Skew-normal distributions: advances in theory and applications

by

Brett William Rowland

Submitted in partial fulfillment of the requirements for the degree

Magister Scientiae (Mathematical Statistics)

In the Department of Statistics
In the Faculty of Natural and Agricultural Sciences
University of Pretoria

August 2017

Abstract

The normal distribution is popular in many statistical contexts. However, due to its symmetry and tail behavior it may not necessarily be the best choice to use in many real world applications. In order to alleviate the aforementioned issues, a symmetric generalised normal distribution that exhibits flexibility in its tail behavior is proposed as candidate to apply existing skewing methodology to. Methods to approximate the characteristics of this new distribution and a corresponding stochastic representation is derived. The skewed version of the generalised normal distribution, along with other distributions, is used in a distribution fitting context and to approximate particular binomial distributions as an application.

Keywords: Approximating binomial distribution, Distribution fitting, Skew generalised-normal, Stochastic representation.

Declaration

I, *Brett William Rowland*, declare that this mini-dissertation, which I hereby submit for the degree Magister Scientiae 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.

____
Brett William Rowland

____
Prof. A. Bekker

____
Prof. M. Arashi

____
Mr. J. T. Ferreira

____
Date

Contents

Acknowledgments

I would like to thank the following people who joined me on this journey:

- My parents, Tim and Jeanne Rowland, for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.
- I am grateful to my supervisor, Prof. A. Bekker, whose expertise, understanding, generous guidance and support made it possible to work on a topic that was of great interest to me. It was a pleasure working with you.
- I would like to express my gratitude to Prof. M. Arashi for his valuable time and advice.
- I would like to thank Mr. J. T. Ferreira for the moral support, encouragement and direction.

I would like to thank acknowledge the financial assistance I received:

- The Crisis in Academic Statistics program of National Research Foundation (NRF) of South Africa and South African Statistical Association (SASA). Opinions expressed and conclusions arrived at in this study, and those of the authors and are not necessarily to be attributed to the NRF.
- STATOMET (Department of Statistics, University of Pretoria).
- National Research Foundation of South Africa (ref CPRR 13090132066 grant nr 91497).

Chapter 1

Introduction

1.1 Background and motivation

The normal distribution is frequently used in a variety of fields: quantitative finance, clinical studies, environmental risk analysis and banking, insurance and investment models. This distribution is prominent in statistics and forms the foundation of various statistical techniques such as analysis of variance and discriminant analysis. It also possesses attractive statistical characteristics including symmetry, infinite support and ease of computation; the drawback, however, is that it is inappropriate to use in scenarios where a non-negligible degree of skewness is present. The gamma-, folded normal-, and Weibull distributions can be used when skewness is present, however, these distributions do not have infinite real support and may not be suitable for certain scenarios.

For example, normality of errors is a popular assumption in linear regression models, but this is indeed only an assumption. If the distribution of the error term (which requires infinite support) exhibits some characteristic of the normal distribution (approximately bell shaped), but the distribution is slightly asymmetrical with non-normal tail behavior, then a distribution that is more general in terms of its symmetry and tails needs to be considered. It is therefore expedient to skew the normal distribution using reliable methodology in instances where the assumption of symmetry is violated. In most cases, the point of departure in this framework is the standard normal distribution. It is therefore of interest to consider generalising the standard normal distribution and using skewing methodology to enhance the flexibility of the existing skew-normal distribution and extensions thereof. This motivates the interest in developing distributions that retain normal-like characteristics but allow primarily for varying levels of skewness.

1.2 Literature review

Azzalini [5] introduced the skew-normal distribution, which includes the standard normal distribution as a special case, and studied the basic mathematical properties thereof. The skewing methodology that is used to skew existing symmetric probability density functions (PDFs) is stated in Proposition 1.

Proposition 1. *Denote by $f_0(\cdot)$ a probability density function (PDF) on $\mathbb{R}^d$, by $G_0(\cdot)$ a continuous cumulative distribution function (CDF) on $\mathbb{R}$, and by $w(\cdot)$ a real-valued function on $\mathbb{R}^d$, such that $f_0(-x) = f_0(x)$, $w(-x) = -w(x)$ and $G_0(-y) = 1 - G_0(y)$ for all $x \in \mathbb{R}^d$, $y \in \mathbb{R}$. Then*

$$f_X(x) \quad = \quad 2f_0(x)\, G_0\{w(x)\} \tag{1.1}$$

is a PDF on $\mathbb{R}^d$ [7].

Note that f_0 is termed the *symmetric base PDF*, $2G_0\{w(x)\}$ is termed the *skewing mechanism* and f_X is termed the skewed version of the symmetric base PDF.

Interestingly, as discussed by Pourahmadi [25], skew-normal random variables inherit some of the properties of the normal distribution, whilst also displaying particular characteristics that are not attributed to the normal distribution. For example, the square of a skew-normal random variable is distributed chi-square with one degree of freedom, however, the sum of two independent skew-normal random variables is generally not skew-normally distributed.

Pewsey [24] highlighted inferential issues regarding Azzalini's skew-normal distribution and discussed reasons as to why a different parameterisation of the skew-normal distribution may be more appropriate.

Arellano-Valle et. al. [3] noted that a limitation of the skew-normal distribution is that for moderate values of the skewness parameter nearly all the mass lies to the right or left of 0, as determined by the parameter. To mitigate this, the authors introduced a skew-generalised normal distribution which contains Azzalini's skew-normal as a special case and displays enhanced flexibility in modeling skewness. The extended skew-generalised normal distribution, a further extension of the skew-generalised normal distribution [3], was derived by Venegas et. al. [30]. Gomez et. al. [15] derived a skew flexible-normal distribution, an extension to the skew-normal model [5], which supports both unimodal and bimodal distributions.

A flexible skew-generalised normal distribution which contains the normal, Azzalini's skew-normal [5], the skew-generalised normal [3] and the skew flexible-normal [15] distributions as special cases, was developed by Bahrami and Qasemi [9]. Skew-elliptical families and semi-parametric extensions were investigated by Azzalini [6].

Gomez et. al. [16] introduced a family of asymmetric distributions which includes Azzalini's skew-normal [5] and is more flexible in its shape (i.e. skewness and kurtosis).

Salinas et. al. [27] introduced an extended family of skew distributions which is more flexible in terms of both skewness and kurtosis than those defined by Azzalini's skewing methodology. This family can be stochastically represented as the product of two random variables which is crucial for simulation studies and derivation of theoretical properties. Salinas et. al. [27] then used this extended family together with the exponential power distribution, as introduced by Subbotin [28], to derive what is termed the extended skew exponential-power distribution. This distribution includes Azzalini's skew-normal [5] and the generalised skew-normal distribution of [16].

Balakrishnan [11] generalised Azzalini's skew-normal [5] in [4]. This distribution is known as the Balakrishnan skew-normal distribution and includes Azzalini's skew-normal [5] distribution. Yadegari et. al. [31] provided a generalisation of the Balakrishnan skew-normal distribution and a multivariate extension was also presented.

Hasanalipour and Sharafi [19] introduced a new generalised Balakrishnan skew-normal distribution which contains Azzalini's skew-normal [5] and Arellano-Valle et. al. [3] skew-generalised normal. A method to simulate from this generalised Balakrishnan skew-normal distribution was also presented.

Mameli and Musio [22] provided generalisations of the Balakrishnan skew-normal distribution. The authors also introduced the beta skew-normal, which generalises Azzalini's skew-normal model [5] by considering the distribution of order statistics of the skew-normal distribution [5]. Abtahi et. al. [1] derived an empirical version of skew-normal density which employed kernel density estimation of the function responsible for skewing the original symmetric distribution.

Multivariate and matrix extensions of the models discussed have been investigated. The univariate skew-normal distribution [5] was extended to the multivariate case by Azzalini and Dalla Valle [8]. The univariate skew-normal distribution [5] was extended to another multivariate case (which includes Azzalini and Dalla Valle's [8] multivariate extension) that is coherent with the joint distribution of univariate skew normal random variables by Gupta and Chen [17]. A matrix variate skew-normal distribution was derived by Harrar and Gupta [18].

For clarity, the univariate distributions mentioned in this literature review that stem from Azzalini's skew-normal distribution [5] are summarised in Figure 1.1.

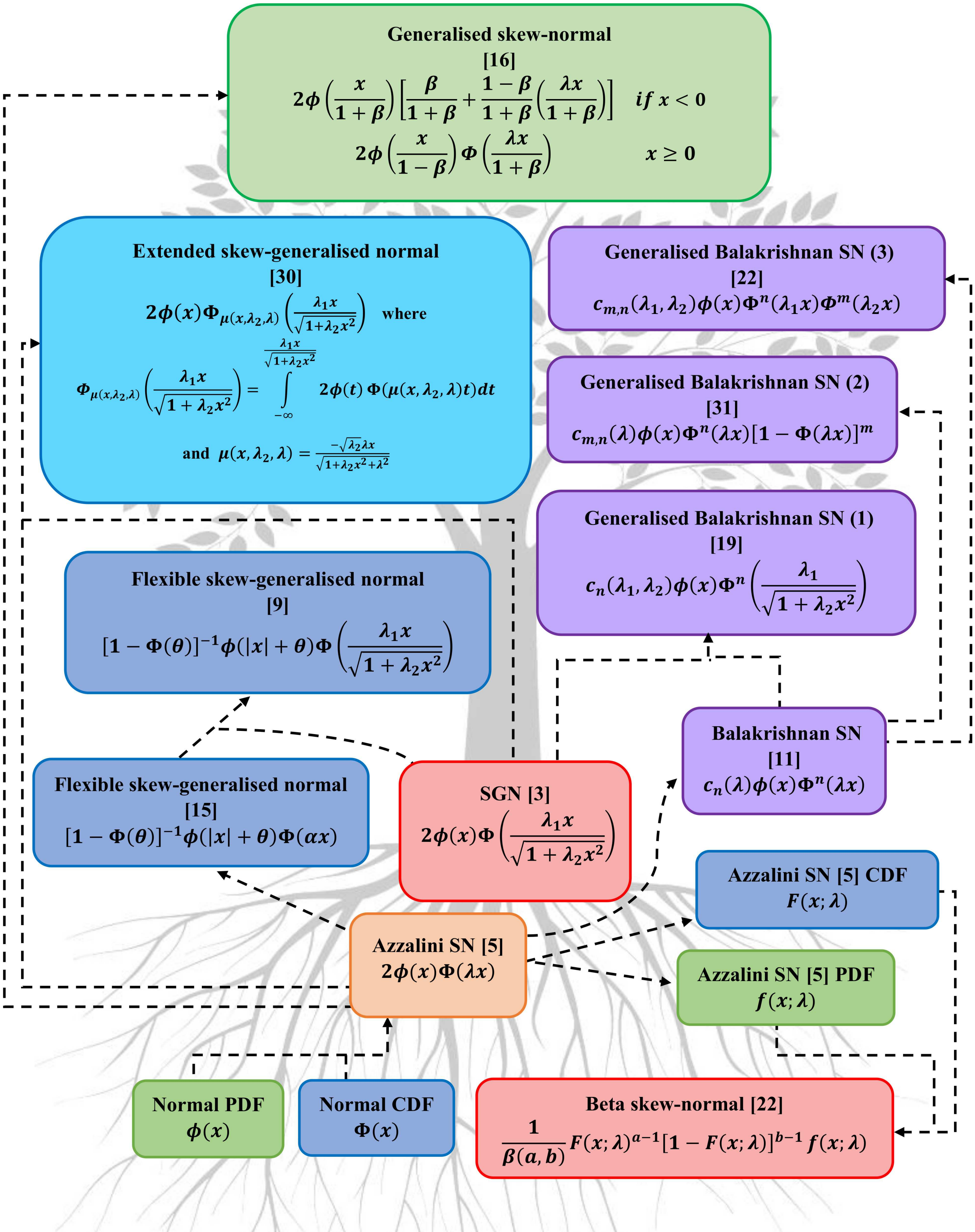


Figure 1.1: A summary of the distributions of referred to in the literature review.

1.3 Objectives

The aim of this study is to:

- Investigate and understand the methodology provided by Azzalini [5] in skewing a symmetric distribution (with a focus on the normal distribution);
- Review and revisit some existing generalisations of Azzalini's skew-normal distribution;
- Justify the use of the generalised normal distribution due to Subbotin [28] as a candidate in this framework;
- Investigate the properties of the newly proposed skew-symmetric version of the generalised normal distribution and compare with existing distributions;
- Investigate the effect of the mechanism responsible for skewing a given symmetric distribution;
- Develop a stochastic representation of the skew-symmetric version of the generalised normal distribution;
- Apply developed theory in a distribution fitting context and to approximate the binomial distribution.

1.4 Outline of study

- In Chapter 2 the skew-normal distribution, as proposed by Azzalini [5], is investigated. Characteristics of this distribution are revisited and a sampling scheme to generate random variates from this distribution is provided. Furthermore, existing extensions of the skew-normal distribution, relevant to this study, are presented and are summarised in Figure 1.2 and Figure 1.3. The distributions in green will be extended in Chapter 3 and Chapter 4.

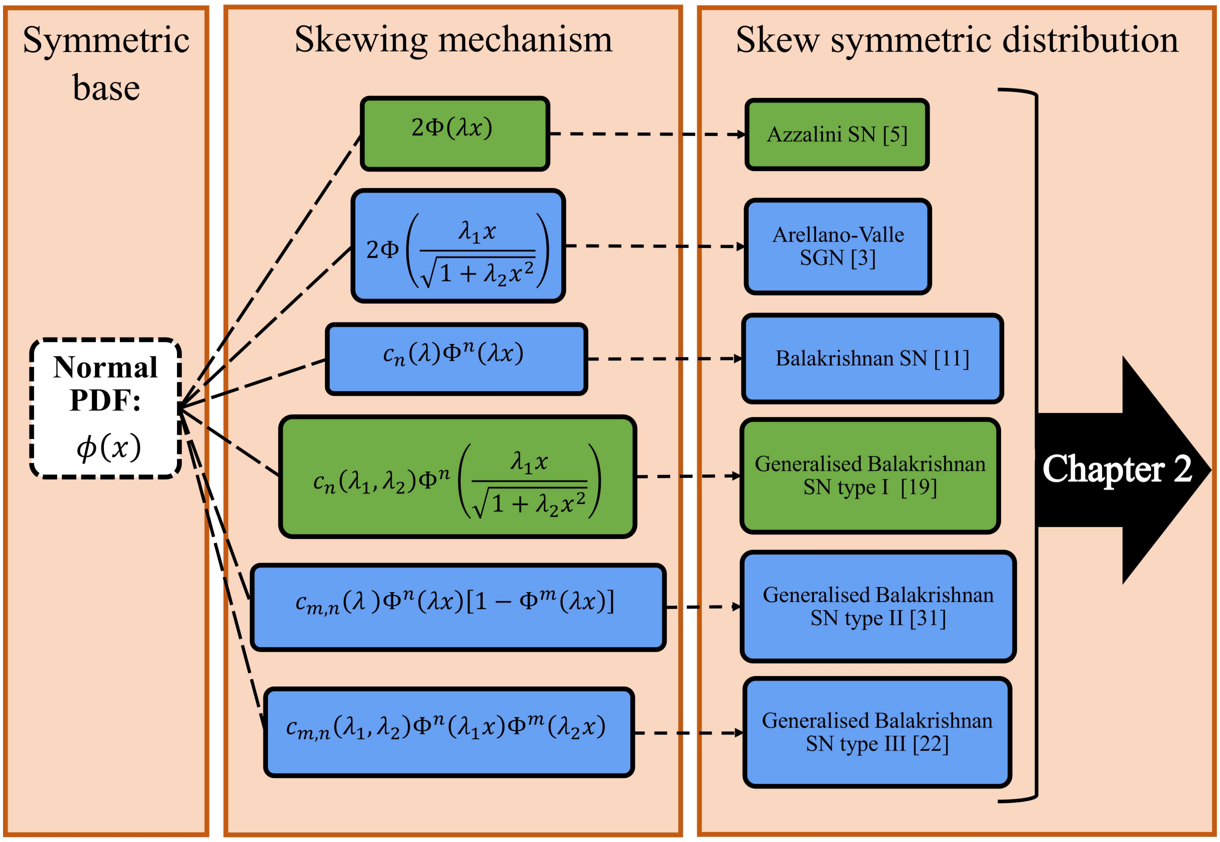


Figure 1.2: A summary of the distributions investigated in Chapter 2.

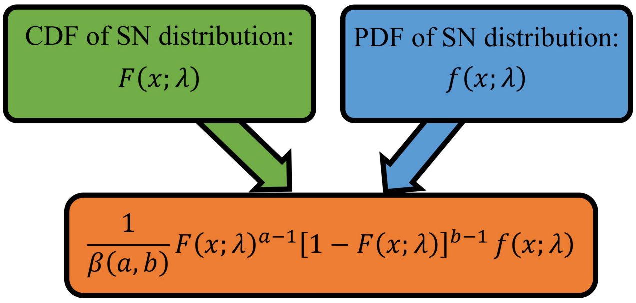


Figure 1.3: The beta skew-normal distribution investigated in Chapter 2 using methodology in [22].

- In Chapter 3 the generalised normal distribution [28] is considered as a new candidate to apply the existing skewing methodology to. The proposed skew-symmetric version of the generalised normal distribution is developed using the methodology discussed in Proposition 1, Section 1.2 and is termed the skew generalised-normal type I distribution. The added flexibility in modeling skewness with the skew generalised-normal type I distribution is graphically and numerically investigated. An acceptance-rejection sampling scheme is used to draw random samples from the skew generalised-normal type I distribution and two different approaches are used to investigate the characteristics of this distribution. Finally, a stochastic representation of the skew generalised-normal type I distribution is derived.

- Chapter 4 explores the effect of the skewing mechanism as outlined in Proposition 1, Section 1.2, the difference being that the generalised normal distribution is used as both the symmetric base PDF and also skewing mechanism. This chapter also uses the generalised normal distribution (see Chapter 3.1) as the symmetric base PDF (see Proposition 1, Section 1.2) to extend the generalised Balakrishnan skew-normal type I distribution in Chapter 2. The beta skew-normal in Chapter 2 (see [22]) is generalised using the proposed skew generalised-normal type I distribution (see Chapter 3) in the definition of a beta generated distribution [22].

The distributions of interest in Chapter 3 and Chapter 4 are summarised in Figure 1.4 and Figure 1.5 respectively where $\phi^*(\cdot)$ and $\Phi^*(\cdot)$ denote the PDF and CDF of the generalised normal distribution.

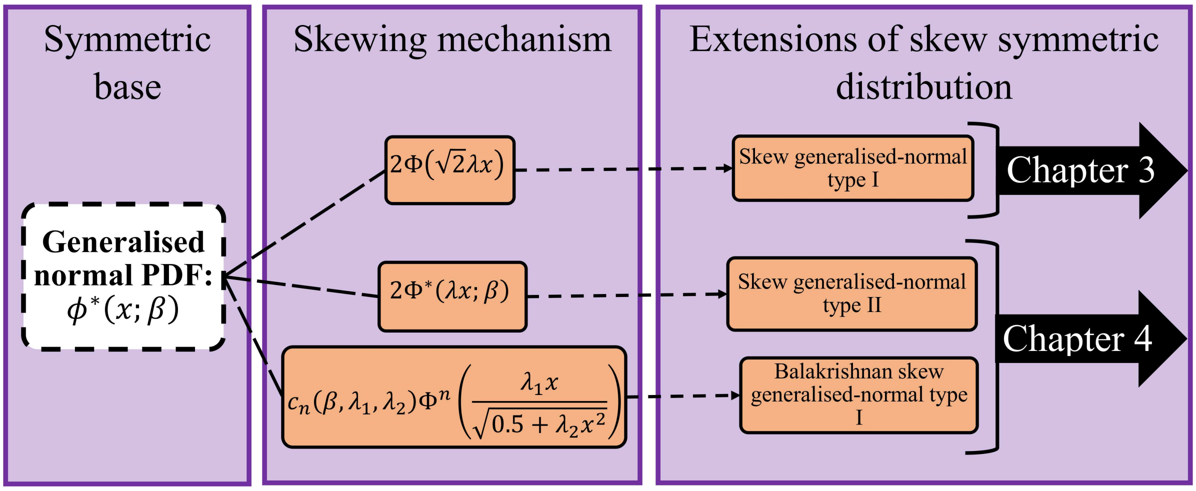


Figure 1.4: A summary of the distributions investigated Chapter 3 and Chapter 4.

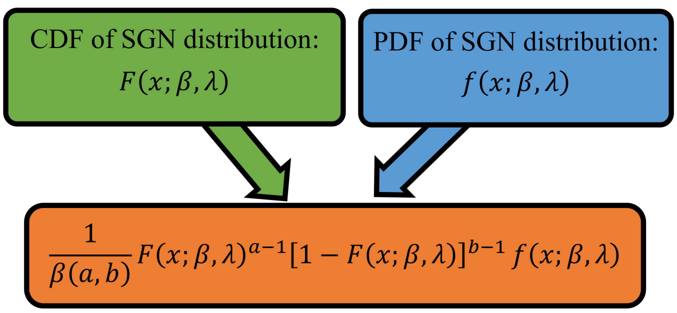


Figure 1.5: The beta skew generalised-normal type I distribution investigated in Chapter 4.

- Chapter 5 applies some of the developed theory in a distribution fitting context. Some of the developed distributions are used to approximate the binomial distribution as a further application.

- Chapter 6 concludes the study.
- Appendix A contains a list of symbols and notation used throughout.
- Appendix B contains a list of additional definitions and results referenced in this study.
- Appendix C contains code used in this study.

Chapter 2

The skew-normal distribution and extensions

In Section 2.1, the skewing methodology (see Proposition 1, Section 1.2) is applied to the normal distribution yielding Azzalini's skew-normal distribution [5]. Characteristics (i.e. expected value, variance, skewness and kurtosis) of the skew-normal distribution are revisited in Section 2.2. A stochastic representation, including a visualisation of the corresponding sampling scheme is presented in Section 2.3. In Section 2.4 the effect of the skewness parameter on the characteristics of the $\mathcal{SN}$ distribution is investigated. In Section 2.5 existing generalisations of the skew-normal model, relevant to this study, are presented.

2.1 The skew-normal ($\mathcal{SN}$) distribution

In this section, using the same notation defined in Proposition 1, Section 1.2, the case where $f_0 = \phi$, (with ϕ and Φ representing the standard normal PDF and CDF respectively) and where $w(x) = \lambda x$ for $\lambda \in \mathbb{R}$ is investigated. The following corollaries result from Proposition 1.

Corollary 1. *A random variable X has the skew-normal distribution if its PDF is given by*

$$f_X(x;\lambda) = 2\phi(x)\Phi(\lambda x), \; x \in \mathbb{R} \tag{2.1}$$

where $\lambda \in \mathbb{R}$. This is denoted by $X \sim \mathcal{SN}(\lambda)$ [5].

Corollary 2. *A random variable Y has the skew-normal distribution with location parameter μ and scale parameter σ if its PDF is given by*

$$f_Y(y;\mu,\sigma,\lambda) = \frac{2}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi\left(\lambda\left(\frac{y-\mu}{\sigma}\right)\right), \; y \in \mathbb{R} \tag{2.2}$$

where $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$ and $\lambda \in \mathbb{R}$. This is denoted by $Y \sim \mathcal{SN}(\mu, \sigma^2, \lambda)$.

Proof. Let $X \sim \mathcal{SN}(\lambda)$ with PDF (2.1). Consider the random variable $Y = \mu + \sigma X$, where the location and scale parameters are denoted $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ respectively.

If $y = \mu + \sigma x$ then $\frac{d}{dy}u^{-1}(y) = \frac{1}{\sigma}$, and it follows that

$$
\begin{aligned}
f_Y(y;\mu,\sigma,\lambda) &= f_X\left(u^{-1}(y);\lambda\right)\left|\frac{d}{dy}u^{-1}(y)\right| \\
&= 2\phi\left(u^{-1}(y)\right)\Phi\left(\lambda u^{-1}(y)\right)\left|\frac{d}{dy}\left(u^{-1}(y)\right)\right| \\
&= 2\phi\left(\frac{y-\mu}{\sigma}\right)\Phi\left(\lambda\left(\frac{y-\mu}{\sigma}\right)\right)\left|\frac{1}{\sigma}\right| \\
&= \frac{2}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi\left(\lambda\left(\frac{y-\mu}{\sigma}\right)\right).
\end{aligned}
$$

□

2.2 Characteristics of skew-normal distribution

2.2.1 Moment generating function

In this section the moment generating function (MGF) of the $\mathcal{SN}\left(\mu,\sigma^2,\lambda\right)$ distribution with PDF (2.1) is derived.

Lemma 1. *If $U \sim \mathcal{N}(0,1)$ then $\mathbb{E}_U[\Phi(hU+k)] = \Phi\left(\frac{k}{\sqrt{1+h^2}}\right)$ for $h,k \in \mathbb{R}$ [7], where $\Phi(\cdot)$ denotes the CDF of the standard normal distribution.*

Proof. By definition

$$
\begin{aligned}
\mathbb{E}_U[\Phi(hU+k)] &= \int_{\mathbb{R}} \Phi(hu+k)\,\phi(u)\,du \\
&= \int_{\mathbb{R}}\int_{-\infty}^{hu+k} \phi(b)\,\phi(u)\,dbdu \\
&= \int_{-\infty}^{hu+k}\int_{\mathbb{R}} \phi(b)\,\phi(u)\,dudb \\
&= \int_{-\infty}^{hu+k} \phi(b)\left(\int_{\mathbb{R}}\phi(u)\,du\right)db \\
&= \int_{-\infty}^{hu+k} \phi(b)\,db \\
&= \mathbb{P}[B < hU+k] \\
&= \mathbb{P}[B - hU < k] \qquad (2.3)
\end{aligned}
$$

where B is independent of U.

Using well-known results [10] it follows that the distribution of $W = B - hU$ is again normally distributed with

$$\begin{aligned}\mathbb{E}[W] &= \mathbb{E}[B-hU] \\ &= \mathbb{E}[B]-h\mathbb{E}[U] \\ &= 0.\end{aligned}$$

Since B and U are independent, $cov(B,U)=0$ and it follows that

$$\begin{aligned}var[W] &= var[B-hU] \\ &= var[B]+h^2 var[U] \\ &= 1+h^2.\end{aligned}$$

Therefore $W \sim \mathcal{N}\left(0,1+h^2\right)$.

Then from (2.3) it follows that

$$\begin{aligned}\mathbb{E}_U[\Phi(hU+k)] &= \mathbb{P}[B-hU<k] \\ &= \mathbb{P}[W<k] \\ &= \mathbb{P}\left[\frac{W-\mathbb{E}[W]}{\sqrt{var[W]}}<\frac{k-E[W]}{\sqrt{var[W]}}\right] \\ &= \mathbb{P}\left[Z<\frac{k}{\sqrt{1+h^2}}\right], \text{ where } Z\sim\mathcal{N}(0,1) \\ &= \Phi\left(\frac{k}{\sqrt{1+h^2}}\right)\end{aligned}$$

which concludes the proof. □

Theorem 1. *The MGF of random variable $Y=\mu+\sigma X$ with PDF (2.1) is given by*

$$M_Y(t) = 2e^{t\mu+\frac{1}{2}t^2\sigma^2}\Phi(\delta\sigma t), \quad t\in\mathbb{R} \tag{2.4}$$

where $\delta=\frac{\lambda}{\sqrt{1+\lambda^2}}$ and $\Phi(\cdot)$ denotes the standard normal CDF [7].

Proof. From (2.1) it follows that

$$\begin{aligned}M_Y(t) &= \mathbb{E}\left[e^{tY}\right] \\ &= \mathbb{E}\left[e^{t(\mu+\sigma X)}\right] \\ &= \int_{\mathbb{R}} e^{t\mu+t\sigma x}2\phi(x)\Phi(\lambda x)\,dx \\ &= 2e^{t\mu}\int_{\mathbb{R}} e^{t\sigma x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\Phi(\lambda x)\,dx \\ &= 2e^{t\mu}\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(x^2-2t\sigma x\right)}\Phi(\lambda x)\,dx.\end{aligned}$$

Note that $(x-t\sigma)^2 = x^2 - 2t\sigma x + t^2\sigma^2 \implies x^2 - 2t\sigma x = (x-t\sigma)^2 - t^2\sigma^2$.

Therefore

$$\begin{aligned} M_Y(t) &= 2e^{t\mu}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left((x-t\sigma)^2 - t^2\sigma^2\right)} \Phi(\lambda x)\,dx \\ &= 2e^{t\mu + \frac{1}{2}t^2\sigma^2}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t\sigma)^2} \Phi(\lambda x)\,dx. \end{aligned}$$

If $p = x - t\sigma$ then $x = p + t\sigma$ and $\frac{dx}{dp} = 1$.

Let $\phi(\cdot)$ denote the standard normal PDF; then:

$$\begin{aligned} M_Y(t) &= 2e^{t\mu + \frac{1}{2}t^2\sigma^2}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}p^2} \Phi(\lambda(p+t\sigma))\,dp \\ &= 2e^{t\mu + \frac{1}{2}t^2\sigma^2}\int_{\mathbb{R}} \phi(p)\,\Phi(\lambda(p+t\sigma))\,dp \\ &= 2e^{t\mu + \frac{1}{2}t^2\sigma^2}\mathbb{E}_P\left[\Phi(\lambda P + \lambda t\sigma)\right] \end{aligned} \tag{2.5}$$

where $P \sim \mathcal{N}(0,1)$.

Applying Lemma 1 it follows from (2.5) that

$$\begin{aligned} M_Y(t) &= 2e^{t\mu + \frac{1}{2}t^2\sigma^2}\Phi\left(\frac{\lambda t\sigma}{\sqrt{1+\lambda^2}}\right) \\ &= 2e^{t\mu + \frac{1}{2}t^2\sigma^2}\Phi(\delta\sigma t) \end{aligned}$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$. □

2.2.2 Moments

To compute the central moments of $Y \sim \mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ with PDF (2.2), the method as in [7] which uses the cumulant generating function, $K_Y(t)$, will be used (see Definition 14, Appendix B.1).

In order to simplify the derivation of the central moments, the inverse Mills ratio, $m(\cdot)$, (see Definition 13, Appendix B.1) will be used. The following properties of the inverse Mills ratio are derived freely using the quotient, product and chain rules of differentiation. Let $m'(\cdot)$ and $m''(\cdot)$ respectively denote the 1^{st} and 2^{nd} derivatives of $m(\cdot)$.

Property 1:

$$m(0) = \sqrt{\frac{2}{\pi}}$$

Proof.

$$\begin{aligned} m(0) &= \frac{\phi(0)}{\Phi(0)} \\ &= \frac{\frac{1}{\sqrt{2\pi}}}{\frac{1}{2}} \\ &= \sqrt{\frac{2}{\pi}}. \end{aligned}$$

□

Property 2:

$$m'(x) = \frac{d}{dx}m(x) = -m(x)\left[x + m(x)\right]$$

Proof.

$$\begin{aligned}
\frac{d}{dx}m(x) &= \frac{d}{dx}\left(\frac{\phi(x)}{\Phi(x)}\right) \\
&= \frac{\left(\frac{d}{dx}\phi(x)\right)\Phi(x) - \phi(x)\left(\frac{d}{dx}\Phi(x)\right)}{\left(\Phi(x)\right)^2} \\
&= \frac{-x\phi(x)\Phi(x) - \phi(x)\phi(x)}{\left(\Phi(x)\right)^2} \\
&= -\left(\frac{x\phi(x)}{\Phi(x)} + \left(\frac{\phi(x)}{\Phi(x)}\right)^2\right) \\
&= -\frac{\phi(x)}{\Phi(x)}\left[x + \frac{\phi(x)}{\Phi(x)}\right] \\
&= -m(x)\left[x + m(x)\right].
\end{aligned}$$

□

Property 3:

$$m''(x) = \frac{d^2}{dx^2}m(x) = -m(x) + x^2 m(x) + 3x\left(m(x)\right)^2 + 2\left(m(x)\right)^3$$

Proof.

$$\begin{aligned}
\frac{d^2}{dx^2}m(x) &= \frac{d}{dx}\left(\frac{d}{dx}m(x)\right) \\
&= \frac{d}{dx}\left(-m(x)\left[x + m(x)\right]\right) \\
&= -\frac{d}{dx}\left(xm(x) + \left(m(x)\right)^2\right) \\
&= -\left(\frac{d}{dx}\left(xm(x)\right) + \frac{d}{dx}\left(m(x)\right)^2\right) \\
&= -\left(\left(\frac{d}{dx}x\right)m(x) + x\left(\frac{d}{dx}m(x)\right) + 2m(x)\left(\frac{d}{dx}m(x)\right)\right) \\
&= -\left(m(x) + x\left(-m(x)\left[x + m(x)\right]\right) + 2m(x)\left(-m(x)\left[x + m(x)\right]\right)\right) \\
&= -\left(m(x) - x^2 m(x) - x\left(m(x)\right)^2 - 2x\left(m(x)\right)^2 - 2\left(m(x)\right)^3\right) \\
&= -\left(m(x) - x^2 m(x) - 3x\left(m(x)\right)^2 - 2\left(m(x)\right)^3\right) \\
&= -m(x) + x^2 m(x) + 3x\left(m(x)\right)^2 + 2\left(m(x)\right)^3.
\end{aligned}$$

□

Using the derived properties of the inverse Mills ratio, some characteristics of the skew-normal distribution are derived.

Expected value

Theorem 2. *Consider* $Y \sim \mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ *with MGF* (2.4), *then*

$$\mathbb{E}\left[Y\right] \quad = \quad \mu + \delta\sigma\sqrt{\frac{2}{\pi}}$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

Proof. From Theorem 1 it follows that

$$\begin{aligned}
\mathbb{E}\left[Y\right] &= \left.\frac{d}{dt}K_Y\left(t\right)\right|_{t=0}^{[1]} \\
&= \left.\frac{d}{dt}\left(\log M_Y\left(t\right)\right)\right|_{t=0} \\
&= \left.\frac{d}{dt}\left(\log\left(2e^{t\mu+\frac{1}{2}t^2\sigma^2}\Phi\left(\delta\sigma t\right)\right)\right)\right|_{t=0} \\
&= \left.\frac{d}{dt}\left(t\mu+\frac{1}{2}t^2\sigma^2+\log\left(2\Phi\left(\delta\sigma t\right)\right)\right)\right|_{t=0} \\
&= \left.\mu+\sigma^2 t+\frac{\frac{d}{dt}2\Phi\left(\delta\sigma t\right)}{2\Phi\left(\delta\sigma t\right)}\right|_{t=0} \\
&= \left.\mu+\sigma^2 t+\frac{2\delta\sigma\phi\left(\delta\sigma t\right)}{2\Phi\left(\delta\sigma t\right)}\right|_{t=0} \\
&= \left.\mu+\sigma^2 t+\delta\sigma m\left(\delta\sigma t\right)\right|_{t=0}{}^{[2]} \\
&= \mu+\delta\sigma m\left(0\right) \\
&= \mu+\delta\sigma\sqrt{\frac{2}{\pi}}\ ^{[3]}
\end{aligned}$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

[1]Applying Definition 14, Appendix B.1;

[2]Applying Definition 13, Appendix B.1;

[3]Using Property 1. □

Variance

Theorem 3. *Consider* $Y \sim \mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ *with MGF* (2.4), *then*

$$var\left[Y\right] \quad = \quad \sigma^2\left(1-\frac{2}{\pi}\delta^2\right)$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

Proof. From Theorem 1 it follows that

$$\begin{aligned}
var\left[Y\right] &= \left.\frac{d^2}{dt^2}K_Y\left(t\right)\right|_{t=0}^{[1]} \\
&= \left.\frac{d}{dt}\left(\frac{d}{dt}\log M_X\left(t\right)\right)\right|_{t=0} \\
&= \left.\frac{d}{dt}\left(\frac{d}{dt}\left(t\mu+\frac{1}{2}t^2\sigma^2+\log\left(2\Phi\left(\delta\sigma t\right)\right)\right)\right)\right|_{t=0} \\
&= \left.\frac{d}{dt}\left(\mu+\sigma^2 t+\frac{\delta\sigma\phi\left(\delta\sigma t\right)}{\Phi\left(\delta\sigma t\right)}\right)\right|_{t=0} \\
&= \left.\sigma^2+\delta\sigma\frac{d}{dt}m\left(\delta\sigma t\right)\right|_{t=0} \quad [2] \\
&= \left.\sigma^2+\delta\sigma\left(-m\left(\delta\sigma t\right)\left[\delta\sigma t+m\left(\delta\sigma t\right)\right]\frac{d}{dt}\left(\delta\sigma t\right)\right)\right|_{t=0} \quad [3] \\
&= \left.\sigma^2+\left(\delta\sigma\right)^2\left(-m\left(\delta\sigma t\right)\left[\delta\sigma t+m\left(\delta\sigma t\right)\right]\right)\right|_{t=0} \\
&= \sigma^2+\left(\delta\sigma\right)^2\left(-\left(m\left(0\right)\right)^2\right) \\
&= \sigma^2+\left(\delta\sigma\right)^2\left(-\frac{2}{\pi}\right) \quad [4] \\
&= \sigma^2\left(1-\frac{2}{\pi}\delta^2\right)
\end{aligned}$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

[1]Applying Definition 14, Appendix B.1;

[2]Applying Definition 13, Appendix B.1;

[3]Using Property 2;

[4]Using Property 1. □

3^{rd} central moment

Theorem 4. *Consider* $Y \sim \mathcal{SN}\left(\mu,\sigma^2,\lambda\right)$ *with MGF* (2.4)*, then*

$$\mathbb{E}\left[\left(Y-\mathbb{E}\left[Y\right]\right)^3\right] = \frac{1}{2}\left(4-\pi\right)\left(\delta\sigma\sqrt{\frac{2}{\pi}}\right)^3$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

Proof. From Theorem 1 it follows that

$$\begin{aligned}
\mathbb{E}\left[(Y-\mathbb{E}[Y])^3\right] &= \left.\frac{d^3}{dt^3}K_Y(t)\right|_{t=0}^{[1]} \\
&= \left.\frac{d}{dt}\left(\frac{d^2}{dt^2}\log M_X(t)\right)\right|_{t=0} \\
&= \left.\frac{d}{dt}\left(\frac{d^2}{dt^2}\log\left(2e^{t\mu+\frac{1}{2}t^2\sigma^2}\Phi(\delta\sigma t)\right)\right)\right|_{t=0} \\
&= \left.\frac{d}{dt}\left(\frac{d^2}{dt^2}\left(t\mu+\frac{1}{2}t^2\sigma^2+\log(2\Phi(\delta\sigma t))\right)\right)\right|_{t=0} \\
&= \left.\frac{d}{dt}\left(\frac{d}{dt}\left(\mu+\sigma^2 t+\frac{\delta\sigma\phi(\delta\sigma t)}{\Phi(\delta\sigma t)}\right)\right)\right|_{t=0} \\
&= \left.\frac{d}{dt}\left(\frac{d}{dt}(\mu+\sigma^2 t+\delta\sigma m(\delta\sigma t))\right)\right|_{t=0} \quad [2] \\
&= \left.\frac{d}{dt}\left(\sigma^2+\delta\sigma\frac{d}{dt}m(\delta\sigma t)\right)\right|_{t=0} \\
&= \left.(\delta\sigma)\left(\frac{d^2}{dt^2}m(\delta\sigma t)\right)\right|_{t=0} \\
&= \left.(\delta\sigma)\frac{d}{dt}(m'(\delta\sigma t)\,\delta\sigma)\right|_{t=0} \\
&= \left.(\delta\sigma)\left(\left(\frac{d}{dt}m'(\delta\sigma t)\right)\delta\sigma+m'(\delta\sigma t)\left(\frac{d}{dt}\delta\sigma\right)\right)\right|_{t=0} \\
&= \left.(\delta\sigma)\left(((m''(\delta\sigma t))\,\delta\sigma)\,\delta\sigma+0\right)\right|_{t=0} \\
&= \left.(\delta\sigma)^3 m''(\delta\sigma t)\right|_{t=0} \\
&= \left.(\delta\sigma)^3\left(-m(\delta\sigma t)+(\delta\sigma t)^2 m(\delta\sigma t)+3(\delta\sigma t)(m(\delta\sigma t))^2+2(m(\delta\sigma t))^3\right)\right|_{t=0} \quad [3] \\
&= (\delta\sigma)^3\left(-m(0)+2(m(0))^3\right) \\
&= (\delta\sigma)^3\left(-\sqrt{\frac{2}{\pi}}+2\left(\sqrt{\frac{2}{\pi}}\right)^3\right) \quad [4] \\
&= (\delta\sigma)^3\left(\frac{2}{\pi}\right)^{3/2}\left(2-\left(\frac{2}{\pi}\right)^{-1}\right) \\
&= \left(\delta\sigma\sqrt{\frac{2}{\pi}}\right)^3\frac{1}{2}(4-\pi) \\
&= \frac{1}{2}(4-\pi)\left(\delta\sigma\sqrt{\frac{2}{\pi}}\right)^3
\end{aligned}$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

[1]Applying Definition 14, Appendix B.1;

[2]Applying Definition 13, Appendix B.1;

[3]Using Property 3;

[4]Using Property 1. □

4^{th} central moment [7]

This stated without proof since the derivation follows similarly as before.

Theorem 5. *Consider $Y \sim \mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$, with MGF (2.4), then*

$$\mathbb{E}\left[\left(Y - \mathbb{E}\left[Y\right]\right)^4\right] = 2\left(\pi - 3\right)\left(\delta\sigma\sqrt{\frac{2}{\pi}}\right)^4$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

Standardisation of the 3^{rd} and 4^{th} central moments produces the commonly used measures of **skewness** (γ_1) and **kurtosis** (γ_2) of random variable Y given by

$$\gamma_1 = \frac{\mathbb{E}\left[\left(Y - \mathbb{E}\left[Y\right]\right)^3\right]}{\left(var\left[Y\right]\right)^{\frac{3}{2}}} = \frac{\frac{1}{2}\left(4-\pi\right)\left(\delta\sigma\sqrt{\frac{2}{\pi}}\right)^3}{\left(\sigma^2\left(1-\frac{2}{\pi}\delta^2\right)\right)^{\frac{3}{2}}}$$

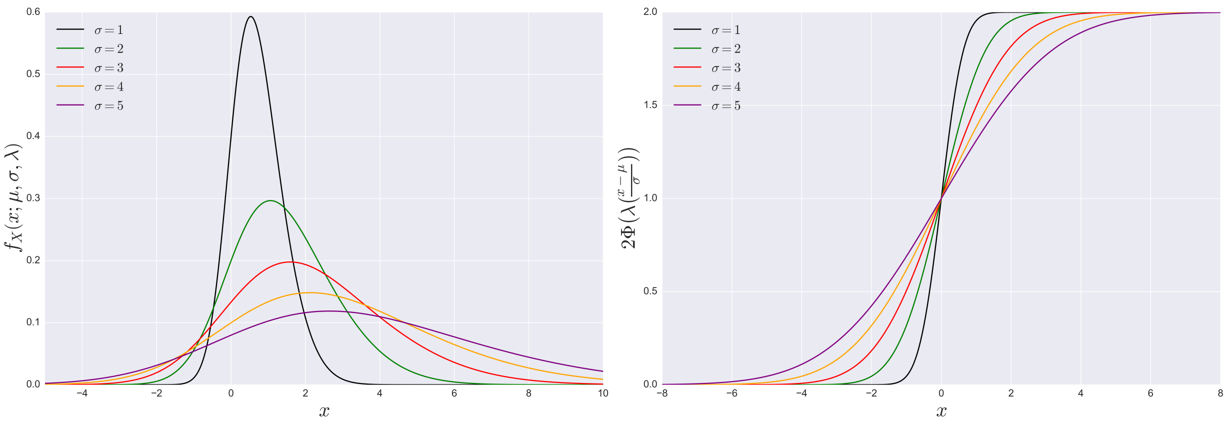
and

$$\gamma_2 = \frac{\mathbb{E}\left[\left(Y - \mathbb{E}\left[Y\right]\right)^4\right]}{\left(var\left[Y\right]\right)^{2}} = \frac{2\left(\pi-3\right)\left(\delta\sigma\sqrt{\frac{2}{\pi}}\right)^4}{\left(\sigma^2\left(1-\frac{2}{\pi}\delta^2\right)\right)^{2}}$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

2.2.3 $\mathcal{SN}$ PDF and skewing mechanism

Figure 2.1 and Figure 2.3 depict the PDF of the $\mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ distribution i.e. $f_X\left(x; \mu, \sigma, \lambda\right)$ as given in (2.2) and the corresponding skewing mechanism, $2\Phi\left(\lambda\left(\frac{x-\mu}{\sigma}\right)\right)$, for varying parameter values.



