

BOUNDS FOR THE DISTANCE BETWEEN THE DISTRIBUTIONS OF SUMS OF ABSOLUTELY CONTINUOUS I.I.D. CONVEX-ORDERED RANDOM VARIABLES WITH APPLICATIONS

TASOS C. CHRISTOFIDES,* *University of Cyprus*

EUTICHIA VAGGELATOU,** *University of Athens*

Abstract

Let X_1, X_2, \dots and Y_1, Y_2, \dots be two sequences of absolutely continuous, independent and identically distributed (i.i.d.) random variables with equal means $E(X_i) = E(Y_i)$, $i = 1, 2, \dots$. In this work we provide upper bounds for the total variation and Kolmogorov distances between the distributions of the partial sums $\sum_{i=1}^n X_i$ and $\sum_{i=1}^n Y_i$. In the case where the distributions of the X_i s and the Y_i s are compared with respect to the convex order, the proposed upper bounds are further refined. Finally, in order to illustrate the applicability of the results presented, we consider specific examples concerning gamma and normal approximations.

Keywords: Zolotarev's ideal metric of order 2; Kolmogorov distance; total variation distance; convex order; NBUE/NWUE; Weibull distribution; Student distribution; gamma approximation; normal approximation

2000 Mathematics Subject Classification: Primary 60E15; 62E17; 62E20
Secondary 60G50; 60K10

1. Introduction

Suppose that X_1, X_2, \dots is a collection of independent and identically distributed (i.i.d.) random variables and that Y_1, Y_2, \dots is another collection of i.i.d. random variables, different from the first one. It is often of interest to provide an estimate for the closeness of the distributions of the partial sums $\sum_{i=1}^n X_i$ and $\sum_{i=1}^n Y_i$ in the form of an upper bound for the total variation distance and the Kolmogorov distance, in the case where the marginal distributions of X_i and Y_i are compared with respect to certain classes of stochastic orderings. Such an interest also arises from the fact that the distribution of the partial sum of one of the collections is often unknown, yet the distribution of the partial sum of the other is not only known, but indeed, is one of the most frequently used in statistics for inferential purposes. For instance, it is not clear what the distribution of the partial sum of i.i.d. random variables from the Weibull distribution is. However, it is known that the Weibull distribution can be compared with the exponential distribution with respect to the so-called *convex order*. In addition, the sum of i.i.d. exponential random variables follows the gamma distribution. Therefore, it would be reasonable to expect that, when the difference between the variances of the marginal distributions is small,

Received 17 June 2008; revision received 9 January 2009.

* Postal address: Department of Mathematics and Statistics, University of Cyprus, P.O. Box 20537, Nicosia, CY 1678, Cyprus. Email address: tastos@ucy.ac.cy

** Postal address: Section of Statistics and Operations Research, Department of Mathematics, University of Athens, Panepistemiopolis, Athens 15784, Greece. Email address: evagel@math.uoa.gr

Partially supported by the University of Athens research grant 70/4/8810.

the distribution of the sum of i.i.d. Weibull random variables can be well approximated by a gamma law. As demonstrated in Section 3, employing the convex order relation, upper bounds for the total variation distance and the Kolmogorov distance between the distribution of that partial sum and the distribution of the partial sum of i.i.d. exponentially distributed random variables are readily available. We would also like to emphasize that we are mainly interested in the case of fixed n , for otherwise, classical central limit theory provides almost all the answers regarding the closeness of the distributions of the partial sums.

In this work, our study will be focused mainly on distributions which are compared with respect to the convex order. Vaggelatou (2009) gave upper bounds for the total variation distance and the Kolmogorov distance between the distributions of two different collections of independent, integer-valued, convex-ordered random variables. For absolutely continuous random variables, in order to obtain results analogous to the discrete case, further assumptions are needed. In this paper we consider the case of absolutely continuous random variables with their partial sums having absolutely continuous densities.

In Section 2 we provide the main machinery for obtaining upper bounds. Theorem 1, which can be thought of as a ‘source’ result, gives upper bounds in *general* forms. The bounds contain quantities such as the Kolmogorov distance, the total variation distance, and the Zolotarev ζ_2 metric between representative observations from the two collections. When comparing the distributions of the X_i s and the Y_i s with respect to the convex order, Corollary 1 provides upper bounds which are further refined.

In Section 3 we exploit the results of Section 2 to obtain upper bounds for the total variation distance and the Kolmogorov distance between the distributions of the partial sums of i.i.d. random variables from important specific distributions with convexly ordered marginal laws. In particular, cases considered are those of distances between the distribution of the partial sum of i.i.d. observations from the Weibull distribution and the gamma distribution. Furthermore, we consider the case of observations from more general classes of distributions, namely the so-called new better than used in expectation (NBUE) and new worse than used in expectation (NWUE) distributions. Specifically, we provide upper bounds for the case of approximating the distribution of the sum of i.i.d. NBUE and NWUE random variables by a suitable gamma distribution. Finally, we consider the case of the distribution of the partial sum of i.i.d. observations from the Student t -distribution and the normal distribution. The results obtained are compared with other known results in the literature.

2. Main results

For any two random variables X and Y defined on the same probability space (Ω, \mathcal{P}) and with values in \mathbb{R} , the *total variation* distance between their distributions is, by definition,

$$\rho_{\text{TV}}(X, Y) := \sup_{A \in \mathcal{B}(\mathbb{R})} |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|.$$

In the case in which the random variables X and Y have densities f_X and f_Y , respectively, the total variation distance is

$$\rho_{\text{TV}}(X, Y) = \frac{1}{2} \int_{\mathbb{R}} |f_X(t) - f_Y(t)| dt.$$

Moreover, the *Kolmogorov* distance is defined by

$$\rho_{\text{K}}(X, Y) := \sup_{t \in \mathbb{R}} |\mathbb{P}(X \leq t) - \mathbb{P}(Y \leq t)|.$$

For random variables X and Y with finite second moments, i.e. $E(X^2) < \infty$ and $E(Y^2) < \infty$, the total variation distance of order 2 or Zolotarev ζ_2 -metric (see Zolotarev (1983)) is defined by

$$\zeta_2(X, Y) := \int_{\mathbb{R}} |E(X - t)_+ - E(Y - t)_+| dt,$$

where $(x)_+ = \max(x, 0)$. Here, we note that if X and Y have expectations $E(X)$ and $E(Y)$ then the finiteness of $\zeta_2(X, Y)$ implies that $E(X) = E(Y)$, and ζ_2 admits the following representation (see Rachev (1991, p. 258)):

$$\begin{aligned} \zeta_2(X, Y) &= \int_{\mathbb{R}} \left| \int_{-\infty}^t (F_X(u) - F_Y(u)) du \right| dt \\ &= \int_{\mathbb{R}} \left| \int_t^{+\infty} (F_X(u) - F_Y(u)) du \right| dt, \end{aligned}$$

where F_X and F_Y denote the cumulative distribution function (CDF) of the random variables X and Y , respectively. It is then important to note that, from now on, whenever the distance $\zeta_2(X, Y)$ appears, it will be assumed that X and Y have finite second moments and that $E(X) = E(Y)$. Note also that, as we will see later on, the equality of means is a natural situation that results from the convex order relation between two distributions (see Remark 2, below).

