A Performance Comparison of Various Bootstrap Methods for Diffusion Processes

Jung S. You[1] & Minsoo Jeong[2]
[1]Department of Economics, California State University-East Bay, USA
[2]Department of Economics, Yonsei University, Wonju, Republic of Korea
jung.you@csueastbay.edu, mssjong@yonsei.ac.kr

Abstract: In this paper, we compare the finite sample performances of various bootstrap methods for diffusion processes. Though diffusion processes are widely used to analyze stocks, bonds, and many other financial derivatives, they are known to heavily suffer from size distortions of hypothesis tests. While there are many bootstrap methods applicable to diffusion models to reduce such size distortions, their finite sample performances are yet to be investigated. We perform a Monte Carlo simulation comparing the finite sample properties, and our results show that the strong Taylor approximation method produces the best performance, followed by the Hermite expansion method.

Keywords: *Bootstrap; diffusion process; strong Taylor approximation; stationarity; Monte Carlo simulation.*

1. Introduction

Markov processes play a central role in financial analyses, granted that prices efficiently reflect information in the market. Diffusion processes are continuous-time Markov processes that have a semi-martingale property, and they have been widely used to model stocks, bonds, and many other financial derivatives (Chan et al., 1992). However, diffusion models are known to heavily suffer from size distortions of hypothesis tests. While there are many bootstrap methods applicable to diffusion models to reduce such size distortions, their finite sample performances are yet to be investigated. Thus, we compare their performance by using Monte Carlo simulations.

It is well known that bootstraps enhance finite sample performance in various statistical analyses, such as parameter value estimations and hypothesis testing. For example, the bootstrap can be used to correct biases in estimation as seen in Tang & Chen (2009). However, bootstrapping the diffusion processes is not straightforward because of the dependence structure residing in the time series data (Kreiss & Lahiri, 2012; Horowitz, 2003; Härdle et al., 2003; Pan & Politis, 2016a; Pan & Politis, 2016b). There are various methods for circumventing this problem. For example, when the parametric dependence structure of the diffusion model is well known and easy to construct, we can generate the bootstrap samples using the parametric bootstraps. Conversely, if the parametric structure of the model is unknown or difficult to construct, we may need to utilize nonparametric approaches. In this study, we consider various parametric and nonparametric bootstrap resampling methods in practice.

Among the parametric approaches, the strong Taylor approximation is straightforward and easy to implement, so it is preferred by many researchers. Having been used for a long time, it is still one of the most widely implemented methods, and many extensions are still proposed, as seen in Gyöngy & Rásonyi (2011) and Mikulevičius & Zhang (2015). A relatively new approach utilizes the Hermite expansion suggested by Aït-Sahalia (1999). Due to its sound theoretical background and good performance in estimation, parameter estimation using the Hermite expansion has gained popularity since its introduction. In this study, we utilize the Hermite expansion method to generate bootstrap samples.

Nonparametric approaches utilize a nonparametrically estimated conditional distribution. The conditional distribution contains essential characteristics of the diffusion process; therefore, it plays an important role in various diffusion analyses, as seen in Chen et al. (2008) and Bhardwaj et al. (2008). Due to its versatility and robustness to misspecification, the nonparametric approach is also preferred by many researchers. As for the bootstrap resampling, Horowitz (2003) suggested using the Markov conditional bootstrap (MCB), in which random samples are generated from the non-parametrically estimated conditional distributions. These parametric and nonparametric approaches have their own strengths and weaknesses, and their finite sample performances in hypothesis testing are yet to be investigated.

This paper compares the accuracy of the coverage probabilities of various bootstrap methods by utilizing Monte Carlo experiments. Overall, we find that the strong Taylor approximation method performs the best, followed by the Hermite expansion method. The nonparametric MCB shows relatively larger size distortions, but it still performs better than the first-order asymptotics. We conjecture that this is because the nonparametric methods involve many user-dependent factors, such as the choice of kernel function and bandwidth. However, as it clearly

outperforms the first-order asymptotics, it is worthwhile to consider the MCB when the nonparametric method is inevitable.

In Section 2, we introduce various bootstrap resampling methods that generate samples either from the parametrically constructed or from nonparametrically constructed transition densities. Section 3 presents the results of a Monte Carlo simulation on the performance of the bootstrap methods. Section 4 presents concluding remarks.

