

A New Bivariate Distribution Obtained by Compounding the Bivariate Normal and Geometric Distributions

Eisa Mahmoudi*

*Department of Statistics, Yazd University,
P.O. Box 89175- 741, Yazd, Iran
emahmoudi@yazd.ac.ir*

Hamed Mahmoodian

*Department of Statistics, Yazd University,
P.O. Box 89175- 741, Yazd, Iran
hamed_-mahmoodian@yahoo.com*

Received 31 August 2016

Accepted 10 October 2016

Recently, Mahmoudi and Mahmoodian [7] introduced a new class of distributions which contains univariate normal–geometric distribution as a special case. This class of distributions are very flexible and can be used quite effectively to analyze skewed data. In this paper we propose a new bivariate distribution with the normal–geometric distribution marginals. Different properties of this new bivariate distribution have been studied. This distribution has five unknown parameters. The EM algorithm is used to determine the maximum likelihood estimates of the parameters. We analyze one series of real data set for illustrative purposes.

Keywords: Normal distribution; Geometric distribution; EM algorithm; Maximum likelihood estimation.

2000 Mathematics Subject Classification: 60E05, 62F10, 62P99

1. Introduction

In recent years, different techniques for extending the family of normal distributions have been proposed. The most usual procedure derived from Azzalini [1] consists in deriving a skew-normal distribution through the transformation $\phi(z; \lambda) = 2\phi(z)\Phi(\lambda z)$, $z, \lambda \in \mathbb{R}$. Recently, Mahmoudi and Mahmoodian [7] considered univariate normal–power series class of distributions which contains normal–geometric distribution as a special case. They, by compounding normal and power series class of distributions, introduced an alternative skewed model. The normal–geometric distribution is defined as follows: The univariate random variable X is said to have a normal–geometric (NG)

*Corresponding author.

distribution, denoted by $X \sim NG(\mu, \sigma, \theta)$, where $\mu \in \mathbb{R}$, $\sigma > 0$ and $\theta < 1$, if its pdf is given by

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

where $\phi(\cdot; \mu, \sigma)$ and $\Phi(\cdot; \mu, \sigma)$ are pdf and cdf of normal distribution with mean μ and variance σ^2 . Different properties of the NG distribution have been discussed by Mahmoudi and Mahmoodian [7].

Recently, Jafari and Roozgar [4] introduced a new bivariate distribution by compounding a bivariate linear failure rate with a power series class of distributions. The bivariate Weibull–power series introduced and studied by Roozegar and Nadarajah [9].

In this paper, we introduce bivariate normal–geometric distribution with the normal–geometric distribution marginals. Different properties of this new distribution have been investigated. It is observed that the generation of random samples from the proposed bivariate model is very simple, hence simulation experiments can be performed quite conveniently. The bivariate normal–geometric distribution has five parameters. The maximum likelihood estimators (MLEs) of the unknown parameters cannot be obtained in explicit forms. One needs to solve five 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. The analysis of one series of real data has been performed for illustrative purposes, and it is observed that the proposed model provides a good fit to the real data set.

To begin with, we shall use the following notation throughout this paper: $\phi(\cdot)$ and $\Phi(\cdot)$ for the standard normal probability density and cumulative distribution function, respectively, $\phi_n(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ 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 $\boldsymbol{\mu} = \mathbf{0}$. Furthermore, for $r, k \in \mathbb{N}$, let $\mathbf{1}_r$, \mathbf{I}_r and $\mathbf{0}_{r \times k}$ denote the vector of ones, the identity matrix of dimension r , and $r \times k$ zero matrix, respectively, and let \mathbf{X}_{-i} be the vector obtained from \mathbf{X} by deleting its i th component.

The rest of the paper is organized as follows. In Section 2, we introduce the bivariate normal–geometric distribution. In Section 3, different properties of the proposed bivariate normal–geometric are discussed. EM algorithm is presented in Section 4. Simulation study is given in Section 5. Application to one real data set is given in Section 6. Finally, Section 7 concludes the paper.

2. Bivariate Normal–Geometric Distribution

Bivariate normal–geometric (BNG) distribution can be constructed as follows. Let $\{X_1, \dots, X_N\}$ and $\{Y_1, \dots, Y_N\}$ be sequence of mutually independent and identically distributed (i.i.d.) $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively. Also N has a geometric distribution with the probability mass function

$$P(N = n) = (1 - \theta)\theta^{n-1}, \quad n = 1, \dots, \quad 0 < \theta < 1.$$

Moreover, N is independent of X_i 's and Y_i 's. Let

$$U_1 = \max(X_1, \dots, X_N) \quad \text{and} \quad U_2 = \max(Y_1, \dots, Y_N).$$

The joint cdf of (U_1, U_2) is

$$\begin{aligned} F_{U_1, U_2}(u_1, u_2) &= P(U_1 \leq u_1, U_2 \leq u_2) \\ &= \sum_{n=1}^{\infty} (1 - \theta)\theta^{n-1} [\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2)]^n \\ &= \frac{(1 - \theta)\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2)}{1 - \theta\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2)}. \end{aligned} \tag{2.1}$$

