Applied Mathematical Sciences, Vol. 16, 2022, no. 4, 173 - 185
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/ams.2022.916763

On Stable Distribution of Multivariate Data

Phuc Ho Dang [1]

Institute of Mathematics, Vietnam Academy of Science and Technology
18 Hoang Quoc Viet, Cau Giay, Ha Noi, Vietnam

Truc Giang Vo Thi

Tien Giang University
119 Ap Bac, My Tho City, Vietnam

Abstract

To investigate the multidimensional distributions of stable random vectors, the paper deals with a transformation in multidimensional space, which turns a given stable random vector that has a positive density function into a sub-Gaussian random vector. The result can be used to perform a procedure of testing the stable distribution of multivariate data. A dataset collected from the Nasdaq stock market is used to illustrate the proposed procedure.

Keywords: Primary 62H10, 62H15; Secondary 62P05

Keywords: Heavy tailed distribution, Random vector, Stock market, Portfolio selection

Introduction

The traditional statistical analysis methods were developed mostly under the normality assumptions. However, normality is only a poor approximation of reality. Whilst normal distributions are always symmetric around their mean, most of the quantities usually concerned in empirical studies do not have symmetric distributions. Moreover, normal distributions do not allow heavy tails,

[1] Corresponding author

which are common in the reality. Stable distributions are asymmetric heavy-tailed extensions of normal distributions and have attracted a lot of attention in applied research ([1], [6], [7], [9], [14], [15]). Presently, the univariate stable distributions are accessible by several methods to estimate stable parameters and reliable programs to compute stable densities, cumulative distribution functions, and quantiles for stable random variables ([1], [5], [8], [11]). However, the use of the heavy-tailed models in practice has been restricted by the lack of tools for multivariate stable distributions.

Currently, computations are more accessible for elliptically contoured stable distributions [16] which are scale mixtures of multivariate normal distributions. The tools for the special class of stable distributions were applied in several empirical studies ([6], [10]). However, the method is available only for a narrow subclass of multivariate stable distributions. The problem of multivariate distribution function calculation for stable random vectors remains open in the general cases. Besides, in many studies on portfolio selection and asset allocation, analysts must determine the distribution function of a linear combination of several stable random variables. Thus, the problem of testing the stability in distribution of a random vector plays an important role in application. That convinces the aim of this paper to create a new procedure of goodness-of-fit testing for a broader family of multivariate stable distributions.

The rest of the paper is organized as follows. Section 1 presents the main notation and some concepts related to multidimensional stable distributions. Section 2 is devoted to creating a bijective transformation in multidimensional space, that turns a given stable random vector into a sub-Gaussian random vector. Section 3 presents the procedure based on the result obtained in Section 2 to conduct the goodness-of-fit testing on stable distribution of multivariate data. An example of dataset collected from the Nasdaq stock market is used to illustrate the practicability of the procedure.

1 Preliminaries

Given a *random vector* $\mathbf{X} = (X_1, ..., X_d)$ taking values in Euclidean space $\mathbb{R}^d$, its *cumulative distribution function* (cdf or *distribution*) and *probability density function* (pdf hereafter) are denoted by $F_\mathbf{X}$ and $f_\mathbf{X}$, respectively. The coordinates $X_1, ..., X_d$ are called *marginals*, simultaneously $F_{X_1}, ..., F_{X_d}$ and $f_{X_1}, ..., f_{X_d}$ are called *marginal cdf's* and *marginal pdf's* of $\mathbf{X}$, respectively. The *characteristic function* $\varphi_\mathbf{X}$ of $\mathbf{X}$ is defined by

$$\varphi_\mathbf{X}(\mathbf{t}) := \mathbb{E}\exp\{i\langle \mathbf{X}, \mathbf{t}\rangle\},$$

for $\mathbf{t} = (t_1, ..., t_d) \in \mathbb{R}^d$, where $\langle \mathbf{x}, \mathbf{t}\rangle = x_1t_1 + ... + x_dt_d$ is the inner product between $\mathbf{x} = (x_1, ..., x_d) \in \mathbb{R}^d$ and $\mathbf{t}$.

A random vector $\mathbf{X}$ has *stable distribution* if for every pair $(\mathbf{X}', \mathbf{X}'')$ of independent random vectors identically distributed as $\mathbf{X}$, for every pair (a, b) of positive numbers, there always exist a positive number c and a vector $\mathbf{d} \in \mathbb{R}^d$ such that $a\mathbf{X}' + b\mathbf{X}''$ has the same distribution as $c\mathbf{X} + \mathbf{d}$. It is shown that the constant c is uniquely determined by the pair (a, b). Namely, there is a number $\alpha \in (0; 2]$ called stability index that satisfies $a^\alpha + b^\alpha = c^\alpha$. Then $\mathbf{X}$ is said to be α-*stable*. It is well known (Theorem 2.3.1[16]) that $\mathbf{X}$ is determined by a spectral measure Λ (a finite Borel measure on the unit sphere $\mathbb{S}_d$ in $\mathbb{R}^d$) and a shift vector $\boldsymbol{\delta} = (\delta_1, ..., \delta_d) \in \mathbb{R}^d$ through the representation

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp(-\int_{\mathbb{S}_d} \psi_\alpha(\langle \mathbf{s}, \mathbf{t}\rangle)\Lambda(d\mathbf{s}) + i\langle \boldsymbol{\delta}, \mathbf{t}\rangle)\,, \tag{1.1}$$

where

$$\psi_\alpha(u) = \begin{cases} |u|^\alpha(1 - i\mathrm{sign}(u)\tan\frac{\pi\alpha}{2}) & \alpha \neq 1 \\ |u|(1 + i\frac{2}{\pi}\mathrm{sign}(u)\ln|u|) & \alpha = 1. \end{cases}$$

We denote $\mathbf{X} \sim S(\alpha; \Lambda; \boldsymbol{\delta})$ to mark (1.1) is valid.

