

MOMENT INEQUALITIES CONNECTED WITH ACCOMPANYING POISSON LAWS IN ABELIAN GROUPS

I. S. BORISOV

Received 25 November 2002

We obtain exact inequalities which connect moments of some functions of sums of independent random variables taking values in a measurable Abelian group and those for the accompanying infinitely divisible laws. Some applications to empirical processes are studied.

2000 Mathematics Subject Classification: 60B15, 60B11, 60E15.

1. Statement of the main results. Let X_1, X_2, \dots be independent random variables taking values in a measurable Abelian group $(\mathcal{G}, \mathcal{A})$ with respective distributions P_1, P_2, \dots . Moreover, if these random variables are identically distributed (the i.i.d. case), then we denote by P their common distribution. We suppose that $\{0\} \in \mathcal{A}$ and the group operation “+” is measurable. In other words, a sum of arbitrary random variables in $(\mathcal{G}, \mathcal{A})$ is a random variable too.

Denote by $\text{Pois}(\mu)$ the generalized Poisson distribution with the Lévy measure μ

$$\text{Pois}(\mu) := e^{-\mu(\mathcal{G})} \sum_{k=0}^{\infty} \frac{\mu^{*k}}{k!}, \quad (1.1)$$

where μ^{*k} is the k -fold convolution of a finite measure μ with itself; μ^{*0} is the unit mass concentrated at zero. Under the measurability conditions above the convolution is well defined because we can define the convolution of probability (i.e., normed finite) measures.

Put $S_n := \sum_{i \leq n} X_i$. Generalized Poisson distribution with the Lévy measure $\mu := \sum_{i \leq n} P_i$ is called the accompanying Poisson law for S_n (see, e.g., Araujo and Giné [1]). We will denote by τ_μ a random variable having this distribution.

The main goal of this paper is to obtain sharp moment inequalities for some measurable functions of S_n via the analogous moments of the accompanying Poisson law. Results of such kind are connected with the Kolmogorov problem of approximation of the sum distributions by infinitely divisible laws as well as with an improvement of the classical probability inequalities for the sums.

For an arbitrary measurable function f satisfying the condition $\mathbf{E}|f(\tau_\mu)| < \infty$, introduce the following notations:

$$\phi(k) := \mathbf{E}f(S_k), \quad \phi_{m,z}(k) := \mathbf{E}f(S_{m,k} + z), \tag{1.2}$$

where $S_{m,k} := \sum_{i \leq k} X_{m,i}$, $S_{m,0} = S_0 = 0$, and $\{X_{m,i} : i \geq 1\}$ are independent copies of the random variable X_m . We assume that all the sequences $\{X_i\}$, $\{X_{1,i}\}$, $\{X_{2,i}\}, \dots$ are independent. Note that, under the moment condition above, the functions $\phi(k)$ exist and the functions $\phi_{m,z}(k)$ are well defined at least for almost all z with respect to the distribution of $S_{j,k}$ for each $j \neq m$ and integer $k \geq 0$ (for details, see Section 3).

We say that a function $g(k)$ is convex if the difference $\Delta g(k) := g(k+1) - g(k)$ is nondecreasing.

THEOREM 1.1. *Let one of the following two conditions be fulfilled:*

- (a) *the random variables $\{X_i\}$ are identically distributed and $\phi(k)$ is a convex function;*
- (b) *for all z and m , all the functions $\phi_{m,z}(k)$ are convex.*

Then, for each n ,

$$\mathbf{E}f(S_n) \leq \mathbf{E}f(\tau_\mu). \tag{1.3}$$

For the initial random variables which are nondegenerate at zero, let $\{X_i^0\}$ be independent random variables with respective distributions

$$P_i^0 := \mathcal{L}(X_i \mid X_i \neq 0). \tag{1.4}$$

For this sequence, we introduce the notations S_k^0 , $S_{m,k}^0$, $\phi^0(k)$, and $\phi_{m,z}^0(k)$ as above.

PROPOSITION 1.2. *Convexity of the functions $\phi^0(k)$ or $\phi_{m,z}^0(k)$ implies convexity of the functions $\phi(k)$ or $\phi_{m,z}(k)$, respectively. The converse implication is false.*

REMARK 1.3. If the functions in the conditions of the above two assertions are concave, then inequality (1.3) is changed to the opposite. This follows from the well-known connection between convex and concave functions.

A simple sufficient condition for the functions $\phi(k)$ and $\phi_{m,z}(k)$ as well as $\phi^0(k)$ and $\phi_{m,z}^0(k)$ to be convex is as follows: for all $x \in \mathcal{G}$ and all $z, h \in \bigcup_{i \leq n} \text{supp } X_i$, the function f satisfies the inequality

$$f(x+h) - f(x) \leq f(x+h+z) - f(x+z), \tag{1.5}$$

where $\text{supp } X_i$ denotes a measurable subset such that $X_i \in \text{supp } X_i$ with probability 1.

For example, in the i.i.d. case, the convexity (say, of $\phi(k)$) follows easily from (1.5):

$$\phi(k+1) - \phi(k) \leq \mathbf{E}(f(S_{k+2}) - f(S_k + X_{k+2})) = \phi(k+2) - \phi(k+1). \quad (1.6)$$

For the Banach-space-valued summands, the following result is valid.

THEOREM 1.4. *Let \mathcal{G} be a separable Banach space. Suppose that at least one of the following two conditions is fulfilled:*

- (1) *the function f is continuously differentiable in Fréchet sense (i.e., $f'(x)[h]$ is continuous in x for each fixed h), and, for each $x \in \mathcal{G}$ and every $z, h \in \bigcup_{i \leq n} \text{supp } X_i$,*

$$f'(x)[h] \leq f'(x+z)[h]; \quad (1.7)$$

- (2) *$\mathbf{E}X_k = 0$ for all k , f is twice continuously differentiable in Fréchet sense, and $f''(x)[h, h]$ is convex in x for each fixed $h \in \bigcup_{i \leq n} \text{supp } X_i$.*

Then all the functions in the conditions of Theorem 1.1 and in Proposition 1.2 are convex.

