CHEBYSHEV-TYPE INEQUALITIES FOR SCALE MIXTURES

Villő Csiszár[1], Tamás F. Móri[2], and Gábor J. Székely[3]

[1,2]Loránd Eötvös University, Budapest, Hungary
[3]Bowling Green State University, Bowling Green, OH

Abstract. For important classes of symmetrically distributed random variables X the smallest constants C_α are computed on the right-hand side of Chebyshev's inequality $P(|X| \geq t) \leq C_\alpha E|X|^\alpha / t^\alpha$. For example if the distribution of X is a scale mixture of centered normal random variables, then the smallest $C_2 = 0.331\ldots$ and, as $\alpha \to \infty$, the smallest $C_\alpha \downarrow 0$ and $\alpha C_\alpha \to \sqrt{2/\pi}$.

Introduction

Chebyshev's classical inequality says that

$$P(|X - E(X)| \geq t) \leq \frac{\mathbf{Var}(X)}{t^2}$$

holds for every positive t. This general inequality is valid for all random variables X with finite variance. For the sake of historical correctness we should call this inequality the Bienaymé–Chebyshev inequality (see [1], [3] for the original works and [9] and [10] for excellent historical accounts). The price of this generality is that in many cases the inequality is far from being sharp. For example, when the distribution of X is symmetric unimodal, Chebyshev's inequality can be improved as follows

$$P(|X| \geq t) \leq \frac{4}{9} \cdot \frac{\mathbf{Var}(X)}{t^2}$$

(see [7]). It is interesting to note that Gauss submitted this inequality to the Royal Scientific Society of Göttingen in 1821, in the same year when Chebyshev was born. Gauss' paper was published in 1823. Some of the most recent nice papers on this inequality are [5], [6], [12], [13]. The aim of the present note is to improve Chebyshev's result in particular cases that still preserve the possibility of important applications. An important class of distributions is the scale mixture of (centered) normal distributions. This class contains many heavy tailed distributions (e.g. all symmetric stable distributions) frequently applied. Another class is the scale mixture of Student's t-distributions. According to [2] and [8], the distributions of logarithmic asset returns can often be fitted extremely well by Student's t-distribution.

2000 *Mathematics Subject Classification.* Primary 60E15.

Key words and phrases. Convexity, scale mixtures, Bienaymé–Chebyshev inequality.

Research supported by the Hungarian National Foundation for Scientific Research, Grant No. T-43037

Let F be the cumulative distribution function (cdf) of a fixed symmetric distribution and denote the corresponding survival function by $\overline{F} = 1 - F$. Our goal is to find a Chebyshev type inequality for the random variable X whose cdf is a scale mixture of F. We are interested in the smallest C such that

$$P(|X| \ge t) \le C\,\frac{\mathbf{Var}(X)}{t^2},$$

for every positive t.

More generally, since we do not necessarily want to require finite variance, let us rather, with given $\alpha > 0$, look for the minimal value of C_α satisfying

$$P(|X| \ge t) \le C_\alpha\,\frac{E|X|^\alpha}{t^\alpha} \tag{1}$$

for every $t > 0$. When doing so, we assume that the moment of order α of the base distribution F,

$$M_\alpha = \int_{-\infty}^{+\infty} |z|^\alpha dF(z),$$

is finite.

Even more generally, for given t and $x = E|X|^\alpha$ we seek the maximum of $y = P(|X| \ge t)$.

In the following C_α always denotes the smallest possible constant in the right hand side of (1) for which (1) holds. We call C_α the *Chebyshev constant.*

Our main result is the following Chebyshev-type inequality.

Theorem 1. *Suppose that F has a density f, which is continuous and positive over a finite or infinite interval, and 0 outside of it. Suppose further, that for $z > 0$, $z^{1+\alpha}f(z)$ is initially increasing, then decreasing. Let z_α be the smallest positive root of the equation*

$$\frac{zf(z)}{\overline{F}(z)} = \alpha. \tag{2}$$

[z_α is the point where $f(z)/\overline{F}(z)$, the force of mortality in actuarial science or the failure rate in reliability theory, equals α/z.] In terms of this z_α and

$$\varrho = \left(\frac{E|X|^\alpha}{M_\alpha}\right)^{1/\alpha}$$

we have

$$\max P(|X| \ge t\varrho) = \begin{cases} 2\,\overline{F}(z_\alpha)z_\alpha^\alpha\,t^{-\alpha}, & \text{if } t \ge z_\alpha, \\ \\ 2\,\overline{F}(t) & \text{otherwise.} \end{cases} \tag{3}$$

Consequently the Chebyshev constant in (1) is

$$C_\alpha = \frac{2\,\overline{F}(z_\alpha)z_\alpha^\alpha}{M_\alpha}. \tag{4}$$

If f is differentiable on $(0, +\infty)$, and $\dfrac{zf'(z)}{f(z)}$ is a strictly decreasing function of $z > 0$, then our condition is satisfied, and the positive root of equation (2) is unique.

