Model Assisted Statistics and Applications 10 (2015) 53–61
DOI 10.3233/MAS-140311
IOS Press

Asymptotic distribution of the bootstrap parameter estimator for AR(1) process

Bambang Suprihatin[a,*], Suryo Guritno[b] and Sri Haryatmi[b]
[a]*Mathematics Department at University of Sriwijaya, Palembang, Indonesia*
[b]*Mathematics Department at University of Gadjah Mada, Yogyakarta, Indonesia*

Abstract. In this paper we investigated the asymptotic distribution of the bootstrap parameter estimator of a first order autoregressive AR(1) model. We described the asymptotic distribution of such estimator by applying the delta method and employing two different approaches, and concluded that the two approaches lead to the same conclusion, viz. both results converge in distribution to a normal distribution. We also presented the Monte Carlo simulation of the residuals bootstrap and application with real data was carried out in order to yield apparent conclusions.

Keywords: Asymptotic distribution, autoregressive, bootstrap, delta method, Monte Carlo simulation

1. Introduction

Consider the following first order autoregressive AR(1) process:

$$X_t = \theta X_{t-1} + \varepsilon_t,$$

where $\{\varepsilon_t\}$ is a zero mean white noise process with constant variance σ^2. Let $\hat{\theta}$ be the estimator of the parameter θ. Studying of estimation of the unknown parameter θ involves:

(i) what estimator $\hat{\theta}$ should be used?
(ii) having choosen to use particular $\hat{\theta}$, is this estimator consistent to the population parameter θ?
(iii) how accurate is $\hat{\theta}$ as an estimator for true parameter θ?
(iv) the interesting one is, what is the asymptotic behaviour of such estimator?

Bootstrap is a general methodology for answering the second and third questions, while the delta method is one of tools used to answer the last question. Consistency theory is needed to ensure that the estimator is consistent to the actual parameter as desired, and thereof the asymptotic behaviour of such estimator will be studied.

Let θ be a parameter, i.e. coefficient of stationary AR(1) process. The estimator for θ is $\hat{\theta} = \hat{\rho}_n(1) = \sum_{t=2}^{n} X_{t-1}X_t / \sum_{t=1}^{n} X_t^2$. The consistency theories of $\hat{\theta}$ have studied in [3,5,10], and for bootstrap version of the same topic, see [1,4,6,7]. They deal with the bootstrap approximation in various senses (e.g., consistency of estimator, asymptotic normality, applying of Edgeworth expansions, etc.), and they reported that the bootstrap works usually very well. Bose [2] studied the accuracy of the bootstrapping method for autoregressive model. He proved that the parameter estimates of the autoregressive model can be bootstrapped with accuracy $o\left(n^{-1/2}\right)$ a.s., thus, it outperforms the normal approximation, which has accuracy "only" of order $O\left(n^{-1/2}\right)$. Suprihatin et al. [8] studied the advantage of bootstrap by simulating the data that fits to the AR(1) process, and the results gave a good accuracy.

*Corresponding author: Bambang Suprihatin, Mathematics Department at University of Sriwijaya, Palembang, Indonesia. E-mail: bambangs@unsri.ac.id.

Zhao and Rahardja [11] studied the application of bootstrap on the nonlinear structural case, and showed that the bootstrap works very well.

A good perform of the bootstrap estimator is applied to study the asymptotic distribution of $\hat{\theta}^*$, i.e., bootstrap estimator for parameter of the AR(1) process, using delta method. We describe the asymptotic distribution of the autocovariance function and investigate the bootstrap limiting distribution of $\hat{\theta}^*$. Section 2 reviews the consistency of bootstrap estimate for mean under Kolmogorov metric and describe the estimation of autocovariance function. Section 3 deals with asymptotic distribution of $\hat{\theta}^*$ by applying the delta method. Section 4 discusses the results of Monte Carlo simulations involve bootstrap standard errors and density estimation for $\hat{\theta}^*$. Section 5, the last section, briefly describes the conclusions of the paper.

2. Consistency of estimation of the autocovariance function

Let $(X_1, X_2, \ldots, X_n)$ be a random sample of size n from a population with common distribution F and let $T(X_1, X_2, \ldots, X_n; F)$ be the specified random variable or statistic of interest, possibly depending upon the unknown distribution F. Let F_n denote the empirical distribution function of $(X_1, X_2, \ldots, X_n)$, i.e., the distribution putting probability $1/n$ at each of the points $X_1, X_2, \ldots, X_n$. A bootstrap sample is defined to be a random sample of size n drawn from F_n, say $X^* = (X_1^*, X_2^*, \ldots, X_n^*)$. The bootstrap sample at first bootstrapping is usually denoted by X^{*1}, at second bootstrapping is denoted by X^{*2}, and so on. In general, the bootstrap sample at Bth bootstrapping is denoted by X^{*B}. The bootstrap data set $X^{*b} = \left(X_1^{*b}, X_2^{*b}, \ldots, X_n^{*b}\right), b = 1, 2, \ldots, B$ consists of members of the original data set $(X_1, X_2, \ldots, X_n)$, some appearing zero times, some appearing once, some appearing twice, etc. The bootstrap method is to approximate the distribution of $T(X_1, X_2, \ldots, X_n; F)$ under F by that of $T(X_1^*, X_2^*, \ldots, X_n^*; F_n)$ under F_n.

