UDC 519.213, 519.214

On Multivariate Geometric Random Sums

I. V. Zolotukhin*

* *St.-Petersburg Department of P.P.Shirshov Institute of Oceanology of RAS, 30, Liniya 1, Vasilievskiy Ostrov, St.-Petersburg, Russia, 199053.*

Abstract. In this paper, we use multivariate geometric distribution to generalize the notion of geometric random sum to the multidimensional case.

The characteristic functions of multivariate random sums are found as well as their projections on an arbitrary coordinate hyperplane. The sufficient conditions for weak convergence of these sums to the Marshall-Olkin multivariate exponential distribution and to the multivariate generalized Laplace distribution are given.

Keywords: multivariate geometric distribution, Marshall-Olkin multivariate exponential distribution, multivariate generalized Laplace distribution, characteristic function.

1. Introduction

Geometric summation arises naturally in the number of fields such as economics, physics, biology, queuing theories. Properties of geometric random sums of the form

$$\sum_{j=1}^{M} X_j,$$

where X_j are i.i.d. random variables, M is a random variable with the geometric distribution: $P(M = n) = pq^{n-1},\ q = 1 - p,\ (n = 1, 2, \dots)$; M and X_j $(j = 1, 2, \dots)$ are independent, have been well studied [1].

It was found that the asymptotic behavior of these sums depends on the expectation $a = \mathsf{E}X_j$. If $X_j \geq 0$ and $0 < a < \infty$, then the limit distribution is exponential, and if $a = 0$ and $0 < \mathsf{E}X_j^2 < \infty$, then the limit distribution is the Laplace distribution (see [2], p.86), both after a suitable normalization.

In practice, often occur time series of random vectors. To date have been studied those limit distributions that are approximated by geometric sums in the form (1) for i.i.d. random vectors (see, for example, [3] and [4]). However, in the time series of random vectors occurring, for example, in models of financial mathematics, the number of random variables can be different for different components, and these random numbers are dependent on each other. For example, if we consider the time series the value of portfolios investments, or other assets, then each component

E-mail: igor.zolotukhin@gmail.com.

of the daily changes of these values is the sum of a random number of random changes, some of which affect only one of the components, while others affect to several ones. The result could be a vector sum in the form

$$Z = (Z_1, \ldots, Z_k) = \left(\sum_{j=1}^{M_1} X_1^{(j)}, \ldots, \sum_{j=1}^{M_k} X_k^{(j)} \right),$$

wherein the random indices M_l $(l = 1, \ldots, k)$ dependent with each other. We consider this case. We introduce a new class of multivariate geometric random sums.

Let $\mathcal{E} = \{\varepsilon\}$ is the set of k - dimensional indices $\varepsilon = (\varepsilon^{(1)}, \ldots, \varepsilon^{(k)})$, each coordinate of which is equal to 0 or 1, and $\mathcal{E}_l$ is the set of indices for which $\varepsilon^{(l)} = 1$.

Let N_ε are the independent geometrically distributed random variables:

$$P(N_\varepsilon = n) = p_\varepsilon q_\varepsilon^{n-1}, n = 1, 2, \ldots, p_\varepsilon = 1 - q_\varepsilon,$$

$$M_l = \min_{\varepsilon \in \mathcal{E}_l} \{N_\varepsilon\}, \ l = 1, \ \ldots, \ k.$$

The distribution of the vector $M = (M_1, \ \ldots, \ M_k)$ is introduced and studied in [5], and is called **multivariate geometric distribution (MVG)**.

Survival function of vector M is

$$\bar{P}(\boldsymbol{m}) = \bar{P}(m_1, \ldots, m_k) = P\left(\min_{\varepsilon \in \mathcal{E}_1} \{N_\varepsilon\} > m_1, \ldots, \min_{\varepsilon \in \mathcal{E}_k} \{N_\varepsilon\} > m_k \right) =$$
$$= \prod_{\varepsilon \in \mathcal{E}} q_\varepsilon^{\max_i \varepsilon \boldsymbol{m}}.$$

Here $\boldsymbol{m} = (m_1, \ldots, m_k)$, m_j $(j = 1, \ldots, k)$ are integers, $\varepsilon \boldsymbol{m}$ is the coordinate-wise product of vectors ε and $\boldsymbol{m}$ (here and below, a couple of signs like $\alpha\beta$, standing side by side, will mean their coordinate-wise product).