Especially, every α-stable random variable has a characteristic function of the form

$$\varphi_X(u) = \mathbb{E}\exp(iuX) = \begin{cases} \exp(-\gamma^\alpha|u|^\alpha[1 - i\beta(\tan\frac{\pi\alpha}{2})\mathrm{sign}(u)] + i\delta u) & \alpha \neq 1 \\ \exp(-\gamma|u|[1 + i\beta\frac{2}{\pi}\mathrm{sign}(u)\ln|u|] + i\delta u) & \alpha = 1, \end{cases}$$

with fixed $\beta \in [-1; 1], \gamma > 0$ and $\delta \in \mathbb{R}$. Then the parameters α, β, γ, and δ uniquely determine the distribution of X, we write $X \sim S(\alpha; \beta; \gamma; \delta)$. Usually, α is called the *stable index*, meanwhile β, γ and δ are named as the *skewness*, the *scale* and the *location* parameters of X, respectively.

2 Stable random vector and sub-Gaussian random vector

Below we investigate a method to create a transformation in multidimensional space, which turns a given stable random vector into a sub-Gaussian random vector. For fixed $\alpha \in (0, 2)$ let $A \sim S(\alpha/2; 1; (\cos\frac{\pi\alpha}{4})^{2/\alpha}; 0)$ be a positive $\alpha/2$ - stable random variable and $\mathbf{G} = (G_1, ..., G_d)$ be a zero-mean Gaussian vector independent of A. Then the random vector

$$\mathbf{X} = (A^{1/2}G_1, ..., A^{1/2}G_d)$$

is called a *sub-Gaussian random vector.* By virtue of Theorems 1.3.1 and 2.1.5 [16], $\mathbf{X}$ is an α-stable random vector.

It is clear that marginals of every sub-Gaussian random vector are symmetric. Therefore, stable random vectors with symmetric marginals are worthy to be considered. The copula tool can be used for that. For a given cdf $G : \mathbb{R} \to [0;1]$ let $G^{\leftarrow}(y) = \inf\{x : G(x) \geq y\}$ be its generalized inverse. The *copula* of r.v. $\mathbf{X}$, denoted by $C_{\mathbf{X}}$, can be defined by

$$C_{\mathbf{X}}(t_1, ..., t_d) = F_{\mathbf{X}}(F^{\leftarrow}_{X_1}(t_1), ..., F^{\leftarrow}_{X_d}(t_d)) \,,$$

for $0 \leq t_1, ..., t_d \leq 1$. Then we have (see also Sklar's Theorem [17])

$$F_{\mathbf{X}}(x_1, ..., x_d) = C_{\mathbf{X}}(F_{X_1}(x_1), ..., F_{X_d}(x_d)) \,, \tag{2.1}$$

for $x_1, ..., x_d \in \bar{\mathbb{R}} = [-\infty; +\infty]$. Moreover, when the random vector $\mathbf{X}$ is continuous, its pdf $f_{\mathbf{X}}$ and marginal pdf's $f_{X_1}, ..., f_{X_d}$ exist, simultaneously $F^{\leftarrow}_{X_k} = F^{-1}_{X_k}$ for $k = 1, ..., d$. From (2.1) it can be easy to prove that if C is a copula of any stable random vector and $X_1, ..., X_d$ are stable random variables then $F(x_1, ..., x_d) = C(F_{X_1}(x_1), ..., F_{X_d}(x_d))$ is a cdf of a stable random vector.

It is evident that the stability of a random vector is invariant under all linear rotations around the origin $\mathbf{0}$ of $\mathbb{R}^d$. The next lemma confirms that the statement remains true for symmetric stable random vectors and some "non-linear rotations" around the origin $\mathbf{0}$.

Lemma 2.1. *Let $\alpha \in (0;2)$ and an α-stable random vector $\mathbf{Z} = (Z_1, ..., Z_d)$ be symmetric, $(R, \Theta_1, ..., \Theta_{d-1}) = B(\mathbf{Z})$. Suppose that there exists an invertible differentiable transformation $Q : \mathbb{I}^{d-1} \to \mathbb{I}^{d-1}$ such that the random vector $\mathbf{Y} = D(R, Q(\Theta_1, ..., \Theta_{d-1}))$ is symmetric. Then $\mathbf{Y}$ has also an α-stable distribution.*

