Anais da Academia Brasileira de Ciências (2018) 90(3): 2553-2577
(Annals of the Brazilian Academy of Sciences)
Printed version ISSN 0001-3765 / Online version ISSN 1678-2690
http://dx.doi.org/10.1590/0001-3765201820170423
www.scielo.br/aabc | www.fb.com/aabcjournal

The Exponentiated Power Generalized Weibull: Properties and Applications

FERNANDO A. PEÑA-RAMÍREZ[1], RENATA R. GUERRA[1], GAUSS M. CORDEIRO[2] and PEDRO R.D. MARINHO[3]

[1]Departamento de Estatística, Universidade Federal de Santa Maria, Cidade Universitária,
Av. Roraima, 1000, 97105-900 Santa Maria, RS, Brazil
[2]Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária,
Av. Prof. Moraes Rego, 1235, 50740-540 Recife, PE, Brazil
[3]Departamento de Estatística, Universidade Federal da Paraíba, Cidade Universitária,
Conj. Pres. Castelo Branco III, s/n, 58051-900 João Pessoa, PB, Brazil

Manuscript received on June 20, 2017; accepted for publication on November 1, 2017

ABSTRACT

We propose a new lifetime model called the exponentiated power generalized Weibull (EPGW) distribution, which is obtained from the exponentiated family applied to the power generalized Weibull (PGW) distribution. It can also be derived from a power transform on an exponentiated Nadarajah-Haghighi random variable. Since several structural properties of the PGW distribution have not been studied, they can be obtained from those of the EPGW distribution. The model is very flexible for modeling all common types of hazard rate functions. It is a very competitive model to the well-known Weibull, exponentiated exponential and exponentiated Weibull distributions, among others. We also give a physical motivation for the new distribution if the power parameter is an integer. Some of its mathematical properties are investigated. We discuss estimation of the model parameters by maximum likelihood and provide two applications to real data. A simulation study is performed in order to examine the accuracy of the maximum likelihood estimators of the model parameters.

Key words: Exponential distribution, lifetime data, Nadarajah-Haghighi distribution, power generalized Weibull distribution, survival function.

1 - INTRODUCTION

There has been an increased interest in defining new continuous distributions by adding shape parameters to an existing baseline model. One of the most widely-accepted methods on this parameter induction is the exponentiated-G (exp-G) class. Let $G(y)$ and $g(y)$ be the baseline cumulative distribution function (cdf) and the probability density function (pdf) of a random variable W, respectively. We obtain the exp-G cdf by

Correspondence to: Fernando A. Peña-Ramírez
E-mail: fernando.p.ramirez@ufsm.br

raising $G(y)$ to a positive power shape parameter to the baseline model. Thus, a random variable Y has an exp-G distribution if its cdf is given by

$$F(y) = G(y)^{\beta},$$

for $y \in \mathscr{D} \subseteq \mathbb{R}$, where $\beta > 0$ represents the additional parameter. The corresponding pdf is given by

$$f(y) = \beta\, g(x)\, G(y)^{\beta-1}.$$

Tahir and Nadarajah (2015) traced this approach back to the first half of the nineteenth century and found twenty-eight different exp-G models published in recent literature. Most of these models are motivated by their usefulness in exploring tail properties and also for improving the goodness-of-fit in comparison with their baselines. Another current reason for introducing exp-G distributions is their applications in lifetime data analysis.

Thus, the classical lifetime distributions have received a great deal of attention as baselines on the exp-G class, among other generated families. Using the exponential lifetime model as baseline, Gupta et al. (1998) pioneered the exponentiated exponential (EE) distribution.

The exponentiated Weibull (EW) distribution was introduced by **?**.

Another model that has been considered for modeling lifetime data is the Nadarajah-Haghighi (NH) distribution pioneered by Nadarajah and Haghighi (2011). The NH model is a generalization of the exponential distribution with cdf given by (for $z > 0$)

$$G(z) = 1 - \exp\{1 - (1 + \lambda z)^{\alpha}\}, \tag{1}$$

where λ and α are the scale and shape parameters, respectively. If Z has the cdf (1), we write $Z \sim \mathrm{NH}(\alpha, \lambda)$. The pdf of Z is given by

$$g(z) = \alpha\lambda\,(1 + \lambda z)^{\alpha-1} \exp\{1 - (1 + \lambda z)^{\alpha}\}. \tag{2}$$

The motivations for studying the NH model are: the relationship between the pdf (2) and its hazard rate function (hrf), the ability (or inability) to model data with mode fixed at zero and the fact that it can be interpreted as a truncated Weibull distribution. Further details and general properties can be found in Nadarajah and Haghighi (2011). The exponentiated Nadarajah-Haghighi (ENH) model was proposed by Lemonte (2013).

The exponential, NH and Weibull distributions are all special cases of the power generalized Weibull (PGW) distribution proposed by Bagdonavicius and Nikulin (2002) in the context of accelerated failure time models. The original PGW cdf is given by (for $t > 0$)

$$G(t) = 1 - \mathrm{e}^{1-\left[1+\left(\frac{t}{\sigma}\right)^{\beta}\right]^{\frac{1}{\gamma}}},$$

where σ, β and $\gamma > 0$. By setting $\lambda = \sigma^{-\beta}$ and $\alpha = \gamma^{-1}$, the cdf, pdf and hrf of this distribution reduce to

$$G(t) = 1 - \exp\{1 - (1 + \lambda t^{\gamma})^{\alpha}\}, \tag{3}$$

$$g(t) = \alpha\lambda\gamma t^{\gamma-1}(1 + \lambda t^{\gamma})^{\alpha-1} \exp\{1 - (1 + \lambda t^{\gamma})^{\alpha}\} \tag{4}$$

and

$$h(t) = \alpha\lambda\gamma t^{\gamma-1}(1 + \lambda t^{\gamma})^{\alpha-1},$$

respectively. Dimitrakopoulou et al. (2007) presented a lifetime model that has the same formulation as that one in (3) with motivation in competing risk scenario. Lai (2013) described the PGW among the Weibull generalizations that are often required to prescribe the nonmonotonic nature of the empirical hazard rates.

Nikulin and Haghighi (2006) introduced a chi-square statistic for testing the validity of the PGW distribution and presented an application to censored survival times of cancer patients. Nikulin and Haghighi (2009) presented shape analysis for the PGW pdf and hrf. They also obtained a series representation for the sth moment of this distribution, but only for integer values of s/γ. They do not provide a general expression for the PGW ordinary moments. We also note that there is a lack of other structural properties of the PGW distribution like incomplete moments, skewness, mean deviations, Bonferroni and Lorenz curves and Rényi entropy.

In this paper, we use the concept of exponentiated distributions for introducing a new four-parameter Weibull-type family, so-called the *exponentiated power generalized Weibull* (EPGW) distribution. The proposed distribution is obtained by considering the PGW model as baseline in the exp-G family. Thus, the EPGW cdf and pdf are given by (for $t > 0$)

$$F(t) = [1 - \exp\{1 - (1 + \lambda t^{\gamma})^{\alpha}\}]^{\beta}, \tag{5}$$

and

$$f(t) = \alpha\beta\lambda\gamma t^{\gamma-1} \frac{(1+\lambda t^{\gamma})^{\alpha-1} \exp\{1-(1+\lambda t^{\gamma})^{\alpha}\}}{[1-\exp\{1-(1+\lambda t^{\gamma})^{\alpha}\}]^{1-\beta}}, \tag{6}$$

respectively. Here, λ is the scale parameter and γ, α and β are shape parameters. Henceforth, we denote by T a random variable having pdf (6), say $T \sim \text{EPGW}(\alpha, \beta, \lambda, \gamma)$. Identifiability is a property which a model must satisfy for precise inference to be possible, which refers to whether the unknown parameters in the model can be uniquely estimated. Equation (5) is clearly identifiable.

The hrf of T is given by

$$\begin{aligned} h(t) =& \alpha\beta\lambda\gamma t^{\gamma-1}(1+\lambda t^{\gamma})^{\alpha-1} \\ & \times \frac{\exp\{1-(1+\lambda t^{\gamma})^{\alpha}\}[1-\exp\{1-(1+\lambda t^{\gamma})^{\alpha}\}]^{\beta-1}}{1-[1-\exp\{1-(1+\lambda t^{\gamma})^{\alpha}\}]^{\beta}}. \end{aligned} \tag{7}$$

By inverting (5), we obtain an explicit expression for the quantile function (qf) of the EPGW distribution, say $Q(u)$, given by

$$Q(u) = \lambda^{-1/\gamma} \left\{ \left[1 - \log(1 - u^{1/\beta})\right]^{1/\alpha} - 1 \right\}^{1/\gamma}, \quad u \in (0,1). \tag{8}$$

Its median M follows by setting $u = 1/2$ in (8). The simulation of the EPGW random variable is straightforward. If $U \sim U(0,1)$, then the random variable $T = Q(U)$ follows the EPGW distribution given by (6).

Some motivations for introducing the EPGW distribution are:

- The new distribution is quite flexible because it contains several well-known lifetime distributions as special models, see Table I. This feature is also suitable for testing the goodness-of-fit of these distributions.

TABLE I
Some special models of the EPGW distribution.

α	β	λ	γ	Distribution
1	1	-	1	Exponential
1	-	-	1	Exponentiated exponential
1	1	-	2	Rayleigh
1	-	-	2	Burr type X
1	1	-	-	Weibull
1	-	-	-	Exponentiated Weibull
-	1	-	1	Nadarajah-Haghighi
-	-	-	1	Exponentiated Nadarajah-Haghighi
-	1	-	-	Power generalized Weibull

- The current distribution can also be derived from a power transform on an ENH random variable. If $Y \sim$ENH(α,β,λ), the cdf of Y is given by

$$F_Y(y) = [1-\exp\{1-(1+\lambda y)^{\alpha}\}]^{\beta}, \quad y>0,$$

  where $\alpha > 0$, $\beta > 0$ and $\lambda > 0$. Consider the transformation $T = Y^{1/\gamma}$, where $\gamma > 0$. Thus, the cdf of T has the form $F(t) = F_Y(t^{\gamma})$ given by (5). A similar approach was addressed by Gomes et al. (2008). They proposed a new method of estimation for the generalized gamma distribution through the power transformation $W = X^c$, where X is a generalized gamma random variable and W has the gamma distribution.

- Once several structural properties of the PGW distribution have not been studied, they shall be obtained from those of the EPGW distribution.

- By pioneering a PGW generalization of the exp-G family, it is also possible to obtain several properties of other generated families based on linear combinations from those of the EPGW distribution. For example, for the beta-G famil (Eugene et al. 2002), the density function can be expressed as a linear combination of exp-G pdfs for any baseline G. Similar results can also be demonstrated for the Kumaraswamy-G introduced by Cordeiro and Castro (2011), among several others generated families of distributions.

