PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 130, Number 7, Pages 2133–2138
S 0002-9939(01)06368-7
Article electronically published on December 31, 2001

A LARGE DEVIATION PRINCIPLE FOR BOOTSTRAPPED SAMPLE MEANS

DELI LI, ANDREW ROSALSKY, AND DHAIFALLA K. AL-MUTAIRI

(Communicated by Richard A. Davis)

ABSTRACT. A large deviation principle for bootstrapped sample means is established. It relies on the Bolthausen large deviation principle for sums of i.i.d. Banach space valued random variables. The rate function of the large deviation principle for bootstrapped sample means is the same as the classical one.

1. INTRODUCTION

Let $\{X, X_n;\ n \geq 1\}$ be a sequence of independent and identically distributed (i.i.d.) real valued random variables defined on a complete probability space $(\Omega, \mathcal{F}, P)$. Let $\mu = \mathcal{L}(X)$ denote the probability law of X. For $\omega \in \Omega$ and $n \geq 1$, let $\mu_n(\mathbf{A}, \omega) = n^{-1}\text{Card}\,\{i; X_i(\omega) \in \mathbf{A}, 1 \leq i \leq n\}, \mathbf{A} \subseteq \mathbf{R} = (-\infty, \infty)$ denote the empirical measure. Let $\{m(n); n \geq 1\}$ be a sequence of positive integers and for each $n \geq 1$, let the random variables $\{\hat{X}_{n,j}; 1 \leq j \leq m(n)\}$ result by sampling $m(n)$ times with replacement from the n observations $X_1, ..., X_n$ such that for each of the $m(n)$ selections, each X_i has probability n^{-1} of being chosen. More precisely, for each $n \geq 1$, let $\hat{X}_{n,j} = X_{Z(n,j)}$ where $\{Z(n,j); 1 \leq j \leq m(n)\}$ are independent random variables uniformly distributed over $\{1, \ldots, n\}$ and independent of $\{X_1, \ldots, X_n\}$. Then for each $n \geq 1, \{\hat{X}_{n,j}; 1 \leq j \leq m(n)\}$ are conditionally i.i.d. given $\{X_1, \ldots, X_n\}$ with $P\{\hat{X}_{n,1} = X_i | X_1, \ldots, X_n\} = n^{-1}$ almost surely (a.s.), $1 \leq i \leq n$. For each $n \geq 1$, $\{\hat{X}_{n,j}; 1 \leq j \leq m(n)\}$ is the so-called Efron [16] bootstrap sample from $X_1, ..., X_n$ with bootstrap sample size $m(n)$.

The main classes of limit theorems of classical probability theory for sample means $\sum_{i=1}^{n} X_i/n, n \geq 1$, have counterparts for bootstrapped sample means $\sum_{j=1}^{m(n)} \hat{X}_{n,j}/m(n), n \geq 1$. References to these bootstrap counterparts of the various classes of classical limit theorems are listed as follows:

- central limit theorem (Bickel and Freedman [7], Singh [24], Giné and Zinn [19]),
- weak law of large numbers (Bickel and Freedman [7], Athreya, Ghosh, Low, and Sen [5], Csörgő [13], Arenal-Gutiérrez, Matrán, and Cuesta-Albertos [2]),

Received by the editors February 3, 2000 and, in revised form, February 15, 2001.
1991 *Mathematics Subject Classification.* Primary 60F10, 62G09; Secondary 60B12, 62G20.
Key words and phrases. Bootstrapped sample means, large deviation principle, Banach space valued random variables.
The research of the first author was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

- strong law of large numbers (Athreya [4], Athreya, Ghosh, Low, and Sen [5], Csörgő [13], Mikosch [22], Arenal-Gutiérrez, Matrán, and Cuesta-Albertos [3], Csörgő and Wu [14]),
- moderate deviation theorem (Hall [20]),
- law of the iterated logarithm (Mikosch [22], Ahmed, Li, Rosalsky, and Volodin [1]),
- complete convergence theorem (Li, Rosalsky, Ahmed [21], Csörgő and Wu [14]).

