

A New Weighted Rayleigh Distribution: Properties and Applications on Lifetime Time Data

Shakila Bashir*, Mujahid Rasul

Department of Statistics, Forman Christian College, Lahore, Pakistan

Email: *shakilabashir@fccollege.edu.pk

How to cite this paper: Bashir, S. and Rasul, M. (2018) A New Weighted Rayleigh Distribution: Properties and Applications on Lifetime Time Data. *Open Journal of Statistics*, 8, 640-650.

<https://doi.org/10.4236/ojs.2018.83041>

Received: April 25, 2018

Accepted: June 25, 2018

Published: June 28, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Rayleigh distribution (RD) has wide applications in many real life situations especially life testing, reliability analysis, medicines etc. In this paper a new weighted Rayleigh distribution named area-biased Rayleigh distribution (ARD) is introduced. Some mathematical properties of the (ARD) including cumulative distribution function, moments, skewness, kurtosis, median, mode, entropy, reliability measures as survival function and hazard function have been derived. Parameter of the ARD is estimated by method of moments (MOM), maximum likelihood (ML), and Bayesian. Properties of the estimators are developed. It is proved that the ML estimator attains the Cramer Rao lower bound. Applications of the ARD provided for some life time data sets. Kolmogorov Smirnov (K-S) test statistics is applied to check the good fit of the model.

Keywords

RD, ARD, MOM, ML, Bayesian

1. Introduction

In many real life fields such as medication, engineering and business, among others, modeling and examine lifetime data are crucial. Numerous lifetime distributions have been used to model lifetime data sets [1]. The quality of the procedures used in a statistical analysis depends heavily on the assumed probability model or distributions. Because of this, a number of standard probability distributions along with relevant statistical methodologies are presented in literature. But, there still remain several problems where the real data set does not follow any of the classical or standard probability models. In this article we present a

new form of the Rayleigh distribution called the area-biased Rayleigh distribution. Rayleigh [2] derived Rayleigh distribution from the ambit of noise resultant from many vital sources. The Rayleigh distribution has a variety of applications including life testing, reliability analysis, applied statistics and clinical studies. The beginning and other characteristic of this distribution can be found in Siddiqui [3], and Hirano [4]. Howlader [5] demonstrated the importance of this distribution in communication engineering. Lalitha and Mishra [6] presented modified maximum likelihood estimation for scalar parameter of Rayleigh distribution. Abd Elfattah *et al.* [7] Studied the effect of different methods of sampling schemes on the estimation of parameter for Rayleigh distribution. Further importance of Rayleigh distribution can be observed from Merovci [8] transmuted Rayleigh distribution, Das and Roy [9] length biased form of the Weighted Generalized Rayleigh distribution, Hoffman and Karst [10] properties of the Rayleigh distribution and applications of Rayleigh distribution to the analysis of the responses of marine vehicles to wave excitation.

A random variable X is said to have the Rayleigh distribution (RD) with parameter σ if its probability density function is given by

$$f(x) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2}, \quad x > 0; \sigma > 0. \quad (1)$$

while the cumulative distribution function of the Rayleigh distribution

$$F(x) = 1 - e^{-x^2/2\sigma^2} \quad (2)$$

where σ denote the scale parameter.

$$\text{Mean} = \mu'_1 = \sigma \sqrt{\frac{\pi}{2}} \quad (3)$$

$$\mu'_2 = 2\sigma^2 \quad (4)$$

One of the generalized Rayleigh distribution is given by

$$f(x) = \frac{2}{(2\sigma^2)^{N/2} \Gamma(N/2)} x^{N-1} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x > 0; \sigma > 0 \quad (5)$$

Pdf in Equation (5) is also named as chi-squared distribution with N degree of freedom and scale parameter σ .

The concept of weighted distributions was initially introduced by Fisher [11] to the study of effect of methods of ascertainment upon estimation of frequencies. On the other hand, Rao [12] presented a unified theory of weighted distributions. Rao [12] identified various real life situations that can be modeled by weighted distributions, where the observations cannot be arise from the original distributions. These situations may occur due to non-observable of some events or damage caused to the original observation ensuing in a reduced value, or arises in practice when observations from a sample are recorded with unequal probabilities.

