JEAN-MARC DESHOUILLERS
GREGORY A. FREIMAN
ALEXANDER A. YUDIN

On bounds for the concentration function. 1

Astérisque, tome 258 (1999), p. 425-436
<http://www.numdam.org/item?id=AST_1999__258__425_0>

Astérisque **258**, 1999, p. 425–436

ON BOUNDS FOR THE CONCENTRATION FUNCTION. 1

by

Jean-Marc Deshouillers, Gregory A. Freiman & Alexander A. Yudin

Abstract. — We give an upper bound for the concentration function of a sum of independent identically distributed integral valued random variables in terms of a lower bound for their tail, under the necessary extra condition that the random variables are not essentially supported in a proper arithmetic progression.

1. Introduction

Let $X_1, \ldots, X_k, \ldots$ be independent real random variables and $S_n = \sum_{k=1}^{n} X_k$. It is well known that, in general, the distribution of S_n spreads out as n grows. When all the X_k's are square-integrable, the relation $\sigma^2(S_n) = \sum_{k=1}^{n} \sigma^2(X_k)$ is a way to express this fact. In the general case, Doeblin and Lévy **[2]** were the first to measure this phenomenon in terms of concentration functions. The concentration function of a real random variable X is defined by

$$Q(X;\lambda) = \sup_t P\{t < X \le t + \lambda\} \text{ for } \lambda \ge 0 \ .$$

The results of Doeblin and Lévy have been successively improved by Kolmogorov **[6]**, Rogozin **[12]** and Kesten **[5]**. Let us quote a corollary to Kesten's result, for the case when the X_k's are identically distributed.

***Theorem* (Kesten [5], Corollary 1, p. 134)).** — *There exists an absolute constant C such that for any set of independent identically distributed random variables $X_1, \ldots, X_n$*

1991 ***Mathematics Subject Classification.*** — 60 E15, 60 E10, 11 B05, 11 Z05.
Key words and phrases. — Sums of discrete random variables; concentration function, DLKRK-inequalities, characteristic function, additive number theory.

J.-M. D: Cette recherche a benificié du soutien du CNRS (UMR 9936, Université Bordeaux 1) et de l'Université Victor Segalen Bordeaux 2.
A. Y: This work was supported in part by RFFI GRANT N 93-01-00260.

and any $0 < \lambda \le 2L$ *we have*

$$Q(S_n; L) \le C \frac{L}{\lambda} \frac{Q(X_1; L)}{\sqrt{n(1 - Q(X_1; \lambda))}} . \tag{1.1}$$

Let us consider the case when the X_k's follow a Cauchy law $\mathcal{C}(1)$,where the Cauchy law $\mathcal{C}(a)$ with parameter $a > 0$ has density $a/\big(\pi(t^2 + a^2)\big)$. One readily sees that for $L = 1$ and $0 < \lambda \le 2$, the right hand side of (1.1) has order of magnitude $(\lambda\sqrt{n})^{-1}$ and is never $o(1/\sqrt{n})$. However, the random variable S_n follows the law $\mathcal{C}(n)$, and so

$$Q(S_n; 1) = \frac{2}{\pi} \arctan\left(\frac{1}{2n}\right) = \frac{1}{n\pi}\big(1 + o(1)\big) .$$

The dispersion (in the standard sense) of S_n is due to the dispersion of the X_k's themselves; but the dispersion of the X_k's is not reflected in a small concentration $Q(X_1; \lambda)$ for *small* λ's, but indeed for *large* λ's: the law of X_1 has a large tail, as can be seen from the fact that X_1 is not integrable.

A connection between the moments of the X_k's and the concentration of their sums has been provided by Esséen **[3]**, who proves that the integrability of $|X_1|^r$ for some $0 < r \le 2$ implies the *lower bound*

$$Q(S_n; L) \ge K(r)L\big(L + (n\mu_r)^{1/r}\big)^{-1} ,$$

where $\mu_r = \inf_a E(|X_1 - a|^r)$ and $K(r)$ is an explicitly given expression that only depends on r.

