*Asian Journal of Statistical Sciences*
Vol. 3, No-1, (2023), pp. 01-12

ARF INDIA
Academic Open Access Publishing
Journal homepage link:
http://arfjournals.com/ajss

# Model Selection using Bayes Factors for the Black-Karasinski Models

Jaya P. N. Bishwal

*Department of Mathematics and Statistics University of North Carolina at Charlotte*

**ARTICLE HISTORY**
Compiled August 28, 2023

Received 1 January 2022; Accepted 18 July 2023

**ABSTRACT**
The paper studies model selection through consistency of several Bayes factors for the Black-Karasinski models with proper informative prior and noninformative prior. We consider continuous and discrete (both dense and fixed time intervals) observations of the process.

**KEYWORDS**
Black-Karasinski model, Ornstein-Uhlenbeck process, Vasicek model, Bayes factor, mean reversion, model choice, consistency, proper prior, noninformative prior, continuous observation, discrete observations.

*2020 Mathematics Subject Classification:* 60F05, 60H10, 62F12, 62F15, 62M05, 91G30, 91G60

## 1. Introduction

The Ornstein-Uhlenbeck (O-U) process, also called the Vasicek model in finance, is being extensively used in finance these days as one-factor short term interest rate model and also as a stochastic volatility model with a positive Levy process driving term, see Bishwal (2022a). Nonhomogeneous extensions with time varying drift functions of the O-U process is known as Ho-Lee model (volatility is constant) and Hull-White model (volatility is time-varying) in finance, see Hull (2015). In the homogeneous case, it is well known that this simple O-U model shows altogether different behaviors in different parts of the parameter space, the negative reals, the positive reals and zero. In view of this, there naturally arises the problem of model selection. We handle this task within a Bayesian paradigm. Note that the study of the behaviour of Bayes factors for competing dependent models has not been paid much attention yet in finance literature. This is an alternative approach to classical hypothesis test. The Bayes factor can be calculated via Girsanov's formula and approximations. The advantages of Bayes factor rather than a classical hypothesis test has been widely discussed, see Berger and Sellke (1987). There is no need to compute the asymptotic distributions of the functionals

CONTACT Jaya P. N. Bishwal Email: J.Bishwal@uncc.edu

of the process or to compute the conditional distribution given the ancillary statistics. Liptser and Shiryayev (1978) studied testing of hypotheses of zero versus non-zero drift of the O-U process. Sequential test for the drift of a Wiener process with smooth prior was studied in Simons *et al.* (1989). For the asymptotic distribution approach, Bernstein-von Mises theorem along with asymptotic behavior of the Bayes estimators and the rate of convergence for the Ornstein-Uhlenbeck process was studied in Bishwal (2000, 2001, 2008). Bishwal (2011a) studied simulated Milstein approximation of posterior density of diffusions along with Bernstein-von Mises theorem and properties of Bayes estimators. Berry-Esseen bounds of approximate Bayes estimators and approximate maximum a posteriori estimators for the discretely observed Ornstein-Uhlenbeck process were studied in Bishwal (2021a).

On a stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$, consider the Ornstein-Uhlenbeck process satisfying the stochastic differential equation

$$dX_t = \theta X_t dt + \sigma dW_t,\ t \geq 0, \quad X_0 = \xi$$

where $\{W_t\}_{t\geq 0}$ is a standard Wiener process adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and $\sigma > 0$ and $\theta \in \mathbb{R}$ are the unknown parameters. Note that the behaviour of the process depends on both the initial condition $\xi$ and the parameter space. Classically, it has been assumed that $\xi$ is either has a normal distribution or a nonzero constant and $\theta < 0$ which makes the process stationary with Gaussian invariant distribution. If $\xi = 0$ is with $\theta < 0$, then the process is asymptotically stationary and ergodic. In the above two cases the model satisfies the LAN (local asymptotic normality) property. With $\xi$ a nonzero constant and $\theta > 0$ the process is transient and satisfies the LAMN (local asymptotic mixed normality) property. With $\theta = 0$, the process is nonstationary and satisfies the LABF (local asymptotic Brownian functional) property. For all $\theta \in \mathbb{R}$, the model satisfies the LABF property, see Bishwal (2018) for the definitions of these LAN, LAMN and LABF properties.