The bivariate random vector (U_1, U_2) is said to have a BNG, denoted by $BNG(\mu_1, \mu_2, \sigma_1, \sigma_2, \theta)$, if (U_1, U_2) has the joint cdf in (2.1). The joint pdf and survival functions of (U_1, U_2) are

$$f_{U_1, U_2}(u_1, u_2) = \frac{(1 - \theta)\phi(u_1; \mu_1, \sigma_1)\phi(u_2; \mu_2, \sigma_2)}{(1 - \theta\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2))^2} \times \left[1 + \frac{2\theta\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2)}{(1 - \theta\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2))} \right],$$

and

$$S_{U_1, U_2}(u_1, u_2) = 1 - \frac{(1 - \theta)\Phi(u_1; \mu_1, \sigma_1)}{1 - \theta\Phi(u_1; \mu_1, \sigma_1)} - \frac{(1 - \theta)\Phi(u_2; \mu_2, \sigma_2)}{1 - \theta\Phi(u_2; \mu_2, \sigma_2)} + \frac{(1 - \theta)\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2)}{1 - \theta\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2)}.$$

respectively. Figure 1 shows the BNG density function for selected values θ where $\mu_1 = \mu_2 = 0$ and $\sigma_1 = \sigma_2 = 1$. From Figure 1, it is quite apparent that it can take different shapes, and it is unimodal for different values of the parameters.

The following theorem provides the marginal distribution and conditional cdf of the BNG distribution.

Theorem 2.1. *If $(U_1, U_2) \sim BNG(\mu_1, \mu_2, \sigma_1, \sigma_2, \theta)$, then*

(i) $U_i \sim NG(\mu_i, \sigma_i, \theta)$, $i = 1, 2$.

(ii) *The conditional pdf of U_1 given $U_2 = u_2$ is*

$$\begin{aligned} f_{U_1|U_2}(u_1|u_2) &= \phi(u_1; \mu_1, \sigma_1) \left(\frac{\theta\Phi(u_2; \mu_2, \sigma_2) - 1}{\theta\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2) - 1} \right)^2 \\ &\quad - \frac{2\theta\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2)\phi(u_1; \mu_1, \sigma_1)(\theta\Phi(u_2; \mu_2, \sigma_2) - 1)^2}{(\theta\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2) - 1)^3}. \end{aligned}$$

(iii) *The conditional cdf of U_1 given $U_2 = u_2$ is*

$$P(U_1 \leq u_1 | U_2 = u_2) = \Phi(u_1; \mu_1, \sigma_1) \times \left(\frac{1 - \theta\Phi(u_2; \mu_2, \sigma_2)}{1 - \theta\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2)} \right)^2$$

