mathematics

Article

Poissonization Inequalities for Sums of Independent Random Variables in Banach Spaces with Applications to Empirical Processes

Igor Borisov

Sobolev Institute of Mathematics, 630090 Novosibirsk, Russia; sibam@math.nsc.ru

Abstract: Inequalities are obtained which connect the probability tails and moments of functions of the nth partial sums of independent random variables taking values in a separable Banach space and those for the accompanying infinitely divisible laws. Some applications to empirical processes are studied.

Keywords: sums of independent random variables; moment inequalities; accompanying infinitely divisible law; convex function; empirical process

MSC: 62G08

check for updates

Citation: Borisov, I. Poissonization Inequalities for Sums of Independent Random Variables in Banach Spaces with Applications to Empirical Processes. *Mathematics* **2024**, *12*, 2803. https://doi.org/10.3390/math12182803

Academic Editors: Irina Shevtsova and Victor Korolev

Received: 27 July 2024
Revised: 2 September 2024
Accepted: 4 September 2024
Published: 10 September 2024

1. Introduction and the Main Results

Let $X_1, X_2, \ldots$ be independent random variables (r.v.s) taking values in a separable Banach space $(\mathcal{B}, \|\cdot\|)$ with respective distributions $P_1, P_2, \ldots$. In the i.i.d. case, we will denote by P the common distribution.

$Pois(\mu)$ denotes the compound Poisson distribution with Lévy measure μ:

$$Pois(\mu) := e^{-\mu(\mathcal{B})} \sum_{k=0}^{\infty} \frac{\mu^{*k}}{k!}, \tag{1}$$

where μ^{*k} is the k-fold convolution of a finite measure μ with itself; μ^{*0} is the Dirac measure with the atom at zero. $S_n := \Sigma_{i \le n} X_i$, where $S_0 = 0$ by definition. The compound Poisson distribution with Lévy measure $\mu \equiv \mu_n := \sum_{i \le n} P_i$ is called the accompanying infinitely divisible law for $\mathcal{L}(S_n)$ (see [1]); here and everywhere in the future, the symbol $\mathcal{L}(\xi)$ denotes the distribution of a random variable (r.v.) ξ. We denote by τ_{μ_n} a r.v. having this distribution.

For every natural $m \le n$, let $\{X_{m,i};\, i \ge 1\}$ be independent copies of the random variable X_m. We assume that all the sequences $\{X_i\}$, $\{X_{1,i}\}$, $\{X_{2,i}\}, \ldots$ are independent. Additionally, let $\pi(t), \pi_1(t), \ldots, \pi_n(t)$, $t \ge 0$ be independent Poisson random processes with unit intensity which do not depend on the sequences of r.vs above. From (1), it follows that

$$Pois(P) = \mathcal{L}\Big(S_{\pi(1)}\Big). \tag{2}$$

The characteristic functional of a $\mathcal{B}$-valued r.v. ζ is defined as follows:

$$\varphi_\zeta(l) := \mathbf{E}e^{il(\zeta)}, \quad l \in \mathcal{B}^*,$$

where $l(\cdot)$ is a bounded linear functional on $\mathcal{B}$, i.e., it is an element of the conjugate space $\mathcal{B}^*$. So, one obtains

$$\varphi_{\tau_P}(l) := \mathbf{E}e^{il(\tau_P)} = \exp\{\varphi_{X_1}(l) - 1)\}.$$

Next, the characteristic functional of the accompanying infinitely devisable law is calculated by the formula

$$\varphi_{\tau_{\mu_n}}(l) := \mathbf{E}e^{il(\tau_{\mu_n})} = \exp\left\{\sum_{i=1}^{n}\Big(\varphi_{X_{1,i}}(l) - 1\Big)\right\}. \tag{3}$$

In other words,

$$\mathcal{L}(\tau_{\mu_n}) = \mathcal{L}\left(\sum_{m=1}^{n} S_{m,\pi_m(1)}\right), \tag{4}$$

where the sums $S_{m,\pi_m(1)} := \sum_{i=1}^{\pi_m(1)} X_{m,i}$, $m = 1, \ldots, n$ (with $\sum_{i=1}^{0} = 0$) are independent, and $\mathcal{L}\Big(S_{m,\pi_m(1)}\Big) = Pois(P_m)$ by virtue of (1). In the i.i.d. case, from (3), we obtain that Formula (4) can be rewritten as follows:

$$Pois(\mu_n) \equiv Pois(nP) = \mathcal{L}\Big(S_{\pi(n)}\Big). \tag{5}$$

The main goal of this paper is to obtain upper and lower moment inequalities for some measurable functions of S_n or of the collection $S_1, \ldots, S_n$ via the analogous moments of the accompanying compound Poisson laws as well as to obtain upper bounds for the probability tail of these functionals. Results of this kind are related to Kolmogorov's problem of approximating sums of independent r.vs by various infinitely divisible laws, in particular, by the accompanying ones, as well as with an improvement of the classical probability inequalities for these sums.

In what follows, we consider functions of one or several (say, n) $\mathcal{B}$-valued arguments. In the latter case, we consider functions of the n-variate argument $\bar{z} := (z_1, \ldots, z_n)$ from the new Banach space $\mathcal{B}^n := \mathcal{B} \times \cdots \times \mathcal{B}$ with the norm $\|\bar{z}\|_* := \big(\|z_1\|^2 + \ldots + \|z_n\|^2\big)^{1/2}$. So, for arbitrary Borel functions $f(z)$, $G(\bar{z})$ and $F(\bar{z})$, with $z \in \mathcal{B}$ and $\bar{z} \in \mathcal{B}^n$, introduce the following notation under appropriate moment conditions:

$$\phi(n) := \mathbf{E}f(S_n), \tag{6}$$

$$\Phi_F(\bar{k}_n) := \mathbf{E}F(S_{1,k_1}, \ldots, S_{n,k_n}), \tag{7}$$

$$g_n := \mathbf{E}G(S_1, \ldots, S_n), \tag{8}$$

where $\bar{k}_n := (k_1, \ldots, k_n)$, with $k_j \in Z_+$ (Z_+ is the set of all nonnegative integers). It is clear that the function $\phi(n)$ is a particular example of the function g_n. In turn, the latter function is a particular case of $\Phi_F(\bar{k}_n)$ if $k_1 = \ldots = k_n = 1$.

We say that a function $\psi : Z_+ \to R$ is convex (concave) if the difference $\Delta\psi(n) := \psi(n+1) - \psi(n)$ is a nondecreasing (nonincreasing, respectively) function in n.

The following two theorems, in particular, contain some results from [2] together with some new results.

Theorem 1. *The following assertions are valid*:

1. *For all $z \in \mathcal{B}$ and naturals m, let the functions $\phi_{m,z}(n) := \mathbf{E}f(S_{m,n} + z)$ be convex. Then,*

$$\mathbf{E}f(S_n) \le \mathbf{E}f(\tau_{\mu_n}) \tag{9}$$

provided that the expectation on the right-hand side of this inequality exists.

In the i.i.d. case, inequality (9) holds if only the function $\phi(n)$ is convex.

2. *Let the function $\Phi_F(\bar{k}_n)$ be convex with respect to each coordinate $k_j \in Z_+$. Then, for every vector $\bar{k} \in Z_+^n$,*

$$\Phi_F(\bar{k}_n) \le \mathbf{E}F(\tau^{(1)}_{k_1 P_1}, \ldots, \tau^{(n)}_{k_n P_n}) \tag{10}$$

if the expectation on the right-hand side of (10) *exists, where* $\{\tau^{(j)}_{k_jP_j}, j = 1, \ldots, n\}$ *are independent r.v-s with respective distributions* $\{Pois(k_jP_j), j = 1, \ldots, n\}$.

Remark 1. *If the functions* $\Phi_F(\bar{k}_n)$ *and* $\phi_{m,z}(n)$ *in Theorem 1 are concave, then, inequalities (9) and (10) are changed to the opposite. It follows from the well-known connection between convex and concave functions: concave =* $-$*convex.*

For a r.v. ζ with values in $\mathcal{B}$, $supp\,\zeta$ denotes the minimal closed subset of $\mathcal{B}$ such that $\zeta \in supp\,\zeta$ with probability 1. We need the notion of convexity in the direction determined by a subset of $\mathcal{B}$. We say that a measurable function f is convex in direction $\bigcup_{i\geq 1} suppX_i$ if for all $x \in \bigcup_{i\geq 1} supp\,S_i$ and all $z, h \in \bigcup_{i\geq 1} suppX_i$, this function satisfies the inequality

$$f(x+h) - f(x) \leq f(x+h+z) - f(x+z). \tag{11}$$

Notice that, in the one-dimensional case, convexity in direction R_+ (nonnegative summands) or Z_+ (integer-valued nonnegative summands) is the classical convexity. But in the multivariate case, the convexity of f does not imply the relation (11). As a counterexample, we consider the three-dimensional case and the convex function $f(x_1, x_2, x_3) = \max_{i\leq 3}|x_i|$. Put $x = (1, 0, 0)$, $h = (0, 2, 0)$ and z=(0,0,3). It is clear that inequality (11) for these parameters is false. However, this function satisfies the relation (11) in the direction determined by any one-dimensional subspace of $\mathbb{R}^3$.

