On Comparing Different Methods of Estimation for the
Parameters of a Pathological Distribution with
Application to Climate Data

Farouq Mohammad A. Alam
Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract

In statistical literature, various probability distributions exist with advantageous properties, while others are considered pathological since their properties are counterintuitive. A well-known pathological probability distribution is the Cauchy distribution, and it has applications in areas related to environmental and financial research. Both the log-Cauchy and half-Cauchy distributions, which have close connections to the Cauchy distribution, are pathological distributions. This paper considers another pathological model called the Cauchy Birnbaum-Saunders distribution. Some of the statistical properties of this distribution are discussed briefly, and its parameters are estimated using eight frequentist estimation methods, including the maximum likelihood, least-squares-based, and minimum distance estimation methods. Monte Carlo simulations are carried out to compare and examine the performance of each estimator numerically. Furthermore, a recent climate data set is analyzed to show the practical applicability of this model.

Keywords: Cauchy Birnbaum-Saunders distribution, Birnbaum-Saunders distribution, maximum likelihood estimation, least-square-based estimation, minimum distance estimation.

1. Introduction

In the statistical literature, pathological distributions, such as the log-Cauchy distribution and the half-Cauchy distribution, have scarce applicability in modeling real-life phenomena since they violate the assumptions of the central limit theorem. In practice, finite moments of the first order and above do not exist for these probabilistic models. Consequently, they are not commonly used in reliability analysis. Nevertheless, they still have applications in other areas of science, including statistical inference and analysis of data with extremes. In fact, the half-Cauchy distribution can be used in a Bayesian framework when a proper prior is necessary; see, for example, Gelman (2006); Polson and Scott (2012). On the other hand, the log-Cauchy distribution can be considered to model survival processes when data contamination exists (Lindsey et al. 2001). From a statistical perspective, data contamination means a data set with outliers, extremes, or both. This study considers another important pathological distribution, namely, the Cauchy Birnbaum-Saunders (CBS) distribution. Recently, Wang et al. (2020) mathematically discussed the shape of both density function and failure rate
function of the CBS distribution. They also studied quasi-maximum likelihood estimator for the shape parameter by assuming that the point estimate of the scale parameter of the CBS distribution is the sample median. Furthermore, they acquired regression and logarithmic moment estimations for the model parameters. By contrast, this study considers several frequentist estimators which are different from those considered by Wang et al. (2020).

In material sciences, researchers study materials’ fatigue of certain substances to determine their reliability. Material fatigue is defined as the weakening of a material caused by subjecting it to periodic stress factors such as pressure and cyclic loading. Birnbaum and Saunders (1969a,b) proposed a distribution that bear their names to model fatigue data. The Birnbaum-Saunders (BS) distribution is a non-negative, skewed probability distribution with a non-monotonic (hazard) failure rate function. Non-monotonic hazard rate can be observed in practice; see, for example, Langlands et al. (1979). A non-negative continuous random variable $X$ is said to follow the BS distribution based with shape parameter $\alpha > 0$, scale parameter $\beta > 0$, if the corresponding cumulative distribution function (CDF) is given by:

$$F(x; \alpha, \beta) = \Phi \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right] \right), \quad x > 0, \quad \alpha, \beta > 0,$$

where $\Phi(\cdot)$ is the CDF of the standard Gaussian distribution. In Desmond (1985), a more general derivation for the BS distribution was provided based on a biological model which strengthened the physical rationalization for the use of the BS distribution by relaxing the original presumptions of Birnbaum and Saunders (1969a,b). Similar to the log-Gaussian distribution, there is a relation between the BS distribution and the Gaussian distribution. Thus, it received considerable attention in statistical literature. In fact, at least a couple of hundred papers and a single research monograph have already appeared describing all properties and developments of this distribution. For more details, see the comprehensive review Balakrishnan and Kundu (2019) in this connection. The following are some recent applications of the BS distribution. In Bourguignon et al. (2020), control charts were discussed for monitoring the median of the BS distribution. In practice, the scale parameter of the BS distribution represents the median of failure time. In Hassani et al. (2020), an extension for the BS distribution was proposed to model bicoid gene expression data. The repetitive group sampling plan under BS distribution was developed in Kannan et al. (2020). Recently, EL-Sharkawy and Ismail (2021) studied the problem of selecting the number of components arising from a mixture of BS distributions.

During the last decade, a generalization of the BS distribution was proposed in Díaz-García and Leiva-Sánchez (2005, 2007) by replacing the Gaussian kernel in Eq. (1) with other elliptically and symmetrically kernels such as those of Student’s $t$, Cauchy, and Laplace distributions. Later on, sampling plans from truncated life tests assuming the GBS distribution were developed in Balakrishnan et al. (2007), while the GBS distribution was used to analyze air pollutant concentration in Leiva et al. (2008). For further details in this connection, see Sanhueza et al. (2008). Student’s $t$ BS ($t$BS) distribution is a particulate case from the generalized BS family of distributions. A non-negative continuous random variable $X$ is said to follow $t$BS distribution with shape parameter $\alpha > 0$, scale parameter $\beta > 0$, and degrees of freedom $\nu > 0$, if the corresponding cumulative distribution function (CDF) is given by:

$$F(x; \alpha, \beta) = \Phi_{\nu} \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right] \right), \quad x > 0, \quad \alpha, \beta > 0,$$

where $\Phi_{\nu}(\cdot)$ is the CDF of the standard $t$ distribution with degrees of freedom $\nu$. Clearly, both the CBS distribution ($\nu = 1$) and the traditional BS distribution ($\nu \to \infty$) are sub-models of $t$BS distribution. Hence, one can use this family of probability distributions to model fatigue lifetime data since its special case; namely, the BS distribution is used to analyze such data. Consequently, the importance of the distribution under investigation (i.e., the CBS distribution) lies in the fact that one might use it to analyze fatigue information instead of the
conventional BS distribution when data contamination exists. Balakrishnan et al. (2009) represented the $t$BS distribution as a BS distribution conditioned on scale mixtures of Gaussian random variables. Azevedo et al. (2012) studied some mathematical aspects of the hazard rate function of the $t$BS distribution; namely, they examined the shape of the failure rate function of this distribution alongside the change point and discussed associated estimation. Recently, the parameters of the $t$BS distribution were estimated using the maximum likelihood theory and the expectation-maximization (EM) algorithm in Balakrishnan and Alam (2019). They reported a few anomalies in their simulation outcomes, such as the unsatisfactory efficiency of both the maximum likelihood estimation and the EM algorithm in the case of the $t$BS distribution with a unity degree of freedom, i.e., the case of the CBS distribution. This sub-model is, in fact, a pathological probabilistic model with undefined distributional properties. Consequently, estimation procedures such as the moments method or one of its variants are impossible to consider here.

In practice, estimating the distributional properties of a probability distribution (e.g., reliability, hazard rate, etc.) requires finding suitable estimates for the model parameters using appropriate estimation methods. Estimating the probability distribution parameters is of great interest to statisticians and has received much attention in the statistical literature. In fact, many researchers conducted comparative studies over time to numerically examine the performance of various estimation methods from different perspectives; see, for example, Gupta and Kundu (2001); Alkasasbeh and Raqab (2009); Mazucheli et al. (2013); do Espirito Santo and Mazucheli (2014); Qoshja and Hoxha (2016); Dey et al. (2017); Balakrishnan and Alam (2019); Alam and Nassar (2021), among other contributions. This paper investigates the performance of eight frequentist estimation methods for the model parameters and examines their robustness since data contamination is expected to exist due to the CBS distribution nature. Data contamination motivated many researchers to propose robust estimators for various distributions; see, for example, Lawson et al. (1997); Boudt et al. (2009); Agostinelli et al. (2014); Wang et al. (2015), among other papers. The obtained estimators for the model parameters in this study are the maximum likelihood estimators (MLEs), the least-squares estimators (LSEs), the weighted least-squares estimators (WLSEs), the percentile estimators (PCEs), the maximum product of spacings estimators (MPSEs), the Cramér-von Mises estimators (CVMEs), the Anderson-Darling estimators (ADEs), the right-tailed Anderson-Darling estimators (RADEs). As previously mentioned, some aspects of the CBS distribution are counterintuitive. Hence, estimation methods, such as the modified moments methods can not be used as indicated by Balakrishnan and Alam (2019), while other estimation methods might not thrive in this case. Therefore, the aim of this study is to address this scientific computation challenge by performing numerical comparison between several estimators for the model parameters to identify which of them perform superiorly in terms of estimation efficiency and robustness. The comparison is based on Monte Carlo simulations and the outcomes of a real data analysis. The originality of this contribution comes from the fact that there has been no previous work about comparing different estimation methods to estimate the parameters of the CBS distribution. The remaining part of this article is organized as follows. Sections 2 and 3 discusses the distributional aspects of the CBS distribution and the eight frequentist estimation methods of interest, respectively. Section 4 reports the outcomes of extensive Monte Carlo simulations to compare the behavior of each estimator under different settings. Section 5 illustrates the practical application of the discussed estimators using a real data set. Finally, the paper is concluded with remarks and future research directions in Section 6.

