iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** IX **Month of publication:** September 2020

DOI: https://doi.org/10.22214/ijraset.2020.31745

International Journal For Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IX Sep 2020- Available AT www.ijraset.com

Lindley Half Cauchy Distribution: Properties and Applications

Arun Kumar Chaudhary[1], Vijay Kumar[2]
[1]Department of Management Science, Nepal Commerce Campus, Tribhuwan University,
[2]Department of Mathematics and Statistics, DDU Gorakhpur University,

Abstract: A new two-parameter Lindley half-Cauchy distribution using Lindley family of distribution is introduced. The mathematical and statistical properties of the new distribution such as probability density function, cumulative distribution function, quantiles, the measure of skewness and kurtosis are presented. The parameter of the new distribution is estimated using three widely used estimation methods namely maximum likelihood estimators (MLE), least-square (LSE) and Cramer-Von-Mises (CVM) methods. By using the maximum likelihood estimate, we have constructed the asymptotic confidence interval for the model parameters. A real data set is taken and we have compared LHC distribution with some selected distributions namely weighted Lindley, Chen, Gompertz, and Lindley. It is proven empirically that the proposed distribution is more flexible and performs better than underling distributions.
Keywords: Contour plot, Chen distribution, Estimation, Gompertz distribution, Half-Cauchy distribution, Lindley distribution, Weighted Lindley distribution.

I. INTRODUCTION

The statistics literature is filled with hundreds of continuous univariate distributions. Many classical distributions have been widely used over the past decades for modeling data in several areas such as actuarial, environmental and medical sciences, life sciences, demography, economics, finance, and insurance. However, in many applied sectors like survival analysis, insurance and finance, there is a clear necessity for modified forms of more flexible distributions to model real data that can address a high degree of skewness and kurtosis.

The half-Cauchy distribution is derived from the Cauchy distribution by reflecting the curve on the origin so that only positive values can be detected. Since it can predict more common long-distance dispersal events, the half-Cauchy distribution has been used as an substitute to model dispersal distances (Shaw, 1995), as a heavy-tailed distribution. Furthermore, Paradis *et al.* (2002) implemented the half-Cauchy distribution to model ringing data on two species of tits in Ireland and Britain. Its cumulative distribution function (CDF) and probability density function (PDF) respectively are,

$$G(x) = \frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right), \quad \lambda > 0, \; x > 0 \tag{1.1}$$

and

$$g(x) = \frac{2}{\pi}\left(\frac{\lambda}{\lambda^2 + x^2}\right), \quad \lambda > 0, \; x > 0 \tag{1.2}$$

Last few decades the half-Cauchy distribution has been used by many researchers as a baseline distribution. Cordeiro & Lemonte (2011) has introduced the beta-half-Cauchy distribution, Jacob & Jayakumar (2012) has presented the "On half-Cauchy distribution and process", and Polson & Scott (2012) have made an extensive study on "On the half-Cauchy prior for a global scale parameter". The Kumaraswamy-half-Cauchy distribution was introduced by (Ghosh, 2014). Alzaatreh et al. (2016) has introduced the gamma half-Cauchy distribution. Cordeiro et al (2017) has created the generalized odd half-Cauchy family of distributions. Hence we are motivated to introduce Lindley half-Cauchy distribution.

The one parameter Lindley distribution was developed by (Lindley, 1958) in the context of Bayesian statistics, as a counterexample to fiducial statistics. In recent years, many studies have been focused to obtain various modified forms of the baseline distribution using Lindley family presented by Zografas and Balakrishnan (2009) with more flexible density and hazard rate functions. A detailed study on the Lindley distribution was done by (Ghitany et al., 2008).

Consider a random variable X follows Lindley distribution with parameter θ and its probability density function (PDF) is given by

$$f(x)=\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x};\ x>0,\theta>0 \tag{1.3}$$

And its cumulative distribution function (CDF) is

$$F(x)=1-\frac{1+\theta+\theta x}{1+\theta}e^{-\theta x};\ x>0,\theta>0 \tag{1.4}$$

Ghitany et al. (2008a) has made some of the modifications in the literature of Lindley distribution, which is quite similar to the exponential distribution. Gupta and Singh (2013), used hybrid censored data to investigate the estimation of the parameters. The estimation of the model parameters for censored samples by (Krishna & Kumar, 2011), Reyes et al. (2019) has introduced the Slash Lindley-Weibull Distribution, and Hassan & Nassr (2019) has created the Power Lindley-G family of distributions. Ieren et al. (2020) has introduced the Odd Lindley-Rayleigh distribution and its properties and applications to simulated and real life datasets.
The main objective of this study is to launch a new probability model by inserting only one additional parameter and hence it is more flexible. The rest of the article is organized as follows. In Section 2, the proposed Lindley half-Cauchy distribution is derived and we obtain some properties of the LHC distribution such as a reliability function, hazard rate function, quantile function, and skewness and kurtosis. In Section3, we introduced the different methods for estimating the model parameters namely maximum likelihood estimators (MLE), least-square (LSE) and Cramer-Von-Mises (CVM) methods. The estimation of the model parameters for uncensored data is discussed in Section 4. Also we have compared the LHC distribution with some selected weighted Lindley, Chen, Gompertz, and Lindley distribution. Finally, conclusions about the proposed model are presented in Section 5.

