Submetido para TEMA

Binomial-exponential 2 distribution: Different estimation methods and weather applications

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 frequentist approaches such as the method of moments, modified moments, ordinary least-squares estimation, weighted least-squares estimation, percentile, maximum product of spacings, Cramer-von Mises type minimum distance, Anderson-Darling and Right-tail Anderson-Darling, and compare them using extensive numerical simulations. We apply our proposed methodology to three real data sets related to the total monthly rainfall during April, May and September at Sao 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. (2014) 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.

Estimating the parameters of the BE2 distribution was discussed by Bakouch et al. (2014) considering only the maximum likelihood estimation (MLE) method. However it is of interest to compare the MLE method with other estimation procedures such as the method of moments, modified moments, ordinary least-squares estimation (OLSE), weighted least-squares estimation (WLSE), percentile (PCE), maximum product of spacings (MPS), Cramér-von-Mises type minimum distance (CME), Anderson-Darling (ADE) and Right-tail Anderson-Darling (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 another 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., Gupta and Kundu (2001), Kundu and Raqab (2005), Alkasasbeh and Raqab (2009), Mazucheli et al. (2013), Teimouri et al.(2013) and Dey et al.(2014).

The main goal of this paper is twofold: First is to show how different frequentist estimators of proposed distribution perform for different sample sizes and second is to show that the distribution outperforms at least as well as 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, its common in dry periods the non occurrence of precipitation, therefore the occurrence of zero value can be observed in different measures such as in the average, maximum and minimum. Popular models such as Gamma, Weibull, Lognormal and Generalized Exponential distribution not accommodate such characteristic. In this paper we demonstrated 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 ten 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 Sao Carlos, Brazil. Some final comments are presented in Section 6.

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 pdf 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. $f(x;p,\lambda) = p\ \lambda^2 x e^{-\lambda x} + (1-p)\ \lambda e^{-\lambda x}$,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

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

Its important to point out, that a simple extension can be obtained for (1.1), considering that x take value on 0, in these case,

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

where $f_X(0; \theta, \lambda) \geq 0$ for all $0 \leq \theta \leq 1$ and $\lambda > 0$. This result allows that the BE2 distribution become a simple alternative to be used in problems with occurrence of zero value.

3. Methods of estimation

In this section, we describe the ten considered estimation methods to obtain the estimators of the parameters θ and λ of the BE2 distribution.

3.1. Maximum Likelihood Estimation

The method of maximum likelihood is the most frequently used method of parameter estimation (Casella and Berger, 2002). The method's success stems no doubt from its many desirable properties including consistency, asymptotic efficiency, normality, invariance and simply its intuitive appeal. Let $x_1, \cdots, x_n$ 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

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

For ease of notation, we will denote the first partial derivatives of any function $f(x, y)$ by f_x and f_y. Now setting $\ell_\theta = 0$ and $\ell_\lambda = 0$, we have

$$\ell_\theta = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \frac{\theta x_i}{(2 - 2\theta + \lambda\theta x_i)} = 0 \tag{3.2}$$

and

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

The maximum likelihood estimator $\hat{\theta}$ and $\hat{\lambda}$ are obtained solving the non-linear equations (3.2) and (3.3). Its 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 first two theoretical moments of (1.1) with the sample moments $\frac{1}{n}\sum_{i=1}^{n} x_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i^2$ respectively,

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{2}{\lambda(2-\theta)} \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{2}{\lambda^2}\left(1 + \frac{2\theta}{2-\theta}\right). \tag{3.4}$$

After some algebraic manipulation the estimator $\hat{\theta}_{ME}$ for θ and $\hat{\lambda}_{ME}$ for λ, can be obtained by solving

$$\hat{\theta}_{ME} = \sqrt{4 - \frac{2}{\bar{x}^2 n}\sum_{i=1}^{n} x_i^2} \quad \text{and} \quad \hat{\lambda}_{ME} = \frac{2}{\bar{x}\left(2 - \hat{\theta}_{ME}\right)}. \tag{3.5}$$

3.3. Method of Modified Moments Estimators

A simple modification can be made in the method of moments for estimating the parameters of the BE2 distribution. Instead of equating the first two theoretical moments consider that

$$E(X|\theta,\lambda) = \frac{2}{\lambda(2-\theta)} \quad \text{and} \quad Var(X|\theta,\lambda) = \frac{2(2-\theta^2)}{\lambda^2(2-\theta)^2}. \tag{3.6}$$

