RESEARCH　　　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

CrossMark

The gamma extended Weibull distribution

Gauss M. Cordeiro[1], Maria do Carmo S. Lima[1], Antonio E. Gomes[2*] , Cibele Q. da-Silva[2]
and Edwin M. M. Ortega[3]

*Correspondence:
aegomes@unb.br
[2]Departamento de Estatística,
Universidade de Brasília, 70910-900
Brasília, DF, Brazil
Full list of author information is
available at the end of the article

Abstract

The two-parameter Weibull has been the most popular distribution for modeling lifetime data. We propose a four-parameter gamma extended Weibull model, which generalizes the Weibull and extended Weibull distributions, among several other models. We obtain explicit expressions for the ordinary and incomplete moments, generating and quantile functions and mean deviations. We employ the method of maximum likelihood for estimating the model parameters. We propose a log-gamma extended Weibull regression model with censored data. The applicability of the new models is well justified by means of two real data sets.

Keywords: Extended Weibull distribution, Gamma extended Weibull distribution, Generating function, Maximum likelihood estimation, Mean deviation, Quantile function

MSC: 62E10, 62E20, 62F10

1 Introduction

There are hundreds of continuous univariate distributions and recent developments focus on constructing general distributions from classic ones. Many distributions have been used as models to make inferences about a population based on a set of empirical data from that population. Determining an adequate model to make inferences is a very important problem. The Weibull distribution is commonly used for modeling systems with monotone failure rates. However, the data sets in reliability analysis, especially over the life-cycle of the product, can involve high initial failure rates and eventual high failure rates due to aging and wear out, indicating a bathtub failure rate. The distributions that allow only monotone failure rates might not be adequate for modeling populations giving rise to such data. The major weakness of the Weibull distribution is its inability to accommodate non-monotone hazard rates, which has led to new generalizations of this distribution. One of the first extensions allowing for non-monotone hazard rates, including the bathtub shaped hazard rate function (hrf), is the *exponentiated Weibull* (ExpW) (Mudholkar and Srivastava 1993; Mudholkar et al. 1995; and Mudholkar et al. 1996) distribution. It has been well established in the literature that the ExpW distribution provides significantly better fits than the well-known exponential, gamma, Weibull and log-normal distributions. In the last paper, the authors presented a more flexible three-parameter *extended Weibull* (EW) model. Further, Shao et al. (2004) used this distribution to study flood frequency and Hao and Singh (2008) described some of its

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 2 of 19

applications in hydrology. In this paper, we propose a further generalization by taking the EW distribution as the baseline model.

The three-parameter EW distribution is defined by the probability density function (pdf) and cumulative distribution function (cdf) (Mudholkar et al. 1996)

$$g_{\lambda,\alpha,\beta}(x) = \lambda\,\beta\,x^{\beta-1}\,\left(1+\alpha\lambda\,x^\beta\right)^{-\frac{1}{\alpha}-1},\, \alpha > 0 \quad \text{and} \quad g_{\lambda,\beta}(x) = \lambda\,\beta\,x^{\beta-1}\,\mathrm{e}^{-\lambda x^\beta},\, \alpha = 0,$$

(1)

and

$$G_{\lambda,\alpha,\beta}(x) = 1 - \left(1+\alpha\lambda x^\beta\right)^{-\frac{1}{\alpha}},\, \alpha > 0 \quad \text{and} \quad G_{\lambda,\beta}(x) = 1 - \mathrm{e}^{-\lambda x^\beta},\, \alpha = 0,$$

(2)

respectively, where $\lambda > 0$ is a scale parameter and $\alpha \geq 0$ and $\beta > 0$ are shape parameters. The support of the EW distribution is $\mathbb{R}^+$. The forms of the pdf and cdf when α goes to zero tend to those ones when $\alpha = 0$. Clearly, the cdf (2) extends the Weibull cdf and this fact justifies the terminology EW model. Due to the shape parameter α, more flexibility can be incorporated in model (1), which is very useful for lifetime data. The survival function associated to (1) is $S_{\lambda,\alpha,\beta}(x) = 1 - G_{\lambda,\alpha,\beta}(x)$ for $\alpha > 0$ and $S_{\lambda,\beta}(x) = 1 - G_{\lambda,\beta}(x)$ for $\alpha = 0$.

A family of univariate distributions generated by gamma random variables was proposed by Zografos and Balakrishnan (2009) and Ristic and Balakrishnan (2012). For any baseline cdf $G(x)$, $x \in \mathbb{R}$, they defined the *gamma-G* (denoted with the prefix "GG" for short) model with an extra parameter $a > 0$ by the pdf and cdf

$$f(x) = \frac{g(x)}{\Gamma(a)}\,\left\{-\log\left[1-G(x)\right]\right\}^{a-1}$$

(3)

and

$$F(x) = \frac{1}{\Gamma(a)}\,\int_0^{-\log[1-G(x)]} t^{a-1}\,\mathrm{e}^{-t}\,dt = \gamma_1(a, -\log\left[1-G(x)\right]),$$

(4)

respectively, where $g(x) = dG(x)/dx$, $\Gamma(a) = \int_0^\infty t^{a-1}\,\mathrm{e}^{-t}dt$ is the gamma function, $\gamma(a,z) = \int_0^z t^{a-1}\,\mathrm{e}^{-t}dt$ denotes the incomplete gamma function and $\gamma_1(a,z) = \gamma(a,z)/\Gamma(a)$ is the incomplete gamma function ratio.

The GG model has the same parameters of the parent G distribution plus one extra shape parameter $a > 0$. Each new GG distribution can be determined from a specified G model. For $a = 1$, the G distribution is a basic exemplar with a continuous crossover towards cases with different shapes (for example, a particular combination of skewness and kurtosis).

We introduce a new four-parameter model called the *"gamma extended Weibull"* ("GEW") distribution, which contains as special models some distributions such as the EW distribution. In fact, this model represents only a basic exemplar of the GEW distribution. We also study some of its mathematical properties. The paper is outlined as follows. In Section 2, we define the GEW distribution and provide some of its special cases. Further, two useful expansions for its density and cumulative distributions are derived in Section 3. In Section 4, we obtain its quantile function (qf). The ordinary moments and mean deviations are investigated in Section 5. Maximum likelihood estimation of the model parameters and some inferential tools are discussed in Section 6. In Section 7, we propose the *log-gamma extended Weibull* regression model, which can be applied for lifetime analysis. The importance of the new models is shown empirically by means of two real data sets in Section 8. Some conclusions are offered in Section 9.

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 3 of 19

2   The GEW distribution

By taking the pdf (1) and cdf (2) of the EW distribution with scale parameter $\lambda > 0$ and shape parameters $\alpha \geq 0$ and $\beta > 0$, the pdf and cdf of the GEW distribution can be obtained from Eqs. (3) and (4) (for $x > 0$) as

$$f(x) = f(x; \tau, a) = \begin{cases} \frac{\lambda \beta x^{\beta-1} \left(1+\alpha\lambda x^{\beta}\right)^{-\frac{1}{\alpha}-1}}{\alpha^{a-1} \Gamma(a)} \left\{\log\left[\left(1+\alpha\lambda x^{\beta}\right)\right]\right\}^{a-1} &, \alpha > 0, \\ \frac{\lambda^a \beta x^{\beta a-1} e^{-\lambda x^{\beta}}}{\Gamma(a)} &, \alpha = 0, \end{cases} \tag{5}$$

$$F(x) = F(x; \tau, a) = \begin{cases} \gamma_1\left(a, \frac{1}{\alpha}\log\left[1+\alpha\lambda x^{\beta}\right]\right) &, \alpha > 0, \\ \gamma_1\left(a, \lambda x^{\beta}\right), &, \alpha = 0, \end{cases} \tag{6}$$

where $\tau = (\alpha, \beta, \lambda)$. Clearly, if $\alpha \to 0$, the first expressions in (5) and (6) tend to the second ones in these equations. For $\alpha = 0$, the GEW model is identical to the *generalized gamma* distribution pioneered by Stacy (1962). Hereafter, a random variable X having pdf (5) is denoted by $X \sim GEW(\tau, a)$. Evidently, the density function (5) does not involve any complicated function and the EW distribution arises as the basic exemplar for $a = 1$. It is a positive point for the current generalization. The GEW model has several sub-models: the EW when $a = 1$, gamma Weibull (GW) when $\alpha = 0$, gamma extended exponential (GEE) when $\beta = 1$, extended exponential (EE) when $\beta = a = 1$, gamma exponential (GE) when $\alpha = 0$ and $\beta = 1$, Weibull (W) when $\alpha = 0$ and $a = 1$, and exponential when $\alpha = 0$ and $a = \beta = 1$. We show empirically the flexibility of the GEW distribution by comparing it with some of its sub-models by means of a real data set in Section 8. Plots of the density function and hrf of X for selected parameter values are displayed in Figs. 1 and 2, respectively.

