New Class of Spherical Pearson Type Family Distributions to Model Asymmetric Spherical Data

Mousa Golalizadeh[1] **and Meisam Moghimbeygi**[2]

[1] **Department of Statistics, Tarbiat Modares University, Tehran, Iran**

[2] **Department of Mathematics, Faculty of Mathematics and Computer Science, Kharazmi University, Tehran, Iran**

Abstract: The Pearson type family densities are among the most important classes of distributions that play a key role in the directional statistics. Their particular structures make them suitable candidates to analysis data on non-Euclidean space, such as sphere. To model data scattered asymmetrically on such spaces, the researchers confined themselves to extend particular distributions from the class of the Pearson type family densities. Those specific distributions are symmetric in nature but their extended versions are usually heavy tailed. In this paper, we introduce some alternative probability density functions in the class of Pearson type distributions on the sphere having the spherical Student's t, Fisher and Chi-square densities as the subfamilies. Via investigating various theoretical properties of this new subclass, we show that it is inheritably asymmetric. To further evaluating this subclass, some simulation studies are conducted. Also, modeling of two real-life data using the proposed densities and then comparing their fitting consequences with those resulted from invoking other common spherical distributions are considered.

Keywords: Spherical distribution, Pearson type family, Heavy-tailed distributions, Gaussian hypergeometric function, Asymmetric.

1 Introduction

It is known from elementary statistics that the Student's t-distribution plays a key role in statistical modeling and inferences, as does the normal density. However, it is not a suitable choice to model heterogeneity among data. Instead, a rich family of distributions, called the Pearson family and initially proposed by Pearson (1895) to model the observations on the Euclidean space, can take into account heterogeneity. Moreover, this family can be used to construct some continuous

[1] golalizadeh@modares.ac.ir

probability distribution functions to deal with skewness. Historically, this system was formerly defined in the context of numerical analysis leading to a valid solution of an associated differential equation (Johnson et al., 1995). The solutions for some of those particular equations are well-known densities such as the Student's t, Fisher and Chi-square distributions. Many progresses have been made in making statistical inference through invoking the Pearson type family distributions so far. However, most research was confined to the densities defined on the Euclidean space. There is an interest to study some subset of this family on the non-Euclidean space in the case of dealing with the data to be modelled using the directional statistics tools.

Many progress has been made over three decades or so to propose probability distributions on the sphere. Most densities defined on the sphere were symmetric in their forms. However, there are some rare exceptions. See, for example, Fisher et al. (1987) and Mardia and Jupp (2000) to consulate a comprehensive treatments of distributions on the non-Euclidean spaces in general and the unit sphere in particular. Let us assume there is an interest to model spherical data having some degrees of skewness.

Generally, there are two well-known methods to construct the spherical distributions. Following Mardia and Jupp (2000), the methods are categorized as the conditional and marginal approaches. For instance, the von Mises-Fisher (vMF) distribution is considered as a density derived using the conditional approach. Interestingly, most of the multivariate distributions can be conditioned in a proper way so that the corresponding densities will lead to some distributions on the sphere. Consequently, the resulted distributions on the sphere inherit relatively the same properties as the initial densities on the Euclidean space. It should be noted that most distributions derived in this way poses the rotational symmetry property. This important feature is relatively crucial while dealing with spherical densities. See, for example, Mardia and Jupp (2000), to consulate theoretical aspect of this property in the context of the directional statistics. However, there is rare activity on modeling skewness of the spherical data. Now, consider the case in which one is also interested in studying the spherical densities having bimodality property.

Some directional distributions can also be derived through wrapping, which is a method of the marginal approach. Following this methodology, some well-known densities such as the wrapped cauchy, wrapped normal and wrapped t-student distributions have been proposed in the literature as well as the popular book written by Mardia and Jupp (2000). The wrapped Cauchy distribution,

belonging to the Pearson family VII and being as a subclass of the Jones-Pewsey distribution, has been extensively considered in many studies. However, the most popular distribution on the circle is the wrapped normal distribution, again a density in the subclass of Pearson type VII.

Most studies treating the Pearson family distribution on the non-Euclidean space have concentrated on the particular cases; circle or sphere. For instance, Kato and Shimizu (2004) studied various aspects of the t-distribution on the sphere. Pewsey et al. (2013) gave a comprehensive treatment of different wrapped distributions along with their important properties all on the circle. Generally, there is an rare research on constructing spherical densities based on the Euclidean version of the Pearson type family distributions. An exception is the research done by Shimizu and Iida (2002), who studied the t-student distribution on the sphere as a subclass of Pearson type VII family.

Following Pewsey et al. (2007), who introduced some skew circular distributions on the unit circle, we propose some spherical densities that simultaneously poss skewness and bimodality properties. To do so, we construct the spherical t-distribution using conditional approach and via invoking the multivariate Student's t-distribution. Then, recalling the relationship between the Student's t, Fisher and Chi-square distributions on the Euclidean space, we propose their counterparts on the sphere as a subclass of the spherical Pearson type distributions. Also, statistical inferences on these densities along with a procedure to generate samples from the proposed distributions are provided. Moreover, an application of fitting the proposed distributions on real-life data and comparing the results with fitting them with some other spherical densities are presented.

The remainder of this paper is organized as follows. We first propose a new class of the skewed spherical densities in Section 2. We then consider making the statistical inference using two common approaches, namely method of moment and maximum likelihood method, in Section 3. Section 4 includes some simulation studies aiming at comparing the proposed distributions with the general class of vMF densities. An analysis of a real-life data is provided in Section 5. The paper concludes with general discussion and some insights on further possible research on this topic.

2 Pearson type family distributions on sphere

There are many cases where taking into account heavy rather than light tailed densities is preferred. Those distributions have particular structures and the way to construct them is also typical. As pointed out by Pearson (1916), those heavy tailed densities with some reasonable propoerties are in fact the members of type VII family of the Pearson system.

On the other hands, it is very common to use the vMF distribution to model data lie on the surface of the unit sphere. One of the main reason to do so is partly due to the fact that this density plays the same role as the normal distribution does on the Euclidean space (Mardia and Jupp, 2000). Following Shimizu and Iida (2002), the spherical t-distribution on the unit sphere is a proper candidate to study the heavy tailed spherical data. In fact, they proposed this density as a subclass of the Pearson type VII distributions. Here, we first review some important properties of this distribution using the standard spherical coordinate systems and then extend it to propose our new Pearson type VI and III densities and study their special sub-classes.

Let us consider random vector $X \in \mathbb{S}^{p-1}$ represented by its equivalent angle in the spherical coordinate systems, i.e., $\boldsymbol{\theta} = (\theta_1, ..., \theta_{p-1})'$. Note that, here, $\mathbb{S}^m$ denotes m-dimensional unit sphere. Following Mardia and Jupp (2000), the connection between two set of variables are given by

$$\begin{aligned} X &= (X_1, X_2, \ldots, X_{p-1}, X_p)' \\ &= (\cos\theta_1,\ \sin\theta_1\cos\theta_2,\ \ldots,\ \sin\theta_1\cdots\sin\theta_{p-2}\cos\theta_{p-1},\ \sin\theta_1\cdots\sin\theta_{p-2}\sin\theta_{p-1})'. \end{aligned}$$

Similarly, a mean direction vector with parameters $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_{p-1})'$, can be defined as

$$\begin{aligned} \mu_\circ &= (\mu_1^\circ, \mu_2^\circ, \ldots, \mu_{p-1}^\circ, \mu_p^\circ)' \\ &= (\cos\alpha_1,\ \sin\alpha_1\cos\alpha_2,\ \ldots,\ \sin\alpha_1\cdots\sin\alpha_{p-2}\cos\alpha_{p-1},\ \sin\alpha_1\cdots\sin\alpha_{p-2}\sin\alpha_{p-1})'. \end{aligned}$$

Using this typical definition of angular random variable, the spherical t-distribution can be derived via invoking its multivariate counterpart on the Euclidean space R^p. More details are given in next section.

