Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v10n2p512

Normal-power series class of distributions: model, properties and applications

By Mahmoudi, Mahmoodian

Published: 14 October 2017

Electronic Journal of Applied Statistical Analysis
Vol. 10, Issue 02, October 2017, 512-541
DOI: 10.1285/i20705948v10n2p512

Normal-power series class of distributions: model, properties and applications

Eisa Mahmoudi*and Hamed Mahmoodian

Department of Statistics, Yazd University, P.O. Box 89175-741, Yazd, Iran

Published: 14 October 2017

In this paper, a new class of distributions, called as normal-power series (NPS), which contains the normal one as a particular case, is introduced. This class, which is obtained by compounding the normal and power series distributions, is presented as an alternative to the class of skew-normal and Balakrishnan skew-normal distributions, among others. The density and distribution functions of this family of distributions, are given by a closed expression which allows us to easily compute probabilities, moments and related measurements. The maximum likelihood method via an EM-algorithm is used to estimate the unknown parameters. Finally, some applications are given to show the flexibility of the new class of distributions.

keywords: Normal distribution, Maximum likelihood, Power series class of distributions, EM-algorithm.

1 Introduction

Recently, many distributions to model lifetime data have been studied and generalized by compounding of some discrete and important lifetime distributions. Adamidis and Loukas (1998), introduced exponential-geometric (EG) distribution by compounding the exponential and geometric distributions. In the similar manner, exponential-Poisson (EP), exponential-logarithmic (EL), exponential-power series (EPS), Weibull-geometric (WG), Weibull-power series (WPS), generalized exponential-power series (GEPS), linear failure rate-power series (LFRPS), exponentiated extended Weibull-power series

*Corresponding author:emahmoudi@yazd.ac.ir.

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

distributions (EEWPS), quadratic hazard rate power Series (QHRPS) and The complementary exponentiated Burr XII Poisson (CEBXIIP) were introduced by Tahmasbi and Rezaei (2008), Chahkandi and Ganjali (2009), Barreto-Souza et al. (2011), Morais and Barreto-Souza (2011), Mahmoudi and Jafari (2012), Mahmoudi and Jafari (2017), Tahmasbi and Jafari (2015), Roozegar and Nadarajah (2016) and Muhammad (2017), respectively.

In recent years, techniques for extending the family of normal distributions have been proposed. The method applied here can be considered as an alternative to the well-known skew-normal distribution (Azzalin, 1985), whose properties (Azzalini, 1986; Azzalini and Chiogna, 2004), estimation (Gupta and Gupta, 2008), diagnostics (Xie et al., 2009), generalization (Gupta and Gupta, 2004) and multivariate extension (Azzalini and Valle, 1996; Azzalini and Capitanio, 1999; Arnold and Beaver, 2002) have been widely developed. Other ways to obtaining skewed normal distributions have also been introduced, such as the Balakrishnan skew-normal density (Sharafi and Behboodian, 2008), the variance-gamma process (Fung and Seneta, 2007) and the generalized normal distribution (Nadarajah, 2005), among others. Whenever the Fisher information matrix of this skew-normal model is singular for values of the added skew parameter λ, and the maximum likelihood estimate of this parameter can be infinite with a positive probability, an alternative model would be desirable.

In this paper, we introduce a new generalization of the normal distribution which is called the NPS class of distributions and is denoted by $NPS(\mu, \sigma, \theta)$. The method used to insert the new shape parameter θ, is described in Marshall and Olkin (1997) for the first time, where it was applied to the exponential and Weibull families. This method enables us to obtain explicit expressions for the probability density and survival functions and allow us to estimate the model parameters via an EM-algorithm.

Note that all of these distributions have support $(0, \infty)$, but in this paper we introduced the new generalization of normal distribution, which has range on $(-\infty, \infty)$. To begin with, we shall use the following notation throughout this paper: $\phi(\cdot)$ for the standard normal probability density function (pdf), $\phi_n(\cdot\,; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ for the pdf of $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ (n -variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$), $\Phi_n(\cdot\,; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ for the cdf of $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ (in both singular and non-singular cases), simply $\Phi_n(\cdot\,; \boldsymbol{\Sigma})$ for the case when $\boldsymbol{\mu} = \mathbf{0}$.

The rest of the paper is organized as follows. In Section 2, we introduce the NPS class of distributions. The density, hazard rate and survival functions and some of their properties are given in this section. In Section 3, we derive moments of NPS by two methods. In Section 4, we present some special distributions which are studied in details. Some properties of sub-model of NPS distristributions are studied in Section 5. Estimation of the parameters by maximum likelihood method and inference for large samples are presented in Section 6. A method for evaluating the standard errors from the EM-algorithm is presented in Section 7. Simulation study is given in Section 8. Applications to two real data sets are given in Section 9. Finally, Section 10 concludes the paper.

Table 1: Quantities for power series distributions

Distribution	a_n	$C(\theta)$	$C'(\theta)$	$C''(\theta)$	$C'''(\theta)$	s
Geometric	1	$\theta(1-\theta)^{-1}$	$(1-\theta)^{-2}$	$2(1-\theta)^{-3}$	$6(1-\theta)^{-4}$	1
Poisson	$n!^{-1}$	$e^{\theta}-1$	e^{θ}	e^{θ}	e^{θ}	∞
Logarithmic	n^{-1}	$-\log(1-\theta)$	$(1-\theta)^{-1}$	$(1-\theta)^{-2}$	$2(1-\theta)^{-3}$	1
Binomial	$\binom{k}{n}$	$(1+\theta)^k-1$	$\frac{k}{(1+\theta)^{1-k}}$	$\frac{k(k-1)}{(1+\theta)^{2-k}}$	$\frac{k(k-1)(k-2)}{(1+\theta)^{3-k}}$	∞
Neg. Binomial	$\binom{n-1}{k-1}$	$\frac{\theta^k}{(1-\theta)^k}$	$\frac{k\theta^{k-1}}{(1-\theta)^{k+1}}$	$\frac{k(k+2\theta-1)}{\theta^{2-k}(1-\theta)^{k+2}}$	$\frac{k(k^2+6k\theta+6\theta^2-3k-6\theta+2)}{\theta^{3-k}(1-\theta)^{k+3}}$	1

2 The NPS class of distributions

Given N, let $X_1, .., X_N$ be a random sample from normal distribution with mean μ and variance σ^2. Here the random variable N, independent of X_i's, belongs to a power series distributions (truncated at zero) with the probability mass function

$$P(N=n) = \frac{a_n \theta^n}{C(\theta)},$$

where $a_n \geq 0$ depends only on n, $C(\theta) = \sum_{n=1}^{\infty} a_n \theta^n$ and $\theta \in (0, s)$ (s can be ∞) is such that $C(\theta)$ is finite. Table 1 lists some particular cases of the truncated (at zero) power series distributions (geometric, Poisson, logarithmic, binomial and negative binomial). Detailed properties of power series distributions can be found in Noack (1950). Here, $C'(\theta)$, $C''(\theta)$ and $C'''(\theta)$ denote the first, second and third derivatives of $C(\theta)$ with respect to θ, respectively.

Let $Y = X_{(N)} = \max(X_1, .., X_N)$, then the conditional cdf of $Y|N=n$ is given by

$$G_{Y|N=n}(y) = \left(\Phi\left(\frac{y-\mu}{\sigma}\right)\right)^n,$$

where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution.
The cdf of NPS class of distributions is defined by the marginal cdf of Y, i.e.,

$$F(y;\mu,\sigma,\theta) = \sum_{n=1}^{\infty} \frac{a_n\theta^n}{C(\theta)} \left(\Phi\left(\frac{y-\mu}{\sigma}\right)\right)^n = \frac{C\left(\theta\Phi\left(\frac{y-\mu}{\sigma}\right)\right)}{C(\theta)}, \tag{2.1}$$

where $y \in \mathbb{R}$, $\mu \in \mathbb{R}$ and $\sigma > 0$. We denote a random variable Y follows NPS distributions by notation $NPS(\mu, \sigma, \theta)$. The density function of NPS follows immediately as

$$f(y;\mu,\sigma,\theta) = \frac{\theta}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\frac{C'(\theta\Phi\left(\frac{y-\mu}{\sigma}\right))}{C(\theta)}. \tag{2.2}$$

The corresponding survival and hazard rate functions are

$$S(y;\mu,\sigma,\theta) = 1 - \frac{C(\theta\Phi\left(\frac{y-\mu}{\sigma}\right))}{C(\theta)},$$

and

$$h(y;\mu,\sigma,\theta) = \frac{\theta}{\sigma}\frac{\phi\left(\frac{y-\mu}{\sigma}\right)C'(\theta\Phi\left(\frac{y-\mu}{\sigma}\right))}{C(\theta) - C(\theta\Phi\left(\frac{y-\mu}{\sigma}\right))},$$

respectively.
One can put $\mu = 0$ and $\sigma = 1$ and obtains the standard version on NPS class of distributions denote by notation $NPS(0,1,\theta)$. Figure 1 shows the probability density function and hazard rate function of the classical normal distribution and the NPS distributions proposed in this paper for some choice of $C(\theta)$. It can be seen that the new model is very versatile and that the value of θ has a substantial effect on the skewness of the probability density function.

Proposition 1. *Let* $c = \min\{n \in \mathbb{N}\colon a_n > 0\}$. *As* $\theta \to 0^+$ *we have*

$$\begin{aligned}\lim_{\theta\to 0^+} F(y;\mu,\sigma,\theta) &= \lim_{\theta\to 0^+}\frac{C(\theta\Phi\left(\frac{y-\mu}{\sigma}\right))}{C(\theta)} = \lim_{\theta\to 0^+}\frac{\sum_{n=1}^{\infty} a_n\theta^n(\Phi\left(\frac{y-\mu}{\sigma}\right))^n}{\sum_{n=1}^{\infty} a_n\theta^n}\\ &= \lim_{\theta\to 0^+}\frac{\sum_{n=1}^{c-1} a_n\theta^n(\Phi\left(\frac{y-\mu}{\sigma}\right))^n + a_c\theta^c(\Phi\left(\frac{y-\mu}{\sigma}\right))^c + \sum_{n=c+1}^{\infty} a_n\theta^n(\Phi\left(\frac{y-\mu}{\sigma}\right))^n}{\sum_{n=1}^{c-1} a_n\theta^n + a_c\theta^c + \sum_{n=c+1}^{\infty} a_n\theta^n}\\ &= \lim_{\theta\to 0^+}\frac{(\Phi\left(\frac{y-\mu}{\sigma}\right))^c + a_c^{-1}\sum_{n=c+1}^{\infty} a_n\theta^{n-c}(\Phi\left(\frac{y-\mu}{\sigma}\right))^n}{1 + a_c^{-1}\sum_{n=c+1}^{\infty} a_n\theta^n} = \left(\Phi\left(\frac{y-\mu}{\sigma}\right)\right)^c.\end{aligned}$$

Proposition 2. *The densities of NPS class of distributions can be written as infinite number of linear combination of density of order statistics. We know that* $C'(\theta) = \sum_{n=1}^{\infty} na_n\theta^{n-1}$, *therefore*

$$f(y;\mu,\sigma,\theta) = \frac{\theta}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\frac{C'(\theta\Phi\left(\frac{y-\mu}{\sigma}\right))}{C(\theta)} = \sum_{n=1}^{\infty} g_{(n)}(y;n)P(N=n),$$

where $g_{(n)}(y;n)$ *denotes the density function of* $X_{(n)} = \max(X_1,..,X_n)$.

Proposition 3. *The* γ*th quantile of the NPS class of distributions is given by*

$$y_\gamma = \sigma\Phi^{-1}\left(\frac{C^{-1}(\gamma C(\theta))}{\theta}\right) + \mu.$$

One can use this expression for generating a random sample from NPS distributions with generating data from uniform distribution.

3 Moments of the NPS

In this section we give two methods to obtain the moments of the NPS distributions.

