Shahid Bahonar University of Kerman
Journal of Mahani Mathematical Research
Print ISSN: 2251-7952 Online ISSN: 2645-4505
JMMR 2012

ANALYZING SKEWED FINANCIAL DATA USING SKEW SCALE-SHAP MIXTURES OF MULTIVARIATE NORMAL DISTRIBUTIONS

M. Tamandi ✉ and M. Amiri

Special issue Dedicated to memory of professor Mahbanoo Tata
Article type: Research Article
(Received: 29 December 2023, Received in revised form 27 March 2024)
(Accepted: 14 June 2024, Published Online: 15 June 2024)

Abstract. This paper introduces an innovative family of statistical models called the multivariate skew scale-shape mixtures of normal distributions. These models serve as a versatile tool in statistical analysis by efficiently characterizing the skewed and leptokurtic nature commonly observed in multivariate datasets. Their applicability shines in real-world scenarios where data often deviate from standard statistical assumptions due to the presence of outliers. We present an EM-type algorithm designed for maximizing likelihood estimation and evaluate the model's effectiveness through real-world data applications. Through rigorous testing against various datasets, we assess the performance and practicality of the proposed algorithm in real statistical scenarios. The results demonstrate the remarkable performance of this new family of distributions.

Keywords: Shape mixtures, Scale Mixtures, EM-type algorithm, Multivariate distributions, Stock Markets
2020 MSC: Primary 62E15.

1. Introduction

Multivariate normal distributions are foundational models in statistics, portraying the collective behavior of correlated variables (Härdle and Simar, 2015). Each variable adheres to a normal (Gaussian) distribution, and their joint distribution characterizes the multivariate structure (Kotz, Balakrishnan, and Johnson, 2004). Defined by a mean vector and covariance matrix, these distributions encapsulate variable averages and relationships (Mardia, Kent, and Bibby, 1979). Conversely, multivariate skew normal distribution offer a nuanced extension, accommodating asymmetry in data (Azzalini and Capitanio, 2014). This distribution introduces skewness, allowing a flexible representation of non-symmetrical datasets (Mondal, 2023). Multivariate skew normal distribution includes skewness matrix parameters alongside means and covariance matrices, enabling modeling of asymmetrical tendencies (Arellano-Valle and Genton, 2005). This adaptability goes beyond symmetric distributions,

✉ tamandi@vru.ac.ir, ORCID: 0000-0003-0056-4642
https://doi.org/10.22103/jmmr.2024.22735.1562

Publisher: Shahid Bahonar University of Kerman
How to cite: M. Tamandi, M. Amiri, *Analyzing skewed financial data using skew scale-shap mixtures of multivariate normal distributions*, J. Mahani Math. Res. 2024; 13(3): 71 - 90.

providing richer depictions of diverse multivariate data structures. In recent years, several authors have introduced various generalizations of multivariate normal and skew-normal distributions, delving into their properties. For additional literature on related distributions, readers are directed to studies such as scale-shape mixtures of skew-normal distributions (Jamalizadeh and Lin, 2017), the multivariate flexible skew-symmetric-normal distribution (Mahdavi et al., 2021), scale-shape mixtures of multivariate skew-normal distributions and their stochastic ordering (Amiri and Balakrishnan, 2022), and generalized location-scale mixtures of elliptical distributions (Pu et al., 2023).

Multivariate distributions frequently use in financial and economic analyses, serving to describe the joint behavior of multiple variables. These distributions find extensive applications in portfolio theory (Fabozzi, 2008), asset pricing models (Fama and French, 1993), and risk management (Jorion, 2007). They provide a reliable framework for depicting the interrelationships among various financial assets, enabling efficient diversification strategies and risk assessments (Bodie and Kane 2020). Conversely, multivariate skew normal distributions have gained traction in financial modeling due to their ability to capture asymmetry in asset returns and economic variables (Harvey and Siddique, 2000). These distributions accommodate skewed and leptokurtic data, often observed in financial markets, making them valuable for modeling stock returns, volatility, and other financial variables (Cont, 2007). Their applications extend to risk modeling in insurance (Gómez-Déniz et al., 2022) and macroeconomic analyses involving skewed economic indicators (Denkowska and Wanat, 2020). The incorporation of skewness enhances the accuracy of financial models, offering a more comprehensive depiction of the complexities inherent in economic data structures.

This paper endeavors to introduce a novel skew scale-shape mixtures of the multivariate normal distribution, designed to effectively encapsulate the skewness and leptokurtosis inherent in financial datasets. Moreover, our study introduces a groundbreaking innovation in multivariate distribution modeling by incorporating a novel approach: treating the shape parameter of the skew-normal distribution as a random variable. This departure from conventional methodologies allows our model to dynamically adapt and incorporate varying levels of skewness and kurtosis, surpassing the capabilities of previous approaches. Building upon seminal works in this field (Azzalini and Capitanio, 2014), our methodology represents a paradigm shift, offering unparalleled flexibility and precision in capturing the asymmetrical patterns and heavy-tailed behavior inherent in financial datasets. Unlike traditional methods that impose fixed or predetermined shape parameters, our approach empowers the distribution to dynamically adjust, enabling it to more accurately reflect the complexities of real-world financial data. Moreover, The introduced family encompasses shape mixtures of multivariate skew-symmetric normal distributions which the symmetric part is a scale mixture of multivariate normal distribution like symmetric generalized hyperbolic, slash, contaminated normal and logistic distributions. It offers greater flexibility compared to these subclasses in fitting multivariate data, owing to its broader coverage of skewness and kurtosis. This fundamental innovation not only distinguishes our study from previous contributions (Mondal et al., 2023) but also positions it as a pioneering advancement in the field of multivariate distribution modeling. By pushing the boundaries of statistical methodology, our research opens up new horizons for advanced financial modeling, promising enhanced risk assessment, asset pricing, and portfolio optimization strategies. Furthermore, our empirical validation in the context of financial stock markets serves as compelling evidence of the efficacy and superiority of our methodology in capturing the nuances of real-world data. Through meticulous exploration of specific cases within our novel framework and a detailed exposition of the EM algorithm's application, we demonstrate the practical utility and methodological rigor of our approach. In essence, our paper represents a transformative contribution to the field, offering a robust and adaptable framework that is poised to revolutionize the landscape of financial modeling.

Section 2 delves into the exploration of this new family of distributions, analyzing specific cases within it. The EM algorithm, pivotal in this context, is detailed in Section 3. A simulation study and the practical application of these models in financial Stock Markets are investigated in Sections 4 and 5, respectively. Some conclusions are made in Section 6.

2. Theory and methods

Within this section, we introduce a formulation of skew scale mixtures of the multivariate normal distribution utilizing established guidelines. Initially, we outline foundational concepts derived from multivariate normal and skew normal distributions to lay the groundwork for our discussion.

2.1. **Preliminary framework.** A random vector $\mathbf{Y}$ is said to follow a multivariate skew normal distribution (MSN) with location vector $\boldsymbol{\xi} = (\xi_1, ..., \xi_p)^\top$, $p{\times}p$ scale covariance matrix $\boldsymbol{\Sigma}$ and skewness parameter vector $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_p)^\top$, denoted by $\mathbf{Y} \sim SN_p(\boldsymbol{\xi}, \boldsymbol{\Sigma}, \boldsymbol{\lambda})$, if it has density

$$f_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\xi}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}) = 2\phi_p(\mathbf{y}; \boldsymbol{\xi}, \boldsymbol{\Sigma})\Phi\left(\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1/2} (\boldsymbol{y} - \boldsymbol{\xi})\right), \tag{1}$$

where $\phi_p(\cdot; \boldsymbol{\xi}, \boldsymbol{\Sigma})$ is the probability density function (pdf) of $N_p(\boldsymbol{\xi}, \boldsymbol{\Sigma})$, and $\Phi(\cdot)$ is the cumulative distribution function (cdf) of univariate standard normal distribution.

According to Arellano-Valle et al. (2005), the MSN distribution has a convenient stochastic representation

$$\mathbf{Y} = \boldsymbol{\xi} + \boldsymbol{\Sigma}^{1/2}\left\{\boldsymbol{\delta}\,|Z_0| + \left(\mathbf{I}_p - \boldsymbol{\delta}\boldsymbol{\delta}^T\right)^{1/2} \mathbf{Z}_1\right\}, \tag{2}$$

where $\boldsymbol{\delta} = \boldsymbol{\lambda}/\sqrt{1+\boldsymbol{\lambda}^T\boldsymbol{\lambda}}$, $Z_0 \sim N(0,1)$ and $\mathbf{Z}_1 \sim N_p\left(\mathbf{0}, \mathbf{I}_p\right)$ independently.

