

The Gamma1-Epsilon Distribution: Its Statistical Properties and Applications

Isaac Esbond Gongsin^{1,*}, Funmilayo Westnand Oshogboye Saporu²

¹Department of Mathematical Sciences, University of Maiduguri, Maiduguri, Nigeria

²National Mathematical Centre, Kwali, Abuja, Nigeria

Email address:

gongsin2013@gmail.com (I. E. Gongsin), saporuf@gmail.com (F. W. O. Saporu)

*Corresponding author

To cite this article:

Isaac Esbond Gongsin, Funmilayo Westnand Oshogboye Saporu. The Gamma1-Epsilon Distribution: Its Statistical Properties and Applications. *International Journal of Statistical Distributions and Applications*. Vol. 6, No. 4, 2020, pp. 65-70.

doi: 10.11648/j.ijstd.20200604.11

Received: September 4, 2020; **Accepted:** September 24, 2020; **Published:** October 17, 2020

Abstract: The construction of new probability distributions is an active field of research. It provides the opportunity for dynamic system modelers to choose the best model from a plethora of probability distributions that provide good fit to some data set using model selection criteria. In this study a new probability distribution function is constructed based on the gamma type-I generator, called the gamma1-epsilon distribution. Its statistical properties are described. The area under the curve of the density plots is shown through numerical integration to equal one with very minimal error margins. The density plots show shapes that are similar to many standard lifetime distributions. This implies that it is flexible to assume different shapes that can model many different random phenomena. Its hazard rate function plots also show varying shapes, namely J-shaped and bathtub-shaped, that indicate its possible use as a model for the study of survival life of biological organisms, electrical and mechanical components. The distribution is applied to the time to death of women with temporary disabilities, remission time of cancer patients and wind speed. The fits to these datasets are good with precise parameter estimates. Its compatibility with data from these dissimilar processes shows it holds a good prospect for real life application.

Keywords: Gamma-G Family, Epsilon Distribution, Hazard Rate Function, Survival Time, Remission Time, Wind Energy

1. Introduction

The dynamic nature of human life comes with several challenges requiring new models that are able to capture and fully describe the information that characterizes the system. Probability distributions play central role in providing such information and many existing distributions that were, hitherto, the best for describing random phenomena now fall short of the requirements of new systems and problem situations that are evolving. Hence, the current problems of dynamic system modelers are to come up with new distributions that are flexible enough to fit and describe new data generation processes that are unfolding. Thus, the construction of probability distributions is an active field of research with several methods in application.

One of the most applied methods of constructing probability distributions is the use of existing lifetime

distributions as generators. Example of such lifetime distributions used as generators include the beta-G due to Eugene et al [6]; the gamma-G types I, II, and III, credited to Zografos and Balakrishnan [17], Ristic and Balakrishnan [14] and Torabi and Montezari [16], respectively; Kumaraswamy-G due to Cordeiro et al [4]; exponentiated-G of Gupta et al [10]; Weibull-G due to Bourguignon et al [3]; and beta exponential G due to Alzaatreh et al. [2]. The G in each generator is an arbitrary cumulative distribution function called the parent distribution. These generators have been used to construct probability distributions that are more flexible than their respective parent distributions [7].

Johnson et al. [11, 12] provide comprehensive study of some of the new continuous and discrete probability distributions with their applications. Also, Lai [13] and Forbes et al. [7] provide a survey of some newly generated probability distributions and their applications. Recently, some new probability distributions have been constructed

from the parent epsilon distribution of Dombi et al. [5]; for example, Kumaraswamy-epsilon [8] and exponentiated-epsilon [9] distributions. These have been applied in modeling several data generation processes.

In this study, a new distribution called Type-I gamma-epsilon distribution, hereafter referred to as gamma1-epsilon distribution, is constructed. The other objective is to apply the new distribution to real life datasets.

