

Penultimate gamma approximation in the CLT for skewed distributions

Michael V. Boutsikas

Department of Statistics and Insurance Science, University of Piraeus, Greece

Abstract

Sharp upper bounds are offered for the total variation distance between the distribution of a sum of n independent random variables following a skewed distribution with an absolutely continuous part, and an appropriate shifted gamma distribution. These bounds vanish at a rate $O(n^{-1})$ as $n \rightarrow \infty$ while the corresponding distance to the normal distribution vanishes at a rate $O(n^{-1/2})$, implying that, for skewed summands, pre-asymptotic (penultimate) gamma approximation is much more accurate than the usual normal approximation. Two illustrative examples concerning lognormal and Pareto summands are presented along with numerical comparisons confirming the aforementioned ascertainment.

Abbreviated Title: Penultimate gamma approximation.

Key words and phrases: Central limit theorem, gamma approximation, lognormal sums, rates of convergence, Pareto sums, total variation distance, quantile approximation.

AMS 2010 subject classification: Primary: 60F05
Secondary: 60G50; 62E17; 60E15

1 Introduction

The Central Limit Theorem (CLT) plays a significant role in probability and statistics providing a very simple approximation for the unknown or complicated distribution of a sum of n independent random variables (r.v.'s). It is well known (via asymptotic expansions, e.g. see Petrov (1995)) that the corresponding rate of convergence to the normal distribution is of order $O(n^{-1/2})$ for skewed summands, while it is of order at least $O(n^{-1})$ for smooth symmetric ones. This is due to the fact that the approximation error is of order $O(n^{-(r-1)/2})$ when the standardized distribution of the summands possesses an absolutely continuous part and has the same r ($r \geq 2$) first moments with the standard normal distribution $\mathcal{N}$. Therefore, when X_i 's are smooth and non symmetric, it seems natural to pursue a more accurate than normal approximation by employing an appropriate (preferably infinitely divisible) distribution which has the same first three moments as the summands, i.e. belongs to a class of distributions with location, scale and shape parameters. Perhaps the most tractable such distribution is the (shifted) gamma distribution. A similar observation was also made by Hall (1983) who derived (via Edgeworth expansions) asymptotic expansions for the distribution function of the sum of n independent r.v.'s using upper critical points of the chi-square distribution.

The purpose of this paper is to investigate the above idea of pre-asymptotic (penultimate) approximation in the CLT by constructing sharp recursive, closed form and asymptotic bounds for the total variation distance between the distribution of the sum of n independent r.v.'s and an appropriate shifted gamma distribution. More specifically, let $X, X_1, X_2, \dots$ be independent r.v.'s following a

skewed distribution F , with $\mathbb{E}X = 0, \mathbb{E}X^2 = 1, \mathbb{E}X^3 = 2/\lambda, \mathbb{E}X^4 < \infty$ for some $\lambda > 0$. The case of negatively skewed X_i 's can be treated by considering the $-X_i$'s. Assume also that F has an absolutely continuous part. Denoting by $\mathcal{SG}_\theta$ the (shifted by $-\theta$) gamma distribution with scale parameter θ and shape parameter θ^2 , we show that the total variation distance, $\mathbf{d}(F_n, \mathcal{SG}_{\sqrt{n}\lambda})$, between the distribution F_n of the standardized sum $n^{-1/2} \sum_{i=1}^n X_i$ and $\mathcal{SG}_{\sqrt{n}\lambda}$, is bounded above by a closed form upper bound $\mathbf{D}_n$ (cf. Theorem 4) which is of order $O(n^{-1})$. It is further proved (cf. Theorem 5 and (13)) that for a large class of distributions of the summands, the same distance is asymptotically bounded above by

$$\frac{\|\varphi^{(4)}\|}{16n} \left| 1 + \frac{(\mathbb{E}X^3)^2}{2} - \frac{\mathbb{E}X^4}{3} \right|, \quad (\|\varphi^{(4)}\| = 4e^{\frac{-3-\sqrt{6}}{2}} \frac{e^{\sqrt{6}} \sqrt{3-\sqrt{6}+\sqrt{3+\sqrt{6}}}}{\sqrt{\pi/3}} \approx 2.8)$$

where $\|\varphi^{(k)}\|$ denotes the L_1 norm of the k -th order derivative of the density φ of the standard normal law $\mathcal{N}$. Note that the corresponding distance between F_n and $\mathcal{N}$ is (cf. Sirazhdinov and Mamatov (1962))

$$\mathbf{d}(F_n, \mathcal{N}) \sim \frac{\|\varphi^{(3)}\|}{12\sqrt{n}} |\mathbb{E}X^3|, \quad (\|\varphi^{(3)}\| = \frac{2+8e^{-3/2}}{\sqrt{2\pi}} \approx 1.51) \quad (1)$$

(we write $a_n \sim b_n$ when $a_n/b_n \rightarrow 1$) and therefore F_n approaches $\mathcal{SG}_{\sqrt{n}\lambda}$ with an approximation error of order $O(n^{-1})$, while both $\mathcal{SG}_{\sqrt{n}\lambda}$ and F_n converge slower to $\mathcal{N}$, at a rate $O(n^{-1/2})$. Consequently, in practice it may be preferable to consider a gamma ($\mathcal{SG}_{\sqrt{n}\lambda}, \lambda = 2/\mathbb{E}X^3$) instead of normal approximation for standardized sums of skewed r.v.'s, especially when the summands follow heavy tailed distributions (but with finite first four moments), and for relatively small and moderate values of n , since for large n , the distributions $\mathcal{SG}_{\sqrt{n}\lambda}$ and $\mathcal{N}$ are practically identical.

It is worth mentioning that F_n could also be approximated by appropriate asymptotic expansions with the same $O(n^{-1})$ or higher order (cf. e.g. Petrov (1995)). The advantage of our approach is that we use a proper probability distribution ($\mathcal{SG}_{\sqrt{n}\lambda}$) and that we offer explicit (closed form) error bounds whereas the approximation error of the asymptotic expansions do not possess a closed form.

The idea of using the first three or four moments of a sum of r.v.'s in order to approximate its unknown skewed distribution has been employed by many researchers in diverse practical applications. Usually a three or four parameter distribution is proposed in order to fit given data by equating the sample and distribution moments. For example, in the actuarial literature various approximations have been proposed for the distribution of the aggregate claims in a nonlife insurance portfolio. These include shifted gamma, inverse Gaussian (IG), and gamma-IG (e.g. cf. Reijnen et al. (2005) and the references therein). The results of the present article theoretically justify the use of a shifted gamma distribution by relying on the first three moments of F_n . It seems that if we could exploit more than the first three moments, we could obtain results with even higher approximation order, but the whole analysis would then be much more complicated.

The organization of the paper is as follows. In the following Section 2 we present necessary notations and auxiliary results, concerning properties (i) of specific probability metrics and (ii) of the standardized gamma distribution. In Section 3 we offer recursive, closed form and asymptotic upper bounds for the total variation distance of interest. Our methodology is based on elementary properties of probability metrics and certain smoothing inequalities (cf. Lemma 1). In the second part of this section we demonstrate how the shifted gamma distribution can be used for quantile approximations, and include an example concerning the noncentral chi-squared distribution. Finally, in Section 4, we present two illustrative examples concerning sums of lognormal and Pareto r.v.'s, along with numerical comparisons that reveal the applicability and the performance of our main results.

Note that all the results are formulated using the total variation distance $\mathbf{d}$, but with a proper modification of our methodology, analogous results could also be established for other distances, e.g. for the Kolmogorov distance ρ . Nevertheless, since $\rho \leq \mathbf{d}$, all the upper bounds for $\mathbf{d}$ are also upper bounds for ρ .

2 Preliminaries - Notations

(i) Probability metrics. We denote by $\mathcal{L}Y$, F_Y and f_Y the distribution, the cumulative distribution function (c.d.f.) and the probability density function (p.d.f.) respectively of a r.v. Y . The *total variation* and *Zolotarev's distance* (ζ -metrics of order $s \in \mathbb{N}$) between the distributions of two r.v.'s W, Y are defined respectively by

$$\begin{aligned} \mathbf{d}(\mathcal{L}W, \mathcal{L}Y) &= \sup_{A \in \mathcal{B}(\mathbb{R})} |P(W \in A) - P(Y \in A)| = \frac{1}{2} \int_{\mathbb{R}} |f_W(x) - f_Y(x)| dx, \\ \zeta_s(\mathcal{L}W, \mathcal{L}Y) &= \frac{1}{(s-1)!} \int_{-\infty}^{\infty} |\mathbb{E}(W-t)_+^{s-1} - \mathbb{E}(Y-t)_+^{s-1}| dt, \end{aligned} \quad (2)$$

where $\mathcal{B}(\mathbb{R})$ denotes the Borel sets of $\mathbb{R}$, and $y_+^s \stackrel{\text{def}}{=} (\max\{y, 0\})^s$. The second equality in (2) holds true when the densities f_W, f_Y exist. It follows readily from (2) that $\mathbf{d}(\mathcal{L}g(W), \mathcal{L}g(Y)) \leq \mathbf{d}(\mathcal{L}W, \mathcal{L}Y)$ for every measurable g , and thus the law of $g(W)$ approximates the law of $g(Y)$ with (at least) the same accuracy as the law of W approximates the law of Y . Note also that the distance $\zeta_s(\mathcal{L}W, \mathcal{L}Y)$ can be used only when $\mathbb{E}W^i = \mathbb{E}Y^i$, $i = 1, 2, \dots, s-1$ (and in this case, it is bounded above by $(\mathbb{E}|W|^s + \mathbb{E}|Y|^s)/s!$) because otherwise it is not finite. In what follows, for simplicity, we shall write Y instead of $\mathcal{L}Y$ in $\mathbf{d}$ or ζ_s .

