J. Nat. Acad. Math.
Vol. 33 (2019), pp.51-65

ISSN 0970-5228

Half-Cauchy Gompertz Distribution : Different Methods of Estimation

By

Lal Babu Sah Telee and Vijay Kumar

Abstract

In this paper, a new distribution called half Cauchy Gompertz distribution is introduced. We have derived some important mathematical properties of the new distribution like hazard function, probability density function, survival function, cumulative distribution function, cumulative hazard function, survival function, quantiles, the measures of skewness based on quartiles and coefficient of kurtosis based on octiles. To estimate the parameters of the new distribution we have applied the three commonly used estimation method namely Cramer-Von-Mises (CVM), maximum likelihood estimators (MLE), and least-square (LSE) methods. For the assessment of potentiality of the new distribution we have consider a real dataset and compared the goodness-of-fit attained by proposed distribution with some competing distribution. It has been observed that the proposed model fits the data well and more flexible as compared to some other models.

Keywords : Half-Cauchy, Gompertz distribution, Estimation, MLE, LSE and CVME.
2010 AMS Subject Classification : 62F15, 65C05.

1. Introduction

The Gompertz model is one of the extensively used probability model having survival function based on laws of mortality. This model can be used in modeling life time data related to human mortality and investigating actuarial tables. Gompertz [7] has defined the Gompertz distribution and it has been employed as a growth model and also used to fit the tumor growth. The function of Gompertz model can be used to reduce a significant gathering of data in life tables into a single function. Initially, this feature was designed to explain human mortality, based on the assumption that as an individual age, exponential decrease in mortality rate is seen.

Received : October 22, 2019; Accepted : December 21, 2019

The cumulative distribution function (CDF) and probability density function (PDF) of Gompertz distribution are

$$G(x) = 1 - \exp\left\{\frac{\lambda}{\alpha}\left(1 - e^{\alpha x}\right)\right\};\ x > 0,\ \alpha > 0, \lambda > 0, \tag{1.1}$$

and

$$g\left(x\right) = \lambda e^{\alpha x} \exp\left\{\frac{\lambda}{\alpha}\left(1 - e^{\alpha x}\right)\right\};\ x > 0,\ \alpha > 0, \lambda > 0, \tag{1.2}$$

respectively. The extensive survey and applications of the Gompertz distribution can be found in Ahuja and Nash [2]. Cooray and Ananda [4] have presented a family of the Gompertz-sinh and used to analyze the reliability data with highly negatively skewed distribution. El-Gohary et al. [6] have introduced a flexible model called the generalized Gompertz distribution with decreasing or increasing or constant or bathtub curve failure rate depending upon the shape parameter. Different method of estimation for the exponentiated Gompertz distribution has been studied by Abu-Zinadah and Al-Oufi [1]. The inverse generalized Gompertz has been introduced in Chaudhary and Kumar [3].

Therefore, we are interested to extend the Gompertz distribution using half-Cauchy family of distribution. Let X be a positive random variable that follows the half-Cauchy distribution and its CDF can be written as

$$\kappa\left(t\right) = \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\theta}\right),\ t > 0, \theta > 0. \tag{1.3}$$

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

$$r\left(t\right) = \frac{2}{\pi}\left(\frac{\theta}{\theta^2 + x^2}\right),\ t > 0, \theta > 0. \tag{1.4}$$

The extending family of distribution has developed by Zografas and Balakrishnan [19] and CDF of family of distribution is

$$F\left(x\right) = \int_0^{-\ln[1-G(x)]} r\left(t\right)\ dt, \tag{1.5}$$

here $G(x)$ is the CDF of any baseline distribution and $r(t)$ is the PDF of any distribution. The family of half-Cauchy distribution whose CDF can be obtained by using $r(t)$ as PDF of half-Cauchy distribution defined in (1.4) and expressed as

$$F(x) = \int_{0}^{-\ln[1-G(x)]} \frac{2}{\pi} \frac{\theta}{\theta^2 + t^2} \, dt = \frac{2}{\pi} \arctan\left(-\frac{1}{\theta} \ln\left[1 - G(x)\right]\right); \; x > 0, \theta > 0. \tag{1.6}$$

The PDF corresponding to (1.6) can be expressed as

$$f(x) = \frac{2}{\pi\theta} \frac{g(x)}{1 - G(x)} \left[1 + \left\{-\frac{1}{\theta} \log\left[1 - G(x)\right]\right\}^2\right]^{-1}; x > 0, \theta > 0. \tag{1.7}$$

The inspiration of this study is to put forward a more flexible distribution by inserting just one extra parameter to the Gompertz distribution to achieve a better fit to the real data. We study the properties of the half Cauchy Gompertz distribution and explore its potentiality and applicability.

