Construction and parameter estimation of wrapped normal models

by

Hannaline Roux

Submitted in partial fulfillment of the requirements for the degree
Magister Scientiae (Mathematical Statistics)
in the Faculty of Natural and Agricultural Sciences
University of Pretoria, Pretoria

August 2019

Publication data:

Hannaline Roux. Construction and parameter estimation of wrapped normal models. Master's dissertation, University of Pretoria, Department of Statistics, Pretoria, South Africa, August 2019.

Construction and parameter estimation of wrapped normal models

by

Hannaline Roux
E-mail: hannalineroux@gmail.com

Abstract

If a known distribution on a real line is given, it can be wrapped on the circumference of a unit circle. This research entails the study of a univariate skew-normal distribution where the skew-normal distribution is generalised for the case of bimodality. Both the skew-normal and flexible generalised skew-normal distributions are wrapped onto a unit circle, consequently referred to as a wrapped skew-normal and a wrapped flexible generalised skew-normal distribution respectively. For each of these distributions a simulation study is conducted, where the performance of maximum likelihood estimation is evaluated. Skew scale mixtures of normal distributions with the wrapped version of these distributions are proposed and graphical representations are provided. These distributions are also compared in an application to wind direction data.

Keywords: Wrapped skew-normal, Wrapped flexible generalised skew-normal, Trigonometric moments, Maximum likelihood estimation, Skew scale mixtures of normal model.

Supervisors : Dr. M. T. Loots
Prof. A. Bekker
Department : Statistics
Degree : Magister Scientiae

Declaration

I, Hannaline Roux, declare that the mini-dissertation, which I hereby submit for the degree MSc in Mathematical Statistics at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

.....
Hannaline Roux

.....
Dr MT Loots

.....
Prof A Bekker

.....
Date

Acknowledgements

A very special thank you to my supervisors, Dr Theodor Loots and Professor Andriëtte Bekker for their patient guidance and support. In particular I would also like to thank Dr Mehrdad Naderi for his suggestions and responses to all my questions and queries; your support and encouragement have been invaluable. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF. Lastly, I would like to thank the Statomet Internship for their financial support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Circular data can be described by data related to directions, which include a variety of research fields, such as: earth sciences, biology, medicine and psychology. Several statistical models have been proposed for analysing circular data by means of wrapping the distribution around a unit circle. The term wrapping refers to taking data on a real line and wrapping it around a unit circle. The most popular and frequently used distribution is the normal distribution, which is known for its simplicity. It is therefore of interest to consider the wrapped normal (WN) distribution along with distributions that can incorporate skewness and bimodality, such as the wrapped skew-normal (WSN) by Pewsey [25] and the wrapped flexible generalised skew-normal (WFGSN) distribution by Hernández-Sánchez and Scarpa [12]. Data applied in practice are often asymmetric and bimodal which also motivates the study of the WSN and WFGSN distribution.
The scale mixtures of normal (SMN) distributions present a group of heavy tailed distributions that are frequently used as a statistical procedure for symmetrical data. However, since theory and application provide a large amount of directional data that are skewed with heavy tails, it motivates the study of the skew scale mixtures of normal (SSMN) distributions as well as the wrapped version of these distributions. It is therefore important to consider the characteristics of each of the mentioned distributions to compare how well these models will perform not only in theory, but also in application.

1.2 Literature review

Mardia and Jupp [19] defined circular data and proposed models and statistical methodology for analysing circular data, such as wrapped distributions in general as well as the WN distribution. The characteristic function (CF) of a circular random variable is also defined. Azzalini [3] introduced the skew-normal (SN) which includes the standard normal distribution as a special case. He also defined the skewing methodology used to skew existing symmetric probability distribution functions. Pewsey [25] proposed the WSN for circular data where centred parameterisation of the distribution is introduced as well as the relationship between the parameters and the direct ones. Methods of moments estimation is also considered. A flexible class of skew-symmetric distributions have been proposed by Ma and Genton [18], where these distributions can capture skewness and bimodality. The probability distribution function (PDF) has the form of a product of the skewing mechanism as well as the PDF of a symmetric distribution. They also discussed and illustrated the flexible generalised skew-normal (FGSN) distribution as an example of a distribution that has shape flexibility and multimodality. The WFGSN distribution is then proposed by Hernández-Sánchez and Scarpa [12] where the parameters of the proposed model are estimated by maximum likelihood estimation (MLE). They also concluded that the WFGSN distribution outperforms the WSN distribution. Da Silva Ferreira et al. [8] defined the SMN distributions as well as the SSMN distributions. They also discussed the skew-Student-t normal (StN), the skew-slash (SSL) and the skew-contaminated normal (SCN) distribution. Several probabilistic and inferential properties for these distributions are also defined.

1.3 Objectives

The aim of this study is to:

- Review and revisit the concept of circular data discussed by Mardia and Jupp [19].
- Understand the term wrapping by investigating the WN distribution.
- Investigate Azzalini's SN distribution [3].
- Investigate the WSN distribution proposed by Pewsey [25].
- Compare the method of MLE and trigonometric moments with a simulation study.
- Understand the FGSN distribution by Ma and Genton [18] as well as the WFGSN distribution.
- Investigate examples of SSMN distributions by Da Silva Ferreira et al. [8] and consider the proposed wrapped versions thereof.
- Apply these wrapped distributions to a data set for comparison purposes.

1.4 Outline of study

- In Chapter 2 a wrapped distribution is revisited and the WN distribution is also discussed. The SN distribution with a stochastic representation is revisited as well as the direct parameterisation as a basis for estimation is discussed. The WSN distribution is revisited where the CF with the trigonometric moments are also investigated with a Monte Carlo approximation for the trigonometric moments. A simulation study is conducted to compare MLE and the method of trigonometric moments.

- In Chapter 3 the FGSN distribution is revisited and the WFGSN distribution with examples are presented. The method of MLE is discussed and a simulation study is conducted. These distributions are fitted to a data set for comparison purposes.

- In Chapter 4 the SMN distributions are revisited as well as the SSMN distributions. Examples of the SSMN distributions are provided, such as the StN, SSL and the SCN distributions. Lastly, the wrapped versions of these distributions are compared to the WN, WSN and WFGSN distributions to investigate if a better fit can be obtained for the data set.

The above outline is summarised in Figure 1.1.

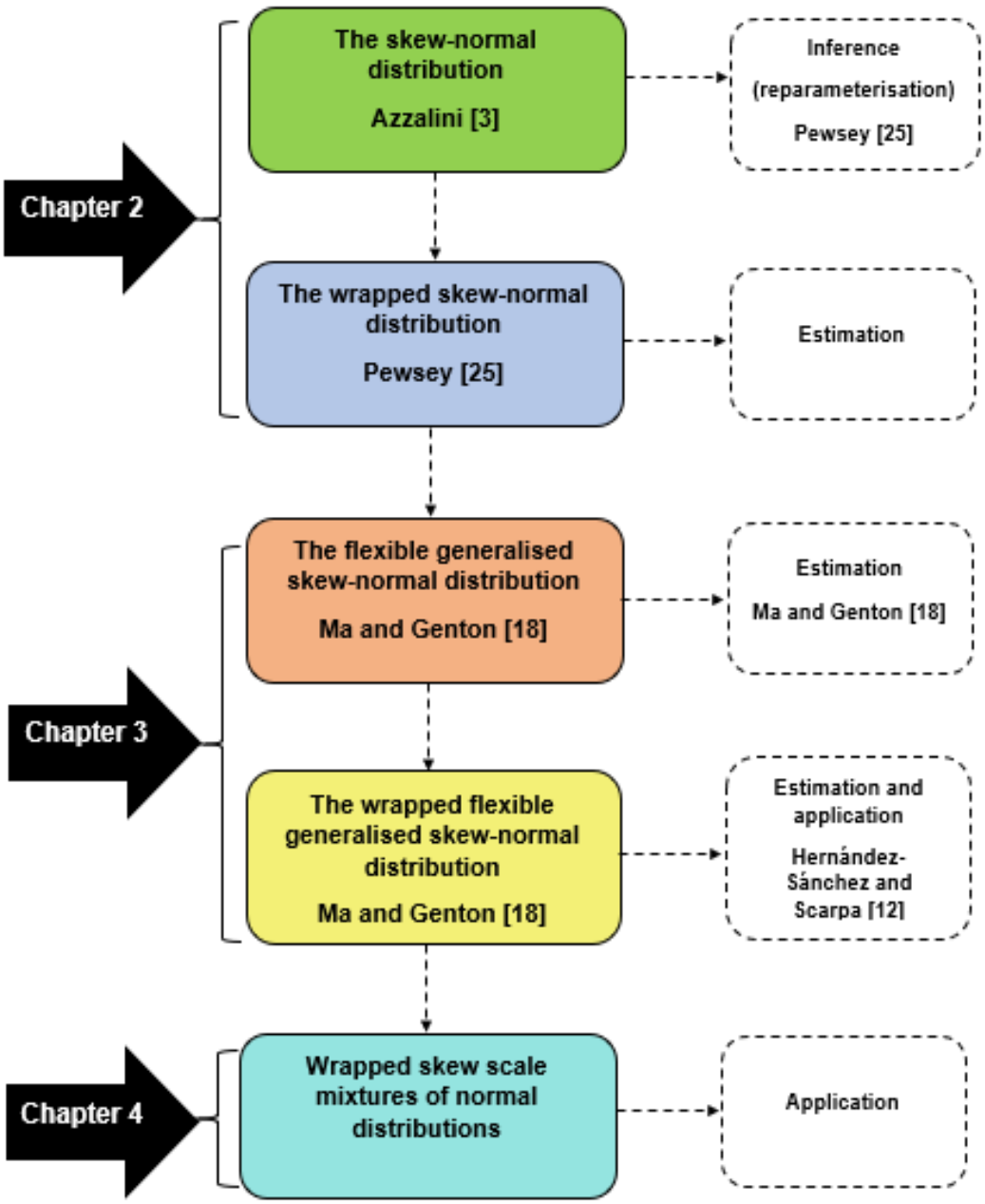


Figure 1.1: Outline of study.

- Chapter 5 concludes the study.
- Appendix A contains a list of additional results and definitions referenced in this

study.

- Appendix B contains code used in this study.
- Appendix C contains a list of acronyms and symbols used throughout the study as well as an index.

Chapter 2

The skew-normal and wrapped skew-normal distributions

In Section 2.1, the PDF of a wrapped distribution is defined and the WN distribution is also discussed. The SN distribution with a stochastic representation is defined in Section 2.2. In Section 2.3 the direct parameterisation as a basis for estimation is discussed where in Section 2.4 the estimation of the centred parameterisation is investigated. The WSN distribution is defined with representations of different parameter values in Section 2.5. The CF with the trigonometric moments are also investigated with a Monte Carlo approximation for the trigonometric moments. In Section 2.6, the method of MLE and the method of trigonometric moments are discussed where a simulation study is then conducted to compare the two methods of estimation. The above outline is summarised in Figure 2.1.

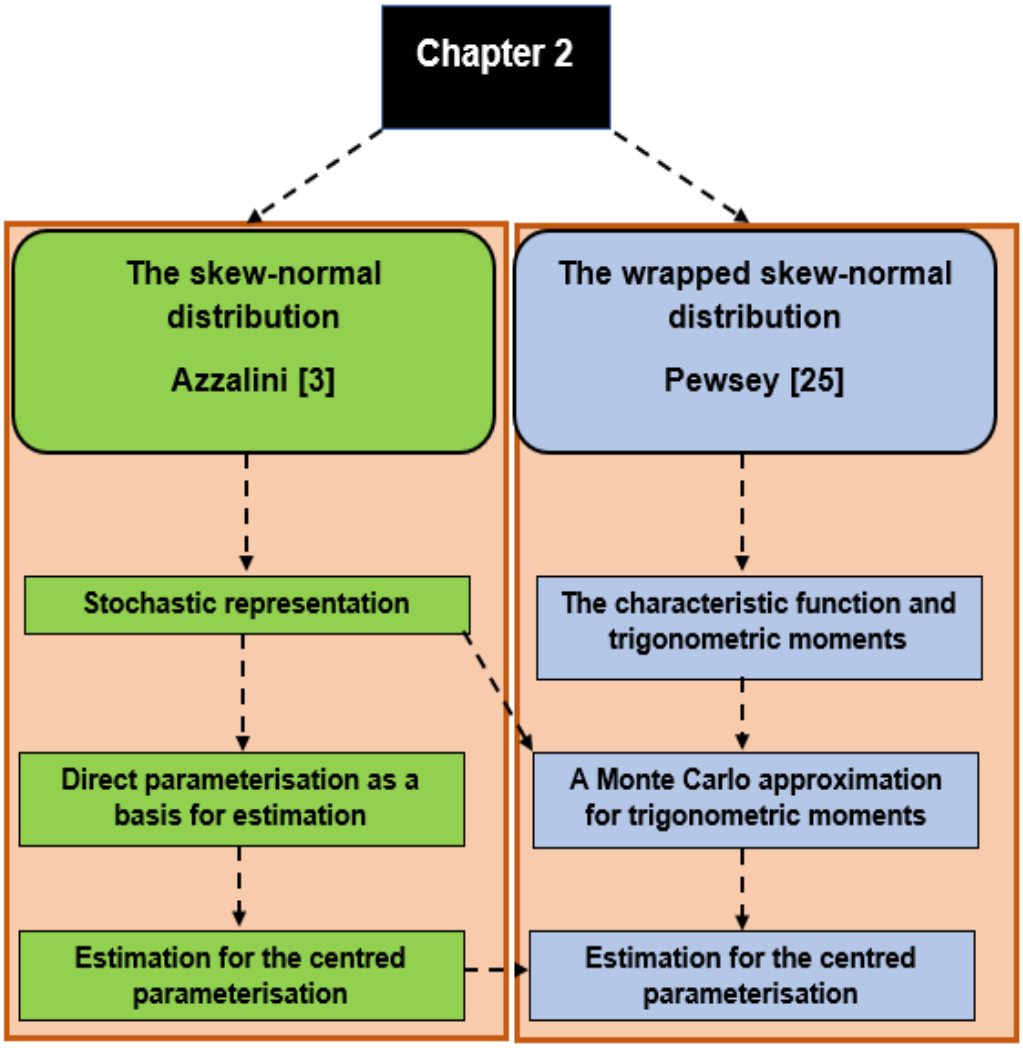


Figure 2.1: Outline of Chapter 2.

2.1 Background

The term wrapping refers to taking data on a real line and wrapping it around a unit circle. This type of data can be referred to as circular data [19]. Circular data can also be described by data related to directions, which include a variety of research fields, such as: earth sciences, biology, medicine and psychology [19]. If X is a linear random variable, with a PDF $g(\cdot)$, then the corresponding 'wrapped' circular version of X can be defined as

$$\theta = X(mod 2\pi).$$

The random variable θ has the following PDF

$$f(\theta) = \sum_{k=-\infty}^{\infty} g(\theta + 2\pi(k)), \tag{2.1}$$

obtained by 'wrapping' g around the unit circle [17]. The distribution function of θ can also be defined as

$$F(\theta) = \sum_{k=-\infty}^{\infty} [g(\theta + 2\pi(k)) - g(2\pi(k))], \quad k = 0, \pm 1, \pm 2, \pm 3, ...,$$

where probability is accumulated over all the overlapping points $x = \theta, \theta \pm 2\theta, \theta \pm 4\theta...$, using the above approach [15].

2.1.1 Wrapped normal (WN) distribution

One of the most frequently used distributions is the normal distribution which is known for its simplicity and the fact that it forms a basis for various statistical techniques. Examples include measurement errors in scientific experiments, scores on various tests and numerous economic measures and indicators [9]. Therefore, by wrapping the normal distribution, $X \sim N(\mu, \sigma^2)$, around the unit circle a WN distribution $\theta \sim WN(\mu, \rho)$ is obtained, where

$$\rho = \exp(\frac{-\sigma^2}{2}),$$

i.e.

$$\sigma^2 = -2\log\rho.$$

If $X \sim N(\mu, \sigma^2)$ then the PDF of the normal distribution is defined as:

$$g(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\exp(\frac{-(x-\mu)^2}{2\sigma^2}), \quad -\infty < x < \infty.$$

The PDF of $\theta = X(mod 2\pi)$, i.e. the PDF of the WN distribution is

$$\phi(\theta;\mu,\sigma^2) = \sum_{k=-\infty}^{\infty} g(\theta + 2\pi(k)) = \frac{1}{\sigma\sqrt{2\pi}}\sum_{k=-\infty}^{\infty} exp(\frac{-(\theta-\mu+2\pi(k))^2}{2\sigma^2}). \tag{2.2}$$

The CF of X is

$$\begin{aligned}\psi(t) &= E[\exp(itX)] \\ &= \exp(i\mu t - \frac{t^2\sigma^2}{2}).\end{aligned} \tag{2.3}$$

It then follows from Theorem 10, Appendix A.2 and Definition 8, Appendix A.1 that the CF of a circular random variable $\Theta = X(mod 2\pi)$ is

$$\psi(p) = \exp(i\mu p - \frac{p^2\sigma^2}{2}), \quad p = 0, \pm1, \pm2, ...,$$

so that

$$\begin{aligned}E(\exp(ip\Theta)) &= \exp(-\frac{p^2\sigma^2}{2})\exp(i\mu p) \\ &= \alpha_p + i\beta_p\end{aligned}$$

where the cosine moment is defined as

$$\alpha_p = \exp(-\frac{p^2\sigma^2}{2})\cos(p\mu), \tag{2.4}$$

and the sine moment

$$\beta_p = \exp(-\frac{p^2\sigma^2}{2})\sin(p\mu), \tag{2.5}$$

in which μ refers to the mean direction $(mod 2\pi)$, ρ refers to the mean resultant length (the length of the average of random vectors on the unit circle - see Definition 6, Appendix A.1) and σ refers to the standard deviation of the unwrapped normal distribution [19].

2.2 The skew-normal (SN) distribution

Suppose a random variable X has a standard SN distribution with a skewness parameter λ where $X \sim SN(\lambda)$ [3, 25]. Under this framework, the PDF is as follows

$$g(x;\lambda) = 2\phi(x)\Phi(\lambda(x)), \quad -\infty < x < \infty, -\infty < \lambda < \infty. \tag{2.6}$$

The SN distribution, introduced by Azzalini [3], includes the standard normal distribution as a special case. The skewing methodology used to skew existing symmetric probability distribution functions can be defined as follows

$$f(x) = 2f_0(x)G(w(x)), \tag{2.7}$$

where $f(x)$ is the PDF for any odd function w, where $w(x) = \lambda x$ and $f_0 = \phi$. This PDF holds for any symmetric PDF f_0 and distribution function G [3, 12]. $2G_0(w(x))$ can be referred to as the skewing mechanism [3]. The class can also be generalised by including ξ and ω^2 as the location and scale parameters [12, 25]. Hence, if $X \sim SN(\lambda)$ then

$$Y = \xi + \omega X, \tag{2.8}$$

is a univariate SN distribution with the following PDF [3, 25]:

$$f(y;\xi,\omega^2,\lambda) = \frac{2}{\omega}\phi(\frac{y-\xi}{\omega})\Phi(\lambda(\frac{y-\xi}{\omega})), \tag{2.9}$$

where $-\infty < y < \infty, -\infty < \lambda < \infty, -\infty < \xi < \infty, \omega^2 \in \mathbb{R}^+$. ξ, ω^2 and λ can be referred to as the direct parameters and denote the distribution of Y as $SN(\xi,\omega^2,\lambda)$ [3, 25].

Figure 2.2 shows an overlay of the SN PDF in Equation 2.6 with $\xi = 0, \omega^2 = 1$ and combinations of λ.

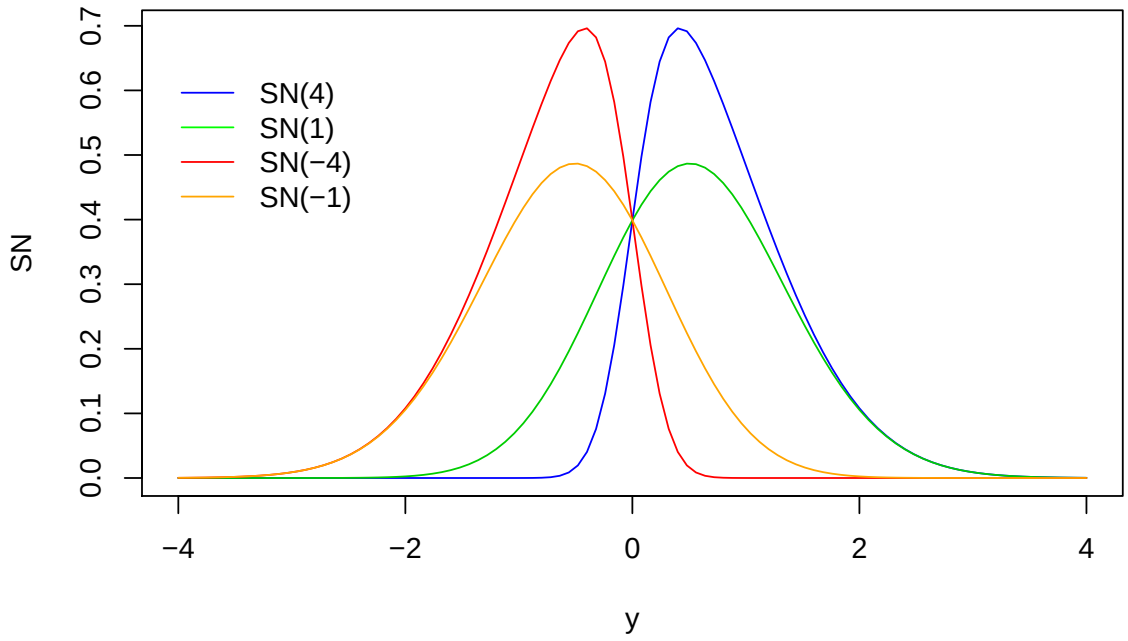


Figure 2.2: An overlay of the SN PDF (2.6) with $\xi = 0, \omega^2 = 1$ and combinations of λ.

2.2.1 Stochastic representation of the SN distribution

Following the approach of Azzalini [3] and Mastrantonio et al. [20], a stochastic representation is revisited that is useful for generating random numbers from a $SN(\xi, \omega^2, \lambda)$ distribution. This provides a method to generate random numbers from the $Y \sim SN(\xi, \omega^2, \lambda)$ with PDF (2.9).

Mastrantonio et al. [20] define U and W as two independent standard normal variables, where $\omega^2 \in \mathbb{R}^+$ and $-\infty < \lambda < \infty$, then

$$Y = \xi + \frac{\omega\lambda}{\sqrt{1+\lambda^2}}|U| + \frac{\omega}{\sqrt{1+\lambda^2}}W - \frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}}. \tag{2.10}$$

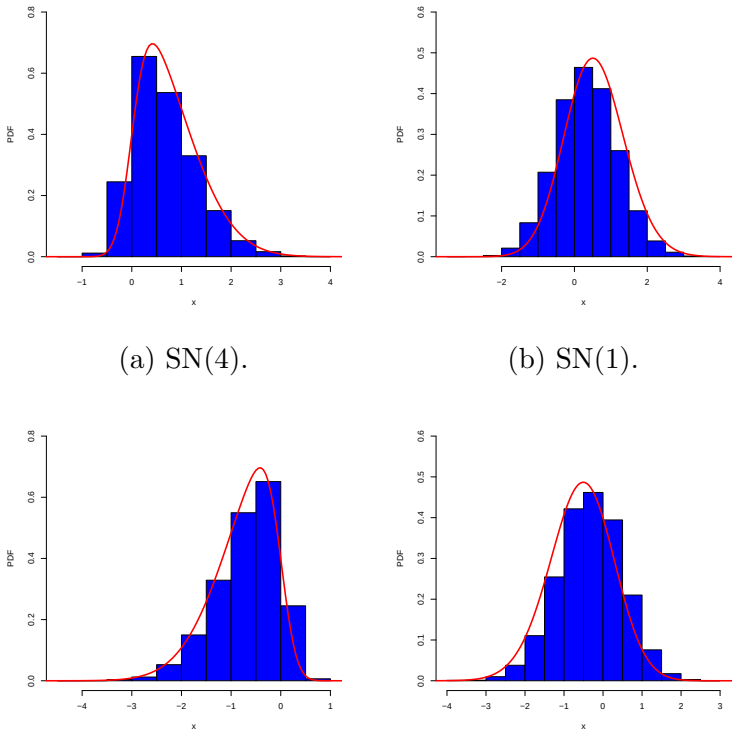
The PDF of Y is defined as follows:

$$f(y;\xi,\omega^2,\lambda) = \frac{2}{\omega}\phi\Big(\frac{y-\xi+\frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}}}{\omega}\Big)\Phi\Big(\lambda\Big(\frac{y-\xi+\frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}}}{\omega}\Big)\Big),$$

where the mean and variance of Y are given as ξ and $\omega^2\lambda^2/(1+\lambda^2)(1-2/\pi)+\omega^2/(1+\lambda^2)$ respectively [20].

Software is already available to generate normal distributed random variables, therefore Section 2.2.1 provides a representation to easily generate random numbers from a SN distribution.

2.2.2 Visualisation of the SN sampling scheme in Section 2.2.1



(a) SN(4). (b) SN(1).

(c) SN(-4). (d) SN(-1).

Figure 2.3: Histograms of random samples of size 10000 taken from SN densities with an overlay of the corresponding theoretical PDF (2.6), for $\xi = 0, \omega^2 = 1$ and different combinations of λ.

Figure 2.3 represent histograms of random samples taken from $Y \sim SN(\xi, \omega^2, \lambda)$ using the stochastic representation in Equation 2.10 with an overlay of the corresponding theoretical PDF (2.6).

2.3 Direct parameterisation as a basis for estimation

In this section, a short overview of MLE as a method of estimation is presented, where the PDF (2.9) in Section 2.2 is used to obtain the likelihood function. The Fisher information matrix is also presented which form part of the asymptotic theory of MLE.

2.3.1 Maximum likelihood estimation (MLE)

In order to use maximum likelihood (ML) as a method of estimation, let $y_1, ..., y_n$ be a random sample of size n from $SN(\xi, \omega^2, \lambda)$ (2.9). The likelihood function of the parameters of the $SN(\xi, \omega^2, \lambda)$ distribution is given by

$$\ell(\xi, \omega^2, \lambda) = (\frac{2}{\omega})^n \prod_{i=1}^{n} \phi(\frac{y_i - \xi}{\omega}) \Phi(\lambda(\frac{y_i - \xi}{\omega})). \tag{2.11}$$

Fisher information matrix

Azzalini [3] gives the Fisher information matrix for the direct parameterisation, generated by the linear transformation represented in Section 2.2, where

$$Y = \xi + \omega X,$$

for $\omega > 0$. The Fisher information for the parameters (ξ, ω^2, λ) is as follows, where $X \sim SN(\lambda)$

$$I_\lambda = \begin{bmatrix} \frac{(1+\lambda^2 a_0)}{\omega^2} & \frac{E(X)\frac{1+2\lambda^2}{1+\lambda^2} + \lambda^2 a_1}{\omega^2} & \frac{\frac{b}{(1+\lambda^2)^{3/2}} - \lambda a_1}{\omega} \\ \frac{E(X)\frac{1+2\lambda^2}{1+\lambda^2} + \lambda^2 a_1}{\omega^2} & \frac{(2+\lambda^2 a_2)}{\omega^2} & \frac{-\lambda a_2}{\omega} \\ \frac{\frac{b}{(1+\lambda^2)^{3/2}} - \lambda a_1}{\omega} & \frac{-\lambda a_2}{\omega} & a_2 \end{bmatrix}$$

where

$$a_k = a_k(\lambda) = E(X^k (\frac{\phi(\lambda X)}{\Phi(\lambda X)})^2), \quad k = 0, 1, 2.$$

The above matrix becomes singular as $\lambda \to 0$. Therefore, in order to attempt the maximisation of the log-likelihood function from Equation 2.11 for this parameterisation using numerical techniques, no unique solution exists and the results could be highly misleading [24].

