Half Cauchy Exponential Geometric Distribution: Model, Properties and Application

Lal Babu Sah Telee [*1]
Murari Karki [*2]
Vijay Kumar [*3]

Article History: *Received 28 Sep. 2021; Reviewed 2 Oct. 2021; Revised 25 Nov. 2021; Accepted 28 Dec. 2021.*

Abstract

In this article, a new distribution called Half Cauchy Exponential Geometric Distribution is introduced. We have derived some important mathematical properties like hazard function, probability density function, survival function, quantiles, the measures of skewness based on quartiles and coefficient of kurtosis based on octiles of the new distribution. To estimate the parameters of the new distribution, we have applied the three commonly used estimation methods namely maximum likelihood estimation (MLE), least-square (LSE) method and Cramer-Von-Mises (CVM) method. We have used R programming as well as analytical methods for data analysis. For the assessment of potentiality of the new distribution, we have considered a real dataset and the goodness-of-fit attained by proposed distribution is compared with some competing distributions.It is found that the proposed model fits the data very well and more flexible as compared to some other models.

Keywords: *Survival function, Maximum lilkelihood Estimation, Quantile function, Hazard function.*

Introduction

Exponential–Geometric distribution proposed by Adamidis and Loukas[A lifetime distribution with decreasing failure rate, Statist. Probab.Lett. 39 (1998), pp. 35–42] accommodate unimodal hazard function with increasing and decreasing failure rate. This distribution explains that lifetime associated with a particular risk is not observable but only the minimum lifetime value among all the risks. Let x be a non-negative continuous random variable denoting the lifetime of a component, then the random variable x is said to have an Exponential Geometric (EG) distribution [E2G, Francisco Lauzada, Vitor A. A. Marchi and Mari Roman] if its probability distribution function is given as

$$G(x,\alpha,\lambda)=\frac{1-e^{-\lambda x}}{1-(1-\alpha)e^{-\lambda x}} \tag{1}$$

[*1] Assistant Professor, Deprtment of Statistics, Nepal Commerce campus, TU, Nepal
[*2] Assistant Professor, Department of Statistics, Sarsawati Multiple Campus, TU, Nepal
[*3] Department of Mathematics and Statistics, DDU Gorakhpur , Corresponding Email: karkimk@gmail.com

where α is shape parameter and λ is scale parameter. PDF of above distribution is given as

$$g(x,\alpha,\lambda)=\frac{\lambda\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)^2} \tag{2}$$

Here we are interested to extend EG distribution using half Cauchy family of distribution. Let X is a positive random variable that follows the half-Cauchy distribution with CDF as

$$G(x;\theta)=\frac{2}{\pi}\tan^{-1}\left(\frac{x}{\theta}\right),\ \ x>0,\ \ \theta>0. \tag{3}$$

The probability density function (PDF) of half Cauchy distribution is

$$g(x;\theta)=\frac{2}{\pi}\left(\frac{\theta}{\theta^2+x^2}\right),\ \ x>0,\quad \theta>0. \tag{4}$$

The extending family of distribution has developed by *Zografas&Balakrishnan*, 2009 [2] and CDF of family of distribution is

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

Here $G(x)$ is the CDF of any baseline distribution and $r(t)$ is the PDF of any distribution given as

$$r(t)=\frac{2}{\pi}\frac{\theta}{\left(\theta^2+t^2\right)};\quad \theta>0$$

The family of half-Cauchy distribution with cumulative distribution function $F(x)$ can be obtained by using $r(t)$ in equation (5) is expressed as

$$F(x)=\int_0^{-\ln[1-G(x)]}\frac{2}{\pi}\frac{\theta}{\theta^2+t^2}\ dt=\frac{2}{\pi}\arctan\left(-\frac{1}{\theta}\ln\left[1-G(x)\right]\right);\ x>0,\theta>0 \tag{6}$$

The PDF corresponding to (6) can be expressed as

$$f(x)=\frac{2}{\pi\theta}\frac{g(x)}{1-G(x)}\left[1+\left\{-\frac{1}{\theta}\log\left[1-G(x)\right]\right\}^2\right]^{-1};x>0,\theta>0 \tag{7}$$

The inspiration of this study is to put forward a more flexible distribution by inserting just one extra parameter to the Exponential Geometric distribution to achieve a better fit to the real data.

We study the properties of the half Cauchy Exponential Geometric distribution and explore its potentiality and applicability.

