On the Kolmogorov exponential inequality for negatively dependent random variables

Andrei Volodin

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

e-mail: volodin@math.uregina.ca

Abstract. The classical exponential Kolmogorov inequality is slightly improved and generalized on the case of negatively dependent random variables.
Key words: Exponential inequalities, negative dependence.

AMS subject classifications: 60E15, 60F15.

Dedicated to Professor Syed Ejaz Ahmed

1 Introduction

It is a great pleasure for me to contribute to this issue in honour of Professor Syed Ejaz Ahmed.

Bosrgnia, Patterson and Taylor [2] mentioned that in many stochastic models, the assumption that random variables are independent is not plausible. Increases in some random variables are often related to decreases in other random variables so an assumption of negative dependence is more appropriate than an assumption of independence. Lehmann [3] investigated various concepts of positive and negative dependence in the bivariate case. In this paper, we present an improvement and generalization of the classical Kolmogorov exponential inequality (cf. for example, Stout [5], p.260 - 263) on the case of negative dependent random variables. The result is of course true for independent random variables.

Definition. The random variables $X_1, \cdots, X_n$ are said to be negatively dependent if we have

$$P\left\{\cup_{j=1}^{n}(X_j \le x_j)\right\} \le \Pi_{j=1}^{n} P\left\{X_j \le x_j\right\},$$

and

$$P\left\{\cup_{j=1}^{n}(X_j \ge x_j)\right\} \le \Pi_{j=1}^{n} P\left\{X_j \ge x_j\right\}$$

for all real $x_1, \cdots, x_n$.

One of the most interesting and useful examples of negative dependent random variables arises in the situation of a sample from finite population without of replacement. Hence we can apply our result to so-called *dependent bootstrap*, that is, the sample drawn without replacement form the collection of items made up of copies of sample observations. Smith and Taylor [4] obtained consistency of the bootstrap mean. With the help of the present paper we can prove the law of iterated logarithm type results for dependent bootstrap analogous to the results of Ahmed, Li, Rosalsky and Volodin [1] for independent bootstrap.

2 Lemata

The following two lemmas are used to obtain the main result in the next section. The first lemma is a simple corollary of the observation that if $X_1, \cdots, X_n$ is a sequence of negatively dependent random variables, then $e^{X_1}, \cdots, e^{X_n}$ is also negatively dependent. The same argument is used in Bosrgnia, Patterson and Taylor ([2] p.1167) in the proof of Lemma 3.3.

Lemma 1. *If $X_1, \cdots, X_n$ is a sequence of negatively dependent random variables, then*

$$E\exp\left\{\sum_{k=1}^{n} X_k\right\} \le \Pi_{k=1}^{n} E\exp\{X_k\}.$$

The second lemma is only a technical result that will help us to improve a constant in the Kolmogorov exponential inequality.

Lemma 2. *Let $a > 0$ and $0 < \alpha \le \frac{a^3}{2(e^a - 1 - a - a^2/2)}$. Then*

$$e^x - 1 - x - \frac{x^2}{2} \le \frac{x^3}{2\alpha}$$

for all $0 \le x \le a$.

Proof. Consider the function

$$f(x,\alpha) = \ln\left(1 + x + \frac{x^2}{2} + \frac{x^3}{2\alpha}\right) - x.$$

We need to prove that $f(x,\alpha) \ge 0$ for all $0 < \alpha \le \frac{a^3}{2(e^a-1-a-a^2/2)}$ and $0 \le x \le a$.

Take the derivative

$$\frac{\partial f}{\partial x} = -\frac{x^2\,(x-(3-\alpha))}{2\alpha\,(1+x+x^2/2+x^3/(2\alpha))}.$$

Hence f is increasing by x on the interval $(0, 3-\alpha)$ and decreasing on the interval $(3-\alpha, a)$.

Note that $f(0,\alpha) = 0$ and $f(a,\alpha) \ge 0$ since $\alpha \le \frac{a^3}{2(e^a-1-a-a^2/2)}$.□

3 Main result

Now we can formulate and prove our main result.

Theorem. *Let $X_1, \cdots, X_n$ be a sequence of negatively dependent random variables with zero means and finite variances. Let $s_n^2 = \sum_{k=1}^{n} EX_k^2$ and assume that $|X_k| \le Cs_n$ almost surely for each $1 \le k \le n$ and $n \ge 1$. Then for each $a > 0$ and $n \ge 1$, the assumptions $\epsilon C \le a$ and $0 < \alpha \le \frac{a^3}{2(e^a-1-a-a^2/2)}$ imply that*

$$P\{S_n/s_n > \epsilon\} \le \exp\left\{-\frac{\epsilon^2}{2}\left(1 - \frac{\epsilon C}{\alpha}\right)\right\},$$

where $S_n = \sum_{k=1}^{n} X_k$ as usual.

