

**BIVARIATE COMPOUND DISTRIBUTION BASED
ON POISSON MAXIMA OF GAMMA VARIATES
AND RELATED APPLICATIONS**

R.J. Abdelghani¹, M.A. Meraou², M.Z. Raqab^{3 §}

^{1,2,3}Department of Mathematics
The University of Jordan
Amman 11942, JORDAN

Abstract: In this paper we introduce a new bivariate model called bivariate compound Poisson-gamma model. The corresponding variable of this model are based on compounding the Poisson number of occurrences and maximum of independent identically distributed gamma variates. For this proposed model, several distributional properties have been established. We implement the EM algorithm based on the missing value principle to find the maximum likelihood estimators. Moreover, we use the observed Fisher information matrix to construct approximate confidence intervals. The performance of the EM type algorithm is illustrated via numerical simulation studies. Finally, a natural environment data have been analyzed to see how the proposed model and the respective methods work in practice.

AMS Subject Classification: 62E15, 62F10

Key Words: bivariate distribution; EM algorithm; Gamma variates; maximum likelihood estimators; compound Poisson distribution

1. Introduction

The construction, study and applications of bivariate distributions is one of the classical active fields of research in statistics, this is due to the uses of bivariate distribution in many fields such as hydrology, medical sciences and environmental studies. In the environmental filed, many climate studies and

Received: February 15, 2021

© 2021 Academic Publications

§Correspondence author

water resources focus on the understanding the properties of hydrology and environmental data. In the dependence between rainfall and runoff see Mathai and Moschopoulos [12]. It is more natural to tackle accurate measurements of the depth snow and the amount falling. This allows us to study the change in ecosystem and the availability of water resources in the coming decades.

Although many bivariate models are considered in the literature, a few of these models can be generalized from the univariate distributions. The univariate models and their bivariate versions can be seen in Cancho et al. [3], Mahmoudi and Jafari [11], and Kundu [8]. In the bivariate compound distributions front, Kzubowski and Panorska [5] formulated a bivariate compound random variable $(\sum_{i=1}^N X_i, N)$ when N has a geometric distribution and $X_i, i = 1, 2, \dots, N$ are independent identically distributed exponential random variables. Applications on this model range from finance and actuarial science to hydrology and climate. Kzubowski et al. [7] showed that the joint distribution of (X, N) when N and X are defined previously, is infinitely divisible. Therefore, it induced a bivariate Lévy process. Kzubowski and panorska [6] explored a bivariate compound distribution based on a geometric maximum of exponential variables. In the context of bivariate gamma distribution, one may refer to Becker and Roux [2] and Barreto-Souza [1]. In fact, Barreto-Souza [1] introduced the bivariate compound random variables (r.v.'s) (X, N) when N has a geometric distribution and X is the sum of iid gamma r.v.'s. Recently, Raqab et al. [13], developed a bivariate compound distribution based on a zero-truncated Poisson sum of independent identically normal random variables.

In the present paper, we propose a new bivariate compound gamma distribution, namely the bivariate compound Poisson gamma (BCP-G) distribution. The corresponding variate of this distribution represents the maximum of Poisson random number of iid gamma random variates. We derive some basic prosperities of the model, including joint and conditional distributions, joint moment generating function, and corresponding correlation coefficient. Moreover, the zero-truncated Poisson distribution can be obtained as a special case of its discrete marginal. In particular, the proposed distribution can be applied in analyzing snowfall data. It can also be applied in other climate variables, such as temperature, air pressure, wind speed and humidity. Now, let $\{X_i : i = 1, 2, \dots\}$ be gamma r.v.'s and let $M = \sum_{i=1}^N I_i$, where the primary distribution N is a Poisson r.v., and the secondary distributions $I_i, i = 1, 2, \dots, N$ are Bernoulli random variables. Assume that all elements of the sequence $\{X_i : i = 1, 2, \dots\}$ are independent identically to each other and independent with N . Next, we define a random variable $U_M = \max\{X_1, \dots, X_M\}$. The new bivariate compound r.v., say, $(U_M, N), M \geq 1$, has a BCP-G distribution. Further, estimation of

the unknown parameters is very important topic in any statistical model. The MLEs of the parameters involved in BCP-G model cannot be obtained in closed forms, one has to solve four non-linear equations system simultaneously. To avoid that, we employed the expectation-maximization (EM) algorithm, which is used in Kundu and Nekoukhou [9] and Raqab et al. [13]. In addition, the confidence intervals of the unknown parameters are obtained using Louis [10] approach.

The layout of the paper are described as follows. In Section 2, we define the BCP-G distribution and derive its related distributional properties. In Section 3, we derive the MLEs of the model parameters using the EM type algorithm, and present their asymptotic properties. In Section 4, we perform a numerical simulation experiment to assess the methods of estimation. In Section 5, we analyze a real data set representing a snowfall real data set for illustrative purposes and to see how the proposed model works in practice when compared to other well-known distributions. Finally, we conclude the paper in Section 6.

2. Model Description and Related Characteristics

The BCP-G distribution is defined by the stochastic representation:

$$(U_M, L) \stackrel{D}{=} \left\{ (\max\{X_1, \dots, X_M\}, N) \mid M \geq 1 \right\},$$

where $\{X_1, X_2, \dots, X_N\}$ is a sequence of iid distributed gamma variables, $G(\alpha, \lambda)$ with a probability density function (PDF)

$$f_X(x; \alpha, \lambda) = h_{\alpha, \lambda}(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, \quad x > 0; \alpha > 0, \lambda > 0, \quad (1)$$

where α and λ are shape and scale parameters. Its cumulative distribution function (CDF) is represented by

$$F(x; \alpha, \lambda) = H_{\alpha, \lambda}(v) = \frac{\Gamma(\alpha; \lambda v)}{\Gamma(\alpha)}, \quad (2)$$

with $\Gamma(\alpha; \lambda v)$ being the incomplete gamma function of the form

$$\Gamma(\alpha; \lambda v) = \int_0^{\lambda v} x^{\alpha-1} e^{-x} dx.$$

Note that $\Gamma(\alpha)$ is the complete gamma function, such that $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$ for all real numbers and $\Gamma(\alpha) = (\alpha - 1)!$ for a positive integer α . Additionally,

one can easily show that N is a Poisson random variable with parameter μ , independent of $\{X_i\}$, with probability mass function (PMF) being

$$p(N = n) = \frac{\mu^n e^{-\mu}}{n!}, \quad n = 0, 1, 2, \dots; \mu > 0. \quad (3)$$

Further, M is a compound random variable defined as

$$M = \sum_{i=1}^N I_i = I_1 + I_2 + \dots + I_N,$$

where $I_i = 0$ means no new event occurred, and $I_i = 1$ means that a new event will be added. So that, $P(I_i = 1) = 1 - P(I_i = 0) = p$, $0 \leq p \leq 1$. The $\stackrel{D}{=}$ stands for equal in distribution.

