Journal of Nepal Mathematical Society (JNMS)
ISSN: 2616-0153 (Print), 2616-0161 (Online)
Vol. 7, Issue 1, 2024 (July): 1-13
DOI: https://doi.org/10.3126/jnms.v7i1.67483

Research Article
Received Date: April 04, 2024
Accepted Date: June 13, 2024
Published Date: July 03, 2024

Model and Properties of Cauchy Modified Inverse Gompertz Distribution with Application to a Real Data Set

Arun Kumar Chaudhary[1,*], Lal Babu Sah Telee[1]

[1]Department of Management Science, Nepal Commerce Campus, Tribhuvan University, Kathmandu, Nepal

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

Abstract: *In this study, we introduce the Cauchy modified inverse Gompertz distribution as a new probability model. Utilizing the modified inverse Gompertz distribution as its baseline distribution, this model blends the Cauchy family of distributions. Our aim is to utilize this model for lifetime data analysis. We have inferred formulas for some basic properties of the model. We have also included graphic representations of the hazard rate and probability density curves. We observed that the probability density function displays positive skewness, while the hazard rate function plot shows an increasing-decreasing pattern. We used the Cramer-Von Mises method, the least squares approach, and maximum likelihood estimation to estimate the model parameters. We employed several statistical criteria to validate our model, including the corrected Akaike's, the Bayesian, the Hannan-Quinn, as well as Akaike's information criterion. We also utilized Q-Q and P-P graphs for further confirmation. To assess goodness of fit, we used the Kolmogorov-Smirnov, Anderson-Darlin, and Cramer-von Mises tests. The empirical results of the study show that it gives a better fit to the real data set. All numerical computations were conducted using the R programming language.*

Keywords: Cauchy family of distribution, Estimation methods, Failure rate function, Inverse Gompertz distribution, Survival function

1 Introduction

The Gompertz model stands as one of the extensively employed probability models, drawing from the laws of mortality to formulate its survival function. It finds application in analyzing lifetime data related to human mortality and behavioral sciences data, enabling exploration of actuarial tables. Originally introduced by Gompertz [18], the Gompertz distribution has found utility not only in mortality analysis but also as a growth model and in fitting tumor growth patterns. By leveraging the Gompertz model's function, a significant array of life table data can be synthesized into a single function. This function facilitates the calculation of the number of individuals expected to survive to a specific age, as a function of their live age. This calculation involves utilizing the Gompertz function, which provides insights into the likelihood of individuals reaching various ages. By applying this model, researchers and actuaries can gain a deeper understanding of mortality patterns and make informed decisions regarding insurance, healthcare, and other related fields. The Gompertz model's versatility and accuracy make it a valuable tool in demographic analysis and risk assessment, contributing to advancements in public health and social policy. The article presented here is the formulation and study of a new probability model. Here we have discussed some important characteristics of the probability distribution relating to proposed model. Model is based on the scientific study as the proposed model and their features follow the properties of scientific probability models. This study will help the researchers in study of probability as well as formulation of new models in future. In this study, we have discussed a real data set and the proposed model may be suitable to analyze other real-life datasets.

The Gompertz distribution is particularly suited for modeling phenomena where the risk of failure or mortality increases exponentially with age or time. With a scale parameter of α and a shape parameter of λ, let x be the non-negative random variable that follows the Gompertz distribution. Consequently, the Gompertz distribution's cumulative distribution function (CDF) is provided by

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

However, in numerous practical fields like lifetime analysis, behavioral sciences, and describing adult mortality patterns for insurance purposes, there is a clear need for more comprehensive versions of these categories.

This is because traditional distributions frequently fall short in accurately capturing real-world data across different practical situations. As a result, there is a rising interest in developing more flexible models, with one notable approach being the "Inverse Distribution." Eliwa et al. [15] introduced new distribution called the inverse Gompertz distribution (IGD) with an upside-down bathtub-shaped hazard rate function. Let x be the non-negative random variable. Then the distribution function of Inverse Gompertz distribution (IGD) described by Eliwa et al.[15] is given by

$$H_{IGD}(x) = exp\left\{\frac{\lambda}{\alpha}\left(1 - e^{\alpha/x}\right)\right\}; x > 0, \alpha > 0, \lambda > 0 \tag{2}$$

The Gompertz distribution has been extensively explored and applied, as detailed in the work of Ahuja and Nash [6]. Gompertz-Sinh distributions are a family of distributions that Cooray and Ananda [12] established and used to examine reliability data with distributions that were extremely negatively skewed. The flexible model known as the generalized Gompertz distribution was introduced by El-Gohary et al. [14]. Depending on the shape parameter, this model can display failure rates that are declining, increasing, constant, or bathtub-shaped. Abu-Zinadah and Al-Oufi [2] have investigated various estimation methods for the exponentiated Gompertz distribution. Additionally, the inverse generalized Gompertz distribution has been introduced by Chaudhary and Kumar [9]. Telee and Kumar [32] developed new distribution called Half-Cauchy Gompertz distribution by compounding Gompertz distribution with Half- Cauchy family of distribution. The Kumaraswamy inverse Gompertz distribution is an expansion of the inverse Gompertz distribution first presented by El-Morshedy et al. [16]. It includes a hazard rate function shaped like an upside-down bathtub, determined by its shape parameters. This extension provides a refined framework for modeling data with characteristics that align with the inverse Gompertz distribution, offering valuable insights and applications in various fields. Chaudhary et al. [10] created the Half-Cauchy inverse Gompertz distribution by combining the inverse Gompertz distribution with the Half-Cauchy family of distributions. This improves the accuracy of utilizing the inverse Gompertz distribution to fit real data. Some further Gompertz distribution extensions found in the literature are the gamma-Gompertz distribution [28], the transmuted generalized Gompertz distribution [22], the Marshall Olkin exponential Gompertz distribution [21], the exponentialized generalized inverted Gompertz distribution [17]. On the inverse power Gompertz distribution [1], Bayesian and classical estimations of transmuted inverse Gompertz distribution [4], the inverted Gompertz-Fréchet distribution [5], Topp-Leone inverse Gompertz distribution [3]. A generalized Gompertz distribution with hazard power parameter and its bivariate extension [25]. New generalized Weibull inverse Gompertz distribution [7]. On the moments of the 3-parameter Gompertz distribution [20] and a weighted Gompertz-G family of distributions for reliability and lifetime data analysis [13]. Chaudhary et al. [11] also created novel model called Cauchy modified generalized exponential distribution. To enhance the accuracy of fitting real data with the inverse Gompertz distribution, we propose the Cauchy modified inverse Gompertz distribution. This alternative distribution offers greater flexibility with two additional parameters. We analyze the characteristics of the Cauchy modified inverse Gompertz distribution and demonstrate its suitability and potential applications.

