

A new extension of Chris-Jerry distribution with its properties¹

C. SUBRAMANIAN AND M. SUBHASHREE

Abstract

In this article, we proposed a new probability distribution called length-biased Chris-Jerry distribution in which length-biased distribution is a specific case of weighted distribution. Also, we determined the structural properties of length-biased Chris-Jerry distribution and the unknown parameters were estimated by maximum likelihood estimator and tested by likelihood ratio test. Finally, we emphasized a real lifetime data set to reveal how the proposed distribution worked in it.

Mathematics Subject Classification 2000: 62F10, 62E15, 60E05

Key Words: Length-biased distribution, Chris-Jerry distribution, Reliability analysis, Moments, Maximum likelihood estimator.

1. INTRODUCTION

The concept of weighted distribution was first proposed by Sir Ronald Aylmer Fisher (1934). The postulate of weighted distribution is an assemblage of model specification and data interpretation. Also, the notion of weighted distribution can be traced to the probe of “the estimation of frequencies based on the effective methods of ascertainment”. Then, it was expanded by Rao (1965), as the weighted distribution was applied in many fields like medicine, reliability, ecology, behavioral sciences, finance, insurance, etc. for the development of statistical models. The idea of length-biased was first suggested by Cox (1969) and Zelen (1974). The length-biased distribution results when the probability of observing a positive-valued random variable is proportional to the value of the variable. The length-biased distribution is a specific case of weighted distribution. So, the length-biased distribution it also applied in various research fields like medicine, reliability, ecology, behavioral sciences, finance, insurance, etc. The size-biased distribution was acknowledged independently by Van Deusen (1986). The unit sample is selected with probability which is proportional to any unit size measure is called the size-biased distribution.

¹ This research was supported by the Funding Agency.

Lindley distribution by Ghitany, M.E., Atieh, B. and Nadarajah, S., (2008) is formed using exponential with scale parameter θ and a gamma distribution having shape parameter 2 and scale parameter θ , $p = \frac{\theta}{\theta+1}$ and the statistical properties like moments, reliability analysis, maximum likelihood estimation of parameters and entropies are discussed. Gupta, R.D. and Kundu, D., (2009) given a weighted distribution as a new model by its probability density function's (p.d.f.) shape is very close to the shape of the p.d.f.'s of Weibull, Gamma and generalized exponential distributions. The analysis of statistical properties for various samples has been demonstrated. The two parameters weighted Shanker distribution has been introduced by Shanker, R. and Shukla, K.K., (2017). It is demonstrated with real lifetime datasets. This fit is found to be quite satisfactory when compared with other distributions. The weighted length- biased generalized uniform distribution is a new extension of generalized uniform distribution given by Rather, A.A. and Subramanian, C., (2018). The model of two combined distributions of Poisson- weighted Lindley distribution was introduced by Shanker, R. and Shukla, K.K., (2019) and had compared its better fitness with other distributions. Later, Subramanian, C. and Rather, A.A., (2020) extended the weighted Shanker distribution as the length-biased Shanker distribution and explained its structural properties and also illustrated real life time data to show its fit of distribution.

2. CHRIS-JERRY DISTRIBUTION (CJD)

The scale parameter θ and shape parameter 3 of the exponential distribution and Gamma distribution are used to form a new one-parameter lifetime distribution called Chris- Jerry distribution introduced by Chrisogonus. K. Onyekwere and Okechukwu. J. Obulezi (2022). Then, the p.d.f. and c.d.f. are given by

$$f_{CJ}(x, \theta) = \frac{\theta^2}{\theta + 2} (1 + \theta x^2) e^{-\theta x}; x > 0, \theta > 0 \quad (2.1)$$

$$F_{CJ}(x, \theta) = 1 - \left[1 + \frac{\theta x(\theta x + 2)}{\theta + 2} \right] e^{-\theta x}; x > 0, \theta > 0 \quad (2.2)$$