Proposition 1. *In the i.i.d. case, let a measurable function* f *satisfy* (11)*. Then, under the moment conditions above, inequality* (9) *holds.*

Example 1. *If* $X_i \geq 0$ *a.s., and* f *is an arbitrary convex function on* $[0, \infty)$*, then, inequalities (9) and (11) are valid.*

We now consider some particular cases of the scheme described in Proposition 1. Let $F^*_n(t)$ be the empirical distribution function based on a sample $\omega_1, \ldots, \omega_n$ from the $[0, 1]$-uniform distribution. Then, the normalized empirical process $\nu_n(t) := nF^*_n(t)$ can be represented as the nth partial sum $\sum\limits_{i=1}^{n} X_i$ of the indicator-type i.i.d. random processes $X_i := I_{\{s:\,s\leq t\}}(\omega_i)$ taking values in a Banach space, say, $L_2[0, 1]$. It is well known that the accompanying compound Poisson r.v. $S_{\pi(n)}$ for this sum is a Poisson random process with intensity n, which coincides in distribution with the Poisson random process $\pi(nt)$, $t \in [0, 1]$. Notice that the finite-dimensional distributions of the random process $\nu_n(\cdot)$ are multinomial. In particular, for each $t \in (0, 1)$, the distribution $\mathcal{L}(\nu_n(t))$ is binomial with parameters (n, t).

As consequences of Proposition 1, we obtain the following two assertions.

Corollary 1. *Let* $f : Z_+ \to R$ *be a convex function. Then, for any* $t \in [0, 1]$,

$$\mathbf{E}f(\nu_n(t)) \leq \mathbf{E}f(\pi(nt)), \tag{12}$$

whenever the right-hand side in (12) *is well-defined.*

Corollary 2. *Let* G *and* f *be nondecreasing convex functions on* R*. Then,*

$$\mathbf{E}G\left(\int\limits_0^1 f(\nu_n(t))\lambda(dt)\right) \leq \mathbf{E}G\left(\int\limits_0^1 f(\pi(nt))\lambda(dt)\right)$$

if the right-hand side of this inequality is well defined, where $\lambda(\cdot)$ *is an arbitrary finite measure on* $[0, 1]$.

One can slightly weaken the convexity property in Corollary 1 by studying power moments of the r.vs under consideration.

Proposition 2. *For every $t \in [0,1]$ and any naturals n and m, the following inequalities hold*:

$$\mathbf{E}(\nu_n(t)+x)^{2m-1} \leq \mathbf{E}(\pi(nt)+x)^{2m-1}, \quad \forall x \geq -n, \tag{13}$$

and

$$|\mathbf{E}(\nu_n(t)+x)^{2m-1}| \leq \mathbf{E}(\pi(nt)+x)^{2m-1}, \quad \forall x \geq -nt. \tag{14}$$

Remark 2. *It is worth noting that, from Corollary 2, one can easily obtain similar inequalities for all even moments and $x \in R$. Additionally, for all $x \geq 0$, inequalities (13) and (14) coincide and also follow from (12). So, the only nontrivial cases in (13) and (14) are $x \in [-n, 0)$ and $x \in [-nt, 0)$, respectively. We note that, for $x < -nt$, the right-hand sides in (13) and (14) may be negative (say, for $m = 1$).*

A direct consequence of Proposition 2 is as follows.

Corollary 3. *Let $f(x)$ be an entire function on $[0,\infty)$.*

1. *Assume that there is a point $x_0 \geq 0$ such that, for all $k \geq 2$, the values of k-th derivatives $f^{(k)}(x_0)$ at the point x_0 are nonnegative. Then, for every $t \in [0,1]$ and all $n \geq x_0$,*

$$\mathbf{E}f(\nu_n(t)) \leq \mathbf{E}f(\pi(nt)) \tag{15}$$

provided that the expectation on the right-hand side of (15) *is well defined.*

2. *Assume that there is a point $x^* \geq 0$ such that*

$$f^*(x) := \sum_{k\geq 0} \frac{|f^{(k)}(x^*)|}{k!}(x-x^*)^k$$

is an entire function on $[0,\infty)$ as well. Then, for every $(0,1]$ and all $n \geq x^*/t$,

$$\mathbf{E}f(\nu_n(t)) \leq \mathbf{E}f^*(\pi(nt)) \tag{16}$$

provided that the expectation on the right-hand side of (16) *is well defined.*

Example 2. *Let $f(x) := x^3 - 3rx^2$, $x \geq 0$, where $r > 0$. Put $x_0 = r$. Then, the conditions in item 1 of Corollary 3 are fulfilled, and inequality (15) is valid for all $n \geq r$. But the function $f(x)$ is convex only for $x \geq r$; otherwise, it is concave.*

Theorem 2. *Suppose that at least one of the following two conditions is fulfilled*:

1. *The function f is continuously differentiable in the Fréchet sense (i.e., $f'(x)[h]$ is continuous in x for each fixed h), and for each $x \in \bigcup_{i\geq 1} supp\, S_i$ and all $z, h \in \bigcup_{i\geq 1} supp\, X_i$,*

$$f'(x+th)[h] \leq f'(x+z+th)[h] \quad \forall t \in [0,1]; \tag{17}$$

2. $\mathbf{E}X_k = 0$ *for all k, f is twice continuously differentiable in the Fréchet sense, and $f''(x)[h,h]$ is convex in x for each fixed $h \in \bigcup_{i\geq 1} supp\, X_i$.*

Then, the function $\phi(n)$ is convex, i.e., inequality (9) *is valid.*

Corollary 4. *If $X_i \geq 0$ a.s. and $f : R_+ \to R$ is a convex function, then, inequality* (11) *is valid. If X_i are random vectors in R^k, $k \geq 2$, or 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* (11) *and* (17)*. For the zero-mean Hilbert-space-valued r.vs X_i, the function $f(x) := \|x\|^\beta$, where $\beta = 2, 4$ or $\beta \geq 6$, satisfies condition 2 of Theorem 2. Therefore, in these cases, inequality* (9) *holds under the restriction $\mathbf{E}|f(\tau_{\mu_n})| < \infty$.*

Remark 3. *There exist functions $f(x)$ which do not satisfy the conditions of Theorem 2 but the corresponding function $\phi(n)$ is convex. For example, in the i.i.d. one-dimensional case, let us consider the function $f(x) := x^5$ and the centered summands $\{X_i\}$. It is clear that the conditions of Theorem 2 are not fulfilled. In this case, we have*

$$\phi(n) = \mathbf{E}(\sum_{i=1}^{n} X_i)^5 = n\mathbf{E}X_1^5 + 10n(n-1)\mathbf{E}X_1^3\mathbf{E}X_1^2,$$

i.e., it is a quadratic function with respect to the variable n. Thus, if $\mathbf{E}X_1^3 \geq 0$, then, the function $\phi(n)$ is convex; otherwise, it is concave. In other words, in this case, we obtain upper and lower Poissonization inequalities in dependence on the sign of the moment $\mathbf{E}X_1^3$.

The exactness of inequality (9) is characterized by the following two assertions.

Corollary 5. *For independent one-dimensional centered r.v.s $\{X_i\}$, consider the function $f(x) := x^3$. Since, for any fixed $z \in R$, the second derivative of the function $F_z(x) = (z + x)^3$ is convex and concave simultaneously, then, by item 2 of Theorem 2,*

$$\mathbf{E}S_n^3 = \mathbf{E}\tau_{\mu_n}^3. \tag{18}$$

Given a finite measure μ on $\mathcal{B}$ satisfying the condition $\mu(\{0\}) = 0$, we denote by $\phi_\mu(n)$ the function $\phi(n)$ in (6) defined in the i.i.d. case for the summand distribution $\mu(\cdot)/\mu(\mathcal{B})$.

Theorem 3 ([2])**.** *In the i.i.d. case, let the function $\phi_\mu(k)$ be convex. Then,*

$$\sup_{n,P} \mathbf{E}f(S_n) = \mathbf{E}f(\tau_\mu) \tag{19}$$

whenever the expectation on the right-hand side of (5) is well defined, where $\mathcal{L}(\tau_\mu) = Pois(\mu)$, and the supremum is taken over all n and P such that $nP(A \setminus \{0\}) = \mu(A)$ for all Borel subsets $A \subseteq \mathcal{B}$.

