Geometric Skew Normal Distribution

Debasis Kundu[1]

Abstract

In this article we introduce a new three parameter skewed distribution of which normal distribution is a special case. This distribution is obtained by using geometric sum of independent identically distributed normal random variables. We call this distribution as the geometric skew normal distribution. Different properties of this new distribution have been investigated. The probability density function of geometric skew normal distribution can be unimodal or multimodal, and it always has an increasing hazard rate function. It is an infinite divisible distribution, and it can have heavier tails. The maximum likelihood estimators cannot be obtained in explicit forms. The EM algorithm has been proposed to compute the maximum likelihood estimators of the unknown parameters. One data analysis has been performed for illustrative purposes. We further consider multivariate geometric skew normal distribution and explore its different properties. The proposed multivariate model induces a multivariate Lévy process, and some properties of this multivariate process have been investigated. Finally we conclude the paper.

Key Words and Phrases: Characteristic function; moment generating function; infinite divisible; maximum likelihood estimators; EM algorithm; Fisher information matrix; Lévy process.

[1] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. E-mail: kundu@iitk.ac.in, Phone no. 91-512-2597141, Fax no. 91-512-2597500.

1 Introduction

In recent times, the skew normal distribution proposed by Azzalini (1985) has received considerable attention because of its flexibility. The probability density function (PDF) of Azzalini's skew normal (ASN) distribution has the following form;

$$f(x) = 2\phi(x)\Phi(\lambda x); \quad -\infty < x < \infty, \quad -\infty < \lambda < \infty, \tag{1}$$

where $\phi(x)$ and $\Phi(x)$ denote the standard normal PDF and the standard normal cumulative distribution function (CDF), respectively. It has several interesting properties, and normal distribution becomes a particular member of this class of distributions. It has an unimodal density function having both positive and negative skewness present. Moreover, ASN law has a very nice interpretation in terms of hidden truncation model, see for example Arnold et al. (1993) and also Arnold and Beaver (2000) in this respect. Due to flexibility of its PDF, this model has been used quite effectively in analyzing non-symmetric data sets from different fields. The ASN model has a natural multivariate generalization. It may be mentioned that although ASN distribution has several interesting properties, it exhibits one problem in developing statistical inference procedure. It is observed in many cases that the maximum likelihood estimators (MLEs) of the unknown parameters of the ASN model may not exist, see for example Gupta and Gupta (2004). The problem is more severe for the multivariate case.

In this paper, we consider a new three parameter skewed normal distribution based on the geometric and normal distributions. The basic idea is as follows. Consider a random variable X, such that

$$X \stackrel{d}{=} \sum_{i=1}^{N} X_i, \tag{2}$$

where '$\stackrel{d}{=}$' means equal in distribution, $\{X_i : i = 1, 2, \ldots, \}$ is a sequence of independent and identically distributed (i.i.d.) normal random variables, and N is a geometric distribution with support on the positive integers only. Moreover, N and X_i's are independently distributed. We call this new distribution as geometric skew normal (GSN) distribution. The idea came from Kuzobowski and Panorska (2005) and Barreto-Souza (2012), where the authors introduced bivariate exponential geometric and bivariate gamma geometric distributions, respectively, along the same line.

We discuss properties of the proposed GSN distribution. It is a skewed version of the normal distribution of which normal distribution is a particular member. The PDF of GSN distribution can be unimodal or multimodal. The GSN distribution can be written as an infinite mixture of normal distributions, and it always has an increasing hazard function. The moment generating function can be obtained in explicit forms, and all the moments can be expressed in terms of the moments of normal distributions. It is an infinitely divisible distribution, and geometric stable. The PDF of GSN distribution can be symmetric with heavier tails. The generation of random samples from a GSN distribution is quite straight forward, hence the simulation experiments can be performed quite easily.

The proposed GSN distribution has three parameters. The maximum likelihood estimators (MLEs) of the unknown parameters of a GSN distribution cannot be obtained in explicit forms. One needs to solve three non-linear equations simultaneously. We propose to use EM algorithm to compute the MLEs of the unknown parameters. It is observed that the EM algorithm can be implemented quite conveniently. At each 'E'-step, the corresponding 'M'-step can be obtained in explicit form. We address some testing of hypotheses problems. The analysis of one real data set has been performed for illustrative purposes, and it is observed that the proposed model provides a good fit to the data set.

We further extend the model to the multivariate case. Along the same line, we define multivariate geometric skew normal (MGSN) distribution. Different properties of MGSN have been explored. It is observed that multivariate normal distribution is a special case of

MGSN. It is also an infinite divisible distribution, and it is geometric stable. The estimation of the unknown parameters using EM algorithm also can be obtained along the same line. It is observed that MGSN distribution can be a good alternative to the multivariate skew normal distribution proposed by Azzalini and Dalla Valle (1996). Since the MGSN distribution is an infinite divisible distribution, it induces a multivariate Lévy process with marginals as normal Lévy processes and which has MGSN motion as a special case. We discuss properties of the multivariate Lévy process, and finally conclude the paper.

Rest of the paper is organized as follows. In Section 2, we discuss different properties of the GSN distribution. Inference procedures of the unknown parameters are discussed in Section 3. The analysis of a real data set is performed in Section 4. In Section 5, we introduce MGSN distribution, and the induced multivariate Lévy process and discuss some of their properties. Finally in Section 6, we conclude the paper.

2 Geometric Skew Normal Distribution

2.1 Definition, PDF, CDF, Generation

We will use the following notations in this paper. A normal random variable with mean μ and variance σ^2 will be denoted by $\mathrm{N}(\mu, \sigma^2)$. A geometric random variable with parameter p will be denoted by $\mathrm{GE}(p)$, and it has the probability mass function (PMF); $p(1-p)^{n-1}$, for $n = 1, 2, \ldots$. Now we define GSN distribution with parameters μ, σ and p as follows.

Definition: Suppose $N \sim \mathrm{GE}(p)$, $\{X_i : i = 1, \ldots, \}$ are i.i.d. $\mathrm{N}(\mu, \sigma^2)$ random variables, and N and X_i's are independently distributed. Define

$$X \stackrel{d}{=} \sum_{i=1}^{N} X_i,$$

then X is said to be GSN random variable with parameters μ, σ and p. It will be denoted as GSN(μ, σ, p).

Throughout this paper we will be using the convention $0^0 = 1$. The joint PDF, $f_{X,N}(\cdot,\cdot)$ of (X, N) is given by

$$f_{X,N}(x,n) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi n}} e^{-\frac{1}{2n\sigma^2}(x-n\mu)^2} p(1-p)^{n-1} & \text{if} \quad 0 < p < 1 \\ \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} & \text{if} \quad p = 1; \end{cases} \tag{3}$$

for $-\infty < x < \infty$, $\sigma > 0$ and for any positive integer n. The joint CDF associated with (3) becomes

$$\begin{aligned} P(X \le x, N \le n) &= \sum_{k=1}^{n} P(X \le x, N = k) \\ &= \sum_{k=1}^{n} P(X \le x | N = k) P(N = k) \\ &= p \sum_{k=1}^{n} \Phi\left(\frac{x - k\mu}{\sigma\sqrt{k}}\right)(1-p)^{k-1}. \end{aligned} \tag{4}$$

The CDF of X can be obtained as

$$F_X(x) = P(X \le x) = \sum_{k=1}^{\infty} P(X \le x, N = k) = p \sum_{k=1}^{\infty} \Phi\left(\frac{x - k\mu}{\sigma\sqrt{k}}\right)(1-p)^{k-1}. \tag{5}$$

Hence the PDF of X becomes

$$f_X(x) = \sum_{k=1}^{\infty} \frac{p}{\sigma\sqrt{k}} \phi\left(\frac{x - k\mu}{\sigma\sqrt{k}}\right)(1-p)^{k-1}. \tag{6}$$

When $\mu = 0$ and $\sigma = 1$, we say that X has standard GSN distribution with PDF

$$f_X(x) = p \sum_{k=1}^{\infty} \frac{1}{\sqrt{k}} \phi\left(\frac{x}{\sqrt{k}}\right)(1-p)^{k-1}. \tag{7}$$

The standard GSN distribution is a symmetric distribution around 0, for all values of p. When $p = 1$, $X \sim \mathrm{N}(\mu, \sigma^2)$. From (6) it follows that the GSN law is a geometric mixture of normal random variables. The PDFs of GSN can take different shapes. Figure 1 provides the PDFs of GN distribution for different values of μ and p, when $\sigma = 1$. The PDF of standard GSN distribution for different values of p are provided in Figure 2.