(i) First method

Jamalizadeh and Balakrishnan (2010) considered unified skew-elliptical (SUE) distribution which contains unified skew-normal (SUN) distribution as a special case. The univariate random variable $Z_{k,\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{\Omega})\,, \boldsymbol{\lambda}, \boldsymbol{\gamma} \in \mathbb{R}^k$ and $\boldsymbol{\Omega} \in \mathbb{R}^{k\times k}$ is a positive definite dispersion matrix, is said to have a unified skew-normal distribution, denoted by $Z_{k,\boldsymbol{\theta}} \sim SUN\,(k, \boldsymbol{\theta})$, if its pdf is given by

$$\phi_{SUN}\,(z; k, \boldsymbol{\theta}) = \frac{\phi\,(z)\,\Phi_k\,(\boldsymbol{\lambda} z + \boldsymbol{\gamma}; \boldsymbol{\Omega})}{\Phi_{_k}\left(\boldsymbol{\gamma}; \boldsymbol{\Omega}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T\right)}.$$

Furthermore, the moment generating function of $Z_{k,\boldsymbol{\theta}}$ for $s \in \mathbb{R}$, is given by

$$M_{SUN}\,(s; k, \boldsymbol{\theta}) = \frac{\exp\left(\frac{1}{2}s^2\right)\Phi_{_k}\left(\boldsymbol{\lambda} s + \boldsymbol{\gamma}; \boldsymbol{\Omega}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T\right)}{\Phi_{_k}\left(\boldsymbol{\gamma}; \boldsymbol{\Omega}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T\right)}.$$

The mean of $Z_{k,\boldsymbol{\theta}}$ can be obtained as the following lemma.

Lemma 1. *If $Z_{k,\boldsymbol{\theta}} \sim SUN\,(k, \boldsymbol{\theta})$, then*

$$\begin{aligned} E\left(Z_{k,\boldsymbol{\theta}}\right) &= \frac{1}{\Phi_{_k}\left(\boldsymbol{\gamma}; \boldsymbol{\Omega}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T\right)} \sum_{i=1}^{k} \frac{\lambda_i}{\sqrt{\omega_{ii} + \lambda_i^2}}\phi\left(\frac{\gamma_i}{\sqrt{\omega_{ii} + \lambda_i^2}}\right) \\ &\quad \times \Phi_{k-1}\left(\boldsymbol{\gamma}_{-i} - \frac{\gamma_i}{\omega_{ii} + \lambda_i^2}\left(\boldsymbol{\omega}_{-ii} + \lambda_i \boldsymbol{\lambda}_{-i}\right); \left(\boldsymbol{\Omega}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T\right)_{-i|i}\right), \end{aligned}$$

where, for some i,

$$\boldsymbol{\lambda} = \begin{pmatrix} \lambda_i \\ \boldsymbol{\lambda}_{-i} \end{pmatrix}, \ \boldsymbol{\gamma} = \begin{pmatrix} \gamma_i \\ \boldsymbol{\gamma}_{-i} \end{pmatrix}, \ \boldsymbol{\Omega} = \begin{pmatrix} \omega_{ii} & \boldsymbol{\omega}_{-ii}^T \\ \boldsymbol{\omega}_{-ii} & \boldsymbol{\Omega}_{-i-i} \end{pmatrix},$$

with $\left(\boldsymbol{\Omega}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T\right)_{-i|i} = \boldsymbol{\Omega}_{-i-i} + \boldsymbol{\lambda}_{-i}\boldsymbol{\lambda}_{-i}^T - \frac{(\boldsymbol{\omega}_{-ii}+\lambda_i\boldsymbol{\lambda}_{-i})(\boldsymbol{\omega}_{-ii}+\lambda_i\boldsymbol{\lambda}_{-i})^T}{\omega_{ii}+\lambda_i^2}$.

In the special case when $\boldsymbol{\gamma} = \mathbf{0}$, the moments can be determined rather easily. If in this case the unified skew-normal is denoted by $Z_{k,\lambda,\boldsymbol{\Omega}}$, we then have

$$E\left(Z_{k,\boldsymbol{\lambda},\boldsymbol{\Omega}}\right) = \frac{1}{\Phi_{_k}\left(\mathbf{0}; \boldsymbol{\Omega}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T\right)\sqrt{2\pi}} \sum_{i=1}^{k} \frac{\lambda_i}{\sqrt{\omega_{ii} + \lambda_i^2}}\ \Phi_{k-1}\left(\mathbf{0}; \left(\boldsymbol{\Omega}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T\right)_{-i|i}\right).$$

In this case, a recurrence formula for the moments of $Z_{k,\boldsymbol{\lambda},\boldsymbol{\Omega}} \sim SN\,(k, \boldsymbol{\lambda}, \boldsymbol{\Omega})$ was obtained. For simplicity, in the following lemma, this recurrence formula is presented when $\boldsymbol{\Omega}$ is the correlation matrix.

Lemma 2. *We have, for $m = 1, 2, \cdots$,*

$$\begin{aligned} E\left(Z_{k,\boldsymbol{\lambda},\boldsymbol{\Omega}}^{m+1}\right) &= mE\left(Z_{k,\boldsymbol{\lambda},\boldsymbol{\Omega}}^{m-1}\right) \\ &+ \frac{1}{\sqrt{2\pi}\Phi_{_k}\left(\mathbf{0}; \boldsymbol{\Omega}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T\right)} \sum_{i=1}^{k} \frac{\lambda_i \Phi_{_{k-1}}\left(\mathbf{0}; \boldsymbol{\Omega}_{-i|i}+\boldsymbol{\lambda}_i^*\boldsymbol{\lambda}_i^{*T}\right)}{\left(1+\lambda_i^2\right)^{\frac{m+1}{2}}} E\left(Z_{k-1,\boldsymbol{\lambda}_i^*,\boldsymbol{\Omega}_{-i|i}}^m\right), \end{aligned}$$

where $\boldsymbol{\lambda}_i^* = \frac{\boldsymbol{\lambda}_{-i}}{\sqrt{1+\lambda_i^2}}$.

In addition when $\mathbf{X} \sim N_n\left(\mu\mathbf{1}_n, \sigma^2\{(1-\rho)\mathbf{I}_n + \rho\mathbf{1}_n\mathbf{1}_n^T\}\right)$,
$\mu \in \mathbb{R},\ \sigma > 0,\ -\frac{1}{n-1} < \rho < 1$, then

$$\frac{X_{(r)} - \mu}{\sigma} \sim SUN(n-1, \boldsymbol{\theta}), \tag{3.1}$$

where

$$\boldsymbol{\theta} = \left(\sigma(1-\rho)\mathbf{J}_{n-1}, \mathbf{0}, \sigma^2(1-\rho)\{\mathbf{I}_{n-1} + \rho\mathbf{J}_{n-1}\mathbf{J}_{n-1}^T\}\right),$$

with $\mathbf{J}_{n-1} = \left(\mathbf{1}_{r-1}^T, -\mathbf{1}_{n-r}^T\right)^T$ and $\mathbf{I}_{n-1} \in \mathbb{R}^{n-1 \times n-1}$.
We use the above lemmas and equation (3.1), without loss of generality, to obtain the moment generating function, kth moment and the first moment of $NPS(\mu, \sigma, \theta)$, when $\mu = 0$ and $\sigma = 1$ in the following proposition.

Proposition 4. *If* $Y \sim NPS(0, 1, \theta)$*, then the moment generating function,* k*th moment and mean of* Y *are given by*

$$\begin{aligned}
M_Y(t) &= \sum_{n=1}^{\infty} \frac{a_n\theta^n}{C(\theta)} M_{X_{(n)}}(t) \\
&= \exp\left(\frac{1}{2}t^2\right) \sum_{n=1}^{\infty} \frac{a_n\theta^n}{C(\theta)} \times n\Phi_{n-1}(\mathbf{1}_{n-1}t; \mathbf{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}_{n-1}^T) \\
&= \exp\left(\frac{1}{2}t^2\right) E(N\Phi_{N-1}(\mathbf{1}_{N-1}t; \mathbf{I}_{N-1} + \mathbf{1}_{N-1}\mathbf{1}_{N-1}^T)), \\
E(Y^{k+1}) &= \sum_{n=1}^{\infty} \frac{a_n\theta^n}{C(\theta)} kE(Y^{k-1}) \\
&\quad + \sum_{n=1}^{\infty} \frac{a_n\theta^n}{C(\theta)} \times \frac{(n-1)\Phi_{n-2}\left(\mathbf{0}; \mathbf{I}_{n-2} + \frac{1}{2}\mathbf{1}_{n-2}\mathbf{1}_{n-2}^T\right)}{2\sqrt{\pi}\Phi_{n-1}\left(\mathbf{0}; \mathbf{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}_{n-1}^T\right)} \\
&\quad \times E\left(Z^k_{n-2, \frac{1}{\sqrt{2}}\mathbf{1}_{n-2}, \mathbf{I}_{n-2}}\right),
\end{aligned}$$

and

$$E(Y) = \frac{1}{2\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{a_n\theta^n}{C(\theta)} \times n(n-1)\Phi_{n-2}\left(\mathbf{0}; \mathbf{I}_{n-2} + \frac{1}{2}\mathbf{1}_{n-2}\mathbf{1}_{n-2}^T\right),$$

respectively.

One can derive the second moment of NPS class of distributions as

$$E(Y^2) = 1 + \frac{1}{4\sqrt{3}\pi} \sum_{n=1}^{\infty} \frac{a_n \theta^n}{C(\theta)} \times n(n-1)(n-2)\Phi_{n-3}\left(\mathbf{0}; \mathbf{I}_{n-3} + \frac{1}{3}\mathbf{1}_{n-3}\mathbf{1}_{n-3}^T\right).$$

(ii) Second method

In the following proposition, we present another formulation for calculate the kth moment around the origin of the random variable $Y \sim NPS(\mu, \sigma, \theta)$. First, we give two well-known relationship, which are necessary in the following proposition. If $\Phi(x; \mu, \sigma)$ denotes the cdf of $N(\mu, \sigma^2)$ distribution, then we have

$$\Phi(x; \mu, \sigma) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right], \tag{3.2}$$

and

$$\Phi^{-1}(t; \mu, \sigma) = \mu + \sigma\sqrt{2}\,\text{erf}^{-1}(2t-1). \tag{3.3}$$

Proposition 5. *We have*

$$E(Y^k) = \int_0^1 \left\{\mu + \sigma\sqrt{2}\,\text{erf}^{-1}\left(\frac{2C^{-1}\left(C(\theta)u\right)}{\theta} - 1\right)\right\}^k du.$$

Proof. We know that $E(Y^k) = \int_{-\infty}^{\infty} y^k f(y; \mu, \sigma, \theta)dy$. Then, substituting $f(y; \mu, \sigma, \theta)$ from (2.2) and change of variable $\Phi\left(y; \mu, \sigma\right) = t$, gives

$$E(Y^k) = \frac{\theta}{C(\theta)} \int_0^1 \left[\Phi^{-1}\left(t; \mu, \sigma\right)\right]^k C'(\theta t)dt.$$

Now, by changing the variable to $u = \frac{C(t\theta)}{C(\theta)}$, we have

$$E(Y^k) = \int_0^1 \left\{\Phi^{-1}\left(\frac{C^{-1}(C(\theta)u)}{\theta}; \mu, \sigma\right)\right\}^k du,$$

thus, the result follows from the Equation (3.3). □

4 Special cases of NPS class of distributions

In this section four important sub-models of NPS class of distributions are studied in details. These models are normal-geometric (NG), normal-Poisson (NP), normal-logarithmic (NL) and normal-binomial (NB) distributions.

4.1 Normal-geometric distribution

Using Table 1, the NPS distributions contain normal-geometric (NG) distribution when $a_n = 1$ and $C(\theta) = \frac{\theta}{1-\theta}(0 < \theta < 1)$. Using Equation (2.1), the cdf of NG is given by

$$F(y;\mu,\sigma,\theta) = \frac{(1-\theta)\Phi(\frac{y-\mu}{\sigma})}{1-\theta\Phi(\frac{y-\mu}{\sigma})}. \tag{4.1}$$

The pdf and hazard rate function of NG distribution are

$$f(y;\mu,\sigma,\theta) = \frac{(1-\theta)\phi(\frac{y-\mu}{\sigma})}{\sigma(1-\theta\Phi(\frac{y-\mu}{\sigma}))^2}, \tag{4.2}$$

and

$$h(y;\mu,\sigma,\theta) = \frac{(1-\theta)\phi(\frac{y-\mu}{\sigma})}{\sigma(1-\Phi(\frac{y-\mu}{\sigma}))(1-\theta\Phi(\frac{y-\mu}{\sigma}))}, \tag{4.3}$$

respectively, where $y \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma > 0$ and $0 < \theta < 1$. We use the notation $Y \sim NG(\mu,\sigma,\theta)$ when the random variable Y has NG distribution with location μ, scale σ and shape parameter θ.

