

Normal–Poisson distribution as a lifetime distribution of a series system

Eisa Mahmoudi*¹, Hamed Mahmoodian² e Fatemeh Esfandiari³

^{1,2,3}Department of Statistics, Yazd University, Yazd, Iran

Abstract

In this paper, we introduce a new three parameter skewed distribution. This new class which is obtained by compounding the normal and Poisson distributions, is presented as an alternative to the class of skew-normal and normal distributions, among others. Different properties of this new distribution have been investigated. The density and distribution functions of proposed distribution, are given by a closed expression which allows us to easily compute probabilities, moments and related measurements. Estimation of the parameters of this new model using maximum likelihood method via an EM-algorithm is given. Finally, some applications of this new distribution to real data are given.

Keywords: Normal distribution, Poisson distribution, EM-algorithm, Maximum likelihood estimation.

1 Introduction

The normal distribution is probably the most well known statistical distribution and widely used to model many phenomena. Since the normal distribution is symmetric, skew-normal distributions have been proposed, studied and generalized by many authors. Azzalini (1985) proposed the skew normal (SN) distribution with the pdf $\phi(z; \lambda) = 2\phi(z)\Phi(\lambda z)$, for $z, \lambda \in \mathbb{R}$. This distribution and its variations have been discussed by several authors including Azzalini (1986), Henze (2008), Branco and Dey (2011), Loperfido (2010), Arnold and Beaver (2002), Balakrishnan (2002), Azzalini and Chiogna (2004), Sharafi and Behboodian (2008), Elal-Olivero (2010). Azzalini and Valle (1996), Azzalini and Capitanio (1999) and Azzalini and Chiogna (2004) have all discussed various multivariate forms of skew-normal distributions.

The number of components in a system can be fixed number n . recently, some researcher introduced a system when the number of components is a random variable N with support $1, 2, \dots$. In such systems, components may be parallel or series. If the lifetime of i th component is the continuous random variable X_i , then the lifetime of such a system is defined by $Y = \max_{1 \leq i \leq N} X_i$ or $Y = \min_{1 \leq i \leq N} X_i$, based on whether the components are parallel or series.

By taking a system with series components in which the random variable N has Poisson distribution truncated as zero and the random variable X_i follows the normal distribution, we introduce a new generalization of the normal distribution.

Recently, some authors focused on these new compounding distribution; Mahmoudi and Mahmoodian (2017) introduced the normal power series class of distributions in a system with parallel components. This introduced class of distributions contains normal-geometric, normal-Poisson, normal-logarithmic and normal-binomial models as special case. Roozegar and Nadarajah (2017) also introduced the power series skew normal class of distributions using this fact that the i th component X_i has SN distribution and consider $Y = \sum_{i=1}^N X_i$.

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 $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, simply $\Phi_n(\cdot; \boldsymbol{\Sigma})$ for the case when $\boldsymbol{\mu} = \mathbf{0}$. Furthermore, for $r \in \mathbb{N}$, let $\mathbf{1}_r$ and \mathbf{I}_r denote the vector of ones and the identity matrix of dimension r , respectively.

The rest of this paper is organized as follows: In Section 2, we define the class of normal–Poisson (NP) distribution. The density, hazard rate, survival functions and some of their properties are given in this Section. We derive moments of NP distribution in Section 3. In Section 4, we derive an expansion for the density of the order statistics. In Section 5, the Shannon entropy is derived. In Section 6, estimates of the parameters based on a random sample coming from this family of distributions are derived via an EM algorithm. Applications to two real data sets are given in Section 7. Finally, Section 8 concludes the paper.

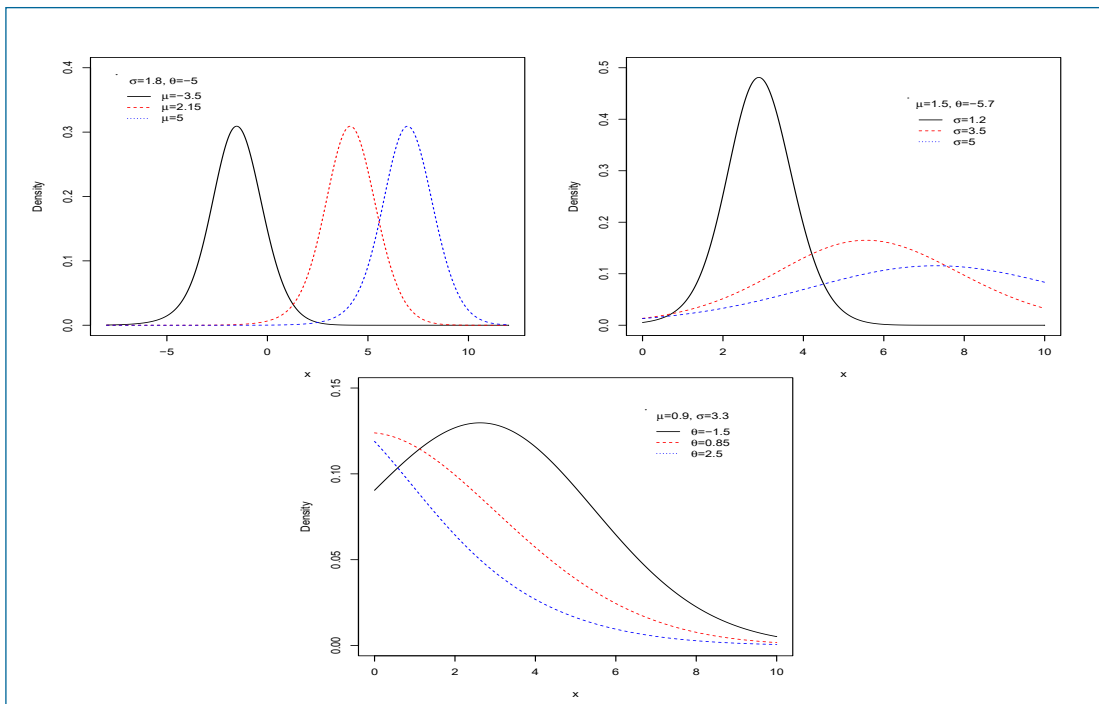


Figure 1: Plots of the NP density function for selected parameter values.