The random vector $\mathbf{Y}$ follows a scale mixture of normal (SMN) distribution if $\mathbf{Y}|\{U=u\} \sim N_p(\boldsymbol{\xi}, u^{-1}\boldsymbol{\Sigma})$ where and U is a positive variable and has pdf $h(u;\boldsymbol{\nu})$, indexed by parameters $\boldsymbol{\nu}$.

Ferreira et al. (2016) introduced a family of skew scale mixture of normal (SSMN) distributions by multiplying the density of SMN distribution by the cdf of the normal distribution (as the skewing function). A random vector $\mathbf{Y}$ has SSMN distribution

and it is denoted by $\mathbf{Y} \sim SSMN(\boldsymbol{\xi}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}; h)$ when its pdf is

$$f(\mathbf{y}) = 2f_0(\mathbf{y})\Phi(\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{y}-\boldsymbol{\xi})), \tag{3}$$

where $f_0(\mathbf{y})$ is the pdf of a SMN distribution. Some familiar skewed distributions belong to this family, including multivariate skew-t-normal (MSTN), skew-slash (MSSL) and skew-contaminated-normal (MSCN). In this paper, we consider a more general class of skew scale-shape mixtures of normal (SKSSMN) distributions and present feasible EM-type algorithms for the computation of maximum likelihood (ML) estimates of parameters. A random variable $\mathbf{Y}$ is said to follow a SKSSMN distribution if it has the following representation

$$\begin{aligned} \mathbf{Y} \mid (U=u, \mathbf{V}=\mathbf{v}) \quad &\sim \quad SN_p\left(\boldsymbol{\xi}, u^{-1}\boldsymbol{\Sigma}, u^{-1/2}\mathbf{v}\right), \\ U \sim h(u;\boldsymbol{\nu}) \quad &\perp \quad \mathbf{V} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Omega}), \end{aligned} \tag{4}$$

where $\perp$ means U and $\mathbf{V}$ are independent. The model in (3) is a special case of the model in (4) when $\mathbf{V}$ is a fixed parameter as $\boldsymbol{\lambda}$. So the new model can be generalized all models discussed in Ferreira et al. (2016).

Let $\mathbf{Y}$ be a random vector following the representation in (4). Then, we say that $\mathbf{Y}$ follows a SKSSMN distribution and write $\mathbf{Y} \sim SKSSMN(\boldsymbol{\xi}, \boldsymbol{\Sigma}, \boldsymbol{\mu}, \boldsymbol{\Omega}; h)$ for short.

Proposition 2.1. *If $\boldsymbol{Y} \sim SKSSMN(\boldsymbol{\xi}, \boldsymbol{\Sigma}, \boldsymbol{\mu}, \boldsymbol{\Omega}; h)$ the pdf of $\boldsymbol{Y}$ is given by*

$$f(\mathbf{y}) = 2f_0(\mathbf{y})\Phi\left(\frac{\boldsymbol{\mu}^\top\mathbf{d}}{\sqrt{1+\mathbf{d}^\top\boldsymbol{\Omega}\mathbf{d}}}\right), \ \mathbf{y} \in \mathbb{R}^p, \tag{5}$$

where $\mathbf{d} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{y}-\boldsymbol{\xi})$ and $f_0(\mathbf{y}) = \int_0^\infty \phi_p\left(\mathbf{y}; \boldsymbol{\xi}, u^{-1}\boldsymbol{\Sigma}\right) h(u;\boldsymbol{\nu})du$.

Proof. The proof is straightforward using the hierarchical representation in (4) and some integration. □

Proposition 2.2. *If $\boldsymbol{Y} \sim SKSSMN(\boldsymbol{\xi}, \boldsymbol{\Sigma}, \boldsymbol{\mu}, \boldsymbol{\Omega}; h)$ a further stochastic representation for $\boldsymbol{Y}$ is*

$$\mathbf{Y} = \boldsymbol{\xi} + \boldsymbol{\Sigma}^{1/2}\left\{\frac{u^{-1/2}\mathbf{V}}{\sqrt{u+\mathbf{V}^\top\mathbf{V}}}|Z_0| + \left(u\mathbf{I}_p + \mathbf{V}\mathbf{V}^\top\right)^{-1/2}\mathbf{Z}_1\right\}. \tag{6}$$

Proof. The proposition is proved using representation (2) for conditional distribution in (4). □

Let $W = \sqrt{1 + u^{-1}\mathbf{V}^\top\mathbf{V}}$ be a reparameterized latent variable. A further hierarchical representation of the SKSSMN distribution can be written as

$$\mathbf{Y} \mid (W = w, \mathbf{V} = \mathbf{v}, U = u) \sim N_p\left(\boldsymbol{\xi} + \frac{w}{u + \mathbf{v}^\top\mathbf{v}}\boldsymbol{\Sigma}^{1/2}\mathbf{v}, \boldsymbol{\Sigma}^{1/2}\left(u\mathbf{I}_p + \mathbf{v}\mathbf{v}^\top\right)^{-1}\boldsymbol{\Sigma}^{1/2}\right),$$
$$W \mid (\mathbf{V} = \mathbf{v}, U = u) \sim TN\left(0, \frac{u + \mathbf{v}^\top\mathbf{v}}{u}; (0, \infty)\right),$$
$$\mathbf{V} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Omega}) \perp U \sim h(u; \boldsymbol{\nu}), \tag{7}$$

where $TN(\mu, \sigma^2; (a, b))$ represents the truncated normal distribution for $N(\mu, \sigma^2)$ lying within the truncated interval (a, b).

Using (7), the joint pdf of $\mathbf{Y}, \mathbf{V}, W$ and U is given by

$$f(\mathbf{y}, \mathbf{v}, w, u) = \frac{2(2\pi)^{-(2p+1)/2}u^{p/2}}{|\boldsymbol{\Omega}|^{1/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left\{-\frac{u}{2}\mathbf{d}^\top\mathbf{d} - \frac{\left(w - \mathbf{d}^\top\mathbf{v}\right)^2}{2}\right\}$$
$$\times \exp\left\{-\frac{1}{2}(\mathbf{v} - \boldsymbol{\mu})^\top\boldsymbol{\Omega}^{-1}(\mathbf{v} - \boldsymbol{\mu})\right\}h(u; \boldsymbol{\nu}), \tag{8}$$

where $\mathbf{d} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{y} - \boldsymbol{\xi})$ and $|\mathbf{A}|$ denotes the determinant of matrix $\mathbf{A}$.

Integrating out w in (8), we get

$$f(\mathbf{y}, \mathbf{v}, u) = \int_0^\infty f(\mathbf{y}, \mathbf{v}, w, u)dw = \frac{2(2\pi)^{-p}}{|\boldsymbol{\Omega}|^{1/2}|\boldsymbol{\Sigma}|^{1/2}}u^{p/2}e^{-\frac{u}{2}\mathbf{d}^\top\mathbf{d}}$$
$$\times \exp\left\{-\frac{1}{2}(\mathbf{v} - \boldsymbol{\mu})^\top\boldsymbol{\Omega}^{-1}(\mathbf{v} - \boldsymbol{\mu})\right\}h(u; \boldsymbol{\nu})\Phi\left(\mathbf{d}^\top\mathbf{v}\right). \tag{9}$$

So dividing (8) by (9) gives

$$f(w|\mathbf{y}, \mathbf{v}, u) = \frac{\phi\left(w - \mathbf{d}^\top\mathbf{v}\right)}{\Phi\left(\mathbf{d}^\top\mathbf{v}\right)} = f(w|\mathbf{y}, \mathbf{v}). \tag{10}$$

Implying that W and U are conditionally independent given $\mathbf{Y} = \mathbf{y}$ and $\mathbf{V} = \mathbf{v}$. It follows from (10) that the conditional distribution of W given $\mathbf{Y} = \mathbf{y}$ and $\mathbf{V} = \mathbf{v}$ is

$$W|(\mathbf{y}, \mathbf{v}) \sim TN\left(\mathbf{d}^\top\mathbf{v}, 1; (0, \infty)\right). \tag{11}$$