II. THE LINDLEY HALF CAUCHY (LHC) DISTRIBUTION

Zografas and Balakrishnan (2009) has defined the CDF of any new modified distribution is

$$F(x)=\int_0^{-\ln[1-G(x)]} r(t)\ dt \tag{2.1}$$

Using $r(t)$ as PDF of Lindley distribution (1.3) and the baseline distribution G(x) as CDF of half-Cauchy distribution (1.1) then the CDF of Lindley half-Cauchy is obtained as,

$$F(x)=1-\left\{1-\frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right\}^{\theta}\left\{1-\left(\frac{\theta}{1+\theta}\right)\ln\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right]\right\};\ \theta,\lambda>0,\ x>0 \tag{2.2}$$

and its corresponding PDF is obtained as,

$$f(x)=\frac{2}{\pi}\left(\frac{\theta^2}{1+\theta}\right)\left(\frac{\lambda}{\lambda^2+x^2}\right)\left\{1-\frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right\}^{\theta-1}\left\{1-\ln\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right]\right\};\ \theta,\lambda>0,\ x>0 \tag{2.3}$$

The reliability/survival function of Lindley half-Cauchy distribution is

$$R(x)=1-F(x)$$

$$=\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right]^{\theta}\left\{1-\left(\frac{\theta}{1+\theta}\right)\ln\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right]\right\};\ \theta,\lambda>0,\ x>0 \tag{2.4}$$

And its hazard rate function is

$$h(x)=\frac{f(x)}{R(x)}=\frac{2\theta^2\lambda\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right]}{\pi(\lambda^2+x^2)\left\{\theta+\left\{1-\ln\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right]\right\}^{-1}\right\}};\ \theta,\lambda>0,\ x>0 \tag{2.5}$$

International Journal For Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IX Sep 2020- Available AT www.ijraset.com

Figure 1 displays the plots for the PDF and hazard function of LHC distribution for numerous values of the parameters λ and θ.

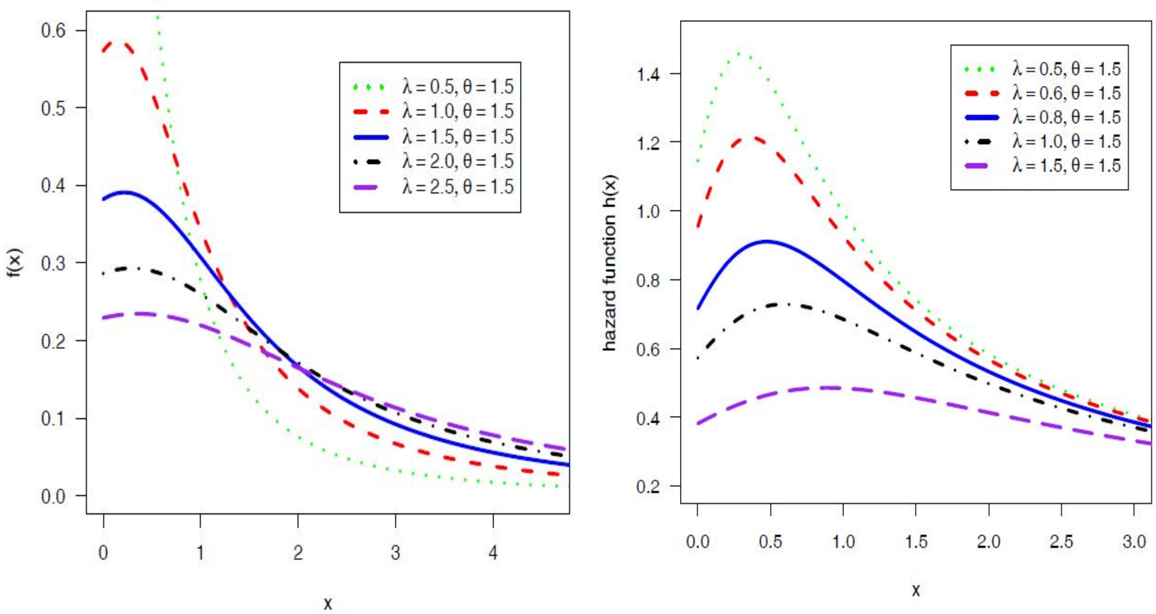


Figure 1. Graph of PDF (left panel) and hazard function (right panel) for different values of λ and θ.

A. Quantile Function of LHC Distribution

In statistics and probability, the quantile function, related with a probability distribution of a random variable, identifies the value of the random variable such that the probability of the variable being less than or equal to that value equals the given probability. It is also named the inverse cumulative distribution function or percent-point function.

$$Q(p) = F^{-1}(p)$$

Hence the quantile function can be written as,

$$p - 1 + \left\{1 - \frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right\}^{\theta}\left\{1 - \left(\frac{\theta}{1+\theta}\right)\ln\left[1 - \frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right]\right\} = 0\ ;\ 0 < p < 1 \qquad (2.6)$$

The random numbers can be generated for the LHC distribution, for this let, simulating values of random variable X with the CDF (2.2). Let U denote a uniform random variable in (0,1), then the simulated values of X are obtained by setting,

$$u - 1 + \left\{1 - \frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right\}^{\theta}\left\{1 - \left(\frac{\theta}{1+\theta}\right)\ln\left[1 - \frac{2}{\pi}\tan^{-1}\left(\frac{x}{\lambda}\right)\right]\right\} = 0\ ;\ 0 < u < 1 \qquad (2.7)$$

and solving for x.

B. Skewness and Kurtosis

These measures are used mostly in data analysis to study the shape of the distribution or data set. Skewness and Kurtosis based on quantile function are

$$Skewness(B) = \frac{Q(0.75) + Q(0.25) - 2Q(0.5)}{Q(3/4) - Q(1/4)}, \text{ and}$$

Coefficient of kurtosis based on octiles given by (Moors, 1988) is

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

International Journal For Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IX Sep 2020- Available AT www.ijraset.com

III.METHODS OF ESTIMATION

In statistics, the estimation theory is an important branch which deals with estimating the values of parameters based on measured empirical data which has a arbitrary component. The parameters describe an underlying physical setting in such a way that their value affects the distribution of the measured data. An estimator tries to approximate the unknown parameters with the help of the measurements. Commonly used estimators (estimation methods) are listed blew,

1) Maximum likelihood estimators (MLE)
2) Method of moments estimators
3) Minimum mean squared error (MMSE), also called Bayes least squared error (BLSE)
4) Cramer-von Mises estimator (CVM)
5) Least-squares estimators (LSE)
6) The maximum product of spacings (MPS) method
7) Cramér–Rao bound
8) Bayes estimators
9) Markov Chain Monte Carlo (MCMC)

We have considered different estimation procedures for the unknown parameters of the LHC distribution. We introduce three types of estimators such as the maximum likelihood (MLE), ordinary least squares (LSE), and Cramer-von Mises (CVM) estimators.

