Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/050203

Log-Logistic Generated Weibull Distribution: Model, Properties, Applications and Estimations Under Progressive Type-II Censoring

Alaa H. Abdel-Hamid[1,*] *and Nahla A. Albasuoni*[2]

[1] Department of Mathematics and Computer Science, Faculty of Science, Beni-Suef University, Egypt
[2] Faculty of Science, University of Tripoli, Libya

Received: 1 Jan. 2016, Revised: 2 Mar. 2016, Accepted: 3 Mar. 2016
Published online: 1 Jul. 2016

Abstract: Composition of distribution functions is considered one of the methods that generate new distributions. In this paper, a new distribution is obtained by composing a log-logistic distribution with a Weibull distribution. The new distribution has decreasing and unimodal shapes of the hazard rate function which make it suitable to fit several real data. Some properties of the new distribution are investigated. Based on progressive type-II censoring, the maximum likelihood, moments and probability weighted moments estimation methods for the involved parameters are studied and compared through a simulation study. The asymptotic confidence intervals for the parameters are also obtained based on asymptotic variance-covariance matrix. Finally, real data are used to compare the new distribution with four lifetime distributions based on five comparison criteria. The comparison shows that the new distribution is better to fit the data than the other four distributions.

Keywords: Composition of distributions, progressive type-II censoring, log-logistic and Weibull distributions, maximum likelihood estimation, moments and probability weighted moments estimations, simulation.

1 Introduction

Generation of new distributions may be of great importance when the known classical distributions are unable to give a good fit to some real data. It could be achieved by some methods such as: (i) Compounding of distributions, see for example, [1]-[6]. (ii) Adding one or more parameters to a cumulative distribution function (CDF) or a survival function (SF), for example by exponentiating the CDF to a new parameter. See, for example [7,8,9]. (iii) Composition of CDF (H) with another CDF G to obtain a CDF F, given by

$$F(x) = H(G(x)), \tag{1}$$

(iv) Composition of a symmetric probability density function (PDF) with transformation of scale. Jones [10] generated a PDF f by composing a symmetric PDF g with transformation of scale $t(x)$, in such a way that

$$f(x) = 2g[t(x)], \quad -\infty < x < \infty.$$

Baker [11] suggested two of such transformations:

$$t_1(x) = x - b/x, \quad b > 0, x > 0,$$

and

$$t_2(x) = \frac{1}{a}\ln(e^{ax} - 1), \quad a > 0, x > 0.$$

More transformations were suggested by Jones [10].

* Corresponding author e-mail: hamid_alh@science.bsu.edu.eg

(v) Composition of a CDF with a function of another CDF. This procedure could be achieved by composing H with $\overline{\eta}$ to get F or H with η to get $R_F(x)$, see [12], as follows

$$F(x) = H(\overline{\eta}(x)) \ \text{ or } \ R_F(x) = H(\eta(x)), \tag{2}$$

where $R_F = 1 - F$ is the SF of F,

$$\eta(x) = -\ln(G(x)), \ \overline{\eta}(x) = -\ln(R_G(x)). \tag{3}$$

More parameters in the generated distribution make it more flexible to analyzing data in the sense of having better fit and more shapes of hazard rate function (HRF).

If H and G are absolutely continuous CDFs, then (1) and (2) can be written as

$$F(x) = H(G(x)) = \int_0^{G(x)} h(y)\mathrm{d}y, \tag{4}$$

and

$$F(x) = H(\overline{\eta}(x)) = \int_0^{-\ln(R_G(x))} h(y)\mathrm{d}y \ \text{ or } \ R_F(x) = H(\eta(x)) = \int_0^{-\ln(G(x))} h(y)\mathrm{d}y, \tag{5}$$

where $h(.)$ is the PDF that corresponds CDF $H(.)$.

If, in (4), h subjects to beta(a,b) distribution on the interval $(0,1)$, then (4) becomes

$$F(x) = \frac{1}{\mathrm{B}(a,b)} \int_0^{G(x)} y^{a-1}(1-y)^{b-1}\mathrm{d}y. \tag{6}$$

CDF (6) is called the beta-G distribution, which was studied by several authors through specifying different forms of G. Among others, Eugene et al. [13] specified G to be normal to obtain the beta-normal distribution. Some authors such as Nadarajah and Gupta [14], Nadarajah and Kotz [15], Barrito-Souza et al. [16] and Cordeiro and Brito [17] considered the beta-Frechét, beta-exponential, beta-exponentiated exponential and beta-power distributions, respectively. AL-Hussaini and Abdel-Hamid [12] introduced a new distribution called half-logistic generated Weibull distribution (HLGWD) and studied some of its properties.

In this paper, we introduce a new distribution called log-logistic generated Weibull distribution (LLGWD) by composing log-logistic CDF (H) with Weibull CDF G. The new distribution has decreasing and unimodal failure rates which make it more suitable to fit several real data.

The rest of the paper is organized as follows: The LLGWD is derived in Section 2. Some properties of the LLGWD are investigated in Section 3. Maximum likelihood (ML), moments and probability weighted moments (PWM) estimation methods are studied in Section 4. Simulation study is worked done in Section 5. Finally, Section 6 presents some concluding remarks.

2 Formulation of the Model

The log-logistic distribution (LLD) (known as the Fisk distribution in economics) is used in survival analysis as a parametric model for events whose rate increases initially and decreases later, for example mortality rate from cancer following diagnosis or treatment. It has been used also in hydrology to model stream flow and precipitation, and in economics as a simple model of the distribution of wealth or income. The LLD is similar in shape to the log-normal distribution but has heavier tails. Unlike the log-normal, its CDF can be written in closed form.

The PDF, CDF and HRF of the LLD are given, respectively, by

$$h(x) = \frac{\gamma x^{\gamma-1}}{(1+x^\gamma)^2}, \quad x > 0, \quad (\gamma > 0), \tag{7}$$

$$H(x) = \frac{x^\gamma}{1+x^\gamma}, \quad x > 0, \quad (\gamma > 0), \tag{8}$$

$$\lambda_1(x) = \frac{\gamma x^{\gamma-1}}{1+x^\gamma}. \tag{9}$$

HRF (9) is unimodal (decreasing) when $\gamma > 1$ ($\gamma \leq 1$).

