We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

From Asymptotic Normality to Heavy-Tailedness via Limit Theorems for Random Sums and Statistics with Random Sample Sizes

Victor Korolev and Alexander Zeifman

Abstract

This chapter contains a possible explanation of the emergence of heavy-tailed distributions observed in practice instead of the expected normal laws. The bases for this explanation are limit theorems for random sums and statistics constructed from samples with random sizes. As examples of the application of general theorems, conditions are presented for the convergence of the distributions of random sums of independent random vectors *with finite covariance matrices* to multivariate elliptically contoured stable and Linnik distributions. Also, conditions are presented for the convergence of the distributions of asymptotically normal (in the traditional sense) statistics to multivariate Student distributions. The joint asymptotic behavior of sample quantiles is also considered.

Keywords: random sum, random sample size, multivariate normal mixtures, heavy-tailed distributions, multivariate stable distribution, multivariate Linnik distribution, Mittag-Leffler distribution, multivariate Student distribution, sample quantiles
AMS 2000 Subject Classification: 60F05, 60G50, 60G55, 62E20, 62G30

1. Introduction

In many situations related to experimental data analysis, one often comes across the following phenomenon: although conventional reasoning based on the central limit theorem of probability theory concludes that the expected distribution of observations should be normal, instead, the statistical procedures expose the noticeable non-normality of real distributions. Moreover, as a rule, the observed non-normal distributions are more leptokurtic than the normal law, having sharper vertices and heavier tails. These situations are typical in the financial data analysis (see, e.g., Chapter 4 in [1] or Chapter 8 in [2] and references therein), in experimental physics (see, e.g., [3]), and other fields dealing with statistical analysis of experimental data. Many attempts were undertaken to explain this heavy-tailedness. Most significant theoretical breakthrough is usually associated with the

IntechOpen

results of B. Mandelbrot and others who proposed, instead of the standard central limit theorem, to use reasoning based on limit theorems for sums of random summands with infinite variances (see, e.g., [4]) resulting in non-normal stable laws as heavy-tailed models of the distributions of experimental data. However, first, in most cases the key assumption within this approach, the infiniteness of the variances of elementary summands can hardly be believed to hold in practice and, second, although more heavy-tailed than the normal law, the real distributions often turn out to be more light-tailed than the stable laws.

In this work, in order to give a more realistic explanation of the observed non-normality of the distributions of real data, an alternative approach based on limit theorems for statistics constructed from samples with random sizes is developed. Within this approach, it becomes possible to obtain arbitrarily heavy tails of the data distributions without assuming the non-existence of the moments of the observed characteristics.

This work was inspired by the publication of the paper [5] in which, based on the results of [6], a particular case of random sums was considered. One more reason for writing this work was the recent publication [7], the authors of which reproduced some results of [8, 9] without citing these earlier papers.

Here we give a more general description of the transformation of the limit distribution of a sum of independent random variables or another statistic (i.e., of a measurable function of a sample) under the replacement of the non-random number of summands or the sample size by a random variable. General limit theorems are proved (Section 3). Section 4 contains some comments on heavy-tailedness of scale mixtures of normal distributions. As examples of the application of general theorems, conditions are presented for the convergence of the distributions of random sums of independent random vectors *with finite covariance matrices* to multivariate elliptically contoured stable and Linnik distributions (Section 5). Also, conditions are presented for the convergence of the distributions of asymptotically normal (in the traditional sense) statistics to multivariate Student distributions (Section 6).

In Section 7, the joint asymptotic behavior of sample quantiles is considered. In applied researches related to risk analysis, such characteristic as VaR (Value-at-Risk) is very popular. Formally, VaR is a certain quantile of the observed risky value. Therefore, the joint asymptotic behavior of sample quantiles in samples with random sizes is considered in detail in Section 7 as one more example of the application of the general theorem proved in Section 3. In this section, we show how the proposed technique can be applied to the continuous-time case assuming that the sample size increases in time following a Cox process. One more interpretation of this setting is related with an important case where the sample size has the mixed Poisson distribution.

In classical problems of mathematical statistics, the size of the available sample, that is, the number of available observations, is traditionally assumed to be deterministic. In the asymptotic settings, it plays the role of infinitely increasing *known* parameter. At the same time, in practice very often the data to be analyzed are collected or registered during a certain period of time and the flow of informative events each of which brings a next observation forms a random point process. Therefore, the number of available observations is unknown till the end of the process of their registration and also must be treated as a (random) observation. For example, this is so in insurance statistics where, during different accounting periods, different numbers of insurance events (insurance claims and/or insurance contracts) occur and in high-frequency financial statistics where the number of events in a limit order book during a time unit essentially depends on the intensity of order flows. Moreover, contemporary statistical procedures of insurance and

financial mathematics do take this circumstance into consideration as one of possible ways of dealing with heavy tails. However, in other fields such as medical statistics or quality control, this approach has not become conventional; yet, the number of patients with a certain disease varies from month to month due to seasonal factors or from year to year due to some epidemic reasons and the number of failed items varies from lot to lot. In these cases, the number of available observations as well as the observations themselves is unknown beforehand and should be treated as random to avoid underestimation of risks or error probabilities.

Therefore, it is quite reasonable to study the asymptotic behavior of general statistics constructed from samples with random sizes for the purpose of construction of suitable and reasonable asymptotic approximations. As this is so, to obtain non-trivial asymptotic distributions in limit theorems of probability theory and mathematical statistics, an appropriate centering and normalization of random variables and vectors under consideration must be used. It should be especially noted that to obtain reasonable approximation to the distribution of the basic statistics, both centering and normalizing values should be non-random. Otherwise, the approximate distribution becomes random itself and, for example, the problem of evaluation of quantiles or significance levels becomes senseless.

In asymptotic settings, statistics constructed from samples with random sizes are special cases of random sequences with random indices. The randomness of indices usually leads to the limit distributions for the corresponding random sequences being heavy-tailed even in the situations where the distributions of non-randomly indexed random sequences are asymptotically normal (see, e.g., [2, 8, 10]).

Many authors noted that the asymptotic properties of statistics constructed from samples with random samples differ from those of the asymptotically normal statistics in the classical sense. To illustrate this, we will repeatedly cite [11] where the following example is given. Let $X_{(1)}, \ldots, X_{(n)}$ be order statistics constructed from the sample $X_1, \ldots, X_n$. It is well known (see, e.g., [12]) that in the standard situation the sample median is asymptotically normal. At the same time, in [11] it was demonstrated that if the sample size N_n has the geometric distribution with expectation n, then the normalized sample median $\sqrt{n}\left(X_{([N_n/2]+1)} - \operatorname{med} X_1\right)$ has the limit distribution function

$$\Psi(x) = \frac{1}{2}\left(1 + \frac{x}{\sqrt{2+x^2}}\right) \tag{1}$$

(the Student distribution with two degrees of freedom) which has such heavy tails that its moments of orders $\delta \geq 2$ do not exist. In general, as it was shown in [8], if a statistic that is asymptotically normal in the traditional sense is constructed on the basis of a sample with random size having negative binomial distribution, then instead of the expected normal law, the Student distribution with power-type decreasing heavy tails appears as an asymptotic law for this statistic.

2. Notation and definitions: auxiliary results

Let $r \in \mathbb{N}$. We will consider random elements taking values in the r-dimensional Euclidean space $\mathbb{R}^r$.

Assume that all the random variables and random vectors are defined on one and the same probability space $(\Omega, \mathcal{A}, \mathrm{P})$. By the measurability of a random field, we will mean its measurability as a function of two variates, an elementary outcome

and a parameter, with respect to the Cartesian product of the σ-algebra $\mathcal{A}$ and the Borel σ-algebra $\mathcal{B}(\mathbb{R}^r)$ of subsets of $\mathbb{R}^r$.

The distribution of a random vector ξ with respect to the measure P will be denoted $\mathcal{L}(\xi)$. The weak convergence, the coincidence of distributions, and the convergence in probability with respect to a specified probability measure will be denoted by the symbols $\Rightarrow$, $\stackrel{d}{=}$, and $\stackrel{P}{\to}$, respectively.