In this paper, we present a large deviation principle (LDP) for bootstrapped sample means. The study of large deviations began with Cramér [12] and Chernoff [10] who investigated the convergence of

$$\frac{1}{n}\log P\left\{\frac{\sum_{i=1}^{n} X_i}{n} \in \mathbf{A}\right\},$$

where $\mathbf{A}$ is an interval. Donsker and Varadhan [15] and Bahadur and Zabell [6] established a LDP for sums of i.i.d. Banach space valued random variables. Bolthausen [9] extended the Donsker-Varadhan-Bahadur-Zabell LDP when the laws of the random variables converge weakly and satisfy a uniform exponential integrability condition.

As an application of the Bolthausen LDP, we shall prove the following LDP for bootstrapped sample means. We are not assuming that the bootstrap samples $\left\{\{\hat{X}_{n,j}; 1 \le j \le m(n)\}; n \ge 1\right\}$ are conditionally independent given $\{X_n; n \ge 1\}$.

Theorem 1.1. *Let $\{X, X_n; n \ge 1\}$ be a sequence of i.i.d. real valued random variables such that*

$$M(t) \equiv \int_{-\infty}^{\infty} \exp(tx)\mu(dx) < \infty \text{ for all } t \in \mathbf{R}. \tag{1.1}$$

For $x \in \mathbf{R}$ and $\mathbf{A} \subseteq \mathbf{R}$, write

$$\lambda(x) = \sup_{t\in\mathbf{R}} (tx - \log M(t)) \text{ and } \Lambda(\mathbf{A}) = \inf_{x\in\mathbf{A}} \lambda(x).$$

Then, for every sequence of positive integers $m(n) \to \infty$, we have the LDPs:

1. *For every closed set $\mathbf{A} \subseteq \mathbf{R}$,*

$$\limsup_{n\to\infty} \frac{1}{m(n)} \log P\left\{\frac{\sum_{j=1}^{m(n)} \hat{X}_{n,j}}{m(n)} \in \mathbf{A} \middle| X_1, \ldots, X_n\right\} \le -\Lambda(\mathbf{A}) \text{ a.s.}$$

2. *For every open set $\mathbf{A} \subseteq \mathbf{R}$,*

$$\liminf_{n\to\infty} \frac{1}{m(n)} \log P\left\{\frac{\sum_{j=1}^{m(n)} \hat{X}_{n,j}}{m(n)} \in \mathbf{A} \middle| X_1, \ldots, X_n\right\} \ge -\Lambda(\mathbf{A}) \text{ a.s.}$$

The rate function $\Lambda(\cdot)$ of this LDP for bootstrapped sample means is the same as the classical one.

It should be mentioned that in the case $m(n) = n, n \ge 1$, Hall [20] proved a large deviation theorem for

$$P\left\{\frac{\sum_{j=1}^{n}\left(\hat{X}_{n,j} - \overline{X}_n\right)}{n^{1/2}\hat{\sigma}_n} > x_n \middle| X_1, \ldots, X_n\right\}$$

where $\overline{X}_n = \sum_{i=1}^n X_i/n, \hat{\sigma}_n^2 = n^{-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2, n \geq 1$, and $0 \leq x_n = o(n^{1/2})$, but his work and the current work do not entail each other.

Section 2 is devoted to the proof of Theorem 1.1. The proof given there consists of an application of Theorem 2.2 below which is a modification of Theorem 2.1 (the Bolthausen [9] LDP).

2. Proof of Theorem 1.1

In this section we prove Theorem 1.1.

Let $(\mathbf{B}, \|\cdot\|)$ be a real separable Banach space, equipped with the Borel σ-field $\mathcal{B}$, let $\mathcal{P}$ be the set of probability measures on $(\mathbf{B}, \mathcal{B})$, and let $\mathbf{B}^*$ denote the (topological) dual of $\mathbf{B}$.