Logarithmic interest rate $X_t = \log R_t$ evolves according to the following models under the hypotheses $H_0, H_1$ and $H_2$:

$$H_0 : dX_t = \theta dt + \sigma dW_t,\ t \geq 0,$$

$$H_1 : dX_t = -\alpha X_t dt + \sigma dW_t,\ t \geq 0,$$

$$H_2 : dX_t = \{\theta - \alpha(X_t - \theta t)\}dt + \sigma dW_t,\ t \geq 0.$$

Here $\alpha > 0$ is known as the mean reversion speed and $\theta$ is known as the mean reversion level. Under $H_0$, the interest rate $R_t$ is the Randleman-Bartter (R-B) model. Under $H_1$, the interest rate $R_t$ is the Geometric Ornstein-Uhlnbeck (G-O-U) model. Under $H_2$, the interest rate $R_t$ is the Black-Karasinski (B-K) model. See Black and Karasinski (1991) and Hull (2015). The main disadvantage of R-B model is that it does not capture the mean reversion of the interest rate. For $\theta = 0$, the B-K model becomes the G-O-U model, for $\alpha = 0$, the B-K model becomes the R-B model. Recall that the process $X_t$ under model $H_0$ follows the famous Bachelier model, also known as arithmetic Brownian motion. Turfus (2019) provided a systematic derivation of an Arrow-Debrew pricing formula for European option of the B-K model using a Green's function approach. As an inverse problem to pricing, Bishwal (2022b) obtained Berry-Esseen inequalities of estimators for the fractional Black-Karasinski model of term structure of interest rates which incorporates long memory.

Let the continuous realization $\{X_t, 0 \le t \le T\}$ be denoted by $X_0^T$. The marginal belief (marginal law) under the hypothesis $H_i$ is defined as

$$P^T(X_0^T|H_i) = \int L_T(\theta|H_i)p(\theta|H_i)d\theta, \quad i = 0, 1, 2$$

where $L_T(\theta|H_i)$ is the likelihood under $H_i$ and $p(\theta|H_i)$ is the prior density under the hypothesis $H_i$.

The Bayes factor (Jeffreys (1961)) between two competing models $H_i$ and $H_j$ is given by the ratio of the marginal beliefs defined as

$$BF_{ij}(X_0^T) := \frac{P^T(X_0^T|H_i)}{P^T(X_0^T|H_j)}, \quad i, j = 0, 1, 2.$$

First, we consider proper prior and obtain consistency or inconsistency of the models. We model the drift parameter in the continuous observation case. Then we model both the drift and the volatility parameters in the discrete observation case. Our approach to discrete observations is approximating the Bayes factor between competing continuous models.

The rest of the paper is organized as follows: In section 2, we study consistency of the Bayes factor for continuous observation. In section 3, we study consistency of the Bayes factor for discrete observations. In section 4, we study Bayes factor for Ornstein-Uhlenbeck process with mean reversion. In Section 5, we consider intrinsic Bayes factors for both simple and mean-reverting Ornstein-Uhlenbeck processes. In Section 6, we study fractional Bayes factors. Section 7 concludes.

## 2. Continuous Observation

Let the realization $\{X_t, 0 \le t \le T\}$ be denoted by $X_0^T$. The unown parameter $\sigma$ is almost surely determined by the continuous observation of the process. Hence we assume it to be known and without loss of generality assume $\sigma = 1$. We will return to the inference on $\sigma$ when we consider discrete observations in the next section. Hence the only unknown parameter is $\theta$. Let $P_\theta^T$ be the measure generated on the space $(C_T, \mathcal{B}_T)$ of continuous functions on $[0, T]$ with the Borel-$\sigma$ algebra $\mathcal{B}_T$ under the supremum norm and let $P_0^T$ be the standard Wiener measure. It is well known that $P_\theta^T \ll P_0^T$ and the Radon-Nikodym derivative (likelihood) under $H_0$ is given by

$$\frac{dP_\theta^T}{dP_0^T}(X_0^T) := L_T(\theta) = \exp\left\{\theta \int_0^T dX_t - \frac{\theta^2}{2}\int_0^T dt\right\},$$