2. The Gamma 1 - Epsilon Distribution

The Type-I gamma-G family of distributions is defined [17] by the distribution and density functions given, respectively, by

$$F(x) = \frac{1}{\Gamma(\alpha)} \int_0^v t^{\alpha-1} e^{-t} dt \quad (1)$$

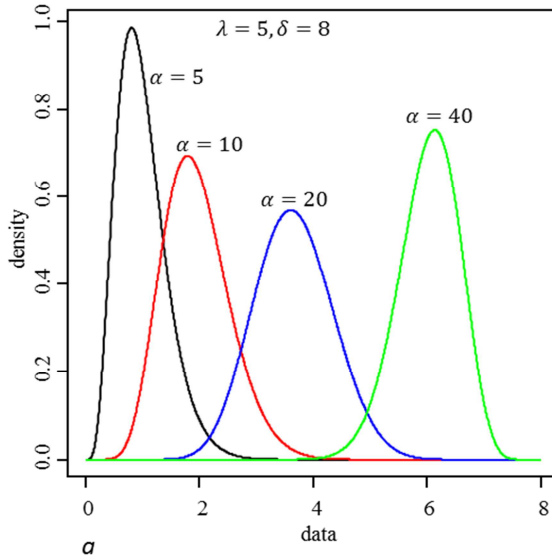
and

$$f(x) = \frac{1}{\Gamma(\alpha)} g(x) v^{\alpha-1} \quad (2)$$

where $v = -\log(1 - G(x))$ and $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$. $\alpha > 0$ determines the shape of the distribution. The range of values of the random variable, X , given in (1) and (2), for which the new distribution is defined is determined by its range of values in the parent distribution, $G(x)$.

The epsilon distribution is specified by the distribution and density functions [5] given, respectively, by

$$G(x) = 1 - \left(\frac{x+\delta}{\delta-x} \right)^{-\frac{\lambda\delta}{2}} \quad (3)$$



and

$$g(x) = \lambda \frac{\delta^2}{\delta^2 - x^2} \left(\frac{x+\delta}{\delta-x} \right)^{-\frac{\lambda\delta}{2}} \quad (4)$$

where $0 < x < \delta$, $\delta > 0$, $\lambda > 0$.

A log-transformation of the parent epsilon distribution function (3), gives

$$v = \frac{\lambda\delta}{2} \log \left(\frac{x+\delta}{\delta-x} \right) \quad (5)$$

Substituting (4) and (5) into (1) and (2) gives the Type-I gamma-epsilon distribution and density functions, given, respectively, by

$$F(x) = \frac{1}{\Gamma(\alpha)} \int_0^v t^{\alpha-1} e^{-t} dt = \frac{\gamma[\alpha, v]}{\Gamma(\alpha)} \quad (6)$$

and

$$f(x) = \frac{\lambda}{\Gamma(\alpha)} \frac{\delta^2}{\delta^2 - x^2} \left(\frac{x+\delta}{\delta-x} \right)^{-\frac{\lambda\delta}{2}} \left[\frac{\lambda\delta}{2} \log \left(\frac{x+\delta}{\delta-x} \right) \right]^{\alpha-1} \quad (7)$$

where $\gamma[\alpha, v]$ is the incomplete gamma function defined by $\gamma[\alpha, v] = \int_0^v t^{\alpha-1} e^{-t} dt$.

Plots of the gamma1-epsilon distribution at varying parameter values are presented in Figure 1 below.

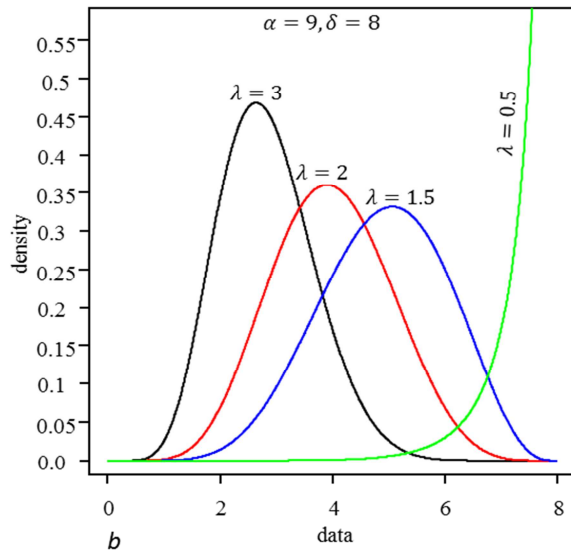


Figure 1. Density plots of gamma1-epsilon distribution at varying parameter values.