Remark 1. *Even when $\theta \leq 0$, Equation (4.2) is also a density function. We can then define the NG distribution by Equation (4.2) for any $\theta < 1$. Some special sub-models of the NG distribution are obtained as follows: If $\theta = 0$, we have the normal distribution. When $\theta \to 1^-$, the NG distribution tends to a distribution degenerated at zero. Hence, the parameter θ can be interpreted as a concentration parameter.*

Figures 1 and 2 show the NG density function and hazard rate function for selected values $\theta < 1$ where $\mu = 0$ and $\sigma = 1$.

Theorem 1. *Suppose that $Y_1 \sim NG(0,1,\theta_1)$ and $Y_2 \sim NG(0,1,\theta_2)$. If $\theta_1 > \theta_2$, then $Y_2 <_{LR} Y_1$.*

Proof. The logarithm of the likelihood ratio is given by

$$\upsilon(y) = \log\frac{1-\theta_1}{1-\theta_2} + 2\log(1-\theta_2\Phi(y)) - 2\log(1-\theta_1\Phi(y)),$$

which is an increasing function of y if $\theta_1 > \theta_2$, since

$$\upsilon'(y) = \frac{2(\theta_1-\theta_2)\phi(y)}{(1-\theta_2\Phi(y))(1-\theta_1\Phi(y))} > 0,$$

for all y. Therefore, the NG has the likelihood ratio ordering, which implies it has the failure rate ordering as well as the stochastic ordering and the mean residual life ordering. □

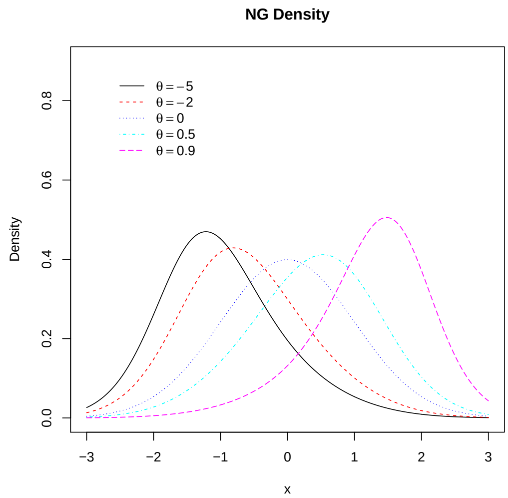

Figure 1: Plots of density function of NG distribution for selected parameter values $\theta < 1$, $\mu = 0$ and $\sigma = 1$.

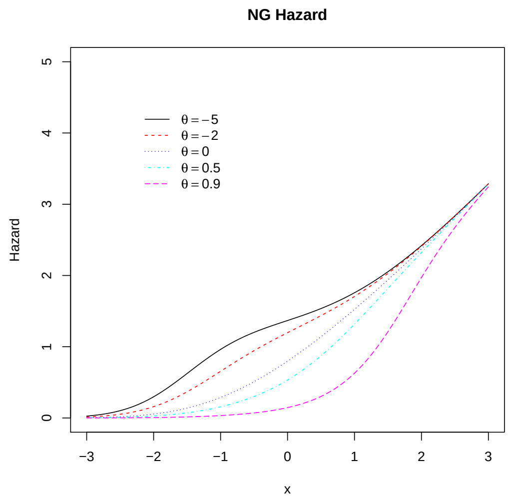

Figure 2: Plots of hazard rate function of NG distribution for selected parameter values $\theta < 1$, $\mu = 0$ and $\sigma = 1$.

Proposition 6. *The moment generating function, mean and second moment of NG are*

Table 2: The first four moments, variance, skewness and kurtusis of NG distribution for $\mu = 0,\ \sigma = 1$

	$\theta=-5$	$\theta=-2$	$\theta=-0.5$	$\theta=0$	$\theta=0.3$	$\theta=0.5$	$\theta=0.8$	$\theta=0.9$
$E(Y)$	-0.9841	-0.6134	-0.2284	0	0.2010	0.3894	0.8884	1.2445
$E(Y^2)$	1.8465	1.3270	1.0452	1	1.0350	1.1315	1.6887	2.3609
$E(Y^3)$	-3.1487	-1.6981	-0.5795	0	0.5083	1.0155	2.7254	4.5206
$E(Y^4)$	7.2110	4.4974	3.1974	3	3.1526	3.5829	6.3424	10.313
VAR	0.8781	0.9508	0.9930	1	0.9946	0.9799	0.8995	0.8123
SK	0.4821	0.3046	0.1141	0	-0.1004	-0.1942	-0.4371	-0.5999
KUR	3.5440	3.2104	3.0291	3	3.0225	3.0846	3.4429	3.8702

given by

$$\begin{aligned} M_Y(t) &= \exp\left(\frac{1}{2}t^2\right)\sum_{n=1}^{\infty} n(1-\theta)\theta^{n-1}\Phi_{n-1}(\mathbf{1}_{n-1}t;\mathbf{I}_{n-1}+\mathbf{1}_{n-1}\mathbf{1}_{n-1}^T), \\ E(Y) &= \frac{1}{2\sqrt{\pi}}\sum_{n=1}^{\infty} n(n-1)(1-\theta)\theta^{n-1}\Phi_{n-2}\left(\mathbf{0};\mathbf{I}_{n-2}+\frac{1}{2}\mathbf{1}_{n-2}\mathbf{1}_{n-2}^T\right), \end{aligned}$$

and

$$E(Y^2) = 1 + \frac{1}{4\sqrt{3}\pi}\sum_{n=1}^{\infty} n(n-1)(n-2)(1-\theta)\theta^{n-1}\Phi_{n-3}\left(\mathbf{0};\mathbf{I}_{n-3}+\frac{1}{3}\mathbf{1}_{n-3}\mathbf{1}_{n-3}^T\right),$$

respectively.

Table 2 gives the first four moments, variance, skewness and kurtusis of the $NG(0,1,\theta)$ for different values $\theta < 1$.

Figure 3 shows the skewness and kurtusis plot of the $NG(0,1,\theta)$ for different values $\theta < 1$ with $\mu = 0$ and $\sigma = 1$.

4.2 Normal-Poisson distribution

The normal-Poisson (NP) distribution is obtained when $a_n = \frac{1}{n!}$ and $C(\theta) = e^{\theta} - 1$. The cdf, pdf and hazard rate function of NP distribution are given by

$$F(y;\mu,\sigma,\theta) = \frac{e^{\theta\Phi(\frac{y-\mu}{\sigma})} - 1}{e^{\theta} - 1},$$

$$f(y;\mu,\sigma,\theta) = \frac{\theta}{\sigma}\frac{\phi(\frac{y-\mu}{\sigma})e^{\theta\Phi(\frac{y-\mu}{\sigma})}}{e^{\theta} - 1}, \tag{4.4}$$

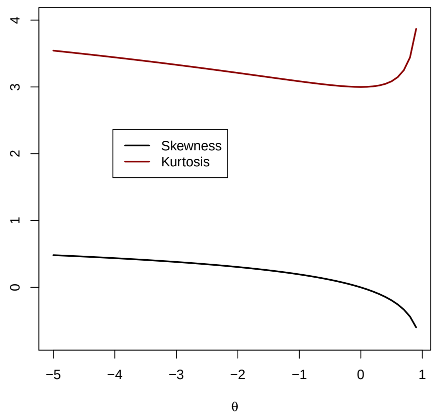

Figure 3: Plots of skewness and kurtosis of NG distribution for selected parameter values $\theta < 1$ with $\mu = 0,\ \sigma = 1$.

and

$$h(y;\mu,\sigma,\theta) = \frac{\theta}{\sigma}\frac{\phi(\frac{y-\mu}{\sigma})e^{\theta\Phi(\frac{y-\mu}{\sigma})}}{e^{\theta} - e^{\theta\Phi(\frac{y-\mu}{\sigma})}},$$

respectively, where $y \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma > 0$ and $\theta \in (0,\infty)$. We use the notation $Y \sim NP(\mu,\sigma,\theta)$ when the random variable Y has NP distribution with location μ, scale σ and shape parameter θ.

Remark 2. *Even when $\theta < 0$, Equation (4.4) is also a density function. We can then define the NP distribution by Equation (4.4) for any $\theta \in \mathbb{R}$-$\{0\}$.*

Figures 4 and 5 show the NP density function and hazard rate function for selected values θ where $\mu = 0$ and $\sigma = 1$.

Theorem 2. *Suppose that $Y_1 \sim NP(\mu,\sigma,\theta_1)$ and $Y_2 \sim NP(\mu,\sigma,\theta_2)$. If $\theta_1 > \theta_2$, then $Y_2 <_{LR} Y_1$.*

Proof. The proof is similar to the proof of Theorem 1 and is omitted. □

Proposition 7. *The moment generating function, mean and second central moment of NP are given by*

$$M_Y(t) = \exp\left(\frac{1}{2}t^2\right)\sum_{n=1}^{\infty}\frac{\theta^n}{n!(e^{\theta}-1)} \times n\Phi_{n-1}(\mathbf{1}_{n-1}t;\mathbf{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}_{n-1}^T),$$

$$E(Y) = \frac{1}{2\sqrt{\pi}}\sum_{n=1}^{\infty}\frac{n(n-1)\theta^n}{n!(e^{\theta}-1)}\Phi_{n-2}\left(\mathbf{0};\mathbf{I}_{n-2} + \frac{1}{2}\mathbf{1}_{n-2}\mathbf{1}_{n-2}^T\right),$$

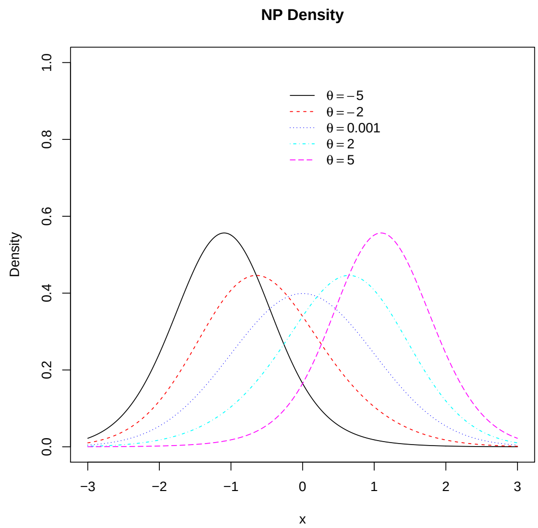

Figure 4: Plots of density function of NP distribution for selected parameter values θ with $\mu = 0$ and $\sigma = 1$.

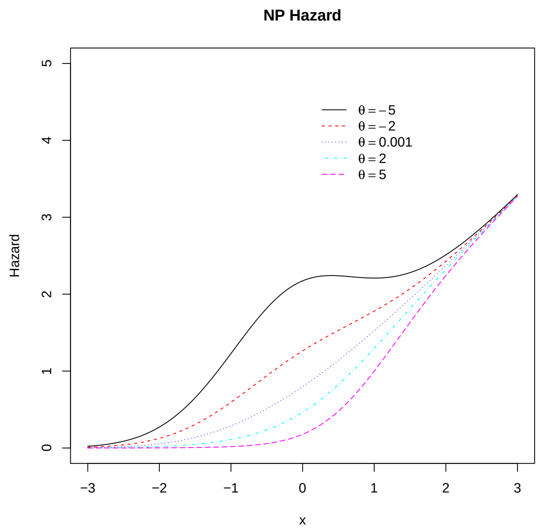

Figure 5: Plots of hazard rate function of NP distribution for selected parameter values θ with $\mu = 0$ and $\sigma = 1$.

$$\begin{aligned} E(Y^2) &= 1 + \left[\frac{1}{4\sqrt{3}\pi} \sum_{n=1}^{\infty} \frac{\theta^n}{n!(e^{\theta}-1)} \times n(n-1)(n-2)\right] \\ &\quad \times \Phi_{n-3}\left(\mathbf{0}; \mathbf{I}_{n-3} + \frac{1}{3}\mathbf{1}_{n-3}\mathbf{1}_{n-3}^T\right). \end{aligned}$$

Table 3 gives the first four moments, variance, skewness and kurtusis of the $NP(0,1,\theta)$ for different values θ. Figure 6 shows the skewness and kurtusis plot of the $NP(0,1,\theta)$ for different values θ.

