Probab. Theory Relat. Fields (2010) 147:449–479
DOI 10.1007/s00440-009-0212-z

Non-central limit theorems for random selections

István Berkes · Lajos Horváth · Johannes Schauer

Received: 7 May 2008 / Revised: 20 November 2008 / Published online: 29 April 2009

Abstract Selection from finite sets is a basic procedure of statistics and the partial sum behavior of selected elements is completely known under the "uniform asymptotic negligibility" condition of central limit theory. The purpose of the present paper is to determine the asymptotic behavior of partial sums when the central limit theorem fails. As an application, we describe the limiting properties of permutation and bootstrap statistics in case of infinite variance.

Keywords Random selection · Uniform asymptotic negligibility · Bootstrap · Functional limit theorems · Permutation statistics

Mathematics Subject Classification (2000) Primary 60F17 · 62F40

Dedicated to the memory of Sándor Csörgő.

I. Berkes's research was supported by OTKA grants K 61052, K 67961, FWF grant S 9603-N23, and NSF-OTKA grant INT-0223262. L. Horváth's research was supported by NSF grant DMS 0604670.

I. Berkes (✉) · J. Schauer
Institute of Statistics, Graz University of Technology,
Münzgrabenstrasse 11, 8010 Graz, Austria
e-mail: berkes@tugraz.at

J. Schauer
e-mail: johannes.schauer@tugraz.at

L. Horváth
Department of Mathematics, University of Utah, 155 South 1440 East,
Salt Lake City, UT 84112-0090, USA
e-mail: horvath@math.utah.edu

1 Introduction

Selection from a finite population is a basic procedure of statistics and large sample properties of many classical tests and estimators are closely connected with the asymptotic behavior of sampling variables. Typical examples are bootstrap and permutation statistics, both of which assume a sample $X_1, X_2, \ldots, X_n$ of i.i.d. random variables with distribution function F and then drawing, with or without replacement, $m = m(n)$ elements from the finite set $\{X_1, \ldots, X_n\}$. The usefulness of this procedure is due to the fact that the asymptotic properties of many important functionals of the random variables $X_1^{(n)}, \ldots, X_m^{(n)}$ obtained by resampling are similar to those of the functionals of the original sample $X_1, \ldots, X_n$. Permutation and bootstrap statistics can be used, for example, to simulate critical values in statistical tests where the limit distribution of the test statistic contains unknown parameters or the convergence is too slow to use asymptotic results.

In the case when the the random variables obtained in the selection procedure satisfy the uniform asymptotic negligibility condition of classical central limit theory, the limiting behavior of their partial sums can be described easily. For each n let

$$x_{1,n} \le x_{2,n} \le \cdots \le x_{n,n}$$

be a sequence of real numbers and denote by $X_1^{(n)}, X_2^{(n)}, \ldots, X_m^{(n)}$ the random variables obtained by drawing, with or without replacement, m elements from the set $\{x_{1,n}, \ldots, x_{n,n}\}$. Define the partial sum process

$$Z_{n,m}(t) = \sum_{j=1}^{\lfloor mt \rfloor} X_j^{(n)} \quad \text{for } 0 \le t \le 1, \tag{1.1}$$

where $\lfloor \cdot \rfloor$ denotes integral part. Let $\xrightarrow{\mathcal{D}[0,1]}$ denote convergence in the space $\mathcal{D}[0, 1]$ of càdlàg functions equipped with the Skorokhod J_1-topology. The following two results are well known.

Theorem A *Let*

$$\sum_{j=1}^{n} x_{j,n} = 0, \quad \sum_{j=1}^{n} x_{j,n}^2 = 1 \tag{1.2}$$

and

$$\max_{1 \le j \le n} |x_{j,n}| \longrightarrow 0 \tag{1.3}$$

and draw $m = m(n)$ elements from the set $\{x_{1,n}, \ldots, x_{n,n}\}$ with replacement, where

$$m/n \to c \quad \textit{for some } c > 0. \tag{1.4}$$

Then

$$Z_{n,m}(t) \xrightarrow{\mathcal{D}[0,1]} W(ct) \quad \textit{for } n \to \infty,$$

where $\{W(t), 0 \le t \le 1\}$ *is a Wiener process.*

Theorem B *Assume* (1.2) *and* (1.3) *and draw* $m = m(n)$ *elements from the set* $\{x_{1,n}, \dots, x_{n,n}\}$ *without replacement, where* $m \le n$ *and* (1.4) *holds. Then*

$$Z_{n,m}(t) \xrightarrow{\mathcal{D}[0,1]} B(ct) \quad \textit{for } n \to \infty,$$

where $\{B(t), 0 \le t \le 1\}$ *is a Brownian bridge.*

In the case of Theorem A the random variables $X_1^{(n)}, \dots, X_m^{(n)}$ are i.i.d. with mean 0 and variance $1/n$ and they satisfy the Lindeberg condition

$$\lim_{n\to\infty} \sum_{j=1}^{m} E[(X_j^{(n)})^2 I\{|X_j^{(n)}| \ge \varepsilon\}] = 0 \quad \text{for any } \varepsilon > 0, \tag{1.5}$$

since the sum on the left hand side is 0 for $n \ge n_0(\varepsilon)$ by the uniform asymptotic negligibility condition (1.3). Thus Theorem A is an immediate consequence of the classical functional central limit theorem for sums of independent random variables (see [20]). Theorem B, due to Rosén [18], describes a different situation: if we sample without replacement, the r.v.'s $X_1^{(n)}, \dots, X_m^{(n)}$ are dependent and the partial sum process $Z_{n,m}(t)$ converges weakly to a process with dependent (actually negatively correlated) increments.

Typical applications of Theorems A and B include bootstrap and permutation statistics. Let $X_1, X_2, \dots$ be i.i.d. random variables with distribution function F with mean 0 and variance 1. Let $\{X_1^{(n)}, \dots, X_m^{(n)}\}$ be the bootstrap sample obtained by drawing $m = m(n)$ elements from the set $\{X_1, \dots, X_n\}$ with replacement. Clearly, $X_1^{(n)}, \dots, X_m^{(n)}$ are independent random variables with common distribution $F_n(t) = n^{-1}\sum_{i=1}^n I\{X_i \le t\}$, the empirical distribution function of the sample $X_1, \dots, X_n$. Define

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k \quad \text{and} \quad \sigma_n^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \overline{X}_n)^2$$

and apply Theorem A with the random finite set

$$\left\{ \frac{X_1 - \overline{X}_n}{\sigma_n\sqrt{n}}, \dots, \frac{X_n - \overline{X}_n}{\sigma_n\sqrt{n}} \right\}, \tag{1.6}$$

where the selection process is independent of the sequence $X_1, X_2, \dots$. It is easily checked that the conditions of Theorem A are satisfied and it follows that if (1.4) holds, then conditionally on $\mathbf{X} = (X_1, X_2, \ldots)$, for almost all paths $(X_1, X_2, \ldots)$,

$$P_{\mathbf{X}} \left\{ \frac{1}{\sigma_n \sqrt{n}} \sum_{k=1}^{\lfloor mt \rfloor} (X_k^{(n)} - \overline{X}_n) \xrightarrow{\mathcal{D}[0,1]} W(ct) \right\} = 1.$$

This fundamental limit theorem for the bootstrap is due to Bickel and Freedman [5]. On the other hand, drawing n elements from the set $\{X_1, \ldots, X_n\}$ without replacement, we get a random permutation of $X_1, \ldots, X_n$ which we denote by $X_{\pi(1)}, \ldots, X_{\pi(n)}$. Again we assume that the selection process is independent of $X_1, X_2, \ldots$. It is clear that all $n!$ permutations of $(1, 2, \ldots, n)$ are equally likely. Applying now Theorem B with the set (1.6), we get that for almost all paths $\mathbf{X} = (X_1, X_2, \ldots)$,

$$P_{\mathbf{X}} \left\{ \frac{1}{\sigma_n \sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} (X_{\pi(k)} - \overline{X}_n) \xrightarrow{\mathcal{D}[0,1]} B(t) \right\} = 1,$$

an important fact about permutation statistics.

The aim of the present paper is to prove analogues of Theorems A and B in the case when the uniform asymptotic negligibility condition (1.3) does not hold, i.e. the elements of the set $\{x_{1,n}, \ldots, x_{n,n}\}$ are not any more "small". This happens in statistical inference if the underlying distribution has infinite variance. Clearly, in this case the limiting behavior of the partial sums of the selected elements will be quite different. If, for example, the largest element $x_{n,n}$ of the set $\{x_{1,n}, \ldots, x_{n,n}\}$ does not tend to 0 as $n \to \infty$, then the contribution of $x_{n,n}$ in the partial sums of a sample of size n taken from this set clearly will not be negligible and thus the limit distribution of such sums (if it exists) will depend on this largest element. A similar effect is well known in classical central limit theory (see e.g. Bergström [4]), but the present situation will exhibit substantial additional difficulties. To simplify the formulas, we will assume throughout this paper that

$$\sum_{j=1}^{n} x_{j,n} = 0, \tag{1.7}$$

$$x_{j,n} \longrightarrow y_j \quad \text{and} \quad x_{n-j+1,n} \longrightarrow z_j \quad \text{for any fixed } j \text{ as } n \to \infty \tag{1.8}$$

for some numbers y_j, z_j, $j = 1, 2, \ldots$. We will also require that

$$\lim_{K \to \infty} \limsup_{n \to \infty} \sum_{j=K+1}^{n-K} x_{j,n}^2 = 0. \tag{1.9}$$

Condition (1.8) is no essential restriction of generality: if we assume only that the sequences $\{x_{j,n},\ n \geq 1\}$, $\{x_{n-j+1,n},\ n \geq 1\}$ are bounded for any fixed j, then by a diagonal argument we can find a subsequence of n's along which (1.8) holds. Then along this subsequence our theorems will apply and if the limiting numbers y_j, z_j are different along different subsequences, the processes $Z_{n,m}(t)$ will also have different limits along different subsequences. This seems to be rather pathological behavior, but it can happen even in simple i.i.d. situations, see Corollary 1.4 below.

The role of condition (1.9) is to exclude a Wiener or Brownian bridge component in the limiting process, as it occurs in Theorems A and B. To see this more clearly, let $r = r(n)$ denote the median of the set $\{1, 2, \dots, n\}$ and assume that $x_{j,n} = 1/\sqrt{n}$ for $[\log n] < j < r$, $x_{j,n} = -1/\sqrt{n}$ for $r < j \le n - [\log n]$ and otherwise $x_{j,n} = 0$. Then (1.7) and (1.8) are valid with $y_j = z_j = 0$, $j = 1, 2, \dots$, but from Theorem A it follows that if we select with replacement, $Z_{n,n}(t)$ converges weakly to the Wiener process.

We are now ready to formulate our results. We start with the case of selection without replacement, since the limiting process is simpler in this case.

Theorem 1.1 *Let, for each $n = 1, 2, \dots,$*

$$x_{1,n} \le x_{2,n} \le \cdots \le x_{n,n} \tag{1.10}$$

be a finite set satisfying (1.7)–(1.9) and

$$\sum_{j=1}^{\infty} y_j^2 < \infty \quad \text{and} \quad \sum_{j=1}^{\infty} z_j^2 < \infty. \tag{1.11}$$

Let $X_1^{(n)}, \dots, X_m^{(n)}$ be the random elements obtained by drawing $m = m(n) \le n$ elements from the set (1.10) without replacement, where (1.4) holds. Then for the processes $Z_{n,m}(t)$ defined by (1.1) we have

$$Z_{n,m}(t) \xrightarrow{\mathcal{D}[0,1]} R(ct) \quad \text{for } n \to \infty$$

where

$$R(t) = \sum_{j=1}^{\infty} y_j(\delta_j(t) - t) + \sum_{j=1}^{\infty} z_j(\delta_j^*(t) - t)$$

and $\{\delta_j(t),\ 0 \le t \le 1\}, \{\delta_j^*(t),\ 0 \le t \le 1\},\ j = 1, 2, \dots$ are independent jump processes, each making a single jump from 0 to 1 at a random point uniformly distributed in $(0, 1)$.