It is expedient to show that the Type-I gamma-epsilon probability density function specified in (7) is a true probability density function. That is,

$$\int_0^\delta \frac{\lambda}{\Gamma(\alpha)} \frac{\delta^2}{\delta^2 - x^2} \left(\frac{x+\delta}{\delta-x} \right)^{-\frac{\lambda\delta}{2}} \left[\frac{\lambda\delta}{2} \log \left(\frac{x+\delta}{\delta-x} \right) \right]^{\alpha-1} dx = 1$$

Direct integration of the density function is difficult. However, numerical integration can be employed to illustrate this integral at various parameter values in order to give an indication. Using the parameter values for the density plots in Figure 1, the approximate values of the integral and their absolute errors are presented in Table 1 (a and b) below.

Table 1. Results of numerical integration based on parameters of Figure 1.

1a				
Parameter values			Integral value	Absolute error
α	λ	δ		
5	5	8	1	3.2372e-09
10	5	8	1	2.0585e-05
20	5	8	1	1.6266e-07
40	5	8	1	3.3556e-09

1b				
Parameter values			Integral value	Absolute error
α	λ	δ		
9	3	8	1	4.8478e-08
9	2	8	1	2.2579e-06
9	1.5	8	1	1.0402e-05
9	0.5	8	1	1.0459e-04

3. Moments of Gamma 1 - Epsilon Distribution

For a continuous random variable, X , from a gammal-epsilon distribution, the r^{th} moment is given by

$$E(X^r) = \frac{\lambda \delta^2}{\Gamma(\alpha)} \int_0^\delta \frac{x^r}{\delta^2 - x^2} \left(\frac{x+\delta}{\delta-x} \right)^{-\frac{\lambda \delta}{2}} \left[\frac{\lambda \delta}{2} \log \left(\frac{x+\delta}{\delta-x} \right) \right]^{\alpha-1} dx \quad (8)$$

where $r = 1, 2, 3, \dots$

Although the expression for the moments of the gammal-epsilon distribution is not explicit, numerical integration can be employed to determine any moment for estimated parameter values. In this case, parameters of the gammal-epsilon distribution should be first estimated.

4. Survival Function

The survival function denoted by $S(t)$ gives the probability of surviving after t . That is

$$S(t) = \text{Prob}(T > t)$$

For the gammal-epsilon distribution, this is derived as

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

$$= 1 - \frac{\lambda}{\Gamma(\alpha)} \int_0^t \frac{\delta^2}{\delta^2 - x^2} \left(\frac{x+\delta}{\delta-x} \right)^{-\frac{\lambda \delta}{2}} \left[\frac{\lambda \delta}{2} \log \left(\frac{x+\delta}{\delta-x} \right) \right]^{\alpha-1} dx$$

$$= 1 - \frac{\gamma[\alpha, \zeta]}{\Gamma(\alpha)} \quad (9)$$

where $\gamma[\alpha, \zeta] = \int_0^\zeta s^{\alpha-1} e^{-s} ds$, $\zeta = \frac{\lambda \delta}{2} \log \left(\frac{t+\delta}{\delta-t} \right)$.

Two functions relate to the survival function; namely, hazard rate and mean residual life (MRL) functions. These are derived for the gammal-epsilon distribution below.

