Southeast Asian Bulletin of Mathematics (2010) 34: 43–50

Southeast Asian
Bulletin of
Mathematics

Asymptotic Probability for Weighted Deviations of Dependent Bootstrap Means from the Sample Mean

Kamon Budsaba* and Tipaval Phatthanangkul
Mathematics and Statistics Department, Thammasat University Rangsit Center, Pathumthani 12121, Thailand.
Email: kamon@mathstat.sci.tu.ac.th; tipaval@mathstat.sci.tu.ac.th

Jiraroj Tosasukul
Department of Mathematics, Naresuan University, Phitsanulok 65000, Thailand.
Email: nuunji@hotmail.com

Andrei Volodin
Department of Mathematics and Statistics, University of Regina Regina, Saskatchewan, Canada S4S 0A2
Email: andrei@math.uregina.ca

Received 31 July 2007
Accepted 15 June 2008

Communicated by Y. Kemprasit

AMS Mathematics Subject Classification(2000): 60F05, 60F25, 60G42

Abstract. In this paper, the asymptotic probability for the weighted deviations of dependent bootstrap means from the sample mean is obtained, without imposing any assumptions on joint distribution of the original sequence of random variables from which the dependent bootstrap sample is selected. A nonrestrictive assumption of stochastic domination by a random variable is imposed on the marginal distributions of this sequence.

Keywords: Dependent bootstrap; Bootstrap means; Asymptotic probability for deviations; Exponential inequalities; Weighted sums; Weak law of large numbers; Stochastic domination.

*Corresponding author. Fax: +66.02.564-4489

1. Introduction

The main focus of the present investigation is to obtain asymptotic results for the probability of the weighted deviations of dependent bootstrap means from the sample mean.

Work on the validity of bootstrap estimators has received much attention in recent years due to a growing demand for the procedure, both theoretically and practically. As is mentioned in Mikosch [3], the sample mean is fundamental for parameter estimation in statistics. Therefore, most of the recent literature on the bootstrap is devoted to statistics of this type. This literature is mainly concerned with bootstrap validity; that is, with showing that a statistic and its bootstrap version have the similar asymptotic distributional behaviour. However, the limiting behaviour of bootstrap statistics is also of interest since it is by no means clear whether the bootstrap version of a consistent estimator is itself consistent.

We call the reader's attention to the special issue of the journal *Statistical Science* (2003) Volume 18, Number 2 devoted to the Silver Anniversary of the Bootstrap, where the wide applications of the bootstrap procedure to diverse areas of statistics are discussed. We also refer the reader to the recent expository paper by Csörgő and Rosalsky [2] where a detailed and comprehensive survey of limit laws for bootstrap sums is given.

The notion of the dependent bootstrap procedure was introduced by Smith and Taylor [4] where some important properties were also established. Let $\{X_n, n \geq 1\}$ be a sequence of random variables (which are not necessarily independent or identically distributed) defined on a probability space (Ω, F, P). Let $\{m(n), n \geq 1\}$ and $\{k(n), n \geq 1\}$ be two sequences of positive integers such that $m(n) \leq nk(n)$ for all $n \geq 1$. For $\omega \in \Omega$ and $n \geq 1$, the *dependent bootstrap* is defined to be the sample of size $m(n)$, denoted $\{\hat{X}^{(\omega)}_{n,j}, 1 \leq j \leq m(n)\}$, drawn without replacement from the collection of $nk(n)$ items made up of $k(n)$ copies each of the sample observations $X_1(\omega), \cdots, X_n(\omega)$.

This dependent bootstrap procedure is proposed as a procedure to reduce variation of estimators and to obtain better confidence intervals. We refer to Smith and Taylor [5] for details and where simulated confidence intervals are obtained to examine possible gains in coverage probabilities and interval lengths.

We may consider the dependent bootstrap procedure as a more general procedure than the classical Efron independent bootstrap. If we take $k(n) = \infty$ for all $n \geq 1$, then the dependent bootstrap reduces to the classical Efron independent bootstrap. The main results presented in this paper do not require any assumptions on $k(n)$; they are certainly true for the independent bootstrap as well.

Henceforth we let $\{\hat{X}^{(\omega)}_{n,j}, 1 \leq j \leq m(n)\}$ denote the dependent bootstrap sample from $X_1, \cdots, X_n$.

We refer the reader to the paper Ahmed *et al.* [1] where some important general properties of the dependent bootstrap are presented and a discussion of results in the literature pertaining to the classical (independent) and dependent bootstrap of the mean is given.

