The Proportional Hazard Generalized Power Weibull Distribution: Properties, Applications and Regression Model

Abstract

We introduce the proportional hazard generalized power Weibull (PHGPW) model in which hazard rate function can assume increasing, decreasing, unimodal or (upside bathtub) and constant. Some of its mathematical properties are studied including the power series for the quantile function. Monte Carlo simulation was performed to determine the finite sample behaviour of the maximum likelihood estimates of the parameters. The flexibility of the PHGPW distribution compared with some other existing distributions is proved empirically by means of two sets of real data related to remission times of bladder cancer patients and strike duration of manufacturing company. A new regression model was defined based on the PHGPW distribution. The performance of the regression model is proved empirically using real data set.

Keywords: Generalized power Weibull, regression, estimation, proportional hazard.

1 Introduction

Of late, there has been many discoveries of new distributions by researchers. The need to develop these new distributions comes as a result of theoretical or practical applications to lifetime data analysis or both. The flexibility of some distributions in modeling some applications and data in practical terms is greatly improved by the extension of these distributions. Several moves have been made towards the generalization of some well-known and existing distributions and their successful applications to problems in survival data analysis, insurance etc.

In modeling monotonic hazard rates, it is realistic to use exponential, Weibull, lognormal, the generalized gamma etc. However, for non-monotonic hazard rates which usually exhibits different forms of shapes such as bathtub shape or upside bathtub, the above-mentioned distributions are not realistic to model this.

In recent times, there has many extensions of existing distributions by researchers. The aim is to improve the flexibility of these distributions in modeling data with different forms of failure rates. Lai (2014) studied the generalized Weibull distribution which used the Weibull as a baseline distribution. To improve the flexibility of the Weibull distribution, Sarhan and Zaindin (2009) proposed and studied the modified Weibull distribution and realized that it can model increasing, decreasing and constant hazard rate functions over the Weibull distribution. Gusmao et al. (2009) also studied the generalized inverse Weibull (GIW) distribution and indicated that it was more flexible than the inverse Weibull distribution and could model data with decreasing, increasing and unimodal hazard rate shapes. Nikulin and Haghighi (2006) added an additional shape parameter to the Weibull distribution and named it the generalized power Weibull (GPW) distribution. Oguntunde et al. (2015) in their bid to improve the flexibility of the Weibull distribution, studied a-four parameter distribution known as exponentiated generalized Weibull (EGW) distribution. The EGW distribution proved suitable for modeling real life events with inverted bathtub shapes.

The generalized power Weibull distribution has also seen some extension in literature. Notable among them include: Selim and Badr (2016) proposed the Kumaraswamy generalized power Weibull distribution, Pu et al. (2016) also proposed and studied generalized class of exponentiated modified Weibull distribution with application, Selim (2018) proposed the generalized power generalized Weibull distribution and Broderick et al. (2020) proposed the exponentiated generalized power series family of distributions and studied their theory, properties and applications.

Also, the proportional hazard models have attracted the attention of researchers. Martinez-Florez et al. (2013) studied the properties and inference for the proportional hazard model, Aboukhamseen et al. (2016) also proposed and studied the proportional hazard inverse Weibull distribution and Moreno-Arenas et al. (2016) presented a paper on proportional Birnbaum-Saunders distribution.

In this paper, we proposed a new four-parameter generalization of the generalized power Weibull (GPW) distribution called the proportional hazard generalized power Weibull (PHGPW) distribution. Considering an arbitrary baseline cumulative distribution function (CDF), the CDF of the GPW distribution is defined as:

$$G_{gpw}(x)=1-\exp\left\{1-\left(1+\left(\frac{x}{\lambda}\right)^{\beta}\right)^{\gamma}\right\}, \beta>0, \gamma>0, \lambda>0 \text{ and } x>0, \qquad (1)$$

and the Corresponding Probability Density Function (PDF) defined as:

$$g_{gpw}(x)=\frac{\gamma\beta}{\lambda^{\beta}}x^{\beta-1}\left(1+\left(\frac{x}{\lambda}\right)^{\beta}\right)^{\gamma-1}\exp\left\{1-\left(1+\left(\frac{x}{\lambda}\right)^{\beta}\right)^{\gamma}\right\}, x>0. \qquad (2)$$

The paper is organized as follow: introduction is found in section 1. The model description is presented in section 2 while some properties of the new distribution are discussed in section 3. Section 4 describes five methods of parameter estimation. In section 5, log-PHGPW location-scale regression model is presented while in section 6, a simulation study is performed to compare the performance of these methods of estimation for the new distribution. We demonstrate the application of this new distribution by means of two real data sets in section 7. The application of the LPHGPW regression is presented in section 8 while section 9 provides some concluding remarks.

2 Proportional Hazard Generalized Power Weibull Distribution

Suppose X is a random variable that follows the generalized power Weibull distribution with parameters $\beta, \gamma, \lambda > 0$. Then the CDF and PDF related to X are respectively expressed in equation (1) and (2).

Therefore, a random variable X is said to possess the PHGPW distribution if its CDF is expressed as:

$$G(x)=1-e^{\alpha\left(1-\left(1+\left(\frac{x}{\lambda}\right)^{\beta}\right)^{\gamma}\right)} \qquad (3)$$

where $\alpha > 0$ is a positive integer.

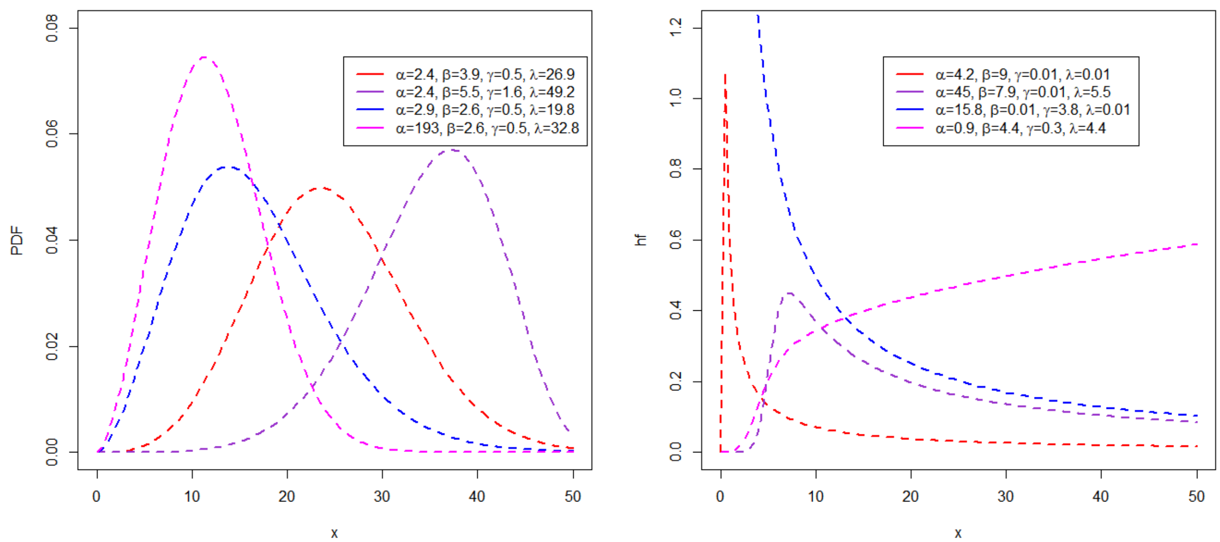
Its corresponding PDF which is given as $g(x)=\alpha f(x)\left[1-F(x)\right]^{\alpha-1}$ can be obtained as:

$$g(x)=\frac{\alpha\beta\gamma}{\lambda^{\beta}}x^{\beta-1}\left(1+\left(\frac{x}{\lambda}\right)^{\beta}\right)^{\gamma-1}e^{\alpha\left(1-\left(1+\left(\frac{x}{\lambda}\right)^{\beta}\right)^{\gamma}\right)} \qquad (4)$$

for $\alpha,\beta,\gamma,\lambda>0$ and $x>0$. The hazard rate function is expressed as $h(x)=g(x)/1-G(x)$.

Hence,

$$h(x)=\frac{\alpha\beta\gamma}{\lambda^{\beta}}x^{\beta-1}\left(1+\left(\frac{x}{\lambda}\right)^{\beta}\right)^{\gamma-1},\ x>0. \qquad (5)$$



(a) (b)

Figure 1: Density and hazard function of the PHGPW distribution

Figures 1(a) and (b) shows the behaviour of the density and hazard rate functions of the PHGPW distribution respectively for some parameter values. Clearly it can be seen from Figure 1(a) that the PDF can be uni-modal, approximately normal, positively and negatively skewed. Figure 1(b) indicates that the hazard rate function of the PHGPW is very flexible. It can have the following shapes; increasing, decreasing, unimodal or (upside bathtub) and constant.

Special cases of the PHGPW distribution are shown in Table 1. The new distribution is quite flexible as it contains several well-known lifetime distributions as special models depending the values of the parameters.

Table 1: Special sub-models of the PHGPW distribution.

Distribution	**Parameter**				**Author**
	α	β	γ	λ	
Proportional hazard Nadarajah-Haghighi (PHNH)	α	1	γ	λ	
Nadarajah-Haghighi (NH)	1	1	γ	λ	Nadaraja and Haghighi (2011)
PH exponential	α	1	1	λ	
Exponential	1	1	1	λ	Gupta and Wu (2010)
PH Weibull	α	β	1	λ	
Weibull	1	β	1	λ	Weibull (1951)
Generalized power Weibull (GPW)	1	β	γ	λ	Nikulin and Haghighi (2006)
PH Rayleigh	α	2	1	λ	
Rayleigh	1	2	1	λ	Rayleigh (1880)

3 Properties of the PHGPW Distribution

In this section, some mathematical properties of the new distribution are presented. These include the quantiles, moments, moment generating function, conditional moment, stochastic ordering and order statistics.