It is clear from Figure 1 that the PDFs of GSN law can take different shapes depending on the values of p and μ. For $\mu > 0$ it is positively skewed, and for $\mu < 0$ it is negatively skewed. If p is very small it is more skewed either positive or negative depending on the values of μ, and as p increases the skewness decreases. If p is 1, it is the normal PDF. The shape of the PDF of GSN distribution is very similar with the shape of the PDF of ASN distribution in some cases. The GSN distribution can have bimodal or multimodal PDF. This is different from the ASN distribution, which is always unimodal. From Figure 2 it is clear that for the standard SGN distribution the PDF is always symmetric and for smaller p it has heavier tails. It seems that GSN is a more flexible than ASN distribution.

The generation from a GSN distribution is quite straight forward using the geometric and normal distribution. The following algorithm can be easily used to generate samples from a GSN(μ, σ, p).

- Step 1: Generate m from GE(p).
- Step 2: Generate x from N$(m\mu, m\sigma^2)$, and x is the required sample.

The hazard function of GSN distribution is an increasing function for all values of μ, σ and p. It simply follows as the hazard function of a normal distribution is an increasing function, and GSN is an infinite mixture of normal distributions.

2.2 Moment Generating Function and Infinite Divisibility

If $X \sim$ GSN$(\mu\sigma, p)$, then the moment generating function of X becomes,

$$M_X(t) = Ee^{tX} = E\left[E(e^{tX}|N)\right] = E\left[e^{N\mu t + \frac{\sigma^2 N t^2}{2}}\right] = \frac{pe^{(\mu t + \frac{\sigma^2 t^2}{2})}}{1-(1-p)e^{(\mu t + \frac{\sigma^2 t^2}{2})}}, \quad t \in \mathbb{R}. \tag{8}$$

Using (8), every moments can be obtained. For example,

$$E(X) = \frac{\mu}{p}, \qquad V(X) = \frac{\sigma^2 p + \mu^2(1-p)}{p^2}, \tag{9}$$

$$E(X - E(X))^3 = \frac{1-p}{p^3}(\mu^3(2p^2 - p + 2) + 2\mu^2 p^2 + \mu\sigma^2(3-p)p). \tag{10}$$

Alternatively, it can be obtained directly using (3) in terms if infinite series as

$$E(X^m) = p\sum_{n=1}^{\infty}(1-p)^{n-1}c_m(n\mu, n\sigma^2). \tag{11}$$

Here $c_m(n\mu, n\sigma^2) = E(Y^m)$, where $Y \sim \mathrm{N}(n\mu, n\sigma^2)$. Note that c_m can be obtained using confluent hyper geometric function, see for example Johnson, Kotz and Balakrishnan (1995). If $\mu = 0$ and $\sigma = 1$,

$$E(X^m) = pd_m\sum_{n=1}^{\infty}(1-p)^{n-1}n^{m/2}, \tag{12}$$

where $d_m = E(Z^m)$, if $Z \sim \mathrm{N}(0,1)$, and

$$d_m = \begin{cases} 0 & \text{if} \quad m \quad \text{is odd} \\ \frac{2^{m/2}\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{\pi}} & \text{if} \quad m \quad \text{is even.} \end{cases}$$

If $X \sim \mathrm{GSN}(\mu, \sigma, p)$, then skewness can be obtained as

$$\gamma_1 = \frac{(1-p)(\mu^3(2p^2 - p + 2) + 2\mu^2 p^2 + \mu\sigma^2(3-p)p)}{(\sigma^2 p + \mu^2(1-p))^{3/2}}. \tag{13}$$

Now we will show that the GSN law is infinitely divisible. Consider the following random vector when $r = 1/n$,

$$R \stackrel{d}{=} \sum_{i=1}^{1+nT} Y_i, \tag{14}$$

where Y_i's are i.i.d., $Y_i \sim \mathrm{N}(\mu/n, \sigma^2/n)$, and T follows a negative binomial $\mathrm{NB}(r,p)$ distribution with the probability mass function

$$P(T = k) = \frac{\Gamma(k+r)}{k!\Gamma(r)}p^r(1-p)^k, \quad k = 0, 1, 2, \ldots. \tag{15}$$

The moment generating function of R is given by

$$\begin{aligned} M_R(t) = E\left(e^{tR}\right) &= E\left(Ee^{t\sum_{i=1}^{1+nT} Y_i}|T\right) \\ &= \left[\frac{pe^{\mu t + \frac{\sigma^2 t^2}{2}}}{1-(1-p)e^{\mu t + \frac{\sigma^2 t^2}{2}}}\right]^{1/n} = (M_X(t))^{1/n} \quad \text{for} \quad t \in \mathbb{R}, \end{aligned} \tag{16}$$

where $M_X(t)$ is same as defined in (8). Therefore, GSN law is infinitely divisible.

Now we will show that GSN law has geometric stability property. Suppose $\{X_i : i = 1, 2, \ldots\}$ is a sequence of i.i.d. random variables, following the distribution GSN$(\mu, \sigma, \tilde{p})$, and M is an independent GE(q), with $0 < q < 1$, random variable. The moment generating function of

$$\sum_{i=1}^{M} X_i \overset{d}{=} X$$

becomes

$$E\left(e^{tX}\right) = q \sum_{m=1}^{\infty} (1-q)^{m-1} \left[\frac{\tilde{p} e^{\mu t + \frac{\sigma^2 t^2}{2}}}{1 - (1-\tilde{p}) e^{\mu t + \frac{\sigma^2 t^2}{2}}}\right]^m = \left[\frac{\tilde{p} q e^{\mu t + \frac{\sigma^2 t^2}{2}}}{1 - (1-\tilde{p} q) e^{\mu t + \frac{\sigma^2 t^2}{2}}}\right],$$

which is the moment generating function of GSN$(\mu, \sigma, \tilde{p}q)$. Hence $X \sim$ GSN$(\mu, \sigma, \tilde{p}q)$.

The following decomposition of GSN is also possible. Suppose $X \sim$ GSN(μ, σ, p), then

$$X \overset{d}{=} Y + \sum_{i=1}^{Q} Y_i. \tag{17}$$

Here, Q is a Poisson random variable with parameter λ, and it is independent of Z_i's, where $\{Z_i, i = 1, 2, \ldots\}$ is a sequence of i.i.d. random variables having logarithmic distribution with the probability mass function

$$P(Z_1 = k) = \frac{(1-p)^k}{\lambda k}, \quad k = 1, 2, \ldots, \quad \lambda = -\ln p.$$

Moreover, given the sequence of random variables $\{Z_i, i = 1, 2, \ldots\}$, $Y_i | Z_i \sim$ N$(\mu Z_i, \sigma^2 Z_i)$ for $i = 1, 2, \ldots$, and they are independently distributed, $Y \sim$ N(μ, σ^2) and it is independent of all the previous random variables.