Given $N = n$, M has a binomial $B(n, p)$ distribution, and thus M is a Poisson with parameter $\theta = \mu p$. To see that, consider the moment generating function (MGF) of M

$$\begin{aligned} \Phi_M(t) = E(e^{tM}) &= E_N E_{M|N}(e^{tM} | N = n) \\ &= e^{\theta(e^t - 1)}. \end{aligned}$$

In our model, we confine ourselves to the achieved events, which means that $M = \sum_{i=1}^N I_i$ takes at least one guaranteed event. Then M follows a zero-truncated Poisson distribution. In our set-up, N is observed but M is not observed. The distribution of (U_M, L) is said to be BCP-G distribution. For simplicity, we use BCP-G(α, λ, μ, p) to denote the BCP-G with parameters α , λ , μ , and p . For notational convenience, we use V variate instead of U_M variate.

The following theorem provides the joint CDF and PDF of (V, L) , for $v > 0$ and $l = 1, 2, \dots$.

Theorem 1. *If (V, L) have BCP-G(α, λ, μ, p) distribution, then the joint CDF and PDF can be obtained, respectively, by*

$$F_{V,L}(v, l) = \frac{e^{-\mu}}{1 - e^{-\theta}} \sum_{n=1}^l \{[(H_{\alpha,\lambda}(v) - 1)p + 1]^n - (1 - p)^n\} \frac{\mu^n}{n!}, \quad (4)$$

and

$$f_{V,L}(v, l) = \frac{e^{-\mu} p}{1 - e^{-\theta}} \frac{\lambda^\alpha v^{\alpha-1} e^{-\lambda v}}{\Gamma(\alpha)} [(H_{\alpha,\lambda}(v) - 1)p + 1]^{l-1} \frac{\mu^l}{(l-1)!}, \quad (5)$$

where $H_{\alpha,\lambda}(v)$ is the CDF corresponding to a gamma random variable defined by 2.

Proof. The joint CDF of (V, L) can be derived through a conditioning argument. It can be described as follows:

$$\begin{aligned} F_{V,L}(v, l) &= P(V \leq v, N \leq l | M \geq 1) \\ &= \frac{1}{1 - e^{-\theta}} \sum_{n=1}^l P(V \leq v, N = n, M \geq 1) \\ &= \frac{e^{-\mu}}{1 - e^{-\theta}} \sum_{n=1}^l \{[(H_{\alpha, \lambda}(v) - 1)p + 1]^n - (1 - p)^n\} \frac{\mu^n}{n!}. \end{aligned}$$

Now, we have for each $L = 1, 2, \dots$,

$$P(V \leq v, L = l) = \frac{e^{-\mu}}{1 - e^{-\theta}} \{[(H_{\alpha, \lambda}(v) - 1)p + 1]^l - (1 - p)^l\} \frac{\mu^l}{l!}. \quad (6)$$

Thus, the joint PDF of (V, L) is

$$f_{V,L}(v, l) = \frac{e^{-\mu} p}{1 - e^{-\theta}} \frac{\lambda^\alpha v^{\alpha-1} e^{-\lambda v}}{\Gamma(\alpha)} [(H_{\alpha, \lambda}(v) - 1)p + 1]^{l-1} \frac{\mu^l}{(l-1)!}.$$

□

Theorem 2. If (V, L) is BCP-MG(α, λ, μ, p) distribution, then the marginal CDF of V , and the marginal PDF of L are given respectively by

$$F_V(v) = \frac{e^{\theta[H_{\alpha, \lambda}(v)]} - 1}{e^\theta - 1}, \quad (7)$$

and

$$P(L = l) = \frac{e^{-\mu} [\mu^l - (\mu - \theta)^l]}{l!(1 - e^{-\theta})}. \quad (8)$$

Proof. The marginal CDF of V

$$\begin{aligned} F_V(v) &= \sum_{l=1}^{\infty} P(V \leq v, L = l) \\ &= \frac{e^{\theta[H_{\alpha, \lambda}(v)]} - 1}{e^\theta - 1}. \end{aligned}$$

The marginal PDF of L is

$$P(L = l) = \lim_{v \rightarrow \infty} P(V \leq v, L = l)$$

$$= \frac{e^{-\mu} [\mu^l - (\mu - \theta)^l]}{l!(1 - e^{-\theta})}.$$

□

It is evident that for $p = 1$, i.e. $\theta = \mu$, L tends to be a zero-truncated Poisson distribution with PMF

$$P(L = l) = \frac{e^{-\mu} \mu^l}{l!(1 - e^{-\mu})}. \quad (9)$$

Theorem 3. If (V, L) is BCP-G(α, λ, μ, p) distribution, then the conditional CDF and PDF given L for $v \geq 0$ are respectively, given by

$$F_{V|L=l}(v) = \frac{[(H_{\alpha,\lambda}(v) - 1)p + 1]^l - (1 - p)^l}{1 - (1 - p)^l}, \quad (10)$$

and

$$f_{V|L=l}(v) = \frac{\lambda^\alpha v^{\alpha-1} e^{-\lambda v} p l}{\Gamma(\alpha) [1 - (1 - p)^l]} [(H_{\alpha,\lambda}(v) - 1)p + 1]^{(l-1)}. \quad (11)$$

Proof. The conditional CDF of V given L for $v \geq 0$ is

$$\begin{aligned} P(V \leq v | L = l) &= \frac{\frac{e^{-\mu}}{1 - e^{-\theta}} \{[(H_{\alpha,\lambda}(v) - 1)p + 1]^l - (1 - p)^l\} \frac{\mu^l}{l!}}{\frac{e^{-\mu} [\mu^l - (\mu - \theta)^l]}{l!(1 - e^{-\theta})}} \\ &= \frac{[(H_{\alpha,\lambda}(v) - 1)p + 1]^l - (1 - p)^l}{1 - (1 - p)^l}. \end{aligned}$$

Equation (11) follows directly from (10). □

Remark 4. It is easily seen that the conditional distribution of V given L can be represented as

$$\{V | L = l\} \stackrel{D}{=} \max_{1 \leq i \leq M} Z_i, \quad (12)$$

where Z_1, \dots, Z_l are iid $G(\alpha, \lambda)$ distribution, and M is a zero-truncated binomial random variable with parameters l and p .

