RT&A, No.3 (74)
Volume 18, September 2023

HALF CAUCHY - EXPONENTIAL DISTRIBUTION: ESTIMATION AND APPLICATIONS

K.Jayakumar

•

Department of Statistics, University of Calicut, Kerala -673 635, India
jkumar19@rediffmail.com

Fasna.K

•

Department of Statistics, University of Calicut, Kerala -673 635, India
fasna.asc@gmail.com

Abstract

In this paper, we introduce a new two-parameter distribution called the new Half Cauchy - exponential distribution (HCE) for modeling lifetime data. The structural properties of the new distribution are discussed. Expressions for the quantiles, mode, mean deviation, and distribution of order statistics are derived. The model parameters of HCE distribution are estimated by the method of maximum likelihood, method of least square, method of Cramer-von-Mises, and Anderson-Darling methods. The existence and uniqueness of maximum likelihood estimates are proved. The importance of the new distribution is proved empirically by real-life data set.

Keywords: Half-Cauchy distribution, Method of least-squares, Method of Cramer-von-Mises, Maximum likelihood estimation, T-X family

1. Introduction

The Cauchy distribution named after Augustin Cauchy is a continuous probability distribution and is also known as the Lorentz distribution or Breit-Wigner distribution. It is also the distribution of the ratio of two independent normally distributed random variables with zero mean. Cauchy distribution is unimodal, symmetric, and bell-shaped with much heavier tails than normal distribution. It is also used for the analysis when outliers are presented in the data. Cauchy distribution has received applications in many areas including physics, mathematics, econometrics, engineering, spectroscopy, biological analysis, clinical trials, stochastic modeling of decreasing failure rate life components, queueing theory, and reliability. For more details and discussion, the reader is referred to [20], [3], and [8].
For the Cauchy distribution, the finite moments of order greater than or equal to one do not exist and hence the central limit theorem does not hold. Further, the maximum likelihood estimation of its parameters is not ideal because of no closed-form solution of the likelihood equations. The use of Edgeworth expansion to construct an accurate approximation to the sampling distribution of the maximum likelihood estimator of parameters of Cauchy distribution suggested by [1]. The method of moments is also not possible for this distribution.
Because of these facts, the Cauchy distribution serves as a counterexample for some well-accepted results and concepts in Statistics. This also makes the choice of the Cauchy distribution as an unrealistic model. That is why, modification of Cauchy distribution have been suggested in the

literature to overcome the problem of the moments and other useful properties.
The Half-Cauchy (HC) distribution is the folded standard Cauchy distribution around the origin so that positive values are observed.
A random variable X has the HC distribution with scale parameter $\sigma > 0$, if its cumulative distribution function (cdf) is given by

$$R(x) = \frac{2}{\pi} \arctan\left(\frac{x}{\sigma}\right), \qquad x > 0 \tag{1}$$

The probability density function (pdf) corresponding to 1 is

$$r(x) = \frac{2}{\pi\sigma} \left[1 + \left(\frac{x}{\sigma}\right)^2\right]^{-1}. \tag{2}$$

Although some applications of the half Cauchy distribution exist in the literature, the fact that the finite moments of order greater than or equal to one do not exist, the central limit theorem does not hold. This fact reduces the applicability of this distribution in modeling real-life scenarios. As a heavy-tailed distribution, the HC distribution has been used as an alternative to exponential distribution to model dispersal distances by [4], as the former predicts more frequent long-distance dispersed events than the later. The HC distribution to model ringing data on two species of tits (Parus caeruleus and Parus major) in Britain and Ireland used by [6].

In the real situation, we come across non-normal data sets frequently. One usual way of dealing with non-normal data is to find a suitable transformation that makes the data more normal-like and to apply standard normal-based methods to the transformed data. Finding a suitable transformation can be difficult with data and it is often preferable to work with data without changing the original scale as the easy way of interpretation. These difficulties motivated for more-flexible parametric families of distributions to model non-normal data.

Our focus in this article is on continuous non-normal data. Because real data often deviate from normality in the tails or exhibit asymmetry in the distribution, there has been a growing interest in distributions with additional parameters regulating asymmetry and tails directly. Traditionally, log-normal or gamma distributions are used to model positively skewed data. As a viable and flexible alternative, in this study, we propose Half-Cauchy Exponential (HCE) distribution. In a number of domains such as medical applications, atmospheric sciences, microbiology, environmental science, and reliability theory among others, data are positive and right skewed. The suitable models used by researchers and practitioners to deal with this kind of data are usually parametric distributions such as log-normal, gamma, and Weibull. However, these distributions are not always enough to reach a good fit of the data. This has motivated the interest in the development of more flexible and better-adapted distributions, which have been generated using different strategies as the combination of known distributions.

In the last two decades, there has been an increased interest in defining new generators for univariate continuous distributions to model data in several areas such as engineering, actuarial, medical sciences, biological studies, demography, economics, finance, and insurance. However, in many applied areas like lifetime analysis, finance, and insurance, there is a clear need for extended forms of these distributions, that is, new distributions which are more flexible to model real data. The addition of parameters has been proved useful in exploring skewness and tail properties, and for improving the goodness-of-fit. Thus motivated in to introduce an extended form of HC distribution.
In this paper, we propose a new lifetime model using the technique introduced by [19].
A family of distributions generated by gamma random variables have introduced by [12]. This family of distributions has its cumulative distribution function (cdf) as

$$G(x) = \int_a^{-\ln[1-F(x)]} r(t)dt. \tag{3}$$

Using similiar approach [19] introduced a new family of distributions with cdf given by

$$G(x) = 1 - \int_{a}^{-\ln[F(x)]} r(t)dt. \tag{4}$$

In this paper the T-X family defined by [19] is used to create the half-Cauchy X family of distribution. Let T be a random variable having HC distribution with pdf , $r(t) = \frac{2}{\pi\sigma}\left[1+(\frac{t}{\sigma})^2\right]^{-1}, t > 0$. Then, the pdf of the Half Cauchy - X family of distributions from equation (4) is

$$g(x) = \frac{2f(x)}{\pi\sigma F(x)}\left[1+\left(\frac{-\ln(F(x))}{\sigma}\right)^2\right]^{-1}. \tag{5}$$