The metrics $\mathbf{d}$ and ζ_s possess the *regularity*, the *homogeneity* and the *subadditivity* property for independent summands (e.g., see Rachev (1991), p.264). Thus, for ζ_s it can be proved that,

$$\zeta_s(c \sum_{i=1}^n W_i, c \sum_{i=1}^n Y_i) \leq c^s \sum_{i=1}^n \zeta_s(W_i, Y_i), \quad n, s \in \mathbb{N}, \quad c > 0, \quad (3)$$

for independent W_i 's and independent Y_i 's. It is also worth mentioning that when the distributions of two r.v.'s W, Y are s -convex ordered, the distance $\zeta_s(W, Y)$ takes on a very simple form. Specifically, if $W \leq_{s-cx} Y$, for some $s \in \mathbb{N}$ (i.e. $\mathbb{E}\phi(W) \leq \mathbb{E}\phi(Y)$ for all regular s -convex functions ϕ for which the expectations exist; cf. Denuit, Lefèvre and Shaked (1998)), then (cf. Boutsikas and Vaggelatos (2002)),

$$\zeta_s(W, Y) = \frac{\mathbb{E}Y^s - \mathbb{E}W^s}{s!}. \quad (4)$$

The following Lemma 1 offers two smoothing inequalities for the metric $\mathbf{d}$. Part (a) has been employed in the past for the derivation of Berry-Esseen-type or Poisson approximation results. Its proof can be found in Rachev (1991), p.274 (see also Boutsikas and Vaggelatos (2009) for a more general result). Part (b) was given by Boutsikas (2011) while similar inequalities can be found in a number of papers in the literature (e.g. see Rachev (1991), p. 325, Zolotarev (1983), p. 294 or Rachev and Ruschendorf (1990)). Recall that the L_1 norm of a function f is $\|f\| = \int_{-\infty}^{\infty} |f(x)| dx$.

Lemma 1 (a) *If the r.v.'s Y, Y' are independent of the r.v.'s W, W' then*

$$\mathbf{d}(W + Y, W + Y') \leq 2\mathbf{d}(Y, Y')\mathbf{d}(W, W') + \mathbf{d}(W' + Y, W' + Y').$$

(b) Let Z be a r.v., independent of the r.v.'s W, Y , with Lebesgue density f_Z , k -times differentiable on $\mathbb{R}$, and denote by $f_Z^{(k)}$ its k -th order derivative ($f^0 = f$). If $\zeta_k(W, Y) < \infty$, $k \geq 1$, then,

$$\mathbf{d}(W + Z, Y + Z) \leq \frac{1}{2} \left\| f_Z^{(k)} \right\| \zeta_k(W, Y).$$

(ii) **The shifted gamma distribution.** We denote by $\mathcal{G}(\kappa, \lambda, c)$ the shifted gamma distribution with p.d.f.,

$$f_{\mathcal{G}(\kappa, \lambda, c)}(x) = \frac{\lambda^\kappa}{\Gamma(\kappa)} (x - c)^{\kappa-1} e^{-\lambda(x-c)}, \quad x \geq c,$$

where $\kappa > 0, \lambda > 0$ and $c \in \mathbb{R}$ are the shape, scale and location parameters respectively. Obviously, the ordinary gamma distribution simply is $\mathcal{G}(\kappa, \lambda, 0) \equiv \mathcal{G}(\kappa, \lambda)$ and if a r.v. $G \sim \mathcal{G}(\kappa, \lambda)$ then $Y = G + c \sim \mathcal{G}(\kappa, \lambda, c)$ and for $k = 1, 2, \dots$,

$$\mathbb{E}G^k = \frac{\prod_{j=0}^{k-1} (\kappa + j)}{\lambda^k}, \quad \mathbb{E}Y^k = \mathbb{E}(G + c)^k = c^k + \sum_{i=1}^k \binom{k}{i} \frac{\prod_{j=0}^{i-1} (\kappa + j)}{\lambda^i} c^{k-i}.$$

It is easy to verify that, if the independent r.v.'s $Y_1, Y_2, \dots, Y_n \sim \mathcal{G}(\kappa, \lambda, c)$ then the sum $S = \sum_{i=1}^n (bY_i + d)$ for $b > 0, d \in \mathbb{R}$, follows the shifted gamma distribution $\mathcal{G}(n\kappa, \lambda/b, n(d + bc))$. We observe that, if $X \sim \mathcal{G}(\lambda^2, \lambda, -\lambda)$ then

$$\mathbb{E}X = 0, \mathbb{E}X^2 = 1, \mathbb{E}X^3 = \frac{2}{\lambda}, \mathbb{E}X^4 = 3 + \frac{6}{\lambda^2}.$$

We call the above distribution *standardized gamma* with parameter λ and denote it by $\mathcal{SG}_\lambda \equiv \mathcal{G}(\lambda^2, \lambda, -\lambda)$. It can be easily verified through the CLT that $\mathcal{SG}_\lambda$ converges to the standard normal distribution, $\mathcal{N}$, as $\lambda \rightarrow \infty$.

We also denote by $\varphi(x) = (2\pi)^{-1/2} e^{-x^2/2}$ the standard normal density, and by $f_{\mathcal{G}(\kappa, \lambda, c)}^{(k)}$ and $\varphi^{(k)}$ the k -th order derivative ($k = 1, 2, \dots$) of $f_{\mathcal{G}(\kappa, \lambda, c)}$ and φ respectively. In the next lemma we offer a simple in form upper bound for the norm $\left\| f_{\mathcal{G}(\kappa, \lambda, c)}^{(4)} \right\| = \left\| f_{\mathcal{G}(\kappa, \lambda)}^{(4)} \right\|$ and derive its asymptotic relation with the norm $\left\| \varphi^{(4)} \right\| \approx 2.80061$.

Lemma 2 (a) $\left\| f_{\mathcal{G}(\kappa, \lambda)}^{(4)} \right\| < \frac{6}{\kappa^2} \lambda^4, \kappa \geq 9$. (b) $\kappa^2 \left\| f_{\mathcal{G}(\kappa, \lambda)}^{(4)} \right\| \rightarrow \left\| \varphi^{(4)} \right\| \lambda^4$ as $\kappa \rightarrow \infty$.

Proof. see Appendix. ■

Throughout, we shall assume that $\sum_{i=j}^k a_i = 0$ when $j > k$.

3 Main Results

3.1 Bounds for the total variation distance

Let $X, X_1, X_2, \dots$ be a sequence of independent and identically distributed (i.i.d.) r.v.'s with $\mathbb{E}X = 0, \mathbb{E}X^2 = 1, \mathbb{E}X^3 = 2/\lambda > 0, \mathbb{E}X^4 < \infty$. The restriction of a positively skewed distribution for X_i 's could be relaxed to $\mathbb{E}X^3 \neq 0$ because when $\mathbb{E}X^3 < 0$ we can work with $-X_i$'s. Denote by

$$\mathbf{d}_n = \mathbf{d}\left(F_n, \mathcal{SG}_{\sqrt{n}\lambda}\right), \quad \lambda = \frac{2}{\mathbb{E}X^3},$$

the total variation distance between the distribution F_n of $\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ and the standardized gamma distribution $\mathcal{SG}_{\sqrt{n}\lambda}$. Note that $\mathcal{SG}_{\sqrt{n}\lambda}$ has the same first three moments as F_n . The next result offers a recursive bound for the above distance.

Lemma 3 *The total variation distance $\mathbf{d}_n$ between the distribution of $\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ and the standardized gamma distribution $\mathcal{SG}_{\sqrt{n}\lambda}$, satisfies the following recursive bound*

$$\mathbf{d}_n \leq \beta \sum_{j=2}^{n-1} \left\| f_{\mathcal{G}((n-j)\lambda^2, \lambda)}^{(4)} \right\| \mathbf{d}_{j-1} + \frac{n}{2} \left\| f_{\mathcal{G}((n-1)\lambda^2, \lambda)}^{(4)} \right\| \beta + 2\mathbf{d}_1 \mathbf{d}_{n-1} \quad (5)$$

where $\beta = \zeta_4(X, \mathcal{SG}_\lambda) < \infty$.

Proof. Let $Z_1, Z_2, \dots$ be a sequence of i.i.d. r.v.'s, independent also of X_i 's, following the standardized gamma distribution $\mathcal{SG}_\lambda$. From the triangle inequality we deduce that,

$$\mathbf{d}_n = \mathbf{d} \left(\sum_{i=1}^n X_i, \sum_{i=1}^n Z_i \right) \leq \sum_{j=1}^n \mathbf{d}(W_j + Y_j, W_j + Y'_j)$$

where $Y_j = X_j + \sum_{i=j+1}^n Z_i$, $Y'_j = \sum_{i=j}^n Z_i$, $W_j = \sum_{i=1}^{j-1} X_i$, $W'_j = \sum_{i=1}^{j-1} Z_i$. The r.v.'s Y_j, Y'_j are independent of W_j, W'_j and by employing Lemma 1(a) we get for $n \geq 2, j = 1, 2, \dots, n$,

$$\begin{aligned} \mathbf{d}_n &\leq \sum_{j=1}^n (2\mathbf{d}(Y_j, Y'_j) \mathbf{d}(W_j, W'_j) + \mathbf{d}(W'_j + Y_j, W'_j + Y'_j)) \\ &= 2 \sum_{j=2}^{n-1} \mathbf{d} \left(X_j + \sum_{i=j+1}^n Z_i, Z_j + \sum_{i=j+1}^n Z_i \right) \mathbf{d}_{j-1} + 2\mathbf{d}_1 \mathbf{d}_{n-1} + \sum_{j=1}^n \mathbf{d} \left(X_j + \sum_{i \in A_{n,j}} Z_i, \sum_{i=1}^n Z_i \right) \end{aligned} \quad (6)$$

where $A_{n,j} = \{1, \dots, n\} - \{j\}$. Invoking now Lemma 1(b) for $k = 4$, and taking also into account that $\sum_{i=j+1}^n Z_i \sim \mathcal{G}((n-j)\lambda^2, \lambda, -(n-j)\lambda)$, the distances appearing in the sums of (6) are bounded above as follows

$$\begin{aligned} \mathbf{d} \left(X_j + \sum_{i=j+1}^n Z_i, Z_j + \sum_{i=j+1}^n Z_i \right) &\leq \frac{1}{2} \left\| f_{\mathcal{G}((n-j)\lambda^2, \lambda)}^{(4)} \right\| \zeta_4(X_j, Z_j), \\ \mathbf{d} \left(X_j + \sum_{i \in A_{n,j}} Z_i, \sum_{i=1}^n Z_i \right) &\leq \frac{1}{2} \left\| f_{\mathcal{G}((n-1)\lambda^2, \lambda)}^{(4)} \right\| \zeta_4(X_j, Z_j). \end{aligned}$$

