UNDER PEER REVIEW

Original Research Article

A Finite Mixture of Convex Combination of Probability Models: Properties and Application

***Abstract*— In this paper, a new continuous probability distribution is developed by using a mixture of exponential and Rayleigh distributions for modeling lifetime data. The forms of the pdf and cdf along with statistical properties such as moments, incomplete moments, survival function, hazard function, mean residual life, stochastic ordering, order statistics, and stress strength reliability of the proposed distribution are explained. We also obtained the Bonferroni index and Lorenz curve of the proposed distribution. The parameters of the proposed distribution are estimated using the maximum likelihood technique. Finally, data analysis is performed using real-time data to illustrate the suitability of the proposed distribution.**

Keywords—Lifetime distribution, Hazard function, Mean residual life function, Order statistic, Parameter estimation.

I. INTRODUCTION

In statistics, data is expressed as a frequency distribution function that displays the range of potential values for a variable together with its frequency. Practically speaking, not all real data sets can be well-fitted by standard probability distributions. Such type of data sets creates a necessity for developing a new class of flexible probability distributions. So, Statisticians created a variety of probability distributions that are more flexible than traditional distributions in various methods. One conventional method is the mixing of probability distributions. There are several other methods available for creating a new family of probability distribution such as the transmutation method, α-power transformation, and so on; In this paper, we use the finite mixture of probability models using exponential and Rayleigh distribution.

A proven method of statistical modeling of a large variety of random events has been made possible by a finite mixture of probability distributions. The data comes from a population that has two or more different natures of subpopulations and is modeled using a finite mixture of the model. Finite mixture models have recently gained a lot of attention both theoretically and practically due to their versatility. Agriculture, astronomy, biology, genetics, medicine, psychology, economics, engineering, and marketing are just a few fields where mixture models have been successfully utilized. Finite mixture models underpin a variety of techniques in the major area of statistics, including cluster and latent class analyses, discriminant analysis, image analysis, and survival analysis, in addition to their more direct role in data analysis and inference of providing descriptive models for distributions where a single component distribution is insufficient.

Let $X = \{x_i, i = 1, 2, ..., n\}$ be a random sample of size n from an m-component finite mixture.

$$f(x_i;\theta) = \sum_{i=1}^{m} w_i g_i(x_i,\theta)$$

where, $g_i(x_i,\theta)$ = probability density or mass function

w_i are non negative quantities (1)

such that $w_1 + w_2 + ... + w_m = 1$

(i.e) $0 \le w_i \le 1$ for i=1,2,...,m

Here, a two-component mixture model is

$$f(x) = w_1 g_1(x) + w_2 g_2(x) \quad (2)$$

One of the first significant analyses that used finite mixture models was given by Karl Pearson (1894), a well-known biometrician, who fitted a proportional mixture of two normal probability density functions with different means μ_1 and μ_2 and different variances σ_1^2 and σ_2^2 in proportions π_1 and π_2 to some crab data provided by his colleague, the evolutionary biologist Weldon (1892, 1893). Many authors afterward used various distributions to fit mixture distributions. In that way, Lindley (1958) provides a distribution that is a mixture of an exponential distribution with a scale parameter of θ and a gamma distribution having a shape parameter of 2 and a scale parameter of θ with their mixing proportions, $\frac{\theta}{\theta+1}$, $\frac{1}{\theta+1}$ respectively, and the pdf and cdf of the Lindley distribution is

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

$$F(x) = 1 - \left[1 + \frac{\theta x}{\theta+1}\right]e^{-\theta x}; x>0, \theta>0 \quad (4)$$

Rama Shanker et al., (2013) used the finite mixture model to propose the Sushila distribution for modeling lifetime data, which is described in its pdf.

$$f(x;\alpha,\theta) = \frac{\theta^2}{\alpha(\theta+1)}\left(1+\frac{x}{\alpha}\right)e^{-\frac{\theta}{\alpha}x}; x>0, \theta>0, \alpha>0 \quad (5)$$

When α=1 it gives Lindley distribution. Where the mixing proportion for Sushila distribution is $w_1 = \frac{\theta}{\theta+1}$ and $w_2 = \frac{1}{\theta+1}$

Here, $g_1(x)$ and $g_2(x)$ denotes pdf of exponential (θ/α) and gamma (2, θ/α) distributions respectively. And also, Shanker

(2015 a) used the finite mixture model to propose the Akash distribution for modeling lifetime data, which is described by its pdf and cdf.

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

$$F(x)=1-\left[1+\frac{\theta x\left(\theta x+2\right)}{\theta^2+2}\right]e^{-\theta x};x>0,\theta>0 \quad (7)$$

Where the mixing proportion for Akash distribution is $w_1=\frac{\theta^2}{\theta^2+2}$ and $w_2=\frac{2}{\theta^2+2}$. Here, $g_1(x)$ and $g_2(x)$ denotes pdf of exponential (θ) and gamma (3, θ) distributions respectively.

Some other distributions which are made by mixture models are Janardan distribution- Shanker (2013), Shanker distribution-Shanker (2015), Aradhana distribution- Rama Shanker (2016), Sujatha distribution-Shanker (2016a), Garima distribution -Shanker (2016b), Amarendra distribution - Shanker (2016c), Devya distribution -Shanker (2016d), Rani distribution-Shanker (2017a), Akshaya distribution -Shanker (2017b), Rama distribution -Shanker (2017), Ishita distribution- Shanker and Shukla (2017), Prakaamy distribution- Shukla (2018), Pranav distribution-Shukla (2018), Ram Awadh distribution- Shukla (2018), Om distribution- Shanker and Shukla (2018), Odama distribution-Odama and Ijomath (2019), Shukla distribution- Kamlesh Kumar Shukla and Rama Shanker (2019), Rama Kamalesh distribution – Shanker and Shukla (2019), Darna distribution-Shraa and Al-Omari (2019), Gharaibeh distribution-Gharaibeh (2021), Alzoubi distribution- Benrabia and Alzoubi (2021). In this paper, a probability distribution is developed by using a mixture of exponential and Rayleigh distributions for modeling lifetime data.

