Convex and radially concave contoured distributions

Wolf-Dieter Richter,

University of Rostock, Germany
E-mail:wolf-dieter.richter@uni-rostock.de
http://www.mathematik.uni-rostock.de/

October 23, 2015

Abstract

Integral representations of the locally defined star-generalized surface content measures on star spheres are derived for boundary spheres of balls being convex or radially concave w.r.t. a fan in $\mathbb{R}^n$. As a result, the general geometric measure representation of star-shaped probability distributions and the general stochastic representation of the corresponding random vectors allow additional specific interpretations in the two mentioned cases. Applications to estimating and testing hypotheses on scaling parameters are presented, and two-dimensional sample clouds are simulated.

Key words: Symmetric distribution, norm contoured distribution, antinorm contoured distribution, dual norm geometry, semi-antinorm geometry, surface content measure, non-Euclidean surface measure, geometric measure representation, light tails, heavy tails, stochastic representation, asymmetric distribution, scaling model, sample cloud, simulation.

2010 Mathematics Subject Classification 60E05, 60D05, 28A50, 28A75, 26B15, 51F99, 62E15, 62F25, 62F03, 62F35

1 Introduction

The families of multivariate Gaussian and elliptically contoured distributions have served for a long time as the main basis of numerous probabilistic models and their many successful applications. Basics of estimation theory of elliptically contoured distributions can be found in [5],[1],[3] and, e.g., [9]. Advancing needs of statistical practice as well as longstanding challenging mathematical questions stimulated the development of larger classes of probability laws containing many well known distributions as particular elements. We note that [2] surveys a big part of the distribution theory in $\mathbb{R}^2$. Numerous authors contributed to establishing the class of multivariate star-shaped distributions. For a recent review of this development, see [14]. Estimating level sets of star-shaped densities has been dealt with in [15],[16], [17] and [4].

Several aspects of analyzing a cloud of sample points may be of importance for the process of defining a class of probability laws. The visual impression of the appearance of star-shaped figures built by the points of a sample cloud may lead to the idea that the boundaries of star-bodies, henceforth called star-spheres, represent density level sets of a probability law.

1

Counting the sample points belonging to thin layers about star-spheres then leads to the idea that a certain function that assigns a nonnegative number to every such star-sphere serves as the density generating function (dgf) of a non-negative random variable (rv), or more generally, as a function being proportional to the Radon-Nikodym density of a multivariate probability law w.r.t. a certain σ-finite measure defined on the sample space. We call such a function an univariate level density function of the multivariate probability distribution.

The combination of the aspects of defining level sets of a multivariate density and of assigning a non-negative level to every such set will be reflected here in a new method of integration. This method may be considered as a hightening and generalization of the classical principle of Cavalieri which was modified by Torricelli. Combining integration on level sets with that along the levels may also be considered as a geometric disintegration method. This method is essentially based upon certain non-Euclidean surface measures on star-spheres. It is one of the main aims of this paper to further develop this theory for two important types of star-spheres. Convex bodies and star-bodies being radially concave w.r.t. a fan in $\mathbb{R}^n$ build these two classes of star-bodies. Thus, in this paper, the focus is on considering probability laws having the boundary of such sets as their density level sets or their contour sets. Actually, the results in Section 3 are mainly restricted to, possibly shifted, symmetric contour sets being norm or antinorm spheres.

There are different ways to introduce a dgf of a continuous probability law. Looking through the statistical and mathematical literature, one finds many interesting non-negative and suitably integrable functions which may serve as a dgf. Another way to introduce a dgf is to analyze the structure of a known multivariate density and to extract from it, if possible, a function which does not depend on the surface measure on the star-spheres but depends exclusively on the levels of the multivariate density.

It is well known that the definition of a dgf is not unique, and how to deal with this circumstance. Densities with heavy tails may be of interest in (re-)insurance, and densities with light tails may be of interest in reliability theory. Both types of densities can be modeled using each time a suitable dgf.

Modeling the density level sets and the univariate level density of a multivariate distribution can be done in a combined way or separate from each other. Sometimes a parameter may influence both the level density and the density contour sets of a multivariate distribution. Another parameter may be only for one of these two aspects of importance.

The paper is organized as follows. Some basic facts from the theory of star-shaped distributions, with an emphasis on geometric measure representations, are collected in Section 2. New geometric descriptions of surface measures on boundary spheres of balls being radially concave w.r.t. a fan in $\mathbb{R}^n$, or convex, are presented in Section 3. Moreover, new statistical applications of geometric measure representations of norm and antinorm contoured distributions and of stochastic representations of correspondingly distributed random vectors are discussed there. In particular, distributions are illustrated by simulated sample clouds. The basics for estimating and testing hypothesis on scaling parameters are presented at the end of Section 3. Section 4 deals with proving the new results, and a discussion of the results can be found in the final Section 5.

2 Star-shaped distributions

Geometric measure representations and stochastic representations of corresponding random vectors have been proved in [14] for general star-shaped distributions making essentially use of the notion of a star-generalized surface content measure. The latter is defined in a local way by taking derivatives of sector volumes and is known to be equivalently defined in an integral (in dimension two even explicitly differential geometric) way in the cases of $l_{n,p}$-spheres and ellipsoids. For recent results and a survey of their probabilistic and statistical applications we refer to [14]. Here, some basic facts from star-shaped distribution theory and its applications will be summarized.

Let a random vector Y follow the probability density function (pdf)

$$\varphi_{g,K,\nu}(x) = C(g,K)g(d_K(x - \nu)), x \in \mathbb{R}^n$$

where $\nu \in \mathbb{R}^n$ is a vector of location, $K \subset \mathbb{R}^n$ is a star-body having the origin as an inner point, d_K is the distance function, or Minkowski functional, of the star-body K,

$$d_K(x) = \inf\{\lambda \geq 0 : x \in \lambda K\}, \, x \in \mathbb{R}^n,$$

the function $g : [0, \infty) \to [0, \infty)$ satisfies $0 < I(g) < \infty$ where $I(g) = \int\limits_0^\infty r^{n-1}g(r)dr$, and the normalizing constant allows the representation

$$C(g,K) = \frac{1}{\mathfrak{O}_S(S)I(g)}.$$

Assuming that the technical Assumption 1 in [14] is satisfied which deals with a certain smoothness property of the boundary S of K, $\mathfrak{O}_S$ denotes the star-generalized surface content measure defined on the Borel subsets of S. The probability measure corresponding to $\varphi_{g,K,\nu}$ allows the geometric measure representation or disintegration formula

$$\Phi_{g,K,\nu}(B) = C(g,K) \int\limits_0^\infty r^{n-1}g(r)\mathfrak{O}_S([\frac{1}{r}(B - \nu)] \cap S)dr, B \in \mathfrak{B}^n. \tag{1}$$

K is called the density contour defining star body of this distribution, and any g under consideration a density generating function (dgf). The sets $(B - \nu) \cap S(r)$ with $S(r) = rS$ may be considered playing the role of the indivisibles within a generalized principle of Cavalieri (which was modified by Torricelli). The random vector Y satisfies the stochastic representation

$$Y \stackrel{d}{=} R \cdot U_S \tag{2}$$

where R and U_S are stochastically independent, R has the pdf

$$f_R(r) = I_{(0,\infty)}(r)\frac{1}{I(g)}r^{n-1}g(r),$$

and U_S follows the star-generalized uniform probability distribution ω_S on the Borel-σ-field $\mathcal{B}_S = \mathcal{B}^n \cap S$ over S, $U_S \sim \omega_S$,

$$P(U_S \in A) = \frac{\mathfrak{O}_S(A)}{\mathfrak{O}_S(S)}, A \in \mathcal{B}_S. \tag{3}$$

3

Because of (3), U_S is called the star-generalized uniform basis of Y. The symbol $X \overset{d}{=} Y$ means that the random vectors X and Y follow the same probability law while $X \sim Q$ indicates that the random vector X follows the probability distribution Q.