A. *Maximum Likelihood Estimation (MLE)*

In this subsection, we discuss the maximum likelihood estimators (MLE's) of the LHC distribution.

Let $\underset{\sim}{x} = (x_1,\ldots..,x_n)$ denote a random sample of size 'n' from LHC(λ, θ), then the likelihood function L(λ, θ/ $\underset{\sim}{x}$) can be expressed as,

$$L(\lambda,\theta\,|\,\underset{\sim}{x}) = \frac{2}{\pi}\left(\frac{\theta^2}{1+\theta}\right)\prod_{i=1}^{n}\left(\frac{\lambda}{\lambda^2+x_i^2}\right)\left\{1-\frac{2}{\pi}\tan^{-1}\left(\frac{x_i}{\lambda}\right)\right\}^{\theta-1}\left\{1-\ln\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{x_i}{\lambda}\right)\right]\right\} \qquad (3.1.1)$$

The log-likelihood density of (3.1) is

$$\begin{aligned} l &= 2n\ln\theta - n\ln(1+\theta) + n\ln 2 - n\ln\pi + n\ln\lambda - n\ln(\lambda^2+x^2) + \\ &(\theta-1)\sum_{i=1}^{n}\ln\left\{1-\frac{2}{\pi}\tan^{-1}\left(\frac{x_i}{\lambda}\right)\right\} + \sum_{i=1}^{n}\ln\left\{1-\ln\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{x_i}{\lambda}\right)\right]\right\} \end{aligned} \qquad (3.1.2)$$

Differentiating (3.2) with respect to λ and θ we get,

$$\frac{\partial l}{\partial\theta} = \frac{2n}{\theta} - \frac{n}{1+\theta} + \sum_{i=1}^{n}\ln\left\{1-\frac{2}{\pi}\tan^{-1}\left(\frac{x_i}{\lambda}\right)\right\} \qquad (3.1.3)$$

$$\frac{\partial l}{\partial\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n}\frac{2n\lambda}{(\lambda^2+x_i^2)} - \sum_{i=1}^{n}\frac{2\lambda(\theta-1)}{\pi(\lambda^2+x_i^2)A(x_i)} + \sum_{i=1}^{n}\frac{2\lambda}{\pi(\lambda^2+x_i^2)\{1-\ln A(x_i)\}A(x_i)} \qquad (3.1.4)$$

$$\frac{\partial^2 l}{\partial\theta^2} = -\frac{2n}{\theta^2} + \frac{n}{(1+\theta)^2}$$

$$\begin{aligned} \frac{\partial^2 l}{\partial\lambda^2} &= -\frac{n}{\lambda^2} + 2\sum_{i=1}^{n}\frac{(\lambda^2-x_i^2)}{(\lambda^2+x_i^2)^2} + \frac{2}{\pi^2}\sum_{i=1}^{n}\frac{1}{(\lambda^2+x_i^2)^2A(x_i)}\times \\ &\left\{\{1-\ln A(x_i)\}^{-2} - \tfrac{1}{A(x_i)}\left[(\theta-1)+\{1-\ln A(x_i)\}^{-1}\right]\left[2\lambda^2-\pi(\lambda^2-x_i^2)\right]\right\} \end{aligned}$$

$$\frac{\partial^2 l}{\partial\theta\partial\lambda} = -\frac{2\lambda}{\pi}\sum_{i=1}^{n}\frac{1}{(\lambda^2+x_i^2)A(x_i)}$$

$$\frac{\partial^2 l}{\partial\lambda\partial\theta} = -\frac{2\lambda}{\pi}\sum_{i=1}^{n}\frac{1}{(\lambda^2+x_i^2)A(x_i)}$$

Where $A(x_i) = 1 - \frac{2}{\pi}\tan^{-1}\left(\frac{x_i}{\lambda}\right)$

By solving equations (3.3) and (3.4) we get the estimated values of the parameters of the Lindley half-Cauchy distribution. Since is difficult to solve them manually but one can use computer programming to solve them numerically. Consider $\underline{\Upsilon} = (\lambda, \theta)$ denote the parameter space and the corresponding MLE of $\underline{\Upsilon}$ as $\hat{\underline{\Upsilon}} = (\hat{\lambda}, \hat{\theta})$, then the asymptotic normality results in, $\left(\hat{\underline{\Upsilon}} - \underline{\Upsilon}\right) \rightarrow N_2\left[0, \left(I\left(\underline{\Upsilon}\right)\right)^{-1}\right]$ where $I\left(\underline{\Upsilon}\right)$ is the Fisher's information matrix defined as

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

In practice, $\left(I\left(\underline{\Upsilon}\right)\right)^{-1}$ the MLE has asymptotic variance is unknown because we don't know $\underline{\Upsilon}$. So, we approximate the asymptotic variance by substituting the estimated value of the parameters.

The common procedure is to use the observed Fisher information matrix $O(\hat{\underline{\Upsilon}})$ as an estimate of the information matrix $I\left(\underline{\Upsilon}\right)$ given by

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

where H stands for Hessian matrix.