This paper is also structured in the following way. Section 2 introduces the Exp-Rayleigh distribution. Section 3 presents the standard moments and other measurements for the Exp-Rayleigh distribution. Section 4 is concerned with reliability analysis. Section 5 derives the log-odds Rate of the proposed distribution. Section 6 takes a look into Entropy. Section 7 contains the stochastic ordering. Section 8 gives the order statistics for the Exp-Rayleigh distribution. The Bonferroni and Lorenz curves are seen in Section 9. Section 10 computed stress-strength reliability. In section 11, the parameters of the Exp-Rayleigh distribution were estimated using the maximum likelihood technique. Finally, in section 12, real-time data was employed for the suggested distribution as an application.

II. Exp-Rayleigh Distribution

The probability density function and cumulative distribution function for the Exp-Rayleigh distribution are

$$f(x)=\theta\lambda e^{-\lambda x}+(1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}} \quad (8)$$

$$F(x)=(\theta-1)e^{-\frac{x^2}{2\sigma^2}}+e^{-\lambda x}(e^{\lambda x}-\theta) \quad (9)$$

for, $x\geq 0, 0\leq\theta\leq 1, \lambda\geq 0, \sigma\geq 0$

The following figures illustrate some of the possible shapes of the pdf and cdf of a Rayleigh distribution for chosen values of parameters. The exponential and Rayleigh distributions are special cases of the Exp-Rayleigh distribution when *θ=1 and θ=0* respectively. According to Fig 1, the Exp-Rayleigh distribution can capture a variety of pdf patterns, including right-skewed, unimodal, and reversed-J-shaped, pdfs, depending on the parameter values.

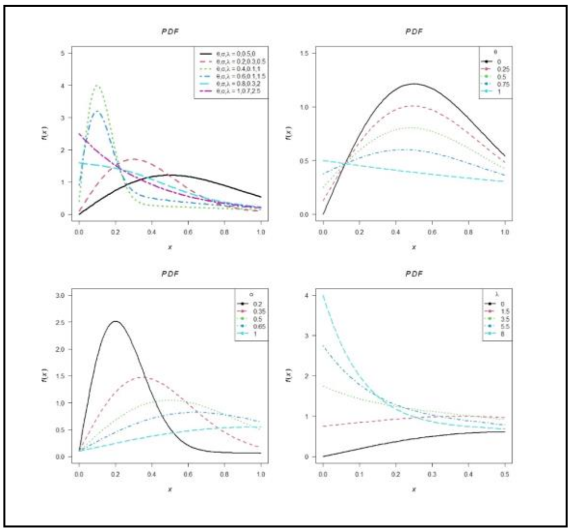

Fig 1: visual displays of pdf of an Exp- Rayleigh distribution for different parameter values.

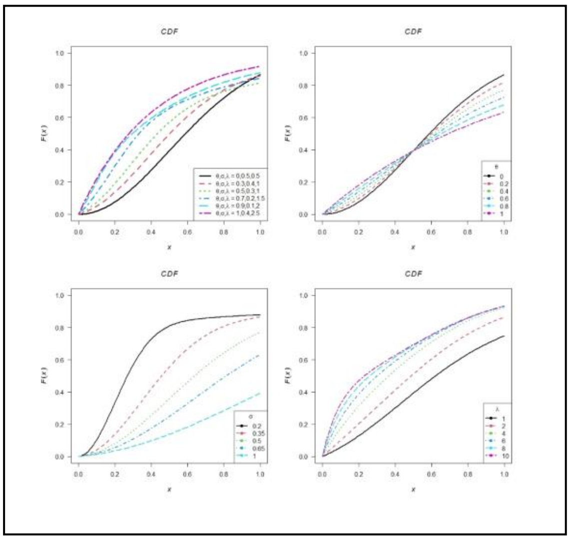

Fig 2: visual displays of cdf of an Exp- Rayleigh distribution for different parameter values.

III. Moments and Related Measures

The r^{th} Moment about the origin (raw moments) has been obtained as

$$E(X^r)=\int_0^\infty x^r f(x)dx$$

$$=\int_0^\infty x^r\left(\theta\lambda e^{-\lambda x}+(1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right)dx \quad (10)$$

$$E(X^r)=2^{r/2}\Gamma\left(\frac{r+2}{2}\right)\sigma^r-\frac{\left(\lambda^r 2^{r/2}\Gamma\left(\frac{r+2}{2}\right)\sigma^r-\Gamma\left(r+1\right)\right)\theta}{\lambda^r}$$

when $r=1,2,3,4$ then the results follow,

The Exp-Rayleigh distribution's first four moments are:

$$Mean(\mu)=E(X)=\frac{\sqrt{\pi}\left(\sqrt{2}\lambda\sigma-\sqrt{2}\lambda\sigma\theta\right)-2\theta}{2\lambda}$$

$$E(X^2)=2\sigma^2-\frac{(2\lambda^2\sigma^2-2)\theta}{\lambda^2}$$

$$E(X^3)=\frac{\sqrt{\pi}\left(3\sqrt{2}\lambda^3\sigma^3-3\sqrt{2}\lambda^3\sigma^3\theta\right)-12\theta}{2\lambda^3}$$

$$E(X^4)=8\sigma^4-\frac{(8\lambda^4\sigma^4-24)\theta}{\lambda^4}$$

Thus, the variance of the Exp-Rayleigh distribution is obtained as

$$Var=\frac{8\lambda^2\sigma^2(1-\theta)-8\theta-(2\pi\lambda^2\sigma^2(1-\theta)^2+4\theta^2-2(\sqrt{\pi}\sqrt{2}\lambda\sigma(1-\theta))}{4\lambda^2}$$