2. Model description

A non-negative continuous random variable $X$ is said to follow the CBS distribution based with shape parameter $\alpha > 0$, scale parameter $\beta > 0$, if the corresponding cumulative distri-
bution function (CDF), probability density function (PDF) are given by:

\[
F(x; \alpha, \beta) = \Phi_1 \left( \frac{A(x; \beta)}{\alpha} \right) = \frac{1}{\pi} \arctan \left( \frac{A(x; \beta)}{\alpha} \right) + \frac{1}{2},
\]

and

\[
f(x; \alpha, \beta) = \frac{a(x; \beta)}{2\alpha x} \phi_1 \left( \frac{A(x; \beta)}{\alpha} \right) = \frac{a(x; \beta)}{\alpha \pi \left( 1 + \left[ \frac{A(x; \beta)}{\alpha} \right]^2 \right)},
\]

respectively, such that \( \Phi_1(\cdot) \) and \( \phi_1(\cdot) \) are the CDF and PDF of the standard Cauchy distribution, respectively, where

\[
A(x; \beta) = \frac{x - \beta}{\sqrt{\beta x}},
\]

and

\[
a(x; \beta) = \frac{x + \beta}{\sqrt{\beta x}}.
\]

Figure 1 illustrates the PDF of the CBS distribution, while Figure 2 shows the hazard function of the CBS distribution. Addressing the mathematical aspects of these non-monotonic distributional features are beyond the scope of this study. The following properties for this distribution are obtained from Balakrishnan and Alam (2019) straightforwardly.

**Property 1.** Suppose that \( X \sim \text{CBS} (\alpha, \beta) \), the quantile function (QF) can be straightforwardly obtained as

\[
Q(u, \alpha, \beta) = \frac{\beta}{4} \left\{ \alpha q(u) + \sqrt{[\alpha q(u)]^2 + 4} \right\}^2, \quad 0 < u < 1,
\]

such that \( q(u) = \tan (\pi [u - 0.5]) \).

**Property 2.** Suppose that \( T \) follows the standard Cauchy distribution, i.e. \( T \sim \text{Cauchy}(0, 1) \), then

\[
X = \frac{\beta}{4} \left( \alpha T + \sqrt{[\alpha T]^2 + 4} \right)^2 \sim \text{CBS} (\alpha, \beta).
\]

**Property 3.** The CBS distribution is a pathological distribution since key distributional properties (e.g., the mean, variance, etc.) are undefined.

**Property 4.** If \( X \sim \text{CBS} (\alpha, \beta) \), then:

1. \( A(x; \beta) \sim \text{Cauchy}(0, \alpha) \), such that \( A(x; \beta) \) is given by (5);
2. \( cX \sim \text{CBS} (\alpha, c\beta) \) given that \( c > 0 \);
3. \( X^{-1} \sim \text{CBS} (\alpha, \beta^{-1}) \).

**Property 5.** If \( X|y \sim \text{BS} \left( \frac{\alpha}{\sqrt{y}}, \beta \right) \), and \( Y \sim \chi^2_1 \), then

\[
X \sim \text{CBS} (\alpha, \beta) \quad \text{and} \quad Y|x \sim \text{Exponential} \left( \text{rate} = \frac{1 + A^2(x; \beta)}{2} \right),
\]

such that \( A(x; \beta) \) is given by (5).
On compar. diff. methods of est. for the par. of a path. dist. with app. to climate data

Figure 1: The PDF of the CBS distribution for different value of $\alpha$ and a unit scale parameter

Figure 2: The HF of the CBS distribution for different value of $\alpha$ and a unit scale parameter
3. Estimation methods

In this section, eight estimation methods for the model parameters are discussed.

3.1. Maximum likelihood estimation

Let \( \mathbf{x} = (x_1, \ldots, x_n) \) be the realization of \( n \) independent and identically distributed random variables (i.e., an observed random sample with \( n \) observations) from the CBS distribution with probability density function (4). The corresponding MLEs of \( \alpha \) and \( \beta \); say, \( \hat{\alpha}_{\text{MLE}} \) and \( \hat{\beta}_{\text{MLE}} \), respectively, are obtained by maximizing the following objective function:

\[
\ell(\alpha, \beta|\mathbf{x}) = -n \log(\alpha) - \frac{n}{2} \log(\beta) + \sum_{i=1}^{n} \log(x_i + \beta) - \sum_{i=1}^{n} \log \left( 1 + A_i^2 \right),
\]

where \( \ell(\alpha, \beta|\mathbf{x}) \) is the log-likelihood function of the CBS distribution (without the constant term), such that \( A_i = A(x_i; \beta) \) is given by Eq. (5), where \( i = 1, \ldots, n \). One possible approach to obtain the MLEs numerically is by using a suitable optimization method which may require the following non-linear log-likelihood (normal) equations:

\[
\frac{\partial \ell}{\partial \alpha} = -\frac{n}{\alpha} + 2 \sum_{i=1}^{n} \frac{A_i^2}{1 + [\alpha^{-1} A_i]^2} = 0
\]

and

\[
\frac{\partial \ell}{\partial \beta} = -\frac{n}{2\beta} + \sum_{i=1}^{n} \frac{1}{x_i + \beta} + \frac{1}{\alpha^2 \beta} \sum_{i=1}^{n} \frac{a_i A_i}{1 + [\alpha^{-1} A_i]^2} = 0,
\]

where \( a_i = a(x_i; \beta) \) is given by Eq. (6), such that \( i = 1, \ldots, n \).

3.2. Least-squares estimation

In Swain et al. (1988), LSEs was used to estimate the parameters of the beta distribution. Suppose that \( x_{1:n} < \ldots < x_{n:n} \) are the observed order statistics from an observed random sample \( \mathbf{x} \) of size \( n \) from an arbitrary probability distribution with CDF \( F(x) \). From Arnold et al. (2008), it is known that:

\[
E[F(x_{i:n})] = \frac{i}{n+1},
\]

since \( F(x_{1:n}), \ldots, F(x_{n:n}) \) are order statistics form a standard uniform distribution. The LSEs of \( \alpha \) and \( \beta \); say, \( \hat{\alpha}_{\text{LSE}} \) and \( \hat{\beta}_{\text{LSE}} \), respectively, are acquired by minimizing:

\[
S(\alpha, \beta) = \sum_{i=1}^{n} \left[ \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{\beta}{x_{i:n}}} \right] \right) - \frac{i}{n+1} \right]^2.
\]

The first-order derivatives of Eq. (9) with respect to \( \alpha \) and \( \beta \) are given by:

\[
\frac{\partial S}{\partial \alpha} = 2 \sum_{i=1}^{n} \left[ \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{\beta}{x_{i:n}}} \right] \right) - \frac{i}{n+1} \right] \varpi_{\alpha}(x_{i:n}),
\]

and

\[
\frac{\partial S}{\partial \beta} = 2 \sum_{i=1}^{n} \left[ \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{\beta}{x_{i:n}}} \right] \right) - \frac{i}{n+1} \right] \varpi_{\beta}(x_{i:n}),
\]

respectively, such that

\[
\varpi_{\alpha}(x) = -\frac{A(x; \beta)}{\alpha^2} \phi_1 \left( \frac{A(x; \beta)}{\alpha} \right),
\]

and

\[
\varpi_{\beta}(x) = -\frac{a(x; \beta)}{2\alpha^2} \phi_1 \left( \frac{A(x; \beta)}{\alpha} \right),
\]
where $A(x; \beta)$ and $a(x; \beta)$ are given by Eq. (5) and Eq. (6), respectively. It is to be noted that these expressions are frequently used in the following subsection.