The following are the discussion of the remaining sections: in section 2, we present a new distribution known as the Cauchy modified inverse Gompertz distribution and explore some statistical features. We use three widely used estimation techniques to estimate the parameters listed in section 3, namely, LSE, CVM, and MLE. A real-time dataset is used in section 4 to illustrate the applicability of the proposed model. Here, we assess the goodness-of-fit attained by the Cauchy modified inverse Gompertz distribution using statistics like AIC, AICC, HQIC, and BIC. Lastly, in section 5, we present some conclusions drawn from our analysis.

2 Model Formulation

Adding an extra scale parameter β to modify the inverse Gompertz distribution (2), resulting a new probability distribution called modified inverse Gompertz distribution (MIG) is provided by

$$G_{MIG}(x) = exp\left\{\frac{\lambda}{\alpha}\left(1 - \exp\left(\frac{\alpha e^{-\beta x}}{x}\right)\right)\right\}; x > 0, \alpha > 0, \lambda > 0, \beta > 0 \tag{3}$$

Here, we have used Cauchy family of distribution to introduce a new probability distribution called Cauchy Modified Inverse Gompertz (CMIG) distribution taking distribution function $G_{MIG}(x)$ of the modified inverse Gompertz distribution as base function. Distribution function of the Cauchy family of distribution is provided by

$$F(x) = 1 - \frac{2}{\pi} \arctan\left(-\frac{1}{\theta} \log G_{MIG}(x)\right); \; x > 0, \theta > 0 \tag{4}$$

The new distribution called Cauchy modified inverse Gompertz (CMIG) distribution is formulated by combining Cauchy family distribution with modified inverse Gompertz distribution as the base line distribution. The CMIG distribution's cumulative distribution function is obtained by applying $G_{MIG}(x)$ to $F(x)$ in equation (4). Given a non-negative random variable x that follows $CMIG$ (α, β, λ, θ), the CDF can be found by

$$F(x; \alpha, \beta, \lambda, \theta) = 1 - \frac{2}{\pi} \arctan\left\{\frac{\lambda}{\alpha\theta}\left(1 - \exp\left(\frac{\alpha e^{-\beta x}}{x}\right)\right)\right\}; \; x > 0, \alpha, \beta, \lambda, \theta > 0 \tag{5}$$

The suggested model CMIG's corresponding density function is provided by

$$\begin{aligned} f(x; \alpha, \beta, \lambda, \theta) &= \frac{2\lambda}{\pi\theta x^2} e^{-\beta x} (1 + \beta x) \exp\left(\frac{\alpha e^{-\beta x}}{x}\right) \\ &\left[1 + \left\{\frac{\lambda}{\alpha\theta}\left(1 - \exp\left(\frac{\alpha e^{-\beta x}}{x}\right)\right)^2\right\}\right]^{-1}; \; x > 0, \alpha, \beta, \lambda, \theta > 0 \end{aligned} \tag{6}$$

2.1 Survival function

The CMIG's survival function is

$$S(x) = \frac{2}{\pi} \arctan\left\{\frac{\lambda}{\alpha\theta}\left(1 - e^{\alpha e^{-\beta x}/x}\right)\right\}; \; x > 0, \alpha, \beta, \lambda, \theta > 0. \tag{7}$$

2.2 Hazard rate function

The CMIG's Hazard rate function is

$$\begin{aligned} h(x) &= \frac{\lambda}{\theta x^2} e^{-\beta x} (1 + \beta x) \exp\left(\frac{\alpha e^{-\beta x}}{x}\right) \\ &\left[\arctan\left\{\frac{\lambda}{\alpha\theta}\left(1 - \exp\left(\frac{\alpha e^{-\beta x}}{x}\right)\right)\right\}\left\{1 + \left\{\frac{\lambda}{\alpha\theta}\left(1 - \exp\left(\frac{\alpha e^{-\beta x}}{x}\right)\right)^2\right\}\right\}\right]^{-1}; \; x > 0, \alpha, \beta, \theta > 0 \end{aligned} \tag{8}$$

2.3 Reversed hazard rate function

The Reversed hazard rate function is

$$\begin{aligned} h_{rev}(x) &= \frac{2\lambda}{\pi\theta x^2} e^{-\beta x} (1 + \beta x) e^{\alpha e^{-\beta x}/x} \left[1 + \left\{\frac{\lambda}{\alpha\theta}\left(1 - \exp\left(\frac{\alpha e^{-\beta x}}{x}\right)\right)^2\right\}\right]^{-1} \\ &\left[1 - \frac{2}{\pi} \arctan\left\{\frac{\lambda}{\alpha\theta}\left(1 - \exp\left(\frac{\alpha e^{-\beta x}}{x}\right)\right)\right\}\right]^{-1}; \; x > 0, \alpha, \beta, \lambda, \theta > 0 \end{aligned} \tag{9}$$

2.4 Cumulative hazard rate function

: The suggested model's cumulative hazard rate function is given by

$$H(x) = -\log\left[\frac{2}{\pi} \arctan\left\{\frac{\lambda}{\alpha\theta}\left(1 - \exp\left(\frac{\alpha e^{-\beta x}}{x}\right)\right)\right\}\right] \tag{10}$$

The various shapes of PDF and HRF of with various possible parameter values at $\alpha = 56.71$ and θ =0.0065 are displayed in Figure [1]. The PDF graph illustrates a distribution that peaks once with a positive skew. In contrast, the hazard rate function plot demonstrates a trend of increasing followed by decreasing.