The new distribution is easily simulated as follows: if V is a gamma random variable with parameter a, then

$$X = \left(\alpha^{-1} \lambda^{-1} \{\exp(\alpha V) - 1\}\right)^{1/\beta}$$

has the $GEW(\tau, a)$ distribution. This generating scheme is straightforward because of the existence of fast generators for gamma random variables.
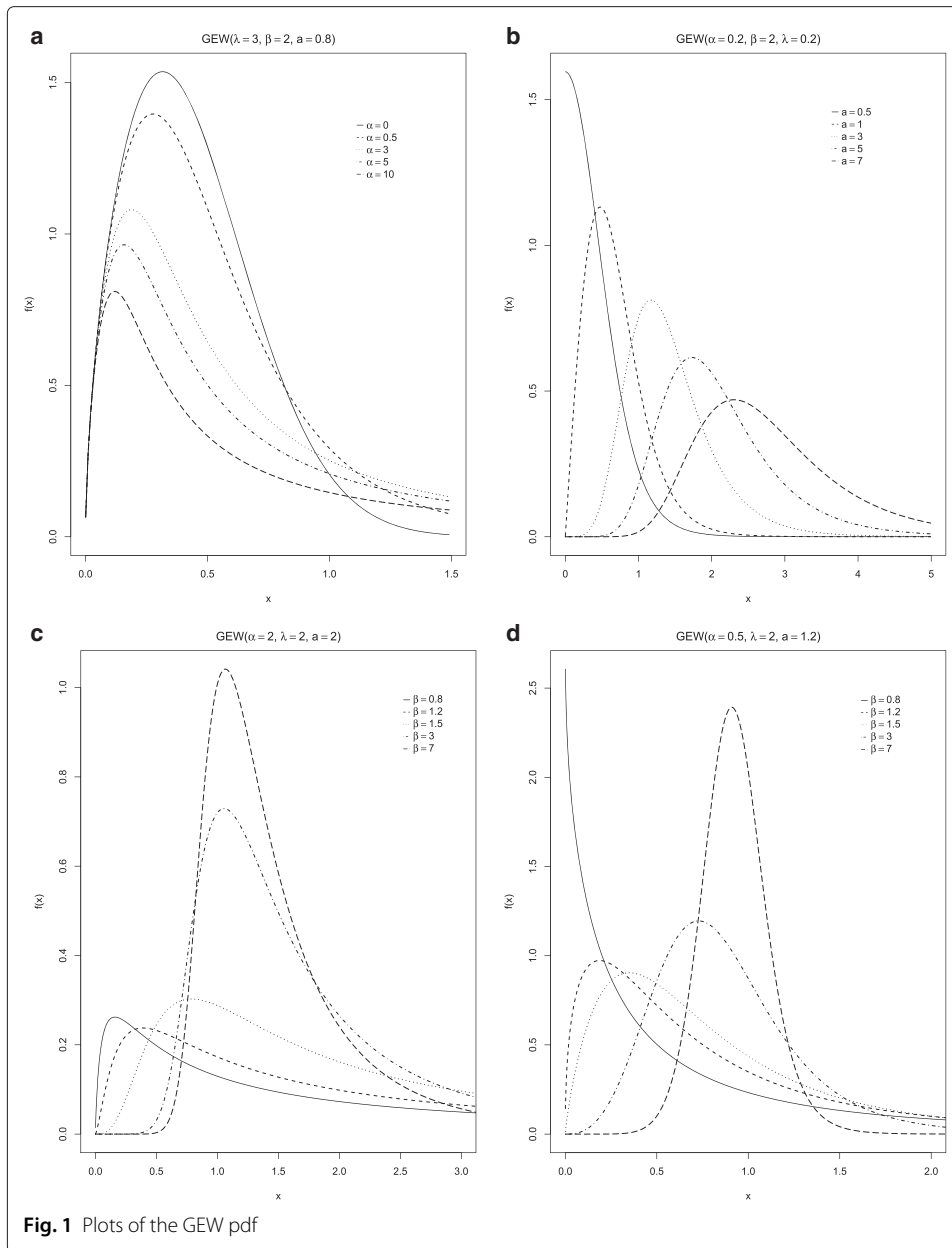
3   Useful expansions

Expansions for (5) and (6) can be derived using the concept of exponentiated distributions. The *exponentiated extended Weibull* (EEW) distribution is defined by raising the cdf (2) to a power $a > 0$ and the associated random variable is denoted by $Y \sim \text{EEW}(\tau, a)$. The cdf and pdf of Y are given by

$$H_a(x; \tau) = \begin{cases} \left[1 - \left(1 + \alpha\lambda x^{\beta}\right)^{-\frac{1}{\alpha}}\right]^a, & \alpha > 0, \\ \left[1 - \exp\left(-\lambda x^{\beta}\right)\right]^a, & \alpha = 0, \end{cases}$$

and

$$h_a(x; \tau) = \begin{cases} a\lambda\beta x^{\beta-1} \left(1 + \alpha\lambda x^{\beta}\right)^{-\frac{1}{\alpha}-1} \left[1 - \left(1 + \alpha\lambda x^{\beta}\right)^{-\frac{1}{\alpha}}\right]^{a-1}, & \alpha > 0, \\ a\lambda\beta x^{\beta-1} e^{-\lambda x^{\beta}} \left[1 - \exp\left(-\lambda x^{\beta}\right)\right]^{a-1} & , \alpha = 0, \end{cases} \tag{7}$$
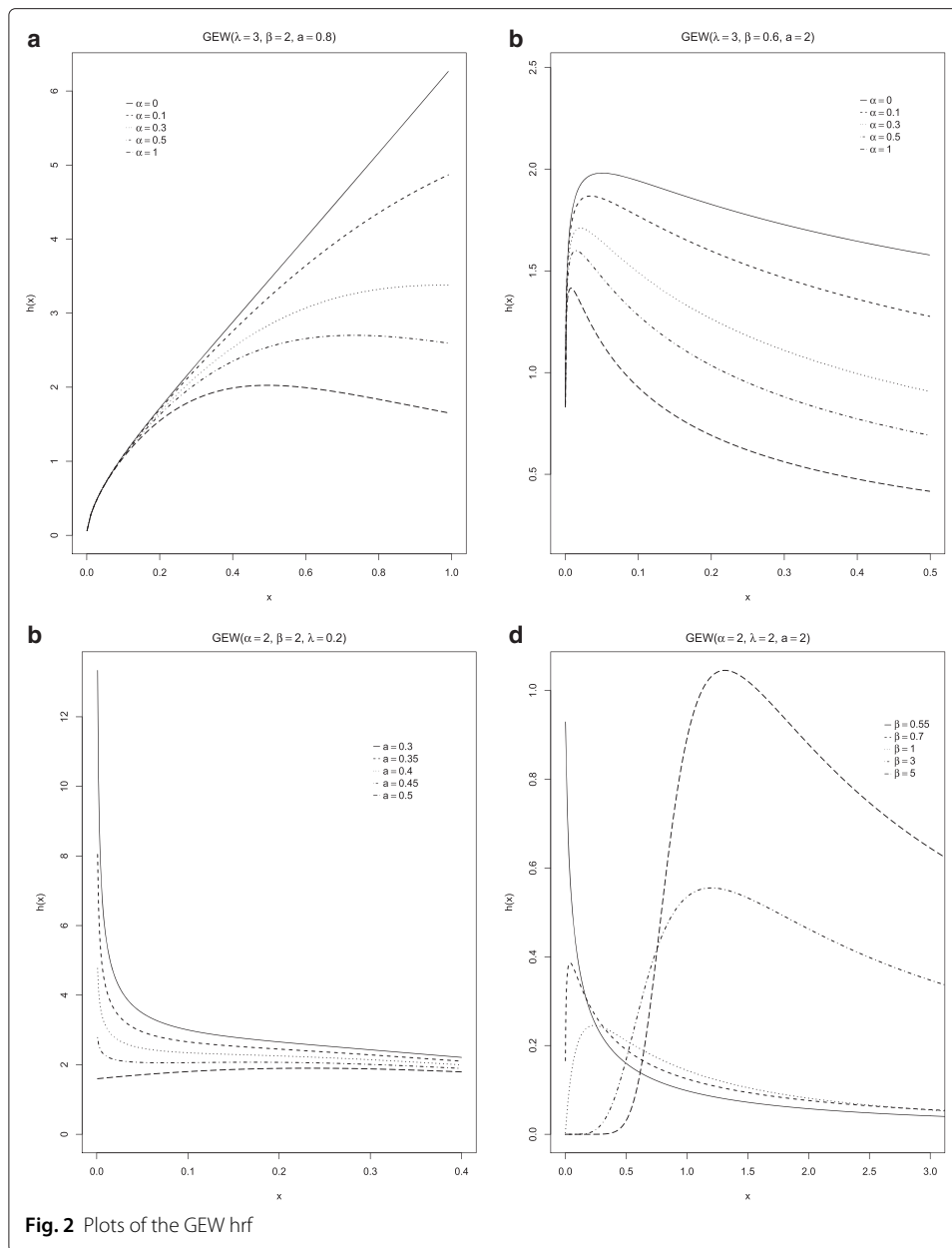
respectively. The properties of some exponentiated distributions have been investigated by several authors, see Mudholkar and Srivastava (1993) and Mudholkar et al. (1995) for exponentiated Weibull (ExpW), Gupta et al. (1998) for exponentiated Pareto, Gupta and

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 4 of 19

Fig. 1 Plots of the GEW pdf

Kundu (2001) for exponentiated exponential (ExpE) and Nadarajah and Gupta (2007) for exponentiated gamma (ExpG) distributions. More recently, Cordeiro et al. (2011) studied these properties for the exponentiated generalized gamma (ExpGG) distribution.

