mathematics

Article

Poissonization Principle for a Class of Additive Statistics

Igor Borisov * and Maman Jetpisbaev

Laboratory of Probability Theory and Mathematical Statistics, Sobolev Institute of Mathematics, Novosibirsk State University, 630090 Novosibirsk, Russia
* Correspondence: sibam@math.nsc.ru

Abstract: In this paper, we consider a class of additive functionals of a finite or countable collection of the group frequencies of an empirical point process that corresponds to, at most, a countable partition of the sample space. Under broad conditions, it is shown that the asymptotic behavior of the distributions of such functionals is similar to the behavior of the distributions of the same functionals of the accompanying Poisson point process. However, the Poisson versions of the additive functionals under consideration, unlike the original ones, have the structure of sums (finite or infinite) of independent random variables that allows us to reduce the asymptotic analysis of the distributions of additive functionals of an empirical point process to classical problems of the theory of summation of independent random variables.

Keywords: empirical point process; Poisson point process; Poissonization; group frequency; additive functional

MSC: 60F05

Citation: Borisov, I.; Jetpisbaev, M. Poissonization Principle for a Class of Additive Statistics. *Mathematics* **2022**, *10*, 4084. https://doi.org/10.3390/math10214084

Academic Editors: Alexander Tikhomirov and Vladimir Ulyanov

Received: 5 September 2022
Accepted: 29 October 2022
Published: 2 November 2022

1. Introduction

In this paper, we study a class of additive functionals (statistics) of a finite or countable collection of group frequencies constructed by a sample of size n with a finite or countable partition of the sample space. Under broad conditions, it is shown that, as $n \to \infty$, the asymptotic behavior of distributions of the additive functionals under consideration is completely similar to the behavior of distributions of the same functionals of the accompanying Poisson point process. From here it is easy to establish that the above-mentioned weak convergence is equivalent to that for the same additive functionals but with independent group frequencies, which are constructed, respectively, using a finite or countable collection of independent copies of the original sample, when we fix in the i-th partition element only the points from the i-th independent copy of the original sample. In other words, in the case under consideration, we remove the dependence of the initial group frequencies with a multinomial distribution. This phenomenon makes it possible to directly use the diverse tool of the summation theory of independent random variables to study the limiting behavior of the additive statistics being considered.

The structure of this paper is as follows. In Section 2, we introduce the empirical and accompanying Poisson vector point processes and formulate some important results regarding their connection. In Section 3, we introduce a class of additive statistics and give a number of examples. Section 4 contains the main result of the paper, i.e., a duality theorem, which states that an original additive statistic with some normalizing and centering constants weakly converges to a limit if, and only if, their Poisson version with the same normalizing and centering constants weakly converges to the same limit. In Section 5, we discuss some applications of the duality theorem. In Section 6, we present moment inequalities connecting the original additive statistics and their Poisson versions. Section 7 is devoted to asymptotic analysis of first two moments of additive statistics connected with an infinite multinomial urn model. Section 8 contains proofs of all results of the paper. Finally, in Section 9, we summarize the results and discuss some their extensions.

2. Empirical and Poisson Point Processes

Let $\{X_i^{(k)}, i \geq 1\}$, $k = \overline{1,m}$ be a finite set of independent copies of a sequence of independent identically distributed random variables with values in an arbitrary measurable space $(\mathfrak{X}, \mathcal{A})$ and distribution P. For any natural $n_1, \ldots, n_m$, consider m independent empirical point processes based on respective samples $X_1^{(k)}, \ldots, X_{n_k}^{(k)}$, $k = \overline{1,m}$:

$$V_{n_k}^{(k)}(A) := \sum_{i=1}^{n_k} I_A(X_i^{(k)}), \quad k = \overline{1,m}, \quad A \in \mathcal{A}.$$

Define the m independent accompanying Poisson point processes as

$$\Pi_{n_k}^{(k)}(A) := \sum_{i=1}^{\pi_k(n_k)} I_A(X_i^{(k)}), \quad k = \overline{1,m}, \quad A \in \mathcal{A},$$

where $\pi_k(t)$, $k = \overline{1,m}$, are independent standard Poisson processes on the positive half-line, which do not depend on all sequences $\{X_i^{(k)}; i \geq 1\}$, $k = \overline{1,m}$. In other words, $\Pi_{n_k}(A) = V_{\pi_k(n_k)}(A)$ for all $k = \overline{1,m}$. We consider the point processes $V_{n_k}(\cdot)$ and $\Pi_{n_k}(\cdot)$ as stochastic processes with trajectories from the measurable space $(\mathbb{B}^{\mathcal{A}}, \mathcal{C})$ of all bounded functions indexed by the elements of the set $\mathcal{A}$, with the σ-algebra $\mathcal{C}$ of all cylindrical subsets of the space $\mathbb{B}^{\mathcal{A}}$. The distributions of stochastic processes $V_{n_k}(\cdot)$ and $\Pi_{n_k}(\cdot)$ on $\mathcal{C}$ are defined in a standard way.

Now, we introduce the vector-valued empirical and accompanying Poisson point processes

$$\overline{V}_{\bar{n}}(A) := (V_{n_1}^{(1)}(A), \ldots, V_{n_m}^{(m)}(A)) \equiv \overline{V}_{\bar{n}},$$

$$\overline{\Pi}_{\bar{n}}(A) := (\Pi_{n_1}^{(1)}(A), \ldots, \Pi_{n_m}^{(m)}(A)) \equiv \overline{\Pi}_{\bar{n}},$$

where $\bar{n} = (n_1, n_2, \ldots, n_m)$. The vector-valued point processes $\overline{V}_{\bar{n}}$ and $\overline{\Pi}_{\bar{n}}$ are considered as random elements with values in the measurable space $((\mathbb{B}^{\mathcal{A}})^m, \mathcal{C}^m)$.

Let $A_0 \in \mathcal{A}$ with $p := P(A_0) \in (0,1)$. Consider the restrictions of the vector point processes $\overline{V}_{\bar{n}}$ and $\overline{\Pi}_{\bar{n}}$ to the set

$$\mathcal{A}_0 := \{A \in \mathcal{A} : A \subseteq A_0\}. \tag{1}$$

These so-called $\mathcal{A}_0$-restrictions are denoted by $\overline{V}_{\bar{n}}^0$ and $\overline{\Pi}_{\bar{n}}^0$, respectively. For the distributions $\mathcal{L}(\overline{V}_{\bar{n}}^0)$ and $\mathcal{L}(\overline{\Pi}_{\bar{n}}^0)$ in the measurable space $((\mathbb{B}^{\mathcal{A}})^m, \mathcal{C}^m)$, there are the following three assertions (some particular versions of these assertions have been proved in [1,2]).

Theorem 1. *The following inequality is valid*:

$$\mathcal{L}(\overline{V}_{\bar{n}}^0) \leq \frac{1}{(1-p)^m} \mathcal{L}(\overline{\Pi}_{\bar{n}}^0). \tag{2}$$

Corollary 1. *For any non-negative measurable functional F defined on $((\mathbb{B}^{\mathcal{A}})^m, \mathcal{C}^m)$,*

$$\mathbf{E}F(\overline{V}_{\bar{n}}^0) \leq \frac{1}{(1-p)^m} \mathbf{E}F(\overline{\Pi}_{\bar{n}}^0); \tag{3}$$

the expectation on the right-hand side of (3) may be infinite at that.

The following result plays an essential role in proving the main result of the paper—a duality limit theorem for the distributions $\mathcal{L}(\overline{V}_{\bar{n}})$ and $\mathcal{L}(\overline{\Pi}_{\bar{n}})$ (see Theorem 3 below).

Theorem 2. *For each multi-index $\bar{n}$, one can define some vector point processes $\overline{V}_{\bar{n}}^{0*}$ and $\overline{\Pi}_{\bar{n}}^{0*}$ on a common probability space so that they coincide in distribution with the point processes $\overline{V}_{\bar{n}}^0$ and $\overline{\Pi}_{\bar{n}}^0$, respectively, and*

$$\sup_{\mathcal{A}_c \subseteq \mathcal{A}_0} \mathbf{P}\left(\sup_{A \in \mathcal{A}_c} \left\| \overline{V}^{0*}_{\bar{n}}(A) - \overline{\Pi}^{0*}_{\bar{n}}(A) \right\| \neq 0 \right) \leq 1 - (1-p)^m < mp, \tag{4}$$

where $\|(z_1, \ldots, z_m)\| := \max_{k \leq m} |z_k|$, *and the outer supremum is taken over all at most countable families* $\mathcal{A}_c$ *of sets from* $\mathcal{A}_0$.

Remark 1. *In Theorem 2, the sup-seminorm* $\sup_{A \in \mathcal{A}_c} \| \cdot \|$ *is obviously measurable with respect to the cylindrical* σ*-algebra* $\mathcal{C}^m$. *If instead of* $\mathcal{A}_c$ *we substitute the entire class* $\mathcal{A}_0$ *(possibly uncountable) then this measurability may no longer exist (unless, of course, the point processes under consideration do not have the separability property). Nevertheless, the assertion of Theorem 2 remains valid in this case if the probability* $\mathbf{P}$ *is replaced by the outer probability* $\mathbf{P}^*(N_o) := \inf_{N \in \mathcal{C}^m : N \supseteq N_o} \mathbf{P}(N)$. *However, the outer probability has only the property of semiadditivity, which makes it difficult to use.*

Let measurable sets $\Delta_1, \Delta_2, \ldots$ *form a finite or countable partition of the sample space under the condition* $p_i := P(\Delta_i) > 0$ *for all* i. *Without loss of generality, we can assume that the sequence* $\{p_i\}$ *is monotonically nonincreasing. Denoted by* $\nu^{(k)}_{n_k 1}, \nu^{(k)}_{n_k 2}, \ldots$, $k = \overline{1,m}$, *the corresponding group frequencies are defined by the sample* $X^{(k)}_1, \ldots, X^{(k)}_{n_k}$. *Put*

$$\bar{v}_{i\bar{n}} := \overline{V}_{\bar{n}}(\Delta_i) = \left(\nu^{(1)}_{n_1 i}, \ldots, \nu^{(m)}_{n_m i} \right), \;\; i = 1, 2, \ldots.$$

Let us agree that everywhere below the limit relation $\bar{n} \to \infty$ *will be understood as* $n_k \to \infty$ *for all* $k = \overline{1,m}$.

3. Additive Statistics: Examples

In the paper, we consider a class of additive statistics of the form

$$\Phi_f(\overline{V}_{\bar{n}}) := \sum_{i \geq 1} f_{i\bar{n}}(\bar{v}_{i\bar{n}}), \tag{5}$$

where $f \equiv \{f_{i\bar{n}}\}$ is an array of arbitrary finite functions defined on $\mathbb{Z}^m_+$ under the condition

$$\sum_{i \geq 1} |f_{i\bar{n}}(0, \ldots, 0)| < \infty \;\; \forall n, \tag{6}$$

which ensures the correct definition of the functional $\Phi_f(\overline{V}_{\bar{n}})$ in the case of a countable partition of the sample space, since the sum under consideration contains only a finite set of nonzero random vectors $\bar{v}_{i\bar{n}}$. In the case of a finite partition and $m = 1$, additive functionals of the form (5) were considered in [3–5].