Proof. From the symmetry of $\mathbf{Z}$ and $\mathbf{Y}$ we see their characteristic function $\varphi_{\mathbf{Z}}$ and $\varphi_{\mathbf{Y}}$ taking only real values and

$$\varphi_{\mathbf{Z}}(\mathbf{t}) = \mathbb{E}\cos\langle \mathbf{Z}, \mathbf{t}\rangle \quad , \quad \varphi_{\mathbf{Y}}(\mathbf{t}) = \mathbb{E}\cos\langle \mathbf{Y}, \mathbf{t}\rangle \,, \tag{2.2}$$

for $\mathbf{t} = (t_1, ..., t_d) \in \mathbb{R}^d$, where $\langle \mathbf{z}, \mathbf{t}\rangle = z_1 t_1 + ... + z_d t_d$ for $\mathbf{z} = (z_1, ..., z_d)$. Meantime, the spectral representation of symmetric stable multivariate distributions (see e.g. Theorem 2.4.3 [16]) confirms that

$$\varphi_{\mathbf{Z}}(\mathbf{t}) = \exp(-\int_{\mathbb{S}_d} |\langle \mathbf{s}, \mathbf{t}\rangle|^{\alpha} \Lambda(d\mathbf{s})) \tag{2.3}$$

with a spectral measure Λ on $\mathbb{S}_d$. Looking at the polar representations

$$\begin{aligned}
z_1 &= r\cos\theta_1 \,, & t_1 &= u\cos\eta_1 \,, \\
z_2 &= r\sin\theta_1\cos\theta_2 \,, & t_2 &= u\sin\eta_1\cos\eta_2 \,, \\
&\;\;\cdots & &\;\;\cdots \\
z_{d-1} &= r\sin\theta_1 ... \sin\theta_{d-2}\cos\theta_{d-1} \,, & t_{d-1} &= u\sin\eta_1 ... \sin\eta_{d-2}\cos\eta_{d-1} \,, \\
z_d &= r\sin\theta_1 ... \sin\theta_{d-2}\sin\theta_{d-1} \,, & t_d &= u\sin\eta_1 ... \sin\eta_{d-2}\sin\eta_{d-1} \,,
\end{aligned}$$

where $0 \leq \theta_1 < \pi$, ..., $0 \leq \theta_{d-2} < \pi$, $0 \leq \theta_{d-1} < 2\pi$; $0 \leq \eta_1 < \pi$, ..., $0 \leq \eta_{d-2} < \pi$, $0 \leq \eta_{d-1} < 2\pi$ and $r = ||\mathbf{z}||$; $u = ||\mathbf{t}||$, we see

$$\langle \mathbf{z}, \mathbf{t} \rangle = ru \cdot v(\boldsymbol{\theta}, \boldsymbol{\eta}) \tag{2.4}$$

and the function v depends only of the arguments $\boldsymbol{\theta} = (\theta_1, ..., \theta_{d-1})$, $\boldsymbol{\eta} = (\eta_1, ..., \eta_{d-1})$ from $\mathbb{I}^{d-1}$. The existence of pdf $f_{\mathbf{Z}}$ of the stable random vector $\mathbf{Z}$ and the assumption that Q is differentiable imply the pdf $f_{\mathbf{Y}}$ of the random vector $\mathbf{Y}$ exists, that yields

$$f^*_{\mathbf{Y}}(r, \theta_1, ..., \theta_{d-1}) = f^*_{\mathbf{Z}}(r, Q(\boldsymbol{\theta}))J_Q\,,$$

where f^* means the polar coordinates form of a multivariate function f and J_Q denotes the Jacobian of Q, that is dependent only on the arguments $\theta_1, ..., \theta_{d-1}$. Therefore (2.2) can be rewritten as

$$\varphi_{\mathbf{Z}}(\mathbf{t}) = \int_{\mathbb{I}^{d-1}} \int_0^{\infty} \cos(ru \cdot v(\boldsymbol{\theta}, \boldsymbol{\eta}))f^*_{\mathbf{Z}}(r, \boldsymbol{\theta})drd\boldsymbol{\theta} \tag{2.5}$$

and

$$\varphi_{\mathbf{Y}}(\mathbf{t}) = \int_{\mathbb{I}^{d-1}} \int_0^{\infty} \cos(ru \cdot v(Q(\boldsymbol{\theta}), \boldsymbol{\eta}))f^*_{\mathbf{Z}}(r, Q(\boldsymbol{\theta}))J_Q drd\boldsymbol{\theta}\,. \tag{2.6}$$

Simultaneously, (2.3) implies

$$\varphi_{\mathbf{Z}}(\mathbf{t}) = \exp(-\int_{\mathbb{I}^{d-1}} |u \cdot v(\boldsymbol{\theta}, \boldsymbol{\eta})|^{\alpha} \Lambda_*(d\boldsymbol{\theta}))\,,$$

where Λ_* denotes the polar representation form of Λ. Combining the above equality with (2.4), (2.5), and (2.6), it is easily pointed out that

$$\varphi_{\mathbf{Y}}(\mathbf{t}) = \exp(-\int_{\mathbb{I}^{d-1}} |u \cdot v(\boldsymbol{\theta}, \boldsymbol{\eta})|^{\alpha} \Lambda^1_*(d\boldsymbol{\theta})) = \exp(-\int_{\mathbb{S}_d} |\langle \mathbf{s}, \mathbf{t} \rangle|^{\alpha} \Lambda^1(d\mathbf{s}))$$

for some spectral measure Λ^1 defined on $\mathbb{S}_d$. Consequently, the random vector $\mathbf{Y}$ is α-stable by virtue of Theorem 2.4.3 [16]. $\square$

In the following theorem we build up invertible transformations that turn random vectors with stable distribution into random vectors with sub-Gaussian stable distribution.