The Weibull distribution (WD) is considered one of the most popular distributions in analyzing skewed data. The PDF, CDF and HRF of the WD with scale and shape parameters α and β are given, respectively, by

$$g(x) = \alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}}, \quad x>0, \quad (\alpha,\beta>0), \tag{10}$$

$$G(x) = 1-e^{-\alpha x^{\beta}}, \quad x>0, \quad (\alpha,\beta>0), \tag{11}$$

$$\lambda_2(x) = \alpha\beta x^{\beta-1}, \quad x>0, \quad (\alpha,\beta>0). \tag{12}$$

It can be easily seen that HRF (12) is increasing (decreasing) [constant] when $\beta>1$ $(\beta<1)$ $[\beta=1]$.

By substitution from (11) in (3), then $\overline{\eta}(x) = \alpha x^{\beta}$. Using PDF (7), then F in Equation (2) becomes

$$\begin{aligned} F(x) &= \gamma\int_0^{\alpha x^{\beta}} \frac{u^{\gamma-1}}{(1+u^{\gamma})^2}\mathrm{d}u \\ &= 1-\frac{1}{1+(\alpha x^{\beta})^{\gamma}}, \quad x>0, \quad (\alpha,\beta,\gamma>0). \end{aligned} \tag{13}$$

We will call CDF (13) as the CDF of LLGWD. The closed form of CDF (13) is considered one of the advantages of LLGWD, since many distributions that arise from composition of CDFs do not have CDFs in closed forms.

The PDF corresponding to (13) is given by

$$f(x) = \frac{\alpha^{\gamma}\beta\gamma x^{\beta\gamma-1}}{[1+(\alpha x^{\beta})^{\gamma}]^2}. \tag{14}$$

From (13) and (14), the HRF and proportional reversed HRF are given, respectively, by

$$\lambda_F(x) = \frac{f(x)}{1-F(x)} = \frac{\alpha^{\gamma}\beta\gamma x^{\beta\gamma-1}}{1+(\alpha x^{\beta})^{\gamma}}, \quad x>0, \tag{15}$$

$$\lambda_F^{\star}(x) = \frac{f(x)}{F(x)} = \frac{\beta\gamma}{x(1+(\alpha x^{\beta})^{\gamma})}, \quad x>0. \tag{16}$$

Different shapes of PDF (14) and HRF (15) are plotted in Figures 1 and 2, respectively, for different values of α, β and γ.

AL-Hussaini and Hussein [18] showed that any CDF F could be written in terms of $\lambda_F(x)$ and $\lambda_F^{\star}(x)$ as follows

$$F(x) = \frac{\lambda_F(x)}{\lambda_F(x)+\lambda_F^{\star}(x)}.$$

The corresponding PDF is given by

$$f(x) = \lambda_F(x)(1-F(x)) = \frac{\lambda_F(x)\lambda_F^{\star}(x)}{\lambda_F(x)+\lambda_F^{\star}(x)}.$$

3 Properties of LLGWD

Theorem 3.1. HRF (15) is

(i) decreasing if $\beta\gamma\leq 1$,

(ii) unimodal with mode at $x_0 = \left(\frac{\beta\gamma-1}{\alpha^{\gamma}}\right)^{1/(\beta\gamma)}$ if $\beta\gamma>1$.

Proof. The first derivative of HRF (15) with respect to x is given by

$$\lambda'(x) = \frac{\alpha^{\gamma}\beta\gamma x^{\beta\gamma-2}}{(1+\alpha^{\gamma}x^{\beta\gamma})^2}[\beta\gamma-1-\alpha^{\gamma}x^{\beta\gamma}]. \tag{17}$$

It is easy to see that if $\beta\gamma \leq 1$, then $\lambda'(x) < 0$ and hence HRF (15) is decreasing. This proves (i).

Now, suppose that $\beta\gamma > 1$. Since $x > 0$, $\lambda'(x) = 0$ at unique $x \equiv x_0 = \left(\frac{\beta\gamma - 1}{\alpha^\gamma}\right)^{1/(\beta\gamma)}$. The second derivative of $\lambda(x)$ at $x = x_0$ is given by

$$\lambda''(x_0) = -\frac{(\alpha^\gamma\beta\gamma)^2 x_0^{2\beta\gamma-3}}{(1+\alpha^\gamma x_0^{\beta\gamma})^4},$$

which is always negative. Therefore, x_0 maximizes Equation (15) and hence the unimodality of (15) is achieved at mode $x_0 = \left(\frac{\beta\gamma - 1}{\alpha^\gamma}\right)^{1/(\beta\gamma)}$. This proves (ii). □.

Corollary 3.1. PDF (14) is decreasing if $\beta\gamma \leq 1$ and unimodal with mode at

$$x_0 = \left(\frac{\beta\gamma - 1}{\alpha^\gamma}\right)^{1/(\beta\gamma)} \quad \text{if} \quad \beta\gamma > 1.$$

Proof. The proof is simple and hence it is omitted. □.

Proposition 3.1. The p-th quantile of LLGWD (13) is given by

$$x_p = \left(\frac{1}{\alpha}\left[\frac{p}{1-p}\right]^{1/\gamma}\right)^{1/\beta}, \quad 0 < p < 1.$$

The median (second quartile) is achieved at $p = 1/2$.

Proof. The proof arises directly from $F(x_p) = p$, where $F(.)$ is CDF (13). □.

Proposition 3.2. If X is a random variable with PDF (14), then the r-th moment of X is given, for $r = 1, 2, \ldots,$ by

$$\mu_r' = \sum_{\zeta=1}^{\aleph} \varpi_\zeta \frac{2}{(1-z_\zeta)^2}\left(\frac{1+z_\zeta}{1-z_\zeta}\right)^r f\left(\frac{1+z_\zeta}{1-z_\zeta}\right),$$

where z_ζ, ϖ_ζ are the zeros and the corresponding Christoffel numbers of the Legendre-Gauss quadrature formula on the interval (-1, 1).