Integrating from (9) over $\mathbf{v}$ and u we have

$$f(\mathbf{y}) = \frac{2(2\pi)^{-p/2}}{|\boldsymbol{\Sigma}|^{1/2}}\Phi\left(\frac{\mathbf{d}^\top\boldsymbol{\mu}}{\sqrt{1 + \mathbf{d}^\top\boldsymbol{\Omega}\mathbf{d}}}\right)\int_0^\infty u^{p/2}e^{-\frac{u}{2}\mathbf{d}^\top\mathbf{d}}h(u; \boldsymbol{\nu})du$$
$$= 2f_0(\mathbf{y})\Phi\left(\frac{\mathbf{d}^\top\boldsymbol{\mu}}{\sqrt{1 + \mathbf{d}^\top\boldsymbol{\Omega}\mathbf{d}}}\right). \tag{12}$$

Observing (9), we have the following relation

$$f(\mathbf{v}, u \mid \mathbf{y}) = f(\mathbf{v} \mid \mathbf{y})f(u \mid \mathbf{y}). \tag{13}$$

Implying that $\mathbf{V}$ and U are conditionally independent given $\mathbf{Y} = \mathbf{y}$. From (12) and (13), we can get the following consequence

$$f(\mathbf{v} \mid \mathbf{y}) = \phi_p(\mathbf{v}; \boldsymbol{\mu}, \boldsymbol{\Omega}) \frac{\Phi\left(\mathbf{d}^\top \mathbf{v}\right)}{\Phi\left(\frac{\mathbf{d}^\top \boldsymbol{\mu}}{\sqrt{1+\mathbf{d}^\top \boldsymbol{\Omega} \mathbf{d}}}\right)}.$$

In particular, $\mathbf{V} \mid \{\mathbf{Y} = \mathbf{y}\}$ follows the extended SN (ESN) distribution (Azzalini 2014), denoted by $ESN_p\left(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\Omega}^{1/2}\mathbf{d}, \mathbf{d}^\top \boldsymbol{\mu}\right)$.

Using the properties of ESN distribution, we have

$$E(\mathbf{V}|\mathbf{y}) = \boldsymbol{\mu} + \frac{\boldsymbol{\Omega}\mathbf{d}}{\sqrt{1+\mathbf{d}^\top \boldsymbol{\Omega}\mathbf{d}}} R\left(\frac{\boldsymbol{\mu}^\top \mathbf{d}}{\sqrt{1+\mathbf{d}^\top \boldsymbol{\Omega}\mathbf{d}}}\right), \tag{14}$$

where $R(x) = \frac{\phi(x)}{\Phi(x)}$ and

$$\begin{aligned} E(\mathbf{V}\mathbf{V}^T|\mathbf{y}) &= \boldsymbol{\mu}\boldsymbol{\mu}^\top + \boldsymbol{\Omega} + \frac{1}{\sqrt{1+\mathbf{d}^\top \boldsymbol{\Omega}\boldsymbol{d}}} R\left(\frac{\boldsymbol{\mu}^\top \mathbf{d}}{\sqrt{1+\mathbf{d}^\top \boldsymbol{\Omega}\mathbf{d}}}\right) \\ &\times \left(\boldsymbol{\Omega}\boldsymbol{d}\boldsymbol{\mu}^\top + \boldsymbol{\mu}\boldsymbol{d}^\top\boldsymbol{\Omega} - \frac{\boldsymbol{\mu}^\top \mathbf{d}}{1+\mathbf{d}^\top \boldsymbol{\Omega}\boldsymbol{d}} \left(\boldsymbol{\Omega}\mathbf{d}\mathbf{d}^\top \boldsymbol{\Omega}\right)\right). \end{aligned} \tag{15}$$

From (11) we have

$$E(W|\mathbf{y}, \mathbf{v}) = \mathbf{d}^\top \mathbf{v} + R(\mathbf{d}^\top \mathbf{v}).$$

Thus

$$E(W|\mathbf{y}) = E(E(W|\mathbf{y}, \mathbf{v})) = \mathbf{d}^\top E(\mathbf{V} \mid \mathbf{y}) + E\left(R(\mathbf{d}^\top \mathbf{V}) \mid \mathbf{y}\right),$$

and

$$E(W\mathbf{V}|\mathbf{y}) = E(\mathbf{V}E(W|\mathbf{y}, \mathbf{v})|\mathbf{y}) = E(\mathbf{V}\mathbf{V}^T|\mathbf{y})\mathbf{d} + E\left(\mathbf{V}R(\mathbf{d}^\top \mathbf{V}) \mid \mathbf{y}\right), \tag{16}$$

but

$$\begin{aligned} E\left(\mathbf{V}R(\mathbf{d}^\top \mathbf{V}) \mid \mathbf{y}\right) &= \int_{\mathbb{R}^p} \mathbf{v}\phi(\mathbf{d}^\top \mathbf{v})\phi_p(\mathbf{v}; \boldsymbol{\mu}, \boldsymbol{\Omega}) d\mathbf{v} \\ &= \frac{|\boldsymbol{\Omega}|^{-1/2}}{\Phi\left(\frac{\boldsymbol{\mu}^T \mathbf{d}}{\sqrt{1+\mathbf{d}^\top \boldsymbol{\Omega}\boldsymbol{d}}}\right)} (2\pi)^{-\frac{p+1}{2}} e^{-\frac{\boldsymbol{\mu}^\top \boldsymbol{\Omega}^{-1}\boldsymbol{\mu}}{2}} \\ &\times \int_{\mathbb{R}^p} \mathbf{v} e^{\boldsymbol{\mu}^\top \boldsymbol{\Omega}^{-1}\mathbf{v}} e^{-\frac{1}{2}\left\{\mathbf{v}^T\left(\mathbf{d}\mathbf{d}^T + \boldsymbol{\Omega}^{-1}\right)\mathbf{v}\right\}} d\mathbf{v} \\ &= \frac{\left|\boldsymbol{I} + \boldsymbol{\Omega}\boldsymbol{d}\boldsymbol{d}^\top\right|^{-1/2}}{\sqrt{2\pi}\Phi\left(\frac{\boldsymbol{\mu}^\top \mathbf{d}}{\sqrt{1+\mathbf{d}^\top \boldsymbol{\Omega}\boldsymbol{d}}}\right)} (\boldsymbol{I} + \boldsymbol{\Omega}\boldsymbol{d}\boldsymbol{d}^\top)^{-1}\boldsymbol{\mu} \\ &\times e^{-\frac{1}{2}\left\{\boldsymbol{\mu}^\top\left(I - (I + \boldsymbol{\Omega}\mathbf{d}\mathbf{d}^\top)^{-1}\right)\boldsymbol{\mu}\right\}}. \end{aligned}$$

2.2. **Some special cases.** In this subsection, we study some special cases of SKSSMN distribution.

2.2.1. *Shape mixture of multivariate SCN.* If U takes one of two states $\{\nu_2, 1\}$ with probabilities $\{\nu_1, 1-\nu_1\}$, then $f_0(\mathbf{y})$ in (5) is the pdf of multivariate contaminated normal distribution (Ferriera et al., 2016). Then, the random vector $\mathbf{Y}$ achieves a shape mixture of multivariate skew-contaminated-normal (MSMSCN) distribution and its pdf, for $\mathbf{y} \in \mathbb{R}^p$, is

$$f(\mathbf{y}) = 2\{\nu_1\phi_p(\mathbf{y};\boldsymbol{\xi}, \nu_2^{-1}\boldsymbol{\Sigma}) + (1-\nu_1)\phi_p(\mathbf{y};\boldsymbol{\xi},\boldsymbol{\Sigma})\}\Phi\left(\frac{\mathbf{d}^\top\boldsymbol{\mu}}{\sqrt{1+\mathbf{d}^\top\boldsymbol{\Omega}\mathbf{d}}}\right). \tag{17}$$

From Ferreira et al. (2016) we have

$$E(U \mid \mathbf{y}) = \frac{1-\nu_1+\nu_1\nu_2^{\frac{p}{2}+1}\exp\{(1-\nu_2)\mathbf{d}^\top\mathbf{d}/2\}}{1-\nu_1+\nu_1\nu_2^{p/2}\exp\{(1-\nu_2)\mathbf{d}^\top\mathbf{d}/2\}}.$$

2.2.2. *Shape mixture of multivariate SSL.* If $U \sim Beta(\nu, 1)$ then the pdf (5), for $\mathbf{y} \in \mathbb{R}^p$, is updated as

$$f(\mathbf{y}) = 2\nu\int_0^1 u^{\nu-1}\phi_p(\mathbf{y};\boldsymbol{\xi}, u^{-1}\boldsymbol{\Sigma})du\Phi\left(\frac{\mathbf{d}^\top\boldsymbol{\mu}}{\sqrt{1+\mathbf{d}^\top\boldsymbol{\Omega}\mathbf{d}}}\right). \tag{18}$$