3.3. Weighted least squares estimation

In Swain et al. (1988), the WLSEs were also considered to estimate the parameters of the beta distribution alongside the LSEs. Both LSEs and WLSEs were considered by several researchers, see; for example, Gupta and Kundu (2001); Alkasasbeh and Raqab (2009); Mazucheli et al. (2013); do Espirito Santo and Mazucheli (2014). Recall that $x_{1:n} < \ldots < x_{n:n}$ are the observed order statistics. Since $F(x_{1:n}), \ldots, F(x_{n:n})$ are order statistics form a standard uniform distribution, then from Arnold et al. (2008), it is known that:

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

The WLSEs of $\alpha$ and $\beta$; say, $\hat{\alpha}_{\text{WLSE}}$ and $\hat{\beta}_{\text{WLSE}}$, respectively, are obtained by minimizing the following weighted objective function:

$$S_w(\alpha, \beta) = \sum_{i=1}^{n} w_i \left[ \Phi_1 \left( \frac{1}{\alpha} \left[ \frac{x_{i:n}}{\beta} - \sqrt[4]{\frac{\beta}{x_{i:n}}} \right] \right) - \frac{i}{n+1} \right]^2, \tag{12}$$

where $w_1, \ldots, w_{n:n}$ are the weights of $x_{1:n} < \ldots < x_{n:n}$, respectively, such that $w_{i:n} = \frac{(n+1)^2(n+1)}{i(n-i+1)}$.

The first-order derivatives of Eq. (12) with respect to $\alpha$ and $\beta$ are given by:

$$\frac{\partial S_w}{\partial \alpha} = 2 \sum_{i=1}^{n} w_{i:n} \left[ \Phi_1 \left( \frac{1}{\alpha} \left[ \frac{x_{i:n}}{\beta} - \sqrt[4]{\frac{\beta}{x_{i:n}}} \right] \right) - \frac{i}{n+1} \right] \varpi_\alpha(x_{i:n}),$$

and

$$\frac{\partial S_w}{\partial \beta} = 2 \sum_{i=1}^{n} w_{i:n} \left[ \Phi_1 \left( \frac{1}{\alpha} \left[ \frac{x_{i:n}}{\beta} - \sqrt[4]{\frac{\beta}{x_{i:n}}} \right] \right) - \frac{i}{n+1} \right] \varpi_\beta(x_{i:n}),$$

respectively.

3.4. Percentile estimation

Percentile estimation methods is also a nonlinear regression-based method and it is commonly used when the CDF and QF have closed-form expressions. The idea of this method is to fit a linear model to the theoretical percentiles obtained from the QF and the sample percentiles, see Kao (1958, 1959) in this connection. The percentile methods is considered by other researchers, see; for example, Gupta and Kundu (2001); Akkasasbeh and Raqab (2009); Al-Mofleh et al. (2020). Recall that $x_{1:n} < \ldots < x_{n:n}$ are the realizations of the order statistics from a random sample $x_1, \ldots, x_n$ of size $n$. Based on Eq. (7), and if $F(x_{i:n}; \alpha, \beta)$ are estimated by $p_{i:n} = \text{E} [F(x_{i:n})]$ for $i = 1, \ldots, n$, then the percentile estimators of $\alpha$ and $\beta$; say, $\hat{\alpha}_{\text{PCE}}$ and $\hat{\beta}_{\text{PCE}}$, can be obtained by minimizing:

$$S_p = \sum_{i=1}^{n} \left( x_{i:n} - \frac{\beta}{4} \left[ \frac{\alpha \Phi^{-1}(p_{i:n}) + \sqrt{4 + \{\alpha \Phi^{-1}(p_{i:n})\}^2}}{4 + \{\alpha \Phi^{-1}(p_{i:n})\}^2} \right] \right)^2, \tag{13}$$

where $p_{i:n} = \frac{i}{n+1}$. The first-order derivatives of Eq. (13) with respect to $\alpha$ and $\beta$ are given by:

$$\frac{\partial S_p}{\partial \alpha} = 2 \sum_{i=1}^{n} \left( x_{i:n} - \frac{\beta}{4} \left[ \frac{\alpha \Phi^{-1}(p_{i:n}) + \sqrt{4 + \{\alpha \Phi^{-1}(p_{i:n})\}^2}}{4 + \{\alpha \Phi^{-1}(p_{i:n})\}^2} \right] \right) \varpi_\alpha(p_{i:n}),$$

and

$$\frac{\partial S_p}{\partial \beta} = 2 \sum_{i=1}^{n} \left( x_{i:n} - \frac{\beta}{4} \left[ \frac{\alpha \Phi^{-1}(p_{i:n}) + \sqrt{4 + \{\alpha \Phi^{-1}(p_{i:n})\}^2}}{4 + \{\alpha \Phi^{-1}(p_{i:n})\}^2} \right] \right) \varpi_\beta(p_{i:n}),$$

respectively.
and 
\[ \frac{\partial S_p}{\partial \beta} = 2 \sum_{i=1}^{n} \left\{ x_{in} - \frac{\beta}{4} \left[ \alpha \Phi^{-1}(p_{i:n}) + \sqrt{4 + \{\alpha \Phi^{-1}(p_{i:n})\}^2} \right] \right\} \omega(\beta), \]
respectively, such that
\[ \omega_\alpha(u) = -2 \frac{\Phi_1^{-1}(u)Q(u; \alpha, \beta)}{\sqrt{4 + \{\alpha \Phi_1^{-1}(u)\}^2}} \]
and
\[ \omega_\beta(u) = -Q(u; \alpha, 1), \]
where \(Q(u; \alpha, \beta)\) is given by Eq. (7). Due to the fact that the CBS distribution is a heavy-tailed distribution, it is expected that this method will perform poorly compared to the other methods.

3.5. Maximum produce of spacings estimation

Recent research indicates that MPSEs currently rivals MLEs in terms of estimation efficiency. It was formally introduced in Cheng and Amin (1979, 1983); Ranneby (1984). MPSEs belong to a class of a more general estimation method using spacings; see in this connection Ghosh and Jammalamadaka (2001). If \(x_{1:n} < \ldots < x_{n:n}\) are the observed order statistics based on a random sample of size \(n\), then the MPSEs for of \(\alpha\) and \(\beta\); say, \(\hat{\alpha}_{\text{MPSE}}\) and \(\hat{\beta}_{\text{MPSE}}\), respectively, are acquired by maximizing:
\[ P = \frac{1}{n+1} \sum_{i=1}^{n+1} \log \nabla_i, \quad (14) \]
such that
\[ \nabla_i = \begin{cases} 
\Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{x_{1:n}}{\beta}} \right] \right) & \text{if } i = 1 \\
\Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{x_{1:n}}{\beta}} \right] \right) - \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i-1:n}}{\beta}} - \sqrt{\frac{x_{i-1:n}}{\beta}} \right] \right) & \text{if } 1 < i \leq n \\
1 - \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{x_{1:n}}{\beta}} \right] \right) & \text{if } i = n+1 
\end{cases} \]
The first-order derivatives of Eq. (14) with respect to the model parameters are as follows:
\[ \frac{\partial P}{\partial \alpha} = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{\omega_\alpha \nabla_i}{\nabla_i}, \]
and
\[ \frac{\partial P}{\partial \beta} = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{\omega_\beta \nabla_i}{\nabla_i}, \]
where
\[ \omega_\alpha \nabla_i = \begin{cases} 
\omega_\alpha(x_{1:n}) & \text{if } i = 1 \\
\omega_\alpha(x_{1:n}) - \omega_\alpha(x_{i-1:n}) & \text{if } 1 < i \leq n \\
-\omega_\alpha(x_{n:n}) & \text{if } i = n+1 
\end{cases} \]
and
\[ \omega_\beta \nabla_i = \begin{cases} 
\omega_\beta(x_{1:n}) & \text{if } i = 1 \\
\omega_\beta(x_{1:n}) - \omega_\beta(x_{i-1:n}) & \text{if } 1 < i \leq n \\
-\omega_\beta(x_{n:n}) & \text{if } i = n+1 
\end{cases} \]

