Ann. Inst. Statist. Math.
Vol. 54, No. 3, 500–516 (2002)

ON THE ASYMPTOTIC STABILITY OF THE INTRINSIC AND FRACTIONAL BAYES FACTORS FOR TESTING SOME DIFFUSION MODELS

S. SIVAGANESAN[1] AND RAMA T. LINGHAM[2]

[1] *Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221, U.S.A.*
[2] *Division of Statistics, Northern Illinois University, DeKalb, IL 60115, U.S.A.*

(Received May 17, 1999; revised May 14, 2001)

Abstract. Random processes, from which a single sample path data are available on a fine time scale, abound in many areas including finance and genetics. An effective way to model such data is to consider a suitable continuous-time-scale analog, X_t say, for the underlying process. We consider three diffusion models for the process X_t and address model selection under improper priors. Specifically, fractional and intrinsic Bayes factors (FBF and IBF) for model selection are considered. Here, we focus on the asymptotic stability of the IBF's and FBF's for comparing these models. Specifically, we propose to employ certain novel transformations of the data in order to ensure the asymptotic stability of the IBF's. While we use different transformations for pairwise comparisons of the models, we also show that a single common transformation can be used when simultaneously comparing all three models. We then demonstrate that, when FBF's are used to compare these models, we may have to employ different, model-specific training fractions in order to achieve asymptotic stability of the FBF's.

Key words and phrases: Fractional Bayes factor, Girsanov formula, intrinsic Bayes factor, Jeffreys prior, local asymptotic normality, mean-reversion, Wiener process.

1. Introduction

In many fields, including finance and genetics, a useful family of stochastic models for an observable time series, X_t say, stipulates that the process X_t evolves according to

$$dX_t = \{\theta f(t, X_t) + g(t, X_t)\}dt + \sigma dB_t, \tag{1.1}$$

where the functions f and g together model the drift of X_t, σ is the constant diffusion of the process and B_t is the standard Brownian motion. See Polson and Roberts (1994) for the scope of this stochastic equation. Three special cases of (1.1) considered by Polson and Roberts (1994) are:

$$M_0 : dX_t = \sigma dB_t, \tag{1.2}$$

$$M_1 : dX_t = \theta dt + \sigma dB_t, \tag{1.3}$$

$$M_2 : dX_t = \{\theta - \alpha(X_t - \theta t)\}dt + \sigma dB_t, \tag{1.4}$$

where α and θ are real valued parameters with $\alpha > 0$. The first two are Brownian motions respectively without and with drift, and the third is mean reversion with drift. Here, we will assume that α is known, and θ is unknown.

Suppose now that the diffusion process X_t is fully observed over the interval $0 \leq t \leq T$, where $T > 0$ is a pre-fixed time point. Under our continuous sampling scheme, it is well known (see Polson and Roberts (1994)) that σ can be estimated perfectly, with no error. Hence, we shall hereinafter assume that σ is known, leaving θ as the only unknown parameter. Note that our data, denoted hereinafter as $X^T = \{X_t : 0 \leq t \leq T\}$, idealizes discrete data that is often available on a fine time scale.

Polson and Roberts (1994) have derived conventional Bayes factors (CBF) for testing (pairs of) the foregoing models, under normal priors for θ. In this paper, we put Jeffreys-type non-informative prior on the drift parameter θ and derive tests for the afore-mentioned models using Intrinsic Bayes Factors (IBF), see Berger and Pericchi (1996), and Fractional Bayes Factors (FBF), see O'Hagan (1995). These partial Bayes factors (PBF) are two popular ways to circumvent the vexing dependence of the CBF on the arbitrary constants in the (improper) prior. For instance, let c_i denote the arbitrary constant in the improper prior for model M_i, $i = 0, 1, 2$ (see Section 2). Then, for testing a model M_i versus another model M_j, a PBF can generically be written as

$$PBF_{ij} = \frac{m_i(X^T)}{m_j(X^T)} CF_{ji},$$

where $m_i(X^T)$ and $m_j(X^T)$, respectively, are the marginals under models M_i and M_j, based on the full sample data X^T, and CF_{ji} is a correction factor that is chosen (see Section 3) to cancel the indeterminate, c_i/c_j, that is present in the first term of the PBF. The PBF methods have been successfully applied in many situations. See, for instance, Lingham and Sivaganesan (1997) and Conigliani and O'Hagan (1998).

While these partial Bayes factors are not actual Bayes factors with respect to any priors, it would be desirable, as recommended in Berger and Pericchi (1996), to seek partial Bayes factors which correspond (*asymptotically*) *to the use of plausible default* (*proper*) *priors*. Berger and Pericchi have espoused the above as a main guiding principle in performing the selection of models using partial Bayes factors. De Santis and Spezzaferri (1997) have also discussed the asymptotic properties of the IBF and FBF. For a related discussion on Berger and Pericchi's principle, see O'Hagan (1997). In this paper, we seek IBF and FBF for comparing the above models that adhere to the above recommendation, and hence will pay particular attention to showing that, as the period, T, of observing the process X_t extends indefinitely, the 'correction' terms, CF_{ji}, converge to positive constants. Specifically, we show that data-transformations may have to precede IBF computations and that the transformations depend on the subset of the above models that are being tested. Also, we show that the choice of a training fraction for the FBF may need to be model specific, but otherwise does not depend on the alternative models that are entertained in the subset, in order to achieve the desired asymptotic stability.

The paper is organized as follows. In Section 2, we derive the non-informative priors under the models. In Section 3, we give a brief outline of the IBF and FBF methods and derive the marginals under these models. We implement the IBF approach in Section 4, including the related asymptotics for pairwise testing of M_0, M_1, and M_2, and then comment on simultaneously testing these three models. In Section 5, we consider the testing issues that are covered in Section 4 using the FBF methodology. Our concluding remarks are given in Section 6. We relegate the proofs of our results to the Appendix.

2. Non-informative priors

Towards deriving default priors for θ in (1.1), we note that the likelihood function $L(\theta)$ of X^T with respect to the measure induced by $dX_t = \sigma dB_t$ is given, using Girsanov's formula (see Øksendal (1985) and Polson and Roberts (1994)), by

$$L(\theta) = \exp\left[\int_0^T \frac{\theta f(t,X_t) + g(t,X_t)}{\sigma^2} dX_t - \frac{1}{2}\int_0^T \left\{\frac{\theta f(t,X_t) + g(t,X_t)}{\sigma}\right\}^2 dt\right]. \tag{2.1}$$

We shall now implement a proposal due to Polson and Roberts (1993) and derive priors for θ. Their method is a type of Jeffreys prior for continuous-time models and can be implemented easily in situations where the family of measures, $\{P_{X^T} \mid \theta, M\}$, of X^T under a model M, is locally asymptotically normal (LAN). Thus, we begin with the following result.