3.1 Quantile Function

The quantile function which is also known as the inverse CDF of a random variable is very useful when generating random numbers from a given probability distribution. It can also be used to describe some properties of a distribution such as the median, kurtosis and skewness.

The quantile function $Q_x(u)$ of the PHGPW is given by

$$Q_x(u) = \lambda\left[\left(1-\tfrac{1}{\alpha}\log(1-u)\right)^{1/\gamma} - 1\right]^{1/\beta}, \quad u \in [0,1]. \qquad (6)$$

3.2 Moments

The moments of a random variable are important in making statistical inference. They are used to study essential characteristics of a distribution such as the measures of central tendency, measures of dispersion and measures of shapes. The r^{th} non-central moment of the PHGPW random variable is derived here.

Proposition. If $X \sim$ PHGPW$(x;\alpha,\beta,\lambda,\gamma)$ then, the r^{th} non-central moment of X can be written as:

$$\mu_r^{'} = \lambda^r e^{\alpha} \sum_{i=0}^{\infty} \frac{(-1)^i}{\alpha^{\frac{(r-\beta i)}{r\beta}}} \binom{r/\beta}{i} \Gamma\left(\frac{r-\beta(i-\gamma)}{\gamma\beta}, \alpha\right), \text{ r =1,2, ...} \qquad (7)$$

where $\Gamma(a,x) = \int_x^{\infty} z^{\alpha-1} e^{-z} dz$ is the upper incomplete gamma function.

Proof. By definition, the r^{th} non-central moment of a continuous random variable X with an interval $(0,\infty)$ is defined as

$$\mu_r^{'} = E\left(X^r\right) = \int_0^{\infty} x^r f(x) dx.$$

Thus, substituting the expanded form of the density function into this definition results

$$\mu_r^{'} = \frac{\alpha\beta\gamma}{\lambda^{\beta}}\int_0^{\infty} x^r.x^{\beta-1}\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma-1} e^{\alpha\left(1-\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma}\right)}dx,$$

$$\mu_r^{'} = \lambda^r e^{\alpha}\sum_{i=0}^{\infty}\frac{(-1)^i}{\alpha^{\frac{(r-\beta i)}{r\beta}}}\binom{r/\beta}{i}\Gamma\left(\frac{r-\beta(i-\gamma)}{\gamma\beta},\alpha\right), \text{ r} =1,2, \ldots$$

The proof is completed.

3.3 Moment Generating Function

The moment generating function (MGF) of the PHGPW distribution is given by the following proposition.

Proposition. If $X \sim PHGPW(x;\alpha,\beta,\gamma,\lambda)$, then the MGF is given by

$$M_X(t) = e^{\alpha}\sum_{r=0}^{\infty}\sum_{i=0}^{\infty}\frac{(-1)^i\lambda^r t^r}{r!\alpha^{\frac{(r-\beta i)}{\gamma\beta}}}\binom{r/\beta}{i}\Gamma\left(\frac{r-\beta(i-\gamma)}{\gamma\beta},\alpha\right). \quad (8)$$

Proof. By definition, the moment generating function of the PHGPW function with a random variable X is given by $M_x(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} f(x)dx$. Substituting the expanded form of the PDF of the PHGPW into the definition above, yields

$$M_X(t) = e^{\alpha}\sum_{r=0}^{\infty}\sum_{i=0}^{\infty}\frac{(-1)^i\lambda^r t^r}{r!\alpha^{\frac{(r-\beta i)}{\gamma\beta}}}\binom{r/\beta}{i}\Gamma\left(\frac{r-\beta(i-\gamma)}{\gamma\beta},\alpha\right).$$

This completes the proof of the MGF.

3.4 Conditional Moment

The conditional moment, $E(X^n|X > x)$ of the PHGPW distribution can be expressed as

$$E(X^n|X > x) = \frac{1}{S(x)}J_n(x),$$

where

$$J_n(x) = \int_x^\infty y^n f(y)\,dy$$

$$= \lambda^n e^\alpha \sum_{i=0}^{\infty} \frac{(-1)^i}{\alpha^{\frac{(n-\beta i)}{\gamma\beta}}} \binom{n/\beta}{i} \Gamma\left(\frac{(n-\beta(i-\gamma))}{\gamma\beta}, \alpha\left(1+\left(x/\lambda\right)^\beta\right)^\gamma\right). \quad n = 1, 2, \ldots.$$

Substituting this into the definition, gives

$$E\left(X^n \middle| X > x\right) = \frac{\lambda^n e^\alpha}{e^{\alpha\left\{1-\left(1+\left(x/\lambda\right)^\beta\right)^\gamma\right\}}} \sum_{i=0}^{\infty} \frac{(-1)^i}{\alpha^{\frac{(n-\beta i)}{\gamma\beta}}} \binom{n/\beta}{i} \Gamma\left(\frac{(n-\beta(i-\gamma))}{\gamma\beta}, \alpha\left(1+\left(x/\lambda\right)^\beta\right)^\gamma\right), \qquad (9)$$

where $\Gamma(a,x) = \int_x^\infty z^{\alpha-1} e^{-z}\,dz$ is the upper incomplete gamma function. An application of the conditional moments can be seen in the mean residual life and the mean deviation about the mean and median.

3.5 Stochastic Ordering

Stochastic ordering is the commonest way to show ordering mechanism in lifetime distributions. Suppose a random $X_1 \sim \text{PHGPW}(\alpha_1, \beta, \lambda, \gamma)$ and $X_2 \sim \text{PHGPW}(\alpha_2, \beta, \lambda, \gamma)$, then X_1 is said to be stochastically smaller than X_2 in the

i. stochastic order $(X_1 \leq_{st} X_2)$ if the associated CDFs satisfy: $\mathrm{F}_{X_1} \geq \mathrm{F}_{X_2}$ for all x.

ii. hazard rate order $(X_1 \leq_{hr} X_2)$ if the associated hazard rate functions satisfy: $\mathrm{h}_{X_1} \geq \mathrm{h}_{X_2}$ for all x.

iii. likelihood ratio order $(X_1 \leq_{lr} X_2)$ if the ratio of the associated PDFs defined as $\frac{f_{X_1}(x)}{f_{X_2}(x)}$ decreases in x.

When X_1 and X_2 have a common finite left end-point support, the following implications are true

$$X_1 \leq_{lr} X_2 \Rightarrow X_1 \leq_{hr} X_2 \Rightarrow X_1 \leq_{st} X_2.$$

Suppose that the densities of X_1 and X_2 are

$$f_{X_1}(x)=\frac{\alpha_1\beta\gamma}{\lambda^{\beta}}x^{\beta-1}\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma-1}\exp\left(\alpha_1\left(1-\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma}\right)\right), x>0,$$

and

$$f_{X_2}(x)=\frac{\alpha_2\beta\gamma}{\lambda^{\beta}}x^{\beta-1}\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma-1}\exp\left(\alpha_2\left(1-\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma}\right)\right), x>0$$

respectively. Then, the ratio of the two densities can be expressed as

$$\frac{f_{X_1}(x)}{f_{X_2}(x)}=\frac{\frac{\alpha_1\beta\gamma}{\lambda^{\beta}}x^{\beta-1}\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma-1}\exp\left(\alpha_1\left(1-\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma}\right)\right)}{\frac{\alpha_2\beta\gamma}{\lambda^{\beta}}x^{\beta-1}\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma-1}\exp\left(\alpha_2\left(1-\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma}\right)\right)}$$

$$=\frac{\alpha_1}{\alpha_2}\exp\left(\left(\alpha_1-\alpha_2\right)\left(1-\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma}\right)\right), x>0.$$

Taking the first derivative of the ratio of the two densities yields

$$\frac{d}{dx}\frac{f_{X_1}(x)}{f_{X_2}(x)}=-\frac{\alpha_1\beta\gamma}{\alpha_2\lambda^{\beta}}x^{\beta-1}\left(\alpha_1-\alpha_2\right)\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma-1}\exp\left(\left(\alpha_1-\alpha_2\right)\left(1-\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma}\right)\right). \tag{10}$$

If $\alpha_1>\alpha_2, \frac{d}{dx}\frac{f_{X_1}(x)}{f_{X_2}(x)}<0$, which implies $\left(X_1 \le_{lr} X_2\right)$.

3.6 Order Statistics

Order statistics are important implement used in non-parametric statistics and inference. They play a vital role in areas such as quality control testing and reliability to determine the failure of future items based on the times of few early failures. Suppose $X_1, X_2, X_3, ..., X_n$ are independent and identically distributed random sample of size n taken from the PHGPW distribution with

CDF $G(x)$ and PDF $f(x)$. Let $X_{1:n} \le X_{2:n} \le X_{3:n} \le ... \le X_{n:n}$ denote the order statistics obtained from the sample. The PDF of the k^{th} order statistic for the PHGPW distribution is given by

$$g_{X_{k:n}}(x) = \frac{n!}{(k-1)!(n-k)!} \sum_{j=0}^{n-k} \sum_{i=0}^{j+k-1} \frac{(-1)^{i+j}}{(i+1)} \binom{n-k}{j} \binom{j+k-1}{i} f\left(x;\alpha(i+1),\beta,\lambda,\gamma\right), \tag{11}$$

where $f\left(x;\alpha(i+1),\beta,\lambda,\gamma\right)$ is the PDF of the PHGPW distribution with parameters $\alpha(i+1),\beta,\lambda$ and γ.

4 Methods of Estimation of Model Parameters

In this section, five methods of estimation for estimating the parameters α,β,γ and λ of the PHGPW distribution are presented. These include: maximum likelihood estimation, ordinary and weighted least-squares, Cramér-von Mises and maximum product spacing.

