A Compounded Probability Model for Decreasing Hazard and its Inferential Properties

Brijesh P. Singh, Utpal Dhar Das* & Sandeep Singh

•

Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi-221005 INDIA
utpal.statmath@gmail.com

Abstract

Early failures are generally observed due to latent defects within a product caused by faulty components, faulty assembly, transportation damage and installation damage. Also early life (infant mortality) failures tend to exhibit a decreasing failure rate over time. Such type of problems can be modelled either by a complex distribution having more than one parameter or by finite mixture of some distribution. In this article a single parameter continuous compounded distribution is proposed to model such type of problems. Some important properties of the proposed distribution such as distribution function, survival function, hazard function and cumulative hazard function, entropies, stochastic ordering are derived. The maximum likelihood estimate of the parameter is obtained which is not in closed form, thus iteration procedure is used to obtain the estimate of parameter. The moments of the proposed distribution does not exist. Some real data sets are used to see the performance of proposed distribution with comparison of some other competent distributions of decreasing hazard using Likelihood, AIC, AICc, BIC and KS statistics.

Keywords: Entropy, Hazard function, KS, MLE, Order Statistics, Quantile function.

I. Introduction

Normal, exponential, gamma and weibull distributions are the basic distributions that demonstrated in a number of theoretical results in the distributions theory. Particularly, exponential distribution is an invariable example for a number of theoretical concepts in reliability studies. It is characterized as constant hazard rate. In case of necessity for an increasing/decreasing failure rate model ordinarily the choice falls on weibull distribution. Lindley distribution is an increasing hazard rate distribution and has its own importance as a life testing distribution. The lindley distribution is one parameter distribution that is a mixture of exponential and gamma distributions and was proposed by Lindley [16]. The lindley distribution is used to explain the lifetime phenomenon such as engineering, biology, medicine, ecology and finance. Ghitany et al. [10]. Lindley distribution has generated little attention in excess of the exponential distribution because of its decreasing mean residual life function and increasing hazard rate however exponential distribution has constant mean residual life function and hazard rate.

Adamidis & Loukas [1] introduced a two-parameter exponential-geometric distribution with decreasing hazard rate and Barreto-Souza et al. [5] introduced a decreasing failure rate model, compounding exponential and poisson-lindley distribution (EPL) and the probability density function is given as

$$f_{epl}(x;\beta,\theta) = \frac{\beta\theta^2(1+\theta)^2 e^{-\beta x}}{(1+3\theta+\theta^2)} \frac{(3+\theta-e^{-\beta x})}{(1+\theta-e^{-\beta x})^3}; x > 0, \beta > 0, \theta > 0 \tag{1}$$

Another idea was proposed by Kuş [15] and Tahmasbi & Rezaei [27]. They introduced the exponential Poisson (EP) and exponential logarithmic (EL) distributions and the pdf is given by

$$f_{ep}(x;\beta,\lambda) = \frac{\lambda\beta}{1-e^{-\lambda}} e^{-\lambda-\beta x+\lambda e^{-\beta x}}; x>0, \beta>0, \lambda>0 \tag{2}$$

$$f_{el}(x;\beta,p) = \frac{1}{-\log p} \frac{\beta(1-p)e^{-\beta x}}{1-(1-p)e^{-\beta x}}; x>0, \beta>0, p\in(0,1) \tag{3}$$

Chahkandi & Ganjali [8] introduced a class of distributions, which is exponential power series distributions (EPS), where compounding procedure follows the same way that was previously given by Adamidis & Loukas [1]. Weibull [29] a Swedish mathematician describe the weibull distribution that is usefull for increasing as well as decresing hazard and the pdf is defined as

$$f_w(x;\beta,\alpha) = \alpha\beta^{\alpha}x^{\alpha-1}e^{-\beta x}; x>0, \beta>0, \alpha>0 \tag{4}$$

Natural mixing of exponential populations, giving rise to a decreasing hazard rate distribution, were first introduce by Proschan [23]. Subsequently other distributions with decreasing hazard rates of practical interest were discussed by Cozzolino [7]. The distributions with decreasing failure rate (DFR) are discussed in the works of Lomax [18], Barlow et al. [4], Barlow & Marshall [2, 3], Marshall & Proschan [19], Dahiya & Gurland [9], Saunders & Myhre [25], Nassar [21], Gleser [12], Gurland & Sethuraman [13]. Keeping these ideas in view, in this study, an attempt has been made to develop a new lifetime distribution by compounding exponential and lindley distribution and named as compounded exponential-lindley (CEL) distribution. The distributional properties, estimation of parameters, Fisher information, entropies, stochastic ordering, quantile function, order statistics and simulation study for the proposed distribution have been discussed in detail.

II. Proposed Distribution

Let $X_1, X_2, ..., X_n$ be a random sample from following exponential distribution with scale parameter $\lambda > 0$ and the probability density function (pdf) is in the form

$$f(x|\lambda) = \lambda e^{-\lambda x}; \quad x>0, \lambda>0 \tag{5}$$

The parameter $\lambda > 0$ of the above distribution takes continuous value and hazard of the distribution is constant. Now we assume the parameter λ is a random variable follows lindley distribution with pdf given as

$$\phi(\lambda;\theta) = \frac{\theta^2}{(\theta+1)}(1+\lambda)e^{-\theta\lambda}; \quad \theta>0, \lambda>0 \tag{6}$$

Now the pdf of the proposed distribution CEL is given by

$$\begin{aligned} g(x;\theta) &= \int_0^{\infty} f(x|\lambda)\phi(\lambda;\theta)d\lambda = \frac{\theta^2}{(\theta+1)}\int_0^{\infty}(\lambda+\lambda^2)e^{-\lambda(x+\theta)}d\lambda \\ &= \frac{\theta^2}{(\theta+1)}\frac{(x+\theta+2)}{(x+\theta)^3}; \quad x>0, \theta>0 \end{aligned} \tag{7}$$

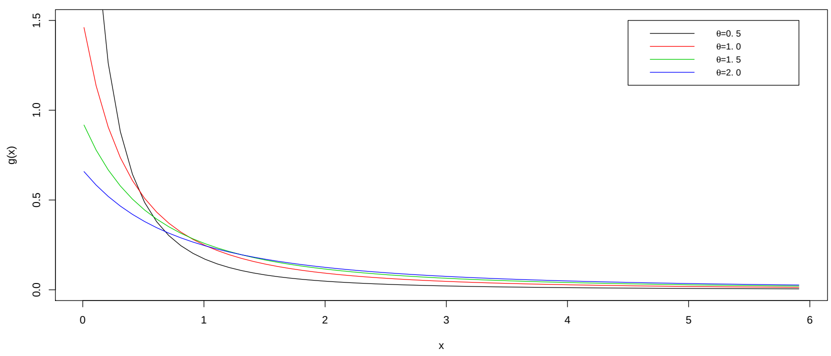

Figure 1: *Probability density function of CEL distribution*

and the cumulative distribution function (cdf) of CEL is obtained as

$$G(x;\theta) = \frac{x\left[x(\theta+1)+\theta(\theta+2)\right]}{(\theta+1)(x+\theta)^2}; x > 0, \theta > 0 \tag{8}$$

