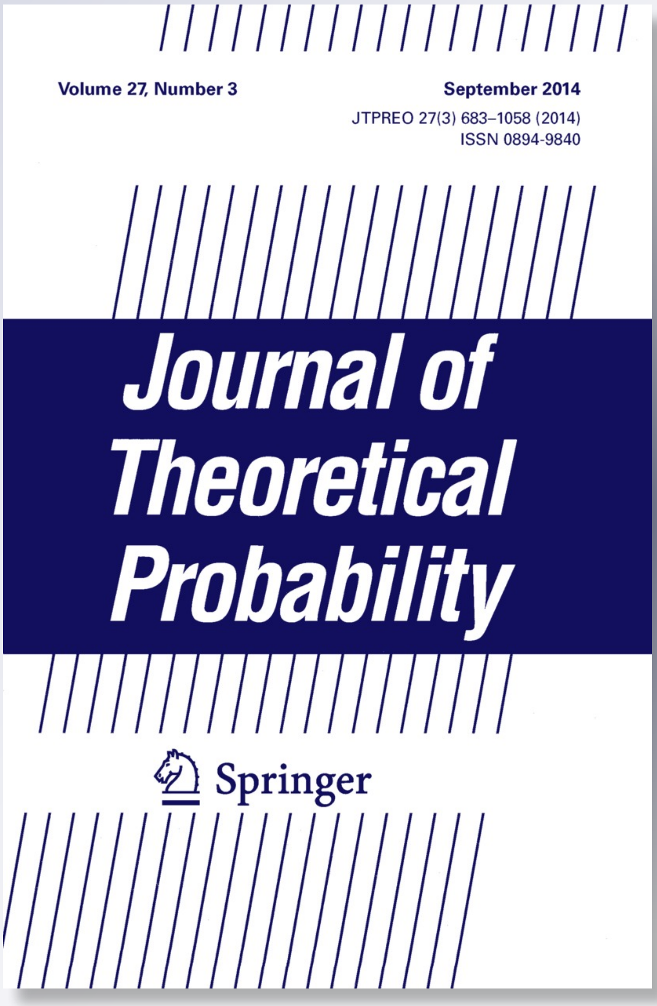
Journal of Theoretical Probability

ISSN 0894-9840
Volume 27
Number 3

J Theor Probab (2014) 27:725-753
DOI 10.1007/s10959-012-0454-z



Volume 27, Number 3 September 2014
JTPREO 27(3) 683–1058 (2014)
ISSN 0894-9840

Journal of Theoretical Probability

Springer

J Theor Probab (2014) 27:725–753
DOI 10.1007/s10959-012-0454-z

On Edgeworth Expansions in Generalized Urn Models

S.M. Mirakhmedov · S. Rao Jammalamadaka · Ibrahim B. Mohamed

Received: 21 February 2012 / Revised: 21 September 2012 / Published online: 17 October 2012

Abstract The random vector of frequencies in a generalized urn model can be viewed as conditionally independent random variables, given their sum. Such a representation is exploited here to derive Edgeworth expansions for a "sum of functions of such frequencies," which are also called "decomposable statistics." Applying these results to urn models such as with- and without-replacement sampling schemes as well as the multicolor Pólya–Egenberger model, new results are obtained for the chi-square statistic, for the sample sum in a without-replacement scheme, and for the so-called Dixon statistic that is useful in comparing two samples.

Keywords Edgeworth expansion · Urn models · Sampling with and without replacement · Pólya–Egenberger model · Poisson distribution · Binomial distribution · Negative binomial distribution · Chi-square statistic · Sample sum · Dixon statistic

Mathematics Subject Classification (2000) 62G20 · 60F05

S.M. Mirakhmedov (✉)
Institute of Mathematics and Information Technologies, Tashkent, Uzbekistan
e-mail: shmirakhmedov@yahoo.com

S. Rao Jammalamadaka
University of California, Santa Barbara, USA
e-mail: rao@pstat.ucsb.edu

I.B. Mohamed
University of Malaya, Kuala Lumpur, Malaysia
e-mail: imohamed@um.edu.my

1 Introduction

Many combinatorial problems in probability and statistics can be formulated and indeed better understood by using appropriate urn models, which are also known as random allocation schemes. Such models naturally arise in statistical mechanics, clinical trials, cryptography, etc. Properties of several types of urn models have been extensively studied in both probability and statistics literature; see e.g. the books by Johnson and Kotz [15], Kolchin et al. [18], and survey papers by Ivanov et al. [12], Kotz and Balakrishnan [19].

One of the more common urn models is the sampling scheme with replacement from a finite population which contains N objects, labeled 1 through N; the probability that the mth object will be selected in each of the sampling steps is equal to $p_m > 0$, $p_1 + \cdots + p_N = 1$. If η_m stands for the frequency of the mth object after n independent selections (i.e. in a sample of size n), then the random vector (henceforth r.v.) $(\eta_1, \dots, \eta_N)$ has a multinomial distribution with parameters $(n, p_1, \dots, p_N)$. As is well known, one important and useful property of such a multinomial r.v. is that its distribution can be represented as the joint conditional distribution of independent random variables $(\xi_1, \dots, \xi_N)$ given their sum $\xi_1 + \cdots + \xi_N = n$, where ξ_m is Poisson (υp_m) for an arbitrary positive real υ. Such a conditional representation is indeed a characteristic property of many urn models, and thus the following definition includes several commonly known urn models as special cases.

Let $\xi = (\xi_1, \dots, \xi_N)$ be an r.v. with independent and non-negative integer components such that $P\{\xi_1 + \cdots + \xi_N = n\} > 0$, for a given integer $n > 1$. Also let $\eta = (\eta_1, \dots, \eta_N)$ be an r.v. whose distribution is defined by

$$\mathcal{L}(\eta_1, \dots, \eta_N) = \mathcal{L}(\xi_1, \dots, \xi_N | \xi_1 + \cdots + \xi_N = n), \tag{1.1}$$

where $\mathcal{L}(X)$ here, and in what follows, stands for the distribution of an r.v. X. Note that (1.1) implies that $P\{\eta_1 + \cdots + \eta_N = n\} = 1$. The model defined in (1.1) is what we will call a "generalized urn model" (GUM): when a sample of size n is drawn from an urn containing N types of objects and η_m represents the number of mth type of object appearing in the sample, the distribution of the r.v. ξ defines the sample scheme through (1.1). We are interested in the following general class of statistics:

$$R_N(\eta) = \sum_{m=1}^{N} f_{m,N}(\eta_m), \tag{1.2}$$

where $f_{1,N}(x), \dots, f_{N,N}(x)$ are Borel functions defined for non-negative x. The functions $f_{m,N}$ can also be allowed to be random, in which case we will assume that the r.v. $(f_{1,N}(x_N), \dots, f_{N,N}(x_N))$ for any collection of real non-negative $x_1, \dots, x_N$ does not depend on the r.v. ξ. A statistic of the type (1.2) is called a "decomposable statistic" (DS) in the literature. For the case when the kernel functions $f_{m,N}$ are also random, the statistic (1.2) is called a "randomized DS" (see for instance [12, 20, 22]). Although the terminology DS is usually reserved for the special case when $f_{m,N}$ are not random, we will use it here for either of these cases. The following three special cases of the GUMs and related DSs are most common in applications.

A. Sample Scheme with Replacement Let $\mathcal{L}(\xi_m) = \mathrm{Poi}(\upsilon p_m)$ be a Poisson distribution with expectation υp_m, where $\upsilon \in (0, \infty)$ is arbitrary, $p_m > 0$, $m = 1, \dots, N$ and $p_1 + \dots + p_N = 1$; then the r.v. η has the multinomial distribution $M(n, p_1, \dots, p_N)$ and we have a sample scheme with replacement. This scheme is associated with the random allocation of n particles to N cells: the cells are labeled 1 through N, particles are allocated to cells independently of each other and the probability of a particle falling into mth cell is p_m, $m = 1, \dots, N$. The classical chi-square, likelihood-ratio statistic, and the empty-cells statistic are examples of the type (1.2) mentioned above.

B. Sample Scheme Without Replacement Suppose $\mathcal{L}(\xi_m) = \mathrm{Bi}(\omega_m, \upsilon)$ is a binomial distribution with parameters $\omega_m > 0$ and arbitrary $\upsilon \in (0, 1)$, $m = 1, \dots, N$, then the r.v. η has the multi-dimensional hypergeometric distribution:

$$P\{\eta_1 = k_1, \dots, \eta_N = k_N\} = \binom{\Omega_N}{n}^{-1} \prod_{m=1}^{N} \binom{\omega_m}{k_m},$$

where $\Omega_N = \omega_1 + \dots + \omega_N$, $k_1 + \dots + k_N = n$ and $0 \le k_m \le \omega_m$, $m = 1, \dots, N$. This GUM corresponds to a sampling scheme without replacement from a stratified finite population of size Ω_N. For instance, the sample sum and the standard sample-based Estimate of the Population Total, are examples of DSs of the form (1.2).

C. Multicolor Pólya–Egenberger Urn Model Let $\mathcal{L}(\xi_m) = \mathrm{NB}(d_m, \upsilon)$ be negative binomial distribution with $d_m > 0$ and arbitrary $\upsilon \in (0, 1)$, $m = 1, \dots, N$. Then

$$P\{\eta_1 = k_1, \dots, \eta_N = k_N\} = \binom{D_N + n - 1}{n}^{-1} \prod_{m=1}^{N} \binom{d_m + k_m - 1}{k_m}, \qquad (1.3)$$

where $D_N = d_1 + \dots + d_N$, is the generalized Pólya–Egenberger distribution; such a specification of the GUM corresponds to the multicolor Pólya–Egenberger urn model (see e.g. [19, Chap. 40]). For example, the number of colors that appear in the sample exactly r times and the number of pairs having the same color, are statistics of the type (1.2). We note that sum of functions of "spacing-frequencies" under the hypothesis of homogeneity of two samples can be formulated as a DS in this GUM; see, for instance, [10, 35], for further details and important applications to testing hypotheses.

There is extensive literature on DSs, much of it related to sampling with and without replacement from a finite population. We specifically mention a few: Mirakhmedov [26] obtains a bound for the remainder term in CLT and Cramer's type large deviation result for a special class of GUM; Mirakhmedov [24] and Ivchenko and Mirakhmedov [14] consider a 2-term expansion with applications to some special cases of DS in a multinomial scheme under somewhat restrictive conditions; Babu and Bai [1] obtain Edgeworth expansion for mixtures of global and local distributions—results that can be used when the DS is a linear function of frequencies and a GUM is defined by *identically* distributed r.v.s ξ_m. Such results are clearly very restrictive on the parameters of the urn model and on the kernel functions $f_{m,N}$.

The aim of this paper is threefold: First, we present a general approach that allows one to obtain an Edgeworth asymptotic expansion to *any number of terms*, for the distribution of a DS in a GUM. Second, this general approach is used to extend known results for classes of DS in the three special cases of GUM just mentioned. Third, we illustrate these results by obtaining general Edgeworth expansions for three special and interesting cases of DS, viz.:

(i) the chi-square statistic in Case A,
(ii) sample-sum in a sample scheme without replacement, i.e. in Case B, and
(iii) the Dixon spacing-frequencies statistic in Case C.

The chi-square statistic is considered for the case when the number of groups increases along with the sample size, a situation that has been considered by some authors, including [8, 29, 33, 34] and [23]. We obtain here a 3-term asymptotic expansion under very general conditions on the parameters, generalizing the results in [23] and [14]. The result in (ii) improves the main results of [5, 21, 36, 37], as well as parts of Theorem 1 of [11]. Asymptotic expansions for a DS in the multicolor Pólya–Egenberger urn model and for the Dixon statistic as a special case are obtained here for the first time.

It should be remarked that although we confine our discussion in this paper to the above three examples of GUM and related DS for illustrative purposes as well as to keep the length of the paper reasonable, it should be mentioned that the results derived in this paper are generally applicable to any DS in other specifications of GUM, for instance to the context of specified random forests, random cyclic substitutions (cf. [17, 31]).

The paper is organized as follows. In Sect. 2 we present a systematic procedure for obtaining an asymptotic expansion for the characteristic function of a DS, to terms of any order. Our general approach is based on the so-called Bartlett's type integral formula and provides a simpler and more streamlined way of obtaining higher-order approximations than what previous authors have used. The main results are presented in Sect. 3. For the sake of completeness and to connect to Bartlett's type formula, we also present two theorems on asymptotic normality and Berry–Esseen type bounds, showing how the current formulation helps simplifying similar results obtained in [25, 26]. Applications to the special DSs (i), (ii), and (iii) are given in Sect. 4, while the proofs of the main results are postponed to an Appendix.

