
−0.5	10	0.35804	0.27618	0.35124	0.28216	0.04589	0.07144	0.06403	0.05131
	20	0.35069	0.32739	0.40949	0.27490	0.17343	0.03968	0.03484	0.19367
−0.3	10	0.25965	0.19318	0.25164	0.19982	0.08531	0.12240	0.11353	0.09220
	20	0.31022	0.23795	0.30468	0.24270	0.04482	0.06782	0.06188	0.04921
0.0	10	0.16793	0.11489	0.16199	0.11935	0.10122	0.14046	0.13299	0.10712
	20	0.12829	0.08866	0.12461	0.09139	0.07539	0.10608	0.10190	0.07858
0.3	10	0.09664	0.06155	0.09140	0.06520	0.20220	0.25350	0.24543	0.20924
	20	0.07049	0.04567	0.06582	0.04899	0.23785	0.29462	0.28840	0.24335
0.5	10	0.13373	0.08937	0.12722	0.09415	0.18530	0.23472	0.22497	0.19376
	20	0.04147	0.06336	0.03788	0.06919	0.32378	0.37224	0.38237	0.31442
1.0	10	0.03933	0.02314	0.03653	0.02494	0.40031	0.44921	0.44306	0.40631
	20	0.00425	0.00254	0.00386	0.00280	0.86302	0.87681	0.87565	0.86428
1.5	10	0.01054	0.00672	0.01087	0.00652	0.76989	0.77583	0.76263	0.77302
	20	0.00009	0.00008	0.00013	0.00006	0.99593	0.99395	0.99391	0.99394

Asymmetric priors. Prior 1: $(\tau_1 = |\mu| + \alpha,\ \tau_2 = 2(|\mu| + \alpha))$; Prior 2: $(\tau_1 = 2(|\mu| + \alpha),\ \tau_2 = |\mu| + \alpha)$.

Symmetric priors. Prior 3: $(\tau_1 = \tau_2 = |\mu| + \alpha)$; Prior 4: $(\tau_1 = \tau_2 = 2(|\mu| + \alpha)$, where $\alpha = .5I_{\Omega_0}(\mu))$.

The table notes a systematic pattern that, when true mean is $\mu < 0$, the asymmetric priors with $\tau_1 < \tau_2$ yield the highest $\Pr(H_1 \mid x)$. On the other hand, when $\mu > 0$, those with $\tau_1 > \tau_2$ yield the highest $\Pr(H_2 \mid x)$. These phenomenon is shown to be consistent with the true mean value (μ) and the sample size (n). Thus, this fact can be used as a criterion for the choice of the asymmetric priors. Finally, from Tables 1 and 2, we can deduce that, if we have different prior attitudes with respect to the hypotheses, the asymmetric Bayes factors leads to more correct test than the usual Bayes factor. In this comparison, we see that the posterior probabilities for Case 1 reveals the same implications as those in Table 2, and hence we omit them from the presentation.

5. Test of serial correlation for simple regression model

5.1 *The simple method*

Consider a model with a disturbance term generated by a first-order autoregressive process; that is,

$$(5.1)\qquad \begin{aligned} y_t &= \beta x_t + u_t \\ u_t &= \rho u_{t-1} + \epsilon_t, \qquad t = 1, \ldots, T, \end{aligned}$$

where ρ is restricted to $A = \{\rho : -1 \le \rho \le 1\}$. It is assumed that the ϵ_t are normally and independently distributed with zero means and common variance $1/\delta$. Note that if $\rho = 0$ would reduce to a simple regression model. Therefore, our interest in this section is to construct the simple method for testing of $H_0 : \rho = \rho_0$, $H_1 : 1 \le \rho \le \rho_0$ and $H_2 : \rho_0 < \rho \le 1$ with different prior attitudes with respect to a sensible choice of $\Omega_0 = \{\rho = \rho_0\}$, $\Omega_1 = \{-1 \le \rho < \rho_0\}$ and $\Omega_2 = \{\rho_0 < \rho \le 1\}$. Difficulties of the test arising from the frequentist approaches are well illustrated in Marr and Quesenberry (1991). Let overall prior distribution function of (ρ, β, δ) be