Theorem 5. If $(V, L) \sim \text{BCP-G}(\alpha, \lambda, \mu, p)$ distribution, then the joint MGF of V and L for $t < 1/\lambda$ and $-\infty < s < \infty$ is

$$\Phi_{V,L}(t, s) = \frac{\theta e^{-\mu+s+e^s(\mu-\theta)} \lambda^\alpha}{\Gamma(\alpha)(1-e^{-\theta})} \int_0^\infty e^{(t-\lambda)v+e^s\theta H_{\alpha,\lambda}(v)} v^{\alpha-1} dv.$$

Proof. The joint MGF of V and L for $t < 1/\lambda$ and $-\infty < s < \infty$ can be obtained as follows:

$$\begin{aligned} \Phi_{V,L}(t, s) &= E_{V,L}(e^{tV+sL}) \\ &= \int_0^\infty \sum_{l=1}^\infty e^{tv+sl} \frac{e^{-\mu} p}{1-e^{-\theta}} \frac{\lambda^\alpha v^{\alpha-1} e^{-\lambda v}}{\Gamma(\alpha)} \\ &\quad \times [(H_{\alpha,\lambda}(v) - 1)p + 1]^{l-1} \frac{\mu^l}{(l-1)!} dv \\ &= \frac{\theta e^{-\mu+s+e^s(\mu-\theta)} \lambda^\alpha}{\Gamma(\alpha)(1-e^{-\theta})} \int_0^\infty e^{(t-\lambda)v+e^s\theta H_{\alpha,\lambda}(v)} v^{\alpha-1} dv. \end{aligned}$$

□

Unfortunately, it is not easy to get a closed form for the joint MGF of V and L . The MGF of L for $-\infty < s < \infty$ is given by

$$\Phi_L(s) = \frac{e^{-\mu}}{1-e^{-\theta}} (e^{\mu e^s} - e^{(\mu-\theta)e^s}). \quad (13)$$

Using (13) the first two moment of L are

$$E(L) = \mu + \frac{\theta}{e^\theta - 1}. \quad (14)$$

$$E(L^2) = \mu + \mu^2 + \frac{\theta(1+2\mu-\theta)}{e^\theta - 1}. \quad (15)$$

We now readily obtain $E(LV)$, $E(V)$, and $E(V^2)$. First let X be a $G(\alpha, \lambda)$ r.v. defined in 1. Next, we consider the following representations:

$$\Psi_1(\theta, \alpha, \lambda) = E_X \left[X H_{\alpha,\lambda}(X) e^{\theta H_{\alpha,\lambda}(X)} \right],$$

$$\Psi_2(\theta, \alpha, \lambda) = E_X \left[X e^{\theta H_{\alpha,\lambda}(X)} \right],$$

and

$$\Psi_3(\theta, \alpha, \lambda) = E_X \left[X^2 e^{\theta H_{\alpha,\lambda}(X)} \right].$$

The following two theorems represent the first two moments of V , product moment of V and L , and the corresponding correlation coefficient between V and L . For simplify, let $\Psi_1 = \Psi_1(\theta, \alpha, \lambda)$, $\Psi_2 = \Psi_2(\theta, \alpha, \lambda)$ and $\Psi_3 = \Psi_3(\theta, \alpha, \lambda)$.

Theorem 6. *The product moment of V and L is*

$$E(VL) = \frac{\theta}{e^\theta - 1} \{\Psi_1 + (\mu + 1 - \theta)\Psi_2\}. \quad (16)$$

Proof. The product moment of V and L can be expressed as

$$\begin{aligned} E(VL) &= \int_0^\infty \sum_{l=1}^\infty \frac{v l e^{-\mu} p \lambda^\alpha v^{\alpha-1} e^{-\lambda v} [p H_{\alpha, \lambda}(v) - p + 1]^{l-1} \mu^l}{(1 - e^{-\theta}) \Gamma(\alpha) (l-1)!} dv \\ &= \int_0^\infty \frac{v e^{-\mu}}{1 - e^{-\theta}} \sum_{l=1}^\infty \frac{\partial}{\partial v} \frac{\{[(H_{\alpha, \lambda}(v) - 1)p + 1] \mu\}^l}{(l-1)!} dv \\ &= \frac{\theta}{e^\theta - 1} \{\Psi_1 + (\mu + 1 - \theta)\Psi_2\}. \end{aligned}$$

□

The first two moments of V are given by

$$E(V) = \frac{\theta}{e^\theta - 1} \Psi_2, \quad (17)$$

and

$$E(V^2) = \frac{\theta}{e^\theta - 1} \Psi_3. \quad (18)$$

Corollary 7. *If $(V, L) \sim BCP-G(\alpha, \lambda, \mu, p)$ distribution, then the correlation coefficient between V and L , $\rho(V, L)$, can be represented by:*

$$\frac{(e^\theta - 1)\Psi_1 + (e^\theta - \theta e^\theta - 1)\Psi_2}{\sqrt{[(e^\theta - 1)\Psi_3 - \theta \Psi_2^2] [(e^\theta - 1)^2 + \mu] p^{-1} + e^\theta - \theta e^\theta - 1}}. \quad (19)$$

Proof. Let us consider

$$\begin{aligned} Cov(V, L) &= E(VL) - E(V)E(L) \\ &= \frac{\theta ((e^\theta - 1)\Psi_1 + (e^\theta - \theta e^\theta - 1)\Psi_2)}{(e^\theta - 1)^2}. \end{aligned}$$