2.1 Spherical t-distribution

Assume a p-dimensional random vector $\mathbf{T}$ follows the multivariate distribution with ν degree of freedom, i.e., $\mathbf{T} \sim t_\nu$. Then, for an outcome t of $\mathbf{T}$, the probability density function (pdf) is given

by

$$C[1+\frac{1}{\nu}(\mathbf{t}-\mu_{\circ}^{'})^{'}\Sigma^{-1}(\mathbf{t}-\mu_{\circ}^{'})]^{-\frac{\nu+p}{2}}, \tag{2.1}$$

where C, $'$ and Σ are the normalizing constant, the transpose operator and the covariance matrix of $\mathbf{T}$, respectively. To impose the constraints $||\mathbf{T}|| = R^2$, $||\mu_{\circ}|| = r^2$, where $||\,||$ is the Euclidean norm, and $\Sigma = \sigma I$, and to let $X = \mathbf{T}/R$ and $\mu = \mu_{\circ}/r$, the *pdf* given in (2.1) turns to

$$C[1+\frac{1}{\sigma\nu}(R^2+r^2-2Rrx^{'}\mu)]^{-\frac{\nu+p}{2}}. \tag{2.2}$$

Shimizu and Iida (2002) called this *pdf* the spherical t-distribution on $\mathbb{S}^{p-1}$. To set further constraints $R = r = \sigma = 1$ and to define $\kappa = 2/\nu$, the spherical t-density in (2.2) is simplified as

$$f_1(x|\mu,\kappa) = C_p(\kappa)[1+\kappa(1-x^{'}\mu)]^{-\frac{1}{\kappa}-\frac{p}{2}} \tag{2.3}$$

where

$$C_p^{-1}(\kappa) = \frac{2^{p-1}\pi^{\frac{p-1}{2}}\Gamma(\frac{p-1}{2})}{(1+2\kappa)^{(1/\kappa+p/2)}\Gamma(p-1)}\,{}_2F_1(\frac{1}{\kappa}+\frac{p}{2},\frac{p-1}{2};p-1;\frac{2\kappa}{1+\kappa})$$

and ${}_2F_1(a,b;c;z)$ is the Gaussian hypergeometric function (Abramowitz and Stegun, 1965). Usually, the interest is on the special case $p = 3$, i.e. the unit sphere. Then, the normalized constant of the density for this particular case is simplified as

$$C_3(\kappa) = \frac{1}{2\pi}\frac{1+\kappa/2}{1-(1+2\kappa)^{-1/\kappa-1/2}}.$$

To shorten the notations for our latter computational tasks, we denote the spherical t-distribution along with its parameters as $\mathrm{ST}_p(\mu,\kappa)$, where μ represents the mean direction and κ is known as the concentration parameter. See, e.g. Mardia and Jupp (2000) for more details on the concepts and interpretations of these parameters.

Using the tangent-normal decomposition, see e.g., Fisher et al. (1987), the random variables θ_1 and $(\theta_2,...,\theta_{p-1})$ are independent on the pole. Hence, the spherical Pearson type densities on the pole can, indeed, be expressed as the product of two independent densities. Moreover, it can be shown that the random vector $(\theta_2,...,\theta_{p-1})$ is distributed uniformly on $\mathbb{S}^{p-2}$. Therefore, the random variable θ_1 plays a key role in expressing the $\mathrm{ST}_p(\mu,\kappa)$ distribution on the pole.

Now, suppose one is interested in deriving the moments of the first component of the spherical variable on $\mathbb{S}^{p-1}$, i.e. X_1. Based upon the earlier discussion on the properties of the $\mathrm{ST}_p(\mu,\kappa)$, it

is straightforward to convey this objective in terms of the density of θ_1. To do so, let us denote the m-th moment of X_1, by $\rho_m(\kappa)$. Then,

$$\begin{aligned}
\rho_m(\kappa) &= E(\cos^m\theta_1) \\
&= C^{-1}_{\theta_1,p}(\kappa)\int_0^{\pi}\cos^m\theta_1[1+\kappa-\kappa\cos\theta_1]^{-\frac{1}{\kappa}-\frac{p}{2}}\sin^{p-2}\theta_1 d\theta_1 \\
&= C^{-1}_{\theta_1,p}(\kappa)\int_{-1}^{1} t^m[1+\kappa-\kappa t]^{-\frac{1}{\kappa}-\frac{p}{2}}(1-t^2)^{\frac{p-3}{2}}dt \\
&= \frac{2^{p-2}C^{-1}_{\theta_1}}{(1+2\kappa)^{\frac{1}{\kappa}+\frac{p}{2}}}\int_0^1(2u-1)^m(1-\frac{2\kappa}{1+2\kappa}u)^{-\frac{1}{\kappa}-\frac{p}{2}}(1-u)^{(p-3)/2}u^{\frac{p-3}{2}}du \\
&= \sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i}\frac{2^{i+p-2}C^{-1}_{\theta_1}}{(1+2\kappa)^{\frac{1}{\kappa}+\frac{p}{2}}}\int_0^1(1-\frac{2\kappa}{1+2\kappa}u)^{-\frac{1}{\kappa}-\frac{p}{2}}(1-u)^{\frac{p-3}{2}}u^{i+\frac{p-3}{2}}du \\
&= \sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i}2^i\frac{\Gamma(i+\frac{p-1}{2})\Gamma(p-1)}{\Gamma(i+p-1)\Gamma(\frac{p-1}{2})}\frac{{}_2F_1(\frac{1}{\kappa}+\frac{p}{2},i+\frac{p-1}{2};i+p-1;\frac{2\kappa}{1+2\kappa})}{{}_2F_1(\frac{1}{\kappa}+\frac{p}{2},\frac{p-1}{2};p-1;\frac{2\kappa}{1+2\kappa})}, \quad (2.4)
\end{aligned}$$

where $C_{\theta_1,p}(\kappa)$ is the normalized constant of the density of the angular variable θ_1.

To recall the relationship between the Student's t and Fisher distributions on the plane, we shall obtain the spherical Fisher distribution in the next section.

2.2 Spherical Fisher distribution

It is known from the elementary statistics that if a random variable, say U, follows the Fisher distribution with parameters α and β, then its probability density function, for outcome u, will be proportional to

$$u^{\frac{p}{2}-1}(1+\frac{\alpha}{\beta}u)^{-\frac{\alpha+\beta}{2}}.$$

Considering the special case $\alpha = 2$ and $\beta = 2/\kappa$, it is then seen that the Fisher distribution is closely linked to the $\mathrm{ST}_2(\mu,\kappa)$. This can be identified via considering the equality $U = 1 - X^{'}\mu$. On the other hand, the truncated $\mathrm{F}(2,\kappa)$-distribution with parameters α_1 and α_2 has the following *pdf*,

$$f(x|\alpha_1,\alpha_2,\sigma) \propto (1-x^{'}\mu)^{\alpha_1-1}(1+\frac{1-x^{'}\mu}{\sigma})^{-(\alpha_1+\alpha_2)}. \quad (2.5)$$

Note that the truncated $\mathrm{F}(2,\kappa)$-density is also known as the Beta Prime distribution (Johnson et al., 1995) and is considered as a subclass of the Pearson type VI distributions.

This last connection motivated us to extend the $\mathrm{ST}_p(\mu,\kappa)$ density to derive the spherical Fisher distribution, as a proper candidate to analyze directional data. Let us assume the *pdf* of this new density is denoted by $f_2(x|\mu,\kappa)$. We can then write the spherical Fisher distribution along with its

parameters μ and κ; identified by the notation $\mathrm{SF}_p(\mu,\kappa)$, expressed by following *pdf*,

$$f_2(x|\mu,\kappa)=C_p(\kappa)(1-x^{'}\mu)^{\frac{p}{2}-1}(1+\frac{p\kappa}{2}(1-x^{'}\mu))^{-\frac{p}{2}-\frac{1}{\kappa}}$$

where

$${C_p}^{-1}(\kappa)=\frac{2^{\frac{3p}{2}-2}\pi^{\frac{p-1}{2}}\Gamma(p-\frac{3}{2})}{\Gamma(\frac{3p}{2}-2)}(1+p\kappa)^{-\frac{p}{2}-\frac{1}{\kappa}}{}_2F_1(\frac{p}{2}+\frac{1}{\kappa},\frac{p-1}{2},\frac{3p-4}{2},\frac{p\kappa}{p\kappa+1}).$$