THEOREM 2.1. (i) *Consider a model M for X_t, which solves* (1.1). *Then, the Fisher information based on the data X^T is*

$$I_T(\theta \mid M) = \frac{1}{\sigma^2}\int_0^T E(f^2(t,X_t))dt. \tag{2.2}$$

(ii) *For any model M_i, $i = 1, 2$, introduced in* (1.3) *and* (1.4), *and any $-\infty < v < \infty$, define the Likelihood Ratio Process (LRP) under model M_i, by*

$$\begin{aligned} LRP(M_i) &= LRP(\theta, v, T \mid M_i) \\ &= \log \frac{dP_{X^T \mid \theta + (v I_T^{-1/2}(\theta)), M_i}}{dP_{X^T \mid \theta, M_i}}(X^T). \end{aligned} \tag{2.3}$$

Then, for $\psi_i \sim N(0,1)$,

$$LRP(M_i) = v\psi_i - \frac{1}{2}v^2. \tag{2.4}$$

Note that the property (2.4) is exact $\forall\ T > 0$ and is therefore stronger than the LAN condition that requires (2.4) to hold only as $T \to \infty$. In view of (2.4), we can apply the proposal of Polson and Roberts (1993) and show that a non-informative prior of θ, under model M_i, is

$$\pi_i^J(\theta) \propto c_i, \tag{2.5}$$

where c_i are given by

$$c_1 = I_T(\theta \mid M_1) = T/\sigma^2, \tag{2.6}$$

$$c_2 = I_T(\theta \mid M_2) = \left(T + \alpha T^2 + \frac{\alpha^2 T^3}{3}\right) \Big/ \sigma^2. \tag{2.7}$$

Remark. (i) It can be shown that the priors in (2.5) are indeed *probability exact matching priors*. Thus, these priors can also be justified from the view point of their frequentist validity.

(ii) The different orders of information in (2.6) and (2.7), in terms of T, reflect the different probability structures of the X_t process under M_1 and M_2; for instance the process X_t has much more variability, under M_1 than under M_2, for large t as reflected by their variances (see (4.13)).

3. Bayes factors

3.1 *Intrinsic and fractional Bayes factors*

The intrinsic Bayes factor (IBF). The IBF for testing a model M_i versus another model M_j is given by

$$IBF_{ij} = \frac{m_i(X^T)}{m_j(X^T)} CFI_{ji},$$

where $m_i(X^T)$ and $m_j(X^T)$, respectively, are the marginals under models M_i and M_j, based on the full sample data X^T, and CFI_{ji} is a correction factor based on a set of minimum training samples (MTS). For example, CFI_{ji} is the arithmetic average of the ratios m_j^*/m_i^*, where m_i^* and m_j^* are the marginals under M_i and M_j, respectively, based on a MTS. Typically, a MTS is a sub-sample of size equal to the number of unknown parameters in a model. For more details, see Berger and Pericchi (1996).

The fractional Bayes factor (FBF). For the models M_i and M_j, let $L_i(\theta)$ and $L_j(\theta)$, respectively, represent their likelihoods. Also, let

$$\begin{aligned} m_i^{(b_i)} &= \int L_i^{b_i(T)}(\theta)\pi_i(\theta)d\theta \\ &= c_i \int L_i^{b_i(T)}(\theta)d\theta \end{aligned} \tag{3.1}$$

where $b_i(T)$ is a fractional power to be chosen, possibly depending on the model M_i. Let $m_j^{(b_j)}$ be similarly defined, for model M_j. Then, the fractional Bayes factor for testing M_i v.s. M_j is (see O'Hagan (1995))

$$FBF_{ij} = \frac{m_i}{m_j} CFF_{ji}, \tag{3.2}$$

where CFF_{ji} is (a correction factor for FBF) given by

$$CFF_{ji} = \frac{m_j^{(b_j)}}{m_i^{(b_i)}}. \tag{3.3}$$

Note that, both IBF and FBF are products of the term m_i/m_j and a correction factor term, CFI and CFF respectively. In the context of IBF, Berger and Pericchi (1996) advocated a principle that requires the IBF to provide asymptotically an actual Bayesian analysis with default (proper) priors. DeSantis and Spezzaferri (1997) have discussed similar properties of the FBF. Thus, the correction factors must necessarily converge almost surely to positive constants as the sample-size (information) grows indefinitely large. In this paper, we follow that recommendation and seek IBF/FBFs that conform to the above principle.

Remark. When introducing the FBF, O'Hagan (1995) used the same fractional power, b, for both models that are tested by the FBF. Here, we allow the possibility that the fractional power may be different for the models that we entertain; for motivation for such a choice, see Section 5.

3.2 *The marginals under models M_0, M_1 and M_2*

First, suppose that the process X_t follows the model M_0 in (1.2). Then, the likelihood of X^T, see (2.1), is given by $L_0 = 1$; recall that the likelihood is in fact calculated with respect to the measure σdB_t itself. Hence, the marginal $m_0(X^T)$ under model M_0, is 1.