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

$$\hat{\theta}_{MME} = \sqrt{2 - 2\left(\frac{s}{\bar{x}}\right)^2}, \quad \text{and} \quad \hat{\lambda}_{MME} = \frac{2}{\bar{x}\left(2 - \sqrt{2 - 2\left(\frac{s}{\bar{x}}\right)^2}\right)} \tag{3.7}$$

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

3.4. Least-Square Estimators

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

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

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

$$\begin{aligned} \sum_{i=1}^{n} \left[F(x_{i:n} \mid \theta, \lambda) - \frac{i}{n+1} \right] \eta_1(x_{i:n} \mid \theta, \lambda) = 0, \\ \sum_{i=1}^{n} \left[F(x_{i:n} \mid \theta, \lambda) - \frac{i}{n+1} \right] \eta_2(x_{i:n} \mid \theta, \lambda) = 0. \end{aligned}$$

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

$$W(\theta, \lambda) = \sum_{i=1}^{n} \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(x_{i:n} \mid \theta, \lambda) - \frac{i}{n+1} \right]^2.$$

These estimators can also be obtained by solving:

$$\begin{aligned} \sum_{i=1}^{n} \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(x_{i:n} \mid \theta, \lambda) - \frac{i}{n+1} \right] \eta_1(x_{i:n} \mid \theta, \lambda) = 0, \\ \sum_{i=1}^{n} \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(x_{i:n} \mid \theta, \lambda) - \frac{i}{n+1} \right] \eta_2(x_{i:n} \mid \theta, \lambda) = 0, \end{aligned}$$

where

$$\eta_1(x_{i:n} \mid \theta, \lambda) = \frac{-2\lambda x_{i:n} e^{-\lambda x_{i:n}}}{(2-\theta)^2}, \tag{3.8}$$

and

$$\eta_2(x_{i:n} \mid \theta, \lambda) = x_{i:n} e^{-\lambda x_{i:n}} (1 + \frac{\lambda \theta x_{i:n}}{2-\theta}) - \frac{\theta x_{i:n} e^{-\lambda x_{i:n}}}{2-\theta}. \tag{3.9}$$

3.5. Percentile Estimators

The percentile estimators is a method of statistical inference originally suggested by Kao (1958, 1959). This method is commonly used to estimate the unknown parameters from the distribution functions that has a closed form for the quantile function. The percentile estimates (PCEŠs) are 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

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

Let $X_{(j)}$ be the jth order statistic, i.e., $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$. If p_j denotes some estimators of $F(x_{(j)};\theta,\lambda)$, then the estimators of θ and λ can be obtained by minimizing $\sum_{j=1}^{n}\left(x_{(j)} - \frac{1}{\lambda}\ln\frac{2-\theta+\lambda\theta x_p}{(2-\theta)(1-p_j)}\right)^2$ 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

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

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

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

3.6. Method of Maximum Product of Spacings

The maximum product spacing (MPS) method has been introduced by Cheng and Amin (1979, 1983) as an alternative to MLE for the estimation of the unknown parameters parameters of continuous univariate distributions. The MPS method was also derived independently by Ranneby (1984) as an approximation to the Kullback-Leibler measure of information. To motivate our choice, Cheng and Amin (1983) proved that this method is as efficient as the MLE estimators and consistent under more general conditions.

Using the same notations in subsection 3.4., define the uniform spacings of a random sample from the BE2 distribution as:

$$D_i(\theta,\lambda) = F\left(x_{i:n} \mid \theta,\lambda\right) - F\left(x_{i-1:n} \mid \theta,\lambda\right), \qquad i = 1,2,\ldots,n,$$

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

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

$$G(\theta,\lambda) = \left[\prod_{i=1}^{n+1} D_i(\theta,\lambda)\right]^{\frac{1}{n+1}}, \tag{3.10}$$

or, equivalently, by maximizing the function

$$H(\theta, \lambda) = \frac{1}{n+1} \sum_{i=1}^{n+1} \log D_i(\theta, \lambda). \tag{3.11}$$