Remark 4. *Taking inequality (9) into account, we can easily reformulate Theorem 2 for the non-i.i.d. case. The idea of employing compound Poisson distributions for constructing upper bounds for the moments of the sums was first proposed by Prokhorov ([3,4]). In particular, relations (9) and (19) were obtained in [4] for the functions $f(x) := x^{2m}$ (m is an arbitrary natural) and $f(x) := \mathrm{ch}(tx)$, $t \in \mathrm{R}$, in the case of one-dimensional symmetric $\{X_i\}$. Moreover, in the case of zero-mean one-dimensional summands, these relations for the functions $f(x) := \exp(hx)$, $h \geq 0$, can be easily deduced from [3] (see also [5]).*

A more general result in this direction was obtained by Utev [6]. Under condition 2 of Theorem 2, he proved extremal equality (19) for nonnegative functions $f(x)$ having an exponential majorant, using a technique by Kemperman [7]. In our opinion, the proof of item 2 of Theorem 2 (see Section 3) is much simpler than that in [6] and needs no additional restrictions on $f(x)$ and the sample space.

Relations like (9) and (19) can also be applied for obtaining sharp moments and the tail probability inequalities for sums of independent r.vs (for details, see [5–12]).

We now consider the centered empirical point process $\bar{\nu}_n(t) := \nu_n(t) - nt$, $t \in [0,1]$, that one can interpret as a sum of n i.i.d. centered r.vs $X_i := I_{\{s:\, s \leq t\}}(\omega_i) - t$ taking values, say, in the Hilbert space $L_2([0,1], \lambda)$, where $\lambda(\cdot)$ is an arbitrary finite measure on $[0,1]$. The accompanying compound Poisson random process can be represented in the form $\pi_n^o(t) := \pi(nt) - t\pi(n)$, $t \in [0,1]$, which may be called as a “Poissonian bridge” with intensity n on the unit interval. By Corollary 5, we then obtain

$$\mathbf{E}\left(\int_0^1 (\bar{\nu}_n(t))^2\lambda(dt)\right)^\gamma \le \mathbf{E}\left(\int_0^1 (\pi_n^o(t))^2\lambda(dt)\right)^\gamma, \tag{20}$$

where $\gamma = 1, 2$ or $\gamma \ge 3$. If $\lambda(\cdot)$ is the Dirac measure with atom at a point t, then, a univariate analog of inequality (20) is as follows:

$$\mathbf{E}|\bar{\nu}_n(t)|^\gamma \le \mathbf{E}|\pi_n^o(t)|^\gamma, \tag{21}$$

with an arbitrary $\gamma \ge 3$ or $\gamma = 2$ and any $t \in [0,1]$. But compared to (20), we have here less restrictive conditions on γ due to item 2 of Theorem 2. It is clear that we can replace the power functions in (21) with any function f having a convex second derivative under appropriate moment conditions:

$$\mathbf{E}f(\bar{\nu}_n(t)) \le \mathbf{E}f(\pi_n^o(t)). \tag{22}$$

It is interesting to compare inequalities (21) and (22) with (12) and (13) taking Remark 2 into account and setting $x = -nt$ in (13). Put $\bar{\pi}(t) := \pi(t) - t$.

Proposition 3. *For every $t \in [0,1]$ and any even convex function f on R, the following two-sided inequality is valid:*

$$\max\{\mathbf{E}f(t\bar{\pi}(n(1-t)), \mathbf{E}f((1-t)\bar{\pi}(nt))\} \\ \le \mathbf{E}f(\pi_n^o(t)) \le \max\{\mathbf{E}f(\bar{\pi}(n(1-t)), \mathbf{E}f(\bar{\pi}(nt))\} \tag{23}$$

if only the Poissonian moments exist. Moreover, if $t \in [1/2, 1]$, then,

$$\mathbf{E}f(\pi_n^o(t)) \le \mathbf{E}f(\bar{\pi}(nt)). \tag{24}$$

Proposition 4. *For any $x \ge 0$, $t \in [0,1]$ and every natural number m,*

$$\mathbf{E}(\pi_n^o(t) + x)^m \le \mathbf{E}(\bar{\pi}(nt) + x)^m. \tag{25}$$

Thus, inequalities (21)–(25) improve the estimates (12)–(15).

We supplemented Corollary 2 and Theorem 2 with an example of an infinitely dimensional function space $\mathcal{B}$. Let $\mathcal{B} = C[0,1]$, with $||x|| := \sup_{0 \le t \le 1} |x(t)|$. Consider an integral-type functional of the form

$$f(x) := \int_0^1 g(x(t))\lambda(dt), \quad x \in C[0,1],$$

where $g(z)$ is a smooth function on R. In this case, the first two Fréchet derivatives of f are defined as follows:

$$f^{(1)}(x)[h] := \int_0^1 g'(x(t))h(t)\lambda(dt),$$

$$f^{(2)}(x)[h_1, h_2] := \int_0^1 g''(x(t))h_1(t)h_2(t)\lambda(dt), \quad h, h_1, h_2 \in C[0,1].$$

For example, if the continuous random processes $X_i = X_i(t)$, $t \in [0,1]$, are nonnegative and the function g is convex (or the first derivative $g'(x)$ is nondecreasing in the positive direction), then, item 1 of Theorem 2 will be fulfilled. On the other hand, if $X_i(t)$ are centered random processes on $[0,1]$ and the second derivative $g''(z)$ is convex on R, then, item 2 of Theorem 2 will also be satisfied.

We can easily reformulate condition (11) and Theorem 2 for the functions $F(\bar{z})$ in (7) if for any $j = 1, \ldots, n$ and fixed $z_i \in \mathcal{B}$, we put

$$F_{\bar{z},j}(x) := F(z_1, \ldots z_{j-1}, x, z_{j+1}, \ldots, z_n) \tag{26}$$

and, under the conditions of Theorem 2, replace the function $f(x)$ with $F_{\bar{z},j}(x)$ for $\bar{z}$ from an appropriate subset.

Corollary 6. *For every fixed $j = 1,\ldots,n$, let the functions $F_{\bar{z},j}(x)$ satisfy (11) for all $x,h \in \bigcup_{i\geq 1} \operatorname{supp} S_{j,i}$ and $z_k \in \bigcup_{i\geq 1} \operatorname{supp} S_{k,i}$, $k \neq j$. Then, under the moment conditions above, inequality (10) holds.*

In (7), put

$$F(\bar{z}) := G(z_1, z_1 + z_2, \ldots, z_1 + \ldots + z_n). \tag{27}$$

Corollary 7. *Let the functions $F_{\bar{z},j}(x)$, defined in (26) by the function $F(\bar{z})$ in (27), satisfy the conditions of Theorem 2. Then,*

$$g_n \leq \mathbf{E}G(\tau^{(1)}_{P_1}, \tau^{(1)}_{P_1} + \tau^{(2)}_{P_2}, \ldots, \sum_{i=1}^{n} \tau^{(i)}_{P_i}), \tag{28}$$

where the independent r.v-s $\{\tau^{(i)}_{P_i}\}$ are defined in Theorem 1.

The above results deal with some type of convexity. However, one can obtain moment inequalities close to those mentioned above without any convexity conditions. The following result is valid for the r.vs $\{X_i\}$ such that $0 < \mathbf{Pr}(X_i = 0) < 1$ for all i.

Theorem 4. *In the i.i.d. case, for every nonnegative measurable function f, the following inequality holds*:

$$\mathbf{E}f(S_n) \leq \frac{1}{1-p}\mathbf{E}f(\tau_\mu), \tag{29}$$

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

Corollary 8. *For any measurable nonnegative function $F(\bar{z})$ in (7),*

$$\Phi_F(\bar{1}) \leq A_n \mathbf{E}F(\tau^{(1)}_{P_1}, \ldots, \tau^{(n)}_{P_n}), \tag{30}$$

with $A_n := \exp\{\sum_{i=1}^n p_i\}$, where $p_i := \mathbf{Pr}(X_i \neq 0)$. In particular, in the non-i.i.d. case, the factor $(1-p)^{-1}$ in (29) may be replaced with A_n.

For an arbitrary vector $\bar{k}_n \neq \bar{1}$,

$$\Phi_F(\bar{k}_n) \leq A_n^* \mathbf{E}F(\tau^{(1)}_{k_1 P_1}, \ldots, \tau^{(n)}_{k_n P_n}), \tag{31}$$

where $A_n^* := \prod_{i=1}^n (1-p_i)^{-1} < \exp\{\sum_{i=1}^n p_i(1-p_i)^{-1}\}$.

In Theorem 4 and Corollary 8, we do not require the existence of the expectations considering their values on the extended real line. It is clear that, in the non-i.i.d. case, inequalities (29) and (30) provide 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 such that the constant A_n is not too large.

Notice that some particular cases of inequality (29) are contained in [1,13].