It is seen that for the particular case $p=3$,

$${C_3}^{-1}(\kappa)=2\pi(\frac{2}{3\kappa})^{\frac{3}{2}}B(\frac{3\kappa}{3\kappa+1};\frac{3}{2},\frac{1}{\kappa})$$

where $B(\cdot)$ is incomplete beta function (Abramowitz and Stegun, 1965) represented by

$$B(x;a,b)=\int_0^x t^{a-1}(1-t)^{b-1}dt.$$

Similar to discussion about the $\mathrm{ST}_p(\mu,\kappa)$, the m-th moment of X_1, while the spherical variable X follows the $\mathrm{SF}_p(\mu,\kappa)$ on the pole, is obtained using the following expressions:

$$\begin{aligned}
\rho_m(\kappa) &= E(\cos^m\theta_1)\\
&= C^{-1}_{\theta_1,p}(\kappa)\int_0^{\pi}\cos^m\theta_1(1-\cos\theta_1)^{\frac{p}{2}-1}(1+\frac{p\kappa}{2}(1-\cos\theta_1))^{-\frac{p}{2}-\frac{1}{\kappa}}\sin^{p-2}\theta_1 d\theta_1\\
&= C^{-1}_{\theta_1,p}(\kappa)\int_{-1}^{1}t^m(1-t)^{\frac{p}{2}-1}(1+\frac{p\kappa}{2}(1-t))^{-\frac{p}{2}-\frac{1}{\kappa}}(1-t^2)^{\frac{p-3}{2}}dt\\
&= 2^{\frac{3p}{2}-3}C^{-1}_{\theta_1,p}(\kappa)(1+p\kappa)^{-\frac{p}{2}-\frac{1}{\kappa}}\int_0^1(2u-1)^m(1-\frac{p\kappa}{p\kappa+1}u)^{-\frac{p}{2}-\frac{1}{\kappa}}(1-u)^{p-\frac{5}{2}}u^{\frac{p-3}{2}}du\\
&= \sum_{i=0}^{m}\frac{\binom{m}{i}(-1)^{m-i}2^{\frac{3p}{2}-3+i}C^{-1}_{\theta_1}}{(1+p\kappa)^{\frac{p}{2}+\frac{1}{\kappa}}}\int_0^1(1-\frac{p\kappa}{p\kappa+1}u)^{-\frac{p}{2}-\frac{1}{\kappa}}(1-u)^{p-\frac{5}{2}}u^{i+\frac{p-3}{2}}du\\
&= \sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i}2^i\frac{\Gamma(\frac{p-1}{2}+i)\Gamma(\frac{3p-4}{2})}{\Gamma(\frac{3p-4}{2}+i)\Gamma(\frac{p-1}{2})}\frac{{}_2F_1(\frac{p}{2}+\frac{1}{\kappa},i+\frac{p-1}{2};i+\frac{3p-4}{2};\frac{p\kappa}{p\kappa+1})}{{}_2F_1(\frac{p}{2}+\frac{1}{\kappa},\frac{p-1}{2};\frac{3p-4}{2};\frac{p\kappa}{p\kappa+1})} \qquad (2.6)
\end{aligned}$$

where $C_{\theta_1,p}(\kappa)$ is the normalized constant of density. Once again, it is seen that for the particular case $p=3$,

$$\rho_1(\kappa)=\frac{4}{5}\frac{{}_2F_1(\frac{1}{\kappa}+\frac{3}{2},2;\frac{7}{2};\frac{3\kappa}{3\kappa+1})}{{}_2F_1(\frac{1}{\kappa}+\frac{3}{2},1;\frac{5}{2};\frac{3\kappa}{3\kappa+1})}-1.$$

The connections between the Pearson type family distributions help us to propose other spherical distributions too. For instance, the Pearson type VI and III distributions, have strong link with the Chi-square distribution on the plane. Hence, we invoke this connection to introduce the spherical Chi-square distribution in subsequent section.

2.3 Spherical Chi-square distribution

Our objective here is to impose some restrictions in the density given by (2.5) to derive a new distribution useful in directional statistics. Let us fix $\alpha_1 = \kappa/2, \sigma = \alpha_2$ and allow α_2 to get a very large number. Then, it can be shown that the resulted density is a new distribution in the class of the Pearson type III family with the following *pdf*

$$f_3(x|\mu,\kappa) = \frac{\Gamma(\frac{\kappa}{2}+p-2)(1-x^{'}\mu)^{\frac{\kappa}{2}-1}e^{x^{'}\mu+1}}{2^{\frac{\kappa}{2}+p-2}\pi^{\frac{p-1}{2}}\Gamma(\frac{\kappa}{2}+p-3)M(\frac{p-1}{2},\frac{\kappa}{2}+p-2,2)} \tag{2.7}$$

where $M(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function (Abramowitz and Stegun, 1965). We call this new density the Spherical Chi-square distribution and denote this statement with the notation $\mathrm{SC}_p(\mu,\kappa)$. It is straightforward to show that for particular case $p = 3$, the *pdf* in (2.7) reduces to

$$f_3(x|\mu,\kappa) = \frac{(1-x^{'}\mu)^{\frac{\kappa}{2}-1}e^{x^{'}\mu-1}}{2\pi\gamma(\frac{\kappa}{2},2)}$$

where $\gamma(\cdot,\cdot)$ is the lower incomplete gamma function.

Similar to discussion provided in dealing with $\mathrm{ST}_p(\mu,\kappa)$, the m-th moment of X_1, can be obtained through the following sequel equalities, provided the spherical variable X follows $\mathrm{SC}_p(\mu,\kappa)$

$$\begin{aligned}
\rho_m(\kappa) &= E(\cos^m\theta_1) \\
&= C^{-1}_{\theta_1,p}(\kappa)\int_0^{\pi}\cos^m\theta_1(1-\cos\theta_1)^{\frac{\kappa}{2}-1}e^{\cos\theta_1}\sin^{p-2}\theta_1 d\theta_1 \\
&= C^{-1}_{\theta_1,p}(\kappa)\int_{-1}^{1} t^m(1-t)^{\frac{\kappa}{2}-1}e^{t}(1-t^2)^{\frac{p-3}{2}}dt \\
&= 2^{p-3+\frac{\kappa}{2}}C^{-1}_{\theta_1,p}(\kappa)\int_0^1(2u-1)^m(1-u)^{\frac{p-3+\kappa}{2}-1}u^{\frac{p-3}{2}}e^{2u-1}du \\
&= \sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i}2^i\frac{\Gamma(\frac{p-1}{2}+i)\Gamma(p-2+\frac{\kappa}{2})M(\frac{p-1}{2}+i,p-2+\frac{\kappa}{2}+i,2)}{\Gamma(\frac{p-1}{2})\Gamma(p-2+\frac{\kappa}{2}+i)M(\frac{p-1}{2},p-2+\frac{\kappa}{2},2)}
\end{aligned} \tag{2.8}$$

where $C_{\theta_1,p}(\kappa)$ is the normalized constant. It is seen that, for the particular case $p = 3$, we have

$$\rho_1 = \frac{2M(2,2+\frac{\kappa}{2},2)}{(1+\frac{\kappa}{2})M(1,1+\frac{\kappa}{2},2)} - 1.$$

To have consistent notation throughout this paper, we use the notation $\mathrm{SPT}_p(\mu,\kappa)$, stands for the Spherical Pearson Type distributions including the t-student, Fisher and chi-square densities. However, if we wants to deal with a particular Pearson type distribution, we will provide explicit notation, as discussed earlier, along with expressing *pdf* of corresponding density.

3 Statistical Inference on the Parameters of $\mathbf{SPT}_p(\mu, \kappa)$

Now, we provide the way of making the statistical inference on the parameters of the spherical Pearson type distributions. We shall deal with both the method of moment and maximum likelihood approaches.

Let us assume the spherical random variables X_i for $i = 1, 2, ..., n$, are n independent observations following the spherical Pearson type distributions with the mean direction μ and the concentration parameter κ on $\mathbb{S}^{p-1}$. The statistical inference about μ and κ, are given below using two approaches.

3.1 Method of Moment Estimator

In order to calculate moments of the SPT distribution, we invoke its rotational symmetry property. The property for our considered density can simply be conveyed as follows:

If Y is distributed as $SPT_p(\mu_0, \kappa)$ and M is a rotation matrix, which rotats μ_0 to μ, then $X = MY$ is a random vector distributed as $SPT_p(\mu, \kappa)$.

As seen, this fruitful property shows that the rotated variable preserves the initial form of the density with just a different directional mean compared with the initial (raw) variable. We shall see that this interesting property will later help us in dealing with theoretical computations tasks greatly. To consult more properties of the rotational symmetry, one can see, for example, Mardia and Jupp (2000).

Using the rotational symmetry property, we can write

$$\begin{aligned} E(X) &= M \times E(Y) \\ &= M \times E\big\{(\cos\theta_1, \sin\theta_1 \cos\theta_2, ..., \sin\theta_1 \cdots \sin\theta_{d-2} \cos\theta_{d-1}, \sin\theta_1 \cdots \sin\theta_{d-2} \sin\theta_{d-1})\big\} \\ &= \rho_1(\kappa) M \times \mu_0 = \rho_1(\kappa)\mu \end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}(X) &= M \times \mathrm{Var}(Y) \times M^{'} \\
&= M \times \{E(YY^{'}) - E(Y)E(Y^{'})\} \times M^{'} \\
&= M\{\mathrm{diag}[\rho_2(\kappa), \frac{1-\rho_2(\kappa)}{p-1}, ..., \frac{1-\rho_2(\kappa)}{p-1}] - \rho_1^2(\kappa)\mu_0\mu_0^{'}\}M^{'} \\
&= M\{\rho_2(\kappa)\mu_0\mu_0^{'} + \frac{1-\rho_2(\kappa)}{p-1}(I_p - \mu_0\mu_0^{'}) - \rho_1^2(\kappa)\mu_0\mu_0^{'}\}M^{'} \\
&= (\rho_2(\kappa) - \rho_1^2(\kappa))\mu\mu^{'} + \frac{1-\rho_2(\kappa)}{p-1}(I_p - \mu\mu^{'})
\end{aligned}$$

where, I_p is the identity matrix of order p.