COROLLARY 1.5. *If $X_i \geq 0$ a.s. and f is an arbitrary convex function on $[0, \infty)$, then inequality (1.5) is true. Moreover, if X_i are random vectors in \mathbb{R}^k , $k \geq 2$, (as well as in the Hilbert space l_2) with nonnegative coordinates, then the function $f(x) := \|x\|^{2+\alpha}$, where $\|\cdot\|$ is the corresponding Euclidean norm and $\alpha \geq 0$, satisfies inequalities (1.5) and (1.7). For the mean zero Hilbert-space-valued summands, the function $f(x) := \|x\|^\beta$, where $\beta = 2, 4$ or $\beta \geq 6$, satisfies Theorem 1.4(2). Therefore, in these cases, inequality (1.3) holds under the additional necessary restriction $\mathbf{E}|f(\tau_\mu)| < \infty$.*

REMARK 1.6. In the multivariate case, conditions (1.5) and (1.7) are slightly stronger than convexity. In particular, in general, the Euclidean norm does not satisfy these conditions.

REMARK 1.7. There exist functions $f(x)$ which do not satisfy the conditions of Theorem 1.4, but the corresponding functions in Theorem 1.1 and Proposition 1.2 are convex. For example, in the i.i.d. one-dimensional case, we consider the function $f(x) := x^5$ and the centered summands $\{X_i\}$. It is clear that the conditions of Theorem 1.4 are not fulfilled. In this case we have

$$\phi(k) = \mathbf{E} \left(\sum_{i=1}^k X_i \right)^5 = k\mathbf{E}X_1^5 + 10k(k-1)\mathbf{E}X_1^3\mathbf{E}X_1^2. \quad (1.8)$$

Thus, if $\mathbf{E}X_1^3 \geq 0$, then the function $\phi(k)$ (as well as the function $\phi^0(k)$) is convex, otherwise it is concave. In other words, in this case we have various

inequality signs in (1.3) depending on positivity or negativity of the third moment of the summands.

Given a finite measure μ on $(\mathcal{G}, \mathcal{A})$ satisfying the condition $\mu(\{0\}) = 0$, we denote by $\phi_\mu(k)$ the function $\phi(k)$ in (1.2) computed in the i.i.d. case for the summand distribution $\mu(\cdot)/\mu(\mathcal{G})$. Exactness of inequality (1.3) is characterized by the following result.

THEOREM 1.8. *In the i.i.d. case, let the function $\phi_\mu(k)$ be convex. Then*

$$\sup_{n,P} \mathbf{E}f(S_n) = \mathbf{E}f(\tau_\mu) \quad (1.9)$$

whenever the expectation on the right-hand side of (1.9) is well defined, where $\mathcal{L}(\tau_\mu) = \text{Pois}(\mu)$ and the supremum is taken over all n and P such that $nP(A \setminus \{0\}) = \mu(A)$ for all $A \in \mathcal{A}$.

REMARK 1.9. Taking inequality (1.3) into account, we can easily reformulate Theorem 1.8 for the non-i.i.d. case. Perhaps, for the first time, the idea of employing generalized Poisson distributions for constructing upper bounds for moments of the sums was proposed by Prokhorov [25, 26]. In particular, relations (1.3) and (1.9) were obtained by Prokhorov [26] for the functions $f(x) := x^{2m}$ (m is an arbitrary natural number) and $f(x) := \text{ch}(tx)$, $t \in \mathbb{R}$, and for one-dimensional symmetric $\{X_i\}$. Moreover, in the case of mean zero one-dimensional summands, these relations for the functions $f(x) := \exp(hx)$, $h \geq 0$, can be easily deduced from Prokhorov [25] (see also Pinelis and Utev [24]).

The most general result in this direction was obtained by Utev [28] who, in fact, rediscovered and essentially employed some results of Cox and Kemperman [11] regarding lower bounds for moments of sums of independent centered random variables. Under Theorem 1.4(2), he proved extremal equality (1.9) for nonnegative functions $f(x)$ having an exponential majorant. Moreover, he required some additional unnecessary restrictions on the sample Banach space. In our opinion, the corresponding proofs of the present paper are simpler than that of Utev and need no additional restrictions on $f(x)$ and the sample space.

Relations like (1.3) and (1.9) can also be applied for obtaining sharp moment and tail probability inequalities for sums of independent random variables (for details, see Kemperman [17], Pinelis and Utev [23, 24], Utev [27, 28], and Ibragimov and Sharakhmetov [15, 16]).

The above results deal with some type of convexity. However, we can obtain moment inequalities close to those mentioned above without any convexity conditions.

THEOREM 1.10. *In the i.i.d. case, for every nonnegative measurable function f , the following inequality holds:*

$$E f(S_n) \leq \frac{1}{1-p} E f(\tau_\mu), \tag{1.10}$$

where $p := \Pr(X_1 \neq 0)$.

In the non-i.i.d. case, the factor $(1-p)^{-1}$ in (1.10) should be replaced by $\exp(\sum_{i \leq n} p_i)$, where $p_i := \Pr(X_i \neq 0)$.

It is clear that inequality (1.10) provides a sufficiently good upper bound under the so-called Poissonian setting when the summand distributions have large atoms at zero (i.e., the probabilities p_i are small enough). Some particular cases of inequality (1.10) are contained in Araujo and Giné [1] and in Giné et al. [14].

