

RESEARCH

Open Access



Statistical reasoning in dependent p -generalized elliptically contoured distributions and beyond

Testing scaling parameters, the role semi-inner products play, and simulating star-shaped distributed random vectors

Wolf-Dieter Richter 

Correspondence:
wolf-dieter.richter@uni.rostock.de
University of Rostock, Institute of
Mathematics, Ulmenstraße 69, Haus
3, 18057 Rostock, Germany

Abstract

First, likelihood ratio statistics for checking the hypothesis of equal variances of two-dimensional Gaussian vectors are derived both under the standard $(\sigma_1^2, \sigma_2^2, \rho)$ -parametrization and under the geometric (a, b, α) -parametrization where a^2 and b^2 are the variances of the principle components and α is an angle of rotation. Then, the likelihood ratio statistics for checking the hypothesis of equal scaling parameters of principle components of p -power exponentially distributed two-dimensional vectors are considered both under independence and under rotational or correlation type dependence. Moreover, the role semi-inner products play when establishing various likelihood equations is demonstrated. Finally, the dependent p -generalized polar method and the dependent p -generalized rejection-acceptance method for simulating star-shaped distributed vectors are presented.

Keywords: Modeling with correlation and variances of Euclidean coordinates, Modeling with rotation and variances of principle components, Geometric parametrization, Likelihood ratio, p -generalized Fisher distribution, Semi-inner product, Dependent p -generalized polar method, Dependent p -generalized rejection-acceptance simulation, Star-shaped distributions

PACS: 02.50.-r, 02.50.Ng, 02.70.-Rr, 02.70.Uv

Mathematics Subject Classification: 62F03, 62F10, 62H15, 62E10

1 Introduction

One of the classical statistical problems deals with comparing variances. Results for quite arbitrary pairs of dependent random variables are due to Morgan (1939), Pitman (1939), Tiku and Balakrishnan (1986), McCulloch (1987), Wilcox (1990) and Mudholkar et al. (2003) and are recently reviewed in Wilcox (2015). For a survey of various practical applications in Snedecor and Cochran (1967), Lord and Novick (1968), Games et al. (1972), Levy (1976) and Rothstein et al. (1981), see again Wilcox (2015). A closely related but nevertheless rather different study in Richter (2016) compares scaling parameters of two-dimensional axes-aligned p -generalized elliptically contoured distributions. Such distributions show independence if the density generating function is that of the p -generalized

Gaussian law and l_p -dependence if the density generating function is of another type, but they show no rotational or correlation type dependence. The present paper is aimed now to study correlation type dependence modeling within the family of p -power exponential distribution laws.

It is well known that two jointly Gaussian distributed random variables are independent if and only if they are uncorrelated. The density level sets of such a vector are axes-aligned ellipses. If the components of a two-dimensional Gaussian vector are not independent then the vector may be constructed by rotating through its distribution center an axes-aligned elliptically contoured distributed Gaussian vector that has heteroscedastic components. The correlation coefficient may be expressed in such situation in terms of the angle of rotation and the ratio of variances, see Dietrich et al. (2013). This type of dependence between two random variables is called here a rotational or correlation type dependence. Basic facts on modeling two-dimensional Gaussian vectors with correlation and variances of Euclidean coordinates on the one hand and with rotation and variances of principle components on the other hand will be summarized in the presented paper. Considering these two models side by side demonstrates different aspects of 'standard' modeling with the more stochastically interpretable parameters and of 'flexible' modeling with the more geometrically motivated parameters.

It is outlined in Wilcox (2015) that "seemingly the best-known technique for testing $H_0 : \sigma_1^2 = \sigma_2^2$ is a method derived by Morgan (1939) and Pitman (1939). Letting $U = X + Y$ and $V = X - Y$, if the null hypothesis is true, then ρ_{UV} , Pearson's correlation between U and V , is zero. So testing H_0 can be accomplished by testing $\rho_{UV} = 0$." We do not consider here the test problem in the same full generality as in Wilcox (2015) where the joint distribution of X and Y is not basically restricted to belong to the families of p -generalized elliptically contoured or star-shaped distributions. While it is proved in Wilcox (2015) that certain heteroscedastic consistent estimators perform well in certain cases of heavy tailed distributions, here we use case sensitive estimators depending on the given value of the shape-tail parameter p , see Sections 3 and 4 and recognize the consequences drawn in Section 5. Note that cases of heavier and lighter than Gaussian distribution tails are observed here in dependence of whether $p \in (0, 2)$ or $p > 2$, respectively. A study demonstrating far and narrow tail effects when sampling from those distribution classes can be seen in Richter (2015a).

In the case of a two-dimensional Gaussian distribution $\Phi_{\mu, \Sigma}$, it turns out that the class of distributions satisfying H_0 is the union of the following two subsets. The elements of the first one are the spherical Gaussian distributions and the second one contains all elliptically contoured Gaussian distributions having the lines $y = +(-)x$ as the main axes of their density level ellipses. Any number from the interval $(-1, 1)$ is attained by the correlation coefficient of a suitably chosen element from the latter subset. Thus, H_0 covers two quite different cases of correlation and uncorrelation. We modify here the null hypothesis in a way that one of these two subsets is not included.

One of the likelihood equations needed to be solved for constructing the likelihood ratio statistic for testing the just mentioned modified hypothesis is formulated here on using a so-called semi-inner product in the sample space. This rises the question whether this analytical tool plays also a role in estimating location. We give a positive answer to this question in the case of axes-aligned p -power exponential distributions.

The paper is structured as follows. Gaussian correlation models and likelihood ratio tests for checking equality of variances of two dependent random variables are studied in Section 2. The content of this section is of some interest of its own although it might be partly known to the reader. Testing equality of scaling parameters of axes-aligned p -power exponential distributions is dealt with in Section 3. The more general results are presented in Sections 4–6. Section 4 deals with testing equality of scaling parameters of principal components of general, i.e. arbitrarily rotated, p -generalized elliptically contoured p -power exponential distributions. Derivations are omitted in the sections on Gaussian and axes-aligned p -generalized elliptically contoured distributions. They can be considered being standard and follow also from proving the more general results in Section 4. Throughout Sections 2–4, we restrict our consideration to the case of known expectations. Practical examples of this type are given in Richter (2016). Section 5 gives a new geometric-analytical insight into estimating the location parameter of the p -power exponential, or p -generalized Gaussian or Laplace, law using semi-inner products in the sample space.

Differently from the situation of statistics in Gaussian sample distributions, many statistical questions in p -generalized Gaussian and more general star-shaped sample distributions cannot yet fully be answered in a theoretical way. For intermittent empirical studies, and much beyond it, methods for simulating such distributions are needed. Generalizing the methods in Kalke and Richter (2013), Section 6 presents corresponding direct and acceptance-rejection methods and indicates how to extend to the dependent p -generalized multivariate case the classical and the rejecting polar methods in Box and Muller (1958) and Marsaglia and Bray (1964), respectively.

2 Likelihood ratio tests for scaling parameters in two-dimensional Gaussian distributions

Testing equality of scaling parameters can be interpreted in Gaussian models at least in two different ways. We deal here with equality of variances of the marginal variables or Euclidean coordinates if the Gaussian density is given in the classical $(\sigma_1^2, \sigma_2^2, \varrho)$ variances-correlation parametrization, and with equality of variances of principal components if the Gaussian density is given in the geometric (a, b, α) -parametrization from Dietrich et al. (2013) where a^2 and b^2 are the variances of the principle components and α is an angle of rotation.

2.1 The common $(\sigma_1^2, \sigma_2^2, \varrho)$ -parametrization

In this section, we consider the marginal variables variances-correlation (mvv-c) model. Likelihood ratio tests with respect to the equality of two variances will be given separately for the cases of a known and an unknown correlation coefficient. Let $(X_i, Y_i)^T, i = 1, \dots, n$ be independent Gaussian random vectors following the density $\varphi_{\mu, \Sigma}(\cdot, \cdot) = \varphi(\cdot, \cdot | \sigma_1, \sigma_2, \varrho)$ where $\mu = (\mu_1, \mu_2)^T$ is a vector from \mathbb{R}^2 and $\Sigma = \begin{pmatrix} \sigma_1^2 & \varrho\sigma_1\sigma_2 \\ \varrho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ is a positive definite matrix, and let $(x_i, y_i)^T, i = 1, \dots, n$ be a corresponding concrete sample. We introduce the likelihood function

$$L(\sigma_1, \sigma_2, \varrho) = \prod_{i=1}^n \varphi(x_i, y_i | \sigma_1, \sigma_2, \varrho)$$

and its restriction to the case of equal variances, $\tilde{L}(\sigma, \varrho) = L(\sigma, \sigma, \varrho)$.

2.1.1 The case of an unknown correlation coefficient

We intend now to decide between the two hypotheses

$$H_0 : \sigma_1 = \sigma_2 \quad \text{and} \quad H_A : \sigma_1 \neq \sigma_2$$

using the likelihood ratio test statistic $Q = \tilde{L}(\tilde{\sigma}, \tilde{\varrho}) / L(\hat{\sigma}_1, \hat{\sigma}_2, \hat{\varrho})$ where $(\hat{\sigma}_1, \hat{\sigma}_2, \hat{\varrho}) = mle(\sigma_1, \sigma_2, \varrho)$ and $(\tilde{\sigma}, \tilde{\varrho}) = mle|_{H_0}(\sigma, \varrho)$ are maximum likelihood and restricted to H_0 such estimators, respectively. Standard calculations show that Q allows the representation

$$\frac{Q^{2/n}}{4} = \frac{\Sigma_x^2 \Sigma_y^2 - \Sigma_{xy}^2}{\left(\Sigma_x^2 + \Sigma_y^2\right)^2 - 4\Sigma_{xy}^2} \quad (1)$$

where

$$\Sigma_x^2 = \sum_1^n (x_i - \mu_1)^2, \Sigma_y^2 = \sum_1^n (y_i - \mu_2)^2, \Sigma_{xy} = \sum_1^n (x_i - \mu_1)(y_i - \mu_2).$$

Let $\alpha \in (0, 1)$. According to the likelihood ratio rule, H_0 will be rejected if $Q < t_\alpha$ where t_α is chosen from the interval $(0, 1)$ in a way such that

$$P(Q < t_\alpha) |_{H_0} = \alpha.$$

A restatement of this size α -test is based upon the following alternative representation of the likelihood ratio,

$$Q^{2/n} = \left(\frac{\hat{\sigma}_1 \hat{\sigma}_2}{\tilde{\sigma}^2}\right)^2 \frac{1 - \hat{\varrho}^2}{1 - \tilde{\varrho}^2},$$

where

$$\hat{\sigma}_1^2 = \Sigma_x^2/n, \hat{\sigma}_2^2 = \Sigma_y^2/n, \hat{\varrho} = \Sigma_{xy}/(\Sigma_x \Sigma_y)$$

and

$$\tilde{\sigma}^2 = (\Sigma_x^2 + \Sigma_y^2)/(2n), \tilde{\varrho} = \Sigma_{xy}/(n\tilde{\sigma}^2).$$

Rewording the corresponding likelihood ratio decision rule, it is then a size $(\alpha_1 + \alpha_2)$ -test to reject H_0 if

$$\frac{\Sigma_x^2}{\Sigma_y^2} \in (0, \lambda_{\alpha_1}] \cup [\lambda_{1-\alpha_2}, \infty)$$

where $\lambda_q, q \in (0, 1)$, is suitably chosen from $(0, \infty)$ such that

$$P\left(\frac{\Sigma_x^2}{\Sigma_y^2} < \lambda_q\right) |_{H_0} = q.$$

Here, Σ_x^2/Σ_y^2 is the ratio of two dependent Chi-squared distributed random variables. The distributions of all statistics considered here and in later sections may be simulated using the methods presented in Section 6. Alternatively, the geometric measure representation in Richter (2014) may be used to establish the exact distributions of several of these statistics, or at least to derive suitable approximations.

2.1.2 The case of a known correlation coefficient

Let $\varrho = \varrho_0$ be a known number and put $L(\sigma_1, \sigma_2) = L(\sigma_1, \sigma_2, \varrho_0)$ and $\tilde{L}(\sigma) = L(\sigma, \sigma, \varrho_0)$.