3.6. Cramér-von Mises estimation

The Cramér-von Mises method belongs to the class of minimum distance methods, and the corresponding CVMEs of of \(\alpha\) and \(\beta\); say, \(\hat{\alpha}_{\text{CVME}}\) and \(\hat{\beta}_{\text{CVME}}\), respectively, are found by
minimizing:

\[ C = \frac{1}{12n} + \sum_{i=1}^{n} \left[ \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{\beta}{x_{i:n}}} \right] \right) - \frac{2i-1}{2n} \right]^2, \]  

(15)
such that \( x_{1:n} < \ldots < x_{n:n} \) are the observed order statistics of a random sample of size \( n \). Moreover, the first-order derivatives of Eq. (15) with respect to the model parameters are as follows:

\[
\frac{\partial C}{\partial \alpha} = 2 \sum_{i=1}^{n} \left[ \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{\beta}{x_{i:n}}} \right] \right) - \frac{2i-1}{2n} \right] \varpi_{\alpha}(x_{i:n}),
\]

and

\[
\frac{\partial C}{\partial \beta} = 2 \sum_{i=1}^{n} \left[ \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{\beta}{x_{i:n}}} \right] \right) - \frac{2i-1}{2n} \right] \varpi_{\beta}(x_{i:n}).
\]

3.7. Anderson-Darling estimation

Anderson-Darling estimation and its right-tailed version also belong to the class of minimum distance methods. In this subsection, the former is considered, while the latter is discussed in the following subsection. Based on the \( x_{1:n} < \ldots < x_{n:n} \), then the ADEs for of \( \alpha \) and \( \beta \); say, \( \hat{\alpha}_{\text{ADE}} \) and \( \hat{\beta}_{\text{ADE}} \), respectively, are determined by minimizing:

\[
A = n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \log \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{n-i+1:n}}{\beta}} - \sqrt{\frac{\beta}{x_{n-i+1:n}}} \right] \right)
- \frac{1}{n} \sum_{i=1}^{n} (2i-1) \log \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{\beta}{x_{i:n}}} \right] \right).
\]  

(16)
The first-order derivatives of Eq. (16) with respect to the model parameters are as follows:

\[
\frac{\partial A}{\partial \alpha} = \frac{1}{n} \sum_{i=1}^{n} (2i-1) \frac{\varpi_{\alpha}(x_{n-i+1:n})}{\Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{n-i+1:n}}{\beta}} - \sqrt{\frac{\beta}{x_{n-i+1:n}}} \right] \right)}
- \frac{1}{n} \sum_{i=1}^{n} (2i-1) \frac{\varpi_{\alpha}(x_{i:n})}{\Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{\beta}{x_{i:n}}} \right] \right)},
\]

and

\[
\frac{\partial A}{\partial \beta} = \frac{1}{n} \sum_{i=1}^{n} (2i-1) \frac{\varpi_{\beta}(x_{n-i+1:n})}{\Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{n-i+1:n}}{\beta}} - \sqrt{\frac{\beta}{x_{n-i+1:n}}} \right] \right)}
- \frac{1}{n} \sum_{i=1}^{n} (2i-1) \frac{\varpi_{\beta}(x_{i:n})}{\Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{\beta}{x_{i:n}}} \right] \right)}.
\]

3.8. Right-tailed Anderson-Darling estimation

From a random sample of size \( n \), and based on the corresponding observed order statistics \( x_{1:n} < \ldots < x_{n:n} \), then the RADEs for of \( \alpha \) and \( \beta \); say, \( \hat{\alpha}_{\text{RADE}} \) and \( \hat{\beta}_{\text{RADE}} \), respectively, are determined by minimizing:

\[
A_R = \frac{n}{2} - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \log \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{n-i+1:n}}{\beta}} - \sqrt{\frac{\beta}{x_{n-i+1:n}}} \right] \right)
- 2 \sum_{i=1}^{n} \Phi_1 \left( \frac{1}{\alpha} \left[ \sqrt{\frac{x_{i:n}}{\beta}} - \sqrt{\frac{\beta}{x_{i:n}}} \right] \right).
\]  

(17)
The first-order derivatives of Eq. (17) with respect to the model parameters are as follows:

\[
\frac{\partial A_R}{\partial \alpha} = \frac{1}{n} \sum_{i=1}^{n} (2i - 1) \frac{\varpi_\alpha(x_{n-i+1:n})}{\Phi_1 \left( \frac{-1}{\alpha} \left[ \sqrt{\frac{x_{n-i+1:n}}{\beta}} - \sqrt{\frac{x_{n-i+1:n}}{\beta}} \right] \right)} - 2 \sum_{i=1}^{n} \varpi_\alpha(x_{i:n}),
\]

and

\[
\frac{\partial A_R}{\partial \beta} = \frac{1}{n} \sum_{i=1}^{n} (2i - 1) \frac{\varpi_\beta(x_{n-i+1:n})}{\Phi_1 \left( \frac{-1}{\alpha} \left[ \sqrt{\frac{x_{n-i+1:n}}{\beta}} - \sqrt{\frac{x_{n-i+1:n}}{\beta}} \right] \right)} - 2 \sum_{i=1}^{n} \varpi_\beta(x_{i:n}).
\]

4. Numerical results and discussions

This section consists of the numerical results from Monte Carlo simulations alongside relevant discussions. It is to be noted that all numerical computations reported in this section were obtained using R, an environment for statistical computing (R Core Team 2021).

4.1. Illustrative numerical example

In this section, a random sample with size \(n = 75\) is simulated from CBS(2, 1), and the model parameters are estimated accordingly. To check the existence and uniqueness of the eight estimators, a three-dimensional (3D) plot for the profile of each objective function of each estimation method is established based on extensive Monte Carlo simulations. Mathematically proving the existence and uniqueness of the estimators is beyond the scope of this study. A 3D profile plot can be established by the following steps:

1. Choose computational domains in which the estimates are expected to be within.
2. Randomly simulate a massive number of uniform random variates from these domains.
3. Evaluate the objective function at the generated uniform random variates.
4. Establish a 3D scatter plot such that the \(z\)-axis is the values of the objective function evaluated at the generated uniform random variates.

Figure 3 shows the established 3D charts, which clearly indicate that global extrema exist and are unique except for the PCAs. The profile of the percentile method’s objective function indicates that its optimization might be challenging; thus, one must carefully deal with this method. Table 2 reports the eight estimators for the model parameters, and it indicates that MLEs and CVMEs provided the closest estimators for the true values of the model parameters.

4.2. Simulation outcomes

To numerically examine and compare the performances of the proposed estimators, a bipartite Monte Carlo simulation study is conducted. The first part of the study examines the estimation efficiency, while the second part focuses on goodness-of-fit analyses for the considered estimators. For each part, the sample size is taken to be \(n = 10(10)100\), and the values of the shape parameter \(\alpha = 0.5, 1.0, 2.0, 4.0\). Since \(\beta\) represents the scale parameter, its value is kept fixed at \(\beta = 1.0\), without losing any generality. The simulation outcomes are based on \(M = 10,000\) simulation runs, and are reported in the following subsections.

