Brazilian Journal of Probability and Statistics
2020, Vol. 34, No. 3, 478–494
https://doi.org/10.1214/19-BJPS438

Moments of truncated scale mixtures of skew-normal distributions

Víctor H. Lachos[a], **Aldo M. Garay**[b] **and Celso R. B. Cabral**[c]

[a]*University of Connecticut*
[b]*Federal University of Pernambuco*
[c]*Federal University of Amazonas*

Abstract. In this work, we consider the problem of finding the moments of a doubly truncated member of the class of scale mixtures of skew-normal (TSMSN) distributions. We obtain a general result and then use it to derive the moments in the case of doubly truncated versions of skew-normal, skew-t, skew-slash and skew-contaminated normal distributions. Many properties of the TSMSN family are studied, inference procedures are developed and a simulation study is performed to assess the procedures. Two applications are also provided, one of them in the context of censored regression models and another in the field of actuarial sciences.

1 Introduction

The scale mixtures of skew-normal (SMSN) family of distributions, introduced by Branco and Dey (2001), is a very flexible class of distributions which takes into account at the same time skewness and heavy tails. Besides this, it has a stochastic representation that facilitates the study of many properties. The skew-t, skew-slash, skew-contaminated normal and all the symmetric class of scale mixtures of normal (SMN) distributions defined by Andrews and Mallows (1974) belong to the SMSN family. However, applications (through simulation or experimentation) often generate a large number of datasets that can be skewed-heavy-tailed with values restricted to a fixed interval. For example, variables such as pH, grades, viral load in HIV studies and humidity in environmental studies have upper and lower bounds, and the support of their distributions is restricted to some interval.

Some broadly related proposals and results have appeared in the literature under the concept of the truncated distribution. Kim (2008) presented the moments of a doubly truncated generalized Student-t distribution and showed its utility for solving statistical problems. Genç (2013) considered the problem of finding the moments of a doubly truncated member of the symmetrical class of SMN distributions. He obtained a general result and then used it to derive the moments in the case of doubly truncated versions of the Pearson type VII, slash, contaminated normal, double exponential and variance gamma distributions. He applied the results to some actuarial data. In the context of truncated skew distributions, Jamalizadeh, Pourmousa and Balakrishnan (2009) obtained the first two moments of the the truncated skew-normal and truncated skew-t distributions and Flecher, Allard and Naveau (2010) obtained a general recursive formula for the moments of the truncated skew-normal distribution and applied the results to model the relative humidity data. In this work, our objective is to combine the results of Flecher, Allard and Naveau (2010) and Genç (2013) to derive the moments of the truncated SMSN (TSMSN) distributions. We then particularize the general result for some common SMSN distributions mentioned above. Since the first four moments are most useful, we give expressions for them. Our proposal generalizes the results obtained by Kim (2008), Flecher, Allard and Naveau (2010), Genç (2013) and Garay et al. (2017).

Key words and phrases. Kurtosis, moments, truncated distributions, scale mixtures of skew-normal distributions, skewness.
Received May 2017; accepted March 2019.

The rest of the paper is organized as follows. In Section 2, we give a brief introduction of the SMSN and TSMSN distributions. In Section 3, we outline the main results related to the moments of the TSMSN distributions. Section 4 deals with particular cases of the TSMSN distributions. Section 5 discusses two applications, one of them in the context of censored regression models and another using real data in the field of actuarial sciences. Section 6 presents a simulation study to verify the performance of our proposed method and Section 7 concludes with some discussion and possible directions for future research.

2 Scale mixtures of skew-normal (SMSN) distributions

Throughout this paper, $\mathrm{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the p-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, $\phi_p(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\boldsymbol{\Phi}_p(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denote its probability density function (pdf) and cumulative distribution function (cdf), respectively. When $p = 1$ we drop the index p. In this case, if $\mu = 0$ and $\sigma^2 = 1$ (the standard case), we write $\phi(\cdot)$ for the pdf and $\Phi(\cdot)$ for the cdf.

We start by defining the skew-normal (SN) distribution and then we introduce some useful properties. As defined by Azzalini (1985), a random variable Z has a skew-normal distribution with location parameter μ, scale parameter σ^2 and skewness parameter λ, denoted by $Z \sim \mathrm{SN}(\mu, \sigma^2, \lambda)$, if its pdf is given by $\phi_{\mathrm{SN}}(z|\mu, \sigma^2, \lambda) = 2\phi(z|\mu, \sigma^2)\Phi(\lambda(z-\mu)/\sigma)$. We denote the cdf of Z by $\Phi_{\mathrm{SN}}(\cdot|\mu, \sigma^2, \lambda)$. In the standard case we use $\phi_{\mathrm{SN}}(\cdot|\lambda)$ and $\Phi_{\mathrm{SN}}(\cdot|\lambda)$ for the pdf and cdf, respectively. As proved by Azzalini and Dalla Valle (1996, eqn. 2.11), we have

$$\Phi_{\mathrm{SN}}(z|\mu, \sigma^2, \lambda) = 2\boldsymbol{\Phi}_2\left(\frac{z-\mu}{\sigma}\mathbf{e}_1\middle|\mathbf{0}, \boldsymbol{\Sigma}\right), \tag{2.1}$$

where $\mathbf{e}_1 = (1, 0)^\top$ and

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}, \quad \text{with } \delta = \lambda/\sqrt{1+\lambda^2}. \tag{2.2}$$

If $Z \sim \mathrm{SN}(\mu, \sigma^2, \lambda)$, then a convenient stochastic representation is given by $Z = \mu + \Delta|T_0| + \Gamma^{1/2}T_1$, where $\Delta = \sigma\delta$, $\Gamma = (1-\delta^2)\sigma^2$, T_0 and T_1 are independent standard normal random variables and $|\cdot|$ denotes the absolute value. This stochastic representation is useful to generate random samples and to obtain moments and other related properties.

Definition 1. Let $Z \sim \mathrm{SN}(0, \sigma^2, \lambda)$ and U a positive random variable with cdf $H(\cdot|\boldsymbol{\nu})$. Suppose that Z and U are independent. We say that the distribution of $Y = \mu + U^{-1/2}Z$ is a scale mixture of skew-normal distributions with location parameter μ, scale parameter σ^2, shape parameter λ, and mixture distribution $H(\cdot|\boldsymbol{\nu})$.

In this case, we use the notation $Y \sim \mathrm{SMSN}(\mu, \sigma^2, \lambda, \boldsymbol{\nu})$. The random variable U is known as *the scale factor* and the parameter $\boldsymbol{\nu}$ can be a vector. The form of a SMSN distribution is determined by the distribution of U. For example if $P(U = 1) = 1$, we have $Y \sim \mathrm{SN}(\mu, \sigma^2, \lambda)$. If $U \sim \mathrm{Gamma}(\nu/2, \nu/2)$, $\nu > 0$, where $\mathrm{Gamma}(a, b)$ denotes the gamma distribution with mean a/b, we have the skew-t distribution, denoted by $Y \sim \mathrm{ST}(\mu, \sigma^2, \lambda, \nu)$. Other examples will be given in Section 4. Note that when $\lambda = 0$, the family reduces to the class of scale mixtures of normal (SMN) distributions.

Using Definition 1, we observe that

$$Y|U = u \sim \mathrm{SN}(\mu, u^{-1}\sigma^2, \lambda) \tag{2.3}$$

and integrating out U from the joint density of Y and U leads to the following marginal pdf of Y:

$$\phi_{\mathrm{SMSN}}(y|\mu,\sigma^2,\lambda,\boldsymbol{\nu}) = 2\int_0^\infty \phi(y|\mu,u^{-1}\sigma^2)\Phi\left(u^{1/2}\frac{\lambda(y-\mu)}{\sigma}\right)dH(u|\boldsymbol{\nu}).$$

Let $\Phi_{\mathrm{SMSN}}(y|\mu,\sigma^2,\lambda,\boldsymbol{\nu})$ be the cdf of Y. Using (2.3) and (2.1) we obtain the following expression for this cdf in the standard case:

$$\Phi_{\mathrm{SMSN}}(y|\lambda,\boldsymbol{\nu}) = \mathrm{E}\left[\Phi_{\mathrm{SN}}\left(U^{1/2}y|\lambda\right)\right] = 2\mathrm{E}\left[\boldsymbol{\Phi}_2\left(U^{1/2}y\mathbf{e}_1|\mathbf{0},\boldsymbol{\Sigma}\right)\right]. \tag{2.4}$$

Now, we introduce a key concept to our theory, namely the truncated SMSN distribution.

Definition 2. Let $Y \sim \mathrm{SMSN}(\mu,\sigma^2,\lambda,\boldsymbol{\nu})$, with $P(a < Y < b) > 0$ for some fixed $a < b$. A random variable X has a truncated SMSN distribution in the interval $\lfloor a,b\rfloor$, denoted by $X \sim \mathrm{TSMSN}_{\lfloor a,b\rfloor}(\mu,\sigma^2,\lambda,\boldsymbol{\nu})$, if it has the same distribution as $Y|Y \in \lfloor a,b\rfloor$. Here $\lfloor a,b\rfloor$ means that each extreme of the interval can be either open or closed.