the Radon-Nikodym derivative (likelihood) under $H_1$ is given by

$$\frac{dP_\alpha^T}{dP_0^T}(X_0^T) := L_T(\alpha) = \exp\left\{-\alpha \int_0^T X_t dX_t - \frac{\alpha^2}{2}\int_0^T X_t^2 dt\right\},$$

the Radon-Nikodym derivative (likelihood) under $H_2$ is given by

$$\frac{dP^T_{\theta,\alpha}}{dP^T_0}(X^T_0) := L_T(\theta,\alpha) = \exp\left\{\int_0^T [\theta - \alpha(X_t - \theta t)]dX_t - \frac{1}{2}\int_0^T [\theta - \alpha(X_t - \theta t)]^2 dt\right\},$$

see Liptser and Shiryayev (1978). We assume normality of the prior. For the G-O-U model $dX_t = \theta X_t dt + \sigma dW_t$, when $\theta|H_i \sim \mathcal{N}(\gamma_i, \tau_i^2)$ with known hyperparameters $\gamma_i, \tau_i$, the marginal beliefs are given by (see Polson and Roberts (1994))

$$2\log P^T(X_0^T|H_i) = \frac{1}{1+\tau_i^2 A_i}(\tau_i^2 B_i^2 + 2\gamma_i B_i - \gamma_i^2 A_i) + 2C_i - \log(\tau_i^2 A_i + 1))$$

where

$$A_i = \int_0^T X_t^2 dt,\ B_i = \int_0^T X_t dX_t - \int_0^T X_t dt,\ C_i = X_T - \frac{1}{2}\int_0^T X_t^2 dt.$$

Under a normal prior $\theta \sim N(\gamma,\ \tau^2)$, the posterior distribution can be calculated via Bayes theorem as

$$\theta|X_0^T \sim N\left(\frac{\int_0^T X_t dX_t + \gamma\tau^{-2}}{\int_0^T X_t^2 dt + \tau^{-2}},\ \frac{1}{\int_0^T X_t^2 dt + \tau^{-2}}\right).$$

The marginal beliefs

$$P^T(X_0^T|H_i) = \int L_T(\theta|H_i)p(\theta|H_i)d\theta, \quad i = 0,1,2$$

can be computed as follows:

$$P^T(X_0^T|H_i) = \frac{\exp(C_i)}{\sqrt{2\pi}\tau_i}\int_{\mathbb{R}} \exp\left[-\frac{1}{2\tau_i^2}\left\{\theta^2\tau_i^2 A_i - 2\tau_i^2 B_i + (\theta-\gamma_i)^2\right\}\right] d\theta$$

using the Gisanov likelihood, the normality of the prior and the definitions of $(A_i, B_i, C_i)$. When $\theta \sim N(0, \tau_i^2)$, i.e., $\gamma_i = 0$,

$$2\log P^T(X_0^T|H_i) = \frac{\tau_i^2 \widetilde{B}_i^2}{1+\tau_i^2\widetilde{A}_i} - \log(1+\tau_i^2\widetilde{A}_i)$$

where $\widetilde{A}_i = A_i$ and $\widetilde{B}_i = \int_0^T X_t dX_t = \frac{1}{2}(X_T^2 - T)$, due to Itô formula. We use quadrature based approximation to estimate $A_i$ (see Bishwal (2006)), in order to obtain approximate Bayes factor. The Bayes factor can be computed, by noting that under $H_0$, the dominating measure is the Wiener measure. Hence for the R-B model

$$2\log BF(X_0^T) = \frac{(X_T - X_0)^2 + 2\gamma\tau^{-2}(X_T - X_0) - \gamma^2\tau^{-2}T}{(T+\tau^{-2})} - \log(1+\tau^2 T).$$

The posterior odds of $H_i$ versus $H_j$ is given by

$$\frac{P(H_i|X_0^T)}{P(H_j|X_0^T)} = (1+\tau^2 T)^{-1/2} \exp\left\{\frac{(X_T - X_0)^2 + 2\gamma\tau^{-2}(X_T - X_0) - \gamma^2\tau^{-2}T}{2(T+\tau^{-2})}\right\} \frac{P(H_i)}{P(H_j)}$$

where $\frac{P(H_i)}{P(H_j)}$ is a priori odds ratio.

