mathematics

MDPI

Article

Convergence in Total Variation of Random Sums

Luca Pratelli [1] and Pietro Rigo [2,*]

[1] Accademia Navale, viale Italia 72, 57100 Livorno, Italy; pratel@mail.dm.unipi.it
[2] Dipartimento di Scienze Statistiche "P. Fortunati", Università di Bologna, via delle Belle Arti 41, 40126 Bologna, Italy
* Correspondence: pietro.rigo@unibo.it

Abstract: Let (X_n) be a sequence of real random variables, (T_n) a sequence of random indices, and (τ_n) a sequence of constants such that $\tau_n \to \infty$. The asymptotic behavior of $L_n = (1/\tau_n) \sum_{i=1}^{T_n} X_i$, as $n \to \infty$, is investigated when (X_n) is exchangeable and independent of (T_n). We give conditions for $M_n = \sqrt{\tau_n}\,(L_n - L) \longrightarrow M$ in distribution, where L and M are suitable random variables. Moreover, when (X_n) is i.i.d., we find constants a_n and b_n such that $\sup_{A \in \mathcal{B}(\mathbb{R})} |P(L_n \in A) - P(L \in A)| \leq a_n$ and $\sup_{A \in \mathcal{B}(\mathbb{R})} |P(M_n \in A) - P(M \in A)| \leq b_n$ for every n. In particular, $L_n \to L$ or $M_n \to M$ in total variation distance provided $a_n \to 0$ or $b_n \to 0$, as it happens in some situations.

Keywords: exchangeability; random sum; rate of convergence; stable convergence; total variation distance

MSC: 60F05; 60G50; 60B10; 60G09

1. Introduction

All random elements appearing in this paper are defined on the same probability space, say $(\Omega, \mathcal{A}, P)$.

A *random sum* is a quantity such as $\sum_{i=1}^{T_n} X_i$, where $(X_n : n \geq 1)$ is a sequence of real random variables and $(T_n : n \geq 1)$ a sequence of $\mathbb{N}$-valued random indices. In the sequel, in addition to (X_n) and (T_n), we fix a sequence $(\tau_n : n \geq 1)$ of positive constants such that $\tau_n \to \infty$ and we let

$$L_n = \frac{\sum_{i=1}^{T_n} X_i}{\tau_n}.$$

Random sums find applications in a number of frameworks, including statistical inference, risk theory and insurance, reliability theory, economics, finance, and forecasting of market changes. Accordingly, the asymptotic behavior of L_n, as $n \to \infty$, is a classical topic in probability theory. The related literature is huge and we do not try to summarize it here. We just mention a general text book [1] and some useful recent references: [2–10].

In this paper, the asymptotic behavior of L_n is investigated in the (important) special case where (X_n) is exchangeable and independent of (T_n). More precisely, we assume that:

(i) (X_n) is exchangeable;

(ii) (X_n) is independent of (T_n);

(iii) $\frac{T_n}{\tau_n} \xrightarrow{P} V$ for some random variable $V > 0$.

Under such conditions, we prove a weak law of large numbers (WLLN), a central limit theorem (CLT), and we investigate the rate of convergence with respect to the total variation distance.

check for updates

Citation: Pratelli, L.; Rigo, P. Convergence in Total Variation of Random Sums. *Mathematics* **2021**, *9*, 194. https://doi.org/10.3390/math9020194

Received: 21 December 2020
Accepted: 15 January 2021
Published: 19 January 2021

Suppose in fact $E|X_1| < \infty$ and conditions (i)-(ii)-(iii) hold. Define

$$L = V\,E(X_1 \mid \mathcal{T}) \quad \text{and} \quad M_n = \sqrt{\tau_n}\,(L_n - L),$$

where V is the random variable involved in condition (iii) and $\mathcal{T}$ the tail σ-field of (X_n). Then, it is not hard to show that $L_n \xrightarrow{P} L$. To obtain a CLT, instead, is not straightforward. In Section 3, we prove that $M_n \to M$ in distribution, where M is a suitable random variable, provided $E(X_1^2) < \infty$ and $\sqrt{\tau_n}\left\{\frac{T_n}{\tau_n} - V\right\}$ converges stably. Finally, in Section 4, assuming (X_n) i.i.d. and some additional conditions, we find constants a_n and b_n such that

$$\sup_{A \in \mathcal{B}(\mathbb{R})} |P(L_n \in A) - P(L \in A)| \leq a_n \qquad \text{and}$$
$$\sup_{A \in \mathcal{B}(\mathbb{R})} |P(M_n \in A) - P(M \in A)| \leq b_n \quad \text{for every } n \geq 1.$$

In particular, $L_n \to L$ or $M_n \to M$ in *total variation distance* provided $a_n \to 0$ or $b_n \to 0$, as it happens in some situations.

A last note is that, to our knowledge, random sums have been rarely investigated when (X_n) is exchangeable. Similarly, convergence of L_n or M_n in total variation distance is usually not taken into account. This paper contributes to fill this gap.

