LOG-LEHMANN TYPE II WEIGHTED WEIBULL (LLWW) REGRESSION MODEL: THEORY AND METHOD

Badmus, N. Idowu & Bamiduro, T. Adebayo

Yaba College of Technology, P.M.B 2011, Yaba, Lagos, Nigeria.

Redeemer’s University, Ede, Osun State, Nigeria

Abstract
We present four-parameter log-lehmann type II weighted weibull distribution based on the Lehmann type II weighted weibull distribution (Badmus et al, 2014). We introduce the new model (Theory and method) which could be used more effectively in the analysis of survival data. Some properties of the newly proposed model are discuss including moments and moment generating function, survival and hazard rate function, and maximum likelihood estimates (MLEs). The LWW distribution, due to its flexibility in accommodating (increasing, decreasing, bathtub and unimodal) many forms of the risk function. However, we hope that the model may give more realistic fits than other models.

Keywords: Log-Lehmann, Hazard rate, Moment Generating Function, Regression Model, Survival rate, Weighted Weibull,

Background to the Study
Numerous works/researches have been done on regression model by different authors/researchers but not specifically on the new propose (LLWW) distribution. This model is very scanty in the literature. Recently, several research areas in univaritae parametric model family have been introduced e.g hydrology, engineering and survival analysis due to lack of mathematics assumption in non-parametric and semi-parametric to handle survival analysis. For instance, Famoye et al, (2005) introduced the beta weibull distribution using the logit of beta function, Nadarajah and Gupta (2007) applied generalized gamma distribution to drought
data, Cox et al, (2007) studied a parametric survival analysis and taxonomy of
generalized gamma hazard functions ,Gupta and Kundu, (2009) worked on
weighted exponential distribution, Shahbaz et al (2010) based their study on
weighted weibull distribution. Cordeiro et al, (2010) introduced the exponentiated
generalized gamma distribution, Cordeiro et al, (2011) discussed various
mathematical properties of the BW geometric distribution. Badmus and Bamiduro,
(2014) investigated on life length of components with the beta-weighted weibull
distribution.

Moreover, survival data is always non-normal data that needs more flexible, robust
and skew distribution. In spite of this, many researchers have been carried out on
transformation and parameterization. For example, the generalized log-gamma
distribution is studied by Lawless (2003),Carrasco et al, (2008) introduced a
regression model considering the modified weibull distribution, Ortega et al, (2009)
proposed a modified generalized log-gamma regression model to allow the
possibility that long-term survivors may be presented in the data. Hashimoto et al
(2010) studied the Log-Exponentiated Weibull Model for Interval-Censored Data,
Edwin et al, (Accepted Manuscript) worked on log-beta weibull regression model
with application to predict recurrence of prostate cancer. Therefore, we introduce
the new distribution due to its flexibility in accommodating several forms of the risk
function and it could be used more effectively in the analysis of survival data.

The paper is arranged as follows: in section 2 we define and derive the LLWW
distribution, section 3 contains some properties of the LLWW distribution. We
express moment and generating function in section 4, while section 5 discuss
LLWW regression model and section 6 gives the concluding remarks.

**Log-Lehmann Type II Weighted Weibull Distribution**

This work is extended to LLWW distribution defined by logarithm of the lehmann
weighted weibull (LWW) random variable. Several research have been done on
generalized weibull distributions in reliability literature to give better fitting of
certain data set than traditional two or more parameter weibull models. Now,
having the LWW density function (Badmus, et al; 2014) with three parameters  $b, a$
and $\beta > 0$ is given by for $t > 0$

$$f(t) = \left[ \frac{a + 1}{a} \beta \ t^{\beta - 1} \ exp( -t^\beta ) (1 - exp ( -at^\beta )) \right] b \left[ 1 $$

$$ - \frac{a + 1}{a} \left\{ (1 - exp ( -at^\beta )) \frac{1}{a + 1} \left( 1 - exp \left( (1 + a)t^\beta \right) \right) \right\}^{b+1} \right] $$

(1)
where, \( a \) is the weight parameter and \( \beta \) is the shape parameter. Then, \( b \) is the additional shape parameter to the weighted weibull (WW) distribution to model the skewness and kurtosis of the data. The expression of weighted weibull distribution can be re-written in another weibull version for the sake of simplicity, we have

\[
f(x) = \frac{\gamma + 1}{\gamma} \left( \frac{\alpha}{\beta} \right)^{\alpha} x^{\alpha-1} \exp\left( -\left( \frac{\alpha}{\beta} \right) x \right) \left( 1 - \exp\left( -\gamma \left( \frac{\alpha}{\beta} \right)^{\alpha} x \right) \right)
\]  