We aim at giving an *upper bound* for the concentration function of S_n in terms of the *tail* of the distribution of the X_k's. There is however a difficulty that will be better seen on discrete random variables. Let us consider an integer $q > 1$ and two integral valued random variables X_1 and X'_1 such that

$$P\{X_1 = 0\} = P\{X'_1 = 0\} = 1/2 ,$$
$$P\{X'_1 = \ell\} = \begin{cases} P\{X_1 = \ell/q\} \neq 0 & \text{when } q \text{ divides } \ell, \\ 0 & \text{otherwise .} \end{cases}$$

We clearly have $Q(X_1; 1) = Q(X'_1; 1) = 1/2$ and the tail of the distribution of X'_1 is heavier than that of X_1. However, if we consider two sets $X_1, \ldots, X_n$ and $X'_1, \ldots, X'_n$ of n independent identically distributed random variables, their sums S_n and S'_n are such that $Q(S_n; 1) = Q(S'_n; 1)$; we have indeed $P\{S_n = N\} = P\{S'_n = qN\}$ and so

$$Q(S_n; 1) = \max_N P\{S_n = N\} = \max_M P\{S'_n = M\} = Q(S'_n; 1) .$$

We give in this paper an upper bound for the concentration function of a sum of independent identically distributed integral valued random variables in terms of the measure of their tail, under the assumption that the support of the random variables is not essentially contained in a proper arithmetic progression.

Theorem 1. — *Let* $\frac{\log 4}{\log 3} < \sigma < 2$, $\varepsilon > 0$, $A \ge 1$ *and* $a > 0$ *be given real numbers. Let* n *be a positive integer and* $X_1, \ldots, X_n$ *a set of independent identically distributed* *integral valued random variables such that*

$$\max_{q\geq 2} \max_{s \bmod q} \sum_{\ell\equiv s (\bmod q)} P\{X_1 = \ell\} \leq 1 - \varepsilon \ , \tag{1.2}$$

$$\forall L \geq A : Q(X_1; L) \leq 1 - aL^{-\sigma} \ . \tag{1.3}$$

Then we have

$$Q(S_n; 1) \leq cn^{-1/\sigma} \ , \tag{1.4}$$

where c depends on σ, ε, A and a at most.

The main aim of this paper being to illustrate the use of inverse additive results to probability theory, we kept the statement and proof of our main result as simple as possible. We have thus restricted our attention to integral valued random variables, have not considered the general case when $0 < \sigma < 2$, and have not made explicit the dependence of c on the parameters ε, A and a. Let us simply notice here that Theorem 1 is valid under the condition $1 < \sigma < 2$: this depends on the fact that, under iterated applications of Lemma 3, the constant 3^k that arise may be improved to $(4 - \epsilon)^k$, an observation which is basically due to Lev. However, when $\sigma < 1$, new phenomena enter the matter (generalized arithmetic progressions); we shall soon return to this topic.

The statement of Theorem 1 becomes false if condition (1.2) is suppressed. Of course, if the constant c in (1.4) is allowed to depend on the law of X_1, then condition (1.2) is no longer necessary.

The proof of this theorem may be summarized as follows. The concentration $Q(S_n; 1)$ is majorized by the mean value of the modulus of the characteristic function of S_n; this latter is the n-th power of that of X_1, which we call φ, so that the problem reduces to the study of the large values of φ. Here we use two ideas that have been introduced by Freiman, Moskvin and Yudin in **[4]** in the context of local limit theorems. The first one, which can be seen as a consequence of Bochner's theorem, is that $\varphi(t_1 + t_2)$ is large as soon as both $\varphi(t_1)$ and $\varphi(t_2)$ are large. The second one comes from the structure theory of set addition: either the set E of the arguments of the large values of φ is small, or it has a structure. In the first case, φ cannot be too large, and so we get (1.4). It remains to exclude the second case; were it to occur, then, as we shall see, either E would contain the vertices of a regular polygon, which would violate (1.2), or it would contain a large interval around 0, which would contradict (1.3).