Theorem 2.2. *Let $\mathbf{X}$ be a random vector from a stable distribution in $\mathbb{R}^d$ such that $f_{\mathbf{X}}$ is positive in whole $\mathbb{R}^d$. Then there exists an invertible differentiable transformation $K : \mathbb{R}^d \to \mathbb{R}^d$ such that the random vector*

$$\mathbf{Y} = K(\mathbf{X})\,,$$

is a sub-Gaussian random vector.

Proof. Let $\mathbf{X} = (X_1, ..., X_n)$ be a stable random vector of marginals $X_k \sim S(\alpha;\beta_k;\gamma_k;\delta_k)$, $0 < \alpha < 2$; $-1 < \beta_k < 1$; $0 < \gamma_k < \infty$; $-\infty < \delta_k < \infty$, $k = 1, ..., d$. Let $X_0 \sim S(\alpha;0;1;0)$ denote a standard stable random variable. From the well known-fact (Property 1.2.14 [16]) that if a stable random variable has skewness parameter different from ± 1 then its cdf is positive on whole $\mathbb{R}$, the pdf's $f_{X_1}, ..., f_{X_d}$ and f_{X_0} are positive on whole $\mathbb{R}$, the cdf's $F_{X_1}, ..., F_{X_d}$ and F_{X_0} are strictly increasing. Then the functions $T_k : \mathbb{R} \to \mathbb{R}$, $k = 1, ..., d$, defined by $T_k(u) = F_{X_0}^{-1}(F_{X_k}(u))$, are strictly increasing functions. Besides, $T_k'(u) = f_{X_k}(u)/f_{X_0}(T_k(u))$, $k = 1, ..., d$, are positive functions. This implies $f_{X_0}(T_k(u))T_k'(u)du = f_{X_k}(u)du$, that yields

$$F_{X_0}(T_k(u)) = \int_{-\infty}^{T_k(u)} f_{X_0}(T_k(u))T_k'(u)du = \int_{-\infty}^{u} f_{X_k}(u)du = F_{X_k}(u)\,. \tag{2.7}$$

On the other hand, for every $t \in \mathbb{R}$ we have

$$F_{T_k \circ X_k}(t) = \mathbb{P}\{\omega : T_k(X_k(\omega)) \le t\} = \mathbb{P}\{\omega : X_k(\omega) \le T_k^{-1}(t)\} = F_{X_k}(T_k^{-1}(t))\,.$$

Compared the above with (2.7), putting $t = T_k(u)$, we get $F_{X_0}(t) = F_{T_k \circ X_k}(t)$. This confirms the two random variables X_0 and $T_k \circ X_k$ have the same distribution. Then, by virtue of Proposition 5.6 [2] and the remark after (2.1), the random vector $\mathbf{Z} = (T_1(X_1), ..., T_d(X_d))$ is stable with symmetric $S(\alpha;0;1;0)$-distributed marginals.

We attempt to design an invertible transformation $U : \mathbb{R}^d \to \mathbb{R}^d$ such that the random vector $U(\mathbf{Z})$ has isotropic distribution, which means $U(\mathbf{Z}) \stackrel{d}{=} MU(\mathbf{Z})$ for all linear rotations M around the origin $\mathbf{0}$. To construct the desired transformation U, we use several times the couple of the polar representation mappings $B = (B_0, B_1, ..., B_{d-1})$ and $D = (D_1, ..., D_d)$ defined by (2.8) and (2.9) as the follows. For $\mathbf{x} = (x_1, ..., x_d) \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, let

$$\begin{aligned} r \quad &= B_0(\mathbf{x}) \quad = ||\mathbf{x}|| = \sqrt{x_1^2 + \cdots + x_d^2}\,, \\ \theta_i \quad &= B_i(\mathbf{x}) \quad = \operatorname{arccot}\left(x_i/\sqrt{x_{i+1}^2 + \cdots + x_d^2} \right), \;\; i = 1, 2, ..., d-2 \\ \theta_{d-1} &= B_{d-1}(\mathbf{x}) = 2\operatorname{arccot}\left(\left[x_{d-1} + \sqrt{x_{d-1}^2 + x_d^2} \right] / x_d \right). \end{aligned} \tag{2.8}$$

Then $B = (B_0, B_1, ..., B_{d-1}) : \mathbb{R}^d \setminus \{\mathbf{0}\} \to \mathbb{R}^+ \times \mathbb{I}^{d-1}$ is a bijective mapping,

$$\mathbb{I}^{d-1} := \underbrace{[0,\pi) \times \cdots \times [0,\pi)}_{d-2} \times [0, 2\pi) \subset \mathbb{R}^{d-1}\,.$$

Inversely, each point $\mathbf{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ is of the form

$$\begin{aligned}
&x_1 = D_1(r,\theta_1,\theta_2,...,\theta_{d-1}) = r\cos\theta_1\,,\\
&x_2 = D_2(r,\theta_1,\theta_2,...,\theta_{d-1}) = r\sin\theta_1\cos\theta_2\,,\\
&\;\;\cdots\\
&x_{d-1} = D_{d-1}(r,\theta_1,\theta_2,...,\theta_{d-1}) = r\sin\theta_1\sin\theta_2...\sin\theta_{d-2}\cos\theta_{d-1}\,,\\
&x_d = D_d(r,\theta_1,\theta_2,...,\theta_{d-1}) = r\sin\theta_1\sin\theta_2...\sin\theta_{d-2}\sin\theta_{d-1}\,,
\end{aligned} \tag{2.9}$$

with $0 \le \theta_1 < \pi$, ..., $0 \le \theta_{d-2} < \pi$, $0 \le \theta_{d-1} < 2\pi$. Then the mapping $D = (D_1, ..., D_d) : \mathbb{R}^+ \times \mathbb{I}^{d-1} \to \mathbb{R}^d \setminus \{\mathbf{0}\}$ is the inverse transformation of B.