The contents of this paper are managed as follows. The new half Cauchy Exponential Geometric distribution is introduced and several distributional properties are discussed in Section 2. Three mostly used estimation approaches are used to estimate the parameters namely maximum likelihood estimation (MLE), least-square estimation (LSE), and Cramer-Von-Mises estimation (CVM) methods are presented in Section 3. In Section 4, a real life dataset have been considered to investigate the applications and suitability of the proposed distribution. In this section, we have calculated the approximate confidence intervals of the ML estimators of the parameters and also AIC, AICC, BIC, HQIC are calculated to evaluate the goodness-of-fit of the half Cauchy Exponential Geometric distribution. Finally, some concluding remarks are presented.

Half Cauchy Exponential Geometric (HCEG) distribution

In this section the new distribution named half Cauchy Exponential Geometric distribution is defined. Substituting (1) and (2) in (6) and (7) we get the CDF and PDF of HCEG distribution. Let X be a non negative random variable following half Cauchy Exponential Geometric distribution $HCEG(\alpha,\lambda,\theta)$ if its CDF can be written as,

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

And PDF corresponding to (8) can be written as,

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

Reliability/survival function:

The reliability function of $HCEG(\alpha,\lambda,\theta)$ distribution is defined as

$$R(x)=1-F(x)=1-\frac{2}{\pi}\arctan\left[-\frac{1}{\theta}\log\left\{\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right\}\right] \qquad (10)$$

Hazard rate function (HRF):

The **HRF** of $HCEG(\alpha,\lambda,\theta)$ distribution can be defined as,

$$h(x)=\frac{f(x)}{R(x)}=\frac{2\lambda}{\pi\theta\left(1-(1-\alpha)e^{-\lambda x}\right)}\left[1+\left\{-\frac{1}{\theta}\log\left\{\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right\}\right\}^2\right]^{-1}$$

$$\left[1-\frac{2}{\pi}\arctan\left[-\frac{1}{\theta}\log\left\{\frac{\alpha e^{-\lambda x}}{\left(1-\left(1-\alpha\right)e^{-\lambda x}\right)}\right\}\right]\right]^{-1} \tag{11}$$

Quantile function:

The quantile function of $HCEG(\alpha,\lambda,\theta)$ can be expressed as,

$$Q(p) = \quad -\frac{1}{\lambda}\ln\left[\left(e^{-\theta\tan\left(\frac{\pi p}{2}\right)}\right)\left\{\alpha+\left(1-\alpha\right)e^{-\theta\tan\left(\frac{\pi p}{2}\right)}\right\}^{-1}\right];\;\; 0<p<1\cdot \tag{12}$$

The Random Deviate Generation:

The random numbers can be drawn from $HCEG(\alpha,\lambda,\theta)$ by

$$x = -\frac{1}{\lambda}\ln\left[\left(e^{-\theta\tan\left(\frac{\pi u}{2}\right)}\right)\left\{\alpha+\left(1-\alpha\right)e^{-\theta\tan\left(\frac{\pi u}{2}\right)}\right\}^{-1}\right];\quad 0<u<1 \tag{13}$$

Plot of PDF and HRF of $HCEG(\alpha,\lambda,\theta)$ are displayed in Figure 1.

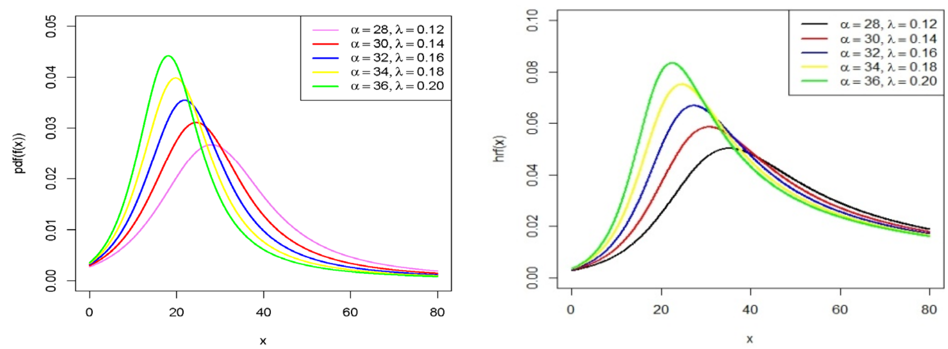

Figure 1:*Plots of pdf (left panel) and HRF (right panel) for different values of α and θ for λ=1.*