The following formula holds (http://functions.wolfram.com/ElementaryFunctions/Log/06/01/04/03/)

$$\left\{\frac{1}{\alpha} \log\left[1 - \alpha\lambda\, x^{\beta}\right]\right\}^{a-1}$$
$$= (a-1) \sum_{k=0}^{\infty} \binom{k+1-a}{k} \sum_{j=0}^{k} \frac{(-1)^{j+k}\, p_{j,k}}{(a-1-j)} \binom{k}{j} \left[1 - \left(1 - \alpha\lambda\, x^{\beta}\right)^{-\frac{1}{\alpha}}\right]^{a+k-1},$$

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 5 of 19

Fig. 2 Plots of the GEW hrf

where $a > 0$ is a real parameter and the constants $p_{j,k}$ can be determined recursively by

$$p_{j,k} = k^{-1} \sum_{m=1}^{k} \frac{(-1)^m \left[m(j+1) - k \right]}{(m+1)} \, p_{j,k-m}, \, j \geq 0,$$

for $k = 1, 2, \ldots$ and $p_{j,0} = 1$. Further, for any real parameter $a > 0$, we define

$$b_k = \frac{\binom{k+1-a}{k}}{(a+k)\Gamma(a-1)} \sum_{j=0}^{k} \frac{(-1)^{j+k} p_{j,k}}{(a-1-j)} \binom{k}{j}$$

and then the first equation in (5) can be expressed as

$$f(x) = \sum_{k=0}^{\infty} b_k \, h_{a+k}(x), \tag{8}$$

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 6 of 19

where $h_{a+k}(x) = (a+k)\,\lambda\,\beta\,x^{\beta-1}\left(1+\alpha\lambda\,x^\beta\right)^{-1/\alpha-1}\left[1-\left(1+\alpha\lambda\,x^\beta\right)^{-\frac{1}{\alpha}}\right]^{a+k-1}$ denotes the EEW$(\tau, a+k)$ density function. So, the GEW density function is a linear mixture of EEW densities.

Similarly, we can derive an expansion for the GW density when $\alpha = 0$. Using Eq. (5) for $\alpha = 0$, we obtain the same expression in (8), but the function $h_{a+k}(x)$ denotes, when $\alpha = 0$, the GEW$(\tau, a+k)$ pdf given by the second equation of (7).

Inserting (7) in Eq. (8), we obtain

$$f_{\text{GEW}}(x; \tau, a) = \begin{cases} \sum\limits_{j=0}^{\infty} e_j\, g_{\lambda^\star, \alpha^\star, \beta}(x)\ , & \text{if}\quad \alpha > 0\,, \\[2ex] \sum\limits_{j=0}^{\infty} e_j\, g_{\lambda^\star, \beta}(x)\ \ , & \text{if}\quad \alpha = 0\,, \end{cases} \tag{9}$$

where $e_j = \sum_{k=0}^{\infty} (-1)^j\,(a+k)\,b_k\binom{a+k-1}{j}$, $g_{\lambda^\star, \alpha^\star, \beta}(x)$ denotes the EW pdf with parameters $\lambda^\star = (j+1)\lambda$, $\alpha^* = \alpha/(j+1)$ and β and $g_{\lambda^\star, \beta}(x)$ denotes the Weibull pdf with parameters $\lambda^\star$ and β.

Equation (9) is the main result of this section. It reveals that the GEW density function is a linear mixture of EW (when $\alpha > 0$) and Weibull (when $\alpha = 0$) densities. So, several GEW structural properties can be obtained from those properties of the EW and Weibull distributions.

4 Quantile function

First, we consider the general case $\alpha > 0$. The GEW qf, say $Q(u) = F^{-1}(u)$, can be expressed in terms of the EW qf, say $Q_{EW}(\cdot)$. Inverting $F(x) = u$ given by (6), the qf of X (for $0 < u < 1$) follows as

$$F^{-1}(u) = Q_{GEW}(u) = Q_{EW}\left\{1 - \exp[-Q^{-1}(a, 1-u)]\right\}, \tag{10}$$

where $Q^{-1}(a, 1-u)$ is the inverse function of $Q(a, z) = 1 - \gamma_1(a, 1-u)$. A power series for it is given in the Wolfram website as

$$z = Q^{-1}(a, 1-u) = \sum_{i=0}^{\infty} a_i\, u^{i/a},$$

where $a_0 = 0$, $a_1 = \Gamma(a+1)^{1/a}$, $a_2 = \Gamma(a+1)^{2/a}/(a+1)$, etc. Quantities of interest can be obtained from (10) by substituting appropriate values for u.

Further, the EW qf is given by

$$Q_{EW}(u) = \left\{\frac{1 - (1-u)^\alpha}{\alpha\lambda(1-u)^\alpha}\right\}^{1/\beta}. \tag{11}$$

5 Moments and mean deviations

Some of the most important features and characteristics of a distribution can be studied through moments. We can write from Eq. (8)

$$\mu'_r = E(X^r) = \begin{cases} \alpha^{-\left(\frac{r}{\beta}+1\right)}\lambda^{-\frac{r}{\beta}}\sum\limits_{j=0}^{\infty} e_j^\star\, B\left(\frac{r}{\beta}+1, \frac{j+1}{\alpha} - \frac{r}{\beta}\right), & \text{if}\ \alpha > 0\ \text{and}\ r < \beta/\alpha\,, \\[2ex] \lambda^{-\frac{r}{\beta}}\,\Gamma\left(\frac{r}{\beta}+1\right)\sum\limits_{j=0}^{\infty}\frac{e_j}{(j+1)^{r/\beta}}, & \text{if}\ \alpha = 0\,, \end{cases} \tag{12}$$

where $e_j^\star = (j+1)\,e_j$, $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function and $\Gamma(a) = \int_0^\infty z^{a-1}e^{-z}dz$ is the gamma function.

Established algebraic expansions to determine μ_r' can be more efficient than computing these moments directly by numerical integration of (5), which can be prone to rounding errors among others. The skewness γ_1 and kurtosis γ_2 of X are easily obtained from the ordinary moments using well-known formulae. Figure 3 displays some plots for the skewness and kurtosis of the GEW model.

For lifetime models, it is of interest to know the rth incomplete moment of X defined by $T_r(y) = \int_0^y x^r f(x)dx$. Based on (9), we obtain (for $\alpha > 0$)
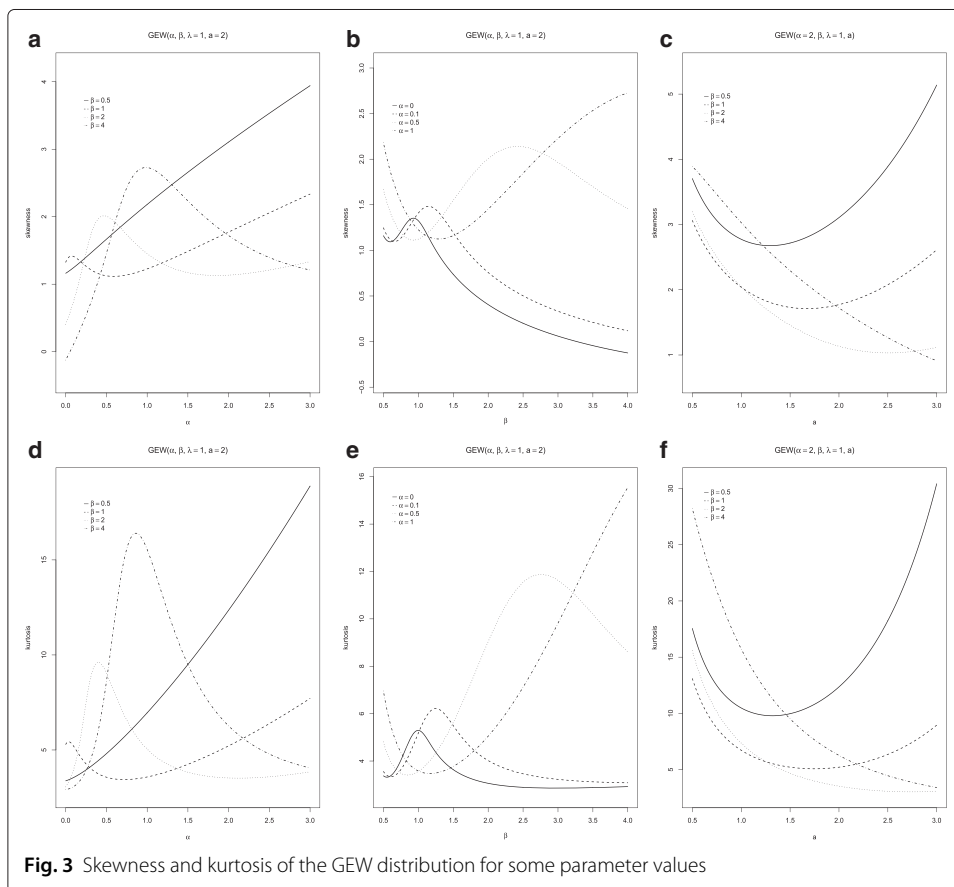
$$T_r(y) = \sum_{j=0}^\infty e_j\,\lambda^\star\,\beta\,\rho\left(y;r,\alpha^\star\lambda^\star,\alpha^{\star-1}\right),\tag{13}$$

where $\rho(y;r,p,q) = \int_0^y x^r\,(1+px)^{-q-1}dx$ for $r,p,q > 0$. Using Maple, this integral can be determined as

$$\rho(y;r,p,q) = A(r,p,q)\left\{2\,y^{r-q}\,{}_2F_1\left(q^{-r},q+1;q+1-r,-\frac{1}{py}\right)\left[B(r,p,q)+C(r,p,q)\right]\right\},$$

where ${}_2F_1$ is the hypergeometric function defined by

$$_2F_1(a,b;c;x) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{j=0}^\infty\frac{\Gamma(a+j)\Gamma(b+j)}{\Gamma(c+j)}\frac{x^j}{j!},$$

Fig. 3 Skewness and kurtosis of the GEW distribution for some parameter values

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 8 of 19

$$A(r, p, q) = \{p \sin[\pi(q - r)] \ (q - r) \ \Gamma(q + 1) \ \Gamma(r + 1 - q)\}^{-1},$$

$$B(r, p, q) = p^{-q} \ \Gamma(q + 1) \ \Gamma(r + 1 - q) \left[\cos(q\pi) \ \sin(\pi r) - \sin(q\pi) \ \cos(\pi r) \right],$$

and

$$C(r, p, q) = \pi \ p^{-r} \ \Gamma(r + 1)(q - r).$$

We obtain the incomplete moments by combining (13) and the expression for $\rho(y; r, p, q)$.