For $\nu \in \mathcal{P}$, $f \in \mathbf{B}^*$, $a \in \mathbf{B}$, and $\mathbf{A} \subseteq \mathbf{B}$, write

$$M(f|\nu) = \int_{\mathbf{B}} \exp(f(x))\nu(dx)$$

and

$$\lambda(a|\nu) = \sup_{f \in \mathbf{B}^*} \{f(a) - \log M(f|\nu)\}, \Lambda(\mathbf{A}|\nu) = \inf_{a \in \mathbf{A}} \lambda(a|\nu).$$

For each $n \geq 1$ and $m \geq 1$, let ν_n^{*m} denote the m-fold convolution of ν_n and let $n\mathbf{A} = \{na; a \in \mathbf{A}\}$.

The following remarkable result is the Bolthausen [9] LDP for sums of i.i.d. Banach space valued random variables.

Theorem 2.1. *Let $\nu, \nu_n \in \mathcal{P}, n \geq 1$ be such that $\nu_n \xrightarrow{w} \nu$ and*

$$\sup_{n \geq 1} \int_{\mathbf{B}} \exp(t\|x\|)\nu_n(dx) < \infty \text{ for all } t > 0.$$

Then

$$\text{if } \mathbf{A} \subseteq \mathbf{B} \text{ is closed, then } \limsup_{n\to\infty} n^{-1} \log \nu_n^{*n}(n\mathbf{A}) \leq -\Lambda(\mathbf{A}|\nu)$$

and

$$\text{if } \mathbf{A} \subseteq \mathbf{B} \text{ is open, then } \liminf_{n\to\infty} n^{-1} \log \nu_n^{*n}(n\mathbf{A}) \geq -\Lambda(\mathbf{A}|\nu).$$

When $\nu_n = \nu, n \geq 1$, Theorem 2.1 reduces to the Donsker-Varadhan-Bahadur-Zabell LDP for sums of i.i.d. Banach space valued random variables. (See Donsker and Varadhan [15] and Bahadur and Zabell [6].) The special case where the ν_n are Gaussian is due to Ellis and Rosen [17] and Chevet [11].

To prove Theorem 1.1, we use the following modification of the Bolthausen [9] LDP. In view of Theorem 2.1, Theorem 2.2 is clearly true for $m(n) \uparrow \infty$ and subsequently for $m(n) \to \infty$.

Theorem 2.2. *Under the conditions of Theorem 2.1, for every sequence of positive integers $m(n) \to \infty$,*

$$\text{if } \mathbf{A} \subseteq \mathbf{B} \text{ is closed, then } \limsup_{n\to\infty} (1/m(n)) \log \nu_n^{*m(n)}(m(n)\mathbf{A}) \leq -\Lambda(\mathbf{A}|\nu)$$

and

$$\text{if } \mathbf{A} \subseteq \mathbf{B} \text{ is open, then } \liminf_{n\to\infty} (1/m(n)) \log \nu_n^{*m(n)}(m(n)\mathbf{A}) \geq -\Lambda(\mathbf{A}|\nu).$$

Proof of Theorem 1.1. Let

$$\begin{aligned}
\Omega_1 &= \left\{\omega; \{\mu_n(\cdot,\omega); n \geq 1\} \text{ converges weakly to } \mu(\cdot)\right\}, \\
\Omega_2 &= \left\{\omega; \sup_{n\geq 1} \int_{-\infty}^{\infty} \exp(t|x|)\mu_n(dx,\omega)\right. \\
&\qquad \left. = \sup_{n\geq 1} \frac{\sum_{i=1}^{n} \exp\left(t\left|X_i(\omega)\right|\right)}{n} < \infty \text{ for all } t > 0\right\} \\
&= \bigcap_{k=1}^{\infty} \left\{\omega; \sup_{n\geq 1} \frac{\sum_{i=1}^{n} \exp\left(k\left|X_i(\omega)\right|\right)}{n} < \infty\right\},
\end{aligned}$$

