The Beta-Weighted Weibull Distribution: Some Properties and Application to Bladder Cancer Data

Badmus N Idowu1,*, Ikegwu and Emmanuel M2

1Department of Statistics, Abraham Adesanya Polytechnic, Ijebu-Igbo, Nigeria
2Department of Statistics, Yaba College of Technology, Yaba, Lagos

Abstract

Combination of weighted weibull distribution proposed by Ramadan and beta distribution gives a better distribution (beta-weighted weibull distribution) than each of them individually in terms of the estimate of their characteristics in their parameters. The combination was done by using the logit of beta function by Jones. We investigated in the new proposed beta-weighted weibull distribution some basic properties including moments, moment generating functions, survival rate function, and hazard rate function, skewness, and kurtosis, coefficient of variation, asymptotic behaviors and estimation of parameters. The proposed model is much more flexibility and has better representation of data than weighted weibull distribution. A real data set is used to illustrate the importance the potentiality of the new model.

Keywords: Beta-weighted weibull; Weighted weibull; Moment generating function; Kurtosis; Skewness

Introduction

The beta distribution is one of the skewed distributions used on describing uncertainty or random variation on a system. The weighted weibull distribution is used to adjust the probabilities of the events as observed and recorded; while the weibull distribution has been a powerful probability distribution in reliability analysis. Azzalini [1] proposed a model that can be used as an alternative to Gamma and weibull distribution; and the model is slightly modifying and given additional parameter called “sensitive skewness parameter” by Ramadan [2]. The sensitive parameter governs essentially the shape of the probability density function of the $W\alpha,\beta,\alpha$ distribution. The new class of weighted distribution is given by

$$f(y) = \frac{\lambda \beta (1 + \alpha^\beta) y^\alpha e^{-\lambda y^\beta}}{\alpha^\beta}, \quad y > 0$$  \hspace{1cm} (1)

and the distribution function is given by

$$F(y) = \frac{\left(1 + e^{-\lambda y^\beta}\right) + e^{-\lambda^\beta(1+y^\beta)} - 1}{\alpha^\beta}$$  \hspace{1cm} (2)

Studies on generalized forms of weighted weibull distribution are few. The article is outlined as follows: In section 2, we introduce the new proposed beta weighted weibull distribution (BWW) including the model and weibull distribution; and the weighted distribution is used to adjust the probabilities of the events describing uncertainty or random variation on a system. The weighted distribution is a powerful probability distribution in reliability analysis. Azzalini as observed and recorded; while the weibull distribution has been used to adjust the probabilities of the events describing uncertainty or random variation on a system. The weighted distribution is a powerful probability distribution in reliability analysis.

Methods

The new proposed beta weighted weibull distribution

Attention on beta family of distribution came to limelight some years back and several works have been done concerning beta distribution combined with other distributions, in particular after recent works of Eugene et al. [3] and Jones [4], beta Log-Logistic (Lemonte [5], beta-hyperbolic secant (Fischer and Vaughan [6]), beta-Gumbel [7], moments of the beta weibull [8], beta weibull [9], beta exponential [7], beta Pareto [10], beta Rayleigh, beta modified Weibull, beta Nakagami [11] beta-beta among others.

Therefore, suppose $Y$ is a random variable form the distribution with parameters and defined in (1) using the logit of beta function by Jones [4]

$$f(y) = B(a, b)\left(1 + e^{-\lambda y^\beta}\right) + e^{-\lambda^\beta(1+y^\beta)} - 1$$  \hspace{1cm} (3)

The Beta Weighted Weibull BWW ($a, b, \lambda, \alpha, \beta$) Distribution is obtained as follows:

$$f_{BWW}(y) = B(a, b) \left(1 + \alpha^\beta\left(1 - e^{-\lambda y^\beta}\right) + e^{-\lambda^\beta(1+y^\beta)} - 1\right)$$  \hspace{1cm} (4)

$$\lambda e^\alpha = \frac{\lambda \beta (1 + \alpha^\beta) y^\alpha e^{-\lambda y^\beta}}{\alpha^\beta}, \quad y > 0$$  \hspace{1cm} (5)

Where $F(y)$ and $f(y)$, $a, b, \lambda, \alpha, \beta$ and $y > 0$ - $BWW(a, b, \lambda, \alpha, \beta)$.