Weighted distributions had been frequently used in research related to relia-

bility, bio-medicine, meta-analysis, econometrics, survival analysis, renewal processes, physics, ecology and branching processes can be observed in Patil and Ord [13], Patil and Rao [14], Gupta and Keating [15].

Suppose X is a non-negative random variable with its pdf $f(x; \theta)$, θ is a parameter, then $f_w(x; \theta)$ distribution is weighted version of $f(x; \theta)$, and is defined as

$$f_w(x; \theta) = \frac{w(x)f(x; \theta)}{E[w(x)]} \tag{6}$$

where $w(x)$ is an arbitrary non-negative function. For $w(x) = x^\alpha, \alpha = 1$ or 2 it is called size biased and area biased distributions respectively. The pdf of size-biased Rayleigh distribution is

$$f(x) = \frac{x^2}{\sigma^3 \sqrt{\frac{\pi}{2}}} e^{-x^2/2\sigma^2}, \quad x > 0; \sigma > 0 \tag{7}$$

2. Area Biased Rayleigh Distribution

Using Equation (1) and Equation (6), pdf of the area biased Rayleigh distribution (ARD) is

$$f(x) = \frac{x^3}{2\sigma^4} e^{-x^2/2\sigma^2}, \quad x > 0; \sigma > 0. \tag{8}$$

Rayleigh distribution in (1) and size-biased Rayleigh distribution in (7) are special cases of the generalized Rayleigh distribution in (6) for $N = 2$ and $N = 3$ respectively. Moreover, the newly derived area-biased Rayleigh distribution is also a special case of generalized Rayleigh distribution given in (6) for $N = 4$.

Cumulative distribution function (cdf) of the ARD

$$F(x) = \gamma(2, x'), \quad \text{where } x' = \frac{x^2}{2\sigma^2}, \tag{9}$$

$$\int_0^x t^{n-1} e^{-t} dt = \gamma(n, x)$$

is lower incomplete gamma function.

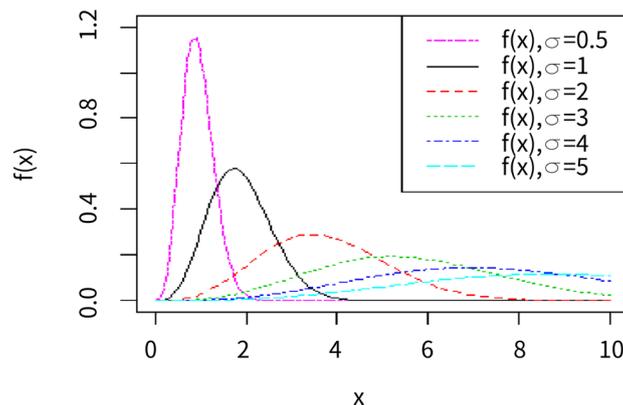


Figure 1. Pdf graph for different values of σ .

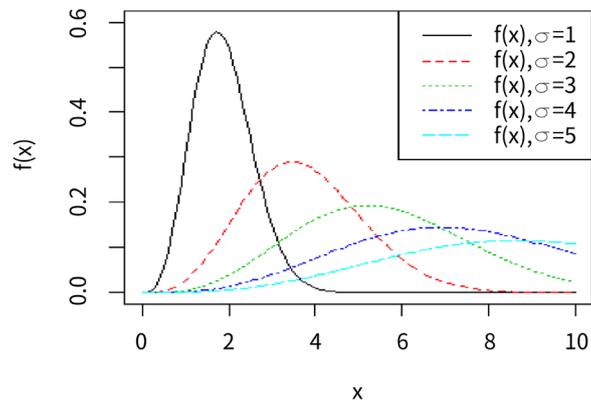


Figure 2. Pdf graph for different values of σ .