We now give some examples of such statistics.

(1) Consider a finite partition $\{\Delta_i;\, i = 1, \ldots, N\}$ of the sample space. Put $f_{i\bar{n}}(\bar{x}) := \frac{|\bar{x} - \bar{n} p_i|^2}{|\bar{n} p_i|}$, $i = 1, \ldots, N$, where $|\cdot|$ is the standard Euclidean norm in $\mathbb{R}^m$. Then the functional

$$\Phi_{\chi^2}(\overline{V}_{\bar{n}}) := \sum_{i=1}^{N} \frac{|\bar{v}_{i\bar{n}} - \bar{n} p_i|^2}{|\bar{n} p_i|} \tag{7}$$

is an m-variate version of a well-known χ^2-statistic. Note that, in the present paper, we are primarily interested in the case where $N \equiv N(\bar{n}) \to \infty$ as $\bar{n} \to \infty$.

(2) Let now the sizes of all m samples be equal: $n_j = n$, $j = 1, \ldots, m$. In an equivalent reformulation of the original problem, we consider a sample of m-dimensional observations $\{(X^1_i, \ldots, X^m_i);\, i \leq n\}$ under the main hypothesis that the sample vector coordinates are independent and have the same N-atomic distribution with unknown masses $p_1, \ldots, p_N$. In this case, the log-likelihood function can be represented as the additive functional

$$\Phi_{\log}(\overline{V}_{\bar{n}}) := \sum_{i=1}^{N} (\bar{v}_{i\bar{n}}, \bar{1}) \log p_i,$$

where $\bar{1}$ is the unit vector in $\mathbb{R}^m$ and $(\cdot,\cdot)$ is the Euclidean inner product.

(3) Consider a finite or countable partition $\{\Delta_i;\, i \geq 1\}$. Let $f_{i\bar{n}}(\bar{x}) \equiv f(\bar{x}) := I_B(\bar{x})$ be the indicator function of some subset $B \subset \mathbb{Z}_+^m$. Then the functional

$$\Phi_{I_B}(\overline{V}_{\bar{n}}) := \sum_{i \geq 1} I_B(\bar{\nu}_{i\bar{n}}) \tag{8}$$

counts the number of partition elements (cells) containing any number of vector sample observations from the range B in a multinomial scheme (finite or infinite) of placing particles into cells (see [6–12]). Note that in the case of an infinite multinomial scheme in (8), it is additionally assumed that $0 \notin B$.

In the case $m = 2$ and $B = \{(x,y) \in \mathbb{Z}_+^2 : x = 0, y > 0\}$, the two-sample statistic (8) counts the number of nonempty cells after second ("additional") series of trials ("future" sample), which were empty in the first series ("original" sample). Statistics of such a kind play an important role in the theory of species sampling (for example, see [13,14]). In this case the functional (8) is called the number of unseen species in the original sample.

(4) In the case $m = 1$, consider the joint distribution (see [10]) of the random variables

$$\Phi_{I_B}(V_{n_1}), \Phi_{I_B}(V_{n_1+n_2}), \ldots, \Phi_{I_B}(V_{n_1+\ldots+n_m})$$

defined in (8) by the sample $(X_1, \ldots, X_N)$, with $N = n_1 + \ldots + n_m$. It is clear that studying the asymptotic behavior of the joint distribution of these random variables (for example, proving the multidimensional central limit theorem) can be reduced to the study of the limit distributions of the linear combinations of the form

$$a_1\Phi_{I_B}(V_{n_1}) + a_2\Phi_{I_B}(V_{n_1+n_2}) + \ldots + a_m\Phi_{I_B}(V_{n_1+\ldots+n_m})$$

for almost all vectors $(a_1, \ldots, a_m)$ with respect to the Lebesgue measure on $\mathbb{R}^m$. It is easy to see that, for any natural $j \leq m$,

$$V_{n_1+\ldots+n_j} = V_{n_1}^{(1)} + \ldots + V_{n_j}^{(j)},$$

where the empirical point processes $V_{n_1}^{(1)}, \ldots, V_{n_j}^{(j)}$ are defined by the above-mentioned independent subsamples. So, in this case, we deal with a functional of the form (5) defined by m independent empirical point processes corresponding to the m independent subsamples $(X_1, \ldots, X_{n_1})$, $(X_{n_1+1}, \ldots, X_{n_1+n_2})$,..., $(X_{N-n_m+1}, \ldots, X_N)$, and with the array of functions

$$f_{i\bar{n}}(\bar{x}) \equiv f(x_1, \ldots, x_m) := a_1 I_B(x_1) + a_2 I_B(x_1 + x_2) + \ldots + a_m I_B(x_1 + \ldots + x_m). \tag{9}$$

(5) Consider the stochastic process $\{\Phi_{I_B}(\overline{V}_{\bar{n}});\, B \subset \mathbb{Z}_+^m\}$ indexed by all subsets of $\mathbb{Z}_+^m$. As was noted above, studying the asymptotic behavior of the joint distributions of this process can be reduced to studying the asymptotic behavior of the distributions of any linear combinations of corresponding one-dimensional projections of this process, i.e., to studying the asymptotic behavior of the distributions of functionals of the form (5) for $m = 1$ and the array of functions

$$f_{i\bar{n}}(x) \equiv f(x) := a_1 I_{B_1}(x) + a_2 I_{B_2}(x) + \ldots + a_r I_{B_r}(x) \tag{10}$$

for almost all vectors $(a_1, \ldots, a_r)$. For one-point sets, the asymptotic analysis of the above-mentioned joint distributions can be found, for example, in [7–12].

(6) Consider the case $m = 1$ and the functional

$$\Phi_f(V_n) := \sum_{i \geq 1} np_i I_B(\nu_{in}), \tag{11}$$

which counts the sampling ratio of the cells containing any number of particles from the range B. For the one-point set $B = \{0\}$, such functional was considered in [9]. In general, if instead of np_i in (11) we consider arbitrary weights $g(n,i) > 0$ (under condition (6)) with one or another interpretation, the functional $\Phi_f(V_n)$ in this case will be interpreted as the total weight of the corresponding cells.

4. Poissonization: Duality Theorem

In this section, we present the main result of the paper—a duality theorem for additive statistics under consideration. First of all, we explain the term "Poissonization". It means that studying the limit behavior of the original additive statistics, we reduce the problem to studying the following "Poissonian version" of the functional (5) under condition (6):

$$\Phi_f(\Pi_{\bar{n}}) := \sum_{i\geq 1} f_{i\bar{n}}(\bar{\pi}_{i\bar{n}}), \tag{12}$$

where $\bar{\pi}_{i\bar{n}} = \left(\pi^{(1)}_{n_1 i}, \ldots, \pi^{(m)}_{n_m i}\right)$, $\pi^{(k)}_{n_k i} := \Pi_{n_k}(\Delta_i)$, $i \geq 1$, is a sequence of independent Poisson random variables with respective parameters $n_k p_i$. It is clear that the functional (12) is well defined with probability 1 since only a finite number of the vectors $\{\bar{\pi}_{i\bar{n}}\}$ differ from the zero vector. Independence of the summands is a crucial difference of the Poisson version of an additive functional from the original one. Some elements of Poissonization for additive functionals of the form (8) and (10) are contained, for example, in [9,12]. In [9], the author used the well-known representation of an empirical point process as the conditional Poisson point process under the condition that the number of atoms of the accompanying Poisson point process equals n. Moreover, in [9], the simple known representation $\pi(n) = n + O_p(\sqrt{n})$ was employed, where $O_p(\sqrt{n})$ denotes a random variable such that $O_p(\sqrt{n})/\sqrt{n}$ is bounded in probability as $n \to \infty$. In [12], proving the multivariate central limit theorem for the above-mentioned joint distributions (in fact, for functionals of the form (10) in the case of one-point subsets $\{B_i\}$), the authors applied a reduction to the joint distributions of the Poissonian versions of additive functionals using known upper bounds for a multivariate Poisson approximation to a multinomial distribution (see also [15]). The main goal of the paper is to establish a duality theorem, which demonstrates absolute identity of the asymptotic behavior of the distributions of the additive functionals under consideration and their Poissonian versions.

First, we formulate a crucial auxiliary assertion in proving the main result.

Lemma 1. *Let $\{\Delta_{\bar{n}}\}$ be an arbitrary scalar array satisfying the condition $f_{i\bar{n}}(\pi_{i\bar{n}})\Delta_{\bar{n}} \xrightarrow{p} 0$ as $\bar{n} \to \infty$ for every fixed i. Then, for each multiindex $\bar{n}$, one can define on a common probability space a pair of point processes $\overline{V}^*_{\bar{n},\Delta_{\bar{n}}}$ and $\overline{\Pi}^*_{\bar{n},\Delta_{\bar{n}}}$ such that $\mathcal{L}(\overline{V}^*_{\bar{n},\Delta_{\bar{n}}}) = \mathcal{L}(\overline{V}_{\bar{n}})$, $\mathcal{L}(\overline{\Pi}^*_{\bar{n},\Delta_{\bar{n}}}) = \mathcal{L}(\overline{\Pi}_{\bar{n}})$, and for any $\varepsilon > 0$,*

$$\mathbf{P}\Big(|\Delta_{\bar{n}}|\Big|\Phi_f(\overline{V}^*_{\bar{n},\Delta_{\bar{n}}}) - \Phi_f(\overline{\Pi}^*_{\bar{n},\Delta_{\bar{n}}})\Big| > \varepsilon\Big) \to 0 \ \ as \ \ \bar{n} \to \infty. \tag{13}$$

Remark 2. *Lemma 1 only asserts that the marginal distributions (that is, for each $\bar{n}$ separately) of the arrays $\{\overline{V}^*_{\bar{n},\Delta_{\bar{n}}},\ \bar{n} \in \mathbb{Z}^m_+\}$ and $\{\overline{V}_{\bar{n}},\ \bar{n} \in \mathbb{Z}^m_+\}$, and also $\{\overline{\Pi}^*_{\bar{n},\Delta_{\bar{n}}},\ \bar{n} \in \mathbb{Z}^m_+\}$ and $\{\overline{\Pi}_{\bar{n}},\ \bar{n} \in \mathbb{Z}^m_+\}$. Note that the probability in (13) is precisely determined by the marginal distributions of the mentioned random arrays, i.e., formally, it also depends on $\bar{n}$. Without loss of generality, we can assume that pairs of point processes $(\overline{V}^*_{\bar{n},\Delta_{\bar{n}}}, \overline{\Pi}^*_{\bar{n},\Delta_{barn}})$ are independent in $\bar{n}$, and on this extended probability space, the universal probability measure $\mathbf{P}$ in (13) is given in the standard way, which no longer depends on $\bar{n}$. In this case it is correct to speak about the convergence to zero in probability of the sequence of random variables in (13).*

Lemma 1 gives the key to the proof of the following duality theorem, a criterion for the weak convergence of distributions of functionals of the point processes under consideration. The essence of this result is that the asymptotic behavior of the distributions of additive functionals of the point processes $\overline{V}_{\bar{n}}$ and $\overline{\Pi}_{\bar{n}}$ is exactly the same. In addition, one can also indicate a third class of additive functionals (under condition (6)) that has the same property:

$$\Phi_f^* := \sum_{i\geq 1} f_{i\bar{n}}(\bar{v}_{i\bar{n}}^*),$$

where $\{\bar{v}_{i\bar{n}}^*, i \geq 1\}$ is a sequence of independent random vectors such that $\mathcal{L}(\bar{v}_{i\bar{n}}^*) = \mathcal{L}(\bar{v}_{i\bar{n}})$ for all i. The functional Φ_f^* is well defined due to the Borel–Cantelli lemma and the simple estimate $\mathbf{P}(\bar{v}_{i\bar{n}}^* \neq 0) = \mathbf{P}(\bar{v}_{i\bar{n}} \neq 0) \leq m\|\bar{n}\|p_i$.