Let $\mathbf{e}_1 = (1,0,0,...,0)$, $\mathbf{e}_2 = (0,1,0,...,0)$, ..., $\mathbf{e}_d = (0,0,0,...,1)$ be the unit vectors in the basis of the Euclidean space $\mathbb{R}^d$. Let $\mathcal{P}_{(\mathbf{e}_i,\mathbf{e}_j)}$ denote the two-dimensional plan generated by $\{\mathbf{e}_i,\mathbf{e}_j\}$ for $1 \le i,j \le d-1$. It is clear that every shift (modulo 2π) of θ_{d-1} in the interval $[0;2\pi)$ corresponds to one rotation of $D(r,\theta_1,...,\theta_{d-1})$ around the origin in $\mathcal{P}_{(\mathbf{e}_{d-1},\mathbf{e}_d)}$.

Firstly, considering $(R,\Theta_1,...,\Theta_{d-1}) = B(\mathbf{Z})$, we see the random variable $F_{\Theta_{d-1}} \circ \Theta_{d-1}$ is uniformly distributed on $[0;1)$. Consequently, the random variable $2\pi F_{\Theta_{d-1}} \circ \Theta_{d-1}$ is uniformly distributed on $[0;2\pi)$, its distribution is invariant against every shift (modulo 2π) in the interval $[0;2\pi)$. Therefore, the distribution of the new defined random vector

$$U^{(1)}(\mathbf{Z}) = \mathbf{Z}^{(1)} = (Z_1^{(1)}, Z_2^{(1)}, ..., Z_d^{(1)}) := D(R,\Theta_1,...,\Theta_{d-2}, 2\pi F_{\Theta_{d-1}} \circ \Theta_{d-1})$$

is symmetric and invariant under all rotations around the origin in $\mathcal{P}_{(\mathbf{e}_{d-1},\mathbf{e}_d)}$.

In the second step, we change the coordinates of $(Z_1^{(1)}, Z_2^{(1)}, ..., Z_d^{(1)})$ by moving the first coordinate to the end and shifting the others ahead by one place (that means the basis $(\mathbf{e}_1,\mathbf{e}_2,...,\mathbf{e}_{d-1},\mathbf{e}_d)$ in $\mathbb{R}^d$ is replaced by the basis $(\mathbf{e}_2,\mathbf{e}_3,...,\mathbf{e}_d,\mathbf{e}_1)$). With the new random vector, we get

$$(R,\Theta_1^{(1)},...,\Theta_{d-2}^{(1)},\Theta_{d-1}^{(1)}) = B(Z_2^{(1)}, Z_3^{(1)}, ..., Z_d^{(1)}, Z_1^{(1)})\,.$$

The second transformation $U^{(2)}(\mathbf{Z}^{(1)}) = \mathbf{Z}^{(2)}$ is defined by

$$\mathbf{Z}^{(2)} = (Z_1^{(2)}, Z_2^{(2)}, ..., Z_d^{(2)}) := D(R,\Theta_1^{(1)},...,\Theta_{d-2}^{(1)}, 2\pi F_{\Theta_{d-1}^{(1)}} \circ \Theta_{d-1}^{(1)})\,.$$

Due to the fact that the random variable $2\pi F_{\Theta_{d-1}^{(1)}} \circ \Theta_{d-1}^{(1)}$ is uniformly distributed on $[0;2\pi)$, by a similar argument as the above, we can confirm that the distribution of the random vector $\mathbf{Z}^{(2)}$ is symmetric and invariant under all rotations around the origin in $\mathcal{P}_{(\mathbf{e}_d,\mathbf{e}_1)}$. Simultaneously, $\mathbf{Z}^{(2)}$ is also invariant under all rotations around the origin in $\mathcal{P}_{(\mathbf{e}_{d-1},\mathbf{e}_d)}$ because $\mathbf{Z}^{(1)}$ is invariant under all rotations around the origin in $\mathcal{P}_{(\mathbf{e}_{d-1},\mathbf{e}_d)}$.

Continuing the above process to the d-th step, the basis $(\mathbf{e}_{d-1},\mathbf{e}_d,...,\mathbf{e}_{d-3},\mathbf{e}_{d-2})$ in $\mathbb{R}^d$ is replaced by the basis $(\mathbf{e}_d,\mathbf{e}_1,...,\mathbf{e}_{d-2},\mathbf{e}_{d-1})$, the random vector obtained after the (d-1)-th step $(Z_1^{(d-2)}, Z_2^{(d-2)}, ..., Z_{d-1}^{(d-2)}, Z_d^{(d-2)})$ is rearranged into the new random vector $(Z_2^{(d-2)}, Z_3^{(d-2)}, ..., Z_d^{(d-2)}, Z_1^{(d-2)})$. The polar representation of the new random vector is

$$B(Z_2^{(d-2)}, ..., Z_d^{(d-2)}, Z_1^{(d-2)}) = (R, \Theta_1^{(d-2)}, ..., \Theta_{d-2}^{(d-2)}, \Theta_{d-1}^{(d-2)})\,.$$