Theorem 1.2 *Let, for each $n = 1, 2, \dots,$*

$$x_{1,n} \le x_{2,n} \le \cdots \le x_{n,n} \tag{1.12}$$

be a finite set satisfying (1.7)–(1.9) and (1.11). Let $X_1^{(n)}, \dots, X_m^{(n)}$ be the random elements obtained by drawing $m = m(n)$ elements from the set (1.12) with replacement, *where* (1.4) *holds. Then for the processes* $Z_{n,m}(t)$ *defined by* (1.1) *we have*

$$Z_{n,m}(t) \xrightarrow{\mathcal{D}[0,1]} R(ct) \quad \textit{for} \;\; n \to \infty$$

where

$$R(t) = \sum_{j=1}^{\infty} y_j(\delta_j(t) - t) + \sum_{j=1}^{\infty} z_j(\delta_j^*(t) - t)$$

and $\{\delta_j(t),\ t \geq 0\}, \{\delta_j^*(t),\ t \geq 0\},\ j = 1, 2, \ldots$ *are independent Poisson processes with parameter* 1.

We now give several applications of Theorems 1.1 and 1.2. Let $X_1, X_2, \ldots$ belong to the domain of attraction of a stable r.v. ξ_α with parameter $0 < \alpha < 2$. That is, letting $S_n = \sum_{k=1}^n X_k$, we have

$$(S_n - a_n)/b_n \xrightarrow{d} \xi_\alpha \tag{1.13}$$

for some numerical sequences (a_n), (b_n). The necessary and sufficient condition for this is

$$P(X_1 > t) \sim p\,L(t)\,t^{-\alpha}, \qquad P(X_1 < -t) \sim q\,L(t)\,t^{-\alpha} \quad \text{as } t \to \infty \tag{1.14}$$

for some numbers $p \geq 0, q \geq 0, p + q = 1$ and a slowly varying function L. Let $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{n,n}$ be the ordered sample of $\{X_1, \ldots, X_n\}$ and apply Theorems 1.1 and 1.2 for the random set

$$\left\{ \frac{X_{1,n} - \overline{X}_n}{T_n}, \ldots, \frac{X_{n,n} - \overline{X}_n}{T_n} \right\}, \tag{1.15}$$

where $\overline{X}_n = \frac{1}{n}\sum_{k=1}^n X_k$ is the sample mean and $T_n = \max_{1 \leq k \leq n} |X_k|$. The normalization T_n is due to the fact that the random variables X_j are outside of the domain of attraction of the normal law. Then we get

Corollary 1.1 *Let* $X_1, X_2, \ldots$ *be i.i.d. random variables with partial sums* S_n *satisfying* (1.13) *with some* $(a_n), (b_n)$ *and a stable random variable* $\xi_\alpha, 0 < \alpha < 2$. *Let* $X_1^{(n)}, \ldots, X_m^{(n)}$ *be the variables obtained by drawing* $m = m(n) \leq n$ *times without replacement from the set* $\{X_1, \ldots, X_n\}$ *such that the selection process is independent of* $X_1, X_2, \ldots$ *and* (1.4) *holds. Let*

$$Z^*_{n,m}(t) = \frac{1}{T_n} \sum_{j=1}^{\lfloor mt \rfloor} (X_j^{(n)} - \overline{X}_n) \quad \textit{for } t \in [0, 1]. \tag{1.16}$$

Then

$$P_{\mathbf{X}}\left(Z^*_{n,m}(t) \le x\right) \xrightarrow{d} P_{\mathbf{Z}}(R(ct) \le x) \quad \text{for } n \to \infty \tag{1.17}$$

for any real x, where

$$R(t) = \frac{1}{M}\left[-q^{1/\alpha}\sum_{j=1}^{\infty}\frac{1}{Z_j^{1/\alpha}}(\delta_j(t) - t) + p^{1/\alpha}\sum_{j=1}^{\infty}\frac{1}{(Z_j^*)^{1/\alpha}}(\delta_j^*(t) - t)\right].$$

Here $Z_j = \eta_1 + \cdots + \eta_j$, $Z_j^* = \eta_1^* + \cdots + \eta_j^*$, where $\{\eta_j,\ \eta_j^*,\ j \in \mathbb{N}\}$ are independent $\exp(1)$ *random variables, $\mathbf{Z} = (Z_1, Z_1^*, Z_2, Z_2^*, \ldots)$,*

$$M = \max\left\{(q/Z_1)^{1/\alpha}, (p/Z_1^*)^{1/\alpha}\right\} \tag{1.18}$$

and $\{\delta_j(t),\ 0 \le t \le 1\}$, $\{\delta_j^*(t),\ 0 \le t \le 1\}$, $j = 1, 2, \ldots$ are independent jump processes, each making a single jump from 0 to 1 at a random point uniformly distributed in $(0, 1)$*, also independent of $\{Z_j, Z_j^*,\ j \in \mathbb{N}\}$.*

Corollary 1.2 *Corollary 1.1 remains valid if $X_1^{(n)}, \ldots, X_m^{(n)}$ are obtained by drawing with replacement from the set $\{X_1, \ldots, X_n\}$. In this case $\{\delta_j(t),\ t \ge 0\}$, $\{\delta_j^*(t),\ t \ge 0\}$, $j = 1, 2, \ldots$ will be independent Poisson processes with parameter* 1.

Note that the right hand side of (1.17) is a random variable, defined on a possibly different probability space than the r.v.'s $X_1, X_2, \ldots$. In other words, the limit distribution in Corollaries 1.1 and 1.2 is random. In case of the bootstrap statistics, this phenomenon was first noted by Athreya [1], who proved Corollary 1.2 in the case $m = n, t = 1$ (with a different representation of the limit). Another representation of the limit in the bootstrap case (still different from ours) was given by Hall [13]. In the case $m = n$ and under additional regularity assumptions on the centering sequence a_n, Corollary 1.1 was obtained in Aue et al. [2]. Note that the process $\{R(t), 0 \le t \le 1\}$ obtained in the case of selection without replacement satisfies $R(0) = R(1) = 0$ and thus it gives a nongaussian "bridge", having the same covariance (up to a constant) as Brownian bridge. Similarly, in the case of selection with replacement, $R(t)$ has the same covariance as a constant multiple of $W(t)$.

Corollaries 1.1 and 1.2 determine the limit of $P_{\mathbf{X}}\{Z^*_{n,m}(t) \le x\}$ computed conditionally on $\mathbf{X}$, i.e. under fixed sample elements $X_1, X_2, \ldots$. It is natural to ask if $Z^*_{n,m}(t)$ converges unconditionally as well. In case of selection without replacement, the answer is obvious: continuing the selection until all elements of the set are drawn (i.e. $m = n$), the vector $(X_1^{(n)}, \ldots, X_n^{(n)})$ is a random permutation of the vector $(X_1, X_2, \ldots, X_n)$ and thus its distribution is the same as that of $(X_1, \ldots, X_n)$. Consequently, $T_n^{-1}\sum_{k \le nt}(X_k^{(n)} - \overline{X}_n)$ converges weakly to a stable process, more precisely to the α-stable analogue of the Brownian bridge. In the bootstrap case the unconditional limit process of the normalized partial sums is an α-stable process with independent stationary increments. This can be proved by direct calculations using characteristic functions; an elegant proof follows from the theory of infinite dimensional stable distributions, see Ledoux and Talagrand [15, Chapter 5]. The authors thank Professor Thomas Mikosch for this observation.

It is worth mentioning that Corollaries 1.1 and 1.2 remain valid in the limiting case $\alpha = 0$, i.e. when X_1 has slowly varying tails. Let $X_1, X_2, \ldots$ be i.i.d. random variables satisfying

$$P(X_1 > t) \sim p\,L(t) \quad \text{and} \quad P(X_1 \le -t) \sim (1-p)\,L(t) \tag{1.19}$$

for $t \to \infty$ and some nonincreasing slowly varying function $L(t)$ with $\lim_{t\to\infty} L(t) = 0$. Using Theorems 1.1 and 1.2 we get the following:

Corollary 1.3 *Let $X_1, X_2, \ldots$ be i.i.d. random variables with slowly varying tails satisfying (1.19). Let $X_1^{(n)}, \ldots, X_m^{(n)}$ be the variables obtained by drawing $m = m(n) \le n$ times without replacement from the set $\{X_1, \ldots, X_n\}$, where (1.4) holds. Define $Z^*_{n,m}(t)$ by (1.16). Then*

$$P_{\mathbf{X}}(Z^*_{n,m}(t) \le x) \xrightarrow{d} P_U(R(ct) \le x)$$

for any real x with

$$R(t) = -I\{U > p\}(\delta(t) - t) + I\{U \le p\}(\delta^*(t) - t).$$

Here U is a uniform r.v. on $(0,1)$ and $\{\delta(t),\ 0 \le t \le 1\}$ and $\{\delta^*(t),\ 0 \le t \le 1\}$ are independent jump processes, both making a single jump from 0 to 1 at a uniformly distributed point on $(0,1)$, independent of U.

Here again, as throughout in our paper, the selection process is independent of the sample $(X_1, \ldots, X_n)$.

Corollary 1.3 remains valid if we sample with replacement. Then, however, $\delta(t)$ and $\delta^*(t)$ are independent Poisson processes with parameter 1, independent of the uniform r.v. U.

Our next corollary describes a situation when relation (1.8) fails, i.e. the sequences $x_{j,n}$ and $x_{n-j+1,n}$ do not converge for fixed j. Let $X_1, X_2, \ldots$ be i.i.d. symmetric random variables with the distribution

$$P(X_1 = \pm 2^k) = 2^{-(k+1)} \quad k = 1, 2, \ldots \tag{1.20}$$

This is the two-sided version of the St. Petersburg distribution. The distribution function $F(x)$ of X_1 satisfies

$$1 - F(x) = 2^{-k} \quad \text{for } 2^{k-1} \le x < 2^k$$

which shows that $G(x) = x(1 - F(x))$ is logarithmically periodic: if x runs through the interval $[2^k, 2^{k+1})$, then $G(x)$ runs through all values in $[1/2, 1)$ and $G(\log_2 x)$ is periodic with period 1. Thus (1.14) fails and consequently F does not belong to the domain of attraction of a stable law. The partial sums $S_k = \sum_{k=1}^{n} X_k$ have a remarkable behavior: for any fixed $1 \le c < 2$, the normed sums $n^{-1} S_n$ converge weakly along the subsequence $n_k = \lfloor c2^k \rfloor$ to an infinitely divisible distribution F_c such that $F_c = F_1^{*c}$ and $F_2 = F_1$. The class $\mathcal{F} = \{F_c, 1 \le c \le 2\}$, can be considered a 'circle', and in each interval $[2^k, 2^{k+1})$, the distribution of $n^{-1} S_n$ essentially runs around this circle in the sense that $n^{-1} S_n$ is close in distribution to F_c with $c = n/2^k$. This phenomenon was discovered by Csörgő [10], who called this quasiperiodic behavior 'merging'. As the following corollary shows, merging will also take place in the behavior of permutation and bootstrap statistics. For simplicity, we consider the case when we draw n elements from the sample $(X_1, \dots, X_n)$. Let $\Psi(x)$, $0 < x < \infty$ denote the function which increases linearly from 1/2 to 1 on each interval $(2^j, 2^{j+1}]$, $j = 0, \pm 1, \pm 2, \dots$.