To obtain the marginal, $m_1(X^T)$ under model M_1 in (1.3), note first that the likelihood $L_1(\theta)$ is obtained by putting $f(t, X_t) = 1$ and $g(t, X_t) = 0$ in (2.1). The marginal $m_1(X^T)$, with the improper prior $\pi_1(\theta) = c_1$, is then given by

$$\begin{aligned} m_1(X^T) &= \int c_1 L_1(\theta) d\theta \\ &= c_1 \sqrt{\frac{2\pi\sigma^2}{T}} \exp\left\{\frac{(X_T - X_0)^2}{\sigma^2 T}\right\}. \end{aligned} \tag{3.4}$$

Now, suppose that the process X_t evolves according to the model M_2 given by (1.4). Then the likelihood $L_2(\theta)$ is obtained by putting $f(t, X_t) = (1 + \alpha t)$ and $g(t, X_t) = -\alpha X_t$ in (2.1). Hence, the marginal $m_2(X^T)$,under model M_2 with the improper prior $\pi_2(\theta) = c_2$, is given by

$$m_2(X^T) = c_2 \sqrt{\frac{2\pi}{A}} \exp\left\{\frac{1}{2}\frac{B^2}{A} + C\right\} \tag{3.5}$$

where

$$\begin{aligned} \sigma^2 A &= T\left(1 + \alpha T + \alpha^2 \frac{T^2}{3}\right) \\ \sigma^2 B &= X_T - X_0 + \alpha \int_0^T t dX_t + \alpha \int_0^T (1 + \alpha t) X_t dt \\ \sigma^2 C &= -\frac{\alpha}{2}(X_T^2 - X_0^2 - \sigma^2 T) - \frac{\alpha^2}{2} \int_0^T X_t^2 dt. \end{aligned} \tag{3.6}$$

Remark. When applying these formulas for the marginal beliefs to an actual discrete observation set, $\{X_{t_i}\}$ say, that is recorded on a fine-time scale, one may use the following approximations of stochastic integrals:

$$\int_0^T (1 + \alpha t) dX_t \simeq \Sigma_{t_i} (1 + \alpha t_{i-1})(X_{t_i} - X_{t_{i-1}}), \tag{3.7}$$

$$\int_0^T X_t^2 dt \simeq \Sigma_{t_i} X_{t_{i-1}}^2 (t_i - t_{i-1}). \tag{3.8}$$

4. Intrinsic Bayes factors for testing M_0, M_1, M_2

In this section, we derive the arithmetic IBF for pairwise testing of the three models and then comment on the simultaneous testing of the models.

IBF for testing M_1 v.s. M_0

Using the prior in (2.5), the marginals of a single observation at time t, i.e., (t, x_t), under the models M_i, $i = 0, 1$ are given, respectively, by

$$m^{(0)}(x_t) = \frac{1}{\sqrt{2\pi t\sigma^2}} \exp\left[-\frac{(x_t - x_0)^2}{2t\sigma^2}\right], \tag{4.1}$$

$$m^{(1)}(x_t) = \frac{c_1}{t}. \tag{4.2}$$

Thus, the Bayes factor for testing M_0 v.s. M_1, which is based on the data at time t, i.e., (t, x_t), is given by,

$$B_{01}(x_t) \propto \sqrt{\frac{t}{2\pi\sigma^2}} \exp\left[-\frac{(x_t - x_0)^2}{2t\sigma^2}\right]. \tag{4.3}$$

Minimum training samples

To define a suitable minimum training sample, note that the observed (full) data is in the form of a continuous path $\mathcal{S} = \{(t, x_t) : 0 \le t \le T\}$. Let $S \subseteq [0, T]$ be a non-empty and measurable set, and

$$x(S) = \{(t, x_t) : t \in S\}. \tag{4.4}$$

Thus we can regard $x(S)$ as a subsample of $\mathcal{S}$. Moreover, it is easy to show that the marginals based on $x(S)$, $m_i(x(S))$, satisfy the property

$$0 < m_i(x(S)) < \infty \quad \text{for} \quad i = 0, 1, \tag{4.5}$$

for all non-empty measurable subsets S of $[0, T]$. Hence, $x(S)$ is a training sample (TS), for any non-empty measurable subset S of $[0, T]$. Consequently, the set of all minimum training samples (MTS) is given by

$$\mathcal{M} \stackrel{def}{=} \{x(S) : S \subseteq [0, T] \text{ \& there is no } S_0 \subset S \text{ so that } x(S_0) \text{ is a TS}\}. \tag{4.6}$$

It is now immediately clear from (4.6) and (4.5) that for an $x(S)$ to be a minimum training sample, S has to be a singleton set. Thus, the set of MTS's is given by the (uncountably) infinite set $\mathcal{M} = \{(t, x_t) : 0 < t \le T\}$.

Correction factors

We recall that the Bayes factor corresponding to an MTS (t, x_t) is given by $B_{01}(x_t)$, as in (4.3). A correction factor, CF_{01}, is therefore a suitable summary of $B_{01}(x_t)$'s over the set of MTS's $\mathcal{M}$, or equivalently, over t. To this end, we describe below three possible approaches, depending on the MTS's used, to calculating the correction factors. The first two use continuous and discrete versions of x_t, while the third uses a transformed (i.e., differenced) version of x_t.

Use all MTS's (t, x_t).

In this case, we may define the correction factor as an integral of $B_{01}(x_t)$ over t, with respect to a suitable measure ν. Thus,

$$CFI_{01}^{(1)} = \int_0^T B_{01}(x_t)\nu(dt), \tag{4.7}$$

where, for example, one may choose $\nu(dt) = dt/T$.

Use a finite number of MTS's.

Let $n = n(T) > 0$, and $0 = t_0 < t_1 < t_2 < \cdots < t_n \leq T$ be such that $t_n \to \infty$ as $T \to \infty$. Also, let $Y_k = X_{t_k}$. Then, $B_{01}(y_k) = B_{01}(x_{t_k})$ is defined through (4.3), where $k = 1, \ldots, n$. Here, we define the correction factor by

$$CFI_{01}^{(2)} = \frac{1}{n} \sum_{k=1}^{n} B_{01}(y_k). \tag{4.8}$$

Use a finite number of MTS's based on the transformed data.

Following the above notation, the first differences of the raw data are given by

$$Z_1 = Y_1 \quad \text{and} \quad Z_k = Y_k - Y_{k-1}, \quad k \geq 2. \tag{4.9}$$

Then, each z_k is an MTS. However, the Bayes factors based on these z_k's have a common form except for z_1 and hence we drop the latter from further consideration. The Bayes factor based on z_k, namely $B_{01}(z_k)$ is the same as in (4.3), with z_k in place of $x_t - x_0$, and $\Delta_k = t_k - t_{k-1}$, in place of t. Now, we define the correction factor based on z_k's by

$$CFI_{01}^{(3)} = \frac{1}{n} \sum_{k=2}^{n} B_{01}(z_k). \tag{4.10}$$

We propose that the choice of which of the foregoing correction factors may be used in testing M_0 v.s. M_1 be based on the principle that, as t_n and $T \to \infty$, the correction factors converge to positive numbers almost surely (a.s, hereinafter) under both models. It can be argued that, for $CFI_{01}^{(1)}$ to converge in expectation to a positive number, one needs to choose a $\nu(dt)$ that depends on θ or the models themselves. Also, it can be shown that, as $\Delta_k \to$ a positive common Δ, $\forall\theta$, $CFI_{01}^{(2)}$ does not converge to a positive number a.s. Moreover, the corresponding IBF, denoted $IBF_{10}^{(2)}$, computed using this correction factor is not consistent. Specifically, we have