The above inequalities, combined with (6), readily lead to (5). Finally, since $\mathbb{E}X_j^i = \mathbb{E}Z_j^i$, $i = 1, 2, 3$, and $\mathbb{E}X_j^4 < \infty$ we have that

$$\beta = \zeta_4(X_j, Z_j) \leq \frac{\mathbb{E}X_j^4 + \mathbb{E}Z_j^4}{4!} = \frac{\mathbb{E}X_j^4 + 3 + \frac{6}{\lambda^2}}{4!} < \infty.$$

■

Employing Lemma 2(a), the above result induce that, for $\lambda \geq 3$, the total variation distance $\mathbf{d}_n$ satisfies the following recursive bound

$$\mathbf{d}_n \leq 6\beta \sum_{j=2}^{n-1} \frac{\mathbf{d}_{j-1}}{(n-j)^2} + 3\beta \frac{n}{(n-1)^2} + 2\mathbf{d}_1 \mathbf{d}_{n-1}, \quad n \geq 2. \quad (7)$$

Using the bounds (5) or (7) we can construct recursive upper bounds for $\mathbf{d}_n$ in terms of $\mathbf{d}_1 = \mathbf{d}(X, \mathcal{SG}_\lambda)$ and $\beta = \zeta_4(X, \mathcal{SG}_\lambda)$. For example from (7) we can get,

$$\mathbf{d}_2 \leq 6\beta + 2\mathbf{d}_1^2, \quad \mathbf{d}_3 \leq \left(\frac{9}{4} + 18\mathbf{d}_1\right)\beta + 4\mathbf{d}_1^3, \quad \mathbf{d}_4 \leq \left(6\mathbf{d}_1 + 36\beta + 48\mathbf{d}_1^2 + \frac{4}{3}\right)\beta + 8\mathbf{d}_1^4,$$

and so forth. If, for instance, $\mathbf{d}_1 \leq 0.05$ and $\beta \leq 0.02$, then $\mathbf{d}_5 \leq 0.0358$, $\mathbf{d}_{10} \leq 0.0134$, $\mathbf{d}_{30} \leq 0.0033$, and $\mathbf{d}_{100} \leq 0.0009$. The recursive bound for $\mathbf{d}_n$ that is derived via (7) is worse (i.e. larger) but computationally simpler than the one that follows from (5). Obviously, a closed form upper bound for $\mathbf{d}_n$ would be more convenient. Such a bound is offered (under certain conditions) in the next theorem by employing (7).

Theorem 4 *If $\mathbf{d}_1 = \mathbf{d}(X, \mathcal{SG}_\lambda) < 1/8$, $\lambda \geq 3$ and $\varepsilon = \frac{8\mathbf{d}_1}{1+\delta\mathbf{d}_1} + 27\beta < 1$, where $\delta = 296/45$, $\beta = \zeta_4(X, \mathcal{SG}_\lambda)$, then*

$$\mathbf{d}_n = \mathbf{d}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i, \mathcal{SG}_{\sqrt{n}\lambda}\right) \leq \frac{n}{(n-1)^2} \frac{3\beta}{(1-\varepsilon)} + \frac{1}{2}(2\mathbf{d}_1)^n \stackrel{\text{def}}{=} \mathbf{D}_n, \quad n \geq 2. \quad (8)$$

Proof. We shall use induction. For $n = 2$, inequality (7) leads to

$$\mathbf{d}_2 \leq 6\beta + 2\mathbf{d}_1^2 \leq \frac{6\beta}{1-\varepsilon} + 2\mathbf{d}_1^2 = \mathbf{D}_2 \quad (9)$$

and therefore (8) is valid for $n = 2$. Likewise, for $n = 3$, inequality (7) along with (9) yields

$$\mathbf{d}_3 \leq 6\beta\mathbf{d}_1 + 2\mathbf{d}_1\mathbf{d}_2 + \frac{9\beta}{4} \leq 6\beta\mathbf{d}_1 + 2\mathbf{d}_1(6\beta + 2\mathbf{d}_1^2) + \frac{9\beta}{4}$$

which is easily verified to be less than or equal to $\mathbf{D}_3$, and thus (8) is valid also for $n = 3$. Assume now that (8) holds for $n = 2, 3, \dots, m-1$ ($m \geq 4$). We shall show that it also holds for $n = m$. Invoking again (7) we get

$$\mathbf{d}_m \leq \frac{6\beta\mathbf{d}_1}{(m-2)^2} + 6\beta \sum_{j=3}^{m-1} \frac{\mathbf{d}_{j-1}}{(m-j)^2} + 2\mathbf{d}_1\mathbf{d}_{m-1} + \frac{3\beta m}{(m-1)^2},$$

and hence, using the induction assumption (i.e. replacing all $\mathbf{d}_n, n \leq m-1$ with their upper bound $\mathbf{D}_n$), we get

$$\begin{aligned} \mathbf{d}_m &\leq \frac{6\beta\mathbf{d}_1}{(m-2)^2} + \sum_{j=3}^{m-1} \frac{6\beta}{(m-j)^2} \left(\frac{j-1}{(j-2)^2} \frac{3\beta}{(1-\varepsilon)} + \frac{(2\mathbf{d}_1)^{j-1}}{2} \right) \\ &\quad + \frac{\mathbf{d}_1(m-1)}{(m-2)^2} \frac{6\beta}{1-\varepsilon} + \frac{1}{2}(2\mathbf{d}_1)^m + \frac{3\beta m}{(m-1)^2}, \quad m \geq 4, \end{aligned} \quad (10)$$

which is proved to be upper bounded by $\mathbf{D}_m$ as follows. Initially it will be shown that,

$$\sum_{j=3}^{m-1} \frac{(2\mathbf{d}_1)^{j-2}}{(m-j)^2} (1-\varepsilon) + \frac{(m-\varepsilon)}{(m-2)^2} \leq \frac{4m}{(m-1)^2(1+\delta\mathbf{d}_1)}. \quad (11)$$

For $m = 4$, the above inequality reduces to

$$2\mathbf{d}_1(1-\varepsilon) + \frac{4-\varepsilon}{4} \leq \frac{16}{9(1+\delta\mathbf{d}_1)}$$

which, taking also into account that $\varepsilon \geq \frac{8\mathbf{d}_1}{1+\delta\mathbf{d}_1}$, reduces to $1 + (2\delta - 16)\mathbf{d}_1^2 + \delta\mathbf{d}_1 \leq 16/9$. The latter is easily verified to be true for $\mathbf{d}_1 \leq 2^{-3}$. For $m \geq 5$, we invoke Lemma 19 of Boutsikas (2011) (for $k = 4$), according to which $\sum_{j=3}^{m-1} 4^{2-j}(m-j)^{-2} \leq 2m(m-1)^{-2}/5, m \geq 5$. By this inequality and the assumption $\mathbf{d}_1 \leq 2^{-3}$, we ascertain that the left part of inequality (11) is bounded above by

$$\begin{aligned} \sum_{j=3}^{m-1} \frac{(\frac{1}{4})^{j-2}}{(m-j)^2} (1 - \varepsilon) + \frac{m}{(m-2)^2} &\leq \frac{2m}{5(m-1)^2} \left(1 - \frac{8\mathbf{d}_1}{1 + \delta\mathbf{d}_1}\right) + \frac{m}{(m-2)^2} \\ &\leq \frac{(\frac{2}{5}(1 + (\delta - 8)\mathbf{d}_1) + \frac{16}{9}(1 + \delta\mathbf{d}_1))m}{(m-1)^2(1 + \delta\mathbf{d}_1)} \leq \frac{4m}{(m-1)^2(1 + \delta\mathbf{d}_1)}, \end{aligned}$$

and thus (11) is valid for all $m \geq 4$.

Next, applying the inequality $\sum_{j=3}^{m-1} \frac{(j-1)}{(m-j)^2(j-2)^2} \leq \frac{9}{2}m(m-1)^{-2}, m \geq 4$ (cf. Lemma 18 of Boutsikas (2011), for $k = 4$), we get that, for $m \geq 4$,

$$\frac{1}{1 - \varepsilon} \sum_{j=3}^{m-1} \frac{(j-1)6^2\beta^2}{2(m-j)^2(j-2)^2} + \frac{3\beta m}{(m-1)^2} + \frac{(2\mathbf{d}_1)^m}{2} \leq \frac{9}{1 - \varepsilon} \frac{m6^2\beta^2}{4(m-1)^2} + \frac{3\beta m}{(m-1)^2} + \frac{(2\mathbf{d}_1)^m}{2}. \quad (12)$$

Adding by parts (11) (after multiplying both sides by $6\beta\mathbf{d}_1/(1 - \varepsilon)$) and (12) we finally get that (10) is bounded above by $\mathbf{D}_m$ and the proof is completed. ■

Remark 1. It is worth mentioning that the distance $\beta = \zeta_4(X, \mathcal{SG}_\lambda)$ appearing in the above expressions can be easily evaluated when $X \leq_{4-cx} Y$ or $Y \leq_{4-cx} X$ (cf. (4)), where $Y \sim \mathcal{SG}_\lambda$, since in this case,

$$\beta = \frac{|\mathbb{E}Y^4 - \mathbb{E}X^4|}{4!} = \frac{1}{4!} \left| 3 + \frac{6}{\lambda^2} - \mathbb{E}X^4 \right| = \frac{1}{8} \left| 1 + \frac{(\mathbb{E}X^3)^2}{2} - \frac{\mathbb{E}X^4}{3} \right|. \quad (13)$$

In the following theorem we prove an asymptotic result for the total variation distance between the distribution of $n^{-1/2} \sum_{i=1}^n X_i$ and $\mathcal{SG}_{\sqrt{n}\lambda}$ that holds true when the common c.d.f. F of X_i 's has an absolutely continuous part. Its proof is based on Lemma 3 and Theorem 4 but the additional assumptions imposed in Theorem 4 are no longer necessary. Note that every c.d.f. F can be written in the form $F = pF_{ac} + (1 - p)F_s$, $p \in [0, 1]$ where F_s is a singular c.d.f. (discrete and/or singular continuous) and F_{ac} is an absolutely continuous c.d.f. A c.d.f. F has an absolutely continuous part when $p > 0$.

