STATISTICS IN TRANSITION new series, September 2019
Vol. 20, No. 3, pp. 155–170, DOI 10.21307/stattrans-2019-029
Submitted – 04.03.2019; Paper ready for publication – 14.05.2019

POWER GENERALIZATION OF CHEBYSHEV'S INEQUALITY – MULTIVARIATE CASE

Katarzyna Budny [1]

ABSTRACT

In the paper some multivariate power generalizations of Chebyshev's inequality and their improvements will be presented with extension to a random vector with singular covariance matrix. Moreover, for these generalizations, the cases of the multivariate normal and the multivariate *t* distributions will be considered. Additionally, some financial application will be presented.

Key words: multivariate Chebyshev's inequality, Mahalanobis distance, multivariate normal distribution, multivariate *t* distribution.

1. Introduction

Chebyshev's inequality yields a bound on the probability of a univariate random variable taking values close to the mean expressed by its variance. Pearson (1919) proposed its univariate power generalization presenting bounds by the central moments of a random variable of even orders.

Theorem 1.1. (Pearson, 1919). If we take a random variable $\xi : \Omega \to R$ with finite central moments of *2s* order (μ_{2s}), then for all $\varepsilon > 0$

$$P\left(\left|\xi - E(\xi)\right| \geq \varepsilon\sigma\right) \leq \frac{\mu_{2s}}{\varepsilon^{2s}\sigma^{2s}} . \tag{1.1}$$

There also exist multivariate generalizations of Chebyshev's inequality (see, e.g. Olkin and Pratt, 1958, Marshall and Olkin, 1960, Osiewalski and Tatar, 1999).

In the paper we present one of those providing upper bounds on the probability that the Mahalanobis distance of a random vector from its mean is greater or equal than the fixed value. These bounds will be given by the power transformations and will constitute the multivariate extension of (1.1).

There are many applications of the Mahalanobis distance in statistical analysis. In particular, this is used in classification methods and in cluster analysis. The multivariate power generalization of Chebyshev's inequality presented below can be exploited to detect outliers.

[1] Department of Mathematics, Cracow University of Economics, Poland.
E-mail: budnyk@uek.krakow.pl. ORCID ID: https://orcid.org/0000-0002-3683-0327.

2. Multivariate power generalization of Chebyshev's inequality

We begin by recalling the inequality which is given by the measure of multivariate kurtosis.

Theorem 2.1. (Mardia, 1970) Let $\mathbf{X}:\Omega \to R^n$ be a random vector with nonsingular covariance matrix $\mathbf{\Sigma}$ and finite fourth–order moments. Then, for any $\varepsilon > 0$ the following inequality holds

$$P\left((\mathbf{X}-E(\mathbf{X}))^T \mathbf{\Sigma}^{-1}(\mathbf{X}-E(\mathbf{X})) \geq \varepsilon\right) \leq \frac{\beta_{2,n}(\mathbf{X})}{\varepsilon^2}, \tag{2.1}$$

where $\beta_{2,n}(\mathbf{X}) = E\left[\left((\mathbf{X}-E(\mathbf{X}))^T \mathbf{\Sigma}^{-1}(\mathbf{X}-E(\mathbf{X}))\right)^2\right]$ is Mardia's kurtosis of a random vector (Mardia, 1970).

Chen (2007, 2011) proposed a tight upper bound (see Navarro, 2014) in the case of a random vector for which only mean and covariance matrix are known.

Theorem 2.2. (Chen, 2007, Chen, 2011) Assume that $\mathbf{X}:\Omega \to R^n$ is a random vector with positive covariance matrix $\mathbf{\Sigma}$. Then, for all $\varepsilon > 0$ we get

$$P\left((\mathbf{X}-E(\mathbf{X}))^T \mathbf{\Sigma}^{-1}(\mathbf{X}-E(\mathbf{X})) \geq \varepsilon\right) \leq \frac{n}{\varepsilon}. \tag{2.2}$$

Budny (2014) obtained the multivariate power generalization of Chebyshev's inequality.

Theorem 2.3. (Budny, 2014) Suppose that $\mathbf{X}:\Omega \to R^n$ is a random vector with nonsingular covariance matrix $\mathbf{\Sigma}$. Let us consider any $s > 0$ such that

$$I_{s,n}(\mathbf{X}) = E\left[\left((\mathbf{X}-E(\mathbf{X}))^T \mathbf{\Sigma}^{-1}(\mathbf{X}-E(\mathbf{X}))\right)^s\right]$$

exists. Then, for all $\varepsilon > 0$

$$P\left((\mathbf{X}-E(\mathbf{X}))^T \mathbf{\Sigma}^{-1}(\mathbf{X}-E(\mathbf{X})) \geq \varepsilon\right) \leq \frac{I_{s,n}(\mathbf{X})}{\varepsilon^s}. \tag{2.3}$$

Remark 2.1. (Budny, 2014) Observe that theorems 2.1 and 2.2 can be considered as the special cases of theorem 2.3. Taking $s = 1$ we get (2.2) and for $s = 2$ we obtain (2.1).

Budny (2016), following Navarro (2016), extended (2.3) to the case of a random vector with singular covariance matrix by using the spectral decomposition.

Assume that $\mathbf{X}:\Omega \to R^n$ is a random vector with covariance matrix $\mathbf{\Sigma}$, $\operatorname{rank}\mathbf{\Sigma} = m$, $m \in \{1,...,n\}$. Let $\mathbf{\Sigma} = P\Lambda P^T$ be a spectral decomposition of a covariance matrix, i.e. P is an orthogonal matrix such that $PP^T = P^T P = I_n$ and $\Lambda = \operatorname{diag}(\lambda_1,...,\lambda_m,0,...,0)$ is the diagonal matrix with the ordered eigenvalues

$$\lambda_1 \geq ... \geq \lambda_m > \lambda_{m+1} = ... = \lambda_n = 0.$$

Hence, the Moore-Penrose generalized inverse matrix of $\mathbf{\Sigma}$ is of the form $\mathbf{\Sigma}^+ = PCP^T$, where $C = \operatorname{diag}\left(\lambda_1^{-1},...,\lambda_m^{-1},0,...,0\right)$.