Estimation efficiency

In this part of the simulation study, the simulated bias and simulated root mean-squared-error (RMSE) are calculated for each estimator according to the combinations of \((n, \alpha, \beta)\) as

\[
\text{Bias } (\hat{\alpha}) = \frac{1}{M} \sum_{i=1}^{M} (\hat{\alpha}_i - \alpha), \quad \text{Bias } (\hat{\beta}) = \frac{1}{M} \sum_{i=1}^{M} (\hat{\beta}_i - \beta),
\]
Table 1: Simulated data from CBS(2,1)

<table>
<thead>
<tr>
<th></th>
<th>MLE</th>
<th>LSE</th>
<th>WLSE</th>
<th>PCE</th>
<th>MPSE</th>
<th>CVME</th>
<th>ADE</th>
<th>RADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.322605</td>
<td>0.236778</td>
<td>48.25141</td>
<td>0.471352</td>
<td>0.686427</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.006556</td>
<td>214.6494</td>
<td>0.761809</td>
<td>185.4404</td>
<td>0.082021</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.63048</td>
<td>1.303483</td>
<td>13.0251</td>
<td>0.750632</td>
<td>0.492427</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.06229</td>
<td>0.115464</td>
<td>0.004807</td>
<td>0.022191</td>
<td>8.4578</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.09539</td>
<td>0.25059</td>
<td>1.167432</td>
<td>821.3805</td>
<td>0.184787</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.531835</td>
<td>54.83529</td>
<td>53.78276</td>
<td>22.89717</td>
<td>2.709339</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.742459</td>
<td>0.396702</td>
<td>1.335046</td>
<td>120.6464</td>
<td>0.257455</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.278792</td>
<td>0.176664</td>
<td>0.510822</td>
<td>25.05264</td>
<td>0.070405</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.29035</td>
<td>152.2387</td>
<td>0.280479</td>
<td>0.285464</td>
<td>0.248384</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.117441</td>
<td>1.003933</td>
<td>664.529</td>
<td>1.33272</td>
<td>0.001952</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.501978</td>
<td>0.013131</td>
<td>1.931444</td>
<td>0.527254</td>
<td>0.097636</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.050339</td>
<td>0.964635</td>
<td>2.609408</td>
<td>0.792273</td>
<td>0.535933</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.386313</td>
<td>11.5986</td>
<td>4.972546</td>
<td>2.462499</td>
<td>5.03194</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>349.1527</td>
<td>2.193183</td>
<td>0.009222</td>
<td>4.05383</td>
<td>1.831104</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29.43906</td>
<td>7.30382</td>
<td>111.2841</td>
<td>0.178484</td>
<td>0.04236</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Estimates for the model parameters using eight methods based on data in Table 1

<table>
<thead>
<tr>
<th>Method</th>
<th>( \hat{\alpha} )</th>
<th>( \hat{\beta} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>2.023782</td>
<td>1.152192</td>
</tr>
<tr>
<td>LSE</td>
<td>2.072389</td>
<td>1.151752</td>
</tr>
<tr>
<td>WLSE</td>
<td>2.047549</td>
<td>1.199592</td>
</tr>
<tr>
<td>PCE</td>
<td>0.235308</td>
<td>27.90558</td>
</tr>
<tr>
<td>MPSE</td>
<td>2.079402</td>
<td>1.165348</td>
</tr>
<tr>
<td>CVME</td>
<td>2.017814</td>
<td>1.148368</td>
</tr>
<tr>
<td>ADE</td>
<td>2.020881</td>
<td>1.199761</td>
</tr>
<tr>
<td>RADE</td>
<td>2.172715</td>
<td>1.168362</td>
</tr>
</tbody>
</table>
Figure 3: 3D plots of the objective functions based on data in Table 1

Figure 3: 3D plots of the objective functions based on data in Table 1 (cont.)
RMSE($\hat{\alpha}$) = $\sqrt{\frac{1}{M} \sum_{i=1}^{M} (\hat{\alpha}_i - \alpha)^2}$ and RMSE($\hat{\beta}$) = $\sqrt{\frac{1}{M} \sum_{i=1}^{M} (\hat{\beta}_i - \beta)^2}$,

such that $\hat{\alpha}_i$ ($\hat{\beta}_i$) is an estimator of the model parameter $\alpha$ ($\beta$) based on simulation run number $i$.

The outcomes of this part of the simulation study is discussed as follows. First and foremost, Table 3 shows the bias and RMSE of the PCEs as well as $D_{\text{abs}}$ and $D_{\text{max}}$ which are two goodness-of-fit criteria that will be defined in the following subsection. It indicates that PCEs of both model parameters were the weakest among their counterparts, which was expected, as previously mentioned. Consequently, the following discussions in this section exclude the outcomes related to PCEs.

Overall, the remaining estimators’ performance was satisfactory since their biases and RMSEs start to decrease as the sample size increases to the point that they become similar, except for the PCEs whose indicators increase as the sample size increases. The simulation outcomes of estimation efficiency for all estimators are provided from Figure 4 to Figure 7. Note that the latter figures consist of log scaled RMSE values for visualization purposes; i.e., the aim of using a log-scale approach is to reduce the potential inflation of some of the values of RMSEs, which are much larger than the remaining bulk of information. From these figures, one can notice the following observations.

Regarding the biases of the estimators of $\alpha$, on the one hand, if $\alpha \leq 1$, then MLEs provide the smallest bias regardless of the sample size. On the contrary, if $\alpha > 1$, RADEs provide the smallest bias for all sample sizes. Regarding the biases of the estimators of $\beta$, both ADEs and CVMEs provided close values for the biases followed by WLSEs for large sample sizes and $1 \leq \alpha \leq 2$. For large sample sizes and $\alpha \geq 4$, MPSEs of $\beta$ provide the smallest biases.

In terms of RMSEs, MLEs of $\alpha$ provided the smallest RMSEs for $\alpha \leq 0.5$ regardless of the sample size. For $\alpha = 1$, the MLEs maintained their performance for $n \leq 30$. If $\alpha = 1$ and $n > 30$, then the CVMEs of $\alpha$ provided the smallest RMSEs. They retained their efficiency for $\alpha > 1$ regardless of the sample size. For the estimators of $\beta$, CVMEs of $\beta$ had the smallest RMSEs given that the sample size was small and $\alpha \leq 2$. For $\alpha = 2$ and large sample sizes, ADEs and WLSEs of $\beta$ had the smallest RMSEs. For the remaining cases, MLEs and MPSEs of $\beta$ yielded the smallest RMSEs.
Table 3: Performance statistics for PCEs of $\alpha$ and $\beta$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n$</th>
<th>Bias (\hat{\alpha})</th>
<th>RMSE (\hat{\alpha})</th>
<th>Bias (\hat{\beta})</th>
<th>RMSE (\hat{\beta})</th>
<th>$D_{abs}$</th>
<th>$D_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>10</td>
<td>0.195</td>
<td>0.832</td>
<td>0.288</td>
<td>9.214</td>
<td>0.110</td>
<td>0.209</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.069</td>
<td>0.837</td>
<td>0.424</td>
<td>3.608</td>
<td>0.119</td>
<td>0.237</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.032</td>
<td>2.284</td>
<td>0.609</td>
<td>3.326</td>
<td>0.139</td>
<td>0.279</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>-0.021</td>
<td>2.651</td>
<td>0.786</td>
<td>3.676</td>
<td>0.161</td>
<td>0.324</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>-0.086</td>
<td>2.695</td>
<td>0.926</td>
<td>3.255</td>
<td>0.182</td>
<td>0.366</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>-0.073</td>
<td>3.405</td>
<td>1.124</td>
<td>4.197</td>
<td>0.205</td>
<td>0.413</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>-0.089</td>
<td>3.261</td>
<td>1.316</td>
<td>3.750</td>
<td>0.218</td>
<td>0.439</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>-0.108</td>
<td>3.522</td>
<td>1.429</td>
<td>3.768</td>
<td>0.231</td>
<td>0.465</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>-0.079</td>
<td>3.711</td>
<td>1.667</td>
<td>4.526</td>
<td>0.240</td>
<td>0.484</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>-0.141</td>
<td>2.997</td>
<td>1.862</td>
<td>4.521</td>
<td>0.254</td>
<td>0.512</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>-0.024</td>
<td>2.208</td>
<td>1.618</td>
<td>14.534</td>
<td>0.124</td>
<td>0.236</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>-0.201</td>
<td>3.066</td>
<td>2.241</td>
<td>10.526</td>
<td>0.151</td>
<td>0.298</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>-0.313</td>
<td>3.127</td>
<td>2.848</td>
<td>10.204</td>
<td>0.185</td>
<td>0.365</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>-0.381</td>
<td>4.200</td>
<td>3.378</td>
<td>8.887</td>
<td>0.222</td>
<td>0.437</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>-0.464</td>
<td>4.525</td>
<td>3.780</td>
<td>8.978</td>
<td>0.258</td>
<td>0.506</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>-0.499</td>
<td>5.753</td>
<td>4.087</td>
<td>10.704</td>
<td>0.295</td>
<td>0.578</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>-0.523</td>
<td>6.245</td>
<td>4.642</td>
<td>11.562</td>
<td>0.325</td>
<td>0.641</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>-0.609</td>
<td>5.619</td>
<td>4.564</td>
<td>10.724</td>
<td>0.342</td>
<td>0.675</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>-0.541</td>
<td>6.294</td>
<td>5.288</td>
<td>12.470</td>
<td>0.355</td>
<td>0.701</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>-0.568</td>
<td>6.021</td>
<td>5.350</td>
<td>12.013</td>
<td>0.367</td>
<td>0.726</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>-0.755</td>
<td>6.956</td>
<td>7.156</td>
<td>38.252</td>
<td>0.156</td>
<td>0.295</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>-1.063</td>
<td>6.279</td>
<td>8.309</td>
<td>31.387</td>
<td>0.203</td>
<td>0.391</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>-1.137</td>
<td>8.378</td>
<td>9.863</td>
<td>23.516</td>
<td>0.242</td>
<td>0.471</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>-0.963</td>
<td>9.509</td>
<td>12.433</td>
<td>28.798</td>
<td>0.266</td>
<td>0.522</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>-0.946</td>
<td>9.836</td>
<td>14.213</td>
<td>31.393</td>
<td>0.286</td>
<td>0.564</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>-0.809</td>
<td>11.106</td>
<td>15.135</td>
<td>28.177</td>
<td>0.300</td>
<td>0.594</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>-0.635</td>
<td>11.898</td>
<td>17.392</td>
<td>31.432</td>
<td>0.312</td>
<td>0.620</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>-0.318</td>
<td>13.608</td>
<td>19.044</td>
<td>32.693</td>
<td>0.322</td>
<td>0.641</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>-0.239</td>
<td>13.801</td>
<td>20.525</td>
<td>33.884</td>
<td>0.327</td>
<td>0.653</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.151</td>
<td>13.862</td>
<td>20.756</td>
<td>32.335</td>
<td>0.333</td>
<td>0.667</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>-2.302</td>
<td>13.431</td>
<td>26.444</td>
<td>146.116</td>
<td>0.197</td>
<td>0.372</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>-2.785</td>
<td>13.234</td>
<td>28.029</td>
<td>67.394</td>
<td>0.243</td>
<td>0.469</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>-2.548</td>
<td>13.503</td>
<td>32.049</td>
<td>68.343</td>
<td>0.269</td>
<td>0.522</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>-1.783</td>
<td>19.667</td>
<td>38.061</td>
<td>73.440</td>
<td>0.287</td>
<td>0.558</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>-1.241</td>
<td>21.738</td>
<td>42.070</td>
<td>79.559</td>
<td>0.296</td>
<td>0.577</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>-0.967</td>
<td>20.259</td>
<td>42.258</td>
<td>78.638</td>
<td>0.294</td>
<td>0.574</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>-0.367</td>
<td>19.391</td>
<td>42.612</td>
<td>71.328</td>
<td>0.290</td>
<td>0.568</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>0.536</td>
<td>22.991</td>
<td>42.417</td>
<td>72.663</td>
<td>0.277</td>
<td>0.542</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>1.218</td>
<td>26.499</td>
<td>44.514</td>
<td>75.835</td>
<td>0.268</td>
<td>0.526</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1.069</td>
<td>21.073</td>
<td>45.360</td>
<td>78.144</td>
<td>0.257</td>
<td>0.505</td>
</tr>
</tbody>
</table>
Figure 4: The simulated biases of the estimators of $\alpha$ (light blue lines represent the smallest values)

