NJBAS

Available online at http://www.ajol.info/index.php/njbas/index
Nigerian Journal of Basic and Applied Science (December, 2017), 25(2): 130-142
DOI: http://dx.doi.org/10.4314/njbas.v25i2.14

ISSN 0794-5698

Full Length Research Article

The Gamma-Rayleigh Distribution and Applications to Survival Data

*[1]E. E. E. Akarawak, [2]I. A. Adeleke, [3]R. O. Okafor
[1,3]Department of Mathematics, University of Lagos, Akoka (Nigeria)
[2]Department of Actuarial Science and Insurance, University of Lagos, Akoka (Nigeria)
[*Corresponding Author: Email: eakarawak@unilag.edu.ng; ☎: +2348137678288]

ABSTRACT

Studies on probability distribution functions and their properties are needful as they are very important in modeling random phenomena. However, research has shown that some real life data can be modeled more adequately by distributions obtained as combination of two random variables with known probability distributions. This paper introduces the Gamma-Rayleigh distribution (GRD) as a new member of the Gamma-*X* family of generalized distributions. The Transformed-Transformer method is used to combine the Gamma and Rayleigh distributions. Various properties of the resulting two-parameter Gamma-Rayleigh distribution, including moments, moment generating function, survival function and hazard function are derived. Results of simulation study reveals that the distribution is unimodal, skewed and normal-type for some values of the shape parameter. The distribution is also found to relate with the Gamma, Rayleigh and Generalized-Gamma distributions. The method of maximum likelihood has been used to estimate the shape and scale parameters of the distribution. To illustrate its adequacy in modelling real life data the distribution is fitted to two survival data sets. The results show that the distribution produced fits that are competitive and compared better, in some cases, to the Gamma, Rayleigh, Weibull and Lognormal distributions.

Keywords: Gamma-*X* family, Gamma-Rayleigh distribution, Maximum Likelihood estimators, Survival data.

INTRODUCTION

Studies on probability distribution functions and their properties are needful as they are very important in modeling random phenomena. However, research has shown that some real life data that cannot be modeled adequately by existing standard distributions are sometimes found to follow distributions of some combinations of two or more random variables with known probability distributions.

Developments on generating new distributions to model naturally occurring phenomena have led to a number of new distributions being defined and studied. Some of the earlier works include those of Marsaglia (1965), Press (1969), Basu and Lochner (1972), and Lee *et al.* (1979). The trend has been on the increase in recent years due to increasing need for adequate distributions to model some data arising in practice (Gupta and Kundu, 1999; Eugene *et al.*, 2004; Famoye *et al.*, 2005; Akinsete *et al.*, 2008; Alzaatreh *et al.*, 2013b, Adeleke *et al.*, 2013; Akarawak *et al.*, 2013 and Akarawak *et al.*, 2015). Furthermore, several studies (Mudholkar and Srivastava, 1993; Gupta and Kundu, 2001; Pal *et al.*, 2006) have shown that distributions of combined random variables are more flexible, perform better and have wider applicability.

Motivated by the recent developments in generating new distributions and the need for continuous extension and generalizations to more complex situations, this research paper introduces and studies the Gamma-Rayleigh Distribution (GRD). The Gamma and Rayleigh distributions are well-known survival distributions; however, there might be some survival data situations in which the two distributions may not fit so well. Combining them might therefore yield better results. Since estimation of parameters of a newly generated distribution is fundamental to application, the method of maximum likelihood is used in this work to estimate parameters of GRD. The distribution is then applied to two survival data sets. MATLAB R2011b has been used for implementations.

Methodology
This article is organized as follows: derivation and presentation of aspects of the Gamma-Rayleigh Distribution; parameter estimation of the Gamma-Rayleigh Distribution and study of the simulation and applications of the distribution.

THE GAMMA-RAYLEIGH DISTRIBUTION (GRD)

Method: The Transformed-Transformer Technique of generating New Family of Continuous Distributions

The Transformed-Transformer method was recently introduced by Alzaatreh *et al.* (2013a) for generating families of continuous distributions. Let $F(v)$ be the cumulative distribution function (cdf) of any random variable V and $p(y)$ the probability density function (pdf) of a random variable Y defined on $[0,\infty)$. The cdf of a generalized family of distributions is given as

$$G(v)=\int_0^{-\log(1-F(v))} p(y)dy\,, \qquad (2.1)$$

The family of distributions defined by (2.1) is called the 'Transformed-Transformer' family (or the *Y-V* family), where the random variable *Y* is being transformed by another random variable *V* (the transformer) into a new random variable with the support $[0,\infty)$. According to Alzaatreh *et al.* (2013a), the corresponding pdf of the generalized distribution in (2.1) is given by

$$g(v)=\frac{f(v)}{1-F(v)}\,p(-\log[1-F(v)]) \qquad (2.2)$$

$$=h(v)p(-\log[1-F(v)]) \qquad (2.3)$$

$$=h(v)p[H(v)], \qquad (2.4)$$

Where *h(v)* is the hazard function and *H(v)* is the cumulative hazard function of *V* with the cdf*F(v)*, which makes *g(v)* as arising from a weighted hazard function. One of the *Y-V* families defined by Alzaatreh *et al*. (2013a) is the Gamma-*V* family, obtained by letting *Y* follow the Gamma distribution with parameters α and β and having the pdf,

$$f(y)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}y^{\alpha-1}e^{-\beta y}; \; y>0;\ \alpha,\ \lambda>0. \qquad (2.5)$$

According to Alzaatreh (2013a), the pdf of the Gamma-*V* generalized family of distribution is given as:

$$g_V(v)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}f(v)[-\ln\{1-F(v)\}]^{\alpha-1}[1-F(v)]^{\beta-1} \qquad (2.6)$$

where *f*(*v*) and *F*(*v*) are the pdf and cdf of any random variable *V*.

2.2 Derivation of the Gamma-Rayleigh Distribution (GRD)

2.2.1 Derivation of the pdf and cdf of GRD

The pdf and cdf of the Gamma-Rayleigh distribution is derived in this section as a class of Gamma-*V* family of generalized distributions.