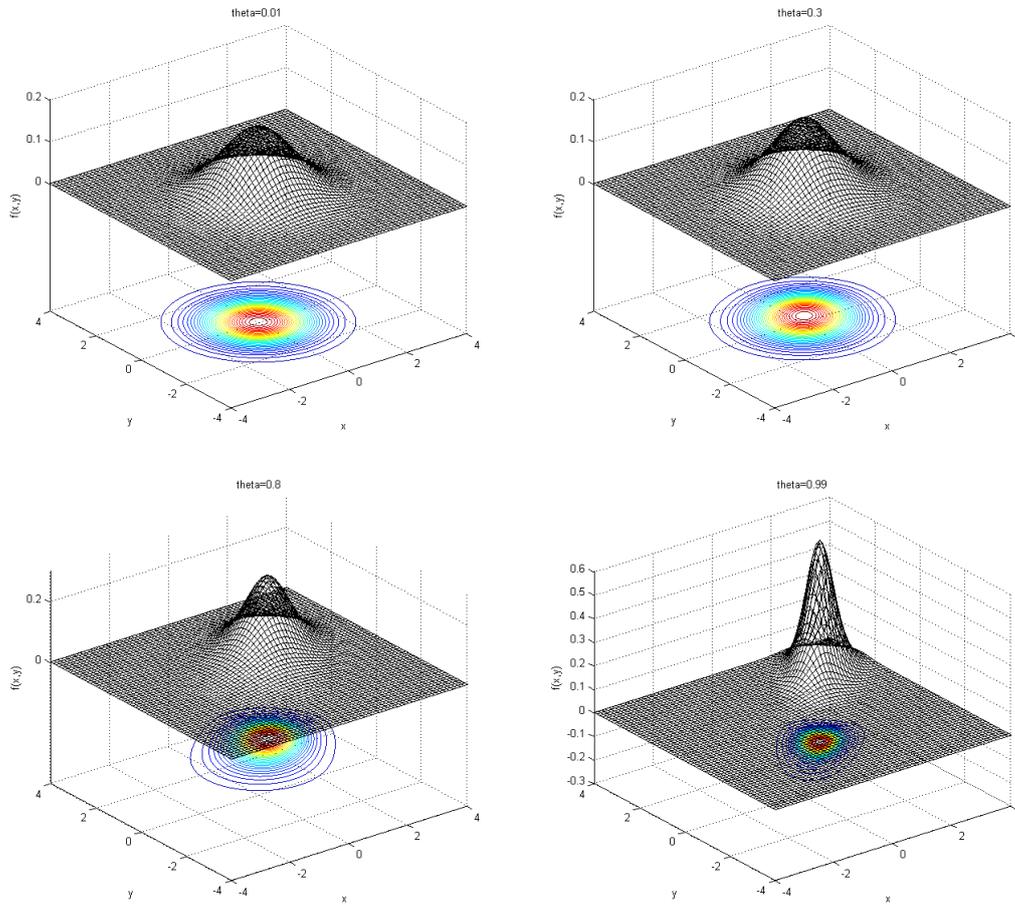


Fig. 1. The pdf of BNG distribution when $\mu_1 = \mu_1 = 0$ and $\sigma_1 = \sigma_1 = 1$ and $\theta = 0.01$ (left top), $\theta = 0.3$ (right top), $\theta = 0.8$ (left bottom), $\theta = 0.99$ (right bottom)

Proof. The proof of (i) and (ii) can be obtained in a routine matter. For part (iii), we can write

$$\begin{aligned}
 P(U_1 \leq u_1 | U_2 = u_2) &= \frac{P(U_1 \leq u_1, U_2 = u_2)}{P(U_2 = u_2)} \\
 &= \sum_{n=1}^{\infty} P(U_1 \leq u_1 | N = n, U_2 = u_2) \times P(N = n | U_2 = u_2) \\
 &= \sum_{n=1}^{\infty} (\Phi(u_1; \mu_1, \sigma_1))^n n \theta^{n-1} (\Phi(u_2; \mu_2, \sigma_2))^{n-1} \\
 &= \Phi(u_1; \mu_1, \sigma_1) (1 - \theta \Phi(u_2; \mu_2, \sigma_2))^2 \sum_{n=1}^{\infty} n \theta^{n-1} (\Phi(u_1; \mu_1, \sigma_1) \Phi(u_2; \mu_2, \sigma_2))^{n-1} \\
 &= \Phi(u_1; \mu_1, \sigma_1) \times \left(\frac{1 - \theta \Phi(u_2; \mu_2, \sigma_2)}{1 - \theta \Phi(u_1; \mu_1, \sigma_1) \Phi(u_2; \mu_2, \sigma_2)} \right)^2.
 \end{aligned}$$

□

3. Properties

To better motivate the results developed in this section, we first provide a brief definition of the multivariate unified skew-normal (SUN) distributions. Let \mathbf{V}_1 and \mathbf{V}_2 be two random vectors of dimensions m and n , respectively, and

$$\begin{pmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{pmatrix} \sim N_{m+n} \left(\begin{pmatrix} \boldsymbol{\eta} \\ \boldsymbol{\xi} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Gamma} & \boldsymbol{\Lambda}^T \\ \boldsymbol{\Lambda} & \boldsymbol{\Omega} \end{pmatrix} \right).$$

The n -dimensional random vector \mathbf{Z} is said to have the SUN distribution with parameter $\boldsymbol{\alpha} = (\boldsymbol{\xi}, \boldsymbol{\eta}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\Lambda})$, where $\boldsymbol{\xi} \in \mathbb{R}^n$ and $\boldsymbol{\eta} \in \mathbb{R}^m$ are location vectors, $\boldsymbol{\Omega} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{\Gamma} \in \mathbb{R}^{m \times m}$ are dispersion matrices, and $\boldsymbol{\Lambda} \in \mathbb{R}^{n \times m}$ is a skewness/shape matrix, denoted by $\mathbf{Z} \sim SUN_{n,m}(\boldsymbol{\xi}, \boldsymbol{\eta}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\Lambda})$ or simply by $\mathbf{Z} \sim SUN_{n,m}(\boldsymbol{\alpha})$, if

$$\mathbf{Z} \stackrel{d}{=} \mathbf{V}_2 \mid (\mathbf{V}_1 > 0).$$

The density function of \mathbf{Z} is given by [see [2]]

$$f_{SUN_{n,m}}(\mathbf{z}; \boldsymbol{\alpha}) = \frac{\phi_n(\mathbf{z}; \boldsymbol{\xi}, \boldsymbol{\Omega}) \Phi_m(\boldsymbol{\eta} + \boldsymbol{\Lambda}^T \boldsymbol{\Omega}^{-1}(\mathbf{z} - \boldsymbol{\xi}); \boldsymbol{\Gamma} - \boldsymbol{\Lambda}^T \boldsymbol{\Omega}^{-1} \boldsymbol{\Lambda})}{\Phi_m(\boldsymbol{\eta}; \boldsymbol{\Gamma})}.$$

Furthermore, when $\mathbf{Z} \sim SUN_{n,m}(\boldsymbol{\alpha})$, the mgf of \mathbf{Z} is available in an explicit form and is given by

$$M_{SUN_{n,m}}(\mathbf{s}; \boldsymbol{\alpha}) = \frac{\exp\left(\boldsymbol{\xi}^T \mathbf{s} + \frac{1}{2} \mathbf{s}^T \boldsymbol{\Omega} \mathbf{s}\right) \Phi_m(\boldsymbol{\eta} + \boldsymbol{\Lambda}^T \mathbf{s}; \boldsymbol{\Gamma})}{\Phi_m(\boldsymbol{\eta}; \boldsymbol{\Gamma})}. \quad (3.1)$$