For $A \in \mathcal{B}_S$, we introduce the central projection cone

$$CPC(A) = \{x \in \mathbb{R}^n : \frac{x}{h_K(x)} \in A\}$$

and the star sector of star radius r,

$$sector(A, r) = CPC(A) \cap K(r)$$

where

$$K(r) = r \cdot K = \{(rx_1, ..., rx_n)^T : (x_1, ..., x_n)^T \in K\}$$

is a star ball of star radius r. Let μ be the Lebesgue measure in $\mathbb{R}^n$. Then the star-generalized surface measure is defined on $r\mathcal{B}_S$ in a local approach by

$$\mathfrak{O}_S(A) = f'(r) \text{ where } f(r) = \mu(sector(A, r)). \tag{4}$$

Making use of the star-sphere intersection proportion function (ipf) of a set A,

$$\mathfrak{F}_S(A, r) = \omega_S([\frac{1}{r}A] \cap S), r > 0,$$

the disintegration representation of $\Phi_{g,K,\nu}$ may be written

$$\Phi_{g,K,\nu}(B) = \frac{1}{I(g)} \int_0^\infty r^{n-1}g(r)\mathfrak{F}_S(B - \nu, r)dr, \ B \in \mathcal{B}^n. \tag{5}$$

The most immediate applications of this formula appear in cases where the ipf is a constant or an indicator of an interval. If, for a certain set B, the ipf takes a constant value, C say, then $\Phi_{g,K,\nu}(B)$ is just equal to this value C.

If, for a statistic T, $B(t) = \{T < t\}, t \in \mathbb{R}$, and the ipfs of all sets $B(t)$ take the constant value, $C(t)$ say, then the statistic T is robust w.r.t. the dgf g, i.e. the distribution of T does not depend on g.

If, for a certain set $A(x)$, the ipf is the indicator function of an interval, $[0, b(x)]$ say, then

$$\frac{d}{dx}\Phi_{g,K,\nu}(A(x)) = \frac{b'(x)b(x)^{n-1}g(b(x))}{I(g)}. \tag{6}$$

For specific statistical examples of such type we refer to Section 3. Applications of (5) in cases where the ipf is more structured are often more involved. Such a situation will be considered in Example 3.4.

The main aim of this paper, however, is not only to give attractive examples where the geometric measure representation applies but is also to give non-trivial explanations of the locally defined surface measure $\mathfrak{O}_S$ on the basis of an integral (or differential-geometric) approach. This will be done in the first two parts of Section 3 for the two important cases where K is a norm or antinorm ball. As a result, in formulas (1)-(3) and (5), $\mathfrak{O}_S$ will

4

afterwards allow additional specific integral (or differential geometric) interpretations in the two mentioned cases. The class

$$CStSh = \{\Phi_{g,K,\nu} : \nu \in \mathbb{R}^n, K \text{ is a star body with } 0 \in \text{int} K, g \text{ is a dgf}\}$$

where $\text{int} K$ means the interior of K is called the class of continuous star-shaped distributions. A random vector Y is said in [14] to belong to the bigger class of star-shaped distributions $StSh^{(n)}$ if there are a vector $\nu \in \mathbb{R}^n$, a star body K with $0 \in \text{int} K$ (and boundary S), a non-negative random variable (rv) R with cumulative distribution function (cdf) F such that $Y - \nu \stackrel{d}{=} R \cdot U_S$ where $U_S \sim \omega_S$, and R and U_S are independent. In this case, we write

$$Y \sim \Psi_{F,K,\nu} \tag{7}$$

and

$$StSh^{(n)} = \{\Psi_{F,K,\nu} : F \text{ is the cdf of a non-negative rv}, K \text{ is a star body}, 0 \in \text{int} K, \nu \in \mathbb{R}^n\}.$$

The random vector U_S is called the star-generalized uniform basis of the class $StSh^{(n)}$. If K is symmetric w.r.t. the origin, $K = -K$, then the functional d_K is a norm, $d_K = ||.||_K$, if K is convex, and is an antinorm, $d_K = \dagger . \dagger_K$, if K is radially concave w.r.t. a fan in $\mathbb{R}^n$. For the latter notions, see [11]. Note that $l_{n,p}$-symmetric distributions are norm or antinorm contoured if $p \geq 1$ or $0 < p \leq 1$, respectively. We shall study general convex or norm contoured distributions in Section 3.1 and distributions being radially concave w.r.t. a fan in $\mathbb{R}^n$, or antinorm contoured, in Section 3.2. The main aim of these two sections is to give closer descriptions of $\mathfrak{O}_S$ being basic for both the general stochastic representation of Y in (2) and the specific geometric measure representations of ω_S in (3) and $\Phi_{g,K,\nu}$ in (1) and (5) in case Y has a density. Moreover, two-dimensional distributions are illustrated by graphics showing simulated sample clouds. Applications to estimating and testing hypotheses on scaling parameters are demonstrated in Section 3.3. The proofs of the results from Sections 3.1 and 3.2 will be presented in Section 4, and a final discussion of the results follows in Section 5.

3   Results

We start the presentation of new results with a remark on asymmetric distribution laws which seems to be very useful: a distribution beeing star-shaped w.r.t. a fan $\mathfrak{F}$ may be restricted to arbitrary unions of elements of $\mathfrak{F}$.

Remark 3.1. *Let* $C_i \in \mathfrak{F}, i \in I$ *and* $C = \bigcup_{i \in I} C_i$. *Then*

$$\Phi_{g,K,\nu}^C(B) = \frac{\mathfrak{O}_S(S)}{\mathfrak{O}_S(S \cap C)} \frac{1}{I(g)} \int_0^\infty r^{n-1} g(r) \mathfrak{F}_S(B - \nu, r) dr$$

is a probability law on $C \cap \mathfrak{B}^n$.

Here, I is a suitable index set. The proof of this result follows immediately by conditioning.

We call the collection of all such distributions the class of fan restricted star laws and denote it by StL,

$$StL = \{\Phi_{g,K,\nu}^{C},\ C = \bigcup_{i\in I} C_i,\ C_i \in \mathfrak{F},\ \forall i\}.$$

Elements of this distribution class are not symmetric, in general.

3.1 Norm contoured distributions

Let K be convex and symmetric w.r.t. the origin throughout this section. Our consideration is restricted therefore here to the class of norm contoured distributions,

$$\mathcal{NC} = \{\Psi_{F,K,\nu} \in StSh^{(n)} : K \text{ is the unit ball of a norm in } \mathbb{R}^n\}.$$

Let the system of Borel sets from the upper half of the sphere S be $\mathcal{B}_S^+$. For $A \in \mathcal{B}_S^+$, put

$$G(A) = \{\vartheta \in \mathbb{R}^{n-1} : \exists \eta = \eta(\vartheta) \text{ with } (\vartheta^T, \eta)^T \in A\},$$

and denote, where ever it exists, the outer normal vector to the norm sphere S at the point $(\vartheta^T, \eta)^T \in S$ by $N(\vartheta)$. Where ever the outer normal vector is not defined, let $N(\vartheta)$ denote the zero element of $\mathbb{R}^n$. Note that the set of boundary points of K where $\nabla \eta$ does not exist is countable and thus without of any influence onto the value of the integral in the following theorem. We recall that the surface content measure $\mathfrak{O}_S$ was locally defined in (4).

