Journal of Statistics Applications &Probability Letters
An International Journal

http://dx.doi.org/10.18576/jsapl/040302

Characterization and Information Measures of Weighted Erlang Distribution

Sofi Mudasir and S. P. Ahmad

Department of Statistics, University of Kashmir, Srinagar, India.

Received: 11 Jan. 2017, Revised: 27 Apr. 2017, Accepted: 2 May 2017
Published online: 1 Sep. 2017

Abstract: In this paper, a new class of weighted Erlang distribution is introduced. The characterizing properties of the model are discussed and derived. Also information measures of a new model are derived and studied.

Keywords: Weighted Erlang distribution, moments and moment generating function, Hazard rate function, Reliability function and Shannon's entropy.

1 Introduction

The Erlang distribution is continuous probability distribution that has wide range of applicability, primarily due to its relation to the exponential and Gamma distributions. The Erlang distribution was developed by A. K. Erlang (1909) to scrutinize the number of telephone calls that could be made at the same time to switching station operators. This effort on telephone traffic engineering has been expanded to consider waiting times in queuing systems in general. Queuing theory originated when Erlang (1909) published his fundamental paper which is related to the study of telephone traffic obstruction. If the durations of individual calls are exponentially distributed, then the duration of successive calls is the Erlang distribution. The probability function of the Erlang distribution is

$$f(x;\lambda,\beta)=\frac{1}{(\lambda-1)!\,\beta^{\lambda}}x^{\lambda-1}\exp\left(-\frac{x}{\beta}\right),x>0,\beta>0,\lambda=1,2,3,\ldots$$

Where λ and β are the shape and the scale parameters, respectively, such that λ is a positive integer number.
Mean and variance of Erlang distribution is given as

$$\text{Mean} = \lambda\beta \quad \text{and} \quad \text{Variance} = \lambda\beta^2$$

The Concept of weighted distributions can be traced to the work of Fisher (1934), in relation with his studies on how methods of ascertainment can persuade the form of distribution of recorded observations. Later it was introduced and formulated in general form by Rao (1965), in connection with modeling statistical data where the usual practice of using standard distributions for the purpose was not found to be apposite. Rao in his paper identified different situations that can be modeled by weighted distributions. These situations refer to instances where the recorded observations cannot be considered as a random sample from the original distributions. This may occur due to non-observability of some events or damage caused to the original observation resulting in a reduced value, or adoption of a sampling procedure which gives unequal chances tothe units in the original. The usefulness and applications of weighted distributions to biased samples in various areas including medicine, ecology, reliability, and branching processes can be seen in Patil and Rao (1978), Gupta and Kirmani (1990), Gupta and Keating(1985). For more important results of weighted distribution you can see also (Oluyede and George (2000), Ghitany and Al-Mutairi (2008), Ahmed, Reshi and Mir (2013), Broderick X. S., Oluyede and Pararai (2012), Oluyede and Terbeche M (2007), Sofi Mudasir and S.P Ahmad (2015), Ghitany et al. (2016), Zahida et al. (2016).

Suppose X is a non-negative random variable with its probability density function $f(x)$, then the

[*]Correspondingauthore-mail: sofimudasir3806@gmail.com

probability density function of the weighted random variable X_w is given by

$$f_w(x) = \frac{w(x) f(x)}{\mu_w}, x > 0, \quad (1)$$

$w(x)$ is a non-negative weight function and $\mu_w = \int w(x) f(x;\gamma) dx$ for continuous series and $\mu_w = \sum w(x) f(x;\gamma)$ for discrete series. The probability function of the Erlang distribution is

$$f(x;\lambda,\beta) = \frac{1}{(\lambda-1)!\beta^{\lambda}} x^{\lambda-1} \exp\left(-\frac{x}{\beta}\right), x > 0, \lambda = 1,2,3,\ldots, \beta > 0 \quad (2)$$

Let $w(x) = x^{\theta}, \theta \in R$ (3)