Let us consider any $s > 0$ such that

$$I_{s,m}(\mathbf{X}) = E\left[\left((\mathbf{X} - E(\mathbf{X}))^T \mathbf{\Sigma}^+ (\mathbf{X} - E(\mathbf{X}))\right)^s\right]$$

exists.

Theorem 2.4. (Budny, 2016) Under the above assumptions, for any $\varepsilon > 0$, we have

$$P\left((\mathbf{X} - E(\mathbf{X}))^T \mathbf{\Sigma}^+ (\mathbf{X} - E(\mathbf{X})) \geq \varepsilon\right) \leq \frac{I_{s,m}(\mathbf{X})}{\varepsilon^s}. \tag{2.4}$$

We will denote by S the set of all $s > 0$ such that $I_{s,m}(\mathbf{X})$ exists. Let us define, for fixed $\varepsilon > 0$, the function $Bd : S \to R_+$:

$$Bd(s) = \frac{I_{s,m}(\mathbf{X})}{\varepsilon^s}, \qquad s \in S.$$

It is easily seen that for $s_1, s_2 \in S$ if $s_1 < s_2$, then the following conditions are equivalent:

$$Bd(s_1) > Bd(s_2) \quad \Leftrightarrow \quad \varepsilon > \left(\frac{I_{s_2,m}(\mathbf{X})}{I_{s_1,m}(\mathbf{X})}\right)^{\frac{1}{s_2 - s_1}}$$

and

$$Bd(s_1) < Bd(s_2) \quad \Leftrightarrow \quad 0 < \varepsilon < \left(\frac{I_{s_2,m}(\mathbf{X})}{I_{s_1,m}(\mathbf{X})}\right)^{\frac{1}{s_2 - s_1}}.$$

Summarizing, we get following remark.

Remark 2.2. For $s_1, s_2 \in S$ if $s_1 < s_2$, then the upper bound $Bd(s_2)$ of $P\left((\mathbf{X}-E(\mathbf{X}))^T \mathbf{\Sigma}^+ (\mathbf{X}-E(\mathbf{X})) \geq \varepsilon\right)$ is better than $Bd(s_1)$ for all $\varepsilon > \left(\frac{I_{s_2,m}(\mathbf{X})}{I_{s_1,m}(\mathbf{X})}\right)^{\frac{1}{s_2-s_1}}$. On the contrary, the upper bound $Bd(s_1)$ is better than $Bd(s_2)$ for all $\varepsilon \in \left(0, \left(\frac{I_{s_2,m}(\mathbf{X})}{I_{s_1,m}(\mathbf{X})}\right)^{\frac{1}{s_2-s_1}}\right)$.

In particular, if we consider $s_1 = 1$, $s_2 = 2$ and $\operatorname{rank}\mathbf{\Sigma} = n$, then the upper bound $Bd(s_2)$ is better than $Bd(s_1)$ for all $\varepsilon > \frac{\beta_{2,n}(\mathbf{X})}{n}$. Conversely, the upper bound $Bd(s_1)$ is better than $Bd(s_2)$ for all $\varepsilon \in \left(0, \frac{\beta_{2,n}(\mathbf{X})}{n}\right)$.

3. The case of the multivariate normal distribution

Budny (2016) proposed the form of the multivariate power generalization of Chebyshev's inequality for a normally distributed random vector for all $s \in N \setminus \{0\}$. In the next theorem we extend this result to the case of any real $s > 0$.

Theorem 3.1. Let $\mathbf{X}: \Omega \to R^n$ be a normally distributed random vector with mean μ and covariance matrix $\mathbf{\Sigma}$, $\mathbf{X} \sim N_n(\mu, \mathbf{\Sigma})$. Suppose that $\operatorname{rank}\mathbf{\Sigma} = m$, $m \in \{1, \ldots, n\}$. Then, for all $\varepsilon > 0$ and $s > 0$ we obtain

$$P\left((\mathrm{X}-\mu)^T \mathbf{\Sigma}^+ (\mathrm{X}-\mu) \geq \varepsilon\right) \leq \left(\frac{2}{\varepsilon}\right)^s \cdot \frac{\Gamma\left(\frac{m}{2}+s\right)}{\Gamma\left(\frac{m}{2}\right)}. \tag{3.1}$$

Proof: The proof is similar to that presented for theorem 3.1 in Budny (2016). A slight change is that we consider *s*th uncorrected moment (*s*th moment about zero) of a chi-square distribution with *m* degrees of freedom for any real $s > 0$ (not only for $s \in N \setminus \{0\}$). Hence, for $s > 0$:

$$I_{s,m}(\mathbf{X}) = \frac{2^s \Gamma\left(\frac{m}{2}+s\right)}{\Gamma\left(\frac{m}{2}\right)}$$

(Johnson, Kotz and Balakrishnan, 1994, p. 420) and it completes the proof.

Remark 3.1. For $s \in N \setminus \{0\}$ the inequality (3.1) takes the following form

$$P\left((\mathbf{X}-\mu)^T \mathbf{\Sigma}^+ (\mathbf{X}-\mu) \geq \varepsilon\right) \leq \frac{m \cdot (m+2) \cdot \ldots \cdot (m+2(s-1))}{\varepsilon^s} \quad \text{(Budny, 2016).}$$

Remark 3.2. On account of remark 2.2, if $s_1 < s_2$, then the upper bound $Bd(s_2)$ is better than $Bd(s_1)$ for all $\varepsilon > 2 \cdot \left(\Gamma\left(\frac{m}{2} + s_2\right) \Big/ \Gamma\left(\frac{m}{2} + s_1\right) \right)^{\frac{1}{s_2 - s_1}}$. Conversely, the upper bound $Bd(s_1)$ is better than $Bd(s_2)$ for all $\varepsilon \in \left(0,\, 2 \cdot \left(\Gamma\left(\frac{m}{2} + s_2\right) \Big/ \Gamma\left(\frac{m}{2} + s_1\right) \right)^{\frac{1}{s_2 - s_1}} \right)$.