4.1 Method of Maximum Likelihood Estimation

The method of maximum likelihood is the most widely and frequently used method for model parameter estimation. It has many desirable characteristics such as asymptotic efficiency, consistency, and invariant property. Let $X_1, X_2, ..., X_n$ represent a random sample of size n, then the log-likelihood function can be expressed as

$$\ell = n\log\left(\frac{\alpha\beta\gamma}{\lambda^{\beta}}\right) + (\beta-1)\sum_{i=1}^{n}\log(x) + (\gamma-1)\sum_{i=1}^{n}\log\left(1+\left(x/\lambda\right)^{\beta}\right) + \alpha\sum_{i=1}^{n}\left(1-\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma}\right). \tag{12}$$

Taking the first derivatives of the log-likelihood function with respect to the various parameters will yield;

$$\frac{\partial(\alpha,\beta,\gamma,\lambda)}{\partial\alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n}\left(1-\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma}\right), \tag{13}$$

$$\frac{\partial(\alpha,\beta,\gamma,\lambda)}{\partial\beta}=n\left(\frac{1}{\beta}-\log(\lambda)\right)+\sum_{i=1}^{n}\log(x)+(\gamma-1)\sum_{i=1}^{n}\log\frac{\left(x/\lambda\right)\left(x/\lambda\right)^{\beta}}{1+\left(x_i/\lambda\right)^{\beta}}$$

$$-\alpha\sum_{i=1}^{n}\gamma\log\left(x/\lambda\right)\left(x/\lambda\right)^{\beta}\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma-1}, \qquad (14)$$

$$\frac{\partial(\alpha,\beta,\gamma,\lambda)}{\partial\gamma} = \frac{n}{\gamma}+\sum_{i=1}^{n}\log\left(1+\left(x/\lambda\right)^{\beta}\right)-\alpha\sum_{i=1}^{n}\log\left(1+\left(x/\lambda\right)^{\beta}\right)\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma}, \qquad (15)$$

$$\frac{\partial(\alpha,\beta,\gamma,\lambda)}{\partial\lambda} = -\frac{n\beta}{\lambda}-(\gamma-1)\sum_{i=1}^{n}\frac{\beta x\left(x/\lambda\right)^{\beta-1}}{\lambda^{2}\left(1+\left(x/\lambda\right)^{\beta}\right)}+\alpha\sum_{i=1}^{n}\frac{\beta\gamma x\left(x/\lambda\right)^{\beta-1}\left(1+\left(x/\lambda\right)^{\beta}\right)^{\gamma-1}}{\lambda^{2}}. \qquad (16)$$

The maximum likelihood estimates can be obtained as the simultaneous solutions of the equations $\frac{\partial\ell(\alpha,\beta,\gamma,\lambda)}{\partial\alpha}=0$, $\frac{\partial\ell(\alpha,\beta,\gamma,\lambda)}{\partial\beta}=0$, $\frac{\partial\ell(\alpha,\beta,\gamma,\lambda)}{\partial\gamma}=0$ and $\frac{\partial\ell(\alpha,\beta,\gamma,\lambda)}{\partial\lambda}=0$. The solutions of these four nonlinear equations must be obtained using a numerical method. The Newton-Raphson algorithm is one of the standard techniques used in solving these equations.

4.2 Method of Ordinary and Weighted Least-Squares

Swain et al. (1988) discovered the estimators of the ordinary least squares and weighted least squares to estimate the parameters of the beta distribution. Suppose $X_{1:n} \le X_{2:n} \le X_{3:n} \le \ldots \le X_{n:n}$ is the order statistics of a random sample of size n from the PDF of the PHGPW distribution, then the least squares estimators (LSE) of the unknown parameters α,β,γ and λ of the PHGPW distribution can be obtained by minimizing the equation

$$LSE(\alpha,\beta,\gamma,\lambda)=\sum_{i=1}^{n}\left[\left(1-e^{\alpha\left\{1-\left(1+\left(x_{(i)}/\lambda\right)^{\beta}\right)^{\gamma}\right\}}\right)-\frac{i}{n+1}\right]^{2}, \qquad (17)$$

with respect to the parameters α,β,γ and λ.

The weighted least square estimators (WLSE) of the parameters, α, β, γ and λ can also be obtained by minimizing the following function

$$WLSE = \sum_{i=1}^{n} \frac{(n+1)^2 (n+2)}{(n-i+1)} \left[\left(1 - e^{\alpha \left\{ 1 - \left(1 + \left(x_{(i)} / \lambda \right)^{\beta} \right)^{\gamma} \right\}} \right) - \frac{i}{n+1} \right]^2, \qquad (18)$$

with respect to the unknown parameters.

4.3 Method of Cramér-von Mises

MacDonald (1971) defines the Cramér-von Mises (CVM) method as the least minimum distance estimation method which is considered to have the smallest bias as compared to the other minimum distance estimators.

The CVM parameters estimates of the PHGPW distribution are obtained by minimizing equation (19) with respect to the various unknown parameters as:

$$\text{CVM} = \frac{1}{12n} + \sum_{i=1}^{n} \left[\left(1 - e^{\alpha \left\{ 1 - \left(1 + \left(x_{(1)} / \lambda \right)^{\beta} \right)^{\gamma} \right\}} - \frac{2i-1}{2n} \right) \right]^2. \qquad (19)$$

4.4 Method of Maximum Product Spacing

Suppose $X_1, X_2, ..., X_n$ are random samples drawn from the PHGPW distribution with CDF $F(x)$ and $X_{(1)} \le X_{(2)} \le, ..., X_{(n)}$ denotes the order statistics of the sample. Then, spacing D_i is defined as

$$D_i = F(x_i) - F\left(x_{(i-1)}\right) \text{ for } i = 1, 2, 3, ..., n+1,$$

where $F\left(x_{(0)}\right) = 0$ and $F\left(x_{(n+1)}\right) = 1$. The maximum product spacing (MPS) estimates can be obtained by maximizing the logarithm of the geometric mean spacing

$$\mathrm{T} = \frac{1}{n+1}\sum_{i=1}^{n+1}\log_e\left(D_i\right),$$

with respect to the parameters α, β, γ and λ. Therefore, for the estimates of the PHGPW distribution, we maximize equation (20),

$$\mathrm{MPS} = \frac{1}{n+1}\sum_{i=1}^{n+1}\log\left[e^{\alpha\left\{1-\left(1+\left(x_{(i-1)}/\lambda\right)^{\beta}\right)^{\gamma}\right\}} - e^{\alpha\left\{1-\left(1+\left(x_{(i)}/\lambda\right)^{\beta}\right)^{\gamma}\right\}}\right]. \qquad (20)$$

5.0 Log PHGPW Location-Scale Regression Model

The log-PHGPW regression model is presented in this section. It can be obtained by using the transformation $Y = \log(X)$. Suppose a random variable X follows the PHGPW distribution, then the random variable $Y = \log(X)$ follows the log-proportional hazard generalized power Weibull (LPHGPW) distribution.

From the transformation $Y = \log(X)$ and supposing $\lambda^{\beta} = e^{\mu/\sigma}$ and $\beta = 1/\sigma$, the PDF of the LPHGPW is defined as

$$f_y(y) = \frac{\alpha\gamma}{\sigma}\exp\left(\frac{y-\mu}{\sigma}\right)\left(1+\exp\left(\frac{y-\mu}{\sigma}\right)\right)^{\gamma-1}\exp\left(\alpha\left(1-\left(1+\exp\left(\frac{y-\mu}{\sigma}\right)\right)^{\gamma}\right)\right), y \in \mathbb{R}, \qquad (21)$$

where $\mu \in \mathbb{R}$ is the location parameter, $\alpha, \gamma > 0$ are both shape parameters and $\sigma > 0$ is the scale parameter.

The corresponding survival function $S_y(y)$ can be expressed as

$$S_y(y) = \exp\left(\alpha\left(1-\left(1+\exp\left(\frac{y-\mu}{\sigma}\right)\right)^{\gamma}\right)\right), y \in \mathbb{R}. \qquad (22)$$

Figure 2 shows the density function of the LPHGPW regression model. It exhibits various forms of shapes such as left skewed, right skewed and symmetric for some values of the parameters.

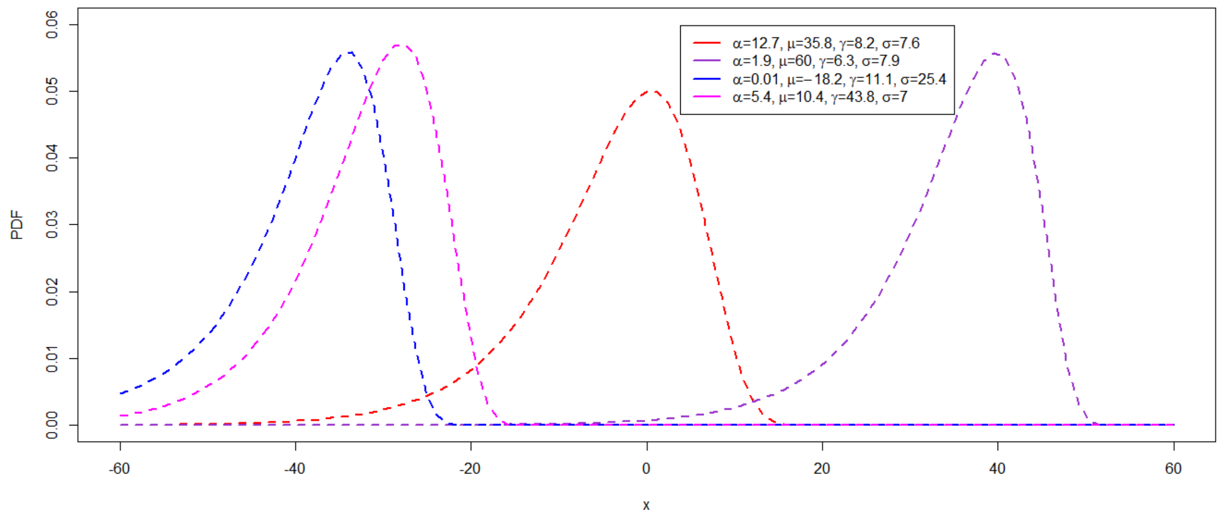


Figure 2: Density function of the LPHGPW regression model

From the density function expressed in equation (21), the log PHGPW location-scale regression model can be expressed with the following regression structure

$$y_i = \boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{\beta} + \sigma z_i, i = 1, \ldots, n \qquad (23),$$

where $\mu = \boldsymbol{x}_{\boldsymbol{i}}^{\boldsymbol{T}} \boldsymbol{\beta}$ is the location parameter that depends on a set of covariates, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_w)^T$ represents the regression parameters, X defines the number of covariates, $\boldsymbol{x}_{\boldsymbol{i}} = (x_{i1}, x_{i2}, \ldots, x_{ik})^{\mathrm{T}}$ also defines the covariates and z_i represents the random error term following the PDF of the LPHGPW.