The cdf corresponding to (6) is given by

$$G(x) = 1 - \frac{2}{\pi}\arctan\left[\frac{-\ln(F(x))}{\sigma}\right]. \tag{6}$$

One of the main benefits of the Half Cauchy - X family is its ability of fitting skewed data that cannot be properly fitted by existing distributions.

The paper is organized as follows. In Section 2, we proposed Half Cauchy-Exponential (HCE) model, and discuss the shape of the density function and distribution function of the model. We derive the quantiles, mode, and Mean deviation. Analytical shapes of the reliability functions of the model under study and pdf of order statistics and their moments are derived in Section 3. In Section 4, the method of maximum likelihood estimation(MLE), method of least square (LSE), method of Cramer-von-Mises (CVME), and Anderson-Darling methods (ADE) are discussed. We explore the usefulness of the proposed distribution by means of real data set and estimation techniques are applied to calculate the model parameters in Section 5. In Section 6, concluding remarks are presented.

2. Half Cauchy-Exponential Distribution

In this section, we consider the case where f follows exponential distribution with parameter $\theta > 0$ and the cdf and pdf are respectively $F(x) = 1 - e^{-\theta x}$ and $f(x) = \theta e^{-\theta x}$; $x > 0, \theta > 0$.
The cdf and pdf of this new distribution are respectively, given by

$$G(x) = 1 - \frac{2}{\pi}\arctan\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right), \quad x > 0,\ \theta > 0,\ \sigma > 0 \tag{7}$$

and

$$g(x) = \frac{2\theta}{\pi\sigma}\frac{e^{-\theta x}}{1-e^{-\theta x}}\left[1+\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)^2\right]^{-1}. \tag{8}$$

We call this new distribution Half Cauchy Exponential (HCE) distribution with parameters θ and σ. Evidently, the density function (8) does not involve any complicated function. Also, there is no functional relationship between the parameters. We denote the random variable X having pdf (8) as $HCE(\theta, \sigma)$.
The pdf plots of $HCE(\theta, \sigma)$ for various values of the parameters are presented in Figure 1. From the figure, it can be seen that the HCE distribution is well-suited for modelling right-skewed data.
The cdf plots of $HCE(\theta, \sigma)$ for various choices of the values of parameters are presented in Figure 2.

3. Properties of the Half Cauchy - Exponential Distribution

Lemma 1. The q^{th} quantile x_q of the HCE random variable is given by

$$x_q = -\frac{1}{\theta}\ln\left(1 - exp^{-\sigma\tan\left(\frac{\pi(1-q)}{2}\right)}\right). \tag{9}$$

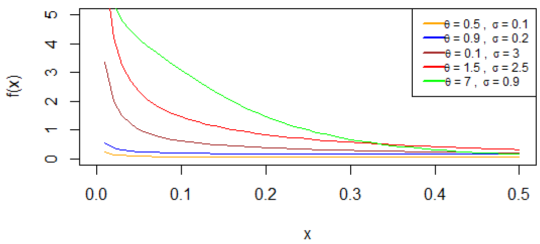


Figure 1: *Plots of the pdf of $HCE(\theta, \sigma)$ distribution.*

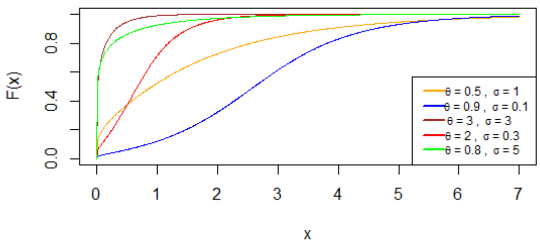


Figure 2: *Plots of the cdf of $HCE(\theta, \sigma)$ distribution.*

Proof. The q^{th} quantile x_q of the HCE random variable is defined as

$$q = P(X \leq x_q) = G(x_q), \quad x_q > 0$$

Using the distribution function of the HCE distribution, we have

$$q = G(x_q) = 1 - \frac{2}{\pi} \arctan\left(\frac{-\ln(1 - e^{-\theta x})}{\sigma}\right)$$

That is,

$$\arctan\left(\frac{-\ln(1 - e^{-\theta x})}{\sigma}\right) = \frac{\pi(1-q)}{2}$$

Hence

$$x_q = -\frac{1}{\theta} \ln\left(1 - exp^{-\sigma \tan\left(\frac{\pi(1-q)}{2}\right)}\right).$$

This completes the proof. ■

Using the usual inverse transformation method, random numbers can be sampled from the proposed model. Let U be a random number drawn from a uniform distribution on (0, 1). Then a random number X following $HCE(\theta, \sigma)$ distribution is obtained by the equation (9) .