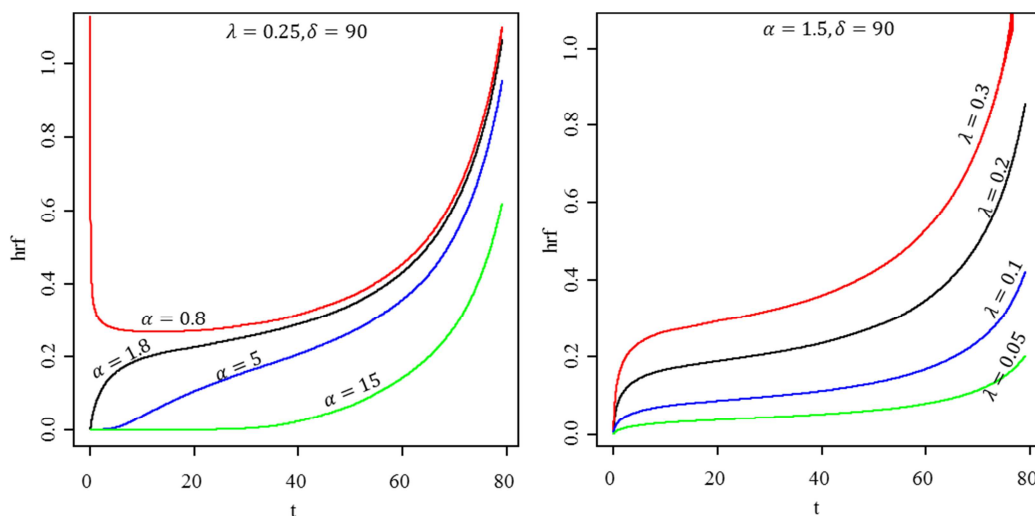
4.1. Hazard Rate Function

The hazard rate function describes the instantaneous failure rate of a component that has survived up to time t . For the survival time variable, T , of components characterized by the gamma-epsilon distribution, its hazard rate function is given by

$$hrf(t) = \frac{f(t)}{S(t)} = \frac{\lambda \frac{\delta^2}{\delta^2 - t^2} \left(\frac{t+\delta}{\delta-t} \right)^{-\frac{\lambda \delta}{2}} \zeta^{\alpha-1}}{\Gamma(\alpha) - \gamma[\alpha, \zeta]} \quad (10)$$

where $\zeta = \frac{\lambda \delta}{2} \log \left(\frac{t+\delta}{\delta-t} \right)$

Plots of the hazard rate function at various parameter values are presented in Figure 2 below. The J-shaped function depicted in the figure show that the gammal-epsilon distribution can be useful in describing systems with zero mortality rate at the beginning of their lifetimes. The bathtub-shaped functions describe systems with high mortality in their early lives, minimal mortality in useful life and high mortality as the system ages.

**Figure 2.** Plots of the gammal-epsilon hazard rate function at varying parameter values.

4.2. Mean Residual Life Function

The mean residual life gives the expected lifetime of a

component after time t . Its function is defined by the relation

$$\mu(t) = E(T - t | T > t)$$

$$= \frac{1}{S(t)} \int_t^{\infty} S(y) dy$$

For a random variable, T , specified by the gamma1-epsilon distribution, the mean residual life function is given by

$$\mu(t) = \frac{Z}{\Gamma(\alpha) - \gamma[\alpha, \zeta]} \quad (11)$$

Where $Z = (\delta - t)\Gamma(\alpha) - \int_t^{\delta} \int_0^{\eta} y^{\alpha-1} e^{-y} dy ds$, $\zeta = \frac{\lambda\delta}{2} \log\left(\frac{t+\delta}{\delta-t}\right)$, $\eta = \frac{\lambda\delta}{2} \log\left(\frac{s+\delta}{\delta-s}\right)$, $0 < t < \delta$, $\alpha > 0$, $\delta > 0$, $\lambda > 0$. From equation (11), when $t = 0$ we obtain the mean time to failure of the component, given by

$$\mu(0) = \delta - \frac{1}{\Gamma(\alpha)} \int_0^{\delta} \int_0^{\eta} y^{\alpha-1} e^{-y} dy ds \quad (12)$$

where η is defined above.