2. Applications to empirical processes. In this section, we formulate some consequences of the above theorems as well as some new analogous results for empirical processes. For the sake of simplicity, we study the empirical processes with one-dimensional time parameter although the results below can be reformulated for empirical processes indexed by subsets of an arbitrary measurable space (moreover, for abstract empirical processes indexed by a family of measurable functions). These results are a basis for the so-called Poissonization method for empirical processes. Sometimes it is more convenient to replace an empirical process under study by the corresponding accompanying Poisson point process having a simpler structure for analysis (e.g., independent “increments”). Some versions of this sufficiently popular and very effective method can be found in many papers. In particular, some probability inequalities connecting the distributions of empirical processes (in various settings) and those of the corresponding Poisson processes are contained in Borisov [3, 4, 5], Einmahl [13], Deheuvels and Mason [12], Giné et al. [14], and others.

Introduce the so-called tail (or local) empirical process on the interval $[0, n]$

$$v_n(t) := nF_n\left(\frac{t}{n}\right), \tag{2.1}$$

where $F_n(\cdot)$ is the empirical distribution function (right-continuous version) based on a sample of size n from the $(0, 1)$ -uniform distribution. We consider v_n as a random variable in the space $LS([0, n])$ which is defined as the linear span of the set of all piecewise constant right-continuous functions on $[0, n]$ with finitely many jumps, endowed with the cylinder σ -field. It is easy to verify that the standard Poisson process $\pi(t)$ on $[0, n]$ (with right-continuous paths) has the accompanying Poisson distribution for v_n in this space.

THEOREM 2.1. *Let $\Phi(\cdot)$ be a convex nonnegative functional on $LS([0, n])$ which is nondecreasing on the subset of all nonnegative functions with respect*

to the standard partial order in function spaces. Suppose that, for each function $x(\cdot) \in \text{LS}([0, n])$, the following relation holds: $\lim_{m \rightarrow \infty} \Phi(x^{(m)}) = \Phi(x)$, where $x^{(m)}(t) = x(\lfloor mt \rfloor / m)$, with $\lfloor \cdot \rfloor$ the integer part of a number. Moreover, if $\mathbf{E}\Phi(\pi) < \infty$, then

$$\mathbf{E}\Phi(v_n) \leq \mathbf{E}\Phi(\pi). \tag{2.2}$$

REMARK 2.2. It is well known that if a convex functional defined on a topological linear space (say, on a Banach space) is bounded in a neighborhood of some point, then it is continuous (see, e.g., Kutateladze [20]). Thus, if the functional in Theorem 2.1 is defined, say, on $L_m([0, n], \lambda)$, where λ is a finite measure, and satisfies the local boundedness condition, then the continuity condition connected with the step functions $x^{(m)}(t)$ can be omitted.

In the sequel, in the case of Banach-space-valued random variables, we consider only continuous convex functionals. For example, the functional $\Phi(x) := \|x\|_m^q \equiv (\int_0^n |x(t)|^m \lambda(dx))^{q/m}$ with arbitrary parameters $m \geq 1$ and $q \geq 1$, where λ is an arbitrary finite measure on $[0, n]$, satisfies the conditions of Theorem 2.1.

Note that the accompanying Poisson process for the centered empirical process $v_n^0(t) := v_n(t) - t$, say, in $L_m([0, n], \lambda)$, differs from the corresponding centered Poisson process. This process can be defined as $\pi^0(t) := \pi(t) - \pi(n)t/n$ and, by analogy with the definition of a Brownian bridge, can be called a Poissonian bridge on $[0, n]$. For such processes, Theorem 1.4(2) can be reformulated as follows.

COROLLARY 2.3. *Let $\Phi(x)$ be a functional on $L_m([0, n], \lambda)$ having convex second Fréchet derivative. Then*

$$\mathbf{E}\Phi(v_n^0) \leq \mathbf{E}\Phi(\pi^0) \tag{2.3}$$

whenever the expectation on the right-hand side of (2.5) exists.

As an example of such a functional we can consider $\Phi(x) := \|x\|_m^{mq}$ for any $m \geq 2$ and $q \geq 3$ or $q = 1, 2$.

If we consider the processes v_n and π as random elements in $\text{LS}([0, \delta n])$, where $\delta < 1$, then the following direct consequence of Theorems 1.10 and 2.1 above and Lemma 3.1 and Corollary 3.4 below holds.

COROLLARY 2.4. *For every measurable functional Φ on $\text{LS}([0, \delta n])$ under the minimal restriction $\mathbf{E}|\Phi(\pi)| < \infty$, the following inequality holds:*

$$\mathbf{E}|\Phi(v_n)| \leq \frac{1}{1-\delta} \mathbf{E}|\Phi(\pi)|. \tag{2.4}$$

Moreover, if $\delta = N/n$, N does not depend on n , and the functional Φ satisfies the conditions of [Theorem 2.1](#), then

$$\sup_n \mathbf{E}\Phi(v_n) = \lim_{n \rightarrow \infty} \mathbf{E}\Phi(v_n) = \mathbf{E}\Phi(\pi). \tag{2.5}$$

Finally, we formulate some useful moment inequalities which deal with one-dimensional projections of the processes $v_n(\cdot)$ and $\pi(\cdot)$. A direct consequence of [Corollary 1.5](#) is as follows.

COROLLARY 2.5. *For every natural n and m and every $t \geq 0$, the following inequality holds:*

$$\mathbf{E}(v_n(t) + x)^m \leq \mathbf{E}(\pi(t) + x)^m, \tag{2.6}$$

where x is arbitrary for even m and $x \geq 0$ for odd m .

In the following assertion which complements this inequality, the above-mentioned convexity conditions need not be fulfilled.

THEOREM 2.6. *For every natural n and m and every $t \geq 0$, the following inequality holds:*

$$|\mathbf{E}(v_n(t) + x)^{2m-1}| \leq \mathbf{E}(\pi(t) + x)^{2m-1}, \tag{2.7}$$

where $x \in [-t, 0)$ is arbitrary.

