Half-Cauchy Generalized Exponential Distribution: Theory and Application

Arun Kumar Chaudhary1,∗, Laxmi Prasad Sapkota2, Vijay Kumar3

1 Department of Management Science, Nepal Commerce Campus, Tribhuvan University, Nepal
2 Department of Statistics, Tribhuvan Multiple Campus, Tribhuvan University, Palpa, Nepal
3 Department of Mathematics and Statistics, DDU Gorakhpur University, Gorakhpur, India

*Correspondence to: Arun Kumar Chaudhary, Email: akchaudhary1@yahoo.com

Abstract: Half-Cauchy generalized exponential (HCGE) distribution is a novel distribution that we have proposed on in this paper. The quantiles, the measures of skewness based on quartiles, and the measures of kurtosis based on octiles, survival function, the probability density function, hazard function, cumulative distribution function, cumulative hazard function, are just a few of the crucial statistical properties we have derived for the proposed distribution. To estimate the parameters of the half-Cauchy generalized exponential distribution, the maximum likelihood estimation method has been applied. For the evaluation of the new distribution’s potential, we have considered a real dataset and compared the goodness-of-fit attained by proposed distribution with some competing distribution. The suggested model fits the data much better and is more adaptable than some other models.

Keywords: Estimation, Generalized exponential distribution, Half-Cauchy distribution, Quantile function, Hazard function

DOI: https://doi.org/10.3126/jnms.v5i2.50018

1 Introduction

There are many continuous probability distributions found in literature related to probability and applied statistics. The areas like environmental, actuarial, and medical sciences, economics, life sciences, demography, finance, and insurance where most of the classical distributions have been extensively used for modeling real datasets for many years. However, there is a definite need for customized forms of more adaptable models for simulating real datasets that can handle a significant amount of kurtosis and skewness in a variety of practical areas like banking, survival analysis, and insurance for detail see Galambos and Kotz [11].

By bending the curve at the origin such that only non-negative values can be shown, we have taken into consideration the half-Cauchy distribution that results from the Cauchy distribution in this study. As it can forecast more frequent long-distance spreading events, the heavy-tailed half-Cauchy distribution was utilized by Shaw [25] as a modeling substitute for spreading distances. In order to simulate ringing data on two species in Ireland and Britain, Paradis et al. [19] also employed the half-Cauchy distribution. Assume that \( X \) has a half-Cauchy distribution and is a non-negative random variable. Consequently, its cumulative distribution function (CDF) can be stated as

\[
G(x; \theta) = \frac{2}{\pi} \tan^{-1}\left(\frac{x}{\theta}\right), \quad x > 0, \theta > 0 \tag{1}
\]

and the probability density function (PDF) corresponding to (1) is

\[
g(x; \theta) = \frac{2}{\pi} \left(\frac{\theta}{\theta^2 + x^2}\right), \quad x > 0, \theta > 0 \tag{2}
\]

Last some years many researchers have used the half-Cauchy distribution as a parent model. The beta-half-Cauchy distribution is a modification of the half-Cauchy distribution developed by Cordeiro and Lemonte.
Barreto-Souza et al. [2] have defined a new distribution using generalized exponential distribution called the beta generalized exponential distribution. Similarly, Maiti and Pramanik [16] have presented the odds generalized exponential-exponential distribution. Marshall-Olkin generalized exponential distribution was developed by Ristic and Kundu [23]. Another extension of generalized exponential distribution was presented by Dey et al. [10]. Chaudhary and Kumar [4] have presented logistic modified exponential distribution and further the extension of exponential distribution have been presented by Chaudhary and Kumar [5] called the logistic exponential extension distribution. New exponentiated exponential extension distribution was also presented by Chaudhary and Kumar [6].

The major objective of this work is to improve the GE distribution’s flexibility by simply adding one additional parameter in order to get a satisfactory fit to real datasets. The article’s remaining sections are organized as follows. Half-Cauchy generalized exponential distribution is defined in Section 2, and some of its statistical and mathematical features are covered and examined. Maximum likelihood estimators (MLE) method commonly used estimation technique that is employed in Section 3 to estimate the parameters of the proposed model. Section 4 presents the application of the suggested model. In Section 5, some concluding remarks are given.