In particular, the median is given by,

$$x_{0.5} = -\frac{1}{\theta} \ln(1 - exp^{-\sigma}). \tag{10}$$

Theorem 1. The mode of the $HCE(\theta, \sigma)$ is the solution of the equation $k(x) = 0$, where

$$k(x) = 2e^{-\theta x} \ln(1 - e^{-\theta x}) - \sigma^2 \left[1 + \left(\frac{-\ln(1 - e^{-\theta x})}{\sigma}\right)^2\right].$$

Proof. The critical point of the HCE density function are the roots of the equation:

$$\frac{\partial \log(g(x))}{\partial x} = 0$$

That is

$$\frac{\partial \log(g(x))}{\partial x} = -\theta - \frac{\theta e^{-\theta x}}{1-e^{-\theta x}} + \frac{2\theta e^{-\theta x}\ln(1-e^{-\theta x})}{\sigma(1-e^{-\theta x})\left[1+\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)^2\right]}. \tag{11}$$

The critical values of (11) are the solution of $k(x) = 0$.
Hence the proof. ∎

3.1. Mean Deviation

The mean deviation about the median can be used as a measure of the degree of scattering in a population. Let M be the median of the HCE distribution given by (10).
The mean deviation about the median can be calculated as

$$\delta(X) = E|X-M| = \int_{-\infty}^{\infty} |x-M| g(x)dx,$$

Hence we obtain the following equation $\delta = \mu - 2J(M)$ where J(q) is

$$J(q) = \frac{2\theta}{\pi\sigma}\int_{-\infty}^{q} x\frac{e^{-\theta x}}{1-e^{-\theta x}}\left[1+\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)^2\right]^{-1} dx. \tag{12}$$

One can easily compute this integral numerically in software such as MATLAB, Mathcad, R, and others and hence obtain the mean deviation about the median as desired.

3.2. Stochastic Ordering

Stochastic orders have been used during the last forty years, at an accelerated rate, in many diverse areas of probability and statistics. Such areas include reliability theory, survival analysis, queueing theory, biology, economics, insurance, and actuarial science (see, [5]). Let X and Y be two random variables having cdf's F and G respectively, and denote by $\bar{F} = 1-F$ and $\bar{G} = 1-G$ their respective survival functions, with corresponding pdf's f,g. The random variable X is said to be smaller than Y in the:

1. stochastic order (denoted as $X \leq_{st} Y$) if $\bar{F}(x) \leq \bar{G}(x)$ for all x;
2. likelihood ratio order (denoted as $X \leq_{lr} Y$) if $\frac{f(x)}{g(x)}$ is decreasing in $x \geq 0$;
3. hazard rate order (denoted as $X \leq_{hr} Y$)if $\frac{\bar{F}(x)}{\bar{G}(x)}$is decreasing in $x \geq 0$;
4. reversed hazard rate order (denoted as $X \leq_{rhr} Y$) if $\frac{F(x)}{G(x)}$ is decreasing in $x \geq 0$.

The four stochastic orders defined above are related to each other, have the following implications (see, [5]):

$$X \leq_{rhr} Y \Leftarrow X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{st} Y. \tag{13}$$

The HCE is ordered with respect to the strongest likelihood ratio ordering as shown in the following theorem. It shows the flexibility of the two-parameter HCE distribution.

Theorem 2. Let $X \sim HCE(\theta_1, \sigma_1)$ and $Y \sim HCE(\theta_1, \sigma_1)$. If $\theta_1 = \theta_2 = \theta$ and $\sigma_1 < \sigma_2$; then $X \leq_{lr} Y$ hence $X \leq_{rhr} Y$,$X \leq_{hr} Y$ and $X \leq_{st} Y$.

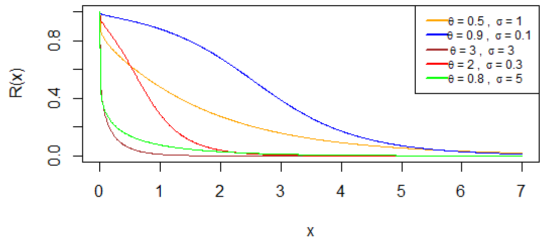


Figure 3: *Plots of reliability function of the $HCE(\theta, \sigma)$ distribution.*

Proof. The likelihood ratio is

$$\frac{g_X(x)}{g_Y(x)} = \frac{\sigma_2}{\sigma_1} \frac{\left[1 + \left(\frac{-\ln(1-e^{-\theta x})}{\sigma_2}\right)^2\right]}{\left[1 + \left(\frac{-\ln(1-e^{-\theta x})}{\sigma_1}\right)^2\right]}$$

Thus,

$$\frac{d}{dx}\left[\frac{g_X(x)}{g_Y(x)}\right] = \frac{2\theta e^{-\theta x}}{1-e^{-\theta x}}\left[\frac{1}{{\sigma_2}^2 + -\ln(1-e^{-\theta x})} - \frac{1}{{\sigma_1}^2 + -\ln(1-e^{-\theta x})}\right]$$

Now, if $\theta_1 = \theta_2 = \theta$ and $\sigma_1 < \sigma_2$, then $\frac{d}{dx}\left[\frac{g_X(y)}{g_Y(y)}\right] < 0$, which implies that $X \leq_{lr} Y$ hence $X \leq_{rhr} Y$,$X \leq_{hr} Y$ and $X \leq_{st} Y$. ■

Lemma 2. If a random variable Y follows the standard exponential distribution, then $X = -\frac{1}{\theta}\ln\left(1 - exp^{-\sigma \tan\left(\frac{\pi(1-e^{-y})}{2}\right)}\right) \sim HCE(\theta, \sigma)$.