Using the above moments, the coefficient of variation and index of dispersion of the Exp-Rayleigh distribution are obtained in closed-form expressions, and also, we can obtain the coefficient of skewness and kurtosis. The index of dispersion (DI) is defined as the variance-to-mean ratio. If the DI value is less than 1, then the model is suitable for under-dispersed datasets. If the DI value is greater than 1, then the model is suitable for over-dispersed datasets.

$$C.V=\frac{\sigma}{\mu_1^1}=\frac{\left(8\lambda^2\sigma^2(1-\theta)-8\theta-(2\pi\lambda^2\sigma^2(1-\theta)^2+4\theta^2-2(\sqrt{\pi}\sqrt{2}\lambda\sigma(1-\theta))\right)^{\frac{1}{2}}}{\left(\sqrt{\pi}\left(\sqrt{2}\lambda\sigma-\sqrt{2}\lambda\sigma\theta\right)-2\theta\right)}$$

$$DI\ (\gamma)=\frac{\sigma^2}{\mu_1^1}=\frac{8\lambda^2\sigma^2(1-\theta)-8\theta-(2\pi\lambda^2\sigma^2(1-\theta)^2+4\theta^2-2(\sqrt{\pi}\sqrt{2}\lambda\sigma(1-\theta))}{2\lambda\left(\sqrt{\pi}\left(\sqrt{2}\lambda\sigma-\sqrt{2}\lambda\sigma\theta\right)-2\theta\right)}$$

The r^{th} Incomplete moment for Exp- Rayleigh distribution has been obtained as

$$\phi_r(x)=\int_0^t x^r f(x)dx$$

$$=\int_0^t x^r\left(\theta\lambda e^{-\lambda x}+(1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right)dx$$

$$\phi_r(x)=\frac{\left[\begin{array}{c}\lambda^r 2^{r/2}\left(\Gamma\left(\frac{r+2}{2},\frac{t^2}{2\sigma^2}\right)-\Gamma\left(\frac{r+2}{2},0\right)\right)|\sigma|^r\\ -\Gamma(r+1,\lambda t)+\Gamma(r+1,0)\end{array}\right]\theta}{\lambda^r}-2^{r/2}\left(\Gamma\left(\frac{r+2}{2},\frac{t^2}{2\sigma^2}\right)-\Gamma\left(\frac{r+2}{2},0\right)\right)|\sigma|^r \tag{11}$$

The first incomplete moment of the Exp- Rayleigh distribution is

$$\phi_1(x)=\frac{\theta}{\lambda}-\frac{e^{\frac{-t^2}{2\sigma^2}-\lambda t}\left[\begin{array}{c}\sqrt{\pi}\left(\sqrt{2}\lambda\sigma\theta-\sqrt{2}\lambda\sigma\right)e^{\frac{t^2}{2\sigma^2}+\lambda t}erf\left(\frac{t}{\sqrt{2}\sigma}\right)+\\ \left(2\lambda\theta t+2\theta\right)e^{\frac{t^2}{2\sigma^2}}+\left(2\lambda-2\lambda\theta\right)te^{\lambda t}\end{array}\right]}{2\lambda} \tag{12}$$

The moment-generating function of the Exp- Rayleigh distribution

$$M_X(t)=E(e^{tX})=\int_0^\infty e^{tX}f(x)dx$$

$$=\sum_{i=0}^{\infty}\frac{t^i}{i!}\left(2^{i/2}\Gamma\left(\frac{i+2}{2}\right)\sigma^i-\frac{\left(\lambda^i 2^{i/2}\Gamma\left(\frac{i+2}{2}\right)\sigma^i-\Gamma(i+1)\right)\theta}{\lambda^i}\right) \tag{13}$$

The characteristic function of the Exp- Rayleigh distribution

$$\phi_X(t)=E(e^{itX})=\int_0^\infty e^{itX}f(x)dx$$

$$=\sum_{i=0}^{\infty}\frac{it^k}{k!}\left(2^{k/2}\Gamma\left(\frac{k+2}{2}\right)\sigma^k-\frac{\left(\lambda^k 2^{k/2}\Gamma\left(\frac{k+2}{2}\right)\sigma^k-\Gamma(k+1)\right)\theta}{\lambda^k}\right) \tag{14}$$

The corresponding cumulant generating function of the Exp-Rayleigh distribution

$$K_X(t)=\log_e M_X(t)$$

$$=\prod_{i=0}^{\infty}\log_e\left(\frac{t^i}{i!}\left(2^{i/2}\Gamma\left(\frac{i+2}{2}\right)\sigma^i-\frac{\left(\lambda^i 2^{i/2}\Gamma\left(\frac{i+2}{2}\right)\sigma^i-\Gamma(i+1)\right)\theta}{\lambda^i}\right)\right) \tag{15}$$

IV. RELIABILITY ANALYSIS

A. *Survival function*

The survival function *S(x)* is the likelihood that an item will not fail before *x*.

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

$$=1-\left[(\theta-1)e^{-\frac{x^2}{2\sigma^2}}+e^{-\lambda x}(e^{\lambda x}-\theta)\right] \tag{16}$$

$$S(x)=\theta e^{-\lambda x}-(\theta-1)e^{-\frac{x^2}{2\sigma^2}}$$

B. *Hazard rate function*

The hazard function and the mean residual life function of *X* are

$$h(x)=\frac{\theta\lambda e^{-\lambda x}+(1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}}{\theta e^{-\lambda x}-(\theta-1)e^{-\frac{x^2}{2\sigma^2}}} \tag{17}$$

C. *Mean residual life function*

$$m(x)=E\left[X-x/X>x\right]=\frac{1}{1-F(x)}\int_x^\infty\left[1-F(t)\right]dt$$

$$=\frac{1}{\theta e^{-\lambda x}-(\theta-1)e^{-\frac{x^2}{2\sigma^2}}}\times\int_t^\infty\left(\theta e^{-\lambda x}-(\theta-1)e^{-\frac{x^2}{2\sigma^2}}\right)dx \tag{18}$$

$$m(x)=\frac{\sqrt{2}\lambda\sqrt{\pi}\sigma(\theta-1)\left(erf\left(\frac{t}{\sqrt{2}\sigma}\right)-1\right)+2\theta e^{-\lambda t}}{2\lambda\left(\theta e^{-\lambda x}-(\theta-1)e^{-\frac{x^2}{2\sigma^2}}\right)}$$

The Exp-Rayleigh distribution's hazard function can capture different patterns: decreasing HF, unimodal HF, constant HF, and increasing HF. Furthermore, the Mean Residual Life Function is an increasing function.