Let a functional T is defined as $T(X_1, X_2, \ldots, X_n; F) = \sqrt{n}(\hat{\theta} - \theta)$ where $\hat{\theta}$ is the estimator for the coefficient θ of stationary AR(1) process. The bootstrap version of T is $T(X_1^*, X_2^*, \ldots, X_n^*; F_n) = \sqrt{n}(\hat{\theta}^* - \hat{\theta})$, where $\hat{\theta}^*$ is a bootstrap version of $\hat{\theta}$ computed from sample bootstrap $X^* = (X_1^*, X_2^*, \ldots, X_n^*)$. In bootstrap view, the key of bootstrap terminology says that the population is to the sample as the sample is to the bootstrap samples. Therefore, when we want to investigate the asymptotic distribution of bootstrap estimator for $\hat{\theta}$, we investigate the distribution of $\sqrt{n}(\hat{\theta}^* - \hat{\theta})$ contrast to $\sqrt{n}(\hat{\theta} - \theta)$. Thus, the bootstrap method is a device for estimating $P_F(\sqrt{n}(\hat{\theta} - \theta) \leqslant x)$ by $P_{F_n}(\sqrt{n}(\hat{\theta}^* - \hat{\theta}) \leqslant x)$.

Suppose we have the observed values $X_1, X_2, \ldots, X_n$ from the stationary AR(1) process. A natural estimators for parameters mean, covariance and correlation function are

$$\hat{\mu}_n = \bar{X}_n = \frac{1}{n}\sum_{t=1}^{n} X_t, \hat{\gamma}_n(h) = \frac{1}{n}\sum_{t=1}^{n-h}(X_{t+h} - \bar{X}_n)(X_t - \bar{X}_n),$$

and $\hat{\rho}_n(h) = \hat{\gamma}_n(h)/\hat{\gamma}_n(0)$ respectively. These all three estimators are consistent (see, e.g., [3,10]). If the series X_t is replaced by the centered series $X_t - \mu_X$, then the autocovariance function does not change. Therefore, studying the asymptotic properties of the sample autocovariance function $\hat{\gamma}_n(h)$, it is not a loss of generality to assume that $\mu_X = 0$. The sample autocovariance function can be written as

$$\hat{\gamma}_n(h) = \frac{1}{n}\sum_{t=1}^{n-h} X_{t+h}X_t - \bar{X}_n\left(\frac{1}{n}\sum_{t=1}^{n-h} X_t\right) - \left(\frac{1}{n}\sum_{t=1}^{n} X_t\right)\bar{X}_n + \left(\bar{X}_n\right)^2. \quad (1)$$

Under some conditions (see, e.g., [10]), the last three terms in Eq. (1) are of the order $O_p(1/n)$. Thus, under the assumption that $\mu_X = 0$, we can write Eq. (1) as

$$\hat{\gamma}_n(h) = \frac{1}{n}\sum_{t=1}^{n-h} X_{t+h}X_t + O_p(1/n).$$

The asymptotic behaviour of the sequence $\sqrt{n}\left(\hat{\gamma}_n(h) - \gamma_X(h)\right)$ depends only on $n^{-1}\sum_{t=1}^{n-h} X_{t+h}X_t$. Note that a change of $n-h$ by n is asymptotically negligible, so that, for simplicity of notation, we can equivalently study the average

$$\tilde{\gamma}_n(h) = \frac{1}{n}\sum_{t=1}^{n} X_{t+h}X_t.$$

Both $\hat{\gamma}_n(h)$ and $\tilde{\gamma}_n(h)$ are unbiased estimators of $E\left(X_{t+h}X_t\right) = \gamma_X(h)$, under the condition that $\mu_X = 0$. Their asymptotic distribution then can be derived by applying a central limit theorem to the averages $\bar{Y}_n$ of the variables $Y_t = X_{t+h}X_t$. The asymptotic variance takes the form $\sum_g \gamma_Y(g)$ and in general depends on fourth order moments of the type $E\left(X_{t+g+h}X_{t+g}X_{t+h}X_t\right)$ as well as on the autocovariance function of the series X_t. Van der Vaart [10] showed that the autocovariance function of the series $Y_t = X_{t+h}X_t$ can be written as

$$V_{h,h} = \kappa_4(\varepsilon)\gamma_X(h)^2 + \sum_g \gamma_X(g)^2 + \sum_g \gamma_X(g+h)\gamma_X(g-h), \tag{2}$$

where $\kappa_4(\varepsilon) = \frac{E\left(\varepsilon_1^4\right)}{E\left(\varepsilon_1^2\right)^2} - 3$, the fourth cumulant of ε_t. The following theorem gives the asymptotic distribution of the sequence $\sqrt{n}\left(\hat{\gamma}_n(h) - \gamma_X(h)\right)$.

Theorem 1. If $X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}$ holds for an i.i.d. sequence ε_t with mean zero and $E\left(\varepsilon_t^4\right) < \infty$ and numbers ψ_j with $\sum_j |\psi_j| < \infty$, then $\sqrt{n}\left(\hat{\gamma}_n(h) - \gamma_X(h)\right) \to_d N\left(0, V_{h,h}\right)$.