Theorem 3.1. *In formulas (1)-(3) and (5), the surface content measure $\mathfrak{O}_S$ allows the representation*

$$\mathfrak{O}_S(A) = \int_{G(A)} d_{K^*}(N(\vartheta))d\vartheta,\ A \in \mathcal{B}_S^+$$

where K^* is the unit ball of the norm $||.||^*$ being dual to $d_K = ||.||_K$.

We shall refer to this result as to the integral or differential geometric approach to measuring surface content on a norm sphere based upon the dual norm geometry. We mention that a similar representation of $\mathfrak{O}_S(A)$ follows for arbitrary $A \in \mathcal{B}_S$. Due to Theorem 3.1, if K is convex, the surface measure $\mathfrak{O}_S$ henceforth allows both the local and the integral interpretation in formulas (1)-(3) and (5). Moreover, Theorem 3.1 reflects a certain specific aspect of duality theory for norms.

In the next section we will deal in an analogous way with balls being radially concave w.r.t. a fan in $\mathbb{R}^n$.

Figures 1-3 show sample clouds of size $k = 2000$ of p-generalized Gaussian distributed two-dimensional ($n = 2$) random vectors for different choices of p, $p \geq 1$. Notice that the six frames reflect different scaling of the clouds due to different values of p. While the sample cloud in Figure 1(a) might seem to be similar to the illustration of the Gaussian case the shape of the sample cloud approaches that of an axes-aligned square if p increases (or even tends to infinity). At the same time, the cloud (probability mass) becomes more and more concentrated. If, however, $p \geq 1$ is tending to one then the shape of the sample cloud approaches that of the diamond. At the same time, probability mass becomes much less concentrated and the contour of the sample cloud appears to be not as sharp as in the opposite
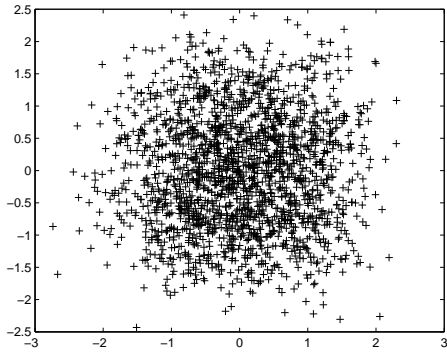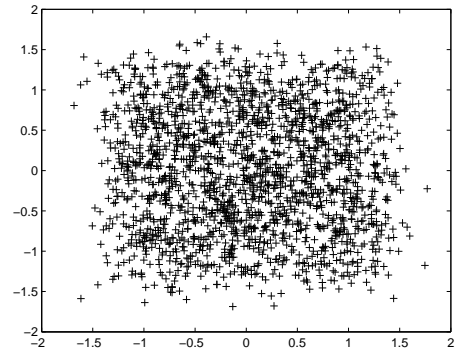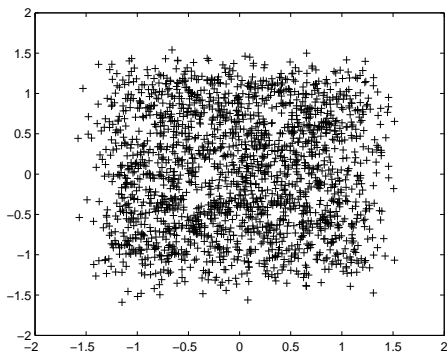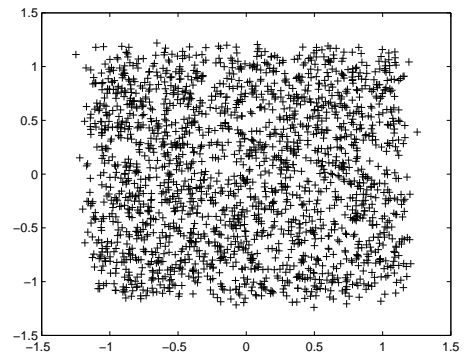
(a) $p = 3$ (b) $p = 6$

Figure 1: Sample clouds of convex contoured p-generalized normal distributions (n=2)
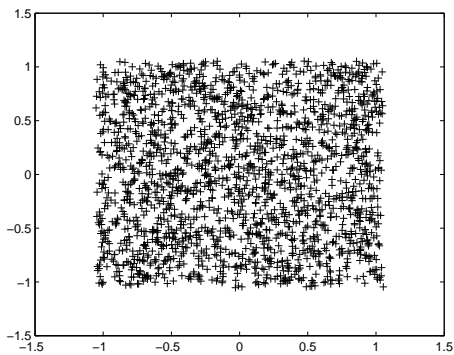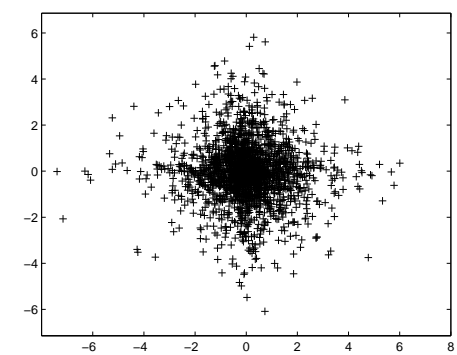

(a) $p = 8$ (b) $p = 20$

Figure 2: Convex contoured cases far from the normal one

(a) $p = 100$ (b) $p = 1$

Figure 3: Two extremal convex cases

7

case. Hence, the parameter p of such a distribution might be called a shape-concentration parameter. Note that Figures 6-8 also present sample clouds of convex contoured distributions but where emphasis is, inter alia, on the effect forced by an increasing sample size k.

Finally, let us remark that Figure 3(b) with equal rights could be presented in the next section because $l_{n,p}$-balls are both convex and radially concave w.r.t. the standard fan in $\mathbb{R}^n$ if $p = 1$.

3.2 Antinorm contoured distributions

Throughout the present section, let K denote a star body having a positive and continuous radial function and being symmetric w.r.t. the origin and radially concave w.r.t. a fan $\mathfrak{F} = \{C_1, C_2, ...\}$ in $\mathbb{R}^n$. The Minkowski functional or distance function of K is then an antinorm, $d_K = \dagger . \dagger$, i.e. a continuous, positively homogeneous, non-degenerate, and in $\mathfrak{F}$ super additive function. For more details we refer to [11]. According to all the assumptions made so far, our consideration is restricted here to the class of antinorm contoured distributions,

$$\mathcal{ANC} = \{\Psi_{F,K,\nu} \in StSh^{(n)} : K \text{ is the unit ball of an antinorm in } \mathbb{R}^n\}.$$

Moreover, we assume that K belongs to a particular class of antinorm balls, $AN1$, meaning that

1) for every i there is a 1-1-map $x_{S,i} : S^{(n-1)} \cap C_i \to S \cap C_i$
   where $S^{(n-1)}$ denotes the Euclidean unit sphere in $\mathbb{R}^n$

and

2) for every i, for all $u \in S^{(n-1)} \cap C_i$ there is a hyperplane $T(u)$ satisfying $T(u) \perp u$
   and being an inner tangent plane to (the boundary of) K at $x_{S,i}(u)$.