(a) $\mathcal{SN}$ PDF (b) $\mathcal{SN}$ skewing mechanism

Figure 2.1: The $\mathcal{SN}$ PDF (2.2) and corresponding skewing mechanism for varying σ and arbitrary $\mu = 0$ and $\lambda = 2$.

Remarks

In Figure 2.1b, it is observed that σ has an effect on the skewing mechanism for all σ. Since in this case $\lambda = 2 > 0$, the resulting $\mathcal{SN}$ distributions are all positively skewed.

The observation made in Figure 2.1b is stated in general terms:

Remark 1. Let x_0 (x_1) be the value of x for which the skewing mechanism approximately attains its minimum (maximum) value. The minimum and maximum values of the skewing mechanism will be denoted θ_{x_0} and θ_{x_1} respectively. Then to obtain the skew-symmetric version of the normal distribution:

- for $x \in (x_0, x_1)$ the skewing mechanism multiplies the original symmetric normal PDF by a value in the interval $(\theta_{x_0}, \theta_{x_1})$;
- for $x < x_0$, the $\mathcal{SN}$ PDF (2.2) is approximately θ_{x_0} and;
- for $x > x_1$, the $\mathcal{SN}$ PDF (2.2) is approximately θ_{x_1} multiplied by the original symmetric normal PDF.

Remark 2. The behavior of the skewing mechanism in the interval (x_0, x_1) determines how the skewness is introduced into the original skew-symmetric distribution. This interval is solely responsible for the resulting skewness and will be referred to as the *skewing window*.

This terminology will be used through the study and is summarised in Figure 2.2.

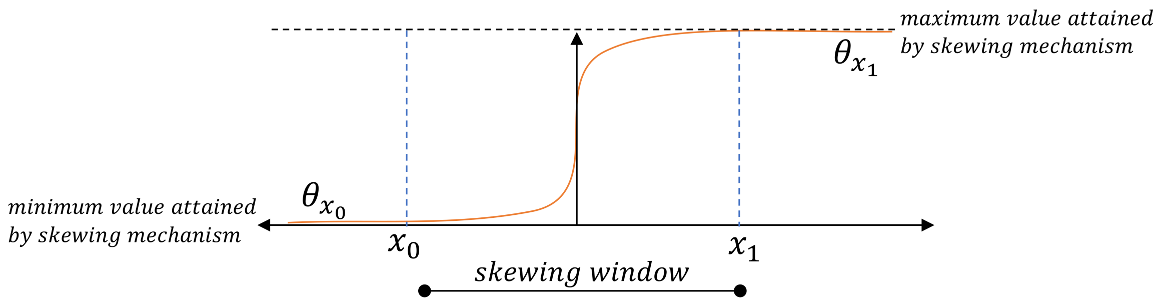


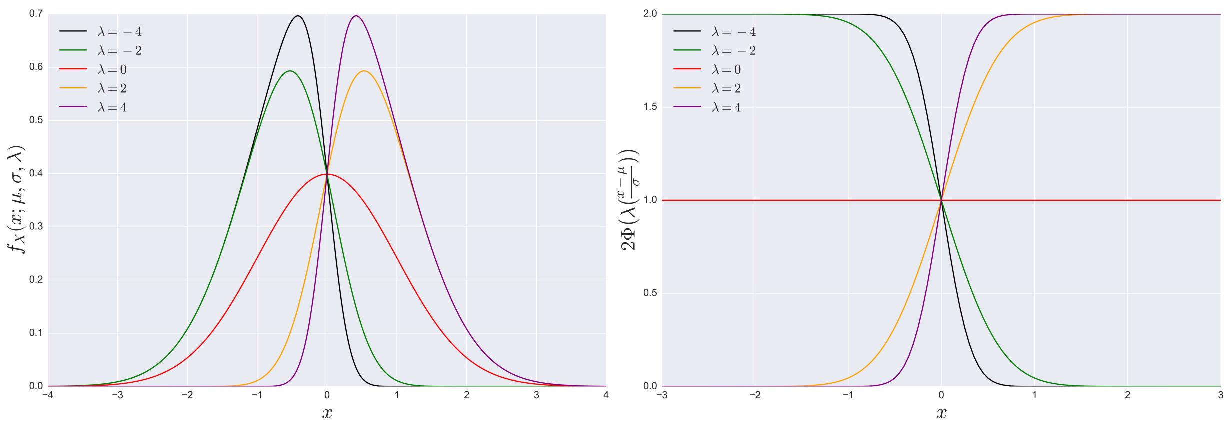
Figure 2.2: The skewing window exhibited by a particular skewing mechanism.

Example 1. Consider Figure 2.1b. When $\sigma = 3$, by inspection of the red curve $x_0 \approx -4$, $x_1 \approx 4$, $\theta_{x_0} \approx 0$ and $\theta_{x_1} \approx 2$, and so:

- for $x \in (-4, 4)$ the skewing mechanism multiplies original symmetric normal PDF by a value in the interval $(0, 2)$;
- for $x < -4$, the $\mathcal{SN}$ PDF is approximately 0;

- for $x > 4$, the $\mathcal{SN}$ PDF is approximately 2 multiplied by the original symmetric normal PDF;

Remark 3. It is important to note that $\theta_{x_0} \approx 0$ and $\theta_{x_1} \approx 2$ will always be the case when the skewing mechanism is $2G_0(\cdot)$ where $G_0(\cdot)$ is as defined in Proposition 1, Section 1.2.



(a) $\mathcal{SN}$ PDF (b) $\mathcal{SN}$ skewing mechanism

Figure 2.3: The $\mathcal{SN}$ PDF (2.2) and skewing mechanism, $2\Phi\left(\lambda\left(\frac{x-\mu}{\sigma}\right)\right)$, for varying λ and arbitrary $\mu = 0$ and $\sigma = 1$.

Remarks

1. In Figure 2.3a, it is observed that for $\lambda > 0$ and $\lambda < 0$ the resulting $\mathcal{SN}$ distributions are respectively positively and negatively skewed;

2. In Figure 2.3b, for $\lambda = 0$ it is observed that the skewing mechanism has a value of 1 and therefore the resulting distribution is simply the original symmetric normal distribution;

3. As can be deduced from Figure 2.3b, for increasing $|\lambda|$, the skewing window becomes narrower which implies that the resulting skewness is obtained by multiplying the original symmetric normal PDF by a value in the interval $(0, 2)$ over a narrower range of x resulting in a $\mathcal{SN}$ PDF (2.2) with peaks attaining a higher probabilities as in Figure 2.3a.

2.2.4 $\mathcal{SN}$ CDF

Consider $X \sim \mathcal{SN}(\lambda)$ with PDF (2.1). Denote the CDF of X by $F_X(x;\lambda)$. Then

$$\begin{aligned} F_X(x;\lambda) &= \int_{-\infty}^{x} 2\phi(t)\,\Phi(\lambda t)\,dt \\ &= \int_{-\infty}^{x} 2\phi(t) \int_{-\infty}^{\lambda t} \phi(u)\,dudt \\ &= 2\int_{-\infty}^{x}\int_{-\infty}^{\lambda t} \phi(t)\,\phi(u)\,dudt \\ &= \Phi(x) - T(x,\lambda) \end{aligned} \tag{2.6}$$

for $x \in \mathbb{R}$, $\lambda \in \mathbb{R}$, where $T(x,\lambda)$ is Owen's T function [23] which is defined as

$$T(x,\lambda) = \frac{1}{\sqrt{2\pi}} \int_0^{\lambda} \frac{e^{-\frac{1}{2}x^2\left(1+t^2\right)}}{1+t^2}dt$$

for $x, \lambda \in \mathbb{R}$.

Proof. The proof is given in [5]. □

2.3 Stochastic representation of $\mathcal{SN}$ distribution

Following the approach of Henze [20], a stochastic representation is revisited that is useful for generating random numbers from a $\mathcal{SN}\left(\mu,\sigma^2,\lambda\right)$ distribution. This provides a method to generate random numbers from $X \sim \mathcal{SN}\left(\mu,\sigma^2,\lambda\right)$ with PDF (2.2).

Theorem 6. *If $U_1 \sim \mathcal{N}(0,1)$ and $U_2 \sim \mathcal{N}(0,1)$ are two independent normal distributed random variables, then*

$$X = \frac{\lambda|U_1| + U_2}{\sqrt{1+\lambda^2}} \sim \mathcal{SN}(\lambda).$$

Proof. Let $U_1 \sim \mathcal{N}(0,1)$ and $U_2 \sim \mathcal{N}(0,1)$ be independent and let $a = \frac{\lambda}{\sqrt{1+\lambda^2}}$, $b = \frac{1}{\sqrt{1+\lambda^2}}$ and $X = \frac{\lambda|U_1|+U_2}{\sqrt{1+\lambda^2}} = a|U_1| + bU_2$.

Then

$$\begin{aligned} \mathbb{P}[X \le x] &= \mathbb{E}_{U_1}\left[\mathbb{P}[X \le x | U_1 = u_1]\right] \\ &= \int_{\mathbb{R}} \mathbb{P}[a|u_1| + bU_2 \le x]\,\phi(u_1)\,du_1 \\ &= \int_{-\infty}^{0} \mathbb{P}[a|u_1| + bU_2 \le x]\,\phi(u_1)\,du_1 + \int_0^{\infty} \mathbb{P}[a|u_1| + bU_2 \le x]\,\phi(u_1)\,du_1. \end{aligned} \tag{2.7}$$

U_1 is symmetric about $u_1 = 0$ therefore

$$\begin{aligned}\int_{-\infty}^{0} \mathbb{P}\left[a|u_1| + bU_2 \leq x\right] \phi\left(u_1\right) du_1 &= \int_{0}^{\infty} \mathbb{P}\left[a|u_1| + bU_2 \leq x\right] \phi\left(u_1\right) du_1 \\ &= \int_{0}^{\infty} \mathbb{P}\left[au_1 + bU_2 \leq x\right] \phi\left(u_1\right) du_1. \end{aligned} \tag{2.8}$$

It follows from (2.7) and (2.8) that

$$\begin{aligned}\mathbb{P}\left[X \leq x\right] &= 2\int_{0}^{\infty} \mathbb{P}\left[au_1 + bU_2 \leq x\right] \phi\left(u_1\right) du_1 \\ &= 2\int_{0}^{\infty} \mathbb{P}\left[U_2 \leq \frac{x - au_1}{b}\right] \phi\left(u_1\right) du_1 \\ &= 2\int_{0}^{\infty} \Phi\left(\frac{x - au_1}{b}\right) \phi\left(u_1\right) du_1. \end{aligned} \tag{2.9}$$

Applying a standard statistical result [10] (see Theorem 13, Appendix B.2) it follows from (2.9) that

$$\begin{aligned}f_X\left(x\right) &= \frac{d}{dx}\mathbb{P}\left[X \leq x\right] \\ &= 2\int_{0}^{\infty} \frac{d}{dx}\Phi\left(\frac{x - au_1}{b}\right) \phi\left(u_1\right) du_1 \\ &= 2\int_{0}^{\infty} \phi\left(\frac{x - au_1}{b}\right) \frac{1}{b}\phi\left(u_1\right) du_1 \\ &= 2\int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-au_1}{b}\right)^2} \frac{1}{b}\phi\left(u_1\right) du_1 \\ &= 2\int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(x^2-2xau_1+a^2u_1^2\right)}{2b^2}} \frac{1}{b}\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u_1^2} du_1 \\ &= 2\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2b^2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi b^2}} e^{-\frac{u_1^2}{2}-\frac{\left(-2xau_1+a^2u_1^2\right)}{2b^2}} du_1 \\ &= 2\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2b^2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi b^2}} e^{-\frac{\left(u_1^2b^2-2xau_1+a^2u_1^2\right)}{2b^2}} du_1 \\ &= 2\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2b^2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi b^2}} e^{-\frac{\left(u_1^2\left(a^2+b^2\right)-2xau_1\right)}{2b^2}} du_1, \text{ since } a^2+b^2=1 \\ &= 2\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2b^2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi b^2}} e^{-\frac{\left(u_1^2-2xau_1+x^2a^2\right)}{2b^2}} e^{\frac{x^2a^2}{2b^2}} du_1 \\ &= 2\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2b^2}+\frac{x^2a^2}{2b^2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi b^2}} e^{-\frac{(u_1-ax)^2}{2b^2}} du_1 \\ &= 2\frac{1}{\sqrt{2\pi}} e^{-\frac{\left(x^2-x^2a^2\right)}{2b^2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi b^2}} e^{-\frac{(u_1-ax)^2}{2b^2}} du_1 \\ &= 2\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2\left(1-a^2\right)}{2b^2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi b^2}} e^{-\frac{(u_1-ax)^2}{2b^2}} du_1 \\ &= 2\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2b^2}{2b^2}} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi b^2}} e^{-\frac{(u_1-ax)^2}{2b^2}} du_1 \\ &= 2\phi\left(x\right) \int_{0}^{\infty} \frac{1}{\sqrt{2\pi b^2}} e^{-\frac{(u_1-ax)^2}{2b^2}} du_1. \end{aligned} \tag{2.10}$$

If $w = \frac{u_1 - ax}{b}$ then $u^{-1}(w) = wb + ax$ and $\frac{d}{dw}u^{-1}(w) = b$. Note the change in the bounds of the integral. If $u_1 = 0 \implies w = \frac{-ax}{b}$. The upper bound does not change.

Applying the transformation it follows from (2.10) that

$$\begin{aligned} f_X(x) &= 2\phi(x)\int_{\frac{-ax}{b}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw \\ &= 2\phi(x)\int_{\frac{-ax}{b}}^{\infty}\phi(w)\,dw \\ &= 2\phi(x)\lim_{k\to\infty}\Phi(w)|_{\frac{-ax}{b}}^{k} \\ &= 2\phi(x)\left(\lim_{k\to\infty}\Phi(k) - \lim_{k\to\infty}\Phi\left(\frac{-ax}{b}\right)\right) \\ &= 2\phi(x)\left(1 - \Phi\left(\frac{-ax}{b}\right)\right) \\ &= 2\phi(x)\,\Phi\left(\frac{ax}{b}\right). \end{aligned} \tag{2.11}$$

Since $\frac{a}{b} = \frac{\frac{\lambda}{\sqrt{1+\lambda^2}}}{\frac{1}{\sqrt{1+\lambda^2}}} = \lambda$ it follows from (2.11) that

$$f_X(x;\lambda) = 2\phi(x)\,\Phi(\lambda x).$$

Therefore $X \sim \mathcal{SN}(\lambda)$ distribution with PDF $f_X(x;\lambda)$ as given in (2.1). □

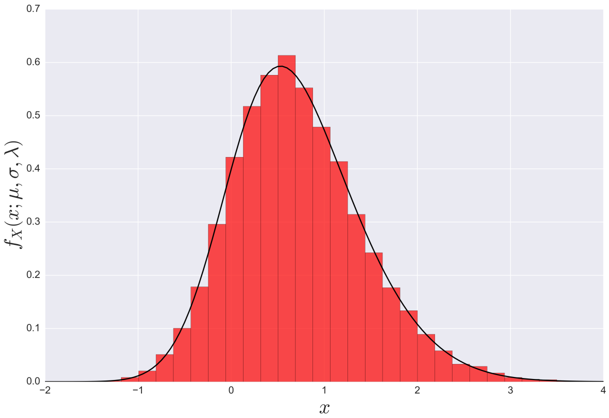
Corollary 3. *If $U_1 \sim \mathcal{N}(0,1)$ and $U_2 \sim \mathcal{N}(0,1)$ are two independent normal distributed random variables, then*

$$\begin{aligned} Y &= \mu + \sigma X \\ &= \mu + \sigma\frac{\lambda|U_1| + U_2}{\sqrt{1+\lambda^2}} \sim \mathcal{SN}\left(\mu,\sigma^2,\lambda\right) \end{aligned}$$

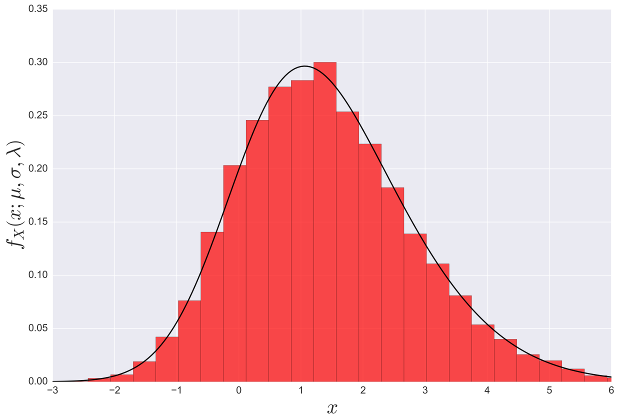
with PDF (2.2).

Since software that can generate normal distributed random variables is readily available, Theorem 6 and Corollary 3 provide a representation to easily generate random numbers from a $\mathcal{SN}$ distribution.

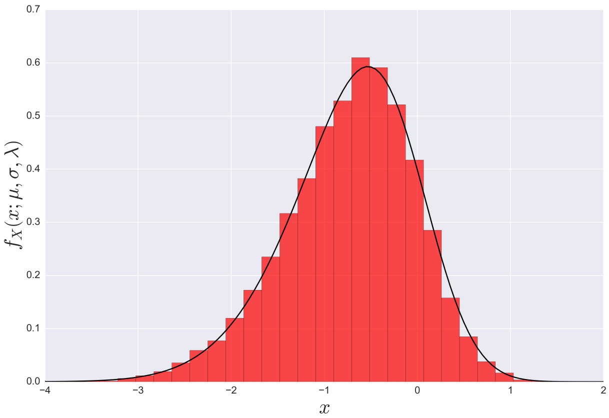
2.3.1 Visualisation of $\mathcal{SN}$ sampling scheme derived in Section 2.3



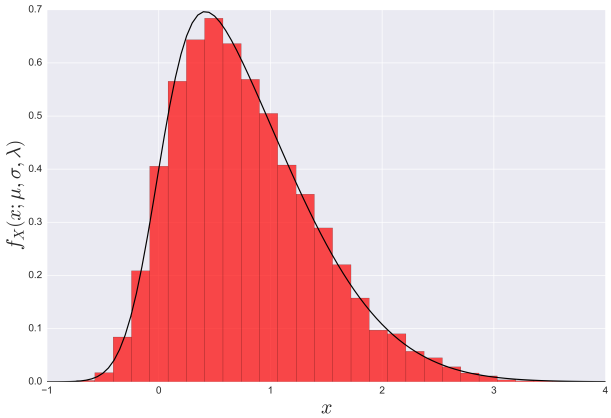
(a) Parameter set: $(\mu = 0, \sigma^2 = 1, \lambda = 2)$



(b) Parameter set: $(\mu = 0, \sigma^2 = 2, \lambda = 2)$



(c) Parameter set: $(\mu = 0, \sigma^2 = 1, \lambda = -2)$



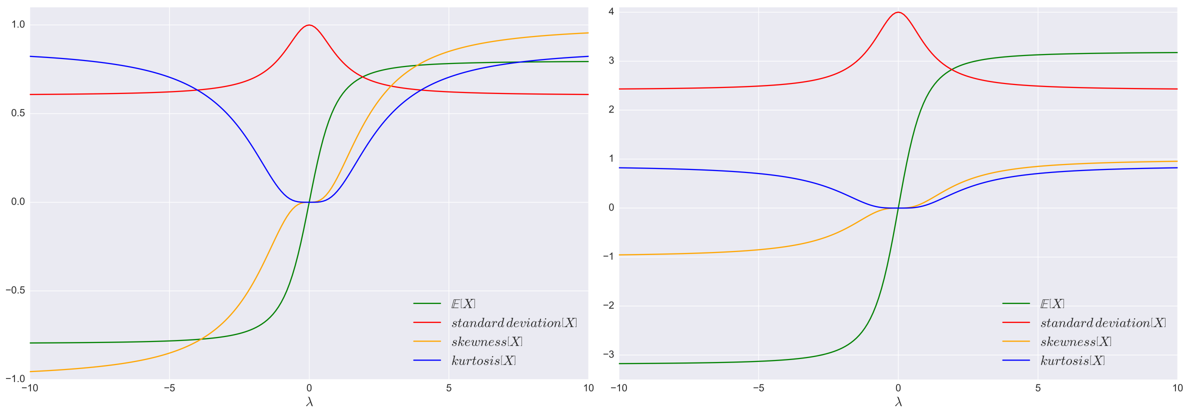
(d) Parameter set: $(\mu = 0, \sigma^2 = 1, \lambda = 4)$

Figure 2.4: Histograms of realised random samples of size 10 000 taken from $X \sim \mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ with the corresponding theoretical PDF (2.2), overlaid for different values of μ, σ^2 and λ.

Figure 2.4 shows histograms of the random samples taken from $X \sim \mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ using the stochastic representation in Corollary 3 with the corresponding theoretical PDF (2.2) overlaid.

2.4 Examining the effect of λ on the characteristics of $\mathcal{SN}$ distribution

Consider random variable $X \sim \mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ with PDF (2.2). The effect of the parameter λ on the characteristics (i.e. expected value, standard deviation, skewness and kurtosis) of the distribution shows the effect that the introduction of λ has on the distribution. Figure 2.5 shows characteristics of the $\mathcal{SN}(\mu, \sigma, \lambda)$ distribution for varying λ. The characteristics are theoretically calculated using the results obtained in Section 2.2.2.



(a) Parameter set: $(\mu = 0, \sigma^2 = 1)$ (b) Parameter set: $(\mu = 0, \sigma^2 = 16)$

Figure 2.5: Characteristics of $\mathcal{SN}$ distribution with PDF (2.2) for varying λ and specified μ and σ^2.

Similarities between Figure 2.5a and Figure 2.5b:

- expected value increases (decreases) monotonically with increasing (decreasing) values of λ;
- standard deviation is at a maximum when λ is zero and decreases towards a minimum as $|\lambda|$ increases;
- skewness is zero when $\lambda = 0$ and increases (decreases) monotonically towards a maximum (minimum) for increasing (decreasing) values of λ (see Figure 2.3a);
- kurtosis is zero when $\lambda = 0$ and increases towards a maximum as $|\lambda|$ increases (see Figure 2.3a).

Effect of increasing σ from Figure 2.5a to Figure 2.5b:

- increasing σ does not effect the respective skewness and kurtosis of the two distributions;
- the expected value in Figure 2.5b is higher (lower) for any $\lambda > 0$ $(\lambda < 0)$ than the expected value in Figure 2.5a;
- the standard deviation in Figure 2.5b is higher (lower) for any $\lambda > 0$ $(\lambda < 0)$ than the expected value in Figure 2.5a.

For any $\lambda \neq 0$, the resulting $\mathcal{SN}$ distribution will have positive kurtosis, indicating that the PDF of $\mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ as given in (2.2) will a sharper peak than the normal distribution with same mean and variance i.e. $\mathcal{N}\left(\mu, \sigma^2\right)$.

2.5 Extensions of the skew-normal model

In this section, a short overview of the extensions of the skew-normal model as illustrated in Figure 1.2 and Figure 1.3 is presented. The following structure summarises the extensions investigated in Section 2.5.1 through to Section 2.5.5.

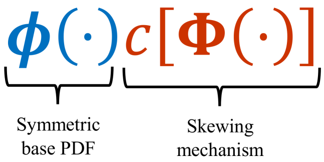


Figure 2.6: The symmetric base PDF and skewing mechanism

Figure 2.6 emphasises that the extensions of the skew-normal distribution presented in in Section 2.5.1 through to Section 2.5.5 use the normal distribution PDF as the symmetric base PDF and a particular function of the CDF of the normal distribution as the skewing mechanism. In Section 2.5.6 the PDF (2.1) and CDF (2.6) of the $\mathcal{SN}$ distribution is used in the definition of a beta generated distribution as illustrated in Figure 1.3.

2.5.1 Skew-generalised normal distribution

Arellano-Valle et. al. [3] introduced the following skew-generalised normal distribution.

Definition 1. A random variable X has the skew-generalised normal distribution if its PDF is given by

$$f_X\left(x; \lambda_1, \lambda_2\right) \quad = \quad 2\phi\left(x\right)\Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right), \; x \in \mathbb{R} \tag{2.12}$$

where $\lambda_1 \in \mathbb{R}$ and $\lambda_2 \in \mathbb{R}^+$. This is denoted by $X \sim \mathcal{SGN}\left(\lambda_1, \lambda_2\right)$.

Special cases:

- $\lambda_2 = 0$ corresponds to the $\mathcal{SN}\left(\lambda_1\right)$;
- $\lambda_1 = 0$ corresponds to the standard $N\left(0, 1\right)$ distribution;

The parameter, λ_2, can be used to adapt the tail-length of the distribution.

2.5.2 Balakrishnan skew-normal distribution

Balakrishnan [11] introduced the following generalisation of Azzalini's skew-normal [5] distribution.

Definition 2. A random variable X has the Balakrishnan skew-normal distribution if its PDF is given by

$$f_X(x; n, \lambda) = c_n(\lambda)\,\phi(x)\,\Phi^n(\lambda x),\ x \in \mathbb{R} \tag{2.13}$$

where $n \in \mathbb{R}^+$, $\lambda \in \mathbb{R}$, and

$$\begin{aligned} c_n(\lambda) &= \frac{1}{\int_{\mathbb{R}} \phi(x)\,\Phi^n(\lambda x)\,dx} \\ &= \frac{1}{\mathbb{E}_U\left[\Phi^n(\lambda U)\right]} \end{aligned}$$

with $U \sim \mathcal{N}(0,1)$. This is denoted by $X \sim \mathcal{BSN}(n, \lambda)$.

Note that for $n = 0$ and $n = 1$, the above PDF reduces to the standard normal and the Azzalini's skew-normal [5] distributions, respectively.

2.5.3 Generalised Balakrishnan skew-normal type I distribution

Hasanalipour and Sharafi [19] generalised the Balakrishnan skew-normal distribution (see Definition 2).

Definition 3. A random variable X has the generalised Balakrishnan skew-normal type I distribution if its PDF is given by

$$f_X(x; n, \lambda_1, \lambda_2) = c_n(\lambda_1, \lambda_2)\,\phi(x)\,\Phi^n\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right),\ x \in \mathbb{R} \tag{2.14}$$

where $n \in \mathbb{R}^+$, $\lambda_1 \in \mathbb{R}$, $\lambda_2 \in \mathbb{R}^+$ and

$$c_n(\lambda_1, \lambda_2) = \frac{1}{\int_{\mathbb{R}} \phi(x)\,\Phi^n\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right) dx} \tag{2.15}$$

$$= \frac{1}{\mathbb{E}_{B_1}\left[\Phi^n\left(\frac{\lambda_1 B_1}{\sqrt{1+\lambda_2 B_1^2}}\right)\right]} \tag{2.16}$$

with $B_1 \sim \mathcal{N}(0,1)$. This is denoted by $X \sim \mathcal{GBSN}_1(n, \lambda_1, \lambda_2)$.

Corollary 4. *A random variable Y has the generalised Balakrishnan skew-normal type I distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma \in \mathbb{R}^+$ if its PDF is given by*

$$f_Y(y; \mu, \sigma, n, \lambda_1, \lambda_2) = \frac{c_n(\mu,\sigma,\lambda_1,\lambda_2)}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\sigma^2+\lambda_2(y-\mu)^2}}\right),\ y \in \mathbb{R} \tag{2.17}$$

where $n \in \mathbb{R}^+$, $\lambda_1 \in \mathbb{R}$, $\lambda_2 \in \mathbb{R}^+$ *and*

$$\begin{aligned} c_n(\mu,\sigma,\lambda_1,\lambda_2) &= \frac{1}{\int_{\mathbb{R}} \frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\sigma^2+\lambda_2(y-\mu)^2}}\right)dy} \\ &= \frac{1}{\mathbb{E}_{B_2}\left[\Phi^n\left(\frac{\lambda_1(B_2-\mu)}{\sqrt{\sigma^2+\lambda_2(B_2-\mu)^2}}\right)\right]} \end{aligned}$$

with $B_2 \sim \mathcal{N}(\mu,\sigma^2)$. *This is denoted by* $Y \sim \mathcal{GBSN}_1(\mu,\sigma^2,n,\lambda_1,\lambda_2)$.

Proof. Let $X \sim \mathcal{GBSN}_1(n,\lambda_1,\lambda_2)$ with PDF (2.14). Consider the random variable $Y = \mu + \sigma X$, where the location and scale parameters are denoted $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ respectively.

If $y = \mu + \sigma x$ then $u^{-1}(y) = \frac{y-\mu}{\sigma}$. Then $\frac{d}{dy}u^{-1}(y) = \frac{1}{\sigma}$, and it follows that

$$\begin{aligned} f_Y(y;\mu,\sigma,\lambda_1,\lambda_2) &= f_X\left(u^{-1}(y);n,\lambda_1,\lambda_2\right)\left|\frac{d}{dy}u^{-1}(y)\right| \\ &= \frac{1}{\int_{\mathbb{R}}\phi\left(u^{-1}(y)\right)\Phi^n\left(\frac{\lambda_1 u^{-1}(y)}{\sqrt{1+\lambda_2(u^{-1}(y))^2}}\right)\left|\frac{d}{dy}\left(u^{-1}(y)\right)\right|dy}\phi\left(u^{-1}(y)\right) \\ &\quad \times\Phi^n\left(\frac{\lambda_1 u^{-1}(y)}{\sqrt{1+\lambda_2\left(u^{-1}(y)\right)^2}}\right)\left|\frac{d}{dy}\left(u^{-1}(y)\right)\right| \\ &= \frac{1}{\int_{\mathbb{R}}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi^n\left(\frac{\lambda_1\left(\frac{y-\mu}{\sigma}\right)}{\sqrt{1+\lambda_2\left(\frac{y-\mu}{\sigma}\right)^2}}\right)\left|\frac{1}{\sigma}\right|dy}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi^n\left(\frac{\lambda_1\left(\frac{y-\mu}{\sigma}\right)}{\sqrt{1+\lambda_2\left(\frac{y-\mu}{\sigma}\right)^2}}\right)\left|\frac{1}{\sigma}\right| \\ &= \frac{1}{\int_{\mathbb{R}}\frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\sigma^2+\lambda_2(y-\mu)^2}}\right)dy}\frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\sigma^2+\lambda_2(y-\mu)^2}}\right) \\ &= \frac{c_n(\mu,\sigma,\lambda_1,\lambda_2)}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\sigma^2+\lambda_2(y-\mu)^2}}\right) \end{aligned}$$

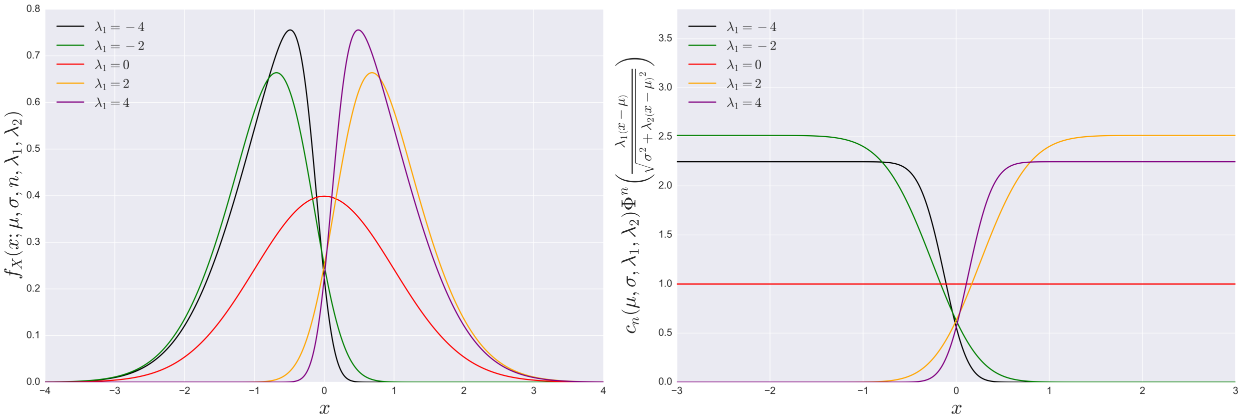
where

$$\begin{aligned} c_n(\mu,\sigma,\lambda_1,\lambda_2) &= \frac{1}{\int_{\mathbb{R}} \frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\sigma^2+\lambda_2(y-\mu)^2}}\right)dy} \\ &= \frac{1}{\mathbb{E}_{B_2}\left[\Phi^n\left(\frac{\lambda_1(B_2-\mu)}{\sqrt{\sigma^2+\lambda_2(B_2-\mu)^2}}\right)\right]} \end{aligned}$$

with $B_2 \sim \mathcal{N}(\mu,\sigma^2)$. □

$\mathcal{GBSN}_1$ PDF and skewing mechanism

Figure 2.7 - Figure 2.9 depict the PDF of the $\mathcal{GBSN}_1(\mu,\sigma^2,n,\lambda_1,\lambda_2)$ distribution i.e. $f_X(x;\mu,\sigma,n,\lambda_1,\lambda_2)$ as given in (2.17) and the corresponding skewing mechanism, $c_n(\mu,\sigma,\lambda_1,\lambda_2)\Phi^n\left(\frac{\lambda_1(x-\mu)}{\sqrt{\sigma^2+\lambda_2(x-\mu)^2}}\right)$, for varying parameter values.



(a) $\mathcal{GBSN}_1$ PDF

(b) $\mathcal{GBSN}_1$ skewing mechanism

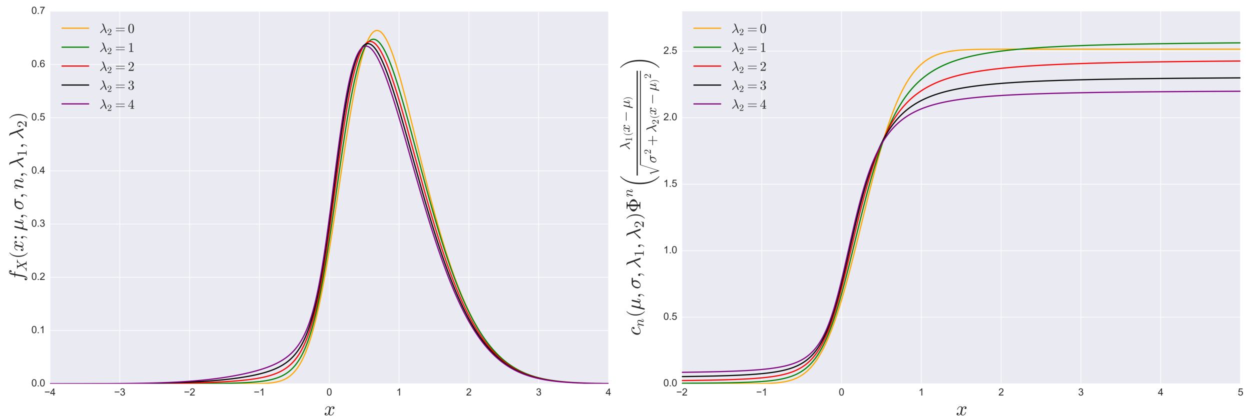
Figure 2.7: The $\mathcal{GBSN}_1$ PDF (2.17) and skewing mechanism, $c_n\left(\mu, \sigma, \lambda_1, \lambda_2\right) \Phi^n\left(\frac{\lambda_1(x-\mu)}{\sqrt{\sigma^2+\lambda_2(x-\mu)^2}}\right)$, for varying λ_1 and arbitrary $\mu = 0$, $\sigma = 1$, $n = 2$ and $\lambda_2 = 0$ (i.e. (2.13)).

Remarks

1. Comparing Figure 2.3b and Figure 2.7b, we can see that by increasing n, θ_{x_1} (as defined in Figure 2.2) increases and the skewing window (see Figure 2.2) becomes narrower resulting in the peaks of the $\mathcal{GBSN}_I$ PDF (2.17) (Figure 2.3a) attaining higher probabilities than that of the $\mathcal{SN}$ PDF (2.1) (see Figure 2.3b).

2. In Figure 2.7b, for increasing $|\lambda_1|$, the skewing window (see Figure 2.2) becomes narrower which implies that the resulting skewness is obtained by multiplying the original symmetric normal PDF by a value in the interval $(\theta_{x_0}, \theta_{x_1})$ (where x_0, x_1, θ_{x_0} and θ_{x_1} are defined in Figure 2.2) over a narrower range of x resulting in a $\mathcal{GBSN}_I$ PDF (2.17) with peaks attaining a higher probabilities.

3. For $\lambda_1 > 0$ ($\lambda_1 < 0$), a larger θ_{x_1} has the effect of pulling the left (right) tail of the $\mathcal{GBSN}_I$ PDF (2.17) closer to zero, resulting in a lighter left tail.

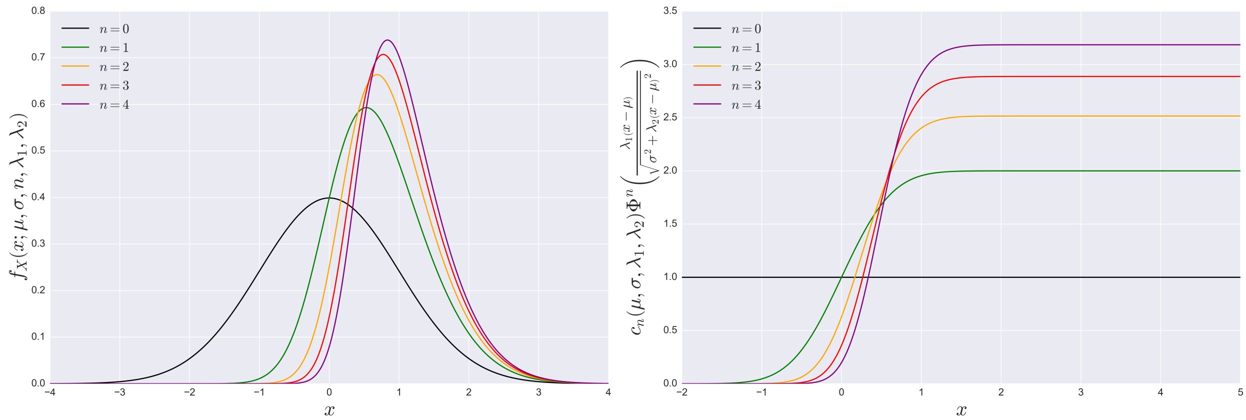


(a) $\mathcal{GBSN}_1$ PDF (b) $\mathcal{GBSN}_1$ skewing mechanism

Figure 2.8: The $\mathcal{GBSN}_1$ PDF (2.17) and skewing mechanism, $c_n(\mu, \sigma, \lambda_1, \lambda_2)\,\Phi^n\left(\frac{\lambda_1(x-\mu)}{\sqrt{\sigma^2+\lambda_2(x-\mu)^2}}\right)$, for varying λ_2 and arbitrary $\mu = 0$, $\sigma = 1$, $n = 2$ and $\lambda_1 = 2$.

Remarks

1. With $\lambda_1 > 0$ in this case, with decreasing λ_2 a larger θ_{x_1} has the effect of pulling the left tail of the $\mathcal{GBSN}_I$ PDF (see (2.17)) closer to zero, resulting in a lighter left tail.



(a) $\mathcal{GBSN}_1$ PDF (b) $\mathcal{GBSN}_1$ skewing mechanism

Figure 2.9: The $\mathcal{GBSN}_1$ PDF (2.17) and skewing mechanism, $c_n(\mu, \sigma, \lambda_1, \lambda_2)\,\Phi^n\left(\frac{\lambda_1(x-\mu)}{\sqrt{\sigma^2+\lambda_2(x-\mu)^2}}\right)$, for varying n and arbitrary $\mu = 0$, $\sigma = 1$, $\lambda_1 = 2$ and $\lambda_2 = 0$.

Remarks

1. As in Figure 2.9b, with $\lambda_1 > 0$ in this case, with increasing n a larger θ_{x_1} has the effect of pulling the left tail of the $\mathcal{GBSN}_I$ PDF (2.17) closer to zero, resulting in a shorter left tail as in Figure 2.9a.

2.5.4 Generalised Balakrishnan skew-normal type II distribution

Yadegari et. al. [31] also generalised the Balakrishnan skew-normal distribution (see Definition 2).

Definition 4. A random variable X has the generalised Balakrishnan skew-normal type II distribution if its PDF is given by

$$f_X(x;m,n,\lambda) = c_{m,n}(\lambda)\,\phi(x)\,\Phi^n(\lambda x)\,[1-\Phi(\lambda x)]^m\,,\; x\in\mathbb{R} \tag{2.18}$$

where $m,n \in \mathbb{R}^+$, $\lambda \in \mathbb{R}$ and

$$\begin{aligned} c_{m,n}(\lambda) &= \frac{1}{\int_{\mathbb{R}} \phi(x)\,\Phi^n(\lambda x)\,[1-\Phi(\lambda x)]^m\,dx} \\ &= \frac{1}{\sum_{i=0}^m \binom{m}{i}(-1)^i \int_{\mathbb{R}} \phi(x)\,[\Phi(\lambda x)]^{n+i}\,dx} \\ &= \frac{1}{\sum_{i=0}^m \binom{m}{i}(-1)^i\,\mathbb{E}_{B_3}\left[[\Phi(\lambda U)]^{n+i}\right]} \end{aligned}$$

by using a binomial expansion (refer to Theorem 15, Appendix B.2) with $B_3 \sim \mathcal{N}(0,1)$. This is denoted by $X \sim \mathcal{GBSN}_2(m,n,\lambda)$.