$$(5.2)\qquad \begin{aligned} F(\rho, \beta, \delta) = {} & \pi_0 I_{[\rho_0, 1]}(\rho) \int_0^\delta \int_{-\infty}^\beta p_0(t_2, t_3) dt_3 dt_2 \\ & + \pi_1 \int_{-1}^\rho \int_0^\delta \int_{-\infty}^\beta g_1(t_1 \mid t_2, t_3) p(t_2, t_3) dt_3 dt_2 dt_1 \\ & + \pi_2 \int_{-1}^\rho \int_0^\delta \int_{-\infty}^\beta g_2(t_1 \mid t_2, t_3) p(t_2, t_3) dt_3 dt_2 dt_1, \end{aligned}$$

where $g_i(\rho \mid \beta, \delta) = g_i(\rho)$ is the pdf of $N_{\Omega_i}(\rho_0, r_i)$, denoting a normal distribution $N(\rho_0, r_i)$, truncated to the set Ω_i and $p(\beta, \delta) = p_0(\beta, \delta) \propto \delta^{1/2}$ (cf. Zellner (1971)). Note that, in this case, $\tilde{g}$ defined in (2.6) with $K = 2$ is a unimodal density having a discontinuity point in the mode $\rho = \rho_0$. Assuming the asymmetric prior setting, we have, for given data $D = (y_1, \ldots, y_T, x_1, \ldots, x_T)$, joint posterior density for β, ρ and δ given by

$$(5.3)\qquad \begin{aligned} & p(\beta, \rho, \delta \mid D) \\ & \propto \delta^{(T+1)/2} \exp\left\{ -\frac{\delta}{2} \sum_{t=2}^{T} [y_t - \rho y_{t-1} - \beta(x_t - \rho x_{t-1})]^2 \right. \\ & \qquad \left. - \frac{(\rho - \rho_0)^2}{2r_i} \right\} I_{\Omega_i}(\rho), \end{aligned}$$

where $-\infty < \beta < \infty$ and $\delta > 0$. Thus we have the following Gibbs sampler for δ, β and rho:

$$\begin{aligned} & \delta \mid \beta, \rho, D \sim Gamma(a, b), \qquad i = 1, 2, \\ & \beta \mid \delta, \rho, D \sim N(Q_2(\rho)/Q_1(\rho), (Q_1(\rho)\delta)^{-1}), \\ & \rho \mid \beta, \delta, D \sim N_{\Omega_i}\left(\frac{r_i Q_2(\beta) + \rho_0 \delta^{-1}}{r_i Q_1(\beta) + \delta^{-1}}, \frac{r_i \delta^{-1}}{r_i Q_1(\beta) + \delta^{-1}} \right), \qquad \text{for} \quad \rho \in \Omega_i \end{aligned}$$

where

$$a = (T+3)/2, \qquad b = 2\Big/\left\{\sum_{t=2}^{T}(y_t - \rho y_{t-1} - \beta(x_t - \rho x_{t-1}))\right\}^2,$$

$$Q_1(\rho) = \sum_{t=2}^{T}(y_t - \rho y_{t-1})^2, \qquad Q_2(\rho) = \sum_{t=2}^{T}(y_t - \rho y_{t-1})(x_t - \rho x_{t-1}),$$

$$Q_1(\beta) = \sum_{t=2}^{T}(y_{t-1} - \beta x_{t-1})^2, \qquad Q_2(\beta) = \sum_{t=2}^{T}(y_t - \beta x_{t-1})(y_t - \beta x_t).$$

Given the prior independence assumption that $p(\beta, \delta \mid \rho) = p(\beta, \delta)$, the asymmetric Bayes factor of Ω_0 versus Ω_i in (3.2) reduces to

$$(5.4) \qquad B_i = \frac{p_i(\rho_0 \mid D)}{g_i(\rho_0)} = p_i(\rho_0 \mid D)(2\pi r_i)^{1/2}\int_{\Omega_i}(2\pi r_i)^{-1/2}\exp\{-(\rho-\rho_0)^2/(2r_i)\}d\rho, \quad i = 1, 2.$$