Skewness and Kurtosis of HCEG distribution:

Skewness and kurtosis are the measure that describes the nature of distribution .Bowely'sskewness of the $HCEG(\alpha,\lambda,\theta)$ distribution based on quartiles has form

$$Sk(B) = \frac{Q(3/4) - 2Q(1/2) + Q(1/4)}{Q(3/4) - Q(1/4)}, \tag{14}$$

Kurtosis of the $HCEG(\alpha, \lambda, \theta)$ distribution based on octiles has form (Moors, 1988).

$$K(moors) = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)}, \tag{15}$$

Methodology

To estimate the parameters of the new distribution, we have applied the three commonly used estimation methods namely maximum likelihood estimation (MLE), least-square (LSE) method and Cramer-Von-Mises (CVM) method. We have used R programming as well as analytical methods for data analysis.

Maximum Likelihood Estimation (MLE)

In this section, we have presented the ML estimators (M.L.E.'s) of the HCEG distribution. Let $\underline{x} = (x_1, \ldots, x_n)$ be a random sample of size 'n' from $HCEG(\alpha, \lambda, \theta)$ then the log likelihood function can be written as,

$$\begin{aligned} \ell(\alpha, \lambda, \theta \mid \underline{x}) = n \ln\left(\frac{2}{\pi}\right) + n \ln \lambda - n \ln \theta - \sum_{i=1}^{n} \ln\left\{1 - (1-\alpha) e^{-\lambda x}\right\} \\ - \sum_{i=1}^{n} \left[1 - \left\{ -\frac{1}{\theta} \ln \left\{ \frac{\alpha e^{-\lambda x}}{\left(1 - (1-\alpha) e^{-\lambda x}\right)} \right\} \right\}^2 \right] \end{aligned} \tag{16}$$

After differentiating (16) with respect to parameters α, λ and θ, we get

$$\frac{\partial \ell}{\partial \alpha} = -\sum_{i=1}^{n} \left\{1 - (1-\alpha) e^{-\lambda x}\right\}^{-1} - \sum_{i=1}^{n} \left[\frac{2}{\theta} \left\{ \frac{\alpha e^{-\lambda x}}{\left(1 - (1-\alpha) e^{-\lambda x}\right)} \right\}^{-1} \left\{ \frac{e^{-\lambda x}\left(1 - e^{-\lambda x}\right)}{\left(1 - (1-\alpha) e^{-\lambda x}\right)^2} \right\} \right] \tag{17}$$

$$\begin{aligned} \frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} - (1-\alpha) \sum_{i=1}^{n} x e^{-\lambda x} \left\{1 - (1-\alpha) e^{-\lambda x}\right\}^{-1} + \frac{2}{\theta^2} \sum_{i=1}^{n} \left\{ \frac{\alpha e^{-\lambda x}}{\left(1 - (1-\alpha) e^{-\lambda x}\right)} \right\}^{-1} \\ \ln \left\{ \frac{\alpha e^{-\lambda x}}{\left(1 - (1-\alpha) e^{-\lambda x}\right)} \right\} \left\{ \frac{x \alpha e^{-\lambda x}}{\left(1 - (1-\alpha) e^{-\lambda x}\right)^2} \right\} \end{aligned} \tag{18}$$

$$\frac{\partial \ell}{\partial \theta} = -\frac{n}{\theta} - \sum_{i=1}^{n} \ln\left\{1-(1-\alpha)e^{-\lambda x}\right\} + \frac{2}{\theta^2}\sum_{i=1}^{n} \ln\left\{\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right\} \tag{19}$$

By setting $\frac{\partial \ell}{\partial \alpha} = \frac{\partial \ell}{\partial \lambda} = \frac{\partial \ell}{\partial \theta} = 0$ and solving them for α, λ and θ we get the ML estimators of the $HCEG(\alpha,\lambda,\theta)$ distribution. But normally, it is not possible to solve non-linear equations (16) so with the aid of suitable computer software one can solve them easily. Let $\underline{\Theta} = (\alpha,\lambda,\theta)$ denote the parameter vector of $HCEG(\alpha,\lambda,\theta)$ and the corresponding MLE of $\underline{\Theta}$ as $\hat{\underline{\Theta}} = (\hat{\alpha},\hat{\lambda},\hat{\theta})$, then the asymptotic normality results in, $\left(\hat{\underline{\Theta}} - \underline{\Theta}\right) \to N_3\left[0,\left(I\left(\underline{\Theta}\right)\right)^{-1}\right]$ where $I\left(\underline{\Theta}\right)$ is the Fisher's information matrix given by,