2. Preliminaries

In the sequel, the probability distribution of any random element U is denoted by $\mathcal{L}(U)$. If S is a topological space, $\mathcal{B}(S)$ is the Borel σ-field on S and $C_b(S)$ the space of real bounded continuous functions on S. The total variation distance between two probability measures on $\mathcal{B}(S)$, say μ and ν, is

$$d_{TV}(\mu, \nu) = \sup_{A \in \mathcal{B}(S)} |\mu(A) - \nu(A)|.$$

With a slight abuse of notation, if X and Y are S-valued random variables, we write $d_{TV}(X, Y)$ instead of $d_{TV}\big[\mathcal{L}(X),\, \mathcal{L}(Y)\big]$, namely

$$d_{TV}(X, Y) = \sup_{A \in \mathcal{B}(S)} |P(X \in A) - P(Y \in A)|.$$

If X is a real random variable, we say that $\mathcal{L}(X)$ is *absolutely continuous* to mean that $\mathcal{L}(X)$ is absolutely continuous with respect to Lebesgue measure. The following technical fact is useful in Section 4.

Lemma 1. *Let X be a strictly positive random variable. Then,*

$$\lim_n d_{TV}\Big(X + q_n\sqrt{X},\, X\Big) = 0$$

provided the q_n are constants such that $q_n \to 0$ and $\mathcal{L}(X)$ is absolutely continuous.

Proof. Let f be a density of X. Since $\lim_n \int_{-\infty}^{\infty} |f_n(x) - f(x)|\,dx = 0$, for some sequence f_n of continuous densities, it can be assumed that f is continuous. Furthermore, since $X > 0$, for each $\epsilon > 0$ there is $b > 0$ such that $P(X < b) < \epsilon$. For such a b, one obtains

$$d_{TV}\Big(X + q_n\sqrt{X},\, X\Big) \leq \epsilon + \sup_{A \in \mathcal{B}(\mathbb{R})} \Big| P\big(X + q_n\sqrt{X} \in A \mid X \geq b\big) - P\big(X \in A \mid X \geq b\big) \Big|.$$

Hence, it can be also assumed $X \geq b$ a.s. for some $b > 0$.

Let g_n be a density of $X + q_n\sqrt{X}$. Since

$$d_{TV}\Big(X + q_n\sqrt{X},\, X\Big) = \int_{-\infty}^{\infty} \big[f(x) - g_n(x)\big]^+ dx = \int_b^{\infty} \big[f(x) - g_n(x)\big]^+ dx,$$

it suffices to show that $f(x) = \lim_n g_n(x)$ for each $x > b$. To prove the latter fact, define $\phi_n(x) = x + q_n\sqrt{x}$. For large n, one obtains $4\,q_n^2 < b$. In this case, $\phi_n' > 0$ on (b, ∞) and g_n can be written as

$$g_n(x) = f\big[\phi_n^{-1}(x)\big]\,\frac{2\sqrt{\phi_n^{-1}(x)}}{q_n + 2\sqrt{\phi_n^{-1}(x)}}.$$

Therefore, $f(x) = \lim_n g_n(x)$ follows from the continuity of f and

$$\phi_n^{-1}(x) = x + \frac{q_n^2}{2} - \frac{q_n}{2}\sqrt{q_n^2 + 4x} \longrightarrow x.$$

□

2.1. Stable Convergence

Stable convergence, introduced by Renyi in [11], is a strong form of convergence in distribution. It actually occurs in a number of frameworks, including the classical CLT, and thus it quickly became popular; see, e.g., [12] and references therein. Here, we just recall the basic definition.

Let S be a metric space, (Y_n) a sequence of S-valued random variables, and K a *kernel* (or a *random probability measure*) on S. The latter is a map K on Ω such that $K(\omega)$ is a probability measure on $\mathcal{B}(S)$, for each $\omega \in \Omega$, and $\omega \mapsto K(\omega)(B)$ is $\mathcal{A}$-measurable for each $B \in \mathcal{B}(S)$. Say that Y_n *converges stably to* K if

$$\lim_n E\big[f(Y_n) \mid H\big] = E\big[K(\cdot)(f) \mid H\big], \tag{1}$$

for all $f \in C_b(S)$ and $H \in \mathcal{A}$ with $P(H) > 0$, where $K(\cdot)(f) = \int f(x)\,K(\cdot)(dx)$.

More generally, take a sub-σ-field $\mathcal{G} \subset \mathcal{A}$ and suppose K is $\mathcal{G}$-measurable (i.e., $\omega \mapsto K(\omega)(B)$ is $\mathcal{G}$-measurable for fixed $B \in \mathcal{B}(S)$). Then, Y_n *converges* $\mathcal{G}$*-stably to* K if condition (1) holds whenever $H \in \mathcal{G}$ and $P(H) > 0$.

An important special case is when K is a trivial kernel, in the sense that

$$K(\omega) = \nu \qquad \text{for all } \omega \in \Omega$$

where ν is a fixed probability measure on $\mathcal{B}(S)$. In this case, Y_n converges $\mathcal{G}$-stably to ν if and only if

$$\lim_n E\big\{G\,f(Y_n)\big\} = E(G)\int f\,d\nu$$

whenever $f \in C_b(S)$ and $G : \Omega \to \mathbb{R}$ is bounded and $\mathcal{G}$-measurable.