For the B-K model, under a normal prior $\theta \sim N(\gamma,\ \tau^2)$, the marginal beliefs are given by

$$2\log P^T(X_0^T|H_2) = \frac{\tau^2 B_2^2}{1+\tau^2 A_2} - \log(1+\tau^2 A_2)$$

where

$$\sigma^2 A_2 = T\left(1 + \alpha T + \frac{\alpha^2 T^2}{3}\right),$$

$$\sigma^2 B_2 = X_T - X_0 + \alpha \int_0^T t dX_t + \alpha \int_0^T (1 + \alpha t X_t dt),$$

$$\sigma^2 C_2 = \frac{-\alpha}{2}(X_T^2 - X_0^2 - \sigma^2 T) - \frac{\alpha^2}{2}\int_0^T X_t^2 dt.$$

If the Bayes factor $BF_{21}$ is greater than one, then there is evidence in the data against the model $H_1$ in favor of model $H_2$. One can also access the probability of mean reversion. Suppose that we specify a priori odds $O = P(H_2)/P(H_0)$ of mean reversion versus no mean reversion. Typically $O = 1$. After the data has been collected, these odds can be updated by the Bayes factor to posterior odds. The probability of mean reversion is given by $P(H_2|Data) = BF/O + BF$.

*Consistency:* Under the true probability distribution, consistency property of Bayes factor refers to stochastic convergence of $BF_{12}$, under the true probability distribution, $BF_{21} \to \infty$ if $H_2$ is the best model, $BF_{12} \to 0$ if $H_2$ is the best model. The following proposition shows consistency of Bayes factor.

*Coherency:* If $H_i, H_j$ and $H_k$ are three models under consideration, models are said to be coherent if $B_{ij} = 1/B_{ji}$ and $B_{ij}/B_{kj} = B_{ik}$.

**Proposition 2.1** Suppose that under $H_0$, $H_1$, and $H_2$, $\theta \sim \mathcal{N}(\gamma, \tau^2)$ and under $H_1$, and $H_2$ $\alpha \sim$ Exponential $(\beta)$ with known hyperparameters $\gamma, \tau, \beta$. Then
(a) $BF_{10} \to \infty$ a.s. as $T \to \infty$.
(b) $BF_{20} \to \infty$ a.s. as $T \to \infty$.
(c) $BF_{21} \to \infty$ a.s. as $T \to \infty$.

## 3. Discrete Observations

In practice, observations are always discrete, though the model is continuous. Let the process $\{X_t\}$ be observed at $0 = t_0 < t_1 < t_2 < \ldots < t_n$. Denote $\Delta := t_i - t_{i-1}, 1 \leq i \leq n$. Note that $\Delta$ could be random having some prior distribution, see Bishwal (2010). However, we will not consider random sampling in this paper. In the following, DBF denotes Bayes factor based on discrete observations.

We will consider two types of equally spaced deterministic sampling : $\Delta$ fixed (low frequency sampling) and $\Delta \to 0$ as $n \to \infty$ (high frequency sampling).

### 3.1. Fixed-$\Delta$ case

In the stationary case, the O-U process is a nonlinearly parametrized autoregressive process (AR(1) model) having the representation

$$X_{t_i} = \exp(\theta\Delta)X_{t_{i-1}} + \epsilon_{t_{i-1}}, i \geq 1$$

where $\epsilon_{t_{i-1}} = \int_{t_{i-1}}^{t_i} e^{(t_i - s)} dW_s,\ i \geq 1$, which is $\mathcal{N}(0, \psi(\theta))$ where $\psi(\theta) := \frac{e^{2\theta\Delta}-1}{2\theta}$, see Bishwal (2011b).

Thus the process has a Gaussian density and the likelihood is given by

$$L_n(\theta) = -\frac{n}{2}\log 2\pi\psi(\theta) - \frac{1}{2}\sum_{i=1}^{n} \frac{\left(X_{t_i} - \exp(\theta\Delta)X_{t_{i-1}}\right)^2}{\psi(\theta)}.$$

The following proposition gives the consistency of the discrete Bayes factors.

**Proposition 3.1** Suppose that under $H_0$, $H_1$, and $H_2$, $\theta \sim \mathcal{N}(\gamma, \tau^2)$ and under $H_1$, and $H_2$ $\alpha \sim$ Exponential $(\beta)$ with known hyperparameters. Then with the likelihood $L_n(\theta)$, we have
(a) $DBF_{10} - BF_{10} = O_P(n^{-1/2})$.
(b) $DBF_{20} - BF_{20} = O_P(n^{-1/2})$.
(c) $DBF_{21} - BF_{21} = O_P(n^{-1/2})$.