Proof.

$$\begin{aligned}\mu_r' = E(X^r) &= \int_0^\infty x^r f(x)\mathrm{d}x, \\ &= \int_{-1}^{1} \frac{2}{(1-z)^2}\left(\frac{1+z}{1-z}\right)^r f\left(\frac{1+z}{1-z}\right)\mathrm{d}z.\end{aligned} \tag{18}$$

The integral in (18) can be approximated by using Legendre-Gauss quadrature formula as

$$\mu_r' = \sum_{\zeta=1}^{\aleph} \varpi_\zeta \frac{2}{(1-z_\zeta)^2}\left(\frac{1+z_\zeta}{1-z_\zeta}\right)^r f\left(\frac{1+z_\zeta}{1-z_\zeta}\right), \tag{19}$$

where $\varpi_\zeta = \frac{2}{(1-z_\zeta^2)[\mathbb{L}'_{\aleph+1}(z_\zeta)]^2}$ and $\mathbb{L}'_{\aleph+1}(z_\zeta) = \frac{\mathrm{d}\mathbb{L}_{\aleph+1}(z)}{\mathrm{d}z}$ at $z = z_\zeta$. and $\mathbb{L}_\aleph(.)$ is the Legendre polynomial of degree $\aleph$, see [19]. □.

The skewness ($\mathscr{S}$) and kurtosis ($\mathscr{K}$) of the LLGWD can be computed as

$$\mathscr{S} = \frac{\mu_3}{\mu_2^{3/2}}, \quad \mathscr{K} = \frac{\mu_4}{\mu_2^2} - 3,$$

where μ_2, μ_3 and μ_4 denote the second, third and fourth central moments, respectively. Table 1 displays the first six moments of X in addition to the skewness and kurtosis for $\alpha = 2.0$, $\beta = 3.0$ and $\gamma = 5.0$.

Table 1: The first six moments, skewness and kurtosis for $\alpha = 2.0$, $\beta = 3.0$ and $\gamma = 5.0$.

r	1	2	3	4	5	6	$\mathscr{S}$	$\mathscr{K}$
$E(X^r)$	0.7995	0.6487	0.5344	0.4473	0.3808	0.3303	0.5989	3.1411

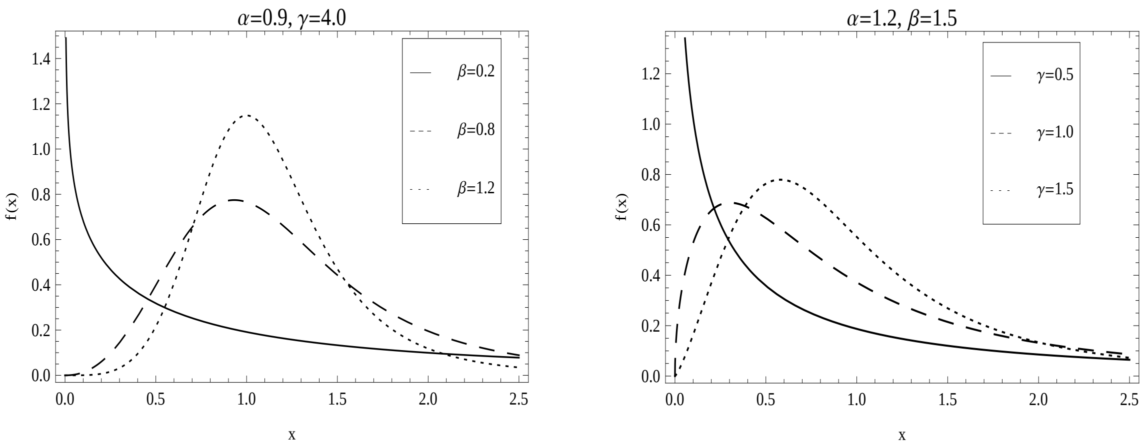

Fig. 1: Left panel: PDF of LLGWD for fixed α, γ and different values of β. Right panel: PDF LLGWD for fixed α, β and different values of γ.

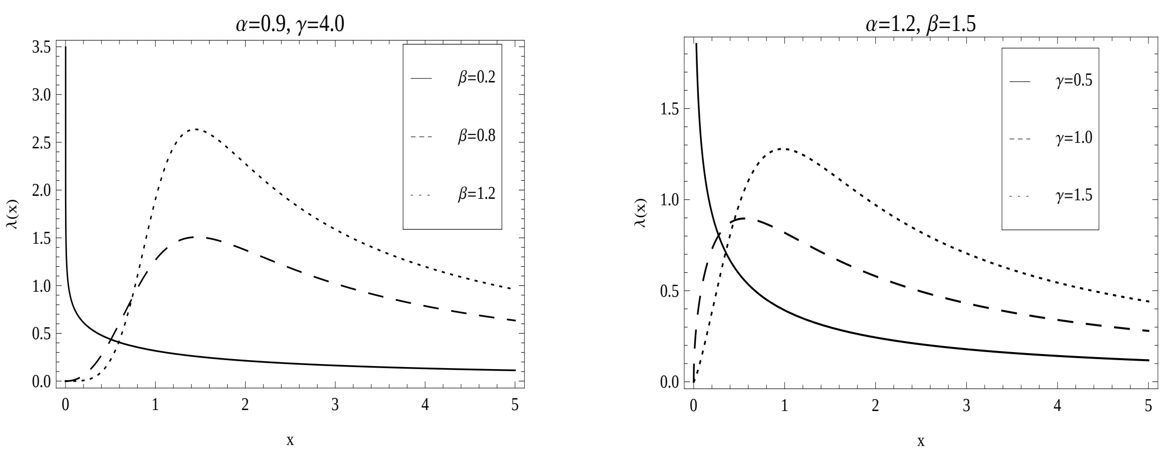

Fig. 2: Left panel: Hazard rate of LLGWD for fixed α, γ and different values of β. Right panel: Hazard rate of LLGWD for fixed α, β and different values of γ.

4 Estimation Methods Under Progressive Type-II Censoring

Censoring is considered in reliability experiments. It usually applies when the experimenter is unable to get total information on lifetimes for each unit or reducing the total test time and the associated cost. Type-I and type-II censoring schemes are most commonly used, see for example [20]. Progressive type-II censoring is considered a generalization of type-II censoring. It gives flexibility to the experimenter to remove units from a life test at several stages during the experiment. Live units removed early on can be readily used in other tests.