Applying D to the polar representation random vector replaced the last coordinate $\Theta_{d-1}^{(d-2)}$ by $2\pi F_{\Theta_{d-1}^{(d-2)}} \circ \Theta_{d-1}^{(d-2)}$, we have $U^{(d-1)}(\mathbf{Z}^{(d-2)}) = \mathbf{Z}^{(d-1)}$, where

$$\mathbf{Z}^{(d-1)} = (Z_1^{(d-1)}, ..., Z_d^{(d-1)}) = D(R, \Theta_1^{(d-2)}, ..., \Theta_{d-2}^{(d-2)}, 2\pi F_{\Theta_{d-1}^{(d-2)}} \circ \Theta_{d-1}^{(d-2)})\,.$$

Then it is clear that the random vector $\mathbf{Z}^{(d-1)}$ has symmetric distribution. For the same reason as presented above, we can conclude the distribution of the random vector $\mathbf{Z}^{(d-1)}$ is invariant under every two-dimensional rotation around the origin in $\mathcal{P}_{(\mathbf{e}_{d-2}, \mathbf{e}_{d-1})}$. Consequently, the distribution is invariant under every two-dimensional rotation around the origin in each of the two-dimensional plans $\mathcal{P}_{(\mathbf{e}_{d-1}, \mathbf{e}_d)}, \mathcal{P}_{(\mathbf{e}_d, \mathbf{e}_1)}, ..., \mathcal{P}_{(\mathbf{e}_{d-2}, \mathbf{e}_{d-1})}$. This confirms the fact that the distribution of the random vector $\mathbf{Z}^{(d-1)}$ is invariant under all linear rotations around the origin $\mathbf{0}$ in the whole $\mathbb{R}^d$, which means the random vector has isotropic distribution.

Let denote by V the transformation of coordinates' rearrangement in $\mathbb{R}^d$ such that $V(x_d, x_1, ..., x_{d-2}, x_{d-1}) = (x_1, x_2, .., x_{d-1}, x_d)$ and define

$$U = V \circ U^{(d-1)} \circ U^{(d-2)} \circ ... \circ U^{(2)} \circ U^{(1)}.$$

Then we see all the transformations $U^{(1)}, U^{(2)}, ..., U^{(d-1)}$ are essentially based on the random variables defined in $\mathbb{I}^{d-1}$. That implies the existence of an invertible differentiable transformation $Q : \mathbb{I}^{d-1} \to \mathbb{I}^{d-1}$ such that $U = D \circ Q \circ B$. Besides, the above argument ensures that the random vector $\mathbf{Y} = U(\mathbf{Z}) = U(T(\mathbf{X}))$ is an isotropic random vector. This together with Lemma 2.1 show $\mathbf{Y}$ is a sub-Gaussian random vector. Therefore, we can confirm $K = U \circ T$ is the desired transformation. □

3 Application to test on the stable distribution of multivariate data

This section presents a procedure of combining the result of Theorem 2.1 and the Cramér - von Mises goodness-of-fit test based on the Kendall functions [3, 4, 7] to check the stable distribution of multivariate data. The transformation $K : \mathbb{R}^d \to \mathbb{R}^d$ determined by Theorem 2.1 turns a data set extracted from a multivariate stable distribution with positive density function into a new data set that has sub-Gaussian multivariate distribution. Then, a suitably tailored non-parametric test of Cramér - von Mises type can be used to check the goodness-of-fit of the transformed data set to sub-Gaussian multivariate distribution, and indirectly to check the stability of the original distribution.

Let $\mathbf{x} = \{x_{ij} : i = 1, 2, ..., d; j = 1, 2, ..., n\}$ be a sample dataset collected from a random vector $\mathbf{X} = (X_1, ..., X_d)$. By using Theorem 2.1, a procedure based on Cramér - von Mises type of tests based on the Kendall functions to test the hypothesis of stable distribution of the random vector can be made by the following steps:

Step 1. For $i = 1, 2, ..., d$, estimate the stable parameters $(\alpha_i; \beta_i; \gamma_i; \delta_i)$ of the data marginal $\mathbf{x}_i := \{x_{ij} : j = 1, 2, ..., n\}$. The hypothesis test aims to check if the random vector $\mathbf{X}$ is $\bar{\alpha}$-stable with

$$\bar{\alpha} = \frac{(\alpha_1 + \alpha_2 + ... + \alpha_d)}{d}.$$

Step 2. Use the operator K defined by Theorem 2.1 to get

$$\mathbf{y} = \{y_{ij} : i = 1, 2, ..., d; j = 1, 2, ..., n\},$$

where

$$\mathbf{y} = K(\mathbf{x}) = U(T(\mathbf{x})).$$

Step 3. Generate a hypothetical dataset

$$\mathbf{y}^{(0)} = \{y_{ij}^{(0)} : i = 1, 2, ..., d; j = 1, 2, ..., n\}$$

that is the dataset extracted from the random vector with sub-Gaussian $\bar{\alpha}$-stable distribution $\mathbf{Y}^* = A^{1/2}\mathbf{G}$ (see Subsection 2.2 [13]), where $\mathbf{G} \sim N(\mathbf{0}; \mathbf{I})$ is a Gaussian random vector with expectation $\mathbf{0}$ and covariance matrix $\mathbf{I}$, the unit matrix of the size $d \times d$, $A \sim S(\bar{\alpha}/2; 1; (\cos(\pi\bar{\alpha}/4))^{2/\bar{\alpha}}; 0)$, A and $\mathbf{G}$ are independent. We denote $\mathbf{Y}^* = (Y_1^*, Y_2^*, ..., Y_d^*)$.