5. Applications

5.1. The Data

To illustrate its potential in modeling real life processes,

Table 2. Gamma1-epsilon and Weibull-Dagum fit results of the age at death dataset.

Distribution	Parameter	Estimate (std error)	LL (AIC)	KS (cv)	Remark	LDP
Gamma1-Epsilon	α	13.3031 (1.3915)	-1053.257 (2112.534)	0.0805 (0.0813)	Good fit	NA
	λ	0.2498 (0.0290)				
	δ	97.1203 (5.4298)				
Weibull-Dagum*	λ	1282.2665 (881.4295)	-1050.590 (2109.180)	0.0687 (0.0813)	Good fit	
	δ	3.9888 (0.1683)				
	β	2183.6861 (2072.9163)				

KS=Kolmogorov-Smirnov statistic value. cv=critical value. LL=log-likelihood value. LDP=likelihood gain per data point of Gamma1-Epsilon relative to Weibull-Dagum distribution. * obtained from Tahir et al (2014). NA=not applicable (exponents of LLs equal zero).

The plot of the gamma1-epsilon probability density fit to the age at death dataset for the estimated parameter values is presented in Figure 3.

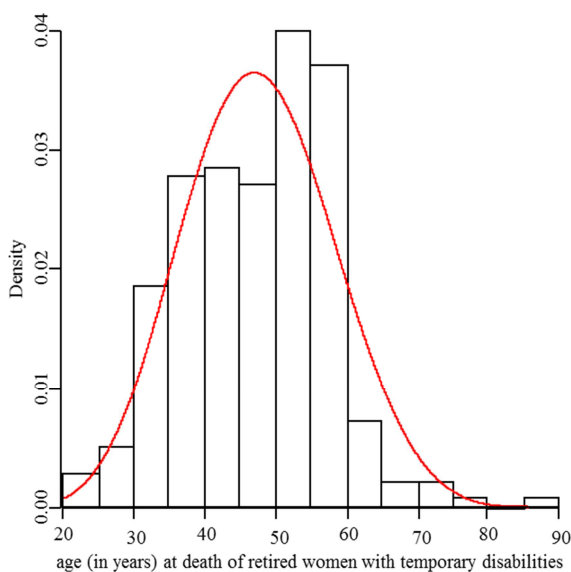


Figure 3. Gamma1-epsilon density plot of age (years) of women with temporary disabilities.

the gamma1-epsilon distribution is applied to three real life datasets. The first data is the age (in years) of 280 retired women with temporary disabilities who died in 2004; the data was obtained from Tahir et al [15]. The second dataset is the remission times (in months) of a random sample of 128 bladder cancer patients obtained from Almheidat et al [1]. The third dataset is the daily wind speed data collected at the Christmas Island Aero Station, Australia between October 1, 2018 and November 28, 2019. The wind speed data, originally recorded in km/h and converted here to m/s, were obtained at www.bom.gov.au/climate/dwo/.

5.2. Parameter Estimation

The gamma1-epsilon distribution was fitted to the three datasets using the optim package in R statistical programming language. The results are given below. Also, the results of parameter estimation by other authors for the respective datasets are shown for the purpose of comparison.

5.2.1. Mortality Data

The fit result of the gamma1-epsilon probability density function (7) to the age at death dataset is given in Table 1 below.

The hazard rate function plot for the mortality of women with temporary disabilities is given in Figure 4, below.