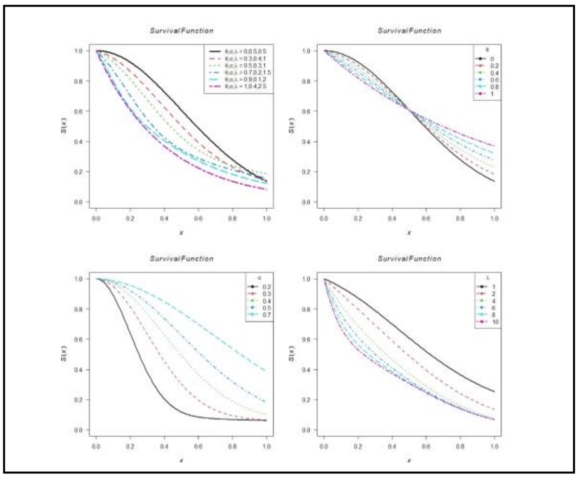

Fig 3: visual displays of the survival function of an Exp-Rayleigh distribution for different parameter values.

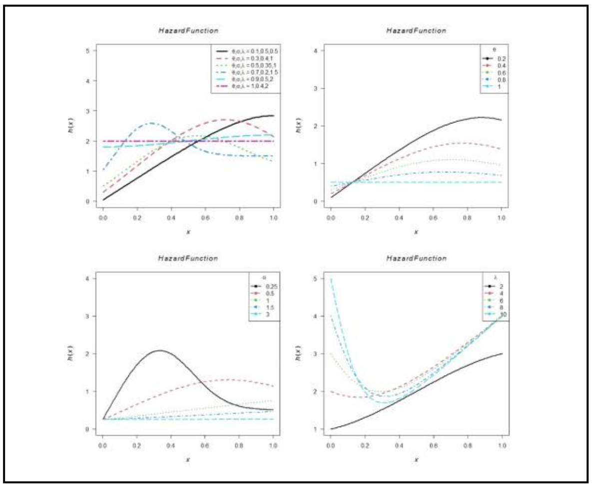

Fig 4: visual displays of the hazard function of an Exp-Rayleigh distribution for different parameter values.

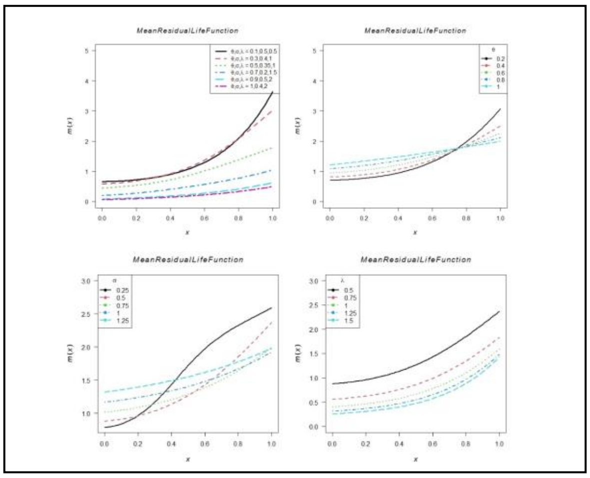

Fig 5: visual displays of mean residual life function of an Exp- Rayleigh distribution for different parameter values.

D. Mean Inactivity time

The mean inactive time is the amount of time that has passed after the failure of an item on the assumption that the failure happened in (0, t).

$$\psi_x(t) = E(X - t / X < t) = t - \frac{\phi_1(t)}{F(t)}$$

$$= t - \frac{2\theta - e^{\frac{-t^2}{2\sigma^2} - \lambda t}\left[\begin{array}{l}\sqrt{\pi}\left(\sqrt{2}\lambda\sigma\theta - \sqrt{2}\lambda\sigma\right)e^{\frac{t^2}{2\sigma^2}+\lambda t} erf\left(\frac{t}{\sqrt{2}\sigma}\right) + \\ \left(2\lambda\theta t + 2\theta\right)e^{\frac{t^2}{2\sigma^2}} + \left(2\lambda - 2\lambda\theta\right)te^{\lambda t}\end{array}\right]}{2\lambda\left((\theta-1)e^{-\frac{x^2}{2\sigma^2}} + e^{-\lambda x}(e^{\lambda x} - \theta)\right)} \tag{19}$$

E. Cumulative Hazard

The cumulative hazard function is given by

$$H(x) = -\log(1 - F(x))$$

$$= -\log\left(\theta e^{-\lambda x} - (\theta - 1)e^{-\frac{x^2}{2\sigma^2}}\right) \tag{20}$$

F. Reversed hazard rate

Reversed Hazard Rate is given by

$$\tau(x) = \frac{f(x)}{F(x)}$$

$$= \frac{\theta\lambda e^{-\lambda x} + (1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}}{(\theta-1)e^{-\frac{x^2}{2\sigma^2}} + e^{-\lambda x}(e^{\lambda x} - \theta)} \tag{21}$$

V. Log-odds Rate

Wang et al. (2003) presented a model for time to failure based on the log-odds rate, as well as some characterization of failure time distributions using the log-odds rate. The model may be used to examine the distribution of time to failure by modeling the failure process in terms of the log odds rate.

