Delta Method on Bootstrapping of Autoregressive Process

Bambang Suprihatin[1,2], Suryo Guritno[3], Sri Haryatmi[4]

[1]University of Sriwijaya, Palembang, Indonesia
[3,4]University of Gadjah Mada, Yogyakarta, Indonesia
[2]Correspondending author: Bambang Suprihatin, e-mail: bambangs@unsri.ac.id

Abstract

Let $\{X_t, t \in T\}$ be autoregressive time series, where T is discrete time, and let $X_1, X_2, \ldots, X_n$ be the sample that satisfies the AR(1) process. Thus, the sample follows the relation $X_t = \theta X_{t-1} + \varepsilon_t$ where $\{\varepsilon_t\}$ is a zero mean white noise process with constant variance σ^2. Let $\hat{\theta}$ be the estimator for parameter θ. Brockwell and Davis (1991) showed that $\hat{\theta} \rightarrow_p \theta$ and $\sqrt{n}\left(\hat{\theta} - \theta\right) \rightarrow_d N\left(0, \sigma^2\right)$. Meantime, by some assumptions, can be showed that the distribution of $\sqrt{n}\left(\overline{X} - \mu\right)$ converges to normal distribution with mean *0* and variance σ^2 as $n \rightarrow \infty$. 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 consistency of the interesting bootstrap estimator for sample mean, we investigate the distribution of $\sqrt{n}\left(\overline{X}^* - \overline{X}\right)$ contrast to $\sqrt{n}\left(\overline{X} - \mu\right)$, where $\overline{X}^*$ is bootstrap version of $\overline{X}$ computed from sample bootstrap X^*. Asymptotic theory of the bootstrap sample mean is useful to study the consistency for many other statistics. Let $\hat{\theta}^*$ be the bootstrap estimator for $\hat{\theta}$. In this paper we investigate the consistency of $\hat{\theta}^*$ using delta method and applying the residuals bootstrap. We also present the Monte Carlo simulations in regard to yield apparent conclusions.

Keywords: Bootstrap, consistency, delta method, Monte Carlo simulations, time series

1. Introduction

Studying of estimation of the unknown parameter θ involves: (1) what estimator $\hat{\theta}$ should be used? (2) having choosen to use particular $\hat{\theta}$, is this estimator consistent to the population parameter θ? (3) how accurate is $\hat{\theta}$ as an estimator of true parameter θ? The bootstrap is a general methodology for answering the second and third questions. Consistency theory is needed to ensure that the estimator is consistent to the actual parameter as desired.

Consider the parameter θ is the population mean. The consistent estimator for θ is the sample mean $\hat{\theta} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. The consistency theory is then extended to the consistency of bootstrap estimator for mean. According to the bootstrap terminology, if we want to investigate the consistency of bootstrap estimator for mean, we investigate the distribution of $\sqrt{n}\left(\overline{X} - \mu\right)$ and $\sqrt{n}\left(\overline{X}^* - \overline{X}\right)$. The consistency of bootstrap under Kolmogorov metric is defined as

$$\sup_x \left| P_F\left(\sqrt{n}\left(\overline{X}-\mu\right)\le x\right) - P_{F_n}\left(\sqrt{n}\left(\overline{X}^* - \overline{X}\right)\le x\right)\right|. \tag{1}$$

Bickel and Freedman (1981) and Singh (1981) showed that (1) converges almost surely to 0 as $n \to \infty$. Meanwhile, Suprihatin, *et.el* (2011) complete the results by giving nice ilustrations for this case. The consistecy of bootstrap for mean is a worthy tool for studying the consistency of other statistics. In this paper, we study the consistecy of bootstrap estimator for parameter of the AR(1) process.

The consistency of bootstrap estimator for mean is then applied to study the consistency of bootstrap estmate for parameter of the AR(1) process using delta method. We describe the consistency of bootstrap estimates for mean and parameter of the AR(1) process. Section 2 reviews the consistency of bootstrap estimate for mean under Kolmogorov metric. Section 3 deal with the consistency of bootstrap estimate for parameter of the AR(1) process using delta method. Section 4 discuss the results of Monte Carlo simulations involve bootstrap standard errors and density estmation for mean and parameter of the AR(1) process. Section 5, is the last section, briefly describes the conclusions of the paper.