The main focus of Ahmed *et al.* [1] was to obtain asymptotic results for the following probability

$$P\left\{\frac{1}{n^{1/\alpha}}\sum_{j=1}^{m(n)}\left(\hat{X}_{n,j}^{(\omega)}-\overline{X}_n(\omega)\right)\geq\epsilon\right\}$$

as $n\to\infty$ where $\epsilon>0$, $0<\alpha<2$. The objective of the investigation resulting in the present paper is to extend the results of Ahmed *et al.* [1] to weighted deviations of the dependent bootstrap means from the sample mean. That is, the main focus of this paper is to obtain asymptotic results for the following probability

$$P\left\{\frac{1}{b_n}\sum_{j=1}^{m(n)}a_j\left(\hat{X}_{n,j}^{(\omega)}-\overline{X}_n(\omega)\right)\geq\epsilon\right\}$$

as $n\to\infty$ where $\epsilon>0$, $\{a_n,n\geq 1\}$ and $\{b_n,n\geq 1\}$ are two sequences of numbers such that $0<b_n\uparrow\infty$, and $\{\hat{X}_{n,j}^{(\omega)},1\leq j\leq m(n)\}$ is the dependent bootstrap sample from $X_1,\cdots,X_n$.

The following notion is well known. We recall that a sequence of random variables $\{X_n,\ n\geq 1\}$ is *stochastically dominated* by a random variable X if there exists a constant $D>0$ such that $P\{|X_n|>t\}\leq DP\{D|X|>t\}$ for all $t\geq 0$ and all $n\geq 1$.

2. Exponential Inequality

The exponential inequality presented in the lemma below is the key tool used in establishing the asymptotic probability for the weighted deviations of dependent bootstrap means from the sample mean that will be presented in the next section. It is a dependent bootstrap analog of the Mikosch exponential inequality (Mikosch [3], Lemma 5.1). We mention that this result was proved by Mikosch [3] under the assumption that $\{X_n,n\geq 1\}$ is a sequence of i.i.d. random variables, for supremum (not partial sums) of bootstrap random variables, and for the independent bootstrap procedure. Moreover, this result was proved in Ahmed *et al.* [1] Lemma 6 for normed sums (not weighted sums). Also, in Lemma 6 of Ahmed *et al.* [1], $c(n)\equiv m(n)$, while in our result $\{c(n),n\geq 1\}$ may be arbitrary positive sequence of constants.

Lemma 1. *Let $\{a_n,n\geq 1\}$, $\{b_n,n\geq 1\}$, $\{c(n),n\geq 1\}$, and $\{h_n,n\geq 1\}$ be sequences of nonnegative real numbers and $\kappa(n)=\max_{1\leq j\leq m(n)}a_j,n\geq 1$. Let $\{X_n,n\geq 1\}$ be a sequence of (not necessary independent or identically distributed) random variables. Then for $\omega\in\Omega$ and $n\geq 1$ such that $h_n\kappa(n)M_n(\omega)<$*

1, *the following inequality holds for all* $\epsilon > 0$*:*

$$P\left\{\sum_{j=1}^{m(n)} a_j \left|\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right| \geq \epsilon b_n\right\}$$
$$\leq 2\exp\left\{-\epsilon\frac{h_n b_n}{c(n)} + \frac{m(n)h_n^2\kappa(n)^2 B_n(\omega)}{2c(n)\left(1 - h_n\kappa(n)M_n(\omega)\right)}\right\},$$

where $M_n(\omega) = \frac{1}{c(n)}\max_{1\leq i\leq n}|X_i(\omega) - \overline{X}_n(\omega)|$ *and* $B_n(\omega) = \frac{1}{nc(n)}\sum_{i=1}^{n}(X_i(\omega) - \overline{X}_n(\omega))^2$.

Proof. By the Markov inequality,

$$P\left\{\left|\sum_{j=1}^{m(n)} a_j\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right| \geq \epsilon b_n\right\}$$
$$= P\left\{\frac{\left|\sum_{j=1}^{m(n)} a_j\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right|}{c(n)} \geq \frac{\epsilon b_n}{c(n)}\right\}$$
$$\leq \exp\left\{-\frac{\epsilon h_n b_n}{c(n)}\right\} E\exp\left\{\frac{h_n}{c(n)}\left|\sum_{j=1}^{m(n)} a_j\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right|\right\}$$
$$\leq \exp\left\{-\frac{\epsilon h_n b_n}{c(n)}\right\} E\exp\left\{\frac{h_n}{c(n)}\left(\sum_{j=1}^{m(n)} a_j\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right)\right\}$$
$$+\exp\left\{-\frac{\epsilon h_n b_n}{c(n)}\right\} E\exp\left\{-\frac{h_n}{c(n)}\left(\sum_{j=1}^{m(n)} a_j\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right)\right\}.$$

We will estimate only the expectation in the first term of the last expression; the same bound is valid for the second expectation.