Now, Let \mathbf{X} and \mathbf{Y} be two random vectors of dimensions n , and $(\mathbf{X}^T, \mathbf{Y}^T)$ having a multivariate normal distribution

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim N_{2n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\mu} = \begin{pmatrix} \mathbf{1}_n \mu_1 \\ \mathbf{1}_n \mu_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 \mathbf{I}_n & \mathbf{0}_{n \times n} \\ \mathbf{0}_{n \times n} & \sigma_2^2 \mathbf{I}_n \end{pmatrix}.$$

If $X_{(n)} = \max(X_1, \dots, X_n)$ and $Y_{(n)} = \max(Y_1, \dots, Y_n)$ be the largest order statistics obtained from \mathbf{X} and \mathbf{Y} , respectively, then the joint pdf of $(X_{(n)}, Y_{(n)})$ is given by (see [8])

$$f_{X_{(n)}, Y_{(n)}}(u_1, u_2) = f_{SUN_{2,2n-2}}(u_1, u_2; \alpha), \tag{3.2}$$

where $\alpha = (\xi, \mathbf{0}, \mathbf{\Omega}, \mathbf{\Gamma}, \mathbf{\Lambda})$, with

$$\begin{aligned} \xi &= \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \mathbf{\Omega} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}, \\ \mathbf{\Gamma} &= \begin{pmatrix} \sigma_1^2 (\mathbf{I}_{n-1} + \mathbf{1}_{n-1} \mathbf{1}_{n-1}^T) & \mathbf{0}_{n-1 \times n-1} \\ \sigma_2^2 (\mathbf{I}_{n-1} + \mathbf{1}_{n-1} \mathbf{1}_{n-1}^T) & \end{pmatrix}, \\ \mathbf{\Lambda} &= \begin{pmatrix} \mathbf{1}_{n-1} \sigma_1^2 & \mathbf{0}_{n-1 \times n-1} \\ \mathbf{1}_{n-1} \sigma_2^2 & \end{pmatrix}. \end{aligned}$$

In the following proposition, we first present the mixture representation for $f_{U_1, U_2}(u_1, u_2)$.

Proposition 3.1. *The densities of BNG class can be written as follows*

$$f_{U_1, U_2}(u_1, u_2) = \sum_{n=1}^{\infty} P(N = n) f_{X_{(n)}, Y_{(n)}}(u_1, u_2), \tag{3.3}$$

where $f_{X_{(n)}, Y_{(n)}}(u_1, u_2)$ is the joint density function of $(X_{(n)}, Y_{(n)})$ in (3.2).

Proposition 3.2. *If the random vector $(U_1, U_2) \sim BNG(\mu_1, \mu_2, \sigma_1, \sigma_2, \theta)$, then*

(i) *the mgf of (U_1, U_2) is available and for $\mathbf{s} \in \mathbb{R}^2$, is given by*

$$\begin{aligned} M_{U_1, U_2}(\mathbf{s}) &= \sum_{n=1}^{\infty} (1 - \theta) \theta^{n-1} M_{SUN_{2,2n-2}}(\mathbf{s}; \alpha) \\ &= \sum_{n=1}^{\infty} (1 - \theta) \theta^{n-1} \times \frac{\exp\left(\xi^T \mathbf{s} + \frac{1}{2} \mathbf{s}^T \mathbf{\Omega} \mathbf{s}\right) \Phi_{2n-2}(\mathbf{\Lambda}^T \mathbf{s}; \mathbf{\Gamma})}{\Phi_{2n-2}(\mathbf{0}; \mathbf{\Gamma})}. \end{aligned}$$

(ii) *The product moment $E(U_1 U_2)$ is given by*

$$\begin{aligned} E(U_1 U_2) &= \\ &= \sum_{n=1}^{\infty} (1 - \theta) \theta^{n-1} \left[\mu_1 + \frac{n(n-1)\sigma_1}{2\sqrt{\pi}} \Phi_{n-2}\left(\mathbf{0}; \mathbf{I}_{n-2} + \frac{1}{2} \mathbf{1}_{n-2} \mathbf{1}_{n-2}^T\right) \right] \\ &\times \left[\mu_2 + \frac{n(n-1)\sigma_2}{2\sqrt{\pi}} \Phi_{n-2}\left(\mathbf{0}; \mathbf{I}_{n-2} + \frac{1}{2} \mathbf{1}_{n-2} \mathbf{1}_{n-2}^T\right) \right]. \end{aligned}$$

Proof. we use (3.1) and (3.3) to proof part (i). For part (ii), we can write

$$\begin{aligned} E(U_1 U_2) &= E(E(X_{(N)} Y_{(N)} | N = n)) = \sum_{n=1}^{\infty} (1 - \theta) \theta^{n-1} E(X_{(n)} Y_{(n)}) \\ &= \sum_{n=1}^{\infty} (1 - \theta) \theta^{n-1} E(X_{(n)}) E(Y_{(n)}) \end{aligned}$$

and based on [5], we have

$$E(X_{(n)}) = \mu_1 + \frac{n(n-1)\sigma_1}{2\sqrt{\pi}} \Phi_{n-2}\left(\mathbf{0}; \mathbf{I}_{n-2} + \frac{1}{2} \mathbf{1}_{n-2} \mathbf{1}_{n-2}^T\right),$$

and

$$E(Y_{(n)}) = \mu_2 + \frac{n(n-1)\sigma_2}{2\sqrt{\pi}} \Phi_{n-2} \left(\mathbf{0}; \mathbf{I}_{n-2} + \frac{1}{2} \mathbf{1}_{n-2} \mathbf{1}_{n-2}^T \right),$$

which completes the proof. \square

The stress-strength parameter, $R = P(U_1 < U_2)$, is useful for data analysis purposes. The following result gives the stress-strength parameter of BNG model.