The variances of V and L can be readily derived as

$$Var(V) = \frac{\theta}{(e^\theta - 1)^2} \left[(e^\theta - 1)\Psi_3 - \theta(\Psi_2)^2 \right],$$

and

$$Var(L) = \frac{\theta}{(e^\theta - 1)^2} \left[\left((e^\theta - 1)^2 + \mu \right) p^{-1} + (e^\theta - \theta e^\theta - 1) \right].$$

Using the above relations we obtain the result. \square

3. Estimation and Inference

In this section, we consider the estimation of the BCP-G parameters α , λ , μ and p . Let $(V_1, L_1), \dots, (V_k, L_k)$ be a random sample of size k from BCP-G(α, λ, μ, p) distribution. For simplicity, the vector of the parameters is denoted by $\varphi = (\alpha, \lambda, \mu, p)$. The joint log-likelihood function $\ell(\varphi)$ can be expressed in the form of

$$\begin{aligned} \ell(\varphi) &= \sum_{i=1}^k \ln f_{V_i|L_i}(v_i) P(L_i = l_i) \\ &= -k\mu + \alpha k \ln(\lambda) + k \ln(p) - k \ln(\Gamma(\alpha)) - k \ln(1 - e^{-p\mu}) \\ &+ (\alpha - 1) \sum_{i=1}^k \ln(v_i) - \lambda \sum_{i=1}^k v_i + \left(\sum_{i=1}^k l_i \right) \ln(\mu) \\ &+ \sum_{i=1}^k (l_i - 1) \ln [(H_{\alpha, \lambda}(v_i) - 1)p + 1]. \end{aligned}$$

The score function $\phi(\varphi) = \left(\frac{\partial \ell}{\partial \mu}, \frac{\partial \ell}{\partial p}, \frac{\partial \ell}{\partial \alpha}, \frac{\partial \ell}{\partial \lambda} \right)$ for the log likelihood function is given by

$$\begin{aligned} \frac{\partial \ell}{\partial \mu} &= \frac{\sum_{i=1}^k l_i}{\mu} - \frac{kp}{e^{\mu p} - 1} - k, \\ \frac{\partial \ell}{\partial p} &= \frac{k}{p} - \frac{k\mu}{e^{\mu p} - 1} + \sum_{i=1}^k (l_i - 1) \frac{H_{\alpha, \lambda}(v_i) - 1}{(H_{\alpha, \lambda}(v_i) - 1)p + 1}, \end{aligned}$$

$$\begin{aligned}\frac{\partial \ell}{\partial \alpha} &= k \ln(\lambda) - k\psi(\alpha) + \sum_{i=1}^k \ln(v_i) \\ &\quad + p \sum_{i=1}^k (l_i - 1) \frac{\Lambda_{\alpha, \lambda}(v_i) - H_{\alpha, \lambda}(v_i) \psi(\alpha)}{(H_{\alpha, \lambda}(v_i) - 1) p + 1}, \\ \frac{\partial \ell}{\partial \lambda} &= \frac{k\alpha}{\lambda} + p \sum_{i=1}^k (l_i - 1) \frac{\lambda^{\alpha-1} v_i^{\alpha} e^{-\lambda v_i}}{\Gamma(\alpha) [(H_{\alpha, \lambda}(v_i) - 1) p + 1]} - \sum_{i=1}^k v_i,\end{aligned}$$

where $\psi(\alpha) = \partial \log \Gamma(\alpha) / \partial \alpha$ is a digamma function, and

$$\Lambda_{\alpha, \lambda}(v_i) = \frac{1}{\Gamma(\alpha)} \int_0^{\lambda v_i} x^{\alpha-1} \ln x e^{-x} dx.$$

The MLEs of the unknown parameter cannot be obtained in closed forms since we need to solve a 4-dimensional nonlinear system of equations. To avoid the complication, nonlinear optimization algorithms such as EM type algorithm by the missing information principal will be used.

To achieve this, we consider $\{(v_1, l_1, m_1), \dots, (v_k, l_k, m_k)\}$ a random sample of size k from (V, L, M) , where V, L and M are defined as before. In fact, we observe only $\{(v_1, l_1), \dots, (v_k, l_k)\}$, and $\{m_1, \dots, m_k\}$ are missing observations. In our set-up we have

$$\begin{aligned}\{V \mid L = l, M = m\} &\sim f_{V|L=l, M=m}(V) \\ &= m [H_{\alpha, \lambda}(v)]^{m-1} \frac{\lambda^{\alpha} v^{\alpha-1} e^{-\lambda v}}{\Gamma(\alpha)},\end{aligned}\quad (20)$$

$$P(M = m \mid L = l) = \frac{1}{1 - (1-p)^l} \binom{l}{m} p^m (1-p)^{l-m}; \quad m = 1, \dots, l, \quad (21)$$

$$P(L = l) = \frac{e^{-\mu} \mu^l}{l! (1 - e^{-\mu})}; \quad l = 1, 2, \dots \quad (22)$$

The log-likelihood function based on the observed triple sample is

$$\begin{aligned}\ell_{pseudo}(\varphi) &\propto \sum_{i=1}^k (m_i - 1) \ln H_{\alpha, \lambda}(v_i) - \ln(\Gamma(\alpha))^k + \sum_{i=1}^k \ln(v_i)^{(\alpha-1)} \\ &\quad + \ln(\lambda)^{\alpha k} - \sum_{i=1}^k (\ln(1 - (1-p)^{l_i}) - \lambda v_i) + \sum_{i=1}^k \ln(p)^{m_i} \\ &\quad + \sum_{i=1}^k (\ln(1-p)^{(l_i-m_i)} + \ln(\mu)^{l_i}) - k\mu - \ln(1 - e^{-\mu})^k.\end{aligned}$$

Note, we have observed $\{(v_1, l_1), \dots, (v_k, l_k)\}$ but $\{m_1, \dots, m_k\}$ are missing. We propose the EM type algorithm to compute the MLEs of $\varphi = (\alpha, \lambda, \mu, p)$. Suppose that $\varphi^{(j)} = (\alpha^{(j)}, \lambda^{(j)}, \mu^{(j)}, p^{(j)})$ denotes the estimated vector of $\varphi = (\alpha, \lambda, \mu, p)$ at the j -th stage of the EM-type algorithm, and $m_i^{(j)}$ denotes the missing value of m_i at the j -th stage of the EM-type algorithm.