3. WLLN and CLT for Random Sums

In this section, we still let

$$L_n = \frac{\sum_{i=1}^{T_n} X_i}{\tau_n}, \quad L = V\,E(X_1 \mid \mathcal{T}) \quad \text{and} \quad M_n = \sqrt{\tau_n}\,(L_n - L),$$

where V is the random variable involved in condition (iii) and

$$\mathcal{T} = \bigcap_n \sigma(X_n, X_{n+1}, \ldots)$$

is the tail σ-field of (X_n). Recall that $V > 0$. Recall also that, by de Finetti's theorem, (X_n) is exchangeable if and only if is i.i.d. conditionally on $\mathcal{T}$, namely

$$P\big(X_1 \in A_1, \dots, X_n \in A_n \mid \mathcal{T}\big) = \prod_{i=1}^{n} P\big(X_1 \in A_i \mid \mathcal{T}\big) \qquad \text{a.s.}$$

for all $n \geq 1$ and all $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R})$.

The following WLLN is straightforward.

Theorem 1. *If $E|X_1| < \infty$ and conditions (i) and (iii) hold, then $L_n \xrightarrow{P} L$.*

Proof. Recall that, if Y_n and Y are any real random variables, $Y_n \xrightarrow{P} Y$ if and only if, for each subsequence (n'), there is a sub-subsequence $(n'') \subset (n')$ such that $Y_{n''} \xrightarrow{a.s.} Y$. Fix a subsequence (n'). Then, by (iii),

$$\frac{T_{n''}}{\tau_{n''}} \xrightarrow{a.s.} V$$

along a suitable sub-subsequence $(n'') \subset (n')$. Since $V > 0$, then $T_{n''} \xrightarrow{a.s.} \infty$. As a result of the SLLN for exchangeable sequences, $(1/n) \sum_{i=1}^{n} X_i \xrightarrow{a.s.} E(X_1 \mid \mathcal{T})$. Therefore,

$$L_{n''} = \frac{T_{n''}}{\tau_{n''}} \frac{\sum_{i=1}^{n''} X_i}{T_{n''}} \xrightarrow{a.s.} V\, E(X_1 \mid \mathcal{T}) = L.$$

□

For definiteness, Theorem 1 has been stated in terms of convergence in probability, but other analogous results are available. As an example, suppose that $E|X_1| < \infty$ and conditions (i)–(ii) are satisfied. Then, $L_n \to L$ in distribution provided $\frac{T_n}{\tau_n} \to V$ in distribution. This follows from Skorohod representation theorem and the current version of Theorem 1. Similarly, $L_n \xrightarrow{a.s.} L$ or $L_n \xrightarrow{L_1} L$ whenever $\frac{T_n}{\tau_n} \xrightarrow{a.s.} V$ or $\frac{T_n}{\tau_n} \xrightarrow{L_1} V$.

We also note that, as implicit in the proof of Theorem 1, condition (iii) implies $T_n \xrightarrow{P} \infty$ or equivalently

$$\lim_n P(T_n \leq c) = 0 \qquad \text{for every fixed } c > 0.$$

We next turn to the CLT. It is convenient to begin with the i.i.d. case. From now on, U and Z are two real random variables such that

$$Z \sim \mathcal{N}(0,1), \quad U \text{ is independent of } Z \quad \text{and} \\ (U, Z) \text{ is independent of } (X_n, T_n : n \geq 1). \tag{2}$$

We also let

$$a = E(X_1) \quad \text{and} \quad \sigma^2 = \text{var}(X_1).$$

Theorem 2. *Suppose (X_n) is i.i.d., $E(X_1^2) < \infty$, condition (ii) holds, and*

$$\sqrt{\tau_n} \left\{ \frac{T_n}{\tau_n} - V \right\} \quad \text{converges stably to } \mathcal{L}(U). \tag{3}$$

Then,

$$M_n \longrightarrow \sigma \sqrt{V}\, Z + a\, U \ \text{ in distribution.}$$

Proof. Let

$$W_n = a\,\sqrt{\tau_n}\left\{\frac{T_n}{\tau_n} - V\right\} + \sqrt{\frac{V}{T_n}}\,\sum_{i=1}^{T_n}(X_i - a).$$

Since (X_n) is i.i.d., $E(X_1 \mid \mathcal{T}) = E(X_1) = a$ a.s. Since $E\Big\{\Big(\frac{\sum_{i=1}^{T_n}(X_i-a)}{\sqrt{T_n}}\Big)^2\Big\} = \sigma^2$ for every n, the sequence $\frac{\sum_{i=1}^{T_n}(X_i-a)}{\sqrt{T_n}}$ is L_2-bounded, and this implies

$$W_n - M_n = W_n - \sqrt{\tau_n}\,(L_n - aV) = \frac{\sum_{i=1}^{T_n}(X_i - a)}{\sqrt{T_n}}\,\left(\sqrt{V} - \sqrt{\frac{T_n}{\tau_n}}\right) \xrightarrow{P} 0.$$

Therefore, it suffices to prove $W_n \longrightarrow \sigma\sqrt{V}\,Z + a\,U$ in distribution. We prove the latter fact by means of characteristic functions.