Progressive type-II censoring can be applied as follows: Suppose that $m(< n)$ and $R_1, R_2, \ldots, R_m$ are fixed before the experiment. R_1 surviving units are randomly removed from the test when the first failure time occurs and R_2 surviving units are randomly removed from the test when the second failure time occurs. The test continues in the same manner until the m-th failure at which all the remaining surviving units $R_m = n - m - \sum_{j=1}^{m-1} R_j$ are removed from the test, thereby terminating the life test. The data from progressively type-II censored samples are as follows: $(x_{1:m:n};R_1)$, ..., $(x_{m:m:n};R_m)$ where $x_{1:m:n} < \ldots < x_{m:m:n}$ denote the m ordered observed failure times and $R_1, \ldots, R_m$ denote the number of units removed from the experiment at failure times $x_{1:m:n}, \ldots, x_{m:m:n}$. For more details on progressive censoring, see [21].

In the following three subsections, some estimation methods are considered.

4.1 Maximum likelihood estimation

If $((x_{1:m:n};R_1), \ldots, (x_{m:m:n};R_m))$ is a progressively type-II censored random sample from a population with CDF (13) and PDF (14), then the likelihood function is given by

$$L(\alpha,\beta,\gamma;\mathbf{x}) \propto \prod_{j=1}^{m} f(x_{j:m:n})[1-F(x_{j:m:n})]^{R_j}, \tag{20}$$

where $\mathbf{x} = (x_1,\ldots,x_m)$.

Based on Equations (13) and (14), logarithm of (20), $\ell(\alpha,\beta,\gamma;\mathbf{x}) = \ln(L(\alpha,\beta,\gamma;\mathbf{x}))$, is given by

$$\ell(\alpha,\beta,\gamma;\mathbf{x}) = m\ln(\alpha^{\gamma}\beta\gamma) + (\beta\gamma-1)\sum_{j=1}^{m}\ln(x_j) - \sum_{j=1}^{m}(R_j+2)\ln\left[1+(\alpha x_j^{\beta})^{\gamma}\right]. \tag{21}$$

The maximum likelihood estimates (MLEs) $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ of α, β and γ could be obtained by solving the likelihood equations, $\frac{\partial \ell}{\partial \alpha} = 0$, $\frac{\partial \ell}{\partial \beta} = 0$ and $\frac{\partial \ell}{\partial \gamma} = 0$, with respect to α, β and γ. These MLEs can not be obtained in closed forms and hence a numerical iteration method for the likelihood equations should be used.

4.2 Approximate confidence interval

The observed Fisher information matrix, $\mathbf{F}$, for MLEs $(\hat{\alpha},\hat{\beta},\hat{\gamma})$ is the 3×3 symmetric matrix of negative second partial derivatives of $\ell(\alpha,\beta,\gamma)$ with respect to α, β and γ, see [22].

$$\mathbf{F} = \begin{pmatrix} -\frac{\partial^2\hat{\ell}}{\partial\alpha^2} & -\frac{\partial^2\hat{\ell}}{\partial\alpha\partial\beta} & -\frac{\partial^2\hat{\ell}}{\partial\alpha\partial\gamma} \\ -\frac{\partial^2\hat{\ell}}{\partial\beta\partial\alpha} & -\frac{\partial^2\hat{\ell}}{\partial\beta^2} & -\frac{\partial^2\hat{\ell}}{\partial\beta\partial\gamma} \\ -\frac{\partial^2\hat{\ell}}{\partial\gamma\partial\alpha} & -\frac{\partial^2\hat{\ell}}{\partial\gamma\partial\beta} & -\frac{\partial^2\hat{\ell}}{\partial\gamma^2} \end{pmatrix},$$

where the caret ˆ indicates that the derivative is calculated at $(\hat{\alpha},\hat{\beta},\hat{\gamma})$. The elements of the matrix $\mathbf{F}$ can be easily obtained.

The inverse of $\mathbf{F}$ is the local estimate $\mathbf{V}$ of the asymptotic variance-covariance matrix of $(\hat{\alpha},\hat{\beta},\hat{\gamma})$. That is

$$\mathbf{V} = \mathbf{F}^{-1} = \begin{pmatrix} \mathrm{var}(\hat{\alpha}) & \mathrm{cov}(\hat{\alpha},\hat{\beta}) & \mathrm{cov}(\hat{\alpha},\hat{\gamma}) \\ \mathrm{cov}(\hat{\beta},\hat{\alpha}) & \mathrm{var}(\hat{\beta}) & \mathrm{cov}(\hat{\beta},\hat{\gamma}) \\ \mathrm{cov}(\hat{\gamma},\hat{\alpha}) & \mathrm{cov}(\hat{\gamma},\hat{\beta}) & \mathrm{var}(\hat{\gamma}) \end{pmatrix}. \tag{22}$$

Following the general asymptotic theory of MLEs, the sampling distribution of $\frac{\hat{\alpha}-\alpha}{\sqrt{\mathrm{var}(\hat{\alpha})}}$, $\frac{\hat{\beta}-\beta}{\sqrt{\mathrm{var}(\hat{\beta})}}$ and $\frac{\hat{\gamma}-\gamma}{\sqrt{\mathrm{var}(\hat{\gamma})}}$ can be approximated by a standard normal distribution which is useful in constructing confidence intervals (CIs) for the unknown parameters.

A two-sided $(1-\alpha)100\%$ normal approximation CIs for the parameters α, β and γ can then be constructed as

$$\hat{\alpha} \pm z_{\alpha^\star/2}\sqrt{\mathrm{var}(\hat{\alpha})}, \quad \hat{\beta} \pm z_{\alpha^\star/2}\sqrt{\mathrm{var}(\hat{\beta})} \quad \text{and} \quad \hat{\gamma} \pm z_{\alpha^\star/2}\sqrt{\mathrm{var}(\hat{\gamma})}.$$

where $z_{\alpha^\star/2}$ is the percentile of standard normal distribution with right-tale probability of $\alpha/2$ and $\sqrt{\mathrm{var}(\hat{\alpha})}$, $\sqrt{\mathrm{var}(\hat{\beta})}$ and $\sqrt{\mathrm{var}(\hat{\gamma})}$ can be obtained from (22).