The likelihood ratio $Q = \sup_{\sigma} \tilde{L}(\sigma) / \sup_{\sigma_1, \sigma_2} L(\sigma_1, \sigma_2)$ allows the representation

$$\frac{Q^{1/n}}{2} = \frac{\Sigma_x \Sigma_y - \varrho_0 \Sigma_{xy}}{\Sigma_x^2 + \Sigma_y^2 - 2\varrho_0 \Sigma_{xy}}.$$

Let $\alpha \in (0, 1)$. The likelihood ratio decision rule leads to rejecting H_0 if $Q < t_\alpha$, $\alpha \in (0, 1)$ satisfies $P(Q < t_\alpha) |_{H_0} = \alpha$ or, equivalently, if $Q^{1/n}/2 < z_\alpha$ for a suitably chosen $z_\alpha = z(t_\alpha) \in (0, 1/2)$, that is, if

$$1 - \varrho_0 (1 - 2z_\alpha) \frac{\Sigma_{xy}}{\Sigma_x \Sigma_y} < z_\alpha \left(\frac{\Sigma_x}{\Sigma_y} + \frac{\Sigma_y}{\Sigma_x} \right).$$

2.2 The geometric (a, b, α) -parametrization

We consider now the principal components variances-rotation (pcv-r) model. For simplicity, we assume that $\mu_1 = \mu_2 = 0$. The geometric parametrization of the Gaussian density is then

$$\varphi^*(x, y | a, b, \alpha) = \frac{1}{2ab\pi} \exp \left\{ -\frac{1}{2} \left[\left(\frac{x \cos \alpha + y \sin \alpha}{a} \right)^2 + \left(\frac{-x \sin \alpha + y \cos \alpha}{b} \right)^2 \right] \right\},$$

see Dietrich et al. (2013). Here,

$$a = (\sigma_1^2 \cos^2 \alpha + \sigma_2^2 \sin^2 \alpha + 2\varrho \sigma_1 \sigma_2 \sin \alpha \cos \alpha)^{1/2},$$

$$b = (\sigma_2^2 \cos^2 \alpha + \sigma_1^2 \sin^2 \alpha - 2\varrho \sigma_1 \sigma_2 \sin \alpha \cos \alpha)^{1/2},$$

and

$$\alpha = \begin{cases} 0 & \text{if } \varrho = 0 \\ \gamma + \begin{cases} 0 & \text{if } \varrho(\sigma_1 - \sigma_2) > 0 \\ \pi/2 & \text{if } \varrho(\sigma_1 - \sigma_2) < 0 \end{cases} & \text{if } \varrho \neq 0, \sigma_1 \neq \sigma_2 \\ \pi/4 & \text{if } \varrho \neq 0, \sigma_1 = \sigma_2 \end{cases}$$

with $\gamma = \frac{1}{2} \arctan(2\varrho \sigma_1 \sigma_2 / (\sigma_1^2 - \sigma_2^2))$. We put $\arctan(+(-)\infty) = +(-)\pi/2$ and remark that a^2 and b^2 are the variances of principal components of the related Gaussian random vector. The Euclidean coordinates of such a vector are correlated if $\varrho \neq 0$ and may then also be called rotational dependent because then $\alpha \neq 0$.

For testing equality of variances of principle components

$$H_0 : a = b \quad \text{vs.} \quad H_A : a \neq b,$$

we introduce the likelihood function

$$L^*(a, b, \alpha) = \prod \varphi^*(x_i, y_i | a, b, \alpha)$$

and its restriction to H_0 , $\tilde{L}^*(a, \alpha) = L^*(a, a, \alpha)$.

2.2.1 The case of an unknown α

The likelihood ratio statistic $Q^* = \max_{a, \alpha} \tilde{L}^*(a, \alpha) |_{H_0} / \max_{a, b, \alpha} L^*(a, b, \alpha)$ in case α is to be estimated, allows the representation

$$\frac{(Q^*)^{2/n}}{4} = \frac{I(\hat{\alpha}) J(\hat{\alpha})}{(\Sigma_x^2 + \Sigma_y^2)^2}$$

where

$$I(\alpha) = (\cos \alpha)^2 \Sigma_x^2 + (\sin \alpha)^2 \Sigma_y^2 + 2(\sin \alpha \cos \alpha) \Sigma_{xy},$$

$$J(\alpha) = (\sin \alpha)^2 \Sigma_x^2 + (\cos \alpha)^2 \Sigma_y^2 - 2(\sin \alpha \cos \alpha) \Sigma_{xy}$$

and the maximum likelihood estimator $\hat{\alpha} = \text{mle}(\alpha)$ is

$$\hat{\alpha} = \frac{1}{2} \arctan \left(2 \frac{\Sigma_{xy}}{\Sigma_y^2 - \Sigma_x^2} \right).$$

If H_0 is true then the correlation or rotational dependence of the Euclidean coordinates is zero. Correspondingly, no restricted under H_0 estimator of the angle of rotation α has any effect onto the statistic Q^* .

2.2.2 The case of a known α

If the angle of rotation α is known, the likelihood ratio allows the representation

$$\frac{(Q^*)^{2/n}}{4} = \frac{I(\alpha)J(\alpha)}{(\Sigma_x^2 + \Sigma_y^2)^2}.$$

The plug-in version of this statistic where, for unknown α , $\alpha = \hat{\alpha} = \text{mle}(\alpha)$, is just the statistic from the previous section. Differently from this situation, the likelihood ratio statistic in Section 2.1.1 using both the unrestricted and the restricted maximum likelihood estimators of α , is not such an immediate plug-in version of the statistic considered in Section 2.1.2.

3 Likelihood ratio test for scaling parameters in axes-aligned p -generalized elliptically contoured distributions

The present section is aimed to shortly summarize some results from the axes-aligned or independence case. To start with, we recall that the univariate p -power exponential distribution has the density

$$f_p(x; \mu, \sigma) = \frac{C_p}{\sigma} \exp \left\{ -\frac{|x - \mu|^p}{p\sigma^p} \right\}, \quad x \in R,$$

which is also called p -generalized Gaussian or Laplace density, $p > 0$. The parameter p controls both the shape of the density and the tail behaviour of the distribution and may therefore be called a shape-tail parameter. Note that $C_p = p^{1-1/p}/(2\Gamma(1/p))$ and the first and second order moments of a correspondingly distributed random variable X are

$$\mathbb{E}X = \mu \in R \quad \text{and} \quad V(X) = \sigma^2 \frac{\Gamma(3/p)}{\Gamma(1/p)}.$$

Moreover, such random variable X allows the stochastic representation

$$X \stackrel{d}{=} \mu + \sigma X_0$$

where X_0 follows the standard p -power exponential density, i.e. $X_0 \sim f_p(\cdot; 0, 1)$. Because of this representation, σ is called a scaling parameter. Note that $E|X - \mu|^p = \sigma^p$. Two independent such variables follow the joint product density

$$f(x, y | \sigma_1, \sigma_2) = \frac{C_p^2}{\sigma_1 \sigma_2} \exp \left\{ -\frac{|x - \mu_1|^p}{p\sigma_1^p} - \frac{|y - \mu_2|^p}{p\sigma_2^p} \right\}, \quad (x, y) \in R^2$$

having the distribution center $(\mu_1, \mu_2)^T \in R^2$ and whose level sets are axes-aligned p -generalized ellipses.

Note that the axes-aligned p -generalized elliptically contoured p -power exponential densities introduced this way should not be confused with functions of the type $f_{(X,Y)}(x,y) = C \exp \{-Q(x,y)^{p/2}\}$ with Q being a quadratic form. The latter type of densities has been considered in Kuwana and Kariya (1991), Gómez et al. (1998), Gómez-Villegas et al. (2011) and Dang et al. (2015) and may also be called elliptically contoured p -power exponential densities. The corresponding type of distributions may be considered as a particular Kotz type distribution within the broad family of elliptically contoured distributions, see Fang et al. (1990) and Nadarajah (2003).

Testing

$$H_0 : \sigma_1 = \sigma_2 \quad \text{vs.} \quad H_A : \sigma_1 \neq \sigma_2$$

in the model of the present section means checking equality of scaling parameters. Let $(X_i, Y_i), i = 1, \dots, n$ be independent random vectors following the density $f(\cdot, \cdot | \sigma_1, \sigma_2)$ and put $X_{(n)} = (X_1, \dots, X_n)^T$, $Y_{(n)} = (Y_1, \dots, Y_n)^T$. We still assume that the expectations μ_1 and μ_2 are known. In case of a true hypothesis H_0 , the test statistic

$$T = \frac{(\text{mle}(\sigma_1)/\sigma_1)^p}{(\text{mle}(\sigma_2)/\sigma_2)^p} = \left(\frac{\sigma_2}{\sigma_1} \right)^p \frac{|X_{(n)} - \mu_1 \mathbf{1}_n|_p^p}{|Y_{(n)} - \mu_2 \mathbf{1}_n|_p^p}$$

follows the p -generalized Fisher distribution with (n, n) d.f., $T|_{H_0} \sim F_{n,n}(p)$. The latter distribution was derived in Richter (2009). It can be considered as the distribution of the ratio of independent p -generalized Chi-squared distributed variables that were introduced in Richter (2007). The density of the p -generalized Fisher distribution with (n, n) degrees of freedom is according to Richter (2009)

$$f_{n,n,p}(t) = \frac{\Gamma(2n/p)}{(\Gamma(n/p))^2} \frac{t^{n/p-1}}{(1+t)^{2n/p}}, t > 0,$$

see Fig. 1. With the notation

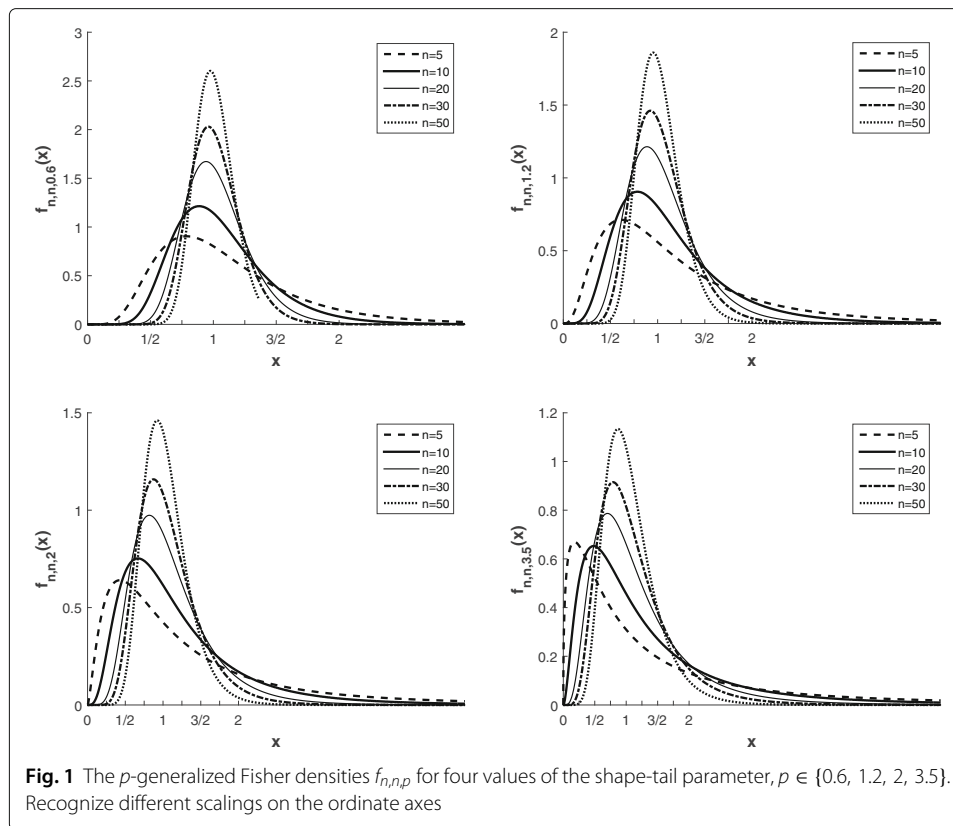
$$\Sigma_{p,X} = \left(\sum_{i=1}^n |X_i - \mu_1|^p \right)^{1/p} \quad \text{and} \quad \Sigma_{p,Y} = \left(\sum_{i=1}^n |Y_i - \mu_2|^p \right)^{1/p},$$

the statistic T can alternatively be represented as

$$T = \left(\frac{\sigma_2}{\sigma_1} \right)^p \frac{\Sigma_{p,X}^p}{\Sigma_{p,Y}^p}.$$

The decision rule according to which one rejects H_0 if $T < F_{n,n,\alpha_2}(p)$ or $T > F_{n,n,1-\alpha_1}(p)$ performs an exact size $(\alpha_1 + \alpha_2)$ -test. This test turns out to be the corresponding likelihood ratio test. Figures 2, 3 and 4 deal with the performance of this test showing histograms of simulation results for the test statistic T , under H_0 . To this end, a random vector $(X, Y)^T$ following the distribution $\Phi_{(1,1),p,(0,0),I_2}$ was simulated $n \times N$ -times, and the value of the statistic T was calculated N -times based upon this sample. The choices of the values of n and p allows direct comparisons with Fig. 1.