Also, $\mu_w = \int_0^{\infty} w(x) f(x;\lambda,\beta) dx$

$$= \int_0^{\infty} \frac{1}{(\lambda-1)!\beta^{\lambda}} x^{\lambda+\theta-1} \exp\left(-\frac{x}{\beta}\right) dx$$

$$= \frac{\beta^{\theta}\Gamma(\theta+\lambda)}{(\lambda-1)!} \quad (4)$$

Substitute the value of equations (2), (3), and (4) in equation (1), we get

$$f_w(x;\lambda,\beta,\theta) = \frac{1}{\Gamma(\lambda+\theta)\beta^{\lambda+\theta}} x^{\lambda+\theta-1} \exp\left(-\frac{x}{\beta}\right) \quad (5)$$

This is the required probability density function of weighted Erlang distribution, Where λ and β are the shape and scale parameters of the distribution.The cumulative distribution function corresponding to (5) is given as

$$F_X(x) = \int_0^{x} f_w(x;\lambda,\beta,\theta) dx$$

$$= \frac{1}{\Gamma(\lambda+\theta)} \gamma\left(\lambda+\theta, \frac{x}{\beta}\right) \quad (6)$$

2 Some particular cases of weighted Erlang distribution

i) If θ=0 in (5), we get Erlang distribution

$$f(x;\lambda,\beta) = \frac{1}{\Gamma(\lambda)\beta^{\lambda}} x^{\lambda-1} \exp\left(-\frac{x}{\beta}\right)$$

ii) If θ=1 in (5), we get length biased Erlang distribution

$$f_l(x;\lambda,\beta) = \frac{1}{\Gamma(\lambda+1)\beta^{\lambda+1}} x^{\lambda} \exp\left(-\frac{x}{\beta}\right)$$

iii) If θ=2 in (5), we get area biased Erlang distribution

$$f_a(x;\lambda,\beta)=\frac{1}{\Gamma(\lambda+2)\beta^{\lambda+2}}x^{\lambda+1}\exp\left(-\frac{x}{\beta}\right)$$