3. Asymptotic distribution of bootstrap estimate for parameter of AR(1) process using delta method

The delta method consists of using a Taylor expansion to approximate a random vector of the form $\phi\left(T_n\right)$ by the polynomial $\phi\left(\theta\right) + \phi'\left(\theta\right)\left(T_n - \theta\right) + \ldots$ in $T_n - \theta$. This method is useful to deduce the limit law of $\phi\left(T_n\right) - \phi\left(\theta\right)$ from that of $T_n - \theta$, which is guaranteed by the following theorem.

Theorem 2. Let $\phi : \Re^k \to \Re^m$ be a map defined on a subset of $\Re^k$ and differentiable at θ. Let T_n be random vectors taking their values in the domain of ϕ. If $r_n\left(T_n - \theta\right) \to_d T$ for numbers $r_n \to \infty$, then $r_n\left(\phi\left(T_n\right) - \phi(\theta)\right) \to_d \phi'_\theta(T)$. Moreover, the difference between $r_n\left(\phi\left(T_n\right) - \phi(\theta)\right)$ and $\phi'_\theta\left(r_n\left(T_n - \theta\right)\right)$ converges to zero in probability.

Assume that $\hat{\theta}_n$ is a statistic, and that ϕ is a given differentiable map. The bootstrap estimator for the distribution of $\phi(\hat{\theta}_n) - \phi(\theta)$ is $\phi(\hat{\theta}_n^*) - \phi(\hat{\theta}_n)$. If the bootstrap is consistent for estimating the distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$, then it is also consistent for estimating the distribution of $\sqrt{n}(\phi(\hat{\theta}_n) - \phi(\theta))$, as given in the following theorem. The proof of the theorem is due to [10].

Theorem 3. (Delta Method For Bootstrap) Let $\phi : \Re^k \to \Re^m$ be a measurable map defined and continuously differentiable in a neighborhood of θ. Let $\hat{\theta}_n$ be random vectors taking their values in the domain of ϕ that converge almost surely to θ. If $\sqrt{n}(\hat{\theta}_n - \theta) \to_d T$ and $\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}) \to_d T$ conditionally almost surely, then both $\sqrt{n}(\phi(\hat{\theta}_n) - \phi(\theta)) \to_d \phi'_\theta(T)$ and $\sqrt{n}(\phi(\hat{\theta}_n^*) - \phi(\hat{\theta}_n)) \to_d \phi'_\theta(T)$ conditionally almost surely.

Proof By applying the mean value theorem, the difference $\phi(\hat{\theta}_n^*) - \phi(\hat{\theta}_n)$ can be written as $\phi'_{\bar{\theta}_n}(\hat{\theta}_n^* - \hat{\theta}_n)$ for a point $\bar{\theta}_n$ between $\hat{\theta}_n^*$ and $\hat{\theta}_n$, if the latter two points are in the ball around θ in which ϕ is continuously differentiable. By the continuity of the derivative, there exists a constant $\delta > 0$ for every $\eta > 0$ such that $\|\phi'_{\theta'} - \phi'_\theta h\| < \eta\|h\|$ for every h and every $\|\hat{\theta}_n - \theta\| \leqslant \delta$. If n is sufficiently large, δ sufficiently small, $\sqrt{n}\|\hat{\theta}_n^* - \hat{\theta}_n\| \leqslant M$, and $\|\hat{\theta}_n - \theta\| \leqslant \delta$, then

$$\begin{aligned} R_n :&= \left\|\sqrt{n}\left(\phi(\hat{\theta}_n^*) - \phi(\hat{\theta}_n)\right) - \phi'_\theta\sqrt{n}\left(\hat{\theta}_n^* - \hat{\theta}_n\right)\right\| \\ &= \left|\left(\phi'_{\bar{\theta}_n} - \phi'_\theta\right)\sqrt{n}\left(\hat{\theta}_n^* - \hat{\theta}_n\right)\right| \leqslant \eta M. \end{aligned}$$

Fix a number $\varepsilon > 0$ and a large number M. For η sufficiently small to ensure that $\eta M < \varepsilon$,

$$P\left(R_n > \varepsilon|\hat{P}_n\right) \leqslant P\left(\sqrt{n}\|\hat{\theta}_n^* - \hat{\theta}_n\| > M \text{ or } \|\hat{\theta}_n - \theta\| > \delta|\hat{P}_n\right). \tag{3}$$

Since $\hat{\theta}_n \to_{a.s.} \theta$, the right side of Eq. (3) converges almost surely to $P(\|T\| \geqslant M)$ for every continuity point M of $\|T\|$. This can be made arbitrarily small by choice of M. Conclude that the left side of Eq. (3) converges to zero almost surely, and hence $\sqrt{n}(\phi(\hat{\theta}_n^*) - \phi(\hat{\theta}_n)) - \phi'_\theta\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n) \to_{a.s.} 0$. Because the random vectors $\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}) \to_d T$ and the matrix ϕ'_θ is continuous, by applying the continuous-mapping theorem we conclude that $\phi'_\theta\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n) \to_d \phi'_\theta(T)$. By applying the Slutsky's lemma, we obtain $\sqrt{n}(\phi(\hat{\theta}_n^*) - \phi(\hat{\theta}_n)) \to_{a.s.} \phi'_\theta\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n)$, and by the earlier conclusion we also conclude that $\sqrt{n}(\phi(\hat{\theta}_n^*) - \phi(\hat{\theta}_n)) \to_d \phi'_\theta(T)$, completing the proof. □