2. Various Bootstrap Methods for Diffusion Models

In this section, we introduce various bootstrap resampling methods that can be applied to diffusion models. For the analysis, we consider a stationary diffusion process X whose stochastic differential equation (SDE) is given as the following:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,$$

where W is the standard Brownian motion.

Since closed-form expressions of exact transition densities are not generally available, we first consider the strong Taylor approximation. The Euler and Milstein schemes are popular strong Taylor approximation methods. In addition to these, we also consider other more sophisticated methods, such as the Heun method and the order 1.5 strong Taylor scheme. Furthermore, we also consider an approximation utilizing the Hermite expansion of the transition density, as suggested by Aït-Sahalia (1999). Although it is one of the most popular methods used for parameter estimation, it is not very common for random sample generation. Lastly, we consider the MCB, a nonparametric bootstrap method proposed by Horowitz (2003). For the financial analysis using diffusion models, it is worthwhile to consider using nonparametric methods, since parametric diffusion models are often imprecisely specified for simplification.

Exact Transition Density: In general, closed-form expressions of exact transition densities are not available except for very few diffusion processes. For stationary diffusions, only the following two models are known to have closed-form transition densities.

- Vasicek model (Ornstein-Uhlenbeck process), attributed to Vasicek (1977). The SDE of the Vasicek model is given as

$$dX_t = \kappa(\alpha - X_t)dt + \sigma dW_t$$

for $\kappa, \sigma \in R_+$ and $\alpha \in R$. The conditional distribution of $X_\Delta | X_0 = x$ is the normal distribution, whose density function is

$$p(y|x) = \varphi(y; m(\Delta, x), v(\Delta, x))$$

where φ is the density function of the normal distribution, whose conditional mean and variance are, respectively,

$$m(t, x) = E(X_t|X_0 = x) = e^{\kappa t}x + \alpha(1 - e^{-\kappa t}), \text{ and}$$
$$v(t, x) = var(X_t|X_0 = x) = \frac{\sigma^2(1 - e^{-2\kappa t})}{2\kappa}.$$

- The Cox-Ingersoll-Ross model (Feller's square root process), attributed to Cox et al. (1985). The SDE is given as

$$dX_t = \kappa(\alpha - X_t)dt + \sigma\sqrt{X_t}dWt$$

for $\kappa, \alpha, \sigma \in R_+$ such that $2\kappa\alpha > \sigma^2$. In this case $X_\Delta | X_0 = x$ follows a noncentral χ^2- distribution. The transition density is $p(y|x) = c \exp(-u - v)\left(\frac{v}{u}\right)^{q/2} I_q(2\sqrt{uv})$ for $x, y \in R_+$ where *c=2κ/(σ2 (1-exp(-κΔ))), q=2κα/σ2 – 1, u=cx exp(-κΔ), v=cy,* and I_q is the modified Bessel function of order q.

Strong Taylor Approximation: The use of a strong Taylor approximation can be considered for bootstrap sampling. Once we obtain the parameter estimates of the diffusion processes, we can generate a stochastic process from the strong Taylor approximation. In order to define the strong Taylor approximation, we first introduce the absolute error criterion $E(|X_T - Y(T)|)$, where Y is an approximation of X. Next, Y^δ, a discrete-time approximated version of X on the time interval $(\tau)_\delta = \{\tau_n: n = 0, 1, \cdots\}$, is considered strongly converging with the order of γ, if there exist C and δ_0 such that $E(|X_T - Y^\delta(T)|) \leq C\delta^\gamma$ for any $\delta \in (0, \delta_0)$. This definition can be naturally understood as a generalization of the deterministic version convergence criterion. The higher value of γ implies a sharper order of convergence. For more details, see Platen (1999). For more convergence results of the strong Taylor

approximation, a discrete-time approximation Y^δ is strongly consistent if, for some nonnegative function $c(\delta)$ with $\lim_{\delta \downarrow 0} c(\delta) = 0$, we have

