

The Exponentiated Generalized Class of Distributions

Gauss M. Cordeiro¹, Edwin M. M. Ortega^{2*} and Daniel C. C. da Cunha¹

¹Universidade Federal de Pernambuco and ²Universidade de São Paulo

Abstract: We propose a new method of adding two parameters to a continuous distribution that extends the idea first introduced by Lehmann (1953) and studied by Nadarajah and Kotz (2006). This method leads to a new class of exponentiated generalized distributions that can be interpreted as a double construction of Lehmann alternatives. Some special models are discussed. We derive some mathematical properties of this class including the ordinary moments, generating function, mean deviations and order statistics. Maximum likelihood estimation is investigated and four applications to real data are presented.

Key words: Exponentiated distribution, exponentiated generalized beta distribution, Kampé de Fériet function, Lauricella function, probability weighted moment.

1. Introduction

Gupta *et al.* (1998) first proposed a generalization of the standard exponential distribution, called the exponentiated exponential (EE) distribution, defined by the cumulative distribution function (cdf) $F(x) = (1 - e^{-\lambda x})^\alpha$ for $x > 0$, $\lambda > 0$ and $\alpha > 0$. This equation is simply the α th power of the standard exponential cumulative distribution. For a full discussion and some of its mathematical properties, see Gupta and Kundu (2001). In a similar manner, Nadarajah and Kotz (2006) proposed the exponentiated gamma (EG), exponentiated Fréchet (EF) and exponentiated Gumbel (EGu) distributions, although the way they defined the cdf of the last two distributions is slightly different. For instance, the EGu cumulative distribution (for $-\infty < x < \infty$) is defined by

$$F(x) = 1 - [1 - \exp\{-\exp(-\frac{x - \mu}{\sigma})\}]^\alpha,$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$.

*Corresponding author.

In this article, we propose a new class of distributions that extend the exponentiated type distributions and obtain some of its structural properties. Given a continuous cdf $G(x)$, we define the exponentiated generalized (EG) class of distributions by

$$F(x) = [1 - \{1 - G(x)\}^\alpha]^\beta, \quad (1)$$

where $\alpha > 0$ and $\beta > 0$ are two additional shape parameters. We note that there is no complicated function in (1) in contrast with the beta generalized family (Eugene *et al.*, 2002), which also includes two extra parameters but involves the beta incomplete function. (1) has tractable properties especially for simulations, since its quantile function takes a simple form, namely $x = Q_G([1 - (1 - u^{1/\beta})^{1/\alpha}])$, where $Q_G(u)$ is the baseline quantile function.

The baseline distribution $G(x)$ is clearly a special case of (1) when $\alpha = \beta = 1$. Setting $\alpha = 1$ gives the exponentiated type distributions defined by Gupta *et al.* (1998). Further, the EE and E Γ distributions are obtained by taking $G(x)$ to be the exponential and gamma cumulative distributions, respectively. For $\beta = 1$ and if $G(x)$ is the Gumbel and Fréchet cumulative distributions, we obtain the EGu and EF distributions, respectively, as defined by Nadarajah and Kotz (2006). Thus, the class of distributions (1) extends both exponentiated type distributions. The probability density function (pdf) of the new class has the form

$$f(x) = \alpha\beta \{1 - G(x)\}^{\alpha-1} [1 - \{1 - G(x)\}^\alpha]^{\beta-1} g(x). \quad (2)$$

The EG family of densities (2) allows for greater flexibility of its tails and can be widely applied in many areas of engineering and biology. We study some mathematical properties of the class (2) because it extends several well-known distributions in the literature. Note that even if $g(x)$ is a symmetric distribution, the distribution $f(x)$ will not be a symmetric distribution. The two extra parameters in (2) can control both tail weights and possibly adding entropy to the center of the EG density function.

Hereafter, we define the exponentiated-G (“Exp-G” for short) distribution for an arbitrary parent distribution $G(x)$, say $X \sim \text{Exp}^c G$, if X has cumulative and density functions given by $H_c(x) = G(x)^c$ and $h_c(x) = c g(x) G(x)^{c-1}$, respectively. This is also called the Lehmann type I distribution, say $\text{Exp}^c(G)$. For $c > 1$ and $c < 1$ and for larger values of x , the multiplicative factor $c G(x)^{c-1}$ is greater and smaller than one, respectively. The reverse assertion is also true for smaller values of x . The latter immediately implies that the ordinary moments associated with the density function $h_c(x)$ are strictly larger (smaller) than those associated with the density $g(x)$ when $c > 1$ ($c < 1$).