Setting $t(y) = \frac{\left(1 + \alpha^\beta\right)\left(1 - e^{-\lambda y^\beta}\right) + e^{-\lambda^\beta(1+y^\beta)} - 1}{\alpha^\beta}$

that is $\frac{\partial t}{\partial y} = \frac{\lambda \beta (1 + \alpha^\beta) y^\alpha e^{-\lambda y^\beta}}{\alpha^\beta}, \quad y > 0$
Equation (4) the probability density function of BWW distribution can be expressed as

\[
f_{BWW}(y) = B(a, b)\frac{t^{a-1}[1-t]^{b-1}\partial t}{\partial y} \tag{6}
\]

The new proposed model can accommodate more parameters because there are two additional parameters \((a, b)\) that made it new and different distribution from the other ones (weighted weibull) and this gives the model more flexibility than other ones. The graph below shows the shape of the pdf of the beta weighted weibull distribution at several values of the parameters Figures 1-3 shows BWW distribution at several values of the parameters and it is clear that indeed it is rightly skewed.

**Cumulative distribution function:** The cumulative distribution function (cdf) is derived and expressed as:

\[
F_{BWW}(y) = P(Y \leq y) = \int_{0}^{y} B(a, b)[(1-e^{-t/\alpha})^{1-(1-e^{-\beta t})}]^{-1} \left[\int_{0}^{e^{-\beta t}} \alpha^{a-1} \beta^{b-1} e^{-\beta t}\partial t\right] \partial y \tag{7}
\]

Putting (5) in (7), we get

\[
F_{BWW}(y) = P(Y \leq y) = \int_{0}^{y} B(a, b)\frac{t^{a-1}[1-t]^{b-1}\partial t}{\partial y} \tag{8}
\]

where \(B(t; a, b)\) is called an incomplete beta function.

Following Jones [4] and Shittu and Adepoju [11], the above expression can be written as:

\[
F_{BWW}(y) = B(a, b)^{-1}\alpha^{-1}(1+\alpha 1^{\alpha-1})t + \ldots + \frac{(1-h)(2-h)(n-h)t^{n}}{n!(a + n)} \tag{9}
\]

**Survival rate function:** The survival rate function of a random variable \(y\) with cumulative distribution function \(F(y)\) is given by

\[
S_{BWW}(y) = 1 - \int_{0}^{y} f(t)\partial t = 1 - F_{BWW}(y) \text{ where } F_{BWW}(y) \text{ equal to (8)},
\]

therefore

\[
S_{BWW}(y) = \frac{B(a, b) - B(t; a, b)}{B(a, b)} \tag{10}
\]
Hazard rate function:

\[ h_{\text{BWW}}(y) = \frac{f_{\text{BWW}}(y)}{1 - F_{\text{BWW}}(y)} \]

where \( f_{\text{BWW}}(y) = B(a,b) t^{a-1} (1-t)^{b-1} \) and \( F_{\text{BWW}}(y) \) as in (8)

Substituting \( f_{\text{BWW}}(y) \) and \( F_{\text{BWW}}(y) \) in the above expression of hazard function, we obtained the hazard rate function of the BWW distribution as given below:

\[ h_{\text{BWW}}(y) = \frac{B(a,b) t^{a-1} (1-t)^{b-1} \lambda}{B(a) - B(t; a,b)} \] (11)

where \( t \) is the distribution in (5)

The asymptotic behavior:

The asymptotic properties of the BWW distribution are examined by considering the behavior of and \( \lim \lim f_{\text{BWW}}(y) \) \( \lim \lim f_{\text{BWW}}(y) \) as follows:

\[ \lim_{y \to -\infty} f_{\text{BWW}}(y) = 0 \]

\[ \lim_{y \to +\infty} B(a,b) \left[ \frac{[1 + \alpha^\beta] \left[ 1 - e^{-\beta(y)} \right] + e^{-\alpha^\beta[y]} - 1}{\alpha^\beta} \right]^{y} \]