**Proposition 3.2** Suppose that under $H_0$, $H_1$, and $H_2$ $\theta \sim \mathcal{N}(\gamma, \tau^2)$, under $H_1$, and $H_2$ $\alpha \sim$ Exponential $(\beta)$ with known hyperparameters. Then with the likelihood $L_n(\theta)$, we have
(a) $DBF_{10} \to \infty$ a.s. as $n \to \infty$.
(b) $DBF_{20} \to \infty$ a.s. as $n \to \infty$.
(c) $DBF_{21} \to \infty$ a.s. as $n \to \infty$.

### 3.2. Small-$\Delta$ case

Here we assume $\Delta = \Delta_n = \frac{T}{n}$. Also $\frac{T}{n} \to 0$ as $T \to \infty$ and $n \to \infty$. For the G-O-U model, taking an Euler type approximation of the continuous Girsanov likelihood, we obtain an approximate likelihood,

$$L_{n,T,1}(\theta) = \exp\left(\sum_{i=1}^{n} X_{t_{i-1}}(X_{t_i} - X_{t_{i-1}}) - \frac{1}{2}X_{t_{i-1}}^2(t_i - t_{i-1})\right).$$

Using Itô formula and applying the rectangular approximation in continuous Girsanov likelihood which is equivalent to Stratonovich approximation), we obtain another approximate likelihood:

$$L_{n,T,2}(\theta) = \exp\left(\frac{1}{2}(X_T^2 - X_0^2 - T) - \frac{1}{2}X_{t_{i-1}}^2(t_i - t_{i-1})\right).$$

Based on contrast function approach, we consider the contrast

$$L_{n,T,3}(\theta) = \exp\left(-\frac{T}{2} - \frac{1}{2}X_{t_{i-1}}^2(t_i - t_{i-1})\right).$$

For the B-K nonhomegeneous mean reversion model, the first order Euler type approximations are:

$$\int_0^T (1 + \alpha t)dX_t \approx \sum_{i=1}^{n}(1 + \alpha t_{i-1})(X_{t_i} - X_{t_{i-1}}),$$

$$\int_0^T X_t^2 dt \approx \sum_{i=1}^{n} X_{t_{i-1}}^2(t_i - t_{i-1}).$$

The second order Stratonovich type approximations (see Bishwal (2021b)) are:

$$\int_0^T (1 + \alpha t)dX_t \approx \sum_{i=1}^{n}(1 + \frac{\alpha}{2}(t_{i-1} + t_i))(X_{t_i} - X_{t_{i-1}}),$$

$$\int_0^T X_t^2 dt \approx \frac{1}{2}\sum_{i=1}^{n}(X_{t_i}^2 + X_{t_{i-1}}^2)(t_i - t_{i-1}).$$

We first obtain the rate of convergence of approximate Bayes factor (ABF) to the continuous time Bayes factors for fixed $T$. This measures the loss of information due to a particular discretization. We will show that the rates of consistency of the ABF using $L_{n,T,2}$ and $L_{n,T,3}$ are faster than the rate using $L_{n,T,1}$. Let ABF$k$, $k = 1, 2, 3$ denote approximate Bayes factor based on the approximate likelihoods $L_{n,T,k}$, $k = 1, 2, 3$ respectively. The proofs depend on the calculations in Bishwal (2021a). We omit the details.

**Proposition 3.3** Suppose that under $H_0$, $H_1$, and $H_2$ $\theta \sim \mathcal{N}(\gamma, \tau^2)$ and under $H_1$, and $H_2$ $\alpha \sim$ Exponential $(\beta)$ with known hyperparameters. Then with the

approximate likelihood $L_{n,T,1}$, we have
(a) $ABF1_{10} - BF_{10} = O_P(\Delta^{1/2})$.
(b) $ABF1_{20} - BF_{20} = O_P(\Delta^{1/2})$.
(c) $ABF1_{21} - BF_{21} = O_P(\Delta^{1/2})$.