It should be mentioned that we are dealing with triangular arrays where all the parameters of a GUM vary (including the distribution of the r.v.s ξ_m) when both n and N tend to infinity, formally through a non-decreasing sequence of positive integers $\{n_v\}$, $\{N_v\}$, as $v \to \infty$; hence it is important to express the remainder terms in our asymptotic expansions which show their explicit dependence on the n, N, distributions of the r.v.s ξ_m and the kernel functions $f_{m,N}$.

In what follows, c and C with or without index are universal positive constants which may depend on the argument and may be different at different places; all asymptotic relations and limits are considered as $n \to \infty$, and $N = N(n) \to \infty$.

2 Bartlett's Type Formula and Asymptotic Expansion of the Characteristic Function of a DS

We now define the following quantities:

$$\begin{aligned}
&A_N=\sum_{m=1}^N E\xi_m, \qquad B_N^2=\sum_{m=1}^N \operatorname{Var}\xi_m, \qquad x_N=(n-A_N)/B_N,\\
&\Lambda_N=\sum_{m=1}^N Ef_m(\xi_m), \qquad \gamma_N=\frac{1}{B_N^2}\sum_{m=1}^N \operatorname{cov}\bigl(f_m(\xi_m),\xi_m\bigr),\\
&g_m(y)=f_m(y)-Ef_m(\xi_m)-\gamma_N(y-E\xi_m), \qquad \hat R_N(\eta)=\sum_{m=1}^N g_m(\eta_m),\\
&\sigma_N^2=\sum_{m=1}^N \operatorname{Var}\, g_m(\xi_m)=\sum_{m=1}^N \operatorname{Var} f_m(\xi_m)-B_N^2\gamma_N^2.
\end{aligned} \tag{2.1}$$

Under some mild conditions (see for instance [13]), one can show that as $n\to\infty$ and $N=N(n)\to\infty$,

$$\begin{aligned}
&\mathrm{ER}_N(\eta)=\Lambda_N+x_NB_N\gamma_N-\frac{1-x_N^2}{2B_N^2}\sum_{m=1}^N \mathrm{Eg}_m(\xi_m)(\xi_m-E\xi_m)^2\bigl(1+o(1)\bigr),\\
&\operatorname{Var} R_N(\eta)=\sigma_N^2\bigl(1+o(1)\bigr).
\end{aligned}$$

(These expressions can also be derived by putting formally $t=0$, $j=1$ and $j=2$ in Proposition 2.1 below; see Remark 2.1.) Also, $\hat R_N(\eta)=R_N(\eta)-\Lambda_N-x_NB_N\gamma_N$ and

$$\sum_{m=1}^N \mathrm{Eg}_m(\xi_m)=0, \qquad \sum_{m=1}^N \operatorname{cov}\bigl(g_m(\xi_m),\xi_m\bigr)=0. \tag{2.2}$$

Let ϕ be a measurable function such that $E|\phi(\xi_1,\xi_2,\dots,\xi_N)|<\infty$ and $\zeta_N=\xi_1+\cdots+\xi_N$. We have $E(\phi(\xi_1,\dots,\xi_N)|\zeta_N=n)=E\phi(\eta_1,\dots,\eta_N)$, because of (1.1). This, together with

$$E\bigl(\phi(\xi_1,\dots,\xi_N)e^{i\tau(\zeta_N-n)}\bigr)=\sum_{k=0}^\infty e^{i\tau(k-n)}P\{\zeta_N=k\}E\bigl(\phi(\xi_1,\dots,\xi_N)|\zeta_N=k\bigr),$$

implies, by Fourier inversion,

$$E\phi(\eta_1,\eta_2,\dots,\eta_N)=\frac{1}{2\pi P\{\zeta_N=n\}}\int_{-\pi}^{\pi} E\phi(\xi_1,\xi_2,\dots,\xi_N)\exp\bigl\{i\tau(\zeta_N-n)\bigr\}\,d\tau. \tag{2.3}$$

Set

$$\Theta_N(t,x_N)=\frac{1}{\sqrt{2\pi}}\int_{-\pi B_N}^{\pi B_N} e^{-i\tau x_N}\Psi_N(t,\tau)\,d\tau, \tag{2.4}$$

where

$$\Psi_N(t,\tau)=\prod_{m=1}^{N} E\exp\bigl\{it\sigma_N^{-1}g_m(\xi_m)+i\tau B_N^{-1}(\xi_m-E\xi_m)\bigr\}.$$

Then (2.3) together with the inversion formula for the local probability $P\{\zeta_N=n\}$ gives us the following Bartlett's type formula (cf. Bartlett [3]):

$$\varphi_N(t,x_N)=:Ee^{it\sigma_N^{-1}\hat{R}_N(\eta)}=\frac{\Theta_N(t,x_N)}{\Theta_N(0,x_N)} \tag{2.5}$$

which provides the crucial formula of interest. Special formulations of this formula show up in literature; see e.g. [9, 22–24, 33]. Also, a very special case of (2.5) is the most commonly used formula of [6] for investigating the sample sum in a without-replacement scheme (see e.g. [2, 11, 38]). Formula (2.3) is also useful in studying large deviation problems (see e.g. [26]).

A formal construction of the asymptotic expansion for $\varphi_N(t,x_N)$ defined in (2.5), proceeds as follows: The integrand $\Psi_N(t,\tau)$is the characteristic function (ch.f.) of the sum of N independent two-dimensional r.v.s (g_m,ξ_m). Because of (2.2), this sum has zero expectation, a unit covariance matrix and uncorrelated components. From [4, Chap. 2], it is well known that under suitable conditions, this ch.f. $\Psi_N(t,\tau)$ can be approximated by a power-series in $N^{-1/2}$ whose coefficients are polynomials in t and τ containing the common factor $\exp\{-(t^2+\tau^2)/2\}$. Hence the series can be integrated wrt τ over the interval $(-\infty,\infty)$. As a result of this integration, we get a power series, say $\mathrm{H}_N(t,x_N)$, in $N^{-1/2}$. Next, we replace $\Theta_N(0,x_N)$ by its series approximation, which is $\mathrm{H}_N(0,x_N)$. Finally, we get the asymptotic expansion of $\varphi_N(t,x_N)$ by dividing $\mathrm{H}_N(t,x_N)$ by $\mathrm{H}_N(0,x_N)$.

The above algorithm, although manageable, needs long and complex calculations as we show below. Assume that $E|g_m(\xi_m)|^s<\infty$ and $E|\xi_m|^s<\infty$ for some $s\geq 3$. Let $P_{m,N}(t,\tau)$, $m=1,2,\ldots$, be the well-known polynomials in t and τ from the theory of the asymptotic expansion of the ch.f. of the sum of independent random vectors (see (7.3), (7.6) of [4], p. 52), in our case for the quantity $(g_1,\xi_2)+\cdots+(g_N,\xi_N)$; the degree of $P_{m,N}(t,\tau)$ is $3m$ and the minimal degree is $m+2$; the coefficients of $P_{m,N}(t,\tau)$ only involve the cumulants of the r.v.s $(g_1,\xi_2),\ldots,(g_N,\xi_N)$ of order $m+2$ and less. Define polynomials (in t) of $G_{k,N}(t,x_N)$ as

$$G_{k,N}(t,x_N)=\frac{e^{x_N^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}P_{k,N}(t,\tau)\exp\left\{-i\tau x_N-\frac{\tau^2}{2}\right\}d\tau,\quad k=0,1,2,\ldots. \tag{2.6}$$

Now define $Q_{j,N}(x_N)$ from the equation

$$\sum_{k=0}^{\infty}(-1)^k\left(\sum_{v=0}^{s-3}N^{-v/2}G_{v,N}(0,x_N)\right)^k=\sum_{j=0}^{\infty}N^{-j/2}Q_{j,N}(x_N).$$

Then

$$Q_{j,N}(x_N) = j!\sum\prod_{i=1}^{s-3}\frac{1}{j_i!}G_{i,N}^{j_i}(0,x_N), \tag{2.7}$$

where the summation is over all $(s-3)$-tuples $(j_1, j_2, \dots, j_{s-3})$ with non-negative integers j_i such that $j_1 + 2j_2 + \dots + (s-3)j_{s-3} = j$. Let

$$W_N^{(s)}(t,x_N) = \sum_{m=0}^{s-3} N^{-m/2}\sum_{v=0}^{m} G_{v,N}(t,x_N)Q_{m-v,N}(x_N). \tag{2.8}$$

Note that $G_{0,N}(t,x_N) = 1$, $Q_{0,N}(x_N) = 1$ so that $W_N^{(3)}(t,x_N) = 1$. For example

$$\begin{aligned} W_N^{(5)}(t,x_N) = 1 &+ \frac{1}{\sqrt{N}}\big(G_{1,N}(t,x_N) - G_{1,N}(0,x_N)\big) \\ &+ \frac{1}{N}\big(G_{2,N}(t,x_N) - G_{2,N}(0,x_N) \\ &- G_{1,N}(0,x_N)\big(G_{1,N}(t,x_N) - G_{1,N}(0,x_N)\big)\big). \end{aligned} \tag{2.9}$$

In what follows, we will need the following additional notation:

$$\begin{aligned} &\hat\sigma_N^2 = N^{-1}\sigma_N^2, \qquad \hat B_N^2 = N^{-1}B_N^2, \qquad \hat g_m = g_m(\xi_m)/\hat\sigma_N, \\ &\hat\xi_m = (\xi_m - E\xi_m)/\hat B_N, \\ &\beta_{j,N} = N^{-j/2}\sum_{m=1}^{N} E|\hat g_m|^j, \qquad \kappa_{j,N} = N^{-j/2}\sum_{m=1}^{N} E|\hat\xi_m|^j, \\ &M_N(T) = \inf_{T\le|\tau|\le\pi}\sum_{m=1}^{N}\big(1 - \big|E\exp\{i\tau\xi_m\}\big|^2\big) \quad \text{if } T\le\pi, \quad \text{else} \quad M_N(T) = \infty, \\ &\Upsilon_{s,N} = \beta_{s,N} + \kappa_{s,N} + B_N^2\exp\left\{-\frac{1}{8}M_N\big(0.3(B_N\kappa_{3,N})^{-1}\big)\right\}, \\ &T_N = \min\big(\beta_{3,N}^{-1}, \mathcal{E}_N^{-1}(1)\big), \\ &\mathcal{E}_N(\delta) = \frac{1}{(M_N(0.3(B_N\kappa_{2+\delta,N}^{1/\delta})^{-1}))^{1/2}} + \frac{\min(B_N,\sqrt{N})}{M_N(0.3(B_N\kappa_{2+\delta,N}^{1/\delta})^{-1})}, \quad 0<\delta\le 1. \end{aligned} \tag{2.10}$$

Throughout the paper we assume that $|x_N| \le c$, although the method used here allows us to let x_N to increase at a rate of $O(\sqrt{\log N})$ (see e.g. [25]). In the above-listed three examples of GUM the parameter υ can be chosen such that $x_N = 0$ (see also the beginning of Sect. 4). We now have the following result which is proved in the Appendix:

Proposition 2.1 *Let $E|g_m(\xi_m)|^s < \infty$, for some $s \geq 3$, $m = 1, \ldots, N$ and $\Upsilon_{s,N} \leq 0.01$. There exist constants c and C such that if $|t| \leq cT_N$, then for $j = 0, 1$,*

$$\left| \frac{\partial^j}{\partial t^j} \left(\varphi_N(t, x_N) - e^{-\frac{t^2}{2}} W_N^{(s)}(t, x_N) \right) \right| \leq C e^{-\frac{t^2}{8}} \Upsilon_{s,N}.$$

Remark 2.1 It may be remarked that Proposition 2.1 can be further extended for any $j = 0, 1, \ldots, s$, but at the expense of added complexity in the proof. Such an extension of Proposition 2.1 can be used to derive asymptotic expansion of the moments of the statistic $R_N(\eta)$.

3 Main Results

We use the notation defined in Sect. 2. Theorems 3.1 and 3.2 follow from Theorems 1 and 2 of Mirakhmedov [25] and are presented here for the sake of completeness, and to connect to Bartlett's type formula (2.5); also, their application to DS in our examples of GUM gives weaker conditions for asymptotic normality and improved Berry–Esseen type bound.

Let $\mathbf{I}\{A\}$ stand for the indicator function of the set A and

$$\begin{aligned} \mathcal{L}_{1,N}(\varepsilon) &= \frac{1}{N^{3/2}} \sum_{m=1}^{N} E|\hat{\xi}_m|^3 \mathbf{I}\{|\hat{\xi}_m| \leq \varepsilon\}, \\ \mathcal{L}_{2,N}(\varepsilon) &= \frac{1}{N} \sum_{m=1}^{N} E\hat{\xi}_m^2 \mathbf{I}\{|\hat{\xi}_m| > \varepsilon\}, \qquad (3.1) \\ L_{2,N}(\varepsilon) &= \frac{1}{N} \sum_{m=1}^{N} E\hat{g}_m^2 \mathbf{I}\{|\hat{g}_m| > \varepsilon\}. \end{aligned}$$

Theorem 3.1 *If for arbitrary $\varepsilon > 0$*

(i) $L_{2,N}(\varepsilon) \to 0$,
(ii) $\mathcal{L}_{2,N}(\varepsilon) \to 0$,
(iii) $M_N(\pi(4B_N\mathcal{L}_{1,N}(\varepsilon))^{-1}) \to \infty$,
(iv) $\min(B_N, \sqrt{N}) = o(M_N(\pi(4B_N\mathcal{L}_{1,N}(\varepsilon))^{-1}))$,

then the statistic $R_N(\eta)$ has an asymptotic normal distribution with expectation $\Lambda_N + x_N B_N \gamma_N$ and variance σ_N^2, given in (2.1).