The contents of this paper are managed as follows. The new half Cauchy Gompertz distribution is introduced and several distributional properties are discussed in Section 2. Three mostly used estimation approaches are used to estimate the parameters namely least-square (LSE), maximum likelihood estimators (MLE), and Cramer-Von-Mises (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 calculate the approximate confidence intervals of the ML estimators of the parameters and also AIC, CAIC, BIC, HQIC are calculated to evaluate the goodness-of-fit of the half Cauchy Gompertz distribution. Finally, some concluding remarks are presented in Section 5.

2. Half-Cauchy Gompertz Distribution

In this section the new distribution named half Cauchy Gompertz distribution is defined. Substituting (1.1) and (1.2) in (1.6) and (1.7) we get the CDF and PDF of $HCGZ$ distribution. Let X be a non negative random variable follows the $HCGZ(\alpha, \lambda, \theta)$ if its CDF can be written as,

$$F(x) = \frac{2}{\pi} \arctan\left\{-\frac{\lambda}{\alpha\theta}\left(1 - e^{\alpha x}\right)\right\}; \; x > 0, (\alpha, \lambda, \theta) > 0, \tag{2.1}$$

and its PDF may be expressed as,

$$f(x) = \frac{2}{\pi} \frac{\lambda e^{\alpha x}}{\theta} \left[1 + \left\{-\frac{\lambda}{\alpha\theta}\left(1 - e^{\alpha x}\right)\right\}^2\right]^{-1}; \; x > 0. \tag{2.2}$$

Survival function:

$$R(x) = 1 - \frac{2}{\pi} \arctan\left\{-\frac{\lambda}{\alpha\theta}\left(1 - e^{\alpha x}\right)\right\} \tag{2.3}$$

Hazard function:

$$h(x) = \frac{2}{\pi}\frac{\lambda e^{\alpha x}}{\theta}\left[1 - \frac{2}{\pi}\arctan\left\{-\frac{\lambda}{\alpha\theta}\left(1 - e^{\alpha x}\right)\right\}\right]^{-1}\left[1 + \left\{-\frac{\lambda}{\alpha\theta}\left(1 - e^{\alpha x}\right)\right\}^2\right]^{-1}. \tag{2.4}$$

Quantile function:

$$Q(p) = \frac{1}{\alpha}\ln\left\{1 + \frac{\alpha\theta}{\lambda}\tan\left(\frac{\pi p}{2}\right)\right\}; 0 < p < 1. \tag{2.5}$$

The Random Deviate Generation:

$$x = \frac{1}{\alpha}\ln\left\{1 + \frac{\alpha\theta}{\lambda}\tan\left(\frac{\pi u}{2}\right)\right\}; 0 < u < 1. \tag{2.6}$$

where $u \sim U(0, 1)$.

PDF and HRF of $HCGZ(\alpha, \lambda, \theta)$ with numerous values of parameters α and θ for $\lambda = 1$ are plotted which are displayed in Figure 1.

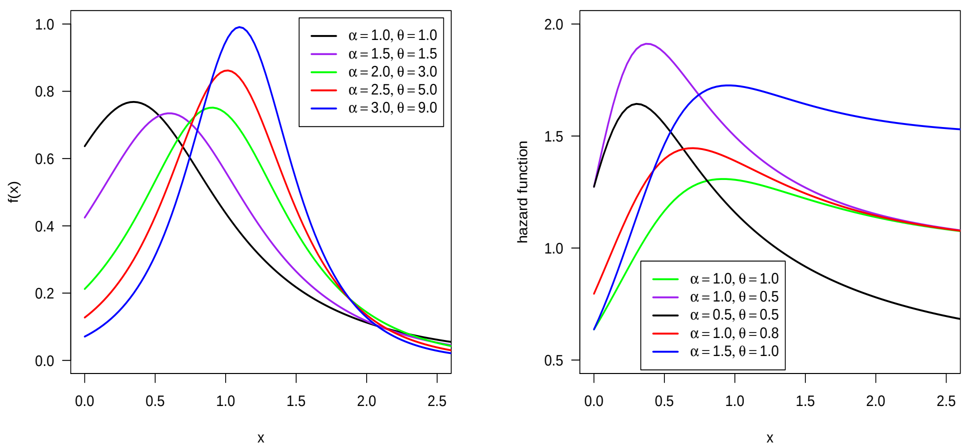


Figure 1. Plots of the probability density function(left panel) and hazard function (right panel), for λ=1 and different values of α and θ.