We define the anti-support function of K w.r.t. $\mathfrak{F}$ by

$$h_K^{\mathfrak{F}}(u) = \sum_i I_{C_i}(u) \inf\{u^T y : y \in S \cap C_i\}, \ u \in \mathbb{R}^n,$$

and the anti-polar set of K,

$$K^o = \{\lambda(u)u : 0 \le \lambda(u)h_K^{\mathfrak{F}}(u) \le 1, u \in S_E^{(n-1)}\}$$

where $0 \le \lambda(u) \le \infty$ if $h_K^{\mathfrak{F}}(u) = 0$. Let $N(\vartheta)$ be the inner normal vector to S at $(\vartheta^T, \eta)^T \in S$. In the sense of Remark 1 in [14], in what follows we will simply write

$$\int_{G(A)} d_{K^o}(N(\vartheta))d\vartheta \quad \text{instead of} \quad \sum_i \int_{G(A \cap C_i)} d_{K^o}(N(\vartheta))d\vartheta.$$

Theorem 3.2. *In formulas (1)-(3) and (5), the surface content measure $\mathfrak{O}_S$ allows the representation*

$$\mathfrak{O}_S(A) = \int_{G(A)} d_{K^o}(N(\vartheta))d\vartheta, \ A \in \mathcal{B}_S.$$

Additionally to the general local definition in formula (4), the result of this theorem allows the integral or differential geometric interpretation of the surface measure $\mathfrak{O}_S$ in the geometry

having K^o as its unit ball. Note that the special case where surface content of subsets of the boundary of the antinorm ball $K = B_{a,p} = \{x \in \mathbb{R}^n : (\sum_{i=1}^n |\frac{x_i}{a_i}|^p)^{1/p} \leq 1\}, 0 < p \leq 1, a_i > 0, \forall i$ is measured based upon the corresponding semi-antinorm geometry has been dealt with already in [14].

Figures 5 and 6 show sample clouds of the same sample size and from the same distribution class as in Section 3.1 but with parameter p chosen from the interval $(0, 1)$. A big proportion of probability mass can be observed tending to the far tails of such distribution if p is approaching zero. At the same time one can identify several points which could be considered being outliers if they had appeared under other circumstances. Hence, the parameter p of such distribution might be called a shape-tail parameter.
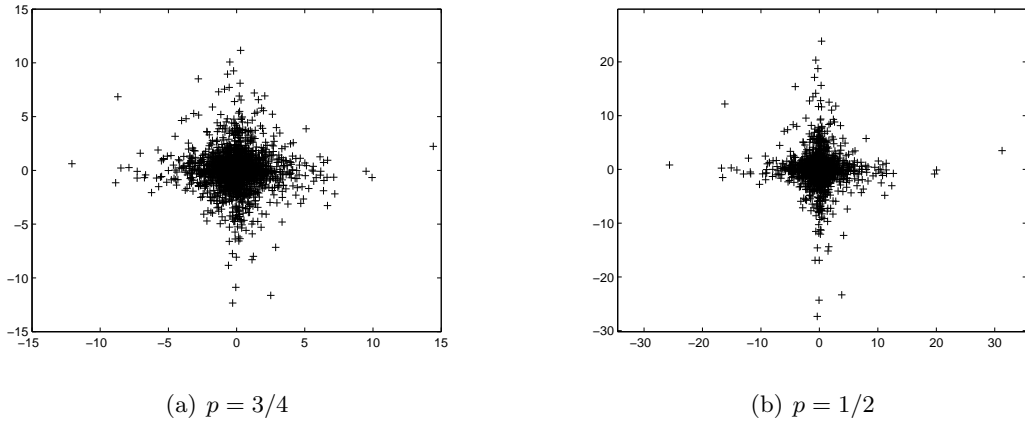

(a) $p = 3/4$ (b) $p = 1/2$

Figure 4: Sample clouds of radially concave contoured p-generalized normal distributions

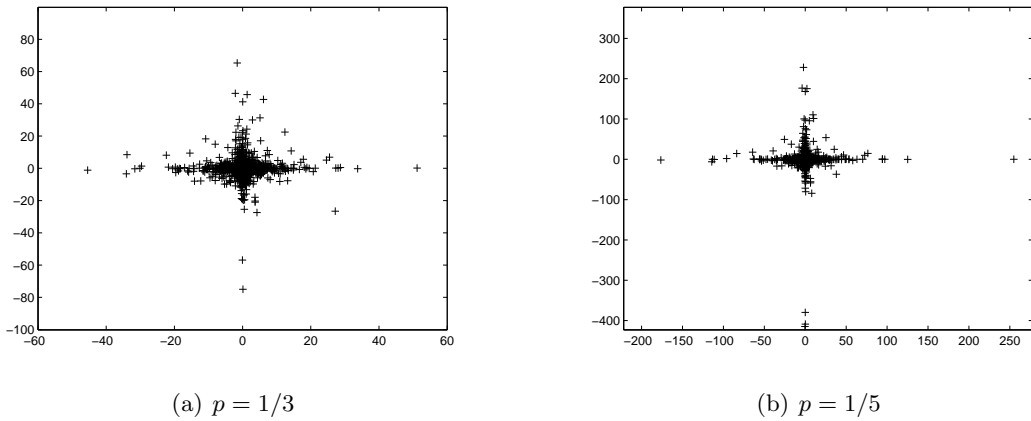
(a) $p = 1/3$ (b) $p = 1/5$

Figure 5: Far reaching tails

9

3.3 Statistical applications

This section deals with several examples where formula (5) applies. The first three examples present relatively immediate applications while the last example concerns a more advanced situation for calculating the ipf.

Example 3.1. *Let $X_1, ..., X_n$ be independent rv's following the common density of the power-exponential (or p-generalized Gaussian, or p-generalized Laplace) distribution,*

$$f_p(x; \mu, \sigma^2) = \frac{C_p}{\sigma} \exp\{-\frac{|x-\mu|^p}{p\sigma^p}\}, x \in \mathbb{R}, \quad C_p = \frac{p^{1-1/p}}{2\Gamma(1/p)}$$

where the location parameter $\mu \in \mathbb{R}$ and the shape-concentration or shape-tail parameter $p > 0$ are known and the scaling parameter σ is unknown. The maximum-likelihood estimator (mle) of σ is

$$\widehat{\sigma} = (\frac{1}{n} \sum_{i=1}^{n} |X_i - \mu|^p)^{1/p}$$

and the random vector $X_{(n)} = (X_1, ..., X_n)^T$ follows the density $\varphi_{g_{PE}, K, \nu}$ where $\nu = (\mu, ..., \mu)^T$, $K = \sigma B_p$, $B_p = \{x \in \mathbb{R}^n : |x|_p \leq 1\}$, $|x|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}$, and

$$g_{PE}(r) = \exp\{-\frac{r^p}{p}\}, r \geq 0.$$

The set B_p is convex if $p \geq 1$, and radially concave w.r.t. the standard fan in $\mathbb{R}^n$ if $0 < p \leq 1$. Hence,

$$\frac{1}{\sigma}(X_{(n)} - \nu) \sim \Phi_{g_{PE}, B_p, 0_n}$$

follows a centered convex or radially concave contoured distribution if $p \geq 1$ or $0 < p \leq 1$, respectively. It turns out that

$$P(n(\frac{\widehat{\sigma}}{\sigma})^p < t) = \Phi_{g_{PE}, B_p, 0_n}(t^{1/p} B_p),$$

thus

$$P(n(\frac{\widehat{\sigma}}{\sigma})^p < t) = \frac{1}{I(g_{PE})} \int_0^\infty r^{n-1} g_{PE}(r) \mathfrak{F}_{S_p}(t^{1/p} B_p, r) dr$$

where $I(g_{PE}) = p^{n/p-1} \Gamma(n/p), \Gamma$ denotes the Gamma function, $S_p = \partial B_p$ is the p-sphere, i.e. the set of boundary points of B_p and the p-sphere ipf of $t^{1/p} B_p$ is

$$\mathfrak{F}_{S_p}(t^{1/p} B_p, r) = I_{[0, t^{1/p}]}(r), r \geq 0.$$

Consequently,

$$P(n(\frac{\widehat{\sigma}}{\sigma})^p < t) = \frac{1}{I(g_{PE})} \int_0^{t^{1/p}} r^{n-1} g_{PE}(r) dr,$$

i.e. $n(\frac{\widehat{\sigma}}{\sigma})^p$ *follows the* $\chi^p_{g_{PE}}$*-density with n d.f. introduced in [12],*

$$t \mapsto \frac{t^{n/p-1}e^{-t/p}}{p^{n/p}\Gamma(n/p)} = f^\chi_{n,g_{PE},p}(t), t > 0.$$

This exact distributional result allows constructing confidence intervals for and testing hypotheses on the scaling parameter σ.

