Chilean Journal of Statistics
Month 200x

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

Reinaldo B. Arellano-Valle[1] and Wolf-Dieter Richter[2,*]

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

(Received: 03 November 2011 · Accepted in final form: xx Month 2012)

Abstract

The general methods from skewed distributions theory 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: $l_{n,p}$-symmetric distributions. Conditional and marginal distributions. Dirichlet distribution. Scale mixture of the $N_{n,p}$-distribution. p-generalized Student-t-distribution. Skewed distributions. Stochastic representations.

Mathematics Subject Classification: Primary 60E99.

1. Introduction

The univariate skew-normal and its extension to a univariate skew-symmetric distribution were introduced first in Azzalini (1985) and Azzalini (1985), respectively. Many authors extended these considerations under various aspects and in different ways. E.g., a multivariate extension of the skew-normal distribution and its main properties are discussed first in Azzalini and Dalla-Valle (1996) and then in Azzalini and Capitanio (1999). Then, 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-elliptically contoured distributions are considered in Arellano-Valle and Azzalini (2006). Genton (2004) gives an overview of these efforts. The concept of fundamental skew distributions which unifies all at this time known approaches has been developed in Arellano-Valle and Genton (2005). The authors of Arellano-Valle et al. (2006b) bring a certain new structure into the widespread field and unify 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 there at once by generalizing the Fisher distribution, and also for

*Corresponding author. Email: wolf-dieter.richter@uni-rostock.de

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

http://www.soche.cl/chjs

several special cases in Richter (2007) and Kalke et al. (2011).

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. (2006b) applies, and finally we follow the general concept in Arellano-Valle and Azzalini (2006).

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

2. Preliminaries

2.1 The p-generalized normal distribution

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

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

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

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

$$f_{n,p}(x) = \frac{I_{(0,\infty)}(x)}{p^{\frac{n}{p}-1}\Gamma(\frac{n}{p})} x^{n-1} e^{-\frac{x^p}{p}}.$$

As in Richter (2007), we refer this distribution by $R_p \sim \chi(p, n)$. In particular, we have $E(R_p^k) = p^{k/p}\Gamma[(n+k)/p]/\Gamma(n/p)$ for all $k \geq 0$. In addition, $\frac{1}{p}|X_i|^p \overset{iid}{\sim} Ga\,(1/p, 1)\,,\ i = 1, ..., n$, following that $\frac{1}{p}\sum_{i=l}^{l+k-1} |X_i|^p \sim G\,(k/p, 1)$ and $\frac{1}{p}R_p^p = \frac{1}{p}\sum_{i=1}^n |X_i|^p \sim G\,(n/p, 1)$. Moreover, since $|X|_p^p = \sum_{i=1}^n |X_i|^p$ we have straightforwardly that

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

where $D_{m+1}(\alpha_1, ..., \alpha_m, \alpha_{m+1})$, $\alpha_i > 0$, $i = 1, \ldots, m+1$, denotes de Dirichlet distribution. Similarly, the sub-random vector $\left(\frac{|X_1|^p}{|X|_p^p}, ..., \frac{|X_k|^p}{|X|_p^p}, 1 - \sum\limits_{1}^{k} \frac{|X_j|^p}{|X|_p^p}\right)^T$ follows a Dirichlet

$D_{k+1}\left(\frac{1}{p},...,\frac{1}{p},\frac{n-k}{p}\right)$ distribution, $k\in\{1,2,...,n-1\}$, and the sub-vector $(Y_1,\ldots,Y_k)^T=\left(\frac{|X_1|^p}{|X|_p^p},...,\frac{|X_k|^p}{|X|_p^p}\right)^T$ has a density

$$h_k(y_1,...,y_k)=\frac{\Gamma\left(\frac{n}{p}\right)}{\Gamma\left(\frac{1}{p}\right)^k\Gamma\left(\frac{n-k}{p}\right)}\prod_{i=1}^k y_i^{\frac{1}{p}-1}\left(1-\sum_{i=1}^k y_i\right)^{\frac{n-k}{p}-1},\quad y_1>0,...,y_k>0,\ \sum_1^k y_i<1.$$

Hence, the density of $(Z_1,\ldots,Z_k)^T=\left(\frac{|X_1|}{|X|_p^p},...,\frac{|X_k|}{|X|_p^p}\right)^T$ is

$$g_k(z_1,\ldots,z_k)=\frac{\partial^k}{\partial y_1...\partial y_k}P\left(\frac{|X_i|}{|X|_p}\le y_i,i=1,...,k\right)=h_k(z_1^p,...,z_k^p)\prod_{i=1}^k(py_i^{p-1}),$$

and the following lemma has thus been proved.