To prove (17), the following results will be useful. The probability generating function of Q and Z_1 are as follows;

$$E(t^Q) = e^{\lambda(t-1)}, \quad t \in \mathbb{R} \qquad \text{and} \quad E(t^{Z_1}) = \frac{\ln(1-(1-p)t)}{\ln p}, \quad t < (1-p)^{-1}. \tag{18}$$

The moment generating function of the right hand side of (17) can be derived as

$$\begin{aligned}
E\left[e^{t(Y+\sum_{i=1}^{Q} Y_i)}\right] &= e^{\mu t+\frac{\sigma^2 t^2}{2}} \times E\left[e^{t\sum_{i=1}^{Q} Y_i}\right] \\
&= e^{\mu t+\frac{\sigma^2 t^2}{2}} \times E\left[e^{(\mu t+\frac{\sigma^2 t^2}{2})\sum_{i=1}^{Q} Z_i}\right] \\
&= e^{\mu t+\frac{\sigma^2 t^2}{2}} \times E\left[\frac{\ln(1-(1-p)e^{\mu t+\frac{\sigma^2 t^2}{2}})}{\ln p}\right]^Q = \frac{e^{\mu t+\frac{\sigma^2 t^2}{2}}}{1-(1-p)e^{\mu t+\frac{\sigma^2 t^2}{2}}}
\end{aligned}$$

∎

2.3 Conditional Distributions

Now we provide different conditional distributions which will be useful for further development. Consider (X, N) which has the joint PDF as given in (3), and suppose $m \le n$, be positive integers. The conditional CDF of (X, N) given $N \le n$ is

$$P(X \le x, N \le m | N \le n) = \frac{P(X \le x, N \le m)}{P(N \le n)} = \frac{p}{1-(1-p)^n} \sum_{k=1}^{m} \Phi\left(\frac{x-k\mu}{\sigma\sqrt{k}}\right)(1-p)^{k-1}. \tag{19}$$

From (19), we obtain

$$P(X \le x | N \le n) = \frac{p}{1-(1-p)^n} \sum_{k=1}^{n} \Phi\left(\frac{x-k\mu}{\sigma\sqrt{k}}\right)(1-p)^{k-1}.$$

We further have for $0 \le x \le y$, and $n \in \mathbb{N}$, the conditional CDF of (X, N) given $X \le y$ is

$$P(X \le x, N \le n | X \le y) = \frac{P(X \le x, N \le n)}{P(X \le y)} = \frac{\sum_{k=1}^{n}(1-p)^k \Phi\left(\frac{x-k\mu}{\sigma\sqrt{k}}\right)}{\sum_{k=1}^{\infty}(1-p)^k \Phi\left(\frac{y-k\mu}{\sigma\sqrt{k}}\right)}. \tag{20}$$

We obtain from (20) that

$$P(N \le n | X \le y) = \frac{\sum_{k=1}^{n}(1-p)^k \Phi\left(\frac{y-k\mu}{\sqrt{k}\sigma}\right)}{\sum_{k=1}^{\infty}(1-p)^k \Phi\left(\frac{y-k\mu}{\sqrt{k}\sigma}\right)}. \tag{21}$$

The conditional probability mass function of N given $X = x$, is

$$P(N = n | X = x) = \frac{(1-p)^{n-1} e^{-\frac{1}{2\sigma^2 n}(x-n\mu)^2}/\sqrt{n}}{\sum_{k=1}^{\infty}(1-p)^{k-1} e^{-\frac{1}{2\sigma^2 k}(x-k\mu)^2}/\sqrt{k}}. \tag{22}$$

The conditional expectations become

$$E(N|X = x) = \frac{\sum_{n=1}^{\infty}(1-p)^{n-1}e^{-\frac{1}{2\sigma^2 n}(x-n\mu)^2}/\sqrt{n}}{\sum_{k=1}^{\infty}(1-p)^{k-1}e^{-\frac{1}{2\sigma^2 k}(x-k\mu)^2}/\sqrt{k}}, \tag{23}$$

and

$$E(N^{-1}|X = x) = \frac{\sum_{n=1}^{\infty}(1-p)^{n-1}e^{-\frac{1}{2\sigma^2 n}(x-n\mu)^2}/n^{3/2}}{\sum_{k=1}^{\infty}(1-p)^{k-1}e^{-\frac{1}{2\sigma^2 k}(x-k\mu)^2}/\sqrt{k}}. \tag{24}$$

3 Statistical Inference

3.1 Maximum Likelihood Estimators

Suppose $\{x_1, \ldots, x_n\}$ is a random sample of size n from GSN(μ, σ, p), the log-likelihood function becomes

$$l(\mu, \sigma, p) = \sum_{i=1}^{n} \ln f_X(x_i) = \sum_{i=1}^{n} \ln \left[\sum_{k=1}^{\infty} \frac{p}{\sigma\sqrt{k}} \phi\left(\frac{x_i - k\mu}{\sigma\sqrt{k}}\right)(1-p)^{k-1} \right]. \tag{25}$$

The maximum likelihood estimators (MLEs) of the unknown parameters can be obtained by maximizing the log-likelihood function with respect to the unknown parameters. The normal equations can be obtained by taking derivatives of $l(\mu, \sigma, p)$ with respect to μ, σ and p, respectively, and equating them to 0. Clearly, MLEs cannot be obtained in explicit forms. We propose to use EM algorithm to compute the MLEs.

The basic idea is as follows. Suppose, $\{(x_1, m_1), \ldots, (x_n, m_n)\}$ is a random sample of size n from (X, N). The log-likelihood function without the additive constant, based on the complete sample becomes

$$l_c(\mu, \sigma, p) = -n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \frac{(x_i - m_i\mu)^2}{m_i} + n \ln p + \ln(1-p) \sum_{i=1}^{n}(m_i - 1). \tag{26}$$

Therefore, based on the complete sample, the MLEs of the unknown parameters are as follows:

$$\widehat{\mu} = \frac{\sum_{i=1}^{n} x_i}{\sum_{k=1}^{n} m_k}, \qquad \widehat{\sigma^2} = \frac{1}{n} \sum_{i=1}^{n} \frac{(x_i - m_i\widehat{\mu})^2}{m_i}, \qquad \widehat{p} = \frac{n}{K+n}, \tag{27}$$

where $K = \sum_{i=1}^{n} m_i$. Therefore, it is immediate that based on the complete samples, the MLEs of the unknown parameters can be obtained in explicit forms. Based on the above observations the EM algorithm can be obtained as follows. Let us denote $\mu^{(k)}$, $\sigma^{(k)}$ and $p^{(k)}$ as the estimates of μ, σ and p, respectively, at the k-th stage of the EM algorithm. At the 'E'-step, the 'pseudo' log-likelihood function at the k-th stage can be formed by replacing the missing values with their expectations, and it is as follows;

$$l_s^{(k)}(\mu,\sigma,p) = -n\ln\sigma - \frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} x_i^2 c_i^{(k)} - 2\mu\sum_{i=1}^{n} x_i + \mu^2\sum_{i=1}^{n} c_i^{(k)}\right) + n\ln p + \ln(1-p)\sum_{i=1}^{n}(d_i^{(k)}-1), \tag{28}$$