Thus, the findings for MLE, according to Azzalini [3] and Pewsey [24], shows that the direct parameterisation cannot be used as a general basis for estimation.

2.4 Estimation for the centred parameterisation

It is possible to parameterise the SN distribution by using the mean, variance and skewness index by means of centred parameterisation [3]. Centred parameterisation can also overcome some estimation problems in certain scenarios, which is caused by direct parameterisation [3]. Azzalini [3] introduces the centred parameterisation with parameters, μ, σ and γ_1 where Pewsey [26] then defines a skew-normal random variable Y_C as

$$Y_C = \mu + \frac{\sigma(X - E(X))}{\sqrt{var(X)}}, \quad -\infty < \mu < \infty, \quad \sigma > 0,$$

where $X \sim SN(\lambda), E(Y_C) = \mu$ and $Var(Y_C) = \sigma^2$ [3, 24]. The subscript "C" refers to the centred parameterisation. Since $E(Y_C) = \mu$, this parameterisation is not parameter redundant for the normal case [24]. The parameter γ_1, denotes the coefficient of skewness of X, therefore also that of Y_C. The distribution of Y_C can therefore be denoted as $SN_C(\mu, \sigma^2, \gamma_1)$ [3].

According to Pewsey [26], the direct parameters can also be written in terms of the centred parameters, where

$$\begin{aligned} \xi &= \mu - c\gamma_1{}^{\frac{1}{3}}\sigma, \\ \omega &= \sigma\sqrt{1 + c^2\gamma_1{}^{\frac{2}{3}}}, \\ \lambda &= \frac{c\gamma_1{}^{\frac{1}{3}}}{\sqrt{\frac{2}{\pi} + c^2(\frac{2}{\pi} - 1)\gamma_1{}^{\frac{2}{3}}}}, \end{aligned} \tag{2.12}$$

and $c = (2/(4-\pi))^{\frac{1}{3}}$. Using the relations (2.12) in Equation 2.9, the PDF of Y_C is given by

$$
\begin{aligned}
f(y;\mu,\sigma^2,\gamma_1) &= \frac{2}{\sigma\sqrt{1+c^2{\gamma_1}^{\frac{2}{3}}}}\phi\Big[\frac{y-\mu+c{\gamma_1}^{\frac{1}{3}}\sigma}{\sigma\sqrt{1+c^2{\gamma_1}^{\frac{2}{3}}}}\Big] \\
&\times \Phi\Big[\frac{c{\gamma_1}^{\frac{1}{3}}}{\sqrt{\frac{2}{\pi}+c^2(\frac{2}{\pi}-1){\gamma_1}^{\frac{2}{3}}}}\Big(\frac{y-\mu+c{\gamma_1}^{\frac{1}{3}}\sigma}{\sigma\sqrt{1+c^2{\gamma_1}^{\frac{2}{3}}}}\Big)\Big] \\
&= \frac{2}{\sigma\sqrt{1+c^2{\gamma_1}^{\frac{2}{3}}}}\phi[\frac{1}{\sqrt{1+c^2{\gamma_1}^{\frac{2}{3}}}}\{(\frac{y-\mu}{\sigma})+c{\gamma_1}^{\frac{1}{3}}\}] \\
&\times \Phi[\frac{c{\gamma_1}^{\frac{1}{3}}}{\sqrt{\frac{2}{\pi}+c^2(\frac{2}{\pi}-1){\gamma_1}^{\frac{2}{3}}(1+c^2{\gamma_1}^{\frac{2}{3}})}}\{(\frac{y-\mu}{\sigma})+c{\gamma_1}^{\frac{1}{3}}\}]. \qquad (2.13)
\end{aligned}
$$

2.4.1 Maximum likelihood estimation (MLE)

From Equation 2.13 the log-likelihood function for a random sample $y_1,...,y_n$, of size n, from $Y_C \sim SN_C(\mu,\sigma^2,\gamma_1)$, is given by

$$
\begin{aligned}
\ell(\mu,\sigma^2,\gamma_1) &= n\log 2 - n\log\sigma - \frac{n}{2}\log(1+c^2{\gamma_1}^{\frac{2}{3}}) + \sum_{i=1}^{n}\log[\phi[\frac{1}{\sqrt{1+c^2{\gamma_1}^{\frac{2}{3}}}}\{(\frac{y_i-\mu}{\sigma})+c{\gamma_1}^{\frac{1}{3}}\}]] \\
&\times \Phi[\frac{c{\gamma_1}^{\frac{1}{3}}}{\sqrt{\frac{2}{\pi}+c^2(\frac{2}{\pi}-1){\gamma_1}^{\frac{2}{3}}(1+c^2{\gamma_1}^{\frac{2}{3}})}}\{(\frac{y_i-\mu}{\sigma})+c{\gamma_1}^{\frac{1}{3}}\}], \qquad (2.14)
\end{aligned}
$$

where $-\infty < \mu < \infty$ and $-0.99527 < \gamma_1 < 0.99527$ [26]. In order to maximise the log-likelihood function in Equation 2.14, numerical optimisation techniques should be used. Pewsey [26] recommend the simplex algorithm of Nelder and Mead [21].

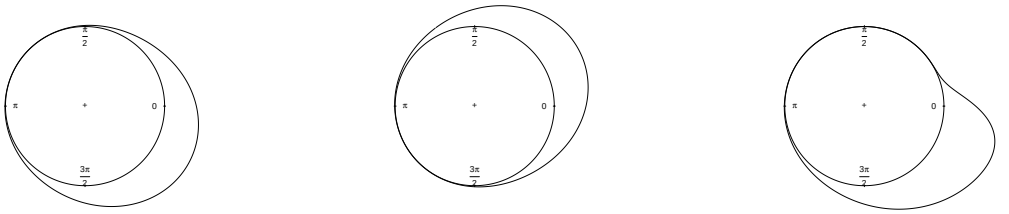
2.5 Wrapped skew-normal (WSN) distribution

In order to use the SN distribution, $Y \sim SN(\xi,\omega^2,\lambda)$ for circular data, Pewsey [25] wrapped Y onto a unit circle where a circular random variable can be defined as $\Theta = Y(mod 2\pi)$, which has the following PDF

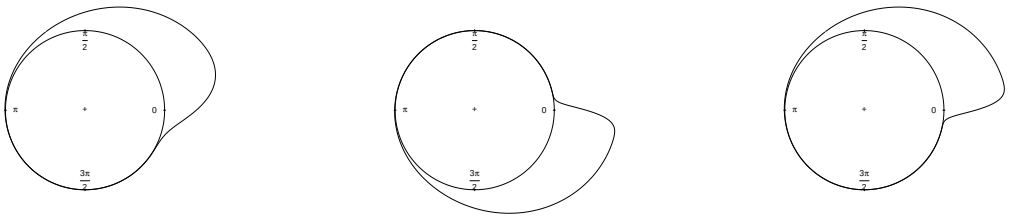
$$
f(\theta;\xi,\omega^2,\lambda) = \frac{2}{\omega}\sum_{r=-\infty}^{\infty}\phi(\frac{\theta+2\pi r-\xi}{\omega})\Phi(\lambda(\frac{\theta+2\pi r-\xi}{\omega})), \qquad (2.15)
$$

for $0 \leq \theta \leq 2\pi$ [25]. The WSN distribution can therefore be denoted as $\Theta \sim WSN(\xi, \omega^2, \lambda)$ [25].

Figure 2.4 shows the shape of the PDF given in Equation 2.15 for $\xi = 0$, $\omega = 1$ and various combinations of λ. The command 'curve.circular' in the R package 'circular' [1] was used to create the circular plots.



(a) WSN(0,1,-1). (b) WSN(0,1,1). (c) WSN(0,1,-5).



(d) WSN(0,1,5). (e) WSN(0,1,-15). (f) WSN(0,1,15).

Figure 2.4: Examples of the WSN PDF, given in Equation 2.15, with $\xi = 0, \omega = 1$ and various combinations of λ.

2.5.1 The characteristic function (CF) and trigonometric moments

In this section the CF of $\Theta \sim WSN(\xi, \omega^2, \lambda)$, by Pewsey [25] is derived, but firstly the focus is on the CF of Y where $Y \sim SN(\xi, \omega^2, \lambda)$.

Theorem 1 *The CF of the SN distribution, $\psi_Y(t)$ where $Y \sim SN(\xi, \omega^2, \lambda)$, by Azzalini [4] is given by:*

$$(1) \quad \psi_Y(t) = 2\exp(i\xi t - \omega^2 t^2/2)\Phi(i\delta\omega t). \tag{2.16}$$

Alternatively, the CF of the SN distribution by Pewsey [25], is given by:

$$(2) \quad \psi_Y(t) = \exp(i\xi t - \omega^2 t^2/2)\{1 + i\tau(\delta\omega t)\}, \tag{2.17}$$

where $\delta = \lambda/\sqrt{1+\lambda^2}$ and

$$\tau(y) = \int_0^y \sqrt{2/\pi}\exp(u^2/2)du, \quad y > 0,$$

and $\tau(-y) = -\tau(y)$.

Proof. (1) From Azzalini [4] (see Theorem 9, Appendix A.2) and $X \sim SN(0, 1, \lambda)$ (Equation 2.6 and 2.8), it follows that

$$\begin{aligned}
\psi_Y(t) &= M_Y(it) \\
&= E[\exp(itY)] \\
&= E[\exp(i\xi t + i\omega x t)] \\
&= \int_{\mathbb{R}} \exp(i\xi t + i\omega x t) 2\phi(x)\Phi(\lambda x)dx \\
&= 2\exp(i\xi t)\int_{\mathbb{R}} \exp(i\omega x t)\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)\Phi(\lambda x)dx.
\end{aligned}$$

Note that $(x - i\omega t)^2 = x^2 - 2i\omega xt - \omega^2 t^2 \Rightarrow x^2 - 2i\omega xt = (x - i\omega t)^2 + \omega^2 t^2$.
Therefore,

$$\begin{aligned}
\psi_Y(t) &= 2\exp(i\xi t)\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{1}{2}(x^2 - 2i\omega xt)\Big)\Phi(\lambda x)dx \\
&= 2\exp(i\xi t)\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{1}{2}((x - i\omega t)^2 + \omega^2 t^2)\Big)\Phi(\lambda x)dx \\
&= 2\exp(i\xi t - \omega^2 t^2/2)\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{1}{2}(x - i\omega t)^2\Big)\Phi(\lambda x)dx \\
&= 2\exp(i\xi t - \omega^2 t^2/2)\int_{\mathbb{R}} \phi(x - i\omega t)\Phi(\lambda x)dx^{(1)} \\
&= 2\exp(i\xi t - \omega^2 t^2/2)\Phi\Big(\frac{\lambda\omega t}{\sqrt{1+\lambda^2}}\Big) \\
&= 2\exp(i\xi t - \omega^2 t^2/2)\Phi(\delta\omega t) \\
&\equiv M_Y(it), \qquad (2.18)
\end{aligned}$$

where $\delta = \lambda/\sqrt{1+\lambda^2}$.
(1) Applying Theorem 9, Appendix A.2.

Proof. (2) Applying Definition 10, Appendix A.1 it follows that

$$\begin{aligned}
M_Y(it) &= 2\exp(i\xi t - \omega^2 t^2/2)\Phi(\delta\omega t) \\
&= 2\exp(i\xi t - \omega^2 t^2/2)\Big(\frac{1}{2} + \int_0^{\delta\omega t}\phi(x)dx\Big),\ t \in \mathbb{R}.
\end{aligned}$$

By applying Azzalini [4], let γ be the line segment linking 0 and $\delta\omega ti$, namely, γ consists of points $x = xi$ where x takes values from 0 to $\delta\omega t$. Then, the CF is given by

$$\begin{aligned}
\psi_Y(t) &= M_Y(it) \\
&= 2\exp(i\xi t - \omega^2 t^2/2)\Big(\frac{1}{2} + \int_0^{\delta\omega t}\phi(x)dx\Big) \\
&= \exp(i\xi t - \omega^2 t^2/2)\Big(1 + 2i\int_0^{\delta\omega t}\frac{1}{\sqrt{2\pi}}\exp(\frac{x^2}{2})dx\Big)^{(1)} \\
&= \psi_Y(t) \\
&= \exp(i\xi t - \omega^2 t^2/2)\{1 + i\tau(\delta\omega t)\}. \qquad (2.19)
\end{aligned}$$

(1) Applying Theorem 8, Appendix A.2.

□

Similar for the WSN distribution, the CF by Pewsey [25] is derived in the following theorem.

Theorem 2 *The CF of* $\Theta \sim WSN(\xi, \omega^2, \lambda)$, ψ_p, *is given by*

$$\psi_p = E[\exp(ip\Theta)] = \psi_Y(p) = \exp(i\xi p - \omega^2 p^2/2)\{1 + i\tau(\delta\omega p)\},$$

for $p = 0, 1, \ldots$.

Proof. Let $\Theta = Y(mod 2\pi)$, then from Theorem 1 it follows that the CF of Θ is given by:

$$\begin{aligned} \psi_p &= E[\exp(ip\Theta)] \\ &\equiv \psi_Y(p)^{(1)} \\ &= \exp(i\xi p - \omega^2 p^2/2)\{1 + i\tau(\delta\omega p)\}, \end{aligned} \tag{2.20}$$

for $p = 0, 1, 2, \ldots$.
(1) Applying Theorem 10, Appendix A.2.

□

The cosine and sine moments, defined by Pewsey [25] as α_p and β_p, are derived in the following theorem.

Theorem 3 *The cosine and sine moments,* α_p *and* β_p, *for* $\Theta \sim WSN(\xi, \omega^2, \lambda)$ *are given by*

$$(i) \quad \alpha_p \equiv E[\cos p\Theta] = \exp(-\omega^2 p^2/2)(\cos p\xi - \tau(\delta\omega p)\sin p\xi),$$

$$(ii) \quad \beta_p \equiv E[\sin p\Theta] = exp(-\omega^2 p^2/2)(\sin p\xi + \tau(\delta\omega p)cosp\xi).$$

Proof. (i) From Equation 2.20 and Definition 2, Appendix A.1, it follows that

$$\begin{aligned}\psi(p) &= \exp(i\xi p)\Big\{\exp(-\omega^2p^2/2) + i\exp(-\omega^2p^2/2)\tau(\delta\omega p)\Big\}\\ &= (\cos p\xi + i\sin p\xi)\Big\{\exp(-\omega^2p^2/2) + i\exp(-\omega^2p^2/2)\tau(\delta\omega p)\Big\}\\ &= \exp(-\omega^2p^2/2)\cos p\xi + i\exp(-\omega^2p^2/2)\tau(\delta\omega p)cosp\xi + i\exp(-\omega^2p^2/2)\sin p\xi\\ &- \exp(-\omega^2p^2/2)\tau(\delta\omega p)\sin p\xi. \qquad (2.21)\end{aligned}$$

Therefore, it follows from Definition 8, Appendix A.1, that the cosine moment is

$$\begin{aligned}\alpha_p &= E[\cos p\Theta]\\ &= \exp(-\omega^2p^2/2)\cos p\xi - \exp(-\omega^2p^2/2)\tau(\delta\omega p)\sin p\xi\\ &= \exp(-\omega^2p^2/2)(\cos p\xi - \tau(\delta\omega p)\sin p\xi). \qquad (2.22)\end{aligned}$$

(ii) Similarly, from Equation 2.21 and Definition 8, Appendix A.1, the sine moment is given by

$$\begin{aligned}\beta_p &= E[\sin p\Theta]\\ &= i\exp(-\omega^2p^2/2)\sin p\xi + i\exp(-\omega^2p^2/2)\tau(\delta\omega p)cosp\xi\\ &= exp(-\omega^2p^2/2)(\sin p\xi + \tau(\delta\omega p)cosp\xi). \qquad (2.23)\end{aligned}$$

□

Using these trigonometric moments, an alternative representation for the PDF of Θ is derived.

Theorem 4 *An alternative representation for the PDF of* $\Theta \sim WSN(\xi, \omega^2, \lambda)$ *is given by*

$$f(\theta;\xi,\omega^2,\lambda) = \frac{1}{2\pi}[1+2\sum_{p=1}^{\infty}\exp(-\omega^2p^2/2)\{\cos p(\theta-\xi)+\tau(\delta\omega p)\sin p(\theta-\xi)\}], \quad 0 \le \theta \le 2\pi.$$

Proof.

$$
\begin{aligned}
f(\theta;\xi,\omega^2,\lambda) &= \sum_{k=-\infty}^{\infty} g(\theta+2\pi k)^{(1)} \\
&= \frac{1}{2\pi}\sum_{p=-\infty}^{\infty} \psi_p \exp(-ip\theta)^{(2)} \\
&= \frac{1}{2\pi}\Big[1+2\sum_{p=1}^{\infty}(\alpha_p\cos p\theta+\beta_p\sin p\theta)\Big]^{(2)} \\
&= \frac{1}{2\pi}\Big[1+2\sum_{p=1}^{\infty}\exp(-\omega^2p^2/2)\{(\cos p\xi-\tau(\delta\omega p)\sin p\xi)\cos p\theta \\
&+ (\sin p\xi+\tau(\delta\omega p)\cos p\xi)\sin p\theta\}\Big]^{(3)} \\
&= \frac{1}{2\pi}\Big[1+2\sum_{p=1}^{\infty}\exp(-\omega^2p^2/2)\{\cos p\theta\cos p\xi-\tau(\delta\omega p)\cos p\theta\sin p\xi \\
&+ \sin p\theta\sin p\xi+\tau(\delta\omega p)\sin p\theta\cos p\xi\}\Big] \\
&= \frac{1}{2\pi}\Big[1+2\sum_{p=1}^{\infty}\exp(-\omega^2p^2/2)\{\cos(p\theta-p\xi)+\tau(\delta\omega p)\sin(p\theta-p\xi)\}\Big]^{(4)} \\
&= \frac{1}{2\pi}\Big[1+2\sum_{p=1}^{\infty}\exp(-\omega^2p^2/2)\{\cos p(\theta-\xi)+\tau(\delta\omega p)\sin p(\theta-\xi)\}\Big]. \quad (2.24)
\end{aligned}
$$

(1) Applying Equation 2.1;

(2) Applying Definition 9, Appendix A.1;

(3) Applying Equation 2.22 and 2.23;

(4) Applying Definition 3, Appendix A.1.

2.5.2 A Monte Carlo approximation for the trigonometric moments

With the transformation $\theta = Y(mod 2\pi)$, implying $0 \leq \theta \leq 2\pi$, where Y is defined in Equation 2.10, a random variable with support on the unit circle is obtained [20]. The inline variable can then be expressed as $Y = \theta + 2\pi K$, where K assumes values in $Y = \{0, \pm 1, \pm 2, ...\}$. Then, the PDF of θ is defined by using Equation 2.15

$$f(\theta; \xi, \omega^2, \lambda) = \sum_{k \in Y} \frac{2}{\omega} \phi\Big(\frac{\theta + 2\pi k - \xi + \frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}}}{\omega}\Big) \Phi\Big(\lambda\Big(\frac{\theta + 2\pi k - \xi + \frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}}}{\omega}\Big)\Big). \tag{2.25}$$

According to Mastrantonio et al. [20], an accurate approximation can be obtained by truncating the sum. Also, let K be a random variable, the PDF inside the sum in Equation 2.25, is the joint PDF of (θ, K), which is marginalised over K to obtain the PDF of the circular variable [20].

In order to obtain an approximation for the trigonometric moments, the cosine and sine moments for the WSN distribution are described in the following corollaries where the derivation follows similarly as before.

Corollary 1 *The cosine moment for* $\Theta \sim WSN(\xi, \omega^2, \lambda)$*, is given by*

$$\begin{aligned} \alpha_p &= E[\cos p\Theta | \Psi] \\ &= \exp(-\omega^2 p^2/2)(\cos p\xi^* - \tau(\frac{\lambda\omega p}{\sqrt{1+\lambda^2}}) \sin p\xi^*), \end{aligned} \tag{2.26}$$

where ξ *is substituted with* $\xi^* = \xi - \frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}}$ *in Equation 2.22 and* Ψ *denotes the vector of parameters* (ξ, ω^2, λ)*.*

Corollary 2 *The sine moment for* $\Theta \sim WSN(\xi, \omega^2, \lambda)$*, is given by*

$$\begin{aligned} \beta_p &= E[\sin p\Theta | \Psi] \\ &= \exp(-\omega^2 p^2/2)(\sin p\xi^* + \tau(\frac{\lambda\omega p}{\sqrt{1+\lambda^2}}) \cos p\xi^*), \end{aligned} \tag{2.27}$$

where ξ *is substituted with* $\xi^* = \xi - \frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}}$ *in Equation 2.23 and* Ψ *denotes the vector of parameters* (ξ, ω^2, λ)*.*

In order to compute $\tau(\cdot)$, Mastrantonio et al. [20] suggest the use of Monte Carlo approximation for α_p and β_p, which is defined as $\hat{\alpha}_p$ and $\hat{\beta}_p$, since $\tau(.)$ is not available in closed form.

Corollary 3 *The Monte Carlo approximation for the cosine and sine moments, $\hat{\alpha}_p$ and $\hat{\beta}_p$, are given by*

$$\hat{\alpha}_p \approx \frac{\exp(-\frac{\omega^2}{1+\lambda^2}p^2/2)}{N}\sum_{n=1}^{N}\cos(p(\xi+\frac{\omega\lambda}{\sqrt{1+\lambda^2}}|U|-\frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}})),$$
$$\hat{\beta}_p \approx \frac{\exp(-\frac{\omega^2}{1+\lambda^2}p^2/2)}{N}\sum_{n=1}^{N}\sin(p(\xi+\frac{\omega\lambda}{\sqrt{1+\lambda^2}}|U|-\frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}})).$$

Proof. From Equation 2.10, it follows that

$$Y|U \sim N\Big(\xi+\frac{\omega\lambda}{\sqrt{1+\lambda^2}}|U|-\frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}},\frac{\omega^2}{1+\lambda^2}\Big), \tag{2.28}$$

and therefore,

$$\Theta|U \sim WN\Big(\xi+\frac{\omega\lambda}{\sqrt{1+\lambda^2}}|U|-\frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}},\frac{\omega^2}{1+\lambda^2}\Big), \tag{2.29}$$

where the skewness parameter, λ, is then equal to 0.
To obtain the Monte Carlo approximation, let $\{U^n\}_{n=1}^N$ be a set of N samples from the distribution of U [20]. The cosine moment is defined by applying Definition 5, Appendix A.1,

$$\begin{aligned}\alpha_p &= E[\cos p\Theta]\\ &= E_U E_{\Theta|U}[\cos p\Theta|U],\end{aligned}$$

since the cosine moment of $\Theta|U$ is $E_{\Theta|U}[\cos p\Theta|U]$, then a Monte Carlo approximation of α_p, using Equation 2.4 and 2.29, is

$$\hat{\alpha}_p \approx \frac{\exp(-\frac{\omega^2}{1+\lambda^2}p^2/2)}{N}\sum_{n=1}^{N}\cos\Big(p\Big(\xi+\frac{\omega\lambda}{\sqrt{1+\lambda^2}}|U|-\frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}}\Big)\Big).$$

Similarly, by applying Definition 5, Appendix A.1, the sine moments can be written as

$$\begin{aligned}\beta_p &= E[\sin p\Theta] \\ &= E_U E_{\Theta|U}[\sin p\Theta|U],\end{aligned}$$

since the sine moment of $\Theta|U$ is $E_{\Theta|U}[sinp\Theta|U]$, then a Monte Carlo approximation of β_p, using Equation 2.5 and 2.29, is

$$\hat{\beta}_p \approx \frac{\exp(-\frac{\omega^2}{1+\lambda^2}p^2/2)}{N}\sum_{n=1}^{N}\sin\Big(p\Big(\xi+\frac{\omega\lambda}{\sqrt{1+\lambda^2}}|U|-\frac{\omega\lambda\sqrt{2}}{\sqrt{\pi(1+\lambda^2)}}\Big)\Big).$$

□

The approximation for the circular mean and concentration [20] follows from Definition 11, Appendix A.1,

$$\begin{aligned}\hat{\bar{\xi}} &= \arctan(\frac{\hat{\alpha}_1}{\hat{\beta}_1}), \\ \hat{\bar{c}} &= \sqrt{\hat{\alpha}_1^2+\hat{\beta}_1^2}.\end{aligned}$$

2.6 Estimation for the centred parameterisation

The implication for the use of MLE for the WSN distribution is exactly the same as those given for the linear case [25]. The problems caused by the direct parameterisation can be avoided by using the centred parameterisation of Azzalini [3] for the MLE, but for the wrapped distributions there exist no equivalent standardisation to that used for the linear case [25]. However, Pewsey [25] continued by using the same logic as before of the centred parameterisation, by denoting parameters that reflect the mean, variance and skewness of the WSN distribution. By wrapping Y_C onto the unit circle, the circular random variable is denoted as Θ_C with centred parameters μ, σ^2 and γ_1 [26]. Then, by using the relations (2.12) in Equation 2.15, the PDF of $\Theta_C \sim WSN_C(\mu, \sigma^2, \gamma_1)$ is given

by

$$f(\theta;\mu,\sigma,\gamma_1) = \frac{2}{\sigma\sqrt{1+c^2\gamma_1^{\frac{2}{3}}}}\sum_{r=-\infty}^{\infty}\phi\{\frac{1}{\sqrt{1+c^2\gamma_1^{\frac{2}{3}}}}((\frac{\theta+2\pi r-\mu}{\sigma})+c\gamma_1^{\frac{1}{3}})\}$$
$$\times\,\Phi[\frac{c\gamma_1^{\frac{1}{3}}}{\sqrt{\frac{2}{\pi}+c^2(\frac{2}{\pi}-1)\gamma_1^{\frac{2}{3}}(1+c^2\gamma_1^{\frac{2}{3}})}}((\frac{\theta+2\pi r-\mu}{\sigma})+c\gamma_1^{\frac{1}{3}})]. \tag{2.30}$$

where $0 \leq \theta \leq 2\pi$. According to Pewsey [26], there are some important limiting cases of $\Theta_C \sim WSN_C(\mu,\sigma^2,\gamma_1)$. As $\omega \to 0(\sigma \to 0)$, Θ_C degenerates to a point. Also, as $\omega \to \infty(\sigma \to \infty)$, the limiting distribution is the circular uniform [26]. Many procedures are available for identifying a point distribution, which are discussed in Mardia and Jupp [19].