Theorem 5 *If the distribution of X_i 's has an absolutely continuous part, then*

$$\limsup_{n \rightarrow \infty} n \cdot \mathbf{d} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i, \mathcal{SG}_{\sqrt{n}\lambda} \right) \leq \frac{1}{2} \left\| \varphi^{(4)} \right\| \beta.$$

where $\beta = \zeta_4(X, \mathcal{SG}_\lambda)$.

Proof. Consider the sequence of i.i.d. r.v.'s $X_{i,r} = \frac{1}{\sqrt{r}} \sum_{j=(i-1)r+1}^{ir} X_j$, $i = 1, 2, \dots$ for some positive integer r . It is easy to verify that $\mathbb{E}X_{i,r} = 0, \mathbb{E}X_{i,r}^2 = 1, \mathbb{E}X_{i,r}^3 = 2/\lambda_r > 0, \mathbb{E}X_{i,r}^4 < \infty$ where $\lambda_r = \lambda r^{1/2}$. Applying (5) for the sequence of r.v.'s $X_{1,r}, X_{2,r}, \dots$ along with Lemma 2(a) we get

$$\mathbf{d}_{m,r} \leq 6\beta_r \sum_{j=2}^{m-1} \frac{\mathbf{d}_{j-1,r}}{(m-j)^2} + \left\| f_{\mathcal{G}((m-1)\lambda_r^2, \lambda_r)}^{(4)} \right\| \frac{m\beta_r}{2} + 2\mathbf{d}_{1,r}\mathbf{d}_{m-1,r}, \quad m \geq 2, \quad (14)$$

where, also using known properties of ζ_4 (cf. (3)),

$$\begin{aligned}\beta_r &= \zeta_4(X_{1,r}, \mathcal{SG}_{\lambda_r}) = \zeta_4\left(\frac{1}{\sqrt{r}} \sum_{i=1}^r X_i, \mathcal{SG}_{\sqrt{r}\lambda}\right) \leq \frac{\zeta_4(X_1, \mathcal{SG}_\lambda)}{r} = \frac{\beta}{r}, \\ \mathbf{d}_{m,r} &= \mathbf{d}\left(\frac{1}{\sqrt{m}} \sum_{i=1}^m X_{i,r}, \mathcal{SG}_{\sqrt{m}\lambda_r}\right) = \mathbf{d}\left(\frac{1}{\sqrt{mr}} \sum_{i=1}^{mr} X_i, \mathcal{SG}_{\sqrt{mr}\lambda}\right) = \mathbf{d}_{mr}, \quad m \geq 1.\end{aligned}$$

Moreover, applying Theorem 4 for the sequence $X_{1,r}, X_{2,r}, \dots$, we also get the inequality

$$\mathbf{d}_{n,r} \leq \frac{n}{(n-1)^2} \frac{3\beta_r}{(1-\varepsilon_r)} + \frac{1}{2}(2\mathbf{d}_{1,r})^n, \quad n \geq 2. \quad (15)$$

provided that $\varepsilon_r = \frac{8\mathbf{d}_{1,r}}{1+\delta\mathbf{d}_{1,r}} + 27\beta_r < 1$ ($\delta = 296/45$), $\lambda_r \geq 3$ and $\mathbf{d}_{1,r} < 1/8$. Combining now inequalities (14) and (15) we deduce, for $m \geq 3$,

$$\begin{aligned}\mathbf{d}_{m,r} &\leq \frac{6\beta_r \mathbf{d}_{1,r}}{(m-2)^2} + 6\beta_r \sum_{j=3}^{m-1} \frac{\frac{(j-1)3\beta_r}{(j-2)^2(1-\varepsilon_r)} + \frac{(2\mathbf{d}_{1,r})^{j-1}}{2}}{(m-j)^2} \\ &\quad + \left\| f_{\mathcal{G}((m-1)\lambda_r^2, \lambda_r)}^{(4)} \right\| \frac{m\beta_r}{2} + 2\mathbf{d}_{1,r} \left(\frac{(m-1)3\beta_r}{(m-2)^2(1-\varepsilon_r)} + \frac{(2\mathbf{d}_{1,r})^{m-1}}{2} \right).\end{aligned} \quad (16)$$

Applying the inequality $\sum_{j=3}^{m-1} \frac{(j-1)}{(m-j)^2(j-2)^2} \leq \frac{9}{2}m(m-1)^{-2}$, $m \geq 4$ (cf. Lemma 18 of Boutsikas (2011), for $k = 4$) and taking into account that $\beta_r = \beta/r$, $\mathbf{d}_{i,r} = \mathbf{d}_{ir}$, $\lambda_r = \lambda\sqrt{r}$, inequality (16) readily leads to

$$\begin{aligned}\mathbf{d}_{mr} &\leq \frac{6\mathbf{d}_r\beta}{r(m-2)^2} + \frac{81m\beta^2}{r^2(1-\varepsilon_r)}(m-1)^{-2} + 3\frac{\beta}{r} \sum_{j=3}^{m-1} \frac{(2\mathbf{d}_r)^{j-1}}{(m-j)^2} \\ &\quad + \left\| f_{\mathcal{G}((m-1)\lambda^2r, \lambda\sqrt{r})}^{(4)} \right\| \frac{m\beta}{2r} + \frac{6\mathbf{d}_r(m-1)\beta}{r(m-2)^2(1-\varepsilon_r)} + \frac{(2\mathbf{d}_r)^m}{2},\end{aligned} \quad (17)$$

for $m \geq 4, r \geq 1$, provided that $\varepsilon_r = \frac{8\mathbf{d}_r}{1+\delta\mathbf{d}_r} + 27\frac{\beta}{r} < 1$, $\lambda\sqrt{r} \geq 3$ and $\mathbf{d}_r < 1/8$.

Observe now that every integer n can be written in the form $n = mr + j$ with $m = r = \lfloor \sqrt{n} \rfloor$ (the integer part of $\sqrt{n}$) and some integer $j \in \{0, 1, \dots, 2\lfloor \sqrt{n} \rfloor\}$. If $Z_1, Z_2, \dots$ is a sequence of i.i.d. r.v.'s, independent also of X_i 's, following the standardized gamma distribution $\mathcal{SG}_\lambda$ then from the triangle inequality we deduce that,

$$\mathbf{d}_n = \mathbf{d}\left(\sum_{i=1}^{mr+j} X_i, \sum_{i=1}^{mr+j} Z_i\right) \leq \mathbf{d}\left(\sum_{i=1}^{mr+j} X_i, \sum_{i=1}^{mr} Z_i + \sum_{i=mr+1}^{mr+j} X_i\right) + \mathbf{d}\left(\sum_{i=1}^{mr} Z_i + \sum_{i=mr+1}^{mr+j} X_i, \sum_{i=1}^{mr+j} Z_i\right). \quad (18)$$

Invoking the regularity property of the metric $\mathbf{d}$ we realize that the first term in the right part of (18) is less than or equal to $\mathbf{d}_{mr}$. Applying also Lemma 1(b) for the second term in the right part, and then employing Lemma 2(a) (for $mr\lambda^2 \geq 9$) and the subadditivity property of the metric ζ_s , we get

$$\begin{aligned}\mathbf{d}_n &\leq \mathbf{d}_{mr} + \frac{1}{2} \left\| f_{\mathcal{G}(mr\lambda^2, \lambda)}^{(4)} \right\| \zeta_4\left(\sum_{i=mr+1}^{mr+j} X_i, \sum_{i=mr+1}^{mr+j} Z_i\right) \leq \mathbf{d}_{mr} + \frac{3}{(mr)^2} \sum_{i=mr+1}^{mr+j} \zeta_4(X_i, Z_i) \\ &= \mathbf{d}_{mr} + \frac{3j\beta}{(mr)^2}.\end{aligned} \quad (19)$$

Notice also that $\mathbf{d}\left(\frac{1}{\sqrt{r}}\sum_{i=1}^r X_i, \mathcal{N}\right) \rightarrow_{r \rightarrow \infty} 0$, provided that $\mathcal{L}X$ has an absolutely continuous part (see Prokhorov (1952)) and thus, as $r \rightarrow \infty$,

$$\mathbf{d}_r = \mathbf{d}\left(\frac{1}{\sqrt{r}}\sum_{i=1}^r X_i, \mathcal{S}\mathcal{G}_{\sqrt{r}\lambda}\right) \leq \mathbf{d}\left(\frac{1}{\sqrt{r}}\sum_{i=1}^r X_i, \mathcal{N}\right) + \mathbf{d}\left(\mathcal{N}, \mathcal{S}\mathcal{G}_{\sqrt{r}\lambda}\right) \rightarrow 0. \quad (20)$$

Therefore, for large enough n such that $m \geq 4, \varepsilon_r < 1, \lambda\sqrt{r} \geq 3, \mathbf{d}_r < 1/8$, and $mr\lambda^2 \geq 9$, where $m = r = \lfloor \sqrt{n} \rfloor$, relations (17) and (19) yield

$$\begin{aligned} n\mathbf{d}_n &\leq \frac{6n\mathbf{d}_r\beta}{r(m-2)^2} + \frac{81nm\beta^2}{r^2(1-\varepsilon_r)(m-1)^2} + 3\frac{n\beta}{r} \sum_{j=3}^{m-1} \frac{(2\mathbf{d}_r)^{j-1}}{(m-j)^2} \\ &\quad + \left\| f_{\mathcal{G}((m-1)\lambda^2r, \lambda\sqrt{r})}^{(4)} \right\| \frac{nm\beta}{2r} + \frac{6n\mathbf{d}_r(m-1)\beta}{r(m-2)^2(1-\varepsilon_r)} + \frac{n(2\mathbf{d}_r)^m}{2} + \frac{3nj\beta}{(mr)^2}. \end{aligned} \quad (21)$$