The estimators $\widehat{\theta}_{MPS}$ and $\widehat{\lambda}_{MPS}$ of the parameters θ and λ can be obtained by solving the nonlinear equations

$$\frac{\partial H(\theta, \lambda)}{\partial \theta} = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\theta, \lambda)} \left[\eta_1(x_{i:n}|\theta, \lambda) - \eta_1(x_{i-1:n}|\theta, \lambda)\right] = 0, \tag{3.12}$$

$$\frac{\partial}{\partial \lambda} H(\theta, \lambda) = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\theta, \lambda)} \left[\eta_2(x_{i:n}|\theta, \lambda) - \eta_2(x_{i-1:n}|\theta, \lambda)\right] = 0, \tag{3.13}$$

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

3.7. Methods of Minimum Distances

In this subsection we present three minimum distance estimators (also called maximum goodness-of-fit estimators) for λ and θ. This class of estimators are based on minimizing any empirical distribution function (EDF) statistics with respect to the unknown parameters (see, D'Agostino and Stephens, 1986; Luceño, 2006).

3.7.1. Method of Cramér-von-Mises

To motivate our choice of Cramér-von-Mises 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 Cramér-von Mises estimators $\widehat{\theta}_{CME}$ and $\widehat{\lambda}_{CME}$ of the parameters θ and λ are obtained by minimizing, with respect to θ and λ, the function:

$$C(\theta, \lambda) = \frac{1}{12n} + \sum_{i=1}^{n} \left(F(x_{i:n} \mid \theta, \lambda) - \frac{2i-1}{2n}\right)^2. \tag{3.14}$$

These estimators can also be obtained by solving the non-linear equations:

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

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

3.7.2. 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, 1952, 1954). Luceño (2006) provides some motivation about using such statistic and also introduces a modification, namely, Right-tail Anderson-Darling statistics. The Anderson-Darling estimators $\widehat{\theta}_{ADE}$ and $\widehat{\lambda}_{ADE}$ of the parameters θ and λ are obtained by minimizing, with respect to θ and λ, the function:

$$A(\theta,\lambda)=-n-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\left\{\log F\left(x_{i:n}\mid\theta,\lambda\right)+\log\overline{F}\left(x_{n+1-i:n}\mid\theta,\lambda\right)\right\}. \quad (3.15)$$

These estimators can also be obtained by solving the non-linear equations:

$$\begin{aligned}\sum_{i=1}^{n}(2i-1)\left[\frac{\eta_1\left(x_{i:n}\mid\theta,\lambda\right)}{F\left(x_{i:n}\mid\theta,\lambda\right)}-\frac{\eta_1\left(x_{n+1-i:n}\mid\theta,\lambda\right)}{\overline{F}\left(x_{n+1-i:n}\mid\theta,\lambda\right)}\right] &= 0,\\ \sum_{i=1}^{n}(2i-1)\left[\frac{\eta_2\left(x_{i:n}\mid\theta,\lambda\right)}{F\left(x_{i:n}\mid\theta,\lambda\right)}-\frac{\eta_2\left(x_{n+1-i:n}\mid\theta,\lambda\right)}{\overline{F}\left(x_{n+1-i:n}\mid\theta,\lambda\right)}\right] &= 0,\end{aligned}$$

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

The Right-tail Anderson-Darling estimators $\widehat{\theta}_{RTADE}$ and $\widehat{\lambda}_{RTADE}$ of the parameters θ and λ are obtained by minimizing, with respect to θ and λ, the function:

$$R(\theta,\lambda)=\frac{n}{2}-2\sum_{i=1}^{n}F\left(x_{i:n}\mid\theta,\lambda\right)-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\log\overline{F}\left(x_{n+1-i:n}\mid\theta,\lambda\right). \quad (3.16)$$

These estimators can also be obtained by solving the non-linear equations:

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

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

4. Simulation Study

In this section we develop a simulation study via Monte Carlo methods. The main goal of these simulations is to compare the efficiency of the different estimation methods for the parameters of the BE2 distribution. The following procedure was adopted:

1. Set the sample size n and the parameter values $\boldsymbol{\alpha}$.

2. Generate values of the $BE2(\theta, \lambda)$ with size n.
3. Using the values obtained in step 2, calculate $\hat{\theta}$ and $\hat{\lambda}$ via MLE, ME, MME, LSE, WLSE, PCE, MPS, CME, ADE, RTADE.
4. Repeat the steps 2 and 3 N times.
5. Using $\hat{\boldsymbol{\alpha}}$ and $\boldsymbol{\alpha}$, compute the mean relative estimates (MRE) $\sum_{i=1}^{N} \dfrac{\hat{\alpha}_i/\alpha}{N}$ and the mean square errors (MSE) $\sum_{i=1}^{N} \dfrac{(\hat{\alpha}_i - \alpha_i)^2}{N}$.