- Let $\beta > 0$ be an integer. Thus, $F(t)$ given in (5) represents the cdf of the maximum value on a β-variate random sample from the PGW distribution, say: $T = \max\{T_1,\ldots,T_{\beta}\}$. In other words, the EPGW distribution can be used to model the maximum lifetime of a random sample from the PGW distribution with size β. Further, as part of the exp-G family, the EPGW distribution has the following physical interpretation. Consider a parallel system consisting of $\beta = n$ components, which means that the system works if at least one of the n-components works. If the lifetime distributions of the components are independent and identically distributed PGW random variables, then the lifetime distribution of the system becomes the EPGW cdf with power parameter $\beta = n$.

- The EPGW may provide consistently 'better fits' than other Weibull generalizations including its special models. This fact is shown by fitting the proposed distribution to two data sets in Section 13. The applications illustrate that the EPGW distribution can also be very competitive to other widely known lifetime models.

The paper is outlined as follows. Some mathematical properties of the new distribution are provided in Sections 3-10. They include ordinary and incomplete moments, mean deviations about the mean and the median, Bonferroni and Lorenz curves, Rényi entropy, reliability and order statistics. In Section 11, we present the maximum likelihood method to estimate the model parameters. In Section 12, a simulation study evaluates the performance of the maximum likelihood estimators (MLEs). Applications to two real data sets are presented in Section 13. Section 14 offers some concluding remarks.

2 - DENSITY AND HAZARD SHAPES

Note that the pdf (6) can be expressed in terms of the cdf and pdf given in (3) and (4), respectively, in the form $f(t) = \beta G(t)^{\beta-1} g(t)$. Thus, the multiplicative factor $\beta G(t)^{\beta-1}$ is greater (smaller) than one for $\beta > 1$ ($\beta < 1$) and for larger values of t, and the opposite occurs for smaller values of t. The inclusion of the extra shape parameter β provides greater flexibility in terms of skewness and kurtosis of the new distribution. The pdf (6) can take various forms depending on the values of the shape parameters α, β and γ. It is easy to verify that

$$\lim_{t\to 0} f(t) = \begin{cases} \infty & \text{if } \beta < 1, \\ \alpha\lambda\gamma & \text{if } \beta = 1, \\ 0 & \text{if } \beta > 1, \end{cases}$$

and $\lim_{t\to\infty} f(t) = 0$.

Seting $z = (1+\lambda t^{\gamma})^{\alpha}$, we can rewrite the EPGW pdf as

$$\psi(z) = \alpha\beta\lambda^{1/\gamma}\gamma z^{(\alpha-1)/\alpha}(z^{1/\alpha}-1)^{(\gamma-1)/\gamma} e^{1-z}(1-e^{1-z})^{\beta-1}.$$

Differentiating twice $\log\psi(z)$ with respect to z, we obtain

$$\frac{d^2 \log\psi(z)}{dz^2} = -\left[\frac{\alpha-1}{\alpha z^2} + \frac{(\beta-1)e^{1-z}}{(1-e^{1-z})^2} + \frac{z^{\frac{1}{\alpha}-2}(\gamma-1)\left[1+\alpha\left(z^{1/\alpha}-1\right)\right]}{\alpha^2\gamma\left(z^{1/\alpha}-1\right)^2}\right].$$

Note that $z = (1+\lambda t^{\gamma})^{\alpha}$ implies that $z > 1$. Thus, we can verify that for $t > 0$, $\alpha < 1$, $\beta < 1$ and $\gamma < 1$, $[d^2 \log\psi(z)/dz^2] > 0$. This implies that the EPGW pdf is log-convex. Further, for $t > 0$, $\alpha > 1$, $\beta > 1$ and $\gamma > 1$, $[d^2 \log\psi(z)/dz^2] < 0$, which implies that the EPGW pdf is log-concave. Figure 1 displays plots of the pdf (6) for some parameter values. It illustrates the flexibility of the EPGW density, which allows modeling skewed and asymmetrical data.

Analogously, the EPGW hrf can be rewritten as

$$\phi(z) = \alpha\beta\lambda^{1/\gamma}\gamma z^{(\alpha-1)/\alpha}(z^{1/\alpha}-1)^{(\gamma-1)/\gamma}\frac{e^{1-z}(1-e^{1-z})^{-1}}{(1-e^{1-z})^{-\beta}-1}.$$

The critical point are obtained from

$$\frac{d\log\phi(z)}{dz} = \frac{\alpha-1}{\alpha z} + \frac{(\gamma-1)z^{(1-\alpha)/\alpha}}{\alpha\gamma(z^{1/\alpha}-1)} + \frac{\beta e^{1-z}}{(1-e^{1-z})[1-(1-e^{1-z})^{\beta}]} - \frac{e^{1-z}}{1-e^{1-z}} - 1 = 0.$$

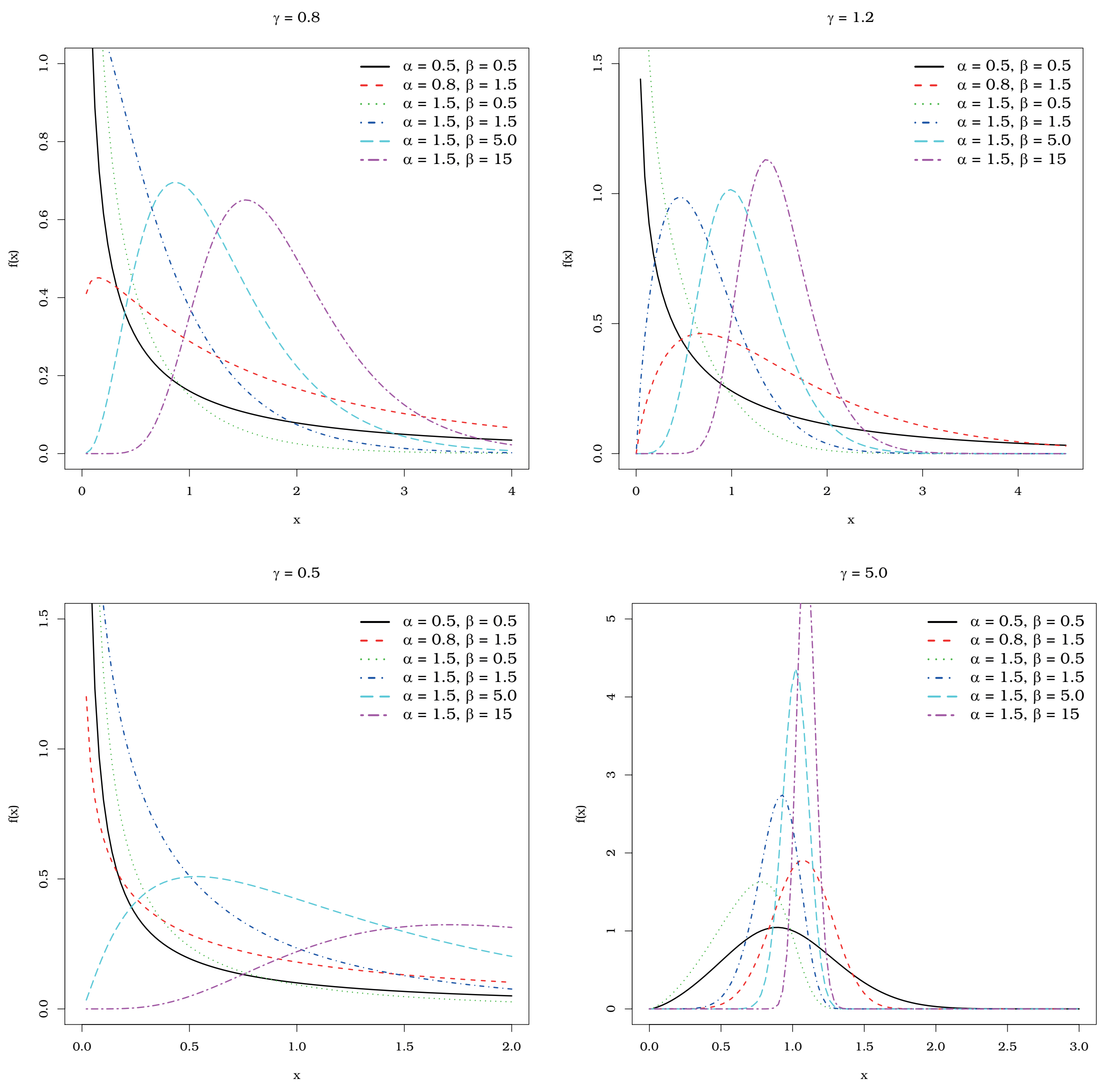


Figure 1 - Plots of the EPGW density for $\lambda = 1$.

For $\alpha = \beta = \gamma = 1$, $\mathrm{d}\log\phi(z)/\mathrm{d}z = 0$ and the hrf is constant. For $\alpha < 1$, $\gamma < 1$ and $\beta < 1$, $\mathrm{d}\log\phi(z)/\mathrm{d}z < 0$ and the hrf is decreasing. There may be more than one root to this equation.

Figure 2 provides plots of the hrf (7) for some parameter values. Figure 2 reveals that the EPGW distribution can have decreasing, increasing, upside-down bathtub and bathtub-shaped hazard functions. This feature makes the new distribution very attractive to model lifetime data. For example, according to Nadarajah et al. (2011) most empirical life systems have bathtub shapes for their hrfs.

3 - MOMENTS

The sth ordinary moment of T is obtained as $\mathrm{E}(T^s) = \int_0^\infty t^s f(t)\mathrm{d}t$, with $f(t)$ from (6). For illustrative purposes, we provide a small numerical study by computing the first six moments for some scenarios. Each one considers a different parametrization for γ and β, with fixed $\alpha = 1.5$ and $\lambda = 1$.