$$E(|E(\frac{Y^\delta_{\tau_{n+1}} - Y^\delta_{\tau_n}}{\Delta_n}|A_{\tau_n}) - \mu(Y^\delta_{\tau_n})|^2) \le c(\delta) \text{ and } E(\frac{1}{\Delta_n}|Y^\delta_{\tau_{n+1}} - Y^\delta_{\tau_n} - E(Y^\delta_{\tau_{n+1}} - Y^\delta_{\tau_n}|A_{\tau_n}) - \sigma(Y^\delta_{\tau_n})\Delta W_n|^2) \le c(\delta)$$

for all n=0,1, ..., where $\{A_t, t \ge 0\}$ is a family of σ-algebras. The above two conditions roughly imply that Y^δ converges to X in terms of the mean and variance. Thus, if an approximation is strongly consistent, then the two processes are pathwisely close to each other. Under some regularity conditions such as the Lipschitz condition and the linear growth bound condition, we formulate the following theorem.

Theorem 1 (Kloeden & Platen, 1992). A strongly consistent equidistant time discrete approximation Y^δ of a one-dimensional autonomous Ito process X with $Y^\delta(0) = X_0$ converges strongly to X.

When we derive a discrete-time approximation in the strong convergence criterion, we refer to it as a strong Taylor approximation. As shown in the following, the convergence order is determined by how many terms we include in the expansion.

The Euler–Maruyama scheme, also called the Euler scheme, is the simplest strong Taylor approximation, and in general, it attains the order of strong convergence $\gamma = 0.5$. The Euler scheme is given by

$$Y_{\tau_{n+1}} = Y_{\tau_n} + (Y_{\tau_n})\Delta + \sigma(Y_{\tau_n})\Delta W$$

where Δ is the length of the interval $[\tau_n, \tau_{n+1}]$ and $\Delta W = W_{\tau_{n+1}} - W_{\tau_n}$ is the increment of the standard Brownian motion W, on $[\tau_n, \tau_{n+1}]$.

For the Euler scheme, $Y_{\tau_{n+1}} = Y_{\tau_n} + \mu(Y_{\tau_n})\Delta + \sigma(Y_{\tau_n})\Delta W$, we evaluate the right-hand side of the equation at the beginning of each interval $\tau_n < t < \tau_{n+1}$. We can obtain a more accurate approximation when we include more information of the process from elsewhere, for example, when we use the average of the values at both τ_n and τ_{n+1}. In this case, we have

$$Y_{\tau_{n+1}} = Y_{\tau_n} + \frac{1}{2}\{\mu(Y_{\tau_n}) + \mu(Y_{\tau_{n+1}})\}\Delta + \sigma(Y_{\tau_n})\Delta W$$

This method is not feasible because the unknown quantity $Y_{\tau_{n+1}}$ appears on both. To address this issue, we use the Euler scheme to replace the $Y_{\tau_{n+1}}$ term on the right-hand side. Accordingly, we obtain that

$$\overline{Y}_{\tau_{n+1}} = Y_{\tau_n} + \mu(Y_{\tau_n})\Delta,$$
$$Y_{\tau_{n+1}} = Y_{\tau_n} + \frac{1}{2}\{\mu(Y_{\tau_n}) + \mu(\overline{Y}_{\tau_{n+1}})\}\Delta + \sigma(Y_{\tau_n})\Delta W,$$

or

$$Y_{\tau_{n+1}} = Y_{\tau_n} + \frac{1}{2}\{\mu(Y_{\tau_n}) + \mu(Y_{\tau_n} + \mu(Y_{\tau_n})\Delta)\}\Delta + \sigma(Y_{\tau_n})\Delta W.$$

This approximation is called the improved Euler scheme or the Heun method.

The Milstein scheme is an order 1.0 strong Taylor approximation method. To obtain the Milstein scheme

$$Y_{\tau_{n+1}} = Y_{\tau_n} + \mu\Delta + \sigma\Delta W + \frac{1}{2}\sigma\sigma'\{(\Delta W)^2 - \Delta\}$$

where, $\mu = \mu(Y_{\tau_n})$, $\sigma = \sigma(Y_{\tau_n})$, and $\sigma' = \sigma'(Y_{\tau_n})$. We add the term $\frac{1}{2}\sigma\sigma'\{(\Delta W)^2 - \Delta\}$ to the Euler scheme from the Ito-Taylor expansion.