Let us agree that the symbol «$\Longrightarrow$» in what follows will denote the weak convergence of distributions. The main result of the paper is as follows.

Theorem 3. *Under the conditions of Lemma* 1*, the following three limit relations are equivalent as* $\bar{n} \to \infty$:

$$(1)\ \mathcal{L}\Big(\Phi_f(\overline{V}_{\bar{n}})\Delta_{\bar{n}} - M_{\bar{n}}\Big) \Longrightarrow \mathcal{L}(\gamma),$$

$$(2)\ \mathcal{L}\Big(\Phi_f(\overline{\Pi}_{\bar{n}})\Delta_{\bar{n}} - M_{\bar{n}}\Big) \Longrightarrow \mathcal{L}(\gamma),$$

$$(3)\ \mathcal{L}\Big(\Phi_f^*\Delta_{\bar{n}} - M_{\bar{n}}\Big) \Longrightarrow \mathcal{L}(\gamma),$$

where $M_{\bar{n}}$ and $\Delta_{\bar{n}}$ are some scalar arrays and γ is some random variable.

5. Applications

Theorem 3 allows us to reduce the asymptotic analysis of the distributions of the additive functionals under consideration to a similar analysis of their Poissonian versions, i.e., to the asymptotic analysis of distributions of sums (finite or infinite) of independent random variables, or to reduce the problem to studying the limit behavior of the distributions $\mathcal{L}\Big(\Phi_f(\overline{V}_{\bar{n}}\Big)$, absolutely ignoring the dependence of the random variables $\{\bar{v}_{i\bar{n}}, i \geq 1\}$. Note also that, under some rather broad assumptions, the law $\mathcal{L}(\gamma)$ will be infinitely divisible. A detailed analysis of such conditions and corresponding examples will be considered in a separate paper. Here we present only a few of these corollaries, focusing our attention on the equivalence of the first two relations of Theorem 3.

First of all, we note one useful property of the expectations of the functionals under consideration as functions of $\bar{n}$.

Lemma 2. *Let* $\max_{\bar{n}} \sup_{\bar{x}} |f_{i\bar{n}}(\bar{x})| \leq C_i$*,* $\sum\limits_{i\geq 1} C_i p_i < \infty$*, and*

$$\sum_{i\geq 1} \mathbf{E}|f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})| < \infty \ \ \forall \bar{n}. \tag{14}$$

Then the relations $\lim\limits_{\bar{n}\to\infty} |\mathbf{E}\Phi_f(\overline{V}_{\bar{n}})| = \infty$ *and* $\lim\limits_{\bar{n}\to\infty} |\mathbf{E}\Phi_f(\overline{\Pi}_{\bar{n}})| = \infty$ *are equivalent. In the case of infinite limits,*

$$\mathbf{E}\Phi_f(\overline{V}_{\bar{n}}) \sim \mathbf{E}\Phi_f(\overline{\Pi}_{\bar{n}}) \ \ as \ \ \bar{n} \to \infty.$$

Remark 3. *For functionals of the form (8) in an infinite multinomial scheme, the conditions of Lemma 2 are typical. Let $m = 1$ and $B := \{j : j > k\}$ for any $k \geq 0$. Then*

$$\lim_{n\to\infty} \mathbf{E}\Phi_f(V_n) = \lim_{n\to\infty} \sum_{i\geq 1} \mathbf{P}(\nu_{in} > k) = \infty$$

since, by virtue of the law of large numbers, $\lim\limits_{n\to\infty} \mathbf{P}(\nu_{in} > k) \to 1$ *for every fixed i. Moreover, in the case under consideration, obviously,* $\mathbf{E}\Phi_f(V_n) \leq n$*. Similarly, without any restrictions on the probabilities $\{p_i\}$, the infinite limits in Lemma 2 for functionals of the form (8) (and even more so for (11)) also hold for the set B consisting of all odd natural numbers. Here the limit relation* $\lim\limits_{n\to\infty} \mathbf{E}\Phi_f(\overline{\Pi}_{\bar{n}}) \equiv \lim\limits_{n\to\infty} \sum\limits_{i\ ge1} \mathbf{P}(\pi_{in} \in B) = \infty$ *follows immediately from the equality* $\mathbf{P}(\pi_{in} \in B) = \frac{1}{2}(1 - e^{-2np_i})$.

It is also worth noting that for some sets B the main contribution to the limit behavior of the series $\sum_{i\geq 1} \mathbf{P}(\pi_{in} \in B)$ can be made not only by their initial segments but also tails. For example, this will be the case for any one-point sets $B_k := \{k\}$ for $k > 0$ if the group probabilities are given as $p_i = Ci^{-1-b}$ or $p_i = ce^{-C_o i^\alpha}$ for some constants $c, C, C_o, b > 0$ and $\alpha \in (0,1)$. In this case, for any subset B of natural numbers in the definition of the functionals (8) and (11), the expectation limits indicated in Lemma 2 will be infinite (see Section 7 and [9,12]). On the other hand, if $p_i = ce^{-C_o i}$, then for any one-point set the expectations mentioned will be bounded uniformly in n (see Section 7 and [9,12]). For more complex functionals with kernels (9) or (10) for the above-mentioned distributions $\{p_i\}$, one can find sufficiently broad conditions that ensure unbounded increase in their expectations and variances as $\bar{n} \to \infty$ for almost all vectors $(a_1, \ldots, a_r) \in \mathbb{R}^r$ (see Section 7).

Now we present one of the corollaries of Theorem 3, namely, the law of large numbers for the additive functionals under consideration, setting in this theorem $\Delta_{\bar{n}} := (\mathbf{E}\Phi_f(\overline{\Pi}_{\bar{n}}))^{-1}$, $M_{\bar{n}} := 0$, and $\gamma := 1$.

Corollary 2. *Let the conditions of Lemma 2 be fulfilled. If* $|\mathbf{E}\Phi_f(\Pi_{\bar{n}})| \to \infty$ *as* $\bar{n} \to \infty$ *then the following criterion holds*:

$$\frac{\Phi_f(\overline{V}_{\bar{n}})}{\mathbf{E}\Phi_f(\overline{V}_{\bar{n}})} \xrightarrow{p} 1 \quad \textit{iff} \quad \frac{\Phi_f(\overline{\Pi}_{\bar{n}})}{\mathbf{E}\Phi_f(\overline{\Pi}_{\bar{n}})} \xrightarrow{p} 1;$$

in this case, the normalizations $\mathbf{E}\Phi_f(\overline{V}_{\bar{n}})$ *and* $\mathbf{E}\Phi_f(\overline{\Pi}_{\bar{n}})$ *can be swapped.*

Remark 4. *In consideration of Chebyshev's inequality, a sufficient condition for the limit relations in Corollary 2 is as follows:*

$$\frac{\sum_{i\geq 1} \mathbf{D} f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})}{\left(\sum_{i\geq 1} \mathbf{E} f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})\right)^2} \to 0.$$

For example, let $f_{i\bar{n}}(\cdot) \geq 0$ *and* $\sup_{\bar{x},i,\bar{n}} f_{i\bar{n}}(\bar{x}) \leq C_0$. *Then* $\mathbf{D} f_{i\bar{n}}(\bar{\pi}_{i\bar{n}}) \leq C_0 \mathbf{E} f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})$ *and*

$$\frac{\sum_{i\geq 1} \mathbf{D} f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})}{\left(\sum_{i\geq 1} \mathbf{E} f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})\right)^2} \leq C_0 \left|\sum_{i\geq 1} \mathbf{E} f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})\right|^{-1} \to 0.$$

In particular, this estimate is valid in the case $f_{i\bar{n}}(\bar{x}) \equiv f(\bar{x}) := I_B(\bar{x})$, *with* $0 \notin B$, *if only* $\mathbf{E}\Phi_f(\overline{\Pi}_{\bar{n}}) = \sum_{i\geq 1} \mathbf{P}(\bar{\pi}_{i\bar{n}} \in B) \to \infty$.

We now formulate an analog of Lemma 2 for the variances of the functionals under consideration.

Lemma 3. *Under the conditions* $\max_{\bar{n}} \sup_{\bar{x}} |f_{i\bar{n}}(\bar{x})| \leq C_i\ \forall i$ *and* $\sum_{i\geq 1} C_i^2 p_i < \infty$ *the limit relation* $\lim_{\bar{n}\to\infty} \mathbf{D}\Phi_f(\overline{V}_{\bar{n}}) = \infty$ *holds if and only if* $\lim_{\bar{n}\to\infty} \mathbf{D}\Phi_f(\overline{\Pi}_{\bar{n}}) = \infty$. *In the case of infinite limit the following equivalence is valid:* $\mathbf{D}\Phi_f(\overline{V}_{\bar{n}}) \sim \mathbf{D}\Phi_f(\overline{\Pi}_{\bar{n}})$ *as* $\bar{n} \to \infty$.

Lemma 3 and Theorem 3 imply the following important criterion, which allows us to reduce proving the central limit theorem for additive functionals $\Phi_f(\overline{V}_{\bar{n}})$ to proving the same assertion for the Poissonian version $\Phi_f(\overline{\Pi}_{\bar{n}})$.

