Moments of truncated skew-normal/independent distributions

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

[a] *Departamento de Estatística, Universidade Federal de Pernambuco, Brazil*
[b] *Departamento de Estatística, Universidade Estatual de Campinas, Campinas, São Paulo, Brazil*
[c] *Departamento de Estatística, Universidade Federal do Amazonas, Brazil*

Abstract

In this work we consider the problem of finding the moments of a doubly truncated member of the class of skew-normal/independent (TSNI) distributions. We obtained 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 TSNI family are studied, inference procedures are developed and a simulation study is performed to assess the procedures. An application in the context of censored regression models is also provided.

Keywords: Kurtosis; Moments; Truncated distributions; Skew-normal/independent distributions; Skewness

1. Introduction

The skew-normal/independent (SNI) 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 normal/independent (NI) distributions defined by Andrews and Mallows (1974) belong to the SNI family. However, the theory and 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 pdf's 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 normal/independent distributions. He obtained a general result and then used it to derive the moments in the case of

*Address for correspondence: Aldo M. Garay, Department of Statistics, Federal University of Pernambuco, Brazil. Av. Prof. Moraes Rego, 1235 - Cidade Universitária, Recife-PE. Brazil

Email address: hlachos@ime.unicamp.br agaray@de.ufpe.br celsoromulo@ufam.edu.br (Victor H. Lachos[b] Aldo M. Garay[a] and Celso R. B. Cabral[c])

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, Flecher et al. (2010) obtained expressions for the moments of truncated skew-normal distributions and applied the results to model the relative humidity data. However, to the best of our knowledge, only the work of Flecher et al. (2010) is dedicated to constructing truncated models based on the skew-normal family. In this work our objective is to combine the results of Flecher et al. (2010) and Genç (2013) to derive the moments of the doubly truncated version of a skew-normal/independent (TSNI) distribution. We then particularize the general result for some common skew-normal/independent distributions mentioned above. Since the first four moments are most useful, we give expressions for them. Computational aspects to generate random samples of the TSNI family of distributions are also discussed. Our proposal generalizes the results obtained by Kim (2008), Flecher et al. (2010), Genç (2013) and Garay et al. (2015).

The rest of the paper is organized as follows. In Section 2 we give a brief introduction of the SNI and TSNI distributions. In Section 3 we outline the main results related to the moments of the TSNI distributions. Section 4 deals with particular cases of the TSNI distributions. Section 5 discusses a practical application in the context of censored regression models. 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. Skew-normal independent (SNI) distributions

Throughout this paper, $X \sim \mathrm{N}(\mu, \sigma^2)$ denotes a random variable X with normal distribution with mean μ and variance σ^2 and $\phi\left(\cdot|\mu, \sigma^2\right)$ denotes its probability density function (pdf). In turn, $\phi(\cdot)$ and $\Phi(\cdot)$ denote the pdf and the cumulative distribution function (cdf) of the standard normal distribution, respectively. In general, we use the traditional convention of denoting a random variable (or a random vector) by an upper-case letter and its realization by the corresponding lower-case letter. Random vectors and matrices are denoted by boldface letters. $\mathbf{X}^\top$ is the transpose of $\mathbf{X}$. $X \perp Y$ indicates that the random variables X and Y are independent.

We start by defining the skew-normal (SN) distribution and then we introduce some useful properties. Thus, 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 $Z \sim SN(\mu, \sigma^2, \lambda)$, if its pdf is given by:

$$\phi_{SN}(z|\mu, \sigma^2, \lambda) = 2\phi(z|\mu, \sigma^2)\Phi\left(\frac{\lambda(z-\mu)}{\sigma}\right). \tag{1}$$

In the following Lemma, we present the pdf of a SN random variable.

Lemma 1. *Let a random variable $Z \sim SN(\mu, \sigma^2, \lambda)$, with pdf defined in (1). Then, the cdf of Z is given by:*

$$\Phi_{SN}(z|\mu, \sigma^2, \lambda) = 2\mathbf{\Phi}_2\Big(\mathbf{z}^*|\mathbf{0}, \mathbf{\Sigma}\Big), \tag{2}$$

where $\mathbf{\Phi}_m(\cdot \mid \boldsymbol{\mu}, \mathbf{\Sigma})$ denotes the cdf of the $m-$variate normal distribution with mean vector $\boldsymbol{\mu}$ and the covariance matrix $\mathbf{\Sigma}$, $\mathbf{0}$ denotes the zero vector, $\mathbf{z}^* = \left(\frac{z-\mu}{\sigma}, 0\right)^\top$ and $\mathbf{\Sigma} = \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}$, with $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

Proof. The proof of this Lemma is given in Appendix A. ☐

Next we present the stochastic representations of a random variable with SN distribution, which is useful to generate random samples and to obtain the moments and other related properties. If $Z \sim SN(\mu, \sigma^2, \lambda)$, then a convenient stochastic representation is given by:

$$Z = \mu + \Delta|T_0| + \Gamma^{1/2} T_1, \tag{3}$$

where $\Delta = \sigma\delta$, $\Gamma = (1-\delta^2)\sigma^2$, $T_0 \perp T_1$ and $|\cdot|$ denotes the absolute value.

It is important to stress that when $\mu = 0$ and $\sigma^2 = 1$, we have the standard SN distribution, whose pdf and cdf will be denoted by $\phi_{SN}(\cdot|\lambda)$ and $\Phi_{SN}(\cdot|\lambda)$, respectively.

The relation between the SNI class and the SN distribution is given in the next definition.

Definition 1. *A random variable Y has a SNI distribution with location parameter μ, scale parameter σ^2 and skewness parameter λ, denoted by SNI$(\mu, \sigma^2, \lambda; H)$, if it has the following stochastic representation:*

$$Y = \mu + U^{-1/2} Z, \quad U \perp Z, \tag{4}$$

where $Z \sim SN(0, \sigma^2, \lambda)$, U is a positive random variable with cdf $H(\cdot|\boldsymbol{\nu})$ indexed by a scalar or vector parameter $\boldsymbol{\nu}$.

The random variable U is known as *the scale factor* and its cdf $H(\cdot|\boldsymbol{\nu})$ is called the *mixing distribution function*. Note that when $\lambda = 0$, the SNI family reduces to the symmetric class of normal independent (NI) distributions.

Using the stochastic representation given in Equation (4), we observe that

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

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

$$\phi_{SNI}\left(y|\mu, \sigma^2, \lambda, \boldsymbol{\nu}\right) = 2\int_0^\infty \phi(y|\mu, u^{-1}\sigma^2)\Phi\left(u^{1/2}\frac{\lambda(y-\mu)}{\sigma}\right) dH(u|\boldsymbol{\nu}). \tag{5}$$

Thus, the cdf of a SNI random variable is given in the following Lemma:

Lemma 2. *Let the random variable $Y \sim$ SNI$(\mu, \sigma^2, \lambda; H)$. Then, the cdf of Y is given by:*

$$\Phi_{SNI}\left(y|\mu, \sigma^2, \lambda, \boldsymbol{\nu}\right) = \int_0^\infty 2\mathbf{\Phi}_2\Big(\boldsymbol{y}(u)^*|\mathbf{0}, \mathbf{\Sigma}\Big) dH(u), \tag{6}$$

where

$$\boldsymbol{y}(u)^* = \left(\sqrt{u}\,\frac{y-\mu}{\sigma}, 0\right)^\top, \quad \mathbf{\Sigma} = \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix} \text{ and } \delta = \frac{\lambda}{\sqrt{1+\lambda^2}}. \tag{7}$$

Proof. The proof is given in Appendix B. ☐

When $\mu = 0$ and $\sigma^2 = 1$ we have the standard SNI distribution, whose pdf and cdf will be denoted by $\phi_{SNI}(\cdot|\lambda, \boldsymbol{\nu})$ and $\Phi_{SNI}(\cdot|\lambda, \boldsymbol{\nu})$, respectively.

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

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

As an obvious consequence of Definition 2, we have that the pdf of the random variable $X \sim \mathrm{TSNI}_{\lfloor a,b \rfloor}(\mu, \sigma^2, \lambda; H)$ is given by:

$$\phi_{TSNI}\left(x \mid \mu, \sigma^2, \lambda, \boldsymbol{\nu}\ ;\ \lfloor a, b \rfloor\right) = \frac{\phi_{SNI}\left(x|\mu, \sigma^2, \lambda, \boldsymbol{\nu}\right)}{\Phi_{SNI}\left(b|\mu, \sigma^2, \lambda, \boldsymbol{\nu}\right) - \Phi_{SNI}\left(a|\mu, \sigma^2, \lambda, \boldsymbol{\nu}\right)} \mathbb{I}_{\lfloor a,b \rfloor}(x), \tag{8}$$

where $\mathbb{I}_{\mathbb{A}}(y)$ denotes the indicator function, that is, $\mathbb{I}_{\mathbb{A}}(y) = 1$ if $y \in \mathbb{A}$ and $\mathbb{I}_{\mathbb{A}}(y) = 0$ otherwise. When $\mu = 0$ and $\sigma^2 = 1$ we have the standard TSNI distribution, whose pdf and cdf will be denoted by $\phi_{TSNI}(\cdot|\lambda, \boldsymbol{\nu}; \lfloor a, b \rfloor)$ and $\Phi_{TSNI}(\cdot|\lambda, \boldsymbol{\nu}; \lfloor a, b \rfloor)$, respectively.