As is common in directional statistics, we use cosine moments to estimate the parameters of the SPT distributions by the method of moment. Following this perspective, the mean direction is easily estimated as

$$\tilde{\mu} = \frac{\sum_{i=1}^{n} x_i}{||\sum_{i=1}^{n} x_i||}.$$

To estimate the concentration parameter (κ), one can again recall the first moment of the SPT. Using the equation (3.1), we have

$$E(X^{'})\mu = \rho_1(\kappa).$$

Then, the estimate of κ can be written as

$$\tilde{\kappa} = \rho_1^{-1}(\frac{1}{n}\sum_{i=1}^{n} x_i^{'}\tilde{\mu}).$$

As seen, one needs to follow some numerically iterated methods to derive the estimates of the parameters using the method of moment because the expressions involving $\tilde{\mu}$ and $\tilde{\kappa}$ are interconnected. There are many algorithms to follow this task. We shall see an implementation of such technique when we conduct and present our simulation studies.

3.2 Maximum Likelihood Estimator

To invoke the maximum likelihood estimate (MLE) approach, it is common to maximize the (logarithm of the) likelihood function of the parameter(s), which is associated with the *pdf* of the corresponding variables. To this end, we use Lagrange functions $\mathcal{L}_j$, for $j = 1, 2, 3$, which are in fact the logarithm of the likelihood functions, i.e.

$$\mathcal{L}_j(\mu, \kappa) = \ell_j(\mu, \kappa) + \lambda(1 - \mu^{'}\mu),$$

where $\ell_j(\mu,\kappa)=\sum_{i=1}^{n}\log f_j(x_i|\mu,\kappa)$, with the obvious constraint $\mu^{'}\mu=1$.

To differentiate $\mathcal{L}_j(\mu,\kappa)$ with respect to μ and κ, the maximum likelihood estimations are given by the following expressions:

$$\hat{\mu} = \Big\{\frac{\partial\ell(\mu,\kappa)/\partial\mu}{\ell(\mu,\kappa)}\Big\}\Big/\|\frac{\partial\ell(\mu,\kappa)/\partial\mu}{\ell(\mu,\kappa)}\| \tag{3.1}$$

$$\hat{\kappa} = \arg\max_{\kappa}\ell(\mu,\kappa). \tag{3.2}$$

It is clear that the MLEs of the parameters should be derived by some iterative algorithms through invoking the equations (3.1) and (3.2). Before describing a general flowchart of such algorithm, it worths pointing out an important assertion useful to derive the MLEs.

Using the density function (2.5), it can be shown that for a small value of σ or equivalently large value of κ, the estimation of μ is approximated by

$$\hat{\mu}=\frac{\sum_{i=1}^{n}\frac{x_i}{1-x_i^{'}\tilde{\mu}}}{||\sum_{i=1}^{n}\frac{x_i}{1-x_i^{'}\tilde{\mu}}||}, \tag{3.3}$$

where $\tilde{\mu}$ is the estimate of μ given by the MM approach. So, this estimation can be used as an initial value while invoking any iterative algorithm to derive the parameter estimates.

In summary, the following simple iterative steps can be followed to obtain the parameters estimate of the $\mathrm{SPT}_p(\mu,\kappa)$ using the ML method:

Step 0. Compute the initial value for $\hat{\mu}$, using the expression (3.3).

Step 1. Derive an estimate for κ using the equation (3.2).

Step 2. Having $\hat{\kappa}$ from **Step 1**, obtain an estimate for μ using the equation (3.1).

Step 3. Iterate between **Step 1** and **Step 2** using the updated parameters until reaching an reasonable rate of convergence; fixed before running the algorithm.

4 Simulation Study

Due to both popularity and importance of three dimensional sphere in many real applications, we confine ourselves to conduct the simulation studies on $\mathbb{S}^2$. Our investigations will be performed in two different scenarios. First, we compare the performances of the MM and MLE estimators in different settings. Then, we study fitting the SPT distributions in the cases where the generated

data are imposed via contaminating observations with an specific degree of noises to constitute some outliers. Let us, first, explain how to simulate the spherical data come from the SPT density.

In order to simulate the data from $\mathrm{ST}_3(\mu,\kappa)$, $\mathrm{SF}_3(\mu,\kappa)$ and $\mathrm{SC}_3(\mu,\kappa)$, we invoke the distribution functions of $\theta \in (0,\pi)$ and $\phi \in (0,2\pi)$ on the pole, where these two variables are independent. In particular, while being on the pole, ϕ is distributed uniformly on the unit circle, independent from θ. So, the distribution function of ϕ is the same for all three densities. In particular, not being concerned about the specific distribution of the SPT, we have

$$G(\phi) = \frac{\phi}{2\pi}, \qquad 0 < \phi < 2\pi.$$

On the other hands, the density function of θ, following three aforementioned distributions are, respectively,

$$\begin{aligned} F_{ST}(\theta) &= \frac{1-(1+\kappa-\kappa\cos\theta)^{-\frac{1}{\kappa}-\frac{1}{2}}}{1-(1+2\kappa)^{-\frac{1}{\kappa}-\frac{1}{2}}}, \\ F_{SF}(\theta) &= \frac{B(\frac{3(1-\cos\theta)\kappa}{3(1-\cos\theta)\kappa+2};\frac{3}{2},\frac{1}{\kappa})}{B(\frac{3\kappa}{3\kappa+1};\frac{3}{2},\frac{1}{\kappa})}, \\ F_{SC}(\theta) &= \frac{\gamma(\frac{\kappa}{2},1-\cos\theta)}{\gamma(\frac{\kappa}{2},2)}. \end{aligned}$$

As seen, to simulate θ from either densities is straightforward using available technological tools.

A remark to note is that the ultimate pair (ϕ,θ) generated then is a point on the pole. But, in real application the spherical data might live somewhere else on the sphere than the pole. In particular, the data are assumed to have unknown spherical mean μ. This problem can also be circumvented using a simple rotation matrix. There are many rotation matrices for this purpose. But, a simple one, proposed by Mardia and Jupp (2000), which rotates the spherical point m_1 to its counterpart m_2 or vice versa, is given by

$$M(m_1,m_2) = \frac{(m_1+m_2)(m_1+m_2)'}{1+m_1'm_2} - I_3.$$

Now, we are at the position to simulate spherical data from three proposed distributions described in the previous sections. This along with some comments on the relevant parameters are given next.

We initiate our simulation study by generating data from the $\mathrm{SPT}_3(\mu,\kappa)$ distributions. This is done via altering the location parameters to be one of the pairs $(\alpha,\beta) = \{(\pi/4,\pi/4),(\pi/3,\pi/6)\}$ while the concentration parameters (κ) is fixed at $\kappa = 1.8$. Also, to investigate the effects of sample

size in estimating parameters, we consider four scenarios $n = 20, 50, 100, 1000$. Furthermore, we iterate each simulation run 1000 times. Note that the large sample case gives us insights on the asymptotic behaviour of our estimators though the theoretical aspects of this issue needs to be comprehensively studied. To evaluate the accuracy of the proposed estimators, we calculate the mean, relative bias (RB) and mean square errors (MSE) for every simulation scenario. The entire results of the simulation studies based upon different scenarios are reported in tables 1, 2 and 3. It worths to mention that the means of parameters are highlighted by a dashed line over the estimates of corresponding parameters in each table. Also, for brevity of notation, we write MLE and MM, indicating the maximum likelihood and method of moment estimators, respectively.

As seen in all three tables, the estimations provided by the MM procedure are generally performing better than those derived through the MLE method in the small sample size situations. But, this is not the case when the numbers of samples are increasing which leads to better performance of the MLEs. Following the results reported in Tables 1, the accuracy of the estimation of the directional means for the ST distribution is improved by increasing either the concentration parameter (κ) or the sample size (n). However, this pattern is not continuing while estimating the concentration parameter. That means, if either the sample size or the directional mean is increased then the estimators proposed for estimating κ, are doing worth.