The regression parameters of the model are estimated using the maximum likelihood estimation method. The log-likelihood function of the LPHGPW regression model is expressed as

$$\ell = n \log\left({}^{\alpha\gamma}\!\big/_{\sigma} \right) + \sum_{i=1}^{n} z_i + (\gamma - 1) \sum_{i=1}^{n} \left(1 + \exp(z_i)\right) + \alpha \sum_{i=1}^{n} \left(1 - \left(1 + \exp(z_i)\right)^{\gamma}\right), \qquad (24)$$

where $z_i = \frac{y_i - x_i^T}{\sigma}$ and n is the number of observations. The parameters estimate of the LPHGPW are estimated by maximizing the log-likelihood function given in equation (24). The adequacy of the LPHGPW regression model is measured using the Cox-Snell residuals (Cox and Snell, 1968). The Cox-Snell residuals of the LPHGPW regression model is defined as

$\hat{r}_i = -\log\left(S\left(y_i \middle| \hat{\alpha}, \hat{\gamma}, \hat{\sigma}, \hat{\mu}\right)\right), i = 1, 2, ..., n,$ where $S\left(y_i \middle| \hat{\alpha}, \hat{\gamma}, \hat{\sigma}, \hat{\mu}\right)$ is expressed in equation (22).

If the LPHGPW regression model fits the given data set well, its Cox-Snell residuals are expected to follow the standard exponential distribution.

6 Simulation Study

To examine the accuracy of the methods of parameter estimation, a Monte Carlos simulation is performed in this section. The quantile function of the PHGPW distribution is used to generate a random sample of observations with size $n = 30, 60, 90, 120, 150, 180$ and 210 with two set of parameter values: $\alpha = 0.1, \beta = 0.3, \gamma = 0.7$, $\lambda = 0.2$ and $\alpha = 0.5, \beta = 1, \gamma = 0.5, \lambda = 0.4$ and the number of replications at $N = 1000$ and during each replication, the average bias (AB), root mean square error (RMSE) and average estimates (AV) are computed.

The results indicated in Table 2 and Table 3 showed that the AV becomes closer to the true parameter value as the size of the sample increases while the RMSE and AB decreases as the sample size increases as expected. The maximum likelihood method is considered for estimating the parameters of the PHGPW distribution as it provides consistent estimation for the parameters. However, the parameter λ fluctuates as the sample size increases as shown in Table 2 and Table 3.

Table 2: Simulation results for α=0.1, β=0.3, ɣ=0.7 and λ=0.2

		AB					**RMSE**					**AV**				
	n	MLE	OLS	WLS	CVM	MPS	MLE	OLS	WLS	CVM	MPS	MLE	OLS	WLS	CVM	MPS
α=0.1	30	0.0710	0.1465	0.1198	0.1239	0.0643	0.0982	0.2424	0.1960	0.2084	0.0849	0.0991	0.1975	0.1735	0.1694	0.0885
	60	0.6350	0.0994	0.0823	0.0904	0.0587	0.0799	0.1468	0.1178	0.1340	0.0704	0.0960	0.1672	0.1534	0.1534	0.0791
	90	0.0609	0.0847	0.0671	0.0797	0.0588	0.0755	0.1147	0.0924	0.1075	0.0702	0.0990	0.1630	0.1444	0.1564	0.0759
	120	0.0593	0.0768	0.0632	0.0723	0.0583	0.0701	0.1034	0.0847	0.0984	0.0675	0.0987	0.1595	0.1426	0.1511	0.0777
	150	0.0562	0.0669	0.0560	0.0656	0.0574	0.0672	0.0886	0.0744	0.0876	0.0656	0.0980	0.1512	0.1385	0.1477	0.0755
	180	0.0563	0.0675	0.0550	0.0627	0.0546	0.0667	0.0880	0.0711	0.0814	0.0628	0.1016	0.1508	0.1370	0.1456	0.0769
	210	0.0528	0.0622	0.0570	0.0641	0.0568	0.0622	0.0803	0.0641	0.0833	0.0641	0.0984	0.1495	0.1341	0.1481	0.0752
β=0.3	30	0.4287	0.2766	0.2534	0.3257	0.5389	0.5174	0.3645	0.3467	0.4128	0.5963	0.6545	0.4719	0.4421	0.5259	0.7944
	60	0.3596	0.2173	0.1893	0.2518	0.4782	0.4656	0.3024	0.2759	0.3410	0.5542	0.6003	0.4466	0.4165	0.4782	0.7407
	90	0.3134	0.1733	0.1432	0.1936	0.4551	0.4309	0.2574	0.2239	0.2742	0.5387	0.5686	0.4227	0.3877	0.4441	0.7275
	120	0.2704	0.1464	0.1370	0.1676	0.3989	0.3907	0.2242	0.2179	0.2514	0.4957	0.5332	0.4034	0.3886	0.4200	0.6733
	150	0.2485	0.1270	0.1088	0.1329	0.3784	0.3714	0.1940	0.1782	0.2000	0.4824	0.5189	0.3889	0.3638	0.3940	0.6592
	180	0.2091	0.1008	0.0959	0.1211	0.3447	0.3288	0.1510	0.1544	0.1815	0.4508	0.4793	0.3604	0.3526	0.3817	0.6249
	210	0.1910	0.1036	0.0898	0.1098	0.3259	0.3116	0.1649	0.1479	0.1672	0.4389	0.4605	0.3687	0.3514	0.3764	0.6065
ɣ=0.7	30	0.6103	0.8191	0.7562	0.8340	0.5748	0.7790	1.1682	1.1077	1.1618	0.7047	0.7180	1.0324	1.0339	1.0192	0.5551
	60	0.4746	0.5745	0.4655	0.6136	0.4869	0.5988	0.8957	0.7198	0.9273	0.5665	0.6457	0.8361	0.7939	0.8575	0.5134
	90	0.3703	0.4181	0.3403	0.4353	0.4247	0.4528	0.6792	0.5364	0.6917	0.4741	0.5802	0.7240	0.7229	0.7173	0.4583
	120	0.3236	0.3389	0.2912	0.3740	0.3807	0.3920	0.5460	0.4223	0.5892	0.4206	0.5710	0.6806	0.6759	0.7022	0.4695

	150	0.2856	0.2942	0.2454	0.3073	0.3458	0.3507	0.4636	0.3567	0.4916	0.3893	0.5622	0.6654	0.6848	0.6757	0.4607
	180	0.2571	0.2717	0.2236	0.2955	0.3291	0.3162	0.4354	0.3108	0.4694	0.3737	0.5786	0.6849	0.6821	0.6803	0.4745
	210	0.2357	0.2485	0.1971	0.2613	0.3098	0.2919	0.3884	0.2623	0.4008	0.3576	0.5940	0.6604	0.6665	0.6566	0.4870
	30	0.4572	0.6872	0.6651	0.6639	0.4146	0.5542	0.7292	0.7143	0.7145	0.5212	0.5045	0.8502	0.8208	0.8212	0.4618
	60	0.4684	0.6565	0.6553	0.6464	0.3959	0.5620	0.7098	0.7077	0.7029	0.4956	0.5036	0.8150	0.8051	0.8057	0.3965
	90	0.4625	0.6582	0.6357	0.6698	0.3600	0.5553	0.7105	0.6944	0.7176	0.4585	0.4874	0.8197	0.7783	0.8326	0.3295
λ=0.2	120	0.4612	0.6439	0.6210	0.6349	0.3757	0.5552	0.7013	0.6835	0.6945	0.4747	0.4889	0.8009	0.7510	0.7896	0.3696
	150	0.4629	0.6498	0.6382	0.6386	0.3688	0.5558	0.7046	0.6957	0.6972	0.4651	0.4918	0.8024	0.7799	0.7933	0.3338
	180	0.4736	0.6546	0.6270	0.6419	0.3631	0.5665	0.7069	0.6880	0.6994	0.4627	0.5126	0.8058	0.7647	0.7962	0.3376
	210	0.4670	0.6444	0.6131	0.6420	0.3734	0.5602	0.7022	0.6781	0.6986	0.4719	0.5044	0.7966	0.7425	0.7934	0.3403

Table 3: Simulation results for $\alpha = 0.5$, $\beta = 1$, $\gamma = 0.5$ and $\lambda = 0.4$