2 The Half-Cauchy Generalized Exponential (HCGE) Distribution

We have used the half Cauchy family of distributions to extend the GE distribution in this study. The $T$-family defined by Ristic and Balakrishnan [22] is used to create the half-Cauchy family of distribution and its CDF may be written as

$$F(x) = 1 - \int_0^{-\ln[G(x)]} r(t) \, dt.$$  (5)

In this case, $G(x)$ is the CDF of any baseline distribution and $r(t)$ is the PDF of any distribution. By using $r(t)$ as the PDF of the half-Cauchy distribution specified in (2) as below, we can define the half-Cauchy family of distribution

$$F(x) = 1 - \int_0^{-\ln[G(x)]} \frac{\theta}{2\pi \theta^2 + t^2} dt$$

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

The PDF that corresponds to (6) can be written as

$$f(x) = \frac{2}{\pi \theta G(x)} \left[ 1 + \left\{ -\frac{1}{\theta} \log G(x) \right\}^2 \right]^{-1}.$$  (7)
The CDF and PDF of the HCGE distribution are now obtained by substituting (3) and (4) in (6) and (2), and they are, respectively, as follows

$$F(x) = 1 - \frac{2}{\pi} \arctan \left( -\frac{\alpha}{\theta} \ln \left( 1 - e^{-\lambda x} \right) \right); \quad x > 0, \alpha, \lambda, \theta > 0$$  (8)

$$f(x) = \frac{2\alpha\lambda}{\pi\theta} e^{-\lambda x} \left( 1 - e^{-\lambda x} \right)^{-1} \left[ 1 + \left( -\frac{\alpha}{\theta} \ln \left( 1 - e^{-\lambda x} \right)^2 \right)^{-1} \right]; \quad x > 0, \alpha, \lambda, \theta > 0$$  (9)

It is clear from the Figure[1] that the PDF can have various types of shapes for different values of the parameters.

2.1 Reliability function

The HCGE distribution’s reliability function is

$$R(x) = \frac{2}{\pi} \arctan \left( -\frac{\alpha}{\theta} \ln \left( 1 - e^{-\lambda x} \right) \right); \quad x > 0, \alpha, \lambda, \theta > 0$$  (10)

2.2 Hazard rate function (HRF)

The HCGE’s hazard rate function is

$$h(x) = \frac{\alpha\lambda}{\theta} e^{-\lambda x} \left( 1 - e^{-\lambda x} \right)^{-1} \left[ \arctan \left( -\frac{\alpha}{\theta} \ln \left( 1 - e^{-\lambda x} \right) \right) \right]^{-1} \left[ 1 + \left( -\frac{\alpha}{\theta} \ln \left( 1 - e^{-\lambda x} \right)^2 \right) \right]^{-1}$$  (11)

We have plotted the graph of HRF of HCGE distribution in Figure[2] and found that it can have increasing or j-shaped or bathtub or up-side-down bathtub according to different values of the parameters.

2.3 Reversed hazard rate function (RHRF)

The reversed hazard rate function of HCGE distribution can be defined as

$$R_{HRF} = \frac{2\alpha\lambda}{\pi\theta} e^{-\lambda x} \left( 1 - e^{-\lambda x} \right)^{-1} \left( 1 - \frac{2}{\pi} \arctan \left( -\frac{\alpha}{\theta} \ln \left( 1 - e^{-\lambda x} \right) \right) \right)^{-1} \left[ 1 + \left( -\frac{\alpha}{\theta} \ln \left( 1 - e^{-\lambda x} \right)^2 \right) \right]^{-1}$$  (12)

Figure[3] shows the multiple PDF forms, we have plotted for varying HCGE distribution parameter values. We have also plotted the various shapes of HRF for different values of the parameters of HCGE distribution in Figure[2].

2.4 Cumulative hazard function (CHF)

The suggested model’s CHF is defined as

$$H(x) = \int_{-\infty}^{x} h(t) dt$$

$$= -\log \left[ 1 - F(x) \right]$$

$$= -\log \left[ \frac{2}{\pi} \arctan \left( -\frac{\alpha}{\theta} \ln \left( 1 - e^{-\lambda x} \right) \right) \right]$$  (13)

2.5 Quantile function of the HCGE distribution

The quantile function can be obtained by inverting the CDF defined in (8) as

$$Q(u) = F^{-1}(u).$$
Figure 1: The graphs of PDF for different values of $\alpha$, $\lambda$ and $\theta$.