2 The Normal–Poisson distribution and Some Properties

Let X_1, X_2, \dots be a sample from a normal distribution with mean μ and variance σ^2 . Let N be distributed according to a Poisson distribution truncated at zero, with the probability mass function

$$P(N = n) = \frac{\theta^n}{n!(e^\theta - 1)},$$

where $\theta > 0$. Moreover, N is independent of X_i 's.

Definition 2.1. A random variable X is said to have a normal–Poisson distribution, denoted by $X \sim NP(\mu, \sigma, \theta)$, if

$$X = \min(X_1, \dots, X_N). \quad (2.1)$$

From the definition in (2.1), we have, for $x \in \mathbb{R}$

$$\begin{aligned} F(x; \mu, \sigma, \theta) &= P(X \leq x) = \sum_{n=1}^{\infty} P(X \leq x \mid N = n)P(N = n) \\ &= \sum_{n=1}^{\infty} P(\min(X_1, \dots, X_N) \leq x \mid N = n)P(N = n) \\ &= 1 - \sum_{n=1}^{\infty} \frac{(1 - \Phi(\frac{x-\mu}{\sigma}))^n \theta^n}{n!(e^\theta - 1)} = \frac{e^\theta - e^{\theta(1-\Phi(\frac{x-\mu}{\sigma}))}}{e^\theta - 1} = \frac{1 - e^{-\theta\Phi(\frac{x-\mu}{\sigma})}}{1 - e^{-\theta}}. \end{aligned} \quad (2.2)$$

Hence the pdf of X is given by

$$f(x; \mu, \sigma, \theta) = \frac{\theta\phi(\frac{x-\mu}{\sigma})e^{\theta(1-\Phi(\frac{x-\mu}{\sigma}))}}{\sigma(e^\theta - 1)}, \quad (2.3)$$

where $\mu \in \mathbb{R}$ is the location parameter, $\sigma > 0$ is the scale parameter and $\theta > 0$ is the shape parameter, which characterize the skewness, kurtosis, and unimodality of the distribution.

Remark 2.1. Even when $\theta < 0$, Equation (2.3) is also a density function. We can then define the NP distribution by Equation (2.3) for any $\theta \in \mathbb{R} - \{0\}$.

Plots of the NP density function for selected parameter values are given in Figure 1. An important characteristic of the NP distribution is that its density function can be unimodal that makes this distribution many advantages in modelling lifetime data. The normal distribution with mean μ and variance σ^2 is a special case of the NP distribution when $\theta \rightarrow 0$. When $\mu = 0$ and $\sigma = 1$, we say that X has standard NP distribution with pdf

$$f(x; \theta) = \frac{\theta\phi(x)e^{\theta(1-\Phi(x))}}{(e^\theta - 1)}.$$

Proposition 2.1. If $X \sim NP(0, 1, \theta)$, then we have

$$F(x; 0, 1, \theta) = 1 - F(-x; 0, 1, -\theta).$$

Proof. From (2.2), it can be found that

$$1 - F(-x; 0, 1, -\theta) = 1 - \frac{1 - e^{\theta(1-\Phi(x))}}{1 - e^{-\theta}} = \frac{1 - e^{-\theta\Phi(x)}}{1 - e^{-\theta}} = F(x; 0, 1, \theta)$$

□

Proposition 2.2. If $X_1 \sim NP(0, 1, \theta_1)$ and $X_2 \sim NP(0, 1, \theta_2)$ are independent random variables, then the stress-strength parameter, $R = P(X_1 < X_2)$, is given by

$$R = \frac{e^{-\theta_1}\theta_2}{(e^{\theta_1} - 1)(e^{\theta_2} - 1)} \left[\frac{e^{-\theta_1}}{(e^{\theta_1} - 1)} + \frac{e^{\theta_2}}{\theta_1 + \theta_2} (e^{\theta_1 + \theta_2} - 1) \right].$$

Proof. The stress-strength parameter is given by

$$\begin{aligned}
 R &= P(X_1 < X_2) = \int_{-\infty}^{+\infty} F(y; 0, 1, \theta_1) f(y; 0, 1, \theta_2) dy \\
 &= \int_{-\infty}^{+\infty} \frac{1 - e^{-\theta_1 \Phi(y)}}{1 - e^{-\theta_1}} \times \frac{\theta_2 \phi(y) e^{\theta_2(1-\Phi(y))}}{(e^{\theta_2} - 1)} dy.
 \end{aligned}$$

By change of variable $\Phi(y) = t$ and some simple calculation, we have

$$R = \frac{e^{-\theta_1} \theta_2}{(e^{\theta_1} - 1)(e^{\theta_2} - 1)} \left[\frac{e^{-\theta_1}}{(e^{\theta_1} - 1)} + \frac{e^{\theta_2}}{\theta_1 + \theta_2} (e^{\theta_1 + \theta_2} - 1) \right].$$

□

Proposition 2.3. The densities of NP distribution can be written as infinite number of linear combination of density of order statistics. We know that $e^\theta = \sum_{n=1}^{\infty} \frac{\theta^{n-1}}{(n-1)!}$, therefore

$$f(x; \mu, \sigma, \theta) = \sum_{n=1}^{\infty} \frac{\theta^n}{n! (e^\theta - 1)} g_{X_{(1)}}(x; n), \tag{2.4}$$

where $g_{X_{(1)}}(x; n)$ denotes the density function of $X_{(1)} = \min(X_1, \dots, X_n)$.

The survival and failure rate functions of the NP distribution are given respectively by

$$S(y; \mu, \sigma, \theta) = \frac{e^{\theta(1-\Phi(\frac{y-\mu}{\sigma}))} - 1}{e^\theta - 1}, \tag{2.5}$$

and

$$h(y; \mu, \sigma, \theta) = \frac{\theta \phi(\frac{x-\mu}{\sigma}) e^{\theta(1-\Phi(\frac{x-\mu}{\sigma}))}}{\sigma (e^{\theta(1-\Phi(\frac{y-\mu}{\sigma}))} - 1)}. \tag{2.6}$$

Plots of the NP failure rate function for selected parameter values are given in Figure 2. An important characteristic of the NP distribution is that its failure rate function can be increasing, unimodal-bathtub shaped that makes this distribution flexible in modeling different types of lifetime data.