The marginal posterior $p_i(\rho_0 \mid D)$, $i = 1, 2$, can be estimated by running the Gibbs sampler for each of the Ω_i separately with m times to produce

$$(5.5) \qquad \hat{p}_i(\rho_0 \mid D) = \frac{1}{m}\sum_{k=1}^{m}\frac{\phi\left(\rho_0; \dfrac{r_i Q_2(\beta_k) + \rho_0\delta_k^{-1}}{r_i Q_1(\beta_k) + \delta_k^{-1}}, \dfrac{r_i\delta_k^{-1}}{r_i Q_1(\beta_k) + \delta_k^{-1}}\right)}{\int_{\Omega_i}\phi\left(\rho_0; \dfrac{r_i Q_2(\beta_k) + \rho_0\delta_k^{-1}}{r_i Q_1(\beta_k) + \delta_k^{-1}}, \dfrac{r_i\delta_k^{-1}}{r_i Q_1(\beta_k) + \delta_k^{-1}}\right)d\rho},$$

where $\phi(\gamma; \gamma_1, \gamma_2)$ denotes the pdf of $N(\gamma_1, \gamma_2)$ with variable γ, and β_k and δ_k are respective values of β and δ at the k-th replication of the posterior simulation from the Gibbs sampler. Finally, we estimate the asymmetric Bayes factor of H_0 versus H_i by

$$(5.6) \qquad \hat{B}_i = \hat{p}_i(\rho_0 \mid D)\int_{\Omega_i}(2\pi r_i)^{-1/2}\exp\{-(\rho-\rho_0)^2/(2r_i)\}d\rho, \qquad i = 1, 2.$$

Thus the posterior probabilities $\Pr(H_0 \mid D)$, $\Pr(H_1 \mid D)$ and $\Pr(H_2 \mid D)$ calculated from the estimated asymmetric Bayes factors lead to testing the hypotheses.

5.2 *An empirical data example*

Consider the data in Table 3 on a set of 113 measurements of diameters of certain automatic-transmission parts made during the life of one cutting tool having an automatic compensator. The data (reported in the study by Marr and Quesenberry (1991)) include compensator setting (Setting) the diameter (y) and part number (x). The objective is to test the existence of serial correlation of y for each of three compensator settings (87, 107, 112) by means of the simple method

Table 3. The empirical data set.

Setting	y	x	Setting	y	x	Setting	y	x
87	26.791	1	87	26.809	40	87	26.831	78
87	26.793	2	87	26.805	41	107	26.816	90
87	26.798	3	87	26.808	42	107	26.821	91
87	26.800	4	87	26.808	43	107	26.813	92
87	26.791	5	87	26.810	44	107	26.817	93
87	26.806	6	87	26.808	45	107	26.821	94
87	26.794	7	87	26.810	46	107	26.820	95
87	26.802	8	87	26.810	47	107	26.823	96
87	26.796	9	87	26.808	48	107	26.818	97
87	26.803	10	87	26.808	49	107	26.824	98
87	26.805	11	87	26.811	50	107	26.824	99
87	26.796	12	87	26.808	51	107	26.824	100
87	26.798	13	87	26.812	52	107	26.821	101
87	26.792	14	87	26.819	53	107	26.828	103
87	26.797	15	87	26.810	54	107	26.826	104
87	26.801	16	87	26.814	55	107	26.833	105
87	26.796	17	87	26.805	56	107	26.830	106
87	26.800	18	87	26.819	57	107	26.822	107
87	26.811	20	87	26.816	58	107	26.823	108
87	26.810	21	87	26.808	59	107	26.813	109
87	26.805	22	87	26.814	60	107	26.820	110
87	26.813	23	87	26.808	61	112	26.815	111
87	26.805	24	87	26.808	62	112	26.816	112
87	26.811	25	87	26.806	63	112	26.813	113
87	26.815	26	87	26.800	64	112	26.826	114
87	26.821	27	87	26.811	65	112	26.821	115
87	26.805	28	87	26.821	66	112	26.827	116
87	26.803	29	87	26.821	67	112	26.829	117
87	26.806	30	87	26.817	68	112	26.828	118
87	26.802	31	87	26.814	69	112	26.826	119
87	26.804	32	87	26.816	70	112	26.826	120
87	26.821	33	87	26.823	71	112	26.833	121
87	26.808	34	87	26.823	72	112	26.835	122
87	26.803	35	87	26.816	73	112	26.826	123
87	26.801	36	87	26.828	74	112	26.827	124
87	26.805	37	87	26.826	75	112	26.825	125
87	26.808	38	87	26.825	76	112	26.823	126
87	26.804	39	87	26.814	77	—	—	—

developed earlier. The first-order autocorrelation of the disturbance term can be explained by the ρ in (5.1). Test results of the simple method are summarized in Table 3, where for each of three settings, estimated Bayes factors are reported

Table 4. Test results for the empirical data set.