We are using the Newton-Raphson algorithm in order to maximize the likelihood and give the observed information matrix. Therefore, the variance-covariance matrix is given by,

$$\left[-H\left(\underline{\Upsilon}\right)_{|(\underline{\Upsilon} = \hat{\underline{\Upsilon}})}\right]^{-1} = \begin{pmatrix} \operatorname{var}(\hat{\lambda}) & \operatorname{cov}(\hat{\lambda}, \hat{\theta}) \\ \operatorname{cov}(\hat{\lambda}, \hat{\theta}) & \operatorname{var}(\hat{\theta}) \end{pmatrix} \quad (3.1.5)$$

Hence from the asymptotic normality of MLEs, approximate 100(1-α) % confidence intervals for λ and θ can be constructed as, $\hat{\lambda} \pm Z_{\alpha/2} SE(\hat{\lambda})$ and $\hat{\theta} \pm Z_{\alpha/2} SE(\hat{\theta})$ where $Z_{\alpha/2}$ is the upper percentile of standard normal variate.

B. Method of Least-Square Estimation (LSE)

The ordinary least square estimators and weighted least square estimators are established by (Swain et al., 1988) to estimate the parameters of Beta distributions. The least-square estimators of the unknown parameters λ and θ of the NHC distribution can be attained by minimizing

$$\omega\left(X; \lambda, \theta\right) = \sum_{i=1}^{n}\left[G(X_i) - \frac{i}{n+1}\right]^2 \quad (3.2.1)$$

with respect to unknown parameters λ and θ.

Let $G(X_{(i)})$ represent the cumulative distribution function of the ordered random variables $X_{(1)} < X_{(2)} < \ldots\ldots < X_{(n)}$, where $\left\{X_1, X_2, \ldots\ldots, X_n\right\}$ is a random sample of size n from a CDF G(.). Therefore, the least square estimators of λ and θ say $\hat{\lambda}$ and $\hat{\theta}$ respectively, can be obtained by minimizing

$$\omega\left(X; \lambda, \theta\right) = \sum_{i=1}^{n}\left[1 - \left\{1 - \frac{2}{\pi}\tan^{-1}\left(\frac{x_i}{\lambda}\right)\right\}^{\theta}\left\{1 - \left(\frac{\theta}{1+\theta}\right)\ln\left[1 - \frac{2}{\pi}\tan^{-1}\left(\frac{x_i}{\lambda}\right)\right]\right\} - \frac{i}{n+1}\right]^2 \quad (3.2.2)$$

International Journal For Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IX Sep 2020- Available AT www.ijraset.com

with respect to λ and θ.

To obtain the least square estimators, we have to solve the following two nonlinear equations equating to zero,

$$\frac{\partial\omega}{\partial\lambda}=-2\sum_{i=1}^{n}\left[1-\{U_i\}^{\theta}\left\{1-\left(\frac{\theta}{1+\theta}\right)\ln[U_i]\right\}-\frac{i}{n+1}\right]$$

$$\left\{\left[\frac{2\theta U_i}{\pi}\frac{x_i}{\lambda^2+x_i^2}\right]\left[1-\left(\frac{\theta}{1+\theta}\right)\ln U_i\right]-[U_i]^{\theta}\left[\frac{2\theta}{\pi(1+\theta)U_i}\frac{x_i}{\lambda^2+x_i^2}\right]\right\}$$

$$\frac{\partial\omega}{\partial\theta}=-2\sum_{i=1}^{n}\left[1-\{U_i\}^{\theta}\left\{1-\left(\frac{\theta}{1+\theta}\right)\ln[U_i]\right\}-\frac{i}{n+1}\right]\left\{U_i^{\theta}\ln(U_i)\left[1-\left(\frac{\theta}{1+\theta}\right)\ln U_i\right]-[U_i]^{\theta}\frac{\ln U_i}{(1+\theta)^2}\right\}$$

Where $U_i=1-\frac{2}{\pi}\tan^{-1}\left(\frac{x_i}{\lambda}\right)$

C. Method of Cramer-Von-Mises (CVM)

One of the important estimation methods is Cramér-von-Mises type minimum distance estimators, (Macdonald 1971) because it provides empirical evidence that the bias of the estimator is smaller than the other minimum distance estimators. The CVM estimators of λ and θ are attained by minimizing the function,

$$C(\lambda,\theta)=\frac{1}{12n}+\sum_{j=1}^{n}\left[F(x_{j:n}\mid\lambda,\theta)-\frac{2j-1}{2n}\right]^2 \qquad (3.3.1)$$

$$=\frac{1}{12n}+\sum_{j=1}^{n}\left[1-\left\{1-\frac{2}{\pi}\tan^{-1}\left(\frac{x_j}{\lambda}\right)\right\}^{\theta}\left\{1-\left(\frac{\theta}{1+\theta}\right)\ln\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{x_j}{\lambda}\right)\right]\right\}-\frac{2j-1}{2n}\right]^2$$

IV. APPLICATION TO A REAL DATASET

In this part, we have demonstrated the applicability of LHC distribution using a real data set used by previous researchers. We have compared the LHC distribution with the weighted Lindley, Chen, Gompertz, and Lindley distribution, which are listed below,

A. Gompertz Distribution (G)

The probability density function of Gompertz (Murthy et al., 2003) distribution with parameters α and θ is

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

B. Weighted Lindley Distribution (WL)

The WL distribution has introduced by (Ghitany et al., 2011) whose PDF is

$$f_{WL}(x)=\frac{\theta^{\alpha+1}}{(\alpha+\theta)\Gamma(\alpha)}x^{\alpha-1}(1+x)e^{-\theta x}\quad ;x\geq 0,\alpha>0,\theta>0.$$

C. Chen Distribution (C)

The probability density function (PDF) of Chain distribution (Chen, 2000) can be expressed as

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

International Journal For Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IX Sep 2020- Available AT www.ijraset.com

V. LINDLEY DISTRIBUTION

The probability density function (PDF) of Lindley distribution (Lindley, 1958) can be expressed as

$$f_{LD}(x) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x} \quad ; x \geq 0, \theta > 0.$$

By using MLE method we estimate the parameter of each of these distributions. For the goodness of fit purpose we use negative log-likelihood (-LL), Akaike information criterion (AIC), Bayesian information criterion (BIC), Corrected Akaike Information criterion (CAIC) and Hannan-Quinn information criterion (HQIC), statistic to select the best model among selected models. The expressions to calculate AIC, BIC, CAIC and HQIC are listed below:

A. $AIC = -2l(\hat{\theta}) + 2k$

B. $BIC = -2l(\hat{\theta}) + k\log(n)$

C. $CAIC = AIC + \frac{2k(k+1)}{n-k-1}$

D. $HQIC = -2l(\hat{\theta}) + 2k\log[\log(n)]$

where k is the number of parameters and n is the size of the sample in the model under consideration.