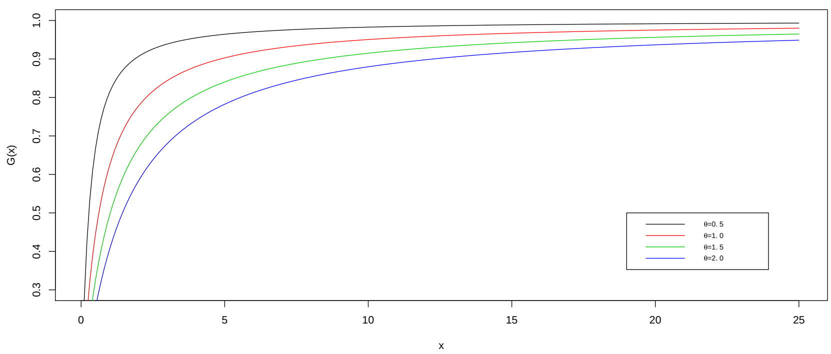

Figure 2: *Cumulative distribution function of CEL distribution*

From the figure 1 and 2, it is clear that the distribution is early failure distribution for smaller value of θ. The survival or reliability function $S(x)$ of CEL having pdf (7), is given as

$$S(x) = \frac{\theta^2(x+\theta+1)}{(\theta+1)(x+\theta)^2} \tag{9}$$

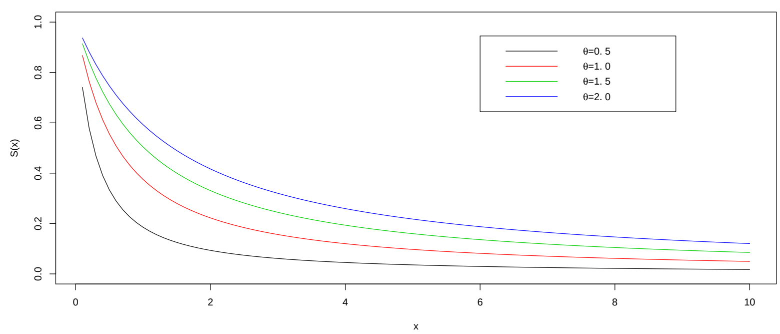

Figure 3: *Survival function of CEL distribution*

The hazard function is defined as

$$h(x) = \frac{g(x)}{1 - G(x)} = \frac{g(x)}{S(x)} = \frac{(x + \theta + 2)}{(x + \theta)(x + \theta + 1)} \tag{10}$$

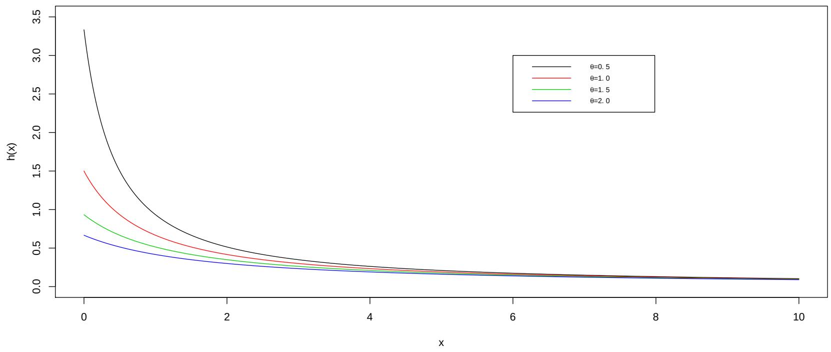

Figure 4: *Hazard rate function of CEL distribution*

According to Glaser [11], $g(t)$ is density function, $g'(t)$ is the first order derivative and $\eta(t) = -\frac{g'(t)}{g(t)}$. If $\eta'(t) > 0 \quad \forall \quad t > 0$, then the distribution has increasing failure rate (IFR) and if $\eta'(t) < 0 \quad \forall \quad t > 0$, then the distribution has decreasing failure rate (DFR). For the proposed CEL distribution

$$\eta(t) = \frac{2t + 3\theta + 3}{(t + \theta)(t + \theta + 2)} \tag{11}$$

Differentiating $\eta(t)$ with respect to t we get

$$\eta'(t) = -\frac{2}{(t + \theta)(t + \theta + 2)} - \frac{4}{(t + \theta)^2(t + \theta + 2)} - \frac{2(\theta - 1)}{(t + \theta)(t + \theta + 2)^2} - \frac{2(\theta - 1)}{[(t + \theta)(t + \theta + 2)]^2} \tag{12}$$

Now from the equation (12) we have $\eta'(t) < 0$ for all $t > 0$, hence distribution has DFR. Also the hazard function of the CEL distribution is

$$h(x) = \frac{(x + \theta + 2)}{(x + \theta)(x + \theta + 1)} = \frac{2}{(x + \theta)} - \frac{1}{(x + \theta + 1)}$$

After differentiating (10) with respect to x we get

$$h'(x) = -\frac{2}{(x + \theta)^2} + \frac{1}{(x + \theta + 1)^2}$$
$$\lim_{x \to 0} h'(x) = -\frac{2}{\theta^2} + \frac{1}{(\theta + 1)^2} < 0 \quad \forall \quad \theta > 0 \tag{13}$$

Therefore $h'(0) < 0 \quad \forall \quad \theta > 0$, Hence CEL distribution is a distribution of monotonic decreasing hazard with increasing time.
Now Cumulative hazard function $H(t)$ is defined as

$$H(t) = \int_0^t h(x)dx = \log\left[\left(\frac{\theta + 1}{t + \theta + 1}\right)\left(\frac{t + \theta}{\theta}\right)^2\right] \tag{14}$$

Theorem 1. *The moments of the $CEL(\theta)$ distribution does not exists.*

Proof: Suppose the random variable X comes from $CEL(\theta)$ then the r^{th} moment is given by

$$E(X^r) = \int_0^\infty x^r g(x)dx = \frac{\theta^2}{\theta+1}\int_0^\infty x^r \frac{x+\theta+2}{(x+\theta)^3}dx$$

Now

$$\frac{1}{\theta+1}\int_0^\infty \frac{x^r}{(1+\frac{x}{\theta})^2}dx + \frac{2}{\theta(\theta+1)}\int_0^\infty \frac{x^r}{(1+\frac{x}{\theta})^3}dx$$

Let $\frac{x}{\theta} = z; \quad dx = \theta dz; \quad x \to 0, z \to 0,$ and $\quad x \to \infty, z \to \infty$ above integral become

$$\frac{\theta^{r+1}}{\theta+1}\int_0^\infty \frac{z^r}{(1+z)^2}dz + \frac{2\theta^{r+1}}{\theta(\theta+1)}\int_0^\infty \frac{z^r}{(1+z)^3}dz$$

using Beta integral of second kind i.e $\int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}}dx = B(m,n) \quad ; m > 0; n > 0$, we get

$$\begin{aligned} E(X^r) &= \frac{\theta^{r+1}}{\theta+1}B(r+1,1-r) + \frac{2\theta^{r+1}}{\theta(\theta+1)}B(r+1,2-r) \\ &= \frac{\theta^{r+1}}{\theta+1}\left[B(r+1,1-r) + \frac{2}{\theta}B(r+1,2-r)\right] \end{aligned} \tag{15}$$

Here range is $-1 < r < 1$. But range of r should be $r \geq 1$. Hence $E(X^r)$ does not exists. Therefore mean, variance, SD as well as higher order moments does not edxists for $CEL(\theta)$.