4.3 Method of moments

One of the oldest methods for estimating the parameters of several univariate continuous distributions is the method of moments (MOM). Based on progressively type-II censored sample $(x_1,\ldots,x_m)$ and according to MOM, the parameters (α,β,γ) are estimated by equating the first three sample moments with the first three population moments and then solving the resulting three equations with respect to α, β and γ to get the moment estimates $(\breve{\alpha},\breve{\beta},\breve{\gamma})$. The first three sample moments are given by

$$E(X^r)=\frac{1}{m}\sum_{i=1}^{m}x_i^r, \quad r=1,2,3, \tag{23}$$

where $m=n-\sum_{i=1}^{m}R_i$.

The first three population moments can be obtained from (19) by putting $r=1,2,3$.

In spite of the simplicity of MOM, cubing of the sample observations can increase the sampling errors in the case of heavy-tailed situations. Outliers (extreme observations) may also exist in the sample causing considerable distortion of the results.

4.4 Probability weighted moments

Greenwood et al. [23] proposed a class of moments called probability weighted moments (PWM). This class seems to be of interest as a method for estimating parameters and quantiles of distributions which can be written in inverse form. Such distributions include the Gumbel, Weibull and logistic, among others.

One of the main advantages of using PWM is that their higher order values can be accurately estimated from small samples. Also, PWM are shown to be fairly insensitive to outliers in the data, because they are linear combinations of the observed data values.

The PWM of random variable X with CDF $F(x)$ are defined as

$$M_{a,b,c}=E[X^a(F(X))^b(1-F(X))^c], \tag{24}$$

where a, b, and c are nonnegative integers and the a-th moment of X is assumed to be finite.

For many distributions, it is most useful to consider the moments

$$M_{1,0,c}\equiv\Psi_c=E[X(1-F(X))^c]. \tag{25}$$

The number of parameters of the distribution that need to be estimated governs the number of PWM to be used.

Based on (13) and (14), Equation (25) takes the form

$$\begin{aligned}\Psi_c&=\int_0^{\infty}x(1-F(x))^c f(x)\mathrm{d}x, \quad c=0,1,2\\&=\int_{-1}^{1}\frac{2(1+z)}{(1-z)^3}\left[1-F\left(\frac{1+z}{1-z}\right)\right]^c f\left(\frac{1+z}{1-z}\right)\mathrm{d}z\\&=\sum_{\zeta=1}^{\mathbb{N}}\varpi_\zeta\frac{2(1+z_\zeta)}{(1-z_\zeta)^3}\left[1-F\left(\frac{1+z_\zeta}{1-z_\zeta}\right)\right]^c f\left(\frac{1+z_\zeta}{1-z_\zeta}\right),\end{aligned} \tag{26}$$

where z_ζ,ϖ_ζ are as defined in Proposition 3.2.

The quantities Ψ_0, Ψ_1 and Ψ_2 are then replaced by suitable estimates denoted by

$$\varphi_c=\frac{1}{m}\sum_{j=1}^{m}x_{j:m}(1-\breve{F}_{j:m})^c, \quad c=0,1,2,$$

where $\breve{F}_{j:m}$ is the empirical CDF which can be written as, see [24],

$$\breve{F}_{j:m}=1-\prod_{i=1}^{j}(1-\widehat{p}_i), \quad j=1,\ldots,m,$$

where

$$\widehat{p}_i=\frac{1}{n-\left[\sum_{k=2}^{i}R_{k-1}\right]-i+1}, \quad i=1,\ldots,m,$$

where $\sum_{k=2}^{i}R_{k-1}$ is equal zero if $k>i$.

The estimates $(\tilde{\alpha},\tilde{\beta},\tilde{\gamma})$ due to PWM can be obtained by solving numerically the three equations included in (26) (after replacing Ψ_c by φ_c, $c=0,1,2$) with respect to α, β and γ.

5 Simulation Study

In this section, the MLEs, moments and PWM estimates of the considered parameters are determined through a Monte Carlo simulation study. The performance of the estimates is investigated through the mean squared errors (MSEs) and relative absolute biases (RABs). The following steps may be applied to generate progressively type-II censored samples from CDF (13) and calculate the estimates of the parameters:

1. For given values of α, β, γ, n and m ($1<m<n$), generate a random sample of size n from Uniform(0,1) distribution, say $(u_1, \ldots, u_n)$.
2. Apply the algorithm presented in [25] to Step 1 to generate progressively type-II censored random sample of size m, $(u_1^\star, \ldots, u_m^\star)$.
3. Generate progressively type-II censored random sample $(x_{1:m:n}, \ldots, x_{m:m:n})$ from CDF (13) where, for $i = 1, \ldots, m$,

$$x_{i:m:n} \equiv x_i = \left(\frac{\alpha^{-\gamma}}{1-u_i^\star} \right)^{1/(\beta\gamma)}.$$

4. The MLEs, moments and PWM estimates of the parameters α, β and γ are computed as shown in Section 4.
5. Repeat the above steps $N(= 1000)$ times.
6. Calculate the averages of estimates, MSEs and RABs of $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ over the N samples as follows:

$$\bar{\hat{\alpha}} = \frac{1}{N}\sum_{i=1}^{N} \hat{\alpha}_i, \quad \bar{\hat{\beta}} = \frac{1}{N}\sum_{i=1}^{N} \hat{\beta}_i, \quad \bar{\hat{\gamma}} = \frac{1}{N}\sum_{i=1}^{N} \hat{\gamma}_i,$$

$$\mathrm{MSE}(\hat{\alpha}) = \frac{1}{N}\sum_{i=1}^{N} (\hat{\alpha}_i - \alpha)^2, \quad \mathrm{MSE}(\hat{\beta}) = \frac{1}{N}\sum_{i=1}^{N} (\hat{\beta}_i - \beta)^2, \quad \mathrm{MSE}(\hat{\gamma}) = \frac{1}{N}\sum_{i=1}^{N} (\hat{\gamma}_i - \gamma)^2,$$

$$\mathrm{RAB}(\hat{\alpha}) = \frac{|\bar{\hat{\alpha}} - \alpha|}{\alpha}, \quad \mathrm{RAB}(\hat{\beta}) = \frac{|\bar{\hat{\beta}} - \beta|}{\beta}, \quad \mathrm{RAB}(\hat{\gamma}) = \frac{|\bar{\hat{\gamma}} - \gamma|}{\gamma}.$$

7. Calculate the CIs of the parameters and then calculate the average interval lengths (AILs) of them. Calculate also the coverage probabilities (COVPs) of the parameters α, β and γ.