Figure 5 shows the influence an increasing simulation sample size N has onto the accuracy of the estimation of the density of the test statistic if the null hypothesis is true. In the case $n = 30, p = 2$ and for four different values of the simulation sample size N , Table 1 presents the correspondingly calculated percentiles of orders 5 and 95, respectively, and the exact Fisher quantiles $F_{30,30,q} = F_{30,30,q}(2), q \in \{0.05, 0.95\}$.



The likelihood ratio test can be equivalently reformulated as to reject H_0 if, for a suitably chosen c , the likelihood ratio Q satisfies $Q < c$. Let

$$L(\sigma_1, \sigma_2) = \prod_{i=1}^n f(x_i, y_i | \sigma_1, \sigma_2)$$

and $\tilde{L}(\sigma) = L(\sigma, \sigma)$, and denote unrestricted and restricted under H_0 mle's of σ_1, σ_2 and $\sigma_1 = \sigma_2 = \sigma$, say by $\hat{\sigma}_1, \hat{\sigma}_2$ and $\tilde{\sigma}$, respectively. The likelihood ratio statistic $Q = \tilde{L}(\tilde{\sigma}) / L(\hat{\sigma}_1, \hat{\sigma}_2)$ allows the representations

$$\frac{Q^{p/n}}{4} = \frac{\Sigma_{p,X}^p \Sigma_{p,Y}^p}{\left(\Sigma_{p,X}^p + \Sigma_{p,Y}^p\right)^2} = \frac{|X_{(n)} - \mu_1 1_n|_p^p |Y_{(n)} - \mu_2 1_n|_p^p}{(|X_{(n)} - \mu_1 1_n|_p^p + |Y_{(n)} - \mu_2 1_n|_p^p)^2}.$$

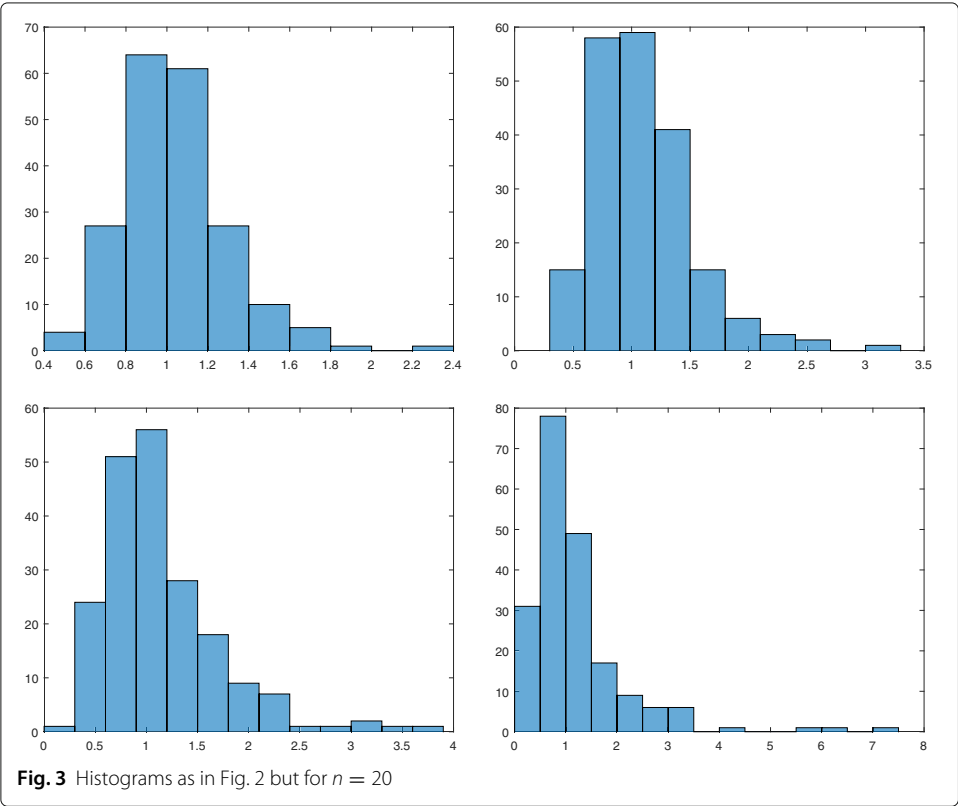
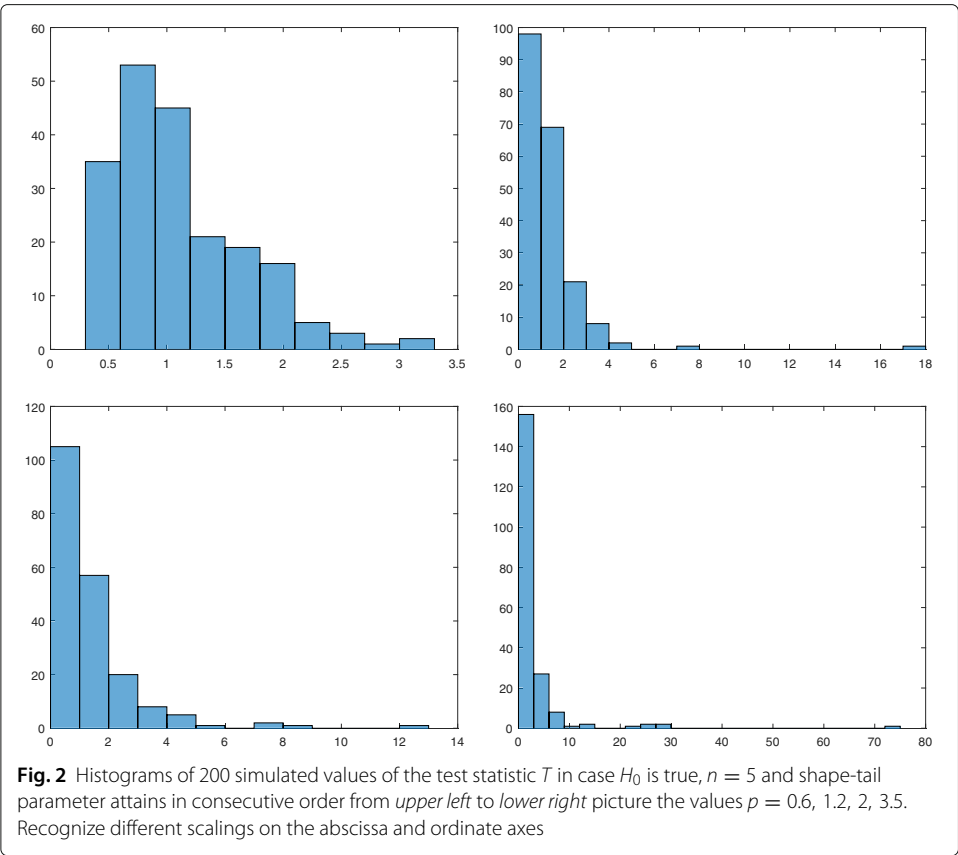
According to the general geometric measure-theoretical methodology of investigation in Richter (2014) and papers referred to there, the restricted distribution function of T if H_0 is true is

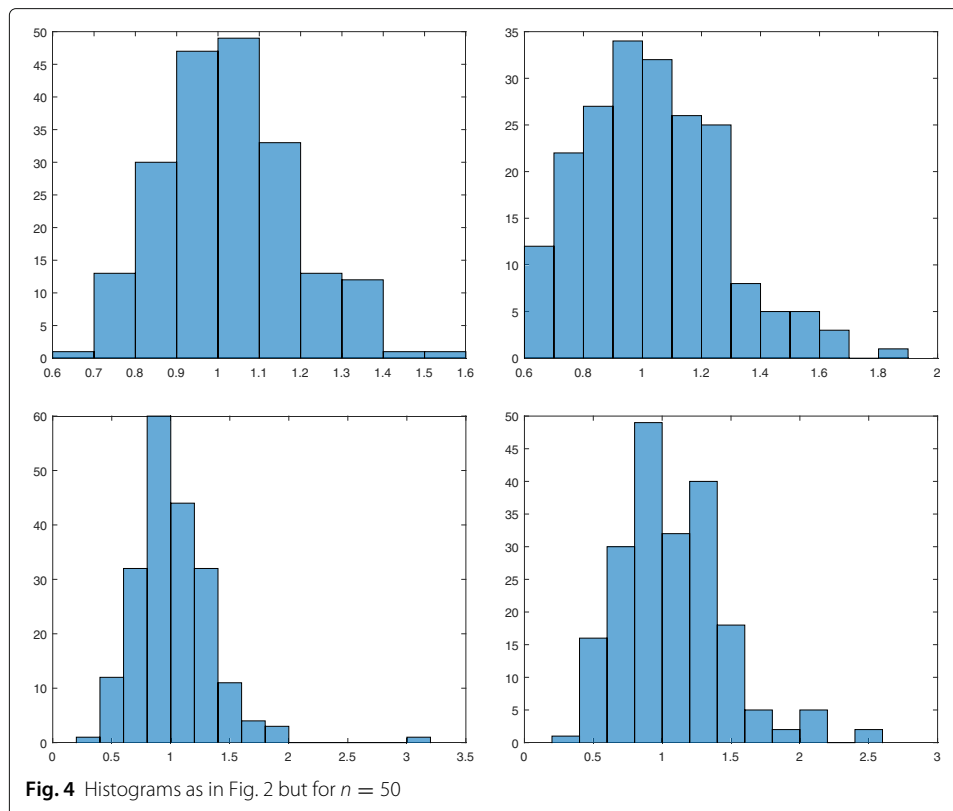
$$P(T < t) |_{H_0} = P\left(\left(X_{(n)}^T Y_{(n)}^T\right)^T \in C_{n,p}(t)\right) |_{H_0}, \quad t \in R$$

where

$$C_{n,p}(t) = \left\{ (x^T y^T)^T \in R^n \times R^n : \frac{|x|_p^p}{|y|_p^p} < t \right\}$$

is a cone with vertex in $0 \in R^{2n}$ and $|z|_p = |z|_{(1,1),p}$. A geometric measure representation of the standardized p -power exponential law applies to show that this distribution is the p -generalized Fisher or $F_{n,n}(p)$ -distribution. As mentioned before, this method may





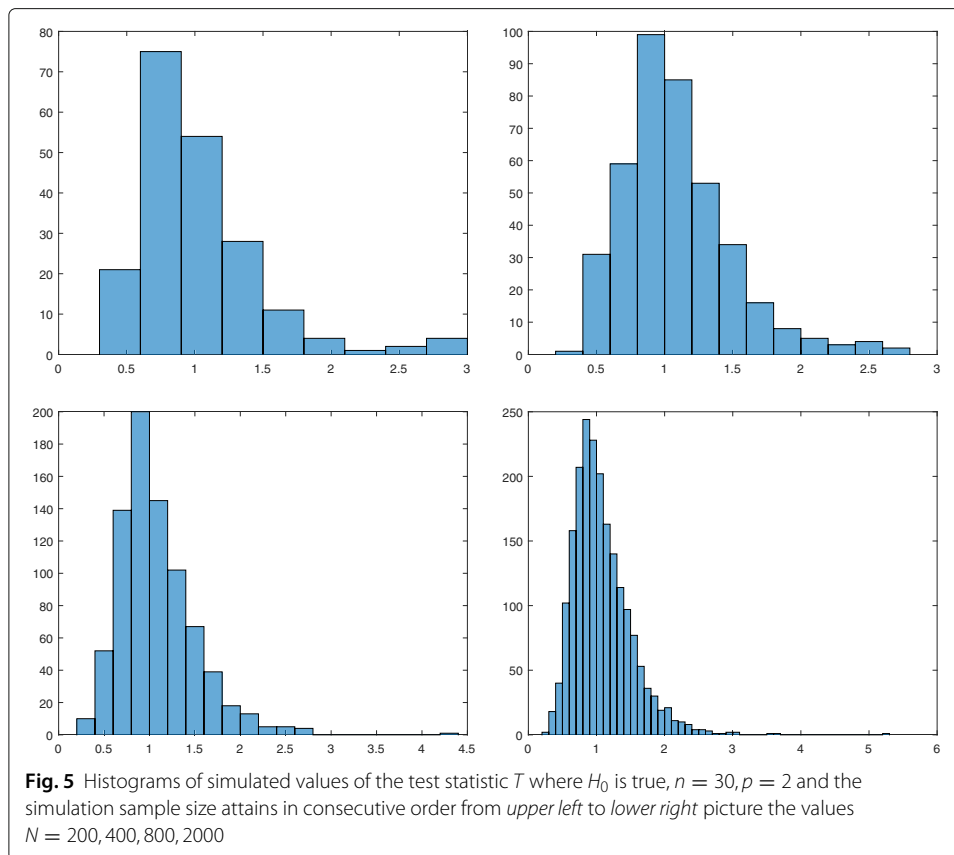
also be used to derive exact distributions of other statistics. Differently from what was considered so far, throughout the following section the vectors $(X_i, Y_i)^T$ are allowed to be rotated through $(\mu_1, \mu_2)^T$, in other words, the variables X_i, Y_i are allowed to be rotational dependent or correlation type dependent, $i = 1, 2, \dots$.