Further, in order to evaluate the fits of the LHC distribution with some selected distributions we have taken the Kolmogorov-Simnorov (KS), the Anderson-Darling (W) and the Cramer-Von Mises (A^2) statistic. These statistics are widely used to compare non-nested models and to illustrate how closely a specific CDF fits the empirical distribution of a given data set. These statistics are calculated as

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

$$W = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\ln d_i + \ln\left(1 - d_{n+1-i}\right)\right]$$

$$A^2 = \frac{1}{12n} + \sum_{i=1}^{n}\left[\frac{(2i-1)}{2n} - d_i\right]^2$$

where $d_i = CDF(x_i)$; the x_i's being the ordered observations.

In this segment, we demonstrate the applicability of LHC Distribution by considering a real dataset. The data presented below represents time interval between failures (in thousands of hours) of secondary reactor pumps (Suprawhardana, et.al, 1999):

0.062, 0.070, 0.101, 0.150, 0.199, 0.273, 0.347, 0.358, 0.402, 0.491, 0.605, 0.614, 0.746, 0.954, 1.060, 1.359, 1.921, 2.160, 3.465, 4.082, 4.992, 5.320, 6.560

In Figure 2 we have displayed the Contour plot and the fitted CDF with empirical distribution function (EDF) (Kumar & Ligges, 2011).

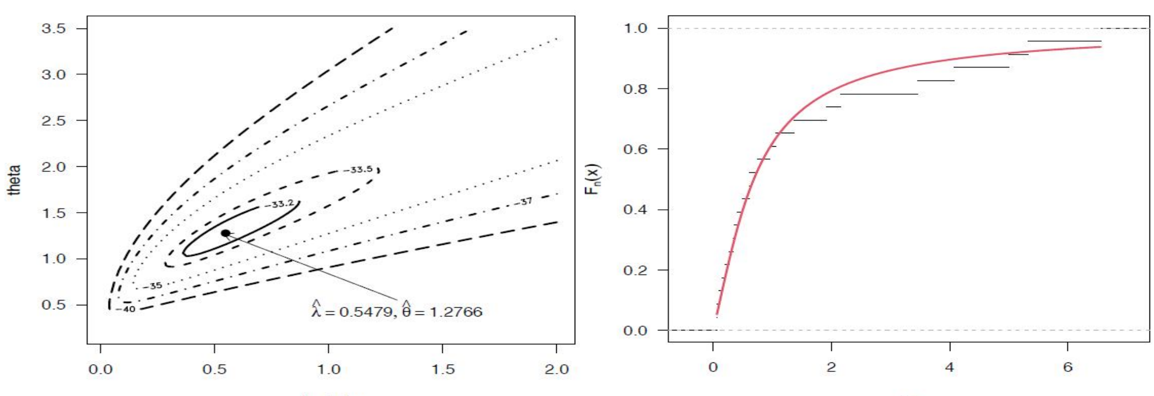


Figure 2. Contour plot (left panel) and the fitted CDF with empirical distribution function (right panel).

International Journal For Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IX Sep 2020- Available AT www.ijraset.com

We have used the log-likelihood function (3.1.2) to compute the MLE directly by using optim() function in R software (R Core Team, 2020). By using the maximum likelihood estimation method for the above data set, we have obtained $\hat{\lambda}$ = 0.5479 and $\hat{\theta}$ = 1.2766 and its corresponding Log-Likelihood value is -33.00769. In Table 1 we have presented the MLE's with their standard errors (SE) and 95% confidence interval for λ and θ.

Table 1
MLE, SE and 95% confidence interval

Parameter	MLE	SE	95% ACI	t-value	Pr(>t)
Lambda	0.5479	0.3830	(-0.20278, 1.29858)	1.430	0.15260
Theta	1.2766	0.4778	(0.340112, 2.213088)	2.672	0.00754

Hence the Hessian variance-covariance matrix is obtained as,

$$\left[-H(\underline{\Upsilon})\big|_{(\Upsilon=\hat{\Upsilon})}\right]^{-1}=\begin{pmatrix}\text{var}(\hat{\lambda}) & \text{cov}(\hat{\lambda},\hat{\theta})\\ \text{cov}(\hat{\lambda},\hat{\theta}) & \text{var}(\hat{\theta})\end{pmatrix}=\begin{pmatrix}0.14671 & 0.1653\\ 0.1653 & 0.2283\end{pmatrix}$$

In Figure 3, the Profile log-likelihood functions of parameters λ and θ are displayed. It can be concluded that the estimated parameters using the MLE method are unique.

