

The Topp-Leone Generated Weibull Distribution: Regression Model, Characterizations and Applications

Gokarna R. Aryal¹, Edwin M. Ortega², G. G. Hamedani³, & Haitham M. Yousof⁴

¹ Department of Mathematics, Statistics and CS, Purdue University Northwest, USA

² Departamento de Ciências Exatas, Universidade de São Paulo, Piracicaba, SP, Brazil

³ Department of Mathematics, Statistics and Computer Science, Marquette University, USA

⁴ Department of Statistics, Mathematics and Insurance, Benha University, Egypt

Correspondence: Gokarna R. Aryal, Department of Mathematics, Statistics and Computer Science, Purdue University Northwest, Hammond, IN, USA. Tel: 1-219-9892701. E-mail: aryalg@pnw.edu

Received: October 13, 2016 Accepted: December 10, 2016 Online Published: December 27, 2016

doi:10.5539/ijsp.v6n1p126 URL: <https://doi.org/10.5539/ijsp.v6n1p126>

Abstract

This paper introduces a new four-parameter lifetime model called the Topp Leone Generated Weibull (TLGW) distribution. This distribution is a generalization of the two parameter Weibull distribution using the genesis of Topp-Leone distribution. We derive many of its structural properties including ordinary and incomplete moments, quantile and generating functions and order statistics. Parameter estimation using maximum likelihood method and simulation results to assess effectiveness of the distribution are discussed. Also, for the first time, we introduce a regression model based on the new distribution. We prove empirically the importance and flexibility of the new model in modeling various types of real data sets.

Keywords: Topp-Leone distribution, Weibull distribution, order statistics, parameter estimation, regression model, simulation

1. Introduction

The Weibull distribution has an undeniable popularity in probability and statistics due to its versatility of modeling real world data. Yet there are many cases where the classical Weibull distribution is unable to capture true phenomenon under study. Therefore, several of its generalizations have been proposed and studied. A generalized form of Weibull distribution is obtained by inducting one or more parameter(s) to the 2-parameter Weibull distribution. It has been proven that several of these generalized distribution are more flexible and are capable of modeling real world data better than the classical Weibull distribution. A state-of-the-art survey on the class of such generalized Weibull distributions can be found in Lai, Xie, & Murthy (2001) and Nadarajah (2009).

Some generalization of the Weibull distribution studied in the literature includes, but are not limited to, exponentiated Weibull (Mudholkar & Srivastava, 1993; Mudholkar, Srivastava, & Freimer, 1995; Mudholkar, Srivastava, & Kollia, 1996), additive Weibull (Xie & Lai, 1995), Marshall-Olkin extended Weibull (Ghitany, Al-Hussaini, & Al-Jarallah, 2005), beta Weibull (Famoye, Lee, & Olumolade, 2005), modified Weibull (Sarhan & Zaindin, 2009), beta modified Weibull (Silva, Ortega, & Cordeiro, 2010), transmuted Weibull (Aryal & Tsokos, 2011), extended Weibull (Xie, Tang, & Goh, 2002), modified Weibull (Lai, Xie, & Murthy, 2003), Kumaraswamy Weibull (Cordeiro, Ortega, & Nadarajah, 2010), Kumaraswamy modified Weibull (Cordeiro, Ortega, & Silva, 2012), Kumaraswamy inverse Weibull (Shahbaz, Shazbaz, & Butt, 2012), exponentiated generalized Weibull (Cordeiro, Ortega, & Cunha 2013), McDonald modified Weibull (Merovci & Elbatal, 2013), beta inverse Weibull (Hanook, Shahbaz, Mohsin, & Kibria, 2013), transmuted exponentiated generalized Weibull (Yousof et al., 2015), McDonald Weibull (Cordeiro, Hashimoto, & Ortega, 2014), gamma Weibull (Provost, Saboor, & Ahmad, 2011), transmuted modified Weibull (Khan & King, 2013), beta Weibull (Lee, Famoye, & Olumolade, 2007), generalized transmuted Weibull (Nofal, Afify, Yousof, & Cordeiro, 2015), transmuted additive Weibull (Elbatal & Aryal, 2013), exponentiated generalized modified Weibull (Aryal & Elbatl, 2015), transmuted exponentiated additive Weibull (Nofal, Afify, Yousof, & Louzada, 2016), Marshall Olkin additive Weibull (Afify, Cordeiro, Yousof, Saboor, & Ortega, 2016) and Kumaraswamy transmuted exponentiated additive Weibull (Nofal, Afify, Yousof, Granzotto, & Louzada, 2016) distributions.

In this paper we introduce a new generalization of the Weibull distribution using the genesis of the Topp-Leone distribution and is named as Topp-Leone Generated Weibull (TLGW) distribution. Consider the Topp-Leone generated family of distributions proposed by Rezaei, Sadr, Alizadeh & Nadarajah (2016) with its probability density function (pdf) and

cumulative distribution function (cdf) given by,

$$f_{TLG}(x) = 2\alpha\theta g(x)[G(x)]^{\theta\alpha-1} [1 - G(x)]^\theta [2 - G(x)]^{\alpha-1}, \quad x \geq 0 \quad (1)$$

$$F_{TLG}(x) = \{G(x)^\theta [2 - G(x)]^\theta\}^\alpha, \quad x \geq 0 \quad (2)$$

respectively. In this paper we will use this generalization to the Weibull (W) distribution whose pdf and cdf are given, respectively, by

$$g(x) = \beta\eta^\beta x^{\beta-1} e^{-(\eta x)^\beta} \quad (3)$$

and

$$G(x) = 1 - e^{-(\eta x)^\beta}. \quad (4)$$

By inserting (3) and (4) into (1), we can write the pdf of the TLGW distribution as

$$f_{TLGW}(x) = 2\alpha\theta\beta\eta^\beta x^{\beta-1} e^{-(\eta x)^\beta} \left(1 - e^{-(\eta x)^\beta}\right)^{\theta\alpha-1} \left[1 - \left(1 - e^{-(\eta x)^\beta}\right)^\theta\right] \left[2 - \left(1 - e^{-(\eta x)^\beta}\right)^\theta\right]^{\alpha-1} \quad (5)$$

The corresponding cdf of TLGW distribution is given by

$$F_{TLGW}(x) = \left(1 - e^{-(\eta x)^\beta}\right)^{\alpha\theta} \left[2 - \left(1 - e^{-(\eta x)^\beta}\right)^\theta\right]^\alpha \quad (6)$$

Figure 1 illustrates the graphical behavior of the pdf of TLGW distribution for selected parameter values. As we shall see in the sequel, this is a rather flexible family compared to the Weibull distribution.

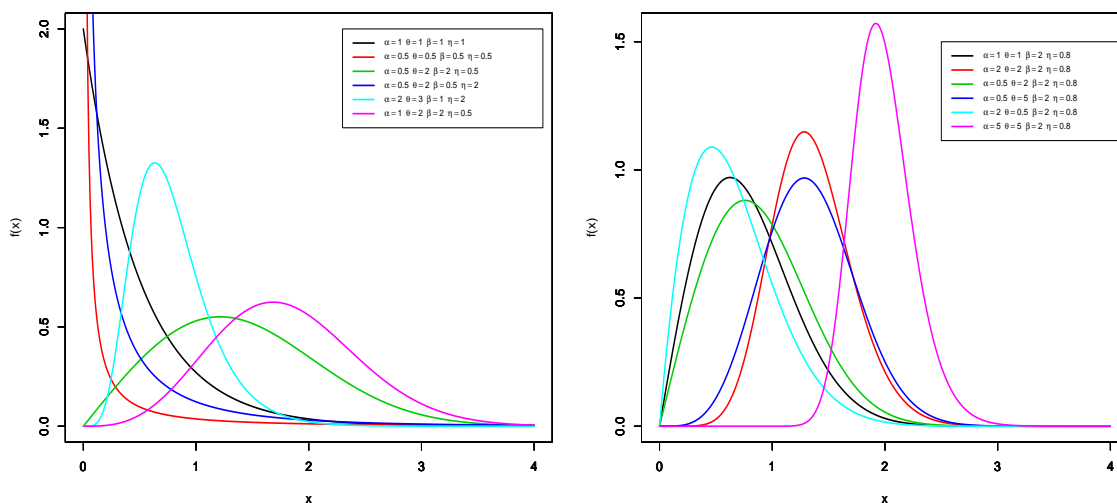


Figure 1. Probability density function of TLGW distribution

In order to derive mathematical and statistical properties of the TLGW distribution, the series expansion of its pdf and cdf will be useful. The cdf in (6) can be expressed as

$$F(x) = \sum_{k=0}^{\infty} \Upsilon_k \Pi_{[(\alpha+k)\theta]}(x) \quad (7)$$

where $\Upsilon_k = (-1)^k 2^{\alpha-k} \binom{\alpha}{k}$ and $\Pi_{(\alpha+k)\theta}(x) = \left(1 - e^{-(\eta x)^\beta}\right)^{(\alpha+k)\theta}$. Observe that $\Pi_\gamma(x) = \left(1 - e^{-(\eta x)^\beta}\right)^\gamma$ is the cdf of exponentiated-Weibull (exp-W) distribution with power parameter γ . This means the TLGW distribution can be expressed as a linear mixture of the exp-W distribution. Similarly, pdf (5) can be expressed as

$$f(x) = \sum_{k=0}^{\infty} \Upsilon_k \pi_{[(\alpha+k)\theta]}(x) \quad (8)$$

with

$$\pi_\gamma(x) = \underbrace{\gamma \beta \eta^\beta x^{\beta-1} e^{-(\eta x)^\beta}}_{g(x;\beta,\eta)} \underbrace{\left[1 - e^{-(\eta x)^\beta}\right]^{\gamma-1}}_{G(x;\beta,\eta)^{\gamma-1}}.$$

Equation (8) reveals that the density of X can be expressed as a linear mixture of exp-W densities. So, several mathematical properties of the new family can be obtained from those of the exp-W distribution.