	$(r_1 = \alpha,\, r_2 = 2\alpha)$			$(r_1 = 2\alpha,\, r_2 = \alpha)$			$(r_1 = r_2 = \alpha)$	
Setting	B_1	B_2	$\frac{\Pr(H_0 \mid x)}{\Pr(H_2 \mid x)}$	B_1	B_2	$\frac{\Pr(H_0 \mid x)}{\Pr(H_2 \mid x)}$	B	$\Pr(H_0 \mid x)$
87	20.947	.160	.229/.766	18.929	.168	.235/.758	.304	.0221
	(3.106)	(.085)		(2.860)	(.092)		(.166)	
107	9.215	.218	.260/.723	8.457	.207	.249/.733	.413	.258
	(2.474)	(.166)		(2.347)	(.163)		(.307)	
122	8.006	.170	.204/.779	7.564	.156	.192/.791	.302	.197
	(2.552)	(.169)		(2.398)	(.157)		(.289)	

along with the posterior probabilities. The values in parentheses are the estimated standard errors of the estimates. The simple method is performed with Gibbs sequence $t = 10$ and $m = 1000$ repeated runs and with the hypotheses $H_0 : \rho = 0$, $H_1 : -1 \le \rho < 0$ and $H_2 : 0 < \rho \le 1$.

From Table 4 it can be noted that the asymmetric Bayes factors precisely indicate the existence of positive serial correlation of y in each of three compensator settings, while the usual Bayes factor provides practically no evidence either way so far as comparing H_1 and H_2 in three settings is concerned. In comparison with classical Durbin-Watson test, we see that, except for the setting 87 with test statistic value 1.441, the classical serial correlation test fails to test the setting 107 and the setting 112 because their test statistic values (1.220 and 1.039) are located in "inconclusive bounds" for the test with significance level .05.

6. Concluding remarks

In this paper we have suggested a method that applies the concept of multiple Bayes factor to the comparison of hypotheses in the presence of a nuisance parameter (extendible to the case of multiple nuisance parameters). The comparison uses posterior probability (calculated from the asymmetric Bayes factors) that accounts for different prior attitudes about suitably chosen subsets of the parameter space. For simple estimation of the multiple Bayes factor, the asymmetric Bayes factors are expressed in the form of modified Savage-Dickey density ratios which can take advantage of posterior simulation to compute the Bayes factors. The simplification is implied by suitable restrictions on the prior distribution of the nuisance parameter conditional on the parameter of interest. It is applicable whenever the null hypothesis is a nested hypothesis. A couple of applications are adopted to describe how Bayesian test of nested point null hypothesis can be implemented straightforwardly by means of the suggested method. It is shown that the method can eliminate complicate numerical integrations over high dimensional sets defined by complex restrictions. Rather, it requires only sampling from univariate full conditional distributions, restricted to easily described subsets of one dimensional space.

Using a Monte Carlo simulation and an empirical data analysis, we examined and demonstrated performance of the suggested method. It is shown that much stronger inference can be made by the method when different prior attitudes with respect to the subsets of a parameter space are taken into consideration in testing the point null hypothesis: (i) The suggested test by posterior probabilities obtained from the asymmetric Bayes factors yields more accurate and flexible test for point null hypothesis than the test based upon those from the usual Bayes factor. Moreover, this fact is shown to be robust with respect to the choice of the priors considered in the study. This coincides with the implication of Bertolino *et al.* (1995), i.e. asymmetric Bayes factors are useful in a robust perspective. (ii) The method can be used for a multiple decision problem in a sense that the estimated asymmetric Bayes factors make a separate analysis for the subsets of the parameter space possible.