COROLLARY 2.7. *Let $f(x)$ be an entire function on $[0, \infty)$, that is, an analytic function which admits Taylor expansion at all points with a converging power series on the whole positive half-line. Assume that, for a point $x_0 \geq 0$, the k th derivative of this function at x_0 is nonnegative for each $k \geq 2$. Then, for every $t \geq x_0$,*

$$\mathbf{E}f(v_n(t)) \leq \mathbf{E}f(\pi(t)). \tag{2.8}$$

3. Proof of the results. First we formulate two important lemmas which play a key role in proving the above results.

LEMMA 3.1. *In the i.i.d. case, under the above notations, the following relations hold:*

$$\text{Pois}(n\mathcal{L}(X_1)) = \mathcal{L}(S_{\pi(n)}), \tag{3.1}$$

where the standard Poisson process $\pi(\cdot)$ is independent of $\{X_i\}$,

$$\mathcal{L}(S_n) = \mathcal{L}(S_{v(n,p)}^0), \quad \text{Pois}(n\mathcal{L}(X_1)) = \mathcal{L}(S_{\pi(np)}^0), \tag{3.2}$$

where $p := \mathbf{Pr}(X_1 \neq 0)$, $\mathcal{L}(v(n, p)) = B_{n,p}$ is the binomial distribution with parameters n and p ; the pair $(v(n, p), \pi(np))$ does not depend on the sequence $\{X_i^0\}$.

The relation (3.1) is well known. It immediately follows from the above-mentioned definition of a generalized Poisson distribution: the probability law $\text{Pois}(\mu)$ may be interpreted as the distribution of $\sum_{i \leq \pi_\mu} Y_i$, where $\{Y_i\}$ are i.i.d. random variables with the common distribution $\mu(\cdot)/\mu(\mathcal{G})$ and π_μ is a Poissonian random variable with parameter $\mu(\mathcal{G})$, which is independent of the sequence $\{Y_i\}$.

The equalities in (3.2), which are more convenient in studying accuracy of Poisson approximation of the sums, are contained in various forms in many papers (see, e.g., Khintchine [18, 19], Le Cam [21, 22], Borovkov [9], Borisov [6, 7], and others). Actually, these relations also represent versions of the total probability formula and are easily proven.

Taking into account the representations in Lemma 3.1, we can reduce the problem to the simplest one-dimensional case when we estimate the analogous moments of the binomial distribution introduced in Remark 1.3. However, in this case, we can obtain sufficiently exact inequalities for moments of arbitrary functions using the following lemma.

LEMMA 3.2. *For each $p \in (0, 1)$,*

$$\sup_{n,j} \frac{B_{n,p}(j)}{\mathcal{L}(\pi(np))(j)} \leq \frac{1}{1-p}. \tag{3.3}$$

PROOF. For every nonnegative integer $j \leq n$, we have

$$\begin{aligned} & \frac{\mathbf{P}(\nu(n,p) = j)}{\mathbf{P}(\pi(np) = j)} \\ &= \frac{n(n-1) \cdots (n-j+1)}{n^j(1-p)^j} (1-p)^n e^{np} \\ &= \exp \left\{ n(p + \log(1-p)) - j \log(1-p) + \sum_{i=0}^{j-1} \log \left(1 - \frac{i}{n} \right) \right\} \\ &\leq \exp \left\{ -\log(1-p) + n(p + \log(1-p)) \right. \\ &\quad \left. - (j-1) \log(1-p) + n \int_0^{(j-1)/n} \log(1-x) dx \right\} \\ &\leq \exp \left\{ -\log(1-p) - nH_p \left(\frac{j-1}{n} \right) \right\}, \end{aligned} \tag{3.4}$$

where $H_p(x) = -p + x + (1-x) \log((1-x)/(1-p))$. The following properties of H_p are obvious:

$$H_p(1) = 1-p, \quad H_p(p) = 0, \quad \frac{d}{dx} H_p(p) = 0, \quad \frac{d^2}{dx^2} H_p(x) = \frac{1}{1-x}, \tag{3.5}$$

which implies $H_p(x) \geq 0$ if $x \leq 1$, that is, inequality (3.3) is proven. □

REMARK 3.3. Inequality (3.3) is a part of a more general result in Borisov and Ruzankin (see [8, Lemma 1, page 1663]). It is worth noting that this upper bound is an estimate for the so-called Radon-Nikodym derivative of a binomial distribution with respect to the accompanying Poisson law. This problem was studied by a number of authors (Le Cam [21], Chen [10], Barbour et al. [2], and others). In particular, under some additional restriction on n and p , a slightly stronger estimate is contained in Le Cam [21]. However, in general, estimate (3.3) cannot be essentially improved. Under some restrictions on n and p , a lower bound for the left-hand side of (3.3) has the form $(1 - cp)^{-1}$, where c is an absolute positive constant.