Multivariate geometric distribution has properties similar to those of one-dimensional geometrical laws. In particular, it has property of absence of aftereffect at the shift of all coordinates on the same value n:

$$P(M > \boldsymbol{n} + \boldsymbol{m} | M > \boldsymbol{n}) = P(M > \boldsymbol{m}), \ \boldsymbol{n} = (n, \ldots, n).$$

Generalized multivariate geometric random sum is called a random vector sum of the form

$$Z = (Z_1, \ldots, Z_k) = \left(\sum_{j=1}^{M_1} X_1^{(j)}, \ldots, \sum_{j=1}^{M_k} X_k^{(j)} \right),$$

where M_l are defined above in (2), $X_r^{(j)}$ $(r = 1, \ldots, k)$ are independent random variables identically distributed for each l with the known characteristic function

$$\mathsf{E}\,\exp(i\,t_r\,X_r) = \phi_r(t_r),$$

and M_l and $X_r^{(j)}$ are independent.

Note that the dependance between the components of the vector Z in the formula (3) is due to dependence in the summation indices M_l, and not because of the dependence between the coordinates of the vectors $X_l^{(j)}$.

Multivariate geometric random sums include the two extreme cases.

For $M_l = N_{\mathbf{1}}$ $(l = 1, \ldots, k)$, $N_\varepsilon = 0$, $\varepsilon \neq \mathbf{1}$, we have the standard geometric vector sums.

And for $M_l = N_{\varepsilon_l}$, where $\varepsilon_l = (0, \ldots, 0, \underset{l}{1}, 0, \ldots, 0)$, $N_\varepsilon = 0$ for $\varepsilon \neq \varepsilon_l$, each component will be a univariate geometric random sum in the form (1), and components of vector Z are independent.

2. Some details about Multivariate Exponential Distribution and Multivariate Laplace Distribution

The random vector $V = (V_1, \ldots, V_k)$ having Marshall-Olkin multivariate exponential distribution (see [6]) is given by the survival function

$$\bar{F}(z) = P(V_1 > z_1, \ldots, V_k > z_k) = \exp\left(-\sum_{\varepsilon\in\mathcal{E}} \lambda_\varepsilon \max_{1\le i\le k} \varepsilon z\right), z_i > 0,$$

here $(\lambda_\varepsilon \geq 0, \varepsilon \in \mathcal{E})$ is a compact notation for the distribution parameters. The class of these distributions will be denoted as $MVE(\lambda_\varepsilon, \varepsilon \in \mathcal{E})$.

Explicit expressions for the characteristic function of the random vector $V \in MVE(\lambda_\varepsilon, \varepsilon \in \mathcal{E})$ as well as its projections to arbitrary coordinate hyperplane, have been received in [8]. If to use notations of the present paper, they will be written down as:

$$\Psi_V(\boldsymbol{t}) = \frac{1}{\sum\limits_{\varepsilon\in\mathcal{E}} \lambda_\varepsilon - i(\boldsymbol{t}, \mathbf{1})} \sum_{\varepsilon\in\mathcal{E}} \lambda_\varepsilon \Psi_V(\bar{\varepsilon}\boldsymbol{t}),$$

$$\Psi_V(\varepsilon\boldsymbol{t}) = \frac{1}{\sum\limits_{\delta:\delta\varepsilon>\mathbf{0}} \lambda_\delta - i(\boldsymbol{t}, \varepsilon)} \sum_{\delta:\delta\varepsilon>\mathbf{0}} \lambda_\delta \Psi_V(\bar{\delta}\varepsilon\boldsymbol{t}).$$

The multivariate generalized Laplace distribution was introduced in [7]. This is a mixture by parameter scale for the k-dimensional normal vector with independent components and zero expectations, if mixing distribution is the Marshall-Olkin multivariate exponential distribution. Random vector with this distribution can be defined as follows:

Let $Y = (Y_1, Y_2, \ldots, Y_k)$ is a normal vector with zero expectation and identity covariance matrix, and $V = (V_1, \ldots, V_k)$, $V \in \mathrm{MVE}\ (\lambda_\varepsilon, \varepsilon \in \mathcal{E})$.

Then the distribution of the vector $W = (\sqrt{V_1}Y_1, \ldots, \sqrt{V_k}Y_k)$ shell be the **multivariate generalized Laplace distribution** denoted by $MGLD(\lambda_\varepsilon, \varepsilon \in \mathcal{E})$.

The characteristic function of the vector W is

$$\Psi_W(\boldsymbol{t}) = \frac{1}{\sum\limits_{\varepsilon\in\mathcal{E}} \lambda_\varepsilon + \frac{1}{2}(\boldsymbol{t}, \boldsymbol{t})} \sum_{\varepsilon\in\mathcal{E}} \lambda_\varepsilon \Psi_W(\bar{\varepsilon}\boldsymbol{t}).$$

Let denote

$$\Phi(t) = \frac{\sum\limits_{\varepsilon\in\mathcal{E}} \lambda_\varepsilon}{\sum\limits_{\varepsilon\in\mathcal{E}} \lambda_\varepsilon + \frac{1}{2}(\boldsymbol{t}, \boldsymbol{t})}.$$

It is the characteristic function of the multivariate symmetric Laplace distribution.