The moment estimator for the parameter of stationary AR(1) process is obtained from the Yule-Walker equation, i.e. $\hat{\theta} = \hat{\rho}_n(1)$ where $\hat{\rho}_n(1)$ be the lag time 1 of the sample autocorrelation

$$\hat{\rho}_n(1) = \frac{\sum_{t=2}^n X_{t-1}X_t}{\sum_{t=1}^n X_t^2}. \tag{4}$$

According to Davison and Hinkley [4], the estimate of standard error of parameter $\hat{\theta}$ is $\widehat{se}(\hat{\theta}) = \sqrt{(1-\hat{\theta}^2)/n}$. Meanwhile, the bootstrap version of standard error was introduced in [5]. In Section 4 we demonstrate the results of the Monte Carlo simulation along with standard errors and brief comments.

In accordance with the Theorem 3, we construct a measurable function ϕ which is defined to be continuous and differentiable in the neighborhood of θ. In Suprihatin et al. [8], the function ϕ was obtained by iterating Eq. (4) as follows:

$$\begin{aligned}\hat{\rho}_n(1) &= \frac{\sum_{t=2}^n X_{t-1}\left(\theta X_{t-1} + \varepsilon_t\right)}{\sum_{t=1}^n X_t^2} \\ &= \frac{\theta\sum_{t=2}^n X_{t-1}^2 + \sum_{t=2}^n X_{t-1}\varepsilon_t}{\sum_{t=1}^n X_t^2} \\ &= \frac{\theta\left(\sum_{t=2}^{n+1} X_{t-1}^2 - X_n^2\right) + \sum_{t=2}^n X_{t-1}\varepsilon_t}{\sum_{t=1}^n X_t^2} \\ &= \frac{\frac{\theta}{n}\sum_{t=1}^n X_t^2 - \frac{\theta}{n}X_n^2 + \frac{1}{n}\sum_{t=2}^n X_{t-1}\varepsilon_t}{\frac{1}{n}\sum_{t=1}^n X_t^2}.\end{aligned}$$

Brockwell and Davis [3] showed that $\hat{\rho}_n(1)$ was a consistent estimator of the true parameter $\theta = \rho_X(1)$. Kolmogorov strong law of large number asserts that $\frac{1}{n}\sum_{t=2}^n X_{t-1}\varepsilon_t \to_{a.s.} E\left(X_{t-1}\varepsilon_t\right)$. Since X_{t-1} is independent of ε_t, then $E\left(X_{t-1}\varepsilon_t\right) = 0$ and hence $\frac{1}{n}\sum_{t=2}^n X_{t-1}\varepsilon_t \to_{a.s.} 0$. By applying the Slutsky's lemma, the last display is approximated by $\tilde{\rho}_n(1) = (\theta\overline{X^2} - \frac{\theta}{n}X_n^2)/\overline{X^2}$. Thus, for $n \to \infty$ we obtain $\hat{\theta} \to_{a.s.} \tilde{\rho}_n(1)$. Moreover, we can see that $\tilde{\rho}_n(1)$ equals to $\phi(\overline{X^2})$ for a function $\phi(x) = \left(\theta x - \frac{\theta}{n}X_n^2\right)/x$. Since ϕ is continous and hence is measurable. Note that the stationarity of the AR(1) process guarantees that the AR(1) process is a causal one, so that the process can be expressed as a moving average of infinite order or MA(∞) process. Hence, Theorem 1 can be applied to the AR(1) process. By the central limit theorem and applying the Theorem 1 we conclude that

$$\sqrt{n}\left(\overline{X^2} - \gamma_X(0)\right) \to_d N\left(0, V_{0,0}\right),$$

where $V_{0,0} = \kappa_4(\varepsilon)\gamma_X(0)^2 + 2\sum_g \gamma_X(g)^2$ as in Eq. (2) for $h = 0$. The map ϕ is differentiable at the point $\gamma_X(0)$, with derivative $\phi' = \frac{\partial}{\partial x}\phi(x) = \frac{\theta X_n^2}{nx^2}$ and $\phi'_{\gamma_X(0)} = \frac{\theta X_n^2}{n\gamma_X(0)^2}$. Theorem 2 says that

$$\sqrt{n}\left(\phi\left(\overline{X^2}\right) - \phi(\gamma_X(0))\right) = \phi'_{\gamma_X(0)}\left(\sqrt{n}\left(\frac{1}{n}\sum_{t=1}^n X_t^2 - \gamma_X(0)\right)\right) + o_p(1)$$

$$= \frac{\theta X_n^2}{n\gamma_X(0)^2}\sqrt{n}\left(\frac{1}{n}\sum_t X_t^2 - \gamma_X(0)\right) + o_p(1).$$

In view of Theorem 2, if T possesses the normal distribution with mean 0 and variance $V_{0,0}$, then

$$\sqrt{n}\left(\hat{\theta}-\theta\right) = \sqrt{n}\left(\phi\left(\overline{X^2}\right) - \phi(\gamma_X(0))\right) \to_d \frac{\theta X_n^2}{n\gamma_X(0)^2}T \sim N\left(0, \left(\frac{\theta X_n^2}{n\gamma_X(0)^2}\right)^2 V_{0,0}\right).$$

Meantime, the bootstrap version of $\hat{\theta}$, denoted by $\hat{\theta}^*$ can be obtained as follows (see, e.g., [5,6]):

(i) Define the residuals $\hat{\varepsilon}_t = X_t - \hat{\theta}X_{t-1}$ for $t = 2, 3, \ldots, n$.