3.3. Reliability Analysis

The reliability function is the characterization of an explanatory that maps a set of events, usually associated with the failure of some system onto time. It is the probability that the system will survive beyond a specified time, which is defined by $R(t) = 1 - G(t)$. The Reliability function of $HCE(\theta, \sigma)$ is given by,

$$R(t) = \frac{2}{\pi}\arctan\left(\frac{-\ln(1-e^{-\theta t})}{\sigma}\right). \tag{14}$$

The reliability behaviour of $HCE(\theta, \sigma)$ for various choices of the values of the parameters is presented in Figure 3. The other characteristic of interest of a random variable is the hazard rate function defined by

$$h(t) = \frac{g(t)}{1-G(t)}$$

The hazard rate function of $HCE(\theta, \sigma)$ is given by,

$$h(t) = \frac{\frac{2\theta}{\pi\sigma}\frac{e^{-\theta t}}{1-e^{-\theta t}}\left[1 + \left(\frac{-\ln(1-e^{-\theta t})}{\sigma}\right)^2\right]^{-1}}{\frac{2}{\pi}\arctan\left(\frac{-\ln(1-e^{-\theta t})}{\sigma}\right)}. \tag{15}$$

The behaviour of the hazard rate function of $HCE(\theta, \sigma)$ for various choices of the values of the parameters is presented in Figure 4. The cumulative hazard rate function of HCE distribution,

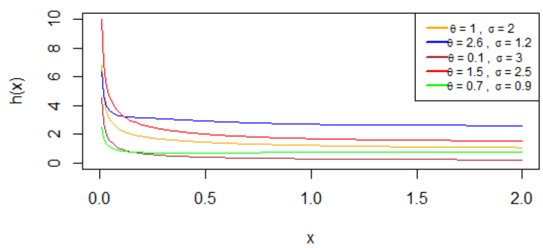


Figure 4: *Plots of hazard rate function of the $HCE(\theta, \sigma)$ distribution.*

$H(t)$ is given by,

$$H(t) = -\ln R(t)$$
$$= -\ln \frac{2}{\pi} \arctan\left(\frac{-\ln(1-e^{-\theta t})}{\sigma}\right). \tag{16}$$

It is important to know that the units for H(t) are the cumulative probability of failure per unit of time, distance, or cycles.

3.4. Order Statistics

Let $X_1, X_2, ..., X_n$ be a random sample from $HCE(\theta, \sigma)$. Also, let $X_{(1)}, X_{(2)}, ..., X_{(n)}$, denote the corresponding order statistics. Then the pdf and cdf of k^{th} order statistics, are given by

$$f_X(x) = \frac{n!}{(k-1)!(n-k)!}[G(x)]^{k-1}[1-G(x)]^{n-k} g(x)$$
$$= \frac{n!}{(k-1)!(n-k)!}\left[1-\frac{2}{\pi}\arctan\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)\right]^{k-1}\left[\frac{2}{\pi}\arctan\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)\right]^{n-k}$$
$$\frac{2\theta}{\pi\sigma}\frac{e^{-\theta x}}{1-e^{-\theta x}}\left[1+\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)^2\right]^{-1} \tag{17}$$

and

$$F_X(x) = \sum_{j=k}^{n}\binom{n}{j}[G(x)]^j[1-G(x)]^{n-j}$$
$$= \sum_{j=k}^{n}\binom{n}{j}\frac{n!}{(k-1)!(n-k)!}\left[1-\left[\frac{2}{\pi}\arctan\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)\right]\right]^j\left[\frac{2}{\pi}\arctan\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)\right]^{n-j} \tag{18}$$

respectively.
The pdf of the minimum is,

$$f_{X_{(1)}}(x) = n\frac{2\theta}{\pi\sigma}\frac{e^{-\theta x}}{1-e^{-\theta x}}\left[\frac{2}{\pi}\arctan\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)\right]^{n-1}\left[1+\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)^2\right]^{-1} \tag{19}$$

and the pdf of the maximum is,

$$f_{X_{(n)}}(x) = n\frac{2\theta}{\pi\sigma}\frac{e^{-\theta x}}{1-e^{-\theta x}}\left[1-\left[\frac{2}{\pi}\arctan\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)\right]\right]^{n-1}\left[1+\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)^2\right]^{-1}. \tag{20}$$

4. Parameter Estimation

In this section, we describe the maximum likelihood estimation procedure, method of least squares, method of Cramer-von-Mises, Anderson-Darling methods, and Method of maximum product of spacings to estimate the parameters θ and σ, in the HCE distribution. We assume throughout that $x_1, x_2, \ldots, x_n$ is a random sample of size n from the HCE distribution both parameters θ and σ are unknown.

4.1. Maximum Likelihood Estimation(MLE)

Here, we consider the estimation of the unknown parameters of the new distribution by the maximum likelihood method. Consider a random sample $(x_1, x_2, \ldots, x_n)$ of size n, from the $HCE(\theta, \sigma)$ distribution. Then, the log likelihood function is given by,

$$\log L = n\log(2\theta) - n\log(\pi\sigma) - \theta\sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\ln(1-e^{-\theta x_i}) - \sum_{i=1}^{n}\left(1+\left(\frac{-\ln(1-e^{-\theta x})}{\sigma}\right)^2\right)$$

The likelihood equations are,

$$\frac{\partial \log L}{\partial\theta} = \frac{n}{\theta} - \sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\frac{x_i e^{-\theta}}{(1-e^{-\theta x_i})} - 2\sum_{i=1}^{n}\frac{x_i e^{-\theta x_i}}{\sigma^2(1-e^{-\theta x_i})}\frac{\ln(1-e^{-\theta x})}{(1+\frac{-\ln(1-e^{-\theta x})}{\sigma})^2}$$
$$= 0, \tag{21}$$

and

$$\frac{\partial \log L}{\partial\sigma} = -\frac{n}{\sigma} - 2\sum_{i=1}^{n}\frac{\ln(1-e^{-\theta x})(1-e^{-\theta x_i})}{\sigma^3(1+\frac{-\ln(1-e^{-\theta x})}{\sigma})^2}$$
$$= 0. \tag{22}$$