and set $\Omega_0 = \Omega_1 \cap \Omega_2$. The Glivenko-Cantelli theorem (see, e.g., Billingsley [8], p. 275) ensures that $P(\Omega_1) = 1$. Now by the Kolmogorov strong law of large numbers and (1.1), $P\left\{\sup_{n\geq 1} n^{-1}\sum_{i=1}^{n}\exp\left(k\left|X_i\right|\right) < \infty\right\} = 1$ for all $k \geq 1$ whence $P(\Omega_2) = 1$. Thus, $P(\Omega_0) = 1$. Then we have by Theorem 2.2 the LDPs:

1. For every closed set $\mathbf{A} \subseteq \mathbf{R}$,

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{m(n)} \log P\left\{\frac{\sum_{j=1}^{m(n)} \hat{X}_{n,j}}{m(n)} \in \mathbf{A} \middle| X_1, \dots, X_n\right\} \\
= &\limsup_{n\to\infty} \frac{1}{m(n)} \log \mu_n^{*m(n)}(m(n)\mathbf{A}, \cdot) \\
\leq & -\Lambda(\mathbf{A}) \text{ a.s.}
\end{aligned}$$

and

2. For every open set $\mathbf{A} \subseteq \mathbf{R}$,

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{m(n)} \log P\left\{\frac{\sum_{j=1}^{m(n)} \hat{X}_{n,j}}{m(n)} \in \mathbf{A} \middle| X_1, \dots, X_n\right\} \\
= &\liminf_{n\to\infty} \frac{1}{m(n)} \log \mu_n^{*m(n)}(m(n)\mathbf{A}, \cdot) \\
\geq & -\Lambda(\mathbf{A}) \text{ a.s.},
\end{aligned}$$

where $\mu_n^{*m(n)}(\cdot,\cdot)$ is the $m(n)$-fold convolution of $\mu_n(\cdot,\cdot), n \geq 1$. The theorem is proved. □

3. Some final comments

Hall [20] applied his bootstrap LDP to compare the relative error in bootstrap and Edgeworth approximation formulae. As far as application of the current work is concerned, we see that the performance of the bootstrap LDP in Theorem 1.1 for estimating tail probabilities of the form

$$P\left\{\frac{\sum_{j=1}^{n} \hat{X}_{n,j}}{n} \geq a \middle| X_1, \dots, X_n\right\}$$

is the same as that of the Cramér [12] and Chernoff [10] LDP for estimating

$$P\left\{\frac{\sum_{i=1}^{n} X_i}{n} \geq a\right\}.$$

Consequently, it appears that Theorem 1.1 can play a major role in obtaining a bootstrap counterpart to the Shepp-Erdös-Rényi strong law of large numbers (see Erdös and Rényi [18] and Shepp [23]).

Acknowledgement

The authors are grateful to the referee for carefully reading the manuscript, for pointing out some obscureness in the presentation, and for offering substantial comments and suggestions on ways to improve the paper.

References

[1] S.E. Ahmed, D. Li, A. Rosalsky, and A.I. Volodin, *Almost sure lim sup behavior of bootstrapped means with applications to pairwise i.i.d. sequences and stationary ergodic sequences,* J. Statist. Plann. Inference **98** (2001), 1-14.