Particularly if we take $s_1 = 1$ and $s_2 = 2$, then the upper bound $Bd(s_2)$ is better than $Bd(s_1)$ for all $\varepsilon > m + 2$. Conversely, the upper bound $Bd(s_1)$ is better than $Bd(s_2)$ for all $\varepsilon \in (0, m+2)$.

Example 3.1. Let us consider normally distributed random vector $\mathbf{X}: \Omega \to R^n$ with mean μ and covariance matrix $\mathbf{\Sigma}$, $\mathbf{X} \sim N_n(\mu, \mathbf{\Sigma})$. Assume that $\mathrm{rank}\mathbf{\Sigma} = m = 3$.

A random variable $(\mathbf{X}-\mu)^T \mathbf{\Sigma}^+ (\mathbf{X}-\mu)$ has a chi-square distribution with *m* degrees of freedom (Kotz, Balakrishnan and Johnson, 2000, p. 110, Budny, 2016), hence we know the exact value of P.

From remark 3.2 for $s_1 = 1$ and $s_2 = 2$ we get that the upper bound $Bd(s_2)$ is better than $Bd(s_1)$ for all $\varepsilon > m + 2 = 5$ and the upper bound $Bd(s_1)$ is better than $Bd(s_2)$ for all $\varepsilon \in (0, 5)$ (see Figure 3.1).

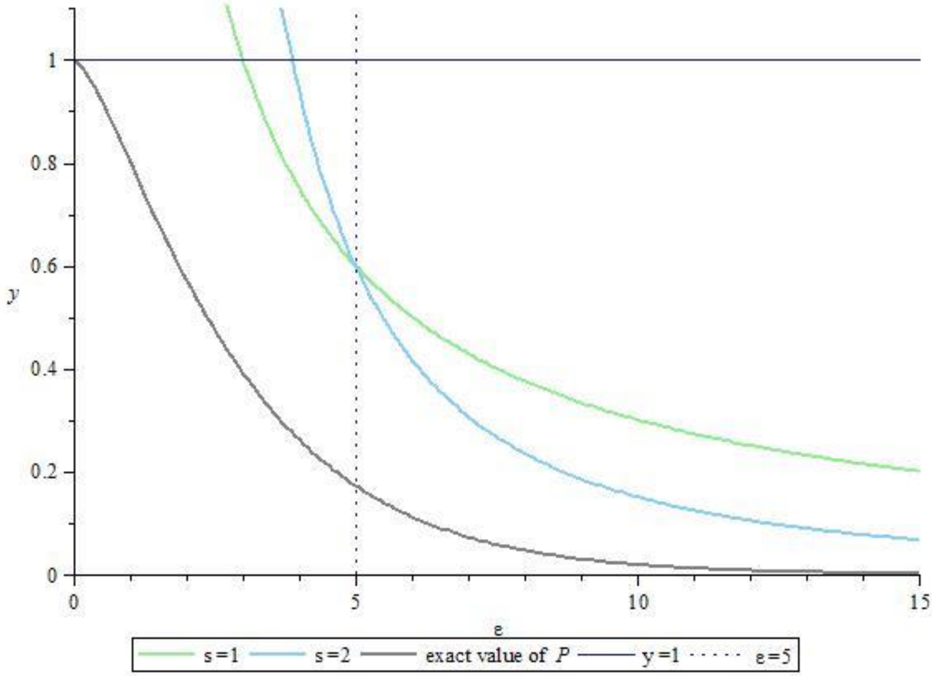

Figure 3.1. The upper bounds ($s = 1$, $s = 2$) and exact value of P for $\mathbf{X} \sim N_n(\mu, \mathbf{\Sigma})$, rank $\mathbf{\Sigma} = 3$.

In turn Figure 3.2 shows the upper bounds of $P\left((\mathbf{X} - \mu)^T \mathbf{\Sigma}^{-1}(\mathbf{X} - \mu) \geq \varepsilon\right)$ for various values of s.

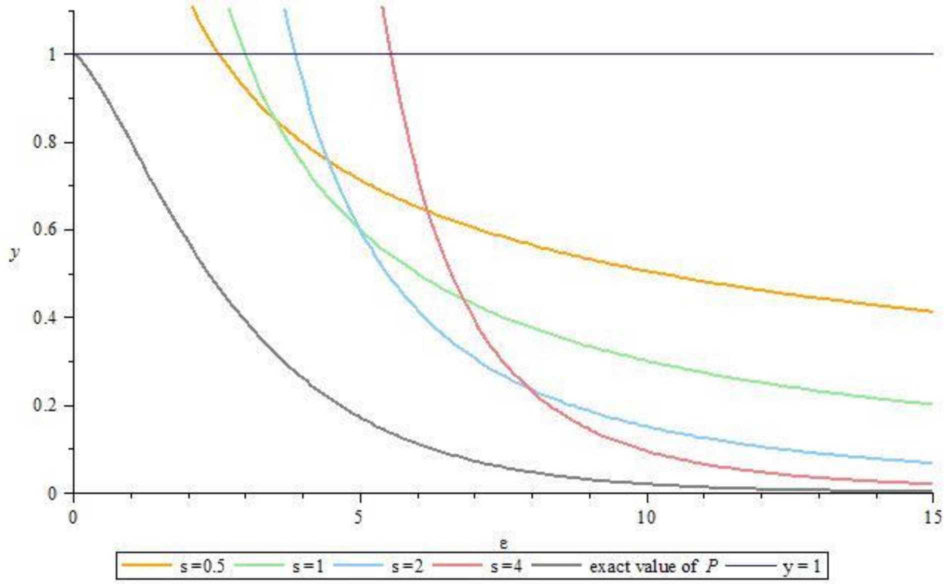

Figure 3.2. The upper bounds ($s=0.5$, $s=1$, $s=2$, $s=4$) and exact value of P for $\mathbf{X} \sim N_n(\mu, \mathbf{\Sigma})$, rank $\mathbf{\Sigma}=3$.