Remark 3.1 For all the three examples of GUM we consider, conditions (ii), (iii) and (iv), being conditions on the parameters of the urn model, are automatically satisfied under very general set-up (see Sect. 4), so that all we need is to check the Lindeberg's condition (i) for ensuring the asymptotic normality of the DS.

Let $E|g_m(\xi_m)|^s < \infty$, for some $s \geq 3$. Define $\mathbb{W}_N^{(s)}(u, x_N)$ so that

$$\int_{-\infty}^{\infty} e^{itu}\, d\mathbb{W}_N^{(s)}(u, x_N) = W_N^{(s)}(t, x_N) e^{-\frac{t^2}{2}}. \qquad (3.2)$$

The function $\mathbb{W}_N^{(s)}(u, x_N)$ can be obtained by formally substituting

$$(-1)^v \frac{d^v}{du^v}\Phi(u) = -e^{-u^2/2}H_{v-1}(u)/\sqrt{2\pi}, \quad \text{where } \Phi(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-\frac{t^2}{2}}\,dt,$$

for $(it)^\nu e^{-t^2/2}$ for each ν in the expression for $W_N^{(s)}(t, x_N)e^{-t^2/2}$ (see Lemma 7.2 of [4, p. 53]), where $H_v(x)$is the νth order Hermite–Chebishev polynomial. Note that $\mathbb{W}_N^{(3)}(u, x_N) = \Phi(u)$. Set

$$\Delta_N^{(s)} = \sup_{-\infty<u<\infty}\left|P\left\{R_N(\eta) < u\sigma_N + \Lambda_N + x_N B_N \gamma_N\right\} - \mathbb{W}_N^{(s)}(u, x_N)\right|,$$

$$\chi_N(a, b) = \mathbf{I}\{a < b\}\int_{a\le|t|\le b}\left|\frac{\varphi_N(t, x_N)}{t}\right|dt.$$

Theorem 3.2 *Let* $0 < \delta \le 1$. *Then, there exists a constant* C *such that* $\Delta_N^{(3)} \le C(\beta_{2+\delta,N} + \kappa_{2+\delta,N} + \mathcal{E}_N(\delta))$.

Theorem 3.3 *Let* $E|g_m(\xi_m)|^s < \infty$, *for some* $s \ge 3$, $m = 1, 2, \ldots, N$. *There exist constants* c *and* C *such that* $\Delta_N^{(s)} \le C\Upsilon_{s,N} + \chi_N(cT_N, \beta_{s,N}^{-1})$.

Theorem 3.4 *Let the statistic* $R_N(\eta)$ *be a lattice r.v. with span* h *and a set of possible values in* $\Re$. *If* $E|g(\xi_m)|^s < \infty$, *for some* $s \ge 3$, $m = 1, 2, \ldots, N$, *then there exist constants* c *and* C *such that uniformly in* $z \in \Re$

$$\sup_{z\in\Re}\left|\frac{\sigma_N}{h}P\left\{R_N(\eta) = z\right\} - \frac{d}{du_z}\mathbb{W}_N^{(s)}(u_z, x_N)\right| \le C\Upsilon_{s,N} + \bar{\chi}_N(cT_N, \pi\sigma_N/h),$$

where $u_z = (z - \Lambda_N - x_N B_N \gamma_N)/\sigma_N$ *and*

$$\bar{\chi}_N(a, b) = \mathbf{I}\{a < b\}\int_{a\le|t|\le b}\left|\varphi_N(t, x_N)\right|dt.$$

The following general bounds for $\chi_N(a, b)$ are useful in applications. Write

$$\psi_m(t, \tau) = E\exp\left\{itf_{m,N}(\xi_m) + i\tau\xi_m\right\},$$

$$d_N(a, b) = 1 - \sup_{\substack{a\sigma_N^{-1}\le|t|\le b\sigma_N^{-1}\\ |\tau|\le\pi}} N^{-1}\sum_{m=1}^{N}\left|\psi_m(t, \tau)\right|^2, \tag{3.3}$$

$$H_m(t, \tau) = E\left(tf_{m,N}^*(\xi_m) + \tau\xi_m^*\right)^2,$$

$$\bar{H}_N(a, b) = \inf_{\substack{a\sigma_N^{-1}\le|t|\le b\sigma_N^{-1}\\ |\tau|\le\pi}} \frac{1}{N}\sum_{m=1}^{N} H_m\left(\frac{t}{2\pi}, \frac{\tau}{2\pi}\right), \tag{3.4}$$

where $\langle a\rangle$ stands for the distance between real a and integers. Here and in what follows, for a given r.v. ζ we define $\zeta^* = \zeta - \zeta'$, where ζ' is an independent copy of ζ. Then

$$\chi_N(a,b) \le C B_N \ln\bigl(b\sigma_N^{-1}\bigr)\exp\biggl\{-\frac{1}{2}N d_N(a,b)\biggr\}, \tag{3.5}$$

$$\chi_N(a,b) \le C B_N \ln\bigl(b\sigma_N^{-1}\bigr)\exp\biggl\{-\frac{1}{2}N \bar{H}_N(a,b)\biggr\} \tag{3.6}$$

and

$$\bar{\chi}_N(a,b) \le C\sigma_N B_N \exp\biggl\{-\frac{1}{2}N \bar{H}_N(a,b)\biggr\}. \tag{3.7}$$

These inequalities (3.5)–(3.7) follow from the following arguments: From formula (2.3) it follows that for the ch.f. $\varphi_N(t,x_N)$ in (2.5), one can write the product $\prod_{m=1}^N \psi_m(t,\tau)$instead of $\Psi_N(t,\tau)$. Since $\Theta_N(0,x_N) \ge c$ (cf. (5.7) below), inequality (3.5) follows by using the fact that $x < e^{(x^2-1)/2}$. On the other hand, by Lemma 4 of [30], we have

$$4H_m\biggl(\frac{t}{2\pi},\frac{\tau}{2\pi}\biggr) \le 1 - \bigl|\psi_m(t,\tau)\bigr|^2 \le 2\pi^2 H_m\biggl(\frac{t}{2\pi},\frac{\tau}{2\pi}\biggr).$$

This inequality together with (3.5) implies the inequalities (3.6) and (3.7).

Remark 3.2 A DS of the special form

$$X_N^2 = \sum_{m=1}^N \eta_m^2$$

arises in many problems in statistics and in discrete probability (see e.g. Sect. 4, and [31]). This DS is a lattice r.v. with span equal to 2. Also,

$$H_m(t,\tau) = \sum_{k,l}\langle v_{k,l}\rangle^2 P(\xi_m = k)P\bigl(\xi_m' = l\bigr), \tag{3.8}$$

where $v_{k,l} = (k-l)((k+l)t+\tau)$. As in Lemma 2 of [31], one can prove that for all real t and τ such that $|t| \le 1/4$, $|\tau| \le 1/2$ and any non-negative integer k and l,

$$\max\bigl\{\langle v_{k,l}\rangle, \langle v_{k+1,l}\rangle, \langle v_{k+2,l}\rangle\bigr\} \ge \frac{|t|}{2}.$$

From this it follows that if

$$\sum_{l=0}^\infty\sum_{j=0}^\infty P\{\xi_m = k_{j,l}\}P\bigl\{\xi_m' = l\bigr\} \ge c > 0, \tag{3.9}$$

then $\bar{H}_N(a,\pi\sigma_N/2) \ge a^2/4\sigma_N^2$, where for each $l = 0,1,2,\dots,k_{j,l}$ is defined such that $\max\{\langle v_{3j,l}\rangle, \langle v_{3j+1,l}\rangle, \langle v_{3j+2,l}\rangle\} = \langle v_{k_{j,l},l}\rangle$, $j = 0,1,2,\dots$.

4 Applications

In what follows, we will use the notation of the preceding sections, keeping in mind that the distribution of the r.v. ξ_m is what is relevant for the particular GUM under consideration. Note that in all our examples of the GUM, the distributions of the r.v.'s ξ_m depend on an arbitrary parameter υ, which can be chosen in a suitably convenient manner. We will thus choose the parameter υ such that $A_N = n$, in which case $x_N = 0$, and hence the terms of asymptotic expansion, i.e. the function $\mathbb{W}_N^{(s)}(u, 0)$ are considerably simplified. For example, it is known that (see [4, pp. 52 and 55])

$$P_{0,N}(t,\tau) = 1,$$

$$P_{1,N}(t,\tau) = \frac{i^3}{6N}\sum_{m=1}^{N} E(t\hat{g}_m + \tau\hat{\xi}_m)^3,$$

$$P_{2,N}(t,\tau) = \frac{i^4}{24N}\sum_{m=1}^{N}\big(E(t\hat{g}_m + \tau\hat{\xi}_m)^4 - 3\big(E(t\hat{g}_m + \tau\hat{\xi}_m)^2\big)^2\big) + \frac{1}{2}P_{1,N}^2(t,\tau).$$

Therefore, from (2.6), (2.9), and (3.2),

$$\begin{aligned}\mathbb{W}_N^{(5)}(u,0) = \varPhi(u) &- \frac{e^{-u^2/2}}{\sqrt{2\pi N}}\left(\frac{u^2-1}{6}\alpha_{3,0,N} - \frac{1}{2}\alpha_{1,2,N}\right)\\ &- \frac{e^{-u^2/2}}{\sqrt{2\pi}N}\left\{\frac{u^5 - 10u^3 + 15u}{72}\alpha_{3,0,N}^2\right.\\ &+ \frac{u^3-3u}{24}\left(\alpha_{4,0,N} - \frac{3}{N}\sum_{m=1}^{N}\hat{\alpha}_{20m}^2 - 3\alpha_{2,1,N}^2 - 2\alpha_{3,0,N}\alpha_{1,2,N}\right)\\ &+ \frac{u}{8}\left(3\alpha_{1,2,N}^2 + 2\alpha_{2,1,N}\alpha_{0,3,N} - 2\alpha_{2,2,N}\right.\\ &\left.\left.+ \frac{4}{N}\sum_{m=1}^{N}\hat{\alpha}_{11m}^2 + \frac{2}{N}\sum_{m=1}^{N}\hat{\alpha}_{20m}\hat{\alpha}_{02m}\right)\right\},\end{aligned}\tag{4.1}$$

where $\hat{\alpha}_{ijm} = E\hat{g}_m^i\hat{\xi}_m^j$, $\alpha_{i,j,N} = N^{-1}\sum_{m=1}^{N}\hat{\alpha}_{ijm}$.

In what follows, we will restrict ourselves to such a 3-term asymptotic expansion given above, just to keep our calculations simple.

4.1 Example A

The r.v. $\eta = (\eta_1, \ldots, \eta_N)$ has the multinomial distribution $M(n, p_1, \ldots, p_N)$, $p_m > 0$, $m = 1, \ldots, N$, $p_1 + \cdots + p_N = 1$, and we take $\mathcal{L}(\xi_m) = \mathrm{Poi}(np_m)$. We assume that $N = N(n) \to \infty$, $\max_{1\le m\le N} p_m \to 0$ as $n \to \infty$. We take $\lambda = n/N$, $\lambda_m = np_m$ and $\mathrm{P}_{iN} = p_1^i + \cdots + p_N^i$.

In this classical scheme, since the best conditions for asymptotic normality and the Berry–Esseen type bound of DS are already given in [26, 28], we concentrate our attention on the asymptotic expansion results.