Step 4. For $k = 1, 2, ..., n$, let

$$M_k = \frac{\#\{j \neq k : y_{1j}^{(0)} < y_{1k}^{(0)}, y_{2j}^{(0)} < y_{2k}^{(0)}, ..., y_{dj}^{(0)} < y_{dk}^{(0)}\}}{n}.$$

Based on the formula of the Kendall functions [12] of the random vector

$$K_{\mathbf{Y}^*}(t) = \mathbb{P}(F_{\mathbf{Y}^*}(Y_1^*, Y_2^*, ..., Y_d^*)) \leq t),$$

to estimate the values of the Kendall function at M_k for the hypothetical dataset and the transformed sample dataset,

$$K_{\mathbf{y}^{(0)}}(M_k) = \frac{\#\{j : \bar{F}_{\mathbf{y}^{(0)}}(y_{1j}^{(0)}, y_{2j}^{(0)}, ..., y_{dj}^{(0)}) \leq M_k\}}{n},$$

$$K_{\mathbf{y}}(M_k) = \frac{\#\{j : \bar{F}_{\mathbf{y}}(y_{1j}, y_{2j}, ..., y_{dj}) \leq M_k\}}{n},$$

where $\bar{F}$ denotes the empirical distribution function. Use the above values of the Kendall function to determine the test statistic

$$D = \sum_{k=1}^{n} (K_{\mathbf{y}^{(0)}}(M_k) - K_{\mathbf{y}}(M_k))^2 .$$

Step 5. Apply the Monte-Carlo sampling procedure by repeating Step 4 to construct 1000 hypothetical datasets

$$\mathbf{y}^{(m)} = \{y_{ij}^{(0)} : i = 1, 2, ..., d; j = 1, 2, ..., n\}, \ m = 1, 2, ..., 1000,$$

then repeatedly conduct Step 4 with the transformed sample dataset $\mathbf{y}$ replaced by each of the hypothetical datasets to get the values

$$d_m = \sum_{k=1}^{n} (K_{\mathbf{y}^{(0)}}(M_k) - K_{\mathbf{y}^{(m)}}(M_k))^2 .$$

With a given probability value $p \in (0; 1)$, let q_{1-p} be the $(1-p)$-quantile of the set $\{d_m, m = 1, 2, ..., 1000\}$ (usually p is taken equal to 0.1, 0.05, 0.01, 0.005, or 0.001). Set $L_p = q_{1-p}$ as the critical value of test. Compare the test statistic D to the critical value L_p. Reject the hypothesis H if $D \geq L_p$ and accept the hypothesis if $D < L_p$. The test procedure completes.

In the following we present two examples of the application of the above goodness-of-fit test procedure to examine the multivariate stable distribution of the daily return data of Nasdaq Finance. The daily return data from the 10 stocks FA (Facebook); AMC (AMC Entertainment Holdings); AXP (American Express); NFLX (Netflix); ZM (Zoom); JNJ (Johnson & Johnson); XOM (Exxon Mobil Corporation Common); FB (Meta Platforms); HD (Home Depot); and PPG (PPG Industries), contain a sample from 22/4/2019 to 31/12/2020 to imply 430 observations. Continuously compounded percentage returns are considered, i.e. daily returns are measured by the log-differences of closing pricing multiplied by 100.

Example 1. Concerning the 5-dimensional data of the stocks FA; AMC; AXP; NFLX; and ZM, the functions *McCullochParametersEstim* and *ks.test* in the R software package are used to estimate the stable parameters of each coordinate, and then to verify the hypothesis of stable distribution goodness-of-fit, with $\bar{\alpha}$ (the average of the five values of α) taken as the common shape parameter for all marginals. The results are presented in Table 1.

Table 1. Stable parameters of the marginals NFLX, ZM, AMC, AXP, and FA

Coordinate	α	$\bar{\alpha}$	β	γ	δ	p-value
X_1 (NFLX)	1.771	1.478	-0.149	1.550	0.006	0.5699
X_2 (ZM)	1.461	1.478	0.000	2.090	-0.169	0.3648
X_3 (AMC)	1.392	1.478	0.046	2.719	-0.264	0.6843
X_4 (AXP)	1.388	1.478	-0.152	1.128	0.029	0.6843
X_5 (FA)	1.378	1.478	0.134	1.229	1.229	0.5699

In Table 1, all the p-values are greater than 0.05, indicate all coordinate data fit to the stable distributions $S(\bar{\alpha}; \beta; \gamma; \delta)$ with the corresponding parameters. Simultaneously, all the skewness parameters β are different from ± 1. Then we conduct Step 2 followed by Step 3, Step 4 and Step 5 to get the test statistic $D = 26.4866$ and the test critical values represented in Table 2.

Table 2. Test critical values on sub-Gaussian 1.478-stable distribution in $\mathbb{R}^5$

Significance p	0.1	0.05	0.01	0.005	0.001
Critical value L_p	87.6010	126.6208	188.3399	236.3742	302.8049

Because the test statistic $D = 26.4866$ is smaller than the critical values in Table 2, the hypothesis H is accepted and we can conclude that the returns' dataset of NFLX, ZM, AMC, AXP, and FA fits to the 5-dimensional $\bar{\alpha}$-stable distribution.