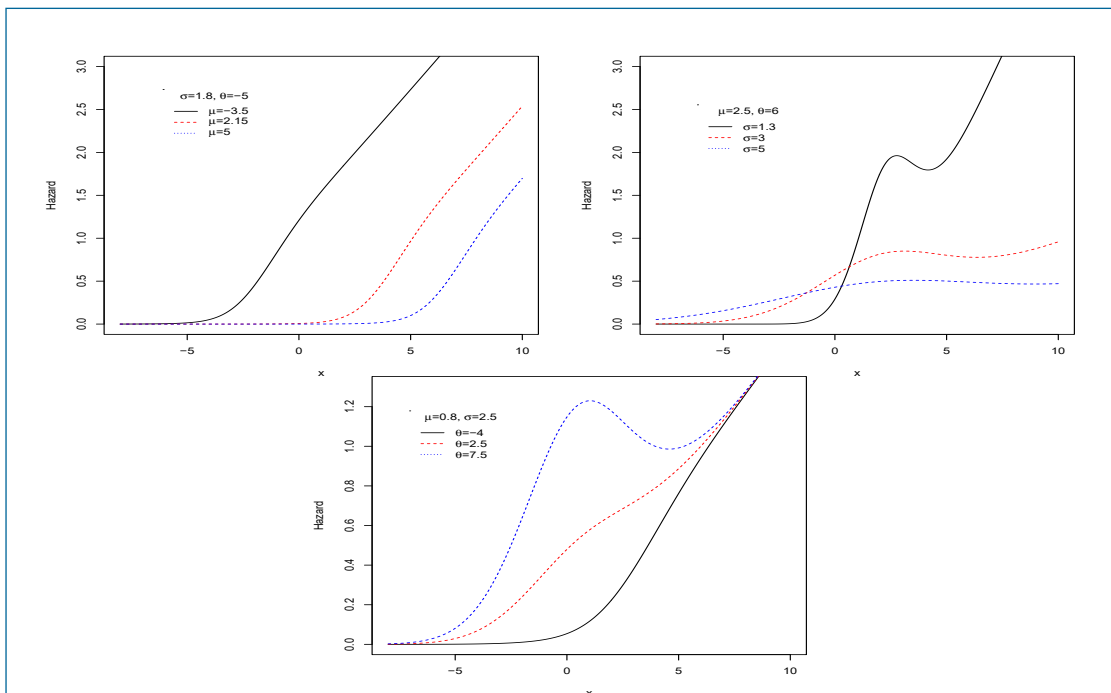


Figure 2: Plots of the NP failure rate function for selected parameter values.

3 Quantiles and Moments

The p th quantile of the NP distributions is given by

$$x_p = \mu + \sigma \Phi^{-1} \left[1 - \frac{\log(p(e^\theta - 1) + 1)}{\theta} \right].$$

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

Now, by using equation (2.4), we derive moment generating function of $X \sim NP(\mu, \sigma, \theta)$. The moment generating function is given by

$$M_X(t) = \sum_{n=1}^{\infty} \frac{\theta^n}{n! (e^\theta - 1)} M_{X_{(1)}}(t). \tag{3.1}$$

and based on Jamalizadeh and Balakrishnan (2010)

$$M_{X_{(1)}}(t) = n \exp \left(\frac{1}{2} \sigma^2 t^2 + \mu t \right) \Phi_{n-1}(-\mathbf{1}_{n-1} \sigma t; \mathbf{I}_{n-1} + \mathbf{1}_{n-1} \mathbf{1}_{n-1}^T).$$

Therefore

$$M_X(t) = \exp \left(\frac{1}{2} \sigma^2 t^2 + \mu t \right) \sum_{n=1}^{\infty} \frac{\theta^n}{n! (e^\theta - 1)} \times n \Phi_{n-1}(-\mathbf{1}_{n-1} \sigma t; \mathbf{I}_{n-1} + \mathbf{1}_{n-1} \mathbf{1}_{n-1}^T). \tag{3.2}$$

Using (3.2), k -th moment of a random variable X can be obtained. But from the direct and simple calculation, we have

$$\mu_k = \int_{-\infty}^{+\infty} x^k \times \frac{\theta \phi(x; \mu, \sigma) e^{\theta(1-\Phi(x; \mu, \sigma))}}{(e^\theta - 1)} dx.$$

By changing of variable to $t = \Phi(x; \mu, \sigma)$, we have

$$\mu_k = \int_0^1 (\mu + \sigma \Phi^{-1}(t))^k \frac{\theta e^{\theta(1-t)}}{e^\theta - 1} dt.$$

Thus, first two moments are as follows:

$$\mu_1 = E(X) = \frac{\mu e^{-\theta}(1 + \sigma)}{e^\theta - 1} [1 - e^{-\theta}] + \frac{\theta e^{-\theta}}{e^\theta - 1} \sqrt{2} \sigma^2 \times \int_0^1 \text{erf}^{-1}(2t - 1) e^{-\theta t} dt, \tag{3.3}$$

$$\begin{aligned} \mu_2 = E(X^2) &= \frac{\theta e^{-\theta}}{e^\theta - 1} \left[\frac{\mu^2(1 + \sigma^2) + 2\mu\sigma}{\theta} [-e^{-\theta}] + 2\sqrt{2}\mu(\sigma^2 + \sigma^3) \right. \\ &\quad \left. \times \int_0^1 \text{erf}^{-1}(2t - 1) e^{-\theta t} dt + 2\sigma^4 \int_0^1 [\text{erf}^{-1}(2t - 1)]^2 e^{-\theta t} dt \right], \end{aligned} \tag{3.4}$$

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

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

	$\theta = -3$	$\theta = -1$	$\theta = -0.3$	$\theta = 0.01$	$\theta = 0.3$	$\theta = 1$	$\theta = 3$	$\theta = 10$
$E(X)$	0.7541	-0.2781	0.0845	-0.0028	-0.0845	-0.2781	-0.7541	-1.5045
$E(X^2)$	1.3477	1.0450	1.0041	1.0000	1.0041	1.0450	1.3477	2.6533
$E(X^3)$	2.0013	-0.7003	0.2114	-0.0071	-0.2114	-0.7003	-2.0013	-5.2127
$E(X^4)$	4.5372	3.1954	3.0179	3.0000	3.0179	3.1954	4.5372	11.2262
VAR	0.7790	0.9677	0.9970	1.0000	0.9970	0.9677	0.7790	0.3898
SK	-0.2764	0.1349	-0.0421	0.0014	0.0421	0.1349	0.2764	-0.1973
KUR	3.5076	3.0792	3.0074	3.0000	3.0074	3.0792	3.5076	3.4236