		AB					RMSE					AV				
	n	MLE	OLS	WLS	CVM	MPS	MLE	OLS	WLS	CVM	MPS	MLE	OLS	WLS	CVM	MPS
$\alpha = 0.5$	30	0.4697	0.6580	0.6368	0.5978	0.3862	0.6196	0.8384	0.8134	0.7743	0.4981	0.6289	0.8639	0.8581	0.7838	0.5047
	60	0.4535	0.5935	0.5698	0.5521	0.3926	0.5917	0.7657	0.7375	0.7163	0.5045	0.6583	0.8657	0.8620	0.8098	0.5416
	90	0.4668	0.5966	0.5686	0.5529	0.3939	0.6119	0.7702	0.7301	0.7197	0.5115	0.6112	0.9147	0.8936	0.8546	0.5880
	120	0.4535	0.5722	0.5503	0.5536	0.3872	0.5902	0.7398	0.7013	0.7187	0.4902	0.6107	0.9043	0.8763	0.8583	0.6009
	150	0.4357	0.5491	0.5132	0.5467	0.3596	0.5628	0.7035	0.6629	0.6971	0.4635	0.6086	0.8690	0.8521	0.8589	0.5751
	180	0.4318	0.5436	0.4816	0.5072	0.3584	0.5495	0.6978	0.6114	0.652	0.4506	0.5225	0.8811	0.8222	0.8313	0.5881
	210	0.4175	0.5344	0.4942	0.5433	0.3566	0.5331	0.6856	0.6337	0.6933	0.4521	0.5123	0.8579	0.8271	0.8770	0.5733
$\beta = 1$	30	0.0886	0.1767	0.1569	0.1527	0.0547	0.1841	0.2935	0.2710	0.2682	0.1388	0.9114	0.8232	0.8430	0.8472	0.9452
	60	0.0633	0.1020	0.0910	0.0897	0.0334	0.1409	0.2041	0.1746	0.1906	0.1023	0.9367	0.8979	0.9089	0.9102	0.9665
	90	0.0423	0.0689	0.0600	0.0630	0.0218	0.1007	0.1503	0.1237	0.1456	0.0731	0.9577	0.9310	0.9399	0.9369	0.9781
	120	0.0373	0.0634	0.0493	0.0520	0.0199	0.0866	0.1329	0.1039	0.1203	0.0647	0.9626	0.9365	0.9506	0.9479	0.9800
	150	0.0335	0.0491	0.0418	0.0457	0.0164	0.0746	0.1048	0.0868	0.1053	0.0523	0.9664	0.9508	0.9581	0.9542	0.9835
	180	0.0357	0.0462	0.0437	0.0395	0.0179	0.0753	0.0971	0.0857	0.0914	0.0518	0.9642	0.9537	0.9562	0.9604	0.9820
	210	0.0323	0.0472	0.0389	0.0395	0.0164	0.0684	0.0990	0.0785	0.0879	0.0488	0.9676	0.9527	0.9611	0.9604	0.9835
$\gamma = 0.5$	30	0.3101	0.6178	0.5324	0.5803	0.3093	0.5416	1.0399	0.9424	0.9817	0.5146	0.7293	0.9521	0.8811	0.9434	0.7624
	60	0.2005	0.3233	0.2506	0.3067	0.1923	0.3391	0.6170	0.4646	0.5823	0.3069	0.6268	0.6897	0.6285	0.6921	0.6450
	90	0.1389	0.2117	0.1610	0.2133	0.1313	0.2111	0.3956	0.2757	0.4063	0.1917	0.5528	0.5707	0.5397	0.5911	0.5751
	120	0.1141	0.1775	0.1351	0.1726	0.1085	0.1594	0.2941	0.2014	0.2926	0.152	0.5325	0.5469	0.5189	0.5529	0.5552
	150	0.1018	0.1420	0.1142	0.1497	0.0960	0.1339	0.2178	0.1547	0.2497	0.1256	0.5268	0.5220	0.5065	0.5330	0.5487
	180	0.0967	0.1307	0.1066	0.1270	0.0914	0.1331	0.1928	0.1419	0.1903	0.1184	0.5231	0.5119	0.5107	0.5208	0.5429
	210	0.092	0.1309	0.1038	0.1227	0.0898	0.1182	0.1946	0.1407	0.1828	0.1132	0.5192	0.5149	0.5027	0.5062	0.5397
$\lambda = 0.4$	30	0.4001	0.4230	0.4285	0.4311	0.3829	0.4463	0.4688	0.4721	0.4747	0.4323	0.6035	0.7088	0.7002	0.7105	0.5916
	60	0.3921	0.4369	0.4364	0.4356	0.3701	0.4389	0.4779	0.4787	0.4766	0.4196	0.5914	0.7070	0.7078	0.6994	0.5473
	90	0.3727	0.4267	0.4327	0.4219	0.3381	0.4222	0.4706	0.4752	0.4665	0.3899	0.5779	0.7007	0.7031	0.6916	0.5246

120	0.3642	0.4335	0.4263	0.4194	0.3354	0.4187	0.4755	0.4696	0.4629	0.388	0.5844	0.7118	0.6888	0.6782	0.5205
150	0.3533	0.4111	0.3981	0.4092	0.3083	0.4098	0.4577	0.4511	0.4556	0.3668	0.5760	0.6693	0.6651	0.6676	0.4938
180	0.3581	0.4164	0.4022	0.4073	0.3119	0.414	0.4641	0.454	0.4565	0.3688	0.5941	0.6873	0.6708	0.6671	0.5091
210	0.3407	0.4044	0.3897	0.4095	0.2947	0.3979	0.454	0.4418	0.4589	0.3504	0.5722	0.6628	0.6498	0.6756	0.4812

7.0 Univariate Applications

This section presents univariate application to demonstrate the flexibility of the PHGPW model by means of two data sets. The performance of the PHGPW model is compared with some of its sub-models and other related models. The fit of PHGPW is compared with following distributions:

Generalized power generalized Weibull (GPGW) distribution with PDF as

$$f(x)=\alpha\theta\lambda b x^{\alpha-1}\left(1+\lambda x^{\alpha}\right)^{\theta-1}\exp\left(b\left(1-\left(1+\lambda x^{\alpha}\right)^{\theta}\right)\right),\alpha,\theta,\lambda,b>0,x>0. \qquad (25)$$

Power inverted Nadarajah-Haghighi (PINH) distribution (Ahsan-ul-Haq et al., 2022) with PDF given as

$$f(x)=\alpha\beta\gamma x^{-(\gamma+1)}\left(1+\beta/x^{\gamma}\right)^{\alpha-1}\exp\left(1-\left(1+\beta/x^{\gamma}\right)^{\alpha}\right),\alpha,\beta,\gamma>0,x>0. \qquad (26)$$

Exponentiated Chen (EC) distribution proposed by (Dey et al., 2017) with its PDF as

$$f(x)=\alpha\beta\lambda x^{\beta-1}\exp(x\beta)\exp\left(\lambda\left(1-\exp\left(x^{\beta}\right)\right)\right)\left(1-\exp\left(1-\exp\left(x^{\beta}\right)\right)\right)^{\alpha-1},\alpha,\beta,\lambda>0,x>0. \qquad (27)$$

Exponentiated exponential Weibull (EEW) distribution (Al-Sulami, 2020) with PDF defined as

$$f(x)=\alpha\lambda c/\beta\left(x/\beta\right)^{c-1}\exp\left(-\lambda\left(x/\beta\right)^{c}\right)\left(1-\exp\left(-\lambda\left(x/\beta\right)^{c}\right)\right)^{\alpha-1},\alpha,\beta,\lambda,c>0,x>0. \qquad (28)$$

Nadaraja-Haghighi (NH) distribution (Nadaraja and Haghighi, 2011) with its PDF as

$$f(x)=\alpha\lambda(1+\lambda x)^{\alpha-1}\exp\left(1-(1+\lambda x)^{\alpha}\right),\alpha,\lambda>0,x>0. \qquad (29)$$

The exponentiated power generalized Weibull (EPGW) distribution (Pena-Ramirez et al. 2017) with PDF as

$$f(x)=\alpha\beta\gamma\lambda x^{\gamma-1}\frac{\left(1+\lambda x^{\gamma}\right)^{\alpha-1}\exp\left(1-\left(1+\lambda x^{\gamma}\right)^{\alpha}\right)}{\left(1-\exp\left(1-\left(1+\lambda x^{\gamma}\right)^{\alpha}\right)\right)^{1-\beta}},\alpha,\beta,\gamma,\lambda>0,x>0. \qquad (30)$$

The inverse Weibull (IW) distribution proposed by Khan et al. (2008) with its PDF as

$$f(x)=\frac{\beta}{\theta}\left(\frac{1}{x-x_0}\right)^{\alpha-1}\exp\left(-\frac{1}{\theta}\left(\frac{1}{x-x_0}\right)^{\alpha}\right),\alpha,\theta>0,x_0>0,-\infty<x_0<x. \qquad (31)$$

The exponentiated Nadaraja-Haghighi (ENH) distribution (Lemote, 2013) with PDF as

$$f(x)=\alpha\beta\lambda\frac{(1+\lambda x)^{\alpha-1}\exp\left(1-(1+\lambda x)^{\alpha}\right)}{\left(1-\exp\left(1-(1+\lambda x)^{\alpha}\right)\right)^{1-\beta}},\alpha,\beta,\lambda>0,x>0. \qquad (32)$$

The generalized Rayleigh distribution (GRD) proposed by Raqab and Madi (2011) with PDF defined as

$$f(x)=2\alpha\lambda^2 x\exp\left(-(\lambda x)^2\right)\left(1-\exp\left(-(\lambda x)^2\right)\right)^{\alpha-1},\alpha,\lambda>0,x>0. \qquad (33)$$

Exponentiated generalized inverse Weibull (EGIW) distribution (Elbatal and Muhammed, 2014) with PDF as

$$f(x)=\alpha\beta\theta\lambda x^{-(\theta+1)}\exp\left(-\left(\frac{\lambda}{x}\right)^{\theta}\right)\left(1-\exp\left(-\left(\frac{\lambda}{x}\right)^{\theta}\right)\right)^{\alpha-1}\times$$
$$\left(1-\left(1-\exp\left(-\left(\frac{\lambda}{x}\right)^{\theta}\right)\right)\right)^{\beta-1},\alpha,\beta,\theta,\lambda>0,x>0. \qquad (34)$$

The maximum likelihood estimates of the unknown parameters of these models are estimated by maximizing the log-likelihood function. The performance of the PHGPW model is then compared with the other distributions using the Akaike information criterion (AIC), corrected Akaike information criterion (AICc) and Bayesian information criterion (BIC). Kolmogorov-Smirnov (KS), Anderson-Darling (AD) and Cramér-von Mises (CVM) test of goodness-of-fit were also performed. The distribution with the smallest of these measures is considered the best model.