   The following three CSs are applied in the generation of the samples:

– CS1:

$$\begin{aligned} R_i &= n - m, & i &= 1, \\ R_i &= 0, & &\text{otherwise}, \end{aligned}$$

which means that we remove $n - m$ units after the first observed failure.

– CS2:

$$\begin{aligned} R_i &= 1, & i &= 1, \ldots, n-m, \\ R_i &= 0, & &\text{otherwise}, \end{aligned}$$

which means that we remove one unit after each observed failure of the first $n - m$ failures.

– CS3:

$$\begin{aligned} R_i &= n - m, & i &= \frac{m}{2}, \\ R_i &= 0, & &\text{otherwise}, \end{aligned}$$

which means that we remove $n - m$ units after the middle observed failure.

Through the simulation procedure, the values of n have been taken to be 25, 50 and 100 while the sizes of progressively type-II censored samples, m, have been chosen to represent 80% of the sample size in addition to the complete sample case.

Table 2: MLEs for (α,β,γ) with their MSEs and RABs in addition to the COVPs (in 100%) and AILs of 95% CIs based on 1000 repetitions. Population parameter values: $\alpha = 3.50$, $\beta = 3.50$, $\gamma = 1.50$.

n	m	CS	$\hat{\alpha}$ $\hat{\beta}$ $\hat{\gamma}$	MSE($\hat{\alpha}$) MSE($\hat{\beta}$) MSE($\hat{\gamma}$)	RAB($\hat{\alpha}$) RAB($\hat{\beta}$) RAB($\hat{\gamma}$)	AIL(α) AIL(β) AIL(γ)	COVP(α) COVP(β) COVP(γ)
25	20	1	3.6250	1.2645	0.0643	8.4196	99.9
			3.4237	0.3930	0.0353	6.7464	97.7
			1.4303	0.0593	0.0122	2.9922	95.3
		2	3.7579	1.2893	0.0737	8.4730	99.9
			3.6445	0.6210	0.0413	6.7266	98.4
			1.4077	0.0726	0.0135	2.8863	95.4
		3	3.6698	1.3522	0.0771	8.7710	99.9
			3.4779	0.6263	0.0508	7.0306	99.1
			1.2861	0.0609	0.0173	2.9622	96.2
	25		3.6041	1.2037	0.0569	7.8519	99.6
			3.3006	0.3037	0.0252	6.7784	98.4
			1.2609	0.0391	0.0096	3.6405	95.7
50	40	1	3.5584	0.5225	0.0167	9.6935	99.9
			3.4879	0.3394	0.0034	8.2540	99.9
			1.3139	0.0814	0.0493	3.8042	99.7
		2	3.5878	0.4829	0.0249	9.8345	99.9
			3.5114	0.3266	0.0032	8.3003	99.9
			1.4046	0.0515	0.0344	3.7300	99.4
		3	3.6370	0.6537	0.0392	10.0769	99.8
			3.5020	0.4930	0.0206	8.4814	99.9
			1.4111	0.0785	0.0441	3.7133	98.9
	50		3.4065	0.4315	0.0097	7.3721	99.6
			3.2573	0.2539	0.0015	6.5397	98.0
			1.2001	0.0248	0.0207	2.4601	95.9
100	80	1	3.4054	0.4924	0.0556	10.7545	99.7
			3.1924	0.2882	0.0309	9.8145	99.9
			1.3692	0.0952	0.0260	4.5176	99.7
		2	3.3568	0.4677	0.0523	10.6224	99.4
			3.4346	0.3786	0.0187	9.8127	99.9
			1.3014	0.0777	0.0143	4.4176	99.7
		3	3.3250	0.3809	0.0541	11.0452	99.7
			3.3672	0.1988	0.0379	9.9622	99.9
			1.2291	0.0647	0.0327	4.6246	99.9
	100		3.2416	0.2955	0.0307	7.0366	99.2
			3.0147	0.1863	0.0205	6.2462	95.0
			1.1563	0.0211	0.0158	2.2089	91.6

5.1 Simulation results

A Monte Carlo simulation study is carried out, based on 1000 simulations, in order to calculate the MLEs, moments and PWM estimates. The performance of the estimates is studied using the MSEs and RABs. The COVPs and AILs of 95% CIs for the parameters (α,β,γ) are also calculated based on different sample sizes, censored sample sizes and three different CSs. Table 2 displays the MLEs with their MSEs and RABs of the parameters (α,β,γ) in addition to the COVPs and AILs of 95% CIs of them. Table 3 displays the estimates of (α,β,γ) with their MSEs and RABs using the MOM and PWM method for $n = 25,50,100$. The population parameter values have been taken to be $\alpha = 3.50$, $\beta = 3.50$ and $\gamma = 1.50$.

From the numerical results presented in Tables 2 and 3 we can observe the following:

1. For fixed values of n, by increasing m the MSEs, RABs and AILs decrease.
2. By increasing n the MSEs, RABs and AILs decrease.
3. The estimates using the PWM method are better than those using MOM through the MSEs and RABs, while the MLEs are better than those using MOM and PWM method.

Table 3: Average moments and PWM estimates of (α,β,γ) with their MSEs and RABs for different sample sizes and CSs, based on 1000 repetitions. Population parameter values: $\alpha=3.50$, $\beta=3.50$, $\gamma=1.50$.