For a function $f: \mathbb{R} \rightarrow \mathbb{R}$, the symbol $f^{(k)}$ will denote the k th-order derivative of f , i.e.

$$f^{(k)}(x) := \frac{d^k f(x)}{d^k x}, \quad x \in \mathbb{R}.$$

The result that follows offers a relationship between the Kolmogorov distance and the metric ζ_2 for the case of real-valued random variables with densities and equal first moments.

Lemma 1. *Let X and Y be absolutely continuous random variables with $E(X^2) < \infty$, $E(Y^2) < \infty$, and $E(X) = E(Y)$. Also, let Z be an absolutely continuous random variable independent of X and Y . If the density f_Z of Z is an absolutely continuous function with $\|f_Z^{(1)}\|_{\infty} := \sup_{x \in A} |f_Z^{(1)}(x)| < \infty$, where $A = \{x \in \mathbb{R} : f_Z^{(1)}(x) \text{ exists}\}$, then*

$$\rho_K(X + Z, Y + Z) \leq \|f_Z^{(1)}\|_{\infty} \zeta_2(X, Y).$$

Proof. Let F_X and F_Y denote the CDFs of X and Y , respectively. We then have

$$\begin{aligned} \rho_K(X + Z, Y + Z) &= \sup_{t \in \mathbb{R}} |P(X + Z \leq t) - P(Y + Z \leq t)| \\ &= \sup_{t \in \mathbb{R}} \left| \int_{\mathbb{R}} (F_X(t - z) - F_Y(t - z)) f_Z(z) dz \right|, \end{aligned} \tag{1}$$

where the second equality follows from the independence assumption. Furthermore, using integration by parts, we derive

$$\begin{aligned} &\int_{\mathbb{R}} (F_X(t - z) - F_Y(t - z)) f_Z(z) dz \\ &= - \int_{\mathbb{R}} f_Z^{(1)}(z) \int_{-\infty}^z (F_X(t - u) - F_Y(t - u)) du dz, \end{aligned} \tag{2}$$

because

$$\lim_{z \rightarrow \pm\infty} f_Z(z) = 0,$$

$$\int_{\mathbb{R}} (F_X(t - u) - F_Y(t - u)) \, du = E(Y) - E(X) = 0.$$

Thus, substituting (2) into (1), we finally obtain

$$\begin{aligned} \rho_K(X + Z, Y + Z) &= \sup_{t \in \mathbb{R}} \left| \int_{\mathbb{R}} f_Z^{(1)}(z) \int_{-\infty}^z (F_X(t - u) - F_Y(t - u)) \, du \, dz \right| \\ &\leq \sup_{t \in \mathbb{R}} \int_{\mathbb{R}} |f_Z^{(1)}(z)| \left| \int_{-\infty}^z [F_X(t - u) - F_Y(t - u)] \, du \right| \, dz \\ &\leq \|f_Z^{(1)}\|_{\infty} \zeta_2(X, Y), \end{aligned}$$

and the proof is complete.

Next, we prove an analogous inequality for the total variation distance.

Lemma 2. *Let X and Y be absolutely continuous random variables with $E(X^2) < \infty$, $E(Y^2) < \infty$, and $E(X) = E(Y)$. Also, let Z be an absolutely continuous random variable independent of X and Y . If the density f_Z of Z and its derivative $f_Z^{(1)}$ are absolutely continuous functions then*

$$\rho_{TV}(X + Z, Y + Z) \leq \frac{1}{2} \|f_Z^{(2)}\|_1 \zeta_2(X, Y),$$

where $\|f_Z^{(2)}\|_1 := \int_{\mathbb{R}} |f_Z^{(2)}(x)| \, dx < \infty$.

Proof. If $X + Z$ and $Y + Z$ have densities f_{X+Z} and f_{Y+Z} , respectively, then

$$\begin{aligned} \rho_{TV}(X + Z, Y + Z) &= \frac{1}{2} \int_{\mathbb{R}} |f_{X+Z}(u) - f_{Y+Z}(u)| \, du \\ &= \frac{1}{2} \int_{\mathbb{R}} \left| \int_{\mathbb{R}} (f_X(t) - f_Y(t)) f_Z(u - t) \, dt \right| \, du, \end{aligned} \tag{3}$$

where the second equality follows from the independence assumption.

Applying integration by parts twice, we obtain

$$\begin{aligned} \int_{\mathbb{R}} (f_X(t) - f_Y(t)) f_Z(u - t) \, dt &= \int_{\mathbb{R}} f_Z(u - t) \, d(F_X(t) - F_Y(t)) \\ &= - \int_{\mathbb{R}} f_Z^{(1)}(u - t) (F_X(t) - F_Y(t)) \, dt \\ &= \int_{\mathbb{R}} \int_{-\infty}^t (F_X(s) - F_Y(s)) f_Z^{(2)}(u - t) \, ds \, dt. \end{aligned} \tag{4}$$

Substituting (4) into (3), we obtain

$$\begin{aligned} \rho_{TV}(X + Z, Y + Z) &= \frac{1}{2} \int_{\mathbb{R}} \left| \int_{\mathbb{R}} f_Z^{(2)}(u - t) \int_{-\infty}^t (F_X(s) - F_Y(s)) \, ds \, dt \right| \, du \\ &\leq \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} |f_Z^{(2)}(u - t)| \left| \int_{-\infty}^t (F_X(s) - F_Y(s)) \, ds \right| \, dt \, du \\ &= \frac{1}{2} \|f_Z^{(2)}\|_1 \zeta_2(X, Y). \end{aligned}$$

This completes the proof.

Throughout this work, we deal with i.i.d. random variables. Therefore, in order to simplify the notation used, for any i.i.d. collection U_1, U_2, \dots, U_n , it is convenient to set

$$\mu_U := E(U_i), \quad \sigma_U^2 := \text{var}(U_i), \quad i = 1, 2, \dots, n,$$

where $E(U_i)$ and $\text{var}(U_i)$ are the mean and variance of U_i , respectively. Furthermore, $S_n^U := \sum_{i=1}^n U_i$. In addition, for any i.i.d. collections X_1, X_2, \dots, X_n and Y_1, Y_2, \dots, Y_n , we set

$$\rho_k := \rho_K(X_i, Y_i), \quad \rho_{\text{TV}} := \rho_{\text{TV}}(X_i, Y_i), \quad \zeta := \zeta_2(X_i, Y_i), \quad (5)$$

for all $i = 1, 2, \dots, n$, assuming, of course, that these quantities are positive, for otherwise, there is nothing to show.

Finally, for any density f , let f_n denote the n th convolution of f , i.e.

$$\underbrace{f * f * \dots * f}_{n \text{ times}}$$

We may now state the following general result.

