

Binomial-exponential 2 Distribution: Different Estimation Methods with Weather Applications

H.S. BAKOUCH¹, S. DEY², P.L. RAMOS^{3*} and F. LOUZADA³

Received on September 8, 2016 / Accepted on May 13, 2017

ABSTRACT. In this paper, we have considered different estimation methods of the unknown parameters of a binomial-exponential 2 distribution. First, we briefly describe different methods of estimation such as maximum likelihood, method of moments, percentile based estimation, least squares, method of maximum product of spacings, method of Cramér-von-Mises, methods of Anderson-Darling and right-tail Anderson-Darling, and compare them using extensive simulations studies. Finally, the potentiality of the model is studied using three real data sets related to the total monthly rainfall during April, May and September at São Carlos, Brazil.

Keywords: binomial-exponential 2, maximum likelihood estimation, Cramér-von-Mises type minimum distance estimators, right-tail Anderson-Darling estimators.

1 INTRODUCTION

The binomial-exponential 2 (*BE2*) distribution has been introduced by Bakouch et al. [6] as a distribution of a random sum of independent exponential random variables when the sample size has a zero truncated binomial distribution. The *BE2* distribution has the probability density function (pdf)

$$f(x; \theta, \lambda) = \left(1 + \frac{(\lambda x - 1)\theta}{2 - \theta}\right) \lambda e^{-\lambda x}, \tag{1.1}$$

and the cumulative distribution function (cdf)

$$F(x; \theta, \lambda) = 1 - \left(1 + \frac{\lambda \theta x}{2 - \theta}\right) e^{-\lambda x}, \tag{1.2}$$

where $0 \leq \theta \leq 1$ is the shape parameter and $\lambda > 0$ is the scale parameter. The *BE2* distribution has an increasing and constant failure rate property. A generalization of the *BE2* distribution was discussed by Asgharzadeh et al. [4].

*Corresponding author: Pedro Luiz Ramos – E-mail: pedrolramos@usp.br

¹Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt. E-mail: hnbakouch@yahoo.com

²Department of Statistics, St. Anthony’s College, Shillong, Meghalaya, India. E-mail: sankud66@gmail.com

³Instituto de Matemática e Ciência da Computação, Universidade de São Paulo, São Carlos, SP, Brasil.
 E-mail: louzada@icmc.usp.br

Bakouch et al. [6] in their paper only considered the maximum likelihood estimation (MLE) method to estimate the parameters of the BE2 distribution. However it is of interest to compare the MLE method with other estimation procedures such as the method of moments, ordinary least-squares estimation (OLSE), weighted least-squares estimation (WLSE), percentile estimation (PCE), maximum product of spacings estimation (MPS), Cram er-von-Mises type minimum distance estimation (CME), Anderson-Darling (ADE) and Right-tail Anderson-Darling estimation (RADE).

We have several estimation methods available for the parametric distribution in the literature, some of the estimation methods are well researched on theoretical aspect. However, it is worth noting that in the case of small samples, there is often evidence that the maximum likelihood method does not perform well. Therefore, other estimating methods have recently been developed. The appeal of the estimation methods vary from user to user and area of application. For instance, one may prefer to use the moment estimator even when it does not have a closed form expression. The objective of the article is to develop a guideline for choosing the best estimation method for the BE2 distribution, which would be of interest to applied statisticians. Comparisons of estimation methods for other distributions have been investigated in the literature, see e.g., [1,5,11,13,17,19,21,26].

The main goal of this paper is two fold: First is to show how different frequentist estimators of the proposed distribution perform for different sample sizes and second is to show that the distribution outperforms at least two-parameter distributions with respect to three real data sets.

Other motivation to use the BE2 distribution comes from the fact that stochastic models that accommodate zero value has vast importance in practical applications, for example in forecast models when we observe the monthly rainfall precipitation, it is common in dry periods the non occurrence of precipitation, therefore the occurrence of zero value can be observed in different measures such as the average, maximum and minimum. Popular models such as Gamma, Weibull, Lognormal and Generalized Exponential distributions do not accommodate such characteristic. In this paper we demonstrate that the BE2 distribution allows the occurrence of zero value, becoming a simple alternative to be used in weather forecast models.

The paper is organized as follows. In Section 2, we present some notes and properties for the model. In Section 3, we discuss the nine estimation methods considered in this paper. In Section 4 a simulation study is presented in order to identify the most efficient estimators. In Section 5 we apply our proposed methodology to three real data sets related to the total monthly rainfall during April, May and September at S ao Carlos, Brazil. Finally in Section 6 we conclude the paper.

2 NOTES AND PROPERTIES

Note that the family of Lindley distributions is a subfamily of the BE2 family for $\theta = \frac{2}{2+\lambda}$. Also, for another motivation, recall that the p.d.f. of the BE2 distribution can be expressed as a two-component mixture of an exponential distribution (with scale parameter λ) and a gamma

distribution (with shape 2 and scale λ), i.e. $f(x; p, \lambda) = p\lambda^2xe^{-\lambda x} + (1 - p)\lambda e^{-\lambda x}$, where the mixing proportion $p = \frac{\theta}{2-\theta}$.

Let $X \sim BE2(\theta, \lambda)$, the raw moments of X about the origin is given by

$$E(X^r) = \frac{r!}{\lambda^r} \left(1 + \frac{r\theta}{2 - \theta} \right), \tag{2.1}$$

and the survival function of X is given by

$$S(x; \theta, \lambda) = \left(1 + \frac{\lambda\theta x}{2 - \theta} \right) e^{-\lambda x}.$$

Many distributions such as Gamma, Weibull, Lognormal, to list a few, do not allow occurrence of zero values. The following proposition prove that the BE2 distribution can be used as a model with occurrence of zero value.

Proposition 2.1. *Let X be a random variable with BE2 distribution then $f_X(0; \theta, \lambda) \geq 0$ for all $0 \leq \theta \leq 1$ and $\lambda > 0$.*

Proof. Note that

$$f_X(0; \theta, \lambda) = \frac{d}{dx}F_X(x; \theta, \lambda)\Big|_0 = \left(\frac{2 - 2\theta}{2 - \theta} \right) \lambda \tag{2.2}$$

where $f_X(0; \theta, \lambda) \geq 0$ for all $0 \leq \theta \leq 1$ and $\lambda > 0$.

This result allows us to use the BE2 distribution as a simple alternative in the problems with occurrence of zero value.

3 METHODS OF ESTIMATION

In this section, nine estimation procedures are discussed to obtain the estimates of the BE2 distribution parameters.