Corollary 3. *Under the conditions of Lemma 3 and* $\mathbf{D}\Phi_f(\overline{\Pi}_{\bar{n}}) \to \infty$ *as* $\bar{n} \to \infty$ *the limit relation*

$$\mathcal{L}\left(\frac{\Phi_f(\overline{V}_{\bar{n}}) - \mathbf{E}\Phi_f(\overline{V}_{\bar{n}})}{\mathbf{D}^{1/2}\Phi_f(\overline{V}_{\bar{n}})}\right) \Longrightarrow N(0,1) \quad as \;\; \bar{n} \to \infty,$$

is valid if, and only if,

$$\mathcal{L}\left(\frac{\Phi_f(\overline{\Pi}_{\bar{n}}) - \mathbf{E}\Phi_f(\overline{\Pi}_{\bar{n}})}{\mathbf{D}^{1/2}\Phi_f(\overline{\Pi}_{\bar{n}})}\right) \Longrightarrow \mathcal{N}(0,1) \quad as \;\; \bar{n} \to \infty,$$

where $\mathcal{N}(0,1)$ is the standard normal distribution. In this case, the normalizing and centering sequences in these two limit relations can be, respectively, swapped.

In order to prove this corollary we should put in Theorem 3 $\Delta_{\bar{n}} := \mathbf{D}^{-1/2}\Phi_f(\overline{\Pi}_{\bar{n}})$, $M_{\bar{n}} := \mathbf{E}\Phi_f(\overline{V}_{\bar{n}})\mathbf{D}^{-1/2}\Phi_f(\overline{\Pi}_{\bar{n}})$, and $\mathcal{L}(\gamma) := \mathcal{N}(0,1)$. In this case, Lemma 3 allows us only to replace the normalizing and centering sequences in Theorem 3 with some equivalent sequences.

Remark 5. *The validity of the central limit theorem for the sequence $\Phi_f(\overline{\Pi}_{\bar{n}})$ in Theorem 3 will be justified if, say, the third-order Lyapunov condition is met:*

$$\frac{\sum\limits_{i\geq 1} \mathbf{E}|f_{i\bar{n}}(\bar{\pi}_{i\bar{n}}) - \mathbf{E}f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})|^3}{\left(\sum\limits_{i\geq 1} \mathbf{D}f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})\right)^{3/2}} \to 0 \;\; as \;\; \bar{n} \to \infty.$$

For example, let $\sup\limits_{\bar{x},i,\bar{n}} |f_{i\bar{n}}(\bar{x})| \leq C_0$. Then it is easy to see that

$$\sum_{i\geq 1} \mathbf{E}|f_{i\bar{n}}(\bar{\pi}_{i\bar{n}}) - \mathbf{E}f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})|^3 \leq 2C_0 \sum_{i\geq 1} \mathbf{D}f_{i\bar{n}}(\bar{\pi}_{i\bar{n}}).$$

Thus, if $\mathbf{D}\Phi_f(\overline{\Pi}_{\bar{n}}) \to \infty$as $\bar{n} \to \infty$, then the Lyapunov condition will be met and the approval of the above investigation will take place. So an important special case $f_{i\bar{n}}(\bar{x}) := I_B(\bar{x})$ is included in the scheme at issue if

$$\mathbf{D}\Phi_{I_B}(\overline{\Pi}_{\bar{n}}) = \sum_{i\geq 1} \mathbf{P}(\bar{\pi}_{i\bar{n}} \in B)(1 - \mathbf{P}(\bar{\pi}_{i\bar{n}} \in B)) \to \infty \;\; as \;\; \bar{n} \to \infty.$$

Note that examples for which the specified variance property takes place or is violated are given, for example, in [9].

Finally, here is another consequence of Theorem 3, relating to the asymptotic behavior of χ^2-statistics in (7) at $m = 1$ and $N \equiv N(n) \to \infty$. First of all, note that

$$\mathbf{E}\Phi_{\chi^2}(\Pi_n) = N,$$

$$D_n := \mathbf{D}\Phi_{\chi^2}(\Pi_n) = 2N + \sum_{i=1}^{N} \frac{1}{np_i}.$$

Corollary 4. *Let $N \equiv N(n) \to \infty$ as $n \to \infty$. Then the following two asymptotic relations are equivalent:*

$$\mathcal{L}\left(\frac{\Phi_{\chi^2}(V_n) - N}{{D_n}^{1/2}}\right) \Longrightarrow \mathcal{N}(0,1), \tag{15}$$

$$\mathcal{L}\left(\frac{\Phi_{\chi^2}(\Pi_n) - N}{{D_n}^{1/2}}\right) \Longrightarrow \mathcal{N}(0,1). \tag{16}$$

Note that in the present case, the requirement of Lemma 1 is met, since each term $\frac{(\nu_{in}-np_i)^2}{np_i}$ (as a sequence of n) is bounded in probability due to Markov's inequality, and therefore, with the normalizing sequence $\Delta_n := D_n^{-1}$, this term will tend to zero in probability as $n \to \infty$.

Remark 6. *In the relations (15) and (16) we can say just about the double limit when $N, n \to \infty$ because this assertion is missing restrictions on the rate of increase in the sequence $N(n)$. The proposed formulation in Corollary 4, equivalent to the one just mentioned, is more convenient to refer to Theorem 3. Note that the centering sequence E_n can be replaced with its equivalent sequence $\mathbf{E}\Phi_{\chi^2}(V_n) = N - 1$. Replacement in the normalization in (15) the variance D_n with the variance of the χ^2-statistic itself, i.e., by the term (for example, see [16])*

$$\mathbf{D}\Phi_{\chi^2}(V_n) = 2N + \frac{1}{N}\sum_{i=1}^{N}\frac{1}{np_i} - \frac{3N-2}{n},$$

is possible only if these two variances are equivalent. For example, this would be the case if $\min_{i \le N} np_i \to \infty$. This means that the growth rate of the sequence $N \equiv N(n)$ is subject to appropriate constraints, which is not the case in the above consequence. So, in this assertion we can talk about a double limit as $n, N \to \infty$.

The formulated criterion allows us to establish a fairly general sufficient condition for the asymptotic normality of χ^2-statistics with an increasing number of groups.

Theorem 4. *Let $N \equiv N(n) \to \infty$ as $n \to \infty$. Then the asymptotic relation (15) is valid if*

$$\frac{\sum_{i=1}^{N}(np_i)^{-2}}{\left(N + \sum_{i=1}^{N}(np_i)^{-1}\right)^{3/2}} \longrightarrow 0 \tag{17}$$

as $n \to \infty$.

The problem of finding more or less broad sufficient conditions for asymptotic normality χ^2-statistics with a growing number of groups were studied by many authors in the second half of the last century (for example, see [3–5,16–18]). Note that all known sufficient conditions for the above weak convergence imply fulfillment of the asymptotic relation (17). For example, the condition $\min_{i \le N} np_i \to \infty$ along with $N \to \infty$ (see [17,18]), obviously immediately entails relation (17). It is equally obvious that the requirement of the so-called regularity of multinomial models (see [3–5]), i.e.,

$$0 < c_1 \le \min_{i \le N} Np_i, \;\; \max_{i \le N} Np_i < c_2 < \infty,$$

where the constants c_1 and c_2 are independent of N, also implies (17). On the other hand, it is easy to construct examples in which the regularity requirement of the multinomial model is violated but relation (17) is valid. For example, let $p_i := C_N i^{-1-b}$, $i = 1, \ldots, N$, where $b > 0$ and $C_N := \left(\sum_{i \le N} i^{-1-b}\right)^{-1}$. It is easy to see that, as $N \to \infty$, the sums $\sum_{i=1}^{N} p_i^{-2}$ and $\sum_{i=1}^{N} p_i^{-1}$ increase as N^{3+2b} and N^{2+b}, respectively. Therefore, as $n, N \to \infty$, the ratio in (17) is equivalent to

$$\frac{N^{3+2b}}{\sqrt{n}(N^{2+b})^{3/2}} = \frac{N^{b/2}}{\sqrt{n}}$$

up to a constant factor. So, here we already need to measure the growth rate N with n. Obviously, in this case, in order to fulfill condition (17), you need to require that $N = o(n^{1/b})$. If the probabilities p_i decrease exponentially then the growth rate zone for N narrows to $o(\log n)$. It is worth to note that for the above-mentioned power-type probabilities at issue the condition $\min_{i \le N} np_i \to \infty$ implies the asymptotic relation $N = o(n^{1/(b+1)})$ that is more restrictive than the above constraint.

6. Probability and Moment Inequalities

The next theorem is related to estimation of the distribution tails of additive functionals.

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

$$\mathbf{P}(\Phi_f(\overline{V}_{\bar{n}}) \geq x) \leq 2C^* \mathbf{P}(\Phi_f(\overline{\Pi}_{\bar{n}}) \geq x/2), \tag{18}$$

where $C^* := \min_{j \geq 1} \max\{(\sum_{i \leq j} p_i)^{-1}, (\sum_{i > j} p_i)^{-1}\}$. If additionally $\sup_x f_{1\bar{n}}(x) \leq c_0$ then

$$\mathbf{P}(\Phi_f(\overline{V}_{\bar{n}}) \geq x) \leq p_1^{-1} \mathbf{P}(\Phi_f(\overline{\Pi}_{\bar{n}}) \geq x - c_0). \tag{19}$$

Remark 7. *In (19), the constant c_0 may depend on $\bar{n}$. What is more, we can use the truncation of the random variable $f_{1\bar{n}}(\nu_{i_{\bar{n}}})$ at the level c_0, while adding to the right-hand side of inequality (19) the probability $\mathbf{P}(f_{1\bar{n}}(\nu_{i_{\bar{n}}}) > c_0)$.*

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

$$\mathbf{E}F(\Phi_f(\overline{V}_{\bar{n}})) \leq 2C^* \mathbf{E}F(2\Phi_f(\overline{\Pi}_{\bar{n}})). \tag{20}$$

As an example, consider the functional $\Phi_{I_B}(\overline{V}_{\bar{n}})$ defined in (8). Then, as a consequence of (19) and Chernoff's upper bound [19] for the distribution tail of a sum of independent nonidentically distributed Bernoulli random variables (the transition from finite sums to series in this case is obvious), we obtain the following result.

Corollary 6. *Put $M_n(B) := \mathbf{E}\Phi_{I_B}(\overline{\Pi}_{\bar{n}}) = \sum_{i \geq 1} \mathbf{P}(\pi_{in} \in B)$. Then for any $\varepsilon > (M_n(B))^{-1}$ the following inequality holds*:

$$\mathbf{P}\left(\left| \frac{\Phi_{I_B}(\overline{V}_{\bar{n}})}{M_n(B)} - 1 \right| > \varepsilon \right) \leq 2p_1^{-1} e^{-\frac{\delta^2 M_n(B)}{2+\delta}}, \tag{21}$$

where $\delta := \varepsilon - \frac{1}{M_n(B)} > 0$.