Problems of estimating the measure of the set of large values of the characteristic function have also been studied by Arak and Zaitsev **[1]**. This gave them the possibility to solve a famous problem of Kolmogorov on the estimation of the approximation of the n-th convolution of any probability distribution by that of an infinitely divisible law.

As a warm up, and in order to introduce some tools and techniques, we devote the second paragraph to prove a special case of the Doeblin-Lévy-Kolmogorov-Rogozin-Kesten (DLKRK) inequality which stems from the same ideas and follows **[10]**, **[11]**.

The interested reader will find questions of a similar flavour in the classical monographs by Petrov **[9]** and the more recent one by Ledoux and Talagrand **[7]**.

2. A DLKRK inequality for discrete random variables

***Theorem 2* (DLKRK).** — *Let $X_1, \dots, X_n$ be independent identically distributed integral valued random variables, and let S_n be their sum and $p = \max_N P\{X_1 = N\}$. For every integer N, we have*

$$P\{S_n = N\} \leq 40 \frac{p}{\sqrt{n(1-p)}} .$$

Let us start by giving some notation that will be used in this paragraph and the next. We let

$$\begin{aligned} &p_\ell = P\{X_1 = \ell\} \text{ for any } \ell \in \mathbb{Z} , \\ &\varphi(t) = \sum_{\ell \in \mathbb{Z}} p_\ell \exp(2\pi i t \ell) \text{ for } t \in \mathbb{T} = \mathbb{R}/\mathbb{Z} , \\ &E(\theta) = \{t \in \mathbb{T} : |\varphi(t)| \geq \cos\theta\} \text{ for } 0 \leq \theta \leq \pi/2 , \\ &\theta^* \text{ be such that } \cos\theta^* = \min |\varphi(t)| \text{ and } 0 \leq \theta^* \leq \pi/2 . \end{aligned}$$

The proof of Theorem 2 will be based on the following two results, for the first of which we give a sketch of a proof.

***Lemma 1* (cf. [4]).** — *For $\theta_1 \geq 0$, $\theta_2 \geq 0$ and $\theta_1 + \theta_2 \leq \frac{\pi}{2}$, we have*

$$E(\theta_1) + E(\theta_2) \subset E(\theta_1 + \theta_2) .$$

Proof. — For $j = 1, 2$, we consider t_j in $E(\theta_j)$, and let $\alpha_j = \arg^t \varphi(t_j)$ and $\lambda_j = \sqrt{1 - |\varphi(t_{3-j})|^2} e^{-i\alpha_j}$. We use the Cauchy inequality to get an upper bound on

$$|\lambda_1 \varphi(t_1) + \lambda_2 \varphi(-t_2)|^2 = \Big| \sum_\ell \sqrt{p_\ell} \left(\lambda_1 \sqrt{p_\ell} e^{2\pi i \ell t_1} + \lambda_2 \sqrt{p_\ell} e^{-2\pi i \ell t_2} \right) \Big|^2 .$$

□

***Lemma 2* (Macbeath-Kneser Theorem, cf. [8], p. 13-14).** — *Let E_1 and E_2 be two non-empty closed sets in $\mathbb{T}$. We have*

$$|E_1 + E_2| \geq \min(1, |E_1| + |E_2|) ,$$

where $|A|$ represents the Haar measure of A in $\mathbb{T}$.

Proof of Theorem 2. — We may of course assume that p is strictly less that 1 and so θ^* is strictly positive. Our first task is to show that

$$|E(\theta)| \leq 12 \frac{\theta p}{\sqrt{1-p}}, \text{ for } \theta \in]0, \theta^*/2[, \tag{2.1}$$

and

$$|E(\theta)| \leq p/\cos^2\theta, \text{ for } \theta \in]0, \frac{\pi}{2}] . \tag{2.2}$$