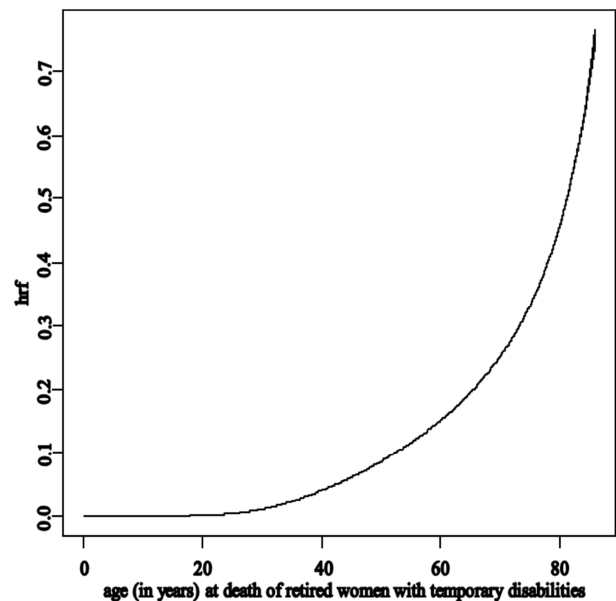


Figure 4. Gamma1-epsilon hazard rate function of the mortality data.

5.2.2. Remission Time Data

The fit results of the remission time (months) of cancer

patients are given in Table 3 below. The density plot is represented in Figure 5.

Table 3. Gamma1-epsilon and CW{L} fit results of the remission time dataset.

Distribution	Parameter	Estimate (std error)	LL (AIC)	KS (cv)	Remark	LDP
Gamma1-Epsilon	α	1.1255 (0.1280)	-415.203 (836.405)	0.0817 (0.1202)	Good fit	3.2%
	λ	0.1148 (0.0166)				
	δ	95.5076 (8.9085)				
3-Parameter CW{L}*	β	1.5623 (0.3614)	-419.233 (844.5)	0.0826 (0.1202)	Good fit	
	k	1.7888 (0.3215)				
	λ	7.3836 (0.6402)				
4-Parameter CW{L}*	α	-2.3040 (1.0937)	-416.0965 (840.2)	0.0667 (0.1202)	Good fit	0.7%
	β	2.0205 (0.4585)				
	k	3.0673 (0.7319)				
	λ	12.663 (2.6326)				

KS=Kolmogorov-Smirnov statistic value. cv=critical value. LL=log-likelihood value. LDP=percent likelihood gain per data point of Gamma1-Epsilon relative to Cauchy-Weibull {Logistic} (CW{L}) distributions. * obtained from Almheda et al (2015).

The mean remission time estimated from the fitted gamma1-epsilon distribution is 9.63 months with absolute error less than 0.00034.

5.2.3. Wind Speed Data

The fit results of the wind speed data (m/s) measured at the Christmas Island, Australia, are presented in Table 4.

Table 4. Gamma1-epsilon and Weibull fit results of the wind speed dataset.

Distribution	Parameter	Estimate (std error)	LL (AIC)	KS (cv)	Remark	LDP
Gamma1-Epsilon	α	5.9685 (0.4834)	-722.731 (1451.46)	0.0466 (0.0672)	Good fit	1.72%
	λ	0.8927 (0.0812)				
	δ	8.6557 (0.1323)				
Weibull	α	4.2295 (0.1685)	-729.710 (1463.42)	0.0685 (0.0672)	Good fit	
	β	5.8616 (0.0720)				

KS=Kolmogorov-Smirnov statistic value. cv=critical value. LL=log-likelihood value. LDP=percent likelihood gain per data point of Gamma1-Epsilon relative to Weibull distribution.

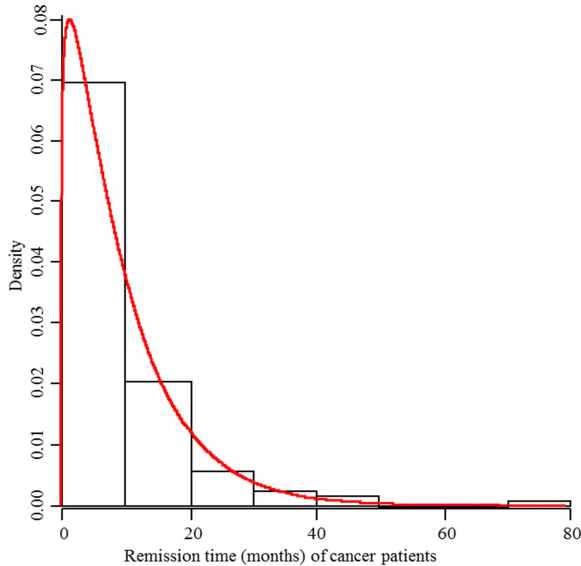


Figure 5. Gamma1-epsilon density plot of remission time (months) of cancer patients.