It is expected that for this approach the MRE's are closer to one with smaller MSE's. The results were computed using the software R (R Core Development Team). The seed used to generate the random values was 2015. The chosen values to perform this procedure were $\boldsymbol{\alpha} = (1, 0.8)$, $N = 10000$ and $n = (15, 20, 25, \cdots, 130)$.

Figure 1 shows the MRE's, MSE's from the estimates of λ and θ obtained using different estimation methods for N simulated samples and considering different values of n. The horizontal lines in Fig. 1 corresponds to MRE's and MSE's being respectively one and zero. We have presented results only for $\lambda = 1$ and $\theta = 0.8$ for reasons of space. But the following results were similar for other choices for λ and θ.

From these figure, we observe that the MSE of all estimators of the parameters tend to zero for large n and also, as expected, the values of MRE's tend to 1, i.e. the estimates are asymptotically unbiased for the parameters.

In the case of λ the CME, ADE and RTADE behave better than MLE for small sample sizes in terms of MRE's and MSE's. In the case of θ, the CME and RTADE indicate a good performance, however the MLE allow to get better estimates of $\hat{\theta}$ than the other methods for all sample sizes. The Percentile estimator has largest MSE's among all the considered estimators even for a large sample size.

Comparing all these we observe that, for small sample sizes, the CME and the RTADE estimators are highly competitive methods compared to the maximum likelihood method for estimating the parameters of the BE2 distribution.

5. Applications

5.1. Data sets

Located in southeastern Brazil, Sao 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 of high strategic and economic importance for the regional development. From Figure 2 we observe that the city has rainy periods from October to March, and from June to August exhibit more dry periods.

Consequently predict the behaviour of the transition periods in rainy sessions (April, May and September) enables that agriculturists be prepared against different

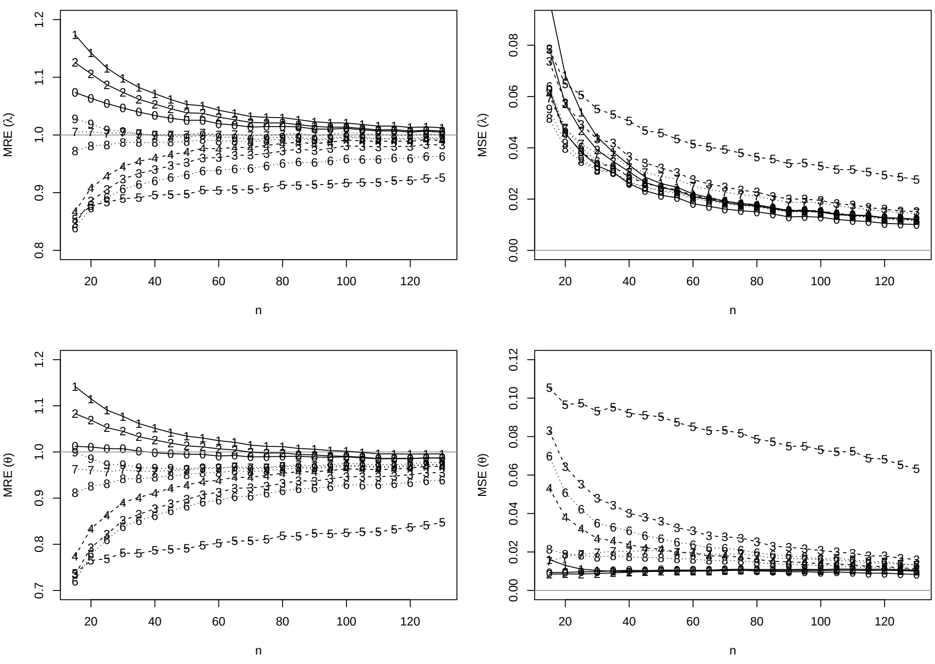

Figura 1: MRE's, MSE's related from the estimates of λ and θ for N simulated samples, considering different values of n obtained using the following estimation method 0-MLE, 1-ME, 2-MME, 3-LSE, 4-WLSE, 5-PCE, 6-MPS, 7-CME, 8-ADE, 9-RTADE.

problems (see Barbieri, 2007, for more details), 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 Sao 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 Sao Paulo including a period from 1960 to 2014.