(ii) A bootstrap sample $X_1^*, X_2^*, \ldots, X_n^*$ is created by sampling $\varepsilon_2^*, \varepsilon_3^*, \ldots, \varepsilon_n^*$ with replacement from the residuals. Letting $X_1^* = X_1$ as an initial bootstrap sample and $X_t^* = \hat{\theta}X_{t-1}^* + \varepsilon_t^*$, $t = 2, 3, \ldots, n$.

(iii) Finally, after centering the bootstrap time series $X_1^*, X_2^*, \ldots, X_n^*$ i.e. X_i^* is replaced by $X_i^* - \bar{X}^*$ where $\bar{X}^* = \frac{1}{n}\sum_{t=1}^n X_t^*$. Using the *plug-in* principle we obtain the bootstrap estimate, $\hat{\theta}^* = \hat{\rho}_n^*(1) = \frac{\sum_{t=2}^n X_{t-1}^* X_t^*}{\sum_{t=1}^n X_t^{*2}}$ computed from the bootstrap sample $X_1^*, X_2^*, \ldots, X_n^*$.

Analog with the previous discussion, we obtain the bootstrap version for the counterpart of $\tilde{\rho}_n(1)$, is that $\tilde{\rho}_n^*(1)$ which equals to $\phi(\overline{X^{*2}})$ for a measurable function $\phi(x) = \left(\theta x - \frac{\theta}{n}X_n^2\right)/x$. Thus, in view of Theorem 3 we conclude that $\tilde{\rho}_n^*(1)$ converges to $\tilde{\rho}_n(1)$ conditionally almost surely. By the Glivenko-Cantelli lemma and applying the *plug-in* principle, we obtain

$$\sqrt{n}\left(\overline{X^{*2}} - \hat{\gamma}_n(0)\right) \to_d N\left(0, V_{0,0}^*\right)$$

and

$$\sqrt{n}\left(\hat{\theta}^* - \hat{\theta}\right) = \sqrt{n}\left(\phi\left(\overline{X^{*2}}\right) - \phi(\hat{\gamma}_n(0))\right) \to_d N\left(0, \left(\frac{\theta X_n^{*2}}{n\hat{\gamma}_n^*(0)^2}\right)^2 V_{0,0}^*\right).$$

However, this method for finding a measurable function ϕ is only efficient for AR(1) model and in general is inefficient for AR(p) model with $p \geqslant 2$. This inefficiency is due to the complicated iterations.

Another way to describe the function ϕ and asymptotic behaviour of $\hat{\rho}_n(1)$ as follows. Since $\hat{\rho}_n(1) = \sum_{t=2}^n X_{t-1}X_t / \sum_{t=1}^n X_t^2$, then we can rewrite this expression as

$$\hat{\rho}_n(1) = \phi\left(\frac{1}{n}\sum_{t=1}^n X_t^2, \frac{1}{n}\sum_{t=2}^n X_{t-1}X_t\right),$$

for the function $\phi : \Re^2 \to \Re$ with $\phi(u, v) = v/u$. According to Theorem 1, the multivariate central limit theorem for $\left(\frac{1}{n}\sum_{t=1}^n X_t^2, \frac{1}{n}\sum_{t=2}^n X_{t-1}X_t\right)^T$ is

$$\sqrt{n}\left(\begin{pmatrix} \frac{1}{n}\sum_{t=1}^n X_t^2 \\ \frac{1}{n}\sum_{t=2}^n X_{t-1}X_t \end{pmatrix} - \begin{pmatrix} \gamma_X(0) \\ \gamma_X(1) \end{pmatrix}\right) \to_d N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} V_{0,0} & V_{0,1} \\ V_{1,0} & V_{1,1} \end{pmatrix}\right). \quad (5)$$

The map ϕ is differentiable with the matrix

$$\phi' = \begin{pmatrix} \frac{\partial}{\partial u}\phi(u, v) & \frac{\partial}{\partial v}\phi(u, v) \end{pmatrix} = \begin{pmatrix} \frac{-v}{u^2} & \frac{1}{u} \end{pmatrix},$$

and

$$\phi'_{(\,\gamma_X(0)\ \gamma_X(1)\,)} = \begin{pmatrix} \dfrac{-\gamma_X(1)}{\gamma_X(0)^2} & \dfrac{1}{\gamma_X(0)} \end{pmatrix}.$$