Theorem 2. *The moment generating function of $CEL(\theta)$ does not exists.*

Proof: Let X be the random variable from $NWEL(\theta)$ distribution then the moment generating function (mgf) is given by

$$\begin{aligned} E(e^{tx}) &= \int_0^\infty e^{tx} g(x)dx = \frac{\theta^2}{\theta+1}\int_0^\infty e^{tx}\frac{x+\theta+2}{(x+\theta)^3}dx \\ &= \frac{\theta^2}{\theta+1}\left[\int_0^\infty \frac{e^{tx}}{(x+\theta)^2}dx + \int_0^\infty \frac{2e^{tx}}{(x+\theta)^3}dx\right] \end{aligned} \tag{16}$$

Now

$$\begin{aligned} \int_0^\infty \frac{e^{tx}}{(x+\theta)^2}dx &= \left[\frac{e^{tx}}{-(x+\theta)}\right]_0^\infty + t\int_0^\infty \frac{e^{tx}}{(x+\theta)}dx \\ &= \frac{1}{\theta} + \lim_{\epsilon\to\infty}\left[t\int_0^\epsilon \frac{e^{tx}}{(x+\theta)}dx\right] \end{aligned} \tag{17}$$

Now applying L'Hospital rules we get

$$\lim_{x\to\infty}\frac{e^{tx}}{(x+\theta)} = \lim_{x\to\infty}\frac{te^{tx}}{1} = \infty$$

Hence integrand is divergent, as well as the function is not integrable over R we conclude thta $E(e^{tx})$ does not exists. The characteristic function of CEL distribution is defined as

$$\Phi_x(t) = \int_0^\infty e^{itx}g(x)dx = \frac{1}{\theta+1}\sum_{k=0}^\infty (-1)^k \frac{(k+1)!}{(it)^{k+1}}\left[1 + \frac{2}{\theta}(k+2)\right] \tag{18}$$

III. Entropies

An entropy is a measure of randomness occured in any system. Entropy is an important property of probability distributions and it measures the uncertainty in a probability distribution.

I. Rényi Entropye

An entropy is a measure of variation of the uncertainty, Rényi [24] gave an expression of the Entropy function defined by

$$e(\eta) = \frac{1}{1-\eta} \log \left[\int_0^\infty g^\eta(x) dx \right]$$

where $0 < \eta < 1$, Substituting the value of g(x) from (7)

$$\begin{aligned} e(\eta) &= \frac{1}{1-\eta} \log \left[\int_0^\infty \left(\frac{\theta^2}{(\theta+1)} \frac{(x+\theta+2)}{(x+\theta)^3} \right)^\eta dx \right] \\ &= \frac{1}{1-\eta} \log \left[\left(\frac{\theta^2}{\theta+1} \right)^\eta \int_0^\infty \left\{ \frac{1}{(x+\theta)^2} + \frac{2}{(x+\theta)^3} \right\} dx \right] \end{aligned}$$

Now applying Binomial expansion $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$ we get

$$\frac{1}{1-\eta} \log \left[\left(\frac{\theta^2}{\theta+1} \right)^\eta \int_0^\infty \sum_{k=0}^{\eta} \binom{\eta}{k} \left(\frac{1}{x+\theta} \right)^{2k} \left(\frac{2}{(x+\theta)^3} \right)^{\eta-k} dx \right]$$

after simlification we get the Renyi entropy as

$$e(\eta) = \frac{\eta}{1-\eta} \log \left(\frac{\theta^2}{\theta+1} \right) + \frac{1}{1-\eta} \log \left[\sum_{k=0}^{\eta} \binom{\eta}{k} \frac{2^{\eta-k}}{(3\eta-k-1)\theta^{(3\eta-k-1)}} \right] \tag{19}$$

where $0 < \eta < 1, \quad \theta > 0, \quad x > 0$

II. Tsallis Entropy

This is introduced by Tsallis [28] as a basis for generalizing the standard statistical mechanics

$$\begin{aligned} S_\lambda &= \frac{1}{1-\lambda} \left[1 - \int_0^\infty g^\lambda(x) dx \right] \\ &= \frac{1}{1-\lambda} \left[1 - \left(\frac{\theta^2}{(\theta+1)} \right)^\lambda \int_0^\infty \left(\frac{(x+\theta+2)}{(x+\theta)^3} \right)^\lambda dx \right] \end{aligned}$$

Now applying Binomial expansion $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$ and simplifying we get Tsallis Entropy as in (20).

$$e(\eta) = \frac{1}{1-\lambda} \left[1 - \left(\frac{\theta^2}{\theta+1} \right)^\lambda \sum_{k=0}^{\lambda} \binom{\lambda}{k} \frac{2^{\lambda-k}}{(3\lambda-k-1)\theta^{(3\lambda-k-1)}} \right] \tag{20}$$

IV. Quantile Function

The quantile function for CEL distribution is defined in the form $x_q = Q(u) = G^{-1}(u)$ where $Q(u)$ is the quantile function of $G(x)$ in the range $0 < u < 1$. Taking $G(x)$ is the cdf of CEL distribution and inverting it as above will give us the quantile function as follows

$$G(x) = \frac{x\left[x(\theta+1)+\theta(\theta+2)\right]}{(\theta+1)(x+\theta)^2} = u \tag{21}$$

Simplifying equation (21) above gives the following:

$$\left(\frac{x}{x+\theta}\right)^2 + \frac{x\theta(\theta+2)}{(x+\theta)^2} = u$$

Now let $\frac{x}{x+\theta} = z$ we get from above

$$z^2 + \left(\frac{\theta+2}{\theta+1}\right) z(1-z) = u$$
$$z^2 - z(\theta+2) + u(\theta+1) = 0 \tag{22}$$

This is a quadratic equation and after solving we get the solution for x as

$$z = \frac{x}{x+\theta} = \frac{(\theta+2) \pm \sqrt{(\theta+2)^2 - 4u(\theta+1)}}{2}$$
$$Q(u) = \theta\left[\frac{2}{-\theta \pm \sqrt{(\theta+2)^2 - 4u(\theta+1)}} - 1\right] \tag{23}$$

where u is a uniform variate on the unit interval (0,1).
The median of X from the CEL distribution is simply obtained by setting $u = 0.5$ and this substitution of $u = 0.5$ in the above equation (23) gives.

$$Median = \theta\left[\frac{2}{-\theta + \sqrt{(\theta+1)^2+1}} - 1\right] \tag{24}$$

Bowley's measure of skewness based on quartiles is defined as:

$$SK = \frac{Q(\frac{3}{4}) - 2Q(\frac{1}{2}) + Q(\frac{1}{4})}{Q(\frac{3}{4}) - Q(\frac{1}{4})} \tag{25}$$

and [20] presented the Moors' kurtosis based on octiles by

$$KT = \frac{Q(\frac{7}{8}) - Q(\frac{5}{8}) - Q(\frac{3}{8}) + Q(\frac{1}{8})}{Q(\frac{6}{8}) - Q(\frac{1}{8})} \tag{26}$$

where Q(.) is calculated by using the quantile function from equation (23).