For $\alpha = 0$, the quantity $T_r(y)$ reduces to

$$T_r(y) = \lambda^{-r} \sum_{j=0}^{\infty} \frac{e_j}{(j + 1)^r} \ \gamma(r + 1, (j + 1)\lambda y).$$

An important application of $T_1(y)$ refers to the mean deviations about the mean and about the median, which can be used as measures of spread in a population. They are given by $\delta_1 = E\left(|X - \mu'_1|\right) = 2\mu'_1 F\left(\mu'_1\right) - 2 T_1\left(\mu'_1\right)$ and $\delta_2 = E(|X - m|) = \mu'_1 - 2 T_1(m)$, respectively, where the mean μ'_1 is determined from (12) and $T_1(\cdot)$ comes from (13) with $r = 1$ as

$$T_1(y) = \lambda \sum_{j=0}^{\infty} e_j \ \rho\left(y; 1, \alpha^{\star}\lambda^{\star}, \alpha^{\star^{-1}}\right).$$

Here, $\rho(y; 1, p, q)$ becomes

$$\rho(y; 1, p, q) = -\left[q(q - 1)p^2\right]^{-1} \left[\frac{(pqy + 1)}{(1 + py)^q} - 1\right].$$

Further, for $\alpha = 0$, we have

$$T_1(q) = \sum_{j=0}^{\infty} e_j \left(\frac{1}{\lambda^{\star}}\right)^{1/\beta} \int_0^{\lambda^{\star}q^{\beta}} v^{1/\beta} \ e^{-v} dv$$

and then $T_1(q)$ reduces to

$$T_1(q) = \sum_{j=0}^{\infty} e_j \left(\frac{1}{(j + 1)\lambda}\right)^{1/\beta} \left[\Gamma\left(1 + \frac{1}{\beta}\right) - \gamma\left(1 + \frac{1}{\beta}, (j + 1)\lambda q^{\beta}\right)\right].$$
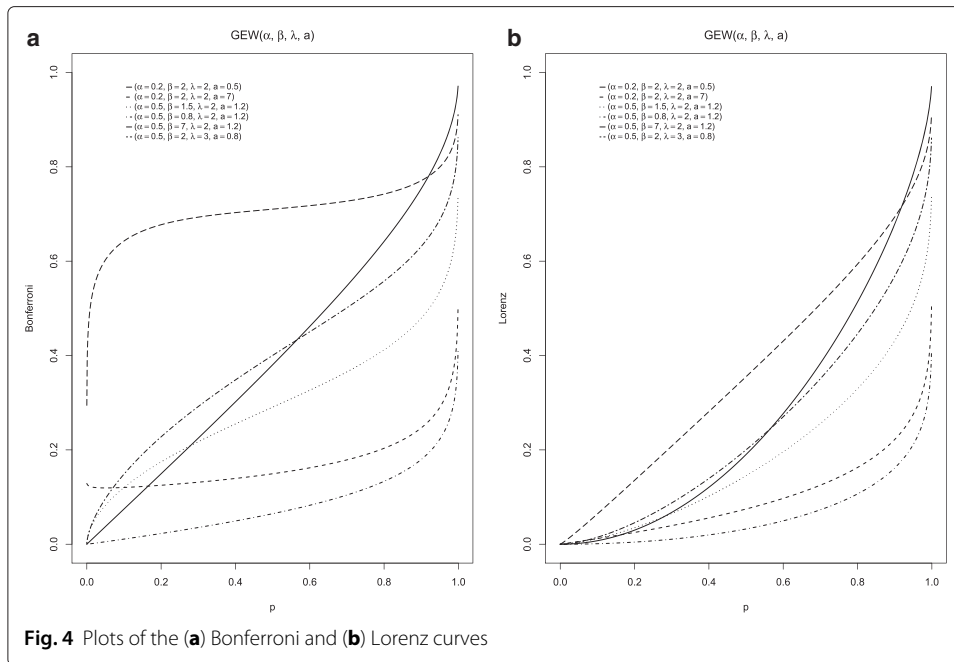
For a given probability π, both equations for $T_1(\cdot)$ can be used to determine the Lorenz and Bonferroni curves defined by $L(\pi) = T_1(q)/\mu'_1$ and $B(\pi) = T_1(q)/\left(\pi \ \mu'_1\right)$, respectively, where $q = Q_{GEW}(\pi)$ comes from Eq. (10) (see Fig. 4).

6 Estimation

The maximum likelihood method is used for estimating the parameters of the GEW model. We determine the maximum likelihood estimates (MLEs) from complete samples only. Let $x_1, \ldots, x_n$ be a sample of size n from the $GEW(\alpha, \beta, \lambda, a)$ distribution. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (\alpha, \lambda, \beta, a)^T$ is given by $l(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log[f(x_i)] = \sum_{i=1}^{n} l_i(\boldsymbol{\theta})$, where

$$l_i(\boldsymbol{\theta}) = \begin{cases} \log(\lambda) + \log(\beta) - \log[\Gamma(a)] - (a - 1)\log(\alpha) + (\beta - 1)\log(x_i) \\ - (1 + 1/\alpha)\log\left(1 + \alpha\lambda x_i^{\beta}\right) + (a - 1)\log\left[\log\left(1 + \alpha\lambda x_i^{\beta}\right)\right] & , \text{if} \quad \alpha > 0, \\ a\log(\lambda) + \log(\beta) - \log[\Gamma(a)] + (\beta a - 1)\log(x_i) - \lambda x_i^{\beta} & , \text{if} \quad \alpha = 0. \end{cases}$$

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 9 of 19

Fig. 4 Plots of the (**a**) Bonferroni and (**b**) Lorenz curves

The expression for $l(\boldsymbol{\theta})$ when $\alpha = 0$ gives the log-likelihood function for the GG distribution, a sub-model of the GEW distribution. Next, we maximize $l(\boldsymbol{\theta}) = \sum_{i=1}^{n} l_i(\boldsymbol{\theta})$ for $\alpha > 0$.

The components of the score vector $U(\boldsymbol{\theta})$ are given by $\sum_{i=1}^{n} \partial l_i(\boldsymbol{\theta}) / \partial \theta_j$, $j = 1, \ldots, 4$, where the θ_j's denote the parameters α, β, λ or a. The elements $\partial l_i(\boldsymbol{\theta}) / \partial \theta_j$, $j = 1, \ldots, 4$, are given below

$$\partial l_i(\boldsymbol{\theta}) / \partial \alpha = -\frac{a-1}{\alpha} + \frac{\log\left(1 + \alpha\,\lambda\,x_i^{\beta}\right)}{\alpha^2} - \frac{\left(\frac{1}{\alpha}+1\right)\lambda\,x_i^{\beta}}{1 + \alpha\,\lambda\,x_i^{\beta}} + \frac{(a-1)\,\lambda\,x_i^{\beta}}{\left(1 + \alpha\,\lambda\,x_i^{\beta}\right)\log\left(1 + \alpha\,\lambda\,x_i^{\beta}\right)}\,,$$

$$\partial l_i(\boldsymbol{\theta}) / \partial \beta = \frac{1}{\beta} + \log(x_i) - \frac{\left(\frac{1}{\alpha}+1\right)\alpha\,\lambda\,x_i^{\beta}\,\log(x_i)}{1 + \alpha\,\lambda\,x_i^{\beta}} + \frac{(a-1)\,\alpha\,\lambda\,x_i^{\beta}\,\log(x_i)}{\left(1 + \alpha\,\lambda\,x_i^{\beta}\right)\log\left(1 + \alpha\,\lambda\,x_i^{\beta}\right)}\,,$$

$$\partial l_i(\boldsymbol{\theta}) / \partial \lambda = \frac{1}{\lambda} - \frac{\left(\frac{1}{\alpha}+1\right)\alpha\,x_i^{\beta}}{1 + \alpha\,\lambda\,x_i^{\beta}} + \frac{(a-1)\,\alpha\,x_i^{\beta}}{\left(1 + \alpha\,\lambda\,x_i^{\beta}\right)\log\left(1 + \alpha\,\lambda\,x_i^{\beta}\right)}\,,$$

$$\partial l_i(\boldsymbol{\theta}) / \partial a = -\psi(a) - \log(\alpha) + \log\left[\log\left(1 + \alpha\,\lambda\,x_i^{\beta}\right)\right]\,,$$

where $\psi(\cdot)$ is the digamma function.

Setting these equations to zero and solving them simultaneously yield the MLEs of the four parameters. They can be solved numerically by using the R-language or any iterative methods such as the NR (Newton-Raphson), BFGS (Broyden-Fletcher-Goldfarb-Shanno), BHHH (Berndt-Hall-Hall-Hausman), NM (Nelder-Mead), SANN (Simulated-Annealing) and Limited-Memory Quasi-Newton code for Bound-Constrained Optimization (L-BFGS-B). Regarding the initial values for the parameters, several authors (Ugray et al. 2007; Glover 1998; Dixon and Szegö 1978; Varadhan and Gilbert 2009) suggest to find local maxima starting from widely varying starting points and then pick the maximum out of these, i.e., they suggest the use of multiple starting points (commonly referred to as

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 10 of 19

multistart optimization). We followed their advice in this paper. As a useful tool we suggest the use of the R rotine `multiStart` of R package `BB` (for solving and optimizing large-scale nonlinear systems) for dealing with multiple starting points to obtain multiple solutions and to test sensitivity to starting values.