Thus we say $\mathbf{Y}$ has a shape mixture of skew slash (MSMSSL) distribution. In this case $U \mid \mathbf{y} \sim TG(\nu+p/2, \mathbf{d}^\top\mathbf{d}/2, 1)$, where $TG(a,b,t)$ is the right truncated gamma distribution, so we have

$$E(U \mid y) = \frac{2}{\mathbf{d}^\top\mathbf{d}}\frac{\gamma(\nu+\frac{p}{2}+1, \frac{\mathbf{d}^\top\mathbf{d}}{2})}{\gamma(\nu+\frac{p}{2}, \frac{\mathbf{d}^\top\mathbf{d}}{2})},$$

and

$$E(\log U \mid y) = \left(\frac{2}{\mathbf{d}^\top\mathbf{d}}\right)^{\nu+\frac{p}{2}+1}\left(\log\left(\frac{2}{\mathbf{d}^\top\mathbf{d}}\right) + \frac{\partial}{\partial\nu}\log\gamma\left(\nu+\frac{p}{2}, \frac{\mathbf{d}^\top\mathbf{d}}{2}\right)\right),$$

where $\gamma(a,b) = \int_0^b x^{a-1}e^{-x}dx$, denotes the incomplete gamma function.

2.2.3. *Shape mixture of skew-generalized-hyperbolic-normal (SGHN).* Let $S = 1/U$ has a generalized inverse Gaussian (GIG) distribution with pdf

$$\frac{\left(\frac{\psi}{\chi}\right)^{\frac{\lambda}{2}}}{2K_\lambda\left(\sqrt{\chi\psi}\right)}\, s^{\lambda-1}e^{-\frac{1}{2}\left(\chi s^{-1}+\psi s\right)}, \quad s>0,$$

where

$$K_\lambda(x) = \frac{1}{2}\int_0^{+\infty} y^{\lambda-1}e^{-\frac{x}{2}\left(y+y^{-1}\right)}dy,\ x>0, \tag{19}$$

being the Bessel function of the third kind with index λ. The parameters λ, χ and ψ satisfy $\chi > 0, \psi \geq 0$ if $\lambda < 0$, $\chi > 0, \psi > 0$ if $\lambda = 0$ and $\chi \geq 0, \psi > 0$ if $\lambda > 0$. Then, the pdf $f_0(.)$ in (5) reduces to the pdf of multivariate symmetric generalized hyperbolic distribution which is an elliptical distribution (See McNeil et al. 2015) which is

$$
(20) \qquad f_0(\mathbf{y}) = \frac{\sqrt{\chi\psi}^{-\lambda}\psi^{\frac{p}{2}}}{(2\pi)^{\frac{p}{2}}\,|\mathbf{\Sigma}|^{\frac{1}{2}}K_\lambda\left(\sqrt{\chi\psi}\right)} \frac{K_{\lambda-\frac{p}{2}}\left(\sqrt{\psi\left(\chi+q\left(\mathbf{x}\right)\right)}\right)}{\sqrt{\psi\left(\chi+q\left(\mathbf{y}\right)\right)}^{\frac{p}{2}-\lambda}},
$$

where $\mathbf{y} \in \mathbb{R}^p$, $q(\mathbf{y}) = (\mathbf{y}-\boldsymbol{\xi})^T \mathbf{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\xi})$ and $K_\lambda(.)$ is defined in(19). Now, the shape mixture of skew-generalized-hyperbolic-normal (SMSGHN) distribution follows from Proposition 2.1 by substituting (20) into (5).

The following remark states some other special cases of the SKSSMN distributions.

Remark 2.3. Let $f_0(.)$ in (20). Then, it becomes: the symmetric normal inverse Gaussian (NIG) distribution if $\lambda = -0.5$, the symmetric variance gamma distribution (VG) if $\lambda > 0$ and $\chi = 0$, the Student's t distribution if $\lambda < 0$ and $\psi = 0$, the symmetric hyperbolic distribution if $\lambda = (p+1)/2$. These choices generate, respectively, the shape mixtures of skew-symmetric-NIG-normal, skew-symmetric-VG-normal, skew-t- normal and skew-symmetric-hyperbolic-normal distributions. Also, some other distributions for U can be considered to generate more SMN distributions for $f_0(.)$ in (5) as introduced in Section 3 of Branco and Dey (2001), for example logistic, stable, exponential power and Pearson type II distributions. Considering these cases for $f_0(.)$ in Proposition 2.1, respectively, follow the shape mixtures of skew- logistic-normal, skew-stable-normal, skew-exponential-power-normal, and skew-Pearson type II-normal distributions. Two other examples for $f_0(.)$ are symmetric forms of distributions introduced by Pourmousa et al. (2015) and Naderi et al. (2018) by supposing the Birnbaum-Saunders and Lindley distributions for $X = 1/U$, respectively.

From Remark 1, If $\lambda = -\nu/2$, $\chi = \nu/2$ and $\psi = 0$, then $U \sim \Gamma(\frac{\nu}{2}, \frac{\nu}{2})$ and we have a shape mixture of multivariate skew-t-normal (MSMSTN) distribution with pdf

$$
(21) \qquad f(\mathbf{y}) = 2t_p(\mathbf{y};\boldsymbol{\xi},\mathbf{\Sigma},\nu)\Phi\left(\frac{\mathbf{d}^\top\boldsymbol{\mu}}{\sqrt{1+\mathbf{d}^\top\mathbf{\Omega}\mathbf{d}}}\right), \ \mathbf{y} \in \mathbb{R}^p.
$$

In this case, $U|\mathbf{y} \sim \Gamma\left(\frac{\nu+p}{2}, \frac{\nu+\mathbf{d}^\top\mathbf{d}}{2}\right)$ and thus $E(U \mid \mathbf{y}) = \frac{\nu+p}{\nu+\mathbf{d}^\top\mathbf{d}}$ and

$$
E(\log U \mid \mathbf{y}) = DG\left(\frac{\nu+p}{2}\right) - \log\left(\frac{\nu+\mathbf{d}^\top\mathbf{d}}{2}\right),
$$

where $\mathrm{DG}(x) = \frac{d}{dx}\log\Gamma(x)$ is the digamma function.

Figure 1 shows the density surface and contour plots of bivariate distributions MSMSTN (for $\nu = 4$), MSMSCN (for $\nu_1 = 0.2$ and $\nu_2 = 0.8$) and MSMSSL (for $\nu = 4$) with pdfs, respectively, in (21), (17), and (18), and each has common location vectors $\boldsymbol{\xi}$ and $\boldsymbol{\mu}$ and common dispersion matrices $\mathbf{\Sigma}$ and

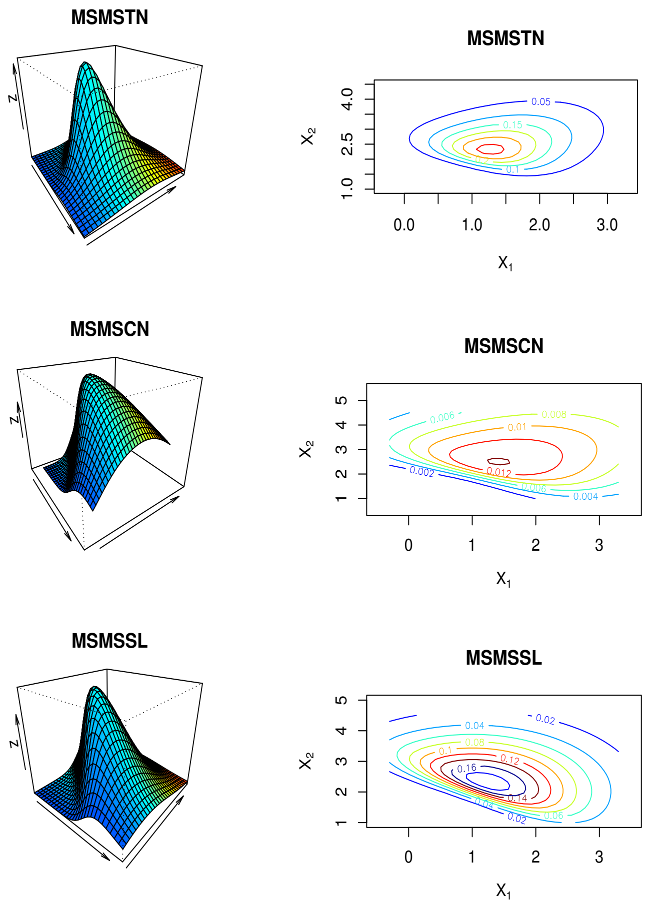

FIGURE 1. Density and contour plots of bivariate MSMSTN, MSMSCN and MSMSSL distributions.