Corollary 1.4 *Let $X_1, X_2, \dots$ be i.i.d. random variables with the distribution (1.20). Let $X_1^{(n)}, \dots, X_n^{(n)}$ be the elements obtained by drawing n times with replacement from the set $\{X_1, \dots, X_n\}$ and let $Z_n^*(t)$ be defined by (1.16) with $m = n$. Let $1 \le c < 2$. Then for $n_k = \lfloor c\, 2^k \rfloor$ we have*

$$P_{\mathbf{X}}(Z_{n_k}^*(t) \le x) \xrightarrow{d} P_{\mathbf{Z}}(R_c(t) \le x)$$

for any real x, where

$$R_c(t) = \frac{1}{M}\left[-\sum_{j=1}^{\infty} \frac{1}{Z_j} \Psi\left(\frac{Z_j}{c}\right)(\delta_j(t) - t) + \sum_{j=1}^{\infty} \frac{1}{Z_j^*} \Psi\left(\frac{Z_j^*}{c}\right)(\delta_j^*(t) - t) \right]$$

with

$$M = \max\left\{ \frac{\Psi(Z_1/c)}{Z_1}, \frac{\Psi\left(Z_1^*/c\right)}{Z_1^*} \right\}.$$

Here $Z_j = \eta_1 + \dots + \eta_j$ and $Z_j^* = \eta_1^* + \dots + \eta_j^*$, where $\{\eta_j, \eta_j^*, j \in \mathbb{N}\}$ are i.i.d. exp(1) *random variables and $\{\delta_j(t), 0 \le t \le 1\}$, $\{\delta_j^*(t), 0 \le t \le 1\}$, $j = 1, 2, \dots$ are independent jump processes, each making a single jump from 0 to 1 at a uniformly distributed point in* (0, 1).

Just like in the case of partial sums, the class R_c of limiting processes is logarithmically periodic, namely $R_{2c} = R_c$ and for a fixed n with $2^k \le n < 2^{k+1}$ the conditional distribution of $Z_n^*(t)$ is close to that of $R_c(t)$ with $c = n/2^k$.

Corollary 1.4 remains valid if we draw $X_1^{(n)}, \dots, X_n^{(n)}$ without replacement from the set $\{X_1, \dots, X_n\}$. Then $\delta_j(t)$ and $\delta_j^*(t)$ are independent Poisson processes with parameter 1.

2 Proofs

We will prove Theorems 1.1 and 1.2 in the case $m = n$; the proofs in the general case require only trivial changes. For studying the sample $(X_1^{(n)}, \dots, X_n^{(n)})$ we introduce random variables $\{\varepsilon_1^{(n)}(t), \ldots, \varepsilon_n^{(n)}(t)\}$, where $\varepsilon_j^{(n)}(t)$ counts how many times $x_{j,n}$ has been chosen among the first $\lfloor nt \rfloor$ sampled elements:

$$\varepsilon_j^{(n)}(t) = k \quad \text{if } x_{j,n} \text{ is chosen } k \text{ times among the first } \lfloor nt \rfloor \text{ elements}$$

for $j = 1, \ldots, n$ and $k \in \{0, 1, \ldots, \lfloor nt \rfloor\}$ when drawing with replacement and $k \in \{0, 1\}$ when drawing without replacement. Obviously the distribution of the $\varepsilon_j^{(n)}(t)$ depends on the selection method. If we draw without replacement, the $\varepsilon_j^{(n)}(t)$ only take the values 0 or 1, as an element can be chosen at most once during some time interval. Clearly

$$Z_n(t) = \sum_{j=1}^{\lfloor nt \rfloor} X_j^{(n)} = \sum_{j=1}^{n} x_{j,n} \varepsilon_j^{(n)}(t) = \sum_{j=1}^{n} x_{j,n} \bar{\varepsilon}_j^{(n)}(t), \tag{2.1}$$

as $\sum_{j=1}^n x_{j,n} = 0$ and the $\varepsilon_j^{(n)}$ are equidistributed. Here $\bar{\varepsilon}_j^{(n)}(t) = \varepsilon_j^{(n)}(t) - E\varepsilon_j^{(n)}(t)$ is the centered version of $\varepsilon_j^{(n)}(t)$.

Note that in Theorems 1.1 and 1.2 we did not assume that the elements of the set (1.10) are different. If, e.g. $x_{1,n} = x_{2,n}$, in the selection procedure $x_{1,n}, x_{2,n}$ should be considered different elements of the set (1.10) and $\varepsilon_1^{(n)}(t)$ and $\varepsilon_2^{(n)}(t)$ denote how many times these (otherwise equal) elements were selected in the first $\lfloor nt \rfloor$ steps. Clearly, the representation (2.1) remains valid in this case.

The following two subsections cover the two different sampling methods.

2.1 Selection with replacement

Since we draw $\lfloor nt \rfloor$ times with replacement from a set with n elements, the vector $\{\varepsilon_1^{(n)}(t), \ldots, \varepsilon_n^{(n)}(t)\}$ follows a multinomial distribution with $p_j = 1/n$ for $j = 1, \ldots, n$ and $\lfloor nt \rfloor$ draws. Obviously the marginal distribution of $\varepsilon_j^{(n)}(t)$ is binomial with parameters $1/n$ and $\lfloor nt \rfloor$. In particular, this implies $E\varepsilon_j^{(n)}(t) = \lfloor nt \rfloor / n$.

In order to prove Theorem 1.2 we need several lemmas, which will be stated next. We first show that under the assumptions of Theorem 1.2 only the very small and very large order statistics will contribute asymptotically to $Z_n(t)$. This will be shown in Lemmas 2.1 and 2.5.

Lemma 2.1 *If* (1.9) *holds, then*

$$\lim_{K \to \infty} \limsup_{n \to \infty} P\left(\left| \sum_{j=K+1}^{n-K} x_{j,n} \bar{\varepsilon}_j^{(n)}(t) \right| \geq \delta \right) = 0$$

for all $\delta > 0$ *and* $0 \leq t \leq 1$.

Proof We get, using the properties of the binomial and multinomial distribution,

$$E\left(\left(\bar{\varepsilon}_j^{(n)}(t)\right)^2\right) = \mathrm{Var}\left(\varepsilon_j^{(n)}(t)\right) = \lfloor nt\rfloor\frac{n-1}{n^2} \quad \text{for } j = 1,\ldots,n$$

and

$$E\left(\bar{\varepsilon}_j^{(n)}(t)\bar{\varepsilon}_k^{(n)}(t)\right) = \mathrm{Cov}\left(\varepsilon_j^{(n)}(t)\varepsilon_k^{(n)}(t)\right) = -\frac{\lfloor nt\rfloor}{n^2} \quad \text{for } 1 \le j < k \le n.$$

Furthermore we have

$$E\left(\sum_{j=K+1}^{n-K} x_{j,n}\bar{\varepsilon}_j^{(n)}(t)\right) = 0$$

and

$$\begin{aligned}
\mathrm{Var}\left(\sum_{j=K+1}^{n-K} x_{j,n}\bar{\varepsilon}_j^{(n)}(t)\right) &= \sum_{j,k=K+1}^{n-K} x_{j,n}x_{k,n}E\left(\bar{\varepsilon}_j^{(n)}(t)\bar{\varepsilon}_k^{(n)}(t)\right)\\
&= \lfloor nt\rfloor\frac{n-1}{n^2}\sum_{j=K+1}^{n-K} x_{j,n}^2 - \frac{\lfloor nt\rfloor}{n^2}\sum_{\substack{j,k=K+1\\k\neq j}}^{n-K} x_{j,n}x_{k,n}\\
&= \frac{\lfloor nt\rfloor}{n}\sum_{j=K+1}^{n-K} x_{j,n}^2 - \frac{\lfloor nt\rfloor}{n^2}\sum_{j,k=K+1}^{n-K} x_{j,n}x_{k,n}\\
&= \frac{\lfloor nt\rfloor}{n}\sum_{j=K+1}^{n-K} x_{j,n}^2 - \frac{\lfloor nt\rfloor}{n^2}\left(\sum_{j=K+1}^{n-K} x_{j,n}\right)^2\\
&\le \sum_{j=K+1}^{n-K} x_{j,n}^2.
\end{aligned}$$

The statement of Lemma 2.1 now follows using the Markov inequality and (1.9). □

The following consequence of the proof of Lemma 2.1 and relation (2.1) will be convenient in applications of our theorems.

Corollary 2.1 *Let $\tilde{Z}_n(t)$ denote the analogue of $Z_n(t)$ when the set $\{x_{1,n},\ldots,x_{n,n}\}$ is replaced by another set $\{\tilde{x}_{1,n},\ldots,\tilde{x}_{n,n}\}$. Then assuming* (1.7) *for both sets* (*but without assuming* (1.8) *or* (1.9)), *we have for any* $0 \le t \le 1$

$$E(\tilde{Z}_n(t) - Z_n(t))^2 \le \sum_{j=1}^{n}(x_{j,n} - \tilde{x}_{j,n})^2.$$

An analogous statement holds in the case of selection without replacement.

In what follows, let dist(X) and dist($X|Y$) denote, respectively, the distribution of the random vector X and its conditional distribution relative to the random vector Y. The following lemma, due to Berkes and Philipp [3], will be crucial for our approximation argument.

Lemma 2.2 *Let $\{X_k, k \geq 1\}$ be a sequence of random vectors with values in $\mathbb{Z}^d$ defined on an atomless probability space. Suppose that*

$$P(\rho(dist(X_k|X_1, \dots, X_{k-1}), dist(X_k)) \geq \varepsilon_k) \leq \varepsilon_k \quad \text{for all } k \geq 1, \tag{2.2}$$

where ρ denotes the Prokhorov distance. Then there exist independent random vectors $\{Y_k, k \geq 1\}$ with values in $\mathbb{Z}^d$ such that $X_k \stackrel{d}{=} Y_k$ and

$$P(|X_k - Y_k| > 6\varepsilon_k) < 6\varepsilon_k \quad \text{for all } k \geq 1.$$

Lemma 2.2 is implicit in Theorem 2 of Berkes and Philipp [3], which assumes a mixing condition, but the proof uses only (2.2).

The following lemma shows that in the index range $j \in \{1, \dots, K, n-K+1, \dots, n\}$ for K "not too large" (roughly for $K \leq \sqrt{n}$) the dependent random variables $\varepsilon_j^{(n)}(t)$ can be approximated by independent binomial random variables.

Lemma 2.3 *If (1.9) holds, then for every $0 < t < 1$ and $n \geq 2/t$ there exist independent random variables $\delta_j^{(n)}(t)$, $j = 1, 2, \dots n$, with binomial distribution $B(\lfloor nt \rfloor, 1/n)$, such that*

$$P\left(\sum_{j=1}^{K} x_{j,n}(\varepsilon_j^{(n)}(t) - \delta_j^{(n)}(t)) \neq 0\right) \leq Ct^{-1}K^2 \frac{(\log n)^3}{n} \tag{2.3}$$

and

$$P\left(\sum_{j=n-K+1}^{n} x_{j,n}\left(\varepsilon_j^{(n)}(t) - \delta_j^{(n)}(t)\right) \neq 0\right) \leq Ct^{-1}K^2 \frac{(\log n)^3}{n} \tag{2.4}$$

for all $K = 1, \dots, \lfloor n/4 \rfloor$, where C is an absolute constant.

Technically, relations (2.3) and (2.4) are valid for $1 \leq K \leq \lfloor n/4 \rfloor$, but they give a trivial bound for $K \geq \text{const}\sqrt{n}/(\log n)^{3/2}$. We will use the lemma for constant K.

Proof Let, for $j \geq 1$, $\gamma_{2j-1} = \varepsilon_j^{(n)}(t)$ and $\gamma_{2j} = \varepsilon_{n-j+1}^{(n)}(t)$. To approximate the dependent $\varepsilon_j^{(n)}(t)$ with the independent $\delta_j^{(n)}(t)$, we use Lemma 2.2 and thus we need an estimate for the difference

$$|P(\gamma_{k+1} = a_{k+1}|\gamma_1 = a_1, \dots, \gamma_k = a_k) - P(\gamma_{k+1} = a_{k+1})|.$$

We are drawing $\lfloor nt \rfloor$ times with replacement from the set (1.10) and γ_k $(k = 1, \dots, n)$ counts how many times the corresponding element was drawn from the set and follows a binomial distribution:

$$P(\gamma_k = a_k) = \binom{\lfloor nt \rfloor}{a_k} \left(\frac{1}{n}\right)^{a_k} \left(\frac{n-1}{n}\right)^{\lfloor nt \rfloor - a_k} \quad \text{for } 0 \le a_k \le \lfloor nt \rfloor.$$

On the other hand, letting $a^{(k)} = \sum_{i=1}^{k} a_i$ we see that

$$\begin{aligned} &P(\gamma_{k+1} = a_{k+1} | \gamma_1 = a_1, \dots, \gamma_k = a_k) \\ &\quad = \binom{\lfloor nt \rfloor - a^{(k)}}{a_{k+1}} \left(\frac{1}{n-k}\right)^{a_{k+1}} \left(\frac{n-k-1}{n-k}\right)^{\lfloor nt \rfloor - a^{(k)} - a_{k+1}} \end{aligned}$$

for $0 \le a_{k+1} \le \lfloor nt \rfloor - a^{(k)}$. Obviously $\gamma_{k+1} = 0$ in the case of $a^{(k)} = \lfloor nt \rfloor$.