Hence we can expressed the quantile function as

$$Q(u) = -\frac{1}{\lambda} \ln \left[ 1 - \exp \left\{ -\frac{\theta}{\alpha} \tan \left( \frac{\pi (1 - u)}{2} \right) \right\} \right] ; \quad u \in (0, 1) \quad (14)$$

Here $u$ is the random variable that follows $U(0,1)$. To generate the random numbers for HCGE distribution, we can utilize the following expression

$$x = -\frac{1}{\lambda} \ln \left[ 1 - \exp \left( -\frac{\theta}{\alpha} \tan \left( \frac{\pi (1 - v)}{2} \right) \right) \right] ; \quad v \in (0,1).$$

2.6 Skewness and kurtosis for HCGE distribution

**Skewness:** The Bowley’s coefficient of skewness based on quantiles can be achieved as

$$S(B) = \frac{Q(0.75) + Q(0.25) - 2Q(1/2)}{Q(3/4) - Q(0.25)}.$$

**Kurtosis:** Moors [17] used octiles to define the coefficient of kurtosis as

$$K_u(M) = \frac{Q(0.875) + Q(0.375) - Q(0.625) - Q(0.125)}{Q(3/4) - Q(1/4)}.$$

3 Parameter Estimation

3.1 Maximum likelihood estimation (MLE)

Here, we have discussed the ML estimators (MLE’s) of the HCGE model are estimated by using MLE method. Let $x = (x_1, \ldots, x_n)$ be a random sample drawn from a population of size $n$ follows $HCGE(\alpha, \lambda, \theta)$. Then, the log likelihood function is

$$\ell(\alpha, \lambda, \theta | x) = n \ln(2/\pi) + n \ln \left( \frac{\alpha \lambda}{\theta} \right) - \lambda \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \ln C(x_i) - \sum_{i=1}^{n} \ln \left[ 1 + \left( -\frac{\alpha}{\theta} \ln C(x_i) \right)^2 \right] \quad (15)$$
Figure 2: The graphs of HRF for different values of α, λ and θ.

Now, differentiating (15) with respect to α, λ and θ, we get

\[
\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} - \frac{2\alpha}{\theta^2} \sum_{i=1}^{n} \left[ \ln C(x_i) \right]^2 \left[ 1 + \left( -\frac{\alpha}{\theta} \ln C(x_i) \right) \right]^{-1}
\]

\[
\frac{\partial \ell}{\partial \lambda} = -\frac{n}{\lambda} - \sum_{i=1}^{n} x_i e^{-\lambda x_i} C(x_i) - 2 \left( \frac{\alpha}{\theta} \right)^2 \sum_{i=1}^{n} x_i e^{-\lambda x_i} C(x_i)^{-1} \ln C(x_i) \left[ 1 + \left( -\frac{\alpha}{\theta} \ln C(x_i) \right) \right]^{-1}
\]

\[
\frac{\partial \ell}{\partial \theta} = -\frac{n}{\theta} + 2\frac{\alpha^2}{\theta^3} \sum_{i=1}^{n} \left[ C(x_i) \ln C(x_i) \left[ 1 + \left( -\frac{\alpha}{\theta} \ln C(x_i) \right) \right]^{-1} \right],
\]

where \( C(x_i) = 1 - e^{-\lambda x_i} \). The ML estimators for the \( HCGE(\alpha, \lambda, \theta) \) model are obtained by solving for α, λ and θ, using the three non-linear equations \( \frac{\partial \ell}{\partial \alpha} = \frac{\partial \ell}{\partial \lambda} = \frac{\partial \ell}{\partial \theta} = 0 \). But normally, it is not possible to solve non-linear equations above, so with the aid of a suitable piece of software, one can solve them easily. If the parameter vector of \( HCGE(\alpha, \lambda, \theta) \) is denoted by \( \Theta = (\alpha, \lambda, \theta) \) and the associated MLE of \( \Theta \) as \( \hat{\Theta} = (\hat{\alpha}, \hat{\lambda}, \hat{\theta}) \) then, the resulting asymptotic normality is, \( (\hat{\Theta} - \Theta) \rightarrow \mathcal{N}_3 \left( 0, (I(\Theta))^{-1} \right) \). The Fisher’s information matrix, denoted by \( I(\Theta) \) here, is provided by

\[
I(\Theta) = -\begin{pmatrix}
E\left( \frac{\partial^2 l}{\partial \alpha^2} \right) & E\left( \frac{\partial^2 l}{\partial \alpha \partial \lambda} \right) & E\left( \frac{\partial^2 l}{\partial \alpha \partial \theta} \right) \\
E\left( \frac{\partial^2 l}{\partial \lambda \partial \alpha} \right) & E\left( \frac{\partial^2 l}{\partial \lambda^2} \right) & E\left( \frac{\partial^2 l}{\partial \lambda \partial \theta} \right) \\
E\left( \frac{\partial^2 l}{\partial \theta \partial \alpha} \right) & E\left( \frac{\partial^2 l}{\partial \theta \partial \lambda} \right) & E\left( \frac{\partial^2 l}{\partial \theta^2} \right)
\end{pmatrix}.
\]

