Lobachevskii Journal of Mathematics
http://ljm.ksu.ru
Vol. 18, 2005, 3–20

S. Ejaz Ahmed, Deli Li, Andrew Rosalsky, and Andrei Volodin

ON THE ASYMPTOTIC PROBABILITY FOR THE DEVIATIONS OF DEPENDENT BOOTSTRAP MEANS FROM THE SAMPLE MEAN

(submitted by F. G. Avkhadiev)

Abstract. In this paper, the asymptotic probability for the deviations of dependent bootstrap means from the sample mean is obtained without imposing any conditions on the joint distributions associated with the original sequence of random variables from which the dependent bootstrap sample is selected. The mild condition of stochastic domination by a random variable is imposed on the marginal distributions of the random variables in this sequence.

Dedicated to Professor Mushtari
on the occasion of his 60th birthday

1. Introduction.

It is a great pleasure for us to contribute this paper in honour of Professor Daniar Khamidovich Mushtari on the occasion of his 60th birthday.

2000 Mathematical Subject Classification. 60F05, 60F25, 60G42.
Key words and phrases. Dependent bootstrap; Bootstrap means; Asymptotic probability for deviations; Exponential inequalities; Strong law of large numbers; Stochastic domination.
The work of A. Volodin is supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

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

The 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 (1994), 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. From our point of view, this explains the usefulness and impact on statistical inference of deviations from the sample means for "exogenously generated" bootstrap samples. Furthermore, asymptotic probabilities for the deviations of bootstrap means are a quite useful tool for the study of bootstrap moments. It is important to note that exponential inequalities are of practical use in establishing the strong asymptotic validity of bootstrap means.

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 begin with a brief discussion of results in the literature pertaining to a sequence of independent and identically distributed (i.i.d.) random variables and to the classical (independent) bootstrap of the mean. Let $\{X, X_n, n \geq 1\}$ be a sequence of i.i.d. random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. For $\omega \in \Omega$ and $n \geq 1$, let $P_n(\omega) = n^{-1} \sum_{i=1}^{n} \delta_{X_i(\omega)}$ denote the empirical measure and let $\{\hat{X}_{n,j}^{(\omega)}, 1 \leq j \leq m(n)\}$ be i.i.d. random variables with law $P_n(\omega)$ where $\{m(n), n \geq 1\}$ is a sequence of positive integers. In other words, the random variables $\{\hat{X}_{n,j}^{(\omega)}, 1 \leq j \leq m(n)\}$ result by sampling $m(n)$ times with replacement from the n observations $X_1(\omega), \cdots, X_n(\omega)$ such that for each of the $m(n)$ selections, each $X_j(\omega)$ has probability n^{-1} of being chosen.

For each $n \geq 1$, $\{\hat{X}_{n,j}^{(\omega)}, 1 \leq j \leq m(n)\}$ is the so-called Efron (1979) *independent bootstrap sample* from $X_1, \cdots, X_n$ with *bootstrap sample size* $m(n)$. Let $\overline{X}_n(\omega) = \frac{1}{n} \sum_{j=1}^{n} X_j(\omega)$ denote the sample mean of $\{X_j(\omega), 1 \leq j \leq n\}, n \geq 1$.

When X is nondegenerate and $EX^2 < \infty$, Bickel and Freedman (1981) showed that for almost every $\omega \in \Omega$ the central limit theorem (CLT)

$$n^{1/2}\left(\frac{1}{n}\sum_{j=1}^{n}\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right) \xrightarrow{d} N(0,\sigma^2)$$

obtains. Here and below and $\sigma^2 = \text{Var } X$. Note that by the Glivenko-Cantelli theorem $P_n(\omega)$ is close to $L(X)$ for almost every $\omega \in \Omega$ and all large n, and by the classical Lévy CLT

$$n^{1/2}\left(\frac{1}{n}\sum_{j=1}^{n}X_j - EX\right) \xrightarrow{d} N(0,\sigma^2).$$

It follows that for almost every $\omega \in \Omega$, the bootstrap statistic

$$n^{1/2}\left(\frac{1}{n}\sum_{j=1}^{n}\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)$$

is close in distribution to that of

$$n^{1/2}\left(\frac{1}{n}\sum_{j=1}^{n}X_j - EX\right)$$

for all large n. This is the basic idea behind the bootstrap. See the pioneering work of Efron (1979) where this nice idea is made explicit and where it is substantiated with several important examples.