Consider first the case $a_i \le C_1 \log n$, $1 \le i \le n$ for some positive constant C_1. By the assumption $1 \le K \le n/4$ of the lemma, it suffices to consider the case $k \le n/2$. Letting

$$T = \frac{(\lfloor nt \rfloor - a^{(k)}) \cdots (\lfloor nt \rfloor - a^{(k)} - a_{k+1} + 1)}{\lfloor nt \rfloor \cdots (\lfloor nt \rfloor - a_{k+1} + 1)}$$

we get

$$\begin{aligned} &|P(\gamma_{k+1} = a_{k+1} | \gamma_1 = a_1, \dots, \gamma_k = a_k) - P(\gamma_{k+1} = a_{k+1})| \\ &\quad = \left| \binom{\lfloor nt \rfloor - a^{(k)}}{a_{k+1}} \left(\frac{1}{n-k}\right)^{a_{k+1}} \left(\frac{n-k-1}{n-k}\right)^{\lfloor nt \rfloor - a^{(k)} - a_{k+1}} \right. \\ &\qquad \left. - \binom{\lfloor nt \rfloor}{a_{k+1}} \left(\frac{1}{n}\right)^{a_{k+1}} \left(\frac{n-1}{n}\right)^{\lfloor nt \rfloor - a_{k+1}} \right| \\ &\quad = \binom{\lfloor nt \rfloor}{a_{k+1}} \left| T \left(\frac{1}{n-k}\right)^{a_{k+1}} \left(\frac{n-k-1}{n-k}\right)^{\lfloor nt \rfloor - a^{(k)} - a_{k+1}} \right. \\ &\qquad \left. - \left(\frac{1}{n}\right)^{a_{k+1}} \left(\frac{n-1}{n}\right)^{\lfloor nt \rfloor - a_{k+1}} \right| \\ &\quad \le \binom{\lfloor nt \rfloor}{a_{k+1}} \left(\frac{n-k-1}{n-k}\right)^{\lfloor nt \rfloor - a^{(k)} - a_{k+1}} \left| T \frac{1}{(n-k)^{a_{k+1}}} - \frac{1}{n^{a_{k+1}}} \right| \\ &\qquad + \binom{\lfloor nt \rfloor}{a_{k+1}} \left(\frac{1}{n}\right)^{a_{k+1}} \left| \left(\frac{n-k-1}{n-k}\right)^{\lfloor nt \rfloor - a^{(k)} - a_{k+1}} - \left(\frac{n-1}{n}\right)^{\lfloor nt \rfloor - a_{k+1}} \right|. \end{aligned}$$

Using the mean value theorem, we get

$$\begin{aligned}
\left| T\frac{1}{(n-k)^{a_{k+1}}} - \frac{1}{n^{a_{k+1}}}\right| &\le \frac{1}{(n-k)^{a_{k+1}}}|T-1| + \left|\frac{1}{(n-k)^{a_{k+1}}} - \frac{1}{n^{a_{k+1}}}\right| \\
&\le \frac{1}{(n-k)^{a_{k+1}}}\left(1-\left(\frac{\lfloor nt\rfloor - a^{(k)} - a_{k+1}+1}{\lfloor nt\rfloor}\right)^{a_{k+1}}\right) \\
&\quad + \frac{|n^{a_{k+1}} - (n-k)^{a_{k+1}}|}{(n-k)^{a_{k+1}} n^{a_{k+1}}} \\
&\le \frac{1}{(n-k)^{a_{k+1}}}\left(1-\left(\frac{\lfloor nt\rfloor - (k+1)\,C_1\log n}{\lfloor nt\rfloor}\right)^{C_1\log n}\right) \\
&\quad + \frac{a_{k+1} n^{a_{k+1}-1} k}{(n-k)^{a_{k+1}} n^{a_{k+1}}} \\
&\le \frac{1}{(n-k)^{a_{k+1}}}\, C_1\log n \frac{(k+1)\,C_1\log n}{\lfloor nt\rfloor} + \frac{k\,C_1\log n}{n(n-k)^{a_{k+1}}} \\
&\le \frac{1}{(n/2)^{a_{k+1}}}\frac{6k\,C_1^2(\log n)^2}{nt}.
\end{aligned}$$

Similarly, we get

$$\begin{aligned}
&\left|\left(\frac{n-k-1}{n-k}\right)^{\lfloor nt\rfloor - a^{(k)} - a_{k+1}} - \left(\frac{n-1}{n}\right)^{\lfloor nt\rfloor - a_{k+1}}\right| \\
&\quad \le \left|\left(\frac{n-k-1}{n-k}\right)^{\lfloor nt\rfloor - a^{(k)} - a_{k+1}} - \left(\frac{n-1}{n}\right)^{\lfloor nt\rfloor - a^{(k)} - a_{k+1}}\right| \\
&\qquad + \left|\left(\frac{n-1}{n}\right)^{\lfloor nt\rfloor - a^{(k)} - a_{k+1}} - \left(\frac{n-1}{n}\right)^{\lfloor nt\rfloor - a_{k+1}}\right| \\
&\quad \le (\lfloor nt\rfloor - a^{(k)} - a_{k+1})\left(\frac{1}{n-k} - \frac{1}{n}\right) + \left(\frac{n-1}{n}\right)^{\lfloor nt\rfloor - a^{(k)} - a_{k+1}}\left(1-\left(\frac{n-1}{n}\right)^{a^{(k)}}\right) \\
&\quad \le (\lfloor nt\rfloor - a^{(k)} - a_{k+1})\frac{k}{n(n-k)} + \frac{a^{(k)}}{n}.
\end{aligned}$$

Putting together the previous estimates and using

$$\binom{\lfloor nt\rfloor}{a_{k+1}} \le n^{a_{k+1}}/a_{k+1}! \le 2(n/2)^{a_{k+1}}$$

we obtain for $k \le n/2$, $a_i \le C_1\log n$, $1 \le i \le n$

$$|P(\gamma_{k+1} = a_{k+1} | \gamma_1 = a_1, \ldots, \gamma_k = a_k) - P(\gamma_{k+1} = a_{k+1})| \le C_2\, t^{-1} k \frac{(\log n)^2}{n}. \quad (2.5)$$

Observe now that

$$\begin{aligned}
&P\left(\gamma_{k+1} > C_1 \log n | \gamma_1 = a_1, \ldots, \gamma_k = a_k\right)\\
&\quad = \sum_{j > C_1 \log n} \binom{\lfloor nt \rfloor - a^{(k)}}{j} \left(\frac{1}{n-k}\right)^j \left(\frac{n-k-1}{n-k}\right)^{\lfloor nt \rfloor - a^{(k)} - j}\\
&\quad \le \sum_{j > C_1 \log n} \frac{n^j}{j!} \frac{1}{(n/2)^j} \le \sum_{j > C_1 \log n} 2^{-j} \le n^{-2}
\end{aligned} \tag{2.6}$$

provided the constant C_1 is large enough. Similarly we get

$$P(\gamma_j > C_1 \log n) \le n^{-2} \quad 1 \le j \le n. \tag{2.7}$$

Putting together (2.5)–(2.7) we obtain

$$\begin{aligned}
&\sum_j |P(\gamma_{k+1} = j | \gamma_1 = a_1, \ldots, \gamma_k = a_k) - P(\gamma_{k+1} = j)|\\
&\quad = \sum_{j \le C_1 \log n} + \sum_{j > C_1 \log n} \le C_3 t^{-1} k \frac{(\log n)^3}{n} + n^{-2} \le C_4 t^{-1} k \frac{(\log n)^3}{n}.
\end{aligned}$$

This implies that for any atom $A = \{\gamma_1 = a_1, \ldots, \gamma_k = a_k\}$ with $a_i \le C_1 \log n$ $(1 \le i \le k)$, the Prokhorov distance of $\operatorname{dist}(\gamma_{k+1}|A)$ and $\operatorname{dist}(\gamma_{k+1})$ is at most $C_4 t^{-1} k (\log n)^3 n^{-1}$. Letting B denote the union of such atoms, by (2.7) we have

$$P(B^c) \le P\left(\max(\gamma_1, \ldots, \gamma_k) > C_1 \log n\right) \le k n^{-2} \le t^{-1} k \frac{(\log n)^3}{n}.$$

Thus we proved that for $k \le n/2$

$$P\left(\rho\left(\operatorname{dist}(\gamma_{k+1}|\gamma_k, \ldots, \gamma_1), \operatorname{dist}(\gamma_{k+1})\right) \ge C_5 t^{-1} k \frac{(\log n)^3}{n}\right) \le C_5 t^{-1} k \frac{(\log n)^3}{n}.$$

Clearly, for $n/2 < k \le n-1$ we have

$$\rho\left(\operatorname{dist}(\gamma_{k+1}|\gamma_k, \ldots, \gamma_1), \operatorname{dist}(\gamma_{k+1})\right) \le 1$$

and thus applying Lemma 2.2 we get that there exist independent random variables γ_k^*, $k = 1, \ldots, n$, such that $\gamma_k^* \stackrel{d}{=} \gamma_k$ and

$$P\left(|\gamma_k - \gamma_k^*| > C_6 t^{-1} k \frac{(\log n)^3}{n}\right) \le C_6 t^{-1} k \frac{(\log n)^3}{n} \quad 1 \le k \le n/2.$$

Since the variables γ_k and γ_k^* take only nonnegative integer values, the last relation implies

$$P(\gamma_k \neq \gamma_k^*) \leq C_6 t^{-1} k \frac{(\log n)^3}{n} \quad 1 \leq k \leq n/2,$$

as one can see separately in the cases when $C_6 t^{-1} k (\log n)^3 n^{-1}$ is < 1 or not. Letting $\delta_j^{(n)}(t) = \gamma_{2j-1}^*$ and $\delta_{n-j+1}^{(n)}(t) = \gamma_{2j}^*$, we get the statement of the lemma. □

The next lemma is a generalization of Lemma 2.3 for the finite dimensional distributions of $\varepsilon_j^{(n)}(t)$. We will formulate it in the simpler form that will be needed in the proof of Theorem 1.2.

Lemma 2.4 *If* (1.9) *holds, then for every n and all $0 < t_1 < \cdots < t_d < 1$ there exist independent, identically distributed random vectors* $\left(\delta_j^{(n)}(t_1), \ldots, \delta_j^{(n)}(t_d)\right)$, *$j = 1, \ldots, n$ such that*

$$\left(\delta_j^{(n)}(t_1), \ldots, \delta_j^{(n)}(t_d)\right) \stackrel{d}{=} (Q(U, n, t_1), \ldots, Q(U, n, t_d))$$

and

$$P\left(\max_{1 \leq l \leq d} \left| \sum_{j=1}^{K} x_{j,n} \left(\varepsilon_j^{(n)}(t_l) - \delta_j^{(n)}(t_l)\right) \right| \geq \delta \right) \to 0$$

$$P\left(\max_{1 \leq l \leq d} \left| \sum_{j=n-K+1}^{n} x_{j,n} \left(\varepsilon_j^{(n)}(t_l) - \delta_j^{(n)}(t_l)\right) \right| \geq \delta \right) \to 0$$

for all $\delta > 0$, $K \geq 1$ and $n \to \infty$. Here U is a uniform random variable on $[0, 1]$ *and $Q(u, n, t)$ is the quantile function of the $B(\lfloor nt \rfloor, 1/n)$ distribution.*

An explicit formula for $Q(u, n, t)$ is

$$Q(u, n, t) = \begin{cases} 0 & \text{if } u \leq p_0(n, t) \\ 1 & \text{if } p_0(n, t) < u \leq p_1(n, t) \\ \vdots & \vdots \\ \lfloor nt \rfloor & \text{if } p_{\lfloor nt \rfloor - 1}(n, t) < u \end{cases}$$

with $p_k(n, t) = \sum_{j=0}^{k} \binom{\lfloor nt \rfloor}{j} \left(\frac{1}{n}\right)^j \left(\frac{n-1}{n}\right)^{\lfloor nt \rfloor - j}$, but we will not need this fact.