COROLLARY 3.4. *Let g be an arbitrary function satisfying the condition $E|g(\pi(\lambda))| < \infty$ for some λ . Then, for every n and p satisfying the condition $np \leq \lambda$, the following inequality holds:*

$$E|g(v(n, p))| \leq \frac{e^{\lambda - np}}{1 - p} E|g(\pi(\lambda))|. \tag{3.6}$$

Moreover,

$$\lim_{n \rightarrow \infty, np \rightarrow \lambda - 0} E g(v(n, p)) = E g(\pi(\lambda)). \tag{3.7}$$

PROOF. Inequality (3.6) follows from Lemma 3.2 and the simple estimate

$$\sup_j \frac{P(\pi(np) = j)}{P(\pi(\lambda) = j)} \leq e^{\lambda - np}. \tag{3.8}$$

Relation (3.7) follows from the classical Poisson limit theorem and inequality (3.6) which provides fulfillment of the uniform integrability condition. The corollary is proven. \square

REMARK 3.5. Inequality (1.10) in Theorem 1.10 immediately follows from Corollary 3.4 and representations (3.2). In the case $n = 1$ in Lemma 3.2 there exists a slightly stronger upper bound for the Radon-Nikodym derivative. It is easy to see that, in this case, the right-hand side of (3.3) can be replaced by e^p . In the non-i.i.d. case, evaluation of the moment $E f(S_n)$ can be reduced to that for a new function of n independent Bernoulli random variables $v_1(1, p), \dots, v_n(1, p)$ (for details, see the proof of Theorem 1.1 below). In this case, the approximating moment is calculated by independent Poisson random variables $\pi_1(p), \dots, \pi_n(p)$ with the same parameter p . Thus, the corresponding upper bound for the Radon-Nikodym derivative (as well as the corresponding factor on the right-hand side of (1.10)) equals $\exp(\sum_{i \leq n} p_i)$. However, in the special case when $S_n = \sum_{i \leq n} v_i(1, p)$, there exist better upper bounds for this derivative. For example, in this case we can replace the factor $\exp(\sum_{i \leq n} p_i)$ by $(1 - \tilde{p})^{-2}$, where $\tilde{p} = \max\{p_i : i \leq n\}$ (see Barbour et al. [2], Borisov and Ruzankin [8]).

It is worth noting that, under the minimal moment condition above, we cannot replace the one-sided double limit in (3.7) by the classical double limit as well as the condition $np \leq \lambda$ in (3.6) cannot be omitted. For example, the function $g(k) = (1 \vee (k-2))! \lambda^{-k}$ satisfies the above-mentioned moment condition; however, it is easy to prove the relation

$$\lim_{n \rightarrow \infty} \sup_{np \leq \lambda} \mathbf{E}g(v(n, p)) = \infty. \quad (3.9)$$

PROOF OF THEOREM 1.1. In the i.i.d. case, inequality (1.3) is a simple consequence of relation (3.1) and the classical Jensen inequality

$$\mathbf{E}f(\tau_\mu) = \mathbf{E}\phi(\pi(n)) \geq \phi(n) = \mathbf{E}f(S_n). \quad (3.10)$$

In order to prove inequality (1.3) in the non-i.i.d. case, we introduce the sequence of i.i.d. random variables $\{\pi_i : i \geq 1\}$ having Poisson distribution with parameter 1, which is independent of all the sequences of random variables introduced in (1.2) (including the initial random variables). Then we can define the random variable τ_μ in the following way:

$$\tau_\mu := \sum_{m=1}^n S_{m, \pi_m}, \quad (3.11)$$

where $S_{m,k}$ are defined in (1.2). The further reasoning is quite analogous to the above. Put $z_1 := \sum_{m=2}^n S_{m, \pi_m}$. Using the above arguments, we have

$$\mathbf{E}f(\tau_\mu) = \mathbf{E}\mathbf{E}_{z_1} \phi_{1, z_1}(\pi_1) \geq \mathbf{E}\mathbf{E}_{z_1} \phi_{1, z_1}(1) = \mathbf{E}f(X_1 + z_1), \quad (3.12)$$

where the symbol \mathbf{E}_{z_1} denotes the conditional expectation given z_1 . Now we put $z_2 := X_1 + \sum_{m=3}^n S_{m, \pi_m}$. Then, repeating the same calculation, we obtain the estimate

$$\mathbf{E}f(X_1 + z_1) = \mathbf{E}\mathbf{E}_{z_2} \phi_{2, z_2}(\pi_2) \geq \mathbf{E}\mathbf{E}_{z_2} \phi_{2, z_2}(1) = \mathbf{E}f\left(X_1 + X_2 + \sum_{m=3}^n S_{m, \pi_m}\right). \quad (3.13)$$

Continuing the calculations in this way, we finally obtain inequality (1.3). **Theorem 1.1** is proven. \square

PROOF OF PROPOSITION 1.2. The first assertion is easily verified. Indeed, by **Corollary 1.5** and **Lemma 3.1** (see (3.2)) we have

$$\begin{aligned} \phi(k+1) - \phi(k) &= \mathbf{E}\phi^0(v(k+1, p)) - \mathbf{E}\phi^0(v(k, p)) \\ &\leq \phi(k+2) - \phi(k+1). \end{aligned} \quad (3.14)$$

The analogous inequality holds for the functions $\phi_{m,z}(k)$.

In order to prove the second assertion of this proposition, we consider the subclass of random variables satisfying the conditions $EX_1^4 < \infty$, $P(X_1 = 0) \neq 0$, and $EX_1 \neq 0$. Put $f(x) := x^4$. Then

$$\phi(k) = A_k^4 (EX_1)^4 + 6A_k^3 EX_1^2 (EX_1)^2 + 3A_k^2 (EX_1^2)^2 + 4A_k^2 EX_1 EX_1^3 + kEX_1^4, \tag{3.15}$$

where $A_k^m := k(k-1) \cdots (k-m+1)$. The second derivative has the form

$$\phi''(k) = Ak^2 + Bk + C, \tag{3.16}$$

where $A := 12(EX_1)^4$, $B := 36EX_1^2(EX_1^2 - (EX_1)^2)$, and

$$C := 22(EX_1)^4 + 6(EX_1^2)^2 + 8EX_1 EX_1^3 - 36EX_1^2 (EX_1)^2. \tag{3.17}$$

Because of positivity of A and B , the function $\phi(k)$ in (3.15) is convex for $C \geq 0$. If $C < 0$ and at least $\phi''(2) < 0$, then the function $\phi(k)$ replaces concavity with convexity. The same representations with the above comments hold for the function $\phi^0(k)$ (with the replacement of X_1 by X_1^0 and C by C^0 in (3.15) and (3.17)).