These equations do not have explicit solutions and they have to be obtained numerically using statistical software like the *nlm* package in R programming.
If the parameter vector of $HCE(\theta, \sigma)$ be $\Theta = (\theta, \sigma)$ and the associated MLE for Θ is $\hat{\Theta} = (\hat{\theta}, \hat{\sigma})$, then the resulting asymptotic normality is $(\hat{\Theta} - \Theta) \to N(0, (I(\Theta))^{-1})$. Where the observed Fisher's information matrix $(I(\Theta))$ is given by,

$$I(\Theta) \approx \begin{bmatrix} -E(\frac{\partial^2 \log L}{\partial\theta^2}) & -E(\frac{\partial^2 \log L}{\partial\theta\partial\sigma}) \\ -E(\frac{\partial^2 \log L}{\partial\theta\partial\sigma}) & -E(\frac{\partial^2 \log L}{\partial\sigma^2}) \end{bmatrix},$$

and hence the variance covariance matrix would be $I^{-1}(\Theta)$.
As a result of MLEs' asymptotic normality, we may construct approximate $100(1-\alpha)\%$ confidence intervals for θ and σ of $HCE(\theta, \sigma)$ as below:

$$\hat{\theta} \pm Z_{\frac{\alpha}{2}} SE(\hat{\theta}), \hat{\sigma} \pm Z_{\frac{\alpha}{2}} SE(\hat{\sigma})$$

Theorem 3. Let $g_1(\theta; \sigma, x)$ denote the function on the right-hand side (RHS) of equation (21), where σ is the true value of the parameter. Then there exists a unique solution for $g_1(\theta; \sigma, x) = 0$, for $\hat{\theta}\epsilon(0, \infty)$.

Proof. We have

$$g_1(\theta; \sigma, x) = \frac{n}{\theta} - \sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\frac{x_i e^{-\theta}}{(1-e^{-\theta x_i})} - 2\sum_{i=1}^{n}\frac{x_i e^{-\theta x_i}}{\sigma^2(1-e^{-\theta x_i})}\frac{\ln(1-e^{-\theta x})}{(1+\frac{-\ln(1-e^{-\theta x})}{\sigma})^2}$$

Now

$$\lim_{\theta \to 0} g_1(\theta;\sigma,x) = \infty,$$

On the other hand

$$\lim_{\theta \to \infty} g_1(\theta;\sigma,,x) < 0.$$

Therefore there exist at least one root, say $\hat{\theta}\epsilon(0,\infty)$ such that $g_1(\theta;\sigma,x) = 0$
To show uniqueness, the first derivative of $g_1(\theta;\sigma,x) = 0$ is

$$\frac{\partial g_1(\theta;\sigma,x)}{\partial \theta} < 0.$$

Hence there exist a solution for $g_1(\theta;\sigma,,x) = 0$, and root, $\hat{\theta}$ is unique. ■

Theorem 4. Let $g_2(\sigma,\theta,x) = 0$ denote the function on the right hand side (RHS) of equation (22), where θ is the true value of the parameter. Then there exists a unique solution for $g_2(\sigma,\theta,x) = 0$, for $\hat{\sigma}\epsilon(0,\infty)$.

Proof. We have

$$g_2(\sigma,\theta,x) = -\frac{n}{\sigma} - 2\sum_{i=1}^{n} \frac{\ln(1-e^{-\theta x})(1-e^{-\theta x_i})}{\sigma^3(1+\frac{-\ln(1-e^{-\theta x})}{\sigma})^2}.$$

Now

$$\lim_{\sigma \to 0} g_2(\sigma,\theta,x) = -\infty,$$

On the other hand

$$\lim_{\sigma \to \infty} g_2(\sigma,\theta,x) > 0.$$

Therefore there exist atleast one root, say $\hat{\sigma}\epsilon(0,\infty)$ such that $g_2(\sigma,\theta,x) = 0$
To show uniqueness, the first derivative of $g_2(\sigma,\theta,x) = 0$ is

$$\frac{\partial g_2(\sigma,\theta,x)}{\partial \sigma} < 0.$$

Hence there exist a solution for $g_2(\sigma,\theta,x) = 0$, and root, $\hat{\sigma}$ is unique. ■

4.2. Method of Cramer-von Mises

Cramer-von-Mises type minimum distance estimators are based on minimizing the distance between the theoretical and empirical cumulative distribution functions. In [7] provided empirical evidence that the bias of these estimators is smaller than the bias of other minimum distance estimators. The Cramer-von-Mises estimators, $\hat{\theta}_{CME}$ and $\hat{\sigma}_{CME}$ are the values of θ and σ minimizing

$$C(\theta,\sigma) = \frac{1}{12n} + \sum_{i=1}^{n} \left[G(t_i \mid \theta,\sigma) - \frac{2i-1}{2n} \right]^2.$$

Differentiating the above equation partially, with respect to the parameters θ and σ respectively and equating them to zero, we get the normal equations. Since the normal equations are non-linear, we can use iterative method to obtain the solution.

4.3. Method of Anderson-Darling

The method of Anderson-Darling test was developed by [9] as an alternative to statistical tests for detecting sample distributions departure from normality.