Theorem 4.1 *Let the statistic $R_N(\eta)$ be a lattice r.v. with span h and a set of possible values $\Re$. If $E|g(\xi_m)|^5 < \infty, m = 1, 2, \ldots, N$ and $\mathrm{P}_{2N} \le (10 \ln n)^{-1}$, then uniformly in $z \in \Re$*

$$\left|\frac{\sigma_N}{h} P\bigl\{R_N(\eta) = z\bigr\} - \frac{d}{du_z}\mathbb{W}_N^{(5)}(u_z, 0)\right|$$
$$\le C\biggl(\beta_{5,N} + \bigl(\mathrm{P}_{3N} + n^{-2}\bigr)^{3/4} + \sigma_N\sqrt{n}\exp\biggl\{-\frac{1}{2}N\bar{H}_N\biggl(cT_N, \frac{\pi\sigma_N}{h}\biggr)\biggr\}\biggr),$$

where $u_z = (z - \Lambda_N)/\sigma_N$ and T_N is defined as in Sect. 3 with

$$\mathcal{E}_N(1) = \sqrt{n^{-1} + \mathrm{P}_{2N}}\bigl(1 + \min\bigl(1, \lambda^{-1/2}\bigr)\sqrt{1 + n\mathrm{P}_{2N}}\bigr).$$

The particular DS $X_N^k := \sum_{m=1}^N \eta_m^k$ for any integer $k > 1$ is a special case of Theorem 4.1. We shall focus on the most important application, the chi-square type statistic X_N^2. As stated before, X_N^2 is the lattice with span equal to 2; also in this case $g_m(\xi_m) = (\xi_m^2 - \lambda_m(\lambda_m + 1)) - (2n\mathrm{P}_{2N} + 1)(\xi_m - \lambda_m)$. Hence

$$\begin{aligned}
\Lambda_N &= n(1 + n\mathrm{P}_{2N}),\\
\sigma_N^2 &= 2n^2\mathrm{P}_{2N} + 4n^3\bigl(\mathrm{P}_{3N} - \mathrm{P}_{2N}^2\bigr) = N\bigl(2n\lambda\mathrm{P}_{2N} + 4n^2\lambda\bigl(\mathrm{P}_{3N} - \mathrm{P}_{2N}^2\bigr)\bigr) := N\hat{\sigma}_N^2,\\
\alpha_{12N} &= 2\lambda\hat{\sigma}_N^{-1}\mathrm{P}_{2N}, \qquad \alpha_{21N} = 4\sqrt{n}\lambda\hat{\sigma}_N^{-2}\bigl(\mathrm{P}_{2N} + 12n\bigl(\mathrm{P}_{3N} - \mathrm{P}_{2N}^2\bigr)\bigr),\\
\alpha_{30N} &= n\lambda\hat{\sigma}_N^{-3}\bigl[4\mathrm{P}_{2N} + 2n\bigl(16\mathrm{P}_{3N} - 9\mathrm{P}_{2N}^2\bigr) + 8n^2\bigl(4\mathrm{P}_{4N} - 9\mathrm{P}_{2N}\mathrm{P}_{3N} + 5\mathrm{P}_{2N}^3\bigr)\bigr],\\
\alpha_{40N} &= \hat{\sigma}_N^{-4}n\lambda\bigl[8\mathrm{P}_{2N} + n\bigl(164\mathrm{P}_{2N}^2 - 17\mathrm{P}_{3N}\bigr)\\
&\quad + n^2\bigl(636\mathrm{P}_{4N} - 768\mathrm{P}_{2N}\mathrm{P}_{3N} + 192\mathrm{P}_{2N}^3\bigr)\\
&\quad + n^3\bigl(448\mathrm{P}_{5N} - 1120\mathrm{P}_{2N}\mathrm{P}_{4N} + 912\mathrm{P}_{2N}^2\mathrm{P}_{3N} - 240\mathrm{P}_{2N}^4\bigr)\\
&\quad + 48n^4\bigl(\mathrm{P}_{6N} - 4\mathrm{P}_{2N}\mathrm{P}_{5N} + 6\mathrm{P}_{2N}^2\mathrm{P}_{4N} - 4\mathrm{P}_{2N}^3\mathrm{P}_{3N} + \mathrm{P}_{2N}^5\bigr)\bigr],\\
\alpha_{22N} &= \bigl(\lambda\hat{\sigma}_N^2\bigr)^{-1}\lambda\bigl[8n\mathrm{P}_{2N} + 2n^2\bigl(19\mathrm{P}_{3N} - 14\mathrm{P}_{2N}^2\bigr)\\
&\quad + 12n^3\bigl(\mathrm{P}_{4N} - 2\mathrm{P}_{2N}\mathrm{P}_{3N} + \mathrm{P}_{2N}^3\bigr) - 1\bigr],\\
\frac{1}{N}\sum_{m=1}^N \tilde{\alpha}_{20m}^2 &= \hat{\sigma}_N^{-4}n\lambda\bigl[3\mathrm{P}_{2N} - 2N\mathrm{P}_{3N} + 4N^2\mathrm{P}_{4N}\\
&\quad + 16N^3(\mathrm{P}_{5N} - 2\mathrm{P}_{4N}\mathrm{P}_{2N} + \mathrm{P}_{3N}\mathrm{P}_{2N})\\
&\quad + 16N^4\bigl(\mathrm{P}_{6N} - 4\mathrm{P}_{2N}\mathrm{P}_{5N} + 6\mathrm{P}_{4N}\mathrm{P}_{2N}^2 - 4\mathrm{P}_{3N}\mathrm{P}_{2N}^3 + \mathrm{P}_{2N}^5\bigr)\bigr],
\end{aligned}$$

$$\frac{4}{N}\sum_{m=1}^{N}\tilde{\alpha}_{11m}^{2}+\frac{2}{N}\sum_{m=1}^{N}\tilde{\alpha}_{20m}\tilde{\alpha}_{02m}$$
$$=\big(\lambda\hat{\sigma}_N^2\big)^{-1}\big[2n^2\lambda \mathrm{P}_{3N}+2n^3\lambda\big(2\mathrm{P}_{4N}-6\mathrm{P}_{3N}\mathrm{P}_{2N}+3\mathrm{P}_{2N}^3\big)\big].$$

Corollary 4.1 *Let $c_1 \le Np_m \le c_2$ for some positive c_1, c_2 and all $m = 1, \dots, N$; then uniformly in $b \in \{n+2k, k = 0, 1, \dots, n(n-1)/2\}$, the set of possible values of the r.v. X_N^2, one has*

$$\left|\frac{\sigma_N}{2}P\big\{X_N^2=b\big\}-\frac{d}{du_b}\mathbb{W}_N^{(5)}(u_b,0)\right|$$
$$\le C\left(\frac{1}{N^{3/2}}+\frac{1}{(n\lambda)^{3/2}}+n\lambda\exp\left\{-\frac{cN}{\lambda\max(1,\lambda)}\right\}\right),$$

where $u_b = (b - \Lambda_N)/\sigma_N$, and the exact formulae for Λ_N, σ_N^2 and the terms of $\mathbb{W}_N^{(5)}(u_b, 0)$ are given above.

Corollary 4.2 follows from Corollary 4.1 by using the Euler–Maclaurin summation formula.

We state just a 2-term asymptotic expansion to keep the expressions simple.

Corollary 4.2 *Let $c_1 \le Np_m \le c_2$ for some positive c_1, c_2 and all $m = 1, \dots, N$. Then*

$$\left|P\big\{X_N^2<u\sigma_N+\Lambda_N\big\}-\Phi(u)-\frac{e^{-u^2/2}}{\sqrt{2\pi N}}\left[\frac{1-u^2}{6}\alpha_{30N}+\frac{\lambda\mathrm{P}_{2N}}{\hat{\sigma}_N}\right.\right.$$
$$\left.\left.+\frac{2}{\hat{\sigma}_N}S_1\left(\frac{1}{2}(u\sigma_N+\Lambda_N)\right)\right]\right|\le C\left(\frac{1}{N}+\frac{1}{n\lambda}+n\lambda\exp\left\{-\frac{cN}{\lambda\max(1,\lambda)}\right\}\right), \quad (4.2)$$

where $S_1(x) = x - [x] + 1/2$ is a well-known periodic function of period one (see for instance [4, p. 254]), *and comes up here due to the Euler–Maclaurin summation formula.*

We may remark here that Corollary 4.2 is already a considerable improvement over Theorem 5 of [14], which states inequality (4.2) with $\exp\{-N\lambda^l e^{-2\lambda}\}$, $l > 0$, instead of the exponential term, and makes sense under the additional restriction $\lambda = O(\ln N)$.

Remark 4.1 Application of Theorems 3.3 and 3.4 to the log-likelihood statistic $L_N = \sum_{m=1}^N \eta_m \ln \eta_m$, and to the count-statistics $\mu_r = \sum_{m=1}^N \mathbf{I}\{\eta_m = r\}$, gives results similar to Theorems 4 and 6, respectively, of [24], but our results can be used to obtain additional terms in the expansions they provide.

A DS with kernel functions $f_{m,N} = f_N$ for all $m = 1, 2, \dots, N$ is called a "symmetric DS"; for example, the X_N^2, μ_r and L_N are all symmetric DS. It is well-known (see e.g. [8, 23, 34]) that the chi-square test is asymptotically most powerful (AMP) within the class of symmetric tests, i.e. among tests based on symmetric DS, for testing the hypothesis of uniformity against the sequence of alternatives H_{1n} given by

$$p_m = \frac{1}{N}\left(1 + \frac{\vartheta_m}{(n\lambda)^{1/4}}\right), \quad m = 1, 2, \ldots, N;$$

$$\sum_{m=1}^{N} \vartheta_m = 0 \quad \text{and} \quad 0 < C_1 \le \frac{1}{N}\sum_{m=1}^{N} \vartheta_m^2 \le C_2 < \infty.$$

Moreover, the chi-square test is the unique AMP test for λ bounded away from zero and infinity; on the other hand, if $\lambda \to 0$ or $\lambda \to \infty$ then there exist other AMP symmetric tests, for example, the empty cells test when $\lambda \to 0$, and the log-likelihood test when $\lambda \to \infty$. In view of this, Ivchenko and Mirakhmedov [14] introduced and studied the "second order asymptotic efficiency" (SOAE) of symmetric tests wrt the chi-square test. Investigation of the SOAE is based on the asymptotic expansion of the power function of such tests. In the case $\lambda \to 0$ they have shown that SOAE may arise only if $n = O(N^{3/4})$; for example, the empty-cells test based on the statistic μ_0 is SOAE for this situation; for the case $\lambda \to \infty$ they could only note that when $\lambda = O(\ln N)$, the SOAE test does not exist, because of the restrictive choice of $\lambda = O(\ln N)$ needed in their asymptotic expansions. Therefore, they point out that the SOAE problem is open when $\lambda \to \infty$. Corollary 4.2 does resolve this problem showing that the chi-square test is still optimal in the sense of SOAE if $n = o(N^{3/2})$; it is also SOAE wrt the log-likelihood test for $n \ge N^{3/2}$ (cf. [24], for further discussion).

4.2 Example B

Now we consider the sample scheme without replacement from a stratified population of size Ω_N; the strata are indexed by $m = 1, \ldots, N$; ω_m is the size of the mth stratum, with $\Omega_N = \omega_1 + \cdots + \omega_N$; and η_m is the number of elements of the mth stratum appearing in a sample of size n. In this scheme, $\mathcal{L}(\xi_m) = \mathrm{Bi}(\omega_m, \upsilon)$, where $\upsilon \in (0, 1)$ is arbitrary. We choose $\upsilon = p =: n/\Omega_N$, $q = 1 - p$, so that $x_N = 0$. Set $\bar{\omega}_N = \max_{1\le m\le N} \omega_m$, $\Omega_{2,N} = \omega_1^2 + \cdots + \omega_N^2$. We consider the case where the strata sizes ω_m may increase together with N but satisfy the following condition

$$\bar{\omega}_N = o\big((nq)^{1/4}\big). \tag{4.3}$$

Theorem 4.2 *If the Lindeberg's condition* (3.1) *is satisfied along with condition* (4.3), *then as* $nq \to \infty$, $R_N(\eta)$ *has the asymptotic normal distribution with expectation* Λ_N *and variance* σ_N^2 *as given in* (2.1).

Theorem 4.3 *For arbitrary* $\delta \in (0, 1]$ *there exists a constant* C *such that*

$$\Delta_N^{(3)} \le C\left(\beta_{2+\delta,N} + \left(\frac{\bar{\omega}_N}{nq}\right)^{\delta/2} + \frac{\bar{\omega}_N^2}{\sqrt{nq}}\right).$$

Remark 4.2 The term $\bar{\omega}_N^2/\sqrt{nq}$ can be replaced by $\bar{\omega}_N \max(1-6pq+3nq\Omega_{2,N}\Omega_N^{-2})/\sqrt{nq}$. If $\bar{\omega}_N \le (nq)^{(1-\delta)/(4-\delta)}$ then the second term on the rhs dominates the third one.

Theorem 4.4 *Let* $E|g_m(\xi_m)|^5 < \infty$; *then there exist constants* c *and* C *such that*

$$\Delta_N^{(5)} \le C\biggl(\beta_{5,N} + \biggl(\frac{\bar{\omega}_N}{nq}\biggr)^{\frac{3}{2}} + \chi_N\bigl(cT_N, \beta_{5,N}^{-1}\bigr)\biggr).$$

Let now the elements of mth stratum be independent r.v.s $X_{m,1}, \dots, X_{m,\omega_m}$, $m = 1, \dots, N$. We draw a sample of size n without replacement from the entire population. Define the indicator r.v.s η_{mi} which are equal to one if an element X_{mi} of the mth stratum appears in the sample, or else it equals to zero, so that $\eta_{m1} + \cdots + \eta_{m\omega_m} = \eta_m$. Then $S_{n,N}^{(m)} = \sum_{i=1}^{\omega_m} X_{mi}\eta_{mi}$ represents the sum of elements of the mth stratum which appear in the sample, and the sum of all the elements in the sample, the "sample-sum," given by $S_{n,N} = \sum_{m=1}^{N} S_{n,N}^{(m)}$, is a DS.