The asymptotic variance \( (I(\Theta))^{-1} \) of the MLE has no practical significance because we do not know what \( \Theta \) is. The estimated parameter values are therefore plugged into the approximate asymptotic variance. The information matrix \( I(\Theta) \) provided by the observed Fisher information matrix \( \hat{\Theta} \) is employed as a

\[
O(\hat{\Theta}) = -\begin{pmatrix}
\frac{\partial^2 l}{\partial \alpha^2} & \frac{\partial^2 l}{\partial \alpha \partial \lambda} & \frac{\partial^2 l}{\partial \alpha \partial \theta} \\
\frac{\partial^2 l}{\partial \lambda \partial \alpha} & \frac{\partial^2 l}{\partial \lambda^2} & \frac{\partial^2 l}{\partial \lambda \partial \theta} \\
\frac{\partial^2 l}{\partial \theta \partial \alpha} & \frac{\partial^2 l}{\partial \theta \partial \lambda} & \frac{\partial^2 l}{\partial \theta^2}
\end{pmatrix}_{(\alpha, \lambda, \theta)} = -H(\Theta)|_{(\alpha, \lambda, \theta)}.
\]
Here, $H$ denotes the Hessian matrix. The observed information matrix is built using the Newton-Raphson method to maximize likelihood, and the resulting variance-covariance matrix is as follows

$$
\left[ -H (\Theta) \right]^{-1} = \begin{pmatrix}
\text{var}(\hat{\alpha}) & \text{cov}(\hat{\alpha}, \hat{\lambda}) & \text{cov}(\hat{\alpha}, \hat{\theta}) \\
\text{cov}(\hat{\alpha}, \hat{\lambda}) & \text{var}(\hat{\lambda}) & \text{cov}(\hat{\lambda}, \hat{\theta}) \\
\text{cov}(\hat{\alpha}, \hat{\theta}) & \text{cov}(\hat{\lambda}, \hat{\theta}) & \text{var}(\hat{\theta})
\end{pmatrix}
$$

Let $Z_{c/2}$ denotes the upper percentile of the standard variate. So, using the asymptotic normality of MLEs, the following approximations of $100(1-c)\%$ confidence intervals for $\alpha$, $\lambda$ and $\theta$ can be created:

$$
\hat{\alpha} \pm Z_{c/2} \sqrt{\text{var}(\hat{\alpha})}, \quad \hat{\lambda} \pm Z_{c/2} \sqrt{\text{var}(\hat{\lambda})}, \quad \hat{\theta} \pm Z_{c/2} \sqrt{\text{var}(\hat{\theta})}.
$$

## 4 Application to Real Dataset

In this part, we have used the real data set from a test that involved accelerated life for 59 conductors performed by Nelson and Doganaksoy [18]. Electro-migration, or the movement of atoms within the conductors of a circuit, is a cause of failures in microcircuits. No observations have been censored, and failure times are given in hours. The R software’s `optim()` function is used to determine the MLEs of the proposed distribution for detail see [15, 21]. The log likelihood value that we have determined is $l = -111.7792$. For the parameters $\alpha$, $\lambda$ and $\theta$, we have shown the MLEs in Table 1 together with their standard errors (SE). The ML estimates are generated exclusively, as shown by the graphs of the profile log-likelihood function in Figure 3 for $\alpha$, $\lambda$ and $\theta$. It is evident from the $P-P$ and $Q-Q$ plots in Figure 4 that the suggested model closely matches the data. Using a real dataset that was previously utilized by researchers, we have demonstrated in this section the goodness-of-fit for the half-Cauchy generalized exponential model.