2. Consistency of Bootstrap Estimator For Mean

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 $X_1, X_2, \ldots, X_n$. 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 whrere $(X_1^*, X_2^*, \ldots, X_n^*)$ denotes a bootstrapping random sample of size *n* from F_n.

We start with definition of consistency. Let *F* and *G* be two distribution functions on sample space ***X***. Let $\rho(F,G)$ be a metric on the space of distribution on ***X***. For $X_1, X_2, \ldots, X_n$ i.i.d from *F*, and a given functional $T(X_1, X_2, \ldots, X_n; F)$, let

$$H_n(x) = P_F\left(T(X_1, X_2, \ldots, X_n; F) \le x\right),$$
$$H_{Boot}(x) = P_*\left(T\left(X_1^*, X_2^*, \ldots, X_n^*; F_n\right) \le x\right).$$

We say that the bootstrap is consistent (strongly) under ρ for T if $\rho(H_n, H_{Boot}) \to 0$ *a.s.*

Let functional *T* is defined as $T(X_1, X_2, \ldots, X_n; F) = \sqrt{n}\left(\overline{X} - \mu\right)$ where $\overline{X}$ and μ are sample mean and population mean respectively. Bootstrap version of T is $T\left(X_1^*, X_2^*, \ldots, X_n^*; F_n\right) = \sqrt{n}\left(\overline{X}^* - \overline{X}\right)$, where $\overline{X}^*$ is boostrapping sample mean. Bootstrap method is a device for estimating $P_F\left(\sqrt{n}\left(\overline{X}-\mu\right)\le x\right)$ by $P_{F_n}\left(\sqrt{n}\left(\overline{X}^* - \overline{X}\right)\le x\right)$. We will investigate the consistency of bootstrap under Kolmogorov metric which is defined as

$$K(F,G) = \sup_x \left|F(x) - G(x)\right| = \sup_x \left| P_F\left(\sqrt{n}\left(\overline{X}-\mu\right)\le x\right) - P_{F_n}\left(\sqrt{n}\left(\overline{X}^* - \overline{X}\right)\le x\right)\right|.$$

Some theorems and lemma which are needed to show that $K(H_n, H_{Boot}) \to 0$ *a.s.* taken from Hall (1992), Serfling (1980) and van der Vaart (2000), such as Khintchine-Kolmogorov Convergence Theorem, Berry-Essen Theorem, and Zygmund-

Marcinkiewicz SLLN. The consistency of H_{Boot} under Kolmogorov metric have shown by Sigh (1981) and DasGupta (2008). The crux result is that $\bar{X}^* \rightarrow \bar{X}$ a.s. Suprihatin, *et.al* (2011) give nice simulations for this result.

3. Consistency 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 $\phi(T_n)$ by the polynomial $\phi(\theta)+\phi'(\theta)(T_n-\theta)+\cdots$ in $T_n-\theta$. This method is useful to deduce the limit law of $\phi(T_n)-\phi(\theta)$ from that of $T_n-\theta$. This method is also valid in bootstrap view, which is given in the following theorem.

Theorem 1 (Delta Method For Bootstrap) *Let* $\phi:\Re^k \rightarrow \Re^m$ *be a measurable map defined and continously 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* $\sqrt{n}\left(\hat{\theta}_n-\theta\right)\xrightarrow{d} T$ *and* $\sqrt{n}\left(\hat{\theta}_n^*-\hat{\theta}\right)\xrightarrow{d} T$ *conditionally almost surely, then both* $\sqrt{n}\left(\phi(\hat{\theta}_n)-\phi(\theta)\right)\xrightarrow{d}\phi'_\theta(T)$ *and* $\sqrt{n}\left(\phi(\hat{\theta}_n^*)-\phi(\hat{\theta}_n)\right)\xrightarrow{d}\phi'_\theta(T)$ *conditionally almost surely.*