V. Stochastic Orderings

Stochastic ordering of a continuous random variable is an important tool to judging their comparative behaviour. A random variable X is said to be smaller than a random variable Y.
(i) Stochastic order $X \leq_{st} Y$ if $F_X(x) \geq F_Y(x)$ for all x.
(ii) Hazard rate order $X \leq_{hr} Y$ if $h_X(x) \geq h_Y(x)$ for all x.
(iii) Mean residual life order $X \leq_{mrl} Y$ if $m_X(x) \geq m_Y(x)$ for all x.
(iv) Likelihood ratio order $X \leq_{lr} Y$ if $\frac{f_X(x)}{f_Y(x)}$ decreases in x.

The following results by Shaked & Shanthikumar [26] are well known for introducing stochastic ordering of distributions

$$X \leq_{lr} Y \implies X \leq_{hr} Y \implies X \leq_{mrl} Y$$
$$i.e \qquad X \leq_{st} Y$$

with the help of following theorem we claim that CEL distribution is ordered with respect to strongest likelihood ratio ordering

Theorem 3. *Let $X \sim CEL(\theta_1)$ distribution and $Y \sim CEL(\theta_2)$ distribution. If $\theta_1 > \theta_2$ then $X \leq_{lr} Y$ and therefore $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.*

Proof: We have

$$\frac{f_X(x)}{f_Y(x)} = \frac{\theta_1^2(\theta_2+1)}{\theta_2^2(\theta_1+1)}\left(\frac{x+\theta_1+2}{x+\theta_2+2}\right)\left(\frac{x+\theta_2}{x+\theta_1}\right)^3; \qquad x > 0$$

Now taking log both side we get

$$\log\left[\frac{f_X(x)}{f_Y(x)}\right] = \log\left[\frac{\theta_1^2(\theta_2+1)}{\theta_2^2(\theta_1+1)}\right] + \log\left(\frac{x+\theta_1+2}{x+\theta_2+2}\right) + 3\log\left(\frac{x+\theta_2}{x+\theta_1}\right)$$

By differentiating both side we get

$$\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)} = \frac{\theta_2-\theta_1}{(2+\theta_1+x)(2+\theta_2+x)} + \frac{3(\theta_2-\theta_1)}{(x+\theta_1)(x+\theta_2)}$$

Thus for $\theta_1 > \theta_2$, $\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)} < 0$.This means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

VI. Distribution of order statistics

Let $X_1, X_2, ..., X_m$ be a random sample of size m from CEL distribution and let $X_{1;m} \leq X_{2;m} \leq ... \leq X_{m;m}$ represent the corresponding order statistics. The pdf of $X_{m;m}$ i.e r^{th} order statistics is given by

$$g_{(r:m)}(x) = \frac{m!}{(r-1)!(m-r)!}G^{r-1}(x)\left[1-G(x)\right]^{m-r}g(x)$$
$$= Z\sum_{l=0}^{m-r}\binom{m-r}{l}(-1)^l G^{r+l-1}(x)g(x) \tag{27}$$

where $Z = \frac{m!}{(r-1)!(m-r)!}$ and $g(x)$ and $G(x)$ are pdf and cdf of CEL distribution defined in (7) and (8) respectively.
Substituting for $G(x)$ and $g(x)$ in (27) and applying the general binomial expansion, we have

$$g_{(r:m)}(x) = Z\sum_{l=0}^{m-r}\binom{m-r}{l}(-1)^l\left[\frac{x\left[x(\theta+1)+\theta(\theta+2)\right]}{(\theta+1)(x+\theta)^2}\right]^{r+l-1}\frac{\theta^2}{(\theta+1)}\frac{(x+\theta+2)}{(x+\theta)^3}$$
$$= Z\sum_{l=0}^{m-r}\sum_{k=0}^{(r+l-1)}\binom{m-r}{l}\binom{r+l-1}{k}C_{l;k}\frac{x^{2r+2l-k-2}(x+\theta+2)}{(x+\theta)^{2r+2l+1}} \tag{28}$$

where $C_{l;k} = (-1)^l\left(\frac{\theta^2}{(\theta+1)}\right)^{k+1}\left(\frac{\theta+2}{\theta}\right)^k$.
Hence, the pdf of the minimum order statistic $X_{(1)}$ and maximum order statistic $X_{(n)}$ of the CEL distribution are respectively given by, respectively given by

$$g_{(1:m)}(x) = Z\sum_{l=0}^{m-1}\sum_{k=0}^{l}\binom{m-1}{l}\binom{l}{k}C_{l;k}\frac{x^{2l-k}(x+\theta+2)}{(x+\theta)^{2l+3}} \tag{29}$$

VII. Estimation of the Parameter of CEL Distribution

Suppose $X = (X_1, X_2, X_3, ..., X_n)$ be an independently and identically distributed (iid) random variables of size n with pdf (7) from CEL(θ). Then, the likelihood function based on observed sample $X = (x_1, x_2, x_3, ..., x_n)$ is defined as

$$L(\theta; x) = \left(\frac{\theta^2}{\theta+1}\right)^n \prod_{i=0}^{n} \frac{x_i+\theta+2}{(x_i+\theta)^3} \tag{30}$$

The log-likelihood function corresponding to (30) is given by

$$\log L = 2n\log\theta - n\log(\theta+1) + \sum_{i=0}^{n}\{\log(x_i+\theta+2) - 3\log(x_i+\theta)\} \tag{31}$$

Hence, the log-likelihood equation for estimating θ is

$$\frac{2n}{\theta} - \frac{n}{(\theta+1)} + \sum_{i=0}^{n}\left\{\frac{1}{(x_i+\theta+2)} - \frac{3}{(x_i+\theta)}\right\} = 0 \tag{32}$$

Above equation is not solvable analytically for θ. Thus numerical iteration technique is used to get its numerical solution. Fisher Information matrix can be estimated by

$$\begin{aligned} I(\hat{\theta}) &= \left[\frac{-\partial^2}{\partial\theta^2}\log L\right]_{\theta=\hat{\theta}} \\ \frac{\partial^2}{\partial\theta^2}\log L &= -\frac{2n}{\theta^2} + \frac{n}{(\theta+1)^2} + \sum_{i=0}^{n}\left\{\frac{3}{(x_i+\theta)^2} - \frac{1}{(x_i+\theta+2)^2}\right\} \end{aligned} \tag{33}$$

For large samples, we can obtain the confidence intervals based on Fisher information matrix $I^{-1}(\hat{\theta})$ which provides the estimated asymptotic variance for the parameter θ. Thus, a two-sided $100(1-\alpha)$% confidence interval of θ and it is defined as $\hat{\theta} \pm Z_\alpha/2\sqrt{var\hat{\theta}}$. Where $Z_\alpha/2$ denotes the upper α-th percentile of the standard normal distribution.