Note that there is a dual transformation $\text{Exp}^c(1 - G)$, referred to as the Lehmann type II distribution corresponding to the cdf $F(x) = 1 - \{1 - G(x)\}^c$.

Thus, (1) encompasses both Lehmann type I ($\text{Exp}^\beta G$ for $\alpha = 1$) and Lehmann type II ($\text{Exp}^\alpha(1 - G)$ for $\beta = 1$) distributions (Lehmann, 1953). Clearly, the double construction $\text{Exp}^\beta[\text{Exp}^\alpha(1 - G)]$ generates the EG class of distributions. The derivations of several properties of the EG class can be facilitated by this double transformation.

The class of EG distributions shares an attractive physical interpretation whenever α and β are positive integers. Consider a device made of β independent components in a parallel system. Furthermore, each component is made of α independent subcomponents identically distributed according to $G(x)$ in a series system. The device fails if all β components fail and each component fails if any subcomponent fails. Let $X_{j1}, \dots, X_{j\alpha}$ denote the lifetimes of the subcomponents within the j th component, $j = 1, \dots, \beta$, with common cdf $G(x)$. Let X_j denote the lifetime of the j th component and let X denote the lifetime of the device. Thus, the cdf of X is

$$\begin{aligned} P(X \leq x) &= P(X_1 \leq x, \dots, X_\beta \leq x) = P(X_1 \leq x)^\beta = [1 - P(X_1 > x)]^\beta \\ &= [1 - P(X_{11} > x, \dots, X_{1\alpha} > x)]^\beta = [1 - P(X_{11} > x)^\alpha]^\beta \\ &= [1 - \{1 - P(X_{11} \leq x)\}^\alpha]^\beta. \end{aligned}$$

So, the lifetime of the device obeys the EG family of distributions.

The rest of the article is organized as follows. In Section 2, we present four special models of the EG class corresponding to the Fréchet, normal, gamma and Gumbel distributions. Section 3 provides some general useful expansions for the EG density function. Moments of the EG class are derived in Section 4 with applications to eight special models. Generating function and mean deviations are derived in Sections 5 and 6, respectively. Order statistics are studied in Section 7. Maximum likelihood estimation is investigated in Section 8. Applications to four real data sets are performed in Section 9. Some concluding remarks are given in Section 10.

2. Special Models

Here, we discuss some special EG distributions. The density function (2) will be most tractable when the cdf $G(x)$ and the pdf $g(x)$ have simple analytic expressions.

2.1 Exponentiated Generalized Fréchet

The cdf of the Fréchet distribution (for $x > 0$) is $G_{\sigma,\lambda}(x) = \exp\{-(\sigma/x)^\lambda\}$, where $\lambda > 0$ and $\sigma > 0$. Then, we define the exponentiated generalized Fréchet

(EGF) cumulative distribution (for $x > 0$) from (1) by

$$F(x) = [1 - (1 - \exp\{-(\frac{\sigma}{x})^\lambda\})^\alpha]^\beta, \quad (3)$$

where $\lambda > 0$, $\sigma > 0$, $\alpha > 0$ and $\beta > 0$. The EGF density function can be obtained from (2) as

$$f(x) = \alpha \beta \lambda \sigma^\lambda x^{-(\lambda+1)} \exp[-(\frac{\sigma}{x})^\lambda] \{1 - \exp[-(\frac{\sigma}{x})^\lambda]\}^{\alpha-1} \\ \times [1 - \{1 - \exp[-(\frac{\sigma}{x})^\lambda]\}^\alpha]^{\beta-1}.$$

Plots of the EGF density function for selected parameter values are given in Figure 1.