The odds function is given by

$$\pi_O(x) = \frac{F(x)}{S(x)}$$

$$= \frac{(\theta-1)e^{-\frac{x^2}{2\sigma^2}} + e^{-\lambda x}(e^{\lambda x} - \theta)}{\theta e^{-\lambda x} - (\theta-1)e^{-\frac{x^2}{2\sigma^2}}} \tag{22}$$

The log-odds function is given by

$$LO(x) = \log\frac{F(x)}{1 - F(x)}$$

$$= \log\left((\theta-1)e^{-\frac{x^2}{2\sigma^2}} + e^{-\lambda x}(e^{\lambda x} - \theta)\right) \tag{23}$$

$$-\log\left(\theta e^{-\lambda x} - (\theta-1)e^{-\frac{x^2}{2\sigma^2}}\right)$$

The log-odds rate is defined as

$$LOR(x)=\frac{h(x)}{\bar{F}(x)}$$

$$=\frac{\theta\lambda e^{-\lambda x}+(1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}}{\left(\theta e^{-\lambda x}-(\theta-1)e^{-\frac{x^2}{2\sigma^2}}\right)^2} \quad (24)$$

VI. ENTROPY

Entropy is a measure of uncertainty in a random variable X for the probability density function derived from the lifetime distribution

A. Renyi Entropy

Renyi entropy of a random variable *Exp-Rayleigh(θ,λ,σ)* with pdf is defined as

$$I_R(\eta)=\frac{1}{1-\eta}\log\int_0^\infty f^{\eta}(x)dx; \quad \eta>0,\eta\neq 1$$
$$=\frac{1}{1-\eta}\log\int_0^\infty\left(\theta\lambda e^{-\lambda x}+(1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right)^{\eta}dx \quad (25)$$

B. Shannon Entropy

The Shannon Entropy of *Exp-Rayleigh(θ,λ,σ)* is given by

$$E\left[-\log f(X)\right]=E\left[-\log\left(\theta\lambda e^{-\lambda x}+(1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right)\right]$$
$$=-E\left[\log\left(\theta\lambda e^{-\lambda x}+(1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right)\right] \quad (26)$$

C. Generalized Entropy

The Generalized Entropy of *Exp-Rayleigh(θ,λ,σ)* is given by

$$GE(w,\delta)=\frac{(2\lambda)^{\delta}}{\delta(\delta-1)\left(\sqrt{\pi}\left(\sqrt{2}\lambda\sigma-\sqrt{2}\lambda\sigma\theta\right)-2\theta\right)^{\delta}}$$
$$\times\left[\int_0^\infty x^{\delta}\left(\theta\lambda e^{-\lambda x}+(1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right)dx\right]-1 \quad (27)$$

VII. STOCHASTIC ORDERING

The use of stochastic ordering to judge the comparative behavior of positive continuous random variables is very useful. A random variable X is smaller than a random variable Y.

(i) Stochastic order $(X\leq_{st}Y)$ *if* $F_X(x)\geq F_Y(y)$ for all x

(ii) Hazard rate order $(X\leq_{hr}Y)$ *if* $h_X(x)\geq h_Y(y)$ for all x

(iii) Mean residual life order $(X\leq_{mrl}Y)$ *if* $m_X(x)\geq m_Y(y)$ for all x

(iv)Likelihood ratio order $(X\leq_{lr}Y)$ *if* $\frac{f_X(x)}{f_Y(y)}$ decreases in x

Shaked and Shanthi Kumar (1994) established the stochastic ordering of distributions with the following conclusions.

$$X\leq_{lr}Y\Rightarrow X\leq_{hr}Y\Rightarrow X\leq_{mrl}Y$$
$$\Downarrow$$
$$X\leq_{st}Y$$

The Exp-Rayleigh distribution is sorted according to the strongest 'likelihood ratio'. Let *X~ Exp-Rayleigh* $(\theta_1,\lambda_1,\sigma_1)$ and *Y~ Exp- Rayleigh* $(\theta_2,\lambda_2,\sigma_2)$. If, $\sigma_1\geq\sigma_2$, then $X\leq_{lr}Y$ hence $X\leq_{hr}Y$, $X\leq_{mlr}Y$ and $X\leq_{st}Y$. we have

$$\frac{f_X(x)}{f_Y(x)}=\frac{\theta_1\lambda_1 e^{-\lambda_1 x}+(1-\theta_1)\frac{x}{\sigma_1^2}e^{-\frac{x^2}{2\sigma_1^2}}}{\theta_2\lambda_2 e^{-\lambda_2 x}+(1-\theta_2)\frac{x}{\sigma_2^2}e^{-\frac{x^2}{2\sigma_2^2}}}$$

$$\log\frac{f_X(x)}{f_Y(x)}=\log\left[\frac{\theta_1\lambda_1 e^{-\lambda_1 x}+(1-\theta_1)\frac{x}{\sigma_1^2}e^{-\frac{x^2}{2\sigma_1^2}}}{\theta_2\lambda_2 e^{-\lambda_2 x}+(1-\theta_2)\frac{x}{\sigma_2^2}e^{-\frac{x^2}{2\sigma_2^2}}}\right]$$
$$=\log\left[\theta_1\lambda_1 e^{-\lambda_1 x}+(1-\theta_1)\frac{x}{\sigma_1^2}e^{-\frac{x^2}{2\sigma_1^2}}\right]-$$
$$\log\left[\theta_2\lambda_2 e^{-\lambda_2 x}+(1-\theta_2)\frac{x}{\sigma_2^2}e^{-\frac{x^2}{2\sigma_2^2}}\right]$$