Remark 8. *one can replace the Poissonian mean $M_n(B)$ in (21) 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 [15,20]). Further, if the condition $M_n(B) \to \infty$ is met as $n \to \infty$ then from (21) 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) (see Section 7). 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 occurs at any speed of increasing the sequence $M_n(B)$ to infinity. This follows from estimate (21), the monotonicity of the functions $I_{B_k}(x)$, and the simple technique in proving SLLN in [9,21].*

7. Asymptotic Analysis of the Means and Variances of Additive Statistics

In the previous section, it was noted that when proving certain limit theorems for the introduced additive functionals, it is extremely important to have information about the behavior of their means and variances. In this section, for additive statistics (8)–(11), we demonstrate exactly how the asymptotic behavior of these moments is studied. To simplify the notation, we will consider here the case $m = 1$. The subsequent asymptotic analysis is based on the following elementary assertion, which is presented in one way or another in many papers on this topic.

Lemma 4. *Let $f_n(x)$ be a sequence of non-negative, integrable, and piecewise monotonic functions defined on $\mathbb{R}_+$. Suppose that each $f_n(x)$ has M monotonicity intervals, where M is independent of n. Finally, assume that, as $n \to \infty$,*

$$\int_0^\infty f_n(x)dx \to \infty, \quad \sup_{x\geq 0} f_n(x) = o\left(\int_0^\infty f_n(x)dx\right).$$

Then, as $n \to \infty$,

$$\sum_{j>0} f_n(j) \sim \int_0^\infty f_n(x)dx.$$

We now give a few examples of calculating the asymptotics we need.

(1) Let $B_k := \{i : i > k\}$ for any $k \in \mathbb{Z}_+$. In Remark 3 it was already noted that $M_n(B_k) \to \infty$ due to the strong law of large numbers for binomially distributed random variables. However, for specific classes of distributions $\{p_i\}$, one can estimate the growth rate of the sequence $\{M_n(B_k)\}$. For example, let $p_i := Ci^{-1-b}$, where $b > 0$, $i = 1, 2, \ldots$. Then, using Lemma 4 and the well-known connection between the tail of a Poisson distribution and the corresponding gamma distribution, we obtain after integration by parts and a change of the integration variable:

$$M_n(B_k) \equiv \sum_{i\geq 1} \mathbf{P}(\pi_{in} > k) = \sum_{i\geq 1} \gamma_{k+1,1}(np_i)$$
$$\sim (Cn)^{\frac{1}{1+b}} \int_0^\infty \gamma_{k+1,1}(y^{-1-b})dy = \frac{(Cn)^{\frac{1}{1+b}}}{k!}\Gamma\left(k + \frac{b}{1+b}\right), \quad (22)$$

where $\gamma_{k+1,1}(z) := \int_0^z \frac{t^k}{k!}e^{-t}dt$, $\Gamma(z) := \int_0^\infty t^{z-1}e^{-t}dt$, $z > 0$, are the distribution function of the gamma-distribution with parameters $(k+1, 1)$, and the gamma-function, , respectively. For example, if $k = 0$ then the asymptotics of the expectation of the number of nonempty cells is as follows (see [6,9]):

$$M_n(B_0) \sim (Cn)^{\frac{1}{1+b}} \int_0^\infty (1 - e^{-y^{-1-b}})dy = (Cn)^{\frac{1}{1+b}}\Gamma\left(\frac{b}{1+b}\right). \quad (23)$$

By analogy to the arguments in proving (22), after an appropriate change of the integration variable, we obtain for the one-point sets the following asymptotics:

$$M_n(\{k\}) \sim (Cn)^{\frac{1}{1+b}} \int_0^\infty \frac{y^{-k(1+b)}}{k!} e^{-y^{-1-b}} dy$$
$$= \frac{(Cn)^{\frac{1}{1+b}}}{(1+b)k!} \int_0^\infty x^{k-1-\frac{1}{1+b}} e^{-x} dx = \frac{(Cn)^{\frac{1}{1+b}}}{(1+b)k!}\Gamma\left(k - \frac{1}{1+b}\right). \quad (24)$$

Thus, from (24) it follows that for *any subset B of the natural series* in the case under consideration of a power-law decrease in $\{p_i\}$ the following asymptotic representation is true:

$$M_n(B) \sim \frac{(Cn)^{\frac{1}{1+b}}}{(1+b)} \sum_{k\in B} \frac{1}{k!}\Gamma\left(k - \frac{1}{1+b}\right). \quad (25)$$

Note that, due to the countable additivity of the finite measure $M_n(\cdot)$ and the relations (22)–(24), the sum (possibly infinite) in (25) will always be finite.

Remark 9. *Inequality (21), relation (25), and the Borel–Cantelli lemma guarantee that the strong law of large numbers holds for the sequence $\{M_n(B)\}$ for any subsets B of the natural series in the case of a power-law decrease in the probabilities $\{p_i\}$. Moreover, what has been said and the above asymptotics are also preserved for probabilities of the form $p_i := C(i)i^{-1-b}$, where $C(x)$ is a slowly varying function under certain minimal constraints (see [9,12]). In this case, in the asymptotic relations (22)–(25) instead of C one should substitute $C(n)$.*

Asymptotic behavior of the variances of the functionals $\Phi_{I_B}(\overline{\Pi}_n)$ for some B and broad conditions on the rate of decrease in the sequence $\{p_i\}$ is given in [9]. Here we only demonstrate how this variance is calculated for *arbitrary* subsets B of the natural series under the above conditions on $\{p_i\}$. Analogously with (22) we have for the infinite intervals B_k:

$$D_n(B_k) := \mathbf{D}\Phi_{I_{B_k}}(\overline{\Pi}_n) = \sum_{i\geq 1}\mathbf{P}(\pi_{in} > k) - \sum_{i\geq 1}\mathbf{P}^2(\pi_{in} > k)$$

$$= \sum_{i\geq 1}\gamma_{k+1,1}(np_i) - \sum_{i\geq 1}\gamma^2_{k+1,1}(np_i) \sim (Cn)^{\frac{1}{1+b}}\int\limits_0^\infty\Big(\gamma_{k+1,1}(y^{-1-b}) - \gamma^2_{k+1,1}(y^{-1-b})\Big)dy. \quad (26)$$

Similarly to proving (24), we derive the asymptotics of the variance for the one-point sets:

$$\begin{aligned} D_n(\{k\}) &= \sum_{i\geq 1}\mathbf{P}(\pi_{in} = k) - \sum_{i\geq 1}\mathbf{P}^2(\pi_{in} = k) \\ &= \frac{(Cn)^{\frac{1}{1+b}}}{(1+b)}\left(\int\limits_0^\infty \frac{1}{k!}x^{k-1-\frac{1}{1+b}}e^{-x}dx - \int\limits_0^\infty \frac{1}{(k!)^2}x^{2k-1-\frac{1}{1+b}}e^{-2x}dx\right) \\ &= \frac{(Cn)^{\frac{1}{1+b}}}{(1+b)k!}\left(\Gamma\left(k-\frac{1}{1+b}\right) - \frac{2^{\frac{1}{1+b}-2k}}{k!}\Gamma\left(2k-\frac{1}{1+b}\right)\right). \end{aligned} \quad (27)$$

Although the set function $D_n(\cdot)$ is not additive, the extension to arbitrary subsets B of the natural series of computing the asymptotics of $D_n(B)$ presents no difficulty. Along with formula (25), which gives one term in the resulting asymptotics, we use the following representation for the second sum:

$$\begin{aligned} \sum_{i\geq 1}\mathbf{P}^2(\pi_{in}\in B) &\sim \frac{(Cn)^{\frac{1}{1+b}}}{1+b}\int\limits_0^\infty\left(\sum_{k\in B}\frac{x^k}{k!}\right)^2 x^{-1-\frac{1}{1+b}}e^{-2x}dx \\ &= \frac{(Cn)^{\frac{1}{1+b}}}{1+b}\sum_{k,l\in B}\frac{2^{\frac{1}{1+b}-k-l}}{k!l!}\Gamma\left(k+l-\frac{1}{1+b}\right). \end{aligned} \quad (28)$$

Thus, the difference between the right-hand sides of (25) and (28) determines the asymptotic of $D_n(B)$ for any subset of the natural series.

(2) The asymptotics of the first two moments for the functionals (10) for pairwise disjoint sets $\{B_j\}$ is derived in exactly the same way. In the case of one-point sets $B_j := \{k_j\}$, the asymptotic behavior of the first moment immediately follows from the previous calculations. As for the variance, we should first note that, due to the orthogonality of the indicator random variables under consideration, we have

$$\mathbf{D}\sum_{s=1}^r a_s I_{B_s}(\pi_{in}) = \sum_{s=1}^r a_s^2\mathbf{P}(\pi_{in}=k_s) - \left(\sum_{s=1}^r a_s\mathbf{P}(\pi_{in}=k_s)\right)^2$$

$$= \sum_{s=1}^r a_s^2\mathbf{P}(\pi_{in}=k_s) - \sum_{j,s=1}^r a_s a_j\mathbf{P}(\pi_{in}=k_s)\mathbf{P}(\pi_{in}=k_j).$$

Summation over i of the resulting expression and the previous calculations give the desired asymptotics:

$$\mathbf{D}\Phi_f(\Pi_n) \sim \frac{(Cn)^{\frac{1}{1+b}}}{b+1}\sum_{s,j=1}^r\left[\frac{a_s^2}{rk_s!}\Gamma\left(k_s-\frac{1}{b+1}\right) - \frac{2^{\frac{1}{b+1}-k_s-k_j}a_s a_j}{k_s!k_j!}\Gamma\left(k_s+k_j-\frac{1}{b+1}\right)\right].$$

We note the resulting representation can vanish on the set of vectors $(a_1, \ldots, a_r)$ of zero Lebesgue measure in $\mathbb{R}^r$, i.e., on the surface defined by the relation $\sum\limits_{s,j=1}^{r} B_{s,j} a_s a_j = 0$ for some coefficients $\{B_{s,j}\}$.

For infinite intervals of the form $B_j := \{i : i > k_j\}$, the variance is studied in a similar way. We assume without loss of generality that $k_1 \le k_2 \le \ldots \le k_r$. To calculate the variance of this functional, it suffices for us to restrict ourselves to the second moment, since the asymptotics of the first one has already been studied. We have

$$\mathbf{E}\left(\sum_{s=1}^{r} a_s I(\pi_{in} > k_s)\right)^2 = \sum_{s=1}^{r} a_s^2 \mathbf{P}(\pi_{in} > k_s) + 2\mathbf{E}\sum_{j=1}^{r-1} a_j I(\pi_{in} > k_j) \sum_{s>j}^{r} a_s I(\pi_{in} > k_s)$$

$$= \sum_{s=1}^{r} a_s^2 \mathbf{P}(\pi_{in} > k_s) + 2\mathbf{E}\sum_{j=1}^{r-1} a_j \sum_{s>j}^{r} a_s I(\pi_{in} > k_s)$$

$$= \sum_{s=1}^{r} a_s^2 \mathbf{P}(\pi_{in} > k_s) + 2\sum_{j=1}^{r-1} a_j \sum_{s>j}^{r} a_s \mathbf{P}(\pi_{in} > k_s).$$

Further calculations in essence have already been made earlier. So, finally we obtain

$$\mathbf{D}\Phi_f(\Pi_n) \sim (Cn)^{\frac{1}{1+b}} \sum_{s,j=1}^{r}\left[\frac{a_s^2}{r}\int_0^\infty \Gamma_{k_s+1,1}(v^{-1-b})dv - a_s a_j \int_0^\infty \Gamma_{k_s+1,1}(v^{-1-b})\Gamma_{k_j+1,1}(v^{-1-b})dv\right]$$

with comments similar to the above regarding the zeroing of the double sum.