(2)

Then, we transform (2) in order to have the Log-Weighted Weibull (LWW) distribution by letting \( Y = \log(x) \) implies \( x = e^Y \), \( a = e^u \) and \( u = \log(\beta) \) implies \( \beta = e^u \) and substituting the transformation in (2), we obtain

\[
f(y) = \frac{\gamma + 1}{\gamma} \exp\left( \frac{\nu - u}{\sigma} \right) \exp\left( -\exp\left( \frac{\nu - u}{\sigma} \right) \right) \left( 1 - \exp\left( -\gamma \exp\left( \frac{\nu - u}{\sigma} \right) \right) \right)
\]  

(3)

(3) is the Log Weighted Weibull (LWW) distribution; and we can equally write the lehmann type II weighted weibull (LLWW) distribution by adding the beta function to the existing expression in (3).

\[
f(t) = \left[ \frac{\gamma + 1}{\gamma} \left( \frac{\alpha}{\beta} \right)^{\alpha} t^{\alpha-1} \exp\left( -\left( \frac{t}{\beta} \right)^{\alpha} \right) \left( 1 - \exp\left( -\gamma \left( \frac{t}{\beta} \right)^{\alpha} \right) \right) \right] b \left[ 1 - \frac{\gamma + 1}{\gamma} \left( 1 - \exp\left( -\left( \frac{t}{\beta} \right)^{\alpha} \right) \right) \right] \left( \frac{1}{\gamma + 1} \exp\left( -\gamma \left( \frac{t}{\beta} \right)^{\alpha} \right) \right) + \gamma \left( \frac{t}{\beta} \right)^{\alpha} \right]^{b-1}
\]  

(4)

\( t \sim LWW \) \((b, \alpha, \gamma)\) Distribution; where \( Y \) is the weighted parameter, \( a \) is the shape parameter, \( \beta \) is the scale parameter and \( b \) is an additional shape parameter and (4) becomes four parameters LWW distribution.

The most important characteristic of LWW distribution is that it contains, as special sub-models, the WW \( b = \beta = 1, Y = a \) and \( a = \beta \) (Shahbaz et al, 2010), WW \( b = 1, a = \beta \), \( Y = a \) (Ramadan, 2013) and various other distributions. Moreover, the corresponding cumulative distribution function, the survival rate and hazard rate function are as follows.
\[ F(t) = \frac{1}{B(1,b)} \int_0^{F(t)} (1 - K)^{b-1} dk = I_{F(t)}(1,b) \]

where \( K \) and \( F(t) \) is the parent cumulative distribution function.

\[ S(t) = 1 - b \int_0^{F(t)} (1 - K)^{b-1} dk = 1 - I_{F(t)}(1,b) \]

and

\[ h(t) = \frac{b(1 - K)^{b-1}K'}{1 - [1 - (1 - t)^b]} \]

respectively, and where \( K' \) is the pdf of the parent distribution

\[ I_y(1,b) = \frac{1}{B(1,b)} \int_0^{y} (1 - K)^{b-1} dk \]

is the incomplete beta function ratio.

By letting \( T \) be a random variable having the lehmann weighted weibull density function in (2). We then investigate the mathematical properties of the LLWW distribution also defined by the random variable \( Y = \log(T) \). Again, the density function of \( Y \) is parameterized in terms of \( \alpha = \frac{1}{\sigma} \) and \( \mu = \log(\beta) \). The density function of \( Y \) can be expressed as

\[
f(y; b, Y, \mu, \sigma) = \frac{Y + 1}{\sigma \gamma B(1,b)} \exp \left( \frac{Y - \mu}{\sigma} \right) \exp \left( -\exp \left( \frac{Y - \mu}{\sigma} \right) \right) \left[ 1 \right. \\
\left. - \frac{Y + 1}{Y} \left( 1 - \exp \left( -\exp \left( \frac{Y - \mu}{\sigma} \right) \right) \right) \right]^{b-1} = (5) \]

Where \(-\infty < y < \infty, \sigma > 0 \) and \(-\infty < \mu < \infty \) and (5) is the new model as LLWW distribution. Say \( X \sim LLWW(\mu, \sigma, Y, b) \) where, \( u \) is the location parameter, \( \sigma \) is a dispersion parameter, \( Y \) is the weighted parameter and \( a \) and \( b \) are shape parameters. Therefore, these results hold: if \( X \sim LWW(\alpha, \beta, Y, b) \) then \( Y = \log(x) \sim LLWW(b, \gamma, \mu, \sigma) \).