Theorem 1. *Let X_1, X_2, \dots, X_n and Y_1, Y_2, \dots, Y_n , $n \geq 2$, be two collections where each one consists of absolutely continuous i.i.d. random variables with finite second moments. Moreover, let f denote the density of either the X_i s or Y_i s. If $\mu_X = \mu_Y$ and the convolution f_{n-1} and its derivative $f_{n-1}^{(1)}$ are absolutely continuous functions, then*

$$\rho_{\text{TV}}(S_n^X, S_n^Y) \leq \left(2\rho_{\text{TV}}^2 + 2\gamma_{\text{TV}} + \frac{\gamma_{\text{TV}}}{2\rho_{\text{TV}}}\right)(1 + 2\rho_{\text{TV}})^{n-2} - \frac{\gamma_{\text{TV}}}{2\rho_{\text{TV}}} \quad (6)$$

and

$$\rho_K(S_n^X, S_n^Y) \leq \left(2\rho_{\text{TV}}\rho_k + 2\gamma_k + \frac{\gamma_k}{2\rho_{\text{TV}}}\right)(1 + 2\rho_{\text{TV}})^{n-2} - \frac{\gamma_k}{2\rho_{\text{TV}}}, \quad (7)$$

where $\gamma_k := \zeta \|f_{n-1}^{(1)}\|_\infty$, $\gamma_{\text{TV}} := \frac{1}{2}\zeta \|f_{n-1}^{(2)}\|_1$, $f_{n-1}^{(1)}$ and $f_{n-1}^{(2)}$ are the first and second derivatives of the density f_{n-1} of the sum $\sum_{i=1, i \neq s}^n Y_i$, and ρ_k, ρ_{TV} , and ζ are given by (5).

In addition, if the density f is bounded then, for the Kolmogorov distance, we have

$$\rho_K(S_n^X, S_n^Y) \leq \left(2\rho_{\text{TV}}\delta + 2\gamma_k + \frac{\gamma_k}{2\rho_{\text{TV}}}\right)(1 + 2\rho_{\text{TV}})^{n-2} - \frac{\gamma_k}{2\rho_{\text{TV}}}, \quad (8)$$

where $\delta = \min\{\rho_{\text{TV}}, 3M^{2/3}\zeta^{1/3}\}$ with $M := \sup_{y \in \mathbb{R}} f(y)$.

Proof. Let the notation ρ stand for either the Kolmogorov or the total variation distance. Using the Lindeberg decomposition, we obtain

$$\begin{aligned} & \rho\left(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i\right) \\ & \leq \rho\left(\sum_{i=1}^n X_i, \sum_{i=1}^{n-1} X_i + Y_n\right) + \rho\left(\sum_{i=1}^{n-1} X_i + Y_n, \sum_{i=1}^{n-2} X_i + \sum_{i=n-1}^n Y_i\right) \\ & \quad + \dots + \rho\left(\sum_{i=1}^2 X_i + \sum_{i=3}^n Y_i, X_1 + \sum_{i=2}^n Y_i\right) + \rho\left(X_1 + \sum_{i=2}^n Y_i, \sum_{i=1}^n Y_i\right) \\ & = \sum_{s=1}^n \rho\left(\sum_{i=1}^s X_i + \sum_{i=s+1}^n Y_i, \sum_{i=1}^{s-1} X_i + \sum_{i=s}^n Y_i\right). \end{aligned} \quad (9)$$

Without loss of generality, we may assume that Y_1, Y_2, \dots, Y_n are independent of X_1, X_2, \dots, X_n . This assumption enables us to apply the next smoothing inequality valid for (Z, W) independent of (U, V) :

$$\rho(Z + U, W + U) \leq 2\rho_{\text{TV}}(Z, W)\rho(U, V) + \rho(Z + V, W + V) \tag{10}$$

(cf. Proposition 2.9 of Rachev and Rüschendorf (1990) and Lemmas 14.3.2 and 17.1.7 of Rachev (1991)).

Now, applying (10) to each term in (9) with

$$Z = X_s + \sum_{i=s+1}^n Y_i, \quad W = \sum_{i=s}^n Y_i, \quad U = \sum_{i=1}^{s-1} X_i, \quad V = \sum_{i=1}^{s-1} Y_i,$$

we derive

$$\begin{aligned} \rho\left(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i\right) &\leq 2 \sum_{s=1}^n \rho_{\text{TV}}\left(X_s + \sum_{i=s+1}^n Y_i, \sum_{i=s}^n Y_i\right) \rho\left(\sum_{i=1}^{s-1} X_i, \sum_{i=1}^{s-1} Y_i\right) \\ &\quad + \sum_{s=1}^n \rho\left(X_s + \sum_{i=1, i \neq s}^n Y_i, \sum_{i=1}^n Y_i\right) \\ &= 2 \sum_{s=2}^n a_s b_s + \sum_{s=1}^n c_s \end{aligned} \tag{11}$$

(the summation for $s = 1$ is empty), where

$$\begin{aligned} a_s &= \rho_{\text{TV}}\left(X_s + \sum_{i=s+1}^n Y_i, \sum_{i=s}^n Y_i\right), \quad b_s = \rho\left(\sum_{i=1}^{s-1} X_i, \sum_{i=1}^{s-1} Y_i\right), \\ c_s &= \rho\left(X_s + \sum_{i=1, i \neq s}^n Y_i, \sum_{i=1}^n Y_i\right). \end{aligned}$$

Next, we will bound each a_s and c_s term.

By the subadditivity property of the total variation distance we obtain

$$a_s = \rho_{\text{TV}}\left(X_s + \sum_{i=s+1}^n Y_i, Y_s + \sum_{i=s+1}^n Y_i\right) \leq \rho_{\text{TV}}(X_s, Y_s) = \rho_{\text{tv}}. \tag{12}$$

Furthermore, applying Lemmas 1 and 2 on c_s , we respectively derive the bounds

$$c_s = \rho\left(X_s + \sum_{i=1, i \neq s}^n Y_i, \sum_{i=1}^n Y_i\right) \leq \begin{cases} \gamma_{\text{K}} & \text{for } \rho = \rho_{\text{K}}, \\ \gamma_{\text{TV}} & \text{for } \rho = \rho_{\text{TV}}, \end{cases} \tag{13}$$

where

$$\gamma_{\text{K}} = \zeta \|f_{n-1}^{(1)}\|_{\infty} \quad \text{and} \quad \gamma_{\text{TV}} = \frac{1}{2} \zeta \|f_{n-1}^{(2)}\|_1.$$

Using inequalities (12) and (13) in (11), we obtain the following relation for the total variation distance:

$$b_{n+1} = \rho_{\text{TV}}(S_n^X, S_n^Y) \leq 2\rho_{\text{tv}} \sum_{s=2}^n b_s + n\gamma_{\text{TV}}. \tag{14}$$

Applying (14) for $n = 2$ we can bound the quantity b_3 in terms of b_2 , and by using (14) again we can bound the quantity b_4 in terms of b_2 etc.; thus, iteratively, we can bound all the b_s s in terms of b_2 , i.e. in terms of ρ_{tv} . Having done that, standard algebraic manipulation in (14) gives the following inequality for the total variation distance:

$$\rho_{TV}(S_n^X, S_n^Y) \leq \left(2\rho_{tv}^2 + 2\gamma_{tv} + \frac{\gamma_{tv}}{2\rho_{tv}} \right) (1 + 2\rho_{tv})^{n-2} - \frac{\gamma_{tv}}{2\rho_{tv}},$$

i.e. inequality (6). The above inequality can be easily verified by induction.

To establish (7), we apply a similar reasoning. Now the b_s and c_s correspond to the Kolmogorov distance. Hence, using inequalities (12) and (13) in (11) for $\rho = \rho_K$, we obtain the recursive inequality

$$b_{n+1} = \rho_K(S_n^X, S_n^Y) \leq 2\rho_{tv} \sum_{s=2}^n b_s + n\gamma_K,$$

and the b_s s can be found recursively in the same way as in the total variation case, finally giving

$$\rho_K(S_n^X, S_n^Y) \leq \left(2\rho_{tv}\rho_K + 2\gamma_K + \frac{\gamma_K}{2\rho_{tv}} \right) (1 + 2\rho_{tv})^{n-2} - \frac{\gamma_K}{2\rho_{tv}},$$

i.e. inequality (7).