VIII. Simulation study

In this section we evaluate the performance of the MLEs of the model parameter for the CEL distribution. We generate random variables from $CEL(\theta)$ and then obtain m.l.e. of the parameter θ, Now for $\theta = 1.5, 2, 2.5, 3$ we generate the sample size 20, 30, 50, 90, 150, 200. The program is replicated N= 2,500 times to get the maximum likelihood estimate of θ. The simulation results are reported in Table (1).

Table 1: *Simulation results for different values of θ*

	n	Bias	MSE	Var.	Est.
θ=1.5	20	0.07273	0.25756	0.26034	1.57273
	30	0.06864	0.16356	0.16447	1.56864
	50	0.03938	0.09339	0.09211	1.53938
	90	0.01632	0.04662	0.04854	1.51632
	150	0.01270	0.02856	0.02858	1.51270
	200	0.01126	0.02597	0.02152	1.51126
θ=2	20	0.11554	0.53363	0.50319	2.11554
	30	0.07354	0.29175	0.30196	2.07354
	50	0.04340	0.16161	0.16965	2.04340
	90	0.01612	0.08701	0.08982	2.01612
	150	0.01415	0.05453	0.05321	2.01415
	200	0.00896	0.03761	0.03949	2.00896
θ=2.5	20	0.15021	0.82604	0.81882	2.65021
	30	0.11545	0.50885	0.50463	2.61544
	50	0.06234	0.27557	0.27921	2.56234
	90	0.02329	0.14411	0.14678	2.52329
	150	0.02114	0.08766	0.08675	2.52114
	200	-0.00545	0.06452	0.06337	2.49455
θ=2.5	20	0.21267	1.23999	1.25054	3.21267
	30	0.16488	0.82261	0.77061	3.16488
	50	0.09941	0.40727	0.42342	3.09941
	90	0.06733	0.22374	0.22481	3.06733
	150	0.03938	0.12947	0.13061	3.03938
	200	0.03598	0.09335	0.09728	3.03598

It is clearly observed from the Table (1) that the values of bias and mean square error (MSE) of the parameter estimates decreases as the sample size n increases. It indicates the consistency of the estimator.

IX. Goodness of fit

The application of goodness of fit of proposed CEL distribution has been discussed with two real data sets. First data set presents the results of a life-test experiment in which specimens of a type of electrical insulating fluid were subject to a constant voltage stress (34 KV/minutes), this data set is reported by Nelson [22] and other data is represents 30 failure times of the air conditioning system of an airplane has been reported in a paper by Linhart & Zucchini [17] and has also been analyzed by Barreto-Souza & Bakouch [6] and so on. For comparing the suitability of the model, we have considered following criterion's; namely AIC (Akaike Information Criterion), BIC (Bayesian information criterion), AICc (Corrected Akaike information criterion) and KS statistics with associated p-value of the fitted distributions are presented in Table (2) and Table (3).The AIC, BIC, AICc and KS Statistics are computed using the following formulae

$$AIC = -2loglik + 2k, \qquad BIC = -2loglik + k\log n$$
$$AICc = AIC + \frac{2k^2 + 2k}{n-k-1}, \qquad D = \sup_x |F_n(x) - F_0(x)|$$

where k= the number of parameters, n= the sample size, and the $F_n(x)$=empirical distribution function and $F_0(x)$ is the theoretical cumulativedistribution function.

Table 2: *MLE's, - 2ln L, AIC, KS and p-values of the fitted distributions for the 1st dataset.*

Distribution	Estimate	-2LL	AIC	BIC	AICc	KS	p-value
CEL(θ)	7.0385	137.98	139.98	140.92	140.21	0.1131	0.9458
EPL(β, θ)	(0.0334, 0.5521)	136.18	140.18	142.06	140.93	0.1500	0.7312
EL(β, p)	(0.0393, 0.0982)	135.98	139.98	141.87	140.73	0.1382	0.8137
EP(β, λ)	(0.0409, 2.2112)	136.89	140.89	142.78	141.64	0.1611	0.6497
Weibull(β, θ)	(0.0818, 0.7708)	136.77	140.77	142.66	141.52	0.1613	0.6482
Gamma(β, θ)	(0.0480, 0.6897)	137.23	141.23	143.12	141.98	0.1846	0.4802