Lemma 2.1 The density of $(Z_1,\ldots,Z_k)^T=\left(\frac{|X_1|}{|X|_p^p},...,\frac{|X_k|}{|X|_p^p}\right)^T$, where $X=(X_1,\ldots,X_n)^T\sim N_{n,p}$, is

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

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 $X\sim N_{n,p}$ allows the stochastic representation

$$X\overset{d}{=}RU_p$$

where $R\overset{d}{=}R_p$ and is independent of $U_p\overset{d}{=}X/R_p$ 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 $\mathcal{S}_{n,p}=\{x\in\mathbb{R}^n:|x|_p=1\}$). Consider now the partition of X

$$X=(X^{(1)T},X^{(2)T})^T,$$

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

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

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

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

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

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

Since $\frac{1}{p}|X^{(1)}|_p^p \sim G(k/p,1)$ and $\frac{1}{p}|X^{(2)}|_p^p \sim G((n-k)/p,1)$ 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., $R_{k,n}^{(p)} = \frac{|X^{(1)}|_p}{|X|_p}$ or as any random variable following the same distribution as $\frac{|X^{(1)}|_p}{|X|_p}$. This result generalizes Lemma 2 in Cambanis et al. (1981) to the case of arbitrary $p > 0$.

The partition $(X^{(1)T}, X^{(2)T})$ of X^T 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), $R_{k,n}^{(p)}$ is 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 $(X^{(1)\,T}, 0^T)^T$, $R_{k,n}^{(p)} = \cos_p(\phi)$. Note that ϕ takes its values only in the interval $[0, \pi/2]$.

2.3 Moments

Generalizing well known results from Fang et al. (1990) to the case of arbitrary $p > 0$, in this section we compute some multivariate moments of a p-generalized normal vector $X \sim N_{n,p}$. For this, we need first some preliminary notations. We denote the sign of X by $\mathrm{sgn}(X) = (\mathrm{sgn}(X_1), \ldots, \mathrm{sgn}(X_n))^T$ and its absolute value by $|X| = (|X_1|, \ldots, |X_n|)^T$. Here, for any random variable Z which is a.s. different from 0, 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 symmetry that the random vectors $|X|$ and $\mathrm{sgn}(X)$ are independent, and that $\mathrm{sgn}(X)$ has uniform distribution on $\{-1,+1\}^n$. 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 $X \sim N_{n,p}$, then $\mathrm{sgn}(X)$ and $|X|$ are independent random vectors, with $\mathrm{sgn}(X) \sim U(\{-1,+1\}^n)$ and $f_{|X|}(t) = 2^n C_p^n e^{-\frac{1}{p}\sum_{i=1}^n t_i^p}$, $t = (t_1,\ldots,t_n)^T \in \mathbb{R}_+^n$.

For any vector $s = (s_1,\ldots,s_n)^T$, 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 $X \sim N_{n,p}$, 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 $X \sim N_{n,p}$, then for any integers $r_i \geq 0$, $i = 1,\ldots,n$,

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

Proof By Lemma 2.3 and the independence property, $E\left(\prod_{i=1}^n X_i^{r_i}\right) = \prod_{i=1}^n E\left(S_i^{r_i}\right) E\left(T_i^{r_i}\right)$, where $E\left(S_i^{r_i}\right)$ equals 0 for r_i odd and 1 for r_i even, and the proof follows by using that $E\left(T_i^{r_i}\right) = p^{r_i/p}\Gamma[(r_i+1)/p]/\Gamma(1/p)$. ■

Corollary 2.6 If $X \sim N_{n,p}$, then $E(X) = 0$ and $E(XX^T) = \sigma_p^2 I_n$, 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 $U_p = (U_1,\ldots,U_n)^T$ be a p-generalized uniform vector on $\mathcal{S}_{n,p}$. Then, for any integer $r_i \geq 0$, $1,\ldots,n$,

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

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

$$E\left(\prod_{i=1}^n X_i^{r_i}\right) = E\left(R_p^{\sum_{i=1}^n r_i}\right) E\left(\prod_{i=1}^n U_i^{r_i}\right),$$

from where the proof follows by Theorem 2.5 and $E(R_p^s) = p^{s/p}\Gamma[(n+s)/p]/\Gamma(n/p)$ 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, $E(U_p) = 0$ and $E(U_p U_p^T) = \tau_{n,p} I_n$, 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 $\tau_{n,p} = 1/n$ following thus the well-known result that $Var(U_p) = (1/n)I_n$.