$\boldsymbol{\Omega}$ specified as

$$\boldsymbol{\xi} = (1,2)^{\top}, \qquad \boldsymbol{\mu} = (3,4)^{\top},$$
$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\Omega} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

These plots show the asymmetry and kurtosis of the introduced distributions well. Therefore, using them to fit economic data is justifiable.

3. ECME algorithm

The EM algorithm (Dempster et al., 1977) serves as a versatile tool for maximum likelihood estimation in models involving missing data or latent variables. Its strength lies in maintaining implementation simplicity and ensuring monotonic convergence. However, directly applying the EM algorithm to estimate the SMSTN model faces challenges due to intractable computations in the M-step.

To overcome this limitation, we propose employing the expectation-conditional maximization (ECM) algorithm (Meng and Rubin, 1993). The ECM algorithm replaces the M-step of EM with a series of simpler conditional maximization steps. Each step maximizes the constrained function over $\mathbf{\Theta}$ while keeping some partitions fixed at their previous updates. Although the ECM algorithm retains the favorable convergence properties of EM and converges to a stationary point, it may exhibit slow convergence in certain scenarios.

To enhance the convergence rate compared to ECM, we suggest utilizing the Expectation Conditional Maximization Either (ECME) algorithm (Liu and Rubin, 1994). In ECME, some conditional maximization (CM) steps of ECM are replaced with the CML-step, which maximizes the correspondingly constrained actual-likelihood function. This modification aims to expedite the convergence process while preserving the desirable properties inherited from the ECM algorithm.

Let $\mathbf{y}_1, \cdots, \mathbf{y}_n$ be a random sample of size n from $SKSSMN(\mathbf{\Theta})$, where $\mathbf{\Theta} = (\boldsymbol{\xi}, \mathbf{\Sigma}, \boldsymbol{\mu}, \mathbf{\Omega}, \boldsymbol{\nu})$. Also assume $\mathbf{V} = (\mathbf{v}_1, ..., \mathbf{v}_n)$, $\boldsymbol{U} = (u_1, ..., u_n)$ and $\mathbf{W} = (w_1, ..., w_n)$, using (7), the complete data log-likelihood function of $\mathbf{\Theta}$ is given by

$$
\begin{aligned}
(22)\qquad \ell_c(\mathbf{\Theta}|\mathbf{y}, \mathbf{v}, w, u) &= \sum_{i=1}^{n} \ln f(\mathbf{y}_i, w_i, \mathbf{v}_i, u_i) \\
&= -\frac{n}{2} \log |\mathbf{\Omega}| \\
&- \frac{n}{2} \log |\mathbf{\Sigma}| \\
&- \frac{1}{2} \sum_{i=1}^{n} u_i \mathbf{d}_i^{\top} \mathbf{d}_i \\
&- \frac{1}{2} \sum_{i=1}^{n} \left(w_i - \mathbf{d}_i^{\top} \mathbf{v}_i\right)^2 \\
&- \frac{1}{2} \sum_{i=1}^{n} \left(\mathbf{v}_i - \boldsymbol{\mu}\right)^T \mathbf{\Omega}^{-1} \left(\mathbf{v}_i - \boldsymbol{\mu}\right) \\
&+ \sum_{i=1}^{n} \log h(u_i; \boldsymbol{\nu}).
\end{aligned}
$$

The expected value of the complete log-likelihood function given observed data, evaluated with $\mathbf{\Theta} = \hat{\mathbf{\Theta}}^{(k)}$, which we shall denote by the Q-function, is

$$\begin{aligned} Q\left(\mathbf{\Theta}|\hat{\mathbf{\Theta}}^{(k)}\right) &= E\left(\ell_c(\mathbf{\Theta})|\mathbf{y},\mathbf{v},w,u\right) \\ &= d - \frac{n}{2}\log|\mathbf{\Omega}| - \frac{n}{2}\log|\mathbf{\Sigma}| - \frac{1}{2}\sum_{i=1}^{n}\widehat{s}_{1i}^{(k)}\mathbf{d}_i^T\mathbf{d}_i + \boldsymbol{\mu}^T\mathbf{\Omega}^{-1}\sum_{i=1}^{n}\widehat{\mathbf{s}}_{2i}^{(k)} \\ &- \frac{n}{2}\boldsymbol{\mu}^T\mathbf{\Omega}^{-1}\boldsymbol{\mu} - \frac{1}{2}tr\left(\mathbf{\Omega}^{-1}\sum_{i=1}^{n}\widehat{\mathbf{s}}_{3i}^{(k)}\right) - \frac{1}{2}\sum_{i=1}^{n}\mathbf{d}_i^T\widehat{\mathbf{s}}_{3i}^{(k)}\mathbf{d}_i \\ &+ \sum_{i=1}^{n}\mathbf{d}_i^T\widehat{\mathbf{s}}_{4i}^{(k)} + \sum_{i=1}^{n}E\left(\log h(u_i;\boldsymbol{\nu})|\mathbf{y}_i,\hat{\mathbf{\Theta}}^{(k)}\right), \end{aligned} \tag{23}$$

where

$$\begin{aligned} \widehat{s}_{1i}^{(k)} &= E\left(u_i|\mathbf{y}_i,\hat{\boldsymbol{\theta}}^{(k)}\right),\ \widehat{\mathbf{s}}_{2i}^{(k)} = E\left(\mathbf{V_i}|\mathbf{y}_i,\hat{\boldsymbol{\theta}}^{(k)}\right), \\ \widehat{\mathbf{s}}_{3i}^{(k)} &= E\left(\mathbf{V_i}\mathbf{V_i^T}|\mathbf{y}_i,\hat{\boldsymbol{\theta}}^{(k)}\right),\ \widehat{\mathbf{s}}_{4i}^{(k)} = E\left(W_i\mathbf{V}_i|\mathbf{y}_i,\hat{\boldsymbol{\theta}}^{(k)}\right). \end{aligned} \tag{24}$$

Let $\mathbf{\Gamma} = \mathbf{\Sigma}^{-1/2}$, proposed ECM algorithm for the SKSSMN distribution, in general version, consists of one E-step and five CM-steps as described below:

E-step: Given $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}^{(k)}$, compute $\widehat{s}_{1i}^{(k)}, \widehat{\mathbf{s}}_{2i}^{(k)}, \widehat{\mathbf{s}}_{3i}^{(k)}$ and $\widehat{\mathbf{s}}_{4i}^{(k)}$ in Eq. (24) using expectations (14), (15) and (16), for $i = 1, ..., n$.

CM-steps:

1. Update $\hat{\boldsymbol{\xi}}^{(k)}$ by maximizing (23) over $\boldsymbol{\xi}$, which leads to

$$\hat{\boldsymbol{\xi}}^{(k+1)} = \left(\sum_{i=1}^{n}\left(\hat{s}_{1i}^{(k)}\mathbf{I}_p + \hat{\mathbf{s}}_{3i}^{(k)}\right)\right)^{-1}\left\{\sum_{i=1}^{n}\left((\hat{\mathbf{s}}_{1i}^{(k)}\mathbf{I}_p + \hat{\mathbf{s}}_{3i}^{(k)})\mathbf{y}_i - \mathbf{\Gamma}^{-1}\hat{\mathbf{s}}_{4i}^{(k)}\right)\right\}.$$

2. Fix $\boldsymbol{\xi} = \hat{\boldsymbol{\xi}}^{(k+1)}$. For updating $\hat{\mathbf{\Gamma}}^{(k)}$, we optimize (23) over $\mathbf{\Gamma}$, which is equivalent to solving the following equation:

$$\sum_{i=1}^{n}A_i\mathbf{\Gamma}B_i\mathbf{\Gamma} - \frac{1}{2}\left(\sum_{i=1}^{n}C_i + \sum_{i=1}^{n}C_i^\top\right)\mathbf{\Gamma} - n\mathbf{I}_p = 0, \tag{25}$$

where $A_i = (\hat{\mathbf{s}}_{1i}^{(k)}\mathbf{I}_p + \hat{\mathbf{s}}_{3i}^{(k)})$, $B_i = (\mathbf{y}_i - \hat{\boldsymbol{\xi}}^{(k+1)})(\mathbf{y}_i - \hat{\boldsymbol{\xi}}^{(k+1)})^\top$, and $C_i = (\mathbf{y}_i - \hat{\boldsymbol{\xi}}^{(k+1)})\hat{\mathbf{s}}_{4i}^{(k)^\top}$.