Let Σ be a positive definite matrix. The normal distribution in $\mathbb{R}^r$ with zero vector of expectations and covariance matrix Σ will be denoted Φ_Σ. This distribution is defined by its density

$$\phi(\mathbf{x}) = \frac{\exp\left\{-\frac{1}{2}\mathbf{x}^T\Sigma^{-1}\mathbf{x}\right\}}{(2\pi)^{r/2}|\Sigma|^{1/2}}, \qquad \mathbf{x}\in\mathbb{R}^r.$$

The characteristic function $\mathfrak{f}^Y(\mathbf{t})$ of a random variable Y such that $\mathcal{L}(Y) = \Phi_\Sigma$ has the form

$$\mathfrak{f}^Y(\mathbf{t}) \equiv \mathrm{E}\exp\left\{i\mathbf{t}^\top Y\right\} = \exp\left\{-\frac{1}{2}\mathbf{t}^\top\Sigma\mathbf{t}\right\}, \quad \mathbf{t}\in\mathbb{R}^r. \tag{2}$$

Consider a sequence $\{S_n\}_{n\geq 1}$ of random elements taking values in $\mathbb{R}^r$. Let $\Xi(\mathbb{R}^r)$ be the set of all nonsingular linear operators acting from $\mathbb{R}^r$ to $\mathbb{R}^r$. The identity operator acting from $\mathbb{R}^r$ to $\mathbb{R}^r$ will be denoted I_r. Assume that there exist sequences $\{B_n\}_{n\geq 1}$ of operators from $\Xi(\mathbb{R}^r)$ and $\{a_n\}_{n\geq 1}$ of elements from $\mathbb{R}^r$ such that

$$Y_n \equiv B_n^{-1}(S_n - a_n) \Rightarrow Y \quad (n\to\infty) \tag{3}$$

where Y is a random element whose distribution with respect to P will be denoted H, $H = \mathcal{L}(Y)$.

Along with $\{S_n\}_{n\geq 1}$, consider a sequence of integer-valued positive random variables $\{N_n\}_{n\geq 1}$ such that for each $n\geq 1$ the random variable N_n is independent of the sequence $\{S_k\}_{k\geq 1}$. Let $c_n\in\mathbb{R}^r$, $D_n\in\Xi(\mathbb{R}^r)$, and $n\geq 1$. Now, we will formulate sufficient conditions for the weak convergence of the distributions of the random elements $Z_n = D_n^{-1}(S_{N_n} - c_n)$ as $n\to\infty$.

For $g\in\mathbb{R}^r$, denote $W_n(g) = D_n^{-1}(B_{N_n}g + a_{N_n} - c_n)$. In [13, 14], the following theorem was proved, which establishes sufficient conditions of the weak convergence of multivariate random sequences with independent random indices under operator normalization.

Theorem 1 [14]. *Let $\|D_n^{-1}\|\to\infty$ as $n\to\infty$ and let the sequence of random variables $\left\{\|D_n^{-1}B_{N_n}\|\right\}_{n\geq 1}$ be tight. Assume that there exist a random element Y with distribution H and an r-dimensional random field $W(g)$, $g\in\mathbb{R}^r$, such that* (3) *holds and*

$$W_n(g)\Rightarrow W(g) \qquad (n\to\infty)$$

for H-almost all $g\in\mathbb{R}^r$. Then the random field $W(g)$ is measurable, linearly depends on g and

$$Z_n\Rightarrow W(Y) \qquad (n\to\infty),$$

where the random field $W(\cdot)$ and the random element Y are independent.

Now, consider an auxiliary statement dealing with the identifiability of a special family of mixtures of multivariate normal distributions. Let U be a nonnegative random variable. The symbol $\mathrm{E}\Phi_{U\Sigma}(\cdot)$ will denote the distribution which for each Borel set A in $\mathbb{R}^r$ is defined as

$$\mathrm{E}\Phi_{U\Sigma}(A) = \int_0^\infty \Phi_{u\Sigma}(A) d\mathrm{P}(U < u).$$

Let $\mathcal{U}$ be the set of all nonnegative random variables.

It is easy to see that if Y is a random vector such that $\mathcal{L}(Y) = \Phi_\Sigma$ independent of U, then $\mathrm{E}\Phi_{U\Sigma} = \mathcal{L}(\sqrt{U}Y)$.

Lemma 1. *Whatever nonsingular covariance matrix* Σ *is, the family of distributions* $\{\mathrm{E}\Phi_{U\Sigma}(\cdot) : U \in \mathcal{U}\}$ *is identifiable in the sense that if* $U_1 \in \mathcal{U}$, $U_2 \in \mathcal{U}$, *and*

$$\mathrm{E}\Phi_{U_1\Sigma}(A) = \mathrm{E}\Phi_{U_2\Sigma}(A) \tag{4}$$

for any set $A \in \mathcal{B}(\mathbb{R}^r)$, then $U_1 \stackrel{d}{=} U_2$.

The proof of this lemma is very simple. If $U \in \mathcal{U}$, then the characteristic function $\mathfrak{v}^{(U)}(\mathbf{t})$ corresponding to the distribution $\mathrm{E}\Phi_{U\Sigma}(\cdot)$ has the form

$$\begin{aligned}\mathfrak{v}^{(U)}(\mathbf{t}) &= \int_0^\infty \exp\left\{-\frac{1}{2}\mathbf{t}^T(u\Sigma)\mathbf{t}\right\} d\mathrm{P}(U < u) = \int_0^\infty \exp\left\{-\frac{1}{2}u\mathbf{t}^T\Sigma\mathbf{t}\right\} d\mathrm{P}(U < u) \\ &= \int_0^\infty \exp\{-us\} d\mathrm{P}(U < u), \quad s = \frac{1}{2}\mathbf{t}^T\Sigma\mathbf{t}, \quad \mathbf{t} \in \mathbb{R}^r,\end{aligned} \tag{5}$$

But on the right-hand side of (5), there is the Laplace-Stieltjes transform of the random variable U. From (4), it follows that $\mathfrak{v}^{(U_1)}(\mathbf{t}) \equiv \mathfrak{v}^{(U_2)}(\mathbf{t})$ whence by virtue of (5) the Laplace-Stieltjes transforms of the random variables U_1 and U_2 coincide, whence, in turn, it follows that $U_1 \stackrel{d}{=} U_2$. The lemma is proved.

Remark 1. When proving Lemma 1, we established a simple but useful by-product result: if $\psi(s)$ is the Laplace-Stieltjes transform of the random variable U, then the characteristic function $\mathfrak{v}^{(U)}(\mathbf{t})$ corresponding to the distribution $\mathrm{E}\Phi_{U\Sigma}$ has the form

$$\mathfrak{v}^{(U)}(\mathbf{t}) = \psi\left(\frac{1}{2}\mathbf{t}^\top\Sigma\mathbf{t}\right), \quad \mathbf{t} \in \mathbb{R}^r. \tag{6}$$

3. General theorems

First, consider the case where the random vectors $\{S_n\}_{n\geq 1}$ are formed as growing sums of independent random variables. Namely, let $X_1, X_2, \ldots$ be independent r-valued random vectors, and for $n \in \mathbb{N}$ let

$$S_n = X_1 + \ldots + X_n.$$

Consider a sequence of integer-valued positive random variables $\{N_n\}_{n\geq 1}$ such that for each $n \geq 1$ the random variable N_n is independent of the sequence $\{S_k\}_{k\geq 1}$. Let $\{b_n\}_{n\geq 1}$ be an infinitely increasing sequence of positive numbers such that

$$\mathcal{L}\left(\frac{S_n}{\sqrt{b_n}}\right) \Rightarrow \Phi_\Sigma \tag{7}$$

as $n \to \infty$, where Σ is some positive definite matrix.

Let $\{d_n\}_{n\geq 1}$ be an infinitely increasing sequence of positive numbers. As Z_n take the scalar normalized random vector

$$Z_n = \frac{S_{N_n}}{\sqrt{d_n}}.$$

Theorem 2. *Let $N_n \to \infty$ in probability as $n \to \infty$. Assume that the random variables $X_1, X_2, \ldots$ satisfy condition (6) with an asymptotic covariance matrix Σ. Then a distribution F such that*

$$L(Z_n) \Rightarrow F \qquad (n \to \infty), \tag{8}$$

exists if and only if there exists a distribution function $V(x)$ satisfying the conditions

i. $V(x) = 0$ *for* $x < 0$;

ii. *for any* $A \in \mathcal{B}(\mathbb{R}^r)$,

$$F(A) = \mathrm{E}\Phi_{U\Sigma}(A) = \int_0^\infty \Phi_{u\Sigma}(A)dV(u), \quad x \in \mathbb{R}^1;$$

iii. $\mathrm{P}(b_{N_n} < d_n x) \Rightarrow V(x)$, $n \to \infty$.