Figure 5: The simulated biases of the estimators of $\beta$ (light blue lines represent the smallest values)
Figure 6: The simulated log(RMSEs) of the estimators of $\alpha$ (light blue lines represent the smallest values)

Figure 7: The simulated log(RMSEs) of the estimators of $\beta$ (light blue lines represent the smallest values)
Goodness-of-fit analysis

In this part of the simulation study, the estimation methods are compared according to two simulated goodness-of-fit criteria; namely, the average absolute difference between the true and estimate CDFs ($D_{abs}$), and the maximum absolute difference between the true and estimate CDFs ($D_{max}$). These criteria are computed as

$$D_{abs} = \frac{1}{n \times M} \sum_{i=1}^{M} \sum_{j=1}^{n} \left| F(x_j; \alpha, \beta) - F(x_j; \hat{\alpha}_i, \hat{\beta}_i) \right|,$$

and

$$D_{max} = \frac{1}{M} \sum_{i=1}^{M} \max_{j=1,...,n} \left| F(x_j; \alpha, \beta) - F(x_j; \hat{\alpha}_i, \hat{\beta}_i) \right|,$$

respectively, such that $\hat{\alpha}_i$ ($\hat{\beta}_i$) is an estimator of the model parameter $\alpha$ ($\beta$) based on simulation run number $i$. Recall that the $D_{abs}$ and $D_{max}$ PCEs are reported in Table 3 and their values are unsatisfactory. Now, for each combinations of $(n, \alpha, \beta)$, the above statistics are obtained for the remaining estimators as shown in Figure 8 and Figure 9, respectively. The results of this part of the simulation study is discussed as follows. MPSEs provided the smallest values for $D_{abs}$ and $D_{max}$ for $\alpha \leq 0.5$ regardless of the sample size. In contrast, when $\alpha = 1$, MPSEs yielded the smallest values for $D_{abs}$ for all sample sizes, and the least values of $D_{max}$ given that $n > 20$. For $\alpha = 2$, LSEs provided the optimal values for $D_{abs}$ and $D_{max}$ for each considered sample size. Nevertheless, for $\alpha \geq 4$, ADEs, MLEs, MPSEs, and WLSEs start to compete in terms of the values of $D_{abs}$ and $D_{max}$. Overall, as the sample sizes increases, all methods have $D_{abs}$ and $D_{max}$.

![Figure 8](image_url)

Figure 8: The simulated average absolute difference between the true and estimate CDFs (light blue lines represent the smallest values)
Figure 9: The simulated maximum absolute difference between the true and estimate CDFs (light blue lines represent the smallest values)
5. Applications

This section provides applications to rainfall data to illustrate the considered estimation methods and the practical applicability of the CBS distribution. In practice, researchers sometimes use the Cauchy distribution to fit extreme events such as floods and rainfalls. Recent precipitation data sets are considered for analysis using the CBS distribution due to its close relationship to the Cauchy distribution.

It is important to mention that the Kolmogorov-Smirnov (KS) distance test statistic has been considered as a goodness-of-fit criterion when analyzing the data. Since the parameters are estimated, and ties exist in the data, the corresponding $p$-value is calculated based on 1,000 parametric bootstrap samples. Using the same bootstrap samples, the bias and the RMSE are approximated for the obtained estimates.

5.1. Analyzing rainfalls at Saudi Arabia

The first data set of interest is the annual maximum one-month rainfalls in Saudi Arabia between 2009 and 2019. The data are transformed from millimeters to meters for computational convenience. The data set is as follows: 0.0925, 0.0990, 0.3185, 0.0736, 0.1115, 0.0612, 0.0884, 0.2439, 0.1834, 0.1503, 0.1587. The original data set is published by the Saudi Arabian General Authority for Statistics periodically, and it can be accessed through the KAPSARC data portal.

Table 4 summarizes the estimated model parameters using the discussed eight methods alongside the KS distance statistics and their corresponding $p$-values are calculated. Using the estimators, the CBS distribution’s theoretical percentiles are estimated and compared to their sample counterparts, as shown in Table 5. Moreover, the fitted PDFs are overlayed over the observed histogram of the data, as shown in Figure 10.

From the acquired tables, all methods have close biases and RMSEs. Nevertheless, based on the KS test statistic and its bootstrap $p$-value, the ADEs provided the best-fitted model. The latter conclusion is strengthened by reviewing Figure 10. In terms of the sample percentiles and their theoretical counterparts, PCEs provided the best-fitted model followed by the one fitted by ADEs. Overall, all estimators performed well and have close performance indicators.