Proposition 3.3. *If $(U_1, U_2) \sim BNG(\mu_1, \mu_2, \sigma_1, \sigma_2, \theta)$, then*

$$P(U_1 < U_2) = \sum_{n=1}^{\infty} (1-\theta)\theta^{n-1} F_{SUN_{1,2n-2}}(\mathbf{0}; \boldsymbol{\alpha}),$$

where $F_{SUN_{1,2n-2}}(\cdot; \boldsymbol{\alpha})$ is the cdf of the univariate $SUN_{1,2n-2}(\boldsymbol{\alpha})$ distribution, and $\boldsymbol{\alpha} = (\mu_1 - \mu_2, \mathbf{0}, \sigma_1^2 + \sigma_2^2, \boldsymbol{\Gamma}, \boldsymbol{\Lambda})$, with

$$\boldsymbol{\Gamma} = \begin{pmatrix} \sigma_1^2 (\mathbf{I}_{n-1} + \mathbf{1}_{n-1} \mathbf{1}_{n-1}^T) & \mathbf{0}_{n-1 \times n-1} \\ \mathbf{0}_{n-1 \times n-1} & \sigma_2^2 (\mathbf{I}_{n-1} + \mathbf{1}_{n-1} \mathbf{1}_{n-1}^T) \end{pmatrix},$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} \mathbf{1}_{n-1} \sigma_1^2 & \mathbf{0}_{n-1 \times n-1} \\ \mathbf{0}_{n-1 \times n-1} & \mathbf{1}_{n-1} \sigma_2^2 \end{pmatrix}.$$

Proof. We have

$$P(U_1 < U_2) = P(X_{(N)} < Y_{(N)}) = \sum_{n=1}^{\infty} (1-\theta)\theta^{n-1} P(X_{(n)} < Y_{(n)})$$

Now, we compute $P(X_{(n)} < Y_{(n)})$.

$$P(X_{(n)} < Y_{(n)}) = \sum_{i=1}^n \sum_{j=1}^n P(\mathbf{1}_{n-1} X_i - \mathbf{X}_{-i} > \mathbf{0}, \mathbf{1}_{n-1} Y_j - \mathbf{Y}_{-j} > \mathbf{0}) \times P(X_i - Y_j \leq 0 \mid \mathbf{1}_{n-1} X_i - \mathbf{X}_{-i} > \mathbf{0}, \mathbf{1}_{n-1} Y_j - \mathbf{Y}_{-j} > \mathbf{0}).$$

Since, for $i = 1, \dots, n$ and $j = 1, \dots, n$

$$\begin{pmatrix} \mathbf{1}_{n-1} X_i - \mathbf{X}_{-i} \\ \mathbf{1}_{n-1} Y_j - \mathbf{Y}_{-j} \\ X_i - Y_j \end{pmatrix} \sim N_{2n} \left(\begin{pmatrix} \mathbf{0} \\ \mu_1 - \mu_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Gamma} & \boldsymbol{\Lambda} \\ \boldsymbol{\Lambda} & \sigma_1^2 + \sigma_2^2 \end{pmatrix} \right),$$

by using the definition of the univariate SUN distribution, we have

$$P(X_i - Y_j \leq 0 \mid \mathbf{1}_{n-1} X_i - \mathbf{X}_{-i} > \mathbf{0}, \mathbf{1}_{n-1} Y_j - \mathbf{Y}_{-j} > \mathbf{0}) = F_{SUN_{1,2n-2}}(\mathbf{0}; \boldsymbol{\alpha}),$$

and

$$\sum_{i=1}^n \sum_{j=1}^n P(\mathbf{1}_{n-1} X_i - \mathbf{X}_{-i} > \mathbf{0}, \mathbf{1}_{n-1} Y_j - \mathbf{Y}_{-j} > \mathbf{0}) = 1.$$

which completes the proof. \square

Proposition 3.4. If $(U_1, U_2) \sim \text{BNG}(\mu_1, \mu_2, \sigma_1, \sigma_2, \theta)$, then