In defining the asymmetric Bayes factors, we considered asymmetric conditional priors $p_i(\omega \mid \psi)$ over $\omega \in \Omega_i$, remaining $p_i(\psi)$'s (the priors of the nuisance parameter) common to each i. It would be appropriate to extend the asymmetry concerns to the priors of nuisance parameter to coming at more flexible Bayes factors. A study pertaining to this problem is left as a future research of interest.

Acknowledgements

The author would like to thank the editor and the referees whose comments helped him a great deal in the revision. This paper was supported by a grant from Dongguk University, Korea.

References

Berger, J. O. and Sellke, T. (1987). Testing a point null hypothesis: the irreconcilability of p values and evidence, *J. Amer. Statist. Assoc.*, **82**, 112–139.

Bertolino, F., Piccinato, L. and Racugno, W. (1995). Multiple Bayes factors for testing hypotheses, *J. Amer. Statist. Assoc.*, **90**, 213–219.

Chen, M. H. (1992). Importance weighted marginal posterior density estimation, Technical Report 92-22, Purdue University, Department of Statistics, Indiana.

Chib, S. (1995). Marginal likelihood from the Gibbs output, *J. Amer. Statist. Assoc.*, **90**, 1313–1321.

Devroye, L. (1986). *Non-uniform Random Variate Generation*, Springer, New York.

Dickey, J. (1971). The weighted likelihood ratio, linear hypotheses on normal location parameters, *Ann. Statist.*, **42**, 204–223.

Ferguson, T. S. (1967). *Mathematical Statistics, A Decision Theoretic Approach*, Academic Press, New York.

Gelfand, A. and Dey, D. (1994). Bayesian model choice: asymptotics and exact calculations, *J. Roy. Statist. Soc. Ser. B*, **56**, 501–514.

Gelfand, A. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities, *J. Amer. Statist. Assoc.*, **85**, 398–409.

Gelfand, A. E., Smith, A. F. M. and Lee, T. M. (1992). Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling, *J. Amer. Statist. Assoc.*, **87**, 523–532.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation, *Statist. Sci.*, **7**, 457–511.

Geyer, C. J. (1992). Practical Markov chain Monte Carlo (with discussion), *Statist. Sci.*, **7**, 473–511.

Hsiao, C. K. (1997). Approximate Bayes factors when a mode occurs on the boundary, *J. Amer. Statist. Assoc.*, **92**, 656–663.

Jeffreys, H. (1961). *Theory of Probability*, 3rd ed., Clarendon Press, Oxford, U.K.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors, *J. Amer. Statist. Assoc.*, **90**, 773–795.

Lee, P. M. (1988). *Bayesian Statistics: An Introduction*, Wiley, New York.

Lewis, S. M. and Raftery, A. E. (1997). Estimating Bayes factors via posterior simulation with the Laplace-Metropolis estimator, *J. Amer. Statist. Assoc.*, **92**, 648–655.

Marr, R. L. and Quesenberry, C. P. (1991). A NU test for serial correlation of residuals from one or more regression regimes, *Technometrics*, **33**, 441–457.

Newton, M. A. and Raftery, A. E. (1994). Approximate Bayesian inference with the weighted likelihood bootstrap, *J. Roy. Statist. Soc. Ser. B*, **56**, 3–48.

Ripley, B. (1987). *Stochastic Simulation*, Wiley, New York.

Tierney, L. (1994). Markov chains for exploring posterior distributions, *Ann. Statist.*, **22**, 1701–1762.

Tierney, L., Kass, R. E. and Kadane, J. B. (1989). Fully exponential Laplace approximations to expectations and variances of nonpositive functions, *J. Amer. Statist. Assoc.*, **84**, 710–716.

Verdinelli, I. and Wasserman, L. (1995). Computing Bayes factors using a generalization of the Savage-Dickey density ratio, *J. Amer. Statist. Assoc.*, **90**, 614–618.

Verdinelli, I. and Wasserman, L. (1996). Bayes factors, nuisance parameters and imprecise tests, *Bayesian Statistics 5* (eds. J. M. Bernardo, J. O. Berger, A. P. Dawid and A. M. F. Smith), 765–771, Oxford University Press, New York.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.

Zellner, A. (1987). Comment on "Testing precise hypotheses" by J. O. Berger and M. Delampady, *Statist. Sci.*, **2**, 339–341.