By applying the Theorem 2,

$$\sqrt{n}\left(\phi\left(\frac{1}{n}\sum_{t=1}^{n} X_t^2, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}X_t\right) - \phi\left(\gamma_X(0), \gamma_X(1)\right)\right)$$

$$= \phi'_{\left(\gamma_X(0)\ \gamma_X(1)\right)} \begin{pmatrix} \sqrt{n}\left(\dfrac{1}{n}\displaystyle\sum_{t=1}^{n} X_t^2 - \gamma_X(0)\right) \\ \sqrt{n}\left(\dfrac{1}{n}\displaystyle\sum_{t=2}^{n} X_{t-1}X_t - \gamma_X(1)\right) \end{pmatrix} + o_p(1)$$

$$= \frac{-\gamma_X(1)}{\gamma_X(0)^2}\sqrt{n}\left(\frac{1}{n}\sum_{t=1}^{n} X_t^2 - \gamma_X(0)\right) + \frac{1}{\gamma_X(0)}\sqrt{n}\left(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}X_t - \gamma_X(1)\right) + o_p(1).$$

In view of Theorem 2, if $(Z_1, Z_2)^T$ possesses the normal distribution as in Eq. (5), then

$$\frac{-\gamma_X(1)}{\gamma_X(0)^2}\sqrt{n}\left(\frac{1}{n}\sum_{t=1}^{n} X_t^2 - \gamma_X(0)\right) + \frac{1}{\gamma_X(0)}\sqrt{n}\left(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}X_t - \gamma_X(1)\right)$$

$$\rightarrow_d \frac{-\gamma_X(1)}{\gamma_X(0)^2} Z_1 + \frac{1}{\gamma_X(0)} Z_2 \sim N\left(0, \tau^2\right),$$

where

$$\tau^2 = Var\left(\frac{-\gamma_X(1)}{\gamma_X(0)^2} Z_1 + \frac{1}{\gamma_X(0)} Z_2\right)$$

$$= \left(\frac{\gamma_X(1)}{\gamma_X(0)^2}\right)^2 Var(Z_1) + \frac{1}{\gamma_X(0)^2} Var(Z_2) - \frac{2\gamma_X(1)}{\gamma_X(0)^3} Cov(Z_1, Z_2)$$

$$= \left(\frac{\gamma_X(1)}{\gamma_X(0)^2}\right)^2 V_{0,0} + \frac{1}{\gamma_X(0)^2} V_{1,1} - \frac{2\gamma_X(1)}{\gamma_X(0)^3} V_{0,1}.$$

Thus, by Theorem 2 we conclude that

$$\sqrt{n}\left(\phi\left(\frac{1}{n}\sum_{t=1}^{n} X_t^2, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}X_t\right) - \phi\left(\gamma_X(0), \gamma_X(1)\right)\right) \rightarrow_d N\left(0, \tau^2\right).$$

Similarly and by applying the *plug-in* principle, Theorem 3 gives the bootstrap version for the above result,

$$\sqrt{n}\left(\begin{pmatrix} \dfrac{1}{n}\displaystyle\sum_{t=1}^{n} X_t^{*2} \\ \dfrac{1}{n}\displaystyle\sum_{t=2}^{n} X_{t-1}^{*}X_t^{*} \end{pmatrix} - \begin{pmatrix} \hat{\gamma}_n(0) \\ \hat{\gamma}_n(1) \end{pmatrix}\right) \rightarrow_d N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} V_{0,0}^{*} & V_{0,1}^{*} \\ V_{1,0}^{*} & V_{1,1}^{*} \end{pmatrix}\right)$$

and

$$\sqrt{n}(\hat{\theta}^{*} - \hat{\theta}) = \sqrt{n}\left(\phi\left(\frac{1}{n}\sum_{t=1}^{n} X_t^{*2}, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^{*}X_t^{*}\right) - \phi\left(\frac{1}{n}\sum_{t=1}^{n} X_t^2, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}X_t\right)\right) \rightarrow_d N\left(0, \tau^{2*}\right),$$

where

$$\tau^{2*} = \left(\frac{\hat{\gamma}_n^{*}(1)}{\hat{\gamma}_n^{*}(0)^2}\right)^2 V_{0,0}^{*} + \frac{1}{\hat{\gamma}_n^{*}(0)^2} V_{1,1}^{*} - \frac{2\hat{\gamma}_n^{*}(1)}{\hat{\gamma}_n^{*}(0)^3} V_{0,1}^{*}. \tag{6}$$

The results which are obtained by these two ways (methods) lead to the same conclusion, i.e. both results converge in distribution to a normal distribution, but the variances of both normal distributions are different as they are obtained by two different approaches.

Table 1
The estimates of $\hat{\theta}^*$ and $\widehat{se}(\hat{\theta})$ as compared to the estimate of $\hat{\theta}$ and $\widehat{se}(\hat{\theta})$ respectively, for various sample size n and bootstrap sample size B

		B				
		50	200	1,000	2,000	
$n = 20$	$\hat{\theta}^*$	0.5947	0.5937	0.6044	0.6224	$\hat{\theta} = 0.7126$
	$\widehat{se}(\hat{\theta})$	0.1836	0.1758	0.1793	0.1725	$\widehat{se}(\hat{\theta}) = 0.1569$
$n = 30$	$\hat{\theta}^*$	0.6484	0.6223	0.6026	0.6280	$\hat{\theta} = 0.7321$
	$\widehat{se}(\hat{\theta})$	0.1496	0.1522	0.1453	0.1408	$\widehat{se}(\hat{\theta}) = 0.1244$
$n = 50$	$\hat{\theta}^*$	0.5975	0.6051	0.5792	0.6002	$\hat{\theta} = 0.6823$
	$\widehat{se}(\hat{\theta})$	0.1162	0.1178	0.1093	0.1103	$\widehat{se}(\hat{\theta}) = 0.1034$
$n = 100$	$\hat{\theta}^*$	0.6242	0.6104	0.6310	0.6197	$\hat{\theta} = 0.6884$
	$\widehat{se}(\hat{\theta})$	0.0962	0.1006	0.0994	0.0986	$\widehat{se}(\hat{\theta}) = 0.0736$

