



HAL
open science

Sharp large deviations under Bernstein's condition

Xiequan Fan, Ion Grama, Quansheng Liu

► **To cite this version:**

Xiequan Fan, Ion Grama, Quansheng Liu. Sharp large deviations under Bernstein's condition. Comptes rendus hebdomadaires des séances de l'Académie des sciences, 2013, 351, pp.845-848. 10.1016/j.crma.2013.10.015 . hal-00905527

HAL Id: hal-00905527

<https://hal.science/hal-00905527v1>

Submitted on 18 Nov 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sharp large deviations under Bernstein's condition

Xiequan Fan^{a,b} Ion Grama^a Quansheng Liu^a

^a *Univ. Bretagne-Sud, UMR 6205, LMBA, F-56000 Vannes, France*

^b *Regularity Team, Inria and MAS Laboratory, Ecole Centrale Paris - Grande Voie des Vignes, 92295 Châtenay-Malabry, France*

Received *****; accepted after revision +++++

Presented by

Abstract

We improve Bernstein's inequality for sums of non-bounded random variables. In particular, we establish a sharp large deviation expansion similar to that of Cramér and Bahadur-Rao. *To cite this article: A. Name1, A. Name2, C. R. Acad. Sci. Paris, Ser. I 340 (2005).*

Résumé

Grandes déviations précises sous la condition de Bernstein. Nous améliorons l'inégalité de Bernstein pour les sommes de variables aléatoires non bornées. En particulier, nous établissons un développement de grandes déviations précises de type Cramér et Bahadur-Rao. *Pour citer cet article : A. Name1, A. Name2, C. R. Acad. Sci. Paris, Ser. I 340 (2005).*

Version française abrégée

Soit $(\xi_i)_{i=1,\dots,n}$ une suite de variables aléatoires (v.a.) indépendantes centrées. Notons $\sigma_i^2 = \mathbb{E}\xi_i^2$, $S_n = \sum_{i=1}^n \xi_i$ et $\sigma^2 = \sum_{i=1}^n \sigma_i^2$. Supposons que $(\xi_i)_{i=1,\dots,n}$ satisfasse la condition de Bernstein [6] suivante, pour une constante $\varepsilon > 0$,

$$|\mathbb{E}\xi_i^k| \leq \frac{1}{2}k!\varepsilon^{k-2}\mathbb{E}\xi_i^2, \quad \text{pour tout } k \geq 2 \text{ et tout } i = 1, \dots, n. \quad (1)$$

Bernstein [6] a démontré que, pour tout $x > 0$,

Email addresses: fanxiequan@hotmail.com (Xiequan Fan), ion.grama@univ-ubs.fr (Ion Grama), quansheng.liu@univ-ubs.fr (Quansheng Liu).

$$\mathbb{P}(S_n > x\sigma) \leq \inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - x\sigma)}. \quad (2)$$

Le but de cette note est de présenter une amélioration de l'inégalité de Bernstein (2) en ajoutant un facteur manquant dans l'esprit de Talagrand [14] sous la condition (1). Dans le théorème qui suit, on obtient un développement de grandes déviations précises de type Cramér [7] et Bahadur-Rao [2].

Théorème 0.1 *Sous la condition de Bernstein, pour tout $0 \leq x < \frac{1}{12} \frac{\sigma}{\varepsilon}$,*

$$\mathbb{P}(S_n > x\sigma) = \inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - x\sigma)} \left(M(x) + 28 \theta R(4x\varepsilon/\sigma) \frac{\varepsilon}{\sigma} \right), \quad (3)$$

où $\sqrt{2\pi}M(x)$ est le ratio de Mills,

$$R(t) = \frac{(1-t+6t^2)^3}{(1-3t)^{3/2}(1-t)^7}, \quad 0 \leq t < \frac{1}{3}, \quad (4)$$

et $|\theta| \leq 1$. En particulier, dans le cas *i.i.d.*, pour tout $0 \leq x = o(\sqrt{n})$, $n \rightarrow \infty$,

$$\left| \mathbb{P}(S_n > x\sigma) - M(x) \inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - x\sigma)} \right| = O\left(\frac{1}{\sqrt{n}} \inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - x\sigma)} \right) \quad (5)$$

donc

$$\frac{\mathbb{P}(S_n > x\sigma)}{M(x) \inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - x\sigma)}} = 1 + o(1). \quad (6)$$

1. Introduction

Let $(\xi_i)_{i=1, \dots, n}$ be a sequence of independent and centered random variables satisfying Bernstein's condition, for a constant $\varepsilon > 0$,

$$|\mathbb{E} \xi_i^k| \leq \frac{1}{2} k! \varepsilon^{k-2} \mathbb{E} \xi_i^2, \quad \text{for all } k \geq 2 \text{ and all } i = 1, \dots, n. \quad (7)$$

Denote by

$$S_n = \sum_{i=1}^n \xi_i \quad \text{and} \quad \sigma^2 = \sum_{i=1}^n \mathbb{E} \xi_i^2. \quad (8)$$