Some important characteristics of the wind speed at the Christmas Island can be summarized as in Table 5 below.

Table 5. Energy characteristics of Wind Speed at Christmas Island, Australia as obtained from fitting Gamma1-epsilon distribution.

Mean wind speed	Absolute error	Energy Potential
5.32 m/s	6.4e-04	92.07 W/m ²

Plots of the gamma1-epsilon and Weibull probability density functions for the wind speed dataset at the estimated parameter values are presented in Figure 6 below.

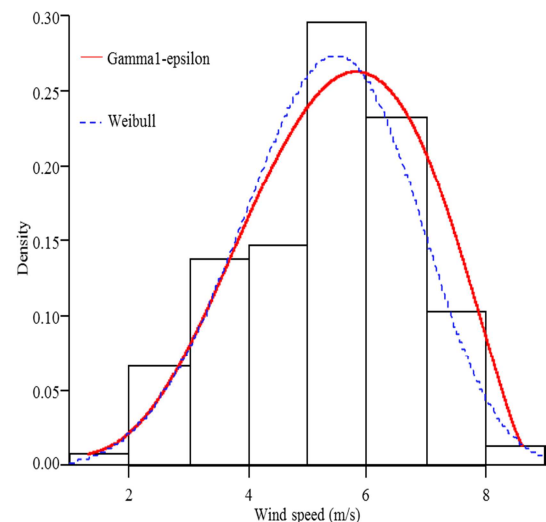


Figure 6. Gamma1-epsilon density plot of wind speed (m/s) dataset.

6. Discussion

From the plots of the density function presented in Figure 1 for different parameter values, the various shapes show that the gamma1-epsilon distribution can be employed as a model for the study of lifetime data generation processes. Density curves skewed to both right- and left-hand directions depicts

the model's flexibility in providing fit to different random phenomena. In addition, the model exhibits symmetrical properties within its parameter space.

The hazard rate function curves in Figure 2 also shows the gamma1-epsilon distribution has the potential to model systems with different lifetime behaviour. For example, the bathtub shape represents system with infant mortality (high early mortality that decreases sharply), low and constant hazard rate that characterizes the useful life of the system and an increasing hazard rate appropriate for systems experiencing wear and tear.

The results of the log-likelihood values and percent gain in likelihood per data point given in Tables 2, 3, and 4 show that the fit of the Gamma1-epsilon distribution favourably compares to those of other distributions fitted to the three datasets by other authors.

For gamma1-epsilon distribution, the results of the parameter estimate in all three cases were precise as seen in Tables 2, 3 and 4. The overall goodness-of-fit test show the model is appropriate for describing the processes that generated the datasets. These are also depicted by the density plots superimposed on the histograms in Figures 3, 5 and 6, respectively.

There are some inference that are derivable from fitting the gamma1-epsilon distribution to the three datasets. Specifically, the hazard rate function plot for the mortality data shows that the number of deaths slowly increases until age 40 and thereafter picks up rapidly. This is a more realistic trend. The mean remission time for the cancer patients to recuperate is 9.63 months after treatment. For the wind speed datasets, the average wind speed is 5.32 which is capable of generating 92.07 W/m² of electrical energy at Christmas Island, Australia.

7. Conclusion

In this study, a new probability distribution called the gamma-epsilon Type-I (referred to as gamma1-epsilon) distribution is constructed. It is a flexible distribution with the potential to model diverse lifetime data generation processes. It also exhibits properties of symmetric density curves, bathtub-shaped and J-shaped hazard rate functions. The model is compatible with three different lifetime datasets and comparable with fits of other existing distributions for the same datasets. Consequently, it is hoped that the new probability distribution will attract wide applications in engineering, reliability analysis, biological system studies and other areas of research.