A strong law of large numbers (SLLN) was first proved by Athreya (1983) for bootstrap means from the classical bootstrap. Arenal-Gutiérrez, Matrán, and Cuesta-Albertos (1996) analyzed the work of Athreya (1983) and, by taking into account different growth rates for the bootstrap sample size $m(n)$, they gave new and simple proofs of even more general results. They also provided examples that show that the sizes of resampling required by their results to ensure almost sure (a.s.) convergence are not far from optimal.

An article which is important for this paper is that of Mikosch (1994). He established a series of exponential inequalities (cf. Lemma 6 below) that are an important tool for deriving results on the consistency of the bootstrap mean. Based on these exponential inequalities, he proved an a.s. convergence result for bootstrap means (Theorem 1 below). Next, using the same exponential inequalities, the Baum - Katz / Erdös / Hsu - Robbins / Spitzer type complete convergence result for bootstrap means (Theorem 2 below) and a moment result for the supremum of normed bootstrap sums were established in Li, Rosalsky, and Ahmed (1999).

The following bootstrap counterpart to the Marcinkiewicz-Zygmund SLLN was obtained by Mikosch (1994), Proposition 3.3.

Theorem 1. *Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d. random variables and let $0 < \alpha < 2$. If $m(n) \equiv n$ and*

$$E|X_1|^\alpha \, |\log |X_1||^\alpha < \infty,$$

then for almost every $\omega \in \Omega$

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

where $\{\hat{X}_{n,j}^{(\omega)}, 1 \leq j \leq m(n)\}$ is Efron's independent bootstrap sample from $X_1, \cdots, X_n$.

We note that the classical Efron independent bootstrap sample $\{\hat{X}_{n,j}^{(\omega)}, \ 1 \leq j \leq m(n), \ n \geq 1\}$ can of course be defined in the same manner even if the original sequence $\{X_n, n \geq 1\}$ is not comprised of independent or identically distributed random variables. The following result was proved by Li, Rosalsky, and Ahmed (1999), Theorem 2.1 and Remark 2.4.

Theorem 2. *Let $\{X_n, n \geq 1\}$ be a sequence of pairwise i.i.d. random variables and let $0 < \alpha < 2$. If $m(n) \equiv n$ and*

$$E|X_1|^\alpha | \log |X_1||^\alpha < \infty,$$

then for every real number q, every $\epsilon > 0$, and almost every $\omega \in \Omega$

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

where $\{\hat{X}_{n,j}^{(\omega)}, 1 \leq j \leq m(n)\}$ is Efron's independent bootstrap sample from $X_1, \cdots, X_n$.

Remark 1. Taking $q = 0$, it follows from the Borel-Cantelli lemma and Theorem 2 that for almost every $\omega \in \Omega$,

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

We also refer the reader to the recent expository paper by Csörgő and Rosalsky (2003) 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 (2001a and 2001b) for a sequence of i.i.d. random variables where some important properties were also established. However, the dependent bootstrap procedure can be defined as follows for an arbitrary sequence of random variables. 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 $(\Omega, \mathcal{F}, P)$. 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 $\{\hat{X}_{n,j}^{(\omega)}, 1 \leq j \leq m(n)\}$, 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 (2001b) 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 $\{\hat{X}_{n,j}^{(\omega)}, 1 \leq j \leq m(n)\}$ denote the **dependent** bootstrap sample from $X_1, \cdots, X_n$.

The following result from Volodin, Ordóñez Cabrera, and Hu (2005) extends the above cited result of Li, Rosalsky, and Ahmed (1999) to the case of the dependent bootstrap. The content of Remark 1 also pertains to Theorem 3.

Theorem 3. *Let $\{X_n, n \geq 1\}$ be a sequence of identically distributed (not necessary independent) random variables and let $0 < \alpha < 2$. If*

$$E|X_1|^\alpha \left|\log |X_1|\right|^\alpha < \infty,$$

then for every real number q, every $\epsilon > 0$, and almost every $\omega \in \Omega$

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

The initial objective of the investigation resulting in the present paper was only to extend the results of Hu, Ordóñez Cabrera, and Volodin (2005) on the SLLN to the dependent bootstrap of the mean. But we are even able to establish in Theorem 4 a more general result than Theorem 3. Notice that Theorem 3 has the moment assumption $E|X_1|^\alpha |\log|X_1||^\alpha < \infty$ whereas our result has much more general moment assumption. The no independence condition in Theorem 3 is noteworthy and it also prevails in Theorem 4; the identical distributions assumption is relaxed in Theorem 4 to stochastic domination by a random variable.