The well-known Bernstein inequality [6] states that, for all $x > 0$,

$$\mathbb{P}(S_n > x\sigma) \leq \inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - x\sigma)}. \quad (9)$$

When $(\xi_i)_{i=1, \dots, n}$ are normal random variables, we find that Bernstein's bound $\inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - x\sigma)} = \exp\{-x^2/2\}$. This suggests that Bernstein's inequality (9) can be substantially refined by adding a missing factor $M(x)$, where $\sqrt{2\pi}M(x)$ is Mills' ratio, to be precise,

$$M(x) = \left(1 - \Phi(x)\right) \exp\left\{\frac{x^2}{2}\right\} = O\left(\frac{1}{x}\right), \quad x \rightarrow \infty, \quad (10)$$

with

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

In the case where the summands ξ_i are assumed to be bounded, result of such type have been obtained by Talagrand (cf. (1.6) of [14]). Talagrand showed that if the random variables ξ_i satisfy $-b \leq \xi_i \leq 1$ for a constant $b \geq 1$, then there exists an universal constant K such that, for all $0 \leq x \leq \frac{\sigma}{Kb}$,

$$\mathbb{P}(S_n > x\sigma) \leq \inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - x\sigma)} \left(M(x) + K \frac{b}{\sigma} \right). \quad (11)$$

Since $M(x) = O(x^{-1})$, $x \rightarrow \infty$, inequality (11) improves on Bernstein's bound $\inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - x\sigma)}$ by adding a factor of order x^{-1} for all $0 < x \leq \frac{\sigma}{Kb}$.

The scope of this paper is to present an improvement of Bernstein's bound (9) under Bernstein's condition (7), in particular, by adding a missing factor in the spirit of Talagrand's inequalities (11) for non-bounded random variables. Our result is similar to the following sharp large deviation results of Cramér and Bahadur-Rao.

In the i.i.d. case, Cramér [7] has established a large deviation expansion under the condition $\mathbb{E} e^{|\xi_1|} < \infty$. For all $0 \leq x = o(\sqrt{n})$, one has

$$\frac{\mathbb{P}(S_n > x\sigma)}{1 - \Phi(x)} = e^{\frac{x^3}{\sqrt{n}} \lambda\left(\frac{x}{\sqrt{n}}\right)} \left[1 + O\left(\frac{1+x}{\sqrt{n}}\right) \right], \quad n \rightarrow \infty, \quad (12)$$

where $\lambda(\cdot) = c_1 + c_2 \frac{x}{\sqrt{n}} + \dots$ is the Cramér series and the values c_1, c_2, \dots depend on the distribution of ξ_1 , see also Petrov [13] and Nagaev [12].

Bahadur-Rao [2], see also Bercu and Rouault [5] for the Ornstein-Uhlenbeck process and Joutard [10,11] for nonparametric estimation, proved the following sharp large deviations similar to (12). Assume Cramér's condition. Then, for given $y > 0$, there is a constant c_y depending on the distribution of ξ_1 and y such that

$$\mathbb{P}\left(\frac{S_n}{n} > y\right) = \frac{\inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - yn)}}{\sigma_y t_y \sqrt{2\pi n}} \left[1 + O\left(\frac{c_y}{n}\right) \right], \quad n \rightarrow \infty, \quad (13)$$

where t_y, σ_y and c_y depend on the distribution of ξ_1 and y .

2. Main result

The following theorem is our main result.

Theorem 2.1 *Assume Bernstein's condition. Then, for all $0 \leq x < \frac{1}{12} \frac{\sigma}{\varepsilon}$,*

$$\mathbb{P}(S_n > x\sigma) = \inf_{\lambda \geq 0} \mathbb{E} e^{\lambda(S_n - x\sigma)} F\left(x, \frac{\varepsilon}{\sigma}\right), \quad (14)$$

where

$$F\left(x, \frac{\varepsilon}{\sigma}\right) = M(x) + 28 \theta R(4x\varepsilon/\sigma) \frac{\varepsilon}{\sigma}, \quad (15)$$

the function R is defined by (4) and $|\theta| \leq 1$. In particular, in the i.i.d. case, for all $0 \leq x = o(\sqrt{n})$, $n \rightarrow \infty$,

$$\left| \mathbb{P}(S_n > x\sigma) - M(x) \inf_{\lambda \geq 0} \mathbb{E}e^{\lambda(S_n - x\sigma)} \right| = O\left(\frac{1}{\sqrt{n}} \inf_{\lambda \geq 0} \mathbb{E}e^{\lambda(S_n - x\sigma)}\right) \quad (16)$$

and thus

$$\frac{\mathbb{P}(S_n > x\sigma)}{M(x) \inf_{\lambda \geq 0} \mathbb{E}e^{\lambda(S_n - x\sigma)}} = 1 + o(1). \quad (17)$$

Notice that $\theta \geq -1$. Equality (14) completes Talagrand's upper bound (11) by giving a sharp lower bound. Since the bounded random variables satisfy Bernstein's condition, Theorem 2.1 generalizes the Talagrand inequality (11).