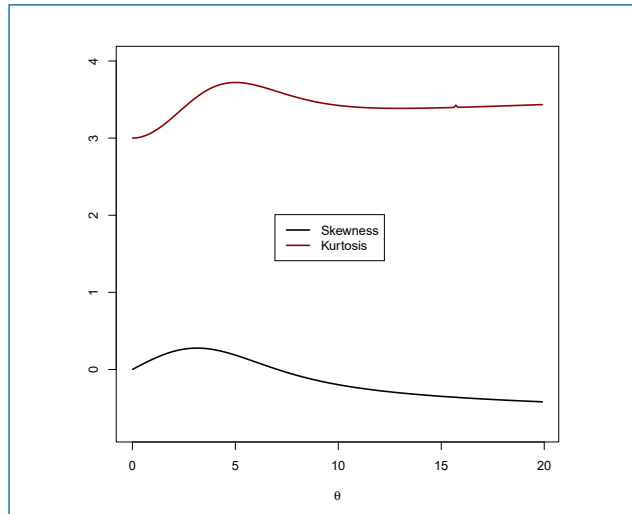


Figure 3: Plots of skewness and kurtosis of NP distribution for selected parameter values θ .

4 Order statistics

Let X_1, \dots, X_n be a random sample from NP distribution. The pdf of the i th order statistic, $X_{(i)}$, is given by

$$f_{(i)}(x; \mu, \sigma, \theta) = \frac{n! f(x; \mu, \sigma, \theta)}{(n-i)!(i-1)!} [1 - F(x; \mu, \sigma, \theta)]^{n-i} [F(x; \mu, \sigma, \theta)]^{i-1}. \quad (4.1)$$

By using the binomial series expansion we have

$$f_{(i)}(x) = \frac{n! f(x; \mu, \sigma, \theta)}{(n-i)!(i-1)!} \sum_{j=0}^{\infty} (-1)^j \binom{n-i}{j} [F(x; \mu, \sigma, \theta)]^{j+i-1}.$$

The moments of the probability density function (4.1) cannot be obtained in a closed form but the values are derived by using numerical computation. Figure 4 shows the expected values of the order statistics plot from $NP(0,1,\theta)$ for $n = 40, i = 1,10,20$ and different values θ .

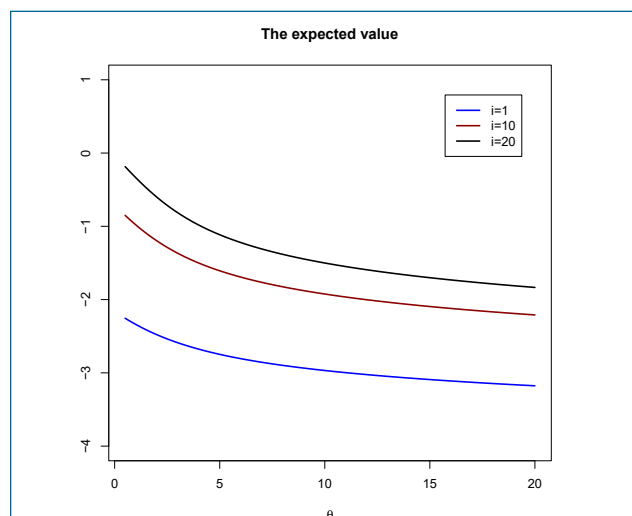


Figure 4: Plots of the expected values of the order statistics from $NP(0,1,\theta)$ for $n = 40, i = 1,10,20$.

5 Shannon entropy

Let X be a random variable from $NP(\mu, \sigma, \theta)$. Then, the Shannon entropy of X is given by

$$\begin{aligned} H(x) &= -E(\log(f(X; \mu, \sigma, \theta))) = -\int_{-\infty}^{+\infty} f(x; \mu, \sigma, \theta) \times \log \left[\frac{\theta \phi \left(\frac{x-\mu}{\sigma} \right) e^{\theta(1-\Phi \left(\frac{x-\mu}{\sigma} \right))}}{\sigma(e^\theta - 1)} \right] dx \\ &= -\log(\theta) + \log(\sigma(e^\theta - 1)) - \int_{-\infty}^{+\infty} f(x; \mu, \sigma, \theta) \log \left[\phi \left(\frac{x-\mu}{\sigma} \right) \right] dx \\ &\quad - \theta \int_{-\infty}^{+\infty} f(x; \mu, \sigma, \theta) \times \left[1 - \Phi \left(\frac{x-\mu}{\sigma} \right) \right] dx. \end{aligned}$$

We can write the first integral as

$$\int_{-\infty}^{+\infty} f(x; \mu, \sigma, \theta) \log \left[\phi \left(\frac{x-\mu}{\sigma} \right) \right] dx = -\log(\sqrt{2\pi}) - \frac{1}{2\sigma^2} [E(X^2) - 2\mu E(X) + \mu^2],$$

Where $E(X^2)$ and $E(X)$ follow from (3.3) and (3.4). We can write the second integral as

$$\int_{-\infty}^{+\infty} f(x; \mu, \sigma, \theta) \times \left(1 - \Phi \left(\frac{x-\mu}{\sigma} \right) \right) dx = \frac{e^\theta (\theta - 1) - 1}{\theta(e^\theta - 1)}.$$

Finally, we obtain

$$\begin{aligned} H(x) &= -\log(\theta) + \log(\sigma(e^\theta - 1)) + \log(\sqrt{2\pi}) \\ &\quad + \frac{1}{2\sigma^2} [E(X^2) - 2\mu E(X) + \mu^2] - \frac{e^\theta (\theta - 1) - 1}{(\theta(e^\theta - 1))}. \end{aligned}$$