Moments and Shannon Entropy

The r th moments of the ARD are

$$\mu'_r = \frac{r(r+1)(2\sigma^2)^{r/2}}{4} \Gamma\left(\frac{r}{2}\right), \quad (10)$$

For $r = 1, 2, 3, 4$ in Equation (10), the first four moments of the ARD are

$$\text{Mean} = \mu'_1 = \frac{3\sigma\sqrt{2\pi}}{4}, \quad (11)$$

$$\mu'_2 = 4\sigma^2, \quad \mu'_3 = \frac{15\sigma^3}{2} \sqrt{\frac{\pi}{2}}, \quad \mu'_4 = 24\sigma^4, \quad (12)$$

First four mean moments of the ARD are

$$\text{Variance} = \mu_2 = \sigma^2 \left(\frac{32 - 9\pi}{8} \right) \quad (13)$$

$$\mu_3 = \frac{\sigma^3 \sqrt{\pi} (27\pi - 84)}{8\sqrt{2}}, \quad \mu_4 = \frac{3\sigma^4 (512 + 96\pi - 81\pi^2)}{64} \quad (14)$$

$$\beta_1 = \frac{4\pi(27\pi - 84)^2}{(32 - 9\pi)}, \quad \beta_2 = \frac{3(512 + 96\pi - 81\pi^2)}{(32 - 9\pi)^2} \quad (15)$$

As the expressions of β_1 & β_2 in Equation (15) are independent of σ^2 so, applying value of π , $\beta_1 \rightarrow 0.1646$ & $\beta_2 \rightarrow 3.0593$. So ARD is positively skewed and leptokurtic.

Median of the ARD is

$$\gamma(2, m') = \frac{1}{2}, \quad \text{where } m' = \frac{m^2}{2\sigma^2}, \quad (16)$$

$\int_0^x t^{n-1} e^{-t} dt = \gamma(n, x)$ is lower incomplete gamma function.

Mode of the ARD is

$$\text{mode} = \sigma\sqrt{3}, \quad (17)$$

The Shannon entropy of the ARD is

$$H(x) = \ln(\sigma) - 1.77058 \quad (18)$$

3. Estimation of Parameters

In this section parameter of ARD is estimated through method of moments (MOM) and maximum likelihood estimator (MLE).

3.1. MOM

Equating $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ and Equation (11) as $\bar{x} = \frac{3\sigma\sqrt{2\pi}}{4}$ we get MOM estimator of σ

$$\tilde{\sigma} = \frac{4\bar{X}}{3\sqrt{2\pi}}. \quad (19)$$

3.2. MLE

The likelihood function of (8)

$$L(\sigma^2) = \prod_{i=1}^n f(x_i/\sigma) = \prod_{i=1}^n \frac{x_i^3}{2\sigma^4} e^{-x_i^2/2\sigma^2}$$

Applying natural logarithm as \log_e

$$\log \prod_{i=1}^n f(x_i/\sigma) = -n \log 2 - n \log \sigma^4 - \frac{\sum_{i=1}^n x_i^2}{2\sigma^2} + \sum_{i=1}^n \log x_i^3 \quad (20)$$

Taking derivative of the Equation (20), we get

$$\frac{d}{d\sigma^2} \log \prod_{i=1}^n f(x_i/\sigma^2) = -\frac{2n}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n x_i^2, \quad (21)$$

Equating (19) to zero and simplifying we get the MLE estimator of σ

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n X_i^2}{4n} \quad (22)$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n X_i^2}{4n}} \quad (23)$$

Theorem 3.1: If X_1, X_2, \dots, X_n follows the ARD then MOM $\tilde{\sigma}$ of σ is unbiased and have minimum variance.

Proof: Applying expectation on (19) and simplifying it we get

$$\begin{aligned} E(\tilde{\sigma}) &= \frac{4}{3\sqrt{2\pi}} E(X) \\ E(\tilde{\sigma}) &= \sigma, \end{aligned} \quad (24)$$

So $\tilde{\sigma}$ is unbiased estimator of σ . Applying variance on (19), we get

$$Var(\tilde{\sigma}) = Var\left(\frac{4\bar{X}}{3\sqrt{2\pi}}\right)$$

After some simplifications we get,

$$Var(\tilde{\sigma}) = \frac{\sigma^2(32-9\pi)}{9\pi n}, \quad (25)$$

As $n \rightarrow \infty, \text{Var}(\tilde{\sigma}) \rightarrow 0$. So for large “ n ”, MOM $\tilde{\sigma}$ estimator of σ have minimum variance.