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 $C > 0$ such that

$$P\{|X_n| > t\} \leq CP\{|X| > t\}$$

for all $t \geq 0$ and all $n \geq 1$.

The main focus of this paper is to obtain in Theorem 4 asymptotic results for

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

as $n \to \infty$ where $\epsilon > 0$, $0 < \alpha < 2$, and $\{\hat{X}_{n,j}^{(\omega)}, 1 \leq j \leq m(n)\}$ is the dependent bootstrap sample from $X_1, \cdots, X_n$. The sequence $\{X_n, n \geq \ 1\}$ is not necessary a sequence of independent random variables but it is assumed to be stochastically dominated.

2. Some general results on the dependent bootstrap

The results from this section are modifications, generalizations, or extensions of the results of Smith and Taylor (2001a and 2001b) for the dependent bootstrap from a sequence of random variables $\{X_n, n \geq 1\}$ which are not necessarily i.i.d. We note again that Smith and Taylor (2001a and 2001b) consider only the i.i.d. case. The results in this section are of general interest and play a role in establishing the asymptotic results discussed above.

The first proposition gives the joint distribution of the random variables in the dependent bootstrap sample. We need the following notation.

For $\omega \in \Omega, n \geq 1$, and a real number x, denote

$$\tau(x) = \sum_{j=1}^{n} I\{X_j(\omega) \leq x\} \text{ and } \mu(x) = \sum_{j=1}^{n} I\{X_j(\omega) > x\}$$

where $I(\cdot)$ is the indicator function. Hence, $\tau(x)$ is the random variable that counts the number of observations $X_j(\omega), 1 \leq j \leq n$, that are less than or equal to x, while $\mu(x)$ is the random variable that counts the number of observations $X_j(\omega), 1 \le j \le n$, that are strictly greater than x. Certainly, $\tau(x) + \mu(x) = n$ for every x.

For a finite sequence $\{x_1, x_2, \cdots, x_m\}$ of real numbers, denote

$$\{x_{(1)}, x_{(2)}, \cdots, x_{(m)}\}$$

its nondecreasing rearrangement, that is $x_{(1)} \le x_{(2)} \le \cdots \le x_{(m)}$ and for any $1 \le j \le m$ there exists $1 \le i \le m$ such that $x_i = x_{(j)}$.

Proposition 1. *For $\omega \in \Omega, n \ge 1$, and a sequence $\{x_1, x_2, \cdots, x_{m(n)}\}$ of real numbers:*
1) If $k(n)\tau(x_{(j)}) \ge j$ for all $1 \le j \le m(n)$, then

$$P\{\hat{X}^{(\omega)}_{n,1} \le x_1, \cdots, \hat{X}^{(\omega)}_{n,m(n)} \le x_{m(n)}\} = \prod_{j=1}^{m(n)} \frac{k(n)\tau(x_{(j)}) - (j-1)}{k(n)n - (j-1)}.$$

If $k(n)\tau(x_{(j)}) < j$ for at least one $1 \le j \le m(n)$, then the above probability is 0.
2) If $k(n)\mu(x_{(i)}) \ge m(n) - i$ for all $1 \le i \le m(n)$, then

$$P\{\hat{X}^{(\omega)}_{n,1} > x_1, \cdots, \hat{X}^{(\omega)}_{n,m(n)} > x_{m(n)}\} = \prod_{i=1}^{m(n)} \frac{k(n)\mu(x_{(i)}) - (m(n) - i + 1)}{k(n)n - (m(n) - i + 1)}.$$

If $k(n)\mu(x_{(i)}) < m(n) - i$ for at least one $1 \le i \le m(n)$, then the above probability is 0.

Proof. Let π be the reordering of $\{1, 2, \cdots, m(n)\}$ such that $\pi(j) = i$ for $x_i = x_{(j)}$.

For the proof of the first statement of Proposition 1, note that

$$\begin{aligned} & P\{\hat{X}^{(\omega)}_{n,1} \le x_1, \cdots, \hat{X}^{(\omega)}_{n,m(n)} \le x_{m(n)}\} \\ = \; & P\{\hat{X}^{(\omega)}_{n,\pi(1)} \le x_{(1)}, \cdots, \hat{X}^{(\omega)}_{n,\pi(m(n))} \le x_{(m(n))}\} \\ = \; & P\{\hat{X}^{(\omega)}_{n,\pi(1)} \le x_{(1)}\} \times P\{\hat{X}^{(\omega)}_{n,\pi(2)} \le x_{(2)} | \hat{X}^{(\omega)}_{n,\pi(1)} \le x_{((1))}\} \times \cdots \end{aligned}$$

$$\times P\{\hat{X}^{(\omega)}_{n,\pi(m(n))} \le x_{(m(n))} | \hat{X}^{(\omega)}_{n,\pi(1)} \le x_{(1)}, \cdots, \hat{X}^{(\omega)}_{n,\pi(m(n)-1)} \le x_{(m(n)-1)}\}$$

$$= \prod_{j=1}^{m(n)} \frac{k(n)\tau(x_{(j)}) - (j-1)}{k(n)n - (j-1)}$$

if $k(n)\tau(x_{(j)}) \ge j$ for all $1 \le j \le m(n)$. The second part of the first statement is obvious.