Additionally, assuming that Y_i has a bounded density f , we may employ relation (14.1.16) of Rachev (1991) to obtain an upper bound for the Kolmogorov distance in terms of ζ_2 , namely,

$$\rho_K = \rho_K(X_s, Y_s) \leq 3M^{2/3}(\zeta_2(X_s, Y_s))^{1/3} = 3M^{2/3}\zeta^{1/3}, \tag{15}$$

where $M := \sup_{y \in \mathbb{R}} f(y)$. Using (15) and the fact that $\rho_K \leq \rho_{tv}$, we finally obtain (8).

Remark 1. For computational purposes, especially in cases where the quantity ρ_{tv} is not known explicitly, but rather we have a suitable upper bound for it, it would be preferable to express the bounds in (6), (7), and (8) as

$$\rho_{TV}(S_n^X, S_n^Y) \leq 2(\rho_{tv}^2 + \gamma_{tv})(1 + 2\rho_{tv})^{n-2} + \gamma_{tv} \sum_{k=1}^{n-2} \binom{n-2}{k} (2\rho_{tv})^{k-1}, \tag{16}$$

$$\rho_K(S_n^X, S_n^Y) \leq 2(\rho_{tv}\rho_K + \gamma_K)(1 + 2\rho_{tv})^{n-2} + \gamma_K \sum_{k=1}^{n-2} \binom{n-2}{k} (2\rho_{tv})^{k-1}, \tag{17}$$

and

$$\rho_K(S_n^X, S_n^Y) \leq 2(\rho_{tv}\delta + \gamma_K)(1 + 2\rho_{tv})^{n-2} + \gamma_K \sum_{k=1}^{n-2} \binom{n-2}{k} (2\rho_{tv})^{k-1}, \tag{18}$$

respectively. The expressions above follow from the binomial expansion of the term $(1 + 2\rho_{tv})^{n-2}$.

Next, we recall the definition of the convex order.

Definition 1. A random variable X is smaller or larger than a random variable Y with respect to the convex order if $E f(X) \leq E f(Y)$ or, respectively, $E f(X) \geq E f(Y)$ for all convex functions f for which the expectations exist. In such cases we write $X \preceq_{cx} Y$ or $X \succeq_{cx} Y$, respectively.

For more details on the properties and applications of the convex order, we refer the reader to Szekli (1995), Müller and Stoyan (2002), and Shaked and Shantikumar (2007).

Remark 2. It should be noted that the convex relation $X \preceq_{cx} Y$ or $X \succeq_{cx} Y$ implies that $E(X) = E(Y)$ and $\text{var}(X) \leq \text{var}(Y)$ or, respectively, $\text{var}(X) \geq \text{var}(Y)$. Moreover, if $X \preceq_{cx} Y$ or $X \succeq_{cx} Y$ then $X =_{st} Y$ if and only if $\text{var}(X) = \text{var}(Y)$. Therefore, the closeness of the distributions of X and Y can be measured in terms of the difference between their variances.

Assuming that the distributions of the X_i s and Y_i s are compared with respect to the convex order as defined above, we can obtain upper bounds for the total variation and Kolmogorov distances in terms of the difference between their variances, $|\sigma_Y^2 - \sigma_X^2|$. This is due to the fact that, for random variables X and Y such that $X \preceq_{cx} Y$ or $X \succeq_{cx} Y$, Kaas (1993) proved that the metric $\zeta_2(X, Y)$ is equal to

$$\zeta_2(X, Y) = \frac{|\sigma_Y^2 - \sigma_X^2|}{2}.$$

Thus, we have the following result, the proof of which is immediate.

Corollary 1. Let X_1, X_2, \dots, X_n and $Y_1, Y_2, \dots, Y_n, n \geq 2$, be two collections where each one consists of absolutely continuous i.i.d. random variables with finite second moments. Moreover, let f denote the density of either the X_i s or Y_i s. If

$$X_i \preceq_{cx} Y_i \quad \text{or} \quad X_i \succeq_{cx} Y_i \quad \text{for all } i = 1, 2, \dots, n,$$

and the convolution f_{n-1} and its derivative $f_{n-1}^{(1)}$ are absolutely continuous functions, then inequalities (6), (7), (8), (16), (17), and (18) are valid with

$$\zeta = \frac{1}{2}|\sigma_Y^2 - \sigma_X^2|, \quad \rho_{TV} = \frac{1}{4}|\sigma_Y^2 - \sigma_X^2| \|f_{n-1}^{(2)}\|_1, \quad \text{and} \quad \gamma_k = \frac{1}{2}|\sigma_Y^2 - \sigma_X^2| \|f_{n-1}^{(1)}\|_\infty.$$

The bounds of Theorem 1 and Corollary 1 are expressed in terms of the total variation distance ρ_{TV} between the distributions of the coordinates X_i and Y_i . Unfortunately, an analogue of the inequality in (15) (cf. proof of Theorem 1) for the total variation distance does not hold true. This is, in fact, a general problem for the continuous distributions. In the discrete case, the total variation distance is always upper bounded by twice the Zolotarev’s metric ζ_2 , i.e. $\rho_{TV}(X, Y) \leq 2\zeta_2(X, Y)$ for X and Y taking values in \mathbb{Z} (cf. Lemma 1 of Boutsikas and Vaggelatou (2008)). However, for real-valued random variables with density, ρ_{TV} cannot be compared with ζ_2 (in particular, there are cases where $\zeta_2 \rightarrow 0$ while ρ_{TV} remains constant). Therefore, the distance ρ_{TV} must be evaluated or upper bounded individually according to the distributions involved each time in the problem of interest. The latter is a situation that usually occurs in the continuous-distribution approximation error estimates (for example, in the Berry–Esseen-type results).

3. Applications

In this section we illustrate the previous results with specific applications concerning gamma and normal approximations.

3.1. Distance between the sum of i.i.d. Weibull-distributed random variables and a gamma-distributed random variable

Let X_1, X_2, \dots, X_n be i.i.d. random variables following the Weibull distribution $\mathcal{W}(\alpha, \theta)$ with density

$$h(x) = \alpha\theta(\theta x)^{\alpha-1} \exp\{-(\theta x)^\alpha\}, \quad x > 0, \alpha > 0, \theta > 0 \tag{19}$$

(α and θ are the shape and scale parameters, respectively). The mean and variance of $\mathcal{W}(\alpha, \theta)$ are given by

$$\mu_X = \frac{1}{\theta} \Gamma\left(1 + \frac{1}{\alpha}\right), \quad \sigma_X^2 = \frac{1}{\theta^2} \Gamma\left(1 + \frac{2}{\alpha}\right) - \mu_X^2,$$

where $\Gamma(\cdot)$ denotes the gamma function, i.e. $\Gamma(\beta) = \int_0^\infty t^{\beta-1} e^{-t} dt$, $\beta > 0$. We also consider a collection of i.i.d. random variables Y_1, Y_2, \dots, Y_n distributed according to the exponential distribution $\mathcal{E}(1/\mu_X)$ with parameter $1/\mu_X$, i.e. with density

$$g(x) = \frac{\theta}{\Gamma(1 + 1/\alpha)} \exp\left\{-\frac{\theta}{\Gamma(1 + 1/\alpha)} x\right\}, \quad x > 0. \tag{20}$$

It is known that $\mathcal{W}(\alpha, \theta) \leq_{cx} \mathcal{E}(1/\mu_X)$ for $\alpha > 1$, $\mathcal{W}(\alpha, \theta) \geq_{cx} \mathcal{E}(1/\mu_X)$ for $\alpha < 1$, and, clearly, $\mathcal{W}(\alpha, \theta) =_{st} \mathcal{E}(1/\mu_X)$ for $\alpha = 1$.