Fix $t \in \mathbb{R}$. Let $\mu_{n,j}(\cdot) = P(V \in \cdot \mid T_n = j)$ be the probability distribution of V under $P(\cdot \mid T_n = j)$ and

$$\phi_j(s) = E\left\{\exp\left(i\,s\,\frac{\sum_{i=1}^{j}(X_i - a)}{\sqrt{j}}\right)\right\} \qquad \text{for all } s \in \mathbb{R}.$$

Then,

$$E\{\exp(i\,t\,W_n)\} = \sum_{j=1}^{\infty} P(T_n = j)\int \exp\left(i\,t\,a\,\sqrt{\tau_n}\left(\frac{j}{\tau_n} - v\right)\right)\phi_j(\sqrt{v}\,t)\,\mu_{n,j}(dv).$$

In addition, for each $c > 0$, the classical CLT yields

$$\lim_{j\to\infty}\sup_{0<v\le c}\left|\phi_j(\sqrt{v}\,t) - \exp\Big(-\frac{t^2\sigma^2 v}{2}\Big)\right| = 0. \tag{4}$$

Since condition (3) implies condition (iii), $\lim_n P(T_n \le b) = 0$ for all $b > 0$. Given $\epsilon > 0$, take $c > 0$ such that $P(V > c) < \epsilon$. As a result of (4), one can find an integer m such that

$$\begin{aligned}\Big|\,E\{\exp(i\,t\,W_n)\} - E\Big\{\exp\Big(i\,t\,a\,\sqrt{\tau_n}\Big(\frac{T_n}{\tau_n} - V\Big)\Big)\,\exp\Big(-\frac{t^2\sigma^2 V}{2}\Big)\Big\}\Big| \le \\ \le \epsilon + 2P(T_n \le m) + 2P(V > c) < 3\,\epsilon + 2P(T_n \le m).\end{aligned}$$

Since ϵ is arbitrary and $\lim_n P(T_n \le m) = 0$, it follows that

$$\limsup_n \Big|\,E\{\exp(i\,t\,W_n)\} - E\Big\{\exp\Big(i\,t\,a\,\sqrt{\tau_n}\Big(\frac{T_n}{\tau_n} - V\Big)\Big)\,\exp\Big(-\frac{t^2\sigma^2 V}{2}\Big)\Big\}\Big| = 0.$$

Finally, since $Z \sim \mathcal{N}(0,1)$ and Z is independent of V,

$$E\{\exp(i\,t\,\sigma\sqrt{V}\,Z)\} = E\Big\{\exp\Big(-\frac{t^2\sigma^2 V}{2}\Big)\Big\}.$$

Therefore,

$$
\begin{aligned}
E\{\exp(i\,t\,\sigma\,\sqrt{V}\,Z + i\,t\,a\,U)\} &= E\Big\{\exp\Big(-\frac{t^2\sigma^2 V}{2}\Big)\Big\}\,E\{\exp(i\,t\,a\,U)\} \\
&= \lim_n E\Big\{\exp\Big(-\frac{t^2\sigma^2 V}{2}\Big)\ \exp\Big(i\,t\,a\,\sqrt{\tau_n}\,\Big(\frac{T_n}{\tau_n} - V\Big)\Big)\Big\} \\
&= \lim_n E\{\exp(i\,t\,W_n)\}
\end{aligned}
$$

where the second equality is due to condition (3). Hence, $W_n \longrightarrow \sigma\,\sqrt{V}\,Z + a\,U$ in distribution, and this concludes the proof. □

The argument used in the proof of Theorem 2 yields a little bit more. Let $\nu = \mathcal{L}(\sigma\,\sqrt{V}\,Z + a\,U)$ and $\mathcal{G} = \sigma(V, X_1, X_2, \ldots)$. Then, M_n converges $\mathcal{G}$-stably (and not only in distribution) to ν. Among other things, since $L_n \xrightarrow{P} L$, this implies that $(L_n, M_n) \to (L, R)$ in distribution, where R denotes a random variable independent of L such that $R \sim \nu$. Moreover, condition (3) can be weakened into $\sqrt{\tau_n}\,\Big\{\frac{T_n}{\tau_n} - V\Big\}$ converges $\sigma(V)$-stably to $\mathcal{L}(U)$.

We also note that, under some extra assumptions, Theorem 2 could be given a simpler proof based on some version of Anscombe's theorem; see, e.g., [13] and references therein.

Finally, we adapt Theorem 2 to the exchangeable case. Let

$$W = E(X_1^2 \mid \mathcal{T}) - E(X_1 \mid \mathcal{T})^2 \quad \text{and} \quad M = \sqrt{W\,V}\,Z + U\,E(X_1 \mid \mathcal{T}).$$

To introduce the next result, it may be useful to recall that

$$\sqrt{n}\,\Big\{\frac{\sum_{i=1}^n X_i}{n} - E(X_1 \mid \mathcal{T})\Big\} \longrightarrow \mathcal{N}(0,\,W) \quad \text{stably}$$

provided (X_n) is exchangeable and $E(X_1^2) < \infty$, where $\mathcal{N}(0,\,W)$ is the Gaussian kernel with mean 0 and random variance W (with $\mathcal{N}(0,0) = \delta_0$); see, e.g., ([14] Th. 3.1).

