Turnitin_Paper SeamsGMU_Bambang S

by Bambang Suprihatin

CONSISTENCY OF THE BOOTSTRAP ESTIMATOR FOR MEAN UNDER KOLMOGOROV METRIC AND ITS IMPLEMENTATION ON DELTA METHOD

[1]Bambang Suprihatin, Suryo Guritno, Sri Haryatmi

Mathematics Department, University of Sriwijaya
Email: bambangs@unsri.ac.id

Abstract. It is known that by Strong Law of Large Number, the sample mean $\overline{X}$ converges almost surely to the population sample μ. Central Limit Theorem asserts that the distribution of $\sqrt{n}\left(\overline{X}-\mu\right)$ converges to Normal distribution with mean *0* and variance σ^2 as $n \to \infty$. 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 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. Thereupon some authors call $\sqrt{n}\left(\overline{X}-\mu\right)$ as pivotal statistic. Here are two out of some ways in proving the consistency of bootstrap estimator. Firstly, the consistency was under Mallow-Wasserstein metric was studied by Bickel and Freedman [2]. The other consistency is using Kolmogorov metric, which is a part of paper in Singh [9]. In our paper, we investigate the consistency of mean under Kolmogorov metric comprehensively and use this result to study the consistency of bootstrap variance using delta Method. The accuracy of the bootstrap estimator using Edgeworth expansion is discussed as well. Results of simulations show that the bootstrap gives good estmates of standard error, which agree to the theory. All results of Monte Carlo simulations are also presented in regard to yield apparent conclusions.

Keywords and phrases: Bootstrap, consistency, Kolmogorov metric, delta method, Edgeworth expansion, Monte Carlo simulations

1. INTRODUCTION

Some questions are usually arise in study of estimation of the unknown parameter θ involves the estimation: (1) what estimator $\hat{\theta}$ should be used or choosen? (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 θ? The bootstrap is a general methodology

---

2010 Mathematics Subject Classification: 62F40

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 $\sqrt{n}\left(\overline{X}^* - \overline{X}\right)$. We will investigate the consistency of bootstrap under Kolmogorov metric which is defined as

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

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

2. CONSISTENCY OF BOOTSTRAP ESTIMATOR FOR MEAN

Let $\left(X_1, X_2, \ldots, X_n\right)$ be a random sample of size *n* from a population with common distribution F and let $T\left(X_1, X_2, \ldots, X_n; F\right)$ 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 $\left(X_1, X_2, \ldots, X_n\right)$, 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 $T\left(X_1, X_2, \ldots, X_n; F\right)$ under F by that of $T\left(X_1^*, X_2^*, \ldots, X_n^*; F_n\right)$ under F_n whrere $\left(X_1^*, X_2^*, \ldots, X_n^*\right)$ 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 $\boldsymbol{X}$. Let $\rho(F,G)$ be a metric on the space of distribution on $\boldsymbol{X}$. 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

$$H_n(x) = P_F(T(X_1, X_2, \ldots, X_n; F) \le x),$$

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

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

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

We state some theorems and lemma which are needed to show that $K(H_n, H_{Boot}) \to 0 \; a.s.$

Theorem 1 (KHINTCHINE-KOLMOGOROV CONVERGENCE THEOREM) *Suppose* $X_1, X_2, \ldots$ *are independent with mean* 0 *such that* $\sum_n \operatorname{var}(X_n) < \infty$. *Then,* $\sum_n X_n < \infty$ *a.s., i.e.* $S_n = \sum_{i=1}^n X_i$ *converges a.s. to* $\sum_{n=1}^{\infty} X_n$.

Kronecker Lemma *Suppose* $a_n > 0$ *and* $a_n \uparrow \infty$. *Then* $\sum_n X_n / a_n < \infty$ *implies* $\sum_{j=1}^n X_j / a_n \to 0$.