Using (20) and the fact that $m^2 \left\| f_{\mathcal{G}((m-1)\lambda^2r, \lambda\sqrt{r})}^{(4)} \right\| = \lambda^4 r^2 m^2 \left\| f_{\mathcal{G}((m-1)\lambda^2r, 1)}^{(4)} \right\| \rightarrow_{n \rightarrow \infty} \left\| \varphi^{(4)} \right\|$ (cf. Lemma 2(b) and relation (30)), it can be verified that the right part of inequality (21), tends to $\frac{1}{2} \left\| \varphi^{(4)} \right\| \beta$ as $n \rightarrow \infty$ and hence finally

$$\limsup_{n \rightarrow \infty} n \mathbf{d}_n \leq \frac{1}{2} \left\| \varphi^{(4)} \right\| \beta.$$

■

From the above theorem it is evident that, when $\mathcal{L}X$ possesses an absolutely continuous part, the distribution of $\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i$ approaches $\mathcal{S}\mathcal{G}_{\sqrt{n}\lambda}$, $\lambda = 2/\mathbb{E}X^3$, and consequently the distribution of the sum $\sum_{i=1}^n Y_i$ of the non-standardized $Y_i = \mu + \sigma X_i, i = 1, 2, \dots, n$, can be approximated by $\mathcal{G}(n\lambda^2, \frac{\lambda}{\sigma}, n\mu - \sigma n\lambda)$ with an approximation error of order $O(n^{-1})$. In particular, $\mathbf{d}_n$ is asymptotically upper bounded by

$$\mathbf{D}_n^* = \frac{1}{2n} \left\| \varphi^{(4)} \right\| \beta = \frac{\left\| \varphi^{(4)} \right\|}{16n} \left| 1 + \frac{(\mathbb{E}X^3)^2}{2} - \frac{\mathbb{E}X^4}{3} \right|,$$

where $\left\| \varphi^{(4)} \right\| \approx 2.8006$, and the last equality holds true when $X \leq_{4-cx} \mathcal{S}\mathcal{G}_\lambda$ or $X \geq_{4-cx} \mathcal{S}\mathcal{G}_\lambda$. It is noteworthy that, at least in the cases treated in the last section, the distance $\mathbf{d}_n$ is almost equal to its asymptotic bound $\mathbf{D}_n^*$, sometimes even for relatively small values of n (see Tables 1,2,3). Note that the closed form upper bound $\mathbf{D}_n \approx \frac{3}{(1-\varepsilon)n} \beta$ that follows from Theorem 4 is usually 2 to 4 times larger than $\mathbf{D}_n^*$.

Remark 2. In view of the conditions imposed in Theorem 4, the bound (8) is valid when $\lambda \geq 3$ and the distances $\beta = \zeta_4(X, \mathcal{S}\mathcal{G}_\lambda)$ and $\mathbf{d}_1 = \mathbf{d}(X, \mathcal{S}\mathcal{G}_\lambda)$ are relatively small, that is, the distribution $\mathcal{L}X$ of the summands is relatively close to $\mathcal{S}\mathcal{G}_\lambda$. However, we can always overcome these requirements as was accomplished in the proof of Theorem 5, provided that $\mathcal{L}X$ has an absolutely continuous part. More specifically, we can apply Theorem 4 for the i.i.d. r.v.'s $X_{i,r} = \frac{1}{\sqrt{r}}\sum_{j=(i-1)r+1}^{ir} X_j, i = 1, 2, \dots$ for appropriately large r . For these r.v.'s we observe that $\mathbb{E}X_{i,r} = 0, \mathbb{E}X_{i,r}^2 = 1, \mathbb{E}X_{i,r}^3 = 2\lambda_r^{-1}, \lambda_r = \sqrt{r}\lambda$ and $\mathbf{d}_{1,r} = \mathbf{d}(X_{i,r}, \mathcal{S}\mathcal{G}_{\lambda_r}) = O(r^{-1})$ (cf. Theorem 5) and $\beta_r = \zeta_4(X_{i,r}, \mathcal{S}\mathcal{G}_{\lambda_r}) \leq \frac{1}{r}\beta$ (see (3)). Therefore, we can always find r large enough such that $\mathbf{d}_{1,r} < 1/8, \lambda_r \geq 3$ and $\varepsilon_r = \frac{8\mathbf{d}_{1,r}}{1+\delta\mathbf{d}_{1,r}} + 27\beta_r < 1$,

and employ Theorem 4 for the sequence $X_{i,r}$, $i = 1, 2, \dots$, to deduce for $\mathbf{d}_n, n = rm$, the inequality

$$\begin{aligned} \mathbf{d}_n &= \mathbf{d} \left(\frac{1}{\sqrt{rm}} \sum_{i=1}^{rm} X_i, \mathcal{SG}_{\sqrt{rm}\lambda} \right) = \mathbf{d} \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m X_{i,r}, \mathcal{SG}_{\sqrt{m}\lambda_r} \right) \\ &\leq \frac{m}{(m-1)^2} \frac{3\beta_r}{(1-\varepsilon_r)} + \frac{1}{2} (2\mathbf{d}_{1,r})^m, \quad \text{for } m \geq 2. \end{aligned} \quad (22)$$

3.2 Distribution and Quantile approximation

In this subsection we demonstrate how we can use the above shifted gamma distribution to approximate the c.d.f. and the a -quantile of $\sum_{i=1}^n U_i$, where $U_1, U_2, \dots$ is a sequence of positively skewed i.i.d. r.v.s. Let $\mu = \mathbb{E}U_i$, $\sigma^2 = \mathbb{V}U_i$, and $X_i = (U_i - \mu)/\sigma$, $i = 1, 2, \dots$. Let also G be a r.v. that follows a $\mathcal{G}(\lambda_n^2, \lambda_n)$ distribution with $\lambda_n = \lambda\sqrt{n} = 2\sqrt{n}/\mathbb{E}X^3$. It follows that $G - \lambda_n \sim \mathcal{G}(\lambda_n^2, \lambda_n, \lambda_n) = \mathcal{SG}_{\sqrt{n}\lambda}$. Therefore, invoking Theorem 4 and the properties of the total variation distance (cf. preliminaries section) we deduce that, under appropriate conditions,

$$\left| P \left(g \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i \right) \in A \right) - P(g(G - \lambda_n) \in A) \right| \leq \mathbf{D}_n \quad (23)$$

for every Borel set A and every measurable function g . On choosing $A = (-\infty, (x - n\mu)/(\sigma\sqrt{n})]$ and $g(t) = t$, we readily get from (23) that, for every $x \in R$,

$$\left| P \left(\sum_{i=1}^n U_i \leq x \right) - F_{\mathcal{G}(\lambda_n^2, \lambda_n)} \left(\frac{x - n\mu}{\sigma\sqrt{n}} + \lambda_n \right) \right| \leq \mathbf{D}_n, \quad \lambda_n = \frac{2\sqrt{n}}{\mathbb{E}X^3}, \quad (24)$$

where $F_{\mathcal{G}(\lambda_n^2, \lambda_n)}$ denotes the c.d.f. of a gamma distribution with shape parameter λ_n^2 and scale parameter λ_n . Hence, the c.d.f. of $\sum_{i=1}^n U_i$ can be approximated at x by $F_{\mathcal{G}(\lambda_n^2, \lambda_n)}((x - n\mu)/(\sigma\sqrt{n}) + \lambda_n)$. Recall that the respective Normal approximation is $\Phi((x - n\mu)/(\sigma\sqrt{n}))$. Utilizing the above inequality (24) we can also extract an approximation for the a -quantile of $\sum_{i=1}^n U_i$. Indeed, if we denote by x_a the upper $(1 - a)$ -level critical point of this distribution, i.e. $P(\sum_{i=1}^n U_i \leq x_a) = 1 - a$, then, by setting $x = x_a$ in (24) we readily deduce that,

$$n\mu + \sigma\sqrt{n} \left(F_{\mathcal{G}(\lambda_n^2, \lambda_n)}^{-1}(1 - a - \mathbf{D}_n) - \lambda_n \right) \leq x_a \leq n\mu + \sigma\sqrt{n} \left(F_{\mathcal{G}(\lambda_n^2, \lambda_n)}^{-1}(1 - a + \mathbf{D}_n) - \lambda_n \right)$$

where $F_{\mathcal{G}(\lambda_n^2, \lambda_n)}^{-1}$ is the inverse of the c.d.f. $F_{\mathcal{G}(\lambda_n^2, \lambda_n)}$. The above implies that the upper $(1 - a)$ -level critical point (referred to as *Value at Risk* when $\sum_{i=1}^n U_i$ expresses the total loss on a financial portfolio) can be approximated by,

$$x_a \approx x_a^{\mathcal{SG}} = n\mu + \sigma\sqrt{n} \left(F_{\mathcal{G}(\lambda_n^2, \lambda_n)}^{-1}(1 - a) - \lambda_n \right). \quad (25)$$

The respective normal approximation is $x_a \approx x_a^{\mathcal{N}} = n\mu + \sigma\sqrt{n}z_a$, where $z_a = \Phi^{-1}(1 - a)$.

Example (*Noncentral chi-squared distribution - numerical comparison with the results of Hall (1983)*). As already noted in the introduction, Hall (1983) has also proposed a penultimate approximation in the CLT, by employing an appropriate (standardized) chi squared distribution. Hall's approximation yields an error of the same order as our bound, $\mathbf{D}_n$ (see (24)), while his approach is based on asymptotic expansions with remainder terms that are expressed via $o(n^{-p})$. As an application, he proposed an approximation for the quantiles of the noncentral chi-squared distribution and proceeded to a numerical evaluation of his approximation for specific values of the parameters (cf.

Table 1 in that paper). In what follows we shall also treat the same problem in order to illustrate the quantile approximation presented above and also for comparison reasons.