Assume that the r.v.s $X_{m,1}, \dots, X_{m,\omega_m}$ have a common distribution, the same as that of an r.v. Y_m, $m = 1, \dots, N$. We also assume that $Y_1, \dots, Y_N$ are independent r.v.s. Then the r.v. $S_{n,N}$ is distributionally equal to a DS with $f_{mN}(0) = 0$, $f_{mN}(j) = X_{m,1} + \cdots + X_{m,j}$, $m = 1, \dots, N$:

$$\mathcal{L}(S_{n,N}) = \mathcal{L}\Biggl(\sum_{m=1}^{N}\Biggl(\sum_{j=1}^{\eta_m} X_{m,j}\mathbf{I}\{\eta_m \ge 1\}\Biggr)\Biggr). \tag{4.4}$$

Suppose $E|Y_m|^s < \infty$ for some $s \ge 3$. Then the expressions in (2.1) have the following form:

$$f_{mN}(\xi_m) = \sum_{j=1}^{\xi_m} X_{m,j}\mathbf{I}\{\xi_m \ge 1\}, \tag{4.5}$$

$$g_m(\xi_m) = \sum_{j=1}^{\xi_m} \mathbf{I}\{\xi_m \ge 1\}(X_{m,j} - \gamma_N) - \omega_m p(EY_m - \gamma_N), \quad m = 1, \dots, N.$$

$$\gamma_N = \frac{1}{\Omega_N}\sum_{m=1}^{N} \omega_m EY_m, \tag{4.6}$$

$$\sigma_N^2 = p\sum_{m=1}^{N} \omega_m\bigl(E(Y_{mN} - \gamma_N)^2 - p\bigl(E(Y_{mN} - \gamma_N)\bigr)^2\bigr).$$

From Theorems 4.3 and 4.4, we immediately have the following corollary.

Corollary 4.3 *If* (4.3) *is satisfied, then for arbitrary* $\delta \in (0, 1]$ *there exists a constant* C *such that*

$$\sup_{-\infty \le u \le \infty}\bigl|P\{S_{n,N} < u\sigma_N + n\gamma_N\} - \Phi(u)\bigr| \le C\biggl(\beta_{2+\delta,N} + \frac{1}{(nq)^{\delta/2}}\biggr),$$

where

$$\begin{aligned}\beta_{2+\delta,N} &= \sigma_N^{-(2+\delta)} \sum_{m=1}^{N} E\big|g_m(\xi_m)\big|^{2+\delta}\\ &\le \frac{2^{2+\delta} p(1+p^{1+\delta})}{\sigma_N^{2+\delta}} \sum_{m=1}^{N} \omega_m^{2+\delta} E|Y_{m,N}-\gamma_N|^{2+\delta}. \qquad (4.7)\end{aligned}$$

Corollary 4.4 *If* (4.3) *is satisfied, then there exist positive constants c and C such that*

$$\Delta_N^{(5)} \le C\Big(\beta_{5,N} + \Big(\frac{\bar{\omega}_N}{nq}\Big)^{\frac{3}{2}} + \chi_N\big(c\tilde{T}_N, \beta_{5,N}^{-1}\big)\Big),$$

where $\tilde{T}_N = \min(\beta_{3N}^{-1}, \sqrt{nq}/\bar{\omega}^2)$, *and the terms of the* $\mathbb{W}_N^{(5)}(u,0)$ *in* (4.1) *have the following forms*:

$$\begin{aligned}
\alpha_{1,2,N} &= 0, \qquad \alpha_{2,1,N} = \sqrt{\frac{q}{n}}\frac{\sum_{m=1}^{N}\omega_m(\alpha_{2,m}-2p\alpha_{1,m}^2)}{\sum_{m=1}^{N}\omega_m(\alpha_{2,m}-p\alpha_{1,m}^2)},\\
\alpha_{0,3,N} &= \frac{1-2q}{\sqrt{nq}},\\
\alpha_{2,2,N} &= \frac{\sum_{m=1}^{N}\omega_m(\alpha_{2,m}(1+(\omega_m-2)p)-\alpha_{1,m}^2(\omega_m-2)p(1-3q))}{\Omega_N p\sum_{m=1}^{N}\omega_m(\alpha_{2,m}-p\alpha_{1,m}^2)},\\
\sum_{m=1}^{N}\alpha_{11m}^2 &= \frac{q\sum_{m=1}^{N}\omega_m^2\alpha_{1,m}^2}{\Omega_N\sum_{m=1}^{N}\omega_m(\alpha_{2,m}-p\alpha_{1,m}^2)},\\
\sum_{m=1}^{N}\alpha_{20m}\alpha_{02m} &= \frac{\sum_{m=1}^{N}\omega_m^2(\alpha_{2,m}-p\alpha_{1,m}^2)}{\Omega_N\sum_{m=1}^{N}\omega_m(\alpha_{2,m}-p\alpha_{1,m}^2)},\\
\alpha_{3,0,N} &= \sum_{m=1}^{N}\omega_m\big(\alpha_{3,m}-3p\alpha_{1,m}\alpha_{2,m}-2p^2\alpha_{1,m}^3\big)\\
&\quad\times\Big(p^{2/3}\sum_{m=1}^{N}\omega_m\big(\alpha_{2,m}-p\alpha_{1,m}^2\big)\Big)^{-3/2},\\
\alpha_{4,0,N} &= \sum_{m=1}^{N}\omega_m\big(\alpha_{4,m}-4p\alpha_{1,m}\alpha_{3,m}\\
&\quad+3(\omega_m-1)p\alpha_{2,m}^2-6(\omega_m-2)p^2\alpha_{1,m}^2\alpha_{2,m}\\
&\quad-3(3\omega_m-2)p^3\alpha_{1,m}^4\big)\Big(\sqrt{p}\sum_{m=1}^{N}\omega_m\big(\alpha_{2,m}-p\alpha_{1,m}^2\big)\Big)^{-2},
\end{aligned}$$

where $\alpha_{i,m} = E(Y_m - \gamma_N)^i$; *also*

$$\chi_N\big(c\tilde{T}_N, \beta_{5,N}^{-1}\big) \le C\sqrt{n}\ln\big(\sigma_N^{-1}\beta_{5,N}^{-1}\big) \times \exp\left\{-n\left(1 - \sup_{(c\sigma_N \tilde{T}_N)^{-1} \le |t| \le (\sigma_N \beta_{5,N})^{-1}} \frac{1}{N}\sum_{m=1}^{N}\big|Ee^{itY_m}\big|\right)\right\}, \tag{4.8}$$

and

$$\chi_N\big(c\tilde{T}_N, \beta_{5,N}^{-1}\big) \le C\sqrt{n}\ln\big(\sigma_N^{-1}\beta_{5,N}^{-1}\big) \times \exp\left\{-2nq\left(1 - \sup_{(c\sigma_N \tilde{T}_N)^{-1} \le |t| \le (\sigma_N \beta_{5,N})^{-1}} \frac{1}{\Omega_N}\left|\sum_{m=1}^{N}\omega_m Ee^{itY_m}\right|\right)\right\}. \tag{4.9}$$

When we take $\omega_1 = \cdots = \omega_N = 1$, our Corollary 4.3 improves a result of [22], and a recent result of [38] for the case when $(nq)^{-1/2} \le \Delta_1$ (in their notation). Note that

$$\beta_{3N} = \frac{p}{\sigma_N^3}\sum_{m=1}^{N} E|Y_m - pEY_m - q\gamma_N|^3 + \frac{qp^3}{\sigma_N^3}\sum_{m=1}^{N}\big|E(Y_m - \gamma_N)\big|^3,$$

which provides a natural expression, showing the exact dependence of the bound on $p = n/N$, and moments of the elements of population, instead of the formula for Δ_2^* in [38]. This fact is confirmed by the second term in $\mathbb{W}_N^{(5)}(u,0)$ (see (4.1)), and that $\alpha_{1,2,N} = 0$. Also, in this case the 3-term asymptotic expansion, i.e. $\mathbb{W}_N^{(5)}(u,0)$, coincides with that given by Mirakhmedov [21]; further, from our Corollary 4.4 follows the main result of [5, 36], and it extends Theorem 1 of [11] giving an additional term in their asymptotic expansion for the case when p is bounded away from one; this case is the most interesting in a sample scheme without replacement.

4.3 Example C

For this case we assume that $\mathcal{L}(\xi_m) = \mathrm{NB}(d_m, p)$, with $p = n/(n + D_{1N})$, $m = 1, \ldots, N$, where $D_{jN} = d_1^j + \cdots + d_N^j$; then $x_N = 0$. Putting $\rho = p/(1-p) = n/D_{1N}$ we get $B_N^2 = D_{1N}\rho(1+\rho)$, $\kappa_{4N} = (1 + 3\rho(1+\rho)(2 + D_{2N}D_{1N}^{-1}))(D_{1N}\rho(1+\rho))^{-1}$.

Theorem 4.5 *Let* $D_{2N}D_{1N}^{-2} = o(N^{-1/2})$. *If the Lindeberg's condition* (3.1) *is satisfied, then the DS* $R_N(\eta)$ *has an asymptotic normal distribution with mean* Λ_N *and variance* σ_N^2 *as given in* (2.1).

Theorem 4.6 *There exists a constant C such that* $\Delta_N^{(3)} \le C(\beta_{3N} + E_N)$, *where*

$$E_N = \frac{1}{\sqrt{n(1+\rho)}} + \sqrt{\frac{3}{D_{1N}}\left(2 + \frac{D_{2N}}{D_{1N}}\right)}\left(1 + \sqrt{\frac{3N}{D_{1N}}\left(2 + \frac{D_{2N}}{D_{1N}}\right)}\right).$$

Remark 4.3 Using the fact that $D_{1N}^2 \le N D_{2N}$ we have

$$E_N \le \frac{1}{\sqrt{n(1+\rho)}} + \frac{\sqrt{3}}{\sqrt{N}}\sqrt{1+\frac{2N}{D_{1N}}}\left(1+\sqrt{1+\frac{2N}{D_{1N}}}\right).$$

Theorems 4.5 and 4.6 do improve as well as correct Theorems 13 and 14 of [26].

Theorem 4.7 *Let the statistic $R_N(\eta)$be a lattice r.v. with span h and a set of possible values $\Re$. If $E|g(\xi_m)|^5 < \infty$, $m = 1, 2, \ldots, N$, then uniformly in $z \in \Re$*

$$\begin{aligned}&\left|\frac{\sigma_N}{h} P\{R_N(\eta) = z\} - \frac{d}{du_z}\mathbb{W}_N^{(5)}(u_z, 0)\right|\\ &\quad\le C\left(\beta_{5,N} + \left(\frac{D_{3N}}{D_{1N}^3} + \frac{D_{2N}}{D_{1N}^2 n(1+\rho)}\right)^{3/4}\right.\\ &\qquad\left.+\,\sigma_N\sqrt{D_{1N}\rho(1+\rho)}\exp\left\{-\frac{1}{2}N\bar{H}_N\left(cT_N, \frac{\pi\sigma_N}{h}\right)\right\}\right),\end{aligned}$$

where $u_z = (z - \Lambda_N)/\sigma_N$ and T_N are defined as in Sect. 3 with $\mathcal{E}_N(1) = E_N$.

Now consider the following practical and important two-sample problem: Let $X_1, \ldots, X_{M-1}$ and $Y_1, \ldots, Y_n$ be two samples from continuous distributions F and G, respectively, defined on the same $A \subset R$. The classical two-sample problem is to test the null hypothesis of homogeneity $H_0 : F = G$. Define the r.v.s

$$\eta_{m,k} = \sum_{i=1}^{n} \mathbf{I}\{Y_i \in [X_{(m\cdot k)}, X_{(m\cdot k-k)}]\},$$

where $m = 1, \ldots, N$, $N = \lfloor M/k \rfloor$ is the largest integer that does not exceed M/k, integer $k \ge 1$, $X_{(1)}, \ldots, X_{(M-1)}$ are the order-statistics of the first sample $X_1, \ldots, X_{M-1}$. The r.v. $(\eta_{1,k}, \ldots, \eta_{N,k})$ are called the "spacing-frequencies," i.e. frequencies of the second sample falling in-between the spacings created by the first sample. A wide class of test statistics for testing H_0 can be expressed in the form (see [10] and [7])

$$V_N = \sum_{m=1}^{N} f_{m,N}(\eta_{m,k})$$

where the $f_{m,N}$ are real valued functions. It is easy to check that under H_0 the r.v. $(\eta_{1,k}, \ldots, \eta_{N,k})$ satisfies (1.1) with $\mathcal{L}(\xi_m) = \mathrm{NB}(k, p)$, $p = n/(n+M)$, i.e. the statistic V_N is DS defined in the Pólya–Egenberger urn model. Hence, Theorems 4.5 and 4.6 immediately lead to the following Corollaries 4.5 and 4.6, by putting $d_m = k$, $m = 1, \ldots, N$ and $\rho = n/M$.