Proof of Theorem 1

We apply the convexity method, which has often proved to be a good tool for moment type inequalities concerning mixtures (see [4] or [11]). Let us consider the curve of the cdf $F(t/\sigma)$, where $\sigma > 0$ is arbitrary. Since

$$x = M_\alpha \sigma^\alpha, \quad y = 2\,\overline{F}(t/\sigma),$$

the explicit form of our function is

$$y = 2\,\overline{F}(tM_\alpha^{1/\alpha} x^{-1/\alpha}), \quad x > 0.$$

What we need is the convex hull of the domain below the graph of this function. Let us analyze the course of the function. Introduce $z = tM_\alpha^{1/\alpha} x^{-1/\alpha}$. Then

$$\frac{dy}{dx} = \frac{dy}{dz}\frac{dz}{dx} = -2f(z) \cdot \left(-\frac{z}{\alpha x}\right) = \frac{2}{\alpha M_\alpha t^\alpha} \cdot z^{1+\alpha} f(z),$$

which implies, by the condition imposed on f, that the function is convex for large z (that is for small x) and thereafter it is concave (i.e. the graph of our function is S-shaped). The function cannot be concave starting from the origin, because this would mean that $z^{1+\alpha} f(z)$ is monotone increasing, so $f(z) \ge const \cdot z^{-1-\alpha}$ if $z \ge z_0$, but then M_α could not be finite.

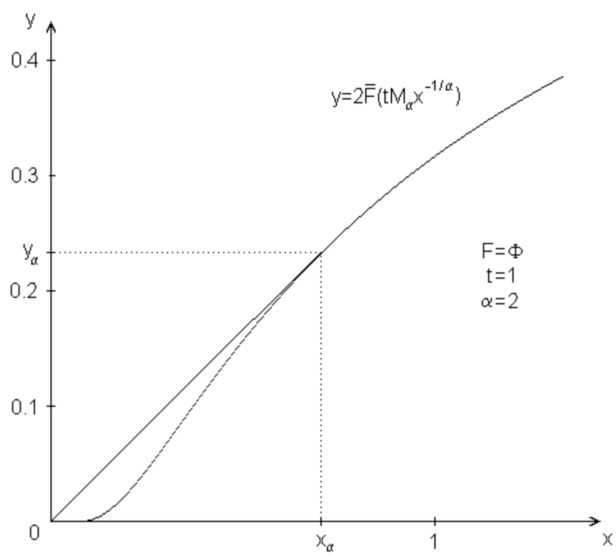

Figure 1

From the origin, draw a tangent line to the positive part of the curve (so that the tangent point on the curve is not the origin itself). Let the tangent point be (x_α, y_α). Then the upper boundary of the convex hull is either the tangent line (if $x \le x_\alpha$) or the curve itself (if $x > x_\alpha$). The tangent point x_α is determined by the equation $\dfrac{dy}{dx} = \dfrac{y}{x}$, that is

$$\frac{2\,\overline{F}(z)}{x} = \frac{2f(z)z}{\alpha x}.$$

Consequently, $x_\alpha = M_\alpha t^\alpha z_\alpha^{-\alpha}$, where z_α satisfies equation (2), and further

$$\frac{y_\alpha}{x_\alpha} = \frac{2f(z_\alpha)z_\alpha}{\alpha x_\alpha} = \frac{2f(z_\alpha)z_\alpha^{1+\alpha}}{\alpha M_\alpha t^\alpha} = \frac{2\,\overline{F}(z_\alpha)z_\alpha^\alpha}{M_\alpha t^\alpha}.$$

From these (3) and (4) will immediately follow.

When $zf'(z)/f(z)$ is strictly decreasing, we have

$$\frac{d}{dz}\log\left(z^{1+\alpha}f(z)\right) = \frac{1}{z}\left(1+\alpha+\frac{zf'(z)}{f(z)}\right).$$

Thus $\log\left(z^{1+\alpha}f(z)\right)$ is concave, therefore f satisfies the conditions of Theorem 1. Finally, since the positive roots of (2) correspond to the tangent points of an S-shaped curve, they form a closed interval. In that interval $\overline{F}(z) = c\cdot z^{-\alpha}$, thus $zf'(z)/f(z) = -(1+\alpha)$. □

Remarks.

1. This result remains valid for non-symmetric base distributions if in the formulae $2\overline{F}(z)$ is replaced by $F(-z)+1-F(z)$ and $2f(z)$ is replaced by $f(-z)+f(z)$. If we suppose that the expectation of F is 0, then the same holds for its scale mixtures, thus our inequality remains centered.

2. We obtained the value of C_α as an immediate corollary of (3). It is, however, easy to see directly that

$$C_\alpha = \frac{2}{M_\alpha}\sup\left\{t^\alpha\,\overline{F}(t) : t>0\right\}. \tag{5}$$

Indeed, any scale mixture of F can be obtained as the distribution of $X = \sigma Y$, where σ is a positive random variable, independent of Y, which is of the base distribution F. First, let σ be a constant. Then, by (5) we have

$$P(|\sigma Y| > t) = 2\overline{F}(t/\sigma) \le C_\alpha\,\frac{\sigma^\alpha M_\alpha}{t^\alpha}.$$

When σ is a random variable, the same inequality holds for the *conditional* probability:

$$P\big(|\sigma Y| > t \mid \sigma\big) \le C_\alpha\,\frac{\sigma^\alpha M_\alpha}{t^\alpha}.$$

By taking expectations on both sides we obtain

$$P(|\sigma Y| > t) \le C_\alpha\,\frac{E|\sigma Y|^\alpha}{t^\alpha}.$$

Since the right-hand side is a continuous function of t, we can also write $P(|\sigma Y| \ge t)$ on the left-hand side, as needed.

Finding the maximum in (5) leads to the equation (2), and finally we arrive at (4). The convexity method, however, produces (3), which is better than (4) because it provides a nontrivial upper bound even in the case where the right-hand side of (4) exceeds 1.

Special cases

The family of symmetric distributions satisfying the conditions of Theorem 1 is still sufficiently wide to have useful properties for modelling real life phenomena. For instance, it is closed under taking powers, in the following sense. Let Y denote a random variable with distribution F, and define $Y' = \operatorname{sign}(Y)\,|Y|^{\beta}$, for some $\beta \neq 0$. Let G and g denote the cdf and the density of Y', resp. If F satisfies the conditions of Theorem 1, then so does G. Indeed, for positive z

$$\frac{zg'(z)}{g(z)} = \frac{1}{\beta} - 1 + \frac{1}{\beta} \cdot \frac{z^{1/\beta} f'\left(z^{1/\beta}\right)}{f\left(z^{1/\beta}\right)},$$

which is increasing by supposition (no matter what the sign of β is).

In the following four examples the particular choice of the base distribution leads to widely known families of symmetric probability measures.

Example 1. Suppose the base distribution is uniform $(-1, 1)$. Then the condition imposed on the density obviously holds, and the family of scale mixtures consists of the symmetric unimodal distributions. Easy calculation yields

$$M_\alpha = \frac{1}{\alpha+1}, \quad z_\alpha = \frac{\alpha}{\alpha+1}, \quad C_\alpha = \left(\frac{\alpha}{\alpha+1}\right)^{\alpha}.$$

Gauss discovered the special case $C_2 = (2/3)^2$.

Example 2. Suppose the base distribution is standard normal with density function $f(z) = (2\pi)^{-1/2} \exp(-z^2/2)$. Then $\dfrac{zf'(z)}{f(z)} = -z^2$, which is indeed decreasing; moreover

$$M_\alpha = \frac{2^{\,\alpha/2}}{\sqrt{\pi}}\,\Gamma\left(\frac{\alpha+1}{2}\right), \quad C_\alpha = \frac{2\sqrt{\pi}\, f(z_\alpha) z_\alpha^{1+\alpha}}{2^{\,\alpha/2} \alpha\, \Gamma\left(\frac{\alpha+1}{2}\right)}. \tag{6}$$

In particular, taking $\alpha = 2$ yields $z_2 = 1.19$ and $C_2 = f(z_2) z_2^3 = 0.331\ldots$, so for simplicity one might take $C_2 \approx 1/3$.

Example 3. Suppose the original distribution is Student's t-distribution with $\nu > 0$ degrees of freedom, i.e.

$$f(z) = \frac{1}{\sqrt{\pi\nu}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{(1+z^2/\nu)^{\frac{\nu+1}{2}}}.$$

Then all moments of order $\alpha < \nu$ are finite. Our condition is satisfied, since

$$\frac{zf'(z)}{f(z)} = -(\nu+1)\,\frac{z^2}{\nu+z^2}$$

is monotone decreasing. Here

$$M_\alpha = \frac{\Gamma\left(\frac{\alpha+1}{2}\right)\Gamma\left(\frac{\nu-\alpha}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\,\nu^{\alpha/2}, \quad \text{in particular} \quad M_2 = \frac{\nu}{\nu-2}\,. \tag{7}$$

Example 4. Let the base distribution be two-sided gamma, i.e.,

$$f(z) = \frac{1}{2\,\Gamma(\nu+1)}\,|z|^{\nu}\,e^{-|z|},$$

where ν is a nonnegative parameter. We obtain the Laplace distribution for $\nu = 0$, and a bimodal density for $\nu > 0$. Here $\dfrac{z f'(z)}{f(z)} = \nu - z$ is strictly decreasing, and $M_\alpha = \dfrac{\Gamma(\alpha+\nu+1)}{\Gamma(\nu+1)}$. Particularly, in the Laplace case $z_\alpha = \alpha$, and $C_\alpha = \dfrac{\alpha^\alpha e^{-\alpha}}{\Gamma(\alpha+1)}$.

At the end of the paper there are tables displaying the Chebyshev constants in the normal case for different values of α (Table 1), in the Student case for different values of ν and fixed $\alpha = 2$ (Table 2), in the Laplace case for different values of α (Table 3), and in the two-sided gamma case for different values of ν and fixed $\alpha = 2$ (Table 4).

Properties of the Chebyshev constants

First we show that the Chebyshev constant $C_\alpha = C_\alpha(F)$ is continuous both in α and the base distribution F.

Theorem 2. *C_α is a continuous function of α, and $\lim_{\alpha\to 0} C_\alpha = 1$ for arbitrary base distribution.*

Let us fix α and vary the base distribution. Suppose that $F_n \to F$ weakly, with $M_\alpha(F_n)$ converging to $M_\alpha(F)$, as $n \to \infty$. Then $C_\alpha(F_n) \to C_\alpha(F)$.

Proof. Since $M_\alpha \to 2\overline{F}(0)$ as $\alpha \to 0$, with arbitrary $\varepsilon > 0$ we have

$$1 \ge \limsup_{\alpha\to 0} C_\alpha \ge \liminf_{\alpha\to 0} C_\alpha \ge \lim_{\alpha\to 0} \varepsilon^\alpha \overline{F}(\varepsilon)/\overline{F}(0) = \overline{F}(\varepsilon)/\overline{F}(0),$$

which can be arbitrarily close to 1 if ε is small enough. Further, since

$$\sup\left\{2\,t^\alpha \overline{F}(t) : 0 \le t < b\right\} \le b^\alpha, \quad \sup\left\{2\,t^\alpha \overline{F}(t) : B < t\right\} \le \int_B^\infty |t|^\alpha dF(t),$$

we have

$$\sup\left\{2\,t^\alpha \overline{F}(t) : t > 0\right\} = \sup\left\{2\,t^\alpha \overline{F}(t) : b \le t \le B\right\}$$

if $b > 0$ is sufficiently small and $B > b$ is large enough. Now, continuity of $\alpha \mapsto C_\alpha$ is implied by that of M_α and the following facts.

$$\lim_{h\to 0} b^{\alpha+h} = b^\alpha, \quad \lim_{h\to 0} \int_B^\infty |t|^{\alpha+h} dF(t) = \int_B^\infty |t|^\alpha dF(t),$$
$$\lim_{h\to 0} \sup\left\{2\,t^{\alpha+h}\overline{F}(t) : b \le t \le B\right\} = \sup\left\{2\,t^\alpha \overline{F}(t) : b \le t \le B\right\}.$$

Next, suppose that $F_n \to F$ weakly. Let δ_n denote the Levy distance of F_n and F. If $\delta_n < b \le t \le B$, then we have

$$2\,t^\alpha \overline{F}_n(t) \le 2\,t^\alpha\left[\overline{F}(t-\delta_n) + \delta_n\right] \le \left(\frac{b}{b-\delta_n}\right)^\alpha \sup\left\{2\,t^\alpha \overline{F}(t) : t > 0\right\} + 2B^\alpha \delta_n,$$

thus

$$\limsup_{n\to\infty}\ \sup\left\{2\,t^\alpha \overline{F}_n(t) : b \le t \le B\right\} \le \sup\left\{2\,t^\alpha \overline{F}(t) : t > 0\right\}.$$

On the other hand,

$$\limsup_{n\to\infty}\ \sup\left\{2\,t^\alpha \overline{F}_n(t) : 0 < t < b\right\} \le b^\alpha,$$

$$\limsup_{n\to\infty}\ \sup\left\{2\,t^\alpha \overline{F}_n(t) : B < t\right\} \le \int_B^\infty |t|^\alpha dF(t),$$

hence $\limsup_{n\to\infty} C_\alpha(F_n) \le C_\alpha(F)$.

For the opposite direction, let $t_n \in [b, B]$ such that

$$2\,t_n^\alpha \overline{F}(t_n) \ge \sup\left\{2\,t^\alpha \overline{F}(t) : t > 0\right\} - \delta_n.$$

Then we have

$$2\,(t-\delta_n)^\alpha \overline{F}_n(t-\delta_n) \ge 2\,(t-\delta_n)^\alpha \left[\overline{F}(t) - \delta_n\right] \ge$$
$$\ge \left(\frac{b-\delta_n}{b}\right)^\alpha \sup\left\{2\,t^\alpha \overline{F}(t) : t > 0\right\} - \delta_n - 2B^\alpha \delta_n,$$

consequently, $\liminf_{n\to\infty} C_\alpha(F_n) \ge C_\alpha(F)$. $\square$

Let us denote the constant corresponding to the Student distribution with ν degrees of freedom by $C_\alpha(\nu)$, and the one corresponding to the normal distribution by $C_\alpha(\infty)$.

Theorem 3. *For fixed $\alpha > 0$, the function $C_\alpha : (\alpha, +\infty] \to \mathbb{R}$ is continuous, monotone increasing, and*

$$\lim_{\nu\to\alpha} \frac{C_\alpha(\nu)}{\nu-\alpha} = \alpha^{(\alpha-2)/2}. \tag{8}$$

For fixed $\nu \in (0, \infty]$, the function $0 < \alpha < \nu$, $\alpha \mapsto C_\alpha(\nu)$ is monotone decreasing, and

$$\lim_{\alpha\to\nu} \frac{C_\alpha(\nu)}{\nu-\alpha} = \nu^{(\nu-2)/2}\,, \ \text{if} \ \ \nu < \infty, \tag{9}$$

$$\lim_{\alpha\to\infty} \alpha\, C_\alpha(\infty) = \sqrt{2/\pi}. \tag{10}$$

Proof. Let α be fixed. Monotonicity of $C_\alpha(.)$ is implied by the observation that the Student distribution with ν degrees of freedom is a scale mixture of the one with μ degrees of freedom if $\nu < \mu$; and further they are all scale mixtures of the standard normal distribution. Indeed, let X_μ be Student-distributed with μ degrees of freedom, and Y be independent of it with $Beta\left(\frac{\nu}{2}, \frac{\mu-\nu}{2}\right)$ distribution. Then the distribution of $X_\nu = X_\mu\sqrt{\nu/\mu Y}$ is Student with ν degrees of freedom. Similarly, if X_∞ is standard normal, and Y is independent of it and $Gamma\left(\frac{\nu}{2}, \frac{1}{2}\right)$ distributed, then the distribution of $X_\nu = X_\infty\sqrt{\nu/Y}$ is again Student with ν degrees of freedom.

Continuity of $C_\alpha(.)$ is a corollary to Theorem 2, because the parametrization of the family of Student distributions is weakly continuous, and the Student distribution is asymptotically standard normal as $\nu \to \infty$.

For the asymptotic analysis of $C_\alpha(\nu)$ let us start from the formula

$$\frac{\overline{F}_\nu(z)}{f_\nu(z)} = \Big(1+\frac{z^2}{\nu}\Big)^{\frac{\nu+1}{2}} \int_z^\infty \Big(1+\frac{y^2}{\nu}\Big)^{-\frac{\nu+1}{2}} dy. \tag{11}$$

By introducing $t = \Big(1+\frac{z^2}{\nu}\Big)^{-1}$ and $s = \Big(1+\frac{y^2}{\nu}\Big)^{-1}$, with which $z = \sqrt{\frac{\nu(1-t)}{t}}$ and $2dy = s^{-3/2}(1-s)^{-1/2}\sqrt{\nu}\,ds$, we get

$$\begin{aligned}\frac{\overline{F}_\nu(z)}{f_\nu(z)} &= \frac{\sqrt{\nu}}{2}\,t^{-(\nu+1)/2}\int_0^t s^{\nu/2-1}(1-s)^{-1/2}ds\\ &= \frac{\sqrt{\nu}}{2}\,t^{-(\nu+1)/2}\int_0^t \sum_{k=0}^\infty (-1)^k\binom{-1/2}{k} s^{k+\nu/2-1}ds\\ &= \sqrt{\nu}\sum_{k=0}^\infty \frac{u_k}{2k+\nu}\cdot t^{k-1/2},\end{aligned}$$

where $u_k = (-1)^k\binom{-1/2}{k} = 2^{-2k}\binom{2k}{k}$, particularly, $u_0 = 1$, $u_1 = 1/2$. On the other hand,

$$z = \sqrt{\nu}\sum_{k=0}^\infty (-1)^k\binom{1/2}{k}t^{k-1/2} = \sqrt{\nu}\sum_{k=0}^\infty \frac{u_k}{1-2k}\cdot t^{k-1/2},$$

hence equation (2) is read as

$$\sum_{k=0}^\infty\Big(\frac{1}{2k+\nu}+\frac{1}{\alpha(2k-1)}\Big)u_k\,t^k = 0.$$

For small positive t we obtain $\Big(\frac{1}{\nu}-\frac{1}{\alpha}\Big)+\Big(\frac{1}{2+\nu}+\frac{1}{\alpha}\Big)\frac{t}{2}+o(t) = 0$, which implies that $\nu\to\alpha \Rightarrow t\to 0$. Hence $t \sim \frac{(2+\alpha)(\nu-\alpha)}{\alpha(1+\alpha)}$ as $\nu\downarrow\alpha$, thus $z_\alpha \sim \alpha\sqrt{\frac{1+\alpha}{2+\alpha}}\cdot\frac{1}{\sqrt{\nu-\alpha}}$. The proof of (8) can be completed by plugging this and (7) into (4).

Now, let ν be fixed. Let us deal with the problem of monotonicity of Chebyshev constants in the general case of arbitrary symmetric base distribution. We start from the formula

$$\log C_\alpha = \log 2 + \alpha\log z_\alpha + \log\overline{F}(z_\alpha) - \log M_\alpha$$

(see (5)). Let us differentiate with respect to α.

$$(\log C_\alpha)' = \log z_\alpha + \frac{\alpha}{z_\alpha}\,(z_\alpha)' - \frac{f(z_\alpha)}{\overline{F}(z_\alpha)}\,(z_\alpha)' - (\log M_\alpha)' = \log z_\alpha - (\log M_\alpha)'.$$

If this is negative, C_α is decreasing.

In the case of standard normal base distribution ($\nu = \infty$) we have $\overline{F}(z) < f(z)/z$, hence $z_\alpha < \sqrt{\alpha}$. By (6) it suffices to show that $\frac{1}{2}\log\alpha < \frac{1}{2}\log 2 + \frac{1}{2}\psi\left(\frac{\alpha+1}{2}\right)$, where ψ denotes the digamma function (i.e., the derivative of $\log\Gamma$). We show that

$$\psi(t) < \log t < \psi\left(\frac{1}{2} + t\right) \qquad t > 0. \tag{12}$$

This holds for large values of t, because $\psi(t) = \log t - \dfrac{1}{2t} - \dfrac{1}{12t^2} + O\left(\dfrac{1}{t^2}\right)$, as $t \to \infty$; hence

$$\psi\left(\frac{1}{2} + t\right) = \log t + \frac{1}{24t^2} + O\left(\frac{1}{t^2}\right).$$

On the other hand, by the monotonicity and convexity of the function $x \mapsto x^{-2}$ we have

$$\psi'\left(\frac{1}{2} + t\right) = \sum_{n=0}^{\infty} \frac{1}{(n+t+1/2)^2} < \sum_{n=0}^{\infty} \int_{n+t}^{n+1+t} \frac{dx}{x^2} = \frac{1}{t} < \sum_{n=0}^{\infty} \frac{1}{(n+t)^2} = \psi'(t).$$

Consequently, $\psi(\frac{1}{2} + t) - \log t$ and $\log t - \psi(t)$ are decreasing, and therefore they are positive for $t > 0$. Thus $\log z_\alpha < (\log M_\alpha)'$, indeed.

Suppose that $\alpha \to \infty$. Since $\dfrac{z f(z)}{\overline{F}(z)} = z^2 + 1 + o(1)$ as $z \to \infty$, we have $z_\alpha = \sqrt{\alpha - 1}\,\bigl(1 + o(\alpha^{-1})\bigr)$. Therefore

$$\alpha\, C_\alpha = \frac{2\, z_\alpha^{\alpha+1} f(z_\alpha)}{2^{\alpha/2}\Gamma\left(\frac{\alpha+1}{2}\right)\pi^{-1/2}} \sim \frac{2\,(\alpha-1)^{(\alpha+1)/2}\,(2\pi)^{-1/2}\,e^{-(\alpha-1)/2}}{2^{\alpha/2}\left(\frac{\alpha-1}{2}\right)^{\alpha/2} e^{-(\alpha-1)/2}\,\alpha^{1/2}} \sim \sqrt{2/\pi},$$

as claimed in (10).

Let us turn to the case of Student base distribution ($\nu < \infty$). We clearly have

$$\left(1 + \frac{y^2}{\nu}\right)^{-(\nu+1)/2} < \left(1 + \frac{1}{y^2}\right)\left(1 + \frac{y^2}{\nu}\right)^{-(\nu+1)/2} = \frac{d}{dy}\left(-\frac{1}{y}\left(1 + \frac{y^2}{\nu}\right)^{-(\nu-1)/2}\right).$$

Integrate this to obtain

$$\int_z^{\infty} \left(1 + \frac{y^2}{\nu}\right)^{-(\nu+1)/2} dy < \frac{1}{z}\left(1 + \frac{z^2}{\nu}\right)^{-(\nu-1)/2}.$$

By (11) it follows that

$$\frac{\overline{F}_\nu(z)}{z f_\nu(z)} < \frac{1}{z^2}\left(1 + \frac{z^2}{\nu}\right) = \frac{1}{z^2} + \frac{1}{\nu},$$

which implies $z_\alpha < \sqrt{\dfrac{\nu\alpha}{\nu - \alpha}}$. By (7) it is sufficient to prove that

$$\frac{1}{2}\left(\log\nu + \log\alpha - \log(\nu - \alpha)\right) < \frac{1}{2}\left(\psi\left(\frac{\alpha+1}{2}\right) - \psi\left(\frac{\nu - \alpha}{2}\right) + \log\nu\right),$$

which is an immediate corollary of (12).

Finally, the proof of (8) yields (9), too. □

Next, let $C_\alpha(\nu)$ denote the Chebyshev constant corresponding to the two-sided gamma base distribution with parameter ν.

Theorem 4. *For fixed $\alpha > 0$, the function $C_\alpha : [0,\infty) \to \mathbb{R}$ is continuous, monotone increasing, and $\lim_{\nu\to\infty} C_\alpha(\nu) = 1$. For fixed $\nu \geq 0$, the function $0 < \alpha \mapsto C_\alpha(\nu)$ is monotone decreasing, and*

$$\lim_{\alpha\to\infty} \sqrt{\alpha}\, C_\alpha(\nu) = \frac{1}{\sqrt{2\pi}}. \tag{13}$$

Proof. Let α be fixed and apply the idea of the proof of Theorem 3. Monotonicity of $C_\alpha(.)$ is implied again by the observation that the $Gamma(\nu,1)$ distribution is a scale mixture of $Gamma(\mu,1)$ if $\nu < \mu$. To see this observe that the distribution of $X_\mu Y$ is $Gamma(\nu,1)$, provided that X_ν and Y are independent with distributions $Gamma(\mu,1)$ and $Beta(\nu,\mu-\nu)$, resp.

Continuity of $C_\alpha(.)$ follows from Theorem 2. Distributions that are scale transforms of each other obviously generate the same family of scale mixtures, thus the base cdf $F_\nu(z)$ can be replaced by $F_\nu(\nu z)$. This latter converges weakly, as $\nu \to \infty$, to the discrete distribution that puts weight $1/2$ to each of the values ± 1; and all of its absolute moments tend to 1, too. Hence $C_\alpha(\nu)$ converges to the corresponding Chebyshev constant of the limit distribution, which can be computed easily by using (5). The result is 1, showing that in the family of all symmetric distribution the classical Chebyshev inequality cannot be improved.

Now, let ν be fixed. For the monotonicity of $C_\alpha(\nu)$ we show that $\log z_\alpha < (M_\alpha)'$. Let $z > \nu$; then by substituting $y = (1+s)z$ we get

$$\frac{\overline{F}(z)}{z f(z)} = \frac{1}{z}\int_z^\infty \left(\frac{y}{z}\right)^\nu e^{-y+z}\, dy = \int_0^\infty (1+s)^\nu e^{-zs}\, ds \leq \int_0^\infty e^{-(z-\nu)s}\, ds = \frac{1}{z-\nu},$$

from which $z_\alpha \leq \nu + \alpha$ follows. Hence $\log z_\alpha \leq \log(\nu+\alpha) < \psi(\nu+\alpha+1) = (M_\alpha)'$, by (12).

On the other hand, from the line above it can be easily seen that z_α is an increasing function of ν for fixed α. Thus $z_\alpha \geq \alpha$. Hence, for fixed ν, as $\alpha \to \infty$,

$$C_\alpha = \frac{2\, z_\alpha^{1+\alpha} f(z_\alpha)}{\alpha\, M_\alpha} = \frac{{z_\alpha}^{1+\alpha+\nu} e^{-z_\alpha}}{\alpha\,\Gamma(1+\alpha+\nu)} \leq \frac{(\alpha+\nu)^{1+\alpha+\nu} e^{-(\alpha+\nu)}}{\alpha\,\Gamma(1+\alpha+\nu)} \sim \frac{1}{\sqrt{2\pi\alpha}},$$

and also

$$C_\alpha \geq \frac{\alpha^{\alpha+\nu} e^{-\alpha}}{\Gamma(1+\alpha+\nu)} \sim \frac{1}{\sqrt{2\pi\alpha}}. \quad \square$$

Open problems

1. *What conditions guarantee that the Chebyshev constants C_α decrease as α increases? Are the conditions of our Theorem 1 sufficient for this?*

In this respect notice the following.

(i) If the base distribution has finite support, then C_α does not decrease for large α. Indeed, let $0 < x_1 < \cdots < x_k$ be the sequence of positive discontinuity points of F, with jumps (probabilities) $p_1,\dots,p_k$, resp. If $q_i = p_1 + \cdots + p_k$, then

$$C_\alpha = \frac{\max\{\, q_i\, x_i^\alpha : 1 \leq i \leq k \,\}}{\sum_{i=1}^k p_i\, x_i^\alpha}.$$

When α is small, the maximum in the numerator is attained for $i = 1$, thus C_α decreases. On the other hand, for large values of α the numerator equals $p_k x_k^\alpha$, hence C_α increases eventually and converges to 1 as $\alpha \to \infty$. From Theorem 2 it is clear that not even an arbitrarily smooth density can guarantee the monotonicity of C_α.

(ii) Let Y denote a random variable with distribution F. If F belongs to the multiplicative class L, then the cdf of $Y' = \text{sign}(Y)\,|Y|^\beta$, $\beta > 1$, is a scale mixture of the cdf of Y. Since $C_\alpha(Y') = C_{\alpha\beta}(Y)$ by (5), it follows that the Chebyshev constants decrease as α increases.

2. *Find conditions for* $C_\alpha \downarrow 0$ *as* $\alpha \nearrow \alpha_{\max} = \sup\{\alpha : M_\alpha < \infty\}$. According to (i) above, distributions with finite support do not belong to this class. On the other hand smoothness itself is not sufficient either; not even when $\alpha_{\max} = \infty$. In Example 1 $C_\alpha = \left(\frac{\alpha}{\alpha+1}\right)^\alpha \downarrow 1/e$, and NOT to 0, as $\alpha \to \infty$.

Tables of Chebyshev constants

α	z_α	M_α	C_α	α	z_α	M_α	C_α
0.5	0.452	0.67598	0.53255	1.7	1.073	0.94374	0.35235
0.6	0.519	0.70192	0.50372	1.8	1.113	0.96285	0.34490
0.7	0.582	0.72701	0.47973	1.9	1.153	0.98159	0.33795
0.8	0.642	0.75133	0.45930	2.0	1.191	1.00000	0.33143
0.9	0.698	0.77493	0.44160	3.0	1.528	1.16858	0.28284
1.0	0.752	0.79788	0.42605	4.0	1.812	1.31607	0.25149
1.1	0.803	0.82023	0.41223	5.0	2.060	1.44879	0.22897
1.2	0.852	0.84201	0.39983	6.0	2.284	1.57042	0.21173
1.3	0.899	0.86328	0.38861	7.0	2.489	1.68333	0.19795
1.4	0.945	0.88405	0.37838	8.0	2.679	1.78916	0.18659
1.5	0.989	0.90424	0.36908	9.0	2.857	1.88909	0.17702
1.6	1.032	0.92426	0.36035	10.0	3.025	1.98401	0.16880