$$\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)}=\frac{\theta_2\lambda_2^2 e^{-\lambda_2 x}-\frac{(1-\theta_2)}{\sigma_2^2}\left(e^{-\frac{x^2}{2\sigma_2^2}}-\frac{x^2e^{-\frac{x^2}{2\sigma_2^2}}}{\sigma_2^2}\right)}{\left[\theta_2\lambda_2 e^{-\lambda_2 x}+(1-\theta_2)\frac{x}{\sigma_2^2}e^{-\frac{x^2}{2\sigma_2^2}}\right]}-$$
$$\frac{\theta_1\lambda_1^2 e^{-\lambda_1 x}-\frac{(1-\theta_1)}{\sigma_1^2}\left(e^{-\frac{x^2}{2\sigma_1^2}}-\frac{x^2e^{-\frac{x^2}{2\sigma_1^2}}}{\sigma_1^2}\right)}{\left[\theta_1\lambda_1 e^{-\lambda_1 x}+(1-\theta_1)\frac{x}{\sigma_1^2}e^{-\frac{x^2}{2\sigma_1^2}}\right]} \quad (28)$$

Now if $\theta_1=\theta_2=\theta$, $\lambda_1=\lambda_2=\lambda$ $\sigma_1\geq\sigma_2$, then it implies $\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)}\leq 0$. This means that $X\leq_{lr}Y$ and hence $X\leq_{hr}Y$, $X\leq_{mlr}Y$ and $X\leq_{st}Y$.

VIII. ORDER STATISTICS

If $X_{(1)}\leq X_{(2)}\leq\ldots\leq X_{(n)}$ denotes the order statistic of a random sample $X_1,X_2,\ldots,X_n$ from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$ then the pdf $X_{(r)}$ is given by

$$f_{X_{(r)}}(x)=\frac{n!}{(r-1)!(n-r)!}f_X(x)\left[F_X(x)\right]^{(r-1)}\left[1-F_X(x)\right]^{(n-r)}$$

For, $r=1,2,\ldots,n$. The pdf of the r^{th} order statistic for the Exp-Rayleigh distribution is calculated, and the pdf of the largest order statistic $X_{(n)}$ and smallest order statistic $X_{(1)}$ are given below.

n^{th} order statistics

$$f_{X_{(n)}}(x)=nf_X(x)\left[F_X(x)\right]^{(n-1)}$$
$$=n\left(\theta\lambda e^{-\lambda x}+(1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right)\left[(\theta-1)e^{-\frac{x^2}{2\sigma^2}}+e^{-\lambda x}(e^{\lambda x}-\theta)\right]^{(n-1)} \quad (29)$$

1^{st} order statistics

$$f_{X_{(1)}}(x) = nf_X(x)\left[1 - F_X(x)\right]^{(n-1)}$$

$$= n\left(\theta\lambda e^{-\lambda x} + (1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right)\left[\theta e^{-\lambda x} - (\theta-1)e^{-\frac{x^2}{2\sigma^2}}\right]^{(n-1)} \quad (30)$$

The pdf of median order statistics

$$f_{m+1:n}(x) = \frac{(2m+1)}{m!m!} f_X(x)\left[F_X(x)\right]^m \left[1-F_X(x)\right]^m$$

$$= \frac{(2m+1)}{m!m!}\left(\theta\lambda e^{-\lambda x} + (1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right)\left[(\theta-1)e^{-\frac{x^2}{2\sigma^2}} + e^{-\lambda x}(e^{\lambda x}-\theta)\right]^m \times \left[\theta e^{-\lambda x} - (\theta-1)e^{-\frac{x^2}{2\sigma^2}}\right]^m \quad (31)$$

IX. Bonferroni and Lorenz Curve

The Bonferroni and Lorenz curves (Bonferroni, 1930) are used in a variety of sectors, including economics, demography, insurance, and medicine. The Bonferroni and Lorenz curves of Exp-Rayleigh distributions are calculated as follows:

$$B_o(x) = \frac{1}{\mu F(x)}\int_0^t xf(x)dx$$

$$= \frac{L_0(x)}{F(x)}$$

$$= \frac{\left[2\theta - e^{\frac{-t^2}{2\sigma^2}-\lambda t}\left[\begin{array}{r}\sqrt{\pi}\left(\sqrt{2}\lambda\sigma\theta - \sqrt{2}\lambda\sigma\right)e^{\frac{t^2}{2\sigma^2}+\lambda t} erf\left(\frac{t}{\sqrt{2}\sigma}\right) + \\ \left(2\lambda\theta t + 2\theta\right)e^{\frac{t^2}{2\sigma^2}} + \left(2\lambda - 2\lambda\theta\right)te^{\lambda t}\end{array}\right]\right]}{2\lambda\mu\left((\theta-1)e^{-\frac{x^2}{2\sigma^2}} + e^{-\lambda x}(e^{\lambda x}-\theta)\right)} \quad (32)$$

$$L_o(x) = \frac{1}{\mu}\int_0^t xf(x)dx$$

$$= \frac{\phi_1(x)}{E(x)}$$

$$= \frac{1}{\mu}\left[\frac{\theta}{\lambda} - \frac{e^{\frac{-t^2}{2\sigma^2}-\lambda t}\left[\begin{array}{r}\sqrt{\pi}\left(\sqrt{2}\lambda\sigma\theta - \sqrt{2}\lambda\sigma\right)e^{\frac{t^2}{2\sigma^2}+\lambda t} erf\left(\frac{t}{\sqrt{2}\sigma}\right) + \\ \left(2\lambda\theta t + 2\theta\right)e^{\frac{t^2}{2\sigma^2}} + \left(2\lambda - 2\lambda\theta\right)te^{\lambda t}\end{array}\right]}{2\lambda}\right] \quad (33)$$

X. Stress Strength Reliability

The life span of a component with an uncertain strength (X) and uncertain stress (Y) is described by stress-strength reliability. The component will work as intended until $X > Y$, at which point it will break immediately when the applied stress exceeds the component's strength. The stress strength parameter in particular is measured by $R=P(Y < X)$ in the statistical literature as a measure of component reliability. It is widely used in virtually every field of knowledge, particularly engineering, where it is used to study things like structures, the aging of concrete pressure vessels, the degeneration of rocket motors, static fatigue of ceramic components, etc.