Skewness and Kurtosis:

The skewness and kurtosis measures are used in statistical analyses to characterize a distribution or a data set. The Bowley's skewness measure based on quartiles is given by

$$S_k = \frac{Q\left(3/4\right) - 2Q\left(1/2\right) + Q\left(1/4\right)}{Q\left(3/4\right) - Q\left(1/4\right)},$$

and the Moors's kurtosis measure based on octiles, Moors [11], is given by

$$K_u = \frac{Q\left(7/8\right) - Q\left(5/8\right) + Q\left(3/8\right) - Q\left(1/8\right)}{Q\left(6/8\right) - Q\left(2/8\right)},$$

where the $Q\left(.\right)$ is the quantile function.The skewness and kurtosis measures based on quantiles like Bowley's skewness and Moors's kurtosis have a number of advantages compared to the classical measures of skewness and kurtosis, e.g. they are less sensitive to outliers and they exist for the distributions even without defined the moments.

3. Methods of Estimation

The object of estimation is to evaluate a model parameter value based on sample information. The estimation theory deals with the basic problem of inferring some relevant features of a chance experiment centered on the observation of the experiment outcomes. There are so many methods which are used to evaluate values of parameters. Three kinds of parameter estimation methods have been considered, such as MLE, LSE, and the Cramer-von Mises (CVM) methods.

(a) Maximum Likelihood Estimation

Let us consider the $\underline{x} = (x_1, ..., x_n)$ of size n be the experiential values from $HCGZ(\alpha, \lambda, \theta)$ then the likelihood function for the parameter vector is expressed as,

$$L\left(\alpha, \lambda, \theta|\underline{x}\right) = \left(\frac{2}{\pi}\right)^n (\lambda\theta)^n \exp\left(\alpha\sum_{i=1}^{n} x_i\right) \prod_{i=1}^{n}\left[\theta^2 + \left\{-\frac{\lambda}{\alpha}\left(1 - e^{\alpha x_i}\right)\right\}^2\right]^{-1}.$$

Taking logarithms on both sides, we get

$$\ell(\alpha,\lambda,\theta|\underline{x}) = n\ln\left(\frac{2}{\pi}\right) + n\ln\lambda + n\ln\theta + \alpha\sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\ln\left[\theta^2 + \left\{-\frac{\lambda}{\alpha}\left(1-e^{\alpha x_i}\right)\right\}^2\right]. \tag{3.1}$$

The elements of the score function are obtained as

$$\begin{aligned}
\frac{\partial\ell}{\partial\alpha} &= \sum_{i=1}^{n} x_i + \frac{2\lambda^2}{\alpha^3}\sum_{i=1}^{n}\left(1-e^{\alpha x_i}\right)\left(1-e^{\alpha x}+\alpha^2 e^{\alpha x_i}\right) \left[1+\left\{-\frac{\lambda}{\alpha\theta}\left(1-e^{\alpha x_i}\right)\right\}^2\right]^{-1} \\
\frac{\partial\ell}{\partial\lambda} &= \frac{n}{\lambda} + \frac{2\lambda}{\alpha}\sum_{i=1}^{n} e^{\alpha x_i}\left(1-e^{\alpha x_i}\right)\left[1+\left\{-\frac{\lambda}{\alpha\theta}\left(1-e^{\alpha x_i}\right)\right\}^2\right]^{-1} \\
\frac{\partial\ell}{\partial\theta} &= \frac{n}{\theta} - 2\theta\sum_{i=1}^{n}\left[1+\left\{-\frac{\lambda}{\alpha\theta}\left(1-e^{\alpha x_i}\right)\right\}^2\right]^{-1}
\end{aligned} \tag{3.2}$$

Equating $\frac{\partial\ell}{\partial\alpha}$, $\frac{\partial\ell}{\partial\lambda}$ and $\frac{\partial\ell}{\partial\theta}$ to zero.
Equating B_α, B_λ *and* B_θ to zero and solve these non-linear equations simultaneously which gives the MLE $\hat{\Theta} = \left(\hat{\alpha},\hat{\lambda},\hat{\theta}\right)$ of $\Theta = (\alpha,\lambda,\theta)^T$. Manually we cannot solve these equations so by using the computer software R, Mathematica, Matlab, or any other programs and Newton-Raphson's iteration method, one can solve these equations. Let $\Theta = (\alpha,\lambda,\theta)^T$ be the parameter vector and associated MLE of Θ as $\hat{\Theta} = \left(\hat{\alpha},\hat{\lambda},\hat{\theta}\right)$, and then using the result of asymptotic normality we have, $\left(\hat{\Theta}-\Theta\right) \to N_3\left[0,\left(I\left(\Theta\right)\right)^{-1}\right]$, where $I\left(\Theta\right)$ is the information matrix of Fisher.

(b) Least-Square Estimation (LSE) Method

The least-square estimators of the unknown parameters α, λ and θ of HCGZ distribution can be obtained by minimizing

$$S\left(X;\alpha,\lambda,\theta\right) = \sum_{i=1}^{n}\left[F(X_{(i)}) - \frac{i}{n+1}\right]^2, \tag{3.4}$$

with respect to (w.r.t.) α, λ and θ, Swain et al.[18].