To conclude this section, we give an example where the above-mentioned moments of the functional under consideration do not tend to infinity as n grows. We put $p_j = e^{-Cj}$, with $C := \log 2$. Let us show that

$$\sup_n \sum_{j\ge 1} \mathbf{P}(\pi_{nj} = k) < \infty.$$

This estimate obviously implies that the first two moments of the functional $\Phi_{I_B}(\Pi_n)$ are uniformly bounded in n for $B := \{k\}$. Indeed, one has

$$\sum_{j\ge 1} \mathbf{P}(\pi_{nj} = k) = \frac{n^k}{k!}\sum_{j\ge 1} e^{-ne^{-Cj}} e^{-Ckj} \le \frac{e^{Ck} n^k}{k!}\int_1^\infty e^{-ne^{-Cx}} e^{-Ckx} dx$$

$$= \frac{e^{Ck} n^k}{Ck!}\int_0^{e^{-C}} e^{-nt} t^{k-1} dt = \frac{e^{Ck}}{Ck!}\int_0^{ne^{-C}} e^{-u} u^{k-1} du;$$

here we used the estimate $e^{-ne^{-Cj}} e^{-Ckj} \le e^{Ck} e^{-ne^{-Cx}} e^{-Ckx}$ for all $x \in [j, j+1]$, also representing the integral over the semiaxis $[0, \infty)$ as a series of integrals over the indicated segments of unit length. If $n \to \infty$ then the integral in the last expression converges monotonically to the quantity $\Gamma(k)$, which proves our assertion. Note also that a similar example is given in [9].

8. Proofs

Proof of Theorem 1. The assertion of the theorem is essentially a consequence of some results from [1,2,22,23] . First we introduce the necessary notation and recall the assertions from [22,23] we need.

Let $\{Y_i\}$ be a sequence of independent identically distributed random elements taking values in a measurable Abelian group $(\mathcal{G}, \mathcal{A})$ with measurable operation «+». Assume that the zero (neutral) element 0, as a one-point set, belongs to σ-algebra $\mathcal{A}$ and $p := \mathbf{P}(Y_1 \ne 0) \in (0,1)$. Denote by $\{Y_i^0\}$ a sequence of independent identically distributed random variables with marginal distribution

$$\mathcal{L}(Y_1^0) = \mathcal{L}(Y_1 | Y_1 \neq 0),$$

and also put $S_n := \Sigma_{i=1}^n Y_i$ and $S_n^0 := \Sigma_{i=1}^n Y_i^0$. In [1,2,22], the following assertion was obtained. □

Lemma 5. *For any natural n, the following representations are valid*:

$$\mathcal{L}(S_n) = \mathcal{L}(S^0_{\nu(n,p)}), \ \ \mathcal{L}(S_{\pi(n)}) = \mathcal{L}(S^0_{\pi(np)}), \tag{29}$$

where $\mathcal{L}(\nu(n,p)) \equiv B_{n,p}$, is the binomial distribution with parameters n and p, $\pi(t)$ is a standard Poisson process; wherein the pair $(\nu(n,p),\ \pi(np))$ does not depend on the sequence $\{X_i^0\}$.

The second important assertion gives an estimate for the Radon–Nikodim derivative of the binomial distribution with respect to the accompanying Poisson law (see [23]).

Lemma 6. *For all $p \in (0,1)$ and natural n, the following estimate holds:*

$$\sup_{k \geq 0} \frac{B_{n,p}(k)}{\mathcal{L}(\pi(np))(k)} \leq \frac{1}{1-p}. \tag{30}$$

Remark 10. *There are other estimates for this Radon–Nikodim derivative. For example, in [24], it was established that*

$$\sup_{k \geq 0} \frac{B_{n,p}(k)}{\mathcal{L}(\pi(np))(k)} \leq \frac{2}{\sqrt{1-p}}$$

for any n and $p \in (0,1)$. Note that for $p \geq 3/4$ this estimate is more accurate than (30).

It is clear that it is enough to prove the assertion for $m = 1$. A proof of the general case is carried out by induction on m and immediately follows from the total probability formula and an estimate for the conditional probability when $m-1$ coordinates of the vector $\overline{V}_{\bar{n}}$ are fixed. From (29) and (30) and the total probability formula (when the sequence $\{Y_i^0\}$ is fixed) we obtain the inequality

$$\mathcal{L}(S_n) \leq \frac{1}{1-p}\mathcal{L}(S_{\pi(n)}). \tag{31}$$

Now we put $Y_i := I_A(X_i^{(1)})$, $A \in \mathcal{A}_0$, where $\mathcal{A}_0$ is defined in (1). Consider the Abelian group

$$\mathcal{G} := \left\{ \sum_{i=1}^{k} e_i I_A(z_i),\ A \in \mathcal{A}_0;\ \forall k \geq 1,\ \forall z_i \in \mathfrak{X},\ \forall e_i = -1, 1 \right\}$$

and equip this group with the cylindric σ-algebra. It is clear that $Y_i \in \mathcal{G}$ and the following is true: $\mathbf{P}(Y_1 \neq 0) = P(A_0) = p \in (0,1)$. So, inequality (2) follows from (31) and the above-mentioned induction on m.

Proof of Theorem 2. We will carry out our reasoning in the generality and notation of the proof of Theorem 1. Both relations (29) will be the basis of construction where the sequence $\{Y_i^0\}$ is assumed to be the same in constructing the sums S_n^0 and $S^0_{\pi(n)}$ on a common probability space. So, to prove the first two assertion of the theorem, we only need to construct on the common probability space the random variables $\nu(n,p)$ and π_{np} so that they would be as close as possible to each other. The resulting probability space will be the direct product of the two probability spaces where are, respectively, defined the sequence of independent identically distributed random variables $\{Y_i^0\}$ and the above-mentioned pair of scalar indices. For the optimal definition of random indices $\nu(n,p)$ and π_{np} on a common probability space, we use Dobrushin's theorem (see [25]), which guarantees the existence of marginal copies $\nu^*(n,p)$ and π^*_{np} of the mentioned random indices defined on a common probability space so that

$$\mathbf{P}(\nu^*(n,p) \neq \pi^*_{np}) = d_{TV}(\mathcal{L}(\nu(n,p), \mathcal{L}(\pi_{np})), \tag{32}$$

where $d_{TV}(\cdot,\cdot)$ is the total variation distance between distributions. Now we use the well-known estimate of Poisson approximation to a binomial distribution (see [15,20]):

$$d_{TV}(\mathcal{L}(\nu(n,p), \mathcal{L}(\pi_{np})) \leq p \wedge (np^2) \leq p. \tag{33}$$

Using the described construction to each of the m independent coordinates of the vector point processes under consideration, we easily obtain from (32) and (33) the assertion of the theorem. □

Proof of Lemma 1. Fix a multi-index $\bar{n}$. Let us assume that the point processes $\overline{V}_{\bar{n}}$ and $\overline{\Pi}_{\bar{n}}$ are defined on the same probability space in one way or another. Then for any natural k we have the estimate

$$|\Phi_f(\overline{V}_{\bar{n}}) - \Phi_f(\overline{\Pi}_{\bar{n}})| \leq \sum_{i\geq k} |f_{i\bar{n}}(\bar{v}_{i\bar{n}}) - f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})| + \zeta_{k\bar{n}}, \tag{34}$$

where $\zeta_{k\bar{n}} := \sum\limits_{i<k} |f_{i\bar{n}}(\bar{v}_{i\bar{n}})| + \sum\limits_{i<k} |f_{i\bar{n}}(\bar{\pi}_{i\bar{n}})|$. Put $A_0 := \bigcup\limits_{i\geq k} \Delta_i$, $p(k) := \mathbf{P}(A_0) = \sum\limits_{i\geq k} p_i$. Note that the tail of the series on the right-hand side of inequality (34) is a functional of the $\mathcal{A}_0$-restrictions of the studied vector point processes defined on common probability space. So we can use Theorem 2, which guarantees the existence of an absolute coupling (depending on k) of the mentioned $\mathcal{A}_0$-restrictions with the following lower bound for the coincidence probability (see (4); here, in order not to clutter up the notation, we omit the upper symbol «*»):

$$\mathbf{P}\begin{pmatrix} (\nu^{(1)}_{n_1 k}, \nu^{(1)}_{n_1 k+1}, \ldots) = (\pi^{(1)}_{n_1 k}, \pi^{(1)}_{n_1 k+1}, \ldots) \\ (\nu^{(2)}_{n_2 k}, \nu^{(2)}_{n_2 k+1}, \ldots) = (\pi^{(2)}_{n_2 k}, \pi^{(2)}_{n_2 k+1}, \ldots) \\ \ldots \qquad \ldots \qquad \ldots \\ (\nu^{(m)}_{n_m k}, \nu^{(m)}_{n_m k+1}, \ldots) = (\pi^{(m)}_{n_m k}, \pi^{(m)}_{n_m k+1}, \ldots) \end{pmatrix}$$

$$= \mathbf{P}\left(\sup_{\Delta_j, j\geq k} \left\| V^0_{\bar{n}}(\Delta_j,\ldots,\Delta_j) - \Pi^0_{\bar{n}}(\Delta_j,\ldots,\Delta_j) \right\| = 0 \right) \geq (1-p(k))^m. \tag{35}$$

Hence, the coupling method of Theorem 2 vanishes the first term on the right-hand side of (34) with a probability no less than $(1-p(k))^m$.