Proof. *The "if" part*. We will essentially exploit Theorem 1. For each $n \geq 1$, set $a_n = c_n = 0$, $B_n = D_n = \sqrt{d_n} I_r$. For the convenience of notation, introduce a random variable U with the distribution function $V(x)$. Note that the conditions of the theorem guarantee the tightness of the sequence of random variables

$$\|D_n^{-1} B_{N_n}\| = \sqrt{\frac{b_{N_n}}{d_n}}, \quad n = 1, 2, \ldots$$

implied by its weak convergence to the random variable $\sqrt{U}$. Further, in the case under consideration, we have $W_n(g) = \sqrt{b_{N_n}/d_n} \cdot g, g \in \mathbb{R}^r$. Therefore, the condition $N_n/d_n \Rightarrow U$ implies $W_n(g) \Rightarrow \sqrt{U}g$ for all $g \in \mathbb{R}^r$. Condition (7) means that in the case under consideration, $H = \Phi_\Sigma$. Hence, by Theorem 1, $Z_n \Rightarrow \sqrt{U}Y$ where Y is a random element with the distribution Φ_Σ independent of the random variable U. It is easy to see that the distribution of the random element $\sqrt{U}Y$ coincides with $\mathrm{E}\Phi_{U\Sigma}(\cdot)$ where the matrix Σ satisfies (7).

The "only if" part. Let condition (8) hold. Make sure that the sequence $\left\{\|D_n^{-1} B_{N_n}\|\right\}_{n \geq 1}$ is tight. Let Y be a random element with the distribution Φ_Σ. There exist $\delta > 0$ and $R > 0$ such that

$$\mathrm{P}(\|Y\| > R) > \delta. \tag{9}$$

For R specified above and an arbitrary $x > 0$, we have

$$\mathrm{P}(\|Z_n\| > x) \geq \mathrm{P}\left(\left\|\frac{S_{N_n}}{\sqrt{d_n}}\right\| > x; \left\|\frac{S_{N_n}}{\sqrt{b_{N_n}}}\right\| > R\right) =$$

$$= \mathrm{P}\left(\sqrt{\frac{b_{N_n}}{d_n}} > x \cdot \left\|\frac{S_{N_n}}{\sqrt{b_{N_n}}}\right\|^{-1}; \left\|\frac{S_{N_n}}{\sqrt{b_{N_n}}}\right\| > R\right) \geq \mathrm{P}\left(\sqrt{\frac{b_{N_n}}{d_n}} > \frac{x}{R}; \left\|\frac{S_{N_n}}{\sqrt{b_{N_n}}}\right\| > R\right) =$$

$$= \sum_{k=1}^\infty \mathrm{P}(N_n = k)\,\mathrm{P}\left(\sqrt{\frac{b_k}{d_n}} > \frac{x}{R}; \left\|\frac{S_k}{\sqrt{b_k}}\right\| > R\right) = \sum_{k=1}^\infty \mathrm{P}(N_n = k)\,\mathrm{P}\left(\sqrt{\frac{b_k}{d_n}} > \frac{x}{R}\right) P\left(\left\|\frac{S_k}{\sqrt{b_k}}\right\| > R\right) \tag{10}$$

(the last equality holds since any constant is independent of any random variable). Since by (7) the convergence $S_k/\sqrt{b_k} \Rightarrow Y$ takes place as $k \to \infty$, from (9) it follows that there exists a number $k_0 = k_0(R, \delta)$ such that

$$\mathrm{P}\left(\left\|\frac{S_k}{\sqrt{b_k}}\right\| > R\right) > \delta/2$$

for all $k > k_0$. Therefore, continuing (10) we obtain

$$\mathrm{P}(\|Z_n\| > x) \geq \frac{\delta}{2}\sum_{k=k_0+1}^{\infty}\mathrm{P}(N_n = k)\mathrm{P}\left(\sqrt{\frac{b_k}{d_n}} > \frac{x}{R}\right) =$$

$$= \frac{\delta}{2}\left[\mathrm{P}\left(\sqrt{\frac{b_{N_n}}{d_n}} > \frac{x}{R}\right) - \sum_{k=1}^{k_0}\mathrm{P}(N_n = k)\mathrm{P}\left(\sqrt{\frac{b_k}{d_n}} > \frac{x}{R}\right)\right] \geq \frac{\delta}{2}\left[\mathrm{P}\left(\sqrt{\frac{b_{N_n}}{d_n}} > \frac{x}{R}\right) - \mathrm{P}(N_n \leq k_0)\right].$$

Hence,

$$\mathrm{P}\left(\sqrt{\frac{b_{N_n}}{d_n}} > \frac{x}{R}\right) \leq \frac{2}{\delta}\mathrm{P}(\|Z_n\| > x) + \mathrm{P}(N_n \leq k_0). \tag{11}$$

From the condition $N_n \xrightarrow{P} \infty$ as $n \to \infty$, it follows that for any $\epsilon > 0$ there exists an $n_0 = n_0(\epsilon)$ such that $\mathrm{P}(N_n \leq n_0) < \epsilon$ for all $n \geq n_0$. Therefore, with the account of the tightness of the sequence $\{Z_n\}_{n\geq 1}$ that follows from its weak convergence to the random element Z with $\mathcal{L}(Z) = F$ implied by (8), relation (11) implies

$$\lim_{x\to\infty} sup_{n\geq n_0(\epsilon)}\mathrm{P}\left(\sqrt{\frac{b_{N_n}}{d_n}} > \frac{x}{R}\right) \leq \epsilon, \tag{12}$$

whatever $\epsilon > 0$ is. Now assume that the sequence

$$\|D_n^{-1}B_{N_n}\| = \sqrt{\frac{b_{N_n}}{d_n}}, \quad n = 1, 2, \ldots$$

is not tight. In that case, there exists an $\alpha > 0$ and sequences $\mathcal{N}$ of natural and $\{x_n\}_{n\in\mathcal{N}}$ of real numbers satisfying the conditions $x_n \uparrow \infty$ $(n \to \infty, n \in \mathcal{N})$ and

$$\mathrm{P}\left(\sqrt{\frac{b_{N_n}}{d_n}} > x_n\right) > \alpha, \quad n \in \mathcal{N}. \tag{13}$$

But, according to (12), for any $\epsilon > 0$ there exist $M = M(\epsilon)$ and $n_0 = n_0(\epsilon)$ such that

$$sup_{n\geq n_0(\epsilon)}\mathrm{P}\left(\sqrt{\frac{b_{N_n}}{d_n}} > M(\epsilon)\right) \leq 2\epsilon. \tag{14}$$

Choose $\epsilon < \alpha/2$ where α is the number from (13). Then for all $n \in \mathcal{N}$ large enough, in accordance with (13), the inequality opposite to (14) must hold. The obtained contradiction by the Prokhorov theorem proves the tightness of the sequence $\left\{\|D_n^{-1}B_{N_n}\|\right\}_{n\geq 1}$ or, which in this case is the same as that, of the sequence $\{b_{N_n}/d_n\}_{n\geq 1}$.

Introduce the set $W(Z)$ containing all nonnegative random variables U such that $\mathrm{P}(Z \in A) = \mathrm{E}\Phi_{U\Sigma}(A)$ for any $A \in \mathcal{B}(\mathbb{R}^r)$. Let $L(\cdot,\cdot)$ be any probability metric that metrizes weak convergence in the space of random variables, or, which is the same in this context, n the space of distribution functions, say, the Lévy metric or the smoothed Kolmogorov distance. If X_1 and X_2 are random variables with the distribution functions F_1 and F_2 respectively, then we identify $L(X_1, X_2)$ and $L(F_1, F_2)$. Show that there exists a sequence of random variables $\{U_n\}_{n\geq 1}$, $U_n \in \mathcal{W}(Z)$, such that

$$L\left(\frac{b_{N_n}}{d_n}, U_n\right) \to 0 \quad (n \to \infty). \tag{15}$$

Denote

$$\beta_n = \inf\left\{L\left(\frac{b_{N_n}}{d_n}, U\right) : \quad U \in \mathcal{W}(Z)\right\}.$$

Prove that $\beta_n \to 0$ as $n \to \infty$. Assume the contrary. In that case, $\beta_n \geq \delta$ for some $\delta > 0$ and all n from some subsequence $\mathcal{N}$ of natural numbers. Choose a subsequence $\mathcal{N}_1 \subseteq \mathcal{N}$ so that the sequence $\{b_{N_n}/d_n\}_{n\in\mathcal{N}_1}$ weakly converges to a random variable U (this is possible due to the tightness of the family $\{b_{N_n}/d_n\}_{n\geq 1}$ established above). But then $W_n(g) \Rightarrow \sqrt{U}g$ $(n \to \infty, n \in \mathcal{N}_1)$ for any $g \in \mathbb{R}^r$. Applying Theorem 1 to $n \in \mathcal{N}_1$ with condition (7) playing the role of condition (3), we make sure that $U \in \mathcal{W}(Z)$, since condition (8) provides the coincidence of the limits of all weakly convergent subsequences. So, we arrive at the contradiction to the assumption that $\beta_n \geq \delta$ for all $n \in \mathcal{N}_1$. Hence, $\beta_n \to 0$ as $n \to \infty$.