6 Estimation and inference

Let x_1, \dots, x_n be n observations from $NP(\mu, \sigma, \theta)$ and $\Psi = (\mu, \sigma, \theta)^T$ be the parameter vector. The log-likelihood function is given by

$$l_n(\Psi) = n \log(\theta) - n \log(\sigma(e^\theta - 1)) + \sum_{i=1}^n \log \left(\phi \left(\frac{x_i - \mu}{\sigma} \right) \right) + \theta \sum_{i=1}^n \left(1 - \Phi \left(\frac{x_i - \mu}{\sigma} \right) \right).$$

The maximum likelihood estimation (MLE) of Ψ , say $\hat{\Psi}$, 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

$$\frac{\partial l_n(\Psi)}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) + \frac{\theta}{\sigma} \sum_{i=1}^n \phi \left(\frac{x_i - \mu}{\sigma} \right), \quad (6.1)$$

$$\frac{\partial l_n(\Psi)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2 + \frac{\theta}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \phi \left(\frac{x_i - \mu}{\sigma} \right), \quad (6.2)$$

$$\frac{\partial l_n(\Psi)}{\partial \theta} = \frac{n}{\theta} - n(1 - e^{-\theta})^{-1} + \sum_{i=1}^n \left(1 - \Phi \left(\frac{x_i - \mu}{\sigma} \right) \right). \quad (6.3)$$

Clearly, MLEs cannot be obtained in closed forms. The observed information matrix is obtained for approximate confidence intervals and hypothesis tests of on the model parameters. The 3×3 observed information matrix is given by

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

where

$$\begin{aligned}
 I_{\mu\mu} &= -\frac{n}{\sigma^2} + \frac{\theta}{\sigma^3} \sum_{i=1}^n (x_i - \mu) \phi\left(\frac{x_i - \mu}{\sigma}\right), \\
 I_{\mu\sigma} &= -\frac{2}{\sigma^3} \sum_{i=1}^n (x_i - \mu) - \frac{\theta}{\sigma^2} \sum_{i=1}^n \phi\left(\frac{x_i - \mu}{\sigma}\right) + \frac{\theta}{\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 \phi\left(\frac{x_i - \mu}{\sigma}\right), \\
 I_{\mu\theta} &= \frac{1}{\sigma} \sum_{i=1}^n \phi\left(\frac{x_i - \mu}{\sigma}\right), \\
 I_{\sigma\sigma} &= \frac{n}{\sigma^2} - \frac{3}{\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 - \frac{2\theta}{\sigma^3} \sum_{i=1}^n (x_i - \mu) \phi\left(\frac{x_i - \mu}{\sigma}\right) + \frac{\theta}{\sigma^5} \sum_{i=1}^n (x_i - \mu)^3 \phi\left(\frac{x_i - \mu}{\sigma}\right), \\
 I_{\sigma\theta} &= \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \phi\left(\frac{x_i - \mu}{\sigma}\right), \\
 I_{\theta\theta} &= -\frac{n}{\theta^2} + ne^{-\theta}(1 - e^{-\theta})^{-2}.
 \end{aligned}$$

It is well-known that under regularity conditions, the asymptotic distribution of $\sqrt{n}(\hat{\Psi} - \Psi)$ is $N_3(0, J_n(\Psi)^{-1})$, where $J_n(\Psi) = \lim_{n \rightarrow \infty} n^{-1} I_n(\Psi)$. Therefore, an $100(1 - \gamma)$ asymptotic confidence interval for each parameter Ψ_r is given by

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

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

The solution of the three non-linear normal equations in (6.1)-(6.3) is needed using a numerical method. We may use the standard methods such as Newton-Raphson, Nelder-Mead and BFGS, but they have their usual problem of convergence. If the initial guesses are not close to the optimal value, the iteration may not converge. Due to this reason, we propose to use EM-algorithm to compute the MLEs. Suppose, $\{(x_1, z_1), \dots, (x_n, z_n)\}$ is a random sample of size n from (X, N) . We define a hypothetical complete-data distribution with a joint probability density function in the form

$$g(z, x; \Psi) = \frac{\theta^z}{\sigma z! (e^\theta - 1)} z \phi\left(\frac{x - \mu}{\sigma}\right) \left(1 - \Phi\left(\frac{x - \mu}{\sigma}\right)\right)^{z-1},$$

where $\mu \in \mathbb{R}$, $\sigma > 0$, $\theta \in \mathbb{R} - \{0\}$, $x \in \mathbb{R}$ and $z \in \mathbb{N}$. The probability density function of Z given $X = x$ is given by

$$g(z | x) = \frac{g(z, x; \Psi)}{f(x)} = \frac{\theta^{z-1} z \left(1 - \Phi\left(\frac{x - \mu}{\sigma}\right)\right)^{z-1}}{z! e^{\theta(1 - \Phi(\frac{x - \mu}{\sigma}))}},$$

and since $(1 + \theta)e^\theta = \sum_{z=1}^{\infty} \frac{z^2 \theta^{z-1}}{z!}$, the expected value of $Z | X = x$, is given by

$$\begin{aligned}
 E(Z | X = x) &= \sum_{z=1}^{\infty} \frac{\theta^{z-1} z^2 \left(1 - \Phi\left(\frac{x - \mu}{\sigma}\right)\right)^{z-1}}{z! e^{\theta(1 - \Phi(\frac{x - \mu}{\sigma}))}} \\
 &= \frac{1}{e^{\theta(1 - \Phi(\frac{x - \mu}{\sigma}))}} \sum_{z=1}^{\infty} \frac{z^2 \left(\theta \left(1 - \Phi\left(\frac{x - \mu}{\sigma}\right)\right)\right)^{z-1}}{z!} \\
 &= 1 + \theta \left(1 - \Phi\left(\frac{x - \mu}{\sigma}\right)\right).
 \end{aligned}$$

By using the maximum likelihood estimation over Ψ , with the missing Z 's replaced by their conditional expectations given above, the M-step of EM cycle is completed. The log-likelihood of the model parameters for the complete data set is

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

The components of the score function $\left(\frac{\partial l_n^*}{\partial \mu}, \frac{\partial l_n^*}{\partial \sigma}, \frac{\partial l_n^*}{\partial \theta}\right)$ are given by