3. Continuous $l_{n,p}$-symmetric distributions

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 distribution function (d.f.) of a positive random variable and put

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

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 $g : \mathbb{R}^+ \to \mathbb{R}^+$ 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 $g(r) = e^{-r/p} I_{(0,\infty)}(r)$. In this case, we have $1/I_{n,g,p} = p^{1-n/p}/\Gamma(n/p)$. 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 shall use the notation

$$X \sim S_{n,p}(g)$$

and for its 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 $X \overset{d}{=} RU_p$, where R is a non-negative random variable with density

$$f_R(r) = r^{n-1} g^{(n)}(r^p), \quad r > 0,$$

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

Lemma 3.1 Let $X = (X_1, ..., X_n)^T \sim S_{n,p}(g)$. Then, $X^{(1)} = (X_1, ..., X_k)^T \sim S_{k,p}(g)$ and has density

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

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

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

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

$$\begin{aligned}
P(X_i \leq x_i, i = 1, ..., k) &= P\left(U_i \leq \frac{x_i}{R}, i = 1, ..., k\right) \\
&= \int_0^{\infty} P\left(U_i \leq \frac{x_i}{r}, i = 1, ..., k\right) P(R \in dr) \\
&= \int_0^{\infty} \int_{-1}^{\frac{x_1}{r}} ... \int_{-1}^{\frac{x_k}{r}} \frac{\partial^k}{\partial \tilde{y}_1 ... \partial \tilde{y}_k} P(U_i \leq \tilde{y}_i, i = 1, ..., k) d\tilde{y}_1 ... d\tilde{y}_k P(R \in dr).
\end{aligned}$$

It follows from Lemma 2.1 that

$$P(X_i \leq x_i, i = 1, ..., k) = C \int_0^{\infty} \int_{-1}^{\frac{x_1}{r}} ... \int_{-1}^{\frac{x_k}{r}} I_{\{y_{(k)}: \sum_{i=1}^{k} |y_i|^p \leq 1\}}(\tilde{y}_{(k)}) \left(1 - \sum_{i=1}^{k} |\tilde{y}_i|^p\right)^{\frac{n-k}{p}-1} d\tilde{y}_1 ... d\tilde{y}_k dF(r),$$

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

$$\begin{aligned}
\frac{\partial^k}{\partial x_1 ... \partial x_k} P(X_i \leq x_i, i = 1, ..., k) &= C \int_0^{\infty} I_{\{y_{(k)}: \sum_{i=1}^{k} |y_i|^p \leq 1\}} \left(\frac{x_{(k)}}{r}\right) \left(1 - \sum_{i=1}^{k} |\frac{x_i}{r}|^p\right)^{\frac{n-k}{p}-1} r^{-k} dF(r) \\
&= C \int_{(\sum_{i=1}^{k} |x_i|^p)^{1/p}}^{\infty} r^{-(n-p)} \left(r^p - \sum_{i=1}^{k} |x_i|^p\right)^{\frac{n-k}{p}-1} dF(r) \\
&= g^{(k)} \left(\sum_{i=1}^{k} |x_i|^p\right),
\end{aligned}$$

where

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

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

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

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

Consider again the partition $X = (X^{(1)T}, X^{(2)T})^T$, where as before $X^{(1)}$ and $X^{(2)}$ take values in $\mathbb{R}^k$ $(0 < m < n)$ and $\mathbb{R}^{n-k}$, respectively. We are interested now 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 4.1 that $X^{(2)}$ follows a continuous $l_{n-k,p}$-symmetric distribution with a density generator $g^{(n-k)}$ satisfying the representation

$$g^{(n-k)}(u) = \frac{2^k \Gamma^k\left(\frac{1}{p}\right)}{p^k \Gamma\left(\frac{k}{p}\right)} \int_u^\infty g^{(n)}(y)(y-u)^{\frac{k}{p}-1} dy = \frac{2^k \Gamma^k\left(\frac{1}{p}\right)}{p^k \Gamma\left(\frac{k}{p}\right)} \int_0^\infty g^{(n)}(z+u) z^{\frac{k}{p}-1} dz.$$