Table 4: Outcomes of fitted models based on different estimators alongside the corresponding goodness-of-fit statistics based on rainfall data at Saudi Arabia

<table>
<thead>
<tr>
<th>Method</th>
<th>$\hat{\alpha}$</th>
<th>Bias$_B(\hat{\alpha})$</th>
<th>RMSE$_B(\hat{\alpha})$</th>
<th>$\hat{\beta}$</th>
<th>Bias$_B(\hat{\beta})$</th>
<th>RMSE$_B(\hat{\beta})$</th>
<th>KS $p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>0.332</td>
<td>0.009</td>
<td>0.173</td>
<td>0.111</td>
<td>0.003</td>
<td>0.021</td>
<td>0.191</td>
</tr>
<tr>
<td>LSE</td>
<td>0.405</td>
<td>0.012</td>
<td>0.213</td>
<td>0.121</td>
<td>0.004</td>
<td>0.030</td>
<td>0.167</td>
</tr>
<tr>
<td>WLSE</td>
<td>0.363</td>
<td>0.010</td>
<td>0.190</td>
<td>0.122</td>
<td>0.003</td>
<td>0.026</td>
<td>0.151</td>
</tr>
<tr>
<td>PCE</td>
<td>0.267</td>
<td>0.008</td>
<td>0.138</td>
<td>0.129</td>
<td>0.002</td>
<td>0.019</td>
<td>0.204</td>
</tr>
<tr>
<td>MPSE</td>
<td>0.385</td>
<td>0.011</td>
<td>0.202</td>
<td>0.114</td>
<td>0.004</td>
<td>0.026</td>
<td>0.174</td>
</tr>
<tr>
<td>CVME</td>
<td>0.353</td>
<td>0.010</td>
<td>0.185</td>
<td>0.121</td>
<td>0.003</td>
<td>0.025</td>
<td>0.150</td>
</tr>
<tr>
<td>ADE</td>
<td>0.326</td>
<td>0.009</td>
<td>0.170</td>
<td>0.121</td>
<td>0.003</td>
<td>0.023</td>
<td>0.140</td>
</tr>
<tr>
<td>RADE</td>
<td>0.342</td>
<td>0.010</td>
<td>0.179</td>
<td>0.122</td>
<td>0.003</td>
<td>0.024</td>
<td>0.144</td>
</tr>
</tbody>
</table>
Table 5: Sample percentiles vs. theoretical percentiles estimated using different estimators based on rainfall data at Saudi Arabia

<table>
<thead>
<tr>
<th></th>
<th>$P_{10}$</th>
<th>$P_{25}$</th>
<th>$P_{50}$</th>
<th>$P_{75}$</th>
<th>$P_{90}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>0.074</td>
<td>0.090</td>
<td>0.112</td>
<td>0.171</td>
<td>0.244</td>
</tr>
<tr>
<td>MLE</td>
<td>0.042</td>
<td>0.080</td>
<td>0.111</td>
<td>0.154</td>
<td>0.296</td>
</tr>
<tr>
<td>LSE</td>
<td>0.037</td>
<td>0.081</td>
<td>0.121</td>
<td>0.182</td>
<td>0.394</td>
</tr>
<tr>
<td>WLSE</td>
<td>0.042</td>
<td>0.085</td>
<td>0.122</td>
<td>0.176</td>
<td>0.356</td>
</tr>
<tr>
<td>PCE</td>
<td>0.058</td>
<td>0.099</td>
<td>0.129</td>
<td>0.168</td>
<td>0.287</td>
</tr>
<tr>
<td>MPSE</td>
<td>0.037</td>
<td>0.078</td>
<td>0.114</td>
<td>0.167</td>
<td>0.351</td>
</tr>
<tr>
<td>CVME</td>
<td>0.043</td>
<td>0.085</td>
<td>0.121</td>
<td>0.172</td>
<td>0.341</td>
</tr>
<tr>
<td>ADE</td>
<td>0.046</td>
<td>0.088</td>
<td>0.121</td>
<td>0.168</td>
<td>0.319</td>
</tr>
<tr>
<td>RADE</td>
<td>0.044</td>
<td>0.087</td>
<td>0.122</td>
<td>0.171</td>
<td>0.334</td>
</tr>
</tbody>
</table>

Figure 10: Data histogram vs. fitted PDFs using different estimators based on rainfall data at Saudi Arabia
5.2. Analyzing rainfalls at Florida, USA

The second data set under consideration consists of annual maximum one-day rainfalls at Florida Atlantic University from PRISM Climate Group, Oregon State University. This data set contains 39 data values and it was recently considered by Ball et al. (2021).

Again, point estimates are obtained for the model parameters using the considered estimation method as well as the KS statistics and their \( p \)-values as shown in Table 6. Also, Table 7 gives the estimated theoretical and sample percentiles, while Figure 11 includes the fitted PDFs overlapped the observed histogram of the data.

From the outcomes of the analysis, all methods have close biases and RMSEs except for the PCEs which provided the smallest values. However, based on the KS test statistics and the bootstrap \( p \)-values, the RADEs, ADEs, and CVMEs provided the best estimates. See also Figure 11 to confirm the latter conclusion. PCEs have provided estimated theoretical percentiles close to the sample percentiles. All estimators in general performed similarly.

Table 6: Outcomes of fitted models based on different estimators alongside the corresponding goodness-of-fit statistics based on rainfall data at Florida, USA

<table>
<thead>
<tr>
<th>Method</th>
<th>( \hat{\alpha} )</th>
<th>Bias(_{B}(\hat{\alpha}))</th>
<th>RMSE(_{B}(\hat{\alpha}))</th>
<th>( \hat{\beta} )</th>
<th>Bias(_{B}(\hat{\beta}))</th>
<th>RMSE(_{B}(\hat{\beta}))</th>
<th>KS</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>0.267</td>
<td>0.002</td>
<td>0.064</td>
<td>77.381</td>
<td>0.449</td>
<td>5.041</td>
<td>0.096</td>
<td>0.945</td>
</tr>
<tr>
<td>LSE</td>
<td>0.279</td>
<td>0.002</td>
<td>0.067</td>
<td>79.653</td>
<td>0.492</td>
<td>5.442</td>
<td>0.095</td>
<td>0.959</td>
</tr>
<tr>
<td>WLSE</td>
<td>0.238</td>
<td>0.001</td>
<td>0.057</td>
<td>79.604</td>
<td>0.394</td>
<td>4.624</td>
<td>0.092</td>
<td>0.958</td>
</tr>
<tr>
<td>PCE</td>
<td>0.092</td>
<td>0.001</td>
<td>0.022</td>
<td>85.659</td>
<td>0.130</td>
<td>1.897</td>
<td>0.314</td>
<td>0.546</td>
</tr>
<tr>
<td>MPSE</td>
<td>0.317</td>
<td>0.002</td>
<td>0.076</td>
<td>79.850</td>
<td>0.590</td>
<td>6.207</td>
<td>0.111</td>
<td>0.810</td>
</tr>
<tr>
<td>CVME</td>
<td>0.267</td>
<td>0.002</td>
<td>0.064</td>
<td>79.580</td>
<td>0.461</td>
<td>5.190</td>
<td>0.090</td>
<td>0.990</td>
</tr>
<tr>
<td>ADE</td>
<td>0.248</td>
<td>0.001</td>
<td>0.060</td>
<td>79.745</td>
<td>0.417</td>
<td>4.825</td>
<td>0.086</td>
<td>0.996</td>
</tr>
<tr>
<td>RADE</td>
<td>0.258</td>
<td>0.002</td>
<td>0.062</td>
<td>79.994</td>
<td>0.443</td>
<td>5.046</td>
<td>0.086</td>
<td>0.999</td>
</tr>
</tbody>
</table>

Table 7: Sample percentiles vs. theoretical percentiles estimated using different estimators based on rainfall data at Florida, USA