Table 3: The first four moments, variance, skewness and kurtusis of NP distribution for $\mu = 0,\ \sigma = 1$.

	$\theta = 0.01$	$\theta = 0.3$	$\theta = 0.5$	$\theta = 0.8$	$\theta = 1$	$\theta = 3$	$\theta = 6$	$\theta = 10$
$E(Y)$	0.0028	0.0845	0.1405	0.2236	0.2781	0.7541	1.1997	1.5045
$E(Y^2)$	1.0000	1.0041	1.0114	1.0290	1.0450	1.3477	1.9673	2.6533
$E(Y^3)$	0.0071	0.2114	0.3520	0.5617	0.7003	2.0013	3.5904	5.2127
$E(Y^4)$	3.0000	3.0179	3.0495	3.1259	3.1954	4.5372	7.4821	11.2262
VAR	1.0000	0.9970	0.9917	0.9790	0.9677	0.7790	0.5279	0.3898
SK	-0.0014	-0.0421	-0.0697	-0.1097	-0.1349	-0.2764	-0.0956	0.1973
KUR	3.0000	3.0074	3.0204	3.0515	3.0792	3.5076	3.6846	3.4236

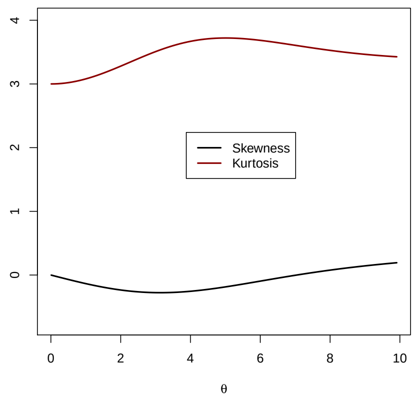

Figure 6: Plots of skewness and kurtusis of NP distribution for selected parameter values θ.

4.3 Normal-binomial distribution

When $a_n = \binom{m}{n}$ and $C(\theta) = (\theta+1)^m - 1(\theta > 0)$, where m $(n \le m)$ is the number of replicas, we obtain the normal-binomial (NB) distribution with cdf

$$F(y;\mu,\sigma,\theta) = \frac{(\theta\Phi(\frac{y-\mu}{\sigma}) + 1)^m - 1}{(\theta+1)^m - 1}.$$

The pdf and hazard rate function are

$$f(y;\mu,\sigma,\theta) = \frac{\theta}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\frac{m(\theta\Phi(\frac{y-\mu}{\sigma})+1)^{m-1}}{(\theta+1)^m-1}, \tag{4.5}$$

and

$$h(y;\mu,\sigma,\theta) = \frac{\theta}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\frac{m(\theta\Phi\left(\frac{y-\mu}{\sigma}\right)+1)^{m-1}}{(\theta+1)^m-(\theta\Phi\left(\frac{y-\mu}{\sigma}\right)+1)^m},$$

respectively, where $y \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma > 0$ and $\theta \in (0,\infty)$. We use the notation $Y \sim NB(\mu,\sigma,\theta)$ when the random variable Y has NB distribution with location μ, scale σ and shape parameter θ. Figures 7 and 8 show the NB density and hazard rate function for selected values θ where $\mu = 0$ and $\sigma = 1$.

Proposition 8. *The moment generating function, mean and second central moment of NB are given by*

$$M_Y(t) = \exp\left(\frac{1}{2}t^2\right)\sum_{n=1}^{\infty}\binom{m}{n}\frac{\theta^n}{(\theta+1)^m-1}\times n\Phi_{n-1}(\mathbf{1}_{n-1}t;\mathbf{I}_{n-1}+\mathbf{1}_{n-1}\mathbf{1}_{n-1}^T),$$

$$\begin{aligned} E(Y) &= \frac{1}{2\sqrt{\pi}}\sum_{n=1}^{\infty}\binom{m}{n}\frac{\theta^n}{(\theta+1)^m-1}n(n-1) \\ &\quad \times\Phi_{n-2}\left(\mathbf{0};\mathbf{I}_{n-2}+\frac{1}{2}\mathbf{1}_{n-2}\mathbf{1}_{n-2}^T\right), \end{aligned}$$

$$\begin{aligned} E(Y^2) &= 1+\Big[\frac{1}{4\sqrt{3}\pi}\sum_{n=1}^{\infty}\binom{m}{n}\frac{n(n-1)(n-2)\theta^n}{(\theta+1)^m-1}\Big] \\ &\quad \times\Phi_{n-3}\left(\mathbf{0};\mathbf{I}_{n-3}+\frac{1}{3}\mathbf{1}_{n-3}\mathbf{1}_{n-3}^T\right). \end{aligned}$$

4.4 Normal-logarithmic distribution

When $a_n = \frac{1}{n}$ and $C(\theta) = -\log(1-\theta)(0 < \theta < 1)$, we obtain the normal-logarithmic (NL) distribution with cdf

$$F(y;\mu,\sigma,\theta) = \frac{\log(1-\theta\Phi\left(\frac{y-\mu}{\sigma}\right))}{\log(1-\theta)}.$$

The pdf and hazard rate function are given by

$$f(y;\mu,\sigma,\theta) = \frac{\frac{\theta}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)}{(\theta\Phi\left(\frac{y-\mu}{\sigma}\right)-1)\log(1-\theta)}, \tag{4.6}$$

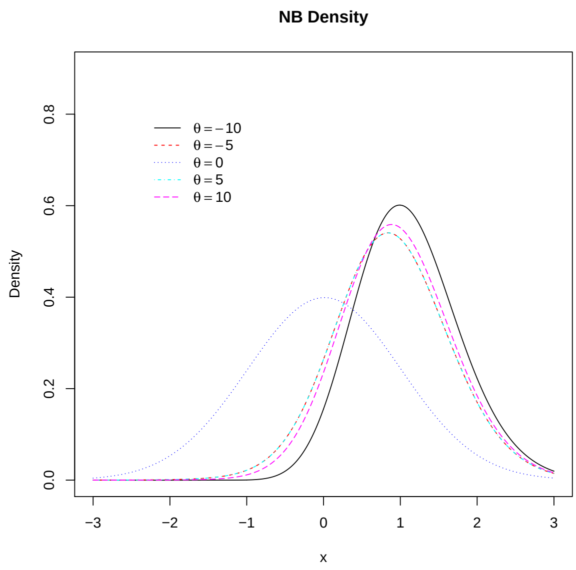

Figure 7: Plots of density function of NB distribution for selected parameter values θ with $\mu = 0$ and $\sigma = 1$.

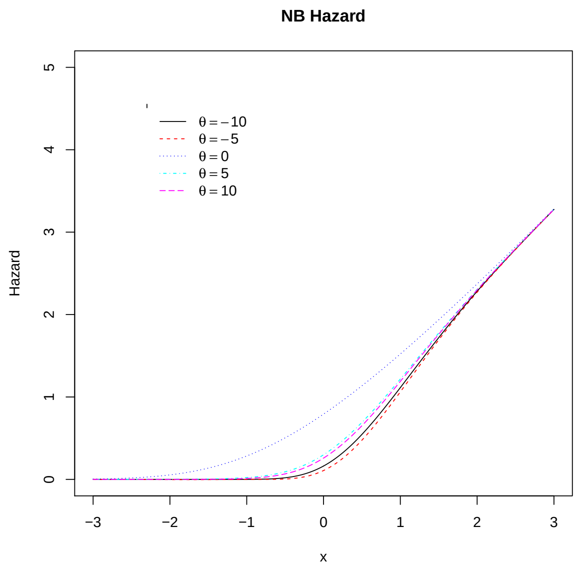

Figure 8: Plots of hazard rate function of NB distribution for selected parameter values θ with $\mu = 0$ and $\sigma = 1$.

and

$$h(y;\mu,\sigma,\theta) = \frac{\frac{\theta}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)}{\left(\theta\Phi\left(\frac{y-\mu}{\sigma}\right)-1\right)\log\frac{1-\theta}{1-\theta\Phi\left(\frac{y-\mu}{\sigma}\right)}},$$

respectively, where $y \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma > 0$ and $\theta \in (0,1)$. We use the notation $Y \sim NL(\mu, \sigma, \theta)$ when the random variable Y has NL distribution with location μ, scale σ and shape parameter θ.

Remark 3. *Even when $\theta < 0$, Equation (4.6) is also a density function. We can then define the NL distribution by Equation (4.6) for any $\theta \in (-\infty, 0) \cup (0,1)$.*

Figures 9 and 10 show the NL density and hazard rate functions for selected values θ where $\mu = 0$ and $\sigma = 1$.

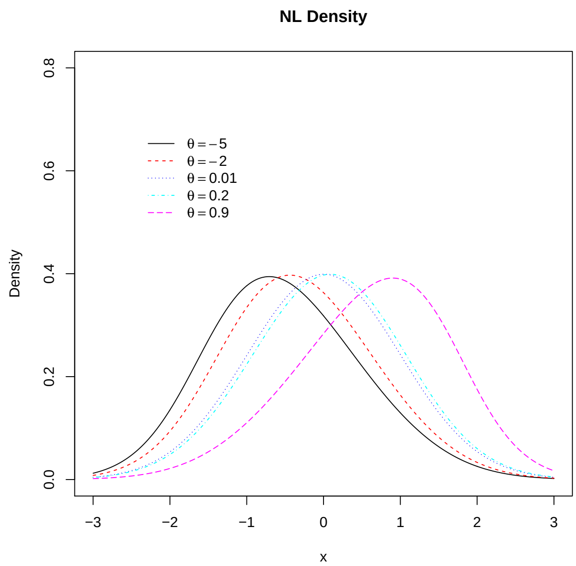

Figure 9: Plots of density function of NL distribution for selected parameter values θ with $\mu = 0$ and$\sigma = 1$.

Proposition 9. *The moment generating function, mean and second central moment of NL are given by*

$$M_Y(t) = \exp(\frac{1}{2}t^2) \sum_{n=1}^{\infty} \frac{\theta^n}{\log(1-\theta)} \times \Phi_{n-1}(\mathbf{1}_{n-1}t; \mathbf{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}_{n-1}^T),$$

$$E(Y) = -\frac{1}{2\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{\theta^n}{n\log(1-\theta)} (n-1)\Phi_{n-2}\left(\mathbf{0}; \mathbf{I}_{n-2} + \frac{1}{2}\mathbf{1}_{n-2}\mathbf{1}_{n-2}^T\right),$$

$$E(Y^2) = 1 - \left[\frac{1}{4\sqrt{3}\pi} \sum_{n=1}^{\infty} \frac{n(n-1)(n-2)\theta^n}{n\log(1-\theta)}\right] \times \Phi_{n-3}\left(\mathbf{0}; \mathbf{I}_{n-3} + \frac{1}{3}\mathbf{1}_{n-3}\mathbf{1}_{n-3}^T\right),$$

respectively.

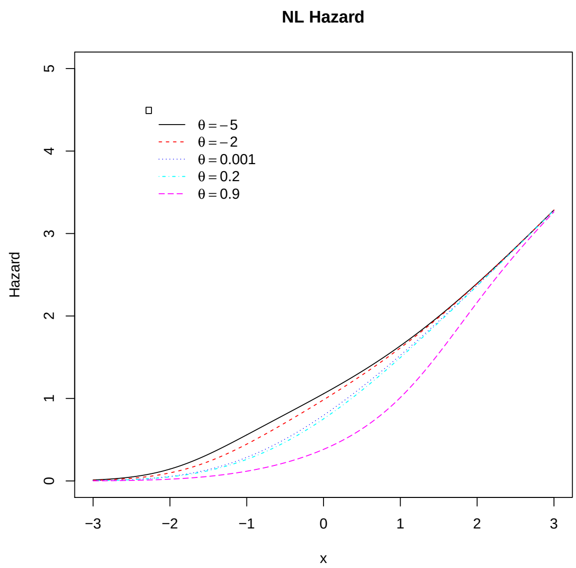

Figure 10: Plots of hazard rate function of NL distribution for selected parameter values θ with $\mu = 0$ and$\sigma = 1$.