The root of equation (25) is obtained numerically. Hence, $\hat{\mathbf{\Sigma}}^{(k+1)}$ is simply updated by $\hat{\mathbf{\Sigma}}^{(k+1)} = \hat{\mathbf{\Gamma}}^{-1^{(k+1)}}\hat{\mathbf{\Gamma}}^{-1^{(k+1)}}$.

3. Update $\hat{\boldsymbol{\mu}}^{(k)}$ by maximizing (23) over $\boldsymbol{\mu}$, which leads to

$$\hat{\boldsymbol{\mu}}^{(k+1)} = \frac{1}{n}\sum_{i=1}^{n}\hat{\mathbf{s}}_{2i}^{(k)}.$$

4. Fix $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}^{(k+1)}$, update $\hat{\boldsymbol{\Omega}}^{(k)}$ by maximizing (23) over $\boldsymbol{\Omega}$, which leads to

$$\hat{\boldsymbol{\Omega}}^{(k+1)} = \frac{1}{n}\sum_{i=1}^{n} \hat{\mathbf{s}}_{3i}^{(k)} - \hat{\boldsymbol{\mu}}^{(k+1)}\hat{\boldsymbol{\mu}}^{(k+1)^\top},$$

4′. In the case with $\boldsymbol{\Omega}$ assumed to be diagonal, say $\boldsymbol{\Omega} = Diag(\omega_1^2, ..., \omega_p^2) = Diag(\boldsymbol{\eta})$, then update $\hat{\boldsymbol{\eta}}^{(k)}$ by

$$\hat{\boldsymbol{\eta}}^{(k+1)} = Diag\left(\frac{1}{n}\sum_{i=1}^{n} \hat{\mathbf{s}}_{3i}^{(k)} - \hat{\boldsymbol{\mu}}^{(k+1)}\hat{\boldsymbol{\mu}}^{(k+1)^\top}\right),$$

4″. In the case with $\boldsymbol{\Omega}$ assumed to be $\boldsymbol{\Omega} = \tau \mathbf{I}_p$, then update $\hat{\tau}^{(k)}$ by

$$\hat{\tau}^{(k+1)} = \frac{1}{np}\left\{tr(\sum_{i=1}^{n} \hat{\mathbf{s}}_{3i}^{(k)}) - n\hat{\boldsymbol{\mu}}^{(k+1)^\top}\hat{\boldsymbol{\mu}}^{(k+1)}\right\},$$

where $tr(\boldsymbol{A})$ denotes the trace of matrix $\boldsymbol{A}$.

Updating $\hat{\boldsymbol{\nu}}^{(k)}$ is related to the form of $h(u_i; \boldsymbol{\nu})$. Since the conditional expectation $E(\log h(u_i; \boldsymbol{\nu})|\mathbf{y}_i, \hat{\boldsymbol{\theta}}^{(k)})$ may be difficult to evaluate, one may resort to maximize the restricted actual log-likelihood function, leading to the following 'CML-step':

CML-step: Update $\hat{\boldsymbol{\nu}}^{(k+1)}$ by

$$\hat{\boldsymbol{\nu}}^{(k+1)} = \arg\max_{\boldsymbol{\nu}} \sum_{i=1}^{n} \log f_{SKSSMN}\left(\mathbf{y}_i; \hat{\boldsymbol{\xi}}^{(k+1)}, \hat{\boldsymbol{\Sigma}}^{(k+1)}, \hat{\boldsymbol{\mu}}^{(k+1)}, \hat{\boldsymbol{\Omega}}^{(k+1)}, \boldsymbol{\nu}\right),$$

which is equal to maximize $\sum_{i=1}^{n} \log f_0(\boldsymbol{y}_i)$ with respect to $\boldsymbol{\nu}$. The iterations of the above algorithm are repeated until a suitable convergence rule is satisfied, e.g., $\left\|\widehat{\boldsymbol{\theta}}^{(k+1)} - \widehat{\boldsymbol{\theta}}^{(k)}\right\|$ is sufficiently small.

4. Simulation

In this section, we present two simulation experiments aimed at assessing the efficacy of the proposed models in fitting the data. Prior to delving into these experiments, it is necessary to outline the simulation method based on the introduced distributions. To simulate, we employ the following algorithm using convolution-type representation in Equation (6).

Algorithm to generate random samples from SKSSMN:

(1) Generate u from $h(u; \nu)$;

(2) Generate $x_1, \cdots, x_{2p+1}$, independently from $N(0, 1)$;

  (a) Let $z_0 = x_{2p+1}$ and $\mathbf{z}_1 = (x_1, \cdots, x_p)^\top$,

  (b) Let $\mathbf{z}_2 = (x_{p+1}, \cdots, x_{2p})^\top$ and $\mathbf{v} = \boldsymbol{\mu} + \boldsymbol{\Omega}^{1/2}\mathbf{z}_2$,

(c) Compute

$$\boldsymbol{y} = \boldsymbol{\xi} + \boldsymbol{\Sigma}^{1/2} \times \left\{ \frac{u^{-1/2}\mathbf{v}}{\sqrt{u + \mathbf{v}^\top \mathbf{v}}} |Z_0| + \left(u\mathbf{I}_p + \mathbf{v}\mathbf{v}^\top\right)^{-1/2} \mathbf{z}_1 \right\}.$$

Repeat steps (1) and (2) n times to get $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n$.

4.1. **Experiment 1.** In this experiment, we first generate 500 observations from each of the distributions MSMSTN (for $\nu = 4$), MSMSCN (for $\nu_1 = 0.2$ and $\nu_2 = 0.8$) and MSMSSL (for $\nu = 4$), respectively, in (21), (17), and (18), and for the other true parameters, we set the same location vectors and dispersion matrices as follows:

$$\boldsymbol{\xi} = (0, 0)^\top, \qquad \boldsymbol{\mu} = (1, -1)^\top,$$
$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\Omega} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Subsequently, employing the algorithm outlined in the preceding section, we proceed to fit the introduced distributions to the generating data. Figure 2 showcases scatter plots depicting the generated data from each distribution alongside their respective contour plots.

In Figure 2, it is evident that the bivariate distributions effectively capture the data. Among these distributions, the MSMSCN distribution demonstrates superior fitting to the generated data.

4.2. **Experiment 2.** In the second experiment, our objective is to examine the impact of outliers on distribution fitting. To do so, we create 100 data

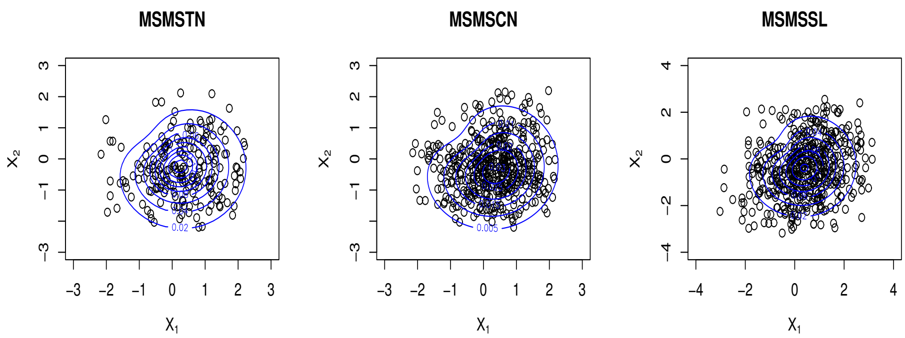

FIGURE 2. Scatter plots for generated data and contour lines of bivariate MSMSTN, MSMSCN and MSMSSL distributions.