For any $n = 1, 2, \ldots$, choose a random variable U_n from $\mathcal{W}(Z)$ satisfying the condition

$$L\left(\frac{b_{N_n}}{d_n}, U_n\right) \leq \beta_n + \frac{1}{n}.$$

This sequence obviously satisfies condition (15). Now consider the structure of the set $\mathcal{W}(Z)$. This set contains all the random variable's defining the family of special mixtures of multivariate normal laws considered in Lemma 1, according to which this family is identifiable. So, whatever a random element Z is, the set $\mathcal{W}(Z)$ contains at most one element. Therefore, actually condition (15) is equivalent to

$$\frac{b_{N_n}}{d_n} \Rightarrow U \qquad (n \to \infty),$$

that is, to condition (iii) of the theorem. The theorem is proved.

Corollary 1. *Under the conditions of Theorem* 2, *non-randomly normalized random sums* S_{N_n}/d_n *are asymptotically normal with some covariance matrix* Σ' *if and only if there exists a number* $c > 0$ *such that*

$$\frac{b_{N_n}}{d_n} \Rightarrow c \quad (n \to \infty).$$

Moreover, in this case, $\Sigma' = c\Sigma$.

This statement immediately follows from Theorem 2 with the account of Lemma 1.

Now consider a formally more general setting.

Let $N_1, N_2, \ldots$ and $W_1, W_2, \ldots$ be random variables and random vectors, respectively, such that for each $n \geq 1$ the random variable N_n takes only natural values and is independent of the sequence $W_1, W_2, \ldots$. Let

$$T_n = T_n(W_1, \ldots, W_n) = (T_{n,1}(W_1, \ldots, W_n), \ldots, T_{n,r}(W_1, \ldots, W_n))$$

be a statistic taking values in $\mathbb{R}^r$, $r \geq 1$. For each $n \geq 1$ define a random vector (random element) T_{N_n} by setting

$$T_{N_n}(\omega) = T_{N_n(\omega)}\big(W_1(\omega), \ldots, W_{N_n(\omega)}(\omega)\big)$$

for every elementary outcome $\omega \in \Omega$.

We shall say that a statistic T_n is *asymptotically normal* with the asymptotic covariance matrix Σ if there exists a non-random r-dimensional vector t such that

$$\mathcal{L}\big(\sqrt{n}(T_n - t)\big) \Rightarrow \Phi_\Sigma \quad (n \to \infty). \tag{16}$$

Examples of asymptotically normal statistics are well known. Under certain conditions, the property of asymptotic normality is inherent in maximum likelihood estimators, sample moments, sample quantiles, etc.

Our nearest aim is to describe the asymptotic behavior of the random elements T_{N_n}, that is, of statistics constructed from samples with random sizes N_n.

Again let $\{d_n\}_{n \geq 1}$ be an infinitely increasing sequence of positive numbers. Now set

$$Z_n = \sqrt{d_n}(T_{N_n} - t).$$

Theorem 3. *Let $N_n \to \infty$ in probability as $n \to \infty$. Assume that a statistic T_n is asymptotically normal in the sense of* (16) *with an asymptotic covariance matrix Σ. Then a distribution F such that*

$$\mathcal{L}(Z_n) \Rightarrow F \qquad (n \to \infty),$$

exists if and only if there exists a distribution function $V(x)$ satisfying the conditions.

(i) $V(x) = 0$ *for* $x < 0$;

(ii) *for any* $A \in \mathcal{B}(\mathbb{R}^r)$

$$F(A) = \int_0^\infty \Phi_{u^{-1}\Sigma}(A)\, dV(u), \quad x \in \mathbb{R}^1;$$

(iii) $\mathrm{P}(N_n < d_n x) \Rightarrow V(x)$, $n \to \infty$.

The proof of Theorem 3 relies on Theorem 1 with (16) playing the role of (3) and Lemma 1 and differs from the proof of Theorem 2 only by that b_{N_n}/d_n is replaced by d_n/N_n.

Corollary 2. *Under the conditions of Theorem* 3 *the statistic T_{N_n} is asymptotically normal with some covariance matrix Σ' if and only if there exists a number $c > 0$ such that*

$$\frac{N_n}{d_n} \Rightarrow c \quad (n \to \infty).$$

Moreover, in this case, $\Sigma' = c^{-1}\Sigma$.

This statement immediately follows from Theorem 2 with the account of Lemma 1.

4. Some remarks on the heavy-tailedness of scale mixtures of normals

The one-dimensional marginals of the multivariate limit law in Theorems 2 and 3 are scale mixtures of normals with zero means of the form $\mathrm{E}\Phi(x/U)$, $x \in \mathbb{R}$, where $\Phi(x)$ is the standard normal distribution function and U is a nonnegative random variable. It turns out, although absolutely not so evident, that these distributions are *always* leptokurtic having sharper vertex and heavier tails than the normal law itself.

It is easy to see that

$$\mathrm{E}\Phi(x/U) = \mathrm{P}(X \cdot U < x), \quad x \in \mathbb{R},$$

where X is a standard normal variable independent of U. First, as a measure of leptokurtosity, consider the excess coefficient which is traditionally used in (descriptive) statistics. Recall that for a random variable Y with $\mathrm{E}Y^4 < \infty$, the excess coefficient (kurtosis) $\kappa(Y)$ is defined as

$$\kappa(Y) = \mathrm{E}\left(\frac{Y - \mathrm{E}Y}{\sqrt{\mathrm{D}Y}}\right)^4.$$

If $\mathrm{P}(X < x) = \Phi(x)$, then $\kappa(X) = 3$. Densities with sharper vertices (and, respectively, with heavier tails) than the normal density, have $\kappa > 3$, and $\kappa < 3$ for densities with more flat vertices.

Lemma 2. *Let X and U be independent random variables with finite fourth moments; moreover, let* $\mathrm{E}X = 0$ *and* $\mathrm{P}(U \geq 0) = 1$. *Then*

$$\kappa(XU) \ \geq \ \kappa(X).$$

Furthermore, $\kappa(XU) = \kappa(X)$ if and only if $\mathrm{P}(U = \mathrm{const}) = 1$.

For the proof see [10].

So, if X is a standard normal random variable and U is a nonnegative random variable with $\mathrm{E}U^4 < \infty$ independent of X, then $\kappa(X \cdot U) \geq 3$ and $\kappa(X \cdot U) = 3$ if and only if U is non-random.

Using the Jensen inequality, we can easily obtain one more inequality directly connecting the tails of the normal mixtures with the tails of the normal distribution.

Lemma 3. *Assume that the random variable U satisfies the normalization condition* $\mathrm{E}U^{-1} = 1$. *Then*

$$1 - \mathrm{E}\Phi(x/U) \geq 1 - \Phi(x), \qquad x > 0.$$

From Lemma 3, it follows that if X is the standard normal random variable and U is a nonnegative random variable independent of X with $\mathrm{E}U^{-1} = 1$, then for any $x \geq 0$

$$\mathrm{P}(|X \cdot U| \geq x) \ \geq \ \mathrm{P}(|X| \geq x) \ \ (= 2[1 - \Phi(x)]),$$

that is, scale mixtures of normal laws are always more leptokurtic and have heavier tails than normal laws themselves.

The class of scale mixtures of normal laws is very rich and involves distributions with various character of decrease of tails. For example, this class contains Student distributions with arbitrary (not necessarily integer) number of degrees of freedom (and the Cauchy distribution included), symmetric stable distributions (see the "multiplication theorem" 3.3.1 in [15]), symmetric fractional stable distributions (see [16]), symmetrized gamma distributions with arbitrary shape and scale parameters (see [10]), and symmetrized Weibull distributions with shape parameters belonging to the interval $(0, 1]$ (see [17, 18]). As an example, in the next section, we will discuss the conditions for the convergence of the distributions of the statistics constructed from samples with random sizes to the multivariate Student distribution.