4. The case of the multivariate *t* distribution

A $n-$variate random vector $\mathbf{X}:\Omega \to R^n$ is said to have multivariate *t* distribution with degrees of freedom ν, mean μ and nonsingular covariance matrix $\frac{\nu}{\nu-2}\mathbf{R}$, $\nu > 2$, denoted by $t_\nu(\mu, \mathbf{R}, n)$, if its joint probability density function (pdf) is given by

$$f(x) = \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{(\pi\nu)^{n/2}\Gamma\left(\frac{\nu}{2}\right)|R|^{1/2}}\left[1+\frac{1}{\nu}(x-\mu)^T R^{-1}(x-\mu)\right]^{-(\nu+n)/2} \qquad \left(x \in R^n\right).$$

If $\mathbf{X} \sim t_\nu(\mu, \mathbf{R}, n)$, then the random variable $\xi = \frac{(\mathbf{X}-\mu)^T \mathbf{R}^{-1}(\mathbf{X}-\mu)}{n}$ has a central *F*-distribution with n, ν degrees of freedom, $\xi \sim F(n, \nu)$ (Lin, 1972).

It follows that for any $s < \frac{\nu}{2}$ we get

$$E\left[\xi^s\right] = \left(\frac{\nu}{n}\right)^s \cdot \frac{\Gamma\left(\frac{n}{2}+s\right)\Gamma\left(\frac{\nu}{2}-s\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} \tag{4.1}$$

(Johnson, Kotz and Balakrishnan, 1995, p. 349).

The power generalization of Chebyshev's inequality for multivariate *t* distribution is established by our next theorem.

Theorem 4.1. Assume that $\mathbf{X} \sim t_\nu(\mu, \mathbf{R}, n)$, rank $\mathbf{R} = n$. Then, for any $\varepsilon > 0$ the inequality (2.3) takes the following form

$$P\left((\mathbf{X}-\mu)^T \mathbf{\Sigma}^{-1}(\mathbf{X}-\mu) \geq \varepsilon\right) \leq \left(\frac{\nu-2}{\varepsilon}\right)^s \cdot \frac{\mathbf{\Gamma}\left(\frac{n}{2}+s\right)\mathbf{\Gamma}\left(\frac{\nu}{2}-s\right)}{\mathbf{\Gamma}\left(\frac{n}{2}\right)\mathbf{\Gamma}\left(\frac{\nu}{2}\right)} \tag{4.2}$$

for any $s > 0$ such that $s < \frac{\nu}{2}$.

Proof: We first observe that $\mathbf{\Sigma}^{-1} = \frac{\nu-2}{\nu}\mathbf{R}^{-1}$. From this it is obvious that

$$I_{s,n}(\mathbf{X}) = \left(\frac{\nu-2}{\nu}\right)^s \cdot n^s \cdot E\left[\xi^s\right]. \tag{4.3}$$

Substituting (4.1) into (4.3) yields

$$I_{s,n}(\mathbf{X}) = (\nu-2)^s \cdot \frac{\Gamma\left(\frac{n}{2}+s\right)\Gamma\left(\frac{\nu}{2}-s\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}. \tag{4.4}$$

This establishes the inequality (4.2).

Remark 4.1. For $s \in N \setminus \{0\}$, $s < \frac{\nu}{2}$ from (4.4) we get

$$I_{s,n}(\mathbf{X}) = \frac{(\nu-2)^s \cdot n \cdot (n+2) \cdot \ldots \cdot (n+2(s-1))}{(\nu-2)\cdot(\nu-4)\cdot\ldots\cdot(\nu-2s)}. \tag{4.5}$$

Hence, the inequality (4.2) is of the form

$$P\left((\mathrm{X}-\mu)^T \mathbf{\Sigma}^{-1}(\mathrm{X}-\mu) \geq \varepsilon\right) \leq \left(\frac{\nu-2}{\varepsilon}\right)^s \frac{n \cdot (n+2) \cdot \ldots \cdot (n+2(s-1))}{(\nu-2)\cdot(\nu-4)\cdot\ldots\cdot(\nu-2s)}.$$

Remark 4.2. According to remark 2.2, if $s_1 < s_2 < \frac{\nu}{2}$, then the upper bound $Bd(s_2)$ is better than $Bd(s_1)$ for all $\varepsilon > (\nu - 2)\left(\frac{\Gamma\left(\frac{n}{2}+s_2\right)\Gamma\left(\frac{\nu}{2}-s_2\right)}{\Gamma\left(\frac{n}{2}+s_1\right)\Gamma\left(\frac{\nu}{2}-s_1\right)} \right)^{\frac{1}{s_2-s_1}}$.

On the contrary, the upper bound $Bd(s_1)$ is better than $Bd(s_2)$ for all

$$\varepsilon \in \left(0, (\nu-2)\left(\frac{\Gamma\left(\frac{n}{2}+s_2\right)\Gamma\left(\frac{\nu}{2}-s_2\right)}{\Gamma\left(\frac{n}{2}+s_1\right)\Gamma\left(\frac{\nu}{2}-s_1\right)} \right)^{\frac{1}{s_2-s_1}} \right).$$

In particular, if we consider $s_1 = 1$, $s_2 = 2$ and $\nu > 4$, then from (4.5) the upper bound $Bd(s_2)$ is better than $Bd(s_1)$ for all $\varepsilon > \frac{(\nu-2)\cdot(n+2)}{(\nu-4)}$. Conversely, the upper bound $Bd(s_1)$ is better than $Bd(s_2)$ for all $\varepsilon \in \left(0, \frac{(\nu-2)\cdot(n+2)}{(\nu-4)}\right)$.