Now by Proposition 2 from Ahmed *et al.* [1] the dependent bootstrap random variables $\{\hat{X}_{n,j}^{(\omega)}, 1 \leq j \leq m(n)\}, n \geq 1$ are negatively dependent and exchangeable. Hence, by Lemma 1(1) from Ahmed *et al.* [1] the random variables

$$\left\{\exp\left\{\frac{h_n a_j}{m(n)}\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right\}, 1 \leq j \leq m(n)\right\}$$

are negatively dependent.

Therefore

$$E\exp\left\{\frac{h_n}{c(n)}\left(\sum_{j=1}^{m(n)} a_j\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right)\right\}$$

$$
\begin{aligned}
&= E\left[\prod_{j=1}^{m(n)} \exp\left\{\frac{h_n a_j}{c(n)}\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right\}\right] \\
&\le \prod_{j=1}^{m(n)} E\exp\left\{\frac{h_n a_j}{c(n)}\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right\} \quad \text{(by [1, Lemma 1(2)])} \\
&= \prod_{j=1}^{m(n)} \left[E\exp\left\{\frac{h_n a_j}{c(n)}\left(\hat{X}_{n,1}^{(\omega)} - \overline{X}_n(\omega)\right)\right\}\right] \ \text{(by identical distribution)} \\
&= \prod_{j=1}^{m(n)} \left[\frac{1}{n}\sum_{i=1}^{n} \exp\left\{\frac{h_n a_j}{c(n)}\left(X_i(\omega) - \overline{X}_n(\omega)\right)\right\}\right] \\
&= \prod_{j=1}^{m(n)} \left[\frac{1}{n}\sum_{i=1}^{n} \exp\left\{\frac{h_n a_j}{c(n)}\left|X_i(\omega) - \overline{X}_n(\omega)\right|\right\}\right] \\
&\le \prod_{j=1}^{m(n)} \left[\frac{1}{n}\sum_{i=1}^{n} \exp\left\{\frac{h_n \kappa(n)}{c(n)}\left|X_i(\omega) - \overline{X}_n(\omega)\right|\right\}\right] \\
&= \left[\frac{1}{n}\sum_{i=1}^{n} \exp\left\{\frac{h_n \kappa(n)}{c(n)}\left|X_i(\omega) - \overline{X}_n(\omega)\right|\right\}\right]^{m(n)} \\
&= \left[1 + \frac{1}{n}\sum_{i=1}^{n}\left(\frac{h_n^2\kappa(n)^2}{2!c(n)^2}\left|X_i(\omega) - \overline{X}_n(\omega)\right|^2 + \frac{h_n^3\kappa(n)^3}{3!c(n)^3}\left|X_i(\omega) - \overline{X}_n(\omega)\right|^3\right.\right. \\
&\qquad \left.\left. + \ \frac{h_n^4\kappa(n)^4}{4!c(n)^4}\left|X_i(\omega) - \overline{X}_n(\omega)\right|^4 + \cdots\right)\right]^{m(n)} \\
&= \left[1 + \frac{h_n^2\kappa(n)^2}{c(n)}\sum_{i=1}^{n}\frac{\left|X_i(\omega) - \overline{X}_n(\omega)\right|^2}{nc(n)}\left[\frac{1}{2!} + \frac{h_n\kappa(n)}{3!}\ \left|\frac{X_i(\omega) - \overline{X}_n(\omega)}{c(n)}\right|\right.\right. \\
&\qquad \left.\left. + \ \frac{h_n^2\kappa(n)^2}{4!}\left|\frac{X_i(\omega) - \overline{X}_n(\omega)}{c(n)}\right|^2 + \cdots\right]\right]^{m(n)} \\
&\le \left[1 + \frac{h_n^2\kappa(n)^2}{c(n)}\frac{B_n(\omega)}{2}\left(1 + h_n\kappa(n)M_n(\omega) + \left(h_n\kappa(n)M_n(\omega)\right)^2 + \cdots\right)\right]^{m(n)} \\
&= \left[1 + \frac{h_n^2\kappa(n)^2}{2c(n)}\frac{B_n(\omega)}{\left(1 - h_n\kappa(n)M_n(\omega)\right)}\right]^{m(n)} \quad \text{(since } h_n\kappa(n)M_n(\omega) < 1\text{)} \\
&\le \left[\exp\left\{\frac{h_n^2\kappa(n)^2}{2c(n)}\frac{B_n(\omega)}{\left(1 - h_n\kappa(n)M_n(\omega)\right)}\right\}\right]^{m(n)} \\
&= \exp\left\{\frac{m(n)h_n^2\kappa(n)^2 B_n(\omega)}{2c(n)\left(1 - h_n\kappa(n)M_n(\omega)\right)}\right\}.
\end{aligned}
$$