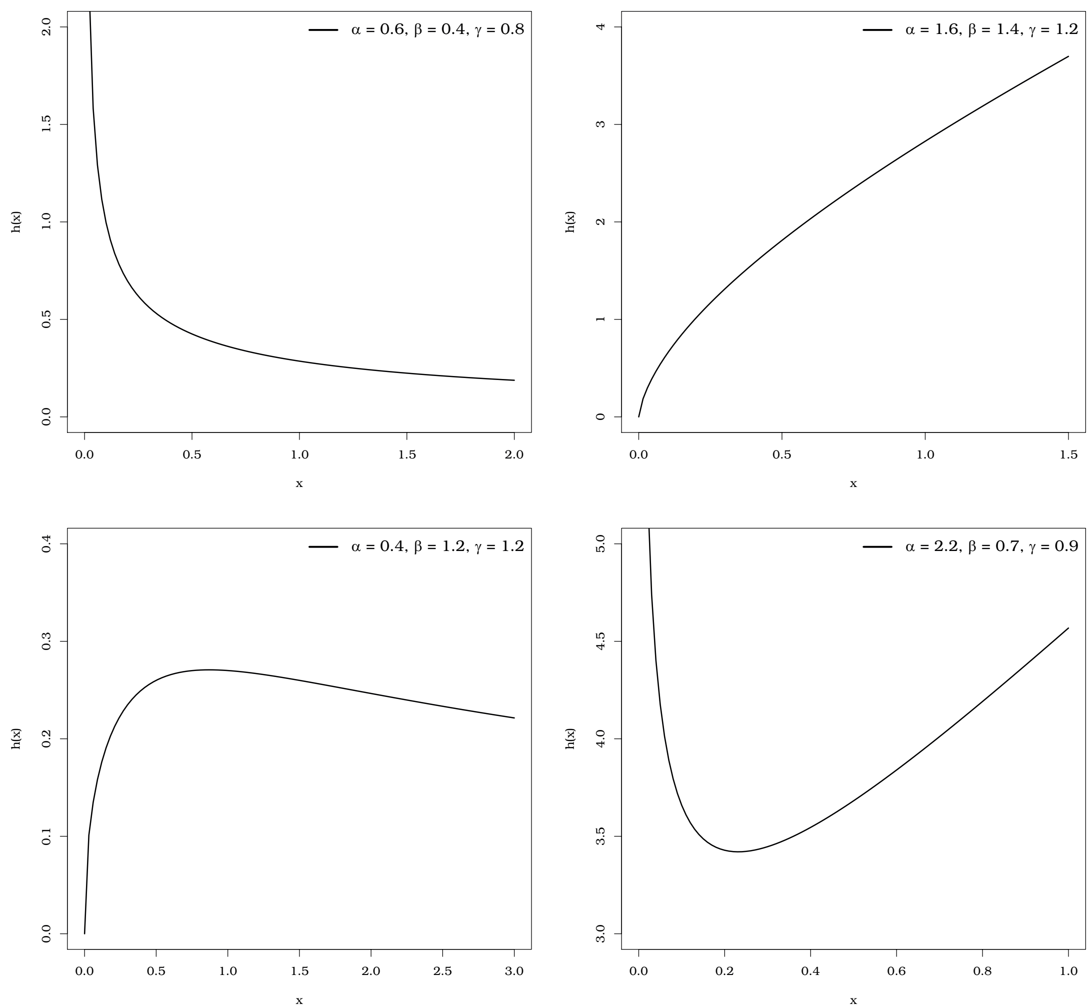


Figure 2 - Plots of the EPGW hrf for $\lambda = 1$.

These results are presented in Table II. All computations are obtained using R software, which have numerical integration routines with great precision. Based on these values, we can note that, for fixed γ, the additional parameter β has large impact on the moments of T. Note that the moments increase as β increases. For β fixed and ≤ 1, the moments decreases when γ increases.

The sth moment of T can also be determined from equation (8). After some algebra, we can write

$$\mu_s' = \mathbb{E}(T^s) = \beta\lambda^{-s/\gamma} I_s(\alpha,\beta,\gamma),$$

where $I_s(\alpha,\beta,\gamma) = \int_0^1 \{[1-\log(1-u)]^{1/\alpha} - 1\}^{s/\gamma} u^{\beta-1} \mathrm{d}u$ is an integral to be evaluated numerically.

Using the binomial expansion since $0 < \mathrm{e}^{1-(1+\lambda t^{\gamma})^{\alpha}} < 1$, the inverse of the denominator of (6) can be expressed as

$$[1-\exp\{1-(1+\lambda t^{\gamma})^{\alpha}\}]^{\beta-1} = \sum_{j=0}^{\infty} (-1)^j \binom{\beta-1}{j} \mathrm{e}^{j[1-(1+\lambda t^{\gamma})^{\alpha}]}.$$

Further, we can rewrite μ_s' as

$$\mu_s' = \alpha\beta\lambda\gamma \sum_{j=0}^{\infty} (-1)^j \binom{\beta-1}{j} \mathrm{e}^{j+1} \int_0^{\infty} t^{s+\gamma-1}(1+\lambda t^{\gamma})^{\alpha-1} \mathrm{e}^{-(j+1)(1+\lambda t^{\gamma})^{\alpha}} \mathrm{d}t. \tag{9}$$

We consider the integral

$$J = \int_0^{\infty} t^{s+\gamma-1}(1+\lambda t^{\gamma})^{\alpha-1} \, \mathrm{e}^{-(j+1)(1+\lambda t^{\gamma})^{\alpha}} \mathrm{d}t.$$

Setting $u = (j+1)(1+\lambda t^{\gamma})^{\alpha}$, we have

$$t = \left\{ \lambda^{-1} \left[\left(\frac{u}{j+1} \right)^{1/\alpha} - 1 \right] \right\}^{1/\gamma}.$$

Hence, after some algebra, we obtain

$$J = \left(\frac{1}{\lambda} \right)^{s/\gamma} \int_{j+1}^{\infty} \left[\left(\frac{u}{j+1} \right)^{1/\alpha} - 1 \right]^{s/\gamma} \frac{\mathrm{e}^{-u}}{\alpha\lambda\gamma(j+1)} \, \mathrm{d}u. \tag{10}$$

The most general case of the binomial theorem is the power series identity

$$(x+a)^{\nu} = \sum_{k=0}^{\infty} \binom{\nu}{k} x^k a^{\nu-k}, \tag{11}$$

where $\binom{\nu}{k}$ is a binomial coefficient and ν is a real number. This power series converges for $\nu \geq 0$ an integer, or $|x/a| < 1$. This general form is given by Graham (1994). By using (11) in equation (10), since $\left|[u/(j+1)]^{1/\alpha}\right| < 1$, it follows from (9) that

$$\mu_s' = \beta\lambda^{-s/\gamma} \sum_{i,j=0}^{\infty} \frac{(-1)^{i+j}\,\mathrm{e}^{j+1}}{(j+1)^{[s-\gamma(i-\alpha)]/\alpha\gamma}} \binom{\beta-1}{j} \binom{s/\gamma}{i} \Gamma\left(\frac{s-\gamma(i-\alpha)}{\alpha\gamma}, \, j+1 \right), \tag{12}$$

where $\Gamma(a,x) = \int_x^{\infty} z^{a-1}\,\mathrm{e}^{-z}\mathrm{d}z$ denotes the complementary incomplete gamma function, which is defined for all real numbers except the negative integers. Equation (12) is the main result of this section.

4 - SKEWNESS

The central moments (μ_s) and cumulants (κ_s) of T can be expressed recursively from equation (12) as

$$\mu_s = \sum_{k=0}^{s} (-1)^k \binom{s}{k} \mu_1'^k \mu_{s-k}' \qquad \text{and} \qquad \kappa_s = \sum_{k=0}^{s-1} \binom{s-1}{k-1} \kappa_k \, \mu_{s-k}',$$

respectively, where $\kappa_1 = \mu_1'$. Thus, $\kappa_2 = \mu_2' - \mu_1'^2$, $\kappa_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$, etc. The skewness $\gamma_1 = \kappa_3/\kappa_2^{3/2}$ and kurtosis $\gamma_2 = \kappa_4/\kappa_2^2$ can be determined from the third and fourth standardized cumulants.

The MacGillivray (1986) skewness function of T is given by

$$\rho(u) = \rho(u;\alpha,\beta,\gamma) = \frac{\rho_{(1)}(u;\alpha,\beta,\gamma)}{\rho_{(2)}(u;\alpha,\beta,)} = \frac{Q(1-u)+Q(u)-2Q(1/2)}{Q(1-u)-Q(u)},$$

TABLE II
First six moments for some scenarios of γ and β, with fixed $\alpha = 1.5$ and $\lambda = 1$.

γ	β	E(T)	E(T^2)	E(T^3)	E(T^4)	E(T^5)	E(T^6)
0.5	0.5	0.2989	0.6249	2.7433	19.2657	191.2419	2498.7853
	0.8	0.4488	0.9827	4.3653	30.7679	305.7880	3997.1257
	1.0	0.5397	1.2149	5.4372	38.4129	382.0702	4995.6341
	2.0	0.9176	2.3100	10.6904	76.3690	762.5177	9983.6001
	3.0	1.2116	3.3145	15.7832	113.8992	1141.4011	14964.0482
	4.0	1.4537	4.2472	20.7336	151.0300	1518.7741	19937.1200
0.8	0.5	0.3162	0.3259	0.5425	1.2207	3.4227	11.4158
	0.8	0.4515	0.4986	0.8507	1.9344	5.4503	18.2210
	1.0	0.5278	0.6062	1.0498	2.4030	6.7917	22.7400
	2.0	0.8109	1.0755	1.9810	4.6677	13.3819	45.1295
	3.0	1.0030	1.4633	2.8250	6.8184	19.7945	67.1961
	4.0	1.1481	1.7964	3.6014	8.8727	26.0480	88.9632
1.0	0.5	0.3450	0.2989	0.3824	0.6249	1.2207	2.7433
	0.8	0.4789	0.4488	0.5924	0.9827	1.9344	4.3653
	1.0	0.5517	0.5397	0.7257	1.2149	2.4030	5.4372
	2.0	0.8066	0.9176	1.3275	2.3100	4.6677	10.6904
	3.0	0.9688	1.2116	1.8477	3.3145	6.8184	15.7832
	4.0	1.0868	1.4537	2.3092	4.2472	8.8727	20.7336
2.0	0.5	0.4818	0.3450	0.3012	0.2989	0.3259	0.3824
	0.8	0.6092	0.4789	0.4391	0.4488	0.4986	0.5924
	1.0	0.6702	0.5517	0.5194	0.5397	0.6062	0.7257
	2.0	0.8518	0.8066	0.8308	0.9176	1.0755	1.3275
	3.0	0.9488	0.9688	1.0538	1.2116	1.4633	1.8477
	4.0	1.0129	1.0868	1.2278	1.4537	1.7964	2.3092
3.0	0.5	0.5773	0.4186	0.3450	0.3099	0.2968	0.2989
	0.8	0.6929	0.5504	0.4789	0.4459	0.4377	0.4488
	1.0	0.7445	0.6168	0.5517	0.5234	0.5210	0.5397
	2.0	0.8868	0.8277	0.8066	0.8158	0.8527	0.9176
	3.0	0.9571	0.9482	0.9688	1.0181	1.0976	1.2116
	4.0	1.0018	1.0306	1.0868	1.1725	1.2925	1.4537
4.0	0.5	0.6440	0.4818	0.3950	0.3450	0.3162	0.3012
	0.8	0.7478	0.6092	0.5276	0.4789	0.4515	0.4391
	1.0	0.7922	0.6702	0.5961	0.5517	0.5278	0.5194
	2.0	0.9091	0.8518	0.8193	0.8066	0.8109	0.8308
	3.0	0.9644	0.9488	0.9507	0.9688	1.0030	1.0538
	4.0	0.9987	1.0129	1.0420	1.0868	1.1481	1.2278

where $u \in (0, 1)$, $Q(\cdot)$ is the qf defined in (8),

$$\begin{aligned}\rho_{(1)}(u;\alpha,\beta,\gamma) &= \left\{\left[1-\log(1-(1-u)^{1/\beta})\right]^{1/\alpha}-1\right\}^{1/\gamma}+\left\{\left[1-\log(1-u^{1/\beta})\right]^{1/\alpha}-1\right\}^{1/\gamma}\\ &- 2\left\{\left[1+\beta^{-1}\log(2)-\log(2^{1/\beta}-1)\right]^{1/\alpha}-1\right\}^{1/\gamma}\end{aligned}$$

and

$$\rho_{(2)}(u;\alpha,\beta,\gamma) = \left\{\left[1-\log(1-(1-u)^{1/\beta})\right]^{1/\alpha}-1\right\}^{1/\gamma}-\left\{\left[1-\log(1-u^{1/\beta})\right]^{1/\alpha}-1\right\}^{1/\gamma}.$$

It is based on quantiles and can illustrate the effects of the shape parameters α, β and γ on the skewness of T. Plots of $\rho(u)$ for some parameter values are displayed in Figure 3. These plots reveal that when the parameters β and γ increase, the function $\rho(u)$ converges to zero. The closer $\rho(u)$ is to the horizontal line $\rho(u) = 0$, the density becomes more symmetrical. The quantity $\rho(u)$ does not depend on the parameter λ since it is a scale parameter.