THEOREM 4.1. *Under M_0, $IBF_{10}^{(2)}$ does not converge to 0 in probability as n goes to ∞.*

PROOF. In the Appendix. □

Hence, we shall employ the correction factor $CFI_{01}^{(3)}$. In view of the sufficiency of the single observation at time T, namely x_T, we define the IBF for testing M_1 v.s. M_0 as

$$IBF_{10} = \frac{m^{(1)}(x_T)}{m^{(0)}(x_T)} CFI_{01}^{(3)}, \tag{4.11}$$

where $m^{(i)}(x_T)$'s are given by (4.1) and (4.2).

Remark. Using arguments similar to the proof of Theorem 4.1, it can be shown that the IBF in (4.11) is in fact consistent. Also, in this IBF, one can substitute other summaries like the median of $B_{01}(z_k)$'s for the correction factor. But we do not pursue these other options here.

We now turn to testing the model M_2 in (1.4) against one or both of models M_1 and M_0. We first observe that the solution to (1.4), X_t, is

$$X_t = X_0 e^{-\alpha t} + \theta t + \sigma e^{-\alpha t}\int_0^t e^{\alpha s} dB_s. \tag{4.12}$$

We will assume that X_0 has a known normal distribution (possibly degenerate at 0) that does not depend on θ. Thus, for instance, if $X_0 \sim N(0,1)$, then X_t is a Gaussian process with $E(X_t) = \theta t$, $V(X_t) = e^{-2\alpha t} + \sigma^2(1 - e^{-2\alpha t})/2\alpha$, and for $s \le t$,

$$\mathrm{Cov}(X_t, X_s) = e^{-\alpha(t+s)} + \sigma^2 e^{-\alpha(t-s)}(1 - e^{-2\alpha s})/2\alpha. \tag{4.13}$$

IBF for Testing M_1 v.s. M_2

With the improper prior $\pi_2(\theta) \propto c_2$ under M_2, as given by (2.5), we have that any single x_t is a minimum training sample (MTS). The marginal under M_2 of an MTS x_t, say $\hat{m}_2$, is given by $\hat{m}_2 = c_2/t$. Similarly, the marginal under M_1 of an MTS x_t, say $\hat{m}_1$, is given by $\hat{m}_1 = c_1/t$. Thus, the Bayes factor for testing M_2 v.s. M_1, which is based on the data at time t, i.e., (t, x_t), is given by $B_{21}(x_t) = c_2/c_1$. The correction factor of the arithmetic IBF is

$$CFI_{21} = \frac{c_2}{c_1},$$

and the IBF for testing M_1 v.s. M_2 is

$$IBF_{12} = \frac{m_1}{m_2}, \tag{4.14}$$

where m_1, m_2 are respectively given by (3.4) and (3.5). Thus, in testing M_1 v.s. M_2, the CBF (ignoring constants c_i's) is "well-calibrated", as might have been anticipated.

IBF for testing M_0 v.s. M_2

In this case, we find that use of x_t's or even the differences z_k in (4.9) as the training samples does not necessarily lead to a correction factor with desirable asymptotics. For instance, choosing z_k's as the training samples, we see that the correction factor is given by

$$CFI_{02}^{(z)} = \frac{1}{n}\sum_{k=2}^{n} B_{02}(z_k), \tag{4.15}$$

where

$$B_{02}(z_k) = \frac{1}{c_2}\sqrt{\frac{\Delta_k}{2\pi\sigma^2}}\exp\left[-\frac{z_k^2}{2\Delta_k\sigma^2}\right]. \tag{4.16}$$

Under model M_0, z_k's are independent and $CFI_{02}^{(z)}$ will converge to a positive constant. But, in view of (4.13), the z_k's are still correlated under M_2.

It is difficult to determine whether $CFI_{02}^{(z)}$ converges to a positive constant a.s. as n goes to ∞, and we suspect it does not. To achieve asymptotic stability of correction factors, we now propose a more general transformation than differencing of Y_k. To this end, in order to simplify presentation, we assume that $\Delta_k = \Delta$ for all k and hence $t_k = k\Delta$. (Final conclusions are still valid in the more general case provided Δ_k converges to a positive number as k goes to ∞.) An exploration of the structure of Y_k's under model M_2 reveals that we should consider the W_k's given by

$$U_k = z_k - e^{-\alpha\Delta} z_{k-1}, k \geq 2 \quad \text{and} \quad W_k = U_{2k+1},\ k \geq 1. \tag{4.17}$$

Note that, when $\alpha = 0$, the above transformation reduces to (4.9), the one meant for M_1, which is natural as M_2 reduces to M_1 for $\alpha = 0$. Moreover, as α is a known parameter in our setup, the data given by the transformation in (4.17) can be used in our model selection procedures. We now state the probabilistic structure of the transformed data, W_k, $k \geq 1$, under all three models, the proof of which is given in the Appendix.

THEOREM 4.2. (i) *Under M_2 with $\alpha > 0$, W_k's are i.i.d. random variables having normal distribution with mean $\theta\mu(\alpha)$ and variance $\sigma^2 v(\alpha)$ where*

$$\mu(\alpha) = \Delta(1 - e^{-\alpha\Delta}) \quad \textit{and} \quad v(\alpha) = (1 - e^{-2\alpha\Delta})/\alpha$$

(ii) *Under M_1, W_k's are i.i.d. with $N(\theta\mu(\alpha), \sigma^2\Delta(1 + e^{-2\alpha\Delta}))$ distribution.*
(iii) *Under M_0, W_k's are i.i.d. with $N(0, \sigma^2\Delta(1 + e^{-2\alpha\Delta}))$ distribution.*

Now, using w_k as the training samples, the correction factor for testing M_2 v.s. M_0 may be defined by

$$CFI_{02} = \frac{1}{n}\sum_{k=1}^{n} B_{02}(w_k), \tag{4.18}$$

where, in view of Theorem 4.2, one can show that

$$B_{02}(w_k) = \frac{1}{c_2}\sqrt{\frac{\Delta g(\alpha)}{2\pi\sigma^2}} \exp\left[-\frac{w_k^2}{2\Delta\sigma^2(1 + e^{-2\alpha\Delta})}\right], \tag{4.19}$$

where $g(\alpha) = (1 - 2e^{-\alpha\Delta}(1 + e^{-2\alpha\Delta})^{-1})$. Again using Theorem 4.2, it is now easy to verify that CFI_{02} does tend to a positive number as n goes to ∞. The IBF for testing M_0 v.s. M_2 is then given by

$$IBF_{02} = \frac{1}{IBF_{20}}, \tag{4.20}$$

where $IBF_{20} = \frac{m_2}{m_0} CFI_{02}$, $m_0 = 1$ and m_2 is given by (3.5).