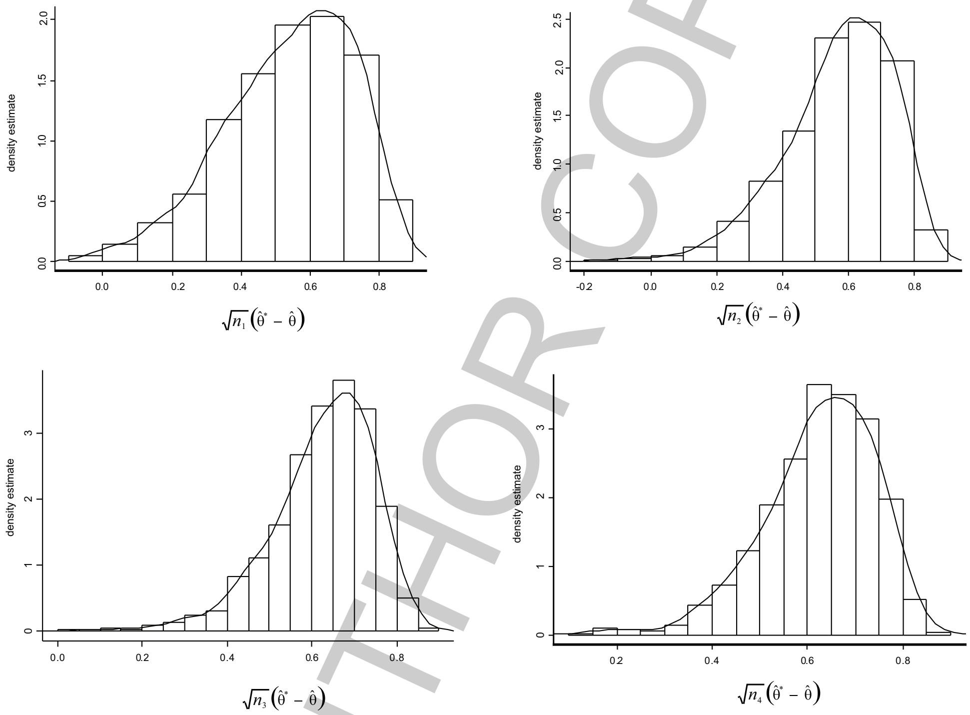

Fig. 1. Histogram and plot of density estimates of 1,000 bootstrap random samples $\sqrt{n_i}(\hat{\theta}^* - \hat{\theta}), i = 1, 2, 3, 4$, with sample of size $n_1 = 20$ (top left), $n_2 = 30$ (top right), $n_3 = 50$ (buttom left), and $n_4 = 100$ (buttom right).

4. The results of Monte Carlo simulations

In this section, we present the results of the Monte Carlo simulation to illustrate an application of the bootstrap to an AR(1) process. The simulation is conducted using S-Plus and the data sets used are 20, 30, 50, and 100 time series data of the exchange rate of US Dollar as compared to Indonesian Rupiah. Let $n_i, i = 1, 2, 3, 4$ be the size of ith data set respectively. The data is taken from the authorized website of Bank Indonesia for transactions during May up to August 2012. Let the count for the tth transaction be X_t and identified as a sample. After centering the four data sets (replacing X_t by $X_t - \bar{X}_t$), then we fit an AR(1) model $X_t = \theta X_{t-1} + \varepsilon_t, t = 1, 2, \ldots, n_i$, and $i = 1, 2, 3, 4$, where $\varepsilon_t \sim \text{WN}\left(0, \sigma^2\right)$. For the data of size $n_1 = 20$, the simulation gives the estimate of $\hat{\theta}$ turned out to be 0.7126 with an estimated standard error $\widehat{se}(\hat{\theta}) = 0.1569$. The simulation shows that larger the n smaller is the estimated standard error, so larger n means a better estimate of θ, as is seen in Table 1.