For interval estimation of the model parameters, we require the 4×4 total observed information matrix $J(\theta)$. The elements of $J(\theta) = \{J_{rs}\}$, where $r, s \in \{\alpha, \beta, \lambda, a\}$, can be obtained from the authors upon request. The multivariate normal $N_4\left(0, J(\widehat{\theta})^{-1}\right)$ distribution, where $J(\widehat{\theta})^{-1}$ is the inverse observed matrix evaluated at $\theta = \widehat{\theta}$, can be used to construct approximate confidence intervals for the parameters.
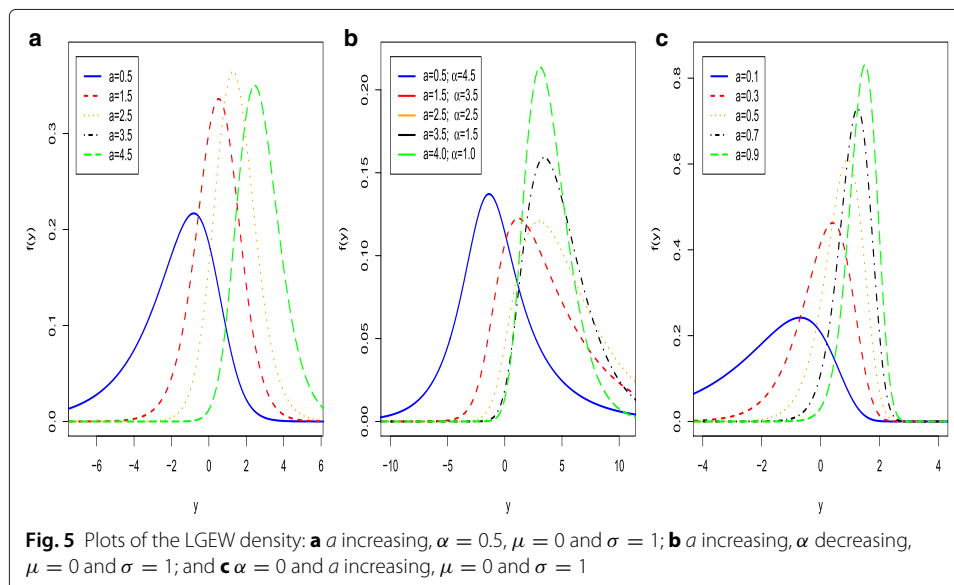
The likelihood ratio (LR) statistic is useful for comparing the GEW distribution with some of its special models. We can evaluate the maximum values of the unrestricted and restricted log-likelihoods to obtain LR statistics for testing some of its sub-models. In any case, hypothesis tests of the type $H_0 : \boldsymbol{\psi} = \boldsymbol{\psi}_0$ versus $H : \boldsymbol{\psi} \neq \boldsymbol{\psi}_0$, where $\boldsymbol{\psi}$ is a vector formed with some components of $\boldsymbol{\theta}$ and $\boldsymbol{\psi}_0$ is a specified vector, can be performed using LR statistics.

7 The LGEW regression model

If X is a random variable having the GEW density function (5), then $Y = \log(X)$ has the *log-gamma extended Weibull* (LGEW) distribution. The density function of Y, parameterized in terms of $\beta = \sigma^{-1}$ and $\lambda = e^{-\mu/\sigma}$, is given by

$$f(y; a, \alpha, \mu, \sigma) = \begin{cases} \frac{\exp\left(\frac{y-\mu}{\sigma}\right)\left[1+\alpha\exp\left(\frac{y-\mu}{\sigma}\right)\right]^{\left(-\frac{1}{\alpha}-1\right)}}{\sigma\,\alpha^{a-1}\Gamma(a)} \left\{\log\left[1+\alpha\exp\left(\frac{y-\mu}{\sigma}\right)\right]\right\}^{a-1} & , \text{if } \alpha > 0, \\ \frac{\exp\left[a\left(\frac{y-\mu}{\sigma}\right)-\exp\left(\frac{y-\mu}{\sigma}\right)\right]}{\sigma\,\Gamma(a)} & , \text{if } \alpha = 0, \end{cases} \tag{14}$$

where $y \in \mathbb{R}$, $a > 0$ and $\alpha > 0$ are shape parameters, $\sigma > 0$ and $\mu \in \mathbb{R}$. If Y is a random variable having density function (14), we write $Y \sim \text{LGEW}(a, \alpha, \mu, \sigma)$. Plots of the density function of Y for selected parameter values are displayed in Fig. 5. They reveal great flexibility for different values of the shape parameters a and α.

Fig. 5 Plots of the LGEW density: **a** *a* increasing, $\alpha = 0.5$, $\mu = 0$ and $\sigma = 1$; **b** *a* increasing, α decreasing, $\mu = 0$ and $\sigma = 1$; and **c** $\alpha = 0$ and *a* increasing, $\mu = 0$ and $\sigma = 1$

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 11 of 19

Thus,

$$\text{if} \quad X \sim \text{GEW}(a, \alpha, \beta, \lambda) \quad \text{then} \quad Y = \log(X) \sim \text{LGEW}(a, \alpha, \mu, \sigma).$$

The corresponding survival function of Y is given by

$$S(y; a, \alpha, \mu, \sigma) = \begin{cases} 1 - \gamma_1\left(a, \frac{1}{\alpha}\log\left[1 + \alpha\exp\left(\frac{y-\mu}{\sigma}\right)\right]\right) & , \text{if} \quad \alpha > 0, \\ 1 - \gamma_1\left(a, \exp\left(\frac{y-\mu}{\sigma}\right)\right) & , \text{if} \quad \alpha = 0. \end{cases}$$

The pdf of the standardized random variable $Z = (Y - \mu)/\sigma$ is given by

$$\pi(z; a, \alpha) = \begin{cases} \frac{\exp(z)[1 + \alpha\exp(z)]^{-\frac{1}{\alpha}-1}}{\alpha^{a-1}\Gamma(a)}\left\{\log[\,1 + \alpha\exp(z)\,]\right\}^{a-1} & , \text{if} \quad \alpha > 0, \\ \frac{\exp[a\,z - \exp(z)]}{\Gamma(a)} & , \text{if} \quad \alpha = 0. \end{cases} \qquad (15)$$

The special case $\alpha = 0$ and $a = 1$ refers to the log-Weibull (LW) (or extreme-value) distribution and, for $\alpha = 0$, we obtain the log-gamma-Weibull (LGW) model.

In many practical applications, the lifetimes x_i are affected by explanatory variables such as the cholesterol level, blood pressure and many others. Let $\mathbf{v}_i = (v_{i1}, \ldots, v_{ip})^T$ be the explanatory variable vector associated with the ith response variable y_i for $i = 1, \ldots, n$. Consider a sample $(y_1, \mathbf{v}_1), \ldots, (y_n, \mathbf{v}_n)$ of n independent observations, where each random response is defined by $y_i = \min\{\log(x_i), \log(c_i)\}$, and $\log(x_i)$ and $\log(c_i)$ are the log-lifetime and log-censoring, respectively. We consider non-informative censoring such that the observed lifetimes and censoring times are independent.

We construct a linear regression model for the response variable y_i based on the LGEW distribution given by

$$y_i = \mathbf{v}_i^T \boldsymbol{\beta} + \sigma z_i, \ i = 1, \ldots, n, \qquad (16)$$

where the random error z_i has the density function (15), $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$, $\sigma > 0$, $a > 0$ and $\alpha > 0$ are unknown scalar parameters and $\mathbf{v}_i$ is the vector of explanatory variables modeling the location parameter $\mu_i = \mathbf{v}_i^T \boldsymbol{\beta}$. Hence, the location parameter vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^T$ of the LGEW model has a linear structure $\boldsymbol{\mu} = \mathbf{v}^T \boldsymbol{\beta}$, where $\mathbf{v} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)^T$ is a known model matrix. The log-Weibull (or the extreme value) regression model is defined by (16) with $\alpha = 0$ and $a = 1$.

Let F and C be the sets of individuals for which y_i is the log-lifetime or log-censoring, respectively. The total log-likelihood function for the model parameters $\boldsymbol{\theta} = \left(a, \alpha, \sigma, \boldsymbol{\beta}^T\right)^T$ can be obtained from Eqs. (15) and (16) as

$$l(\boldsymbol{\theta}) = \begin{cases} -q\{\log(\sigma) + (a-1)\log(\alpha) - \log[\,\Gamma(a)\,]\} + \sum\limits_{i \in F} z_i \\ \quad -\left(\frac{1}{\alpha} + 1\right)\sum\limits_{i \in F}\log\left[1 + \alpha\exp(z_i)\right] \\ \quad +(a-1)\sum\limits_{i \in F}\log\{\log[\,1 + \alpha\exp(z_i)\,]\} \\ \quad +\sum\limits_{i \in C}\log\left\{1 - \gamma_1\left(a, \frac{1}{\alpha}\log[\,1 + \alpha\exp(z_i)\,]\right)\right\} & , \text{if} \quad \alpha > 0, \\ -q\{\log(\sigma) + \log[\,\Gamma(a)\,]\} + a\sum\limits_{i \in F} z_i - \sum\limits_{i \in F}\exp(z_i) \\ \quad +\sum\limits_{i \in C}\log\{1 - \gamma_1(a, \exp[\,z_i\,])\} & , \text{if} \quad \alpha = 0, \end{cases} \qquad (17)$$

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 12 of 19

where q is the observed number of failures and $z_i = \left(y_i - \mathbf{v}_i^T \boldsymbol{\beta}\right)/\sigma$. The MLE $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ can be obtained by maximizing the log-likelihood function (17). We use the procedure NLMixed in SAS to evaluate the estimate $\widehat{\boldsymbol{\theta}}$. Initial values for $\boldsymbol{\beta}$ and σ are taken from the fit of the LW regression model with $\alpha = 0$ and $a = 1$.