4 Tests for equal scaling parameters in correlational dependent p -generalized elliptically contoured distributions

This section is aimed to generalize the results presented in Section 3 for the case that two random variables may be rotational or correlation type dependent. To this end, we start in Section 4.1 with a p -generalization of the (a, b, α) -representation of the Gaussian law. Section 4.2 is aimed to give a geometric explanation of correlation in the particular case of a two-dimensional Gaussian distribution. Roughly spoken, correlation is interpreted by rotation under heteroscedasticity. Section 4.3 presents a test for checking homoscedasticity of principal components.

4.1 The geometric $((a, b), p, \alpha)$ -parametrization

Let a random vector follow a rotational dependent p -generalized elliptically contoured p -power exponential distribution, $(X, Y)^T \sim \Phi_{(a,b),p,(\mu_1,\mu_2),D(\alpha)}$ where a, b, p are positive parameters, $(\mu_1, \mu_2)^T \in \mathbb{R}^2$ and $D(\alpha) = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$ with $0 \leq \alpha < 2\pi$ is an orthogonal matrix causing a clockwise rotation around the origin through an angle of size α . Such random vector has according to Richter (2014) and Richter (2015a), (33), the density



$$f_{(X,Y)}(x,y) = \frac{C_p^2}{ab} \exp \left\{ -\frac{1}{p} \left| D(\alpha) \begin{pmatrix} x - \mu_1 \\ y - \mu_2 \end{pmatrix} \right|_{(a,b),p}^p \right\}$$

where the functional

$$|z|_{(a,b),p} = (|\frac{z_1}{a}|^p + |\frac{z_2}{b}|^p)^{1/p}, (x,y)^T \in \mathbb{R}^2$$

is a norm if $p \geq 1$ and an antinorm if $0 < p \leq 1$. For the latter notion, see Moszyńska and Richter (2012). The level sets of the density $f_{(X,Y)}$ are p -generalized ellipses being not necessarily axes-aligned but centered at the point $(\mu_1, \mu_2)^T$.

Moreover, the stochastic representation

$$D(\alpha) \begin{pmatrix} X - \mu_1 \\ Y - \mu_2 \end{pmatrix} \stackrel{d}{=} RU$$

holds true where $R \geq 0$ and $U \sim \omega_{E(a,b),p}$ are independent, R follows the density

$$f_R(r) = re^{-\frac{r^p}{p}} / \int_0^\infty re^{-\frac{r^p}{p}} dr, r > 0$$

Table 1 Simulating quantiles $F_{30,30,q}(2), q = 0.05, q = 0.95$

Simulation sample size N :	200	400	800	2000	$F_{30,30,q}(2)$
5-percentile	0.541	0.523	0.546	0.545	0.543
95-percentile	1.933	1.984	1.864	1.858	1.841=1/0.543

and U follows the $E_{(a,b),p}$ -star-generalized uniform distribution on the Borel σ -field of the p -generalized axes-aligned ellipse having main axes of lengths $2a, 2b$,

$$E_{(a,b),p} = \left\{ (x, y)^T \in \mathbb{R}^2 : \left| (x, y)^T \right|_{(a,b),p} = 1 \right\}.$$

Thus,

$$P(U \in A) = \omega_{E_{(a,b),p}}(A) = \mathfrak{U}(A)/\mathfrak{U}(E_{(a,b),p}), \quad A \in \mathfrak{B}^2 \cap E_{(a,b),p}$$

where \mathfrak{U} denotes the $E_{(a,b),p}$ -generalized arc-length measure.

Let us finally remark that another definition of a bivariate p -generalized error density is given in Taguchi (1978).

4.2 Geometry of variance homogeneity

In this section, we exploit the fact that under heteroscedasticity a rotation causes a particular type of dependence, and give a new geometric interpretation of the hypothesis of variance homogeneity. To this end, we restrict our consideration once again to the Gaussian case. Let us assume that $(X, Y)^T$ is an anti-clockwise rotated axes-aligned Gaussian vector

$$\begin{pmatrix} X \\ Y \end{pmatrix} = D^T(\alpha) \begin{pmatrix} \xi \\ \eta \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Sigma \right), \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \varrho \sigma_1 \sigma_2 \\ \varrho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}$$

where

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} \sim N \left(\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix} \right).$$

Then,

$$\begin{aligned} f_{(X,Y)}(x, y) &= f_{(\xi,\eta)}(D(\alpha)(x, y)^T) \\ &= \frac{1}{2\pi ab} \exp \left\{ -\frac{1}{2} \begin{pmatrix} x - \mu_1 \\ y - \mu_2 \end{pmatrix}^T D(\alpha)^T \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix}^{-1} D(\alpha) \begin{pmatrix} x - \mu_1 \\ y - \mu_2 \end{pmatrix} \right\} \\ &= \frac{1}{2\pi ab} \exp \left\{ -\frac{1}{2} \left(\frac{(x - \mu_1) \cos \alpha + (y - \mu_2) \sin \alpha}{a} \right)^2 \right. \\ &\quad \left. - \frac{1}{2} \left(\frac{-(x - \mu_1) \sin \alpha + (y - \mu_2) \cos \alpha}{b} \right)^2 \right\}. \end{aligned}$$

According to Dietrich et al. (2013), one can represent the parameters $\sigma_1, \sigma_2, \varrho$ in terms of the parameters a, b, α as follows:

$$\sigma_i = \sigma_i(\alpha) = [V((\xi, \eta) \theta_i)]^{1/2} > 0, \quad i = 1, 2$$

and

$$\varrho = \varrho(\alpha) = \frac{\varrho_*}{2} \sin(2\alpha) \left(\frac{\sigma_1(\alpha)}{\sigma_2(\alpha)} + \frac{\sigma_2(\alpha)}{\sigma_1(\alpha)} \right)$$

where

$$\theta_1 = \theta_1(\alpha) = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}, \quad \theta_2 = \theta_2(\alpha) = \begin{pmatrix} \cos(\alpha + \frac{\pi}{2}) \\ \sin(\alpha + \frac{\pi}{2}) \end{pmatrix}$$

and

$$\varrho_* = \frac{a^2 - b^2}{a^2 + b^2}.$$

Let us try now to geometrically understand the hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ that is often supposed to hold in the statistical literature. It follows from the representations

$$\sigma_1^2 = a^2 \cos^2 \alpha + b^2 \sin^2 \alpha \quad \text{and} \quad \sigma_2^2 = a^2 \sin^2 \alpha + b^2 \cos^2 \alpha$$

that hypothesis H_0 means that

- either $\alpha \in \left\{ \frac{\pi}{4}, \frac{3\pi}{4} \right\}$ with arbitrary a, b
- or $\alpha \notin \left\{ \frac{\pi}{4}, \frac{3\pi}{4} \right\}$ and $a = b$.

If $a = b$ then $\varrho = 0$ and $\Sigma = \sigma^2 I_{2n}$ thus the density level sets of $(\xi, \eta)^T$ and $(X, Y)^T$ are Euclidean circles. If $\alpha \in \left\{ \frac{\pi}{4}, \frac{3\pi}{4} \right\}$ then these level sets are arbitrary ellipses with main axes belonging to the lines $\{(x, y)^T \in \mathbb{R}^2 : y = x\}$ and $\{(x, y)^T \in \mathbb{R}^2 : y = -x\}$, and the correlation attains any value from the interval $(-1, 1)$. Thus, H_0 is not focussing, or is wavering, with respect to the parameters a, b, α and the shape of the density level ellipses and might therefore not always being primarily of interest, from this geometric point of view. If one presumes just the hypothesis

$$H_0^* : a = b$$

then the sample distribution and with it the distributions of all statistics derived from this sample vector are the same as in the axes-aligned and homoscedastic case.

Let us finally consider the following well known particular case of homoscedasticity. If $(X, Y)^T \sim \Phi_{\mu, \Sigma}$ then the random vector

$$(\xi, \eta)^T = D\left(\frac{\pi}{4}\right)(X, Y)^T = \frac{1}{2}(X + Y, X - Y)^T = \frac{1}{2}(U, V)^T$$

has the covariance matrix

$$\Psi = D\left(\frac{\pi}{4}\right) \Sigma D^T\left(\frac{\pi}{4}\right) = \frac{1}{2} \begin{pmatrix} \sigma_1^2 + \sigma_2^2 + 2\varrho\sigma_1\sigma_2 & \sigma_2^2 - \sigma_1^2 \\ \sigma_2^2 - \sigma_1^2 & \sigma_1^2 + \sigma_2^2 - 2\varrho\sigma_1\sigma_2 \end{pmatrix}.$$

$$\text{Thus, if } \sigma_1 = \sigma_2 = \sigma, \text{ say, then } \Psi = 2\sigma^2 \begin{pmatrix} 1 + \varrho & 0 \\ 0 & 1 - \varrho \end{pmatrix}.$$

4.3 Testing homoscedasticity of principal components

We are well motivated now for testing equality of scaling parameters in the $((a, b), p, \alpha)$ -parameterized model by checking the hypothesis $H_0 : a = b$ vs. the alternative $H_A : a \neq b$. Throughout this section, let

$$\hat{\alpha} = \begin{cases} \alpha & \text{if } \alpha \text{ is known} \\ \text{mle}(\alpha) & \text{if } \alpha \text{ is unknown.} \end{cases}$$

Let us further be given a concrete sample $\mathbf{x}_i = (x_i, y_i)^T, i = 1, \dots, n$ from independent identically and according to $\Phi_{(a,b),p,(0,0),D(\alpha)}$ distributed random vectors. Then, the likelihood function $L(a, b, \alpha)$ satisfies the equation

$$\begin{aligned} & (ab)^n L(a, b, \alpha) / C_p^{2n} \\ &= \exp \left\{ -\frac{1}{pa^p} \sum_1^n |x_i \cos \alpha + y_i \sin \alpha|^p - \frac{1}{pb^p} \sum_1^n |-x_i \sin \alpha + y_i \cos \alpha|^p \right\}. \end{aligned}$$

Let us consider the first two of the three likelihood equations. The partial derivatives of $\ln L$ with respect to a and b attain the value zero if $a = \hat{a}(\hat{\alpha})$ and $b = \hat{b}(\hat{\alpha})$, respectively,

where

$$\widehat{a}(\alpha)^p = \frac{1}{n} \sum_{i=1}^n |x_i \cos \alpha + y_i \sin \alpha|^p \text{ and } \widehat{b}(\alpha)^p = \frac{1}{n} \sum_{i=1}^n |y_i \cos \alpha - x_i \sin \alpha|^p,$$

what ever the value of α is. The resulting equation

$$L(\hat{a}, \hat{b}, \hat{\alpha}) = \frac{C_p^{2n} e^{-2n/p}}{(\hat{a}(\hat{\alpha}) \hat{b}(\hat{\alpha}))^n}$$

will be used later for constructing the likelihood ratio statistic. An angle $\hat{\alpha}$ solves the third likelihood equation if

$$\frac{\sum_{i=1}^n |\mathbf{x}_i^T \theta_1(\hat{\alpha})|^{p-2} (\mathbf{x}_i^T \theta_1(\hat{\alpha})) (\mathbf{x}_i^T \theta_2(\hat{\alpha}))}{\sum_{i=1}^n |\mathbf{x}_i^T \theta_1(\hat{\alpha})|^p} - \frac{\sum_{i=1}^n |\mathbf{x}_i^T \theta_2(\hat{\alpha})|^{p-2} (\mathbf{x}_i^T \theta_2(\hat{\alpha})) (\mathbf{x}_i^T \theta_1(\hat{\alpha}))}{\sum_{i=1}^n |\mathbf{x}_i^T \theta_2(\hat{\alpha})|^p} = 0.$$