Theorem 3. *If $E(X_1^2) < \infty$ and conditions (i)–(ii) and (3) hold, then $M_n \to M$ in distribution.*

Proof. Just note that (X_n) is i.i.d. conditionally on $\mathcal{T}$, with mean $E(X_1 \mid \mathcal{T})$ and variance W. Hence, for each $f \in C_b(\mathbb{R})$, Theorem 2 yields

$$E\{f(M_n) \mid \mathcal{T}\} \xrightarrow{a.s.} E\{f(M) \mid \mathcal{T}\},$$

which in turn implies

$$E\{f(M)\} = E\Big\{\lim_n E\{f(M_n) \mid \mathcal{T}\}\Big\} = \lim_n\, E\Big\{E\{f(M_n) \mid \mathcal{T}\}\Big\} = \lim_n E\{f(M_n)\}.$$

□

4. Rate of Convergence with Respect to Total Variation Distance

To obtain upper bounds for $d_{TV}(L_n, L)$ and $d_{TV}(M_n, M)$, some additional assumptions are needed. In particular, in this section, (X_n) is i.i.d. (with the exception of Remark 1). Hence, L and M reduce to $L = a\,V$ and $M = \sigma\,\sqrt{V}\,Z + a\,U$, where $a = E(X_1)$, $\sigma^2 = \text{var}(X_1)$ and (U, Z) satisfies condition (2).

We begin with a rough estimate for $d_{TV}(L_n, L)$.

Theorem 4. *Suppose that conditions (ii)–(iii) hold, (X_n) is i.i.d., $E(|X_1|^3) < \infty$ and $\mathcal{L}(X_1)$ has an absolutely continuous part. Then,*

$$d_{TV}(L_n, L) \leq P(T_n \leq m) + \frac{c}{\sqrt{m+1}} + d_{TV}\left(L + \sigma \sqrt{\frac{V}{\tau_n}}\, Z,\, L\right) + \\ + E\left[\frac{|\sqrt{V} - \sqrt{\frac{T_n}{\tau_n}}|}{\max(\sqrt{V},\, \sqrt{\frac{T_n}{\tau_n}})}\right] + \frac{|a|\,\sqrt{\tau_n}}{\sigma}\, E\left[\frac{|V - \frac{T_n}{\tau_n}|}{\max(\sqrt{V},\, \sqrt{\frac{T_n}{\tau_n}})}\right]$$

for all $m,\, n \geq 1$, where $c > 0$ is a constant independent of m and n.

In order to prove Theorem 4, we recall that

$$d_{TV}\Big(\mathcal{N}(a_1, b_1),\, \mathcal{N}(a_2, b_2)\Big) \;\leq\; \frac{|\sqrt{b_1} - \sqrt{b_2}| \,+\, |a_1 - a_2|}{\sqrt{\max(b_1, b_2)}} \tag{5}$$

for all $a_1,\, a_2 \in \mathbb{R}$ and $b_1,\, b_2 > 0$; see, e.g., ([15] Lem. 3).

Proof of Theorem 4. Fix $m,\, n \geq 1$. By ([16] Lem. 2.1), up to enlarging the underlying probability space $(\Omega, \mathcal{A}, P)$, there is a sequence $((S_j, Z_j) : j \geq 1)$ of random variables, independent of (T_n, V), such that

$$S_j \sim \sum_{i=1}^{j} X_i, \quad Z_j \sim \mathcal{N}(0,1), \quad P(S_j \neq a\, j + \sigma\, \sqrt{j}\, Z_j) = d_{TV}\Big(S_j,\, a\, j + \sigma\, \sqrt{j}\, Z_j\Big).$$

In addition, by ([17] Th. 2.6), there is a constant $c > 0$ depending only on $E(|X_1|^3)$ such that

$$d_{TV}\Big(S_j,\, a\, j + \sigma\, \sqrt{j}\, Z_j\Big) = d_{TV}\left(\frac{S_j - a\, j}{\sigma\, \sqrt{j}}\,,\, Z_j\right) \leq \frac{c}{\sqrt{m+1}} \qquad \text{for all } j > m.$$

Having noted these facts, define

$$L_n^* = \frac{a\, T_n + \sigma\, \sqrt{T_n}\, Z_{T_n}}{\tau_n}.$$

Then,

$$\begin{aligned} d_{TV}(L_n, L_n^*) &\leq P(T_n \leq m) + \sum_{j>m} P(T_n = j)\, d_{TV}\Big[P(L_n \in \cdot \mid T_n = j)\,, P(L_n^* \in \cdot \mid T_n = j)\Big] \\ &\leq P(T_n \leq m) + \sup_{j>m}\, d_{TV}\Big[P(L_n \in \cdot \mid T_n = j)\,, P(L_n^* \in \cdot \mid T_n = j)\Big] \\ &= P(T_n \leq m) + \sup_{j>m}\, d_{TV}\left[\frac{\sum_{i=1}^{j} X_i}{\tau_n}\,,\, \frac{a\, j + \sigma\, \sqrt{j}\, Z_j}{\tau_n}\right] \\ &= P(T_n \leq m) + \sup_{j>m}\, d_{TV}\Big(S_j,\, a\, j + \sigma\, \sqrt{j}\, Z_j\Big) \\ &\leq P(T_n \leq m) + \frac{c}{\sqrt{m+1}}. \end{aligned}$$

Next, since $Z_{T_n} \sim \mathcal{N}(0,1)$, by conditioning on (L_n, V) and applying inequality (5), one obtains

$$d_{TV}\left(L_n^*,\, aV + \sigma \sqrt{\frac{V}{\tau_n}}\, Z_{T_n}\right) \leq E\left[\frac{|\sqrt{V} - \sqrt{\frac{T_n}{\tau_n}}|}{\max(\sqrt{V},\, \sqrt{\frac{T_n}{\tau_n}})}\right] + \frac{|a|\,\sqrt{\tau_n}}{\sigma}\; E\left[\frac{|V - \frac{T_n}{\tau_n}|}{\max(\sqrt{V},\, \sqrt{\frac{T_n}{\tau_n}})}\right].$$