From the fitted model (16), the survival function for y_i can be estimated by

$$
\widehat{S}\left(y_i; \hat{a}, \hat{\alpha}, \hat{\sigma}, \widehat{\boldsymbol{\beta}}^T\right) =
\begin{cases}
1 - \gamma_1 \left(\hat{a}, \frac{1}{\hat{\alpha}} \log\left[1 + \hat{\alpha} \exp\left(\frac{y_i - \mathbf{v}_i^T \widehat{\boldsymbol{\beta}}}{\hat{\sigma}}\right)\right]\right) , & \text{if} \quad \alpha > 0, \\[4mm]
1 - \gamma_1 \left(\hat{a}, \exp\left[\frac{y_i - \mathbf{v}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right]\right) & \text{if} \quad \alpha = 0.
\end{cases}
\tag{18}
$$

Under general regularity conditions, the asymptotic distribution of $(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is multivariate normal $N_{p+3}\left(0, K(\boldsymbol{\theta})^{-1}\right)$, where $K(\boldsymbol{\theta})$ is the $(p+3) \times (p+3)$ expected information matrix. The asymptotic covariance matrix $K(\boldsymbol{\theta})^{-1}$ of $\widehat{\boldsymbol{\theta}}$ can be estimated by $\widehat{\boldsymbol{\theta}}^{-1}$ and then the inference on the parameter vector $\boldsymbol{\theta}$ can be based on the normal approximation $N_{p+3}\left(0, J(\widehat{\boldsymbol{\theta}})^{-1}\right)$ for $\widehat{\boldsymbol{\theta}}$. In fact, an $100(1 - \eta)\%$ asymptotic confidence interval for each parameter θ_r is given by

$$
ACI_r = \left(\hat{\theta}_r - z_{\eta/2}\sqrt{-\widehat{J}^{r,r}}, \hat{\theta}_r + z_{\eta/2}\sqrt{-\widehat{J}^{r,r}}\right),
$$

where $-\widehat{J}^{r,r}$ represents the rth diagonal element of $J(\widehat{\boldsymbol{\theta}})^{-1}$ and $z_{\eta/2}$ is the quantile $1 - \eta/2$ of the standard normal distribution. The LR statistic can be used to discriminate between the LGEW and LW regression models since they are nested models. In this case, the hypotheses to be tested are $H_0 : (a, \alpha)^T = (1, 0)^T$ versus $H_1 : H_0$ *is not true*, and the LR statistic reduces to $w = 2\left\{l(\widehat{\boldsymbol{\theta}}) - l(\widetilde{\boldsymbol{\theta}})\right\}$, where $\widetilde{\boldsymbol{\theta}}$ is the MLE of $\boldsymbol{\theta}$ under H_0. The null hypothesis is rejected if $w > \chi^2_{2,1-\eta}$, where $\chi^2_{2,1-\eta}$ is the quantile of the chi-square distribution with two degrees of freedom.

8 Applications

8.1 The GEW model

In the first application, we use the warp break data for six types of weaving warps discussed by Tippett (1950, p. 105). We describe his experiment: "The results of a weaving experiment was conducted in a factory. There were 6 lots of warp yarn labelled respectively AL, AM, etc. They were spun from two growths of cotton, A and B, and each cotton was spun to three twists (i.e., the number of turns in the yarn per inch): low (L), medium (M), and high (H). The combination of these three factors give 6 kinds of yarn, which are the experimental treatments. From each yarn were prepared 9 warps (a warp is a quantity of warp yarn that goes into one loom as a unit), and, as a loom came available in the course of events, a warp chosen at random from the 54 was assigned to it, until ultimately all 54 were disposed of. More than one warp was woven in some looms, but that did not affect the randomness of the distribution. The number of warp threads that broke during the weaving of each warp was counted and expressed as a rate of so many breaks per unit of warp."

We analyze the warp breakage rates for individual warps disregarding the factors. First, we fit the GEW model and some of its sub-models to these data by the method of maximum likelihood. Afterwards, we compare the GEW model with some four-parameter competitive models.

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 13 of 19

The GEW model, its sub-models and some competitors

We compare the GEW model with some of its sub-models and we also compare it with some competitor models: the Burr XII geometric (BXIIG) (Silva and Cordeiro 2015), the beta log-normal (BLN) (Castellares et al. 2013) and the exponentiated Weibull-Poisson (EWP) (Mahmoudi and Sepahdar 2013) distributions. The MLEs of the models parameters and the Akaike Information Criterion (AIC) statistic for some models fitted to the data are listed in Table 1. In such table the parameters used for the competitor models are the same in the cited references. However, for the BXIIG we included the scale parameter s that was not considered in Silva and Cordeiro (2015). Even though it is expected that the GEW model will provide a better fit than its sub-models, we can verify whether or not there are significant differences. Some model comparison is given in Table 2.

The histogram of the data and the plots of the fitted densities for some models are displayed in Fig. 6. They indicate that the new distribution provides a better fit than other sub-models. A comparison of the new distribution with four sub-models using LR statistics is performed in Table 2. These statistics indicate that the new distribution is the most adequate model to describe the data.

In the case of the competitor models, the AIC values suggest that the GEW model is quite comparable to the BLN and EWP models (being slightly superior) and it is a bit better than the BXIIG model. Figure 7 displays the histogram of the data and the fitted GEW, BXIIG, BLN and EWP densities. The last comments about the AIC are in agreement with the general notion that the GEW density provides the best fit for the current data, although all of them are very close.

The required numerical evaluations are implemented by using an R script (sub-routine `nlminb` that can be found at `https://cran.r-project.org`). The `nlminb()` is a derivative-free method for function minimization as it requires only the function to be

Table 1 MLEs of the model parameters for the warp breakage rate data, the corresponding SEs (given in parentheses) and the AIC measure

Model	α	λ	β	a	AIC
GEW	0.895137	0.000026	9.792369	23.609396	420.257020
	(0.042691)	(0.000011)	(0.714861)	(2.615331)	
GW	0	0.008219	1.622223	2.061746	425.080529
	(-)	(0.000971)	(0.025583)	(0.071981)	
EW	0.104458	0.000192	2.499730	1	427.270840
	(0.025452)	(0.000032)	(0.052776)	(-)	
GEE	0.000961	0.190944	1	5.364537	421.924911
	(0.013883)	(0.026556)	(-)	(0.502296)	
GE	0	0.250448	1	6.896945	422.114535
	(-)	(0.006646)	(-)	(0.176429)	
Model	k	c	θ	s	AIC
BXIIG	1.430692	3.850236	0.885467	49.00841	422.653656
	(0.530019)	(0.094952)	(0.055033)	(11.093854)	
Model	μ	σ	a	b	AIC
BLN	1.290503	1.322159	30.985635	2.704164	420.316246
	(0.668315)	(0.373892)	(21.297427)	(1.327947)	
Model	α	θ	β	γ	AIC
EWP	14.50818	0.90275	0.13294	0.84268	420.51673
	(2.69087)	(0.27133)	(0.02132)	(0.04587)	

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 14 of 19

Table 2 LR tests for the warp breakage rate data

Model	Hypotheses	Statistic LR	*p*-value
GEW vs *GW*	$H_0 : \alpha = 0$ vs $H_1 : H_0$ is false	6.823509	0.008996563
GEW vs *EW*	$H_0 : a = 1$ vs $H_1 : H_0$ is false	9.013820	0.002679457
GEW vs *GEE*	$H_0 : \beta = 1$ vs $H_1 : H_0$ is false	3.667891	0.055470343
GEW vs *GE*	$H_0 : \alpha = 0, \beta = 1$ vs $H_1 : H_0$ is false	5.857515	0.053463433

evaluated. Other R routine choices such as the `optimx` package (Nash and Varadhan, 2011) provided similar results. The R rotine `multiStart` of R package `BB` is also used in this work. The data set `warpbreaks` is available in the R data frame.
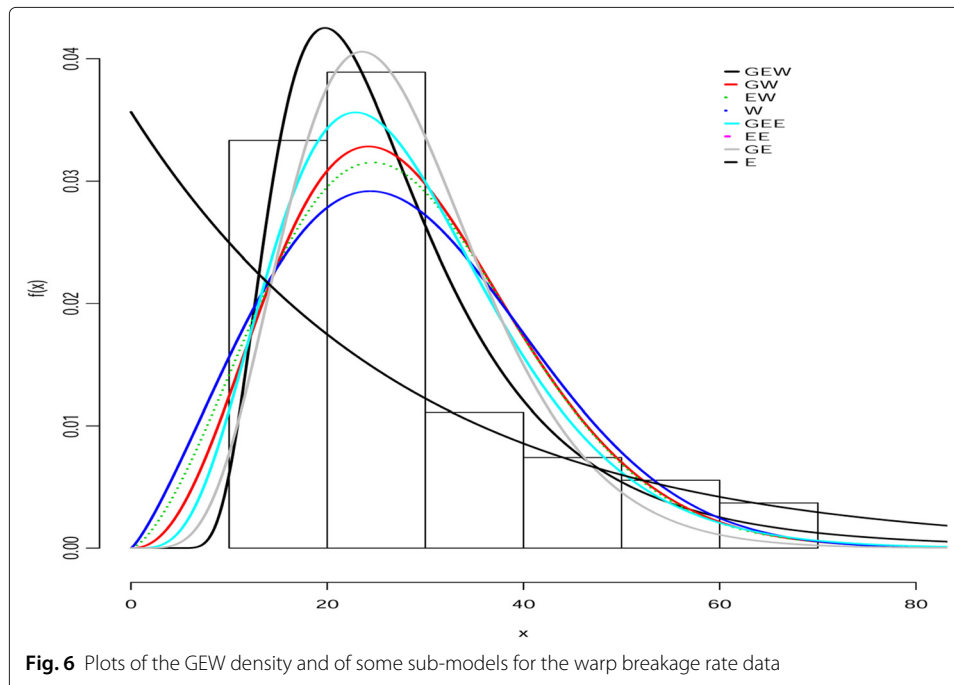
Some estimation issues