Example 2. Regarding the 5-dimensional data of the stocks JNJ; XOM; FB; HD; and PPG, we proceed with the same procedure as that of Example 1. The functions *McCullochParametersEstim* and *ks.test* in the R software package are used in the first step to estimate the stable parameters and to check the stability of marginal distributions, giving the results in Table 3.

Table 3. Stable parameters of the marginals JNJ, XOM, FB, HD, and PPG

Coordinate	α	$\bar{\alpha}$	β	γ	δ	p-value
X_1 (JNJ)	1.385	1.4444	-0.313	0.642	0.098	0.2462
X_2 (XOM)	1.329	1.4444	0.061	1.146	-0.129	0.1847
X_3 (FB)	1.580	1.4444	-0.264	1.261	0.127	0.2821
X_4 (HD)	1.465	1.4444	-0.150	0.808	0.110	0.6843
X_5 (PPG)	1.463	1.4444	-0.003	0.953	0.083	0.8900

All the p-values In Table 3 are greater than 0.05, confirming the $\bar{\alpha}$-stability of all marginal distributions of the concerned data, where $\bar{\alpha} = 1.4444$. Then

Step 2, Step 3, Step 4 and Step 5 of the proposed procedure are sequentially proceeded to provide the test statistic $D = 342.6804$ and Table 2 containing the test critical values of the test on sub-Gaussian distribution in $\mathbb{R}^5$.

Table 4. Test critical values on sub-Gaussian 1.444-stable distribution in $\mathbb{R}^5$

Significance p	0.1	0.05	0.01	0.005	0.001
Critical value L_p	115.2128	131.7771	168.7637	173.2055	186.3036

In this example, the test statistic $D = 342.6804$ is greater than all critical values in Table 3, therefore we can realize that the returns' dataset of JNJ, XOM, FB, HD, and PPG does not fit to the 1.4444-stable 5-dimensional distribution.

References

[1] R. J. Adler, R. E. Feldman and M. S. Taqqu, *A Practical Guide to Heavy Tailed Data*, Birkhäuser, Boston, 1998.

[2] P. Embrechts, F. Lindskog and A. McNeil, *Modeling Dependence with Copulas and Applications to Risk Management*, Handbook of Heavy Tailed Distributions in Finance, 2003, ed. Rachev S., Elsevier, Chapter 8, 2001, 329–384. https://doi.org/10.1016/b978-044450896-6.50010-8

[3] C. Genest, J.-F. Quessy and B. Rémillard, Goodness-of-fit procedures for copula models based on the probability integral transformation, *Scand. J. Statist.*, **33** (2006), 337–366.
https://doi.org/10.1111/j.1467-9469.2006.00470.x

[4] C. Genest and L.-P. Rivest, Statistical inference procedures for bivariate Archimedean copulas, *J. Amer. Statist. Assoc.*, **88** (1993), 1034–1043.
https://doi.org/10.1080/01621459.1993.10476372

[5] S. M. Kogon and D. B. Williams,*Characteristic function based estimation of stable parameters*, in Adler R., Feldman R. and Taqqu M. (eds.) A Practical Guide to Heavy Tailed Data, Birkhäuser, Boston, MA 1998, 311–335.

[6] F. Lamantia, S. Ortobelli and S. Rachev, VaR, CVaR and Time Rules with Elliptical and Asymmetric Stable Distributed Returns, *Investment Management and Financial Innovations*, **3** (2006), no. 4, 19–39.

[7] S. Majoros and A. Zempléni, Multivariate stable distributions and their application in modeling returns, arXiv:1810.09521 (2018) [stat.AP].

[8] J. H. McCulloch, Simple consistent estimators of stable distribution parameters, *Commun. Statist. Simul.*, **15** (1986), 1109–1136. https://doi.org/10.1080/03610918608812563

[9] J. H. McCulloch, *Financial Applications of Stable Distributions*, Handbook of Statistics 14, ed. Maddala G. and Rao C., Elsevier Science Publishers, North-Holland (1996), 393–425.

[10] J. P. Nolan, Multivariate elliptically contoured stable distributions: theory and estimation, *Computational Statistics*, **28** (2013), no. 5, 2067–2089. https://doi.org/10.1007/s00180-013-0396-7

[11] J. P. Nolan, *Stable Distributions - Models for Heavy Tailed Data*, Birkhauser, Boston, MA, USA, 2016.

[12] R. B. Nelsena, J. J. Quesada-Molina, J. A. Rodriguez-Lallena and M. Ubeda-Flores, Kendall distribution functions, *Statistics & Probability Letters*, **65** (2003), no. 3, 263–268. https://doi.org/10.1016/j.spl.2003.08.002

[13] V. Omelchenko, Parameter estimation of sub-Gaussian stable distributions, *Kybernetika*, **50** (2014), no. 6, 929–949. https://doi.org/10.14736/kyb-2014-6-0929

[14] K. J. Palmer, M. S. Ridout and B. J. T. Morgan, Modeling cell generation times using the tempered stable distribution, *Journal of the Royal Statistical Society*, **57** (2008), 379–397. https://doi.org/10.1111/j.1467-9876.2008.00625.x

[15] S. I. Resnik, *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer, New York, 2007.

[16] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes*, New York, NY, Chapman & Hall, 1994.

[17] A. Sklar, Fonctions de rèpartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Universitè de Paris*, **8** (1959), 229–231.

Received: April 19, 2022; Published: May 21, 2022