Following the results reported in Table 2, we see that the performance of two methods in estimating the directional mean using the SF distribution is relatively the same. Note that, if the concentration parameter is increased, unlike the RB measure, the values of the MSEs are increasing too. This shows critical impact of the concentration parameter on other parameters of spherical Pearson type family densities studied in this paper. Note that the similar patterns are also observed in Table 3 while the SC distribution is invoked in both generating the spherical data and estimating the parameters.

To have an visual inspection of efficiency of the methods in different situations as well as in the case of various distributions proposed in this paper, we provided the comparative plots in 1, 2 and 3. In all plots, the first and second cases refer to the simulation study in which the parameters of the SPT densities are fixed as $(\alpha, \beta, \kappa) = (\frac{\pi}{4}, \frac{\pi}{4}, 1)$, and $(\alpha, \beta, \kappa) = (\frac{\pi}{3}, \frac{\pi}{6}, 8)$, respectively. Also, the approximate 95% confidence intervals have been computed for each density and particular sample size. The computed intervals are drawn as solid lines with the average of the estimates at the

middle, identified by either the ML or MM approaches.

Many remarks can be driven from the plots. As seen, to increase the sample size leads to the accurate estimates regardless of the type of the distribution considered for fitting the data. Moreover, two methods of estimations, i.e., ML and MM, are performing the same in most scenarios. However, there are some cases in which the MM estimators are doing better than ML, particularly when the concentration parameter (κ) is small.

The approximate confidence intervals for the concentration parameter (κ) while using small sample size and the ST, if the real values are $(\alpha, \beta, \kappa) = (\frac{\pi}{4}, \frac{\pi}{4}, 1)$, contains some negative values which are not feasible. This shows that the standard errors of estimator in this scenario are too big. This situation occurs for both the ML and MM estimators. Also, we just see this odd outcome for the ST density while other two distributions are at the safe side. Amazingly, the approximate confidence intervals for the concentration parameter (κ) are too narrow if the SC density is used in the first case, i.e., $(\alpha, \beta, \kappa) = (\frac{\pi}{4}, \frac{\pi}{4}, 1)$.

Under the scenario $(\alpha, \beta, \kappa) = (\frac{\pi}{3}, \frac{\pi}{6}, 8)$, the performance of either estimators using different distributions are relatively reasonable. Although the estimators for α and β display appropriate patterns for all three considered densities, that for κ varies too much. We mean the estimator of κ neither follows the asymptotic manner nor shows equal pattern for three distributions. It might be due to large values of its real value, $\kappa = 8$ imposing remarkable variation in each sample path. Do remember that the spherical data might display multimodality on the empirical density in this case. Hence, the common procedures proposed in this paper fail to make sensible inference. This topic can then be studied in further research.

Table 1: Mean, RB and MSE of the parameter estimations using the ML and MM approaches under different scenarios. Dashed lines over the estimators indicate the mean of the estimations in 1000 runs derived in using from ST distribution. Real values of the parameters are on the top of each panel.

$(\alpha, \beta, \kappa) = (\frac{\pi}{4}, \frac{\pi}{4}, 1)$

Method	n	$\bar{\hat{\alpha}}$	$\text{RB}_{\hat{\alpha}}$	$\text{MSE}_{\hat{\alpha}}$	$\bar{\hat{\beta}}$	$\text{RB}_{\hat{\beta}}$	$\text{MSE}_{\hat{\beta}}$	$\bar{\hat{\kappa}}$	$\text{RB}_{\hat{\kappa}}$	$\text{MSE}_{\hat{\kappa}}$
ML	20	0.847	0.078	0.058	0.713	-0.092	0.214	2.366	1.366	4.601
	50	0.798	0.016	0.028	0.789	0.005	0.061	1.580	0.580	1.041
	100	0.791	0.007	0.014	0.772	-0.018	0.030	1.278	0.278	0.354
	1000	0.787	0.002	0.001	0.783	-0.003	0.003	1.009	0.009	0.029
MM	20	0.837	0.065	0.062	0.726	-0.076	0.218	2.367	1.367	5.369
	50	0.798	0.016	0.030	0.786	0.001	0.067	1.590	0.590	1.590
	100	0.793	0.010	0.016	0.773	-0.016	0.031	1.285	0.285	0.552
	1000	0.788	0.003	0.002	0.783	-0.003	0.003	1.011	0.011	0.040
$(\alpha, \beta, \kappa) = (\frac{\pi}{3}, \frac{\pi}{6}, 8)$										
ML	20	1.045	-0.002	0.016	0.521	-0.005	0.025	8.692	0.087	3.365
	50	1.046	-0.002	0.007	0.525	0.003	0.008	8.633	0.079	3.602
	100	1.048	0.001	0.003	0.524	0.001	0.004	8.335	0.042	2.710
	1000	1.047	0.00035	0.00032	0.524	-0.00001	0.00044	8.094	0.01173	0.5053
MM	20	1.051	0.004	0.025	0.520	-0.007	0.036	7.338	-0.083	5.843
	50	1.051	0.003	0.010	0.526	0.005	0.013	7.478	-0.065	5.377
	100	1.049	0.002	0.005	0.527	0.006	0.006	7.878	-0.015	4.188
	1000	1.047	-0.00018	0.00045	0.523	-0.00035	0.00069	8.131	0.01633	0.9492

Table 2: Mean, RB and MSE of the parameter estimations using the ML and MM approaches under different scenarios. Dashed lines over the estimators indicate the mean of the estimations in 1000 runs using the from SF distribution. Real values of the parameters are on the top of each panel.

$(\alpha, \beta, \kappa) = (\frac{\pi}{4}, \frac{\pi}{4}, 1)$

Method	n	$\bar{\hat{\alpha}}$	$\text{RB}_{\hat{\alpha}}$	$\text{MSE}_{\hat{\alpha}}$	$\bar{\hat{\beta}}$	$\text{RB}_{\hat{\beta}}$	$\text{MSE}_{\hat{\beta}}$	$\bar{\hat{\kappa}}$	$\text{RB}_{\hat{\kappa}}$	$\text{MSE}_{\hat{\kappa}}$
ML	20	0.925	0.178	0.208	0.404	-0.485	0.651	1.131	0.131	0.980
	50	0.788	0.003	0.068	0.685	-0.128	0.321	1.277	0.277	0.839
	100	0.790	0.005	0.045	0.737	-0.061	0.259	1.219	0.219	0.551
	1000	0.783	-0.003	0.005	0.779	-0.008	0.011	1.049	0.049	0.060
MM	20	0.926	0.179	0.196	0.388	-0.506	0.661	1.291	0.291	0.808
	50	0.790	0.005	0.059	0.691	-0.120	0.292	1.341	0.341	0.753
	100	0.784	-0.002	0.035	0.790	0.006	0.142	1.342	0.342	0.737
	1000	0.781	-0.006	0.004	0.781	-0.005	0.008	1.071	0.071	0.097
$(\alpha, \beta, \kappa) = (\frac{\pi}{3}, \frac{\pi}{6}, 8)$										
ML	20	1.052	0.005	0.074	0.467	-0.108	0.148	7.978	-0.003	1.406
	50	1.058	0.010	0.031	0.516	-0.015	0.041	8.035	0.004	1.179
	100	1.056	0.008	0.017	0.520	-0.007	0.023	8.000	0.000	1.317
	1000	1.047	0.000	0.001	0.514	-0.017	0.002	7.984	-0.002	0.893
MM	20	1.066	0.018	0.065	0.488	-0.068	0.110	8.011	0.001	1.514
	50	1.062	0.014	0.030	0.520	-0.008	0.043	7.956	-0.005	1.496
	100	1.053	0.006	0.017	0.521	-0.004	0.021	7.840	-0.020	1.267
	1000	1.046	-0.002	0.001	0.514	-0.019	0.002	8.074	0.009	0.889

Table 3: Mean, RB and MSE of the parameter estimations. The data are generated from SC distribution using the ML and MM approaches under different scenarios. Dashed lines over the estimators indicate the mean of the estimations in 1000 runs using the from SC distribution. Real values of the parameters are on the top of each panel.