5.2. Discrimination Criterion Methods

We firstly consider different discrimination criterion methods based on log-likelihood function evaluated at the MLEs. Let k be the number of parameters to be fitted and $\hat{\boldsymbol{\alpha}}$ the MLE's of $\boldsymbol{\alpha}$, the discrimination criterion methods are respectively: Akaike information criterion (AIC) computed through $AIC = -2l(\hat{\boldsymbol{\alpha}};\boldsymbol{x}) + 2k$, Corrected Akaike information criterion $AICC = AIC + \frac{2\,k\,(k+1)}{(n-k-1)}$, Hannan-Quinn information criterion $HQIC = -2\,l(\hat{\boldsymbol{\alpha}};\boldsymbol{x}) + 2\,k\,\log\left(\log(n)\right)$ and the consistent Akaike information criterion $CAIC = -2\,l(\hat{\boldsymbol{\alpha}};\boldsymbol{x}) + k\,(\log(n) + 1)$. The best model is the one which provides the minimum values of those criteria.

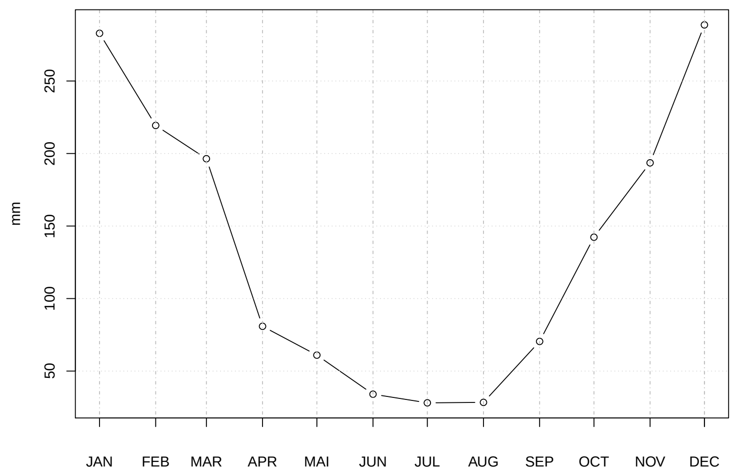

Figura 2: Average of the total monthly rainfall from January to December at Sao Carlos, Brazil.

For sake of comparison, to analyze the rainfall data, the results obtained from the BE2 distribution will be compared with the, Weibull, Gamma, Lognormal, Gumbel (Lawless, 2002) and Generalized Exponential (Gupta & Kundu, 1999, 2001) distributions and nonparametric survival function.

5.3. Results

Firstly, the data set related to May and September has the occurrence of zero value. Therefore to compute the MLE's of the Gamma, Weibull, Lognormal and Generalized Exponential distribution we approximate such values to 0.1. In the case of BE2 and Gumbel distributions we use the original values. Table 1 presents the results from AIC, AICC, HQIC and CAIC criteria, for different probability distributions. In the Figure 3, we have the survival function adjusted by different distributions and non-parametric survival estimator.

Comparing the empirical survival function with the adjusted distributions it can be observed a better fit for the BE2 distribution among the chosen models. This result is confirmed from AIC, AICC, HQIC and CAIC criteria as the BE2 distribution has the minimum values. Table 2 displays the MLE's and 95% confidence intervals for θ and λ of the BE2 distribution.

The discrimination criterion methods select the BE2 distribution as the best model among the compared models. The quantile-quantile (Q-Q) plot is a graphical technique that also provide an assessment of goodness of fit. If the data set come from the proposed distribution the points should fall approximately along the 45-degree reference line. In the Figure 4, we have the Q-Q plot from the proposed data set.