An even more accurate approximation can be obtained by including more stochastic Taylor expansion terms. These additional terms consist of stochastic integrals that carry more information about the process. These additional stochastic integral terms play an important role in improving the accuracy of the approximation since they represent the difference between the stochastic differential equations and the deterministic differential equations. To obtain a strong Taylor scheme of order $\gamma = 1.5$, we add more terms to the Milstein scheme utilizing the Ito lemma. Kloeden & Platen (1992) obtain the order 1.5 strong Taylor scheme as follows:

$$Y_{\tau_{n+1}} = Y_{\tau_n} + \mu\Delta + \sigma\Delta W + \frac{1}{2}\sigma\sigma'\{(\Delta W)^2 - \Delta\} + \mu'\sigma\Delta Z + \frac{1}{2}\{\mu\mu' + \frac{1}{2}\sigma^2\mu''\}\Delta^2 + (\mu\sigma' + \frac{1}{2}\sigma^2\sigma'')\{\Delta W\Delta - \Delta Z\} + \frac{1}{2}\sigma(\sigma\sigma'' + (\sigma')^2)\{\frac{1}{3}(\Delta W)^2 - \Delta\}\Delta W,$$

where, $\mu = \mu(Y_{\tau_n})$, $\mu' = \mu'(Y_{\tau_n})$, $\mu'' = \mu''(Y_{\tau_n})$, $\sigma = \sigma(Y_{\tau_n})$, $\sigma' = \sigma'(Y_{\tau_n})$, $\sigma'' = \sigma''(Y_{\tau_n})$ and ΔZ is defined as

$$\Delta Z = \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{s_2} dW_{s_1} ds_2.$$

Hermite Expansion: The Hermite expansion method was first introduced by Aït-Sahalia (1999). For this method, we apply the Lamperti transformation on the original diffusion process so that the transition density of the transformed model would become more suitable for the Hermite expansion. In the below brief overview of the transformations, let $D_x(\underline{x}, \overline{x})$ be the domain of the process X. For the approximation, we first transform X into Y such that

$$Y_t \equiv f(X_t; \theta) = \int_{x^\#}^{X_t} \frac{1}{\sigma(u)} du,$$

where $x^\#$ is an arbitrary point in the domain D_X. Note that as $\sigma > 0$, the transformation function f is strictly increasing and invertible. We denote $D_Y = (\underline{y}, \overline{y})$ as the domain of Y where $\underline{y} \equiv \lim_{x \to \underline{x}} f(x)$ and $\overline{y} \equiv \lim_{x \to \overline{x}} f(x)$. Then from Ito's lemma, we obtain $dY_t = \mu_Y(Y_t)dt + dW_t$ where

$$\mu_Y(y) = \frac{\mu(f^{-1}(y))}{\sigma(f^{-1}(y))} - \frac{1}{2}\frac{\partial \sigma}{\partial x}(f^{-1}(y)).$$

Next, we transform Y into Z such that $Z_t = \Delta^{-\frac{1}{2}}(Y_t - y_0)$ where $y_0 = f(x_0)$. For the processes X, Y, and Z, Aït-Sahalia (1999) derived their transition densities as follows:

$$p_Z(z|y_0) = \Delta^{1/2} p_Y(\Delta^{1/2} z + y_0|y_0),$$
$$p_Y(y|y_0) = \Delta^{-1/2} p_Z(\Delta^{-1/2}(y - y_0)|y_0),$$
$$p_x(x|x_0) = \frac{p_Y(f(x)|f(x_0))}{\sigma(x)},$$

and

$$p_Y(y|y_0) = \sigma(f^{-1}(y))p_x(f^{-1}(y)|f^{-1}(y_0)).$$

We approximate the transition density of Z as

$$p_Z^{(J)}(z|y_0) \equiv \phi(z) \sum_{j=0}^{J} \eta_j(\Delta, y_0)H_j(z),$$

where φ is the standard normal density and

$$\eta_j(\Delta, y_0) \equiv \frac{1}{j!} \int_{-\infty}^{\infty} H_j(z) p_Z^{(J)}(z|y_0) \, dz.$$