here $c_i^{(k)}$ and $d_i^{(k)}$ can be obtained from (24) and (23), by replacing x, μ, σ, p with x_i, $\mu^{(k)}$, $\sigma^{(k)}$, $p^{(k)}$, respectively. The 'M'-step can be obtained by maximizing (28) with respect to the unknown parameters. Therefore, $\mu^{(k+1)}$, $\sigma^{(k+1)}$, $p^{(k+1)}$, can be obtained as

$$\mu^{(k+1)} = \frac{\sum_{i=1}^{n} x_i}{\sum_{j=1}^{n} c_j^{(k)}}, \quad \sigma^{(k+1)} = \frac{1}{\sqrt{n}} \times \sqrt{\sum_{i=1}^{n} x_i^2 c_i^{(k)} - 2\mu^{(k+1)}\sum_{i=1}^{n} x_i + (\mu^{(k+1)})^2\sum_{i=1}^{n} c_i^{(k)}}. \tag{29}$$

and

$$p^{(k+1)} = \frac{n}{\sum_{i=1}^{n} d_i^{(k)} + n}. \tag{30}$$

The iteration process should be continued until the convergence is met. The asymptotic distribution of the MLEs can be obtained in a routine manner that is, if $\widehat{\mu}$, $\widehat{\sigma}$ and $\widehat{p}$ denote the MLEs of μ, σ and p, respectively, then

$$\sqrt{n}(\widehat{\mu}-\mu, \widehat{\sigma}-\sigma, \widehat{p}-p) \xrightarrow{d} N_3(0, \mathbf{F}^{-1}), \tag{31}$$

where '$\xrightarrow{d}$' denotes convergence in distribution, and the 3×3 matrix $\mathbf{F}$ is the expected Fisher information matrix.

3.2 Observed Information Matrix

In this section we provide the observed information matrix, which will be useful to construct the asymptotic confidence intervals of the unknown parameters. The observed information matrix is obtained from the EM algorithm using the idea of Louis (1982). We use the same notation as that of Louis (1982). Here the matrix B denotes the negative second derivative of the 'pseudo' log-likelihood function and S is the derivative vector.

$$F_{obs} = B - SS^T.$$

Now we provide the elements of the matrix B and the vector S.

$$\begin{aligned}
B(1,1) &= \frac{\sum_{i=1}^n c_i^{(k)}}{\sigma^2}, \quad B(2,2) = \frac{3}{\sigma^4}\left(\sum_{i=1}^n x_i^2 c_i^{(k)} - 2\mu \sum_{i=1}^n x_i + \mu^2 \sum_{i=1}^n c_i^{(k)}\right) - \frac{n}{\sigma^2} \\
B(3,3) &= \frac{n}{p^2} + \frac{1}{(1-p)^2}\sum_{i=1}^n (d_i^{(k)} - 1), \quad B(1,2) = B(2,1) = \frac{2}{\sigma^3}\left(\sum_{i=1}^n x_i - \mu \sum_{i=1}^n c_i^{(k)}\right) \\
B(1,3) &= B(3,1) = B(2,3) = B(3,2) = 0. \\
S(1) &= \frac{1}{\sigma^2}\left(\sum_{i=1}^n x_i - \mu \sum_{i=1}^n c_i^{(k)}\right), \quad S(2) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\left(\sum_{i=1}^n x_i^2 c_i^{(k)} - 2\mu \sum_{i=1}^n x_i + \mu^2 \sum_{i=1}^n c_i^{(k)}\right) \\
S(3) &= \frac{n}{p} - \frac{1}{1-p}\sum_{i=1}^n (d_i^{(k)} - 1).
\end{aligned}$$

3.3 Testing of Hypotheses

In this section we discuss likelihood ratio test for some hypotheses of interest. We consider the following specific testing that will be of interest.

Test I: $H_0 : \mu = 0$ vs. $H_1 : \mu \neq 0$.

The problem is of interest as it tests whether the distribution is symmetric or not. In this the MLEs of the unknown parameters can be obtained using the EM algorithm as before. Under the null hypothesis, the 'pseudo' log-likelihood function becomes

$$l_{sI}^{(k)}(\sigma, p) = -n \ln \sigma - \frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2 c_i^{(k)} + n \ln p + \ln(1-p)\sum_{i=1}^n (d_i^{(k)} - 1), \tag{32}$$

here $c_i^{(k)}$ and $d_i^{(k)}$ can be obtained from (24) and (23), by replacing x, μ, σ, p with x_i, 0, $\sigma^{(k)}$, $p^{(k)}$, respectively. The 'M'-step can be obtained by maximizing (32) with respect to the unknown parameters. Therefore, $\sigma^{(k+1)}$, $p^{(k+1)}$, can be obtained as

$$\sigma^{(k+1)} = \frac{1}{\sqrt{n}} \times \sqrt{\sum_{i=1}^{n} x_i^2 c_i^{(k)}} \quad \text{and} \quad p^{(k+1)} = \frac{n}{\sum_{i=1}^{n} d_i^{(k)} + n}. \tag{33}$$

Therefore, if $\widehat{\mu}$, $\widehat{\sigma}$ and $\widehat{p}$ denote the MLEs of μ, σ and p, and $\widetilde{\sigma}$ and $\widetilde{p}$ denote the MLEs of σ and p, respectively under the restriction H_0, then

$$2(l(\widehat{\mu}, \widehat{\sigma}, \widehat{p}) - l(0, \widetilde{\sigma}, \widetilde{p})) \longrightarrow \chi_1^2.$$

Test II: $H_0 : p = 1$ vs. $H_1 : p < 1$.

The problem is of interest as it tests whether the distribution is symmetric or not. In this case under the null hypothesis the MLEs of μ and σ become

$$\widetilde{\mu} = \frac{\sum_{i=1}^{n} x_i}{n}, \quad \text{and} \quad \widetilde{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \widetilde{\mu})^2}{n}}.$$

In this case p is in the boundary under the null hypothesis, the standard results do not work. But using Theorem 3 of Self and Liang (1987), it follows that

$$2(l(\widehat{\mu}, \widehat{\sigma}, \widehat{p}) - l(\widetilde{\mu}, \widetilde{\sigma}, 1)) \longrightarrow \frac{1}{2} + \frac{1}{2}\chi_1^2.$$

4 Data Analysis

In this section we analyze one data set to see the effectiveness of the proposed model. This data set represents the survival times of guinea pigs injected with different doses of tubercle bacilli, and it has been obtained from Bjerkedal (1960). Typically guinea pigs are chosen for tuberculosis experiments because of their high susceptibility. The 72 observations are presented in Table 1.

The mean, standard deviation and the coefficient of skewness are calculated as 99.82, 80.55 and 1.80, respectively. The skewness measure indicates that the data are positively

Table 1: Guinea pig data set.

12	15	22	24	24	32	32	33	34
38	38	43	44	48	52	53	54	54
55	56	57	58	58	59	60	60	60
60	61	62	63	65	65	67	68	70
70	72	73	75	76	76	81	83	84
85	87	91	95	96	98	99	109	110
121	127	129	131	143	146	146	175	175
211	233	258	258	263	297	341	341	376

skewed. We have plotted the histogram in Figure 3. The histogram clearly indicates that the data are right skewed, and the sample skewness also indicates that.