For the proof of the second statement of Proposition 1, note that

$$\begin{aligned}
& P\{\hat{X}^{(\omega)}_{n,1} > x_1, \cdots, \hat{X}^{(\omega)}_{n,m(n)} > x_{m(n)}\} \\
= \;& P\{\hat{X}^{(\omega)}_{n,\pi(m(n))} > x_{(m(n))}, \cdots, \hat{X}^{(\omega)}_{n,\pi(1)} > x_{(1)}\} \\
= \;& P\{\hat{X}^{(\omega)}_{n,\pi(m(n))} > x_{(m(n))}\}
\end{aligned}$$

$$\times P\{\hat{X}^{(\omega)}_{n,\pi(m(n)-1)} > x_{(m(n)-1)} | \hat{X}^{(\omega)}_{n,\pi(m(n))} > x_{((m(n)))}\} \times \cdots$$

$$\times P\{\hat{X}^{(\omega)}_{n,\pi(1)} > x_{(1)} | \hat{X}^{(\omega)}_{n,\pi(m(n))} > x_{(m(n))}, \cdots, \hat{X}^{(\omega)}_{n,\pi(2)} > x_{(2)}\}$$

$$= \prod_{j=1}^{m(n)} \frac{k(n)\mu(x_{(m(n)-j+1)}) - (j-1)}{k(n)n - (j-1)}$$

$$= \prod_{i=1}^{m(n)} \frac{k(n)\mu(x_{(i)}) - (m(n)-i+1)}{k(n)n - (m(n)-i+1)}$$

if $k(n)\mu(x_{(i)}) \geq m(n) - i$ for all $1 \leq i \leq m(n)$. The second part of the second statement is obvious.

Of course, the dependent bootstrap random variables $\{\hat{X}^{(\omega)}_{n,j},\ 1 \leq j \leq m(n)\}$ are indeed dependent. They obey the so-called negatively dependent property; this property will be established in Proposition 2. The concept of negatively dependent random variables was introduced by Lehmann (1966) as follows.

Random variables $Y_1, Y_2, \cdots$ are said to be *negatively dependent* if for each $n \geq 2$, the following two inequalities hold:

$$P\{Y_1 \leq y_1, \cdots, Y_n \leq y_n\} \leq \prod_{i=1}^{n} P\{Y_i \leq y_i\}$$

and

$$P\{Y_1 > y_1, \cdots, Y_n > y_n\} \leq \prod_{i=1}^{n} P\{Y_i > y_i\}$$

for every sequence $\{y_1, \cdots, y_n\}$ of real numbers.

Proposition 2. *For $\omega \in \Omega$ and $n \geq 1$, the dependent bootstrap random variables $\{\hat{X}^{(\omega)}_{n,j}, 1 \leq j \leq m(n)\}$ are negatively dependent and exchangeable.*

Proof. Let $\{x_1, x_2, \cdots, x_{m(n)}\}$ be a sequence of real numbers. For the first inequality of the negative dependence property, we note that we only need to consider the case $k(n)\tau(x_{(j)}) \geq j$ for all $1 \leq j \leq m(n)$. By Proposition 1(1)

$$\begin{aligned} & P\{\hat{X}^{(\omega)}_{n,1} \leq x_1, \cdots, \hat{X}^{(\omega)}_{n,m(n)} \leq x_{m(n)}\} \\ = & \prod_{j=1}^{m(n)} \frac{k(n)\tau(x_{(j)}) - (j-1)}{k(n)n - (j-1)} \\ \leq & \prod_{j=1}^{m(n)} \frac{k(n)\tau(x_{(j)})}{k(n)n} = \prod_{j=1}^{m(n)} P\{\hat{X}^{(\omega)}_{n,j} \leq x_j\}. \end{aligned}$$