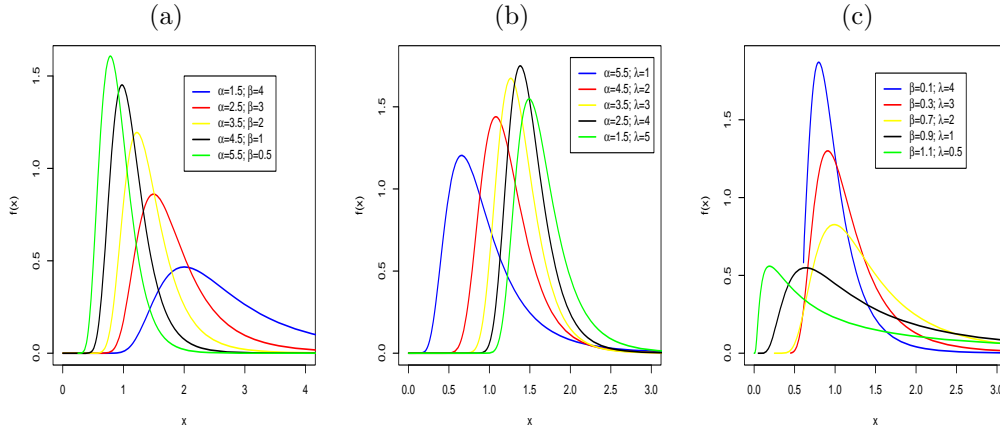


Figure 1: Plots of the EGF density function for some parameter values. (a) For $\sigma = 1.5$ and $\lambda = 2.0$. (b) For $\beta = 1.5$ and $\sigma = 1.5$. (c) For $\alpha = 1.5$ and $\sigma = 1.5$

2.2 Exponentiated Generalized Normal

Let $\Phi(\cdot)$ and $\phi(\cdot)$ denote the standard normal cumulative and density functions, respectively. The exponentiated generalized normal (EGN) cumulative distribution is

$$F(x) = [1 - \{1 - \Phi(\frac{x - \mu}{\sigma})\}^\alpha]^\beta, \quad (4)$$

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, $\alpha > 0$ and $\beta > 0$. The EGN density function becomes

$$f(x) = \alpha \beta \sigma^{-1} \{1 - \Phi(\frac{x - \mu}{\sigma})\}^{\alpha-1} [1 - \{1 - \Phi(\frac{x - \mu}{\sigma})\}^\alpha]^{\beta-1} \phi(\frac{x - \mu}{\sigma}). \quad (5)$$

Plots of the EGN density function for some parameter values are given in Figure 2.

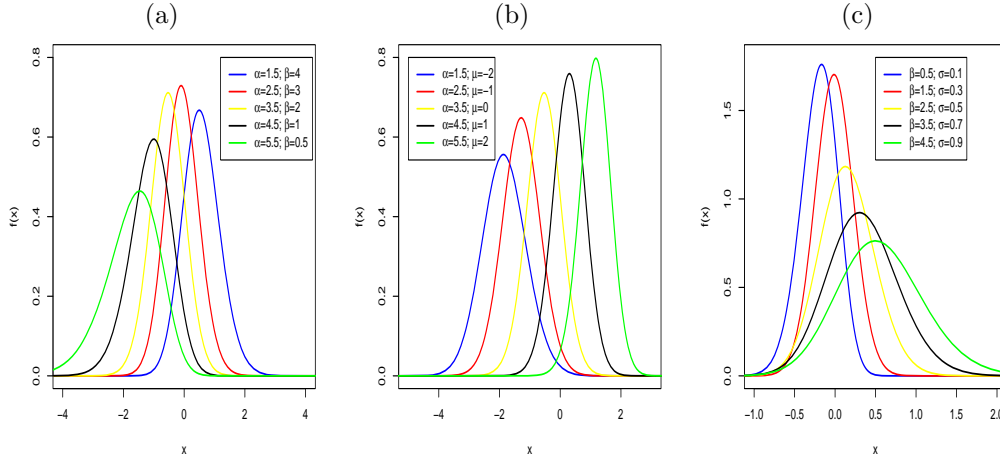


Figure 2: Plots of the EGN density function for some parameter values. (a) For $\mu = 0$ and $\sigma = 1.0$. (b) For $\beta = 2.0$ and $\sigma = 1.0$. (c) For $\alpha = 1.5$ and $\mu = 0$

2.3 Exponentiated Generalized Gamma

The gamma cumulative distribution (for $x > 0$) with shape parameter $a > 0$ and scale parameter $b > 0$ is $G_{a,b}(x) = \gamma(a, bx)/\Gamma(a)$, where $\Gamma(a) = \int_0^\infty w^{a-1} e^{-w} dw$ is the gamma function and $\gamma(a, x) = \int_0^x w^{a-1} e^{-w} dw$ is the incomplete gamma function. The exponentiated generalized gamma (EGGa) cumulative distribution becomes

$$F(x) = \{1 - (1 - \frac{\gamma(a, bx)}{\Gamma(a)})^\alpha\}^\beta, \quad (6)$$

and the associated density function reduces to

$$f(x) = \frac{\alpha\beta b^a x^{a-1} e^{-bx}}{\Gamma(a)} \{1 - \frac{\gamma(a, bx)}{\Gamma(a)}\}^{\alpha-1} \{1 - (1 - \frac{\gamma(a, bx)}{\Gamma(a)})^\alpha\}^{\beta-1}. \quad (7)$$

Plots of the density function (7) for selected parameter values are given in Figure 3.