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

In practice, we don't know $\underline{\Theta}$ hence it is useless that the MLE has an asymptotic variance $\left(I\left(\underline{\Theta}\right)\right)^{-1}$. Hence we approximate the asymptotic variance by plugging in the estimated value of the parameters. The observed fisher information matrix $O(\hat{\underline{\Theta}})$ is used as an estimate of the information matrix $I\left(\underline{\Theta}\right)$ given by

$$O(\hat{\underline{\Theta}}) = -\begin{pmatrix} \frac{\partial^2 l}{\partial \alpha^2} & \frac{\partial^2 l}{\partial \hat{\alpha} \partial \hat{\lambda}} & \frac{\partial^2 l}{\partial \hat{\alpha} \partial \hat{\theta}} \\ \frac{\partial^2 l}{\partial \hat{\alpha} \partial \hat{\lambda}} & \frac{\partial^2 l}{\partial \hat{\lambda}^2} & \frac{\partial^2 l}{\partial \hat{\theta} \partial \hat{\lambda}} \\ \frac{\partial^2 l}{\partial \hat{\alpha} \partial \hat{\theta}} & \frac{\partial^2 l}{\partial \hat{\theta} \partial \hat{\lambda}} & \frac{\partial^2 l}{\partial \hat{\theta}^2} \end{pmatrix}_{\left(\hat{\alpha},\hat{\lambda},\hat{\theta}\right)} = -H\left(\underline{\Theta}\right)_{\left(\underline{\Theta}=\hat{\underline{\Theta}}\right)}$$

where H is the Hessian matrix.

The Newton-Raphson algorithm to maximize the likelihood produces the observed information matrix. Therefore, the variance-covariance matrix is given by,

$$\left[-H\left(\underline{\Theta}\right)_{\left|(\underline{\Theta}=\hat{\underline{\Theta}})\right.}\right]^{-1}=\begin{pmatrix}\operatorname{var}(\hat{\alpha}) & \operatorname{cov}(\hat{\alpha},\hat{\lambda}) & \operatorname{cov}(\hat{\alpha},\hat{\theta})\\ \operatorname{cov}(\hat{\alpha},\hat{\lambda}) & \operatorname{var}(\hat{\lambda}) & \operatorname{cov}(\hat{\lambda},\hat{\theta})\\ \operatorname{cov}(\hat{\alpha},\hat{\theta}) & \operatorname{cov}(\hat{\lambda},\hat{\theta}) & \operatorname{var}(\hat{\theta})\end{pmatrix} \tag{20}$$

Hence from the asymptotic normality of MLEs, approximate *100(1-a) %* confidence intervals for α, λ and θ can be constructed as,

$$\hat{\alpha}\pm Z_{a/2}\sqrt{\operatorname{var}(\hat{\alpha})}\,,\ \hat{\lambda}\pm Z_{a/2}\sqrt{\operatorname{var}(\hat{\lambda})}\ \text{and}\ \hat{\theta}\pm Z_{a/2}\sqrt{\operatorname{var}(\hat{\theta})}\,.$$

where $Z_{a/2}$ is the upper percentile of standard normal variate.

Method of Least-Square Estimation (LSE)

The least-square estimators of the unknown parameters α, λ and θ of $HCEG(\alpha,\lambda,\theta)$ distribution can be obtained by minimizing

$$A\left(x;\alpha,\lambda,\theta\right)=\sum_{i=1}^{n}\left[F(X_i)-\frac{i}{n+1}\right]^2 \tag{21}$$

with respect to unknown parameters α, λ and θ.