Remark 5. *In the case $n = 1$, there exists a slightly better upper bound than that in (29). In this case, the factor $(1-p)^{-1}$ on the right-hand side of (29) can be replaced with e^p. However, in the special case when $S_n = \sum_{i\leq n} \nu^{(i)}_1(p_i)$, where $\{\nu^{(i)}_1(p_i)\}$ are independent Bernoulli r.vs with respective parameters $\{p_i\}$, there exists a better upper bound than that in (29). In this case, we can replace the factor A_n with $(1-\tilde{p})^{-2}$, where $\tilde{p} = \max\{p_i;\, i \leq n\}$ (see [14,15]).*

Corollary 9. *Let g be a nonnegative function satisfying the condition $\mathbf{E}g(\pi(\lambda)) < \infty$ for some λ. Then, for every n and p satisfying the condition $np \leq \lambda$, the following inequality holds*:

$$\mathbf{E}g(\nu_n(p)) \leq \frac{e^{\lambda-np}}{1-p}\mathbf{E}g(\pi(\lambda)). \tag{32}$$

Moreover,

$$\lim_{n\to\infty,\, np\to\lambda-0} \mathbf{E}g(\nu_n(p)) = \mathbf{E}g(\pi(\lambda)). \tag{33}$$

Remark 6. *It is worth noting that, under the minimal moment condition above, we cannot replace the one-sided double limit in (33) with the classical double limit, and the condition $np \leq \lambda$ in (32) cannot be omitted. Moreover, there exists a nonnegative function $g(k)$ (see Section 3) such that $\mathbf{E}g(\pi(\lambda)) < \infty$ and*

$$\lim \sup_{n\to\infty,\, np\to\lambda+0} \mathbf{E}g(\nu_n(p)) = \infty. \tag{34}$$

2. Applications to Empirical Point Processes

In this section, we formulate some consequences of the above theorems as well as some new similar results for empirical point processes indexed by subsets of a measurable space. These processes generalize the scheme of univariate empirical point processes $\nu_n(t)$ from the previous section. These results are a basis for the so-called Poissonization method for generalizing empirical point processes. Sometimes, it is more convenient to replace an empirical point process under study with the corresponding accompanying Poisson point process having a simpler structure for analysis (for example, 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 [13,16–18], etc.

Let $x_1, x_2, \ldots$ be i.i.d. r.vs taking values in an arbitrary measurable space $(\mathfrak{X}, \mathcal{A})$ and having a common distribution P. The empirical point process is introduced as

$$V_n(A) := \sum_{i=1}^{n} I_A(x_i), \quad A \in \mathcal{A}_c,$$

and the accompanying Poisson point process

$$\Pi_n(A) := \sum_{i=1}^{\pi(n)} I_A(x_i), \quad A \in \mathcal{A}_c,$$

where $\mathcal{A}_c \equiv \{A_i\}$ is a countable family of measurable sets, and a standard Poisson random process $\pi(\cdot)$ is independent of the collection $\{x_i\}$.

We will consider these processes as r.vs taking values in the separable Banach space $B_q(\mathcal{A}_c)$ of all functions $Y(\cdot)$ on $\mathcal{A}_c$ such that $\sum_{i\geq 1} |Y(A_i)|^q 2^{-i} < \infty$ for some $q \geq 1$, endowed with the norm

$$\|Y\|_q = \left(\sum_{i\geq 1} \frac{|Y(A_i)|^q}{2^i} \right)^{1/q}.$$

In the case $q = \infty$, we deal with the supnorm $\|Y\|_\infty := \sup_i |Y(A_i)|$. It is clear that the Banach space $B_q(\mathcal{A}_c)$ is isomorphic to the Banach space $L_q[\mathbb{N}, \lambda]$, where $\mathbb{N}$ is the set of natural numbers, and λ is a discrete probability measure on $\mathbb{N}$ with $\lambda(\{k\}) = 2^{-k}$. So, the point process $\Pi_n(\cdot)$ is the accompanying compound Poisson process for the point process $V_n(\cdot)$ in the Banach space $B_q(\mathcal{A}_c)$.

As a direct consequence of Proposition 1, the following assertion is valid.

Corollary 10. *Let $\Phi(\cdot)$ be a measurable functional on $B_1(\mathcal{A}_c)$, which is convex in the positive direction with respect to the standard pointwise partial order in function spaces. Then, under appropriate moment conditions, the following inequality holds*:

$$\mathbf{E}\Phi(V_n) \le \mathbf{E}\Phi(\Pi_n). \tag{35}$$

As examples, one can consider functionals of the form

$$\Phi_{G,f}(Y) := G\left(\sum_{i\ge1} f_i(Y(A_i))\right),$$

where G and $\{f_i\}$ are nondecreasing convex functions on $\mathbb{R}$ provided that $\sum_{i\ge1} |f_i(x)| < \infty$ for all $x \in \mathbb{R}$. For such functionals, it is easy to verify the conditions of Corollary 10 (see the proof of Corollary 2).

By analogy with the univariate case, the centered empirical point process $V_n^o(A) := V_n(A) - nP(A)$ and the corresponding accompanying compound Poisson point process $\Pi_n^o(A) := \Pi_n(A) - nP(A)\Pi_n(\mathfrak{X})$ are introduced. For such processes, the second assertion of Theorem 2 can be reformulated as follows:

Corollary 11. *Let $\Phi(x)$ be a measurable functional on $B_1(\mathcal{A}_c)$ having a convex second Fréchet derivative. Then,*

$$\mathbf{E}\Phi(V_n^o) \le \mathbf{E}\Phi(\Pi_n^o), \tag{36}$$

whenever the expectation on the right-hand side of this inequality exists.

As examples of such functionals, one can cite $\Phi_{G,f}(Y)$.

We now introduce the so-called *restricted empirical point processes.* Let $A_0 \in \mathcal{A}$ and $p := \mathbf{P}(x_1 \in A_0) \in (0,1)$. Consider the restrictions of the point processes $V_n(A)$ and $\Pi_n(A)$ on the set $\mathcal{A}_0 := \{A \in \mathcal{A}_c : A \subseteq A_0\}$, which is denoted by $V_n^*(A)$ and $\Pi_n^*(A)$, respectively. We call these processes $\mathcal{A}_0$-*restricted point processes*. In this case, $V_n^*(A) \equiv S_n = \sum_{i=1}^n X_i$, where $X_i := \{I_A(x_i);\, A \in \mathcal{A}_0\}$, $i = 1,\ldots,n$, are i.i.d. stochastic processes indexed by the elements from the family $\mathcal{A}_0$, with $p := \mathbf{P}(X_1 \ne 0) \in (0,1)$. We may consider $\{X_i\}$ as i.i.d. r.vs taking values in the Banach space $B_\infty(\mathcal{A}_c)$. As a direct consequence of Theorem 4, we then obtain

Corollary 12. *The following inequalities are valid*:

$$\mathcal{L}(V_n^*) \le \frac{1}{1-p}\mathcal{L}(\Pi_n^*), \tag{37}$$

$$\mathbf{E}F(V_n^*) \le \frac{1}{1-p}\mathbf{E}F(\Pi_n^*),$$

where $F(\cdot) \ge 0$ and the expectations take their values on the extended real line.

We now introduce a class of additive statistics of the empirical point processes. Let $\Delta_1, \Delta_2, \ldots$ be a finite or countable measurable partition of the sample space. We assume that $p_i := P(\Delta_i) > 0$ for all i and $p_1 \ge p_2 \ge p_3 \ge \ldots$. Denote $\nu_{in} := V_n(\Delta_i)$. We study a class of additive functionals of the form

$$\Phi_f(V_n) := \sum_{i\ge1} f_{in}(\nu_{in}), \tag{38}$$

where $\{f_{in}\}$ is an array of functions on $\mathbb{Z}_+$, with $\sum_{i\ge1} |f_{in}(0)| < \infty$.

Example 3. *We now give a few examples of additive statistics.*

(1) Given a finite partition $\{\Delta_i;\, i = 1, \ldots, m\}$*, put* $f_{in}(x) := \frac{(x-np_i)^2}{np_i}$*,* $i = 1, \ldots, m$*. Then, we deal with a* χ^2*-statistic of the form*

$$\Phi_{\chi^2}(V_n) = \sum_{i=1}^{m} \frac{(\nu_{in} - np_i)^2}{np_i}.$$

(2) The log-likelihood function can be represented as the following linear functional:

$$\Phi_l(V_n) := \sum_{i=1}^{N} \nu_{in} \log p_i.$$

(3) Fix a countable partition $\{\Delta_i;\, i \geq 1\}$*. Let* $f_{in}(x) \equiv f(x) := I_A(x)$*. (For countable partitions, we assume that* $0 \notin A$*). Then, the functional*

$$\Phi_I(V_n) = \sum_{i \geq 1} I_A(\nu_{in}) \tag{39}$$

is the number of cells Δ_i*, each of them contains a number of the sample points* X_i *from the range defined by a subset* A *of naturals. It is the so-called infinite multinomial scheme of placing particles (balls) in cells (boxes) (for example, see [19–24]).*

For additive functionals (38), one can also obtain Poissonization inequalities using the above mentioned inequalities for restricted empirical point processes. The next theorem is related to estimating the distribution tails of additive functionals (38) via the probability tails of the same functionals of the accompanying Poisson point process $\Pi_n(\cdot)$. The main property of the functionals $\Phi_f(\Pi_n)$ is that they have a structure of sums of independent r.vs.