Before analyzing the data we divide all the observations by 50 for computational purposes, it is not going to affect the inference procedure. We obtain the MLEs of the unknown parameters of the GSN model and they are as follows;

$$\widehat{p} = 0.5657, \quad \widehat{\sigma} = 0.5975 \quad \widehat{\mu} = 1.1311,$$

and the associated log-likelihood value becomes -113.4698. The corresponding 95% confidence intervals of p, σ and μ become (0.5657$\mp$0.1987), (0.5975$\mp$0.2312) and (1.1311$\mp$0.4217), respectively.

Now the natural question is whether the proposed model is a good fit to the data set or not. We provide the empirical survival function and the fitted survival function in Figure 4 and also the histogram of the data along with the fitted PDF in Figure 5. The Kolmogorov-Smirnov (KS) test statistic and the associated p values are 0.1118 and 0.3283, respectively. All of these indicate that the proposed model provides a good fit to the data set. Now for comparison purposes, we have fitted three parameter ASN model to the same data set. The three-parameter ASN model has the following PDF

$$f(x; \lambda, \sigma, \mu) = 2\phi((x-\mu)/\sigma)\Phi(\lambda(x-\mu)/\sigma). \tag{34}$$

The MLEs of the unknown parameters are as follows: $\widehat{\lambda} = 19.7001$, $\widehat{\sigma} = 2.3299$ and $\widehat{\mu} = 0.3099$. The associated log-likelihood value becomes -115.5862. The Kolmogorov-Smirnov distance between the empirical and the fitted distribution functions is 0.1314, and the corresponding p value is 0.2887. Hence based on the log-likelihood and KS test statistic values we can conclude that GSN provides a better fit than the ASN model to this data set.

5 Generalizations

5.1 Multivariate Geometric Skew Normal Distribution

In this section we introduce the multivariate geometric skew normal (MGSN) distribution, which can be a good alternative to the Azzalini's multivariate skew normal (AMSN) distribution. We use the following notation. An m-variate random vector with mean vector $\boldsymbol{\mu}$ and the dispersion matrix $\boldsymbol{\Sigma}$ will be denoted by $\mathrm{N}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The corresponding PDF and CDF at the point $\boldsymbol{x}$, will be denoted by $\phi_m(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\Phi_m(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$, respectively. Now we define MGSN distribution.

Definition: Suppose $N \sim \mathrm{GE}(p)$, $\{\boldsymbol{X}_i : i = 1, 2, \ldots, \}$ are i.i.d. $\mathrm{N}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ random vectors, N and $\boldsymbol{X}_i$'s are independently distributed. Define

$$\boldsymbol{X} \stackrel{d}{=} \sum_{i=1}^{N} \boldsymbol{X}_i,$$

then $\boldsymbol{X}$ is said to be m-variate geometric skew normal distribution with parameters p, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, and it will be denoted by $\mathrm{MGSN}(m, p, \boldsymbol{\mu}, \boldsymbol{\Sigma})$.

The joint PDF, $f_{\boldsymbol{X},N}(\cdot, \cdot)$ of $(\boldsymbol{X}, N)$ is given by

$$f_{\boldsymbol{X},N}(\boldsymbol{x}, n) = \frac{p(1-p)^{n-1}}{(2\pi)^{m/2}|\boldsymbol{\Sigma}|^{1/2}\sqrt{n}} \; e^{-\frac{1}{2n}(\boldsymbol{x}-n\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-n\boldsymbol{\mu})} \; ;$$
$$\text{for} \quad \boldsymbol{x} \in \mathbb{R}^m, \quad \boldsymbol{\mu} \in \mathbb{R}^m, \quad \boldsymbol{\sigma} > 0, \quad 0 < p \leq 1. \qquad (35)$$

Therefore, the PDF of $\boldsymbol{X}$ becomes

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \sum_{k=1}^{\infty} f_{\boldsymbol{X},N}(\boldsymbol{x}, k) = \sum_{k=1}^{\infty} \frac{p(1-p)^{k-1}}{(2\pi)^{m/2}|\boldsymbol{\Sigma}|^{1/2}\sqrt{k}} \; e^{-\frac{1}{2k}(\boldsymbol{x}-k\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-k\boldsymbol{\mu})}$$
$$= \sum_{k=1}^{\infty} p(1-p)^{k-1}\phi_m(\boldsymbol{x}; k\boldsymbol{\mu}, k\boldsymbol{\Sigma}). \qquad (36)$$

If $\boldsymbol{\mu} = \boldsymbol{0}$ and $\boldsymbol{\Sigma} = \boldsymbol{I}$, we say that $\boldsymbol{X}$ has standard MGSN distribution with PDF

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \sum_{k=1}^{\infty} p(1-p)^{k-1}\phi_m(\boldsymbol{x}; \boldsymbol{0}, k\boldsymbol{I}). \qquad (37)$$

In the multivariate case also, when $p = 1$, $\boldsymbol{X} \sim \mathrm{N}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, i.e. it coincides with the multivariate normal distribution. Further, the generation from a MGSN distribution is quite simple, and it can be performed along the same line as the univariate GSN model.

In Figure 6, we provide the contour plots of bivariate geometric skew normal distribution (BGSN) for different values of μ_1, μ_2, σ_1, σ_2, p and ρ, where

$$\boldsymbol{\mu} = (\mu_1, \mu_2)^T \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{bmatrix}.$$

From the contour plots of the BGSN distribution, it is clear that the PDF of BGSN can take different shapes depending on the different parameter values. It can be unimodal or multimodal, and it can be skewed in any directions. It can be symmetric with heavier tails.

Now we discuss the marginals and conditional distributions. We use the following notations.

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}, \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

Here the vectors $\boldsymbol{X}_1$ and $\boldsymbol{\mu}_1$ are of the orders m_1 each, and the matrix $\boldsymbol{\Sigma}_{11}$ is of the order $m_1 \times m_1$. Rest of the quantities are defined so that they are compatible, and we define $m_2 = m - m_1$. We have the following results.

Theorem 5.1: If $\boldsymbol{X} \sim \mathrm{MGSN}(m, p, \boldsymbol{\mu}, \boldsymbol{\Sigma})$, then

(a) $\boldsymbol{X}_1 \sim \mathrm{MGSN}(m_1, p, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ and $\boldsymbol{X}_2 \sim \mathrm{MGSN}(m_2, p, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$.

(b) The conditional PDF of $\boldsymbol{X}_1$, given $\boldsymbol{X}_2 = \boldsymbol{x}_2$ becomes

$$f_{\boldsymbol{X}_1|\boldsymbol{X}_2=\boldsymbol{x}_2}(\boldsymbol{x}_1) = \frac{f_{\boldsymbol{X}}(\boldsymbol{x})}{f_{\boldsymbol{X}_2}(\boldsymbol{x}_2)}$$

$$= \frac{\sum_{k=1}^{\infty}(1-p)^{k-1}\phi_{m_1}(\boldsymbol{x}_1; k\boldsymbol{\mu}_1 + k\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), k(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}))\phi_{m_2}(\boldsymbol{x}_2, k\boldsymbol{\mu}_2, k\boldsymbol{\Sigma}_{22})}{\sum_{k=1}^{\infty}(1-p)^{k-1}\phi_{m_2}(\boldsymbol{x}_2, k\boldsymbol{\mu}_2, k\boldsymbol{\Sigma}_{22})}.$$