From a distribution function $F(.)$, consider random sample be denoted by $\{X_1,\ldots,X_n\}$ with sample size is n where $F(X_i)$ represents the distribution function of the random variables ordered $\mathrm{X}_{(1)} < \mathrm{X}_{(2)} < \ldots < \mathrm{X}_{(n)}$. Then LSE

($\tilde{\alpha}, \tilde{\lambda}$ and $\tilde{\theta}$) is acquired with minimization of

$$S(X;\alpha,\lambda,\theta)=\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left\{-\frac{\lambda}{\alpha\theta}\left(1-e^{\alpha x_{(i)}}\right)\right\}-\frac{i}{n+1}\right]^{2}, \qquad (3.5)$$

with respect to α, λ and θ. Differentiation of (3.5) with respect to α, λ and θ yields,

$$\frac{\partial S}{\partial \alpha}=\frac{4}{\pi}\frac{\lambda}{\alpha^{2}\theta}\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left\{-\frac{\lambda}{\alpha\theta}A(x_{(i)})\right\}-\frac{i}{n+1}\right]$$
$$\left\{1+e^{\alpha x_{(i)}}\left(\alpha x_{(i)}-1\right)\right\}T(x_{(i)}),$$
$$\frac{\partial S}{\partial \lambda}=-\frac{4}{\pi\alpha\theta}\sum_{i=1}^{n}A(x_{(i)})\left[\frac{2}{\pi}\arctan\left\{-\frac{\lambda}{\alpha\theta}A(x_{(i)})\right\}-\frac{i}{n+1}\right]T(x_{(i)}),$$
$$\frac{\partial S}{\partial \theta}=\frac{4}{\pi\alpha\theta^{2}}\sum_{i=1}^{n}A(x_{(i)})\left[\frac{2}{\pi}\arctan\left\{-\frac{\lambda}{\alpha\theta}A(x_{(i)})\right\}-\frac{i}{n+1}\right]T(x_{(i)}).$$

where

$$T(x_{(i)})=\left[1+\left\{-\frac{\lambda}{\alpha\theta}A(x_{(i)})\right\}^{2}\right]^{-1} \text{ and } A(x_{(i)})=1-e^{\alpha x_{(i)}}.$$

Likewise, the weighted LSEs can be found with minimization w.r.t. α, λ and θ.

$$D(X;\alpha,\lambda,\theta)=\sum_{i=1}^{n}w_{i}\left[F(X_{(i)})-\frac{i}{n+1}\right]^{2}.$$

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 LSEs of α, λ and θ can be found respectively by minimizing following function w.r.t. α, λ and θ.

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

(c) Cramer-Von-Mises estimation (CVME)

The CVM estimators for α, λ and θ are obtained by minimization of

$$C\left(X;\alpha,\lambda,\theta\right)=\frac{1}{12n}+\sum_{i=1}^{n}\left[F\left(x_{(i)}|\alpha,\lambda,\theta\right)-\frac{2i-1}{2n}\right]^2$$
$$=\frac{1}{12n}+\sum_{i=1}^{n}\left[\frac{2}{\pi}\arctan\left\{-\frac{\lambda}{\alpha\theta}\left(1-e^{\alpha x_{(i)}}\right)\right\}-\frac{2i-1}{2n}\right]^2. \tag{3.6}$$

Differentiating (3.6) with respect to α, λ and θ we get,

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

The CVM estimators can be found by solving

$$\frac{\partial C}{\partial\alpha}=0,\frac{\partial C}{\partial\lambda}=0\ and\ \frac{\partial C}{\partial\theta}=0. \tag{3.7}$$

simultaneously.

4. Real data Application

We establish the applicability of HCGZ model in this section by using a real dataset used by earlier investigators. The data is on the breaking stress of 66 carbon fibres of 50 mm length (GPa). The data has been previously used by Nichols and Padgett [13].

0.39, 0.85, 1.08, 1.25, 1.47, 1.57, 1.61, 1.61, 1.69, 1.80, 1.84,
1.87, 1.89, 2.03, 2.03, 2.05, 2.12, 2.35, 2.41, 2.43, 2.48, 2.50,
2.53, 2.55, 2.55, 2.56, 2.59, 2.67, 2.73, 2.74, 2.79, 2.81, 2.82,
2.85, 2.87, 2.88, 2.93, 2.95, 2.96, 2.97, 3.09, 3.11, 3.11, 3.15,
3.15, 3.19, 3.22, 3.22, 3.27, 3.28, 3.31, 3.31, 3.33, 3.39, 3.39,
3.56, 3.60, 3.65, 3.68, 3.70, 3.75, 4.20, 4.38, 4.42, 4.70, 4.90