The truncated normal distribution is obtained when $Y \sim \mathrm{N}(\mu,\sigma^2)$. In this case, we use the notation $X \sim \mathrm{TN}_{\lfloor a,b\rfloor}(\mu,\sigma^2)$. Analogously, we define the truncated SMN ($\mathrm{TSMN}_{\lfloor a,b\rfloor}(\mu,\sigma^2,\boldsymbol{\nu})$), the truncated skew-normal ($\mathrm{TSN}_{\lfloor a,b\rfloor}(\mu,\sigma^2,\lambda)$) and the truncated skew-t ($\mathrm{TST}_{\lfloor a,b\rfloor}(\mu,\sigma^2,\lambda,\nu)$) distributions. As an obvious consequence of the definition, we have that the pdf of $X \sim \mathrm{TSMSN}_{\lfloor a,b\rfloor}(\mu,\sigma^2,\lambda,\boldsymbol{\nu})$ is given by:

$$\phi_{\mathrm{TSMSN}}\left(x|\mu,\sigma^2,\lambda,\boldsymbol{\nu};\lfloor a,b\rfloor\right) = \frac{\phi_{\mathrm{SMSN}}(x|\mu,\sigma^2,\lambda,\boldsymbol{\nu})}{\Phi_{\mathrm{SMSN}}(b|\mu,\sigma^2,\lambda,\boldsymbol{\nu}) - \Phi_{\mathrm{SMSN}}(a|\mu,\sigma^2,\lambda,\boldsymbol{\nu})} \times \mathbb{I}_{\lfloor a,b\rfloor}(x),$$

where $\mathbb{I}_B(y)$ denotes the indicator function, that is, $\mathbb{I}_B(y) = 1$ if $y \in B$ and $\mathbb{I}_B(y) = 0$ otherwise.

Let $X \sim \mathrm{TSMSN}_{\lfloor \alpha,\beta\rfloor}(0,1,\lambda;\boldsymbol{\nu})$. Then it is straightforward to prove that $Y = \mu + \sigma X$ has a $\mathrm{TSMSN}_{\lfloor a,b\rfloor}(\mu,\sigma^2,\lambda;\boldsymbol{\nu})$ distribution, where $a = \mu + \sigma\alpha$ and $b = \mu + \sigma\beta$. So, to compute the moments of Y, it is enough to compute the moments of X. Thus, the n-th moment of Y is given by

$$\mathrm{E}[Y^n] = \sum_{k=0}^{n}\frac{n!}{(n-k)!k!}\sigma^k\mu^{n-k}\mathrm{E}[X^k], \quad \text{for } n = 1,2,3\ldots$$

3 Main results

The derivation of formulas for the moments of the TSMSN distributions can require lengthy calculations. Instead, we propose a general recursive formula and then we get closed form expressions for the moments. Kim (2008, Lemma 2.3) obtained the moments of the TN distribution, a result that was used by Genç (2013) to derive the moments of the truncated scale mixtures of normal (TSMN) distributions. In deriving the moments of the TSMSN, we follow the same strategy used by these authors.

In the following lemma, we present a recursive expression for the moments of the truncated skew-normal distribution. Although, this result was obtained before by Flecher, Allard and Naveau (2010, Proposition 1), here we provide a new proof.

Lemma 1. *If* $X \sim \mathrm{TSN}_{\lfloor a,b \rfloor}(0,1,\lambda)$, *then*

$$(k+1)\mathrm{E}[X^k] - \mathrm{E}[X^{k+2}] = \frac{1}{B(\lambda)}\Big\{ b^{k+1}\phi_{\mathrm{SN}}(b|\lambda) - a^{k+1}\phi_{\mathrm{SN}}(a|\lambda) - \left(\frac{2}{\pi}\right)^{1/2} \frac{\lambda A(\lambda)}{(1+\lambda^2)^{\frac{k+2}{2}}}\mathrm{E}[Z^{k+1}]\Big\}$$

for $k=-1,0,1,2,\ldots$, *where* $Z \sim \mathrm{TN}_{\lfloor a_\lambda, b_\lambda \rfloor}(0,1)$, $a_\lambda = a(1+\lambda^2)^{1/2}$, $b_\lambda = b(1+\lambda^2)^{1/2}$, $A(\lambda) = \Phi(b_\lambda) - \Phi(a_\lambda)$ *and* $B(\lambda) = \Phi_{\mathrm{SN}}(b|\lambda) - \Phi_{\mathrm{SN}}(a|\lambda)$.

Proof. First note that for $k=-1,0,1,2,\ldots$

$$\frac{d(x^{k+1}\phi_{\mathrm{SN}}(x|\lambda))}{dx} = (k+1)x^k\phi_{\mathrm{SN}}(x|\lambda) - x^{k+2}\phi_{\mathrm{SN}}(x|\lambda) + x^{k+1}\lambda\left(\frac{2}{\pi}\right)^{1/2}\phi\big(x(1+\lambda^2)^{1/2}\big).$$

Then,

$$\begin{aligned} &B(\lambda)\{\mathrm{E}[(k+1)X^k] - \mathrm{E}[X^{k+2}]\} \\ &= \int_a^b \{(k+1)x^k - x^{k+2}\}\phi_{\mathrm{SN}}(x|\lambda)\,dx \\ &\quad + \int_a^b x^{k+1}\lambda\left(\frac{2}{\pi}\right)^{1/2}\phi\big(x(1+\lambda^2)^{1/2}\big)\,dx \\ &\quad - \int_a^b x^{k+1}\lambda\left(\frac{2}{\pi}\right)^{1/2}\phi\big(x(1+\lambda^2)^{1/2}\big)\,dx \\ &= b^{k+1}\phi_{\mathrm{SN}}(b|\lambda) - a^{k+1}\phi_{\mathrm{SN}}(a|\lambda) \\ &\quad - \lambda\left(\frac{2}{\pi}\right)^{1/2}\int_a^b x^{k+1}\phi\big(x(1+\lambda^2)^{1/2}\big)\,dx \\ &= b^{k+1}\phi_{\mathrm{SN}}(b|\lambda) - a^{k+1}\phi_{\mathrm{SN}}(a|\lambda) \\ &\quad - \left(\frac{2}{\pi}\right)^{1/2}\frac{\lambda A(\lambda)}{(1+\lambda^2)^{\frac{k+2}{2}}}\int_{a_\lambda}^{b_\lambda}\frac{1}{A(\lambda)}z^{k+1}\phi(z)\,dz. \end{aligned}$$

□

Now we establish the following theorem, which is crucial to the deve-lopment of our proposed theory. This theorem states that the moments of a TSMSN distribution can be computed recursively. It generalizes the results obtained by Kim (2008), Flecher, Allard and Naveau (2010), Genç (2013) and Garay et al. (2017).

Theorem 1. *Let* $Y \sim \mathrm{SMSN}(0,1,\lambda,\boldsymbol{\nu})$. *Then, for* $a<b$, *we have*

$$E[Y^{k+2}|Y \in \lfloor a,b \rfloor] = \tau(a,b) \times \mathrm{E}[U^{-\frac{k+2}{2}}R_{k+2}],$$

where

$$\begin{aligned} \tau(a,b) &= \{\Phi_{\mathrm{SMSN}}(b|\lambda,\boldsymbol{\nu}) - \Phi_{\mathrm{SMSN}}(a|\lambda,\boldsymbol{\nu})\}^{-1}, \\ \mathrm{R}_{k+2} &= \{\Phi_{\mathrm{SN}}(bU^{1/2}|\lambda) - \Phi_{\mathrm{SN}}(aU^{1/2}|\lambda)\}\mathrm{E}[X^{k+2}|U], \\ &\quad \textit{for } k=-1,0,1,2,\ldots, \\ X|U=u &\sim \mathrm{TSN}_{\lfloor au^{1/2}, bu^{1/2} \rfloor}(0,1,\lambda), \quad U \sim H(\cdot|\boldsymbol{\nu}). \end{aligned} \tag{3.1}$$

Proof. We have that $Y|Y \in \lfloor a, b\rfloor \sim \mathrm{TSMSN}_{\lfloor a,b\rfloor}(0,1,\lambda,\boldsymbol{\nu})$. Also, from Definition 1, we have that $Y|U = u \sim \mathrm{SN}(0, u^{-1}, \lambda)$, which implies $Y|U = u, Y \in \lfloor a, b\rfloor \sim \mathrm{TSN}_{\lfloor a,b\rfloor}(0, u^{-1}, \lambda)$. Using the same notation of Definition 1, that is, $Y = U^{-1/2}Z$ with $U \sim H(\cdot|\boldsymbol{\nu})$ and $Z \sim \mathrm{SN}(0,1,\lambda)$ independent, we have:

$$\begin{aligned}
&\mathrm{E}\big[Y^{k+2}|Y \in \lfloor a,b\rfloor\big]\\
&\quad= \mathrm{E}\big[U^{-\frac{k+2}{2}}Z^{k+2}|Y \in \lfloor a,b\rfloor\big]\\
&\quad= \mathrm{E}\big\{\mathrm{E}\big[U^{-\frac{k+2}{2}}Z^{k+2}|U,\big(Y \in \lfloor a,b\rfloor\big)\big]|Y \in \lfloor a,b\rfloor\big\} &(3.2)\\
&\quad= \mathrm{E}\big\{U^{-\frac{k+2}{2}}\mathrm{E}\big[Z^{k+2}|Z \in \lfloor aU^{1/2}, bU^{1/2}\rfloor\big]|Y \in \lfloor a,b\rfloor\big\} &(3.3)\\
&\quad= \int_0^\infty u^{-\frac{k+2}{2}}\mathrm{E}\big[Z^{k+2}|Z \in \lfloor au^{1/2}, bu^{1/2}\rfloor\big]f\big(u|Y \in \lfloor a,b\rfloor\big)\,du, &(3.4)
\end{aligned}$$