Proof: The proofs can be obtained easily, and it is avoided. ∎

If $\boldsymbol{X} \sim \text{MGSN}(m, p, \boldsymbol{\mu}, \boldsymbol{\Sigma})$, then the moment generating function of $\boldsymbol{X}$ becomes;

$$\begin{aligned} M_{\boldsymbol{X}}(\boldsymbol{t}) &= Ee^{\boldsymbol{t}^T\boldsymbol{X}} = E(Ee^{\boldsymbol{t}^T\boldsymbol{X}}|N) = E\left[e^{N\left(\boldsymbol{\mu}^T\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}\right)}\right] \\ &= \frac{pe^{\boldsymbol{\mu}^T\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}}{1-(1-p)e^{\boldsymbol{\mu}^T\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}}; \quad \boldsymbol{t} \in \mathbb{R}^m. \end{aligned}$$

Now we will show that similarly as the univariate case, MGSN distribution is also infinitely divisible. Similarly, as before, consider the following random vector $\boldsymbol{R}$, when $r = \dfrac{1}{n}$.

$$\boldsymbol{R} \stackrel{d}{=} \sum_{i=1}^{1+nT} \boldsymbol{Y}_i, \tag{38}$$

where $\boldsymbol{Y}_i$'s are i.i.d., $\boldsymbol{Y}_1 \sim \text{N}_m\left(\frac{1}{n}\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}\right)$ and $T \sim \text{NB}(r, p)$, as defined before. The moment generating function of $\boldsymbol{R}$ becomes

$$M_{\boldsymbol{R}}(\boldsymbol{t}) = E\left(e^{\boldsymbol{t}^T\boldsymbol{R}}\right) = E\left(e^{\sum_{i=1}^{1+nT}\boldsymbol{t}^T\boldsymbol{Y}_i}|T\right) = \left[\frac{pe^{\boldsymbol{\mu}^T\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}}{1-(1-p)e^{\boldsymbol{\mu}^T\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}}\right]^r = \left[M_{\boldsymbol{X}}(\boldsymbol{t})\right]^r.$$

It implies that the MGSN law is infinitely divisible. It can further be shown that MGSN law has geometric stability property, and it also enjoys the stochastic representation as in (17). Note that the EM algorithm also can be developed along the same line as the univariate case, and it is not pursues further.

5.2 Induced Lévy Process

In this section we will show that the MGSN law induces multivariate Lévy process. We have already observed that MGSN distribution is infinitely divisible. Consider the following random vector

$$(\boldsymbol{R}, V) \stackrel{d}{=} \left(\sum_{i=1}^{1+nT} \boldsymbol{Y}_i, \frac{1}{n} + T \right), \tag{39}$$

here $\boldsymbol{Y}_i$'s are i.i.d., $\boldsymbol{Y}_1 \sim \mathrm{N}_m\left(\frac{1}{n}\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}\right)$, and $T \sim \mathrm{NB}(r, p)$, with $r = \frac{1}{n}$. The moment generating function of $(\boldsymbol{R}, V)$ is given by

$$\begin{aligned} M_{\boldsymbol{R},V}(\boldsymbol{t}, s) &= Ee^{\boldsymbol{t}^T \boldsymbol{R} + sV} = E\left[e^{s(r+T)} E\left(e^{\sum_{i=1}^{1+nT} \boldsymbol{t}^T \boldsymbol{Y}_i} | T \right) \right] \\ &= \left\{ e^s \left[e^{\boldsymbol{\mu}^T \boldsymbol{t} + \frac{1}{2} \boldsymbol{t}^T \boldsymbol{\Sigma} \boldsymbol{t}} \right] \right\}^r \times E\left(e^s \left[e^{\boldsymbol{\mu}^T \boldsymbol{t} + \frac{1}{2} \boldsymbol{t}^T \boldsymbol{\Sigma} \boldsymbol{t}} \right] \right)^T \\ &= \left[\frac{pe^s \left[e^{\boldsymbol{\mu}^T \boldsymbol{t} + \frac{1}{2} \boldsymbol{t}^T \boldsymbol{\Sigma} \boldsymbol{t}} \right]}{1 - (1-p)e^s \left[e^{\boldsymbol{\mu}^T \boldsymbol{t} + \frac{1}{2} \boldsymbol{t}^T \boldsymbol{\Sigma} \boldsymbol{t}} \right]} \right]^r . \end{aligned} \tag{40}$$

Clearly, (40) is a moment generating function for any $r > 0$, and this is associated with the following multivariate random vector;

$$(\boldsymbol{R}(r), V(r)) \stackrel{d}{=} \left(\sum_{i=1}^{T} \boldsymbol{Y}_i + \boldsymbol{Z}, r + T \right), \tag{41}$$

where $\boldsymbol{Y}_i$'s are i.i.d., $\boldsymbol{Y}_1 \sim \mathrm{N}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{Z} \sim \mathrm{N}_m(r\boldsymbol{\mu}, r\boldsymbol{\Sigma})$, $T \sim \mathrm{NB}(r, p)$, and all the associated random variables/ vectors involved here are mutually independent. Hence, it follows that MGSN law induces a Lévy process $\{(\boldsymbol{X}(r), NB(r)), r \geq 0\}$, which has the following stochastic representation

$$\{(\boldsymbol{X}(r), N(r)); r \geq 0\} \stackrel{d}{=} \left\{ \left(\sum_{i=1}^{NB(r)} \boldsymbol{Y}_i + \boldsymbol{Z}(r), r + NB(r) \right); r \geq 0 \right\}, \tag{42}$$

where $\boldsymbol{Y}_i$'s are same as defined above in (41), $\{\boldsymbol{Z}(r) : r \geq 0\}$ is a normal Lévy process, and $\{NB(r); r \geq 0\}$ is a negative binomial Lévy process with moment generating functions given by

$$E\left(e^{\boldsymbol{t}^T \boldsymbol{Z}(r)} \right) = e^{r\boldsymbol{t}^T \boldsymbol{\mu} + r\frac{1}{2} \boldsymbol{t}^T \boldsymbol{\Sigma} \boldsymbol{t}}; \quad \boldsymbol{t} \in \mathbb{R}^m \tag{43}$$

and

$$E\left(e^{sNB(r)} \right) = \left(\frac{p}{1 - (1-p)e^s} \right)^r; \quad s \in \mathbb{R}, \tag{44}$$

respectively.

Further, observe that

$$\sum_{i=1}^{NB(r)} \boldsymbol{Y}_i + \boldsymbol{Z}(r) \mid NB(r) = k \sim \mathrm{N}_m\left((r+k)\boldsymbol{\mu}, (r+k)\boldsymbol{\Sigma}\right),$$

therefore, we can obtain another Lévy process with the following stochastic representation

$$\{(\boldsymbol{Y}(r), NB(r))\,; r \geq 0\} \stackrel{d}{=} \{(\boldsymbol{Z}(r+NB(r)), NB(r))\,; r \geq 0\}. \tag{45}$$

The above result also follows from the stochastic self-similarity property: a normal Lévy process subordinated to a negative binomial process with drift is again a normal process. The moment generating function corresponding to (45) process becomes

$$E\left(e^{\boldsymbol{t}^T\boldsymbol{Y}(r)+sNB(r)}\right) = \left[\frac{pe^{\boldsymbol{t}^T\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}}{1-(1-p)e^{s+\boldsymbol{t}^T\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}}\right]^r. \tag{46}$$

Therefore, the moment generating function of the marginal process $\boldsymbol{X}(r)$ becomes

$$E\left(e^{\boldsymbol{t}^T\boldsymbol{Y}(r)}\right) = \left[\frac{pe^{\boldsymbol{t}^T\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}}{1-(1-p)e^{\boldsymbol{t}^T\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}}\right]^r. \tag{47}$$

The moment generating function (47) corresponds to a random variable whose PDF is an infinite mixture of multivariate normal random variables with negative binomial weights. Estimation procedures of the unknown parameters and other inferential issues will be of interest, and it is not pursued here any more.