Theorem 2.1:
Let the pdf of a gamma distribution be $f(x)=\frac{\lambda}{\Gamma(\alpha)}(\lambda x)^{\alpha-1}e^{-\lambda x}$ and *V* Rayleigh distributed random variable with pdf $f(v)=\frac{v}{\sigma^2}e^{\frac{-v^2}{2\sigma^2}}$, and cdf $F(v)=1-e^{\frac{-v^2}{2\sigma^2}}$. Then the pdf of the Gamma-Rayleigh distribution is given by:

$$g_V(v)=\frac{2\theta^{\alpha}}{\Gamma(\alpha)}v^{2\alpha-1}\exp(-\theta v^2); \quad v>0;\alpha,\theta>0 \,, \qquad (2.7)$$

Where, $\Gamma(\alpha)$ is the gamma function.

Proof:
The pdf of the Gamma-*V* family of distribution is given by:

$$g_V(v)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}f(v)[-\ln\{1-F(v)\}]^{\alpha-1}[1-F(v)]^{\beta-1} \,, \qquad (2.8)$$

Let *V* follow the Rayleigh distribution with pdf $f(v)=\frac{v}{\sigma^2}e^{\frac{-v^2}{2\sigma^2}}$, and cdf $F(v)=1-e^{\frac{-v^2}{2\sigma^2}}$,

such that $1-F(v)=e^{\frac{-v^2}{2\sigma^2}}$. Then the pdf in (2.8) becomes:

$$g_V(v)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\frac{v}{\sigma^2}e^{\frac{-v^2}{2\sigma^2}}\left[-\ln\left\{e^{\frac{-v^2}{2\sigma^2}}\right\}\right]^{\alpha-1}\left[e^{\frac{-v^2}{2\sigma^2}}\right]^{\beta-1},$$

$$=\frac{\beta^{\alpha}}{\Gamma(\alpha)\sigma^2}ve^{\frac{-v^2}{2\sigma^2}}\left[\frac{v^2}{2\sigma^2}\right]^{\alpha-1}\left[e^{\frac{-v^2(\beta-1)}{2\sigma^2}}\right],$$

$$= \frac{\beta^{\alpha}}{(2\sigma^2)^{\alpha-1}\Gamma(\alpha)\sigma^2} v^{2\alpha-1} e^{\frac{-v^2}{2\sigma^2}} e^{\frac{-\beta v^2}{2\sigma^2}} e^{\frac{v^2}{2\sigma^2}},$$

$$= \frac{2\beta^{\alpha}}{2^{\alpha}\Gamma(\alpha)\sigma^{2\alpha}} v^{2\alpha-1} e^{\frac{-\beta v^2}{2\sigma^2}},$$

$$= \left(\frac{\beta}{2\sigma^2}\right)^{\alpha} \frac{2}{\Gamma(\alpha)} v^{2\alpha-1} e^{\frac{-\beta v^2}{2\sigma^2}}, \quad (2.9)$$

Replacing $\frac{\beta}{2\sigma^2}$ by θ, (2.9) becomes:

$$g_V(v) = \frac{2\theta^{\alpha}}{\Gamma(\alpha)} v^{2\alpha-1} e^{-\theta v^2}.$$

Any random variable *V* that has the probability density function given in (2.7) is said to have the Gamma-Rayleigh distribution with shape parameter α and scale parameter θ and written as *V* ~ GRD(α, θ).

Theorem 2.2
The function given in (2.7) is a valid probability density function.

Proof:
Required to prove that $\int_0^{\infty} g(v)dv = 1$.

$$\int_0^{\infty} g(v)dv = \frac{2\theta^{\alpha}}{\Gamma(\alpha)} \int_0^{\infty} v^{2\alpha-1} e^{-\theta v^2} dv, \quad (2.10)$$

Let, $u = \theta v^2 \Rightarrow v = \left(\frac{u}{\theta}\right)^{1/2}$, so that

$$dv = \frac{du}{2\theta^{\frac{1}{2}} u^{\frac{1}{2}}}, \quad (2.11)$$

Then (2.10) reduces to:

$$\int_0^{\infty} g(v)dv = \frac{2\theta^{\alpha}}{\Gamma(\alpha)} \int_0^{\infty} \left(\left(\frac{u}{\theta}\right)^{1/2}\right)^{2\alpha-1} e^{-u} \frac{du}{2\theta^{\frac{1}{2}} u^{\frac{1}{2}}},$$

$$= \frac{1}{\Gamma(\alpha)} \int_0^{\infty} u^{\alpha-1} e^{-u} du = \frac{\Gamma(\alpha)}{\Gamma(\alpha)} = 1. \quad (2.12)$$

Theorem 2.3:
The cumulative distribution function (cdf) of the Gamma-Rayleigh distribution is given by:

$$G(v) = \frac{\gamma(\alpha, \theta v^2)}{\Gamma(\alpha)}; \quad (2.13)$$

where, $\gamma(\alpha, \theta v^2) = \int_0^{\theta v^2} u^{\alpha-1} e^{-u} du$ is the lower incomplete gamma function and $\Gamma(\alpha) = \int_0^{\infty} u^{\alpha-1} e^{-u} du$ is the complete gamma function.

Proof:
We prove the result using the fact that $G(v) = \int_0^{v} g(x)dx$,

$$G(v) = \frac{2\theta^{\alpha}}{\Gamma(\alpha)} \int_0^{v} x^{2\alpha-1} e^{-\theta x^2} dx, \quad (2.14)$$

Changing variable as in (2.11), the result follows.

2.2.2 Relationship with Other Distributions
The Gamma-Rayleigh distribution has two parameters α and θ.

Corollary 2.1: When α = 1 the pdf of Gamma-Rayleigh distribution reduces to the pdf of the Rayleigh distribution with parameter $\sigma^2 = \frac{1}{2\theta}$.

Preposition 2.1: The Gamma-Rayleigh distribution is a special case of the generalized gamma distribution when *p* = 2, *d* = 2α and *a*= $\theta^{-\frac{1}{2}}$.

Proof of Preposition 2.1:
The generalized gamma distribution (Stacy, 1962) has a pdf of the form:

$$f(x) = \frac{\left(\frac{p}{a^d}\right) x^{d-1} \exp\left[\left(-\frac{x}{a}\right)^p\right]}{\Gamma\left(\frac{d}{p}\right)}; \quad x \geq 0; a, p, d > 0. \quad (2.15)$$

When *p* = 2, *d* = 2α and *a*= $\theta^{-\frac{1}{2}}$, (2.15) reduces to the Gamma-Rayleigh pdf given in (2.7).

Corollary 2.2: If the random variable *X* follows the Gamma-Rayleigh distribution with parameters α and θ, then the Random variable $Y = X^2$ follows the gamma distribution with parameters α and θ.

Proof of Corollary 2.2:

Let X ~ GRD (α, θ) and $Y = X^2$ such that $x = y^{\frac{1}{2}}$. Then $\frac{dx}{dy} = \frac{1}{2y^{1/2}}$.

$$f(y) = g(x)\left|\frac{dx}{dy}\right|; x = y^{\frac{1}{2}},$$

$$= \frac{\theta^{\alpha} y^{\alpha-1} e^{-\theta y}}{\Gamma(\alpha)}, \qquad (2.16)$$

where (2.16) is the pdf of the gamma distribution with parameters α and θ.

Plots of the Pdf and Cdf of the Gamma-Rayleigh Distribution

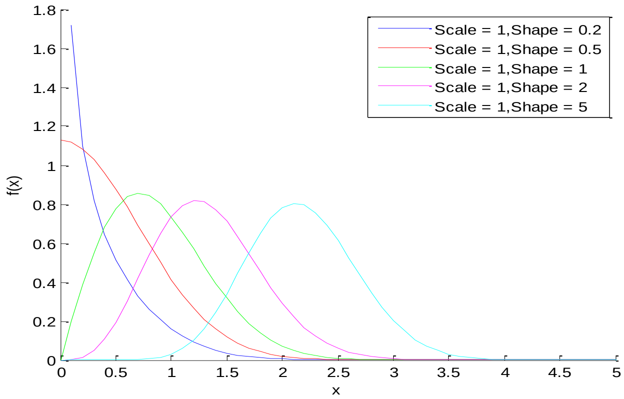

Figure 1: Pdf Plot of GRD for different values of the shape parameter

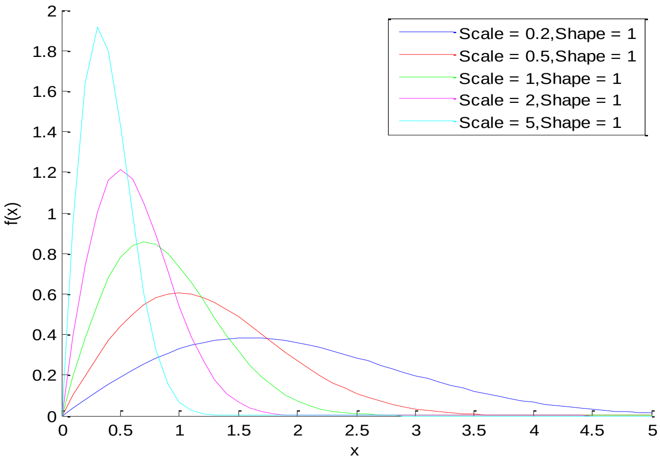

Figure 2: Pdf Plot of GRD for different values of the scale parameter

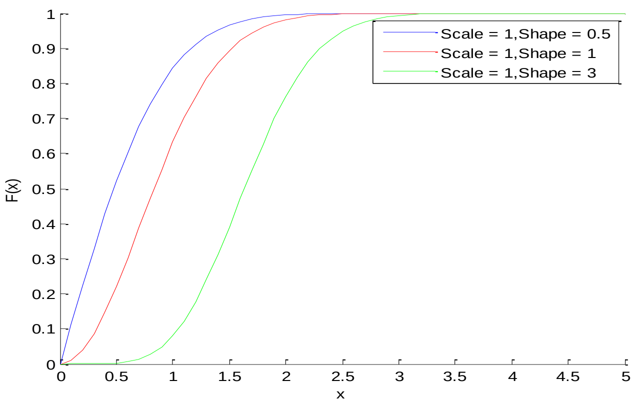

Figure 3: Cdf Plots for GRD

From Figures 1 and 2, it is observed that as the shape parameter varies from 0.5 to 5, the Gamma-Rayleigh distribution increasingly resembles the normal distribution; the spread reduces with increasing scale parameter. Furthermore, the distribution shows positive skewness for different values of the scale parameter. The cdf plots shown in Figure 3 increase from 0 to 1 as *X* increases.

SOME PROPERTIES OF THE GAMMA-RAYLEIGH DISTRIBUTION

Moments

Theorem 2.4

The *r*th non-central moment of a Gamma-Rayleigh random variable *X* is given by:

$$E(X^r) = \frac{\Gamma\left(\alpha + \frac{r}{2}\right)}{\theta^{\frac{r}{2}}\Gamma(\alpha)} \qquad (2.17)$$

Proof:

$$E(X^r) = \frac{2\theta^{\alpha}}{\Gamma(\alpha)}\int_0^{\infty} x^{2\alpha+r-1}e^{-\theta x^2}dx, \qquad (2.18)$$

A change of variable from *x* to *u* reduces (2.18) to:

$$= \frac{1}{\theta^{\frac{r}{2}}\Gamma(\alpha)}\int_0^{\infty} u^{\alpha+\frac{r}{2}-1}e^{-u}du, \qquad (2.19)$$

$$= \frac{\Gamma\left(\alpha + \frac{r}{2}\right)}{\theta^{\frac{r}{2}}\Gamma(\alpha)},$$

From (2.17) the first four moments are given below:

$$u_1' = E(X) = \frac{\Gamma\left(\frac{1}{2}+\alpha\right)}{\theta^{\frac{1}{2}}\Gamma(\alpha)},$$

$$u_2' = E(X^2) = \frac{\Gamma(1+\alpha)}{\theta\Gamma(\alpha)} = \frac{\alpha\Gamma(\alpha)}{\theta\Gamma(\alpha)} = \frac{\alpha}{\theta},$$

(2.20)

$$u_3' = E(X^3) = \frac{\Gamma\left(\frac{3}{2}+\alpha\right)}{\theta^{\frac{3}{2}}\Gamma(\alpha)},$$

$$u_4' = E(X^4) = \frac{\Gamma(2+\alpha)}{\theta^2\Gamma(\alpha)}.$$

(2.21)

and, $Var(X) = E(X^2) - [E(X)]^2$,

$$= \frac{\Gamma(\alpha+1)}{\theta\Gamma(\alpha)} - \frac{\left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2}{\theta(\Gamma(\alpha))^2}$$

$$= \frac{\alpha(\Gamma(\alpha))^2 - \left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2}{\theta(\Gamma(\alpha))^2}. \qquad (2.22)$$

The standard deviation of *X* is given by:

$\sigma = Std(X) = \sqrt{Var(X)}$

$$= \sqrt{\frac{\alpha(\Gamma(\alpha))^2 - \left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2}{\theta(\Gamma(\alpha))^2}}. \qquad (2.23)$$

Coefficient of variation of *X*

Coefficient of variation of *X* is given by:

$$CV = \frac{Std(X)}{E(X)} = \sqrt{\frac{\alpha(\Gamma(\alpha))^2 - \left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2}{\theta(\Gamma(\alpha))^2}}\;\frac{\theta^{\frac{1}{2}}\Gamma(\alpha)}{\Gamma\left(\frac{1}{2}+\alpha\right)},$$

(2.2

$$= \frac{\sqrt{\alpha(\Gamma(\alpha))^2 - \left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2}}{\Gamma\left(\frac{1}{2}+\alpha\right)}. \qquad (2.25)$$

The coefficient of variation given in (2.25) is expressed in terms of α only.

Skewness and Kurtosis

The skewness of a distribution is a measure of its departure from symmetry, while the kurtosis is a measure of its peakedness. The skewness and kurtosis of a distribution are respectively given by:

Skewness = $\dfrac{E(X-\mu)^3}{\sigma^3} = \dfrac{\mu_3}{\sigma^3}$, (2.26)

Kurtosis = $\dfrac{E(X-\mu)^4}{\sigma^4} = \dfrac{\mu_4}{\sigma^4}$. (2.27)

Where, $\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^3$ and $\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4$. (2.28)

For the Gamma-Rayleigh distribution,

$$\mu_3 = \left[\frac{\Gamma\left(\frac{3}{2}+\alpha\right)}{\theta^{\frac{3}{2}}\Gamma(\alpha)} - \frac{3\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma(\alpha+1)}{\theta^{\frac{3}{2}}(\Gamma(\alpha))^2} + \frac{2\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^3}{\theta^{\frac{3}{2}}(\Gamma(\alpha))^3}\right],$$

$$= \frac{(\Gamma(\alpha))^2\Gamma\left(\frac{3}{2}+\alpha\right) - 3\Gamma(\alpha)\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma(\alpha+1) + 2\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^3}{\theta^{\frac{3}{2}}(\Gamma(\alpha))^3}. \quad (2.29)$$

$$\mu_4 = \left[\frac{\Gamma(\alpha+2)}{\theta^2\Gamma(\alpha)} - \frac{4\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma\left(\alpha+\frac{3}{2}\right)}{\theta^2(\Gamma(\alpha))^2} + \frac{6\Gamma(\alpha+1)\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^2}{\theta^2\Gamma(\alpha)^3} - \frac{3\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^4}{\theta^2(\Gamma(\alpha))^4}\right],$$

$$= \frac{(\Gamma(\alpha))^2\Gamma(\alpha+2) - 4(\Gamma(\alpha))^2\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma\left(\alpha+\frac{3}{2}\right) + 6\Gamma(\alpha)\Gamma(\alpha+1)\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^2 - 3\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^4}{\theta^2(\Gamma(\alpha))^4}, \quad (2.30)$$

Then, $\text{Skewness} = \dfrac{\mu_3}{\sigma^3} = \dfrac{\dfrac{(\Gamma(\alpha))^2\Gamma\left(\frac{3}{2}+\alpha\right) - 3\Gamma(\alpha)\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma(\alpha+1) + 2\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^3}{\theta^{\frac{3}{2}}(\Gamma(\alpha))^3}}{\left[\sqrt{\dfrac{\alpha(\Gamma(\alpha))^2 - \left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2}{\theta(\Gamma(\alpha))^2}}\right]^3}$

$$= \frac{(\Gamma(\alpha))^2\Gamma\left(\frac{3}{2}+\alpha\right) - 3\Gamma(\alpha)\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma(\alpha+1) + 2\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^3}{\theta^{\frac{3}{2}}(\Gamma(\alpha))^3} \times \frac{\theta^{\frac{3}{2}}(\Gamma(\alpha))^3}{\left[\alpha(\Gamma(\alpha))^2 - \left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2\right]^{\frac{3}{2}}}, \quad (2.31)$$

$$\text{Kurtosis} = \frac{\mu_4}{\sigma^4} = \frac{(\Gamma(\alpha))^2\Gamma(\alpha+2) - 4(\Gamma(\alpha))^2\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma\left(\alpha+\frac{3}{2}\right) + 6\Gamma(\alpha)\Gamma(\alpha+1)\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^2 - 3\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^4}{\theta^2(\Gamma(\alpha))^4}$$

$$\div\left[\sqrt{\frac{\alpha(\Gamma(\alpha))^2 - \left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2}{\theta(\Gamma(\alpha))^2}}\right]^3$$

$$= \frac{(\Gamma(\alpha))^2\Gamma(\alpha+2) - 4(\Gamma(\alpha)^2)\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma\left(\alpha+\frac{3}{2}\right) + 6\Gamma(\alpha)\Gamma(\alpha+1)\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^2 - 3\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^4}{\theta^2(\Gamma(\alpha))^4} \times \frac{\theta^2(\Gamma(\alpha))^4}{\left[\alpha(\Gamma(\alpha))^2 - \left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2\right]}$$

$$= \frac{(\Gamma(\alpha))^2\Gamma(\alpha+2) - 4(\Gamma(\alpha))^2\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma\left(\alpha+\frac{3}{2}\right) + 6\Gamma(\alpha)\Gamma(\alpha+1)\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^2 - 3\left(\Gamma\left(\alpha+\frac{1}{2}\right)\right)^4}{\left[\alpha(\Gamma(\alpha))^2 - \left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2\right]^4}. \quad (2.32)$$

Corllary 2.1: Moment Generating Function

The moment generating function (mgf) of the Gamma-Rayleigh random variable V is given by:

$$M_V(t) = \sum_{r=0}^{\infty} \frac{\Gamma\left(\alpha+\frac{r}{2}\right)}{\theta^{\frac{r}{2}}\Gamma(\alpha)} \frac{t^r}{r!}. \quad (2.33)$$

Proof:

The moment generating function of a continuous random variable V is defined by:

$M_V(t) = E(e^{tV})$,

$$= \int_0^{\infty} e^{tv} g(v)dv, \quad (2.34)$$

But, $e^{tv} = 1 + tv + \frac{(tv)^2}{2!} + \frac{(tv)^3}{3!} + ... = \sum_{r=0}^{\infty} \frac{t^r r^r}{r!}$, (2.35)

Therefore, (2.34) becomes:

$$M_V(t) = \int_0^{\infty}\left(\sum_{r=0}^{\infty}\frac{t^r v^r}{r!}\right) g(v)dv,$$

$$= \sum_{r=0}^{\infty}\frac{t^r}{r!}\left(\int_0^{\infty} v^r g(v)dv\right) = \sum_{r=0}^{\infty}\frac{t^r \mu_r'}{r!}, \quad (2.36)$$

For GRD random variable V, $\mu_r' = \frac{\Gamma\left(\alpha+\frac{r}{2}\right)}{\theta^{\frac{r}{2}}\Gamma(\alpha)}$, therefore, the moment generating function of V is given by:

$$M_V(t) = \sum_{r=0}^{\infty} \frac{\Gamma\left(\alpha+\frac{r}{2}\right)}{\theta^{\frac{r}{2}}\Gamma(\alpha)} \frac{t^r}{r!}.$$

Corollary 2.2: Cumulants

According to Staurt and Ord (1994), the *rth*cumulant c_r is given in terms of the *rth* non-central moment μ_r' as

$$\sum_{r=1}^{\infty} c_r \frac{t^r}{r!} = \log\left(\sum_{r=0}^{\infty}\mu_r' \frac{t^r}{r!}\right) \quad (2.37)$$

Expanding the logarithm in (2.37) gives c_r in terms of μ_r' and using (2.20) the first two cumulants of GRD are given as:

$$c_1 = \frac{\Gamma\left(\alpha + \frac{r}{2}\right)}{\theta^{\frac{r}{2}}\Gamma(\alpha)},$$

$$c_2 = \frac{\Gamma(\alpha+1)}{\theta\Gamma(\alpha)} - \frac{\left(\Gamma\left(\frac{1}{2}+\alpha\right)\right)^2}{\theta\left(\Gamma(\alpha)\right)^2}. \qquad (2.38)$$

2.3.2 Survival and Hazard Functions

The survival function for the Gamma-Rayleigh distribution is given by:

$$S(v) = 1 - F(v) = 1 - \frac{\gamma(\alpha, \theta v^2)}{\Gamma(\alpha)} \qquad (2.39)$$

The hazard function of the Gamma-Rayleigh distribution is given by:

$$h(v) = \frac{g(v)}{S(v)} = \frac{2\theta^\alpha v^{2\alpha-1} e^{-\theta v^2}\Gamma(\alpha)}{\Gamma(\alpha)[\Gamma(\alpha) - \gamma(\alpha, \theta v^2)]} = \frac{2\theta^\alpha v^{2\alpha-1} e^{-\theta v^2}}{\Gamma(\alpha, \theta v^2)} \qquad (2.40)$$

where, $\Gamma(\alpha, \theta v^2) = \int_{\theta v^2}^{\infty} u^{\alpha-1} e^{-u} du$ is the upper incomplete gamma function such that $\Gamma(\alpha, \theta v^2) + \gamma(\alpha, \theta v^2) = \Gamma(\alpha)$.

Plots of the Survival and Hazard Functions of GRD

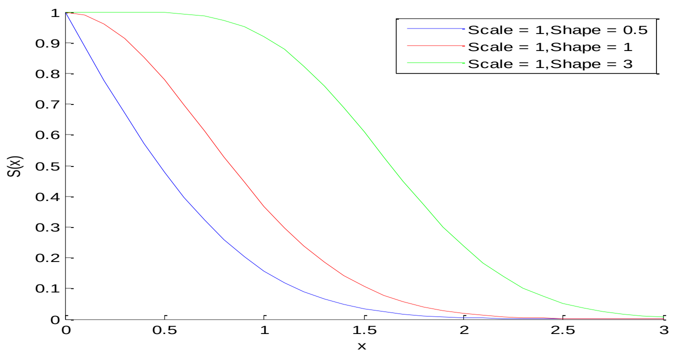

Figure 4: Plots of GRD Survival Function

The plot of the survival function is a decreasing function of *X*.

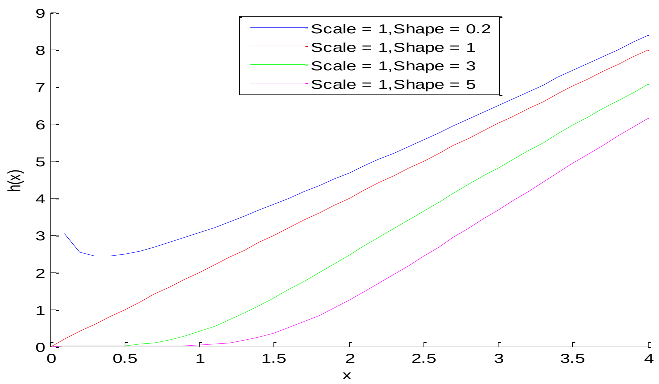

Figure 5: Plot of the Gamma-Rayleigh Hazard Functions

The plots show that GRD hazard function is an increasing function of time.

Estimation of the Parameters Of GRD

In this section, the method of maximum likelihood and method of moments estimation of the shape and scale parameters of the Gamma-Rayleigh distribution (GRD) are presented.

The Method of Maximum Likelihood (ML) Estimation

For estimating an unknown parameter θ, the likelihood principle can be used to obtain the maximum likelihood estimator (MLE) $\hat{\theta}$ (Bai and Fu, 1987). The definition of maximum likelihood estimator is presented below.

Definition 3.1: **Likelihood Function** (Mood *et al*., 1974)
The likelihood function of *n* random variables $X_1, X_2, \ldots, X_n$ is defined to be the joint density of the *n* random variables, say $f_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n;\theta)$, which is considered to be a function of θ. In particular, if $X_1, X_2, \cdots, X_n$ is random sample from the density $f(x;\theta)$, then the likelihood function is given by $L(\theta) = f(x_1;\theta) f(x_1;\theta) \cdots f(x_1;\theta)$.

Definition 3.2: Maximum Likelihood Estimator (MLE)
Let

$$L(\theta) = L(\theta; x_1, \cdots, x_n) = f(x_1,\theta) f(x_2,\theta) \cdots f(x_n,\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

be the likelihood function of the random variables $X_1, X_2, \cdots, X_n$ and $\hat{\theta} = h(X_1, \cdots, X_n)$ a function of the random variables. If the value of $\hat{\theta}$ given by $h(x_1, \cdots, x_n)$ maximizes $L(\theta)$, then $\hat{\theta} = h(X_1, \cdots, X_n)$ is a maximum likelihood estimator of θ. Under certain regularity conditions, the maximum likelihood estimator of *θ* is obtained by solving the likelihood equation:

$$\frac{dL(\theta)}{d\theta} = 0, \quad (3.1)$$

such that $\frac{\partial^2 L}{\partial \hat{\theta}^2} < 0$.

If the likelihood function contains *k* parameters, that is, if

$$L(\theta_1, \cdots, \theta_k) = \prod_{i=1}^{n} f(x_i, \theta_1, \cdots, \theta_k), \quad (3.2)$$

then, the maximum likelihood estimators are the random variables $\hat{\theta}_1 = h_1(X_1, X_2, \cdots, X_n), \hat{\theta}_2 = h_2(X_1, X_2, \cdots, X_n), \cdots, \hat{\theta}_k = h_k(X_1, X_2, \cdots, X_n)$, whose values maximize $L(\theta_1, \cdots, \theta_k)$. If the regularity conditions are satisfied, the point where the likelihood is a maximum is a solution set of the *k* equations:

$$\frac{\partial L(\theta_1, \theta_2, \cdots, \theta_k)}{\partial \theta_1} = 0$$

$$\frac{\partial L(\theta_1, \theta_2, \cdots, \theta_k)}{\partial \theta_2} = 0$$

$$\vdots$$

$$\frac{\partial L(\theta_1, \theta_2, \cdots, \theta_k)}{\partial \theta_k} = 0,$$

Such that, $\frac{\partial^2 L}{\partial \hat{\theta}_i^2} < 0; i = 1, \cdots, k$.

The logarithm of the likelihood function has the same maximum point as the likelihood function and is easier to compute.

Derivation of the ML Estimators for the Parameters of the Gamma-Rayleigh Distribution

Let there be a random sample of independent random variables from GRD each having the pdf, $g_V(v) = \frac{2\theta^\alpha}{\Gamma(\alpha)} v^{2\alpha-1} \exp(-\theta v^2)$, the likelihood function is given by:

$$L(\alpha, \theta; v_i) = \prod_{i=1}^{n} \frac{2\theta^\alpha}{\Gamma(\alpha)} v_i^{2\alpha-1} \exp(-\theta v_i^2), \quad (3.3)$$

$$= \left[\frac{2\theta^\alpha}{\Gamma(\alpha)}\right]^n \prod_{i=1}^{n} v_i^{2\alpha-1} \exp(-\theta \sum_{i=1}^{n} v_i^2),$$

Taking the log of the likelihood function gives:

$$\log L = n\log 2 + n\alpha \log\theta - n\log\Gamma(\alpha) + (2\alpha-1)\sum \log v - \theta \sum_{i=1}^{n} v_i^2,$$

$$\frac{\partial \log L}{\partial \theta} = \frac{n\alpha}{\theta} - \sum v^2,$$

$$\Rightarrow n\alpha - \theta \sum v^2 = 0, \quad (3.4)$$

$$\frac{\partial \log L}{\partial \alpha} = n\log\theta - \frac{n\Gamma'(\alpha)}{\Gamma(\alpha)} - 2\sum \log v,$$

$$\Rightarrow n\log\theta - n\psi(\alpha) - 2\sum \log v = 0, \quad (3.5)$$

(3.4) and (3.5) will be solved simultaneously to obtain the estimates of α and θ. This cannot be done analytically, therefore a numerical technique will be adopted. Consequently, properties of the maximum likelihood estimators could not be derived in this paper.

Simulation Study and Application

In this section, simulated data from the Gamma-Rayleigh distribution are analyzed. The simulation was done to study the behaviour and some properties of the new distribution. Simulations were done using probability integral transform. The probability integral transform equates the cdf of the GRD random variable V to uniform random variates, say u such that $u \sim U(0,1)$. This procedure allows the random values of V (the quantiles) to be obtained.

Also, the distribution is applied to two real life survival data on time to recover from typhoid fever and time to drop out from an insurance policy. The aim is to illustrate how the distribution can be applied to real life data.

All simulations, analysis and plots are done using programme codes written in MATLAB Application R2011b.

RESULT OF SIMULATION STUDY ON GRD

The simulation for GRD was done for different values of α and θ. The results are presented in Table 1.

Table 1: Result of Simulation Study on Gamma-Rayleigh Distribution

Model	Parameters	Mean	StdDev	Variance	SE of Mean	Skewness	Kurtosis	CV
GRD(1,0.5)	α = 1; θ =0.5	1.2462	0.6537	0.4274	0.0092	0.5999	3.1374	0.5246
GRD(1,2)	α = 1; θ = 2	0.6276	0.3311	0.1096	0.0047	0.6125	3.1334	0.5276
GRD(1,5)	α = 1; θ = 5	0.3950	0.2068	0.0428	0.0029	0.6385	3.3093	0.5235
GRD(0.5,1)	α = 0.5; θ =1	0.5661	0.4312	0.1859	0.0061	1.0557	4.1373	0.7617
GRD(2,1)	α = 2; θ = 1	1.3308	0.4755	0.2261	0.0067	0.4103	3.0981	0.3573
GRD(5,1)	α = 5; θ = 1	2.1785	0.4864	0.2366	0.0069	0.2527	3.0152	0.2233

The simulation results in Table 4.1 shows that the skewness reduces as α increases. The distribution shows kurtosis that is not so high, with smaller values of α resulting in higher kurtosis. Also, the variance decreases with increasing values of the scale parameter θ. The distribution shows a coefficient of variation (CV) of less than one for all values of the parameter and therefore, can be classified as a member of the hypo-exponential (CV < 1) class of distributions.

APPLICATIONS TO REAL LIFE DATA

In this section, the new distribution is applied to two data sets to illustrate their adequacies in fitting real life data. The criterion used to check model adequacy and performance is the Akaike Information Criterion (AIC). AIC is given by: AIC = $2k - 2\log L$; where L is the maximized value of the likelihood and k is the number of parameters in the model. A rule of thumb is that a better model should have a smaller AIC (Kwok *et al.*, 2008).

Data Description and Exploration

The two survival data used in this study are secondary data on time to recover from typhoid fever of patients and time to drop out from an insurance policy. The typhoid recovery time data were collected from patients' case notes in General Hospital, Gbagada, Lagos while the policy drop out time data were collected from an Insurance company. Table 2 gives the summary of statistics and histograms for the two sets of survival data.

DISCUSSION

The estimates of the parameters and performance criterion of AIC (figures in bold) for fitting GRD to the two survival data sets are given in Table 4.3 along with other distributions. From the results, GRD is the most adequate in fitting the Policy drop-out time data; followed by WRD, Gamma and then Weibull distribution. The typhoid recovery time data is most fitted by Gamma, closely followed by GRD, Weibull and then Rayleigh.

The results show that the newly introduced distribution is adequate in fitting the two survival data and in some cases produce better fits. It actually performs better than the Rayleigh distribution in both cases. Therefore, combining the Rayleigh distribution with the Gamma distribution is worthwhile.

Table 2: Summary of Statistics for Survival Data

Statistics	N	Mean	Variance	Std Dev.	Skewness	Kurtosis	Coef. of Var.
Recovery Time	150	4.52	5.52	2.12597	1.097	1.809	0.47035
Policy Drop Out	87	5.25	1.121	1.059	0.916	2.328	0.20171

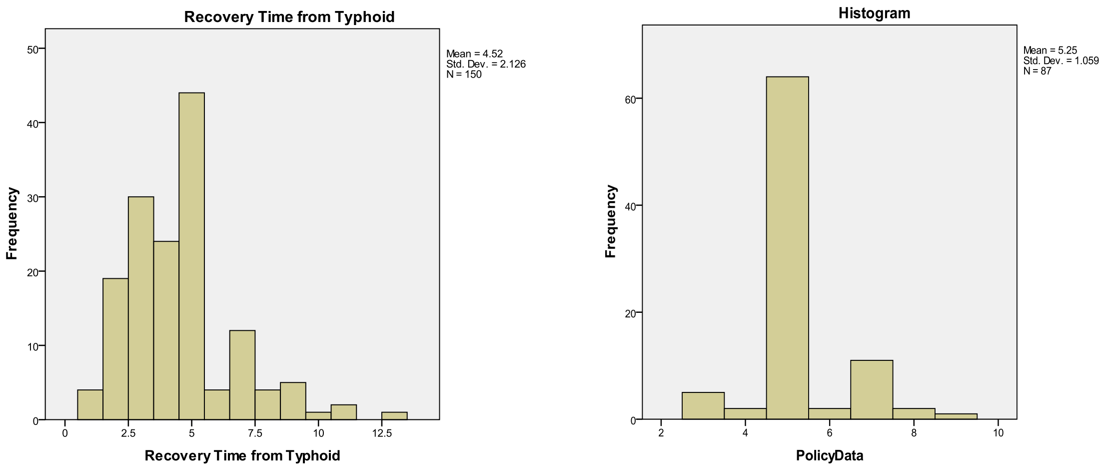

Figure 6: Histograms of Survival Data

Both the descriptive results and histograms indicate that the data is peaked and skewed to the right.

RESULTS

Table 3: Results of Fits of Distributions to Survival Data

Data	Weibull	Rayleigh	Gamma	Lognormal	GRD
Time to Recover from Typhoid	*Sh=2.2545	-	Sh=4.6846	μ=1.3980	Sh=1.3335
	*Sc=5.1129	Sc=3.5299	Sc=0.9649	σ=0.4872	Sc=0.0536
	637.7309	639.3577	628.0819	663.4589	632.3686
Time to drop out of Policy	Sh=4.8280	-	Sh=25.7668	1.6392	Sh=6.5359
	Sc=5.6885	Sc=3.7882	Sc=0.2039	0.1996	Sc=0.2277
	272.6087	354.2708	273.008	267.50	256.6209

N/B: Sh = Shape; Sc = Scale; Figures in bold are AIC

Fitted Density Plots for All Distributions

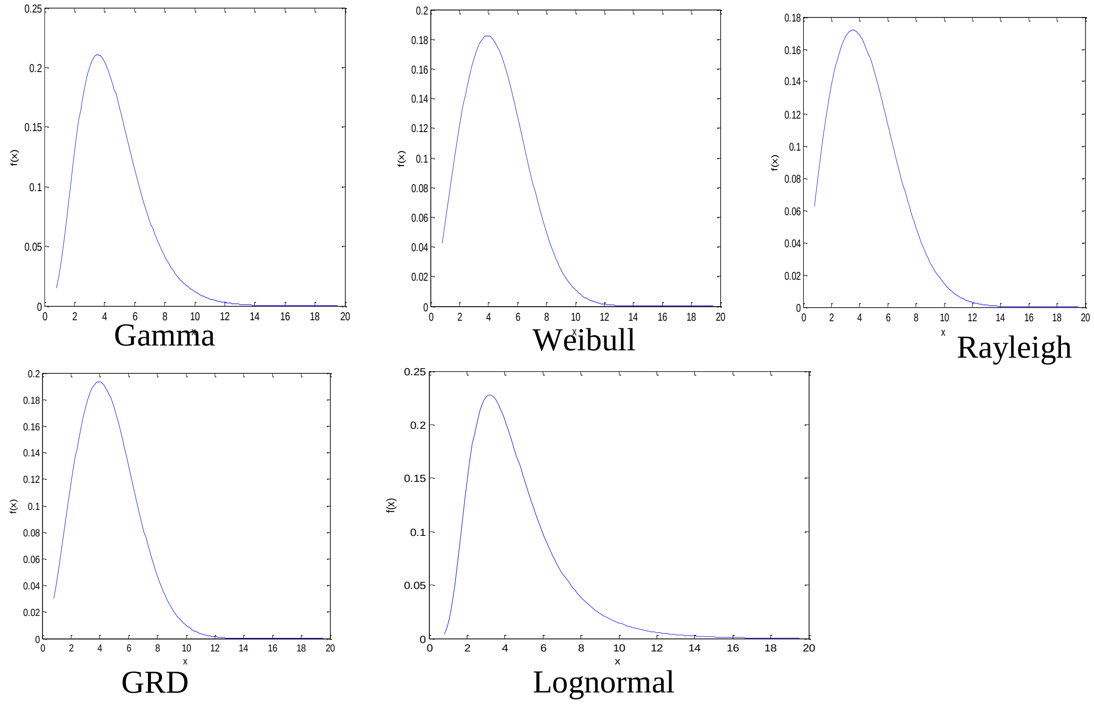

Figure 7: Fitted Density Plots for Typhoid Recovery Time Data

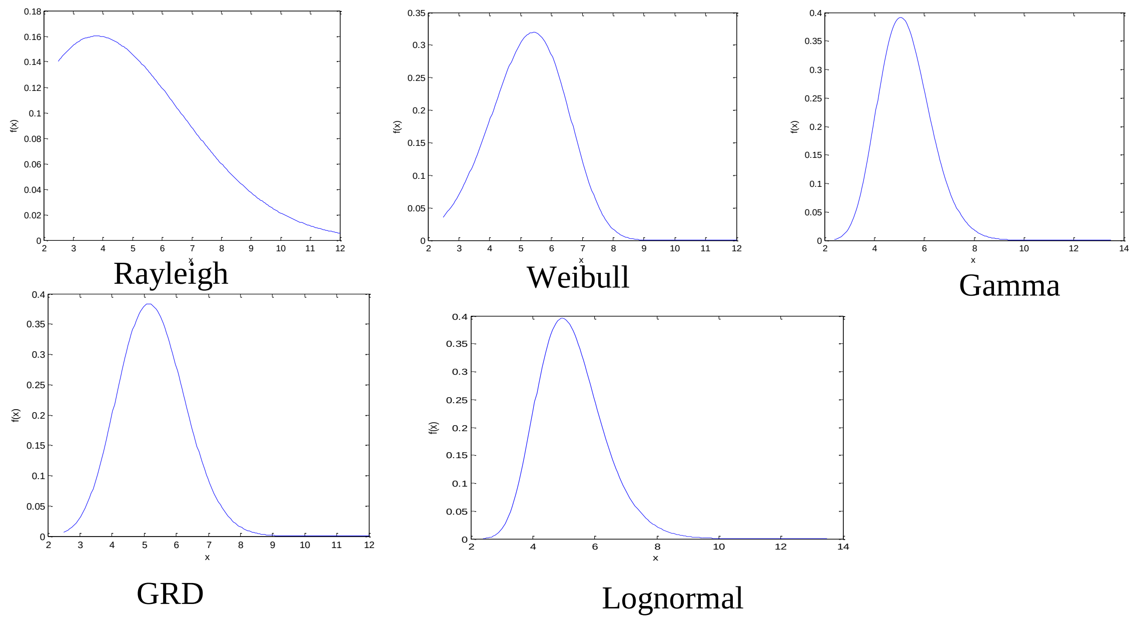

Figure 8: Fitted Density Plots for Policy Drop Out Time Data

CONCLUSION

This research work focused on generation of a new distribution involving gamma and Rayleigh distributions. The resulting two-parameter Gamma-Rayleigh Distribution (GRD) was obtained through the Transformed-Transformer method. Properties of the resulting distribution like the probability density function, cumulative distribution function, *r*th non-central moments, expectation, variances and cumulants were derived. Expressions for its survival and hazard functions were also presented. Plots of the pdf for different values of parameters and simulation studies revealed that the newly derived distribution is unimodal, peaked and skewed. Furthermore, plots of the hazard function reveal that the distributions can be used to model data with increasing hazard rates.

In order to make the distribution applicable and relevant, estimators of the parameters were derived using the method of maximum likelihood. The distribution was used to model two sets of survival data and results show that it is competitively adequate in fitting the survival data compared to the Weibull, Rayleigh, Lognormal and Gamma distributions.

REFERENCES

Adeleke, I. A, Akarawak, E. E. E, and Okafor, R. O. (2013). Investigating the Distribution of the Ratio of Independent Beta and Weibull Random Variables. *Journal of Mathematics and Technology*, **4**(1): 16-22.

Akarawak, E. E. E., Adeleke, I. A. and Okafor, R. O. (2013). The Weibull-Rayleigh Distribution and its Properties. *Journal of Engineering Research*, **18**(1): 56-67.

Akarawak, E. E. E., Adeleke, I. A. and Okafor, R. O. (2015). On the Distribution of the Ratio of Independent Gamma and Rayleigh Random Variables. *Journal of Scientific Research and Developments*, **15**(1): 54-63.

Akinsete, A., Famoye, F. and Lee, C. (2008). The beta-Pareto distribution.*Statistics*, **42**:547-563.

Alzaatreh, A., Lee, C. and Famoye, F. (2013a). A New Method for Generating Families of Continuous Distributions.Metron: *International Journal of Statistics*, **71**(1): 63-79.

Alzaatreh, A., Famoye, F. and Lee, C. (2013b).Weibull-Pareto Distribution and its Applications. *Communication in Statistics*: *Theory and Methods*, 42(9): 1673-1691.

Bai, Z.D. and Fu, J.C. (1987). On the Maximum-Likelihood Estimator for the Location Parameter of a Cauchy Distribution. *The Canadian Journal of Statistics*. **15**(2): 137-146.

Basu, A. P. and Lochner, R. H. (1972). On the Distribution of the Ratio of Two Random Variables Having the Generalized Life Distributions. *Technometrics*. *Journal Storage*, **13**(2): 281-287.

Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications.*Communication in Statistics-Theory and Methods*, **31**(4): 497 – 512.

Famoye, F., Lee, C. and Olugbenga, O. (2005).The Beta-Weibull distribution. *Journal of Statistical Theory and Applications*, **4**(2): 121-138.

Gupta, R. D. and Kundu, D. (1999). Generalized Exponential Distribution. *Australian and New Zealand Journal of Statistics*, 41: 173 – 188.

Gupta, R. D. and Kundu, D. (2001). Exponentiated Exponential Family: An Alternative to Gamma and Weibull Distributions. *Biometrical Journal*, 43(1): 319 – 326.

Kwok, O.M., Underhill, A.T., Berry, J. W., Luo, W., Ellot, T.R. and Yoon, M. (2008). Analyzing Longitudinal Data with Multilevel Models: An Example with Individuals Living with Lower Extremity Intra-articular Fractures. Reliability Psychology, 53(3): 370-386.

Lee, R. Y., Holland, B. S., Flueck, J. A. (1979). Distribution of a Ratio of Correlated Gamma Randon Variables, *SIAM Journal of Applied Mathematics*, **36**:304 – 320.

Marsaglia, G. (1965). Ratios of normal variables and ratios of sums of uniform variables.*Journal of the American Statistical Association*, **60**: 193–204.

Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). Introduction to the Theory of Statistics (3rded). McGraw Hill International Book Company. Tokyo.

Mudholkar,G.S. and Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing Bathtub Failure Data. IEEE Trans. Rel. **42**: 299-302.

Pal, M., Ali, M. M. and Woo, J. (2006). Exponentiated Weibull Distribution. *Statistica*, 66 (2): 139-148.

Press, S. J. (1969). The *t* ratio distribution. *Journal of the American Statistical Association*, **64**: 242–252.

Stacy, E.W. (1962). A generalization of the gamma distribution. *Annals of Mathematical Statistics*, **33**: 1187-1192.

Stuart, A. and Ord, J. K. (1994). Kendall's Advanced Theory of Statistics: Distribution Theory, Vol 1. London: Arnold.