Assume X and Y be independent stress and strength random variables, with Exp-Rayleigh distribution parameters of θ_1 and θ_2, respectively. The stress strength reliability R is then calculated as

$$R = P(Y < X) = \int_0^\infty P(Y < X | X = x) f_X(x)$$

$$= \int_0^\infty f_1(x)F_2(x)dx$$

$$= \int_0^\infty \left(\theta_1\lambda e^{-\lambda x} + (1-\theta_1)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right)\left((\theta_2-1)e^{-\frac{x^2}{2\sigma^2}} + e^{-\lambda x}(e^{\lambda x}-\theta_2)\right)dx$$

$$= -\frac{\left(\left(\begin{array}{l}2\sqrt{\pi}\lambda\sigma e^{\frac{\lambda^2\sigma^2}{2}} erf\left(\frac{\lambda\sigma}{\sqrt{2}}\right) + \\ \left(\left(2\Gamma\left(\frac{1}{2},\frac{\lambda^2\sigma^2}{2}\right) - 2\sqrt{\pi}\right)\lambda\sigma - 2^{\frac{3}{2}}\Gamma\left(1,\frac{\lambda^2\sigma^2}{2}\right)\right)e^{\frac{\lambda^2\sigma^2}{2}} + 2^{\frac{3}{2}}\end{array}\right)\theta_1 + \left(\begin{array}{l}2^{\frac{3}{2}}\Gamma\left(1,\frac{\lambda^2\sigma^2}{2}\right) - \\ 2\Gamma\left(\frac{1}{2},\frac{\lambda^2\sigma^2}{2}\right)\lambda\sigma\end{array}\right)e^{\frac{\lambda^2\sigma^2}{2}} - \sqrt{2}\right)\theta_2 + \left(-2\sqrt{\pi}\lambda\sigma e^{\frac{\lambda^2\sigma^2}{2}} erf\left(\frac{\lambda\sigma}{\sqrt{2}}\right) + 2\sqrt{\pi}\lambda\sigma e^{\frac{\lambda^2\sigma^2}{2}} - \sqrt{2}\right)\theta_1 - \sqrt{2}}{2^{\frac{3}{2}}} \quad (34)$$

XI. Estimation of Parameters

In this section, the parameters θ, λ, and β are estimated using the MLE method. Let $x_1, x_2, ..., x_n$ be a random sample from the Exp-Rayleigh distribution with pdf. Then the log-likelihood function takes the form.

$$g(x) = \theta\lambda e^{-\lambda x} + (1-\theta)\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}$$

$$L(x_i,\theta,\lambda,\sigma) = \prod_{i=1}^{n} g(x_i,\theta,\lambda,\sigma)$$

$$L(x_i,\theta,\lambda,\sigma) = \prod_{i=1}^{n}\left(\theta\lambda e^{-\lambda x_i} + (1-\theta)\frac{x_i}{\sigma^2}e^{-\frac{x_i^2}{2\sigma^2}}\right)$$

The respective sample log-likelihood function is

$$\log L(x_i,\theta,\lambda,\sigma) = \sum_{i=1}^{n}\log\left(\theta\lambda e^{-\lambda x_i} + (1-\theta)\frac{x_i}{\sigma^2}e^{-\frac{x_i^2}{2\sigma^2}}\right)$$

Now, by differentiating w.r.t. θ, λ, and, σ we can write

$$\frac{\partial \log L}{\partial\theta} = \sum_{i=1}^{n}\frac{\lambda e^{-\lambda x_i} - \frac{x_i}{\sigma^2}e^{-\frac{x_i^2}{2\sigma^2}}}{\left(\theta\lambda e^{-\lambda x_i} + (1-\theta)\frac{x_i}{\sigma^2}e^{-\frac{x_i^2}{2\sigma^2}}\right)} = 0$$

$$\frac{\partial \log L}{\partial\lambda} = \sum_{i=1}^{n}\frac{e^{-\lambda x_i} - \lambda x_i e^{-\lambda x_i}}{\left(\theta\lambda e^{-\lambda x_i} + (1-\theta)\frac{x_i}{\sigma^2}e^{-\frac{x_i^2}{2\sigma^2}}\right)} = 0$$

and

$$\frac{\partial \log L}{\partial\sigma} = \sum_{i=1}^{n}\frac{(1-\theta)x_i\left(x_i^2 - 2\sigma^2\right)e^{-\frac{x_i^2}{2\sigma^2}}}{\sigma^5\left(\theta\lambda e^{-\lambda x_i} + (1-\theta)\frac{x_i}{\sigma^2}e^{-\frac{x_i^2}{2\sigma^2}}\right)} = 0$$

This nonlinear system of equations is solved to obtain the MLEs. Nonlinear optimization procedures are frequently more convenient to use to numerically optimize the sample likelihood function. To solve these equations numerically, we can utilize statistical tools like R programming (maxLik package).

XII. APPLICATION

The data set contains the survival times (in days) of guinea pigs given various tubercle bacilli doses (72 observations). This data was analyzed by Kundu and Howlader (2010), Singh, Singh, and Sharma (2013), and Sankudey et al. (2014). Sankudey et al. (2017) gave the data after eliminating the ties. In this paper, we compared the Exp-Rayleigh, Rayleigh, Transmuted Rayleigh, Weibull, and Gamma distributions for the same data set. This section also provides a density comparison graphic.

To compare the goodness of fit we use -2lnL, AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), K-S (Kolmogorov-Smirnov) Statistic, CVM (Cramer-von-Mises), and AD (Anderson-Darling's). For the dataset, the above measures are computed and presented in Table II.