2.6.1 Method of trigonometric moments

By pursuing the same logic as the centred parameterisation, let $\Theta_C \sim WSN_C(\mu,\sigma^2,\gamma_1)$. According to Pewsey [25], the moment estimates are generally excellent starting values for MLE where numerical optimisation procedures are required [25]. The moment estimators of the parameters μ,σ^2,γ_1 [25] are

$$\bar{\theta} = \begin{cases} \tan^{-1}(\frac{b_1}{a_1}), & \text{if } a_1 \geq 0 \\ \pi + \tan^{-1}(\frac{b_1}{a_1}), & \text{otherwise} \end{cases} \tag{2.31}$$

$$\bar{R} = \sqrt{a_1^2+b_1^2}, \tag{2.32}$$

and

$$\bar{b}_2 = \frac{1}{n}\sum_{i=1}^{n}\sin 2(\theta_i-\bar{\theta}), \tag{2.33}$$

respectively, where $\theta_1,...,\theta_n$ represent a random sample of size n from the WSN distribution [25]. Also, $\tan^{-1}(\theta) \in [-\pi/2,\pi/2]$, and

$$a_1 = \frac{1}{n}\sum_{i=1}^{n}\cos\theta_i,$$
$$b_1 = \frac{1}{n}\sum_{i=1}^{n}\sin\theta_i, \tag{2.34}$$

represent the first-order sample trigonometric moments. The parameters $(\mu, \sigma^2, \gamma_1)$ can be referred to as the circular parameterisation of the WSN distribution [25].
It can also be written in terms of the direct parameters, by using Equation 2.22 and 2.23,

$$\begin{aligned}\mu &= \tan^{-1}\left(\frac{\beta_1}{\alpha_1}\right) \\ &= \tan^{-1}\left(\frac{\sin\xi + \tau(\delta\omega)\cos\xi}{\cos\xi - \tau(\delta\omega)\sin\xi}\right),\end{aligned} \tag{2.35}$$

$$\begin{aligned}\sigma &= \sqrt{\alpha_1^2 + \beta_1^2} \\ &= \sqrt{\left((\exp(-\omega^2/2)(\cos\xi - \tau(\delta\omega)\sin\xi)\right)^2 + \left((\exp(-\omega^2/2)(\sin\xi + \tau(\delta\omega)\cos\xi)\right)^2} \\ &= \exp(-\omega^2/2)\sqrt{\cos^2\xi + \tau^2(\delta\omega)\sin^2\xi + \sin^2\xi + \tau^2(\delta\omega)\cos^2\xi} \\ &= \exp(-\omega^2/2)\sqrt{1 + \tau^2(\delta\omega)\sin^2\xi + \tau^2(\delta\omega)\cos^2\xi}^{(1)} \\ &= \exp(-\omega^2/2)\sqrt{1 + \tau^2(\delta\omega)[\sin^2\xi + \cos^2\xi]} \\ &= \exp(-\omega^2/2)\sqrt{1 + \tau^2(\delta\omega)}.\end{aligned} \tag{2.36}$$

(1) Applying Definition 4, Appendix A.1.
By using the following identity defined by Pewsey [25],

$$\gamma_1\sigma^2 = \beta_2(\alpha_1^2 - \beta_1^2) - 2\alpha_1\beta_1\alpha_2,$$

with Equation 2.22 and 2.23, the following is obtained

$$\begin{aligned}\gamma_1\sigma^2 &= \gamma_1(\exp(-\omega^2/2)\sqrt{1 + \tau^2(\delta\omega)})^2 \\ &= \exp(-2\omega^2)(\sin 2\xi + \tau(2\delta\omega)\cos 2\xi)\Big((\exp(-\omega^2/2)(\cos\xi - \tau(\delta\omega)\sin\xi)^2 \\ &- (\exp(-\omega^2/2)(\sin\xi + \tau(\delta\omega)\cos\xi)^2\Big) - 2(\exp(-\omega^2/2)(\cos\xi - \tau(\delta\omega)\sin\xi)) \\ &\times (\exp(-\omega^2/2)(\sin\xi + \tau(\delta\omega)\cos\xi))(\exp(-\omega^2/2)(\cos 2\xi - \tau(2\delta\omega)\sin 2\xi)) \\ &= \exp(-2\omega^2)(\sin 2\xi + \tau(2\delta\omega)\cos 2\xi)(1 - \tau^2(\delta\omega)) - 2(\exp(-\omega^2/2)(\cos\xi - \tau(\delta\omega)\sin\xi)) \\ &\times (\exp(-\omega^2/2)(\sin\xi + \tau(\delta\omega)\cos\xi))(\exp(-2\omega^2)(\cos 2\xi - \tau(2\delta\omega)\sin 2\xi)) \\ &= \exp(-\omega^6/2)[\tau(2\delta\omega)(1 - \tau^2(\delta\omega)) - 2\tau(\delta\omega)].\end{aligned}$$

Therefore, γ_1 is equal to

$$\gamma_1 = \frac{\exp(-\omega^6/2)[\tau(2\delta\omega)(1-\tau^2(\delta\omega)) - 2\tau(\delta\omega)]}{(1+\tau^2(\delta\omega))}, \tag{2.37}$$

where, as before, $\delta = \lambda/\sqrt{1+\lambda^2}$. The values of the direct parameters can be obtained from the circular ones, by solving numerically for $\delta\omega$ in

$$\frac{\gamma_1}{\sigma^4} = \frac{\tau(2\delta\omega)\{1-\tau^2(\delta\omega)\} - 2\tau(\delta\omega)}{\{1+\tau^2(\delta\omega)\}^3}. \tag{2.38}$$

The value of ω can then be obtained using

$$\omega = [-2\log\sigma + \log\{1+\tau^2(\delta\omega)\}]^{\frac{1}{2}}, \tag{2.39}$$

and hence that of δ, where λ can then be determined by

$$\lambda = \sqrt{\delta^2/1-\delta^2}.$$

The value of ξ is then the solution to

$$\xi = \tan^{-1}\{\tan\mu - \tau(\delta\omega)\}/\{1+\tau(\delta\omega)\tan\mu\}, \tag{2.40}$$

which satisfies Equation 2.35.

2.6.2 Maximum likelihood estimation (MLE)

From Equation 2.30 the log-likelihood function for a random sample, $\theta = (\theta_1, ..., \theta_n)$, of size n, $WSN_C(\mu, \sigma^2, \gamma_1)$ is given by

$$\begin{aligned}\ell(\mu,\sigma^2,\gamma_1) = {} & n\log 2 - n\log\sigma - \frac{n}{2}\log(1+c^2{\gamma_1}^{\frac{2}{3}}) \\ & + \sum_{i=1}^{n}\log\sum_{r=-\infty}^{\infty}\phi\{\frac{1}{\sqrt{1+c^2{\gamma_1}^{\frac{2}{3}}}}((\frac{\theta_i+2\pi r-\mu}{\sigma}) + c{\gamma_1}^{\frac{1}{3}})\} \\ & \times\Phi[\frac{c{\gamma_1}^{\frac{1}{3}}}{\sqrt{\frac{2}{\pi}+c^2(\frac{2}{\pi}-1){\gamma_1}^{\frac{2}{3}}(1+c^2{\gamma_1}^{\frac{2}{3}})}}((\frac{\theta_i+2\pi r-\mu}{\sigma}) + c{\gamma_1}^{\frac{1}{3}})].\end{aligned} \tag{2.41}$$

In order to maximise the log-likelihood function (Equation 2.41), numerical optimisation techniques should be used. Pewsey [26] recommend the simplex algorithm of Nelder and Mead [21].

2.6.3 Simulation study

Method of trigonometric moments

In order to evaluate the performance of the method of trigonometric moments in Section 2.6.1, a simulation study is conducted. 1000 samples are simulated from the $WSN_C(0, 1, 0.7)$ and $WSN_C(0, 1, 0.95)$ distribution, using the function 'rwsn' in the R package 'NPCirc' [23]. The simulated samples are plotted in Figure 2.5.

(a) $WSN_C(0, 1, 0.7)$. (b) $WSN_C(0, 1, 0.95)$.

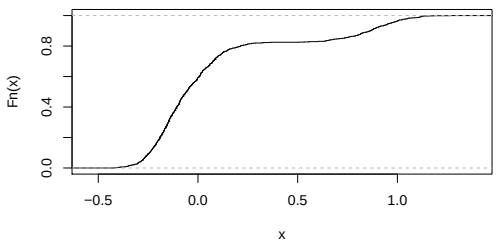
Figure 2.5: Random samples of size 1000 (red dots on the circle) taken from a WSN distribution with $\mu = 0, \sigma^2 = 1$ and different values for γ_1.

The moment estimators of the parameters $(\mu, \sigma^2, \gamma_1)$ can then be calculated by using the simulated values within Equation 2.31, 2.32,2.33 and then solving Equation 2.38, 2.39, 2.40. By repeating the process 1000 times, an average can be calculated for each estimator. The results are shown in table 2.1 and 2.2, along with the bias, standard error and the quantile of the true value.

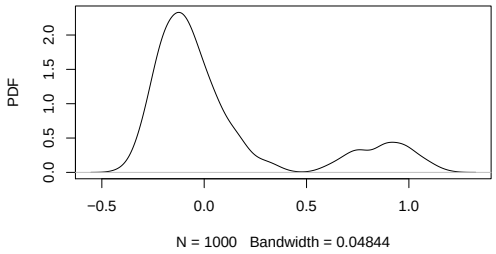
Table 2.1: The trigonometric moment estimators for a $WSN_C(0, 1, 0.7)$.

	μ	σ^2	γ_1
True value	0	1	0.7
Average value	0.08521464	1.00414198	0.94410201
Bias	0.085214644	0.004141982	0.244102014
Standard error	0.39307827	0.08108349	0.33985937
Quantile 0%	-0.4190053	0.8115167	0.1051646
Quantile 25%	-0.1683635	0.9438349	0.6920084
Quantile 50%	-0.0584025	1.0025875	0.9400457
Quantile 75%	0.1194406	1.0611042	1.1803698
Quantile 100%	1.2620864	1.2410292	2.2126766
Quantile of true value	0.594	0.490	0.259

The empirical cumulative distribution function (ECDF) plot and the empirical probability distribution function (EPDF) for $(\mu, \sigma^2, \gamma_1)$ are shown in the following figures:

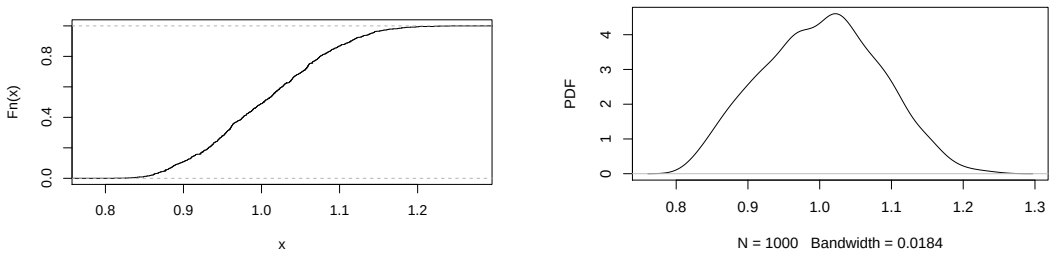


(a) ECDF for μ.



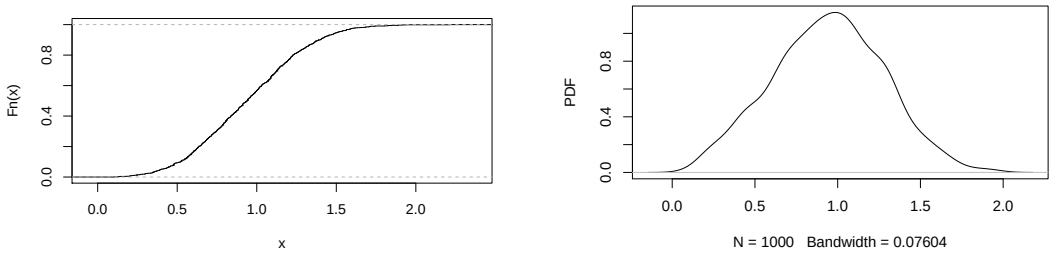
(b) EPDF for μ.

Figure 2.6: The ECDF plot and the EPDF of μ.



(a) ECDF for σ^2. (b) EPDF for σ^2.

Figure 2.7: The ECDF plot and the EPDF of σ^2.



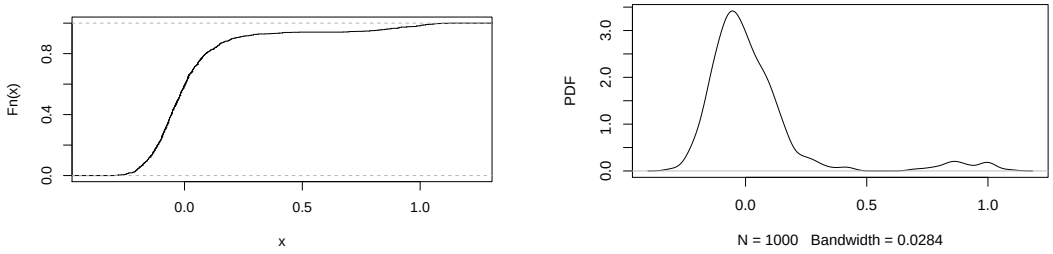
(a) ECDF for γ_1. (b) EPDF for γ_1.

Figure 2.8: The ECDF plot and the EPDF of γ_1.

Table 2.2: The trigonometric moment estimators for a $WSN_C(0, 1, 0.95)$.

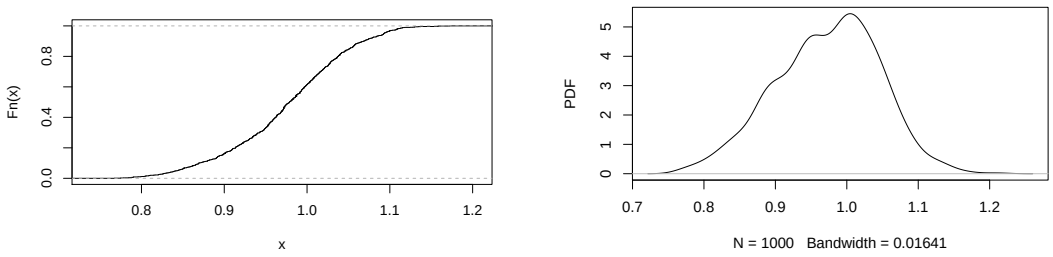
	μ	σ^2	γ_1
True value	0	1	0.95
Average value	0.03281717	0.97515607	1.12888254
Bias	0.03281717	-0.02484393	0.17888254
Standard error	0.2508611	0.0735756	0.3380342
Quantile 0%	-0.29755379	0.7673261	0.1167697
Quantile 25%	-0.09736302	0.9291077	0.9181349
Quantile 50%	-0.02727389	0.9788485	1.1460320
Quantile 75%	0.06145809	1.0266570	1.3362168
Quantile 100%	1.12528785	1.1722558	2.1300459
Quantile of true value	0.593	0.615	0.275

The ECDF plot and the EPDF for $(\mu, \sigma^2, \gamma_1)$ are shown in the following figures:



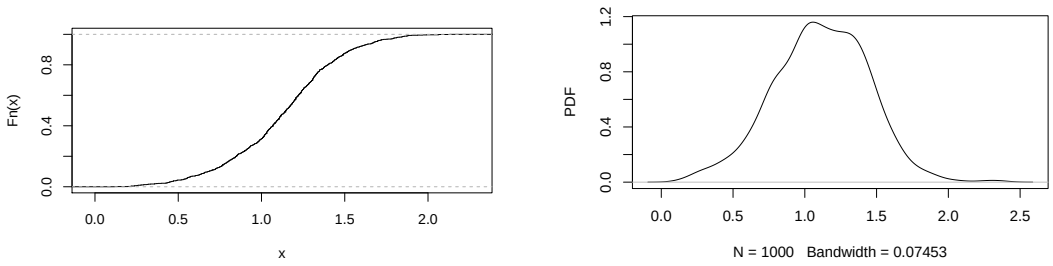
(a) ECDF for μ. (b) EPDF for μ.

Figure 2.9: The ECDF plot and the EPDF of μ.



(a) ECDF for σ^2. (b) EPDF for σ^2.

Figure 2.10: The ECDF plot and the EPDF of σ^2.



(a) ECDF for γ_1. (b) EPDF for γ_1.

Figure 2.11: The ECDF plot and the EPDF of γ_1.

In figures 2.6-2.11 the following are observed:

- The ECDF for μ in Figure 2.9a reaches a point of convergence much sooner than Figure 2.6a.
- The EPDF for μ is skewed to the right, while the EPDF for σ^2 and γ_1 appears to be more symmetric.

From table 2.1 and 2.2 the following are observed:

- The average value obtained for μ and σ^2 is close to the true value of the parameters. The average value obtained for γ_1 is not close to the true value.
- The bias for μ and σ^2 is close to zero but not the bias for γ_1.
- The bias and the standard error for each parameter are reduced by using the Monte Carlo approximation method discussed in Section 2.5.2.
- The quantile of the true value shows that the true value of the parameters lie within a 95% confidence interval.

Maximum likelihood estimation (MLE)

In order to evaluate the performance of MLE, in Section 2.6.2, a simulation study is conducted. Samples are simulated by using the trigonometric moment estimators as starting values since the moment estimators are generally excellent starting values for MLE [25]. Therefore, samples are simulated from the $WSN_C(0.08521464, 1.00414198, 0.94410201)$ and $WSN_C(0.03281717, 0.97515607, 1.12888254)$ distribution, where the starting parameter values are the estimators from the trigonometric moments in Section 2.6.3. For each combination 1000 samples are simulated using the function 'rwsn' in the R package 'NPCirc' [23]. The simulated samples are plotted in Figure 2.12.

(a) $WSN_C(0.0852, 1.0041, 0.9441)$. (b) $WSN_C(0.0328, 0.9752, 1.1289)$.

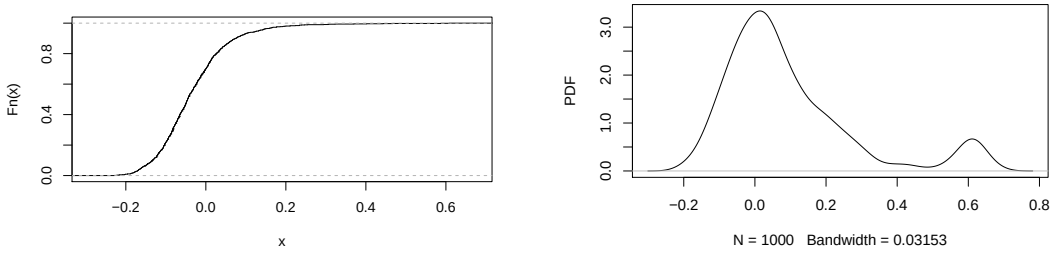
Figure 2.12: Random samples of size 1000 (red dots on the circle) taken from wrapped skew-normal distributions.

The ML estimators of the parameters $(\mu, \sigma^2, \gamma_1)$ are calculated by using the simulated values within the log-likelihood function (Equation 2.41). The log-likelihood function can then be maximised by using the Nelder and Mead optimisation technique. By repeating the process 1000 times, an average can be calculated for each estimator. The results are shown in tables 2.3 and 2.4, along with the bias, standard error and the quantile of the true value.

Table 2.3: The ML estimators for a $WSN_C(0.08521464, 1.00414198, 0.94410201)$.

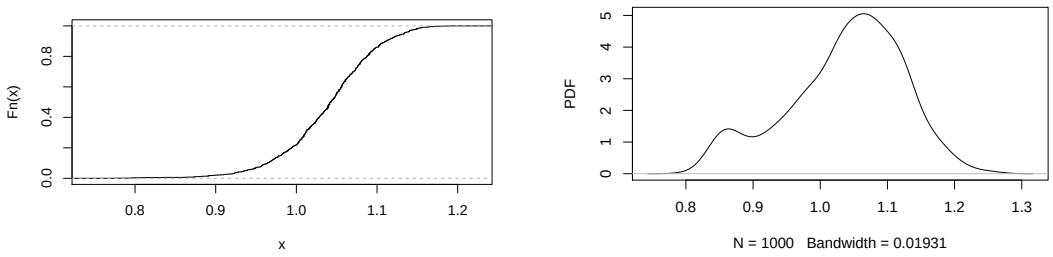
	μ	σ^2	γ_1
True value	0.08521464	1.00414198	0.94410201
Average value	-0.03304795	1.03870157	1.32160501
Bias	-0.11826259	0.03455959	0.37750300
Standard error	0.09341199	0.05939078	0.25975270
Quantile 0%	-0.22853825	0.7744576	0.01350392
Quantile 25%	-0.09303704	1.0043852	1.17627420
Quantile 50%	-0.04555383	1.0422544	1.33554936
Quantile 75%	0.01158395	1.0780974	1.49455458
Quantile 100%	0.60939905	1.1900363	2.00767593
Quantile of true value	0.914	0.249	0.072

The ECDF plot and the EPDF for $(\mu, \sigma^2, \gamma_1)$ are shown in the following figures:



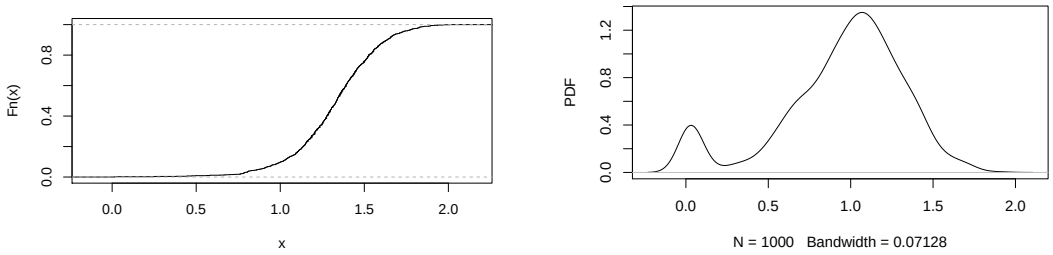
(a) ECDF for μ. (b) EPDF for μ.

Figure 2.13: The ECDF plot and the EPDF of μ.



(a) ECDF for σ^2. (b) EPDF for σ^2.

Figure 2.14: The ECDF plot and the EPDF of σ^2.



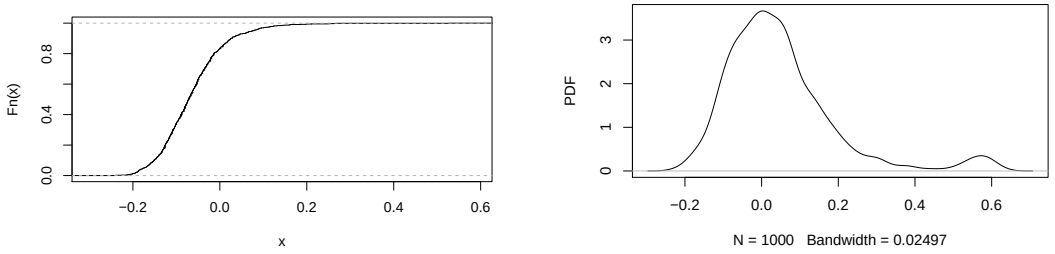
(a) ECDF for γ_1. (b) EPDF for γ_1.

Figure 2.15: The ECDF plot and the EPDF of γ_1.

Table 2.4: The ML estimators for a $WSN_C(0.03281717, 0.97515607, 1.12888254)$.

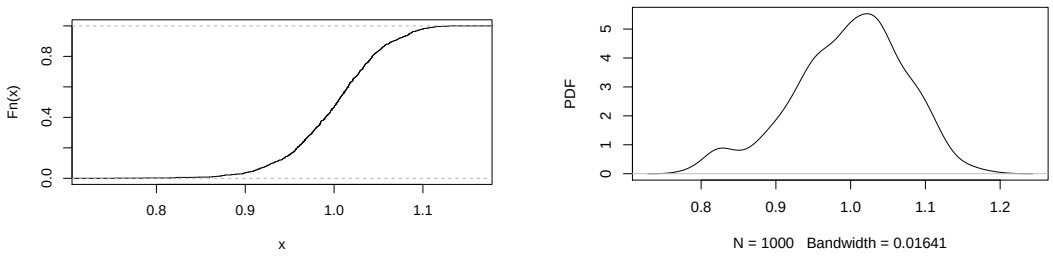
	μ	σ^2	γ_1
True value	0.03281717	0.97515607	1.12888254
Average value	-0.06428091	1.00027645	1.42510848
Bias	-0.09709808	0.02512038	0.29622594
Standard error	0.07697082	0.05316165	0.25182062
Quantile 0%	-0.24251558	0.7527145	0.100584
Quantile 25%	-0.11599322	0.9668246	1.275585
Quantile 50%	-0.07233008	1.0036101	1.436166
Quantile 75%	-0.02384610	1.0352295	1.583849
Quantile 100%	0.52871804	1.1301242	2.084326
Quantile of true value	0.910	0.297	0.112

The ECDF plot and the EPDF for $(\mu, \sigma^2, \gamma_1)$ are shown in the following figures:



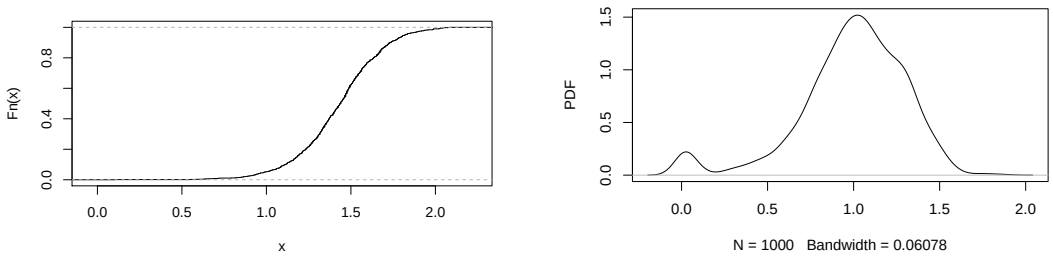
(a) ECDF for μ. (b) EPDF for μ.

Figure 2.16: The ECDF plot and the EPDF of μ.



(a) ECDF for σ^2. (b) EPDF for σ^2.

Figure 2.17: The ECDF plot and the EPDF of σ^2.



(a) ECDF for γ_1. (b) EPDF for γ_1.

Figure 2.18: The ECDF plot and the EPDF of γ_1.

In figures 2.13-2.18 the following are observed:

- The ECDF for μ in Figure 2.13a and 2.16a reach a point of convergence much sooner than the ECDF for σ^2 and γ_1.
- The EPDF for μ is skewed to the right, while the EPDF for σ^2 and γ_1 are skewed to the left.

From table 2.3 and 2.4 the following are observed:

- The average value obtained for μ and σ^2 is close to the true value of the parameters. The average value obtained for γ_1 is not close to the true value.
- The bias for μ and σ^2 is close to zero but not the bias for γ_1.
- The standard error for each parameter is a bit lower than those obtained in Table 2.1 and 2.2, which suggests that MLE give slightly more accurate results than the method of trigonometric moments.
- The quantile of the true value shows that the true value of the parameters lie within a 95% confidence interval.