Note that for $m=0$ the above PDF reduces to the PDF of the $\mathcal{BSN}(n,\lambda)$ distribution i.e. $f_X(x;n,\lambda)$ as given in (2.13).

2.5.5 Generalised Balakrishnan skew-normal type III distribution

As stated in [22] another generalisation of the Balakrishnan skew-normal distribution (see Definition 2) is as follows:

Definition 5. A random variable X has the generalised Balakrishnan skew-normal type III distribution if its PDF is given by

$$f_X(x;m,n\lambda_1,\lambda_2) = c_{m,n}(\lambda_1,\lambda_2)\,\phi(x)\,\Phi^n(\lambda_1 x)\,\Phi^m(\lambda_2 x)\,,\; x\in\mathbb{R} \tag{2.19}$$

where $m,n \in \mathbb{R}^+$; $\lambda_1,\lambda_2 \in \mathbb{R}$ and

$$\begin{aligned} c_{m,n}(\lambda_1,\lambda_2) &= \frac{1}{\int_{\mathbb{R}} \phi(x)\,\Phi^n(\lambda_1 x)\,\Phi^m(\lambda_2 x)\,dx} \\ &= \frac{1}{\mathbb{E}_{B_4}\left[\Phi^n(\lambda_1 B_4)\,\Phi^m(\lambda_2 B_4)\right]} \end{aligned}$$

with $B_4 \sim \mathcal{N}(0,1)$. This is denoted by $X \sim \mathcal{GBSN}_3(m, n, \lambda_1, \lambda_2)$.

Note that for $m = 0$ the above PDF reduces to the PDF of the $\mathcal{BSN}(n, \lambda_1)$ distribution i.e. $f_X(x; n, \lambda_1)$ as given in (2.13).

2.5.6 Beta skew-normal distribution

Using the form of a beta generated distribution, Mameli [22] proposed setting $F(\cdot)$ as the CDF of $\mathcal{SN}(\mu, \sigma, \lambda)$. Then, with $f_X(x; \mu, \sigma, \lambda)$, the corresponding PDF as given in (2.2), the following definition is obtained:

Definition 6. A random variable Y has the beta skew-normal distribution with location parameter μ and scale parameter σ if its PDF is given by

$$f_Y(y; \mu, \sigma, \lambda, a, b) = \frac{1}{B(a,b)} F(y; \mu, \sigma, \lambda)^{a-1} (1 - F(y; \mu, \sigma, \lambda))^{b-1} f_Y(y; \mu, \sigma, \lambda), \ y \in \mathbb{R} \tag{2.20}$$

where $B(a, b)$ denotes the complete beta function (see Definition 15, Appendix B.1), $F(y; \mu, \sigma, \lambda)$ denotes the CDF of the $\mathcal{SN}(\mu, \sigma^2, \lambda)$ distribution with PDF (2.2), $\lambda \in \mathbb{R}$ and $a, b \geq 1$.
This is denoted by $X \sim Beta\mathcal{SN}(\mu, \sigma^2, \lambda, a, b)$.

Remark. The restriction on the parameters a and b in Definition 6 i.e. $a, b \geq 1$ ensures than the $Beta\mathcal{SN}$ PDF (2.21) is unimodal.

Corollary 5. *A random variable X has the standard beta skew-normal distribution if its PDF is given by*

$$f_X(x; \lambda, a, b) = \frac{1}{B(a,b)} F(x; \lambda)^{a-1} (1 - F(x; \lambda))^{b-1} f_X(x; \lambda), \ x \in \mathbb{R} \tag{2.21}$$

where $B(a, b)$ denotes the complete beta function (see Definition 15, Appendix B.1), $F(x; \lambda)$ denotes the CDF of the $\mathcal{SN}(\lambda)$ distribution with PDF $f_X(x; \lambda)$ as given in (2.1), $\lambda \in \mathbb{R}$ and $a, b \geq 1$. This is denoted by $X \sim Beta\mathcal{SN}(\lambda, a, b)$.

$Beta\mathcal{SN}$ PDF

Figure 2.10 and Figure 2.11 depict the PDF of the $Beta\mathcal{SN}(\mu, \sigma^2, \lambda, a, b)$ distribution i.e. $f_Y(y; \mu, \sigma, \lambda, a, b)$ as given in (2.20) for varying parameter values.

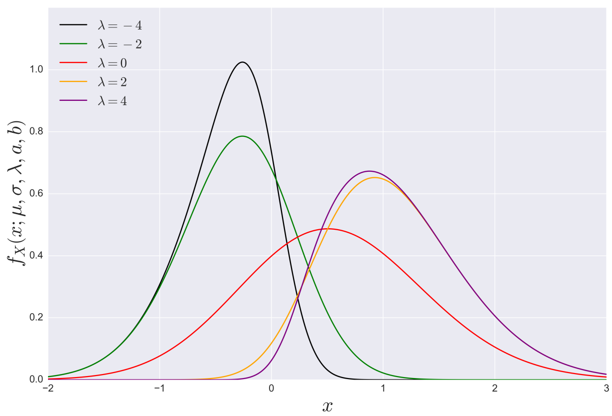


Figure 2.10: The *Beta*$\mathcal{SN}$ PDF (2.20) for varying λ and arbitrary $\mu = 0$, $\sigma^2 = 1$, $a = 2$ and $b = 1$

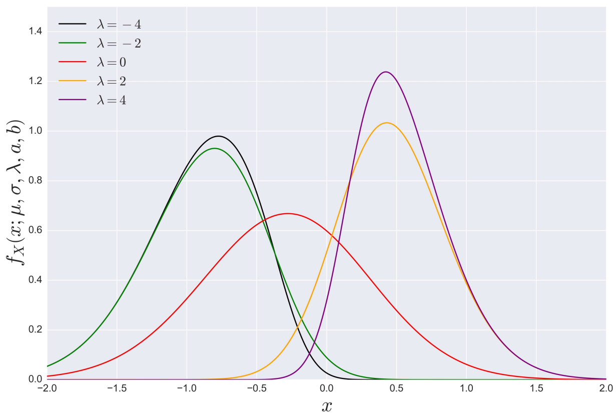


Figure 2.11: The *Beta*$\mathcal{SN}$ PDF (2.20) for varying λ and arbitrary $\mu = 0$, $\sigma^2 = 1$, $a = 2$ and $b = 3$

Remarks

1. When $\lambda = 0$, the *Beta*$\mathcal{SN}$ PDF (2.20) reduced to the beta-normal case as in Eugene et. al. [14];

2. When $a > b$, as in Figure 2.10, having $\lambda < 0$ will result in a *Beta*$\mathcal{SN}$ PDF (2.20) with peaks attaining a higher probability than a corresponding $-\lambda > 0$. For example, when $\lambda = -4$ the peak of the *Beta*$\mathcal{SN}$ PDF (2.20) attains a higher probability than that of a *Beta*$\mathcal{SN}$ PDF with $\lambda = 4$;

3. The opposite is true when $b > a$ as in Figure 2.11. Having $\lambda > 0$ will result in a *Beta*$\mathcal{SN}$ PDF (2.20) with peaks attaining a higher probability than a corresponding $-\lambda < 0$. For example, when $\lambda = 4$ the peak of the *Beta*$\mathcal{SN}$ PDF (2.20) attains a higher probability than that of a *Beta*$\mathcal{SN}$ PDF with $\lambda = -4$.

2.6 Summary

In this chapter, the skewing methodology used to define the $\mathcal{SN}$ distribution, as introduced by Azzalini [5], is investigated. The characteristics (i.e. MGF, expected value, variance, skewness and kurtosis) of this distribution are revisited, and the PDF and CDF of the skew-normal distribution is explored. The skewing mechanism that is used to develop the $\mathcal{SN}$ distribution is also investigated. A stochastic representation of the skew-normal model is developed and using this representation, a sampling scheme is employed to generate random variates from a

$\mathcal{SN}$ distribution with specified parameters. Finally, existing generalisations and extensions of the $\mathcal{SN}$ distribution and associated skewing mechanisms, relevant to this study, are addressed.

Chapter 3

Skew generalised-normal type I distribution

Azzalini [7] remarked that the $\mathcal{SN}$ distribution has short tails making it unsuitable for use when there is a need for a model to have heavier tails than the normal distribution. One method to solve this problem is to use a *symmetric base* PDF f_0, (see Proposition 1, Section 1.2) with heavier tails than the normal distribution. Following this motivation, this study focuses on the generalised normal distribution (introduced by Subbotin [28]) which is flexible enough to allow for tails heavier than that of the normal distribution. In Section 3.1, the generalised normal distribution is given and a sampling scheme to generate random variates from this distribution is proposed. A skew generalised-normal distribution is proposed in Section 3.2 and is termed the skew generalised-normal type I ($\mathcal{SGN}_I$) distribution in order to differentiate from the distribution discussed in Chapter 4.1. The effect of particular parameters on the characteristics of the $\mathcal{SGN}_I$ distribution is investigated in Section 3.3 and Section 3.4 respectively. The acceptance-rejection algorithm is presented and used to sample from the $\mathcal{SGN}_I$ distribution and a visual representation of this method is presented in Section 3.5. Two methods which are used to derive expressions for the characteristics (i.e. expected value, variance, skewness and kurtosis) of this distribution are presented in Section 3.6 and Section 3.7 respectively. In Section 3.8 a numerical study is performed to evaluate the effectiveness of the two methods in approximating the characteristics of the $\mathcal{SGN}_I$ distribution. These results are compared to the characteristics which are calculated from realised random samples (using the approach in Section 3.5) drawn from the $\mathcal{SGN}_I$ distribution. In Section 3.9 the stability and efficiency of the two methods in approximating the characteristics of the $\mathcal{SGN}_I$ distribution is investigated. A stochastic representation of the $\mathcal{SGN}_I$ distribution is derived and visualisation thereof is presented in Section 3.10. In Section 3.11 the sample estimates of characteristics evaluated from a random sample from a $\mathcal{SGN}_I$ distribution versus increasing sample size is investigated.

3.1 Generalised normal distribution

In this section, the generalised normal ($\mathcal{GN}$) distribution attributed to Subbotin [28] is presented and a sampling scheme to generate random variates from this distribution is proposed.

Definition 7. A random variable X has the generalised normal distribution if its PDF is given by

$$f_X(x;\beta) = \frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)} e^{-|x|^{\beta}}, \; x \in \mathbb{R} \tag{3.1}$$

where $\Gamma(\cdot)$ denotes the gamma function (see Definition 18, Appendix B.1). This is denoted by $X \sim \mathcal{GN}(\beta)$.

Corollary 6. *A random variable Y has the generalised normal distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\alpha \in \mathbb{R}^+$ if its PDF is given by*

$$f_Y(y;\mu,\alpha,\beta) = \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)} e^{-\left|\frac{y-\mu}{\alpha}\right|^{\beta}}, \; y \in \mathbb{R} \tag{3.2}$$

where $\beta \in \mathbb{R}^+$. This is denoted by $Y \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$.

Proof. Let $X \sim \mathcal{GN}(\beta)$ with PDF $f_X(x;\beta)$ as given in (3.1). Consider the random variable $Y = \mu + \alpha X$, where the location and scale parameters are denoted $\mu \in \mathbb{R}$ and $\alpha \in \mathbb{R}^+$ respectively. If $y = \mu + \alpha x$ then $x = u^{-1}(y) = \frac{y-\mu}{\alpha}$. Then $\frac{d}{dy}u^{-1}(y) = \frac{1}{\alpha}$, and it follows that

$$\begin{aligned}
f_Y(y;\mu,\alpha,\beta) &= f_X\left(u^{-1}(y);\beta\right)\left|\frac{d}{dy}u^{-1}(y)\right| \\
&= \frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)} e^{-\left|u^{-1}(y)\right|^{\beta}} \left|\frac{d}{dy}\left(u^{-1}(y)\right)\right| \\
&= \frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)} e^{-\left|\frac{y-\mu}{\alpha}\right|^{\beta}} \left|\frac{1}{\alpha}\right| \\
&= \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)} e^{-\left|\frac{y-\mu}{\alpha}\right|^{\beta}}.
\end{aligned}$$

□

Corollary 7. *Note that when $\mu = 0$, $\alpha = \sqrt{2}$ and $\beta = 2$ in (3.2), the PDF simplifies to*

$$\begin{aligned}
f_X(x) &= \frac{2}{2\sqrt{2}\Gamma\left(\frac{1}{2}\right)} e^{-\left|\frac{x-0}{\sqrt{2}}\right|^{2}} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}^{2}} \\
&= \phi(x)
\end{aligned}$$

for $x \in \mathbb{R}$, which is the PDF of the standard normal distribution.

3.1.1 $\mathcal{GN}$ PDF

Figure 3.1 - Figure 3.3 depict the PDF of the $\mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ distribution i.e. $f_X\left(x;\mu,\alpha,\beta\right)$ as given in (3.2), for varying parameter values.

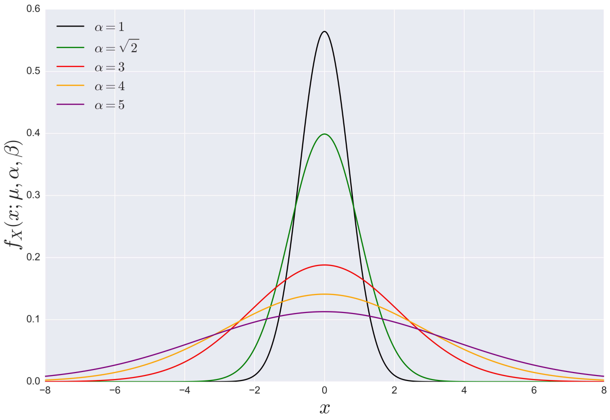


Figure 3.1: The $\mathcal{GN}$ PDF (see (3.2)) for varying α and arbitrary $\mu = 0$ and $\beta = 2$

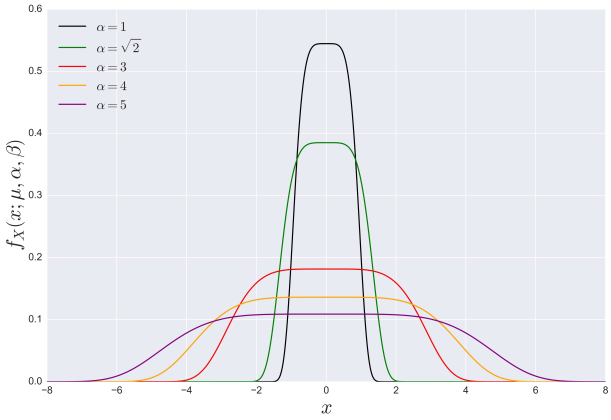


Figure 3.2: The $\mathcal{GN}$ PDF (see (3.2)) for varying α and arbitrary $\mu = 0$ and $\beta = 5$

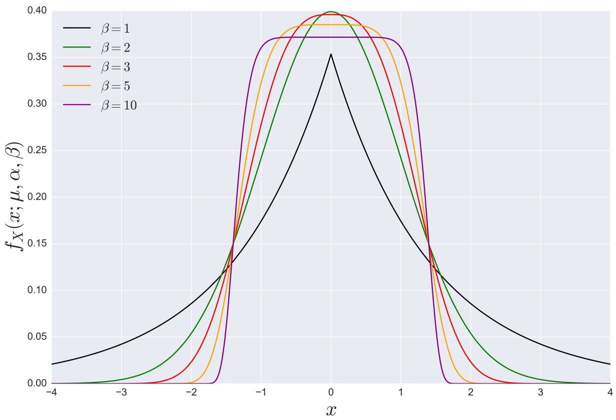


Figure 3.3: The $\mathcal{GN}$ PDF (see (3.2)) for varying β and arbitrary $\mu = 0$ and $\alpha = \sqrt{2}$

Remarks

- β affects the shape of the $\mathcal{GN}$ PDF (see (3.2));
- For $\beta > 2$ the peak of the PDF becomes flatter with shorter tails when compared to the corresponding PDF of the normal distribution, $\mathcal{N}\left(\mu, \alpha^2\right)$;
- For $\beta < 2$ the peak of the PDF becomes more pointed with longer tails when compared to the corresponding PDF of the normal distribution i.e. $\mathcal{N}\left(\mu, \alpha^2\right)$.

3.1.2 $\mathcal{GN}$ CDF

If $X \sim \mathcal{GN}(\beta)$ then the CDF is given as

$$F_X(x;\beta) = \tfrac{1}{2} + \tfrac{\text{sign}(x)}{2\Gamma\left(\frac{1}{\beta}\right)}\gamma\left(\tfrac{1}{\beta}, |x|^\beta\right), \ x \in \mathbb{R} \tag{3.3}$$

(see [2]) where $\beta \in \mathbb{R}^+$, $\Gamma(\cdot)$ denotes the gamma function (see Definition 18, Appendix B.1); $\gamma(\cdot)$ denotes the incomplete gamma function (see Definition 19, Appendix B.1) and sign$(\cdot)$ denotes the function as in Definition 20, Appendix B.1). $F_X(x;\beta)$ will be denoted as $\Phi^*(\cdot)$ from Chapter 4 onward.

3.1.3 Expected value and variance of $\mathcal{GN}$ distribution

Theorem 7. *If $X \sim \mathcal{GN}(\beta)$ with PDF $f_X(x;\beta)$ as in (2.1) then $\mathbb{E}[X] = 0$ and $var[X] = \frac{\Gamma\left(\frac{3}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}$.*

Proof. Consider

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x;\beta)\,dx \\
&= \frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)} \int_{-\infty}^{\infty} x e^{-|x|^\beta} dx \\
&= \frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)} \left(\int_{-\infty}^{0} x e^{-(-x)^\beta} dx + \int_{0}^{\infty} x e^{-x^\beta} dx \right) \\
&= \frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)} \left(\int_{0}^{\infty} -x e^{-x^\beta} dx + \int_{0}^{\infty} x e^{-x^\beta} dx \right) \\
&= \frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)} \int_{0}^{\infty} \left(-x e^{-x^\beta} + x e^{-x^\beta} \right) dx \\
&= 0.
\end{aligned} \tag{3.4}$$

Using $var[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ we calculate $\mathbb{E}[X^2]$ as

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x;\beta)\,dx \\
&= \frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)} \int_{-\infty}^{\infty} x^2 e^{-|x|^\beta} dx \\
&= \frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)} \left(\int_{-\infty}^{0} x^2 e^{-(-x)^\beta} dx + \int_{0}^{\infty} x^2 e^{-x^\beta} dx \right) \\
&= \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)} \left(\int_{0}^{\infty} x^2 e^{-x^\beta} dx \right) \\
&= \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)\beta} \left(\int_{0}^{\infty} \left(x^\beta\right)^{\frac{1}{\beta}} e^{-x^\beta} x^{2-\beta} \beta x^{\beta-1} dx \right).
\end{aligned} \tag{3.5}$$

Let $w = x^{\beta}$ then $x = w^{\frac{1}{\beta}}$ and $\frac{dw}{dx} = \beta x^{\beta-1}$ then it follows from (3.5) that

$$\begin{aligned}
\mathbb{E}\left[X^2\right] &= \frac{1}{\Gamma\left(\frac{1}{\beta}\right)}\left(\int_0^{\infty} w^{\frac{1}{\beta}} e^{-w}\left(w^{\frac{1}{\beta}}\right)^{2-\beta} dw\right) \\
&= \frac{1}{\Gamma\left(\frac{1}{\beta}\right)}\left(\int_0^{\infty} w^{\frac{3}{\beta}-1} e^{-w} dw\right) \\
&= \frac{\Gamma\left(\frac{3}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}^{[1]}. \qquad (3.6)
\end{aligned}$$

[1]Applying Definition 18, Appendix B.1.

If follows from (3.4) and (3.6) that

$$var\left[X\right] = \mathbb{E}\left[X^2\right] - \left(\mathbb{E}\left[X\right]\right)^2 = \frac{\Gamma\left(\frac{3}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}.$$

□

3.1.4 Stochastic representation of $\mathcal{GN}$ distribution

Following a similar approach to that of Azzalini [7], a stochastic representation of the $\mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ distribution is derived. This provides a method to generate random numbers from $X \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ with PDF (3.2).

Theorem 8. *Let $X \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$. Then*

$$X = \begin{cases} \mu + \alpha Z^{\frac{1}{\beta}} & \text{, with probability } \frac{1}{2} \\ \mu - \alpha Z^{\frac{1}{\beta}} & \text{, with probability } \frac{1}{2} \end{cases}$$

where $Z \sim Gamma\left(\frac{1}{\beta}, 1\right)$ (see Definition 16, Appendix B.1).

Proof. Consider $Z = \left|\frac{X-\mu}{\alpha}\right|^{\beta}$. Then $|X-\mu| = \alpha Z^{\frac{1}{\beta}}$which can be written as

$$X = \begin{cases} \mu + \alpha Z^{\frac{1}{\beta}} & , \mathbb{P}\left[X-\mu > 0\right] = \frac{1}{2} \\ \mu - \alpha Z^{\frac{1}{\beta}} & , \mathbb{P}\left[X-\mu < 0\right] = \frac{1}{2} \end{cases}$$

The property, $\mathbb{P}\left[X-\mu > 0\right] = \mathbb{P}\left[X-\mu < 0\right] = \frac{1}{2}$, follows from the symmetry of $X \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ with PDF $f_X\left(x; \mu, \alpha, \beta\right)$ as given in (3.2). Subsequently the PDF of Z needs to be determined. Since this is a many-to-one transformation it is necessary to partition $A = \{x | f_X\left(x; \mu, \alpha, \beta\right) > 0\} = (-\infty, \infty)$ into two disjoint subsets, $A_1 = \{x | x < \mu\} = (-\infty, \mu)$ and $A_2 = \{x | x > \mu\} = (\mu, \infty)$.

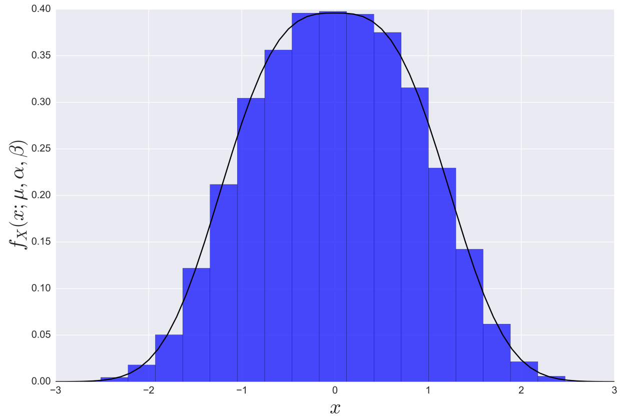
The point $x=\mu$ can be neglected in the partitioning since X is continuous. Then $z=\left|\frac{X-\mu}{\alpha}\right|^{\beta}$ has the unique solutions $x_1 = u_1^{-1}(z) = \mu - \alpha z^{\frac{1}{\beta}}$ and $x_2 = u_2^{-1}(z) = \mu + \alpha z^{\frac{1}{\beta}}$ over these respective intervals. In addition $\frac{d}{dz}u_1^{-1}(z) = -\frac{\alpha}{\beta}z^{\frac{1}{\beta}-1}$ and $\frac{d}{dz}u^{-1}(y) = \frac{\alpha}{\beta}z^{\frac{1}{\beta}-1}$.
It follows from Theorem 14, Appendix B.2 that

$$\begin{aligned}
f_Z(z) &= \sum_{j=1}^{2} f_X\left(u_j^{-1}(z)\right)\left|\frac{d}{dz}u_j^{-1}(z)\right| \\
&= f_X\left(u_1^{-1}(z)\right)\left|\frac{d}{dz}u_1^{-1}(z)\right| + f_X\left(u_2^{-1}(z)\right)\left|\frac{d}{dz}u_2^{-1}(z)\right| \\
&= \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)}e^{-\left|\frac{u_1^{-1}(z)-\mu}{\alpha}\right|^{\beta}}\left|-\frac{\alpha}{\beta}z^{\frac{1}{\beta}-1}\right| + \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)}e^{-\left|\frac{u_2^{-1}(z)-\mu}{\alpha}\right|^{\beta}}\left|\frac{\alpha}{\beta}z^{\frac{1}{\beta}-1}\right| \\
&= \left(\frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)}e^{-\left|\frac{\mu-\alpha z^{\frac{1}{\beta}}-\mu}{\alpha}\right|^{\beta}} + \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)}e^{-\left|\frac{\mu+\alpha z^{\frac{1}{\beta}}-\mu}{\alpha}\right|^{\beta}}\right)\frac{\alpha}{\beta}z^{\frac{1}{\beta}-1} \\
&= \left(\frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)}e^{-z} + \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)}e^{-z}\right)\frac{\alpha}{\beta}z^{\frac{1}{\beta}-1} \\
&= \frac{\beta}{\alpha\Gamma\left(\frac{1}{\beta}\right)}e^{-z}\frac{\alpha}{\beta}z^{\frac{1}{\beta}-1} \\
&= \frac{1^{\frac{1}{\beta}}}{\Gamma\left(\frac{1}{\beta}\right)}e^{-z}z^{\frac{1}{\beta}-1}
\end{aligned}$$

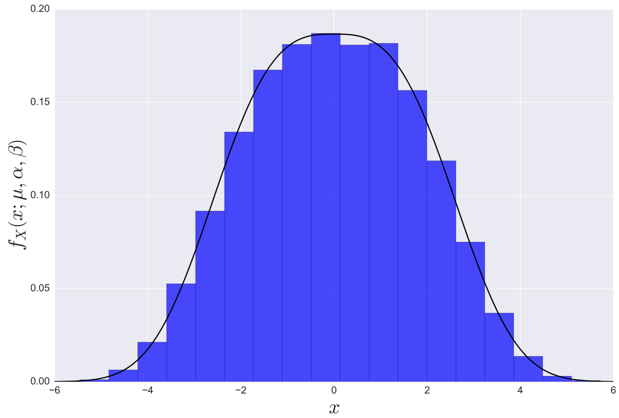
for $z \in \mathbb{R}^+$.
Therefore $Z \sim Gamma\left(\frac{1}{\beta}, 1\right)$ and the result follows. □

Since there is readily available software that can generate gamma distributed random numbers, Theorem 8 provides a representation to easily generate random numbers from a generalised normal distribution.

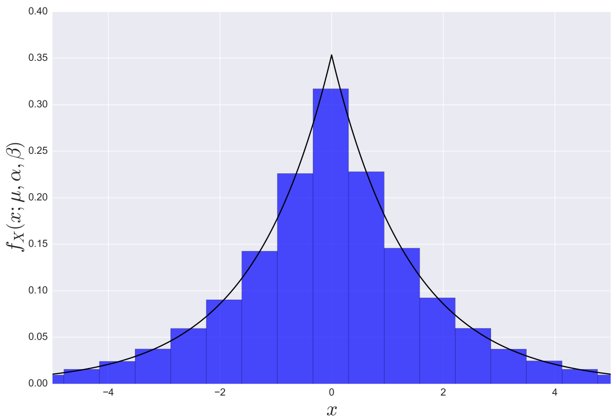
3.1.5 Visualisation of $\mathcal{GN}$ sampling scheme derived in Section 3.1.4



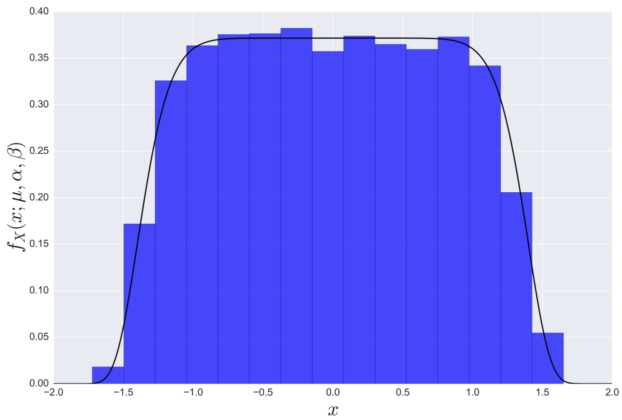
(a) Parameter set: $\left(\mu = 0, \alpha^2 = 2, \beta = 3\right)$ (see Figure 3.3)



(b) Parameter set: $\left(\mu = 0, \alpha^2 = 9, \beta = 3\right)$



(c) Parameter set: $\left(\mu = 0, \alpha^2 = 2, \beta = 1\right)$ (see Figure 3.3)



(d) Parameter set: $\left(\mu = 0, \alpha^2 = 2, \beta = 10\right)$ (see Figure 3.3)

Figure 3.4: Histograms of realised random samples of size 10 000 taken from $X \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ with the corresponding theoretical PDF (3.2), overlaid for different values of μ, α^2 and β.

Figure 3.4 shows histograms of the random samples taken from $X \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ using the stochastic representation in Theorem 8 with the corresponding theoretical PDF (3.2) overlaid.

3.2 The skew generalised-normal type I distribution

In this section, the methodology introduced in Chapter 2 is applied using the generalised normal distribution defined in Section 3.1 as the symmetric base PDF (see Figure (1.4)). Using the same notation defined in Proposition 1, Section 1.2, the case when $f_0 = \phi^*$, $G_0 = \Phi$, where $\phi^*(x;\beta)$ represents the PDF defined in (3.1), $\Phi(\cdot)$ represents the standard normal CDF, and where $w(x) = \sqrt{2}\lambda x$ for $\lambda \in \mathbb{R}$ is investigated and the following definition is obtained:

Definition 8. A random variable X has the standard skew generalised-normal type I distribution if its PDF is given by

$$f_X(x;\beta,\lambda) \;=\; 2\phi^*(x;\beta)\,\Phi\left(\sqrt{2}\lambda x\right) \quad , x\in\mathbb{R} \tag{3.7}$$

where $\beta\in\mathbb{R}^+$ and $\lambda\in\mathbb{R}$. This is denoted by $X\sim\mathcal{SGN}_I(\beta,\lambda)$.

Expanding (3.7) using the definition $\phi^*(\cdot)$ as in (3.2), it follows that

$$\begin{aligned} f_X(x;\beta,\lambda) &= 2\phi^*(x;\beta)\,\Phi\left(\sqrt{2}\lambda x\right) \\ &= 2\frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)}e^{-|x|^\beta}\Phi\left(\sqrt{2}\lambda x\right) \\ &= \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)}e^{-|x|^\beta}\Phi\left(\sqrt{2}\lambda x\right) \end{aligned} \tag{3.8}$$

for $x\in\mathbb{R}$, where $\beta\in\mathbb{R}^+$ and $\lambda\in\mathbb{R}$.

Corollary 8. *A random variable Y has the skew generalised-normal type I distribution with location parameter $\mu\in\mathbb{R}$ and scale parameter $\alpha\in\mathbb{R}^+$ if its PDF is given by*

$$f_Y(y;\mu,\alpha,\beta,\lambda) \;=\; \tfrac{2}{\alpha}\phi^*\left(\tfrac{y-\mu}{\alpha};\beta\right)\Phi\left(\sqrt{2}\lambda\left(\tfrac{y-\mu}{\alpha}\right)\right), y\in\mathbb{R} \tag{3.9}$$

where $\beta\in\mathbb{R}^+$ and $\lambda\in\mathbb{R}$. This is denoted by $Y\sim\mathcal{SGN}_I\left(\mu,\alpha^2,\beta,\lambda\right)$.

Proof. Let $X\sim\mathcal{SGN}_I(\beta,\lambda)$ with PDF (3.7). Consider the random variable $Y=\mu+\alpha X$, where the location and scale parameters are denoted $\mu\in\mathbb{R}$ and $\alpha\in\mathbb{R}^+$ respectively.
If $y=\mu+\alpha x$ then $\frac{d}{dy}u^{-1}(y)=\frac{1}{\alpha}$ and it follows that

$$\begin{aligned} f_Y(y;\mu,\alpha,\beta,\lambda) &= f_X\left(u^{-1}(y)\right)\left|\frac{d}{dy}u^{-1}(y)\right| \\ &= 2\phi^*\left(u^{-1}(y);\beta\right)\Phi\left(\sqrt{2}\lambda u^{-1}(y)\right)\left|\frac{d}{dy}\left(u^{-1}(y)\right)\right| \\ &= 2\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi\left(\sqrt{2}\lambda\left(\frac{y-\mu}{\alpha}\right)\right)\left|\frac{1}{\alpha}\right| \\ &= \frac{2}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi\left(\sqrt{2}\lambda\left(\frac{y-\mu}{\alpha}\right)\right) \end{aligned} \tag{3.10}$$

Expanding (3.9) using the definition $\phi^*(\cdot)$ as in (3.2) it follows that

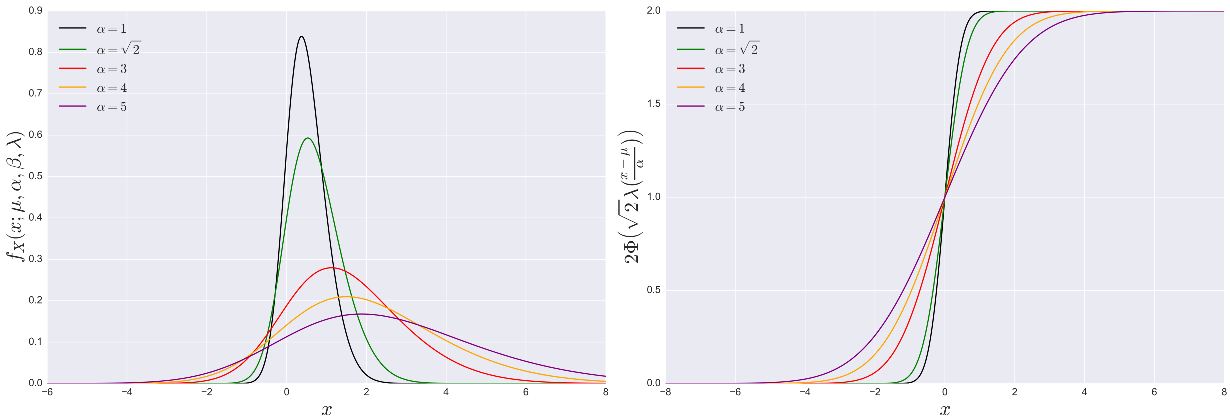
$$\begin{aligned} f_Y(y;\mu,\alpha,\beta,\lambda) &= \frac{2}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha}\right)\Phi\left(\sqrt{2}\lambda\left(\frac{y-\mu}{\alpha}\right)\right) \\ &= = \frac{2}{\alpha}\frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)}e^{-\left|\frac{y-\mu}{\alpha}\right|^\beta}\Phi\left(\sqrt{2}\lambda\left(\frac{y-\mu}{\alpha}\right)\right) \\ &= \frac{\beta}{\alpha\Gamma\left(\frac{1}{\beta}\right)}e^{-\left|\frac{y-\mu}{\alpha}\right|^\beta}\Phi\left(\sqrt{2}\lambda\left(\frac{y-\mu}{\alpha}\right)\right) \end{aligned} \tag{3.11}$$

for $y \in \mathbb{R}$, where $\mu \in \mathbb{R}$, $\alpha, \beta \in \mathbb{R}^+$ and $\lambda \in \mathbb{R}$. This is denoted by $Y \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$. $\square$

Corollary 9. *When $\mu = 0$, $\alpha = \sqrt{2}$ and $\beta = 2$ the $\mathcal{SGN}_I$ distribution with PDF* (3.11) *collapses to that of the $\mathcal{SN}$ distribution with PDF* (2.1).

3.2.1 $\mathcal{SGN}_I$ PDF

Figure 3.5 - Figure 3.8 depict the PDF of the $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ distribution as given in (3.11), and the corresponding skewing mechanism for varying parameter values.



(a) $\mathcal{SGN}_I$ PDF

(b) $\mathcal{SGN}_I$ skewing mechanism

Figure 3.5: The $\mathcal{SGN}_I$ PDF (3.11) and skewing mechanism, $2\Phi\left(\sqrt{2}\lambda\left(\frac{x-\mu}{\alpha}\right)\right)$, for varying α and arbitrary $\mu = 0$, $\beta = 2$ and $\lambda = 2$.

Remarks

1. In Figure 3.5a when $\alpha = \sqrt{2}$, the $\mathcal{SGN}_I$ distribution simplifies to the $\mathcal{SN}$ distribution with PDF (2.1) with $\mu = 0$, $\sigma = 1$ and $\lambda = 2$ (see black curve Figure 2.1a);
2. Figure 3.5a is the $\mathcal{SGN}_I$ PDF (3.11) that results when the skewing mechanism in Figure 3.5b acts on the $\mathcal{GN}$ PDF (3.2) in Figure 3.1.

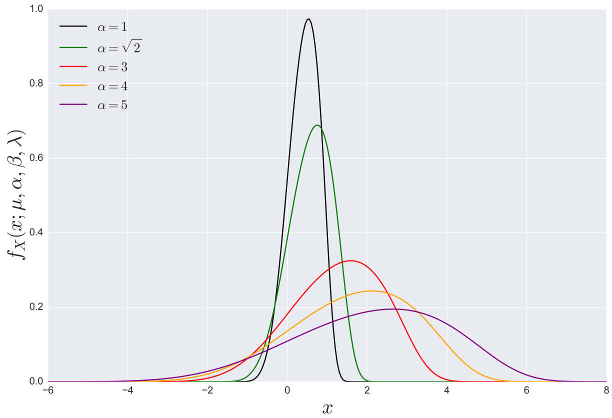


Figure 3.6: The $\mathcal{SGN}_I$ PDF (3.11) for varying α and arbitrary $\mu = 0$, $\beta = 5$ and $\lambda = 2$.

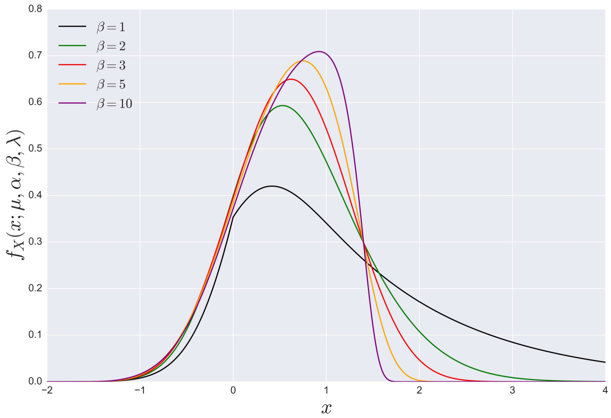
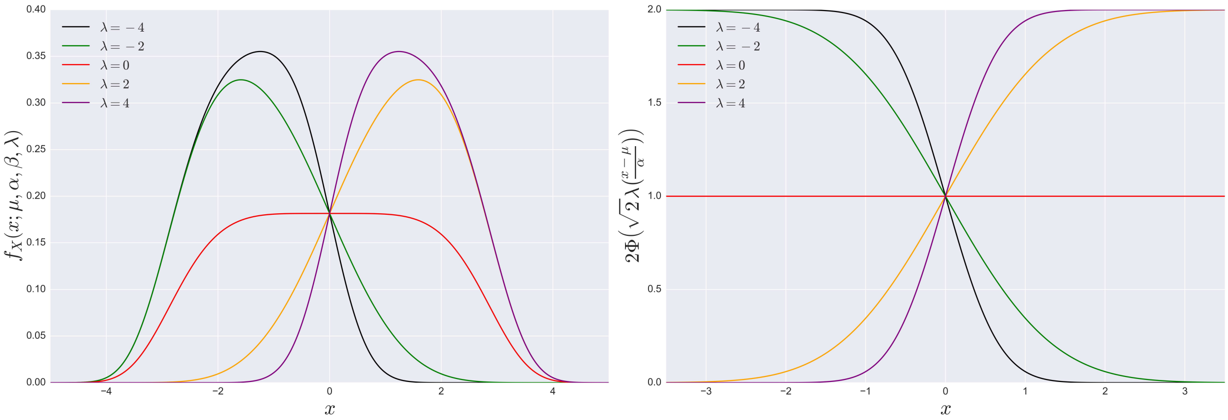


Figure 3.7: The $\mathcal{SGN}_I$ PDF (3.11) for varying β and arbitrary $\mu = 0$, $\alpha = \sqrt{2}$ and $\lambda = 2$.

Remarks

1. The accompanying $\mathcal{SGN}_I$ skewing mechanism for Figure 3.6 is not provided since it would be identical to that of Figure 3.5b since β has no impact on the skewing mechanism. Thus Figure 3.6 is the $\mathcal{SGN}_I$ PDF 3.11 that results when the skewing mechanism in Figure 3.5b acts on the $\mathcal{GN}$ PDF (3.2) in Figure 3.2.

2. The accompanying $\mathcal{SGN}_I$ skewing mechanism for Figure 3.7 is not provided since it would be identical to that of the green curve Figure 3.5b (corresponding to $\alpha = \sqrt{2}$) for all β, since β has no impact on the skewing mechanism. Thus Figure 3.7 is the $\mathcal{SGN}_I$ PDF (see 3.11) that results when the skewing mechanism for $\alpha = \sqrt{2}$ in Figure 3.5b acts on the $\mathcal{GN}$ PDF (3.2) in Figure 3.3.



(a) $\mathcal{SGN}_I$ PDF (b) $\mathcal{SGN}_I$ skewing mechanism

Figure 3.8: The $\mathcal{SGN}_I$ PDF (3.11) and skewing mechanism, $2\Phi\left(\sqrt{2}\lambda\left(\frac{x-\mu}{\alpha}\right)\right)$, for varying λ and arbitrary $\mu = 0$, $\alpha = 3$ and $\beta = 5$.

Remarks

1. Figure 3.8a is the $\mathcal{SGN}_I$ PDF 3.11 that results when the skewing mechanism in Figure 3.8b acts on the $\mathcal{GN}$ PDF (3.2) in Figure 3.2 where $\alpha = 3$;

2. In Figure 3.8b, for $\lambda = 0$ it is observed that the skewing mechanism has a value of 1 and therefore the resulting distribution is simply the original symmetric $\mathcal{GN}$ PDF (3.2) as illustrated in Figure 3.2 where $\alpha = 3$;

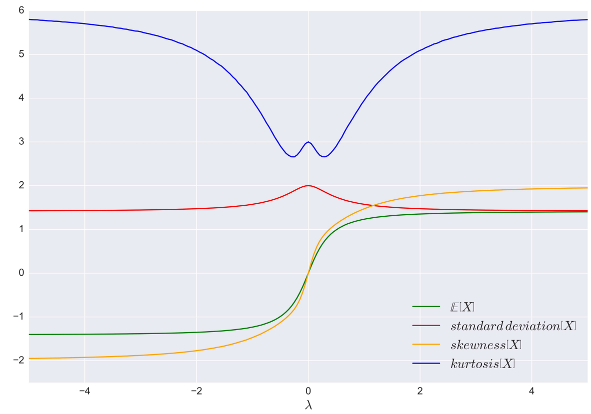
3. As can be deduced from Figure 3.8b, for increasing $|\lambda|$, the skewing window becomes narrower which implies that the resulting skewness is obtained by multiplying the original symmetric $\mathcal{GN}$ PDF (3.2) by a value in the interval $(0, 2)$ over a narrower range of x resulting in a $\mathcal{SGN}_I$ PDF 3.11 with peaks attaining higher probabilities.