5 Properties of sub-models of NPS distristributions

In this section we present some additional and useful properties of the sub-models of NPS distribution.

Proposition 10. *For NG, NB and NL distributions we have*

$$F(y;0,1,\theta) = 1 - F\left(-y;0,1,\frac{\theta}{\theta-1}\right),$$

and for NP distribution we have

$$F(y;0,1,\theta) = 1 - F\left(-y;0,1,-\theta\right),$$

Proof. We shall prove for NG distribution. Proofs of other distributions are similar. From (4.1), it can be found that

$$\begin{aligned} 1 - F\left(-y;0,1,\frac{\theta}{\theta-1}\right) &= 1 + \frac{\frac{1}{\theta-1}\Phi(-y)}{1 - \frac{\theta}{\theta-1}\Phi(-y)} = 1 + \frac{1-\Phi(y)}{\theta\Phi(y)-1} \\ &= \frac{(1-\theta)\Phi(y)}{1-\theta\Phi(y)} = F(y;0,1,\theta), \end{aligned}$$

and hence the proof is compeleted. □

In the following proposition we give approximations for first and second moments around the origin of NG, NP and NB distributions.

Proposition 11. *We have*
(i) If $Y \sim NG(\mu,\sigma,\theta)$, then $E(Y)$ and $E(Y^2)$ are approximated by

$$E(Y) \simeq \frac{1}{2\theta^2}\left\{2\theta^2\mu + \sqrt{2\pi}\sigma\left(2\log(1-\theta)(1-\theta) + 2\theta - \theta^2\right)\right\},$$

$$\begin{aligned} E(Y^2) \;\simeq\; & \frac{1}{4\theta(\theta-1)^2}\Big\{\theta^3\left(\pi\sigma^2 + 2\mu^2 - 2\sqrt{2}\sqrt{\pi}\sigma\mu\right) \\ & +28\sigma\theta^2\left(\sqrt{2\pi}\mu - 2\pi\sigma\right) + 16\pi\theta\sigma^2 \\ & +8\sigma\log(1-\theta)\left(\sqrt{2\pi}\theta\mu(1-\theta) + (\theta(\theta-3)+2)\pi\sigma\right)\Big\}, \end{aligned}$$

(ii) If $Y \sim NP(\mu,\sigma,\theta)$, then $E(Y)$ and $E(Y^2)$ are approximated by

$$E(Y) \simeq \frac{1}{2\theta\sigma\left(e^\theta - 1\right)}\left\{(2 + \theta\sigma\sqrt{2\pi} - 2\theta\mu + (2\theta\mu + (\theta-2)\sigma\sqrt{2\pi})e^\theta\right\},$$

$$\begin{aligned} E(Y^2) \;\simeq\; & \frac{1}{2\theta^2\left(e^\theta - 1\right)}\Big[\sigma\mu\theta\left(4\sqrt{2\pi} + 2\sqrt{2\pi}\theta\right) - 8\pi\sigma^2 - \theta^2\left(\pi\sigma^2 + 2\mu^2\right) \\ & -4\pi\theta\sigma^2 + \Big(2\theta^2\mu^2 + 8\pi\sigma^2 + \pi\theta^2\sigma^2 + 2\sqrt{2\pi}\theta^2\sigma\mu \\ & -4\sqrt{2\pi}\theta\sigma\mu\Big)e^\theta\Big]. \end{aligned}$$

(iii) If $Y \sim NB(\mu,\sigma,\theta)$, then $E(Y)$ and $E(Y^2)$ are approximated by

$$\begin{aligned} E(Y) \;\simeq\; & \frac{1}{2\theta\left((\theta+1)^m - 1\right)(m+1)}\Big\{2\sqrt{2\pi}\sigma + \left(\sqrt{2\pi}\sigma - 2\mu\right)(1+m)\theta \\ & -\Big[\left(2\sqrt{2\pi}\sigma + \sqrt{2\pi}\sigma - 2\mu\right)\theta(1+m) - 2\sqrt{2\pi}m\sigma(\theta+1)\Big](\theta+1)^m\Big\}. \end{aligned}$$

$$\begin{aligned} E(Y^2) \;\simeq\; & \frac{1}{\theta^2(\theta+1)^m - 1(m+2)(m+1)}\Big\{(-2\mu^2 + \sqrt{2\pi}m^2\sigma\mu - \frac{3}{2}\pi m\sigma^2 - \pi\sigma^2 \\ & -3m\mu^2 + 3\sqrt{2\pi}m\sigma\mu - m^2\mu^2 - \frac{1}{2}\pi m^2\sigma^2 + 2\sqrt{2\pi}\sigma\mu)\theta^2 \\ & +(4\sqrt{2\pi}\sigma\mu + 2\sqrt{2\pi}m\sigma\mu - 2\pi m\sigma^2 - 4\pi\sigma^2)\theta + [(2\mu^2 + 3m\mu^2 + \pi\sigma^2 \\ & +\sqrt{2\pi}m\sigma\mu + \sqrt{2\pi}m^2\sigma\mu - 2\sqrt{2\pi}\sigma\mu - \frac{1}{2}\pi m\sigma^2 + \frac{1}{2}\pi m^2\sigma^2 + m^2\mu^2)\theta^2 \\ & +4\pi\sigma^2 + (4\pi\sigma^2 - 2\pi m\sigma^2 - 4\sqrt{2\pi}\sigma\mu - 2\sqrt{2\pi}m\sigma\mu)\theta]\,(\theta+1)^m\Big\}. \end{aligned}$$

Proof. For (i), firstly note that

$$E(Y) = \int_{-\infty}^{\infty} \frac{(1-\theta)y\phi(y;\mu,\sigma,\theta)}{(1-\theta\Phi(y;\mu,\sigma,\theta))^2} dy.$$

After the change of variable $u = 1-\theta\Phi(y;\mu,\sigma,\theta)$, we obtain $y = \Phi^{-1}\left(\frac{1-u}{\theta};\mu,\sigma\right) = \mu + \sqrt{2}\sigma\,\mathrm{erf}^{-1}\left(2\left(\frac{1-u}{\theta}\right)-1\right)$. Now because $\mathrm{erf}^{-1}(z) = z\frac{\sqrt{\pi}}{2} + O(z^3)$, we can write $\mathrm{erf}^{-1}(z) \simeq z\frac{\sqrt{\pi}}{2}$. Therefore we have

$$\mu_1 = -\frac{1-\theta}{\theta}\int_1^{1-\theta} \frac{\mu + \sqrt{\frac{\pi}{2}}\sigma\left(2\left(\frac{1-u}{\theta}\right)-1\right)}{u^2} du.$$

the result is obtained by solving the integral. Finally, $E(Y^2)$ is derived in the same manner after simple computation. Parts (ii) and (iii) follow in a same way. □

6 Estimation and inference

In this section, we discuss the estimation of the parameters of NPS class of distributions. let $Y_1, Y_2, ..., Y_n$ be a random sample with observed values $y_1, y_2, ..., y_n$ from a $NPS(\mu,\sigma,\theta)$ and $\mathbf{\Psi} = (\mu,\sigma,\theta)^T$ be a parameter vector. The total log-likelihood function is given by

$$\begin{aligned} l_n &= l_n(\mathbf{\Psi};\mathbf{y}) = n\log(\theta) - n\log(\sigma) - \frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^n t_i^2 \\ &+\sum_{i=1}^n \log(C'(\theta\Phi(t_i)) - n\log(C(\theta)), \end{aligned}$$

where $t_i = \frac{y_i-\mu}{\sigma}$. The maximum likelihood estimation (MLE) of $\mathbf{\Psi}$, say $\widehat{\mathbf{\Psi}}$, is obtained by solving the nonlinear system of equations $\left(\frac{\partial l_n}{\partial\mu}, \frac{\partial l_n}{\partial\sigma}, \frac{\partial l_n}{\partial\theta}\right)^T = 0$, where

$$\begin{aligned} \frac{\partial l_n}{\partial\mu} &= \frac{1}{\sigma}\sum_{i=1}^n t_i - \frac{\theta}{\sigma}\sum_{i=1}^n \frac{\phi(t_i)C''(\theta\Phi(t_i))}{C'(\theta\Phi(t_i))}, \\ \frac{\partial l_n}{\partial\sigma} &= -\frac{n}{\sigma} + \frac{1}{\sigma}\sum_{i=1}^n t_i^2 - \frac{\theta}{\sigma}\sum_{i=1}^n \frac{t_i\phi(t_i)C''(\theta\Phi(t_i))}{C'(\theta\Phi(t_i))}, \\ \frac{\partial l_n}{\partial\theta} &= \frac{n}{\theta} + \sum_{i=1}^n \frac{\Phi(t_i)C''(\theta\Phi(t_i))}{C'(\theta\Phi(t_i))} - \frac{nC'(\theta)}{C(\theta)}. \end{aligned}$$

The solution of this nonlinear system of equation has not a closed form. The observed information matrix is obtained for approximate confidence intervals and hypothesis tests of the vector. The 3×3 observed information matrix is given by

$$I_n(\mathbf{\Psi}) = -\begin{bmatrix} I_{\mu\mu} & I_{\mu\sigma} & I_{\mu\theta} \\ I_{\mu\sigma} & I_{\sigma\sigma} & I_{\sigma\theta} \\ I_{\mu\theta} & I_{\sigma\theta} & I_{\theta\theta} \end{bmatrix},$$

where

$$
\begin{aligned}
I_{\mu\mu} &= -\frac{n}{\sigma^2} - \frac{\theta}{\sigma^2}\sum_{i=1}^{n}\frac{\left[t_i\phi(t_i)C^{''}(\theta\Phi(t_i)) - \theta C^{'''}(\theta\Phi(t_i)\phi^2(t_i)\right]C^{'}(\theta\Phi(t_i))}{\left(C^{'}(\theta\Phi(t_i))\right)^2} \\
&\quad -\frac{\theta}{\sigma^2}\sum_{i=1}^{n}\frac{\theta\left(C^{''}(\theta\Phi(t_i))\right)^2\phi^2(t_i)}{\left(C^{'}(\theta\Phi(t_i))\right)^2}, \\
I_{\mu\sigma} &= -\frac{2}{\sigma^2}\sum_{i=1}^{n}t_i + \frac{\theta}{\sigma^2}\sum_{i=1}^{n}\frac{\phi(t_i)C^{''}(\theta\Phi(t_i))}{C^{'}(\theta\Phi(t_i))} - \frac{\theta}{\sigma^2}\sum_{i=1}^{n}\frac{\theta t_i\phi^2(t_i)\left(C^{''}(\theta\Phi(t_i)\right)^2}{\left(C^{'}(\theta\Phi(t_i))\right)^2} \\
&\quad -\frac{\theta}{\sigma^2}\sum_{i=1}^{n}\frac{\left[t_i^2\phi(t_i)C^{''}(\theta\Phi(t_i)) - \theta t_i\phi^2(t_i)C^{'''}(\theta\Phi(t_i))\right]C^{'}(\theta\Phi(t_i))}{\left(C^{'}(\theta\Phi(t_i))\right)^2}, \\
I_{\mu\theta} &= -\frac{1}{\sigma}\sum_{i=1}^{n}\frac{\phi(t_i)C^{''}(\theta\Phi(t_i))}{C^{'}(\theta\Phi(t_i))} \\
&\quad -\frac{\theta}{\sigma}\sum_{i=1}^{n}\frac{\Phi(t_i)\phi(t_i)C^{'''}(\theta\Phi(t_i))C^{'}(\theta\Phi(t_i)) - \Phi(t_i)\phi(t_i)\left(C^{''}(\theta\Phi(t_i)\right)^2}{\left(C^{'}(\theta\Phi(t_i)\right)^2},
\end{aligned}
$$