For the second inequality of the negative dependence property, we note that we only need to consider the case $k(n)\mu(x_{(i)}) \geq m(n) - i$ for all $1 \leq i \leq m(n)$. By Proposition 1(2)

$$\begin{aligned} & P\{\hat{X}^{(\omega)}_{n,1} > x_1, \cdots, \hat{X}^{(\omega)}_{n,m(n)} > x_{m(n)}\} \\ = & \prod_{i=1}^{m(n)} \frac{k(n)\mu(x_{(i)}) - (m(n) - i + 1)}{k(n)n - (m(n) - i + 1)} \\ \leq & \prod_{i=1}^{m(n)} \frac{k(n)\mu(x_{(i)})}{k(n)n} = \prod_{i=1}^{m(n)} P\{\hat{X}^{(\omega)}_{n,i} > x_i\}. \end{aligned}$$

The exchangeability is obvious by Proposition 1.

3. Some technical lemmas

In this section we present six technical results that we will use in establishing the main result of this paper including its corollaries. Some of the lemmas are only generalizations and extensions of well-known results. For expository purposes we outline many of their proofs.

For the simplicity, by the log-function in this section we mean the natural logarithm function. The results can be easily generalized to any other logarithm function with base greater than one.

The first lemma is well known and trivial.

Lemma 1. *Let $\{Y_n, n \geq 1\}$ be a sequence of negatively dependent random variables.*
1) If $\{f_n, n \geq 1\}$ is a sequence of real functions all of which are monotone increasing (or all monotone decreasing), then $\{f_n(Y_n), n \geq 1\}$ is a sequence of negatively dependent random variables.
2) For every $n \geq 1, E(\prod_{j=1}^{n} Y_j) \leq \prod_{j=1}^{n} E(Y_j)$ provided the expectations are finite.

Proof. 1) Let f_n^{-1} denote the inverse function of $f_n, n \geq 1$ and assume that all f_n are increasing. Then for any $n \geq 1$ and all real $y_1, \cdots, y_n$ we have by the definition of negative dependence that

$$\begin{aligned} & P\{f_1(Y_1) \leq y_1, \cdots, f_n(Y_n) \leq y_n\} \\ = \;& P\{Y_1 \leq f_1^{-1}(y_1), \cdots, Y_n \leq f_n^{-1}(y_n)\} \\ \leq \;& \prod_{j=1}^{n} P\{Y_j \leq f_j^{-1}(y_j)\} = \prod_{j=1}^{n} P\{f_j(Y_j) \leq y_j\}. \end{aligned}$$

The second inequality also follows from the definition of negative dependence. The case of decreasing functions can be proved in the same manner.

2) Consider the expectation

$$E(\prod_{j=1}^{n} Y_j) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} P\{Y_1 > y_1, \cdots, Y_n > y_n\} \prod_{j=1}^{n} dy_j$$

$$\leq \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \prod_{j=1}^{n} P\{Y_j > y_j\} dy_j = \prod_{j=1}^{n} E(Y_j).$$

The next lemma is in effect the special case $a_n \equiv 1$ of Theorem 2 of Adler and Rosalsky (1987) and we omit the proof. The random variables $\{Y_n, n \geq 1\}$ are not assumed to be independent.

Lemma 2. *Let $\phi(t), t > 0$, be a continuous function that is positive, strictly increasing and satisfying the condition $\phi(t) \to \infty$ as $t \to \infty$. Put $b_n = \phi^{-1}(n), n \geq 1$, where $\phi^{-1}(t)$ is the inverse function of $\phi(t)$. Let $\{Y_n, n \geq 1\}$ be a sequence of random variables which is stochastically dominated by a random variable Y. If*

$$\sum_{n=k}^{\infty} b_n^{-1} = \mathcal{O}(k b_k^{-1}) \text{ and } E\phi(|Y|) < \infty,$$

then

$$\frac{1}{b_n} \sum_{j=1}^{n} Y_j \to 0 \;\; a.s.$$

The third lemma deals with convergence of maxima of random variables and is a generalization of the Corollary to Theorem 3 of Barnes and Tucker (1977). Again, no assumption of independence is made.