Example 3.2. *We consider independent rv's* $X_1, ..., X_n$ *following the densities*

$$f_p(x; \mu_i, \sigma^2 a_i^2) = \frac{C_p}{\sigma a_i}\exp\{-\frac{|x-\mu_i|^p}{p(\sigma a_i)^p}\}, \, x \in \mathbb{R}$$

where $\mu_i \in \mathbb{R}, p > 0, a_i > 0$ *are known and* $\sigma > 0$ *is an unknown common scaling parameter. The mle of* σ *allows the functional representation*

$$\widehat{\sigma} = \frac{1}{n^{1/p}}|X_{(n)} - \nu|_{a,p}$$

where $X_{(n)} = (X_1, ..., X_n)^T, \nu = (\mu_1, ..., \mu_n)^T$ *and*

$$|x|_{a,p} = (\sum_{i=1}^n |\frac{x_i}{a_i}|^p)^{1/p}, \, x \in \mathbb{R}^n.$$

The density of $X_{(n)}$ *is*

$$f_{X_{(n)}}(x) = \frac{C_p^n}{\sigma^n a_1...a_n} g_{PE}(d_{\sigma B_{a,p}}(x - \nu)),$$

thus

$$P(n(\frac{\widehat{\sigma}}{\sigma})^p < t) = \Phi_{g_{PE},B_{a,p},0_n}(t^{1/p}B_{a,p})$$

where $B_{a,p} = \{x \in \mathbb{R}^n : |x|_{a,p} \leq 1\}$. *Note that* $\mathfrak{F}_{S_{a,p}}(t^{1/p}B_{a,p}, r)$ *does not depend on a. Hence,* $n(\frac{\widehat{\sigma}}{\sigma})^p$ *follows the* $\chi^p_{g_{PE}}$*-distribution with n d.f., independently of* $a = (a_1, ..., a_n)^T$.

While Examples 3.1 and 3.2 are restricted to the p-generalized Gaussian or Laplace distribution which is defined using the dgf $g_{PE}, p > 0$, Example 3.3 deals with measuring the same sets as before but with measures having another level distribution, especially allowing lighter and heavier distribution tails. Such tails may be of interest in various types of applications.

Example 3.3. *Let us assume that Y follows the probability distribution law* $\Phi_{g,K,\nu}$ *with* $\nu \in \mathbb{R}^n, K = B_{\sigma a,p}$ *and dgf g. Examples of dgf's are, besides* g_{PE}, *the Kotz type dgf* g_K *defined by* $g_K(r) = r^{M-1}e^{-\beta r^\gamma}, \beta, \gamma > 0, M + n > 1$, *and the Pearson-VII-type dgf* g_{PT7} *defined by* $g_{PT7}(r) = (1 + \frac{r}{m})^{-M}, M > n, m > 0$. *Note that the Student- and Cauchy-type dgfs are special cases of* g_{PT7}, *and that*

$$I(g_{PE}) = p^{n/p-1}\Gamma(n/p), I(g_K) = \frac{\Gamma((n+M-1)/\gamma)}{\gamma\beta^{(n+M-1)/\gamma}}, I(g_{PT7}) = m^n B(M-n, n)$$

where B denotes the Beta function. It follows with $B_{\sigma a,p} = \sigma B_{a,p}$ and

$$P(Y - \nu \in t^{1/p}B_{\sigma a,p}) = \frac{1}{I(g)}\int_0^\infty r^{n-1}g(r)\mathfrak{F}_{S_{a,p}}(t^{1/p}B_{\sigma a,p}, r)dr = \frac{1}{I(g)}\int_0^{t^{1/p}} r^{n-1}g(r)dr$$

that, independently of a,

$$\frac{d}{dt}P(n(\frac{\hat{\sigma}}{\sigma})^p < t) = \frac{d}{dt}P(|\frac{Y_1 - \nu_1}{a_1}|^p + ... + |\frac{Y_n - \nu_n}{a_n}|^p < t) = \frac{t^{n/p-1}g(t^{1/p})}{pI(g)} = f^\chi_{n,g,p}(t).$$

Note that our earlier representation of this density in [13] differs from the present one because of the (slightly) different notation for the dgf g.

Example 3.4. *Let $X_i \sim f_p(.; 0, \sigma_1), i = 1, ..., n_1$ and $Y_i \sim f_p(.; 0, \sigma_2), i = 1, ..., n_2$ where $p > 0$ is known and $\sigma_1, \sigma_2 > 0$ are unknown. These rv's are assumed to be completely independent. We define $X = (X_1, ..., X_{n_1})^T$ and $Y = (Y_1, ..., Y_{n_2})^T$. It follows from the well known theory of exponential families that the statistic $T = (|X|_p^p, |Y|_p^p)$ is sufficient for (σ_1, σ_2). The rv $Z = (X^TY^T)^T$ has the density*

$$f_Z((x^T, y^T)^T) = \frac{C_p^{n_1+n_2}}{\sigma_1^{n_1}\sigma_2^{n_2}}g_{PE}(|(x^T, y^T)^T|_{p,\sigma_1,\sigma_2}), x \in \mathbb{R}^{n_1}, y \in \mathbb{R}^{n_2}$$

with

$$|(x^T, y^T)^T|_{p,\sigma_1,\sigma_2} = (\frac{|x|_p^p}{\sigma_1^p} + \frac{|y|_p^p}{\sigma_2^p})^{1/p}.$$

Thus, for every measurable function $h : \mathbb{R}_+^2 \to \mathbb{R}$,

$$P(h(T) < t) = \Phi_{g_{PE}, B_{p,\sigma_1,\sigma_2}}(B(t)), \tag{8}$$

$B(t) = \{(x^T, y^T)^T : h(|x|_p^p, |y|_p^p) < t\}, B_{p,\sigma_1,\sigma_2} = \{(x^T, y^T)^T \in \mathbb{R}^{n_1+n_2} : |(x^T, y^T)^T|_{p,\sigma_1,\sigma_2 \le 1}\}.$

Applications of formula (8) allow the derivation of geometric measure representations of exact distributions of statistics such as $\alpha|X|_p^p + \beta|Y|_p^p$, $\frac{|X|_p^p}{|Y|_p^p}$ and $\frac{|X|_p^p}{|X|_p^p + |Y|_p^p}$, under the non-standard model assumptions made here.

A non-constant ipf of the set $B(t)$ has been dealt with for different functions h and under different parameter assumptions in earlier papers of the author and several coauthors, see [14].