Let $Z_1, Z_2, \dots, Z_n$ be independent standard normal r.v.s and set $U_i = (Z_i + \sqrt{\xi/n})^2, i = 1, 2, \dots, n$. It is well known that the r.v. $\sum_{i=1}^n U_i$ follows a noncentral chi-squared distribution with n degrees of freedom and noncentrality parameter $\xi > 0$. It can be easily verified that

$$\mu = \mathbb{E}U_i = 1 + \frac{\xi}{n}, \quad \sigma^2 = \mathbb{V}U_i = 2(1 + \frac{2\xi}{n}), \quad \mathbb{E}X_i^3 = \mathbb{E}(\frac{U_i - \mu}{\sigma})^3 = \frac{2^{3/2}(1 + 3\xi/n)}{(1 + 2\xi/n)^{3/2}}.$$

Thus, the upper $(1 - a)$ -level critical point, x_a , of the noncentral chi-squared distribution (n, ξ) can be approximated by (cf. (25)),

$$x_a^{\mathcal{SG}} = n + \xi + \sqrt{2(n + 2\xi)} \left(F_{\mathcal{G}(\lambda_n^2, \lambda_n)}^{-1}(1 - a) - \lambda_n \right), \quad \lambda_n = \frac{2\sqrt{n}(1 + 2\xi/n)^{3/2}}{2^{3/2}(1 + 3\xi/n)}.$$

We numerically compare the above approximating formula, $x_a^{\mathcal{SG}}$, with the one proposed by Hall (denoted by x_a^H). The corresponding values are shown in Table 1, where we have also included the exact value of x_a (up to 4 decimals) for $a = 0.1$. The first four rows of this table are the same as in Hall (1983) (recalculated). The last two rows exhibit the performance of the shifted gamma approximation, $x_a^{\mathcal{SG}}$, and the usual normal approximation, $x_a^{\mathcal{N}} = n\mu + \sigma\sqrt{n}z_a$.

Table 1 (approximating the upper critical points of the noncentral χ^2 distribution, $a = 0.1$)

n	10	15	20	25	10	15	20	25
ξ	2.935	3.599	4.161	4.658	9.432	11.189	12.677	13.992
x_a	20.4833	27.4882	34.1692	40.6465	29.5876	37.6969	45.3151	52.6203
$x_a^H - x_a$	-0.1128	-0.1246	0.0999	0.0750	-0.1493	-0.0564	0.0494	-0.0455
$x_a^{\mathcal{SG}} - x_a$	-0.0128	-0.0083	-0.0061	-0.0047	-0.0285	-0.0210	-0.0166	-0.0136
$x_a^{\mathcal{N}} - x_a$	-0.3283	-0.3501	-0.3629	-0.3715	-0.4185	-0.4274	-0.4325	-0.4359

In Table 2 we numerically compare the same quantities, but this time we choose a very small a in order to show the performance of x_a^H , $x_a^{\mathcal{SG}}$ and $x_a^{\mathcal{N}}$ further in the right tail of the noncentral χ^2 distribution.

Table 2 (approximating the upper critical points of the noncentral χ^2 distribution, $a = 0.001$)

n, ξ	5, 1	10, 2	20, 4	30, 6	40, 8	50, 10	60, 12
x_a	24.0662	34.8901	53.675	70.8733	87.254	103.113	118.602
$x_a^H - x_a$	-0.3348	0.2689	-0.1926	-0.0307	0.0814	0.1690	-0.0688
$x_a^{\mathcal{SG}} - x_a$	0.0643	0.0554	0.0443	0.0380	0.0338	0.0308	0.0285
$x_a^{\mathcal{N}} - x_a$	-6.5036	-6.5381	-6.5498	-6.5508	-6.5501	-6.5489	-6.5477

From the above tables it is evident that the shifted gamma approximation, $x_a^{\mathcal{SG}}$, exhibits a very good overall performance in the right tail, also slightly better than x_a^H . The advantage of x_a^H is that it relies on the upper critical points of the (central) chi-squared distribution which sometimes are more readily available (e.g. when a computer is not ready for use). The second table also shows that the usual normal approximation is remarkably poor at the far right tail of this distribution and in practice should be avoided.

4 Applications

1. The distribution of the sum of lognormal r.v.'s. As a first simple example we consider the case where the summands follow a lognormal distribution, $\mathcal{LN}(\mu, \sigma)$. This distribution appears

in numerous applications in various research areas (e.g. finance, actuarial theory, reliability theory, telecommunications, biology, hydrology etc.). The distribution of the sum of i.i.d. lognormally distributed r.v.'s has no closed-form analytic expression and therefore an effective approximation seems essential (e.g. see Mehta, Wu, Molisch and Zhang (2007), Dufresne (2004) and the references therein). Lognormal is a skewed distribution and therefore normal approximation may not be so efficient, especially for moderate or small values of n . More formally, we consider independent r.v.'s $L_1, L_2, \dots$ with $L_i \sim \mathcal{LN}(\mu_i, \sigma)$ and p.d.f.

$$f_{\mathcal{LN}(\mu_i, \sigma)}(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln x - \mu_i)^2}{2\sigma^2}}, x > 0.$$

We shall work with the standardized i.i.d. r.v.'s $X_i = (L_i - \mathbb{E}L_i)/\sqrt{\mathbb{V}L_i}$, $i = 1, 2, \dots$ for which we have $\mathbb{E}X_i = 0$, $\mathbb{E}X_i^2 = 1$ and

$$\mathbb{E}X_i^3 = (2 + e^{\sigma^2})\sqrt{e^{\sigma^2} - 1} = \frac{2}{\lambda}, \quad \mathbb{E}X_i^4 = -3 + e^{2\sigma^2}(3 + e^{\sigma^2}(2 + e^{\sigma^2})). \quad (26)$$

Note that the distribution of X_i has only one parameter, σ (does not depend on μ_i). It can be easily checked numerically (at least for $\sigma < 1$) that $\mathcal{L}X_i \geq_{4-cx} \mathcal{SG}_\lambda$. Indeed, according to the Karlin-Novikoff criterion (cf. Karlin and Novikoff (1963)), the corresponding p.d.f.'s cross each other at most 4 times and after their last crossing (at right) the p.d.f. of X_i is above the p.d.f. of $\mathcal{SG}_\lambda$. Thus (cf. 13),

$$\beta = \zeta_4(X_i, \mathcal{SG}_\lambda) = \frac{\mathbb{E}X_i^4 - (3 + \frac{6}{\lambda^2})}{4!} = \frac{2e^{4\sigma^2} + e^{3\sigma^2} - 3e^{2\sigma^2}}{2 \cdot 4!}. \quad (27)$$

Employing Theorem 5 we derive that

$$\limsup_{n \rightarrow \infty} n \mathbf{d}_n \leq \frac{\|\varphi^{(4)}\| \beta}{2} = 2e^{\frac{-3-\sqrt{6}}{2}} \frac{e^{\sqrt{6}} \sqrt{3-\sqrt{6}} + \sqrt{3+\sqrt{6}}}{\sqrt{\pi/3}} \frac{2e^{4\sigma^2} + e^{3\sigma^2} - 3e^{2\sigma^2}}{2 \cdot 4!} = n \mathbf{D}_n^*, \quad (28)$$

and therefore $\mathbf{D}_n^*$ is a simple in form asymptotic bound for $\mathbf{d}_n = \mathbf{d}(n^{-1/2} \sum_{i=1}^n X_i, \mathcal{SG}_{\sqrt{n}\lambda})$, $\lambda = 2/\mathbb{E}X_i^3$. Hence, when $\mu_i = \mu$, the distribution of $\sum_{i=1}^n L_i$ can be approximated by $\mathcal{G}(n\lambda^2, \lambda/\sigma_L, n\mu_L - n\lambda\sigma_L)$, where $\mu_L = \mathbb{E}L_i = e^{\mu+\sigma^2/2}$, and $\sigma_L = \sqrt{\mathbb{V}L_i} = \sqrt{e^{\sigma^2} - 1}e^{\mu+\sigma^2/2}$. If, for example, $\sigma = 0.6$ and $n = 30$, we readily compute that $\mathcal{SG}_{\sqrt{n}\lambda} \approx \mathcal{SG}_{6.26}$, and $\mathbf{D}_n^* \approx 0.0050$ while the corresponding distance from the normal distribution is $\mathbf{d}_n^{\mathcal{N}} = \mathbf{d}(n^{-1/2} \sum_{i=1}^n X_i, \mathcal{N}) \approx 0.0519$ (cf. (1)).

Furthermore, should we wish to construct an explicit upper bound for $\mathbf{d}_n$ we can use Theorem 4. For this purpose we need to compute the value of $\mathbf{d}_1 = \mathbf{d}(X_1, \mathcal{SG}_\lambda)$. For illustration purposes we shall examine two cases : (i) $\sigma = 0.2$ (where Theorem 4 is readily applicable) and (ii) $\sigma = 0.5$ (where we have to follow the procedure described in Remark 2).

(i) For $\sigma = 0.2$, we compute $\lambda \approx 3.25577$ (cf. (26)), $\beta \approx 0.00468036$ (cf. (27)), $\mathbf{d}_1 \approx 0.0065625$ (cf. (2)) and $\varepsilon = 8\mathbf{d}_1(1 + \frac{296}{45}\mathbf{d}_1)^{-1} + 27\beta \approx 0.176698 < 1$. Since $\varepsilon < 1$, $\lambda > 3$ and $\mathbf{d}_1 < 1/8$, we can apply Theorem 4, and get

$$\mathbf{d}_n = \mathbf{d}(\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i, \mathcal{SG}_{\sqrt{n}\lambda}) \leq \mathbf{D}_n = \frac{n}{(n-1)^2} \frac{3\beta}{(1-\varepsilon)} + \frac{(2\mathbf{d}_1)^n}{2}, \quad n \geq 2.$$

The performance of the above bound, $\mathbf{D}_n$, is shown in Table 3, where we have also included the asymptotic upper bound $\mathbf{D}_n^* \approx 0.00655391 \cdot n^{-1}$, as well as the exact values of the distances $\mathbf{d}_n$ and $\mathbf{d}_n^{\mathcal{N}}$, which were evaluated numerically (via numerical integration of the n -fold convolution integral). In order to show the performance of the shifted gamma approximation at the far right tail of the

distribution, we present in the same table the Monte Carlo estimated values, $\hat{x}_a$, of the upper 99.9%-level critical point x_a of the distribution of $\sum_{i=1}^n L_i$ for $\mu_i = 0, \sigma = 0.2$ (based on 10^6 samples), along with its shifted gamma (x_a^{SG} , cf. (25)) and its usual normal approximation (x_a^N). Note that when $\sum_{i=1}^n L_i$ expresses the total loss on a portfolio of n financial assets, x_a is also referred to as Value at Risk. The approximation through the chi-square distribution suggested by Hall (1983) is almost identical to x_a^{SG} (they differ only beyond the 2nd decimal place) and thus it is not included in the table. All computations were performed using Wolfram Mathematica Software.