3. LENGTH- BIASED CHRIS-JERRY (LBCJ) DISTRIBUTION

Let X be a non- negative random variable and $w(x)$ be a non- negative weighted function. Let $f(x)$ be a p.d.f of X . Then, the p.d.f of the weighted random variable x_w is

$$f_w(x) = \frac{w(x)f(x)}{E(w(x))} ; x > 0 \quad (3.1)$$

where $w(x)$ be a non-negative weight function and

$$E(w(x)) = \int w(x)f(x)dx < \infty$$

The length-biased version of CJD by taking the weight function as

$$w(x) = x$$

Then, the p.d.f. of length-biased Chris-Jerry (LBCJ) distribution is given as

$$f_l(x) = \frac{xf(x)}{E(x)} ; x > 0 \quad (3.2)$$

where,

$$E(x) = \int_0^{\infty} x f(x; \theta) dx = \frac{(\theta + 6)}{\theta(\theta + 2)} \quad (3.3)$$

Substituting (2.1) and (3.3) in equation (3.2), the requisite p.d.f. of length-biased Chris- Jerry distribution is given as

$$f_l(x; \theta) = \frac{\theta^3 x(1 + \theta x^2)e^{-\theta x}}{(\theta + 6)} ; x > 0, \theta > 0 \quad (3.4)$$

Then, the cumulative density function (c.d.f.) of the length-biased Chris-Jerry (LBCJ) distribution is derived as

$$F_{lCJ}(x; \theta) = \frac{[\theta\gamma(2, \theta x) + \gamma(4, \theta x)]}{(\theta + 6)} ; x > 0, \theta > 0 \quad (3.5)$$

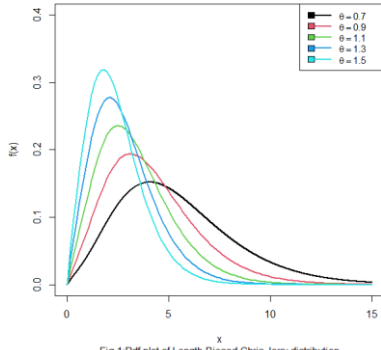


Fig.1 Pdf plot of Length-Biased Chris Jerry distribution

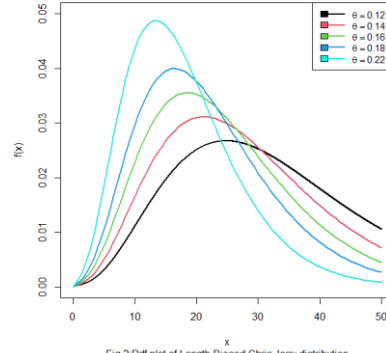


Fig.2 Pdf plot of Length-Biased Chris Jerry distribution

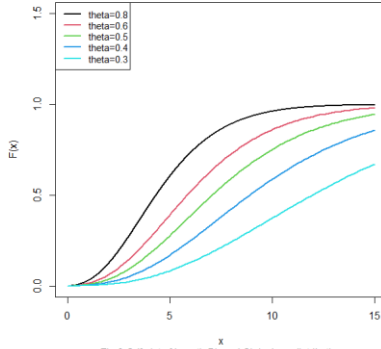


Fig.3 Cdf plot of Length-Biased Chris Jerry distribution

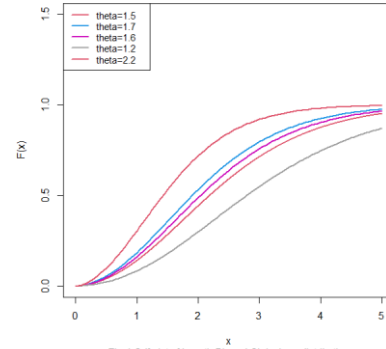


Fig.4 Cdf plot of Length-Biased Chris Jerry distribution

4. RELIABILITY ANALYSIS

In this part, the reliability or survival function, hazard function or failure rate function and reverse hazard function of the LBCJ distribution are given as follows:

$$R(x) = 1 - F(x) = 1 - \frac{[\theta\gamma(2, \theta x) + \gamma(4, \theta x)]}{(\theta + 6)}$$

$$h(x) = \frac{f_l(x; \theta)}{1 - F_l(x; \theta)} = \frac{\theta^3 x(1 + \theta x^2)e^{-\theta x}}{(\theta + 6) - [\theta\gamma(2, \theta x) + \gamma(4, \theta x)]}$$

$$h_r(x) = \frac{f(x)}{F(x)} = \frac{\theta^3 x(1 + \theta x^2)e^{-\theta x}}{[\theta\gamma(2, \theta x) + \gamma(4, \theta x)]}$$