Consider the case in which the first moment of X_1^0 is positive and the third moment equals zero. It is clear that we can choose the distribution of X_1^0 so that the constant C^0 will be negative with its absolute value large enough. In this case the function $\phi^0(k)$ will be of the mixed type. For example, we can define this distribution as follows: given a positive constant K , we put $X_1^0 = K$ with probability $8/9$ and $X_1^0 = -2K$ with probability $1/9$. In this case, $EX_1^0 = 6K/9$, $E(X_1^0)^2 = 4K^2/9$, and $C^0 < -K^4$.

Since $EX_1^k = pE(X_1^0)^k$ for each integer k , given the above-mentioned distribution of X_1^0 , we can consider p as a free parameter. Substituting this representation into (3.17) we conclude that, for sufficiently small p (say, $p \leq 0.1$), the constant C will be positive. **Proposition 1.2** is proven. \square

PROOF OF THEOREM 1.4. The first assertion is trivial because, under condition (1), from Taylor's formula we have

$$\begin{aligned} f(x+h) - f(x) &= \int_0^1 f'(x+th)[h]dt \\ &\leq \int_0^1 f'(x+z+th)[h]dt \\ &= f(x+z+h) - f(x+z) \end{aligned} \tag{3.18}$$

for every $x \in \mathcal{G}$ and $z, h \in \bigcup_{i \leq n} \text{supp } X_i$, that is, inequality (1.5) is fulfilled.

To prove the second assertion we need only to prove this in the i.i.d. case because, using the arguments in the proof of **Theorem 1.1** above, we can reduce the problem to the i.i.d. case. It remains to observe that, under condition (2) and given z , the function $f(x+z)$ has convex second derivative with respect to x . So, we prove the assertion in the i.i.d. case. Taking into account continuity

in x of the function $f''(x)[h, h]$ for any fixed h and using Taylor's formula, we have

$$f(S_{k+1}) - f(S_k) = f'(S_k)[X_{k+1}] + \int_0^1 (1-t)f''(S_k + tX_{k+1})[X_{k+1}, X_{k+1}]dt. \tag{3.19}$$

First we average both sides of (3.19) with respect to the distribution of X_{k+1} and use the fact that, for any centered (in Bochner sense) random variable X and an arbitrary linear continuous functional $l(\cdot)$, the equality $El(X) = 0$ holds. Averaging both sides of this identity with respect to the other distributions, we then obtain (with more convenient representation of the remainder in (3.19)) the equality

$$\phi(k+1) - \phi(k) = \frac{1}{2}E f''(S_k + \zeta X_{k+1})[X_{k+1}, X_{k+1}], \tag{3.20}$$

where ζ is a random variable with the density $2(1-t)$ on the unit interval, which is defined on the main probability space and independent of the sequence $\{X_k\}$ (we may assume here that this space is rich enough). It is worth noting that, because of integrability of the left-hand side of (3.19), the expectation on the right-hand side of (3.20) is well defined due to Fubini's theorem. In the i.i.d. case, by the classical Jensen inequality (for the conditional expectation $E_{\zeta, X_{k+2}}$) we finally obtain the inequality we need:

$$\begin{aligned} \phi(k+1) - \phi(k) &= \frac{1}{2}E E_{\zeta, X_{k+2}} f''(S_k + \zeta X_{k+2})[X_{k+2}, X_{k+2}] \\ &\leq \frac{1}{2}E f''(S_{k+1} + \zeta X_{k+2})[X_{k+2}, X_{k+2}] \\ &= \phi(k+2) - \phi(k+1). \end{aligned} \tag{3.21}$$

The proof of convexity of $\phi^0(k)$ and $\phi_{m,z}^0(k)$ is the same because, for the centered initial summands, $EX_k^0 = 0$. The theorem is proven. \square

PROOF OF THEOREM 1.8. Put $n > \mu(\mathcal{G})$ and consider the independent random variables $X_k \equiv X_k(n)$ with the following common distribution:

$$P(A \setminus \{0\}) = \frac{\mu(A)}{n}, \quad P(\{0\}) = 1 - \frac{\mu(\mathcal{G})}{n}. \tag{3.22}$$

Then the corresponding random variables X_i^0 have the common distribution $P^0(A) = \mu(A)/\mu(\mathcal{G})$. Therefore, for each n , by Proposition 1.2 we have the corresponding inequality for the moments under study. It is easy to see that, in this case, the function $\phi^0(k) \equiv \phi_\mu(k)$ does not depend on n and we can apply

Lemma 3.1 and relation (3.7). Thus, we have

$$\lim_{n \rightarrow \infty} \mathbf{E}f(S_n) = \lim_{n \rightarrow \infty} \mathbf{E}\phi_\mu(v(n, p)) = \mathbf{E}\phi_\mu(\pi(\lambda)) = \mathbf{E}f(\tau_\mu), \tag{3.23}$$

where $p = \mu(\mathcal{G})/n$ and $\lambda = \mu(\mathcal{G})$. The theorem is proven. □

PROOF OF THEOREM 1.10. The claim follows immediately from Lemmas 3.1 and 3.2 and Remark 3.5. □

PROOF OF THEOREM 2.1. Because of the monotonicity and continuity conditions, it is sufficient to prove the assertion for any finite-dimensional projections $v_n^{(m)}$ and $\pi^{(m)}$ of the processes under consideration. To this end, we consider an arbitrary nonnegative function $\psi(x_1, \dots, x_k)$ which is convex and increasing in every coordinate x_i . We will study the moment $\mathbf{E}\psi(v_n(t_1), \dots, v_n(t_k))$, where $t_i \in [0, n]$ are arbitrary points and $t_i < t_{i+1}$ for every $i < k$. We will also assume that the corresponding Poisson moment exists.