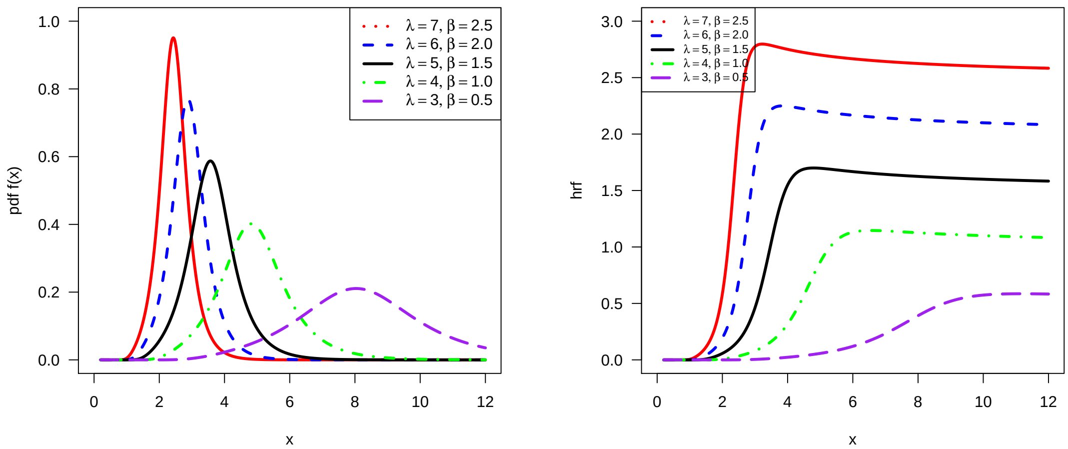

Figure 1: Probability density curves (Left) and hazard rate function (Right).

2.5 Quantile function

Assuming X is a non-negative random variable with CDF $F_x(x)$, the quantile function is

$$Q_X(p) = {F_X}^{-1}(p)$$

$$Q_x(p) = \alpha e^{-\beta x} - x \log\left[1 + \left(\frac{\alpha\theta}{\lambda}\right)\tan(1-p)\left(\frac{\pi}{2}\right)\right];\ 0 \leq p \leq 1 \tag{11}$$

The random deviate generation is

$$\alpha e^{-\beta x} - x \log\left[1 + \left(\frac{\alpha\theta}{\lambda}\right)\tan(1-u)\left(\frac{\pi}{2}\right)\right] = 0;\ 0 < u < 1 \tag{12}$$

2.6 Skewness and Kurtosis

The quantile coefficient of measure of symmetry is

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

Expression for testing normality by using octiles of the data [24] is

$$K - Moors = \left[Q\left(\frac{3}{8}\right) - Q\left(\frac{5}{8}\right) - Q\left(\frac{1}{8}\right) + Q\left(\frac{7}{8}\right)\right]\left[Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)\right]^{-1}$$

3 Parameter Estimation

Maximum Likelihood Estimation

In this subsection, we have shown the CMIG model's MLE approach for parameter estimation. Let $\underline{x}$ =

$(x_1, \ldots, x_n)$ be sample with n items from CMIG. Log likelihood function of CMIG is

$$
\begin{aligned}
l(x;\alpha,\beta,\lambda,\theta) = {} & n\log\left(\frac{2\lambda}{\pi\theta}\right) + 2\sum_{i=1}^{n}\log x_i - \beta\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\log\left(1+\beta x\right) + \alpha\sum_{i=1}^{n} e^{-\beta x}/x \\
& - \sum_{i=1}^{n}\log\left[1+\left\{\frac{\lambda}{\alpha\theta}\left(1-e^{\alpha e^{-\beta x}/x}\right)^2\right\}\right]
\end{aligned}
\tag{13}
$$

The MLE of the $CMIG(\alpha,\beta,\lambda,\theta)$ can be found by taking the first order derivative of (13) with respect to α, β, λ, and θ, and equating $\frac{\partial l}{\partial \alpha} = \frac{\partial l}{\partial \lambda} = \frac{\partial l}{\partial \beta} = \frac{\partial l}{\partial \theta} = 0$, and solving for the α, β, λ, and θ simultaneously. These non-linear equations are normally unsolvable, but we can solve them with ease if we use the suitable computer language. Let $\underset{-}{\Delta} = (\alpha,\beta,\lambda,\theta)$ be the parameter vector and the related maximum likelihood estimation for be represented by $\underset{-}{\hat{\Delta}} = (\hat{\alpha},\hat{\beta},\hat{\lambda},\hat{\theta})$, then resulting asymptotic normality is $\left(\underset{-}{\hat{\Delta}} - \underset{-}{\Delta}\right) \to N_4\left[0, \left(K\left(\underset{-}{\Delta}\right)\right)^{-1}\right]$. Here, Fisher's information matrix is denoted by $K\left(\underset{-}{\Delta}\right)$ which is given by

$$
K\left(\underset{-}{\Delta}\right) = -E\begin{pmatrix}
E\left(\frac{\partial^2 l}{\partial\alpha^2}\right) & E\left(\frac{\partial^2 l}{\partial\alpha\partial\beta}\right) & E\left(\frac{\partial^2 l}{\partial\alpha\partial\lambda}\right) & E\left(\frac{\partial^2 l}{\partial\alpha\partial\theta}\right) \\
E\left(\frac{\partial^2 l}{\partial\beta\partial\alpha}\right) & E\left(\frac{\partial^2 l}{\partial\beta^2}\right) & E\left(\frac{\partial^2 l}{\partial\beta\partial\lambda}\right) & E\left(\frac{\partial^2 l}{\partial\beta\partial\theta}\right) \\
E\left(\frac{\partial^2 l}{\partial\alpha\partial\lambda}\right) & E\left(\frac{\partial^2 l}{\partial\beta\partial\lambda}\right) & E\left(\frac{\partial^2 l}{\partial\lambda^2}\right) & E\left(\frac{\partial^2 l}{\partial\theta\partial\lambda}\right) \\
E\left(\dfrac{\partial^2 l}{\partial\alpha\partial\theta}\right) & E\left(\dfrac{\partial^2 l}{\partial\beta\partial\theta}\right) & E\left(\dfrac{\partial^2 l}{\partial\lambda\partial\theta}\right) & E\left(\dfrac{\partial^2 l}{\partial\theta^2}\right)
\end{pmatrix}
$$