We now discuss some estimation issues related to the GEW distribution. The GEW density is given by (5) and, when $\alpha \rightarrow 0$, the first expression in (5) tends to the second one. For $\alpha = 0$, the GEW distribution is identical to the generalized gamma distribution. Therefore, according to Theorem 2.1 in Mäkeläinen et al. (1981), the GEW density is not constant on the boundary since
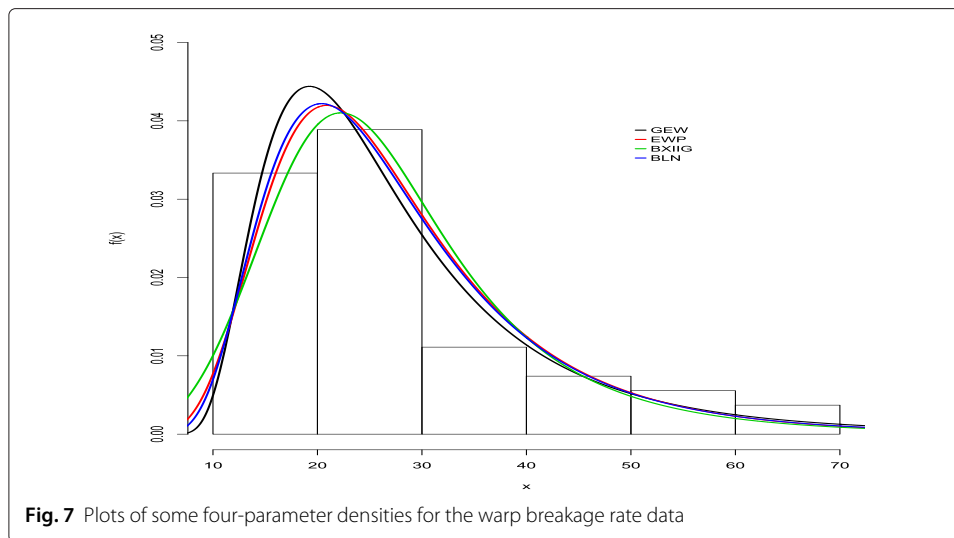
$$\lim_{\alpha \to 0} f(x; \tau, a) = \frac{\lambda^a \, \beta \, x^{\beta a - 1} \, e^{-\lambda x^\beta}}{\Gamma(a)}.$$

This fact implies that the MLE $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta} = (\alpha, \lambda, \beta, a)$ of the GEW model may not be unique.

As mentioned before, in order to handle the case of multiple solutions we adopt different initial points, say *k*, in the maximization procedure and use the estimate which maximizes the log-likelihood among these *k* values. For the current data sets, we have always been able to find a solution.

Fig. 6 Plots of the GEW density and of some sub-models for the warp breakage rate data

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 15 of 19

Fig. 7 Plots of some four-parameter densities for the warp breakage rate data

As mentioned above, the GEW distribution has different expressions depending on whether $\alpha > 0$ or $\alpha = 0$. In order to study possible estimation difficulties for this model when the true value of α is either zero or around it, we simulate data from the gamma generalized distribution, i.e., by taking $\alpha = 0$ for some choices of the parameters a, λ and β and then use the GEW model in the estimation procedure. We note that the GEW model is able to estimate all the parameter values correctly. As expected, for larger samples, the estimated values tend to be closer to the true values. Of course, for any similar problem, the program may fail to converge for some choices of initial values.

We also perform some additional analyses in order to evaluate the GEW robustness. We simulate two different data sets from the beta-Weibull distribution (which does not belong to the GEW family) and use them to fit the GEW distribution. The resulting fits (not shown here) indicate that the GEW distribution is very robust. In fact, the estimated GEW densities capture the main aspects of the beta-Weibull generated data in the sense that they are able to correctly locate the mode of the histograms obtained from the generated data, inflection points and other data characteristics.

Regarding the impact of the sample size on the estimated values of the parameters, a simulation study is conducted in which we generate GEW data considering $(\alpha, \lambda, \beta, a) = (0.895137, 0.000026, 9.792369, 23.609396)$ as the true parameter values. We simulate data by taking $k = 500$ replicates for each of the sample sizes $n = 50, 100, 500, 1000, 5000$ and $n = 10,000$. For a given sample size and $k = 500$ estimated values we evaluate the average of those estimated vectors and the mean squared errors (MSEs). We can conclude that the estimated expected vector does approach the true vector, but the MSE decreases slowly.

8.2 The LGEW model

We now illustrate some of the ideas and methodology of regression models for the LGEW model using the data set from a two-arm clinical trial considered earlier by Efron (1988) and Mudholkar et al. (1996). Efron observed that the empirical hazard functions for both samples start near zero, suggesting an initial high-risk period in the beginning, a decline for a while, and then stabilization after about one year. Specifically, Efron's data from a head and neck cancer clinical trial consist of survival times of 51 patients in arm A who were given radiation therapy and 45 patients in arm B who were given radiation plus

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 16 of 19

chemotherapy. Nine patients in arm A and 14 patients in arm B were lost to follow-up and were regarded as censored. In this paper, we consider only one predictor: ($v1$): two-Arm ($Arm\ A = 0$, $Arm\ B = 1$).

We fit the LGEW regression model

$$y_i = \beta_0 + \beta_1 v_{i1} + \sigma z_i, \tag{19}$$

where the errors $z_1, \ldots, z_{96}$ are independent random variables with density function (15).

An alternative approach for modeling these data can be provided by the LW distribution. There are various extensions of this lifetime distribution; see, for example, the log-beta Weibull (LBW) (Cordeiro et al. 2013) and Kumaraswamy-logistic (KwL) (Santana et al. 2012) distributions, among others. The LBW and KwL density functions are given below:

- **LBW distribution**

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

  where $-\infty < y < \infty$, $\sigma > 0$ and $-\infty < \mu < \infty$. See, for example, more details and properties in Cordeiro et al. (2013) and Ortega et al. (2015).

- **KwL distribution**

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

$$\times \left\{ 1 - \left[1 - \frac{1}{1 + \exp\left(\frac{y-\mu}{\sigma} \right)} \right]^a \right\}^{b-1},$$

  where $-\infty < y < \infty$, $\sigma > 0$ and $-\infty < \mu < \infty$. Some applications of the KwL distribution are discussed in Santana et al. (2012) and Nadarajah et al. (2012).

Taking $\mu = \beta_0 + \beta_1 v_{i1}$ for the location parameter, σ is the dispersion parameter and a and b are shape parameters.

Table 3 MLEs of the parameters from some fitted regression models to the Efron's data, the corresponding SEs (given in parentheses), *p*-values in [.] and the AIC and BIC measures

Model	a	α	σ	β_0	β_1	AIC	BIC
LGEW	0.3101	32.5577	0.1420	5.6000	−0.0868	281.4	294.2
	(0.0904)	(21.3182)	(0.0405)	(0.1706)	(0.1654)		
				[<0.0001]	[0.6010]		
LGW	0.0264	0	0.0510	8.0113	−0.6274	339.7	349.9
	(0.0079)		(0.0152)	(0.1458)	(0.1608)		
				[<0.0001]	[0.0002]		
LW	1	0	1.1800	6.7873	−0.7586	312.6	320.3
			(0.1082)	(0.2088)	(0.2803)		
				[<0.0001]	[0.0089]		
	a	b	σ	β_0	β_1	AIC	BIC
LBW	167.06	14.0059	12.8338	−6.0419	−0.5836	299.4	312.2
	(0.3987)	(0.0682)	(1.1065)	(1.0492)	(0.2726)		
				[<0.0001]	[0.0348]		
KwL	16.2819	312.91	5.8106	1.6912	−0.6834	308.0	320.8
	(2.6189)	(24.86)	(0.3258)	(6.0235)	(0.2792)		
				[0.7795]	[0.0162]		

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 17 of 19

Table 4 LR tests for the Efron's data

Model	Hypotheses	Statistic LR	*p*-value
LGEW vs LGW	$H_0 : \alpha = 0$ vs $H_1 : H_0$ is false	60.3	<0.0001
LGEW vs LW	$H_0 : (a, \alpha)^T = (1, 0)^T$ vs $H_1 : H_0$ is false	35.2	<0.0001

Table 3 lists the MLEs of the parameters for the fitted LGEW, LGW, LW, LBW and KwL regression models to the current data using the NLMixed procedure in SAS. Initial values for β and σ are taken from the fitted LW regression model with $\alpha = 0$ and $a = 1$.

A comparison of the new distribution with two of its sub-models using LR statistics is performed in Table 4. These statistics indicate that the LGEW regression model is the most adequate model to explain the data.