Theorem 5. *Let* $f_{i\tilde{n}}(\cdot) \geq 0$ *for all* i*. Then, for any* $x > 0$*,*

$$\mathbf{P}(\Phi_f(V_n) \geq x) \leq 2C^* \mathbf{P}(\Phi_f(\Pi_n) \geq x/2), \tag{40}$$

where $C^* := \min_{k \geq 1} \max\Big\{ \big(\sum_{i \leq k} p_i\big)^{-1}, \big(\sum_{i > k} p_i\big)^{-1} \Big\}$*. If, additionally,*

$$\sup_x \sum_{i \leq m} f_{i\tilde{n}}(x) \leq C_{m,n}$$

for some natural number m*, then,*

$$\mathbf{P}(\Phi_f(V_n) \geq x) \leq \left(\sum_{i \leq m} p_i \right)^{-1} \mathbf{P}(\Phi_f(\Pi_n) \geq x - C_{m,n}). \tag{41}$$

Remark 7. *It is worth noting that, in (41), the constant* $C_{m,n}$ *may be interpreted as a level of truncation for the r.v.* $\sum_{i \leq m} f_{i\tilde{n}}(\nu_{i_n})$*. In this case, we should add the probability* $\mathbf{P}\big(\sum_{i \leq m} f_{i\tilde{n}}(\nu_{i_n}) > C_{m,n}\big)$ *to the right-hand side of inequality (41).*

Integrating both sides of inequality (40) in x on the positive half-line, we obtain

Corollary 13. *Under the conditions of Theorem 5, let* F *be a nondecreasing function defined on* $\mathbb{R}_+$*, continuous at zero and* $F(0) = 0$*. If* $\mathbf{E}F(2\Phi_f(\overline{\Pi}_{\tilde{n}})) < \infty$*, then,*

$$\mathbf{E}F(\Phi_f(V_n)) \leq 2C^* \mathbf{E}F(2\Phi_f(\Pi_n)). \tag{42}$$

As an example, consider the functional $\Phi_{I_B}(V_n)$ defined in (39). Then, as a consequence of (41) and Chernoff's upper bound (see [25]) for the probability tail of a sum of independent nonidentically distributed Bernoulli r.vs (the transition from finite sums to series in this case is obvious), we obtain the following result.

Corollary 14 ([24])**.** *Put $M_n(B) := \mathbf{E}\Phi_{I_B}(\Pi_{\bar{n}}) = \sum_{i\geq 1}\mathbf{P}(\pi_{in} \in B)$. Then, the following inequality holds for any $\varepsilon > (M_n(B))^{-1}$:*

$$\mathbf{P}\left(\left|\frac{\Phi_{I_B}(V_n)}{M_n(B)} - 1\right| > \varepsilon\right) \leq 2p_1^{-1}e^{-\frac{\delta^2 M_n(B)}{2+\delta}}, \tag{43}$$

where $\delta := \varepsilon - \frac{1}{M_n(B)}$.

Remark 8. *We note that one can replace the Poissonian mean $M_n(B)$ in (43) with the mean $\mathbf{E}\Phi_{I_B}(\overline{V}_{\bar{n}})$, which differs from $M_n(B)$ by no more than 1 due to Barbour–Hall's estimate of the Poisson approximation to a binomial distribution (see [24,26]). Further, if the condition $M_n(B) \to \infty$ is met as $n \to \infty$, then, from (43), we obtain not only the law of large numbers (already formulated in Corollary 2) but at a certain growth rate of the sequence $M_n(B)$, the strong law of large numbers (SLLN). In particular, if $p_i = Ci^{-1-b}$, then, $M_n(B) \sim C(B)n^{\frac{1}{1+b}}$ for any subset of natural numbers B (see [24]). If in the case $m = 1$, we consider the infinite intervals $B \equiv B_k := \{i : i > k\}$ for any $k \in \mathbb{Z}_+$, then, the SLLN is valid only under the condition $p_i > 0$ for all i. This follows from estimate (43), monotonicity of the functions $I_{B_k}(x)$ and simple arguments in proving SLLN in [27] (see also [21]). Moreover, inequality (43) allows us to estimate the rate of convergence in SLLN. If $M_n(B) \to \infty$, this rate of convergence has the order $O(M_n^{-1/2}(B)\log^{1/2} n)$.*

3. Proofs

In this section, we prove some key assertions formulated in the previous two sections.

Proof of Proposition 1. In the i.i.d. case, the convexity of $\phi(k)$ directly follows from (11):

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

□

Proof of Corollary 2. Denote

$$F_{x,h,z} := \int_0^1 f(x(t) + h(t) + z(t))\lambda(dt),$$

where $x(t)$, $h(t)$ and $z(t)$ are nonnegative measurable bounded functions. Due to the convexity and monotonicity of f, one has

$$F_{x,0,0} \leq F_{x,0,z}, \quad F_{x,h,0} - F_{x,0,0} \leq F_{x,h,z} - F_{x,0,z}.$$

From these inequalities and convexity of G, we immediately obtain

$$G(F_{x,h,0}) - G(F_{x,0,0}) \leq G(F_{x,h,z}) - G(F_{x,0,z}).$$

So, condition (11) is fulfilled. □

Proof of Proposition 2. First, we consider the case $n = 1$. In other words, we deal here with the Bernoulli r.v. $\nu_1(t)$ with parameter t.

Lemma 1. *For every natural m, the following inequalities hold:*

$$\mathbf{E}(\nu_1(t) + x)^{2m-1} \leq \mathbf{E}(\pi(t) + x)^{2m-1}, \quad \forall x \geq -1, \tag{44}$$

and

$$|\mathbf{E}(\nu_1(t) + x)^{2m-1}| \leq \mathbf{E}(\pi(t) + x)^{2m-1}, \quad \forall x \geq -t. \tag{45}$$

Proof. In order to prove (44), we first study the case $x = -1$.. We have

$$\mathbf{E}(\nu_1(t) - 1)^{2m-1} = t - 1,$$

$$\mathbf{E}(\pi(t) - 1)^{2m-1} = -e^{-t} + \sum_{k=2}^{\infty} \frac{(k-1)^{2m-1}}{k!} t^k e^{-t}$$

$$> -e^{-t} + \frac{1}{2} \sum_{k=2}^{\infty} \frac{(k-1)^{2m-3}}{(k-2)!} t^k e^{-t} = -e^{-t} + \frac{t^2}{2} \mathbf{E}(1 + \pi(t))^{2m-3}$$

$$> t - 1 - \frac{t^2}{2} + \frac{t^2}{2} \mathbf{E}(1 + \pi(t))^{2m-3} > t - 1, \tag{46}$$

where $m \geq 2$ (in the case $m = 1$, the assertion is trivial). We have proved inequalities of the form (44) for even moments and all $x \in \mathbb{R}$ (see Corollary 1). Therefore, inequality (44) remains true after derivation of both its sides with respect to x. So, inequality (44) follows from this fact and (46).

Taking inequality (44) into account, we conclude that, to prove (45), it suffices to deduce only the inequality

$$\mathbf{E}(t - \nu_1(t))^{2m-1} \leq \mathbf{E}(\pi(t) - t)^{2m-1}. \tag{47}$$

Denote $g_m(t) := \mathbf{E}(\pi(t) - t)^m$. We need the following recurrent relation for $g_m(t)$ (for details, see [16,28]):

$$g_m(t) = t \sum_{k=0}^{m-2} C_{m-1}^k g_k(t), \tag{48}$$

where $m \geq 2$, $g_0(t) \equiv 1$ and $g_1(t) \equiv 0$. From (48), we conclude that, for all naturals m, the functions $g_m(t)$ are nonnegative and nondecreasing on $[0, 1]$.

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

$$\mathbf{E}(t - \nu_1(t))^{2m-1} = t(1-t)(t^{2m-2} - (1-t)^{2m-2}) \leq 0,$$

and (47) holds because of the nonnegativity of the functions $g_m(t)$.

In the case $t > 1/2$, we consider another Bernoulli r.v. $\tilde{\nu}_1(\tilde{t}) := 1 - \nu_1(t)$, with $\tilde{t} := 1 - t$. By (44), we then obtain

$$\mathbf{E}(t - \nu_1(t))^{2m-1} = \mathbf{E}(\tilde{\nu}_1(\tilde{t}) - \tilde{t})^{2m-1} \leq g_{2m-1}(\tilde{t}) \leq g_{2m-1}(t)$$

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

Since $\nu_n(t)$ coincides in distribution with the sum of independent copies of the r.vs $\nu_{n-1}(t)$ and $\nu_1(t)$, the further proof of the theorem can be continued by induction on n (using (22) and the binomial formula). The proposition is proven. □

Proof of Corollary 3. Due to Fubini's theorem and the Taylor expansion of the function f at the point x_0, the existence of the moment $\mathbf{E}f(\pi(nt))$ implies the equality

$$\mathbf{E}f(\pi(nt)) = \sum_{k \geq 0} \frac{f^{(k)}(x_0)}{k!} \mathbf{E}(\pi(nt) - x_0)^k.$$

So, for all $n \geq x_0$, one can apply inequality (13) for every summand with $k \geq 2$ of the series on the right-hand side of the above identity that yields inequality (15). Here, we have taken into account the fact that

$$\mathbf{E}(\pi(nt) - x_0) = \mathbf{E}(\nu_n(t) - x_0),$$

i.e., the first two summands in the series representations of the expectations $\mathbf{E}f(\pi(nt))$ and $\mathbf{E}f(\nu_n(t))$ coincide.