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 $X = (X^{(1)T}, X^{(2)T})^T$ 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

$$g_{[a]}^{(k)}(u) = \frac{p^k \Gamma\left(\frac{k}{p}\right) g^{(n)}(a+u)}{2^k \Gamma\left(\frac{1}{p}\right)^k \int_0^\infty g^{(n)}(z+a) z^{\frac{k}{p}-1} dz}.$$

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

This lemma generalizes a corresponding formula in Section 2.4 of Fang et al. (1990). In the special case of the generalized $N_{n,p}$-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 $X = (X^{(1)T}, X^{(2)T})^T$ 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 $R_1 \stackrel{d}{=} R R_{k,n}^{(p)}$, $R_2 \stackrel{d}{=} R(1 - R_{k,n}^{(p)p})^{1/p}$, 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 is independent of $U_p^{(k)}$.
(c) The vectors $X^{(1)}$ and $X^{(2)}$ are conditionally independent given $|X^{(2)}|_p$,

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

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,

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

and

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

where $R_{[a^p]} U_p^{(k)}$ and $a U_p^{(m)}$ 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). The part (c) is a consequence of (b) and generalizes the same result for spherical distributions (see e.g. Arellano-Valle et al. (2006a)).

3.3 Scale mixture of the $N_{n,p}$-distribution

Let $R = V^{-1/p} R_p$, where $R_p \sim \chi(n, p)$ and is independent of V, which is a non-negative mixing variable with d.f. G which 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 $X \sim N_{n,p}$ and is independent of $V \sim G$. We then have $Y \sim S_{n,p}(g^{(n)})$, where the generator function $g^{(n)}$ will be defined below. The density of $V^{-1/p} R_p$ is

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

This density 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 here by $Y \sim t_{n,p}(\nu)$, for which $V \sim Ga(\nu/p, \nu/p)$. In this case, $V^{-1/p}R_p$ has the density

$$f(u) = \frac{u^{n-1}I_{(0,\infty)(u)}}{p^{\frac{n}{p}-1}\Gamma\left(\frac{n}{p}\right)\Gamma\left(\frac{\nu}{p}\right)}\left(\frac{\nu}{p}\right)^{\frac{\nu}{p}}\int\limits_0^\infty v^{\frac{n+\nu}{p}-1}e^{-\frac{v}{p}(u^p+\nu)}dv.$$

The functions f and $g^{(n)}$ satisfy according to Richter (2009) the equation

$$f(r) = \frac{2^n\Gamma\left(\frac{1}{p}\right)^n}{p^{n-1}\Gamma(\frac{n}{p})}r^{n-1}g^{(n)}(r^p)I_{(0,\infty)}(r).$$

Hence, Y follows the $l_{n,p}$-symmetric density

$$f_Y(y) = \frac{|y|_p^{n-1}I_{(0,\infty)}(y)}{p^{\frac{n}{p}-1}\Gamma\left(\frac{\nu}{p}\right)\Gamma\left(\frac{n}{p}\right)}\int_0^\infty v^{[(n+\nu)/p]-1}e^{-\frac{\nu+|y|_p^p}{p}}dv,$$

that is,

$$f_Y(y) = D_{n,p,\nu}\left\{1+\frac{|y|_p^p}{\nu}\right\}^{-\frac{\nu+n}{p}}, \quad D_{n,p,\nu} = \frac{\left(\frac{p}{2}\right)^n\Gamma\left(\frac{\nu+n}{p}\right)}{\Gamma\left(\frac{\nu}{p}\right)\Gamma\left(\frac{1}{p}\right)^n\nu^{\frac{n}{p}}}.$$

Definition 3.4 The distribution of a random vector Y following the density

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

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

This class of p-generalized Student densities 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 of V as

$$V = \frac{|Z_1|^p + ... + |Z_\nu|^p}{\nu} \quad \text{with} \quad (Z_1, ..., Z_\nu)^T \sim N_{n,p} \quad \text{in} \quad \mathbb{R}^\nu.$$

The following theorem has thus been proved.