The corresponding survival rate function to (5)
\[ S_{(y)} = 1 - b \int_0^f (1 - K)^{b-1} dk = 1 - I_b[z(y)] \quad (6) \]

where \[ F_{(y)} = \left[ \frac{y+1}{y} \left( \frac{1 - \exp\left( -\exp\left( \frac{y-\mu}{\sigma} \right) \right)}{1 - \exp\left( (1 + \gamma) \exp\left( \frac{y-\mu}{\sigma} \right) \right)} \right] \]  

Properties of the LLWW distribution

Now, we want to study some properties of the standardized LLWW random variable defined by \( Z = \frac{Y - \mu}{\sigma} \). The density function of \( Z \) becomes

\[ \pi(z; b, \gamma) = b \left[ \frac{y+1}{y} \left( \frac{\exp(z) \exp(-\exp(z)) \left( -\exp(-\gamma \exp(z)) \right)}{1 - \exp(-\exp(z))} \right) \right] \left[ 1 \right. \]

\[ - \left. \frac{y+1}{y} \left( 1 - \exp(-\exp(z)) \right) \right] \left[ 1 - \exp(-\gamma \exp(z)) \right] \left\{ 1 - \exp(-(1 + \gamma) \exp(z)) \right\}^{b-1} \quad (7) \]

The corresponding cumulative distribution function (cdf) is

\[ F_{Z}^{(z)} = I_{\left[ \frac{y+1}{y} \left( \frac{1 - \exp(-\exp(\frac{y-\mu}{\sigma}))}{1 - \exp(-\gamma \exp(z))} \right) \right]}(b). \]

The basic \( b = 1 \) associated to the standardized weighted weibull distribution.

Linear Combination

By expanding the binomial term in (7), we can write

\[ \pi(z; b, \gamma) = b \sum_{j=0}^{\infty} (-1)^j \binom{b-1}{j} \left[ \frac{y+1}{y} \left( \frac{\exp(z) \exp(-\exp(z)) \left( -\exp(-\gamma \exp(z)) \right)}{1 - \exp(-\exp(z))} \right) \right] \left[ \frac{y+1}{y} \left( 1 - \exp(-\gamma \exp(z)) \right) \right]^{(j+1)-1} \quad (8) \]

The density function \( h_{b} = (b - 1) \left[ 1 - \frac{1}{(b-1)^{y+1}} \left( \frac{\exp(z) \exp(-\exp(z)) \left( -\exp(-\gamma \exp(z)) \right)}{1 - \exp(-\gamma \exp(z))} \right) \right] \) for \( b > 0 \) gives Lehmann type II weighted weibull (Badmus et al, 2014 and Carol Alexander et al, 2011) and its corresponding cumulative function is
Furthermore, the LLWW distribution function can be expressed as a linear combination of LWW densities; and for $b=1$, it becomes the log weighted weibull (LWW) distribution which are new models defined here. The LLWW random variable $z$ can be generated directly from the beta variate $V$ with parameter $b > 0$ by

$$Z = \log[-\log(1 - v) + (-\log(1 - v))]$$

Moments and Generating Function

The $s$th ordinary moment of the LBWW distribution (7) is

$$\mu'_s = E(Z^s) = b \int_{-\infty}^{\infty} Z^s \left[ 1 - \frac{y + 1}{y} \left( \frac{\exp(-\exp(y))}{\exp(-\exp(y))} \right) \right]^{b-1} \, dz$$

We expanding the binomial term and setting $u = e^z$, we get

$$\mu'_s = b \sum_{j=0}^{\infty} (-1)^{j} \binom{b-1}{j} \int_{0}^{\infty} \log(u)^s \left[ \frac{y + 1}{y} \left( \frac{\exp(u)\exp(-\exp(y))}{\exp(-\exp(y))} \right) \right]^{(j+1)-1} \, du$$

$$I_{r,(j+1)} = \left( \frac{\partial}{\partial p} \right)^r \left[ (j + 1) \Gamma(p) \right]_{p=1}$$

and thus

$$\mu'_s = b \sum_{j=0}^{\infty} (-1)^{j} \binom{b-1}{j} I_{r,(j+1)}$$

Equation (9) gives the moments of the LLWW distribution and other measures i.e skewness and kurtosis can be calculated from the ordinary moments using well known relationship and the measures are mainly controlled by the parameter of $b$. 

\[
\mu'_s = b \sum_{j=0}^{\infty} (-1)^{j} \binom{b-1}{j} I_{r,(j+1)}
\]
Meanwhile, the moment (7) can be derived from (10) using simple differentiation.