In what follows, $\mathrm{E}[\cdot]$ denotes expectation, $\mathrm{E}_X[\cdot]$ denotes expectation relative to the distribution of X and, for the sake of notation simplicity, we denote all pdf's by $f(\cdot)$. Thus, for example, $f(u, x)$ denotes the joint pdf of U and X and $f(u|X \in \mathcal{A})$ stands for the pdf of U given the event $\{X \in \mathcal{A}\}$.

Let $X \sim \mathrm{TSNI}_{\lfloor \alpha,\beta \rfloor}(0, 1, \lambda; H)$ be a standard TSNI distribution. Then it is straightforward to prove that $Y = \mu + \sigma X$ has a $\mathrm{TSNI}_{\lfloor a,b \rfloor}(\mu, \sigma^2, \lambda; H)$ 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

$$\mathrm{E}[Y^n] = \sum_{k=0}^{n} \frac{n!}{(n-k)!\ k!}\ \sigma^k \mu^{n-k} \mathrm{E}\left[X^k\right], \tag{9}$$

for $n = 1, 2, 3 \ldots$

3. Main results

The derivation of formulas for the moments of the TSNI distributions can require lengthy calculations. Instead, we propose a general recursive formula and then we get closed form expressions for the moments. In deriving these moments, we follow the same strategy used by Kim (2008) and Genç (2013).

The following Lemma, related to the moments of the truncated normal distribution (TN), was provided by Kim (2008) and used by Genç (2013) to derive the moments of the truncated normal/independent distributions (TNI).

Lemma 3. *If $Z \sim TN_{\lfloor a,b \rfloor}(0,1)$, then*

$$(k+1)\, E\left[Z^k\right] - E\left[Z^{k+2}\right] = \frac{b^{k+1}\phi(b) - a^{k+1}\phi(a)}{\Phi(b) - \Phi(a)},$$

for $k = -1, 0, 1, 2, \ldots$

Proof. See Lemma 2.3 in Kim (2008). □

In the following Lemma we present a recursive expression for the moments of the truncated skew-normal distribution (TSN). Although this result was obtained before by Flecher et al. (2010, Prop. 1), here we provide a new proof.

Lemma 4. *If $X \sim TSN_{\lfloor a,b \rfloor}(0,1,\lambda)$, then*

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

for $k = -1, 0, 1, 2, \ldots$, where $Z \sim TN_{\lfloor \lambda_* a,\, \lambda_* b\rfloor}(0,1)$, $\lambda_* = \sqrt{1+\lambda^2}$, $A(\lambda_*) = \Phi(\lambda_* b) - \Phi(\lambda_* a)$ and $B(\lambda) = \Phi_{SN}\left(b|\lambda\right) - \Phi_{SN}\left(a|\lambda\right)$.

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

$$\frac{d\, x^{k+1}\phi_{SN}(x|\lambda)}{d\, x} = (k+1)x^k\phi_{SN}(x|\lambda) - x^{k+2}\phi_{SN}(x|\lambda) + \frac{2}{\sqrt{2\pi}}\frac{\lambda}{\sqrt{2\pi}}x^{k+1}\exp\left\{-\frac{1}{2}\lambda_*^2x^2\right\}.$$

Then,

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

□

Lemma 5. *Consider a standard SNI distribution with stochastic representation given in Definition 1. Then, $E\left[\Phi_{SN}\left(a\sqrt{U}|\lambda\right)\right] = \Phi_{SNI}\left(a|\lambda,\boldsymbol{\nu}\right)$ for all $a \in \mathbb{R}$.*

Proof. The SNI distribution of the statement is the distribution of $Y = U^{-1/2}Z$, where U has cdf $H(\,\cdot\,|\boldsymbol{\nu})$, Z has cdf $\Phi_{SN}(\cdot|\lambda)$, and they are independent. Let $A = \{(z,u);\ u^{-1/2}z \le a\}$. Then,

$$\begin{aligned}
\Phi_{SNI}\left(a|\lambda,\boldsymbol{\nu}\right) &= \mathbb{P}\left(Y \le a\right) = \mathbb{P}\left(Z \le a\sqrt{U}\right) \\
&= \mathbb{P}((Z,U) \in A) = \int_{-\infty}^{a\sqrt{u}}\int_0^{\infty} f(z|u)dH\left(u|\boldsymbol{\nu}\right)dz \\
&= \int_0^{\infty}\mathbb{P}\left(Z \le a\sqrt{u}|U=u\right)dH\left(u|\boldsymbol{\nu}\right) \\
&= \mathrm{E}\left[\mathbb{P}\left(Z \le a\sqrt{U}\right)\right] = \mathrm{E}\left[\Phi_{SN}\left(a\sqrt{U}|\lambda\right)\right].
\end{aligned}$$

□

Now we establish the following theorem, which is crucial to the development of our proposed theory. This theorem states that the moments of a truncated skew-normal/independent distribution can be computed recursively. It generalizes the results obtained by Kim (2008), Flecher et al. (2010), Genç (2013) and Garay et al. (2015).

Theorem 1. *Let $Y \sim SNI(0,1,\lambda;H)$, with stochastic representation given in Definition 1. Then, for $a<b$, we have*

$$E\left[Y^{k+2}|Y\in\lfloor a,b\rfloor\right]=\tau(a,b)\times E\left[U^{-\frac{k+2}{2}}\left(\Phi_{SN}(b\sqrt{U}|\lambda)-\Phi_{SN}(a\sqrt{U}|\lambda)\right)R^{k+2}\right], \tag{10}$$

where

$$\tau(a,b) = \frac{1}{\Phi_{SNI}(b|\lambda,\boldsymbol{\nu})-\Phi_{SNI}(a|\lambda,\boldsymbol{\nu})}, \tag{11}$$
$$R^{k+2} = E\left[Z^{k+2}|U=u\right] \quad \textit{and} \quad Z|U=u\sim TSN_{\lfloor a\sqrt{u},b\sqrt{u}\rfloor}(0,1,\lambda) \quad \textit{for} \quad k=-1,0,1,2,\ldots$$

Proof. We use the following result for conditional expectations: if X_1 and X_2 are arbitrary random variables and $g(\cdot)$ is a measurable function, then

$$\mathrm{E}[\mathrm{E}(X_1|X_2)|g(X_2)]=\mathrm{E}[X_1|g(X_2)]. \tag{12}$$

For a proof of this, see Ash (2000, Theorem 5.5.10).

We have that $Y|Y\in\lfloor a,b\rfloor\sim \mathrm{TSNI}_{\lfloor a,b\rfloor}(0,1,\lambda;H)$. 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 and Z independent, we have:

$$\begin{aligned}
\mathrm{E}\left[Y^{k+2}|Y\in\lfloor a,b\rfloor\right] &= \mathrm{E}\left[U^{-\frac{k+2}{2}}Z^{k+2}|Y\in\lfloor a,b\rfloor\right]\\
&= \mathrm{E}\left[\,\mathrm{E}\left[U^{-\frac{k+2}{2}}Z^{k+2}|U,(Y\in\lfloor a,b\rfloor)\right]|Y\in\lfloor a,b\rfloor\right] && (13)\\
&= \mathrm{E}\left[U^{-\frac{k+2}{2}}\,\mathrm{E}\left[Z^{k+2}|U,(U^{-1/2}Z\in\lfloor a,b\rfloor)\right]|Y\in\lfloor a,b\rfloor\right]\\
&= \mathrm{E}\left[U^{-\frac{k+2}{2}}\,\mathrm{E}\left[Z^{k+2}|Z\in\lfloor a\sqrt{U},b\sqrt{U}\rfloor\right]|Y\in\lfloor a,b\rfloor\right] && (14)\\
&= \int_0^\infty u^{-\frac{k+2}{2}}\mathrm{E}\left[Z^{k+2}|Z\in\lfloor a\sqrt{u},b\sqrt{u}\rfloor\right]f(u|Y\in\lfloor a,b\rfloor)\,du, && (15)
\end{aligned}$$

where in (13) we used result (12) and in (14) 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:

$$\begin{aligned}
f(u|Y\in\lfloor a,b\rfloor) &= \int f(u,y|Y\in\lfloor a,b\rfloor)\,dy && (16)\\
&= \int f(u|Y=y,Y\in\lfloor a,b\rfloor)\,f(y|Y\in\lfloor a,b\rfloor)\,dy\\
&= \tau(a,b)\int f(u|Y=y,Y\in\lfloor a,b\rfloor)\,f(y)\,\mathbb{I}_{\lfloor a,b\rfloor}(y)dy && (17)\\
&= \tau(a,b)\int f(u,y)\,\mathbb{I}_{\lfloor a,b\rfloor}(y)dy && (18)\\
&= \tau(a,b)\int_a^b f(u)\,\phi_{SN}\left(x|0,u^{-1},\lambda\right)dx && (19)\\
&= \tau(a,b)f(u)\int_{a\sqrt{u}}^{b\sqrt{u}}\phi_{SN}(z|\lambda)\,dz\\
&= \tau(a,b)\,f(u)\left[\Phi_{SN}\left(b\sqrt{u}|\lambda\right)-\Phi_{SN}\left(a\sqrt{u}|\lambda\right)\right], && (20)
\end{aligned}$$

where $\tau(a,b)=1/\left(\Phi_{SNI}(b|\lambda,\boldsymbol{\nu})-\Phi_{SNI}(a|\lambda,\boldsymbol{\nu})\right)$. Equation (18) 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 (17) and (18) are equal to zero. Equation (19) is consequence of $Y|U=u \sim \text{SN}(0, u^{-1}, \lambda)$. Thus, by (15) and (20), we have:

$$\begin{aligned} \text{E}\left[Y^{k+2}|Y \in \lfloor a,b \rfloor\right] &= \int_0^{\infty} u^{-\frac{k+2}{2}} \text{E}\left[Z^{k+2}|Z \in \lfloor a\sqrt{u}, b\sqrt{u} \rfloor\right] \\ &\quad \times \tau(a,b) f(u) \left[\Phi_{SN}\left(b\sqrt{u}|\lambda\right) - \Phi_{SN}\left(a\sqrt{u}|\lambda\right)\right] du \\ &= \tau(a,b) \times \text{E}\left[U^{-\frac{k+2}{2}} \left(\Phi_{SN}(b\sqrt{U}|\lambda) - \Phi_{SN}(a\sqrt{U}|\lambda)\right) \text{R}^{k+2}\right], \end{aligned}$$

where $\text{R}^{k+2} = \text{E}\left[Z^{k+2}|U=u\right]$ and $Z|U=u \sim TSN_{\lfloor a\sqrt{u}, b\sqrt{u} \rfloor}(0,1,\lambda)$ for $k = -1, 0, 1, 2, \ldots$. □

Observe that the expectation $\text{E}\left[Z^{k+2}|U=u\right]$, where $Z|U = u \sim TSN_{\lfloor a\sqrt{u}, b\sqrt{u} \rfloor}(0,1,\lambda)$, can be computed using Lemma 4. These computations involve the expectation $E[Z^{k+1}|U=u]$, where $Z|U = u \sim TN_{\lfloor \lambda_* a\sqrt{u}, \lambda_* b\sqrt{u} \rfloor}(0,1,\lambda)$, which can be computed using Lemma 3. As an important remark, observe that the case $\lambda = 0$ in Theorem 1 corresponds to the moments of the $NI(0,1;H)$ distribution, which were obtained before by Genç (2013, Theorem 1). However, in this case, our expression for the moments does not coincide with that obtained by Genç (2013). They differ by the term $\Phi(b\sqrt{U}) - \Phi(a\sqrt{U})$, which is not inside the expectation sign in Genç's Theorem.

In general, the first four moments $\mu_i = E\left[Y^i|Y \in \lfloor a,b \rfloor\right], (i = 1,\ 2,\ 3,\ 4)$ are most useful. Thus, we have the following corollary:

Corollary 1. *Let $X \sim SNI(0,1,\lambda,\boldsymbol{\nu})$ and $a < b$, we have that the first four moments $\mu_i = E\left[X^i|X \in \lfloor a,b \rfloor\right]$, for $i = 1,2,3,4$ are given by*

$$\mu_1 = \tau(a,b)\left[L(1)\left(E_\Phi(-0.5\,,\,b\lambda^*) - E_\Phi(-0.5, a\lambda^*)\right) - \left(E_{\phi_{SN}}(-0.5, b) - E_{\phi_{SN}}(-0.5, a)\right)\right],$$

$$\begin{aligned} \mu_2 = {} & \tau(a,b)\left[\left(E_{\Phi_{SN}}(-1,b) - E_{\Phi_{SN}}(-1,a)\right) - L(2)\left(E_\phi(-1\,,\,b\lambda^*) - E_\phi(-1, a\lambda^*)\right)\right. \\ & \left. - \left(bE_{\phi_{SN}}(-0.5\,,\,b) - aE_{\phi_{SN}}(-0.5, a)\right)\right], \end{aligned}$$

$$\begin{aligned} \mu_3 = {} & \tau(a,b)\left[2\left(L(1)\left(E_\Phi(-1.5\,,\,b\lambda^*) - E_\Phi(-1.5, a\lambda^*)\right) - \left(E_{\phi_{SN}}(-1.5, b) - E_{\phi_{SN}}(-1.5, a)\right)\right)\right. \\ & - \left(b^2 E_{\phi_{SN}}(-0.5\,,\,b) - a^2 E_{\phi_{SN}}(-0.5, a)\right) + L(3)\left(E_\Phi(-1.5\,,\,b\lambda^*) - E_\Phi(-1.5, a\lambda^*)\right) \\ & \left. - L(2)\,\left(bE_\phi(-1\,,\,b\lambda^*) - aE_\phi(-1, a\lambda^*)\right)\right], \end{aligned}$$

$$\begin{aligned} \mu_4 = {} & \tau(a,b)\left[3\left(\left(E_{\Phi_{SN}}(-2\,,\,b) - E_{\Phi_{SN}}(-2\,,\,a)\right) - L(2)\left(E_\phi(-2\,,\,b\lambda_*) - E_\phi(-2\,,\,a\lambda_*)\right)\right)\right. \\ & \left. - \left(bE_{\phi_{SN}}(-1.5\,,\,b) - aE_{\phi_{SN}}(-1.5\,,\,a)\right)\right) - \left(b^3 E_{\phi_{SN}}(-0.5\,,\,b) - a^3 E_{\phi_{SN}}(-0.5\,,\,a)\right)\right] \\ & - \tau(a,b)L(4)\left[2\left(\left(E_\phi(-2\,,\,b\lambda_*) - E_\phi(-2\,,\,a\lambda_*)\right)\right) + \left((b\lambda_*)^2 E_\phi(-1\,,\,b\lambda_*) - (a\lambda_*)^2 E_\phi(-1\,,\,a\lambda_*)\right)\right], \end{aligned}$$

where $\tau(a,b)$ is defined in (11), $\lambda_* = \sqrt{(1+\lambda^2)}$, $L(s) = \sqrt{\dfrac{2}{\pi}} \dfrac{\lambda}{(\lambda_*)^s}$,

$$E_{\phi_{SN}}(r,q) = E\left[U^r \phi_{SN}(q\,U^{1/2}|\lambda)\right] = \int_0^{\infty} u^r \phi_{SN}(q\,u^{1/2}|\lambda) dH(u|\boldsymbol{\nu}),$$

and $E_{\Phi_{SN}}(r,q)$, $E_\phi(r,q)$ and $E_\Phi(r,q)$ are defined as $E_{\phi_{SN}}(r,q)$, with Φ_{SN}, ϕ and Φ replacing ϕ_{SN}, 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), respectively given by:

$$
\begin{aligned}
S &= (\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3)/(\mu_2 - \mu_1^2)^{3/2}, \\
K &= (\mu_4 - 4\mu_1\mu_3 + 6\mu_2\mu_1^2 - 3\mu_1^4)/(\mu_2 - \mu_1^2)^2, \quad \text{and} \\
CV &= (\mu_2 - \mu_1^2)^{1/2}/\mu_1.
\end{aligned}
$$

4. Particular cases of SNI distributions

The asymmetric class of SNI 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 element of this class, in its standardized form we compute the expected values $E_{\phi_{SN}}(r,q)$, $E_{\Phi_{SN}}(r,q)$, $E_{\phi}(r,q)$ and $E_{\Phi}(r,q)$. For the sake of completeness, a detailed proof of these results is given in Appendix C.

4.1. The skew-t distribution

In this case we consider in Definition 1 $U \sim Gamma(\nu/2, \nu/2)$, $\nu > 0$, where $Gamma(a,b)$ denotes the gamma distribution with mean a/b. The density of Y takes the form (Branco et al., 2013)

$$
\phi_{ST}(y|\mu,\sigma^2,\lambda,\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}\sigma}\left(1+\frac{d}{\nu}\right)^{-\frac{\nu+1}{2}} \mathbf{T}\left(\sqrt{\frac{v+1}{d+\nu}}A|\nu+1\right), \quad y \in \mathbb{R},
$$

where, $A = \dfrac{\lambda(y-\mu)}{\sigma}$, $d = (y-\mu)^2/\sigma^2$ 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 $Y \sim ST(\mu,\sigma^2,\lambda,\nu)$. 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.