Table 3 (*Approximation error estimates ($\sigma = 0.2$), and upper critical points ($a = 0.001$)*)

n	$\mathbf{d}_n^N$	$\mathbf{d}_n$	$\mathbf{D}_n$	$\mathbf{D}_n^*$	$\hat{x}_a$	$\hat{x}_a - x_a^{SG}$	$\hat{x}_a - x_a^N$
2	0.0542657	0.0032673	0.0341953	0.0032769	3.135	0.012	0.194
4	0.0385001	0.0016331	0.0075798	0.0016384	5.546	0.009	0.191
8	0.0272744	0.0008174	0.0027844	0.0008192	10.151	0.006	0.188
16	0.0193058	0.0004091	0.0012127	0.0004096	19.056	0.004	0.185
32	0.0136576	0.0002046	0.0005678	0.0002048	36.433	0.003	0.184
64	0.0096598	0.0001023	0.0002750	0.0001024	70.573	0.004	0.185
128	0.0068314	0.0000512	0.0001353	0.0000512	137.97	-0.002	0.179

(ii) For $\sigma = 0.5$, we similarly compute $\lambda \approx 1.14273$, $\beta \approx 0.0543208$, $\mathbf{d}_1 \approx 0.087255$ and $\varepsilon \approx 1.91016$. In this case we cannot apply directly Theorem 4 since $\varepsilon > 1$ and $\lambda < 3$. According to Remark 2 we can overcome this by considering the r.v.'s $X_{i,r} = r^{-1/2} \sum_{j=(i-1)r+1}^{ir} X_j$, $i = 1, 2, \dots$. By choosing $r = 8$ we have that

$$\lambda_r = \lambda r^{1/2} \approx 3.23214, \quad \mathbf{d}_{1,r} = \mathbf{d}(X_{1,r}, \mathcal{SG}_{\lambda_r}) \approx 0.008691, \quad \beta_r = \frac{\beta}{r} \approx 0.0067901$$

and thus $\varepsilon_r \approx 0.249101$. Now the conditions $\varepsilon_r < 1, \lambda_r > 3, \mathbf{d}_{1,r} < 1/8$ are satisfied for the sequence $X_{1,r}, X_{2,r}, \dots$, and Theorem 4 leads to (cf. (22))

$$\mathbf{d}_n \leq \mathbf{D}_{m,r} = \frac{m}{(m-1)^2} \frac{3\beta}{(1-\varepsilon_r)r} + \frac{1}{2}(2\mathbf{d}_{1,r})^m, \quad n = mr, \quad m \geq 2.$$

Similarly to (i) we construct the following Table 4.

Table 4 (*Approximation error estimates ($\sigma = 0.5$), and upper critical points ($a = 0.001$)*)

$n = mr$	$\mathbf{d}_n^N$	$\mathbf{d}_n$	$\mathbf{D}_{m,r}$	$\mathbf{D}_n^*$	$\hat{x}_a$	$\hat{x}_a - x_a^{SG}$	$\hat{x}_a - x_a^N$
16	0.0540407	0.0044506	0.0544068	0.0047540	27.26	0.14	1.67
32	0.0385208	0.0022781	0.0120569	0.0023770	48.48	0.14	1.66
64	0.0273708	0.0011591	0.0044290	0.0011885	89.05	0.08	1.60
128	0.0193924	0.0005846	0.0019290	0.0005942	167.73	0.06	1.57

From Tables 3,4 above we observe that the asymptotic upper bound $\mathbf{D}_n^*$ is very close to the true value of $\mathbf{d}_n$ even for small values of n . It is evident that the n -fold convolution of the lognormal distribution can be efficiently approximated by the standardized gamma distribution $\mathcal{SG}_\lambda$ with $\lambda = 2(\mathbb{E}X_i^3)^{-1} = 2(2 + e^{\sigma^2})^{-1}(e^{\sigma^2} - 1)^{-1/2}$, while the normal approximation performs rather poorly. The performance of the approximation x_a^{SG} for the upper 99.9%-level critical point x_a seems also quite satisfactory, much superior to x_a^N .

2. The distribution of the sum of Pareto r.v.'s. In this example we consider the case where the independent summands follow a Pareto distribution $Pa(\alpha)$ with p.d.f.

$$f(x) = \frac{\alpha}{x^{\alpha+1}}, \quad x > 1.$$

This distribution appears in many applications (economics, actuarial theory, hydrology, reliability theory etc.). It is also used to model the tail of heavy tailed distributions in extreme value theory. For example, peaks over high thresholds of heavy tailed r.v.'s (having regularly varying tails with index $-\alpha$) follow asymptotically (when the threshold tends to infinity) a Pareto distribution with parameter α . Thus, the sum of exceedances of i.i.d. heavy tailed r.v.'s above a high threshold can be considered as a sum of i.i.d. Pareto r.v.'s. The k -th order moment of $Pa(\alpha)$ is finite only when $k < \alpha$ and therefore we assume $\alpha > 4$ so that $Pa(\alpha)$ possesses finite first four moments. Note that, when $0 < \alpha < 2$, the distribution of the sum of Pareto r.v.'s is usually approximated through the generalized CLT by a α -stable distribution (e.g. see Zaliapin, Kagan and Schoenberg (2005) and the references therein) while for α integer the distribution of this sum can be approximated by inverting its Laplace transform (see Ramsay (2006)).

In what follows we will examine a penultimate gamma distribution approximation when $\alpha > 4$. Let $Y, Y_1, Y_2, \dots$, be independent r.v.'s following a $Pa(\alpha)$ distribution. The moments of Y are $\mathbb{E}Y^k = \frac{\alpha}{\alpha-k}, k < \alpha$. We shall work with the standardized r.v.'s $X_i = (Y_i - \mathbb{E}Y_i)/\sqrt{\mathbb{V}Y_i}$. Obviously, $\mathbb{E}X_i = 0, \mathbb{E}X_i^2 = 1$ and also

$$\mathbb{E}X_i^3 = \frac{2(\alpha+1)\sqrt{\alpha-2}}{\sqrt{\alpha}(\alpha-3)} = \frac{2}{\lambda}, \quad \mathbb{E}X_i^4 = \frac{3(\alpha-2)(2+\alpha+3\alpha^2)}{\alpha(\alpha-3)(\alpha-4)}.$$

Again, it can be easily checked numerically (at least for $\alpha < 10$) that $\mathcal{SG}_\lambda \geq_{4-cx} \mathcal{L}X_i$. Therefore,

$$\beta = \zeta_4(X_i, \mathcal{SG}_\lambda) = \frac{\mathbb{E}X_i^4 - (3 + \frac{6}{\lambda^2})}{4!} = \frac{(\alpha-1)^3}{2(\alpha-4)(\alpha-3)^2\alpha}.$$

We can now approximate $\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ by $\mathcal{SG}_{\sqrt{n}\lambda}$, or equivalently, the sum $\sum_{i=1}^n Y_i$ of Pareto $Pa(\alpha)$ r.v.'s can be approximated by the shifted gamma distribution

$$\mathcal{G}(n\lambda^2, \frac{\lambda}{\sqrt{\mathbb{V}Y}}, n\mathbb{E}Y - n\lambda\sqrt{\mathbb{V}Y}) = \mathcal{G}(\frac{n\alpha(\alpha-3)^2}{(\alpha+1)^2(\alpha-2)}, \frac{(\alpha-3)(\alpha-1)}{(\alpha+1)}, \frac{n\alpha(\alpha-1)}{(\alpha+1)(\alpha-2)})$$

and the value of their distance, $\mathbf{d}_n$, can be approximated by the asymptotic upper bound $\mathbf{D}_n^* = \frac{1}{2} \|\varphi^{(4)}\| \beta/n$ (cf. Theorem 5). This approximation is demonstrated in Table 5 for $\alpha = 7$, where we have also included the exact values of the distances $\mathbf{d}_n$ and $\mathbf{d}_n^N$, which, as in the previous application, were evaluated numerically using Wolfram Mathematica Software. In the same table, we also present the Monte Carlo estimated values $\hat{x}_a$ of the upper 99.9%-level critical points of $\sum_{i=1}^n Y_i$ (using 10^6 generated samples), along with its shifted gamma ($x_a^{\mathcal{SG}}$, cf. (25)) and normal approximation (x_a^N). As was expected, the shifted gamma approximation outperforms the usual normal approximation.