			MOM			PWM		
			$\bar{\breve{\alpha}}$	MSE($\breve{\alpha}$)	RAB($\breve{\alpha}$)	$\bar{\tilde{\alpha}}$	MSE($\tilde{\alpha}$)	RAB($\tilde{\alpha}$)
			$\bar{\breve{\beta}}$	MSE($\breve{\beta}$)	RAB($\breve{\beta}$)	$\bar{\tilde{\beta}}$	MSE($\tilde{\beta}$)	RAB($\tilde{\beta}$)
n	m	CS	$\bar{\breve{\gamma}}$	MSE($\breve{\gamma}$)	RAB($\breve{\gamma}$)	$\bar{\tilde{\gamma}}$	MSE($\tilde{\gamma}$)	RAB($\tilde{\gamma}$)
25	20	1	4.6940	5.3364	0.4503	4.1148	2.4172	0.2654
			3.6778	1.4944	0.1951	3.5043	0.7242	0.1676
			1.4325	0.1676	0.1778	1.7388	0.2088	0.1407
		2	4.7737	4.8333	0.4382	4.3614	2.4438	0.3003
			3.6959	1.4614	0.1717	3.5749	0.6425	0.1378
			1.7559	0.1424	0.1569	1.5162	0.1340	0.1158
		3	4.5692	3.9808	0.3547	4.4986	2.6782	0.3384
			3.5674	1.4537	0.1464	3.5663	0.7183	0.1218
			1.4951	0.1434	0.1153	1.3946	0.1260	0.1048
	25		4.3508	3.1769	0.2692	4.0639	1.5613	0.2419
			3.5277	1.4418	0.1756	3.4360	0.3796	0.1537
			1.6021	0.1505	0.1434	1.4563	0.1395	0.1229
50	40	1	4.4039	2.4657	0.3563	3.6198	1.6926	0.1142
			4.5209	1.7975	0.3166	3.6184	0.3712	0.1947
			2.3005	0.8423	0.5470	1.4724	0.0910	0.1562
		2	4.3224	2.4815	0.3431	3.7608	0.5149	0.1452
			4.4955	1.7761	0.3151	3.7273	0.4524	0.1745
			2.2909	0.8279	0.5544	1.4221	0.0622	0.1443
		3	4.3568	1.7087	0.2898	3.7709	0.9641	0.1221
			4.4670	1.4479	0.2908	3.5019	0.5310	0.1646
			2.3340	0.8255	0.5972	1.4301	0.1504	0.1302
	50		4.1198	2.3742	0.2275	3.8203	1.2511	0.1068
			3.4921	1.2465	0.1476	3.3178	0.2936	0.1314
			1.3838	0.0911	0.1207	1.3543	0.0714	0.1151
100	80	1	4.7274	4.3176	0.4293	3.8428	1.7134	0.2061
			4.0379	1.2236	0.2084	3.7482	0.7677	0.1578
			1.5909	0.1515	0.1692	1.4358	0.1261	0.1472
		2	4.9361	4.6677	0.4604	3.7231	1.8212	0.1546
			4.1889	1.1906	0.2347	3.6341	0.7171	0.1067
			1.5315	0.1041	0.1488	1.3127	0.0868	0.1210
		3	4.6378	2.8811	0.3729	3.6831	1.6437	0.1896
			3.9752	1.3729	0.1990	3.6452	0.4801	0.1343
			1.6094	0.5729	0.1772	1.2437	0.4081	0.1554
	100		4.0564	2.0787	0.2067	3.6193	1.1607	0.0992
			3.3988	1.0268	0.1348	3.2934	0.2550	0.1198
			1.2456	0.0833	0.1078	1.2064	0.0507	0.0933

5.2 Application of LLGWD to a real data set

Consider deep groove ball bearings. The ball bearings are a type of rolling-element bearings which use balls to maintain the separation between the moving parts of the bearings with the purpose of reducing rational friction and supporting radial and axial loads. The number of revolutions in a ball bearing endurance test is usually measured before failure of the ball bearings, see [20].

As indicated in Lawless ([20], p. 98), the following data arise in tests on the endurance of deep groove ball bearings. The observations are the number of million revolutions before failure for each one of 23 ball bearings. The 23 failure times are
17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.40, 51.84, 51.96, 54.12, 55.56, 67.80, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.40.

We compare the fit of the LLGWD with the WD, LLD, Burr-XII (BURR), and half-logistic generated Weibull distribution (HLGWD) due to AL-Hussaini and Abdel-Hamid [12]. For each distribution, the unknown parameters are estimated using the ML method. The validity of these distributions to fit the above data is checked using

Table 4: Comparison among LLGWD, WD, LLD, BURR, and HLGWD based on five comparison criteria.

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	K-S	P-value	AIC	CAIC	BIC
LLGWD(α,β,γ)	0.00061	1.77642	0.85139	0.21931	0.21857	248.361	249.621	251.768
WD(α,β)	0.06066	0.63808	–	0.31798	0.01910	264.279	264.879	266.550
LLD(γ)	–	–	0.37062	0.70088	0.00000	327.486	327.676	328.622
BURR(α,β)	0.69821	0.33855	–	0.47532	0.00006	309.911	310.511	312.182
HLGWD(α,β,γ)	0.06948	0.64309	1.54669	0.37546	0.00305	262.750	264.250	266.157

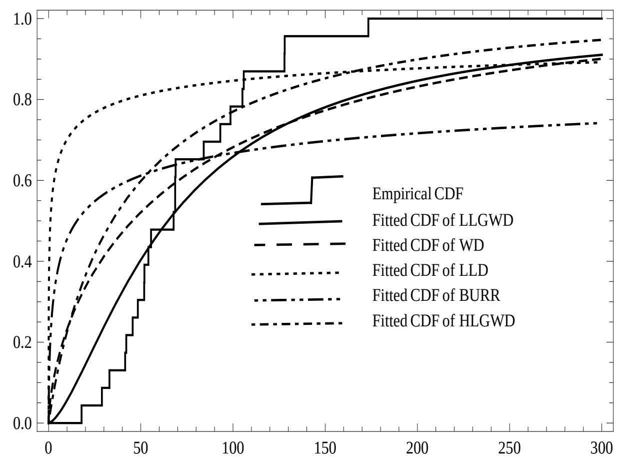

Fig. 3: Empirical CDF versus CDFs of LLGWD, WD, LLD, BURR and HLGWD.