E-step: The E-step of the EM type algorithm is based on the principle that at j -th iterate, we can write the 'pseudo' log likelihood function as

$$\begin{aligned} \ell_{pseudo}(\varphi \mid \varphi^{(j)}) &= \sum_{i=1}^k (m_i^{(j)} - 1) \ln H_{\alpha, \lambda}(v_i) - k \ln \Gamma(\alpha) - \lambda \sum_{i=1}^k v_i \\ &+ (\alpha - 1) \sum_{i=1}^k \ln(v_i) - \sum_{i=1}^k \ln(1 - (1 - p)^{l_i}) - k\mu \\ &+ \left(\sum_{i=1}^k m_i^{(j)} \right) \ln(p) + \sum_{i=1}^k (l_i - m_i^{(j)}) \ln(1 - p) \\ &+ \alpha k \ln(\lambda) + \left(\sum_{i=1}^k l_i \right) \ln(\mu) - k \ln(1 - e^{-\mu}). \end{aligned}$$

Here

$$m_i^{(j)} = \arg \max_m P(M = m \mid L = l_i, p^{(j)}).$$

M-step: The M-step of the EM type algorithm is based on maximizing the 'pseudo' log likelihood function with respect to the unknown parameters to obtain $\varphi^{(j+1)} = (\alpha^{(j+1)}, \lambda^{(j+1)}, \mu^{(j+1)}, p^{(j+1)})$. To obtain $p^{(j+1)}$ we need to solve $\frac{\partial \ell_{pseudo}}{\partial p} = 0$, this is equivalent to solve

$$p = \frac{\sum_{i=1}^k m_i^{(j)}}{\sum_{i=1}^k l_i \left[1 + \frac{(1 - p)^{l_i}}{1 - (1 - p)^{l_i}} \right]}.$$

To obtain $\mu^{(j+1)}$ we need to solve $\frac{\partial \ell_{pseudo}}{\partial \mu} = 0$. That is,

$$\mu = (1 - e^{-\mu}) \bar{L},$$

where $\bar{L} = \frac{\sum_{i=1}^k l_i}{k}$.

To obtain $\alpha^{(j+1)}$ and $\lambda^{(j+1)}$ we need to solve the following two non linear equations:

$$\sum_{i=1}^k (m_i^{(j)} - 1) \left(\frac{\Lambda_{\alpha, \lambda}(v_i)}{H_{\alpha, \lambda}(v_i)} - \psi(\alpha) \right) - k\psi(\alpha) + \sum_{i=1}^k \ln(v_i) + k \ln(\lambda) = 0,$$

and

$$\sum_{i=1}^k (m_i^{(j)} - 1) \frac{v_i^\alpha \lambda^{\alpha-1} e^{-\lambda v_i}}{\Gamma(\alpha) H_{\alpha, \lambda}(v_i)} - \sum_{i=1}^k v_i + \frac{\alpha k}{\lambda} = 0.$$

Now, we are interested in constructing approximate confidence intervals (CIs) for the parameters $\boldsymbol{\varphi} = (\alpha, \lambda, \mu, p)$. For this, we use the method mentioned in Louis [10] to obtain the approximate confidence intervals of the unknown parameters. The MLE $\hat{\boldsymbol{\varphi}}$ of $\boldsymbol{\varphi}$ is just the last step of the EM type algorithm. Then the asymptotic distribution of the MLEs can be obtained as

$$\sqrt{k}(\hat{\boldsymbol{\varphi}} - \boldsymbol{\varphi}) \xrightarrow{d} N_4(0, J_{obs}^{-1}) \quad \text{as } k \rightarrow \infty.$$

Here, J_{obs}^{-1} is the inverse matrix of the Fisher information matrix J_{obs} . The asymptotic multivariate normal distribution $\sqrt{k}(\hat{\boldsymbol{\varphi}} - \boldsymbol{\varphi})$ can be used to construct approximate confidence intervals and confidence regions for the parameters.

The observed Fisher information matrix can be expressed as $J_{obs} = H - SS^T$, where, the matrix $H = ((h_{ij}))$ represents the Hessian matrix of the 'pseudo' log-likelihood function $\ell_{pseudo}(\boldsymbol{\varphi})$ and S is the corresponding gradient vector.

$$H = \begin{pmatrix} \frac{\partial^2 \ell_{pseudo}}{\partial \alpha^2} & \frac{\partial^2 \ell_{pseudo}}{\partial \alpha \partial \lambda} & 0 & 0 \\ \frac{\partial^2 \ell_{pseudo}}{\partial \lambda \partial \alpha} & \frac{\partial^2 \ell_{pseudo}}{\partial \lambda^2} & 0 & 0 \\ 0 & 0 & \frac{\partial^2 \ell_{pseudo}}{\partial \mu^2} & 0 \\ 0 & 0 & 0 & \frac{\partial^2 \ell_{pseudo}}{\partial p^2} \end{pmatrix}.$$

4. Simulation Experiments

In this section, we perform some simulation experiments mainly to see how the proposed EM-type algorithm works in practice. We present a Monte Carlo (MC) simulation results based on samples of sizes $k = 20(20)100$ from BCP-G model with parameters values $\varphi = (0.5, 1, 3.75, 0.55)$, $(0.5, 1, 3.75, 0.65)$, $(1, 1, 3.75, 0.65)$, $(1.5, 1, 3.75, 0.65)$, and $(1.5, 1, 5, 0.65)$. In each particular case, the EM type algorithm is stopped when the convergence occurred such that the absolute differences between successive estimates are less than 10^{-5} for the four estimates. The generation process is repeated 1000 times. For each combination of k and $(\alpha, \lambda, \mu, p)$ we derive the average estimates (AEs) and the mean squared errors (MSEs). Their results are summarized in Tables 1-5. All the numerical computations here are performed using software R.

From Tables 1-5, we see that the MSEs decrease as the sample size increases. It indicates that the proposed EM type algorithm is working well. Also it is noticed that as μ and p increase, the MSEs increase for all the estimates. It verifies the consistency property of the MLEs. It can be checked that when we keep λ , μ , p , fixed and increase α , the MSEs increase for all the estimates. Fixing all parameters and increasing the proportion p , we notice that changes in the MSEs are very small.