Here H denotes the classical Hermite polynomials, which are given by

$$H_j(z) := e^{z^2/2} \frac{d^j}{dz^j} e^{-z^2/2}, j \geq 0.$$

Finally, we derive the approximated transition densities of Y and X as

$$p_Y^{(J)}(y|y_0) = \Delta^{-\frac{1}{2}} p_Z^{(J)}(\Delta^{-\frac{1}{2}}(y - y_0)|y_0) \text{ and } p_X^{(J)}(x|x_0) = \frac{p_Y^{(J)}(f(x)|f(x_0))}{\sigma(x)}.$$

Then the following convergence theorem holds under the conditions assumed in Aït-Sahalia (1999).

Theorem 2 (Aït-Sahalia, 1999). There exists $\overline{\Delta} > 0$ such that for every $\Delta \in (0, \overline{\Delta})$ and $(x_0, x) \in D_X^2$, $p(x|x_0) \to_{J \to \infty} p_X(x|x_0)$.

For the practical implementation of Theorem 1, we first compute $p^{(J)}(z|y_0)$ for a given J. To obtain the coefficients $\eta_j(\Delta, y_0)$ for $j = 0, 1, \ldots, J$, we have

$$\eta_j(\Delta, y_0) = \frac{1}{j!} E\left[H_j\left(\Delta^{-\frac{1}{2}}(Y_{t+\Delta} - y_0)\right)\bigg|Y_t = y_0\right].$$

To calculate this expectation, we use the following lemma, which is obtained from the Taylor approximation.

Lemma 1 (Aït-Sahalia, 1999). Let g be a function such that g and all its derivatives have at most exponential growth. Then, for $\Delta \in (0, \overline{\Delta})$, $y_0 \in R$, there exists $\delta \in [0, \Delta]$ such that

$$E[g(Y_{t+\Delta}|Y_t = y_0)] = \sum_{j=0}^{J} (A^J \cdot g(y_0))\frac{\Delta^j}{j!} + E[A^{J+1} \cdot g(Y_{t+\delta})|Y_t = y_0]\frac{\Delta^{J+1}}{(J+1)!},$$

where A is the infinitesimal operator of the diffusion Y defined by $A: g \to \mu_Y(\cdot)\partial g(\cdot)/\partial y + 1/2\, \partial^2 g(\cdot)/\partial y^2$ and $A^j \cdot$

g (y_0) means A applied j times to the function $y \rightarrow g(y)$ and evaluated at $y = y_0$.

In practice, we need to consider how many terms should be included in this Taylor series. Aït-Sahalia (1999) suggested that one should first choose the order J and then expand the expectation such that the approximation of the transition density $p_z^{(J)}(x|x_0)$ is at most of order $\Delta^{J/2}$.

Markov Conditional Bootstrap: Horowitz (2003) proposed the MCB to conduct bootstrap sampling for Markov processes, and he sought to use a transition density to construct the dependency in time series data. We apply this idea by nonparametrically estimating the diffusion transition density. When it is applied to the diffusion model estimation, we expect in general that the nonparametric approaches will perform less accurately than the parametric approaches, since the nonparametric methods involve additional user-dependent factors. However, there are cases in which asymptotic expansion type approximations may not work well enough, so it is still worthwhile to check the performance of the nonparametric method. The MCB estimates the joint and marginal densities as follows:

$$p(x,y) = \frac{1}{(n-1)h_n^2} \sum_{j=2}^{n} K(\frac{x - X_j}{h_n}, \frac{y - Y_j}{h_n})$$

and

$$p(x) = \frac{1}{(n-1)h_n} \sum_{j=2}^{n} K(\frac{x - X_j}{h_n}),$$

where $K(\cdot)$ is the kernel function and h_n is the bandwidth.