5. STATISTICAL PROPERTIES

5.1. Moments

Let X indicate the random variable of length-biased Chris-Jerry distribution with parameters θ and c . Then, the r^{th} order moment $E(X^r)$ of length-biased Chris-Jerry distribution is derived as

$$\begin{aligned}\mu'_r = E(X^r) &= \int_0^\infty x^r f_l(x; \theta) dx = \int_0^\infty \frac{\theta^3}{(\theta + 6)} x^r x(1 + \theta x^2) e^{-\theta x} dx \\ \mu'_r &= \frac{[\theta \Gamma(r + 2) + \Gamma(r + 4)]}{\theta^r(\theta + 6)}\end{aligned}\quad (5.1)$$

In equation 5.1, when $r=1$, the mean of length-biased Chris-Jerry distribution which is given by

$$\mu'_1 = \frac{2[\theta + 12]}{\theta(\theta + 6)}$$

Similarly, when $r=2$, the moment of length-biased Chris-Jerry distribution becomes

$$\mu'_2 = \frac{6[\theta + 20]}{\theta^2(\theta + 6)}$$

Therefore,

$$Variance = \frac{6[\theta + 20]}{\theta^2(\theta + 6)} - \left[\frac{2[\theta + 12]}{\theta(\theta + 6)} \right]^2$$

The Standard deviation (σ), Coefficient of variation (C.V.) and Index of dispersion (γ) are derived as

$$\begin{aligned}\sigma &= \sqrt{\frac{6[\theta + 20]}{\theta^2(\theta + 6)} - \left[\frac{2[\theta + 12]}{\theta(\theta + 6)} \right]^2} \\ C.V. &= \frac{\sigma}{\mu'_1} = \frac{\sqrt{6(\theta + 20)(\theta + 6) - 4(\theta + 12)^2}}{2(\theta + 12)} \\ \gamma &= \frac{\sigma^2}{\mu'_1} = \frac{6(\theta + 20)(\theta + 6) - 4(\theta + 12)^2}{2(\theta + 12)(\theta + 6)}\end{aligned}$$

5.2. Harmonic Mean

The Harmonic mean of the aspired model can be derived as

$$H.M = E\left[\frac{1}{x}\right] = \int_0^{\infty} \frac{1}{x} f_l(x; \theta) dx = \int_0^{\infty} \frac{1}{x} \frac{\theta^3 x(1 + \theta x^2) e^{-\theta x}}{(\theta + 6)} dx = \frac{\theta(\theta + 2)}{(\theta + 6)}$$

5.3. Moment Generating Function and Characteristic Function

Assume X has a length-biased Chris-Jerry distribution, we will get the moment-generating function of X as

$$M_x(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} f_l(x; \theta) dx$$

Using Taylor's expansion

$$\begin{aligned} &= \int_0^{\infty} \sum_{j=0}^{\infty} \frac{t^j}{j!} x^j f_l(x; \theta) dx = \sum_{j=0}^{\infty} \frac{t^j}{j!} \mu'_j \\ &= \sum_{j=0}^{\infty} \frac{t^j}{j!} \frac{[\theta \Gamma(j+2) + \Gamma(j+4)]}{\theta^j (\theta + 6)} \end{aligned}$$

$$M_x(t) = \frac{1}{(\theta + 6)} \sum_{j=0}^{\infty} \frac{t^j}{j! \theta^j} (\theta \Gamma(j+2) + \Gamma(j+4))$$

In the same way, we will get the characteristics function of length-biased Chris-Jerry distribution can be obtained as

$$\varphi_x(t) = E[e^{itx}] = \frac{1}{(\theta + 6)} \sum_{j=0}^{\infty} \frac{t^j}{j! \theta^j} (\theta \Gamma(j+2) + \Gamma(j+4))$$

6. ORDER STATISTICS

Let $X_1, X_2, \dots, X_n$ be the random variable drawn from the continuous population. Their p.d.f. be $f_x(x)$ and cumulative density function with $F_x(x)$. Then, assume $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be the order statistics of a random sample. Thus, the probability density function of r^{th} order statistics $X_{(r)}$ is given by

$$f_{x(r)}(x; \theta) = \frac{n!}{(n-r)!(r-1)!} f(x) [F(x)]^{r-1} [1 - F(x)]^{n-r} \quad (6.1)$$