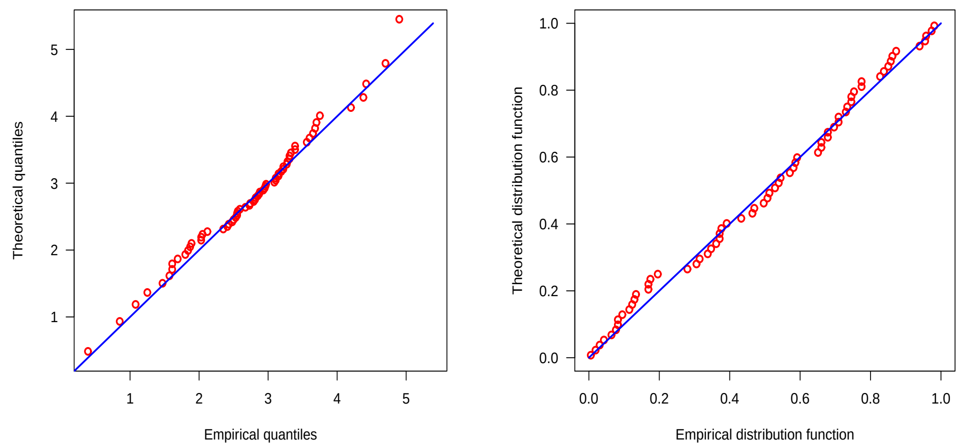
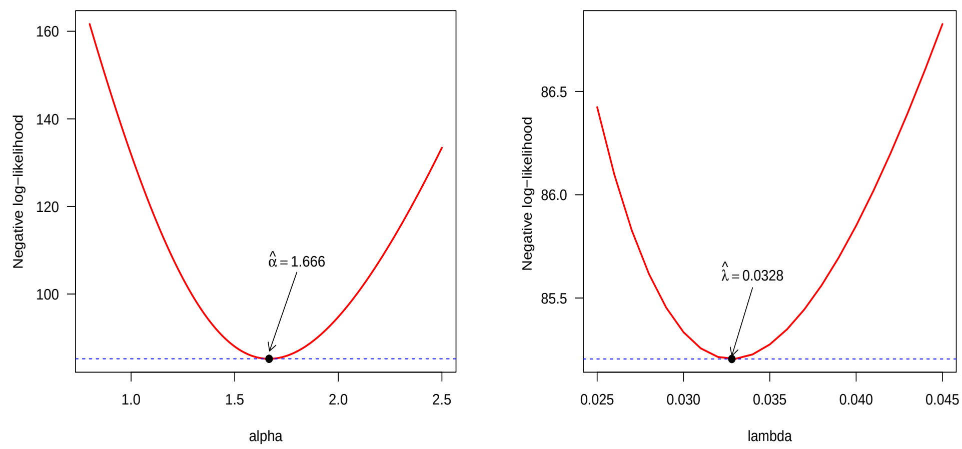


Figure 2. QQ plot(left panel) and the PP plot (right panel).

The HCGZ distribution MLEs are calculated by maximizing the likelihood function (3.1) by using *optim*() command in R platform (R Core Team [14]) and Rizzo [17]. We have obtained the value of log-likelihood is $\ell(\hat{\Theta}) = -85.2050$. The MLEs of α, λ and θ are given in table 1. The Q-Q plot and P-P plot have been plotted in Figure 2, and it is seen that the HCGZ probability model is well adapted to the data given, Kumar and Ligges [8]. Profile log-likelihood function graphs of α, λ and θ are shown in Figure 3 and it is perceived the MLEs are uniquely determined.



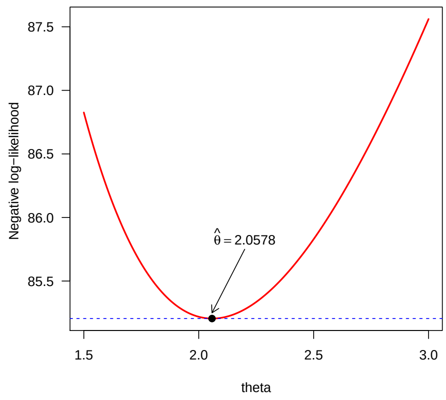


Figure 3. Profile log-likelihood functions of α, λ and θ.

The models are compared via the Akaike Information Criterion (AIC), the Corrected Akaike Information criterion (CAIC), Bayesian information criterion (BIC) and Hanann-Quinn information criterion(HQIC) which are used to select the best model among several models, (D'Agostino and Stephens [5]). The definitions of AIC, BIC, CAIC and HQIC are given below:

$$AIC = -2\,\ell(\hat{\Theta}) + 2\,k$$
$$BIC = -2\,\ell(\hat{\Theta}) + \,k\,\log{(n)}$$
$$CAIC = AIC\, + \,\frac{2k\,(k+1)}{n-k-1}$$
$$HQIC = -2\,\ell(\hat{\Theta})\, + \,2\,k\,log(log(n))$$

where k is the number of parameters in the model under consideration. Table 1 presents the estimated value of the HCGZ distribution parameters using the MLE, LSE and CVME methods and their negative log-likelihood and AIC criteria.