2.7 Summary

In this chapter, the PDF of a wrapped distribution and a WN distribution were revisited. The SN distribution with a stochastic representation was also investigated in order to generate random numbers and visualisations thereof. The direct parameterisation as a basis for estimation was discussed with the estimation of the centred parameterisation. The WSN distribution was revisited with representations of different parameter values, where the CF and trigonometric moments were also investigated. A Monte Carlo approximation for the trigonometric moments, proposed by Mastrantonio et al. [20], was also provided. Finally, MLE and the method of trigonometric moments were investigated, where a simulation study consisting of two examples for each method was conducted for the purpose of comparison. The simulation study suggested that MLE gave slightly more accurate results than the method of trigonometric moments, which was observed by the lower standard error for each parameter. The bias and the standard error of the trigonometric moments can be reduced by using the Monte Carlo approximation method. In the following chapter the flexible generalised skew-normal (FGSN) distribution and the wrapped flexible generalised skew-normal (WFGSN) distribution are revisited. The method of MLE for the parameters of the FGSN distribution and the WFGSN distribution are then discussed. The WN, WSN and WFGSN distribution are also compared with a real data set as an application.

Chapter 3

The flexible generalised and wrapped flexible generalised skew-normal distributions

In Chapter 2 the SN and the WSN distribution were discussed. In Chapter 3 the focus will firstly be on the FGSN distribution where the transformation from Chapter 2 will be applied to obtain the WFGSN distribution. In Section 3.1 the PDF of the FGSN distribution is revisited and proven that it has at most two modes. A stochastic representation of the FGSN distribution is also presented. The method of MLE for the parameters of the FGSN distribution is discussed. In Section 3.2 the WFGSN distribution with examples is presented. The method of MLE for the parameters of the WFGSN distribution is discussed and a simulation study is conducted. The WN, WSN and WFGSN distribution are also compared with a real data set as an application in Section 3.3. The above outline is summarised in Figure 3.1.

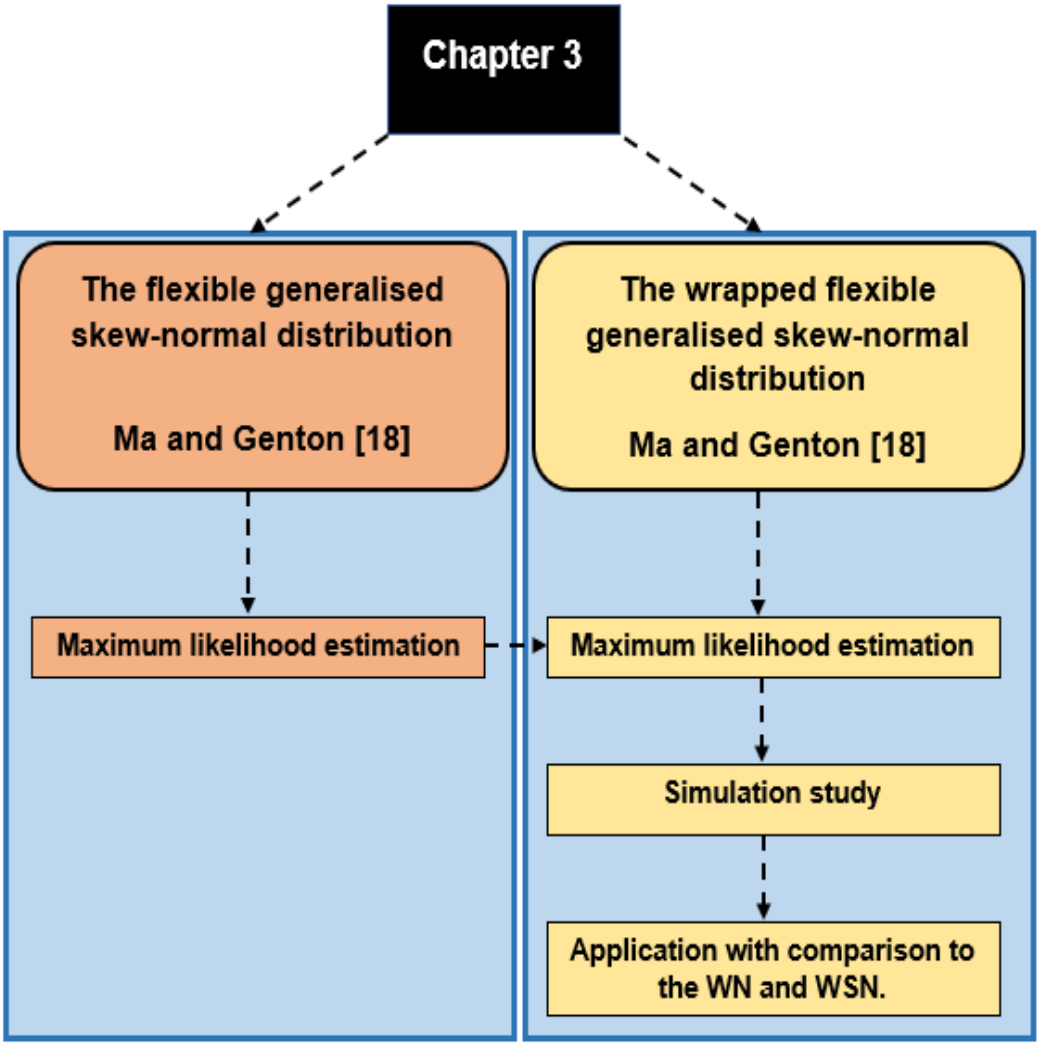


Figure 3.1: Outline of Chapter 3.

3.1 The flexible generalised skew-normal (FGSN) distribution

In this Section, using the same notation as in Section 2.2, Equation 2.7, it follows from Azzalini and Capitanio [5], Ma and Genton [18] and Wang et al. [27] that by setting $w(y) = \lambda(y)+\beta(y)^3$, the FGSN distribution is obtained. The PDF of the random variable $Y \sim FGSN(\xi, \omega^2, \lambda, \beta)$ [12] is defined as

$$f(y;\xi,\omega,\lambda,\beta) = \frac{2}{\omega}\phi(\frac{y-\xi}{\omega})\Phi(\lambda(\frac{y-\xi}{\omega}) + \beta(\frac{y-\xi}{\omega})^3). \tag{3.1}$$

Figure 3.2 shows an overlay of the FGSN PDF in Equation 3.1 with $\xi = 0, \omega^2 = 1$ and combinations of λ and β. This figure can be compared to Figure 2.2 in Section 2.2, where it can be observed that when $\beta = 0$ then $Y \sim SN(\lambda)$. If $\beta \neq 0$ then the PDF in Equation 3.1 shows bimodality as illustrated in Figure 3.2.

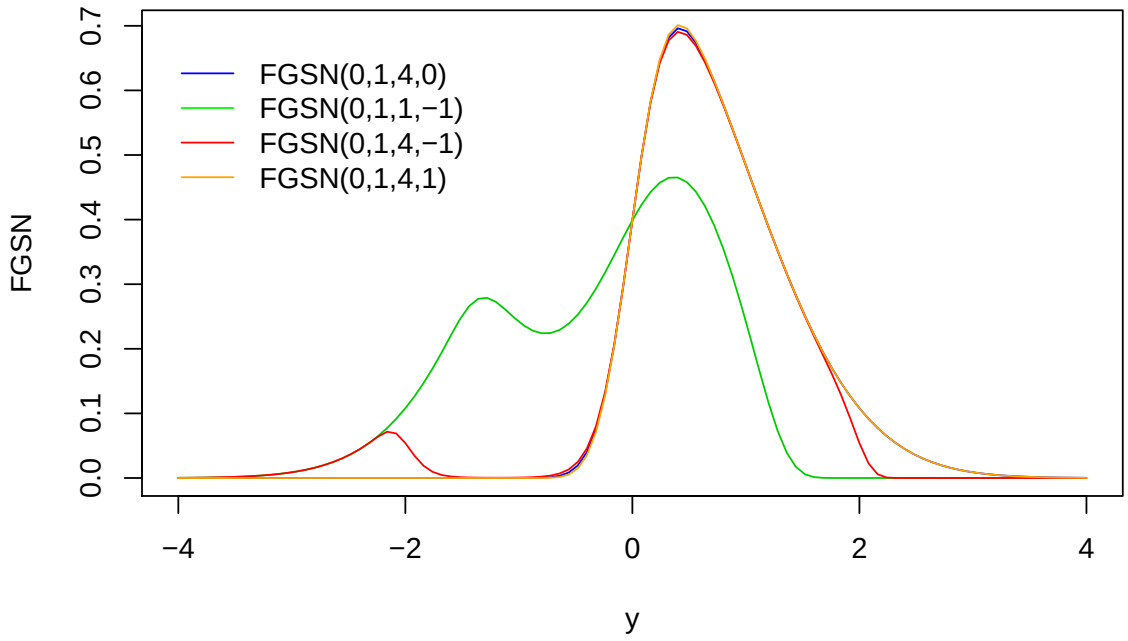


Figure 3.2: An overlay of the FGSN PDF (3.1) with $\xi = 0, \omega^2 = 1$ and combinations of λ and β.

It can therefore also be proven that Equation 3.1 has at most two modes [12].

Theorem 5 *The class of the FGSN distributions with PDF defined in Equation 3.1 has at most two modes.*

Proof. Without loss of generality, let $\xi = 0, \omega^2 = 1$ and assume $\beta > 0$ then it is only necessary to prove that,

$$\psi(y) = 2\phi(y)\Phi(\lambda y + \beta y^3),$$

has at most two modes [18], which can be proven by contradiction. If $\psi(y)$ has more than two modes, then $\psi'(y)$ has at least five zeros. In the following proof from Ma and Genton [18], it will be shown that this cannot be the case. Therefore,

$$\begin{aligned}
\psi^{'}(y) &= \frac{d}{dy}[2\phi(y)\Phi(\lambda y+\beta y^3)] \\
&= \frac{d}{dy}\Big[\frac{2}{\sqrt{2\pi}}\exp(-\frac{1}{2}y^2)\Phi(\lambda y+\beta y^3)\Big] \\
&= \Big[\frac{d}{dy}\Big(\frac{2}{\sqrt{2\pi}}\exp(-\frac{1}{2}y^2)\Big)\Big]\Phi(\lambda y+\beta y^3)+\frac{2}{\sqrt{2\pi}}\exp(-\frac{1}{2}y^2)\frac{d}{dy}\Big(\Phi(\lambda y+\beta y^3)\Big) \\
&= \Big(\frac{2}{\sqrt{2\pi}}(-y)\exp(-\frac{1}{2}y^2)\Phi(\lambda y+\beta y^3)\Big)+\Big(2\phi(y)\Phi(\lambda y+\beta y^3)(\lambda+3\beta y^2)\Big) \\
&= -2y\phi(y)\Phi(\lambda y+\beta y^3)+2\phi(y)\Phi(\lambda y+\beta y^3)(\lambda+3\beta y^2) \\
&= 2\phi(y)\Big[(\lambda+3\beta y^2)\Phi(\lambda y+\beta y^3)-y\Phi(\lambda y+\beta y^3)\Big],
\end{aligned}$$

where the following three cases need to be considered:

(i) Case 1 ($\lambda = 0$)

Let

$$\begin{aligned}
\psi^{'}(y) &= 2\phi(y)(3\beta y^2\Phi(\beta y^3)-y\Phi(\beta y^3)) \\
&= 2y\phi(y)(3\beta y\Phi(\beta y^3)-\Phi(\beta y^3) \\
&= 2y\phi(y)\eta(y),
\end{aligned}$$

where

$$\eta(y) = 3\beta y\Phi(\beta y^3)-\Phi(\beta y^3),$$

then it can be verified that

$$\eta^{'}(y) = -3\beta y(\beta y^3 3\beta y^2\phi(\beta y^3))+3\beta\phi(\beta y^3)-\phi(\beta y^3)3\beta y^2.$$

Let $x = y^2$ then,

$$\begin{aligned}
\eta^{'}(y) &= -3\beta y(\beta y^3 3\beta x\phi(\beta y^3))+3\beta\phi(\beta y^3)-\phi(\beta y^3)3\beta x \\
&= 3\beta\phi(\beta y^3)[-3\beta y\beta y^3 x+1-x] \\
&= 3\beta\phi(\beta y^3)[-3\beta^2 y^4 x+1-x] \\
&= 3\beta\phi(\beta y^3)[-3\beta^2 x^2 x+1-x] \\
&= 3\beta\phi(\beta y^3)[1-x-3\beta^2 x^3] \\
&= 3\beta\phi(\beta y^3)\eta_1(x),
\end{aligned}$$

where $\eta_1(x) = 1 - x - 3\beta^2 x^3$. As $\eta_1(x)$ is a decreasing function on $x \geq 0$, $\eta^{'}(y)$ has at most two zeros. Therefore, $\eta(y)$ has at most three zeros, hence $\psi^{'}(y)$ has at most four zeros.

(ii) Case 2 $(\lambda > 0)$

Notice that $\psi^{'}(y) > 0$ for $y \leq 0$. For

$$\gamma_1(y) = \frac{\psi^{'}(y)}{2y\phi(y)} = \frac{\Phi(\lambda y + \beta y^3)(\lambda + 3\beta y^2)}{y} - \Phi(\lambda y + \beta y^3),$$

the derivative can be obtained as follows:

note that,

$$\begin{aligned}\frac{d}{dy}\Phi(\lambda y + \beta y^3) &= \phi(\lambda y + \beta y^3)\frac{d}{dy}(\lambda y + \beta y^3) \\ &= \phi(\lambda y + \beta y^3)(\lambda + 3\beta y^2),\end{aligned}$$

then for the first component of $\gamma_1(y)$, let

$$\begin{aligned}\frac{d}{dy}\frac{\Phi(\lambda y + \beta y^3)(\lambda + 3\beta y^2)}{y} &= \Big[y\frac{d}{dy}[\Phi(\lambda y + \beta y^3)(\lambda + 3\beta y^2)] \\ &\quad - \Phi(\lambda y + \beta y^3)(\lambda + 3\beta y^2)\frac{d}{dy}y\Big]/(y^2) \\ &= \Big[y\phi(\lambda y + \beta y^3)(\lambda + 3\beta y^2) \\ &\quad - \Phi(\lambda y + \beta y^3)(\lambda + 3\beta y^2)\Big]/(y^2).\end{aligned}$$

Now let $x = \lambda + 3\beta y^2$, then

$$\frac{d}{dy}\frac{\Phi(\lambda y + \beta y^3)(\lambda + 3\beta y^2)}{y} = \Big[y\phi(\lambda y + \beta y^3)x - \Phi(\lambda y + \beta y^3)x\Big]/(y^2),$$

then for the second component of $\gamma_1(y)$ let $x = \lambda + 3\beta y^2$, then

$$\frac{d}{dy}\Phi(\lambda y + \beta y^3) = \phi(\lambda y + \beta y^3)x.$$

By joining the two components the following is obtained

$$\begin{aligned}\gamma_1^{'}(y) &= \frac{y\phi(\lambda y+\beta y^3)x-\Phi(\lambda y+\beta y^3)x}{y^2}-\phi(\lambda y+\beta y^3)x\\ &= \frac{y\phi(\lambda y+\beta y^3)x-\Phi(\lambda y+\beta y^3)x-xy^2\phi(\lambda y+\beta y^3)}{y^2}\\ &= \frac{xy\phi(\lambda y+\beta y^3)[1-y]-\Phi(\lambda y+\beta y^3)x}{y^2},\end{aligned}$$

where

$$x=\lambda+3\beta y^2>0.$$

Which means that $\gamma_1^{'}(y)$ has at most two positive zeros, so $\psi^{'}(y)$ has at most three positive zeros [18].

(iii) Case 3 ($\lambda<0$)

Notice that $\psi^{'}(y)<0$ for

$$y\in[0,\sqrt{-\lambda/(3\beta)}],$$

and $\psi^{'}(y)>0$ for

$$y\in(-\infty,-\sqrt{-\lambda/(3\beta)})].$$

So it is only necessary to look for solutions

$$y\in(\sqrt{-\lambda/(3\beta)},\infty),$$

and

$$y\in(-\sqrt{-\lambda/(3\beta)},0).$$

Let

$$x=\lambda+3\beta y^2.$$

Then, there is a one-to-one mapping between the y in the above range and

$$x\in(\lambda,\infty).$$

Let $\gamma_1(y)$ have the same expression as in Case 2, then $\gamma_1^{'}(y)$ has at most three zeros, therefore $\psi^{'}(y)$ has at most four zeros [18].

□

3.1.1 Stochastic representation of the flexible generalised skew-normal (FGSN) distribution

Elal-Olivero et al. [10] proposed a stochastic representation of the FGSN distribution that provide a method that is useful for simulation from a $FGSN(\xi, \omega^2, \lambda, \beta)$ with PDF defined in Equation 3.1. Let X and W be standard normal independent random variables then

$$Y = \begin{cases} X, & \text{if } W < w(Y), \\ -X, & \text{if } W \geq w(Y). \end{cases} \tag{3.2}$$

When $w(y) = \lambda(y) + \beta(y)^3$ the simulated values of the FGSN distribution, Equation 3.1 by Ma and Genton [18], is obtained [10].

An example from Wang et al. [27] is provided in Theorem 12, Appendix A.2.

3.1.2 Maximum likelihood estimation (MLE)

Parameter estimation for the FGSN distribution can be performed by the ML method. For the FGSN distribution, there is no closed forms available for centred parameterisation [12]. From Equation 3.1 the log-likelihood function for a sample of size $n, y_1, ..., y_n$, from $Y \sim FGSN(\xi, \omega^2, \lambda, \beta)$ is given by

$$\begin{aligned} \ell(\xi, \omega^2, \lambda, \beta) &= n \log 2 - n \log \omega \\ &\quad + \sum_{i=1}^{n} \log(\phi(\frac{y_i - \xi}{\omega})\Phi(\lambda(\frac{y_i - \xi}{\omega}) + \beta(\frac{y_i - \xi}{\omega})^3)). \end{aligned} \tag{3.3}$$

In order to maximise the log-likelihood function (Equation 3.3), numerical optimisation techniques should again be used, such as the simplex algorithm of Nelder and Mead [12, 21].

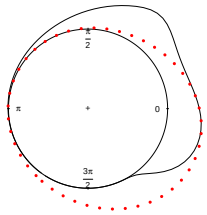
3.2 Wrapped flexible generalised skew-normal (WFGSN) distribution

By wrapping the FGSN distribution in Equation 3.1 onto the unit circle, the following PDF is obtained where $\Theta = Y(mod 2\pi)$ [12]:

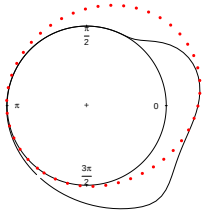
$$f(\theta;\xi,\omega^2,\lambda,\beta) = \frac{2}{\omega}\sum_{r=-\infty}^{\infty}\phi(\frac{\theta+2\pi(r)-\xi}{\omega})\Phi(\lambda(\frac{\theta+2\pi(r)-\xi}{\omega})+\beta(\frac{\theta+2\pi(r)-\xi}{\omega})^3), \tag{3.4}$$

for $0 \leq \theta \leq 2\pi$, $-\infty < \xi < \infty$, $\omega^2 \in \mathbb{R}^+$, $-\infty < \lambda < \infty$ and $-\infty < \beta < \infty$, which is denoted by $\Theta \sim WFGSN(\xi,\omega^2,\lambda,\beta)$. Centred or circular parameterisation are not explicitly available for this distribution, since circular moments of the WFGSN are not available in closed form [12].

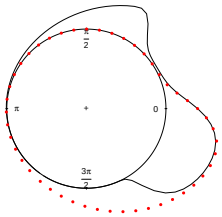
Figure 3.3 shows the shape of the PDF given in Equation 3.4 for $\xi = 0$, $\omega^2 = 1$ and various choices of λ and β, as well as a comparison to the WSN PDF in Figure 2.4. When $\beta \neq 0$ then the PDF in Equation 3.4 shows bimodality as illustrated in Figure 3.3. The command 'curve.circular' in the R package 'circular' [1] was used to create the circular plots.



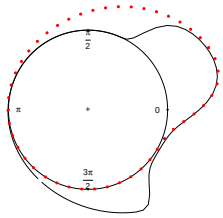
(a) WFGSN(0,1,-1,3).



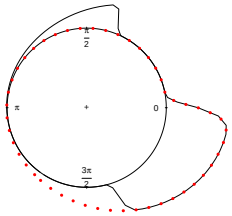
(b) WFGSN(0,1,1,-3).



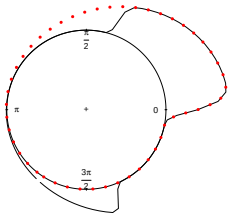
(c) WFGSN(0,1,-5,6).



(d) WFGSN(0,1,5,-6).



(e) WFGSN(0,1,-15,10).



(f) WFGSN(0,1,15,-10).

Figure 3.3: Examples of the WFGSN PDF, given in Equation 3.4, with $\xi = 0, \omega^2 = 1$ and various combinations of λ and β, as well as a comparison to the WSN PDF illustrated in Figure 2.4, which are shown with the dotted line.

3.2.1 Maximum likelihood estimation (MLE)

From Equation 3.4, the log-likelihood function for a sample of size $n, \theta = (\theta_1, ..., \theta_n)$, from $\Theta \sim WFGSN(\xi, \omega^2, \lambda, \beta)$ is given by

$$\ell(\theta; \xi, \omega^2, \lambda, \beta) = n \log 2 - n \log \omega + \sum_{i=1}^{n} \log\Big(\sum_{r=-\infty}^{\infty} \phi\Big(\frac{\theta_i + 2\pi r - \xi}{\omega}\Big)\Phi\Big(\lambda\Big(\frac{\theta_i + 2\pi r - \xi}{\omega}\Big) + \beta\Big(\frac{\theta_i + 2\pi r - \xi}{\omega}\Big)^3\Big)\Big). \tag{3.5}$$

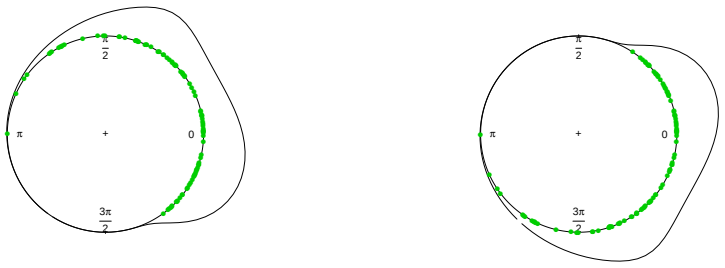
In order to maximise the log-likelihood function (Equation 3.5), numerical optimisation techniques should again be used, such as the simplex algorithm of Nelder and Mead [12, 21].

3.2.2 Simulation study

In this section, only MLE is considered, since centred parameterisations are not explicitly available for the WFGSN distribution [12]. Circular moments of the WFGSN distribution exist, but are not available in closed form, therefore it is not possible to use the method of trigonometric moments for the WFGSN distribution, in comparison to Section 2.6.3.

Maximum likelihood estimation (MLE)

In order to evaluate the performance of MLE in Section 3.2.1, a simulation study is conducted. 100 samples are simulated from the $WFGSN(0, 1, -1, 2)$ and $WFGSN(0, 1, 1, -2)$ distributions by using the stochastic representation of the FGSN distribution in Section 3.1.1, by Elal-Olivero et al. [10], and applying the transformation $\Theta = Y(mod 2\pi)$. The simulated samples are plotted in Figure 3.4.



(a) $WFGSN(0, 1, -1, 2)$. (b) $WFGSN(0, 1, 1, -2)$.

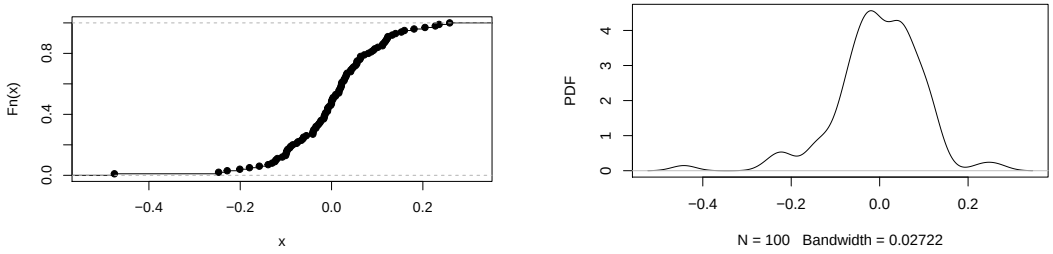
Figure 3.4: Random samples of size 100 (green dots on the circle) taken from a WFGSN distribution with $\xi = 0, \omega^2 = 1$ and different values for λ and β.

The MLE of the parameters $(\xi, \omega^2, \lambda, \beta)$ can then be calculated by using the simulated values within the log-likelihood function (Equation 3.5) where the function can then be maximised by the 'BFGS' optimisation technique, which is a quasi-Newton method (also known as a variable metric algorithm) by Wright et al. [28]. The 'BFGS' optimisation technique improved the results, therefore the Nelder and Mead optimisation technique was not used. By repeating the process, 100 times, an average can be calculated for each estimator. The results are shown in tables 3.1 and 3.2, along with the bias, standard error and the quantile of the true value.

Table 3.1: The ML estimators for a $WFGSN(0, 1, -1, 2)$.

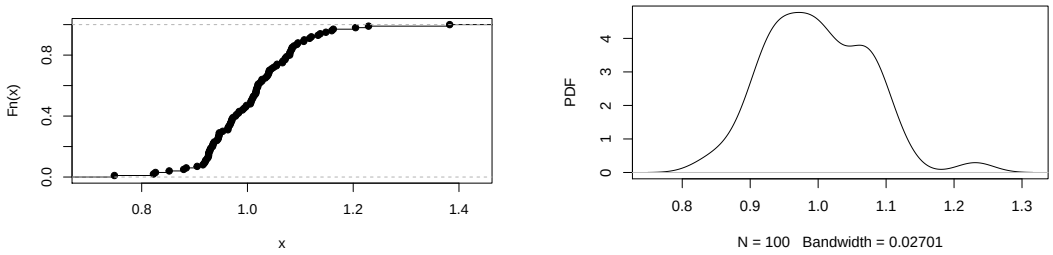
	ξ	ω^2	λ	β
True value	0	1	-1	2
Average value	0.001637939	1.008215490	-1.294772629	3.246210323
Bias	0.001637939	0.008215490	-0.294772629	1.246210323
Standard error	0.10824371	0.09053715	0.76698341	5.64933119
Quantile 0%	-0.475565221	0.7484135	-4.4114704	0.6171127
Quantile 25%	-0.056706795	0.9449470	-1.8367378	1.6119667
Quantile 50%	0.003272851	1.0079604	-1.2010460	2.2187159
Quantile 75%	0.057608409	1.0663819	-0.9177207	3.3151885
Quantile 100%	0.258690904	1.3823947	0.6865652	55.3855385
Quantile of true value	0.47	0.47	0.66	0.43

The ECDF plot and the EPDF for $(\xi, \omega^2, \lambda, \beta)$ are shown in the following figures:



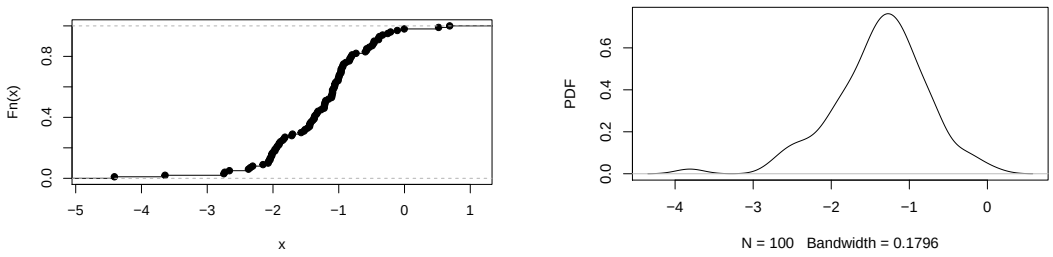
(a) ECDF for ξ. (b) EPDF for ξ.