<table>
<thead>
<tr>
<th></th>
<th>( P_{10} )</th>
<th>( P_{25} )</th>
<th>( P_{50} )</th>
<th>( P_{75} )</th>
<th>( P_{90} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>51.808</td>
<td>61.905</td>
<td>79.85</td>
<td>110.095</td>
<td>133.626</td>
</tr>
<tr>
<td>MLE</td>
<td>34.795</td>
<td>59.314</td>
<td>77.381</td>
<td>100.949</td>
<td>172.088</td>
</tr>
<tr>
<td>LSE</td>
<td>34.556</td>
<td>60.297</td>
<td>79.653</td>
<td>105.222</td>
<td>183.602</td>
</tr>
<tr>
<td>WLSE</td>
<td>35.816</td>
<td>62.753</td>
<td>79.604</td>
<td>109.440</td>
<td>163.253</td>
</tr>
<tr>
<td>PCE</td>
<td>64.631</td>
<td>78.146</td>
<td>85.659</td>
<td>98.394</td>
<td>113.527</td>
</tr>
<tr>
<td>MPSE</td>
<td>31.214</td>
<td>58.260</td>
<td>79.850</td>
<td>109.440</td>
<td>204.267</td>
</tr>
<tr>
<td>CVME</td>
<td>35.753</td>
<td>60.982</td>
<td>79.580</td>
<td>103.850</td>
<td>177.133</td>
</tr>
<tr>
<td>ADE</td>
<td>37.814</td>
<td>62.263</td>
<td>79.745</td>
<td>102.136</td>
<td>168.174</td>
</tr>
<tr>
<td>RADE</td>
<td>36.822</td>
<td>61.819</td>
<td>79.994</td>
<td>103.513</td>
<td>173.783</td>
</tr>
</tbody>
</table>
Figure 11: Data histogram vs. fitted PDFs using different estimators based on rainfall data at Florida, USA
5.3. Analyzing rainfalls at Singapore

The third and final data set contains the highest daily total rainfall for the month recorded at the Changi Climate Station, Singapore between 1982-01-01 and 2021-07-31. The data consists of 475 observations and can be obtained from data.gov.sg website.

Similar to what have been done in the preceding subsections, Table 8 reports the model parameters estimates alongside the KS distance statistics and their corresponding p-values, while Table 9 provides the estimated theoretical percentiles and their sample counterparts. Furthermore, Figure 12 shows the fitted PDFs placed over the observed histogram of the data.

From the previously outcomes, all methods have close biases and RMSEs. Nevertheless, based on Figure 12, and based on each KS test, the WLSEs, ADEs, and ADEs provided the best fitted models. Here, PCEs have provided inaccurate estimated theoretical percentiles compared to the sample percentiles. In general, all estimators performed well except for the PCEs.

Table 8: Outcomes of fitted models based on different estimators alongside the corresponding goodness-of-fit statistics based on rainfall data at Singapore

<table>
<thead>
<tr>
<th>Method</th>
<th>$\hat{\alpha}$</th>
<th>$\text{Bias}_B(\hat{\alpha})$</th>
<th>$\text{RMSE}_B(\hat{\alpha})$</th>
<th>$\hat{\beta}$</th>
<th>$\text{Bias}_B(\hat{\beta})$</th>
<th>$\text{RMSE}_B(\hat{\beta})$</th>
<th>KS</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>0.345</td>
<td>&lt; 0.001</td>
<td>0.022</td>
<td>43.748</td>
<td>0.017</td>
<td>0.975</td>
<td>0.061</td>
<td>0.536</td>
</tr>
<tr>
<td>LSE</td>
<td>0.344</td>
<td>&lt; 0.001</td>
<td>0.022</td>
<td>43.529</td>
<td>0.017</td>
<td>0.966</td>
<td>0.061</td>
<td>0.553</td>
</tr>
<tr>
<td>WLSE</td>
<td>0.293</td>
<td>&lt; 0.001</td>
<td>0.019</td>
<td>43.322</td>
<td>0.013</td>
<td>0.816</td>
<td>0.057</td>
<td>0.728</td>
</tr>
<tr>
<td>PCE</td>
<td>0.016</td>
<td>&lt; 0.001</td>
<td>0.001</td>
<td>43.397</td>
<td>&lt; 0.001</td>
<td>0.046</td>
<td>0.380</td>
<td>0.524</td>
</tr>
<tr>
<td>MPSE</td>
<td>0.365</td>
<td>&lt; 0.001</td>
<td>0.024</td>
<td>43.400</td>
<td>0.019</td>
<td>1.022</td>
<td>0.064</td>
<td>0.494</td>
</tr>
<tr>
<td>CVME</td>
<td>0.342</td>
<td>&lt; 0.001</td>
<td>0.022</td>
<td>43.528</td>
<td>0.017</td>
<td>0.962</td>
<td>0.060</td>
<td>0.557</td>
</tr>
<tr>
<td>ADE</td>
<td>0.323</td>
<td>&lt; 0.001</td>
<td>0.021</td>
<td>43.515</td>
<td>0.016</td>
<td>0.906</td>
<td>0.057</td>
<td>0.701</td>
</tr>
<tr>
<td>RADE</td>
<td>0.319</td>
<td>&lt; 0.001</td>
<td>0.021</td>
<td>43.745</td>
<td>0.016</td>
<td>0.899</td>
<td>0.057</td>
<td>0.721</td>
</tr>
</tbody>
</table>

Table 9: Sample percentiles vs. theoretical percentiles estimated using different estimators based on rainfall data at Singapore

<table>
<thead>
<tr>
<th></th>
<th>$P_{10}$</th>
<th>$P_{25}$</th>
<th>$P_{50}$</th>
<th>$P_{75}$</th>
<th>$P_{90}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>18.8</td>
<td>30.5</td>
<td>43.4</td>
<td>63.25</td>
<td>94.54</td>
</tr>
<tr>
<td>MLE</td>
<td>15.805</td>
<td>31.027</td>
<td>43.748</td>
<td>61.684</td>
<td>121.089</td>
</tr>
<tr>
<td>LSE</td>
<td>15.790</td>
<td>30.918</td>
<td>43.529</td>
<td>61.285</td>
<td>119.997</td>
</tr>
<tr>
<td>WLSE</td>
<td>18.103</td>
<td>32.366</td>
<td>43.322</td>
<td>57.987</td>
<td>103.673</td>
</tr>
<tr>
<td>PCE</td>
<td>41.289</td>
<td>42.701</td>
<td>43.397</td>
<td>44.106</td>
<td>45.614</td>
</tr>
<tr>
<td>MPSE</td>
<td>14.882</td>
<td>30.200</td>
<td>43.400</td>
<td>62.370</td>
<td>126.568</td>
</tr>
<tr>
<td>CVME</td>
<td>15.849</td>
<td>30.959</td>
<td>43.528</td>
<td>61.201</td>
<td>119.547</td>
</tr>
<tr>
<td>ADE</td>
<td>16.708</td>
<td>31.546</td>
<td>43.515</td>
<td>60.024</td>
<td>113.333</td>
</tr>
<tr>
<td>RADE</td>
<td>16.998</td>
<td>31.849</td>
<td>43.745</td>
<td>60.085</td>
<td>112.583</td>
</tr>
</tbody>
</table>
Figure 12: Data histogram vs. fitted PDFs using different estimators based on rainfall data at Singapore
6. Conclusions

In this paper, eight estimators for the model parameters are obtained for the CBS distribution, which is a pathological distribution. Although this means that this distribution violets the assumptions of the central limit theorem, it is expected to have practical application when analyzing data with extremes which are common in financial and climate studies. Numerical results and Monte Carlo simulation outcomes indicated that the choice of the estimator depends on the true value of the shape parameter and the sample size. Nevertheless, all methods performed well except for PCEs, as the sample size increases. The practical application of the considered estimators have been illustrated by analyzing rainfall data. In terms of bootstrap bias and RMSE, the PCEs provided the best fitted model. On the other hand, the ADEs provided the best fitted model for the considered data according to the chosen goodness-of-fit procedure. The analysis also indicated that ADEs and PCEs provided the closet theoretical percentiles to their sample counterparts.

There are three future research direction that are to be considered. First, the behavior of the PDF and HF of the CBS distribution need to be studied from a mathematical perspective. Second, the considered estimators are known as frequentist estimates. Comparing their performances to those of Bayesian counterparts is indeed an important research direction to be addressed. Finally, for any probability distribution, data contamination is expected to negatively affect the performance of some important estimators such as MLEs. Thus, it is of interest to propose nonparametric estimators for the CBS distribution and assess their performances.

References


On comp. diff. methods of est. for the par. of a path. dist. with app. to climate data


**Affiliation:**

Farouq Mohammad A. Alam  
Department of Statistics, Faculty of Science  
King Abdulaziz University  
Jeddah 21589, Kingdom of Saudi Arabia  
E-mail: fmalam@kau.edu.sa  
URL: https://fmalam.kau.edu.sa/Default-0007085-EN