3.3 Examining the effect of λ on the characteristics of $\mathcal{SGN}_I$ distribution

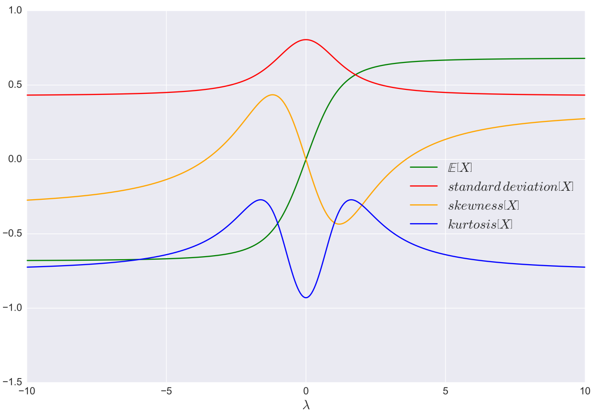
Consider random variable $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11). The effect of the parameter λ on the characteristics (i.e. expected value, standard deviation, skewness and kurtosis) of the $\mathcal{SGN}_I$ distribution is illustrated below. The characteristics are calculated by numerically integrating the function

$$\mathbb{E}\left[X^r\right] = \int_{\mathbb{R}} x^r f_X\left(x;\mu,\alpha,\beta,\lambda\right) dx \tag{3.12}$$

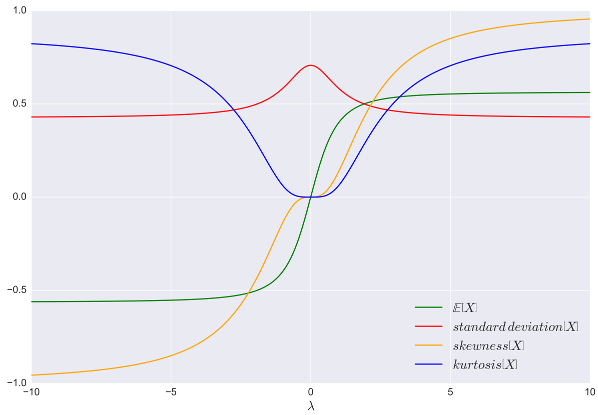
for $r = 1, 2, 3, 4$ to obtain the respective moments. These calculated moments are used in Definitions 22 - 25, Appendix B.1 to obtain the respective characteristics. It is important to note however, that for $\beta = 1$ (see Figure 3.9a), the particular numerical integration technique used in SAS to evaluate (3.12) failed. The Method 2 (as will be discussed in Section 3.7) was instead employed to approximate the characteristics. Figure 3.9 shows the characteristics of the $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ distribution versus λ.



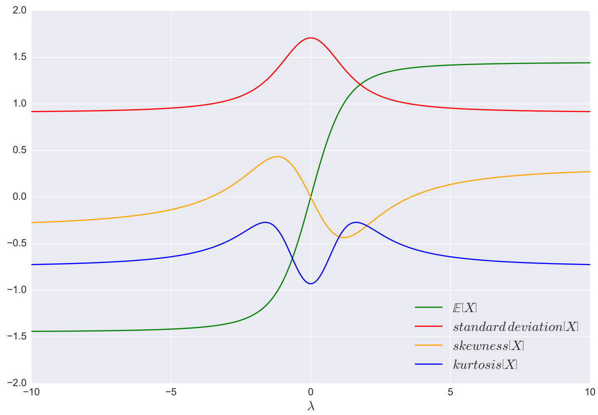
(a) Parameter set: $\left(\mu = 0, \alpha^2 = 2, \beta = 1\right)$



(b) Parameter set: $\left(\mu = 0, \alpha^2 = 2, \beta = 5\right)$



(c) Parameter set: $\left(\mu = 0, \alpha^2 = 1, \beta = 2\right)$



(d) Parameter set: $\left(\mu = 0, \alpha^2 = 9, \beta = 5\right)$

Figure 3.9: Characteristics of the $\mathcal{SGN}_I$ distribution with PDF (3.11) for varying λ and specified μ, α and β .

Remarks

1. If we had $\mu = 1$, $\alpha^2 = 2$ and $\beta = 2$ then the characteristics of $\mathcal{SGN}_I$ distribution with PDF (3.11) for varying λ would be identical to that in Figure 2.5a;

2. As illustrated by Figure 3.9b and 3.9d, the kurtosis obtained for the $\mathcal{SGN}_I$ distribution is negative for all λ, indicating that for $\beta = 5$ the $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ PDF has lighter

tails and a flatter peak than a normal distribution with same mean and variance as the respective $\mathcal{SGN}_I$ distribution;

3. As illustrated by Figure 3.9a, the kurtosis obtained for the $\mathcal{SGN}_I$ distribution is positive for all λ, indicating that for $\beta = 1$ the $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ PDF has heavier tails and a more pronounced peak than a normal distribution with same mean and variance as the respective $\mathcal{SGN}_I$ distribution;

4. As illustrated by Figure 3.9c, the kurtosis obtained for the $\mathcal{SGN}_I$ distribution is positive for $\lambda \neq 0$, indicating that for $\beta = 2$ and $\lambda \neq 0$ the $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ PDF has heavier tails and a more pronounced peak than a normal distribution with same mean and variance as the respective $\mathcal{SGN}_I$ distribution.

3.4 Examining the effect of β on the skewness of $\mathcal{SGN}_I$ distribution

It is important to note that skewness is invariant under location-scale transformations. Therefore, the only parameters that affect the skewness of $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ are β and λ. Given β, the minimum and maximum attainable skewness of $X \sim \mathcal{SGN}_I(\beta, \lambda)$ with PDF $f_X(x; \beta, \lambda)$ as in (3.8) is calculated with the following algorithm.

1. Fix a value β;

2. Use Method 2 (see Section 3.7) to approximate the skewness of $X \sim \mathcal{SGN}_I(\beta, \lambda)$ with PDF (3.8);

3. Use an optimisation strategy (e.g. Nelder–Mead method) to find a two values λ_1, λ_2 in the arbitrary range $[-50, 50]$ such that the skewness of $X \sim \mathcal{SGN}_I(\beta, \lambda)$ obtains its respective minimum and maximum.

Given a value of β, the minimum and maximum skewness attainable by varying the respective values of λ_1 and λ_2 is summarised in Table 3.1.

β	Minimum skewness	λ_1	Maximum skewness	λ_2
0.5	-4.30197	-44.21	4.30197	44.21
1	-1.9994	-50	1.9994	50
1.5	1.33585	-50	1.33585	50
2	-0.99356	-50	0.99356	48.87
2.5	-0.10750	0.6888842	0.10943	-0.606611
3	-0.21079	0.9239452	0.20991	-0.910176
4	-0.35097	1.0671345	0.35022	-1.118644
5	-0.43620	1.2053199	0.43485	-1.192504
10	-0.59082	1.3530452	0.58926	-1.287575
100	-0.66620	1.4154796	0.66644	-1.31521

Table 3.1: Approximate ranges of skewness attainable by the $\mathcal{SGN}_I(\beta, \lambda)$ distribution by varying λ for different β values.

Table 3.1 provides insight into the degree of skewness that is attainable be skewing a symmetric distribution using the methodology presented in Proposition 1, Chapter 1.

- When $\beta = 2$, the PDF (3.11) is reduced to a scaling of the $\mathcal{SN}(\lambda)$ distribution with PDF (2.1). However, as previously noted, scaling does not affect the skewness of a distribution. Thus when $\beta = 2$ the associated range of skewness is that which is attainable by the $\mathcal{SN}(\lambda)$ distribution with PDF (2.1).

- The broader range of the attainable skewness for $\beta < 2$ highlights the added flexibility in modeling skewness that the $\mathcal{SGN}_I(\beta, \lambda)$ distribution with PDF (3.11) provides over the $\mathcal{SN}(\lambda)$ distribution with PDF (2.1).

Remarks

1. When considering the $\mathcal{SN}(\lambda)$ distribution with PDF (2.1), $\lambda > 0$ and $\lambda < 0$ imply that the distribution will be respectively positively and negatively skewed. In this case skewness is a monotonically increasing function of λ as illustrated in Figure 2.5a and Figure 2.5b;

2. For $\beta \leq 2$ the $\mathcal{SGN}_I(\beta, \lambda)$ distribution with PDF (3.11) obtains its minimum and maximum skewness at the respective lower and upper edge of the considered range of λ (i.e. $\lambda \in [50, 50]$). This is due to the skewness of the $\mathcal{SGN}_I(\beta, \lambda)$ distribution with PDF (3.11) being a monotonically increasing function of λ when $\beta \leq 2$ as illustrated in Figure 3.9a and Figure 3.9c;

3. For $\beta > 2$ the $\mathcal{SGN}_I(\beta, \lambda)$ distribution with PDF (3.11) obtains its maximum skewness (i.e. most positively skewed) at a value $\lambda < 0$;

4. For $\beta > 2$ the $\mathcal{SGN}_I(\beta, \lambda)$ distribution with PDF (3.11) obtains its minimum skewness (i.e. most negatively skewed) at a value $\lambda > 0$;

5. The effect described by the two above points above is graphically represented by the yellow curve in Figure 3.9b and Figure 3.9d which can be seen to be non-monotonic functions of λ. For $\beta > 2$ this is the general behavior of skewness of the $\mathcal{SGN}_I(\beta, \lambda)$ distribution with PDF (3.11).

3.5 Sampling from the $\mathcal{SGN}_I$ distribution

The acceptance-rejection (AR) method (similar to [21]) is used to generate random numbers from a $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ distribution with PDF (3.11).

3.5.1 AR algorithm

1. Find a distribution H, with PDF, $h(x)$, with a similar range and shape to $f(x)$, for which there are methods to generate variates from;
2. Find c, the maximum value $\frac{f(x)}{h(x)}$ takes over the range of x i.e. $c = max_x \left\{\frac{f(x)}{h(x)}\right\}$;
3. Let $g(x) = \frac{f(x)}{ch(x)}$;
4. Generate random variate u from uniform distribution that covers the PDF $f(x)$;
5. Generate a random variate y from $h(x)$ independent from U;
6. If $u > g(y)$ repeat Step 4 and 5, else return $x = y$.

It is required to generate random numbers from the distribution, $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11) and the AR algorithm (see 3.5.1) in applied to this framework:

1. Find a distribution H, with PDF, $h(x)$, with a similar range and shape to $f_X(x; \mu, \alpha, \beta, \lambda)$ as given in (3.11), from which random numbers can be generated with known methods. Let $H \sim Uniform(-N, N)$ thus the PDF is, $h(x) = \frac{1}{2N}$. In implementation it is necessary that $N \to \infty$ (i.e. become 'large') in the code so that H covers the range of F.
2. Find c, the maximum value $\frac{f(x)}{h(x)}$ takes over the range i.e. $c = max_x \left\{\frac{f(x)}{h(x)}\right\}$
Therefore $f_X(x; \mu, \alpha, \beta, \lambda)$ is maximised and $h(x)$ is minimised over values of x.

In our case the maximum of $f_X(x;\mu,\alpha,\beta,\lambda)$ as given in (3.11) is found using an optimization routine and is denoted f_{max}.
Since $h(x) = \frac{1}{2N} \geq 0$ does not depend on x, $min_x\{h(x)\} = \frac{1}{2N}$
Thus $c = max_x \left\{ \frac{f_X(x;\mu,\alpha,\beta,\lambda)}{h(x)} \right\} = \frac{max_x\{f_X(x;\mu,\alpha,\beta,\lambda))\}}{min\{h(x)\}} = \frac{f_{max}}{\frac{1}{2N}} = 2Nf_{max}$.

3. Let $g(x) = \frac{f(x)}{ch(x)} = \frac{f(x)}{2Nf_{max} \times \frac{1}{2N}} = \frac{f(x)}{f_{max}}$.

4. Generate random variate u from $U \sim Uniform(0,1)$.

5. Generate a random variate y from H independent from U:

   (a) Generate random variate u^* from $U^* \sim Uniform(0,1)$ distribution;

   (b) Set $y = -N + 2Nu^*$i.e. a realisation from $H \sim Uniform(-N, N)$ distribution.

6. If $u > g(y)$ repeat Steps 4 and 5, else return $x = y$.

3.5.2 Visualisations of the AR algorithm

Figure 3.10 - Figure 3.13 visually show the results of implementing AR algorithm on SAS 9.4 software. The green and red points indicate the random numbers that were respectively accepted and rejected as candidates from the $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution with PDF (3.11). A histogram of the realised random variates drawn from $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ is plotted and the true theoretical PDF as given in (3.11) is overlaid.

Simulation 1

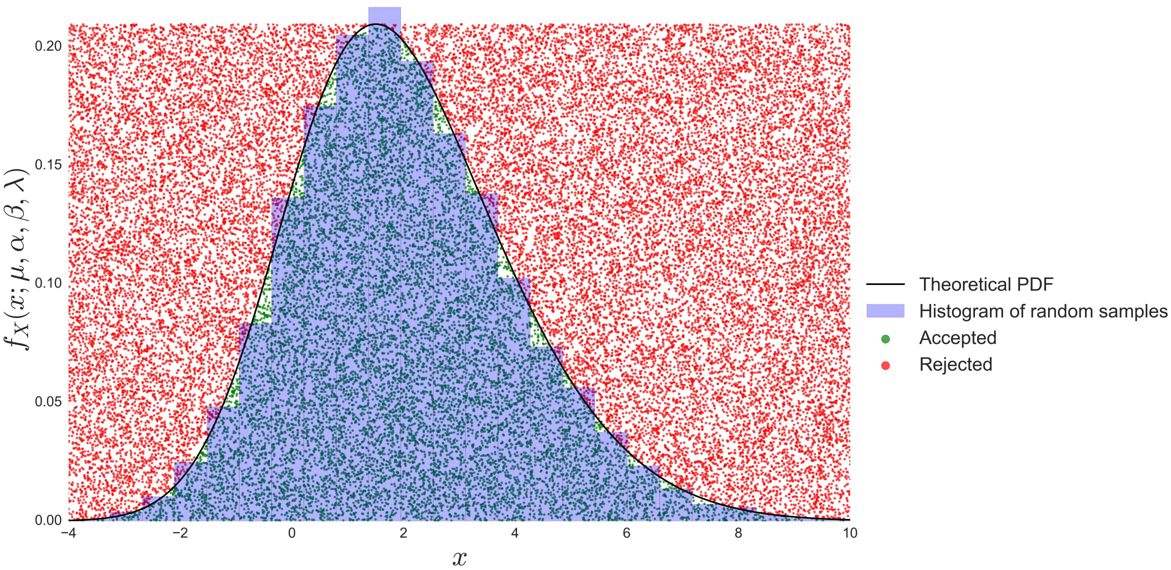


Figure 3.10: Parameter set: $(\mu = 0, \alpha^2 = 16, \beta = 2, \lambda = 2)$

Simulation 2

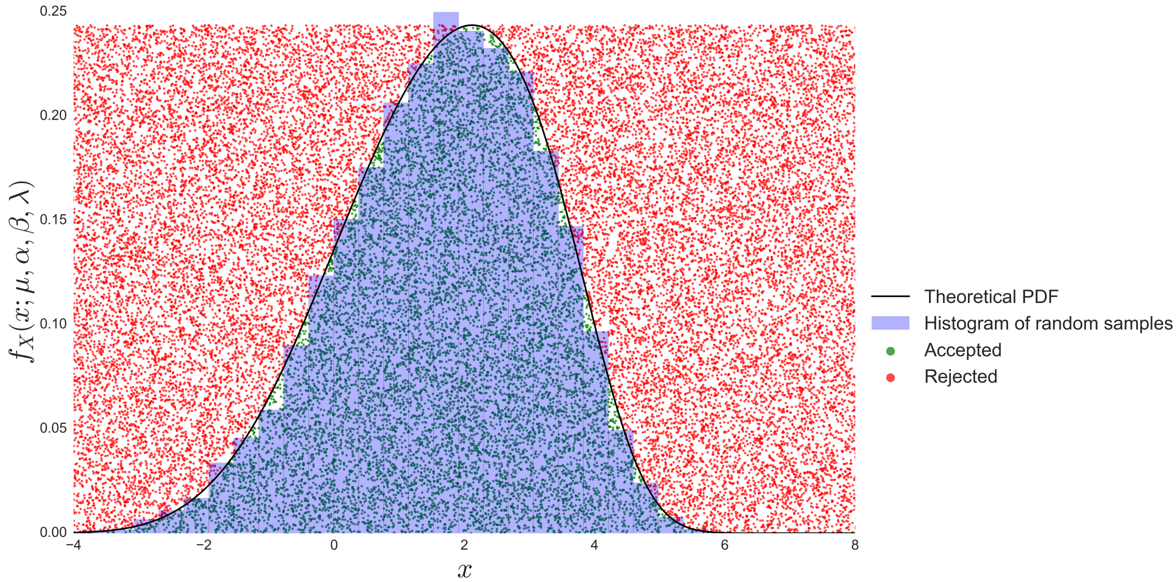


Figure 3.11: Parameter set: $\left(\mu = 0, \alpha^2 = 16, \beta = 5, \lambda = 2\right)$

Simulation 3

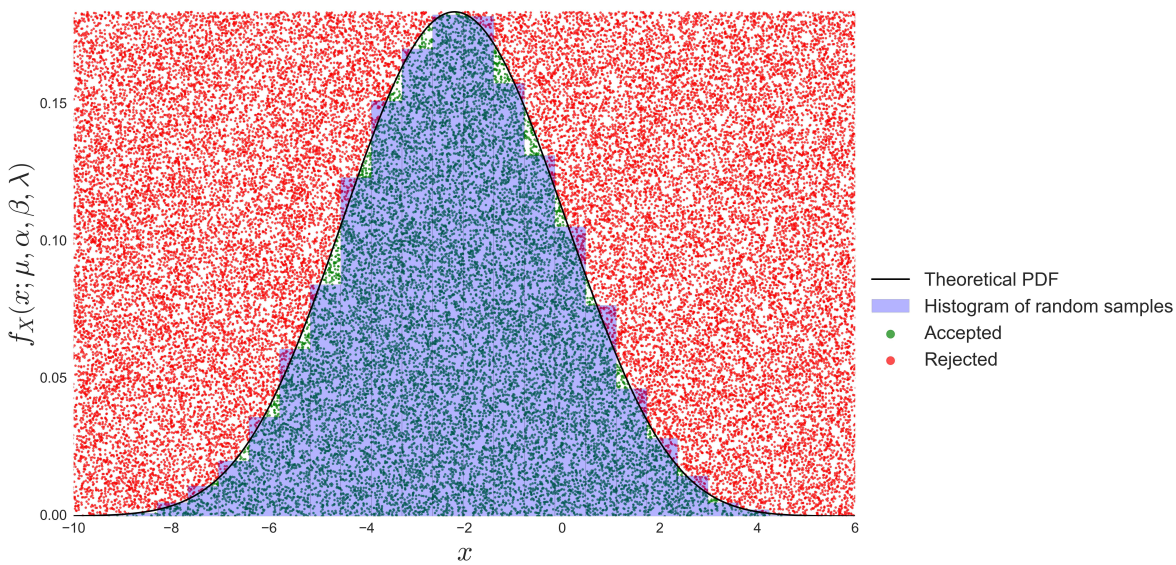


Figure 3.12: Parameter set: $\left(\mu = 0, \alpha^2 = 25, \beta = 3, \lambda = -2\right)$

Simulation 4

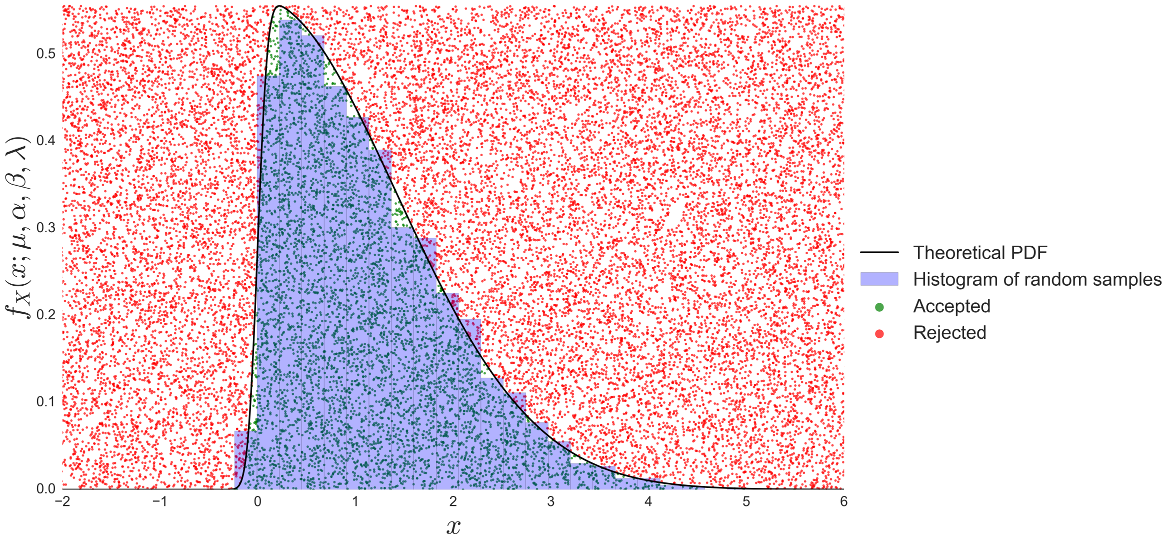


Figure 3.13: Parameter set: $\left(\mu = 0, \alpha^2 = 4, \beta = 2, \lambda = 25\right)$

Table 3.2 summarises the results of the $\mathcal{SGN}_I$ variates obtained in the above simulations.

	Simulation 1	Simulation 2	Simulation 3	Simulation 4
μ	0	0	0	0
α^2	16	16	25	4
β	2	5	3	2
λ	2	2	-2	25
Sample size (n)	13883	12000	15735	5347
Estimated mean	2.0079664	1.6660456	-2.188361	1.1362277
Estimated standard deviation	1.964543	1.5412058	2.1220783	0.8546721
Estimated skewness	0.4353946	-0.322299	-0.017579	0.9116042
Estimated kurtosis	0.2824671	-0.251794	-0.211116	0.5684229
Time taken (seconds)	1.09	0.0940001	0.1100001	0.1099999

Table 3.2: Parameter structure and analysis of $\mathcal{SGN}_I$ variates obtained in Simulation 1 through Simulation 4.

3.6 Characteristics of the $\mathcal{SGN}_I$ distribution (Method 1)

Consider a random variable $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11). Expressions for the moments, expected value, variance, skewness and kurtosis of the distribution of X (see Definition 21 - Definition 25, Appendix B.1) are now derived:

Moments of $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ distribution

Applying Definition 21, Appendix B.1 to the random variable $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11), an expression for the r^{th} non-central moments of X is obtained as follows

$$\begin{aligned}
\mathbb{E}\left[X^r\right] &= \int_{\mathbb{R}} x^r f_X\left(x; \mu, \alpha, \beta, \lambda\right) dx \\
&= \int_{\mathbb{R}} x^r \frac{\beta}{\alpha\Gamma\left(\frac{1}{\beta}\right)} e^{-\left|\frac{x-\mu}{\alpha}\right|^\beta} \Phi\left(\sqrt{2}\lambda\left(\frac{x-\mu}{\alpha}\right)\right) dx \\
&= \int_{\mathbb{R}} 2x^r \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)} e^{-\left|\frac{x-\mu}{\alpha}\right|^\beta} \Phi\left(\sqrt{2}\lambda\left(\frac{x-\mu}{\alpha}\right)\right) dx \\
&= \int_{\mathbb{R}} 2x^r f_X\left(x; \mu, \alpha, \beta\right) \Phi\left(\sqrt{2}\lambda\left(\frac{x-\mu}{\alpha}\right)\right) dx
\end{aligned}$$

where $f_X\left(x; \mu, \alpha, \beta\right)$ denotes the PDF of the $\mathcal{GN}$ distribution as given in (3.2).
Therefore

$$\begin{aligned}
\mathbb{E}\left[X^r\right] &= \int_{\mathbb{R}} 2x^r \Phi\left(\sqrt{2}\lambda\left(\frac{x-\mu}{\alpha}\right)\right) f_X\left(x; \mu, \alpha, \beta\right) dx \\
&= \mathbb{E}_{X_*}\left[2X_*^r \Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right] \qquad (3.13)
\end{aligned}$$

where $X_* \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ has PDF $f_{X_*}\left(x_*; \mu, \alpha, \beta\right)$, as in (3.2).

Expected value

From (3.13) and Definition 22, Appendix B.1 it follows that

$$\mathbb{E}\left[X\right] = \mathbb{E}_{X^*}\left[2X_* \Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]$$

where $X_* \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ has PDF (3.2).

Variance

From (3.13) it follows that

$$\mathbb{E}\left[X^2\right] = \mathbb{E}_{X^*}\left[2X_*^2 \Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]$$

and using Definition 23, Appendix B.1 it follows that

$$\begin{aligned} var\left[X\right] &= \mathbb{E}\left[X^2\right] - \left(\mathbb{E}\left[X\right]\right)^2 \\ &= \mathbb{E}_{X^*}\left[2X_*^2\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right] - \left(\mathbb{E}_{X^*}\left[2X_*\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]\right)^2 \end{aligned}$$

where $X^* \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ has PDF (3.2).

Skewness (γ_1)

From (3.13) it follows that

$$\mathbb{E}\left[X^3\right] = \mathbb{E}_{X^*}\left[2X_*^3\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]$$

and using Definition 24, Appendix B.1 it follows that

$$\begin{aligned} \gamma_1 = &\left\{\mathbb{E}_{X^*}\left[2X_*^3\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]\right. \\ &-3\mathbb{E}_{X^*}\left[2X_*\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]\mathbb{E}_{X^*}\left[2X_*^2\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right] \\ &\left.+2\left(\mathbb{E}_{X^*}\left[2X_*\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]\right)^3\right\}\left(var\left[X\right]\right)^{-\frac{3}{2}} \end{aligned}$$

where $X^* \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ has PDF (3.2).

Kurtosis (γ_2)

From (3.13) it follows that

$$\mathbb{E}\left[X^4\right] = \mathbb{E}_{X^*}\left[2X_*^4\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]$$

and using Definition 25, Appendix B.1 it follows that

$$\begin{aligned} \gamma_2 = &\left\{\mathbb{E}_{X^*}\left[2X_*^4\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]\right. \\ &-4\mathbb{E}_{X^*}\left[2X_*\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]\mathbb{E}_{X^*}\left[2X_*^3\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right] \\ &+6\left(\mathbb{E}_{X^*}\left[2X_*\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]\right)^2\mathbb{E}_{X^*}\left[2X_*^2\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right] \\ &\left.-3\left(\mathbb{E}_{X^*}\left[2X_*\Phi\left(\sqrt{2}\lambda\left(\frac{X_*-\mu}{\alpha}\right)\right)\right]\right)^4\right\}\left(var\left[X\right]\right)^{-2} - 3 \end{aligned}$$

where $X^* \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ has PDF (3.2).

3.7 Characteristics of the $\mathcal{SGN}_I$ distribution (Method 2)

Consider a random variable $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11). Alternative expressions for the moments, expected value, variance, skewness and kurtosis of the distribution of X (see Definition 21 - Definition 25, Appendix B.1) are now derived:

Theorem 9. *If* $X \sim \mathcal{SGN}_I(\beta, \lambda)$ *with PDF* (3.8) *then*

$$\mathbb{E}\left[X^k\right] = \begin{cases} \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} & , \text{for } k \text{ even} \\ \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\left\{2\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]-1\right\} & , \text{for } k \text{ odd} \end{cases} \tag{3.14}$$

where $A \sim Gamma\left(\frac{k+1}{\beta}, 1\right)$ *(see Definition 16, Appendix B.1).*

Proof. Let $X \sim \mathcal{SGN}_I(\beta, \lambda)$ with PDF $f_X(x; \beta, \lambda)$ as defined in (3.8) then

$$\begin{aligned}
\mathbb{E}\left[X^k\right] &= \int_{\mathbb{R}} x^k f_X(x;\lambda)\, dx \\
&= \int_{\mathbb{R}} x^k \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)} e^{-|x|^\beta}\Phi\left(\sqrt{2}\lambda x\right) dx \\
&= \int_0^\infty x^k \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)} e^{-x^\beta}\Phi\left(\sqrt{2}\lambda x\right) dx + \int_{-\infty}^0 x^k \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)} e^{-(-x)^\beta}\Phi\left(\sqrt{2}\lambda x\right) dx \\
&= \int_0^\infty x^k \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)} e^{-x^\beta}\Phi\left(\sqrt{2}\lambda x\right) dx + \int_0^\infty (-x)^k \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)} e^{-x^\beta}\Phi\left(-\sqrt{2}\lambda x\right) dx \\
&= \int_0^\infty x^k \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)} e^{-x^\beta}\Phi\left(\sqrt{2}\lambda x\right) dx + (-1)^r \int_0^\infty x^k \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)} e^{-x^\beta}\Phi\left(-\sqrt{2}\lambda x\right) dx \\
&= I_1 + I_2
\end{aligned}$$

where

$$I_1 = \int_0^\infty x^k \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)} e^{-x^\beta}\Phi\left(\sqrt{2}\lambda x\right) dx$$

and

$$I_2 = (-1)^k \int_0^\infty x^k \frac{\beta}{\Gamma\left(\frac{1}{\beta}\right)} e^{-x^\beta}\Phi\left(-\sqrt{2}\lambda x\right) dx$$

.

Let $a = x^{\beta}$ then $x = a^{\frac{1}{\beta}}$. Therefore, $da = \beta x^{\beta-1}dx$ and it follows that

$$\begin{aligned} I_1 &= \frac{1}{\Gamma\left(\frac{1}{\beta}\right)} \int_0^{\infty} \beta x^{k+(\beta-1)-(\beta-1)} e^{-x^{\beta}} \Phi\left(\sqrt{2}\lambda x\right) dx \\ &= \frac{1}{\Gamma\left(\frac{1}{\beta}\right)} \int_0^{\infty} x^{k-(\beta-1)} e^{-x^{\beta}} \Phi\left(\sqrt{2}\lambda x\right) \beta x^{\beta-1} dx \\ &= \frac{1}{\Gamma\left(\frac{1}{\beta}\right)} \int_0^{\infty} \left(a^{\frac{1}{\beta}}\right)^{k-(\beta-1)} e^{-a} \Phi\left(\sqrt{2}\lambda a^{\frac{1}{\beta}}\right) da \\ &= \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \int_0^{\infty} \frac{1}{\Gamma\left(\frac{n+1}{\beta}\right)} a^{\frac{k+1}{\beta}-1} e^{-a} \Phi\left(\sqrt{2}\lambda a^{\frac{1}{\beta}}\right) da \\ &= \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] \end{aligned} \tag{3.15}$$

where $A \sim Gamma\left(\frac{k+1}{\beta}, 1\right)$.

Similarly,

$$\begin{aligned} I_2 &= \frac{(-1)^k \Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \mathbb{E}_A\left[\Phi\left(-\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] \\ &= \frac{(-1)^k \Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \mathbb{E}_A\left[1-\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] \end{aligned} \tag{3.16}$$

where $A \sim Gamma\left(\frac{k+1}{\beta}, 1\right)$.

Therefore from (3.15) and (3.16) it follows that

$$\begin{aligned} \mathbb{E}\left[X^k\right] &= \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] + \frac{(-1)^k \Gamma\left(\frac{r+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \mathbb{E}_A\left[1-\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] \\ &= \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \left(\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] + (-1)^k \mathbb{E}_A\left[1-\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]\right) \\ &= \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \left(\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] + (-1)^k \left(1-\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]\right)\right) \\ &= \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \left(\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] + (-1)^k + (-1)^{k+1} \mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]\right) \\ &= \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \left\{(-1)^k + \mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]\left(1+(-1)^{k+1}\right)\right\} \\ &= \begin{cases} \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} & \text{, for } k \text{ even} \\ \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \left\{2\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]-1\right\} & \text{, for } k \text{ odd} \end{cases} \end{aligned}$$

where $A \sim Gamma\left(\frac{k+1}{\beta}, 1\right)$ (see Definition 16, Appendix B.1). □

As an additional result $\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]$, as it appears in Theorem 9, is considered.

$$\begin{aligned}
\mathbb{E}_Y\left[\Phi\left(\lambda A^{\frac{1}{\beta}}\right)\right] &= \int_{\mathbb{R}^+} \Phi\left(\sqrt{2}\lambda a^{\frac{1}{\beta}}\right) f_A(a)\, da \\
&= \int_{\mathbb{R}^+} \Phi\left(\sqrt{2}\lambda a^{\frac{1}{\beta}}\right) \frac{1}{\Gamma\left(\frac{k+1}{\beta}\right)} a^{\frac{k+1}{\beta}-1} e^{-a} da \qquad (3.17)
\end{aligned}$$

since $A \sim Gamma\left(\frac{k+1}{\beta}, 1\right)$.

Using Theorem 16, Appendix B.2 it follows from (3.17)

$$\begin{aligned}
\mathbb{E}\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] &= \int_{\mathbb{R}^+} \left(\frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{m=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^m \left(\sqrt{2}\lambda a^{\frac{1}{\beta}}\right)^{2m+1}}{m!\,(2m+1)}\right) \frac{1}{\Gamma\left(\frac{n+1}{\beta}\right)} a^{\frac{k+1}{\beta}-1} e^{-a} da \\
&= \frac{1}{2} \int_{\mathbb{R}^+} \frac{1}{\Gamma\left(\frac{k+1}{\beta}\right)} a^{\frac{k+1}{\beta}-1} e^{-a} da + \int_{\mathbb{R}^+} \frac{1}{\sqrt{2\pi}} \sum_{m=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^m \left(\sqrt{2}\lambda a^{\frac{1}{\beta}}\right)^{2m+1}}{m!\,(2m+1)} \frac{a^{\frac{k+1}{\beta}-1} e^{-a}}{\Gamma\left(\frac{k+1}{\beta}\right)} da \\
&= 0.5 + \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^+} \sum_{m=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^m \left(\sqrt{2}\lambda a^{\frac{1}{\beta}}\right)^{2m+1}}{m!\,(2m+1)} \frac{1}{\Gamma\left(\frac{k+1}{\beta}\right)} a^{\frac{k+1}{\beta}-1} e^{-a} da \\
&= 0.5 + \frac{1}{\sqrt{2\pi}\Gamma\left(\frac{k+1}{\beta}\right)} \sum_{m=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^m \left(\sqrt{2}\lambda\right)^{2m+1}}{m!\,(2k+1)} \int_{\mathbb{R}^+} a^{\frac{2m+1}{\beta}} a^{\frac{k+1}{\beta}-1} e^{-a} da \\
&= 0.5 + \frac{1}{\sqrt{2\pi}\Gamma\left(\frac{k+1}{\beta}\right)} \sum_{m=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^m \left(\sqrt{2}\lambda\right)^{2m+1}}{m!\,(2m+1)} \int_{\mathbb{R}^+} a^{\frac{k+2m+2}{\beta}-1} e^{-a} da \\
&= 0.5 + \frac{1}{\sqrt{2\pi}} \sum_{m=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^m \left(\sqrt{2}\lambda\right)^{2m+1}}{m!\,(2m+1)} \left(\frac{\Gamma\left(\frac{k+2(m+1)}{\beta}\right)}{\Gamma\left(\frac{k+1}{\beta}\right)}\right) \\
&= 0.5 + \frac{1}{\sqrt{2\pi}} \sum_{m=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^m \left(\sqrt{2}\lambda\right)^{2m+1}}{m!\,(2m+1)} \frac{\Gamma\left(\frac{k+1}{\beta} + \frac{2m+1}{\beta}\right)}{\Gamma\left(\frac{k+1}{\beta}\right)} \qquad (3.18)
\end{aligned}$$

since $\int_{\mathbb{R}^+} \frac{1}{\Gamma\left(\frac{n+1}{\beta}\right)} a^{\frac{n+1}{\beta}} e^{-a} da = 1$.

Now

$$\begin{aligned}
2\mathbb{E}\left[\Phi\left(\lambda A^{\frac{1}{\beta}}\right)\right] - 1 &= \left(1 + \frac{2}{\sqrt{2\pi}} \sum_{m=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^m \left(\sqrt{2}\lambda\right)^{2m+1}}{m!\,(2m+1)} \frac{\Gamma\left(\frac{k+1}{\beta} + \frac{2m+1}{\beta}\right)}{\Gamma\left(\frac{k+1}{\beta}\right)}\right) - 1 \\
&= \sqrt{\frac{2}{\pi}} \sum_{m=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^m \left(\sqrt{2}\lambda\right)^{2m+1}}{m!\,(2m+1)} \frac{\Gamma\left(\frac{k+1}{\beta} + \frac{2m+1}{\beta}\right)}{\Gamma\left(\frac{k+1}{\beta}\right)} \qquad (3.19)
\end{aligned}$$

Then from (3.14) and (3.19) it follows that

$$\mathbb{E}\left[X^k\right] = \begin{cases} \frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} & \text{, for } k \text{ even} \\ \sqrt{\frac{2}{\pi}}\frac{\Gamma\left(\frac{k+1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\sum_{m=0}^{\infty}\frac{\left(-\frac{1}{2}\right)^m\left(\sqrt{2}\lambda\right)^{2m+1}}{m!(2m+1)}\frac{\Gamma\left(\frac{k+1}{\beta}+\frac{2m+1}{\beta}\right)}{\Gamma\left(\frac{k+1}{\beta}\right)} & \text{, for } k \text{ odd} \end{cases} \tag{3.20}$$

where $A \sim Gamma\left(\frac{k+1}{\beta}, 1\right)$ (see Definition 16, Appendix B.1).

The sum in (3.19) only converges for certain parameter structures. Generally, β needs to be sufficiently larger than λ in order for the sum in (3.19) to be able to converge with standard computing power. The issue is that the terms $\left(\sqrt{2}\lambda\right)^{2m+1}$ and $\Gamma\left(\frac{k+1}{\beta}+\frac{2m+1}{\beta}\right)$ become too large for a computer to store in memory for large values of k.

The range of λ given β such that the infinite sum converges is obtained by calculating the infinite sum for a given β and increasing λ in increments of 0.01 until the infinite sum fails to converge. The last λ for which the infinite sum given β converges is the upper limit of λ. The lower limit of λ for a given β is calculated similarly. Table 3.3 shows values of λ for which the infinite sum in (3.20) will converge given β :

β	Range of λ such that the sum converges
1	$[-0.07, 0.07]$
2	$[-0.94, 0.94]$
3	$[-2.17, 2.17]$
4	$[-3.14, 3.14]$
5	$[-3.7, 3.7]$
10	$[-4.7, 4.7]$

Table 3.3: Values of λ for which the infinite sum in (3.20) will converge given β.

From Table 3.3 it is easy to see that the range of λ for which can be used is very limited. For this reason it is recommended to use Theorem 9 without representing the term $\left\{2\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]-1\right\}$ as the infinite sum in (3.19).

Theorem 10. *Let $X \sim \mathcal{SGN}_I(\beta,\lambda)$ and $Y = \mu + \alpha X$ then*

$$\mathbb{E}\left[Y^r\right] = \sum_{k=0}^{r}\binom{r}{k}\mu^{r-k}\alpha^k\mathbb{E}\left[X^k\right]$$

with $\mathbb{E}\left[X^k\right]$ as defined in Theorem 9.

Proof. Let $X \sim \mathcal{SGN}_I(\beta,\lambda)$ with PDF (3.8). Consider $Y = \mu+\alpha X$, then $Y \sim \mathcal{SGN}_I\left(\mu,\alpha^2,\beta,\lambda\right)$ with PDF (3.11) and it follows from the binomial theorem (see Definition 15, Appendix B.2) that

$$\begin{aligned}\mathbb{E}\left[Y^r\right] &= \mathbb{E}\left[(\mu+\alpha X)^r\right] \\ &= \mathbb{E}\left[\sum_{k=0}^{r}\binom{r}{k}\mu^{r-k}(\alpha X)^k\right] \\ &= \sum_{k=0}^{r}\binom{r}{k}\mu^{r-k}\mathbb{E}\left[(\alpha X)^k\right] \\ &= \sum_{k=0}^{r}\binom{r}{k}\mu^{r-k}\alpha^k\mathbb{E}\left[X^k\right] \qquad (3.21)\end{aligned}$$

with $\mathbb{E}\left[X^k\right]$ as defined in Theorem 9. □

The characteristics (Definitions 22 - 25, Appendix B.1) of random variable $Y \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF $f_Y\left(y; \mu, \alpha, \beta, \lambda\right)$ as given in (3.11) can be obtained using Theorem 10 and the moments of random variable $X \sim \mathcal{SGN}_I\left(\beta, \lambda\right)$ with PDF $f_X\left(x; \beta, \lambda\right)$, as given in (3.8) which are calculated using Theorem 9.

1^{st} moment

Applying Theorem 10 it follows that

$$\begin{aligned}\mathbb{E}\left[Y\right] &= \sum_{k=0}^{1}\binom{1}{k}\mu^{1-k}\alpha^k\mathbb{E}\left[X^k\right] \\ &= \binom{1}{0}\mu^{1-k}\alpha^0\mathbb{E}\left[X^0\right] + \binom{1}{1}\mu^{1-1}\alpha^1\mathbb{E}\left[X\right] \\ &= \mu + \alpha\mathbb{E}\left[X\right]. \qquad (3.22)\end{aligned}$$

Applying Theorem 9 it follows from (3.22) that

$$\mathbb{E}\left[Y\right] = \mu + \alpha\frac{\Gamma\left(\frac{2}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\left\{2\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] - 1\right\}$$

where $A \sim Gamma\left(\frac{2}{\beta}, 1\right)$.

2^{nd} moment

Applying Theorem 10 it follows that

$$\begin{aligned}\mathbb{E}\left[Y^2\right] &= \sum_{k=0}^{2}\binom{2}{k}\mu^{2-k}\alpha^k\mathbb{E}\left[X^k\right] \\ &= \binom{2}{0}\mu^2\alpha^0\mathbb{E}\left[X^0\right] + \binom{2}{1}\mu^{2-1}\alpha^1\mathbb{E}\left[X^1\right] + \binom{2}{2}\mu^{2-2}\alpha^2\mathbb{E}\left[X^2\right] \\ &= \mu^2 + 2\mu\alpha\mathbb{E}\left[X\right] + \alpha^2\mathbb{E}\left[X^2\right]. \qquad (3.23)\end{aligned}$$

Applying Theorem 9 it follows from (3.23) that

$$\mathbb{E}\left[Y^2\right] = \mu^2 + 2\mu\alpha\frac{\Gamma\left(\frac{2}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\left\{2\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]-1\right\} + \alpha^2\frac{\Gamma\left(\frac{3}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}$$

where $A \sim Gamma\left(\frac{2}{\beta}, 1\right)$.