Hence,

$$P\left\{\sum_{j=1}^{m(n)} a_j \left|\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right| \geq \epsilon b_n\right\}$$
$$\leq 2\exp\left\{-\epsilon\frac{h_n b_n}{c(n)} + \frac{m(n)h_n^2\kappa(n)^2 B_n(\omega)}{2c(n)\left(1 - h_n\kappa(n)M_n(\omega)\right)}\right\}. \qquad \blacksquare$$

3. The Main Result

With the preliminaries accounted for, we can formulate and prove the main result of this paper, that is the asymptotic probability for the weighted deviations of dependent bootstrap means from the sample mean. We emphasize that there are no independence or identical distribution assumptions on the original sequence of random variables $\{X_n,\ n \geq\ 1\}$.

Theorem 2. *Let $\psi(t),\ t \geq 0$ be an increasing function such that*

$$\sum_{j=n}^{\infty} \frac{1}{(\psi^{-1}(j))^2} = \mathcal{O}\left(\frac{n}{(\psi^{-1}(n))^2}\right),\ n \geq 1, \tag{1}$$

where $\psi^{-1}(t)$ is the inverse function of $\psi(t)$. Let $\{a_n, n \geq 1\}$ and $\{b_n, n \geq 1\}$ be sequences of nonnegative real numbers and $\kappa(n) = \max_{1\leq j\leq m(n)} a_j, n \geq 1$. Let moreover $\{X_n, n \geq 1\}$ be a sequence of random variables which is stochastically dominated by a random variables X such that $E\psi(DX) < \infty$ for all $D > 0$. Then for almost every $\omega \in \Omega$, for every $\epsilon > 0$, and for every positive number r,

$$P\left\{\left|\sum_{j=1}^{m(n)} a_j \left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right| \geq \epsilon b_n\right\}$$
$$= \mathcal{O}\left(\exp\left\{-r\frac{b_n}{\psi^{-1}(n)\kappa(n)} + \frac{m(n)}{n}o(1)\right\}\right).$$

Proof. Fix the arbitrary constants $r > 0$ and $\epsilon > 0$ and let $h_n = \frac{rc(n)}{\epsilon\psi^{-1}(n)\kappa(n)}, n \geq 1$. The fact that

$$h_n M_n \leq \frac{r}{\epsilon}\ \frac{2}{\psi^{-1}(n)} \max_{1\leq j\leq n} |X_j| \to 0 \text{ a.s.}$$

follows directly from Lemma 3 of [1].

Next in Lemma 2 of [1], we consider $Y_j = X_j^2, Y = X^2$, and $\phi(t) = \psi(\sqrt{t})$. Then $\phi^{-1}(n) = (\psi^{-1}(n))^2$ and $E\phi(Y) = E\psi(X) < \infty$. By Lemma 2 of Ahmed *et al.* [1],

$$h_n^2 B_n = \frac{r^2}{\epsilon^2}\ \frac{m(n)}{n} \frac{1}{(\psi^{-1}(n))^2} \sum_{j=1}^{n} X_j^2 = \frac{m(n)}{n} o(1) \text{ a.s.}$$

Hence,

$$\frac{h_n^2 B_n}{2(1 - h_n M_n)} = \frac{m(n)}{n} o(1) \quad \text{a.s.}$$

We also note that

$$\epsilon \frac{h_n a_n}{m(n)} = r \frac{a_n}{\psi^{-1}(n)}, n \geq 1.$$

The result then follows directly from Lemma. ■

Remark 3. The conclusion of Theorem is of course stronger the larger r is taken. The constant r does not play a role in any assumptions and it can be taken to be arbitrarily large.

Using different moment assumptions, we can now derive different results on the asymptotic probability for the deviations of dependent bootstrap means from the sample mean. In all corollaries we assume that $m(n) = c(n) = n, b_n = \log n, a_j = \log j, n \geq 1, j \geq 1$.