5 - INCOMPLETE MOMENTS

The sth incomplete moment of T, say $m_s(y) = \int_0^y t^s f(t)\mathrm{d}t$, follows as

$$m_s(y) = \beta\lambda^{-s/\gamma}\int_0^{1-\mathrm{e}^{1-(1+\lambda y^\gamma)^\alpha}}\{[1-\log(1-u)]^{1/\alpha}-1\}^{s/\gamma}u^{\beta-1}\mathrm{d}u.$$

An alternative expression for $m_s(y)$ takes the form

$$\begin{aligned}m_s(y) = \beta\lambda^{-s/\gamma}\sum_{i,j=0}^{\infty}\frac{(-1)^{i+j}\,\mathrm{e}^{j+1}}{(j+1)^{[s-\gamma(i-\alpha)]/\alpha\gamma}}\binom{\beta-1}{j}\binom{s/\gamma}{i}\Bigg[\Gamma\left(\frac{s-\gamma(i-\alpha)}{\alpha\gamma},\ j+1\right)\\ -\Gamma\left(\frac{s-\gamma(i-\alpha)}{\alpha\gamma},\ (j+1)(1+\lambda\,y^\gamma)^\alpha\right)\Bigg].\end{aligned}$$

6 - MEAN DEVIATIONS

The mean deviations about the mean ($\delta_1 = \mathbb{E}(|T-\mu_1'|)$) and about the median ($\delta_2 = \mathbb{E}(|T-M|)$) of T can be expressed as

$$\delta_1 = 2\mu_1' F(\mu_1') - 2\,m_1(\mu_1') \qquad \text{and} \qquad \delta_2 = \mu_1' - 2\,m_1(M),$$

respectively, where $\mu_1' = \mathbb{E}(T)$, $M = \mathrm{Median}(T) = Q(0.5)$ is the median, $F(\mu_1')$ is easily determined from (5) and $m_1(y) = \int_0^y t\,f(t)\mathrm{d}t$ is the first incomplete moment. Hence, we can write

$$m_1(y) = \beta\lambda^{-1/\gamma}\int_0^{1-\mathrm{e}^{1-(1+\lambda y^\gamma)^\alpha}}\{[1-\log(1-u)]^{1/\alpha}-1\}^{1/\gamma}u^{\beta-1}\mathrm{d}u.$$

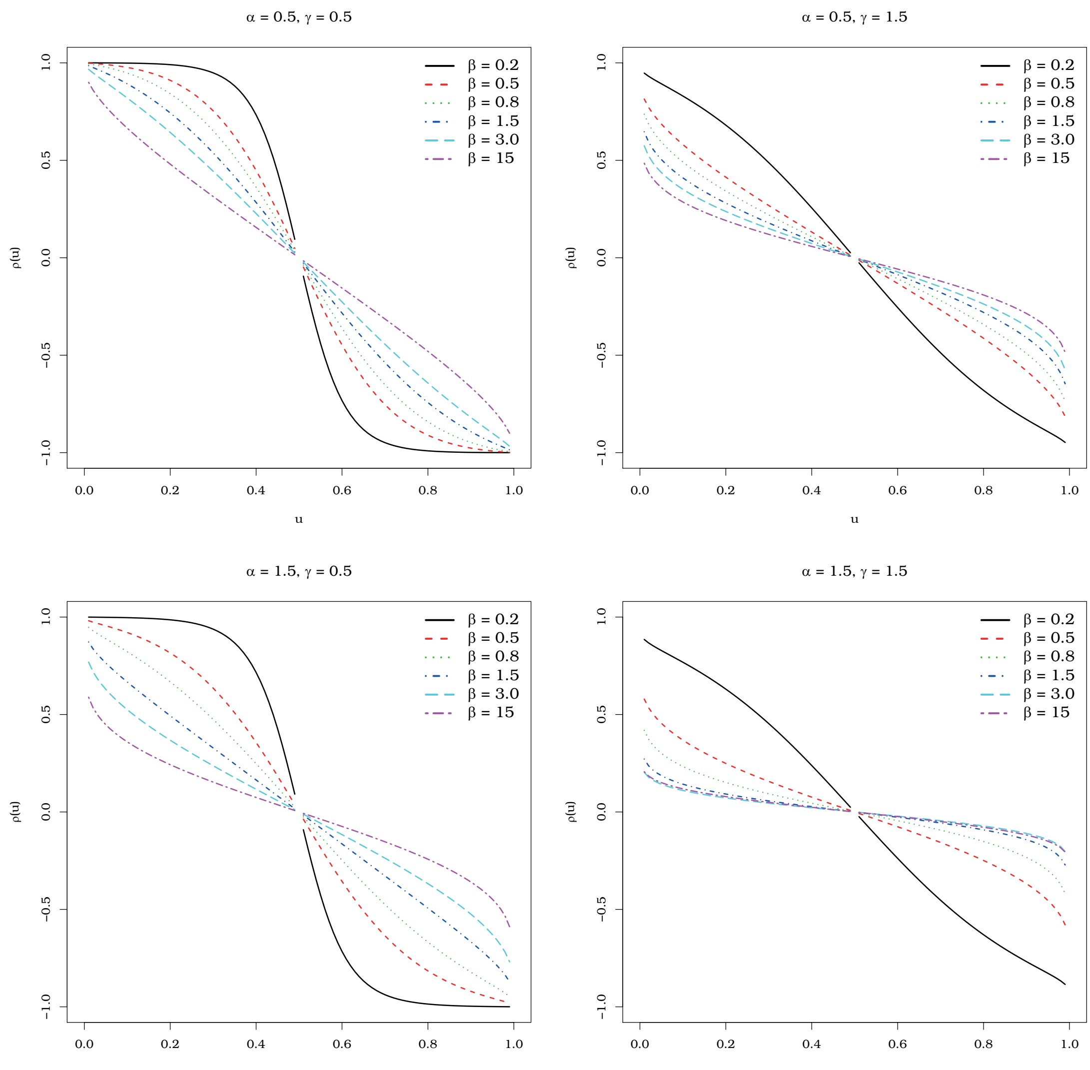


Figure 3 - The MacGillivray's skewness of the EPGW distribution.

Alternatively, we can determine $m_1(y)$ as

$$m_1(y)=\beta\lambda^{-1/\gamma}\sum_{i,j=0}^{\infty}\frac{(-1)^{i+j}\,\mathrm{e}^{j+1}}{(j+1)^{[1-\gamma(i-\alpha)]/\alpha\gamma}}\binom{\beta-1}{j}\binom{1/\gamma}{i}\left[\Gamma\left(\frac{1-\gamma(i-\alpha)}{\alpha\gamma},\,j+1\right)\right.$$
$$\left.-\Gamma\left(\frac{1-\gamma(i-\alpha)}{\alpha\gamma},\,(j+1)(1+\lambda\,y^{\gamma})^{\alpha}\right)\right].$$

7 - BONFERRONI AND LORENZ CURVES

Applications of the previous results to the Bonferroni and Lorenz curves are important in several fields such as economics, demography, insurance and medicine. They are defined, for a given probability π, by $B(\pi)=$

$m_1(q)/(\pi\mu_1')$ and $L(\pi) = m_1(q)/\mu_1'$, respectively, where $q = Q(\pi)$ follows from (8). The Gini concentration (C_G) is defined as the area between the curve $L(\pi)$ and the straight line. Hence,

$$C_G = 1 - 2\int_0^1 L(\pi)\mathrm{d}u.$$

An alternative expression is $C_G = (2\delta - \mu_1')/\mu_1'$, where $\delta = \mathbb{E}[TF(T)]$. The quantity δ is given by

$$\delta = \beta\lambda^{-1/\gamma}\int_0^1 u^{2\beta-1}\{[1-\log(1-u)]^{1/\alpha} - 1\}^{1/\gamma}\mathrm{d}u.$$

This integral can be easily evaluated numerically in softwares such as `R` and `Ox`, among others. An alternative expression for δ takes the form

$$\delta = \beta\lambda^{-1/\gamma}\sum_{i,j=0}^{\infty}\frac{(-1)^{i+j}\,\mathrm{e}^{j+1}}{(j+1)^{[1-\gamma(i-\alpha)]/\alpha\gamma}}\binom{2\beta-1}{j}\binom{1/\gamma}{i} \times\Gamma\left(\frac{1-\gamma(i-\alpha)}{\alpha\gamma},\, j+1\right).$$

For $\gamma = 1$, we can prove that this expression reduces to that one obtained by Lemonte (2013).