7.1 First Application

The first data set represents remission time in months of a random sample of 128 bladder cancer patients as cited in Shanker et al. (2015). Table 4 indicates the maximum likelihood estimates of the distributions with their corresponding standard errors in parentheses for data set I.

Table 4: Parameter estimates and standard errors for data set I

Model	**Parameter Estimates**			
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\lambda}$
PHGPW	61.8739	1.4589	0.0303	10.8552
	(0.0413)	(0.1809)	(0.0131)	(4.7499)
GPGW	2.6178	0.3939	838.1000	0.0071
	(0.9511)	(0.1520)	(0.0003)	(0.0079)
EEW	2.8584	65.8642	7.1592	0.6531
	(1.2773)	(0.0518)	(0.7293)	(0.1322)
EPGW	11.8210	179.0400	0.0899	0.1479
	(8.4057)	(0.0201)	(0.0072)	(0.1126)
EGIW	22.4921	0.6334	0.3651	260.3165
	(4.0733)	(0.2998)	(0.0501)	(0.2214)
GPW	1.5773	0.4181	3.4564	
	(0.2409)	(0.1056)	(0.8652)	
GW	2.8584	0.6531	0.3093	
	(1.2773)	(0.1322)	(0.1726)	
PINH	0.2058	506.4900	2.1864	
	(0.01071)	(0.0000)	(0.0871)	
EC	8.3643	0.2209	0.7280	
	(3.7543)	(0.0273)	(0.2016)	
IW	-3.7186	0.0131	2.0699	
	(1.9819)	(0.0219)	(0.5399)	
ENH	0.3193	5.0312	100.0000	
	(0.0170)	(0.9073)	(0.0219)	
W	0.0946	1.0515		
	(0.0191)	(0.0675)		
NH	0.9243	0.1234		
	(0.1499)	(0.0345)		
GRD	0.3404	0.7726		

	(0.0474)	(0.0881)

Table 5 also indicates the model selection criteria and goodness-of-fit measures for the bladder cancer patient’s data set. It can be seen that the PHGPW distribution has the least statistic in terms of all the measures. This is an indication that the PHGPW distribution provides a better fit to the bladder cancer patient’s data set as compared to the other competing distributions. This is demonstrated by the P-P plots shown in Figure 3. It can be seen that PHGPW distribution provides the better description for this data set. These figures support the results in Table 5.

Table 5: Model selection criteria and Goodness-of-fit statistics of data set I

Model	-2LogL	AIC	AICc	BIC	KS	AD	CVM
PHGPW	**814.0991**	**822.0991**	**822.4243**	**833.5072**	**00343.** **(0.9982)**	**0.1252** **(0.9997)**	**0.0173** **(0.9989)***
GPGW	822.9828	830.9828	831.3080	842.3909	0.0724 (0.5129)	1.0098 (0.3520)	0.1798 (0.3614)
EEW	816.4617	824.4617	824.7869	835.8699	0.0455 (0.9539)	0.3096 (0.9306)	0.0458 (0.9021)
EPGW	834.8782	843.8782	843.2034	854.2864	0.0819 (0.3560)	1.6334 (0.1477)	0.24807 (0.1909)
EGIW	822.3527	830.3527	830.6779	841.7608	0.0758 (0.4539)	1.0826 (0.3166)	0.1873 (0.2938)
GPW	815.6632	822.6632	822.8567	833.2193	0.0392 (0.9893)	0.2494 (0.9707)	0.0359 (0.9532)
GW	816.4617	822.4617	822.6553	831.0178	0.0455 (0.9539)	0.3096 (0.9306)	0.0458 (0.9021)
PINH	826.119	832.119	832.3126	840.6751	0.1040 (0.1253)	2.1729 (0.0741)	0.3908 (0.0763)
EC	816.8311	822.8311	823.0246	831.3872	0.04655 (0.9451)	0.3417 (0.9037)	0.0511 (0.8704)
IW	817.3415	823.3415	823.5351	831.8976	0.1416 (0.0118)	6.3228 (0.0001)	1.0209 (0.0021)
ENH	816.3044	822.3044	822.4979	830.8605	0.0447 (0.9601)	0.3024 (0.9363)	0.0449 (0.9072)
W	823.7849	827.7849	827.8809	833.489	0.9999 (0.0000)	1.0478 (0.3330)	0.16631 (0.3438)
NH	824.1473	828.1473	828.2433	833.8514	0.0948 (0.2005)	1.4535 (0.1880)	0.22381 (0.2258)
GRD	855.5003	859.5003	859.5963	865.2043	0.1593	5.0772	1.0011

	(0.0030)	(0.0026)	(0.0024)

*: Means the model that fits the data well.

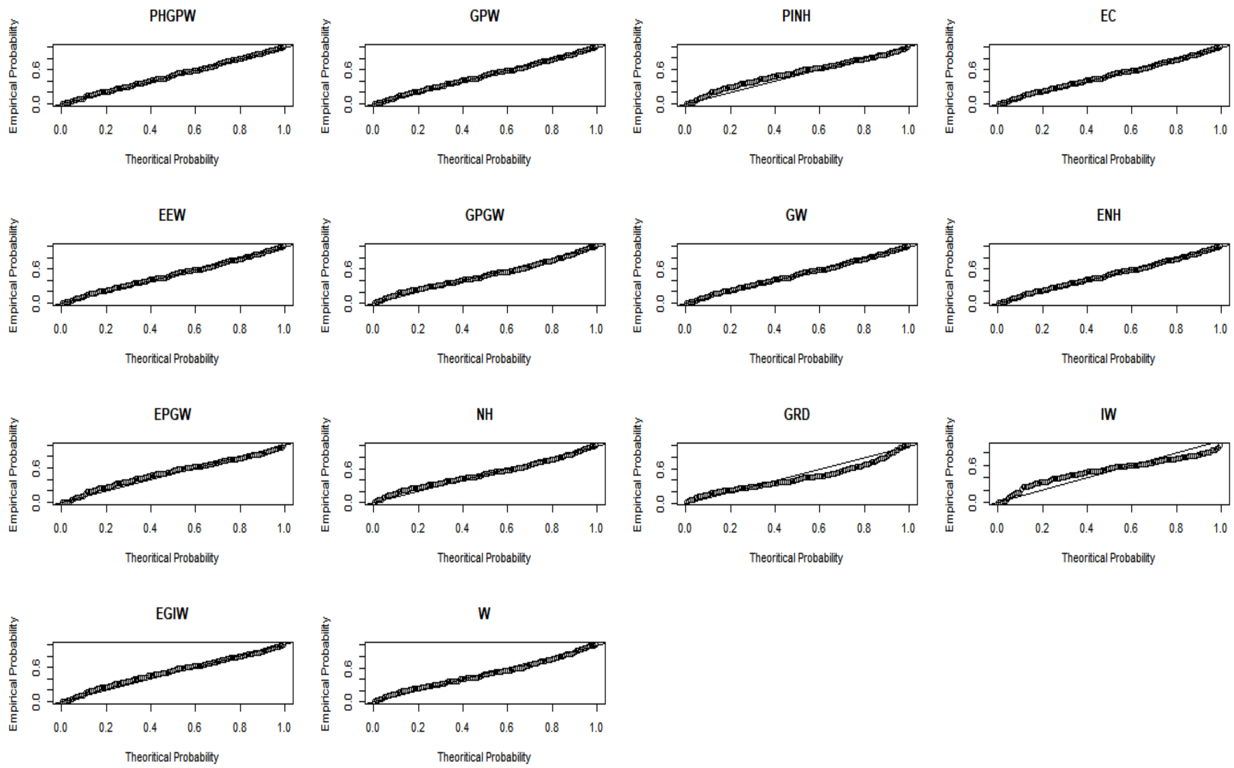


Figure 3: P-P Plots for the fitted distributions for data set I

7.2 Second Application

The second data set taken from Kennan (1985) are made up of 62 observations of strike durations in days for a US manufacturing companies. The parameters estimate of the second data with their corresponding standard errors in parentheses is shown in Table 6.

Table 6: Parameter estimates and standard errors for data set II

MODEL	PARAMETERS			
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\lambda}$
PHGPW	0.0487 (0.0237)	85.8170 (0.0000)	0.0097 (0.0012)	0.9601 (0.0671)
GPGW	100.0000 (0.0000)	0.0083 (0.0011)	38.3970 (0.0000)	0.0489 (0.0234)
EEW	1.5259 (1.2412)	68.2269 (0.0183)	2.0211 (0.8412)	0.7334 (0.3215)
EPGW	0.0733 (0.0811)	4.7299 (3.5932)	3.5719 (4.4808)	64.9133 (0.0594)
EGIW	6.4548 (1.7858)	0.3746 (0.2785)	0.4914 (0.1264)	289.2910 (0.0233)
GPW	1.0118 (0.2557)	0.7512 (0.5075)	26.3635 (26.9754)	
GW	1.5259 (1.2412)	0.7334 (0.3215)	0.0383 (0.0372)	
PINH	0.2955 (0.0242)	380.4700 (0.0000)	1.4141 (0.0730)	
EC	5.1253 (3.9972)	0.1915 (0.0407)	0.3786 (0.2556)	
IW	−7.0468 (5.3174)	0.0183 (0.0262)	1.2876 (0.3395)	
ENH	0.7358 (0.3348)	1.0417 (0.3292)	0.0410 (0.0428)	
W	0.0322 (0.0129)	0.9247 (0.0916)		
NH	0.7720 (0.2062)	0.0364 (0.0179)		
GRD	0.3234 (0.0463)	0.0104 (0.0013)		

The model selection criteria and test statistics of the good-of-fit measures are displayed in Table 7. The results indicates that PHGPW provides a better fit for the data though GPW, GW, EEW, PINH, ENH and EC distributions also compete well in describing this data. This is evident in the P-P plots indicated in Figure 4. It can be seen that PHGPW distribution best describes the data set. These figures support the results in Table 7.