The Anderson-Darling estimators $\hat{\theta}_{ADE}$ and $\hat{\sigma}_{ADE}$ are the values of θ and σ minimizes

$$A(\theta,\sigma) = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\{\log G(t_i \mid \theta,\sigma,\theta) - \log \bar{G}(t_{n+1-i} \mid \theta,\sigma)\}.$$

Differentiating the above equation partially, with respect to the parameters θ and σ respectively and them equating to zero, we get the normal equations. Since the normal equations are non-linear, we can use the iterative method to obtain the solution.

4.4. Method of Least-Square Estimation

The least-square estimators were proposed by [2] to estimate the parameters of Beta distributions. Here, we apply the same technique for the HCE distribution. The least-square estimators of the unknown parameters θ and σ of HCE distribution can be obtained by minimizing

$$\sum_{i=1}^{n}\left[G(t_i \mid \theta,\sigma) - \frac{i}{n+1}\right]^2.$$

with respect to unknown parameters θ and σ.

5. Applications

In this section, we have taken two real life data set to illustrate the importance of the proposed distribution. For each data set, we estimate the unknown parameters of each distribution by the maximum-likelihood method, method of least squares, method of Cramer-von-Mises, Anderson-Darling methods, and Method of maximum product of spacings.With these obtained estimates, we obtain the values of the Akaike information criterion (AIC) and Bayesian information criterion (BIC) as well as Kolmogorov-Smirnov statistic and the corresponding p-value. Here, $AIC = -2\ln(L) + 2k$ and $BIC = -2\ln(L) + k\ln(n)$; where L is the likelihood function evaluated at the maximum likelihood estimates, k is the number of parameters and n is the sample size. The K-S distance $D_n = sup_x|F(x) - F_n(x)|$,where,$F_n(x)$ is the empirical distribution. Kolmogorov-Smirnov (K-S) statistics is computed to compare the fitted models.
The required computations are carried out in the R-language introduced by [10].

5.1. Data set I

We consider the corona-virus cases distribution among the fifteen countries viz.,France, Italy, Spain, US, Germony, UK, Turkey, Iran, Russia, China, Brazil, Canada, Belgium, Netherlands and Switzerland. Data has taken from a website and URL is https://www.worldometers.info/coronavirus/coronavirus-cases/.
Data is given in percentage and the observations are:
5.37,6.56,7.61,32.83,5.24,5.06,3.65,3.03 2.89,2.74,2.10,1.57,1.55,1.27,0.97.
The data is skewed-to-the right with skewness =3.0901 and kurtosis =8.4119
The descriptive statistics of the above data set are given in Table 1. The MLEs for θ and σ are listed in Table 2 along with their standard errors (S.E.). The values in Table 3 show that the HCE distribution leads to a better fit for the other three models. Based on the values of the AIC and BIC criteria as well as the value of the KS-statistic and the corresponding p-value, we observe that the HCE distribution provides the best fit for these data among all the models considered. Table 3, it has been observed that the proposed model is best fit as compared to xgamma distribution

Table 1: *The descriptive statistics of Data set.*

Min	1st Q	Median	Mean	3rd Q	Max
0.970	1.835	3.030	5.496	5.305	32.830

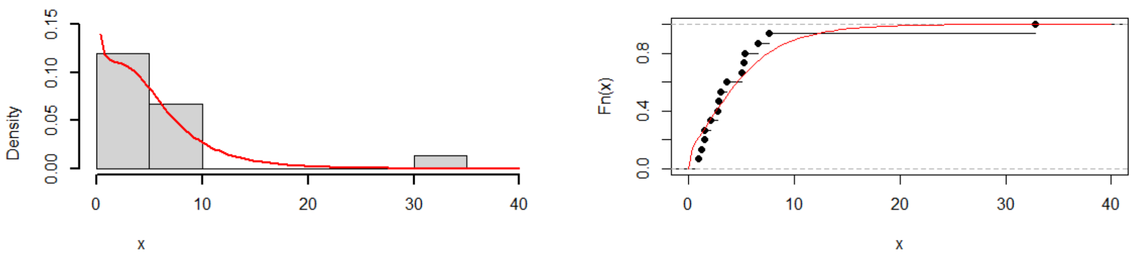
Table 2: *S.E., MLE for θ and σ*

Parameters	MLE	S.E.
θ	0.2359	0.0799
σ	0.5618	0.5618

(XGD) by [11], Akash distribution (AKD) by [21], and exponential power distribution (EPD) by [17].

Figure 5 shows the fitted density curves, Empirical and the fitted cumulative distribution functions for the Data set I.

The goodness-of-fit of the CVME, MLE, LSE, and ADE methods are observed by the test statistic



(a) Fitted density curves for the data set I

(b) Empirical and the fitted cumulative distribution functions for the data set I

Figure 5: *Histogram with fitted pdf's (left) and Empirical cdf with fitted cdf's (right) for the data set I.*

values and their p-values for CVM (Cramer-Von Mises), KS (Kolmogorov-Simnorov), and AD (Anderson-Darling) for the dataset I which are displayed in Table 4.

Figure 6 shows fitted distribution's histogram and the density function having CVME, MLE, LSE, and ADE for the data set I of HCE distribution.

5.2. Data set II

The following data comes from a 59-conductor accelerated life test by [16]. Atomic movement in the circuit's conductors can create failures in microcircuits, which is known as electromigration.