3. Monte Carlo Experiments

This section presents the results of the Monte Carlo simulation, which compares the numerical performance of the introduced bootstrap methods. To examine how well the bootstraps perform, we used the Ornstein-Uhlenbeck (OU) process:

$$dX_t = \kappa(\alpha - X_t)dt + \sigma dW_t.$$

We used this process for our simulation as it is one of the most popularly used diffusion models in practice. Furthermore, its transition density function is known in closed form, and its sampling can be carried out without any cumbersome numerical approximation. We use values of $\kappa = 1.0$, $\alpha = 0.6$, and $\sigma = 0.1$ and sample size $n = 300$. We perform 1000 Monte Carlo replications in an experiment. The critical values are obtained from 300 bootstrap iterations, and the coefficients are estimated with the generalized method of moments (GMM). Though applying the exact maximum likelihood estimation is possible for this OU process, we utilize the GMM estimation because the GMM estimation is widely used in practice, whereas the exact maximum likelihood estimation is available only for a few limited diffusion models. In the simulation, we examine the performance of two-tailed t-tests with a significance level of 0.1.

GMM Moment Conditions and the Test Statistic: For the GMM estimation we discretize our diffusion model as follows:

$$X_{t+\Delta} - X_t = \kappa(\alpha - X_t)\Delta + \varepsilon_{t+\Delta},$$

where $E[\varepsilon_{t+\Delta}] = 0$ and $E[\varepsilon_{t+\Delta}^2] = \sigma^2$ following Brennan & Schwartz (1988), Dietrich-Campbell & Schwartz (1986), and Sanders & Unal (2001). We let $\theta = (\alpha, \beta, \sigma^2)$. Given $\varepsilon_{t+\Delta} = X_{t+\Delta} - X_t - \kappa(\alpha - X_t)\Delta$, the vector $f_t(X, \theta)$ is written as

$$f_t(X, \theta) = \begin{pmatrix} \varepsilon_{t+\Delta} \\ \varepsilon_{t+\Delta} X_t \\ \varepsilon_{t+\Delta}^2 - \sigma^2 \Delta \\ (\varepsilon_{t+\Delta}^2 - \sigma^2 \Delta) X_t \end{pmatrix}.$$

Denoting with θ_0 the true but unknown value of θ, we have $E[f_t(X, \theta_0)] = 0$. The GMM procedure estimates parameter values by finding the values that satisfy the sample version of the moment conditions, where $E[f_t(X, \theta)]$ is replaced with $g_T(X, \theta) = \frac{1}{T}\sum_{t=1}^{T} f_t(X, \theta)$. Then the parameter estimates are given by the minimizer of the quadratic form

$$J_T(\theta) = g'_T(X, \theta) W_T(\theta) g_T(X, \theta),$$

where $W_T(\theta)$ is a positive-definite symmetric matrix. From the matrix differentiation, finding the minimum of $J_T(\theta)$ is equivalent to solving, $D'(\theta) W_T(\theta) g_T(X, \theta) = 0$, where $D(\theta)$ is the Jacobian matrix of $g_T(X, \theta)$ with respect to θ.

The choice of the weighing matrix W_T is also important since the performance of the GMM estimator θ_n depends on how we define W_T. Hansen (1982) showed that if we let $W_T(\theta) = S^{-1}(\theta)$ where

$$S(\theta) = E\big[f_t(X,\theta)f'_t(X,\theta)\big],$$

then the resulting GMM estimator achieves the smallest asymptotic variance. Denoting $S_0(\theta)$ as an estimator of $S(\theta)$, the asymptotic variance matrix of the GMM estimator is

$$\sigma = \frac{1}{T}(D'_0(\theta)S_0^{-1}(\theta)D_0(\theta))^{-1},$$

where $D_0(\theta)$ is the Jacobian matrix at the estimated values.

We denote the consistent estimator of σ with σ_n; the (r, r) component of σ_n with $(\sigma_n)_{rr}$; and the r'th components of θ and θ_n with θ_r and θ_{nr}, respectively. Thus, the *t*-statistic for the null hypothesis $H_0: \theta_r = \theta_0$ is $t_{nr} = n^{1/2}(\theta_{nr} - \theta_{0r})/(\varsigma_n)_{rr}^{1/2}$. To obtain the bootstrap versions of t_{nr}, we define $\hat{f}_t(\cdot,\theta) = f_t(\cdot,\theta) - \hat{E}f_t(\cdot,\theta_n)$ using moment conditions such that $\hat{E}\hat{f}_t(\hat{X},\theta) = 0$ where $\hat{E}$ is the bootstrap expectation and $\hat{X}$ is a bootstrap sample. As in Hall & Horowitz (1996) and Andrews (2004), we apply the recentering technique to the bootstrap version since there is no θ such that $\hat{E}f_t(\hat{X},\theta) = 0$ for an overidentified case. The bootstrap estimator of θ, denoted by $\hat{\theta}_n$, is obtained by replacing f_t with $\hat{f}_t$ and X with $\hat{X}$. The bootstrap version of ς_n, denoted by $\hat{\varsigma}_n$, is obtained by replacing f_t with $\hat{f}_t$, X with $\hat{X}$, and θ_n with $\hat{\theta}_n$ in the expression for ς_n. The bootstrap version of the *t*-statistic is $\hat{t}_{nr} = n^{1/2}(\hat{\theta}_{nr} - \theta_{nr})/(\hat{\varsigma}_n)_{rr}^{1/2}$ where $\hat{\theta}_{nr}$ denotes the r'th component of $\hat{\theta}_n$.

