Chilean Journal of Statistics
Vol. 3, No. 2, September 2012, 193–212

PROBABILISTIC AND INFERENTIAL ASPECTS OF SKEW-SYMMETRIC MODELS
SPECIAL ISSUE: "IV INTERNATIONAL WORKSHOP IN HONOUR OF ADELCHI AZZALINI'S 60TH BIRTHDAY"

On skewed continuous $l_{n,p}$-symmetric distributions

REINALDO B. ARELLANO-VALLE[1] AND WOLF-DIETER RICHTER[2,*]

[1]Departamento de Estadística, Pontificia Universidad Católica de Chile, Santiago, Chile
[2]Institute of Mathematics, University of Rostock, Rostock, Germany

(Received: 05 November 2011 · Accepted in final form: 28 June 2012)

Abstract

The general methods from theory of skewed distributions and from the theory of geometric and stochastic representations of $l_{n,p}$-symmetric distributions are combined here to introduce skewed continuous $l_{n,p}$-symmetric distributions.

Keywords: Conditional and marginal distributions · Dirichlet distribution · $l_{n,p}$-symmetric distributions · Scale mixture of the $N_{n,p}$ distribution · p-generalized Student-t distribution · Skewed distributions · Stochastic representations.

Mathematics Subject Classification: Primary 60E99 · Secondary 62E99.

1. INTRODUCTION

The univariate skew-normal and its extension to a univariate skew-symmetric distribution were introduced in Azzalini (1985) and Azzalini (1986), respectively. Many authors extended these considerations in various aspects and in different ways. For instance, a multivariate extension of the skew-normal distribution and its main properties were discussed first in Azzalini and Dalla-Valle (1996) and then in Azzalini and Capitanio (1999). Thus, several multivariate skew-normal versions and their extensions to skew-elliptical distributions have been introduced; see, e.g., Azzalini and Capitanio (1999) and Branco and Dey (2001).

Multivariate unified skew-normal and skew-elliptical distributions were considered in Arellano-Valle and Azzalini (2006) and Arellano-Valle and Genton (2010). Genton (2004) gave an overview of these efforts. The concept of fundamental skew distributions, which unifies all current known approaches, has been developed in Arellano-Valle and Genton

*Corresponding author. Reinaldo B. Arellano-Valle. Departamento de Estadística, Pontificia Universidad Católica de Chile, V. Mackenna 4860, Santiago, Chile.
Emails: reivalle@mat.puc.cl (R. Arellano-Valle) – wolf-dieter.richter@uni-rostock.de (W-D. Richter)

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

(2005). The authors of Arellano-Valle et al. (2006a) brought a certain new structure to the widespread field and unified many different approaches from a selection point of view.

The Gaussian measure indivisible-representation was first introduced in Richter (1985) and later used in solving several problems in probability theory and mathematical statistics. An overview of such applications is given in Richter (2009). Based upon a generalized method of indivisibles, which makes use of the notion of non-Euclidean surface content, in the same paper, a more general geometric measure representation formula for $l_{n,p}$-symmetric distributions is derived. This formula enables one to derive exact distributions of several types of functions of $l_{n,p}$-symmetrically distributed random vectors. This has been demonstrated by generalizing the Fisher distribution, and also for several special cases in Richter (2007) and Kalke et al. (2012).

Here, we extend the class of skewed distributions for cases where the underlying distribution is an $l_{n,p}$-symmetric one. To this end, we first exploit stochastic representations which are based upon the geometric measure representation formula in Richter (2009) to derive marginal and conditional distributions from $l_{n,p}$-symmetric distributions. Then, the general density formula for skewed distributions from Arellano-Valle et al. (2006a) applies, and then we follow the general concept in Arellano-Valle and Azzalini (2006).

The paper is structured as follows. In Section 2, we introduce the p-generalized normal distribution $\mathrm{N}_{n,p}$ and consider partitions of correspondingly distributed random vectors. Consequently, we generalize some results on Dirichlet distributions and on moments. In Section 3, we deal with continuous $l_{n,p}$-symmetric distributions, where their moments, marginal and conditional densities are also derived and the scale mixture of the $\mathrm{N}_{n,p}$ distribution is considered. Then, we use the general ideas from Arellano-Valle et al. (2006a) and Arellano-Valle and Azzalini (2006) to introduce in the final Section 4 skewed $l_{n,p}$-symmetric densities.

2. Preliminaries

In this section, we present the generalized normal distribution, consider partitions of random vectors and generalize some results on Dirichlet distributions and on moments.

2.1 The p-generalized normal distribution

Let $X = (X_1, \ldots, X_n)^\top$ be a random vector following a p-generalized normal distribution, denoted by $X \sim \mathrm{N}_{n,p}$, which in terms of its density is defined by

$$f_X(x) = C_p^n e^{-\frac{|x|_p^p}{p}}, \quad x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n,$$

where $|x|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ and $C_p = p^{1-1/p}/2\Gamma(1/p), \ \ p > 0$. Clearly, this is equivalent to $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.), with power exponential density $C_p e^{-\frac{1}{p}|x|^p}$, for $x \in \mathbb{R}$.

Let now $R_p = |X|_p$ be the p-functional of the random vector X, which is a norm if $p \geq 1$ and an antinorm if $0 < p < 1$; see Moszyńska and Richter (2012). Since $|X_1|^p, \ldots, |X_n|^p$ are i.i.d. $G(1/p, 1/p)$ random variables, we have $R_p^p = |X|_p^p = \sum_{i=1}^p |X_i|^p \sim G(n/p, 1/p)$, where $G(\alpha, \lambda)$ denotes the gamma distribution with shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$. Hence, the random variable R_p has density given by

$$f(r) = \frac{r^{n-1} e^{-\frac{r^p}{p}}}{p^{\frac{n}{p}-1}\Gamma(\frac{n}{p})}, \quad r > 0. \tag{1}$$

As in Richter (2007), we refer this distribution by $R_p \sim \chi(p,n)$. In particular, we have

$$\mathrm{E}\big[R_p^k\big] = \frac{p^{k/p}\Gamma([n+k]/p)}{\Gamma(n/p)}, \quad \forall\, k \geq 0.$$

In addition, $(1/p)|X_i|^p \overset{\text{iid}}{\sim} G\,(1/p, 1)$, for $i = 1, \ldots, n$, following that

$$\frac{1}{p}\sum_{i=l}^{l+k-1} |X_i|^p \sim G\left(\frac{k}{p}, 1\right) \quad \text{and} \quad \frac{1}{p}R_p^p = \frac{1}{p}\sum_{i=1}^{n} |X_i|^p \sim G\left(\frac{n}{p}, 1\right).$$

Moreover, since $|X|_p^p = \sum_{i=1}^n |X_i|^p$, we straightforwardly have that

$$\left(\frac{|X_1|^p}{|X|_p^p}, \ldots, \frac{|X_n|^p}{|X|_p^p}\right)^\top \sim D_n\left(\frac{1}{p}, \ldots, \frac{1}{p}\right),$$

where $D_{m+1}(\alpha_1, \ldots, \alpha_{m+1})$, with $\alpha_i > 0$, for $i = 1, \ldots, m+1$, denotes the Dirichlet distribution. Similarly,

$$\left(\frac{|X_1|^p}{|X|_p^p}, \ldots, \frac{|X_k|^p}{|X|_p^p}, 1 - \sum_{j=1}^{k} \frac{|X_j|^p}{|X|_p^p}\right)^\top \sim D_{k+1}\left(\frac{1}{p}, \ldots, \frac{1}{p}, \frac{n-k}{p}\right),$$

for $k \in \{1, \ldots, n-1\}$, and the sub-vector

$$(Y_1, \ldots, Y_k)^\top = \left(\frac{|X_1|^p}{|X|_p^p}, \ldots, \frac{|X_k|^p}{|X|_p^p}\right)^\top$$

has a density

$$h_k(y_1, \ldots, y_k) = \frac{\Gamma\left(\frac{n}{p}\right)}{\Gamma\left(\frac{1}{p}\right)^k \Gamma\left(\frac{n-k}{p}\right)} \prod_{i=1}^{k} y_i^{\frac{1}{p}-1} \left(1 - \sum_{i=1}^{k} y_i\right)^{\frac{n-k}{p}-1},$$

for $y_1 > 0, \ldots, y_k > 0$, and $\sum_1^k y_i < 1$. Hence, the density of $(Z_1, \ldots, Z_k)^\top = ((|X_1|)/(|X|_p^p), \ldots, (|X_k|)/(|X|_p^p))^\top$ is

$$g_k(z_1, \ldots, z_k) = \frac{\partial^k}{\partial z_1 \cdots \partial z_k}\mathrm{P}\left(\frac{|X_i|}{|X|_p} \leq z_i, i = 1, \ldots, k\right) = h_k(z_1^p, \ldots, z_k^p)\prod_{i=1}^{k}(p z_i^{p-1}),$$

and the following lemma has thus been proved.