Inequality (16) is proved similarly, using the estimate (14). □

Remark 9. *Inequality (53) is a part of a more general result in [15]. 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 ([14,29,30], and others). In particular, under some additional restriction on n and p, a slightly stronger estimate is contained in [29]. However, in general, the upper bound (53) cannot be essentially improved. Under some restrictions on n and p, a lower bound for the left-hand side of (53) has the form* $(1-cp)^{-1}$*, where* $c<1$ *is an absolute positive constant. For example, for* $n=1$*, an unimprovable upper bound in (53) equals* $e^p<(1-p)^{-1}$*. It is easy to check that* $e^p>(1-p/2)^{-1}$ *for all* $p\le 1/2$.

Proof of Corollary 8. Taking Remark 9 into account, we have a refinement of estimate (53) in the case $n=1$, and as a consequence, we obtain

$$\mathbf{E}g(\nu_1(p))\le e^p g(\pi(p))$$

for any nonnegative function g. The further arguments of proving estimate (30) are quite similar to those in the proof of Theorem 1 below.

Estimate (31) is a direct consequence of estimate (53) and the arguments above. □

Proof of Corollary 9. Inequality (32) follows from Lemma 2 and the simple estimate

$$\sup_j\frac{\mathbf{P}(\pi(np)=j)}{\mathbf{P}(\pi(\lambda)=j)}=\sup_j\left(\frac{np}{\lambda}\right)^j e^{\lambda-np}\le e^{\lambda-np}$$

if $np\le\lambda$ only. Otherwise, there are no uniform upper bounds for the Radon–Nikodym derivative under consideration.

Relation (33) follows from the classical Poisson limit theorem and inequality (32), which provides fulfillment of the uniform integrability condition. The corollary is proven. □

To prove relation (34), we consider the function $g(k):=(1\vee(k-2)!)\lambda^{-k}$. It is clear that $\mathbf{E}g(\pi(\lambda))<\infty$. Otherwise, we have

$$\mathbf{E}g(\nu_n(p))>\frac{1}{k_n(k_n-1)}\prod_{j\le k_n-1}(1-j/n)\left(\frac{np}{\lambda}\right)^{k_n}\left(1-\frac{np}{n}\right)^{n-k_n},$$

where $n\ge 4$ and $k_n:=[\sqrt{n}]$. Further, it easy to see that, as $n\to\infty$ and $np\to\lambda$,

$$\prod_{j\le k_n-1}(1-j/n)=\exp\left\{-\sum_{j\le k_n-1}j/n+O(k_n^3/n^2)\right\}\sim\exp\{-k_n^2/2n\}\sim e^{-1/2}$$

and

$$\left(1-\frac{np}{n}\right)^{n-k_n}\sim e^{-\lambda}.$$

We now suppose that $np/\lambda=1+n^{-\alpha}$ for some $\alpha<1/2$. Then,

$$\left(\frac{np}{\lambda}\right)^{k_n}\sim e^{k_n n^{-\alpha}}\to\infty,$$

which must be proved. □

Proof of Theorem 1. In the i.i.d. case, inequality (9) is a simple consequence of relation (5) and the classical Jensen inequality:

$$\mathbf{E}f(\tau_{\mu_n})=\mathbf{E}\phi(\pi(n))\ge\phi(n)=\mathbf{E}f(S_n).$$

In order to prove inequality (9) in the non-i.i.d. case, taking formula (4) into account, we put

$$\tau_{\mu_n} := \sum_{m=1}^{n} S_{m,\pi_m(1)}. \tag{49}$$

The further reasoning is quite analogous to the above. Put $z_1 := \sum_{m=2}^{n} S_{m,\pi_m(1)}$. Using the above arguments, we have

$$\mathbf{E}f(\tau_{\mu_n}) = \mathbf{E}\mathbf{E}_{z_1}\phi_{1,z_1}(\pi(1)) \geq \mathbf{E}\mathbf{E}_{z_1}\phi_{1,z_1}(1) = \mathbf{E}f(X_1 + z_1),$$

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

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

Continuing calculations in this way, we obtain inequality (9). Theorem 1 is proven. □

Proof of Theorem 2. The first assertion is trivial because, under condition 1, from Taylor's formula, we have

$$f(x+h) - f(x) = \int_0^1 f'(x+th)[h]dt \leq \int_0^1 f'(x+z+th)[h]dt$$
$$= f(x+z+h) - f(x+z)$$

for every $x \in \mathcal{G}$ and $z, h \in \bigcup_{i \leq n} suppX_i$,, that is, inequality (11) is fulfilled.

To prove the second assertion, we only need to prove this in the i.i.d. case because, using the arguments in proving Theorem 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 a 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

$$f(S_{k+1}) - f(S_k) = f'(S_k)[X_{k+1}] + \int_0^1 (1-t)f''(S_k + tX_{k+1})[X_{k+1}, X_{k+1}]dt. \tag{50}$$

First, we average both sides of (50) with respect to the distribution of X_{k+1} and use the fact that, for any centered (in Bochner sense) r.v. X and an arbitrary linear continuous functional $l(\cdot)$, the equality $\mathbf{E}l(X) = 0$ holds. Averaging both sides of this identity with respect to the other distributions, we then obtain the equality (with a probability interpretation of the remainder in (50))

$$\phi(k+1) - \phi(k) = \frac{1}{2}\mathbf{E}f''(S_k + \zeta X_{k+1})[X_{k+1}, X_{k+1}] = \mathbf{E}f''(S_k + \zeta X_{k+2})[X_{k+2}, X_{k+2}] \tag{51}$$

due to the i.i.d. condition of $\{X_k\}$, where ζ is a r.v. 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 reached enough). It is worth noting that, because of the integrability of the left-hand side of (50), the expectation on the right-hand side of (51) is well defined due to Fubini's theorem. In the i.i.d. case, by Jensen's inequality (for the conditional expectation $\mathbf{E}_{\zeta, X_{k+2}}$), we finally obtain from (51) the inequality we need:

$$\phi(k+1) - \phi(k) = \frac{1}{2}\mathbf{E}\mathbf{E}_{\zeta, X_{k+2}} f''(S_k + \zeta X_{k+2})[X_{k+2}, X_{k+2}]$$
$$\leq \frac{1}{2}\mathbf{E}f''(S_{k+1} + \zeta X_{k+2})[X_{k+2}, X_{k+2}] = \phi(k+2) - \phi(k+1).$$

The theorem is proven. □

Proof of Theorem 4. First, we prove two important lemmas which play a key role in proving the theorem. For the initial r.vs which are nondegenerate at zero, let $\{X_i^0\}$ be independent r.vs with respective distributions

$$\mathcal{L}(X_i^0) := \mathcal{L}(X_i | X_i \neq 0),$$

with $p := \mathbf{Pr}(X_1 \neq 0) \in (0,1)$. Denote $S_m^0 := \sum_{i \leq m} X_i^0$.

Lemma 2 ([2,31])**.** *In the i.i.d. case, under the above notation, the following relations hold*:

$$\mathcal{L}(S_n) = \mathcal{L}(S^0_{\nu_n(p)}), \quad Pois(n\mathcal{L}(X_1)) = \mathcal{L}(S^0_{\pi(np)}), \tag{52}$$

where $\mathcal{L}(\nu_n(p))$ is the binomial distribution with parameters n and p; the pair $(\nu_n(p), \pi(np))$ does not depend on the sequence $\{X_i^0\}$.

The equalities in (52), which are very convenient in studying the accuracy of the Poisson approximation of the sums, are contained in various forms in many papers (see, for example, Refs. [29–35], and others). Actually, these relations also represent versions of the total probability formula and are easily proven.

Taking into account the representations in (52), we can reduce the problem to the simplest one-dimensional case when we estimate moments of a binomial distribution using, for example, convexity arguments as above. However, in this case, we can obtain sufficiently exact inequalities for the moments of arbitrary functions without convexity using the following lemma from [15] (see also [2]). For the convenience of the reader, we reproduce the proof of this assertion.