Remarks. (i) In view of (4.17), it is clear that

$$z_k = e^{-(k-1)\alpha\Delta} z_1 + U_k + \sum_{j=1}^{k-2} e^{-j\alpha\Delta} U_{k-j} \tag{4.21}$$

and it is easy to show from this that z_k's are correlated.

(ii) Definitions such as (4.20) have been used in other contexts by Berger and Pericchi (1996).

Simultaneous testing of M_0, M_1 and M_2.

Our interest has so far been in testing the models pairwise. Here, the transformation (4.17), although it would work, would not be very meaningful if we are only testing M_1 v.s. M_0 because, while these two models are free of α, the transformation involves α. Thus, in order to obtain asymptotic stability of the correction factors when using IBF approach, we have proposed the use of different transformations to suit the pairwise testing. Specifically, we have used the transformation (4.9) when testing M_1 v.s. M_0, and (4.17) when testing M_0 v.s. M_2, and did not transform the raw data for testing M_1 v.s. M_2. However, if we are interested in jointly testing all three models, it would be meaningful to have a single transformation that will simultaneously stabilize the correction factors of all three IBF's for testing M_0 v.s. M_1, M_1 v.s. M_2 and M_2 v.s. M_0, as these are needed to jointly evaluate all three models. Fortunately, the transformation (4.17) does provide a unified tool to accomplish this task because, in view of Theorem 4.2, the transformation produces independent data no matter which model is the correct one. Here, we omit further details as they are similar to our earlier discussion of the pairwise testing problem.

5. Fractional Bayes factors for testing M_0, M_1 and M_2

We begin this section with testing M_1 v.s. M_0.

LEMMA 5.1. *For $0 < b \le 1$, we have*

$$m_0^{(b)} = 1$$
$$m_1^{(b)} = c_1\sqrt{\frac{2\pi\sigma^2}{Tb}}\exp\left[\frac{b}{2T\sigma^2}(x_T - x_0)^2\right]$$
$$CFF_{01}^{(b)} = 1/m_1^{(b)}.$$

Using the above expressions, and with $m_i = m_i^{(1)}$, $i = 0, 1$, we obtain $FBF_{10}^{(b)}$, the FBF for testing M_1 v.s. M_0 as

$$FBF_{10}^{(b)} = \frac{m_1}{m_0}CFF_{01}^{(b)}, \tag{5.1}$$

LEMMA 5.2. *Let $b = 1/T$. Then, as $T \to \infty$, under models M_0 and M_1*

$$CFF_{01}^{(b)} \stackrel{as}{\to} \frac{1}{c_1\sqrt{2\pi\sigma^2}}\exp\left(-\frac{\theta^2}{2\sigma^2}\right)$$

PROOF. From (A.3) in the Appendix, as $T \to \infty$, under models M_0 and M_1, we obtain $(X_T - X_0)^2/T^2\sigma^2 \to^{as} \theta^2/\sigma^2$. The result follows readily from this convergence.

Turning to testing M_2 v.s. M_1, M_0, we note that a commonly recommended choice for the fractional power b is m/n, where m and n respectively are the size of a MTS and the size of the full sample. In the context of our continuous sampling from a continuous time process (with one unknown parameter) fully over the interval $0 \le t \le T$, a similar choice for b would be $b = 1/T$. The choice of b can affect the asymptotic behavior of the correction factor CFF, and hence is crucial when desirable asymptotic behavior for CFF (as outlined earlier) is sought.

Indeed, we saw in Lemma 5.2 that the choice of a *common training fraction*, namely $b = 1/T$, was satisfactory for testing M_1 v.s. M_0. Such a choice of b, however, is not satisfactory, as we show below, for testing M_1 v.s. M_2 and M_0 v.s. M_2. In particular, we find it necessary to use different choices of b's in the numerator and denominator of (3.3), in order to ensure that CFF_{21} and CFF_{20} converge a.s. to positive constants. More specifically, we shall see that the choice of b needs to be, in some sense, related to the size of the 'information' in the sample under each model, and that, when this information differs in order of magnitude between the two models being compared, different b's would be required for the two models, to ensure satisfactory convergence of CFF_{21} and CFF_{20}. To this end, note that the Fisher informations for models M_1 and M_2, respectively given by (2.6) and (2.7), are of the orders T and $T + \alpha T^2 + \alpha^2 T^3/3$. It is noteworthy here that fractions other than $b = 1/n$ have indeed been recommended in the literature; for instance, Berger and Pericchi (1997) proposed training fractions other than $1/n$ in the context of the Neyman-Scott problem.

In the following, we derive the asymptotic behavior of the various marginals, and hence show (Theorem 5.2) that, when the *common training fraction* $b(T) = 1/T$ is used for both numerator (models M_0 and M_1) and denominator (model M_2) in (3.3), CFF_{21} and CFF_{20} do not converge to positive constants, and that for the choices of $b_1 = 1/T$ for model M_1 and $b_2 = 1/v(T)$, where $v(T) = T + \alpha T^2 + \alpha^2 T^3/3$ for model M_2, the correction factors do converge a.s. to positive constants.