$$
\begin{aligned}
I_{\sigma\sigma} &= \frac{n}{\sigma^2} - \frac{3}{\sigma^2}\sum_{i=1}^{n}t_i^2 + \frac{\theta}{\sigma^2}\sum_{i=1}^{n}\frac{t_i\phi(t_i)C^{''}(\theta\Phi(t_i))}{C^{'}(\theta\Phi(t_i))} \\
&\quad +\frac{\theta}{\sigma^2}\sum_{i=1}^{n}\frac{\left[\left(t_i^3\phi(t_i) - t_i\phi(t_i)\right)C^{''}(\theta\Phi(t_i) - \theta t_i^2\phi^2(t_i)C^{'''}(\theta\Phi(t_i)\right]C^{'}(\theta\Phi(t_i))}{\left(C^{'}(\theta\Phi(t_i))\right)^2} \\
&\quad -\frac{\theta^2}{\sigma^2}\sum_{i=1}^{n}\frac{t_i^2\phi^2(t_i)\left(C^{''}(\theta\Phi(t_i)\right)^2}{\left(C^{'}(\theta\Phi(t_i))\right)^2}, \\
I_{\sigma\theta} &= -\frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{t_i\phi(t_i)C^{''}(\theta\Phi(t_i))}{C^{'}(\theta\Phi(t_i))} \\
&\quad -\frac{\theta}{\sigma}\sum_{i=1}^{n}\frac{t_i\phi(t_i)\Phi(t_i)C^{'}(\theta\Phi(t_i))C^{'''}(\theta\Phi(t_i) - t_i\phi(t_i)\Phi(t_i)\left(C^{''}(\theta\Phi(t_i)\right)^2}{\left(C^{'}(\theta\Phi(t_i))\right)^2}, \\
I_{\theta\theta} &= -\frac{n}{\theta^2} + \sum_{i=1}^{n}\frac{\Phi^2(t_i)C^{'''}(\theta\Phi(t_i))C^{'}(\theta\Phi(t_i)) - \Phi^2(t_i)\left(C^{''}(\theta\Phi(t_i))\right)^2}{\left(C^{'}(\theta\Phi(t_i))\right)^2} \\
&\quad -\frac{nC^{''}(\theta)}{C(\theta)} + \frac{n\left(C^{'}(\theta)\right)^2}{\left(C(\theta)\right)^2}.
\end{aligned}
$$

It is well-known that under regularity conditions, the asymptotic distribution of $\sqrt{n}\left(\widehat{\boldsymbol{\Psi}} - \boldsymbol{\Psi}\right)$

is $N_3(0, J_n(\mathbf{\Psi})^{-1})$, where $J_n(\mathbf{\Psi}) = \lim_{n\to\infty} = n^{-1}I_n(\mathbf{\Psi})$. Therefore, an $100(1-\gamma)$ asymptotic confidence interval for each parameter $\mathbf{\Psi}_r$ is given by

$$ACI_r = \left(\widehat{\mathbf{\Psi}}_r - Z_{\gamma/2}\sqrt{\widehat{I}^{rr}}, \widehat{\mathbf{\Psi}}_r + Z_{\gamma/2}\sqrt{\widehat{I}^{rr}}\right),$$

where $\widehat{I}^{rr}$ is the (r,r) diagonal element of $I_n(\widehat{\mathbf{\Psi}})^{-1}$ for $r = 1, 2, 3$ and $Z_{\gamma/2}$ is the quantile $1-\gamma/2$ of the standard normal distribution.

7 EM-algorithm

The EM-algorithm is one such elaborate technique. The EM-algorithm is a general method of finding the maximum likelihood estimate of the parameters of an underlying distribution from a given data set when the data is incomplete or has missing values. There are two main applications of the EM-algorithm. The first occurs when the data indeed has missing values, due to problems with or limitations of the observation process. The second occurs when optimizing the likelihood function is analytically intractable but when the likelihood function can be simplified by assuming the existence of values for additional but missing (or hidden) parameters.
We define a hypothetical complete-data distribution with a joint probability density function in the form

$$g(z, y; \mathbf{\Psi}) = \frac{a_z\theta^z}{\sigma C(\theta)} z\phi\left(\frac{y-\mu}{\sigma}\right)\Phi^{z-1}\left(\frac{y-\mu}{\sigma}\right),$$

where $\sigma > 0$, $\theta \in (0, s)$, $y \in \mathbb{R}$ and $z \in \mathbb{N}$. The probability density function of Z given $Y = y$ is given by

$$g(z \mid y) = \frac{g(z, y; \mathbf{\Psi})}{f(y)} = \frac{a_z\theta^{z-1} z\Phi^{z-1}\left(\frac{y-\mu}{\sigma}\right)}{C'\left(\theta\Phi\left(\frac{y-\mu}{\sigma}\right)\right)}.$$

After some simple calculation we have

$$E(Z \mid Y = y) = 1 + \frac{\theta\Phi\left(\frac{y-\mu}{\sigma}\right) C''\left(\theta\Phi\left(\frac{y-\mu}{\sigma}\right)\right)}{C'\left(\theta\Phi\left(\frac{y-\mu}{\sigma}\right)\right)}.$$

The complete-data log- likelihood has the form

$$\begin{aligned} l_n^*\left(\mathbf{y}, \mathbf{z}; \mu, \sigma, \theta\right) \quad \propto \quad & \sum_{i=1}^n z_i \log\theta - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - \mu)^2 \\ & + \sum_{i=1}^n (z_i - 1)\log\left(\Phi\left(\frac{y_i-\mu}{\sigma}\right)\right) - n\log(C(\theta)). \end{aligned}$$

The components of the score function $U_c(\mathbf{y},\mathbf{z};\mathbf{\Psi}) = \left(\frac{\partial l_n^*}{\partial \mu}, \frac{\partial l_n^*}{\partial \sigma}, \frac{\partial l_n^*}{\partial \theta}\right)$, are

$$
\begin{aligned}
\frac{\partial l_n^*}{\partial \mu} &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i-\mu) - \frac{1}{\sigma}\sum_{i=1}^{n}(z_i-1)\frac{\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi\left(\frac{y_i-\mu}{\sigma}\right)},\\
\frac{\partial l_n^*}{\partial \sigma} &= -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(y_i-\mu)^2 - \frac{1}{\sigma^2}\sum_{i=1}^{n}(z_i-1)\frac{(y_i-\mu)\,\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi\left(\frac{y_i-\mu}{\sigma}\right)},\\
\frac{\partial l_n^*}{\partial \theta} &= \frac{1}{\theta}\sum_{i=1}^{n} z_i - n\frac{C'(\theta)}{C(\theta)}.
\end{aligned}
$$

The maximum likelihood estimates can be obtained from the iterative algorithm given by

$$
\begin{aligned}
&\frac{1}{\widehat{\sigma}^{(h)}}\sum_{i=1}^{n}\left(y_i-\widehat{\mu}^{(h+1)}\right) - \sum_{i=1}^{n}\left(\widehat{z}_i^{(h)}-1\right)\frac{\phi\left(\frac{y_i-\widehat{\mu}^{(h+1)}}{\widehat{\sigma}^{(h)}}\right)}{\Phi\left(\frac{y_i-\widehat{\mu}^{(h+1)}}{\widehat{\sigma}^{(h)}}\right)} = 0,\\
&n - \frac{1}{\left(\widehat{\sigma}^{(h+1)}\right)^2}\sum_{i=1}^{n}\left(y_i-\widehat{\mu}^{(h)}\right)^2 + \frac{1}{\widehat{\sigma}^{(h+1)}}\sum_{i=1}^{n}\left(\widehat{z}_i^{(h)}-1\right)\frac{\left(y_i-\widehat{\mu}^{(h)}\right)\phi\left(\frac{y_i-\widehat{\mu}^{(h)}}{\widehat{\sigma}^{(h+1)}}\right)}{\Phi\left(\frac{y_i-\widehat{\mu}^{(h)}}{\widehat{\sigma}^{(h+1)}}\right)} = 0,\\
&\widehat{\theta}^{(h+1)} = \frac{C(\widehat{\theta}^{(h+1)})}{nC'(\widehat{\theta}^{(h+1)})}\sum_{i=1}^{n}\widehat{z}_i^{(h)},
\end{aligned}
$$

where $\widehat{\mu}^{(h)}$, $\widehat{\sigma}^{(h+1)}$ and $\widehat{\theta}^{(h+1)}$ are found numerically. Here, for $i = 1, ..., n$, we have that

$$
\widehat{z}_i^{(h)} = 1 + \frac{\widehat{\theta}^{(h)}\Phi\left(\frac{y_i-\widehat{\mu}^{(h)}}{\widehat{\sigma}^{(h)}}\right)C''\left(\widehat{\theta}^{(h)}\Phi\left(\frac{y_i-\widehat{\mu}^{(h)}}{\widehat{\sigma}^{(h)}}\right)\right)}{C'(\widehat{\theta}^{(h)}\Phi\left(\frac{y_i-\widehat{\mu}^{(h)}}{\widehat{\sigma}^{(h)}}\right))}.
$$

7.1 Evaluation of the standard errors from the EM-algorithm

We use the results of Louis (1982) to obtain the standard errors of the estimators from the EM-algorithm. The elements of the 3×3 observed information matrix $I_c(\Psi;\mathbf{y},\mathbf{z}) = -\left[\frac{\partial U_C(\mathbf{y},\mathbf{z};\mathbf{\Psi})}{\partial \mathbf{\Psi}}\right]$ are given by

$$
\begin{aligned}
\frac{\partial^2 l_n^*}{\partial \mu^2} &= \frac{n}{\sigma^2} + \frac{1}{\sigma^2}\sum_{i=1}^{n}(z_i-1)\frac{\left(\frac{y_i-\mu}{\sigma}\right)\phi\left(\frac{y_i-\mu}{\sigma}\right)\Phi\left(\frac{y_i-\mu}{\sigma}\right)+\phi^2\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi^2\left(\frac{y_i-\mu}{\sigma}\right)}, \\
\frac{\partial^2 l_n^*}{\partial \mu \partial \sigma} &= \frac{\partial^2 l_n^*}{\partial \sigma \partial \mu} = \frac{2}{\sigma^3}\sum_{i=1}^{n}(y_i-\mu) - \frac{1}{\sigma^2}\sum_{i=1}^{n}(z_i-1)\frac{(y_i-\mu)\,\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi\left(\frac{y_i-\mu}{\sigma}\right)} \\
&\quad + \frac{1}{\sigma^2}\sum_{i=1}^{n}(z_i-1)\frac{\left(\frac{y_i-\mu}{\sigma}\right)^2\phi\left(\frac{y_i-\mu}{\sigma}\right)\Phi\left(\frac{y_i-\mu}{\sigma}\right)+\left(\frac{y_i-\mu}{\sigma}\right)\phi^2\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi^2\left(\frac{y_i-\mu}{\sigma}\right)}, \\
\frac{\partial^2 l_n^*}{\partial \sigma^2} &= -\frac{n}{\sigma^2} + \frac{3}{\sigma^4}\sum_{i=1}^{n}(y_i-\mu)^2 - \frac{2}{\sigma^3}\sum_{i=1}^{n}(z_i-1)\frac{(y_i-\mu)\,\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi\left(\frac{y_i-\mu}{\sigma}\right)} \\
&\quad + \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{(z_i-1)\left(\frac{y_i-\mu}{\sigma}\right)^3\phi\left(\frac{y_i-\mu}{\sigma}\right)\Phi\left(\frac{y_i-\mu}{\sigma}\right)-\left(\frac{y_i-\mu}{\sigma}\right)\phi^2\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi^2\left(\frac{y_i-\mu}{\sigma}\right)}, \\
\frac{\partial^2 l_n^*}{\partial \theta^2} &= \frac{1}{\theta^2}\sum_{i=1}^{n} z_i + n\frac{C''(\theta)C(\theta)-(C'(\theta))^2}{C^2(\theta)}, \\
\frac{\partial^2 l_n^*}{\partial \theta \partial \mu} &= \frac{\partial^2 l_n^*}{\partial \mu \partial \theta} = \frac{\partial^2 l_n^*}{\partial \sigma \partial \theta} = \frac{\partial^2 l_n^*}{\partial \sigma \partial \theta} = 0.
\end{aligned}
$$