8 - ENTROPY

The entropy of a random variable is a measure of variation of the uncertainty and has been used in many fields. Several measures of entropy have been studied in the literature. However, we consider the most popular entropy measure: the Rényi entropy of a random variable with pdf $f(x)$ defined by

$$I_R = I_R(\delta) = \frac{1}{1-\delta}\log\left[\int_{-\infty}^{\infty} f^{\delta}(x)\right]\mathrm{d}x,$$

for $\delta > 0$ and $\delta \neq 1$. The Rényi entropy of T can be expressed as

$$I_R = M + \frac{1}{1-\delta}\log\left(\int_1^{\infty}\frac{u^{\alpha^{-1}(\alpha-1)(\delta-1)}(u^{1/\alpha}-1)^{\gamma^{-1}(\gamma-1)(\delta-1)}\mathrm{e}^{\delta(1-u)}}{[1-\mathrm{e}^{1-u}]^{\delta(1-\beta)}}\mathrm{d}u\right),$$

where $M = -\log(\alpha\gamma\lambda^{\gamma}) + \frac{\delta}{1-\delta}\log(\beta)$. The above integral can be evaluated numerically. By expanding the inverse of the denominator using the binomial expansion, we obtain

$$\begin{aligned} I_R &= M + \frac{1}{1-\delta}\log\left[\sum_{j=0}^{\infty}(-1)^j\mathrm{e}^{\delta+j}\binom{\delta(\beta-1)}{j}\right. \\ &\times \left.\left(\int_1^{\infty} u^{\alpha^{-1}(\alpha-1)(\delta-1)}(u^{1/\alpha}-1)^{\gamma^{-1}(\gamma-1)(\delta-1)}\mathrm{e}^{-u(\delta+j)}\mathrm{d}u\right)\right]. \end{aligned}$$

Again, by using the binomial expansion, I_R can be expressed as

$$\begin{aligned} I_R &= M + \frac{1}{1-\delta}\log\left[\sum_{j,k=0}^{\infty}\frac{(-1)^{j+k}\mathrm{e}^{\delta+j}}{(j+\delta)^{[\delta(\gamma\alpha-1)+1]/\gamma\alpha}}\right. \\ &\times \left.\binom{\delta(\beta-1)}{j}\binom{(\gamma-1)(\delta-1)/\gamma}{k}\Gamma\left(\frac{\delta(\gamma\alpha-1)+1}{\gamma\alpha}, j+\delta\right)\right]. \end{aligned}$$

For $\gamma = 1$, the last expression agrees with the result by Lemonte (2013).

9 - STRESS-STRENGTH RELIABILITY

The stress-strength reliability, defined as $R = P(X > Y)$, is a measure that describes the life of a component with a random strength X, which is subjected to a random stress Y. The failure occurs if the stress applied to the component exceeds the strength, i.e. $Y > X$, otherwise it will function satisfactorily. Clearly, this measure is very useful in engineering context such as deterioration of rocket motors and the aging of concrete pressure vessels.

Let X and Y be two independent random variables with EPGW$(\alpha,\beta_1,\lambda,\gamma)$ and EPGW$(\alpha,\beta_2,\lambda,\gamma)$ distributions, respectively. We shall obtain the stress–strength parameter in the form

$$R = \alpha\beta_1\lambda\gamma\int_0^\infty \frac{t^{\gamma-1}(1+\lambda t^\gamma)^{\alpha-1}\exp\{1-(1+\lambda t^\gamma)^\alpha\}}{[1-\exp\{1-(1+\lambda t^\gamma)^\alpha\}]^{1-\beta_1-\beta_2}}\mathrm{d}t.$$

Thus, by taking $u = 1-\exp\{1-(1+\lambda t^\gamma)^\alpha\}$, the above integral can be reduced to

$$R = \beta_1\int_0^1 u^{\beta_2+\beta_1-1}\mathrm{d}u = \frac{\beta_1}{\beta_1+\beta_2}.$$

10 - ORDER STATISTICS

Let $T_1,\cdots,T_n$ be a random sample from the EPGW distribution. Let $T_{i:n}$ denote the ith order statistic. The probability density function of $T_{i:n}$ is

$$f_{i:n}(t) = \frac{1}{B(i,n-i+1)}\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}f(t)\,F(t)^{i+j-1}. \tag{13}$$

By inserting (5) and (6) in (13) and after some algebra, we obtain

$$f_{i:n}(t) = \frac{1}{B(i,n-i+1)}\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}\alpha\beta\lambda\gamma t^{\gamma-1}\frac{(1+\lambda t^\gamma)^{\alpha-1}\exp\{1-(1+\lambda t^\gamma)^\alpha\}}{[1-\exp\{1-(1+\lambda t^\gamma)^\alpha\}]^{1-(i+j)\beta}}.$$

Thus, we can write

$$f_{i:n}(t) = \sum_{j=0}^{n-i}\upsilon_{ij}\,f(t;\alpha,(i+j)\beta,\lambda,\gamma), \tag{14}$$

where

$$\upsilon_{ij} = \frac{(-1)^j}{(i+j)\,B(i,n-i+1)}\binom{n-i}{j},$$

$f(t;\alpha,(i+j)\beta,\lambda,\gamma)$ is the EPGW density function with scale parameter λ and shape parameters γ, α and $(i+j)\beta$. Equation (14) is the main result of this section. Based on this, we can obtain some structural properties of $T_{i:n}$ using similar procedures as presented in the previous sections.

11 - MAXIMUM LIKELIHOOD ESTIMATION

This section addresses the estimation of the unknown parameters of the EPGW distribution by the method of maximum likelihood. Let $t_1,\ldots,t_n$ be a random sample of size n from the EPGW$(\alpha,\beta,\lambda,\gamma)$ distribution. Let $\boldsymbol{\theta}=(\alpha,\beta,\lambda,\gamma)^T$ be the parameter vector of interest. The log-likelihood function for $\boldsymbol{\theta}$ based on this sample is

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= n+n\log(\alpha\,\beta\,\lambda\,\gamma)+(\gamma-1)\sum_{i=1}^{n}\log(t_i)-\sum_{i=1}^{n}\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha} \\ &+ (\alpha-1)\sum_{i=1}^{n}\log\left(1+\lambda\,t_i^{\gamma}\right)+(\beta-1)\sum_{i=1}^{n}\log\left[1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}\right].\end{aligned} \tag{15}$$

The components of the score vector $\boldsymbol{U}(\boldsymbol{\theta})$ are given by

$$\begin{aligned}U_{\alpha}(\boldsymbol{\theta}) &= \frac{n}{\alpha}+\sum_{i=1}^{n}\log\left(1+\lambda\,t_i^{\gamma}\right)-\sum_{i=1}^{n}\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}\log\left(1+\lambda\,t_i^{\gamma}\right) \\ &+(\beta-1)\sum_{i=1}^{n}\frac{\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}\log\left(1+\lambda\,t_i^{\gamma}\right)\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}{1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}},\quad U_{\beta}(\boldsymbol{\theta}) = \frac{n}{\beta}+\sum_{i=1}^{n}\log\left[1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}\right],\end{aligned}$$

$$\begin{aligned}U_{\lambda}(\boldsymbol{\theta}) &= \frac{n}{\lambda}+(\alpha-1)\sum_{i=1}^{n}t_i^{\gamma}\left(1+\lambda\,t_i^{\gamma}\right)^{-1}-\alpha\sum_{i=1}^{n}t_i^{\gamma}\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha-1} \\ &+\alpha\,(\beta-1)\sum_{i=1}^{n}\frac{t_i^{\gamma}\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha-1}\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}{1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}},\end{aligned}$$

and

$$\begin{aligned}U_{\gamma}(\boldsymbol{\theta}) &= \frac{n}{\gamma}+\sum_{i=1}^{n}\log(t_i)-\alpha\lambda\sum_{i=1}^{n}t_i^{\gamma}\log(t_i)\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha-1} \\ &+\lambda(\alpha-1)\sum_{i=1}^{n}t_i^{\gamma}\log(t_i)\left(1+\lambda\,t_i^{\gamma}\right)^{-1} \\ &+\lambda\,\alpha\,(\beta-1)\sum_{i=1}^{n}\frac{t_i^{\gamma}\log(t_i)\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha-1}\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}{1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}.\end{aligned}$$

Setting the above equations to zero, $\boldsymbol{U}(\boldsymbol{\theta})=\boldsymbol{0}$, and solving them simultaneously yields the MLEs of the four parameters. These equations cannot be solved analytically. We have to use iterative techniques such as the quasi-Newton BFGS and Newton-Raphson algorithms. The initial values for the parameters are important but are not hard to obtain from fitting special EPGW sub-models.

Note that, for fixed α, λ and γ, the MLE of β is given by

$$\hat{\beta}(\alpha,\lambda,\gamma)=-\frac{n}{\sum_{i=1}^{n}\log\left[1-\mathrm{e}^{1-\left(1+\lambda t_i^{\gamma}\right)^{\alpha}}\right]}.$$

Thus, it is easily observed that fixed on $t_1,\ldots,t_n$,

- $\hat{\beta} \to 0$ when $\hat{\alpha} \to 0$ and/or $\hat{\lambda} \to 0$
- $\hat{\beta} \to \infty$ when $\hat{\alpha} \to 0$ and/or $\hat{\lambda} \to \infty$
- $\hat{\beta} \to 0$ when $\hat{\gamma} \to \infty$ and $t_i < 1$, for some $i \leq n$
- $\hat{\beta} \to \infty$ when $\hat{\gamma} \to \infty$ and $t_i < 1$, $\forall i \leq n$.

This behavior anticipates that estimates for smaller α and/or λ may require improved estimation procedures.

By replacing β by $\hat{\beta}$ in equation (15) and letting $\boldsymbol{\theta_p} = (\alpha, \lambda, \gamma)$, the profile log-likelihood function for $\boldsymbol{\theta_p}$ can be expressed as

$$\begin{aligned}\ell(\boldsymbol{\theta_p}) = \quad & n \quad \log(n) + n\log(\alpha\,\lambda\,\gamma) + (\gamma-1)\sum_{i=1}^{n}\log(t_i) - \sum_{i=1}^{n}\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha} \\ + \quad & (\alpha-1)\sum_{i=1}^{n}\log\left(1+\lambda\,t_i^{\gamma}\right) - \sum_{i=1}^{n}\log\left[1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}\right] \\ - \quad & n\log\left\{-\sum_{i=1}^{n}\log\left[1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}\right]\right\}. \end{aligned} \tag{16}$$

We assume that the standard regularity conditions for $\ell_p = \ell(\boldsymbol{\theta_p})$ hold.