where (3.2) is due to basic properties of conditional expectation and in (3.3) we used the independence between U and Z. The pdf $f(u|Y \in \lfloor a,b\rfloor)$ in the integral sign takes the following form:

$$\begin{aligned}
f\big(u|Y \in \lfloor a,b\rfloor\big) &= \int f\big(u|Y = y, Y \in \lfloor a,b\rfloor\big)f\big(y|Y \in \lfloor a,b\rfloor\big)\,dy\\
&= \tau(a,b)\int f\big(u|Y = y, Y \in \lfloor a,b\rfloor\big)f(y)\mathbb{I}_{\lfloor a,b\rfloor}(y)\,dy &(3.5)\\
&= \tau(a,b)\int f(u,y)\mathbb{I}_{\lfloor a,b\rfloor}(y)\,dy &(3.6)\\
&= \tau(a,b)\int_a^b f(u)\phi_{\mathrm{SN}}\big(x|0,u^{-1},\lambda\big)\,dx &(3.7)\\
&= \tau(a,b)f(u)\big\{\Phi_{\mathrm{SN}}\big(bu^{1/2}|\lambda\big) - \Phi_{\mathrm{SN}}\big(au^{1/2}|\lambda\big)\big\}. &(3.8)
\end{aligned}$$

Equation (3.6) is consequence of the fact that, if $y \in \lfloor a,b\rfloor$, then $\{Y \in \lfloor a,b\rfloor, Y = y\} = \{Y = y\}$, implying that $f(u,y) = f(u|Y = y)f(y) = f(u|Y = y, Y \in \lfloor a,b\rfloor)f(y)$. If $y \notin \lfloor a,b\rfloor$ then $\mathbb{I}_{\lfloor a,b\rfloor}(y) = 0$ and the integrals in (3.5) and (3.6) are equal to zero. Equation (3.7) is consequence of $Y|U = u \sim \mathrm{SN}(0, u^{-1}, \lambda)$. Thus, by (3.4) and (3.8), we have:

$$\begin{aligned}
\mathrm{E}\big[Y^{k+2}|Y \in \lfloor a,b\rfloor\big] &= \int_0^\infty u^{-\frac{k+2}{2}}\mathrm{E}\big[Z^{k+2}|Z \in \lfloor au^{1/2}, bu^{1/2}\rfloor\big]\\
&\qquad\times \tau(a,b)f(u)\big\{\Phi_{\mathrm{SN}}\big(bu^{1/2}|\lambda\big) - \Phi_{\mathrm{SN}}\big(au^{1/2}|\lambda\big)\big\}\,du\\
&= \tau(a,b)\times \mathrm{E}\big[U^{-\frac{k+2}{2}}R_{k+2}\big].
\end{aligned}$$

□

Observe that the conditional expectation $\mathrm{E}[X^{k+2}|U]$, where $X|U = u \sim \mathrm{TSN}_{\lfloor au^{1/2}, bu^{1/2}\rfloor}(0,1,\lambda)$, can be computed using Lemma 1. These computations involve conditional expectations of the $\mathrm{TN}_{\lfloor a_\lambda u^{1/2}, b_\lambda u^{1/2}\rfloor}(0,1)$ distribution, which can be computed using Lemma 2.3 in Kim (2008).

As an important remark, observe that the case $\lambda = 0$ in Theorem 1 corresponds to the moments of the $\mathrm{TSMN}_{\lfloor a,b\rfloor}(0,1,\boldsymbol{\nu})$ distribution, which were obtained before by Genç (2013, Theorem 1).

In general, the first four moments are most useful. Thus, we have the following corollary.

Corollary 1. *Let* $X \sim \mathrm{SMSN}(0,1,\lambda,\boldsymbol{\nu})$ *and* $\mu_i \equiv E[X^i|X \in \lfloor a,b\rfloor]$, *with* $a < b$. *Then,*

$$\begin{aligned}
\mu_1 = \tau(a,b)\big[&L(1)\big\{\mathcal{E}_\Phi(-0.5, b_\lambda) - \mathcal{E}_\Phi(-0.5, a_\lambda)\big\}\\
&- \big\{\mathcal{E}_{\phi_{\mathrm{SN}}}(-0.5, b) - \mathcal{E}_{\phi_{\mathrm{SN}}}(-0.5, a)\big\}\big],
\end{aligned}$$

$$\begin{aligned}
\mu_2 = \tau(a,b)\big[&\{\mathcal{E}_{\Phi_{\mathrm{SN}}}(-1,b) - \mathcal{E}_{\Phi_{\mathrm{SN}}}(-1,a)\}\\
&- L(2)\{E_{\phi}(-1,b_\lambda) - \mathcal{E}_{\phi}(-1,a_\lambda)\}\\
&- \{b\mathcal{E}_{\phi_{\mathrm{SN}}}(-0.5,b) - a\mathcal{E}_{\phi_{\mathrm{SN}}}(-0.5,a)\}\big],\\
\mu_3 = \tau(a,b)\big[&2L(1)\{\mathcal{E}_{\Phi}(-1.5,b_\lambda) - \mathcal{E}_{\Phi}(-1.5,a_\lambda)\}\\
&- 2\{\mathcal{E}_{\phi_{\mathrm{SN}}}(-1.5,b) - \mathcal{E}_{\phi_{\mathrm{SN}}}(-1.5,a)\}\\
&- \{b^2\mathcal{E}_{\phi_{\mathrm{SN}}}(-0.5,b) - a^2\mathcal{E}_{\phi_{\mathrm{SN}}}(-0.5,a)\}\\
&+ L(3)\{\mathcal{E}_{\Phi}(-1.5,b_\lambda) - \mathcal{E}_{\Phi}(-1.5,a_\lambda)\}\\
&- L(2)\{b\mathcal{E}_{\phi}(-1,b_\lambda) - a\mathcal{E}_{\phi}(-1,a_\lambda)\}\big],\\
\mu_4 = \tau(a,b)\big[&3\{\mathcal{E}_{\Phi_{\mathrm{SN}}}(-2,b) - \mathcal{E}_{\Phi_{\mathrm{SN}}}(-2,a)\}\\
&- 3L(2)\{\mathcal{E}_{\phi}(-2,b_\lambda) - \mathcal{E}_{\phi}(-2,a_\lambda)\}\\
&- 3\{b\mathcal{E}_{\phi_{\mathrm{SN}}}(-1.5,b) - a\mathcal{E}_{\phi_{\mathrm{SN}}}(-1.5,a)\}\\
&- \{b^3\mathcal{E}_{\phi_{\mathrm{SN}}}(-0.5,b) - a^3\mathcal{E}_{\phi_{\mathrm{SN}}}(-0.5,a)\}\big]\\
&- \tau(a,b)L(4)\big[2\{\mathcal{E}_{\phi}(-2,b_\lambda) - \mathcal{E}_{\phi}(-2,a_\lambda)\}\\
&+ \{(b_\lambda)^2\mathcal{E}_{\phi}(-1,b_\lambda) - (a_\lambda)^2\mathcal{E}_{\phi}(-1,a_\lambda)\}\big],
\end{aligned}$$

where a_λ, b_λ and $\tau(a,b)$ are defined in Lemma 1 and (3.1), $L(s) = (2/\pi)^{\frac{1}{2}}\lambda/(1+\lambda^2)^{\frac{s}{2}}$, $\mathcal{E}_{\phi_{\mathrm{SN}}}(r,q) = \mathrm{E}[U^r\phi_{\mathrm{SN}}(qU^{1/2}|\lambda)]$, $\mathcal{E}_{\Phi_{\mathrm{SN}}}(r,q) = \mathrm{E}[U^r\Phi_{\mathrm{SN}}(qU^{1/2}|\lambda)]$, $\mathcal{E}_{\phi}(r,q)$ and $\mathcal{E}_{\Phi}(r,q)$ are defined like $\mathcal{E}_{\phi_{\mathrm{SN}}}(r,q)$ and $\mathcal{E}_{\Phi_{\mathrm{SN}}}(r,q)$, with $\phi(\cdot)$ and $\Phi(\cdot)$ replacing $\phi_{\mathrm{SN}}(\cdot)$ and $\Phi_{\mathrm{SN}}(\cdot)$, respectively. These expected values can be computed by direct integration when the distribution of U is available.

The expressions in Corollary 1 are useful, for example, to compute some distribution measures, like the skewness (S), kurtosis (K) and coefficient of variation (CV), given by $S = (\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3)/(\mu_2 - \mu_1^2)^{3/2}$, $K = (\mu_4 - 4\mu_1\mu_3 + 6\mu_2\mu_1^2 - 3\mu_1^4)/(\mu_2 - \mu_1^2)^2$ and $\mathrm{CV} = (\mu_2 - \mu_1^2)^{1/2}/\mu_1$.