Theorem 3.2: If X_1, X_2, \dots, X_n follows the ARD then MLE $\hat{\sigma}^2$ of σ^2 is unbiased and have minimum variance.

Proof: Applying expectation on (22) and simplifying it we get

$$E(\hat{\sigma})^2 = \frac{1}{4n} \sum_{i=1}^n E(X_i^2), \text{ where } E(X_i^2) = \mu'_2$$

$$E(\hat{\sigma}^2) = \sigma^2 \quad (26)$$

So $\hat{\sigma}^2$ is unbiased estimator of σ^2 and $\hat{\sigma}$ is biased estimator of σ . Now applying variance on (22), we get

$$\text{Var}(\hat{\sigma}^2) = \text{Var}\left(\frac{\sum_{i=1}^n X_i^2}{4n}\right)$$

$$\text{Var}(\hat{\sigma}^2) = \frac{\sigma^4}{2n} \quad (27)$$

As $n \rightarrow \infty, \text{Var}(\hat{\sigma}^2) \rightarrow 0$. So for large “ n ”, MLE $\hat{\sigma}^2$ estimator of σ^2 have minimum variance.

3.3. Cramer Rao Lower Bound

Theorem 3.3: Let $X = (X_1, X_2, \dots, X_n)$ be a random sample from a pdf $f(x/\sigma^2)$ in (8), where σ shape parameter, under regularity conditions on $f(x/\sigma^2)$ for an unbiased estimator $\hat{\sigma}(x)$ of $\sigma(\sigma^2)$ i.e. $E(\hat{\sigma}^2) = \sigma^2$

$$\text{Var}(\sigma^2) \geq \frac{\frac{d}{d\sigma^2} \sigma(\sigma^2)}{I(\sigma^2)} \quad (28)$$

where

$$I(\sigma^2) = -E\left[\frac{d^2}{d(\sigma^2)^2} \log \prod_{i=1}^n f(x_i/\sigma^2)\right]$$

Proof: Taking second derivative of (21), we get

$$\frac{d^2}{d(\sigma^2)^2} \log \prod_{i=1}^n f(x_i/\sigma^2) = \frac{2n}{\sigma^4} - \frac{\sum_{i=1}^n x_i^2}{\sigma^6} \quad (29)$$

Applying expectation on (27) and simplifying it

$$I(\sigma^2) = -E\left[\frac{d^2}{d(\sigma^2)^2} \log \prod_{i=1}^n f(x_i/\sigma^2)\right] = \frac{2n}{\sigma^4} \quad (30)$$

and

$$\sigma(\sigma^2) = \sigma^2, \text{ so } \frac{d}{d\sigma^2} \sigma(\sigma^2) = 1 \quad (31)$$

Substituting (30) and (31) in (28), we get

$$\text{Var}(\sigma^2) \geq \frac{1}{2n/\sigma^4}, \text{ where } \text{Var}(\sigma^2) = \frac{\sigma^4}{2n}$$

So the unbiased estimator $\hat{\sigma}^2$ estimator of σ^2 attains the Cramer Lower Bound.

4. Bayesian Estimation

The posterior probability distribution function can be derived by using

$$f(\sigma/\underline{x}) = \frac{f(\underline{x}/\sigma)g(\sigma)}{\int f(\underline{x}/\sigma)g(\sigma)} \quad (32)$$

Using ARD $f(x/\sigma)$ in (8) and uniform prior $g(\sigma) = 1, 0 < \sigma < 1$, in (32) we get the posterior pdf of ARD as

$$f(\sigma/\underline{x}) = \frac{2}{\Gamma\left(2n - \frac{1}{2}\right) \left(\frac{2}{\sum_{i=1}^n x_i^2}\right)^{2n - \frac{1}{2}}} \exp\left(-\frac{\sum_{i=1}^n x_i^2}{2\sigma^2}\right), \sigma > 0 \quad (33)$$

Using (33) the Bayesian estimator of σ

$$\begin{aligned} \sigma^* &= E(\sigma/\underline{x}) = \int \sigma f(\sigma/x) d\sigma \\ \sigma^* &= \frac{\sqrt{\frac{\sum_{i=1}^n x_i^2}{2}}}{\Gamma\left(2n - \frac{1}{2}\right)} \Gamma(2n - 1) \end{aligned} \quad (34)$$

5. Reliability Measures

The survival function of the ARD

$$S(x) = \Gamma(2, x') \quad (35)$$

where

$$x' = \frac{x^2}{2\sigma^2}, \text{ and } \Gamma(n, x) = \int_x^\infty t^{n-1} e^{-t} dt$$

is upper incomplete gamma function.