We actually do not know $\underset{-}{\Delta}$, thus the asymptotic variance $\left(K\left(\underset{-}{\Delta}\right)\right)^{-1}$ of the MLE is worthless. Therefore, we approximate the asymptotic variance by substituting the calculated parameter values. An approximation of the information matrix $K\left(\underset{-}{\Delta}\right)$ is given by the observed Fisher information matrix, denoted by the symbol $O(\underset{-}{\hat{\Delta}})$. Thus, we approximate the asymptotic variance using the calculated parameter values

$$
O(\underset{-}{\hat{\Delta}}) = -\begin{pmatrix}
\left(\frac{\partial^2 l}{\partial\alpha^2}\right) & \left(\frac{\partial^2 l}{\partial\alpha\partial\beta}\right) & \left(\frac{\partial^2 l}{\partial\alpha\partial\lambda}\right) & \left(\frac{\partial^2 l}{\partial\alpha\partial\theta}\right) \\
\left(\frac{\partial^2 l}{\partial\beta\partial\alpha}\right) & \left(\frac{\partial^2 l}{\partial\beta^2}\right) & \left(\frac{\partial^2 l}{\partial\beta\partial\lambda}\right) & \left(\frac{\partial^2 l}{\partial\beta\partial\theta}\right) \\
\left(\frac{\partial^2 l}{\partial\alpha\partial\lambda}\right) & \left(\frac{\partial^2 l}{\partial\beta\partial\lambda}\right) & \left(\frac{\partial^2 l}{\partial\lambda^2}\right) & \left(\frac{\partial^2 l}{\partial\theta\partial\lambda}\right) \\
\left(\dfrac{\partial^2 l}{\partial\alpha\partial\theta}\right) & \left(\dfrac{\partial^2 l}{\partial\beta\partial\theta}\right) & \left(\dfrac{\partial^2 l}{\partial\lambda\partial\theta}\right) & \left(\dfrac{\partial^2 l}{\partial\theta^2}\right)
\end{pmatrix}_{|(\hat{\alpha},\hat{\beta},\hat{\lambda})} = -H\left(\underset{-}{\Delta}\right)_{|\left(\underset{-}{\Delta}=\underset{-}{\hat{\Delta}}\right)},
$$

where H denotes the Hessian.

To maximize the likelihood, the Newton-Raphson method constructs the observed information matrix. This method iteratively updates the parameter estimates by leveraging the information contained in the observed data matrix, ultimately aiming to optimize the likelihood function. Consequently, it is possible to provide the matrix for variance-covariance through

$$
\left[-H\left(\underset{-}{\Delta}\right)_{|\left(\underset{-}{\Delta}=\underset{-}{\hat{\Delta}}\right)}\right]^{-1} = \begin{pmatrix}
\mathrm{A11} & \mathrm{A12} & \mathrm{A13} & \mathrm{A14} \\
\mathrm{A21} & \mathrm{A22} & \mathrm{A23} & \mathrm{A24} \\
\mathrm{A31} & \mathrm{A32} & \mathrm{A33} & \mathrm{A34} \\
\mathrm{A41} & \mathrm{A42} & \mathrm{A43} & \mathrm{A44}
\end{pmatrix}
$$

where, $A11 = Var\ of\ \hat{\alpha}$, $A22 = Var\ of\ \hat{\beta}$, $A33 = Var\ of\ \hat{\lambda}$, $A44 = Var\ of\ \hat{\theta}$

$A12 =$ Co-var $of\ \hat{\alpha}\ \&\ \hat{\beta}$, $A13 =$ Co-var $of\ \hat{\alpha}\ \&\ \hat{\lambda}$, $A14 =$ Co-var $of\ \hat{\alpha}\ \&\ \hat{\theta}$

$A21 =$ Co-var $of\ \hat{\beta},\ \&\ \hat{\alpha}$, $A23 =$ Co-var $of\ \hat{\beta}\ \&\ \hat{\lambda}$, $A24 =$ Co-var $of\ \hat{\beta}\ \&\ \hat{\theta}$

$A31 =$ Co-var $of\ \hat{\lambda},\ \&\ \hat{\alpha}$, $A32 =$ Co-var $of\ \hat{\lambda}\ \&\ \hat{\beta}$, $A34 =$ Co-var $of\ \hat{\lambda}\ \&\ \hat{\theta}$

$A41 =$ Co-var $of\ \hat{\theta},\ \&\ \hat{\alpha}$, $A42 =$ Co-var $of\ \hat{\theta}\ \&\ \hat{\beta}$, $A43 =$ Co-var $of\ \hat{\theta}\ \&\ \hat{\lambda}$

If $Z_{\delta/2}$ is upper percentile for standard normal distribution, then the following method can be employed to estimate confidence intervals, with a confidence level of 100(1-δ)% for computing α, β, λ, and θ. This approach relies on the asymptotic normality of maximum likelihood estimators (MLEs). The confidence intervals are estimated by $\hat{\alpha} \pm Z_{\delta/2}\sqrt{\text{var}(\hat{\alpha})}$, $\hat{\beta} \pm Z_{\delta/2}\sqrt{[\text{var}(\hat{\beta})]}$, $\hat{\lambda} \pm Z_{\delta/2}\sqrt{[\text{var}(\hat{\lambda})]}$ and $\hat{\theta} \pm Z_{\delta/2}\sqrt{[\text{var}(\hat{\theta})]}$ respectively.

Least-Square Estimation

The function is minimized when the unknown parameters of the CMIG distribution, α, β, λ, and θ, are estimated via least-square estimation

$$B\left(X;\alpha,\beta,\lambda,\theta\right) = \sum_{i=1}^{n}\left[F(X_i) - \frac{i}{n+1}\right]^2 \tag{14}$$

with respect to α, β, λ, and θ. Let us consider $\{X_1, X_2, \ \ldots, X_n\}$ be a random sample with n units from ordered random variable $X_{(1)} < X_{(2)} < \ \ldots < X_{(n)}$ having CDF as $\lambda =$ the LSE of α, β, λ, and θ can be calculated by minimizing the relation (15) with respect to α, β, λ, and θ

$$B\left(X;\alpha,\beta,\lambda,\theta\right) = \sum_{i=1}^{n}\left[1 - \frac{2}{\pi}\arctan\left\{\frac{\lambda}{\alpha\theta}\left(1 - e^{\alpha e^{-\beta x_{(i)}}/x_{(i)}}\right)\right\} - \frac{i}{n+1}\right]^2 \tag{15}$$