An n -dimensional vector-algebraic reformulation of this equation is

$$\frac{\left[\begin{pmatrix} \mathbf{x}_1^T \theta_2(\hat{\alpha}) \\ \vdots \\ \mathbf{x}_n^T \theta_2(\hat{\alpha}) \end{pmatrix}, \begin{pmatrix} \mathbf{x}_1^T \theta_1(\hat{\alpha}) \\ \vdots \\ \mathbf{x}_n^T \theta_1(\hat{\alpha}) \end{pmatrix} \right]_p}{\left| \begin{pmatrix} \mathbf{x}_1^T \theta_1(\hat{\alpha}) \\ \vdots \\ \mathbf{x}_n^T \theta_1(\hat{\alpha}) \end{pmatrix} \right|_p^2} = \frac{\left[\begin{pmatrix} \mathbf{x}_1^T \theta_1(\hat{\alpha}) \\ \vdots \\ \mathbf{x}_n^T \theta_1(\hat{\alpha}) \end{pmatrix}, \begin{pmatrix} \mathbf{x}_1^T \theta_2(\hat{\alpha}) \\ \vdots \\ \mathbf{x}_n^T \theta_2(\hat{\alpha}) \end{pmatrix} \right]_p}{\left| \begin{pmatrix} \mathbf{x}_1^T \theta_2(\hat{\alpha}) \\ \vdots \\ \mathbf{x}_n^T \theta_2(\hat{\alpha}) \end{pmatrix} \right|_p^2}$$

where $[\cdot, \cdot]_p$ denotes a semi-inner-product defined by

$$[x, y]_p = \frac{\sum_{i=1}^n x_i y_i |y_i|^{p-2}}{|y|_p^{p-2}}, x, y \in \mathbb{R}^n.$$

For the theory and applications of semi-inner products we refer to Lumer (1961), Giles (1967), Dragomir (2004) and Horváth et al. (2015). We just mention here that, for all x, y, z from \mathbb{R}^n , $a \in \mathbb{R}$,

$$[x + z, y]_p = [x, y]_p + [z, y]_p, [ax, y]_p = a[x, y]_p, [x, ay]_p = a[x, y]_p,$$

$$[x, x]_p \geq 0, [x, x]_p = 0 \text{ iff } x = 0, \text{ and } [x, y]_p \leq [x, x]_p^{1/2} [y, y]_p^{1/2}.$$

In general, a semi-inner product is not symmetric and non-linear in the second argument. With the notations $\xi_i(\alpha) = (\mathbf{x}_1^T \theta_i(\alpha), \dots, \mathbf{x}_n^T \theta_i(\alpha))^T$, $i = 1, 2$ and $x_{(n)} = (x_1, \dots, x_n)$, $y_{(n)} = (y_1, \dots, y_n)$, $\hat{\alpha}$ solves the equation

$$\frac{[\xi_2(\hat{\alpha}), \xi_1(\hat{\alpha})]_p}{|\xi_1(\hat{\alpha})|_p^2} = \frac{[\xi_1(\hat{\alpha}), \xi_2(\hat{\alpha})]_p}{|\xi_2(\hat{\alpha})|_p^2}$$

or

$$\begin{aligned} & \frac{[(\cos \hat{\alpha}) y_{(n)} - (\sin \hat{\alpha}) x_{(n)}, (\cos \hat{\alpha}) x_{(n)} + (\sin \hat{\alpha}) y_{(n)}]_p}{|(\cos \hat{\alpha}) x_{(n)} + (\sin \hat{\alpha}) y_{(n)}|_p^2} \\ &= \frac{[(\cos \hat{\alpha}) x_{(n)} + (\sin \hat{\alpha}) y_{(n)}, (\cos \hat{\alpha}) y_{(n)} - (\sin \hat{\alpha}) x_{(n)}]_p}{|(\cos \hat{\alpha}) y_{(n)} - (\sin \hat{\alpha}) x_{(n)}|_p^2}. \end{aligned} \quad (2)$$

The Hessian matrix of $\ln L(a, b, \alpha)$ at the critical point $(a, b, \alpha) = (\hat{a}, \hat{b}, \hat{\alpha})$ is

$$HM = \begin{pmatrix} -\frac{np}{\hat{a}^2(\hat{\alpha})} & 0 & \epsilon_1 \\ 0 & -\frac{np}{\hat{b}^2(\hat{\alpha})} & \epsilon_2 \\ \epsilon_1 & \epsilon_2 & \epsilon_3 \end{pmatrix}$$

with

$$\epsilon_1 = \epsilon_1(\hat{\alpha}) = pn^{1+1/p} \frac{[\xi_2(\hat{\alpha}), \xi_1(\hat{\alpha})]_p}{|\xi_1(\hat{\alpha})|_p^3}, \epsilon_2 = \epsilon_2(\hat{\alpha}) = -pn^{1+1/p} \frac{[\xi_1(\hat{\alpha}), \xi_2(\hat{\alpha})]_p}{|\xi_2(\hat{\alpha})|_p^3}$$

and

$$\epsilon_3 = \epsilon_3(\hat{\alpha}) = 2n + n(1-p) \frac{<\xi_2(\hat{\alpha}), \xi_1(\hat{\alpha})>_p}{|\xi_1(\hat{\alpha})|_p^p} + \frac{<\xi_1(\hat{\alpha}), \xi_2(\hat{\alpha})>_p}{|\xi_2(\hat{\alpha})|_p^p}$$

where

$$<\eta, v>_p = \frac{\sum_{i=1}^n |v_i|^{p-2} \eta_i^2}{|v|_p^p}, \quad \eta, v \text{ from } R^n.$$

Obviously, $\Delta_1 < 0$ and $\Delta_2 > 0$ where

$$\Delta_1 = \left(-\frac{np}{\hat{a}^2(\hat{\alpha})}\right) \text{ and } \Delta_2 = \det \begin{pmatrix} -\frac{np}{\hat{a}^2(\hat{\alpha})} & 0 \\ 0 & -\frac{np}{\hat{b}^2(\hat{\alpha})} \end{pmatrix}.$$

Let $\Delta_3 = \det(HM)$. If $\Delta_3 < 0$ then $L(a, b, \alpha)$ attains a local maximum at the point $(a, b, \alpha) = (\hat{a}, \hat{b}, \hat{\alpha})$, see, e.g., Arens et al. (2013), Section 24.6. Under this assumption, $(\hat{a}, \hat{b}, \hat{\alpha}) = \text{mle}(a, b, \alpha)$. Note that $\Delta_3 < 0$ if and only if

$$np\epsilon_3(\hat{\alpha}) + a^2(\hat{\alpha})\epsilon_1^2(\hat{\alpha}) + b^2(\hat{\alpha})\epsilon_2^2(\hat{\alpha}) < 0. \quad (3)$$

Thus, for finding $\text{mle}(a, b, \alpha)$, one has to solve (2) under the constraint (3).

If $p = 2$ then the semi-inner product $[\cdot, \cdot]_p$ is symmetric, thus $\hat{\alpha}$ satisfies either the equation $[\xi_1(\hat{\alpha}), \xi_2(\hat{\alpha})]_2 = 0$, or $|\xi_1(\hat{\alpha})|_2^2 = |\xi_2(\hat{\alpha})|_2^2$. The first and second equations mean that

$$\hat{\alpha} = \frac{1}{2} \arctan \frac{2 \cos \angle(x_{(n)}, y_{(n)})}{\frac{|y_{(n)}|_2}{|x_{(n)}|_2} - \frac{|x_{(n)}|_2}{|y_{(n)}|_2}} \text{ and } \hat{\alpha} = \frac{1}{2} \arctan \frac{\frac{|y_{(n)}|_2}{|x_{(n)}|_2} - \frac{|x_{(n)}|_2}{|y_{(n)}|_2}}{2 \cos \angle(x_{(n)}, y_{(n)})},$$

respectively, where $\angle(\xi, \eta)$ denotes the angle between the vectors ξ and η , and $\arctan(+(-)\infty) = +(-)\pi/2$.

We consider now the the H_0 -restricted likelihood function

$$\tilde{L}(a, \alpha) = L|_{H_0} = L(a, a, \alpha)$$

and put

$$\tilde{\alpha} = \begin{cases} \alpha & \text{if } \alpha \text{ is known} \\ \text{mle}(\alpha)|_{H_0} & \text{if } \alpha \text{ is unknown.} \end{cases}$$

The partial derivative of \tilde{L} with respect to a attains the value zero if $a = \tilde{a}(\alpha)$ where

$$\tilde{a}(\alpha)^p = \frac{1}{2} \left(\hat{a}(\alpha)^p + \hat{b}(\alpha)^p \right)$$

what ever the value of α is. Thus, for suitable choice of $\tilde{\alpha}$, the maximum value of the restricted likelihood function \tilde{L} can be represented as

$$\tilde{L}(\tilde{\alpha}, \tilde{\alpha}) = \frac{C_p^{2n} e^{-2n/p}}{(\tilde{\alpha}^2(\tilde{\alpha}))^n}.$$

An angle $\tilde{\alpha}$ solves the second restricted likelihood equation iff it satisfies the equation

$$[\xi_1(\tilde{\alpha}), \xi_2(\tilde{\alpha})]_p |\xi_2(\tilde{\alpha})|_p^{p-2} = [\xi_2(\tilde{\alpha}), \xi_1(\tilde{\alpha})]_p |\xi_1(\tilde{\alpha})|_p^{p-2}$$

which can be reformulated as

$$\begin{aligned} & [(\cos \tilde{\alpha})y_{(n)} - (\sin \tilde{\alpha})x_{(n)}, (\cos \tilde{\alpha})x_{(n)} + (\sin \tilde{\alpha})y_{(n)}]_p |(\cos \tilde{\alpha})x_{(n)} + (\sin \tilde{\alpha})y_{(n)}|_p^{p-2} \\ &= [(\cos \tilde{\alpha})x_{(n)} + (\sin \tilde{\alpha})y_{(n)}, (\cos \tilde{\alpha})y_{(n)} - (\sin \tilde{\alpha})x_{(n)}]_p |(\cos \tilde{\alpha})y_{(n)} - (\sin \tilde{\alpha})x_{(n)}|_p^{p-2}. \end{aligned}$$

If $p = 2$ then every $\tilde{\alpha} \in [0, \frac{\pi}{2})$ solves this equation. Our test statistic

$$Q = \frac{\tilde{L}(\tilde{\alpha}, \tilde{\alpha})}{L(\hat{a}, \hat{b}, \hat{\alpha})} = \frac{(\hat{a}(\hat{\alpha})\hat{b}(\hat{\alpha}))^n}{(\tilde{\alpha}(\tilde{\alpha}))^{2n}}$$

satisfies the representation

$$\frac{Q^{p/n}}{4} = \frac{\sum_{i=1}^n |x_i^T \theta_1(\hat{\alpha})|^p \sum_{i=1}^n |x_i^T \theta_2(\hat{\alpha})|^p}{\left(\sum_{i=1}^n |x_i^T \theta_1(\tilde{\alpha})|^p + \sum_{i=1}^n |x_i^T \theta_2(\tilde{\alpha})|^p \right)^2}.$$

The likelihood ratio decision rule means to reject H_0 if for some suitably chosen $t \in (0, 1)$ there holds $Q < t$.

We remark that the present statistic becomes the same as that in the axes-aligned p -generalized elliptically contoured case in Section 3 if $\hat{\alpha} = \tilde{\alpha} \in \{0, \frac{\pi}{2}\}$ and $\mu_1 = \mu_2 = 0$. Moreover, the present decision rule can be equivalently reformulated then as to reject H_0 if $T_p = \frac{|Y_{[1]}|_p^p}{|Y_{[2]}|_p^p}$ attains sufficiently small or large values where

$$Y_{[i]}^T = \theta_i(\alpha)^T \begin{pmatrix} X_1 & X_2 & \dots & X_n \\ Y_1 & Y_2 & \dots & Y_n \end{pmatrix}, \quad i = 1, 2.$$

Example 1 • If $\alpha = 0$ then $\theta_1(\alpha) = (1, 0)^T$ and $\theta_2(\alpha) = (0, 1)^T$.

• If α is known (then $\hat{\alpha} = \tilde{\alpha} = \alpha$) then the considered decision rule means in other words to reject H_0 for large values of $\sqrt{R_p} + 1/\sqrt{R_p}$ where

$$R_p = \frac{\sum_{i=1}^n |x_i^T \theta_1(\alpha)|^p}{\sum_{i=1}^n |x_i^T \theta_2(\alpha)|^p}.$$

Note that, for $i = 1, \dots, n$,

$$D(\alpha)x_i \sim \Phi_{(a,b),p,(0,0),I_2}$$

where I_2 denotes the 2×2 -unit matrix. The statistic R_p has therefore independently of the actual value of the angle of rotation α the same p -generalized Fisher distribution as the likelihood ratio statistic in Section 3. Thus, in this case, rotational dependence is without influence onto the null distribution of the likelihood ratio statistic for proving $H_0 : \sigma_1^2 = \sigma_2^2$, or not.