For the next result $\mathbf{T}_p(\cdot|\,\boldsymbol{\mu},\boldsymbol{\Sigma},\nu)$ denotes the cdf of the p-variate Student-t distribution with mean vector $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$ and ν degrees of freedom.

Lemma 6. *Let $\mathbf{X} \sim N_p(\mathbf{0},\boldsymbol{\Sigma})$, with cdf $\boldsymbol{\Phi}_p(\cdot|\mathbf{0},\ \boldsymbol{\Sigma})$, and $U \sim Gamma\,(\alpha,\beta)$ be independent. Then, for any fixed vector $\boldsymbol{w} \in \mathbb{R}^p$,*

$$
E\left[\boldsymbol{\Phi}_p\left(\sqrt{U}\boldsymbol{w} \mid \mathbf{0},\boldsymbol{\Sigma}\right)\right] = \mathbf{T}_p\left(\sqrt{\frac{\alpha}{\beta}}\boldsymbol{w} \mid \mathbf{0},\boldsymbol{\Sigma},2\alpha\right).
$$

Proof. Let $H(\cdot)$ be the cdf of the $Gamma\,(\alpha,\beta)$ distribution and $\phi_p(\cdot|\boldsymbol{\mu},\ \boldsymbol{\Sigma})$ be the pdf of the p-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. In what follows, for $\mathbf{X} = (X_1,\ldots,X_p)^\top$ and $\mathbf{x} = (x_1,\ldots,x)^\top$, $\mathbf{X} \le \mathbf{x}$ is interpreted element wise, that is, $X_i \le x_i$, $i = 1,\ldots,p$. We have that

$$
\begin{aligned}
E\left[\boldsymbol{\Phi}_p\left(\sqrt{U}\boldsymbol{w} \mid \mathbf{0},\boldsymbol{\Sigma}\right)\right] &= \int_0^\infty \mathbb{P}\left(\mathbf{X} \le \boldsymbol{w}\sqrt{u}\right) dH(u) \\
&= \int_0^\infty \mathbb{P}\left(\frac{\mathbf{X}}{(u\beta/\alpha)^{1/2}} \le \sqrt{\frac{\alpha}{\beta}}\boldsymbol{w}|U=u\right) dH(u). \\
&= E\left[\mathbb{P}\left(\frac{\mathbf{X}}{(U\beta/\alpha)^{1/2}} \le \sqrt{\frac{\alpha}{\beta}}\boldsymbol{w}|U\right)\right] \\
&= \mathbb{P}\left(\frac{\mathbf{X}}{(U\beta/\alpha)^{1/2}} \le \sqrt{\frac{\alpha}{\beta}}\boldsymbol{w}\right).
\end{aligned}
$$

It is straightforward to show that the random vector $\alpha^{1/2}\mathbf{X}/(U\beta)^{1/2}$ has a Student-t distribution with mean vector $\mathbf{0}$, scale matrix $\boldsymbol{\Sigma}$ and 2α degrees of freedom, concluding the proof. □

From Lemmas 2 and 6 we obtain an expression for the cdf a skew-t random variable, given by:

$$\Phi_{ST}(y|\mu,\sigma^2,\lambda,\nu) = 2E[\boldsymbol{\Phi}_2(\sqrt{U}\mathbf{y}^*)|\mathbf{0},\ \boldsymbol{\Sigma}] = 2\ \mathbf{T}_2\left(\mathbf{y}^* \mid \mathbf{0},\boldsymbol{\Sigma},\nu\right),$$

where $\mathbf{y}^* = ((y-\mu)/\sigma, 0)^\top$, $U \sim Gamma(\nu/2,\nu/2)$ and $\boldsymbol{\Sigma}$ is given in (7).

The truncated skew-t distribution will be denoted by $TST_{\lfloor a,b \rfloor}(\mu,\sigma^2,\lambda,\nu)$. For this model we have the following result for the expected values involved in the calculation of the moments.

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

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

where $\Gamma(a)$ *is the gamma function,* $\Phi_{PVII}(\cdot|\nu+2r,\nu)$ *is the cdf of the standard Pearson type VII distribution (see Garay et al. (2015) for more details),* $\mathbf{y}^* = (q,0)^\top$ *and* $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}$.

4.2. The skew-slash distribution

In this case we have $U \sim Beta(\nu,1)$ – where $Beta(a,b)$ denotes the beta distribution with parameters a and b – with positive shape parameter ν, and we use the notation $Y \sim SSL(\mu,\sigma^2,\lambda,\nu)$. The density of Y is given by:

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

The cdf of the skew-slash distribution does not have a closed form expression. However, using Lemma 2, we can write it in terms of the following integral, which can be obtained by numerical methods:

$$\Phi_{SSL}(y|\mu,\sigma^2,\lambda,\nu) = \int_0^\infty 2\nu\Phi_2\Big(\boldsymbol{y}(u)^*|\mathbf{0},\boldsymbol{\Sigma}\Big)u^{\nu-1}\,du, \tag{21}$$

where $\boldsymbol{y}(u)^*$ and $\boldsymbol{\Sigma}$ are as in (7). The truncated skew-slash distribution will be denoted by $TSSL_{\lfloor a,b \rfloor}(\mu,\sigma^2,\lambda,\nu)$.

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

$$\begin{aligned}
E_{\phi_{SN}}(r,q) &= \frac{2^{\nu+r+1}\nu\Gamma(r+\nu)}{\sqrt{2\pi}q^{2r+2\nu}}G\left(1|r+\nu,\frac{q^2}{2}\right)E\left[\Phi(\lambda q\sqrt{U'})\right];\\
E_{\Phi_{SN}}(r,q) &= \frac{2\nu}{r+\nu}E\left[\Phi_2(\sqrt{U''}\mathbf{y}^*|0,\ \boldsymbol{\Sigma})\right];\\
E_{\Phi}(r,q) &= \left(\frac{\nu}{\nu+r}\right)\Phi_{SL}(q|\nu+r);\\
E_{\phi}(r,q) &= \frac{\nu}{\sqrt{2\pi}}\left(\frac{q^2}{2}\right)^{-(\nu+r)}\Gamma\left(\nu+r,\frac{q^2}{2}\right),
\end{aligned}$$

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*,* $U^{'}\sim TGamma(r+\nu, q^2/2, \lfloor 0,1\rfloor)$*,* $U^{''}\sim Beta(r+\nu,1)$*,* $\mathbf{y}^*=(q,0)^\top$*,* $\mathbf{\Sigma}=\begin{pmatrix}1 & -\delta\\ -\delta & 1\end{pmatrix}$ *and* $\Phi_{SL}(\cdot|\nu+r)$ *is the cdf of the standard slash distribution (See Garay et al. (2015, p. 7)).*

Observe that $E_{\phi_{SN}}(r,q)$ and $E_{\Phi_{SN}}(r,q)$ do not have closed form expressions, but they can be approximated by Monte Carlo methods, using the R random number generators of the function `rbeta` (for the beta distribution) and the package **truncdist** (for the truncated gamma distribution, as in Novomestky and Nadarajah (2012)).

4.3. The skew-contaminated normal distribution

This distribution is denoted by $Y\sim SCN(\mu,\sigma^2,\lambda,(\gamma,\xi))$. Here U is a discrete random variable taking one of two states, namely:

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

It follows immediately that:

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

and

$$\Phi_{SCN}(y|\mu,\sigma^2,\lambda,(\gamma,\xi))=2\left[\gamma\Phi_2\Big(\sqrt{\xi}\boldsymbol{y}^*|\mathbf{0},\mathbf{\Sigma}\Big)+(1-\gamma)\Phi_2\Big(\boldsymbol{y}^*|\mathbf{0},\mathbf{\Sigma}\Big)\right], \tag{22}$$

where $A=\dfrac{\lambda(y-\mu)}{\sigma}$, $\mathbf{y}^*=((y-\mu)/\sigma,0)^\top$, and $\mathbf{\Sigma}$ is given in (7). In this case, the expected values are given in the following corollary:

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

$$\begin{aligned}
E_{\phi_{SN}}(r,q) &= \gamma\xi^r\phi_{SN}\left(q\xi^{\frac{1}{2}}|\lambda\right)+(1-\gamma)\phi_{SN}\Big(q|\lambda\Big)\\
E_{\Phi_{SN}}(r,q) &= \xi^r\Phi_{SCN}\Big(q|\lambda,(\gamma,\xi)\Big)+2\,(1-\gamma)\,(1-\xi^r)\,\Phi_2\Big(\boldsymbol{y}^*|\mathbf{0},\mathbf{\Sigma}\Big);\\
E_{\Phi}(r,q) &= \xi^r\Phi_{CN}(q|(\gamma,\xi))+(1-\gamma)\,(1-\xi^r)\,\Phi\,(q)\,;\\
E_{\phi}\,(r,q) &= \gamma\xi^r\phi\left(\sqrt{\xi}q\right)+(1-\gamma)\,\phi\,(q)\,,
\end{aligned}$$

where $\mathbf{y}^*=(q,0)^\top$*,* $\mathbf{\Sigma}=\begin{pmatrix}1 & -\delta\\ -\delta & 1\end{pmatrix}$ *and* $\Phi_{CN}(\cdot|(\gamma,\xi))$ *represents the cdf of the standard contaminated normal distribution.*

An important aspect is how to generate random samples of a TSNI distribution. Thus, to be more informative to the reader, we present two methods to generate random samples from $Y\sim \text{TSNI}_{\lfloor a,b\rfloor}(\mu,\sigma^2,\lambda;H)$.