5. Convergence of the distributions of random sums of random vectors with finite covariance matrices to multivariate elliptically contoured stable and Linnik distributions

5.1 Convergence of the distributions of random sums of random vectors to multivariate stable laws

Let Σ be a positive definite $(r \times r)$-matrix, $\alpha \in (0, 2]$. A random vector $Z_{\alpha,\Sigma}$ is said to have the (centered) elliptically contoured stable distribution $G_{\alpha,\Sigma}$ with characteristic exponent α, if its characteristic function $\mathfrak{g}_{\alpha,\Sigma}(\mathbf{t})$ has the form

$$\mathfrak{g}_{\alpha,\Sigma}(\mathbf{t}) \equiv \mathrm{E}\exp\left\{i\mathbf{t}^{\top}X\right\} = \exp\left\{-\left(\mathbf{t}^{\top}\Sigma\mathbf{t}\right)^{\alpha/2}\right\}, \quad \mathbf{t} \in \mathbb{R}^r.$$

Univariate stable distributions are popular examples of heavy-tailed distributions. Their moments of orders $\delta \geq \alpha$ do not exist (the only exception is the normal law corresponding to $\alpha = 2$). Stable laws and only they can be limit distributions for sums of a non-random number of independent identically distributed random variables with infinite variance under linear normalization. Here it will be shown that they also can be limiting for *random* sums of random vectors with *finite covariance matrices*. The result of this subsection generalizes the main theorem of [19] to a multivariate case.

By ζ_{α}, we will denote a positive random variable with the one-sided stable distribution corresponding to the characteristic function

$$\mathfrak{g}_{\alpha}(t) = \exp\left\{-|t|^{\alpha}\exp\left\{-\frac{1}{2}i\pi\alpha sign t\right\}\right\}, \quad t \in \mathbb{R},$$

with $0 < \alpha \leq 1$ (for more details see [15] or [4]).

Let $\alpha \in (0, 2]$. It is known that, if Y is a random vector such that $\mathcal{L}(Y) = \Phi_{\Sigma}$ independent of the random variable $\zeta_{\alpha/2}$, then

$$Z_{\alpha,\Sigma} \stackrel{d}{=} \sqrt{\zeta_{\alpha/2}} \cdot Y \tag{17}$$

(see Proposition 2.5.2 in [4]). In other words,

$$G_{\alpha,\Sigma} = \mathrm{E}\Phi_{\zeta_{\alpha/2}\Sigma}. \tag{18}$$

As in Section 3, let $X_1, X_2, \ldots$ be independent r-valued random vectors. For $n \in \mathbb{N}$, denote $S_n = X_1 + \ldots + X_n$. Consider a sequence of integer-valued positive

random variables $\{N_n\}_{n\geq 1}$ such that for each $n\geq 1$ the random variable N_n is independent of the sequence $\{S_k\}_{k\geq 1}$. Let $\{b_n\}_{n\geq 1}$ be an infinitely increasing sequence of positive numbers providing convergence (6) with some positive definite matrix Σ.

Theorem 4. *Let $N_n \to \infty$ in probability as $n \to \infty$. Assume that the random variables $X_1, X_2, \ldots$ satisfy condition (7) with an asymptotic covariance matrix Σ. Then*

$$L\left(\frac{S_{N_n}}{\sqrt{d_n}}\right) \Rightarrow G_{\alpha,\Sigma} \qquad (n \to \infty)$$

with some infinitely increasing sequence of positive numbers $\{d_n\}_{n\geq 1}$ and some $\alpha\in(0,2]$, if and only if

$$\frac{N_n}{d_n} \Rightarrow \zeta_{\alpha/2,1}$$

as $n \to \infty$.

Proof. This theorem is a direct consequence of Theorem 2 with the account of relations (17) and (18).

5.2 Convergence of the distributions of random sums of random vectors with finite covariance matrices to multivariate elliptically contoured Linnik distributions

In 1953, Yu. V. Linnik [20] introduced the class of univariate symmetric probability distributions defined by the characteristic functions

$$f_\alpha^L(t) = \frac{1}{1+|t|^\alpha}, \quad t\in\mathbb{R},$$

where $\alpha\in(0,2]$. Later, the distributions of this class were called *Linnik distributions* [21] or *α-Laplace distributions* [22]. Here the first term will be used since it has become conventional. With $\alpha = 2$, the Linnik distribution turns into the Laplace distribution corresponding to the density

$$f^\Lambda(x) = \frac{1}{2}e^{-|x|}, \qquad x\in\mathbb{R}.$$

A random variable with the Linnik distribution with parameter α will be denoted $L_{1,\alpha}$.

The Linnik distributions possess many interesting analytic properties (see, e.g., [17, 18] and the references therein) but, perhaps, most often Linnik distributions are recalled as examples of geometric stable distributions often used as heavy-tailed models of some statistical regularities in financial data [23, 24].

The multivariate Linnik distribution was introduced by D. N. Anderson in [25] where it was proved that the function

$$f_{\alpha,\Sigma}^{(L)}(\mathbf{t}) = \frac{1}{1+(\mathbf{t}^\top\Sigma\mathbf{t})^{\alpha/2}}, \quad \mathbf{t}\in\mathbb{R}^r, \quad \alpha\in(0,2), \tag{19}$$

is the characteristic function of an r-variate probability distribution, where Σ is a positive definite $(r\times r)$-matrix. In [25], the distribution corresponding to the

characteristic function (19) was called *the r-variate Linnik distribution*. For the properties of the multivariate Linnik distributions, see [25, 26].

The r-variate Linnik distribution can also be defined in another way. For this purpose, recall that the distribution of a nonnegative random variable M_δ whose Laplace transform is

$$\psi_\delta(s) \equiv \mathrm{E}e^{-sM_\delta} = \frac{1}{1+s^\delta}, \qquad s \geq 0, \tag{20}$$

where $0 < \delta \leq 1$, is called *the Mittag-Leffler distribution*. It is another example of heavy-tailed geometrically stable distributions; for more details see for example, [17, 18] and the references therein. The Mittag-Leffler distributions are of serious theoretical interest in the problems related to thinned (or rarefied) homogeneous flows of events such as renewal processes or anomalous diffusion or relaxation phenomena, see [27, 28] and the references therein. In [18], it was demonstrated that

$$L_{1,\alpha} \stackrel{d}{=} Y_1 \cdot \sqrt{2M_{\alpha/2}}, \tag{21}$$

where Y_1 is a random variable with the standard univariate normal distribution independent of the random variable $M_{\alpha/2}$ with the Mittag-Leffler distribution with parameter $\alpha/2$.

Now let Y be a random vector such that $\mathcal{L}(Y) = \Phi_\Sigma$, where Σ is a positive definite $(r \times r)$-matrix, independent of the random variable $M_{\alpha/2}$. By analogy with (21), introduce the random vector $L_{r,\alpha,\Sigma}$ as

$$L_{r,\alpha,\Sigma} = \sqrt{2M_{\alpha/2}} \cdot Y.$$

Then, in accordance with what has been said in Section 2,

$$\mathcal{L}(L_{r,\alpha,\Sigma}) = E\Phi_{2M_{\alpha/2}\Sigma}. \tag{22}$$

The distribution (14) will be called *the (centered) elliptically contoured multivariate Linnik distribution*.

Using Remark 1, we can easily make sure that the two definitions of the multivariate Linnik distribution coincide. Indeed, with the account of (20), according to Remark 1, the characteristic function of the random vector $L_{r,\alpha,\Sigma}$ defined by (22) has the form

$$\mathrm{E}\exp\left\{i\mathbf{t}^\top L_{r,\alpha,\Sigma}\right\} = \psi_{\alpha/2}\left(\mathbf{t}^\top\Sigma\mathbf{t}\right) = \frac{1}{1+\left(\mathbf{t}^\top\Sigma\mathbf{t}\right)^{\alpha/2}} = \mathrm{f}^{(L)}_{\alpha,\Sigma}(\mathbf{t}), \quad \mathbf{t}\in\mathbb{R}^r,$$

that coincides with Anderson's definition (19).

Our definition (22) together with Theorem 2 opens the way to formulate a theorem stating that the multivariate Linnik distribution can not only be limiting for geometric random sums of independent identically distributed random vectors with infinite second moments [29], but it can also be limiting for random sums of independent random vectors with finite covariance matrices.