For use in the following, we use the notation $\tilde{m}_i$ for $m_i^{(b)}$, $i = 1, 2$, and $CFF_{21}^{(1)}$ for the corresponding correction factor when the same $b = 1/T$ is used (in 3.3) for both models. Likewise, when $b_1 = 1/T$ and $b_2 = 1/v(T)$ are used for models M_1 and M_2 respectively in (3.3), we use the notation m_i^* for $m_i^{(b)}$, and $CFF_{21}^{(2)}$, for the corresponding correction factor. Here, m_1^*, which corresponds to M_1 or $\alpha = 0$, is calculated with the fraction $1/T$ (and hence $\tilde{m}_1 = m_1^*$). Also, m_2^*, which corresponds to M_2 with a known $\alpha > 0$, is calculated with the fraction $1/v(T) = 1/(T + \alpha T^2 + \alpha^2 T^3/3)$. For convenience, we also let $\tilde{m}_i$ and m_i^*, $i = 0, 2$ (with $\tilde{m}_0 = m_0^* = 1$), $CFF_{20}^{(1)}$, and $CFF_{20}^{(2)}$ have similar definitions in testing M_0 v.s. M_2.

THEOREM 5.1. *Under models M_1 and M_2,*

$$\tilde{m}_1 \overset{a.s.}{\to} c_1 \sigma \sqrt{2\pi} e^{\theta^2/2\sigma^2}. \tag{5.2}$$

THEOREM 5.2. (i) *Under M_2, $\tilde{m}_2 \to 0$, in probability, and hence, $CFF_{21}^{(1)} = \tilde{m}_2/\tilde{m}_1$ and $CFF_{20}^{(1)} = \tilde{m}_2/\tilde{m}_0$ converge, in probability, to 0 as T goes to ∞.*

(ii) *Under models M_1 and M_2, $m_2^* \to^{a.s.} c_2 \sigma \sqrt{2\pi}$, and hence $CFF_{21}^{(2)} = \frac{m_2^*}{m_1^*} \to^{a.s.} \frac{c_2}{c_1} e^{-\theta^2/2\sigma^2}$.*

(iii) *Under models M_0 and M_2, $CFF_{20}^{(2)} = \frac{m_2^*}{m_0^*} \to^{a.s.} c_2 \sqrt{2\pi}$.*

Proofs of the theorems are given in the Appendix.

Motivated by the foregoing asymptotic considerations, we suggest that we replace the fraction b in the definition of the FBF by a vector $\boldsymbol{b} = (b_1, b_2)$ to facilitate further discussion of the FBF. Specifically, in our context, we distinguish and recognize the following FBF's for testing M_1 v.s. M_2:

$$FBF_1 = \frac{m_1}{m_2} CFF_{21}^{(1)}, \tag{5.3}$$

where $\boldsymbol{b} = (T, T)$ uses a *common training fraction*, and

$$FBF_2 = \frac{m_1}{m_2} CFF_{21}^{(2)}, \tag{5.4}$$

where $\boldsymbol{b} = (T, (T + \alpha T^2 + \frac{\alpha^2 T^3}{3}))$ consists of *different training fractions* that are proportional to the information numbers for the models being tested. Using $CFF_{20}^{(1)}$ and $CFF_{20}^{(2)}$, one can define similar FBF's for testing M_0 v.s. M_2.

Remarks. Although we do not provide the details here, it can be shown that both FBF_1 and FBF_2 are consistent for testing M_1 v.s. M_2, in the sense that both tend to ∞ (respectively, 0), in probability, under M_1 (respectively, M_2). Thus, these facts combined with Theorem 5.2 suggest that we should prefer using FBF_2 to FBF_1 if we were to require asymptotic stability of the correction term. A similar recommendation can be made in testing M_0 v.s. M_2. Regarding the simultaneous testing of the three models M_0 to M_2, the correction factors in (3.3), where $b_0 = 1/T$, $b_1 = 1/T$, and $b_2 = 1/(T+\alpha T^2 + \frac{\alpha^2 T^3}{3})$, will suffice, as they all converge to positive constants asymptotically. We omit the details, as they are similar to the earlier discussion in this section.

6. Conclusions

In this paper, we have used the IBF and FBF approaches to test, under Jeffreys-type probability matching improper priors, the hypotheses that an observable continuous-time process is a Brownian motion with no drift, Brownian motion with drift and a mean reversion process. In doing so, we have adhered to the Berger-Pericchi's principle that for a Bayes factor to be reasonable it should correspond (at least asymptotically) to some (reasonable default) priors. We have used this principle as a guide in evaluating, and seeking suitable "adjustments" for, the IBF and FBF in order that their correction factors are asymptotically stable. Except when testing the mean-reversion model against the Brownian motion with drift, we have found that an "automatic" use of the IBF methodology does not lead to an asymptotically stable correction factor. In other pairwise and joint tests of the models, we need to first "ungroup" the data using a single transformation that will simultaneously stabilize, under all the relevant models, all the correction factors involved in the (arithmetic) IBF's needed for the testing. We have provided a novel solution to this non-trivial task. Next, whenever the mean-reversion is included in the model selection, the fraction of the data used in the correction term of the FBF needs careful attention. In particular, we have shown that, when *different training fractions* that depend on model-specific Fisher information are used, the correction factor of the FBF is asymptotically stable. It should be noted that these model specific training fractions remain the same no matter what the alternative models considered for comparison are. To sum up, when selecting a model for our correlated diffusion process, careful consideration must typically be given to data transformations and to choosing possibly different training fractions to enable the correction factors in the IBF and FBF to stabilize asymptotically.

Acknowledgements

The authors would like to express their deep gratitude to two anonymous referees for several important comments which have considerably improved the substance and presentation of the article.

Appendix

We now provide proofs of our results in Sections 2 to 5. In the following, the operators P_i, E_i, F_i respectfully refer to the probability, expectation and the cumulative distribution function under model M_i.