The following so-called Markov property of $v_n(\cdot)$ is well known: given the quantity $v_n(t)$ (the number of the sample points to the left of t/n), two new samples constituted by the points to the left and to the right of t/n , respectively, are independent and distributed as samples (of the corresponding sizes) from the uniform distributions on $[0, t/n]$ and $[t/n, 1]$, respectively. In other words, given $v_n(t_1)$, the increment $v_n(t_2) - v_n(t_1)$ coincides in distribution with $v_N^*((t_2 - t_1)N/n)$, where $N := n - v_n(t_1)$, and the process $v_n^*(\cdot)$ is an independent copy of $v_n(\cdot)$. Thus, taking into account Corollary 1.5 and convexity and monotonicity of the function $\psi_1(x) := \mathbf{E}\psi(v_n(t_1), \dots, v_n(t_{k-1}), v_n(t_{k-1}) + x)$, we have

$$\begin{aligned} \mathbf{E}\psi(v_n(t_1), \dots, v_n(t_k)) &= \mathbf{E}\mathbf{E}_N \psi_1(v_N^*((t_k - t_{k-1})N/n)) \\ &\leq \mathbf{E}\mathbf{E}_N \psi_1(\pi((t_k - t_{k-1})N/n)) \\ &\leq \mathbf{E}\psi_1(\pi(t_k) - \pi(t_{k-1})), \end{aligned} \tag{3.24}$$

where $\pi(\cdot)$ is a Poisson process independent of $v_n(\cdot)$.

Therefore, we reduced the problem to evaluating the moment of a function of the analogous $(k - 1)$ -dimensional projection $v_n(t_1), \dots, v_n(t_{k-1})$. It remains to observe that the function $\psi_2(x_1, \dots, x_{k-1}) := \mathbf{E}\psi(x_1, \dots, x_{k-1}, x_{k-1} + \pi(t_k) - \pi(t_{k-1}))$ is convex and monotone too. In other words, to prove the assertion we may use induction on k . The theorem is proven. □

PROOF OF THEOREM 2.6. It is clear that, under the above notations, we deal with the random variable $v(n, p)$ having the binomial distribution $B_{n,p}$. First we consider the case $n = 1$.

LEMMA 3.6. For every natural m , the function $g_m(t) := \mathbf{E}(\pi(t) - t)^m$ is a polynomial on $[0, \infty)$ with nonnegative coefficients and the following inequalities

hold:

$$\mathbf{E}(\nu(1, p) + x)^{2m-1} \leq \mathbf{E}(\pi(p) + x)^{2m-1} \tag{3.25}$$

if $x \geq -1$, and

$$|\mathbf{E}(\nu(1, p) + x)^{2m-1}| \leq \mathbf{E}(\pi(p) + x)^{2m-1} \tag{3.26}$$

if $x \geq -p$.

PROOF. The properties of the functions $g_m(t)$ immediately follow from the relation

$$\begin{aligned} g_m(t) &= \sum_{k=0}^{\infty} (k-t)^{m-1} (k-t) \frac{t^k}{k!} e^{-t} \\ &= \sum_{k=1}^{\infty} (k-t)^{m-1} \frac{t^k}{(k-1)!} e^{-t} - t g_{m-1} \\ &= t \mathbf{E}(\pi(t) - t + 1)^{m-1} - t g_{m-1} \\ &= t \sum_{k=0}^{m-2} \frac{(m-1)!}{k!(m-k-1)!} g_k(t), \end{aligned} \tag{3.27}$$

where $m \geq 2$, $g_0(t) \equiv 1$, and $g_1(t) \equiv 0$.

In order to prove (3.25) we first study the case $x = -1$. We have

$$\begin{aligned} \mathbf{E}(\nu(1, p) - 1)^{2m-1} &= p - 1, \\ \mathbf{E}(\pi(p) - 1)^{2m-1} &= -e^{-p} + \sum_{k=2}^{\infty} \frac{(k-1)^{2m-1}}{k!} p^k e^{-p} \\ &> -e^{-p} + \frac{1}{2} \sum_{k=2}^{\infty} \frac{(k-1)^{2m-3}}{(k-2)!} p^k e^{-p} \\ &= -e^{-p} + \frac{p^2}{2} \mathbf{E}(1 + \pi(p))^{2m-3} \\ &> p - 1 - \frac{p^2}{2} + \frac{p^2}{2} \mathbf{E}(1 + \pi(p))^{2m-3} \\ &> p - 1, \end{aligned} \tag{3.28}$$

where $m \geq 2$ (in the case $m = 1$ the assertion is trivial). Inequality (3.25) follows from (3.28) and the analogous inequality for the corresponding derivatives with respect to x (see Corollary 2.5).

To prove (3.26) we need to deduce only the inequality

$$\mathbf{E}(p - \nu(1, p))^{2m-1} \leq \mathbf{E}(\pi(p) - p)^{2m-1}. \tag{3.29}$$

First we assume that $p \leq 1/2$. Then we have

$$\mathbf{E}(p - \nu(1, p))^{2m-1} = p(1-p)(p^{2m-2} - (1-p)^{2m-2}) \leq 0, \quad (3.30)$$

and (3.29) holds because of nonnegativity of the functions $g_m(t)$.