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

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

$$\begin{aligned} \frac{1}{\hat{\sigma}^{(h)}} \sum_{i=1}^n \left(x_i - \hat{\mu}^{(h+1)}\right) + \sum_{i=1}^n \left(\hat{z}_i^{(h)} - 1\right) \frac{\phi\left(\frac{x_i - \hat{\mu}^{(h+1)}}{\hat{\sigma}^{(h)}}\right)}{1 - \Phi\left(\frac{x_i - \hat{\mu}^{(h+1)}}{\hat{\sigma}^{(h)}}\right)} &= 0, \\ n - \frac{1}{\left(\hat{\sigma}^{(h+1)}\right)^2} \sum_{i=1}^n \left(x_i - \hat{\mu}^{(h)}\right)^2 - \frac{1}{\hat{\sigma}^{(h+1)}} \sum_{i=1}^n \left(\hat{z}_i^{(h)} - 1\right) \frac{\left(x_i - \hat{\mu}^{(h)}\right) \phi\left(\frac{x_i - \hat{\mu}^{(h)}}{\hat{\sigma}^{(h+1)}}\right)}{1 - \Phi\left(\frac{x_i - \hat{\mu}^{(h)}}{\hat{\sigma}^{(h+1)}}\right)} &= 0, \\ \hat{\theta}^{(h+1)} &= \frac{e^{\hat{\theta}^{(h+1)}} - 1}{ne^{\hat{\theta}^{(h+1)}}} \sum_{i=1}^n \hat{z}_i^{(h)}, \end{aligned}$$

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

$$\hat{z}_i^{(h)} = 1 + \hat{\theta}^{(h)} \left(1 - \Phi\left(\frac{x_i - \hat{\mu}^{(h)}}{\hat{\sigma}^{(h)}}\right)\right).$$

In the rest of this section, we verify the performance of the proposed estimator of μ, σ and θ of the proposed EM method for NP distribution. We simulate 1000 times under the NP distribution with different sets of parameters and sample sizes $n = 50, 100, 300$ 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 NP distribution are reported in Tables 2. 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.

Table 2: 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 NP distribution.

n	(μ, σ, θ)	Average estimators			RMSE			SE			Cov		
		$\hat{\mu}$	$\hat{\sigma}$	$\hat{\theta}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\theta}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\theta}$	$(\hat{\mu}, \hat{\sigma})$	$(\hat{\mu}, \hat{\theta})$	$(\hat{\sigma}, \hat{\theta})$
50	(0.0, 1.0, -1.0)	-0.0741	1.0103	-1.2628	0.5032	0.1301	0.2982	0.2227	0.1235	0.6131	-0.0118	0.1376	-0.02821
	(0.0, 1.0, 0.5)	-0.0493	0.8693	0.7323	0.1010	0.1430	0.3171	0.0882	0.0580	0.2160	0.0021	0.0104	0.0020
	(0.0, 1.0, 1.0)	0.0276	0.9979	1.1047	0.2004	0.0747	0.6387	0.1986	0.0747	0.6303	0.0058	0.1080	0.0134
	(0.0, 1.0, 2.0)	-0.0562	0.9922	1.8759	0.4116	0.1322	1.5530	0.4079	0.1321	1.5488	0.0371	0.5852	0.1245
	(0.0, 1.0, 5.0)	-0.0142	0.9864	5.5179	0.5468	0.1795	3.0369	0.5469	0.1790	2.9940	0.0876	1.5472	0.4250
100	(0.0, 1.0, -1.0)	-0.0667	1.0073	-1.1713	0.3127	0.2112	1.454	0.2014	0.0667	0.6234	-0.0073	0.1285	-0.0214
	(0.0, 1.0, 0.5)	-0.0277	0.9054	0.6971	0.0912	0.1012	0.2878	0.0869	0.0360	0.2099	0.0016	0.0128	0.0015
	(0.0, 1.0, 1.0)	0.0180	1.0043	1.0677	0.1710	0.0482	0.5787	0.1701	0.0480	0.5750	0.0044	0.0915	0.0131
	(0.0, 1.0, 2.0)	-0.0283	0.9947	1.9266	0.3463	0.1058	1.3190	0.3453	0.1057	1.3176	0.0287	0.4315	0.1016
	(0.0, 1.0, 5.0)	0.0169	0.9946	5.2860	0.5058	-0.0054	0.2860	0.5057	0.1609	2.6908	0.0761	1.3104	0.3766
300	(0.0, 1.0, -1.0)	-0.0561	1.0016	-1.0981	0.2873	0.1897	1.372	0.1871	0.0424	0.5821	-0.0048	0.1201	-0.0166
	(0.0, 1.0, 0.5)	0.0007	0.9311	0.6768	0.0662	0.0705	0.2596	0.0662	0.0150	0.1902	0.0008	0.0102	0.0016
	(0.0, 1.0, 1.0)	0.0197	1.0040	1.0824	0.1568	0.0373	0.5492	0.1556	0.0371	0.5432	0.0032	0.0810	0.0103
	(0.0, 1.0, 2.0)	0.0035	1.0046	2.0353	0.2687	0.0739	1.0202	0.2688	0.0738	1.0201	0.0167	0.2663	0.0605
	(0.0, 1.0, 5.0)	-0.0534	0.9845	4.8726	0.3311	0.1002	1.6897	0.3270	0.0991	1.6857	0.0313	0.5414	0.1563
500	(0.0, 1.0, -1.0)	-0.0424	1.0040	-1.0665	0.2754	0.1765	1.381	0.1769	0.0382	0.5671	-0.0041	0.1128	-0.0102
	(0.0, 1.0, 0.5)	-0.0068	0.9237	0.6742	0.0554	0.0794	0.2620	0.0550	0.0219	0.1958	0.0006	0.0094	0.0018
	(0.0, 1.0, 1.0)	0.0001	0.9873	1.0392	0.2196	0.1055	0.6364	0.2197	0.1048	0.6355	0.0088	0.1064	0.0161
	(0.0, 1.0, 2.0)	0.0019	1.0033	2.0232	0.2345	0.0632	0.8698	0.2346	0.0631	0.8699	0.0128	0.1996	0.0455
	(0.0, 1.0, 5.0)	-0.0208	0.9937	5.0709	0.3311	0.1002	1.6897	0.3270	0.0991	1.6857	0.0313	0.5414	0.1563