Further, by virtue of estimate (2) we conclude that $\mathcal{L}(\bar{v}_{i\bar{n}})) \leq \frac{1}{(1-p_i)^m} \mathcal{L}(\bar{\pi}_{i\bar{n}})$ for any i. Therefore, by virtue of the conditions of the theorem, we have $\Delta_{\bar{n}} f_{i\bar{n}}(\nu_{i\bar{n}}) \xrightarrow{p} 0$ for any i for $\bar{n} \to \infty$. So, for any given (obviously, such construction exists) random variable $\zeta_{k\bar{n}}$ on the same probability space with the $\mathcal{A}_0$-restrictions of the point processes mentioned above, there is the relation $\Delta_{\bar{n}} \zeta_{k\bar{n}} \xrightarrow{p} 0$ for $\bar{n} \to \infty$ for any fixed k. Therefore, using the diagonal method, one can choose $k \equiv k(\bar{n}) \to \infty$ for $\bar{n} \to \infty$, for which $\Delta_{\bar{n}} \zeta_{k\bar{n}} \xrightarrow{p} 0$ as $\bar{n} \to \infty$. After constructing the point processes under consideration on a common probability space by the method of Theorem 2 for each $\bar{n}$ and already chosen $k(\bar{n})$ (in this case, obviously, $p(k(n)) \to 0$), the limit relation (13) will hold. Lemma 1 is proved. □

Proof of Theorem 3. The equivalence of items 1 and 2 directly follows from Lemma 1 and the evident two-sided estimate

$$\mathbf{P}(\xi \leq x - \varepsilon) - \mathbf{P}(|\xi - \eta| > \varepsilon) \leq \mathbf{P}(\eta \leq x) \leq \mathbf{P}(\xi \leq x + \varepsilon) + \mathbf{P}(|\xi - \eta| > \varepsilon)$$

for any $x \in \mathbb{R}$, $\varepsilon > 0$, and arbitrary random variables ξ and η defined on a common probability space. It remains to put

$$\xi := \Phi_f(\overline{V}^*_{\bar{n},\Delta_{\bar{n}}})\Delta_{\bar{n}} - M_{\bar{n}}, \quad \eta := \Phi_f(\overline{\Pi}^*_{\bar{n},\Delta_{\bar{n}}})\Delta_{\bar{n}} - M_{\bar{n}},$$

where the point processes $V^*_{\bar{n},\Delta_{\bar{n}}}$ and $\overline{\Pi}^*_{\bar{n},\Delta_{\bar{n}}}$ are defined in Lemma 1.

We now prove the equivalence of items 2 and 3 of the theorem. To this end we need to reformulate the assertion in Lemma 1 where we substitute Φ^*_f for the functional $\Phi_f(\overline{V}_{\bar{n}})$. As the resulting probability space in this assertion, we consider the direct product of the probability spaces where ν_{ni} and π_{ni} are defined by Dobrushin's theorem. We only note that, after such construction,

$$\mathbf{P}(\{\bar{\nu}^*_{i\bar{n}},\, i \geq k\} \equiv \{\bar{\pi}_{i\bar{n}},\, i \geq k\}) \geq 1 - m \sum_{i \geq k} p_i \sim 1$$

if only $k \to \infty$. Further, we repeat the corresponding reasoning in the proof of Lemma 1 (using the corresponding analog of (34)) as well as the above-mentioned arguments in proving the equivalence of items 1 and 2. □

Proof of Lemma 2. We restrict ourselves to the case $m = 2$. For an arbitrary m, the assertion can be easily proved by induction on m using analogues of the estimates that will be given below. So we have

$$\mathbf{E}\Phi_f(\overline{V}_{\bar{n}}) = \sum_{i \geq 1} \sum_{k_1,k_2 \geq 0} f_{i\bar{n}}(k_1,k_2)\mathbf{P}(\nu^{(1)}_{in_1} = k_1)\mathbf{P}(\nu^{(2)}_{in_2} = k_2),$$

$$\mathbf{E}\Phi_f(\overline{\Pi}_{\bar{n}}) = \sum_{i \geq 1} \sum_{k_1,k_2 \geq 0} f_{i\bar{n}}(k_1,k_2)\mathbf{P}(\pi^{(1)}_{in_1} = k_1)\mathbf{P}(\pi^{(2)}_{in_2} = k_2);$$

here the introduction of the operator $\mathbf{E}$ under the summation sign in the second formula is legal due to (14) and Fubini's theorem. Now, estimate the total variation distance between the distributions of the vectors $(\nu^{(1)}_{in_1}, \nu^{(2)}_{in_2})$ and $(\pi^{(1)}_{in_1}, \pi^{(2)}_{in_2})$:

$$\begin{aligned}
&\sum_{k_1,k_2 \geq 0} |\mathbf{P}(\nu^{(1)}_{in_1} = k_1)\mathbf{P}(\nu^{(2)}_{in_2} = k_2) - \mathbf{P}(\pi^{(1)}_{in_1} = k_1)\mathbf{P}(\pi^{(2)}_{in_2} = k_2)| \\
&\quad \leq \sum_{k_1,k_2 \geq 0} |\mathbf{P}(\nu^{(1)}_{in_1} = k_1) - \mathbf{P}(\pi^{(1)}_{in_1} = k_1)|\mathbf{P}(\nu^{(2)}_{in_2} = k_2) \\
&\quad + \sum_{k_1,k_2 \geq 0} |\mathbf{P}(\nu^{(2)}_{in_2} = k_2) - \mathbf{P}(\pi^{(2)}_{in_2} = k_2)|\mathbf{P}(\pi^{(1)}_{in_1} = k_1) \\
&= \sum_{k_1 \geq 0} |\mathbf{P}(\nu^{(1)}_{in_1} = k_1) - \mathbf{P}(\pi^{(1)}_{in_1} = k_1)| + \sum_{k_2 \geq 0} |\mathbf{P}(\nu^{(2)}_{in_2} = k_2) - \mathbf{P}(\pi^{(2)}_{in_2} = k_2)|.
\end{aligned}$$

We now use once more Barbour–Hall's upper bound (see [15,20]) for the total variation distance between the distributions $\mathcal{L}\left(\nu^{(j)}_{in_j}\right)$ and $\mathcal{L}\left(\pi^{(j)}_{in_j}\right)$:

$$\sum_{k_j \geq 0} |\mathbf{P}(\nu^{(j)}_{in_j} = k_j) - \mathbf{P}(\pi^{(j)}_{in_j} = k_j)| < 2p_i, \quad j = \overline{1,m}.$$

Then the total variation distance between the distributions of the bivariate vectors under consideration is estimated as follows:

$$\sum_{k_1,k_2 \geq 0} |\mathbf{P}(\nu^{(1)}_{in_1} = k_1)\mathbf{P}(\nu^{(2)}_{in_2} = k_2) - \mathbf{P}(\pi^{(1)}_{in_1} = k_1)\mathbf{P}(\pi^{(2)}_{in_2} = k_2)| \leq 4p_i.$$

Therefore,

$$\Bigg|\sum_{i\geq 1}\sum_{k_1,k_2\geq 0} f_{i\bar{n}}(k_1,k_2)\mathbf{P}(\nu^{(1)}_{in_1}=k_1)\mathbf{P}(\nu^{(2)}_{in_2}=k_2)$$

$$-\sum_{i\geq 1}\sum_{k_1,k_2\geq 0} f_{i\bar{n}}(k_1,k_2)\mathbf{P}(\pi^{(1)}_{in_1}=k_1)\mathbf{P}(\pi^{(2)}_{in_2}=k_2)\Bigg|$$

$$\leq \sum_{i\geq 1} C_i \sum_{k_1,k_2\geq 0}\Big|\mathbf{P}(\nu^{(1)}_{in_1}=k_1)\mathbf{P}(\nu^{(2)}_{in_2}=k_2)-\mathbf{P}(\pi^{(1)}_{in_1}=k_1)\mathbf{P}(\pi^{(2)}_{in_2}=k_2)\Big| \leq 4\sum_{i\geq 1} C_i p_i$$

or

$$|\mathbf{E}\Phi_f(\overline{V}_{\bar{n}})-\mathbf{E}\Phi_f(\overline{\Pi}_{\bar{n}})| \leq 4\sum_{i\geq 1} C_i p_i.$$

From here we obtain the assertion we need. □

Proof of Lemma 3. As in the proof of Lemma 2, we restrict ourselves to the case $m=2$. It is clear that we need to examine two series

$$S_1(\overline{V}_{\bar{n}}) := \sum_{i\geq 1}\sum_{k_1,k_2\geq 0} f^2_{i\bar{n}}(k_1,k_2)\mathbf{P}(\nu^{(1)}_{in_1}=k_1)\mathbf{P}(\nu^{(2)}_{in_2}=k_2),$$

$$S_2(\overline{V}_{\bar{n}}) := \sum_{i\geq 1}\Bigg(\sum_{k_1,k_2\geq 0} f_{i\bar{n}}(k_1,k_2)\mathbf{P}(\nu^{(1)}_{in_1}=k_1)\mathbf{P}(\nu^{(2)}_{in_2}=k_2)\Bigg)^2,$$

In the same way as in the proof of Lemma 1, we obtain

$$|S_1(\overline{V}_{\bar{n}})-S_1(\overline{\Pi}_{\bar{n}})| \leq 4\sum_{i\geq 1} C_i^2 p_i.$$

Similarly,

$$\begin{aligned}&|S_2(\overline{V}_{\bar{n}})-S_2(\overline{\Pi}_{\bar{n}})|\\ &\quad\leq \sum_{i\geq 1} 2C_i \sum_{k_1,k_2\geq 0} |f_{i\bar{n}}(k_1,k_2)|\Big|\mathbf{P}(\nu^{(1)}_{in_1}=k_1)\mathbf{P}(\nu^{(2)}_{in_2}=k_2)-\mathbf{P}(\pi^{(1)}_{in_1}=k_1)\mathbf{P}(\pi^{(2)}_{in_2}=k_2)\Big|\\ &\quad\leq 4\sum_{i\geq 1} C_i^2 p_i.\end{aligned}$$

From these estimates it follows that

$$|\mathbf{D}\Phi_f(\overline{\Pi}_{\bar{n}})-\mathbf{D}\Phi_f(\overline{V}_{\bar{n}})| \leq 8\sum_{i\geq 1} C_i^2 p_i,$$

whence we obtain the assertion of Lemma 2. □

Proof of Theorem 4. By Corollary 4, it suffices to present conditions for the asymptotic normality of the Poisson version of the χ^2-statistic, i.e., conditions for the feasibility of relation (16). As such, we take the Lyapunov condition of third order. Indeed, consider the following scheme of series of independent in each series of centered random variables:

$$\xi_{in} := \frac{(\pi_{in}-np_i)^2}{np_i} - 1, \quad i=1,\ldots,N(n), \quad n\geq 1.$$

The Lyapunov condition of third order, which guarantees the fulfillment of the central limit theorem (16), is as follows:

$$D_n^{-3/2}\sum_{i=1}^{N(n)} \mathbf{E}|\xi_{in}|^3 \to 0 \quad \text{as } n\to\infty. \tag{36}$$

In order to estimate the absolute third moment in (36), we need the well-known recurrence relation for the central moments of the Poisson distribution:

$$\mathbf{E}(\pi_\lambda - \lambda)^n = \lambda \sum_{k=0}^{n-2} C_{n-1}^k \mathbf{E}(\pi_\lambda - \lambda)^k, \quad n \geq 2,$$

where π_λ is a Poisson random variable with parameter λ. From here it follows that

$$\mathbf{E}(\pi_\lambda - \lambda)^6 = 15\lambda^3 + 25\lambda^2 + \lambda,$$

and using the elementary estimate $|a^2 - 1|^3 \leq 4(a^6 + 1)$, we obtain

$$\mathbf{E}|\xi_{in}|^3 \leq \frac{4}{(np_i)^3}\Big(15(np_i)^3 + 25(np_i)^2 + np_i\Big) + 4 = 64 + \frac{100}{np_i} + \frac{4}{(np_i)^2}.$$