Proof By applying the same procedure as in Lemma 2.3 to the random vector $(\varepsilon_j^{(n)}(t_1), \ldots, \varepsilon_j^{(n)}(t_d))$ instead of $\varepsilon_j^{(n)}(t)$, Lemma 2.4 can be proven with some minor changes. □

Let $\bar{\delta}_j^{(n)}(t) = \delta_j^{(n)}(t) - E\delta_j^{(n)}(t)$. The following lemma is the equivalent of Lemma 2.1 for the $\bar{\delta}_j^{(n)}(t)$ (instead of the $\bar{\varepsilon}_j^{(n)}(t)$).

Lemma 2.5 *If* (1.9) *holds, then*

$$\lim_{K\to\infty} \limsup_{n\to\infty} P\left(\left|\sum_{j=K+1}^{n-K} x_{j,n}\bar{\delta}_j^{(n)}(t)\right| \geq \delta\right) = 0$$

for all $\delta > 0$ *and* $0 \leq t \leq 1$.

Proof The proof can be carried out as in the case of Lemma 2.1. □

We are now ready to complete the proof of Theorem 1.2.

Proof of Theorem 1.2 We first prove the convergence of the finite dimensional distributions. To simplify the formulas, we consider the one-dimensional case; the changes needed for the d-dimensional case will be stated at the end of the proof. By Lemma 2.1 we have

$$\limsup_{n\to\infty} \mathcal{L}\left(\operatorname{dist}\sum_{j=1}^{n} x_{j,n}\bar{\varepsilon}_j^{(n)}(t), \operatorname{dist}\sum_{j\in[1,K]\cup[n-K+1,n]} x_{j,n}\bar{\varepsilon}_j^{(n)}(t)\right) = B(K) \qquad (2.8)$$

with $B(K) \longrightarrow 0$ as $K \to \infty$, where $\mathcal{L}$ denotes the Lévy distance. We used the fact that if for two random variables ξ and η we have $P(|\xi - \eta| \geq \varepsilon) \leq \varepsilon$, then we also have $\mathcal{L}(\operatorname{dist}\xi, \operatorname{dist}\eta) \leq \varepsilon$. Lemma 2.5 furthermore tells us that (2.8) remains valid if we replace $\bar{\varepsilon}_j^{(n)}(t)$ with $\bar{\delta}_j^{(n)}(t)$. We know that $\bar{\varepsilon}_j^{(n)}(t) - \bar{\delta}_j^{(n)}(t) = \varepsilon_j^{(n)}(t) - \delta_j^{(n)}(t)$, therefore Lemma 2.3 implies that for any fixed K

$$\mathcal{L}\left(\operatorname{dist}\sum_{j\in[1,K]\cup[n-K+1,n]} x_{j,n}\bar{\varepsilon}_j^{(n)}(t), \operatorname{dist}\sum_{j\in[1,K]\cup[n-K+1,n]} x_{j,n}\bar{\delta}_j^{(n)}(t)\right) \longrightarrow 0 \qquad (2.9)$$

as $n \to \infty$. Observe that for any real sequences $\{c_j\}$, $\{c'_j\}$ and any $\lambda > 0$

$$\begin{aligned} P\left(\left|\sum_{j\in[1,K]\cup[n-K+1,n]} (c_j - c'_j)\bar{\delta}_j^{(n)}(t)\right| \geq \lambda\right) &\leq \frac{1}{\lambda}\sum_{j\in[1,K]\cup[n-K+1,n]} |c_j - c'_j|\, E\left|\bar{\delta}_j^{(n)}(t)\right| \\ &\leq \frac{1}{\lambda}\sum_{j\in[1,K]\cup[n-K+1,n]} 2|c_j - c'_j| \end{aligned}$$

whence by a proper choice of λ we get

$$\begin{aligned} &\mathcal{L}\left(\operatorname{dist}\sum_{j\in[1,K]\cup[n-K+1,n]} c_j\,\bar{\delta}_j^{(n)}(t), \operatorname{dist}\sum_{j\in[1,K]\cup[n-K+1,n]} c'_j\,\bar{\delta}_j^{(n)}(t)\right) \\ &\quad \leq \sqrt{2}\left(\sum_{j\in[1,K]\cup[n-K+1,n]} |c_j - c'_j|\right)^{1/2}. \end{aligned}$$

Consequently, by (1.8)

$$\mathcal{L}\left(\operatorname{dist} \sum_{j\in[1,K]\cup[n-K+1,n]} x_{j,n}\bar{\delta}_j^{(n)}(t),\right.$$
$$\left.\operatorname{dist}\left[\sum_{j=1}^{K} y_j\bar{\delta}_j^{(n)}(t)+\sum_{j=1}^{K} z_j\bar{\delta}_{n-j+1}^{(n)}(t)\right]\right)\longrightarrow 0 \tag{2.10}$$

as $n \to \infty$. Hence, letting $\delta_j(t), \delta_j^*(t),\ j = 1, 2, \ldots$ denote independent Poisson processes with parameter 1, it suffices to show that

$$\mathcal{L}\left(\operatorname{dist}\left[\sum_{j=1}^{K} y_j\bar{\delta}_j^{(n)}(t)+\sum_{j=1}^{K} z_j\bar{\delta}_{n-j+1}^{(n)}(t)\right],\right.$$
$$\left.\operatorname{dist}\left[\sum_{j=1}^{K} y_j(\delta_j(t)-t)+\sum_{j=1}^{K} z_j(\delta_j^*(t)-t)\right]\right)\longrightarrow 0 \tag{2.11}$$

as $n \to \infty$ for any fixed K and that the second distribution in (2.11) converges to the same expression with $K = \infty$. To prove the first statement let us note that a sharpened form of the Poisson approximation of the binomial due to Le Cam [14, p. 187] implies that

$$\sum_{l=0}^{\infty} |P(\delta_j^{(n)}(t)=l)-P(\delta_j(t)=l)| \le C_7 n^{-1} \tag{2.12}$$

for any $j \ge 1, n \ge 1$ with an absolute constant C_7. This implies that

$$|P(\delta_j^{(n)}(t)\in A)-P(\delta_j(t)\in A)| \le C_7 n^{-1}$$

for any set $A \subset \{0, 1, \ldots\}$ and therefore

$$\rho(\operatorname{dist}\delta_j^{(n)}(t), \operatorname{dist}\delta_j(t)) \le C_7 n^{-1},$$

proving (2.11). The convergence of the distributions in the second line of (2.11) is an immediate consequence of condition (1.11), $E(\delta_j(t)-t)^2 = E(\delta_j^*(t)-t)^2 = t$ and the Kolmogorov two series theorem, implying the a.s. convergence of the series

$$\sum_{j=1}^{\infty}\Big[y_j(\delta_j(t)-t)+z_j(\delta_j^*(t)-t)\Big].$$

By a theorem of Lévy (see e.g. Breiman [8, p. 51, Problem 16]), the distribution of the last sum is actually continuous. As weak convergence of distributions to a continuous limit implies that the corresponding distribution functions converge pointwise, we have proved the convergence of the one-dimensional distributions in Theorem 1.2. The proof of the corresponding statement for the multi-dimensional distributions uses the same arguments, we just need Lemma 2.4 instead of Lemma 2.3. Therefore we omit the details. □

So far we proved the convergence of finite dimensional distributions of $Z_n(t)$. To prove the tightness of $Z_n(t)$ in $\mathcal{D}[0, 1]$ it suffices, in view of Theorem 15.6 on p. 128 in Billingsley [6], to show that

$$E(Z_n(s) - Z_n(t_1))^2(Z_n(t_2) - Z_n(s))^2 \leq C(t_2 - t_1)^2 \quad \text{for all } t_1 \leq s \leq t_2. \tag{2.13}$$

To prove (2.13) we first observe that by (1.8) and (1.9) we can find a constant $c > 0$ such that $\sum_{j=1}^{n} x_{j,n}^2 \leq c$ for all n. Letting $l_1 := \lfloor ns \rfloor - \lfloor nt_1 \rfloor$ and $l_2 := \lfloor nt_2 \rfloor - \lfloor ns \rfloor$, we get [recall $EX_j^{(n)} = 0$ by (1.2)],

$$\begin{aligned} E(Z_n(s) - Z_n(t_1))^2 &= E\left(\sum_{j=\lfloor nt_1 \rfloor+1}^{\lfloor ns \rfloor} X_j^{(n)}\right)^2 \\ &= l_1 E\left((X_j^{(n)})^2\right) + l_1(l_1 - 1)(EX_j^{(n)})^2 \\ &= \frac{l_1}{n} \sum_{j=1}^{n} x_{j,n}^2 \leq \frac{l_1}{n} c. \end{aligned}$$

A similar inequality holds for $E(Z_n(t_2) - Z_n(s))^2$ and thus by the independence of the two differences on the left hand side of (2.13) we get

$$E\,(Z_n(s) - Z_n(t_1))^2\,(Z_n(t_2) - Z_n(s))^2 \;\leq \frac{l_1 l_2}{n^2} c^2.$$

In the case of $t_2 - t_1 < \frac{1}{n}$ at least two of the 3 numbers nt_1, ns, nt_2 lie between two consecutive integers and thus one of the differences in (2.13) is 0. If $t_2 - t_1 \geq \frac{1}{n}$ then we get

$$\frac{l_1 l_2}{n^2} \leq \left(\frac{l_1 + l_2}{n}\right)^2 \leq \left(\frac{nt_2 - nt_1 + 1}{n}\right)^2 \leq 4(t_2 - t_1)^2,$$

which completes the proof of (2.13).

2.2 Selection without replacement

In the case of selection without replacement, the variables $\varepsilon_j^{(n)}(t)$ only take the values 0 or 1. Therefore the distribution of a single $\varepsilon_j^{(n)}(t)$ is Bernoulli with $P(\varepsilon_j^{(n)}(t) = 1) = \lfloor nt \rfloor / n$, while for the vector $\left(\varepsilon_1^{(n)}(t), \ldots, \varepsilon_n^{(n)}(t)\right)$ we have

$$P\left(\varepsilon_1^{(n)}(t) = a_1, \ldots, \varepsilon_n^{(n)}(t) = a_n\right) = 1 \Big/ \binom{n}{\lfloor nt \rfloor}$$

provided all a_j's are 0 or 1 and $\sum_{i=1}^{n} a_i = \lfloor nt \rfloor$; otherwise the probability is 0. As in the case of selection with replacement, we get $E\varepsilon_j^{(n)}(t) = \lfloor nt \rfloor/n$ for $1 \le j \le n$; let $\bar{\varepsilon}_j^{(n)}(t) = \varepsilon_j^{(n)}(t) - E\varepsilon_j^{(n)}(t)$. In the sequel we formulate lemmas which are the permutation analogues of the lemmas in the proof of Theorem 1.2.

Lemma 2.6 *If* (1.9) *holds, then*

$$\lim_{K\to\infty} \limsup_{n\to\infty} P\left(\left|\sum_{j=K+1}^{n-K} x_{j,n}\bar{\varepsilon}_j^{(n)}(t)\right| \ge \delta\right) = 0$$

for all $\delta > 0$ *and* $0 \le t \le 1$.