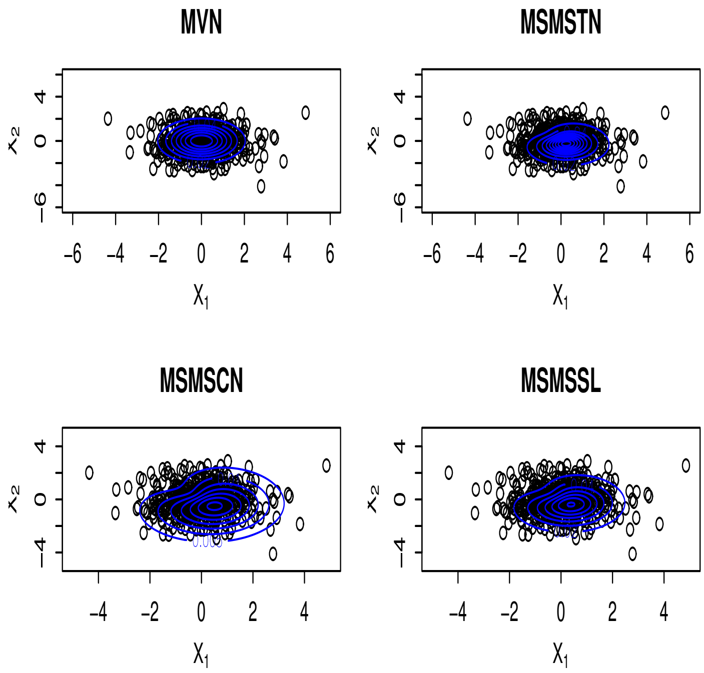

Figure 3. Generative data from bivariate normal with some outliers and contour plots of bivariate MSMSTN, MSMSCN, and MSMSSL distributions.

points from a bivariate normal distribution with true parameters similar to the previous experiment and add 10 additional data points from a uniform distribution within the range of $min(Y) - 1.5$ to $max(Y) + 1.5$ to each component. This results in the generated data exhibiting increased skewness compared to the original distribution. We then fit the MSMSTN, MSMSCN, and MSMSSL distributions to these data points. Figure 3 illustrates both the simulated data and the distributions fitted to them. According to Figure 3, among the introduced distributions, the MSMSCN distribution exhibits the best fit to the generated data, which have higher kurtosis than the normal distribution.

5. **Application**

The methodology implemented in this study involves the analysis of Stock Market data, specifically focusing on the S&P 500 Composite Index of the United States (SP500) and Singapore's Straits Times Index (STI). The data, spanning from January 1, 2000, to December 31, 2022, was sourced from https://finance.yahoo.com. Given the distinctive characteristics of asset returns, including sharper peaks, slight skewness, and heavier tails compared to the normal distribution (Adcock et al., 2015), we recognize the need for innovative approaches to handle financial data effectively. Skewed distributions such as the skew-normal (SN) and its robust extensions present promising solutions by accommodating these non-normal characteristics collectively. In our analysis, we utilized adjusted closing prices to compute daily log returns (r_t) for each share, expressed as a percentage using the formula $r_t = (\log P_t - \log P_{t-1}) \times 100\%$, where P_t and P_{t-1} represent the adjusted closing price of the share at time t and $t-1$, respectively.

Table 1 presents descriptive characteristics, indicating a slight negative skewness but notably high excess kurtosis across the two log-return series. Additionally, we employ the D'Agostino test for skewness and the Anscombe-Glynn test for kurtosis and the results are presented in Table 2. The findings imply that the observed log-returns might be better represented by particular skew distributions instead of the normal distribution.

Table 1. Descriptive statistics for stock market indices.

Stock Index	Sample size	Mean	Standard deviation	Skewness	Kurtosis
SP500	4276	0.010	1.245	–0.194	11.163
STI	4276	0.003	1.159	–0.333	8.952

Table 2. Test of skewness and kurtosis for stock market indices.

Stock Index	Sample size	D'Agostino test (p-value)	Anscombe Glynn test (p–value)
SP500	4276	–5.13 (<1e-4)	25.86 (<1e-4)
STI	4276	–8.68 (<1e-4)	23.28 (<1e-4)

Many researchers have suggested that stock market data are correlated and can be considered as multivariate data. Junior and Franca (2012) discussed on correlations between financial markets in times of crisis. Madan (2020) used

Stock Market data for fitting multivariate distributions. Mata et al. (2021) studied multivariate distributions in the Stock Markets of some countries. We apply the ECM algorithm, as detailed in Section 2, to fit the SKSSMN distributions to the bivariate (SP500, STI) log-returns. Additionally, we fit the MSTN, MSSL, MSCN, MST, and MSN distributions to facilitate comparison. Maximum likelihood estimates of the introduced model parameters were obtained through the ECM algorithm. However, due to the large number of parameters and for the sake of summarizing the article, we won't display them here.

Our comprehensive analysis extends beyond the evaluation of model performance using Table 3, encompassing visual representations of the data patterns through contour plots depicted in Figures 4 and 5. These figures superimpose several fitted bivariate skew densities, allowing for a direct comparison of model fit across different distributions. Notably, our findings from these visualizations align with the quantitative assessments presented in Table 3. This evaluation was based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which are widely utilized measures for model selection. A lower value of AIC or BIC indicates a better fit of the model to the data.

Upon meticulous examination of Table 3, which presents the results of our model evaluations, it is evident that the MSMSCN model emerges as the most favorable choice for the bivariate data series. Specifically, the MSMSCN model exhibits significantly lower AIC and BIC values compared to alternative models considered in our analysis. This suggests that the MSMSCN model not only provides a superior fit to the observed data but also offers a more parsimonious representation of the underlying data structure.

Moreover, the contour plots in Figures 4 and 5 illustrate that the MSMSCN distribution exhibits a more satisfactory fit to the data compared to other competing models. This observation further reinforces the robustness and efficacy of the MSMSCN model in capturing the underlying patterns and dependencies within the bivariate data series. In summary, our results from both quantitative evaluations and visual representations collectively highlight the exceptional performance of the MSMSCN distribution in modeling multivariate data with skewed and heavy-tailed distributions. The alignment of findings across different analytical approaches lends further support to the robustness and efficacy of the MSMSCN model, positioning it as a valuable tool for researchers and practitioners seeking to accurately model and analyze complex multivariate datasets.

6. **Conclusions**

In this paper, we undertake an in-depth examination of SKSSMN distributions, introducing a diverse array of multivariate skewed distributions characterized by significantly elevated levels of kurtosis and skewness compared to the standard multivariate normal distribution. Throughout our investigation, we thoroughly explore various specific instances within this expansive family.

TABLE 3. Information criteria of several fitted distribution to financial market data. ($\ell_{\max}$ is the maximized log-likelihood)

Distribution	ℓ_{max}	AIC	BIC
MSN	-13638.71	27291.43	27335.95
MST	-12677.53	25377.07	25447.04
MSTN	-12596.28	25212.56	25276.17
MSSL	-12648.44	25310.88	25355.41
MSCN	-12687.47	25388.95	25433.47
MSMSTN	-12594.09	25204.17	25255.06
MSMSSL	-12641.78	25301.56	25358.80
MSMSCN	-12589.04	25194.08	25244.97

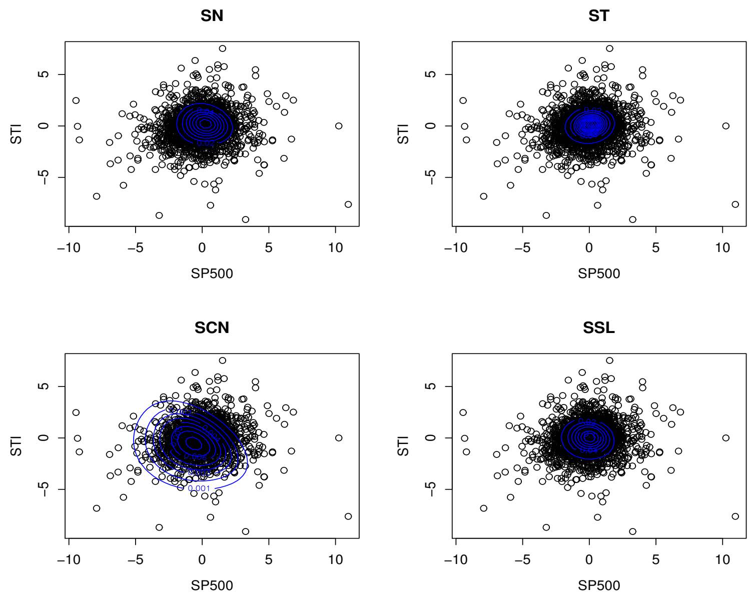

FIGURE 4. Scatter plots for the log-returns of SP500 and STI indices and contour lines of some fitted bivariate SMSN distributions.