Order Statistics
In many areas of statistical theory and application another crucial aspect is order statistics. The density \( f_{\text{lin}}(z) \) of the order statistics \( z_i \) for \( i = 1, \ldots, n \) for independently identically distributed LLWW random variables \( z, \ldots, z_n \) is expressed as

\[
f_{\text{lin}}(z) = \frac{n(z; b, \gamma)}{B(i, n - i + 1)} \sum_{j=0}^{n-i} (-1)^{j} \binom{n-i}{j} \left[ \frac{1}{\gamma} \left( 1 - \exp(-\exp(z)) \right) \right]^{j} \left( 1 - \exp(-(1+y)\exp(z)) \right)^{(j+1)}
\]

(11)

Here we obtain an expansion for the density function of the LLWW order statistics. For instance, we make use of the incomplete beta function expansion for \( r > 0 \) real non-integer.

\[
I_{\gamma+1}^{(b)} \left[ \frac{(1 - \exp(-\exp(z))) - \frac{1}{\gamma+1} (1 - \exp(-(1+y)\exp(z)))}{(1 + m)(b + m)} \right] = \frac{b \sum_{m=0}^{\infty} (1 - y)^m}{\Gamma(f + w)/\Gamma(f)} \text{ is the ascending factorial. We then get}
\]

\[
I_{\gamma+1}^{(b)} \left[ \frac{(1 - \exp(-\exp(z))) - \frac{1}{\gamma+1} (1 - \exp(-(1+y)\exp(z)))}{(1 + m)(b + m)} \right]
\]
\[ = \sum_{w=0}^{\infty} dwl \left[ \frac{\gamma+1}{\gamma} \left( \frac{\exp(-\omega \exp(z))}{(1+\gamma)\exp(-\omega \exp(z))} \right) \right] \]

where the coefficients \( dw \) (for \( w = 0, 1, \ldots \)) are

\[ d_w = \frac{(-1)^w}{b} \sum_{m=0}^{\infty} \frac{(1 - \gamma)_m (1 + m)}{(1 + m)(b + m)} \]

We use the identity \( (\sum_{w=0}^{\infty} x^w)^n = \sum_{w=0}^{\infty} C_{n,w} x^w \) for a positive integer \( n \) in

\[ I \left[ \frac{\gamma+1}{\gamma} \left( \frac{\exp(-\omega \exp(z))}{\exp(-\omega \exp(z))} - \frac{1}{\gamma+1} \left( \exp(-(1+\gamma)\omega \exp(z)) \right) \right) \right] \]

Again, we have

\[ I \left[ \frac{\gamma+1}{\gamma} \left( \frac{\exp(-\omega \exp(z))}{\exp(-\omega \exp(z))} - \frac{1}{\gamma+1} \left( \exp(-(1+\gamma)\omega \exp(z)) \right) \right) \right] = \sum_{w=0}^{\infty} C_{i+j-1,w} \left[ \frac{\gamma+1}{\gamma} \left( \exp(-\omega \exp(z)) \right) - \frac{1}{\gamma+1} \left( \exp(-(1+\gamma)\omega \exp(z)) \right) \right] \]

Where \( C_{i+j-1,0} = d_0^{i+j-1} \) and for \( w = 1, 2, \ldots, \)

Also, following (Edwin et al accepted manuscript and Ortega 2009), we obtain

\[ C_{i+j-1,w} = (Wd_0)^{-1} \sum_{r=1}^{w} [(i+j)r - w] d_r C_{i+j-1,w-r} \]

Putting (8) and (13) in (11), we get

\[ f_{\text{lin}}(z) = \sum_{m,w=0}^{\infty} (-1)^m \left( \frac{b-1}{m} \right) \sum_{r=1}^{w} [(i+j)r - w] d_r C_{i+j-1,w-r} \]

Where \( g_w = \frac{\sum_{j=-1}^{\infty} \left( \frac{\gamma}{\gamma+1} \right) C_{i+j-1,w}}{e^{(i+n+\delta) b}} \)
The moments, mgf, mean deviations of the LLWW order statistics can be derived from (15) by using the compensation for these quantities of the LLWW distribution. For instance, the $5^{th}$ ordinary moment of $Z_m$ is simply expressed

$$E(X_m^5) = \sum_{m,w=0}^{\infty} (-1)^m \binom{b-1}{m} g_{w} l(s,m+w)$$

where $(s, m + w)$ has been defined just before (9).

The Log-Lehmann Type II Weighted Weibell (LLWW) Regression Model.