Kolmogorov-Smirnov (K-S) statistic and the corresponding P-value. The model selection is carried out by inspection of the Akaike information criterion (AIC), $2q-2\ln(L(\hat{\mathfrak{R}}))$, consistent AIC (CAIC), $[2qn/(n-q-1)]-2\ln(L(\hat{\mathfrak{R}}))$ and Bayesian information criterion (BIC), $q\ln(n)-2\ln(L(\hat{\mathfrak{R}}))$. where $\hat{\mathfrak{R}}$ is the vector of parameter values that maximizes likelihood function L, q is the number of parameters in the model and n is the sample size. The best choice for a model is the one with smallest criterion. This is done graphically by plotting the empirical distribution versus the fitted CDFs of LLGWD, WD, LLD, BURR and HLGWD, see Figure 3.

Table 4 lists the MLEs of the parameters for the fitted LLGWD, WD, LLD, BURR and HLGWD in addition to the values of the following statistics: AIC, CAIC, BIC and K-S statistics. The P-value that corresponds the K-S statistic is also included in Table 4. These results indicate that the LLGWD represents a better fit to the data than the other four distributions through the lowest value of the K-S statistic and largest value of the corresponding P-value. It can also be noted that the LLGWD has the lowest values for the AIC, BAIC and BIC criteria among the fitted distributions. Therefore, it could be chosen as the best distribution.

6 Concluding Remarks

Generation of a new distribution is needed if the new distribution is more flexible to analyzing data in the sense of having better fit, more shapes of HRFs, etc. One of the methods that may be used to generate new distributions is the composition of a CDF with another cumulative distribution or function of such distribution. This technique can add at least an extra parameter to the distribution and hence makes it more flexible to fitting data.

In this paper, we have generated a new distribution (LLGWD) by composition of LLD with WD and obtained the conditions under which the HRF is decreasing or unimodal. Among the advantages of the LLGWD are (i) the CDF is obtained in closed form and (ii) the monotonic and non-monotonic shapes of the HRF make it more flexible to fitting data. Moreover, real data have been used to compare the LLGWD with WD, LLD, BURR and HLGWD through K-S test, P-value, AIC, CAIC and BIC. The comparison has showed that LLGWD is better to fit the data than the other four distributions. Based on a simulation study, the ML, moments and PWM estimation methods have been performed to estimate the considered parameters under progressive type-II censoring.

The simulation results have showed that the estimates using PWM are better than those using MOM while the MLEs are better than those using PWM and MOM.

Acknowledgement

The authors would like to thank the editor and referee for their helpful comments and corrections which led to improvements of an earlier version of this article.
They also dedicate the article to the spirit of late Professor Essam K. AL-Hussaini who passed away in Aug., 2015.

References

[1] F. Louzada-Neto, V.G. Cancho, G.D.C Barrigac, Journal of Applied Statistics **38**, 1239-1248 (2011).
[2] F. Louzada-Neto, V. Marchi, J. Carpenter (2013), J. Probab. Statist., Article ID 502159, 12 pages (2013).
[3] E. Mahmoudi, A. Sepahdar, Mathematics and Computers in Simulation **92**, 76-97 (2013).
[4] V.L.D. Tomazella, V.G. Cancho, F. Louzada, Chilean Journal of Statistics **4**, 99-113 (2013).
[5] S.K. Singh, U. Singh, M. Kumar, Journal of Data Science **12**, 157-173 (2014).
[6] B. Al-Zahrani, H. Sagor, Revista Colombiana de Estadística **37**, 223-243 (2014).
[7] R.D. Gupta, D. Kundu, J. Statist. Plann. Inference **137** (11), 3537-3547 (2007).
[8] A.H. Abdel-Hamid, E.K. AL-Hussaini, Comput. Statist. Data Anal. **53**, 1328-1338 (2009).
[9] E.K. AL-Hussaini, M. Ahsanullah, Exponentiated Distributions. Atlantis Press, Paris, France, 2015.
[10] M.C. Jones, Statist. Sinica **24**, 749-771 (2014).
[11] R. Baker, Commun. Statist. Th. Meth. **37**, 2162-2176 (2008).
[12] E.K. AL-Hussaini, A.H. Abdel-Hamid, Commun. Statist. Th. and Meth. (2016). (To appear).
[13] N. Eugene, C. Lee, F. Famoye, Commun. Statist. Th. Meth. **31**, 497-512 (2002).
[14] S. Nadarajah, A.K. Gupta, Commun. Statist. Th. Meth. **34**, 253-256 (2005).
[15] S. Nadarajah, S. Kotz, Rel. Engin. System Saf. **91**, 689-697 (2006).
[16] W. Barreto-Souza, A.H.S. Santos, G.M. Cordeiro, J. Statist. Comput. Simul. **80**, 159-172 (2010).
[17] G.M. Cordeiro, R. Brito, Braz. J. Prob. Statist. **26**, 88-112 (2012).
[18] E.K. AL-Hussaini, M. Hussein, Open J. Statist. **2**, 28-38 (2012).
[19] C. Canuto, M.Y. Hussaini, A. Quarteroni, T.A. Zang, Spectral Methods: Fundamentals in Single Domains. Springer-Verlag, New York, 2006.
[20] J.F. Lawless, Statistical Models and Methods for Lifetime Data. Second edition. Wiley, New York, 2003.
[21] N. Balakrishnan, Test **16**, 211-296 (2007).
[22] W. Nelson, Accelerated Testing: Statistical Models, Test Plans and Data Analysis. Wiley, New York, 1990.
[23] J.A. Greenwood, J.M. Landwehr, N.C. Matalas, J.R. Wallis, Water Resources Research **15**, 1049-1054 (1979).
[24] W.Q. Meeker, L.A. Escobar, Statistical Methods for Reliability Data. Wiley, New York, 1998.
[25] N. Balakrishnan, R.A. Sandhu, Amer. Statist. **49**, 229-230 (1995).

Alaa H. Abdel-Hamid received the PhD degree in Mathematical Statistics from Cairo university, Egypt, in 2004 and obtained the Assistant Professor degree from Beni-Suef university in 2010. His research interests are in the areas of statistical inference (Bayesian estimation and prediction), accelerated life tests, mixture of distribution functions, compounding of distributions and progressive censoring. He has published research articles in reputed international journals of mathematical statistics. He acts as a referee for some statistical journals.

Nahla A. Albasuoni got the B.Sc. in 1989 from department of statistics Faculty of Science, University of Tripoli, Libya. She prepares now for Master degree in mathematical statistics.