iv) If θ=0, λ=1 in (5), we get exponential distribution

$$f(x)=\frac{1}{\beta}\exp\left(-\frac{x}{\beta}\right)$$

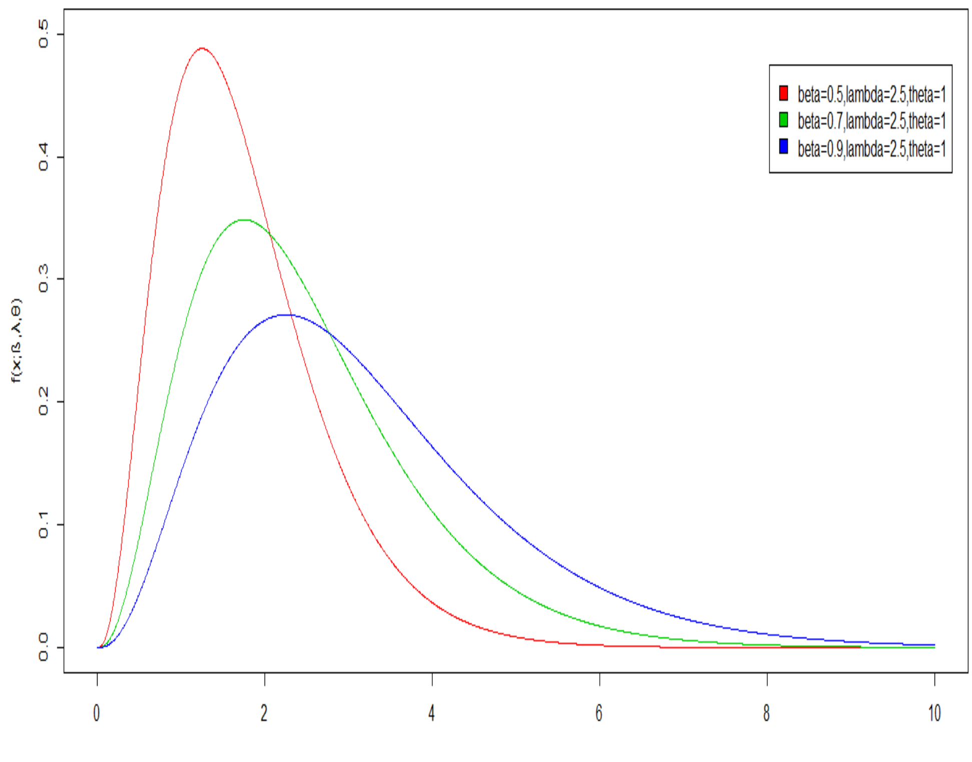

Figure 1. The pdf of weighted Erlang under different values of scale parameter β

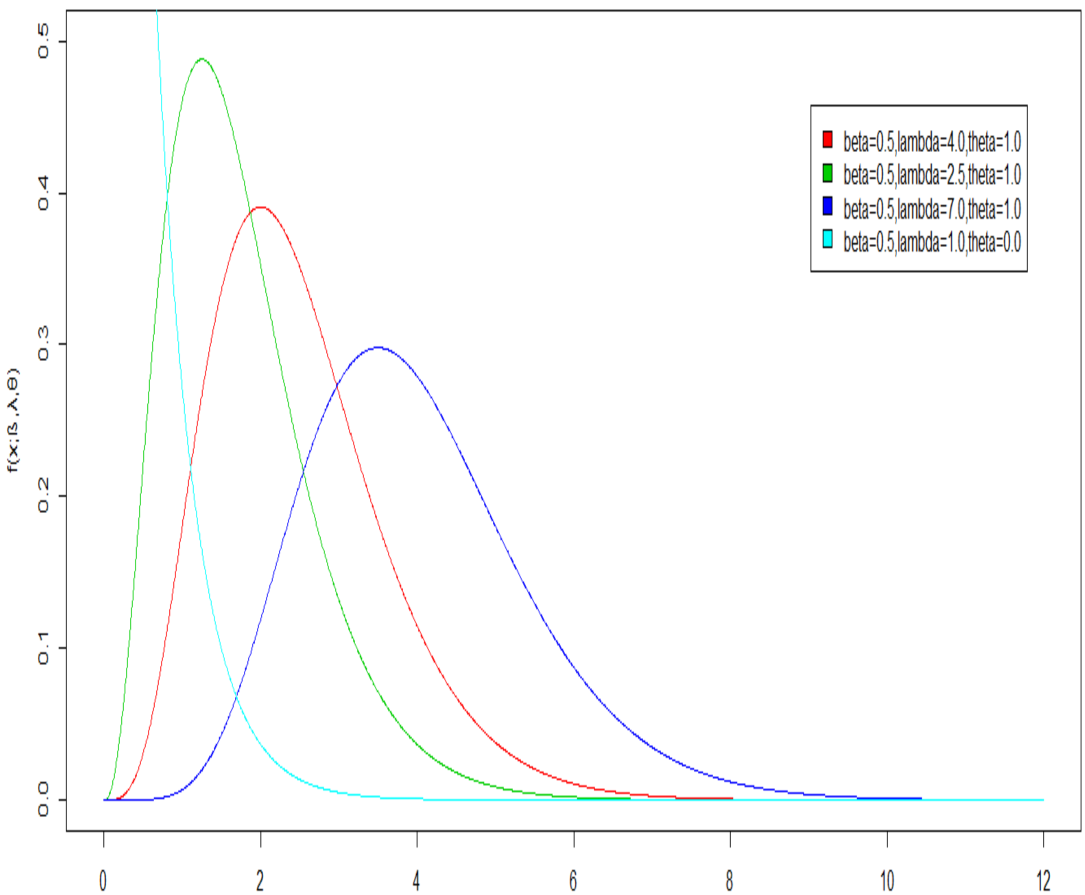

Figure 2. The pdf of weighted Erlang distribution under different values of shape Parameter λ