6 Conclusions

In this paper we have introduced a new three-parameter distribution of which normal distribution is a particular member. The proposed distribution, can be unimodal, multimodal, positively or negatively skewed and symmetric with heavy tails. The model is more flexible than the very popular Azzalini's skew normal distribution, although they have the same number of parameters. We derive different properties of this distribution and develop different inferential issues. Further the model has been generalized to the multivariate case, and it is observed that the multivariate generalized model can be more flexible than Azzalini's multivariate skew normal model. It will be of interest to develop different inferential issues of the multivariate model. More work is needed along that direction.

Acknowledgements:

The author would like to thank the referees for many constructive comments which had helped to improve the manuscript significantly.

References

[1] Arnold, B.C. and Beaver, R.J. (2000), "Hidden truncation models", *Sankhya*, vol. 62, 23 - 35.

[2] Arnold, B.C. and Beaver, R.J., Groeneveld, R.A. and Meeker, W.Q. (1993), "The non-truncated marginal of a truncated bivariate normal distribution", *Psychometrika*, vol. 58, 471 - 488.

[3] Azzalini, A.A. (1985), "A class of distributions which include the normal", *Scandinavian Journal of Statistics*, vol. 12, 171 - 178.

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

[5] Barreto-Souza, W. (2012), "Bivariate gamma-geometric law and its induced Levy process", *Journal of Multivariate Analysis*, vol. 109, 130 - 145.

[6] Bjerkedal, T. (1960), "Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli", *American Journal of Hygiene*, vol. 72, 130-148.

[7] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995), *Continuous Univariate Distribution*, Volume 1, John Wiley and Sons, New York, USA.

[8] Johnson, R.A. and Wichern, D.W. (1992), *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey, USA.

[9] Gupta, R.C. and Gupta, R.D. (2004), "Generalized skew normal model", *TEST*, vol. 13, 1- 24.

[10] Kuzobowski, T.J. and Panorska, A.K. (2005), "A mixed bivariate distribution with exponential and geometric marginals", *Journal of Statistical Planning and Inference*, vol. 134, 501 - 520.

[11] Kuzobowski, T.J., Panorska, A.K. and Podgorski, K. (2008), "A bivariate Levy process with negative binomial and gamma marginals", *Journal of Multivariate Analysis*, vol. 99, 1418 - 1437.

[12] Louis, T.A. (1982), "Finding the observed information matrix when using the EM algorithm", *Journal of the Royal Statistical Society*, Ser. B, vol. 44, 226 - 233.

[13] Self, S.G. and Liang, K-L. (1987), "Asymptotic properties of the maximum likelihood estimators and likelihood ratio test under non-standard conditions", *Journal of the American Statistical Association*, vol. 82, 605 - 610.

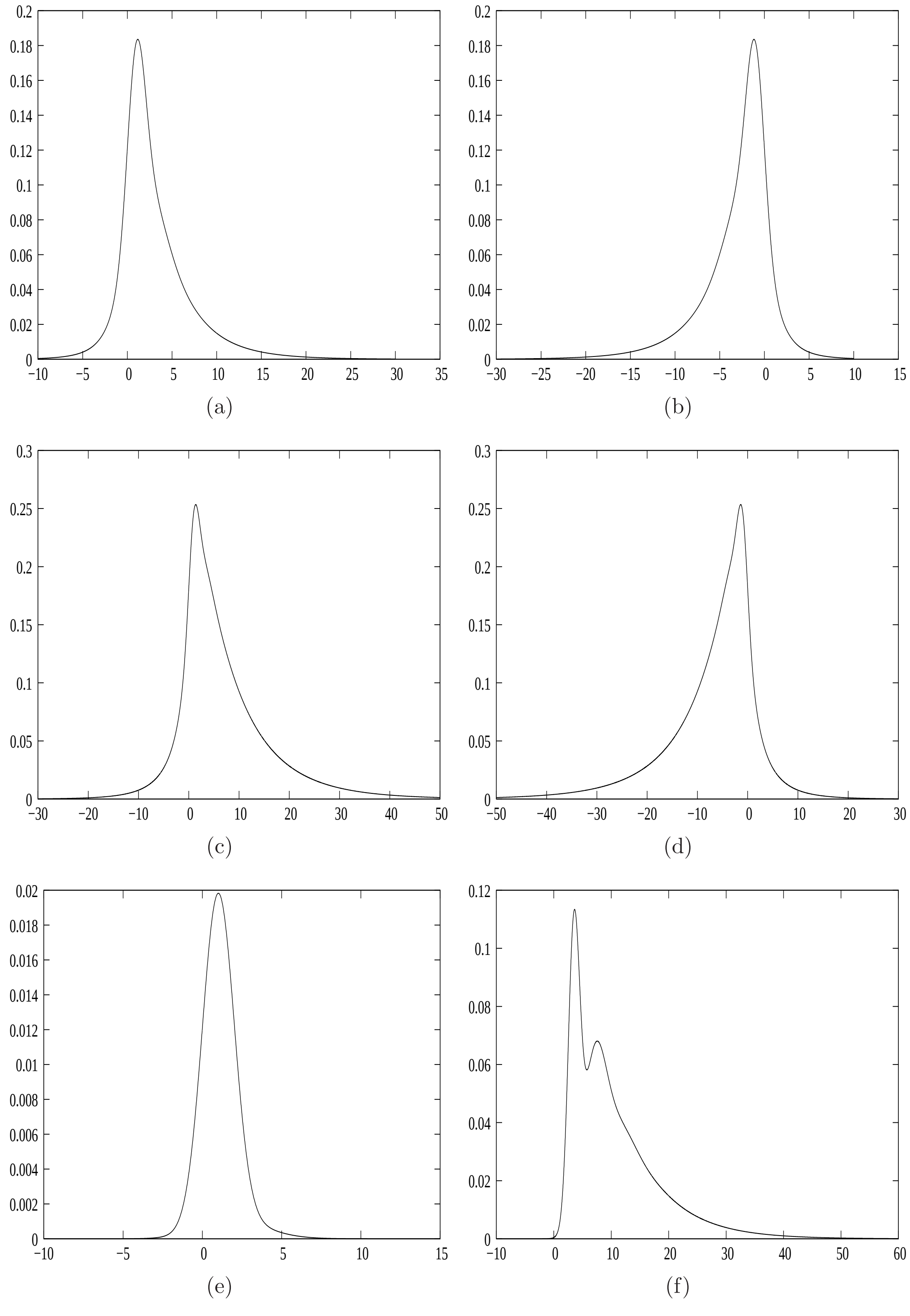

Figure 1: The PDF of the GSN law for different parameter values of μ and p, when $\sigma = 1$, (μ, p) : (a) (1,0.5) (b) (-1,0.5) (c) (1.0, 0.25) (d) (-1.0,0.25) (e)(1.0,0.95) (f) (3.5,0.5)
.