Table 1

Estimated parameters, log-likelihood, AIC and BIC

Method	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\theta}$	$-\ell(\hat{\Theta})$	AIC	BIC
MLE	1.6660	0.0328	2.0578	85.2050	176.4100	182.9790
LSE	1.6086	0.0329	1.8127	85.2473	176.4947	183.0636
CVME	1.6536	0.0306	1.8671	85.2070	176.4139	182.9829

Moreover, perfection of competing models is also tested via the Kolmogrov-Simnorov(K-S), the Anderson-Darling (A^2) and the Cramer-Von Mises (W) statistics. The mathematical expressions for the statistics above are given below

$$KS = \max_{1\leqslant i\leqslant n}\left(\xi_i - \frac{i-1}{n}, \frac{i}{n} - \xi_i\right) ,$$

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left\{\ln \xi_i + \ln(1-\xi_{n+1-i})\right\} ,$$

$$W = \frac{1}{12n} + \sum_{i=1}^{n}\left\{\frac{(2i-1)}{2n} - \xi_i\right\}^2 ,$$

where $\xi_i = CDF\left(x_{(i)}\right)$; the $x_{(i)}$'s being the ordered observations, (D'Agostino and Stephens [5]). In Table 3 we have presented The KS, W and A^2 statistics with their corresponding p-value of MLE, LSE and CVE estimates.

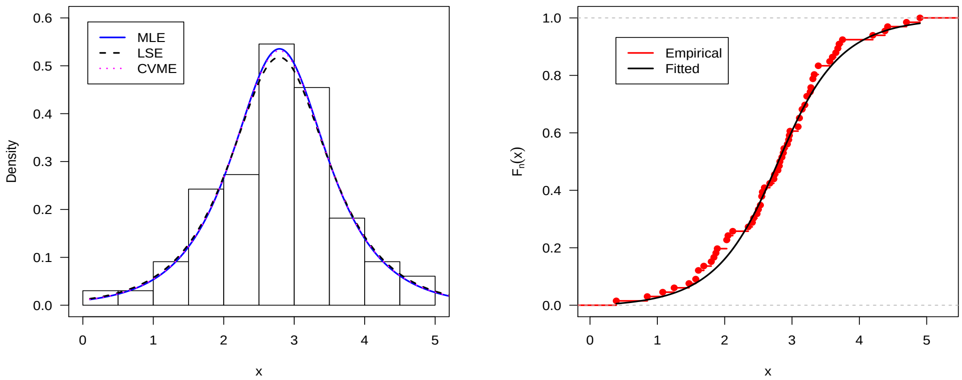


Figure 4. The Histogram and the PDF of fitted distributions for MLE, LSE and CVME methods (left panel) and KS plot of HCGZ distribution (right panel).

Table 2

The KS, A^2 and W statistics with p-value

Method	KS(p-value)	W(p-value)	A^2(p-value)
MLE	0.0677(0.9227)	0.0468(0.8967)	0.3321(0.9120)
LSE	0.0652(0.9418)	0.0477(0.8914)	0.3256(0.9175)
CVME	0.0662(0.9344)	0.0468(0.8972)	0.3287(0.9149)

The values of Kolmogorov-Smirnov(KS), Anderson-Darling(A^2) and Cramer-Von Mises (W) statistic with their respective p-value of different models are reported in Table 3. As we can see in Table 3, the proposed model has the minimum values of the test statistics and higher p-value. Figure 4 (left panel) displays the histogram and the fitted density functions, which support the results in Tables 2 and 3. Also, Figure 4 (right panel) which compares the distribution functions for the different models with the empirical distribution function reveals the same. Therefore, for the given data set shows the proposed distribution gets better fit and more reliable solutions from other alternatives.

Some of the well-known existing distributions are chosen for comparison purposes, which are listed below, to show the goodness-of-fit of the HCGZ model.

(i) Generalized Gompertz distribution:
The probability density function of generalized Gompertz(GGZ) distribution (El-Gohary et al.[6]) with parameters $\alpha > 0, \lambda > 0$ and $\theta > 0$ is

$$f_{GGZ}(x) = \theta\lambda e^{\alpha x} e^{-\frac{\lambda}{\alpha}(e^{\alpha x}-1)} \left[1 - \exp\left(-\frac{\lambda}{\alpha}(e^{\alpha x}-1)\right)\right]^{\theta-1}; x > 0.$$

(ii) Gompertz distribution:
The probability density function of Gompertz distribution(GZ)with parameters α and $\theta > 0$ is given by, (Marshall & Olkin [10])

$$f_{GZ}(x) = \theta \ e^{\alpha x} \ \exp\left\{\frac{\theta}{\alpha}(1 - e^{\alpha x})\right\} \quad ; x \geqslant 0, \ \theta > 0, \ -\infty < \alpha < \infty.$$

(iii) Exponentiated Exponential Poisson (EEP) distribution:
The probability density function of EEP (Ristić & Nadarajah [16]) with parameters $\alpha > 0, \beta > 0$ and $\lambda > 0$ can be expressed as

$$f_{EEP}(x) = \frac{\alpha\beta\lambda}{(1 - e^{-\lambda})} e^{-\beta x} \left(1 - e^{-\beta x}\right)^{\alpha-1} \exp\left\{-\lambda\left(1 - e^{-\beta x}\right)^{\alpha}\right\}; \ x > 0.$$

(iv) Generalized Exponential Extension (GEE) distribution:
The probability density function of GEE introduced by (Lemonte [9]) having upside down bathtub-shaped hazard function distribution with parameters α, β and λ is

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

(v) Power Cauchy distribution:
The PDF of power Cauchy (PC) distribution has been introduced by (Rooks et al. [15]) is

$$f_{PC}(x; \alpha, \lambda) = \frac{2\alpha}{\pi x}\left(\frac{x}{\lambda}\right)^{\alpha}\left\{1 + \left(\frac{x}{\lambda}\right)^{2\alpha}\right\}^{-1}; \ x > 0, \alpha > 0, \lambda > 0.$$

(vi) Exponentiated Weibull distribution:
The PDF of exponentiated Weibull (EW) distribution is introduced by (Mudholkar & Srivastava [12])

$$f_{EW}\left(x\right) = \alpha\beta\lambda\, x^{\beta-1}\exp\left(-\alpha x^{\beta}\right)\left\{1-\exp\left(-\alpha x^{\beta}\right)\right\}^{\lambda-1};x>0.$$

The CAIC, AIC, HQIC, and BIC that are shown in Table 4 have been determined to judge the potentiality of the HCGZ distribution.

Table 3
log-likelihood, AIC, BIC, CAIC and HQIC

Distribution	$-\ell(\hat{\Theta})$	AIC	BIC	CAIC	HQIC
HCGZ	85.2050	176.4100	182.9790	176.7971	179.0057
GGZ	85.6858	177.3716	183.9406	177.7587	179.9673
EW	85.9447	177.8894	184.4584	178.2765	180.4851
EEP	86.6899	179.3798	185.9488	179.7669	181.9755
GEE	87.2704	180.5408	187.1098	180.9279	183.1365
GZ	88.0884	180.1767	184.5560	180.3672	181.9072
PC	90.5126	185.0252	189.4045	185.2157	186.7557

The fitted distribution's PDF and histogram and the empirical distribution function of some selected distributions with the estimated distribution function of the HCGZ distribution are illustrated in Figure 5.

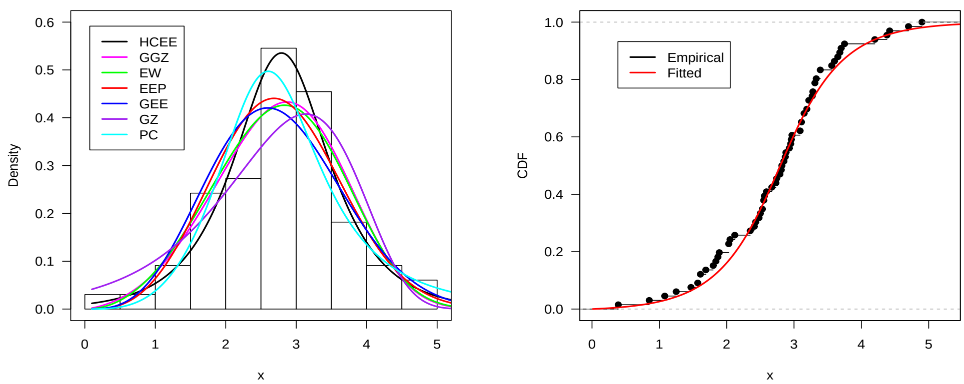


Figure 5. The Histogram and the PDF of fitted distributions (left panel); Empirical CDF with estimated CDF(right panel).

Table 4
Goodness of fit statistic

Distribution	*KS(p-value)*	*W(p-value)*	A^2*(p-value)*
HCGZ	0.0677(0.9227)	0.0468(0.8967)	0.3321(0.9120)
GGZ	0.0833(0.7498)	0.0715(0.7443)	0.4457(0.8020)
EW	0.0809(0.7805)	0.0813(0.6861)	0.4846(0.7620)
EEP	0.0895(0.6662)	0.1014(0.5796)	0.5657(0.6804)
GEE	0.1096(0.4065)	0.1530(0.3812)	0.7816(0.4940)
GZ	0.1120(0.3794)	0.1397(0.4233)	0.9485(0.3851)
PC	0.0963(0.5731)	0.1246(0.4782)	1.0733(0.3207)

The values of the statistics for Anderson-Darling (W), Cramer-Von Mises (A^2) and Kolmogorov-Simnorov (KS) are given in Table 5 in order to evaluate the HCGZ distribution's goodness-of-fit with other rival distributions. The test statistical value of the HCGZ distribution is found to be both minimum and greater p-value, so we achieve that the HCGZ distribution obtains results that are relatively well-fit and steadier and more precise than those taken for comparison.