Putting the equation (3.4) and (3.5) in equation (6.1), the probability density function of r^{th} order statistics $X_{(r)}$ of length-biased Chris-Jerry distribution is given by

$$f_{x(r)}(x; \theta) = \frac{n!}{(n-r)!(r-1)!} \left\{ \frac{\theta^3}{(\theta+6)} x(1+\theta x^2)e^{-\theta x} \right\} x \left\{ \frac{[\theta\gamma(2, \theta x) + \gamma(4, \theta x)]}{(\theta+6)} \right\}^{r-1} \\ \times \left\{ 1 - \frac{[\theta\gamma(2, \theta x) + \gamma(4, \theta x)]}{(\theta+6)} \right\}^{n-r} \quad (6.2)$$

Therefore, the probability density function of higher order statistics $X_{(n)}$ can be obtained as

$$f_{x(n)}(x; \theta) = \left\{ \frac{n\theta^3 x(1+\theta x^2)e^{-\theta x}}{(\theta+6)} \right\} x \left\{ \frac{[\theta\gamma(2, \theta x) + \gamma(4, \theta x)]}{(\theta+6)} \right\}^{n-1}$$

Hence, the probability density function of 1st order statistics $X_{(1)}$ can be obtained as

$$f_{x(1)}(x; \theta) = \left\{ \frac{n\theta^3 x(1+\theta x^2)e^{-\theta x}}{(\theta+6)} \right\} x \left\{ 1 - \frac{[\theta\gamma(2, \theta x) + \gamma(4, \theta x)]}{(\theta+6)} \right\}^{n-1}$$

7. MAXIMUM LIKELIHOOD ESTIMATOR AND FISHER INFORMATION MATRIX

It is one of the best methods for estimating parameters of the distribution compared with the other estimating methods. Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be the random sample of size n drawn from the length-biased Chris-Jerry distribution, then the likelihood function of length-biased Chris-Jerry distribution is given as:

$$L(x; \theta) = \prod_{i=1}^n f_i(x; \theta) = \frac{\theta^{3n}}{[\theta+6]^n} \prod_{i=1}^n x_i(1+\theta x_i^2)e^{-\theta x_i}$$

The natural log likelihood function is

$$\log L = 3n \log \theta - n \log(\theta+6) \\ + \sum_{i=1}^n \log x_i + \sum_{i=1}^n \log(1+\theta x_i^2) - \theta \sum_{i=1}^n x_i \quad (7.1)$$

By differentiating equation (7.1) with respect to θ and c , the maximum likelihood estimates of θ and c can be attained as

$$\frac{\partial \log L}{\partial \theta} = \frac{3n}{\theta} - \frac{n}{(\theta + 6)} + \sum_{i=1}^n \frac{x_i^2}{(1 + \theta x_i^2)} - \sum_{i=1}^n x_i = 0 \quad (7.2)$$

Because of the complicated form of likelihood equation (7.2), algebraically it is very difficult to solve the system of non-linear equation. Therefore, we use R and Wolfram Mathematica for estimating the required parameters.

8. LIKELIHOOD RATIO TEST

Suppose $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be a random sample from the weighted Chris-Jerry distribution. This hypothesis is tested by

$$H_0: f(x) = f(x; \theta) \text{ against } H_1: f(x) = f_i(x; \theta)$$

To check whether the random sample of size n comes from the Chris – Jerry distribution or length-biased Chris – Jerry distribution, we are using this test statistic

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^n \frac{f_i(x; \theta)}{f(x; \theta)} = \left[\frac{\theta(\theta + 2)}{(\theta + 6)} \right]^n \prod_{i=1}^n x_i$$

Therefore, the null hypothesis is rejected if

$$\Delta = \left[\frac{\theta(\theta + 2)}{(\theta + 6)} \right]^n \prod_{i=1}^n x_i > k$$

$$\Delta^* = \prod_{i=1}^n x_i > k^* \text{ where } k^* = k \left[\frac{\theta + 6}{\theta(\theta + 2)} \right]^n > 0$$

We can conclude, for large sample size n , $2 \log$ is distributed as chi-square distribution with one degree of freedom. Also, p-value is attained from the chi-square distribution. If $p(\Delta^* > k^*)$, where $k^* = \prod_{i=1}^n x_i$ is less than the specified level of significance and $\prod_{i=1}^n x_i$ is the observed value of the statistic Δ^* , then, reject null hypothesis.