[2] E. Arenal-Gutiérrez, C. Matrán, and J.A. Cuesta-Albertos, *Unconditional Glivenko-Cantelli-type theorems and weak laws of large numbers for bootstrap,* Statist. Probab. Lett. **26** (1996), 365-375. MR **97c:**60048

[3] E. Arenal-Gutiérrez, C. Matrán, and J.A. Cuesta-Albertos, *On the unconditional strong law of large numbers for the bootstrap mean,* Statist. Probab. Lett. **27** (1996), 49-60. MR **97h:**60020

[4] K.B. Athreya, *Strong law for the bootstrap,* Statist. Probab. Lett. **1** (1983), 147-150. MR **84g:**62026

[5] K.B. Athreya, M. Ghosh, L.Y. Low, and P.K. Sen, *Laws of large numbers for bootstrapped U-statistics,* J. Statist. Plann. Inference **9** (1984), 185-194. MR **85m:**62062

[6] R.R. Bahadur and S.L. Zabell, *Large deviations of the sample mean in general vector spaces,* Ann. Probab. **7** (1979), 587-621. MR **80i:**60031

[7] P.J. Bickel and D.A. Freedman, *Some asymptotic theory for the bootstrap,* Ann. Statist. **9** (1981), 1196-1217. MR **83a:**62051

[8] P. Billingsley, *Probability and Measure*, 2nd ed., Wiley, New York, 1986. MR **87f:**60001

[9] E. Bolthausen, *On the probability of large deviations in Banach spaces,* Ann. Probab. **12** (1984), 427-435. MR **85k:**60011

[10] H. Chernoff, *A measure of asymptotic efficiency for tests of hypothesis based on the sum of observations,* Ann. Math. Statist. **23** (1952), 493-507. MR **15:**241c

[11] S. Chevet, *Gaussian measures and large deviations,* Unpublished manuscript (1982).

[12] H. Cramér, *Sur un nouveau théorème-limite de la théorie des probabilités,* Actualités Sci. Indust. **736** (1938), 5-23.

[13] S. Csörgő, *On the law of large numbers for the bootstrap mean,* Statist. Probab. Lett. **14** (1992), 1-7. MR **93g:**60063

[14] S. Csörgő and W.B. Wu, *Random graphs and the strong convegence of bootstrap means,* Combin. Probab. Comput. **9** (2000), 315-347. CMP 2001:02

[15] M.D. Donsker and S.R.S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time III,* Comm. Pure Appl. Math. **29** (1976), 389-461. MR **55:**1492

[16] B. Efron, *Bootstrap methods: Another look at the jackknife,* Ann. Statist. **7** (1979), 1-26. MR **80b:**62021

[17] R.S. Ellis and J.S. Rosen, *Laplace's method for Gaussian integrals with an application to statistical mechanics,* Ann. Probab. **10** (1982), 47-66. MR **82m:**60010; MR **84e:**60014

[18] P. Erdös and A. Rényi, *On a new law of large numbers,* J. Analyse Math. **23** (1970), 103-111. MR **42:**6907

[19] E. Giné and J. Zinn, *Necessary conditions for the bootstrap of the mean,* Ann. Statist. **17** (1989), 684-691. MR **90e:**62030

[20] P. Hall, *On the relative performance of bootstrap and Edgeworth approximations of a distribution function,* J. Multivariate Anal. **35** (1990), 108-129. MR **91m:**62077

[21] D. Li, A. Rosalsky, and S.E. Ahmed, *Complete convergence of bootstrapped means and moments of the supremum of normed boostrapped sums,* Stochastic Anal. Appl. **17** (1999), 799-814. MR **2001g:**60067

[22] T. Mikosch, *Amost sure convergence of bootstrapped means and U-statistics,* J. Statist. Plann. Inference **41** (1994), 1-19. MR **95h:**60046

[23] L.A. Shepp, *A limit theorem concerning moving averages,* Ann. Math. Statist. **35** (1964), 424-428. MR **29:**4091

[24] K. Singh, *On the asymptotic accuracy of Efron's bootstrap,* Ann. Statist. **9** (1981), 1187-1195. MR **83c:**62047

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

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

Department of Statistics, University of Florida, P.O. Box 118545, Gainesville, Florida 32611

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

Department of Statistics & Operations Research, Kuwait University, P.O. Box 21, Khaldiya 72461, Kuwait

E-mail address: dhaif@kuc01.kuniv.edu.kw