2.4 Exponentiated Generalized Gumbel

The Gumbel cumulative distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$ is $G_{\mu,\sigma}(x) = \exp\{-\exp(-(x - \mu)/\sigma)\}$. The exponentiated generalized Gumbel (EGGu) cumulative distribution becomes

$$F(x) = \{1 - [1 - \exp\{-\exp(-\frac{x - \mu}{\sigma})\}]^\alpha\}^\beta, \quad (8)$$

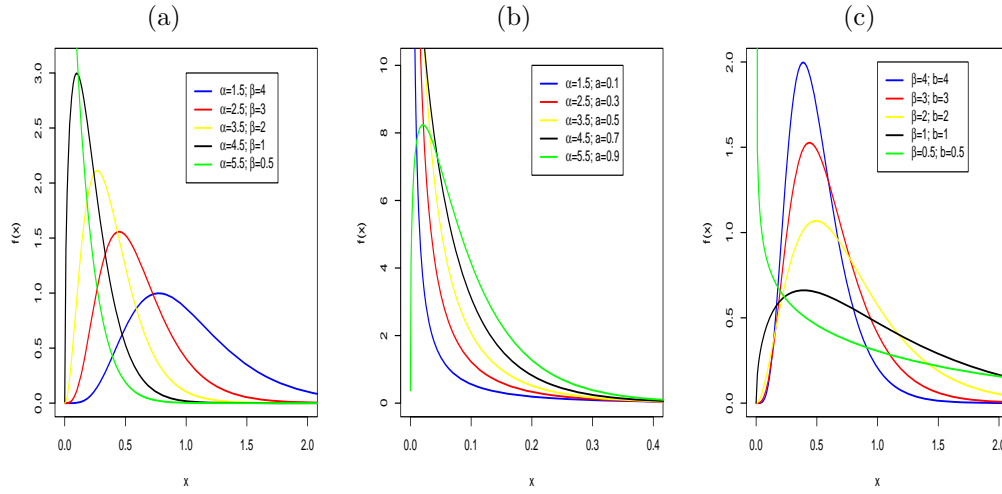


Figure 3: Plots of the EGGa density function for some parameter values. (a) For $a = 1.5$ and $b = 2.0$. (b) For $\beta = 1.5$ and $b = 2.0$. (c) For $\alpha = 1.5$ and $a = 1.5$

and the corresponding pdf reduces to

$$f(x) = \alpha\beta\sigma^{-1} \{1 - [1 - \exp\{-\exp(-\frac{x-\mu}{\sigma})\}]^\alpha\}^{\beta-1} [1 - \exp\{-\exp(-\frac{x-\mu}{\sigma})\}]^{\alpha-1} \times \exp\{-\exp(-\frac{x-\mu}{\sigma})\} \exp\{-\frac{x-\mu}{\sigma}\}. \quad (9)$$

Plots of this density function for selected parameter values are given in Figure 4.

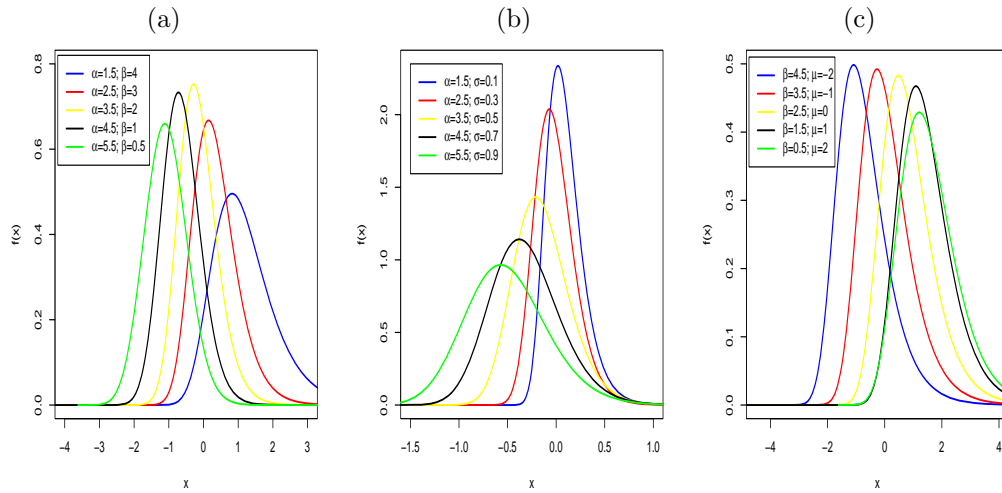


Figure 4: Plots of the EGGu density function for some parameter values. (a) For $\mu = 0$ and $\sigma = 1.0$. (b) For $\beta = 1.5$ and $\mu = 0$. (c) For $\alpha = 1.5$ and $\sigma = 1.0$