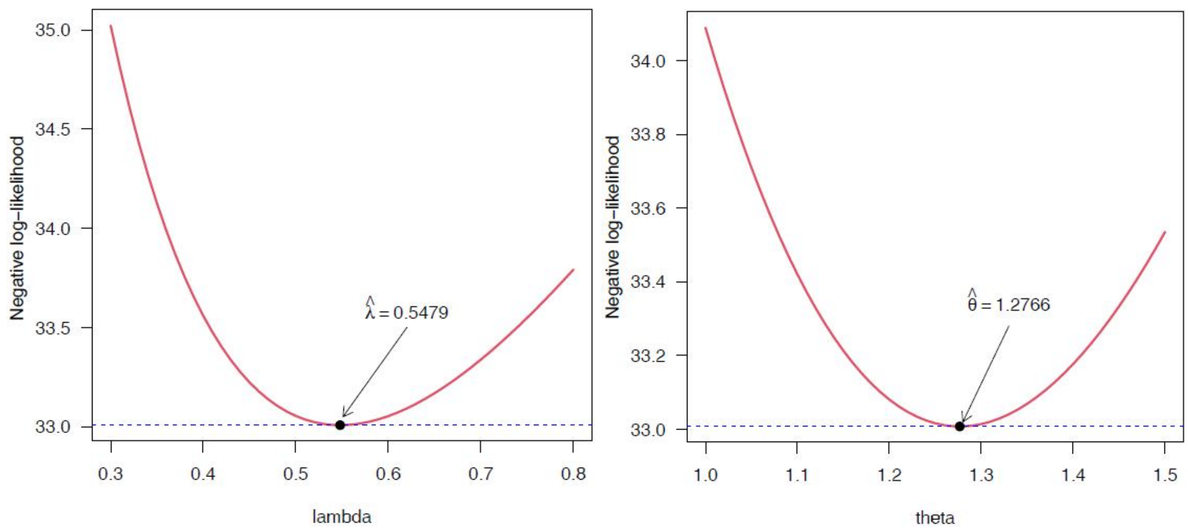


Figure 3. Plots of the Profile log-likelihood functions of the parameters λ and θ

The parameters are estimated by using the maximum likelihood method (MLE), ordinary least squares (LSE), and Cramer-von Mises (CVM) methods. For comparison, we use Negative Log-Likelihood values ($-LL$), the Akaike information criterion (AIC), Bayesian information criterion (BIC) and AICC which are defined by $-2LL+2p$, $-2LL+p\log(n)$, and $AIC + \{2p(1+p)\}/(n-p-1)$ respectively, where p is the number of parameters estimated and n is the sample size.

Table 2
Estimated parameters, log-likelihood, AIC, BIC and AICC

Method	$\hat{\alpha}$	$\hat{\theta}$	-LL	AIC	BIC	AICC
MLE	0.5479	1.2766	33.0077	70.01538	72.28637	70.56084
LSE	0.3155	0.9225	33.4547	70.90936	73.18035	71.45481
CVE	0.3765	1.0156	33.2205	70.44107	72.71205	70.98652

Table 3
Log-likelihood, AIC, BIC, CAIC and HQIC of selected model

Model	-LL	AIC	BIC	CAIC	HQIC
LHC	33.0077	70.0154	72.2864	70.6154	70.5865
G	33.4919	70.9838	73.2548	71.5838	71.5550
WL	33.6207	71.2414	73.5124	71.8414	71.8126
C	33.8402	71.6804	73.9514	72.2804	72.2515
L	35.3054	72.6108	73.7463	72.8013	72.8963

In Figure 4 we have presented the P-P plot (empirical distribution function against theoretical distribution function) and Q-Q plot (empirical quantile against theoretical quantile).

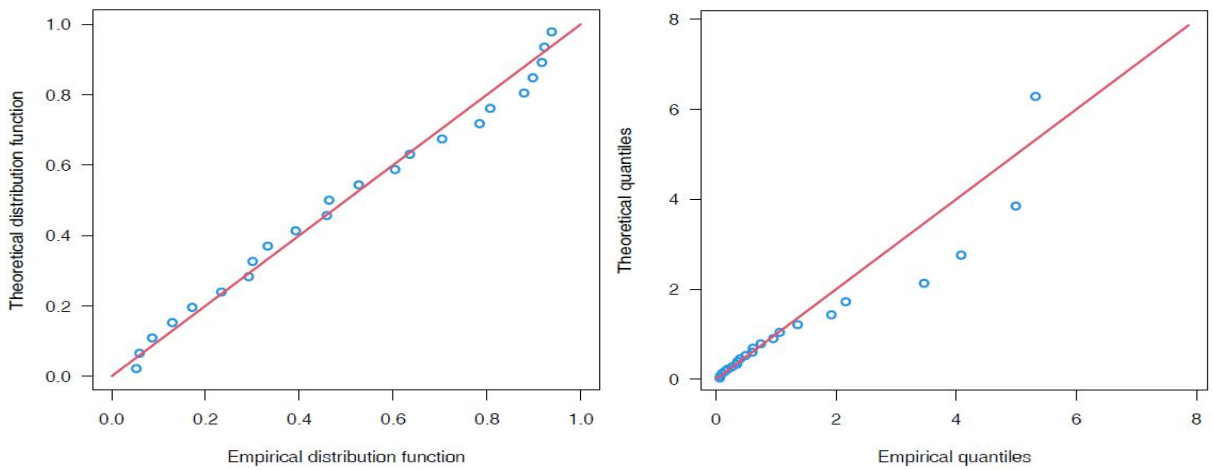


Figure 4. The graph of the P-P plot (left panel) and Q-Q plot (right panel)

The histogram and the fitted density functions and the empirical CDF with estimated CDF are displayed in Figure 5, which compares the distribution functions for the different models with the empirical distribution function produces the same. Therefore, for the given data set we have found that the proposed distribution gets better fit and more reliable results than selected ones.