The bootstrap estimator of $\hat{\theta}$ is usually denoted by $\hat{\theta}^*$. How accurate is $\hat{\theta}^*$ as an estimator for $\hat{\theta}$? To answer the question, we need a bootstrap estimated standard error which is denoted by $\widehat{se}(\hat{\theta}^*)$, as a measure of statistical accuracy. To do so, we resample the data X_t as many B ranging from 50 to 2,000 for each sample of size $n_i, i = 1, 2,\ 3,\ 4$. To produce a good approximation, Davison and Hinkley [4] and Efron and Tibshirani [5] suggested to use the number of bootstrap samples (B) at least 50. As mentioned, the simulation involves a large number B and repeated computation of random numbers, so the Monte Carlo method is most suited to calculation by a computer. Table 1 shows the results of simulation for various size of data sets and the number of bootstrap samples. As we can see, the increasing of the number of bootstrap samples tends to yield the estimates of $\widehat{se}(\hat{\theta}^*)$ close to the estimated standard error, $\widehat{se}(\hat{\theta})$. For example, for small sample of size $n_1 = 20$, the bootstrap shows a fair performance. Using the bootstrap samples $B = 50$, the resulting of its bootstrap standard error $\widehat{se}(\hat{\theta}^*)$ turned out to be 0.1836, while the estimated standard error $\widehat{se}(\hat{\theta}) = 0.1569$. The difference between the two estimates is somewhat large. However, if we employ 1,000 and 2,000 bootstrap samples of size 20 each, the simulation yields $\widehat{se}(\hat{\theta}^*)$ to be 0.1793 and 0.1725 repectively, versus their estimated standard error of 0.1569. This fact shows a better performance of the bootstrap method along with the increasing of the number of bootstrap samples used. A better performance of bootstrap is also shown when we simulate a larger sample, as we can see in Table 1. For $n_4 = 100$ the bootstrap estimate of standard errors are 0.0962 and 0.0986 for $B = 50$ and 2,000 respectively, agreeing nicely with the estimated standard error of 0.1034.

Meanwhile, the four histograms and four plots of density estimates of 1,000 bootstrap samples of $\sqrt{n}(\hat{\theta}^* - \hat{\theta})$ are presented in Fig. 1. The top row of Fig. 1 shows the distribution of random variable $\sqrt{n}(\hat{\theta}^* - \hat{\theta})$ looks skewed because of employing the small size of samples used, i.e. 20 and 30 respectively. Overall, from Fig. 1 we can see that the four resulting histograms are closely related to the probability density of normal random variables. In fact, the four plots of density estimates are resemble a plot of the probability density function (pdf) of an $N(0, \tau^{2*})$ random variable, where τ^{2*} as in Eq. (6). One can see that larger the n closer the density estimates in estimating the pdf of the normal random variable. This result agrees with the result of [2,6].

5. Conclusions

A number of points arise from the study of Sections 2–4, amongst which we state as follows.

(i) Consider an AR(1) process $X_t = \theta X_{t-1} + \varepsilon_t$, with Yule-Walker estimator $\hat{\theta} = \hat{\rho}_n(1)$ of the true parameter $\theta = \rho_X(1)$. The crux result, by applying the delta method we have shown that the asymptotic distribution,

$$\sqrt{n}\left(\hat{\theta}^* - \hat{\theta}\right) = \sqrt{n}\left(\phi\left(\frac{1}{n}\sum_{t=1}^{n} X_t^{*2}, \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^* X_t^*\right) - \phi\left(\frac{1}{n}\sum_{t=1}^{n} X_t^2, \frac{1}{n}\sum_{t=2}^{n} X_{t-1} X_t\right)\right) \\ \to_d N\left(0, \tau^{2*}\right),$$

where τ^{2*} as in Eq. (6). This result leads to the same conclusion with that of using the first way. The difference of both variances is a reasonable property, because the two variances are concluded by two different approaches.

(ii) The results of Monte Carlo simulation show a good performance of the bootstrap method in estimating the parameter of an AR(1) process, as represented by its standard errors and plot of density estimates.

Acknowledgments

The authors thank Prof. A.W. van der Vaart of Mathematics Department, Leiden University and Vrije University the Netherlands, for his valuable discussions during and after first author's visit at Leiden University. The authors also thank the MASA Co-Editor-in-chief Dr. Stan Lipovetsky and two reviewers for insightful comments which led to important improvements over an earlier version of manuscript.

References

[1] P.J. Bickel and D.A. Freedman, Some asymptotic theory for the bootstrap, *Ann Statist* **9** (1981), 1996–1217.
[2] A. Bose, Edgeworth correction by bootstrap in autoregressions, *Ann Statist* **16** (1988), 1709–1722.
[3] P.J. Brockwell and R.A. Davis, *Time Series: Theory and Methods,* Springer, New York, 1991.
[4] A.C. Davison and D.V. Hinkley, *Bootstrap Methods and Their Application,* Cambridge University Press, Cambridge, 2006.
[5] B. Efron and R. Tibshirani, Bootstrap methods for standard errors, confidence intervals, and others measures of statistical accuracy, *Statistical Science* **1** (1986), 54–77.
[6] D.A. Freedman, On bootstrapping two-stage least-squares estimates in stationary linear models, *Ann Statist* **12** (1985), 827–842.
[7] K. Singh, On the asymptotic accuracy of Efron's bootstrap, *Ann Statist* **9** (1981), 1187–1195.
[8] B. Suprihatin, S. Guritno and S. Haryatmi, Delta method for deriving the consistency of bootstrap estimator for parameter of autoregressive model, *Proc of the* 8^{th} *World Congress in Prob and Statist*, Istanbul, (2012).
[9] A.W. van der Vaart, *Asymptotic Statistics,* Cambridge University Press, Cambridge, 2000.
[10] A.W. van der Vaart, *Lecture Notes on Time Series,* Vrije Universiteit, Amsterdam, 2012.
[11] Y.D. Zhao and D. Rahardja, Nonlinear structural equation modeling with distribution-free method, *Model Assisted Statistics and Application* **6** (2011), 39–46.