Some earlier lower bounds on tail probabilities, based on Cramér large deviations, can be found in Arkhangel'skii [1] and Nagaev [12]. In particular, Nagaev has established the following lower bound

$$\mathbf{P}(S_n > x\sigma) \geq \left(1 - \Phi(x)\right) e^{-c_1 x^3 \varepsilon / \sigma} \left(1 - c_2(1+x) \frac{\varepsilon}{\sigma}\right), \quad (18)$$

for two numerical values c_1, c_2 and for all $0 \leq x \leq \frac{1}{25} \frac{\sigma}{\varepsilon}$. It is obvious that (14) is a refinement of Nagaev's bound (18).

Notice that $\inf_{\lambda \geq 0} \mathbb{E}e^{\lambda(S_n - x\sigma)}$ is sometimes written in the form $\exp\{-n\Lambda_n^*(\frac{x\sigma}{n})\}$, where $\Lambda_n^*(x) = \sup_{\lambda \geq 0} \{\lambda x - \frac{1}{n} \log \mathbb{E}e^{\lambda S_n}\}$ is the Fenchel-Legendre transform of the normalized cumulant generating function of S_n . In the i.i.d. case, we call $\Lambda^*(x) = \Lambda_n^*(x)$ the good rate function in the LDP theory (see Dembo and Zeitouni [8]).

To show the relation among equality (14) and the results of Cramér [7] and Bahadur-Rao [2]. Consider the i.i.d. case. Without loss of generality, we take $\sigma_1 = 1$. Our Theorem 2.1 shows that, for all $0 \leq x = o(\sqrt{n})$,

$$\mathbb{P}(S_n > x\sigma) = e^{-n\Lambda^*(x/\sqrt{n})} M(x) \left[1 + O\left(\frac{1+x}{\sqrt{n}}\right)\right], \quad n \rightarrow \infty. \quad (19)$$

Note that the exponential factors of the equalities (19) and (12) are different. Our Theorem 2.1 also shows that, for all $0 \leq y = o(1)$,

$$\mathbb{P}\left(\frac{S_n}{n} > y\right) = e^{-n\Lambda^*(y)} \left[M(y\sqrt{n}) + O\left(\frac{1}{\sqrt{n}}\right)\right], \quad n \rightarrow \infty, \quad (20)$$

and implies that, for given $0 < y = o(1)$,

$$\mathbb{P}\left(\frac{S_n}{n} > y\right) = \frac{e^{-n\Lambda^*(y)}}{y\sqrt{2\pi n}} \left[1 + o(1)\right], \quad n \rightarrow \infty. \quad (21)$$

The advantage of the expansion (21) over the Bahadur-Rao expansion (13) is that we replace the complicated factor t_y and σ_y by the simpler expressions y and 1.

Acknowledgements

We would like to thank the referee for his(her) helpful remarks and suggestions.

References

- [1] A.N. Arkhangel'skii, Lower bounds for probabilities of large deviations for sums of independent random variables. *Theory Probab. Appl.* 34, No. 4 (1989) 565–575.
- [2] R. Bahadur, R. Rao, On deviations of the sample mean, *Ann. Math. Statist.* 31 (1960) 1015–1027.
- [3] G. Bennett, Probability inequalities for the sum of independent random variables, *J. Amer. Statist. Assoc.* 57 (1962) 33–45.
- [4] S.N. Bernstein, *The Theory of Probabilities*, Moscow, Leningrad, 1946.
- [5] B. Bercu, A. Rouault, Sharp large deviations for the Ornstein-Uhlenbeck process, *Theory Probab. Appl.* 46, No. 1 (2006) 1–19.
- [6] S.N. Bernstein, *The Theory of Probabilities*, Moscow, Leningrad, 1946.
- [7] H. Cramér, Sur un nouveau théorème-limite de la théorie des probabilités, *Actualité's Sci. Indust.* 736 (1938) 5–23.
- [8] A. Dembo, O. Zeitouni, *Large deviations techniques and applications* (2nd ed.), Springer, New York, 1998.
- [9] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* 58 (1963) 13–30.
- [10] C. Joutard, Sharp large deviations in nonparametric estimation, *J. Nonparametr. Stat.* 18 (2006) 293–306.
- [11] C. Joutard, Strong large deviations for arbitrary sequences of random variables, *Ann. Inst. Stat. Math.* 65.1 (2013) 49–67.
- [12] S.V. Nagaev, Lower bounds for the probabilities of large deviations of sums of independent random variables. *Theory Probab. Appl.* 46 (2002) 79–102; 728–735.
- [13] V.V. Petrov, *Sums of independent random variables*, Springer-Verlag, Berlin, 1975.
- [14] M. Talagrand, The missing factor in Hoeffding's inequalities, *Ann. Inst. H. Poincaré Probab. Statist.* 31 (1995) 689–702.