Proof From the joint distribution of $\varepsilon_j^{(n)}(t)$ it is easy to obtain

$$E\left(\left(\bar{\varepsilon}_j^{(n)}(t)\right)^2\right) = \frac{\lfloor nt \rfloor}{n} - \left(\frac{\lfloor nt \rfloor}{n}\right)^2 \le 1 \quad \text{for } j = 1, \ldots, n$$

and

$$\begin{aligned} &E\left(\bar{\varepsilon}_j^{(n)}(t)\bar{\varepsilon}_k^{(n)}(t)\right) \\ &= \frac{\lfloor nt \rfloor}{n}\frac{\lfloor nt \rfloor - 1}{n-1} - \left(\frac{\lfloor nt \rfloor}{n}\right)^2 = -\frac{\lfloor nt \rfloor(n - \lfloor nt \rfloor)}{n^2(n-1)} \quad \text{for } 1 \le j < k \le n. \end{aligned}$$

This implies

$$\operatorname{Var}\left(\sum_{j=K+1}^{n-K} x_{j,n}\bar{\varepsilon}_j^{(n)}(t)\right) \le 3 \sum_{j=K+1}^{n-K} x_{j,n}^2$$

and therefore the Markov inequality together with (1.9) yields the statement of the lemma. □

Lemma 2.7 *If* (1.9) *holds, then for every* n *and each* $0 < t < 1$ *there exist independent, identically Bernoulli distributed random variables* $\delta_j^{(n)}(t)$, $j = 1, \ldots, n$ *with* $P(\delta_j^{(n)}(t) = 1) = \lfloor nt \rfloor/n$, *such that*

$$P\left(\sum_{j=1}^{K} x_{j,n}\left(\varepsilon_j^{(n)}(t) - \delta_j^{(n)}(t)\right) \ne 0\right) \le 30\frac{K^2}{n} \tag{2.14}$$

and

$$P\left(\sum_{j=n-K+1}^{n} x_{j,n}\left(\varepsilon_j^{(n)}(t) - \delta_j^{(n)}(t)\right) \ne 0\right) \le 30\frac{K^2}{n} \tag{2.15}$$

for all $K = 1, \ldots, \lfloor n/4 \rfloor$.

Proof We define for $j \geq 1$ the random variables $\gamma_{2j-1} = \varepsilon_j^{(n)}(t)$ and $\gamma_{2j} = \varepsilon_{n-j+1}^{(n)}(t)$. For the approximation of the dependent $\varepsilon_j^{(n)}(t)$ with independent $\delta_j^{(n)}(t)$, we use again Lemma 2.2 and thus we need to estimate the differences

$$|P(\gamma_k = a_k | \gamma_1 = a_1, \ldots, \gamma_{k-1} = a_{k-1}) - P(\gamma_k = a_k)|.$$

Clearly

$$P(\gamma_{k+1} = 1) = \frac{\lfloor nt \rfloor}{n}$$

and

$$P(\gamma_{k+1} = 1 | \gamma_1 = a_1, \ldots, \gamma_k = a_k) = \frac{\lfloor nt \rfloor - a^{(k)}}{n-k}$$

with $a_i \in \{0, 1\}$ for $1 \leq i \leq k$ and $a^{(k)} = \sum_{i=1}^k a_i$ $(\leq \lfloor nt \rfloor)$. By $K \leq \lfloor n/4 \rfloor$ we have $k \leq n/2$, which implies

$$\left| \frac{\lfloor nt \rfloor - a^{(k)}}{n-k} - t \right| \leq \frac{k+1}{n-k} \leq \frac{4k}{n}.$$

Since $|\lfloor nt \rfloor / n - t| \leq 1/n$, we conclude for $a_{k+1} = 1$

$$|P(\gamma_{k+1} = a_{k+1} | \gamma_1 = a_1, \ldots, \gamma_k = a_k) - P(\gamma_{k+1} = a_{k+1})| \leq \frac{5k}{n}$$

and consequently the same is true for $a_{k+1} = 0$. Hence

$$\rho\left(\mathrm{dist}(\gamma_{k+1} | \gamma_k, \ldots, \gamma_1), \mathrm{dist}(\gamma_{k+1})\right) \leq \frac{5k}{n}$$

and thus Lemma 2.2 yields the existence of independent Bernoulli random variables $\delta_j^{(n)}(t)$, $j = 1, \ldots, n$, with $P(\delta_j^{(n)}(t) = 1) = \lfloor nt \rfloor / n$ such that

$$P\left(\left| \varepsilon_j^{(n)}(t) - \delta_j^{(n)}(t) \right| \geq \frac{30k}{n} \right) \leq \frac{30k}{n}$$

and

$$P\left(\left| \varepsilon_{n-j+1}^{(n)}(t) - \delta_{n-j+1}^{(n)}(t) \right| \geq \frac{30k}{n} \right) \leq \frac{30k}{n}.$$

The variables $\varepsilon_j^{(n)}(t)$ and $\delta_j^{(n)}(t)$ only take values in $\{0, 1\}$, hence the statement of the lemma follows. $\square$

As in the case of drawing with replacement, we will now formulate a generalization of Lemma 2.7 for the finite dimensional distributions of $\varepsilon_j^{(n)}(t)$.

Lemma 2.8 *If* (1.9) *holds, then for every n and all $0 < t_1 < \ldots < t_d < 1$ there exist independent, identically distributed random vectors* $\left(\delta_j^{(n)}(t_1), \ldots, \delta_j^{(n)}(t_d)\right)$, *$j = 1, \ldots, n$, such that*

$$\left(\delta_j^{(n)}(t_1), \ldots, \delta_j^{(n)}(t_d)\right) \stackrel{d}{=} \left(I\left\{U \leq \lfloor nt_1 \rfloor / n\right\}, \ldots, I\left\{U \leq \lfloor nt_d \rfloor / n\right\}\right),$$

where U is a uniform random variable on $[0, 1]$ *and furthermore*

$$P\left(\max_{1 \leq l \leq d} \left| \sum_{j=1}^{K} x_{j,n} \left(\varepsilon_j^{(n)}(t_l) - \delta_j^{(n)}(t_l)\right) \right| \geq \delta \right) \to 0$$

and

$$P\left(\max_{1 \leq l \leq d} \left| \sum_{j=n-K+1}^{n} x_{j,n} \left(\varepsilon_j^{(n)}(t_l) - \delta_j^{(n)}(t_l)\right) \right| \geq \delta \right) \to 0$$

are satisfied for all $\delta > 0$, $K \geq 1$ and $n \to \infty$.

Proof As in Sect. 2.1, the application of the same procedure as in Lemma 2.7 to the random vector $\left(\varepsilon_j^{(n)}(t_1), \ldots, \varepsilon_j^{(n)}(t_d)\right)$ instead of $\varepsilon_j^{(n)}(t)$ will prove Lemma 2.8. □

Let $\bar{\delta}_j^{(n)}(t) = \delta_j^{(n)}(t) - E\delta_j^{(n)}$. The next lemma is the analogue of Lemma 2.6 for the independent random variables $\bar{\delta}_j^{(n)}(t)$.

Lemma 2.9 *If* (1.9) *holds, then*

$$\lim_{K \to \infty} \limsup_{n \to \infty} P\left(\left| \sum_{j=K+1}^{n-K} x_{j,n} \bar{\delta}_j^{(n)}(t) \right| \geq \delta \right) = 0$$

for all $\delta > 0$ and $0 \leq t \leq 1$.

Proof The proof can be given as the one to Lemma 2.6. □

Proof of Theorem 1.1 The proof of Theorem 1.1 can be given along the lines of the proof of Theorem 1.2, where Lemmas 2.6–2.9 replace Lemmas 2.1 and 2.3–2.5 and one uses simple modifications due to the different distributions of $\delta_j^{(n)}(t)$ and $\delta_j(t)$. □

The following lemma yields the tightness in Theorem 1.1.

Lemma 2.10 *If* (1.8) *and* (1.9) *hold and* $t_1, t_2 \in [0, 1]$, *then*

$$E\,(Z_n(s) - Z_n(t_1))^2\,(Z_n(t_2) - Z_n(s))^2 \le C(t_2 - t_1)^2 \quad \textit{for all } t_1 \le s \le t_2 \tag{2.16}$$

and $Z_n(t)$ *is tight.*

Proof As in the proof of inequality (2.13), we use $\sum_{j=1}^{n} x_{j,n}^2 \le c$ for all n and consequently $\sum_{j=1}^{n} x_{j,n}^4 \le c^2$. With these inequalities, $\sum_{j=1}^{n} x_{j,n} = 0$ and using $l_1 := \lfloor ns \rfloor - \lfloor nt_1 \rfloor$ and $l_2 := \lfloor nt_2 \rfloor - \lfloor ns \rfloor$ we get

$$\begin{aligned}
&E\,(Z_n(s) - Z_n(t_1))^2\,(Z_n(t_2) - Z_n(s))^2 \\
&\quad = E\left(\sum_{j=\lfloor nt_1 \rfloor+1}^{\lfloor ns \rfloor} X_j^{(n)}\right)^2 \left(\sum_{j=\lfloor ns \rfloor+1}^{\lfloor nt_2 \rfloor} X_j^{(n)}\right)^2 \\
&\quad = \sum_{i,j=\lfloor nt_1 \rfloor+1}^{\lfloor ns \rfloor} \sum_{k,l=\lfloor ns \rfloor+1}^{\lfloor nt_2 \rfloor} E\left(X_i^{(n)} X_j^{(n)} X_k^{(n)} X_l^{(n)}\right) \\
&\quad = l_1 l_2\, E\left((X_1^{(n)})^2 (X_2^{(n)})^2\right) \\
&\qquad + [l_1 l_2 (l_2 - 1) + l_1 (l_1 - 1) l_2]\, E\left((X_1^{(n)})^2 X_2^{(n)} X_3^{(n)}\right) \\
&\qquad + l_1 (l_1 - 1) l_2 (l_2 - 1)\, E\left(X_1^{(n)} X_2^{(n)} X_3^{(n)} X_4^{(n)}\right).
\end{aligned}$$

Now

$$n(n-1)\,E\left((X_1^{(n)})^2 (X_2^{(n)})^2\right) = \sum_{\substack{i,j=1 \\ i \ne j}}^{n} x_{i,n}^2 x_{j,n}^2 \le \left(\sum_{i=1}^{n} x_{i,n}^2\right)^2 \le c^2,$$

$$\begin{aligned}
n(n-1)(n-2)\left|E\left((X_1^{(n)})^2 X_2^{(n)} X_3^{(n)}\right)\right| &= \left|\sum_{\substack{i,j,k=1 \\ i \ne j \ne k}}^{n} x_{i,n}^2 x_{j,n} x_{k,n}\right| \\
&= \left|\sum_{\substack{i,j=1 \\ i \ne j}}^{n} x_{i,n}^2 x_{j,n} (-x_{i,n} - x_{j,n})\right| \\
&= \left|-\sum_{i=1}^{n} x_{i,n}^3 (-x_{i,n}) - \sum_{i=1}^{n} x_{i,n}^2 \sum_{\substack{j=1 \\ j \ne i}}^{n} x_{j,n}^2\right| \le 2c^2,
\end{aligned}$$

$$
\begin{aligned}
&n(n-1)(n-2)(n-3)\left|E\left(X_1^{(n)}X_2^{(n)}X_3^{(n)}X_4^{(n)}\right)\right| \\
&= \left|\sum_{\substack{i,j,k,l=1\\ i\neq j\neq k\neq l}}^{n} x_{i,n}x_{j,n}x_{k,n}x_{l,n}\right| \\
&= \left|\sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{n} x_{i,n}x_{j,n}x_{k,n}(-x_{i,n}-x_{j,n}-x_{k,n})\right| \\
&= \left|3\left(-\sum_{\substack{i,j=1\\ i\neq j}}^{n} x_{i,n}^2 x_{j,n}(-x_{i,n}-x_{j,n})\right)\right| \le 6c^2.
\end{aligned}
$$

The estimates above imply for $n \ge 6$

$$
\begin{aligned}
&E\left(Z_n(s)-Z_n(t_1)\right)^2\left(Z_n(t_2)-Z_n(s)\right)^2 \\
&\le c^2\frac{l_1l_2}{n(n-1)} + 2c^2\frac{l_1l_2(l_1+l_2-2)}{n(n-1)(n-2)} + 6c^2\frac{l_1(l_1-1)l_2(l_2-1)}{n(n-1)(n-2)(n-3)} \\
&\le 2c^2\frac{l_1l_2}{n^2} + 8c^2\frac{l_1l_2(l_1+l_2)}{n^3} + 48c^2\frac{l_1^2l_2^2}{n^4} \le 30c^2\left(\frac{l_1+l_2}{n}\right)^2.
\end{aligned}
$$