Figure 3.5: The ECDF plot and the EPDF for ξ.



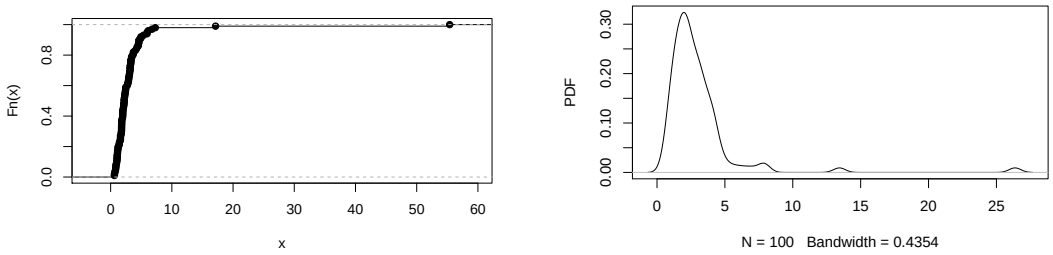
(a) ECDF for ω^2. (b) EPDF for ω^2.

Figure 3.6: The ECDF plot and the EPDF for ω^2.



(a) ECDF for λ. (b) EPDF for λ.

Figure 3.7: The ECDF plot and the EPDF for λ.



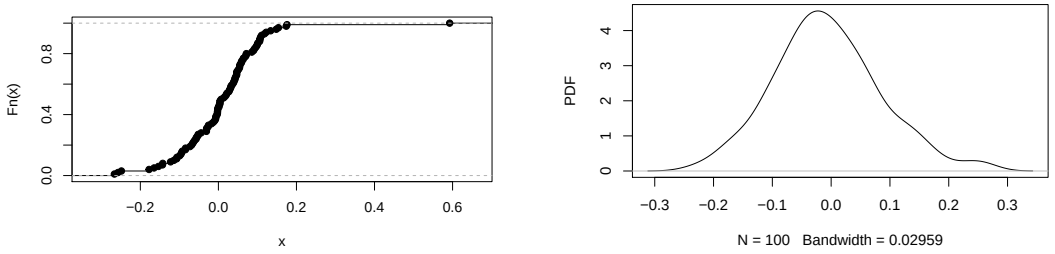
(a) ECDF for β. (b) EPDF for β.

Figure 3.8: The ECDF plot and the EPDF for β.

Table 3.2: The ML estimators for a $WFGSN(0, 1, 1, -2)$.

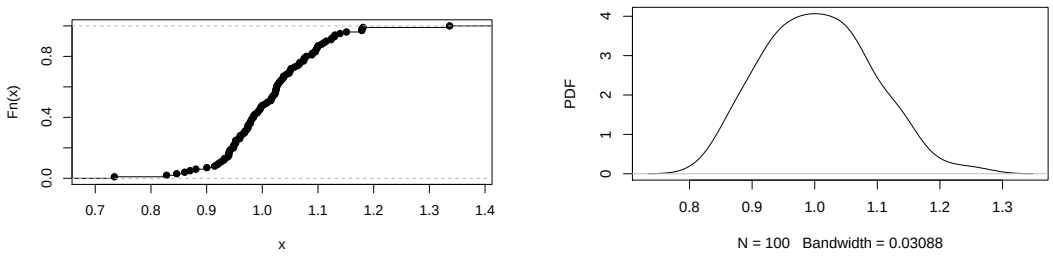
	ξ	ω^2	λ	β
True value	0	1	1	-2
Average value	0.008136449	1.012756408	1.214558936	-2.542347359
Bias	0.008136449	0.012756408	0.214558936	-0.542347359
Standard error	0.10778681	0.08606779	0.80880253	1.28404575
Quantile 0%	-0.26639751	0.7343827	-4.3027661	-7.8674828
Quantile 25%	-0.05353107	0.9576679	0.8616042	-2.8948024
Quantile 50%	0.01028177	1.0126085	1.2017762	-2.3064475
Quantile 75%	0.06029766	1.0670326	1.6866325	-1.6299923
Quantile 100%	0.59253459	1.3364711	2.5678761	-0.6322154
Quantile of true value	0.44	0.47	0.36	0.64

The ECDF plot and the EPDF for $(\xi, \omega^2, \lambda, \beta)$ are shown in the following figures:



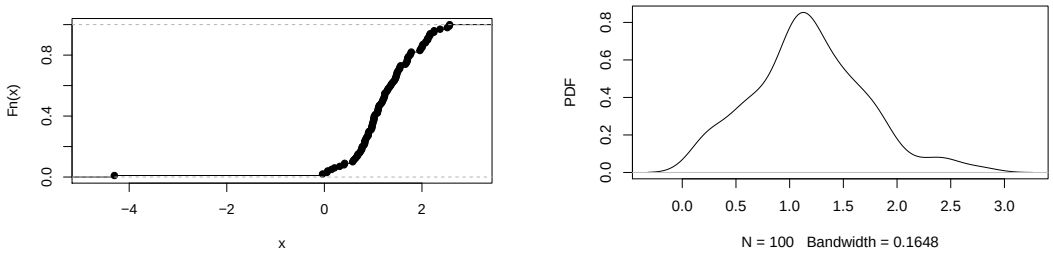
(a) ECDF for ξ. (b) EPDF for ξ.

Figure 3.9: The ECDF plot and the EPDF for ξ.



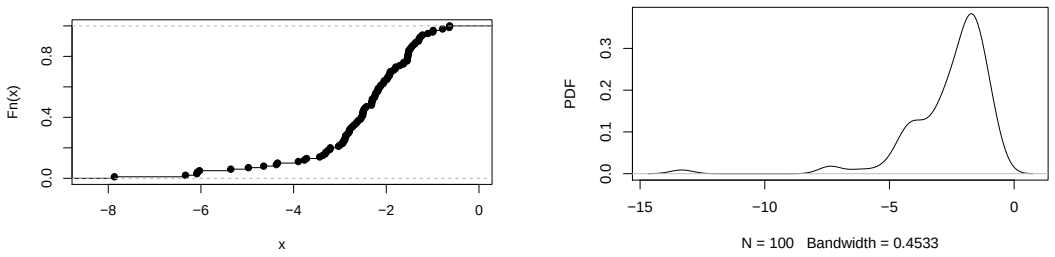
(a) ECDF for ω^2. (b) EPDF for ω^2.

Figure 3.10: The ECDF plot and the EPDF for ω^2.



(a) ECDF for λ. (b) EPDF for λ.

Figure 3.11: The ECDF plot and the EPDF for λ.



(a) ECDF for β. (b) EPDF for β.

Figure 3.12: The ECDF plot and the EPDF for β.

In figures 3.5-3.12 the following are observed:

- The ECDF for ξ in Figure 3.9a reach a point of convergence much sooner than Figure 3.5a.
- The ECDF for λ in Figure 3.11a only starts increasing from 0 onwards compared to Figure 3.7a.
- The EPDF for ξ is symmetric around 0.
- The EPDF for λ and β in Figure 3.7b,3.11b and 3.12b are all skewed to the left but the EPDF of β in Figure 3.8b is skewed to the right.

From table 3.1 and 3.2 the following are observed:

- The average value obtained for each estimate is relatively close to the true value of the parameters.
- The bias for each parameter might improve when increasing the sample size.
- The standard error for each parameter is close to zero except for β, but the value of the standard error for β improved in table 3.2.

- The quantile of the true value shows that the true value of the parameters lie within a 95% confidence interval.

3.3 Application

The dataset used for this application consists of a random sample of 1000 hourly observations of wind direction in the Atlantic coast of Galicia (NW-Spain). The wind direction data is only during the winter season (November to February) from 2003 until 2012 with a total of 19488 hourly observations. The data is registered by a buoy located at longitude -9.21E and latitude 43.500N in the Atlantic Ocean. This dataset 'speed.wind', analysed in Oliveira et al. [22], is from the R package 'NPCirc' [23]. Figure 3.13 represents a map of the Galicia area and the location of the buoy.

Figure 3.13: A map of the Galicia area and the location of the buoy shown in red.

Figure 3.14 illustrates a raw circular data plot and rose diagram for the Galicia wind direction data where the command 'rose.diag' in the R package 'NPCirc' [23] was used to create the rose diagram.

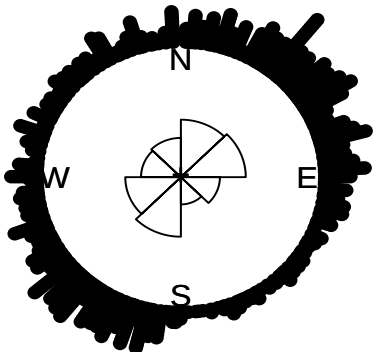


Figure 3.14: Raw circular data plot and rose diagram for the Galicia wind direction data.

The rose diagram in Figure 3.14 reveals that the distribution contains at least one skewed mode. The goal is to compare the WN (Equation 2.2), WSN (Equation 2.15) and WFGSN (Equation 3.4) distributions for the Galicia wind direction data. Estimates of the parameters are obtained using MLE. Several different starting points were used to implement an optimisation algorithm in order to avoid local maxima [12]. A small number of local maxima is obtained when starting from many different points, where the one corresponding to the maximum value of the likelihood function is selected. Table 3.3 represents the parameter estimates and the standard errors for the parameters of the WN, WSN and WFGSN distributions.

Table 3.3: Estimates and standard errors (in parenthesis) for the parameters of the WN, WSN and WFGSN distributions fitted to the Galicia wind direction data.

Distribution	ξ	ω^2	λ	β
WN	-1.895	2.296	-	-
	(0.231)	(0.048)	-	-
WSN	-0.952	2.000	-1.000	-
	(0.053)	(0.113)	(0.294)	-
WFGSN	1.104	2.500	3.480	-15.000
	(0.060)	(0.035)	(0.182)	(2.114)

Figure 3.15 illustrates a raw circular data plot and rose diagram of the Galicia wind direction data together with the WN, WSN and WFGSN PDF fitted by MLE.

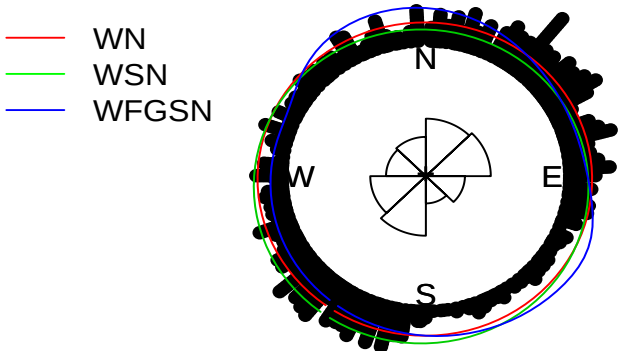


Figure 3.15: Raw circular data plot and rose diagram for the Galicia wind direction data together with the WN, WSN and WFGSN PDF fitted by MLE.

Table 3.4 shows the maximised log-likelihood (ℓ_{max}) and Akaike information criterion (AIC)/Bayesian information criterion (BIC) values for the WN, WSN and WFGSN distributions fitted to the Galicia wind direction data. The AIC is calculated as

$-2(\ell_{max} - p)$ and the BIC is calculated as $-2\ell_{max} + \log(n)p$ where p is the number of parameters estimated and n the sample size [12].

Table 3.4: Maximised log-likelihood (ℓ_{max}) and AIC/BIC values for the WN, WSN and WFGSN distributions fitted to the Galicia wind direction data where the minimum AIC/BIC value is identified using bold type.

Distribution	ℓ_{max}	AIC	BIC
WN	-1803.424	**3602.849**	**3593.069**
WSN	-1842.98	3679.96	3665.285
WFGSN	-1867.77	3727.54	3707.973

The comparison based on the AIC/BIC value has identified that the WN distribution fit the Galicia wind direction data best. However, from Figure 3.15 it seems that the WFGSN distribution describe the shape of the observed distribution relatively well.

3.4 Summary

In this chapter, the PDF of the FGSN distribution was revisited with representations of different parameter values, and proved that it has at most two modes. A stochastic representation of the FGSN distribution was presented. MLE was also investigated as a method of parameter estimation. The PDF of the WFGSN distribution was revisited with representations of different parameter values. A simulation study was also conducted using MLE. The WN, WSN and WFGSN distributions were also compared with wind direction data as an application. A comparison based on the AIC/BIC has identified the WN distribution to fit the Galicia wind direction data best. It is anticipated that the WSN and WFGSN distribution will provide a better fit to circular data that contains a higher level of skewness and bimodality. In the following chapter, examples of skew scale mixtures of normal (SSMN) distributions are provided such as the skew-Student-t normal (StN) distribution, the skew-slash (SSL) distribution and the skew-contaminated normal (SCN) distribution. The wrapped versions of each distribution will be used to

investigate if a better fit can be obtained for the Galicia wind direction data discussed in Section 3.3.

Chapter 4

Skew scale and wrapped skew scale mixtures of normal distributions

In Chapter 3 the WSN and WFGSN distributions were discussed, here the focus will be the skew scale and wrapped skew scale mixtures of normal (WSSMN) distributions. In Section 4.1, the PDF of the scale mixtures of normal (SMN) distributions is revisited as well as the SSMN distributions. In Section 4.2 examples of the SSMN distributions are provided, such as the StN, SSL and SCN distributions. The PDF and the CF of these distributions are revisited where the wrapped version of each distribution is defined. Graphical representations of the PDF and the wrapped PDF of these distributions are provided. In Section 4.3, the wrapped skew-Student-t normal (WStN), wrapped skew-slash (WSSL) and wrapped skew-contaminated normal (WSCN) distributions are compared with the Galicia wind direction data, discussed in Section 3.3, as an application. Lastly, these distributions are then compared to the WN, WSN and WFGSN distributions to investigate if a better fit can be obtained for the Galicia wind direction data. The above outline is summarised in Figure 4.1.

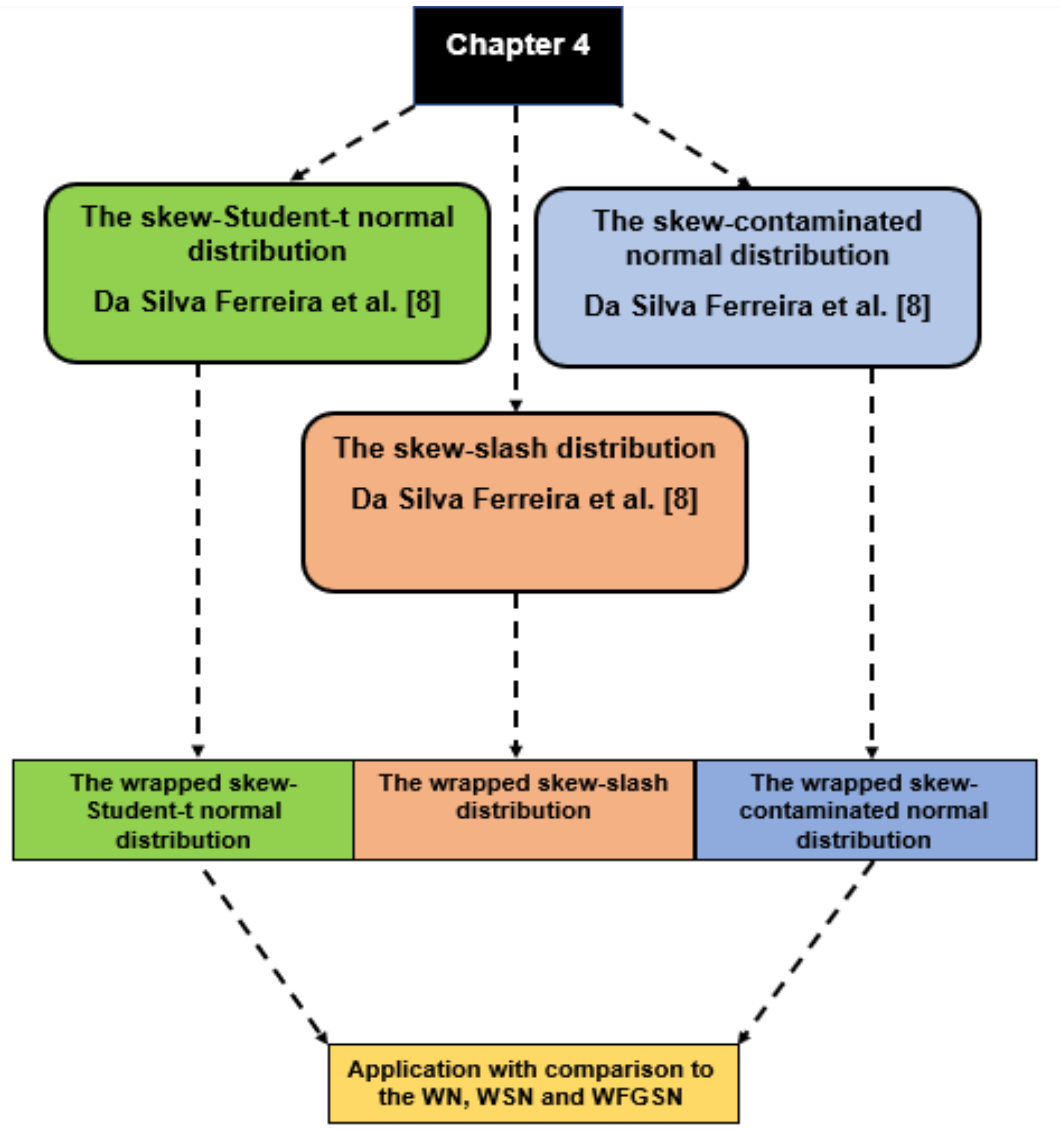


Figure 4.1: Outline of Chapter 4.

4.1 Background

The SMN distributions present a group of heavy tailed distributions [8]. Appropriate distributions for fitting skewed and heavy tailed data can be referred to as SSMN distributions which include distributions such as the StN, the SSL and the SCN distribution [8]. The SSMN distributions is based on the terminology of the SN distribution defined in Section 2.2, Equation 2.7. Jammalamadaka and Kozubowski [14] highlighted the importance of wrapping SMN distributions which generates useful flexible and asymmetric circular models.

4.1.1 Skew scale mixtures of normal (SSMN) distributions

A standardised continuous random variable Z has a SMN distribution if its PDF is as follows:

$$Z = \xi + \kappa(A)^{1/2}V, \tag{4.1}$$

where $V \sim N(0, \omega^2)$ is independent of the positive random variable A, $\kappa(\cdot)$ is a strictly positive function and $H(\cdot;\tau)$ the cumulative distribution function (CDF) which is indexed by the parameter vector τ. The following theorem then follows from da Silva Ferreira et al. [8].

Theorem 6 *A random variable Z follows a SMN distribution with location parameter $-\infty < \xi < \infty$ and scale parameter space of ω^2 if its PDF assumes the form*

$$f_0(z) = \int_0^\infty \phi(z|\xi, \kappa(a)\omega^2) dH(a;\tau), \tag{4.2}$$

with the notation $Z \sim SMN\ (\xi, \omega^2, H; \kappa)$. When $\xi = 0$ and $\omega^2 = 1$, the notation $Z \sim SMN\ (H;\kappa)$ is used.

The following definition introduces the new class of SMN distributions on the basis of Theorem 6 and Section 2.2, Equation 2.7, which will form the basis of this chapter [8].

Definition 1 *From da Silva Ferreira et al. [8], a random variable Z follows a SSMN distribution with location parameter $-\infty < \xi < \infty$, scale parameter ω^2 and skewness parameter $-\infty < \lambda < \infty$ if its PDF is given by*

$$f(z;\xi,\omega^2,\lambda,H;\kappa) = 2f_0(z)\Phi(\lambda\frac{z-\xi}{\omega}), \tag{4.3}$$

where $f_0(z)$ is the PDF of SMN distribution as defined in Equation 4.2. The PDF in Equation 4.3 has the notation $Z \sim SSMN(\xi, \omega^2, \lambda, H; \kappa)$. If $\xi = 0$ and $\omega^2 = 1$, it can be referred to as the standard SSMN distribution where $Z \sim SSMN(\lambda, H; \kappa)$.

When $\lambda = 0$, the SMN distribution defined in Equation 4.2 is obtained. The wrapped skew scale mixtures of normal (WSSMN) distributions is based on the terminology of the WSN distribution defined in Section 2.5, Equation 2.15.

Scale mixtures of skew-normal distributions are defined by the following stochastic representation by Kim and Genton [16]:

$$Z = \xi + W(\eta)^{1/2} Y, \tag{4.4}$$

where $Y \sim SN(0, \omega^2, \lambda)$, η is a mixing variable with a weight function $W(\eta)$, independent of Y, and a CDF $H(\eta)$.

4.2 Examples of skew scale mixtures of normal (SSMN) distributions

In the following section examples of SSMN distributions, such as the StN, the SSL and the SCN distribution are discussed. The wrapped version of the distributions are defined with a graphical representation of the PDF and wrapped PDF of each distribution.

4.2.1 The skew-Student-t normal (StN) distribution

The StN distribution with ν degrees of freedom, $StN(\xi, \omega^2, \lambda, \nu)$, is obtained from the mixture model in Theorem 13, Appendix A.2, with $A \sim$Gamma$(\nu/2, \nu/2)$ (Definition 12, Appendix A.1), $\nu > 0$ and $\kappa(a) = 1/a$. The PDF of Z has the form

$$f(z; \xi, \omega^2, \lambda, \nu) = 2\frac{1}{\omega\sqrt{\nu\pi}}\frac{\Gamma((\nu+1)/2)}{\Gamma(\frac{\nu}{2})}(1 + \frac{d}{\nu})^{-(\frac{\nu+1}{2})}\Phi(\lambda\frac{(z-\xi)}{\omega}), \tag{4.5}$$

where $d = (z - \xi)^2/\omega^2$. Gomez et al. [11] show that the StN distribution can provide a broader asymmetry range than the ordinary SN distribution. If $\nu = 1$ then it follows

that the skew-Cauchy normal distribution is a special case of the StN distribution. Also, when $\nu \to \infty$ the SN distribution is obtained as the limiting case. From Theorem 14, Appendix A.2, the mean and variance of Z are given by

$$E[Z] = \xi + b\omega\lambda(\nu/2)^{1/2}\frac{\Gamma((\nu-1)/2)}{\Gamma(\frac{\nu}{2})}E_D[(D+\lambda^2)^{-1/2}],$$

$$Var[Z] = \omega^2\Big[\frac{\nu}{\nu-2} - \frac{b^2\lambda^2\nu}{2}(\frac{\Gamma((\nu-1)/2)}{\Gamma(\frac{\nu}{2})})^2 E_D[(D+\lambda^2)^{-1/2}]\Big],$$

where $b = \sqrt{\frac{2}{\pi}}$ and $D \sim$Gamma$(\frac{\nu-1}{2}, \frac{\nu}{2})$ [8].

Figure 4.2 shows the PDF of the StN distribution in Equation 4.5 with parameters $\xi = 0, \omega^2 = 1$ and different values for λ and ν.

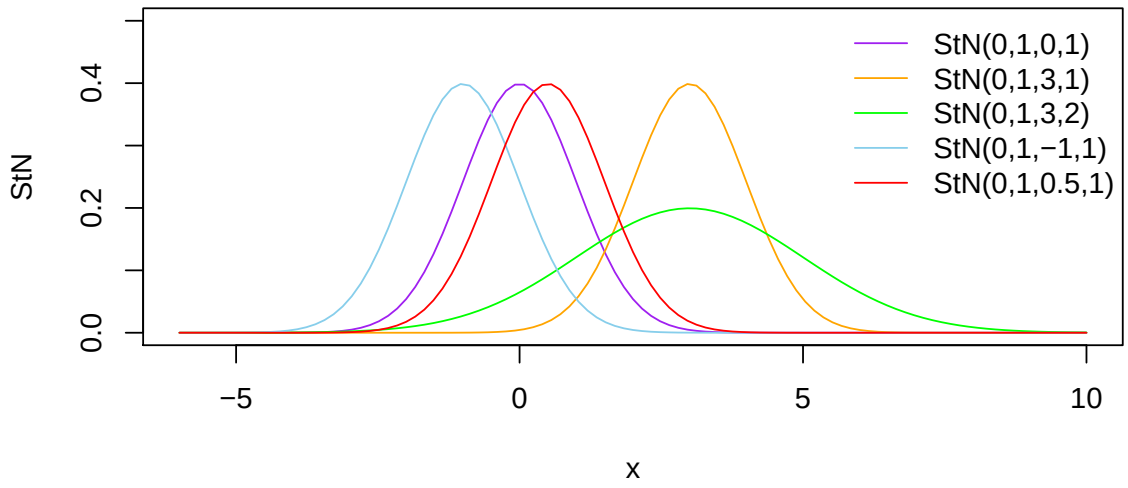


Figure 4.2: The StN PDF in Equation 4.5 with parameters $\xi = 0, \omega^2 = 1$ and combinations of λ and ν.

In Figure 4.2 it can be observed that the tail is heavier when the value of ν increases. Also, when λ becomes negative the PDF shifts to the left and vice versa.

• **The CF of the StN distribution**

The CF of the StN distribution is obtained from the Laplace-Stieltjes transform and the CF of the skew-t distribution using scale mixtures of skew-normal distributions [16]. Equation 4.6 is used to apply the Laplace-Stieltjes transform [16]. The skew-t distribution is related to the SN distribution, from Equation 4.4, by the following stochastic representation:

$$Z = \xi + \eta^{1/2} Y, \tag{4.6}$$

where η has an inverse-Gamma distribution, $\eta \sim IG(\nu/2, \nu/2)$. The PDF of η is

$$\frac{(\nu/2^{\nu/2})}{\Gamma(\nu/2)} \eta^{-\nu/2-1} \exp\{-\nu/(2\eta)\},$$

where $\eta > 0$.