Theorem 5. *Let $N_n \to \infty$ in probability as $n \to \infty$. Assume that the random variables $X_1, X_2, \ldots$ satisfy condition (7) with an asymptotic covariance matrix Σ. Then*

$$L\left(\frac{S_{N_n}}{\sqrt{d_n}}\right) \Rightarrow \mathcal{L}(L_{r,\alpha,\Sigma}) \qquad (n \to \infty)$$

with some infinitely increasing sequence of positive numbers $\{d_n\}_{n\geq 1}$ and some $\alpha \in (0, 2]$, if and only if

$$\frac{N_n}{d_n} \Rightarrow 2M_{\alpha/2}$$

as $n \to \infty$.

Proof. This theorem is a direct consequence of Theorem 2 with the account of relation (22).

6. Convergence of the distributions of asymptotically normal statistics to the multivariate Student distribution

The multivariate Student distribution is described, for example, in [30] (also see [31]). Consider an r-dimensional normal random vector Y with zero vector of expectations and covariance matrix Σ. Assume that a random variable W_γ has the chi-square distribution with parameter (the "number of degrees of freedom") $\gamma > 0$ (not necessarily integer) and is independent of Y. The distribution $P_{\gamma,\Sigma}$ of the random vector

$$Q_{\gamma,\Sigma} = \sqrt{\gamma/W_\gamma} \cdot Y \tag{23}$$

is called the *multivariate Student distribution* (with parameters γ and Σ). For any $\mathbf{x} \in \mathbb{R}^r$ the distribution density of Z has the form

$$p_{\gamma,\Sigma}(\mathbf{x}) = \frac{\Gamma(r+\gamma)/2)}{|\Sigma|^{1/2}\Gamma(\gamma/2)(\pi\gamma)^{r/2}} \cdot \frac{1}{\left(1+\frac{1}{\gamma}\mathbf{x}^T\Sigma^{-1}\mathbf{x}\right)^{(r+\gamma)/2}}.$$

According to Theorem 3, the multivariate Student distribution is the resulting transformation of the limit distribution of an asymptotically normal (in the sense of (16)) statistic under the replacement of the sample size by a random variable whose asymptotic distribution is chi-square. Consider this case in more detail.

Let $G_{m,m}(x)$ be the gamma-distribution function with the shape parameter coinciding with the scale parameter and equal to m:

$$G_{m,m}(x) \;=\; \begin{cases} 0 & \text{if } x \leq 0, \\ \dfrac{m^m}{\Gamma(m)}\displaystyle\int_0^x e^{-my} y^{m-1} dy & \text{if } x > 0. \end{cases}$$

Theorem 6. *Let $\gamma > 0$ be arbitrary, Σ be a positive definite matrix and let $\{d_n\}_{n\geq 1}$ be an infinitely increasing sequence of positive numbers. Assume that $N_n \to \infty$ in probability as $n \to \infty$. Let a statistic T_n be asymptotically normal in the sense of* (16). *Then the convegence*

$$\mathcal{L}\left(\sqrt{d_n}(T_{N_n} - t)\right) \;\Rightarrow P_{\gamma,\Sigma} \qquad (n \to \infty),$$

takes place if and only if

$$\mathrm{P}(N_n < d_n x) \Rightarrow G_{\gamma/2,\gamma/2}(x), \quad n \to \infty,$$

where $G_{\gamma/2,\gamma/2}(x)$ is the gamma-distribution function with coinciding shape and scale parameters equal to $\gamma/2$.

Proof. This statement is a direct consequence of Theorem 3, representation (23) and Lemma 1.

Let $N_{p,m}$ be a random variable with the negative binomial distribution

$$\mathrm{P}\big(N_{p,m} = k\big) \;=\; C_{m+k-2}^{k-1} p^m (1-p)^{k-1}, \qquad k = 1, 2, \ldots \tag{24}$$

Here $m > 0$ and $p \in (0,1)$ are parameters; for non-integer m, the quantity C_{m+k-2}^{k-1} is defined as

$$C_{m+k-2}^{k-1} \;=\; \frac{\Gamma(m+k-1)}{(k-1)! \cdot \Gamma(m)}.$$

In particular, for $m = 1$, relation (24) determines the geometric distribution. It is well known that

$$\mathrm{E}N_{p,m} \;=\; \frac{m(1-p)+p}{p},$$

so that $\mathrm{E}N_{p,m} \to \infty$ as $p \to 0$.

As is known, the negative binomial distribution with natural m admits an illustrative interpretation in terms of Bernoulli trials. Namely, the random variable with distribution (24) is the number of the Bernoulli trials held up to the mth failure, if the probability of the success in a trial is $1-p$.

Lemma 4. *For any fixed $m > 0$*

$$\lim_{p \to 0} sup_{x \in \mathbb{R}} \left| \mathrm{P}\left(\frac{N_{p,m}}{\mathrm{E}N_{p,m}} < x \right) - G_{m,m}(x) \right| \;=\; 0,$$

where $G_{m,m}(x)$ is the gamma-distribution function with the shape parameter coinciding with the scale parameter and equal to m.

The proof is a simple exercise on characteristic functions; for more details, see [8].

Corollary 3. *Let $m > 0$ be arbitrary. Assume that for each $n \geq 1$ the random variable N_n has the negative binomial distribution with parameters $p = \frac{1}{n}$ and m. Let a statistic T_n be asymptotically normal in the sense of* (16). *Then*

$$\mathcal{L}\big(\sqrt{mn}(T_{N_n} - t)\big) \;\Rightarrow P_{2m,\Sigma} \qquad (n \to \infty)$$

where $P_{2m,\Sigma}$ is the r-variate Student distribution with parameters $\gamma = 2m$ and Σ.

Proof. By Lemma 4 we have

$$\frac{N_n}{nm} \;=\; \frac{N_n}{\mathrm{E}N_n} \cdot \frac{\mathrm{E}N_n}{nm} = \frac{N_n}{\mathrm{E}N_n} \cdot \frac{m(n-1)+1}{mr} \;= \frac{N_n}{\mathrm{E}N_n}\left[1 + O\left(\frac{1}{n}\right)\right] \Rightarrow U_m$$

as $n \to \infty$ where U_m is the random variable having the gamma-distribution function with coinciding shape and scale parameters equal to m. Now the desired assertion directly follows from Theorem 6.

Remark 2. The r-variate Cauchy distribution ($\gamma = 1$) appears in the situation described in Corollary 2 when the sample size N_n has the negative binomial distribution with the parameters $p = \frac{1}{n}$, $m = \frac{1}{2}$, and n is large.

Remark 3. In the case where the sample size N_n has the negative binomial distribution with the parameters $p = \frac{1}{n}$, $m = 1$ (that is, the geometric distribution with the parameter $p = \frac{1}{n}$), then, as $n \to \infty$, we obtain the limit r-variate Student distribution with parameters $\gamma = 2$ and Σ. Moreover, if $\Sigma = I_r$ (that is, the r-variate Student distribution is spherically symmetric), then its one-dimensional marginals have the form (1). As we have already noted, distribution (1) was apparently for the first time introduced as a limit distribution for the sample median in a sample with geometrically distributed random size in [11]. It is worth noticing that in the cited paper [11], distribution (1) was not identified as the Student distribution with two degrees of freedom.

Thus, the main conclusion of this section can be formulated as follows. If the number of random factors that determine the observed value of a random variable is random itself with the distribution that can be approximated by the gamma distribution with coinciding shape and scale parameters (e.g., is negative binomial with probability of success close to one, see Lemma 4), then those functions of the random factors that are regarded as asymptotically normal in the classical situation are actually asymptotically Student with considerably heavier tails. Hence, since gamma-models and/or negative binomial models are widely applicable (to confirm this it may be noted that the negative binomial distribution is mixed Poisson with mixing gamma distribution, this fact is widely used in insurance), the Student distribution can be used in descriptive statistics as a rather reasonable heavy-tailed asymptotic approximation.