4.4. Computational aspects to generate samples from the TSNI models

In this section, we describe two simulation methods to generate samples, from the random variable $Y\sim\text{TSNI}_{\lfloor a,b\rfloor}(\mu,\sigma^2,\lambda;H)$. The first is based on the sampling/importance re-sampling method proposed by Rubin (1987) and Rubin et al. (1988) and the other one is based on the stochastic representation of a SNI random variable given in Definition 1.

- **SIR Method**

  The sampling/importance resampling (SIR) method is useful to generate an approximate independent and identically distributed sample of size r from the target density $f(y)$, where $y \in \mathcal{S}_Y \subseteq \mathbb{R}$. Thus, let $g(y)$ be a proposed density with the same support $\mathcal{S}_Y$. The method consists of two steps:

  - **Step 1:** ***Sampling***

    Generate a random sample $Y_1,\ Y_2, \cdots, Y_J$ from $g(y)$ and construct weights

    $$W(Y_j) = \frac{f(Y_j)}{g(Y_j)},\quad j = 1, \ldots, J$$

    and probabilities

    $$\pi_j = \frac{W(Y_j)}{\sum_{j=1}^{J} W(Y_j)},\quad j = 1, \ldots, J.$$

  - **Step 2:** ***Importance resampling***

    Draw r values ($r << J$) $Y_1^*, ..., Y_r^*$ from the J values $Y_1, Y_2, \ldots, Y_J$ with respective probabilities $\pi_1,\ \pi_2, \ldots \pi_J$. In practice, Rubin (1987) suggested $J/r = 20$.

  For instance, for the $\mathrm{TSN}_{\lfloor a,b \rfloor}\left(\mu, \sigma^2, \lambda\right)$ we can consider the truncated normal distribution $\mathrm{TN}_{\lfloor a,b \rfloor}\left(\mu, \sigma^2\right)$ as the proposed density $g(\cdot)$. For the $\mathrm{TST}_{\lfloor a,b \rfloor}\left(\mu, \sigma^2, \lambda, \nu\right)$, we can use the truncated t distribution $\mathrm{Tt}_{\lfloor a,b \rfloor}\left(\mu, \sigma^2, \nu\right)$ as the proposed density $g(\cdot)$.

- **Stochastic Representation Method**

  Another simulation method to generate samples from the truncated SNI distribution is to use the stochastic representation given in Definition 1. Thus, if one is interested in generating a random sample of size r from $Y \sim \mathrm{TSNI}_{\lfloor a,b \rfloor}(\mu, \sigma^2, \lambda; H)$, and since U is a positive random variable with $a < Y < b$, then we have that:

  $$(a-\mu)U^{1/2} < (Y-\mu)U^{1/2} < (b-\mu)U^{1/2}.$$

  Considering the stochastic representation given in (4), we have that $Z = (Y-\mu)U^{1/2}$, where $Z \sim \mathrm{SN}(0, \sigma^2, \lambda)$. Thus,

  $$(a-\mu)U^{1/2} < Z < (b-\mu)U^{1/2}.$$

  Therefore, to generate random samples of size r from a TSNI distribution, we propose the following algorithm:

  - **Step 1.** Generate a random sample $U_1, U_2, \ldots, U_r$ from $H(\,\cdot\,|\boldsymbol{\nu})$.
  - **Step 2.** Generate a random sample $Z_1, Z_2, \ldots, Z_r$ from the $\mathrm{TSN}_{\lfloor \gamma_1, \gamma_2 \rfloor}(0, \sigma^2, \lambda)$ model, where $\gamma_1 = (a-\mu)U^{1/2}$ and $\gamma_2 = (b-\mu)U^{1/2}$ by using the SIR method, for instance.
  - **Step 3.** Finally, using the stochastic representation given in Definition 1, we set $Y = \mu + U^{1/2}Z$.

5. Statistical Application : The SNI 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, i.e., the responses are either left or right censored. For instance, Garay et al. (2015) recently established a new link between the censored regression model and the symmetric class of NI distributions, which extends the normal one by the inclusion of kurtosis. An interesting extension is to consider the asymmetrical class of SNI distributions, which allows capturing skewness and kurtosis in data simultaneously. Thus, in the following we define the censored linear regression model under skew normal/independent distributions, denoted the SNI-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 (iid) according to some SNI distribution, as follows:

$$Y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \sigma\varepsilon_i, \quad \varepsilon_i \overset{\text{iid}}{\sim} \text{SNI}(0,1,\lambda;H), \;\; i=1,\dots,n, \tag{23}$$

  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 $Y_i \sim \text{SNI}(\mathbf{x}_i^\top\boldsymbol{\beta}, \sigma^2, \lambda; H)$. In this application, we are interested in the case where left-censored observations can occur. That is, the observations are of the form:

$$Y_{\text{obs}_i} = \begin{cases} \kappa_i & \text{if} \quad Y_i \le \kappa_i; \\ Y_i & \text{if} \quad Y_i > \kappa_i, \end{cases} \tag{24}$$

  $i = 1,\dots,n$, for some threshold point κ_i. The model defined in (23)-(24) is called the SNI-CR model. See Massuia et al. (2015) for further details.

- **The mean and variance of the SNI–CR models**

  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 SNI–CR models for the $i-$th observed response are defined by:

$$\begin{aligned} \text{E}\left[Y_{\text{obs}_i}\right] &= \text{E}\left[\{\kappa_i\,(1-D_i)\} + Y_i D_i\right] \\ &= \kappa_i \Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right) + \text{E}\left[Y_i D_i\right]. \end{aligned} \tag{25}$$

$$\begin{aligned} \text{Var}\left[Y_{\text{obs}_i}\right] &= \text{Var}\left[\{\kappa_i\,(1-D_i)\} + \; Y_i D_i\right] \\ &= \kappa_i^2 \Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right)\left[1 - \Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right)\right] + \text{Var}\left[Y_i D_i\right]. \end{aligned} \tag{26}$$

  Defining $\kappa_i^* = (\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta})/\sigma$, the expectation in (25) can be obtained in the following way:

$$\begin{aligned} \text{E}\left[Y_i D_i\right] &= \text{E}\left[\,\text{E}(Y_i D_i|D_i)\right] = \text{E}\left[Y_i|Y_i > \kappa_i\right]\mathbb{P}\left(Y_i > \kappa_i\right) \\ &= \left(\mathbf{x}_i^\top\boldsymbol{\beta} + \sigma\text{E}\left[\varepsilon_i|\varepsilon_i > \kappa_i^*\right]\right)\left[1 - \Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right)\right]. \end{aligned} \tag{27}$$

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

The variance in (26) can be obtained as follows:

$$\begin{aligned}
\mathrm{Var}\left[Y_iD_i\right] &= \mathrm{E}\left[\mathrm{Var}\left(Y_iD_i|D_i\right)\right] + \mathrm{Var}\left[\mathrm{E}\left(Y_iD_i|D_i\right)\right] \\
&= \mathrm{Var}\left[Y_i|Y_i > \kappa_i\right]\left[1 - \Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right)\right] \\
&\quad + \left(\mathrm{E}\left[Y_i|Y_i > \kappa_i\right]\right)^2\left[1 - \Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right)\right]\Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right) \\
&= \sigma^2\mathrm{Var}\left[\varepsilon_i|\varepsilon_i > \kappa_i^*\right]\left[1 - \Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right)\right] \\
&\quad + \left(\mathbf{x}_i^\top\boldsymbol{\beta} + \sigma\mathrm{E}\left[\varepsilon_i|\varepsilon_i > \kappa_i^*\right]\right)^2\left[1 - \Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right)\right]\Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right) \\
&= \sigma^2\left[\mathrm{E}\left[\varepsilon_i^2|\varepsilon_i > \kappa_i^*\right] - \left(\mathrm{E}\left[\varepsilon_i|\varepsilon_i > \kappa_i^*\right]\right)^2\right]\left[1 - \Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right)\right] \\
&\quad + \left(\mathbf{x}_i^\top\boldsymbol{\beta} + \sigma\mathrm{E}\left[\varepsilon_i|\varepsilon_i > \kappa_i^*\right]\right)^2\left[1 - \Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right)\right]\Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right),
\end{aligned} \tag{28}$$