They are not restrictive and hold for the models cited in this paper. The score vector corresponding to (16), $\boldsymbol{U}(\boldsymbol{\theta_p})$, has the components

$$\begin{aligned}U_\alpha(\boldsymbol{\theta_p}) = & \frac{n}{\alpha} + \sum_{i=1}^{n}\log\left(1+\lambda\,t_i^{\gamma}\right) - \sum_{i=1}^{n}\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}\log\left(1+\lambda\,t_i^{\gamma}\right) \\ & - n\sum_{i=1}^{n}\frac{\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}\log\left(1+\lambda\,t_i^{\gamma}\right)\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}{1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}\left\{\sum_{i=1}^{n}\log\left[1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}\right]\right\}^{-1} \\ & - \sum_{i=1}^{n}\frac{\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}\log\left(1+\lambda\,t_i^{\gamma}\right)\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}{1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}, \\ U_\lambda(\boldsymbol{\theta_p}) = & \frac{n}{\lambda} + (\alpha-1)\sum_{i=1}^{n}t_i^{\gamma}\left(1+\lambda\,t_i^{\gamma}\right)^{-1} - \alpha\sum_{i=1}^{n}t_i^{\gamma}\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha-1} \\ & - n\alpha\sum_{i=1}^{n}\frac{t_i^{\gamma}\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha-1}\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}{1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}\left\{\sum_{i=1}^{n}\log\left[1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}\right]\right\}^{-1} \\ & - \alpha\sum_{i=1}^{n}\frac{t_i^{\gamma}\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha-1}\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}{1-\mathrm{e}^{1-\left(1+\lambda\,t_i^{\gamma}\right)^{\alpha}}}\end{aligned}$$

and

$$U_\gamma(\boldsymbol{\theta_p}) = \frac{n}{\gamma} + \sum_{i=1}^{n}\log(t_i) + \lambda(\alpha-1)\sum_{i=1}^{n}\frac{t_i^\gamma\log(t_i)}{1+\lambda t_i^\gamma} - \alpha\lambda\sum_{i=1}^{n} t_i^\gamma\log(t_i)\left(1+\lambda t_i^\gamma\right)^{\alpha-1}$$
$$- n\alpha\lambda\sum_{i=1}^{n}\frac{t_i^\gamma\log(t_i)\left(1+\lambda t_i^\gamma\right)^{\alpha-1}\mathrm{e}^{1-\left(1+\lambda t_i^\gamma\right)^\alpha}}{1-\mathrm{e}^{1-\left(1+\lambda t_i^\gamma\right)^\alpha}}\left\{\sum_{i=1}^{n}\log\left[1-\mathrm{e}^{1-\left(1+\lambda t_i^\gamma\right)^\alpha}\right]\right\}^{-1}$$
$$- \alpha\lambda\sum_{i=1}^{n}\frac{t_i^\gamma\log(t_i)\left(1+\lambda t_i^\gamma\right)^{\alpha-1}\mathrm{e}^{1-\left(1+\lambda t_i^\gamma\right)^\alpha}}{1-\mathrm{e}^{1-\left(1+\lambda t_i^\gamma\right)^\alpha}}.$$

Solving the equations in $\boldsymbol{U}(\boldsymbol{\theta_p}) = \mathbf{0}$ simultaneously yields the MLEs of α, λ and γ. The MLE of β is just $\hat{\beta}(\hat{\alpha},\hat{\lambda},\hat{\gamma})$. The maximization of the profile log-likelihood might be simpler since it involves only three parameters. Lemonte (2013) noted a similar result for the ENH model but mentioned that some of the properties that hold for a genuine likelihood do not hold for its profile version.

For interval estimation of the components of $\boldsymbol{\theta}$, we can adopt the observed information matrix $\boldsymbol{J}(\boldsymbol{\theta})$, whose elements can be obtained from the authors upon request. The multivariate normal $N_4(0, \boldsymbol{J}(\widehat{\boldsymbol{\theta}})^{-1})$ distribution can be used to construct approximate confidence intervals for the model parameters.

12 - SIMULATION STUDY

Here, a Monte Carlo simulation experiment is performed in order to examine the accuracy of the MLEs of the model parameters. The simulations are carried out by generating observations from the EPGW distribution using the inverse transformation method for different parameter combinations. The number of observations is set at $n = 100$, 300 and 500 and the number of replications at $10,000$. For maximizing the log-likelihood function, we use the `Optim` function with analytical derivatives in `R`. From the results of the simulations given in Table III, we can verify that the root mean squared errors (RMSEs) of the MLEs of α, β, λ and γ decay toward zero when the sample size n increases, as expected. The mean estimates of the parameters tend to be closer to the true parameter values when n increases.

13 - APPLICATIONS

In this section, we present two applications to illustrate the flexibility of the EPGW distribution. They indicate the potentiality of the new distribution for modeling positive data. The first data set represents the remission times (in months) of 128 patients with bladder cancer (Lee and Wang 2003). The second one corresponds to the 101 observations representing the stress-rupture life of kevlar 49/epoxy strands that are subjected to constant sustained pressure at the 90% stress level until all had failed. Then, we obtain a complete data set with exact failure times. This data set was studied by Andrews and Herzberg (1985). Table IV gives a descriptive summary of the samples. Note that both data sets present positive skewness and that the remission times show higher variance.

TABLE III
Mean estimates and RMSEs of the EPGW distribution for some parameter values.

					Mean estimates				RMSEs			
n	α	β	λ	γ	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\gamma}$
100	0.3	4.0	3.0	1.6	0.355	4.041	3.061	2.140	0.313	1.988	2.084	1.376
	1.7	0.8	0.1	0.2	1.670	0.795	0.105	0.253	0.833	0.434	0.109	0.117
	3.0	2.0	5.0	0.6	3.848	2.212	5.512	0.627	2.027	1.104	3.112	0.157
	3.5	0.9	0.2	0.1	2.729	0.781	0.284	0.166	1.674	0.493	0.276	0.128
	7.0	1.5	5.0	0.2	7.296	1.643	5.215	0.199	1.717	0.504	1.644	0.027
	7.5	1.3	4.0	0.5	8.034	1.436	4.786	0.510	2.488	0.589	2.386	0.100
300	0.3	4.0	3.0	1.6	0.316	4.035	2.990	1.843	0.145	1.515	1.534	0.807
	1.7	0.8	0.1	0.2	1.630	0.762	0.094	0.235	0.657	0.271	0.052	0.086
	3.0	2.0	5.0	0.6	3.454	2.096	5.083	0.603	1.277	0.573	2.189	0.086
	3.5	0.9	0.2	0.1	3.155	0.858	0.233	0.117	1.063	0.278	0.111	0.050
	7.0	1.5	5.0	0.2	7.093	1.556	5.121	0.199	1.130	0.285	1.141	0.018
	7.5	1.3	4.0	0.5	7.774	1.337	4.431	0.507	1.753	0.303	1.645	0.066
500	0.3	4.0	3.0	1.6	0.308	4.005	2.990	1.766	0.105	1.296	1.344	0.613
	1.7	0.8	0.1	0.2	1.639	0.766	0.093	0.226	0.572	0.226	0.041	0.069
	3.0	2.0	5.0	0.6	3.322	2.062	5.030	0.600	0.991	0.429	1.902	0.068
	3.5	0.9	0.2	0.1	3.301	0.883	0.219	0.107	0.789	0.204	0.073	0.025
	7.0	1.5	5.0	0.2	7.090	1.530	5.090	0.200	0.938	0.218	0.964	0.015
	7.5	1.3	4.0	0.5	7.724	1.323	4.272	0.504	1.462	0.231	1.327	0.053
800	0.3	4.0	3.0	1.6	0.304	4.013	2.994	1.718	0.083	1.134	1.167	0.498
	1.7	0.8	0.1	0.2	1.640	0.770	0.095	0.220	0.509	0.192	0.033	0.059
	3.0	2.0	5.0	0.6	3.242	2.044	4.993	0.599	0.812	0.336	1.656	0.056
	3.5	0.9	0.2	0.1	3.383	0.892	0.211	0.103	0.603	0.160	0.051	0.016
	7.0	1.5	5.0	0.2	7.063	1.518	5.061	0.200	0.789	0.171	0.801	0.012
	7.5	1.3	4.0	0.5	7.620	1.318	4.187	0.502	1.215	0.182	1.104	0.043

We fit the EPGW distribution (6) to these data sets and also present a comparative study with the fits of some embedded and not embedded models. One of these models is the Kumaraswamy Weibull (Kw-W) distribution, whose pdf is given by

$$g(t)=\frac{a\,b\,c\,\beta^{c}\,t^{c-1}\exp\{-(\beta t)^{c}\}\,[1-\exp\{-(\beta t)^{c}\}]^{a-1}}{\{1-[1-\exp\{-(\beta t)^{c}\}]^{a}\}^{1-b}},\quad t>0,$$

TABLE IV
Descriptive statistics.

Statistics	**Real data sets**	
	Remission times data	**Stress-rupture data**
Mean	9.3656	1.0248
Median	6.3950	0.8000
Mode	5.0000	0.5000
Variance	110.4250	1.2529
Skewness	3.2865	3.0017
Kurtosis	15.4831	13.7089
Maximum	79.0500	7.8900
Minimum	0.0800	0.0100
n	128	101

where $a>0$, $b>0$, $c>0$ and $\beta>0$. Another model is the beta Weibull (BW) distribution, whose pdf is given by

$$g(t)=\frac{\alpha}{\lambda\,B(a,b)}\left(\frac{t}{\lambda}\right)^{\alpha-1}[1-\exp\{-(t/\lambda)^{\alpha}\}]^{a-1}\exp\{-b\,(t/\lambda)^{\alpha}\},$$

where $a>0$, $b>0$, $\alpha>0$ and $\lambda>0$. The Beta-Fréchet (BFr) distribution, whose pdf is given by

$$g(t)=\frac{\lambda\,\sigma}{B(a,b)}t^{-(\lambda+1)}\exp\left\{a\left(\frac{\sigma}{t}\right)^{\lambda}\right\}\left[1-\exp\left\{a\left(\frac{\sigma}{t}\right)^{\lambda}\right\}\right]^{b-1},$$

where $a>0$, $b>0$, $\lambda>0$ and $\sigma>0$. We also consider the Marshall-Olkin Nadarajah-Haghighi (MONH) model, whose pdf is given by

$$g(t)=\alpha\,\beta\,\lambda\,\frac{(1+\lambda\,t)^{\alpha-1}\exp\{1-(1+\lambda\,t)^{\alpha}\}}{[1-(\beta-1)\exp\{1-(1+\lambda\,t)^{\alpha}\}]^{2}},$$

where $\alpha>0$, $\beta>0$ and $\lambda>0$. The EW distribution, whose pdf is given by

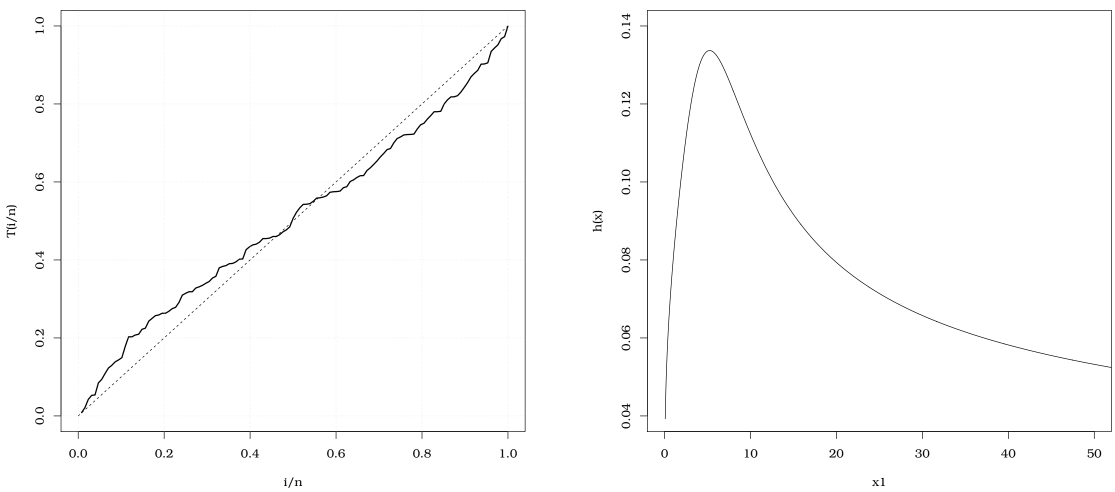
$$g(t)=\alpha\,\beta\,\lambda\,t^{\alpha-1}\exp(-\lambda\,t^{\alpha})\,[1-\exp(-\lambda\,t^{\alpha})]^{\beta-1},\;\; t>0,$$

where $\alpha>0$ and $\beta>0$ are shape parameters and $\lambda>0$ is a scale parameter. This distribution is quite flexible because its hrf presents the classic five forms (constant, decreasing, increasing, upside-down bathtub and bathtub-shaped). The Weibull model is a special case of the EW model when $\beta=1$.