Corollary 4. *Let $\{X_n, n \geq 1\}$ be a sequence of random variables which is stochastically dominated by a random variable X and let $0 < \alpha < 2$. If $E|X|^\alpha < \infty$, then for almost every $\omega \in \Omega$ and every $\epsilon > 0$*

$$P\left\{ \frac{1}{n^{1/\alpha} \log n} \left| \sum_{j=1}^{n} \log j \left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega) \right) \right| \geq \epsilon \right\} = o(1);$$

that is, for almost every $\epsilon > 0$ the weak law of large numbers

$$\frac{1}{n^{1/\alpha} \log n} \sum_{j=1}^{n} \log j \left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega) \right) \to 0 \quad \textit{in probability}$$

obtains.

Proof. Let $\psi(t) = t^\alpha, t > 0$. Then $\psi^{-1}(n) = n^{1/\alpha}, n \geq 1$. The relation (*) holds trivially since $2/\alpha > 1$. If we take $a_n = n^{1/\alpha}$ and $m(n) = n, n \geq 1$, then according to Theorem for every $\epsilon > 0$ and every $r > 0$ and for all sufficiently large n and some constant $C < \infty$,

$$P\left\{ \frac{1}{n^{1/\alpha} \log n} \left| \sum_{j=1}^{n} \log j \left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega) \right) \right| \geq \epsilon \right\} \leq C \exp\{-r\};$$

that is, since $r > 0$ is arbitrary

$$\frac{1}{n^{1/\alpha} \log n} \sum_{j=1}^{n} \log j \left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega) \right) \to 0 \quad \text{in probability.}$$

■

Corollary 5. *Let $\{X_n, n \geq 1\}$ be a sequence of random variables which is stochastically dominated by a random variable X and let $0 < \alpha < 2$. If $E|X|^\alpha |\log|X||^\alpha < \infty$, then for every $\epsilon > 0$, every real number r, and almost every $\omega \in \Omega$*

$$P\left\{\frac{1}{n^{1/\alpha}\log n}\left|\sum_{j=1}^{n}\log j\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right| \geq \epsilon\right\} = \mathcal{O}(n^{-r}).$$

Proof. Let $\psi(t) = t^\alpha \log^\alpha t, t \geq 1$. Then according to Lemma 4 from Ahmed *et al.* [1], the sequence $\psi^{-1}(n)$ is equivalent to $\frac{n^{1/\alpha}}{\log n}$, $n \geq 2$. The relation (*) holds by Lemma 5 Ahmed *et al.* [1] since $2/\alpha > 1$. For fixed r, $\epsilon > 0$, $m(n) = n$, and $a_n = n^{1/\alpha}, n \geq 1$, applying Theorem we obtain the result. ■

Corollary 6. *Let $\{X_n, n \geq 1\}$ be a sequence of random variables which is stochastically dominated by a random variable X and let $0 < \alpha < 2$. If $E|X|^\delta < \infty$ for some $\alpha < \delta < 2$, then for every $\epsilon > 0$, every r, and almost all $\omega \in \Omega$*

$$P\left\{\frac{1}{n^{1/\alpha}\log n}\left|\sum_{j=1}^{n}\log j\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right| \geq \epsilon\right\} = \mathcal{O}\left(\exp\{-rn^{\frac{1}{\alpha}-\frac{1}{\delta}}\}\right).$$

Proof. Let $\psi(t) = t^\delta, t > 0$, then $\psi^{-1}(n) = n^{1/\delta}$. The relation (*) holds trivially since $2/\delta > 1$. For fixed r, $\epsilon > 0$, $m(n) = n$, and $a_n = n^{1/\alpha}\log n, n \geq 1$, applying Theorem we obtain the result. ■

Acknowledgement. The authors are exceptionally grateful to the referee for offering helpful remarks and comments that improved presentation of the paper.

References

[1] S.E. Ahmed, D. Li, A. Rosalsky and A. Volodin, On the asymptotic probability for the deviations of dependent bootstrap means from the sample mean, *Lobachevskii J. Math.* **18** (2005) 3–20.

[2] S. Csörgő and A. Rosalsky, A survey of limit laws for bootstrapped sums, *Internat. J. Math. & Math. Sci.* **2003** 2835–2861 (2003).

[3] T. Mikosch, Almost sure convergence of bootstrapped means and U-statistics, *J. Statist. Plann. Inference* **41** (1994) 1–19.

[4] W.D. Smith and R.L. Taylor, Consistency of dependent bootstrap estimators, *Amer. J. Math. Management Sci.* **21** (2001) 359–382.

[5] W.D. Smith and R.L. Taylor, Dependent bootstrap confidence intervals, in: *Selected Proceedings of the Symposium on Inference for Stochastic Processes (Athens, GA, 2000). IMS Lecture Notes - Monograph Series* **37** (2001) 91–107.