In the case $p > 1/2$ we put $\tilde{\nu}(1, \tilde{p}) := 1 - \nu(1, p)$, where $\tilde{p} := 1 - p$. By (3.29) we then obtain

$$\mathbf{E}(p - \nu(1, p))^{2m-1} = \mathbf{E}(\tilde{\nu}(1, \tilde{p}) - \tilde{p})^{2m-1} \leq g_{2m-1}(\tilde{p}) \leq g_{2m-1}(p) \quad (3.31)$$

due to monotonicity of the functions $g_m(t)$. The lemma is proven. \square

Since $\nu(n, p)$ coincides in distribution with a sum of independent copies of the random variables $\nu(n-1, p)$ and $\nu(1, p)$, the further proof of the theorem can be continued by induction on n (using (3.26) and the binomial formula). The theorem is proven. \square

ACKNOWLEDGMENT. This work was partially supported by the Russian Foundation for Basic Research Grants 00-01-00802 and 03-01-00459 and INTAS Grant 99-01317.

REFERENCES

- [1] A. Araujo and E. M. Giné, *The Central Limit Theorem for Real and Banach Valued Random Variables*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York, 1980.
- [2] A. D. Barbour, L. H. Y. Chen, and K. P. Choi, *Poisson approximation for unbounded functions. I. Independent summands*, Statist. Sinica 5 (1995), no. 2, 749-766.
- [3] I. S. Borisov, *Problem of accuracy of approximation in the central limit theorem for empirical measures*, Siberian Math. J. 24 (1983), no. 6, 833-843 (Russian).
- [4] ———, *Exponential inequalities for the distributions of von Mises and U-statistics*, Probability Theory and Mathematical Statistics, Vol. I (Vilnius, 1989), "Mokslas", Lithuania, 1990, pp. 166-178.
- [5] ———, *Approximation of distributions of von Mises statistics with multidimensional kernels*, Siberian Math. J. 32 (1991), no. 4, 554-566.
- [6] ———, *Strong Poisson and mixed approximations of sums of independent random variables in Banach spaces*, Siberian Adv. Math. 3 (1993), no. 2, 1-13.
- [7] ———, *Poisson approximation of the partial sum process in Banach spaces*, Siberian Math. J. 37 (1996), no. 4, 627-634.
- [8] I. S. Borisov and P. S. Ruzankin, *Poisson approximation for expectations of unbounded functions of independent random variables*, Ann. Probab. 30 (2002), no. 4, 1657-1680.
- [9] K. A. Borovkov, *Refinement of Poisson approximation*, Theory Probab. Appl. 33 (1988), no. 2, 343-347 (Russian).
- [10] L. H. Y. Chen, *An approximation theorem for convolutions of probability measures*, Ann. Probability 3 (1975), no. 6, 992-999.
- [11] D. C. Cox and J. H. B. Kemperman, *Sharp bounds on the absolute moments of a sum of two i.i.d. random variables*, Ann. Probab. 11 (1983), no. 3, 765-771.

- [12] P. Deheuvels and D. M. Mason, *Functional laws of the iterated logarithm for the increments of empirical and quantile processes*, Ann. Probab. **20** (1992), no. 3, 1248–1287.
- [13] J. H. J. Einmahl, *Multivariate Empirical Processes*, CWI Tracts, vol. 32, Centrum voor Wiskunde en Informatica, Amsterdam, 1987.
- [14] E. M. Giné, D. M. Mason, and A. Yu. Zaitsev, *The L_1 -norm density estimator process*, Ann. Probab. **31** (2003), no. 2, 719–768.
- [15] R. Ibragimov and Sh. Sharakhmetov, *On an exact constant for the Rosenthal inequality*, Theory Probab. Appl. **42** (1997), no. 2, 294–302.
- [16] ———, *The exact constant in the Rosenthal inequality for random variables with mean zero*, Teor. Veroyatnost. i Primenen. **46** (2001), no. 1, 134–138 (Russian).
- [17] J. H. B. Kemperman, *On a class of moment problems*, Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif, 1970/1971), Vol. II: Probability Theory, University of California Press, California, 1972, pp. 101–126.
- [18] A. Khintchine, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Julius Springer, Berlin, 1933 (German).
- [19] ———, *Asymptotic Laws of Probability Theory*, United Sci. and Techn. Publ. House of NKTP USSR, Moscow, 1936 (Russian), translated from German.
- [20] S. S. Kutateladze, *Fundamentals of Functional Analysis*, Kluwer Texts in the Mathematical Sciences, vol. 12, Kluwer Academic Publishers, Dordrecht, 1995.
- [21] L. Le Cam, *An approximation theorem for the Poisson binomial distribution*, Pacific J. Math. **10** (1960), no. 4, 1181–1197.
- [22] ———, *On the distribution of sums of independent random variables*, Proc. Internat. Res. Sem., Statist. Lab., Univ. California, Berkeley, Calif, Springer-Verlag, New York, 1965, pp. 179–202.
- [23] I. F. Pinelis and S. A. Utev, *Estimation of the moments of sums of independent random variables*, Theory Probab. Appl. **29** (1985), no. 3, 574–577.
- [24] ———, *Exact exponential bounds for sums of independent random variables*, Theory Probab. Appl. **34** (1989), no. 2, 340–346 (Russian).
- [25] Yu. V. Prokhorov, *An extremal problem in probability theory*, Theory Probab. Appl. **4** (1960), no. 2, 201–203 (Russian).
- [26] ———, *Extremal problems in limit theorems*, Proceedings of the 6th USSR Conference on Probab. Theory and Math. Stat., Gos. Izd-vo Polit. i Nauchn. Literat. Litov. SSR, Lithuania, 1962, pp. 77–84 (Russian).
- [27] S. A. Utev, *Extremal problems in moment inequalities*, Theory Probab. Appl. **29** (1984), no. 2, 411–412.
- [28] ———, *Extremal problems in moment inequalities*, Limit Theorems of Probability Theory, Trudy Inst. Mat., vol. 5, Nauka Sibirsk. Otdel., Novosibirsk, 1985, pp. 56–75 (Russian).

I. S. Borisov: Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia

E-mail address: sibam@math.nsc.ru