PROOF OF THEOREM 2.1. It is straightforward to show that part (i) holds. Turning to part (ii), we note that $LRP(M_i) = v\psi_i - \frac{1}{2}v^2\xi_i$, where

$$\psi_i = \frac{1}{\sigma\sqrt{I_T(\theta)}}\int_0^T f_i(t)dB_t \tag{A.1}$$

and

$$\xi_i = \frac{1}{\sigma^2 I_T(\theta)}\int_0^T f_i^2(t)dt, \tag{A.2}$$

where $f_1(t) = 1$ and $f_2(t) = 1 + \alpha t$. It follows from part (i) and (6.2) that $\xi_i \equiv 1$. Also, since

$$E_i(\psi_i^2) = \frac{1}{\sigma^2 I_T(\theta)}\int_0^T f_i^2 dt = 1,$$

we obtain $\psi_i \sim N(0,1)$. □

PROOF OF THEOREM 4.1. For simplicity, we assume that $t_k = k$, $Y_k = X_{t_k}$ for $k = 0, 1, \ldots, n = T$, and that Y_k are observed for $k = 1, \ldots, n$. Using the sufficiency of Y_n, $IBF_{10}^{(2)}$ is given, as in (4.11), by

$$IBF_{10}^{(2)} = \frac{m^{(1)}(x_n)}{m^{(0)}(x_n)} CFI_{01}^{(2)} \stackrel{def}{=} IBF_n,$$

where $CFI_{01}^{(2)} = \frac{1}{n}\sum_{k=1}^n B_{01}(Y_k)$. Note that

$$\begin{aligned} IBF_n &= \frac{1}{n^{3/2}}\exp\{(Y_n - Y_0)^2/2n\sigma^2\}\sum_{k=1}^n \sqrt{k}\exp\{-(Y_k - Y_0)^2/2k\sigma^2\} \\ &> \frac{1}{n^{3/2}}\sum_{k=1}^n \sqrt{k}\exp\{-(Y_k - Y_0)^2/2k\sigma^2\} \\ &\stackrel{def}{=} B_n. \end{aligned}$$

It is easy to see that $E_0(B_n)$ converges to a positive constant c as n goes to ∞. Thus, for (any) small $\epsilon > 0$, we have for large n,

$$c - \epsilon < \int B_n dF_0(B_n) \le \epsilon + \int_{B_n > \epsilon} B_n dF_0(B_n),$$

where $F_0(B_n)$ denotes the cdf of B_n. Since B_n is positive and bounded above, say, by $K > 0$, we have from the above,

$$c - \epsilon < \epsilon + KP_0(B_n > \epsilon).$$

Now, since $IBF_n > B_n$, we have, for large n,

$$P_0(IBF_n > \epsilon) > P_0(B_n > \epsilon) > \frac{(c-2\epsilon)}{2K}.$$

Hence, $P_0(IBF_n > \epsilon)$ does not converge to 0 as n goes to ∞, proving the desired result. □

PROOF OF THEOREM 4.2. We prove part (i). Proof of parts (ii) and (iii) are straightforward and hence will be omitted. Using the representation (4.12) for $Y_k = X_{k\Delta}$, we can show that

$$V_k \stackrel{def}{=} e^{\alpha k\Delta}U_k = \theta\Delta(1-e^{-\alpha\Delta})e^{\alpha k\Delta} + \sigma\int_{(k-1)\Delta}^{k\Delta} e^{\alpha s}dB_s - \sigma e^{\alpha\Delta}\int_{(k-2)\Delta}^{(k-1)\Delta} e^{\alpha s}dB_s.$$

Now, it is clear that V_{2k+1} are independent and normally distributed for $k = 1, 2, \ldots$. Moreover, using routine calculation, we can show that

$$E_2(V_k) = \theta\Delta(1-e^{-\alpha\Delta})e^{\alpha k\Delta} \quad \text{and} \quad \mathrm{Var}_2(V_k) = \sigma^2 e^{2\alpha k\Delta}(1-e^{2\alpha\Delta})/\alpha.$$

Thus $U_k = e^{-\alpha k\Delta}V_k$ are identically distributed with $N(\theta\mu(\alpha), \sigma^2 v(\alpha))$ distribution, and U_{2k+1} are i.i.d. for k=1,2, ..., proving part (i). □

PROOF OF THEOREM 5.1. Putting $f(t, X_t) = 1$ and $g(t, X_t) = 0$ in (2.1), and using (3.1) with $i = 1$ and $b_1(T) = 1/T$, we get

$$\log \tilde{m}_1 = \log(c_1\sigma\sqrt{2\pi}) + \frac{(X_T - X_0)^2}{2\sigma^2 T^2}.$$

Now, under M_1, $X_T - X_0 = \theta T + B_T$, which gives

$$(X_T - X_0)^2/\sigma^2 T^2 \stackrel{a.s.}{\to} \theta^2/\sigma^2 \tag{A.3}$$

This proves (5.2) under M_1. Proof under M_2 is similar, using (4.12). □

PROOF OF THEOREM 5.2. We only prove the parts (i) and (ii) since the proof of part (iii) follows immediately.

For convenience, we take $\sigma^2 = 1$ and $X_0 = 0$ in the following. First we note that $m_2^{(b)}(X^T)$ can be written as

$$m_2^{(b)}(X^T) = c_2\sqrt{\frac{2\pi}{A_1}}\exp\left\{\frac{1}{2}\frac{B_1^2}{A_1} + C_1\right\} \tag{A.4}$$

where, with A, B, C given by (3.6), $A_1 = A \cdot b(T), B_1 = B \cdot b(T)$ and $C_1 = C \cdot b(T)$. (In the following, except for A and A_1, which are non-random functions of T, we use small-case letters to denote non-random functions of T, and capital letters to denote random functions of T.)

We assume that the process X_t follows the model M_1, and let

$$Y_t = e^{-\alpha t}\int_0^t e^{\alpha s}dB_s. \tag{A.5}$$

Then, using (4.12) and (3.6), we can write, after some simplification, $b(T)^{-1} \cdot B_1 = p(T) + \alpha V_1$, where

$$V_1 = \int_0^T t dY_t + \int_0^T Y_t dt + \alpha \int_0^T tY_t dt + Y_T,$$

and

$$b(T)^{-1} \cdot C_1 = q(T) - \frac{\alpha}{2}\left(2\theta T Y_T + Y_T^2\right) - \frac{\alpha^2}{2}\int_0^T (2\theta t Y_t + Y_t^2) dt,$$

where $p(T) = \theta(T + \alpha T^2 + \alpha^2 \frac{T^3}{3})$ and $q(T) = -\frac{\alpha}{2}(\theta^2 T^2 - T) - \frac{\alpha^2\theta^2}{6}T^3$. Using integration by parts, e.g., see Øksendal (1985), we can re-write V_1 as,