4 Particular cases of SMSN distributions

The class of SMSN distributions includes the skew-t, skew-slash and skew-contaminated normal. All these distributions have heavier tails than the skew-normal and can be used for robust inference. Some of these distributions are described subsequently. For each one, we compute the expected values $\mathcal{E}_{\phi_{\mathrm{SN}}}(r,q)$, $\mathcal{E}_{\Phi_{\mathrm{SN}}}(r,q)$, $\mathcal{E}_{\phi}(r,q)$ and $\mathcal{E}_{\Phi}(r,q)$. For the sake of completeness, detailed proofs of these results are given in the Appendix.

4.1 The skew-t distribution

In this case, we consider $U \sim \mathrm{Gamma}(\nu/2, \nu/2)$, $\nu > 0$, in Definition 1. The density of Y takes the form (Azzalini and Capitanio, 2003)

$$\phi_{\mathrm{ST}}(y|\mu,\sigma^2,\lambda,\nu) = \frac{2\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{1/2}\sigma}\left(1+\frac{d}{\nu}\right)^{-\frac{\nu+1}{2}}\mathbf{T}\left(\left(\frac{\nu+1}{d+\nu}\right)^{1/2}A\Big|\nu+1\right),$$

where $A = \lambda(y-\mu)/\sigma$, $d = (y-\mu)^2/\sigma^2$ and $\mathbf{T}(\cdot|\nu)$ denotes the cdf of the standard Student–t distribution, with location zero, scale one and ν degrees of freedom ($t(0,1,\nu)$). We use the notation $Y \sim \mathrm{ST}(\mu, \sigma^2, \lambda, \nu)$. A particular case of the skew-t distribution is the skew–Cauchy distribution, when $\nu = 1$. Also, when $\nu \to \infty$, we get the skew-normal distribution as the limiting case.

From (2.4), we obtain the following expression for the cdf of a standard skew-t random variable:

$$\begin{aligned}\Phi_{\mathrm{ST}}(y|\lambda,\nu) &= 2\mathrm{E}\big[\boldsymbol{\Phi}_2(U^{1/2}y\mathbf{e}_1)|\mathbf{0},\boldsymbol{\Sigma}\big] = 2\mathrm{E}\big[P\big(\mathbf{X} \le U^{1/2}y\mathbf{e}_1|U\big)\big] \\ &= 2P\left(\frac{\mathbf{X}}{U^{1/2}} \le y\mathbf{e}_1\right) = 2\mathbf{T}_2(y\mathbf{e}_1|\mathbf{0},\boldsymbol{\Sigma},\nu),\end{aligned}$$

where $\mathbf{e}_1 = (1,0)^\top$, $\mathbf{X} \sim \mathbf{N}_2(\mathbf{0}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ is given in (2.2) and $\mathbf{T}_2(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ denotes the cdf of the bivariate Student-t distribution with mean vector $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$ and ν degrees of freedom. For $\mathbf{X} = (X_1, X_2)^\top$ and $\mathbf{x} = (x_1, x_2)^\top$, $\mathbf{X} \le \mathbf{x}$ is interpreted element wise, that is, $X_i \le x_i$, $i = 1, 2$. This expression can also be obtained as a special case of the one given for the multivariate skew-t distribution in Azzalini and Capitanio (2003, Section 4.2.1).

For the truncated skew-t distribution, we have the following result for the expected values involved in the calculation of the moments.

Corollary 2. *Let* $X \sim \mathrm{TST}_{\lfloor a,b \rfloor}(0, 1, \lambda, \nu)$. *Then*

$$\begin{aligned}\mathcal{E}_{\phi_{SN}}(r,q) &= \frac{2^{r+1}\nu^{\nu/2}\Gamma(\frac{\nu+2r}{2})}{\sqrt{2\pi}\,\Gamma(\frac{\nu}{2})(q^2+\nu)^{\frac{\nu+2r}{2}}}\mathbf{T}\left(\left(\frac{2r+\nu}{q^2+\nu}\right)^{1/2}\lambda q \,\middle|\, 2r+\nu\right);\\ \mathcal{E}_{\Phi_{SN}}(r,q) &= \frac{2^{r+1}\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})\nu^r}\mathbf{T}_2\left(\left(\frac{2r+\nu}{\nu}\right)^{1/2} q\mathbf{e}_1 \,\middle|\, \mathbf{0},\boldsymbol{\Sigma},2r+\nu\right);\\ \mathcal{E}_{\Phi}(r,q) &= \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})}\left(\frac{2}{\nu}\right)^r \mathbf{T}\left(\left(\frac{2r+\nu}{\nu}\right)^{1/2} q \,\middle|\, 2r+\nu\right);\\ \mathcal{E}_{\phi}(r,q) &= \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})\sqrt{2\pi}}\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}\left(\frac{q^2+\nu}{2}\right)^{-\frac{(\nu+2r)}{2}},\end{aligned}$$

where $\Gamma(a)$ *is the gamma function, and* $\boldsymbol{\Sigma}$ *is given in* (2.2).

4.2 The skew-slash distribution

In this case, we have $U \sim \mathrm{Beta}(\nu, 1)$—where $\mathrm{Beta}(a, b)$ denotes the beta distribution with parameters a and b—with $\nu > 0$, and we use the notation $Y \sim \mathrm{SSL}(\mu, \sigma^2, \lambda, \nu)$. The density of Y is given by:

$$\phi_{\mathrm{SSL}}(y|\mu,\sigma^2,\lambda,\nu) = 2\nu\int_0^1 u^{\nu-1}\phi(y|\mu,u^{-1}\sigma^2)\Phi(u^{1/2}A)\,du.$$

The cdf of the standard skew-slash distribution does not have a closed form expression. However, using (2.4), we can write it in terms of the following integral, which can be obtained by numerical methods:

$$\Phi_{\mathrm{SSL}}(y|\lambda,\nu) = \int_0^\infty 2\nu\Phi_2(u^{1/2}y\mathbf{e}_1|\mathbf{0},\boldsymbol{\Sigma})u^{\nu-1}\,du, \tag{4.1}$$

where $\boldsymbol{\Sigma}$ is given in (2.2). The truncated skew-slash distribution will be denoted by $\mathrm{TSSL}_{\lfloor a,b \rfloor}(\mu, \sigma^2, \lambda, \nu)$.

Corollary 3. *Let* $X \sim \mathrm{TSSL}_{\lfloor a,b \rfloor}(0, 1, \lambda, \nu)$. *Then*

$$\mathcal{E}_{\phi_{\mathrm{SN}}}(r, q) = \frac{2^{\nu+r+1}\nu\Gamma(r+\nu)}{(2\pi)^{1/2}q^{2r+2\nu}} G\left(1\middle|r+\nu, \frac{q^2}{2}\right) E[\Phi(\lambda q(U')^{1/2})];$$

$$\mathcal{E}_{\Phi_{\mathrm{SN}}}(r, q) = \frac{2\nu}{r+\nu} E[\Phi_2((U'')^{1/2} q\mathbf{e}_1|0, \mathbf{\Sigma})];$$

$$\mathcal{E}_{\Phi}(r, q) = \left(\frac{\nu}{\nu+r}\right)\Phi_{\mathrm{SL}}(q|\nu+r);$$

$$\mathcal{E}_{\phi}(r, q) = \frac{\nu}{\sqrt{2\pi}}\left(\frac{q^2}{2}\right)^{-(\nu+r)}\Gamma\left(\nu+r, \frac{q^2}{2}\right),$$

where $\Gamma(a, b) = \int_0^b e^{-t}t^{a-1}\,dt$ *is the incomplete gamma function*, $G(\cdot|\alpha, \beta)$ *represents the cdf of the Gamma distribution with parameters* α *and* β, $U' \sim \mathrm{TGamma}_{\lfloor 0,1 \rfloor}(r+\nu, q^2/2)$ (*a truncated Gamma distribution*), $U'' \sim \mathrm{Beta}(r+\nu, 1)$, $\mathbf{\Sigma}$ *is given in* (2.2) *and* $\Phi_{\mathrm{SL}}(\cdot|\nu+r)$ *is the cdf of the standard slash distribution—that is*, *when* $\lambda = 0$ *in* (4.1).

Observe that $\mathcal{E}_{\phi_{\mathrm{SN}}}(r, q)$ and $\mathcal{E}_{\Phi_{\mathrm{SN}}}(r, q)$ do not have closed form expressions, but the integrals $\mathrm{E}[\Phi(\lambda q(U')^{1/2})]$ and $\mathrm{E}[\Phi_2((U'')^{1/2}q\mathbf{e}_1|0, \mathbf{\Sigma})]$ can be easily approximated using the R function *integrate*.

4.3 The skew-contaminated normal distribution

This distribution is denoted by $Y \sim \mathrm{SCN}(\mu, \sigma^2, \lambda, (\gamma, \xi))$. Here $U = \xi$ with probability γ and $U = 1$ with probability $1-\gamma$. It follows immediately that its pdf is given by

$$\begin{aligned}&\phi_{\mathrm{SCN}}(y|\mu, \sigma^2, \lambda, (\gamma, \xi))\\ &\quad = 2\{\gamma\phi(y|\mu, \xi^{-1}\sigma^2)\Phi(\xi^{1/2}A) + (1-\gamma)\phi(y|\mu, \sigma^2)\Phi(A)\},\end{aligned}$$

and the cdf in the standard case is

$$\Phi_{\mathrm{SCN}}(y|\lambda, (\gamma, \xi)) = 2\{\gamma\Phi_2(\xi^{1/2}y\mathbf{e}_1|\mathbf{0}, \mathbf{\Sigma}) + (1-\gamma)\Phi_2(y\mathbf{e}_1|\mathbf{0}, \mathbf{\Sigma})\}. \tag{4.2}$$

Denoting the truncated SCN distribution by $\mathrm{TSCN}_{\lfloor a,b \rfloor}(\mu, \sigma^2, \lambda, (\gamma, \xi))$, we get the following expected values.