Theorem 3.5 If $Y = (Y^{(1)T}, Y^{(2)T})^T \sim N_{n+\nu,p}$ where $Y^{(1)}$ and $Y^{(2)}$ take values in $\mathbb{R}^n$ and $\mathbb{R}^\nu$, respectively, then $\frac{\nu^{1/p}}{|Y^{(2)}|_p}Y^{(1)}$ follows the density $t_{n,p}(y;\nu)$, $y\in\mathbb{R}^n$.

This theorem has been proved for $n = 1$ in Richter (2007) and for $p = 2$ in Arellano-Valle and Bolfarine (1995).

If $Y = (Y^{(1)T}, Y^{(2)T})^T \sim t_{n,p}(\nu)$, 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 $X \sim S_{n,p}$, we obtain from the stochastic representation $X \stackrel{d}{=} RU^{(n)}$ that

$$E\left(\prod_{i=1}^{n} X_i^{r_i}\right) = E\left(R^{\sum_{i=1}^{n} r_i}\right) E\left(\prod_{i=1}^{n} U_i^{r_i}\right),$$

provided that $E\left(R^{\sum_{i=1}^{n} r_i}\right)$ is finite, and where $E\left(\prod_{i=1}^{n} U_i^{r_i}\right)$ is given in Corollary 2.7. In particular, by Corollary 2.8 we have $E(X) = 0$ if $E(R)$ is finite and $E(XX^T) = \sigma_{p,g}^2 I_n$, where $\sigma_{p,g}^2 = \tau_p E(R^2)$, if $E(R^2)$ is finite. It is convenient to emphasize here that similarly to the case of $p = 2$, the univariate variance component $\sigma_{p,g}^2 = \tau_p E(R^2)$ 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 $R_p \sim \chi(n,p)$, implying that

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

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

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

In particular, we have $E(X) = 0$ if $\nu > 1$ and $E(XX^T) = \sigma_{p,\nu}^2 I_n$ if $\nu > 2$, where $\sigma_{p,\nu}^2 = \nu\Gamma[(\nu-2)/p]\Gamma(3/p)/\Gamma(\nu/p)\Gamma(1/p)$.

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 $X \sim S_{m,p}(g)$, $\Gamma \in \mathbb{R}^{n\times m}$ 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 function 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 $D_m \cdot LS$ 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 will be called a p-generalized ellipsoid with location vector μ and shape matrix $\Gamma = H_n D_n$.

Since $X \overset{d}{=} RU_p$, we have $Y \overset{d}{=} \mu + R\Gamma U_p$. Y has location vector μ and if $\Gamma = H_n D_n$ we say that the random vector Y has shape matrix Γ. If $E(R^2) < \infty$, then it is straightforward to see that $E(Y) = \mu$ and $Cov(Y) = \sigma^2_{p,g}\Sigma$, where $\Sigma = \Gamma\Gamma^T$ and as was mentioned $\sigma^2_{p,g} = \tau_p E(R^2)$ 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 a density given by

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

Its d.f. $F_Y(y) = P(Y \le y)$ is then

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

where the sign of inequality $\le$ means componentwise inequality. In what follows, we will denote the d.f. of Y by $F_{n,p}(y; \mu, \Sigma, g)$, where $\Sigma = \Gamma\Gamma^T$, 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 elliptically contoured distribution with location vector μ and dispersion matrix Σ and will be commonly denoted by $EC_n(\mu, \Sigma, g)$.

4. Skewed $l_{n,p}$-symmetric distributions

We discuss next 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 in that follows 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

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

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

Proof According to Lemma ,

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

With a matrix $\Lambda : \mathbb{R}^k \to \mathbb{R}^m$, 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

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

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

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

It follows that

$$f_{U_2|U_1=u_1}(u_2) = \frac{g^{(k+m)}(|u_1|_p^p + |\Lambda u_1 - u_2|_p^p)}{g^{(k)}(|u_1|_p^p)} = g^{(m)}_{[|u_1|_p^p]}(|\Lambda u_1 - u_2|_p^p). \tag{1}$$

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

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 hence

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

By the general representation formula for the density of the corresponding conditional distribution in Arellano-Valle and del Pino (2004),

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

where $f_{U_1}(z) = f_{X^{(1)}}(z) = g^{(k)}(|z|_p^p)$.

By (1) and the change of variable $w = \Lambda z - u_2$, we have

$$P(0 < U_2|U_1 = z) = \int\limits_{\mathbb{R}^m_+} g^{(m)}_{[|z|_p^p]}(|\Lambda z - u_2|_p^p) du_2 = F^{(1)}_{m,p}\left(\Lambda z; g^{(m)}_{[|z|_p^p]}\right). \tag{2}$$

Hence,

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

with $1/C_{m,p} = P(0 < U_2)$.