Now we can write down $\Psi_W(t)$ as

$$\Psi_W(\mathbf{t}) = \Phi(\mathbf{t}) \sum_{\varepsilon\in\mathcal{E}} p_\varepsilon \Psi_W(\bar{\varepsilon}\mathbf{t}),$$

where

$$p_\varepsilon = \frac{\lambda_\varepsilon}{\sum\limits_{\varepsilon\in\mathcal{E}} \lambda_\varepsilon}.$$

Hence the distribution of W is the discrete mixture of multivariate symmetrical Laplace distribution with their convolutions with projections of W on the coordinate hyperplanes.

When $\lambda_{\mathbf{1}} > 0$ and $\lambda_\varepsilon = 0$, $\varepsilon \neq \mathbf{1}$, $\varepsilon \in \mathcal{E}$, we have the symmetric Laplace distribution.

When $\lambda_{0,\ldots,0,\underset{l}{1},0,\ldots,0} > 0$ $(l = 1, \ldots, k)$ and $\lambda_\varepsilon = 0$ for other cases $(\varepsilon \in \mathcal{E})$, the coordinates of W will be independent random variables with the univariate Laplace distribution.

The characteristic functions of projection of vector W on the coordinate hyperplane ε are

$$\Psi_W(\varepsilon\boldsymbol{t}) = \frac{1}{\sum\limits_{\delta:\delta\varepsilon>\mathbf{0}} \lambda_\delta + \frac{1}{2}(\varepsilon\boldsymbol{t}, \varepsilon\boldsymbol{t})} \sum_{\delta:\delta\varepsilon>\mathbf{0}} \lambda_\delta \Psi_W(\bar{\delta}\varepsilon\boldsymbol{t}).$$

3. Main Results

Let denote $\Psi_Z(t) = \mathsf{E}\exp(i(t, Z))$ the characteristic function of vector Z; $\ln\phi(\boldsymbol{t}) = (\ln\phi_1(t_1), \ldots, \ln\phi_k(t_k))$ is the coordinate-wise logarithm of the vector function $\phi(\boldsymbol{t}) = (\phi_1(t_1), \ldots, \phi_k(t_k))$; $\boldsymbol{t} = (t_1, \ldots, t_k)$; $\varepsilon_1 \vee \varepsilon_2$ is the vector, each i-th coordinate is $\max\{\varepsilon_1^{(i)}, \varepsilon_2^{(i)}\}$.

Define the partial order on the set $\mathcal{E}$ by the rule:

$$\forall\varepsilon, \delta \in \mathcal{E} \quad \delta \leqslant \varepsilon, \text{ if } \forall j \ (j = 1, \ldots, k) \ \ \delta_j \leqslant \varepsilon_j.$$

Let $\delta < \varepsilon$ if $\delta \leqslant \varepsilon$ and $\delta \neq \varepsilon$. Note that $\forall\delta, \varepsilon \in \mathcal{E} \quad \delta\varepsilon \leqslant \varepsilon$.

Theorem 1. *The characteristic function of the projection of vector Z on the coordinate hyperplane ε is*

$$\Psi_Z(\varepsilon\boldsymbol{t}) = \mathsf{E}\exp(i(\boldsymbol{t},\ \varepsilon Z)) =$$

$$= \frac{\exp(\varepsilon,\ \ln\varphi(\boldsymbol{t}))}{1 - \exp(\varepsilon,\ \ln\varphi(\boldsymbol{t}))\prod\limits_{\gamma:\gamma\varepsilon>\mathbf{0}} q_\gamma} \cdot \sum_{j=1}^{k} \sum_{\substack{\delta_l:\delta_l\varepsilon>\mathbf{0}\\ l=1,\ldots,j}} \prod_{l=1}^{j} p_{\delta_l} \prod_{\substack{\gamma:\gamma\varepsilon>\mathbf{0}\\ \gamma\neq\delta_l\\ l=1,\ldots,j}} q_\gamma \Psi_Z\left(\overline{\bigvee_{l=1}^{j} \delta_l\varepsilon\boldsymbol{t}}\right).$$

Corollary 1.1.