7 Applications

In this section, the NP distribution is fitted to three real data sets and also compared the fitted NP with two relative models, normal (N) and skew-normal (SN) distributions with pdf $\frac{2}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \Phi\left(\alpha \frac{x-\mu}{\sigma}\right)$, to show the superiority of the NP distribution. The first data set concerning the plasma ferritin concentration 102 male and 100 female athletes collected at the Australian Institute of Sport. We estimate parameters by numerically maximizing the likelihood function. The variance covariance matrix of the MLEs under the NP distribution is computed as

$$\begin{bmatrix} 0.2844 & 0.08030 & 0.01995 \\ 0.08030 & 0.0458 & 0.0045 \\ 0.01995 & 0.0045 & 0.0018 \end{bmatrix}.$$

The MLEs of the parameters, -2log-likelihood, AIC (Akaike Information Criterion), the Kolmogorov-Smirnov test statistic (K-S) and the associated p-value are displayed in Table 3 for this data set. The results for these data set show that the NP distribution provides a better fit to this data set than the N and SN distributions. Also this conclusion is confirmed from the plots of the fitted densities in Figure 5.

Table 3: MLEs, -2 Log L, K-S, p-value and AIC for plasma ferritin concentration.

Dist	MLE	-2 Log L	K-S	p-value	AIC
NP	$\hat{\mu} = 111.7220, \hat{\sigma} = 49.9121, \hat{\theta} = 3.0089$	2106.9180	0.0824	0.1288	2112.9180
N	$\hat{\mu} = 76.8762, \hat{\sigma} = 47.3835$	2131.994	0.1217	0.0050	2135.994
SN	$\hat{\mu} = 75.9954, \hat{\sigma} = 47.3919, \hat{\alpha} = 0.0234$	2131.994	0.1217	0.0050	2137.994

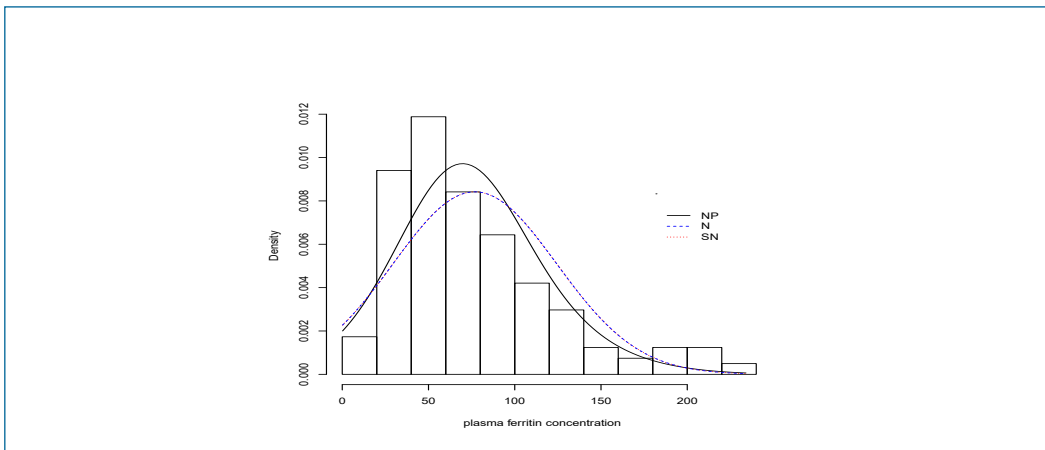


Figure 5: Plots of fitted NP, N and SN for plasma ferritin concentration data.

The second data set is given by Birnbaum and Saunders (1969) that refers to fatigue life of 6061-T6 aluminium coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second. The data set consists of 100 observations. The variance covariance matrix of the MLEs under the NP distribution is computed as

$$\begin{bmatrix} 0.0134 & 0.0042 & 0.0231 \\ 0.0042 & 0.0015 & 0.0068 \\ 0.0231 & 0.0068 & 0.0427 \end{bmatrix}.$$

Table 4 gives the MLEs of the parameters, -2log-likelihood, AIC, the K-S test statistic and the associated P-value for the second data set. The fitted densities functions of NP, N and SN models is displayed in Figure 6. The results for this data set show that the NP distribution is a good competitor for the normal and SN distributions. Also the plots of the densities in Figure 6 confirmed this conclusion.

Table 4: MLEs, $-2\text{Log}L$, K-S, p-value and AIC for time between failures data.

Dist	MLE	-2 Log L	K-S	p-value	AIC
NP	$\hat{\mu} = 9.6705, \hat{\sigma} = 2.7745, \hat{\theta} = 4.6120$	442.0941	0.0481	0.9749	446.0941
N	$\hat{\mu} = 6.8780, \hat{\sigma} = 2.2499$	444.1903	0.0687	0.7325	448.1903
SN	$\hat{\mu} = 6.8362, \hat{\sigma} = 2.2300, \hat{\alpha} = -0.0018$	444.1903	0.0701	0.7012	450.1903

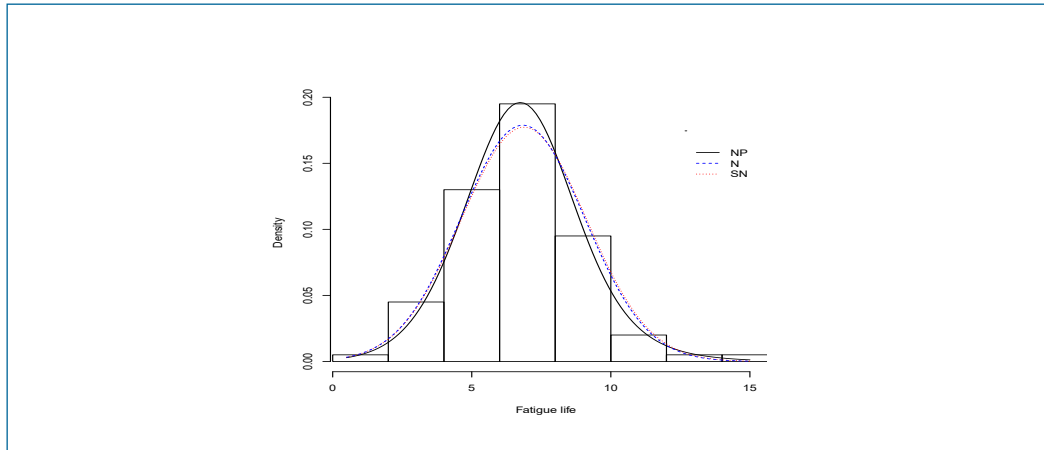


Figure 6: Plots of fitted NP, N and SN for fatigue life data.