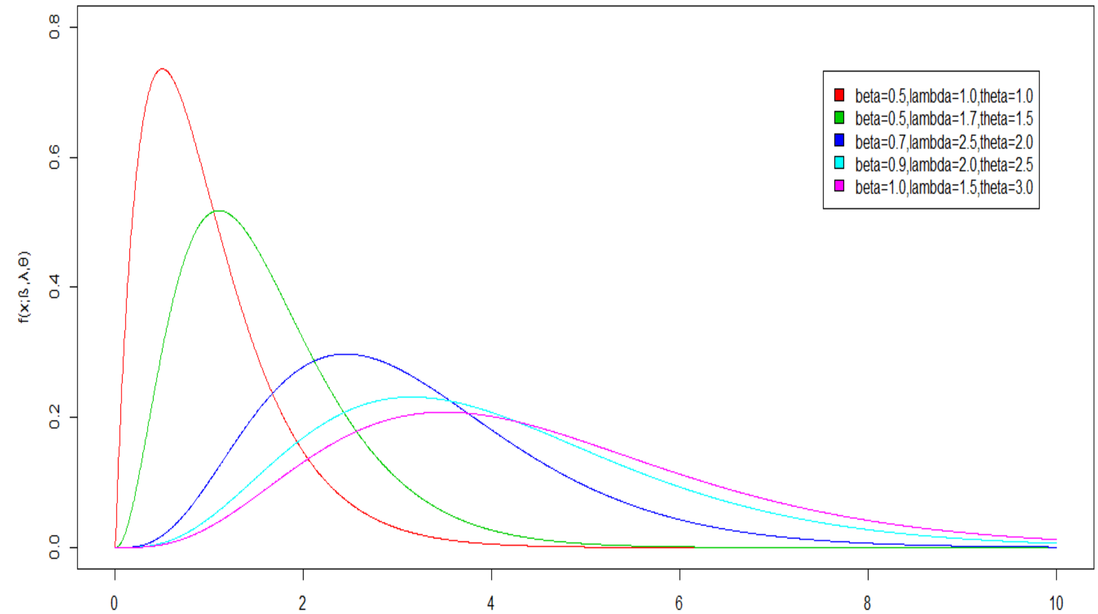

Figure 3. The pdf of weighted Erlang distribution under different values of Parameters

Figure 1 shows the pdf of weighted Erlang distribution under different values of scale parameter. From the figure it is clear that by changing the value of scale parameter, there is only change in the origin (scale) of the distribution and has no effect on the shape of the distribution. From figure 2, it is clear that there is effect on the shape of the distribution on changing the value of shape parameter. Also it is clear from the figure 3 that there is effect on the shape and origin of the distribution on changing the values of the parameters.

3 Reliability Analysis

The reliability function of weighted Erlang distribution is denoted by R(t) and is defined as

$$R(t) = 1 - F(t)$$
$$= 1 - \frac{1}{\Gamma(\lambda+\theta)} \gamma\left(\lambda+\theta, \frac{t}{\beta}\right)$$

One of the characteristic in reliability analysis is the hazard rate function denoted by h(t) and is defined as

$$h(t) = \frac{f(t)}{1 - F(t)}$$

$$= \frac{t^{\theta+\lambda-1} \exp\left\{-\left(\frac{t}{\beta}\right)\right\}}{\beta^{\lambda+\theta}\left[\Gamma(\lambda+\theta) - \gamma\left(\lambda+\theta, \frac{t}{\beta}\right)\right]}$$

4 Moments and Moment Generating Function

In this section, the r^{th} non-central moment and moment generating function have been derived.

Theorem 4.1: If a random variable X has weighted Erlang distribution, then the r^{th} non-central moment is given by *the* following

$$\mu_r' = \beta^r \frac{\Gamma(\lambda+\theta+r)}{\Gamma(\lambda+\theta)}$$

Proof: The r^{th} moment of weighted Erlang distribution about origin is obtained as

$$\mu_r' = \int_0^\infty x^r f_w(x;\lambda,\beta,\theta)dx$$

$$= \int_0^\infty x^r \frac{1}{\Gamma(\lambda+\theta)\beta^{\lambda+\theta}} x^{\lambda+\theta-1} \exp\left(-\frac{x}{\beta}\right)dx$$

$$= \frac{1}{\Gamma(\lambda+\theta)\beta^{\lambda+\theta}} \int_0^\infty x^{\theta+\lambda+r-1} \exp\left(-\frac{x}{\beta}\right)dx$$

On solving the above integral, we

$$\mu_r' = \beta^r \frac{\Gamma(\lambda+\theta+r)}{\Gamma(\lambda+\theta)} \tag{7}$$

This completes the proof.