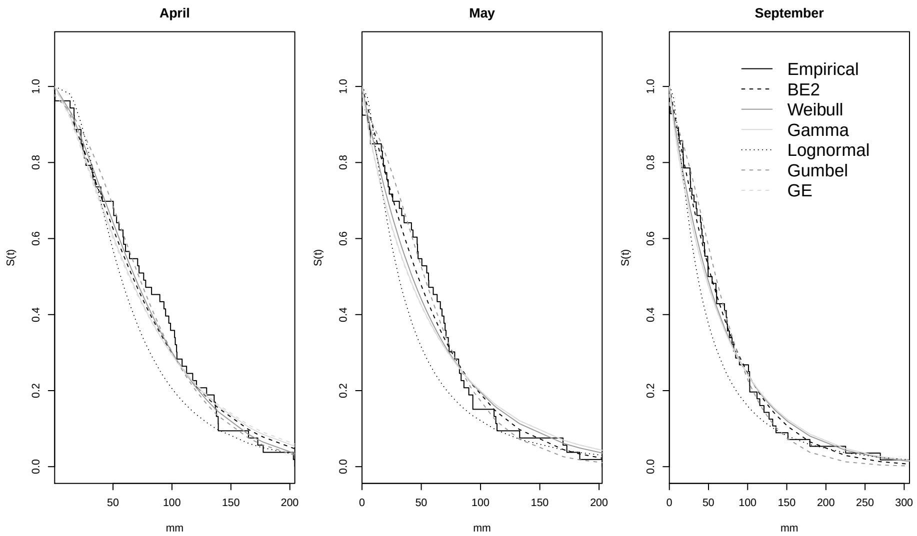

Figura 3: Survival function adjusted by different distributions and a non-parametric method considering the data sets related to the total monthly rainfall during April, May and September at Sao Carlos

Tabela 1: Results of the AIC, AICC, HQIC and CAIC criteria for different probability distributions considering the data sets related to the total monthly rainfall during April, May and September at Sao Carlos.

Month	Test	BE2	Weibull	Gamma	Lognormal	Gumbel	GE
April	AIC	**566.422**	566.814	570.371	595.503	568.860	571.125
	AICC	**566.662**	567.054	570.611	595.743	569.100	571.365
	HQIC	**567.937**	568.329	571.886	597.018	570.375	572.640
	CAIC	**572.362**	572.755	576.312	601.443	574.801	577.066
May	AIC	**543.378**	545.769	544.980	580.391	553.281	544.751
	AICC	**543.618**	546.009	545.220	580.631	553.521	544.991
	HQIC	**544.893**	547.284	546.495	581.906	554.796	546.266
	CAIC	**549.319**	551.710	550.921	586.332	559.222	550.692
September	AIC	**591.569**	592.515	591.911	625.321	604.370	591.780
	AICC	**591.795**	592.741	592.137	625.547	604.596	592.006
	HQIC	**593.139**	594.085	593.481	626.891	605.940	593.350
	CAIC	**597.620**	598.566	597.962	631.372	610.421	597.831

From the Figure 4, we observe that the points are approximately along the reference line. Therefore, through the proposed methodology the data related to

Tabela 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 Sao Carlos.

Month	$\boldsymbol{\alpha}$	MLE	$CI_{95\%}(\alpha)$
April	θ	0.9100	(0.6436;0.9826)
	λ	0.0227	(0.0180;0.0286)
May	θ	0.7097	(0.3561;0.9153)
	λ	0.0254	(0.0186;0.0346)
September	θ	0.6375	(0.1766;0.9351)
	λ	0.0208	(0.0139;0.0313)

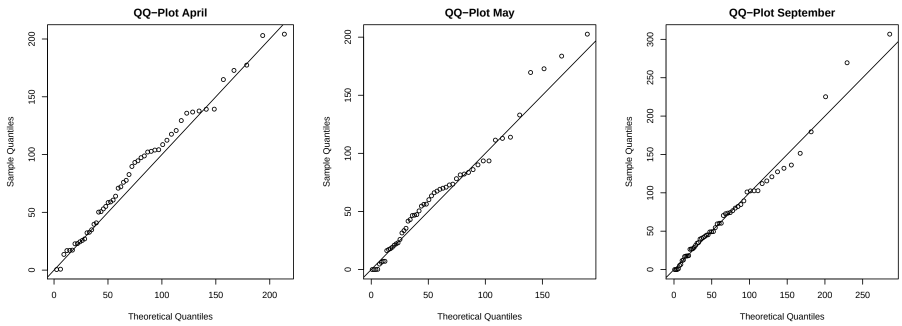

Figura 4: Q-Q plot considering the data sets related to the total monthly rainfall during April, May and September at Sao Carlos.

the total monthly rainfall during April, May and September at Sao Carlos can be described by the Binomial-exponential 2 distribution.