Numerical Results: Table 1 reports the results of the Monte Carlo simulation, listing the differences in the coverage probabilities of the bootstrap tests for nominal 90% confidence intervals. For the comparison of the bootstrap performances, we only focus on the coverage probability of the drift term parameters κ and α, which are known to suffer from large size distortions. Since the Euler and Milstein schemes coincide with each other in the case of the OU process, the simulation of Milstein schemes is not conducted. When the transition density is estimated by a nonparametric method or by the Hermite expansion, it is impossible to generate bootstrap samples with explicit formulae. Therefore, sampling is executed with the accept–reject method.[1] To evaluate the accuracy of each bootstrap method, the summation of the absolute values of differences between the nominal and empirical coverage probabilities in each bootstrap test is shown at the bottom of the table.

As seen in Table 1, the bootstraps utilizing the Hermite expansions of orders 1 and 2 outperform the first-order asymptotic test. However, they do not outperform the Euler approximation, which is unexpected. We conjecture that the accept–reject method brought about some inefficiency in the process of bootstrap sampling. Moreover, the bootstrap method with the best performance is the one utilizing the order 1.5 strong Taylor approximation. The higher the order of strong Taylor approximation, the more accurately the empirical coverage probability from bootstrap is calculated. The nonparametric MCB shows relatively larger errors, as it inevitably involves more factors to adjust, but it at least performs better than first-order asymptotic test even for this simple OU model case.

Table 1: Difference between the Actual and Nominal Coverage Probabilities (Actual − Nominal)

	Asym.	Exact	Tayl. 1.5	Heun	Euler	Herm. 1	Herm. 2	MCB
κ	-0.0818	-0.0313	-0.0263	-0.0313	-0.0364	0.0448	0.0400	0.1000
α	-0.0465	-0.0010	0.0040	0.0040	-0.0010	0.0310	-0.0267	-0.0243
Abs. sum	0.1283	0.0323	0.0303	0.0353	0.0374	0.0758	0.0667	0.1243

4. Conclusion

This paper compares the performance of various bootstrap methods that are applicable to diffusion models. Among the various bootstrap methods, those using the idea of a strong Taylor approximation provide the most precise results. The bootstrap using the Hermite expansion fails to perform well possibly because of sampling errors in accept—reject method. The nonparametric MCB shows relatively poor performance, but it still outperforms the first-order asymptotic test. Based on our results, we suggest using the strong Taylor approximation for bootstrapping diffusion processes. In addition, we also suggest the Hermite expansion method when the diffusion

---

[1] The, accept—reject method works as follows. Given a density function f, first we choose a density function g and a constant M such that $f(x) \leq Mg(x)$. Then we generates $X \sim g, U \sim U_{[0,1]}$. Next we accept $Y = X$ if $U \leq f(X)/Mg(X)$, and reject otherwise. Finally, we return to the first step. If we follow this method, it produces a variable Y distributed according to the density function f.

model is more complicated as we only consider a simple diffusion model here. Though the MCB produces rather an unsatisfactory result, we may also consider it when we are not very certain about the model specification, since it surely outperforms the first-order asymptotics. Furthermore, a future study would be needed to reduce the sampling error in accept—reject method, which would enhance the performance of the bootstraps using the Hermite expansion or the nonparametric transition density estimation.