3^{rd} moment

Applying Theorem 10 it follows that

$$\begin{aligned}\mathbb{E}\left[Y^3\right] &= \sum_{k=0}^{3}\binom{3}{k}\mu^{3-k}\alpha^k\mathbb{E}\left[X^k\right] \\ &= \mu^3 + 3\alpha\mu^2\mathbb{E}\left[X\right] + 3\alpha^2\mu\mathbb{E}\left[X^2\right] + \alpha^3\mathbb{E}\left[X^3\right]. \end{aligned} \tag{3.24}$$

Applying Theorem 9 it follows from (3.24) that

$$\begin{aligned}\mathbb{E}\left[Y^3\right] &= \mu^3 + 3\alpha\mu^2\frac{\Gamma\left(\frac{2}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\left\{2\mathbb{E}_{A_1}\left[\Phi\left(\sqrt{2}\lambda A_1^{\frac{1}{\beta}}\right)\right]-1\right\} + 3\alpha^2\mu\frac{\Gamma\left(\frac{3}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \\ &\quad +\alpha^3\frac{\Gamma\left(\frac{4}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\left\{2\mathbb{E}_{A_2}\left[\Phi\left(\sqrt{2}\lambda A_2^{\frac{1}{\beta}}\right)\right]-1\right\}\end{aligned}$$

where $A_1 \sim Gamma\left(\frac{2}{\beta}, 1\right)$ and $A_2 \sim Gamma\left(\frac{4}{\beta}, 1\right)$ with A_1 and A_2 independently generated.

4^{th} moment

Applying Theorem 10 it follows that

$$\begin{aligned}\mathbb{E}\left[Y^4\right] &= \sum_{k=0}^{4}\binom{4}{k}\mu^{4-k}\alpha^k\mathbb{E}\left[X^k\right] \\ &= \mu^4 + 4\alpha\mu^3\mathbb{E}\left[X\right] + 6\alpha^2\mu^2\mathbb{E}\left[X^2\right] + 4\alpha^3\mu\mathbb{E}\left[X^3\right] + \alpha^4\mathbb{E}\left[X^4\right]. \end{aligned} \tag{3.25}$$

Applying Theorem 9 it follows from (3.25) that

$$\begin{aligned}\mathbb{E}\left[Y^4\right] &= \mu^4 + 4\alpha\mu^3\frac{\Gamma\left(\frac{2}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\left\{2\mathbb{E}_{A_1}\left[\Phi\left(\sqrt{2}\lambda A_1^{\frac{1}{\beta}}\right)\right]-1\right\} + 6\alpha^2\mu^2\frac{\Gamma\left(\frac{3}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \\ &\quad +4\alpha^3\mu\frac{\Gamma\left(\frac{4}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\left\{2\mathbb{E}_{A_2}\left[\Phi\left(\sqrt{2}\lambda A_2^{\frac{1}{\beta}}\right)\right]-1\right\} + \alpha^4\frac{\Gamma\left(\frac{5}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}\end{aligned}$$

where $A_1 \sim Gamma\left(\frac{2}{\beta}, 1\right)$ and $A_2 \sim Gamma\left(\frac{4}{\beta}, 1\right)$ with A_1 and A_2 independently generated.

These derived moments are then used in Definitions 22 - 25, Appendix B.1 to arrive at expressions for the characteristics of random variable Y.

Remark. If the term $\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]$ where $A \sim Gamma\left(\frac{r+1}{\beta}, 1\right)$ in Theorem 9 is calculated using numerical integration as

$$\begin{aligned}\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right] &= \int_0^\infty \Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right) f_A(a)\,da \\ &= \int_0^\infty \Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right) \frac{1}{\Gamma\left(\frac{r+1}{\beta}\right)} a^{\frac{r+1}{\beta}-1} \times e^{-a} da \qquad (3.26)\end{aligned}$$

the moments of $X \sim \mathcal{SGN}_I(\beta, \lambda)$ with PDF (3.8) can be calculated without generating random numbers. For a similar approach see Section 3.3.

3.8 Comparison of Method 1, Method 2 and the AR algorithm

In this section Method 1 and Method 2 (as respectively discussed in Section 3.6 and Section 3.7) are used to approximate the characteristics of the $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11), and are compared with the results obtained using the AR algorithm in Section 3.5.

3.8.1 Numerical results

Table 3.4 summarises the sample statistics of each simulation run in Section 3.5. Note that Table 3.4 presents the same results as in Table 3.2, however, the time taken is included to compare with other methods. Random samples are drawn from $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11) using the AR algorithm as discussed in Section 3.5. The number of iterations of the AR algorithm is predefined and once the algorithm terminates, the sample statistics of the "accepted samples" are calculated. Note that the sample sizes (i.e. number of samples that are accepted in AR algorithm) are not equal - this is due to the nature of the AR algorithm resulting in a different number of "accepted samples" given a certain parameter structure.

Note - in the tables below:

- *"Est." is used as an abbreviation for "estimated";*
- *"Sim." is used as an abbreviation for "simulated";*
- *"std." is used as an abbreviation for "standard deviation";*

	Time taken (seconds)	Sample size	Est. mean	Est. std.	Est. skewness	Est. kurtosis
Sim.1	0.109	13883	2.0079664	1.9645431	0.4353946	0.2824671
Sim.2	0.094	12000	1.6660456	1.5412058	-0.322299	-0.251794
Sim.3	0.11	15735	-2.188361	2.1220783	-0.017579	-0.211116
Sim.4	0.11	5347	1.1362277	0.8546721	0.9116042	0.5684229

Table 3.4: Results using methodology in Section 3.5.

Table 3.5 summarises the approximation of sample statistics of each simulation using the Method 1 discussed in Section 3.6.

	Time taken (seconds)	Est. mean	Est. std.	Est. skewness	Est. kurtosis
Sim.1	33.362	2.0152464	1.9807983	0.4566053	0.318235
Sim.2	22.545	1.6714532	1.5454452	-0.310153	-0.307261
Sim.3	20.018	-2.216299	2.1088914	-0.010097	-0.200622
Sim.4	10.545	1.128395	0.8528525	0.9856522	0.848758

Table 3.5: Comparison of results using Method 1 in Section 3.6.

Table 3.6 summarises the approximation of sample statistics of each simulation using the Method 2 discussed in Section 3.7.

	Time taken (seconds)	Est. mean	Est. std.	Est. skewness	Est. kurtosis
Sim.1	0.094	2.0176304	1.9822128	0.4529445	0.3026684
Sim.2	3.017	1.672128	1.5472937	-0.308348	-0.309481
Sim.3	4.094	-2.209192	2.1098642	-0.012971	-0.206271
Sim.4	0.047	1.1275587	0.8535873	0.9893243	0.8622352

Table 3.6: Comparison of results using Method 2 (with random number generation) in Section 3.7.

Table 3.7 summarises the approximation of sample statistics of each simulation using the Method 2 discussed in Section 3.7 the difference from Table 3.6 being that the term $\mathbb{E}_A\left[\Phi\left(\sqrt{2}\lambda A^{\frac{1}{\beta}}\right)\right]$ where $A \sim Gamma\left(\frac{n+1}{\beta}, 1\right)$ in Theorem 9 is calculated using numerical integration on SAS 9.4. The results are obtained instantaneously with time taken less than 0.001 seconds across all simulations.

	Time taken (seconds)	Est. mean	Est. std.	Est. skewness	Est. kurtosis
Sim.1	<0.001	2.018506	1.9813211	0.4538256	0.3050503
Sim.2	<0.001	1.672798	1.5465694	-0.308567	-0.306072
Sim.3	<0.001	-2.209958	2.1090619	-0.013814	-0.206734
Sim.4	<0.001	1.1274775	0.8536946	0.9887343	0.8615734

Table 3.7: Comparison of results using Method 2 (with numerical integration) in Section 3.7.

The results in (3.7) obtained are similar to Table 3.5 and Table 3.6 barring Simulation 4. The AR algorithm yielded only 5347 sample in Simulation 4. The AR algorithm was run again with a higher number of iterations and yielded the following results which are more consistent with the results summarised in Table 3.5 and Table 3.6:

	Time taken (seconds)	Sample size	Est. mean	Est. std.	Est. skewness	Est. kurtosis
Sim.4	0.219	14946	1.1219374	0.8543913	1.0017251	0.8611195

Table 3.8: Improved results using methodology in Section 3.5 for Simulation 4

3.8.2 Discussion

- The AR algorithm generates a random sample from $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11). The characteristics are directly calculated from this sample. This method is exceptionally quick, however, it must be noted that the AR algorithm cannot draw appropriate samples from $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ given a certain parameter structure. For example, if $\left(\mu = 0, \alpha^2 = 2, \beta = 5, \lambda = -5\right)$, the AR algorithm is unable to sample from $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$. For this reason it is sometimes necessary to use either Method 1 or Method 2 to approximate the characteristics of the $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ distribution.

- **Method 1** (see Section 3.6) applies a particular function (unique for each characteristic, see (3.13)) to a random sample from $X_* \sim \mathcal{GN}\left(\mu, \alpha^2, \beta\right)$ with PDF (3.2) (using Theorem 8). Each function is then averaged over the number of samples taken in order to approximate the characteristics of $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11).

- **Method 2** (see Section 3.7) is applied in two ways:

  - **Method 2a** applies a particular function (unique for each characteristic, see (3.14)) to a random sample from a specific gamma distribution with parameters depending on the characteristic being approximated. Each function is then averaged over the number of samples taken in order to approximate the characteristics of $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11).

  - **Method 2b** uses numerical integration (see 3.26) to avoid generating random numbers to calculate characteristics of $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11).

- It is important to note that in Method 1 and Method 2a the number of samples taken increases until the sample average of the corresponding function converges according to a specified stopping criterion which is described in Section 3.9

- As can be inferred by the tables, the results provided Method 2a and Method 2b converge considerably faster than those provided by Method 1.

- Using Method 2b performed exceptionally well results being obtained instantly. However, it must be noted numerical integration may not converge for certain parameter structures. For example, the numerical integration does not yield results when approximating the characteristics of $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ where $\left\{\mu = 0,\ \alpha^2 = 4\,, \beta = 2,\ \lambda = -27\right\}$.

- Therefore in Section 3.9, Method 1 and Method 2a will be compared.

3.9 Visual comparison of Method 1 and Method 2

Figures 3.14 - 3.17 compare Method 1 and Method 2 (*using random number generation, not numerical integration*) formulated in Section 3.6 and Section 3.7 respectively. The metric that will be used to compare the two methods will be the absolute change in the approximated characteristics over an iteration. As the program executes, the 4 approximated characteristics at iteration i are written into a vector, say $\mathbf{a}_i$. Similarly, the approximated characteristics at iteration $i+1$ are written into a vector, say $\mathbf{a}_{i+1}$. The absolute sum of the elements of $[\mathbf{a}_{i+1}, -\mathbf{a}_i]$ is calculated. This provides a scalar metric that decreases in magnitude as the approximated characteristics converge in value. A stopping criterion is satisfied if the absolute change in the approximated characteristics over an iteration is less than a predefined threshold which is set to 0.0005 in this study and is represented by dashed line in Figures 3.14 - 3.17.

The plots below display the absolute change in approximated characteristics versus n, the number of random variates used to approximate the characteristics. This gives an indication of the efficiency and stability of the competing methods.

Figure 3.14: Simulation 1 - Parameter set: $\left(\mu = 0, \alpha^2 = 4, \beta = 2, \lambda = 25\right)$

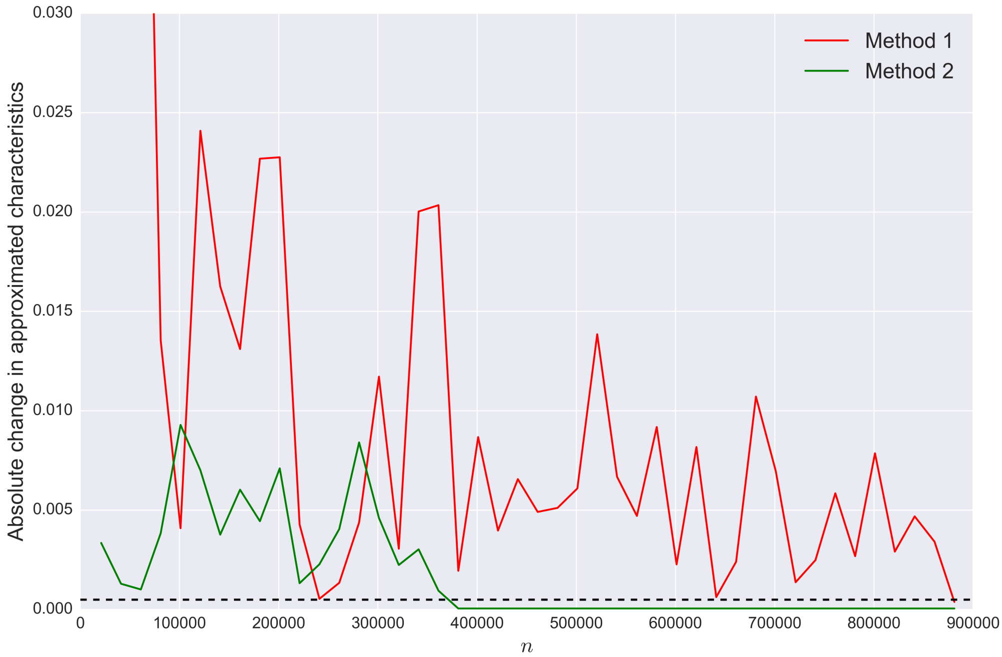


Figure 3.15: Simulation 2 - Parameter set: $\left(\mu = 0, \alpha^2 = 25, \beta = 4, \lambda = 10\right)$

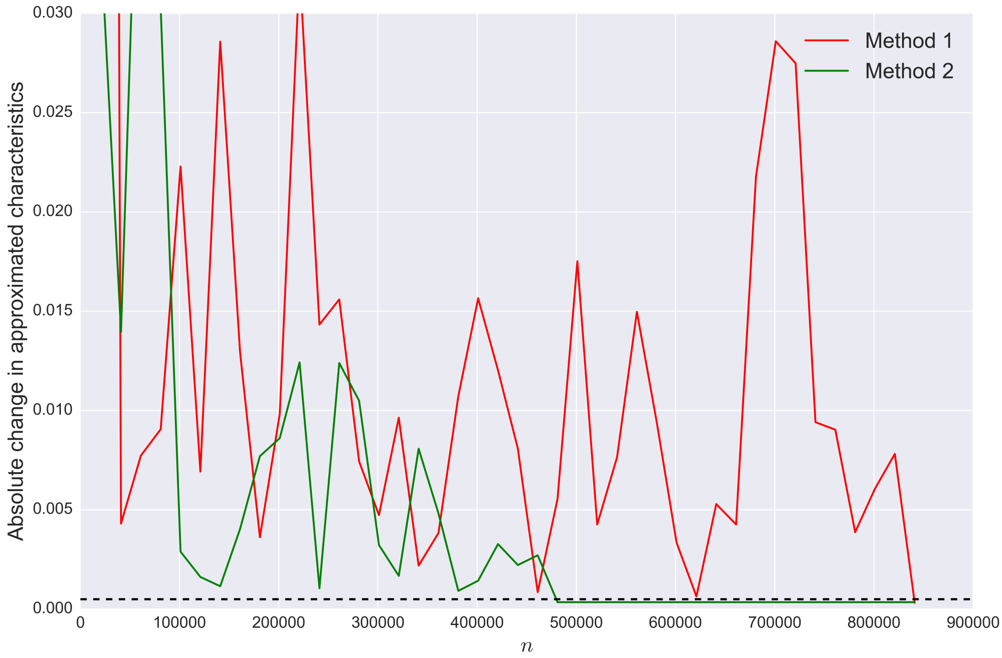


Figure 3.16: Simulation 3 - Parameter set: $\left(\mu = 0, \alpha^2 = 25, \beta = 3, \lambda = -3\right)$

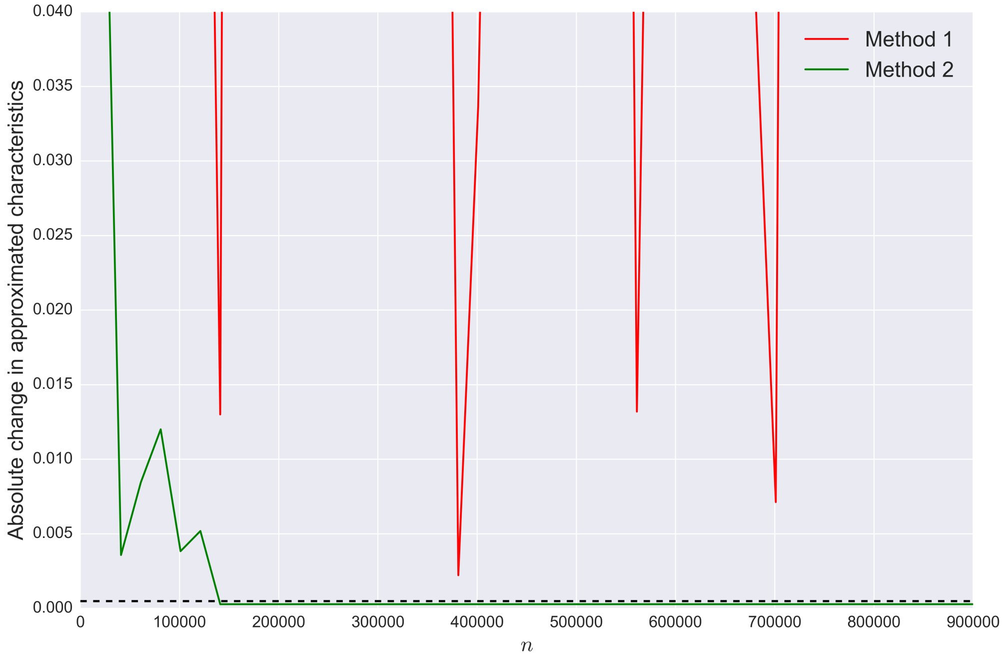


Figure 3.17: Simulation 4 - Parameter set: $\left(\mu = 0, \alpha^2 = 4, \beta = 1, \lambda = 5\right)$

The plots shows that Method 2 consistently outperforms Method 1 and is more stable than the latter. Method 2 also achieves the stopping condition before Method 1 in every simulation i.e. Method 2 approximated characteristics converge faster than those in Method 1.

3.10 Stochastic representation of $\mathcal{SGN}_I$ distribution

After noting in Section 3.5 that the AR algorithm cannot draw appropriate samples from $\mathcal{SGN}_I\left(\mu,\alpha^2,\beta,\lambda\right)$ for certain parameter structures, it was undertaken to investigate a more stable sampling scheme that generates random variates from $X\sim\mathcal{SGN}_I\left(\mu,\alpha^2,\beta,\lambda\right)$ with PDF (3.11). Following a similar approach to that of Hasanalipour [19], a stochastic representation is developed that is useful for generating random numbers from a $\mathcal{SGN}_I\left(\mu,\alpha^2,\beta,\lambda\right)$ distribution. Contrary to the AR method, this stochastic representation is able to draw random samples from $X\sim\mathcal{SGN}_I\left(\mu,\alpha^2,\beta,\lambda\right)$ with PDF (3.11) for any valid parameter structure.

Lemma 2. *Let X and Y be independenet random variables with respective PDFs $f_X(x)$ and $f_Y(y)$ both symmetric about zero. If $W=-\lambda Y$ then*

$$\begin{aligned}\mathbb{P}\left[X+W<0\right]&=\mathbb{P}\left[X-\lambda Y<0\right]\\&=\mathbb{P}\left[X<\lambda Y\right]\\&=\frac{1}{2}.\end{aligned}$$

Proof. Consider the random variable $W=-\lambda Y$ with PDF $f_W(w)$.

Case 1: $\lambda\neq 0$. For one symmetric random variable Y, we have

$$\begin{aligned}\mathbb{P}\left[W\geq w\right]&=\mathbb{P}\left[-\lambda Y\geq w\right]\\&=\mathbb{P}\left[Y\leq-\frac{w}{\lambda}\right]\\&=\mathbb{P}\left[Y\geq\frac{w}{\lambda}\right]\\&=\mathbb{P}\left[\frac{W}{-\lambda}\geq\frac{w}{\lambda}\right]\\&=\mathbb{P}\left[W\leq -w\right].\end{aligned}$$

Therefore, since $\mathbb{P}\left[W\geq w\right]=\mathbb{P}\left[W\leq -w\right]$ we have shown that $f_W(w)$ is symmetric about 0. Let $Z=X+W$ and using the convolution of marginal PDFs we obtain

$$\begin{aligned}f_Z(z)&=f_{X+W}(z)\\&=\int_{\mathbb{R}}f_X(z-w)\,f_W(w)\,dw\\&=\int_{\mathbb{R}}f_X(-(z-w))\,f_W(-w)\,dw^{[3]}\\&=\int_{\mathbb{R}}f_X(-z+w)\,f_W(-w)\,dw.\end{aligned}\tag{3.27}$$

[3] Since X and W are symmetric random variables around zero.

Let $t = -w$ then $\frac{dt}{dw} = -1$ and it follows from (3.27) that

$$\begin{aligned} f_Z(z) &= \int_{\mathbb{R}} f_X(-z-t) f_W(t) \left|-1\right| dt \\ &= \int_{\mathbb{R}} f_X(-z-t) f_W(t)\, dt. \end{aligned} \tag{3.28}$$

Let $t = w$ then $\frac{dt}{dw} = 1$ and it follows from (3.28) that

$$\begin{aligned} f_Z(z) &= \int_{\mathbb{R}} f_X(-z-w) f_W(w)\, dw \\ &= f_{X+W}(-z) \\ &= f_Z(-z). \end{aligned}$$

Therefore, we have shown the $Z = X + W = X - \lambda Y$ is a symmetric random variable around zero and it follows that

$$\begin{aligned} \mathbb{P}[Z < 0] &= \mathbb{P}[X - \lambda Y < 0] \\ &= \mathbb{P}[X < \lambda Y] \\ &= \frac{1}{2}. \end{aligned}$$

Case 2: $\lambda = 0$

$$\begin{aligned} \mathbb{P}[Z < 0] &= \mathbb{P}[X - 0Y < 0] \\ &= \mathbb{P}[X < 0] \\ &= \frac{1}{2}. \end{aligned}$$

since X is symmetric random variables around zero. □

Theorem 11. *Let $U \sim \mathcal{GN}(\beta)$ with PDF $\phi^*(x;\beta)$ as given in (3.1), and $U_1 \sim \mathcal{N}(0,1)$ with U and U_1 independent. If*

$$X = U \quad \textit{whenever} \quad U_1 \leq \sqrt{2}\lambda U$$

then $X \sim \mathcal{SGN}(\beta,\lambda)$ with PDF (3.8).

Proof. Let $X = U|\left\{U_1 \leq \sqrt{2}\lambda U\right\}$. Then

$$\begin{aligned} \mathbb{P}[X \leq x] &= \mathbb{P}\left[U \leq x| \left\{U_1 \leq \sqrt{2}\lambda U\right\}\right] \\ &= \frac{\mathbb{P}\left[U \leq x, U_1 \leq \sqrt{2}\lambda U\right]}{\mathbb{P}\left[U_1 \leq \sqrt{2}\lambda U\right]}. \end{aligned} \tag{3.29}$$

Now

$$\begin{aligned}\mathbb{P}\left[U \leq x, U_1 \leq \sqrt{2}\lambda U\right] &= \int_{u=-\infty}^{x} \int_{u_1=-\infty}^{\sqrt{2}\lambda u} \phi^*(u;\beta)\,\phi(u_1)\,du_1 du \\ &= \int_{-\infty}^{x} \phi^*(u;\beta) \left(\int_{u_1=-\infty}^{\sqrt{2}\lambda u} \phi(u_1)\,du_1 \right) du^{[1]} \\ &= \int_{-\infty}^{x} \phi^*(u;\beta)\,\Phi\left(\sqrt{2}\lambda u\right) du. \qquad (3.30)\end{aligned}$$

[1] Since U and U_1 are independent.

Since $f_U(u)$ and $f_{U_1}(u_1)$ are both symmetric about zero applying Lemma 2 it follows that

$$\mathbb{P}\left[U_1 \leq \sqrt{2}\lambda U\right] = \frac{1}{2} \qquad (3.31)$$

Then using (3.30) and (3.31) in (3.29) it follows that

$$\begin{aligned}\mathbb{P}[X \leq x] &= \frac{\int_{-\infty}^{x} \phi^*(u;\beta)\,\Phi\left(\sqrt{2}\lambda u\right) du}{0.5} \\ &= \int_{-\infty}^{x} 2\phi^*(u;\beta)\,\Phi\left(\sqrt{2}\lambda u\right) du.\end{aligned}$$

Applying a standard statistical result (Theorem 13, Appendix B.2) it follows that the PDF is

$$\begin{aligned}\frac{d}{du}\int_{-\infty}^{x} 2\phi^*(u;\beta)\,\Phi\left(\sqrt{2}\lambda \mu\right) du &= 2\phi^*(u;\beta)\,\Phi\left(\sqrt{2}\lambda u\right)\Big|_{-\infty}^{x} \\ &= 2\phi^*(x;\beta)\,\Phi\left(\sqrt{2}\lambda x\right) - \lim_{k\to-\infty}\left(2\phi^*(k;\beta)\,\Phi\left(\sqrt{2}\lambda k\right)\right) \\ &= 2\phi^*(x;\beta)\,\Phi\left(\sqrt{2}\lambda x\right) - 0 \\ &= 2\phi^*(x;\beta)\,\Phi\left(\sqrt{2}\lambda x\right)\end{aligned}$$

which is the PDF $f_X(x;\beta,\lambda)$ as given in (3.7). □

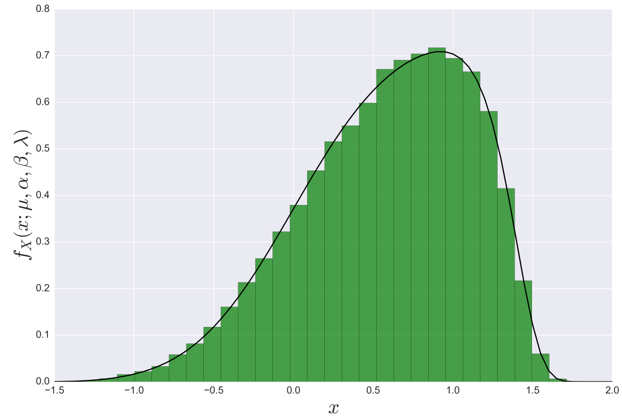
Corollary 10. *If $U \sim \mathcal{GN}(\beta)$ with PDF $\phi^*(\cdot;\beta)$ as given in (3.1) and $U_1 \sim N(0,1)$ with U and U_1 independent. If*

$$Y = \mu + \alpha U \quad \textit{whenever} \quad U_1 \leq \sqrt{2}\lambda U$$

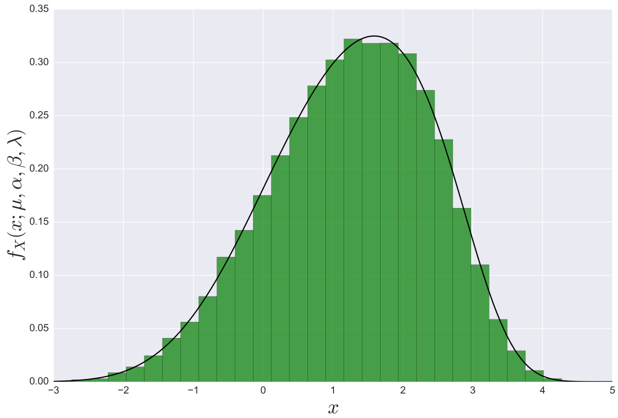
then $Y \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11).

Since there is readily available software that can generate normal distributed random numbers and the sampling scheme in Section 3.1.4 can be used to generate generalised normal random numbers. Theorem 11 and Corollary 10 provide a representation to easily generate random numbers from a $\mathcal{SGN}_I$ distribution.

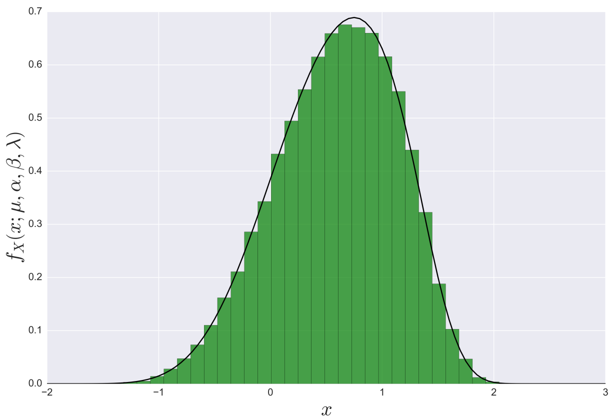
3.10.1 Visualisation of $\mathcal{SGN}_I$ sampling scheme derived in Section 3.10



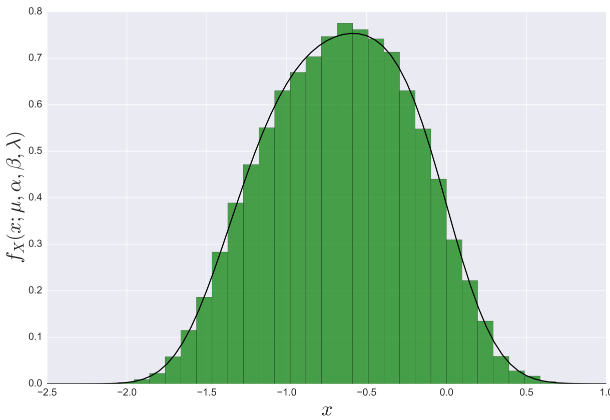
(a) Parameter set: $(\mu = 0, \alpha^2 = 2, \beta = 10, \lambda = 2)$



(b) Parameter set: $(\mu = 0, \alpha^2 = 9, \beta = 5, \lambda = 2)$



(c) Parameter set: $(\mu = 0, \alpha^2 = 2, \beta = 5, \lambda = 2)$



(d) Parameter set: $(\mu = 0, \alpha^2 = 2, \beta = 3, \lambda = -4)$

Figure 3.18: Histograms of realised random samples of size 10 000 taken from $X \sim \mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ with the corresponding theoretical PDF (3.11), overlaid for different values of μ, α^2, β and λ.

Figure 3.18 shows histograms of the random samples taken from $X \sim \mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ using the stochastic representation in Corollary 10 with the corresponding theoretical PDF (3.11) overlaid.

3.11 Convergence of sample statistics of the $\mathcal{SGN}_I$ distribution

Consider a random variable $X \sim \mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF (3.11). The estimated characteristics (i.e. expected value, standard deviation, skewness and kurtosis) of the distribution of X are plotted against the number of random variates that are sampled to estimate the characteristics, say n, to investigate the convergence of using the Method 2 described in Section 3.7.

Figures 3.19 - 3.22 show the sample estimate of a particular characteristic of the $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ versus increasing n which ranges from $n = 1000$ to $n = 100\,000$. The dotted center line of each represents the mean level of that particular characteristic over the whole simulation.

The parameter values that are used in Simulation 3 in Section 3.5.2 are used in this section.

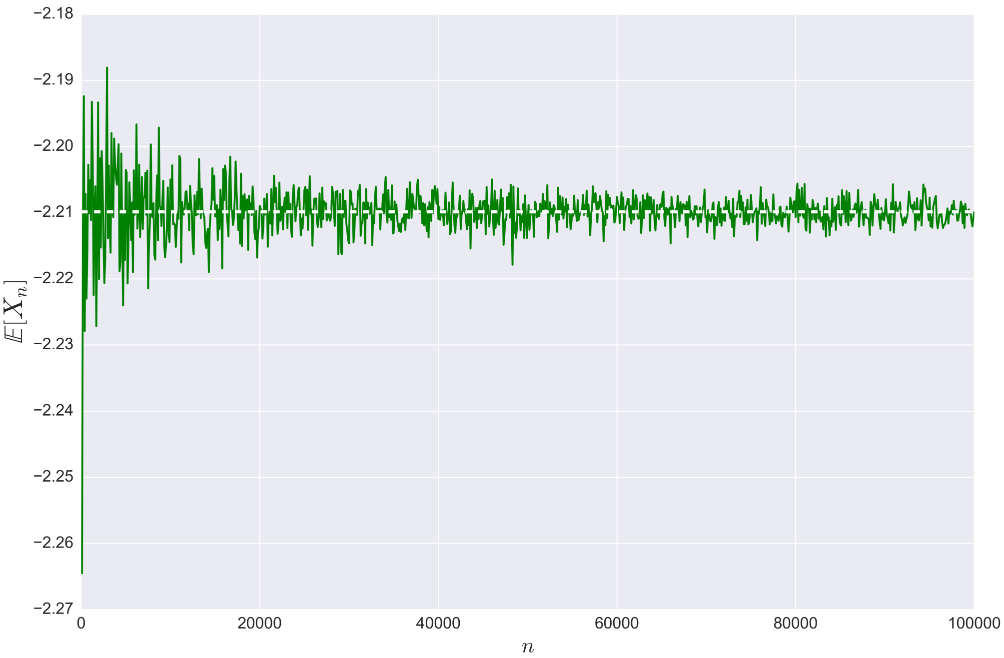


Figure 3.19: The expected value of $\mathcal{SGN}_I\left(\mu = 0, \alpha^2 = 25, \beta = 3, \lambda = -2\right)$ distribution for varying sample size n.

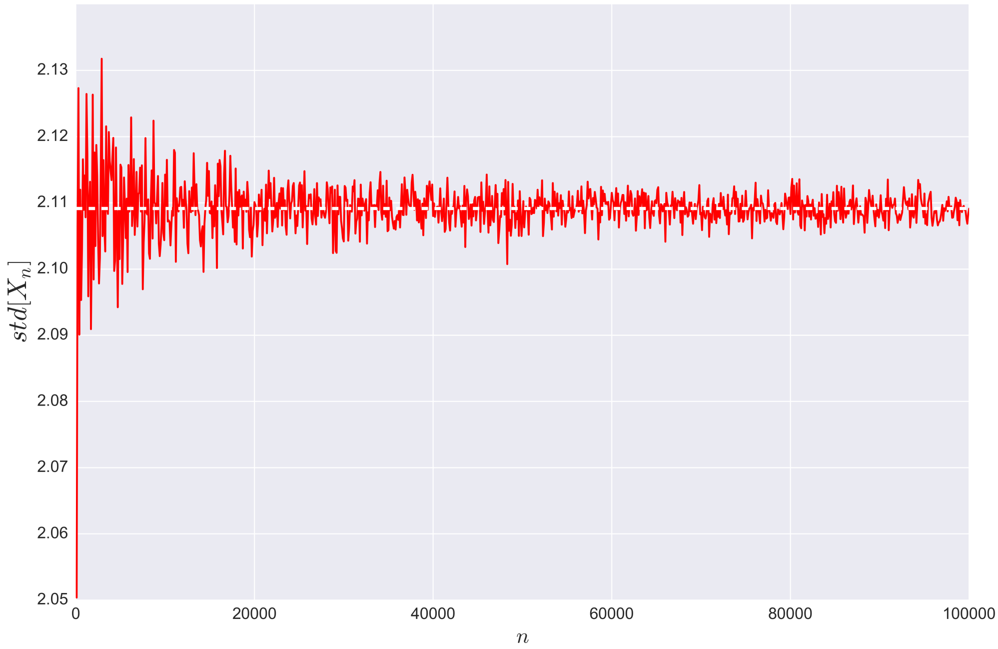


Figure 3.20: The standard deviation of $\mathcal{SGN}_I\left(\mu=0,\alpha^2=25,\beta=3,\lambda=-2\right)$ distribution for varying sample size n.

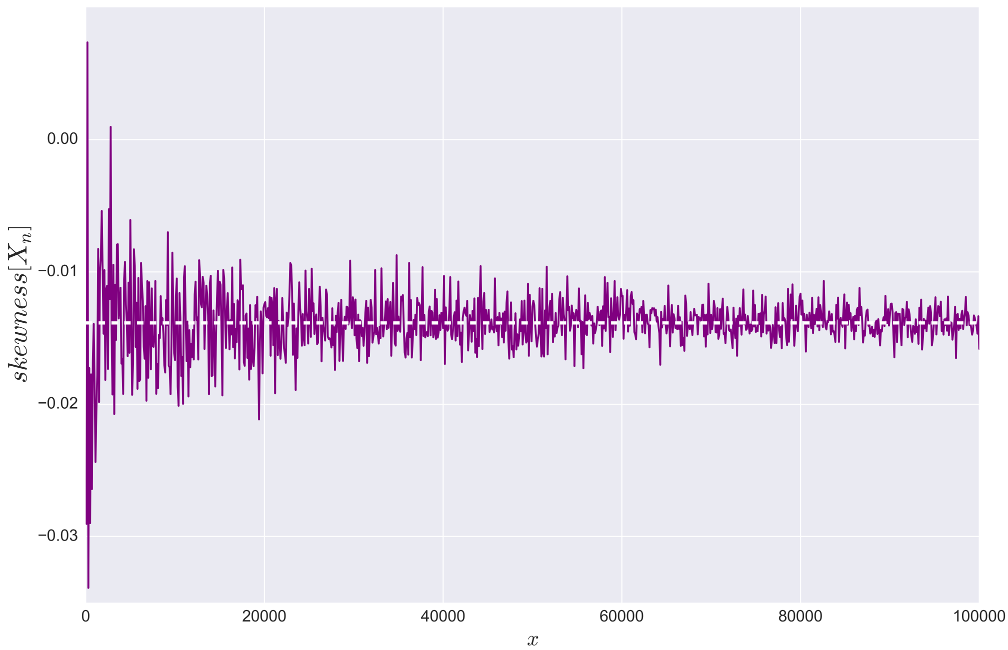


Figure 3.21: The skewness of $\mathcal{SGN}_I\left(\mu=0,\alpha^2=25,\beta=3,\lambda=-2\right)$ distribution for varying sample size n.

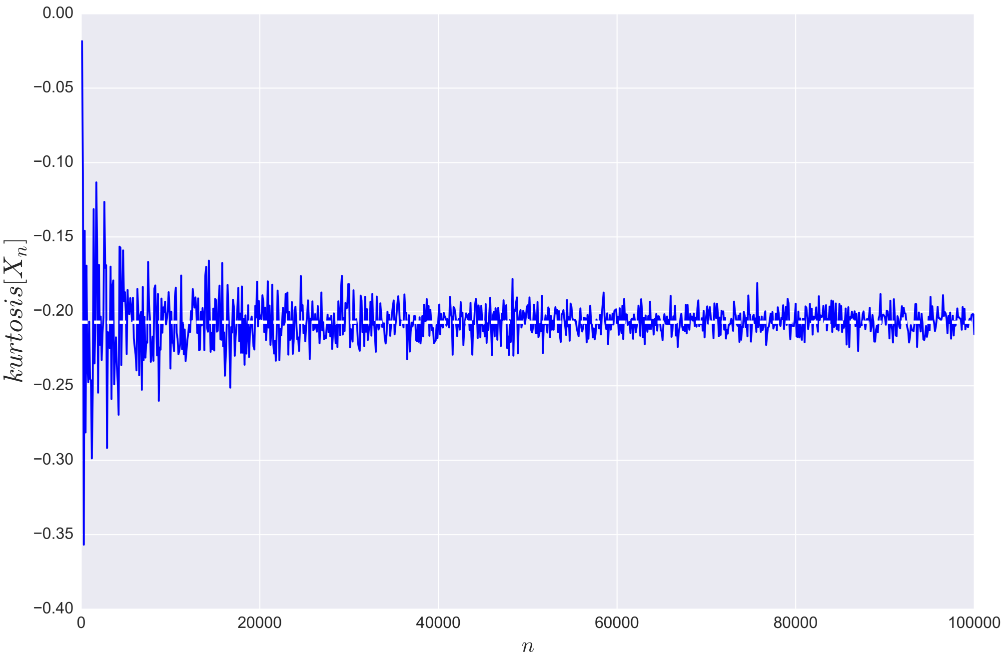


Figure 3.22: The kurtosis of $\mathcal{SGN}_I\left(\mu = 0, \alpha^2 = 25, \beta = 3, \lambda = -2\right)$ distribution for varying sample size n.

As can be seen in the figures, the convergence of the characteristics of the as calculated using Method 2 in Section 3.5.2 are stable and converge to a mean level as indicated by the dotted line. These graphs give an indication of how many random numbers are required to be generated in the corresponding stochastic representations to have confidence in your estimates. For example, in all the figures, there is not a significant improvement in convergence (the fluctuations around the mean level of the characteristic do not have a decreasing variance) for $n > 60\,000$. Therefore, it would be sufficient to have taken $n = 60\,000$ as the number of random numbers required to be generated in the corresponding stochastic representations to be confident in the estimates of the characteristics.

3.12 Summary

In this chapter, the $\mathcal{GN}$ distribution 3.1 is defined and the skewing methodology, as laid out in Proposition 1, Section 1.2, is applied to this distribution. This results in a skew-symmetric version of the $\mathcal{GN}$ distribution. The effect of the parameters β and λ, on the characteristics of $\mathcal{SGN}_I$ distribution is investigated. An acceptance-rejection (AR) method, which samples directly from the $\mathcal{SGN}_I$ distribution, is employed to approximate the characteristics of the $\mathcal{SGN}_I$ distribution. Thereafter, two methods, which do not sample directly from the $\mathcal{SGN}_I$ distribution, are developed to approximate the characteristics of the $\mathcal{SGN}_I$ distribution and are compared. The drawback of the AR algorithm is highlighted and an alternative stochastic sampling scheme is employed to sample directly from the $\mathcal{SGN}_I$ distribution.

Chapter 4

Generalising the extensions of the skew-normal distribution

A new skew generalised-normal type II distribution is presented in Section 4.1. In Section 4.2, the generalised Balakrishnan skew normal type I distribution is generalised by replacing the normal PDF used in these equations with the generalised normal distribution as defined in Section 3.1. In Section 4.3, the beta skew-normal distribution is generalised by using the PDF and CDF of a $\mathcal{SGN}_I$ distribution in the definition of a beta generated distribution [22].

4.1 Skew generalised-normal type II distribution

In this section, using the same notation defined in Section 1.2, Proposition 1, the case where $f_0 = \phi^*$, $G_0 = \Phi^*$, where $\phi^*(x;\beta)$ represents the PDF defined in (3.1), $\Phi^*(x;\beta)$ represents the standard generalised normal CDF defined in (3.3), and where $w(x) = \lambda x$ for $\lambda \in \mathbb{R}$ is investigated. The use of the generalised normal distribution with PDF $\phi^*(x;\beta)$ as defined in (3.1) as the symmetric base PDF is illustrated in the following structure. Unlike Figure 2.6, the skewing mechanism here is function of the CDF of the generalised normal distribution.

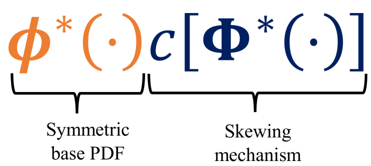


Figure 4.1: The symmetric base PDF and skewing mechanism

The following definition is now obtained:

Definition 9. A random variable X has the skew generalised-normal type II distribution if its

PDF is given by

$$f_X(x;\beta,\lambda) = 2\phi^*(x;\beta)\,\Phi^*(\lambda x;\beta)\,,\ x\in\mathbb{R} \tag{4.1}$$

where $\beta\in\mathbb{R}^+$ and $\lambda\in\mathbb{R}$. This is denoted by $X\sim\mathcal{SGN}_{II}(\beta,\lambda)$.