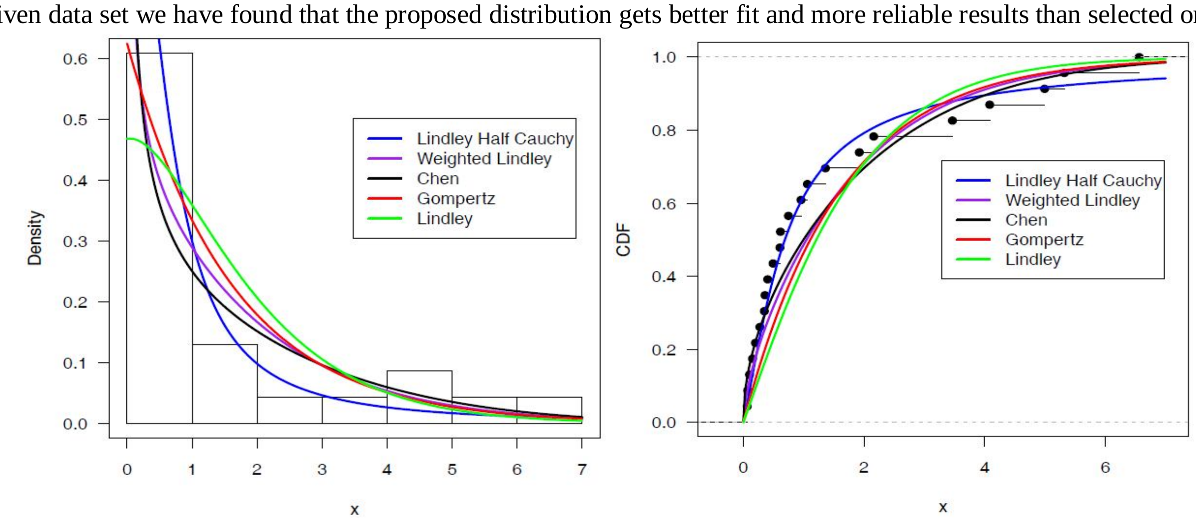


Figure 5. The Histogram and the PDF of fitted distributions (left panel) and Empirical CDF with estimated CDF (right panel).

We have displayed the test statistics and their corresponding p-value of competing models for a selected data set is displayed in Tabl e 4. The result shows that the proposed model has the minimum value of the test statistic and higher *p*-value hence we conclude that the LHC is best in the prospect of goodness-of-fit.

Table 4. The goodness-of-fit statistics and their corresponding *p*-value

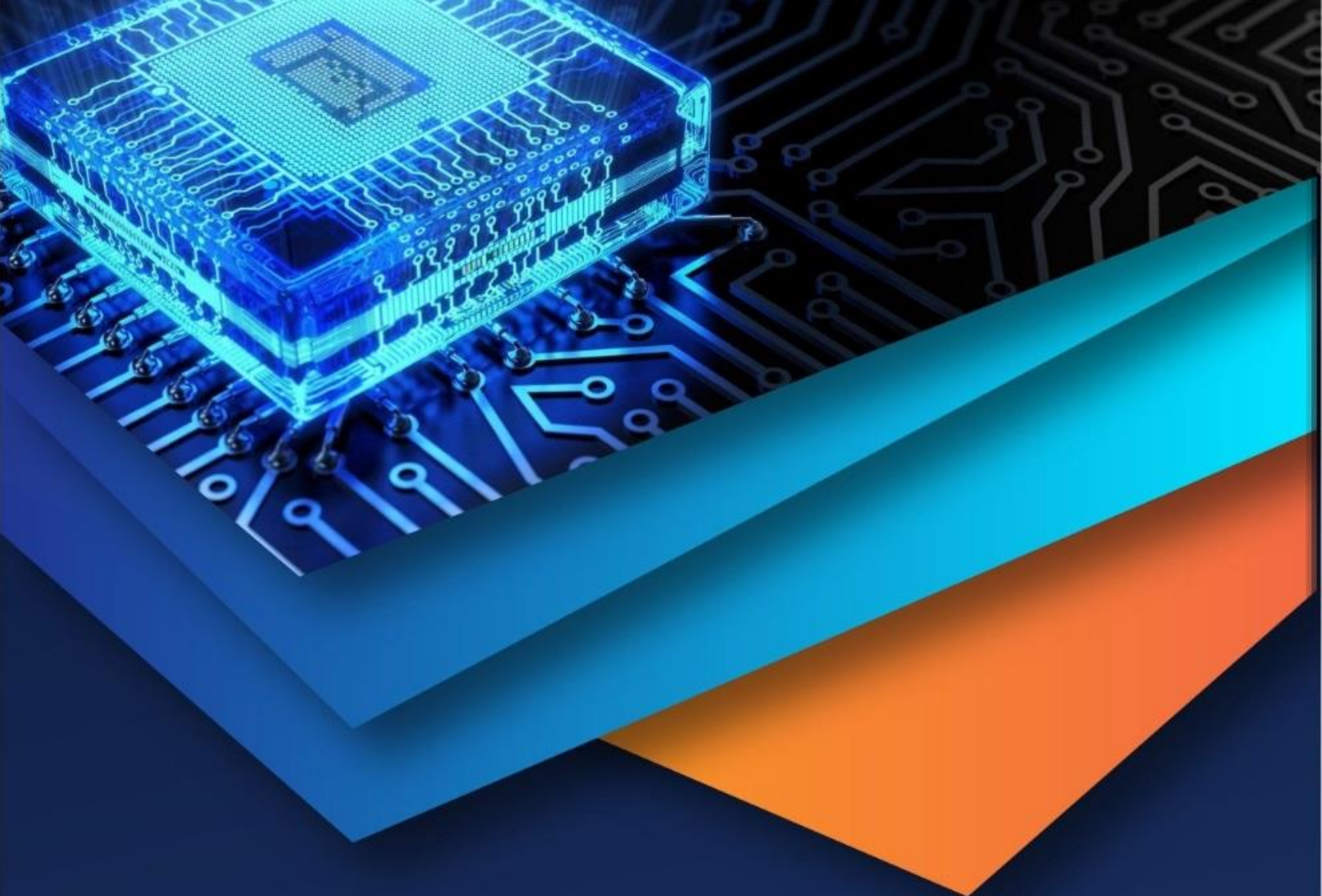
Model	*KS(p-value)*	*W(p-value)*	A^2*(p-value)*
LHC	0.0964(0.9692)	0.0287(0.9824)	0.2559(0.9670)
G	0.2030(0.2620)	0.2413(0.2002)	1.3404(0.2195)
WL	0.1606(0.5403)	0.1336(0.4462)	0.7583(0.5105)
C	0.1364(0.7356)	0.1024(0.5774)	0.6452(0.6045)
L	0.2441(0.1084)	0.3826(0.07960)	2.2994(0.0638)