References

Aït-Sahalia, Y. (1999). Transition Densities for Interest Rate and Other Nonlinear Diffusions. *Journal of Finance*, 54(4), 1361–1395.

Andrews, D. W. K. (2004). The Block–Block Bootstrap: Improved Asymptotic Refinements. *Econometrica*, 72(3), 673–700. https://doi.org/10.1111/j.1468-0262.2004.00509.x

Bhardwaj, G., Corradi, V. & Swanson, N. R. (2008). A Simulation-Based Specification Test for Diffusion Processes. *Journal of Business & Economic Statistics*, 26(2), 176–193. https://doi.org/10.1198/073500107000000412

Brennan, M. & Schwartz, E. S. (1988). The Case for Convertibles. https://doi.org/10.1111/J.1745-6622.1988.TB00166.X

Chan, K. C., Karolyi, G. A., Longstaff, F. A. & Sanders, A. B. (1992). An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. *The Journal of Finance*, 47(3), 1209–1227. https://doi.org/10.1111/j.1540-6261.1992.tb04011.x

Chen, S. X., Gao, J. & Tang, C. Y. (2008). A test for model specification of diffusion processes. *The Annals of Statistics*, 36(1), 167–198. https://doi.org/10.1214/009053607000000659

Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2), 385–407. https://doi.org/10.2307/1911242

Dietrich-Campbell, B. & Schwartz, E. (1986). Valuing debt options: Empirical evidence. *Journal of Financial Economics*, 16(3), 321–343. https://doi.org/10.1016/0304-405X(86)90033-4

Gyöngy, I. & Rásonyi, M. (2011). A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients. *Stochastic Processes and Their Applications*, 121(10), 2189–2200. https://doi.org/10.1016/j.spa.2011.06.008

Hall, P. & Horowitz, J. L. (1996). Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moments Estimators. *Econometrica*, 64(4), 891–916. https://doi.org/10.2307/2171849

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50(4), 1029–1054. https://doi.org/10.2307/1912775

Härdle, W., Horowitz, J. & Kreiss, J. P. (2003). Bootstrap Methods for Time Series. *International Statistical Review / Revue Internationale de Statistique*, 71(2), 435–459.

Horowitz, J. L. (2003). Bootstrap Methods for Markov Processes. *Econometrica*, 71(4), 1049–1082. https://doi.org/10.1111/1468-0262.00439

Kloeden, P. E. & Platen, E. (1992). Numerical Solution of Stochastic Differential Equations. Springer-Verlag. https://doi.org/10.1007/978-3-662-12616-5

Kreiss, J. P. & Lahiri, S. N. (2012). 1-Bootstrap Methods for Time Series. In Subba Rao, T., Subba Rao, S. & Rao, C. R. (Eds.), *Handbook of Statistics*, 30, 3–26. Elsevier. https://doi.org/10.1016/B978-0-444-53858-1.00001-6

Mikulevičius, R. & Zhang, C. (2015). Weak Euler Approximation for Itô Diffusion and Jump Processes. *Stochastic Analysis and Applications*, 33(3), 549–571. https://doi.org/10.1080/07362994.2015.1014102.

Pan, L. & Politis, D. N. (2016a). Bootstrap prediction intervals for Markov processes. *Computational Statistics & Data Analysis*, 100, 467–494. https://doi.org/10.1016/j.csda.2015.05.010

Pan, L. & Politis, D. N. (2016b). Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions. *Journal of Statistical Planning and Inference*, 177, 1–27. https://doi.org/10.1016/j.jspi.2014.10.003

Platen, E. (1999). An introduction to numerical methods for stochastic differential equations. *Acta Numerica*, 8, 197–246. https://doi.org/10.1017/S0962492900002920

Sanders, A. & Unal, H. (2001). On the Intertemporal Behavior of the Short-Term Rate of Interest. *Journal of Financial and Quantitative Analysis*, 23, 417–423. https://doi.org/10.2307/2331080

Tang, C. Y. & Chen, S. X. (2009). Parameter estimation and bias correction for diffusion processes. *Journal of Econometrics*, 149(1), 65–81. https://doi.org/10.1016/j.jeconom.2008.11.001

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2), 177–188. https://doi.org/10.1016/0304-405X(77)90016-2.