Remark. This normalisation constant of 2 in Definition 9 follows from Lemma 2.

Corollary 11. *A random variable Y has the $\mathcal{SGN}_{II}$ distribution with location $\mu\in\mathbb{R}$ and scale $\alpha\in\mathbb{R}^+$ if its PDF is given by*

$$f_Y(y;\mu,\alpha,\beta,\lambda) = \tfrac{2}{\alpha}\phi^*\left(\tfrac{y-\mu}{\alpha};\beta\right)\Phi^*\left(\lambda\left(\tfrac{y-\mu}{\alpha}\right);\beta\right)\quad,\ y\in\mathbb{R} \tag{4.2}$$

where $\beta\in\mathbb{R}^+$ and $\lambda\in\mathbb{R}$. This is denoted by $Y\sim\mathcal{SGN}_{II}\left(\mu,\alpha^2,\beta,\lambda\right)$.

Proof. Let $X\sim\mathcal{SGN}_{II}(\beta,\lambda)$ with PDF (4.1). Consider the random variable $Y=\mu+\sigma X$, where the location and scale parameters are denoted $\mu\in\mathbb{R}$ and $\alpha\in\mathbb{R}^+$ respectively.
If $y=\mu+\alpha x$ then $x=u^{-1}(y)=\frac{y-\mu}{\alpha}$. Then $\frac{d}{dy}u^{-1}(y)=\frac{1}{\alpha}$, and it follows that

$$\begin{aligned}
f_Y(y;\mu,\alpha,\beta,\lambda) &= f_X\left(u^{-1}(y);\lambda\right)\left|\frac{d}{dy}u^{-1}(y)\right| \\
&= 2\phi^*\left(u^{-1}(y);\beta\right)\Phi^*\left(\lambda u^{-1}(y);\beta\right)\left|\frac{d}{dy}\left(u^{-1}(y)\right)\right| \\
&= 2\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi^*\left(\lambda\left(\frac{y-\mu}{\alpha}\right);\beta\right)\left|\frac{1}{\alpha}\right| \\
&= \frac{2}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi^*\left(\lambda\left(\frac{y-\mu}{\alpha}\right);\beta\right).
\end{aligned}$$

□

Corollary 12. *When $\mu=0$, $\alpha=\sqrt{2}$ and $\beta=2$ the $\mathcal{SGN}_{II}$ distribution simplifies to the $\mathcal{SN}$ distribution.*

Proof. Using (4.2)

$$\begin{aligned}
f_Y\left(y;0,\sqrt{2},2,\lambda\right) &= \frac{2}{\sqrt{2}}\frac{2}{2\Gamma\left(\frac{1}{2}\right)}e^{-\left|\frac{y-0}{\sqrt{2}}\right|^2}\int_{-\infty}^{\frac{\lambda y}{\sqrt{2}}}\frac{2}{2\Gamma\left(\frac{1}{2}\right)}e^{-|t|^2}dt \\
&= \frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\int_{-\infty}^{\frac{\lambda y}{\sqrt{2}}}\frac{1}{\sqrt{\pi}}e^{-t^2}dt \\
&= 2\phi(y)\int_{-\infty}^{\frac{\lambda y}{\sqrt{2}}}\frac{1}{\sqrt{2\pi}}e^{-t^2}\sqrt{2}dt \\
&= 2\phi(y)\int_{-\infty}^{\frac{\lambda y}{\sqrt{2}}}\frac{1}{\sqrt{2\pi}}e^{-\left(\frac{\sqrt{2}t}{\sqrt{2}}\right)^2}\sqrt{2}dt.
\end{aligned} \tag{4.3}$$

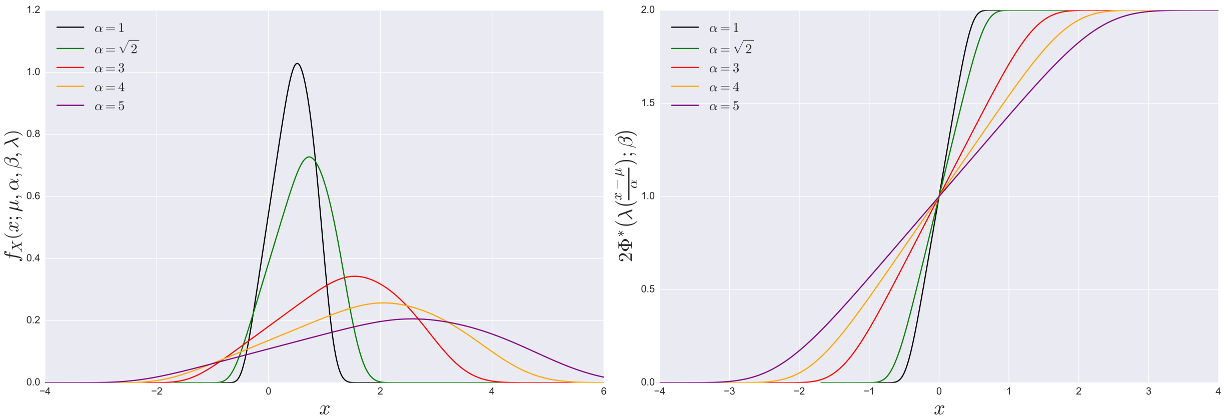
Let $w=\sqrt{2}t$ with $\frac{dw}{dt}=\sqrt{2}$. Then, the upper limit of the integral becomes $\sqrt{2}\frac{\lambda y}{\sqrt{2}}=\lambda y$ and it follows from (4.3) that

$$\begin{aligned}
f_Y(y;\mu,\alpha,\lambda) &= 2\phi(y)\int_{-\infty}^{\lambda y}\frac{1}{\sqrt{2\pi}}e^{-\left(\frac{w}{\sqrt{2}}\right)^2}dw \\
&= 2\phi(y)\,\Phi(\lambda y)
\end{aligned}$$

which is the PDF of the $\mathcal{SN}$ distribution with PDF (2.1). □

4.1.1 $\mathcal{SGN}_{II}$ PDF

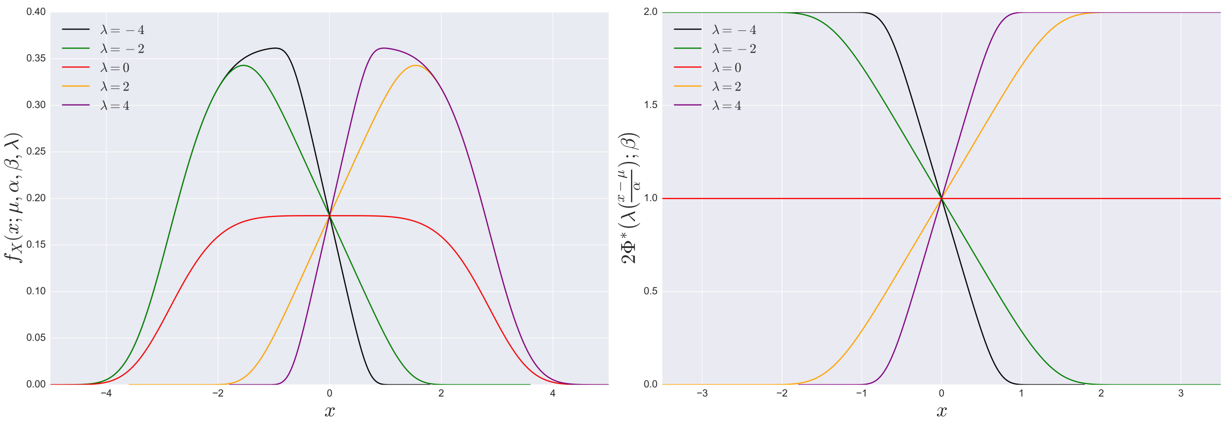
Figure 4.2 and Figure 4.3 depict the PDF of the $\mathcal{SGN}_{II}\left(\mu, \alpha^2, \beta, \lambda\right)$ distribution as given in (4.2), for varying parameter values.



(a) $\mathcal{SGN}_{II}$ PDF

(b) $\mathcal{SGN}_{II}$ skewing mechanism

Figure 4.2: The $\mathcal{SGN}_{II}$ PDF (4.2) and skewing mechanism, $2\Phi^*\left(\lambda\left(\frac{x-\mu}{\alpha}\right);\beta\right)$, for varying α and arbitrary $\mu = 0$, $\beta = 5$ and $\lambda = 2$.



(a) $\mathcal{SGN}_{II}$ PDF

(b) $\mathcal{SGN}_{II}$ skewing mechanism

Figure 4.3: The $\mathcal{SGN}_{II}$ PDF (4.2) and skewing mechanism, $2\Phi^*\left(\lambda\left(\frac{x-\mu}{\alpha}\right);\beta\right)$, for varying λ and arbitrary $\mu = 0$, $\alpha = 3$ and $\beta = 5$.

Remarks

1. Comparing Figure 3.5b and Figure 4.2b it is observed that β now has an effect on the skewing mechanism of the distribution. For increasing β, the skewing window (see Figure 2.2) becomes narrower. Comparing Figure 3.5a and Figure 4.2a it is observed that for

increasing β, the resulting skewness is obtained by multiplying the original symmetric $\mathcal{GN}$ PDF (3.2) by a value in the interval $(0,2)$ over a narrower range of x resulting in a skew-symmetric PDF with its peak attaining a higher probability;

2. Comparing Figure 3.8b and Figure 4.3b it is observed that β and λ jointly have an effect on the skewing mechanism of the distribution. In particular, the same β leads to slightly narrower skewing window (see Figure 2.2). Comparing Figure 3.5a and Figure 4.2a it is observed that the effect of this is that the $\mathcal{SGN}_{II}$ PDF (4.2) has peaks attaining a higher probability than that of the $\mathcal{SGN}_{I}$ PDF (3.11) for the same parameter structure.

4.1.2 Stochastic representation of $\mathcal{SGN}_{II}$ distribution

Theorem 12. *Let $U \sim \mathcal{GN}(\beta)$ and $U_1 \sim \mathcal{GN}(\beta)$ be independent. If*

$$X = U \quad \textit{whenever} \quad U_1 \leq \lambda U$$

then, $X \sim \mathcal{SGN}_{II}(\beta, \lambda)$.

Proof. If $X = U$ whenever $U_1 \leq \lambda U$ then

$$\begin{aligned} \mathbb{P}[X \leq x] &= \mathbb{P}[U \leq x | \{U_1 \leq \lambda U\}] \\ &= \frac{\mathbb{P}[U \leq x, U_1 \leq \lambda U]}{\mathbb{P}[U_1 \leq \lambda U]}. \end{aligned} \tag{4.4}$$

Now

$$\begin{aligned} \mathbb{P}[U \leq x, U_1 \leq \lambda U] &= \int_{u=-\infty}^{x} \int_{u_1=-\infty}^{\lambda u} \phi^*(u;\beta)\,\phi^*(u_1;\beta)\,du_1 du \\ &= \int_{-\infty}^{x} \phi^*(u;\beta) \left(\int_{u_1=-\infty}^{\lambda u} \phi^*(u_1;\beta)\,du_1 \right) du \\ &= \int_{-\infty}^{x} \phi^*(u;\beta)\,\Phi^*(\lambda u;\beta)\,du. \end{aligned} \tag{4.5}$$

Since U and U_1 are both symmetric about zero, applying Lemma 2 it follows that

$$\mathbb{P}[U_1 \leq \lambda U] = \frac{1}{2}. \tag{4.6}$$

Then using (4.5) and (4.6) in (4.4) it follows that

$$\begin{aligned} \mathbb{P}[X \leq x] &= \frac{\int_{-\infty}^{x} \phi^*(u;\beta)\,\Phi^*(\lambda u;\beta)\,du}{0.5} \\ &= \int_{-\infty}^{x} 2\phi^*(u;\beta)\,\Phi^*(\lambda u;\beta)\,du. \end{aligned}$$

Applying a standard statistical result (Result 13, Appendix B.2) it follows that the PDF is

$$\begin{aligned} f_X(x;\beta,\lambda) &= \frac{d}{du}\int_{-\infty}^{x} 2\phi^*(u;\beta)\,\Phi(\lambda\mu)\,du \\ &= 2\phi^*(u;\beta)\,\Phi^*(\lambda\mu;\beta)\big|_{-\infty}^{x} \\ &= 2\phi^*(x;\beta)\,\Phi^*(\lambda x;\beta) - \lim_{k\to-\infty}\left(2\phi^*(k;\beta)\,\Phi^*(\lambda k;\beta)\right) \\ &= 2\phi^*(x;\beta)\,\Phi^*(\lambda x;\beta) - 0 \\ &= 2\phi^*(x;\beta)\,\Phi^*(\lambda x;\beta) \end{aligned}$$

which is the PDF $f_X(x;\beta,\lambda)$ as given in (4.1).

□

Corollary 13. *Let* $U \sim \mathcal{GN}(\beta)$ *and* $U_1 \sim \mathcal{GN}(\beta)$ *be independent. If*

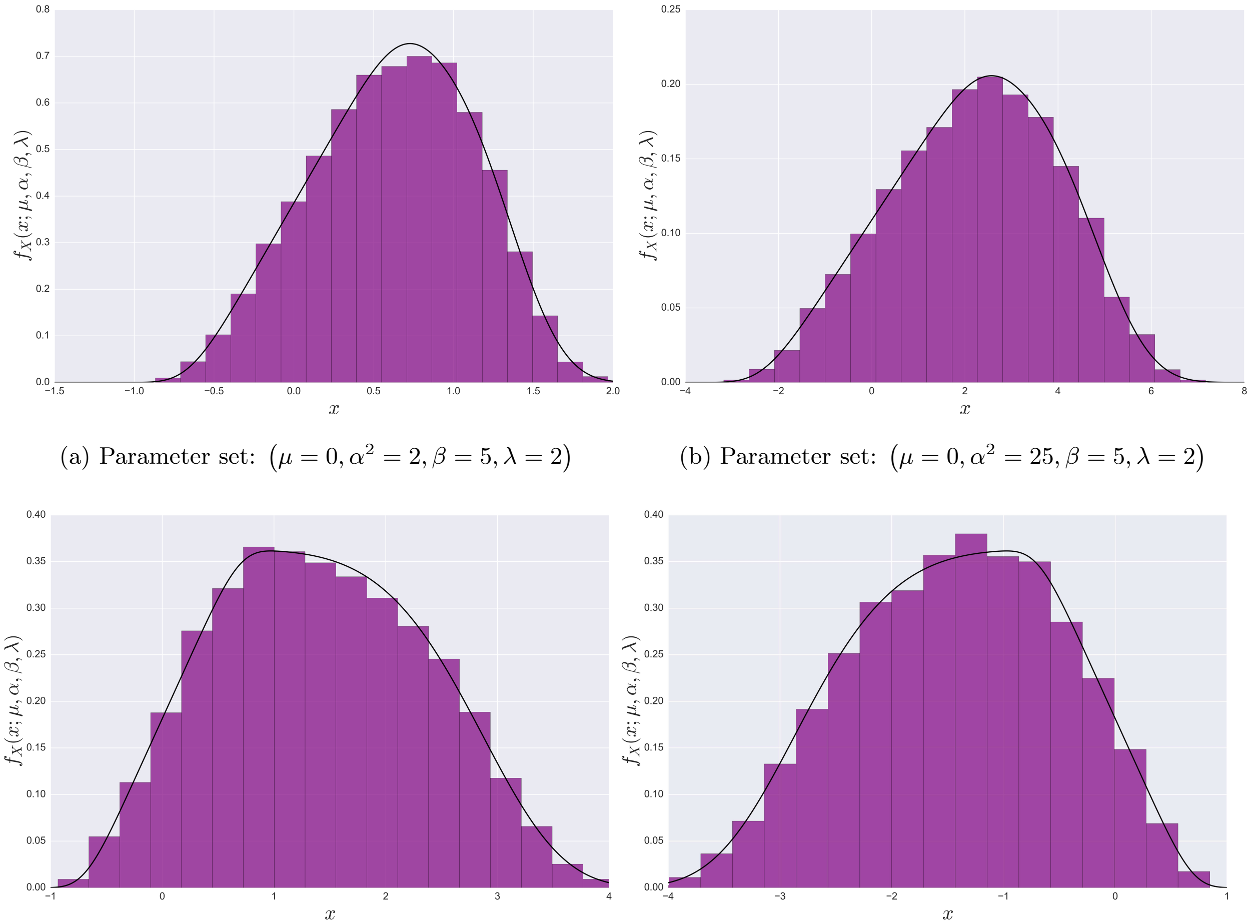
$$Y = \mu + \alpha U \quad \textit{whenever} \quad U_1 \leq \lambda U$$

then $X \sim \mathcal{SGN}_{II}\left(\mu,\alpha^2,\beta,\lambda\right)$.

Proof. The proof is similar to that of Theorem 12. □

4.1.3 Visualisation of $\mathcal{SGN}_{II}$ sampling scheme derived in Section 4.1.2

Figure 4.4 displays the histograms of realised random samples of size approximately 10 000 taken from $X \sim \mathcal{SGN}_{II}\left(\mu,\alpha^2,\beta,\lambda\right)$ with the corresponding theoretical PDF (4.2), overlaid for different values of μ,α^2,β and λ.



(a) Parameter set: $\left(\mu = 0, \alpha^2 = 2, \beta = 5, \lambda = 2\right)$

(b) Parameter set: $\left(\mu = 0, \alpha^2 = 25, \beta = 5, \lambda = 2\right)$

(c) Parameter set: $\left(\mu = 0, \alpha^2 = 9, \beta = 5, \lambda = 4\right)$

(d) Parameter set: $\left(\mu = 0, \alpha^2 = 9, \beta = 3, \lambda = -4\right)$

Figure 4.4: Histograms of realised random samples of size 10 000 taken from $X \sim \mathcal{SGN}_{II}\left(\mu, \alpha^2, \beta, \lambda\right)$ with the corresponding theoretical PDF (4.2), overlaid for different values of μ, α^2, β and λ.

Figure 4.4 shows histograms of the random samples taken from $X \sim \mathcal{SGN}_{II}\left(\mu, \alpha^2, \beta, \lambda\right)$ using the stochastic representation in Corollary 13 with the corresponding theoretical PDF (4.2) overlaid.

4.2 Balakrishnan skew generalised-normal

Consider the generalised Balakrishnan skew-normal distribution as in (2.14). The generalised normal distribution with PDF $\phi^*(x;\beta)$ as defined in (3.1) is now used as the symmetric base PDF. The skewing mechanism remains a function of the CDF of the normal distribution, but with adjustment as indicated in Figure 1.4. The following definition is then obtained:

Definition 10. A random variable X has the generalised Balakrishnan skew generalised-normal distribution if its PDF is given by

$$f_X\left(x;n,\beta,\lambda_1,\lambda_2\right) \quad = \quad c_n\left(\beta,\lambda_1,\lambda_2\right)\phi^*\left(x;\beta\right)\Phi^n\left(\frac{\lambda_1 x}{\sqrt{\frac{1}{2}+\lambda_2 x^2}}\right),\ x\in\mathbb{R} \tag{4.7}$$

where $n\in\mathbb{Z}^+$, $\beta\in\mathbb{R}^+$, $\lambda_1\in\mathbb{R}$, $\lambda_2\in\mathbb{R}^+$ and

$$\begin{aligned} c_n\left(\beta,\lambda_1,\lambda_2\right) &= \frac{1}{\int_{\mathbb{R}}\phi^*\left(x;\beta\right)\Phi^n\left(\frac{\lambda_1 x}{\sqrt{\frac{1}{2}+\lambda_2 x^2}}\right)dx} \\ &= \frac{1}{\mathbb{E}_{B_5}\left[\Phi^n\left(\frac{\lambda_1 B_5}{\sqrt{\frac{1}{2}+\lambda_2 B_5^2}}\right)\right]} \end{aligned}$$

where $B_5\sim\mathcal{GN}\left(\beta\right)$ with PDF (3.1). This is denoted by $X\sim\mathcal{GBSN}_1^*\left(n,\beta,\lambda_1,\lambda_2\right)$.

Corollary 14. *A random variable Y has the generalised Balakrishnan skew generalised-normal distribution with location parameter $\mu\in\mathbb{R}$ and scale parameter $\alpha\in\mathbb{R}^+$ if its PDF is given by*

$$f_Y\left(y;\mu,\alpha,n,\beta,\lambda_1,\lambda_2\right) \quad = \quad \frac{c_n(\mu,\alpha,\beta,\lambda_1,\lambda_2)}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(y-\mu)^2}}\right),\ y\in\mathbb{R} \tag{4.8}$$

where $n\in\mathbb{Z}^+$, $\beta\in\mathbb{R}^+$, $\lambda_1\in\mathbb{R}$, $\lambda_2\in\mathbb{R}^+$ and

$$\begin{aligned} c_n\left(\mu,\alpha,\beta,\lambda_1,\lambda_2\right) &= \frac{1}{\int_{\mathbb{R}}\frac{1}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(y-\mu)^2}}\right)dy} \\ &= \frac{1}{\mathbb{E}_{B_6}\left[\frac{\lambda_1(B_6-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(B_6-\mu)^2}}\right]} \end{aligned}$$

where $B_6\sim\mathcal{GN}\left(\mu,\alpha^2,\beta\right)$ with PDF (3.2).
This is denoted by $Y\sim\mathcal{GBSN}_1^*\left(\mu,\alpha^2,n,\beta,\lambda_1,\lambda_2\right)$.

Proof. Let $X\sim\mathcal{GBSN}_1^*\left(n,\beta,\lambda_1,\lambda_2\right)$ with PDF (4.7). Consider the random variable $Y=\mu+\alpha X$, where the location and scale parameters are denoted $\mu\in\mathbb{R}$ and $\alpha\in\mathbb{R}^+$ respectively.

If $y = \mu + \alpha x$ then $\frac{d}{dy}u^{-1}(y) = \frac{d}{dy}\left(\frac{y-\mu}{\alpha}\right) = \frac{1}{\alpha}$ giving

$$\begin{aligned}
f_Y(y;\mu,\alpha,n,\beta,\lambda_1,\lambda_2) &= f_X\left(u^{-1}(y);\beta,\lambda_1,\lambda_2\right)\left|\frac{d}{dy}u^{-1}(y)\right| \\
&= \frac{1}{\int_{\mathbb{R}}\phi^*\left(u^{-1}(y);\beta\right)\Phi^n\left(\frac{\lambda_1 u^{-1}(y)}{\sqrt{\frac{1}{2}+\lambda_2\left(u^{-1}(y)\right)^2}}\right)\left|\frac{d}{dy}\left(u^{-1}(y)\right)\right|dy}\phi^*\left(u^{-1}(y);\beta\right) \\
&\times \Phi^n\left(\frac{\lambda_1 u^{-1}(y)}{\sqrt{\frac{1}{2}+\lambda_2\left(u^{-1}(y)\right)^2}}\right)\left|\frac{d}{dy}\left(u^{-1}(y)\right)\right| \\
&= \frac{\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)}{\int_{\mathbb{R}}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi^n\left(\frac{\lambda_1\left(\frac{y-\mu}{\alpha}\right)}{\sqrt{\frac{1}{2}+\lambda_2\left(\frac{y-\mu}{\alpha}\right)^2}}\right)\left|\frac{1}{\alpha}\right|dy}\Phi^n\left(\frac{\lambda_1\left(\frac{y-\mu}{\alpha}\right)}{\sqrt{\frac{1}{2}+\lambda_2\left(\frac{y-\mu}{\alpha}\right)^2}}\right)\left|\frac{1}{\alpha}\right| \\
&= \frac{\frac{1}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)}{\int_{\mathbb{R}}\frac{1}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(y-\mu)^2}}\right)dy}\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(y-\mu)^2}}\right) \\
&= \frac{c_n(\mu,\alpha,\beta,\lambda_1,\lambda_2)}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(y-\mu)^2}}\right)
\end{aligned}$$

where

$$\begin{aligned}
c_n(\mu,\alpha,\beta,\lambda_1,\lambda_2) &= \frac{1}{\int_{\mathbb{R}}\frac{1}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi^n\left(\frac{\lambda_1(y-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(y-\mu)^2}}\right)dy} \\
&= \frac{1}{\mathbb{E}_{B_6}\left[\Phi^n\left(\frac{\lambda_1(B_6-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(B_6-\mu)}}\right)\right]}
\end{aligned}$$

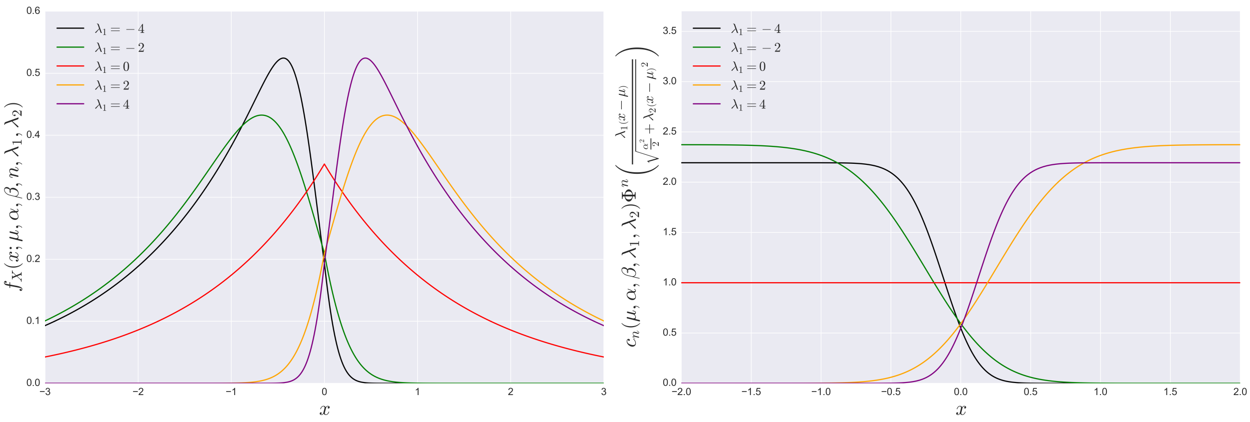
□

which gives the required result.

Corollary 15. *If $\mu = 0$, $\alpha = \sqrt{2}$ and $\beta = 2$ in (4.8) the $\mathcal{GBSN}_1^*(\mu,\alpha,n,\beta,\lambda_1,\lambda_2)$ distribution reduced to the $\mathcal{GBSN}_1$ distribution with PDF (2.14).*

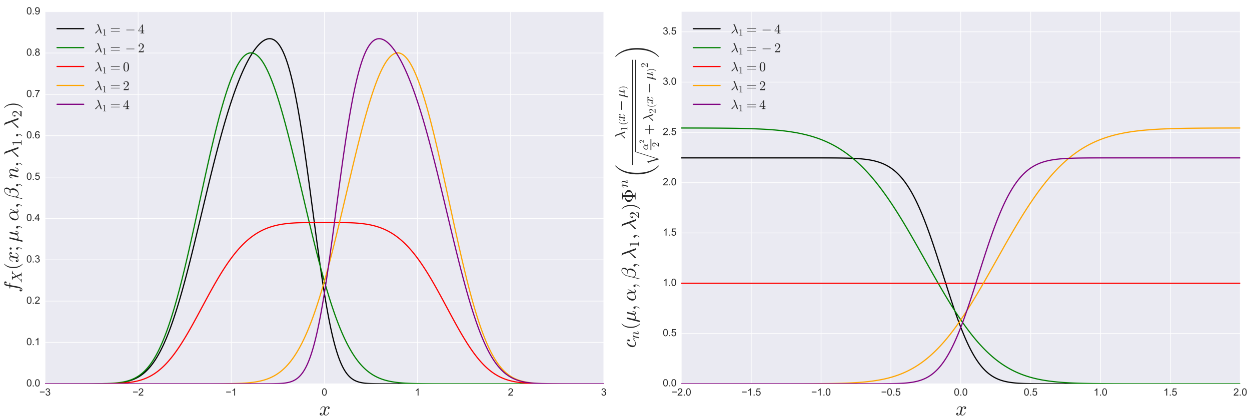
4.2.1 $\mathcal{GBSN}_1^*$ PDF

Figures 4.5 - 4.8 depict the PDF of the $\mathcal{GBSN}_1^*(\mu,\alpha,n,\beta,\lambda_1,\lambda_2)$ as given in (4.8) and associated skewing mechanism, for varying parameter values.



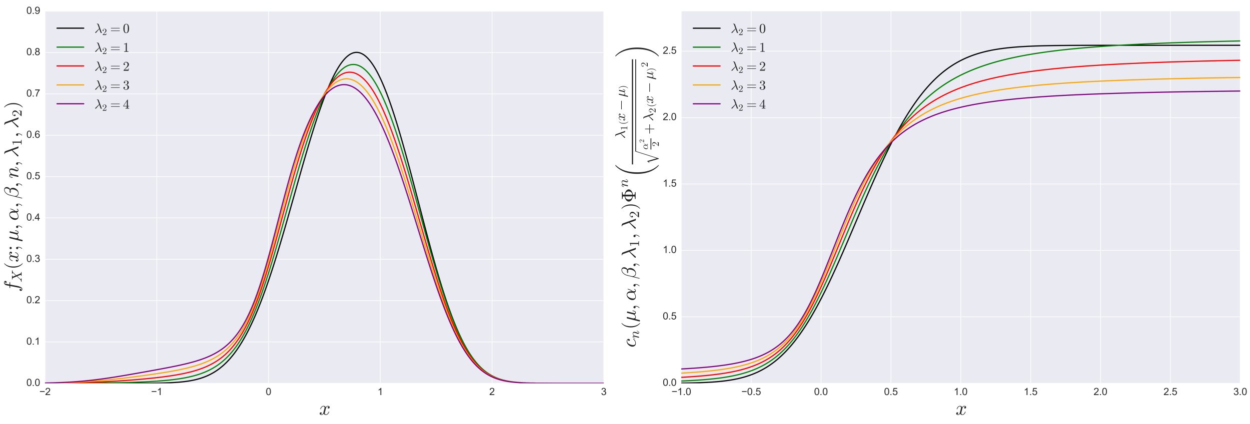
(a) $\mathcal{GBSN}_1^*$PDF (b) $\mathcal{GBSN}_1^*$ skewing mechanism

Figure 4.5: The $\mathcal{GBSN}_1^*$ PDF 4.8 and skewing mechanism, $c_n\left(\mu,\alpha,\beta,\lambda_1,\lambda_2\right)\Phi^n\left(\frac{\lambda_1(x-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(x-\mu)^2}}\right)$, for varying λ_1 and arbitrary $\mu = 0$, $\alpha = \sqrt{2}$, $n = 2$, $\beta = 1$ and $\lambda_2 = 0$.



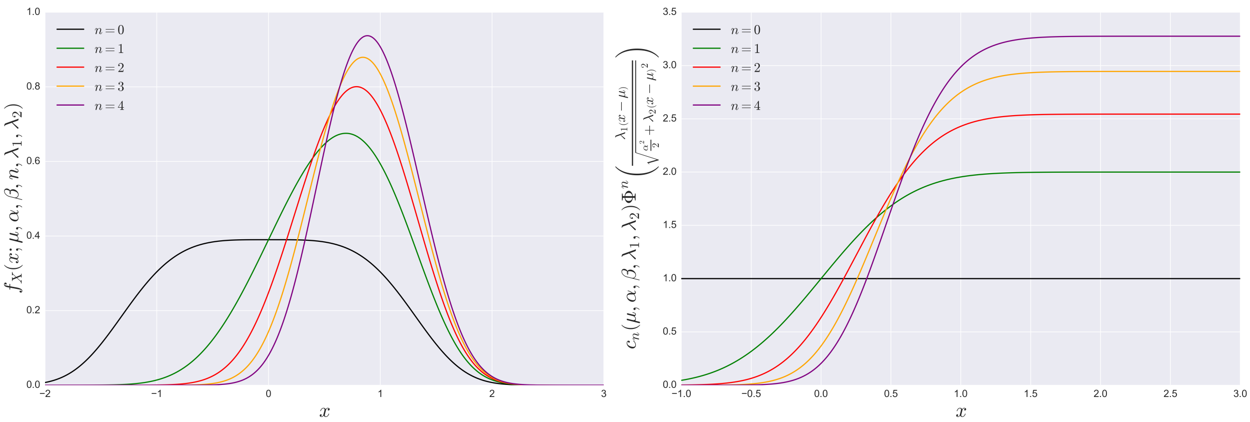
(a) $\mathcal{GBSN}_1^*$PDF (b) $\mathcal{GBSN}_1^*$ skewing mechanism

Figure 4.6: The $\mathcal{GBSN}_1^*$ PDF 4.8 and skewing mechanism, $c_n\left(\mu,\alpha,\beta,\lambda_1,\lambda_2\right)\Phi^n\left(\frac{\lambda_1(x-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(x-\mu)^2}}\right)$, for varying λ_1 and arbitrary $\mu = 0$, $\alpha = \sqrt{2}$, $n = 2$, $\beta = 4$ and $\lambda_2 = 0$.



(a) $\mathcal{GBSN}_1^*$ PDF (b) $\mathcal{GBSN}_1^*$ skewing mechanism

Figure 4.7: The $\mathcal{GBSN}_1^*$ PDF 4.8 and skewing mechanism, $c_n(\mu,\alpha,\beta,\lambda_1,\lambda_2)\,\Phi^n\left(\frac{\lambda_1(x-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(x-\mu)^2}}\right)$, for varying λ_2 and arbitrary $\mu = 0$, $\alpha = \sqrt{2}$, $n = 2$, $\beta = 4$ and $\lambda_1 = 2$.



(a) $\mathcal{GBSN}_1^*$ PDF (b) $\mathcal{GBSN}_1^*$skewing mechanism

Figure 4.8: The $\mathcal{GBSN}_1^*$ PDF 4.8 and skewing mechanism, $c_n(\mu,\alpha,\beta,\lambda_1,\lambda_2)\,\Phi^n\left(\frac{\lambda_1(x-\mu)}{\sqrt{\frac{\alpha^2}{2}+\lambda_2(x-\mu)^2}}\right)$, for varying n and arbitrary $\mu = 0$, $\alpha = \sqrt{2}$, $\beta = 4$, $\lambda_1 = 2$ and $\lambda_2 = 0$.

Remarks

1. Since we have fixed $\alpha = \sqrt{2}$, the skewing mechanism corresponding to the $\mathcal{GBSN}_1^*$ distribution with PDF 4.8 is identical to that of the skewing mechanism corresponding to the $\mathcal{GBSN}_1$ distribution with PDF (2.17) in Section 2.5.3. The comments with respect to the skewing mechanism are thus identical to those outlined in Section 2.5.3.

2. Comparing Figures 4.5a and 4.6a it is clear that by increasing β, the PDF of the $\mathcal{GBSN}_1^*$ distribution as given in (4.8) exhibits shorter/lighter tails.

4.3 The beta skew generalised-normal

One mechanism to generate flexible distributions is to use the kernel of the beta distribution as the generator distribution and compound with the CDF of another distribution as in Mameli [22]. Let $F(\cdot;\theta)$ be a CDF indexed with parameter $\theta \in \Theta$ and corresponding PDF $f(\cdot;\theta)$, termed the baseline distribution. The CDF of the beta generated distribution can be constructed using

$$G(x;\theta,a,b) = \int_0^{F(x;\theta)} u^{a-1}(1-u)^{b-1}\,du.$$

Applying the Leibniz integral rule [26] we get

$$g(x;\theta a,c) = \frac{1}{B(a,b)} f(x;\theta)\,F(x;\theta)^{a-1}(1-F(x;\theta))^{b-1}$$

where $B(a,b)$ denotes the complete beta function (see Definition 15, Appendix B.1). It is proposed to set the baseline distribution function as the CDF of $\mathcal{SGN}_I\left(\mu,\alpha^2\beta,\lambda\right)$ distribution.

Definition 11. A random variable X has the beta skew generalised-normal distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\alpha \in \mathbb{R}^+$ if its PDF is given by

$$f_X(x;\mu,\alpha,\beta,\lambda,a,b) = \tfrac{1}{B(a,b)} F(x;\mu,\alpha,\beta,\lambda)^{a-1}(1-F(x;\mu,\alpha,\beta,\lambda))^{b-1} f(x;\mu,\alpha,\beta,\lambda)\ ,x\in\mathbb{R} \tag{4.9}$$

where $F(x;\mu,\alpha,\beta,\lambda)$ refers to the CDF of the $\mathcal{SGN}_I\left(\mu,\alpha^2,\beta,\lambda\right)$ distribution with PDF (3.11), $\beta \in \mathbb{R}^+$, $\lambda \in \mathbb{R}^+$ and $a,b \geq 1$. This is denoted by $Y \sim Beta\mathcal{SGN}\left(\mu,\alpha^2,\beta,\lambda,a,b\right)$.

Remark. The restriction on the parameters a and b in Definition 11 i.e. $a,b \geq 1$ ensures than the $Beta\mathcal{SGN}$ PDF as in (4.9) is unimodal.

Corollary 16. *Let $\mu = 0$, $\alpha = \sqrt{2}$, $\beta = 2$ in the $Beta\mathcal{SGN}$ PDF (4.9). Then the $Beta\mathcal{SGN}$ distribution reduced to the standard $Beta\mathcal{SN}$ distribution with PDF (2.21)*

Proof. It follows from (4.9) that

$$\begin{aligned} f_X(x;\mu,\alpha,\beta,\lambda,a,b) &= \frac{1}{B(a,b)} F\left(x;0,\sqrt{2},2,\lambda\right)^{a-1}\left(1-F\left(x;0,\sqrt{2},2,\lambda\right)\right)^{b-1} f\left(x;0,\sqrt{2},2,\lambda\right) \\ &= \frac{1}{B(a,b)}\left(\int_{-\infty}^{x} f\left(x;0,\sqrt{2},2,\lambda\right)dx\right)^{a-1}\left(1-\left(\int_{-\infty}^{x} f\left(x;0,\sqrt{2},2,\lambda\right)dx\right)\right)^{b-1} \\ &\quad \times f\left(x;0,\sqrt{2},2,\lambda\right) \end{aligned} \tag{4.10}$$

Applying Corollary 9 if follows from (4.10) that

$$
\begin{aligned}
f_X(x;\mu,\alpha,\beta,\lambda,a,b) &= \frac{1}{B(a,b)} F\left(x;0,\sqrt{2},2,\lambda\right)^{a-1}\left(1-F\left(x;0,\sqrt{2},2,\lambda\right)\right)^{b-1} 2\phi(x)\,\Phi(\lambda x) \\
&= \frac{1}{B(a,b)}\left(\int_{-\infty}^{x} 2\phi(x)\,\Phi(\lambda x)\,dx\right)^{a-1} \\
&\times \left(1-\left(\int_{-\infty}^{x} 2\phi(x)\,\Phi(\lambda x)\,dx\right)\right)^{b-1} 2\phi(x)\,\Phi(\lambda x) \\
&= \frac{1}{B(a,b)} F(x;\lambda)^{a-1}\left(1-F(x;\lambda)\right)^{b-1} f(x;\lambda)
\end{aligned}
$$

where $F(x;\lambda)$ is the CDF of $\mathcal{SN}$ distribution with PDF $f(x;\lambda)$ (2.1).

This is the PDF of the standard $Beta\mathcal{SN}$ with PDF (2.21) and the proof is complete. □

4.3.1 $Beta\mathcal{SGN}$ PDF

Figure 4.9 and Figure 4.10 depict the PDF of the $Beta\mathcal{SGN}\left(\mu,\alpha^2,\beta,\lambda,a,b\right)$ distribution i.e. $f_X(x;\mu,\alpha,\lambda,a,b)$ as given in (4.9), for varying parameter values.

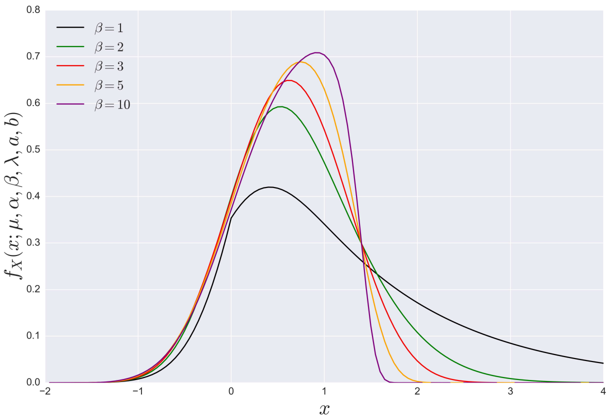


Figure 4.9: The $Beta\mathcal{SGN}$ PDF (4.9) for varying β and arbitrary $\mu = 0$, $\alpha^2 = \sqrt{2}$, $\lambda = 2$, $a = 1$ and $b = 1$.

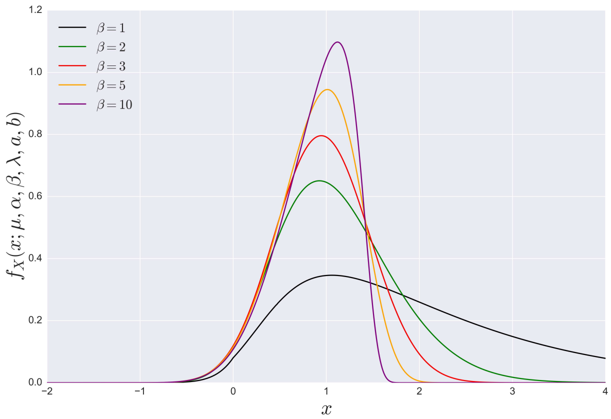


Figure 4.10: The $Beta\mathcal{SGN}$ PDF (4.9) for varying β and arbitrary $\mu = 0$, $\alpha^2 = \sqrt{2}$, $\lambda = 2$, $a = 2$ and $b = 1$.

Remarks

1. It is important to note that the construction of the $Beta\mathcal{SN}$ and $Beta\mathcal{SGN}$ distributions only allow a symmetric PDF if $a = b$ and $\lambda = 0$.

2. In Figure 4.9 and Figure 4.10 it is noted that for $\beta > 2$ the $Beta\mathcal{SGN}$ PDF (4.9) has peaks attaining a higher probability than those of the $Beta\mathcal{SN}$ PDF as in (2.20).

3. In Figure 4.9 and Figure 4.10 it is noted that for $\beta < 2$ the $Beta\mathcal{SGN}$ PDF (4.9) has peaks attaining a lower probability than those of the $Beta\mathcal{SN}$ PDF as in (2.20).

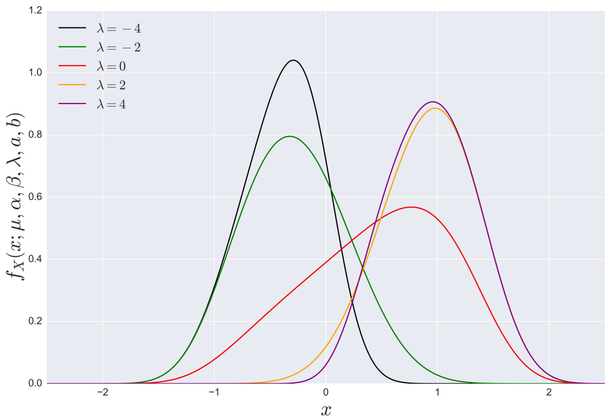


Figure 4.11: The $Beta\mathcal{SGN}$ PDF (see (4.9)) for varying λ and arbitrary $\mu = 0$, $\alpha^2 = \sqrt{2}$, $\beta = 4$, $a = 2$ and $b = 1$.