If r=1 in equation (7), we get

$$\mu_1' = \beta(\lambda+\theta)$$

This is the mean of the weighted Erlang distribution.

If r=2 in equation (7), we get

$$\mu_2' = \beta^2(\lambda+\theta+1)(\lambda+\theta)$$

Therefore, the variance is given by

$$\mu_2 = \mu_2' - \mu_1'^2$$

$$= \beta^2(\lambda+\theta)$$

If r=3, 4 in equation (7), we get

$$\mu_3' = \beta^3(\lambda+\theta+2)(\lambda+\theta+1)(\lambda+\theta)$$

And

$$\mu_3 = 2\beta^3(\theta+\lambda)$$

$$\mu_4' = \beta^4(\lambda+\theta+3)(\lambda+\theta+2)(\lambda+\theta+1)(\lambda+\theta)$$

And

$$\mu_4 = \beta^4(3\lambda+3\theta+6)(\lambda+\theta)$$

Theorem 4.2: let X have a weighted Erlang distribution with pdf given in (5), then the moment generating function (MGF) is given by

$$M_X(t) = \sum_{r=0}^{\infty}\frac{(\beta t)^r}{r!}\frac{\Gamma(\lambda+\theta+r)}{\Gamma(\lambda+\theta)}$$

Proof: By definition

$$M_X(t) = E(e^{tx})$$

$$= \int_0^{\infty} e^{tx} f(x)dx$$

Using Taylor series

$$M_X(t) = \int_0^{\infty}(1+tx+\frac{(tx)^2}{2!}+\frac{(tx)^3}{3!}+\ldots)f(x)dx$$

$$= \int_0^{\infty}\sum_{r=0}^{\infty}\frac{(tx)^r}{r!}f(x)dx$$

$$= \sum_{r=0}^{\infty}\frac{t^r}{r!}\int_0^{\infty}x^r f(x)dx$$

$$= \sum_{r=0}^{\infty}\frac{t^r}{r!}\mu_r'$$

Now by using the equation (7) in the above equation, we get

$$M_X(t)=\sum_{r=0}^{\infty}\frac{(\beta t)^r}{r!}\frac{\Gamma(\lambda+\theta+r)}{\Gamma(\lambda+\theta)}$$

This is the required MGF of weighted Erlang distribution

.

5 Standard Deviation and Coefficient of Variation

Standard deviation is defined as the positive square root of variance. Symbolically, $\sigma=\sqrt{\sigma^2}$. We have

$$\sigma^2=\beta^2(\lambda+\theta)$$

$$\Rightarrow \quad \sigma=\beta\sqrt{(\lambda+\theta)}$$

Also, the coefficient of variation is defined as the ratio of standard deviation to mean. Usually, it is denoted by C.V and is given by

$$C.V=\frac{\sigma}{\mu}$$

$$\Rightarrow \quad C.V=\frac{\beta\sqrt{(\lambda+\theta)}}{\beta(\lambda+\theta)}$$

$$\Rightarrow \quad C.V=\frac{1}{\sqrt{(\lambda+\theta)}}$$

6 Coefficient of Skewness and Kurtosis

Denote it by C.S, the coefficient of skewness enables us to know if the distribution under study is symmetric or not, it is defined by

$$C.S=\frac{{\mu_3}^2}{{\mu_2}^3} \quad (8)$$

The skewness is zero for symmetric distribution, positive for skewed right distribution, and negative if the distribution is skewed towards left.