$$V_1 = TY_T + Y_T + \alpha \int_0^T tY_t dt. \tag{A.6}$$

Now, from (A.4),

$$\begin{aligned} &2\log m_2^{(b)}(X^T) \\ &= \frac{B_1^2}{A_1} + 2C_1 - \log A_1 + 2\log c_2 \\ &= s(T) + 2\alpha p(T) b^2(T)\frac{V_1}{A_1} + \frac{\alpha^2 V_1^2 b^2(T)}{A_1} - 2\alpha\theta T b(T) Y_T - 2\alpha^2\theta b(T)\int_0^T tY_t dt \\ &\quad -b(T)\left(\alpha Y_T^2 + \alpha^2 \int_0^T Y_t^2 dt\right) - \log A_1 + 2\log(c_2\sqrt{2\pi}), \end{aligned} \tag{A.7}$$

where

$$s(T) = b^2(T)\left(\frac{p^2(T)}{A_1} + 2\frac{q(T)}{b(T)}\right).$$

Using (A.6), we can simplify (A.7), and write

$$\begin{aligned} 2\log m_2^{(b)}(X^T) &= s(T) + 2\alpha\theta b(T) Y_T + \frac{\alpha^2 V_1^2 b^2(T)}{A_1} \\ &\quad -b(T)\left(\alpha Y_T^2 + \alpha^2\int_0^T Y_t^2 dt\right) - \log A_1 + 2\log(c_2\sqrt{2\pi}). \end{aligned} \tag{A.8}$$

Now, to prove part (i) of the theorem, we put $b(T) = 1/T$ in the above. Then, we have $s(T) = \theta^2 + \alpha$, $A_1 = A/T$, and

$$\begin{aligned} 2\log \tilde{m}_2(X^T) &= \theta^2 + \alpha + \frac{\alpha^2 V_1^2}{T^2 A_1} + 2\alpha\theta\frac{Y_T}{T} - \log A_1 + 2\log(c_2\sqrt{2\pi}) \\ &\quad -\frac{1}{T}\left(\alpha Y_T^2 + \alpha^2\int_0^T Y_t^2 dt\right). \end{aligned}$$

Thus,

$$2\log \tilde{m}_2(X^T) \le V_2 - \log A_1 + 2\log(c_2\sqrt{2\pi}), \tag{A.9}$$

where

$$V_2 = (\theta^2 + \alpha) + \frac{\alpha^2 V_1^2}{T^2 A_1} + 2\alpha\theta \frac{Y_T}{T}. \tag{A.10}$$

We note, from (A.5), that Y_T has normal distribution with mean 0 and variance $\sigma^2(1 - \exp\{-2\alpha T\}/2\alpha)/(2\alpha)$, and hence the third term in the above, Y_T/T, converges to 0, a.s. w.r.t. $P_2(X^T)$, as T goes to ∞. Now, to see the limit of the second term on the right of (A.10), we note that A_1 is of the order T^2, and consider

$$\frac{V_1}{T^2} = \alpha \frac{Y_T}{T} + \frac{Y_T}{T^2} + \frac{\alpha}{T^2} \int_0^T t Y_t dt. \tag{A.11}$$

As before, we can deduce that the first two terms on the right of (A.11) converges to 0 a.s. as T goes to ∞. It can also be shown, using (A.5), that the third term on the right of (A.11) has normal distribution with mean 0 and variance of the order $1/T$, and hence converges to 0 in probability as T goes to ∞. Thus, the second term on the right of (A.10) converges to 0, in probability, as T goes to ∞, and hence V_2 converges to $(\alpha + \theta^2)$, in probability, as T goes to ∞. We can now write, from (A.9)

$$\tilde{m}_2(X^T) \le c_2 \sqrt{\frac{2\pi}{A_1}} \exp\{V_2/2\},$$

and conclude that $\tilde{m}_2(X^T)$ converges to 0 , in probability, as T goes to ∞, proving (i).

To prove (ii), under M_2, we put $b(T) = 1/v(T) = 1/A$ in (A.9). Then, $A_1 = 1$, and we can verify that $s(T)$ converges to 0 as T goes to ∞. Thus, we have

$$\begin{aligned} 2 \log m_2^*(X^T) &= s(T) + 2\alpha\theta \frac{Y_T}{A} + \frac{\alpha^2 V_1^2}{A^2} \\ &\quad - \alpha \frac{Y_T^2}{A} - \frac{\alpha^2}{A} \int_0^T Y_t^2 dt + 2\log(c_2\sqrt{2\pi}). \end{aligned}$$

We now re-write the above as

$$2 \log m_2^*(X^T) = s(T) + V_3 - \frac{\alpha^2}{A} \int_0^T Y_t^2 dt + 2\log(c_2\sqrt{2\pi}), \tag{A.12}$$

where V_3 can be shown, as before, to converge to 0 a.s., as T goes to ∞. Moreover, we can show, using (A.5), that the third term in (A.12) has an expectation which is of the order $1/T^2$, and hence converges to 0 a.s. as T goes to ∞. Hence, from (A.12), we have that $m_2^*(X^T)$ converges to $c_2\sqrt{2\pi}$ a.s. as T goes to ∞. The proof of (ii) under M_1 is similar and hence omitted. □

References

Berger, J. O. and Pericchi, L. R. (1996). The intrinsic Bayes factor for model selection and prediction, *J. Amer. Statist. Assoc.*, **91**, 109–122.

Berger, J. O. and Pericchi, L. R. (1997). On criticisms and comparisons of default Bayes factors for model selection and hypothesis testing, *Proceedings of the Workshop on Model Selection* (ed. W. Racugno), 1–50, Bologna, Italy.

Conigliani, C. and O'Hagan, A. (1998). Computing fractional Bayes factors via two-dimensional path sampling (preprint).

De Santis, F. and Spezzaferri, F. (1997). Alternative Bayes factors for model selection, *Canad. J. Statist.*, **25**, 503–515.

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**, 693–710.

O'Hagan, A (1995). Fractional Bayes factors for model comparison (with discussion), *J. Roy. Statist. Soc. Ser. B*, **57**(1), 99–138.

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

Øksendal, B. (1985). *Stochastic Differential Equations: An Introduction with Applications*, Springer, Berlin.

Polson, N. G. and Roberts, G. O. (1993). A utility based approach to information for stochastic differential equations, *Stochastic Process. Appl.*, **48**, 341–356.

Polson, N. G. and Roberts, Gareth O. (1994). Bayes factors for discrete observations from diffusion processes, *Biometrika*, **81**, 11–26.

Sivaganesan, S. and Lingham, R. T. (2000). Bayes factors for a test about the drift of the Brownian motion under noninformative priors, *Statist. Probab. Lett.*, **48**, 163–171.