Proof. Set $b_n = \sum_{i=1}^n X_i / a_i$ and $a_0 = b_0 = 0$. Then, $b_n \to b_\infty < \infty$ and $X_n = a_n(b_n - b_{n-1})$. Write

$$\frac{1}{a_n}\sum_{j=1}^n X_j = \frac{1}{a_n}\sum_{j=1}^n a_j(b_j - b_{j-1}) = \frac{1}{a_n}\left(\sum_{j=1}^n a_j b_j - \sum_{j=1}^n a_j b_{j-1}\right)$$

$$= b_n + \frac{1}{a_n}\left(\sum_{j=1}^{n-1} a_j b_j - \sum_{j=1}^n a_j b_{j-1}\right) = b_n + \frac{1}{a_n}\left(\sum_{j=1}^n a_{j-1} b_{j-1} - \sum_{j=1}^n a_j b_{j-1}\right)$$

$$= b_n - \frac{1}{a_n}\sum_{j=1}^{n} b_{j-1}\left(a_j - a_{j-1}\right) \to b_\infty - b_\infty = 0. \qquad \square$$

Theorem 2 (POLYA'S THEOREM) *If* $F_n \xrightarrow{d} F$, *where* F *is a continuous distribution function, then* $\sup_x \left| F_n(x) - F(x) \right| \to 0$ *as* $n \to \infty$.

Theorem 3 (BERRY-ESSEN) *Let* $X_1, X_2, \ldots, X_n$ *be i.i.d. with* $E(X_1) = \mu$, $Var(X_1) = \sigma^2$, *and* $E\left|X_1 - \mu\right|^3 < \infty$. *Then there exists a universal constant* C, *not depending on* n *or the distribution of the* X_i, *such that*

$$\sup_x \left| P\left(\frac{\sqrt{n}\left(\overline{X} - \mu\right)}{\sigma} \le x \right) - \Phi(x) \right| \le \frac{C \cdot E\left|X_1 - \mu\right|^3}{\sigma^3 \sqrt{n}}.$$

Theorem 4 (ZYGMUND-MARCINKIEWICZ SLLN) *Suppose* $X, X_1, X_2, \ldots$ *are i.i.d. and* $E\left(\left|X\right|^p\right) < \infty$ *for some* $0 < p < 1$. *Then,* $\frac{S_n}{n^{1/p}} \to 0$ *a.s.*

Proof. This is consequence of the corrolary following Theorem 1 and Kronecker lemma, as desired. $\square$

Now we show the consistency of H_{Boot} under Kolmogorov metric, which is based on Sigh [9] and DasGupta [4]. We can write that

$$H_{Boot} K\left(H_n, H_{Boot}\right)$$

$$= \sup_x \left| P_F\left(T_n \le x\right) - P_*\left(T_n^* \le x\right) \right|$$

$$= \sup_x \left| P_F\left(\frac{T_n}{\sigma} \le \frac{x}{\sigma} \right) - P_*\left(\frac{T_n^*}{s} \le \frac{x}{s} \right) \right|$$

$$= \sup_x \left| P_F\left(\frac{T_n}{\sigma} \le \frac{x}{\sigma} \right) - \Phi\left(\frac{x}{\sigma}\right) + \Phi\left(\frac{x}{\sigma}\right) - \Phi\left(\frac{x}{s}\right) + \Phi\left(\frac{x}{s}\right) - P_*\left(\frac{T_n^*}{s} \le \frac{x}{s} \right) \right|$$

$$\le \sup_x \left| P_F\left(\frac{T_n}{\sigma} \le \frac{x}{\sigma} \right) - \Phi\left(\frac{x}{\sigma}\right) \right| + \sup_x \left| \Phi\left(\frac{x}{\sigma}\right) - \Phi\left(\frac{x}{s}\right) \right| + \sup_x \left| \Phi\left(\frac{x}{s}\right) - P_*\left(\frac{T_n^*}{s} \le \frac{x}{s} \right) \right|$$

$= A_n + B_n + C_n$, say.

By Polya's theorem, we conclude that $A_n \to 0$. Also, by SLLN, we obtain $s^2 \to \sigma^2$ a.s., and by the continuous mapping theorem, $s \to \sigma$ a.s. Hence, we conclude that $B_n \to 0$ a.s. Finally, by the Berry-Essen theorem,

$$C_n \leq \frac{C \cdot E\left|X_1^* - \bar{X}_n\right|^3}{\sqrt{n}\left(\operatorname{var}_{F_n}\left(X_1^*\right)\right)^{3/2}} = \frac{C \cdot \sum_{i=1}^n \left|X_1 - \bar{X}_n\right|^3}{\sqrt{n} \cdot ns^3}$$

$$\leq \frac{C \cdot \left(\sum_{i=1}^n \left|X_1 - \mu\right|^3 + n\left|\mu - \bar{X}_n\right|^3\right)}{n^{3/2}s^3}$$

$$= \frac{C}{s^3}\left(\frac{1}{n^{3/2}}\sum_{i=1}^n \left|X_1 - \mu\right|^3 + \frac{\left|\mu - \bar{X}_n\right|^3}{\sqrt{n}}\right).$$