Table 1. Normal mixtures

ν	z_2	M_2	C_2	ν	z_2	M_2	C_2
3	2.101	1.73205	0.18605	15	1.283	1.07417	0.31237
4	1.707	1.41421	0.23750	16	1.277	1.06904	0.31366
5	1.554	1.29099	0.26211	17	1.271	1.06458	0.31479
6	1.472	1.22474	0.27654	18	1.267	1.06066	0.31578
7	1.420	1.18322	0.28601	19	1.262	1.05719	0.31667
8	1.384	1.15470	0.29269	20	1.258	1.05409	0.31745
9	1.358	1.13389	0.29767	30	1.235	1.03510	0.32231
10	1.338	1.11803	0.30151	40	1.223	1.02598	0.32467
11	1.323	1.10554	0.30457	50	1.216	1.02062	0.32605
12	1.310	1.09545	0.30707	75	1.208	1.01361	0.32788
13	1.300	1.08711	0.30914	100	1.203	1.01015	0.32878
14	1.291	1.08012	0.31088	∞	1.191	1.00000	0.33143

Table 2. Student mixtures, ν degrees of freedom, $\alpha = 2$

$\alpha = z_\alpha$	M_α	C_α	$\alpha = z_\alpha$	M_α	C_α
0.5	0.88623	0.48394	1.7	1.54469	0.29149
0.6	0.89351	0.45208	1.8	1.67649	0.28403
0.7	0.90864	0.42577	1.9	1.82736	0.27711
0.8	0.93138	0.40356	2.0	2	0.27067
0.9	0.96177	0.38449	3.0	6	0.22404
1.0	1.00000	0.36788	4.0	24	0.19537
1.1	1.04649	0.35324	5.0	120	0.17547
1.2	1.10180	0.34022	6.0	720	0.16062
1.3	1.16671	0.32853	7.0	5040	0.14900
1.4	1.24217	0.31797	8.0	40320	0.13959
1.5	1.32934	0.30838	9.0	362880	0.13176
1.6	1.42962	0.29957	10.0	3628800	0.12511

Table 3. Laplace mixtures

ν	z_2	M_2	C_2	ν	z_2	M_2	C_2
1	2.732	6	0.30218	13	12.233	210	0.46798
2	3.480	12	0.32755	14	13.057	240	0.47544
3	4.240	20	0.34880	15	13.884	272	0.48247
4	5.010	30	0.36708	16	14.715	306	0.48911
5	5.790	42	0.38311	17	15.548	342	0.49540
6	6.576	56	0.39737	18	16.384	380	0.50138
7	7.369	72	0.41021	19	17.222	420	0.50706
8	8.168	90	0.42188	20	18.063	462	0.51248
9	8.972	110	0.43256	30	26.575	992	0.55598
10	9.781	132	0.44240	40	35.233	1722	0.58705
11	10.595	156	0.45153	50	43.996	2652	0.61094
12	11.412	182	0.46003	100	88.752	10302	0.68217

Table 4. Two-sided gamma mixtures with parameter ν, for $\alpha = 2$

References

1. I. J. Bienaymé, *Considérations a áppui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés*, C. R. Acad. Sci., Paris **37** (1853), 309–324, reprinted immediately preceding [3] in Liouville's J. Math. Pures Appl. **12/2** (1867), 158-176.
2. N. H. Bingham and R. Kiesel, *Semi-parametric modelling in finance: theoretical foundations*, Quant. Finance vol 2 (2002), 241-250.
3. P. L. Chebyshev, *Des valeurs moyennes*, Liouville's J. Math. Pures Appl. **12/2** (1867), 177–184, translated into French by N. Hanikov from Russian published in Mat. Sbornik **2/2** (1867), 1–9.
4. V. Csiszár and T. F. Móri, *The convexity method of proving moment-type inequalities*, Statist. Probab. Lett. **66** (2004), 303–313.
5. A. DasGupta, *Best constants in Chebyshev inequalities with various applications*, Metrika **51/3** (2000), 185–200.

6. S. W. Dharmadhikari and K. Joag-Dev, *The Gauss–Tchebysheff inequality for unimodal distributions*, Theory of Probability and Its Applications **30** (1985), 867–887.

7. C. F. Gauss, *Theoria Combinationis Observationum Erroribus Minimus Obnoxiae*, Göttingen, 1821–1823.

8. C. C. Heyde, *A risky asset model with strong dependence through fractal activity time*, J. Appl. Probab. **36** (1999), 1234-1239.

9. C. C. Heyde and E. Seneta, *I. J. Bienaymé, Statistical Theory Anticipated*, Springer, New York, 1977.

10. C. C. Heyde and E. Seneta (eds.), *Statisticians of the centuries*, Springer, New York, 2001.

11. T. F. Móri and G. J. Székely, *A note on the background of several Bonferroni–Galambos type inequalities*, J. Appl. Probab. **22** (1985), 836–843.

12. T. M. Sellke, *Generalized Gauss–Chebyshev inequalities for unimodal distributions*, Metrika **43/2** (1996), 107–121.

13. T. M. Sellke and S. H. Sellke, *Chebyshev inequalities for unimodal distributions*, Amer. Statist. **51/1** (1997), 34–40.

[1,2]Department of Probability Theory and Statistics,
Loránd Eötvös University,
Pázmány P. s. 1/C,
H-1117 Budapest, Hungary

[3]Department of Mathematics and Statistics,
Bowling Green State University,
Bowling Green, OH 43403-0221 USA

E-mail address: villo@ludens.elte.hu, moritamas@ludens.elte.hu, gabors@bgnet.bgsu.edu