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

- **The log-likelihood function**

  Let $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \sigma^2, \lambda, \boldsymbol{\nu})^\top$ be the vector with all parameters in the SNI-CR model defined in (23)-(24). Supposing there are (possibly) m censored values of the characteristic of interest, we can partition the observed sample $\mathbf{y}_{\text{obs}}$ into two subsamples of m censored and $n - m$ uncensored values, such that $\mathbf{y}_{\text{obs}} = \{\kappa_1, \ldots, \kappa_m, y_{m+1}, \ldots, y_n\}$. Then, the log-likelihood function is given by:

$$\ell\left(\boldsymbol{\theta}|\mathbf{y}_{\text{obs}}\right) = \sum_{i=1}^{m}\log\left[\Phi_{SNI}\left(\frac{\kappa_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}|\lambda,\boldsymbol{\nu}\right)\right] + \sum_{i=m+1}^{n}\log\left[\phi_{SNI}\left(y_i|\mathbf{x}_i^\top\boldsymbol{\beta},\sigma^2,\lambda,\boldsymbol{\nu}\right)\right]. \tag{29}$$

  For maximum likelihood estimation of the SNI-CR model, an alternative is to maximize the log-likelihood function (29) directly, via the **optim** routine in the R software (R Development Core Team, 2015), for example. This procedure can be quite cumbersome, and a more reliable one is to implement the EM algorithm (Dempster et al., 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.

6. Simulation Study

In this section we present a simulation study to compare the theoretical moments of the truncated skew-normal/independent distributions along with the empirical moments computed via a Monte Carlo approximation.

Table 1: Simulation study. Comparison between the theoretical and empirical moments of the TSNI distributions, considering different values of the skewness parameter λ.

Skewness Parameter	Models	Theoretical (Empirical) moments			
		μ_1	μ_2	μ_3	μ_4
$\lambda = -3$	**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)
$\lambda = -1$	**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)
$\lambda = 0$	**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)
$\lambda = 1$	**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)
$\lambda = 3$	**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)

Thus, we generated 300 artificial samples of size 1000 from $Y \sim \text{TSNI}_{\lfloor a,b \rfloor}(\mu, \sigma^2, \lambda; H)$, using the simulation methods described in Subsection 4.4. 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 respectively 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 1, shows the comparison between the first four theoretical moments $\text{E}\left[Y^k\right] = \mu^k$, $(k = 1, 2, 3, 4)$, computed using Theorem 1, and the values of the average empirical moments, across 300 replicates (In parentheses), of the TSNI 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 TSNI distributions is reliable.

7. Concluding Remarks and Discussion

In this paper, we have developed a simpler alternative device to formulate an exact expression for the moments of doubly truncated univariate skew-normal/independent distributions (TSNI), which are derived from a stochastic representation of scale mixture of skew normal distributions (SNI). The main formulas are useful in many specific problems and they generalize the results obtained by Kim (2008), Flecher et al. (2010), Genç (2013) and Garay et al. (2015). Furthermore, we have present some efficient simulation methods to generate random samples from the TSNI distributions. We have demonstrated the practicability and usefulness of the proposed techniques with a simulation study and an application.

We conjecture that our method can be extending to the context of multivariate truncated SNI 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 iterative algorithm has been coded and implemented in the R software (R Development Core Team, 2015), which is available from us upon request.

Acknowledgments

Aldo M. Garay acknowledges support from FAPESP-Brazil (Grant 2013/21468-0). Víctor H. Lachos acknowledges support from CNPq-Brazil (Grant 305054/2011-2) and FAPESP-Brazil (Grant 2014/02938-9). Celso R. B. Cabral was supported by CNPq (Grants 308243/2012-9 and 447964/2014-3).

Appendix A. Proof of Lemma 1

In this appendix, we derive the cdf of the random variable $Z \sim \mathrm{SN}(\mu, \sigma^2, \lambda)$

Proof. Let $Z \sim \mathrm{SN}(\mu, \sigma^2, \lambda)$. We have that $Z = \mu + \sigma W$, with $W \sim \mathrm{SN}(0, 1, \lambda)$. Then,

$$\begin{aligned}
\Phi_{SN}\left(z|\mu, \sigma^2, \lambda\right) &= \mathbb{P}\left(Z \leq z|\mu, \sigma^2, \lambda\right) = \mathbb{P}\left(\mu + \sigma W \leq z|\lambda\right) \\
&= \Phi_{SN}\left(\frac{z-\mu}{\sigma}|0, 1, \lambda\right). \qquad \text{(A.1)}
\end{aligned}$$

The cdf of the random variable $W \sim \mathrm{SN}(0, 1, \lambda)$ can be written as

$$\begin{aligned}
\Phi_{SN}\left(w|0, 1, \lambda\right) &= \int_{-\infty}^{w} \phi_{SN}\left(t|0, 1, \lambda\right) dt = \int_{-\infty}^{w} 2\phi\left(t\right)\Phi\left(\lambda t\right) dt = 2\int_{-\infty}^{w}\int_{-\infty}^{\lambda t} \phi\left(t\right)\phi\left(u\right) du\ dt \\
&= 2\int_{-\infty}^{w}\int_{-\infty}^{0} \sqrt{1+\lambda^2}\phi\left(t\right)\phi\left(v\sqrt{1+\lambda^2}+\lambda t\right) dv\ dt \qquad \text{(A.2)} \\
&= 2\int_{-\infty}^{w}\int_{-\infty}^{0} \frac{\sqrt{1+\lambda^2}}{2\pi} exp\left\{-\frac{1}{2}\left[\left(1+\lambda^2\right)t^2 + 2\lambda\sqrt{1+\lambda^2}tv + \left(1+\lambda^2\right)v^2\right]\right\} dv\ dt \\
&= 2\int_{-\infty}^{w}\int_{-\infty}^{0} \frac{\sqrt{1+\lambda^2}}{2\pi} exp\left\{-\frac{1}{2}\left[\begin{pmatrix} t \\ v \end{pmatrix}^{\top}\begin{pmatrix} 1+\lambda^2 & \lambda\sqrt{1+\lambda^2} \\ \lambda\sqrt{1+\lambda^2} & 1+\lambda^2 \end{pmatrix}\begin{pmatrix} t \\ v \end{pmatrix}\right]\right\} dv\ dt \\
&= 2\int_{-\infty}^{w}\int_{-\infty}^{0} \frac{1}{2\pi\sqrt{1-\delta^2}} exp\left\{-\frac{1}{2}\left[\begin{pmatrix} t \\ v \end{pmatrix}^{\top}\begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}^{-1}\begin{pmatrix} t \\ v \end{pmatrix}\right]\right\} dv\ dt \\
&= 2\boldsymbol{\Phi}_2\left(\boldsymbol{w}^*|\mathbf{0}, \boldsymbol{\Sigma}\right), \qquad \text{(A.3)}
\end{aligned}$$

where $\boldsymbol{w}^* = (w, 0)^{\top}$ and $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}$, with $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$. Equation (A.2) is obtained using the transformation $v = \frac{u-\lambda t}{\sqrt{1+\lambda^2}}$. Thus, using (A.3) in (A.1), the cdf of Z is obtained immediately. □

Appendix B. Proof of Lemma 2

In this appendix, we derive the cdf of the random variable $Y \sim \text{SNI}(\mu, \sigma^2, \lambda; H)$, using the same strategy used in Appendix A.

Proof. Let $Y \sim \text{SNI}(\mu, \sigma^2, \lambda; H)$. We have that $Y = \mu + \sigma X$, with $X \sim \text{SNI}(0, 1, \lambda; H)$. Then,

$$\begin{aligned}\Phi_{SNI}\left(y|\mu, \sigma^2, \lambda, \boldsymbol{\nu}\right) &= \mathbb{P}\left(Y \le y|\mu, \sigma^2, \lambda\right) = \mathbb{P}\left(\mu + \sigma X \le y|\lambda\right) \\ &= \Phi_{SNI}\left(\frac{y-\mu}{\sigma}|0, 1, \lambda, \boldsymbol{\nu}\right). \end{aligned} \tag{B.1}$$