According to Ortega et al, (accepted paper), Ortega et al, 2009 that some of the practical applications on lifetimes are affected by explanatory variables such as the weight, blood pressure, cholesterol level and among others. Now, being a family of parametric models it can be applied to provide a good fit to lifetime data tends to yield more pressure estimates of the quantities of interest. Based on LLWW distribution function, the linear location-scale regression model linking the response variable $y_i$ and the explanatory variable vector $X_i = (x_{i1}, \ldots, x_{ip})$ are as follows

$$y_i = X_i^T \beta + \sigma z_i, \quad i = 1, \ldots, n$$

(16)

Where the random error $z_i$ has distribution function (7) and the LLWW model (16) gives possibilities for fitting many difficult and skewed types of data.

Maximum Likelihood Estimation

The likelihood function for the vector of parameters $\theta = (b, \gamma, \sigma, \beta)^T$ from model (16) has the form $l(\theta) = \sum_{i \in \xi} log[f(y_i)] + \sum_{i \in \eta} log[S(y_i)], f(y_i)$ is the density function (5) and $S(y_i)$ is the survival function (6) of $Y_i$.

The log-likelihood function for $\theta$ is reduces to.
Conclusion

Our model (LLWW regression model) can be applied to any survival data since it represents a parametric family of models that includes as special sub-models several widely known regression models and the data can be fitted and analyzed using the subroutine NLMixed in SAS and R soft code. We hope that our model will have better representation of data, flexibility and applicability than some of the regression models.

\[
I(\theta) = -r \log(\log(\sigma) + \log[b]) + \left[ \gamma + \frac{1}{\gamma} \left( \sum_{i \in F} (z_i) + \sum_{i \in F} (-\exp(-\gamma \exp(z_i))) \right) \right] \\
+ \left( b - 1 \right) \sum_{i \in F} \log \left[ \frac{1}{\gamma + 1} \sum_{i \in F} \left( 1 - \exp(-\exp(z_i)) \right) - \frac{1}{\gamma + 1} \sum_{i \in F} \left( 1 - \exp(-(1 + \gamma) \exp(z_i)) \right) \right] \\
+ \sum_{i \in C} \log \left[ \frac{1}{\gamma + \frac{1}{\gamma}} \sum_{i \in C} \left( 1 - \exp(-\exp(z_i)) \right) - \frac{1}{\gamma + \frac{1}{\gamma}} \sum_{i \in C} \left( 1 - \exp(-(1 + \gamma) \exp(z_i)) \right) \right] (b) (17)
\]

Where \( r \) is the number of uncensored observations (failures) and \( z_i = \frac{(y - x_i^T \beta)}{\sigma} \).

The MLE \( \hat{\theta} \) of the vector \( \theta \) of unknown parameters can be calculated by maximizing the by likelihood (17). The LLWW model survival function of \( Y \) for any individual with explanatory vector \( x \) is given as:

\[
S(y; \hat{\beta}, \hat{\gamma}, \hat{\sigma}\hat{\beta}^T) = 1 - I \left[ \frac{1}{\gamma + \frac{1}{\gamma}} \sum_{i \in C} \left( 1 - \exp(-\exp(z_i)) \right) - \frac{1}{\gamma + \frac{1}{\gamma}} \sum_{i \in C} \left( 1 - \exp(-(1 + \gamma) \exp(z_i)) \right) \right] (b) (18)
\]

The invariance property of the MLEs yields the survival function for \( T = \exp(Y) \)

\[
S(y; \hat{\beta}, \hat{\sigma}, \hat{\beta}, \hat{\gamma}, \hat{\sigma}) = 1 - I \left[ \frac{1}{\gamma + \frac{1}{\gamma}} \sum_{i \in C} \left( 1 - \exp(-(\gamma \hat{\beta})) \right) - \frac{1}{\gamma + \frac{1}{\gamma}} \sum_{i \in C} \left( 1 - \exp(-(1 + \gamma) \exp(\gamma \hat{\beta})) \right) \right] (b) (19)
\]

where \( \hat{\sigma} = \frac{1}{\hat{\sigma}} \) and \( \hat{\beta} = \exp(\gamma \hat{\beta}) \).

Conclusion

Our model (LLWW regression model) can be applied to any survival data since it represents a parametric family of models that includes as special sub-models several widely known regression models and the data can be fitted and analyzed using the subroutine NLMixed in SAS and R soft code. We hope that our model will have better representation of data, flexibility and applicability than some of the regression models.
References


Edwin, M. M., Ortega, Gauss M. Cordeiro & Michael W. Kattan. (Accepted Manuscript). “The Log-Beta Weibull Regression Model with Application to Predict Recurrence of Prostate Cancer.”