Example 4.1. We will consider a random vector $\mathbf{X} : \Omega \to R^3$ that has multivariate *t* distribution with degrees of freedom $\nu = 9$, $\mathbf{X} \sim t_9(\mu, \mathbf{R}, 3)$, rank $\mathbf{R} = 3$. Let us observe that

$$P\left((\mathbf{X}-\mu)^T \mathbf{\Sigma}^{-1}(\mathbf{X}-\mu) \geq \varepsilon\right) = P\left(\frac{(\mathbf{X}-\mu)^T \mathbf{R}^{-1}(\mathbf{X}-\mu)}{n} \geq \frac{\nu\varepsilon}{(\nu-2)n} \right),$$

thus we know the exact value of P. For $s_1 = 1$ and $s_2 = 2$ we get $\varepsilon_0 = \frac{(\nu-2)\cdot(n+2)}{(\nu-4)} = 7$. From this it follows that the upper bound $Bd(s_2)$ is better than $Bd(s_1)$ for all $\varepsilon > 7$ and the upper bound $Bd(s_1)$ is better than $Bd(s_2)$ for all $\varepsilon \in (0, 7)$ (see Figure 4.1).

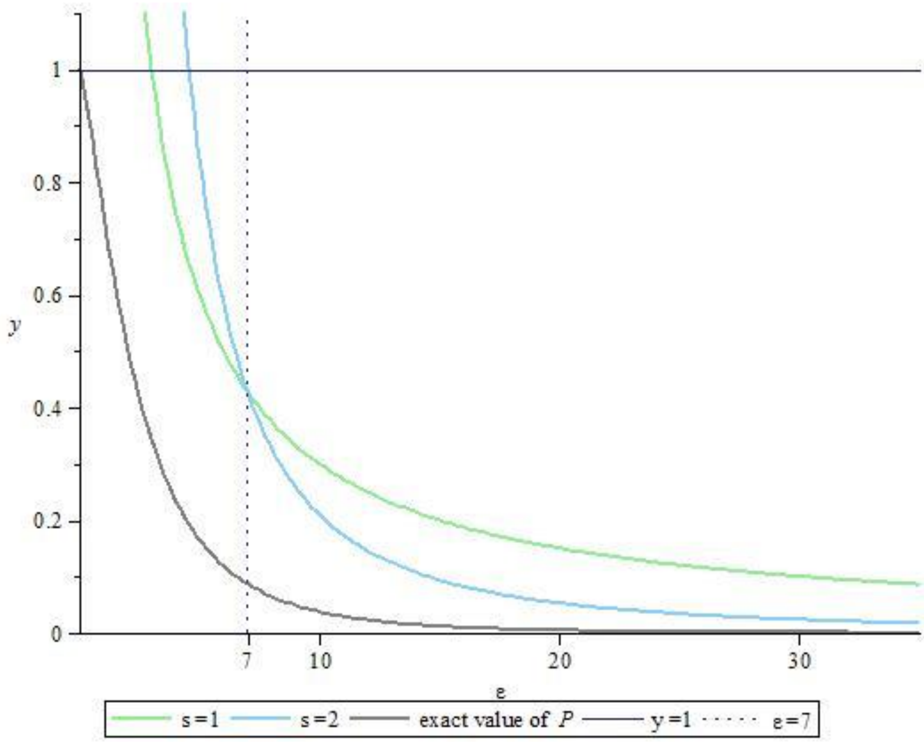

Figure 4.1. The upper bounds ($s=1$, $s=2$) and exact value of P for $\mathbf{X} \sim t_9(\mu, \mathbf{R}, 3)$.

The next figure presents the upper bounds of $P\left((\mathbf{X}-\mu)^T \mathbf{\Sigma}^{-1}(\mathbf{X}-\mu) \geq \varepsilon\right)$ for various values of $s < \frac{\nu}{2} = 4.5$.

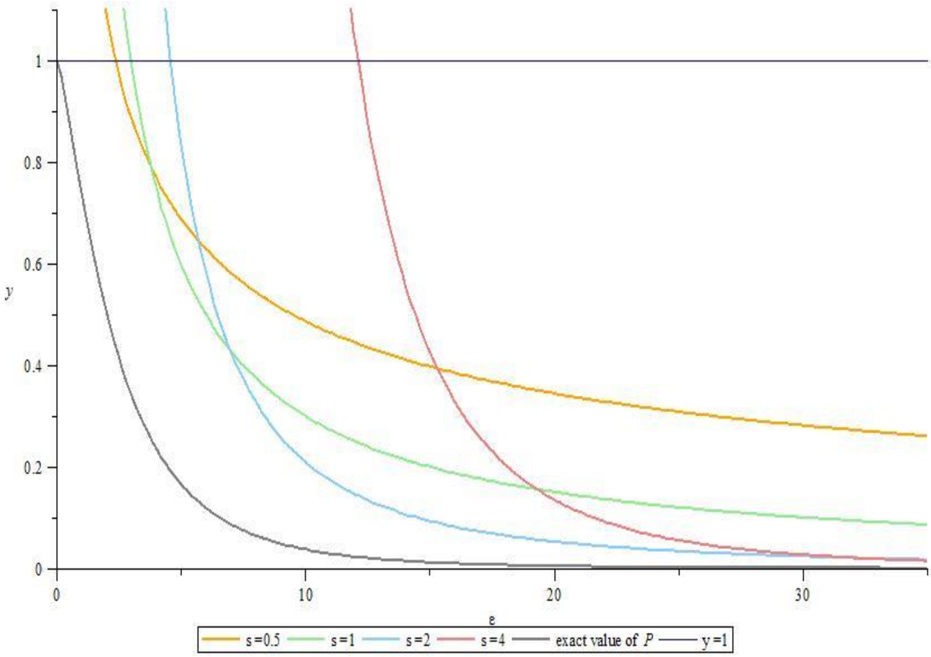

Figure 4.2. The upper bounds ($s = 0.5$, $s = 1$, $s = 2$, $s = 4$) and exact value of P for $\mathbf{X} \sim t_9(\mu, \mathbf{R}, 3)$.