3. Expansions for the Density Function

For any real non-integer β , we consider the power series expansion

$$(1 - z)^{\beta-1} = \sum_{k=0}^{\infty} \frac{(-1)^k \Gamma(\beta)}{\Gamma(\beta - k) k!} z^k, \quad (10)$$

which is valid for $|z| < 1$. Applying (10) in (1) and using the binomial expansion for a positive real power yields

$$F(x) = \sum_{j=0}^{\infty} w_j G(x)^j, \quad (11)$$

where the coefficients $w_j = w_j(\alpha, \beta)$ are

$$w_j = \sum_{k=0}^{\infty} \frac{(-1)^{k+j} \Gamma(\beta + 1) \Gamma(k\alpha + 1)}{\Gamma(\beta - k) k! j!}.$$

(11) gives the generated cdf $F(x)$ distribution as an infinite power series of the parent $G(x)$. Using again the series expansion (10), we can express (2) (for α real non-integer) as

$$f(x) = \alpha\beta g(x) \sum_{j=0}^{\infty} t_j G(x)^j, \quad (12)$$

where the coefficients $t_j = t_j(\alpha, \beta)$ are

$$t_j = \frac{(-1)^j \Gamma(\beta)}{j!} \sum_{k=0}^{\infty} \frac{(-1)^k \Gamma((k+1)\alpha)}{\Gamma(\beta - k) \Gamma((k+1)\alpha - j) k!}.$$

Further, (12) can be rewritten as

$$f(x) = \sum_{j=0}^{\infty} t_j^* h_{j+1}(x), \quad (13)$$

where $t_j^* = \alpha\beta t_j / (k+1)$ and $h_{j+1}(x) = (j+1)g(x)G(x)^j$ is the $\text{Exp}^{j+1}(G)$ distribution.

(13) reveals that the EG density function is a linear combination of Exp-G density functions. Thus, some structural properties of the EG class of distributions, such as ordinary and incomplete moments and generating function, can be obtained from well-established properties of the Exp-G distributions.

The properties of Exp-G distributions have been studied by many authors in recent years, see Mudholkar and Srivastava (1993) and Mudholkar *et al.* (1996) for

exponentiated Weibull, Gupta *et al.* (1998) for exponentiated Pareto, Gupta and Kundu (1999) for exponentiated exponential, Nadarajah (2005) for exponentiated Gumbel, Shirke and Kakade (2006) for exponentiated log-normal and Nadarajah and Gupta (2007) for exponentiated gamma distributions. See, also, Nadarajah and Kotz (2006), among others. (11), (12) and (13) are the main results of this section.

4. Moments

Hereafter, we shall assume that G is the cdf of a random variable X and that F is the cdf of a random variable Y having density function (2). The moments of the EG distribution can be obtained from the (r, j) th probability weighted moment (PWM) of X defined by

$$\tau_{r,j} = E[X^r G(X)^j] = \int_{-\infty}^{\infty} x^r G(x)^j g(x) dx. \quad (14)$$

In fact, we have

$$E(Y^r) = \alpha \beta \sum_{j=0}^{\infty} t_j \tau_{r,j}. \quad (15)$$

Thus, the moments of any EG distribution can be expressed as an infinite weighted sum of PWMs of the parent distribution. A second formula for $\tau_{r,j}$ can be based on the parent quantile function $Q_G(x) = G^{-1}(x)$. Setting $G(x) = u$, we obtain

$$\tau_{r,j} = \int_0^1 Q_G(u)^r u^j du, \quad (16)$$

where the integral is now calculated over $(0, 1)$.

The PWMs for various distributions will be determined in the following sections using alternatively (14) and (16).