VI. CONCLUDING REMARKS

We have introduced a novel lifetime model, called the Lindley half Cauchy distribution that extends the half-Cauchy distribution, and studied some of its general structural properties. Our expressions related to the LHC model are well manageable with the use of modern computer resources with analytic and numerical abilities. We have provided some mathematical treatment of the new distribution including expressions for the reliability function, hazard rate function, quantile function, skewness and kurtosis. The model parameters are estimated by using four well-known estimation methods namely maximum likelihood estimators (MLE), least-square (LSE) and Cramer-Von-Mises (CVM) methods. The helpfulness of the proposed model is demonstrated in an application to real data and also displays P-P plot and Q-Q plot for formal goodness-of-fit. We have concluded that the new model provides a consistently better fit than other competing models.

REFERENCES

[1] Alzaatreh, A., Mansoor, M., Tahir, M. H., Zubair, M., & Ghazali, S. A. (2016). The gamma half-Cauchy distribution: Properties and applications. Hacettepe Journal of Mathematics and Statistics, 45(4), 1143-1159.
[2] Chen, Z. (2000). A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. Statistics & Probability Letters, 49, 155-161.
[3] Cordeiro, G. M., & Lemonte, A. J. (2011). The beta-half-Cauchy distribution. Journal of Probability and Statistics, 2011.
[4] Cordeiro, G. M., Alizadeh, M., Ramires, T. G., & Ortega, E. M. (2017). The generalized odd half-Cauchy family of distributions: properties and applications. Communications in Statistics-Theory and Methods, 46(11), 5685-5705.
[5] Ghitany, M., Al-Mutairi, D. & Nadarajah, S. (2008a). Zero-truncated Poisson–Lindley distribution and its application. Math Comput Simulat, 79. 279-287.
[6] Ghitany, M.E., Atieh, B., & Nadarajah, S. (2008). Lindley distribution and its application. Math. Comput. Simu,. 78, 493–506.
[7] Ghitany, M. E., Alqallaf, F., Al-Mutairi, D. K., & Husain, H. A. (2011). A two-parameter weighted Lindley distribution and its applications to survival data. Mathematics and Computers in simulation, 81(6), 1190-1201.
[8] Ghosh, I. (2014). The Kumaraswamy-half-Cauchy distribution: properties and applications. Journal of Statistical Theory and Applications, 13(2), 122-134.
[9] Gupta, P.K. & Singh, B. (2013). Parameter estimation of Lindley distribution with hybrid censored data. Int J Syst Assur Eng Manag 4, 378-385.
[10] Hassan, A. S., & Nassr, S. G. (2019). Power Lindley-G family of distributions. Annals of Data Science, 6(2), 189-210.
[11] Ieren, T. G., Abdulkadir, S. S., & Issa, A. A. (2020). Odd Lindley-Rayleigh distribution: Its properties and applications to simulated and real life datasets. Journal of Advances in Mathematics and Computer Science, 63-88.
[12] Jacob, E., & Jayakumar, K. (2012). On half-Cauchy distribution and process. International Journal of Statistika and Mathematika, 3(2), 77-81.
[13] Krishna, H. & Kumar, K. (2011). Reliability estimation in Lindley distribution with progressively type II right censored sample. Math Comput Simulat, 82, 281-294.
[14] Kumar, V. and Ligges, U. (2011). reliaR : A package for some probability distributions. http://cran.r-project.org/web/packages/reliaR/index.html.
[15] Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. Journal of the Royal Statistical Society. Series B (Methodological), 102-107.
[16] Macdonald, P. D. M. (1971). Comments and Queries Comment on “An Estimation Procedure for Mixtures of Distributions” by Choi and Bulgren. Journal of the Royal Statistical Society: Series B (Methodological), 33(2), 326-329.
[17] Moors, J. J. A. (1988). A quantile alternative for kurtosis. Journal of the Royal Statistical Society: Series D (The Statistician), 37(1), 25-32.
[18] Murthy, D.N.P., Xie, M. and Jiang, R. (2003). Weibull Models, Wiley, New York.
[19] Paradis, E., Baillie, S.R, Sutherland, W.J. (2002). Modeling large-scale dispersal distances.Ecological Modelling, 151, 279–292.
[20] Polson, N. G., & Scott, J. G. (2012). On the half-Cauchy prior for a global scale parameter. Bayesian Analysis, 7(4), 887-902.
[21] R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
[22] Reyes, J., Iriarte, Y. A., Jodrá, P., & Gómez, H. W. (2019). The Slash Lindley-Weibull Distribution. Methodology and Computing in Applied Probability, 21(1), 235-251.
[23] Salman Suprawhardana M, Prayoto & Sangadji(1999). Total time on test plot analysis for mechanical components of the RSG-GAS reactor. Atom Indones ;25(2)
[24] Shaw, M.W. (1995). Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance. Proceedings of the Royal Society B: Biological Sciences, 259, 243–248.
[25] Swain, J. J., Venkatraman, S. & Wilson, J. R. (1988), ‘Least-squares estimation of distribution functions in johnson’s translation system’, Journal of Statistical Computation and Simulation 29(4), 271–297.
[26] Zografos, K., & Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference. Statistical Methodology, 6(4), 344-362.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY