A Note on Generalized Weibull-Rayleigh Distribution with Its Applications

A. Yahaya[1]; A. A. Yohanna[2]

Department of Statistics,
Ahmadu Bello University,
Zaria, Nigeria.
e-mail: abubakaryahaya@abu.edu.ng[1]; alaku26@yahoo.com[2]

***Abstract* — Rayleigh [14] derived the Rayleigh distribution from the amplitude of sound resulting from many important sources. The Rayleigh distribution is related to have a wide range of applications in diverse areas of human endeavors comprising life testing experiments, reliability analysis, applied statistics and clinical studies. This distribution is a special case of the two parameter Weibull distribution when the shape parameter take the value 2.In this paper we propose Generalized Weibull-Rayleigh distribution using the generalized family proposed by [11]. Some properties of the new distribution such as moments and moment generating function were studied. The estimation of the distribution parameters was conducted using the method of maximum likelihood. We also compared the proposed distribution to some other generalizations of Rayleigh distribution using some lifetime data sets.**

Keywords--*Rayleigh distribution, moments, moment generating function, maximum likelihood.*

I. Introduction

Rayleigh [14] derived the Rayleigh distribution from the amplitude of sound resulting from many important sources. The Rayleigh distribution has a wide range of applications including life testing experiments, reliability analysis, applied statistics and clinical studies. This distribution is a special case of the two parameter Weibull distribution when the shape parameter takes a value 2. A random variable X is said to have a Rayleigh distribution with parameter θ if its probability density function *(pdf)* is given by:

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

And the corresponding cumulative distribution function (*cdf*) is given as

$$G(x) = 1 - e^{-\frac{\theta}{2}x^2} \qquad (2)$$

for $x > 0, \theta > 0$ where θ is the scale parameter.

We have so many generalized families of distributions proposed by different researchers which have been used by others to extend so many standard or classical distributions to produce compound distributions found to be better than the classical ones. Some generalizations of the Rayleigh distribution have led to the development of other compound distributions such as the generalized Rayleigh distribution by [8], Bivariate generalized Rayleigh distribution by [1], Transmuted Rayleigh distribution by [9] and the Transmuted Inverse Rayleigh distribution by [3]. In a similar manner, for any continuous distribution with *cdf* G(x), and *pdf* $g(x)$, [11] defined the generalized Weibull family of distributions (*GW-G*) with two extra parameters $\alpha > 0$ and $\beta > 0$ to have its *pdf* $f(x)$ and *cdf* $F(x)$ respectively given by:

$$f(x) = \alpha\beta \frac{g(x)}{G^{'}(x)} \left(-\log\left[1 - G(x)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-G(x)\right]\right)^{\beta}} \qquad (3)$$

and

$$F(x) = 1 - e^{-\alpha\left(-\log\left[1-G(x)\right]\right)^{\beta}} \qquad (4)$$

where $g(x)$ and $G(x)$ are the *pdf* and the *cdf* of any continuous distribution respectively while $G^{'}(x) = 1 - G(x)$ and $\alpha > 0$ and $\beta > 0$ are the scale and shape parameters respectively.

The aim of this paper is to introduce a new continuous distribution called the Generalized Weibull-Rayleigh distribution (*GWRD)* from the proposed family by [11].

II. Materials and Methods

The G*WRD*

By taking the *pdf* (1) and *cdf* (2) of the Rayleigh distribution. The *cdf* and *pdf* of the *GWRD* are obtained from equation (3) and (4) as

$$F(x)=1-e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}} \tag{5}$$

and

$$f(x)=\frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}} \tag{6}$$

The *pdf* and *cdf* of the G*WRD* at chosen parameter values are displayed in Figures 1 and 2.

III. Analysis

In this section, we defined and discuss some properties of the *GWRD* distribution.

A. The Moments

Let X denote a continuous random variable, the n^{th} moment of X is given by;

$$\mu_n^{'}=E\left[X^n\right]=\int_0^\infty x^n f(x)dx \tag{7}$$

where*f(x)* the *pdf* of the Generalized Weibull-Rayleigh distribution is as given in equation (6).

Hence,

$$\mu_n^{'}=E\left[X^n\right]=\left(\frac{1}{\alpha}\right)^{\frac{n}{2\beta}}\left(\frac{2}{\theta}\right)^{\frac{n}{2}}\Gamma\left(\frac{n}{2\beta}+1\right) \tag{8}$$

The Mean

The mean of the *GWRD* can be obtained from the n^{th} moment of the distribution when *n=1* as follows:

$$E(X)=\left(\frac{1}{\alpha}\right)^{\frac{1}{2\beta}}\left(\frac{2}{\theta}\right)^{\frac{1}{2}}\Gamma\left(\frac{1}{2\beta}+1\right) \tag{9}$$

Also the second moment of the *GWRD* is obtained from the n^{th} moment of the distribution when *n=2* as

$$E(X^2)=\left(\frac{1}{\alpha}\right)^{\frac{1}{\beta}}\left(\frac{2}{\theta}\right)\Gamma\left(\frac{1}{\beta}+1\right) \tag{10}$$

The Variance

The n^{th} central moment or moment about the mean of *X*, say μ_n, can be obtained as

$$\mu_n=E\left[X-\mu_1^{'}\right]^n=\sum_{i=0}^{n}(-1)^i\binom{n}{i}\mu_1^{'i}\mu_{n-i}^{'} \tag{11}$$

The variance of *X* for *GWRD* is obtained from the central moment when *n=2,* that is,

$$Var(X)=E(X^2)-\{E(X)\}^2 \tag{12}$$

$$Var(X)=\left(\frac{1}{\alpha}\right)^{\frac{1}{\beta}}\left(\frac{2}{\theta}\right)\left\{\Gamma\left(\frac{1}{\beta}+1\right)-\left[\Gamma\left(\frac{1}{2\beta}+1\right)\right]^2\right\} \tag{13}$$

The variation, skewness and kurtosis measures can also be calculated from the non-central moments using some well-known relationships.

B. Moment generating function

This is another simple way writing all the moments in one mathematical function. In other words, the *mgf*generates the moments of *X* by differentiation i.e., for any real number say *k*, the k^{th} derivative of $M_X(t)$ evaluated at $t=0$ is the k^{th} moment $\mu_k^{'}$ of *X*.

The *mgf* of a random variable *X* can be obtained by

$$M_x(t)=E(e^{tx})=\int_0^\infty e^{tx}f(x)dx \tag{14}$$

$$M_x(t)=E(e^{tx})=\frac{\alpha\beta\theta^{\beta}}{2^{\beta-1}}\int_0^\infty x^{2\beta-1}e^{tx}e^{-\alpha\left(\frac{\theta}{2}x^2\right)^{\beta}}dx=\sum_{n=0}^{\infty}\frac{t^n}{n!}\mu_n^{'} \tag{15}$$

where

$$\mu_n^{'}=E\left[X^n\right]=\left(\frac{1}{\alpha}\right)^{\frac{n}{2\beta}}\left(\frac{2}{\theta}\right)^{\frac{n}{2}}\Gamma\left(\frac{n}{2\beta}+1\right)$$

is as defined in equation (8) previously.

C. Estimation of Parameters

Let $X_1,\cdots,X_n$ be a sample of size *n*containing independently and identically distributed random variables drawn from the *GWRD* with unknown parameters *α, β,* and *θ*. The *pdf* of the *GWRD* is then given as:

$$f(x)=\frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}} \tag{16}$$

The likelihood function is given as;

$$L\left(X_1,X_2,.....,X_n/\alpha,\beta,\theta\right)=\frac{(\alpha\beta\theta)^n\sum_{i=1}^{n}x_i}{\sum_{i=1}^{n}\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}e^{-\frac{\theta}{2}\sum_{i=1}^{n}x^2}\sum_{i=1}^{n}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} \quad (17)$$

$$\times e^{\sum_{i=1}^{n}\left\{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}\right\}}$$

The log-likelihood function, $l(\varphi)=\log L\left(X_1,X_2,.....,X_n/\alpha,\beta,\theta\right)$, is obtained by:

$$l(\varphi)=n\log\alpha+n\log\beta+n\log\theta+\sum_{i=1}^{n}\log x_i-\frac{\theta}{2}\sum_{i=1}^{n}x_i^2-\sum_{i=1}^{n}\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]$$
$$+(\beta-1)\sum_{i=1}^{n}\log\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)-\alpha\sum_{i=1}^{n}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta} \quad (18)$$

Taking the partial derivative of the log likelihood function, $l(\varphi)$ with respect to *α, β* and *θ* respectively, we get the *MLEs* as follows;

$$\frac{\partial l(\varphi)}{\partial\alpha}=\frac{n}{\alpha}-\sum_{i=1}^{n}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta} \quad (19)$$

$$\frac{\partial l(\varphi)}{\partial\beta}=\frac{n}{\beta}+\sum_{i=1}^{n}\log\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)$$
$$-\alpha\sum_{i=1}^{n}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}\log\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right) \quad (20)$$

$$\frac{\partial l(\varphi)}{\partial\theta}=\frac{n}{\theta}+\frac{(\beta-1)}{2}\sum_{i=1}^{n}\frac{x_i^2e^{-\frac{\theta}{2}x^2}}{\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)\left(1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right)}$$
$$-\frac{\alpha\beta}{2}\sum_{i=1}^{n}\frac{\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1}\left(x_i^2e^{-\frac{\theta}{2}x^2}\right)}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]} \quad (25)$$

Solving the equations of $\frac{dl}{d\alpha}=0,\ \frac{\partial l}{\partial\beta}=0,$ and $\frac{\partial l}{\partial\theta}=0$ for each parameter shall provide the maximum likelihood estimates (*MLEs*) of parameters α,β and θ. However, these solutions cannot be obtained analytically except numerically with the help of some statistical software/packages.

IV. Results

In this section, we have considered the adequacy of the *GWRD* compared to those of three generalizations of the Rayleigh model including the Weibull-Rayleigh distribution (*WRD*), the Transmuted Rayleigh distribution (*TRD*) and the Rayleigh distribution (*RD*) and three real life data sets for fitting the above selected models with their descriptive statistics.

Data set I: This data set represents the strength of 1.5cm glass fibers initially collected by members of staff at the UK national laboratory. It has been used by [1], [4], [5], [13] as well as [16].

Data set II: This data set represents the lifetime's data relating to relief times (in minutes) of 20 patients receiving an analgesic and reported by [6] and has been used by [15].

Data set III: The second data set represents 66 observations of the breaking stress of carbon fibres of 50mm length (in GPa) given by [12].

We also provide some histograms and densities for the three data sets as shown in **Figures 5, 6** and **7**below respectively.

A. Figures and Tables

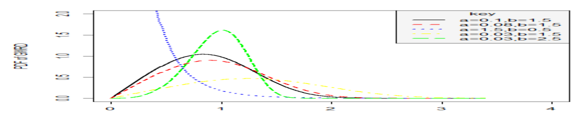

Figure 1: The graph of pdf of the GWRD using different parameter values where $a=\alpha,b=\beta\ for\theta=6.5$.

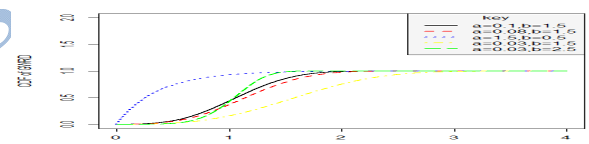

Figure 2: The graph of *cdf* of the G*WRD* at different parameter values where $a=\alpha,b=\beta\ for\theta=6.5$.

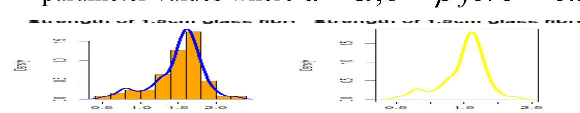

Figure 3: A histogram and density plot for the strength of 1.5cm glass fibres (Data set I)

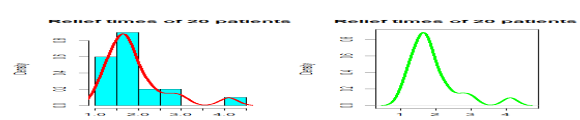

Figure 4: A Histogram and density plot for the Relief times of 20 patients (Data set II)

Figure 5: A Histogram and density plot for the Breaking stress of carbon fibres (Data set III)

Table 1: Performance of the selected models using the *AIC*, *CAIC*, *BIC and HQIC* values of the models evaluated at the *MLEs* based on data set I.

Distributions	Parameter estimates	-ll=(-log-likelihood value)	*AIC*	*CAIC*	*BIC*	*HQIC*	Ranks of models performance
GWRD	$\hat{\theta}$=0.3965 $\hat{\alpha}$=3.6027 $\hat{\beta}$=8.9e+6	-890.6894	-1775.3788	-1775.2826	-1775.9808	-1779.8481	**1**
WRD	$\hat{\theta}$=0.3918 $\hat{\alpha}$=2.2981 $\hat{\beta}$=2.2634	14.6575	35.3150	35.7218	34.7130	30.8457	**2**
TRD	$\hat{\theta}$=1.5061 $\hat{\lambda}$=-1.6251	27.0675	58.1350	58.3350	57.7337	55.1555	**3**
RD	$\hat{\theta}$=0.8426	49.7909	101.5818	101.6474	101.3811	100.0920	**4**

Table 2: Performance of the distribution using the *AIC*, *CAIC*, *BIC and HQIC* values of the models based on data set II.

Distributions	Parameter estimates	-ll=(-log-likelihood value)	*AIC*	*CAIC*	*BIC*	*HQIC*	Ranks of models performance
GWRD	$\hat{\theta}$=0.1166 $\hat{\alpha}$=1.0563 $\hat{\beta}$=1.8e+7	-290.5456	-575.0912	-573.5912	-577.1881	-580.4055	**2**
WRD	$\hat{\theta}$=0.0264 $\hat{\alpha}$=38.4269 $\hat{\beta}$=1.3273	20.9109	47.8218	49.3218	45.7249	42.5075	**3**
TRD	$\hat{\theta}$=1.3652 $\hat{\lambda}$=-1.2e+26	-1154.409	-2308.818	-2304.1121	-2306.2159	-2308.3609	**1**
RD	$\hat{\theta}$=0.4900	22.4788	46.9576	47.1798	46.2586	45.1862	**4**

Table 3: Performance of the distribution using the *AIC*, *CAIC*, *BIC and HQIC* values of the models based on data set III.

Distributions	Parameter estimates	-ll=(-log-likelihood value)	*AIC*	*CAIC*	*BIC*	*HQIC*	Ranks of models performance
GWRD	$\hat{\theta}$=0.1090 $\hat{\alpha}$=0.8462 $\hat{\beta}$=5.0865	50.0391	106.0782	106.4653	105.5368	101.6379	**1**
WRD	$\hat{\theta}$=0.1197 $\hat{\alpha}$=1.3299 $\hat{\beta}$=1.3024	86.5747	179.1494	179.5365	178.6080	174.7092	**2**
TRD	$\hat{\theta}$=0.3480 $\hat{\lambda}$=-0.9587	88.8744	181.7488	181.9393	181.3879	178.7887	**3**
RD	$\hat{\theta}$=0.2382	98.2084	198.4168	198.4793	198.2363	196.9367	**4**

V. Discussion

The plot for the *pdf* shows that the *GWRD* is positively skewed and hence good model for modeling rightly-skewed data sets. From the histograms and densities shown above in Figures 5, 6 and 7 for the three data sets respectively, we observed that the first data set is negatively skewed, the second data set is positively skewed while the third data is approximately normal and therefore suitable for distributions that are skewed to the left, right and symmetry respectively.

Besides the models listed above, these three data sets could also be analyzed using a distribution which has various shapes depending on the values of the parameters and which the generalized Weibull-Rayleigh distribution is a special case. In **Table 1,** the values of the parameter MLEs and the corresponding values of *-ll*, *AIC*, *BIC*, *CAIC* and *HQIC* for each model show that the *GWRD* has better performance compared to the *WRD*, *TRD* and *RD*. This also agrees with the fact that generalizing any continuous distribution provides a distribution with a better fit than the classical distribution.

Table 2 also shows the parameter estimates to each one of the four fitted distributions for the second data set (data set II), the table also provide the values of *-ll*, *AIC*, *BIC*, *CAIC* and *HQIC* of the fitted models. The values in **Table 2** indicate that the *TRD* has better performance with the lowest values of *AIC*, *CAIC*, *BIC* and *HQIC* followed by the *GWRD*, *WRD* and *RD*. The secret behind this performance is that, the second data set has a higher degree of skewness and kurtosis meanwhile, our proposed model has various shapes with a moderate peak and skewness.

Similarly, **Table 3** presentsthe parameter estimates and the values of *-ll*, *AIC*, *BIC*, *CAIC* and *HQIC* for the four fitted models. The values in the above table also provide evidence that the *GWRD* has better performance with the lowest values of *AIC*, *CAIC*, *BIC* and *HQIC* compared to the other three models. This also implies that the *GWRD* could be used to model all kinds of data sets since the above data set is approximately normal. It also shows that the *GWRD* has various shapes as earlier stated and proven by its graph of *pdf* and other properties.

VI. Conclusion

In this paper, a new distribution has been proposed. Some mathematical and statistical properties of the proposed distribution have been studied appropriately. The derivations of some expressions for its moments and moment generating function, have been done appropriately. Some plots of the distribution revealed that it can take any shape depending on values of the parameters.

The model parameters have been estimated using the method of maximum likelihood estimation. The results of the three applications showed that the proposed distribution (Generalized Weibull-Rayleigh distribution) performs better than the Weibull-Rayleigh, transmuted Rayleigh and the Rayleigh distributions irrespective of the nature of the data sets. This implies that the proposed distribution can be used when we have skewed and symmetric datasets.

References

[1] Abdel-Hady, D. H. (2013). Bivariate Generalized Rayleigh Distribution. *Journal of Applied Sciences Research, 9* (9): 5403 – 5411.

[2] Afify, A. Z. and Aryal, G. (2016).The Kumaraswamy exponentiated Frechet distribution. *Journal of Data Science, 6*: 1-19.

[3] Ahmad, A., Ahmad, S. P. and Ahmed, A. (2014).Transmuted Inverse Rayleigh distribution: a generalization of the Inverse Rayleigh distribution. *Mathematical Theory and Modeling,* 4(7): 90 – 98.

[4] Barreto-Souza, W. M., Cordeiro, G. M. and Simas, A. B. (2011). Some results for beta Frechet distribution. *Communication in Statistics: Theory and Methods*. 40: 798-811.

[5] Bourguignon, M., Silva, R. B. and Cordeiro, G. M. (2014). The Weibull-G Family of Probability Distributions. *Journal of Data Science,* 12: 53-68.

[6] Gross A., J and Clark, V., A. (1975). Survival Distributions Reliability Applications in the Biometrical Sciences. John Wiley, New York, USA.

[7] Kenney, J. F. & Keeping, E. S. (1962). Mathematics of Statistics, 3 edn, *Chapman & Hall Ltd*, *New Jersey*.

[8] Kundu, D., and Raqab, M. Z. (2005). Generalized Rayleigh Distribution: Different methods of estimations. *Computational Statistics and Data Analysis, 49*: 187 – 200.

[9] Merovci, F. (2013). The transmuted Rayleigh distribution. *Australian Journal of Statistics*, 22(1): 21–30.

[10] Merovci, F., and Elbatal, I. (2015).Weibull Rayleigh Distribution: Theory and Applications. *Applied Mathematics and Information Science*, 9(5): 1-11.

[11] Cordeiro, G. M., Ortega, E. M. M. and Ramires, T. G. (2015), A new generalized Weibull family of distributions: Mathematical properties and applications; *Journal of Statistical Distributions and Applications* (2015) 2:13DOI 10.1186/s40488-015-0036-6

[12] Nichols, M. D. and Padgett, W. J. (2006). A bootstrap control chart for Weibull percentiles. *Quality Reliability Engineering Int.*, 22: 141-151.

[13] Oguntunde, P. E., Balogun, O. S, Okagbue, H. I, and Bishop, S. A. (2015). The Weibull-Exponential Distribution: Its Properties and Applications. *Journal of Applied Science,* 15(11): 1305-1311.

[14] Rayleigh, J. (1980). On the resultant of a large number of vibrations of the same pitch and of arbitrary phase, *Philosophy and Management*, 10: 73–78.

[15] Shanker, R., Fesshaye, H. and Sharma, S. (2016). On Two - Parameter Lindley Distribution and its Applications to Model Lifetime Data. *Biometrics and Biostatics International Journal* 3(1): 00056. DOI: 10.15406/bbij.2016.03.00056

[16] Smith, R. L. and Naylor, J. C. (1987).A Comparison of Maximum Likelihood and Bayesian Estimators for the Three-Parameter Weibull Distribution. *Journal of Applied Statistics,* 36: 358-369.

.