Table 3: *Goodness of fit for various models fitted for the Data set I.*

Model	-LL	AIC	BIC	K-S(p-value)
XGD	-43.555	89.111	89.819	0.2501(0.2585)
AKD	-44.565	91.131	91.839	0.2687(0.1905)
EPD	-42.940	89.881	91.297	0.2471(0.2709)
HCE	-42.435	88.871	90.286	0.2166(0.4224)

Table 4: *Statistics values and their associated p-values for the dataset I.*

Estimation method	Estimates	K-S(p-value)	CVM(p-value)	AD(p-value)
CVME	0.4402 0.2360	0.1398(0.8924)	0.0446(0.9144)	0.9312(0.3937)
MLE	0.2359 0.5618	0.2166(0.4224)	0.1210(0.4962)	1.0133(0.349)
ADE	0.3529 0.3219	0.1619(0.7698)	0.0570(0.84)	0.8506(0.4438)
LSE	0.3918 0.2866	0.1553(0.8101)	0.0480(0.8951)	0.8731(0.4292)

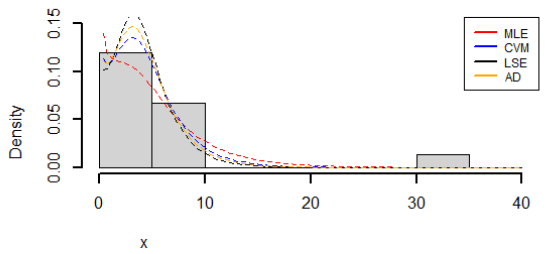


Figure 6: *fitted distribution's histogram and the density function having CVME, MLE, LSE, and ADE for the data set I.*

There are no censored observations, and the failure times are in hours.
5.923, 4.288, 6.522, 4.137, 6.071, 7.495, 6.573, 6.538, 5.589, 6.087, 5.807, 6.725, 8.532, 9.663, 6.545, 10.491, 7.543, 6.956, 6.492, 5.459, 8.120, 4.706, 8.687, 2.997, 8.591, 6.129, 11.038, 5.381, 10.092, 7.496, 4.531, 7.974, 8.799, 7.683, 7.224, 7.365, 6.923, 5.640, 5.434, 7.937, 6.515, 6.476, 6.369, 7.024, 8.336, 9.218, 7.945, 6.869, 6.352, 4.700, 6.948, 9.254, 5.009, 7.489, 7.398, 6.033,7.459, 9.289 , 6.958.
The data is skewed-to-the right with skewness =0.1932 and kurtosis =0.0874
The descriptive statistics of the above data set are given in Table 5. The MLEs for θ and σ are listed in Table 6 along with their standard errors (S.E.). The values in Table 7 show that the *HCE* distribution leads to a better fit for the other five models. Based on the values of the AIC and BIC criteria as well as the value of the KS-statistic and the corresponding p-value, we observe that the HCE distribution provides the best fit for these data among all the models considered. Table 7, it has been observed that the proposed model is best fit as compared to Lindley-Exponential (LE) model by [13], generalized exponential (GE) model by [14], modified Weibull (MW) model by [15], exponential power (EP) model by [17], and Weibull extension (WE) model by [18].

Figure 7 shows the fitted density curves, Empirical and the fitted cumulative distribution functions for the Data set II.

The goodness-of-fit of the CVME, MLE, LSE, and ADE methods are observed by the test statistics values and their p-values for CVM (Cramer-Von Mises), KS (Kolmogorov-Simnorov), and AD (Anderson-Darling) for the dataset II which are displayed in Table 8.

Table 5: *The descriptive statistics of Data set.*

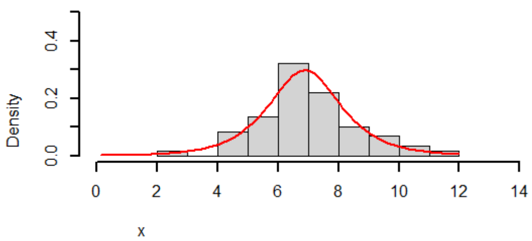
Min	1st Q	Median	Mean	3rd Q	Max
2.997	6.052	6.923	6.980	7.941	11.038

Table 6: *S.E., MLE for θ and σ*

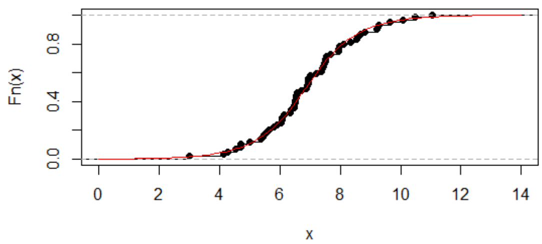
Parameters	MLE	S.E.
θ	0.9352	0.1117
σ	0.0015	0.0012

Table 7: *Goodness of fit for various models fitted for the Data set.*

Model	-LL	AIC	BIC	K-S	p-value
EP	-116.5015	237.0029	237.2098	0.1042	0.5103
LE	-114.9528	233.9055	234.1198	0.1042	0.5099
GE	-114.9473	233.8946	234.1098	0.1365	0.2021
WE	-113.6745	233.3491	233.7855	0.1067	0.4796
MW	-112.5218	231.0435	231.4799	0.0914	0.6738
HCE	-111.7792	227.5584	231.713	0.05806	0.9819



(a) Fitted density curves for the data set II



(b) Empirical and the fitted cumulative distribution functions for the data set II

Figure 7: *Histogram with fitted pdf's (left) and Empirical cdf with fitted cdf's (right) for the data set II.*

Figure 8 shows fitted distribution's histogram and the density function having CVME, MLE, LSE, and ADE for the data set II of HCE distribution.

6. Concluding remarks

In this article, we have introduced and studied a new family of distributions called the Half Cauchy exponential(*HCE*) distribution. we have provided explicit expressions for the quantiles, hazard rates, mean deviation about median, the stochastic ordering and order statistics. The model parameters are estimated by maximum likelihood, least-squares, Cramer-von Mises, and Anderson-Darling. Our formulas related to the *HCE* model are manageable, and with the use of modern computer resources with analytic and numerical capabilities, may turn into adequate tools for a certain purpose of statisticians.