We have

$$\mu_2=\beta^2(\theta+\lambda) \quad (9)$$

$$\mu_3=2\beta^3(\theta+\lambda) \quad (10)$$

Substitute the value of equations (9) and (10) in equation (8), we get

$$C.S = \frac{4}{\theta + \lambda}$$

Denoted by C.K, the coefficient of kurtosis measures the flatness of the top of the curve and is defined by

$$C.K = \frac{\mu_4}{\mu_2^{\;2}} - 3 \qquad (11)$$

The kurtosis is equal to zero for the normal distribution, positive for the more tall and slim curves than the normal one in the neighborhood of the mode, in this case the distribution is said to be leptokurtic. It is negative for platykurtic distributions (flatter than the normal distribution).

Wehave

$$\mu_4 = \beta^4 (3\lambda + 3\theta + 6)(\lambda + \theta) \qquad (12)$$

Substitute the value of equations (9) and (12) in equation (11), we get

$$C.K = \frac{6}{\theta + \lambda}$$

7 Harmonic mean of weighted Erlang distribution

The harmonic mean denoted by H is given as

$$\frac{1}{H} = \int_0^\infty \frac{1}{x} f_w(x;\lambda,\beta,\theta)dx$$

$$= \frac{1}{\Gamma(\lambda+\theta)\beta^{\lambda+\theta}} \int_0^\infty x^{\lambda+\theta-2} \exp\left(-\frac{x}{\beta}\right)dx$$

By setting $u = \frac{x}{\beta}$, we get

$$\frac{1}{H} = \frac{1}{\beta\,\Gamma(\lambda+\theta)} \int_0^\infty u^{\lambda+\theta-2} \exp(-u)du$$

$$\Rightarrow \qquad \frac{1}{H} = \frac{\Gamma(\lambda+\theta-1)}{\beta\,\Gamma(\lambda+\theta)}$$

$$\Rightarrow \qquad H = \beta(\lambda+\theta-1)$$

8 Shannon's Entropy of the Weighted Erlang Distribution

The concept of Shannon's entropy is the central role of information theory, sometimes referred as measure of uncertainty. The entropy of a random variable is defined in terms of its probability distribution and can be shown to be a good measure of randomness or uncertainty. For deriving the Shannon's entropy of the weighted Erlang distribution, we need the following two definitions that more details of this can be found in Thomas J.A. et.al.(1991).

Definition 1: The entropy of a discrete alphabet random variable f defined on the probability space (Ω, β, P) is defined by

$$H_p(f) = -\sum_{a \in A} p(f = a) \log\left(p(f = a)\right)$$

It is obvious that $H_p(f) \geq 0$.

Definition 2: The obvious generalizations of the definition of entropy for a probability density function f defined on the real line as:

$$H(f) = -\int_{-\infty}^{\infty} f(x) \log f(x) dx = E\left(-\log\left(f(x)\right)\right)$$

Provided the integral exists.

Theorem 8.1: let X have a weighted Erlang distribution with the following probability density function

$$f_w(x; \lambda, \beta, \theta) = \frac{1}{\Gamma(\lambda + \theta)\beta^{\lambda+\theta}} x^{\lambda+\theta-1} \exp\left(-\frac{x}{\beta}\right)$$

Then Shannon's entropy of weighted Erlang distribution is

$$H\left[f_w(x; \lambda, \beta, \theta)\right] = (1 - \theta - \lambda)\psi(\theta + \lambda) + \log\left(\beta\Gamma(\lambda + \theta)\right) + (\theta + \lambda)$$

Where $\psi(\theta + \lambda) = \dfrac{\Gamma'(\theta + \lambda)}{\Gamma(\theta + \lambda)}$

Proof: Shannon's entropy is defined as

$$H\left[f_w(x; \lambda, \beta, \theta)\right] = E\left[-\log\left(f_w(x; \lambda, \beta, \theta)\right)\right]$$

$$= E\left[-\log\left(\frac{1}{\Gamma(\lambda + \theta)\beta^{\lambda+\theta}} x^{\lambda+\theta-1} \exp\left(-\frac{x}{\beta}\right)\right)\right]$$

$$=(\lambda+\theta)\log\beta+\log\Gamma(\lambda+\theta)-(\lambda+\theta-1)E(\log x)+\frac{1}{\beta}E(x) \tag{13}$$