The estimated parameters can be found by differentiating the relation (16) with respect to α, β, λ, and θ, and then solving the simultaneous equation. We can minimize following relation to get weighted least square estimators as

$$B\left(X;\alpha,\beta,\lambda,\theta\right) = \sum_{i=1}^{n} w_i\left[1 - \frac{2}{\pi}\arctan\left\{\frac{\lambda}{\alpha\theta}\left(1 - e^{\alpha e^{-\beta x_{(i)}}/x_{(i)}}\right)\right\} - \frac{i}{n+1}\right]^2, \tag{16}$$

where w_i are weight and is given as

$$w_i = \frac{1}{Var(X_{(i)})} = \frac{(n+1)^2\,(n+2)}{i\,(n-i+1)}$$

We can estimate unknown parameters α, β, λ, and θ using the weighted least square technique by minimizing relation (17) with respect to α, β, λ, and θ

$$B\left(X;\alpha,\ \beta,\lambda,\theta\right) = \ \sum_{i=1}^{n}\frac{(n+1)^2\,(n+2)}{i\,(n-i+1)}\left[1 - \frac{2}{\pi}\arctan\left\{\frac{\lambda}{\alpha\theta}\left(1 - e^{\alpha e^{-\beta x_{(i)}}/x_{(i)}}\right)\right\} - \frac{i}{n+1}\right]^2 \tag{17}$$

Cramer-Von Mises estimation

By minimizing relation (18) we can get estimated values of unknown parameters α, β, λ, and θ using Cramer-Von Mises technique of estimation

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

$$= \frac{1}{12n} + \sum_{i=1}^{n}\left[1 - \frac{2}{\pi}\arctan\left\{\frac{\lambda}{\alpha\theta}\left(1 - e^{\alpha e^{-\beta x_{(i)}}/x_{(i)}}\right)\right\} - \frac{2i-1}{2n}\right]^2 \tag{18}$$

The CVM estimators of the parameters can be obtained by differentiating equation (18) with respect to α, β, λ, and θ and by solving the non-linear equations

$$\frac{\partial C}{\partial \alpha} = 0, \quad \frac{\partial C}{\partial \beta} = 0, \quad \frac{\partial C}{\partial \lambda} = 0 \text{ and } \frac{\partial C}{\partial \theta} = 0$$

4 Application to Real Dataset

Data set

This segment includes a study of an real dataset that validates the suggested model. Because atoms in circuits migrate about and create microcircuit failures, electromigration can occasionally occur in a circuit. 59 conductors were used in the accelerated life test [26, 29], where without any observational censoring, failure time in hours was determined as follows

6.545, 4.700, 7.543, 9.289, 6.492, 6.956, 8.120, 5.459, 8.687, 4.706, 8.591, 2.997, 11.038, 6.129, 6.958, 5.381, 6.522, 4.288, 7.459, 4.137, 6.573, 7.495, 5.589, 6.538, 6.725, 5.807, 9.663, 8.532, 7.024, 6.369, 9.218, 8.336, 6.869, 7.945, 6.087, 6.352, 9.254, 6.948, 7.489, 5.009, 6.033, 7.398, 7.496, 10.092, 8.799, 7.974, 7.224, 7.683, 6.923, 7.365, 5.434, 5.640, 6.515, 7.937, 6.071, 10.491, 6.476 ,4.531, 5.923.

Exploratory Data Analysis (EDA)

Exploratory data analysis is crafted by combining various statistical analyses that elucidate and summarize the dataset used in research. Gaining a thorough understanding of the data set being used is the aim of this. In addition to the graphical displays of the data, it could also contain some descriptive statistics. The key metrics that can be incorporated into EDA are listed below.

- The density curve, histogram, boxplot, and so forth are the graphical plots that aid in identifying patterns in the data as well as identifying any odd patterns or findings.
- Certain aspects and characteristics of the data are revealed by measurements such as location, scatter, skewness, kurtosis, and so forth.

The data summary was found using the R programming language, and the results are shown in the table below.

The data is positively skewed and non-normal, according to summary statistics.

Table 1: Descriptive statistics of the data

Min	Q1	Median	Mean	Q3	Max	Skewness	Kurtosis
2.997	6.052	6.923	6.980	7.9417	11.038	0.1931723	3.087389

The boxplot, histogram, and density fit of the suggested CMIG model are shown in Figure [2]. The data set's boxplot and TTT plot are also shown in Figure [2]. Using R software (R Core Team[27]) having optim () function, by maximizing the likelihood function (13), calculations of MLEs of CMIG model is done. We got the value of Log-Likelihood as l = -111.2442. Table [2] contains the estimated parameter values of α, β, λ, and θ. The Probability-Probability plot on the left and the Quantile-Quantile plot on the right of

Table 2: Estimated parameters

Methods	Alpha	Beta	Lambda	Theta
MLE	56.7169	0.7480	7.7269	0.0065
LSE	51.9078	0.7001	7.8161	0.0093
CVME	51.9078	0.7297	7.8161	0.0075

Figure[3] have also been included to assess the model's normality. Using above estimated parameter values by all three methods we have calculated log-likelihood and AIC criterion and are presented in Table [3].

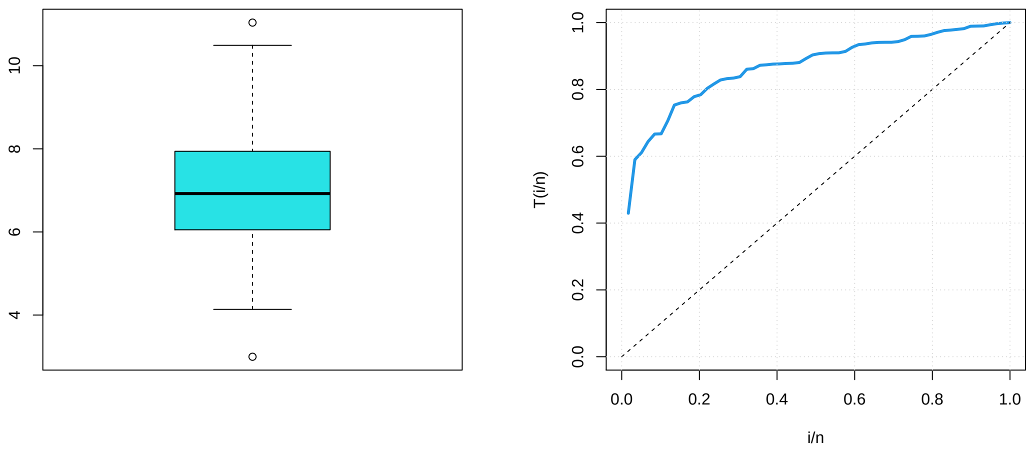

Figure 2: Boxplot (Left) and TTT plot (Right).