Corollary 4.5 *If the Lindeberg's condition (3.1) is satisfied then the statistic V_N has asymptotic normal distribution with expectation Λ_N and variance σ_N^2.*

Corollary 4.6 *There exists a constant $C > 0$ such that*

$$\Delta_N^{(3)} \le C\left(\beta_{3N} + \frac{1}{\sqrt{n(1+\rho)}} + \frac{1}{\sqrt{N}}\right).$$

Consider now the so-called Dixon statistic defined by $\mathcal{D}_N = \sum_{m=1}^N \eta_{m,k}^2$. For this statistic we obtain:

$$\Lambda_N = M\big(1 + (1+k)\rho\big), \qquad \gamma_N = 1 + 2(1+k)\rho,$$

$$\sigma_N^2 = 2M(1+2k)\rho^2(1+\rho)^2,$$

$$g_m(\xi_m) = (\xi_m - k\rho)^2 - k\rho(1+\rho) - (1+2\rho)(\xi_m - k\rho),$$

$$\alpha_{12N} = \frac{\sqrt{2}(k+1)}{\sqrt{k(1+2k)}},$$

$$\alpha_{30N} = \frac{8k^2\rho^3(1+\rho)^3 + k(1+\rho)^2(19 + 76\rho(1+\rho)) + 2(1+\rho)^2(15 - 13\rho + 16\rho^2(1+\rho))}{2\sqrt{2k}(1+2k)^{3/2}\rho^3(1+\rho)^3}.$$

Although the exact formula for $\beta_{4,N} = \alpha_{40N}$ is manageable, it is quite long and therefore we restrict ourselves to its leading term as $k \to \infty$, and obtain the following bounds:

$$\beta_{4N} \le CN^{-1}\max\big(1, \big(k\rho(1+\rho)\big)^{-4}\big), \qquad (\sigma_N\beta_{3N})^{-1} \ge c\big(k\rho(1+\rho)\big)^{-1/2}.$$

The Dixon statistic satisfies the conditions of Theorem 4.7. In particular by evaluating the moments of the r.v. $g_m(\xi_m) = (\xi_m - k\rho)^2 - k\rho(1+\rho) - (1+2\rho)(\xi_m - k\rho)$ and using Corollaries 4.5, 4.6 and Theorem 4.7 along with Remark 3.2, we obtain the following result (we omit the details).

Corollary 4.7

(i) *If $\sqrt{N}k^2\rho^2(1+\rho)^2 \to \infty$, then the Dixon statistic has an asymptotic normal distribution with mean $M(1+(1+k)\rho)$ and variance $2M(1+2k)\rho^2(1+\rho)^2$.*

(ii) $P\big\{\mathcal{D}_N < u\sqrt{2M(1+2k)}\rho(1+\rho) + M(1+(1+k)\rho)\big\}$
$$= \Phi(u) + O\left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{N}k^2\rho^2(1+\rho)^2}\right).$$

(iii) *Let $k \to \infty$ and $k = o(M^{1/3})$; then we have, for any $b = 0, 1, \ldots, n(n-1)/2$,*

$$M(1+2k)\rho^2(1+\rho)^2 P\{\mathcal{D}_N = n + 2b\} = \frac{e^{-u_b^2/2}}{\sqrt{2\pi}}\left(1 + \frac{u_b^2 - 3u_b}{6\sqrt{N}}\alpha_{30N} + \frac{(k+1)u}{\sqrt{2Nk(1+2k)}}\right) + O\big(N^{-1}\big),$$

where $u_b = (n + 2b - M(1+(1+k)\rho))/\sqrt{2M(1+2k)}\rho(1+\rho)$.

Consider the class of symmetric tests (i.e. based on symmetric DS) for testing the hypothesis of homogeneity against some of "smooth" sequence of alternatives which approaches the null at the rate $O((nk)^{-1/4})$. The asymptotic power of symmetric tests increases as k grows; the Dixon statistic is an example of a symmetric DS; it is known to be unique AMP within the class of symmetric tests for any fixed k, the step of spacings; see [35]. The above stated Corollaries 4.4–4.6 allow us to consider the situation when $k \to \infty$; in this case the AMP test is not unique. Comparison of the AMP tests based on their second-order asymptotic efficiencies using the asymptotic expansion of the power function can be done. For this purpose, the asymptotic expansion results presented here are central and such comparisons will be the subject of another investigation.

Acknowledgements We would like to express our gratitude to the referee for a careful reading of the manuscript and for comments that helped improving the paper.

Appendix: Proofs

Proof of Proposition 2.1 We need the following three lemmas to complete the proof of this proposition.

Lemma A.1 *Set* $\ell_{s,N} = \min(\beta_{s,N}^{-1/s}, \kappa_{s,N}^{-1/s})$. *There exist constants* $c > 0$ *and* $C > 0$ *such that if* $\max(|t|, |\tau|) \le c\ell_{s,N}$ *then for* $k = 0$ *and* 1,

$$\left|\frac{\partial^k}{\partial t^k}\left(\Psi_N(t,\tau) - e^{-\frac{t^2+\tau^2}{2}}\left(1 + \sum_{v=1}^{s-3} N^{-v/2} P_v(t,\tau)\right)\right)\right|$$
$$\le C(\beta_{s,N} + \kappa_{s,N})\big(1 + |t|^s + |\tau|^s\big)e^{-\frac{t^2+\tau^2}{4}}.$$

Lemma A.1 follows from Theorem 9.11 of [4] because of (2.2) and the fact that the sum of the r.v.s $(\hat{g}_m, \hat{\xi}_m)$ has unit correlation matrix. □

Lemma A.2 *For any integer* l *satisfying* $0 \le l \le 3v$, *where* $v = 0, 1, \ldots, s-2$, *there exists a constant* $c(l, v) > 0$ *such that*

$$\left|\frac{\partial^l}{\partial t^l} G_{v,N}(t, x_N)\right| \le c(l,v)\big(1 + \big(|t| + |x_N|\big)^{3v-l}\big)(\beta_{v+2,N} + \kappa_{v+2,N}).$$

Proof Similarly to that of Lemma 9.5 of [4, p. 71], the only difference being that in (9.12) of [4, p. 72], we use the inequality $\rho_{j_i+2}/\rho_2^{(j_i+2)/2} \le (\rho_{r+2}/\rho_2^{(r+2)/2})^{j_i/r}$, in their notation, to obtain

$$\begin{aligned}\left|\frac{\partial^l}{\partial t^l} P_{v,N}(t,\tau)\right| &\le c(l,v)\big(1 + \big(|t| + |\tau|\big)^{3v-l}\big)(\beta_{v+2,N} + \kappa_{v+2,N})\\ &\le c(l,v)\big(1 + \big(|t| + |\tau|\big)^{3v-l}\big)\big(\beta_{s,N}^{v/(s-2)} + \kappa_{s,N}^{v/(s-2)}\big). \qquad (5.1)\end{aligned}$$

Lemma A.2 follows from this and (2.6). □

Lemma A.3 *Let* $\max(\beta_{3N},\kappa_{3N})\le 0.01$. *If* $|t|\le 0.3\beta_{3N}^{-1}$ *and* $|\tau|\le 0.3\kappa_{3N}^{-1}$, *then for* $k=0$ *and* 1,

$$\left|\frac{\partial^k}{\partial t^k}\Psi_N(t,\tau)\right|\le \exp\left\{-\frac{t^2+\tau^2}{10}\right\}.$$

Lemma A.3 follows from Lemma A, Part (2) of [27].

Put $T_N(s)=\min(\beta_{s,N}^{-1/s},\kappa_{s,N}^{-1/s},\mathcal{E}_N^{-1}(1))$, where $s\ge 3$. Note that $T_N(s)\le T_N$, since $\beta_{s,N}^{-1/s}\le\beta_{3,N}^{-1}$. Let $|t|\le c_1T_N(s)$, where $c_1>0$ is to be chosen sufficiently small. From (2.4) and (2.6),

$$\begin{aligned}\nabla_N(t)=:&\left|\frac{\partial^k}{\partial t^k}\left(\Theta_N(t,x_N)-e^{-\frac{t^2+x_N^2}{2}}\left(1+\sum_{v=1}^{s-3}N^{-v/2}G_{v,N}(t,x_N)\right)\right)\right|\\ &\le\int_{|\tau|\le c_1\ell_{s,N}}\left|\frac{\partial^k}{\partial t^k}\left(\Psi_N(t,\tau)-e^{-\frac{t^2+\tau^2}{2}}\sum_{v=1}^{s-3}N^{-v/2}P_{v,N}(t,\tau)\right)\right|d\tau\\ &\quad+\int_{c_1\ell_{s,N}\le|\tau|}\left|\frac{\partial^k}{\partial t^k}\left(e^{-\frac{t^2+\tau^2}{2}}\sum_{v=1}^{s-3}N^{-v/2}P_{v,N}(t,\tau)\right)\right|d\tau\\ &\quad+\int_{c_1\ell_{s,N}\le|\tau|\le 0.3\kappa_{3,N}^{-1}}\left|\frac{\partial^k}{\partial t^k}\Psi_N(t,\tau)\right|d\tau\\ &\quad+\int_{0.3\kappa_{3,N}^{-1}\le|\tau|\le\pi B_N}\left|\frac{\partial^k}{\partial t^k}\Psi_N(t,\tau)\right|d\tau=\Im_1+\Im_2+\Im_3+\Im_4.\end{aligned}\tag{5.2}$$

Applying Lemma A.1, (5.1), and Lemma A.3 to $\Im_1$, $\Im_2$ and $\Im_3$, respectively, after some algebraic manipulations we obtain

$$\Im_l\le C(\beta_{s,N}+\kappa_{s,N})\big(1+|t|^s\big)e^{-\frac{t^2}{6}},\quad l=1,2,3.\tag{5.3}$$

Set $\hat\psi_m(t,\tau)=E\exp\{it\hat g_m(\xi_m)+i\tau\hat\xi_m\}$ and recall that $\Psi_N(t,\tau)=\prod_{m=1}^N\hat\psi_m(t,\tau)$. We have

$$\begin{aligned}\big|\hat\psi_m(t,\tau)\big|^2&=\big|\hat\psi_m(0,\tau)\big|^2+E\big[\big(e^{it\hat g_m^*}-1\big)\big(e^{it\hat\xi_m^*}-1\big)\big]+E\big(e^{it\hat g_m^*}-1\big)\\ &\le\big|\hat\psi_m(0,\tau)\big|^2+|t||\tau|E\big|\hat g_m^*\hat\xi_m^*\big|+t^2E\hat g_m^{*^2}\end{aligned}$$

and

$$\big|\hat\psi_m(t,\tau)\big|^2=\big|\hat\psi_m(0,\tau)\big|^2+Ee^{it\hat\xi_m^*}\big(e^{it\hat g_m^*}-1\big)\le\big|\hat\psi_m(0,\tau)\big|^2+2|t|E|\hat g_m|.$$

Using these inequalities, and the fact that $x < \exp\{(x^2-1)/2\}$, and

$$\left|\frac{\partial}{\partial t}\prod_{m=1}^{N}\hat{\psi}_m(t,\tau)\right| \le \sum_{m=1}^{N}\left|\frac{\partial}{\partial t}\hat{\psi}_m(t,\tau)\right|\prod_{l\neq m}\left|\hat{\psi}_l(t,\tau)\right|,$$

$$\sum_{m=1}^{N}\left|\frac{\partial}{\partial t}\hat{\psi}_m(t,\tau)\right| \le |t|+2|\tau|,$$

we find for $k=0,1$ that

$$\left|\frac{\partial^k}{\partial t^k}\prod_{m=1}^{N}\hat{\psi}_m(t,\tau)\right| \le \sqrt{e}\big(|t|+|\tau|\big)^k \exp\Bigg\{\min\big(|t||\tau|+2t^2, 2|t|\beta_{1,N}\big) - \frac{1}{2}\sum_{m=1}^{N}\big(1-\left|\hat{\psi}_m(0,\tau)\right|^2\big)\Bigg\}. \tag{5.4}$$

Choosing c_1 to be sufficiently small, using (5.4) and that $\beta_{1,N}\le\sqrt{N}$, we get for $|t|\le c_1 T_N(s)$,

$$\begin{aligned}\Im_4 &\le C B_N^{k+1}\exp\left\{-\frac{1}{4}M_N\big(0.3(B_N\kappa_{3,N})^{-1}\big)+\min\big(B_N|t|+2t^2, 2|t|\sqrt{N}\big)\right\}\\ &\le C B_N^{k+1}\exp\left\{-\frac{1}{8}M_N\big(0.3(B_N\kappa_{3,N})^{-1}\big)-\frac{t^2}{8}\right\}.\end{aligned} \tag{5.5}$$