Lemma 3. *For each $p \in (0,1)$,*

$$\sup_{n,j} \frac{\mathcal{L}(\nu_n(p))(j)}{\mathcal{L}(\pi(np))(j)} \leq \frac{1}{1-p}. \tag{53}$$

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

$$\begin{aligned} \frac{\mathbf{P}(\nu_n(p)=j)}{\mathbf{P}(\pi(np)=j)} &= \frac{n(n-1)\cdots(n-j+1)}{n^j(1-p)^j}(1-p)^n e^{np} \\ &= \exp\left\{n(p+\log(1-p)) - j\log(1-p) + \sum_{i=0}^{j-1}\log\left(1-\frac{i}{n}\right)\right\} \end{aligned}$$

$$\leq \exp\bigg\{-\log(1-p) + n(p+\log(1-p)) - (j-1)\log(1-p)$$

$$+ n\int_0^{(j-1)/n}\log(1-x)dx\bigg\} \leq \exp\left\{-\log(1-p) - nH_p\left(\frac{j-1}{n}\right)\right\},$$

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

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

which implies $H_p(x) \geq 0$ for all $x \leq 1$ due to the convexity of $H(x)$, i.e., inequality (13) is proven. □

Finally, as a consequence of Lemmas 1 and 2, we obtain the following moment inequality for any nonnegative function $g(\cdot)$:

$$\mathbf{E}g(\nu_n(p)) \leq \frac{1}{1-p} g(\pi(np)),$$

and apply this inequality for the conditional expectation $\mathbf{E}_{\{X_i^0\}} f(S^0_{\nu_n(p)})$, given the sequence $\{X_i^0\}$. Theorem 4 is proven. □

Remark 10. *Inequality (53) is a part of a more general result in [15]. 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 ([14,29,30] and others). In particular, under some additional restriction on n and p, a slightly stronger estimate is contained in [29]. However, in general, the upper bound (53) cannot be essentially improved. Under some restrictions on n and p, a lower bound for the left-hand side of (53) has the form* $(1-cp)^{-1}$*, where* $c < 1$ *is an absolute positive constant. For example, for* $n = 1$*, an unimprovable upper bound in* (53) *equals* $e^p < (1-p)^{-1}$*. It is easy to confirm that* $e^p > (1-p/2)^{-1}$ *for all* $p \leq 1/2$*.*

Proof of Corollary 8. Taking Remark 9 into account, we have a refinement of estimate (53) in the case $n = 1$, and as a consequence, we obtain

$$\mathbf{E}g(\nu_1(p)) \leq e^p g(\pi(p))$$

for any nonnegative function g. The further arguments of proving estimate (30) are quite similar to those in the proof of Theorem 1, applying formulas (52) for $n = 1$ and the above inequality for the corresponding conditional expectations. Estimate (31) is a direct consequence of estimate (53) and the arguments above. □

Proof of Corollary 9. Inequality (32) follows from Lemma 2 and the simple estimate

$$\sup_j \frac{\mathbf{P}(\pi(np) = j)}{\mathbf{P}(\pi(\lambda) = j)} = \sup_j \left(\frac{np}{\lambda}\right)^j e^{\lambda - np} \leq e^{\lambda - np}$$

if $np \leq \lambda$ only. Otherwise, there are no uniform upper bounds for the Radon–Nikodym derivative under consideration.

Relation (33) follows from the classical Poisson limit theorem and inequality (32), which provides fulfillment of the uniform integrability condition. The corollary is proven. □

To prove relation (34), we consider the function $g(k) := (1 \vee (k-2)!)\lambda^{-k}$. It is clear that $\mathbf{E}g(\pi(\lambda)) < \infty$. Otherwise, we have

$$\mathbf{E}g(\nu_n(p)) > \frac{1}{k_n(k_n-1)} \prod_{j \leq k_n - 1} (1 - j/n) \left(\frac{np}{\lambda}\right)^{k_n} \left(1 - \frac{np}{n}\right)^{n-k_n},$$

where $n \geq 4$ and $k_n := [\sqrt{n}]$. Further, it is easy to see that, as $n \to \infty$ and $np \to \lambda$,

$$\prod_{j \leq k_n - 1} (1 - j/n) = \exp\left\{-\sum_{j \leq k_n - 1} j/n + O(k_n^3/n^2)\right\} \sim \exp\{-k_n^2/2n\} \sim e^{-1/2}$$

and

$$\left(1 - \frac{np}{n}\right)^{n-k_n} \sim e^{-\lambda}.$$

We now suppose that $np/\lambda = 1 + n^{-\alpha}$ for some $\alpha < 1/2$. Then,

$$\left(\frac{np}{\lambda}\right)^{k_n} \sim e^{k_n n^{-\alpha}} \to \infty,$$

which was to be proved. □

Proof of Theorem 5. For any natural k, we denote

$$\Phi_f^{(k)}(V_n) := \sum_{i \le k} f_{in}(\nu_{in}).$$

It is clear that

$$\mathbf{P}\Big(\Phi_f(V_n) \ge x\Big) \le \mathbf{P}\Big(\Phi_f^{(k)}(V_n) \ge \frac{x}{2}\Big) + \mathbf{P}\Big(\Phi_f(V_n) - \Phi_f^{(k)}(V_n) \ge \frac{x}{2}\Big). \tag{54}$$

In the notation of Theorem 1, let V_n^* be the restriction of the point process V_n to the set $A_0 := \bigcup_{i \le k} \Delta_i$ with a hit probability $p := \sum_{i \le k} p_i$. Under the sign of the first probability of the right-hand side of inequality (54), we replace the point process V_n with V_n^* and use inequality (37) for the distributions of the restrictions of the corresponding point processes under consideration.

The difference

$$\Phi_f(V_n) - \Phi_f^{(k)}(V_n) = \sum_{i > k} f_{in}(\nu_{in})$$

is an additive functional of the restriction of V_n to the set $\overline{A}_0 := \bigcup_{i > k} \Delta_i$ with hitting probability $p := \sum_{i > k} p_i$. For this functional, we also use estimate (37). As a result, taking into account the nonnegativity of all $f_{in}(\cdot)$, from (54) and Theorem 1, we easily obtain

$$\mathbf{P}\Big(\Phi_f(V_n) \ge x\Big) \le \Big(\sum_{i > k} p_i\Big)^{-1} \mathbf{P}\Big(\Phi_f^{(k)}(\Pi_n) \ge \frac{x}{2}\Big)$$

$$+ \Big(\sum_{i \le k} p_i\Big)^{-1} \mathbf{P}\Big(\Phi_f(\Pi_n) - \Phi_f^{(k)}(\Pi_n) \ge \frac{x}{2}\Big) \le 2C^* \mathbf{P}\Big(\Phi_f(\Pi_n) \ge \frac{x}{2}\Big).$$

Inequality (41) is proved similarly:

$$\mathbf{P}\Big(\Phi_f(V_n) \ge x\Big) \le \mathbf{P}\Big(\sum_{i > m} f_{in}(\nu_{in}) \ge x - C_{m,n}\Big) \le \Big(\sum_{i \le m} p_i\Big)^{-1} \mathbf{P}\Big(\Phi_f(\Pi_n) \ge x - C_{m,n}\Big).$$

The theorem is proven. □

Proof of Proposition 3. First, taking into account the fact that the increments of Poissonian processes are independent, we note that the r.v. $\pi_n^o(t) = \pi(nt) - t\pi(n)$ coincides in distribution with the r.v.

$$Y := (1-t)\bar{\pi}(nt) - t\bar{\pi}_1(n(1-t)). \tag{55}$$

Since the r.vs $\bar{\pi}(nt)$ and $\bar{\pi}_1(n(1-t))$ are independent and centered, the lower bound in (23) immediately follows from Jensen's inequality. The upper bound follows from the convexity and evenness of the function f.

Since the centered random process $\bar{\pi}(t)$ has independent increments, the expectation $\mathbf{E}f(\bar{\pi}(t))$ is a nondecreasing function in t in virtue of Jensen's inequality. So, for $t \in [1/2, 1]$, the right-hand side of (23) coincides with $\mathbf{E}f(\bar{\pi}(nt))$. □

Proof of Proposition 4. It is clear that it suffices to consider the case $x = 0$. Moreover, taking inequality (24) into account, we prove estimate (25) for every $t \in [0, 1/2]$ only. The characteristic function of Y has the form

$$\varphi_Y(s) := \exp\{g(s)\},$$

where $g(s) := nte^{i(1-t)s} + n(1-t)e^{-its} - n$. So, the mth moment of Y is calculated by the formula $\mathbf{E}Y^m = i^{-m}\varphi^{(m)}(0)$. We need the multiple differentiation formula of products:

$$(uv)^{(n)} = \sum_{k=0}^{n} C_n^k u^{(k)} v^{(n-k)},$$

with $\psi^{(0)} = \psi$. We then obtain

$$\varphi_Y^{(m)}(0) = (g'(s)\varphi_Y(s))_{s=0}^{(m-1)} = \sum_{k=0}^{m-1} C_{m-1}^k g^{(k+1)}(0)\varphi_Y^{(m-1-k)}(0). \tag{56}$$

A similar representation is valid for the characteristic function of $\bar{\pi}(nt)$:

$$\varphi_{\bar{\pi}(nt)}^{(m)}(0) = \sum_{k=0}^{m-1} C_{m-1}^k f^{(k+1)}(0)\varphi_{\bar{\pi}(nt)}^{(m-1-k)}(0), \tag{57}$$

where $f(s) := nt(e^{is} - 1) - isnt$ and $\varphi_{\bar{\pi}(nt)}(s) := e^{f(s)}$.