$(\alpha, \beta, \kappa) = (\frac{\pi}{4}, \frac{\pi}{4}, 1)$

Method	n	$\bar{\hat{\alpha}}$	$\text{RB}_{\hat{\alpha}}$	$\text{MSE}_{\hat{\alpha}}$	$\bar{\hat{\beta}}$	$\text{RB}_{\hat{\beta}}$	$\text{MSE}_{\hat{\beta}}$	$\bar{\hat{\kappa}}$	$\text{RB}_{\hat{\kappa}}$	$\text{MSE}_{\hat{\kappa}}$
ML	20	0.763	-0.028	0.020	0.835	0.063	0.054	0.643	-0.357	0.134
	50	0.783	-0.003	0.005	0.785	-0.001	0.011	0.823	-0.177	0.039
	100	0.783	-0.003	0.002	0.794	0.011	0.004	0.994	-0.006	0.007
	1000	0.785	0.000	0.000	0.787	0.002	0.000	1.004	0.004	0.001
MM	20	0.798	0.016	0.029	0.793	0.010	0.068	0.973	-0.027	0.084
	50	0.798	0.016	0.013	0.790	0.006	0.021	0.978	-0.022	0.033
	100	0.789	0.005	0.005	0.786	0.001	0.011	1.001	0.001	0.015
	1000	0.784	-0.002	0.001	0.788	0.003	0.001	1.005	0.005	0.002

$(\alpha, \beta, \kappa) = (\frac{\pi}{3}, \frac{\pi}{6}, 8)$

ML	20	1.054	0.007	0.062	0.496	-0.053	0.082	8.031	0.004	1.618
	50	1.033	-0.014	0.027	0.545	0.042	0.045	8.057	0.007	1.092
	100	1.045	-0.002	0.015	0.515	-0.017	0.020	8.090	0.011	0.498
	1000	1.050	0.002	0.010	0.522	-0.002	0.001	8.000	0.010	0.344
MM	20	1.064	0.016	0.056	0.500	-0.045	0.091	8.107	0.013	1.029
	50	1.011	-0.035	0.030	0.538	0.027	0.042	8.122	0.015	0.848
	100	1.046	-0.001	0.014	0.517	-0.012	0.022	8.101	0.013	0.453
	1000	1.050	0.002	0.014	0.522	-0.001	0.020	8.001	0.012	0.386

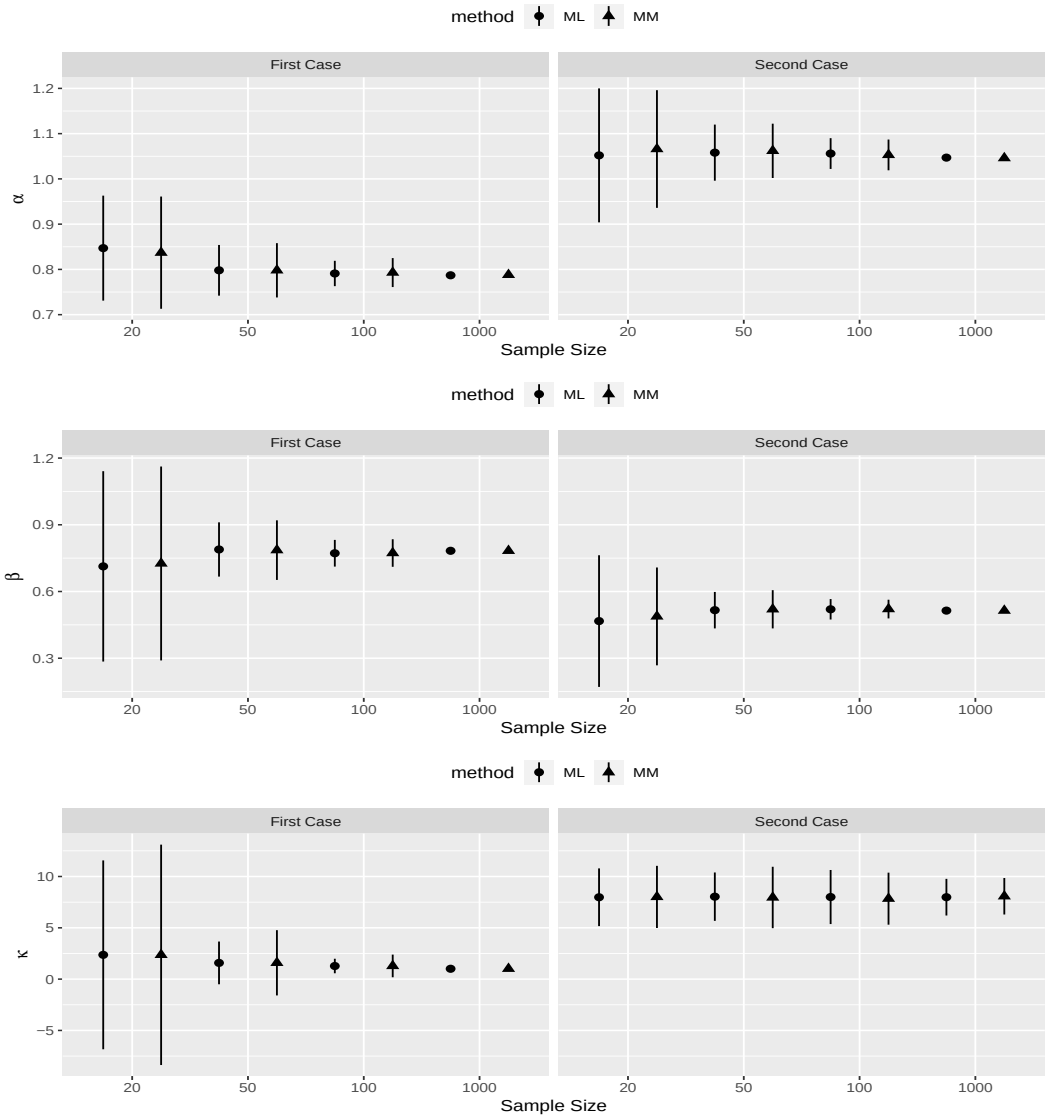

Figure 1: Comparative plots for the estimators of the ST's parameters using two estimation procedures. The approximated %95 CIs are plotted versus different sample sizes. The first and second cases refer to the scenario in which the parameters of the ST density are fixed as $(\alpha, \beta, \kappa) = (\frac{\pi}{4}, \frac{\pi}{4}, 1)$, and $(\alpha, \beta, \kappa) = (\frac{\pi}{3}, \frac{\pi}{6}, 8)$, respectively.