From (5.2), (5.3), and (5.4) it follows that

$$\left|\frac{\partial^k}{\partial t^k}\left(\Theta_N(t,x_N)-e^{-\frac{t^2+x_N^2}{2}}\left(1+\sum_{v=1}^{s-3}N^{-v/2}G_{v,N}(t,x_N)\right)\right)\right| \le Ce^{-\frac{t^2}{8}}\Upsilon_{s,N}. \tag{5.6}$$

In particular (5.6) implies

$$\left|\Theta_N(0,x_N)-e^{-\frac{x_N^2}{2}}\sum_{v=0}^{s-3}N^{-v/2}G_{v,N}(0,x_N)\right| \le C\Upsilon_{s,N}. \tag{5.7}$$

Put

$$\mathcal{G}_N^{(s)}(t,x_N)=e^{-\frac{t^2+x_N^2}{2}}\sum_{v=0}^{s-3}N^{-v/2}G_{v,N}(t,x_N).$$

Then, from (2.5), (5.6) and (5.7) we have

$$\begin{aligned}&\frac{\partial^k}{\partial t^k}\left(\varphi_N(t,x_N)-W_N^{(s)}(t,x_N)\right)\\&\quad=\frac{\partial^k}{\partial t^k}\left(\frac{\mathcal{G}_N^{(s)}(t,x_N)}{\mathcal{G}_N^{(s)}(0,x_N)}-W_N^{(s)}(t,x_N)\right)\\&\qquad+\frac{\theta_2\Upsilon_{s,N}}{\mathcal{G}_N^{(s)}(0,x_N)+\theta_2\Upsilon_{s,N}}\left(\theta_1 t^k e^{-\frac{t^2}{8}}+\frac{1}{\mathcal{G}_N^{(s)}(0,x_N)}\frac{\partial^k}{\partial t^k}\mathcal{G}_N^{(s)}(t,x_N)\right),\end{aligned}\tag{5.8}$$

where, as usual, $|\theta_i|\le 1$. Note that the polynomials $Q_{j,N}(x)$ in (2.7) are actually the result of the expansion of $(\exp\{x_N^2/2\}\mathcal{G}_N^{(s)}(0,x_N))^{-1}$ noting that $G_{0,N}(x)=1$; also, it is clear that $\mathcal{G}_N^{(s)}(0,x_N)\ge c$ for some $c>0$. Using these facts and (2.7), (2.8), Lemma A.2 in (5.8) after some algebra, we complete the proof of Proposition 2.1 for $|t|\le c_1T_N(s)$.

Let now $c_1T_N(s)\le|t|\le c_1T_N$. Then, using Lemma A.2 it is easy to see that

$$\nabla_N(t)\le\left|\frac{\partial^k}{\partial t^k}\left(\Theta_N(t,x_N)-e^{-\frac{t^2+x_N^2}{2}}\right)\right|+Ce^{-\frac{t^2}{8}}\Upsilon_{s,N}.$$

Apply the outlined above technique for the first term in the rhs of this inequality using (2.4), Lemma A.1 with $s=3$, Lemma A.3 and the fact that $|t|\ge c_1T_N(s)$, to complete the proof of Proposition 2.1; the details are omitted.

Proof of Theorem 3.1 In addition to the notations of Sect. 3, define

$$L_{1,N}(\varepsilon)=\frac{1}{N^{3/2}}\sum_{m=1}^{N}E|\hat{g}_m|^3\mathbf{I}\left\{|\hat{g}_m|\le\varepsilon\right\}.$$

Since $|x_N|\le c_0$, Theorem 1 of [26] gives: for arbitrary $\varepsilon>0$ there exists a constant $C>0$ such that

$$\begin{aligned}\Delta_N^{(3)}\le C\Bigg(&L_{1,N}(\varepsilon)+L_{2,N}(\varepsilon)+\mathcal{L}_{1,N}(\varepsilon)+\mathcal{L}_{2,N}(\varepsilon)+B_N^2\mathcal{L}_{1,N}(\varepsilon)\\&\times\exp\left\{-\frac{1}{8}M_N\left(\pi\left(4B_N\mathcal{L}_{1,N}(\varepsilon)\right)^{-1}\right)\right\}\\&+\frac{\max(\sqrt{M_N(\pi(4B_N\mathcal{L}_{1,N}(\varepsilon))^{-1})},\min(B_N,\sqrt{N}))}{M_N(\pi(4B_N\mathcal{L}_{1,N}(\varepsilon))^{-1})}\Bigg).\end{aligned}$$

Since $L_{1,N}(\varepsilon)\le\varepsilon$, $\mathcal{L}_{1,N}(\varepsilon)\le\varepsilon$ and $\varepsilon>0$ is arbitrarily small, Theorem 3.1 follows. □

Proof of Theorem 3.2 Putting $X_{mN} = f_{m,N}(\xi_m)$ and $Y_{mN} = \xi_m$ in Theorem 2 of [25], it can be shown that

$$\begin{aligned}\Delta_N^{(3)} \le C\biggl(&\beta_{2+\delta,N} + \kappa_{2+\delta,N} + B_N^2 \kappa_{2+\delta,N}^{1/\delta}\\ &\times \exp\biggl\{-\frac{1}{8} M_N\bigl(\pi\bigl(4B_N\kappa_{2+\delta,N}^{1/\delta}\bigr)^{-1}\bigr)\biggr\} + \mathcal{E}_N(\delta)\biggr),\end{aligned}$$

since $\mathrm{P}_N(u) = P(g_1(\xi_1) + \cdots + g_N(\xi_N) < u\sigma_N | \zeta_N = n)$. If $M_N(\pi(4B_N\kappa_{2+\delta}^{1/\delta})^{-1}) \le cB_N$ for some $c > 0$ then Theorem 3.2 is true with $C = c$. If $M_N(\pi(4B_N\kappa_{2+\delta,N}^{1/\delta})^{-1}) > cB_N$ then $B_N^2\kappa_{2+\delta,N}^{1/\delta} \exp\{-M_N(\pi(4B_N\kappa_{2+\delta,N}^{1/\delta})^{-1})/8\} \le c_1\kappa_{2+\delta,N}^{1/\delta} \le c_1\kappa_{2+\delta,N}$, since $\delta \in (0, 1]$, and Theorem 3.2 follows. □

Proof of Theorem 3.3 By the well-known Esseen's smoothing inequality we have

$$\begin{aligned}\Delta_N^{(s)} &\le \frac{1}{\pi}\int_{|t|\le\beta_{s,N}^{-1}} \left|\frac{\varphi_N(t,x_N) - e^{-\frac{t^2}{2}}W_N^{(s)}(t,x_N)}{t}\right| dt + \frac{24}{\sqrt{2\pi}}\beta_{s,N}\\ &\le \frac{1}{\pi}\int_{|t|\le cT_N} \left|\frac{\varphi_N(t,x_N) - e^{-\frac{t^2}{2}}W_N^{(s)}(t,x_N)}{t}\right| dt\\ &\quad + \int_{cT_N\le|t|\le\beta_{s,N}^{-1}} \left|\frac{e^{-\frac{t^2}{2}}W_N^{(s)}(t,x_N)}{t}\right| dt\\ &\quad + \int_{cT_N\le|t|\le\beta_{s,N}^{-1}} \left|\frac{\varphi_N(t,x_N)}{t}\right| dt + \frac{24}{\sqrt{2\pi}}\beta_{s,N}.\end{aligned}$$

Also,

$$\left|\varphi_N(t,x_N) - e^{-\frac{t^2}{2}}W_N^{(s)}(t,x_N)\right| \le |t| \max_{|u|\le|t|}\left|\frac{\partial}{\partial u}\Bigl(\varphi_N(u,x_N) - e^{-\frac{u^2}{2}}W_N^{(s)}(u,x_N)\Bigr)\right|.$$

On the other hand, from the definition of $W_N^{(s)}(t,x_N)$, Lemma A.2, and the inequality $\beta_{3,N} \le \beta_{s,N}^{1/(s-2)}$, we observe that

$$\int_{cT_N\le|t|\le\beta_{s,N}^{-1}} \left|\frac{e^{-\frac{t^2}{2}}W_N^{(s)}(t,x_N)}{t}\right| dt \le Ce^{-c_3T_N^2} \le c\Upsilon_{s,N}.$$

Therefore

$$\begin{aligned}\Delta_N^{(s)} \le \frac{1}{\pi}\biggl[&\int_{1\le|t|\le cT_N} \left|\varphi_N(t,x_N) - e^{-\frac{t^2}{2}}W_N^{(s)}(t,x_N)\right| dt\\ &+ \max_{|t|\le1}\left|\frac{\partial}{\partial t}\Bigl(\varphi_N(t,x_N) - e^{-\frac{t^2}{2}}W_N^{(s)}(t,x_N)\Bigr)\right|\biggr] + c\Upsilon_{s,N} + \chi_N\bigl(cT_N, \beta_{s,N}^{-1}\bigr).\end{aligned}$$

Applying Proposition 2.1 here completes the proof of Theorem 3.3. □

Proof of Theorem 3.4 The proof follows by standard methods outlined, as for instance in [32, pp. 204–207], which uses the inversion formula and Propositions 2.1 the details are omitted. □

Proof of Theorem 4.1 To find the central moments of order k of the Poi(λ) r.v. which is a polynomial in λ of order $\lfloor k/2 \rfloor$, for even j, $\kappa_{j,N} \le c(P_{(j-2)/2\,N} + n^{-(j-2)/2})$; for odd j, one can use the well-known inequality $\kappa_{l,N} \le \kappa_{s,N}^{(l-2)/(s-2)}$, $3 \le l \le s$. Similarly, from inequality (53) of [26] we have $M_N(0.3(B_N\kappa_{3N})^{-1}) \ge 0.2n(1+nP_{2N})^{-1}$. Theorem 4.1 follows from these facts and the inequality (3.7). □

Proof of Corollary 4.1 To obtain the set of equalities given before Corollary 4.1, write $g_m(\xi_m) = (\xi_m - \lambda_m)^2 + 2\lambda_m(\xi_m - \lambda_m) - \lambda_m + (2nP_{2N} + 1)(\xi_m - \lambda_m)$; next, to find higher order central moments of Poi(λ) r.v. ξ, we use the following recurrence formula of [16]:

$$E(\xi-\lambda)^{v+1} = v\lambda E(\xi-\lambda)^{v-1} + \lambda\frac{d}{d\lambda}E(\xi-\lambda)^{v}.$$

Considering an r.v. which equals p_m^{l-1} with the probability p_m, $m = 1, \dots, N$, and using well-known inequalities between moments, one can check $\mathrm{P}_{lN}^{l} \le \mathrm{P}_{l+1,N}^{l-1}$, $l = 2, 3, \dots$; with equality iff $p_m = N^{-1}$, $m = 1, \dots, N$. Write $p_m = N^{-1}(1+\varepsilon_m)$, with $\varepsilon_m = Np_m - 1$, and put $\Sigma_N^2 = N^{-1}(\varepsilon_1^2 + \cdots + \varepsilon_N^2)$. It is easy to observe that $\sigma_N^2 = 2n\lambda(1 + c(1+\lambda)\Sigma_N^2)$, also $\alpha_{40N} \le cn^2(1+\lambda^4\Sigma_N^2)/\sigma_N^4$. Considering separately the cases when $\lambda \to 0$, $\lambda \to \infty$, and λ is bounded away from zero and infinity, one can show that $\beta_{4N} = N^{-1}\alpha_{40N} \le c((n\lambda)^{-1} + N^{-1})$, $(\beta_{4N}\sigma_N^2)^{-1} \ge c(1+\lambda^2)^{-1}$ and $T_N \ge c(\sqrt{\lambda}\max(1, \sqrt{\lambda}))^{-1}$; the details are omitted; here $c > 0$ is a constant and it is different in different places. It is evident that the condition (3.9) is fulfilled. Corollary 4.1 follows from Theorem 4.1 and Remark 3.2. □

Proof of Theorems 4.2–4.4 We recall that in this case ξ_m is Bi(ω_m, p) r.v. with $p = n/\Omega_N$. To find the central moments of the r.v. ξ_m we use the following formula: for integer $k \ge 2$,

$$E(\xi_m - \omega_m p)^k = pq\left(\frac{d}{dp}E(\xi_m - \omega_m p)^{k-1} + (k-1)\omega_m E(\xi_m - \omega_m p)^{k-2}\right). \tag{5.9}$$