Suppose $F(X_i)$ denotes the CDF of the ordered random variables $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ where $\{X_1, X_2, \ldots, X_n\}$ is a random sample of size n from a distribution function F(.). The least-square estimators of α, λ and θ say $\hat{\alpha}, \hat{\lambda},\ and\ \hat{\theta}$ respectively, can be obtained by minimizing

$$A\left(x;\alpha,\lambda,\theta\right)=\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left[-\frac{1}{\theta}\log\left\{\frac{\alpha e^{-\lambda x}}{\left(1-\left(1-\alpha\right)e^{-\lambda x}\right)}\right\}\right]-\frac{i}{n+1}\right]^2;\ x>0,(\alpha,\lambda,\theta)>0 \tag{22}$$

with respect to α, λ and θ. Differentiating (22) with respect to α, λ and θ we get,

$$A(x;\alpha,\lambda,\theta) = \sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left[-\frac{1}{\theta}\log\left\{\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right\}\right]-\frac{i}{n+1}\right]^2;\ x>0,(\alpha,\lambda,\theta)>0 \tag{23}$$

$$\frac{\partial A}{\partial \alpha} = -\frac{4}{\pi\theta}\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)\right\}-\frac{i}{n+1}\right]\left[1+\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)-\frac{i}{n+1}\right\}^2\right]^{-1}$$
$$\left\{\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right\}^{-1}\left\{\frac{e^{-\lambda x}\left(1-e^{-\lambda x}\right)}{\left(1-(1-\alpha)e^{-\lambda x}\right)^2}\right\} \tag{24}$$

$$\frac{\partial A}{\partial \lambda} = -\frac{4}{\pi\theta}\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)\right\}-\frac{i}{n+1}\right]$$
$$\left[1+\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)-\frac{i}{n+1}\right\}^2\right]^{-1}\left\{\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right\}^{-1}\left\{\frac{\alpha x e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)^2}\right\} \tag{25}$$

$$\frac{\partial A}{\partial \theta} = -\frac{4}{\pi\theta^2}\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)\right\}-\frac{i}{n+1}\right]$$
$$\left[1+\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)-\frac{i}{n+1}\right\}^2\right]^{-1}\ln\left\{\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right\} \tag{26}$$

Similarly the weighted least square estimators can be obtained by minimizing

$$D(X;\alpha,\lambda,\theta) = \sum_{i=1}^{n} w_i\left[F(X_{(i)})-\frac{i}{n+1}\right]$$

with respect to α, λ and θ. The weights w_i are $w_i = \frac{1}{Var(X_{(i)})} = \frac{(n+1)^2(n+2)}{i(n-i+1)}$

Hence, the weighted least square estimators of α, λ and θ respectively can be obtained by minimizing,

$$D(X;\alpha,\lambda,\theta)=\sum_{i=1}^{n}\frac{(n+1)^2(n+2)}{i(n-i+1)}\left[\frac{2}{\pi}\arctan\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)\right\}-\frac{i}{n+1}\right]^2 \tag{27}$$

with respect to α, λ and θ.

Method of Cramer-Von-Mises estimation (CVME)

The Cramer-Von-Mises estimators of α, λ and θ are obtained by minimizing the function

$$Z(X;\alpha,\lambda,\theta)=\frac{1}{12n}+\sum_{i=1}^{n}\left[F\left(x_{i:n}\mid\alpha,\lambda,\theta\right)-\frac{2i-1}{2n}\right]^2$$

$$=\frac{1}{12n}+\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)\right\}-\frac{2i-1}{2n}\right]^2 \tag{28}$$

Differentiating (28) with respect to α, λ and θ we get,

$$\frac{\partial Z}{\partial\alpha}=\frac{1}{12n}-\frac{4}{\pi\theta}\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)\right\}-\frac{2i-1}{2n}\right]\left[1+\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)-\frac{2i-1}{2n}\right\}^2\right]^{-1}$$
$$\left\{\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right\}^{-1}\left\{\frac{e^{-\lambda x}\left(1-e^{-\lambda x}\right)}{\left(1-(1-\alpha)e^{-\lambda x}\right)^2}\right\} \tag{29}$$

$$\frac{\partial Z}{\partial\lambda}=\frac{1}{12n}-\frac{4}{\pi\theta}\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)\right\}-\frac{2i-1}{2n}\right]\left[1+\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)-\frac{2i-1}{2n}\right\}^2\right]^{-1}$$
$$\left\{\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right\}^{-1}\left\{\frac{\alpha x e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)^2}\right\} \tag{30}$$

$$\frac{\partial Z}{\partial\theta}=\frac{1}{12n}-\frac{4}{\pi\theta^2}\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)\right\}-\frac{2i-1}{2n}\right]$$
$$\left[1+\left\{-\frac{1}{\theta}\ln\left(\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right)-\frac{2i-1}{2n}\right\}^2\right]^{-1}\ln\left\{\frac{\alpha e^{-\lambda x}}{\left(1-(1-\alpha)e^{-\lambda x}\right)}\right\} \tag{31}$$

By solving $\frac{\partial Z}{\partial \alpha} = 0, \frac{\partial Z}{\partial \lambda} = 0 \; and \; \frac{\partial Z}{\partial \theta} = 0$ simultaneously we obtain the CVM estimators.

Application to Real Dataset

In this section, we have demonstrated the applicability of the $HCEG(\alpha, \lambda, \theta)$ distribution using a real dataset. The data represents the remission times (in months) of a random sample of 128 bladder cancer patients [Lee and Wang (2003)]:

0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11,23.63,0.20,2.23,3.52,4.98,6.97,9.02,13.29,0.40,2.26,3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74,14.76,26.31,0.81,2.62,3.82,5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14,79.05,1.35,2.87,5.62,7.87,11.64,17.36,1.40,3.02, 4.34, 5.71, 7.93, 11.79 ,18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13,1.76,3.25,4.50,6.25,8.37,12.02,2.02,3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36,6.76,12.07,21.73,2.07,3.36, 6.93, 8.65,12.63,22.69.

By employing the optim () function in R software [R Core Team, 2020] and [Ming Hui, 2019], we have calculated the MLEs of HCEG distribution by maximizing the likelihood function (16). We have obtained the value of Log-Likelihood is *l* = -410.2938. In Table 1, we have demonstrated the M.L.E.'s with their standard errors (S.E.) for α, λ and θ.

Table 1

MLE and SE α, θ and λ of HCEG

Parameter	MLE	SE
alpha	35.06397791	44.17386118
theta	0.02359592	0.03303572
lambda	0.09566683	0.02741327

In Figure 2 we have plotted the Q-Q plot and P-P plot and it is seen that the HCEG distribution fits the data very well.

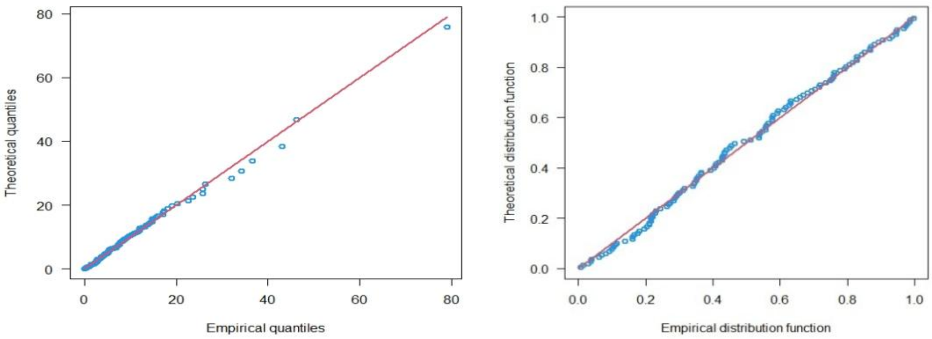

Figure 2: *The Q-Q plot (left panel) and P-P plot (right panel) of the HCEG distribution.*

In Table 2, we have presented the estimated value of the parameters of HCGZ distribution using MLE, LSE and CVE method and their corresponding negative log-likelihood, AIC, BIC and KS statistics with p-value.

Table 2

Estimated parameters, log-likelihood, AIC, BIC and KS statistic

Method of Estimation	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	-LL	AIC	BIC	*KS(p-value)*
MLE	35.0640	0.0236	0.0957	-410.2938	826.5877	835.1438	0.0507 (0.8972)
LSE	33.4331	0.0230	0.0872	-410.4349	826.8698	835.4259	0.0497(0.9094)
CVE	34.1377	0.0295	0.1104	-410.4263	826.8526	835.4087	0.0449(0.9587)

In Figure 3, we have plotted the histogram and the density plot (left panel) using different methods of estimation MLE, LSE and CVME. Similarly, figure in right panel compares the empirical c.d.f. with estimated c.d.f. using different method of estimations.

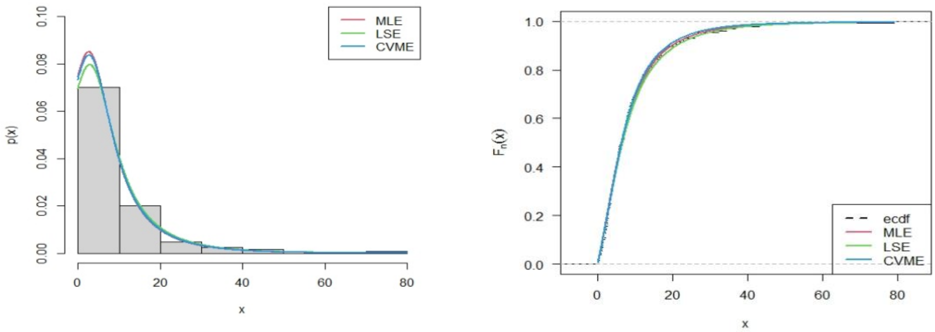

Figure 3: *The Histogram & fitted p.d.f. (left panel) and c.d.f. versus estimated c.d.f. (right panel).*

Here, we have illustrated the applicability of HCEG distribution using the same real dataset comparing to different models used by different researchers. To compare the potentiality of the proposed model, we have considered the following four distributions.

Generalized Exponential Extension (GEE) distribution:

The probability density function of GEE introduced by [Lemonte, 2013] having upside down bathtub-shaped hazard function distribution with parameters α , β and λ is

$$f_{GEE}\left(x\right)=\alpha\beta\lambda\left(1+\lambda x\right)^{\alpha-1} exp\left\{1-\left(1+\lambda x\right)^{\alpha}\right\}\left[1-exp\left\{1-\left(1+\lambda x\right)^{\alpha}\right\}\right]^{\beta-1} ; x>0$$

ExponentiatedWeibull Distribution (EW):

A new family of distributions, namely the *Exponentiated Exponential Distribution* was introduced by Gupta et al. (1998). Here we have considered the *ExponentiatedWeibull Family* that was introduced by Mudholkar and Srivastava (1993). It has a scale parameter and two shape parameters. EW is a generalization of the *Exponentiated Exponential Family* as well as the *Weibull Family*. EW distribution also has a very nice physical interpretation.

$$f_{EW}(x)=\alpha\beta\lambda x^{\beta-1}\exp\left(-\alpha x^{\beta}\right)\left(1-\exp\left(-\alpha x^{\beta}\right)\right)^{\lambda-1}$$

Generalized Weibull Extension (GWE):

Generalized Weibull Extension (GWE) model is very flexible in modeling various types of lifetime distribution. It will be denoted by $GWE(\alpha,\beta,\lambda)$, [Sarhan, A.M. and Apaloo, J.(2013)].

$$f_{GWE}(x)=\alpha\,\beta\left(\lambda x\right)^{\beta-1}\exp\left\{\left(\lambda x\right)^{\beta}+\frac{1}{\lambda}\left(1-\exp\left(\left(\lambda x\right)^{\beta}\right)\right)\right\}$$
$$\left[1-\exp\left\{\frac{1}{\lambda}\left(1-\exp\left(\left(\lambda x\right)^{\beta}\right)\right)\right\}\right]^{\alpha-1};\quad x>0$$

Exponentiated Power Lindley distribution (EPL):

The probability density function of the Exponentiated power Lindley distribution with three parameters is given by,

$$f_{EPL}\left(x\right)=\frac{\alpha\beta\theta^{2}x^{\beta-1}}{\left(\theta+1\right)}\left(1+x^{\beta}\right)e^{-\theta x^{\beta}}\left[1-\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)e^{-\theta x^{\beta}}\right]^{\alpha-1};x>0,$$

where $\alpha>0$, $\beta>0$ and $\lambda>0$ are the parameters. This model is very flexible in modeling various types of lifetime distribution [Ashour and Eltehiwy (2015)].

For the assessment of the fit of the proposed model, we have calculated the Akaike information criterion (AIC), Bayesian information criterion (BIC), Corrected Akaike information criterion (CAIC), and Hannan-Quinn information criterion (HQIC) and presented in Table 3.

Table 3

Log-likelihood (LL), AIC, BIC, CAIC and HQIC

Model	LL	AIC	BIC	CAIC	HQIC
HCEG	-410.2938	826.5877	835.1438	826.7811	821.5587
EPL	-410.4335	826.8670	835.4231	827.0605	821.8381
GEE	-410.6013	827.2026	835.7587	827.3962	822.1737
EW	-410.6801	827.3603	835.9163	827.5538	822.3313
GWE	-410.8070	827.6137	836.1698	827.8076	822.5851

Figure 4 shows the plot of *Histogram* versus the *Density function* of fitted distributions in left panel .In right panel plot of "*Empirical distribution function* "with the "*Estimated distribution function* " of HCEG distribution and some selected distributions are presented.

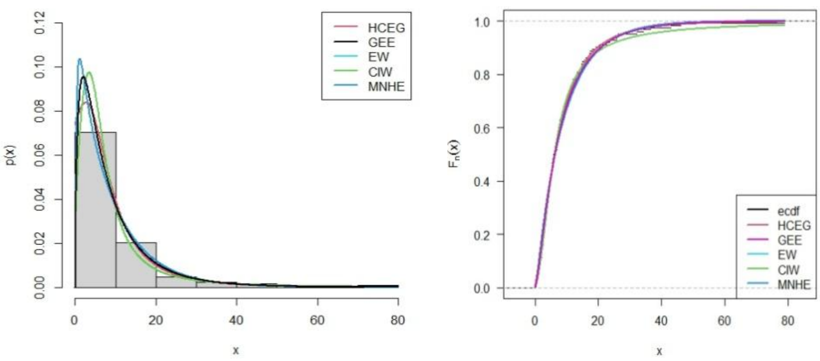

Figure 4: *The Histogram and the density function of fitted distributions (left panel) and Empirical distribution function with estimated distribution function (right panel).*

To compare the goodness-of-fit of the HCEG distribution with other competing distributions, we have presented the value of Anderson-Darling (AD) and the Cramer-Von Mises (CVM) statistics in Table 4. It is observed that the HCEG distribution has the minimum value of the test statistic and higher *p*-value thus we conclude that the HCEG distribution gets better fit and more consistent results than the others taken for comparison.

Table 4

The goodness-of-fit statistics and their corresponding p-value

Model	*AD(p-value)*	*CVM(p-value)*
HCEG	0.24751 (0.9717)	0.033509 (0.9638)
EPL	0.2391(0.9708)	0.035986(0.9532)
GEE	0.26349 (0.9628)	0.039531 (0.9362)
EW	0.27021 (0.9587)	0.040287 (0.9324)
GWE	0.29343(0.9429)	0.044372(0.9102)

Conclusion

In this study, we have introduced a new distribution named Half Cauchy Exponential Geometric Distribution along with some mathematical properties, probability distribution function, survival function, hazard function, skewness, kurtosis, random deviate generation and quantile functions etc. Also, AIC, BIC, CAIC and HQIC are calculated and have smaller values than the other distribution indicating better model. Graphical comparison of empirical distribution functions CDF with estimated distribution function EDF and PDF of proposed model shows better with comparison to other models taken in consideration.

References

Adamidis, K., &Loukas, S. (1998). A lifetime distribution with decreasing failure rate. Statistics & Probability Letters, 39(1), 35-42.

Ashour, S. K., &Eltehiwy, M. A. (2015).Exponentiated power Lindley distribution. Journal of advanced research, 6(6), 895-905.

Elisa T. Lee, John Wenyu Wang (2003), Statistical Methods for Survival Data Analysis, John Wiley & Sons, Inc.

Learn R for Applied Statistics Eric Goh Ming Hui Singapore, Singapore ISBN-13 (pbk): 978-1-4842-4199-8 https://doi.org/10.1007/978-1-4842-4200-1

Lemonte, A. J., &Cordeiro, G. M. (2013).An extended Lomax distribution. Statistics, 47(4), 800-816.

Louzada, F., Marchi, V., & Roman, M. (2014). The exponentiated exponential–geometric distribution: a distribution with decreasing, increasing and unimodal failure rate. Statistics, 48(1), 167–181.

Mudholkar, G. S., &Srivastava, D. K. (1993).ExponentiatedWeibull family for analyzing bathtub failure-rate data.IEEE transactions on reliability, 42(2), 299-302.

R Core Team (2021). R: A language and environment for statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

SaraleesNadarajah, Gauss M. Cordeiro& Edwin M. M. Ortega (2015). The Zografos–Balakrishnan-G Family of Distributions: Mathematical Properties and Applications, Communications in Statistics - Theory and Methods, 44:1, 186-215, DOI: 10.1080/03610926.2012.740127

Sarhan, A. M., &Apaloo, J. (2013).Exponentiated modified Weibull extension distribution. Reliability Engineering & System Safety, 112, 137-144.