The distribution of the sum $S_n^X = \sum_{i=1}^n X_i$ is unknown and, therefore, it would be useful to approximate it by a known distribution. Of course, for $n \rightarrow \infty$, employing the central limit theorem we may use the normal approximation. However, we note that, when $\alpha \rightarrow 1$, $\mathcal{W}(\alpha, \theta)$ gets closer to the exponential distribution $\mathcal{E}(\theta)$ and, therefore, it is expected that, for fixed n and $\alpha \rightarrow 1$, approximation by a gamma distribution would be a more appropriate solution. Here, we observe that the distribution of the sum $S_n^Y = \sum_{i=1}^n Y_i$ is indeed the gamma distribution with parameters n and $1/\mu_X$ (denoted by $\mathcal{G}(n, 1/\mu_X)$), and, thus, exploiting the results of Corollary 1, we may obtain error estimates for the distances $\rho_K(S_n^X, \mathcal{G}(n, 1/\mu_X))$ and $\rho_{TV}(S_n^X, \mathcal{G}(n, 1/\mu_X))$ for fixed values of n .

First of all we need to find the norms $\|f^{(1)}\|_\infty$ and $\|f^{(2)}\|_1$ of the density f relative to the gamma distribution. Tedious calculations lead to the following general result.

Lemma 3. *Let f denote the density of the gamma distribution $\mathcal{G}(v, \lambda)$, i.e.*

$$f(x) = \frac{\lambda^v}{\Gamma(v)} x^{v-1} e^{-\lambda x}, \quad x > 0, v > 0, \lambda > 0.$$

Then, for $v \geq 3$,

$$\|f^{(1)}\|_\infty = \frac{\lambda^2 \sqrt{v-1}}{(v-1)!} b_v^{v-2} \exp\{-b_v\}$$

and

$$\|f^{(2)}\|_1 = \frac{2\lambda^2 \sqrt{v-1}}{(v-1)!} (b_v^{v-2} \exp\{-b_v\} + c_v^{v-2} \exp\{-c_v\}),$$

where

$$b_v := v - 1 - \sqrt{v-1}, \quad c_v := v - 1 + \sqrt{v-1}.$$

We will apply Corollary 1 to obtain the desired error bounds. We have

$$|\sigma_Y^2 - \sigma_X^2| = \left| 2\mu_X^2 - \frac{1}{\theta^2} \Gamma\left(1 + \frac{2}{\alpha}\right) \right| = 2\mu_X^2 \left| 1 - \frac{\Gamma(1 + 2/\alpha)}{2\Gamma^2(1 + 1/\alpha)} \right|. \tag{21}$$

Furthermore, we need to evaluate the distances $\rho_{TV} = \rho_{TV}(\mathcal{W}(\alpha, \theta), \mathcal{E}(1/\mu_X))$ and $\rho_K = \rho_K(\mathcal{W}(\alpha, \theta), \mathcal{E}(1/\mu_X))$. Using (19) and (20), we obtain

$$\begin{aligned} \rho_{TV}\left(\mathcal{W}(\alpha, \theta), \mathcal{E}\left(\frac{1}{\mu_X}\right)\right) &= \frac{1}{2} \int_0^\infty |h(x) - g(x)| \, dx \\ &= \frac{1}{2} \int_0^{r_1} (g(x) - h(x)) \, dx + \frac{1}{2} \int_{r_1}^{r_2} (h(x) - g(x)) \, dx \\ &\quad + \frac{1}{2} \int_{r_2}^\infty (g(x) - h(x)) \, dx \\ &= \exp\left\{-\frac{\theta r_2}{\Gamma(1 + 1/\alpha)}\right\} - \exp\left\{-\frac{\theta r_1}{\Gamma(1 + 1/\alpha)}\right\} \\ &\quad + \exp\{-(\theta r_1)^\alpha\} - \exp\{-(\theta r_2)^\alpha\}, \end{aligned} \tag{22}$$

where r_1 and r_2 are the roots of the nonlinear equation

$$(\theta x)^\alpha - \frac{\theta x}{\Gamma(1 + 1/\alpha)} = (\alpha - 1) \log x + \log\left(\alpha \theta^{\alpha-1} \Gamma\left(1 + \frac{1}{\alpha}\right)\right)$$

with respect to x . In addition, using standard calculus, it can be verified that

$$\rho_K = \max\{\varphi(r_1), \varphi(r_2)\}, \tag{23}$$

where

$$\varphi(x) := \left| \exp\left\{-\frac{\theta x}{\Gamma(1 + 1/\alpha)}\right\} - \exp\{-(\theta x)^\alpha\} \right|.$$

So, finally, applying Corollary 1 and using (21) and Lemma 3, we immediately derive the following error bounds.

Proposition 1. *Let X_1, X_2, \dots, X_n be i.i.d. observations from the Weibull distribution with density (19). For $n \geq 4$,*

$$\rho_{TV}\left(S_n^X, \mathcal{G}\left(n, \frac{1}{\mu_X}\right)\right) \leq B_{TV}^{\mathcal{G}} := \left(2\rho_{TV}^2 + 2\gamma_{TV} + \frac{\gamma_{TV}}{2\rho_{TV}}\right)(1 + 2\rho_{TV})^{n-2} - \frac{\gamma_{TV}}{2\rho_{TV}}$$

and

$$\rho_K\left(S_n^X, \mathcal{G}\left(n, \frac{1}{\mu_X}\right)\right) \leq B_K^{\mathcal{G}} := \left(2\rho_{TV}\rho_K + 2\gamma_K + \frac{\gamma_K}{2\rho_{TV}}\right)(1 + 2\rho_{TV})^{n-2} - \frac{\gamma_K}{2\rho_{TV}},$$

where

$$\begin{aligned} \gamma_{TV} &= \frac{\sqrt{n-2}}{(n-2)!} (b_{n-1}^{n-3} \exp\{-b_{n-1}\} + c_{n-1}^{n-3} \exp\{-c_{n-1}\}) \left| 1 - \frac{\Gamma(1 + 2/\alpha)}{2\Gamma^2(1 + 1/\alpha)} \right|, \\ \gamma_K &= \frac{\sqrt{n-2}}{(n-2)!} b_{n-1}^{n-3} \exp\{-b_{n-1}\} \left| 1 - \frac{\Gamma(1 + 2/\alpha)}{2\Gamma^2(1 + 1/\alpha)} \right|, \\ b_{n-1} &= n - 2 - \sqrt{n-2}, \quad c_{n-1} := n - 2 + \sqrt{n-2}, \end{aligned}$$

and ρ_{TV} and ρ_K are given by (22) and (23), respectively.

Remark 3. Using mathematical packages, we can easily compute the exact values of ρ_{TV} and ρ_K .