Table 5 (Approximation error estimates ($\alpha = 7$), and upper critical points ($a = 0.001$))

n	$\mathbf{d}_n^N$	$\mathbf{d}_n$	$\mathbf{D}_n^*$	$\hat{x}_a$	$\hat{x}_a - x_a^{\mathcal{SG}}$	$\hat{x}_a - x_a^N$
2	0.26703	0.19665	0.22504	4.19	0.07	0.99
4	0.18897	0.08293	0.11252	6.96	0.13	1.07
8	0.13514	0.03696	0.05625	12.16	0.15	1.10
16	0.09785	0.01876	0.02812	22.22	0.15	1.11
32	0.07076	0.00997	0.01406	41.90	0.16	1.12
64	0.05093	0.00539	0.00703	80.64	0.14	1.10

Finally an explicit upper bound for $\mathbf{d}_n$ can be derived via Theorem 4 (and e.g. $\alpha = 7$) by considering the r.v.'s $X_{1,r}, X_{2,r}, \dots$, since $\mathbf{d}_1 \approx 0.362 > 1/8$ (cf. Remark 2). By choosing $r = 32$ we compute

$$\lambda_r = \lambda r^{1/2} \approx 3.34664, \quad \mathbf{d}_{1,r} \approx 0.00997937, \quad \beta_r = \frac{\beta}{r} \approx \frac{0.321429}{32}, \quad \varepsilon_r \approx 0.346123$$

From Theorem 4 we get ($n = mr$)

$$\mathbf{d}_n \leq \frac{m}{(m-1)^2} \frac{3\frac{\beta}{r}}{(1-\varepsilon_r)} + \frac{1}{2}(2\mathbf{d}_{1,r})^m \approx 0.046085 \frac{m}{(m-1)^2}, \quad m \geq 2. \quad (29)$$

5 Appendix

Proof. (of Lemma 2) The fourth derivative of the p.d.f. of $\mathcal{G}(\gamma+1, 1)$ is equal to

$$f_{\mathcal{G}(\gamma+1,1)}^{(4)}(x) = g_\gamma(x) f_{\mathcal{G}(\gamma+1,1)}(x), \text{ where } g_\gamma(x) = \frac{(\gamma-3)(\gamma-2)(\gamma-1)\gamma-4(\gamma-2)(\gamma-1)\gamma x+6(\gamma-1)\gamma x^2-4\gamma x^3+x^4}{x^4}.$$

An analytical form of $\left\| f_{\mathcal{G}(\gamma+1,1)}^{(4)} \right\|$ can be expressed via the roots $\rho_{\gamma,i}, i = 1, 2, 3, 4$, of the polynomial in x at the nominator of g_γ . The explicit form of these roots is

$$\rho_{\gamma,i} = \gamma + r_i \sqrt{\gamma + c_\gamma} + u_i \sqrt{2\gamma - c_\gamma + \frac{r_i 2\gamma}{\sqrt{\gamma + c_\gamma}}}, \text{ where } c_\gamma = -\frac{\gamma}{b_\gamma} + \frac{\gamma^2}{b_\gamma} + \frac{b_\gamma}{2},$$

$b_\gamma = 2^{1/3}(\gamma^3 - \gamma^2 + \sqrt{(2-\gamma)(\gamma-1)^2\gamma^3})^{1/3}$ and (r_i, u_i) equals to $(-1, -1), (-1, 1), (1, -1), (1, 1)$ for $i = 1, 2, 3, 4$ respectively. By denoting $\rho_{\gamma,0} = 0, \rho_{\gamma,5} = \infty$, we can express $\left\| f_{\mathcal{G}(\gamma+1,1)}^{(4)} \right\|$ as follows,

$$\left\| f_{\mathcal{G}(\gamma+1,1)}^{(4)} \right\| = \int_0^\infty \left| f_{\mathcal{G}(\gamma+1,1)}^{(4)}(x) \right| dx = \sum_{i=1}^5 (-1)^{i+1} \int_{\rho_{\gamma,i-1}}^{\rho_{\gamma,i}} f_{\mathcal{G}(\gamma+1,1)}^{(4)}(x) dx = 2 \sum_{i=1}^4 (-1)^{i+1} f_{\mathcal{G}(\gamma+1,1)}^{(3)}(\rho_{\gamma,i}).$$

Unfortunately, the above complicated form cannot be easily upper bounded for all γ . We shall use an alternative way to find a bound for large γ . If a r.v. $Z \sim \mathcal{G}(\gamma+1, 1)$, and $\gamma \geq 8$ then,

$$\begin{aligned} \left\| f_{\mathcal{G}(\gamma+1,1)}^{(4)} \right\| &= \int_0^\infty \left| f_{\mathcal{G}(\gamma+1,1)}^{(4)}(x) \right| dx = \int_0^\infty |g_\gamma(x)| f_{\mathcal{G}(\gamma+1,1)}(x) dx = \mathbb{E} |g_\gamma(Z)| \leq (\mathbb{E} g_\gamma(Z)^2)^{1/2} \\ &= \left(\frac{24(5+\gamma)(42+\gamma(37+\gamma))}{(\gamma-1)\cdots(\gamma-7)} \right)^{1/2} < \frac{6}{(\gamma+1)^2}, \end{aligned}$$

where the last inequality is valid for $\gamma \geq 176$. This is true since the only real root of the polynomial

$$h(\gamma) = 6^2(\gamma-1)\cdots(\gamma-7) - (\gamma+1)^4 24(5+\gamma)(42+\gamma(37+\gamma))$$

is approximately equal to $\rho = 175.342$ and therefore $h(\gamma) > 0$ for $\gamma > \rho$ (because $h(\gamma) \rightarrow \infty$ as $\gamma \rightarrow \infty$).

Finally, since $\left\| f_{\mathcal{G}(\gamma+1,1)}^{(4)} \right\|$ is a smooth function of γ we easily verify numerically that, $\left\| f_{\mathcal{G}(\gamma+1,1)}^{(4)} \right\| < 6/(\gamma+1)^2$ for $8 \leq \gamma < 176$. For the second part of the lemma we observe that $c_\gamma/\gamma \rightarrow (1+\sqrt{3})/2$ as $\gamma \rightarrow \infty$, and thus $(\rho_{\gamma,i} - \gamma)/\sqrt{\gamma} \rightarrow r_i \sqrt{3} + r_i u_i \sqrt{6}, i = 1, 2, 3, 4$. Moreover, it can be shown that $f_{\mathcal{G}(\gamma,1)}^{(3)}(\gamma + x\sqrt{\gamma})\gamma^2 \rightarrow \varphi^{(3)}(x)$ as $\gamma \rightarrow \infty$, and therefore,

$$\gamma^2 \left\| f_{\mathcal{G}(\gamma,1)}^{(4)} \right\| = 2 \sum_{i=1}^4 (-1)^{i+1} f_{\mathcal{G}(\gamma,1)}^{(3)}\left(\gamma + \frac{\rho_{\gamma-1,i} - \gamma}{\sqrt{\gamma}} \sqrt{\gamma}\right) \gamma^2 \xrightarrow{\gamma \rightarrow \infty} 2 \sum_{i=1}^4 (-1)^{i+1} \varphi^{(3)}(r_i \sqrt{3} + r_i u_i \sqrt{6})$$

which can be easily verified that is equal to $\left\| \varphi^{(4)} \right\| = \int_{-\infty}^\infty |\varphi^{(4)}(y)| dy$.

To complete the proof of parts (a),(b) it finally suffices to observe that

$$\left\| f_{\mathcal{G}(\gamma,\lambda)}^{(4)} \right\| = \int_0^\infty \left| f_{\mathcal{G}(\gamma,\lambda)}^{(4)}(x) \right| dx = \int_0^\infty \left| \lambda^5 f_{\mathcal{G}(\gamma,1)}^{(4)}(\lambda x) \right| dx \underset{y=\lambda x}{=} \lambda^4 \int_0^\infty \left| f_{\mathcal{G}(\gamma,1)}^{(4)}(y) \right| dy = \lambda^4 \left\| f_{\mathcal{G}(\gamma,1)}^{(4)} \right\|. \quad (30)$$

■

References

- [1] Boutsikas, M.V. (2011) Asymptotically optimal Berry–Esseen-type bounds for distributions with an absolutely continuous part. *Journal of Statistical Planning and Inference*, **141**, 1250–1268.
- [2] Boutsikas, M.V. and Vaggelatou, E. (2002). On the distance between convex-ordered random variables. *Adv. Appl. Probab.* **34**, 349–374.
- [3] Boutsikas, M.V. and Vaggelatou, E. (2010) A new method for obtaining sharp compound Poisson approximation error estimates for sums of locally dependent random variables. *Bernoulli* **16**, 301–330
- [4] Denuit M., Lefèvre C. and Shaked M. (1998) The s -convex orders among real random variables, with applications. *Mathematical Inequalities & Applications* **1**, 585–613.
- [5] Dufresne, D. (2004) The log-normal approximation in financial and other computations. *Advances in Applied Probability* **36**, 747–773.
- [6] Hall, P. (1983) Chi squared approximations to the distribution of a sum of independent random variables. *The Annals of Probability*, **11**, 1028–1036.
- [7] Karlin, S. and Novikoff, A. (1963) Generalized convex inequalities. *Pacific Journal of Mathematics* **13**, 1251–1279.
- [8] Mehta N., Wu J., Molisch A., Zhang J. (2007) Approximating a Sum of Random Variables with a Lognormal. *IEEE Transactions on Wireless Communications* **6**, 2690–2699.
- [9] Petrov, V.V. (1995). *Limit Theorems of Probability Theory*. Clarendon Press, Oxford.
- [10] Prokhorov Yu V. (1952) On a local limit theorem for densities. *Dokl. Akad. Nauk SSSR*, **83**, 797–800.
- [11] Rachev S.T. (1991) *Probability Metrics and the Stability of Stochastic Models*. Wiley, Chichester, New York.
- [12] Rachev S.T. and Ruschendorf L. (1990) Approximation of sums by compound Poisson distributions with respect to stop-loss distances. *Advances in Applied Probability* **22**, 350–374.
- [13] Ramsay C.M. (2006) The Distribution of Sums of Certain I.I.D. Pareto Variates. *Communications in Statistics - Theory and Methods*, **35:3**, 395–405.
- [14] Reijnen R., Albers W. and Kallenberg W.C.M. (2005) Approximations for stop-loss reinsurance premiums. *Insurance: Mathematics and Economics* **36**, 237–250.
- [15] Sirazhdinov S.Kh. and Mamatov M. (1962) On convergence in the mean for densities. *Theory Probab. Appl.*, **7**, 424–429.
- [16] Zaliapin I.V., Kagan Y.Y. and Schoenberg F.P. (2005) Approximating the Distribution of Pareto Sums. *Pure appl. geophys.* **162**, 1187–1228.
- [17] Zolotarev V.M. (1983) Probability metrics. *Theory Prob. Appl.*, **28**, 278–302.