4.1 EGF

The expansion for (3) reduces to $F(x) = \sum_{j=0}^{\infty} w_j G_{\sigma^*, \lambda}(x)$, where $\sigma^* = \sigma j^{\frac{1}{\lambda}}$ and $G_{\sigma^*, \lambda}(x)$ is the Fréchet cumulative function with parameters σ^* and λ . This equation reveals that the EGF cumulative function can be expressed as an infinite mixture of Fréchet cdf's. Correspondingly, the EGF density function follows a similar mixture

$$f(x) = \sum_{j=0}^{\infty} w_j g_{\sigma^*, \lambda}(x) = \alpha \beta \lambda \sigma^\lambda x^{-(\lambda+1)} \sum_{j=0}^{\infty} t_j \exp\left\{-(j+1)\left(\frac{\sigma}{x}\right)^\lambda\right\},$$

where $g_{\sigma^*,\lambda}(x) = dG_{\sigma^*,\lambda}(x)/dx$.

The (r, j) th PWM of the Fréchet distribution is

$$\tau_{r,j} = \lambda \sigma^\lambda \int_0^\infty x^{r-(\lambda+1)} \exp\left\{-\left(j+1\right) \left(\frac{\sigma}{x}\right)^\lambda\right\} dx.$$

Setting $u = (j+1) (\sigma/x)^\lambda$, $\tau_{r,j}$ reduces to

$$\tau_{r,j} = \frac{\sigma^r}{(j+1)^{1-r/\lambda}} \int_0^\infty u^{-r/\lambda} \exp(-u) du.$$

The integral converges absolutely for $r < \lambda$. Finally, for $r < \lambda$,

$$E(Y^r) = \alpha \beta \sigma^r \Gamma\left(1 - \frac{r}{\lambda}\right) \sum_{j=0}^\infty \frac{t_j}{(j+1)^{1-r/\lambda}}.$$

4.2 EGN

The moments of $X \sim \mathcal{N}(\mu, \sigma)$ can be obtained from the moments of $Z \sim \mathcal{N}(0, 1)$ using $E(X^r) = \sum_{k=0}^r \mu^{r-t} \sigma^r E(Z^r)$, and then we can work with the standard normal distribution. Consider the error function $\text{erf}(\cdot)$ defined by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt.$$

From equation

$$\Phi(x) = \frac{1}{2} \left\{1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right\}, \tag{17}$$

we can expand the EGN cumulative function (4) (with $\mu = 0$ and $\sigma = 1$) as

$$F(x) = \sum_{j=0}^\infty 2^{-j} w_j \sum_{l=0}^j \binom{j}{l} \text{erf}\left(\frac{x}{\sqrt{2}}\right)^l.$$

From the series expansion for the error function $\text{erf}(\cdot)$

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \sum_{m=0}^\infty \frac{(-1)^m x^{2m+1}}{(2m+1) m!},$$

we obtain a series expansion for (5) (with $\mu = 0$ and $\sigma = 1$) given by

$$\begin{aligned} f(x) &= \alpha \beta \exp\left(-\frac{x^2}{2}\right) 2^{1/2} \pi^{1/2} \sum_{j=0}^\infty 2^{-j} t_j \sum_{l=0}^j \binom{j}{l} 2^{l/2} \pi^{-l/2} \\ &\times \sum_{m_1=0}^\infty \cdots \sum_{m_l=0}^\infty \frac{(-1)^{m_1+\cdots+m_l} x^{2(m_1+\cdots+m_l)+l}}{2^{m_1+\cdots+m_l} (2m_1+1) \cdots (2m_l+1) m_1! \cdots m_l!}. \end{aligned}$$

From Cordeiro and Nadarajah (2011, (11)), the normal PWMs can be expressed in terms of the Lauricella functions of type A (Exton, 1978; Aarts, 2000) defined by

$$\begin{aligned} F_A^{(n)}(a; b_1, \dots, b_n; c_1, \dots, c_n; x_1, \dots, x_n) \\ = \sum_{m_1=0}^{\infty} \dots \sum_{m_n=0}^{\infty} \frac{(a)_{m_1+\dots+m_n} (b_1)_{m_1} \dots (b_n)_{m_n} x_1^{m_1} \dots x_n^{m_n}}{(c_1)_{m_1} \dots (c_n)_{m_n} m_1! \dots m_n!}, \end{aligned}$$

where $(a)_i = a(a+1)\dots(a+i-1)$ is the ascending factorial given by (with the convention that $(a)_0 = 1$). The (r, j) th PWM of the normal distribution is

$$\begin{aligned} \tau_{r,j} = 2^{r/2} \pi^{-(j+1/2)} \sum_{\substack{l=0 \\ (r+j-l) \text{ even}}}^j \binom{j}{l} 2^{-l} \pi^l \Gamma\left(\frac{r+j-l+1}{2}\right) \\ \times F_A^{(j-l)}\left(\frac{r+j-l+1}{2}; \frac{1}{2}, \dots, \frac{1}{2}; \frac{3}{2}, \dots, \frac{3}{2}; -1, \dots, -1\right). \end{aligned}$$

This equation holds when $r+j-l$ is even and it vanishes when $r+j-l$ is odd. So, any EGN moment can be expressed as an infinite weighted linear combination of Lauricella functions of type A.