Since $\bar{X} \to \mu$, it is clear that $\frac{\left|\mu - \bar{X}_n\right|^3}{\sqrt{n}s^3} \to 0$ a.s. In the first term, let $Y_i = \left|X_1 - \mu\right|^3$ and take $p = 2/3$, by Zygmund-Marcinkiewicz SLLN yields

$$\frac{1}{n^{3/2}}\sum_{i=1}^n \left|X_1 - \mu\right|^3 = \frac{1}{n^{1/p}}\sum_{i=1}^n Y_i \to 0 \text{ a.s. as } n \to \infty.$$

Thus, $A_n + B_n + C_n \to 0$ a.s. and hence $K(H_n, H_{Boot}) \to 0$ a.s.

Since $\sqrt{n}(\bar{X} - \mu) \xrightarrow{d} N(0, \sigma^2)$ and $K(H_n, H_{Boot}) \to 0$ a.s we could infer that $\sqrt{n}(\bar{X}^* - \bar{X}) \xrightarrow{d} N(0, \sigma^{*2})$, where σ^{*2} is bootstrap version of σ^2. On the other hand, according to the terminology of bootstrap, we conclude that $\bar{X}^* \to \bar{X}$ almost surely as $\bar{X} \to \mu$ a.s. Moreover, by Theorem 2.7 of van der Vaart [10] we conclude the crux result i.e. $\bar{X}^* \to \mu$. Then, a question arises about the use the bootstrap whether the bootstrap has any advantages when a Central Limit Theorem is already available. For our case, suppose $T(X_1, X_2, \ldots, X_n; F) = \sqrt{n}(\bar{X} - \mu)$. Then $\sqrt{n}(\bar{X} - \mu) \xrightarrow{d} N(0, \sigma^2)$ and under Kolmogorov metric $K(H_n, H_{Boot}) \to 0$ almost surely. So, we have two approximations to $P_F\left(\sqrt{n}(\bar{X} - \mu) \leq x\right)$, i.e. $\Phi(x/\sigma)$ and $P_{F_n}\left(\sqrt{n}(\bar{X}^* - \bar{X}) \leq x\right)$. The bootstrap approximation is theoretically more accurate than the approximation provided by the Central Limit Theorem. This is caused by the fact that normal distribution is symmetric such that the Central Limit Theorem can not capture information about the skewness at the finite sample distribution of $T(X_1, X_2, \ldots, X_n; F)$, whereas the bootstrap approximation does so. Thus, the bootstrap can be used in correcting for skewness, as an Edgeworth expansion would do. Babu and Singh [1] discussed the accuracy of bootstrap using one term Edgeworth expansion. Hutson and Ernst [8] studied the exact bootstrap for mean and suggest the bootstrap for variance of an *L*-estmator.

Since $T = \sqrt{n}(\overline{X} - \mu) \xrightarrow{d} N(0, \sigma^2)$, then Edgeworth expansion for T is

$$H(x) = P(T \le x) = \Phi(x/\sigma) + n^{-1/2} p(x/\sigma)\phi(x/\sigma) + O_p(n^{-1}),$$

where Φ is the standard normal distribution function and p is polynomial with coefficients depending on cumulants of $\overline{X} - \mu$. In the comprehensive studies, Hall (1992) showed that $p(x)$ denotes a function whose Fourier-Stieltjes transform $\int_{-\infty}^{\infty} e^{itx} dp(x) = r(it) e^{-t^2/2}$, where $r(it)$ can be derived from Hermite's polynomials $r(-d/dx) = -H_n(x)\phi(x)$ and satisfies $H_n(x) = (-1)^n e^{-x^2/2} \dfrac{d^n}{dx^n} e^{-x^2/2}$. The bootstrap estimate of H admits an analogous expansion