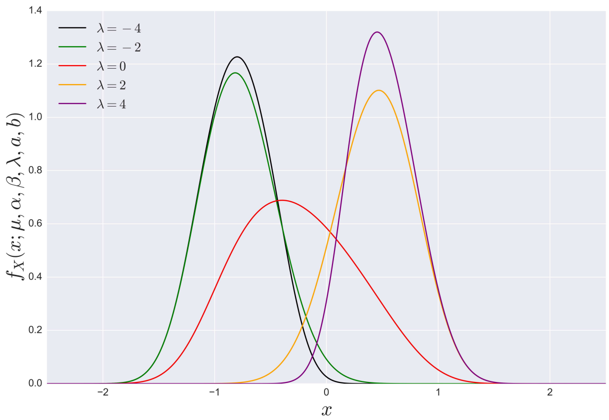


Figure 4.12: The $Beta\mathcal{SGN}$ PDF (see (4.9)) for varying λ and arbitrary $\mu = 0$, $\alpha^2 = \sqrt{2}$, $\beta = 4$, $a = 2$ and $b = 3$.

Remarks

1. It is important to note that the construction of the $Beta\mathcal{SN}$ and $Beta\mathcal{SGN}$ distributions only allow a symmetric PDF if $a = b$ and $\lambda = 0$.

2. Comparing Figure 2.10 and Figure 4.11 it is noted that for $\lambda = 0$, $a > b$ and $\beta > 2$ more of the mass of the $Beta\mathcal{SGN}$ PDF (4.9) lies to the right of $x = 0$ compared to that of the $Beta\mathcal{SN}$ PDF (2.20).

3. Similarly, comparing Figure 2.11 and Figure 4.12 it is noted that when $\lambda = 0$, $a < b$ and $\beta > 2$ more of the mass of the $Beta\mathcal{SGN}$ PDF (4.9) lies to the left of $x = 0$ compared to that of the $Beta\mathcal{SN}$ PDF (2.20).

4.4 Summary

In this chapter, extensions of distributions in Chapter 2 are presented. The skew generalised-normal type II ($\mathcal{SGN}_{II}$) distribution is defined and a stochastic representation is derived. The skewing mechanism associated with $\mathcal{SGN}_{II}$ distribution is compared to that of the $\mathcal{SGN}_I$ distribution. The $\mathcal{GBSN}_1^*$ and $Beta\mathcal{SGN}$ distributions are defined and compared with their counterparts in Chapter 2.

Chapter 5

Application

In this chapter, distribution fitting to real world data along with approximating binomial probabilities are proposed as applications of our new results. The focus is on the $\mathcal{SGN}_I$ distribution only.

5.1 Fitting to data

Here the usefulness of the proposed $\mathcal{SGN}_I$ distribution in fitting to real-world data is illustrated. An Australian athletes data set containing various measurement on athletes specialising in different sports is used [29]. The variable of interest is the caliper measurement obtained from each athlete, which provides an indication of body-fat percentage.

5.1.1 Assessing the suitability of distributions

It is vital to perform a test on whether a distribution is a candidate for fitting to a particular data set. The Kolmogorov-Smirnov $(K - S)$ goodness-of-fit test (see Definition 26, Appendix B.1) is performed to assess the suitability of fitting the $\mathcal{SGN}_I$ distribution to this data set. The standard critical values of the $K - S$ statistic do not apply when any parameters of the candidate distribution ($\mathcal{SGN}_I$ in this case) are estimated from the data. Hence, a Monte Carlo approach must be used instead to construct an appropriate test. We propose the following algorithm to find critical values of the $K - S$ statistic. For our purposes we take the null hypothesis, H_0, that the data are from a $\mathcal{SGN}_I$ distribution.

Algorithm:

1. Fit the $\mathcal{SGN}_I$ distribution to the original data using maximum likelihood estimation;

2. Calculate the $K - S$ distance, d^*, between the data and the fitted $\mathcal{SGN}_I$ distribution;

3. Run a bootstrap re-sampling from the original data 500 times and fit the $\mathcal{SGN}_I$ distribution again;

4. Calculate $K-S$ distance, d_i, between the bootstrapped data and the fitted $\mathcal{SGN}_I$ distribution;

5. Repeat steps $3-4$ M times to obtain the set $d = \{d_1, d_2, \cdots, d_M\}$;

6. Calculate the $(1-\alpha)^{th}$ sample quantiles $q_{1-\alpha}$ of d_M at levels of significance: $\alpha = 0.01,\ 0.05,\ 0.1$;

7. If $d^* < q_{1-\alpha}$ the null hypothesis cannot be rejected at α level of significance and there is not enough evidence to suggest that the data are not from a $\mathcal{SGN}_I$ distribution. If this is the case, the $K-S$ test indicated that the $\mathcal{SGN}_I$ distribution is a suitable candidate to fit to the data;

Using the above algorithm, an approximate $p-value$ can also be obtained as the proportion of time the elements in d_M are greater than d^*.
The test is performed with $M = 500$. The test statistic is calculated as $d^* = 0.1174142$. The critical values at different levels of significance are presented in Table 5.1.

Level of significance, α	Critical value, $q_{1-\alpha}$
0.01	0.138527
0.05	0.145027
0.1	0.1562062

Table 5.1: Simulated critical values of the $K-S$ goodness of fit test at various levels of significance.

Since $d^* < q_{1-\alpha}$ for $\alpha = 0.01, 0.05, 0.1$ and the approximate $p-value$ is calculated as 0.545109, the null hypothesis, that the data are from a $\mathcal{SGN}$ distribution, cannot be rejected. Therefore it is concluded that the $\mathcal{SGN}$ distribution is a suitable candidate to fit to this data set.

5.1.2 Distribution fitting

The following distributions are fitted to the caliper measurement obtained from each athlete:

- the normal distribution, $\mathcal{N}\left(\mu, \sigma^2\right)$;
- $\mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ distribution with PDF $f_X\left(x; \mu, \sigma, \lambda\right)$ as given in (2.2);

- $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ with PDF $f_X\left(x; \mu, \alpha, \beta, \lambda\right)$ as given in (3.11);
- $\mathcal{GBSN}_1\left(\mu, \sigma^2, n, \lambda_1, \lambda_2\right)$ with PDF $f_X\left(x; \mu, \sigma, n, \lambda_1, \lambda_2\right)$ as given in (2.17)

Methodology

1. The method of maximum likelihood was used for the fitting of all distributions;
2. In particular, to fit the $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ distribution, let $x_1, \cdots, x_m$ be a random sample of size m from $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ and maximise the likelihood function given by
$$\mathcal{L}\left(\mu, \alpha^2, \beta, \lambda\right) = \left(\frac{2}{\alpha}\right)^m \prod_{i=1}^m \phi^*\left(\frac{x_i - \mu}{\alpha}; \beta\right) \Phi\left(\sqrt{2}\lambda\left(\frac{x_i - \mu}{\alpha}\right)\right)$$
using well-known optimisation techniques (see Nelder–Mead method) in SAS 9.4.
3. The location and scale parameters of $\mathcal{N}\left(\mu, \sigma^2\right)$ distribution are respectively set as the sample mean and sample standard deviation of the data;
4. The location and scale parameters of the of $\mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ distribution with PDF (2.2) are set as the maximum likelihood estimates of the respective shape and scale parameters of the $\mathcal{N}\left(\mu, \sigma^2\right)$ distribution. A histogram of the data is obtained and it was concluded that the data was positively skewed (see Figure 5.1). Therefore an arbitrary positive value of λ is chosen as an initial value. The fitting of the $\mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ distribution with PDF (2.2) can be repeated using various positive values of λ to test the sensitivity of the initial values on the maximum likelihood estimates.
5. The location and scale parameters of the of $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ distribution are set as the maximum likelihood estimates of the respective location and scale parameters of the $\mathcal{SN}\left(\mu, \sigma^2, \lambda\right)$ distribution. The initial value of the shape parameter β is set to two. Care needs to be taken in setting the initial value of the skewness parameter λ as a positive λ does not necessarily imply that the resulting $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ PDF as in (3.11) will be positively skewed (see Figure 3.9). It is therefore recommended to set the initial value of λ to zero.
6. The adequacy of the fit of the four distributions was assessed by the BIC and AIC information criteria (see Definition 27 and Definition 28, Appendix B.1).

Results

	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\alpha}^2$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{n}$	AIC	BIC
$\mathcal{N}$	22.96	2.86^2	-	-	-	-	-	-	1001.336	997.336
$\mathcal{SN}$	19.97	4.13^2	-	-	2.313	-	-	-	986.199	980.199
$\mathcal{SGN}_I$	20.80	-	3.842^2	1.381	1.074	-	-	-	**984.349**	**976.349**
$\mathcal{GBSN}_I$	19.76	4.288^2	-	-	-	2.535	1.092	1.1613677	989.820	979.820

Table 5.2: AIC and BIC criteria obtained for each of the fitted distributions

As a visual assessment of goodness of fit, the estimated PDFs of the four distributions and the empirical histogram are plotted in Figure 5.1.

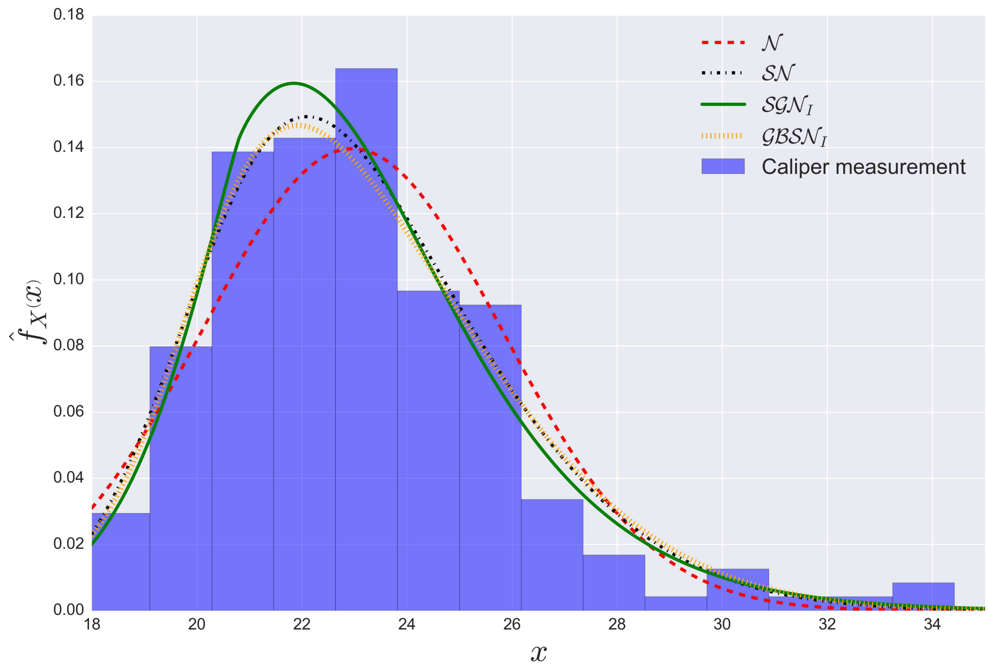


Figure 5.1: Empirical histogram of data with overlaid fitted PDFs

The results in Table 5.2 identify the $\mathcal{SGN}_I\left(\mu, \alpha^2, \beta, \lambda\right)$ distribution with PDF (3.11) as the best fit for the given data.

5.2 Approximating the binomial distribution

Consider a random variable $X \sim Binomial\,(n, p)$ with PDF as given in Definition 17, Appendix B.1. A normal distribution with expected value np and variance $np(1-p)$ is usually used to approximate the binomial distribution when n is large or when p is close to 0.5 (in which case the binomial distribution PDF is approximately symmetric). However, it is well known that the $f_X(x; n, p)$ (Definition 17, Appendix B.1) is not symmetric for $p \neq 0.5$ and exhibits a non-negligible degree of skewness for both large and small values of p. It is therefore of interest to consider approximating the binomial distribution using the $\mathcal{SN}$ (as done by Chang et. al.) [12] and $\mathcal{SGN}_I$ distributions in order to account for the skewness present.

5.2.1 Methodology

The methodology adopted is vastly different than what is done in [12].

1. Let $X \sim Binomial(n, p)$ with PDF $f_X(x; n, p)$ as given Definition 17, Appendix B.1;
2. Let the classic normal approximation to the binomial distribution be $A \sim \mathcal{N}\left(\hat{\mu} = np, \hat{\sigma}^2 = np(1-p)\right)$ with PDF $f_A(a; \hat{\mu}, \hat{\sigma})$ and calculate $d = \max_{0 \leq i \leq n} |f_X(i; n, p) - f_A(i; \hat{\mu}, \hat{\sigma})|$;
3. Let the $\mathcal{SN}$ approximation be $B \sim \mathcal{SN}\left(\hat{\mu}, \hat{\sigma}^2, \hat{\lambda}\right)$ with PDF $f_B\left(b; \hat{\mu}, \hat{\sigma}, \hat{\lambda}\right)$ as in (2.2) with estimated parameters numerically minimising the maximum distance between $f_X(x; n, p)$ and $f_B\left(b; \hat{\mu}, \hat{\sigma}, \hat{\lambda}\right)$ (2.2) (i.e minimising $d = \max_{0 \leq i \leq n} \left|f_X(i; n, p) - f_B\left(i; \hat{\mu}, \hat{\sigma}, \hat{\lambda}\right)\right|$);
4. Let the $\mathcal{SGN}_I$ approximation be $C \sim \mathcal{SGN}_I\left(\hat{\mu}, \hat{\alpha}^2, \hat{\beta}, \hat{\lambda}\right)$ with PDF $f_C\left(c; \hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\lambda}\right)$ as in (3.11) with estimated parameters numerically minimising the maximum distance between $f_X(x; n, p)$ and $f_C\left(c; \hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\lambda}\right)$ (3.11) (i.e minimising $d = \max_{0 \leq i \leq n} \left|f_X(i; n, p) - f_C\left(i; \hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\lambda}\right)\right|$);
5. The initial values in the optimisation algorithm used to estimate the parameters of the $\mathcal{SN}$ approximation are $\{\mu, \sigma, \lambda\} = \left\{np, \sqrt{np(1-p}, 0\right\}$. λ is set zero so that the optimisation algorithm begins with a symmetric distribution.
6. The initial values in the optimisation algorithm used to estimate the parameters of the $\mathcal{SGN}_I$ approximations $\{\mu, \alpha, \beta, \lambda\} = \left\{np, \sqrt{np(1-p}, 2, 0\right\}$. β is set to two and λ is set to zero so that the optimisation algorithm begins with a symmetric distribution with normal tail behavior.
7. In Step 5 and Step 6, the parameters are set in this way to ensure that the algorithm used does not favour any distribution over another.
8. The approximation error is calculated for $i = 0, \cdots, n$ as
   (a) $f_X(i; n, p) - f_A(i; \hat{\mu}, \hat{\sigma})$ for the $\mathcal{N}$ distribution;
   (b) $f_X(i; n, p) - f_B\left(i; \hat{\mu}, \hat{\sigma}, \hat{\lambda}\right)$ for the $\mathcal{SN}$ distribution and
   (c) $f_X(i; n, p) - f_C\left(i; \hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\lambda}\right)$ for the $\mathcal{SGN}_I$ distribution.

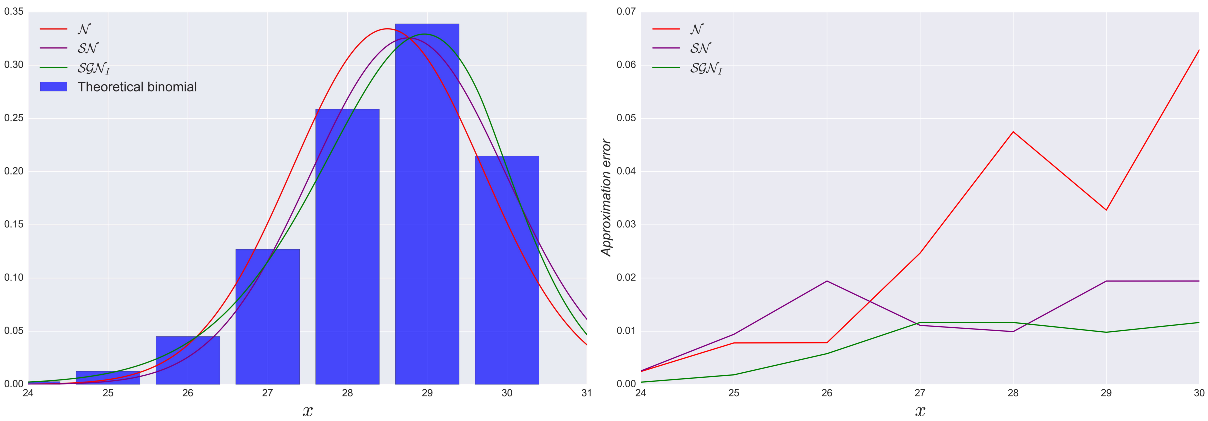
5.2.2 Results

Case 1: $n = 30$, $p = 0.95$

Table 5.3 and Figure 5.2 summarise the results obtained when the $\mathcal{N}$, $\mathcal{SN}$ (see (2.2)) and $\mathcal{SGN}_I$ (see (3.11)) are used to approximate $X \sim Binomial\,(30, 0.95)$:

	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\alpha}^2$	$\hat{\beta}$	$\hat{\lambda}$	d
$\mathcal{N}$	28.5	0.1937^2	-	-	-	0.06289
$\mathcal{SN}$	28.624	1.232^2	-	-	0.140	0.01944
$\mathcal{SGN}_I$	29.672	-	2.083^2	1.696	-1.702	**0.01165**

Table 5.3: d obtained for each of the distributions approximating the $Binomial\,(30, 0.95)$ distribution



(a) $\mathcal{N}$, $\mathcal{SN}$ (2.2) and $\mathcal{SGN}_I$ (3.11) PDFs approximating $f_X\,(x; 30, 0.95)$.

(b) Approximation error

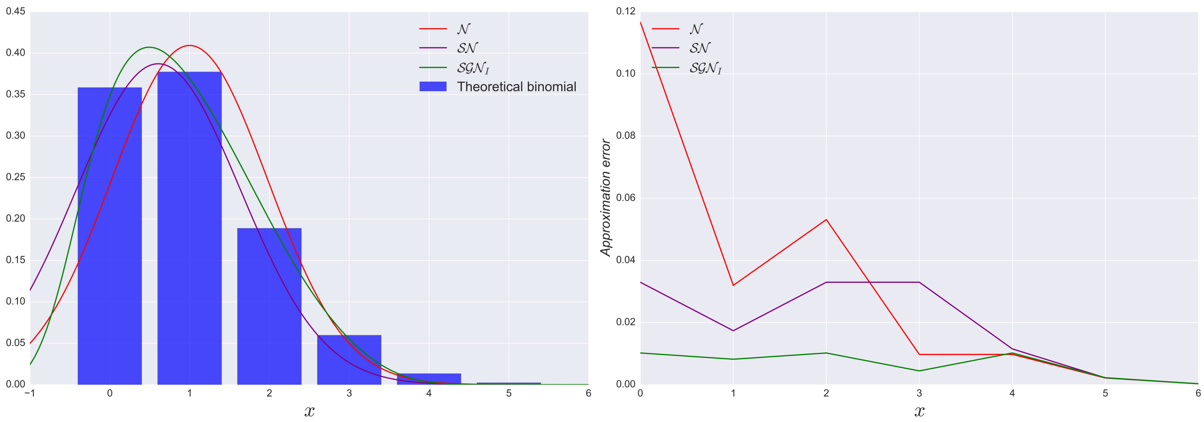
Figure 5.2: $\mathcal{N}$, $\mathcal{SN}$ (2.2) and $\mathcal{SGN}_I$ (3.11) approximations to $Binomial\,(30, 0.95)$ distribution.

Case 2: $n = 20$, $p = 0.05$

Table 5.4 and Figure 5.3 summarise the results obtained when the $\mathcal{N}$, $\mathcal{SN}$ (see (2.2)) and $\mathcal{SGN}_I$ (see (3.11)) distributions are used to approximate $X \sim Binomial\,(20, 0.05)$:

	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\alpha}^2$	$\hat{\beta}$	$\hat{\lambda}$	d
$\mathcal{N}$	1	0.975^2	-	-	-	0.11668
$\mathcal{SN}$	0.318	1.071^2	-	-	0.365	0.03296
$\mathcal{SGN}_I$	1.777	-	2.264^2	5.332	-2.180	**0.01020**

Table 5.4: d obtained for each of the distributions approximating the $Binomial\,(20, 0.05)$ distribution.



(a) $\mathcal{N}$, $\mathcal{SN}$ (2.2) and $\mathcal{SGN}_I$ (3.11) PDFs approximating $f_X\,(x; 20, 0.05)$.

(b) Approximation error

Figure 5.3: $\mathcal{N}$, $\mathcal{SN}$ (2.2) and $\mathcal{SGN}_I$ (3.11) approximations to $Binomial\,(20, 0.05)$ distribution.

It is observed that using the $\mathcal{SGN}_I$ distribution to approximate a binomial distribution with p either large or small results in a overall more accurate approximation compared to both the $\mathcal{N}$ and $\mathcal{SN}$ distributions. This is due to the $\mathcal{SGN}_I$ having two parameters, i.e. β and λ, adding flexibility in accounting for skewness (see Section 3.3 and Section 3.4) of the binomial distribution exhibited when p is large or small. In both cases above, the $\mathcal{SGN}_I$ resulted in the minimum d, and by this measure it is conclude that the $\mathcal{SGN}_I$ distribution outperforms both the $\mathcal{N}$ and $\mathcal{SN}$ distributions in approximating a binomial distribution with p either large or small.

Chapter 6

Conclusion and future work

In Chapter 1 the use and importance of skew-symmetric distributions is motivated and the mechanism that is used to introduced skewness into a symmetric distribution is defined. In Chapter 2 the existing skew-normal distribution and its characteristics are revisited. A existing stochastic representation of the skew-normal distribution is revisited and some existing generalisations of the skew-normal model are presented. In Chapter 3 the basic characteristics and a stochastic representation of the generalised normal distribution is derived. The *skew generalised-normal type I distribution* is then defined and the effect of changing certain parameters of this distribution is examined. An acceptance-rejection algorithm is employed to sample from the skew generalised-normal type I distribution and shortfalls of this approach are noted. Two methods which approximate the characteristics of the skew generalised-normal type I distribution are derived and compared using comprehensive numerical study. Given the shortfalls of the acceptance-rejection algorithm, it was then undertaken to derive a new stochastic representation for the skew generalised-normal type I distribution. In Chapter 4 the *skew generalised-normal type II distribution* is defined and a stochastic representation is derived. The skewing mechanism associated with skew generalised-normal type II distribution is compared to that of the skew generalised-normal type I distribution. The Balakrishnan skew generalised-normal and beta skew generalised-normal distributions are defined. In Chapter 5 a distribution fitting application is presented and it is found that the skew generalised-normal type I distribution was the best fit for the given data. A second application which involves the approximating the binomial distribution using the normal-, skew-normal- and skew generalised-normal type I distribution is also presented. It is determined the the skew generalised-normal type I distribution outperforms the normal and skew-normal distributions in approximating skewed binomial distributions.

6.1 Future work

There are opportunities for future research based on this study:

- Abtahi et. al. [1] introduced a unified skew-normal distribution that replaces the skewing mechanism with a data-driven kernel density estimate. It would be possible to use the generalised normal distribution to extend the unified skew-normal distribution to a unified skew generalised-normal distribution.

- Another possible idea for future work is to extend the skewing mechanism to the elliptical class. For this, the following definition along with a practical lemma is proposed.

Definition 12. A random variable Y has the skew elliptical generalised-normal distribution with location parameter μ, scale parameter α and shape parameter β if its PDF is given by

$$f_Y(y;\mu,\alpha\beta,\lambda) = \frac{2}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\Phi^E\left(\sqrt{2}\lambda\frac{y-\mu}{\alpha}\right)$$

where $\alpha,\beta \in \mathbb{R}^+$, $\lambda \in \mathbb{R}$ and $\Phi^E(\cdot)$ is the CDF of an elliptically contoured distribution [2]. This is denoted by $Y \sim \mathcal{SEGN}\left(\mu,\alpha^2,\beta,\lambda,g\right)$.

Lemma 3. *Let* $Y \sim \mathcal{SEGN}\left(\mu,\alpha^2,\beta,\lambda,g\right)$ *then*

$$f_Y(y;\mu,\alpha\beta,\lambda) = \frac{2}{\alpha}\phi^*\left(\frac{y-\mu}{\alpha};\beta\right)\int_0^\infty \mathcal{W}(t)\,\Phi_{N(0,t^{-1})}\left(\sqrt{2}\lambda\left(\frac{y-\mu}{\alpha}\right)\right)dt \qquad (6.1)$$

where $\beta \in \mathbb{R}^+$, $\lambda \in \mathbb{R}$ and $\Phi_{N(0,t^{-1})}$ is the CDF of a $N\left(0,t^{-1}\right)$ distribution and where $\mathcal{W}(\cdot)$ is a weighting function on $(0,\infty)$.

Proof. Follows from Theorem 17, Appendix B.2 together with (3.9). □

The following table shows the weighting function, $\mathcal{W}(\cdot)$, for three particular skewed distributions that can be explored within this context.

Distribution	$\mathcal{W}(t)$
$\mathcal{SGN}_I$ (see Chapter 3)	$\delta(t-1)$ - Dirac delta function
$\mathcal{SGN}_{\mathcal{T}_\nu}$	$\frac{\nu\left(\frac{\nu t}{2}\right)^{\frac{\nu}{2}-1}}{2\Gamma\left(\frac{\nu}{2}\right)e^{\frac{\nu t}{2}}}$
$\mathcal{SGN}_{Sinusoidal}$	$\frac{(-1)^{\frac{2(\beta-1)}{(\beta+1)}}\sin\left(\frac{t}{2}\right)}{\Gamma\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta\pi}{4}\right)(2t)^{\frac{\beta}{2}}}$

In conclusion it may be of interest to use the generalised normal distribution as the *symmetric base* and extending *skewing mechanism* to the elliptical class to further enhance the skewing mechanism's flexibility (see Proposition 1, Section 1.2).

Bibliography

[1] A. Abtahi, M. Towhidi, and J. Behboodian. An appropriate empirical version of skew-normal density. *Statistical Papers*, 52:469–489, 2011.

[2] M. Arashi and S. Nadarajah. Generalized elliptical distributions. *Communications in Statistics - Theory and Methods*, 46(13):6412–6432, 2017.

[3] R. B. Arellano-Valle, H. Gomez, and F. Quintana. A new class of skew-normal distributions. *Communications in Statistics - Theory and Methods*, 33(7):1465–1480, 2004.

[4] B. C. Arnold, R. J. Beaver, A. Azzalini, N. Balakrishnan, A. Bhaumik, D. K. Dey, C. M. Cuadras, and J. M. Sarabia. Skewed multivariate models related to hidden truncation and/or selective reporting. *Test*, 11(1):7–54, 2002.

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

[6] A. Azzalini. The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, 32(2):159–188, 2005.

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

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

[9] W. Bahrami and E. Qasemi. A flexible skew-generalised normal distribution. *Journal of Statistical Research of Iran*, 11(2):131–145, 2015.

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

[11] N. Balakrishnan. Discussion of "skewed multivariate models related to hidden truncation and/or selective reporting". *Test*, 11:37–39, 2002.

[12] C. Chang, Lin J., N. Pal, and Chiang M. A note on improved approximation of the binomial distribution by the skew-normal distribution. *The American Statistician*, 62(2):167–170, 2008.

[13] K. C. Chu. Estimation and decision for linear system with elliptical random processes. *IEEE Transactions on Automatic Control*, 18:499–505, 1973.

[14] N. Eugene, C. Lee, and F. Famoye. Beta-normal distribution and its appllications. *Communications in Statistics - Theory and Methods*, 31(4):497–512, 2002.

[15] H. W. Gómez, D. Elal-Olivero, H. Salinas, and H. Bolfarine. Bimodal extension based on the skew-normal distribution with application to pollen data. *Environmetrics*, 22(1):50–62, 2011.

[16] H. W. Gómez, H. S. Salinas, and H. Bolfarine. Generalized skew-normal models: properties and inference. *Statistics*, 40(6):495–505, 2006.

[17] A. Gupta and J. Chen. A class of multivariate skew-normal models. *Annals of the Institute of Statistical Mathematics*, 56(2):305, 2004.

[18] S. Harrar and A. Gupta. On matrix variate skew-normal distributions. *Statistics*, 42(2):179–194, 2008.

[19] P. Hasanalipour and M. Sharafi. A new generaslised Balakishnan skew-normal distribution. *Statistical Papers*, 53:219–228, 2012.

[20] N. Henze. A probabilistic representation of the skew-normal distribution. *Scandinavian Journal of Statistics*, 13:271–275, 1986.

[21] K. Lange. *Numerical Analysis for Statisticians*. Springer New York, 2 edition, 2010.

[22] V. Mameli and M. Musio. A generalization of the skew-normal distribution: the beta skew-normal. *Communications in Statistics - Theory and Methods*, 42(12):2229–2244, 2013.

[23] D. B. Owen. Tables for computing bivariate normal probabilities. *The Annals of Mathematical Statistics*, 27(4):1075–1090, 1953.

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

[25] M. Pourahmadi. Construction of skew-normal random variables: Are they linear combinations of normal and half-normal? *Journal of Statistical Theory and Application*, 3:314–328, 2007.

[26] W. Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, 1976.

[27] H. S. Salinas, R. B. Arellano-Valle, and H. W. Gómez. The extended skew-exponential power distribution and its derivation. *Communications in Statistics - Theory and Methods*, 36(9):1673–1689, 2007.

[28] M. Subbotin. On the law of frequency of error. *Matematicheskii Sbornik*, 31(2):296–301, 1923.

[29] R. D. Telford and R.B. Cunningham. Sex, sport and body-size dependency of hematology in highly trained athletes. *Medicine and Science in Sports and Exercise*, 23:788–794, 1991.

[30] O. Venegas, A. Sanhueza, and H. Gomez. An extension of the skew-generalized normal distribution and its derivation. *Proyecciones (Antofagasta)*, 30(3):401–413, 2011.

[31] I. Yadegari, A. Gerami, and M. J. Khaledi. A generalization of the Balakrishnan skew-normal distribution. *Statistics and Probability Letters*, 78(10):1165–1167, 2008.

Appendix A

Notation and symbols used

Symbol	Meaning
$\mathbb{R}$	real number
$\mathbb{R}^d$	$d-$dimensional real space
$\mathbb{R}^+$	positive real number
$\mathbb{Z}^+$	positive integer
$\mathbb{E}\left[\cdot\right]$	expectation operator
$\mathbb{E}_U\left[\cdot\right]$	expectation with respect to random variable U
$\mathbb{P}\left[\cdot\right]$	probability operator
PDF	probability density function
CDF	cumulative distribution function
MGF	moment generating function
γ_1	skewness
γ_2	kurtosis
$\Gamma\left(\cdot\right)$	gamma function
$\gamma\left(\cdot\right)$	incomplete gamma function
$B\left(a,b\right)$	complete beta function

Table A.1: Abbreviation/symbol with corresponding meaning

Symbol	Distribution
$\mathcal{N}$	normal distribution
$\mathcal{SN}$	skew-normal distribution
$Beta\mathcal{SN}$	beta skew-normal distribution
$\mathcal{GN}$	generalised normal distribution
$\mathcal{SGN}_I$	skew generalised-normal type I distribution
$Beta\mathcal{SGN}$	beta skew generalised-normal distribution
$\mathcal{SGN}_{II}$	skew generalised-normal type II distribution
$\phi(\cdot)$	PDF of the standard normal distribution
$\Phi(\cdot)$	CDF of the standard normal distribution
$\phi^*(\cdot)$	PDF of the standard generalised normal type I distribution
$\Phi^*(\cdot)$	CDF of the standard generalised normal type I distribution
$\mathcal{GBSN}_1$	Generalised Balakrishnan skew-normal type I distribution
$\mathcal{GBSN}_1^*$	Balakrishnan skew generalised-normal distribution

Table A.2: Abbreviation/symbol and corresponding distribution

Appendix B

Definitions and results

B.1 Definitions

Definition 13. The inverse Mills ratio is defined as $m(x) = \frac{\phi(x)}{\Phi(x)}$ for $x \in \mathbb{R}$, where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal PDF and CDF respectively [7].

Definition 14. The cumulant generating function, $K_X(t)$, of a random variable X is defined as the logarithm of the moment generating function of X:

$$K_X(t) = \log \mathbb{E}_X\left[e^{tX}\right] = \log M_X(t)$$

Note that $\frac{d^n}{dt^n} K_X(t)\big|_{t=0}$ is the n^{th} central moment of the distribution of X [7].

Definition 15. The complete beta function denoted by $B(a,b)$, is defined as

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

where $a, b \in \mathbb{R}^+$ [10].

Definition 16. A random variable W has the gamma distribution if its PDF is given by

$$f_W(w) = \frac{b^a}{\Gamma(a)} w^{a-1} e^{-bx},\ w > 0 \tag{B.1}$$

where $a, b \in \mathbb{R}^+$. This is denoted by $W \sim Gamma(a,b)$ [10].

Definition 17. A random variable X has the binomial distribution if its PDF is given by

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad , x \in \{0, \cdots, n\} \tag{B.2}$$

where $n \in \mathbb{Z}^+$ and $p \in [0,1]$. This is denoted by $X \sim Binomial(n,p)$ [10].

Definition 18. The gamma function, $\Gamma(\cdot)$, is defined as

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$$

where $z \in \mathbb{R}^+$ [10].

Definition 19. The incomplete gamma function, $\gamma(\cdot)$, is defined as

$$\gamma(a,z) = \int_0^z t^{a-1}e^{-t}dt$$

where $a, z \in \mathbb{R}^+$ [10].

Definition 20. The sign function of $x \in \mathbb{R}$ is defined as [2]

$$sign(x) = \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

Definition 21. The k^{th} moment of a random variable X with PDF $f_X(x)$ is defined as [10]

$$\mathbb{E}\left[X^k\right] = \int_{\mathbb{R}} x^k f_X(x)\,dx.$$

Definition 22. The expected value of a random variable X with PDF $f_X(x)$ is defined as [10]

$$\mathbb{E}[X] = \int_{\mathbb{R}} x f_X(x)\,dx$$

Definition 23. The variance of a random variable X with PDF $f_X(x)$ is defined as [10]

$$var[X] = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2$$

Definition 24. The skewness of a random variable X with PDF $f_X(x)$ is defined as [10]

$$skewness(X) = \gamma_1 = \frac{\mathbb{E}\left[X^3\right] - 3\mathbb{E}[X]\,\mathbb{E}\left[X^2\right] + 2(\mathbb{E}[X])^3}{(var[X])^{\frac{3}{2}}}$$

Definition 25. The kurtosis of a random variable X with PDF $f_X(x)$ is defined as [10]

$$kurtosis(X) = \gamma_2 = \frac{\mathbb{E}\left[X^4\right] - 4\mathbb{E}[X]\,\mathbb{E}\left[X^3\right] + 6(\mathbb{E}[X])^2\,\mathbb{E}\left[X^2\right] - 3(\mathbb{E}[X])^4}{(var[X])^2} - 3$$

Definition 26. Given n data points $y_1, \cdots, y_n$, the one sample Kolmogorov-Smirnov $(K-S)$ distance, d, is defined as the maximum distance between , $F_{Y_i}(y_i)$, the CDF of the theoretical distribution assumed and $\hat{F}_{Y_i}(y_i)$, the empirical distribution function of the data

$$d = \max_{1\le i\le n}\left|F_{Y_i}(y_i) - \hat{F}_{Y_i}(y_i)\right|$$

Definition 27. The Akaike information criterion (AIC) is defined as

$$AIC = 2k - 2\ln \hat{L}$$

where $\hat{L}$ is the maximised value of the likelihood function of the model and k is the number of estimated parameters.

Definition 28. The Bayesian information criterion (AIC) is defined as

$$BIC = \ln(n)\,k - 2\ln \hat{L}$$

where $\hat{L}$ is the maximised value of the likelihood function of the model and k is the number of estimated parameters and n is the number of data points.

B.2 Results

Theorem 13. *Consider random variable X. Let $f(x)$ and $F(x)$ denote the PDF and CDF of X then*

$$f(x) = \frac{d}{dx}F(x)$$

wherever the derivative exists [10].

Theorem 14. *As stated in [10], let X be a random variable with PDF $f_X(\cdot)$ and consider the transformation $Y = u(X)$. Suppose also that the function $u(x)$ is not one-to-one over $A = \{x | f_X(x) > 0\}$. If it is possible to partition A into disjoint subsets $A_1, A_2, \cdots$ such that $u(x)$ is one-to-one over each A_j then for each y in the range of $u(x)$,the equation $y = u(x)$ has a unique solution $x_j = w_j(y)$ over the set A_j and*

$$f_Y(y) = \sum_j f_X(w_j(y)) \left| \frac{d}{dy} w_j(y) \right|.$$

Theorem 15. *The binomial theorem states that it is possible to expand any power of $(x+y)$ into a sum of the form*

$$(x+y)^n = \textstyle\sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k = \displaystyle\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ is a specific positive integer known as a binomial coefficient [10].

Theorem 16. *The CDF of a standard normal random variable can be represented as*

$$\Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^k x^{2k+1}}{k!\,(2k+1)}$$

Proof. Consider

$$\begin{aligned}\Phi(x) &= \frac{1}{2}+\frac{1}{2}\text{erf}\left(\frac{x}{\sqrt{2}}\right)\\ &= \frac{1}{2}+\frac{1}{2}\frac{2}{\sqrt{\pi}}\sum_{k=0}^{\infty}\frac{(-1)^k\left(\frac{x}{\sqrt{2}}\right)^{2k+1}}{k!\,(2k+1)}\\ &= \frac{1}{2}+\frac{1}{\sqrt{2\pi}}\sum_{k=0}^{\infty}\frac{\left(-\frac{1}{2}\right)^k x^{2k+1}}{k!\,(2k+1)}\end{aligned}$$

□

Theorem 17. *If X is an elliptical random variable with parameters μ and σ^2 and PDF denoted $f_X(x)$ then there exists a weighting function $\mathcal{W}(\cdot)$ defined on $(0,\infty)$ such that*

$$f_X(x) = \int_0^{\infty}\mathcal{W}(t)\, f_{\mathcal{N}(\mu,t^{-1}\sigma^2)}(x)\,dt$$

where $f_{\mathcal{N}(\mu,t^{-1}\sigma^2)}(\cdot)$ denotes the PDF of a normal random variable with parameters $\mu\in\mathbb{R}^+$ and $t^{-1}\sigma^2\in\mathbb{R}^+$ [13].