Remark Since the purpose is to test whether H_0 or not, when there is a correlation between two groups, one might like to consider testing the significance of correlation structure prior to testing H_0 or not. In the present situation where there is no rotational correlation, this would mean to test whether the shape-scale parameter satisfies $\tilde{H}_0 : p = 2$ or not. Searching the literature the author was not aware of a significance test for this hypothesis, see for example in González-Farías et al. (2009), Yu et al. (2012), Purczynski and Bednarsz-Okrzynska (2014) and Pascal et al. (2017).

5 The semi-inner product $[\cdot, \cdot]_p$ appears also in estimating location

Many authors were dealing with estimating parameters of the p -power exponential distribution. Without aiming completeness, and without going into any details, we refer to Stacy and Mihram (1965), Harter (1967), Rahman and Gokhale (1996), Varanasi and Aazhang (1989), Do and Vetterli (1988), Mineo and Ruggieri (2005), González-Farías et al. (2009), Saatci and Akan (2010).

It is well known from the Gauss-Markov theorem that orthogonal projections play a fundamental role in estimating parameters in the theory of linear models. The notion of an orthogonal projection is closely connected with that of a scalar product. If the standard Gaussian distribution is the sample distribution in R^n then it is natural to use the Euclidean norm for several statistical calculations. This norm is generated by the Euclidean scalar product in $R^n \times R^n$. If the density of the sample vector $X_{(n)}$ is, for some $p \geq 1$,

$$f_{X_{(n)}}(x) = \prod_{i=1}^n \frac{C_p}{\sigma} \exp \left\{ -\frac{|x_i - \mu|^p}{p\sigma^p} \right\}, x = (x_1, \dots, x_n)^T \in R^n$$

then it is natural to work with the norm $|\cdot|_p$ which is not generated by an inner product if $p \neq 2$,

$$f_{X_{(n)}}(x) = \left(\frac{C_p}{\sigma} \right)^n \exp \left\{ -\frac{|x - \mu 1_n|_p^p}{p\sigma^p} \right\}, x \in R^n, 1_n = (1, \dots, 1)^T \in R^n.$$

It is known, however, that this norm is generated by the semi-inner product $[\cdot, \cdot]_p$ considered in Section 4, $|x|_p = [x, x]_p^{1/2}$. The present section is aimed to verify that this semi-inner product plays also a role in estimating the location parameter of a p -power exponential distribution. Let $L(\mu) = f_{X_{(n)}}(x)$. Maximizing L with respect to μ is equivalent to minimizing the function

$$f(\mu) = |x - \mu 1_n|_p^p, \mu \in R.$$

Let $x_{(1)} \leq \dots \leq x_{(n)}$ be the ordered values of the concrete sample vector $x = (x_1, \dots, x_n)^T$. Given μ , there exists a natural number n_1 such that

$$f(\mu) = \sum_{i=1}^{n_1} (\mu - x_{(i)})^p + \sum_{i=n_1+1}^n (x_{(i)} - \mu)^p.$$

Thus, $f'(\mu) = p(f_1(\mu) - f_2(\mu))$ where

$$f_1(\mu) = \sum_{x_{(i)} < \mu} (\mu - x_{(i)})^{p-1} \text{ and } f_2(\mu) = \sum_{x_{(i)} > \mu} (x_{(i)} - \mu)^{p-1}.$$

If $p \in (1, \infty)$, the functions $f_i, i = 1, 2$ are monotonously increasing/decreasing if $i = 1/2$, respectively. Moreover, these functions are continuous, and satisfy $f_1(x_{(1)}) = 0, f_1(x_{(n)}) > 0$ and $f_2(x_{(n)}) = 0, f_2(x_{(1)}) > 0$. Thus there exists a uniquely

determined $\hat{\mu}$ such that $(\hat{\mu}, f_1(\hat{\mu})) = (\hat{\mu}, f_2(\hat{\mu}))$ is the intersection point of the curves $\{(\mu, f_1(\mu)) : \mu \in [x_{(1)}, x_{(n)}]\}$ and $\{(\mu, f_2(\mu)) : \mu \in [x_{(1)}, x_{(n)}]\}$. Based upon a bisection algorithm, $\hat{\mu}$ can be numerically calculated and is the solution of the equation $f'(\mu) = 0|_{\mu=\hat{\mu}}$, i.e.

$$\sum_{i=1}^n |x_i - \hat{\mu}|^{p-1} \text{sign}(\hat{\mu} - x_i) = 0. \quad (4)$$

Example 2 If $p = 2$ then (4) reads as

$$-\sum_{i=1}^{n_1} (x_{(i)} - \hat{\mu}) + \sum_{i=n_1+1}^n (x_{(i)} - \hat{\mu}) (-1) = 0,$$

thus $\hat{\mu} = \bar{x}_n$.

We consider now two cases excluded so far.

Example 3 In the case $p = 1$,

$$f'(\mu) = \sum_{i=1}^n \text{sign}(\mu - x_i) = \left(\sum_{x_{(i)} < \mu} 1 - \sum_{x_{(i)} > \mu} 1 \right),$$

thus $f'(\hat{\mu}) = 0$ iff $\sharp\{x_i < \hat{\mu}\} = \sharp\{x_i > \hat{\mu}\}$ where $\sharp\{\dots\}$ means the number the event written between the brackets occurs. For odd n , $\hat{\mu} = x_{[n/2]+1}$, and for even n , every $\hat{\mu}$ from $[x_{[n/2]}, x_{[n/2]+1}]$ satisfies this condition, thus $\hat{\mu}$ is the sample median.

Example 4 In the case $p = \infty$, we first define the notion $|\cdot|_\infty$.

(a) Let $f(\mu) = \max\{|x_1 - \mu|, \dots, |x_n - \mu|\}$ and i^* such that $f(\mu) = |x_{i^*} - \mu|$. Then

$$|x_{(n)} - \mu|_p = |x_{i^*} - \mu| \left(1 + \frac{\sum_{i \neq i^*} |x_i - \mu|^p}{|x_{i^*} - \mu|^p} \right)^{1/p}$$

where

$$\left(1 + \frac{\sum_{i \neq i^*} |x_i - \mu|^p}{|x_{i^*} - \mu|^p} \right)^{1/p} \geq 1 \text{ and } \left(1 + \frac{\sum_{i \neq i^*} |x_i - \mu|^p}{|x_{i^*} - \mu|^p} \right)^{1/p} \leq n^{1/p} \rightarrow 1, p \rightarrow \infty.$$

By definition,

$$|x_{(n)} - \mu|_\infty = \lim_{p \rightarrow \infty} |x_{(n)} - \mu|_p.$$

(b) If for some i^* there holds $|x_{i^*} - \hat{\mu}| = \max\{|x_1 - \hat{\mu}|, \dots, |x_n - \hat{\mu}|\}$ then $\hat{\mu}$ is maximum likelihood estimator of μ . The number $\hat{\sigma} = \max |x_i - \hat{\mu}|$ is the smallest number satisfying $-\hat{\sigma} \leq x_i - \hat{\mu} \leq \hat{\sigma}, \forall i$, thus $x_{(n)} - \hat{\sigma} \leq \hat{\mu} \leq x_{(1)} + \hat{\sigma}$. The smallest possible $\hat{\sigma}$ satisfies $2\hat{\sigma} \geq x_{(n)} - x_{(1)}$, thus $\hat{\sigma} = (x_{(n)} - x_{(1)})/2$. It follows that $\hat{\mu} = (x_{(n)} + x_{(1)})/2 = \text{midrange } \{x_1, \dots, x_n\}$.

With

$$f''(\hat{\mu}) = p(p-1) \sum_{i=1}^n |x_i - \hat{\mu}|^{p-2} > 0 \text{ if } p > 1,$$

it follows in the general setting that the uniquely determined solution $\hat{\mu}$ of the equation (4) is a relative maximum point of the likelihood function L , thus $\hat{\mu} = \text{mle}(\mu)$. This means that

$$\sum_{x_i < \hat{\mu}} |x_i - \hat{\mu}|^{p-2} (\hat{\mu} - x_i) = \sum_{x_i > \hat{\mu}} |x_i - \hat{\mu}|^{p-2} (x_i - \hat{\mu})$$

or, equivalently,

$$\hat{\mu} \sum_{i=1}^n |x_i - \hat{\mu}|^{p-2} = \sum_{i=1}^n x_i |x_i - \hat{\mu}|^{p-2}.$$

Thus, on the one hand, $\hat{\mu}$ solves the oscillating fixed point equation

$$\hat{\mu} = \frac{\sum_{i=1}^n x_i |x_i - \hat{\mu}|^{p-2}}{\sum_{i=1}^n |x_i - \hat{\mu}|^{p-2}}.$$

On the other hand, it follows that $\hat{\mu} = \text{mle}(\mu)$ satisfies the equation

$$0 = [1_n, x - \hat{\mu} 1_n]_p \quad (5)$$

which means that $\hat{\mu} = \bar{x}_n$ if $p = 2$. Under suitable assumptions upon the convergence of $\hat{\mu}$ and the limit $\mu^* = \lim_{n \rightarrow \infty} \hat{\mu}$, it follows

$$\lim_{n \rightarrow \infty} n^{-2/p} [1_n, x - \hat{\mu} 1_n]_p = \frac{E(X - \mu^*) |X - \mu^*|^{p-2}}{(E|X - \mu^*|^p)^{(p-2)/p}},$$

thus $\lim_{n \rightarrow \infty} n^{-2/p} [1_n, x - \hat{\mu} 1_n]_p = 0$ can be reformulated as

$$EX|X - \mu^*|^{p-2} - \mu^* E|X - \mu^*|^{p-2} = 0. \quad (6)$$

6 Simulation of star-shaped distributed random vectors

6.1 Preliminary remarks

It may be of interest to determine exact distributions of the statistics dealt with in Sections 2-5. To this end, one might use various analytical tools like, e.g., a geometric measure representation as a starting point of explicit analytical derivations.

As an alternative to such derivations, we present here simulation methods which allow to generate stochastic approximations of statistical distributions.

Let $(X_{1,j}, Y_{1,j})^T, \dots, (X_{n,j}, Y_{n,j})^T, j = 1, \dots, N$ be independent samples of independent random vectors following the rotational dependent p -generalized elliptically contoured density $f_{(X,Y)}$ defined in Section 4.1., and let further

$$T_j = T \left((X_{1,j}, Y_{1,j})^T, \dots, (X_{n,j}, Y_{n,j})^T \right), j = 1, \dots, N$$

be a sample of i.i.d. copies of a real valued statistic T . For sufficiently large N , the probability $P(T < t)$ can be stochastically approximated by the relative frequency $\frac{1}{N} \sum_{i=1}^N I_{(-\infty, t)}(T_i)$. To this end, we present an acceptance-rejection method for simulating random vectors $(X_{i,j}, Y_{i,j})^T$ in Section 6.2, and a generalized polar method in Section 6.3. This will be done even under the much more general assumption that $(X, Y)^T$ follows an arbitrary star-shaped distribution. This class includes that of p -generalized elliptically contoured distributions. For approaches to general distribution classes see

Fernández et al. (1995), Arnold et al. (2008), Kamiya et al. (2008), Sarabia and Gómez-Déniz (2008), Balkema and Nolde (2010). A geometric representation of star-shaped distributions is given in Richter (2014).

We refer to the letter paper for main notions and recall that $(X, Y)^T$ allows the stochastic representation $X \stackrel{d}{=} R \cdot U$ where R and U are stochastically independent, R is a non-negative random variable, and the singular random vector U follows the star-generalized uniform distribution ω_S on the Borel σ -field $\mathfrak{B}(S)$ of the star-sphere S being the topological boundary of a suitably defined star body K ,

$$\omega_S(A) = \frac{O_S(A)}{O_S(S)}, A \in \mathfrak{B}(S).$$

For the distribution considered in Section 4.1, K can be chosen as a rotated through the origin axes-aligned p -generalized ellipsoid, $K = D^T(\alpha)B_{(a_1, a_2), p}$, and O_S means the corresponding star-generalized surface content measure.