9. ENTROPIES

The entropy measures the amount of information present in a variable. It is the primary measure in information theory. The entropy of a random variable is the average level of “information” or “surprise” or “uncertainty” inherent in the variable’s possible outcomes.

9.1. Renyi entropy

It is defined as the index of diversity. Renyi means affable and even-tempered. It is mainly used in statistical physics, particularly in quantum information and equilibrium. For a given probability distribution, Renyi entropy is given by

$$\begin{aligned}
 e(\beta) &= \frac{1}{1-\beta} \log \left[\int_0^\beta f^\beta(x) dx \right]; \beta > 0, \beta \neq 1 \\
 &= \frac{1}{1-\beta} \log \left\{ \left[\frac{\theta^{3\beta}}{(\theta+6)^\beta} \right] \sum_{j=0}^{\infty} \binom{\beta}{j} (\theta x^2)^j \int_0^\beta x^\beta e^{-\beta\theta x} dx \right\} \\
 e(\beta) &= \frac{1}{1-\beta} \log \left\{ \left[\frac{\theta^{3\beta}}{(\theta+6)^\beta} \right] \sum_{j=0}^{\infty} \binom{\beta}{j} \frac{\Gamma(\beta+2j+1)}{(\theta)^{\beta+j+1} (\beta)^{\beta+2j+1}} \right\}
 \end{aligned}$$

9.2. Tsallis entropy

It is a generalization of Boltzmann-Gibbs (B-G) statistics. This entropy is non-extensive. For a continuous random variable is defined as follows

$$\begin{aligned}
 S_\lambda &= \frac{1}{\lambda-1} \left[1 - \int_0^\infty f^\lambda(x) dx \right] \\
 &= \frac{1}{\lambda-1} \left[1 - \left[\frac{\theta^3}{(\theta+6)} \right]^\lambda \sum_{j=0}^{\infty} \binom{\lambda}{j} (\theta x^2)^j \int_0^\infty x^\lambda e^{-\lambda\theta x} dx \right] \\
 S_\lambda &= \frac{1}{\lambda-1} \left[1 - \left[\frac{\theta^{3\lambda}}{(\theta+6)^\lambda} \right] \sum_{j=0}^{\infty} \binom{\lambda}{j} (\theta)^j \frac{\Gamma(\lambda+2j+1)}{(\lambda\theta)^{\lambda+2j+1}} \right]
 \end{aligned}$$

10. DATA ANALYSIS

As we mentioned earlier, the weighted distribution is widely applied in various fields. Similarly, LBCJ distribution can also be applied in various research fields like medicine, bio-medical science and engineering. Let us illustrate a dataset to show that LBCJ distribution can be better than one parameter CJD. Let us represent a data set, which represents the data of fossil CO₂ emissions per capita in India from 1981 to

2020 by Andrew, Robbie. M and Peters, Glen. P in DOI: 10.5281/zenodo.7215364.

The data set is given as follows:

0.441829	0.446234	0.472227	0.473931	0.509576	0.534304
0.558213	0.589751	0.634558	0.664019	0.692245	0.722197
0.731148	0.757746	0.789656	0.839232	0.856934	0.858549
0.913936	0.923058	0.919146	0.930677	0.947477	0.989662
1.026231	1.073585	1.140592	1.211135	1.316904	1.351343
1.414156	1.539909	1.578007	1.673244	1.71655	1.780779
1.798018	1.899518	1.898949	1.750956	1.925088	

In order to compare the distributions, we study the criteria like Akaike Information Criterion (AIC), Akaike Information Criterion Corrected (AICC), Bayesian Information Criterion (BIC) and $-2 \log L$. A distribution is considered superior if it corresponds to lower values of AIC, BIC, AICC and $-2 \log L$.