The LGEW model involves an extra parameter, which gives it more flexibility to fit the data. We note from the fitted LGEW regression model that the dummy variable v_1 is not significant at 5 %. We note that there is no significant difference between "Arm A" and "Arm B" clinical trial for the survival times. The LGEW regression model outperforms the other models irrespective of the criteria and it can be used effectively in the analysis of these data.

A graphical comparison of the fitted LGEW, LGW and LW models (see Fig. 8) indicates that the LGEW model provides a superior fit. The curves displayed in Fig. 8 represent the empirical survival function and the estimated survival functions obtained from Eq. (18). Finally, we conclude that the LGEW regression model provides a good fit to these data.

9 Conclusions

We introduce a new model named the gamma extended Weibull (GEW) distribution and study some of its structural properties. It generalizes some important distributions in the literature and provides means of its continuous extension to still more complex situations. The new model contains several distributions as special models including the extended Weibull (Mudholkar et al. 1996), gamma Weibull (Zografos and Balakrishnan 2009) and

Fig. 8 Estimated survival functions and the empirical survival for Efron's data. **a** The LGEW versus LGW regression models. **b** The LGEW versus LW regression models

Cordeiro *et al. Journal of Statistical Distributions and Applications* (2016) 3:7

Page 18 of 19
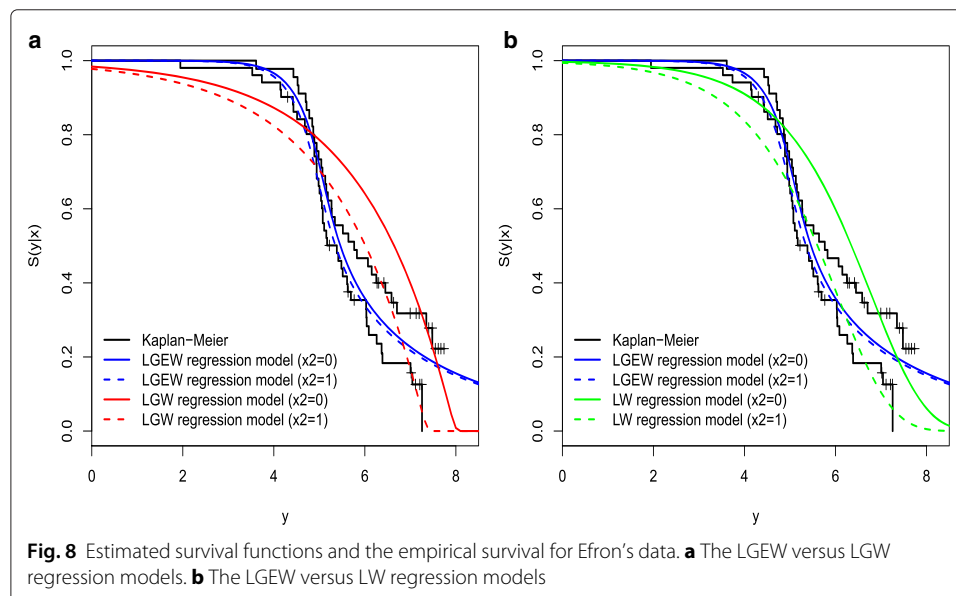
generalized gamma (Stacy 1962). We derive explicit expressions for the density function, ordinary and incomplete moments, quantile function and mean deviations. The model parameters are estimated by maximum likelihood. The usefulness of the GEW distribution is illustrated by means of an application to real data, where we show empirically that it gives a better fit than some of its sub-models. We also propose the log-gamma extended Weibull (LGEW) regression model, which has greater flexibility as shown by means of an application to real data.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
GMC proposed the model, wrote sections 1, 2, 4, 5, 9 (1, 2 and 4 jointly with MCSL, and 5 jointly with AEG) and drafted the manuscript. MCSL wrote and obtained the results in sections 1 to 4 (1, 2 and 4 jointly with GMC). AEG wrote and obtained the results in sections 5 (jointly with GMC), 6 and 8.1 (jointly with CQS). CQS wrote and obtained the results in Sections 6 and 8.1 (jointly with AEG), and the plots in Fig. 1 to 4, 6 and 7. EMMO proposed the log gamma extended Weibull regression model described in section 7, wrote and obtained the results in that section and in section 8.2, and also the plots in Fig. 5 and 8.

Author details
[1]Departamento de Estatística, Universidade Federal de Pernambuco, 52171-900 Recife, PE, Brazil. [2]Departamento de Estatística, Universidade de Brasília, 70910-900 Brasília, DF, Brazil. [3]Departamento de Ciências Exatas, Universidade de São Paulo, 13418-900 Piracicaba, SP, Brazil.

References
Castellares, F, Montenegro, LC, Cordeiro, GM: The beta log-normal distribution. J. Stat. Comput. Simul. **83**(2), 203–228 (2013)
Cordeiro, GM, Nadarajah, S, Ortega, EMM: General results for the beta Weibull distribution. J. Stat. Comput. Sim. **83**, 1082–1114 (2013)
Cordeiro, GM, Ortega, EMM, Silva, GO: The exponentiated generalized gamma distribution with application to lifetime data. J. Stat. Comput. Simul. **81**, 827–842 (2011)
Dixon, LCW, Szegö, GP: The Global Optimization Problem: an Introduction. In: Dixon, LCW, Szegö, GP (eds.) Towards Global Optimisation, Vol. 2, pp. 1–15. North Holland, Amsterdam, (1978)
Efron, B: Logistic regression, survival analysis, and the Kaplan-Meier curve. J. Am. Stat. Assoc. **83**, 414–425 (1988)
Glover, F: A template for scatter search and path relinking. In: Hao, J-K, Lutton, E, Ronald, E, Schoenauer, M, Snyers, D (eds.) Artificial Evolution. Lecture Notes in Computer Science, 1363, pp. 13–54. Springer, Berlin/Heidelberg, (1998)
Gupta, RC, Gupta, RD, Gupta, PL: Modeling failure time data by Lehman alternatives. Commun. Stat. Theor. M. **27**, 887–904 (1998)
Gupta, RD, Kundu, D: Exponentiated exponential family: an alternative to gamma and Weibull distributions. Biometrical J. **43**, 117–130 (2001)
Hao, Z, Singh, VP: Entropy-based parameter estimation for extended Burr XII distribution. Stoch. Env. Res. Risk A. **23**, 1113–1122 (2008)
Mahmoudi, E, Sepahdar, A: Exponentiated Weibull-Poisson distribution: Model, Properties and Applications. Math. Comput. Simul. **92**, 76–97 (2013)
Mäkeläinen, T, Schmidt, K, Styan, GPH: On the existence and uniqueness of the maximum likelihood estimate of a vVector-valued parameter in fixed-size. Ann. Stat. **9**, 758–767 (1981)
Mudholkar, GS, Srivastava, DK: Exponentiated Weibull family for analyzing bathtub failure-rate data. IEEE T. Reliab. **42**, 299–302 (1993)
Mudholkar, GS, Srivastava, DK, Freimer, M: The exponentiated Weibull family: a reanalysis of the bus-motor-failure data. Technometrics. **37**, 436–445 (1995)
Mudholkar, GS, Srivastava, DK, Kollia, GD: A generalization of the Weibull distribution with application to the analysis of survival data. J. Am. Stat. Assoc. **91**, 1575–1583 (1996)
Nadarajah, S, Cordeiro, GM, Ortega, EMM: General results for the Kumaraswamy-G distribution. J. Stat. Comput. Sim. **82**, 951–979 (2012)
Nadarajah, S, Gupta, AK: The exponentiated gamma distribution with application to drought data. Calcutta Stat. Assoc. **59**, 29–54 (2007)
Nash, JC, Varadhan, R: Unifying optimization algorithms to aid software system users: optimx for R. J. Stat. Software. **43**, 1–14 (2011)
Ortega, EMM, Cordeiro, GM, Campelo, AK, Kattam, MW, Cancho, VG: A power series beta Weibull regression model for predicting breast carcinoma. Stat. Med. **34**, 1366–1388 (2015)
Ristic, MM, Balakrishnan, N: The gamma exponentiated exponential distribution. J. Stat. Comput. Simul. **82**, 1191–1206 (2012)
Santana, TVF, Ortega, EMM, Cordeiro, GM, Silva, GO: The Kumaraswamy-Log-Logistic Distribution. J. Stat. Theory Appl. **11**, 265–291 (2012)
Shao, Q, Wong, H, Xia, J, Wai-Cheung, I: Models for extremes using the extended three-parameter Burr XII system with application to flood frequency analysis. Hydrolog. Sci. J. **49**, 685–702 (2004)

Cordeiro *et al. Journal of Statistical Distributions and Applications*   (2016) 3:7

Page 19 of 19

Silva, RB, Cordeiro, GM: The Burr XII power series distributions: A new compounding family. Stat. Braz. J. Probab. Stat. **29**(3), 565–589 (2015)

Stacy, EW: A generalization of the gamma distribution. Ann. Math. Stat. **33**, 1187–1192 (1962)

Tippett, LHC: Technological Applications of Statistics. Wiley, New York (1950)

Ugray, Z, Lasdon, L, Plummer, J, Glover, F, Kelly, J, Marti, R: Scatter search and local NLP solvers: a multistart framework for global optimization. INFORMS J. Comput. **19**(3), 328–340 (2007)

Varadhan, R, Gilbert, PD: BB: An R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function. J. Stat. Softw. **32**, 4 (2009)

Zografos, K, Balakrishnan, N: On families of beta- and generalized gamma-generated distributions and associated inference. Statist. Method. **6**, 344–362 (2009)