6.2 Dependent p -generalized acceptance-rejection method

General aspects of acceptance-rejection or simply rejection methods are studied in Kalke and Richter (2013) and applied there to the p -generalized rejecting polar method. Platonically generalized uniformly distributed and polyhedral star-shaped distributed random vectors are generated this way in Richter and Schicker (2014, 2016a) as well as Richter and Schicker 2016b, respectively. If the star-spheres are represented in a certain analytical way, Nolan (2016) aims to exploit the geometric measure representation approximatively in a sense, not yet explicitly defined. Here, we demonstrate how to generate in four steps star-shaped distributed vectors.

Step 1. To start with, let $C_i, i = 1, \dots, d$ be positive constants, $C_{(d)} = (C_1, \dots, C_d)^T$ and $0_{(d)} = (0, \dots, 0)^T \in \mathcal{R}^d$. We denote by

$$(0_{(d)}, C_{(d)}) = (0, C_1) \times \dots \times (0, C_d)$$

an axes-aligned d -dimensional rectangle and by $O \in \mathcal{R}^{d \times d}$ an orthogonal matrix. Using the further notation

$$O(0_{(d)}, C_{(d)}) = \{Ox : x \in (0_{(d)}, C_{(d)})\},$$

we assume that the random vectors $\xi_n, n = 1, 2, \dots$ follow the uniform distribution on $O(0_{(d)}, C_{(d)})$, and are independent. Because the vector $O^{-1}\xi_n$ follows the product measure of uniform distributions on univariate intervals, $U_{(0, C_1)} \times \dots \times U_{(0, C_d)}$, it can immediately be simulated.

Step 2. Let the acceptance region

$$A \in \mathcal{B}^{(d)} \cap [O(0_{(d)}, C_{(d)})]$$

be a star body having the origin as an interior point. According to Remark A.1 in Kalke and Richter (2013), the stopping time

$$\tau^A = \inf \{n \in \mathcal{N} : \xi_n \in A\}$$

is almost surely finite if $P(\xi_1 \in A) > 0$. The following lemma says that the stopping element

$$\xi_{\tau^A} = \sum_{n=1}^{\infty} I_{\{\tau^A=n\}} \xi_n$$

is uniformly distributed in the acceptance region A , $\xi_{\tau_A} \sim U_A$.

Lemma 1 The stopping element ξ_{τ^A} satisfies the equation

$$P(\xi_{\tau^A} \in M) = U_A(M) \text{ for all } M \in \mathcal{B}^{(d)} \cap [O(0_{(d)}, C_{(d)})].$$

Proof It follows from

$$\begin{aligned} P(\xi_{\tau^A} \in M) &= P\left(\bigcup_{n=1}^{\infty} \{\tau^A = n, \xi_n \in M\}\right) \\ &= P(\xi_1 \in M)[1 + P(\xi_1 \notin A) + P^2(\xi_1 \notin A) + \dots] \end{aligned}$$

that

$$P(\xi_{\tau^A} \in M) = P(\xi_1 \in M) \left[1 + \frac{P(\xi_1 \notin A)}{1 - P(\xi_1 \notin A)} \right]$$

□

Example 5 For simulating a random vector following a p -generalized elliptically contoured distribution law, put $A = OB_{a,p}$ with $B_{a,p} = \left\{ x \in \mathcal{R}^d : \sum_{i=1}^d \left| \frac{x_i}{a_i} \right|^p \leq 1 \right\}$, and $C_i \geq a_i > 0, i = 1, \dots, d, a = (a_1, \dots, a_d)^T$.

Example 6 For simulating norm or antinorm contoured distributed vectors one can chose the acceptance region $A = \{x \in \mathcal{R}^d : \|x\| \leq 1\}$, where $\|\cdot\|$ is an arbitrary norm or antinorm and the constants $C_i > 0$ are chosen such that $A \subset O[0_{(d)}, C_{(d)}]$.

Example 7 If $A = P$ is a star-shaped polyhedron having the origin as an interior point, one can check whether a point belongs to A using the various representations of the Minkowski functional of A given in Richter and Schicker (2016b).

Example 8 If A is as described in Nolan (2016), check the condition given there.

Step 3. It is well known that if A is a star-shaped subset of \mathcal{R}^d having the origin as an interior point then the Minkowski functional $h_A(x) = \inf \{\lambda > 0 : x \in \lambda A\}, x \in \mathcal{R}^d$ is well defined. A normalization of the stopping element based upon this functional is used in the following lemma.

Step 4. Lemma 2 The random element $X_{\partial A} = \xi_{\tau^A}/h_A(\xi_{\tau^A})$ follows the star-generalized uniform distribution on $S = \partial A = \{x \in \mathcal{R}^d : h_A(x) = 1\}$, $X_{\partial A} \sim \omega_S$ for short.

Proof Let $\tilde{M} \in \mathcal{B}^d \cap \partial A$ and $\text{sector}(\tilde{M}) = \{\lambda x : x \in \tilde{M}, 0 \leq \lambda \leq 1\}$. Then

$$P(X_{\partial A} \in \tilde{M}) = P(\xi_{\tau^A} \in \text{sector}(\tilde{M})) = U_A(\text{sector}(\tilde{M}))$$

□

Example 5, continued. The Minkowski functional of the set $B_{a,p} = O^{-1}A$ is $h_{O^{-1}A}(x) = |x|_{a,p} = \left(\sum_{i=1}^d |x_i|^p \right)^{1/p}$ and, with $S = \partial B_{a,p}$,

$$O_S(\tilde{M}) = \int_{\{(x_1, \dots, x_{d-1}) : (x_1, \dots, x_d)^T \in \tilde{M}\}} \frac{d(x_1 \dots x_{d-1})}{(1 - \sum_{i=1}^{d-1} |\frac{x_i}{a_i}|)^{1/p}}, \tilde{M} \in \mathcal{B}^d \cap E_{a,p}.$$

Example 6, continued. For arbitrary norm or antinorm $||\cdot||$, $h_A(x) = ||x||$, and for all $\tilde{M} \in \mathcal{B}^d \cap S$, $S = \{x : ||x|| = 1\}$,

$$O_S(\tilde{M}) = \int_{\{\vartheta = (x_1, \dots, x_{d-1}) : (\vartheta, x_d(\vartheta))^T \in \tilde{M}\}} h_{K^*}(N(\vartheta)) d\vartheta$$

where K^* is the unit ball of the norm $||\cdot||^*$ being dual to the norm $||\cdot||$ in the first case, and the antipolar set of $K = \{x : ||x|| \leq 1\}$ in the second one. Moreover, $N(\vartheta)$ is the outer/inner normal vector to S at the point $(\vartheta, x_d(\vartheta))$, $N(\vartheta) = (\text{grad } x_d(\vartheta), -1)^T$.

Remark 1 Let NonNegSim denote the set of all non-negative random variables for which there is known a simulation method. Extensive overviews of simulation algorithms for non-uniform random variables are given in Rubinstein (1981) and Devroye (1986). If $R \in \text{NonNegSim}$ is independent of $X_{\partial A}$ where $X_{\partial A}$ is a star-generalized uniformly on the star sphere ∂A distributed random vector then the random vector $RX_{\partial A}$ follows a star-shaped distribution centered at the origin, Φ_A say.

As to summarize, Steps 1-4 together constitute an acceptance-rejection algorithm for simulating random vectors following a star-shaped distribution law.

Example 9 In case of a distribution having a density generating function, g say, the cumulative distribution function of $R = R(g)$ is

$$F_{R(g)}(r) = \frac{1}{I(g)O_A(\partial A)} \int_0^r \rho^{d-1} g(\rho) d\rho \text{ where } I(g) = \int_0^\infty r^{d-1} g(r) dr,$$

thus $R(g)$ can be simulated accordingly. To this end, let U be uniformly distributed on $(0, 1)$, then $F_{R(g)}^{-1}(U) \stackrel{d}{=} R(g)$.

Remark 2 If a density generating function g satisfies the equation

$$O_A(\partial A)I(g) = 1$$

then it is called a density generator. Methods of estimating a density generator are described in Liebscher and Richter (2017).

6.3 Dependent p -generalized polar method

The classical polar method is due to Box and Muller (1958). If the acceptance rate of the algorithm described in the previous section is not large enough, or for some other reason, one might seek for a direct star-generalization of the polar method. We just mention here that there are different particular methods for directly generating the star-generalized uniform distribution on a star sphere. For the p -generalized polar method, e.g., such

method has been established in Kalke and Richter (2013) and applied in Richter (2015a). The independent coordinate representation of general two-dimensional norm contoured distributions which is the basis for a norm-generalization of the polar method is proved in Richter (2015b).

Below, we describe an algorithm for an (a, p, O) -generalization of the polar method of Box and Muller where $a = (a_1, \dots, a_d)^T$, $a_i > 0$, $i = 1, \dots, d$; $p > 0$ and O is an orthogonal $d \times d$ -matrix. To this end, we assume that a random vector $X = (X_1, \dots, X_d)^T$ follows an axes-aligned p -generalized elliptically contoured distribution having density generating function g and vector of scaling parameters a , $X \sim \Phi_{g, B_{a,p}}$. Let further $Y = OX + v$ denote a transformation vector, with an orthogonal $d \times d$ -matrix O and $v \in \mathbb{R}^d$, and put

$$h_0(r) = C_0 r^{d-1} g(r) I_{(0, \infty)}(r)$$

and, for $\phi_i \in [0, \pi)$, $i = 1, \dots, d-2$, $\phi_{d-1} \in [0, 2\pi)$,

$$h_i(\phi_i) = C_i \frac{(\sin_{(a_i, a_{i+1}; p)}(\phi_i))^{d-i-1}}{N_{(a_i, a_{i+1}; p)}^2(\phi_i)}$$

where the generalized trigonometric functions $\sin_{(a_i, a_{i+1}; p)}(\phi_i)$, $\cos_{(a_i, a_{i+1}; p)}(\phi_i)$ and the normalizing functions $N_{(a_i, a_{i+1}; p)}(\phi_i)$ are defined in Richter (2014). With suitably chosen constants C_i , the functions h_i are the densities of independent random variables $R, \Phi_1, \dots, \Phi_{d-1}$ jointly satisfying the stochastic representation $X = RU$ where

$$\begin{aligned} U_1 &= a_1 \cos_{(a_1, a_2; p)}(\Phi_1), U_2 = a_2 \sin_{(a_1, a_2; p)}(\Phi_1) \cos_{(a_2, a_3; p)}(\Phi_2), \dots, \\ U_d &= a_d \sin_{(a_1, a_2; p)}(\Phi_1) \dots \sin_{(a_{d-2}, a_{d-1}; p)}(\Phi_{d-2}) \sin_{(a_{d-2}, a_{d-1}; p)}(\Phi_{d-1}), \end{aligned} \quad (7)$$

cf. Definition 4 in the same paper.

Step 1 Start the algorithm by generating a non-negative random number R according to the density h_0 .

Step 2 Generate random numbers $\Phi_1, \dots, \Phi_{d-2}$ from $[0, \pi)$ and Φ_{d-1} from $[0, 2\pi)$ following the densities h_1, \dots, h_{d-2} and h_{d-1} , respectively.

Step 3 Carry out transformation (7).

Step 4 Return $Y = OR(U_1, \dots, U_d)^T + v$.

This algorithm generates a random vector Y following the p -generalized elliptically contoured distribution law $\Phi_{g, a, p, v, O}$, see Theorem 4 and Remark 11 in Richter (2014). The particular case $d = 2$, $a_1 = a_2 = 1$ has been dealt with in Kalke and Richter (2013). Finally, we notice that $X \sim \Phi_{g, a, p, 0_d, I_d} = \Phi_{g, B_{a,p}, 0_d}$.

7 Discussion

Comparing mvv-c with pc-v-r models led to some new aspects in testing equality of variances or scaling parameters. Effects of rotational dependence are outlined. A new geometric interpretation of certain likelihood equations is given in terms of a semi-inner product. Based upon the present results for the more specific models dealt with in Sections 2-5, it could be of some interest to re-consider in the future the more general model in Wilcox (2015) and to possibly draw some new conclusions for this model. Our results might further stimulate a comparison of simulation methods, e.g. for particular cases being in the intersection of the work in Nolan (2016) and in Richter and Schicker

(2016a, b). To this end, one would particularly have to determine the Minkowski functionals of the sets considered in Nolan (2016) and then to compare the approximative simulation method there with the exact method presented in Richter and Schicker (2016a,b). Challenging questions are opened for deriving new exact statistical distributions, e.g. of Σ_X/Σ_Y , from dependent sample distributions, and to compare these results with corresponding simulation results. As another open problem it remains to combine rotational and l_p -dependence. Consequences the latter notion has for the derivation of exact distributions of certain statistics have been studied in Müller and Richter (2015, 2016a,b). There, the effects caused by the deviation of a density generating function from that of the p -power exponential law are studied in various situations.