Taking the conditional expectation of $I_c(\Psi;\mathbf{y},\mathbf{z}) = -\left[\frac{\partial U_C(\mathbf{y},\mathbf{z};\boldsymbol{\Psi})}{\partial \boldsymbol{\Psi}}\right]$ given $\mathbf{y}$, we obtain the 3×3 matrix

$$
l_c(\boldsymbol{\Psi};\mathbf{y},\mathbf{z}) = E\left(I_c(\boldsymbol{\Psi};\mathbf{y},\mathbf{z}) \mid \mathbf{y}\right) = [c_{ij}] \tag{7.1}
$$

where

$$
\begin{aligned}
c_{11} &= \frac{n}{\sigma^2} + \frac{1}{\sigma^2}\sum_{i=1}^{n}(E(Z_i\mid y)-1)\left(\frac{y_i-\mu}{\sigma}\right)\frac{\phi\left(\frac{y_i-\mu}{\sigma}\right)\Phi\left(\frac{y_i-\mu}{\sigma}\right)+\phi^2\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi^2\left(\frac{y_i-\mu}{\sigma}\right)}, \\
c_{21} &= c_{12} = \frac{2}{\sigma^3}\sum_{i=1}^{n}(y_i-\mu) - \frac{1}{\sigma^2}\sum_{i=1}^{n}(E(Z_i\mid y)-1)\frac{(y_i-\mu)\,\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi\left(\frac{y_i-\mu}{\sigma}\right)} \\
&\quad + \frac{1}{\sigma^2}\sum_{i=1}^{n}(E(Z_i\mid y)-1)\frac{\left(\frac{y_i-\mu}{\sigma}\right)^2\phi\left(\frac{y_i-\mu}{\sigma}\right)\Phi\left(\frac{y_i-\mu}{\sigma}\right)-\left(\frac{y_i-\mu}{\sigma}\right)\phi^2\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi^2\left(\frac{y_i-\mu}{\sigma}\right)}, \\
c_{22} &= -\frac{n}{\sigma^2} + \frac{3}{\sigma^4}\sum_{i=1}^{n}(y_i-\mu)^2 - \frac{2}{\sigma^3}\sum_{i=1}^{n}(E(Z_i\mid y)-1)\frac{(y_i-\mu)\,\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi\left(\frac{y_i-\mu}{\sigma}\right)} \\
&\quad + \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{(E(Z_i\mid y)-1)\left(\frac{y_i-\mu}{\sigma}\right)^3\phi\left(\frac{y_i-\mu}{\sigma}\right)\Phi\left(\frac{y_i-\mu}{\sigma}\right)-\left(\frac{y_i-\mu}{\sigma}\right)\phi^2\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi^2\left(\frac{y_i-\mu}{\sigma}\right)}, \\
c_{33} &= \frac{1}{\theta^2}\sum_{i=1}^{n}E(Z_i\mid y) + n\frac{C''(\theta)C(\theta)-(C'(\theta))^2}{C^2(\theta)}, \quad c_{13} = c_{31} = c_{23} = c_{32} = 0,
\end{aligned}
$$

and

$$E(Z_i \mid \mathbf{y}) = 1 + \frac{\theta\Phi\left(\frac{y-\mu}{\sigma}\right) C''\left(\theta\Phi\left(\frac{y-\mu}{\sigma}\right)\right)}{C'\left(\theta\Phi\left(\frac{y-\mu}{\sigma}\right)\right)}.$$

Moving now to the computation of $l_m\left(\mathbf{\Psi};\mathbf{y}\right)$ as

$$l_m\left(\Psi;\mathbf{y}\right) = Var[U_C\left(\mathbf{y},\mathbf{z};\Psi\right) \mid \mathbf{y}] = [v_{ij}], \tag{7.2}$$

where

$$\begin{aligned}
v_{11} &= \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{\phi^2\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi^2\left(\frac{y_i-\mu}{\sigma}\right)} Var[Z_i \mid \mathbf{y}],\\
v_{22} &= \sum_{i=1}^{n}\left(\frac{\left(y_i-\mu\right)\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\sigma^2\Phi\left(\frac{y_i-\mu}{\sigma}\right)}\right)^2 Var[Z_i \mid \mathbf{y}],\\
v_{33} &= \frac{1}{\theta^2}\sum_{i=1}^{n} Var[Z_i \mid \mathbf{y}],\\
v_{21} &= v_{12} = \frac{1}{\sigma^3}\sum_{i=1}^{n}\left(y_i-\mu\right)\left(\frac{\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi\left(\frac{y_i-\mu}{\sigma}\right)}\right)^2 Var[Z_i \mid \mathbf{y}],\\
v_{13} &= v_{31} = -\frac{1}{\sigma\theta}\sum_{i=1}^{n}\frac{\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi\left(\frac{y_i-\mu}{\sigma}\right)}\; Var[Z_i \mid \mathbf{y}],\\
v_{23} &= v_{32} = -\frac{1}{\theta\sigma^2}\sum_{i=1}^{n}\frac{\left(y_i-\mu\right)\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\Phi\left(\frac{y_i-\mu}{\sigma}\right)} Var[Z_i \mid \mathbf{y}].
\end{aligned}$$

and

$$\begin{aligned}
Var\left[Z_i \mid \mathbf{y}\right] &= E\left(Z_i^2 \mid \mathbf{y}\right) - \left(E\left(Z_i \mid \mathbf{y}\right)\right)^2\\
&= \frac{1}{C'\left(\theta_*\right)}\sum_{z=1}^{n} a_z z^3 \theta_*^{z-1} - \frac{\left(C'\left(\theta_*\right) + \theta_* C''\left(\theta_*\right)\right)^2}{\left(C'\left(\theta_*\right)\right)^2}\\
&= \frac{\theta_*^2 C'''\left(\theta_*\right) + C'\left(\theta_*\right) + 3\theta_* C''\left(\theta_*\right)}{C'\left(\theta_*\right)} - \frac{[C'\left(\theta_*\right) + \theta_* C''\left(\theta_*\right)]^2}{\left(C'\left(\theta_*\right)\right)^2},
\end{aligned}$$

in which $\theta_* = \theta\Phi\left(\frac{y_i-\mu}{\sigma}\right)$.
Applying the Equations (7.1) and (7.2), we obtain the observed information as

$$I\left(\widehat{\mathbf{\Psi}};\mathbf{y}\right) = l_c\left(\widehat{\mathbf{\Psi}};\mathbf{y}\right) - l_m\left(\widehat{\mathbf{\Psi}};\mathbf{y}\right).$$

The standard errors of the MLEs of the EM-algorithm are the square root of the diagonal elements of the $I\left(\widehat{\mathbf{\Psi}};\mathbf{y}\right)$.

8 Simulation study

This section provides the results of simulation study. Because of time-consuming simulation, it has been performed in order to investigate the proposed estimator of, μ, σ, θ of the proposed EM method for NG distribution. We simulate 1000 times under the NG distribution with different sets of parameters and sample sizes $n = 50,\ 100,\ 150,\ 300$ and 500. For each sample size, we compute the MLEs by EM-method. We also compute the root of mean square errors (RMSE), standard errors (SE) and covariances of the MLEs of the EM-algorithm. The results for the NG distribution are reported in Tables 4. Some of the points are quite clear from the simulation results: (i) Convergence has been achieved in all cases and this emphasizes the numerical stability of the EM-algorithm. (ii) The differences between the average estimates and the true values are almost small. (iii) These results suggest that the EM estimates have performed consistently. (iv) As the sample size increases, the root of mean square errors and the standard errors of the MLEs decrease.

9 Applications of the NPS class of distributions to two real data sets

In this section, we try to illustrate the better performance of the proposed model. For this perpose, we fit NG, NP and NL models to two real data sets. We also fit the Azzalini's skew-normal (SN) and normal distributions to make a comparison with the NPS models. The first data concerning the heights (in centimeters) of 100 Australian athletes. The data have been previously analyzed in Cook and Weisberg (1994) and are available for download at *http://azzalini.stat.unipd.it/SN/index.html*. We estimate parameters by numerically maximizing the likelihood function. The MLEs of the parameters, the log-likelihood, the AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) for the NG, NP, NL, normal and SN models are given in Table 5.

As is well known, a model with a minimum AIC value is to be preferred. Therefore NG distribution provides a better fit to this data set than the other distributions and hence could be chosen as the best distribution. Also this conclusion is confirmed from the plots of the densities functions in Figure 11.

The second data represent the Oits IQ Scores for 52 non-White males hired by a large insurance company in 1971 given in Roberts (1988). Table 6 gives the MLEs of the parameters, the log-likelihood, the AIC and BIC for the NG, NP, NL, Normal and SN models for the second data set.

The results for the second data set show that the NG distributions yield the best fit among the NPS class of distributions and is a proper competitor for the normal and SN distributions. Also the plots of the densities in Figure 12 confirmed this conclusion.

Table 4: The averages of the 1000 MLE's, mean of the simulated root of mean square errors, mean of simulated standard errors and mean of the simulated covariances of EM estimators for NG distribution.