The third data set from Bjerkedal (1960), represents the survival time in days of 72 guinea pigs infected with virulent tubercle bacilli. The variance covariance matrix of the MLEs under the NP distribution is computed as

$$\begin{bmatrix} 11.4618 & 3.3416 & 0.3748 \\ 3.3416 & 1.8285 & 0.0900 \\ 0.3748 & 0.0900 & 0.0018 \end{bmatrix}.$$

Table 5 gives the MLEs of the parameters, $-2\log$ -likelihood, AIC, the K-S test statistic and the associated P-value for the third data set. The fitted densities functions of NP, N and SN models is displayed in Figure 7. The results for this data set show that the NP distributions yield the best fit among the N and SN distributions. Also the plots of the densities in Figure 6 confirmed this conclusion.

8 Concluding remarks

In this paper based on compounding approach, a new three-parameter normal-Poisson was developed. The proposed NP is an alternative to the Azzalini skew-normal distribution for fitting skewed data. 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 NP distribution to three real data sets to show the potential of the new proposed distribution.

Table 5: MLEs, $-2\text{Log}L$, K-S, P-value and AIC for guinea pigs data.

Dist	MLE	-2 Log L	K-S	P-value	AIC
NP	$\hat{\mu} = 256.3517, \hat{\sigma} = 110.7858, \hat{\theta} = 3.1438$	431.9072	0.09718	0.5062	869.8144
N	$\hat{\mu} = 176.8334, \hat{\sigma} = 102.7445$	435.6852	0.1299	0.1762	875.3704
SN	$\hat{\mu} = 174.9396, \hat{\sigma} = 102.7617, \hat{\alpha} = 0.0232$	435.6852	0.9478	0	877.3703

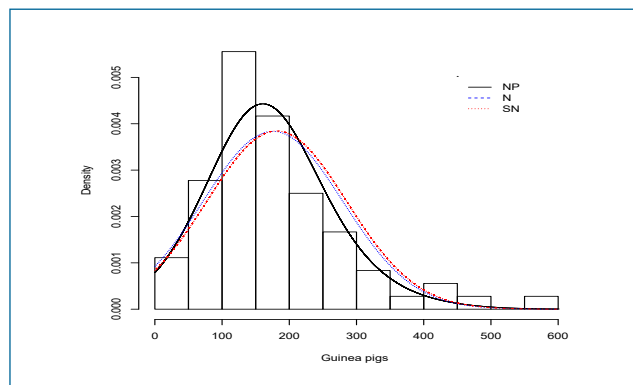


Figure 7: Plots of fitted NP, N and SN for air conditioning system data.

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.

Referências

- B.C. Arnold, R.J. Beaver, Skewed multivariate models related to hidden truncation and/or selective reporting (with discussion), *Test* 11 (1) (2002) 7-5
- A. Azzalini, A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics* 12 (1985) 171-178.
- A. Azzalini, Further results on a class of distributions which includes the normal ones, *Statistica XLVI* (2) (1986) 199-208.
- A. Azzalini, A. Capitanio, Statistical applications of the multivariate skew-normal distribution, *Journal of the Royal Statistical Society Series B* 61 (3) (1999) 579-602.
- A. Azzalini, M. Chiogna, Some results on the stress-strength model for skew-normal variates, *Metron* 3 (2004) 315-326.
- A. Azzalini, A.D. Valle, The multivariate skew-normal distribution, *Biometrika* 83 (4) (1996) 715-726.
- N. Balakrishnan, Discussion on "Skew multivariate normal models related to hidden truncation and / or selective reporting" by B.C. Arnold and R. J. Beaver, *Test* 11 (2002) 37-39.
- T. Bjerkedal, Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *American Journal of Epidemiology* 72 (1960) 130-148.
- Z.W. Birnbaum, S.C. Saunders, Estimation for a family of life distributions with applications to fatigue, *Journal of Applied Probability* 6 (1969) 328-347.
- M. Branco, D.K. Dey, A general class of multivariate elliptical distribution, *Journal of Multivariate Analysis* 79 (2001) 99-113.
- D. DElal-Olivero, Alpha-skew-normal distribution, *Proyecciones* 29 (2010) 224-240.
- N.A. Henze, A probabilistic representation of the skew-normal distribution, *Scandinavian Journal of Statistics* 13 (1986) 271-275.
- A. Jamalizadeh, N. Balakrishnan, 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 (2010) 1412-1427.
- N. Loperfido, Quadratic forms of skew-normal and random vectors, *Statistics and Probability Letters* 54 (2001) 381-387.
- E. Mahmoudi and H. Mahmoodian, Normal power series class of distributions: Model, properties and applications, *Electronic Journal of Applied Statistical Analysis* 10(2) (2017) 512-541.

- R. Roozegar, S. Nadarajah, The power series skew normal class of distributions, *Communication in Statistics-Theory and Methods* 46(22) (2017) 11404-11423.
- S. Nadarajah, V. Nassiri, A. Mohammadpour, Truncated-exponential skew-symmetric distributions, *Statistics* 48 (2014) 872-895.
- A. Noack, A class of random variables with discrete distributions, *Annals of Mathematical Statistics* 21 (1950) 127-132.
- M. Sharafi, J. Behboodian, The Balakrishnan skew-normal density, *Statistical Papers* 49 (2008) 769-778.

Eisa Mahmoudi

Department of Statistics, Yazd University, Yazd, Iran

Email: emahmoudi@yazd.ac.ir

author's contribution:

Hamed Mahmoodian

Department of Statistics, Yazd University, Yazd, Iran

Email: hamed_mahmoodian@yahoo.com

author's contribution:

Fatemeh Esfandiari

Department of Statistics, Yazd University, Yazd, Iran

Email: f.esfandiyari89@gmail.com

author's contribution::