Table 7: Goodness-of-fit statistics fit data set II

Model	-2ℓ	AIC	AICc	BIC	KS	AD	CVM
PHGPW	**584.1434**	**592.1434**	**592.8452**	**600.6520**	**0.0716 (0.9084)**	**0.3286 (0.9149)**	**0.0368 (0.9500)** *
GPGW	584.2139	592.2139	593.8157	601.6225	0.1148 (0.3875)	1.5191 (0.1720)	0.2135 (0.2340)
EEW	588.5051	596.5051	597.2068	605.0136	0.0933 (0.6527)	0.6799 (0.5752)	0.0918 (0.6282)
EPGW	594.0775	602.0775	602.7793	610.5861	0.1213 (0.3214)	1.0797 (0.3177)	0.1809 (0.3081)
EGIW	592.1597	600.1597	600.8615	608.6683	0.1256 (0.2820)	1.1983 (0.2680)	0.2023 (0.2634)
GPW	588.6566	594.6566	595.0704	601.0380	0.0969 (0.6047)	0.7360 (0.5289)	0.0698 (0.7547)
GW	588.6566	594.6566	594.9189	600.8865	0.0357 (0.9552)	0.1697 (0.0563)	3.2531 (0.0205)
PINH	593.1895	599.1895	599.6033	605.5709	0.1187 (0.3467)	1.0653 (0.3245)	0.1738 (0.3252)
EC	588.5383	594.5383	594.9521	600.9197	0.0721 (0.9037)	0.3479 (0.8979)	0.0390 (0.9395)
IW	598.7235	604.7235	605.1373	611.1049	0.1429 (0.1589)	1.9569 (0.0972)	0.3118 (0.1251)
ENH	588.6415	594.6415	595.0553	601.0229	0.0792 (0.8317)	0.4916 (0.7548)	0.0599 (0.8161)
W	588.8054	592.8054	593.0088	597.0596	0.0702 (0.9199)	0.3504 (0.8957)	0.0398 (0.9356)
NH	588.6588	592.6588	592.8622	596.9131	0.8302 $(< 2.2\times10^{-1}$	75.8710 (0.0000)	13.8380 $(< 2.2\times10^{-16})$
GRD	593.5089	597.5089	597.7123	601.7632	0.1309	1.0229	0.1932

	(0.2384)	(0.3452)	(0.2815)

*: Means the model that fits the data well.

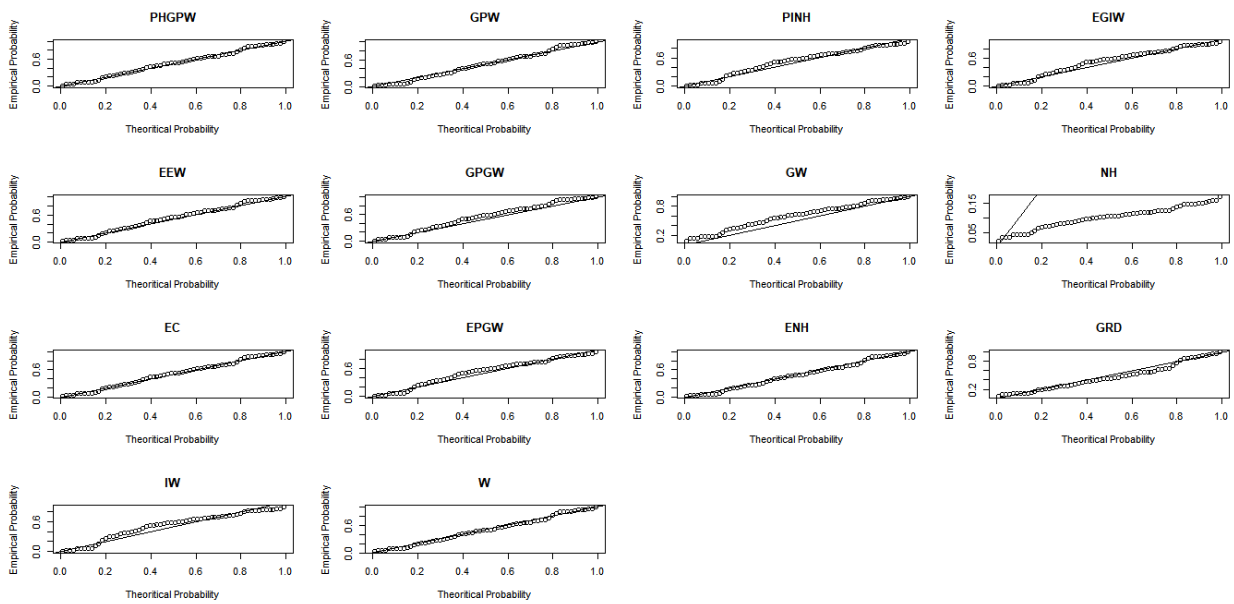


Figure 4. P-P Plots for the fitted distributions for data set II

8.0 Applications of the log-Proportional Hazard Generalized Power Weibull Regression Model

In this section, two real data set are used to demonstrate the application and the usefulness of the LPHGPW regression model. The maximum likelihood estimation technique is used to obtain the parameter estimates for the LPHGPW regression model. The parameter estimates and their standard errors are obtained together with the AIC, AICc and BIC to compare the LPHGPW regression model with some other regression models such the log Weibull (LW), log generalized power Weibull (LGPW) and log proportional hazard Weibull (LPHW).

8.1 First Data Set

The first data set used to illustrate the application of the LPHGPW regression model consist of medical expenses data which contains weekly average medical expenses including cost of drugs for 33 randomly sampled families from 600 families made up of 2700 individuals

(https://vincentarelbundock.github.io/Rdatasets/datasets.html). The following regression model is fitted to the given data set

$$y_i = \beta_0 + \beta_1 x_{1i} + \sigma z_i, i = 1, 2, ..., 33,$$

where $y_i = \log(x_i)$ follows the LPHGPW distribution. Y_i (the response variable) is the average weekly medical expenses per a member of a family and x_i is the number of members in a family.

For this data set, the LW, LGPW, LPHW and the LPHGPW distributions are adjusted for and their performance compared. The maximum likelihood estimates of the parameters of these regression models are shown in Table 8. From the LPHGPW regression model, the family size is significant at 1% and negatively affects the family weekly cost of medication.

Table 8: Estimated parameters of the regression models

Model	Parameter	Estimate	Std Error	P-Value
LPHGPW	α	1.3616	4.2323	0.7477
	γ	0.1676	0.1583	0.2895
	σ	0.8461	0.5958	0.1556
	$\beta 0$	10.4317	3.8959	0.0007
	$\beta 1$	–1.2188	0.3509	0.0005
LGPW	γ	1.1635	0.1292	$< 2.2\times10^{-16}$
	σ	1000000.0000	4.3928×10^{-13}	$< 2.2\times10^{-16}$
	$\beta 0$	0.3670	1.0845×10^{-7}	$< 2.2\times10^{-16}$
	$\beta 1$	0.9671	4.0421×10^{-7}	$< 2.2\times10^{-16}$
LPHW	α	1.0000	0.1741	9.2160×10^{-9}
	σ	1000000.0000	4.0×10^{-7}	$< 2.2\times10^{-16}$
	$\beta 0$	0.1669	0.0007	$< 2.2\times10^{-16}$
	$\beta 1$	0.9669	6.4883×10^{-7}	$< 2.2\times10^{-16}$
LW	σ	2.1893	0.2761	2.2020×10^{-15}
	$\beta 0$	12.9633	0.9670	$< 2.2\times10^{-16}$
	$\beta 1$	–1.0303	0.2397	1.7220×10^{-5}

The LPHGPW regression model's adequacy for this data set was assessed using the Cox-Snell residuals and the results indicated that the LPHGPW follows the standard exponential distribution. This is supported by KS, AD and CVM values displayed in Table 9.

Table 9: Model Selection Criteria and Goodness-of-fit statistics of the regression models

Model	-2ℓ	AIC	AICc	BIC	KS	AD	CVM
LPHGPW	**143.6446**	**153.6446**	**155.8668**	**161.1272**	**0.0601** (**0.9998**)	**0.0965** (**1.0000**)	**0.0133** (**0.9999**) *
LGPW	976.1896	984.1896	985.6181	990.1756	0.7106 6.6610×10^{-15}	19.1940 (1.8180×10^{-5})	4.2138 $(<2.2\times10^{-16})$
LPHW	977.8237	985.8237	987.2523	991.8097	0.6321 7.044×10^{-12}	15.1360 (1.8180×10^{-5})	3.3260 $(<2.2\times10^{-16})$
LW	150.9120	156.9120	157.7396	161.4015	0.1227 (0.7034)	0.7109 (0.5486)	0.1014 (0.5808)

*: Means the model that fits the data well.

According to the model selection criteria, the values of the AIC, AICc and BIC of the LPHGPW regression model for the first data set are smaller as compared to the values of the LW, LGPW and LPHW regression models as indicated in Table 9. Therefore, it can be concluded that the regression model with LPHGPW error distribution provides a better fit for this data set than the regression models of the LW, LGPW and LPHW distributions. The KS, AD and CVM along with their P-values in parentheses, test of goodness-of-fit are also performed and the results indicated in Table 9. The LPHGPW regression model has the smallest test statistic values for the KS, AD and CVM which is an indication that the LPHGPW regression model fit this data set better than the rest of regression models. This also confirm the results of the model selection criteria displayed in Table 9.