It is clear that, to prove relation (36) it suffices to verify that, under the conditions of the theorem,

$$\frac{64N + 100\sum_{i=1}^N \frac{1}{np_i} + 4\sum_{i=1}^N \frac{1}{(np_i)^2}}{\Big(2N + \sum_{i=1}^N \frac{1}{np_i}\Big)^{3/2}}$$

$$\leq 100\left(2N + \sum_{i=1}^N \frac{1}{np_i}\right)^{-1/2} + \frac{4\sum_{i=1}^N \frac{1}{(np_i)^2}}{\Big(N + \sum_{i=1}^N \frac{1}{np_i}\Big)^{3/2}} \to 0,$$

that is true in virtue of (17). □

Proof of Theorem 5. For any natural k, denote

$$\Phi_f^{(k)}(\overline{V}_{\bar{n}}) := \sum_{i\leq k} f_{i\bar{n}}(\bar{v}_{i\bar{n}}).$$

$$\mathbf{P}\Big(\Phi_f(\overline{V}_{\bar{n}}) \geq x\Big) \leq \mathbf{P}\Big(\Phi_f^{(k)}(\overline{V}_{\bar{n}}) \geq \frac{x}{2}\Big) + \mathbf{P}\Big(\Phi_f(\overline{V}_{\bar{n}}) - \Phi_f^{(k)}(\overline{V}_{\bar{n}}) \geq \frac{x}{2}\Big). \tag{37}$$

In the notation of Theorem 1, let $V_{\bar{n}}^0$ be the restriction of the point process $\overline{V}_{\bar{n}}$ to the set $A_0 := \bigcup_{i\leq k} \Delta_i$ with hit probability $p := \sum_{i\leq k} p_i$. Under the sign of the first probability of the right-hand side of inequality (37), instead of the point process $\overline{V}_{\bar{n}}$, we can substitute $V_{\bar{n}}^0$ and use inequality (2) for the distributions of the restrictions of the corresponding point processes.

The difference

$$\Phi_f(\overline{V}_{\bar{n}}) - \Phi_f^{(k)}(\overline{V}_{\bar{n}}) = \sum_{i>k} f_{i\bar{n}}(\bar{v}_{i\bar{n}})$$

is also an additive functional of the restriction of the point process $\overline{V}_{\bar{n}}$ to the additional set $A_0 := \bigcup_{i>k} \Delta_i$ with hit probability $p := \sum_{i>k} p_i$. For this functional, we also use estimate (2). As a result, from (37) and Theorem 1, taking into account the non-negativity of the terms $f_{i\bar{n}}(\cdot)$, we obtain

$$\mathbf{P}\Big(\Phi_f(\overline{V}_{\bar{n}}) \geq x\Big) \leq \left(\sum_{i>k} p_i\right)^{-m} \mathbf{P}\Big(\Phi_f^{(k)}(\overline{\Pi}_{\bar{n}}) \geq \frac{x}{2}\Big)$$

$$+ \left(\sum_{i\leq k} p_i\right)^{-m} \mathbf{P}\Big(\Phi_f(\overline{\Pi}_{\bar{n}}) - \Phi_f^{(k)}(\overline{\Pi}_{\bar{n}}) \geq \frac{x}{2}\Big) \leq 2C^*\mathbf{P}\Big(\Phi_f(\overline{\Pi}_{\bar{n}}) \geq \frac{x}{2}\Big).$$

The theorem is proved. □

Proof of Corollary 5. is based on the following well-known equality. If ζ is a non-negative random variable with finite mean then

$$\mathbf{E}\zeta = \int_0^\infty \mathbf{P}(\zeta \geq x)dx.$$

Using successively this equality for ζ equal to $\Phi_f(\overline{V}_{\bar{n}})$ or $2\Phi_f(\overline{\Pi}_{\bar{n}})$, we easily obtain from (18) the moment inequality (20). $\square$

9. Conclusions

In this paper, we discuss a remarkable asymptotic property of a wide class of additive statistics that allows us to ignore the dependence of the summands in the additive structure of the statistics under consideration and to reduce asymptotic analysis of their distributions to the classical theory of the central limit problem. As consequences, we obtain refinements of certain results concerning the limit behavior of some known classes of additive statistics. Although we limited ourselves only to the law of large numbers and the central limit theorem for the statistics at issue, in the model under consideration it is possible to study sufficient conditions for the weak convergence of their distributions to other infinitely divisible laws as well. In fact, we deal here with a variant of Poisson approximation of empirical point processes, or in other words, with a compound Poisson approximation of an n-th partial sum of independent random variables taking values in some function space. So, in the present paper we deal with the classical subject of Probability Theory and the Poisson approximation of sums of independent multivariate random variables (for example, see [1,12,22,23]).

Moreover, one can reformulate the above-mentioned Poissonization duality theorem for more general U-statistic-type functionals

$$U_f(\overline{V}_n) := \sum_{i_1 \leq \ldots \leq i_m} f_{\bar{n},i_1,\ldots,i_m}(\overline{\nu}_{\bar{n},i_1},\ldots,\overline{\nu}_{\bar{n},i_m}),$$

where $f \equiv \{f_{\bar{n},i_1,\ldots,i_m}(\cdot)\}$ is an array of finite functions defined on Z_+^d, with $d := \sum_{k\leq m} n_k$, satisfying only the restriction

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

For example, in this more general setting, one can study the limit behavior of the functionals

$$U_I(V_n) := \sum_{i\geq 1} I_{\bar{A}}(\nu_{i-1,n}) I_A(\nu_{i,n}) \cdots I_A(\nu_{i+m-1,n}) I_{\bar{A}}(\nu_{i+m,n}),$$

where $\bar{A}$ is the complement of an arbitrary subset $A \subset \mathbb{Z}_+$, with $0 \notin A$, and $\nu_{0n} := 0$. These functionals count the number of success chains of length m in the dependent (finite or infinite) Bernoulli trials $\{I_A(\nu_{i,n});\, i \geq 1\}$.

Author Contributions: Conceptualization, I.B.; formal analysis, I.B. and M.J.; methodology, I.B.; writing—original draft, I.B. and M.J.; writing—review and editing, I.B. All authors have read and agreed to the published version of the manuscript.

Funding: The study of I. Borisov was supported by the Russian Science Foundation, project no. 22-21-00414.

Acknowledgments: The authors thank the anonymous reviewers for careful reading of the paper and insightful comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Borisov, I.S. Poisson approximation of the partial sum process in Banach spaces. *Sib. Math. J.* **1996**, *37*, 627–634. [CrossRef]
2. Borisov, I.S. Moment inequalities connected with accompanying Poisson laws in Abelian groups. *Int. J. Math. Math. Sci.* **2003**, *44*, 2771–2786. [CrossRef]
3. Medvedev, J.I. Some theorems on the asymptotic distribution of the χ^2 statistic. *Dokl. Akad. Nauk SSSR* **1970**, *192*, 987–989. (In Russian)
4. Medvedev, Y.I. Decomposable statistics in a polynomial scheme. I. *Theory Probab. Appl.* **1977**, *22*, 1–15. [CrossRef]
5. Medvedev, Y.I. Decomposable statistics in a polynomial scheme. II. *Theory Probab. Appl.* **1978**, *22*, 607–615. [CrossRef]

6. Bahadur, R.R. On the number of distinct values in a large sample from an infinite discrete distribution. *Proc. Nat. Inst. Sci. India* **1960**, *26A*, 67–75.
7. Kolchin, V.F.; Sevastyanov, B.A.; Chistyakov, V.P. *Random Assignments*; Nauka: Moscow, Russia, 1976.
8. Darling, D.A. Some limit theorems associated with multinomial trials. In *Fifth Berkeley Symposium on Mathematical Statistics and Probability*; Part 1. Berkley–Los Angelos; University of California Press: Berkeley, CA, USA , 1967; Volume 2, pp. 345–350.
9. Karlin, S. Central limit theorems for certain infinite urn schemes. *J. Math. Mech.* **1967**, *17*, 373–401. [CrossRef]
10. Chebunin, M.; Kovalevskii, A.P. Functional central limit theorems for certain statistics in an infinite urn scheme. *Stat. Probab. Lett.* **2016**, *119*, 344–348. [CrossRef]
11. Sevastyanov, B.A.; Chistyakov, V.P. Asymptotic Normality in a Classical Problem with Balls. *Theory Probab. Appl.* **1964**, *9*, 198–211. [CrossRef]
12. Barbour, A.D.; Gnedin, A.V. Small counts in the infinite occupancy scheme. *Electr. J. Probab.* **2009**, *14*, 365–384. [CrossRef]
13. Fisher, R.A.; Corbet, A.S.; Williams, C.B. The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* **1943**, *12*, 42–58 . [CrossRef]
14. Orlitsky, A.; Suresh, A.T.; Wu, Y. Supplementary Information for: Estimating the number of unseen species: A bird in the hand is worth $\log n$ in the bush. *Proc. Natl. Acad. Sci. USA* **2016**, *1511*, 07428.
15. Barbour, A.D.; Holst, L.; Janson, S. *Poisson Approximation*; Oxford University Press: Oxford, UK, 1992.
16. Kruglov, V.M. The asymptotic behavior of the Pearson statistic. *Theory Probab. Appl.* **2001**, *45*, 69–92. [CrossRef]
17. Steck, G.P. *Limit Theorems for Conditional Distributions*; University of California Press: Berkeley, CA, USA, 1957; Volume 2, pp. 237–284.
18. Tumanyan, S.K. Asymptotic distribution of χ^2-criterion when the size of observations and the number of groups simultaneously increase. *Theory Probab. Appl.* **1956**, *1*, 117–131. [CrossRef]
19. Hagerup, T.; Rüb, C. A guided tour of Chernoff bounds. *Inf. Process. Lett.* **1990**, *33*, 305–308. [CrossRef]
20. Barbour, A.D.; Hall, P. On the rate of Poisson convergence. *Math. Proc. Camb. Philos. Soc.* **1984**, *95*, 473–480. [CrossRef]
21. Loéve, M. *Probability Theory*; Nauchnaya Literatura: Moscow, Russia, 1962. (In Russian)
22. Borisov, I.S. Strong Poisson and mixed approximations of sums of independent random variables in Banach spaces. *Sib. Adv. Math.* **1993**, *3*, 1–13.
23. Borisov, I.S.; Ruzankin, P.S. Poisson approximation for expectations of unbounded functions of independent random variables. *Ann. Probab.* **2002**, *30*, 1657–1680. [CrossRef]
24. Borisov, I.S. Approximation of distributions of von Mises statistics with multidimensional kernels. *Sib. Math. J.* **1991**, *32*, 554–566. [CrossRef]
25. Dobrushin, R.L. Definition of random variables by conditional distributions. *Theory Probab. Appl.* **1970**, *15*, 458–486. [CrossRef]