References

- [1] Almheidat, M., Famoye, F., & Lee, C. (2015) Some Generalized Families of Weibull Distribution: Properties and Applications, *International Journal of Statistics and Probability*; Vol. 4, No. 3; pp 18-35, <http://dx.doi.org/10.5539/ijsp.v4n3p18>.
- [2] Alzaatreh A, Lee C, & Famoye F (2013b). "A New Method for Generating Families of Continuous Distributions." *Metron*, 71, 63–79. doi: 10.1007/s40300-013-0007-y.
- [3] Bourguignon, M., Silva, R. B. & Cordeiro, G. M. (2014) The Weibull-G family of probability distributions. *Journal of Data Science*, 12, 53-68.
- [4] Cordeiro, G. M., Ortega, E. M. & da Cunha, D. C. C. (2013) The exponentiated generalized class of distributions. *Journal of Data Science*, 11, 1-27.
- [5] Dombi, J., J'on'as, T. & T'oth, Z., E. (2018) The Epsilon Probability Distribution and its Application in Reliability Theory. *Acta Polytechnica Hungarica*, Vol. 15, No. 1, pp 197-216.
- [6] Eugene, N., Lee, C. & Famoye, F. (2002) Beta-normal distribution and its applications. *Communication in Statistics-Theory and Methods*, 31 (4), 497-512.
- [7] Forbes, C., Evans, M., Hastings, N. & Peacock, B. (2011) *Statistical Distributions*, Fourth Edition, New Jersey, John Wiley & Sons, Inc.
- [8] Gongsin, I. E. & Saporu, F. W. O. (2019a) On the Construction of Kumaraswamy-Epsilon Distribution with Applications, *International Journal of Science and Research (IJSR)*, Volume 8 Issue 11, November 2019. www.ijsr.net.
- [9] Gongsin, I. E. & Saporu, F. W. O. (2019b) The Exponentiated-Epsilon Distribution: its Properties and Applications, *International Journal of Science and Research (IJSR)*, Volume 8 Issue 12, December 2019. www.ijsr.net.
- [10] Gupta, R. C., Gupta, P. L. & Gupta, R. D. (1998) Modeling Failure Time Data by Lehmann Alternatives. *Comm. Statistical Theory and Methods*. 27, 887-904.
- [11] Johnson, N. L., Kotz, S., & Balakrishnan (1994) *Continuous Univariate Distributions*, Vol. 1, Second Edition, New York, John Wiley & Sons, Inc.
- [12] Johnson, N. L., Kotz, S., & Balakrishnan (1995) *Continuous Univariate Distributions*, Vol. 2, Second Edition, New York, John Wiley & Sons, Inc.
- [13] Lai, C. D. (2014). Generalized Weibull Distributions. *Springer Briefs in Statistics*, DOI: 10.1007/978-3-642-39106-4_2.
- [14] Ristic, M. M. & Balakrishnan, N. (2012) The Gamma Exponentiated Exponential Distribution, *Journal of Statistical Computation and Simulation*, 82, 1191–1206. doi: 10.1080/00949655.2011.574633.
- [15] Tahir, M. H., Cordeiro, G. M., Mansoor, M., Zubair, M. and Alizadeh, M. (2014) The Weibull-Dagum Distribution: Properties and Applications, *Communication in Statistics-Theory and Methods*, pp 1-28, DOI: 10.1080/03610926.2014.983610.
- [16] Torabi, H. & Montazeri, N. H. (2012) The Gamma-Uniform Distribution and Its Applications, *Kybernetika*, 48, 16–30.
- [17] Zografos, K. & Balakrishnan, N. (2009) On families of beta and generalized gamma generated distributions and associated inference. *Statistical Methodology*, 6, 344-362.