Table 1: The AE and the associated MSEs of the MLEs for the BCP-G distribution when $\alpha = 0.5$, $\lambda = 1$, $\mu = 3.75$ and $p = 0.55$.

size		α	λ	μ	p
20	AE	0.4905	0.9531	3.7502	0.4455
	MSE	0.0421	0.1232	0.1881	0.0169
40	AE	0.4135	0.8385	3.7449	0.4844
	MSE	0.0321	0.0741	0.0921	0.0039
60	AE	0.3955	0.8160	3.7346	0.5069
	MSE	0.0270	0.0666	0.0659	0.0150
80	AE	0.3715	0.7936	3.7493	0.5230
	MSE	0.0235	0.0639	0.0469	0.0006
100	AE	0.3657	0.7870	3.7551	0.5460
	MSE	0.0226	0.0627	0.0382	0.0003

Table 2: The AE and the associated MSEs of the MLEs for the BCP-G distribution when $\alpha = 0.5$, $\lambda = 1$, $\mu = 3.75$ and $p = 0.65$.

size		α	λ	μ	p
20	AE	0.4188	0.9076	3.7229	0.6352
	MSE	0.0501	0.1271	0.1950	0.0196
40	AE	0.3744	0.8192	3.7641	0.6397
	MSE	0.0404	0.0827	0.0989	0.0069
60	AE	0.3551	0.7885	3.7619	0.6428
	MSE	0.0289	0.0761	0.0668	0.0016
80	AE	0.3489	0.7762	3.7485	0.6415
	MSE	0.0284	0.0732	0.0472	0.0008
100	AE	0.3405	0.7656	3.7586	0.6501
	MSE	0.0274	0.0721	0.0390	0.0003

Table 3: The AE and the associated MSEs of the MLEs for the BCP-G distribution when $\alpha = 1$, $\lambda = 1$, $\mu = 3.75$ and $p = 0.65$.

size		α	λ	μ	p
20	AE	0.8552	0.9023	3.7603	0.6375
	MSE	0.2253	0.1410	0.2022	0.0202
40	AE	0.7224	0.8001	3.7525	0.6377
	MSE	0.1388	0.0868	0.1013	0.0062
60	AE	0.7021	0.7791	3.7416	0.6385
	MSE	0.1234	0.0786	0.0670	0.0018
80	AE	0.6784	0.7641	3.7492	0.6414
	MSE	0.1159	0.0764	0.0527	0.0009
100	AE	0.6779	0.7607	3.7512	0.6492
	MSE	0.0916	0.0745	0.0395	0.0004

Table 4: The AE and the associated MSEs of the MLEs for the BCP-G distribution when $\alpha = 1.5$, $\lambda = 1$, $\mu = 3.75$ and $p = 0.65$.

size		α	λ	μ	p
20	AE	1.2790	0.9049	3.7496	0.6368
	MSE	0.3266	0.1453	0.2198	0.0208
40	AE	1.0995	0.8071	3.7377	0.6388
	MSE	0.2916	0.0918	0.1045	0.0069
60	AE	1.0587	0.7632	3.7422	0.6415
	MSE	0.1990	0.0788	0.0901	0.0026
80	AE	1.0135	0.7540	3.7531	0.6431
	MSE	0.1473	0.0774	0.0686	0.0014
100	AE	1.0101	0.7458	3.7536	0.6494
	MSE	0.0928	0.0752	0.0480	0.0009

Table 5: The AE and the associated MSEs of the MLEs for the BCP-G distribution when $\alpha = 1.5$, $\lambda = 1$, $\mu = 5$ and $p = 0.65$.

size		α	λ	μ	p
20	AE	1.1661	0.8471	4.9968	0.7150
	MSE	0.4227	0.1523	0.2620	0.0215
40	AE	1.1499	0.7725	4.9823	0.7041
	MSE	0.3190	0.0940	0.1215	0.0176
60	AE	1.1297	0.7537	4.9939	0.6891
	MSE	0.2182	0.0892	0.1028	0.0121
80	AE	1.1054	0.7429	5.0068	0.6717
	MSE	0.1545	0.0842	0.0986	0.0093
100	AE	1.0895	0.7404	5.0103	0.6511
	MSE	0.0946	0.0840	0.0522	0.0054

5. Application to Annual Snowfall Records

In this section, we present the analysis of a real data set representing annual maximum snowfall measured in centimeter (cm). This data is taken from several stations located in Canada. The considered data contains the maximum annual record of snowfall (V) and the total number of no snowfall event in each month (L). The data are available in the World Meteorological Organization (WMO) (<https://www.data.un.org>). The WMO provides the basis for better understanding the climatology of severe weather and extreme events. Some basic statistics of the observed value of V , it is presented in Table 6. The histogram with the fitted PDF of the data set is displayed in Figure 1.

Table 6: Basic statistics of of annual maximum snowfall data set

Mean	Median	Std.Dev.	Q_1	Q_3	skewness
40.6	33.8	19.3	26.2	50.8	1.24

From the skewness measure in Table 6 and Figure 1, it is clear that the annual maximum snowfall data is positively skewed.

Now, we would like to fit the BCP-G to the bivariate data set. We have used the EM type algorithm to compute the MLEs of the unknown parameters. The estimates and the corresponding 95% CIs of the unknown parameters and

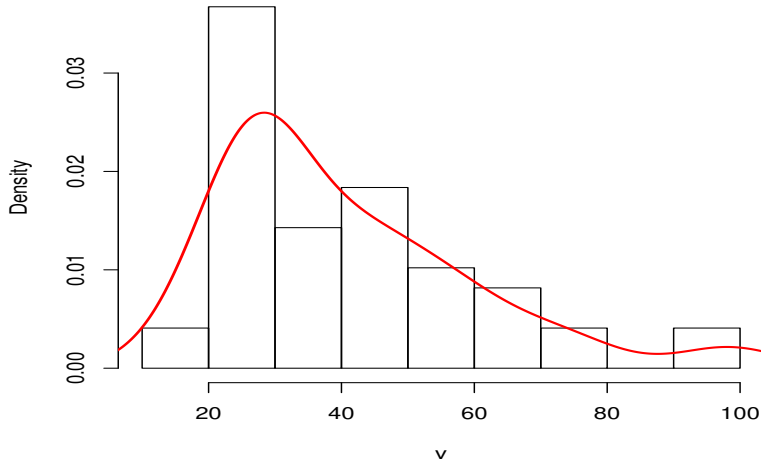


Figure 1: Histogram and fitted PDF for the annual maximum snow-fall data set.

associated log-likelihood (ll) are provided as follows

$$\hat{\alpha} = 3.5095, \quad \hat{\lambda} = 0.1084, \quad \hat{\mu} = 2.9376, \quad \hat{p} = 0.5239, \quad ll = -172.1588$$

$$(3.1495, 3.8696), \quad (0.0922, 0.1246), \quad (2.4267, 3.4485), \quad (0.4304, 0.6174).$$

To check whether the proposed EM type algorithm actually converges to the MLEs, we have performed a grid search method with size 0.001 and it is observed that they match. Also, we have tried the EM type algorithm with some other initial estimates, and it converges to the same set of estimates, which ensures that the proposed EM type algorithm actually converges to the MLEs.

The univariate compound Poisson gamma CP-G(α, λ, θ) distribution is fitted to the univariate data set. More details of the model is presented in Appendix A. The MLEs and the corresponding 95% CIs of α , λ and θ by using the EM type algorithm have been reported as follows

$$\hat{\alpha} = 1.7278, \quad \hat{\lambda} = 0.0889, \quad \hat{\theta} = 6.1115, \quad ll = -203.966$$

$$(1.4058, 2.0498), \quad (0.0700, 0.1079), \quad (4.3243, 7.8986).$$

The gamma distribution is fitted to the univariate data set. The estimates of α and λ and the corresponding 95% CIs value have been reported as

$$\hat{\alpha} = 5.3370, \quad \hat{\lambda} = 0.1312, \quad ll = -206.7671$$

$$(4.6849, 5.9766), \quad (0.1153, 0.1472).$$

We also fit the exponential (EXP) distribution to the univariate data set. The MLE of λ , the associated 95% CIs and the corresponding log-likelihood (ll) values are provided as follows

$$\hat{\lambda} = 0.0246(\mp 0.0069), \quad \text{ll} = -230.4896.$$

For checking the model validity, several criteria are computed such as Akaike Information Criterion (AIC), Akaike Information Criterion correction (AICc), Hannan-Quinn Information Criterion (HQIC), Bayesian Information Criterion (BIC) and the Kolmogorov-Smirnov (K-S) with associated p -value. Table 7 gives the so obtained results. Note that the values of AIC, AICc, HQIC and BIC for BCP-G are smaller than these of CP-G, Gamma and EXP distributions. Thus, BCP-G is indeed an appropriate distribution for the snowfall data set.

Table 7: The goodness of fit tests for annual maximum snowfall data set

Distribution	AIC	AICc	HQIC	BIC	K-S	p -value
BCP-G	352.31	353.22	355.18	359.88	0.1200	0.4830
CP-G	413.93	414.46	416.08	419.60	0.1341	0.3415
Gamma	417.53	417.79	418.96	421.31	0.1381	0.3067
EXP	462.97	463.06	463.69	464.87	0.3599	6.122×10^{-06}

We observe from Table 7 based on the p -value that the proposed BCP-G model provides a better fit than CP-G, Gamma and EXP distributions to the snowfall data set. Figure 2 shows the plots of the fitted PDFs and CDFs with their associated empirical distribution functions. Also, the empirical survival functions (ESFs) and fitted SFs are displayed in Figure 3. All these plots ensure the same conclusion.

6. Conclusions

We have presented a new bivariate compound distribution based on compounding the Poisson number of occurrences and maximum of independent identically distributed gamma variates. The proposed BCP-G model can be quite effectively to model the skewed data sets. We have studied several properties of the proposed model. Since, the MLEs of the unknown parameters cannot be

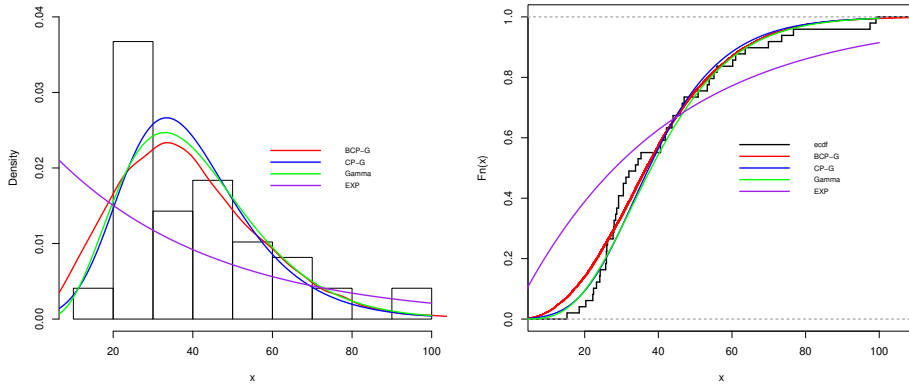


Figure 2: Plots for the PDFs and CDFs of fitted BCP-G, CP-G, Gamma and EXP distributions for annual maximum snowfall data set.

obtained in explicit forms, we have adopted the EM type algorithm to estimate the unknown parameters. It is observed based on extensive simulation experiments that the proposed EM type algorithm works quite well in practice. Further, we fitted BCP-G model to real snowfall data set and found that the model is the best fitting model when compared to other related models.

Appendix A: The Univariate Compound Poisson Gamma Distribution

Here, we present some derivations for CP-G model based on combining the gamma and zero-truncated Poisson distributions. Let us define the random variable Y as a random Poisson maximum of gamma random variables. Precisely,

$$Y = \max\{X_1, X_2, \dots, X_{L^*}\},$$

where, $\{X_1, X_2, \dots\}$ is sequence of iid $G(\alpha, \lambda)$ variables and L^* is zero-truncated Poisson distribution with parameter $\theta = \mu p$, where $p = P(I_i = 1) = 1 - P(I_i = 0); i = 1, 2, \dots, N$.

The joint PDF of CP-G of Y and L^* is given by

$$f_{Y, L^*}(y, l) = \frac{\theta^l \lambda^\alpha y^{\alpha-1} e^{-\lambda y}}{\Gamma(\alpha)(e^\theta - 1)(l-1)!} (H_{\alpha, \lambda}(y))^{l-1}.$$

The CDF of Y for $0 < y < \infty$, is given by

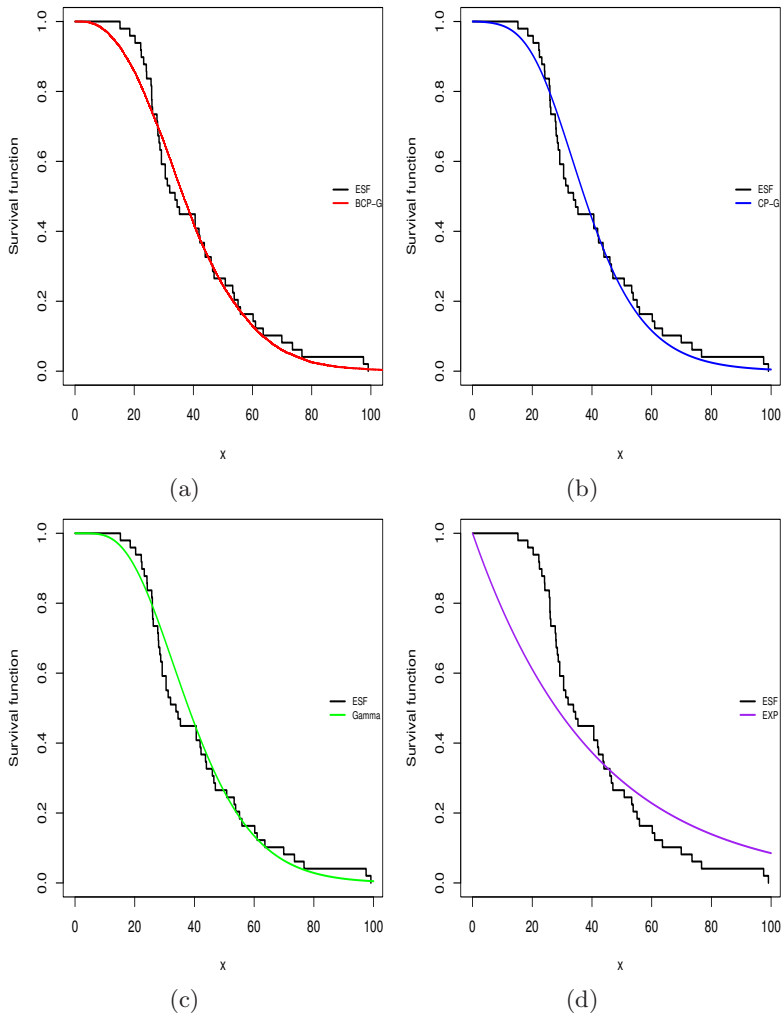


Figure 3: The ESFs and fitted SFs for different distributions: (a) BCP-G (b) CP-G (c) Gamma (d) EXP .

$$\begin{aligned}
 F_Y(y) &= \sum_{l=1}^{\infty} P(Y \leq y, L^* = l) = \frac{1}{e^{\theta} - 1} \sum_{l=1}^{\infty} (H_{\alpha, \lambda}(y))^l \frac{\theta^l}{l!} \\
 &= \frac{e^{\theta H_{\alpha, \lambda}(y)} - 1}{e^{\theta} - 1}.
 \end{aligned}$$

Its respective PDF of Y is obtained to be

$$f_Y(y) = \frac{\theta \lambda^\alpha y^{\alpha-1} e^{-\lambda y}}{(e^\theta - 1)\Gamma(\alpha)} e^{\theta H_{\alpha,\lambda}(y)}, \quad y > 0.$$

Acknowledgements

The authors thank the editor and referees for their comments and helpful suggestions.

References

- [1] W. Barreto-Souza, Bivariate gamma-geometric law and its induced lévy process, *Journal of Multivariate Analysis*, **109** (2012), 130-145.
- [2] P.J. Becker, J.J.J. Roux, A bivariate extension of the gamma distribution, *South African Statistical Journal*, **15** (1981), 1-12.
- [3] V.G. Cancho, F. Louzada-Neto and G.D.C. Barriga, The Poisson-exponential lifetime distribution, *Computational Statistics and Data Analysis*, **55** (2011), 677-686.
- [4] M. Chahkandi, M. Ganjali, On some lifetime distributions with decreasing failure rate, *Computational Statistics and Data Analysis*, **53** (2009), 4433-4440.
- [5] T.J. Kozubowski, A.K. Panorska, A mixed bivariate distribution with exponential and geometric marginals, *Journal of Statistical Planning and Inference*, **134** (2005), 501-520.
- [6] T.J. Kozubowski, A.K. Panorska, A mixed bivariate distribution connected with geometric maxima of exponential variables, *Communications in Statistics-Theory and Methods*, **37** (2008), 2903-2923.
- [7] T.J. Kozubowski, A.K. Panorska and K. Podgórski, A bivariate Lévy process with negative binomial and gamma marginals, *Journal of Multivariate Analysis*, **99** (2008), 1418-1437.
- [8] D. Kundu, Geometric skew normal distribution, *Sankhya*, **B76**, No 2 (2014), 167-189.

- [9] D. Kundu, V. Nekoukhou, Univariate and bivariate geometric discrete generalized exponential distribution, *Journal of Statistical Theory and Practice*, **12**, No 3 (2018), 595-614.
- [10] T.A. Louis, Finding the observed information matrix when using the EM algorithm, *Journal of the Royal Statistical Society*, Ser. **B44** (1982), 226-233.
- [11] E. Mahmoudi, A.A. Jafari, Genaralized exponential-power series distributions, *Computational Statistics and Data Analysis*, **56** (2012), 4047-4066.
- [12] A.M. Mathai, P.G. Moschopoulos, On a multivariate gamma, *Journal of Multivariate Analysis*, **39** (1991), 135-153.
- [13] M.Z. Raqab, D. Kundu and A.F. Fahimah, Compound zero-truncated Poisson normal distribution and its Applications, *Communications in Statistics-Theory and Methods*, **50**, No 13 (2021), 3030–3050.