**Proposition 3.4** Suppose that under $H_0$, $H_1$, and $H_2$, $\theta \sim \mathcal{N}(\gamma, \tau^2)$ and under $H_1$, and $H_2$ $\alpha \sim$ Exponential $(\beta)$ with known hyperparameters. Then with the approximate likelihood $L_{n,T,1}$, we have
(a) $ABF1_{10} \to \infty$ a.s. as $T \to \infty$ and $\Delta \to 0$.
(b) $ABF1_{20} \to \infty$ a.s. as $T \to \infty$ and $\Delta \to 0$.
(c) $ABF1_{21} \to \infty$ a.s. as $T \to \infty$ and $\Delta \to 0$.

**Proposition 3.5** Suppose that under $H_0$, $H_1$, and $H_2$, $\theta \sim \mathcal{N}(\gamma, \tau^2)$ and under $H_1$, and $H_2$ $\alpha \sim$ Exponential $(\beta)$ with known hyperparameters. Then with the approximate likelihood $L_{n,T,2}$, we have
(a) $ABF2_{10} - BF_{10} = O_P(\Delta)$.
(b) $ABF2_{20} - BF_{20} = O_P(\Delta)$.
(b) $ABF2_{21} - BF_{21} = O_P(\Delta)$.

This shows that ABF2 converges to BF faster than that ABF1 converges to BF.

**Proposition 3.6** Suppose that under $H_0$, $H_1$, and $H_2$, $\theta \sim \mathcal{N}(\gamma, \tau^2)$ and under $H_1$, and $H_2$ $\alpha \sim$ Exponential $(\beta)$ with known hyperparameters. Then with the approximate likelihood $L_{n,T,2}$, we have
(a) $ABF2_{10} \to \infty$ a.s. as $T \to \infty$ and $\Delta \to 0$.
(b) $ABF2_{20} \to \infty$ a.s. as $T \to \infty$ and $\Delta \to 0$.
(c) $ABF2_{21} \to \infty$ a.s. as $T \to \infty$ and $\Delta \to 0$.

**Proposition 3.7** Suppose that under $H_0$, $H_1$, and $H_2$, $\theta \sim \mathcal{N}(\gamma, \tau^2)$ and under $H_1$, and $H_2$ $\alpha \sim$ Exponential $(\beta)$ with known hyperparameters. Then with the approximate likelihood $L_{n,T,3}$, we have
(a) $ABF3_{10} - BF_{10} = O_P(\Delta)$.
(b) $ABF3_{20} - BF_{20} = O_P(\Delta)$.
(b) $ABF3_{21} - BF_{21} = O_P(\Delta)$.

This shows that ABF3 converges to BF faster than that ABF1 converges to BF. Another advantage is that the estimators based on $L_{n,T,3}$ are asymptotically efficient while ordinary least squares estimator is inefficient, see Bishwal (2008) and Tanaka (2013).

**Proposition 3.8** Suppose that under $H_0$, $H_1$, and $H_2$, $\theta \sim \mathcal{N}(\gamma, \tau^2)$ and under $H_1$, and $H_2$ $\alpha \sim$ Exponential $(\beta)$ with known hyperparameters. Then with the approximate likelihood $L_{n,T,3}$, we have

(a) $ABF3_{10} \to \infty$ a.s. as $T \to \infty$ and $\Delta \to 0$.
(b) $ABF3_{20} \to \infty$ a.s. as $T \to \infty$ and $\Delta \to 0$.
(c) $ABF3_{21} \to \infty$ a.s. as $T \to \infty$ and $\Delta \to 0$.

## 4. Ornstein-Uhlenbeck Process with Mean Reversion

First consider the model hypotheses:

$$H_0 : dX_t = \theta X_t dt + \sigma dW_t,\ \theta = 0,$$

$$H_1 : dX_t = \theta X_t dt + \sigma dW_t,\ \theta < 0,$$

$$H_2 : dX_t = \theta X_t dt + \sigma dW_t,\ \theta > 0.$$

Note that under $H_0$, the model satisfies LAN condition, under $H_1$, the model satisfies LAMN condition and under $H_2$, the model satisfies LABF condition. The model satisfies LAQ for all $\theta$. This is in contrast to the Black-Scholes model, where the model is LAN for all hypotheses. This provides the Ornstein-Uhlenbeck model more flexibility for modelling in finance.