4.3 EGG

Using the power series expansion for the incomplete gamma function

$$G_{a,b}(x) = \frac{(bx)^a}{\Gamma(a)} \sum_{m=0}^{\infty} \frac{(-bx)^m}{(a+m)m!},$$

we obtain the following series expansion for (7)

$$f(x) = \frac{\alpha \beta b^a x^{a-1} e^{-bx}}{\Gamma(a)} \sum_{j=0}^{\infty} t_j \frac{(bx)^{aj}}{\Gamma(a)^j} \sum_{m_1=0}^{\infty} \dots \sum_{m_j=0}^{\infty} \frac{(-bx)^{m_1+\dots+m_j}}{(a+m_1)\dots(a+m_j) m_1! \dots m_j!},$$

respectively. From (9) of Cordeiro and Nadarajah (2011), the quantities $\tau_{r,j}$ can be determined as

$$\tau_{r,j} = \frac{\Gamma(r+a(j+1))}{a^j b^r \Gamma(a)^{j+1}} F_A^{(j)}(r+a(j+1); a, \dots, a; a+1, \dots, a+1, -1, \dots, -1).$$

4.4 EGGu

The expansion for the density function (9) becomes

$$f(x) = \sum_{j=0}^{\infty} w_j g_{\mu^*, \sigma}(x),$$

where $g_{\mu^*, \sigma}(x)$ is the Gumbel density function with parameters $\mu^* = \mu - \sigma \log(j)$ and σ . So, the EGGu density function can be expressed as an infinite mixture of Gumbel densities. The PWMs of the Gumbel distribution are

$$\tau_{r,j} = \sigma^{-1} \int_{-\infty}^{\infty} x^r \exp\left\{-\left(\frac{x-\mu}{\sigma}\right)\right\} \exp\left\{-(j+1) \exp\left(-\frac{x-\mu}{\sigma}\right)\right\} dx.$$

Setting $u = \exp\{-(x-\mu)/\sigma\}$, they reduce to

$$\tau_{r,j} = \int_0^{\infty} [\mu - \sigma \log(u)]^r \exp[-(j+1)u] du.$$

Using the binomial expansion, we have

$$\tau_{r,j} = \sum_{l=0}^r \binom{r}{l} (-\sigma)^l \mu^{r-l} \int_0^{\infty} \log(u)^l \exp[-(j+1)u] du.$$

By (2.6.21.1) in Prudnikov *et al.* (1986), the integral becomes

$$\int_0^{\infty} \log(u)^l \exp[-(j+1)u] du = \left(\frac{\partial}{\partial a}\right)^l [(j+1)^{-a} \Gamma(a)] \Big|_{a=1}.$$

Finally,

$$E(Y^r) = \alpha \beta \sum_{j=0}^{\infty} t_j \sum_{l=0}^r \binom{r}{l} \mu^{r-l} (-\sigma)^l \left(\frac{\partial}{\partial a}\right)^l [(j+1)^{-a} \Gamma(a)] \Big|_{a=1}.$$

4.5 Exponentiated generalized exponential

For the exponential cumulative function $F(x) = 1 - \exp(-\lambda x)$, $\lambda > 0$, the quantile function is $Q_G(u) = -\lambda^{-1} \log(1-u)$. From (16), $\tau_{r,j}$ becomes

$$\tau_{r,j} = (-1)^r \lambda^{-r} \int_0^1 u^j [\log(1-u)]^r du = r! \lambda^r \sum_{m=0}^{\infty} \frac{(-1)^{m+r} \binom{j}{m}}{(m+1)^{r+1}},$$

and then the moments of the exponential generalized exponential (EGE) distribution can be readily determined from (15).