We have: $B_N^2 = nq$,

$$\kappa_{2+\delta,N} \le \kappa_{4,N}^{\delta/2} \le \left(\frac{1-6pq}{nq} + 3\frac{\Omega_{2,N}}{\Omega_N^2}\right)^{\delta/2}$$
$$\le \left(\frac{1-6pq+3\bar{\omega}_N pq}{nq}\right)^{\delta/2} \le \left(\frac{7\bar{\omega}_N}{4nq}\right)^{\delta/2},$$

$$\kappa_{5,N}\le\kappa_{6,N}^{3/4}<3\sqrt{3}\left(\frac{\bar{\omega}}{n}q\right)^{3/2},\qquad B_N\kappa_{3,N}\le\sqrt{1-6pq+3\bar{\omega}_N pq}\le\sqrt{7\bar{\omega}_N}.$$

Now using the inequalities

$$\begin{aligned}&\big|E\exp\{i\tau\xi_m\}\big|^2\le\exp\big\{-4\omega_m pq\sin^2\tau/2\big\},\\&\sin^2\frac{\tau}{2}\ge\frac{\tau^2}{\pi^2},\quad |\tau|\le\pi,\qquad 1-e^{-u}\ge\frac{1-e^{-c}}{c}u,\quad 0\le u\le c,\end{aligned}\tag{5.10}$$

we get

$$M_N\big(\pi\big(4B_N\kappa_{2+\delta,N}^{1/\delta}\big)^{-1}\big)\ge\frac{(1-e^{-1})nq}{4\bar{\omega}_N(1-6pq+3nq\varOmega_{2,N}\varOmega_N^{-2})}\ge\frac{(1-e^{-1})nq}{4\bar{\omega}_N^2},$$

since $\varOmega_{2,N}=\omega_1^2+\cdots+\omega_N^2$. Finally $\mathcal{L}_{1,N}(\varepsilon)\le\varepsilon^{-1}\kappa_{3,N}\le\varepsilon^{-1}\sqrt{\bar{\omega}_N/nq}$. Theorems 4.2–4.4 follow from Theorems 3.1–3.3, respectively, and the relations given above. □

Proof of Corollaries 4.3 and 4.4 Use inequality $(a_1+\cdots+a_n)^s\le n^{s-1}(a_1^s+\cdots+a_n^s)$, $a_m\ge 0$, $s\ge 1$, to get (4.7). Applying (5.9) we obtain the formulas for α_{ijN}. Recall that $\psi_m(t,\tau)=E\exp\{itf_{mN}(\xi_m)+i\tau\xi_m\}$ and the fact that ξ_m is a sum of $\omega_m$independent Bi$(1,p)$ r.v.s. From (4.5) we have

$$\begin{aligned}\big|\psi_m(t,\tau)\big|&=\left|\sum_{k=0}^{\omega_m}P(\xi_m=k)e^{i\tau k}Ee^{itf_{mN}(k)}\right|\le\sum_{k=0}^{\omega_m}P(\xi_m=k)\big|Ee^{itY_m}\big|^k\\&\le P(\xi_m=0)+\big|Ee^{itY_m}\big|\sum_{k=1}^{\omega_m}P(\xi_m=k)=P(\xi_m=0)\\&\quad+\big|Ee^{itY_m}\big|\big(1-P(\xi_m=0)\big).\end{aligned}$$

Hence

$$\begin{aligned}d_N&\ge\frac{1}{N}\sum_{m=1}^{N}\big(1-(1-p)^{\omega_m}\big)\Big(1-\sup_{(\sigma_N T_N)^{-1}\le|t|\le\sigma_N\beta_{s,N}^{-1}}\big|Ee^{itY_m}\big|\Big)\\&\ge p\left(1-\sup_{(\sigma_N T_N)^{-1}\le|t|\le\sigma_N\beta_{s,N}^{-1}}\frac{1}{N}\sum_{m=1}^{N}\big|Ee^{itY_m}\big|\right),\end{aligned}$$

since $P(\xi_m=0)=(1-p)^{\omega_m}$. Inequality (4.8) follows from this and (3.5). On the other hand

$$\psi_m(t,\tau)=Ee^{(i\tau+\ln Ee^{itY_m})\xi_m}=\big(Ee^{i\tau\varsigma}\big(Ee^{itY_m}\big)^{\varsigma}\big)^{\omega_m}=\big(1+p\big(Ee^{i(\tau+tY_m)}-1\big)\big)^{\omega_m},$$

with $\mathcal{L}(\varsigma) = \mathrm{Bi}(1, p)$. Hence

$$\begin{aligned}\prod_{m=1}^{N} \left|\psi_m(t,\tau)\right|^2 &= \prod_{m=1}^{N} \left|1 + p\left(Ee^{i(\tau+tY_m)} - 1\right)\right|^{2\omega_m} \\ &\le \exp\left\{-2pq \sum_{m=1}^{N} \omega_m \left(1 - E\cos(\tau + tY_m)\right)\right\} \\ &\le \exp\left\{-2\Omega_N pq \left(1 - \frac{1}{\Omega_N}\left|\sum_{m=1}^{N} \omega_m Ee^{i\tau+itY_m}\right|\right)\right\}.\end{aligned}$$

Inequality (4.9) follows. □

Proof of Theorems 4.5 and 4.6 Recall in this case that $\mathcal{L}(\xi_m) = \mathrm{NB}(d_m, p)$, with $p = n/(n + D_{1N})$, $m = 1, \ldots, N$, where $D_{jN} = d_1^j + \cdots + d_N^j$, and $\rho = p/(1-p)$. We use that $Ee^{i\tau\xi_m} = (1-p)^{d_m}(1 - pe^{i\tau})^{-d_m}$ to find the moments of the r.v. ξ_m and that $B_N^2 = D_{1N}\rho(1+\rho)$,

$$\begin{aligned}B_N^2 \kappa_{4N} &= 1 + 3\rho^2(1+\rho)^2\left(2 + D_{2N}D_{1N}^{-1}\right), \\ \left|Ee^{i\tau\xi_m}\right|^2 &= \left(1 + 4\rho(1+\rho)\sin^2\frac{\tau}{2}\right)^{-d_m}.\end{aligned}$$

Therefore, using the inequalities (5.10) we get

$$M_N\left(0.3(B_N\kappa_{3N})^{-1}\right) \ge \frac{3(1 - e^{-1/3})D_{1N}\rho(1+\rho)}{(1 + 3\rho(1+\rho)(2 + D_{2N}D_{1N}^{-1}))} = 3\left(1 - e^{-1/3}\right)\kappa_{4N}^{-1},$$

since $d_m\rho(1+\rho)(1 + 3\rho(1+\rho)(2 + D_{2N}D_{1N}^{-1}))^{-1} < 1/3$. Therefore, Theorems 4.5, 4.6 and 4.7 follow from Theorems 3.1, 3.2 and 3.4, respectively, and the inequality (3.7) by putting $\delta = 1$, and some simple algebra. □

References

1. Babu, G.J., Bai, Z.D.: Mixtures of global and local Edgeworth expansion and their applications. J. Multivar. Anal. **59**, 282–307 (1996)
2. Babu, G.J., Singh, E.: Edgeworth expansions for sampling without replacement from finite populations. J. Multivar. Anal. **17**, 261–278 (1985)
3. Bartlett, M.S.: The characteristic function of a conditional statistic. J. Lond. Math. **13**, 62–67 (1938)
4. Bhattacharya, R.N., Ranga Rao, R.: Normal Approximation and Asymptotic Expansions. Wiley, New York (1976)
5. Bloznelis, M.: One and two term Edgeworth expansion for finite population sample mean. Exact results. I. Lith. Math. J. **40**(3), 213–227 (2000)

6. Erdős, P., Renyi, A.: On the central limit theorem for samples from a finite population. Publ. Math. Inst. Hung. Acad. Sci. **4**, 49–61 (1959)
7. Gatto, R., Jammalamadaka, S.R.: Small sample approximations for spacings statistics. J. Stat. Plan. Inference **69**, 245–261 (1998)
8. Holst, L.: Asymptotic normality and efficiency for certain goodness-of-fit tests. Biometrika **59**, 137–145 (1972)
9. Holst, L.: A unified approach to limit theorems for urn models. J. Appl. Probab. **16**(1), 154–162 (1979)
10. Holst, L., Rao, J.S.: Asymptotic theory for families of two-sample nonparametric statistics. Sankhya **42**(Ser. A), 19–52 (1980)
11. Hu, Z., Robinson, J., Wang, Q.: Edgeworth expansion for a sample sum from a finite set of independent random variables. Electron. J. Probab. **12**, 1402–1417 (2007)
12. Ivanov, V.A., Ivchenko, G.I., Medvedev Yu, I.: Discrete problems of the probability theory (a survey). J. Sov. Math. **31**(2), 3–60 (1985)
13. Ivchenko, G.I.: Moments of separable statistics in a generalized distribution scheme. Math. Notes **39**, 154–159 (1986)
14. Ivchenko, G.I., Mirakhmedov, Sh.A.: On limit theorems for decomposable statistics and efficiency of the corresponding statistical tests. Discrete Math. Appl. **2**, 547–562 (1992)
15. Johnson, N.L., Kotz, S.: Urn Models and Their Applications. Wiley, New York (1977)
16. Kenney, J.F., Keeping, E.S.: Mathematics of Statistics, Part 2. Van Nostrand, Princeton (1953)
17. Kolchin, V.F.: In: Balakrishnan, A.V. (ed.) Random Mappings, Translations Series in Mathematics and Engg. Optimization Software Inc., New York (1985)
18. Kolchin, V.F., Sevast'yanov, B.A., Chistyakov, V.P.: Random Allocations. Winston, Washington (1978)
19. Kotz, S., Balakrishnan, N.: Advances in urn models during the past two decades. In: Advances in Combinatorial Methods and Appl. to Probab. and Statist, pp. 203–257. Birkhauser, Boston (1997)
20. Mikhaylov, A.: Polynomial and polynomial like allocation: recent developments. In: Kolchin, V.F., et al. (eds.) Probab. Methods in Discrete Math., Proc., pp. 40–59. TVP/VSP, Utrecht/Moscow (1993)
21. Mirakhmedov, Sh.A.: An asymptotic expansion for a sample sum from a finite population. Theory Probab. Appl. **28**, 492–502 (1983)
22. Mirakhmedov, Sh.A.: Estimates of proximity to the normal distribution in sampling without replacement. Theory Probab. Appl. **30**, 451–464 (1985)
23. Mirakhmedov, Sh.A.: Approximation of the distribution of multidimensional randomized divisible statistics by normal distribution (multinomial scheme). Theory Probab. Appl. **32**, 696–707 (1987)
24. Mirakhmedov, Sh.A.: Randomized decomposable statistics in the scheme of independent allocating particles into boxes. Discrete Math. Appl. **2**, 91–108 (1992)
25. Mirakhmedov, Sh.A.: Limit theorems for conditional distributions. Discrete Math. Appl. **4**, 519–542 (1994)
26. Mirakhmedov, Sh.A.: Limit theorems on decomposable statistics in a generalized allocation schemes. Discrete Math. Appl. **6**, 379–404 (1996)
27. Mirakhmedov, Sh.A.: Lower estimation of the remainder term in the CLT for a sum of the functions of k-spacings. Stat. Probab. Lett. **73**, 411–424 (2005)
28. Mirakhmedov, Sh.M.[1]: Asymptotic normality associated with generalized occupancy problem. Stat. Probab. Lett. **77**, 1549–1558 (2007)
29. Morris, C.: Central limit theorems for multinomial sums. Ann. Stat. **3**, 165–188 (1975)
30. Mukhin, A.B.: Local limit theorems for lattice random variables. Theory Probab. Appl. **36**, 660–674 (1991)
31. Pavlov, Yu.L., Cherepanova, E.V.: Limit distribution of a number of pairs in the generalized allocation scheme. Discret. Mat. **14**, 149–159 (2002) (Russian)
32. Petrov, V.V.: Sums of Independent Random Variables. Springer, New York (1995)
33. Quine, M.P., Robinson, J.: Normal approximations to sums of scores based on occupancy numbers. Ann. Probab. **13**, 794–804 (1984)
34. Quine, M.P., Robinson, J.: Efficiencies of chi-square and likelihood ratio goodness-of-fit tests. Ann. Stat. **13**, 727–742 (1985)

[1] Mirakhmedov Sh.M. was formerly Mirakhmedov Sh.A.

35. Rao, J.S., Schweitzer, R.L.: On tests for the two-sample problem based on higher order spacing-frequencies. In: Matusita, K. (ed.) Statistical Theory and Data Analysis, pp. 583–618. Nort-Holland, Amsterdam (1985)
36. Robinson, J.: An asymptotic expansion for samples from a finite population. Ann. Stat. **6**, 1004–1011 (1978)
37. Wilks, S.S.: Mathematical Statistics. Wiley, New York (1963)
38. Zhao, L.C., Wu, C.Q., Wang, Q.: Berry–Esseen bound for a sample sum from a finite set of independent random variables. J. Theor. Probab. **17**, 557–572 (2004)