5. Conclusion

In this study, we have presented a new distribution called half Cauchy Gompertz distribution. A comprehensive study of some statistical and mathematical properties of the proposed distribution including the derivation of explicit expressions for its reliability function, survival function, hazard function, the quantile function and skewness and kurtosis. Three well-known estimation methods namely maximum likelihood estimation (MLE), least-square estimation (LSE), and Cramer-Von-Mises estimation (CVME) methods are used for the parameter estimation and we found that the MLEs are relatively good than LSE and CVM methods. The curves of the PDF of the proposed distribution have shown that its shape is increasing-decreasing and right skewed and flexible for modeling real-life data. Also, the graph of the hazard function is monotonically increasing or constant or reverse j-shaped according to the value of the model parameters. The applicability and suitability of the proposed distribution has been evaluated by considering a real-life dataset and the results exposed that the proposed distribution is much flexible as compared to some other fitted distributions.

Department of Mathematics & Statistics
DDU Gorakhpur University, Gorakhpur
e-mail : vkgkp@rediffmail.com
e-mail : lalbabu3131@gmail.com

References

[1] **Abu-Zinadah, H. H. and Al-Oufi, A. S.:** Different Method Estimations of the Three Parameters of Exponentiated Gompertz Distribution, Appl. Math. Inf. Sci., 10(2) (2016), 705-710.

[2] **Ahuja, J. C. and Nash, S.W.:** The generalized Gompertz-Verhulst family of distributions, Sankhya, Part A, 29 (1967), 141-156.

[3] **Chaudhary, A. K. and Kumar, V.:** Inverse Generalized Gompertz Distribution with Properties and Applications, Journal of National Academy of Mathematics, 31 (2017), 99-108.

[4] **Cooray, K. and Ananda, M. M.:** Analyzing survival data with highly negatively skewed distribution: The Gompertz-sinh family, Journal of Applied Statistics, 37(1) (2010), 1-11.

[5] **D'Agostino, R. B. and Stephens, M. A.:** Goodness-of-fit techniques, New York, Marcel Dekker; (1986).

[6] **El-Gohary, A., Alshamrani, A., and Al-Otaibi, A. N.:** The generalized Gompertz distribution, Applied Mathematical Modelling, 37(1-2) (2013), 13-24.

[7] **Gompertz, B.:** On the nature of the function expressive of the law of human mortality and on the new mode of determining the value of life contingencies, Phil. Trans. Royal Soc. A, 115 (1824), 513-580.

[8] **Kumar, V. and Ligges, U. :** reliaR: A package for some probability distributions, (2011), http://cran.r-project.org/web/packages/reliaR/index.html

[9] **Lemonte, A. J.:** A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function, Computational Statistics and Data Analysis, 62 (2013), 149-170.

[10] **Marshall, A. W. and Olkin, I.:** Life Distributions: Structure of Nonparametric, Semiparametric, and Parametric Families, Springer, New York, (2007).

[11] **Moors, J.:** A quantile alternative for kurtosis, The Statistician, 37 (1988), 25-32.

[12] **Mudholkar, G. S. and Srivastava, D. K.:** Exponentiated Weibull family for analyzing bathtub failure-rate data, IEEE Transactions on Reliability, 42(2) (1993), 299-302.

[13] **Nichols, M. D. and Padgett, W. J.:** A bootstrap control chart for Weibull percentiles. Quality and reliability engineering international, 22(2)(2006), 141-151.

[14] **R Core Team:** R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ (2018).

[15] **Rooks, B., Schumacher, A., and Cooray, K.:** The power Cauchy distribution: derivation, description, and composite models. NSF-REU Program Reports, (2010).

[16] **Ristić, M. M. and Nadarajah, S.:** A new lifetime distribution, Journal of Statistical Computation and Simulation, 84(1) (2014), 135-150.

[17] **Rizzo, M. L.:** Statistical computing with R. Chapman and Hall/CRC, (2008).

[18] **Swain, J. J., Venkatraman, S. and Wilson, J. R.:** Least-squares estimation of distribution functions in Johnson's translation system, Journal of Statistical Computation and Simulation, 29(4) (1988), 271-297.

[19] **Zografos, K. and Balakrishnan, N.:** On families of beta-and generalized gamma-generated distributions and associated inference, Statistical Methodology, 6(4)(2009), 344-362.