7. The asymptotic distribution of sample quantiles in samples with sizes generated by a Cox process

Sometimes, when the performance of a technical or financial system is analyzed, a forecast of main characteristics is made on the basis of data accumulated during a certain period of the functioning of the system. As a rule, data are accumulated as a result of some "informative events" that occur during this period. For example, inference concerning the distribution of insurance claims, which is very important for the estimation of, say, the ruin probability of an insurance company, is usually performed on the basis of the statistic $W_1, W_2, \ldots, W_{N(T)}$ of the values of insurance claims arrived within a certain time interval $[0, T]$ (here $N(T)$ denotes the number of claims arrived during the time interval $[0, T]$). Moreover, this inference is typically used for the prediction of the value of the ruin probability for the next period $[T, 2T]$. But it is obvious (at least in the example above) that the observed number of informative events occurred during the time interval $[0, T]$ is actually a realization of a random variable, because both the number of insurance claims arrived within this interval follow a stochastic counting process. If the random character of the number of available observations is not taken into consideration, then all what can be done is the *conditional* forecast. To obtain a complete prediction with the account of the randomness of the number of "informative events," we should use the results similar to Theorems 2 and 3. One of rather realistic and general assumptions concerning $N(t)$, the number of observations accumulated by the time t, is that $N(t)$ is a Cox process. In this section, as an example, we will consider the asymptotic behavior of sample quantiles constructed from a sample whose size is determined by a Cox process. As we have already noted in the introduction, this problem is very

important for the proper application of such risk measures as VaR (Value-at-Risk) in, say, financial engineering.

Let $W_1, \ldots, W_n$, $n \geq 1$, be independent identically distributed random variables with common distribution density $p(x)$ and $W_{(1)}, \ldots, W_{(n)}$ be the corresponding order statistics, $W_{(1)} \leq W_{(2)} \leq \ldots \leq W_{(n)}$. Let $r \in \mathbb{N}$ $\lambda_1, \ldots, \lambda_r$ be some numbers such that $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_r < 1$. The quantiles of orders $\lambda_1, \ldots, \lambda_r$ of the random variable W_1 will be denoted ξ_{λ_i}, $i = 1, \ldots, r$. The sample quantiles of orders $\lambda_1, \ldots, \lambda_r$ are the random variables $W_{([\lambda_i n]+1)}$, $i = 1, \ldots, r$, with $[a]$ denoting the integer part of a number a. The following result due to Mosteller [32] (also see [33], Section 9.2) is classical. Denote

$$Y_{n,j}^{*} = \sqrt{n}\left(W_{([\lambda_j n]+1)} - \xi_{\lambda_j}\right), \qquad j = 1, \ldots, r.$$

Theorem 7 [32]. *If $p(x)$ is differentiable in some neighborhoods of the quantiles ξ_{λ_i} and $p\left(\xi_{\lambda_i}\right) \neq 0$, $i = 1, \ldots, r$, then, as $n \to \infty$, the joint distribution of the normalized sample quantiles $Y_{n,1}^{*}, \ldots, Y_{n,r}^{*}$ weakly converges to the r-variate normal distribution with zero vector of expectations and covariance matrix $\Sigma = \left(\sigma_{ij}\right)$,*

$$\sigma_{ij} = \frac{\lambda_i\left(1-\lambda_j\right)}{p\left(\xi_{\lambda_i}\right)p\left(\xi_{\lambda_j}\right)}, \quad i \leq j.$$

To take into account the randomness of the sample size, consider the sequence $W_1, W_2 \ldots$ of independent identically distributed random variables with common distribution density $p(x)$.

Let $N(t)$, $t \geq 0$, be a Cox process controlled by a process $\Lambda(t)$. Recall the definition of a Cox process. Let $N_1(t)$, $t \geq 0$, be a standard Poisson process (i.e., a homogeneous Poisson process with unit intensity). Let $\Lambda(t)$, $t \geq 0$, be a random process with non-decreasing right-continuous trajectories, $\Lambda(0) = 0$, $\mathrm{P}(\Lambda(t) < \infty) = 1$ for all $t > 0$. Assume that the processes $\Lambda(t)$ and $N_1(t)$ are independent. Set

$$N(t) = N_1(\Lambda(t)), \qquad t \geq 0.$$

The process $N(t)$ is called a doubly stochastic Poisson process (or a Cox process) controlled by the process $\Lambda(t)$. The one-dimensional distributions of a Cox process are mixed Poisson. For example, if $\Lambda(t)$ has the gamma distribution, then $N(t)$ has the negative binomial distribution.

Cox processes are widely used as models of inhomogeneous chaotic flows of events, see, for example, [2].

Assume that all the involved random variables and processes are independent. In this section, under the assumption that $\Lambda(t) \to \infty$ in probability, the asymptotics of the joint distribution of the random variables $W_{([\lambda_i N(t)]+1)}$, $i = 1, \ldots, r$ is considered as $t \to \infty$.

As we have already noted, it was B. V. Gnedenko who drew attention to the essential distinction between the asymptotic properties of sample quantiles constructed from samples with random sizes and the analogous properties of sample quantiles in the standard situation. Briefly recall the history of the problem under consideration. B. V. Gnedenko, S. Stomatovič, and A. Shukri [34] obtained sufficient conditions for the convergence of distribution of the sample median constructed from sample of random size. In the candidate (PhD) thesis of A. K. Shukri, these conditions were extended to quantiles of arbitrary orders. In [35], necessary and sufficient conditions for the weak convergence of the

one-dimensional distributions of sample quantiles in samples with random sizes were obtained.

Our aim here is to give necessary and sufficient conditions for the weak convergence of the *joint* distributions of sample quantiles constructed from samples with random sizes driven by a Cox process and to describe the r-variate limit distributions emerging here, thus extending Mosteller's Theorem 4 to samples with random sizes. The results of this section extend those of [36] to the continuous-time case.

Lemma 5. *Let $N(t)$ be a Cox process controlled by the process $\Lambda(t)$. Then $N(t) \xrightarrow{P} \infty$ $(t \to \infty)$ if and only if $\Lambda(t) \xrightarrow{P} \infty$ $(t \to \infty)$.*

Lemma 6. *Let $N(t)$ be a Cox process controlled by the process $\Lambda(t)$. Let $d(t) > 0$ be a function such that $d(t) \to \infty$ $(t \to \infty)$. Then the following conditions are equivalent:*

1. One-dimensional distributions of the normalized Cox process weakly converge to the distribution of some random variable Z as $t \to \infty$:

$$\frac{N(t)}{d(t)} \Rightarrow Z \qquad (t \to \infty).$$

2. One-dimensional distributions of the controlling process $\Lambda(t)$, appropriately normalized, converge to the same distribution:

$$\frac{\Lambda(t)}{d(t)} \Rightarrow Z \qquad (t \to \infty).$$

For the proof of Lemmas 5 and 6 see [37].

Now we proceed to the main results of this section. In addition to the notation introduced above, for positive integer n set $Q_j(n) = W_{([\lambda_j n]+1)}$, $j = 1, \ldots, r$, $Q(n) = (Q_1(n), \ldots, Q_r(n))$, $\xi = (\xi_{\lambda_1}, \ldots, \xi_{\lambda_r})$. Let $d(t)$ be an infinitely increasing positive function. Set

$$Z(t) = \sqrt{d(t)}(Q(N(t)) - \xi).$$

Theorem 8. *Let $\Lambda(t) \xrightarrow{P} \infty$ as $t \to \infty$. If $p(x)$ is differentiable in neighborhoods of the quantiles ξ_{λ_i} and $p(\xi_{\lambda_i}) \neq 0$, $i = 1, \ldots, r$, then the convergence*

$$Z(t) \Rightarrow Z \quad (t \to \infty),$$

to some random vector Z takes place, if and only if there exists a nonnegative random variable U such that

$$\mathrm{P}(Z \in A) = \mathrm{E}\Phi_{U^{-1}\Sigma}(A), \quad A \in \mathcal{B}(\mathbb{R}^r),$$

where $\Sigma = (\sigma_{ij})$,

$$\sigma_{ij} = \frac{\lambda_i(1-\lambda_j)}{p(\xi_{\lambda_i})p(\xi_{\lambda_j})}, \quad i \le j,$$

and

$$\frac{\Lambda(t)}{d(t)} \Rightarrow U \quad (t \to \infty).$$

The proof is a simple combination of Lemmas 1, 5, and 6 and Theorem 3.

Corollary 4. *Under the conditions of Theorem 8, the joint distribution of the normalized sample quantiles* $\sqrt{d(t)}\left(W_{([\lambda_j N(t)]+1)} - \xi_{\lambda_j}\right)$, $j = 1, \ldots, r$, *weakly converges to the r-variate normal law with zero expectation and covariance matrix* Σ, *if and only if*

$$\frac{\Lambda(t)}{d(t)} \Rightarrow 1 \quad (t \to \infty).$$

This statement immediately follows from Theorem 8 with the account of Lemma 1.