Moreover, since $Z_{T_n} \sim Z$ and both Z_{T_n} and Z are independent of V,

$$d_{TV}\left(aV + \sigma \sqrt{\frac{V}{\tau_n}}\, Z_{T_n},\, L\right) = d_{TV}\left(L + \sigma \sqrt{\frac{V}{\tau_n}}\, Z,\, L\right).$$

Collecting all these facts together, one finally obtains

$$\begin{aligned} d_{TV}(L_n, L) &\leq d_{TV}(L_n, L_n^*) + d_{TV}(L_n^*, L) \\ &\leq P(T_n \leq m) + \frac{c}{\sqrt{m+1}} + d_{TV}\left(L + \sigma \sqrt{\frac{V}{\tau_n}}\, Z,\, L\right) + \\ &+ E\left[\frac{|\sqrt{V} - \sqrt{\frac{T_n}{\tau_n}}|}{\max(\sqrt{V},\, \sqrt{\frac{T_n}{\tau_n}})}\right] + \frac{|a|\,\sqrt{\tau_n}}{\sigma}\; E\left[\frac{|V - \frac{T_n}{\tau_n}|}{\max(\sqrt{V},\, \sqrt{\frac{T_n}{\tau_n}})}\right]. \end{aligned}$$

□

The upper bound provided by Theorem 4 is generally large but it becomes manageable under some further assumptions. For instance, if $V \geq b$ a.s. for some constant $b > 0$, it reduces to

$$\begin{aligned} d_{TV}(L_n, L) \leq P(T_n \leq m) + \frac{c}{\sqrt{m+1}} + d_{TV}\left(L + \sigma \sqrt{\frac{V}{\tau_n}}\, Z,\, L\right) + \\ + \left(\frac{1}{b} + \frac{|a|\,\sqrt{\tau_n}}{\sigma\,\sqrt{b}}\right) E\left[\left|V - \frac{T_n}{\tau_n}\right|\right]. \end{aligned} \tag{6}$$

As an example, we discuss a simple but instructive case.

Example 1. *For each $x \in \mathbb{R}$, denote by $J(x)$ the integer part of x. Suppose $V \geq b$ a.s. for some constant $b > 0$ and define*

$$T_n = J(\tau_n\, V + 1).$$

Suppose also that (X_n) is independent of V and satisfies the other conditions of Theorem 4. Then,

$$T_n > \tau_n\, b \quad \textit{and} \quad \left|V - \frac{T_n}{\tau_n}\right| = \frac{T_n}{\tau_n} - V \leq \frac{1}{\tau_n} \quad \textit{a.s.}$$

Hence, letting $m = J(\tau_n\, b)$, inequality (6) reduces to

$$d_{TV}(L_n, L) \leq \frac{c^*}{\sqrt{\tau_n}} + d_{TV}\left(L + \sigma \sqrt{\frac{V}{\tau_n}}\, Z,\, L\right)$$

for some constant c^*. Finally, $d_{TV}\left(L + \sigma \sqrt{\frac{V}{\tau_n}}\, Z,\, L\right) = O(1/\sqrt{\tau_n})$ if V is bounded above and $\mathcal{L}(V)$ is absolutely continuous with a Lipschitz density. Hence, under the latter condition on V, one obtains

$$d_{TV}(L_n, L) = O(1/\sqrt{\tau_n}).$$

Incidentally, this bound is essentially of the same order as the bound obtained in [6] when T_n has a mixed Poisson distribution and the total variation distance is replaced by the Wasserstein distance.

One more consequence of Theorem 4 is the following.

Corollary 1. *$L_n \to L$ in total variation distance provided the conditions of Theorem 4 hold, $a \neq 0$, $\mathcal{L}(V)$ is absolutely continuous, and*

$$\lim_n \sqrt{\tau_n}\, E\left[\left|V - \frac{T_n}{\tau_n}\right|\right] = 0.$$

Proof. First, assume $V \geq b$ a.s. for some constant $b > 0$. For each $z \in \mathbb{R}$, letting $q_n = \frac{\sigma}{a\sqrt{\tau_n}}\, z$, Lemma 1 implies

$$\limsup_n d_{TV}\left(L + \sigma\sqrt{\frac{V}{\tau_n}}\, z,\, L\right) = \limsup_n d_{TV}\left(V + q_n\sqrt{V},\, V\right) = 0.$$

Conditioning on Z and taking inequality (6) into account, it follows that

$$\begin{aligned}
\limsup_n d_{TV}(L_n, L) &\leq \frac{c}{\sqrt{m+1}} + \limsup_n d_{TV}\left(L + \sigma\sqrt{\frac{V}{\tau_n}}\, Z,\, L\right) \\
&\leq \frac{c}{\sqrt{m+1}} + \limsup_n \int d_{TV}\left(L + \sigma\sqrt{\frac{V}{\tau_n}}\, z,\, L\right) \mathcal{N}(0,1)(dz) \\
&= \frac{c}{\sqrt{m+1}} \qquad \text{for each } m \geq 1.
\end{aligned}$$