8.2 Second Data Set

Another data set used to demonstrate the application of the LPHGPW regression model consist of a group of 200 women of at least 21 years of age, of Pima Indian heritage and living closer to Phoenix, Arizona, were tested for diabetes according to the World Health Organization procedure (Smith et al., 1988). The regression model fitted to this data set is given as

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \sigma z_i, i = 1, 2, ..., 200,$$

where y_i (response variable) represents plasma glucose concentration in an oral tolerance test, x_1 (body mass index in kilograms) and x_2 (age in years). The maximum likelihood estimates of the parameters of the LW, LGPW, LPHW and LPHGPW regression models are shown in Table 10.

Table 10: Estimated parameters of regression models

Model	Parameter	Estimate	Std Error	P-Value
LPHGPW	α	2.8702	1.6471	0.0814
	γ	0.1207	0.0414	0.0036
	σ	12.2531	1.8674	5.3220×10^{-11}
	$\beta 0$	49.7967	11.4994	1.4890×10^{-5}
	$\beta 1$	0.7328	0.3353	0.0288
	$\beta 2$	0.9086	0.1939	2.8000×10^{-6}
LGPW	γ	1.1634	5.2484×10^{-2}	$< 2.20\times10^{-16}$
	σ	1000000.0000	3.8308×10^{-12}	$< 2.20\times10^{-16}$
	$\beta 0$	0.3671	4.4046×10^{-8}	$< 2.20\times10^{-16}$
	$\beta 1$	0.9699	1.4231×10^{-6}	$< 2.20\times10^{-16}$
	$\beta 2$	0.1649	1.4143×10^{-6}	$< 2.20\times10^{-16}$
LPHW	α	0.9999	7.0705×10^{-2}	$< 2.20\times10^{-16}$
	σ	1000000.0000	6.1702×10^{-12}	$< 2.20\times10^{-16}$
	$\beta 0$	0.3669	7.0699×10^{-8}	$< 2.20\times10^{-16}$
	$\beta 1$	0.9657	2.2843×10^{-6}	$< 2.20\times10^{-16}$
	$\beta 2$	0.1597	2.2701×10^{-6}	$< 2.20\times10^{-16}$
LW	σ	30.7412	1.5313	$< 2.20\times10^{-16}$
	$\beta 0$	71.3054	13.3780	9.8190×10^{-8}
	$\beta 1$	1.1932	0.3676	1.1710×10^{-3}
	$\beta 2$	0.9074	0.1945	3.1000×10^{-6}

The Cox-Snell residuals were used to assess the adequacy of the LPHGPW regression model which showed that it followed the standard exponential distribution. This is supported by the results of the goodness-of-fit statistics indicated in Table 11.

The model selection criteria (AIC, AICc and BIC) and the goodness-of-fit statistics (KS, AD and CVM) test values with P-values in parentheses of the LPHGPW regression model are the least as compared to the LW, LGPW and LPHW regression models as displayed in Table 11. Therefore, the LPHGPW regression model provides a better fit for the test of diabetes data than the rest of the regression models (LW, LGPW and LPHW).

Table 11: Goodness-of-fit statistics of regression models

Model	-2ℓ	AIC	AICc	BIC	Ks	AD	CVM
LPHGPW	**1911.0360**	**1923.0360**	**1923.4710**	**1942.8260**	**0.0685 (0.3045)**	**0.7239** $(\mathbf{0.5390})$	**0.1444** $(\mathbf{0.4073})^{*}$
LGPW	5916.3080	5926.3080	5926.6170	5942.7990	0.7106 $(< 2.2\times10^{-16})$	116.3100 (3.0×10^{-6})	25.5360 $(< 2.2\times10^{-16})$
LPHW	5926.2040	5936.2040	5936.5140	5952.6960	0.6321 $(< 2.2\times10^{-16})$	91.7290 (3.0×10^{-6})	20.1570 $(< 2.2\times10^{-16})$
LW	1965.7210	1973.7210	1973.9260	1986.9140	0.1514 (0.0002)	5.7249 (0.0013)	1.0188 (0.0022)

*: Means the model that fits the data well.

9.0 Conclusions

In this paper we introduce the proportional hard generalized power Weibull (PHGPW) model to extend the generalized power Weibull distribution. It has additional parameter and its hazard rate function can assume increasing, decreasing, unimodal or bathtub (upside bathtub) and constant shapes. The proposed distribution contains special models. We have discussed some mathematical properties of the model, estimate its parameters using the maximum likelihood and demonstrate its flexibility with application of two data sets. The new distribution provides a good adjustment in the two applications and it is also quite competitive with other models. The model can be used to effectively provide better fit as compare to the other lifetime distributions.

The LPHGPW regression model was derived. This was supported with real data demonstration which showed that the LPHGPW regression model proves to have a better fit than the LW, LGPW and LPHW regression models. The results of this demonstration indicates that the proposed new distribution is flexible.

Motivation.

The ability of the distribution to model hazard rate function with increasing, decreasing, unimodal or (upside bathtub) and constant.

References

Aboukhamseen, S. M., Ghitany, M. E. and Gupta, R. C. (2016). Proportional hazard inverse Weibull distribution and associated inference. *Journal of Mathematical and Statistics*,**12**(2):86-98.

Ahsan-ul-Haq, M., Ahmed, J., Albassam, M. and Aslam, M. (2022). Power inverted Nadarajah Haghighi distribution: Properties, estimation and applications. *Journal of Mathematics*, 1-10.

Al-Sulami, D. (2020). Exponentiated exponential weibull distribution: Mathematical properties and application. *American Journal of Applied Sciences*, **17**: 188-195.

Broderick, O.O., Mashabe, B., Fagbamigbe, A., Makubate, B. and Wanduku, D. (2020). The exponentiated generalized power series of family of distributions: Theory and application. *Heliyon*, **6**(2020): 1-16.

Cox, D. R. and Snell, E. J. (1968). A general definition of residuals. *Journal of Royal Statistical Society, Series B*, **30**(2):248-275.

Dey, S., Kumar, D., Ramos, P. L. and Louzada, F. (2017). Exponentiated Chen distribution: Properties and estimation. *Communications in Statistics-Simulation and Computation*, **46**(10): 8118-8139.

Elbatal, I. and Muhammed, H. Z. (2014). Exponentiated generalized inverse Weibull distribution. *Applied Mathematical Sciences*, **8**(81): 3997-4012.

Gusmao, F. R. S., Ortega, E. M. M. and Cordeiro, G. M. (2009). The generalized inverse Weibull distribution. *Stat Papers*, 1-29. DOI 10.1007/s00362-009-0271-3.

Kennan, J. (1985). The duration of contract strikes in US manufacturing. *Journal of Econometrics*, **28**:5-28.

Khan, M. S., Pasha, G. R. and Pasha, A. H. (2008). Theoretical analysis of inverse Weibull distribution. *WSEAS Transactions on Mathematics*, **7**(2): 29-38.

Lai, C. (2014). Generalized Weibull distributions. Springer Heidelberg New York Dordrecht London.

Lemote, A. J. (2013). A new exponentiated-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function. *Computational Statistics and Data Analysis*, **62**(2013): 149-170.

MacDonald, P. (1971). An estimation procedure for mixture of distribution. *Journal of the Royal Statistical Society, Series B*, **33**: 326-329.

Martinez-Florez, G., Moreno-Arenas, G. and Vergara-Cardozo, S. (2013). Properties and inference for proportional hazard models. *Revista Colombiana de Estadistica*, **38**(1):95-114.

Moreno-Arenas, G. Martinez-Florez, G. and Barrera-Causil, C. (2016). Proportional hazard Birnbaum-Saunders distribution with application to survival data analysis. *Revista Colombiana de Estadistica*, **39**(1):129-114.

Nadaraja, S. and Haghighi, F. (2011). An extension of the exponential distribution. Statistics: *journal of Theoretical and Applied Statistics*, **45**(6):543-558.

Nikulin, M. and Haghighi, F. (2006). A Chi-Squared Test for the Generalized Power Weibull Family for the Head-and-Neck Cancer Censored Data. *Journal of Mathematical Sciences*, **133**(3): 1333-1341.

Oguntunde, P. E., Odetunmibi, O. A. and Adejumo, A. O. (2015). On the exponentiated generalized Weibull distribution: A generalization of Weibull distribution. *Indian Journal of Science and Technology*, **8**(35):1-7.

Pena-Ramirez, F. A., Guerra, R. R., Cordeiro, G. M. and Marinho, P. R. D. (2017). The exponentiated power generalized Weibull: Properties and applications. *Annal of the Brazilian Academy of Sciences,* 1-26.

Pu, S., Broderick, O.O., Yiu, Q. and Daniel, F. L. (2016). A generalized class of exponentiated modified Weibull distribution with applications. *Journal of Data Science,* **14**(4): 585-614. Source: https://www.researchgate.net/publication.

Raqab, M. Z. and Madi, M. T. (2011). Generalized Rayleigh distribution. *ResearchGate*, 1-6.

Sarhan, A. M. and Zaindin, M. (2009). Modified Weibull distribution. *Applied Sciences*, **11**(2009): 123-139.

Selim, M. A. and Badr, A. M. (2016). The Kumaraswamy generalized power Weibull distribution. *Mathematical Theory and Modeling*, **6**(2): 110-124.

Selim, M.A. (2018). The generalized power generalized Weibull distribution: properties and applications. *ResearchGate*, 1-15.

Shanker, R., Fesshaye, H. and Selvaraj, S. (2015). On modeling of lifetimes data using exponential and lindley distributions. *Biometrics And Biostatistics International Journal,* **2**(5): 140-147.

Smith, J. W., Everhart, J. E., Dickson, W. C., Knowler, W. C. and Jonnaes, R. S. (1988). Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In proceedings of the Symposium on Computer Applications in Medical Care (Washington, 1988), ed. R. A. Greenes, 261-265. Los Alamitos, CA: IEEE Computer Society Press.

Swain, J.J., Venkatraman, S. and Wilson, J.R. (1988). Least-squares estimation of distribution in Johnson's translation system. *Journal Of Statistical Computation and Simulation*, **29**(4):271-297. Source: doi:10.1080/00949658808811068.