		Average estimators			RMSE			SE			Cov		
n	(μ, σ, θ)	$\widehat{\mu}$	$\widehat{\sigma}$	$\widehat{\theta}$	$\widehat{\mu}$	$\widehat{\sigma}$	$\widehat{\theta}$	$\widehat{\mu}$	$\widehat{\sigma}$	$\widehat{\theta}$	$(\widehat{\mu}, \widehat{\sigma}))$	$(\widehat{\mu}, \widehat{\theta})$	$(\widehat{\sigma}, \widehat{\theta})$
50	(0.0, 1.0, -0.8)	-0.4111	0.8925	-0.8727	0.5822	0.2564	0.8772	0.4125	0.1398	0.8747	-0.0088	-0.3018	0.0510
	(0.0, 1.0, -0.5)	0.2847	0.9064	-0.5980	0.4220	0.2269	0.6123	0.3116	0.1282	0.6047	-0.0058	-0.1373	0.0370
	(0.0, 1.0, -0.2)	0.186	0.9099	-0.3034	0.2751	0.2227	0.3740	0.2028	0.1309	0.3596	-0.0003	-0.0373	0.0168
	(0.0, 1.0, 0.05)	-0.0890	0.9839	0.1392	0.3232	0.0990	0.2551	0.3109	0.0978	0.2391	-0.0051	-0.0644	0.0039
	(0.0, 1.0, 0.2)	-0.1418	0.9907	0.3094	0.3754	0.0977	0.2741	0.3477	0.0973	0.2515	-0.0065	-0.0774	0.0034
	(0.0, 1.0, 0.5)	-0.0326	0.9865	0.4558	0.3764	0.0888	0.2648	0.3752	0.0878	0.2612	-0.0091	-0.0844	0.0043
	(0.0, 1.0, 0.8)	0.0747	0.9814	0.6942	0.4699	0.0835	0.2733	0.4642	0.0815	0.2521	-0.0175	-0.1025	0.0086
100	(0.0, 1.0, -0.8)	-0.3214	0.9271	-0.7796	0.4890	0.1769	0.6916	0.3687	0.1002	0.6916	-0.0068	-0.2141	0.0348
	(0.0, 1.0, -0.5)	0.2427	0.93835	-0.4805	0.3610	0.1498	0.3760	0.2673	0.0851	0.3757	-0.0040	-0.0872	0.0135
	(0.0, 1.0, -0.2)	0.1626	0.93785	-0.2165	0.2264	0.1512	0.1706	0.1575	0.0861	0.1699	0.0006	-0.0173	0.0022
	(0.0, 1.0,0.05)	-0.0311	0.9913	0.0789	0.1745	0.0708	0.1422	0.1718	0.0703	0.1393	-0.0008	-0.0191	0.0007
	(0.0, 1.0, 0.2)	-0.0876	0.9983	0.2742	0.2520	0.0698	0.2107	0.2364	0.0698	0.1973	-0.0028	-0.0411	0.0019
	(0.0, 1.0, 0.5)	-0.0295	0.9958	0.4809	0.2864	0.2878	0.0658	0.2153	0.0657	0.2146	-0.0052	-0.0547	0.0027
	(0.0, 1.0, 0.8)	0.0587	0.9906	0.7354	0.3280	0.3330	0.0591	0.1935	0.0584	0.1825	-0.0097	-0.0538	0.0046
150	(0.0, 1.0, -0.8)	-0.2914	0.9419	-0.7396	0.4455	0.1402	0.5687	0.3372	0.0785	0.5658	-0.0078	-0.1727	0.0233
	(0.0, 1.0, -0.5)	0.204	0.9408	-0.4897	0.3132	0.1396	0.3590	0.2378	0.0740	0.3590	-0.0038	-0.0753	0.0112
	(0.0, 1.0, -0.2)	0.1293	0.9478	-0.2265	0.1948	0.1234	0.1429	0.1458	0.0657	0.1405	0.0015	-0.0147	0.0005
	(0.0, 1.0, 0.05)	-0.0176	0.9938	0.0674	0.1420	0.0569	0.1088	0.1410	0.0566	0.1074	-0.0007	-0.0120	0.0005
	(0.0, 1.0, 0.2)	-0.0714	0.9976	0.2605	0.2095	0.0557	0.1838	0.1971	0.0556	0.1736	-0.0017	-0.0302	0.0008
	(0.0, 1.0, 0.5)	-0.0127	0.9993	0.4842	0.2186	0.0545	0.1671	0.2183	0.0545	0.16641	-0.0038	-0.0322	0.0019
	(0.0, 1.0, 0.8)	0.0346	0.9914	0.7648	0.2664	0.0469	0.1347	0.2643	0.0461	0.1300	-0.0061	-0.0310	0.0024
300	(0.0, 1.0, -0.8)	-0.2369	0.9525	-0.7188	-0.7188	0.3814	0.1169	0.4670	0.2990	0.4601	-0.0051	-0.1291	0.0138
	(0.0, 1.0, -0.5)	-0.1766	0.95645	-0.4565	0.2695	0.1044	0.2909	0.2037	0.0575	0.2878	-0.0027	-0.0536	0.0071
	(0.0, 1.0, -0.2)	0.1017	0.9637	-0.2059	0.1463	0.0845	0.0962	0.1053	0.0433	0.0961	0.0004	-0.0068	0.0003
	(0.0, 1.0, 0.05)	0.0017	0.9974	0.0488	0.0570	0.0415	0.0018	0.0570	0.0414	0.0018	-0.0001	0.0002	0.0001
	(0.0, 1.0, 0.2)	-0.0390	0.9985	0.2385	0.1323	0.0408	0.1255	0.1265	0.0407	0.1195	-0.0007	-0.0131	0.0002
	(0.0, 1.0, 0.5)	0.0092	0.9996	0.4944	0.1460	0.0384	0.1159	0.1458	0.0384	0.1158	-0.0015	-0.0149	0.0006
	(0.0, 1.0, 0.8)	0.0230	0.9932	0.7846	0.1680	0.0321	0.0757	0.1665	0.0314	0.0742	-0.0026	-0.0111	0.0009
500	(0.0, 1.0, -0.8)	-0.1982	0.9568	-0.7398	0.3442	0.1026	0.4351	0.2815	0.0554	0.4311	-0.0065	-0.1154	0.0137
	(0.0, 1.0, -0.5)	-0.1627	0.9697	-0.4266	0.2472	0.0767	0.2537	0.1862	0.0471	0.2430	-0.0022	-0.0425	0.0046
	(0.0, 1.0, -0.2)	0.0885	0.9714	-0.1986	0.1149	0.0709	0.0762	0.0733	0.0419	0.0762	-0.0002	-0.0038	0.0004
	(0.0, 1.0, 0.05)	0.0012	0.9997	0.0499	0.0436	0.0325	0.0014	0.0436	0.0326	0.0014	-0.0001	-0.0001	0.0001
	(0.0, 1.0, 0.2)	-0.0157	0.9994	0.2144	0.0756	0.0311	0.0716	0.0739	0.0311	0.0702	-0.0004	-0.0030	0.0001
	(0.0, 1.0, 0.5)	-0.0054	0.9996	0.5001	0.1460	0.0384	0.1159	0.1005	0.0301	0.0796	0.0007-	0.0010	0.0001
	(0.0, 1.0, 0.8)	0.0062	0.9992	0.7946	0.0997	0.0242	0.0397	0.0995	0.0242	0.0393	-0.0012	-0.0034	0.0003

Table 5: Parameter estimates, AIC and BIC for AIS data.

Dist.	Parameter estimates	$-\log(L)$	AIC	BIC
NG	$\widehat{\mu}$=136.001, $\widehat{\sigma}$=13.642, $\widehat{\theta}$=0.998	348.376	702.752	710.567
NP	$\widehat{\mu}$=167.106, $\widehat{\sigma}$=9.208, $\widehat{\theta}$=3.398	349.145	704.291	712.106
NL	$\widehat{\mu}$=169.353, $\widehat{\sigma}$=7.947, $\widehat{\theta}$=0.897	350.872	707.745	715.560
Normal	$\widehat{\mu}$=174.594, $\widehat{\sigma}$=8.209	352.319	708.635	713.846
SN	$\widehat{\mu}$=170.320, $\widehat{\sigma}$=8.002, $\widehat{\theta}$=0.0016	352.032	710.636	718.451

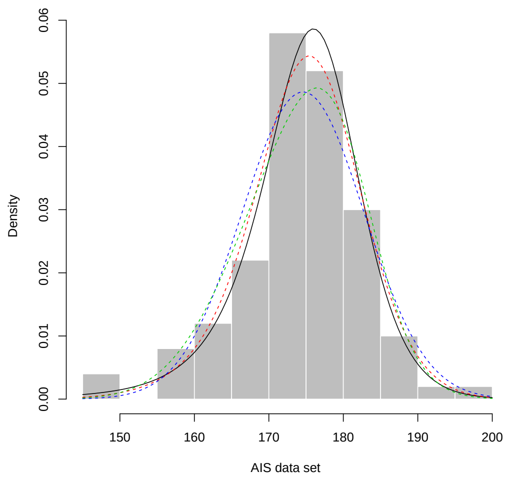

Figure 11: Histogram of heights of 100 Australian athletes.The lines represent distributions fitted using maximum likelihood estimation: NG (Black), NP (Red), NL(Green) and ASN (Blue)

.

10 Conclusion

In this paper we introduce a new three-parameter class of distributions called the normal-power series distributions (NPS), which is an alternative to the Azzalini skew-normal distribution for fitting skewed data. The NPS distributions contain the NG, NP, NB and NL distributions as special cases. We obtain expressions for the moments. The estimation of the unknown parameters of the proposed distribution is approached by the EM-algorithm. Finally, we fitted NPS models to two real data sets to show the potential of the new proposed class.

Table 6: Parameter estimates, AIC and BIC for OTIS IQ scores data.

Dist.	Parameter estimates	$-\log(L)$	AIC	BIC
NG	$\widehat{\mu}$=112.875, $\widehat{\sigma}$=82.313, $\widehat{\theta}$=-2.989	182.313	370.628	376.479
NP	$\widehat{\mu}$=106.263, $\widehat{\sigma}$=8.227, $\widehat{\theta}$=0.0000002	182.313	372.850	376.479
NL	$\widehat{\mu}$=106.308, $\widehat{\sigma}$=7.947, $\widehat{\theta}$=0.0000002	183.433	372.867	378.719
Normal	$\widehat{\mu}$=106.654, $\widehat{\sigma}$=8.230	183.387	370.774	374.676
SN	$\widehat{\mu}$=98.790, $\widehat{\sigma}$=11.380, $\widehat{\theta}$=1.710	182.436	370.872	376.726

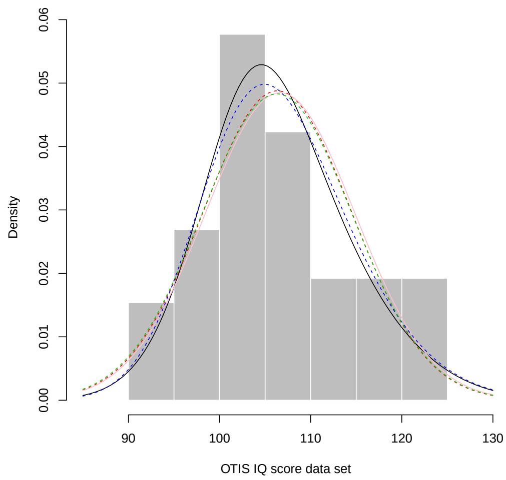

Figure 12: Histogram of OTIS IQ scores. The lines represent distributions fitted using maximum likelihood estimation: NG (Black), NP (Red), NL(Green), normal(pink) and ASN (Blue)

.

Acknowledgement

The authors would like to thank the Referees and the Editor for their valuable comments and suggestions which have contributed to substantially improving the manuscript. The authors are also indebted to Yazd University for supporting this research.

References

Adamidis, A. K. and Loukas, S. (1998). A lifetime distribution with decreasing failure rate. *Statistics and Probability Letters*, 39.

Arnold, B. C. and Beaver, R. J. (2002). Skewed multivariate models related to hidden truncation and selective reporting (with discussion). *Test*, 11(1).

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12.

Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones. *Statistica*, XLVI(2).

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution. *Journal of the Royal Statistical Society, Series B*, 61(3).

Azzalini, A. and Chiogna, M. (2004). Some results on the stress-strength model for skew-normal variates. *Metron*, 3.

Azzalini, A. and Valle, A. D. (1996). The multivariate skew-normal distribution. *Biometrika*, 83(4).

Barreto-Souza, W., Morais, A. L. and Cordeiro, G. M. (2011). The Weibull-geometric distribution. *Journal of Statistical Computation and Simulation*, 81.

Chahkandi, M. and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics and Data Analysis*, 53.

Cook, R. D. and Weisberg, S. (1994). *An Introduction to Regression Graphics.* Wiley, New York.

Fung, T. and Seneta, E. (2007) Tail weight, quantiles and kurtosis: a study of competing distributions. *Operations Research Letters*, 35.

Gupta, R. C. and Gupta, R. D. (2004). Generalized skew normal model. *Test* , 12 (2).

Gupta, R. D. and Gupta, R. C. (2008). Analyzing skewed data by power normal model. *Test*, 17(1).

Jamalizadeh, A. and Balakrishnan, N. (2010). Distributions of order statistics and linear combinations of order statistics from an elliptical distribution as mixtures of unified skew- elliptical distributions. *Journal of Multivariate Analysis*, 101.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 44.

Mahmoudi, E. and Jafari, A. A. (2012). Generalized exponential-power series distributions. *Computational Statistics and Data Analysis*, 56(12).

Mahmoudi, E. and Jafari, A. A. (2017). The compound class of linear failure rate-power series distributions: Model, properties and applications. *Communication in Statistics-Simulation and Computation*, 46(2).

Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84(3).

Morais, A. L. and Barreto-Souza, W. (2011). A compound class of Weibull and power series distributions. *Computational Statistics and Data Analysis*, 55.

Muhammad, M. (2017). The complementary exponentiated Burr XII Poisson distribution: Model, properties and application. *Journal of Statistics Applications Probability*, 6(1).

Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics*, 32(7).

Noack, A. (1950). A class of random variables with discrete distribution. *Annals of Mathematical Statistics*, 21.

Roberts, H. V. (1988). *Data Analysis for Managers with Minitab*. Scientific Press, Redwood City, CA.

Roozegar, R., and Nadarajah, S. (2016). The quadratic hazard rate power series distribution. *Journal of Testing and Evaluation*, 49.

Sharafi, M. and Behboodian, J. (2008). The Balakrishnan skew-normal density. *Statistical Papers*, 45.

Tahmasbi, R. and Rezaei, S. (2008) A two-parameter lifetime distribution with decreasing failure rate. *Computational Statistics and Data Analysis,* 52.

Tahmasbi, S. Jafari, A. A. (2015). Exponentiated extended Weibull-power series class of distributions. *Ciencia e Natura*, 7(2).

Xie, F. G., Lin, J. G. and Wei, B. C. (2009). Diagnostics for skew-normal nonlinear regression models with AR(1) errors. *Computational Statistics and Data Analysis,* 53.