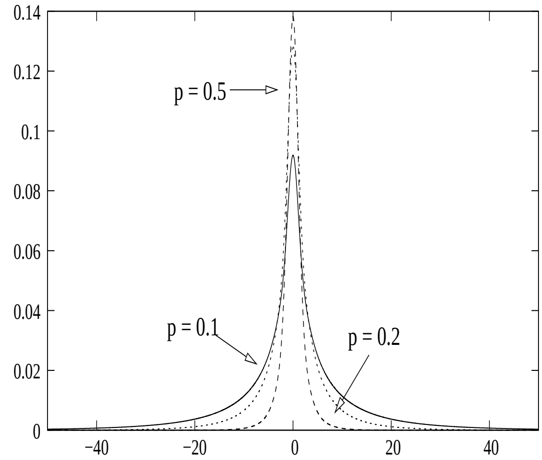

Figure 2: The PDF of the standard GSN law for different values of p, when $\sigma = 1$

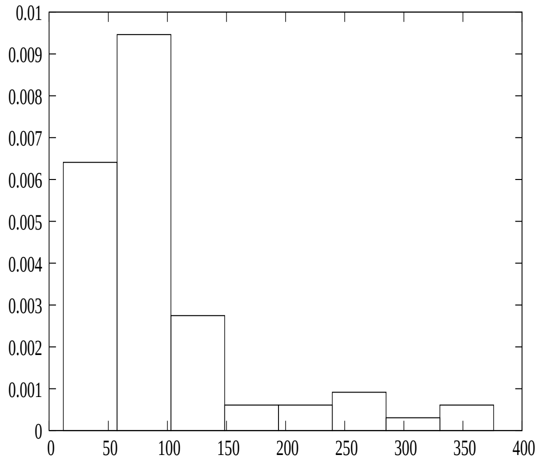

Figure 3: Histogram plot of the guineapig data.

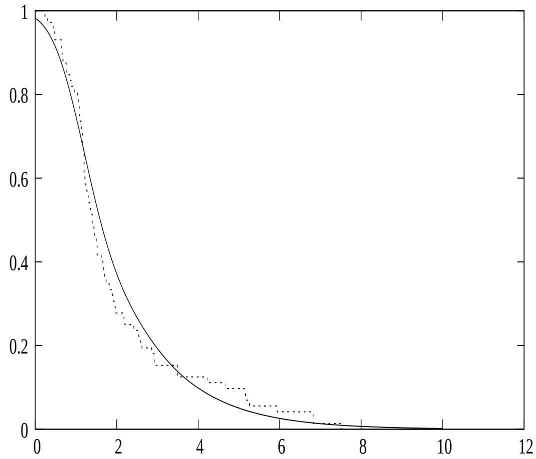

Figure 4: Fitted and the empirical survival function of the guineapig data.

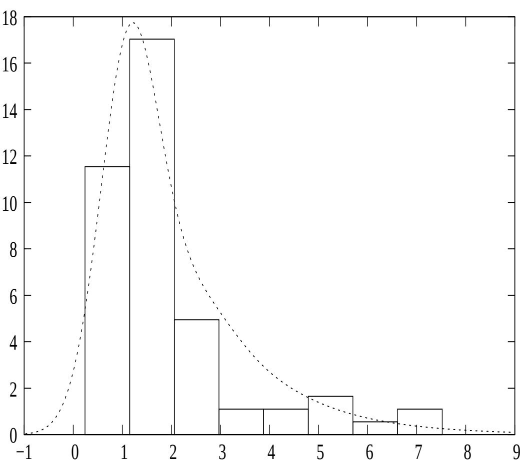

Figure 5: Histogram and the fitted probability density function for the guineapig data.

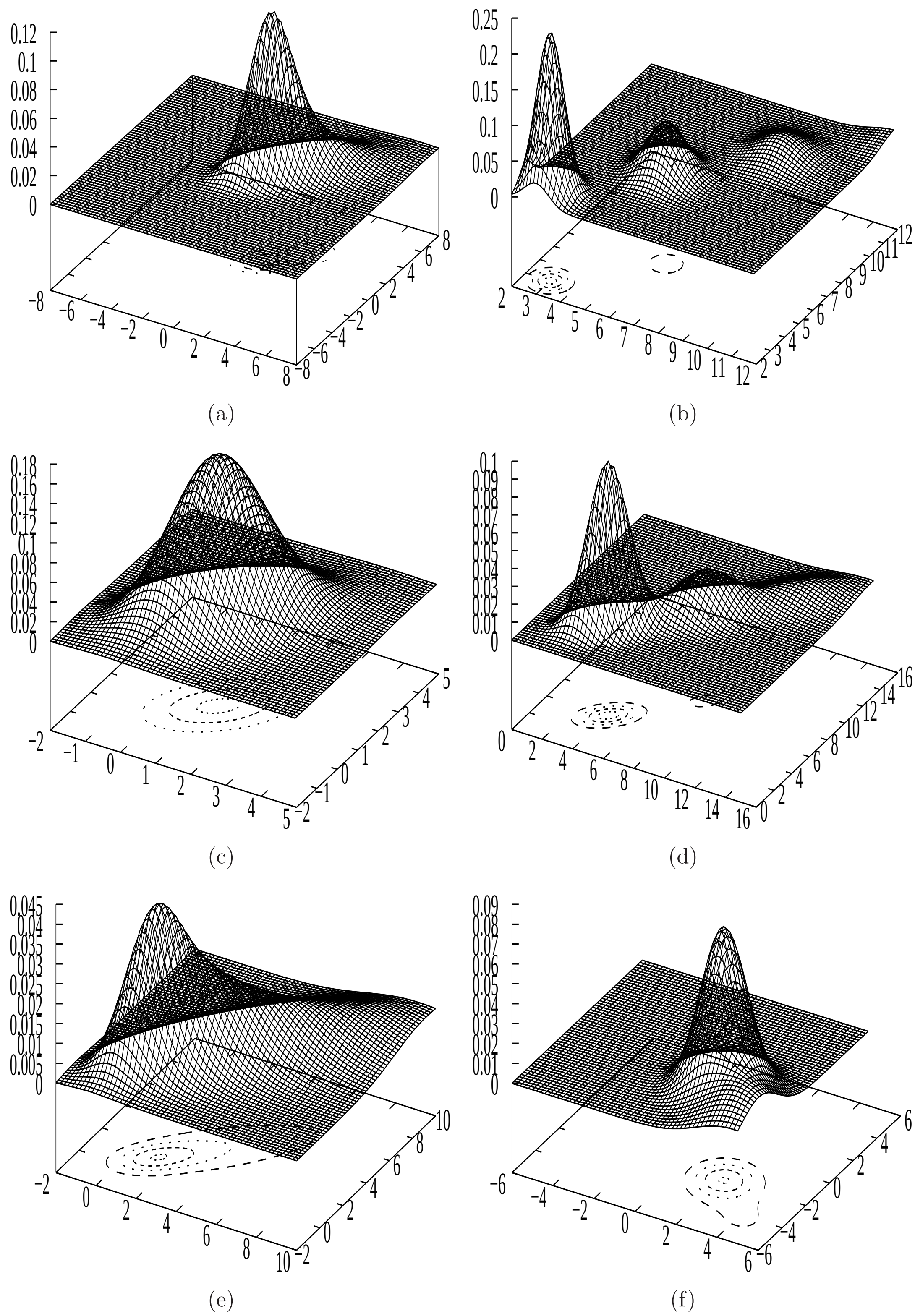

Figure 6: The surface plots of BGSN distribution for different parameter values of $(\mu_1,\ \mu_2,\ \sigma_1,\ \sigma_2,\ p,\ \rho)$: (a) (1.0,1.0,1.0,1.0,0.5,0.5) (b) (3.0,3.0,0.5,0.5,0.35,0.0) (c) (1.0,1.0,1.0, 1.0,0.85,0.5) (d) (4.0,4.0,1.0,1.0,0.5,0.5) (e)(1.0,1.0,1.0,1.0,0.15,0.15) (f) (2.0,-2.0,1.0,1.0,0.5,0.0)