Corollary 4. *Let* $X \sim \mathrm{TSCN}_{\lfloor a,b \rfloor}(0, 1, \lambda, (\gamma, \xi))$ *Then*

$$\mathcal{E}_{\phi_{\mathrm{SN}}}(r, q) = \gamma\xi^r\phi_{\mathrm{SN}}(q\xi^{\frac{1}{2}}|\lambda) + (1-\gamma)\phi_{\mathrm{SN}}(q|\lambda);$$

$$\mathcal{E}_{\Phi_{\mathrm{SN}}}(r, q) = \xi^r\Phi_{\mathrm{SCN}}(q|\lambda, (\gamma, \xi)) + 2(1-\gamma)(1-\xi^r)\Phi_2(q\mathbf{e}_1|\mathbf{0}, \mathbf{\Sigma});$$

$$\mathcal{E}_{\Phi}(r, q) = \xi^r\Phi_{\mathrm{CN}}(q|(\gamma, \xi)) + (1-\gamma)(1-\xi^r)\Phi(q);$$

$$\mathcal{E}_{\phi}(r, q) = \gamma\xi^r\phi(\sqrt{\xi}q) + (1-\gamma)\phi(q),$$

where $\Phi_{\mathrm{CN}}(\cdot|(\gamma, \xi))$ *represents the cdf of the standard contaminated normal distribution—that is*, *when* $\lambda = 0$ *in* (4.2).

5 Statistical applications

5.1 The SMSN censored linear regression model

In recent years, there has been wide concern to find more flexible parametric families of non-normal distributions for robust statistical modeling of linear regression models, when the data collected are subject to some upper and lower detection limits, that is, the responses are either left or right censored. For instance, Garay et al. (2017) recently established a new link between the censored regression model and the symmetric class of SMN distributions, which extends the normal one by the inclusion of kurtosis. An interesting extension is to consider the asymmetrical class of SMSN distributions, which allows capturing skewness and kurtosis in data simultaneously. Thus, in the following we define the censored linear regression model under scale mixtures of skew-normal distributions, denoted the SMSN-CR model, and some properties of this proposed model are derived by using the results presented in this work. Inferential procedures can be also easily implemented.

Description of the model. Consider a linear regression model where the responses are observed with errors which are independent and identically distributed (i.i.d.) according to some SMSN distribution, as follows:

$$Y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \sigma\varepsilon_i, \quad \varepsilon_i \overset{\text{iid}}{\sim} \text{SMSN}(0, 1, \lambda, \boldsymbol{\nu}), i = 1, \dots, n, \tag{5.1}$$

where the Y_i are responses, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ is a vector of regression parameters and $\mathbf{x}_i^\top = (x_{i1}, \dots, x_{ip})$ is a vector such that x_{ij} is the value of the j-th explanatory variable for subject i. With this structure, we have that $Y_i \sim \text{SMSN}(\mathbf{x}_i^\top \boldsymbol{\beta}, \sigma^2, \lambda, \boldsymbol{\nu})$. In this application, we are interested in the case where left-censored observations can occur. That is, the observations are of the form:

$$Y_{\text{obs}_i} = \begin{cases} \kappa_i & \text{if } Y_i \le \kappa_i; \\ Y_i & \text{if } Y_i > \kappa_i, \end{cases} \tag{5.2}$$

$i = 1, \dots, n$, for some threshold point κ_i. The model defined in (5.1) and (5.2) is called the SMSN-CR model. See Massuia et al. (2017) for further details.

The mean and variance of the SMSN–CR model. Let us define the binary random variable $D_i = 1$ if $Y_i > \kappa_i$ and $D_i = 0$ otherwise. Then the mean and variance of the SMSN–CR model for the i-th observed response are given by:

$$\begin{aligned} \text{E}[Y_{\text{obs}_i}] &= \text{E}\big[\kappa_i(1 - D_i) + Y_i D_i\big] \\ &= \kappa_i \Phi_{\text{SMSN}}\left(\frac{\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\Big|\lambda, \boldsymbol{\nu}\right) + \text{E}[Y_i D_i], \\ \text{Var}[Y_{\text{obs}_i}] &= \text{Var}\big[\kappa_i(1 - D_i) + Y_i D_i\big] \\ &= \kappa_i^2 \Phi_{\text{SMSN}}\left(\frac{\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\Big|\lambda, \boldsymbol{\nu}\right)\left\{1 - \Phi_{\text{SMSN}}\left(\frac{\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\Big|\lambda, \boldsymbol{\nu}\right)\right\} \\ &\quad + \text{Var}[Y_i D_i]. \end{aligned} \tag{5.3}$$

Defining $\kappa_i^* = (\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta})/\sigma$, we get

$$\begin{aligned} \text{E}[Y_i D_i] &= \text{E}\big[\text{E}(Y_i D_i | D_i)\big] = \text{E}[Y_i | Y_i > \kappa_i] P(Y_i > \kappa_i) \\ &= \{\mathbf{x}_i^\top \boldsymbol{\beta} + \sigma \text{E}[\varepsilon_i | \varepsilon_i > \kappa_i^*]\}\left\{1 - \Phi_{\text{SMSN}}\left(\frac{\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\Big|\lambda, \boldsymbol{\nu}\right)\right\}. \end{aligned} \tag{5.4}$$

Observe that the conditional expectation $\text{E}[\varepsilon_i | \varepsilon_i > \kappa_i^*]$ is the first moment of the standard TSMSN distribution, which can be easily obtained using Theorem 1 along with Corollary 1.

The variance in (5.3) can be obtained as follows:

$$
\begin{aligned}
\mathrm{Var}[Y_i D_i] &= \mathrm{E}\big[\mathrm{Var}(Y_i D_i | D_i)\big] + \mathrm{Var}\big[\mathrm{E}(Y_i D_i | D_i)\big] \\
&= \mathrm{Var}[Y_i | Y_i > \kappa_i]\bigg\{1 - \Phi_{\mathrm{SMSN}}\bigg(\frac{\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\Big|\lambda, \boldsymbol{\nu}\bigg)\bigg\} \\
&\quad + \big(\mathrm{E}[Y_i | Y_i > \kappa_i]\big)^2\bigg\{1 - \Phi_{\mathrm{SMSN}}\bigg(\frac{\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\Big|\lambda, \boldsymbol{\nu}\bigg)\bigg\} \\
&\quad \times \Phi_{\mathrm{SMSN}}\bigg(\frac{\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\Big|\lambda, \boldsymbol{\nu}\bigg) \\
&= \sigma^2\big\{\mathrm{E}[\varepsilon_i^2 | \varepsilon_i > \kappa_i^*] - \big(\mathrm{E}[\varepsilon_i | \varepsilon_i > \kappa_i^*]\big)^2\big\} \\
&\quad \times \bigg\{1 - \Phi_{\mathrm{SMSN}}\bigg(\frac{\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\Big|\lambda, \boldsymbol{\nu}\bigg)\bigg\} \\
&\quad + \big(\mathbf{x}_i^\top \boldsymbol{\beta} + \sigma \mathrm{E}[\varepsilon_i | \varepsilon_i > \kappa_i^*]\big)^2\bigg\{1 - \Phi_{\mathrm{SMSN}}\bigg(\frac{\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\Big|\lambda, \boldsymbol{\nu}\bigg)\bigg\} \\
&\quad \times \Phi_{\mathrm{SMSN}}\bigg(\frac{\kappa_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\Big|\lambda, \boldsymbol{\nu}\bigg),
\end{aligned} \tag{5.5}
$$

where the conditional expectation $\mathrm{E}[\varepsilon_i^2 | \varepsilon_i > \kappa_i^*]$ is the second moment of the TSMSN distribution, which can be obtained using Theorem 1 along with Corollary 1.

It is important to note that, the results obtained in (5.4) and (5.5) will help to implement the recursive methods to obtain the maximum likelihood estimation of the SMSN-CR model, as for example, the EM algorithm (Dempster, Laird and Rubin, 1977) or some extension like the ECM or the ECME algorithm (Liu and Rubin, 1994). An in-depth investigation of these algorithms and their extensions is beyond the scope of the present paper, but it is an interesting topic for further research.

5.2 Tail conditional expectation (TCE)

Significant changes in the financial markets are giving increasing attention to the need for developing a standard framework for risk measurement. Thus, in the actuarial science, there has been a growing interest among investment experts to focus on the use of a Tail Conditional Expectation (TCE), because it shares properties that are considered desirable and applicable in a variety of situations, see Artzner et al. (1999) and Landsman and Valdez (2003) for more details. The TCE is defined by

$$
\mathrm{TCE}_X(x_q) = \mathrm{E}[X | X > x_q], \tag{5.6}
$$

where x_q denotes the quantile of order q of the distribution of X. This measure is interpreted as the expected worst possible loss, given the loss will exceed a particular value x_q. In particular, this threshold value is called value-at-risk, or simply VaR, which properties were studied and developed by Artzner et al. (1999).