TABLE 1.

n = 5			n = 10			n = 20		
α	B_K^θ	B_{TV}^θ	α	B_K^θ	B_{TV}^θ	α	B_K^θ	B_{TV}^θ
0.90	0.098 077	0.139 942	0.90	0.081 772	0.129 262	0.90	0.127 408	0.210 300
0.95	0.039 789	0.056 967	0.95	0.027 896	0.044 639	0.95	0.030 055	0.050 972
0.97	0.022 029	0.031 578	0.97	0.014 502	0.023 306	0.97	0.013 784	0.023 578
0.99	0.006 788	0.009 741	0.99	0.004 211	0.006 795	0.99	0.003 579	0.006 166
1.01	0.006 567	0.009 425	1.01	0.004 074	0.006 573	1.01	0.003 460	0.005 961
1.03	0.019 932	0.028 575	1.03	0.013 092	0.021 043	1.03	0.012 381	0.021 182
1.05	0.033 625	0.048 157	1.05	0.023 423	0.037 496	1.05	0.024 863	0.042 189
1.10	0.069 440	0.099 209	1.10	0.056 323	0.089 198	1.10	0.082 025	0.135 868
1.15	0.107 745	0.153 609	1.15	0.102 214	0.160 120	1.15	0.208 963	0.337 398

Remark 4. For fixed n and θ , when $\alpha \rightarrow 1$, the aforementioned error estimates tend to 0, as expected.

In Table 1, numerical values for the bounds B_{TV}^θ and B_K^θ of Proposition 1 are presented for several values of α and n . The bounds are not affected by the value of θ because the distances ρ_K and ρ_{TV} are stable under scale changes and, thus, for convenience, we chose $\theta = 1$. Obviously, we can see that, for fixed n , the numerical error estimates tend to 0 when $\alpha \rightarrow 1$, as expected. The bounds were computed using MATHEMATICA® software. We also note that, for the calculations, the exact numerical values of the distances ρ_{TV} and ρ_K were used.

3.2. Distance between sums of NBUE or NWUE i.i.d. observations and a gamma random variable

The Weibull distribution considered in Subsection 3.1 belongs to specific classes of distributions provided by the following definition (cf. Barlow and Proschan (1975)).

Definition 2. The nonnegative random variable X (and its distribution function F) is said to be NBUE if, for all $t \geq 0$,

$$\frac{\int_t^\infty \bar{F}(x) dx}{\bar{F}(t)} \leq E(X),$$

and it is said to be NWUE if, for all $t \geq 0$,

$$\frac{\int_t^\infty \bar{F}(x) dx}{\bar{F}(t)} \geq E(X),$$

where $\bar{F}(x) = 1 - F(x)$.

We can verify that the Weibull distribution is NWUE if $\alpha < 1$ and NBUE for $\alpha > 1$.

The following result (see Theorem 3.A.55 of Shaked and Shanthikumar (2007)) is relevant.

Theorem 2. *If X is an NBUE or NWUE random variable with mean μ and Y is an exponentially distributed random variable with parameter $1/\mu$, then*

$$X \leq_{cx} Y \text{ or, respectively, } X \geq_{cx} Y.$$

Assume now that X_1, X_2, \dots, X_n are i.i.d. observations from a distribution which is NBUE (or NWUE). Let μ_X and σ_X^2 denote the mean and variance, respectively, of the distribution. Assume further that Y_1, Y_2, \dots, Y_n are i.i.d. observations from the exponential distribution with parameter $1/\mu_X$. We can easily verify that

$$|\sigma_Y^2 - \sigma_X^2| = \mu_X^2 |1 - C_X^2|, \tag{24}$$

where $C_X = \sigma_X/\mu_X$ denotes the coefficient of variation of the random variable X .

In what follows, for notational simplicity, we set

$$A_X := \frac{1}{2} |1 - C_X^2|, \tag{25}$$

$$m_n := \frac{\sqrt{n-2}}{(n-2)!} b_{n-1}^{n-3} \exp\{-b_{n-1}\}, \tag{26}$$

$$h_n := \frac{\sqrt{n-2}}{(n-2)!} (b_{n-1}^{n-3} \exp\{-b_{n-1}\} + c_{n-1}^{n-3} \exp\{-c_{n-1}\}), \tag{27}$$

where $b_{n-1} = n - 2 - \sqrt{n-2}$ and $c_{n-1} = n - 2 + \sqrt{n-2}$ (recall Lemma 3).

In view of (24) and Lemma 3, Corollary 1 provides the following gamma approximation bounds for $n \geq 4$:

$$\rho_{TV} \left(S_n^X, \mathcal{G} \left(n, \frac{1}{\mu_X} \right) \right) \leq \left(2\rho_{TV}^2 + 2h_n A_X + \frac{h_n A_X}{2\rho_{TV}} \right) (1 + 2\rho_{TV})^{n-2} - \frac{h_n A_X}{2\rho_{TV}}, \tag{28}$$

$$\rho_K \left(S_n^X, \mathcal{G} \left(n, \frac{1}{\mu_X} \right) \right) \leq \left(2\rho_{TV} \rho_K + 2m_n A_X + \frac{m_n A_X}{2\rho_{TV}} \right) (1 + 2\rho_{TV})^{n-2} - \frac{m_n A_X}{2\rho_{TV}}, \tag{29}$$

where A_X, h_n , and m_n are given by (25), (26), and (27), respectively.

The upper bounds for both distances in (28) and (29) are expressed in terms of the total variation distance ρ_{TV} and the Kolmogorov distance ρ_K . We may use a similar approach to the one described in Subsection 3.1 in order to determine the exact values of ρ_{TV} and ρ_K . However, for NBUE or NWUE random variables, upper bounds for ρ_K were obtained in Daley (1988, Theorems 1 and 2). In particular, Daley showed that if X is NBUE and Y is exponentially distributed with the same mean, then

$$\rho_K(X, Y) \leq 1 - \exp\{-\sqrt{2A_X}\}, \tag{30}$$

whereas, if X is NWUE,

$$\rho_K(X, Y) \leq \sqrt{A_X^2 + 2A_X} - A_X. \tag{31}$$

Furthermore, as the following result shows, the total variation distance ρ_{TV} can be upper bounded by twice the quantities on the right-hand sides of (30) and (31) involving the coefficient of variation of X .

Lemma 4. *Let X be a random variable with mean μ_X and finite variance σ_X^2 . Also, let Y be an exponential random variable with parameter $1/\mu_X$. If X is NBUE then*

$$\rho_{TV}(X, Y) \leq 2(1 - \exp\{-\sqrt{2A_X}\}),$$

and if X is NWUE then

$$\rho_{TV}(X, Y) \leq 2(\sqrt{A_X^2 + 2A_X} - A_X).$$

Proof. Let h and g denote the densities, and let H and G denote the CDFs of the random variables X and Y , respectively. Since X is NBUE or NWUE, by Theorem 2, the distributions of X and Y are ordered according to the convex order, and, therefore, their densities cross each other at exactly two points, say r_1 and r_2 . Then,

$$\begin{aligned} \rho_{TV}(X, Y) &= \frac{1}{2} \int_0^\infty |h(x) - g(x)| \, dx \\ &= \frac{1}{2} \int_0^{r_1} (g(x) - h(x)) \, dx + \frac{1}{2} \int_{r_1}^{r_2} (h(x) - g(x)) \, dx \\ &\quad + \frac{1}{2} \int_{r_2}^\infty (g(x) - h(x)) \, dx \\ &= G(r_1) - G(r_2) - H(r_1) + H(r_2) \\ &= \exp\left\{-\frac{r_2}{\mu_X}\right\} - \exp\left\{-\frac{r_1}{\mu_X}\right\} - H(r_1) + H(r_2) \\ &= \exp\left\{-\frac{r_2}{\mu_X}\right\} - R_X(r_2) + R_X(r_1) - \exp\left\{-\frac{r_1}{\mu_X}\right\}, \end{aligned}$$

where R_X denotes the survival function of X .