This concludes the proof if $V \geq b$ a.s. In general, for each $b > 0$, define

$$\begin{aligned}
V_b &= 1_{\{V>b\}}\, V + 1_{\{V\leq b\}}\, (V+b) \qquad \text{and} \\
T_{n,b} &= J\Big(1_{\{V>b\}}\, T_n + 1_{\{V\leq b\}}\, (1 + \tau_n\,(V+b))\Big)
\end{aligned}$$

where $J(x)$ denotes the integer part of x. Since $\frac{T_{n,b}}{\tau_n} \xrightarrow{P} V_b > b$, the first part of the proof implies

$$\frac{\sum_{i=1}^{T_{n,b}} X_i}{\tau_n} \longrightarrow a\, V_b \qquad \text{in total variation distance.}$$

Finally, since $V > 0$ and

$$d_{TV}(L_n, L) \leq 2\, P(V \leq b) + d_{TV}\left(\frac{\sum_{i=1}^{T_{n,b}} X_i}{\tau_n},\, a\, V_b\right) \quad \text{for all } b > 0,$$

one obtains $\lim_n d_{TV}(L_n, L) = 0$. □

We next turn to $d_{TV}(M_n, M)$. Following [18], our strategy is to estimate $d_{TV}(M_n, M)$ through the Wasserstein distance between $\mathcal{L}(M_n)$ and $\mathcal{L}(M)$.

Recall that, if X and Y are real integrable random variables, the Wasserstein distance between $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ is

$$d_W(X,Y) = \inf_{(H,K)} E|H-K| = \sup_f |E(f(X)) - E(f(Y))|,$$

where inf is over the real random variables H and K such that $H \sim X$ and $K \sim Y$ while sup is over the 1-Lipschitz functions $f : \mathbb{R} \to \mathbb{R}$. Define also

$$l_n = \int |t\,\phi_n(t)|\,dt = 2\int_0^\infty t\,|\phi_n(t)|\,dt$$

where ϕ_n is the characteristic function of M_n.

Theorem 5. *Assume the conditions of Theorem 2 and:*

(iv) $U = \sqrt{V_0}\,Z_0$, *where* $Z_0 \sim \mathcal{N}(0,1)$, $V_0 \geq 0$ *is independent of* Z_0, *and* (V_0, Z_0) *is independent of* (V, Z);

(v) $E(T_{n_0}^2) < \infty$ *for some* n_0 *and*

$$\sup_n \tau_n\, E\Big\{\Big(\frac{T_n}{\tau_n} - V\Big)^2\Big\} < \infty.$$

Then, $d_W(M_n, M) \to 0$. *Moreover, letting* $d_n = d_W(M_n, M)$, *one obtains*

$$d_{TV}(M_n, M) \leq d_n^{1/2} + d_n^{1/2-\alpha} + P\big(\sqrt{\sigma^2 V + a^2 V_0} < d_n^{\alpha}\big) + k\left(l_n\, d_n^{1/2}\right)^{2/3}$$

$$\text{and} \qquad d_{TV}(M_n, M) \leq d_n^{1/2}\left(1 + \frac{1}{\sigma}\,E(V^{-1/2})\right) + k\left(l_n\, d_n^{1/2}\right)^{2/3}$$

for each $n \geq 1$ *and* $\alpha < 1/2$, *where* k *is a constant independent of* n.

Proof. By Theorem 2, $M_n \to M$ in distribution. By condition (iv),

$$M = \sigma\sqrt{V}\,Z + a\sqrt{V_0}\,Z_0 \sim \sqrt{\sigma^2 V + a^2 V_0}\;Z,$$

so that $\mathcal{L}(M)$ is a mixture of centered Gaussian laws. On noting that

$$E\Big\{\Big(\sum_{i=1}^{T_n}(X_i - a)\Big)^2\Big\} = \sigma^2 E(T_n),$$

one obtains

$$\begin{aligned} E(M_n^2) &= \tau_n\, E\left\{\left(\frac{\sum_{i=1}^{T_n}(X_i - a)}{\tau_n} + a\left(\frac{T_n}{\tau_n} - V\right)\right)^2\right\} \\ &\leq \frac{2}{\tau_n}\, E\Big\{\Big(\sum_{i=1}^{T_n}(X_i - a)\Big)^2\Big\} + 2\,a^2\tau_n\, E\left\{\left(\frac{T_n}{\tau_n} - V\right)^2\right\} \\ &= 2\,\sigma^2 E\left(\frac{T_n}{\tau_n}\right) + 2\,a^2\tau_n\, E\left\{\left(\frac{T_n}{\tau_n} - V\right)^2\right\}. \end{aligned}$$

Finally, by condition (v), $\lim_n E\Big(\frac{T_n}{\tau_n}\Big) = E(V) < \infty$ and $\sup_n E(M_n^2) < \infty$. To conclude the proof, it suffices to apply Theorem 1 of [18] (see also the subsequent remark) with $\beta = 2$. □

Theorem 5 gives two upper bounds for $d_{TV}(M_n, M)$ in terms of $d_n = d_W(M_n, M)$ and l_n. To avoid trivialities, suppose $\sigma > 0$. Obviously, the second bound makes sense only if $E(V^{-1/2}) < \infty$. However, since $V > 0$ and $d_n \to 0$, the first bound implies $d_{TV}(M_n, M) \to 0$ if $\lim_n l_n d_n^{1/2} = 0$. In particular, $d_{TV}(M_n, M) \to 0$ if $\sup_n l_n < \infty$.