The ENH distribution can also have the same shapes for the hrf and therefore can be an interesting alternative to the EW distribution in modeling positive data. The ENH density is given by (6) when $\gamma=1$. Further, for $\gamma=\beta=1$, we have as a special model the NH distribution given by (2). We also consider the PGW model, whose pdf is given in (4), which arises from the EPGW model when $\beta=1$.

Xie et al. (2002) proposed a modified Weibull (MW) density given by

$$g(t)=\lambda\beta\left(\frac{t}{\alpha}\right)^{1-\beta}\exp\left\{\left(\frac{t}{\alpha}\right)^{\beta}+\lambda\,\alpha\left(1-\exp\left\{\frac{t}{\alpha}\right\}^{\beta}\right)\right\},\;\; t>0,$$



(a) **(b)**

Figure 4 - The TTT plot **(a)** and EPGW hrf for the remission times data **(b)**.

where $\lambda > 0$, $\beta > 0$ and $\alpha > 0$. For $\alpha = 1$, it becomes the Chen distribution (Chen 2000). The MW and Chen distributions can have increasing or bathtub-shaped failure rate. An extension of the Weibull model proposed by Bebbington et al. (2007) has pdf given by

$$g(t) = \left(\alpha + \frac{\beta}{t^2}\right) \exp\left(\alpha t - \frac{\beta}{t}\right) \exp\left\{-\exp\left(\alpha t - \frac{\beta}{t}\right)\right\}, \; t > 0,$$

where $\alpha > 0$ and $\beta > 0$. We shall use the same terminology by Lemonte (2013) for this distribution, i.e., denote the flexible Weibull (FW) density. The FW model can have increasing or modified bathtub-shaped failure rate.

We use the simulated-annealing method for maximizing the log-likelihood function of the models in the two applications. The MLEs and goodness-of-fit statistics are evaluated using the `AdequacyModel` script in `R` software. Tables V and VI list the MLEs and the corresponding standard errors (SEs) in parentheses of the unknown parameters for the fitted models to remission times data (first data set) and stress-rupture failure times (second data set), respectively.

In applications there is qualitative information about the failure rate shape, which can help for selecting some models. Thus, a device called the total time on test (TTT) plot is useful. The TTT plot is obtained by plotting

$$T\left(\frac{r}{n}\right) = \left[\sum_{i=1}^{r} y_{i:n} + (n-r)\, y_{r:n}\right] \Big/ \sum_{i=1}^{n} y_{i:n},$$

against r/n, where $r = 1, \ldots, n$ and $y_{i:n}\,(i = 1, \ldots, n)$ are the order statistics of the sample.

The figures in Tables V and VI indicate that the MLEs of the EPGW model are precise for both data sets. Figures 4 and 5 provide the TTT plots and plots of the hrf for the EPGW fitted model for the remission times and stress-rupture times data sets, respectively. They reveal that the EPGW hrf has decreasing and decreasing-increasing-decreasing shapes, respectively. This fact is in agreement with the TTT plot based on each data set.

TABLE V
The MLEs of the model parameters for the remission times data and the corresponding SEs in parentheses.

Distributions	Estimates			
EPGW($\alpha,\beta,\lambda,\gamma$)	0.2076	0.4062	0.0047	3.1008
	(0.0347)	(0.0547)	(0.0025)	(0.2866)
Kw-W(a,b,c,β)	3.8071	1.7364	0.5144	0.2904
	(1.3951)	(0.8486)	(0.1157)	(0.1435)
BW(a,b,α,λ)	6.4498	8.3256	0.3909	25.4616
	(2.7464)	(4.3776)	(0.0859)	(18.7415)
BFr(a,b,λ,σ)	4.1513	8.8050	0.3087	8.8088
	(1.4516)	(2.6397)	(0.0460)	(4.0961)
PGW(α,λ,γ)	0.4253	0.1364	1.5564	
	(0.0996)	(0.0359)	(0.2212)	
MONH(λ,α,β)	1.1844	0.5025	6.3566	
	(0.5670)	(0.0480)	(3.1274)	
EW(β,λ,γ)	0.4854	0.5421	3.9736	
	(0.1821)	(0.0619)	(1.0804)	
MW(α,β,λ)	0.0030	0.1979	2.2188	
	(0.0009)	(0.0063)	(0.6607)	
ENH(α,β,λ)	0.6003	0.4002	1.7880	
	(0.0841)	(0.1614)	(0.3419)	
NH(α,λ)	0.9134	0.1236		
	(0.1475)	(0.0344)		
Chen(β,λ)	0.1106	0.3538		
	(0.0152)	(0.0123)		
Weibull(α,λ)	9.5470	1.0490		
	(0.8499)	(0.0676)		
FW(α,β)	0.0325	2.1553		
	(0.0026)	(0.2490)		

Chen and Balakrishnan (1995) constructed the corrected Cramér-von Mises and Anderson-Darling statistics. We adopt these statistics, where we have a random sample $x_1,\ldots,x_n$ with empirical distribution function $F_n(x)$, and require to test if the sample comes from a special distribution. The Cramér-von Mises (W^*) and Anderson-Darling (A^*) statistics are given by

$$\begin{aligned} W^* &= \left\{n\int_{-\infty}^{+\infty}\{F_n(x)-F(x;\widehat{\theta}_n)\}^2 dF(x;\widehat{\theta}_n)\right\}\left(1+\frac{0.5}{n}\right) \\ &= W^2\left(1+\frac{0.5}{n}\right) \end{aligned}$$

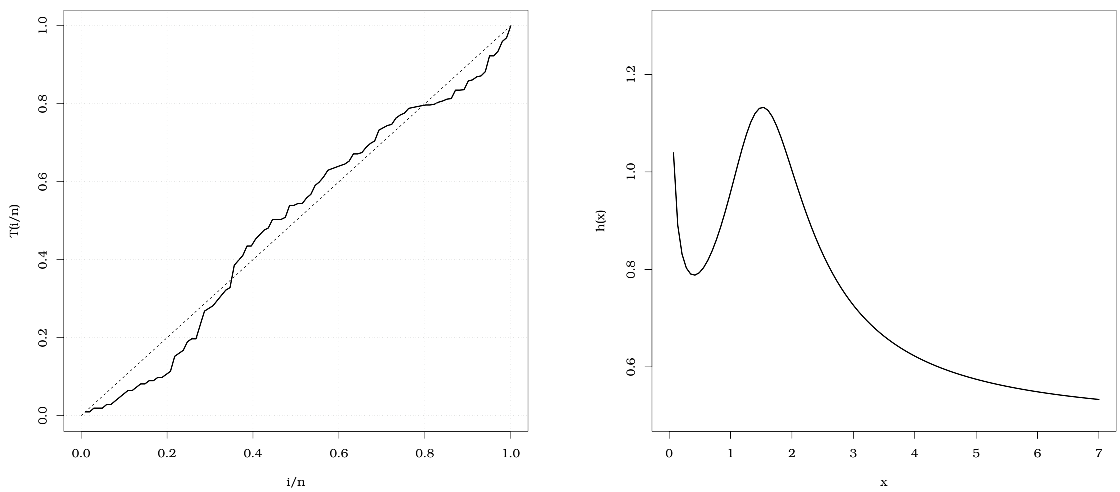
TABLE VI
The MLEs of the model parameters for the stress-rupture data and the corresponding SEs in parentheses.

Distributions	**Estimates**			
EPGW($\alpha,\beta,\lambda,\gamma$)	0.1349	0.1022	0.0415	6.6681
	(0.0171)	(0.0104)	(0.0154)	(0.0136)
Kw-W(a,b,c,β)	0.7029	0.2175	1.0118	4.3625
	(0.1620)	(0.1038)	(0.0027)	(2.1072)
BW(a,b,α,λ)	0.7482	0.2305	1.1275	0.3245
	(0.1608)	(0.0338)	(0.0858)	(0.0856)
BFr(a,b,λ,σ)	0.3984	5.0048	0.4208	9.1684)
	(0.1702)	(1.4399)	(0.0534)	(4.3072)
PGW(α,λ,γ)	1.2659	0.7182	0.8696	
	(0.4483)	(0.3485)	(0.1039)	
MONH(λ,α,β)	0.0146	17.7252	0.2122	
	(0.0067)	(7.7187)	(0.0618)	
EW(β,λ,γ)	0.8488	1.0419	0.8171	
	(0.2981)	(0.2511)	(0.3157)	
MW(α,β,λ)	0.0027	0.2259	7.0190	
	(0.0008)	(0.0076)	(1.5244)	
ENH(α,β,λ)	1.0732	0.7762	0.8426	
	(0.2760)	(0.3582)	(0.1238)	
NH(α,λ)	0.8898	1.1810		
	(0.1853)	(0.4270)		
Chen(β,λ)	0.5410	0.5303		
	(0.0585)	(0.0321)		
Weibull(α,λ)	0.9919	0.9259		
	(0.1121)	(0.0726)		
FW(α,β)	0.3287	0.0838		
	(0.0246)	(0.0133)		

and

$$\begin{aligned} A^* &= \left\{ n\int_{-\infty}^{+\infty} \frac{\{F_n(x)-F(x;\widehat{\theta}_n)\}^2}{\{F(x;\widehat{\theta}_n)[1-F(x;\widehat{\theta}_n)]\}} dF(x;\widehat{\theta}_n) \right\} \left(1+\frac{0.75}{n}+\frac{2.25}{n^2}\right) \\ &= A^2\left(1+\frac{0.75}{n}+\frac{2.25}{n^2}\right), \end{aligned}$$

respectively, where $F_n(x)$ is the empirical distribution function, $F(x;\hat{\theta}_n)$ is the postulated distribution function evaluated at the MLE $\hat{\theta}_n$ of θ. Note that the statistics W^* and A^* are given by the differences of $F_n(x)$



(a) **(b)**

Figure 5 - The TTT plot **(a)** and EPGW hrf for the stress-rupture failure data **(b)**.

and $F(x;\hat{\theta}_n)$. Thus, the lower are these statistics, we have more evidence that $F(x;\hat{\theta}_n)$ generates the sample. The details to evaluate the statistics W^* and A^* are given by Chen and Balakrishnan (1995).