$$\hat{H}(x) = P(T^* \le x \mid X) = \Phi(x/\hat{\sigma}) + n^{-1/2} \hat{p}(x/\hat{\sigma})\phi(x/\hat{\sigma}) + O_p(n^{-1}),$$

where $\hat{p}$ is obtained from p on replacing unknowns by their bootstrap estimate. According to Davison and Hinkley [5], the estimate in the coefficients of $\hat{p}$ are typically distant $O_p(n^{-1/2})$ from their respective value in p, and so $\hat{p} - p = O_p(n^{-1/2})$. Hall [7] also showed that $\hat{\sigma} - \sigma = O_p(n^{-1/2})$ whence $\hat{H}(x) - H(x) = \Phi(x/\hat{\sigma}) - \Phi(x/\sigma) + O_p(n^{-1})$. Thus we can deduce that $\Phi(x/\hat{\sigma}) - \Phi(x/\sigma)$ is generally of size $n^{-1/2}$ not n^{-1}. Hence, $P(T^* \le x \mid X) - P(T \le x) = O_p(n^{-1/2})$. Consistency of the bootstrap sample mean is useful to study the consistency for many other statistics, see e.g. van der Vaart [10] and Cheng and Huang [3].

3. CONSISTENCY OF BOOTSTRAP ESTIMATE FOR VARIACE 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 5 (DELTA METHOD FOR BOOTSTRAP) *Let $\phi:\Re^k\to\Re^m$ 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 $\sqrt{n}\left(\phi\left(\hat{\theta}_n\right)-\phi(\theta)\right)\xrightarrow{d}\phi_\theta'(T)$ and $\sqrt{n}\left(\phi\left(\hat{\theta}_n^*\right)-\phi\left(\hat{\theta}_n\right)\right)\xrightarrow{d}\phi_\theta'(T)$ conditionally almost surely.*

Let $\theta = \mu$ is the population mean, and then $\hat{\theta}_n=\bar{X}$ is the sample mean. The SLLN asserts that $\hat{\theta}_n\to\theta$ 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 for the unbiased sample variance using delta method. Again, the SLLN asserts that unbiased sample variance $s^2=\frac{1}{n-1}\sum_{i=1}^n\left(X_i-\bar{X}\right)^2$ converges almost surely to σ^2. Let $s^{2*}=\frac{1}{n-1}\sum_{i=1}^n\left(X_i^*-\bar{X}^*\right)^2$ is the bootstrap estimate for the sample variance, the counterpart of s^2. Set $s^{2*}=\frac{n}{n-1}\frac{\sum_{i=1}^n\left(X_i^*-\bar{X}^*\right)^2}{n}=\frac{n}{n-1}\left(\frac{\sum X_i^{*2}}{n}-\left(\frac{\sum X_i^*}{n}\right)^2\right)$. The question is the s^{2*} converges a.s. to s^2? We see that s^2 equals to $\phi\left(\bar{X},\overline{X^2}\right)$ and s^{2*} equals to $\phi\left(\overline{X^*},\overline{X^{*2}}\right)$ for the map $\phi(x,y)=\frac{n}{n-1}\left(y-x^2\right)$. Thus, according to Theorem

5 we conclude that s^{2*} converges to s^2 conditionally almost surely. Furthermore, $\sqrt{n}\left(s^{2*} - s^2\right) \xrightarrow{d} T$ where T is a normal distribution.

4. RESULTS OF MONTE CARLO SIMULATIONS

The simulation is conducted using S-Plus and the sample is twenty marks of statistics test for 20 students are taken as follows: 80, 90, 75, 50, 85, 85, 45, 65, 50, 95, 70, 90, 35, 45, 50, 75, 70, 95, 60, 70. It is obvious that sample mean $\overline{X}$ = 69.0 with standard error 18.4. Efron and Tibshirani [6] suggested to conduct simulations using at least B equals 50 for standard errors and that 1000 for confidence intervals due to give good approximations. Using the number of bootstrap samples B = 2000, the resulting of simulation gives $\overline{X}^*$ = 69.12 with estimate for standard error 18.1, which is a good approximation. Figure 1 depicts the densities estimation for the distribution of $\sqrt{n}\left(\overline{X}^* - \overline{X}\right)$ and $\sqrt{n}\left(s^{2*} - s^2\right)$, respectively. From the figure, we could infer that the distributions for both statistics are approximately normal.