$$\Psi_Z(\boldsymbol{t}) = \mathsf{E}\exp(i(\boldsymbol{t},\ Z)) =$$

$$= \frac{\exp(\mathbf{1},\ \ln\varphi(\boldsymbol{t}))}{1 - \exp(\mathbf{1},\ \ln\varphi(\boldsymbol{t}))\prod_{\delta\in\mathcal{E}} q_\gamma} \cdot \sum_{j=1}^{k} \sum_{\substack{\delta_l\in\mathcal{E}\\ l=1,\ldots,j}} \prod_{l=1}^{j} p_{\delta_l} \prod_{\substack{\gamma\in\mathcal{E}\\ \gamma\neq\delta_l\\ l=1,\ldots,j}} q_\gamma \Psi_Z\left(\overline{\bigvee_{l=1}^{j} \delta_l\boldsymbol{t}}\right).$$

Corollary 1.2. *The characteristic function of the vector Z projection on the axis $\varepsilon = (1, 0, \ldots, 0)$ is*

$$\Psi(\varepsilon\boldsymbol{t}) = \mathsf{E}\exp(i(t_1, Z_1)) = \frac{\varphi_1(t_1)\left(1 - \prod_{\delta\in\mathcal{E}_1} q_\delta\right)}{1 - \varphi_1(t_1)\prod_{\delta\in\mathcal{E}_1} q_\delta}.$$

Now let denote $p_\varepsilon = \lambda_\varepsilon p,\ \varepsilon \in \mathcal{E}$. We shall search the limit distributions of the vector Z as $p \to 0$.

Theorem 2. *Let $X_l^{(j)} \geq 0$; $\mathsf{E}X_l^{(j)} = a_l$; $0 < a_l < \infty$; $l = 1, \ldots, k$, then*

$$\left(\frac{p}{a_1}Z_1, \ldots, \frac{p}{a_k}Z_k\right) = \left(\sum_{j=1}^{M_1} \frac{p}{a_1}X_1^{(j)}, \ldots, \sum_{j=1}^{M_k} \frac{p}{a_k}X_k^{(j)}\right) \xrightarrow[p\to 0]{\mathcal{D}} V,$$

$$V \in MVE(\lambda_\varepsilon, \varepsilon \in \mathcal{E}).$$

Theorem 3. *Let* $\mathsf{E}X_l^{(j)} = 0$; $\mathsf{E}X_l^2 = \sigma_l^2$; $0 < \sigma_l^2 < \infty$; $l = 1, \ldots, k$, *then*

$$\left(\frac{\sqrt{p}}{\sigma_1} Z_1, \ldots, \frac{\sqrt{p}}{\sigma_k} Z_k\right) \xrightarrow[p \to 0]{\mathcal{D}} W,\ W \in MGLD(\lambda_\varepsilon,\ \varepsilon \in \mathcal{E}).$$

4. Conclusions

We present the following results:

- The characteristic function of the vector Z as well as projections of Z on any coordinate hyperplane is found.

- It has been shown that the distributions of the vector Z can be recursively restored by distributions of their univariate components.

- It has been shown that, by analogy with the univariate case, after a suitable normalization the limit distributions will be Marshall-Olkin distribution, if $X_l^{(j)} \geq 0$ and $0 < \mathsf{E}X_l^{(j)} < \infty$, and multivariate Laplace distribution, if $\mathsf{E}X_l^{(j)} = 0$ and $0 < \mathsf{E}(X_l^{(j)})^2 < \infty$.

- Thus, it is shown that these distributions possess the property of geometric stability, under the special scheme of geometric summation.

References

1. *Kalashnikov V.* Geometric Sums: Bounds for Rare Events with Applications. — Dordrecht; Boston: Kluwer Academic, Series: Mathematics and Its Applications, 1997. — Vol. 413. — 270 p.
2. *Korolev V.Y., Bening V.E., Shorgin S.Ya.* Mathematical Faundation of Risk Theory. — M.: Fizmatlit, 2007. — 542 p. (in Russian)
3. *Kozubowski T.G., Rachev S.T.* Multivariate geometric stable laws // J. Comput. Anal. Appl. — 1999. — Vol. 1, no. 4. — P. 349-385.
4. *Meerschaert M.M., Scheffler H.P.* Limit distributions for sums of independent random vectors. — Wiley, New York, 2001.
5. *Zolotukhin I.V.* A new class of geometrically distributed random vectors with dependent components // Vestnik NovSU. Phys.-Math. Sci. — 2014. — Vol. 80. — P. 23–26. (in Russian)
6. *Marshall A.W., Olkin I.A.* A multivariate exponential distribution // J. Amer. Statist. Assoc. — 1967. — Vol. 62 (317). — P. 30–44.
7. *Zolotukhin I.V., Zolotukhina L.A.* New Class of Multivariate generalized Laplace Distributions // Trans. of XXIV International Seminar on Stability Problems for Stochastic Models, Riga, 2004.
8. *Zolotukhin I.V.* The Study of the Laplace Transform of Marshall-Olkin Multivariate Exponential Distribution // Topics in Statistical Simulation. — Springer, New York, 2014. — P. 531–539.