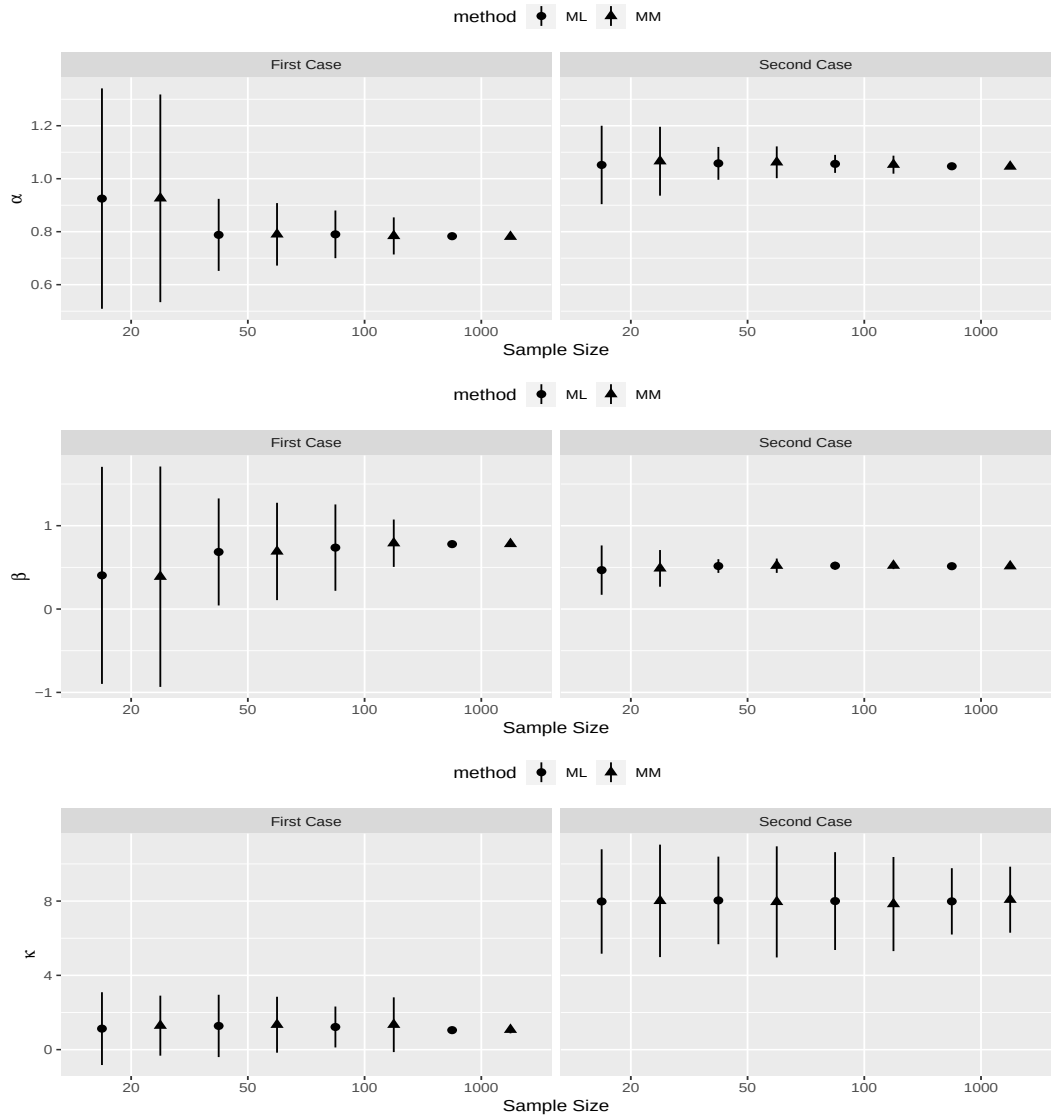

Figure 2: Comparative plots for the estimators of the SF's parameters using two estimation procedures. The approximated %95 CIs are plotted versus different sample sizes. The first and second cases refer to the scenario in which the parameters of the SF density are fixed as $(\alpha, \beta, \kappa) = (\frac{\pi}{4}, \frac{\pi}{4}, 1)$, and $(\alpha, \beta, \kappa) = (\frac{\pi}{3}, \frac{\pi}{6}, 8)$, respectively.

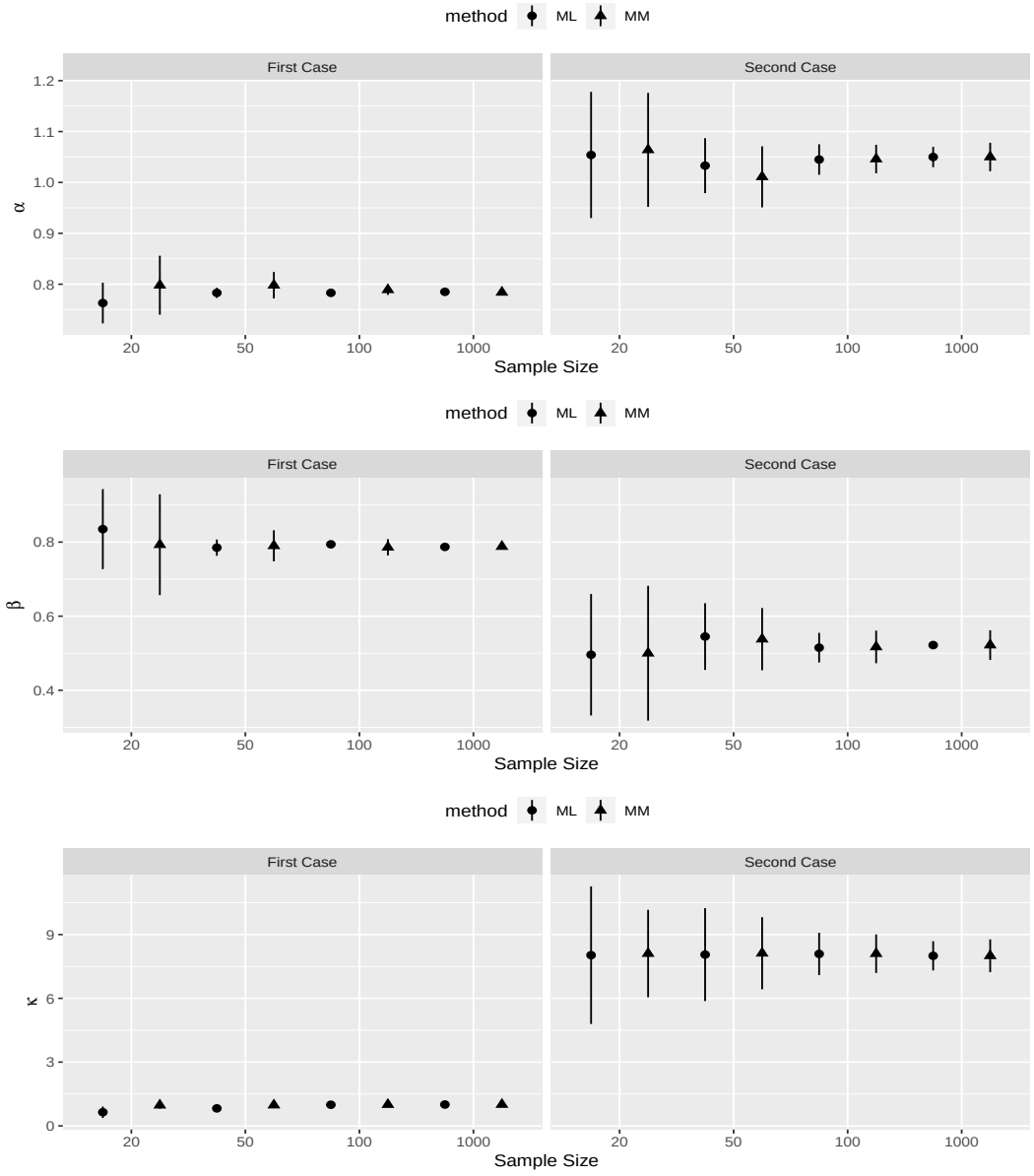

Figure 3: Comparative plots for the estimators of the SC's parameters using two estimation procedures. The approximated %95 CIs are plotted versus different sample sizes. The first and second cases refer to the scenario in which the parameters of the SC density are fixed as $(\alpha, \beta, \kappa) = (\frac{\pi}{4}, \frac{\pi}{4}, 1)$, and $(\alpha, \beta, \kappa) = (\frac{\pi}{3}, \frac{\pi}{6}, 8)$, respectively.

Based upon the discussion provided in the introduction, we expect that the SPT is performing well where the empirical plot of the data are relatively skewed. To evaluate this objective, we were interested in fitting the vMF density as well as various forms of the $\text{SPT}_3(\mu, \kappa)$ distributions where the simulated data are subject to some outliers. To do this end, we initially generated data from the vMF distribution and then contaminated them by adding a fixed value to the latitude on the North pole. Remember that longitudinal is uniformly distributed on the North pole irrespective of the type of distribution, from SPT family, is considered to simulate the data. Hence, one can only concentrated on tracing the sample path of the latitudinal variable being modeled by some stochastic processes. Hence, we now compare the aforementioned models in terms of variability imposed on the latitude. To do so, we generate data from vMF on the North pole and randomly select $5, 10, 20, 50$ percent of observations and contaminate them by adding the constant angle to the latitude. In other words, we add a fixed point to a fixed proportion, say $\Pi\%$, of the latitude angles, mathematically expressed as $\theta_{outlier} = \theta_{real} + \zeta$. Note that the concentration parameter (κ) is allowed to vary in any positive value. To this end, we tried its value at three feasible numbers to cover various amount of concentrations, i.e. low, medium and large.

Moreover, we were interested in comparing the performance of the SPT and vMF distributions under different percentage of contamination. To this end, we derived the difference between the likelihood functions computed under those underlying distributional assumption. Note that the vMF density is shifted using a pre-fixed proportion of angles to cope with skewness feature of its counterpart distribution. In particular, we first generated data from $\text{vMF}_3(\mu_0, \kappa)$ and then contaminated Π percent of their latitudes with the fixed angles ζ. At the end, we recorded the number of times (from 100 simulation runs) of when the SPT likelihood was bigger than the shifted vMF counterpart. The results are reported in Table 4. Some of the remarks, achieved from the results in this table, are as follows:

Based upon discussion above, we could take the value of ζ into account as a degree of contamination. Hence, tolerance of any well performing of the considered distribution after increasing the value of ζ will show the level of robustness. As seen, the SPT distributions are more robust in compare with the shifted vMF for the large value of ζ. This is the case for both densities come from the SPT. We also see that the SPT distribution is more robust if the concentration parameter (κ) is increased.

Table 4: Comparing the SPT and the shifted vMF distribution when the generated samples from vMF(μ_0, κ) are contaminated with Π percent on one angle. Values in each cell show the number of time (over 100 run), the logarithm of the likelihood for each of the SPT family is greater than the shifted vMF. See the text for more details.