The result now follows from the fact that if X is NBUE then, by Cheng and He (1988),

$$\sup_{t>0} \left| R_X(t) - \exp\left\{-\frac{t}{\mu_X}\right\} \right| \leq 1 - \exp\{-\sqrt{2A_X}\},$$

and if X is NWUE then, by He and Cheng (1987),

$$\sup_{t>0} \left| R_X(t) - \exp\left\{-\frac{t}{\mu_X}\right\} \right| \leq \sqrt{A_X^2 + 2A_X} - A_X.$$

This completes the proof.

Now, relations (28) and (29) in combination with (30), (31), and Lemma 4 yield the following error estimates (as a matter of fact, we use (16) and (17) due to the upper bounding of ρ_{TV} and ρ_K , which lead to exactly the same results).

Proposition 2. *Let X_1, X_2, \dots, X_n be i.i.d. observations from a distribution which is NBUE with mean μ_X . Then, for $n \geq 4$,*

$$\begin{aligned} \rho_{TV}\left(S_n^X, \mathcal{G}\left(n, \frac{1}{\mu_X}\right)\right) &\leq \left(8B_X^2 + 2h_n A_X + \frac{h_n A_X}{4B_X}\right)(1 + 4B_X)^{n-2} - \frac{h_n A_X}{4B_X}, \\ \rho_K\left(S_n^X, \mathcal{G}\left(n, \frac{1}{\mu_X}\right)\right) &\leq \left(4B_X^2 + 2m_n A_X + \frac{m_n A_X}{4B_X}\right)(1 + 4B_X)^{n-2} - \frac{m_n A_X}{4B_X}, \end{aligned}$$

where $B_X := 1 - \exp\{-\sqrt{2A_X}\}$ and A_X, h_n , and m_n are given by (25), (26), and (27), respectively.

Proposition 3. *Let X_1, X_2, \dots, X_n be i.i.d. observations from a distribution which is NWUE with mean μ_X . Then, for $n \geq 4$,*

$$\begin{aligned} \rho_{TV}\left(S_n^X, \mathcal{G}\left(n, \frac{1}{\mu_X}\right)\right) &\leq \left(8D_X^2 + 2h_n A_X + \frac{h_n A_X}{4D_X}\right)(1 + 4D_X)^{n-2} - \frac{h_n A_X}{4D_X}, \\ \rho_K\left(S_n^X, \mathcal{G}\left(n, \frac{1}{\mu_X}\right)\right) &\leq \left(4D_X^2 + 2m_n A_X + \frac{m_n A_X}{4D_X}\right)(1 + 4D_X)^{n-2} - \frac{m_n A_X}{4D_X}, \end{aligned}$$

where $D_X := (A_X^2 + A_X)^{1/2} - A_X$ and $A_X, h_n,$ and m_n are given by (25), (26), and (27), respectively.

The error estimates of Propositions 2 and 3 are useful both in theory and practice for fixed n , in cases where we cannot invoke central limit theory to approximate the distribution of the sum S_n^X by the normal distribution.

Remark 5. As the coefficient of variation of X tends to 1, A_X tends to 0 and, consequently, the upper bounds of Propositions 2 and 3 tend to 0. This is reasonable, given that the coefficient of variation of an exponential random variable is equal to 1.

Remark 6. Clearly, Propositions 2 and 3 provide bounds for the Weibull distribution considered in Subsection 3.1 with

$$A_X = \left| 1 - \frac{\Gamma(1 + 2/\alpha)}{2\Gamma^2(1 + 1/\alpha)} \right|.$$

3.3. Distance between the distribution of the sum of i.i.d. Student random variables and a normal distribution

Assume that T_1, T_2, \dots, T_n are i.i.d. random variables following the Student t -distribution with k degrees of freedom, $k > 2$. In this case we have $\mu_T = 0$ and $\sigma_T^2 = k/(k - 2)$. Furthermore, let Z_1, Z_2, \dots, Z_n be i.i.d. normally distributed random variables with $\mu_Z = 0$ and $\sigma_Z^2 = 1$. Obviously, the sum S_n^Z follows the normal distribution $N(0, n)$ and, consequently, $S_n^Z/\sqrt{n} \sim N(0, 1)$. We also have $T_i \geq_{cx} Z_i$ for all $i = 1, 2, \dots, n$.

Let f_{n-1} denote the density of $N(0, n - 1)$. Clearly,

$$|\sigma_Z^2 - \sigma_T^2| = \frac{k}{k - 2} - 1 = \frac{2}{k - 2}, \quad k > 2.$$

In addition, the norms $\|f_{n-1}^{(1)}\|_\infty$ and $\|f_{n-1}^{(2)}\|_1$ of the normal distribution $N(0, n - 1)$ are given by

$$\|f_{n-1}^{(1)}\|_\infty = \frac{1}{(n - 1)\sqrt{2\pi e}} \quad \text{and} \quad \|f_{n-1}^{(2)}\|_1 = \frac{4}{(n - 1)\sqrt{2\pi e}}. \tag{32}$$

Let h and ϕ denote the densities of T_i and Z_i , respectively. Then h and ϕ cross each other at exactly two (symmetric) points, say $-r$ and r , since $T_i \geq_{cx} Z_i$. Thus, it is easy to see that

$$\begin{aligned} \rho_{TV} &= \rho_{TV}(T_i, N(0, 1)) \\ &= \frac{1}{2} \int_{-\infty}^{\infty} |h(t) - \phi(t)| dt \\ &= \frac{1}{2} \int_{-\infty}^{-r} (h(t) - \phi(t)) dt + \frac{1}{2} \int_{-r}^r (\phi(t) - h(t)) dt + \frac{1}{2} \int_r^{\infty} (h(t) - \phi(t)) dt \\ &= 2(F_T(r) - \Phi(r)), \end{aligned}$$

where F_T is the CDF of the Student t -distribution with k degrees of freedom and Φ is the CDF of $N(0, 1)$.

Here, it is important to note that Papadatos and Papathanasiou (1995, relation (5.2)) provided the following simple upper bound for the distance $\rho_{TV}(T_i, N(0, 1))$:

$$\rho_{TV}(T_i, N(0, 1)) \leq \frac{4}{k - 2}, \tag{33}$$

which establishes that the rate of convergence is at least $O(1/k)$. Therefore, since ρ_K and ρ_{TV} are stable under scale changes and by taking into consideration (16), (18), Corollary 1, (32), and (33), we obtain the following error estimates.