Let μ is the population mean, and then $\bar{X}$ is the sample mean. The Kolmogorov SLLN asserts that $\bar{X}\rightarrow\mu$ a.s. and $\sqrt{n}\left(\bar{X}-\mu\right)\xrightarrow{d} N\left(0,\sigma^2\right)$. The resulting of Section 2 shows that $\sqrt{n}\left(\bar{X}^*-\bar{X}\right)\xrightarrow{d} N\left(0,s^2\right)$. Based on the consistency of the bootstrap for the sample mean we investigate the consistency of the bootstrap estimate for parameter of AR(1) process using delta method.

Let $\{X_t, t=1,2,\ldots,n\}$ be time series data which satisfies the AR (1) process, i.e. if $\{X_t, t=1,2,\ldots,n\}$ follows the equaiton $X_t=\theta X_{t-1}+\varepsilon_t$ where $\{\varepsilon_t\}$ be random variable sequence of *white noise* with mean *0* and variance σ^2. The process is stationary if $|\theta|<1$. The comprehensive discussions for time series can be found in Wei (1990) and Brockwell and Davis (1991).

For the AR(1) process, from Yule-Walker equation we obtain the estimate for θ is $\hat{\theta}=\hat{\rho}_1$ where $\hat{\rho}_1$ be the lag 1 sample autocorrelation

$$\hat{\rho}_1=\frac{\sum_{t=2}^{n} X_{t-1}X_t}{\sum_{t=1}^{n} X_t^2}. \qquad (2)$$

According to Wei (1990) and Brockwell and Davis (1991), the estimate of standard error of parameter θ is $\widehat{se}(\theta)=\sqrt{\frac{1-\hat{\theta}^2}{n}}$. Meanwhile, the bootstrap version of standard error was introduced by Efron, B. and Tibshirani, R. (1986). In Section 4 we demonstrate results of Monte Carlo simulations consist the two of standard errors and give brief comments. From (2) we can see that

$$\hat{\rho}_1=\frac{\sum_{t=2}^{n} X_{t-1}\left(\theta X_{t-1}+\varepsilon_t\right)}{\sum_{t=1}^{n} X_t^2}$$

$$= \frac{\theta \sum_{t=2}^{n} X_{t-1}^2 + \sum_{t=2}^{n} X_{t-1}\varepsilon_t}{\sum_{t=1}^{n} X_t^2}$$

$$= \frac{\theta \left(\sum_{t=2}^{n+1} X_{t-1}^2 - X_n^2 \right) + \sum_{t=2}^{n} X_{t-1}\varepsilon_t}{\sum_{t=1}^{n} X_t^2}$$

$$= \frac{\frac{\theta}{n} \sum_{t=1}^{n} X_t^2 - \frac{\theta}{n} X_n^2 + \frac{1}{n} \sum_{t=2}^{n} X_{t-1}\varepsilon_t}{\frac{1}{n} \sum_{t=1}^{n} X_t^2}$$

Brockwell and Davis (1991) have shown that $\hat{\rho}_1$ is consistent estimator of true parameter $\theta = \rho_1$. Kolmogorov SLLN asserts that $\frac{1}{n}\sum_{t=2}^{n} X_{t-1}\varepsilon_t \xrightarrow{a.s} E(X_{t-1}\varepsilon_t)$. Since X_{t-1} is independent of ε_t, then $E(X_{t-1}\varepsilon_t)$ = 0. Hence, $\frac{1}{n}\sum_{t=2}^{n} X_{t-1}\varepsilon_t \xrightarrow{a.s} 0$. Finally, (2) is approximated by $\tilde{\rho}_1 = \frac{\theta \overline{X^2} - \frac{\theta}{n} X_n^2}{\overline{X^2}}$. Thus, for $n \to \infty$ we obtain $\hat{\theta} \to \tilde{\rho}_1$. We see that $\tilde{\rho}_1$ equals to $\phi\left(\overline{X^2}\right)$ for the function $\phi(x) = \frac{\theta x - \frac{\theta}{n} X_n^2}{x}$. Since ϕ is continous and hence is measurable.