Theorem 7 *Let Z follow the skew-t distribution defined in Equation 4.6. Then the CF of Z is*

$$\psi_Z(t) = \exp(i\xi t)\{\Psi_T(\omega t) + i\tau^*(\delta, \omega t)\},$$

where

$$\psi_T(t) = \frac{K_{\nu/2}(\sqrt{\nu} \mid t \mid)(\sqrt{\nu} \mid t \mid)^{\nu/2}}{\Gamma(\nu/2) 2^{\nu/2-1}}, \tag{4.7}$$

for $t \in \mathbb{R}$, $\nu > 0$,

$$\tau^*(\delta, \omega t) = \int_0^\infty \exp(\eta \omega^2 t^2/2) \tau(\delta \sqrt{\eta} dH(\eta)), \qquad \delta t > 0, \tag{4.8}$$

with $\tau^*(\delta, -\omega t) = -\tau^*(\delta, \omega t)$. The modified Bessel function of the third kind has the following integral representation [16]

$$K_\alpha(w) = \frac{1}{2} \int_0^\infty y^{\alpha-1} \exp\Big\{ -\frac{1}{2} w \big(y + \frac{1}{y}\big) \Big\} dy, \qquad w > 0 \quad for \quad -\infty < \alpha < \infty.$$

Proof. The conditional distribution of Z given η follows a SN distribution, i.e.

$Z|\eta \sim SN(\xi, \eta\omega^2, \lambda)$. Then, the CF of Z is

$$\begin{aligned}
\psi_Z(t) &= \int_0^\infty \int_{\mathbb{R}} \exp(itz) f(z|\eta) dz dH(\eta) \\
&= \int_0^\infty \psi_{Z|\eta}(t) dH(\eta) \\
&= \exp(i\xi t) \int_0^\infty \exp(-\eta\omega^2 t^2/2)\{1 + i\tau(\delta\sqrt{\eta}\omega t)\} dH(\eta)^{(1)} \\
&= \exp(i\xi t)\Big\{ L_\eta(\omega^2 t^2/2) + i \int_0^\infty \exp(-\eta\omega^2 t^2/2)\tau(\delta\sqrt{\eta}\omega t) dH(\eta) \Big\} \\
&= \exp(i\xi t)\{\psi_T(\omega t) + i\tau^*(\delta, \omega t)\},
\end{aligned}$$

where $L_\eta(\gamma)$ is the Laplace-Stieltjes transform

$$L_\eta(\gamma) = E\{\exp(-\gamma\eta)\} = \int_0^\infty \exp(-\gamma\eta) dH(\eta),$$

when η is a non-negative random variable.
(1) From Equation 2.17.

From Hurst [13], $L_\eta(\omega^2 t^2/2)$ becomes the CF of a StN distribution, $\psi_T(\omega t)$, after obtaining the CF of a symmetric generalised hyperbolic distribution. Some of the properties of the modified Bessel function of the third kind is then applied [13, 16]. Here, $\psi_T(t)$ (Equation 4.7) is the CF of a StN distribution with ν degrees of freedom and the integrand in Equation 4.8.

4.2.2 The wrapped skew-Student-t normal (WStN) distribution

By wrapping the StN distribution in Equation 4.5 onto the unit circle, the following PDF is defined where $\Theta = Z(mod 2\pi)$

$$f(\theta; \xi, \omega^2, \lambda, \nu) = 2\frac{1}{\omega\sqrt{\nu\pi}} \frac{\Gamma((\nu+1)/2)}{\Gamma(\frac{\nu}{2})} (1 + \frac{d}{\nu})^{-(\frac{\nu+1}{2})} \sum_{r=-\infty}^{\infty} \Phi(\lambda(\frac{\theta + 2\pi(r) - \xi}{\omega})), \quad (4.9)$$

for $0 \leq \theta \leq 2\pi$, $-\infty < \xi < \infty$, $-\infty < \lambda < \infty$ and $\nu > 0$, which is denoted by $\Theta \sim WStN(\xi, \omega^2, \lambda, \nu)$.

Figure 4.3 shows the PDF of the WStN distribution in Equation 4.9 with parameters $\xi = 0, \omega^2 = 1$ and different values for λ and ν.

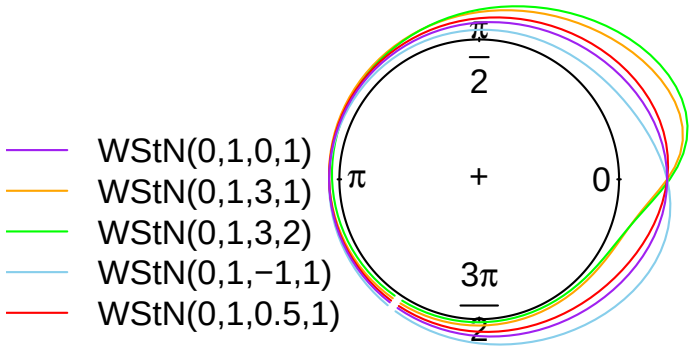


Figure 4.3: The PDF of the WStN in Equation 4.9 with parameters $\xi = 0, \omega^2 = 1$ and combinations of λ and ν.

In Figure 4.3 it can be observed that as the value of λ increases, the PDF develops a single mode.

- **The CF of the WStN distribution**

From Equation 4.9, the WSSMN distributions are defined by the following stochastic representation:

$$Z = \xi + W(\eta)^{1/2}(\theta + 2\pi r), \tag{4.10}$$

where $\theta \sim WSN(0, \omega^2, \lambda)$, η is a mixing variable with a weight function $W(\eta)$, independent of θ, and a CDF $H(\eta)$. The CF of the WStN distribution is obtained from the Laplace-Stieltjes transform and the CF of the wrapped skew-t distribution using wrapped scale mixtures of skew-normal distributions. The wrapped skew-t distribution is related

to the WSN distribution, from Equation 4.10, by the following stochastic representation:

$$Z = \xi + \eta^{1/2}(\theta + 2\pi r), \tag{4.11}$$

where η has an inverse-Gamma distribution, $\eta \sim IG(\nu/2, \nu/2)$. The PDF of η is

$$\frac{(\nu/2^{\nu/2})}{\Gamma(\nu/2)}\eta^{-\nu/2-1}\exp\{-\nu/(2\eta)\},$$

where $\eta > 0$. To apply the Laplace-Stieltjes transform, Equation 4.11 is used.

Let Z follow the wrapped skew-t distribution defined in Equation 4.11. By applying Theorem 10, Appendix A.2, the CF of Z is

$$\begin{aligned}\psi_Z(p) &= \exp(i\xi p)\{\psi_T(\omega p) + i\tau^*(\delta, \omega p)\},\\ \psi_T(p) &= \frac{K_{\nu/2}(\sqrt{\nu} \mid p \mid)(\sqrt{\nu} \mid p \mid)^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}},\end{aligned} \tag{4.12}$$

for $p = 0, 1, ..., \nu > 0$,

$$\tau^*(\delta, \omega p) = \int_0^\infty \exp(\eta\omega^2 p^2/2)\tau(\delta\sqrt{\eta}dH(\eta)), \qquad \delta p > 0,$$

with $\tau^*(\delta, -\omega p) = -\tau^*(\delta, \omega p)$ and the integral representation of the modified Bessel function of the third kind, $K_\alpha(w)$, defined in Theorem 7. Here, $\psi_T(t)$ (Equation 4.12) is the CF of a WStN distribution. This is stated without proof since the derivation is similar to that of Theorem 7.

4.2.3 The skew-slash (SSL) distribution

From da Silva Ferreira et al. [8], the SSL distribution is denoted by SSL$(\xi, \omega^2, \lambda, \nu)$ and arises when $\kappa(a) = 1/a$ where the distribution of A is Beta$(\nu, 1)$, $0 < a < 1$ and $\nu > 0$ [8]. The PDF of the SSL distribution (from Definition 1) is given by

$$f(z; \xi, \omega^2, \lambda, \nu) = 2\nu\Phi(\lambda\frac{(z-\xi)}{\omega})\int_0^1 a^{\nu-1}\phi(z|\xi, \frac{\omega^2}{a})da, \tag{4.13}$$

where $z \in \mathbb{R}$. The SSL distribution reduces to the SN distribution when $\nu \to \infty$. The mean and variance of Z are given by

$$E[Z] = \xi + \frac{b\omega\lambda\nu}{\nu - 1/2}E_D[(D + \lambda^2)^{-1/2}],$$

$$Var[Z] = \omega^2\Big(\frac{\nu}{\nu-1} - \frac{b^2\lambda^2\nu^2}{(\nu-1/2)^2}E_D[(D+\lambda^2)^{-1/2}]\Big),$$

where $D \sim$Beta$(1, \nu - 1/2)$ [8].

Figure 4.4 shows the PDF of the SSL distribution in Equation 4.13 with parameters $\xi = 0, \omega^2 = 1$ and different values for λ and ν.

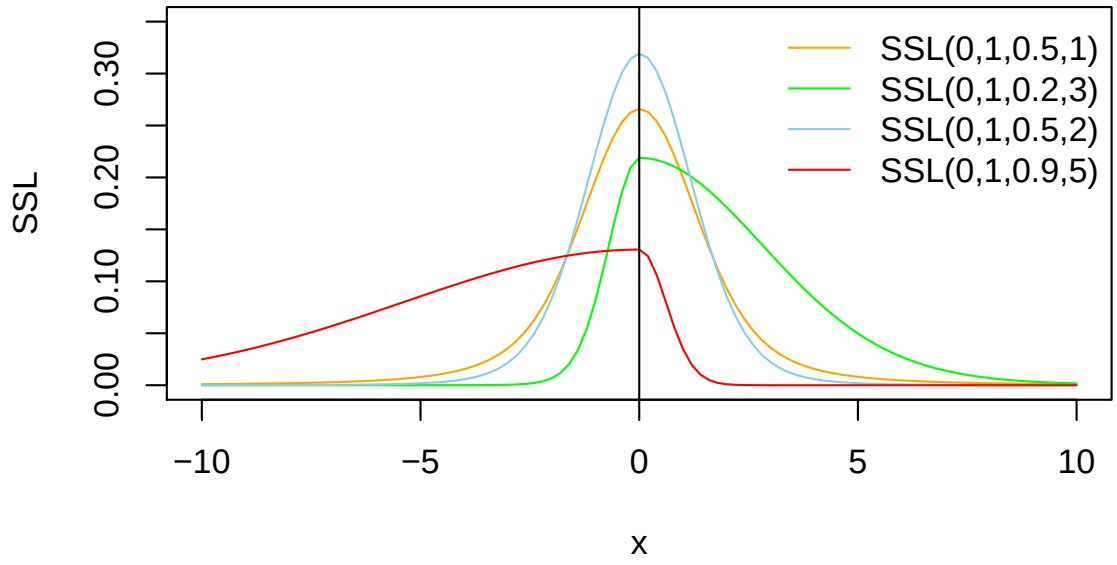


Figure 4.4: The PDF of the SSL distribution in Equation 4.13 with parameters $\xi = 0, \omega^2 = 1$ and combinations of λ and ν.

Figure 4.4 is centred at 0 where it can be observed that the tail is heavier with a lower peak when the value of ν increases.

• **The CF of the SSL distribution**

The CF of the SSL distribution is obtained by the following corollary.

Corollary 4 *From Kim and Genton [16], let Z follow a mixture of SN distributions with PDF from Definition 1. Then, the CF of Z is*

$$\begin{aligned}\psi_Z(t) &= 2\sum_{j=1}^{n} p_j \exp(i\xi t - W(\eta_j)\omega^2 t^2/2)\Phi(i\delta\sqrt{W(\eta_j)}\omega t) \\ &= \sum_{j=1}^{n} p_j \exp(i\xi t - W(\eta_j)\omega^2 t^2/2)\{1 + i\tau(\delta\sqrt{W(\eta_j)}\omega t)^{(1)}\},\end{aligned}$$

where p_j represent the probabilities $p_1, p_2, ..., p_n$ and $W(\eta_j)$ the weight function.
(1) *From Equation 2.17.*

The CF for the StN distribution is obtained from the Laplace-Stieltjes transform and the CF of the skew-t distribution using scale mixtures of skew-normal distributions. For the SSL distribution the weight function is equal to $W(\eta) = \frac{1}{\eta}^{2/(1-p)}$ [16], then the CF for the SSL distribution follows from Corollary 4,

$$\begin{aligned}\psi_Z(t) &= 2p\exp\left(i\xi t - \frac{1}{\eta}^{2/(1-p)}\omega^2 t^2/2\right)\Phi\left(i\delta\sqrt{\frac{1}{\eta}^{2/(1-p)}}\omega t\right) \\ &= p\exp\left(i\xi t - \frac{1}{\eta}^{2/(1-p)}\omega^2 t^2/2\right)\left\{1 + i\tau\left(\delta\sqrt{\frac{1}{\eta}^{2/(1-p)}}\omega t\right)\right\}.\end{aligned}$$

4.2.4 The wrapped skew-slash (WSSL) distribution

By wrapping the SSL distribution in Equation 4.13 onto the unit circle, the following PDF is obtained where $\Theta = Z(mod 2\pi)$

$$f(\theta;\xi,\omega^2,\lambda,\nu) = 2\nu\sum_{r=-\infty}^{\infty}\Phi(\lambda(\frac{\theta+2\pi(r)-\xi}{\omega}))\int_0^1 a^{\nu-1}\phi(\theta+2\pi(r)|\xi,\frac{\omega^2}{a})da, \quad (4.14)$$

for $0 \leq \theta \leq 2\pi$, $-\infty < \xi < \infty$, $-\infty < \lambda < \infty$ and $\nu > 0$, which is denoted by $\Theta \sim WSSL(\xi,\omega^2,\lambda,\nu)$.

Figure 4.5 shows the PDF of the WSSL distribution in Equation 4.14 with parameters $\xi = 0, \omega^2 = 1$ and different values for λ and ν.

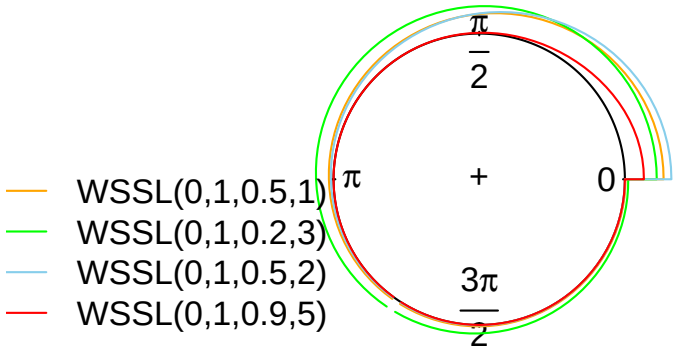


Figure 4.5: The PDF of the WSSL distribution in Equation 4.14 with parameters $\xi = 0, \omega^2 = 1$ and combinations of λ and ν.

In Figure 4.5 it can be observed that as the value of ν increases, the "bulge" of the PDF increases. Also, when λ decreases the PDF tends to be much heavier around the circle.

- **The CF of the WSSL distribution**

From Equation 4.10, the WSSL distribution is related to the WSN distribution by the following stochastic representation:

$$Z = \xi + (1/\eta^{2/(1-p)})^{1/2}(\theta + 2\pi r), \tag{4.15}$$

where η has an inverse-Gamma distribution, $\eta \sim IG(\nu/2, \nu/2)$. The PDF of η is

$$\frac{(\nu/2^{\nu/2})}{\Gamma(\nu/2)}\eta^{-\nu/2-1}\exp\{-\nu/(2\eta)\},$$

where $\eta > 0$. To apply the Laplace-Stieltjes transform, Equation 4.15 is used.

Similarly, from Corollary 4 and by applying Theorem 10, Appendix A.2 the CF for the

WSSL distribution can be defined as follows

$$\psi_Z(p) = 2p\exp\Big(i\xi p - \frac{1}{\eta}^{2/(1-p)}\omega^2 p^2/2\Big)\Phi\Big(i\delta\sqrt{\frac{1}{\eta}^{2/(1-p)}}\omega p\Big)$$
$$= p\exp\Big(i\xi p - \frac{1}{\eta}^{2/(1-p)}\omega^2 p^2/2\Big)\Big\{1 + i\tau\Big(\delta\sqrt{\frac{1}{\eta}^{2/(1-p)}}\omega p\Big)\Big\}^{(1)}$$

for $p = 0, 1, \ldots$.

(1) From Equation 2.17.

4.2.5 The skew-contaminated normal (SCN) distribution

The SCN distribution is denoted by SCN$(\xi, \omega^2, \lambda, \nu, \gamma)$, $0 < \nu < 1$, $0 < \gamma < 1$, where A is now a discrete random variable taking one of two states and $\kappa(a) = 1/a$ [8]. Given the parameter vector $\tau = (\nu, \gamma)^T$, the PDF of A is denoted by $h(a;\tau) = \nu \parallel_{(a=\gamma)} +(1 - \nu) \parallel_{(a=1)}$, $\tau = (\nu, \gamma)^T$. Then, it follows that

$$f(z;\xi,\omega^2,\lambda,\nu,\gamma) = 2\{\nu\phi(z|\xi,\frac{\omega^2}{\gamma}) + (1-\nu)\phi(z|\xi,\omega^2)\}\Phi(\lambda\frac{(z-\xi)}{\omega}). \tag{4.16}$$

The SCN distribution reduces to the SN distribution when $\gamma \to 1$ [8]. From da Silva Ferreira et al. [8], the mean and variance of Z are given by

$$E[Z] = \xi + b\omega\lambda\Big(\frac{\nu}{(\gamma(\gamma+\lambda^2))^{1/2}} + \frac{1-\gamma}{(1+\lambda^2)^{1/2}}\Big),$$

$$Var[Z] = \omega^2\Big[\frac{\nu}{\gamma} + 1 - /nu - b^2\lambda^2\Big(\frac{\nu}{(\gamma(\gamma+\lambda^2))^{1/2}} + \frac{1-\gamma}{(1+\lambda^2)^{1/2}}\Big)^2\Big].$$

Figure 4.6 shows the PDF of the SCN distribution in Equation 4.16 with parameters $\xi = 0, \omega^2 = 1$ and different values for λ , ν and γ.

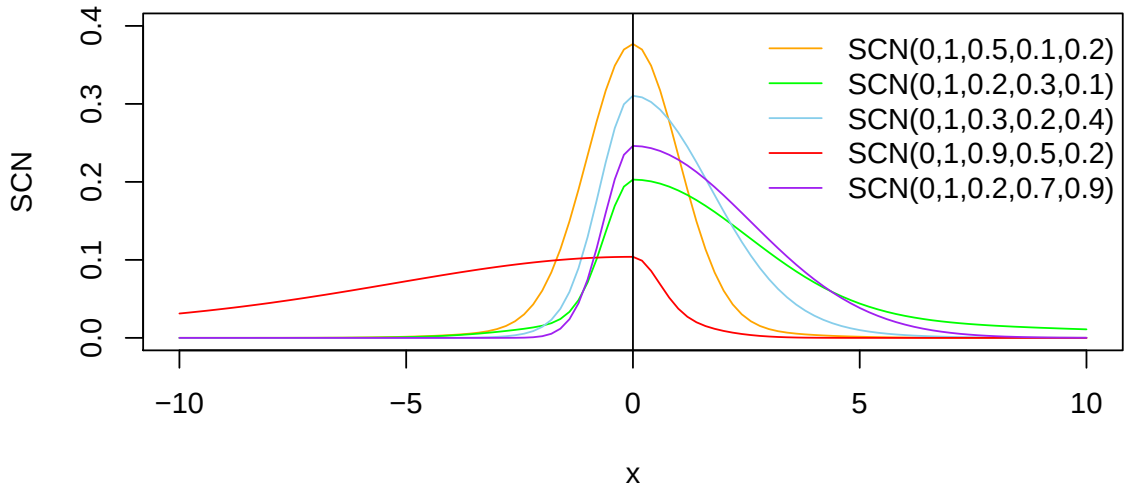


Figure 4.6: The PDF of the SCN distribution in Equation 4.16 with parameters $\xi = 0, \omega^2 = 1$ and combinations of λ , ν and γ.

Figure 4.6 is centred at 0 where it can be observed that the tail is heavier with a lower peak when the value of ν increases. When λ increases the tail becomes much heavier.

- **The CF of the SCN distribution**

The CF of the SCN distribution is a special case of Corollary 4, since $W(\eta) = 1/\eta$ and H is a discrete measure on $\{\eta_1 = \gamma, \eta_2 = 1\}$ with probabilities p and $1 - p$, respectively [16]. The CF of the SCN distribution, using Corollary 4, is then defined as follows

$$\begin{aligned}\psi_Z(t) &= 2\exp(i\xi t)[p\exp(-\gamma^{-1}\omega^2 t^2/2)\Phi(i\delta\gamma^{-1/2}\omega t) + (1-p)\exp(-\omega^2 t^2/2)\Phi(i\delta\omega t)] \\ &= \exp(i\xi t)[p\exp(-\gamma^{-1}\omega^2 t^2/2)\{1 + i\tau(\delta\gamma^{-1/2}\omega t)\} \\ &+ (1-p)\exp(-\omega^2 t^2/2)\{1 + i\tau(\delta\omega t)\}].\end{aligned}$$

4.2.6 The wrapped skew-contaminated normal (WSCN) distribution

By wrapping the SCN distribution in Equation 4.16 onto the unit circle, the following PDF is obtained where $\Theta = Z(mod 2\pi)$

$$f(\theta;\xi,\omega^2,\lambda,\nu,\gamma) = 2\sum_{r=-\infty}^{\infty}\{\nu\phi(\theta+2\pi(r)|\xi,\frac{\omega^2}{\gamma}) + (1-\nu)\phi(\theta+2\pi(r)|\xi,\omega^2)\} \quad (4.17)$$

$$\times\,\Phi(\lambda(\frac{\theta+2\pi(r)-\xi}{\omega})), \quad (4.18)$$

for $0 \leq \theta \leq 2\pi$, $-\infty < \xi < \infty$, $-\infty < \lambda < \infty$, $0 < \nu < 1$ and $0 < \gamma < 1$, which is denoted by $\Theta \sim WSCN(\xi,\omega^2,\lambda,\nu,\gamma)$.

Figure 4.7 shows the PDF of the WSCN distribution in Equation 4.17 with parameters $\xi = 0, \omega^2 = 1$ and different values for λ, ν and γ.

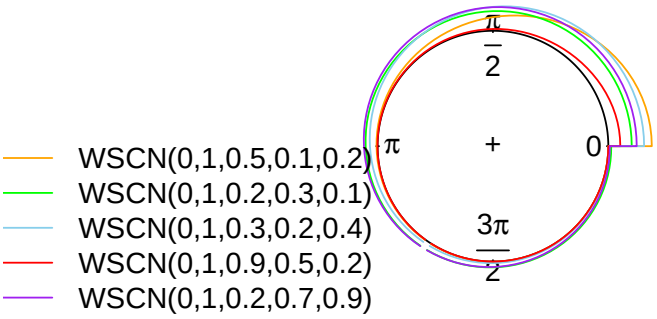


Figure 4.7: The PDF of the WSCN distribution in Equation 4.17 with parameters $\xi = 0, \omega^2 = 1$ and different values (labeled on the figure) for λ, ν and γ.

In Figure 4.7 it can be observed that as the value of ν increases, the "bulge" of the PDF increases. Also, when γ increases the WSCN distribution reduces to the SN distribution.

• **The CF of the WSCN distribution**

The WSCN distribution is related to the WSN distribution, from Equation 4.10, by the following stochastic representation:

$$Z = \xi + (1/\eta)^{1/2}(\theta + 2\pi r), \tag{4.19}$$

where η has an inverse-Gamma distribution, $\eta \sim IG(\nu/2, \nu/2)$. The PDF of η is

$$\frac{(\nu/2^{\nu/2})}{\Gamma(\nu/2)}\eta^{-\nu/2-1}\exp\{-\nu/(2\eta)\},$$

where $\eta > 0$. To apply the Laplace-Stieltjes transform, Equation 4.19 is used.

Similarly, the CF of the WSCN distribution is a special case of Corollary 4, since $W(\eta) = 1/\eta$ and H is a discrete measure on $\{\eta_1 = \gamma, \eta_2 = 1\}$ with probabilities p and $1-p$, respectively [16]. The CF of the WSCN distribution, using Corollary 4 and Theorem 10, Appendix A.2 is then defined as follows

$$\begin{aligned}\psi_Z(p) &= 2\exp(i\xi p)[p\exp(-\gamma^{-1}\omega^2p^2/2)\Phi(i\delta\gamma^{-1/2}\omega p) + (1-p)\exp(-\omega^2p^2/2)\Phi(i\delta\omega p)] \\ &= \exp(i\xi p)[p\exp(-\gamma^{-1}\omega^2p^2/2)\{1 + i\tau(\delta\gamma^{-1/2}\omega p)\} \\ &+ (1-p)\exp(-\omega^2p^2/2)\{1 + i\tau(\delta\omega p)\}]\end{aligned}$$

for $p = 0, 1, \ldots$.

4.3 Application

The dataset used for this application is the same as discussed in Chapter 3, Section 3.3, which consists of a 1000 hourly observations of wind direction in the Atlantic coast of Galicia (NW-Spain). The goal in Section 4.3 is to compare the WN, WSN and WFGSN distributions (from Chapter 3, Section 3.3) to the WStN, WSSL and WSCN distributions (Section 4.2), for the Galicia wind direction data. Estimation of the parameters are obtained using MLE. Circular moments of the WStN, WSSL and WSCN distributions exist, but are not available in closed form, therefore it is not possible to use the method of trigonometric moments in comparison to Section 2.6.3. Several different starting points

were used to implement an optimisation algorithm in order to avoid local maxima [12]. A small number of local maxima is obtained when starting from many different points, where the one corresponding to the maximum value of the likelihood function is selected. Table 4.1 represents the parameter estimates and the standard errors for the parameters of the WStN, WSSL and WSCN distributions.

Table 4.1: Estimates and the standard errors (in parenthesis) for the parameters of the WStN, WSSL and WSCN distributions fitted to the Galicia wind direction data.

Distribution	ξ	ω^2	λ	ν	γ
WStN	-1.459	2.358	-0.250	34.577	-
	(0.231)	(0.048)	(0.141)	(2.699)	-
WSSL	5.200	1.000	0.813	10.000	-
	(1.057)	(0.801)	(0.186)	(3.288)	-
WSCN	5.262	1.000	0.824	0.010	0.990
	(0.493)	(0.406)	(0.088)	(0.121)	(0.012)

Figure 4.8 illustrates a raw circular data plot and rose diagram of the Galicia wind direction data together with the WStN, WSSL and WSCN PDF fitted by MLE. This figure can be compared to the figure in Section 3.3, Figure 3.15.

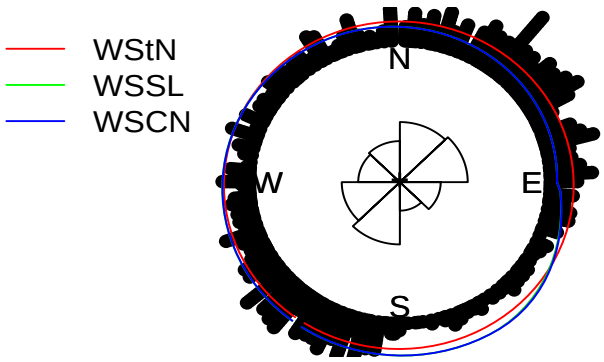


Figure 4.8: Raw circular data plot and rose diagram for the Galicia wind direction data together with the WStN, WSSL and WSCN PDF fitted by MLE.