Proposition 4. *Let T_1, T_2, \dots, T_n be i.i.d. random variables which follow the Student t -distribution with k degrees of freedom, $k > 2$. For $n \geq 2$, we have*

$$\begin{aligned} \rho_{TV}\left(\frac{S_n^T}{\sqrt{n}}, N(0, 1)\right) &\leq B_{TV}^{\mathcal{N}} \\ &:= \left(\frac{32}{(k-2)^2} + \frac{1+16(k-2)^{-1}}{4(n-1)\sqrt{2\pi e}}\right)\left(1 + \frac{8}{k-2}\right)^{n-2} \\ &\quad - \frac{1}{4(n-1)\sqrt{2\pi e}}, \\ \rho_K\left(\frac{S_n^T}{\sqrt{n}}, N(0, 1)\right) &\leq B_K^{\mathcal{N}} \\ &:= \left(\frac{32}{(k-2)^2} + \frac{1+16(k-2)^{-1}}{8(n-1)\sqrt{2\pi e}}\right)\left(1 + \frac{8}{k-2}\right)^{n-2} \\ &\quad - \frac{1}{8(n-1)\sqrt{2\pi e}}. \end{aligned}$$

We note that in Proposition 4, for the distance ρ_K , we used $\delta = \min\{\rho_{TV}, 3M^{2/3}\zeta^{1/3}\}$ with

$$M = \sup_{y \in \mathbb{R}} f(y) = \sup_{z \in \mathbb{R}} \left(\frac{\exp\{-z^2/2\}}{\sqrt{2\pi}}\right) = \frac{1}{\sqrt{2\pi}}$$

(recall (8)). Thus, we obtain $\delta \leq \min\{4/(k-2), 3/(2\pi(k-2))^{1/3}\}$ and the minimum is always equal to $4/(k-2)$ for $k \geq 6$.

Theorem 2.1 of Cacoullos *et al.* (1997) combined with Theorem 2.1 of Cacoullos *et al.* (2001) yields the following upper bound for the total variation distance:

$$\begin{aligned} \rho_{TV}\left(\frac{S_n^T}{\sqrt{n}}, N(0, 1)\right) &\leq B_{CPP} \\ &:= \frac{3}{\sqrt{2n(k-1)(k-4)}} + \frac{4}{k-2} \end{aligned} \tag{34}$$

$$\sim \left(\frac{2.12132}{\sqrt{n}} + 4\right)\frac{1}{k} \tag{35}$$

for fixed n and $k \rightarrow \infty$ (the notation $a_k \sim b_k$ implies that $\lim_{k \rightarrow \infty}(a_k/b_k) = 1$). The bounds $B_{TV}^{\mathcal{N}}$ and $B_K^{\mathcal{N}}$ provided by Proposition 4 are asymptotically equal to

$$B_{TV}^{\mathcal{N}} \sim \frac{2n}{(n-1)\sqrt{2\pi e}} \frac{1}{k}, \tag{36}$$

$$B_K^{\mathcal{N}} \sim \frac{n}{(n-1)\sqrt{2\pi e}} \frac{1}{k},$$

for fixed n and $k \rightarrow \infty$. Both error estimates (35) and (36) are of order $O(1/k)$, but, asymptotically, (36) is obviously smaller than (35) due to the existence of the constant $\sqrt{2\pi e} \simeq 4.13273$ in the denominator of (36).

TABLE 2.

$n = 3$			$n = 5$			$n = 10$		
k	B_{CPP}	$B_{\text{TV}}^{\mathcal{N}}$	k	B_{CPP}	$B_{\text{TV}}^{\mathcal{N}}$	k	B_{CPP}	$B_{\text{TV}}^{\mathcal{N}}$
25	0.228 468	0.120 412	30	0.177 406	0.122 136	35	0.141 875	0.216 775
30	0.187 460	0.083 342	35	0.150 434	0.084 298	40	0.123 166	0.139 503
35	0.158 937	0.062 061	40	0.130 582	0.062 309	45	0.108 817	0.097 159
40	0.137 949	0.048 610	45	0.115 359	0.048 372	50	0.097 463	0.071 707
45	0.121 859	0.039 502	50	0.103 316	0.038 952	55	0.088 254	0.055 300
50	0.109 130	0.033 007	55	0.093 549	0.032 263	60	0.080 636	0.044 132
55	0.098 810	0.028 186	60	0.085 470	0.027 325	65	0.074 228	0.036 191
60	0.090 273	0.024 491	65	0.078 675	0.023 562	70	0.068 764	0.030 342
65	0.083 094	0.021 584	70	0.072 882	0.020 620	80	0.059 939	0.022 455
70	0.076 972	0.019 247	80	0.063 525	0.016 355	90	0.053 122	0.017 501

In Table 2, numerical comparisons between the bounds $B_{\text{TV}}^{\mathcal{N}}$ of Proposition 4 and B_{CPP} of (34) are presented for several values of k and n . Obviously, for fixed n , both error estimates tend to 0 when $k \rightarrow \infty$, as expected. The bounds were computed using MATHEMATICA software. Here, it is important to note that, for the calculations, we used the exact values of the bounds given by Proposition 4 and (34), and not the asymptotic error estimates (35) and (36).

Acknowledgement

The authors are grateful to the anonymous referee, whose suggestions led to a much improved version of the manuscript.

References

- BARLOW, R. E. AND PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart and Winston, New York.
- BOUTSIKAS, M. V. AND VAGGELATOU, E. (2008). A new method for obtaining sharp compound Poisson approximation error estimates for sums of locally dependent random variables. Preprint.
- CACOULOS, T., PAPADATOS, N. AND PAPATHANASIOU, V. (1997). Variance inequalities for covariance kernels and applications to central limit theorems. *Theory Prob. Appl.* **42**, 195–201.
- CACOULOS, T., PAPADATOS, N. AND PAPATHANASIOU, V. (2001). An application of a density transform and the local limit theorem. *Theory Prob. Appl.* **46**, 803–810.
- CHENG, K. AND HE, Z. F. (1988). Exponential approximations in the classes of life distributions. *Acta Math. Appl. Sinica (English Ser.)* **4**, 234–244.
- DALEY, D. J. (1988). Tight bounds on the exponential approximation of some aging distributions. *Ann. Prob.* **16**, 414–423.
- HE, Z. AND CHENG, K. (1987). Exponential approximations for some classes of life distributions. In *Reliability Theory and Applications*, eds S. Osaki *et al.*, World Scientific, Singapore, pp. 113–126.
- KAAS, R. (1993). How to and how not to compute stop-loss premiums in practice. *Insurance Math. Econom.* **13**, 241–254.
- MÜLLER, A. AND STOYAN, D. (2002). *Comparison Methods for Stochastic Models and Risks*. John Wiley, Chichester.
- PAPADATOS, N. AND PAPATHANASIOU, V. (1995). Distance in variation between two arbitrary distributions via the associated w -functions. *Theory Prob. Appl.* **40**, 685–694.
- RACHEV, S. T. (1991). *Probability Metrics and the Stability of Stochastic Models*. John Wiley, Chichester.
- RACHEV, S. T. AND RÜSCHENDORF, L. (1990). Approximation of sums by compound Poisson distributions with respect to stop-loss distances. *Adv. Appl. Prob.* **22**, 350–374.

- SHAKED, M. AND SHANTHIKUMAR, J. G. (2007). *Stochastic Orders*. Springer, New York.
- SZEKLI, R. (1995). *Stochastic Ordering and Dependence in Applied Probability* (Lecture Notes Statist. **97**). Springer, New York.
- VAGGELATOU, E. (2009). A new method for bounding the distance between sums of independent integer-valued random variables. To appear in *Methodology Comput. Appl. Prob.*
- ZOLOTAREV, V. M. (1983). Probability metrics. *Theory Prob. Appl.* **28**, 278–302.