Proof. We will follow the proof of the classical Kolmogorov exponential inequality (cf. Stout [5], p.263). Fix $n \ge 1$ and $a > 0$. Suppose $x = \epsilon C \le a$.

For each $1 \le k \le n$,

$$
\begin{aligned}
E\exp\{\epsilon X_k/s_k\} &= 1 + \frac{\epsilon^2 E X_k^2}{2! s_n^2} + \frac{\epsilon^3 X_k^3}{3! s_n^3} + \cdots \\
&\le 1 + \frac{\epsilon^2 E X_k^2}{2 s_n^2}\left(1 + \frac{\epsilon C}{3} + \frac{\epsilon^2 C^2}{3\cdot 4} + \cdots\right) \\
&\le 1 + \frac{\epsilon^2 E X_k^2}{2 s_n^2}\left(1 + \frac{x}{3} + \frac{x^2}{3\cdot 4} + \cdots\right) \\
&= 1 + \frac{\epsilon^2 E X_k^2}{2 s_n^2}\left(e^x - x - \frac{x^2}{2}\right) \\
&\le 1 + \frac{\epsilon^2 E X_k^2}{2 s_n^2}\left(1 + \frac{x}{\alpha}\right) \text{ by Lemma 2} \\
&\le \exp\left\{\frac{\epsilon^2 E X_k^2}{2 s_n^2}\left(1 + \frac{x}{\alpha}\right)\right\}
\end{aligned}
$$

since $1 + t \le e^t$ for all t. By Lemma 1,

$$
E\exp\{\epsilon S_n/s_n\} \le \exp\left\{\frac{\epsilon^2}{2}\left(1 + \frac{\epsilon C}{\alpha}\right)\right\}.
$$

Thus

$$
\begin{aligned}
P\{S_n/s_n > \epsilon\} &\le \exp\{-\epsilon^2\} E\exp\{\epsilon S_n/s_n\} \\
&\le \exp\left\{-\frac{\epsilon^2}{2}\left(1 + \frac{\epsilon C}{\alpha}\right)\right\}. \square
\end{aligned}
$$

Remarks. 1. Even for $a = 1$, our theorem gives better constant $\alpha = \frac{1}{2e-5} = 2.2906 > 2$, while in the Kolmogorov original inequality we have $\alpha = 2$ (cf. Stout, p.263).
2. If $a \to 0$ then $\alpha \to 3$. As $a \to \infty$ then $\alpha \to 2$. We need $a \to 0$ for the proof of the law of iterated logarithm.
3. Another interesting advantage of the theorem is that we can consider any positive a, while in Kolmogorov inequality $a = 1$. In our inequality, the upper bound involves a fixed (given) c and fixed ϵ and a variable α. Now, α is a function of a, for $a \ge \epsilon c$. Note that the left-hand side of the inequality doesn't

involve a anywhere, whereas the right-hand side is, in effect, a function of a. So the best possible inequality occurs when a is chosen so that $\alpha(a)$ is maximized on the interval $[\epsilon, \infty]$. However, α is a decreasing function of a, so the maximum value of the upper bound occurs when a is as small as possible; i.e., when $a = \epsilon c$.

In short, then technically we have a family of inequalities – one for each value of $a \geq \epsilon c$. However, the special case where $\alpha = \alpha(\epsilon c)$ implies the validity of the inequality for all larger values of α – so there is really only one inequality, for one specific value of α.

Acknowledgments. The author is grateful to his supervisor Professor S. Ejaz Ahmed for carefully reading the manuscript and for offering some suggestions which resulted in an improved presentation. Author is also grateful to Professor A. Rosalsky (University of Florida) and Professor R.J. Tomkins (University of Regina) for the extremely helpful discussions leading to the result.

Reference

[1] Ahmed,S.E., Li, D., Rosalsky, A., and Volodin, A., Almost sure Lim Sup behavior of bootstrapped means with applications to pairwise i.i.d. sequences and stationary ergodic sequences. To appear in *Journal of Statistical Planning and Inference.*
[2] Bozorgnia, A., Patterson, R.F. and Taylor, R.L., Limit theorems for dependent random variables, *Proceedings of WCNA'92,* Tampa, Florida (1996), 1639 - 1650.
[3] Lehmann, E., Some concepts of dependence, *Ann. Math. Statist.* **37** (1966), 1137 - 1153.
[4] Smith, W.S. and Taylor, R.L., Consistency of dependent bootstrap estimators. UGA Technical Report 2000-6.
[5] Stout, W. F., *Almost Sure Convergence.* Academic Press, New York, 1974.