Some authors computed the TCE using different distributions for X, like Landsman and Valdez (2003) who derived explicit formulas assuming that the distribution of X is elliptical and Genç (2013), who developed a recursive formula to estimate the TCE considering the Student-t distribution. Here we estimate the TCE assuming that X has a skew-t distribution, using the recursive formulas developed in Section 3. We consider a dataset consisting of the total damage done by 35 hurricanes (hurricanes data) between the years 1949 and 1980, which was considered before by Hogg and Klugman (1984) and Genç (2013).

Table 1 *Hurricanes data. Estimates of* $\mathrm{TCE}_X(x_q)$ *for various quantiles*

q	x_q	$\widehat{\mathrm{TCE}}_X(x_q)$
0.500	143,610.3	321,988.1
0.750	267,691.1	441,102.6
0.900	402,772.3	615,083.6
0.950	514,784.2	779,157.4
0.975	648,743.1	986,485.2
0.990	879,055.3	1,355,296.3
0.999	1,958,192.6	3,130,346.0

As mentioned by Genç (2013), the last loss value is far away from the rest of the data and so it is a possible discording data. Thus, in order to estimate the TCE measure under the skew-t distribution for the hurricanes data, we propose the following procedure:

Step 1: Fit the skew-t distribution using the maximum likelihood method to obtain the estimates of μ, σ^2, λ, and ν. In this case, we used direct maximization of the log-likelihood through the bult-in function *optim* of R software (R Core Team, 2018), obtaining the following estimates: $\widehat{\mu} = 204{,}900.4$, $\widehat{\sigma} = 174{,}287.6$, $\widehat{\lambda} = -0.4399$ and $\widehat{\nu} = 2.511131$.

Step 2: Find the x_q values using the function *qst* of the R package *sn* (Azzalini, 2018).

Step 3: Estimate the TCE measure using the results presented in Corollary 1 and 2. These values, for several values of q, are presented in Table 1.

6 Simulation study

In this section, we present a simulation study to compare the theoretical moments of the TSMSN distributions along with the empirical moments computed via a Monte Carlo approximation. Thus, we generated 300 artificial samples of size 1000 from $Y \sim \mathrm{TSMSN}_{\lfloor a,b \rfloor}(\mu, \sigma^2, \lambda, \boldsymbol{\nu})$, using the sampling/importance resampling method proposed by Rubin (1987) and Rubin et al. (1988), as well as the stochastic representation of a SMSN random variable given in Definition 1. We adopted the bilateral truncation, with truncation limits $\lfloor a, b \rfloor = \lfloor 3, 10 \rfloor$. The true parameter values were taken as $\mu = 2$ for the the location parameter, $\sigma^2 = 10$ for the scale parameter and five values were adopted for the shape parameter $\lambda = \{\pm 3, \pm 1, 0\}$, corresponding to high and low (negative and positive) levels of skewness and also the symmetric case. For the degrees of freedom, we considered $\nu = 5$ for the TST and TSSL models and $\boldsymbol{\nu}^\top = (0.5, 0.5)^\top$ for the TSCN model.

Table 2, shows the comparison between the first four theoretical moments $\mathrm{E}[Y^k] = \mu^k$, $(k = 1, 2, 3, 4)$, computed using Theorem 1 and the values of the average empirical moments, across 300 replicates (in parentheses), of the TSMSN distributions, considering different values of the skewness parameter λ. We observe that, in general, the values of the first four (theoretical and empirical) moments are very close, in all the models at all levels of skewness, indicating that the proposed recursive formula to obtain the moments of the TSMSN distributions is reliable.

7 Concluding remarks and discussion

In this paper, we have developed exact expressions for the moments of the family of truncated scale mixtures of skew-normal (TSMSN) distributions, generalizing results obtained by Kim (2008), Flecher, Allard and Naveau (2010), Genç (2013) and Garay et al. (2017). For

Table 2 *Simulation study. Comparison between the theoretical and empirical moments of the TSMSN distributions, considering different values of the skewness parameter* λ

Skewness Parameter λ	Models	Theoretical (empirical) moments			
		μ_1	μ_2	μ_3	μ_4
-3	TSN	3.45374 (3.45340)	12.08993 (12.08742)	42.96560 (42.95163)	155.29388 (155.22626)
	TST	3.71294 (3.71247)	14.37196 (14.36854)	58.79881 (58.78245)	258.66044 (258.63532)
	TSSL	3.55715 (3.52059)	12.85195 (12.61450)	48.06670 (46.12137)	182.62574 (172.58123)
	TSCN	3.61100 (3.57975)	13.35080 (13.10054)	50.73327 (49.19340)	198.95424 (190.31305)
-1	TSN	4.25011 (4.25088)	19.17883 (19.18708)	92.49161 (92.56216)	478.51429 (479.10284)
	TST	4.52275 (4.52223)	22.33490 (22.33061)	121.53582 (121.51143)	729.25127 (729.19451)
	TSSL	4.41090 (4.40489)	20.89430 (20.81780)	106.90463 (106.28329)	592.49416 (587.70365)
	TSCN	4.55260 (4.52648)	22.47088 (22.18035)	121.01840 (118.45615)	711.21899 (690.04365)
0	TSN	5.10261 (5.10326)	28.54072 (28.54647)	174.41678 (174.45214)	1153.67033 (1153.88341)
	TST	5.24018 (5.23949)	30.38067 (30.37541)	193.73722 (193.71099)	1341.23141 (1341.22234)
	TSSL	5.27986 (5.27514)	30.70013 (30.63898)	195.26771 (194.69099)	1342.11148 (1337.03559)
	TSCN	5.39429 (5.39269)	32.13358 (32.10981)	209.58456 (209.31021)	1475.64220 (1472.82651)
1	TSN	5.30363 (5.30444)	30.74829 (30.75607)	193.73509 (193.78855)	1312.87505 (1313.22234)
	TST	5.42378 (5.42573)	32.43965 (32.46705)	212.21422 (212.49948)	1497.84276 (1500.56343)
	TSSL	5.49669 (5.49749)	33.13645 (33.15480)	217.22821 (217.40864)	1528.66830 (1530.26625)
	TSCN	5.61479 (5.61648)	34.66502 (34.68217)	232.78708 (232.91258)	1675.90499 (1676.69721)
3	TSN	5.15246 (5.15217)	29.03805 (29.03393)	178.39074 (178.34207)	1183.85266 (1183.36983)
	TST	5.29975 (5.30043)	31.00508 (31.01469)	199.00043 (199.09597)	1383.45657 (1384.33049)
	TSSL	5.34182 (5.34392)	31.34259 (31.35065)	200.58899 (200.58274)	1384.02834 (1383.26223)
	TSCN	5.46705 (5.46743)	32.90002 (32.90361)	216.06651 (216.09109)	1527.73772 (1527.93369)

a reader interested in real-world applications, we show the practicability of our results with a simulation study as well as two applications, the first one is the computation of the mean and variance of a censored regression model based on SMSN distributions and the other one using a real data in the field of actuarial sciences.

We conjecture that our method can be extended to the context of multivariate truncated SMSN distributions, as discussed in Ho et al. (2012). An in-depth investigation of such extension is beyond the scope of the present paper, but it is an interesting topic for further research. Finally, the proposed method has been coded and implemented in the R software (R Core Team, 2018), which is available from us upon request.

Appendix. Computation of $\mathcal{E}_{\phi_{\mathrm{SN}}}(r,q)$, $\mathcal{E}_{\Phi_{\mathrm{SN}}}(r,q)$, $\mathcal{E}_{\phi}(r,q)$ and $\mathcal{E}_{\Phi}(r,q)$ for some TSMSN distributions

In this appendix, we obtain expressions for the expected values $\mathcal{E}_{\phi_{\mathrm{SN}}}(r,q)$, $\mathcal{E}_{\Phi_{\mathrm{SN}}}(r,q)$, $\mathcal{E}_{\phi}(r,q)$ and $\mathcal{E}_{\Phi}(r,q)$ given in Section 4, for specific SMSN distributions.