(i) The cdf and pdf of $\max(U_1, U_2)$ is

$$F_{\max(U_1, U_2)}(u) = P(U_1 \leq u, U_2 \leq u) = \frac{(1 - \theta)\Phi(u; \mu_1, \sigma_1)\Phi(u; \mu_2, \sigma_2)}{1 - \theta\Phi(u; \mu_1, \sigma_1)\Phi(u; \mu_2, \sigma_2)},$$

and

$$f_{\max(U_1, U_2)}(u) = \frac{(1 - \theta)[\theta\phi(u; \mu_1, \sigma_1)\Phi(u; \mu_2, \sigma_2) + \theta\phi(u; \mu_2, \sigma_2)\Phi(u; \mu_1, \sigma_1)]}{\theta(1 - \theta\Phi(u; \mu_1, \sigma_1)\Phi(u; \mu_2, \sigma_2))^2}.$$

(ii) The cdf and pdf of $\min(U_1, U_2)$ is

$$\begin{aligned} F_{\min(U_1, U_2)}(u) &= P(\min(U_1, U_2) \leq u) \\ &= \frac{(1 - \theta)\Phi(u; \mu_1, \sigma_1)}{1 - \theta\Phi(u; \mu_1, \sigma_1)} + \frac{(1 - \theta)\Phi(u; \mu_2, \sigma_2)}{1 - \theta\Phi(u; \mu_2, \sigma_2)} \\ &\quad - \frac{(1 - \theta)\Phi(u; \mu_1, \sigma_1)\Phi(u; \mu_2, \sigma_2)}{1 - \theta\Phi(u; \mu_1, \sigma_1)\Phi(u; \mu_2, \sigma_2)}, \end{aligned}$$

and

$$f_{\min(U_1, U_2)}(u) = \frac{(1 - \theta)\phi(u; \mu_1, \sigma_1)}{(1 - \theta\Phi(u; \mu_1, \sigma_1))^2} + \frac{(1 - \theta)\phi(u; \mu_2, \sigma_2)}{(1 - \theta\Phi(u; \mu_2, \sigma_2))^2} - f_{\max(U_1, U_2)}(u).$$

4. Inference

In this section, we consider estimation of unknown parameters of the BNG distributions. Let $\{(u_{11}, u_{21}), \dots, (u_{1n}, u_{2n})\}$ be a bivariate sample of size n from BNG with parameters $\Psi = (\mu_1, \sigma_1, \mu_2, \sigma_2, \theta)$. The log likelihood function can be written as

$$\begin{aligned} l_n &= l_n(\Psi) = n \log(1 - \theta) - n \log(\sigma_1) - n \log(\sigma_2) - 2n \log(2\pi) \\ &\quad - \frac{1}{2} \sum_{i=1}^n t_{1i}^2 - \frac{1}{2} \sum_{i=1}^n t_{2i}^2 + \sum_{i=1}^n \log \left[1 + \frac{2\theta\Phi(u_{1i}; \mu_1, \sigma_1)\Phi(u_{2i}; \mu_2, \sigma_2)}{(1 - \theta\Phi(u_{1i}; \mu_1, \sigma_1)\Phi(u_{2i}; \mu_2, \sigma_2))} \right], \end{aligned} \tag{4.1}$$

where $t_{1i} = \frac{u_{1i} - \mu_1}{\sigma_1}$ and $t_{2i} = \frac{u_{2i} - \mu_2}{\sigma_2}$. The maximum likelihood estimators (MLEs) can be obtained by maximizing (4.1) with respect to the unknown parameters. Clearly, MLEs cannot be obtained in closed forms. We may use the Newton-Raphson method, but one of the problems is its 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. The EM algorithm is a very powerful tool in handling the incomplete data problem ([3]). Let the complete-data be $(U_{11}, U_{21}), \dots, (U_{1n}, U_{2n})$ with observed values $(u_{11}, u_{21}), \dots, (u_{1n}, u_{2n})$ and the hypothetical random variable Z_1, \dots, Z_n . We define a hypothetical complete-data distribution with a joint probability density function in the form

$$g(z, u_1, u_2; \Psi) = (1 - \theta)\theta^{z-1}z^2\phi(u_1; \mu_1, \sigma_1)\phi(u_2; \mu_2, \sigma_2)[\Phi(u_1; \mu_1, \sigma_1)\Phi(u_2; \mu_2, \sigma_2)]^{z-1},$$

where $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$ and $z \in \mathbb{N}$. Suppose $\Psi^{(k)} = (\mu_1^{(k)}, \sigma_1^{(k)}, \mu_2^{(k)}, \sigma_2^{(k)}, \theta^{(k)})$ is the current estimate (in the k th iteration) of Ψ . Then, the E-step of an EM cycle requires the expectation of

$(Z|U_1, U_2; \Psi)$. Consider $\theta_* = \theta \Phi(u_1; \mu_1, \sigma_1) \Phi(u_2; \mu_2, \sigma_2)$, then the probability density function of Z given $U_1 = u_1, U_2 = u_2$ is given by

$$g(z|u_1, u_2) = \frac{z^2(1 - \theta_*)^3 [\theta_*]^{z-1}}{\theta_* + 1},$$

and its expected value is

$$E(Z|u_1, u_2; \Psi) = \frac{\theta_*^2 + 4\theta_* + 1}{1 - \theta_*^2}.$$

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

$$\begin{aligned} & l_n^*(\mathbf{u}_1, \mathbf{u}_2, \mathbf{z}; \mu_1, \mu_2, \sigma_1, \sigma_2, \theta) \\ & \propto \sum_{i=1}^n z_i \log \theta - n \log \sigma_1 - n \log \sigma_2 - \frac{1}{2\sigma_1^2} \sum_{i=1}^n (u_{1i} - \mu_1)^2 - \frac{1}{2\sigma_2^2} \sum_{i=1}^n (u_{2i} - \mu_2)^2 \\ & + \sum_{i=1}^n (z_i - 1) \log \Phi(u_{1i}; \mu_1, \sigma_1) + \sum_{i=1}^n (z_i - 1) \log \Phi(u_{2i}; \mu_2, \sigma_2) - n \log(\theta) + n \log(1 - \theta). \end{aligned}$$