For more simplification, we take the limit of

\[ \lim_{y \to -\infty} \frac{\lambda \beta(1 + \alpha^\beta) y^{\alpha^\beta-1} e^{-\beta(y)} (1 - e^{-\alpha^\beta y})}{\alpha^\beta} = 0 \] also

\[ \lim_{y \to +\infty} \frac{\lambda \beta(1 + \alpha^\beta) y^{\alpha^\beta-1} e^{-\beta(y)} (1 - e^{-\alpha^\beta y})}{\alpha^\beta} = 0 \]

This has shown that at least one mode exists. According to literature, whenever, \( y \to -\infty \) and \( y \to +\infty \), then the PDF also tends to zero, hence the BWW distribution has mode.

Special models

Special sub-models (New): The density (4) is important since it also includes as special sub-models, some distributions not previously focused on or considered in the existing literature. The new special sub-models are given below:

- when \( a=1 \), density (4) becomes Lehmann Type II Weighted Weibull (LWW) distribution.
- density (4) can also be simplifies to new exponential weighted weibull distribution when \( b=1 \) and the WW distribution being the parent distribution (as exemplar) when \( a=b=1 \).

Moments and moment generating function

Hosking (1990) described that when a random variable following a generalized beta generated distribution i.e. \( x \sim GBG(f,a,b,c) \) then \( \mu_{r} = E[F^{-1}(\frac{i}{j})] \) where \( j \cup B(a,b,c) \) is a constant and \( F^{-1}(y) \) is the inverse of CDF of the weighted weibull distribution, since \( BWW(a,b,x,\alpha) \) distribution is a special form when \( b=1 \). We then derive the moment generating function (mgf) of the proposed distribution \( M(t) = E(e^{yt}) \) and the general \( r^{th} \) moment of a beta generated distribution is defined by

\[ \mu_{r} = \frac{1}{B(a,b)} \int_{0}^{1} F^{-1}(y)^{y^{a^\beta}[1-y]^{b^\beta}} dy \] (16)

Cordeiro et al. [12] discussed another mgf of \( y \) for generated beta distribution

\[ M(t) = \frac{1}{B(a,b)} \sum_{n=0}^{\infty} (-1)^{n} \int_{0}^{1} \rho(t, r; ai-1) f(\alpha^\beta) \] (17)

Where, \( \rho(t, r) = \int_{0}^{1} e^{yt} F(\alpha^\beta) f(y) dy \)

Therefore, \( M_{(r)} = \frac{1}{B(a,b)} \sum_{n=0}^{\infty} (-1)^{n} \int_{0}^{1} \rho(t, r; ai-1) f(\alpha^\beta) \) (18)

Substituting both pdf and cdf (F(y) & f(y)) of the weighted weibull distribution into (18), we have

\[ M_{(r)} = \frac{1}{B(a,b)} \sum_{n=0}^{\infty} (-1)^{n} \int_{0}^{1} e^{yt} \left[ \frac{[1 + \alpha^\beta] \left[ 1 - e^{-t^{\beta}} \right] + e^{-\alpha^\beta y} - 1}{\alpha^\beta} \right]^{y^{\alpha^\beta}} \] (19)
setting \(a=b=1\) in (18) yields the moment generating function of the parent distribution.