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. Hence, $P(0 < U_2) = P(-U_2 < 0) = P(U_2 < 0)$. The d.f. of $U_2 = \Gamma X$ will be denoted by $F_{m,p}^{(2)}\left(u_2; \Sigma, g^{(m)}\right)$, where $\Sigma = \Gamma\Gamma^T = I_m + \Lambda\Lambda^T$, thus $1/C_{m,p} = F_{m,p}^{(2)}\left(0; I_m + \Lambda\Lambda^T, g^{(m)}\right)$. ■

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

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

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

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

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

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 $f_Z(z) = 2g^{(k)}(|z|_p^p)\, F_{1,p}^{(1)}\left(\lambda^T z; g_{[|z|_p^p]}^{(1)}\right)$, $z \in \mathbb{R}^k$, for which $F_{1,p}^{(1)}\left(u; g_{[|z|_p^p]}^{(1)}\right)$ is a univariate d.f. and is immediate to be computed numerically when it has not an explicit expression. For $m \geq 1$, the above definition extends the analoguous 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.

This corollary 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 $N_{n,p}$ distribution, where $g^{(k)}(|x|_p^p) = C_p^k e^{-\frac{1}{p}|x|_p^p} =: \phi_{k,p}(x)$ is $N_{k,p}$ density function and $F_{k,p}^{(1)}\left(x; \Sigma, g^{(k)}\right) = \int_{t<x} \phi_{k,p}(t;\Sigma)dt =: \Phi_{k,p}^{(1)}(x;\Sigma)$, $x \in \mathbb{R}^k$, i.e. the d.f. of a non-singular linear transformation $Y = \Gamma X$, with $X \sim N_{n,p}$ and $\Gamma\Gamma^T = \Sigma$. Denoting accordingly $F_{m,p}^{(2)}(0; I_m + \Lambda\Lambda^T, g^{(m)}) = \Phi_{m,p}^{(2)}(0, I_m + \Lambda\Lambda^T)$, we shall say that random vector Z has k-dimensional skew-$N_{n,p}$ distribution with dimensionality parameter m, density generator g and skewness/shape matrix parameter $\Lambda \in \mathbb{R}^{m\times k}$, denoted by $Z \sim SN_{k,m,p}(\Lambda)$, if its density is given by

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

For $m = 1$ and $p = 2$ we obtain the multivariate skew-normal density $f_Z(z) = 2\phi_{k,p}(z)\,\Phi_{1,p}(\lambda^T z)$, $z \in \mathbb{R}^k$, 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-$N_{n,p}$ random vector $Z = (Z_1, \ldots, Z_k)^T$ 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 $g^{(k)}(|x|_p^p) = D_{k,p,\nu}\left\{1 + \frac{|x|_p^p}{\nu}\right\}^{-(\nu+k)/p} =: t_{k,p}(x;\nu)$ is the $t_{k,p}(\nu)$ density, and $F^{(1)}_{k,p}(x;\Sigma,g) = \int_{t<x} t_{k,p}(t;\Sigma,\nu)dt := T^{(1)}_{k,p}(x;\Sigma,\nu)$ and $T^{(2)}_{m,p}$ is defined accordingly. We shall say that a random vector Z has skew-$t_{k,p}$ distribution with dimensionality parameter m and skewness/shape matrix parameter $\Lambda \in \mathbb{R}^{m\times k}$, denoted by $Z \sim St_{k,m,p}(\Lambda)$, if its density is given by

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

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 as case, we have the more general skew p-generalized $l_{k,p}$-symmetric class of densities defined by

$$f_Z(z) = \frac{1}{F^{(2)}_{m,p}(\tau; I_m + \Lambda\Lambda^T g)}\, g^{(k)}(|z|_p^p)\, F^{(1)}_{m,p}\left(\Lambda z + \tau; g^{(m)}_{[|z|_p^p]}\right), \quad z \in \mathbb{R}^k. \tag{3}$$

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 as SUE-p-generalized family of distributions, and most of the above results could be be explored for this class.