Appendix C

Code

C.1 Chapter 2

C.1.1 The theoretical characteristics of the $\mathcal{SN}(\mu, \sigma^2, \lambda)$ distribution with PDF $f_X(x; \mu, \sigma, \lambda)$ as given in equation (2.2)

```
proc iml;

start charactersistics(lambda) global(mu, sigma);
    delta = lambda/sqrt(1+lambda**2);
    pi = constant('pi');
    expected = mu + delta*sigma*sqrt(2/pi);
    std = sqrt((sigma**2)*(1-(2/pi)*delta**2));
    var = ((sigma**2)*(1-(2/pi)*delta**2));
    third = 0.5*(4-pi)*(delta*sqrt(2/pi))**3;
    fourth = 2*(pi-3)*(delta*sqrt(2/pi))**4;
    skewness = third/((1-(2/pi)*delta**2)**(3/2));
    kurtosis = fourth/((1-(2/pi)*delta**2)**(4/2));

    ans = lambda||expected||std||skewness||kurtosis;
    return (ans);
finish;

*Parameters;
sigma = 1;
mu = 0;
***********;
```

```
lambdaVec = do(-10,10,0.1)`;
do i = 1 to nrow(lambdaVec);
    lambda = lambdaVec[i];
    ans = ans//charactersistics(lambda);
end;

create plot from ans; append form ans; close;

quit;
```

C.1.2 Generation of variates with a $\mathcal{SN}\left(\mu,\sigma^2,\lambda\right)$ distribution with PDF $f_X\left(x;\mu,\sigma,\lambda\right)$ as given in equation (2.2)

```
proc iml;

*Parameters;
n=10000;
lambda =2;
mu = 2;
sigma2 = 2;
sigma = sqrt(sigma2);
************;

*Simulatation of SN random variates;
u1 = rannor(j(n,1,0));
u2 = rannor(j(n,1,0));
data = mu + sigma#((lambda#abs(u1)+u2)/sqrt(1+lambda##2));

*Theoretical PDF;
xDomain = do(-8,8,0.05)`;
density = (2/sigma)#PDF('normal', ((xDomain-mu)/sigma))#CDF('normal', lambda#((xDomain-mu)/sigma));
density = xDomain||density;
************;

create data from data; append from data; close;
create theoretical from density; append from density; close;

quit;
```

C.2 Chapter 3

C.2.1 Generation of variates with a $\mathcal{GN}(\mu, \alpha^2, \beta)$ distribution with PDF $f_X(x; \mu, \alpha, \beta)$ as given in equation (3.1)

```
proc iml;

seed = 0;

start GND_pdf (x) global(mu, alpha, beta);
    pdf = (beta/(2#alpha#gamma(1/beta)))#exp(-(abs((x-mu)/alpha))##beta);
    return pdf;
finish;

*Parameters*;
beta = 3;
mu = 0;
alpah2 = 2;
alpha = sqrt(alpah2);
******************************;

*Theoretical PDF*;
xDomain = do(-8, 8, 0.1)`;
pdf = GND_pdf(xDomain);
plot = xDomain||pdf;
create pdf from plot; append from plot; close;

*Simulation of SGN random variates*;
n=10000;

u = ranuni(j(n,1,seed));
loc1 = (u>0.5);
loc2 = -(u<0.5);
s = loc1 + loc2;
y = rangam(j(n,1,seed+1),1/beta);

x = mu+s#(alpha#(y)##(1/beta));
create data from x; append from x; close;

quit;
```

C.2.2 The characteristics of the $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution using numerical integration for varying λ

```
proc iml;

*Function that will be integrated to obtain moments;
start fx(x) global(beta, lambda, mu, alpha,moment);
   return( (x##moment)#beta/(alpha*gamma(1/beta))#(exp(-((abs(x)-mu)/alpha)##beta))
            #CDF('NORMAL',((x-mu)/alpha)#lambda#sqrt(2)) );
finish;

*Numerical integration;
start EX(w);
   limits = -10||10;
   call quad(w, "fx", limits);
   return(w);
finish;

*Parameters*;
beta = 5;
mu = 0;
alpha2 = 9;
alpha = sqrt(alpha2);
************;

start CharacteristicsOfSGN (characteristic) global(moment);
    do moment = 1 to 4;
       temp = temp||EX(w);
    end;

    ExpectedX  = temp[,1];
    ExpectedX2 = temp[,2];
    ExpectedX3 = temp[,3];
    ExpectedX4 = temp[,4];
```

```
    stdDev =    (ExpectedX2 - ExpectedX##2)##0.5;
    skewness =  (ExpectedX3 - 3#ExpectedX#ExpectedX2+2#ExpectedX##3)/stdDev##3;
    kurtosis =  (ExpectedX4 - 4#ExpectedX#ExpectedX3
                +6#(ExpectedX##2)#ExpectedX2-3#ExpectedX##4);
    kurtosis =  (kurtosis/stdDev##4) - 3;

    characteristic = ExpectedX||stdDev||skewness||kurtosis;
    return (characteristic);
finish;

a = -10; b = 10; step = 0.05;
col = do(a,b,step)‘;

do lambda = a to b by step;
    result = result//CharacteristicsOfSGN(characteristic);
end;

result = col||result;
create result from result; append from result; close;

quit;
```

C.2.3 The characteristics of the $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution for varying λ using Method 2

```
proc iml;

seed = 0;
*Parameters********;
mu=0;
beta = 1;
alpha2 = 2;
alpha = sqrt(alpha2);
********************;
```

```
start characteristics(lambda) global(mu,alpha,beta,seed);
    num=1000000;
    Expected = j(4,1,.);
    do n = 1 to 4;
    y = rangam(j(num,1,seed+1),(n+1)/beta);
    x = (y##(1/beta))#lambda;
    Func = cdf('normal',x*sqrt(2))[:];

    if mod(n,2)=0 then do;
    Expected[n] = Gamma((n+1)/beta)/Gamma(1/beta);
    end;
    if mod(n,2)^=0 then do;
    Expected[n] = (Gamma((n+1)/beta)/Gamma(1/beta))#(2#Func-1);
    end;
    end;

    ExpectedX =  mu + alpha#Expected[1];
    ExpectedX2 = mu##2 + 2#mu#alpha#Expected[1] + (alpha##2)#Expected[2];
    ExpectedX3 = mu##3 + 3#alpha#(mu##2)#Expected[1]
                       + 3#(alpha##2)#mu#Expected[2] + (alpha##3)#Expected[3];
    ExpectedX4 = mu##4 + 4#alpha#(mu##3)#Expected[1] + 6#Expected[2]#(alpha#mu)##2
                       + 4#Expected[3]#(alpha##3)#mu + Expected[4]#alpha##4;

    stdSS =     (ExpectedX2 - ExpectedX##2)##0.5;
    skew =      (ExpectedX3 -3#ExpectedX#ExpectedX2+2#ExpectedX##3)/stdSS##3;
    kurtosis = (ExpectedX4 - 4#ExpectedX#ExpectedX3
               +6#(ExpectedX##2)#ExpectedX2-3#ExpectedX##4);
    kurtosis =  kurtosis/stdSS##4 -3

    result = lambda||ExpectedX||stdSS||skew||kurtosis;
    return(result);
finish;

lambdaVec = do(-10,10,0.01)`;;
```

```
do i = 1 to nrow(lambdaVec);
    lambda = lambdaVec[i];
    ans = ans//characteristics(lambda);
end;
print ans;

create data from ans; append from ans; close;
quit;
```

C.2.4 The range of skewness attainable by $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution

```
proc iml;

seed = 1;
beta = 2.5;

start Func(param) global (beta, seed);
    lambda  = param[1];
    num = 100000;

    Expected = j(3,1,.);
    do n = 1 to 3;
        y = rangam(j(num,1,seed+1),(n+1)/beta);
    x = (y##(1/beta))#lambda#sqrt(2);
    Func = cdf('normal',x)[:];

    if mod(n,2)=0 then do;
        Expected[n] = Gamma((n+1)/beta)/Gamma(1/beta);
    end;
    if mod(n,2)^=0 then do;
    Expected[n] = (Gamma((n+1)/beta)/Gamma(1/beta))#(2#Func-1);
    end;
    end;

    ExpectedX = Expected[1];
```

```
    ExpectedX2 = Expected[2];
    ExpectedX3 = Expected[3];

    stdSS = (ExpectedX2 - ExpectedX##2)##0.5;
    skew =  (ExpectedX3 -3#ExpectedX#ExpectedX2+2#ExpectedX##3)/stdSS##3;
    return ( skew );
finish;

con = {-50, 50};

p = {0};/* initial guess for solution */

opt = {1,1};
call nlpnms(rc, result, "Func", p, opt, con);
print result;

quit;
```

C.2.5 Acceptance-rejection sampling from the $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution

```
proc iml;

start fx(x) global(beta, lambda, mu, alpha);
    return( beta/(alpha#gamma(1/beta))#(exp(-(abs(x-mu)/alpha)##beta))
    #CDF('NORMAL',((x-mu)/alpha)#lambda#sqrt(2)) );
finish;

start gx(x,N,c);
   return( fx(x)/(c*(1/(2*N))) );
finish;

seed=1;

*Parameters;
mu=0;
```

```
alpha=2;
beta = 2;
lambda = 25;
alpha2 = (alpha**2);
*#####################;

*Theoretical PDF;
x = do(-12,12,0.01)`;
f1 = fx(x);
plot = x||f1;
create plot from plot[colname = {'x' 'fx'}];
append from plot; close;
*#####################;

*AR method;
optn = {1,0};
init = 6;
call nlpnra(rc, max, "fx", init, optn);
y_max = fx(max);
N = 12;
c = 2*N*fx(max);

*Number of iterations of the AR algorithm;
num=70000;

u = ranuni(j(num,1,seed));
ustar = ranuni(j(num,1,seed+1));

y = -N + (2*N)*ustar;
gy = gx(y,N,c);

mat = u||y||gy;
loc1 = loc(mat[,1]<=mat[,3]);
loc2 = loc(mat[,1]>mat[,3]);
```

```
accept = y[loc1]||mat[loc1,3]||u[loc1]#y_max;
reject = y[loc2]||mat[loc2,3]||u[loc2]#y_max;

call sort(accept);
call sort(reject);

n=nrow(accept[,1]); * number of random variates generated;
mean = (accept[,1])[:];
var = std(accept[,1]);
skewness = skewness(accept[,1]);
kurtosis = kurtosis(accept[,1]);

print 'Analysis of SGN variates generated using AR method';
print 'sample size = ' n;
print mean var skewness kurtosis;

time1 = time();

timeTaken = time1-time0;
print  timeTaken;

create accept from accept[colname = {'yaccept' 'accept' 'u1'}];
append from accept; close;
create reject from reject[colname = {'yreject' 'reject' 'u2'}];
append from reject; close;

quit;
```

C.2.6 The characteristics of the $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution using Method 1 derived in Chapter 3.6

```
proc iml;
```

```
seed = 1;

*Parameters*;
mu = 0;
alpah2 = 16;
beta = 2;
lambda = 2;
alpha = sqrt(alpah2);
******************************;

n=800000; *number of samples to be drawn;

u = ranuni(j(n,1,seed));
loc1 =  (u>0.5);
loc2 = -(u<0.5);
s = loc1 + loc2;
y = rangam(j(n,1,seed+1),1/beta);

x = mu+s#(alpha#(y)##(1/beta));
create data from x[colname={'data'}]; append from x; close;

ExpectedX =   (2#(x##1)#cdf('normal',sqrt(2)#lambda#(x-mu)/alpha))[:];
ExpectedX2 =  (2#(x##2)#cdf('normal',sqrt(2)#lambda#(x-mu)/alpha))[:];
ExpectedX3 =  (2#(x##3)#cdf('normal',sqrt(2)#lambda#(x-mu)/alpha))[:];
ExpectedX4 =  (2#(x##4)#cdf('normal',sqrt(2)#lambda#(x-mu)/alpha))[:];

stdDev   = (ExpectedX2 - ExpectedX##2)##0.5;
skewness = (ExpectedX3 - 3#ExpectedX#ExpectedX2 + 2#ExpectedX##3)/stdDev##3;
kurtosis = (ExpectedX4 - 4#ExpectedX#ExpectedX3
+6#(ExpectedX##2)#ExpectedX2-3#ExpectedX##4)/stdDev##4 -3;

CharacteristicsSGN = ExpectedX||stdDev||skewness||kurtosis;

quit;
```

C.2.7 The characteristics of the $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution using Method 2 derived in Chapter 3.7 as written in Theorem 9 (random number generation)

```
proc iml;

seed = 0;

*Parameters*;
mu = 0;
alpha2 = 16;
beta = 2;
lambda = 2;
alpha = sqrt(alpha2);
******************************;

num=100000; *number of samples to be drawn;

Expected = j(4,1,.);
do n = 1 to 4;
    y = rangam(j(num,1,seed+1),(n+1)/beta);
    x = (y##(1/beta))#lambda#sqrt(2);
    Func = cdf('normal',x)[:];

    if mod(n,2)=0 then do;
    Expected[n] = Gamma((n+1)/beta)/Gamma(1/beta);
    end;
    if mod(n,2)^=0 then do;
    Expected[n] = (Gamma((n+1)/beta)/Gamma(1/beta))#(2#Func-1);
    end;
end;

ExpectedX =  mu    + alpha#Expected[1];
ExpectedX2 = mu##2 + 2#mu#alpha#Expected[1] + (alpha##2)#Expected[2];
ExpectedX3 = mu##3 + 3#alpha#(mu##2)#Expected[1] + 3#(alpha##2)#mu#Expected[2]
```

```
       + (alpha##3)#Expected[3];
ExpectedX4 = mu##4 + 4#alpha#(mu##3)#Expected[1] + 6#Expected[2]#(alpha#mu)##2
       + 4#Expected[3]#(alpha##3)#mu + Expected[4]#alpha##4;

stdDev =   (ExpectedX2 - ExpectedX##2)##0.5;
skewness = (ExpectedX3 - 3#ExpectedX#ExpectedX2+2#ExpectedX##3)/stdDev##3;
kurtosis = (ExpectedX4 - 4#ExpectedX#ExpectedX3+6#(ExpectedX##2)#ExpectedX2
       - 3#ExpectedX##4)/stdDev##4 -3;

CharacteristicSGN = Expected||stdDev||skewness||kurtosis;
quit;
```

C.2.8 The characteristics of the $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution using Method 2 derived in Chapter 3.7 as written in Theorem 9 (numerical integration)

```
proc iml;
time0 = time();
seed = 0;

*CHANGE*****************************;
mu = 0;
alpha2 = 4;
beta = 3;
lambda = -27;
alpha = sqrt(alpha2);
******************************;

num=100000;

start function(a) global(x,beta,n,lambda);
term = (n+1)/beta;
   func = PDF('Gamma', a,term)#CDF('Normal',sqrt(2)#lambda#(a##(1/beta)));
   return(func);
finish;
```

```
Expected = j(4,1,.);
do n = 1 to 4;
limits = 0 || .P;
   call quad(w, "function", limits);
Func = w;

if mod(n,2)=0 then do;
Expected[n] = Gamma((n+1)/beta)/Gamma(1/beta);
end;
if mod(n,2)^=0 then do;
Expected[n] = (Gamma((n+1)/beta)/Gamma(1/beta))#(2#Func-1);
end;
end;

ExpectedX = mu + alpha#Expected[1];
ExpectedX2 = mu##2 + 2#mu#alpha#Expected[1] + (alpha##2)#Expected[2];
ExpectedX3 = mu##3 + 3#alpha#(mu##2)#Expected[1]
+ 3#(alpha##2)#mu#Expected[2] + (alpha##3)#Expected[3];
ExpectedX4 = mu##4 + 4#alpha#(mu##3)#Expected[1] +
6#Expected[2]#(alpha#mu)##2 + 4#Expected[3]#(alpha##3)#mu
+ Expected[4]#alpha##4;



stdSS = (ExpectedX2 - ExpectedX##2)##0.5;
skew = (ExpectedX3 -3#ExpectedX#ExpectedX2+2#ExpectedX##3)/stdSS##3;
kurtosis = (ExpectedX4 - 4#ExpectedX#ExpectedX3
+6#(ExpectedX##2)#ExpectedX2-3#ExpectedX##4)/stdSS##4 -3;

time1 = time();
print time0 time1;

timetaken = time1-time0;

print 'using Gamma' timetaken ExpectedX stdSS skew kurtosis;
```

```
quit;
```

C.2.9 The characteristics of the $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution using Method 2 derived in Chapter 3.7 as written in (19)

```
proc iml;


start sumApprox(stop) global(mu, alpha, beta, lambda, seed, n, total);
    sum = 0;
    do k = 0 to total by 1;
    const1 = sqrt(2/constant('pi'));
    kthing = (2#k+1)/beta;
    sum1 = (-0.5)##k;
    sum2 = Gamma((n+1)/beta + (2#k+1)/beta)/Gamma((n+1)/beta)#((sqrt(2)#lambda)##(2#k+1));
    sum3 = fact(k)#(2#k+1);
    a = const1#sum1#sum2/sum3;
    sum = sum + a;
    end;
    if abs(a)<0.000001 then print 'converged'; else print "didn't converge";
    return (sum);
finish;



start Method2 (num) global(mu, alpha, beta, lambda, seed, n, stop);
    Expected = j(4,1,.);
    do n = 1 to 4;
    if mod(n,2)=0 then do;
    Expected[n] = Gamma((n+1)/beta)/Gamma(1/beta);
    end;

    if mod(n,2)^=0 then do;
    Expected[n] = (Gamma((n+1)/beta)/Gamma(1/beta))#sumApprox(stop);
    end;
    end;
```

```
    ExpectedX = mu + alpha#Expected[1];
    ExpectedX2 = mu##2 + 2#mu#alpha#Expected[1] + (alpha##2)#Expected[2];
    ExpectedX3 = mu##3 + 3#alpha#(mu##2)#Expected[1] + 3#(alpha##2)#mu#Expected[2]
                       + (alpha##3)#Expected[3];
    ExpectedX4 = mu##4 + 4#alpha#(mu##3)#Expected[1] + 6#Expected[2]#(alpha#mu)##2
                       + 4#Expected[3]#(alpha##3)#mu + Expected[4]#alpha##4;

    stdDev =  (ExpectedX2 - ExpectedX##2)##0.5;
    skewness =  (ExpectedX3 - 3#ExpectedX#ExpectedX2+2#ExpectedX##3)/stdDev##3;
    kurtosis = (ExpectedX4 - 4#ExpectedX#ExpectedX3
               +6#(ExpectedX##2)#ExpectedX2-3#ExpectedX##4)
    kurtosis = (kurtosis/stdDev##4)-3;
    characteristicSGN = ExpectedX||stdDev||skewness||kurtosis;
    return (characteristicSGN);
finish;

*Parameters;
total=50;
mu = 0;
alpha2 = 1 ;
beta =1;
lambda =0.07;
alpha = sqrt(alpha2);
*************;

ans = Method2(1);
print ans;

quit;
```

C.2.10 Visual comparison of Method 1 and Method 2 derived in Section 3.6 and Section 3.7 respectively

```
proc iml;
```

```
start Method1(n) global(mu, alpha, beta, lambda, seed);
u = ranuni(j(n,1,seed));
loc1 =  (u>0.5);
loc2 = -(u<0.5);
s = loc1 + loc2;
y = rangam(j(n,1,seed+1),1/beta);

x = mu+s#(alpha#(y)##(1/beta));

ExpectedX =  (2#(x##1)#cdf('normal',sqrt(2)#lambda#(x-mu)/alpha))[:];
ExpectedX2 = (2#(x##2)#cdf('normal',sqrt(2)#lambda#(x-mu)/alpha))[:];
ExpectedX3 = (2#(x##3)#cdf('normal',sqrt(2)#lambda#(x-mu)/alpha))[:];
ExpectedX4 = (2#(x##4)#cdf('normal',sqrt(2)#lambda#(x-mu)/alpha))[:];

varSS =     (ExpectedX2 - ExpectedX##2);
stdDev =  (ExpectedX2 - ExpectedX##2)##0.5;
deno =        stdDev;
skewness =  (ExpectedX3 -3#ExpectedX#ExpectedX2+2#ExpectedX##3)/stdDev##3;
kurtosis =  (ExpectedX4 - 4#ExpectedX#ExpectedX3
                +6#(ExpectedX##2)#ExpectedX2-3#ExpectedX##4);
kurtosis =  (kurtosis/stdDev##4) - 3;

characteristicSGN = ExpectedX||stdDev||skewness||kurtosis;
return (characteristicSGN);
finish;

start Method2 (num) global(mu, alpha, beta, lambda, seed);
Expected = j(4,1,.);
do n = 1 to 4;
if mod(n,2)=0 then do;
Expected[n] = Gamma((n+1)/beta)/Gamma(1/beta);
end;

if mod(n,2)^=0 then do;
```

```
y = rangam(j(num,1,seed),(n+1)/beta);
x = (y##(1/beta))#lambda#sqrt(2);
Func = cdf('normal',x)[:];
Expected[n] = (Gamma((n+1)/beta)/Gamma(1/beta))#(2#Func-1);
end;
end;

ExpectedX =  mu    + alpha#Expected[1];
ExpectedX2 = mu##2 + 2#mu#alpha#Expected[1] + (alpha##2)#Expected[2];
ExpectedX3 = mu##3 + 3#alpha#(mu##2)#Expected[1] + 3#(alpha##2)#mu#Expected[2]
                       + (alpha##3)#Expected[3];
ExpectedX4 = mu##4 + 4#alpha#(mu##3)#Expected[1] + 6#Expected[2]#(alpha#mu)##2
                   + 4#Expected[3]#(alpha##3)#mu + Expected[4]#alpha##4;

stdDev =  (ExpectedX2 - ExpectedX##2)##0.5;
skewness =  (ExpectedX3 - 3#ExpectedX#ExpectedX2+2#ExpectedX##3)/stdDev##3;
kurtosis =  (ExpectedX4 - 4#ExpectedX#ExpectedX3
                +6#(ExpectedX##2)#ExpectedX2-3#ExpectedX##4);
kurtosis =  kurtosis/stdDev##4-3

characteristicSGN = ExpectedX||stdDev||skewness||kurtosis;
return (characteristicSGN);
finish;

seed = 1;
*Parameters*****************;
mu = 0;
alpha2 = 4;
beta = 2;
lambda = 25;
alpha = sqrt(alpha2);
*****************************;

a = 1000; b=2000000; step = 20000;
```

```
sampleSize = do(a,b,step)`;

t0Method1 = time();
stop1 = 100;
old1 = 100;
do num1 = a to b by step until(stop1 <0.0005);
result1 = Method1(num1);
new1 = abs(result1)[+];
stop1 = abs(new1 - old1);
old1 = new1;
stopSave1 = stopSave1//(num1||stop1);
end;

tMethod1 = time() - t0Method1;
Method1Results =  tMethod1||(num1-step)||result1||stop1;
print Method1Results[colname={'time' 'step' 'expectedValue' 'stdDev' 'skewness' 'kurtosis'
                              'stopCondition'}];

t0Method2 = time();
stop2 = 100;
old2 = 100;
do num2 = a to b by step until(stop2 <0.0005);
result2 = Method2(num2);
new2 = abs(result2)[+];
stop2 = abs(new2 - old2);
old2 = new2;
stopSave2 = stopSave2//(num2||stop2);
end;
tMethod2 = time() - t0Method2;
Method2Results =  tMethod2||(num2-step)||result2||stop2;
print Method2Results[colname={'time' 'step' 'expectedValue' 'stdDev' 'skewness' 'kurtosis'
                              'stopCondition'}];

if (nrow(stopSave2))<(nrow(stopSave1)) then do;
```

```
a = nrow(stopSave1);
b = nrow(stopSave2);
diff = a - b;
append = j(diff, 1, stopSave2[b,2]);
stopSave2 = stopSave2[,2]//append;
end;

stopSave = (stopSave1||stopSave2)[2:nrow(stopSave1),];
create stopSave from stopSave; append from stopSave; close;
quit;
```

C.2.11 Generation of variates with a $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution with PDF $f_X(x; \mu, \alpha\beta, \lambda)$ as given in equation (3.9)

```
proc iml;

mu = 0;
alpha2 = 2;
beta= 5;
lambda =-4;
alpha = sqrt(alpha2);

start SGN(x) global(mu, alpha, beta, lambda);
   return( beta/(alpha#gamma(1/beta))#(exp(-(abs(x-mu)/alpha)##beta))
           #CDF('NORMAL',((x-mu)/alpha)#lambda#sqrt(2)) );
finish;

start ranGN(seed) global(mu, alpha, beta);
    u = ranuni(j(1,1,seed));
    loc1 = (u>0.5);
    loc2 = -(u<0.5);
    s = loc1 + loc2;
    y = rangam(j(1,1,seed),1/beta);
    x = s#((y)##(1/beta));
    return(x);
finish;

*Simulation of SGN random variates;
```

```
do i = 1 to 100000;
    uGN = ranGN(0);
    uMat = rannor(j((1),1,0));
    max = uMat[<>];
    cond = sqrt(2)#lambda#uGN;
    if max < cond then do;
    sample = sample//(uGN#alpha+mu);
    end;
end;

*Theoretical PDF;
x = do(-5, 8, 0.05)`;
SGN = SGN(x);
density = x||SGN;

print (nrow(sample));
create data from sample; append from sample; close;
create pdf from density; append from density; close;

quit;
```

C.2.12 Examining convergence of sample statistics of the $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution for increasing sample size

```
proc iml;

start Method2 (num) global(beta, alpha, lambda, mu, seed);
    Expected = j(4,1,.);
    do n = 1 to 4;
    if mod(n,2)=0 then do;
    Expected[n] = Gamma((n+1)/beta)/Gamma(1/beta);
    end;

    if mod(n,2)^=0 then do;
    y = rangam(j(num,1,seed),(n+1)/beta);
    x = (y##(1/beta))#lambda#sqrt(2);
    Func = cdf('normal',x)[:];
    Expected[n] = (Gamma((n+1)/beta)/Gamma(1/beta))#(2#Func-1);
```

```
    end;
    end;

    ExpectedX = mu + alpha#Expected[1];
    ExpectedX2 = mu##2 + 2#mu#alpha#Expected[1] + (alpha##2)#Expected[2];
    ExpectedX3 = mu##3 + 3#alpha#(mu##2)#Expected[1] + 3#(alpha##2)#mu#Expected[2]
                       + (alpha##3)#Expected[3];
    ExpectedX4 = mu##4 + 4#alpha#(mu##3)#Expected[1] + 6#Expected[2]#(alpha#mu)##2
                       + 4#Expected[3]#(alpha##3)#mu + Expected[4]#alpha##4;

    std =      (ExpectedX2 - ExpectedX##2)##0.5;
    deno =      (ExpectedX2 - ExpectedX##2)##0.5;
    skew =      (ExpectedX3 - 3#ExpectedX#ExpectedX2+2#ExpectedX##3)/deno##3;
    kurtosis = (ExpectedX4 - 4#ExpectedX#ExpectedX3
               +6#(ExpectedX##2)#ExpectedX2-3#ExpectedX##4)
kurtosis = (kurtosis/deno##4)-3;

    ans = ExpectedX||std||skew||kurtosis;

    return (ans);
finish;

seed = 1;
*Parameters*****************************;
beta = 5;
lambda = 2;
mu = 0;
alpha2 = 16;
alpha = sqrt(alpha2);
******************************;

do num = 100 to 100000 by 100;
result1 = result1//(num||Method2(num));
end;
```

```
create data1 from result1; append from result1; close;

quit;
```

C.3 Chapter 4

C.3.1 Theoretical PDF of $\mathcal{SGN}_{II}(\mu, \alpha^2, \beta, \lambda)$ distribution with PDF $f_X(\mu, \alpha, \beta, \lambda)$ as given in equation (4.2)

```
proc iml;

start TrapIntegral(x,y);
   N = nrow(x);
   dx    =   x[2:N] - x[1:N-1];
   meanY = ( y[2:N] + y[1:N-1] )/2;
   return( dx‘ * meanY );
finish;

start ranGN(seed) global(beta);
u = ranuni(j(1,1,seed));
loc1 = (u>0.5);
loc2 = -(u<0.5);
s = loc1 + loc2;
y = rangam(j(1,1,seed),1/beta);
x = s#((y)##(1/beta));
return(x);
finish;

start GND_pdf (x) global(beta);
pdf = (beta/(2#gamma(1/beta)))#exp(-(abs(x))##beta);
return pdf;
finish;

start incompleteGamma(t) global(x,beta);
   func = (t##((1/beta)-1))#exp(-t);
   return(func);
finish;
```

```
start GND_cdf(w) global(x,lambda,beta);
   limits = 0 || abs(x#lambda)##beta;
   call quad(w, "incompleteGamma", limits);
   w = (1/2)+sign(x#lambda)/(2#Gamma(1/beta))#w;
   return(w);
finish;

mu= 0;
alpha2= 2;
alpha = sqrt(alpha2);
beta = 1;
lambdaVec = {-4,-2,0,2,4};

do i = 1 to nrow(lambdaVec);
lambda = lambdaVec[i];
do x = -8 to 8 by 0.01;
pdf = pdf//(((2/alpha)#GND_pdf((x-mu)/alpha)#GND_cdf(w)));
skew = skew//(2#GND_cdf(w));
end;
pdfSave = pdfSave||pdf;
skewSave = skewSave||skew;
free pdf;
free skew;
end;

pdfSave = do(-8,8,0.01)`||pdfSave;
skewSave = do(-8,8,0.01)`||skewSave;

create pdf from pdfSave; append from pdfSave; close;
create skew from skewSave; append from skewSave; close;

quit;
```

C.3.2 Generation of variates with a $\mathcal{SGN}_{II}(\mu, \alpha^2, \beta, \lambda)$ distribution with PDF $f_X(x; \mu, \alpha\beta, \lambda)$ as given in equation (4.1.2)

```
proc iml;

start ranGN(seed) global(beta);
```

```
u = ranuni(j(1,1,seed));
loc1 = (u>0.5);
loc2 = -(u<0.5);
s = loc1 + loc2;
y = rangam(j(1,1,seed),1/beta);
x = s#((y)##(1/beta));
return(x);
finish;

start GND_pdf (x) global(beta);
pdf = (beta/(2#gamma(1/beta)))#exp(-(abs(x))##beta);
return pdf;
finish;

start GND_cdf(w) global(x,lambda,mu,alpha);
    limits = -10 || ((x-mu)/alpha)#lambda;
   call quad(w, "GND_pdf", limits);
   return(w);
finish;

*PARAMETER VALUES;
mu= 5;
alpha = 2;
beta = 1.5;
lambda = 4;

do x = -4 to 10 by 0.05;
pdf = pdf//(x||((2/alpha)#GND_pdf((x-mu)/alpha)#GND_cdf(w)));
end;

*Conditionally sample a skew generalsied normal GN CDF random number;
do i = 1 to 20000;
U = ranGN(0);
uGN = ranGN(0);
cond = lambda#U;
if uGN < cond then do;
sample = sample//(mu + alpha#U);
end;
end;
```

```
quit;
```

C.3.3 Theoretical PDF of $\mathcal{GBSN}_1^*(k, \beta, \lambda_1, \lambda_2)$ distribution with PDF $f_X(k, \beta, \lambda_1, \lambda_2)$ as given in equation (4.7)

```
proc iml;

start TrapIntegral(x,y);
   n = nrow(x);
   dx     =   x[2:n] - x[1:n-1];
   meanY = ( y[2:n] + y[1:n-1] )/2;
   return( dx` * meanY );
finish;

*Generalised normal PDF;
start GND_pdf (x) global(beta);
pdf = (beta/(2#gamma(1/beta)))#exp(-(abs(x)##beta));
return pdf;
finish;

*Normalising constant;
start cnFunc(x) global(mu,alpha,n,beta, lambda1, lambda2);
   return(
           (1/alpha)#GND_pdf((x-mu)/alpha)
            #(CDF('NORMAL', (lambda1#(x-mu))/sqrt(((alpha##2/2)+lambda2#((x-mu)##2)))))##n
         );
finish;

*Numerical integration;
start cn(cnVal);
   limits = .M||.P;
   call quad(cnVal, "cnFunc", limits);
   return(cnVal);
finish;
```

```
*Parameters*;
mu = 0;
alpha = sqrt(2);
n=2;
beta = 1;
lambda1 = 2;
lambda2 = 0;
*****************************;

x = do(-5, 5, 0.01)`;
lambda1Vec = {0,1,2,3,4};
do i = 1 to nrow(lambda1Vec);
lambda1 = lambda1Vec[i];
cn = cn(cnVal);
print cn;
skew = skew||((1/cn)
           #(CDF('NORMAL', (lambda1#(x-mu))/sqrt(alpha##2/2+lambda2#((x-mu)##2))))##n);
BSN = BSN||((1/cn)#(1/alpha)#GND_pdf((x-mu)/alpha)
           #(CDF('NORMAL', (lambda1#(x-mu))/sqrt(alpha##2/2+lambda2#((x-mu)##2))))##n);
end;

density = x||BSN;
skew = x||skew;

print skew;
create skew from skew; append from skew; close;
create pdf from density; append from density; close;
quit;
```

C.3.4 Theoretical PDF of $Beta\mathcal{SGN}\left(\mu,\alpha^2,\beta,\lambda,a,b\right)$ distribution with PDF $f_X\left(\mu,\alpha^2,\beta,\lambda,a,b\right)$ as given in equation (4.9)

```
proc iml;

start SGN_pdf(x) global(mu, alpha, beta, lambda);
fx = 2#beta/(2*alpha*gamma(1/beta))#(exp(-(abs(x-mu)/alpha)##beta))
#cdf('normal',lambda#((x-mu)/alpha));
return(fx);
```

```
finish;

start SGN_cdf(w) global(limits);
   call quad(w, "SGN_pdf", limits);
   return(w);
finish;

*CHANGE****************************;
beta = 2;
lambda = 2;
mu = 0;
alpha2 = 2;
alpha = sqrt(alpha2);
a = 0.2;
b =0.1;
******************************;

do lambda = 0 to 4;
do x = -8 to 8 by 0.05;
limits = .M||x;
BSGN = BSGN//((1/beta(a,b))#((SGN_cdf(w))##(a-1))
#((1-SGN_cdf(w))##(b-1))#SGN_pdf(x));
end;
density = density||BSGN;
free BSGN;
end;

xRange = do(-8,8,0.05)`;
density = xRange||density;
create plot from density; append from density; close;

quit;
```

C.4 Chapter 5

C.4.1 The $K - S$ test to asses the suitability of $\mathcal{SGN}_I(\mu, \alpha^2, \beta, \lambda)$ distribution with PDF $f_X(x; \mu, \alpha\beta, \lambda)$ as given in equation (3.9) to fit to data

```
proc iml;
```

```
*Set numer of iterations;
M = 5000;
do i = 1 to M;
use sasuser.aus; read all into x; close;
x = x[,6];

*Calculate d*;
if i>1 then x = sample(x, nrow(x),'Replace')`;

*Calculated the Empirial distribution functionl;
call sort(x, 1);
ECDF = do(0, 1, 1/(nrow(x)))`;
x_ECDF = (x[1]-0.00001#x[1])//x;

*Define the log-likelihood of the SGN distibution;
start LogLikSGN(param) global (x);
mu  = param[1];
alpha  = param[2];
beta  = param[3];
  lambda  = param[4];
  n = nrow(x);
f = log(beta/(alpha#gamma(1/beta))#(exp(-(abs(x-mu)/alpha)##beta))
#CDF('NORMAL',((x-mu)/alpha)#lambda#sqrt(2)))[+];
return ( f );
finish;

con = {.   0  0.1  .,
       .   .  .   .};

p = {20 5 1 1};/* initial guess for solution */

opt = {1,    /* find maximum of function   */
       0};   /* print a LOT of output      */

call nlpnms(rc, result, "LogLikSGN", p, opt, con);

xDomain = x;
SGN = result[3]/(result[2]#gamma(1/result[3]))#
```

```
(exp(-(abs(xDomain-result[1])/result[2])##result[3]))
#CDF('NORMAL',((xDomain-result[1])/result[2])#result[4]);
SGN_cdf = 0//(cusum(SGN)/SGN[+]);

*Calculate the K-S distance;
KS = max(abs(SGN_cdf-ECDF));
KS_save = KS_save//KS;
end;

*Calculate critical values and p-value;
p={0.9 0.95 0.99};
call qntl(critical_values , KS_save[2:nrow(KS_save)], p);
test_stat = KS_save[1];
pval1 = (KS_save[2:nrow(KS_save)]>test_stat)[:];
print pval1;

print test_stat critical_values
[rowname={'alpha=0.1' 'alpha=0.05' 'alpha=0.01'}];


quit;
```

C.4.2 Maximum likelihood estimation of parameters of distributions fitted to data

```
*NORMAL DISTRIBUTION;
proc iml;

use sasuser.aus; read all into x; close;
x = x[,6];

create data from x; append from x; close;

start LogLikN(param) global (x);
mu  = param[1];
alpha  = param[2];
  n = nrow(x);
```

```
f = log((1/alpha)#PDF('NORMAL',((x-mu)/alpha)))[+];
return ( f );
finish;

con = {.    0     ,
       .    .     };

p = {20 5};/* initial guess for solution */

opt = {1,     /* find maximum of function    */
       1};    /* print a LOT of output       */

call nlpnms(rc, result, "LogLikN", p, opt, con);

xDomain = do(min(x)-1,max(x)+1,0.01)`;
density = (1/result[2])#PDF('NORMAL',((xDomain-result[1])/result[2]));

N = xDomain||(density);

AIC_N = 2*ncol(result) - (2*(-498.66789));

nData = ncol(x);
BIC_N = log(nData)*ncol(result) - (2*(-498.66789));

print AIC_N BIC_N;
print result;

quit;

*SKEW NORMAL DISTRIBUTION;
proc iml;

use sasuser.aus; read all into x; close;
x = x[,6];

start LogLikSN(param) global (x);
mu   = param[1];
alpha  = param[2];
  lambda  = param[3];
```

```
  n = nrow(x);
f = log((2/(alpha))#PDF('NORMAL',((x-mu)/alpha))
#CDF('NORMAL',((x-mu)/alpha)#lambda))[+];
return ( f );
finish;

con = {.   0  .  ,
       .   .  .  };

p = {20 5 1};/* initial guess for solution */

opt = {1,    /* find maximum of function   */
       1};   /* print a LOT of output      */

call nlpnms(rc, result, "LogLikSN", p, opt, con);
print rc;
xDomain = do(min(x)-1,max(x)+1,0.01)`;
density = (2/(result[2]))#PDF('NORMAL',((xDomain-result[1])/result[2]))
#CDF('NORMAL',((xDomain-result[1])/result[2])#result[3]);

SN = xDomain||(density);

AIC_SN = 2*ncol(result) - (2*(-490.09936));

nData = ncol(x);
BIC_SN = log(nData)*ncol(result) - (2*(-490.09936));

print AIC_SN BIC_SN;
print result;
quit;

*SKEW GENERALISED NORMAL DISTRIBUTION;
proc iml;

use sasuser.aus; read all into x; close;
x = x[,6];

start LogLikSGN(param) global (x);
mu  = param[1];
```

```
alpha  = param[2];
beta  = param[3];
  lambda  = param[4];
  n = nrow(x);
f = log(beta/(alpha#gamma(1/beta))
    #(exp(-(abs(x-mu)/alpha)##beta))
      #CDF('NORMAL',((x-mu)/alpha)#lambda#sqrt(2)))[+];
return ( f );
finish;

con = {.   0  0.1  .,
       .   .  .    .};

p = {20 5 1 1};/* initial guess for solution */

opt = {1,    /* find maximum of function   */
       1};   /* print a LOT of output      */

call nlpnms(rc, result, "LogLikSGN", p, opt, con);

xDomain = do(min(x)-1,max(x)+1,0.01)`;
density = result[3]/(result[2]#gamma(1/result[3]))
    #(exp(-(abs(xDomain-result[1])/result[2])##result[3]))
    #CDF('NORMAL',((xDomain-result[1])/result[2])#result[4]#sqrt(2));
SGN = xDomain||(density);



AIC_SGN = 2*ncol(result) - (2*(-488.17452));

nData = ncol(x);
BIC_SGN = log(nData)*ncol(result) - (2*(-488.17452));

print AIC_SGN BIC_SGN;
print result;

quit;

*GENERALISED BALAKRISHNAN SKEW NORMAL DISTRIBUTION;
proc iml;
```

```
use sasuser.aus; read all into x; close;
x = x[,6];
start LogLikGBSK(param) global ( x, mu, alpha, n, lambda1, lambda2);
mu  = param[1];
alpha  = param[2];
n = param[3];
  lambda1 = param[4];
lambda2 = param[5];

u = rannor(j(10000,1,2));
fun = ((CDF('normal', lambda1#(u)/sqrt(1+lambda2#(u##2))))##n)[:];
cn = 1/fun;
f = log((cn/(alpha)) # PDF('NORMAL', (x-mu)/alpha)
        #CDF('NORMAL', lambda1#((x-mu))/sqrt(alpha##2+lambda2#((x-mu))##2))##n)[+];
return ( f );
finish;



con = {.   0  0  . 0,
       .   .  .  . .};



p = {19.97 4.13 1 2.313 0};/* initial guess for solution */

opt = {1,    /* find maximum of function   */
       1};   /* print a LOT of output      */

call nlpnms(rc, result, "LogLikGBSK", p, opt, con);

xDomain = do(min(x)-1,max(x)+1,0.01)`;

mu  = result[1];
alpha  = result[2];
n = result[3];
lambda1 = result[4];
lambda2 = result[5];

u = rannor(j(100000,1,1));
```

```
fun = ((CDF('normal', lambda1#(u)/sqrt(1+lambda2#(u##2))))##n)[:];
cn = 1/fun;

density = (cn/(alpha)) # PDF('NORMAL', (xDomain-mu)/alpha)
          #CDF('NORMAL', lambda1#((xDomain-mu))
          /sqrt(alpha##2+lambda2#((xDomain-mu))##2))##n;
GBSN = xDomain||(density);

create GBSN from GBSN;
append from GBSN; close;

AIC_GBSN = 2*ncol(result) - (2*(-489.9097766));

nData = ncol(x);
BIC_GBSN = log(nData)*ncol(result) - (2*(-489.9097766));

print result;
print AIC_GBSN BIC_GBSN;

quit;
```

C.4.3 Approximating the binomial distribution with $\mathcal{N}$, $\mathcal{SN}$ and $\mathcal{SGN}_I$ distribution

```
proc iml;

start SN(param) global (mu, sigma, beta, lambda, seed, p_prob, n);
    seed = 1;
mu = param[1];
sigma = param[2];
lambda = param[3];
xDomain = do(0,n,1)`;
density = (2/sigma)#PDF('NORMAL',((xDomain-mu)/sigma))
#CDF('NORMAL',((xDomain-mu)/sigma)#lambda);
prob = PDF('Binomial', xDomain , p_prob , n);
```

```
dif =   max(abs(density-prob));
return ( dif );
finish;

start SGN(param) global (target, mu, alpha,beta,lambda,seed,p_prob,n);
    seed = 1;
mu = param[1];
alpha = param[2];
beta = param[3];
lambda = param[4];
xDomain = do(0,n,1)`;
density = beta/(alpha#gamma(1/beta))
    #(exp(-(abs(xDomain-mu)/alpha)##beta))
    #CDF('NORMAL',((xDomain-mu)/alpha)#lambda);
prob = PDF('Binomial', xDomain , p_prob , n);
dif =   max(abs(density-prob));
return ( dif );
finish;



n=20;
p_prob=0.05;

con1 = {.   0.1  .   ,
       .    .   . };
p1 = {1 0.97 0};/* initial guess for solution */

con2 = {.   0.1  0.1 .   ,
       .    .     . .};
p2 = {1 0.97 2 0};/* initial guess for solution */


opt = {0,    /* find maximum of function    */
       1};   /* print a LOT of output       */

call nlpnms(rc, result1, "SN", p1, opt, con1);
print result1;
muSN = result1[1];
sigma = result1[2];
```

```
lambdaSN = result1[3];

call nlpnms(rc, result2, "SGN", p2, opt, con2);
print result2;
muSGN = result2[1];
alpha = result2[2];
beta = result2[3];
lambdaSGN = result2[4];

now = do(0,n,1)`;
binomDens = PDF('Binomial', now , p_prob , n);
Ndens = PDF('Normal', now , n#p_prob , sqrt(n#p_prob#(1-p_prob)));
SNdens =  (2/sigma)#PDF('NORMAL',((now-muSN)/sigma))
#CDF('NORMAL',((now-muSN)/sigma)#lambdaSN);
SGNdens = beta/(alpha#gamma(1/beta))
    #(exp(-(abs(now-muSGN)/alpha)##beta))
    #CDF('NORMAL',((now-muSGN)/alpha)#lambdaSGN);

diffN = abs(Ndens-binomDens);
diffSN = abs(SNdens-binomDens);
diffSGN = abs(SGNdens-binomDens);
diff = now||diffN||diffSN||diffSGN;

diffNmax = diffN[<>];
diffSNmax = diffSN[<>];
diffSGNmax = diffSGN[<>];

print diffNmax diffSNmax diffSGNmax;

xDomain = do(-n,n+10,0.01)`;
Ndens = PDF('Normal', xDomain , n#p_prob , sqrt(n#p_prob#(1-p_prob)));
SNdens =  (2/sigma)#PDF('NORMAL',((xDomain-muSN)/sigma))
#CDF('NORMAL',((xDomain-muSN)/sigma)#lambdaSN);
SGNdens = beta/(alpha#gamma(1/beta))
    #(exp(-(abs(xDomain-muSGN)/alpha)##beta))
    #CDF('NORMAL',((xDomain-muSGN)/alpha)#lambdaSGN);

print diff;
create diff from diff[colname = {'x' 'N' 'SN' 'SGN'}];
```

```
append from diff; close;

binomDens = now||binomDens;
create binomDens from binomDens; append from binomDens; close;

N = xDomain||Ndens;
create N from N; append from N; close;

SN = xDomain||SNdens;
create SN from SN; append from SN; close;

SGN = xDomain||SGNdens;
create SGN from SGN; append from SGN; close;

quit;

proc export data=diff
   outfile="C:\Users\Brett Rowland\Desktop\Full paper\New Binomial pictures\One\diff.csv"
   dbms=csv replace;
run;
proc export data=binomDens
   outfile="C:\Users\Brett Rowland\Desktop\Full paper\New Binomial pictures\One\binomDens.csv"
   dbms=csv replace;
run;

proc export data=N
   outfile="C:\Users\Brett Rowland\Desktop\Full paper\New Binomial pictures\One\N.csv"
   dbms=csv replace;
run;
proc export data=SN
   outfile="C:\Users\Brett Rowland\Desktop\Full paper\New Binomial pictures\One\SN.csv"
   dbms=csv replace;
run;
proc export data=SGN
    outfile="C:\Users\Brett Rowland\Desktop\Full paper\New Binomial pictures\One\SGN.csv"
   dbms=csv replace;
run;
```