Example 3.5. *For data in [8] reporting the profits at the box office and the number of sold home videos, in [10] the authors study fitting a linear regression model with random errors distributed according to an exponential power distribution. When analyzing the residuals they present Q-Q-plots for both the exponential power distribution with the estimated parameter $p = 2, 386877$ and the normal distribution ($p = 2$). The observer's subjective impression after a visual inspection of these plots may be that one cannot really be sure in preferring one of the two error models, this way.*

This example, which was presented by the authors only for the purpose to show the use of the functions implemented by them to fit a linear regression model if errors are possibly exponentially power distributed, gives rise to throw up in a similar two-dimensional situation the following standard question of statistical practice: how large should a sample size be to
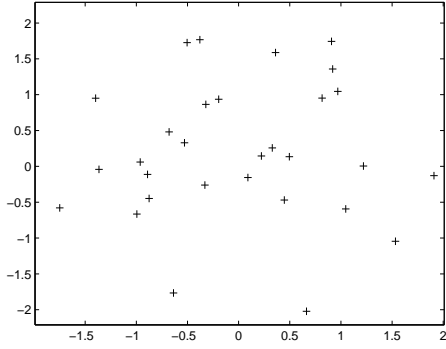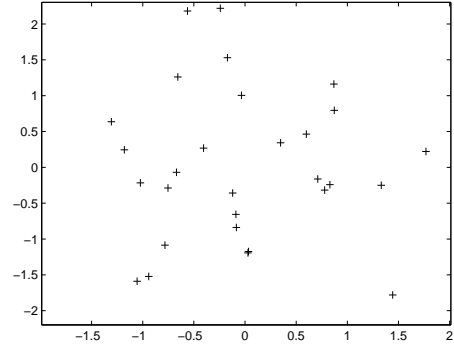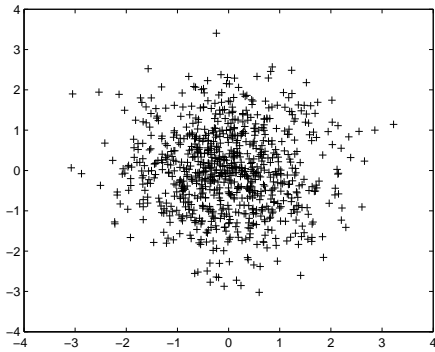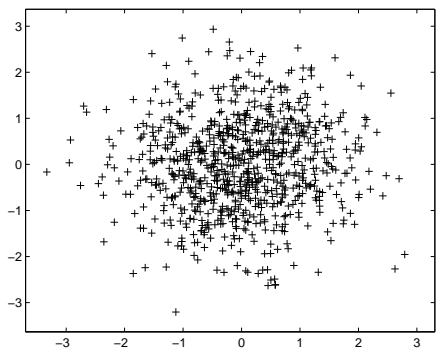
(a) $p = 2$ (b) $p = 2,388677$

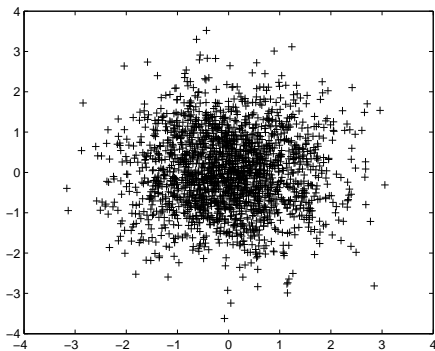Figure 6: Generalized normal distribution: (small) sample size $k = 30$
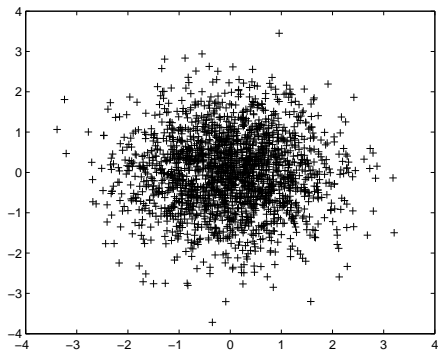

(a) $p = 2$ (b) $p = 2,388677$

Figure 7: (Medium) sample size $k = 850$

(a) $p = 2$ (b) $p = 2,388677$

Figure 8: (Large) sample size $k = 2000$

13

make a practical decision based upon a visual inspection "relatively safe" ? In particular, how large should a sample size be for the observer being able to visually choose between the two two-dimensional p-generalized normal distributions with parameters $p = 2$ and $p = 2,388677$?

This question, clearly, is not formulated in a strong mathematical way, and will not be answered in such way, here. Instead, we present Figures 6-8 showing that one can hardly distinguish this way between the parameters $p = 2$ and $p = 2,388677$ of the two-dimensional p-generalized normal distribution even if sample sizes are large. As a consequence, one may ask, e.g., for a mathematical method yielding a sure decision about the first decimal place of parameter p, say. For a certain general 20-percent sensitivity and g-robustness principle, established when dealing with another particular problem, we refer to Application 2 and Section 3 in [6].

Example 3.6. (a) Simulation in dimension one. *There are several possibilities for simulating the p-generalized normal distribution. The p-generalized polar method and the p-generalized rejecting polar method are established in [7] and compared with several methods known from the literature. Moreover, the resulting recommendation for using which of the methods in which situation is realized by S. Kalke in the R-module 'pgnorm'.*

(b) Simulation in dimension two. *If dimension is $n = 2$, and $K = B_p$ for some $p > 0$ then $S = \{(x_1, x_2)^T \in \mathbb{R}^2 : |x_1|^p + |x_2|^p = 1\}$. A random vector $(U_{p,1}, U_{p,2})^T$ following the star-generalized uniform distribution ω_S on $\mathcal{B}_S$ allows the stochastic representation*

$$(U_{p,1}, U_{p,2})^T \stackrel{d}{=} (\cos_p(\Phi), \sin_p(\Phi))^T$$

where the generalized trigonometric functions

$$\cos_p(\varphi) = \frac{\cos \varphi}{N_p(\varphi)}, \ \sin_p(\varphi) = \frac{\sin \varphi}{N_p(\varphi)} \ \text{with} \ N_p(\varphi) = (|\sin \varphi|^p + |\cos \varphi|^p)^{1/p}$$

are introduced in [12] and applied to the class of $l_{n,p}$-symmetric distributions in [13], and the polar angle Φ has the pdf

$$f_\Phi(\varphi) = \frac{p\Gamma(\frac{2}{p})}{4(\Gamma(\frac{1}{p})N_p(\varphi))^2} I_{(0,2\pi)}(\varphi).$$

For a graphical representation of this function, we refer to [7].

Let Z be a uniformly on the interval $(0,1)$ distributed random variable, $z_1, ..., z_k$ realizations of it in k independent trials, and put $F_\Phi(\varphi) = \int\limits_0^\varphi f_\Phi(\psi)d\psi, 0 < \varphi < 2\pi$. Numerically solving the equations

$$F_\Phi(\varphi_i) = z_i \ \text{with starting value} \ \varphi_{i,0} = 2\pi z_i, \ i = 1, ..., k$$

yields the realizations $(u_{p,1}^{(1)}, u_{p,2}^{(1)}), ..., (u_{p,1}^{(k)}, u_{p,2}^{(k)})$ of $(U_{p,1}, U_{p,2})^T$ from k independent trials.