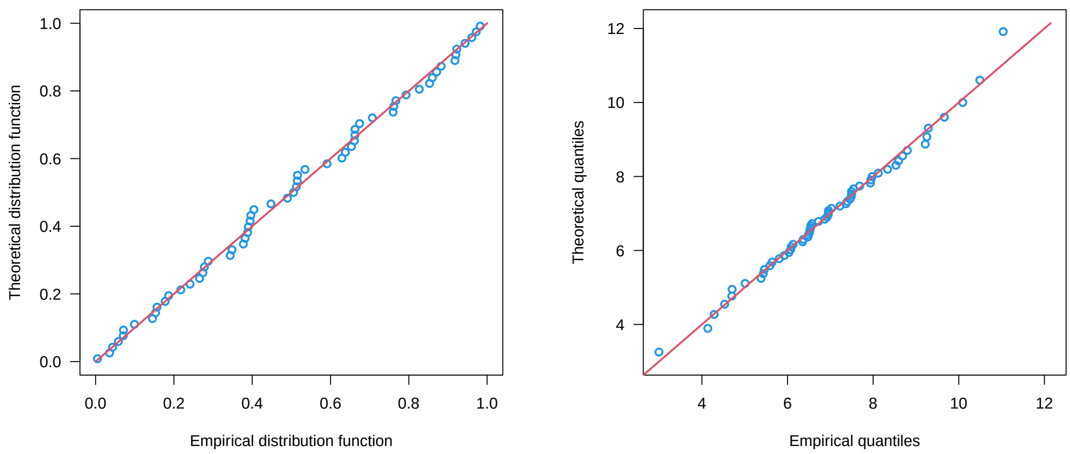

Figure 3: P-P plot (Left) and Q-Q plot (Right).

Table 3: Log likelihood (LL), AIC, BIC, CAIC and HQIC

Model	LL	AIC	BIC	CAIC	HQIC
MLE	-111.259	230.518	238.828	231.259	233.762
LSE	-111.324	230.647	230.647	231.388	233.891
CVM	-111.259	230.518	238.828	231.259	233.762

Table [4] presents the A^2, W and KS statistics together with the corresponding p-values of the LSE, MLE, and CVE estimations for evaluating the model's goodness of fit.

The suggested model CMIG distribution's empirical CDF, fitted CDF plot, density function and histogram of fitted distributions of estimation techniques utilizing LSE, MLE, and CVM are displayed in Figure [4].

Table 4: KS, W, and (A^2) statistics with corresponding p-values

Methods	KS(p-value)	W(p-value)	A^2(p-value)
MLE	0.0531(0.9932)	0.0199(0.9974)	0.1321(0.9995)
LSE	0.0488(0.9977)	0.0205 (0.9967)	0.1329(0.9995)
CVME	0.0515(0.9953)	0.0194(0.9978)	0.1264(0.9997)

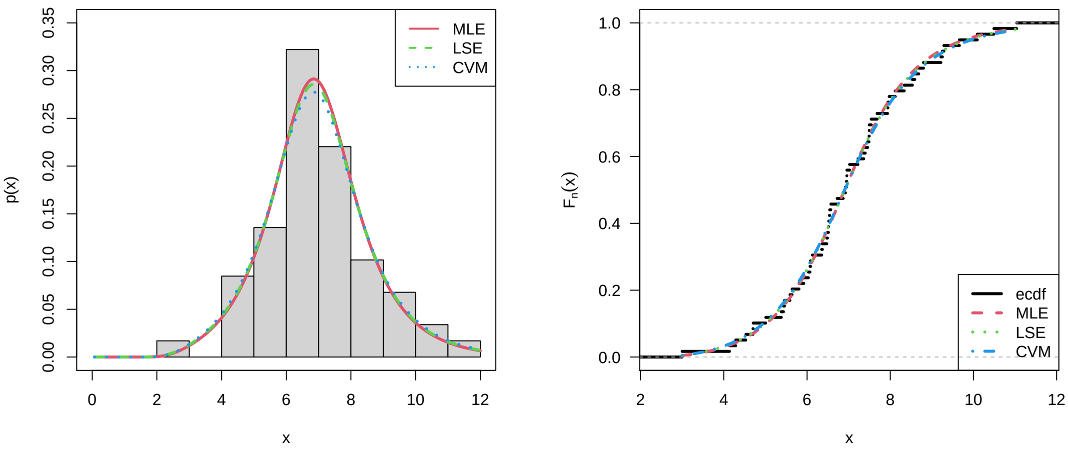

Figure 4: Histogram versus density curve (Left) and ECDF versus CDF (Right) for different estimation methods.

Model Comparison

Here, we've evaluated how well the CMIG model applies compared to other models used by researchers with the same dataset. Here we have chosen five distributions to compare the potentiality of the CMIG. Five distributions considered are Exponential Power (EP) model [30], Generalized Exponential (GE) model [19], Modified Weibull (MW) Model [23], Weibull Extension (WE) Model [31]), and Lindley-Exponential (LE) model [8]. The estimated parameters of both the competing models under investigation and the suggested model are shown in Table[5]. We also displayed the histogram of the dataset, graph for goodness-of-fit of

Table 5: Estimated parameters models

Model	Alpha	Beta	Lambda	Theta
CMIG	56.7170	0.7480	7.7270	0.0066
WE	32.2529	4.6823	26.6786	-
GE	52.4066	-	0.6424	-
LE	-	-	0.6422	53.3296
EP	3.14011	0.1138	-	-
MW	0.0070	0.2216	0.5782	-

CMIG model along with models taken in considerations in left panel of Figure[5]. In right panel of the graph, empirical CDF and fitted CDF are shown. From plotting it is observed that CMIG distribution fits the dataset better than other models taken in consideration. Examples of the HQIC, BIC, AIC, and CAIC can be used to determine the applicability of the CMIG distribution. These values are presented in Table[6]. Validation criteria obtained clearly show that model CMIG has fewer values than the value of other considered models. This suggests that proposed model fit data better than other models taken in considerations.