Let $h(\cdot)$ be the pdf associated to $H(\cdot)$. Then, the cdf of X can be written as

$$\Phi_{SNI}\left(x|0,1,\lambda,\boldsymbol{\nu}\right) = \int_{-\infty}^{x} \phi_{SNI}\left(t|0,1,\lambda,\boldsymbol{\nu}\right) dt = \int_{-\infty}^{x}\int_{0}^{\infty} \phi_{SN}\left(t|0,u^{-1},\lambda\right) h\left(u\right) du\ dt \tag{B.2}$$

$$= \int_{-\infty}^{x}\int_{0}^{\infty} 2\phi\left(t|0,u^{-1}\right)\Phi\left(\lambda t u^{\frac{1}{2}}\right) h\left(u\right) du\ dt$$

$$= \int_{-\infty}^{x}\int_{0}^{\infty}\int_{-\infty}^{\lambda t u^{\frac{1}{2}}} 2\phi\left(t|0,u^{-1}\right)\phi\left(w\right) h\left(u\right) dw\ du\ dt$$

$$= 2\int_{-\infty}^{x}\int_{0}^{\infty}\int_{-\infty}^{0} \sqrt{1+\lambda^2}\phi\left(t|0,u^{-1}\right)\phi\left(v\sqrt{1+\lambda^2}+\lambda t u^{\frac{1}{2}}\right) h\left(u\right) dv\ du\ dt \tag{B.3}$$

$$= 2\int_{-\infty}^{x}\int_{0}^{\infty}\int_{-\infty}^{0} \frac{\sqrt{1+\lambda^2}}{2\pi}u^{-1/2} exp\left\{-\frac{1}{2}\left[\left(1+\lambda^2\right)t^2u + 2\lambda\sqrt{1+\lambda^2}tu^{\frac{1}{2}}v + \left(1+\lambda^2\right)v^2\right]\right\} h\left(u\right) dv\ du\ dt$$

$$= 2\int_{0}^{\infty}\int_{-\infty}^{x}\int_{-\infty}^{0} \frac{\sqrt{1+\lambda^2}}{2\pi}u^{-1/2}$$

$$\times exp\left\{-\frac{1}{2}\left[\begin{pmatrix} tu^{\frac{1}{2}} \\ v\end{pmatrix}^{\top}\begin{pmatrix} 1+\lambda^2 & \lambda\sqrt{1+\lambda^2} \\ \lambda\sqrt{1+\lambda^2} & 1+\lambda^2\end{pmatrix}\begin{pmatrix} tu^{\frac{1}{2}} \\ v\end{pmatrix}\right]\right\} h\left(u\right) dv\ dt\ du$$

$$= 2\int_{0}^{\infty}\int_{-\infty}^{xu^{\frac{1}{2}}}\int_{-\infty}^{0} \frac{1}{2\pi\sqrt{1-\delta^2}} exp\left\{-\frac{1}{2}\left[\begin{pmatrix} t_1 \\ v\end{pmatrix}^{\top}\begin{pmatrix} 1 & -\delta \\ -\delta & 1\end{pmatrix}^{-1}\begin{pmatrix} t_1 \\ v\end{pmatrix}\right]\right\} h\left(u\right) dv\ dt_1\ du \tag{B.4}$$

$$= 2\int_{0}^{\infty} \boldsymbol{\Phi}_2\left(\boldsymbol{x}\left(\boldsymbol{u}\right)^*|\boldsymbol{0},\boldsymbol{\Sigma}\right) h\left(u\right)\ du, \tag{B.5}$$

where $\boldsymbol{x}\left(\boldsymbol{u}\right)^* = (xu^{\frac{1}{2}}, 0)^{\top}$ and $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & -\delta \\ -\delta & 1\end{pmatrix}$, with $\delta = \dfrac{\lambda}{\sqrt{1+\lambda^2}}$. Equation (B.2) is obtained using the pdf of X given in (5). Equation (B.3) is obtained using the transformation $v = \frac{w-\lambda t u^{\frac{1}{2}}}{\sqrt{1+\lambda^2}}$. We obtain Equation (B.4) considering the transformation $t_1 = tu^{\frac{1}{2}}$. The cdf of Y is obtained substituting (B.5) in (B.1). □

Appendix C. Derivations of quantities $\boldsymbol{E}_{\phi_{SN}}\left(\boldsymbol{r},\boldsymbol{q}\right)$, $\boldsymbol{E}_{\Phi_{SN}}\left(\boldsymbol{r},\boldsymbol{q}\right)$, $\boldsymbol{E}_{\phi}\left(\boldsymbol{r},\boldsymbol{q}\right)$ and $\boldsymbol{E}_{\Phi}\left(\boldsymbol{r},\boldsymbol{q}\right)$ for SNI distributions

In this appendix, we calculate the expressions for the expected values $E_{\phi_{SN}}\left(r,q\right)$, $E_{\Phi_{SN}}\left(r,q\right)$, $E_{\phi}\left(r,q\right)$ and $\mathrm{E}_{\Phi}\left(r,q\right)$ given in Section 4, for the SNI distributions.

- **Skew-t distribution**

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

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

where $U' \sim \mathit{Gamma}(\alpha_1,\alpha_2)$, $\mathbf{y}^* = (q,0)^\top$, $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}$. (C.1) is obtained from Lemma 1 and (C.2) is obtained using Lemma 6.

$$
\begin{aligned}
\mathrm{E}_{\Phi}(r,q) &= \mathrm{E}\left[U^r \Phi(q\,U^{1/2})\right] = \int_0^\infty \frac{u^{\frac{2r+\nu}{2}-1}\Phi\left(qu^{\frac{1}{2}}\right)\nu^{\frac{\nu}{2}}}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \exp\left\{-\frac{u\nu}{2}\right\} du \\
&= \frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{2}{\nu}\right)^r \int_0^\infty \Phi\left(qu^{\{\frac{1}{2}\}}\right)\frac{1}{\Gamma(\alpha_1)}\alpha_2^{\alpha_1}u^{\{\alpha_1-1\}}\exp\{-u\alpha_2\}\,du \\
&= \frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{2}{\nu}\right)^r \mathrm{E}\left[\Phi\left(qU'^{\{\frac{1}{2}\}}\right)\right] \\
&= \frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{2}{\nu}\right)^r \Phi_{PVII}(q|\nu+2r,\nu). \qquad \text{(C.3)}
\end{aligned}
$$

Equation (C.3) is obtained using Lemma 3 of Genç (2013), where $\Phi_{PVII}(\cdot)$ represents the cdf of the standard Pearson type VII distribution.

$$
\begin{aligned}
\mathrm{E}_{\phi_{SN}}(r,q) &= E\left[U^r\phi_{SN}\left(q\,U^{\frac{1}{2}}|\lambda\right)\right] \\
&= \int_0^\infty u^r 2\phi\left(qu^{1/2}\right)\Phi\left(\lambda q u^{1/2}\right)\frac{1}{\Gamma\left(\frac{\nu}{2}\right)}u^{\frac{\nu}{2}-1}\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}\exp\left\{-\frac{u\nu}{2}\right\}du \\
&= \frac{2}{\sqrt{2\pi}}\frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}\left(\frac{q^2+\nu}{2}\right)^{-\frac{\nu+2r}{2}}\int_0^\infty \Phi\left(\lambda\,q\,u^{\{\frac{1}{2}\}}\right)\frac{\alpha_3^{\alpha_1}}{\Gamma(\alpha_1)}u^{\{\alpha_1-1\}}\exp\{-\alpha_3 u\}\,du \\
&= \frac{2}{\sqrt{2\pi}}\frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}\left(\frac{q^2+\nu}{2}\right)^{-\frac{\nu+2r}{2}}\mathrm{E}\left[\Phi\left(\lambda\,q\,U''^{\{\frac{1}{2}\}}\right)\right]+ \\
&= \frac{2^{1+r}}{\sqrt{2\pi}}\frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\nu^{\frac{\nu}{2}}\left(q^2+\nu\right)^{-\frac{\nu+2r}{2}}\mathbf{T}\left(\sqrt{\frac{2r+\nu}{q^2+\nu}}\lambda\,q\,|2r+\nu\right), \qquad \text{(C.4)}
\end{aligned}
$$

where $U'' \sim \mathit{Gamma}(\alpha_1,\alpha_3)$. Equation (C.4) is obtained using Lemma 6.

$$
\begin{aligned}
\mathrm{E}_{\phi}(r,q) &= E\left[U^r\phi\left(q\,U^{\frac{1}{2}}\right)\right] = \int_0^\infty \frac{\nu^{\frac{\nu}{2}}u^{\frac{\nu}{2}-1}u^r}{\sqrt{2\pi}\Gamma\left(\frac{\nu}{2}\right)2^{\frac{\nu}{2}}}\exp\left\{-\frac{u(q^2+\nu)}{2}\right\}du \\
&= \frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{\nu^{\frac{\nu}{2}}}{\sqrt{2\pi}2^{\frac{\nu}{2}}}\left(\frac{q^2+\nu}{2}\right)^{-\frac{\nu+2r}{2}}\int_0^\infty \frac{\alpha_2^{\alpha_1}u^{\{\alpha_1-1\}}}{\Gamma(\alpha_1)}\exp\{-\alpha_2 u\}\,du \qquad \text{(C.5)}\\
&= \frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{1}{\sqrt{2\pi}}\left(\frac{\nu}{2}\right)^{\nu/2}\left(\frac{q^2+\nu}{2}\right)^{-\frac{\nu+2r}{2}},
\end{aligned}
$$

where the integrand in (C.5) is the pdf of a random variable with distribution $Gamma\,(\alpha_1, \alpha_2)$.