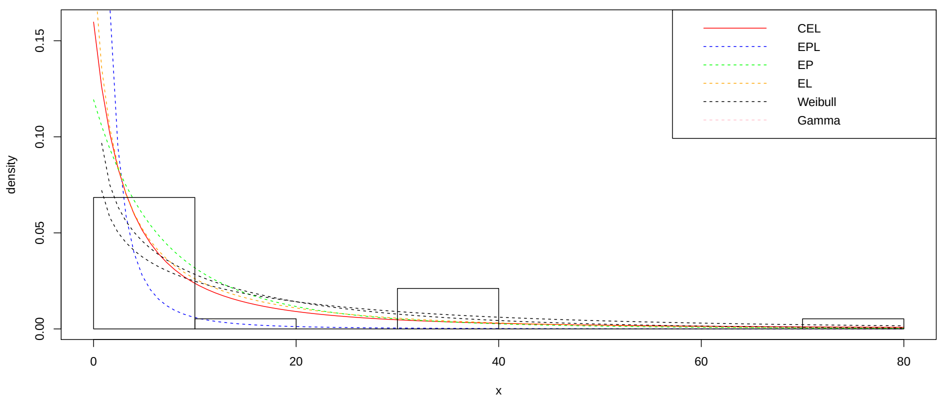

Figure 5: *Fitted pdfs of 1st data set*

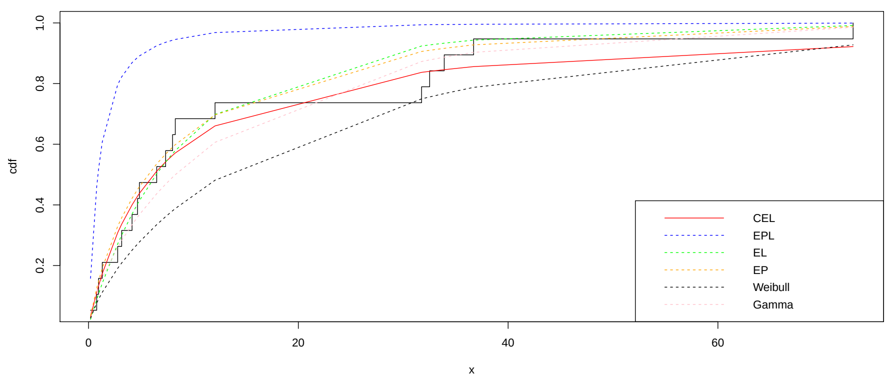

Figure 6: *Fitted cdfs and ecdf of 1st data set*

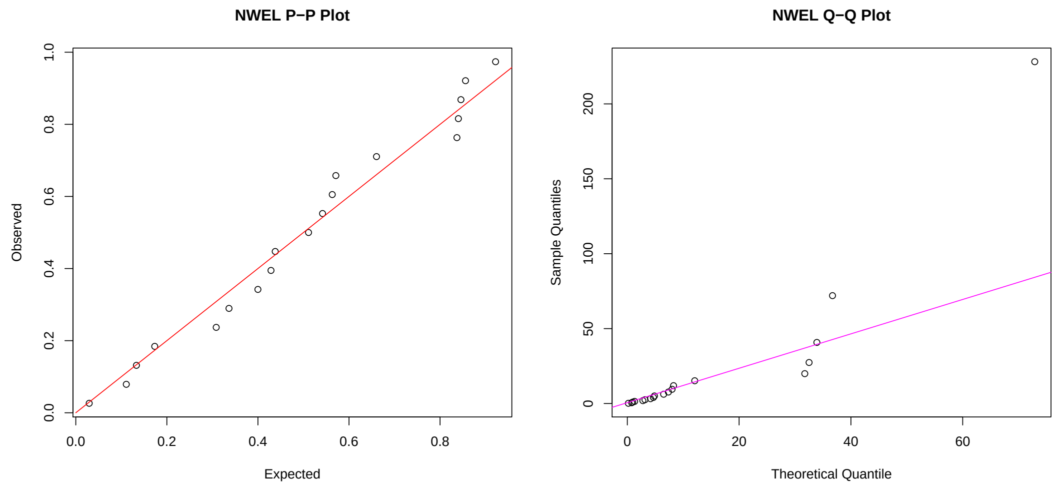

Figure 7: *p-p and q-q plot for the 1st data set.*

Here we notice that all of the considered models fit the data at 5% level of significance but the proposed distribution has minimum KS and maximum p-value among all the fitted models. Therefore, we may say that proposed CEL distribution is the most acceptable model for the present data set among the other considered models. For better visualization of the fitted models the estimated pdfs, cdfs, pp and qq plots are shown in Figure 5, Figure 6, Figure 7 for the first data set.

Table 3: *MLE's, - 2ln L, AIC, KS and p-values of the fitted distributions for the 2nd dataset.*

Distribution	Estimate	-2LL	AIC	BIC	AICc	KS	p-value
CEL(θ)	30.267	307.17	309.17	310.57	309.31	0.1061	0.8695
EPL(β,θ)	(0.0101, 0.9193)	302.87	306.87	309.68	307.32	0.1282	0.7076
EL(β,p)	(0.0111, 0.1932)	302.83	306.83	309.63	307.28	0.1291	0.6986
EP(β,θ)	(0.0105, 1.8243)	303.22	307.22	310.02	307.66	0.1468	0.5375
Weibull(β,θ)	(0.0183, 0.8536)	307.87	310.68	308.32	303.87	0.1534	0.4806
Gamma(β,θ)	(0.0136, 0.8119)	304.33	308.33	311.13	308.78	0.1694	0.3556

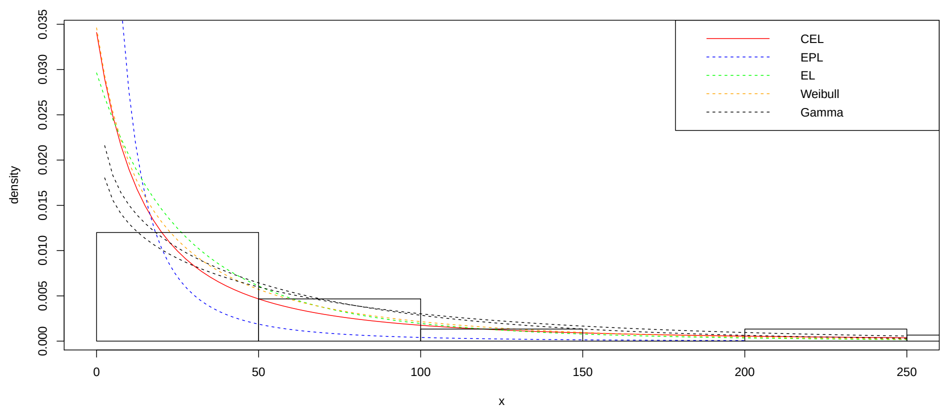

Figure 8: *Fitted pdfs of 2nd data set*

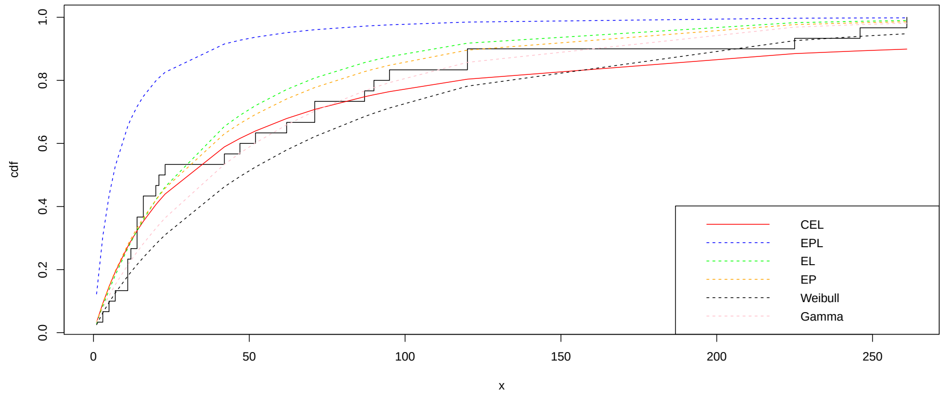

Figure 9: *Fitted cdfs of 2nd data set*

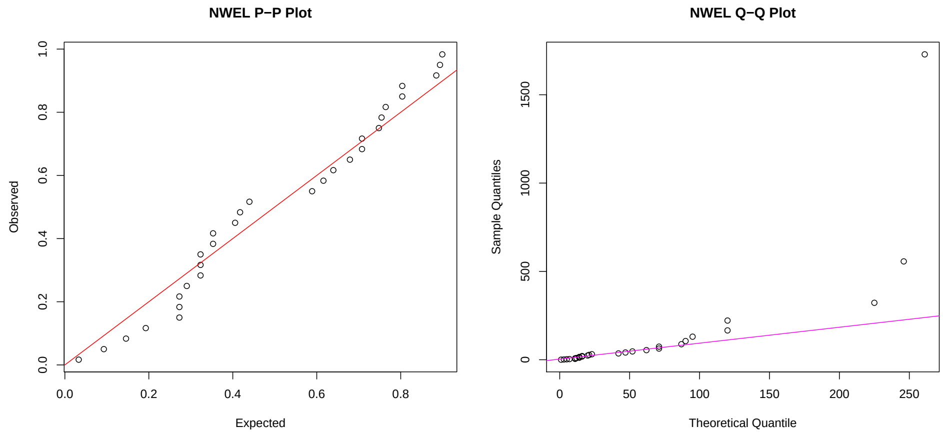

Figure 10: *p-p and q-q plot for the 2nd data set.*

For the second data set also all the considered models fit well. Here also, the value of KS statistics is the minimum for CEL distribution with the maximum p-value. From the above discussion on two real data sets we see that all the considered six decreasing failure models fit to the two data sets. The fitted models the estimated pdfs, cdfs, pp and qq plots are shown in Figure 8, Figure 9, Figure 10 for the second data set.