Skew-t distribution

In this case, in Definition 1, we have that $U \sim \mathrm{Gamma}(\nu/2, \nu/2)$, with $\nu > 0$. To facilitate notation, let us make $\alpha_1 = (\nu + 2r)/2$, $\alpha_2 = \nu/2$ and $\alpha_3 = (q^2 + \nu)/2$. Then,

$$
\begin{aligned}
\mathcal{E}_{\Phi_{\mathrm{SN}}}(r,q) &= \mathrm{E}\big[U^r \Phi_{\mathrm{SN}}\big(qU^{1/2}|\lambda\big)\big] \\
&= \int_0^\infty \frac{u^{\frac{2r+\nu}{2}-1}\Phi_{\mathrm{SN}}(qu^{\frac{1}{2}}|\lambda)\nu^{\frac{\nu}{2}}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} \exp\Big\{-\frac{u\nu}{2}\Big\}\,du \\
&= \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})}\Big(\frac{2}{\nu}\Big)^r \int_0^\infty \Phi_{\mathrm{SN}}\big(qu^{\frac{1}{2}}|\lambda\big)\frac{1}{\Gamma(\alpha_1)}\alpha_2^{\alpha_1}u^{\alpha_1-1}\exp\{-u\alpha_2\}\,du \\
&= \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})}\Big(\frac{2}{\nu}\Big)^r \mathrm{E}\big[\Phi_{\mathrm{SN}}\big(qU'^{\frac{1}{2}}|\lambda\big)\big] \\
&= \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})}\Big(\frac{2}{\nu}\Big)^r \mathrm{E}\big[2\Phi_2\big(U'^{\frac{1}{2}}\boldsymbol{y}^*|\mathbf{0},\boldsymbol{\Sigma}\big)\big] \\
&= 2^{r+1}\frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})}\nu^{-r}\mathbf{T}_2\bigg(\sqrt{\frac{2r+\nu}{\nu}}q\mathbf{e}_1\Big|\mathbf{0},\boldsymbol{\Sigma},2r+\nu\bigg),
\end{aligned}
$$

where $U' \sim \mathrm{Gamma}(\alpha_1, \alpha_2)$, $\mathbf{e}_1 = (1,0)^\top$ and $\boldsymbol{\Sigma} = \big(\begin{smallmatrix} 1 & -\delta \\ -\delta & 1 \end{smallmatrix}\big)$.

$$
\begin{aligned}
\mathcal{E}_{\Phi}(r,q) &= \mathrm{E}\big[U^r \Phi\big(qU^{1/2}\big)\big] \\
&= \int_0^\infty \frac{u^{\frac{2r+\nu}{2}-1}\Phi(qu^{\frac{1}{2}})\nu^{\frac{\nu}{2}}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} \exp\Big\{-\frac{u\nu}{2}\Big\}\,du \\
&= \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})}\Big(\frac{2}{\nu}\Big)^r \int_0^\infty \Phi\big(qu^{\frac{1}{2}}\big)\frac{1}{\Gamma(\alpha_1)}\alpha_2^{\alpha_1}u^{\alpha_1-1}\exp\{-u\alpha_2\}\,du \\
&= \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})}\Big(\frac{2}{\nu}\Big)^r \mathrm{E}\big[\Phi\big(qU'^{\frac{1}{2}}\big)\big] \\
&= \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})}\Big(\frac{2}{\nu}\Big)^r \mathbf{T}\bigg(\Big(\frac{2r+\nu}{\nu}\Big)^{1/2}q\Big|2r+\nu\bigg).
\end{aligned}
\tag{A.1}
$$

Equation (A.1) was obtained using Lemma 3 of Genç (2013).

$$
\begin{aligned}
\mathcal{E}_{\phi_{\mathrm{SN}}}(r,q) &= \mathrm{E}\big[U^r \phi_{\mathrm{SN}}\big(qU^{\frac{1}{2}}|\lambda\big)\big] \\
&= \int_0^\infty u^r 2\phi(qu^{1/2})\Phi(\lambda q u^{1/2}) \frac{1}{\Gamma(\frac{\nu}{2})} u^{\frac{\nu}{2}-1}\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}} \exp\left\{-\frac{u\nu}{2}\right\} du \\
&= \frac{2}{\sqrt{2\pi}} \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})}\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}\left(\frac{q^2+\nu}{2}\right)^{-\frac{\nu+2r}{2}} \int_0^\infty \Phi(\lambda q u^{\frac{1}{2}}) \frac{\alpha_3^{\alpha_1}}{\Gamma(\alpha_1)} \\
&\quad \times u^{\{\alpha_1-1\}} \exp\{-\alpha_3 u\}\, du \\
&= \frac{2}{\sqrt{2\pi}} \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})}\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}\left(\frac{q^2+\nu}{2}\right)^{-\frac{\nu+2r}{2}} \mathrm{E}\big[\Phi\big(\lambda q U''^{\frac{1}{2}}\big)\big] \\
&= \frac{2^{r+1}}{\sqrt{2\pi}} \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})} \nu^{\frac{\nu}{2}} (q^2+\nu)^{-\frac{\nu+2r}{2}} \mathbf{T}\left(\sqrt{\frac{2r+\nu}{q^2+\nu}}\lambda q \,\middle|\, 2r+\nu\right),
\end{aligned} \tag{A.2}
$$

where $U'' \sim \mathrm{Gamma}(\alpha_1, \alpha_3)$. Equation (A.2) was obtained using Lemma 3 of Genç (2013).

$$
\begin{aligned}
\mathcal{E}_{\phi}(r,q) = \mathrm{E}\big[U^r\phi\big(qU^{\frac{1}{2}}\big)\big] &= \int_0^\infty \frac{\nu^{\frac{\nu}{2}} u^{\frac{\nu}{2}-1} u^r}{\sqrt{2\pi}\Gamma(\frac{\nu}{2})2^{\frac{\nu}{2}}} \exp\left\{-\frac{u(q^2+\nu)}{2}\right\} du \\
&= \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})} \frac{\nu^{\frac{\nu}{2}}}{\sqrt{2\pi}2^{\frac{\nu}{2}}} \left(\frac{q^2+\nu}{2}\right)^{-\frac{\nu+2r}{2}} \int_0^\infty \frac{\alpha_2^{\alpha_1} u^{\alpha_1-1}}{\Gamma(\alpha_1)} \exp\{-\alpha_2 u\}\, du \\
&= \frac{\Gamma(\frac{\nu+2r}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{\sqrt{2\pi}} \left(\frac{\nu}{2}\right)^{\nu/2} \left(\frac{q^2+\nu}{2}\right)^{-\frac{\nu+2r}{2}},
\end{aligned} \tag{A.3}
$$

where the integrand in (A.3) is the pdf of a random variable with $\mathrm{Gamma}(\alpha_1, \alpha_2)$ distribution.

Skew-slash distribution

In this case, $U \sim \mathrm{Beta}(\nu, 1)$, with positive shape parameter ν. Thus

$$
\begin{aligned}
\mathcal{E}_{\Phi_{\mathrm{SN}}}(r,q) = \mathrm{E}\big[U^r \Phi_{\mathrm{SN}}\big(qU^{1/2}|\lambda\big)\big] &= \int_0^1 u^r \Phi_{\mathrm{SN}}\big(qu^{\frac{1}{2}}|\lambda\big)\nu u^{\nu-1}\, du \\
&= \left(\frac{\nu}{\nu+r}\right) \int_0^1 \Phi_{\mathrm{SN}}\big(qu^{\frac{1}{2}}|\lambda\big)(\nu+r)u^{(\nu+r)-1}\, du \\
&= \left(\frac{\nu}{\nu+r}\right) \mathrm{E}\big[\Phi_{\mathrm{SN}}\big(qU''^{\frac{1}{2}}|\lambda\big)\big] \\
&= \left(\frac{2\nu}{\nu+r}\right) \mathrm{E}\big[\Phi_2\big(U''^{\frac{1}{2}} q\mathbf{e}_1|\mathbf{0}, \boldsymbol{\Sigma}\big)\big],
\end{aligned}
$$

where $U'' \sim \mathrm{Beta}(\nu + r, 1)$.

$$
\begin{aligned}
\mathcal{E}_{\Phi}(r,q) = \mathrm{E}\big[U^r \Phi\big(qU^{1/2}\big)\big] &= \int_0^1 u^r \Phi\big(qu^{\frac{1}{2}}\big)\nu u^{\nu-1}\, du \\
&= \left(\frac{\nu}{\nu+r}\right) \int_0^1 \Phi\big(qu^{\frac{1}{2}}\big)(\nu+r)u^{(\nu+r)-1}\, du \\
&= \left(\frac{\nu}{\nu+r}\right) \mathrm{E}\big[\Phi\big(qU''^{\frac{1}{2}}\big)\big]
\end{aligned}
$$

$$= \left(\frac{\nu}{\nu + r}\right)\Phi_{\mathrm{SL}}(q|\nu + r), \tag{A.4}$$

with $U'' \sim \mathrm{Beta}(r+\nu, 1)$. Using Lemma 3 of Genç (2013) we obtain Equation (A.4).

$$\begin{aligned}
\mathcal{E}_{\phi_{\mathrm{SN}}}(r,q) &= \mathrm{E}\big[U^r \phi_{\mathrm{SN}}(qU^{\frac{1}{2}}|\lambda)\big] = \int_0^1 u^r \phi_{\mathrm{SN}}(qu^{\frac{1}{2}}|\lambda)\nu u^{\nu-1}\,du \\
&= \int_0^1 2\phi(qu^{\frac{1}{2}})\Phi(\lambda q u^{\frac{1}{2}})\nu u^{(r+\nu)-1}\,du \\
&= \left(\frac{2\nu}{\sqrt{2\pi}}\right)\Gamma(\nu+r)\left(\frac{q^2}{2}\right)^{-(\nu+r)} G\left(1\Big|r+\nu, \frac{q^2}{2}\right) \\
&\quad \times \int_0^1 \frac{1}{G(1|\alpha_1,\alpha_2)}\Phi(\lambda q u^{\frac{1}{2}})\frac{1}{\Gamma(\alpha_1)}(\alpha_2)^{\alpha_1} u^{\alpha_1 - 1}\exp\{-\alpha_2 u\}\,du \\
&= \left(\frac{2\nu}{\sqrt{2\pi}}\right)\Gamma(\nu+r)\left(\frac{q^2}{2}\right)^{-(\nu+r)} G\left(1\Big|r+\nu, \frac{q^2}{2}\right)\mathrm{E}\big[\Phi(\lambda q U'^{\frac{1}{2}})\big],
\end{aligned}$$

where $\alpha_1 = \nu + r$, $\alpha_2 = q^2/2$ and $U' \sim \mathrm{TGamma}_{\lfloor 0,1\rfloor}(\alpha_1, \alpha_2)$.

$$\begin{aligned}
\mathcal{E}_{\phi}(r,q) &= \mathrm{E}\big[U^r \phi(qU^{\frac{1}{2}})\big] = \int_0^1 u^r \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{q^2}{2}u\right\}\nu u^{\nu-1}\,du \\
&= \frac{\nu}{\sqrt{2\pi}}\left(\frac{q^2}{2}\right)^{-(\nu+r)}\Gamma\left(\nu+r, \frac{q^2}{2}\right),
\end{aligned}$$

where $\Gamma(a,x) = \int_0^x e^{-t}t^{a-1}\,dt$.