Table 4.2 shows the maximised log-likelihood (ℓ_{max}) and AIC/BIC values for the WN, WSN, WFGSN, WStN, WSSL and WSCN distributions fitted to the Galicia wind direction data. The AIC/BIC is calculated as $-2(\ell_{max} - p)$ and $-2\ell_{max} + \log(n)p$ respectively, where p is the number of parameters estimated and n the sample size[12].

Table 4.2: ℓ_{max} and AIC/BIC values for the WN, WSN, WFGSN, WStN, WSSL and WSCN distributions fitted to the Galicia wind direction data where the minimum AIC/BIC value is identified using bold type.

Distribution	ℓ_{max}	AIC	BIC
WN	-1803.424	3602.849	3593.069
WSN	-1842.98	3679.96	3665.285
WFGSN	-1867.77	3727.54	3707.973
WStN	-1803.459	**3598.917**	**3579.351**
WSSL	-1923.783	3839.565	3819.999
WSCN	-1921.290	3832.580	3808.122

The comparison based on the AIC/BIC value has identified that the WStN distribution is the best fit for the Galicia wind direction data.

4.4 Summary

In this chapter, the PDF of the SMN distributions was revisited as well as the SSMN distributions. Examples of SSMN distributions were provided, such as the StN distribution, SSL distribution and the SCN distribution. The PDF and CF of these distributions were revisited where the wrapped version of each distribution was defined. Graphical representations of the PDF and the wrapped PDF of these distributions were provided. Lastly, the WStN distribution, WSSL distribution and the WSCN distribution were compared to the WN, WSN and WFGSN distributions in Section 3.3, where the AIC/BIC value has identified that the WStN distribution is the best fit for the Galicia wind direction data. In the following chapter a conclusion of the study is provided.

Chapter 5

Conclusion

Directional data applied in practice are often asymmetric and bimodal which motivates the study of wrapped distributions. In Chapter 1 the use and importance of the WN distribution along with distributions that can incorporate skewness and bimodality, such as the WSN and WFGSN distributions were motivated. The importance of the SSMN distribution as well as the wrapped versions of these distributions were also highlighted.

In Chapter 2 the wrapped distribution and the WN distribution were revisited. The SN and the WSN distributions were also investigated. The MLE and the method of trigonometric moments were also defined, where a simulation study consisting of two examples for each method was conducted for the purpose of comparison. The simulation study suggested that MLE gave slightly more accurate results than the method of trigonometric moments, where the bias and the standard error of the trigonometric moments can be reduced by using the Monte Carlo approximation method.

In Chapter 3 the FGSN and the WFGSN distributions were defined. The WN, WSN and WFGSN distributions were also compared with wind direction data as an application where the estimates of the parameters were obtained using MLE. A comparison based on the AIC/BIC has identified the WN distribution to fit the Galicia wind direction data best. The WSN and WFGSN distribution will provide a better fit to circular data that contains a higher level of skewness and bimodality.

In Chapter 4 examples of SSMN distributions were provided, such as the StN distribution, SSL distribution and the SCN distribution. The wrapped versions of each distribution were used to improve the fit for the Galicia wind direction data. The estimates of the parameters were obtained using MLE where the AIC/BIC value has indeed identified that the WStN distribution is the best fit for the Galicia wind direction data.

For future research it can be feasible to consider another skewing mechanism to improve the fit for the Galicia wind direction data. It would also be possible to wrap other distributions, such as the skew-exponential power distribution discussed by Da Silva Ferreira et al. [8], where the wrapped skew-exponential power distribution might describe the shape of the observed distribution more effectively.

Bibliography

[1] C Agostinelli and U Lund. *R package circular: Circular Statistics (version 0.4-93)*, 2017.

[2] Barry C Arnold and Gwo Dong Lin. Characterizations of the skew-normal and generalized chi distributions. *Sankhyā: The Indian Journal of Statistics*, 66(4):593–606, 2004.

[3] Adelchi Azzalini. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12(2):171–178, 1985.

[4] Adelchi Azzalini. *The skew-normal and related families*, volume 3. Cambridge University Press, 2013.

[5] Adelchi Azzalini and Antonella Capitanio. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):367–389, 2003.

[6] Adelchi Azzalini and A Dalla Valle. The multivariate skew-normal distribution. *Biometrika*, 83(4):715–726, 1996.

[7] L.J. Bain and M. Engelhardt. *Introduction to probability and mathematical statistics*. Thomson Learning, 1992.

[8] Clécio da Silva Ferreira, Heleno Bolfarine, and Víctor H Lachos. Skew scale mixtures of normal distributions: properties and estimation. *Statistical Methodology*, 8(2):154–171, 2011.

[9] Jay L Devore and Kenneth N Berk. *Modern mathematical statistics with applications*. Springer, 2012.

[10] David Elal-Olivero, Héctor W Gómez, and Fernando A Quintana. Bayesian modeling using a class of bimodal skew-elliptical distributions. *Journal of Statistical Planning and Inference*, 139(4):1484–1492, 2009.

[11] Héctor W Gómez, Osvaldo Venegas, and Heleno Bolfarine. Skew-symmetric distributions generated by the distribution function of the normal distribution. *Environmetrics: The official journal of the International Environmetrics Society*, 18(4):395–407, 2007.

[12] Estefanía Hernández-Sánchez and Bruno Scarpa. A wrapped flexible generalized skew-normal model for a bimodal circular distribution of wind directions. *Chilean Journal of Statistics (ChJS)*, 3(2):129–141, 2012.

[13] Simon Hurst. The characteristic function of the student t distribution. *Research report: statistics research report/Centre for mathematics and its applications (Canberra)*, 1995.

[14] S Rao Jammalamadaka and Tomasz J Kozubowski. A general approach for obtaining wrapped circular distributions via mixtures. *Sankhya A*, 79(1):133–157, 2017.

[15] K Jayakumar and Sophy Jacob. Wrapped skew Laplace distribution on integers: A new probability model for circular data. *Open Journal of Statistics*, 2(1):106–114, 2012.

[16] Hyoung-Moon Kim and Marc G Genton. Characteristic functions of scale mixtures of multivariate skew-normal distributions. *Journal of Multivariate Analysis*, 102(7):1105–1117, 2011.

[17] Alan Lee. Circular data. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4):477–486, 2010.

[18] Yanyuan Ma and Marc G Genton. Flexible class of skew-symmetric distributions. *Scandinavian Journal of Statistics*, 31(3):459–468, 2004.

[19] Kanti V Mardia and Peter E Jupp. *Directional statistics*, volume 494. John Wiley & Sons, 2009.

[20] Gianluca Mastrantonio, Alan E Gelfand, and Giovanna Jona Lasinio. The wrapped skew Gaussian process for analyzing spatio-temporal data. *Stochastic Environmental Research and Risk Assessment*, 30(8):2231–2242, 2016.

[21] John A Nelder and Roger Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.

[22] M Oliveira, RM Crujeiras, and A Rodrguez-Casal. Npcirc: Nonparametric circular methods. *R package version*, 2013.

[23] María Oliveira, Rosa M. Crujeiras, and Alberto Rodríguez-Casal. NPCirc: An R package for nonparametric circular methods. *Journal of Statistical Software*, 61(9):1–26, 2014.

[24] Arthur Pewsey. Problems of inference for Azzalini's skew-normal distribution. *Journal of Applied Statistics*, 27(7):859–870, 2000.

[25] Arthur Pewsey. The wrapped skew-normal distribution on the circle. *Communications in Statistics-Theory and Methods*, 29(11):2459–2472, 2000.

[26] Arthur Pewsey. Modelling asymmetrically distributed circular data using the wrapped skew-normal distribution. *Environmental and Ecological Statistics*, 13(3):257–269, 2006.

[27] Jiuzhou Wang, Joseph Boyer, and Marc G Genton. A skew-symmetric representation of multivariate distributions. *Statistica Sinica*, 14(4):1259–1270, 2004.

[28] Stephen Wright and Jorge Nocedal. Numerical optimization. *Springer Science*, 35(67-68):7, 1999.

Appendix A

Definitions and results

This Appendix contains a list of additional results and definitions referenced in this study.

A.1 Definitions

Definition 2 *Euler's formula states that for any real number x,*

$$\exp(i\xi p) = \cos p\xi + i \sin p\xi.$$

Definition 3 *The sine and cosine of a difference is defined as*

$$\cos(A - B) = \cos A \cos B + \sin A \sin B$$
$$\sin(A - B) = \sin A \cos B - \cos A \sin B.$$

Definition 4 *The following trigonometric identity is defined as*

$$\sin^2(x) + \cos^2(x) = 1.$$

Definition 5 *From [7], the expected value of the random variable X, when X is conditional on Y, is equal to the following*

$$E[X] = E[E[X|Y]].$$

Definition 6 *Suppose that we are given unit vectors $x_1, ..., x_n$ with corresponding angles θ_i, $i = 1, ..., n$. The mean direction μ of $\theta_1, ..., \theta_n$ is the direction of the resultant $x_1 + ... + x_n$ of $x_1, ..., x_n$. It is also the direction of the centre of mass $\bar{x}$ of $x_1, ..., x_n$. Since the Cartesian coordinates of x_j, for $j = 1, ..., n$, are $(\cos(\theta), \sin(\theta))$, the Cartesian coordinates of the centre of mass are $(\bar{C}, \bar{S})$ where*

$$\bar{C} = \frac{1}{n}\sum_{j=1}^{n}\cos(\theta)$$

$$\bar{S} = \frac{1}{n}\sum_{j=1}^{n}\sin(\theta)$$

Therefore μ is the solution of the equations

$$\bar{C} = \rho\cos(\theta)$$

$$\bar{S} = \rho\sin(\theta)$$

(provided that $\rho > 0$), where the mean resultant length ρ is given by

$$\rho = (\bar{C}^2 + \bar{S}^2)^{\frac{1}{2}}$$

[19].

Definition 7 *As stated in [19], the moments*

$$\bar{C} = \frac{1}{n}\sum_{j=1}^{n}\cos(\theta_i)$$

$$\bar{S} = \frac{1}{n}\sum_{j=1}^{n}\sin(\theta_i)$$

play key roles in defining the sample mean direction and the sample circular variance. It is useful to combine them into the first trigonometric moment about the zero direction

$$m_1^{'} = \bar{C} + i\bar{S}.$$

Then

$$m_1^{'} = \bar{R}\exp(i\bar{\theta}).$$

Extending the notation, the pth trigonometric moment about the zero direction for $p = 1, 2, ...$ *is defined as*

$$m_p^{'} = a_p + ib_p,$$

where

$$a_p = \frac{1}{n}\sum_{j=1}^{n} \cos p(\theta_j)$$

$$b_p = \frac{1}{n}\sum_{j=1}^{n} \sin p(\theta_j).$$

Then

$$m_p^{'} = \bar{R}_p \exp(i\bar{\theta}_p), \tag{A.1}$$

where $\bar{\theta}_p$ *and* $\bar{R}_p$ *denote the sample mean direction and sample mean resultant length of* $p\theta_1, ..., p\theta_n$*. The pth trigonometric moment about the mean direction is*

$$m_p = a_p + ib_p, \tag{A.2}$$

where

$$a_p = \frac{1}{n}\sum_{j=1}^{n} \cos p(\theta_i - \bar{\theta})$$

$$b_p = \frac{1}{n}\sum_{j=1}^{n} \sin p(\theta_i - \bar{\theta}).$$

Definition 8 *From [19], the pth trigonometric moment about the zero mean direction is defined as*

$$\psi_p = \rho_p \exp(i\mu_p),\ \rho_p \geq 0$$

as the population version of equation (A.1). The pth trigonometric moment about the zero mean direction is defined by analogy with equation (A.2) as

$$\psi_p = \alpha_p + i\beta_p \tag{A.3}$$

where

$$\alpha_p = E[\cos p\Theta] = \int_0^{2\pi} \cos p\theta dF(\Theta),$$

$$\beta_p = E[\sin p\Theta] = \int_0^{2\pi} \sin p\theta dF(\Theta).$$

Definition 9 *As defined in [19], the complex numbers* $\{\psi_p : p = 0, \pm 1, ...\}$*, equation (A.3), are the Fourier coefficients of* F*. When the* ψ_p *are related to* F *by*

$$\psi_p = E[\exp(ip\Theta)] = \int_0^{2\pi} \exp(ip\theta) dF(\theta)\ p = 0, \pm 1, ...,$$

it is usual to write

$$dF(\theta) \sim \frac{1}{2\pi} \sum_{p=-\infty}^{\infty} \psi_p \exp(-ip\theta).$$

If $\sum_{p=1}^{\infty}(\alpha_p^2 + \beta_p^2)$ *is convergent then the random variable* θ *has a PDF* f *which is defined almost everywhere by*

$$f(\theta) = \frac{1}{2\pi} \sum_{p=-\infty}^{\infty} \psi_p \exp(-ip\theta). \tag{A.4}$$

This result is an analogue on the unit circle of the inversion theorem for continuous random variables on the real line. Equation (A.4) can then be written as

$$f(\theta) = \frac{1}{2\pi}[1 + 2\sum_{p=1}^{\infty}(\alpha_p \cos p\theta + \beta_p \sin p\theta)].$$

Definition 10 *As stated in Arnold et al. [2], the moment generating function (MGF)* $M_Y(t)$*, where* $Y \sim SN(\xi, \omega^2, \lambda)$*, can be written as*

$$\begin{aligned} M_Y(t) &= 2\exp\left(\frac{t^2}{2}\right)\Phi(\delta t) \\ &= 2\exp\left(\frac{t^2}{2}\right)\left(\frac{1}{2} + \int_0^{\delta t} \phi(x)dx\right), \end{aligned}$$

where $t \in \mathbb{R}$ *and* $X \sim SN(0, 1, \lambda)$*.*

Definition 11 *The circular mean and concentration of* Θ *are given by [19],*

$$\begin{aligned} \tilde{\xi} &= \arctan(\frac{\alpha_1}{\beta_1}), \\ \tilde{c} &= \sqrt{\alpha_1^2 + \beta_1^2}. \end{aligned}$$

Definition 12 *A random variable* A *has the gamma distribution if its PDF is given by*

$$f_A(a) = \frac{(\nu/2)^{\nu/2}}{\Gamma(\frac{\nu}{2})} a^{\nu/2-1} e^{-\nu/2a}, \quad a > 0 \tag{A.5}$$

where $\nu > 0$*. This is denoted by* $A \sim Gamma(\nu/2, \nu/2)$ *[7].*

A.2 Results

Theorem 8 *If the function $\phi(z), z = x + iy$, is defined by integration along a path in the complex plane parallel with the x-axis from $-\infty + iy$ to $x + iy$, then:*

(i) $\phi(z)$ is convergent;
(ii) $\phi(z) = \phi(x + iy) = \exp(\frac{y^2}{2}) \int_{-\infty}^{x} \exp(-ity)\phi(t)dt$, where $\phi(t)$ is the standard normal PDF;
(iii) When $x = 0$:

$$\phi(iy) = \frac{1}{2} + \frac{i}{\sqrt{\pi}} \int_0^{y/\sqrt{2}} \exp(t^2)dt$$

[16].

Theorem 9 *From [4], if $U \sim N(0, 1)$ then*

$$E\{\Phi(hU + k)\} = \Phi(\frac{k}{\sqrt{1 + h^2}}),$$

where $h, k \in \mathbb{R}$. From this result, the MGF of Y, where $Y \sim SN(\xi, \omega^2, \lambda)$, is readily obtained, that is

$$\begin{aligned} M_Y(t) &= E[\exp(\xi t + \omega x t)] \\ &= 2\exp(\xi t + \omega^2 t^2/2) \int_{\mathbb{R}} \phi(x - \omega t)\Phi(\lambda x)dx \\ &= 2\exp(\xi t + \omega^2 t^2/2)\Phi(\delta \omega t) \end{aligned} \tag{A.6}$$

where

$$\delta = \delta(\lambda) = \lambda/\sqrt{1 + \lambda^2},\ \delta \in (-1, 1),$$

and $X \sim SN(0, 1, \lambda)$. Multiplication of (A.6) by the MGF of the $N(\mu, \sigma^2)$ distribution, $\exp(\mu t + \sigma^2 t^2/2)$ is still a function of type (A.6) [4].

Theorem 10 *As stated in [19], if y is a random variable on the line then the corresponding random variable y_w of the wrapped distribution is given by*

$$x_w = x(mod 2\pi).$$

If x has a distribution function F then the distribution function F_w of y_w is given by

$$F_w(\theta) = \sum_{k=-\infty}^{\infty} [F(\theta + 2\pi(k)) - F(2\pi(k))], \quad 0 \leq \theta \leq 2\pi.$$

If the CF of y is ψ then the characteristic function $\{\psi_p : p = 0, \pm 1, ...\}$ of y_w is given by

$$\psi_p = \psi(p).$$

Proof.

$$\begin{aligned}
\psi_p &= \int_0^{2\pi} \exp(i\Theta p) dF_w(\theta) \\
&= \sum_{k=-\infty}^{\infty} \int_{2\pi_k}^{2\pi k+1} \exp(i\Theta p) dF(\theta) \\
&= \int_{-\infty}^{\infty} \exp(iyp) dF(y) \\
&= \psi(p).
\end{aligned}$$

Theorem 11 *If ψ is integrable then x has a PDF and*

$$\begin{aligned}
f_w(\theta) &= \sum_{k=-\infty}^{\infty} g(\theta + 2\pi k) \\
&= \frac{1}{2\pi}[1 + 2\sum_{p=1}^{\infty}(\alpha_p \cos p\theta + \beta_p \sin p\theta)]
\end{aligned}$$

where $\psi(p) = \alpha_p + i\beta_p$ [19].

Theorem 12 *Azzalini and Capitanio [5] proposed a stochastic representation of the flexible generalised skew-normal distribution that provides a method that is useful for simulation from a $FGSN(\xi, \omega^2, \lambda, \beta)$ with PDF defined in Equation 3.1.*

Let $\boldsymbol{X}$ be a continuous random vector with PDF $f(\boldsymbol{x})$ and let U be a uniform random variable on $(0, 1)$, independent of $\boldsymbol{X}$ [27]. A random vector $\boldsymbol{Y}$ can be simulated with the following representation:

$$\boldsymbol{Y} = \begin{cases} \boldsymbol{X} + \xi, & \text{if } U < \pi(\boldsymbol{X}) \\ -\boldsymbol{X} + \xi, & \text{if } U > \pi(\boldsymbol{X}) \end{cases} \tag{A.7}$$

As an example, from Wang et al. [27], let Y have a flexible generalised skew-normal PDF defined in Equation 3.1 and $\boldsymbol{X}$ have a $n-$dimensional normal PDF with mean $\boldsymbol{0}$ and correlation matrix Ω, denoted as $\phi_n(\boldsymbol{x}; \boldsymbol{0}, \Omega)$. Azzalini and Dalla Valle [6] then provide the application of the probability integral transformation to Equation A.7, where the following is then obtained,

$$Y = \begin{cases} \boldsymbol{X} + \xi, & \text{if } W < \lambda^T \boldsymbol{X} \\ -\boldsymbol{X} + \xi, & \text{if } W > \lambda^T \boldsymbol{X} \end{cases} \tag{A.8}$$

where W is $N(0,1)$, which is independent of $\boldsymbol{X}$ [27].

Theorem 13 *Let $Z \sim SSMN(\xi, \omega^2, \lambda, H; \kappa)$. Then its stochastic representation is given by*

$$Z|A = a \sim SN(\xi, \omega^2\kappa(a), \lambda\sqrt{\kappa(a)}),$$

where $A \sim H(a; \tau)$ [8].

Theorem 14 *Suppose that $Z \sim SSMN(\xi, \omega^2, \lambda, H; \kappa)$. Then, for $b = \sqrt{\frac{2}{\pi}}$:*
(a) $E[Z] = \xi + b\omega\lambda E_A[\frac{\kappa(A)}{\sqrt{1+\lambda^2\kappa(A)}}]$,
(b) $Var[Z] = \omega^2\Big(E_A[\kappa(A)] - b^2\lambda^2 E_A^2[\frac{\kappa(A)}{\sqrt{1+\lambda^2\kappa(A)}}]\Big)$ *[8].*

Appendix B

R code used in this study

This Appendix contains all R code used throughout the study.

B.1 Chapter 2

B.1.1 An overlay of the $SN(\xi, \omega^2, \lambda)$ distribution with PDF as given in Equation 2.9

```
library(circular)
library(Wrapped)
library(NPCirc)
library(sn)

dsn <- function(y,xi,omega,lambda,beta)
(2/omega)*dnorm((y-xi)/omega, 0, 1)*pnorm(lambda*(y-xi)/omega, 0, 1)
wsn <-function(x) dsn(y=x, xi=0, omega=1, lambda=4)
PDF <- Vectorize(wsn)
wsn1 <-function(x) dsn(y=x, xi=0, omega=1, lambda=1)
PDF1 <- Vectorize(wsn1)
wsn2 <-function(x) dsn(y=x, xi=0, omega=1, lambda=-4)
PDF2 <- Vectorize(wsn2)
```

```
wsn3 <-function(x) dsn(y=x, xi=0, omega=1, lambda=-1)
PDF3 <- Vectorize(wsn3)

plot.function(PDF,xlim=c(-4,4),col=4,lwd=1,xlab="y", ylab="SN")
plot(PDF1,xlim=c(-4,4),col=3,lwd=1,add=TRUE)
plot(PDF2,xlim=c(-4,4), col=2,lwd=1,add=TRUE)
plot(PDF3,xlim=c(-4,4), col="orange",lwd=1,add=TRUE)
legend(-4.2,0.65,legend=c("SN(4)","SN(1)","SN(-4)",
"SN(-1)"), bty="n",col=c("blue","green","red","orange"),
lty=c(1,1,1,1) ,ncol=1)
```

B.1.2 Visualisation of the skew-normal sampling scheme in Section 2.2.1

```
x <- rnorm(10000, mean=0, sd=1)
w <- rnorm(10000, mean=0, sd=1)
xi <- 0
omega <- sqrt(1)
lambda <- 4
dat <- xi+(omega*lambda)/sqrt(1+lambda^2)*abs(x)+
(omega/sqrt(1+lambda^2))*w-(omega*lambda*sqrt(2)/(pi*(1+lambda^2)))
theopdf <- function(y,xi,omega,lambda)
(2/omega)*dnorm((y-xi)/omega, 0, 1)*pnorm(lambda*(y-xi)/omega, 0, 1)
hist(dat, probability = TRUE,main = "",xlab="x",ylab="PDF",
ylim = c(0, 0.8),col="blue")
xfit<-seq(-8,8,0.05)
yfit<-theopdf(xfit,xi,omega,lambda)
lines(xfit,yfit,col="red",lwd=3)

xi <- 0
omega <- sqrt(1)
lambda <- 1
```

```
dat <- xi+(omega*lambda)/sqrt(1+lambda^2)*abs(x)+
(omega/sqrt(1+lambda^2))*w-(omega*lambda*sqrt(2)/(pi*(1+lambda^2)))
theopdf <- function(y,xi,omega,lambda)
(2/omega)*dnorm((y-xi)/omega, 0, 1)*pnorm(lambda*(y-xi)/omega, 0, 1)
hist(dat, probability = TRUE,main = "",xlab="x",ylab="PDF",
ylim = c(0, 0.6),col="blue")
xfit<-seq(-8,8,0.05)
yfit<-theopdf(xfit,xi,omega,lambda)
lines(xfit,yfit,col="red",lwd=3)

xi <- 0
omega <- sqrt(1)
lambda <- -4
dat <- xi+(omega*lambda)/sqrt(1+lambda^2)*abs(x)+
(omega/sqrt(1+lambda^2))*w-(omega*lambda*sqrt(2)/(pi*(1+lambda^2)))
theopdf <- function(y,xi,omega,lambda)
(2/omega)*dnorm((y-xi)/omega, 0, 1)*pnorm(lambda*(y-xi)/omega, 0, 1)
hist(dat, probability = TRUE,main = "",xlab="x",ylab="PDF",
ylim = c(0, 0.8),col="blue")
xfit<-seq(-8,8,0.05)
yfit<-theopdf(xfit,xi,omega,lambda)
lines(xfit,yfit,col="red",lwd=3)

xi <- 0
omega <- sqrt(1)
lambda <- -1
dat <- xi+(omega*lambda)/sqrt(1+lambda^2)*abs(x)+
(omega/sqrt(1+lambda^2))*w-(omega*lambda*sqrt(2)/(pi*(1+lambda^2)))
theopdf <- function(y,xi,omega,lambda)
(2/omega)*dnorm((y-xi)/omega, 0, 1)*pnorm(lambda*(y-xi)/omega, 0, 1)
hist(dat, probability = TRUE,main = "",xlab="x",ylab="PDF",
```

```
ylim = c(0, 0.6),col="blue")
xfit<-seq(-8,8,0.05)
yfit<-theopdf(xfit,xi,omega,lambda)
lines(xfit,yfit,col="red",lwd=3)
```

B.1.3 Examples of the $WSN(\xi, \omega^2, \lambda)$ PDF given in Equation 2.15

```
set.seed(2012)
#Only need to change parameter values
wsn <- function(x) dwsn(x, xi=0, eta=1, lambda=-1)
curve.circular(wsn,n=500,xlim=c(-1.65,1.65),main=NULL)
```