Lemma 2.1 The density of $(Z_1, \ldots, Z_k)^\top = ((|X_1|)/(|X|_p^p), \ldots, (|X_k|)/(|X|_p^p))^\top$, where $X = (X_1, \ldots, X_n)^\top \sim \mathrm{N}_{n,p}$, is

$$g_k(z_1, \ldots, z_k) = \frac{\Gamma\left(\frac{n}{p}\right)\left(\frac{p}{2}\right)^k}{\Gamma\left(\frac{1}{p}\right)^k \Gamma\left(\frac{n-k}{p}\right)} \left(1 - \sum_{i=1}^{k} z_i^p\right)^{\frac{n-k}{p}-1}, \quad z_1 > 0, \ldots, z_k > 0,\ \sum_1^k z_i < 1.$$

This is a generalization of Formula (1.26) in Fang et al. (1990).

2.2 Stochastic representation of a partitioned p-generalized normally distributed random vector

It is known from Richter (2009) that $X \sim \mathrm{N}_{n,p}$ allows the stochastic representation

$$X \stackrel{d}{=} RU_p,$$

where $R \stackrel{d}{=} R_p$ and it is independent of $U_p \stackrel{d}{=} X/R_p$, which follows a p-generalized uniform distribution (i.e., the uniform distribution with respect to the p-generalized surface content on the p-generalized unit sphere $\mathcal{S}_{n,p} = \{x \in \mathbb{R}^n\colon |x|_p = 1\}$). Now, consider the partition of X

$$X = (X^{(1)\top}, X^{(2)\top})^\top,$$

where $X^{(1)} \in \mathbb{R}^k$ and $X^{(2)} \in \mathbb{R}^{n-k}$, $0 < k < n$. Similarly, we partition

$$U_p = (U_{p,1}^\top, U_{p,2}^\top)^\top,$$

where $U_{p,1}$ is k-dimensional and so $U_{p,2}$ is $(n-k)$-dimensional.

Lemma 2.2 The random vector U_p allows the stochastic representation

$$(U_{p,1}, U_{p,2}) \stackrel{d}{=} (R_{k,n}^{(p)} U_p^{(k)}, (1 - R_{k,n}^{(p)\,p})^{1/p} U_p^{(n-k)}),$$

where the random elements $R_{k,n}^{(p)}$, $U_p^{(k)}$ and $U_p^{(n-k)}$ are independent, $U_p^{(k)}$ and $U_p^{(n-k)}$ are any p-generalized uniformly distributed random vectors on $\mathcal{S}_{k,p}$ and $\mathcal{S}_{n-k,p}$, respectively, and $R_{k,n}^{(p)\,p}$ is any random variable such that $R_{k,n}^{(p)\,p} \sim B\,(k/p, (n-k)/p)$, where $B(\alpha, \beta)$ denotes the beta distribution with parameters $\alpha > 0$ and $\beta > 0$.

Proof The random elements

$$\frac{X^{(1)}}{|X^{(1)}|_p} = U_p^{(k)}, \quad \frac{X^{(2)}}{|X^{(2)}|_p} = U_p^{(n-k)}, \quad |X^{(1)}|_p, \quad |X^{(2)}|_p,$$

are independent. We put $R_{k,n}^{(p)} = (|X^{(1)}|_p)/(|X|_p)$. Then, $(|X^{(2)}|_p^p)/(|X|_p^p) = 1 - R_{k,n}^{(p)\,p}$ and

$$U_p^\top = (U_{p,1}^\top, U_{p,2}^\top) \stackrel{d}{=} \frac{X^\top}{|X|_p} = (R_{k,n}^{(p)} U_p^{(k)^\top}, (1 - R_{k,n}^{(p)\,p})^{1/p} U_p^{(n-k)^\top}).$$

Since $(1/p)|X^{(1)}|_p^p \sim G(k/p, 1)$ and $(1/p)|X^{(2)}|_p^p \sim G((n-k)/p, 1)$ and they are independent, we then have

$$R_{k,n}^{(p)\,p} = \frac{\frac{1}{p}|X^{(1)}|_p^p}{\frac{1}{p}|X^{(1)}|_p^p + \frac{1}{p}|X^{(2)}|_p^p} \sim B\left(\frac{k}{p}, \frac{n-k}{p}\right).$$

■

Let us remark that one may think of $R_{k,n}^{(p)}$ as, e.g., $R_{k,n}^{(p)} = (|X^{(1)}|_p)/(|X|_p)$ or as any random variable following the same distribution as $(|X^{(1)}|_p)/(|X|_p)$. This result generalizes Lemma 2 in Cambanis et al. (1981) to the case of arbitrary $p > 0$. The partition

$(X^{(1)\top}, X^{(2)\top})$ of $X^\top$ allows according to this lemma the stochastic representation

$$(X^{(1)}, X^{(2)}) \stackrel{d}{=} (RR_{k,n}^{(p)} U_p^{(k)}, R(1 - R_{k,n}^{(p)\,p})^{1/p} U_p^{(n-k)}),$$

where $R, R_{k,n}^{(p)}, U_p^{(k)}$ and $U_p^{(n-k)}$ are independent. The meaning of the nonnegative random variable R is quite different from that of the nonnegative variable $R_{k,n}^{(p)}$.

According to Richter (2007), $R_{k,n}^{(p)} = \cos_p(\phi)$, i.e., the p-generalized cosine-value of the angle ϕ between the two one-dimensional subspaces of $\mathbb{R}^n$ spanned up by $0 \in \mathbb{R}^n$ and one of the vectors X and $(X^{(1)\top}, 0^\top)^\top$. Note that ϕ only takes its values in the interval $[0, \pi/2]$.

2.3 Moments

Generalizing well known results from Fang et al. (1990) to the case of arbitrary $p > 0$, we compute some multivariate moments of a p-generalized normal vector $X \sim \mathrm{N}_{n,p}$. For this end, we first need some preliminary notations. We denote the sign of X by $\mathrm{sgn}(X) = (\mathrm{sgn}(X_1), \ldots, \mathrm{sgn}(X_n))^\top$ and its absolute value by $|X| = (|X_1|, \ldots, |X_n|)^\top$. Here, for any random variable Z, which is a.s. different from zero, the sign of Z is defined by

$$\mathrm{sgn}(Z) = \begin{cases} +1, \text{ if } Z > 0; \\ -1, \text{ if } Z < 0. \end{cases}$$

It is clear by the symmetry that the random vectors $|X|$ and $\mathrm{sgn}(X)$ are independent, and that $\mathrm{sgn}(X)$ has uniform distribution on $\{-1, +1\}^n$, $\mathrm{sgn}(X) \sim \mathrm{U}(\{-1, +1\}^n)$ say. We formalize these results in the following lemma, where the marginal distribution of $|X|$ is also given. For further properties of these random vectors in the context of a more general class of symmetric distributions, see Arellano-Valle et al. (2002) and Arellano-Valle and del Pino (2004).

Lemma 2.3 If $X \sim \mathrm{N}_{n,p}$, then $\mathrm{sgn}(X)$ and $|X|$ are independent random vectors, with $\mathrm{sgn}(X) \sim \mathrm{U}(\{-1, +1\}^n)$ and $f_{|X|}(t) = 2^n C_p^n e^{-\frac{1}{p}\sum_{i=1}^n t_i^p}$, $t = (t_1, \ldots, t_n)^\top \in \mathbb{R}_+^n$.

For any vector $s = (s_1, \ldots, s_n)^\top$, let $D(s)$ be the diagonal $n \times n$ matrix given by $D(s) = \mathrm{diag}(s_1, \ldots, s_n)$.

Lemma 2.4 If $X \sim \mathrm{N}_{n,p}$, then $X \stackrel{d}{=} D(S)T$, where S and T are independent random vectors such that $S \stackrel{d}{=} \mathrm{sgn}(X)$ and $T \stackrel{d}{=} |X|$.

Theorem 2.5 If $X \sim \mathrm{N}_{n,p}$, then for any integers $r_i \geq 0$, with $i = 1, \ldots, n$,

$$\mathrm{E}\left[\prod_{i=1}^n X_i^{r_i}\right] = \begin{cases} \dfrac{p^{\frac{1}{p}\sum_{i=1}^n r_i} \prod_{i=1}^n \Gamma\left(\frac{r_i+1}{p}\right)}{\Gamma\left(\frac{1}{p}\right)^n}, & \text{if } r_i \text{ is even for all } i = 1, \ldots, n; \\ 0, & \text{if } r_i \text{ is odd for some } i = 1, \ldots, n. \end{cases}$$

Proof By Lemma 2.3 and the independence property,

$$\mathrm{E}\left[\prod_{i=1}^n X_i^{r_i}\right] = \prod_{i=1}^n \mathrm{E}\left[S_i^{r_i}\right] \mathrm{E}\left[T_i^{r_i}\right],$$

where $\mathrm{E}\,[S_i^{r_i}]$ equals zero for r_i odd and one for r_i even. The proof follows by using

$$\mathrm{E}\,[T_i^{r_i}] = p^{r_i/p}\frac{\Gamma([r_i+1]/p)}{\Gamma(1/p)}.$$

■

Corollary 2.6 If $X \sim \mathrm{N}_{n,p}$, then $\mathrm{E}[X] = 0$ and $\mathrm{E}[XX^\top] = \sigma_p^2 I_n$, where $\sigma_p^2 = p^{2/p}\Gamma(3/p)/\Gamma(1/p)$.

Obviously for $p=2$, we have $\sigma_p^2 = 1$.

Corollary 2.7 Let $U_p = (U_1,\ldots,U_n)^\top$ be a p-generalized uniform vector on $\mathcal{S}_{n,p}$. Then, for any integer $r_i \geq 0$, with $i=1,\ldots,n$,

$$\mathrm{E}\left[\prod_{i=1}^n U_i^{r_i}\right] = \begin{cases} \dfrac{\Gamma\left(\frac{n}{p}\right)\prod_{i=1}^n\Gamma\left(\frac{r_i+1}{p}\right)}{\Gamma\left(\frac{n+\sum_{i=1}^n r_i}{p}\right)\Gamma^n\left(\frac{1}{p}\right)}, & \text{if } r_i \text{ is even for all } i=1,\ldots,n; \\ 0, & \text{if } r_i \text{ is odd for some } i=1,\ldots,n. \end{cases}$$

Proof Let $X \sim \mathrm{N}_{n,p}$ and $R_p = |X|_p$. According to Subsection 2.1 in Richter (2007), R_p follows the $\chi(p,n)$-density $f(r) = r^{n-1}e^{-\frac{r^p}{p}} / \int_0^\infty r^{n-1}e^{-\frac{r^p}{p}}\,dr$, $r>0$. Since $X = R_pU_p$, where R_p and U_p are independent, we have

$$\mathrm{E}\left[\prod_{i=1}^n X_i^{r_i}\right] = \mathrm{E}\left[R_p^{\sum_{i=1}^n r_i}\right]\mathrm{E}\left[\prod_{i=1}^n U_i^{r_i}\right],$$

and the proof follows by Theorem 2.5 and $\mathrm{E}\left[R_p^s\right] = p^{s/p}\Gamma[(n+s)/p]/\Gamma(n/p)$ for all $p>0$ and $s \geq 0$. ■

This result generalizes one in Theorem 3.3 of Fang et al. (1990).

Corollary 2.8 Let U_p be the p-generalized uniform vector on $\mathcal{S}_{n,p}$. Then, $\mathrm{E}[U_p]=0$ and $\mathrm{E}[U_pU_p^\top] = \tau_{n,p}I_n$, where $\tau_{n,p} = \Gamma(3/p)\Gamma(n/p)/(\Gamma(1/p)\Gamma[(n+2)/p])$.

This result generalizes Theorem 2.7 in Fang et al. (1990). For the proof of this corollary, we refer to Richter (2009). From Corollary 2.8 we can note that if $p=2$, then $\tau_{n,p} = 1/n$, thus following the well-known result that $\mathrm{Var}[U_p] = (1/n)I_n$.

3. Continuous $l_{n,p}$-Symmetric Distributions

In this section, we deal with continuous $l_{n,p}$-symmetric distributions and find their moments, marginal and conditional densities.

3.1 Notations for $l_{n,p}$-spherical distributions

Following the notation in Fang et al. (1990), Henschel and Richter (2002), and Richter (2009), we denote by $\mathcal{R}$ the set of all nonnegative random variables defined on the same probability space as the random variable R_p and which are independent of the p-generalized uniform random vector U_p. Let F be any cumulative distribution function (c.d.f.) of a positive random variable and put

$$L_n(F) = \{X\colon X \stackrel{d}{=} RU_p,\ R \in \mathcal{R} \text{ has distribution function } F,\\ R \text{ and } U_p \text{ are stochastically independent}\}.$$

From now on, let X denote an arbitrary element of $L_n(F)$. The random vector X is called $l_{n,p}$-symmetric or -spherical distributed, or even $l_{n,p}$-norm symmetric distributed if $p \geq 1$, and the corresponding random variable $R \in \mathcal{R}$ is called its generating variate. The assumption $X \in L_n(F)$ implies that X has a density iff R has a density. In this case, the density of X is of the form $C_p(n, g)g(\sum_{i=1}^n |x_i|^p)$, where $C_p(n, g)$ is a suitably chosen normalizing constant and $g\colon \mathbb{R}^+ \to \mathbb{R}^+$ is called the density generating function.

It is assumed that g satisfies the assumption $I_{n+2,g,p} < \infty$, where

$$I_{k,g,p} = \int_0^\infty r^{k-1} g(r^p) dr.$$

This distribution is the p-generalized normal distribution if the density generating function is $g(r) = e^{-r/p}$, for $r > 0$. In this case, we have

$$\frac{1}{I_{n,g,p}} = \frac{p^{1-n/p}}{\Gamma(n/p)}.$$

In what follows, we assume $C_p(n, g) = 1$, that is, X follows an $l_{n,p}$-symmetric distribution with density generator $g = g^{(n)}$. For an $l_{n,p}$-spherical distribution defined in this way, we use the notation $X \sim \mathrm{S}_{n,p}(g)$, and for its c.d.f. we write $F_{n,p}(\cdot\ ; g)$. Equivalently, the distribution of X is determined by the density

$$f_X(x) = g^{(n)}(|x|_p^p), \quad x \in \mathbb{R}^n.$$

It follows by definition that X allows the stochastic representation $X \stackrel{d}{=} RU_p$, where R is a non-negative random variable with density

$$f(r) = \frac{2\left(\frac{2}{p}\right)^{n-1} \Gamma\left(\frac{1}{p}\right)^n}{\Gamma(\frac{n}{p})} r^{n-1} g^{(n)}(r^p), \quad r > 0, \tag{2}$$

which is independent of the p-generalized uniform random vector U_p. The cases $p = 1, 2$ concern the Gaussian distribution and the Laplace distribution, respectively.

3.2 Marginal and conditional densities

Let $X = (X_1, \ldots, X_n)^\top \sim \mathrm{S}_{n,p}(g)$ be a $l_{n,p}$-symmetrically distributed random vector with density generator $g = g^{(n)}$. We are interested in the marginal density of $X^{(1)} = (X_1, \ldots, X_k)^\top$, for $1 \leq k < n$. The following result generalizes Theorem 2.10 and Formula (2.23) in Fang et al. (1990).

LEMMA 3.1 Let $X = (X_1, \ldots, X_n)^\top \sim \mathrm{S}_{n,p}(g)$. Then, $X^{(1)} = (X_1, \ldots, X_k)^\top \sim \mathrm{S}_{k,p}(g)$ and has density

$$\frac{\partial^m}{\partial x_1 \cdots \partial x_k} \mathrm{P}(X_i \le x_i, i = 1, \ldots, k) = g^{(k)} \left(\sum_{i=1}^{k} |x_i|^p \right),$$

where the marginal density generator $g^{(k)}$ is given by

$$g^{(k)}(u) = \frac{\left(\frac{2}{p}\right)^{n-k} \Gamma\left(\frac{1}{p}\right)^{n-k}}{\Gamma\left(\frac{n-k}{p}\right)} \int\limits_{u}^{\infty} g^{(n)}(y)(y-u)^{\frac{n-k}{p}-1} dy.$$

PROOF Since $X \stackrel{d}{=} RU$, where $R = R_p$ and $U = X/R_p$ are independent, we have

$$\begin{aligned}
\mathrm{P}(X_i \le x_i, i = 1, \ldots, k) &= \mathrm{P}\left(U_i \le \frac{x_i}{R}, i = 1, \ldots, k\right) \\
&= \int\limits_0^\infty \mathrm{P}\left(U_i \le \frac{x_i}{r}, i = 1, \ldots, k\right) \mathrm{P}(R \in dr) \\
&= \int\limits_0^\infty \int\limits_{-1}^{\frac{x_1}{r}} \cdots \int\limits_{-1}^{\frac{x_k}{r}} \frac{\partial^k}{\partial u_1 \cdots \partial u_k} \mathrm{P}(U_i \le u_i, i = 1, \ldots, k) \\
&\quad du_1 \cdots du_k \mathrm{P}(R \in dr).
\end{aligned}$$

It follows from Lemma 2.1 that

$$\mathrm{P}(X_i \le x_i, i = 1, \ldots, k) = C \int\limits_0^\infty \int\limits_{-1}^{\frac{x_1}{r}} \cdots \int\limits_{-1}^{\frac{x_k}{r}} I_{\{\sum_{i=1}^k |u_i|^p \le 1\}}(u_1, \ldots, u_k) \left(1 - \sum_{i=1}^{k} |u_i|^p\right)^{\frac{n-k}{p}-1}$$
$$du_1 \cdots du_k dF(r),$$

where F is the c.d.f. of R and $C = \Gamma\left(n/p\right)\left(p/2\right)^k / \Gamma\left(1/p\right)^k \Gamma\left((n-k)/p\right)$. Hence,

$$\begin{aligned}
\frac{\partial^k}{\partial x_1 \cdots \partial x_k} \mathrm{P}(X_i \le x_i, i = 1, \ldots, k) &= C \int\limits_0^\infty I_{\{\sum_{i=1}^k |x_i|^p \le r^p\}} \left(\frac{x_1}{r}, \ldots, \frac{x_k}{r}\right) \\
&\quad \left(1 - \sum_{i=1}^{k} |\frac{x_i}{r}|^p\right)^{\frac{n-k}{p}-1} r^{-k} dF(r) \\
&= C \int\limits_{(\sum_{i=1}^k |x_i|^p)^{1/p}}^{\infty} r^{-(n-p)} \left(r^p - \sum_{i=1}^{k} |x_i|^p\right)^{\frac{n-k}{p}-1} dF(r) \\
&= g^{(k)}\left(\sum_{i=1}^{k} |x_i|^p\right),
\end{aligned}$$

where

$$g^{(k)}(u) = \frac{\Gamma\left(\frac{n}{p}\right)\left(\frac{p}{2}\right)^{k/2}}{\Gamma\left(\frac{1}{p}\right)^{k}\Gamma\left(\frac{n-k}{p}\right)} \int\limits_{u^{1/p}}^{\infty} r^{-(n-p)}(r^p - u)^{\frac{n-k}{p}-1} dF(r).$$

It is known from Richter (2009) that $dF(r) = I_{n,g,p}^{-1} r^{n-1} g^{(n)}(r^p)dr$, $r > 0$. Hence,

$$g^{(k)}(u) = \frac{\left(\frac{p}{2}\right)^{\frac{k}{2}}\Gamma\left(\frac{n}{p}\right)}{\Gamma\left(\frac{1}{p}\right)^{k}\Gamma\left(\frac{n-k}{p}\right) pI_{n,g,p}} \int\limits_{u}^{\infty} y^{-\frac{(n-p)}{p}+\frac{(n-1)}{p}+\frac{(1-p)}{p}}(y-u)^{(n-k)/p-1} g^{(n)}(y)dy.$$

Making use of the equation $I_{n,g,p} = 1/n\pi_n(p)$, where $p \to \pi_n(p)$ and $\pi_n(p) = 2^n\Gamma^n\left(1/p\right)/np^{n-1}\Gamma\left(n/p\right)$ denotes the ball number function in Richter (2011), the result follows. ∎

Consider again the partition $X = (X^{(1)\top}, X^{(2)\top})^\top$, where as before $X^{(1)}$ and $X^{(2)}$ take values in $\mathbb{R}^k$ and $\mathbb{R}^{n-k}$, for $0 < m < n$, respectively. We now are interested in determining the conditional density $f_{X^{(1)}|X^{(2)}=x^{(2)}}(x^{(1)})$ of $X^{(1)}$ given $X^{(2)} = x^{(2)}$.

It follows from Lemma 3.1 that $X^{(2)}$ has a continuous $l_{n-k,p}$-symmetric distribution with a density generator $g^{(n-k)}$ satisfying the representation

$$g^{(n-k)}(u) = \frac{\left(\frac{2}{p}\right)^{k}\Gamma\left(\frac{1}{p}\right)^{k}}{\Gamma\left(\frac{k}{p}\right)} \int\limits_{u}^{\infty} g^{(n)}(y)(y-u)^{\frac{k}{p}-1}dy = \frac{\left(\frac{2}{p}\right)^{k}\Gamma\left(\frac{1}{p}\right)^{k}}{\Gamma\left(\frac{k}{p}\right)} \int\limits_{0}^{\infty} g^{(n)}(z+u)z^{\frac{k}{p}-1}dz.$$

Hence,

$$f_{X^{(1)}|X^{(2)}=x^{(2)}}(x^{(1)}) = \frac{g^{(n)}(|x^{(1)}|_p^p + |x^{(2)}|_p^p)}{g^{(n-k)}(|x^{(2)}|_p^p)} =: g_{[a]}^{(k)}(|x^{(1)}|_p^p),$$

where $a = |x^{(2)}|_p^p$. The following lemma has thus been proved.

Lemma 3.2 Let $X = (X^{(1)\top}, X^{(2)\top})^\top$ follow the $l_{n,p}$-symmetric distribution with the density generator $g^{(n)}$. The conditional density of $X^{(1)}$ given $X^{(2)} = x^{(2)}$ is then a $l_{k,p}$-symmetric density satisfying the representation

$$f_{X^{(1)}|X^{(2)}=x^{(2)}}(x^{(1)}) = g_{[a]}^{(k)}(|x^{(1)}|_p^p), \quad a = |x^{(2)}|_p^p,$$

with the uniquely defined conditional density generator

$$g_{[a]}^{(k)}(u) = \frac{\left(\frac{p}{2}\right)^{k}\Gamma\left(\frac{k}{p}\right)g^{(n)}(a+u)}{\Gamma\left(\frac{1}{p}\right)^{k}\int\limits_{0}^{\infty} g^{(n)}(z+a)z^{\frac{k}{p}-1}dz}.$$

In other words, we have $(X^{(1)}|\, X^{(2)} = x^{(2)}) \sim \mathrm{S}_{k,p}\left(g_{[|x^{(2)}|_p^p]}^{(k)}\right)$.

This lemma generalizes a corresponding formula in Section 2.4 of Fang et al. (1990). In the special case of the generalized $\mathrm{N}_{n,p}$ distribution, $g^{(n)}(u) = C_p^n e^{-\frac{u}{p}}$, $u > 0$, Lemma 3.2 yields $g_{[a]}^{(k)}(u) = g^{(m)}(u) = C_p^m e^{-\frac{u}{p}}$, $u > 0$, for all $a > 0$.

According to the stochastic representation in Subsection 2.2 it may be remarked here that the components $(1-a^p)^{1/p}U_p^{(k)}$ and $aU_p^{(n-k)}$ of the vector $((1-a^p)^{1/p}U_p^{(k)}, aU_p^{(n-k)})$ are obviously independent. Moreover, the stochastic representation from the end of Subsection 2.2 may be reformulated as follows.

Corollary 3.3 If the random vector $X = (X^{(1)\top}, X^{(2)\top})^\top$ follows a continuous $l_{n,p}$-symmetric distribution, then the following statements are true:

(a) The sub-vectors $X^{(1)}$ and $X^{(2)}$ allow the stochastic representations $X^{(1)} \stackrel{d}{=} R_1 U_p^{(k)}$ and $X^{(2)} \stackrel{d}{=} R_2 U_p^{(n-k)}$, where $R_1 \stackrel{d}{=} RR_{k,n}^{(p)}$, $R_2 \stackrel{d}{=} R(1 - R_{k,n}^{(p)p})^{1/p}$, and where (R_1, R_2), $U_p^{(k)}$ and $U_p^{(n-k)}$ are independent.

(b) A random vector following the conditional distribution of $X^{(1)}$ given $X^{(2)} = x^{(2)}$ allows the stochastic representation $(X^{(1)} | X^{(2)} = x^{(2)}) \stackrel{d}{=} R_{[|x^{(2)}|_p^p]} U_p^{(k)}$, where, for each fixed $|x^{(2)}|_p$, the random variable $R_{[|x^{(2)}|_p^p]} \stackrel{d}{=} (R^p - |x^{(2)}|_p^p)^{1/p}$ and it is independent of $U_p^{(k)}$.

(c) The vectors $X^{(1)}$ and $X^{(2)}$ are conditionally independent given $|X^{(2)}|_p$, that is,

$$X^{(1)} \perp\!\!\!\perp X^{(2)} | \, |X^{(2)}|_p.$$

Proof The assertion in (a) is known from Subsection 2.2. Statement (b) is, because of the geometric measure representation theorem in Richter (2009), just a reformulation of the distributional statement in Lemma 3.2. From (b), it follows that

$$(X^{(1)} | \, |X^{(2)}|_p = a) \stackrel{d}{=} R_{[a^p]} U_p^{(k)}.$$

Moreover,

$$(X^{(2)} | \, |X^{(2)}|_p = a) \stackrel{d}{=} aU_p^{(m)}$$

and

$$(X^{(1)}, X^{(2)} | \, |X^{(2)}|_p = a) \stackrel{d}{=} (R_{[a^p]} U_p^{(k)}, aU_p^{(m)}),$$

where $R_{[a^p]} U_p^{(k)}$ and $aU_p^{(m)}$ are independent. ∎

The first part of this corollary generalizes Formula (2.6.9) of Theorem 2.6.6 in Fang and Zhang (1990). The part (b) generalizes (2.29)–(2.30) of Theorem 2.13 in Fang et al. (1990). Part (c) is a consequence of (b) and generalizes the same result for spherical distributions; see, e.g., Arellano-Valle et al. (2006b).

3.3 Scale mixture of the $\mathrm{N}_{n,p}$ distribution

Let $R = V^{-1/p} R_p$, where $R_p \sim \chi(n,p)$ is independent of V, which is a non-negative mixing variable with c.d.f. G that does not depend on n. Suppose that R is independent of $U^{(n)}$, the p-generalized uniform vector of $\mathbb{R}^n$. Then, the random vector defined by $Y = RU^{(n)} = V^{-1/p} R_p U^{(n)} = V^{-1/p} X$, where $X \sim \mathrm{N}_{n,p}$ is independent of $V \sim G$. We then have $Y \sim \mathrm{S}_{n,p}(g^{(n)})$, where the generator function $g^{(n)}$ is defined below. Conditioning on $V = v$, we have by Equation (1) that the density of $R = V^{-1/p} R_p$ is

$$f(r) = \frac{r^{n-1}}{p^{\frac{n}{p}-1}\Gamma(\frac{n}{p})} \int_0^\infty v^{\frac{n}{p}} e^{-\frac{r^p}{p} v} dG(v), \quad r > 0.$$

Since the functions f_R and $g^{(n)}$ satisfy Equation (2), it follows that

$$g^{(n)}(u) = \frac{\left(\frac{p}{2}\right)^n}{p^{\frac{n}{p}} \Gamma\left(\frac{1}{p}\right)^n} \int_0^\infty v^{\frac{n}{p}} e^{-\frac{r^p}{p} v} dG(v).$$

This density generator function defines an important class of $l_{n,p}$-symmetric distributions, which extends the scale mixtures of normal distributions to the scale mixtures of p-generalized normal distributions. An important member is the n-dimensional p-generalized Student-t distribution, with $\nu > 0$ degrees of freedom, denoted by $Y \sim t_{n,p}(\nu)$, for which $V \sim G(\nu/p, \nu/p)$. In this case,

$$g^{(n)}(u) = \frac{\left(\frac{p}{2}\right)^n \left(\frac{\nu}{p}\right)^{\frac{\nu}{p}}}{p^{\frac{n}{p}} \Gamma\left(\frac{1}{p}\right)^n \Gamma\left(\frac{\nu}{p}\right)} \int_0^\infty v^{\frac{n+\nu}{p}-1} e^{-\frac{\nu+u}{p} v} dv.$$

Hence, Y follows a $l_{n,p}$-symmetric distribution with density

$$f_Y(y) = D_{n,p,\nu} \left(1 + \frac{|y|_p^p}{\nu}\right)^{-\frac{\nu+n}{p}}, \quad y \in \mathbb{R}^n,$$

where

$$D_{n,p,\nu} = \frac{\left(\frac{p}{2}\right)^n \Gamma\left(\frac{\nu+n}{p}\right)}{\Gamma\left(\frac{\nu}{p}\right) \Gamma\left(\frac{1}{p}\right)^n \nu^{\frac{n}{p}}}.$$

Definition 3.4 The distribution of a random vector Y with density

$$t_{n,p}(y;\nu) := D_{n,p,\nu} \left(1 + \frac{|y|_p^p}{\nu}\right)^{-\frac{\nu+n}{p}}, \quad y \in \mathbb{R}^n,\ p > 0,\ \nu > 0,$$

is called the n-dimensional p-generalized Student-t distribution with ν degrees of freedom.

The class of p-generalized Student-t densities given in Definition 3.4 was introduced in Richter (2007) for $n = 1$; for $p = 2$, see Arellano-Valle and Bolfarine (1995). It follows from there that, in the case of the p-generalized Student-t distribution, one can think

$$V = \frac{|Z_1|^p + \cdots + |Z_\nu|^p}{\nu}, \quad (Z_1, \ldots, Z_\nu)^\top \sim \mathrm{N}_{n,p} \text{ in } \mathbb{R}^\nu.$$

The following theorem has thus been proved.

THEOREM 3.5 Let $Y = (Y^{(1)\top}, Y^{(2)\top})^\top \sim \mathrm{N}_{n+\nu,p}$, where $Y^{(1)}$ and $Y^{(2)}$ take values in $\mathbb{R}^n$ and $\mathbb{R}^\nu$, respectively. Then, $(\nu^{1/p})/(|Y^{(2)}|_p)Y^{(1)}$ follows the density $t_{n,p}(y;\nu)$, for $y \in \mathbb{R}^n$.

If $Y = (Y^{(1)\top}, Y^{(2)\top})^\top \sim t_{n,p}(\nu)$, where $Y^{(1)} \in \mathbb{R}^k$ and $Y^{(2)} \in \mathbb{R}^{n-k}$ $(0 < k < n)$, then we have by construction that the density generator of $Y^{(1)}$ satisfies the representation

$$g^{(k)}(u) = D_{k,p,\nu}\left(1 + \frac{u}{\nu}\right)^{-(\nu+k)/p},$$

that is, $Y^{(1)} \sim t_{k,p}(\nu)$, with density $t_{k,p}(y^{(1)}, \nu)$.

The conditional density of $Y^{(1)}$ given $Y^{(2)} = y^{(2)}$ is therefore

$$f_{Y^{(1)}|Y^{(2)}=y^{(2)}}(y^{(1)}) = \left(\frac{\nu+n-k}{\nu+a}\right)^{\frac{k}{p}} t_{k,p}\left(\left(\frac{\nu+n-k}{\nu+a}\right)^{\frac{1}{p}} y^{(1)}; \nu+n-k\right),$$

with $a = |y^{(2)}|_p^p$, that is, this conditional density is an $l_{k,p}$-symmetric one, but rescaled by the factor $(\nu+a)^{1/p}/(\nu+n-k)^{1/p}$.

3.4 MOMENTS

To compute the mixed moments of an $l_{n,p}$-symmetric random vector $X \sim \mathrm{S}_{n,p}$, we obtain from the stochastic representation $X \stackrel{d}{=} RU^{(n)}$ that

$$\mathrm{E}\left[\prod_{i=1}^n X_i^{r_i}\right] = \mathrm{E}\left[R^{\sum_{i=1}^n r_i}\right] \mathrm{E}\left[\prod_{i=1}^n U_i^{r_i}\right],$$

provided that $\mathrm{E}\left[R^{\sum_{i=1}^n r_i}\right]$ is finite, and where $\mathrm{E}\left[\prod_{i=1}^n U_i^{r_i}\right]$ is given in Corollary 2.7. In particular, by Corollary 2.8, we have $\mathrm{E}[X] = 0$ if $\mathrm{E}[R]$ is finite and $\mathrm{E}[XX^\top] = \sigma_{p,g}^2 I_n$, where $\sigma_{p,g}^2 = \tau_p \mathrm{E}[R^2]$, if $\mathrm{E}[R^2]$ is finite. It is convenient to emphasize here that similarly to the case of $p = 2$, the univariate variance component $\sigma_{p,g}^2 = \tau_p \mathrm{E}[R^2]$ does not depend on n.

For example, if $X \sim t_{n,p}(\nu)$, we have by Subsection 3.2 that $R = V^{-1/p} R_p$, where $V \sim G(\nu/p, \nu/p)$ and is independent of $R_p \sim \chi(n,p)$, thus implying that

$$\begin{aligned}\mathrm{E}\left[R^{\sum_{i=1}^n r_i}\right] &= \mathrm{E}\left[V^{-\sum_{i=1}^n r_i/p}\right] \mathrm{E}\left[R_p^{\sum_{i=1}^n r_i}\right] \\ &= \frac{\nu^{\frac{\sum_{i=1}^n r_i}{p}} \Gamma\left(\frac{\nu - \sum_{i=1}^n r_i}{p}\right) \Gamma\left(\frac{n + \sum_{i=1}^n r_i}{p}\right)}{\Gamma\left(\frac{\nu}{p}\right) \Gamma\left(\frac{n}{p}\right)}, \quad \nu > \sum_{i=1}^n r_i.\end{aligned}$$

Hence, for the $t_{n,p}(\nu)$-symmetric distribution, we have that, for $\nu > \sum_{i=1}^n r_i$,

$$\mathrm{E}\left[\prod_{i=1}^n X_i^{r_i}\right] = \begin{cases} \dfrac{\nu^{\frac{\sum_{i=1}^n r_i}{p}} \Gamma\left(\frac{\nu - \sum_{i=1}^n r_i}{p}\right) \prod_{i=1}^n \Gamma\left(\frac{r_i+1}{p}\right)}{\Gamma\left(\frac{\nu}{p}\right)\Gamma^n\left(\frac{1}{p}\right)}, & \text{if } r_i \text{ is even for all } i = 1, \ldots, n; \\ 0, & \text{if } r_i \text{ is odd for some } i = 1, \ldots, n. \end{cases}$$

In particular, we have $\mathrm{E}[X] = 0$ if $\nu > 1$ and $\mathrm{E}[XX^\top] = \sigma_{p,\nu}^2 I_n$ if $\nu > 2$, where

$$\sigma_{p,\nu}^2 = \frac{\nu\Gamma([\nu-2]/p)\Gamma(3/p)}{\Gamma(\nu/p)\Gamma(1/p)}.$$

3.5 Linear transformations

A further extension of the family of continuous $l_{n,p}$-spherical distributions follows by considering the distribution of the linear transformation $Y = \mu + \Gamma X$, where $X \sim \mathrm{S}_{m,p}(g)$, $\Gamma \in \mathbb{R}^{n\times m}$ and $\mu \in \mathbb{R}^n$.

We recall that a density level set (LS) is a set of points from the sample space where the density attains one and the same value which is called the density level. In the case of X, every density level set is an $l_{m,p}$-sphere, which is centered at the origin. It is reasonable to call the set $D_m \cdot \mathrm{LS}$ an axes-aligned p-generalized ellipsoid if D_m is an $m \times m$-diagonal matrix consisting of positive elements. Rotating such a set with an orthogonal $m \times m$-matrix H_m and shifting the resulting set, then in the case $m = n$ by μ leads to a set which is called a p-generalized ellipsoid with location vector μ and shape matrix $\Gamma = H_n D_n$.

Since $X \stackrel{d}{=} RU_p$, we have $Y \stackrel{d}{=} \mu + R\Gamma U_p$. The random vector Y has location vector μ and if $\Gamma = H_n D_n$ we say that Y has shape matrix Γ. If $\mathrm{E}[R^2] < \infty$, then it is straightforward to see that $\mathrm{E}[Y] = \mu$ and $\mathrm{Cov}(Y) = \sigma_{p,g}^2 \Sigma$, where $\Sigma = \Gamma\Gamma^\top$ and as was mentioned $\sigma_{p,g}^2 = \tau_p \mathrm{E}[R^2]$ is the univariate variance component induced by the density generator function $g = g^{(n)}$. Also, if $m = n$ still holds, then the random vector Y has density

$$f_Y(y) = |\Gamma|^{-1} g^{(n)}(\|\Gamma^{-1}(y-\mu)\|_p^p), \quad y \in \mathbb{R}.$$

Its c.d.f. $F_Y(y) = \mathrm{P}(Y \leq y)$ is then

$$F_Y(y) = \mathrm{P}(\mu + \Gamma X \leq y) = \int\limits_{\{x\in\mathbb{R}^n:\ \mu+\Gamma x \leq y\}} g^{(n)}(x)dx, \quad y \in \mathbb{R}^n,$$

where the sign of inequality $\leq$ means component-wise inequality. In what follows, we denote the c.d.f. of Y by $F_{n,p}(y; \mu, \Sigma, g)$, where $\Sigma = \Gamma\Gamma^\top$, or by $F_{n,p}(y; \Sigma, g)$ when $\mu = 0$, or simply by $F_{n,p}(y; g)$ when $\mu = 0$ and $\Sigma = I_n$. In the case of $p = 2$, Y has the usual elliptical distribution with location vector μ and dispersion matrix Σ and is commonly denoted by $\mathrm{EC}_n(\mu, \Sigma, g)$.

4. Skewed $l_{n,p}$-Symmetric Distributions

Next, we discuss two ways to construct skewed $l_{n,p}$-symmetric distributions.

4.1 Construction from selection mechanisms

Let $X^{(1)} \in \mathbb{R}^k$ and $X^{(2)} \in \mathbb{R}^m$ be two random vectors following a $l_{k+m,p}$-symmetric joint distribution with density generator $g^{(k+m)}$, i.e., they have joint density

$$f_{X^{(1)},X^{(2)}}(x^{(1)},x^{(2)}) = g^{(k+m)}(|x^{(1)}|_p^p + |x^{(2)}|_p^p), \quad (x^{(1)},x^{(2)}) \in \mathbb{R}^{k+m}.$$

For any fixed matrix $\Lambda \in \mathbb{R}^{m\times k}$, we study the distribution of $X^{(1)}$ when a linear random selection mechanism of the form $X^{(2)} < \Lambda X^{(1)}$ is considered. The following result characterizes the density of this particular selection distribution.

Theorem 4.1 It holds

$$f_{X^{(1)}|X^{(2)}<\Lambda X^{(1)}}(z) = \frac{1}{F_{m,p}^{(2)}\left(0; I_m + \Lambda\Lambda^\top, g^{(m)}\right)} f_{X^{(1)}}(z) F_{m,p}^{(1)}\left(\Lambda z; g_{[|z|_p^p]}^{(m)}\right), \quad z \in \mathbb{R}^m,$$

where $F_{m,p}^{(1)}(x; g_a^{(m)}) = \int_{\mathbb{R}_+^m} g_a^{(m)}(|x-u|_p^p)du$ and $F_{m,p}^{(2)}(x; \Sigma, g^{(m)})$ denotes the c.d.f. of ΓX with $\Gamma = (\Lambda, -I_m)$ and $\Sigma = \Gamma\Gamma^\top = I_m + \Lambda\Lambda^\top$.

Proof According to Lemma 3.1,

$$f_{X^{(1)}}(x^{(1)}) = g^{(k)}(|x^{(1)}|_p^p), \quad x^{(1)} \in \mathbb{R}^k.$$

With a matrix Λ: $\mathbb{R}^k \to \mathbb{R}^m$, we set

$$U_1 = X^{(1)}, \text{ and } U_2 = \Lambda X^{(1)} - X^{(2)},$$

which is equivalent to

$$X^{(1)} = U_1, \text{ and } X^{(2)} = \Lambda U_1 - U_2.$$

The Jacobian of this transformation is

$$J = \begin{vmatrix} I_k & 0 \\ \Lambda & -I_m \end{vmatrix}$$

and hence $|J| = 1$. The joint density of U_1 and U_2 is thus

$$f_{U_1,U_2}(u_1,u_2) = f_{X^{(1)},X^{(2)}}(u_1, \Lambda u_1 - u_2) = g^{(k+m)}(|u_1|_p^p + |\Lambda u_1 - u_2|_p^p).$$

Then, it follows that

$$f_{U_2|U_1=u_1}(u_2) = \frac{g^{(k+m)}(|u_1|_p^p + |\Lambda u_1 - u_2|_p^p)}{g^{(k)}(|u_1|_p^p)} = g_{[|u_1|_p^p]}^{(m)}(|\Lambda u_1 - u_2|_p^p). \tag{3}$$

We note also that an interpretation of Λ follows from the fact that $\mathrm{Cov}(U_2, U_1) = \sigma_{p,g}^2 \Lambda$.

Let Z denote a random vector, which follows the conditional distribution of $X^{(1)}$ under $X^{(2)} < \Lambda X^{(1)}$, $Z \stackrel{d}{=} (X^{(1)}|X^{(2)} < \Lambda X^{(1)})$. Then, $Z \stackrel{d}{=} (U_1|\, 0 < \Lambda X^{(1)} - X^{(2)})$ and

$$Z \stackrel{d}{=} (U_1|\, 0 < U_2).$$

By the general representation formula for the density of the corresponding conditional distribution in Arellano-Valle and del Pino (2004) or Arellano-Valle et al. (2006a),

$$f_Z(z) = f_{U_1}(z)\frac{\mathrm{P}(0 < U_2|U_1 = z)}{\mathrm{P}(0 < U_2)},$$

where $f_{U_1}(z) = f_{X^{(1)}}(z) = g^{(k)}(|z|_p^p)$. By Equation (3) and the change of variable $w = \Lambda z - u_2$, we have

$$\mathrm{P}(0 < U_2|U_1 = z) = \int\limits_{\mathbb{R}_+^m} g^{(m)}_{[|z|_p^p]}(|\Lambda z - u_2|_p^p)du_2 = F^{(1)}_{m,p}\left(\Lambda z; g^{(m)}_{[|z|_p^p]}\right).$$

Hence,

$$f_Z(z) = C_{m,p}g^{(k)}(|z|_p^p)F^{(1)}_{m,p}\left(\Lambda z; g^{(m)}_{[|z|_p^p]}\right),$$

with $1/C_{m,p} = \mathrm{P}(0 < U_2)$. Since $X \stackrel{d}{=} -X$ and $U_2 = \Gamma X$, where $\Gamma = (\Lambda, -I_m)$, then U_2 and $-U_2$ have the same distribution. Thus, $\mathrm{P}(0 < U_2) = \mathrm{P}(-U_2 < 0) = \mathrm{P}(U_2 < 0)$ and the c.d.f. of $U_2 = \Gamma X$ is denoted by $F^{(2)}_{m,p}\left(u_2; \Sigma, g^{(m)}\right)$, where $\Sigma = \Gamma\Gamma^\top = I_m + \Lambda\Lambda^\top$, such that $1/C_{m,p} = F^{(2)}_{m,p}\left(0; I_m + \Lambda\Lambda^\top, g^{(m)}\right)$. ■

Definition 4.2 The distribution of a random vector Z with density of the form

$$f_Z(z) = \frac{1}{F^{(2)}_{m,p}\left(0; I_m + \Lambda\Lambda^\top, g^{(m)}\right)}\, g^{(k)}(|z|_p^p)\, F^{(1)}_{m,p}\left(\Lambda z; g^{(m)}_{[|z|_p^p]}\right), \quad z \in \mathbb{R}^k,$$

is called skewed $l_{k,p}$-symmetric distribution with dimensionality parameter m, density generator g and skewness/shape matrix-parameter Λ. The notation $Z \sim \mathrm{SS}_{k,m,p}(\Lambda, g)$ is used for this distribution.

An important simplification is obtained when the matrix $I_m + \Lambda\Lambda^\top$ is diagonal, where $F^{(2)}_{m,p}\left(0; I_m + \Lambda\Lambda^\top, g\right) = 1/(2^m)$ by symmetry, following thus that

$$f_Z(z) = 2^m g^{(k)}(|z|_p^p)\, F^{(1)}_{m,p}\left(\Lambda z; g^{(m)}_{[|z|_p^p]}\right), \quad z \in \mathbb{R}^k.$$

The skewed $l_{k,p}$-symmetric subclass for $m = 1$ extends the skew-spherical class introduced in Branco and Dey (2001), where $p = 2$. For this subclass, the above density reduces to

$$f_Z(z) = 2g^{(k)}(|z|_p^p)\, F^{(1)}_{1,p}\left(\lambda^\top z; g^{(1)}_{[|z|_p^p]}\right), \; z \in \mathbb{R}^k,$$

for which

$$F^{(1)}_{1,p}\left(u; g^{(1)}_{[|z|_p^p]}\right)$$

is a univariate c.d.f. and is immediate to be computed numerically when it does not have an explicit expression. For $m \geq 1$, the above definition extends the analogous definition in Arellano-Valle and Genton (2005), where $p = 2$.

Corollary 4.3 The conditional distribution of $X^{(1)}$ under $X^{(2)} < \Lambda X^{(1)}$ is skewed $l_{k,p}$-symmetric with dimensionality parameter m, density generator g and skewness/shape matrix-parameter Λ.

Corollary 4.3 extends the corresponding results in Branco and Dey (2001) and Arellano-Valle and Genton (2005), which deal with the cases $m = 1, p = 2$ and $m \geq 1, p = 2$, respectively.

Example 4.4 An important special case is the skewed $\mathrm{N}_{n,p}$ distribution, where

$$g^{(k)}(|x|_p^p) = C_p^k e^{-\frac{1}{p}|x|_p^p} =: \phi_{k,p}(x)$$

is the $\mathrm{N}_{k,p}$ density and

$$F_{k,p}^{(1)}\left(x; \Sigma, g^{(k)}\right) = \int_{t<x} \phi_{k,p}(t; \Sigma) dt =: \Phi_{k,p}^{(1)}(x; \Sigma), \quad x \in \mathbb{R}^k,$$

i.e., the c.d.f. of a non-singular linear transformation $Y = \Gamma X$, with $X \sim \mathrm{N}_{n,p}$ and $\Gamma\Gamma^\top = \Sigma$. Denoting accordingly

$$F_{m,p}^{(2)}(0; I_m + \Lambda\Lambda^\top, g^{(m)}) = \Phi_{m,p}^{(2)}(0; I_m + \Lambda\Lambda^\top),$$

we say that a random vector Z has k-dimensional skew-$\mathrm{N}_{n,p}$ distribution with dimensionality parameter m and skewness/shape matrix parameter $\Lambda \in \mathbb{R}^{m\times k}$, denoted by $Z \sim \mathrm{SN}_{k,m,p}(\Lambda)$, if its density is given by

$$f_Z(z) = \frac{1}{\Phi_{m,p}^{(2)}(0; I_m + \Lambda\Lambda^\top)} \phi_{k,p}(z)\, \Phi_{m,p}^{(1)}(\Lambda z), \quad z \in \mathbb{R}^k.$$

For $m = 1$ and $p = 2$, we obtain the multivariate skew-normal density

$$f_Z(z) = 2\phi_{k,p}(z)\, \Phi_{1,p}(\lambda^\top z),\ z \in \mathbb{R}^k,$$

which was introduced in Azzalini and Dalla-Valle (1996) and studied systematically in Azzalini and Capitanio (1999). For $m = k$ with $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_k)$, the components of the skew-$\mathrm{N}_{n,p}$ random vector $Z = (Z_1, \ldots, Z_k)^\top$ are independent and have marginal densities $f_{Z_i}(z_i) = 2\phi_{1,p}(z_i)\Phi_{1,p}(\lambda_i z_i)$, $i = 1, \ldots, k$.

Example 4.5 Another important special case is the skew-$t_{k,p}(\nu)$ distribution, which is considered next, where

$$g^{(k)}(|x|_p^p) = D_{k,p,\nu}\left(1 + (|x|_p^p)/(\nu)\right)^{-(\nu+k)/p} =: t_{k,p}(x; \nu)$$

is the $t_{k,p}(\nu)$ density, and

$$F_{k,p}^{(1)}(x; \Sigma, g) = \int_{t<x} t_{k,p}(t; \Sigma, \nu) dt := T_{k,p}^{(1)}(x; \Sigma, \nu)$$

and $T_{m,p}^{(2)}$ is defined accordingly.

We say that a random vector Z has a skew-$t_{k,p}$ distribution with dimensionality parameter m and skewness/shape matrix parameter $\Lambda \in \mathbb{R}^{m\times k}$, denoted by $Z \sim \mathrm{St}_{k,m,p}(\Lambda)$, if its density is given by

$$f_Z(z) = \frac{1}{T^{(2)}_{m,p}(0; I_m + \Lambda\Lambda^\top, \nu)}\, t_{k,p}(z;\nu)\, T^{(1)}_{m,p}\left(\left(\frac{\nu+k}{\nu+|z|_p^p}\right)^{1/p}\Lambda z; \nu+k\right),\ z \in \mathbb{R}^k.$$

For $m = 1$ and $p = 2$, we have the multivariate skew-t distribution introduced in Branco and Dey (2001), Gupta (2003) and Azzalini and Capitanio (2003).

A straightforward extension follows when we consider the conditional distribution of $X^{(1)}$ given the selection mechanism $X^{(2)} < \Lambda X^{(1)} + \tau$. In such a case, we have the more general skew p-generalized $l_{k,p}$-symmetric class of densities defined by

$$f_Z(z) = \frac{1}{F^{(2)}_{m,p}\left(\tau; I_m + \Lambda\Lambda^\top g\right)}\, g^{(k)}(|z|_p^p)\, F^{(1)}_{m,p}\left(\Lambda z + \tau; g^{(m)}_{[|z|_p^p]}\right), \quad z \in \mathbb{R}^k.$$

The convenience of this more general class is because it is closed by marginalization and also by conditioning when $p = 2$, while for $\tau = 0$, it does not preserve this last property. This class generalizes the unified skew-elliptical (SUE) family obtained for $p = 2$ and studied systematically in Arellano-Valle and Genton (2010); see also Arellano-Valle and Genton (2005) and Arellano-Valle and Azzalini (2006). We call this last class the SUE-p-generalized family of distributions and most of the above results could be explored for this class.

4.2 Construction from stochastic representations

Now consider the stochastic representation

$$Z \stackrel{d}{=} X^{(1)} + \Delta|X^{(2)}|, \tag{4}$$

where $X^{(1)}$ and $X^{(2)}$ are as before, i.e., with joint $\mathrm{S}_{k+m,p}(g)$ distribution, and where $\Delta \in \mathbb{R}^{k\times m}$ is fixed matrix. Also, consider the linear transformation $W_1 = X^{(1)} + \Delta X^{(2)}$ and $W_2 = X^{(2)}$. Note that W_1 and W_2 have joint density

$$f_{W_1,W_2}(w_1,w_2) = g^{(k+m)}(|w_1 - \Delta w_2|_p^p + |w_2|_p^p),\ (w_1,w_2) \in \mathbb{R}^{k+m}.$$

Moreover, since

$$f_{X^{(1)},|X^{(2)}|}(x,t) = f_{X^{(1)},X^{(2)}|X^{(2)}>0}(x,t) = C g^{(k+m)}(|x|_p^p + |t|_p^p),\ (x,t) \in \mathbb{R}^k \times \mathbb{R}^m_+,$$

we have $(X^{(1)}, |X^{(2)}|) \stackrel{d}{=} (X^{(1)}, X^{(2)}) \mid X^{(2)} > 0$, which is equivalent to (see Arellano-Valle et al., 2002; Arellano-Valle and del Pino, 2004) $X^{(1)} \perp\!\!\!\perp \mathrm{sgn}(X^{(2)}) |\, |X^{(2)}|$. Hence, we have

$$Z \stackrel{d}{=} (X^{(1)} + \Delta X^{(2)}) \mid X^{(2)} > 0 = W_1 \mid W_2 > 0,$$

following that the density of Z is

$$
\begin{aligned}
f_Z(z) &= f_{W_1}(z)\frac{\mathrm{P}(W_2 > 0 \mid W_1 = z)}{\mathrm{P}(W_2 > 0)} \\
&= C f_{W_1}(z)\mathrm{P}(W_2 > 0 \mid W_1 = z) \\
&= C\int_{\mathbb{R}^m_+} g^{(k+m)}(|z - \Delta w|_p^p + |w|_p^p)dw,\ \ z \in \mathbb{R}^k.
\end{aligned}
$$

For $p = 2$, this density reduces to the skew-elliptical density given by

$$
f_Z(z) = 2^m g^{(k)}(Q(z))F_m\left((I_m + \Delta^\top\Delta)^{-1}\Delta^\top z; (I_m + \Delta^\top\Delta)^{-1}, g^{(m)}_{[Q(z)]}\right),
$$

where

$$
Q(z) = z^\top[I_k - \Delta(I_m + \Delta^\top\Delta)^{-1}\Delta^\top]z = z^\top(I_k + \Delta\Delta^\top)^{-1}z.
$$

For $m = k$, this skew-elliptical class of distributions was introduced in Sahu et al. (2003). For extensions of this family and its relation with other skew-elliptical families, see Arellano-Valle and Genton (2005), Arellano-Valle and Azzalini (2006) and Arellano-Valle and Genton (2010).

One of the advantages of this route to obtain multivariate skew-symmetric distributions comes from the stochastic representation given in Equation (4), which among other things allows us to easily compute the moments of Z; see Arellano-Valle et al. (2002) and Arellano-Valle and del Pino (2004). In particular, when the mean vector and covariance matrix of Z exist, we have from Equation (4) that they are given by

$$
\mathrm{E}[Z] = \Delta\mathrm{E}\left[|X^{(2)}|\right] \quad \text{and} \quad \mathrm{Cov}\,(Z) = \mathrm{Cov}(X_1^{(2)}) + \Delta\mathrm{Cov}(|X^{(2)}|)\Delta^\top,
$$

where

$$
\mathrm{Cov}(X_1^{(2)}) = \sigma^2_{p,g} I_k.
$$

To compute $\mathrm{E}[|X^{(2)}|]$ and $\mathrm{Cov}(|X^{(2)}|)$, we can use the following lemma, whose proof is straightforward from the results in Section 2.3.

Lemma 4.6 Let $X = (X_1, \ldots, X_n)^\top \sim \mathrm{S}_{n,p}(g)$ and $R = |X|_p$. Then,

$$
\mathrm{E}\left[|X_i|^r|X_j|^s\right] = \begin{cases} \dfrac{\Gamma(\frac{r+s+1}{p})}{\Gamma(\frac{1}{p})}\dfrac{\Gamma(\frac{n}{p})E(R^{r+s})}{\Gamma(\frac{n+r+s}{p})}, & i = j; \\[2ex] \dfrac{\Gamma(\frac{r+1}{p})\Gamma(\frac{s+1}{p})}{\Gamma^2(\frac{1}{p})}\dfrac{\Gamma(\frac{n}{p})E(R^{r+s})}{\Gamma(\frac{n+r+s}{p})}, & i \neq j; \end{cases}
$$

if $\mathrm{E}\left[R^{r+s}\right]$ is finite.

For the particular case of $X \sim \mathrm{N}_{n,p}$, the moments of the p-generalized normal radial random variable $R_p = |X|_p$ satisfies the relation

$$
\frac{\Gamma(\frac{n}{p})\mathrm{E}\left[R_p^k\right]}{\Gamma(\frac{n+k}{p})} = p^{\frac{k}{p}}.
$$

Hence, for the mean vector and covariance matrix of the corresponding skew-$N_{n,p}$ random vector $Z_p \stackrel{d}{=} X_p^{(1)} + \Delta|X_p^{(2)}|$, we obtain

$$\mathrm{E}\left[Z_p\right] = \frac{p^{\frac{1}{p}}\Gamma(\frac{2}{p})}{\Gamma(\frac{1}{p})}\Delta 1_k, \quad \text{and} \quad \mathrm{Cov}(Z_p) = \frac{p^{\frac{2}{p}}\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})}\left\{I_k + \left(1 - \frac{\Gamma^2(\frac{2}{p})}{\Gamma(\frac{1}{p})\Gamma(\frac{3}{p})}\right)\Delta\Delta^\top\right\}.$$

If Z is a scale-mixture of the skew-$N_{n,p}$ random vector Z_p, then there is a non-negative random variable V which is independent of Z_p such that $Z \stackrel{d}{=} V^{-1/p}Z_p$. Hence, we have

$$\mathrm{E}\left[Z\right] = \mathrm{E}\big[V^{-1/p}\big]\mathrm{E}\left[Z_p\right]$$

if $\mathrm{E}\left[V^{-1/p}\right]$ is finite and

$$\mathrm{E}\left[Z_p Z_p^\top\right] = \mathrm{E}\left[V^{-2/p}\right]\mathrm{E}\left[Z_p Z_p^\top\right]$$

if $\mathrm{E}\left[V^{-2/p}\right]$ is finite, from where we can compute $\mathrm{Cov}(Z)$.

Conclusions

In this paper, we have introduced a class of skewed continuous symmetric distributions combining the theory of skewed distributions and the theory of geometric and stochastic representations of $l_{n,p}$-symmetric distributions. Several properties and results about skewed continuous $l_{n,p}$-symmetric distributions have been provided.

Acknowledgement

The second author would like to express his gratitude to the Pontificia Universidad Católica de Chile for supporting his stay there in March 2010 and to Reinaldo B. Arellano-Valle for his kind hospitality. The first author's research was supported by FONDECYT, Chile grants 1085241 and 1120121. The authors express their thanks to the referee and the guest editors for valuable hints.

References

Arellano-Valle, R.B., Azzalini, A., 2006. On the unification of families of skew-normal distributions. Scandinavian Journal of Statistics, 33, 561–574.

Arellano-Valle, R.B., Bolfarine, H., 1995. On some characterizations of the t-distribution. Statistics and Probability Letters, 25, 79–85.

Arellano-Valle, R.B., Branco, M.D., Genton, M.G., 2006a. A unified view on skewed distributions arising from selections. The Canadian Journal of Statistics, 34, 581–601.

Arellano-Valle, R.B., del Pino, G., 2004. From symmetric to asymmetric distributions: a unified approach. In Genton, M.G. (ed.), Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality. Chapman and Hall/CRC, Boca Raton, FL.

Arellano-Valle, R.B., del Pino, G., Iglesias, P., 2006b. Bayesian inference in spherical linear models: robustness and conjugate analysis. Journal of Multivariate Analysis, 97, 179–197.

Arellano-Valle, R.B., del Pino, G., San Martín, E., 2002. Definition and probabilistic properties of skew distributions. Statistics and Probability Letters, 58, 111–121.

Arellano-Valle, R.B., Genton, M.G., 2005. On fundamental skew distributions. Journal of Multivariate Analysis, 96, 93–116.

Arellano-Valle, R.B., Genton, M.G., 2010. Multivariate unified skew-elliptical distributions. Chilean Journal of Statistics, Special issue "Tribute to Pilar Loreto Iglesias Zuazola", 1, 17–33.

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

Azzalini, A., 1986. Further results on a class of distributions which includes the normal ones. Statistica, 46, 199–208.

Azzalini, A., Capitanio, A., 1999. Statistical applications of the multivariate skew normal distributions. Journal of The Royal Statistical Society Series B - Statistical Methodology, 61, 579–602.

Azzalini, A., 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.

Azzalini, A., Dalla Valle, A., 1996. The multivariate skew-normal distribution. Biometrika, 83, 715–726.

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

Cambanis, S., Huang, S., Simons, G., 1981. On the theory of elliptically contoured distributions. Journal of Multivariate Analysis, 11, 368–385.

Fang, K.T., Kotz, S., Ng, K.W., 1990. Symmetric multivariate and related distributions. Chapman and Hall, York.

Fang, K.T., Zhang, Y.T., 1990. Generalized Multivariate Analysis. Science Press, Biejing, Springer-Verlag, Berlin.

Genton, M.G., (ed.) 2004. Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality. Chapman and Hall/CRC, Boca Raton, FL.

Gupta, A.K., 2003. Multivariate skew-t distribution. Statistica, 37, 359–363.

Henschel, V., Richter, W.-D., 2002. Geometric generalization of the exponential law. Journal of Multivariate Analysis, 81, 189–204.

Kalke, S., Richter, W.-D., Thauer, F., 2012. Linear combinations, products and ratios of simplicial or spherical variates. Communications in Statistics - Theory and Methods (in press).

Moszyńska, M., Richter, W.-D., 2012. Reverse triangle inequality. Antinorms and semi-antinorms. Studia Scientiarum Mathematicarum Hungarica, 49, 120–138.

Richter, W.-D., 1985. Laplace-Gauss integrals, Gaussian measure asymptotic behaviour and probabilities of moderate deviations. Zeitschrift für Analysis und ihre Anwendungen, 4, 257–267.

Richter, W.-D., 2007. Generalized spherical and simplicial coordinates. Journal of Mathematical Analysis and Applications, 336, 1187–1202.

Richter, W.-D., 2009. Continuous $l_{n,p}$-symmetric distributions. Lithuanian Mathematical Journal, 49, 93–108.

Richter, W.-D., 2011. On the ball number function. Lithuanian Mathematical Journal, 51, 440–449.

Sahu, S.K., Dey, D.K., Branco, M., 2003. A new class of multivariate skew distributions with application to Bayesian regression models. The Canadian Journal of Statistics, 31, 129–150.