Consider the mean reverting Ornstein-Uhlenbeck process, also known as Vasicek model in finance,

$$dX_t = (\mu + \theta X_t)dt + \sigma dW_t, \quad X_0 = \xi$$

Here $\theta$ is the mean reversion speed and $\mu$ is the mean reversion level and $\sigma > 0$ is the volatility.

First one can consider proper prior and obtain consistency or inconsistency among the LAN, LAMN and LABF models.

Our model hypotheses are:

$$M_0 : dX_t = (\mu + \theta X_t)dt + \sigma dW_t,\ \mu > 0,\ \theta < 0,$$

$$M_1 : dX_t = (\mu + \theta X_t)dt + \sigma dW_t,\ \mu < 0,\ \theta < 0,$$

$$M_2 : dX_t = (\mu + \theta X_t)dt + \sigma dW_t,\ \mu = 0,\ \theta < 0,$$

$$M_3 : dX_t = (\mu + \theta X_t)dt + \sigma dW_t,\ \mu > 0,\ \theta > 0,$$

$$M_4 : dX_t = (\mu + \theta X_t)dt + \sigma dW_t,\ \mu < 0,\ \theta > 0,$$

$$M_5 : dX_t = (\mu + \theta X_t)dt + \sigma dW_t,\ \mu = 0,\ \theta > 0,$$

$$M_6 : dX_t = (\mu + \theta X_t)dt + \sigma dW_t,\ \mu > 0,\ \theta = 0,$$

$$M_7 : dX_t = (\mu + \theta X_t)dt + \sigma dW_t,\ \mu < 0,\ \theta = 0,$$

$$M_8 : dX_t = (\mu + \theta X_t)dt + \sigma dW_t,\ \mu = 0,\ \theta = 0.$$

The model is also known as Vasicek model in finance literature where $\mu$ is the level of mean reversion and $\theta$ is the speed of mean reversion. The model is used for modeling interest rate and bond pricing. One can obtain BF and ABF for these models similar to the previous sections.

## 5. Intrinsic Bayes Factors

The intrinsic Bayes factor (IBF) between competing models $H_i$ and $H_j$ is defined as

$$IBF_{ij} := BF_{ij} \cdot CF_{ji}, \quad i, j = 1, 2$$

where $CF_{ji}$ is a correction factor, see Berger and Pericchi (1996).

IBF are of two types: Arithmetic IBF (AIBF) and Geometric IBF (GIBF) which are given by

$$AIBF_{ij} := BF_{ij}^N \cdot \frac{1}{K} \sum_{k=1}^{K} BF_{ji}^N (X(t_k)), \quad i, j = 1, 2,$$

$$GIBF_{ij} := BF_{ij}^N \cdot \left( \prod_{k=1}^{K} BF_{ji}^N (X(t_k)) \right), \quad i, j = 1, 2$$

where $X(t_k), k = 1, 2. \ldots, K$ is the minimal training sample. Obviously, due to the inequality between geometric mean and arithmetic mean,

$$GIBF_{ij} \leq AIBF_{ij}.$$

IBF can be calculated for B-K model.

## 6. Fractional Bayes Factors

The fractional Bayes factor (FBF) is defined as

$$FBF_{ij} := \frac{m_i}{m_j} \cdot \frac{m_j^{(\frac{1}{n})}}{m_i^{(\frac{1}{n})}}$$

where

$$m_i^{(r)} = \int L_i^{(r)}(\theta) p_i(\theta) d\theta, \quad 0 \leq r \leq 1,$$

and $m_i = m_i^{(1)}$ is the marginal density using the full likelihood $L_i(\theta)$. The fractional Bayes factor uses a $\frac{1}{n}$-th fraction of the full likelihood. FBF can be calculated for B-K model. Using the three approximations of the full likelihood, one can calculate approximate fractional Bayes factor (AFBF).

## 7. Concluding Remarks

We studied consistency and discretization error rate of Bayes factor for Black-Karasinski model which help model selection. This approach to model selection has not yet been paid much attention in the financial literature earlier. This opens up the gateway to model selection for other financial models. It would be interesting to calculate approximate fractional Bayes factor (AFBF) for fractional Black-Karasinski model of Bishwal (2022b).