X. Applications on infant mortality data

Since the CEL distribution is an early failure distribution then this may be suitable for the data of infant deaths. In this study an attemt has been made to apply CEL distribution for the data of infant deaths taken from the fourth round of National Family Health Survey (NFHS-4) for the most poupulous state of India i.e. Uttar Pradesh conducted in 2015-16 (IIPS and ICF, 2017)[14]. The data on infant deaths for four categories have been extacted and CEL distribution with other compitent distributions considered here have been applied. The fitting, estimate of parameters, KS distance and its p-value are provided in table 4-7. The p-value reveals that the CEL distribution is most appropriate among all considered distributions.

Table 4: *Comperison of Goodness of fit for CEL, EPL, EP, EL, Weibull and Gamma on infant mortality data (Infant deaths of mothers aged 20-25)*

Age at Infant Death	Observed frequencies	Expected frequencies of CEL	Expected frequencies of EPL	Expected frequencies of EP	Expected frequencies of EL	Expected frequencies of Weibull	Expected frequencies of Gamma
0-1	104	90.61	83.12	80.71	76.30	72.18	143.00
1-2	17	30.67	33.66	36.74	34.29	36.55	16.75
2-3	2	15.03	17.46	18.91	19.43	21.51	4.39
3-4	10	98.83	10.35	10.69	12.09	13.18	1.28
4-5	5	5.79	6.66	6.49	7.90	8.26	0.39
5-6	7	4.07	4.53	4.16	5.32	5.26	0.12
6-7	7	3.02	3.19	2.78	3.64	3.39	0.04
7-8	2	2.33	2.31	1.92	2.52	2.21	0.01
8-9	3	1.85	1.71	1.36	1.76	1.45	0.00
9-10	4	1.50	1.28	0.98	1.24	0.95	0.00
10-11	2	1.24	0.97	0.71	0.87	0.63	0.00
11-12	3	1.05	0.74	0.53	0.62	0.42	0.00
Total	166	166.00	166.00	166.00	166.00	166.00	166.00
Estimates of parameter	$\theta = 1.4410$		$\theta = 0.6102$ $\beta = 0.2399$	$\lambda = 2.4852$ $\beta = 0.2700$	$p = 0.2378$ $\beta = 0.3399$	$\alpha = 0.8961$ $\beta = 0.5304$	$\alpha = 0.4745$ $\beta = 0.9554$
K-S Distance	0.0807		0.1257	0.1403	0.1668	0.1917	0.2478
p-value	0.2206		0.0094	0.0025	0.0002	0.0000	0.0000

Table 5: *Comperison of Goodness of fit for CEL, EPL, EP, EL, Weibull and Gamma on infant mortality data (Infant deaths of mothers aged 25-30)*

Age at Infant Death	Observed frequencies	Expected frequencies of CEL	Expected frequencies of EPL	Expected frequencies of EP	Expected frequencies of EL	Expected frequencies of Weibull	Expected frequencies of Gamma
0-1	94	86.94	84.05	84.79	76.14	75.60	104.47
1-2	17	22.52	24.63	27.91	28.19	30.36	22.28
2-3	8	9.93	11.02	11.38	14.02	14.74	6.86
3-4	3	5.51	6.01	5.43	7.72	7.55	2.24
4-5	3	3.48	3.67	2.91	4.45	3.99	0.75
5-6	0	2.40	2.41	1.69	2.64	2.15	0.26
6-7	3	1.74	1.66	1.05	1.58	1.18	0.09
7-8	2	1.33	1.18	0.68	0.96	0.66	0.03
8-9	4	1.04	0.86	0.46	0.58	0.37	0.01
9-10	1	0.84	0.64	0.31	0.36	0.21	0.00
10-11	2	0.69	0.48	0.22	0.22	0.12	0.00
11-12	0	0.57	0.37	0.16	0.13	0.07	0.00
Total	137	137.00	137.00	137.00	137.00	137.00	137.00
Estimates of parameter	$\theta = 1.0624$		$\theta = 0.3689$ $\beta = 0.2355$	$\lambda = 3.4829$ $\beta = 0.3033$	$p = 0.2501$ $\beta = 0.4879$	$\alpha = 0.8868$ $\beta = 0.7795$	$\alpha = 0.7081$ $\beta = 0.9833$
K-S Distance	0.0515		0.0726	0.0672	0.1304	0.1342	0.1150
p-value	0.8509		0.4507	0.5508	0.0171	0.0128	0.0492

Table 6: *Comperison of Goodness of fit for CEL, EPL, EP, EL, Weibull and Gamma on infant mortality data (Infant deaths in year 2003)*

Age at Infant Death	Observed frequencies	Expected frequencies of CEL	Expected frequencies of EPL	Expected frequencies of EP	Expected frequencies of EL	Expected frequencies of Weibull	Expected frequencies of Gamma
0-1	76	66.35	64.10	60.46	55.41	55.08	98.89
1-2	9	18.33	19.84	23.23	20.33	21.54	7.25
2-3	3	8.28	9.03	10.54	11.09	11.82	1.43
3-4	2	4.65	4.99	5.42	6.87	7.05	0.32
4-5	1	2.97	3.09	3.06	4.54	4.41	0.07
5-6	2	2.05	2.07	1.86	3.11	2.84	0.02
6-7	3	1.50	1.46	1.19	2.18	1.87	0.00
7-8	1	1.14	1.07	0.80	1.55	1.25	0.00
8-9	3	0.90	0.81	0.55	1.11	0.85	0.00
9-10	2	0.72	0.63	0.39	0.81	0.59	0.00
10-11	3	0.60	0.50	0.28	0.59	0.41	0.00
11-12	3	0.50	0.40	0.20	0.43	0.29	0.00
Total	108	108.00	108.00	108.00	108.00	108.00	108.00
Estimates of parameter	$\theta = 1.1394$		$\theta = 0.2528$ $\beta = 0.1512$	$\lambda = 3.1272$ $\beta = 0.2785$	$p = 0.1183$ $\beta = 0.3040$	$\alpha = 0.7880$ $\beta = 0.6433$	$\alpha = 0.7081$ $\beta = 0.9833$
K-S Distance	0.0894		0.1102	0.1439	0.1907	0.1937	0.2119
p-value	0.3392		0.1361	0.0204	0.0006	0.0005	0.0000