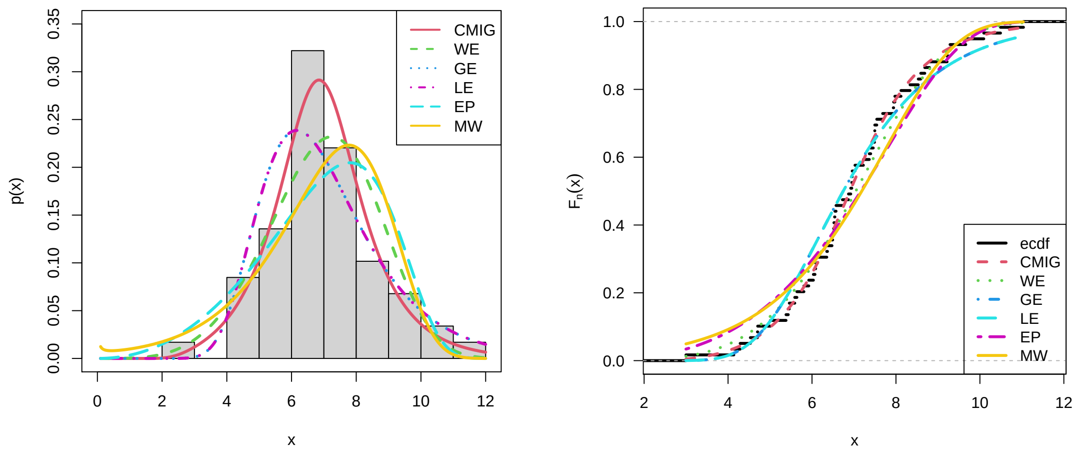

Figure 5: Histogram versus density curves (Left) and ECDF versus CDFs (Right) models.

Table 6: Log likelihood (LL), AIC, BIC, CAIC and HQIC

Model	LL	AIC	BIC	CAIC	HQIC
CMIG	-111.2590	230.5180	238.8284	231.2591	233.7620
WE	-112.5054	231.0108	237.2434	231.447	233.4437
GE	-114.9471	233.8946	238..0497	234.109	235.5166
LE	-114.9525	233.9055	238.0606	234.1198	235.5275
EP	-116.5015	237.0029	241.1580	237.2090	238.6249
MW	-117.3520	240.7041	246.9367	241.1404	243.1370

Testing of goodness of fit of fitted model and comparing it with other model is essential part of statistical modeling. We generated the statistics for the Anderson-Darling (AD), Cramer-Von Mises (CVM), and Kolmogorov-Smirnov (KS) tests in order to assess the goodness-of-fit of the CMIG distribution in comparison to alternative distributions. These test statistic values are displayed in Table[7]. It is found that the CMIG distribution has lesser values of the test statistics and higher p-value in every method of goodness of fit. This concludes that the CMIG distribution fits real data set more consistently compared to other models taken in consideration.

Table 7: KS, W, and A^2 statistics with p-values

Model	KS	W	A^2
CMIG	0.0531(0.9932)	0.0199(0.9973)	0.1321(0.9995)
WE	0.0964(0.6089)	0.0858(0.6606)	0.4845(0.7619)
GE	0.1042(0.5103)	0.1173(0.5079)	0.7368(0.5282)
LE	0.1042(0.5099)	0.1173(0.5077)	0.7373(0.5279)
EP	0.1365(0.2021)	0.2398(0.2021)	1.3735(0.2098)
MW	0.1345(0.2160)	0.2236(0.2263)	1.3349(0.2213)

5 Conclusion

The Cauchy Modified Inverse Gompertz Distribution is a newly introduced distribution that is featured here. Some fundamental statistical properties of the new model are thoroughly examined. We have derived some important expression for the hazard rate, survival, the quantile function etc. We have also calculated summary of the data as mean, median skewness and kurtosis etc. and are displayed. For estimating the unknown parameters, three important and more effective methods of estimation as MLE, LSE and CVME are used. In this case, we discovered that the CVM approach outperforms the MLE and LSE approaches. PDF curve of the proposed model CMIG has different shaped density curve for different values of parameters showing that curve is positively skewed having increasing-decreasing hazard rate curve. Here, we have also taken into account a real dataset and have compared the applicability and suitability of the suggested model with a few other models that have been introduced by previous researchers utilizing the same dataset and parameter estimation techniques. The suggested model outperforms the other models that were considered in terms of data fit. We also considered different methods of model validation techniques to test the validity showing that proposed model has lesser values of information criteria than competing models. So, we can conclude that model CMIG will play significant role in data analysis and new model formulations. The model CMIG is original and new probability model so it can be useful in data analysis and probability model formulation. The main limitation of the study is that we have applied on a single data set. To generalize the model application of model on different data sets may be required.

References

[1] Abdelhady, D. H., and Amer, Y. M., 2021, On the inverse power Gompertz distribution, *Annals of Data Science*, 8(3), 451-473. https://doi.org/10.1007/s40745-020-00246-4

[2] Abu-Zinadah, H. H., and Al-Oufi, A. S., 2016, Different method estimations of the three parameters of exponentiated Gompertz distribution, *Applied Mathematics and Information Sciences*, 10(2), 705-710. http://dx.doi.org/10.18576/amis/100230

[3] Adegoke, T. M., Oladoja, O. M., Bashiru, S. O., Mustapha, A. A., Aderupatan, D. E., and Nzei, L. C., 2023, Topp-Leone inverse Gompertz distribution: Properties and different estimations techniques and applications, *Pakistan Journal of Statistics*, 39(4),433.

[4] Adegoke, T. M., Obisesan, K. O., Oladoja, O. M., and Adegoke, G. K., 2023, Bayesian and classical estimations of transmuted inverse Gompertz distribution, *Reliability: Theory and applications*, 18(2 (73)), 207-222. http://dx.doi.org/10.24412/1932-2321-2023-273-207-222

[5] Akarawak, E. E., Adeyeye, S. J., Khaleel, M. A., Adedotun, A. F., Ogunsanya, A. S., and Amalare, A. A., 2023, The inverted Gompertz-Fréchet distribution with applications, *Scientific African*, 21, e01769. https://doi.org/10.1016/j.sciaf.2023.e01769

[6] Ahuja, J. C., and Nash, S. W., 1967, The generalized Gompertz-Verhulst family of distributions, Sankhyā: *The Indian Journal of Statistics*, Series A, 141-156. https://www.jstor.org/stable/25049460

[7] Baharith, L. A., 2024, New generalized Weibull inverse Gompertz distribution: Properties and applications, *Symmetry*, 16(2), 197. https://doi.org/10.3390/sym16020197

[8] Bhati, D., Malik, M. A., and Vaman, H. J., 2015, Lindley–exponential distribution: Properties and applications, *Metron*, 73, 335-357. https://doi.org/10.1007/s40300-015-0060-9

[9] Chaudhary, A. K., and Kumar, V., 2017, Inverse generalized Gompertz distribution with properties and applications, *Journal of National Academy of Mathematics*, 31, 99-108.