		κ											
		5				10				50			
	Π / μ_0	5	10	20	50	5	10	20	50	5	10	20	50
ST	$\frac{\pi}{6}$	0	0	0	0	0	0	0	0	17	100	100	100
	$\frac{\pi}{4}$	0	0	0	0	0	3	90	10	100	100	100	100
	$\frac{\pi}{3}$	0	33	100	100	12	100	100	100	100	100	100	100
SF	$\frac{\pi}{6}$	0	0	0	0	0	0	0	0	0	0	5	2
	$\frac{\pi}{4}$	0	0	0	0	0	0	0	0	0	100	100	100
	$\frac{\pi}{3}$	0	0	53	95	0	0	100	100	100	100	100	100
SC	$\frac{\pi}{6}$	0	0	0	0	0	0	0	0	0	0	0	0
	$\frac{\pi}{4}$	0	0	0	0	0	0	0	0	0	7	100	100
	$\frac{\pi}{3}$	0	0	6	100	0	0	80	100	0	100	100	100

5 Application

As an application of the models studied in this paper, we consider two real-life data sets in this section. We first investigate fitting the $\mathrm{ST}_3(\mu, \kappa)$, $\mathrm{SF}_3(\mu, \kappa)$ and $\mathrm{SC}_3(\mu, \kappa)$, distributions using those data sets. Moreover, we also fit those data with four popular spherical distributions, namely vMF, Kent, Wood and Angular Central Gaussian (ACG) densities, to compare the performance of all densities. See, for example, Mardia and Jupp (2000) for more details on the properties of those latter densities. The criteria to choose a candidate density that fit our real-life data well are two common statistical measures; Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

Now, we give more details of our real-life data sets. The first data set (Data 1) deals with an sociological study of the attitudes of 48 individuals to 16 different occupations. Following Coxon

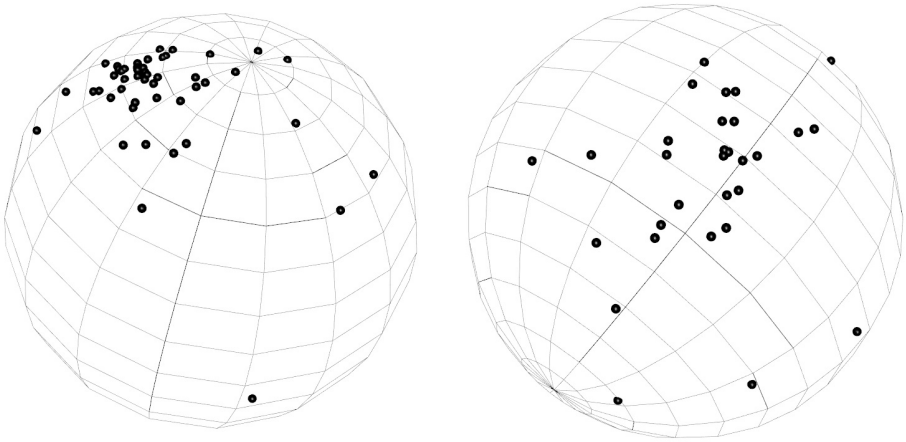

Figure 4: Schematic representation of the real data set (left: Data 1 and right: Data 2) on $\mathbb{S}^2$.

and Jones (1979), individuals were asked to make their judgments on rating or rank ordering of occupations based upon four different criteria namely Earnings, Social Status, Reward, Social Usefulness. Rating and rank ordering were transformed on $\mathbb{S}^2$ through multidimensional scaling method reported in Coxon and Jones (1978). Note that the data, registered as the unit vectors, can be found in Appendix B20 of Fisher et al. (1987).

The second data set (Data 2) is a report on the measurements of the orientation of the dendritic field at various sites in the retinas of 6 cats, in response to different visual stimuli. We used a subset of the original data where 30 responses of cats to horizontally polarized light are provided. Keilson et al. (1983) was first to analyze this data set. The spherical coordinates of this final data set can be found in Appendix B15 of Fisher et al. (1987).

A schematic representation of two real data set is shown in Figure 4. As seen, both data sets constitute some outliers, at least via detecting them with eyes in the specific directions that the spheres are plotted. In particular, a clear sign of outlying one observation is seen among the first data set. As discussed earlier, the SPT distributions are more heavy tailed than other common spherical densities. Hence, we expect that our proposed densities are better choices in modeling two data sets described above.

The values of two model choice criteria for all distributions, described at the earlier of this section, are shown in Table 5. Because the value of AIC and BIC for SPT are less than others densities, we are able to claim that the SPT outperforms its counterparts if two mentioned data sets are used as the samples to fit spherical densities. Moreover, among the family of SPT distributions,

Table 5: The values of the AIC and BIC criteria after fitting the vMF, Kent, Wood, ACG, ST, SF and SC distributions to two real data sets. See text for more details of abbreviates. The least values in each criterion and using each data set, highlighted as boldface, show superiority of the ST density among alternative distributions.

		Spherical Distribution						
Data set	Criterion	vMF	Kent	Wood	ACG	ST	SF	SC
Data 1	AIC	39.38	38.95	37.24	94.34	**3.77**	8.29	20.33
	BIC	44.99	48.30	46.60	98.08	**1.84**	13.90	25.94
Data 2	AIC	52.57	54.81	52.93	106.4	**44.97**	50.62	51.18
	BIC	56.78	61.82	59.94	109.1	**49.18**	54.82	55.39

the ST is the best candidate density to fit both data sets. These results support our initial guesses on suitability of the SPT distributions in compare with other spherical densities.

6 Discussion

It is believed that the outliers usually biase making statistical inference while fitting some non-robust models to analysis data. In the lack of possible approaches to treat outliers prior to the ultimate statistical analysis, one should seek some suitable models to analyze the entire data. To recall some heavy tailed distributions; being robust to the outliers, to fit the data is a possible approach in this case. Such phenomenon is more critical while dealing with spherical data. To provide a solution to this problem, we proposed the spherical t-distribution in this paper. Then, various statistical properties of this density have been investigated. In particular, we expressed the procedures of estimating the parameters of this density using two common statistical inference methods. We also studied some special cases of this distribution in which the data are available in the unit sphere. We illustrated the robustness of the spherical t-distribution while being compared with the popular vMF density using some simulated studies conducted in different scenarios. Moreover, we have showed superiority of our proposed distribution in compare with other alternative spherical densities while employing them to fit two real data sets.

The ST proposed in this paper belongs to the Pearson type family distribution. Following

this point, we also introduced two other densities, called SF and SC, and studied their properties. These latter densities are multimodal in some particular situations and are, indeed, robust to the outliers in compare with the vMF distributions.

According to the terminologies of the spherical distributions, the SPT distributions proposed in this paper are constructed through the conditional approach. One can study driving these densities following other methods to introduce special distributions. As another topic to research in this field, one can conduct the statistical tests on the parameters of the SPT distributions. Finally, one might prefer to impose the skewness on the sphere and then look at the asymmetric distributions based on the Pearson type family that are linked the proposed densities in this paper.

Acknowledgment

This research was in part supported by a grant from Iran National Science Foundation (INSF) [no. 95014574].

References

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York.

Coxon, A. P. and Jones, C. L. (1978). *Images of Occupational Prestige*. Academic Press, London.

Coxon, A. P. M. and Jones, C. L. (1979). *Measurement and Meanings*. Academic Press, London.

Fisher, N. I., Lewis, T., and Embleton, B. J. (1987). *Statistical Analysis of Spherical Data*. Cambridge University Press, Cambridge.

Johnson, N. L., Kotz, S., and Narayanaswamy, B. (1995). *Continuous Univariate Distributions*. John Wiley & Sons, New York.

Kato, S. and Shimizu, K. (2004). *A Further Study of t-Distributions on Spheres*. Technical Report, School of Fundamental Science and Technology, Keio University, Yokohama.

Keilson, J., Petrondas, D., Sumita, U., and Wellner, J. (1983). Significance points for some tests of uniformity on the sphere. *Journal of Statistical Computation and Simulation*, 17(3):195–218.

Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*. John Wiley & Sons, London.

Pearson, K. (1895). Contributions to the mathematical theory of evolution, II: Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London*, 186(Part I):343–424.

Pearson, K. (1916). Mathematical contributions to the theory of evolution, XIX: Second supplement to a memoir on skew variation. *Philosophical Transactions of the Royal Society A*, 216(538–548):429–457–424.

Pewsey, A., Lewis, T., and Jones, M. (2007). The wrapped t family of circular distributions. *Australian & New Zealand Journal of Statistics*, 49(1):79–91.

Pewsey, A., Neuhäuser, M., and Ruxton, G. D. (2013). *Circular Statistics in R*. Oxford University Press, Oxford.

Shimizu, K. and Iida, K. (2002). Pearson type VII distributions on spheres. *Communications in Statistics-Theory and Methods*, 31(4):513–526.