Lemma 3. *Let $\psi(t), t \geq 0$ be an increasing function such that $\psi(t) \to \infty$ as $t \to \infty$ and let $\{b_n,\ n \geq 1\}$ be a sequence of positive numbers such that $b_n = \psi^{-1}(n),\ n \geq 1,$, where $\psi^{-1}(t)$ is the inverse function of $\psi(t)$. Let $\{X_n,\ n \geq 1\}$ be a sequence of positive random variables which is stochastically dominated by a random variable X such that $E\psi(CX) < \infty$ for all $C > 0$. Then*

$$\frac{1}{b_n} \max_{1 \leq j \leq n} X_j \to 0 \ a.s.$$

Proof. We recall at the outset that for any random variable Y, the conditions $E|Y| < \infty$ and $\sum_{n=1}^{\infty} P\{|Y| \geq n\} < \infty$ are equivalent. Hence the assumption

$$E\psi\left(\frac{X}{\epsilon}\right) < \infty \text{ for all } \epsilon > 0$$

is equivalent to

$$\sum_{n=1}^{\infty} P\{X > \epsilon b_n\} < \infty \text{ for all } \epsilon > 0.$$

Then by the stochastic domination hypothesis and the Borel-Cantelli lemma $X_n/b_n \to 0$ a.s. For arbitrary $n \geq k \geq 2$,

$$\begin{aligned} & \frac{1}{b_n} \max_{1 \leq j \leq n} X_j \leq \frac{1}{b_n} \max_{1 \leq j \leq k-1} X_j + \frac{1}{b_n} \max_{k \leq j \leq n} X_j \\ \leq\ & \frac{1}{b_n} \max_{1 \leq j \leq k-1} X_j + \max_{k \leq j \leq n} X_j/b_j \text{ (since } \{b_n, n \geq 1\} \text{ is nondecreasing)} \\ \leq\ & \frac{1}{b_n} \max_{1 \leq j \leq k-1} X_j + \sup_{j \geq k} X_j/b_j \to 0 \end{aligned}$$

as first $n \to \infty$ and then $k \to \infty$.

Unfortunately, it is not possible to find the inverse to the function $\phi(t) = t^{1/\beta}/\log^{\gamma} t, \quad t > 0, \beta > 0,$ and $\gamma > 0$ in closed form. But the following lemma gives a good "approximation" to the inverse function.

Lemma 4. *Let $\phi(t) = t^{1/\beta} \log^{-\gamma/\beta} t$ and $\psi(t) = t^{\beta} \log^{\gamma} t, \quad t \geq e, \beta > 0,$ and $0 < \gamma < e$. Then for any $\epsilon > 0$ and for all sufficiently large t*

$$\beta^{-\gamma}(1-\epsilon)t \leq \psi(\phi(t)) \leq \beta^{-\gamma} t$$

and, consequently, for all sufficiently large t

$$\beta^{-\gamma}(1-\epsilon)\phi^{-1}(t) \leq \psi(t) \leq \beta^{-\gamma}\phi^{-1}(t).$$

Proof. Note that

$$\psi(\phi(t)) = \frac{t}{\beta^\gamma}\left(1 - \gamma\frac{\log\log t}{\log t}\right)^\gamma$$

and

$$\frac{\log\log t}{\log t} \downarrow 0$$

for $t \geq e^e$ which can be established by the differentiation.

It follows from Lemma 4 that for a positive random variable Y, the conditions $E\phi^{-1}(Y) < \infty$ and $E\psi(Y) < \infty$ are equivalent.

Lemma 5. *Let $\beta > 1$ and $b_n = n^\beta \log^{-2} n,\ n \geq 2$, then*

$$\sum_{j=n}^{\infty} b_j^{-1} = \mathcal{O}(nb_n^{-1}).$$

Proof. Note that for all large n, letting C_i $(1 \leq i \leq 4)$ denote positive constants,

$$\begin{aligned}
&\sum_{j=n}^{\infty} \frac{1}{b_j} = \sum_{j=n}^{\infty} \frac{\log^2 j}{j^\beta} = \sum_{m=1}^{\infty} \sum_{j=mn}^{(m+1)n-1} \frac{\log^2 j}{j^\beta} \\
\leq\ & \sum_{m=1}^{\infty} \frac{n\log^2(mn)}{(mn)^\beta} \\
&\quad \text{(since the sequence } \{\frac{\log^2 j}{j^\beta}, j \geq e^{2/\beta}\} \text{ is strictly decreasing)} \\
=\ & \frac{n}{n^\beta}\sum_{m=1}^{\infty}\left(\frac{\log^2 m}{m^\beta} + \frac{2(\log m)(\log n)}{m^\beta} + \frac{\log^2 n}{m^\beta}\right) \\
=\ & \frac{n}{n^\beta}(C_1 + C_2\log n + C_3\log^2 n) \ \ \text{(since } \beta > 1) \\
\leq\ & C_4\frac{n\log^2 n}{n^\beta} = C_4\frac{n}{b_n}.
\end{aligned}$$

The exponential inequality presented in the last lemma is the key tool used in establishing in Theorem 4 the asymptotic probability for the deviations of dependent bootstrap means from the sample mean. It is a dependent bootstrap analog of the Mikosch exponential inequality (Mikosch (1994), Lemma 5.1). We mention that this result was proved by Mikosch (1994) 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.