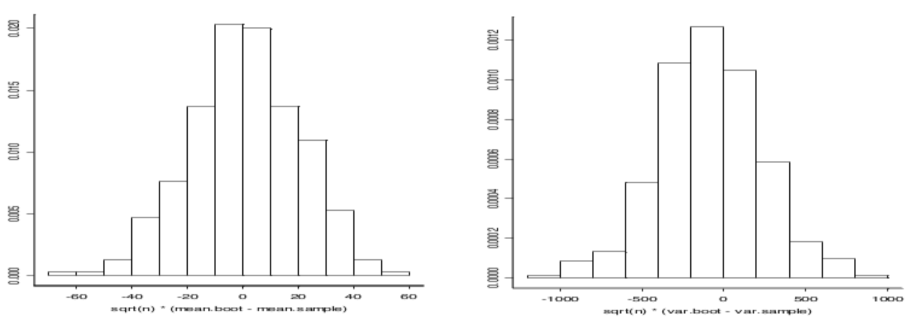

Figure 1 *Left panel:* Plot of Density Estimation for $\sqrt{n}\left(\overline{X}^* - \overline{X}\right)$, *Right panel:* Plot of Density Estimation for $\sqrt{n}\left(s^{2*} - s^2\right)$

5. CONCLUDING REMARK

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

1. Since $\overline{X} \to \mu$ a.s. and $\overline{X}^* \to \overline{X}$ a.s., according to the bootstrap terminology, we conclude that $\overline{X}^*$ is a consistent estimator for μ.
2. So far, by using delta method we have shown that unbiased bootstrap sample variance $s^{2*} \to s^2$ a.s., and it is obvious that for biased version $\hat{s}^{2*} = \frac{\sum_{i=1}^{n}\left(X_i^* - \overline{X}^*\right)^2}{n}$. Accordingly, both s^{2*} and $\hat{s}^{2*}$ are consistent estimators for σ^2.
3. Resulting of Monte Carlo simulation show that the bootstrap estimators are good approximations, as represented by their standard errors and plot of densities estimation.

REFERENCES

[1] BABU, G. J. AND SINGH, K. On one term Edgeworth correction by Efron's bootstrap, *Sankhya,* **46**, 219-232, 1984.

[2] BICKEL, P. J. AND FREEDMAN, D. A. Some asymptotic theory for the bootstrap, *Ann. Statist.,* **9**, 1996-1217, 1981.

[3] CHENG, G. AND HUANG, J. Z. Bootstrap consistency for general semiparametric M-estimation, *Ann. Statist.,* **5**, 2884-2915, 2010.

[4] DASGUPTA, A. *Asymptotic Theory of Statistics and Probability,* Springer, New York, 2008.

[5] DAVISON, A. C. AND HINKLEY, D. V. *Bootstrap Methods and Their Application,* Cambridge University Press, Cambridge, 2006.

[6] EFRON, B. AND TIBSHIRANI, R. Bootstrap methods for standard errors, confidence intervals, and others measures of statistical accuracy, *Statistical Science*, **1**, 54-77, 1986.

[7] HALL, P. *The Bootstrap and Edgeworth Expansion*, Springer-Verlag, New York, 1992.

[8] HUTSON, A. D. AND ERNST, M. D. The exact bootstrap mean and variance of an L-estimator, *J. R. Statist. Soc,* **62**, 89-94, 2000.

[9] SINGH, K. On the asymptotic accuracy of Efron's bootstrap, *Ann. Statist.,* **9**, 1187-1195, 1981.

[10] VAN DER VAART, A. W. *Asymptotic Statistics,* Cambridge University Press, Cambridge, 2000.

AUTHORS:

BAMBANG SUPRIHATIN: University of Sriwijaya, email: bambangs@unsri.ac.id,
SURYO GURITNO: University of Gadjahmada, email: Guritno0@mailcity.com,
SRI HARYATMI: University of Gadjahmada, email: s_kartiko@yahoo.com

Turnitin_Paper SeamsGMU_Bambang S

ORIGINALITY REPORT

21%	20%	11%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

3%

★ usermanual.wiki

Internet Source

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		