Table I. MLE's, $-2 \log L$, AIC, BIC and AICC of the length-biased Chris- Jerry distribution of data set

Table.1: MLE's, $-2 \log L$, AIC, BIC and AICC of the length-biased Chris- Jerry distribution of data set

Distribution	MLE	S.E	$-2 \log L$	AIC	BIC	AICC
Chris – Jerry	$\hat{\theta}=2.0115891$	0.2129982	80.64461	82.64461	84.35818	82.74717
Length-biased Chris – Jerry	$\hat{\theta}=5.609662$	1.050614	48.35771	52.35771	55.78485	52.67349

It is clear that the LBCJ distribution has lower values of AIC, BIC, AICC, and $-2 \log L$ compared to the CJD distribution. Therefore, we can conclude that the LBCJ distribution provides a better fit than the CJD.

11. CONCLUSION

In this study, we have presented a special case of weighted distribution of Chris- Jerry distribution known as length-biased Chris- Jerry distribution with single parameter. The resiliency of the parameter of this distribution makes it provide a better result. The structural properties are very well explained. Also, the parameter is

estimated using maximum likelihood estimator. And tested by using the likelihood ratio test. This distribution has better compatibility than Chris- Jerry distribution in real lifetime data set.

ACKNOWLEDGMENTS

We express sincere thanks to Dr. Sivabalan Arumugam, Professor, who helped in formatting this paper in TEX which improved the paper quality and presentation of mathematical expressions.

REFERENCES

1. Cox, D. R., 1969. Some sampling problems in technology, In *New Development in Survey Sampling*, Johnson, N. L. and Smith, H., Jr (eds.) New York Wiley – Interscience.
2. Chrisogonus, K. O. and Okechukwu, J. O., 2022. Chris – Jerry Distribution and its Applications, *Asian Journal of Probability and Statistics*, Vol. 20, No.1, pp. 16-30.
3. Fisher, R. A., 1934, The effects of methods of ascertainment upon the estimate of frequencies. *Annals of Eugenics*, Vol. 12, No. 1, pp. 13-25.
4. Ghitany, M. E., Atieh, B. and Nadarajah, S., 2008, Lindley Distribution and its Application, *Mathematics Computing and Simulation*, Vol. 78, No. 4, pp. 493-506.
5. Gupta, R. D. and Kundu, D., 2009, A New Class of Weighted Exponential Distribution, *A Journal of theoretical and Applied Statistics*, Vol. 43, No. 6, pp. 621-634.
6. Radhakrishna Rao, C., 1965, *On Discrete Distributions Arising Out of Methods of Ascertainment*, Sankhya - Series A.
7. Rather, A. A. and Subramanian, C., 2018, Characterization and Estimation of Length- biased Weighted Generalized Uniform Distribution, *International Journal of Scientific Research in Mathematical and Statistical sciences*, Vol. 5, No. 5, pp. 72-76.
8. Subramanian, C. and Rather, A. A., 2020, A New Extension of Shanker distribution with Real Life Data, *Journal of Xi'an University of Architecture and Technology*, Vol. 12, No. 3, pp. 2001-2009.
9. Shanker, R., & Shukla, K. K., 2019, on Poisson – Weighted Lindley Distribution and its Applications, *Journal of Scientific Research*, Vol. 11, No. 1, pp. 1-13.
10. Shanker, R., & Shukla, K. K., 2017, Weighted Shanker Distribution and its Applications to Model Life time Data, *Journal of Applied Quantitative Methods*, Vol. 12, No. 2, pp. 245-265.
11. Van Deusen, P. C., 1986, Fitting assumed distributions to horizontal point sample diameters, Vol. 32, No. 1, pp. 146-148.
12. Zelen, M., (1974), Problems in cell kinetic and the early detection of disease, in *Reliability and Biometry*, F. Proschan & R. J. Sering, eds, SIAM, Philadelphia, Vol.6, No.5, pp. 701-706.

Subramanian.C, Professor,
Department of Statistics,
Annamalai University,
Annamalainagar, Chidambaram, Tamilnadu, India- 608002

Subhashree.M, Research Scholar,
Department of Statistics,
Annamalai University,
Annamalainagar, Tamil Nadu, India-608002.

Received April 2023