3.1 Maximum Likelihood Estimation

The method of maximum likelihood is the most frequently used method of parameter estimation. The method's success stems no doubt from its many desirable properties including consistency, asymptotic efficiency, normality, invariance and simply its intuitive appeal. Let x_1, \dots, x_n be a random sample of size n from (1.1), the likelihood function of the density (1.1) is given by

$$L(\theta, \lambda; x) = \prod_{i=1}^n f(x_i, \theta, \lambda) = \lambda^n \exp \left(-\lambda \sum_{i=1}^n x_i \right) \prod_{i=1}^n \left(1 + \frac{(\lambda x_i - 1)\theta}{2 - \theta} \right) \tag{3.1}$$

The log-likelihood function without constant terms is given by

$$\ell(\theta, \lambda; x) = n \log \lambda - \lambda \sum_{i=1}^n x_i - n \log(2 - \theta) + \sum_{i=1}^n \log(2 - 2\theta + \lambda\theta x_i). \tag{3.2}$$

From the expressions $\frac{\partial}{\partial \theta} l(\theta, \lambda; \mathbf{x}) = 0$, $\frac{\partial}{\partial \lambda} l(\theta, \lambda; \mathbf{x}) = 0$, the likelihood equations are

$$\frac{n}{\lambda} - \sum_{i=1}^n x_i + \sum_{i=1}^n \frac{\theta x_i}{(2 - 2\theta + \lambda \theta x_i)} = 0 \tag{3.3}$$

and

$$\frac{n}{2 - \theta} + \sum_{i=1}^n \frac{\lambda x_i - 2}{2 - 2\theta + \lambda \theta x_i} = 0. \tag{3.4}$$

The maximum likelihood estimator $\hat{\theta}$ and $\hat{\lambda}$ are obtained by solving the non-linear equations (3.3) and (3.4). It is important to point out that, non-linear optimization algorithms such as the quasi-Newton algorithm, can be used to maximize directly the likelihood function given in (3.1).

3.2 Moments Estimators

The method of moments is fairly simple procedure and has been widely used for estimating parameters in statistical models. The moments estimators (MEs) of the BE2 distribution can be obtained by equating the theoretical moments of (1.1) with the sample moments. Consider that

$$E(X|\theta, \lambda) = \frac{2}{\lambda(2 - \theta)} \quad \text{and} \quad Var(X|\theta, \lambda) = \frac{2(2 - \theta^2)}{\lambda^2(2 - \theta)^2}. \tag{3.5}$$

are the theoretical moments of the BE2 distribution. Note that, the population coefficient of variation given by

$$CV(X|\theta, \lambda) = \frac{\sqrt{2(2 - \theta^2)}}{2}$$

is independent of the scale parameter λ . So, the estimator $\hat{\theta}_{MME}$ for θ and $\hat{\lambda}_{MME}$ for λ , can be easily obtained by solving

$$\begin{aligned} \hat{\theta}_{MME} &= \sqrt{2 - 2\left(\frac{s}{\bar{x}}\right)^2}, \quad \text{and} \\ \hat{\lambda}_{MME} &= \frac{2}{\bar{x} \left(2 - \sqrt{2 - 2\left(\frac{s}{\bar{x}}\right)^2}\right)} \end{aligned} \tag{3.6}$$

where \bar{x} and s are the sample mean and sample standard deviation respectively.

3.3 Least-Square Estimators

The ordinary least square and the weighted least square are well known methods used for estimating the unknown parameters [25]. Let $F(x)$ be the distribution function of the random variables $\{X_1, X_2, \dots, X_n\}$ and $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ be ordered random variables. The least square estimators of θ and λ , denoted by $\hat{\theta}_{LSE}$ and $\hat{\lambda}_{LSE}$ can be obtained by minimizing the function

$$LS(\theta, \lambda) = \sum_{i=1}^n \left[F(x_{(i)} | \theta, \lambda) - \frac{i}{n+1} \right]^2 \tag{3.7}$$

with respect to θ and λ , where $F(\cdot)$ is given by (1.2). Equivalently, they can be obtained by solving the following non-linear equations:

$$\sum_{i=1}^n \left[F(x_{(i)} | \theta, \lambda) - \frac{i}{n+1} \right] \eta_1(x_{(i)} | \theta, \lambda) = 0,$$

$$\sum_{i=1}^n \left[F(x_{(i)} | \theta, \lambda) - \frac{i}{n+1} \right] \eta_2(x_{(i)} | \theta, \lambda) = 0.$$

Consider the following weighted function (see Gupta & Kundu [13])

$$w_i = \frac{1}{\text{Var}(F(x_{(i)}))} = \frac{(n+1)^2(n+2)}{i(n-i+1)}.$$

The WLSEs, $\hat{\theta}_{WLSE}$ and $\hat{\lambda}_{WLSE}$, can be obtained by minimizing

$$\text{WLS}(\theta, \lambda) = \sum_{i=1}^n \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F(x_{(i)} | \theta, \lambda) - \frac{i}{n+1} \right]^2. \quad (3.8)$$

These estimators can also be obtained by solving:

$$\sum_{i=1}^n \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F(x_{(i)} | \theta, \lambda) - \frac{i}{n+1} \right] \eta_1(x_{(i)} | \theta, \lambda) = 0,$$

$$\sum_{i=1}^n \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F(x_{(i)} | \theta, \lambda) - \frac{i}{n+1} \right] \eta_2(x_{(i)} | \theta, \lambda) = 0,$$

where

$$\eta_1(x_{(i)} | \theta, \lambda) = \frac{-2\lambda x_{(i)} e^{-\lambda x_{(i)}}}{(2-\theta)^2}, \quad (3.9)$$

and

$$\eta_2(x_{(i)} | \theta, \lambda) = x_{(i)} e^{-\lambda x_{(i)}} \left(1 + \frac{\lambda \theta x_{(i)}}{2-\theta} - \frac{\theta x_{(i)} e^{-\lambda x_{(i)}}}{2-\theta} \right). \quad (3.10)$$

3.4 Percentile Estimators

The percentile estimators is originally suggested by Kao [15, 16]. This method is commonly used to estimate the unknown parameters from the distribution functions that has a closed form of the quantile function. The percentile estimates (PCEs) can be obtained by minimizing with respect unknown parameters, the Euclidean distance between the ordered sample points and ordered theoretical points, computed throughout the quantile function. Since,

$$F(x, \theta, \lambda) = 1 - \left(1 + \frac{\lambda \theta x}{2-\theta} \right) e^{-\lambda x}$$

therefore, the quantile function is given by

$$x_p = \frac{1}{\lambda} \log \left(\frac{2 - \theta + \lambda \theta x_p}{(2 - \theta)(1 - p)} \right).$$

Let $X_{(j)}$ be the j th order statistics, i.e., $X_{(1)} < X_{(2)} < \dots < X_{(n)}$. If p_j denotes some estimators of $F(x_{(j)}; \theta, \lambda)$, then the estimators of θ and λ can be obtained by minimizing

$$\sum_{j=1}^n \left(x_{(j)} - \frac{1}{\lambda} \ln \frac{2 - \theta + \lambda \theta x_p}{(2 - \theta)(1 - p_j)} \right)^2 \tag{3.11}$$

with respect to θ and λ . The percentile estimators $\hat{\theta}_{PCE}$ and $\hat{\lambda}_{PCE}$ can be obtained by solving the following nonlinear equations

$$\sum_{j=1}^n \left[x_j - \frac{1}{\lambda} \log \left(\frac{(2 - \theta + \lambda \theta x_p)}{(2 - \theta)(1 - p_j)} \right) \right] \left(\frac{x_p}{(2 - \theta + \lambda \theta x_p)(2 - \theta)} \right) = 0,$$

$$\sum_{j=1}^n \left[x_j - \frac{1}{\lambda} \log \left(\frac{(2 - \theta + \lambda \theta x_p)}{(2 - \theta)(1 - p_j)} \right) \right] \left[\frac{1}{\lambda^2} \log \left(\frac{(2 - \theta + \lambda \theta x_p)}{(2 - \theta)(1 - p_j)} \right) - \frac{1}{\lambda} \frac{\theta x_p}{(2 - \theta + \lambda \theta x_p)} \right] = 0,$$

respectively. In this paper, we consider the estimator of p_j as $p_j = \frac{j}{n+1}$. However, different estimators can be used instead, see for example Mann, et al. (1974).

3.5 Method of Maximum Product of Spacings

The maximum product spacing (MPS) method has been introduced by Cheng & Amin [9] as an alternative to MLE for the estimation of the unknown parameters of continuous univariate distributions. The MPS method was also derived independently by Ranneby [22] as an approximation to the Kullback-Leibler measure of information. This method is as efficient as the MLE estimators and consistent under more general conditions. Using the same notations in subsection 3.3, define the uniform spacings of a random sample from the BE2 distribution as:

$$D_i(\theta, \lambda) = F(x_{i:n} | \theta, \lambda) - F(x_{i-1:n} | \theta, \lambda), \quad i = 1, 2, \dots, n,$$

where $F(x_{0:n} | \theta, \lambda) = 0$ and $F(x_{n+1:n} | \theta, \lambda) = 1$. Clearly $\sum_{i=1}^{n+1} D_i(\theta, \lambda) = 1$.

The maximum product of spacings estimators $\hat{\theta}_{MPS}$ and $\hat{\lambda}_{MPS}$, of the parameters θ and λ are obtained by maximizing the geometric mean of the spacings:

$$G(\theta, \lambda) = \left[\prod_{i=1}^{n+1} D_i(\theta, \lambda) \right]^{\frac{1}{n+1}}, \tag{3.12}$$

or, equivalently, by maximizing the function

$$g(\theta, \lambda) = \frac{1}{n+1} \sum_{i=1}^{n+1} \log D_i(\theta, \lambda). \tag{3.13}$$

with respect to θ and λ . Although Cheng & Amin [9] proved that the MPS is asymptotically equivalent to the MLE, the authors do not present a motivation in maximizing the geometric mean. However, Cheng & Stephens [10] showed that the MPS is also a minimum goodness of fit estimator based on the Moran's statistics given by

$$M(\theta, \lambda) = - \sum_{i=1}^{n+1} \log D_i(\theta, \lambda)$$

i.e., to find the minimum of the Moran's statistics is the same as finding the maximum of the geometric mean of the spacings. Hence, the estimators $\widehat{\theta}_{MPS}$ and $\widehat{\lambda}_{MPS}$ of the parameters θ and λ can be obtained by solving the nonlinear equations

$$\frac{\partial g(\theta, \lambda)}{\partial \theta} = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\theta, \lambda)} [\eta_1(x_{(i)}|\theta, \lambda) - \eta_1(x_{i-1:n}|\theta, \lambda)] = 0, \quad (3.14)$$

$$\frac{\partial g(\theta, \lambda)}{\partial \lambda} = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\theta, \lambda)} [\eta_2(x_{(i)}|\theta, \lambda) - \eta_2(x_{i-1:n}|\theta, \lambda)] = 0, \quad (3.15)$$

where $\eta_1(\cdot | \theta, \lambda)$ and $\eta_2(\cdot | \theta, \lambda)$ are given by (9) and (10), respectively.

3.6 Methods of Minimum Distances

In this subsection, we present three minimum distance estimators (also called maximum goodness-of-fit estimators) for θ and λ . This class of estimators are based on minimizing any empirical distribution function (EDF) statistics with respect to the unknown parameters [18].

3.6.1 Method of Cramér-von-Mises

To motivate our choice of Cramér-von-Mises (CVM) type minimum distance estimators, MacDonald (1971) provided empirical evidence that the bias of the estimator is smaller than the other minimum distance estimators. Thus, the proposed estimators are based on the Cramér-von Mises statistics given by

$$W_n^2 = n \int_{-\infty}^{\infty} (F(x_{(i)}) - E_n(x_{(i)}))^2 dF(x_{(i)})$$

where $E_n(\cdot)$ is the empirical density function. Boos [7] presented a detailed discussion about this CVM estimator. Moreover, the author presented its computational form which is given by

$$C(\theta, \lambda) = \frac{1}{12n} + \sum_{i=1}^n \left(F(x_{(i)} | \theta, \lambda) - \frac{2i-1}{2n} \right)^2. \quad (3.16)$$

Then the CMV estimators are obtained by minimizing (3.16) with respect to θ and λ . These estimators can also be obtained by solving the following non-linear equations:

$$\sum_{i=1}^n \left(F(x_{(i)} | \theta, \lambda) - \frac{2i-1}{2n} \right) \eta_1(x_{(i)} | \theta, \lambda) = 0,$$

$$\sum_{i=1}^n \left(F(x_{(i)} | \theta, \lambda) - \frac{2i-1}{2n} \right) \eta_2(x_{(i)} | \theta, \lambda) = 0,$$

where $\eta_1(\cdot | \theta, \lambda)$ and $\eta_2(\cdot | \theta, \lambda)$ are given by (9) and (10), respectively.

3.6.2 Methods of Anderson-Darling and Right-tail Anderson-Darling

The Anderson-Darling estimator is another type of minimum distance estimator and is based on an Anderson-Darling statistic (Anderson & Darling, [2, 3]). The Anderson-Darling statistic is given by

$$ADS_n^2 = n \int_{-\infty}^{\infty} \frac{(F(x_{(i)}) - E_n(x_{(i)}))^2}{F(x)(1 - F(x))} dF(x_{(i)})$$

Boos [7] also discussed the properties of the AD estimators and presented its computational form which is given by

$$A(\theta, \lambda) = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1) \tag{3.17}$$

$$\{ \log F(x_{(i)} | \theta, \lambda) + \log \bar{F}(x_{(n+1-i)} | \theta, \lambda) \}.$$

Therefore, the Anderson-Darling estimators $\hat{\theta}_{ADE}$ and $\hat{\lambda}_{ADE}$ of the parameters θ and λ are obtained by minimizing (3.17) with respect to θ and λ . Analogously, these estimators can also be obtained by solving the following non-linear equations:

$$\sum_{i=1}^n (2i - 1) \left[\frac{\eta_1(x_{(i)} | \theta, \lambda)}{F(x_{(i)} | \theta, \lambda)} - \frac{\eta_1(x_{(n+1-i)} | \theta, \lambda)}{S(x_{(n+1-i)} | \theta, \lambda)} \right] = 0,$$

$$\sum_{i=1}^n (2i - 1) \left[\frac{\eta_2(x_{(i)} | \theta, \lambda)}{F(x_{(i)} | \theta, \lambda)} - \frac{\eta_2(x_{(n+1-i)} | \theta, \lambda)}{S(x_{(n+1-i)} | \theta, \lambda)} \right] = 0,$$

where $\eta_1(\cdot | \theta, \lambda)$ and $\eta_2(\cdot | \theta, \lambda)$ are given by (9) and (10), respectively.

Further, Luceño [18] discussed modifications of the standard AD statistics. The most used statistic [11, 17, 21] is the Right-tail AD statistics given by

$$RADS_n^2 = n \int_{-\infty}^{\infty} \frac{(F(x_{(i)}) - E_n(x_{(i)}))^2}{1 - F(x_{(i)})} dF(x_{(i)}).$$

Additionally, its computational form was presented in the form of

$$R(\theta, \lambda) = \frac{n}{2} - 2 \sum_{i=1}^n F(x_{(i)} | \theta, \lambda) - \frac{1}{n} \sum_{i=1}^n (2i - 1) \log S(x_{n+1-i:n} | \theta, \lambda). \quad (3.18)$$

Hence, the Right-tail Anderson-Darling estimators $\hat{\theta}_{RTADE}$ and $\hat{\lambda}_{RTADE}$ of the parameters θ and λ are obtained by minimizing (3.18) with respect to θ and λ . These estimators can also be obtained by solving the following non-linear equations:

$$\begin{aligned} -2 \sum_{i=1}^n \frac{\eta_1(x_{(i)} | \theta, \lambda)}{F(x_{(i)} | \theta, \lambda)} + \frac{1}{n} \sum_{i=1}^n (2i - 1) \frac{\eta_1(x_{n+1-i:n} | \theta, \lambda)}{S(x_{n+1-i:n} | \theta, \lambda)} &= 0, \\ -2 \sum_{i=1}^n \frac{\eta_2(x_{(i)} | \theta, \lambda)}{F(x_{(i)} | \lambda, \sigma)} + \frac{1}{n} \sum_{i=1}^n (2i - 1) \frac{\eta_2(x_{n+1-i:n} | \theta, \lambda)}{S(x_{n+1-i:n} | \theta, \lambda)} &= 0, \end{aligned}$$

where $\eta_1(\cdot | \theta, \lambda)$ and $\eta_2(\cdot | \theta, \lambda)$ are given by (9) and (10), respectively.

4 SIMULATION STUDY

In this section, we conduct Monte Carlo simulation studies to compare the performance of the frequentist estimators discussed in the previous sections. Using the mixture representation described in Section 2, the values of the BE2 distribution were generated using the following algorithm:

1. Generate $U_i \sim \text{Uniform}(0, 1)$, $i = 1, \dots, n$;
2. Generate $X_i \sim \text{Gamma}(2, \lambda)$, $i = 1, \dots, n$;
3. Generate $Y_i \sim \text{Exponential}(\lambda)$, $i = 1, \dots, n$;
4. If $U_i \leq p = \theta/(2 - \theta)$, then set $T_i = X_i$, otherwise, set $T_i = Y_i$, $i = 1, \dots, n$.

We evaluate the performance of the estimators of the BE2 distribution based on bias and MSE. The following procedure is adopted to evaluate the performance of the estimators:

1. Generate pseudo random sample with size n of the $BE2(\theta, \lambda)$
2. Using the values obtained in step 1, calculate $\hat{\theta}$ and $\hat{\lambda}$ via MLE, ME, MME, LSE, WLSE, PCE, MPS, CME, ADE, RTADE.
3. Repeat the steps 2 and 3 N times.
4. Using $\hat{\Theta} = (\hat{\theta}, \hat{\lambda})$ and $\Theta = (\theta, \lambda)$, compute the Bias $\frac{1}{N} \sum_{i=1}^N (\hat{\Theta}_{i,j} - \Theta_{i,j})$ and the mean square errors (MSE) $\sum_{i=1}^N \frac{(\hat{\Theta}_{i,j} - \Theta_{i,j})^2}{N}$ for $j = 1, 2$.

It is expected that for this approach the Bias and the MSE are closer to zero. The results are computed using the software R (R Core Development Team). The seed used to generate the random values is 2017. The chosen values to perform this procedure are $\Theta = (1, 0.8)$, $N = 500,000$ and $n = (15, 20, 25, \dots, 130)$.

The estimation methods are put under the same conditions (initial values and random samples). The initial values used to initiate the iterative methods are the true values. The estimates are obtained by applying the maxBFGS function in (3.2), (3.13), (3.17) and (3.18). This function is available in the maxlik package [14]. The nls function available in the stats package and are used in (3.7), (3.8), (3.11) and (3.16). The programs can be obtained, upon request.

In order to present a fair comparison, the estimation procedures are performed under the same conditions. However, for some samples and estimation procedures, the numerical techniques fail in finding the parameters estimates. Hence, a numerical study is conducted to verify the frequency of convergence for each estimation method by counting the number of times that each estimation fails in finding the numerical solution. In Figure 1, we present the proportion of failure of each estimation.

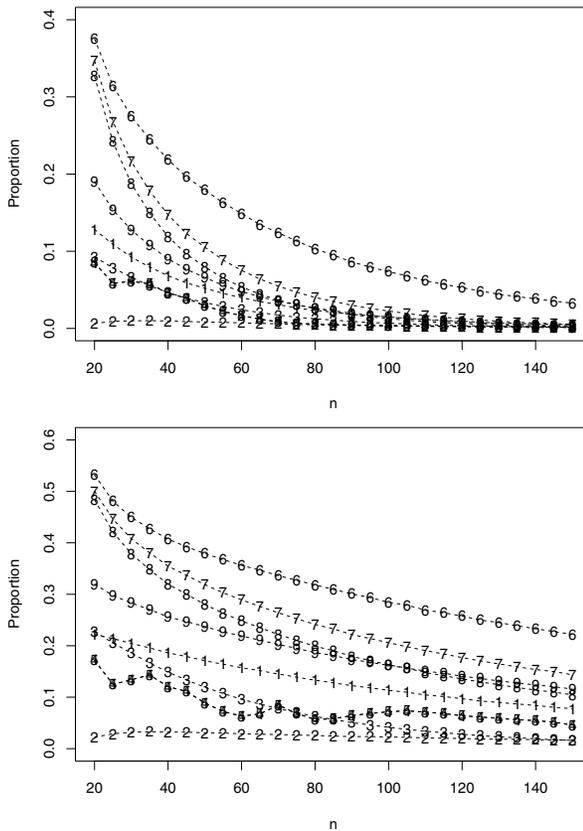


Figure 1: Rate of convergence considering different values of Θ and different estimation procedures: 1-ME, 2-MLE, 3-MPS, 4-ADE, 5-RADE, 6-PCE, 7-LSE, 8-WLSE, 9-CME.

From Figure 1, the PCE, LSE, WLSE and CME estimators show a high proportion of failure in the numerical procedures. As the above estimators result in rate of failures, therefore we discarded these estimators from our simulation study. Figure 2 shows the Bias, MSEs for the estimates of λ and θ obtained by using different estimation methods for 500,000 simulated samples and considering different values of n . We have presented results only for $\lambda = 1$ and $\theta = 0.8$ due to space constraint. But the results are similar for other choices for λ and θ .

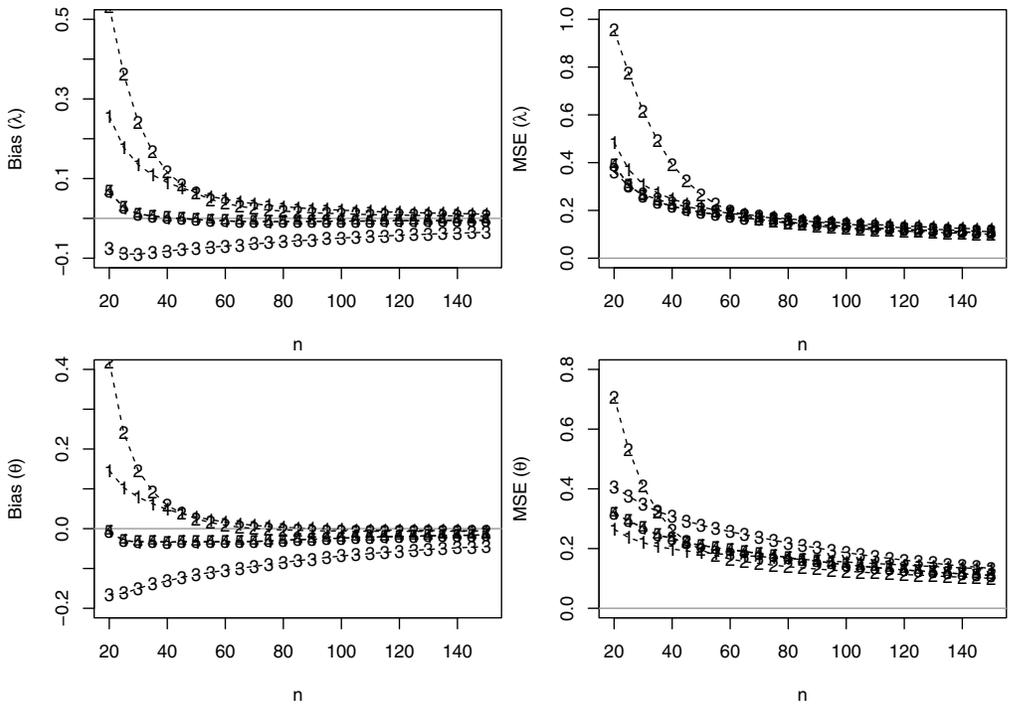


Figure 2: Bias, MSEs related from the estimates of λ and θ for N simulated samples, considering different values of n obtained using the following estimation method 1-ME, 2-MLE, 3-MPS, 4-ADE, 5-RADE.

From the graphs we can see that the Bias and the MSE of all estimators tend to zero for large n , i.e., the estimates are asymptotically unbiased for the parameters. The MPS estimator shown to be superior than the MLE, this result are consistent with other studies [21, 24]. It is worth mentioning that, although the ME has closed form expressions, we may have $2 \left(\frac{\sigma}{\mu}\right)^2 > 2$, i.e., θ is a complex number. The ME and the MLE shows positive bias for both the parameters, while the MPS shows a negative bias for the parameters. Overall, the ADE and RTADE provides superior estimates than their counter parts in terms of Bias and MSEs. Although, both ADE and RTADE shows almost the same results, the ADE has desirable properties such as robustness, consistence and normally asymptotic properties [7,8] and may be used for estimating the BE2 distribution parameters. Rodrigues et al. [23] observed a similar result for the Poisson-exponential distribution where the ADE was the best method of estimation.

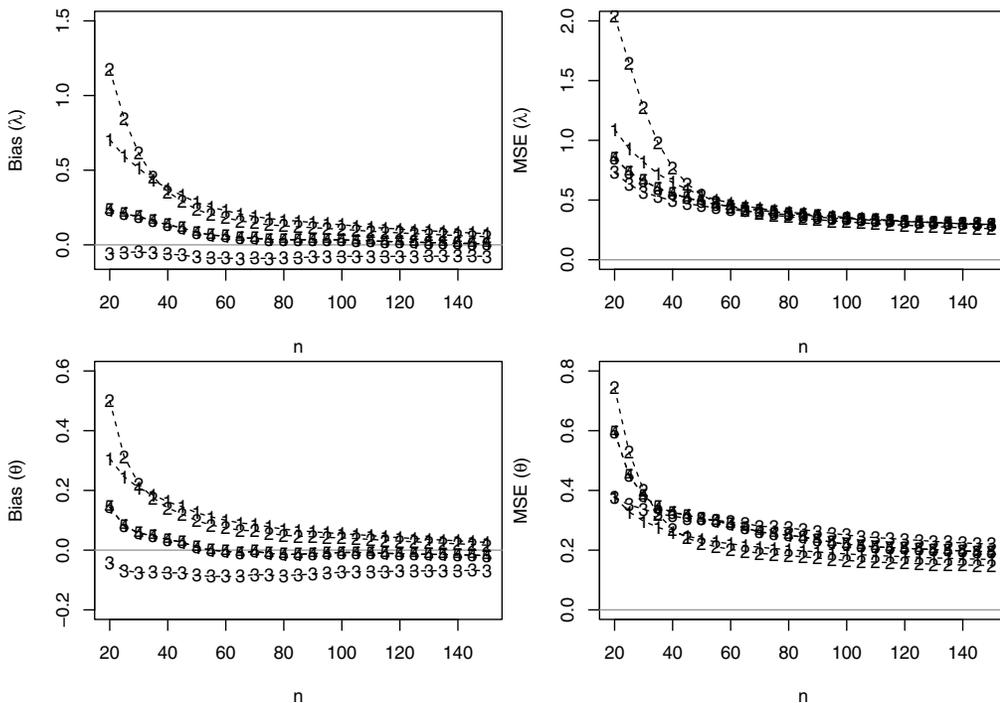


Figure 3: Bias, MSEs related from the estimates of λ and θ for N simulated samples, considering different values of n obtained using the following estimation method 1-ME, 2-MLE, 3-MPS, 4-ADE, 5-RADE.

5 APPLICATIONS

Located in southeastern Brazil, São Carlos is a city of 238,958 inhabitants. The city has an active industrial profile and high agricultural importance. Therefore, the study of the behaviour of dry and wet periods has proved to be strategic and economically significant for the regional development. From Figure 4, we observe that the city has rainy periods from October to March, and from June to August exhibit more dry periods.

Consequently, prediction of the behavior of the transition periods in rainy sessions (April, May and September) enables the agriculturists to be prepared against different problems, such as water scarcity. In this paper, we consider three real data sets related to the total monthly rainfall during April, May and September at São Carlos. The data sets (see the Appendix for more details) was obtained from the Department of Water Resources and Power agency manager of water resources of the State of São Paulo including a period from 1960 to 2014.

5.1 Initial Values

Finding good initial values to start iterative procedures is an important problem in numerical analysis, while in the Section 4 we have used the true values to start the iterative procedure. In

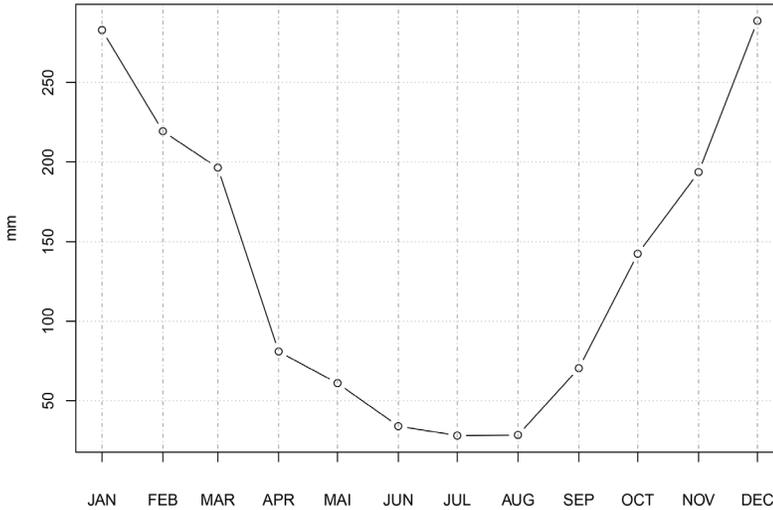


Figure 4: Average of the total monthly rainfall from January to December at São Carlos, Brazil.

any application these values are unknown. A good choice as initial values would be to use the MEs (3.6), since these estimators have closed-form expressions. However, we have shown in Section 4 that the MEs may not be computed in some cases.

Two problems can arise during the application of the MEs. First, we may have $2 \left(\frac{s}{\bar{x}}\right)^2 > 2$, i.e., θ will be a complex number. Second $2 \left(\frac{s}{\bar{x}}\right)^2 < 1$, i.e., θ to be greater than 1. The first problem can be overcome by taking the absolute value among $\left(2 - 2 \left(\frac{s}{\bar{x}}\right)^2\right)$, while the second can be overcome by taking the minimum value between $\sqrt{\left|2 - 2 \left(\frac{s}{\bar{x}}\right)^2\right|}$ and 1. Therefore, we have chosen the modified estimator and is given by

$$\tilde{\theta} = \min \left(\sqrt{\left|2 - 2 \left(\frac{s}{\bar{x}}\right)^2\right|}, 1 \right), \quad \text{and} \quad \tilde{\lambda} = \frac{2}{\tilde{\theta}\bar{x}} \tag{5.1}$$

where $|x|$ is the absolute value of x . In this case λ and θ can be computed without any problem. Note that, here we suggest the use of this estimator as initial value to be used in the iterative methods.

5.2 Discrimination Criterion Methods

Here, different discrimination criterion are considered based on log likelihood function. Let k be the number of parameters to be fitted and $\hat{\Theta}$ the MLEs of Θ , the discrimination criterion methods are respectively: Akaike information criterion (AIC) computed through $AIC = -2l(\hat{\Theta}; \mathbf{x}) + 2k$, Corrected Akaike information criterion $AICC = AIC + (2k(k+1))/(n-k-1)$, Hannan-Quinn information criterion $HQIC = -2l(\hat{\Theta}; \mathbf{x}) + 2k \log(\log(n))$ and the consistent Akaike information criterion $CAIC = -2l(\hat{\Theta}; \mathbf{x}) + k(\log(n) + 1)$. The best model is the one which provides

the minimum values of these criteria. The Kolmogorov-Smirnov (KS) test is also considered in order to check the goodness of the fit for the models. This procedure is based on the KS statistic $D_n = \sup_x |F_n(x) - F(x; \theta, \lambda)|$, where \sup_x is the supremum of the set of distances, $F_n(x)$ is the empirical distribution function and $F(x; \theta, \lambda)$ is c.d.f. Under a significance level of 5% if the data comes from $F(x; \theta, \lambda)$ (null hypothesis), the hypothesis is rejected if the p-value is smaller than 0.05.

For the sake of comparison, the results obtained from the BE2 distribution are compared with the Weibull, Gamma, Lognormal, Gumbel and Generalized Exponential [13] distributions and nonparametric survival function.

5.3 Results

The data sets related to May and September have the occurrence of zero values, i.e., non occurrence of precipitation. This type of data does not allow us to fit popular distributions such as Gamma, Weibull, Lognormal and Generalized Exponential distribution since they are defined only for $x > 0$. To overcome this problem, we approximate 0.0 in the data set to 0.1. Although this is not a standard procedure, yet, without changing this results we will not be able to fit these common distributions. Nadarajah & Haghighi [20] observed that maximum likelihood estimate of the shape parameter is non-unique for the Gamma, Weibull and Generalized exponential distributions if data set consists of zeros and therefore none of these three distributions can fit this kind of data set. On the other hand the BE2 distribution is defined as $x \geq 0$, which allow us to use the original values in the presence of zero.

Table 1 presents the results for AIC, AICC, HQIC and CAIC criteria, for different probability distributions. In the Figure 5, we have the survival function adjusted by different distributions and non-parametric survival estimator.

Comparing the empirical survival function with the adjusted distributions it is observed that the BE2 distribution fits best among the chosen models. This result is confirmed from AIC, AICC, HQIC and CAIC criteria as the BE2 distribution has the minimum values. Considered the parametric bootstrap confidence intervals [12] in order to build the confidence intervals for the parameters of BE2 distribution using the Anderson-Darling estimates.

Table 2 displays the MLEs and 95% confidence intervals for θ and λ of the BE2 distribution.

The quantile-quantile (Q-Q) plot is a graphical technique which provides an assessment of goodness of fit. If the data set comes from the proposed distribution, the points should fall approximately along the 45-degree reference line. Figures 4 and 5 display the histogram and the Q-Q plot from the proposed data set.

From the Figure 4, we observe that the points are approximately along the reference line. Therefore, the proposed methodology suggests that the data related to the total monthly rainfall during April, May and September at São Carlos demonstrate the Binomial-exponential 2 distribution.

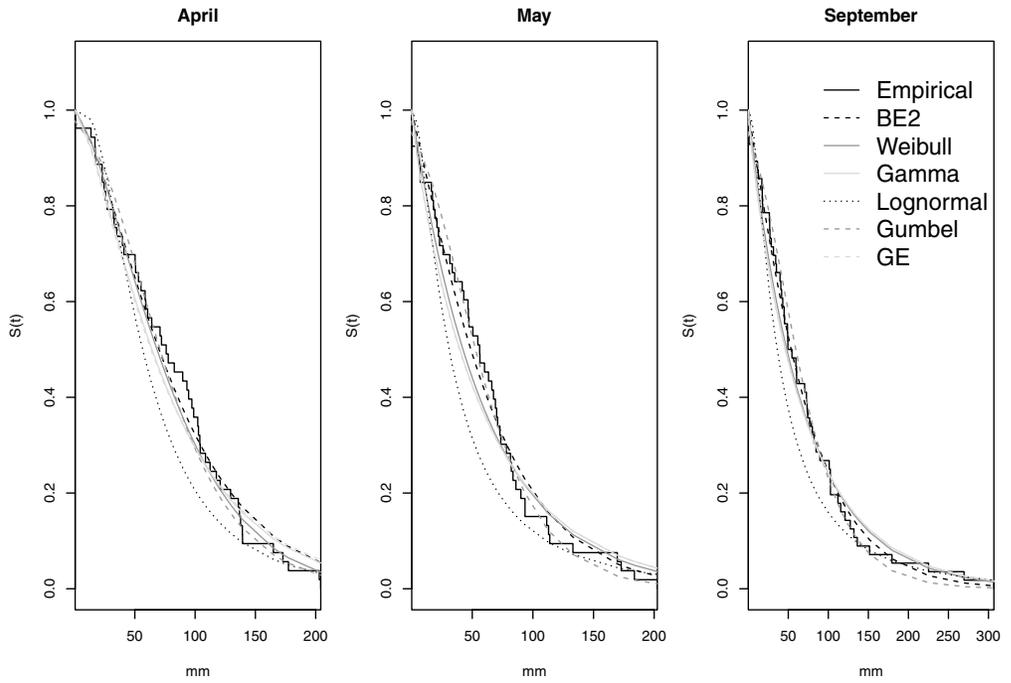


Figure 5: Survival function adjusted by different distributions and a non-parametric method considering the data sets related to the total monthly rainfall during April, May and September at São Carlos.

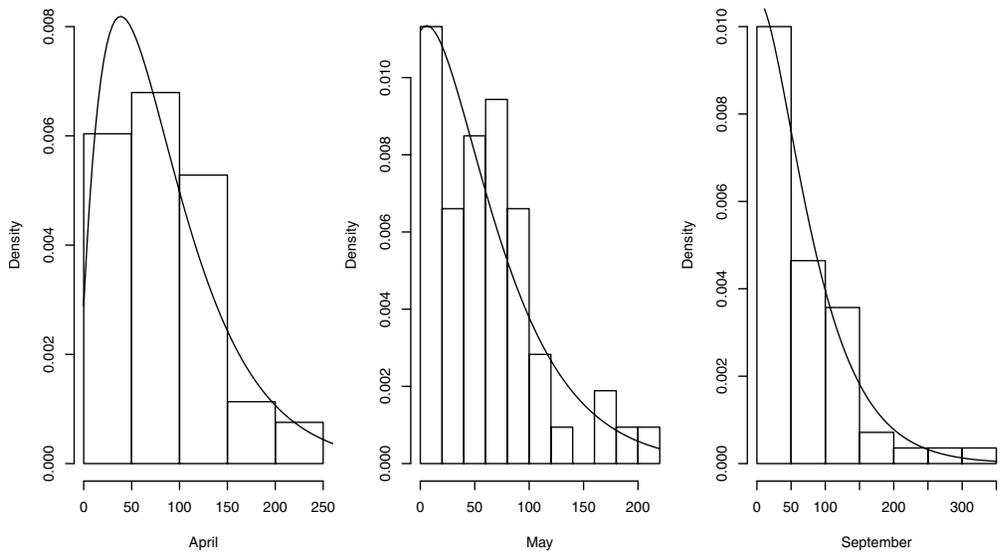


Figure 6: Histogram from the data sets related to the total monthly rainfall.

Table 1: Results of the AIC, AICC, HQIC and CAIC criteria and the p-values of KS statistic for different probability distributions considering the data sets related to the total monthly rainfall during April, May and September at São Carlos.

Month	Test	BE2	Weibull	Gamma	Lognormal	Gumbel	GE
April	AIC	566.697	566.814	570.371	595.503	568.860	571.125
	AICC	566.937	567.054	570.611	595.743	569.100	571.365
	HQIC	568.212	568.329	571.886	597.018	570.375	572.640
	CAIC	572.637	572.755	576.312	601.443	574.801	577.066
	KS*	0.8919	0.6810	0.5284	0.0246	0.7456	0.5148
May	AIC	543.480	545.769	544.980	580.391	553.281	544.751
	AICC	543.720	546.009	545.220	580.631	553.521	544.991
	HQIC	544.990	547.284	546.495	581.906	554.796	546.266
	CAIC	549.420	551.710	550.921	586.332	559.222	550.692
	KS*	0.8446	0.2734	0.1457	0.0013	0.0223	0.1348
September	AIC	591.595	592.515	591.911	625.321	604.370	591.780
	AICC	591.821	592.741	592.137	625.547	604.596	592.006
	HQIC	593.165	594.085	593.481	626.891	605.940	593.350
	CAIC	597.645	598.566	597.962	631.372	610.421	597.831
	KS*	0.9858	0.5679	0.34056	0.0256	0.0001	0.3205

*p-values of KS statistic.

Table 2: MLE, 95% confidence intervals for θ and λ considering the data sets related to the total monthly rainfall during April, May and September at São Carlos.

Month	Θ	ADE	CI _{95%} (Θ)
April	θ	0.9297	(0.611; 0.996)
	λ	0.0220	(0.016; 0.028)
May	θ	0.7097	(0.200; 0.962)
	λ	0.0254	(0.015; 0.030)
September	θ	0.6375	(0.234; 0.955)
	λ	0.0208	(0.017; 0.033)

6 CONCLUSIONS

In this paper, the model parameters of Binomial-exponential 2 are estimated by nine methods of estimation, namely, maximum likelihood, moments, percentile, least squares, weighted least squares, maximum product of spacing, Cramer-von Mises, Anderson-Darling and right tailed Anderson-Darling. As it is not feasible to compare these methods of estimation theoretically,

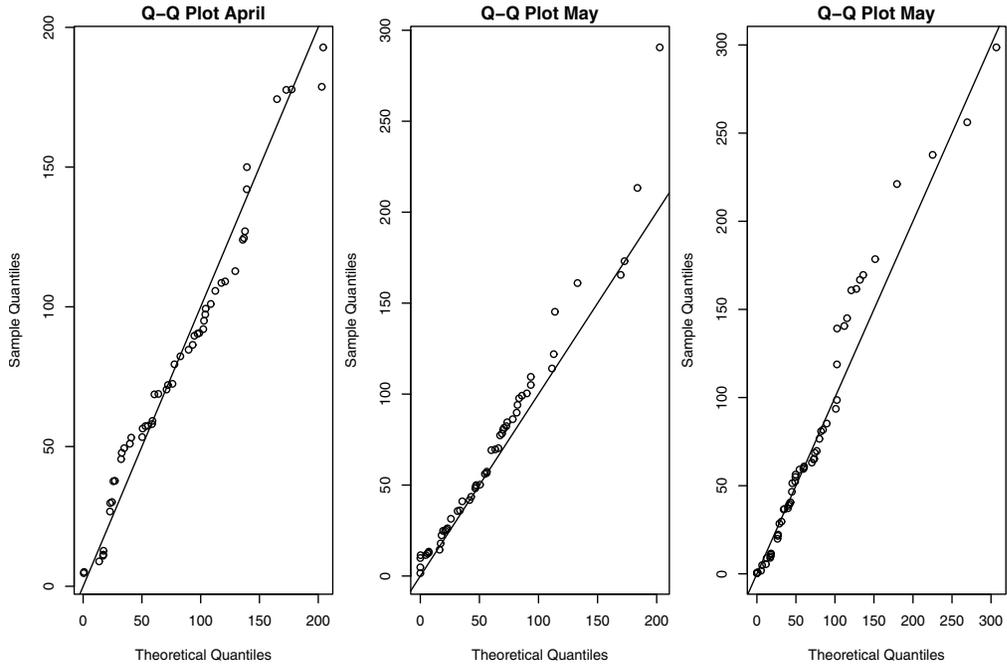


Figure 7: Q-Q plot from the data sets related to the total monthly rainfall.

we have presented the simulation study results in order to identify the most efficient procedure. The simulation results show that the Anderson-Darling estimators outperform other procedures such as the maximum likelihood method for estimating the parameters of the BE2 distribution. The proposed methodology is applied in three real data sets related to the total monthly rainfall during April, May and September at São Carlos, Brazil, demonstrating that the BE2 distribution can be used as alternative to some well known distributions in weather related data.

ACKNOWLEDGEMENTS

The authors are very grateful to the Editor and the reviewers for their helpful and useful comments that improved the manuscript.

7 APPENDIX A – DATA SET

- April: 59.00, 102.20, 17.30, 23.00, 50.60, 27.00, 203.00, 40.90, 53.00, 177.40, 94.60, 129.40, 76.00, 93.20, 22.80, 98.80, 77.70, 204.20, 16.90, 55.10, 103.90, 34.90, 39.70, 137.70, 104.20, 117.60, 17.10, 120.80, 164.90, 50.20, 172.80, 58.50, 112.40, 24.50, 32.80, 64.00, 72.10, 139.30, 0.50, 70.90, 0.80, 82.70, 108.60, 32.30, 13.60, 25.70, 135.80, 136.80, 89.70, 139.20, 102.80, 97.30, 60.60.
- May: 63.40, 41.70, 0.00, 0.00, 47.30, 31.50, 172.80, 93.50, 0.00, 60.10, 23.00, 90.10, 50.50, 67.50, 4.70, 7.10, 93.50, 0.20, 82.20, 112.90, 7.10, 35.50, 81.50, 202.60, 56.10,

19.20, 69.10, 133.00, 111.40, 25.90, 33.50, 46.80, 54.60, 43.00, 46.50, 83.60, 73.50, 18.00, 16.30, 70.00, 56.30, 70.90, 183.70, 78.20, 6.20, 86.00, 66.10, 72.80, 20.90, 17.20, 113.90, 169.60, 22.10.

- September: 26.40, 12.50, 1.00, 44.80, 0.00, 74.20, 179.50, 76.70, 269.50, 49.00, 306.80, 102.70, 73.50, 35.20, 72.70, 28.80, 49.30, 132.00, 151.50, 39.70, 136.20, 112.00, 17.70, 11.60, 225.20, 102.60, 27.10, 17.50, 6.70, 82.20, 40.70, 54.60, 115.50, 89.50, 0.00, 17.00, 127.40, 41.70, 43.10, 84.70, 102.50, 120.90, 80.10, 18.10, 5.30, 59.50, 26.80, 0.00, 34.30, 101.10, 60.30, 31.50, 60.40, 45.30, 49.50, 70.44.

RESUMO. Neste trabalho, apresentamos diferentes métodos de estimação para os parâmetros da distribuição binomial-exponencial 2, tais como, estimador de máxima verossimilhança, método dos momentos, método percentil, estimador de mínimos quadrados, método do máximo produto espaçado, estimador de Cramér-von-Mises, estimador de Anderson-Darling e o estimador de Anderson-Darling com cauda a direita são apresentados. Com base em um estudo de simulação numérica, verificamos que o estimador de Anderson-Darling retorna estimativas mais eficientes se comparado com os outros estimadores. Por fim, nossa proposta é aplicada em três conjuntos de dados relacionados à pluviosidade total mensal ao longo dos meses de Abril, Maio e Setembro em São Carlos, Brasil.

Palavras-chave: distribuição exponencial-binomial 2, estimador de máxima verossimilhança, estimador de Cramér-von-Mises, estimador de Anderson-Darling.

REFERENCES

- [1] M.R. Alkawasbeh & M.Z. Raqab. Estimation of the generalized logistic distribution parameters: Comparative study. *Statistical Methodology*, **6** (2009), 262–279.
- [2] T.W. Anderson & D.A. Darling. Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes. *The Annals of Mathematical Statistics*, **23**(2) (1952), 193–212.
- [3] T.W. Anderson & D.A. Darling. A test of goodness of fit. *Journal of the American Statistical Association*, **49** (1954), 765–769.
- [4] A. Asgharzadeh, H.S. Bakouch & M. Habibi. A generalized binomial exponential 2 distribution: modeling and applications to hydrologic events. *Journal of Applied Statistics*, (2016), 1–20. Available online: <http://dx.doi.org/10.1080/02664763.2016.1254729>
- [5] A. Asgharzadeh, R. Rezaie & M. Abdi. Comparisons of methods of estimation for the half-logistic distribution. *Selcuk Journal of Applied Mathematics*, Special Issue (2011), 93–108.
- [6] H.S. Bakouch, M.A. Jazi, S. Nadarajah, A. Dolati & R. Roozgar. A lifetime model with increasing failure rate. *Applied Mathematical Modelling*, **38** (2014), 5392–5406.
- [7] D.D. Boos. Minimum distance estimators for location and goodness of fit. *Journal of the American Statistical Association*, **76** (1981), 663–670.

- [8] D.D. Boos. Minimum Anderson-Darling estimation. *Communications in Statistics-Theory and Methods*, **11** (1982), 2747–2774.
- [9] R. Cheng & N. Amin. Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society, Series B (Methodological)*, **45** (1983), 394–403.
- [10] R. Cheng & M. Stephens. A goodness-of-fit test using Moran’s statistic with estimated parameters. *Biometrika*, (1989), 385–392.
- [11] S. Dey, T. Dey & D. Kundu. Two-parameter Rayleigh distribution: different methods of estimation. *American Journal of Mathematical and Management Sciences*, **33** (2014), 55–74.
- [12] B. Efron & R.J. Tibshirani. An introduction to the bootstrap. CRC press, (1994), 1st Edition, 436 p.
- [13] R.D. Gupta & D. Kundu. Generalized exponential distribution: Different method of estimations. *Journal of Statistical Computation and Simulation*, **69** (2001), 315–337.
- [14] A. Henningsen & O. Toomet. maxlik: A package for maximum likelihood estimation in R. *Computational Statistics*, **26** (2011), 443–458. <http://dx.doi.org/10.1007/s00180-010-0217-1>.
- [15] J.H. Kao. Computer methods for estimating Weibull parameters in reliability studies. *IRE Transactions on Reliability and Quality Control*, **13** (1958), 15–22.
- [16] J.H. Kao. A graphical estimation of mixed Weibull parameters in life-testing of electron tubes. *Techonometrics*, **1** (1959), 389–407.
- [17] F. Louzada, P.L. Ramos & G.S. Perdoná. Different estimation procedures for the parameters of the extended exponential geometric distribution for medical data. *Computational and Mathematical Methods in Medicine*, **2016** (2016). Article ID 8727951, 12 pages. doi:10.1155/2016/8727951.
- [18] A. Luceño. Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators. *Computational Statistics & Data Analysis*, **51** (2006), 904–917.
- [19] J. Mazucheli, F. Louzada & M. Ghitany. Comparison of estimation methods for the parameters of the weighted Lindley distribution. *Applied Mathematics and Computation*, **220** (2013), 463–471.
- [20] S. Nadarajah & F. Haghghi. An extension of the exponential distribution. *Statistics*, **45** (2011), 543–558.
- [21] P. Ramos & F. Louzada. The generalized weighted Lindley distribution: Properties, estimation and applications. *Cogent Mathematics*, **3** (2016), 1–18.
- [22] B. Ranneby. The maximum spacing method. An estimation method related to the maximum likelihood method. *Scandinavian Journal of Statistics*, **11** (1984), 93–112.
- [23] G.C. Rodrigues, F. Louzada & P.L. Ramos. Poisson-exponential distribution: different methods of estimation. *Journal of Applied Statistics*, (2016), pp. 1–17. Available online: <http://dx.doi.org/10.1080/02664763.2016.1268571>
- [24] V.K. Sharma, S.K. Singh, U. Singh & F. Merovci. The generalized inverse Lindley distribution: A new inverse statistical model for the study of upside-down bathtub data. *Communications in Statistics-Theory and Methods*, **45** (2016), 5709–5729.
- [25] J.J. Swain, S. Venkatraman & J.R. Wilson. Least-squares estimation of distribution functions in Johnson’s translation system. *Journal of Statistical Computation and Simulation*, **29** (1988), 271–297.
- [26] M. Teimouri, S.M. Hoseini & S. Nadarajah. Comparison of estimation methods for the Weibull distribution. *Statistics*, **47** (2013), 93–109.