The hazard function of the ALD is

$$h(x) = \frac{x^3 \exp(-x^2/2\sigma^2)}{2\sigma^4 \Gamma(2, x')} \quad (36)$$

6. Applications

In this section ARD is applied on two life time data sets and compared it with Lindley distribution (LD), Exponential distribution, quasi Lindley distribution (QLD), Rayleigh distribution (RD) and size-biased Rayleigh distribution (SRD) by using Kalmogorov Smirnov (K-S) Statistic. ARD is also compared for survival

function and hazard function with RD and SRD on both data 1 and only with SRD for data 2 as RD is not provided good fit for data 2.

Data set 1: This data set represents the lifetime's data relating to relief times (in minutes) of 20 patients receiving an analgesic and reported by Gross and Clark [16]: 1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3, 1.7, 2.3, 1.6, 2.

Data Set 2: This data set is the strength data of glass of the aircraft window reported by Fuller *et al.* [17]: 18.83, 20.8, 21.657, 23.03, 23.23, 24.05, 24.321, 25.5, 25.52, 25.8, 26.69, 26.77, 26.78, 27.05, 27.67, 29.9, 31.11, 33.2, 33.73, 33.76, 33.89, 34.76, 35.75, 35.91, 36.98, 37.08, 37.09, 39.58, 44.045, 45.29, 45.381.

Form **Table 1**, it can be seen that the K-S value for ARD is lower than the other discussed models so ARD is providing better alternate for the above data sets.

7. Discussion

Form **Figure 3** and **Figure 4** it can be seen that

1) The survival function graphs of the ARD are smoothly decreasing as compare to RD and SRD.

2) The hazard function graphs of the ARD are monotonically increasing in a smooth way as compare to RD and SRD. From **Figure 4** it can be seen that the hazard rate of 20 patients receiving analgesic is monotonically increasing. During an initial period, the risk is low but subsequently increases that may indicate that the patients who are receiving this painkiller drug might be suffering from severe side effects of it. We may conclude that these 20 patients have risk that gradually increases with entire range of life, which may be a result of ineffective treatment.

Form **Figure 5** and **Figure 6** it can be seen that

1) The survival function graphs of the ARD are smoothly decreasing as compare to SRD.

2) The hazard function graphs of the ARD are monotonically increasing; it means that the instantaneous failure for the strength of aircraft window is increasing.

Table 1. Comparison of KS test between distributions Lindley, Exponential, QLD, SRD and ARD.

	Model	Estimation of parameters		K-S Statistics	P-value for K-S Statistics
Data 1	Lindley	$\hat{\theta} = 0.816118$		0.34	0.294
	Exponential	$\hat{\theta} = 0.526316$		0.39	
	Quasi Lindley	$\hat{\theta} = 1.545110$	$\hat{\alpha} = -0.483393$	0.20	
	Rayleigh	$\hat{\sigma}^2 = 1.702$		0.249	
	Size-biased Rayleigh	$\hat{\sigma}^2 = 1.135$		0.165	
	Area-biased Rayleigh	$\hat{\sigma}^2 = 0.851$		0.112	

Continued

	Lindley	$\hat{\theta} = 0.062988$		0.33	
	Exponential	$\hat{\theta} = 0.032455$		0.43	
Data 2	Quasi Lindley	$\hat{\theta} = 0.103985$	$\hat{\alpha} = -0.546267$	0.30	0.240
	Rayleigh	$\hat{\sigma}^2 = 500.129$		0.287	
	Size-biased Rayleigh	$\hat{\sigma}^2 = 333.419$		0.209	
	Area-biased Rayleigh	$\hat{\sigma}^2 = 250.065$		0.157	