6. Conclusions

In this paper, we derived and compared, via intensive simulation experiments, the estimation of the parameters of the Binomial-exponential 2 distribution using ten estimation methods. The simulations show that, for small sample sizes, the Cramer-von Mises type minimum distance estimators and the Right-tail Anderson-Darling estimators are highly competitive methods compared to the maximum likelihood method for estimating the parameters of the BE2 distribution. We also apply our proposed methodology in three real data sets related to the total monthly rainfall during April, May and September at Sao Carlos, Brazil, demonstrating that the BE2 distribution is a simple alternative to be used in weather forecast models.

7. Appendix

- 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.

Referências

[1] Anderson, T.W. and Darling, D.A. (1952). Asymptotic theory of certain "goodness of fit"criteria based on stochastic processes. Ann. Math. Statist., 23,193-212.

[2] Anderson, T.W. and Darling, D.A. (1954). A test of goodness of fit, J. Amer. Stat. Assn., 49, 765-769

[3] Alkasasbeh, M. R. and Raqab, M. Z. (2009). Estimation of the generalized logistic distribution parameters: Comparative study. Statistical Methodology, 6, 262-279.

[4] Bakouch, H.S. Jazi, M.A. Nadarajah, S. Dolati, A. Roozegar, R. (2014). A lifetime model with increasing failure rate, Applied Mathematical Modelling, 38(23), 5392-5406.

[5] Barbieri, P.R.B. (2007). Caracterizacao da estacao chuvosa nas regioes sul e sudeste do Brasil associado com a circulacao atmosferica, São José dos Campos: INPE, 118 p.

[6] Casella, G. Berger, R. (2002). Statistical Inference (2nd ed.), Belmont, CA: Duxbury.

[7] Cheng, R.C.H., Amin, N.A.K. (1979). Maximum product of spacings estimation with application to the lognormal distributions. Math Report 79-1, Department of Mathematics, UWIST, Cardiff.

[8] Cheng, R.C.H., Amin, N.A.K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. J. Roy. Statist. Soc. Ser. B 45, 394-403.

[9] D'Agostino, R. Stephens, M. (1986). Goodness-of-fit techniques. Marcel Dekker, Inc., New York.

[10] Dey, S., Dey, T. and Kundu, D. (2014). Two-parameter Rayleigh distribution: Different methods of estimation. American Journal of Mathematical and Management Sciences, 33, 55–74.

[11] Gupta, R.D. and Kundu, D. (1999). Generalized exponential distributions. Austral. NZ J. Statist., 41, 173-188.

[12] Gupta, R.D. and Kundu, D. (2001). Generalized exponential distributions: Different methods of estimations. J. Statist. Comput. Simul. 69, 315-338.

[13] Kao, J.H.K. (1958). Computer methods for estimating Weibull parameters in reliability studies. Trans. IRE Reliability Quality Control, 13, 15-22.

[14] Kao, J.H.K. (1959). A graphical estimation of mixed Weibull parameters in life testing electron tube. Technometrics, 1, 389 - 407.

[15] Kundu, D. Raqab, M.Z. (2005). Generalized Rayleigh distribution: different methods of estimations. Computational Statistics and Data Analysis 49 (1) 187-200.

[16] Lawless, J. F. (2002). Statistical models and methods for lifetime data, Second Edition, New York: John Wiley and Sons.

[17] Luceno A (2006). Fitting the Generalized Pareto Distribution to Data using Maximum Goodness-of-fit Estimators. Computational Statistics and Data Analysis, 51(2), 904-917.

[18] Macdonald, P.D.M. (1971). An estimation procedure for mixtures of distribution. J. Roy. Statist. Soc. Ser. B, 33, 326-329.

[19] Mann, N. R., Schafer, R. E. and Singpurwalla, N. D.(1974). Methods for Statistical Analysis of Reliability and Life Data, New York, Wiley.

[20] Mazucheli, J. Louzada, F. Ghitany, M.E. (2013). Comparison of estimation methods for the parameters of the weighted Lindley distribution. Applied Mathematics and Computation, 220(1) , 463-471.

[21] Ranneby, B. (1984). The maximum spacing method: An estimation method related to the maximum likelihood method. Scandinavian Journal of Statistics 11, 93-112.

[22] Swain, J.J., Venkataraman, S. Wilson, J.R. (1988). Least squares estimation of distribution functions in Johnson's translation system. Journal of Statistical Computation and Simulation, 29, 271 - 297.

[23] Teimouri, M. Hoseini, S.M. Nadarajah, S. (2013). Comparison of estimation methods for the Weibull distribution. Statistics 47 (1) 93-109.