**Parameter estimation**

We derive the maximum likelihood estimate (MLEs) of the parameters of the BWW distribution. Now, let \(\theta\) be a vector of parameters. Shittu and Adepoju \[11\] gave the log-likelihood function for, where \(\theta = (\alpha, \beta, \lambda)\), and even was used by Shittu and Adepoju \[11\]. as in (1) and (2)

\[
L_i = \log(y; \theta) + (a-1)\log(F(y; \theta)) + (b-1)\log[1-F(y; \theta)]
\]

when \(c=1\), reduces the class of generalized beta distribution to the class of beta generated distribution. Then, we obtain \(\theta = (a, b, 1, \omega)\) given as

\[
L_{BWW} = -n\log[B(a, b)] + \sum_{i=1}^{n} \log f(y; \theta) + (a-1) \sum_{i=1}^{n} \log F(y; \theta) + (b-1) \sum_{i=1}^{n} \log[1-F(y; \theta)]
\]

where \(f(y; \theta)\) and \(F(y; \theta)\) as in (1) and (2) and recall that \(B(a, b) = \Gamma(a) \Gamma(b) / \Gamma(a+b)\)

\[
\frac{\partial L(\theta)}{\partial (\alpha)} = -n\Gamma(a) + nI(\alpha + b) + \sum_{i=1}^{n} \log \left[ \frac{1+\alpha y}{\alpha y} \right] + e^{-\alpha y} - 1
\]

\[
\frac{\partial L(\theta)}{\partial (\beta)} = \frac{n\Gamma(a) + n\Gamma(\alpha + b)}{\Gamma(a+b)} + \sum_{i=1}^{n} \log \left[ \frac{1+\alpha y}{\alpha y} \right] + e^{-\alpha y} - 1
\]

\[
\sum_{i=1}^{n} \log \left[ \frac{1+\alpha y}{\alpha y} \right] + e^{-\alpha y} - 1
\]

\[
\frac{\partial L(\theta)}{\partial (\lambda)} = \sum_{i=1}^{n} \left[ \frac{\partial \log B(a, b)}{\partial \lambda} \right] + \sum_{i=1}^{n} \log \left[ \frac{1+\alpha y}{\alpha y} \right] + e^{-\alpha y} - 1
\]

\[
\sum_{i=1}^{n} \log \left[ \frac{1+\alpha y}{\alpha y} \right] + e^{-\alpha y} - 1
\]

The above equations can be solved using Newton Ralphson method to obtain the \(a, b, \alpha, \beta, \lambda\) the MLE of \((a, b, \alpha, \beta, \lambda)\) respectively.

Taking second derivatives of equations 23, 24, 25, 26 and 27 with respect to the parameters above we can derive the interval estimate and hypothesis tests on the model parameter and inverse of fisher’s information matrix needed.

**Results and Discussion**

**Application to real data**

In this section we applied the data set studied by Lee and Wang \[13\] and Lemonte (BIPS Accepted Manuscript) to compare between the results of the proposed BWW, LWW and sub-special models. The data set is secondary data consists of 128 bladder cancer patients reported by Lee and Wang and used by Lemonte [5]. The Akaike Information Criterion (AIC) of the Beta Weighted Weibull (BWW), Lehmann Type II Weighted Weibull (LWW), Exponentiated Weighted Weibull (EWW) and Weighted Weibull distribution (WW) is by using R software (codes) to obtain. The table below consist the estimate, standard error and t-value of the proposed model and sub-models. From Table 1, one can see that both LWW and EWW are the same throughout and BWWD is better than others. Figure 4 Carries the Histogram of x (Data), and, Figure 5 shows that the data used is non-normal (skewed data).
Figure 4: Carries the Histogram of x (Data).

Figure 5: It shows that the data used is not normal (skewed data).

### Table 1

| S/No | Parameter                  | Estimate | Standard Error | t-value |
|------|----------------------------|----------|----------------|---------|
| 1    | BWWD $(a,b,\alpha,\beta,\lambda)$ | 8.00501  | 0.89508        | 8.9434  |
|      |                            | 8.65608  | 0.98976        | 8.7457  |
|      |                            | 7.54633  | 0.85797        | 8.7979  |
|      |                            | 8.11542  | 0.89559        | 9.0616  |
|      |                            | 9.63819  | 1.08461        | 8.8683  |
| 2    | LWW When $a=1, b, \alpha,\beta,\lambda$ | 8.95941  | 0.89564        | 10.0034 |
|      |                            | 8.17089  | 0.80891        | 10.1011 |
|      |                            | 8.36111  | 0.84003        | 9.9534  |
|      |                            | 7.91621  | 0.78043        | 10.1434 |
| 3    | EWW When $b=1, a, \alpha,\beta,\lambda$ | 8.95941  | 0.89564        | 10.0034 |
|      |                            | 8.17089  | 0.80891        | 10.1011 |
|      |                            | 8.36111  | 0.84003        | 9.9534  |
|      |                            | 7.91621  | 0.78043        | 10.1434 |
| 4    | WW When $a=b=1, \alpha,\beta,\lambda$ | 7.90434  | 0.68954        | 11.463  |
|      |                            | 8.44282  | 0.73013        | 11.563  |
|      |                            | 8.70033  | 0.75332        | 11.549  |