As in the proof of (2.13), the left hand side of (2.16) is equal to 0 if $t_2 - t_1 < \frac{1}{n}$. If $t_2 - t_1 \ge \frac{1}{n}$, we obtain

$$
\frac{l_1+l_2}{n} \le \frac{nt_2 - nt_1 + 1}{n} \le 2(t_2 - t_1)
$$

and therefore (2.16) is shown. By applying the Markov inequality and Theorem 15.6 (p. 128) by Billingsley [6] the proof of the lemma is completed. □

2.3 Proofs of the corollaries

Proof of Corollaries 1.1 and 1.2 Let $X_1, X_2, \dots$ be i.i.d. random variables such that, letting $S_n = \sum_{k=1}^{n} X_k$, we have

$$
(S_n - a_n)/b_n \xrightarrow{d} \xi_\alpha \tag{2.17}
$$

for some numerical sequences (a_n), (b_n) and an α-stable r.v. ξ_α. By the classical theory (see e.g. [12] or [11], Chapter XVII) the sequences (a_n), (b_n) can be chosen to satisfy

$$nb_n^{-2}\int_{-b_n}^{b_n} x^2 dF(x) \to C, \quad a_n = n\int_{-b_n}^{b_n} x dF(x) \tag{2.18}$$

for some $C > 0$. As a consequence, we have

$$a_n/nb_n \to 0 \quad \text{and} \quad a_n^2/nb_n^2 \to 0. \tag{2.19}$$

Clearly, the first relation of (2.19) follows from the second one and the second relation follows from (2.18) upon observing that $(\int_{-t}^{t} x dF(x))^2 = o(\int_{-t}^{t} x^2 dF(x))$ for $t \to \infty$ for any distribution F with infinite second moment. Actually, as is shown in Feller [11, Chapter XVII], in the case $\alpha \neq 1$ we can choose $a_n = 0$ (after centering the X_j at expectations for $\alpha > 1$ which is no restriction of generality), and thus (2.19) becomes trivial, but in the case $\alpha = 1$ the centering factor a_n can be nonlinear and the situation is more delicate.

Let $T_n = \max_{1\le k\le n} |X_k|$, $\overline{X}_n = n^{-1}\sum_{k=1}^n X_k$. Formally, Corollaries 1.1 and 1.2 are obtained by applying Theorems 1.1 and 1.2 for the random set

$$\mathcal{H}_n = \left\{ \frac{X_{1,n} - \overline{X}_n}{T_n}, \ldots, \frac{X_{n,n} - \overline{X}_n}{T_n} \right\}, \tag{2.20}$$

where $X_{1,n} \le X_{2,n} \le \cdots \le X_{n,n}$ is the ordered sample of $(X_1, \ldots, X_n)$. However, to avoid certain technical difficulties in verifying conditions (1.8) and (1.9) for $\mathcal{H}_n$, we will use an indirect approach by first applying Theorems 1.1 and 1.2 for a perturbed version of $\mathcal{H}_n$. As in the proof of Theorems 1.1 and 1.2, we will assume again $m = n$; the general case requires only trivial changes. Let $Z_{n,m}^*(t)$ in Corollaries 1.1 and 1.2 be denoted in this case by $Z_n^*(t)$.

Write

$$\frac{\overline{X}_n}{T_n} = \frac{b_n}{T_n}\left(\frac{S_n - a_n}{nb_n} + \frac{a_n}{nb_n}\right). \tag{2.21}$$

It is known (see e.g. [7]) that both b_n/T_n and T_n/b_n are bounded in probability and thus (2.17), (2.19) and (2.21) imply $\bar{X}_n/T_n \xrightarrow{P} 0$. Fix now an integer $L \ge 1$. By Csörgő et al. [9, p. 109], we have

$$\begin{aligned}
&\frac{1}{b_n}(X_{j,n} : j \in [1, L] \cup [n - L + 1, n]) \\
&\quad \xrightarrow{d} \left(-\frac{q^{1/\alpha}}{Z_1^{1/\alpha}}, \ldots, -\frac{q^{1/\alpha}}{Z_L^{1/\alpha}}, \frac{p^{1/\alpha}}{(Z_L^*)^{1/\alpha}}, \ldots, \frac{p^{1/\alpha}}{(Z_1^*)^{1/\alpha}} \right)
\end{aligned}$$

where $Z_j = \eta_1 + \cdots + \eta_j$, $Z_j^* = \eta_1^* + \cdots + \eta_j^*$ and (η_j) and (η_j^*) are i.i.d. sequences of exp(1) random variables, independent also of each other. The last convergence relation remains valid if we divide both sides by their maximum norm, i.e. with T_n/b_n resp. M, where M is the random variable defined by (1.18). Thus using $\bar{X}_n/T_n \xrightarrow{P} 0$ we get that

$$\left(\frac{X_{j,n}-\overline{X}_n}{T_n} : j \in [1, L] \cup [n-L+1, n]\right) \xrightarrow{d} \frac{1}{M}\left(-\frac{q^{1/\alpha}}{Z_1^{1/\alpha}}, \ldots, -\frac{q^{1/\alpha}}{Z_L^{1/\alpha}}, \frac{p^{1/\alpha}}{(Z_L^*)^{1/\alpha}}, \ldots, \frac{p^{1/\alpha}}{(Z_1^*)^{1/\alpha}}\right). \quad (2.22)$$

Next we show that

$$\lim_{K\to\infty} \limsup_{n\to\infty} P\left(T_n^{-2} \sum_{j=K+1}^{n-K} (X_{j,n}-\overline{X}_n)^2 \geq \varepsilon\right) = 0 \quad \text{for any } \varepsilon > 0. \quad (2.23)$$

We write

$$\begin{aligned}
\sum_{j=K+1}^{n-K} (X_{j,n}-\overline{X}_n)^2 &= \sum_{j=K+1}^{n-K} \left[(X_{j,n}-a_n/n) - (\overline{X}_n - a_n/n)\right]^2 \\
&\leq \sum_{j=K+1}^{n-K} (X_{j,n}-a_n/n)^2 + 2|\overline{X}_n - a_n/n| \left|\sum_{j=K+1}^{n-K} (X_{j,n}-a_n/n)\right| \\
&\quad + n|\overline{X}_n - a_n/n|^2 \\
&=: I_1 + I_2 + I_3.
\end{aligned}$$

Clearly,

$$b_n^{-2} I_3 = \frac{1}{n}\left(\frac{S_n - a_n}{b_n}\right)^2 = o_P(1)$$

by (2.17). To estimate I_1 and I_2, we note that for any fixed $K \geq 1$,

$$\frac{1}{b_n}\left(\sum_{j=K+1}^{n-K} X_{j,n} - a_n\right) \xrightarrow{d} A_1(K), \quad \frac{1}{b_n^2} \sum_{j=K+1}^{n-K} X_{j,n}^2 \xrightarrow{d} A_2(K) \quad (2.24)$$

where $A_1(K) \to 0$ a.s., $A_2(K) \to 0$ a.s. for $K \to \infty$. (See Csörgő et al. [9].) Now by (2.19)

$$b_n^{-2} I_2 = \frac{2}{n}\left|\frac{S_n - a_n}{b_n}\right| \left|\frac{\sum_{j=K+1}^{n-K} X_{j,n} - a_n}{b_n} + o(1)\right|,$$

and

$$
\begin{aligned}
b_n^{-2} I_1 &= \frac{1}{b_n^2} \sum_{j=K+1}^{n-K} X_{j,n}^2 - \frac{2a_n}{nb_n^2} \left(\sum_{j=K+1}^{n-K} X_{j,n} - a_n \right) - \left(\frac{2K}{n} + 1 \right) \frac{a_n^2}{nb_n^2} \\
&:= I_4 + I_5 + I_6.
\end{aligned}
$$

By (2.19) we have $I_6 = o(1)$ for any fixed K, while (2.24) and (2.19) take care of the estimate of I_4 and I_5 and thus (2.23) is proved.

The just proved relation (2.23) implies that for every $\varepsilon > 0$ there exist positive integers $L = L(\varepsilon)$, $n_0 = n_0(\varepsilon)$ and for any $n \geq n_0$ a set $A_n = A_n(\varepsilon)$ in the underlying probability space such that $P(A_n) \geq 1 - \varepsilon$ and

$$
T_n^{-2} \sum_{j=L+1}^{n-L} (X_{j,n} - \overline{X}_n)^2 \leq \varepsilon \quad \text{on } A_n \text{ for } n \geq n_0. \tag{2.25}
$$

By Jensen's inequality, the last relation implies that

$$
\left| T_n^{-1} \sum_{j=L+1}^{n-L} (X_{j,n} - \overline{X}_n) \right| \leq (\varepsilon n)^{1/2} \quad \text{on } A_n \text{ for } n \geq n_0. \tag{2.26}
$$

Let $\mathcal{H}_{n,L}$ denote the set obtained from $\mathcal{H}_n$ in (2.20) by replacing all elements with indices $L+1 \leq j \leq n-L$ by 0. Since the sum of the elements of $\mathcal{H}_n$ is 0, (2.26) shows that the average $\eta_{n,L}$ of the set $\mathcal{H}_{n,L}$ satisfies $|\eta_{n,L}| \leq (\varepsilon/n)^{1/2}$ and consequently $n\eta_{n,L}^2 \leq \varepsilon$ on A_n. Hence if $\mathcal{H}_{n,L}^*$ denotes the set obtained by subtracting $\eta_{n,L}$ from each element of $\mathcal{H}_{n,L}$, the Euclidean distance of the vectors $\mathcal{H}_{n,L}$ and $\mathcal{H}_{n,L}^*$ is at most $\sqrt{\varepsilon}$. By (2.25) the Euclidean distance of $\mathcal{H}_n$ and $\mathcal{H}_{n,L}$ on A_n is also at most $\sqrt{\varepsilon}$. It follows that the distance of $\mathcal{H}_n$ and $\mathcal{H}_{n,L}^*$ is at most $2\sqrt{\varepsilon}$ and thus Corollary 2.1 yields for any $0 \leq t \leq 1$

$$
E_{\mathbf{X}}(Z_n^*(t) - Z_{n,L}^*(t))^2 \leq 4\varepsilon \quad \text{on } A_n \text{ for } n \geq n_0, \tag{2.27}
$$

where $Z_n^*(t)$ and $Z_{n,L}^*(t)$ $(0 \leq t \leq 1)$ denote the partial sum processes of random elements sampled (with or without replacement) from $\mathcal{H}_n$ resp. $\mathcal{H}_{n,L}^*$. Note also that $Z_j \sim j$ a.s. and $Z_j^* \sim j$ a.s. by the strong law of large numbers and thus letting $R_L(t)$ denote the analogue of $R(t)$ in Corollaries 1.1 and 1.2 when the infinite sums are replaced by their Lth partial sums, we have by $0 < \alpha < 2$

$$
E_{\mathbf{Z}}(R_L(t) - R(t))^2 = \mathcal{O}\left(\sum_{j=L+1}^{\infty} (Z_j^{-2/\alpha} + (Z_j^*)^{-2/\alpha}) \right) = o_L(1) \quad \text{a.s.} \tag{2.28}
$$

By (2.27), (2.28) and the Markov inequality, the Lévy distance $\mathcal{L}_{\mathbf{X}}$ of the conditional distributions of $Z_n^*(t)$ and $Z_{n,L}^*(t)$ relative to $\mathbf{X}$ is $\leq 2\varepsilon^{1/3}$ on A_n for $n \geq n_0$ and the

Lévy distance $\mathcal{L}_{\mathbf{Z}}$ of the conditional distributions of $R(t)$ and $R_L(t)$ relative to $\mathbf{Z}$ is $o_L(1)$ a.s. for any fixed $0 \le t \le 1$. Since $\varepsilon > 0$ was arbitrary, it suffices to prove that given any integer $L \ge 1$, Corollaries 1.1 and 1.2 hold with $Z_n^*(t)$ replaced by $Z_{n,L}^*(t)$ and $R(t)$ replaced by $R_L(t)$.