Example 2. *Under the conditions of Theorem 5, suppose also that* $\mathcal{L}(X_1)$ *is absolutely continuous with a density* f *satisfying* $\int |f'(x)|\,dx < \infty$. *Then, conditioning on* T_n *and* V *and arguing as* *in ([18] Ex. 2), it can be shown that* $\sup_n l_n < \infty$. *Hence,* $M_n \to M$ *in total variation distance. Furthermore, if* $E(V^{-1/2}) < \infty$, *the second bound of Theorem 5 yields*

$$d_{TV}(M_n, M) \leq k^* \left(1 \wedge d_n\right)^{1/3}$$

for all $n \geq 1$ *and a suitable constant* k^* *(independent of* n*).*

We close the paper by briefly discussing the exchangeable case.

Remark 1. *Usually, the upper bounds for the total variation distance are preserved under mixtures. Hence, by conditioning on* $\mathcal{T}$ *and making some further assumptions, the results obtained in this section can be extended to the case where* (X_n) *is exchangeable. As an example, define* L *and* M *as in Section 3 and suppose*

$$\left| E\Big\{\exp(i\,t\,X_1) \mid \mathcal{T}\Big\} \right| \leq \frac{Q}{|t|} \quad a.s.$$

for each $t \in \mathbb{R} \setminus \{0\}$ *and for some integrable random variable* Q. *Then, Corollary 1 and Theorem 5 are still valid even if* (X_n) *is exchangeable (and not necessarily i.i.d.) up to replacing* $a \neq 0$ *with* $E(X_1 \mid \mathcal{T}) \neq 0$ *a.s. in Corollary 1.*

Author Contributions: Methodology, L.P. and P.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

References

1. Gnedenko, B.V.; Korolev, V. *Random Summation: Limit Theorems and Applications*; CRC Press: Boca Raton, FL, USA, 1996.
2. Kiche, J.; Ngesa, O.; Orwa, G. On generalized gamma distribution and its application to survival data. *Int. J. Stat. Probab.* **2019**, *8*, 85–102.
3. Korolev, V.; Chertok, A.; Korchagin, A.; Zeifman, A. Modeling high-frequency order flow imbalance by functional limit theorems for two-sided risk processes. *Appl. Math. Comput.* **2015**, *253*, 224–241. [CrossRef]
4. Korolev, V.; Dorofeeva, A. Bounds of the accuracy of the normal approximation to the distributions of random sums under relaxed moment conditions. *Lith. Math. J.* **2017**, *57*, 38–58. [CrossRef]
5. Korolev, V.; Zeifman, A. Generalized negative binomial distributions as mixed geometric laws and related limit theorems. *Lith. Math. J.* **2019**, *59*, 366–388. [CrossRef]
6. Korolev, V.; Zeifman, A. Bounds for convergence rate in laws of large numbers for mixed Poisson random sums. *Stat. Prob. Lett.* **2021**, *168*, 1–8. [CrossRef]
7. Mattner, L.; Shevtsova, I. An optimal Berry-Esseen type theorem for integrals of smooth functions. *ALEA Lat. Am. J. Probab. Math. Stat.* **2019**, *16*, 487–530. [CrossRef]
8. Schluter, C.; Trede, M. Weak convergence to the student and Laplace distributions. *J. Appl. Probab.* **2016**, *53*, 121–129. [CrossRef]
9. Shevtsova, I.; Tselishchev, M. A generalized equilibrium transform with application to error bounds in the Renyi theorem with no support constraints. *Mathematics* **2020**, *8*, 577. [CrossRef]
10. Sheeja, S.; Kumar, S. Negative binomial sum of random variables and modeling financial data. *Int. J. Stat. Appl. Math.* **2017**, *2*, 44–51.
11. Renyi, A. On stable sequences of events. *Sankhya A* **1963**, *25*, 293–302.
12. Nourdin, I.; Nualart, D.; Peccati, G. Quantitative stable limit theorems on the Wiener space. *Ann. Probab.* **2016**, *44*, 1–41. [CrossRef]
13. Berti, P.; Crimaldi, I.; Pratelli, L.; Rigo, P. An Anscombe-type theorem. *J. Math. Sci.* **2014**, *196*, 15–22. [CrossRef]
14. Berti, P.; Pratelli, L.; Rigo, P. Limit theorems for a class of identically distributed random variables. *Ann. Probab.* **2004**, *32*, 2029–2052. [CrossRef]
15. Pratelli, L.; Rigo, P. Total variation bounds for Gaussian functionals. *Stoch. Proc. Appl.* **2019**, *129*, 2231–2248. [CrossRef]
16. Sethuraman J. Some extensions of the Skorohod representation theorem. *Sankhya* **2002**, *64*, 884–893.
17. Bally, V.; Caramellino, L. Asymptotic development for the CLT in total variation distance. *Bernoulli* **2016**, *22*, 2442–2485. [CrossRef]
18. Pratelli, L.; Rigo, P. Convergence in total variation to a mixture of Gaussian laws. *Mathematics* **2018**, *6*, 99. [CrossRef]