The applicability of the model is demonstrated by using two real data set. From Tables 3 and 7, we observed a better performance of our distribution than the existing distributions. Because HCE distribution has the least test statistic value and the highest p value, we can deduce that it has a considerably better fit than the other distributions studied. Based on these findings, the newly suggested model can be considered as a more efficient, flexible, and therefore may be an alternative to other distributions for modeling positive real data sets. Our proposed model may attract wider applications in survival analysis for modeling positive real data sets. Estimation of

Table 8: *Statistics values and their associated p-values for the dataset II.*

Estimation method	Estimates	K-S(p-value)	CVM(p-value)	AD(p-value)
CVME	0.9055 0.0019	0.0563(0.9867)	0.0233(0.9932)	0.1667(0.997)
MLE	0.9352 0.0015	0.0580(0.9819)	0.0246(0.9908)	0.1798(0.995)
ADE	0.9057 0.0018	0.05598(0.9876)	0.0232(0.9932)	0.1672(0.9969)
LSE	0.8797 0.0022	0.0633(0.9596)	0.0279(0.9832)	0.1800(0.995)

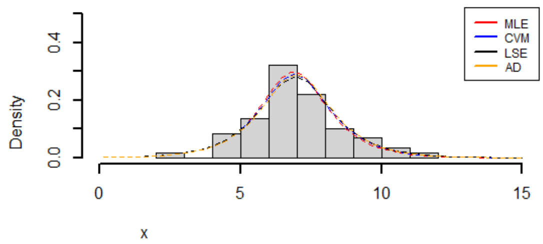


Figure 8: *fitted distribution's histogram and the density function having CVME, MLE, LSE, and ADE for the data set II.*

the model parameters under the Bayesian paradigm is currently underway.

References

[1] Vrbik and Jan. (2011). Sampling distribution of ML estimators: Cauchy example, *Mathematica Journal*, 13:13–19,

[2] Swain, James J and Venkatraman, Sekhar and Wilson, James R. (1988) Least-squares estimation of distribution functions in Johnson's translation system, *Journal of Statistical Computation and Simulation*, 29:271–297.

[3] Krishnamoorthy and Kalimuthu. (2006). Handbook of statistical distributions with applications, *Chapman and Hall/CRC*

[4] Shaw, MW. (1995). Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 259:243–248.

[5] Shaked, Moshe and Shanthikumar, J George. (2007). Stochastic orders, *Springer*

[6] Paradis, Emmanuel and Baillie, Stephen R and Sutherland, William J.(2002) Modeling large-scale dispersal distances, *Ecological Modelling*, 151:279–292.

[7] Macdonald, PDM. (1971). Comments and queries comment on an estimation procedure for mixtures of distributions by choi and bulgren, *Journal of the Royal Statistical Society: Series B (Methodological)*, 33:326–329.

[8] Forbes, Catherine and Evans, Merran and Hastings, Nicholas and Peacock, Brian. (2011). Statistical distributions, author=Forbes, Catherine and Evans, Merran and Hastings, Nicholas and Peacock, Brian, *Wiley New York*,4.

[9] Anderson, Theodore W and Darling, Donald A. (1952). Asymptotic theory of certain" goodness of fit" criteria based on stochastic processes, *The annals of mathematical statistics*, 193–212.

[10] Team, R Develpment Core.(2009) A language and environment for statistical computing, *http://www. R-project. org*.

[11] Sen, Subhradev and Maiti, Sudhansu S and Chandra, N. (2016). The xgamma distribution: statistical properties and application, *Journal of Modern Applied Statistical Methods*, 15:38

[12] Zografos, Konstantinos and Balakrishnan, Narayanaswamy. (2009). On families of beta-and generalized gamma-generated distributions and associated inference, *Statistical methodology*, 6:344–362.

[13] Bhati, Deepesh and Malik, Mohd Aamir and Vaman, HJ. (2015). Lindley–exponential distribution: properties and applications, *Metron*, 73:335–357.

[14] Gupta, Rameshwar D and Kundu, Debasis. (1999). Theory & methods: Generalized exponential distributions, *Australian & New Zealand Journal of Statistics*, 41:173–188.

[15] Lai, CD and Xie, Min and Murthy, DNP. (2003). A modified Weibull distribution, *IEEE Transactions on reliability*, 52:33–37.

[16] Nelson, Wayne and Doganaksoy, Necip. (1995). Statistical analysis of life or strength data from specimens of various sizes using the power-(log) normal model, *Recent Advances in Life-Testing and Reliability*, 377–408.

[17] Smith, Robert M and Bain, Lee J. (1975). An exponential power life-testing distribution, *Communications in Statistics-Theory and Methods*, 4:469–481,

[18] Tang, Y and Xie, M and Goh, TN. (2003). Statistical analysis of a Weibull extension model, *Communications in Statistics-Theory and Methods*, 32:913–928.

[19] Ristić, Miroslav M and Balakrishnan, Narayanaswamy. (2012). The gamma-exponentiated exponential distribution, *Journal of statistical computation and simulation*, 82:1191–1206.

[20] Johnson, Norman L and Kotz, Samuel and Balakrishnan, Narayanaswamy. (1995). Continuous univariate distributions, volume 2, *John wiley & sons*

[21] Shanker, Rama. (2015). Akash distribution and its applications, *International Journal of Probability and Statistics*, 4:65–75.