- **Skew-slash distribution**

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

$$\begin{aligned}
\mathrm{E}_{\Phi_{SN}}(r,q) &= \mathrm{E}\left[U^r \Phi_{SN}(q\,U^{1/2}|\lambda)\right] = \int_0^1 u^r \Phi_{SN}\left(qu^{\frac{1}{2}}|\lambda\right)\nu u^{\nu-1}du \\
&= \left(\frac{\nu}{\nu+r}\right)\int_0^1 \Phi_{SN}\left(qu^{\frac{1}{2}}|\lambda\right)(\nu+r)\,u^{(\nu+r)-1}du \\
&= \left(\frac{\nu}{\nu+r}\right)\mathrm{E}\left[\Phi_{SN}\left(qU'^{\{\frac{1}{2}\}}|\lambda\right)\right] \\
&= \left(\frac{2\nu}{\nu+r}\right)\mathrm{E}\left[\Phi_2\left(U'^{\{\frac{1}{2}\}}\boldsymbol{y}^*|\mathbf{0},\boldsymbol{\Sigma}\right)\right], \qquad \text{(C.6)}
\end{aligned}$$

where $\mathbf{y}^* = (q,0)^\top$, $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}$ and $U' \sim$ Beta$(\nu+r,1)$.

$$\begin{aligned}
E_{\Phi}(r,q) &= \mathrm{E}\left[U^r \Phi(q\,U^{1/2})\right] = \int_0^1 u^r \Phi\left(qu^{\frac{1}{2}}\right)\nu u^{\nu-1}du \\
&= \left(\frac{\nu}{\nu+r}\right)\int_0^1 \Phi\left(qu^{\frac{1}{2}}\right)(\nu+r)\,u^{(\nu+r)-1}du \\
&= \left(\frac{\nu}{\nu+r}\right)\mathrm{E}\left[\Phi\left(qU'^{\{\frac{1}{2}\}}\right)\right] \\
&= \left(\frac{\nu}{\nu+r}\right)\Phi_{SL}(q|\nu+r), \qquad \text{(C.7)}
\end{aligned}$$

with $U' \sim$ Beta$(\nu+r,1)$. Using Lemma 3 of Genç (2013), we obtain Equation (C.7) where $\Phi_{SL}(\cdot|\nu+r)$ is the cdf of the standard slash distribution.

To facilitate the notation, let us make $\alpha_1 = \nu + r$ and $\alpha_2 = q^2/2$. Thus

$$\begin{aligned}
E_{\phi_{SN}}(r,q) &= E\left[U^r \phi_{SN}\left(q\,U^{\frac{1}{2}}|\lambda\right)\right] = \int_0^1 u^r \phi_{SN}\left(q\,u^{\frac{1}{2}}|\lambda\right)\nu u^{\nu-1}du \\
&= \int_0^1 2\phi\left(q\,u^{\frac{1}{2}}\right)\Phi\left(\lambda\,q\,u^{\frac{1}{2}}\right)\nu u^{(r+\nu)-1}du \\
&= \left(\frac{2\nu}{\sqrt{2\pi}}\right)\Gamma(\nu+r)\left(\frac{q^2}{2}\right)^{-(\nu+r)} G\left(1|r+\nu,\frac{q^2}{2}\right) \\
&\times \int_0^1 \frac{1}{G(1|\alpha_1,\alpha_2)}\Phi\left(\lambda\,q\,u^{\frac{1}{2}}\right)\frac{1}{\Gamma(\alpha_1)}(\alpha_2)^{\alpha_1}u^{\alpha_1-1}\exp\{-\alpha_2 u\}\,du \\
&= \left(\frac{2\nu}{\sqrt{2\pi}}\right)\Gamma(\nu+r)\left(\frac{q^2}{2}\right)^{-(\nu+r)} G\left(1|r+\nu,\frac{q^2}{2}\right)\mathrm{E}\left[\Phi\left(\lambda\,q\,U''^{\frac{1}{2}}\right)\right], \qquad \text{(C.8)}
\end{aligned}$$

where $G(\cdot|\nu+r,\frac{q^2}{2})$ represents the cdf of the gamma distribution, with parameters $\nu+r$ and $q^2/2$. $U'' \sim TGamma\,(\alpha_1,\alpha_2,\lfloor 0,1 \rfloor)$.

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

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

$$\begin{aligned}
\mathrm{E}_{\Phi_{SN}}(r,q) &= \mathrm{E}\left[U^r \Phi_{SN}\left(q\,U^{1/2}|\lambda\right)\right] = u^r \Phi_{SN}\left(qu^{\frac{1}{2}}|\lambda\right)\left[\gamma \mathbb{I}_{\{\xi\}}(u) + (1-\gamma)\mathbb{I}_{\{1\}}(u)\right] \\
&= 2\gamma\xi^r \Phi_2\left(\xi^{\frac{1}{2}}\boldsymbol{y}^*|\mathbf{0},\boldsymbol{\Sigma}\right) + 2(1-\gamma)\Phi_2\left(\boldsymbol{y}^*|\mathbf{0},\boldsymbol{\Sigma}\right) \\
&= \xi^r 2\left[\gamma\Phi_2\left(\xi^{\frac{1}{2}}\boldsymbol{y}^*|\mathbf{0},\boldsymbol{\Sigma}\right) + (1-\gamma)\,\Phi_2\left(\boldsymbol{y}^*|\mathbf{0},\boldsymbol{\Sigma}\right)\right] + 2\,(1-\gamma)\,(1-\xi^r)\,\Phi_2\left(\boldsymbol{y}^*|\mathbf{0},\boldsymbol{\Sigma}\right) \\
&= \xi^r \Phi_{SCN}\left(q|\lambda,(\gamma,\xi)\right) + 2\,(1-\gamma)\,(1-\xi^r)\,\Phi_2\left(\boldsymbol{y}^*|\mathbf{0},\boldsymbol{\Sigma}\right), \qquad \text{(C.9)}
\end{aligned}$$

where $\mathbf{y}^* = (q,0)^\top$, $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}$. We used result (22) to obtain Equation (C.9).

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

In this case, $\Phi_{CN}(\cdot|(\gamma,\xi))$ is the cdf of the standard contaminated normal distribution, as presented in Genç (2013, Eqn. (65)) and Garay et al. (2015, Sec. 2).

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

$$\begin{aligned}
\mathrm{E}_{\phi}(r,q) &= \mathrm{E}\left[U^r \phi\left(q\,U^{1/2}\right)\right] = u^r \phi\left(qu^{\frac{1}{2}}\right)\left[\gamma \mathbb{I}_{\{\xi\}}(u) + (1-\gamma)\mathbb{I}_{\{1\}}(u)\right] \\
&= \gamma\xi^r \phi\left(q\xi^{\frac{1}{2}}\right) + (1-\gamma)\phi(q)\,.
\end{aligned}$$

References

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

Ash, R. B., 2000. Probability and Measure Theory. Academic Press, San Diego.

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

Branco, M. D., Dey, D. K., 2001. A general class of multivariate skew-elliptical distributions. Journal of Multivariate Analysis 79, 99–113.

Branco, M. D., Genton, M. G., Liseo, B., 2013. Objective Bayesian analysis of skew-t distributions. Scandinavian Journal of Statistics 40 (1), 63–85.

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

Flecher, C., Allard, D., Naveau, P., 2010. Truncated skew-normal distributions: moments, estimation by weighted moments and application to climatic data. Metron 68, 331–345.

Garay, A. M., Lachos, V. H., Bolfarine, H., Cabral, C. R. B., 2015. Linear censored regression models with scale mixtures of normal distributions. Statistical Papers, DOI: 10.1007/s00362–015–0696–9.

Genç, A. İ., 2013. Moments of truncated normal/independent distributions. Statistical Papers 54, 741–764.

Ho, H. J., Lin, T.-I., Chen, H.-Y., Wang, W.-L., 2012. Some results on the truncated multivariate t distribution. Journal of Statistical Planning and Inference 142 (1), 25–40.

Kim, H.-J., 2008. Moments of truncated Student-t distribution. Journal of the Korean Statistical Society 37, 81–87.

Liu, C., Rubin, D. B., 1994. The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence. Biometrika 80, 267–278.

Massuia, M. B., Cabral, C. R. B., Matos, L. A., Lachos, V. H., 2015. Influence diagnostics for Student-t censored linear regression models. Statistics 49, 1074–1094.

Novomestky, F., Nadarajah, S., 2012. truncdist: Truncated Random Variables. R package version 1.0-1. URL `https://CRAN.R-project.org/package=truncdist`

R Development Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
URL `http://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., et al., 1988. Using the SIR algorithm to simulate posterior distributions. Bayesian statistics 3, 395–402.