B.1.4 Simulation study: Method of trigonometric moments and maximum likelihood estimation

```
#Method of trigonometric moments
rm(list = ls(all= TRUE))
library(circular)
library(sn)
library(NPCirc)
library(CircStats)

xi = 0 #only need to change parameters
omega = 1
lambda = 0.7
n=1000
m=1000

ARB = MSE = estimates = matrix(0, m, 3)

for (i in 1:m) {
theta <- rwsn(n, xi=xi, eta = omega, lambda = lambda)
```

```
#Calculate trig moments
#cos
alpha1 = as.vector(mean(cos(theta)))
#sin
beta1 = as.vector(mean(sin(theta)))
thetabar = ifelse(alpha1<0, pi+atan(beta1/alpha1), atan(beta1/alpha1))
rbar = (alpha1^2+beta1^2)^(1/2)
bbar2 = as.vector(mean(sin(2*(theta-thetabar))))

#Calculate chi
integrand <- function(u) {(2/pi)^(1/2)*exp(u^2/2)}
chiest <- function(chiest){
  ja <- integrate(integrand, lower = 0, upper = chiest)$val
  t=2*chiest
  jb <- integrate(integrand, lower = 0, upper = t)$val
  func <- (jb*(1-ja^2)-2*(ja))/(1+ja^2)^3-(bbar2/(rbar^4))
  return(func)
}
chi <- uniroot(chiest, lower = -10, upper = 15)$root
#chi

#Calculate omega
jchi <- integrate(integrand, lower = 0, upper = chi)$val
newomega=(-2*log(rbar)+log(1+jchi^2))^(1/2)
#newomega

#Calculate lambda
pro = (chi/newomega)^2
newlambda = sqrt(pro/(1-pro))
#newlambda
```

```
#Calculate xi
newxi = atan((tan(thetabar)-jchi)/(1+jchi*tan(thetabar)))
#newxi

#Final estimates
colnames(estimates) <- c('mu', 'sigma', 'gamma1')
estimates[i,] = c(newxi, newomega, newlambda)
}

avg=colMeans(estimates)
avg
trueval <- c(xi, omega, lambda)

bias <- avg-trueval
bias

sdpar <- apply(estimates, 2, sd)
sdpar

Q <- apply(estimates, 2, quantile)
Q

ECDF <- lapply(1:3, function(i) ecdf(estimates[,i]))
trueQ <- c(ECDF[[1]](xi), ECDF[[2]](omega), ECDF[[3]](lambda))
trueQ

Pml1 = ecdf(estimates[,1])
plot(Pml1, main = "")

Pml2 = ecdf(estimates[,2])
```

```
plot(Pml2, main = "")

Pml3 = ecdf(estimates[,3])
plot(Pml3, main = "")

plot(density(estimates[,1]), main = "", ylab="PDF")
plot(density(estimates[,2]), main = "", ylab="PDF")
plot(density(estimates[,3]), main = "", ylab="PDF")

#MLE

#Only need to change parameters
xi <- 0.08521464
omega <- 1.00414198
lambda <- 0.94410201
n <- 1000

a <- rwsn(n, xi=xi, eta=omega, lambda=lambda)

logwsn <- function(par, theta) -sum(log(dwsn(theta,
xi=par[1], eta=par[2], lambda=par[3])))

MLE <- do.call(rbind, lapply(1:n, function(...)
{
  Samp <- sample(n, n, TRUE)
  optim(par=c(xi, omega, lambda), fn=logwsn,
  method=c("Nelder-Mead"), theta=a[Samp])$par
}))


trueVal <- c(xi, omega, lambda)
```

```
avgpar <- apply(MLE, 2, mean)
avgpar

bias <- avgpar-trueVal
bias

sdpar <- apply(MLE, 2, sd)
sdpar

Q <- apply(MLE, 2, quantile)
Q

ECDF <- lapply(1:3, function(i) ecdf(MLE[,i]))
trueQ <- c(ECDF[[1]](xi), ECDF[[2]](omega), ECDF[[3]](lambda))
trueQ

Pml1 = ecdf(MLE[,1])
plot(Pml1, main = "")

Pml2 = ecdf(MLE[,2])
plot(Pml2, main = "")

Pml3 = ecdf(MLE[,3])
plot(Pml3, main = "")

plot(density(MLE[,1]), main = "",ylab="PDF")
plot(density(MLE[,2]), main = "",ylab="PDF")
plot(density(MLE[,3]), main = "",ylab="PDF")
```

B.2 Chapter 3

B.2.1 An overlay of the $FGSN(\xi, \omega^2, \lambda, \beta)$ distribution with PDF as given in Equation 3.1

```
dwgsn <- function(y,xi,omega,lambda,beta)
(2/omega)*dnorm((y-xi)/omega, 0, 1)*pnorm(lambda*(y-xi)/omega+beta*
((y-xi)/omega)^3, 0, 1)
wgnsn <-function(x) dwgsn(y=x, xi=0, omega=1, lambda=4, beta=0)
FGSN <- Vectorize(wgnsn)
wgnsn1 <-function(x) dwgsn(y=x, xi=0, omega=1, lambda=1, beta=-1)
FGSN1 <- Vectorize(wgnsn1)
wgnsn2 <-function(x) dwgsn(y=x, xi=0, omega=1, lambda=4, beta=-1)
FGSN2 <- Vectorize(wgnsn2)
wgnsn3 <-function(x) dwgsn(y=x, xi=0, omega=1, lambda=4, beta=1)
FGSN3 <- Vectorize(wgnsn3)

plot.function(FGSN,xlim=c(-4,4),col=4,lwd=1, xlab="y")
plot(FGSN1,xlim=c(-4,4),col=3,lwd=1,add=TRUE)
plot(FGSN2,xlim=c(-4,4), col=2,lwd=1,add=TRUE)
plot(FGSN3,xlim=c(-4,4), col="orange",lwd=1,add=TRUE)
legend(-4.2,0.68,legend=c("FGSN(0,1,4,0)","FGSN(0,1,1,-1)",
"FGSN(0,1,4,-1)", "FGSN(0,1,4,1)"), bty="n",
col=c("blue","green","red","orange"), lty=c(1,1,1,1))
```

B.2.2 Examples of the $WFGSN(\xi, \omega^2, \lambda, \beta)$ PDF given in Equation 3.4

```
dwgsn <- function(theta,xi,omega,lambda,r)
(2/omega)*sum(dnorm((theta+2*pi*r-xi)/omega, 0, 1)*
pnorm(lambda*(theta+2*pi*r-xi)/omega, 0, 1))
dwgn <- function(theta, xi, omega, lambda, beta, r)
```

```
(2/omega)*sum(dnorm((theta+2*pi*r-xi)/omega, 0, 1)*
pnorm(lambda*(theta+2*pi*r-xi)/omega+
beta*((theta+2*pi*r-xi)/omega)^3, 0, 1))
#Only need to change parameter values
wgnsn <-function(x)
dwgsn(theta=x, xi=0, omega=1, lambda=-1, r=-100:100)
wgn <- function(x)
dwgn(theta=x, xi=0, omega=1, lambda=-1, beta=3, r=-100:100)
vwgnsn <- Vectorize(wgnsn)
vwgn <- Vectorize(wgn)
curve.circular(vwgn,xlim=c(-1.65,1.65))
curve.circular(vwgnsn,xlim=c(-1.65,1.65),add = TRUE,
col=2,lty = 9, lwd = 6)
```

B.2.3 Simulation study: Maximum likelihood estimation

```
rm(list = ls(al=T))
library(NPCirc)
options(scipen = 999)
#Only need to change parameter values
xi <- 0
Omega <- 1
lambda <- 1
beta <- -2

#simulation for FGSN
n=100
x=rnorm(n,0,1)
y=rnorm(n,0,1)
wx=(lambda*x)+(beta*x^3)

z=ifelse(y<wx,x,-x)
```

```
zfin=xi+(Omega*z)

#simulation for WFGSN by applying the transformation
theta = zfin%%(2*pi)

logwfgsn=function (par, x, K = 100, min.k = 20)
{
  x <- conversion.circular(x, units = "radians", zero = 0,
                           rotation = "counter")
  xi <- conversion.circular(xi, units = "radians", zero = 0,
                            rotation = "counter")
  xi = par[1]
  eta = par[2]
  lambda = par[3]
  beta = par[4]
  x <- x[!is.na(x)]
  n <- length(x)
  if (sum(is.na(x)) > 0)
    warning("Missing values were removed")
  if (is.null(K)) {
    range <- abs(xi - x)
    K <- (range + 6 * eta)%/%(2 * pi) + 1
    K <- max(min.k, K)
  }
  else {
   if (!is.numeric(K) | K <= 0) {
    #warning("Argument 'K' must be a positive integer.
    #'K=min.k' was used")
      K <- min.k
    }
  }
```

```
  fx <- numeric(n)
  for (i in 1:n) {
    val <- (x[i] + 2 * pi * seq(-K, K, 1) - xi)/eta
    suma <- sum(dnorm(val) * pnorm(lambda * val + beta * val^3))
    fx[i] <- 2/eta * suma
  }
  return(-sum(log(fx)))
}

MLE <- do.call(rbind, lapply(1:100, function(...)
{
  Samp <- sample(n, n, TRUE)

  x=rnorm(n,0,1)
  y=rnorm(n,0,1)
  wx=(lambda*x)+(beta*x^3)

  z=ifelse(y < wx, x,-x)
  zfin=xi+(Omega*z)
  theta = zfin%%(2*pi)
  #simulation for WFGSN by applying the transformation

  optim(par=c(xi, Omega, lambda, beta),
        fn = logwfgsn, method = c("BFGS"), x = theta)$par
}))

trueVal <- c(xi, Omega, lambda, beta)

avgpar <- apply(MLE, 2, mean)
avgpar
```

```
bias <- avgpar-trueVal
bias

#Calculating standard error
sdpar <- apply(MLE, 2, sd)
sdpar

#Calculating quantiles
Q <- apply(MLE, 2, quantile)
Q

#Calculating the quantile where the true value lies
ECDF <- lapply(1:4, function(i) ecdf(MLE[,i]))
trueQ <- c(ECDF[[1]](xi), ECDF[[2]](Omega),
ECDF[[3]](lambda), ECDF[[4]](beta))
trueQ

Pml1 = ecdf(MLE[,1])
plot(Pml1, main = "")

Pml2 = ecdf(MLE[,2])
plot(Pml2, main = "")

Pml3 = ecdf(MLE[,3])
plot(Pml3, main = "")

Pml4 = ecdf(MLE[,4])
plot(Pml4, main = "")

plot(density(MLE[,1]), main = "",ylab="PDF")
plot(density(MLE[,2]), main = "",ylab="PDF")
```

```
plot(density(MLE[,3]), main = "",ylab="PDF")
plot(density(MLE[,4]), main = "",ylab="PDF")
```

B.2.4 Raw circular data plot and rose diagram for the Galicia wind direction data

```
library(NPCirc)
library(Wrapped)
library(sn)

data("speed.wind")
set.seed(2)
sample_wind <- sample (speed.wind$Direction, size=1000, replace =F)

#Rose diagram of data
dir <- circular(sample_wind, units="degrees", template="geographics")
plot(dir, stack=TRUE,shrink = 1.15)
rose.diag(dir, bins=8, add=TRUE)
```

B.2.5 Estimates, standard errors and AIC values for the WN, WSN and WFGSN

```
#Wrapped Normal"
data("speed.wind")
set.seed(2)
sample_wind <- sample (speed.wind$Direction, size=1000, replace =F)

xi <- 0
omega <- 1

a1 <- sample_wind
a2 <- a1%%(2*pi)
a =na.exclude(a2)
```

```
logwsn <- function(par, theta)
-sum(log(dwrappednormal(x=theta, mu = par[1], sd = par[2])))

Opt = optim(par = c(xi, omega), fn = logwsn, method = c("L-BFGS-B"),
theta = a,hessian = TRUE)
Opt$val
Opt$par

aicwsn <- 2*(Opt$value-2)
aicwsn

bicwn <- 2*(mw$Measures[7])-log(984)*2
bicwn

#Standarderror (excluding NA values)
SE1=mw$Estimates[1,2]/sqrt(984)
SE1
SE2=mw$Estimates[2,2]/sqrt(984)
SE2

#Wrapped skew-normal
xi <- 0
omega <-0.5
lambda <- -1

a1 <- sample_wind
a2 <- a1%%(2*pi)
a =na.exclude(a2)
```

```
logwsn <- function(par, theta)
-sum(log(dwsn(theta, xi = par[1], eta = par[2], lambda = par[3])))

Opt = optim(par = c(xi, omega, lambda), fn = logwsn,
method = c("L-BFGS-B"), theta = a,lower = c(-3,0,-15),
          upper = c(3,2,-1),hessian = TRUE)
Opt$val
Opt$par
#Opt

aicwsn <- 2*(Opt$value-3)
aicwsn

bicwsn <- 2*(Opt$value)-log(984)*3
bicwsn

standarderrors = sqrt(abs(diag(solve(-Opt$hessian))))
standarderrors

#Wrapped flexible generalised skew-normal
xi <-0
Omega <- 1
lambda <- 2
beta <- 2

a <- speed.wind2$Direction
#simulation for WFGSN by applying the transformation
theta = a%%(2*pi)
#a=circular(a1)

logwfgsn=function (par, x, K = 100, min.k = 20)
```

```
{
  x <- conversion.circular(x, units = "radians",
   zero = 0, rotation = "counter")
  xi <- conversion.circular(xi, units = "radians",
   zero = 0, rotation = "counter")
  xi = par[1]
  eta = par[2]
  lambda = par[3]
  beta = par[4]
  x <- x[!is.na(x)]
  n <- length(x)
  if (sum(is.na(x)) > 0)
  #warning("Missing values were removed")
  if (is.null(K)) {
    range <- abs(xi - x)
    K <- (range + 6 * eta)%/%(2 * pi) + 1
    K <- max(min.k, K)
  }
  else {
    if (!is.numeric(K) | K <= 0) {
     #warning("Argument 'K' must be a positive integer.
     #'K=min.k' was used")
      K <- min.k
    }
  }
  fx <- numeric(n)
  for (i in 1:n) {
    val <- (x[i] + 2 * pi * seq(-K, K, 1) - xi)/eta
    suma <- sum(dnorm(val) * pnorm(lambda * val + beta * val^3))
    fx[i] <- 2/eta * suma
  }
```

```
  return(-sum(log(fx)))
}

Optwf =  optim(par=c(xi, Omega, lambda, beta), fn = logwfgsn,
method = c("BFGS"), x = theta, hessian=TRUE,
         lower = c(0,1,0.01,-140), upper = c(4,4,40,150))
Optwf$par
Optwf$value

aicwf <- 2*(Optwf$value-4)
aicwf

bicwf <- 2*(Optwf$value)-log(984)*4
bicwf

standarderrors = sqrt(abs(diag(solve(-Optwf$hessian))))
standarderrors
```

B.2.6 Raw circular data plot with the WN, WSN and WFGSN PDF fitted by MLE

```
library(NPCirc)
library(circular)
library(Wrapped)
data("wind.data")

#Rose diagram of data
dir <- circular(sample_wind, units="degrees", template="geographics")
plot(dir, stack=TRUE, shrink = 1.15)
rose.diag(dir, bins=8, add=TRUE)
wn <- function(x) dwrappedg(x,"norm",mean=-1.895088,sd=2.295792,K=2)
plot.function.circular(wn,add=TRUE,lwd=1, col=2, lty=1 )
```

```
wsn <- function(x) dwsn(x, xi=-0.9516579 , eta=2, lambda= -1)
plot.function.circular(wsn, add=TRUE, lwd=1, col=3, lty=1 )
dwgn <- function(theta, xi, omega, lambda, beta, r)
(2/omega)*sum(dnorm((theta+2*pi*r-xi)/omega, 0, 1)*
pnorm(lambda*(theta+2*pi*r-xi)/omega+beta*
((theta+2*pi*r-xi)/omega)^3, 0, 1))
wgn <- function(x) dwgn(theta=x, xi=1.103749, omega=2.5,
lambda= 3.479915, beta=-15, r=-100:100)
vwgn <- Vectorize(wgn)
curve.circular(vwgn,add=TRUE, lwd=1, col=4, lty=1)
legend("topleft",legend=c("WN","WSN","WFGSN"),
col=c("red","green","blue"),
bty="n", lty=c(1,1,1) ,ncol=1)
```

B.3 Chapter 4

B.3.1 The PDF of the S*t*N, WS*t*N, SSL, WSSL, SCN and WSCN

```
library("sn")
library("Wrapped")
library("circular")
library("BayesCR")
library("lqr")
library("VGAM")

#skew t normal
dstn <- dst(seq(-5,15, by=0.005),xi=0,omega=1,alpha=3,nu=2)
curve(dst(x, 0,1), from = -6, to = 10, ylim=c(0,0.5), col = "purple",
ylab = "StN", lwd = 1)
curve(dst(x, 3,1), from = -6, to = 10, col = "orange",
add = TRUE, lwd = 1)
curve(dst(x, 3,2), from = -6, to = 10, col = "green",
```

```
add = TRUE, lwd = 1)
curve(dst(x, -1,1), from = -6, to = 10, col = "sky blue",
add = TRUE, lwd = 1)
curve(dst(x, 0.5,1), from = -6, to = 10, col = "red",
add = TRUE, lwd = 1)
legend("topright",legend=c("StN(0,1,0,1)","StN(0,1,3,1)",
"StN(0,1,3,2)",
"StN(0,1,-1,1)", "StN(0,1,0.5,1)"), bty="n",
col=c("purple","orange","green","sky blue","red"),
lty=c(1,1,1,1) ,ncol=1)
#abline(v=0,lty=2,add=TRUE)

#wrapped skew t normal
stnwr<-function(x) dwrappedg(x,"st",xi=0,omega=1,alpha=3,nu=1,K=100)
curve.circular(stnwr,col="orange",shrink=1.25,lwd=1,lty=1)
stnwr1<-function(x) dwrappedg(x,"st",xi=0,omega=1,alpha=3,nu=2,K=100)
curve.circular(stnwr1,col="green",lwd=1,add=TRUE,lty=1)
stnwr2<-function(x) dwrappedg(x,"st",xi=0,omega=1,alpha=-1,nu=1,K=100)
curve.circular(stnwr2,col="sky blue",lwd=1,add=TRUE,lty=1)
stnwr3<-function(x) dwrappedg(x,"st",xi=0,omega=1,alpha=0.5,nu=1,K=100)
curve.circular(stnwr3,col="red",lwd=1,add=TRUE,lty=1)
stnwr4<-function(x) dwrappedg(x,"st",xi=0,omega=1,alpha=0,nu=1,K=100)
curve.circular(stnwr4,col="purple",lwd=1,add=TRUE,lty=1)
legend(-3.6,0.5,legend=c("WStN(0,1,0,1)","WStN(0,1,3,1)","WStN(0,1,3,2)",
       "WStN(0,1,-1,1)","WStN(0,1,0.5,1)"), bty="n",
        col=c("purple","orange","green","sky blue","red"),
        lty=c(1,1,1,1) ,ncol=1)

#Skew slash pdf
ssn<-function(y) dSKD(y, mu = 0.01, sigma = 1, p = 0.5,
dist = "slash", nu = 1)
```

```
ssn1<-function(y) dSKD(y, mu = 0.01, sigma = 1, p = 0.2,
dist = "slash", nu = 3)
ssn2<-function(y) dSKD(y, mu = 0.01, sigma = 1, p = 0.5,
dist = "slash", nu = 2)
ssn3<-function(y) dSKD(y, mu = 0.01, sigma = 1, p = 0.9,
dist = "slash", nu = 5)
curve(ssn, from = -10, to = 10, ylim=c(0,0.35), col = "orange",
ylab = "SSL",lwd = 1, type = "l")
curve(ssn1, from = -10, to = 10, col = "green", add = TRUE, lwd = 1)
curve(ssn2, from = -10, to = 10, col = "sky blue", add = TRUE, lwd = 1)
curve(ssn3, from = -10, to = 10, col = "red", add = TRUE, lwd = 1)
abline(v=0,lty=1,add=TRUE,lwd=1)
legend("topright",legend=c("SSL(0,1,0.5,1)",
"SSL(0,1,0.2,3)","SSL(0,1,0.5,2)","SSL(0,1,0.9,5)"),
bty="n",col=c("orange","green","sky blue","red"),
lty=c(1,1,1,1) ,ncol=1)

#wrapped skew slash pdf
curve.circular(ssn,col="orange",shrink=1.2,lwd=1,lty=1)
curve.circular(ssn1,col="green",lwd=1,add=TRUE,lty=1)
curve.circular(ssn2,col="sky blue",lwd=1,add=TRUE, lty=1)
curve.circular(ssn3,col="red",lwd=1,add=TRUE, lty=1)
legend(-3.6,0.2,legend=c("WSSL(0,1,0.5,1)","WSSL(0,1,0.2,3)",
"WSSL(0,1,0.5,2)", "WSSL(0,1,0.9,5)"), bty="n",
col=c("orange","green","sky blue","red"), lty=c(1,1,1,1) ,ncol=1)


#Skew contaminated normal pdf
ssln<-function(y) dSKD(y, mu = 0, sigma = 1, p = 0.5,
 dist = "cont", nu = 0.1, gama=0.2)
ssnl1<-function(y) dSKD(y, mu = 0, sigma = 1, p = 0.2,
```

```
 dist = "cont", nu = 0.3, gama=0.1)
ssnl2<-function(y) dSKD(y, mu = 0, sigma = 1, p = 0.3,
 dist = "cont", nu = 0.2, gama=0.4)
ssnl3<-function(y) dSKD(y, mu = 0, sigma = 1, p = 0.9,
 dist = "cont", nu = 0.5, gama=0.2)
ssnl4<-function(y) dSKD(y, mu = 0, sigma = 1, p = 0.2,
 dist = "cont", nu = 0.7, gama=0.9)
curve(ssln, from = -10, to = 10, ylim=c(0,0.4),
 col = "orange", ylab = "SCN", lwd = 1)
curve(ssnl1, from = -10, to = 10,
 col = "green", add = TRUE, lwd = 1)
curve(ssnl2, from = -10, to = 10,
 col = "sky blue", add = TRUE, lwd = 1)
curve(ssnl3, from = -10, to = 10,
 col = "red", add = TRUE, lwd = 1)
curve(ssnl4, from = -10, to = 10,
 col = "purple", add = TRUE, lwd = 1)
abline(v=0,lty=1,add=TRUE,lwd=1)
legend("topright",legend=c("SCN(0,1,0.5,0.1,0.2)",
"SCN(0,1,0.2,0.3,0.1)",
"SCN(0,1,0.3,0.2,0.4)","SCN(0,1,0.9,0.5,0.2)",
"SCN(0,1,0.2,0.7,0.9)"),
bty="n",col=c("orange","green","sky blue","red", "purple"),
lty=c(1,1,1,1) ,ncol=1)


#wrapped SCN
curve.circular(ssln,col="orange",shrink=1.3,lwd=1,lty=1)
curve.circular(ssnl1,col="green",lwd=1,add=TRUE,lty=1)
curve.circular(ssnl2,col="sky blue",lwd=1,add=TRUE, lty=1)
curve.circular(ssnl3,col="red",lwd=1,add=TRUE, lty=1)
```

```
curve.circular(ssnl4,col="purple",lwd=1,add=TRUE, lty=1)
legend(-4.5,0.2,legend=c("WSCN(0,1,0.5,0.1,0.2)","WSCN(0,1,0.2,0.3,0.1)",
"WSCN(0,1,0.3,0.2,0.4)", "WSCN(0,1,0.9,0.5,0.2)",
"WSCN(0,1,0.2,0.7,0.9)"),
bty="n",col=c("orange","green","sky blue","red","purple"),
lty=c(1,1,1,1) ,ncol=1)
```

B.3.2 Estimates, standard errors and AIC values for the WS*t*N, WSSL and WSCN

```
#WStN
library(NPCirc)
library(Wrapped)
library(sn)
library(circular)
library(BayesCR)
library(lqr)
library(VGAM)



#data from the NPCirc package
data("speed.wind")
set.seed(2)
sample_wind <- sample(speed.wind$Direction, size=1000, replace =F)

x1 <- sample_wind
x2 = x1%%(2*pi)
x =na.exclude(x2)

#wrapped skew student t
mw=mwrappedg("st",data = x, starts = c(0,1,0.4,0.4),
K = 2, method = "BFGS")
```

```
mw

#Standarderror (excluding NA values)
SE1=mw$Estimates[1,2]/sqrt(984)
SE1
SE2=mw$Estimates[2,2]/sqrt(984)
SE2
SE3=mw$Estimates[3,2]/sqrt(984)
SE3
SE4=mw$Estimates[4,2]/sqrt(984)
SE4

#AIC calculation
aicwn <- 2*(mw$Measures[7]-4)
aicwn

#BIC
bicwn <- 2*(mw$Measures[7])-log(984)*4
bicwn

#WSSL

mu <- 0
sigma <- 1
p <- 0.5
nu <- 1

a1 <- sample_wind
a2 <- a1%%(2*pi)
a =na.exclude(x2)
```

```
logwsn <- function(par, theta)
-sum(log(circular(dSKD(theta, mu = par[1], sigma = par[2],
p = par[3], dist = "slash", nu = par[4]))))

Opt = optim(par = c(mu, sigma, p, nu), fn = logwsn,
method = c("L-BFGS-B"), theta = a, lower = c(0, 1, 0.01,0.01),
          upper = c(10,10,0.999,10), hessian = TRUE)

Opt$val
Opt$par

standarderrors = sqrt(abs(diag(solve(-Opt$hessian))))
standarderrors

aicwsn <- 2*(Opt$val-4)
aicwsn

bicwn <- 2*(Opt$val)-log(984)*4
bicwn


#WSCN

mu <- 0
sigma <- 1
p <- 0.1
nu <- 0.1
gama <- 0.1

a1 <- sample_wind
a2 <- a1%%(2*pi)
```

```
a =na.exclude(x2)

logwsn <- function(par, theta)
-sum(log(dSKD(theta, mu = par[1], sigma = par[2], p = par[3],
dist = "cont", nu = par[4], gama = par[5])))

logwsn(par = c(mu, sigma, p, nu, gama), a)

Opt = optim(par = c(mu, sigma, p, nu, gama), fn = logwsn,
            method = c("L-BFGS-B"), theta = a,
            lower = c(0, 1, 0.0001,0.01,0.01),
            upper = c(10,10,0.99,0.99,0.99), hessian = TRUE)
Opt$val
Opt$par

standarderrors = sqrt(abs(diag(solve(-Opt$hessian))))
standarderrors

aicwscn <- 2*(Opt$value-5)
aicwscn

bicwscn <- 2*(Opt$value)-log(984)*5
bicwscn
```

B.3.3 Raw circular data plot with the WS*t*N, WSSL and WSCN PDF fitted by MLE

```
library(NPCirc)
library(circular)
library(Wrapped)

data("speed.wind")
```

```
set.seed(2)
sample_wind <- sample(speed.wind$Direction, size=1000, replace =F)

#Rose diagram of data
dir <- circular(sample_wind, units="degrees", template="geographics")
plot(dir, stack=TRUE,shrink = 1.1)
rose.diag(dir, bins=8, add=TRUE)
wstn <- function(x) dwrappedg(x,"st",xi= -1.4591276,
omega=2.3584247,alpha=-0.2502062,nu=34.5768341,K=100)
plot.function.circular(wstn,add=TRUE,lwd=1, col=2, lty=1 )
wssl<- function(x) dSKD(x, mu =  5.2001005, sigma = 1,
p =0.8130232, dist = "slash", nu = 10)
plot.function.circular(wssl,add=TRUE,lwd=1, col=3, lty=1 )
wscn<- function(x) dSKD(x, mu = 5.2618865, sigma = 1.000,
p = 0.8240806, dist = "cont", nu =  0.010, gama=0.990)
plot.function.circular(wscn,add=TRUE,lwd=1, col=4, lty=1 )

legend(-2.8,1.2,legend=c("WStN","WSSL","WSCN"),
col=c("red","green","blue"),
bty="n", lty=c(1,1,1) ,ncol=1)
```

Appendix C

Acronyms and symbols used

This Appendix contains a list of acronyms and symbols used throughout the study.

Acronyms

AIC	Akaike information criterion
BIC	Bayesian information criterion
CDF	Cumulative distribution function
CF	Characteristic function
ECDF	Empirical cumulative distribution function
EPDF	Empirical probability distribution function
FGSN	Flexible generalised skew-normal
MGF	Moment generating function
ML	Maximum likelihood
MLE	Maximum likelihood estimation
PDF	Probability distribution function

StN	Skew-Student-t normal
SCN	Skew-contaminated normal
SMN	Scale mixtures of normal
SN	Skew-normal
SSL	Skew-slash
SSMN	Skew scale mixtures of normal
WFGSN	Wrapped flexible generalised skew-normal
WN	Wrapped normal
WStN	Wrapped skew-Student-t normal
WSCN	Wrapped skew-contaminated normal
WSN	Wrapped skew-normal
WSSL	Wrapped skew-slash
WSSMN	Wrapped skew scale mixtures of normal

Symbols

$\mathbb{R}$	real number
$\mathbb{R}^+$	positive real number
$(.)^T$	transpose of a vector

Index