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

$$\begin{aligned} & \frac{1}{(\hat{\sigma}_1^{(k)})^2} \sum_{i=1}^n (u_{1i} - \hat{\mu}_1^{(k+1)}) - \sum_{i=1}^n (\hat{z}_i^{(k)} - 1) \frac{\phi(u_{1i}; \hat{\mu}_1^{(k+1)}, \hat{\sigma}_1^{(k)})}{\Phi(u_{1i}; \hat{\mu}_1^{(k+1)}, \hat{\sigma}_1^{(k)})} = 0, \\ & \frac{1}{(\hat{\sigma}_2^{(k)})^2} \sum_{i=1}^n (u_{2i} - \hat{\mu}_2^{(k+1)}) - \sum_{i=1}^n (\hat{z}_i^{(k)} - 1) \frac{\phi(u_{2i}; \hat{\mu}_2^{(k+1)}, \hat{\sigma}_2^{(k)})}{\Phi(u_{2i}; \hat{\mu}_2^{(k+1)}, \hat{\sigma}_2^{(k)})} = 0, \\ & \frac{n}{\hat{\sigma}_1^{(k+1)}} - \frac{1}{(\hat{\sigma}_1^{(k+1)})^3} \sum_{i=1}^n (u_{1i} - \hat{\mu}_1^{(k)})^2 + \frac{1}{\hat{\sigma}_1^{(k+1)}} \sum_{i=1}^n (\hat{z}_i^{(k)} - 1) \frac{(u_{1i} - \hat{\mu}_1^{(k)}) \phi(u_{1i}; \hat{\mu}_1^{(k)}, \hat{\sigma}_1^{(k+1)})}{\Phi(u_{1i}; \hat{\mu}_1^{(k)}, \hat{\sigma}_1^{(k+1)})} = 0, \\ & \frac{n}{\hat{\sigma}_2^{(k+1)}} - \frac{1}{(\hat{\sigma}_2^{(k+1)})^3} \sum_{i=1}^n (u_{2i} - \hat{\mu}_2^{(k)})^2 + \frac{1}{\hat{\sigma}_2^{(k+1)}} \sum_{i=1}^n (\hat{z}_i^{(k)} - 1) \frac{(u_{2i} - \hat{\mu}_2^{(k)}) \phi(u_{2i}; \hat{\mu}_2^{(k)}, \hat{\sigma}_2^{(k+1)})}{\Phi(u_{2i}; \hat{\mu}_2^{(k)}, \hat{\sigma}_2^{(k+1)})} = 0, \\ & \hat{\theta}^{(k+1)} = \frac{1}{n} \hat{\theta}^{(k+1)} (\hat{\theta}^{(k+1)} - 1) \sum_{i=1}^n \hat{z}_i^{(k)}, \end{aligned}$$

where $\hat{\mu}_1^{(k+1)}, \hat{\mu}_2^{(k+1)}, \hat{\sigma}_1^{(k+1)}, \hat{\sigma}_2^{(k+1)}$ and $\hat{\theta}^{(k+1)}$ are found numerically. Here, for $i = 1, \dots, n$, we have that

$$\hat{z}_i^{(k)} = E(Z|U_1 = u_{1i}, U_2 = u_{2i}; \mu_1^{(k)}, \mu_2^{(k)}, \sigma_1^{(k)}, \sigma_2^{(k)}, \theta^{(k)})$$

5. Simulation

This section provides the results of simulation study. Simulations have been performed in order to investigate the proposed estimator of, $\mu_1, \sigma_1, \mu_2, \sigma_2$ and θ of the proposed EM method. Generation from a BNG distribution is straightforward using the definition of the model. First generate N from geometric distribution and call the observed value n , then generate $\{X_1, \dots, X_n\}$ and $\{Y_1, \dots, Y_n\}$ from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, and $U_1 = \max(X_1, \dots, X_n)$ and $U_2 = \max(Y_1, \dots, Y_n)$ is the required sample. We simulate 500 times under the BNG distribution with three different sets of parameters and sample sizes $n = 100, 200$ and 500 . For each sample size, we compute the MLEs by EM method. We also compute the standard error of the MLEs of the EM algorithm. The results for the BNG distribution are shown in Table 1. 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 standard errors of the MLEs decrease.