Now, $E(\log x)=\int_0^\infty \log x f_w(x;\lambda,\beta,\theta)dx$

$$=\int_0^\infty \log x\frac{1}{\Gamma(\lambda+\theta)\beta^{\lambda+\theta}}x^{\lambda+\theta-1}\exp\left(-\frac{x}{\beta}\right)dx$$

$$=\frac{1}{\Gamma(\lambda+\theta)\beta^{\lambda+\theta}}\int_0^\infty \log x\frac{1}{\Gamma(\lambda+\theta)\beta^{\lambda+\theta}}x^{\lambda+\theta-1}\exp\left(-\frac{x}{\beta}\right)dx$$

$$=\psi(\theta+\lambda)+\log\beta \tag{14}$$

Also, $E(x)=\beta(\lambda+\theta)$ (15)

By putting the value of equations (14) and (15) in equation (13), we get

$$H\left[f_w(x;\lambda,\beta,\theta)\right]=(1-\theta-\lambda)\psi(\theta+\lambda)+\log(\beta\Gamma(\lambda+\theta))+(\theta+\lambda)$$

This completes the proof.

9 Entropy Estimation of Weighted Erlang Distribution

In order to introducing an approach for model selection, we remember Akaike and Bayesian information criterion based on entropy estimation.

The Akaike information criterion (AIC) was developed by Hirotugu Akaike, originally under the name “an information criterion”. It was first announced by Akaike at a (1971)symposium, the proceedings of which werepublished in (1973). The AIC is a measure of relative quality of a statistical model for a given set of data. That is, given a collection of models for the data, AIC estimates the quality of each model relative to each of the other models. Hence AIC provides a means for model selection.

Suppose that we have a statistical model of some data. Let L be the maximized value of the likelihood function for the model. Let K be the number of estimated parameters in the model. Then the AIC value of the model is the following

$$AIC=2K-2Log(L) \tag{16}$$

Given a set of candidate models for the data, the preferred model is the one which has the minimum AIC value.
The Bayesian information criterion (BIC) or Schwarz criterion is a criterion for model selection among a finite set of models; the model with the lowest BIC value is preferred. It is based, in part on the likelihood function and it is closely related to the AIC. The BIC was developed by Gideon E. Schwarz and published in a (1978) paper, where he gave a Bayesian argument for adopting it. The BIC is formally defined as

$$BIC = KLog(n) - 2Log(L) \tag{17}$$

Where, n is the number of observations or equivalently the sample size.

Now, likelihood function of equation (5) is given as

$$L(x_1, x_2, \ldots, x_n) = \left(\frac{1}{\beta^{\lambda+\theta}\Gamma(\lambda+\theta)}\right)^n \prod_{i=1}^{n} x_i^{\lambda+\theta-1} \exp\left\{-\left(\frac{\sum_{i=1}^{n} x_i}{\beta}\right)\right\}$$

By taking log on both sides, we get

$$\log L = -n(\lambda+\theta)\log\beta - n\log\Gamma(\lambda+\theta) + (\lambda+\theta-1)\sum_{i=1}^{n}\log x_i - \frac{\sum_{i=1}^{n} x_i}{\beta}$$

$$\Rightarrow \quad -\frac{\log L}{n} = (\lambda+\theta)\log\beta + \log\Gamma(\lambda+\theta) - \frac{(\lambda+\theta-1)}{n}\sum_{i=1}^{n}\log x_i + \frac{\sum_{i=1}^{n} x_i}{n\beta}$$

$$\Rightarrow \quad -\frac{\log L}{n} = (\lambda+\theta)\log\beta + \log\Gamma(\lambda+\theta) - (\lambda+\theta-1)E(\log x) + \frac{1}{\beta}E(x) \tag{18}$$

On comparing equations (13) and (18), we get

$$\log L = -nH\left[f_w(x;\lambda,\beta,\theta)\right]$$

Thus from equations (16) and (17), we have

$$AIC = 2K + 2nH\left[f(x)\right]$$

$$BIC = K\log n + 2nH\left[f(x)\right]$$

10 Numerical Example

This section illustrates the usefulness of weighted Erlang distribution to real data to see how our model works in practice. For analysis we take the sample of 20 observations from Mendenhall andHader (1958) mixture datarecorded about times to failure for ARC-1 VHF communication transmitter receivers of a single commercial airline. Following is the set of 20 observations:

152, 528, 424, 208, 536, 40, 8, 224, 112, 72, 72, 72, 112, 360, 120, 552, 104, 384, 464, 552

Table 1. AIC, BIC and AICC Criteria of weighted Erlang and other Related Distributions.