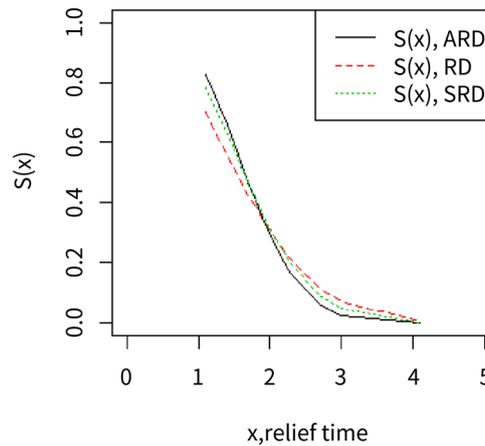


Figure 3. Survival function graph for relief time (in minutes) of 20 patients.

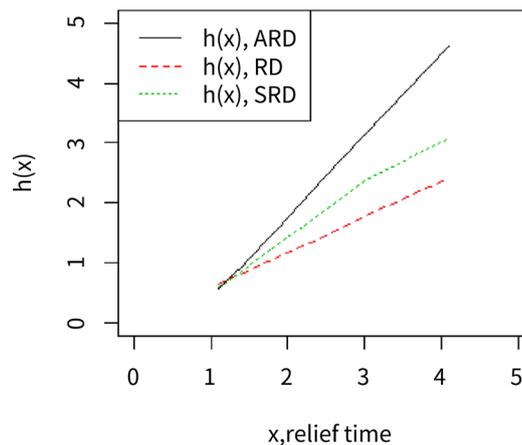


Figure 4. Hazard function graph for relief time (in minutes) of 20 patients.

8. Conclusion

In this article a new weighted single parameter Rayleigh distribution named as area-biased Rayleigh distribution (ARD) is introduced. Various properties of the ARD have been derived. It can be seen from **Figure 1**, **Figure 2** and coefficient of skewness and kurtosis that the ARD is positively skewed. Parameter is estimated by the method of MOM, ML and Bayesian. The properties of the MOM and MLE have been proved. It is shown that the estimated parameter by MOM and MLE is unbiased and having minimum variance for large “ n ”. The ML

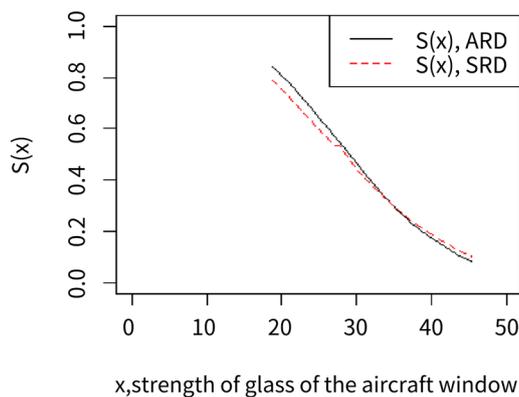


Figure 5. Survival graph of strength of glass of the aircraft window.

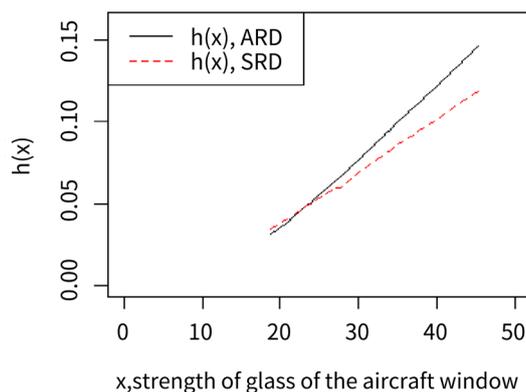


Figure 6. Hazard graph of strength of glass of the aircraft window.

estimator attains the Cramer Rao lower bound. Then the model is applied into two life time data sets. Kolmogorov Smirnov (K-S) test statistic is used to see the fit good on both data sets. The value of K-S for ARD is compared with some other well-known models named Lindley, Exponential, Quasi Lindley, Size-Biased Rayleigh and it is concluded that ARD is showing better fit on such kind of data sets as comparing to these models. At the end ARD, RD and SRD are used to show the graphical trend of the survival function and hazard function on both data sets. The survival function graph of SRD is decreasing more smoothly as comparing to other models. The hazard function graph of ARD for both data sets is increasing gradually. It means that the instantaneous failures are increasing. Overall it can be seen that ARD model can be a best alternative of the other well-known models and it is showing wide applications in the field of medical.