We have chosen the five distributions such as Exponential power (EP) distribution by Smith and Bain [26],

<table>
<thead>
<tr>
<th>Parameter</th>
<th>MLE</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>alpha</td>
<td>6.61405</td>
<td>5.48858</td>
</tr>
<tr>
<td>lambda</td>
<td>0.9352</td>
<td>0.11109</td>
</tr>
<tr>
<td>theta</td>
<td>0.01028</td>
<td>0.01178</td>
</tr>
</tbody>
</table>

Table 1: SE and MLE for $\alpha$, $\lambda$ and $\theta$ of HCGE distribution

Figure 3: Graphs of Profile log-likelihood function for $\alpha$, $\lambda$ and $\theta$. 
Figure 4: The $P$-$P$ (left panel) and $Q$-$Q$ (right panel) plots for the HCGE distribution.

Figure 5: Empirical distribution function with estimated distribution function (right panel) and the Histogram and the density function of fitted distributions (left panel).

Generalized Exponential (GE) distribution by Gupta and Kundu [12], Modified Weibull (MW) by Lai et al. [14], Weibull Extension Model by Tang et al. [27] and Power Cauchy distribution by Rooks et al. [24] to contrast the recommended model’s potential. We have calculated the corrected Akaike information criterion (CAIC), the Hannan-Quinn information criterion (HQIC) and the Bayesian information criterion (BIC) in order to assess the adequacy of the $HCGE(\alpha, \lambda, \theta)$ model. Table 2 displays these results. Figure 5 displays the histogram along with a few selected distributions, as well as the density functions of the fitted distributions, the empirical distribution function, and the estimated distribution function. Table 3 displays the results of the Cramer-Von Mises ($CVM$), Kolmogorov-Smirnov ($KS$), and Anderson-Darling
Table 2: AIC, CAIC, BIC, HQIC, and Log-likelihood (LL)

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>BIC</th>
<th>CAIC</th>
<th>HQIC</th>
<th>LL</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCGE</td>
<td>229.56</td>
<td>235.79</td>
<td>229.99</td>
<td>231.99</td>
<td>-111.78</td>
</tr>
<tr>
<td>PC</td>
<td>228.18</td>
<td>232.33</td>
<td>228.39</td>
<td>229.80</td>
<td>-112.09</td>
</tr>
<tr>
<td>MW</td>
<td>231.04</td>
<td>237.27</td>
<td>231.48</td>
<td>233.47</td>
<td>-112.52</td>
</tr>
<tr>
<td>WE</td>
<td>233.34</td>
<td>239.58</td>
<td>233.78</td>
<td>235.78</td>
<td>-113.67</td>
</tr>
<tr>
<td>GE</td>
<td>233.89</td>
<td>238.05</td>
<td>234.10</td>
<td>235.51</td>
<td>-114.95</td>
</tr>
<tr>
<td>EP</td>
<td>237.00</td>
<td>241.15</td>
<td>237.21</td>
<td>238.62</td>
<td>-116.5</td>
</tr>
</tbody>
</table>

(AD) statistics in order to compare the goodness-of-fit of the suggested distribution to that of competing distributions. We draw the conclusion that the HCGE distribution has a much better fit to the data and more consistent and trustworthy findings when compared to others because it has a higher p-value and the lowest value of the test statistic.

Table 3: The goodness-of-fit statistics and the p-value

<table>
<thead>
<tr>
<th>Model</th>
<th>KS(p-value)</th>
<th>AD(p-value)</th>
<th>CVM(p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCGE</td>
<td>0.0580</td>
<td>0.0247</td>
<td>0.1799</td>
</tr>
<tr>
<td>PC</td>
<td>0.0480</td>
<td>0.0199</td>
<td>0.1780</td>
</tr>
<tr>
<td>MW</td>
<td>0.0914</td>
<td>0.0821</td>
<td>0.4839</td>
</tr>
<tr>
<td>WE</td>
<td>0.1067</td>
<td>0.1154</td>
<td>0.6800</td>
</tr>
<tr>
<td>GE</td>
<td>0.1042</td>
<td>0.1173</td>
<td>0.7368</td>
</tr>
<tr>
<td>EP</td>
<td>0.1365</td>
<td>0.2398</td>
<td>1.3735</td>
</tr>
</tbody>
</table>

5 Conclusions

We have introduced a novel continuous distribution in this paper termed the half-Cauchy generalized exponential distribution. Along with a thorough consideration of some of the new distribution’s mathematical and statistical characteristics, the exact equations for the skewness, and kurtosis, quantile function, survival function, hazard function, cumulative hazard function are addressed and examined. The most often used estimating approach, MLE method, is used to calculate the model parameters. Testing the appropriateness and application of the suggested distribution on a real data set revealed that it is far more flexible than some chosen distributions.

References