Lemma 6. *Let $\{a_n, n \geq 1\}$ and $\{h_n, n \geq 1\}$ be two sequences of positive real numbers and 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 $h_n M_n(\omega) < 1$ the following inequality holds for all $\epsilon > 0$:*

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

where

$$M_n(\omega) = \frac{1}{m(n)}\max_{1\leq j\leq n}|X_j(\omega) - \overline{X}_n(\omega)|$$

and

$$B_n(\omega) = \frac{1}{nm(n)}\sum_{j=1}^{n}\left(X_j(\omega) - \overline{X}_n(\omega)\right)^2.$$

Proof. By the Markov inequality,

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

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 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) the random variables

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

are negatively dependent and identically distributed.

Therefore

$$
\begin{aligned}
& E\exp\left\{h_n\left(\frac{\sum_{j=1}^{m(n)}\hat{X}_{n,j}^{(\omega)}}{m(n)} - \overline{X}_n(\omega)\right)\right\} \\
= \; & E\left[\prod_{j=1}^{m(n)}\exp\left\{\frac{h_n}{m(n)}\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right\}\right]
\end{aligned}
$$

$$
\begin{aligned}
\leq \; & \prod_{j=1}^{m(n)} E\exp\left\{\frac{h_n}{m(n)}\left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega)\right)\right\} \text{ (by Lemma 1(2))} \\
= \; & \left[E\exp\left\{\frac{h_n}{m(n)}\left(\hat{X}_{n,1}^{(\omega)} - \overline{X}_n(\omega)\right)\right\}\right]^{m(n)} \text{ (by identical distribution)} \\
= \; & \left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{\frac{h_n}{m(n)}\left(X_i(\omega) - \overline{X}_n(\omega)\right)\right\}\right]^{m(n)} \\
= \; & \left[1 + \frac{1}{n}\sum_{i=1}^{n}\left(\frac{h_n^2}{2!m(n)^2}\left(X_i(\omega) - \overline{X}_n(\omega)\right)^2 + \frac{h_n^3}{3!m(n)^3}\left(X_i(\omega) - \overline{X}_n(\omega)\right)^3\right.\right. \\
& \left.\left. + \frac{h_n^4}{4!m(n)^4}\left(X_i(\omega) - \overline{X}_n(\omega)\right)^4 + \cdots\right)\right]^{m(n)} \\
= \; & \left[1 + \frac{h_n^2}{m(n)}\sum_{i=1}^{n}\frac{\left(X_i(\omega) - \overline{X}_n(\omega)\right)^2}{nm(n)}\left[\frac{1}{2!} + \frac{h_n}{3!}\left(\frac{X_i(\omega) - \overline{X}_n(\omega)}{m(n)}\right)\right.\right. \\
& \left.\left. + \frac{h_n^2}{4!}\left(\frac{X_i(\omega) - \overline{X}_n(\omega)}{m(n)}\right)^2 + \cdots\right]\right]^{m(n)} \\
\leq \; & \left[1 + \frac{h_n^2}{m(n)}\frac{B_n(\omega)}{2}\left(1 + h_nM_n(\omega) + (h_nM_n(\omega))^2 + \cdots\right)\right]^{m(n)} \\
= \; & \left[1 + \frac{h_n^2}{2m(n)}\frac{B_n(\omega)}{(1 - h_nM_n(\omega))}\right]^{m(n)} \text{ (since } h_nM_n(\omega) < 1) \\
\leq \; & \left[\exp\left\{\frac{h_n^2}{2m(n)}\frac{B_n(\omega)}{(1 - h_nM_n(\omega))}\right\}\right]^{m(n)} \\
= \; & \exp\left\{\frac{h_n^2B_n(\omega)}{2\left(1 - h_nM_n(\omega)\right)}\right\}.
\end{aligned}
$$

Hence,

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

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

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

where $\psi^{-1}(t)$ is the inverse function of $\psi(t)$. Let $\{X_n, n\geq 1\}$ be a sequence of random variables which is stochastically dominated by a random variables X such that $E\psi(CX)<\infty$ for all $C>0$ and let $\{a_n, n\geq 1\}$ be a sequence of positive constants. Then for almost every $\omega\in\Omega$, for every $\epsilon>0$, and for every real number r,

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

Proof. Fix the arbitrary constants r and $\epsilon>0$ and let $h_n=\frac{rm(n)}{\epsilon\psi^{-1}(n)}, n\geq\ 1$. We may assume that $r>0$. 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.