Skew-contaminated normal distribution

In this case, U is a discrete random variable with probability function given by:

$$U = \begin{cases} \xi & \text{with probability } \gamma; \\ 1 & \text{with probability } 1-\gamma. \end{cases}$$

Thus, we have that

$$\begin{aligned}
\mathcal{E}_{\Phi_{\mathrm{SN}}}(r,q) &= \mathrm{E}\big[U^r \Phi_{\mathrm{SN}}(qU^{1/2}|\lambda)\big] \\
&= u^r \Phi_{\mathrm{SN}}(qu^{\frac{1}{2}}|\lambda)\big[\gamma \mathbb{I}_{\{\xi\}}(u) + (1-\gamma)\mathbb{I}_{\{1\}}(u)\big] \\
&= 2\gamma\xi^r \Phi_2(\xi^{\frac{1}{2}} q\mathbf{e}_1|\mathbf{0}, \mathbf{\Sigma}) + 2(1-\gamma)\Phi_2(q\mathbf{e}_1|\mathbf{0}, \mathbf{\Sigma}) \\
&= \xi^r 2\big[\gamma\Phi_2(\xi^{\frac{1}{2}} q\mathbf{e}_1|\mathbf{0}, \mathbf{\Sigma}) + (1-\gamma)\Phi_2(q\mathbf{e}_1|\mathbf{0}, \mathbf{\Sigma})\big] \\
&\quad + 2(1-\gamma)(1-\xi^r)\Phi_2(q\mathbf{e}_1|\mathbf{0}, \mathbf{\Sigma}) \\
&= \xi^r \Phi_{\mathrm{SCN}}\big(q|\lambda, (\gamma, \xi)\big) + 2(1-\gamma)\big(1-\xi^r\big)\Phi_2(q\mathbf{e}_1|\mathbf{0}, \mathbf{\Sigma}), \\
\mathcal{E}_{\Phi}(r,q) &= \mathrm{E}\big[U^r \Phi(qU^{1/2})\big] \\
&= u^r \Phi(qu^{\frac{1}{2}})\big[\gamma \mathbb{I}_{\{\xi\}}(u) + (1-\gamma)\mathbb{I}_{\{1\}}(u)\big] \\
&= \gamma\xi^r \Phi(q\xi^{\frac{1}{2}}) + (1-\gamma)\Phi(q) \\
&= \xi^r\big[\gamma\Phi(q\xi^{\frac{1}{2}}) + (1-\gamma)\Phi(q)\big] + (1-\gamma)\big(1-\xi^r\big)\Phi(q)
\end{aligned}$$

$$
\begin{aligned}
&= \xi^r \Phi_{\mathrm{CN}}(q|(\gamma,\xi)) + (1-\gamma)(1-\xi^r)\Phi(q), \\
\mathcal{E}_{\phi_{\mathrm{SN}}}(r,q) &= \mathrm{E}[U^r \phi_{\mathrm{SN}}(qU^{1/2}|\lambda)] \\
&= u^r \phi_{\mathrm{SN}}(qu^{\frac{1}{2}}|\lambda)[\gamma \mathbb{I}_{\{\xi\}}(u) + (1-\gamma)\mathbb{I}_{\{1\}}(u)] \\
&= \gamma \xi^r \phi_{\mathrm{SN}}(q\xi^{\frac{1}{2}}|\lambda) + (1-\gamma)\phi_{\mathrm{SN}}(q|\lambda). \\
\mathcal{E}_{\phi}(r,q) &= \mathrm{E}[U^r \phi(qU^{1/2})] \\
&= u^r \phi(qu^{\frac{1}{2}})[\gamma \mathbb{I}_{\{\xi\}}(u) + (1-\gamma)\mathbb{I}_{\{1\}}(u)] \\
&= \gamma \xi^r \phi(q\xi^{\frac{1}{2}}) + (1-\gamma)\phi(q).
\end{aligned}
$$

Acknowledgments

We would like to thank the Associate Editor and three reviewers for their constructive comments, which helped to improve this paper substantially. This paper was written while Celso R. B. Cabral was a visiting professor in the Department of Statistics at the University of Campinas, Brazil. Celso R. B. Cabral acknowledges the support from FAPESP (Grant 2015/20922-5) and CNPq-Brazil (Grant 447964/2014-3). The research of Aldo M. Garay was supported by Grant 420082/2016-6 from CNPq-Brazil.

References

Andrews, D. F. and Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society, Series B, Methodological* **36**, 99–102. MR0359122

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance* **9**, 203–228. MR1850791 https://doi.org/10.1111/1467-9965.00068

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* **12**, 171–178. MR0808153

Azzalini, A. (2018). The R package `sn`: The Skew-Normal and Related Distributions such as the Skew-*t* (version 1.5-2). Università di Padova, Italia. http://azzalini.stat.unipd.it/SN.

Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **65**, 367–389. MR1983753 https://doi.org/10.1111/1467-9868.00391

Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika* **83**, 715–726. MR1440039 https://doi.org/10.1093/biomet/83.4.715

Branco, M. D. and Dey, D. K. (2001). A general class of multivariate skew–elliptical distributions. *Journal of Multivariate Analysis* **79**, 99–113. MR1867257 https://doi.org/10.1006/jmva.2000.1960

Dempster, A., Laird, N. and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, Methodological* **39**, 1–38. MR0501537

Flecher, C., Allard, D. and Naveau, P. (2010). Truncated skew-normal distributions: Moments, estimation by weighted moments and application to climatic data. *Metron* **68**, 331–345. MR3041157 https://doi.org/10.1007/BF03263543

Garay, A. M., Lachos, V. H., Bolfarine, H. and Cabral, C. R. B. (2017). Linear censored regression models with scale mixtures of normal distributions. *Statistical Papers* **58**, 247–278. MR3610167 https://doi.org/10.1007/s00362-015-0696-9

Genç, A. İ. (2013). Moments of truncated normal/independent distributions. *Statistical Papers* **54**, 741–764. MR3072898 https://doi.org/10.1007/s00362-012-0459-9

Ho, H. J., Lin, T.-I., Chen, H.-Y. and Wang, W.-L. (2012). Some results on the truncated multivariate t distribution. *Journal of Statistical Planning and Inference* **142**, 25–40. MR2827127 https://doi.org/10.1016/j.jspi.2011.06.006

Hogg, R. V. and Klugman, S. A. (1984). *Loss Distributions*. New York: Wiley. MR0747141 https://doi.org/10.1002/9780470316634

Jamalizadeh, A., Pourmousa, R. and Balakrishnan, N. (2009). Truncated and limited skew-normal and skew-t distributions: Properties and an illustration. *Communications in Statistics Theory and Methods* **38**, 2653–2668. MR2568177 https://doi.org/10.1080/03610910902936109

Kim, H.-J. (2008). Moments of truncated Student-t distribution. *Journal of the Korean Statistical Society* **37**, 81–87. MR2409373 https://doi.org/10.1016/j.jkss.2007.06.001

Landsman, Z. M. and Valdez, E. A. (2003). Tail conditional expectations for elliptical distributions. *North American Actuarial Journal* **7**, 55–71. MR2061237 https://doi.org/10.1080/10920277.2003.10596118

Liu, C. and Rubin, D. B. (1994). The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence. *Biometrika* **81**, 633–648. MR1326414 https://doi.org/10.1093/biomet/81.4.633

Massuia, M. B., Garay, A. M., Cabral, C. R. B. and Lachos, V. H. (2017). Bayesian analysis of censored linear regression models with scale mixtures of skew-normal distributions. *Statistics and its Interface* **10**, 425–439. MR3608552 https://doi.org/10.4310/SII.2017.v10.n3.a7

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org.

Rubin, D. B. (1987). The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The SIR algorithm. *Journal of the American Statistical Association*, 543–546.

Rubin, D. B., Bernardo, J. M., DeGroot, M. H., Lindley, D. V. and Smith, A. F. M. (1988). Using the SIR algorithm to simulate posterior distributions. *Bayesian Statistics* **3**, 395–402.

V. H. Lachos
Department of Statistics
University of Connecticut
Storrs, Connecticut
06269
USA
E-mail: hlachos@uconn.edu

A. M. Garay
Department of Statistics
Federal University of Pernambuco
Recife, Pernambuco
50740-540
Brazil
E-mail: agaray@de.ufpe.br

C. R. B. Cabral
Department of Statistics
Federal University of Amazonas
Manaus, AM
69080-900
Brazil
E-mail: celsoromulo@ufam.edu.br