Our analysis is underpinned by the development and refinement of an ECM algorithm meticulously crafted to precisely estimate the parameters governing these models. Subsequently, we employ these developed models to analyze bivariate Stock Market data. Notably, our comparison with traditional distributions, particularly the scale-shape mixtures of multivariate normal, reveals a superior fit, with the MSMSCN model demonstrating particularly strong performance. Furthermore, our findings demonstrate a significant enhancement in performance compared to scale mixtures of multivariate normal distribution

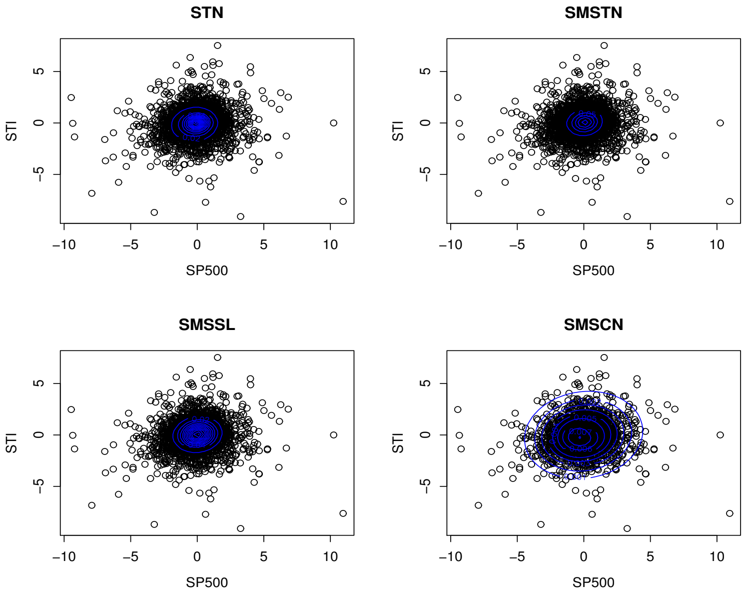

FIGURE 5. Scatter plots for the log-returns of SP500 and STI indices and contour lines of some fitted bivariate SKSSMN distributions.

models, underscoring the superior efficacy of these distributions in capturing the nuanced behaviors inherent in Stock Market data.

Acknowledgments

We are grateful to the Editor and anonymous reviewers for their valuable comments and suggestions which have led to improvement in this paper. Furthermore, we must dedicate this work to and honor the memory of our esteemed professor, Dr. Mahbanoo Tata, who was a distinguished leader in both knowledge and ethics within the Iranian statistical community.

References

[1] Adcock, C., Eling, M., & Loperfido, N. (2015) Skewed distributions in finance and actuarial science: a review. Europian Journal of Finance 21:1–29. https://doi.org/10.1080/1351847X.2012.720269

[2] Amiri, M., & Balakrishnan, N. (2022). Hessian and increasing-Hessian orderings of scale-shape mixtures of multivariate skew-normal distributions and applications. Journal of Computational and Applied Mathematics, 402, 113801. https://doi.org/10.1016/j.cam.2021.113801

[3] Arellano-Valle, R. B., & Genton, M. G. (2005). On fundamental skew distributions. Journal of Multivariate Analysis, 96(1), 93-116. https://doi.org/10.1016/j.jmva.2004.10.002

[4] Azzalini, A., & Capitanio, A. (2014). The Skew-Normal and Related Families. Cambridge University Press.

[5] Bodie, Z., & Kane, A. (2020). Essentials of investments.

[6] Branco, M.D., & Dey, D. k. (2001). A general class of multivariate skew-elliptical distributions. Journal of Multivariate Analysis, 79, 99–113. https://doi.org/10.1006/jmva.2000.1960

[7] Cont, R. (2007). Volatility clustering in financial markets: empirical facts and agent-based models. In Long memory in economics (pp. 289-309). Berlin, Heidelberg: Springer Berlin Heidelberg.

[8] Denkowska, A., & Wanat, S. (2020). A tail dependence-based MST and their topological indicators in modeling systemic risk in the European insurance sector. Risks, 8(2), 39. https://doi.org/10.3390/risks8020039

[9] Fabozzi, F. J., Markowitz, H. M., & Gupta, F. (2008). Portfolio selection. Handbook of finance.

[10] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1), 3-56. https://doi.org/10.1016/0304-405X(93)90023-5

[11] Ferreira, C. S., Bolfarine, H., & Lachos, V. H. (2011). Skew scale mixtures of normal distributions: properties and estimation. Statistical Methodology, 8(2), 154-171. https://doi.org/10.1016/j.stamet.2010.09.001

[12] Ferreira, C. S., Lachos, V. H., & Bolfarine, H. (2016). Likelihood-based inference for multivariate skew scale mixtures of normal distributions. AStA Advances in Statistical Analysis, 100(4), 421-441. https://doi.org/10.1007/s10182-016-0266-z

[13] Gómez-Déniz, E., Calderín-Ojeda, E., & Gómez, H. W. (2022). Symmetric and Asymmetric Distributions: Theoretical Developments and Applications III. Symmetry, 14(10), 21-43. https://doi.org/10.3390/sym14102143

[14] Härdle, W. K., & Simar, L. (2019). Applied multivariate statistical analysis. Springer Nature.

[15] Harvey, C. R., & Siddique, A. (2000). Conditional skewness in asset pricing tests. The Journal of finance, 55(3), 1263-1295. https://doi.org/10.1111/0022-1082.00247

[16] Jamalizadeh, A., & Lin, T. I. (2017). A general class of scale-shape mixtures of skew-normal distributions: properties and estimation. Computational Statistics, 32, 451-474. https://doi.org/10.1007/s00180-016-0691-1

[17] Jorion, P. (2007). Value at risk: the new benchmark for managing financial risk. McGraw-Hill.

[18] Junior, L. S., & Franca, I. D. P. (2012). Correlation of financial markets in times of crisis. Physica A: Statistical Mechanics and its Applications, 391(1-2), 187-208. https://doi.org/10.1016/j.physa.2011.07.023

[19] Kotz, S., Balakrishnan, N., & Johnson, N. L. (2004). Continuous multivariate distributions, Volume 1: Models and applications (Vol. 1). John Wiley & Sons.

[20] Liu, C.H., & Rubin, D.B. (1994). The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence. Biometrika, 81, 633–648. https://doi.org/10.1093/biomet/81.4.633

[21] Madan, D. B. (2020). Multivariate distributions for financial returns. International Journal of Theoretical and Applied Finance, 23(06), 2050041. https://doi.org/10.1142/S0219024920500417

[22] Mahdavi, A., Amirzadeh, V., Jamalizadeh, A., & Lin, T. I. (2021). A Multivariate flexible skew-symmetric-normal distribution: Scale-shape mixtures and parameter estimation via selection representation. Symmetry, 13(8), 1343. https://doi.org/10.3390/sym13081343

[23] Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). Multivariate Analysis. Academic Press.

[24] Mata, L. M., Núñez Mora, J. A., & Serrano Bautista, R. (2021). Multivariate Distribution in the Stock Markets of Brazil, Russia, India, and China. SAGE Open, 11(2), 21582440211009509. https://doi.org/10.1177/21582440211009509

[25] McNeil, A. J., Frey, R., & Embrechts, P. (2015). Quantitative risk management: Concepts, techniques and tools. Princeton university press.

[26] Meng, X. L.,& Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: a general framework. Biometrika, 80, 267-278. https://doi.org/10.1093/biomet/80.2.267

[27] Mondal, S., Arellano-Valle, R. B., & Genton, M. G. (2023). A multivariate modified skew-normal distribution. Statistical Papers, 1-45. https://doi.org/10.1007/s00362-023-01397-1

[28] Naderi, M., Arabpour, A., & Jamalizadeh, A. (2018). Multivariate normal mean-variance mixture distribution based on Lindley distribution. Communications in Statistics-Simulation and Computation, 47, 1179–1192. https://doi.org/10.1080/03610918.2017.1307400

[29] Pourmousa, R., Jamalizadeh, A., & Rezapour, M. (2015). Multivariate normal mean–variance mixture distribution based on Birnbaum–Saunders distribution. Journal of Statistical Computation and Simulation, 85(13), 2736–2749. https://doi.org/10.1080/00949655.2014.937435

[30] Pu, T., Zhang, Y., & Yin, C. (2023). Generalized location-scale mixtures of elliptical distributions: Definitions and stochastic comparisons. Communications in Statistics-Theory and Methods, 1-25. https://doi.org/10.1080/03610926.2023.2165407

Mostafa Tamandi
Orcid number: 0000-0003-0056-4642
Department of Statistics
Vali-e-Asr University of Rafsanjan
Rafsanjan, Iran.
Email address: tamandi@vru.ac.ir

Mehdi Amiri
Orcid number: 0000-0001-9449-8791
Department of Statistics
Faculty of Basic Sciences
University of Hormozgan,
Bandar Abbas, Iran.
Email address: m.amiri@hormozgan.ac.ir