4.2 Construction from stochastic representations

Consider now 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 $S_{k+m,p}(g)$-distribution, and where $\Delta \in \mathbb{R}^{k\times m}$ is fixed matrix. Consider also 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

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

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) and Arellano-Valle 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

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

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

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

where $Q(z) = z^T [I_k - \Delta(I_m + \Delta^T \Delta)^{-1} \Delta^T] z = z^T (I_k + \Delta\Delta^T)^{-1} z$. 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 turns out from the stochastic representation (4), which among other things allows to compute easily the moments of Z (see Arellano-Valle et al. (2002) and Arellano-Valle del Pino (2004)). In particular when the mean vector and covariance matrix of Z exist, we have from (4) that they are given by

$$E(Z) = \Delta E(|X^{(2)}|) \quad \text{and} \quad Cov(Z) = Cov(X_1^{(2)}) + \Delta Cov(|X^{(2)}|) \Delta^T,$$

where $Cov(X_1^{(2)}) = \sigma_{p,g}^2 I_k$. To compute $E(|X^{(2)}|)$ and $Cov(|X^{(2)}|)$ we can use the following lemma, whose proof is straightforward from the results in Section 2.3.

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

$$E(|X_i|^r |X_j|^s) = \begin{cases} \frac{\Gamma(\frac{r+s+1}{p})}{\Gamma(\frac{1}{p})} \frac{\Gamma(\frac{n}{p}) E(R^{r+s})}{\Gamma(\frac{n+r+s}{p})}, & i = j, \\ \frac{\Gamma(\frac{r+1}{p}) \Gamma(\frac{s+1}{p})}{\Gamma^2(\frac{1}{p})} \frac{\Gamma(\frac{n}{p}) E(R^{r+s})}{\Gamma(\frac{n+r+s}{p})}, & i \neq j, \end{cases}$$

if $E(R^{r+s})$ is finite.

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

$$\frac{\Gamma(\frac{n}{p}) E(R_p^k)}{\Gamma(\frac{n+k}{p})} = p^{\frac{k}{p}}.$$

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

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

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 $Z \overset{d}{=} V^{-1/p}Z_p$. Hence, we have $E(Z) = E(V^{-1/p})E(Z_p)$ if $E(V^{-1/p})$ is finite and $E(Z_pZ_p^T) = E(V^{-2/p})E(Z_pZ_p^T)$ if $E(V^{-2/p})$ is finite, from where we can compute $Cov(Z)$.

Acknowledgement

The second author would like to express his gratitude to the Pontificia Universidad Cátolica de Chile for supporting his stay there in March 2010 and to Reinaldo B. Arellano-Valle for his kind hospitality. The research of the first author was supported by grant FONDECYT 1085241-Chile. The authors express their thanks to the referee for valuable hints.

References

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., del Pino, G., San Martín, E., 2002. Definition and probabilistic properties of skew sistributions, Statistics and Probability Letters, 58, 111-121.

Arellano-Valle, R. B., del Pino, G., 2004. From symmetric to asymmetric distributions: A unified approach. Book chapter 7, in Skew-Distributions and Their Applications: A Journey Beyond Normality, edited volume by M.G. Genton, CRC/Chapman-Hall.

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

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., Branco, M. D., Genton, M. G., 2006b. A unified view on skewed distributions arising from selections. The Canadian Journal of Statistics, 34, No.4, 1-21.

Arellano-Valle, R. B., Genton, M. G., 2005. On fundamental skew distributions. Journal of Multivariate Analysis, 96,9 3-116.

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 Royal Statistical Society, Series B **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, 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, London, New York.

Fang, K.T., Zhang, Y.T., 1990. Generalized multivariate analysis. Sicience Press, Biejing, Springer-Verlag, Berlin.

Genton, M.G.,ed., 2004. Skew elliptical distributions and their applications: A journey beyond normality. Chapman Hall CRC, Boca Raton, Florida.

Gupta, A. K., 2003. Multivariate skew-t distribution. Statistics, 37(4), 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 print.

Moszyńska, M., Richter, W.-D., 2012. Reverse triangle inequality. Antinorms and semi-antinorms. Studia Scientarum Mathematicarum Hungarica, in print.

Richter, W.-D., 1985. Laplace-Gauss integrals, Gaussian measure asymptotic behaviour and probabilities of moderate deviations. Z. Anal. Anw., 4(3):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 Mathematics Journal, 49, 1, 93-108.

Richter, W.-D., 2011. On the ball number function. Lithuanian Mathematics Journal, 51, 3, 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.