Meantime, the bootstrap version of $\hat{\theta}$, denoted by $\hat{\theta}^*$ can be obtained as follows [see, *e.g*. Efron dan Tibshirani (1986) and Bose (1988)]:

1. Define the residuals $\hat{\varepsilon}_t = X_t - \hat{\theta} X_{t-1}$ for $t = 2,3,\ldots,n.$
2. 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 dan $X_t^* = \hat{\theta} X_{t-1}^* + \varepsilon_t^*$, $t = 2,3,\ldots,n$.
3. Finally, after centering the bootstrap time series $X_1^*, X_2^*, \ldots, X_n^*$ *i.e.* X_i^* is replaced by $X_i^* - \overline{X}^*$ where $\overline{X}^* = \frac{1}{n}\sum_{t=1}^{n} X_t^*$. Using the *plug-in* principle, we obtain the bootstrap estimator $\hat{\theta}^* = \hat{\rho}_1^* = \frac{\sum_{t=2}^{n} X_{t-1}^* X_t^*}{\sum_{t=1}^{n} X_t^{*2}}$ computed from the sample $X_1^*, X_2^*, \ldots, X_n^*$.

Analog with the previous discussion, we obtain the bootstrap version for counterpart of $\tilde{\rho}_1$, that is measurable map $\tilde{\rho}_1^* = \frac{\theta \overline{X^{*2}} - \frac{\theta}{n} X_n^{*2}}{\overline{X^{*2}}}$. Thus, according to Theorem 1 we conclude that $\tilde{\rho}_1^*$ converges to $\tilde{\rho}_1$ conditionally almost surely. Furthermore, $\sqrt{n}\left(\tilde{\rho}_1^* - \tilde{\rho}_1\right) \xrightarrow{d} T$ and for $n \to \infty$ we obtain

$\sqrt{n}\left(\hat{\rho}_1^* - \hat{\rho}_1\right) \xrightarrow{d} T$ where T is a normal distribution with zero mean and variance $\alpha_4 - \alpha_2^2$ with α_2 and α_4 are second and fourth moments repectively.

4. Results of Monte Carlo Simulations

The simulation is conducted using S-Pus and the sample is the 50 time series data of exchange rate of US dollar compared to Indonesian rupiah. Data is taken from authorized website of Bank Indonesia, i.e. http://www.bi.go.id for fifty days of transactions on March and April 2010. Suprihatin, *et. al*. (2011) has identified that the time series data satisfies the AR(1) procces, such that the data follows the equation

$$X_t = \theta X_{t-1} + \varepsilon_t,\ t = 2,3,\ldots,50,$$

where $\varepsilon_t \sim \mathrm{WN}\left(0,\sigma^2\right)$. The simulation yields the estimator for parameter θ is $\hat{\theta}$ = -0,448 with standard error 0,1999. To produce a good approximation, Efron and Tibshirani (1986) and Davison and Hinkley (2006) suggest to use the number of resamples at least B = 50. Bootstrap version of standard errror using bootstrap samples of size B = 25, 50, 100, 200, 500 and 1000 yielding as presented in Table 1.

Table 1 Estimates for Standard Errors of $\hat{\theta}^*$ for Various B

	B					
	25	50	100	200	500	1000
$se_{\hat{F}}\left(\hat{\theta}^*\right)$	0,2005	0,1981	0,1997	0,1991	0,1972	1,1964

From Table 1 we can see that the values of bootstrap standard errors tend to decrease in term of size of B increase and closed to the value of 0,1999 (actual standard error). These results show that the bootstrap gives a good estimate. Meantime, the histogram and density estimate of $\hat{\theta}^*$ are presented in Figure 1. From Figure 1 we can see that the resulting histogram close related to the normal density. Of course, this result agree to the result of Freedman (1985) and Bose (1988).