TABLE I. Estimated parameter values of the distributions

Model	Parameter Estimate	Log-Lik
Exp-Rayleigh	$\hat{\theta}$ =0.3088, $\hat{\sigma}$ =0.0547 $\hat{\lambda} = 5.9327$	91.48
Rayleigh	$\hat{\theta}$=0.0913	95.85
Transmuted Rayleigh	$\hat{\sigma} = 0.1035, \hat{\lambda} = 0.6476$	99.83
Weibull	$\hat{\lambda}$=1.4058, $\hat{k} = 0.1118$	102.83
Gamma	$\hat{\alpha} = 2.1239, \hat{\beta} = 21.0731$	105.23

TABLE II. Criteria for comparison

Distribution	AIC	BIC	K-S(p)	AD(p)	CVM(p)
Rayleigh	-178.96	-174.41	0.25 (0.00)	6.15 (0.00)	1.28 (0.001)
Transmuted Rayleigh	-187.70	-183.15	0.22 (0.00)	5.2561 (0.00)	1.02 (0.0021)
Weibull	-195.66	-191.10	0.15 (0.88)	2.36 (0.59)	0.41 (0.067)
Gamma	-201.66	-197.11	0.99 (0.00)	781.6 (0.00)	23.99 (0.00)
Exp-Rayleigh	-204.46	-197.63	0.14 (0.14)	1.429 (0.19)	0.23 (0.22)

The good of fit of the relatively better probability distribution is that one corresponds to the lowest values of -2lnL, AIC, AICC, BIC, and K-S Statistics for modeling lifetime data. From Table II, it is concluded that the Exp-Rayleigh distribution provides a better fit to the dataset better than Rayleigh, Transmuted Rayleigh, Weibull, and Gamma distributions.

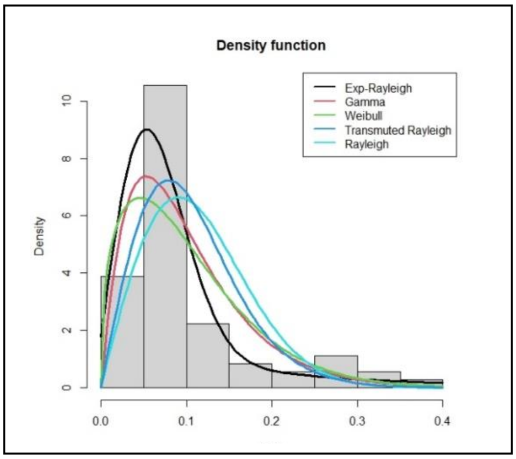

Fig 6: Comparison of fit for the five distributions of the guinea pig survival time data.

CONCLUSION

This paper develops a new weighted three-parameter probability distribution for modeling lifetime data. We derive the expressions for important statistical measures such as mean, variance, moments, and moment-generating functions, etc. Further, the estimation of Exp-Rayleigh distribution parameters is obtained using the maximum likelihood estimation procedure and the study of Exp-Rayleigh distribution characteristics using hazard and reliability functions. Finally, Real-time data is used to illustrate the suitability of the proposed distribution.

REFERENCES

[1] Abdulaziz S. Alghamdi et al., "The discrete power-Ailamujia distribution: properties, inference, and applications". AIMS Mathematics, 7(5) pp: 8344-8360,2022.

[2] M. Ali, A. Khalil, M. Ijaz, N. Saeed, "Alpha-power Exponentiated inverse Rayleigh distribution and its applications to real and simulated data". PLoS ONE. 16(1),2021.

[3] A. Alzaatreh, C. Lee, & F. Famoye, "A new method for generating families of continuous distributions". Metron. 7(11): pp: 63-79,2013.

[4] Anita Abdollahi Nanvapisheh, S.M.T.K MirMostafee, and EmrahAltun, "A new two-parameter distribution: properties and applications". Journal of Mathematical Modeling. 7(1), pp: 35-48.2019.

[5] C. Chesneau, V. Kumar, M. Khetan, M. Arshad. "On a modified weighted Exponential distribution with applications". Math. Comput. Appl. 17(4), 2022.

[6] Debasis Kundu and Ramashwar D. Gupta. "Estimation of P[Y<X] For Weibull distribution". IEEE Transactions on Reliability. 55(2),2006.

[7] M.E. Ghitany, B. Atieh, S. Nadarajah, "Lindley distribution and its application". Mathematics Computing and Simulation, 78. pp: 493-506,2008.

[8] G.M. Giorgi, S. Nadrajah, "Bonferroni and Gini indices for various parametric families of distributions". Metron. 68. pp:23-46, 2010.

[9] M. Ijaz, WK. Mashwani, SBBelhaouari. "A novel family of lifetime distribution with applications to real and simulated data". PLoS ONE. 15(10),2020.

[10] D.V. Lindley, "Fiducial distributions and Bayes' theorem". Journal of the Royal Statistical Society. Series B. 20 pp:102-107, 1958.

[11] R. Shanker, "Akash distribution and its applications". International Journal of Probability and Statistics. 4(3), pp:65-75,2015 a.

[12] R. Shanker, "Shanker distribution and its applications." International Journal of Statistics and Applications. 5(6) pp:338-348,2015 b.

[13] R. Shanker, "Aradhana distribution and its applications." International Journal of Statistics and Applications. 6(1) pp: 23-34, 2016.

[14] Wahid A. M. Shehata, Haitham M. Yousof, Mohamed Aboraya, "A novel generator of continuous probability distributions for the asymmetric left-skewed bimodal real-life data with properties and copulas". Pakistan Journal of Statistics and Operation Research. 17(4), pp:943-961, 2021.

[15] Zhongjie Shen et al., "A new generalized Rayleigh distribution with analysis to big data of an online community". Alexandria Engineering Journal. 61. pp:11523-11535,2022.

*****