Moreover, let the random variable R independently of $(U_{p,1}, U_{p,2})^T$ follow the density $f_R(r) = f(r; 1/3; 1.6, 2.6; 3, 4)$ where with suitably chosen $a < b < c < d$,

$$f(r; p; a, b; c, d) = pI_{(a,b)}(r) + (1 - p)I_{(c,d)}(r), 0 < p < 1.$$

If $r_1, ..., r_k$ are realizations of R in independent trials then $y_i = r_i(u_{p,1}^{(i)}, u_{p,2}^{(i)}), i = 1, ..., k$ are realizations of a so called two-layer $l_{2,p}$-symmetrically distributed random vector Y which are represented in Figure 9(a). Similarly, Figure 9(b) is drawn with $f_R(r) = f(r; 2/3; 1.9, 2.9; 3, 4)$. Still using the latter function but having smaller sample sizes, Figure 10 is generated. It turns out that it becomes impossible to further visually distinguish between the two layers for sample sizes becoming too small.
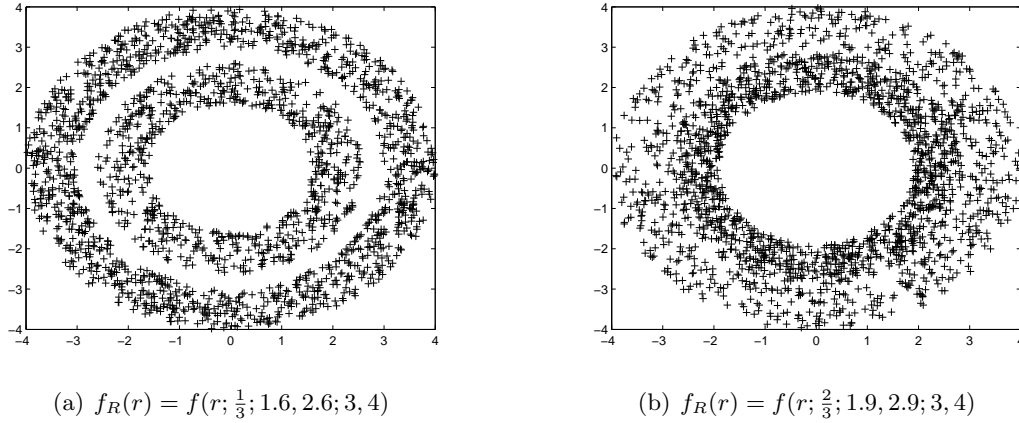

(a) $f_R(r) = f(r; \frac{1}{3}; 1.6, 2.6; 3, 4)$ (b) $f_R(r) = f(r; \frac{2}{3}; 1.9, 2.9; 3, 4)$

Figure 9: Two-layer $l_{2,p}$-symmetric distributions sampled in $k = 2000$ independent trials with $p = 2,388677$
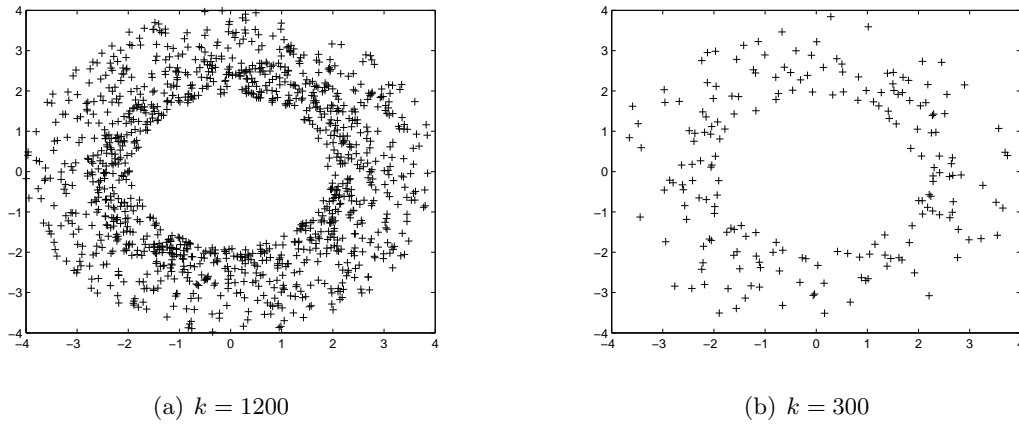

(a) $k = 1200$ (b) $k = 300$

Figure 10: Independent realizations of two-layer $l_{2,p}$-symmetrically distributed random points where p and f_R are as in Figure 9(b)

We remark additionally that Figures 9 and 10 do not represent elliptically contoured distributions as one might argue at first glance. This again supports the above discussion on introducing 20-percent g-robust decisions.

Finally, we notice that Figures 1-8 are generated with the level density function of the

two-dimensional p-generalized normal density,

$$f_R(r) = \frac{1}{p^{\frac{2}{p}-1}\Gamma(\frac{2}{p})} re^{-\frac{r^p}{p}}, \, r > 0.$$

4 Proofs

The general method of proof in this paper can be divided into two main parts. Using the properties of the support function of a convex body, h_K, in the first part it will be shown that the absolute value of the Jacobian of a certain transformation may be interpreted in terms of the normal vector N to the boundary of K. This allows according to Lemma 1 in [14] to represent the surface measure $\mathfrak{O}_S$ as an integral of $h_K(N)$. The second part of the method of proof deals with a relation between the functional $h_K(N)$ and the Minkowski functional of a suitably defined set K^*, or K^o. While, in the convex case, K^* is extensively studied, K^o yet has to be found in the most general case when K is radially concave w.r.t. a fan in $\mathbb{R}^n$.

Proof of Theorem 3.1. The support function of the convex body K is defined as

$$h_K(u) = \sup\{u^T y : y \in K\}, \, u \in \mathbb{R}^n. \tag{9}$$

Recall that if $u \in S^{(n-1)}$ then $h_K(u)$ describes the distance from the origin to the hyperplane with outer normal vector u and supporting K. For compactness and continuity reasons, the supremum is always attained,

$$\forall u \in S^{(n-1)} \exists x_S(u) \in S : h_K(u) = u^T x_S(u). \tag{10}$$

The set of all such points x_S is called the supporting set of K at u. If the norm is smooth then $x_S(u) \in S \cap T(u)$ where $T(u)$ is the tangent hyperplane to S at the point $x_S(u)$ with $T(u)$ being orthogonal to u. If K is strongly convex then the supporting set of K at u consists of just one point, thus $x_S(u)$ being then always uniquely defined. Note that it may happen that a point $\xi \in S$ satisfies $\xi = x_S(u)$ for more than one point $u \in S^{(n-1)}$. To see this, assume that S contains corner points, and let ξ be such a corner point of S. As a consequence, even the union of all supporting sets of a convex body may be finite.

According to Lemma 1 in [14],

$$\mathfrak{O}_S(A) = \int\limits_{G(A)} |(\vartheta^T, \eta(\vartheta)) N(\vartheta)| d\vartheta$$

where the function $\vartheta \mapsto \eta(\vartheta)$ is chosen such that $d_K((\vartheta^T, \eta(\vartheta))^T) = 1$ describes the boundary S of K. Note that here $x = (\vartheta^T, \eta(\vartheta))^T \in S$ and a.e. $N(\vartheta) = (\nabla\eta(\vartheta), -1)^T$ is the outer normal vector of S at $(\vartheta^T, \eta(\vartheta))^T$, thus $(\vartheta^T, \eta(\vartheta)) N(\vartheta) > 0$. It follows from (10) that

$$h_K(n(\vartheta)) = n(\vartheta)^T (\vartheta^T, \eta(\vartheta))^T, \, n(\vartheta) = N(\vartheta)/|N(\vartheta)|_E$$

where $|.|_E$ means the Euclidean norm. By the homogeneity property of h_K,

$$\mathfrak{O}_S(A) = \int\limits_{G(A)} h_K(N(\vartheta)) d\vartheta. \tag{11}$$