References

- [1] Shanker, R., Fesshaye, H. and Sharma, S. (2016) On Quasi Lindley Distribution and Its Applications to Model Lifetime Data. *International Journal of Statistical Distributions and Applications*, **2**, 1-7. <https://doi.org/10.11648/j.ijstd.20160201.11>
- [2] Lord Rayleigh, F.R.S. (1880) On the Resultant of a Large Number of Vibrations of

- the Same Pitch and of Arbitrary Phase. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **10**, 73-78.
<https://doi.org/10.1080/14786448008626893>
- [3] Siddiqui, M.M. (1962) Some Problems Connected with Rayleigh Distributions. *Journal of Research of the National Bureau of Standards*, **60D**, 167-174.
<https://doi.org/10.6028/jres.066D.020>
- [4] Hirano, K. (1986) Rayleigh Distributions. Wiley, New York.
- [5] Howlader, H.A. (1985) HPD Prediction Intervals for Rayleigh Distribution. *IEEE Transactions on Reliability*, **R-34**, 121-123.
<https://doi.org/10.1109/TR.1985.5221968>
- [6] Lalitha, S. and Mishra, A. (1996) Modified Maximum Likelihood Estimation for Rayleigh Distribution. *Communications in Statistics—Theory and Methods*, **25**, 389-401. <https://doi.org/10.1080/03610929608831702>
- [7] Abd-Elfattah, A.M., Hassan, A.S. and Ziedan, D.M. (2006) Efficiency of Maximum Likelihood Estimators under Different Censored Sampling Schemes for Rayleigh Distribution. *InterStat*, No. 1, 1-16.
- [8] Merovci, F. (2013) Transmuted Rayleigh Distribution. *Austrian Journal of Statistics*, **42**, 21-31. <https://doi.org/10.17713/ajs.v42i1.163>
- [9] Das, K.K. and Roy, T.D. (2011) Applicability of Length Biased Weighted Generalized Rayleigh Distribution. *Advances in Applied Science Research*, **2**, 320-327.
- [10] Hoffman, D. and Karst, O.J. (1975) The Theory of the Rayleigh Distribution and Some of Its Applications. *Journal of Ship Research*, **19**, 172-191.
- [11] Fisher, R.A. (1934) The Effect of Models of Ascertainment upon the Estimation of Frequencies. *Annals of Eugenics*, **6**, 13-25.
<https://doi.org/10.1111/j.1469-1809.1934.tb02105.x>
- [12] Rao, C.R. (1965) On Discrete Distributions Arising out of Methods of Ascertainment. *Sankhyā: The Indian Journal of Statistics, Series A*, **27**, 311-324.
- [13] Patil, G.P. and Ord, J.K. (1976) On Size-Biased Sampling and Related Form-Invariant Weighted Distribution. *Sankhyā: The Indian Journal of Statistics, Series B*, **38**, 48-61.
- [14] Patil, G.P. and Rao, C.R. (1978) Weighted Distributions and Size-Biased Sampling with Applications to Wildlife Populations and Human Families. *Biometrics*, **34**, 179-184. <https://doi.org/10.2307/2530008>
- [15] Gupta, R.C. and Keating, J.P. (1985) Relations for Reliability Measures under Length Biased Sampling. *Scandinavian Journal of Statistics*, **13**, 49-56.
- [16] Gross, A.J. and Clark, V.A. (1975) Survival Distributions: Reliability Applications in the Biometrical Sciences. Wiley, New York.
- [17] Fuller, E.J., Frieman, S., Quinn, J., Quinn, G. and Carter, W. (1994) Fracture Mechanics Approach to the Design of Glass Aircraft Windows: A Case Study. *Proceeding SPIE*, **2286**, 419-430.