Corollary 5. *Under the conditions of Theorem 8, the joint distribution of the normalized sample quantiles* $\sqrt{d(t)}\left(W_{([\lambda_j N(t)]+1)} - \xi_{\lambda_j}\right)$, $j = 1, \ldots, r$, *weakly converges to the r-variate Student distribution with parameters* $\gamma > 0$ *and* Σ *defined in Theorem 4, if and only if*

$$\mathrm{P}(\Lambda(t) < xd(t)) \Rightarrow G_{\gamma/2,\gamma/2}(x), \quad t \to \infty,$$

where $G_{\gamma/2,\gamma/2}(x)$ is the gamma-distribution function with coinciding shape and scale parameters equal to $\gamma/2$.

Let $0 < \lambda < 1$ and let ξ_λ be the λ-quantile of the random variable W_1. As above, the standard normal distribution function will be denoted $\Phi(x)$.

8. Conclusion

The purpose of the chapter was to give a possible explanation of the emergence of heavy-tailed distributions that are often observed in practice instead of the expected normal laws. As the base for this explanation, limit theorems for random sums and statistics constructed from samples with random sizes were considered. Within this approach, it becomes possible to obtain arbitrarily heavy tails of the data distributions without assuming the non-existence of the moments of the observed characteristics. Some comments were made on the heavy-tailedness of scale mixtures of normal distributions. Two general theorems presenting necessary and sufficient conditions for the convergence of the distributions of random sums of random vectors and multivariate statistics constructed from samples with random sizes were proved. As examples of the application of these general theorems, conditions were presented for the convergence of the distributions of random sums of independent random vectors with finite covariance matrices to multivariate elliptically contoured stable and Linnik distributions. An alternative definition of the latter was proposed. Also, conditions were presented for the convergence of the distributions of asymptotically normal (in the traditional sense) statistics to multivariate elliptically contoured Student distributions when the sample size is replaced by a random variable. The joint asymptotic behavior of sample quantiles in samples with random sizes was considered. Special attention was paid to the continuous-time case assuming that the sample size increases in time following a Cox process resulting in the sample size having the mixed Poisson distribution.

Acknowledgements

Supported by Russian Science Foundation, project 18-11-00155.

IntechOpen

Author details

Victor Korolev[1,2,3*] and Alexander Zeifman[3,4]

1 Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia

2 Hangzhou Dianzi University, Hangzhou, China

3 Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, Moscow, Russia

4 Vologda State University, Vologda, Russia

*Address all correspondence to: vkorolev@cs.msu.ru

IntechOpen

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] Shiryaev AN. Foundations of Financial Mathematics. Vol. 1. Facts, Models. – Singapore: World Scientific; 1998

[2] Bening V, Korolev V. Generalized Poisson Models and Their Application in Insurance and Finance. Utrecht: VSP; 2002

[3] Meerschaert MM, Scheffler H-P. Limit theorems for continuous-time random walks with infinite mean waiting times. Journal of Applied Probability. 2004;**41**(3):623-638

[4] Samorodnitsky G, Taqqu MS. Stable Non-Gaussian Random Processes, Stochastic Models with Infinite Variance. New York: Chapman and Hall; 1994

[5] Chen J. From the central limit theorem to heavy-tailed distributions. Journal of Applied Probability. 2003; **40**(3):803-806

[6] Korolev VY. Convergence of random sequences with independent random indices. I. Theory of Probability and its Applications. 1994;**39**(2):313-333

[7] Schluter C, Trede M. Weak convergence to the Student and Laplace distributions. Journal of Applied Probability. 2016;**53**:121-129

[8] Bening V, Korolev V. On an application of the Student distribution in the theory of probability and mathematical statistics. Theory of Probability and its Applications. 2005; **49**(3):377-391

[9] Bening V, Korolev V. Some statistical problems related to the Laplace distribution. Informatics and its Apllications. 2008;**2**(2):19-34

[10] Gnedenko BV, Korolev V. Random Summation: Limit Theorems and Applications. Boca Raton: CRC Press; 1996

[11] Gnedenko BV. On estimation of the unknown parameters of distributions from a random number of independent observations. Probability Theory and Mathematical Statistics. Proccedings of Tbilisi Mathematical Institute named after A. M. Razmadze. 1989;**24**:146-150 (in Russian)

[12] Kolmogorov AN. The method of median in the theory of errors. Matematicheskii Sbornik. 1931;**38**(3/4): 47-50

[13] Korolev VY, Kossova EV. On limit distributions of randomly indexed multidimensional random sequences with an operator normalization. Journal of Mathematical Sciences. 1992;**72**(1): 2915-2929

[14] Korolev VY, Kossova EV. Convergence of multidimensional random sequences with independent random indices. Journal of Mathematical Sciences. 1995;**76**(2): 2259-2268

[15] Zolotarev VM. One-Dimensional Stable Distributions. Providence: American Mathematical Society; 1986

[16] Kolokoltsov V, Korolev V, Uchaikin V. Fractional stable distributions. Journal of Mathematical Sciences. 2001;**105**(6):2569-2576

[17] Korolev VY, Zeifman AI. A note on mixture representations for the Linnik and Mittag-Leffler distributions and their applications. Journal of Mathematical Sciences. 2017;**218**(3): 314-327

[18] Korolev VY, Zeifman AI. Convergence of statistics constructed from samples with random sizes to the Linnik and Mittag-Leffler distributions

and their generalizations. Journal of the Korean Statistical Society. 2017;**46**(2): 161-181 Available online 25 July 2016. Also available on arXiv:1602.02480v1 [math.PR]

[19] Korolev VY. On the convergence of distributions of random sums of independent random variables to stable laws. Theory of Probability and its Applications. 1998;**42**(4):695-696

[20] Linnik YV. Linear forms and statistical criteria, I, II. Selected Translations in Mathematical Statistics and Probability. 1963;**3**:41-90 (Original paper appeared in: Ukrainskii Matematicheskii Zhournal, 1953. Vol. 5. pp. 207-243, 247-290)

[21] Kotz S, Kozubowski TJ, Podgorski K. The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance. Boston: Birkhauser; 2001

[22] Pillai RN. Semi-α-Laplace distributions. Communications in Statistical Theory and Methods. 1985;**14**: 991-1000

[23] Kozubowski TJ, Rachev ST. The theory of geometric stable distributions and its use in modeling financial data. European Journal of Operational Research. 1994;**74**(2):310-324

[24] Kozubowski T, Panorska A. Multivariate geometric stable distributions in financial applications 1. Mathematical and Computer Modelling. 1999;**29**:83-92

[25] Anderson DN. A multivariate Linnik distribution. Statistics & Probability Letters. 1992;**14**:333-336

[26] Ostrovskii IV. Analytic and asymptotic properties of multivariate Linnik's distribution. Mathematical Physics, Analysis and Geometry. 1995; **2**(3):436-455

[27] Weron K, Kotulski M. On the Cole-Cole relaxation function and related Mittag-Leffler distributions. Physica A: Statistical Mechanics and its Applications. 1996;**232**:180-188

[28] Gorenflo R, Mainardi F. Continuous time random walk, Mittag-Leffler waiting time and fractional diffusion: Mathematical aspects, chap. 4. In: Klages R, Radons G, Sokolov IM, editors. Anomalous Transport: Foundations and Applications. Weinheim, Germany: Wiley-VCH; 2008. pp. 93-127. Available at: http://arxiv.org/abs/0705.0797

[29] Kozubowski TJ, Rachev ST. Multivariate geometric stable laws. Journal of Computational Analysis and Applications. 1999;**1**(4):349-385

[30] DeGroot MH. Optimal Statistical Decisions. New York, London: McGraw-Hill Company; 1970

[31] Gupta SS. Bibliography on the Multivariate Normal Integrals and Related Topics. Providence: American Mathematical Society; 1963

[32] Mosteller F. On some useful "inefficient" statistics. Annals of Mathematical Statistics. 1946;**17**: 377-408

[33] David HA. Order Statistics. New York: Wiley; 1970

[34] Gnedenko BV, Stomatovič S, Shukri A. On the distribution of the median. Bulletin of Moscow University, Series Mathematics, Mechanics. 1984; **2**:59-63

[35] Selivanova DO. Estimates of Convergence Rate in Limit Theorems for Random Sums. PhD Thesis: Moscow State University; 1995

[36] Korolev VY. Asymptotic properties of sample quantiles constructed from samples with random sizes. Theory of

Probability and its Applications. 2000;
44(2):394-399

[37] Korolev VY. On convergence of distributions of compound Cox processes to stable laws. Theory of Probability and its Applications. 1999;
43(4):644-650

IntechOpen

IntechOpen