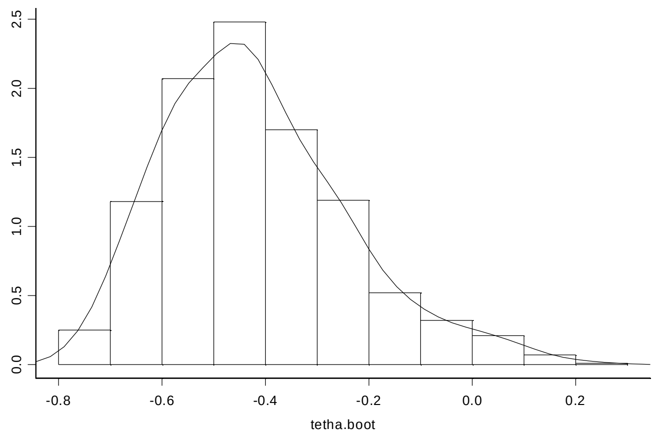

Figure 1 Histogram and Density Estimate of Bootstrap Estimator $\hat{\theta}^*$

5. Conclusions

A number of points arise from the study of Section 2, 3, and 4, amongst which we state as follows.

1. Consider an AR(1) process $X_t = \theta X_{t-1} + \varepsilon_t$ with Yule-Walker estimator $\hat{\theta}$ = $\hat{\rho}_1$ is a consistent estimator for true parameter $\theta = \rho_1$. By using the delta method we have shown that $\tilde{\rho}_1^*$ is also a consistent estimator for $\tilde{\rho}_1$ where $\hat{\theta} \to \tilde{\rho}_1$ for $n \to \infty$. Moreover, we obtain that $\sqrt{n}\left(\tilde{\rho}_1^* - \tilde{\rho}_1\right) \xrightarrow{d} N$ and for $n \to \infty$ the crux result is that $\sqrt{n}\left(\hat{\rho}_1^* - \hat{\rho}_1\right) \xrightarrow{d} N$ where N is a normal distribution.
2. Resulting of Monte Carlo simulations show that the bootstrap esstimators are good approximations, as represented by their standard errors and plot of densities estimation.

References

Bickel, P. J. and Freedman, D. A. (1981) "Some asymptotic theory for the bootstrap," *Ann. Statist.*, **9**, 1996-1217.

Bose, A. (1988) "Edgeworth correction by bootstrap in autoregressions," *Ann. Statist.*, **16**, 1709-1722.

Brockwell, P. J. and Davis, R. A. (1991) *Time Series: Theory and Methods,* Springer, New York.

DasGupta, A. (2008) *Asymptotic Theory of Statistics and Probability,* Springer, New York.

Davison, A. C. and Hinkley, D. V. (2006) *Bootstrap Methods and Their Application*, Cambridge University Press, Cambridge.

Efron, B. and Tibshirani, R. (1986) "Bootstrap methods for standard errors, confidence intervals, and others measures of statistical accuracy," *Statistical Science*, **1**, 54-77.

Freedman, D. A. (1985) "On bootstrapping two-stage least-squares estimates in stationary linear models," *Ann. Statist.*, **12**, 827-842.

Hall, P. (1992) *The Bootstrap and Edgeworth Expansion*, Springer-Verlag, New York.

Serfling, R. J. (1980) *Approximation Theorems of Mathematical Statistics,* John Wiley & Sons, New York.

Singh, K. (1981) "On the asymptotic accuracy of Efron's bootstrap," *Ann. Statist.*, **9**, 1187-1195.

Suprihatin, B., Guritno, S., and Haryatmi, S. (2011) "Consistency of the bootstrap estimator for mean under kolmogorov metric and its implementation on delta method," *Proc. of The* 6th *Seams-GMU Conference*.

van der Vaart, A. W. (2000) *Asymptotic Statistics,* Cambridge University Press, Cambridge.

Wei, W. W. S. (1990) *Time Series Analysis: Univariate and Multivariate Methods*, Addison Wesley, California.