## References

[1] Berger, J.O. (1985). Statistical Decision Theory and Bayesian Analysis, Springer-Verlag, Berlin.

[2] Berger, J.O. and Pericchi, L.R. (1996). The intrinsic Bayes factor for model selection and prediction, Journal of the American Statistical Association, 91(433), 109-122.

[3] Berger, J.O. and Sellke, T. (1987). Testing a point null hypothesis: The irreconcilability of $P$-values and evidence (with discussion), Journal of the American Statistical Association, 82, 112-139.

[4] Bishwal, J.P.N. (2000). Rates of convergence of the posterior distributions and the Bayes estimators in the Ornstein-Uhlenbeck process, Random Operators and Stochastic Equations, 8(1), 51-70.

[5] Bishwal, J.P.N. (2001). Accuracy of normal approximation for the maximum likelihood and the Bayes estimators in the Ornstein-Uhlenbeck process using random norming, Statistics and Probability Letters, 52(4), 427-439.

[6] Bishwal, J.P.N. (2006). Rates of weak convergence of the approximate minimum contrast estimators for the discretely observed Ornstein-Uhlenbeck process, Statistics and Probability Letters, 76(13), 1397-1409.

[7] Bishwal, J.P.N. (2008). Parameter Estimation in Stochastic Differential Equations, Lecture Notes in Mathematics 1923, Springer-Verlag, Berlin.

[8] Bishwal, J.P.N. (2010). Conditional least squares estimation in diffusion processes based on Poisson sampling, Journal of Applied Probability and Statistics, 5(2), 169-180.

[9] Bishwal, J.P.N. (2011a). Milstein approximation of posterior density of diffusions, International Journal of Pure and Applied Mathematics, 68(4), 403-414.

[10] Bishwal, J.P.N. (2011b). Sufficiency and Rao-Blackwellization of Vasicek model, Theory of Stochastic Processes, 17(1), 12-15.

[11] Bishwal, J.P.N. (2018). Sequential maximum likelihood estimation in nonlinear non-Markov diffusion type processes, Dynamic Systemns and Applications, 27(1), 107-124.

[12] Bishwal, J.P.N. (2021a). Berry-Esseen bounds of approximate Bayes estimators for the discretely observed Ornstein-Uhlenbeck process, Asian Journal of Statistical Sciences, 1(2), 83-122.

[13] Bishwal, J.P.N. (2021b). A new algorithm for approximate maximum likelihood estimation in sub-fractional Chan-Karolyi-Longstaff-Sanders model, Asian Journal of Probability and Statistics, 13(3), 62-88.

[14] Bishwal, J.P.N. (2022a). Parameter Estimation in Stochastic Volatility Models, Springer Nature, Cham.

[15] Bishwal, J.P.N. (2022b). Berry-Esseen inequalities for the fractional Black-Karasinski model of term structure of interest rates, *Monte Carlo Methods and Applications*, 28(2), 111-124.

[16] Black, F. and Karasinski, P. (1991). Bond and option pricing when short rates are lognormal, Financial Analysts Journal, **47**(4), 52-59.

[17] Hull, J.C. (2015). Options, Futures and Other Derivatives, Ninth Edition, Prentice Hall.

[18] Jeffreys, H. (1961). Theory of Probability, 3rd ed., Oxford University Press.

[19] Liptser, R.S. and Shiryayev, A.N. (1978). Statistics of Random Processes II : Applications Springer-Verlag, Berlin.

[20] Polson, N.G. and Roberts, G.O. (1994). Bayes factor for discrete observations from diffusion processes, Biometrika, 81, 11-26.

[21] Rendleman, R. and Bartter, B. (1980). The pricing of options on debt securities, Journal of Financial and Quantitative Analysis, 15, 11-24.

[22] Simons, G., Yao, Y.C. and Wu, X. (1989). Sequential tests for the drift of a Wiener process with smooth prior and the heat equation, Annals of Statistics, 17(2), 783-792.

[23] Tanaka, K. (2013). Distributions of the maximum likelihood and minimum contrast estimators associated with the fractional Ornstein-Uhlenbeck process, Statistical Inference for Stochastic Processes, 36, 173-192.

[24] Turfus, C. (2019). Exact Arrow-Debrew pricing for the Black-Karasinski short rate model, SSRN-ID3253839.