Table 7: *Comperison of Goodness of fit for CEL, EPL, EP, EL, Weibull and Gamma on infant mortality data (Infant death in year 2004)*

Age at Infant Death	Observed frequencies	Expected frequencies of CEL	Expected frequencies of EPL	Expected frequencies of EP	Expected frequencies of EL	Expected frequencies of Weibull	Expected frequencies of Gamma
0-1	54	46.83	42.23	42.22	38.71	36.31	75.22
1-2	15	17.27	18.87	20.22	19.41	20.56	10.87
2-3	3	8.79	11.02	10.77	11.40	12.59	3.25
3-4	2	5.28	6.01	6.28	7.21	7.89	1.08
4-5	1	3.51	3.67	4.11	4.74	5.00	0.37
5-6	2	2.49	2.41	2.82	3.19	3.20	0.13
6-7	4	1.86	1.66	2.00	2.18	2.06	0.05
7-8	2	1.44	1.18	1.45	1.51	1.33	0.02
8-9	2	1.15	0.86	1.07	1.05	0.87	0.01
9-10	2	0.94	0.64	0.08	0.73	0.57	0.00
10-11	2	0.78	0.48	0.60	0.56	0.37	0.00
11-12	2	0.66	0.37	0.46	0.36	0.24	0.00
Total	91	91.00	91.00	91.00	91.00	91.00	91.00
Estimates of parameter	$\theta = 1.6062$		$\theta = 0.7679$ $\beta = 0.2496$	$\lambda = 2.6031$ $\beta = 0.2382$	$p = 0.3311$ $\beta = 0.3477$	$\alpha = 0.9427$ $\beta = 0.4853$	$\alpha = 0.5032$ $\beta = 1.0726$
K-S Distance	0.0788		0.1293	0.1294	0.1680	0.1943	0.2332
p-value	0.6064		0.0876	0.0871	0.0102	0.0017	0.0000

XI. Conclusions

A single parameter lifetime distribution $CEL(\theta)$ has been introduced. The $CEL(\theta)$ distribution is mean free distribution and has decreasing hazard. The moment generating function, r^{th} oreder moments does not exists thus mean, variance, cumulant generating function, mean deviation about mean and median, Bonferroni, Gini index, mean residual life function (MRLF) also does not exists. The beauty of CEL distribution is that, this is a single parameter decreasing hazard distribution and explains the phenomenon better than other two parameter models. Although the moments do not exist, but Figure 1 indicates that, the distribution is highly positively skewed distribution. As the value of θ is increasing the density of the distribution becomes flatten. Hence, we can easily conclude that the proposed CEL distribution may be considered as a suitable model for the case of decreasing failure rate scenario with a hope to get better model in various disciplines such as medical, engineering, and social sciences.

References

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

[2] Barlow, R. E. & Marshall, A. W. (1964). Bounds for distributions with monotone hazard rate. *The Annals of Mathematical Statistics,* 35(3), 1234-1257.

[3] Barlow, R. E., & Marshall, A. W. (1965). Tables of bounds for distributions with monotone hazard rate. *Journal of the American Statistical Association*, 60(311), 872-890.

[4] Barlow, R. E., Marshall, A. W. & Proschan, F. (1963). Properties of probability distributions with monotone hazard rate. *The Annals of Mathematical Statistics,* 34(2), 375-389.

[5] Barreto-Souza, W., de Morais, A. L. & Cordeiro, G. M. (2011). The Weibull-geometric distribution. *Journal of Statistical Computation and Simulation,* 81(5): 645-657.

[6] Barreto-Souza, W. & Bakouch, H. S. (2013). A new lifetime model with decreasing failure rate. *Statistics*, 47(2): 465-476.

[7] Cozzolino, J. M. (1968). Probabilistic models of decreasing failure rate processes. *Naval Research Logistics Quarterly*, 15(3), 361-374.

[8] Chahkandi, M. & Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics & Data Analysis*, 53(12): 4433-4440.

[9] Dahiya, R. C. & Gurland, J. (1972). Goodness of fit tests for the gamma and exponential distributions. *Technometrics*, 14(3), 791-801.

[10] 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, 1190-1201.

[11] Glaser, R. E. (1980). Bathtub and related failure rate characterizations. *Journal of the American Statistical Association*, 75(371): 667-672.

[12] Gleser, L. J. (1989). The gamma distribution as a mixture of exponential distributions. *The American Statistician*, 43(2), 115-117.

[13] Gurland, J. & Sethuraman, J. (1994). Reversal of increasing failure rates when pooling failure data. *Technometrics*, 36(4), 416-418.

[14] International Institute for Population Sciences (IIPS) and ICF. (2017). National Family Health Survey (NFHS-4), 2015-16: India. Mumbai: IIPS.

[15] Kuş, C. (2007). A new lifetime distribution. *Computational Statistics & Data Analysis*, 51(9): 4497-4509.

[16] Lindley D. V. (1958). Fiducial distributions and Bayes theorem. *Journal of the Royal Statistical Society, Series B (Methodological)*, 102-107.

[17] Linhart, H. & Zucchini, W. (1986). Model selection. John Wiley & Sons,

[18] Lomax, K. S. (1954). Business failures: Another example of the analysis of failure data. *Journal of the American Statistical Association*, 49(268), 847-852.

[19] Marshall, A. W. & Proschan, F. (1965). Maximum likelihood estimation for distributions with monotone failure rate. *The annals of mathematical statistics*, 36(1), 69-77.

[20] Moors, J. J. A. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 37(1), 25-32.

[21] Nassar, M. M. (1988). Two properties of mixtures of exponential distributions. *IEEE transactions on reliability*, 37(4), 383-385.

[22] Nelson, W. Applied Life Data Analysis. John Wiley & Sons, New York, 1982.

[23] Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics*, 5, 375-383.

[24] Rényi, A. (1961) On measures of entropy and information. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. Contributions to the Theory of Statistics, *Berkeley, California: University of California Press*, (1): 547–561.

[25] Saunders, S. C. & Myhre, J. M. (1983). Maximum likelihood estimation for two-parameter decreasing hazard rate distributions using censored data. *Journal of the American Statistical Association*, 78(383), 664-673.

[26] Shaked, M. & Shanthikumar, J. Stochastic Orders and Their Applications. Boston: Academic Press, 1994.

[27] Tahmasbi, R. & Rezaei, S. (2008). A two-parameter lifetime distribution with decreasing failure rate. *Computational Statistics & Data Analysis*, 2008; 52(8): 3889-3901.

[28] Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of statistical physics*, 52(1-2), 479-487.

[29] Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of applied mechanics*, 18(3): 293-297.