Article

On the Accuracy of the Exponential Approximation to Random Sums of Alternating Random Variables

Irina Shevtsova [1,2,3,4,*] **and Mikhail Tselishchev** [2,4,*]

1   Department of Mathematics, School of Science, Hangzhou Dianzi University, 310018 Hangzhou, China
2   Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics,
    Lomonosov Moscow State University, 119991 Moscow, Russia
3   Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences,
    119333 Moscow, Russia
4   Moscow Center for Fundamental and Applied Mathematics, 119991 Moscow, Russia
*   Correspondence: ishevtsova@cs.msu.ru (I.S.); m.tselishchev@cs.msu.ru (M.T.)

check for
updates

Abstract: Using the generalized stationary renewal distribution (also called the equilibrium transform) for arbitrary distributions with a finite non-zero first moment, we prove moment-type error-bounds in the Kantorovich distance for the exponential approximation to random sums of possibly dependent random variables with positive finite expectations, in particular, to geometric random sums, generalizing the previous results to alternating and dependent random summands. We also extend the notions of new better than used in expectation (NBUE) and new worse than used in expectation (NWUE) distributions to alternating random variables in terms of the corresponding distribution functions and provide a criteria in terms of conditional expectations similar to the classical one. As corollary, we provide simplified error-bounds in the case of NBUE/NWUE conditional distributions of random summands.

Keywords: Rényi theorem; Kantorovich distance; zeta-metrics; Stein's method; stationary renewal distribution; equilibrium transform; geometric random sum; characteristic function; NBUE, NWUE distributions

1. Introduction

According to the generalized Rényi theorem, a geometric random sum of independent identically distributed (i.i.d.) nonnegative random variables (r.v.'s), normalized by its mean, converges in distribution to the exponential law when the expectation of the geometric number of summands tends to infinity. Some numerical bounds for the exponential approximation to geometric random sums, as well as their various applications, can be found in the classical monograph of Kalashnikov [1]. Peköz and Röllin [2] developed Stein's method for the exponential distribution and obtained moment-type estimates for the exponential approximation to geometric and non-geometric random sums with non-negative summands completing Kalashnikov's bounds in the Kantorovich distance. Their method was substantially based on the equilibrium transform (stationary renewal distribution) of non-negative random variables, hence yielding the technical restriction on the support of the random summands under consideration. Moreover, Peköz and Röllin considered dependent random summands with constant conditional expectations and presented some error-bounds in this case. The present authors extended Stein's method to alternating (i.e., taking values of both signs) random summands by generalizing the equilibrium transform to distributions with arbitrary support, and obtained moment-type estimates of the accuracy of the exponential approximation for geometric and non-geometric random sums of independent alternating random variables. The same

paper [3] contains a detailed overview of the estimates of the exponential approximation to geometric random sums.

The aim of the present work is to extend the results of [3] to dependent random summands with constant conditional expectations, also generalizing the results of [2] to alternating random summands.

Recall that the Kantorovich distance ζ_1 between probability distributions of r.v.'s X and Y with distribution functions (d.f.'s) F and G is defined as a simple probability metric with ζ-structure (see [1,4]) as

$$\zeta_1\left(F, G\right) \equiv \zeta_1\left(\mathscr{L}(X), \mathscr{L}(Y)\right) \equiv \zeta_1\left(X, Y\right) := \sup_{h \in \mathrm{Lip}_1^\infty} \left| \int_{\mathbb{R}} h \, dF - \int_{\mathbb{R}} h \, dG \right|, \tag{1}$$

where $\mathrm{Lip}_c^\infty = \left\{ h \in \mathrm{Lip}_c \mid h \text{ is bounded} \right\}$ and

$$\mathrm{Lip}_c := \left\{ h \colon \mathbb{R} \to \mathbb{R} \,\middle|\, |h(x) - h(y)| \leq c \, |x - y| \quad \forall x, y \in \mathbb{R} \right\}, \quad c > 0.$$

If both X and Y are integrable, then $\zeta_1\left(X, Y\right) < \infty$ and the supremum in (1) can be taken over a wider class Lip_1 of Lipschitz functions. In this case, according to the Kantorovich–Rubinshtein theorem, ζ_1 allows several alternative representations

$$\zeta_1\left(F, G\right) = \min_{\mathscr{L}(X',Y') \colon X' \overset{d}{=} X, \, Y' \overset{d}{=} Y} \mathbf{E} |X' - Y'| = \int_0^1 \left| F^{-1}(u) - G^{-1}(u) \right| du = \int_{-\infty}^\infty |F(x) - G(x)| \, dx, \tag{2}$$

where F^{-1} and G^{-1} are generalized inverse functions of F and G, respectively.

We will use a *generalized equilibrium transform* that was introduced and studied in [3]. Given a probability distribution of a r.v. X with d.f. F and finite $a := \mathbf{E}X \neq 0$, its equilibrium transform is defined as a (signed) measure $\mathscr{L}^e(X)$ on $(\mathbb{R}, \mathcal{B})$ with the d.f.

$$F^e(x) := \begin{cases} -\dfrac{1}{a} \displaystyle\int_{-\infty}^x F(y) \, dy, & \text{if } x \leq 0, \\[3mm] -\dfrac{\mathbf{E}X^-}{a} + \dfrac{1}{a} \displaystyle\int_0^x (1 - F(y)) \, dy, & \text{if } x > 0. \end{cases} \tag{3}$$

Observe that $\mathscr{L}^e(X)$ is absolutely continuous (a.c.) with respect to (w.r.t.) the Lebesgue measure with the density

$$p^e(x) = \begin{cases} -\frac{1}{a} F(x), & \text{if } x \leq 0, \\[2mm] \frac{1}{a}(1 - F(x)), & \text{if } x > 0. \end{cases} \tag{4}$$

The characteristic function (ch.f.) of $\mathscr{L}^e(X)$ can be expressed in terms of the original ch.f. f of r.v. X as

$$f^e(t) := \int_{\mathbb{R}} e^{itx} dF^e(x) = \frac{f(t) - 1}{t f'(0)} = \frac{f(t) - 1}{ita}, \quad \text{if } t \neq 0, \quad \text{and} \quad f^e(0) = 1. \tag{5}$$

If X is nonnegative or nonpositive almost surely (a.s.), then $\mathscr{L}^e(X)$ is a probability distribution and it is possible to construct a r.v. $X^e \sim \mathscr{L}^e(X)$.

Below, we list some other properties of the equilibrium transform (see ([3], Theorem 1) for details and proofs) which will be used in the present work:

Homogeneity. For any r.v. X with finite $\mathbf{E}X \neq 0$ and d.f. F_X we have

$$(F_{cX})^e(x) = F_X^e(x/c), \quad \text{for all } c \in \mathbb{R} \setminus \{0\}, \, x \in \mathbb{R}. \tag{6}$$

Moments. If $\mathbf{E}|X|^{r+1} < \infty$ for some $r > 0$, then for all $k \in \mathbb{N} \cap [1, r]$ we have

$$\int_{\mathbb{R}} x^k \, dF^e(x) = \frac{\mathbf{E}X^{k+1}}{(k+1)\mathbf{E}X} \; , \quad \int_{\mathbb{R}} |x|^r \, dF^e(x) = \frac{\mathbf{E}X|X|^r}{(r+1)\mathbf{E}X} \; . \tag{7}$$

We will also use the following inequality from ([3], Theorem 3), which states that the Kantorovich distance to the exponential law is no more than twice greater than distance to the equilibrium transform.

Lemma 1. *Let X be a square integrable r.v. with $\mathbf{E}X = 1$ and $\mathscr{E} \sim \mathrm{Exp}(1)$. Then,*

$$\zeta_1(X, \mathscr{E}) \leq 2\,\zeta_1\big(\mathscr{L}(X), \mathscr{L}^e(X)\big), \tag{8}$$

where $\mathscr{L}^e(X)$ is the equilibrium transform of $\mathscr{L}(X)$.

The r.-h.s.'s of (8), in turn, can be bounded from above by the second moment $\mathbf{E}X^2$ in the following way.

Lemma 2 (see Theorem 2 and Remark 2 in [3]). *For any square-integrable r.v. X with $\mathbf{E}X \neq 0$,*

$$\zeta_1\big(\mathscr{L}(X), \mathscr{L}^e(X)\big) \leq \frac{1}{2} \cdot \frac{\mathbf{E}X^2}{|\mathbf{E}X|} - |\mathbf{E}X| \cdot \mathbf{P}\left(X \cdot \mathrm{sign}\,\mathbf{E}X \leq 0\right). \tag{9}$$

Note the presence of the Kantorovich distance between $\mathscr{L}(X)$ and possibly signed measure $\mathscr{L}^e(X)$ on the r.-h.s.'s of (8) and (9), which requires some extra explanation. As described in [3], it is defined in terms of d.f.'s in the same way as for probability measures in (1) and allows an alternative representation as an area between d.f.'s of its arguments (similar to the last expression in (2)). Moreover, this generalization retains the property of the homogeneity of order 1 (see ([3], Lemma 1)). Namely, if F and G are d.f.'s of (signed) Borel measures on $\mathbb{R}$ with $F(+\infty) = G(+\infty)$ and $F_c(x) := F(cx)$, $G_c(x) := G(cx)$, $x \in \mathbb{R}$, $c > 0$, then

$$\zeta_1(F_c, G_c) = \frac{1}{c}\,\zeta_1(F, G). \tag{10}$$

Using the above notation and techniques, we prove moment-type error bounds in the Kantorovich distance for the exponential approximation to random sums of possibly dependent r.v.'s with positive finite expectations (Theorem 1), which generalize the results of [2] to alternating random summands and results of [3] to dependent random summands.

Moreover, we extend the definitions of *new better than used in expectation* (NBUE) and *new worse than used in expectation* (NWUE) distributions to alternating random variables in terms of the corresponding d.f.'s and provide a criteria in terms of conditional expectations similar to the classical one (Theorem 2). Finally, we provide simplified error-bounds in cases of NBUE/NWUE conditional distributions of random summands, generalizing those obtained in [2].

2. Main Results

Lemma 3. *Let $X_1, X_2, \ldots$ be a sequence of random variables, such that for every $n \geq 2$ there exists a regular conditional probability $\mathscr{L}(X_n \mid X_1, \ldots, X_{n-1})$ with the constant conditional expectation $a_n := \mathbf{E}\left(X_n \mid X_1, \ldots, X_{n-1}\right) \in (0, +\infty)$. Let $S_n := \sum_{i=1}^n X_n$ for $n \in \mathbb{N}$, $S_0 := 0$ and N be a $\mathbb{N}_0 \equiv \mathbb{N} \cup \{0\}$-valued r.v., independent of $\{X_1, X_2, \ldots\}$, with*

$$A := \mathbf{E}S_N = \sum_{n=1}^{\infty} a_n \, \mathbf{P}(N \geq n) < +\infty.$$

Then the characteristic function of $\mathscr{L}^e(S_N)$ is

$$f^e_{S_N}(t) = \sum_{m=1}^{\infty} \mathbf{P}(M = m)\, \mathbf{E}\left(e^{itS_{m-1}} \cdot f^e_m(t \mid X_1, \ldots, X_{m-1})\right), \quad t \in \mathbb{R}, \tag{11}$$

where M is an $\mathbb{N}$-valued r.v. with

$$\mathbf{P}(M = m) = \frac{a_m}{A}\mathbf{P}(N \geq m), \quad m \in \mathbb{N},$$

and $f^e_n(t \mid x_1, \ldots, x_{n-1})$ is the characteristic function of the equilibrium transform $\mathscr{L}^e(X_n \mid X_1 = x_1, \ldots, X_{n-1} = x_{n-1})$ of the conditional distribution $\mathscr{L}(X_n \mid X_1 = x_1, \ldots, X_{n-1} = x_{n-1})$. Or, in terms of (conditional) distribution functions,

$$F^e_{S_N}(x) = \sum_{m=1}^{\infty} \mathbf{P}(M = m) \int_{\mathbb{R}^{m-1}} F^e_m(x - x_1 - \ldots - x_{m-1} \mid x_1, \ldots, x_{m-1})\, dF(x_1, \ldots, x_{m-1}), \tag{12}$$

where $F(x_1, \ldots, x_m)$ denotes the joint d.f. of $X_1, \ldots, X_m$ and $F^e_m(x \mid x_1, \ldots, x_{m-1})$ denotes the conditional d.f. of $\mathscr{L}^e(X_m \mid X_1 = x_1, \ldots, X_{m-1} = x_{m-1})$, $m \in \mathbb{N}$. Here, and in what follows, we assume that $f^e_m(t \mid X_1, \ldots, X_{m-1})$ and $F^e_m(x \mid x_1, \ldots, x_{m-1})$ for $m = 1$ denote unconditional ch.f. and d.f. of $\mathscr{L}^e(X_1)$. A similar notation will be used for other characteristics of distributions.

Remark 1. *If $X_1, X_2, \ldots$ are independent, then* (11)–(12) *reduces to the single summand property of the equilibrium transform (see (Equation (30), [3]).*

Remark 2. *If all $X_n \geq 0$ a.e. and M is independent of $\{X_1, X_2, \ldots\}$, then* (11)–(12) *can be expressed in terms of random variables as*

$$S^e_N \overset{d}{=} S_{M-1} + Z_M,$$

where the sequence $\{Z_1, Z_2, \ldots\}$ is independent of M and the conditional distribution of Z_n given $X_1, \ldots, X_{n-1}$ coincides with $\mathscr{L}^e(X_n \mid X_1, \ldots, X_{n-1})$.

Proof. According to ([3], Lemma 2), for every $t \in \mathbb{R}$ and $n \in \mathbb{N}$ we have

$$\prod_{k=1}^{n} e^{itX_k} - 1 = \sum_{k=1}^{n}\left(e^{itX_k} - 1\right)\prod_{j=1}^{k-1} e^{itX_j},$$

where $\prod_{j=1}^{0} \ldots \equiv 1$. By applying (5) twice, we obtain for $t \neq 0$

$$\begin{aligned}
f^e_{S_N}(t) &= \frac{f_{S_N}(t) - 1}{t f'_{S_N}(0)} = \frac{1}{itA} \sum_{n=1}^{\infty} \mathbf{P}(N = n)\, \mathbf{E}\left(\prod_{k=1}^{n} e^{itX_k} - 1\right) = \\
&= \sum_{n=1}^{\infty} \mathbf{P}(N = n) \sum_{k=1}^{n} \mathbf{E}\left(\frac{e^{itX_k} - 1}{itA} \prod_{j=1}^{k-1} e^{itX_j}\right) = \\
&= \sum_{k=1}^{\infty} \frac{a_k}{A}\mathbf{P}(N \geq k)\, \mathbf{E}\left(\frac{e^{itX_k} - 1}{ita_k} \prod_{j=1}^{k-1} e^{itX_j}\right) = \\
&= \sum_{k=1}^{\infty} \mathbf{P}(M = k)\, \mathbf{E}\left(e^{itS_{k-1}}\mathbf{E}\left(\frac{e^{itX_k} - 1}{ita_k} \,\Big|\, X_1, \ldots, X_{k-1}\right)\right) = \\
&= \sum_{k=1}^{\infty} \mathbf{P}(M = k)\, \mathbf{E}\left(e^{itS_{k-1}} f^e_k(t \mid X_1, \ldots, X_{k-1})\right). \quad \square
\end{aligned}$$

Theorem 1. *Let* $X_1, X_2, \ldots$ *be a sequence of random variables, such that for every* $n \geq 2$ *there exists a regular conditional probability* $\mathscr{L}(X_n \mid X_1, \ldots, X_{n-1})$ *with the constant conditional expectation* $a_n := \mathbf{E}(X_n \mid X_1, \ldots, X_{n-1}) \in (0, +\infty)$. *Let* $S_n := \sum_{i=1}^{n} X_n$ *for* $n \in \mathbb{N}$, $S_0 := 0$ *and* N *be a* $\mathbb{N}_0$-*valued r.v., independent of* $\{X_1, X_2, \ldots\}$, *with*

$$A := \mathbf{E}S_N = \sum_{n=1}^{\infty} a_n \, \mathbf{P}(N \geq n) < +\infty.$$

Let $\mathscr{E} \sim \mathrm{Exp}(1)$, $W := S_N / \mathbf{E}S_N = S_N / A$, *and* M *be a* $\mathbb{N}$-*valued r.v. with*

$$\mathbf{P}(M = m) = \frac{a_m}{A} \mathbf{P}(N \geq m), \quad m \in \mathbb{N}.$$

Then, for any joint distribution of N *and* M *we have*

$$\zeta_1(W, \mathscr{E}) \leq 2A^{-1}\left(\sup_n \mathbf{E}|X_n| \cdot \mathbf{E}|N - M| + D \right), \tag{13}$$

where the first term vanishes in case of $N \stackrel{d}{=} M$ *and*

$$D = \sum_{m \in \mathbb{N}} \mathbf{P}(M = m) \int_{\mathbb{R}^{m-1}} \zeta_1\big(\mathscr{L}(X_m \mid x_1, \ldots, x_{m-1}), \mathscr{L}^e(X_m | x_1, \ldots, x_{m-1})\big) \, dF(x_1, \ldots, x_{m-1}),$$

and both notations $\mathscr{L}(X_m \mid x_1, \ldots, x_{m-1})$, $\mathscr{L}^e(X_m | x_1, \ldots, x_{m-1})$ *stand for the short forms of* $\mathscr{L}(X_m \mid X_1 = x_1, \ldots, X_{m-1} = x_{m-1}), \mathscr{L}^e(X_m | X_1 = x_1, \ldots, X_{m-1} = x_{m-1})$, *respectively.*

Proof. By Lemma 1 and homogeneity of both the Kantorovich distance and the equilibrium transform (see (6) and (10)), we have

$$\zeta_1(W, \mathscr{E}) \leq 2\,\zeta_1\big(\mathscr{L}(W), \mathscr{L}(W^e)\big) = 2A^{-1}\zeta_1\big(\mathscr{L}(S_N), \mathscr{L}^e(S_N)\big). \tag{14}$$

Let us bound $\zeta_1\big(\mathscr{L}(S_N), \mathscr{L}^e(S_N)\big)$ from above.

For a given joint distribution $\mathscr{L}(N, M)$, let $p_{nm} := \mathbf{P}(N = n, M = m)$, $n \in \mathbb{N}_0$, $m \in \mathbb{N}$. Denoting $S_{j,k} := \sum_{i=j}^{k} X_i$ for $j \leq k$, designating $F_m(x \mid x_1, \ldots, x_{m-1})$ and $F_m^e(x \mid x_1, \ldots, x_{m-1})$ the short forms of the conditional d.f.'s of $\mathscr{L}(X_m \mid X_1 = x_1, \ldots, X_{m-1} = x_{m-1})$ and $\mathscr{L}^e(X_m \mid X_1 = x_1, \ldots, X_{m-1} = x_{m-1})$, respectively, $m \in \mathbb{N}$, and using Lemma 3 together with the representation of the Kantorovich distance between (signed) measures as an area between their distribution functions, we obtain

$$\zeta_1\big(\mathscr{L}(S_N), \mathscr{L}^e(S_N)\big) = \int_{\mathbb{R}} \left| F_{S_N}(x) - F_{S_N}^e(x) \right| dx =$$

$$= \int_{\mathbb{R}} \left| F_{S_N}(x) - \sum_{m=1}^{\infty} \mathbf{P}(M=m) \int_{\mathbb{R}^{m-1}} F_m^e(x - x_1 - \ldots - x_{m-1} \mid x_1, \ldots, x_{m-1}) \, dF(x_1, \ldots, x_{m-1}) \right| dx =$$

$$= \int_{\mathbb{R}} \left| \sum_{\substack{n \in \mathbb{N}_0 \\ m \in \mathbb{N}}} p_{nm} \left(F_{S_n}(x) - \int_{\mathbb{R}^{m-1}} F_m^e(x - x_1 - \ldots - x_{m-1} \mid x_1, \ldots, x_{m-1}) \, dF(x_1, \ldots, x_{m-1}) \right) \right| dx \leq$$

$$\leq \sum_{\substack{n \in \mathbb{N}_0 \\ m \in \mathbb{N}}} I_{nm},$$

where

$$I_{nm} = \int_{\mathbb{R}} \left| F_{S_n}(x) - \int_{\mathbb{R}^{m-1}} F_m^e(x - x_1 - \ldots - x_{m-1} \mid x_1, \ldots, x_{m-1}) \, dF(x_1, \ldots, x_{m-1}) \right| dx.$$

For the summands with $n < m$ by Tonelli's theorem we have

$$I_{nm} \le \int_{\mathbb{R}^n} \int_{\mathbb{R}} \left| \mathbb{1}_{\{x_1 + \ldots + x_n < x\}} - \int_{\mathbb{R}^{m-1-n}} F_m^e(x - x_1 - \ldots - x_{m-1} \mid x_1, \ldots, x_{m-1}) \times \right.$$

$$\left. \times dF(x_{n+1}, \ldots, x_{m-1} \mid x_1, \ldots, x_n) \right| dx \, dF(x_1, \ldots, x_n) =$$

$$= \int_{\mathbb{R}^n} \int_{\mathbb{R}} \left| \mathbb{1}_{\{0 < x\}} - \int_{\mathbb{R}^{m-1-n}} F_m^e(x - x_{n+1} - \ldots - x_{m-1} \mid x_1, \ldots, x_{m-1}) \times \right.$$

$$\left. \times dF(x_{n+1}, \ldots, x_{m-1} \mid x_1, \ldots, x_n) \right| dx \, dF(x_1, \ldots, x_n),$$

where $F(x_{n+1}, \ldots, x_{m-1} \mid x_1, \ldots, x_n)$ stands for the conditional joint d.f. of $(X_{n+1}, \ldots, X_{m-1})$ given that $X_1 = x_1, \ldots, X_n = x_n$. By adding and subtracting

$$\int_{\mathbb{R}^{m-1-n}} F_m(x - x_{n+1} - \ldots - x_{m-1} \mid x_1, \ldots, x_{m-1}) \, dF(x_{n+1}, \ldots, x_{m-1} \mid x_1, \ldots, x_n)$$

under the modulus sign and using further the triangle inequality, we obtain

$$I_{nm} \le \int_{\mathbb{R}^n} \zeta_1 \left(\delta_0, \mathscr{L}(S_{n+1,m} \mid x_1, \ldots, x_n) \right) \, dF(x_1, \ldots, x_n) +$$

$$+ \int_{\mathbb{R}^{m-1}} \zeta_1 \left(\mathscr{L}(X_m \mid x_1, \ldots, x_{m-1}), \mathscr{L}^e(X_m \mid x_1, \ldots, x_{m-1}) \right) \, dF(x_1, \ldots, x_{m-1}),$$

where δ_0 is the Dirac measure concentrated in zero.

For the case of $n \ge m$ by Tonelli's theorem, we have

$$I_{nm} \le \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}} \left| F_{S_{m,n}}(x \mid x_1, \ldots, x_{m-1}) - F_m^e(x \mid x_1, \ldots, x_{m-1}) \right| dx \, dF(x_1, \ldots, x_{m-1}),$$

where $F_{S_{m,n}}(x \mid x_1, \ldots, x_{m-1})$ stands for the conditional d.f. $F_{S_{m,n}}(x \mid X_1 = x_1, \ldots, X_{m-1} = x_{m-1})$.

By adding and subtracting $F_m(x \mid x_1, \ldots, x_{m-1})$ in the integrand under the modulus sign and using further the triangle inequality, we obtain

$$I_{nm} \le \int_{\mathbb{R}^m} \zeta_1 \left(\delta_0, \mathscr{L}(S_{m+1,n} \mid x_1, \ldots, x_m) \right) \, dF(x_1, \ldots, x_m) +$$

$$+ \int_{\mathbb{R}^{m-1}} \zeta_1 \left(\mathscr{L}(X_m \mid x_1, \ldots, x_{m-1}), \mathscr{L}^e(X_m \mid x_1, \ldots, x_{m-1}) \right) \, dF(x_1, \ldots, x_{m-1}).$$

Combining both $n < m$ and $n \ge m$ cases and using the fact that $\zeta_1(\delta_0, \mathscr{L}(X)) = \mathbf{E}|X|$, we get

$$\zeta_1 \left(\mathscr{L}(S_N), \mathscr{L}^e(S_N) \right) \le \sum_{n \in \mathbb{N}_0, m \in \mathbb{N}} p_{nm} \mathbf{E} \left| \sum_{i=(n \wedge m)+1}^{n \vee m} X_i \right| +$$

$$+ \sum_{m \in \mathbb{N}} \mathbf{P}(M = m) \int_{\mathbb{R}^{m-1}} \zeta_1 \left(\mathscr{L}(X_m \mid x_1, \ldots, x_{m-1}), \mathscr{L}^e(X_m \mid x_1, \ldots, x_{m-1}) \right) \, dF(x_1, \ldots, x_{m-1}), \quad (15)$$

where the first sum can be bounded from above by

$$\sup_i \mathbf{E}|X_i| \cdot \sum_{n \in \mathbb{N}_0, m \in \mathbb{N}} p_{nm} |n - m| = \sup_i \mathbf{E}|X_i| \cdot \mathbf{E} |N - M|.$$

Substituting the latter bound into (14) yields (13). If $N \overset{d}{=} M$, then we take a comonotonic pair (N, N) as (N, M), which eliminates the first term on the r.-h.s. of (15). $\quad \square$

Remark 3. *Theorem 1 reduces to ([2], Theorem 3.1) in case of nonnegative $\{X_n\}$ and to ([3], Theorem 6) in case of independent $\{X_n\}$.*

If both expectations $\mathbf{E}N$ and $\mathbf{E}M$ are finite, then $\mathbf{E}|N - M|$ in (13) can be replaced with $\zeta_1(N, M)$ due to the dual representation of the ζ_1-metric as

$$\zeta_1(N, M) = \inf_{\mathscr{L}(N',M'):\ N' \stackrel{d}{=} N, M' \stackrel{d}{=} M} \mathbf{E}|N' - M'|.$$

Moreover, if N and M are stochastically ordered (that is, $F_N(x) \leq F_M(x)$ for all $x \in \mathbb{R}$ or vice versa), then

$$\zeta_1(N, M) = \int_{\mathbb{R}} |F_N(x) - F_M(x)|\ dx = \left| \int_{\mathbb{R}} F_N(x) - F_M(x)\ dx \right| = |\mathbf{E}N - \mathbf{E}M|.$$

If, in addition, all $\mathbf{E}X_n = a$, then

$$\mathbf{E}M = \sum_{m=1}^{\infty} \frac{m}{\mathbf{E}N} \mathbf{P}(N \geq m) = \frac{1}{\mathbf{E}N} \sum_{n=1}^{\infty} \sum_{m=1}^{n} m\,\mathbf{P}(N = n) = \frac{1}{\mathbf{E}N} \mathbf{E}\frac{N(N+1)}{2} = \frac{1}{2}\left(\frac{\mathbf{E}N^2}{\mathbf{E}N} + 1\right),$$

and the first term on the r.-h.s of (13) can be bounded from above as

$$2A^{-1} \sup_n \mathbf{E}|X_n| \cdot |\mathbf{E}N - \mathbf{E}M| \leq \frac{2}{a\,\mathbf{E}N} \sup_n \mathbf{E}|X_n| \cdot \left| \frac{1}{2}\left(\frac{\mathbf{E}N^2}{\mathbf{E}N} + 1\right) - \mathbf{E}N \right| =$$

$$= \frac{1}{a} \sup_n \mathbf{E}|X_n| \cdot \left| \frac{\mathbf{E}N^2}{(\mathbf{E}N)^2} + \frac{1}{\mathbf{E}N} - 2 \right|.$$

Hence, we arrive at the following.

Corollary 1. *Let, in addition to the conditions of Theorem 1, $\mathbf{E}X_n = a$ for all $n \in \mathbb{N}$ and the r.v.'s N and M be stochastically ordered with finite expectations. Then*

$$\zeta_1(W, \mathscr{E}) \leq \frac{1}{a} \sup_n \mathbf{E}|X_n| \cdot \left| \frac{\mathbf{E}N^2}{(\mathbf{E}N)^2} + \frac{1}{\mathbf{E}N} - 2 \right| + \frac{2D}{a\mathbf{E}N}.$$

Remark 4. *If $N \sim \mathrm{Geom}(p)$, $p \in (0, 1)$, that is $\mathbf{P}(N = n) = (1 - p)^{n-1}p$, $n \in \mathbb{N}$, then*

$$A = \mathbf{E}S_N = \sum_{n=1}^{\infty} a_n \mathbf{P}(N \geq n) = \sum_{n=1}^{\infty} a_n(1 - p)^{n-1} = \frac{1}{p}\mathbf{E}a_N.$$

In this case, for every $h \in \mathrm{Lip}_1$ with $\mathbf{E}|h(M)| < \infty$ we have

$$\mathbf{E}h(M) = \sum_{m=1}^{\infty} h(m)\frac{a_m}{A}(1 - p)^{m-1} = \frac{\mathbf{E}h(N)a_N}{\mathbf{E}a_N}.$$

Therefore, by the Cauchy–Bunyakovsky–Schwarz inequality, we have

$$\zeta_1(N, M) = \sup_{\substack{h \in \mathrm{Lip}_1 \\ h(0)=0}} |\mathbf{E}h(M) - \mathbf{E}h(N)| \leq \sup_{\substack{h \in \mathrm{Lip}_1 \\ h(0)=0}} \mathbf{E}\left| \left(\frac{a_N}{\mathbf{E}a_N} - 1\right) h(N) \right| \leq \mathbf{E}\left| \left(\frac{a_N}{\mathbf{E}a_N} - 1\right) N \right| \leq$$

$$\leq \frac{\sqrt{\mathbf{Var}\,a_N \cdot \mathbf{E}N^2}}{\mathbf{E}a_N} = \frac{\sqrt{(2 - p)\mathbf{Var}\,a_N}}{p\,\mathbf{E}a_N} < \frac{\sqrt{2\mathbf{Var}\,a_N}}{p\,\mathbf{E}a_N}.$$

Thus, the first term on the r.-h.s of (13) *can be bounded from above as*

$$2A^{-1} \sup_n \mathbf{E}|X_n| \cdot \zeta_1(N, M) \leq 2\sqrt{2} \sup_n \mathbf{E}|X_n| \cdot \frac{\sqrt{\mathbf{Var}\, a_N}}{(\mathbf{E}a_N)^2}.$$

This means that in case of $\sup_n \mathbf{E}|X_n| < \infty$ *and* $\inf_n a_n > 0$, *the first term on the r.-h.s. of* (13) *is, at most, of order* $O\left(\sqrt{\mathbf{Var}\, a_N}\right)$ *as* $p \to +0$.

If $N \sim \text{Geom}(p)$, $p \in (0, 1)$ *and* $\mathbf{E}X_n = a$ *for all* $n \in \mathbb{N}$, *then* $M \sim \text{Geom}(p)$, *and thus,* $\zeta_1(N, M) = 0$. *Therefore, if* $\sup_n \mathbf{E}|X_n| < \infty$, *then the first term on the r.-h.s. of* (13) *vanishes.*

If $N + 1 \sim \text{Geom}(p)$, $p \in (0, 1)$ *and* $\mathbf{E}X_n = a$ *for all* $n \in \mathbb{N}$, *then* $M \sim \text{Geom}(p)$ *as well, and thus,* $\zeta_1(N, M) = 1$.

Next, let us simplify the second term D in (13).

Corollary 2. *Let, in addition to the conditions of Theorem* 1, $b_n = \mathbf{E}X_n^2 < \infty$ *for every* $n \in \mathbb{N}$ *and the r.v.* M *be independent of* $\{X_1, X_2, \ldots\}$. *Then*

$$\zeta_1(W, \mathscr{E}) \leq A^{-1}\left[2 \sup_n \mathbf{E}|X_n| \cdot \mathbf{E}|N - M| + \mathbf{E}\left(\frac{b_M}{a_M} - 2a_M \mathbf{P}(X_M \leq 0|M)\right)\right].$$

Proof. By Lemma 2, we have

$$D \leq \sum_{m \in \mathbb{N}} \mathbf{P}(M{=}m) \int_{\mathbb{R}^{m-1}} \left(\frac{\mathbf{E}\left(X_m^2 \mid x_1, \ldots, x_{m-1}\right)}{2a_m} - a_m \mathbf{P}(X_m \leq 0|x_1, \ldots, x_{m-1})\right) dF(x_1, \ldots, x_{m-1}) =$$

$$= \sum_{m \in \mathbb{N}} \mathbf{P}(M{=}m)\left(\frac{\mathbf{E}X_m^2}{2a_m} - a_m \mathbf{P}(X_m \leq 0)\right) = \frac{1}{2}\mathbf{E}\left(\frac{b_M}{a_M} - 2a_M \mathbf{P}(X_M \leq 0|M)\right),$$

which proves the statement of the corollary. $\square$

Recall that a nonnegative r.v. X with finite $\mathbf{E}X > 0$ is said to be new better than used in expectation (NBUE), if

$$\mathbf{E}X \geq \mathbf{E}\left(X - t \mid X > t\right) \quad \text{for all} \quad t \geq 0,$$

and new worse than used in expectation (NWUE), if

$$\mathbf{E}X \leq \mathbf{E}\left(X - t \mid X > t\right) \quad \text{for all} \quad t \geq 0.$$

Using Tonelli's theorem, it can be ascertained that X is NBUE if and only if X stochastically dominates its equilibrium transform X^e, that is, $F(x) \leq F^e(x)$ for all $x \geq 0$. Similarly, X is NWUE if and only if X^e stochastically dominates X. We will show that the same results hold true if we extend both NBUE and NWUE notions to the case of r.v.s without support constraints.

Definition 1. *We say that a (possibly alternating) r.v.* X *with d.f.* F *and* $\mathbf{E}X \in (0, +\infty)$ *is NBUE, if* $F(x) \leq F^e(x)$ *for all* $x \in \mathbb{R}$, *where* F^e *is the equilibrium transform w.r.t.* F. *Similarly, we say that the r.v.* X *with d.f.* F *and* $\mathbf{E}X \in (0, +\infty)$ *is NWUE (new worse than used in expectation), if* $F(x) \geq F^e(x)$ *for all* $x \in \mathbb{R}$.

Theorem 2. *A r.v.* X *with finite* $\mathbf{E}X > 0$ *is NBUE if and only if*

$$\mathbf{E}X \geq \mathbf{E}\left(X - t \mid X > t\right) \quad \text{for all} \quad t \in [0, \text{ess sup } X). \tag{16}$$

Moreover, (16) *implies that* $X > 0$ *a.s. The r.v.* X *with finite* $\mathbf{E}X > 0$ *is NWUE if and only if*

$$\mathbf{E}X \leq \mathbf{E}\left(X - t \mid X > t\right) \quad \text{for all} \quad t \in [0, \text{ess sup } X). \tag{17}$$

Proof. By Tonelli's theorem, for every $t < \text{ess sup } X$, we have

$$
\mathbf{E}\,(X - t \mid X > t) = \frac{1}{\mathbf{P}(X > t)} \int_{(t,+\infty)} (x - t)\, dF(x) =
$$
$$
= \frac{1}{\mathbf{P}(X > t)} \int_{(t,+\infty)} \left(\int_{(t,+\infty)} \mathbb{1}\{x > y\}\, dy \right) dF(x) =
$$
$$
= \frac{1}{\mathbf{P}(X > t)} \int_{(t,+\infty)} \left(\int_{(t,+\infty)} \mathbb{1}\{x > y\}\, dF(x) \right) dy =
$$
$$
= \frac{1}{\mathbf{P}(X > t)} \int_{(t,+\infty)} (1 - F(y))\, dy. \tag{18}
$$

Note that the same chain of equalities holds true with the event $\{X \geq t\}$ in place of $\{X > t\}$. If $t \in [0, \text{ess sup } X)$, then $\int_{(t,+\infty)} (1 - F(y))\, dy = \mathbf{E}X \cdot (1 - F^e(t))$ and (18) turns into

$$
\mathbf{E}\,(X - t \mid X > t) = \frac{\mathbf{E}X\,(1 - F^e(t))}{1 - F(t + 0)} \quad \text{for all} \quad t \in [0, \text{ess sup } X). \tag{19}
$$

Let X be NBUE, i.e., $F(x) \leq F^e(x)$ for all $x \in \mathbb{R}$. This implies that $1 - F(t + 0) \geq 1 - F^e(t)$ due to the absolute continuity of F^e, and hence, with the account of (19), we obtain (16).

Conversely, let (16) hold true. For $t = 0$, we have

$$
\mathbf{E}X^+ \geq \mathbf{E}X \geq \mathbf{E}\,(X \mid X > 0) = \frac{1}{\mathbf{P}(X > 0)} \mathbf{E}X^+,
$$

which is possible if and only if $\mathbf{P}(X > 0) = 1$, i.e., $X > 0$ a.s. Hence, $F(t) = F^e(t) = 0$ for $t \leq 0$. For positive t, inequality (16) together with Equation (19) yields $1 - F(t) \geq 1 - F^e(t)$. Therefore, X is NBUE.

Let X be NWUE, i.e., $F(x) \geq F^e(x)$ for all $x \in \mathbb{R}$. This yields $1 - F(t + 0) \leq 1 - F^e(t)$, and hence, with the account of (19), we obtain (17).

Conversely, let (17) hold true. For $t \leq 0$ we have $F(t) \geq 0 \geq F^e(t)$, since $\mathscr{L}^e(X)$ has negative density on the negative half-line. Finally, (17) and (19) yield $F(t) \geq F^e(t)$ for positive t. □

If X is NBUE or NWUE with $\mathbf{E}X > 0$ and $\mathbf{E}X^2 < \infty$, then

$$
\zeta_1\,(\mathscr{L}(X), \mathscr{L}^e(X)) = \int_{\mathbb{R}} |F(x) - F^e(x)|\, dx = \left| \int_{\mathbb{R}} (F(x) - F^e(x))\, dx \right| =
$$
$$
= \left| \int_{\mathbb{R}} x\, dF^e(x) - \mathbf{E}X \right| = \left| \frac{\mathbf{E}X^2}{2\,\mathbf{E}X} - \mathbf{E}X \right|,
$$

where the last equality follows from (7). Hence, if for all $m \in \mathbb{N}$ and $x_1, x_2 \ldots \in \mathbb{R}$ the conditional distribution $\mathscr{L}(X_m \mid x_1, \ldots, x_{m-1})$ is NBUE or NWUE, then the second term on the r.-h.s. of (13) takes the form

$$
D = \sum_{m \in \mathbb{N}} \mathbf{P}(M{=}m) \int_{\mathbb{R}^{m-1}} \left| \frac{\mathbf{E}\,(X_m^2 \mid x_1, \ldots, x_{m-1})}{2a_m} - a_m \right| dF(x_1, \ldots, x_{m-1}). \tag{20}
$$

If M is independent of $\{X_1, X_2, \ldots\}$, then the latter expression can be bounded from above with the help of the conditional Jensen's inequality

$$
D \leq \sum_{m \in \mathbb{N}} \mathbf{P}(M{=}m)\, \mathbf{E} \left| \frac{X_m^2}{2a_m} - a_m \right| = \sum_{m \in \mathbb{N}} \mathbf{P}(M{=}m)\, \frac{\mathbf{E}\,|X_m^2 - 2a_m^2|}{2a_m} = \mathbf{E}\,\frac{|X_M^2 - 2a_M^2|}{2a_M} \leq \mathbf{E}\,\frac{b_M + 2a_M^2}{2a_M},
$$

where we used the notation $b_m = \mathbf{E}X_m^2$ as before. If all the r.v.'s $\{M, X_1, X_2, \ldots\}$ are independent, then (20) may be simplified as

$$D = \sum_{m \in \mathbb{N}} \mathbf{P}(M{=}m) \left| \frac{\mathbf{E}X_m^2}{2a_m} - a_m \right| = \mathbf{E} \frac{|b_M - 2a_M^2|}{2a_M}.$$

Hence, we get the following

Corollary 3. *Let, in addition to the conditions of Theorem 1, $b_n = \mathbf{E}X_n^2 < \infty$ for every $n \in \mathbb{N}$, the conditional distributions $\mathscr{L}(X_n \mid x_1, \ldots, x_{n-1})$ be NBUE or NWUE for all $n \in \mathbb{N}$ and $x_1, x_2 \ldots \in \mathbb{R}$, and the r.v. M be independent of $\{X_1, X_2, \ldots\}$. Then*

$$\zeta_1(W, \mathscr{E}) \leq A^{-1} \left(2 \sup_n \mathbf{E}|X_n| \cdot \mathbf{E}|N - M| + \mathbf{E} \frac{|X_M^2 - 2a_M^2|}{a_M} \right) \leq$$

$$\leq A^{-1} \left(2 \sup_n \mathbf{E}|X_n| \cdot \mathbf{E}|N - M| + \mathbf{E} \frac{b_M + 2a_M^2}{a_M} \right).$$

Moreover, if all the r.v.'s $\{M, X_1, X_2, \ldots\}$ are independent, then

$$\zeta_1(W, \mathscr{E}) \leq A^{-1} \left(2 \sup_n \mathbf{E}|X_n| \cdot \mathbf{E}|N - M| + \mathbf{E} \frac{|b_M - 2a_M^2|}{a_M} \right).$$

Corollary 3 reduces to ([2], Corollary 3.1) in the case of all $X_n \geq 0$ being independent, all $\mathbf{E}X_n = 1$ and the r.v.s N, M being stochastically ordered (cf. Corollary 1).

Abbreviations

The following abbreviations are used in this manuscript:

r.v.   random variable
i.i.d.   independent identically distributed
d.f.   distribution function
ch.f.   characteristic function
a.s.   almost sure
a.c.   absolute continuity, absolutely continuous
w.r.t.   with respect to
r.-h.s.   right-hand side

References

1. Kalashnikov, V.V. *Geometric Sums: Bounds for Rare Events with Applications: Risk Analysis, Reliability, Queueing;* Mathematics and Its Applications; Springer: Dordrecht, The Netherlands, 1997.
2. Peköz, E.A.; Röllin, A. New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **2011**, *39*, 587–608. doi:10.1016/0047-259X(78)90029-5. [CrossRef]

3. Shevtsova, I.; Tselishchev, M. A Generalized Equilibrium Transform with Application to Error Bounds in the Rényi Theorem with No Support Constraints. *Mathematics* **2020**, *8*, 577. [CrossRef]

4. Zolotarev, V.M. Probability metrics. *Theory Probab. Appl.* **1984**, *28*, 278–302. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.