We now compare the coefficients in the sums in (56) and (57). One has

$$g^{(k)}(0) = i^k nt(1-t)[(1-t)^{k-1} + (-1)^k t^{k-1}],$$

$$f'(0) = 0, \;\; f^{(k)}(0) = i^k nt, \;\; \forall k \geq 2.$$

Moreover, for each $t \in [0, 1/2]$ and all naturals k,

$$0 \leq (1-t)[(1-t)^{k-1} + (-1)^k t^{k-1}] \leq 1.$$

In other words, for all naturals k,

$$0 \leq i^{-k} g^{(k)}(0) \leq i^{-k} f^{(k)}(0). \tag{58}$$

Since

$$i^{-m}\varphi_Y^{(m)}(0) = \sum_{k=0}^{m-1} C_{m-1}^k i^{-k-1} g^{(k+1)}(0) i^{k+1-m} \varphi_Y^{(m-1-k)}(0) \tag{59}$$

then, all the values $i^{-k}\varphi_Y^{(k)}(0)$ are nonnegative. Therefore, the inequalities

$$i^{-m}\varphi_Y^{(m)}(0) \leq i^{-m}\varphi_{\bar{\pi}(nt)}^{(m)}(0)$$

are easily proved by induction on m using the relations (56)–(59).

Thus, the proposition is proven. □

4. Conclusions

In the present paper, some inequalities were obtained for the distributions of sums of independent B-space-valued r.vs in terms of the accompanying infinitely divisible laws. As consequences of these results, similar inequalities were obtained for the distributions of empirical and accompanying Poisson point processes.

It is worth noting that the above arguments for additive statistics are also transferred to more general additive functionals of the U-statistic structure of empirical group frequencies:

$$U_{f,n}(V_n) := \sum_{1 \leq i_1 < \ldots < i_m} f_{n,i_1,\ldots,i_m}(\nu_{n,i_1}, \ldots, \nu_{n,i_m}),$$

where $\{f_{n,i_1,\ldots,i_m}(\cdot)\}$ is an array of finite functions defined on Z_+^m and satisfying only the restriction

$$\sum_{1 \leq i_1 < \ldots < i_m} |f_{n,i_1,\ldots,i_m}(0, \ldots, 0)| < \infty \quad \forall n,$$

with $\nu_{n,i} := V_n(\Delta_i)$ and finite or countable measurable partition $\{\Delta_i\}$ of $\mathcal{B}$. In this case, the problem is reduced to studying the distribution of the Poissonian version $U_{f,n}(\Pi_n)$ where one can use a martingale approach for estimating its moments and probability tail.

Funding: This study is supported by the Basic Research Program of the Siberian Branch of the Russian Academy of Sciences, project no. FWNF-2024-0001.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable for studies that did not create any new data.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Araujo, A.; Giné, E. *The Central Limit Theorem for Real and Banach Valued Random Variables*; Wiley: New York, NY, USA, 1980.
2. Borisov, I.S. Moment inequalities connected with accompanying Poisson laws in Abelian groups. *Intern. J. Math. Math. Sci.* **2003**, *44*, 2771–2786. [CrossRef]
3. Prokhorov, Y.V. An extremal problem in probability theory. *Theory Probab. Appl.* **1960**, *4*, 201–203. [CrossRef]
4. Prokhorov, Y.V. Extremal problems in limit theorems. In Proceedings of the 6th USSR Conference on Probability Theory and Mathematical Statistics, Vilnius, Lithuania, 5–14 September 1962; pp. 77–84. (In Russian)
5. Pinelis, I.F.; Utev, S.A. Exact exponential bounds for sums of independent random variables. *Theory Probab. Appl.* **1989**, *34*, 340–346. [CrossRef]
6. Utev, S.A. Inequalities for sums of weakly dependent random variables and rates of convergence in the invariance principle. *Mat. Tr.* **1985**, *5*, 56–75. (In Russian)
7. Kemperman, J.H.B. On a class of moments problems. In *Proceedings of the 6th Berkley Symposium on Mathematical Statistics and Probability, Berkley, CA, USA, 9–12 April 1971*; University California Press: Berkley, CA, USA, 1972; Volume 2.
8. Cox, D.C.; Kemperman, J.H.B. Sharp bounds on the absolute moments of a sum of two i.i.d. random variables. *Ann. Probab.* **1983**, *11*, 765–771. [CrossRef]
9. Pinelis, I.F.; Utev, S.A. Estimation of the moments of sums of independent random variables. *Theory Probab. Appl.* **1985**, *29*, 574–577. [CrossRef]
10. Utev, S.A. Extremal problems in moment inequalities. *Theory Probab. Appl.* **1984**, *29*, 411–412.
11. Ibragimov, R.; Sharakhmetov, S. On the exact constant for the Rosentahl inequality. *Theory Probab. Appl.* **1997**, *42*, 294–302. [CrossRef]
12. Ibragimov, R.; Sharakhmetov, S. The exact constant in the Rosentahl inequality for random variables with mean zero. *Theory Probab. Appl.* **2001**, *46*, 134–138. (In Russian)
13. Giné, E.M.; Mason, D.M.; Zaitsev, A.Y. The L_1-norm density estimator process. *Ann. Probab.* **2003**, *31*, 719–768. [CrossRef]
14. Barbour, A.D.; Chen, L.H.Y.; Choi, K.P. Poisson approximation for unbounded functions, I: Independent summands. *Stat. Sin.* **1995**, *5*, 749–766.
15. Borisov, I.S.; Ruzankin, P.S. Poisson approximation for expectation of unbounded functions of independent random variables. *Ann. Probab.* **2002**, *30*, 1657–1680. [CrossRef]
16. Borisov, I.S. Approximation of von Mises statistics with multidimensional kernels. *Siberian Math. J.* **1991**, *32*, 554–566. [CrossRef]
17. Deheuvels, P.; Mason, D.M. Functional laws of the iterated logarithm for increments of empirical and quantile processes. *Ann. Probab.* **1992**, *20*, 1248–1287. [CrossRef]
18. Einmahl, J.H.J. *Multivariate Empirical Processes*; CWI Tract 33; Centrum voor Wiskunde en Informatica: Amsterdam, The Netherlands, 1987.
19. Bahadur, R.R. On the number of distinct values in a large sample from an infinite discrete distribution. *Proceed. Nation. Inst. Sci. India* **1960**, *26A* (Suppl. II), 67–75.
20. Barbour, A.D.; Gnedin, A.V. Small counts in the infinite occupancy scheme. *Electr. J. Probab.* **2009**, *14*, 365–384. [CrossRef]
21. Karlin, S. Central limit theorems for certain infinite urn schemes. *J. Math. Mech.* **1967**, *17*, 373–401. [CrossRef]
22. Darling, D.A. Some limit theorems associated with multinomial trials. In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkley, CA, USA, 21 June–18 July 1965; Volume 2(A), pp. 345–350.
23. Kolchin, V.F.; Sevastyanov, B.A.; Chistyakov, V.P. *Random Assignments*; Nauka: Moscow, Russia, 1976. (In Russian)
24. Borisov, I.S.; Jetpisbaev, M.J. Poissonization principle for additive functionals of empirical point processes. *Mathematics* **2022**, *10*, 4084. [CrossRef]
25. Hagerup, T.; Rüb, C. A guided tour of Chernoff bounds. *Inform. Process. Lett.* **1990**, *33*, 305–308. [CrossRef]
26. Barbour, A.D.; Hall, P. On the rate of Poisson convergence. *Math. Proc. Camb. Philos. Soc.* **1984**, *95*, 473–480. [CrossRef]
27. Loeve, M. *Probability Theory*; Russian Translation; Nauchnaya Literatura: Moscow, Russia, 1962.
28. Kendall, M.G. *The Advanced Theory of Statistics*; Griffin: London, UK, 1943; Volume 1.
29. Le Cam, L. An approximation theorem for the Poisson binomial distribution. *Pacific J. Math.* **1960**, *10*, 1181–1197. [CrossRef]
30. Chen, L.H.Y. An approximation theorem for convolutions of probability measures. *Ann. Probab.* **1975**, *3*, 992–999. [CrossRef]
31. Borisov, I.S. Poisson approximation of the partial sum process in Banach spaces. *Sib. Math. J.* **1996**, *37*, 627–634. [CrossRef]

32. Khintchine, A. *Asymptotishe Gesetze der Wahrscheinlichkeitsrechnung*; Springer: Berlin/Heidelberg, Germany, 1933.
33. Khintchine, A.Y. *Asymptotic Laws of Probability Theory*; Translation from German; United Scientific and Technical Publishing House of NKTP USSR: Moscow/Leningrad, Russia, 1936. (In Russian)
34. Le Cam, L. On the distribution of sums of independent random variables. In *Bernoulli, Bayes, Laplace (Anniversary Volume)*; Springer: Berlin/Heidelberg, Germany; New York, NY, USA, 1965; pp. 179–202.
35. Deheuvels, P.; Pfiefer, D. Poisson approximation of multinomial distributions and point processes. *J. Multivar. Anal.* **1988**, *25*, 65–89. [CrossRef]