Table 1: LWW and EWW are the same throughout and BWWD is better than others.

### Conclusions

The study of skew models is useful in modeling skew data that brings about the new proposed distribution which generalizes the weighted weibull distribution and the new distribution includes as special sub-models other distribution. We call the new model as the beta weighted weibull (BWW) distribution and it was studied mathematically and with some properties. We also derive the density and distribution function, survival rate, hazard rate, asymptotic behaviours, moments and moment generating function and show graphically in Figure 1 above shows the shape of the pdf through several values, also the shape is rightly skewed. The parameters of the propose distribution were estimated and inverse of fisher information matrix is derived. Application to real data shows that apart from the Beta Weighted Weibull is more flexible, it is also has better representation of data than weighted weibull model. Furthermore, the new model may be applicable to many areas such as survival analysis, economics, pollution, engineering, environmental etc.
Authors' contributions

The new proposed distribution which generalizes the weighted Weibull distribution and the new distribution includes as special sub-models other distribution. We call the new model as the beta weighted Weibull (BWW) distribution and it was studied mathematically and with some properties. The combination that brings about the beta weighted Weibull model was done by using the logic of beta function by Jones [4]. We studied skew models in order to use the new model to analyze skew data. We then applied the data set studied by Lee and Wang [13] and Lemonte [5] to compare between the results of the proposed BWW and WW. The data set (secondary data) consists of 128 bladder cancer patients reported by Lee and Wang and used by Lemonte [4,13] with the help of R software (codes) for the maximum likelihood estimates and the maximized log-likelihood of both distributions (BWW and WW).

References

1. Azzalini A (1985) A class of distributions which include the normal ones. Scand J Stat 12: 171-178.
2. Ramadan MM (2013) A class of weighted Weibull distributions and Its Properties. Studies in Mathematics Sciences, Canada 1: 35-45.
3. Eugene NC, Famoye F (2002) Beta-Normal distribution and its applications. Communications in Statistics--Theory and Methods 31: 497-512.
4. Jones MC (2004) Families of distributions arising from distributions of order statistics test 13: 1-43.
5. Lemonte J Artur The Beta Log-logistic distribution. BJPS 1-21.
6. Fischer JM, Vaughan D (2010) The Beta-Hyperbolic Secant Distribution. Austrian Journal of Statistics 39: 245-258.
7. Nadarajah S, Kotz S (2005) The beta exponential distribution. Reliability Engineering and System Safety 91: 689-697.
8. Cordeiro GM, Simas AB, Stosic DB (2008) Explicit expressions for moments of the beta Weibull distribution.
9. Famoye F, Lee C, Olugbenga O (2005) The beta-weibull distribution. Journal of Statistical Theory and Applications 4: 121-138.
10. Akinsete A, Famoye F, Lee C (2008) The Beta-Pareto distribution. Statistics 42: 547-563.
11. Shittu OI, Adepoju AK (2013) On the Beta-Nakagami Distribution. Progress in Applied Mathematics 5: 49-58.
12. Cordeiro GM, Alexandra C Ortega, Edwin MM, Sarabia JM (2011) Generalized Beta Generated distributions. ICMA Centre, Discussion Papers in Finance 1-29.
13. Lee ET, Wang JW (2003) Statistical Methods for Survival Data Analysis, 3rd ed. Willy: New York.