Next, in Lemma 2 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

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

Hence,

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

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 6.

Remark 2. The conclusion of Theorem 4 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 arbitrary 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.

Corollary 1. *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 $\epsilon > 0$
$$P\left\{ \frac{1}{n^{1/\alpha}} \left| \sum_{j=1}^{n} \left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega) \right) \right| \geq \epsilon \right\} = o(1);$$
that is, for almost every $\omega \in \Omega$ the weak law of large numbers
$$\frac{1}{n^{1/\alpha}} \sum_{j=1}^{n} \left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega) \right) \to 0 \quad \textit{in probability}$$
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 4 for almost every $\omega \in \Omega$, for every $\epsilon > 0$ and every $r > 0$, and for all sufficiently large n and some constant $C < \infty$,
$$P\left\{ \frac{1}{n^{1/\alpha}} \left| \sum_{j=1}^{n} \left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega) \right) \right| \geq \epsilon \right\} \leq C \exp\{-r\};$$
that is, since $r > 0$ is arbitrary
$$\frac{1}{n^{1/\alpha}} \sum_{j=1}^{n} \left(\hat{X}_{n,j}^{(\omega)} - \overline{X}_n(\omega) \right) \to 0 \quad \text{in probability.}$$

Corollary 2. *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* $\omega \in \Omega$

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

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

Remark 3. Theorem 3 easily follows from Corollary 2. To see this, for any constant q from Theorem 3, let $r = q + 2$ and apply Corollary 2.

Corollary 3. *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* $\omega \in \Omega$

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

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 $a_n = n^{1/\alpha}$, $n \geq 1$, applying Theorem 4 we obtain the result.

References

[1] Adler, A., Rosalsky, A. Some general strong laws for weighted sums of stochastically dominated random variables. Stochastic Anal. Appl. 5, (1987), 1-16.

[2] Arenal-Gutiérrez, E., Matrán, C., Cuesta-Albertos, J.A. On the unconditional strong law of large numbers for the bootstrap mean. Statist. Probab. Lett. 27, (1996), 49-60.

[3] Athreya, K.B. Strong law for the bootstrap. Statist. Probab. Lett. 1, (1983), 147-150.

[4] Barnes, G.R., Tucker, H.G. On almost sure convergence of normed maxima of independent random variables. J. London Math. Soc. (2), 16, (1977), 377-383.

[5] Bickel, P.J., Freedman, D.A. Some asymptotic theory for the bootstrap. Ann. Statist. 9, (1981), 1196-1217.

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

[7] Efron, B. Bootstrap methods: another look at the jackknife. Ann Statist. 7, (1979), 1-26.

[8] Hu, T.C. Strong consistencies of the bootstrap moments. Internat. J. Math. & Math. Sci. 14, (1991), 797-802.

[9] Hu, T.-C., Ordóñez Cabrera, M., Volodin, A. Asymptotic probability for the bootstrapped means deviations from the sample mean. To appear in Statist. Probab. Lett. (2005).

[10] Lehmann, E. Some concepts of dependence. Ann. Math. Statist., 37, (1966), 1137-1153.

[11] Li, D., Rosalsky, A., Ahmed, S. E. Complete convergence of bootstrapped means and moments of the supremum of normed bootstrapped sums. Stochastic Anal. Appl. 17, (1999), 799-814.

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

[13] Smith, W.D., Taylor, R.L. Consistency of dependent bootstrap estimators. Statistics, combinatorics and related areas (Mobile, AL, 1999). Amer. J. Math. Management Sci., 21 no. 3-4, (2001a), 359-382.

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

[15] Volodin, A., Ordóñez Cabrera M., Hu T.-C. Convergence rate of the dependent bootstrapped means. To appear in Theory Probab. Appl. 50, (2005).

Department of Mathematics and Statistics, University of Windsor, Windsor, Ontario, Canada N9B 3P4

E-mail address: seahmed@uwindsor.ca

Department of Mathematics and Statistics, Lakehead University, Thunder Bay, Ontario, Canada P7B 5E1

E-mail address: dli@sleet.lakeheadu.ca

Department of Statistics, University of Florida, Gainesville, Florida 32611 USA

E-mail address: rosalsky@stat.ufl.edu

Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, Canada S4S 0A2

E-mail address: andrei@math.uregina.ca

Received May 5, 2005