The W^*, A^* and Kolmogorov-Smirnov (KS) statistics for these models are given in Tables VII and VIII for both data sets. We emphasize that the EPGW model fits the remission times and stress-rupture failure data better than the other models according to these statistics. They indicate that the EPGW distribution yields the best fits in both applications.

More information is provided by the histogram of the data and some fitted density functions for both data sets given in Figure 6. Clearly, in both applications, the new distribution provides a closer fit to the histogram than the other competitive models. The fitted cdfs of these models are also displayed in Figure 7. Finally, we can conclude in the two situations that the EPGW distribution is quite competitive to other well-known and widely used distributions such as the Kw-W, EW and Weibull models.

14 - CONCLUSIONS

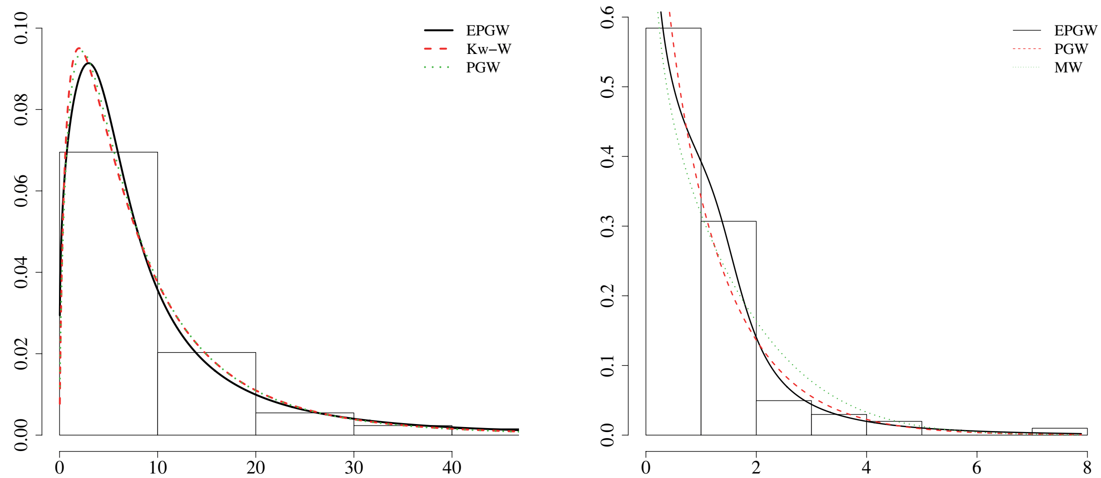
In this paper, we introduce the exponentiated power generalized Weibull (EPGW) model to extend the Weibull distribution. It has a power parameter and its hazard rate function allows constant, decreasing, increasing, upside-down bathtub or bathtub-shaped forms. The proposed distribution contains as special models several well-known lifetime distributions. It can also be derived from a power transform on an exponentiated Nadarajah-Haghighi random variable. Several structural properties of the power generalized Weibull (PGW) distribution have not been studied. However, they can be determined from those of the EPGW distribution. It can also be useful to obtain the properties for other generated families under the PGW baseline. We give a physical motivation for introducing the new distribution if the power parameter is an integer. We obtain some mathematical properties of the EPGW distribution, estimate the model parameters by maximum likelihood and prove empirically its flexibility in two applications to real data. In fact, the new distribution yields a good adjustment in both applications. We note that the EPGW distribution is quite competitive with other lifetime models and can be used effectively to provide better fits than other usual lifetime distributions.

TABLE VII
Goodness-of-fit statistics for the models fitted to the remission times data.

	Statistics		
Distributions	**W***	**A***	**KS**
EPGW	**0.0166**	**0.1148**	**0.0384**
Kw-W	0.0408	0.2744	0.0447
BW	0.0484	0.3313	0.0543
BFr	0.2370	1.5485	0.0987
PGW	0.0352	0.2345	0.0431
MONH	0.0832	0.4864	0.0641
EW	0.0458	0.3163	0.0483
MW	0.3636	2.1554	0.1067
ENH	0.0413	0.2781	0.0422
NH	0.0997	0.6017	0.0927
Chen	0.4430	2.6114	0.1411
Weibull	0.1318	0.7890	0.0695
FW	1.4134	7.8093	0.2085

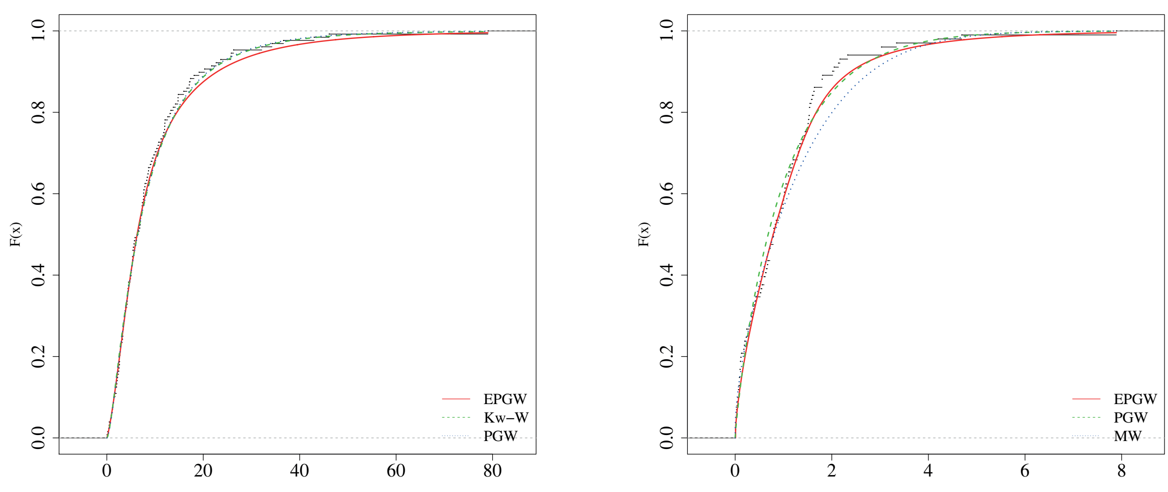
TABLE VIII
Goodness-of-fit statistics for the models fitted to the stress-rupture data.

	Statistics		
Distributions	**W***	**A***	**KS**
EPGW	**0.0722**	**0.4672**	**0.0699**
Kw-W	0.1400	0.8478	0.1017
BW	0.2753	1.5190	0.1005
BFr	0.7116	3.8276	0.1914
PGW	0.1730	0.9930	0.0833
MONH	1.1054	5.9604	0.3068
EW	0.1686	0.9736	0.0875
MW	0.0980	0.7596	0.1292
ENH	0.1670	0.9667	0.0837
NH	0.2053	1.1434	0.0819
Chen	0.1207	0.8756	0.0973
Weibull	0.1987	1.1115	0.0900
FW	1.1130	5.9971	0.3054



(a) (b)

Figure 6 - Histogram and estimated densities of the **(a)** EPGW, Kw-W and PGW models for the remission times data; **(b)** EPGW, PGW and MW models for the stress-rupture data.



(a) (b)

Figure 7 - Estimated and empirical cdfs for **(a)** EPGW, Kw-W and PGW models for the remission times data; **(b)** EPGW, PGW and MW models for the stress-rupture data.

REFERENCES

ANDREWS DF AND HERZBERG AM. 1985. Data: A Collection of Problems from Many Fields for the Student and Research Worker. New York: Springer Series in Statistics, 442 p.

BAGDONAVICIUS V AND NIKULIN M. 2002. Accelerated life models: modeling and statistical analysis. Boca Raton: Chapman and Hall/CRC, 360 p.

BEBBINGTON M, LAI CD AND ZITIKIS R. 2007. A flexible Weibull extension. Reliab Eng and Syst Safe 92: 719-726.

CHEN G AND BALAKRISHNAN N. 1995. A general purpose approximate goodness-of-fit test. J Qual Technol 27: 154-161.

CHEN Z. 2000. A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. Stat Probabil Lett 49: 155-161.

CORDEIRO GM AND CASTRO MD. 2011. A new family of generalized distributions. J Stat Comput Sim 81: 883-898.

DIMITRAKOPOULOU T, ADAMIDIS K AND LOUKAS S. 2007. A lifetime distribution with an upside-down bathtub-shaped hazard function. IEEE T Reliab 121: 308-311.

EUGENE N, LEE C AND FAMOYE F. 2002. Beta-normal distribution and its applications. Commun Stat Theory Methods 31: 497-512.

GOMES O, COMBES C AND DUSSAUCHOY A. 2008. Parameter estimation of the generalized gamma distribution. Math Comput Simul 79: 963-995.

GRAHAM RL. 1994. Concrete mathematics: a foundation for computer science, 2nd ed., Boston: Addison-Wesley, 672 p.

GUPTA R, GUPTA P AND GUPTA R. 1998. Modeling failure time data by Lehman alternatives. Commun Stat Theory Methods 27: 887-904.

LAI CD. 2013. Constructions and applications of lifetime distributions. Appl Stoch Model Bus 29: 127-140.

LEE E AND WANG J. 2003. Statistical Methods for Survival Data Analysis, 3rd ed., New York: Wiley, 534 p.

LEMONTE AJ. 2013. A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function. Comput Stat Data An 62: 149-170.

MACGILLIVRAY H. 1986. Skewness and asymmetry: measures and orderings. Ann Stat 14: 994-1011.

NADARAJAH S, BAKOUCH H AND TAHMASBI R. 2011. A generalized Lindley distribution. Sankhya Ser B 73: 331-359.

NADARAJAH S AND HAGHIGHI F. 2011. An extension of the exponential distribution. Statistics 45: 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. J Math Sci 133: 1333-1341.

NIKULIN M AND HAGHIGHI F. 2009. On the power generalized Weibull family. Metron 67: 75-86.

TAHIR MH AND NADARAJAH S. 2015. Parameter induction in continuous univariate distributions: Well-established G families. An Acad Bras Cienc 87: 539-568.

XIE M, TANG Y AND GOH T. 2002. A modified Weibull extension with bathtub-shaped failure rate function. Reliab Eng and Syst Safe 76: 279-285.