The paper is unfolded as follows. In Section 2, we obtain some mathematical properties including moments, cumulants, generating function, residual, reversed residual life functions, stress-strength model and order of the proposed distribution. In Section 3, we provide a useful characterization of the new distribution. In Section 4, the model parameters are estimated by maximum likelihood and a simulation study is performed. In Section 5, we present a regression model based on the TLGW distribution with censored data. In Section 6, the usefulness of the new distribution is illustrated by means of four real data sets, where we prove empirically that it outperforms some well-known lifetime distributions. Finally, Section 7 offers some concluding remarks.

2. Mathematical Properties

In this section we will provide some mathematical properties of the TLGW distribution including the moments, incomplete moments, mean deviations, order statistic etc.

2.1 Moments, Cumulants and Generating Function

The r th ordinary moment of X is given by $\mu'_r = E(X^r) = \int_{-\infty}^{\infty} x^r f_{TLGW}(x) dx$. For any $r > -\beta$, we have

$$\mu'_r = \sum_{k,h=0}^{\infty} \Psi_{k,h} \Gamma\left(1 + \frac{r}{\beta}\right), \quad (9)$$

where

$$\Psi_{k,h} = \Upsilon_k \frac{(-1)^h \Gamma((\alpha+k)\theta+1)}{h! \eta^r \Gamma((\alpha+k)\theta+1-h)(h+1)^{\frac{r+\beta}{\beta}}}.$$

Henceforth, $Y_{[(\alpha+k)\theta]}$ denotes the exp-W distribution with power parameter $[(\alpha+k)\theta]$. Setting $r = 1$ in (9), we have the mean of X . The r th central moment of X , say M_r , follows as $M_r = E(X - \mu)^r = \sum_{h=0}^r (-1)^h \binom{r}{h} (\mu'_1)^r \mu'_{r-h}$. The skewness and kurtosis measures also can be calculated from the ordinary moments using well-known relationships. The cumulants (κ_n) of X follow recursively from $\kappa_n = \mu'_n - \sum_{r=0}^{n-1} \binom{n-1}{r-1} \kappa_r \mu'_{n-r}$, where $\kappa_1 = \mu'_1$, $\kappa_2 = \mu'_2 - \mu_1'^2$, $\kappa_3 = \mu'_3 - 3\mu'_2\mu'_1 + \mu_1'^3$, etc. The moment generating function (mgf) of X is given by $M_X(t) = E(e^{tX})$. Clearly, It can be derived using equation (9), for $r > -\beta$, as

$$M_X(t) = \sum_{k=0}^{\infty} \Upsilon_k M_{[(\alpha+k)\theta]}(t) = \sum_{k=0}^{\infty} \Upsilon_k \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu'_r = \sum_{r,k,h=0}^{\infty} \Psi_{k,h} \frac{t^r}{r!} \Gamma\left(1 + \frac{r}{\beta}\right).$$

The effect of the parameters α and θ on mean, variance, skewness and kurtosis, for given values of $\beta = 2$ and $\eta = 2$, are displayed in Figures 2 and 3.

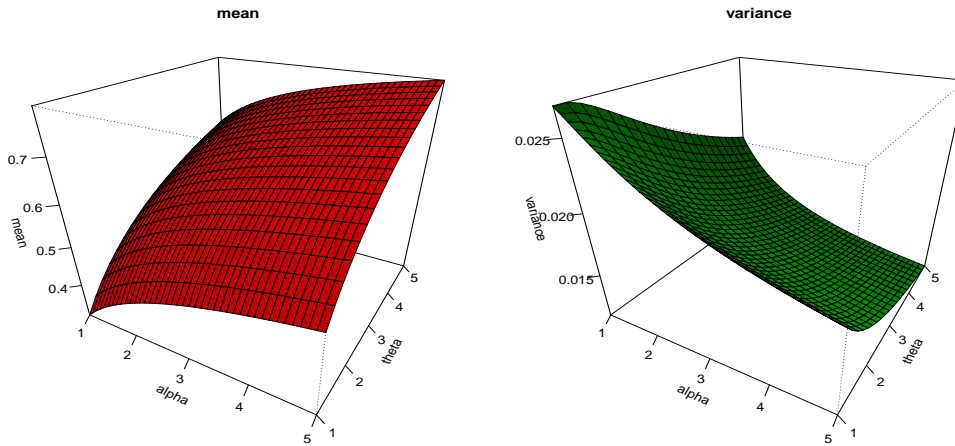


Figure 2. Plots of mean and variance of TLGW distribution

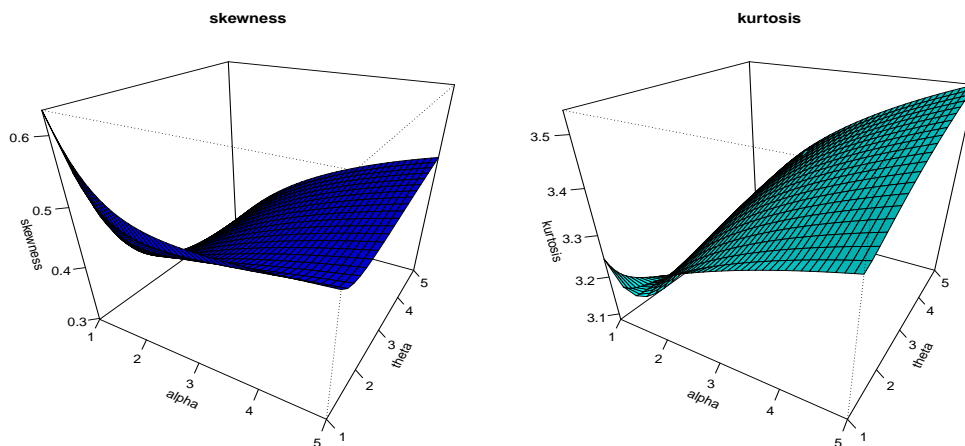


Figure 3. Plots of skewness and kurtosis of TLGW distribution

2.2 Incomplete Moments and Mean Deviations

The main applications of the first incomplete moment refer to the mean deviations and the Bonferroni and Lorenz curves. These curves are very useful in economics, reliability, demography, insurance and medicine. The s th incomplete moment, say $\varphi_s(t)$, of X for $s > -\beta$ can be expressed from (8) as

$$\begin{aligned}\varphi_s(t) &= \sum_{k=0}^{\infty} \Upsilon_k \int_{-\infty}^t x^s \pi_{[(\alpha+k)\theta]}(x) dx \\ &= \sum_{k,h=0}^{\infty} \Upsilon_k \frac{(-1)^h \Gamma((\alpha+k)\theta+1)}{h! \eta^s \Gamma((\alpha+k)\theta+1-h) (h+1)^{\frac{s+\beta}{\beta}}} \Gamma\left(1 + \frac{s}{\beta}, \left(\frac{\eta}{y}\right)^{\beta}\right).\end{aligned}\quad (10)$$

The mean deviations about the mean [$\delta_1 = E(|X - \mu'_1|)$] and about the median [$\delta_2 = E(|X - M|)$] of X are given by $\delta_1 = 2\mu'_1 F(\mu'_1) - 2\varphi_1(\mu'_1)$ and $\delta_2 = \mu'_1 - 2\varphi_1(M)$, respectively, where $\mu'_1 = E(X)$, $M = \text{Median}(X) = Q(0.5)$ is the median, $F(\mu'_1)$ is easily calculated from (4) and $\varphi_1(t)$ is the first incomplete moment given by (10) with $s = 1$. A general equation

for $\varphi_1(t)$ can be derived from (10) as

$$\varphi_1(t) = \sum_{k,h=0}^{\infty} \Upsilon_k \frac{(-1)^h \Gamma((\alpha+k)\theta+1)}{h! \eta \Gamma((\alpha+k)\theta+1-h)(h+1)^{\frac{1+\beta}{\beta}}} \Gamma\left(1 + \frac{1}{\beta}, \left(\frac{\eta}{y}\right)^{\beta}\right).$$

2.3 Residual and Reversed Residual Life Functions

The n th moment of the residual life, say $m_n(t) = E[(X-t)^n | X > t]$, $n = 1, 2, \dots$, uniquely determines $F(x)$. The n th moment of the residual life of X is given by $m_n(t) = \frac{1}{R(t)} \int_t^{\infty} (x-t)^n dF(x)$. Therefore,

$$m_n(t) = \frac{1}{R(t)} \sum_{k,h=0}^{\infty} \Upsilon_k^* \frac{(-1)^h \Gamma((\alpha+k)\theta+1)}{h! \eta^n \Gamma((\alpha+k)\theta+1-h)(h+1)^{\frac{n+\beta}{\beta}}} \Gamma\left(1 + \frac{n}{\beta}, \left(\frac{\eta}{t}\right)^{\beta}\right),$$

where $\Upsilon_k^* = \Upsilon_k \sum_{r=0}^n \binom{n}{r} (-t)^{n-r}$. Another interesting function is the mean residual life (MRL) function or the life expectation at age t defined by $m_1(t) = E[(X-t) | X > t]$, which represents the expected additional life length for a unit which is alive at age t . The MRL of X can be obtained by setting $n = 1$ in the last equation. The n th moment of the reversed residual life, say $M_n(t) = E[(t-X)^n | X \leq t]$ for $t > 0$ and $n = 1, 2, \dots$, uniquely determines $F(x)$. We obtain $M_n(t) = \frac{1}{F(t)} \int_0^t (t-x)^n dF(x)$. Then, the n th moment of the reversed residual life of X becomes

$$M_n(t) = \frac{1}{F(t)} \sum_{k,h=0}^{\infty} \Upsilon_k^{**} \frac{(-1)^h \Gamma((\alpha+k)\theta+1)}{h! \eta^n \Gamma((\alpha+k)\theta+1-h)(h+1)^{\frac{n+\beta}{\beta}}} \Gamma\left(1 + \frac{n}{\beta}, \left(\frac{\eta}{t}\right)^{\beta}\right),$$

where $\Upsilon_k^{**} = \Upsilon_k \sum_{r=0}^n (-1)^r \binom{n}{r} t^{n-r}$. The mean inactivity time (MIT) or mean waiting time (MWT), also called the mean reversed residual life function, is given by $M_1(t) = E[(t-X) | X \leq t]$ and it represents the waiting time elapsed since the failure of an item on condition that this failure had occurred in $(0, t)$. The MIT of the TLGW distributions can be obtained easily by setting $n = 1$ in the above equation.

2.4 A Stress-Strength Model

Stress-strength model is the most widely approach used for reliability estimation. This model is used in many applications of physics and engineering such as strength failure and system collapse. In stress-strength modeling, $\mathbf{R} = \Pr(X_2 < X_1)$ is a measure of reliability of the system when it is subjected to random stress X_2 and has strength X_1 . The system fails if and only if the applied stress is greater than its strength and the component will function satisfactorily whenever $X_1 > X_2$. R can be considered as a measure of system performance and naturally arise in electrical and electronic systems. Other interpretation can be that, the reliability, say \mathbf{R} , of the system is the probability that the system is strong enough to overcome the stress imposed on it. Let X_1 and X_2 be two independent random variables have TLGW($\alpha_1, \theta_1, \eta, \beta$) and TLGW($\alpha_2, \theta_2, \eta, \beta$) distributions. Then, we can write

$$\mathbf{R} = \int_0^{\infty} f_1(x; \alpha_1, \theta_1, \eta, \beta) F_2(x; \alpha_2, \theta_2, \eta, \beta) dx = \sum_{k,j=0}^{\infty} \Omega_{k,j} \int_0^{\infty} \pi_{[(\alpha_1+k)\theta_1]+[(\alpha_2+j)\theta_2]}(x) dx,$$

where

$$\Omega_{k,j} = \frac{(-1)^{k+j} 2^{\alpha_1+\alpha_2-k-j} \binom{\alpha_1}{k} \binom{\alpha_2}{j} ((\alpha_1+k)\theta_1)}{[(\alpha_1+k)\theta_1 + (\alpha_2+j)\theta_2]}$$

and

$$\pi_{[(\alpha_1+k)\theta_1]+[(\alpha_2+j)\theta_2]}(x) = [(\alpha_1+k)\theta_1 + (\alpha_2+j)\theta_2] \beta \eta^{\beta} x^{\beta-1} e^{-(\eta x)^{\beta}} \left[1 - e^{-(\eta x)^{\beta}}\right]^{[(\alpha_1+k)\theta_1]+[(\alpha_2+j)\theta_2]-1}.$$

Thus, \mathbf{R} can be simply expressed as

$$\mathbf{R} = \sum_{k,j=0}^{\infty} \Omega_{k,j}.$$

2.5 Order Statistics

Let X_1, \dots, X_n be a random sample from the TLGW distribution and let $X_{(1)}, \dots, X_{(n)}$ be the corresponding order statistics. The pdf of i th order statistic can be written as

$$f_{i:n}(x) = \frac{f(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F^{j+i-1}(x), \quad (11)$$

where $B(\cdot, \cdot)$ is the beta function. Substituting (5) and (6) in (11) the pdf of $X_{i:n}$ can be expressed as the pdf of $X_{i:n}$ can be expressed as

$$f_{i:n}(x) = \sum_{j=0}^{n-i} \sum_{r,k=0}^{\infty} \mathbf{m}_{j,r,k} \pi_{r+k}(x).$$

where

$$\mathbf{m}_{j,r,k} = \frac{r(-1)^j \Upsilon_r \mathbf{f}_{j+i-1,k}}{(r+k) B(i, n-i+1)}$$

and $\mathbf{f}_{j+i-1,k}$ can be obtained recursively from

$$\mathbf{f}_{j+i-1,k} = \frac{1}{(kb_0)} \sum_{m=0}^k [m(j+i)-k] \Upsilon_m \mathbf{f}_{j+i-1,k-m},$$

for $k \geq 1$ with $\mathbf{f}_{j+i-1,0} = b_0^{j+i-1}$. Then, the density function of the TLGW order statistics is a mixture of exp-W density. Based on the last equation, we note that the properties of $X_{i:n}$ follow from those of Y_{r+k} . For example, the moments of $X_{i:n}$ can be expressed as, for $q > -\beta$,

$$E(X_{i:n}^q) = \sum_{j=0}^{n-i} \sum_{r,k=0}^{\infty} \mathbf{m}_{j,r,k} E(Y_{r+k}^q) = \sum_{j=0}^{n-i} \sum_{r,k,h=0}^{\infty} \mathbf{m}_{j,r,k,h}^* \Gamma\left(1 + \frac{q}{\beta}\right), \quad (12)$$

where

$$\mathbf{m}_{j,r,k,h}^* = \mathbf{m}_{j,r,k} \frac{(-1)^h \Gamma(r+k+1)}{h! \eta^q \Gamma(r+k+1-h) (h+1)^{\frac{q+\beta}{\beta}}}.$$

The L-moments are analogous to the ordinary moments but can be estimated by linear combinations of order statistics. They exist whenever the mean of the distribution exists, even though some higher moments may not exist and are relatively robust to the effects of outliers. Based upon the moments in equation (12), we can derive explicit expressions for the L-moments of X as infinite weighted linear combinations of the means of suitable TLGW order statistics. They are linear functions of expected order statistics defined by

$$\lambda_r = \frac{1}{r} \sum_{d=0}^{r-1} (-1)^d \binom{r-1}{d} E(X_{r-d:r}), \quad r \geq 1.$$

3. Characterizations

In this section we present certain characterizations of TLGW distribution. These characterizations are based on the ratio of two truncated moments. Due to the nature of this distribution, we believe that these may be the only possible characterizations of TLGW distribution. Our first characterization result employs a theorem due to Glänzel (1987) see Theorem 1 in Appendix A. Note that the result holds also when the interval H is not closed. Moreover, it could also be applied when the cdf F does not have a closed form. As shown in Glänzel (1990), this characterization is stable in the sense of weak convergence.

Proposition 3.1. Let $X : \Omega \rightarrow (0, \infty)$ be a continuous random variable and let $q_1(x) = (1 - e^{-(\eta x)^\beta})^{\theta(1-\alpha)} \left[2 - (1 - e^{-(\eta x)^\beta})^\theta \right]^{1-\alpha}$ and $q_2(x) = q_1(x) \left[1 - (1 - e^{-(\eta x)^\beta})^\theta \right]$ for $x > 0$. The random variable X belongs to the family (5) if and only if the function ξ defined in Theorem 1 has the form

$$\xi(x) = \frac{2}{3} \left[1 - (1 - e^{-(\eta x)^\beta})^\theta \right], \quad x > 0.$$

Proof. Let X be a random variable with pdf (5), then

$$(1 - F(x)) E[q_1(x) | X \geq x] = \alpha \left[1 - (1 - e^{-(\eta x)^\beta})^\theta \right]^2, \quad x > 0,$$

and

$$[1 - F(x)] E[q_2(x) | X \geq x] = \frac{2\alpha}{3} \left[1 - \left(1 - e^{-(\eta x)^\beta} \right)^\theta \right]^3, \quad x > 0,$$

and finally

$$\xi(x) q_1(x) - q_2(x) = -\frac{1}{3} q_1(x) \left[1 - \left(1 - e^{-(\eta x)^\beta} \right)^\theta \right] < 0 \quad \text{for } x > 0.$$

Conversely, if ξ is given as above, then

$$s'(x) = \frac{\xi'(x) q_1(x)}{\xi(x) q_1(x) - q_2(x)} = \frac{2\theta\beta\eta^\beta x^{\beta-1} e^{-(\eta x)^\beta} \left(1 - e^{-(\eta x)^\beta} \right)^{\theta-1}}{1 - \left(1 - e^{-(\eta x)^\beta} \right)^\theta}, \quad x > 0,$$

and hence

$$s(x) = \ln \left\{ \left[1 - \left(1 - e^{-(\eta x)^\beta} \right)^\theta \right]^{-2} \right\}, \quad x > 0.$$

Now, in view of Theorem 1, X has density (5).

Corollary 3.1. Let $X : \Omega \rightarrow (0, \infty)$ be a continuous random variable and let $q_1(x)$ be as in Proposition 3.1. Then X has pdf (5) if and only if there exist functions q_2 and ξ defined in Theorem 1 satisfying the differential equation

$$\frac{\xi'(x) q_1(x)}{\xi(x) q_1(x) - q_2(x)} = \frac{2\theta\beta\eta^\beta x^{\beta-1} e^{-(\eta x)^\beta} \left(1 - e^{-(\eta x)^\beta} \right)^{\theta-1}}{1 - \left(1 - e^{-(\eta x)^\beta} \right)^\theta}, \quad x > 0.$$

Corollary 3.2. The general solution of the differential equation in Corollary 3.1 is

$$\xi(x) = \left[1 - \left(1 - e^{-(\eta x)^\beta} \right)^\theta \right]^{-1} \left[- \int 2\theta\beta\eta^\beta x^{\beta-1} e^{-(\eta x)^\beta} \left(1 - e^{-(\eta x)^\beta} \right)^{\theta-1} (q_1(x))^{-1} q_2(x) dx + D \right],$$

where D is a constant. Note that a set of functions satisfying the differential equation in Corollary 3.1, is given in Proposition 3.1 with $D = 0$. However, it should be also noted that there are other triplets (q_1, q_2, ξ) satisfying the conditions of Theorem 1.

4. Parameter Estimation

Subsection 4.1 provides procedures for maximum likelihood estimation of the TLGW distribution. Subsection 4.2 assesses the performance of the maximum likelihood estimators (MLEs) in terms of biases and mean squared errors by means of a simulation study.

4.1 Parameter Estimation

Several methods for parameter estimation have been proposed in the literature but the maximum likelihood method is the most commonly employed one. The maximum likelihood estimators enjoy desirable properties and can be used for constructing confidence intervals and regions and also in test statistics. The normal approximation for these estimators, in large samples, can be easily handled either analytically or numerically. So, we consider the estimation of the unknown parameters of this family from complete samples only by maximum likelihood. Let x_1, \dots, x_n be a random sample from the TLGW distribution. Let $\tau = (\alpha, \theta, \beta, \eta)^T$ be the 4×1 parameter vector. To determine the MLE of τ , we use the log-likelihood function (ℓ) of TLGW distribution given by

$$\begin{aligned} \ell(\tau) = & n \ln 2 + n \ln \alpha + n \ln \theta + n \ln \beta + n \beta \ln \eta + (\beta - 1) \sum_{i=1}^n \ln(x_i) - \sum_{i=1}^n (\eta x_i)^\beta \\ & + (\theta\alpha - 1) \sum_{i=1}^n \ln s_i + \sum_{i=1}^n \ln(1 - s_i^\theta) + (\alpha - 1) \sum_{i=1}^n \ln(2 - s_i^\theta), \end{aligned}$$

where $s_i = \left(1 - e^{-(\eta x_i)^\beta}\right)$. Let $z_i = \frac{\partial s_i}{\partial \beta} = (\eta x_i)^\beta e^{-(\eta x_i)^\beta} \ln(\eta x_i)$ and $q_i = \beta \eta^{\beta-1} x_i^\beta e^{-(\eta x_i)^\beta}$. Then the components of the score vector are given by

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\tau})}{\partial \alpha} &= \frac{n}{\alpha} + \theta \sum_{i=1}^n \ln s_i + \sum_{i=1}^n \ln(2 - s_i^\theta), \\ \frac{\partial \ell(\boldsymbol{\tau})}{\partial \theta} &= \frac{n}{\theta} + \alpha \sum_{i=1}^n \ln s_i - \sum_{i=1}^n \frac{s_i^\theta \ln s_i}{(1 - s_i^\theta)} - (\alpha - 1) \sum_{i=1}^n \frac{s_i^\theta \ln s_i}{(2 - s_i^\theta)}, \\ \frac{\partial \ell(\boldsymbol{\tau})}{\partial \beta} &= \frac{n}{\beta} + n \ln \eta + \sum_{i=1}^n \ln(x_i) - \sum_{i=1}^n \frac{\ln(\eta x_i)}{(\eta x_i)^{-\beta}} + (\theta \alpha - 1) \sum_{i=1}^n \frac{z_i}{s_i} - \sum_{i=1}^n \frac{\theta z_i s_i^{\theta-1}}{1 - s_i^\theta} - (\alpha - 1) \sum_{i=1}^n \frac{\theta z_i s_i^{\theta-1}}{2 - s_i^\theta}\end{aligned}$$

and

$$\frac{\partial \ell(\boldsymbol{\tau})}{\partial \eta} = \frac{n\beta}{\eta} - \beta \eta^{\beta-1} \sum_{i=1}^n x_i^\beta + (\theta \alpha - 1) \sum_{i=1}^n \frac{q_i}{s_i} - \theta \sum_{i=1}^n \frac{q_i s_i^{\theta-1}}{(1 - s_i^\theta)} - (\alpha - 1) \sum_{i=1}^n \frac{q_i s_i^{\theta-1}}{(2 - s_i^\theta)}.$$

Now, setting the nonlinear system of equations $\frac{\partial \ell(\boldsymbol{\tau})}{\partial \alpha} = 0$, $\frac{\partial \ell(\boldsymbol{\tau})}{\partial \theta} = 0$, $\frac{\partial \ell(\boldsymbol{\tau})}{\partial \beta} = 0$ and $\frac{\partial \ell(\boldsymbol{\tau})}{\partial \eta} = 0$ and solving them simultaneously yields the MLE $\widehat{\boldsymbol{\tau}} = (\widehat{\alpha}, \widehat{\theta}, \widehat{\beta}, \widehat{\eta})^T$. To solve these equations, it is usually more convenient to use nonlinear optimization methods such as the quasi-Newton algorithm to numerically maximize ℓ . For interval estimation of the parameters, we obtain the 4×4 observed information matrix $J(\boldsymbol{\tau}) = \left\{ \frac{\partial^2 \ell}{\partial r \partial s} \right\}$ (for $r, s = \alpha, \theta, \beta, \eta$), whose elements can be computed numerically. Under standard regularity conditions when $n \rightarrow \infty$, the distribution of $\widehat{\boldsymbol{\tau}}$ can be approximated by a multivariate normal $N_4(0, J(\widehat{\boldsymbol{\tau}})^{-1})$ distribution to construct approximate confidence intervals for the parameters. Here, $J(\widehat{\boldsymbol{\tau}})$ is the total observed information matrix evaluated at $\widehat{\boldsymbol{\tau}}$.

4.2 Simulation Study

In this section, we present some simulations for different sample sizes to assess the accuracy of the MLEs. Simulating random variables from well defined probability distributions has been discussed in the computational statistics literature, e.g; the inverse transformation method, the rejection and acceptance sampling technique, etc. An ideal technique for simulating from the TLGW distribution is the inversion method. We can simulate random variable X by

$$X = \frac{1}{\eta} \left[-\ln \left\{ 1 - \left(1 - \sqrt{1 - U^{\frac{1}{\alpha}}} \right)^{\frac{1}{\theta}} \right\} \right]^{\frac{1}{\beta}},$$

where U is a uniform random number in $(0, 1)$. For selected combinations of θ, α, β and η we generate samples of sizes $n = 50, 100, 200, 300, 500$ and 1000 from the TLGW distribution. We repeat the simulations $N = 1,000$ times and evaluate the mean estimates and the root mean squared errors (RMSEs). The empirical results obtained using the statistical computing software R are given in Tables 1 and 2. It can be observed that as sample size increases the mean squared error decreases. Therefore, the maximum likelihood method works very well to estimate the parameters of TLGW distribution.

Table 1. Empirical means and the RMSEs of TLGW distribution

	$\theta = 0.5$	$\alpha = 1.5$	$\beta = 2.0$	$\eta = 2.0$
n	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\eta}$
50	0.6557 (1.7665)	3.5869 (4.6465)	3.6344 (3.0606)	2.4805 (6.6074)
100	0.6102 (0.7534)	2.9638 (3.0763)	2.9702 (2.0948)	2.0982 (0.8570)
200	0.5622 (0.5062)	2.4859 (2.3428)	2.5936 (1.3669)	1.9948 (0.4574)
300	0.5528 (0.4576)	2.3607 (2.0398)	2.4361 (1.0502)	1.9857 (0.3922)
500	0.5865 (0.4487)	2.1773 (1.7359)	2.2545 (0.7393)	2.0171 (0.3690)
1000	0.5599 (0.3763)	1.9159 (1.1891)	2.1635 (0.5083)	2.0049 (0.2988)

Table 2. Empirical means and the RMSEs of TLGW distribution

	$\theta = 2.0$	$\alpha = 0.5$	$\beta = 1.5$	$\eta = 2.5$
n	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\eta}$
50	3.0078 (3.4059)	1.3099 (10.4155)	2.1327 (1.8721)	2.9353 (2.3342)
100	2.9044 (2.8306)	0.6648 (0.6846)	1.7716 (0.9444)	2.6559 (0.9228)
200	2.8347 (2.2698)	0.5538 (0.4076)	1.5785 (0.4318)	2.6469 (0.5841)
300	2.7806 (2.0799)	0.5285 (0.3081)	1.5396 (0.3188)	2.6439 (0.5508)
500	2.6320 (1.7402)	0.5084 (0.2567)	1.5199 (0.2367)	2.6043 (0.4239)
1000	2.4849 (1.4294)	0.5019 (0.2098)	1.5100 (0.1781)	2.5806 (0.3650)

5. The Log-Topp Leone Generated Weibull Regression Model With Censored Data

Let X be a random variable having the density function (5). The random variable $Y = \log(X)$ has a log-Topp Leone Generated Weibul (LTLGW) distribution, whose density function (parameterized in terms of $\beta = 1/\sigma$ and $\eta = \exp(-\mu)$) can be expressed as

$$f(y) = \frac{2\alpha\theta}{\sigma} \exp\left[\left(\frac{y-\mu}{\sigma}\right) - \exp\left(\frac{y-\mu}{\sigma}\right)\right] \left[1 - \left\{1 - \exp\left[-\exp\left(\frac{y-\mu}{\sigma}\right)\right]\right\}^\theta\right] \times \left\{1 - \exp\left[-\exp\left(\frac{y-\mu}{\sigma}\right)\right]\right\}^{\alpha\theta-1} \left[2 - \left\{1 - \exp\left[-\exp\left(\frac{y-\mu}{\sigma}\right)\right]\right\}^\theta\right]^{\alpha-1}, \quad (13)$$

where $-\infty < y < \infty$, $\sigma > 0$ and $-\infty < \mu < \infty$.

We refer to equation (13) as the (new) LTLGW distribution, say $X \sim \text{TLGW}(\alpha, \theta, \sigma, \mu)$, where μ is a location parameter, σ is a dispersion parameter and α and θ are shape parameters. Thus,

$$\text{if } X \sim \text{LTLGW}(\alpha, \theta, \eta, \beta) \text{ then } Y = \log(X) \sim \text{LTLGW}(\alpha, \theta, \sigma, \mu).$$

The plots of (13) in Figure 5 for selected parameter values show great flexibility of the density function in terms of the parameters α and θ .

The survival function corresponding to (13) becomes

$$S(y) = 1 - \left\{1 - \exp\left[-\exp\left(\frac{y-\mu}{\sigma}\right)\right]\right\}^{\alpha\theta} \left[2 - \left\{1 - \exp\left[-\exp\left(\frac{y-\mu}{\sigma}\right)\right]\right\}^\theta\right]^\alpha. \quad (14)$$

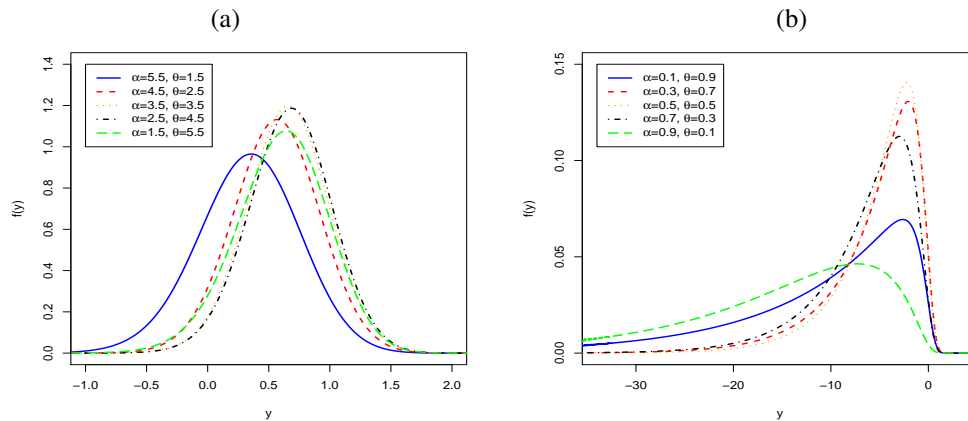


Figure 4. Plots of the LTLGW density for some parameter values. (a) For different values of $\alpha \in (1.0, 5.0)$ and $\beta \in (1.0, 5.0)$ for $\mu = 0$ and $\sigma = 0$. (b) For different values of $\alpha \in (0, 1.0)$ and $\beta \in (0, 1.0)$ for $\mu = 0$ and $\sigma = 0$.

We define the standardized random variable $Z = (Y - \mu)/\sigma$ with density function

$$\pi(z; \alpha, \theta) = 2\alpha\theta \exp[z - \exp(z)] \{1 - \exp[-\exp(z)]\}^{\alpha\theta-1} \left[1 - \{1 - \exp[-\exp(z)]\}^\theta\right] \left[2 - \{1 - \exp[-\exp(z)]\}^\theta\right]^{\alpha-1}. \quad (15)$$

In many practical applications, the lifetimes are affected by explanatory variables such as the cholesterol level, blood pressure, weight and many others. Parametric regression models to estimate univariate survival functions for censored data are widely used. A parametric model that provides a good fit to lifetime data tends to yield more precise estimates of the quantities of interest. Based on the LTLGW density function, we propose a linear location-scale regression model for censored data linking the response variable y_i and the explanatory vector $\mathbf{v}_i^T = (v_{i1}, \dots, v_{ip})$ as follows

$$y_i = \mathbf{v}_i^T \boldsymbol{\gamma} + \sigma z_i, \quad i = 1, \dots, n, \quad (16)$$

where the random error z_i has density function (15), $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_p)^T$, $\sigma > 0$, $a > 0$ and $b > 0$ are unknown parameters. The parameter $\mu_i = \mathbf{v}_i^T \boldsymbol{\gamma}$ is the location of y_i . The location parameter vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^T$ is given by a linear model $\boldsymbol{\mu} = \mathbf{V} \boldsymbol{\gamma}$, where $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n)^T$ is a known model matrix. The LTLGW regression model (16) opens new possibilities for fitting many different types of censored data. It is an extension of an accelerated failure time model using the TLGW distribution for censored data.

Consider a sample $(y_1, \mathbf{v}_1), \dots, (y_n, \mathbf{v}_n)$ of n independent observations, where each random response is defined by $y_i = \min\{\log(x_i), \log(c_i)\}$. We assume non-informative censoring such that the observed lifetimes and censoring times are independent. Let F and C be the sets of individuals for which y_i is the log-lifetime or log-censoring, respectively. Conventional likelihood estimation techniques can be applied here. The log-likelihood function for the vector of parameters $\boldsymbol{\tau} = (\alpha, \theta, \sigma, \boldsymbol{\gamma}^T)^T$ from model (16) has the form

$$l(\boldsymbol{\tau}) = \sum_{i \in F} l_i(\boldsymbol{\tau}) + \sum_{i \in C} l_i^{(c)}(\boldsymbol{\tau}),$$

where $l_i(\boldsymbol{\tau}) = \log[f(y_i|\mathbf{v}_i)]$, $l_i^{(c)}(\boldsymbol{\tau}) = \log[S(y_i|\mathbf{v}_i)]$, $f(y_i|\mathbf{v}_i)$ is the density (13) and $S(y_i|\mathbf{v}_i)$ is the survival function (14) of Y_i . The total log-likelihood function for $\boldsymbol{\tau}$ reduces to

$$\begin{aligned} l(\boldsymbol{\tau}) = & r \log\left(\frac{2\alpha\theta}{\sigma}\right) + \sum_{i \in F} z_i - \sum_{i \in F} \exp(z_i) + (\alpha\theta - 1) \sum_{i \in F} \log\{1 - \exp[-\exp(z_i)]\} + \sum_{i \in F} \log\left[1 - \{1 - \exp[-\exp(z_i)]\}^\theta\right] \\ & + (\alpha - 1) \sum_{i \in F} \log\left[2 - \{1 - \exp[-\exp(z_i)]\}^\theta\right] + \sum_{i \in C} \log\left\{1 - \{1 - \exp[-\exp(z_i)]\}^{\alpha\theta} \left[2 - \{1 - \exp[-\exp(z_i)]\}^\theta\right]^\alpha\right\}, \quad (17) \end{aligned}$$

where $z_i = (y_i - \mathbf{v}_i^T \boldsymbol{\gamma})/\sigma$ and r is the number of uncensored observations (failures). The score functions for the parameters α , β , σ and $\boldsymbol{\gamma}$ are given by $U_\alpha(\boldsymbol{\tau})$, $U_\theta(\boldsymbol{\tau})$, $U_\sigma(\boldsymbol{\tau})$ and $U_{\gamma_j}(\boldsymbol{\tau})$ for $j = 1, \dots, p$.

The MLE $\hat{\boldsymbol{\tau}}$ of $\boldsymbol{\tau}$ is obtained by solving the nonlinear equations $U_\alpha(\boldsymbol{\tau}) = 0$, $U_\theta(\boldsymbol{\tau}) = 0$, $U_\sigma(\boldsymbol{\tau}) = 0$ and $U_{\gamma_j}(\boldsymbol{\tau}) = 0$ simultaneously. These equations cannot be solved analytically and statistical software can be used to solve them numerically.

We can use iterative techniques such as a Newton-Raphson type algorithm to calculate the estimate $\widehat{\tau}$.

We use the subroutine NLMixed in SAS to compute $\widehat{\tau}$. Initial values for γ and σ are taken from the fit of the log-Weibull regression model. The fit of the LTLGW model produces the estimated survival function for y_i given by

$$S(y_i; \hat{\alpha}, \hat{\theta}, \hat{\sigma}, \widehat{\gamma}^T) = 1 - \{1 - \exp[-\exp(\hat{z}_i)]\}^{\hat{\alpha}\hat{\theta}} \left[2 - \{1 - \exp[-\exp(\hat{z}_i)]\}^{\hat{\theta}} \right]^{\hat{\alpha}}, \quad (18)$$

where $\hat{z}_i = (y_i - \mathbf{v}_i^T \widehat{\gamma}) / \hat{\sigma}$.

Under standard regularity conditions, the approximate multivariate normal distribution $N_{p+4}(0, J(\tau)^{-1})$ for $\widehat{\tau}$ can be used in the classical way to construct confidence intervals for the parameters in τ , where $J(\tau)$ is the observed information matrix. Further, we can use LR statistics for comparing some special sub-models with the LTLGW model. We consider the partition $\tau = (\tau_1^T, \tau_2^T)^T$, where τ_1 is a subset of parameters of interest and τ_2 is a subset of the remaining parameters. The LR statistic for testing the null hypothesis $H_0 : \tau_1 = \tau_1^{(0)}$ versus the alternative hypothesis $H_1 : \tau_1 \neq \tau_1^{(0)}$ is $w = 2\{\ell(\widehat{\tau}) - \ell(\widetilde{\tau})\}$, where $\widetilde{\tau}$ and $\widehat{\tau}$ are the estimates under the null and alternative hypotheses, respectively. The statistic w is asymptotically (as $n \rightarrow \infty$) distributed as χ_k^2 , where k is the dimension of the subset of parameters τ_1 of interest.

6. Applications

In this section, we compare the fit of the TLGW distribution to two real uncensored data sets. Further, we fit the new regression model based on the LTLGW distribution to the censored data set studied by Efron (1988). All the computations were performed using the procedure NLMixed in SAS and the R statistical software.

6.1 Application 1: Uncensored Data Sets

In this subsection we will provide some applications of TLGW distribution to model real data set. We will compare the fit of the model with another generalization proposed and studied by Alzaatreh, Lee, & Famoye (2013) (we call it ALF generator). The corresponding cdf of ALF generator is given by $F(x) = 1 - e^{-\alpha\{-\log[1-G(x)]\}^\theta}$. Therefore the cdf of ALFW distribution is $F_{ALFW}(x) = 1 - e^{-\alpha[(\eta x)^\beta]^\theta}$.

The measures of goodness-of-fit including the log-likelihood function evaluated at the MLEs ($\hat{\ell}$), Akaike information criterion (AIC), Anderson-Darling (A^*), Cramer-von Mises (W^*) and Kolmogorov-Smirnov (K-S), are calculated to compare the fitted models. In general, the smaller the values of these statistics, the better the fit to the data.

Data set 1: This data set consists of the waiting times(in seconds), between 65 successive eruptions of the Kiama Blow-hole. These values were recorded with the aid of digital watch on July 12, 1998 by Jim Irish and has been referenced by several authors including da Silva, Thiago, Maciel, Campos, & Cordeiro(2013) and Pinho, Cordeiro & Nobre(2015). The actual data are:

83, 51, 87, 60, 28, 95, 8, 27, 15, 10, 18, 16, 29, 54, 91, 8, 17, 55, 10, 35, 47, 77, 36, 17, 21, 36, 18, 40, 10, 7, 34, 27, 28, 56, 8, 25, 68, 146, 89, 18, 73, 69, 9, 37, 10, 82, 29, 8, 60, 61, 61, 18, 169, 25, 8, 26, 11, 83, 11, 42, 17, 14, 9, 12.

In Table 3 we provide the MLEs and their corresponding standard errors (in parentheses) of the model parameters. The values of $\widehat{\ell}$, AIC, A^* , W^* and K-S statistic are provided in Table 4.

Table 3. Estimated parameters- Kiama Blowhole data

model	α	β	θ	η
TLGW	4.7964 (2.9824)	0.5371 (0.2172)	1.3044 (0.8547)	0.0503 (0.0585)
ALFW	0.6500 (1.7765)	0.9525 (3.7659)	1.3378 (5.2894)	0.0324 (0.0695)

Table 4. The $\widehat{\ell}$, AIC, A^* , W^* and K-S Statistics-Kiama Blowhole data

model	$-\widehat{\ell}$	AIC	A^*	W^*	K-S	p-value(K-S)
TLGW	294.3426	600.0873	0.8042	0.1102	0.0929	0.6373
ALFW	296.9001	601.8003	1.0079	0.1471	0.1113	0.4060

Data set 2: This data set includes an active repair time (in hours) for an airborne communication transceiver reported by Balakrishnan, Leiva, Sanhueza, & Cabrera (2009), which was originally given by Chhikara & Folks (1989). The actual observations are listed below.

0.2, 0.3, 0.5, 0.5, 0.5, 0.5, 0.6, 0.6, 0.7, 0.7, 0.7, 0.8, 0.8, 1.0, 1.0, 1.0, 1.0, 1.1, 1.3, 1.5, 1.5, 1.5, 1.5, 2.0, 2.0, 2.2, 2.5, 2.7, 3.0, 3.0, 3.3, 3.3, 4.0, 4.0, 4.5, 4.7, 5.0, 5.4, 5.4, 7.0, 7.5, 8.8, 9.0, 10.3, 22.0, 24.5.

In Table 5 we provide the MLEs and the corresponding standard errors (in parentheses) of the model parameters. The values of $\hat{\ell}$, AIC , W^* , A^* and K-S statistic are provided in Table 6.

Table 5. Estimated parameters- repair time data

model	α	β	θ	η
TLGW	4.0709 (3.0202)	0.3368 (0.1359)	2.3032 (1.9328)	2.1797 (4.5138)
ALFW	0.4004 (2.591)	0.2324 (1.4755)	3.8658 (24.5386)	0.8163 (5.8762)

Table 6. The $\hat{\ell}$, AIC , A^* , W^* and K-S Statistic-repair time data

model	$-\hat{\ell}$	AIC	A^*	W^*	K-S	p-value(K-S)
TLGW	100.3413	208.6826	0.3699	0.0578	0.0919	0.8313
ALFW	104.4697	216.9394	0.9009	0.1298	0.1204	0.5174

From Tables 4 and 6 it is noted that the values for $-\hat{\ell}$, AIC , A^* , W^* and K-S test statistic are lower for TLGW distribution than for ALFW distribution indicating that the TLGW distribution fits both the Kiama Blowhole data and repair time data better than ALFW distribution. Graphical comparisons of TLGW and ALFW models for the Kiama Blowhole data are provided in Figure 5. Similarly, the graphical comparisons of repair time data are provided in Figure 6. It is evident that the TLGW distribution fits both of these data better than ALFW distribution.

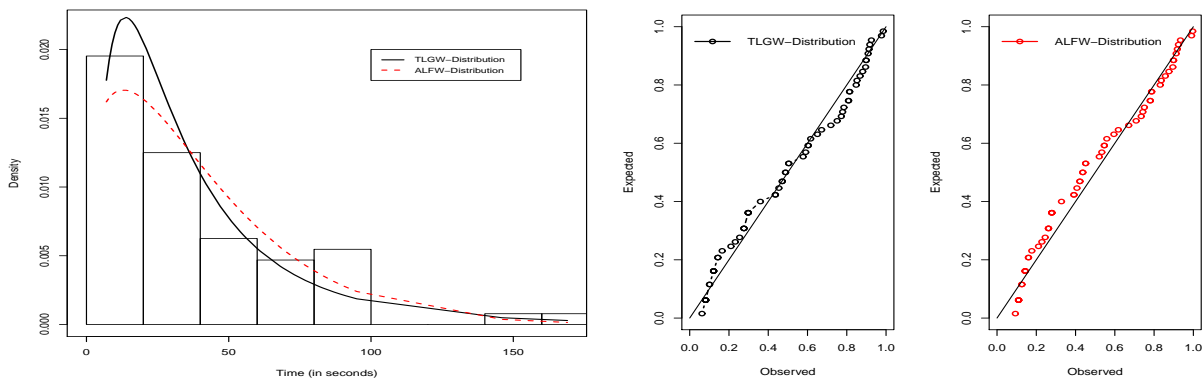


Figure 5. Fitted pdf and QQ plots of TLGW and ALFW distributions for the Kiama Blowhole data

6.2 Application 2: Regression Model With Censored Data

For an application of the LTLGW regression model, we consider the data from a two-arm clinical trial discussed earlier by Efron (1988). Efron (1988) and Mudholkar et al. (1996) observed that the empirical hazard functions for both samples start near zero, suggesting an initial high-risk period in the beginning, a decline for a while and then stabilization after about one year. Specifically, Efron's data from a head and neck cancer clinical trial consist of survival times of 51 patients in arm A (17.6% censored data) who were given radiation therapy and 45 patients in arm B (31.1% censored data) who were given radiation plus chemotherapy. Mudholkar et al. (1996) analyzed these data separately using the exponentiated Weibull distribution. Here, we use the LTLGW regression model.

Let x_i be the survival time (in days) for the i th observation and v_{i1} be the two-Arm clinical trial (0=Arm A, 1=Arm B)

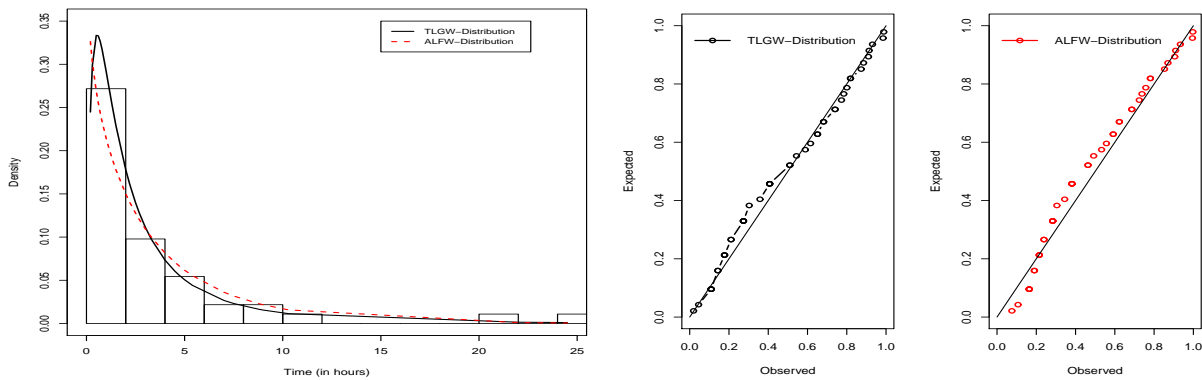


Figure 6. Fitted pdf and QQ plots of TLGW and ALFW distributions for the repair time data

(for other details, see Mudholkar et al. (1996). We propose the model

$$y_i = \gamma_0 + \gamma_1 v_{i1} + \sigma z_i,$$

where the random variable $y_i = \log(x_i)$ follows the LTLGW distribution (13) for $i = 1, 2, \dots, 96$.

An alternative approach for modeling these data can be provided by the log-Weibull (LW) distribution. There are various extensions of this lifetime distribution; see, for example, the log-beta Weibull (LBW) (Ortega, et al. (2011)) distribution. The LBW density function is given by

$$f(y; a, b, \sigma, \mu) = \frac{1}{\sigma B(a, b)} \exp \left\{ \left(\frac{y - \mu}{\sigma} \right) - b \exp \left(\frac{y - \mu}{\sigma} \right) \right\} \left\{ 1 - \exp \left[- \exp \left(\frac{y - \mu}{\sigma} \right) \right] \right\}^{a-1}, \quad (19)$$

where $-\infty < y < \infty$, $\sigma > 0$ and $-\infty < \mu = \gamma_0 + \gamma_1 v_1 < \infty$. The special case $a = b = 1$ corresponds to the LW distribution.

The MLEs of the model parameters are computed using the procedure NLMixed in SAS. As initial values for γ and σ in the iterative algorithm for maximizing the log-likelihood function (17), we adopt the fitted values obtained by fitting the LW regression model. The MLEs of the parameters and the AIC, CAIC and BIC statistics for some models are listed in Table 6.2.

Table 7. MLEs of the parameters from the LTLGW regression model fitted to the Efron data, the corresponding SEs (given in parentheses), p-values in [.] and the basic statistics.

Model	α	θ	σ	β_0	β_1	AIC	CAIC	BIC
LTLGW	25.2819 (5.6059)	45.4687 (8.0643)	11.9181 (2.6627)	-14.9062 (3.5488) [0.5282]	0.5202 (0.2646) [0.0522]	295.3	296.0	308.2
LBW	244.20 (0.7614)	162.98 (0.4759)	19.4117 (2.2409)	7.2415 (0.2525) [<0.0001]	0.5905 (0.4340) [0.1768]	300.0	300.7	312.8
LW	1	1	1.1800 (0.1081)	6.0387 (0.1828) [<0.0001]	0.7486 (0.2772) [0.0082]	312.6	312.8	320.3

These results from Table 6.2 indicate that the LTLGW model has the lowest AIC, CAIC and BIC values among the fitted models, so the LTLGW model provides an appropriate fit for these data.

Further, the fitted LTLGW regression model suggests that x_1 is significant at 6% and that there is a significant difference between the two-Arm clinical trial. We plot in Figure 6.2 the empirical survival function and the estimated survival functions for the LTLGW, LBW and LW models. These plots suggest that the LTLGW model provides a suitable fit.

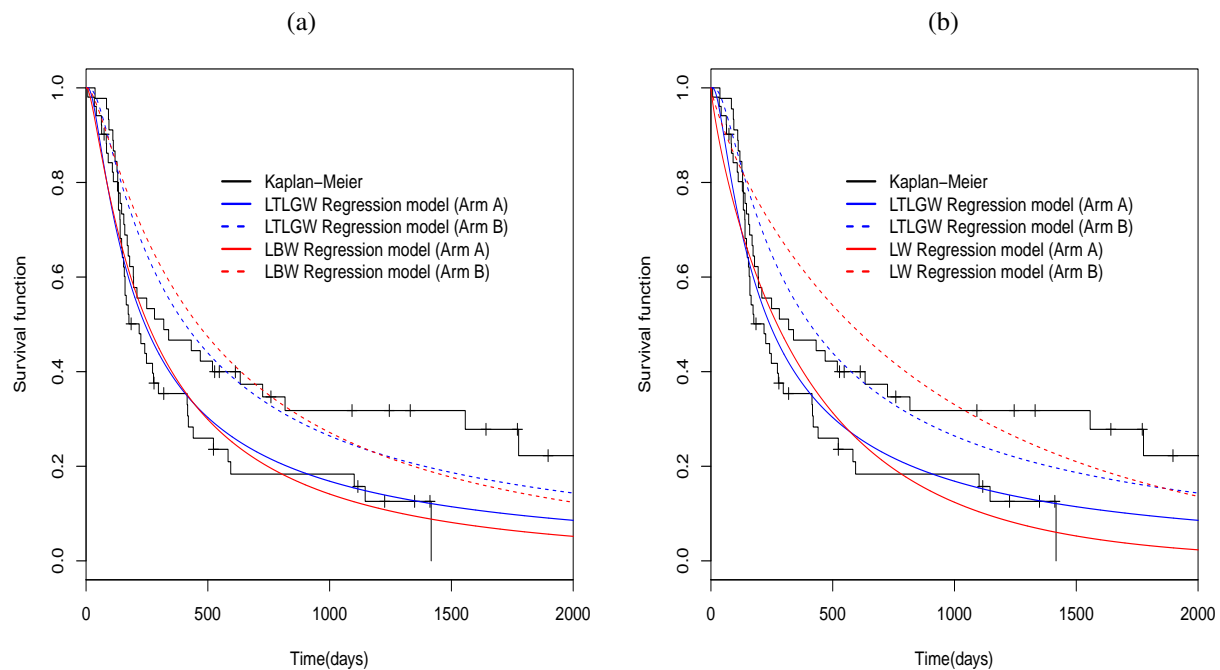


Figure 7. Estimated survival functions and the empirical survival for the Efron data. (a) LTLGW versus LBW regression model. (b) LTLGW versus LW regression model.

7. Concluding Remarks

In this paper, we propose a new four parameter model, called the *Topp-Leone Generated Weibull* (TLGW) distribution, which extends the Weibull distribution. In fact, the TLGW distribution is motivated by the extensive use of the Weibull distribution in many applied areas and also for the fact that the generalization provides more flexibility to analyze real data. The TLGW density function can be expressed as a mixture of epx-W densities. We derive explicit expressions for the ordinary and incomplete moments, moment generating function, entropies, mean residual life and mean inactivity time. We also provide some useful characterizations of the new model. We discuss the maximum likelihood estimation of the model parameters. We introduce the so-called *log-Topp-Leone Generated Weibull* (LTLGW) distribution, based on this new distribution, we propose a LTLGW regression model very suitable for modeling censored and uncensored lifetime data. Applications to real data modeling has been discussed and it has been shown that the proposed model provides consistently better fit than the other competitive models.

References

- Afify, A. Z., Cordeiro, G. M., Yousof, H. M., Saboor, A., & Ortega, E. M. M. (2016). The Marshall-Olkin additive Weibull distribution with variable shapes for the hazard rate. *Hacettepe Journal of Mathematics and Statistics*, Forthcoming.
- Alzaatreh, A., Lee, C., & Famoye, F. A. (2013). A new method for generating families of continuous distributions, *Metron* 71, 63-79. <http://dx.doi.org/10.1007/s40300-013-0007-y>
- Aryal, G. R., & Tsokos, C. P. (2011). Transmuted Weibull distribution: a generalization of the Weibull probability distribution. *European Journal of Pure and Applied Mathematics*, 4, 89-102.
- Aryal, G., & Elbatal, I. (2015). On the Exponentiated Generalized Modified Weibull Distribution. *Communications for Statistical Applications and Methods*, 22, 333-348. <http://dx.doi.org/10.5351/csam.2015.22.4.333>
- Balakrishnan, N., Leiva, V., Sanhueza, A., & Cabrera, E. (2009). Mixture inverse Gaussian distributions and its transformations, moments and applications. *Statistics*, 43, 91-104. <http://dx.doi.org/10.1080/02331880701829948>
- Chhikara, R. S., & Folks, J. L. (1989). *The inverse Gaussian distribution*, Marcel Dekker, New York.
- Cordeiro, G. M., Hashimoto, E. M., & Ortega, E. M. M. (2014). The McDonald Weibull model. *Statistics: A Journal of Theoretical and Applied Statistics*, 48, 256-278. <http://dx.doi.org/10.1080/02331888.2012.748769>
- 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.

- Cordeiro, G. M., Ortega, E. M., & Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute*, 347, 1399-1429. <http://dx.doi.org/10.1016/j.jfranklin.2010.06.010>
- Cordeiro, G. M., Ortega, E. M., & Silva, G. O. (2012). The Kumaraswamy modified Weibull distribution: theory and applications. *Journal of Statistical Computation and Simulation*, 84, 1387-1411. <http://dx.doi.org/10.1080/00949655.2012.745125>
- da Silva, R., Thiago, A., Maciel, D., Campos, R., & Cordeiro, G. (2013). A new lifetime model: the gamma extended Fréchet distribution. *Journal of Statistical Theory and Applications*, 12, 39-54. <http://dx.doi.org/10.2991/jsta.2013.12.1.4>
- Efron, B. (1988). Logistic regression, survival analysis and the Kaplan-Meier curve. *Journal of the American Statistical Association*, 83, 414-425. <http://dx.doi.org/10.2307/2288857>
- Elbatal, I., & Aryal, G. (2013). On the transmuted additive Weibull distribution. *Austrian Journal of Statistics*, 42, 117-132.
- Famoye, F., Lee, C., & Olumolade, O. (2005). The Beta-Weibull Distribution, *Journal of Statistical Theory and Applications*, 4, 121-136.
- Ghitany, M. E., Al-Hussaini, E. K., & Al-Jarallah, R. A. (2005). Marshall-Olkin extended Weibull distribution and its application to censored data. *Journal of Applied Statistics*, 32, 1025-1034. <http://dx.doi.org/10.1080/02664760500165008>
- Glänzel, W. (1987). A characterization theorem based on truncated moments and its application to some distribution families, *Mathematical Statistics and Probability Theory (Bad Tatzmannsdorf, 1986)*, Vol. B, Reidel, Dordrecht, 75-84. http://dx.doi.org/10.1007/978-94-009-3965-3_8
- Glänzel, W. (1990). Some consequences of a characterization theorem based on truncated moments, *Statistics: A Journal of Theoretical and Applied Statistics*, 21, 613-618. <http://dx.doi.org/10.1080/02331889008802273>
- Hanook, S., Shahbaz, M. Q., Mohsin, M., & Kibria, G. (2013). A Note on Beta Inverse Weibull Distribution, *Communications in Statistics - Theory and Methods*, 42, 320-335. <http://dx.doi.org/10.1080/03610926.2011.581788>
- Khan, M. S., & King, R. (2013). Transmuted modified Weibull distribution: a generalization of the modified Weibull probability distribution. *European Journal of Pure and Applied Mathematics*, 6, 66-88.
- Lai, C. D., Xie, M., & Murthy, D. N. P. (2001). Bathtub-shaped failure rate life distributions, *Chapter 3, in Advances in Reliability, vol. 20 of Handbook of Statistics*, pp. 69-704. [http://dx.doi.org/10.1016/s0169-7161\(01\)20005-4](http://dx.doi.org/10.1016/s0169-7161(01)20005-4)
- Lai, C. D., Xie, M., & Murthy, D. N. P. (2003). A modified Weibull distribution. *IEEE Transactions on Reliability*, 52, 33-37.
- Lee, C., Famoye, F., & Olumolade, O. (2007). Beta-Weibull distribution: some properties and applications to censored data. *Journal of modern applied statistical methods*, 6, 17.
- Merovci, F. and Elbatal, I. (2013). The McDonald modified Weibull distribution: properties and applications. arXiv preprint arXiv:1309.2961.
- Mudholkar, G. S., & Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-real data. *IEEE Transactions on Reliability*, 42, 299-302. <http://dx.doi.org/10.1109/24.229504>
- Mudholkar, G. S., Srivastava, D. K., & Freimer, M. (1995). The exponentiated Weibull family: a reanalysis of the bus-motor-failure data. *Technometrics*, 37, 436-445. <http://dx.doi.org/10.1080/00401706.1995.10484376>
- Mudholkar, G. S., Srivastava, D. K., & Kollia, G. D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*, 91, 1575-1583. <http://dx.doi.org/10.1080/01621459.1996.10476725>
- Nadarajah, S. (2009). Bathtub-shaped failure rate functions. *Quality and Quantity*, 43, 855- 863. <http://dx.doi.org/10.1007/s11135-007-9152-9>
- Nofal, Z. M., Afify, A. Z., Yousof, H. M., & Cordeiro, G. M. (2015). The generalized transmuted-G family of distributions. *Communications in Statistics-Theory and Methods*, Forthcoming. <http://dx.doi.org/10.1080/03610926.2015.1078478>
- Nofal, Z. M., Afify, A. Z., Yousof, H. M., & Louzada, F.(2016). Transmuted exponentiated additive Weibull distribution: properties and applications, Forthcoming.
- Nofal, Z. M., Afify, A. Z., Yousof, H. M., Granzotto, D. C. T., & Louzada, F. (2016). Kumaraswamy transmuted exponen-

- tiated additive Weibull distribution. *International Journal of Statistics and Probability*, 5, 78-99. <http://dx.doi.org/10.5539/ijsp.v5n2p78>
- Ortega, E. M. M., Cordeiro, G. M., & Hashimoto, E. M. (2011). A log-linear regression model for the Beta-Weibull distribution. *Communications in Statistics- Simulation and Computations*, 40, 1206-1235. <http://dx.doi.org/10.1080/03610918.2011.568150>
- Pinho, L. G., Cordeiro, G. M., & Nobre, J. S. (2015). The Harris extended exponential distribution. *Communications in Statistics- Theory and methods*, 44, 3486-3502. <http://dx.doi.org/10.1080/03610926.2013.851221>
- Provost, S. B., Saboor, A., & Ahmad, M. (2011). The gamma-Weibull distribution, *Pak. Journal Stat.*, 27, 111-131.
- Rezaei, S., Sadr, B. B., Alizadeh, M., & Nadarajah, S. (2016). Topp-Leone generated family of distributions: Properties and applications. *Communications in Statistics- Theory and Methods*, Forthcoming. <http://dx.doi.org/10.1080/03610926.2015.1053935>
- Sarhan, A. M., & Zaindin, M. (2009). Modified Weibull distribution, *Applied Sciences*, 11, 123-136.
- Shahbaz, M. G., Shahbaz, S., & Butt, N. M. (2102). The Kumaraswamy-inverse Weibull distribution. *Pak. Journal Stat. Oper. Res.*, 8, 479-489. <http://dx.doi.org/10.18187/pjsor.v8i3.520>
- Silva, G. O., Ortega, E. M. M., & Cordeiro, G. M. (2010). The beta modified Weibull distribution. *Lifetime Data Analysis*, 16, 409-430. <http://dx.doi.org/10.1007/s10985-010-9161-1>
- Xie, M., Tang, Y., & Goh, T. N. (2002). A modified Weibull extension with bathtub failure rate function. *Reliability Engineering and System Safety*, 76, 279-285. [http://dx.doi.org/10.1016/s0951-8320\(02\)00022-4](http://dx.doi.org/10.1016/s0951-8320(02)00022-4)
- Xie, M. & Lai, C. D. (1995). Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function. *Reliability Engineering and System Safety*, 52, 87-93. [http://dx.doi.org/10.1016/0951-8320\(95\)00149-2](http://dx.doi.org/10.1016/0951-8320(95)00149-2)
- Yousof, H. M., Afify, A. Z., Alizadeh, M., Butt, N. S., Hamedani, G. G., & Ali, M. M. (2015). The transmuted exponentiated generalized-G family of distributions. *Pak. J. Stat. Oper. Res.*, 11, 441-464. <http://dx.doi.org/10.18187/pjsor.v11i4.1164>

Appendix A

Theorem 1. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a given probability space and let $H = [d, e]$ be an interval for some $d < e$ ($d = -\infty$, $e = \infty$ might as well be allowed). Let $X : \Omega \rightarrow H$ be a continuous random variable with the distribution function F and let q_1 and q_2 be two real functions defined on H such that

$$\mathbf{E}[q_2(X) | X \geq x] = \mathbf{E}[q_1(X) | X \geq x] \xi(x), \quad x \in H,$$

is defined with some real function ξ . Assume that $q_1, q_2 \in C^1(H)$, $\xi \in C^2(H)$ and F is twice continuously differentiable and strictly monotone function on the set H . Finally, assume that the equation $\xi q_1 = q_2$ has no real solution in the interior of H . Then F is uniquely determined by the functions q_1, q_2 and ξ , particularly

$$F(x) = \int_a^x C \left| \frac{\xi'(u)}{\xi(u) q_1(u) - q_2(u)} \right| \exp(-s(u)) du,$$

where the function s is a solution of the differential equation $s' = \frac{\xi' q_1}{\xi q_1 - q_2}$ and C is the normalization constant, such that $\int_H dF = 1$.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).