Table 1. The averages of the MLE's, mean of the simulated standard errors of the BNG distribution.

n	$(\mu_1, \sigma_1, \mu_2, \sigma_2, \theta)$	AE					Std				
		$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}$
100	(0.0,1.0,0.0,1.0,0.1)	-0.0317	0.9896	-0.0334	0.9931	0.1409	0.0150	0.0070	0.0140	0.0070	0.0140
	(0.0,1.0,0.0,1.0,0.3)	0.0214	0.9894	0.0175	0.9949	0.2652	0.0160	0.0070	0.0160	0.0070	0.0160
	(0.0,1.0,0.0,1.0,0.7)	0.1287	0.9794	0.1519	0.9749	0.6113	0.0230	0.0070	0.0220	0.0080	0.0120
200	(0.0,1.0,0.0,1.0,0.1)	-0.0129	0.9956	-0.0158	0.9969	0.1170	0.0073	0.0035	0.0068	0.0036	0.0076
	(0.0,1.0,0.0,1.0,0.3)	0.0203	0.9977	0.0258	0.9904	0.2689	0.0091	0.0036	0.0080	0.0036	0.0086
	(0.0,1.0,0.0,1.0,0.7)	0.1426	0.9791	0.1658	0.9686	0.6139	0.0123	0.0038	0.0109	0.0040	0.0060
500	(0.0,1.0,0.0,1.0,0.1)	-0.0031	0.9973	-0.0058	0.9978	0.1052	0.0030	0.0015	0.0029	0.0014	0.0036
	(0.0,1.0,0.0,1.0,0.3)	0.0094	0.9976	0.0162	0.9977	0.2861	0.0036	0.0015	0.0033	0.0014	0.0036
	(0.0,1.0,0.0,1.0,0.7)	0.1394	0.9795	0.1587	0.9762	0.6216	0.0054	0.0014	0.0049	0.0016	0.0027

6. Application

In this section, we try to illustrate the better performance of the proposed model. For this end, we fit BNG model to a real data set. We also fit the bivariate normal (BN) distribution to make a comparison with the BNG model. This data, taken from [6], are related to air pollution. Here, we consider two variables of these data, viz., Solar rad (U_1) and O_3 (U_2). We first test the marginal distributions. For this purpose, We compute the Kolmogorov-Smirnov distance between the empirical and fitted distribution functions, and the associated p -value (in parentheses) for U_1 and U_2 in Table 2. These results suggest that BNG and BN may be used for analyzing this bivariate data set.

Table 2. Kolmogorov-Smirnov distance and the associated p value.

Dist.	U_1	U_2
NG	0.0637 (0.9956)	0.1466 (0.3275)
N	0.1319 (0.4577)	0.1491 (0.3079)

For comparison purposes, we estimate parameters by numerically maximizing the likelihood function. The MLEs of the parameters, the maximized loglikelihood, the AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) for the BNG and BN models are given in Table 3. As is well-known, a model with a minimum AIC value is to be preferred. Therefore BNG distribution provides a better fit to this data set than BN.

Table 3. Parameter estimates, AIC and BIC for air pollution data.

Dist.	Parameter estimates	$-\log(L)$	AIC	BIC
BNG	$\hat{\mu}_1=-76.90, \hat{\sigma}_1=40.62, \hat{\mu}_2=-35.69, \hat{\sigma}_2=11.99, \hat{\theta}=0.99$	305.2414	620.4827	629.1711
BN	$\hat{\mu}_1=73.85, \hat{\sigma}_1=17.33, \hat{\mu}_2=9.40, \hat{\sigma}_2=5.56, \hat{\rho}=0.319$	307.8487	625.6973	634.3857

7. Conclusion

In this paper we have introduced the bivariate normal–geometric distribution whose marginals are normal–geometric distributions. Several statistical properties of this new bivariate distribution have been studied. The estimation of the unknown parameters of the proposed distribution is approached by the EM algorithm. Finally, we fitted BNG models to a real data set to show the potential of the new proposed class.

Acknowledgments

The authors would like to thank the Referees and the Editors 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

- [1] A. Azzalini, A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **12** (1985) 171–178.
- [2] R. B. Arellano-Valle, and A. Azzalini, On the unification of families of skew-normal distributions, *Scandinavian Journal of Statistics* **33** (2006) 561–574.
- [3] A.P. Dempster, N.M. Laird, and D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm (with discussion), *Journal of the Royal Statistical Society Series B* **39** (1977) 10–38.
- [4] A. A. Jafari and R. Roozgar, On bivariate generalized exponential-power series class of distributions, *arXiv preprint arXiv: (2015) 1508.00219*.
- [5] A. Jamalizadeh and 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.
- [6] R.A. Johnson and D.W. Wichern, *Applied multivariate statistical analysis*, 3-rd edition, (Prentice Hall, New Jersey, 1992).
- [7] E. Mahmoudi and H. Mahmoodian, Normal power series class of distributions: Model, properties and applications, *arXiv preprint arXiv: (2015) 1510.07180*.
- [8] R. Pourmousa and A. Jamalizadeh, On Bivariate Order Statistics from Elliptical Distributions, *Communications in Statistics- Theory and Methods* **43** (2014) 2183–2198.
- [9] R. Roozegar and S. Nadarajah, Bivariate Weibull-power series class of distribution, *Hacettepe Journal of Mathematics and Statistics* (2016) Doi: 10.15672/HJMS.201610714835.