[10] Chaudhary, A. K., Yadav, R. S., and Kumar, V., 2022, Half-Cauchy inverse Gompertz distribution: Theory and applications, *International Journal of Statistics and Applied Mathematics*, 7(5), 94-102. https://doi.org/10.22271/maths.2022.v7.i5b.885

[11] Chaudhary, A. K., Telee, L. B. S., and Kumar, V., 2024, Cauchy modified generalized exponential distribution: Estimation and applications, *International Journal of Statistics and Applied Mathematics*, 9(2): 56-65. https://doi.org/10.22271/maths.2024.v9.i2a.1682

[12] Cooray, K., and Ananda, M. M., 2010, Analyzing survival data with highly negatively skewed distribution: The Gompertz-Sinh family, *Journal of Applied Statistics*, 37(1), 1-11. https://doi.org/10.1080/02664760802663072

[13] Eghwerıdo, J. T., and Nzeı, L., 2024, A weighted Gompertz-G family of distributions for reliability and lifetime data analysis, *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 73(1), 235-258. https://doi.org/10.31801/cfsuasmas.1210841

[14] El-Gohary, A., Alshamrani, A., and Al-Otaibi, A. N., 2013, The generalized Gompertz distribution, *Applied Mathematical Modelling*, 37(1-2), 13-24. https://doi.org/10.1016/j.apm.2011.05.017

[15] Eliwa, M. S., El-Morshedy, M., and Ibrahim, M., 2019, Inverse Gompertz distribution: Properties and different estimation methods with application to complete and censored data, *Annals of data science*, 6, 321-339. https://doi.org/10.1007/s40745-018-0173-0

[16] El-Morshedy, M., El-Faheem, A. A., and El-Dawoody, M., 2020, Kumaraswamy inverse Gompertz distribution: Properties and engineering applications to complete, type-II right censored and upper record data, *Plos one*, 15(12), e0241970. https://doi.org/10.1371/journal.pone.0241970

[17] El-Morshedy, M., El-Faheem, A. A., Al-Bossly, A., and El-Dawoody, M., 2021, Exponentiated generalized inverted Gompertz distribution: Properties and estimation methods with applications to symmetric and asymmetric data, *Symmetry*, 13(10), 1868. https://doi.org/10.3390/sym13101868

[18] Gompertz, B., 1825, On the nature of the function expressive of the law of human mortality and on a new mode of determining the value of life contingencies, *Philosophical Transactions of the Royal Society London*, 115, 513-585. https://doi.org/10.1098/rstl.1825.0026

[19] Gupta, R. D., and Kundu, D., 1999, Theory and methods: Generalized exponential distributions, *Australian & New Zealand Journal of Statistics*, 41(2), 173-188. https://doi.org/10.1111/1467-842X.00072

[20] Hakiki, M., Adipradana, D., Ahmad, I., and Putri, L., 2024, On the moments of the 3-Parameter Gompertz distribution, Barekeng: *Journal Ilmu Matematika Dan Terapan*, 18(2), 1023-1036. https://doi.org/10.30598/barekengvol18iss2pp1023-1036

[21] Khaleel, M. A., Al-Noor, N. H., and Abdal-Hameed, M. K., 2020, Marshall Olkin exponential Gompertz distribution: Properties and applications, *Periodicals of Engineering and Natural Sciences*, 8(1), 298-312. http://dx.doi.org/10.21533/pen.v8i1.1152

[22] Khan, M. S., King, R., and Hudson, I. L., 2017, Transmuted generalized Gompertz distribution with application, *Journal of Statistical Theory and Applications*, 16(1), 65-80. https://doi.org/10.2991/jsta.2017.16.1.6

[23] Lai, C. D., Xie, M., and Murthy, D. N. P., 2003, A modified Weibull distribution, *IEEE Transactions on reliability*, 52(1), 33-37. https://doi.org/10.1109/TR.2002.805788

[24] Moors, J. J. A., 1988, A quantile alternative for kurtosis, *Journal of the Royal Statistical Society: Series D (The Statistician)*, 37 (1), 25-32. https://doi.org/10.2307/2348376

[25] Muhammed, H. Z., 2024, A generalized Gompertz distribution with hazard power parameter and its bivariate Extension: Properties and applications, *Annals of Data Science*, 11(2), 759-784. https://doi.org/10.1007/s40745-022-00420-w

[26] Nelson, W., and Doganaksoy, N., 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. https://doi.org/10.1201/9781003418313-21

[27] R Core Team, 2023, R: *A language and environment for statistical computing*, (Version 4.1) [Computer software], retrieved from `https://cran.r-project.org` (R packages retrieved from CRAN snapshot 2023-04-07).

[28] Shama, M. S., Dey, S., Altun, E., and Afify, A. Z., 2022, The gamma–Gompertz distribution: Theory and applications, *Mathematics and Computers in Simulation*, 193, 689-712. `https://doi.org/10.1016/j.matcom.2021.10.024`

[29] Schafft, H. A., Station, T. C., Mandel, J., and Shott, J. D., 1987, Reproducibility of electro-migration measurements, *IEEE Transactions on Electron Devices*, ED-34, 673-681. `https://doi.org/10.1109/T-ED.1987.22979`

[30] Smith, R. M., and Bain, L. J., 1975, An exponential power life-testing distribution, *Communications in Statistics-Theory and Methods*, 4(5), 469-481. `https://doi.org/10.1080/03610927508827263`

[31] Tang, Y., Xie, M., and Goh, T. N., 2003, Statistical analysis of a Weibull extension model, *Communications in Statistics-Theory and Methods*, 32(5), 913-928. `https://doi.org/10.1081/STA-120019952`

[32] Telee, L. B. S., and Kumar, V., 2019, Half-Cauchy Gompertz distribution: Different methods of estimation, *Journal of National Academy of Mathematics*, 33, 51-65.