5. Improvement of some multivariate power generalization of Chebyshev's inequality – extension to a random vector with a singular covariance matrix

In this section we will consider some improvement of multivariate power generalization of Chebyshev's inequality. We should mention that this improvement will be given by restricting the range of ε, it means for ε sufficiently large.

Loperfido (2014) proposed improvement of the inequality (2.4) for $s = 2$ in the case of a random vector with nonsingular covariance matrix.

Theorem 5.1. (Loperfido, 2014) Let $\beta_{2,n}(\mathbf{X})$ be Mardia's kurtosis of a random vector $\mathbf{X} : \Omega \to R^n$ with nonsingular covariance matrix $\mathbf{\Sigma}$ and finite fourth–order moments. Then for any $\varepsilon > n$ the following inequalities hold

$$P\left((\mathbf{X} - E(\mathbf{X}))^T \mathbf{\Sigma}^{-1} (\mathbf{X} - E(\mathbf{X})) \geq \varepsilon\right) \leq \frac{\beta_{2,n}(\mathbf{X}) - n^2}{\varepsilon^2 - 2n\varepsilon + \beta_{2,n}(\mathbf{X})} \leq \frac{\beta_{2,n}(\mathbf{X})}{\varepsilon^2}. \qquad (5.1)$$

In the next theorem we will extend (5.1) to the case of a random vector with singular covariance matrix.

Theorem 5.2. Assume that $\mathbf{X}:\Omega \to R^n$ is a random vector with covariance matrix $\mathbf{\Sigma}$, $\mathrm{rank}\mathbf{\Sigma} = m$, $m \in \{1,..., n\}$. Let Y denote the random variable $Y = (\mathbf{X} - E(\mathbf{X}))^T \mathbf{\Sigma}^+ (\mathbf{X} - E(\mathbf{X}))$. Let us consider Y such that $I_{2,m}(\mathbf{X}) = E(Y^2)$ is finite. Then, for all $\varepsilon > m$ we obtain

$$P\left((\mathbf{X} - E(\mathbf{X}))^T \mathbf{\Sigma}^+ (\mathbf{X} - E(\mathbf{X})) \geq \varepsilon\right) \leq \frac{I_{2,m}(\mathbf{X}) - m^2}{\varepsilon^2 - 2m\varepsilon + I_{2,m}(\mathbf{X})} \leq \frac{I_{2,m}(\mathbf{X})}{\varepsilon^2}. \qquad (5.2)$$

Proof: At the beginning we should mention that the proof will be similar to that presented in Loperfido (Loperfido 2014, proof of Theorem 1) for $s = 2$.

We first consider the random variable $Z = \left\{(Y - m)(\varepsilon - m) + I_{2,m}(\mathbf{X}) - m^2\right\}^2$. Let us observe that

$$E(Z) = (\varepsilon - m)^2 E\left((Y - m)^2\right) + \left(I_{2,m}(\mathbf{X}) - m^2\right)^2 = \left(I_{2,m}(\mathbf{X}) - m^2\right) \cdot \left(\varepsilon^2 - 2m\varepsilon + I_{2,m}(\mathbf{X})\right)$$

.

This gives that expected value of a nonnegative random variable

$$W = \frac{Z}{\left\{\varepsilon^2 - 2m\varepsilon + I_{2,m}(\mathbf{X})\right\}^2}$$

exists for all $\varepsilon > m$ and takes the form $E(W) = \dfrac{I_{2,m}(\mathbf{X}) - m^2}{\varepsilon^2 - 2m\varepsilon + I_{2,m}(\mathbf{X})}$. Markov's inequality implies that

$$P(W \geq 1) \leq \frac{I_{2,m}(\mathbf{X}) - m^2}{\varepsilon^2 - 2m\varepsilon + I_{2,m}(\mathbf{X})}.$$

The assumption $\varepsilon > m$, in turn, leads to equality $P(Y \geq \varepsilon) = P(W \geq 1)$. Indeed, the set $D = \{Y \geq \varepsilon\}$ takes the form $D = \{Y - m \geq \varepsilon - m\}$. It follows that for $\varepsilon > m$:

$$D = \left\{(Y - m)(\varepsilon - m) + \left(I_{2,m}(\mathbf{X}) - m^2\right) \geq \varepsilon^2 - 2m\varepsilon + I_{2,m}(\mathbf{X})\right\}.$$

Thus, from Jensen's inequality we have:

$$D = \left\{\ Z \geq \left\{\varepsilon^2 - 2m\varepsilon + I_{2,m}(\mathbf{X})\right\}^2\ \right\} = \{W \geq 1\}.$$

Finally, we get

$$P\left((\mathbf{X} - E(\mathbf{X}))^T \mathbf{\Sigma}^+ (\mathbf{X} - E(\mathbf{X})) \geq \varepsilon\right) \leq \frac{I_{2,m}(\mathbf{X}) - m^2}{\varepsilon^2 - 2m\varepsilon + I_{2,m}(\mathbf{X})}.$$

To prove the second inequality $\frac{I_{2,m}(\mathbf{X})}{\varepsilon^2} \geq \frac{I_{2,m}(\mathbf{X}) - m^2}{\varepsilon^2 - 2m\varepsilon + I_{2,m}(\mathbf{X})}$, let us observe that simple transformations lead to equivalent form $\frac{\left(I_{2,m}(\mathbf{X}) - m\varepsilon\right)^2}{\varepsilon^2\left\{(\varepsilon - m)^2 + I_{2,m}(\mathbf{X}) - m^2\right\}} \geq 0$, which holds for all $\varepsilon > m$, since $I_{2,m}(\mathbf{X}) \geq m^2$.

Remark 5.1. If $\operatorname{rank}\mathbf{\Sigma} = m = n$, then $\mathbf{\Sigma}^+ = \mathbf{\Sigma}^{-1}$ and $I_{2,m}(\mathbf{X}) = \beta_{2,n}(\mathbf{X})$. Hence, we obtain the inequalities (5.1).

Remark 5.2. Let $\mathbf{X}: \Omega \to R^n$ be normally distributed random vector with mean μ and covariance matrix Σ, $\mathbf{X} \sim N_n(\mu, \mathbf{\Sigma})$. Assume that $\operatorname{rank}\mathbf{\Sigma} = m$, $m \in \{1, \ldots, n\}$. Then, for any $\varepsilon > m$ we get:

$$P\left((\mathbf{X} - \mu)^T \mathbf{\Sigma}^+ (\mathbf{X} - \mu) \geq \varepsilon\right) \leq \frac{m}{\varepsilon^2 - 2m\varepsilon + m \cdot (m+2)} \leq \frac{m \cdot (m+2)}{\varepsilon^2}. \qquad (5.3)$$

We will illustrate the improvement (5.3) for $m = 3$ with Figure 5.1.

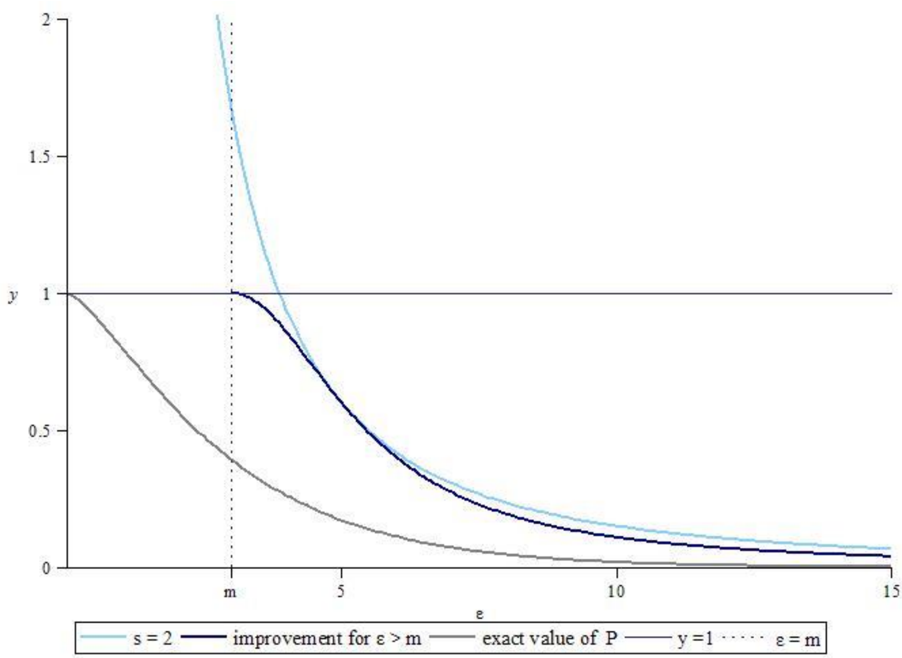

Figure 5.1. The upper bound $s = 2$, its improvement and exact value of P for $\mathbf{X} \sim N_n(\mu, \Sigma)$, $\operatorname{rank}\mathbf{\Sigma} = 3$.

Remark 5.3. Let us consider $\mathbf{X} \sim t_\nu(\mu, \mathbf{R}, n)$, rank $\mathbf{R} = n$, $\nu > 4$. Then, from (4.5) for any $\varepsilon > n$ the inequalities (5.1) take the following form

$$P\left((\mathbf{X}-\mu)^T \mathbf{\Sigma}^{-1}(\mathbf{X}-\mu) \geq \varepsilon\right) \leq \frac{\left(\frac{\nu-2}{\nu-4}\right)\cdot n \cdot (n+2) - n^2}{\varepsilon^2 - 2n\varepsilon + \left(\frac{\nu-2}{\nu-4}\right)\cdot n \cdot (n+2)} \leq \frac{\left(\frac{\nu-2}{\nu-4}\right)\cdot n \cdot (n+2)}{\varepsilon^2} \tag{5.4}$$

As the example, we will consider a 3-variate random vector $\mathbf{X} \sim t_9(\mu, \mathbf{R}, 3)$, rank $\mathbf{R} = 3$ (see example 4.1). The inequalities (5.4) are presented in Figure 5.2.

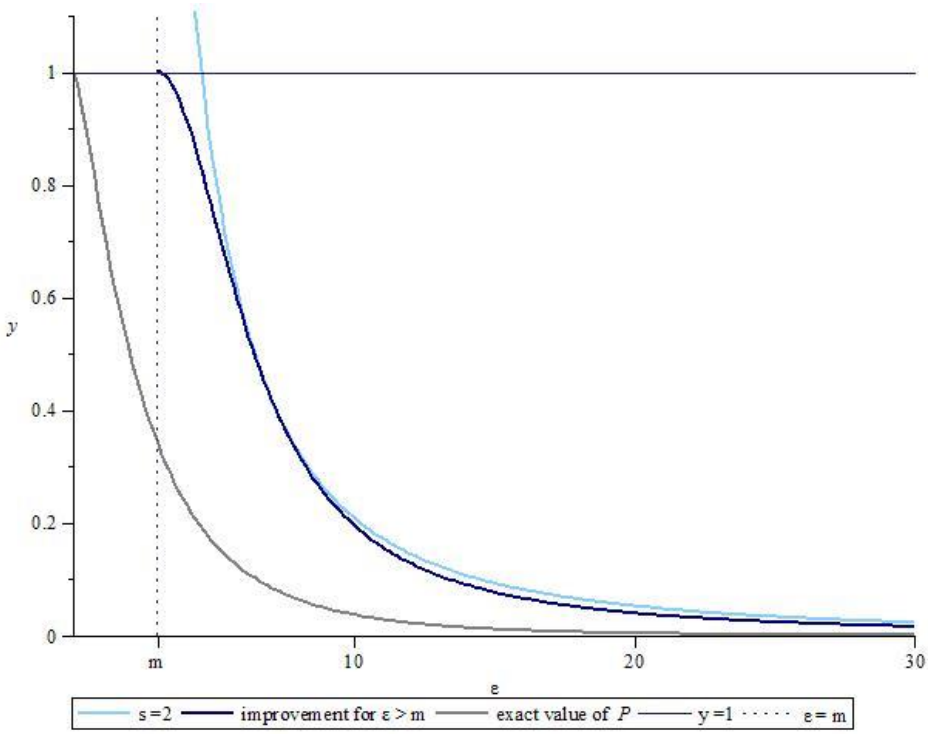

Figure 5.2. The upper bound $s = 2$, its improvement and exact value of P for $\mathbf{X} \sim t_9(\mu, \mathbf{R}, 3)$.

6. Applications in finance

Let us take a random vector r_t of *n* assets returns on a specific day t with mean μ (sample mean vector of historical returns) and covariance matrix $\mathbf{\Sigma}$ (sample covariance matrix of historical returns). Kritzman and Li (2010) propose to use the Mahalanobis distance as a measure of financial turbulence, which is understood as occurrence of unusual multivariate financial data. They defined (the so-called “the turbulence index”) turbulence for a particular time t as:

$$d_t = (r_t - \mu)^T \mathbf{\Sigma}^{-1} (r_t - \mu).$$

In the examples presented in section 3 and 4, for any $\varepsilon > 0$, we know the exact value of $P\left(d_t = (r_t - \mu)^T \mathbf{\Sigma}^{-1} (r_t - \mu) \geq \varepsilon\right)$. In the general case, this probability may not be easy to compute and if we are able to calculate the upper bounds (5.2), then we can estimate the exact value of *P*.

Other financial applications of the Mahalanobis distance were presented by Stöckl and Hanke (2014).

Acknowledgement

The publication was financed from the funds granted to the Faculty of Finance and Law at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

REFERENCES

BUDNY, K., (2014). A generalization of Chebyshev's inequality for Hilbert-space-valued random elements. Statistics and Probability Letters, 88, pp. 62–65.

BUDNY, K., (2016). An extension of the multivariate Chebyshev’s inequality to a random vector with a singular covariance matrix, Communications in Statistics – Theory and Methods, 45 (17), pp. 5220–5223.

CHEN, X., (2007). A new generalization of Chebyshev inequality for random vectors. Available at: <https://arxiv.org/abs/0707.0805>[Accessed 5 July 2007].

CHEN, X., (2011). A new generalization of Chebyshev inequality for random vectors. Available at: <https://arxiv.org/abs/0707.0805v2>[Accessed 24 June 2011].

JOHNSON, N.L., KOTZ, S., BALAKRISHNAN, N., (1994). Continuous univariate distribution. Vol. 1, 2nd ed. John Wiley & Sons Inc.

JOHNSON, N.L., KOTZ, S., BALAKRISHNAN, N., (1995). Continuous univariate distribution, Vol. 2, 2nd ed. John Wiley & Sons Inc.

KRITZMAN, M., Li, Y., (2010). Skulls, financial turbulence, and risk management. Financial Analysts Journal, 66 (5), pp. 30–41.

KOTZ, S., BALAKRISHNAN, N., JOHNSON, N.L., (2000). Continuous multivariate distribution, Vol. 1: Models and applications, 2nd ed. John Wiley & Sons Inc.

LIN, P., (1972). Some characterizations of the multivariate *t* distribution, Journal of Multivariate Analysis, 2, pp. 339–344.

LOPERFIDO, N., (2014). A probability inequality related to Mardia's kurtosis. In: C. Perna, M. Sibillo (eds.). Mathematical and statistical methods of actuarial science and finance, Springer: Springer International Publishing Switzerland 201. pp. 129–132.

MARDIA, K.V., (1970). Measures of multivariate skewness and kurtosis with applications, Biometrika, 57 (3), pp. 519–530.

MARSHALL, A., OLKIN, I., (1960). Multivariate Chebyshev inequalities, The Annals of Mathematical Statistics, 31, pp. 1001–1014.

NAVARRO, J., (2014). Can the bounds in the multivariate Chebyshev inequality be attained? Statistics and Probability Letters, 91, pp. 1–5.

NAVARRO, J., (2016). Avery simple proof of the multivariate Chebyshev's inequality. Communications in Statistics – Theory and Methods, 45 (12), pp. 3458–3463.

OLKIN, I., PRATT, J.W., (1958). A multivariate Tchebycheff inequality. The Annals of Mathematical Statistics, 29, pp. 226–234.

OSIEWALSKI, J., TATAR, J., (1999). Multivariate Chebyshev inequality based on a new definition of moments of a random vector, Przegląd Statystyczny (Stat. Rev.), 2, pp. 257–260.

PEARSON, K., (1919). On generalised Tchebycheff theorems in the mathematical theory of statistics, Biometrika,12(3–4), pp. 284–296.

STÖCKL, S., HANKE, M., (2014). Financial applications of the Mahalanobis distance, Applied Economics and Finance, 1 (2), pp. 78–84.