Model	estimates			-2logl	AIC	BIC
	lambda	beta	theta			
Weighted Erlang distribution	0.15	74.26	2.67	271.02	277.02	274.92
Erlang distribution	0.40	58.37	3.00	279.65	283.65	282.25
Exponential distribution	2.33	64.86	0.70	275.72	277.72	277.02

11 Conclusions

We consider the entropy estimation of weighted Erlang distribution and based on this calculate theAIC and BIC using the real data set. From the above tables, it has been observed that the weighted Erlang distribution have the smallest AIC and BIC values as compared to Erlang and exponential distribution. Hence we can conclude that the weighted Erlang distribution gives better results and estimates as compared to exponential and Erlang distributions for the data set which we used in this paper.

Acknowledgement

We are grateful to the anonymous reviewers for their valuable comments and suggestions in improving the manuscript.

References

[1] A. K. Erlang, The theory of probabilities and telephone conversations. NytTidsskriftfor Matematik, B, 20(6), 87-98 (1909).

[2] B. O. Oluyede, and E. O. George, On stochastic inequalities and comparisons ofreliabilitymeasuresfor weighted distributions. Mathematical Problems in Engineering, 8, 1-13 (2000)

[3] B. O. Oluyede, and M. Terbeche, On energy and expected uncertainty measures in weighteddistribution .International Mathematical Forum, 2(20), 947-956 (2007).

[4] C.R. Rao, On discrete distributions arising out of methods of ascertainment in classical and contagious discrete distributions. Pergamonpress and statistical publishing society, Calcutta, 320-332 (1965).

[5] G. P. Patil, & C. R. Rao, Weighted distributions and size-biased sampling with applications to wildlife populations and human families. Biometrics, 34, 179–184 (1978).

[6] K.A. Mir, A. Ahmad, and J.A. Reshi, Structural properties of length biased Beta distribution of first kind. American Journal of Engineering Research, vol-02, 01-06 (2013).

[7] M. E. Ghitany, and D. K. Al-Mutairi, Size-biased Poisson-Lindleydistribution and its application.METRON-International Journal of Statistics, LXVI(3), 299-311 (2008).

[8] M. Evans, N. Hastings, and B. Peacock, Statistical Distributions. 3rd Edition, Wiley, New York (2000).

[9] M.E. Ghitany, S.M. Aboukhamseen, and E.A.S. Mohammad, Weighted half exponential powerdistributionand associated inference, 10, 91-108 (2016).

[10] R. C. Gupta, and J. P. Keating, Relations for reliability measures under length biasedsampling. Scan. J. Statist., 13, 49-56 (1985)
[11] R. C. Gupta, and S. N. U. A. Kirmani, The role of weighted distributions in stochastic modeling.Communication Statistics, 19(9), 3147-3162 (1990).
[12] R.A. Fisher, The effects of methods of ascertainment upon the estimation of frequencies. The Annals of Eugenics, 6, 13–25 (1934).
[13] Sofi Mudasir and S.P. Ahmad, Structural properties of length biased Nakagami distribution. International Journal Modern Mathematical Sciences, 13(3), 217-227 (2015).
[14] T. M. Cover, J.A. Thomas, Elements of information theory. New York, Wiley (1991).
[15] X.S. Broderick, B.O. Oluyede, and M. Pararai, Theoretical properties of weighted generalized Rayleigh and related distributions. Theoretical Mathematics & Applications, 2(2), 45-62 (2012).
[16] Z. Parveen, Z. Ahmad, and M. Ahmad, on size-biased double weighted exponential distribution. open journal of statistics, 6, 917-930 (2016).