Fix $L \ge 1$. By (2.22) and the Skorokhod–Dudley–Wichura theorem (see e.g. [19]), on a suitable probability space one can redefine the vectors $\mathbf{X}^{(n)} = (X_{1,n}, \dots, X_{n,n})$, $n = 1, 2, \dots$, without changing their distributions, together with a sequence $\eta_1, \eta_1^*, \eta_2, \eta_2^* \dots$ of independent exp(1) random variables such that the convergence in distribution in relation (2.22) can be replaced by a.s. convergence. Define $\mathbf{X}$ after the redefinition as $\mathbf{X} = (\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots)$. Clearly, this redefinition does not change the distribution of the r.v. on the left hand side of (1.17). Since after the redefinition the first and last L elements of $\mathcal{H}_n$ converge a.s. to finite limits, the average $\eta_{n,L}$ of $\mathcal{H}_{n,L}$ satisfies $|\eta_{n,L}| = O(1/n)$ a.s. Thus the set $\mathcal{H}_{n,L}^*$ satisfies conditions (1.7) and (1.8) with probability 1, with limiting numbers

$$y_j = -M^{-1} q^{1/\alpha} Z_j^{-1/\alpha}, \quad z_j = M^{-1} p^{1/\alpha} (Z_j^*)^{-1/\alpha}, \quad j = 1, \dots, L,$$

where M is defined by (1.18) and all the other y_j, z_j are equal to 0. Also, since the elements of the set $\mathcal{H}_{n,L}^*$ with indices $L + 1 \le j \le n - L$ are equal to $-\eta_{n,L}$, the sum of the squares of these elements is at most $n\eta_{n,L}^2 = o(1)$ a.s. and thus condition (1.9) is also satisfied a.s. Applying Theorems 1.1 and 1.2 for the set $\mathcal{H}_{n,L}^*$ completes the proof. □

Proof of Corollary 1.3 Let $X_{1,n} \le X_{2,n} \le \cdots \le X_{n,n}$ be the order statistics of the sample $(X_1, \dots, X_n)$ and let $M_{1,n}, M_{2,n}, \dots, M_{n,n}$ be the elements of the sample $(X_1, \dots, X_n)$ arranged in decreasing order of absolute value. Since $1 - F(t) + F(-t)$ is slowly varying, Theorem 2.3 in Maller and Resnick [17] implies that

$$M_{2,n}/M_{1,n} \xrightarrow{P} 0.$$

Consequently, for any fixed integer $K \ge 1$ and $n \to \infty$ we have

$$\frac{1}{M_{1,n}} \left(X_{1,n}, \dots, X_{K,n}, X_{n-K+1,n}, \dots, X_{n,n}\right) \xrightarrow{d} (I\{U > p\}, 0, \dots, 0, I\{U \le p\})$$

where U is a uniform random variable on the unit interval (0, 1) and I denotes indicator function. The last relation implies

$$\left(\frac{X_{1,n}}{T_n}, \frac{X_{2,n}}{T_n}, \dots, \frac{X_{K,n}}{T_n}, \frac{X_{n-K+1,n}}{T_n}, \dots, \frac{X_{n,n}}{T_n}\right) \xrightarrow{d} (-I\{U > p\}, 0, \dots, 0, I\{U \le p\}). \tag{2.29}$$

By the Skorokhod-Dudley-Wichura theorem the vectors in (2.29) can be redefined such that we get almost sure convergence. By Theorem 2.3 in Maller and Resnick [17] the partial sums $S_n = \sum_{k=1}^n X_k$ satisfy $S_n/M_{1,n} \xrightarrow{P} 1$ and consequently $\bar{X}_n/T_n \xrightarrow{P} 0$. The proof can now be completed as in the case of Corollary 1.1. □

Proof of Corollary 1.4 Let Ψ be the function defined before the formulation of Corollary 1.4 and let $F^{-1}(x) = \sup\{t : F(t) \le x\}$ be the inverse of the distribution function F of X_1. Clearly

$$F^{-1}(x) = -2^{k-1} = -x^{-1}\Psi(x) \quad \text{for } x \in [2^{-k}, 2^{-(k-1)}), \quad k \ge 2$$

and by the symmetry of the distribution of X_1

$$F^{-1}(x) = (1-x)^{-1}\Psi(1-x) \quad \text{for } 1/2 < x < 1$$

at points of continuity of F. We also have $\Psi(2^{-k}x) = \Psi(x)$ for all $k \in \mathbb{Z}$. Let, as in the formulation of the Corollary, $Z_j = \eta_1 + \cdots + \eta_j$, $Z_j^* = \eta_1^* + \cdots + \eta_j^*$, where $\eta_1, \eta_2, \ldots$ and $\eta_1^*, \eta_2^*, \ldots$ are independent sequences of i.i.d. exp(1) random variables and put

$$X_{j,n}^* = F^{-1}\left(\frac{Z_j}{Z_{n+1}}\right), \quad 1 \le j \le n.$$

Let further $X_{1,n} \le \ldots \le X_{n,n}$ be the ordered sample of $X_1, \ldots, X_n$. As is well known (see e.g. [16]), the vectors $(X_{1,n}, \ldots, X_{n,n})$ and $(X_{1,n}^*, \ldots, X_{n,n}^*)$ have the same distribution. Now

$$X_{1,n}^* = F^{-1}\left(\frac{Z_1}{Z_{n+1}}\right) = -\frac{Z_{n+1}}{Z_1}\Psi\left(\frac{Z_1}{Z_{n+1}}\right)$$

provided $Z_1/Z_{n+1} < 1/2$ which holds for any fixed ω if $n \ge n_0(\omega)$. By the strong law of large numbers $Z_{n+1}/n \to 1$ a.s. whence it follows

$$X_{1,n}^* = -\frac{n}{Z_1}\Psi\left(\frac{Z_1}{n}\right)(1+o(1)) \quad \text{a.s.} \tag{2.30}$$

and similarly, for every fixed ω for $n \ge n_0(\omega)$

$$\begin{aligned} X_{n,n}^* = F^{-1}\left(\frac{Z_n}{Z_{n+1}}\right) &= \left(1 - \frac{Z_n}{Z_{n+1}}\right)^{-1}\Psi\left(1 - \frac{Z_n}{Z_{n+1}}\right) = \frac{Z_{n+1}}{\eta_{n+1}}\Psi\left(\frac{\eta_{n+1}}{Z_{n+1}}\right) \\ &= \frac{n}{\eta_{n+1}}\Psi\left(\frac{\eta_{n+1}}{n}\right)(1+o(1)) \quad \text{a.s.} \end{aligned} \tag{2.31}$$

Since $Z_1 = \eta_1$ and η_{n+1} are independent, relations (2.30) and (2.31) show that $-X_{1,n}^*/n$ and $X_{n,n}^*/n$ are asymptotically independent and equidistributed; the same holds for the vectors $-\frac{1}{n}(X_{1,n}^*, \ldots, X_{K,n}^*)$ and $\frac{1}{n}(X_{n-K+1,n}^*, \ldots, X_{n,n}^*)$ for any fixed $K \ge 1$ as $n \to \infty$.

Now along the subsequence $n = n_k = \lfloor c\, 2^k \rfloor$ for some constant $1 \le c < 2$ we obtain, using (2.30) and $\Psi(2^{-k}x) = \Psi(x)$,

$$\frac{X^*_{1,n_k}}{n_k} \to -\frac{1}{Z_1}\Psi\left(\frac{Z_1}{c}\right) \quad \text{a.s.}$$

Similar formulas apply for $X^*_{j,n_k}/n_k$ for any fixed $j \ge 1$ and also for the variables $X^*_{n_k-j+1,n_k}/n_k$ on the other end for any fixed j; in the latter case, Z_j should be replaced by Z^*_j and a.s. convergence by convergence in distribution. Thus letting $T_n = \max_{1\le j\le n}|X_j|$ and using the asymptotic independence of the ordered sample elements at the two ends of the sample, we get for any fixed $K \ge 1$

$$\begin{aligned}&\frac{1}{T_{n_k}}(X_{1,n_k}, \ldots, X_{K,n_k}, X_{n_k-K+1,n_k}, \ldots, X_{n_k,n_k})\\ &\quad \xrightarrow{d} \frac{1}{M}\left(-\frac{\Psi\left(\frac{Z_1}{c}\right)}{Z_1}, \ldots, -\frac{\Psi\left(\frac{Z_K}{c}\right)}{Z_K}, \frac{\Psi\left(\frac{Z^*_K}{c}\right)}{Z^*_K}, \ldots, \frac{\Psi\left(\frac{Z^*_1}{c}\right)}{Z^*_1}\right), \qquad (2.32)\end{aligned}$$

where

$$M = \max\left\{\frac{\Psi\left(Z_1/c\right)}{Z_1}, \frac{\Psi\left(Z^*_1/c\right)}{Z^*_1}\right\}.$$

Observing also that $\bar{X}_n/T_n \xrightarrow{P} 0$, the proof can be completed exactly as in the case of Corollary 1.1. □

Acknowledgment We would like to thank three anonymous referees for their valuable comments on the paper, leading to a considerable improvement of the presentation.

References

1. Athreya, K.: Bootstrap of the mean in the infinite variance case. Ann. Stat. **15**, 724–731 (1987)
2. Aue, A., Berkes, I., Horváth, L.: Selection from a stable box. Bernoulli **14**, 125–139 (2008)
3. Berkes, I., Philipp, W.: Approximation theorems for independent and weakly dependent random vectors. Ann. Probab. **7**, 29–54 (1979)
4. Bergström, H.: The limit problem for sums of independent random variables which are not uniformly asymptotically negligible. Transactions of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes. Technical University of Prague, pp. 125–135. Academia, Prague (1971)
5. Bickel, P., Freedman, D.: Some asymptotic theory for the bootstrap. Ann. Stat. **9**, 1196–1217 (1981)
6. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
7. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation. Cambridge University Press, Cambridge (1987)
8. Breiman, L.: Probability. Addison-Wesley, Reading (1968)
9. Csörgő, M., Csörgő, S., Horváth, L., Mason, D.M.: Normal and stable convergence of integral functions of the empirical distribution function. Ann. Probab. **14**, 86–118 (1986)
10. Csörgő, S.: Rates of merge in generalized St. Petersburg games. Acta Sci. Math. (Szeged) **68**, 815–847 (2002)

11. Feller, W.: An Introduction to Probability Theory and its Applications, vol. 2. Wiley, New York (1966)
12. Gnedenko, B.V., Kolmogorov, A.N.: Limit Distributions of Sums of Independent Random Variables. Addison-Wesley, Reading (1968)
13. Hall, P.: Asymptotic properties of the bootstrap for heavy-tailed distributions. Ann. Probab. **18**, 1342–1360 (1990)
14. Le Cam, L.: On the distribution of sums of independent random variables. In: Neyman, J., Le Cam, L. (eds.), Bernoulli, Bayes, Laplace: Anniversary Volume, pp. 179–202. Springer, Heidelberg (1965)
15. Ledoux, M., Talagrand, M.: Probability in Banach Spaces. Springer, New York (1991)
16. LePage, R., Woodroofe, M., Zinn, J.: Convergence to a stable distribution via order statistics. Ann. Probab. **9**, 624–632 (1981)
17. Maller, R., Resnick, S.: Limiting behaviour of sums and the term of maximum modulus. Proc. Lond. Math. Soc. **49**, 385–422 (1984)
18. Rosén, B.: Limit theorems for sampling from finite populations. Ark. Mat. **5**, 383–424 (1965)
19. Shorack, G., Wellner, J.: Empirical Processes with Applications to Statistics. Wiley, New York (1986)
20. Skorokhod, A.V.: Limit theorems for stochastic processes. Theory Probab. Appl. **1**, 261–290 (1956)