Acknowledgements

The author is grateful to the Associate Editor for his valuable comments leading to an improvement of the paper.

Competing interests

The author declares that he has no competing interest.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 27 December 2016 Accepted: 4 August 2017

Published online: 20 September 2017

References

- Arens, T, Hettlich, F, Karpfinger, C, Kockelhorn, U, Lichtenegger, K, Stachel, H: Mathematik, Spektrum, Heidelberg (2013)
- Arnold, BC, Castillo, E, Sarabia, JM: Multivariate distributions defined in terms of contours. *J. Stat. Plann. Inference.* **138**(12), 4158–4171 (2008). doi:10.1016/j.jspi.2008.03.033
- Balkema, G, Nolde, N: Asymptotic independence for unimodal densities. *Adv. Appl. Probab.* **42**(2), 411–432 (2010). doi:10.1239/aap/1275055236
- Box, GEP, Muller, ME: A note on the generation of random normal deviates. *Ann. Math. Statist.* **29**(2), 610–611 (1958)
- Dang, UJ, Browne, RP, McNicholas, PD: Mixtures of multivariate power exponential distributions. *Biometrics.* **71**(4), 1081–1089 (2015)
- Devroye, L: Non-Uniform Random Variate Generation. Springer, Berlin (1986)
- Dietrich, T, Kalke, S, Richter, W-D: Stochastic representations and a geometric parametrization of the two-dimensional Gaussian law. *Chil J Stat.* **4**(2), 27–59 (2013)
- Do, MN, Vetterli, M: Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE Trans. Image. Process.* **11**(2), 146–158 (1988). doi:10.1017/S0266466600013384
- Dragomir, S: Semi-inner products and applications. 6th edn. Nova Science Publishers, Hauppauge, New York (2004)
- Fang, KT, Kotz, S, Ng, KW: Symmetric Multivariate And Related Distributions. Chapman and Hall, London etc (1990)
- Fernández, C, Osiewalski, J, Steel, MF: Modeling and inference with ν -spherical distributions. *J. Am. Stat. Assoc.* **90**(432), 1331–1340 (1995). doi:10.2307/2291523
- Games, PA, Winkler, HB, Probert, DA: Robust tests for homogeneity of variance. *Educ. Psychol. Meas.* **32**(4), 887–909 (1972)
- Giles, J: Classes of semi-inner product spaces. *Trans. Amer. Math. Soc.* **129**(6), 436–446 (1967). doi:10.1080/03610920600672120
- Gómez, E, Gómez-Villegas, MA, Marín, JM: A multivariate generalization of the power exponential family of distributions. *Commun. Stat. Theory Methods.* **27**, 589–600 (1998). doi:10.1080/03610929808832115
- Gómez-Villegas, MA, Gómez-Sánchez-Manzano, E, Man, P, Navarro, H: The effect of non-normality in the power exponential distributions. In: Pardo, L, Balakrishnan, N, Gil, MA (eds.), pp. 119–129. Berlin: Springer-Verlag, 972-985, Taylor & Francis, Philadelphia, PA, (2011). doi:10.1080/03610920701762754
- González-Farías, G, Domínguez-Molina, A, Rodríguez-Dagnino, RM: Efficiency of the approximated shape parameter estimator in the generalized Gaussian distribution. *IEEE Trans. vehic. techn.* **58**(8), 4214–4223 (2009). doi:10.1016/j.jspi.2003.09.008
- Harter, L: Maximum-likelihood estimation of the parameters of a four-parameter generalized gamma population from complete and censored samples. *Technometrics.* **9**(1), 159–165 (1967). doi:10.1016/j.jspi.2003.09.008
- Horváth, AG, Lángi, Z, Spirova, M: Semi-inner products and the concept of semi-polarity. **126**(2), 16 (2015). arXiv:13080974v4. doi:10.1016/j.jspi.2003.09.008
- Kalke, S, Richter, W-D: Simulation of the p-generalized Gaussian distribution. *J. Statist. Comput. Simulation.* **83**(4), 639–665 (2013). doi:10.1080/00949655.2011.631187
- Kamiya, H, Takemura, A, Kuriki, S: Star-shaped distributions and their generalizations. *J. Stat. Plann. Inference.* **138**(11), 3429–3447 (2008). doi:10.1016/j.jspi.2006.03.016
- Kuwana, Y, Kariya, T: LBI tests for multivariate normality in exponential power distributions. *J. Multivariate. Anal.* **39**, 117–134 (1991)
- Levy, KJ: A procedure for testing the equality of p correlated variances. *Br. J. Math. Stat. Psychol.* **29** (1976). doi:10.1111/j.2044-8317.1976.tb00705.x, <http://dx.doi.org/10.1111/j.2044-8317.1976.tb00705.x>
- Liebscher, E, Richter, W-D: Estimation of star-shaped distributions. *Risks* **xx**(x) (2017). doi:10.1111/j.2044-8317.1976.tb00705.x, <http://dx.doi.org/10.1111/j.2044-8317.1976.tb00705.x>

- Lord, F, Novick, M: Statistical theories of mental test scores. Addison-Wesley, Reading, MA (1968)
- Lumer, G: Semi-inner product spaces. *Trans. Amer. Math. Soc.* **100**(1), 29–43 (1961). doi:10.1103/PhysRevE.65.036127
- Marsaglia, G, Bray, T: A convenient method for generating normal variables. *SIAM Rev.* **6**, 260–264 (1964). doi:10.1109/TAC.1973.1100374
- McCulloch, CE: Tests for equality of variance for paired data. *Commun. Stat. Theory Methods.* **16** (1987). doi:10.1080/03610928708829445, <http://dx.doi.org/10.1080/03610928708829445>
- Mineo, AM, Ruggieri, M: A software tool for the exponential power distribution: The normalp package. *J. Stat. Softw.* **12**(4), 24 (2005). doi:10.1017/S0266466600013384
- Morgan, WA: A test for the significance of the difference between two variances in a sample from a normal bivariate population. *Biometrika.* **31** (1939)
- Moszyńska, M, Richter, W-D: Reverse triangle inequality. *Antinorms and semi-antinorms. Stud. Sci. Math. Hung.* **49**(1), 120–138 (2012). doi:10.1556/SScMath.49.2012.1.1192
- Mudholkar, GS, Wilding, GE, Mielowski, WL: Robustness properties of the pitman-morgan test. *Commun. Stat. Theory Methods.* **32** (2003). doi:10.1081/STA-120022710, <http://dx.doi.org/10.1081/STA-120022710>
- Müller, K, Richter, W-D: Exact extreme value, product, and ratio distributions under non-standard assumptions. *ASTA Adv. stat. Anal.* **99**(1), 1–30 (2015). doi:10.1007/s10182-014-0228-2
- Müller, K, Richter, W-D: Exact distributions of order statistics of dependent random variables from $I_{n,p}$ -symmetric sample distributions, $n \in \{3, 4\}$. *Depend Model.* **4**, 1–29 (2016a). doi:10.1515/demo-2016-0001
- Müller, K, Richter, W-D: Extreme value distributions for dependent jointly $I_{n,p}$ -symmetrically distributed random variables. *Depend Model.* **4**, 30–62 (2016b). doi:10.1515/demo-2016-0002
- Nadarajah, S: The Kotz-type distribution with applications. *Statistics.* **37**(4), 341–358 (2003). doi:10.1080/0233188031000078060
- Nolan, JP: An R package for modeling and simulating generalized spherical and related distributions. *J. Stat. Distrib. Appl.* **3**(14) (2016)
- Pascal, F, Bombrun, L, Tournet, JY, Berthoumieu, Y: Parameter estimation for multivariate generalized gaussian distribution. *IEEE Trans. Signal Process.* **61**(23), 5960–5971 (2017). arXiv 1302.6498v2
- Pitman, EJG: A note on normal correlation. *Biometrika.* **31** (1939). doi:10.1093/biomet/31.1-2.9, <http://dx.doi.org/10.1093/biomet/31.1-2.9>
- Purczynski, J, Bednarz-Okrzyńska, K: Estimation of the shape parameter of GED distribution for small sample size. *Folia Oeconomica Stetinensia.* **80**(2), 1–46 (2014)
- Rahman, M, Gokhale, D: On estimation of parameters of the exponential power family of distributions. *Commun. Stat. Simul. Comput.* **25**(2), 291–299 (1996). doi:10.1016/j.jspi.2003.09.008
- Richter, W-D: Generalized spherical and simplicial coordinates. *J. Math. Anal. Appl.* **336**, 1187–1202 (2007). doi:10.1016/j.jmaa.2007.03.047
- Richter, W-D: Continuous $I_{n,p}$ -symmetric distributions. *Lith. Math. J.* **49**(1), 93–108 (2009). doi:10.1007/s10986-009-9030-3
- Richter, W-D: Geometric disintegration and star-shaped distributions. *J. Stat. Distrib. Appl.* **1**(20), 1–24 (2014). doi:10.1186/s40488-014-0020-6
- Richter, W-D: Convex and radially concave contoured distributions. *J. Probab. Stat.* **2015**, 1–12 (2015a). doi:10.1155/2015/165468, article ID 165468
- Richter, W-D: Norm contoured distributions in \mathbb{R}^2 . In: *Lecture notes of Seminario Interdisciplinare di Matematica. Vol. XII*, vol 12, pp. 179–199. Potenza: Seminario Interdisciplinare di Matematica (S.I.M.), University of Basilicata, Italy, (2015b)
- Richter, W-D: Exact inference on scaling parameters in norm and antinorm contoured sample distributions. *J. Stat. Distrib. Appl.* **3**(8), 1–24 (2016). doi:10.1186/s40488-016-0046-z
- Richter, WD, Schicker, K: Ball numbers of platonic bodies. *J. Math. Anal. Appl.* **416**(2), 783–799 (2014). doi:10.1016/j.jmaa.2014.03.007
- Richter W-D, Schicker, K: Circle numbers of regular convex polygons. *Result Math.* **69**(3–4), 521–538 (2016a). doi:10.1007/s00025-016-0534-y
- Richter, W-D, Schicker, K: Polyhedral star-shaped distributions. *J. Probab. Stat.* (2016b). Volume 2017, <https://doi.org/10.1155/2017/7176897>
- Rothstein, SM, Bell, WD, Patrick, JA, Miller, H: A jackknife test of homogeneity of variance with paired replicates of data. *Psychometrika.* **46** (1981). doi:10.1007/BF02293916, <http://dx.doi.org/10.1007/BF02293916>
- Rubinstein, R: *Simulation and the Monte Carlo Method*. 1st ed. Wiley, New York (1981)
- Saatci, E, Akan, A: Respiratory parameter estimation in non-invasive ventilation based on generalized Gaussian noise models. *Signal Process.* **90**(2), 480–489 (2010). doi:10.1016/j.sigpro.2009.07.015
- Sarabia, JM, Gómez-Déniz, E: Construction of multivariate distributions: a review of some recent results. *SORT.* **32**(1), 3–36 (2008). with discussion by M.C. Pardo and J. Navarro
- Snedecor, GW, Cochran, W: *Statistical Methods*. University Press, Ames, IA (1967)
- Stacy, E, Mihram, G: Parameter estimation for a generalized Gamma distribution. *Technometrics.* **7**(3), 349–358 (1965). doi:10.1016/j.jspi.2003.09.008
- Taguchi, T: On a generalization of Gaussian distribution. *Ann. Inst. Stat. Math.* **30**, 211–242 (1978). doi:10.1007/BF02480215
- Tiku, ML, Balakrishnan, N: A robust test for testing the correlation coefficient. *Commun. Stat. Simul. Comput.* **15** (1986). doi:10.1080/03610918608812554, <http://dx.doi.org/10.1080/03610918608812554>
- Varanasi, M, Aazhang, B: Parametric generalized Gaussian density estimation. *J. Acoust. Soc. Amer.* **86**(4), 1404–1415 (1989). doi:10.1016/j.jspi.2003.09.008
- Wilcox, R: Comparing the variances of two dependent groups. *J. Educ. Stat.* **15**, 237–247 (1990). doi:10.1186/s40488-015-0030-z, <http://dx.doi.org/10.2307/1165033>
- Wilcox, R: Comparing the variances of two dependent variables. *J Stat Distrib Appl.* **2**(7), 8 (2015). doi:10.1186/s40488-015-0030-z
- Yu, S, Zhang, A, Li, H: A review of estimating the shape parameter of generalized Gaussian distribution. *J. Comput. Inform. Syst.* **8**(21), 9055–9064 (2012)