The theorem follows by the well known fact that $h_K = d_{K^*}$ □

16

Proof of Theorem 3.2. We consider $(\vartheta^T, \eta(\vartheta))^T \in S \cap C_i$ and denote the (a.e. defined) inner normal vector to the boundary S of the antinorm ball K at $(\vartheta^T, \eta(\vartheta))^T$ by $N(\vartheta)$. Further we put $n(\vartheta) = N(\vartheta)/|N(\vartheta)|_2$ and $u = -n(\vartheta)$. Then $u \in S^{(n-1)} \cap C_i$ and because $K \in AN1$,

$$|(\vartheta^T, \eta(\vartheta))^T n(\vartheta)| = |x_{S,i}^T(n(\vartheta)) n(\vartheta)| = \inf\{n^T(\vartheta) y : y \in S \cap C_i\}.$$

Thus, $|(\vartheta^T, \eta(\vartheta))^T n(\vartheta)| = h_K^{\mathfrak{F}}(n(\vartheta))$, and

$$\mathfrak{O}_S(A) = \int\limits_{G(A)} h_K^{\mathfrak{F}}(N(\vartheta)) d\vartheta. \tag{12}$$

The proof will be finished by the following lemma.

$\square$

Lemma 4.1. *The anti-support function of K w.r.t. $\mathfrak{F}$ is equal to the distance function of K^o, $h_K^{\mathfrak{F}} = d_{K^o}$.*

Proof. The radial function of the radially concave star-shaped set K^o is, on the one hand, by definition

$$\varrho_{K^o}(u) = \sup\{\lambda \geq 0 : \lambda u \in K^o\} = \frac{1}{h_K^{\mathfrak{F}}(u)}, u \in S^{(n-1)},$$

and allows, on the other hand, the representation

$$\varrho_{K^o}(u) = \sup\{\lambda \geq 0 : u \in \frac{1}{\lambda} K^o\} = 1/\inf\{\mu \geq 0 : u \in \mu K^o\} = 1/d_{K^o}(u).$$

Thus $h_K^{\mathfrak{F}} = d_{K^o}$.

$\square$

5 Discussion

To make both the similarity and the difference between Theorems 3.1 and 3.2 more visible, let us remark that K^* allows a representation looking similar to that of K^o,

$$K^* = \{y \in \mathbb{R}^n : y^T x \leq 1, \forall x \in K\} = \{\lambda(u) u : 0 \leq \lambda(u) h_K(u) \leq 1, u \in S^{(n-1)}\}.$$

For dealing with a combined example where both Theorems 3.1 and 3.2 apply, we recall that the function $x \mapsto |x|_{a,p}$ is a norm if $p \geq 1$ and, according to [11], an antinorm if $0 < p \leq 1$. Thus, $K = B_{a,p}$ is convex if $p \geq 1$ and radially concave w.r.t. the standard fan $\mathfrak{F}$ if $0 < p \leq 1$. Let q be defined by the equation $\frac{1}{p} + \frac{1}{q} = 1$, and $\frac{1}{a} = (\frac{1}{a_1}, ..., \frac{1}{a_n})^T$. Note that if $p > 1$ then $K^* = B_{\frac{1}{a}, q}$, and if $0 < p < 1$ then $q < 0$ and

$$K^o = \{x \in \mathbb{R}^n : \sum_{i=1}^n \frac{1}{|a_i x_i|^{|q|}} \geq 1\}$$

is a semi-antinorm ball. For an illustration of such sets, see [11]. Note that a and p are independently dealt with when constructing the sets K^* and K^o.

17

Remark 5.1. *In the applications of Theorem 3.1, the surface content of $[\frac{1}{r}(B-\nu)] \cap S$ (which may be considered as an indivisible of the set $B-\nu$) is always measured w.r.t. the metric d_{K^*}. Successful applications of this generalized method of indivisibles are surveyed in [14]. The representation of Theorem 3.1 generalizes those presented in the latter and several earlier papers.*

Remark 5.2. *The set K^o in Section 3.2 is radially concave, thus d_{K^o} is a semi-antinorm.*

Proof. We show that if x_1 and x_2 are from $\bar{K}^o \cap C$ for some $C \in \mathfrak{F}$ then $\lambda x_1 + (1-\lambda)x_2 \in \bar{K}^o \cap C$ for $0 < \lambda < 1$ where $\bar{K}^o$ means the complement of the set K^o. Let $x_1 = \lambda_1 u_1, x_2 = \lambda_2 u_2$ with $u_i \in S^{(n-1)}, \lambda_i \geq \frac{1}{h_K^{\mathfrak{F}}(u_i)}, i = 1, 2$. Then

$$h_K^{\mathfrak{F}}(\lambda x_1 + (1-\lambda)x_2) = h_K^{\mathfrak{F}}(\lambda \lambda_1 u_1) + h_K^{\mathfrak{F}}((1-\lambda)\lambda_2 u_2) \geq \frac{\lambda \lambda_1}{\lambda_1} + \frac{(1-\lambda)\lambda_2}{\lambda_2} = 1$$

$\square$

References

[1] T. W. Anderson, Kai-Tai Fang, Theory and applications of elliptically contoured and related distributions. Technical report 24, Department of Statistics, Stanford University, (1990).

[2] N. Balakrishnan, Chin-Diew Lai, Continuous bivariate distributions. Springer (2009).

[3] M. Bilodeau, D.Brenner, Theory of Multivariate Statistics. N.Y.: Springer (1999).

[4] A. Cuevas, M. Febrero and R. Freiman, Estimating the number of clusters. Canad. J. Statist. **28**, 367-382,(2000).

[5] K.-T. Fang, Y. Zhang, Generalized Multivariate Analysis. Springer (1990).

[6] C. Ittrich, D. Krause and W.-D. Richter, Probabilities and large quantiles of noncentral generalized chi-square distributions. Statistics **34**: 53-101, (2000).

[7] S. Kalke, W.-D. Richter, Simulation of the p-generalized Gaussian distribution. Journal of Statistical Computation and Simulation **83**:4, 639-665, (2013).

[8] D. M. Levine, T. C. Krehbiel and M. L. Berenson, Business statistics: A first course. Prentice-Hall, (2000).

[9] E. Liebscher, A semiparametric density estimator based on elliptical distributions. J. Multiv. Anal.**92**,205-225 (2005).

[10] A. M. Mineo, M. Ruggieri, A software tool for the exponential power distribution: The normal package. Journal of Statistical Software **12**, 4 (2005), 1-24.

[11] M. Moszyńska, W.-D. Richter, Reverse triangle inequality. Antinorms and semi-antinorms. Studia Scientiarum Mathematicarum Hungarica **49** (1), 120-138 (2012).

[12] W.-D. Richter, Generalized spherical and simplicial coordinates. J. Math. Anal. Appl. **336**, 1187-1202 (2007).

[13] W.-D. Richter, Continuous $l_{n,p}$-symmetric distributions. Lithuanian Math. J. **49**,1, 93-108 (2009).

[14] W.-D. Richter, Geometric disintegration and star-shaped distributions. J. Stat. Distr. Appl. (2014), **1**: 20.

[15] M. Rudemo, H. Stryhn, Boundary estimation for star-shaped objects. In: Change-point Problems, eds. E. Carlstein, H.-G. Müller and D. Siegmund. Institute of Mathematical Statistics, Hayward, CA, 276-283 (1994).

[16] A. B. Tsybakov, On nonparametric estimation of density level sets. Ann. Statist. **25**, 948-969 (1997).

[17] G. Walther, Granulometric smoothing. Ann. Statist. **25**, 2273-2299 (1997).