4.6 Exponentiated generalized logistic

For the exponentiated generalized logistic (EGL) distribution, the parent cumulative function is $G(x) = \{1 + \exp(-x)\}^{-1}$. Using a result from Prudnikov *et al.* (1986, Section 2.6.13, (4)), we can write from (16)

$$\tau_{r,j} = \left(\frac{\partial}{\partial t} \right)^r B(t+j+1, 1-t) \Big|_{t=0},$$

where $B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function. The r th moment of the EGL distribution comes from (15) as

$$E(Y^r) = \alpha \beta \sum_{j=0}^{\infty} t_j \left(\frac{\partial}{\partial t} \right)^r B(t+j+1, 1-t) \Big|_{t=0}.$$

4.7 Exponentiated Generalized Pareto

For the exponentiated generalized Pareto (EGPa) distribution, where the baseline cumulative function is $G(x) = 1 - (1+x)^{-\nu}$ (for $\nu > 0$), we immediately obtain from (15) and (16)

$$E(Y^r) = \alpha \beta \sum_{j,m=0}^{\infty} (-1)^{r+m} \binom{r}{m} t_j B(j, 1 - \frac{m}{\nu}).$$

4.8 Exponentiated Generalized Beta

The cdf of the beta distribution (for $0 < x < 1$) is

$$G_{a,b}(x) = I_x(a, b) = \frac{1}{B(a, b)} \int_0^x t^{a-1} (1-t)^{b-1} dt,$$

where $a > 0$ and $b > 0$. The exponentiated generalized beta (EGB) cumulative distribution is $F(x) = [1 - \{1 - I_x(a, b)\}^\alpha]^\beta$ and the corresponding density function becomes

$$f(x) = \frac{\alpha \beta x^{a-1} (1-x)^{b-1}}{B(a, b)} \{1 - I_x(a, b)\}^{\alpha-1} [1 - \{1 - I_x(a, b)\}^\alpha]^{\beta-1}.$$

Using the incomplete beta power series for real non-integer $b > 0$

$$I_x(a, b) = \frac{x^a}{B(a, b)} \sum_{m=0}^{\infty} \frac{(1-b)_m x^m}{(a+m) m!},$$

we can obtain an expansion for $f(x)$ as follows

$$f(x) = \frac{\alpha \beta x^{a-1} (1-x)^{b-1}}{B(a,b)} \sum_{j=0}^{\infty} \frac{t_j x^{ja}}{B(a,b)^j} \sum_{m_1=0}^{\infty} \cdots \sum_{m_j=0}^{\infty} \frac{(1-b)_{m_1} \cdots (1-b)_{m_j} x^{m_1+\cdots+m_j}}{(a+m_1) \cdots (a+m_j) m_1! \cdots m_j!}.$$

The PWMs of the beta distribution can be expressed in terms of the generalized Kampé de Fériet function (Exton, 1978; Mathai, 1993; Aarts, 2000; Chaudhry and Zubair, 2002). They are given by (Cordeiro and Nadarajah, 2011)

$$\tau_{r,j} = a^{-j} B(a,b)^{-(j+1)} B(b,r+a(j+1)) F_{1:1}^{1:2}((r+a(j+1)) : (1-b,a); \cdots ; (1-b,a) : (b+r+a(j+1)) : (a+1); \cdots ; (a+1); 1, \cdots, 1),$$

where the generalized Kampé de Fériet function is defined by

$$F_{C:D}^{A:B}((a) : (b_1); \cdots ; (b_n); (c) : (d_1); \cdots ; (d_n); x_1, \cdots, x_n) = \sum_{m=1=0}^{\infty} \cdots \sum_{m_n=0}^{\infty} \frac{((a))_{m_1+\cdots+m_n} ((b_1))_{m_1} \cdots ((b_n))_{m_n} x_1^{m_1} \cdots x_n^{m_n}}{((c))_{m_1+\cdots+m_n} ((d_1))_{m_1} \cdots ((d_n))_{m_n} m_1! \cdots m_n!},$$

where $a = (a_1, \cdots, a_A)$, $b_i = (b_{i,1}, \cdots, b_{i,B})$, $c = (c_1, \cdots, c_C)$, $d_i = (d_{i,1}, \cdots, d_{i,D})$ for $i = 1, \cdots, n$ and

$$((f))_k = ((f_1, \cdots, f_p))_k = (f_1)_k \cdots (f_p)_k.$$

Hence, the EGB moments can be expressed as an infinite weighted linear combination of generalized Kampé de Fériet functions.

5. Generating Function

Here, we provide three formulae for the moment generating function (mgf) $M(s) = E[\exp(sY)]$ of Y . Clearly, the first one is simply

$$M(s) = \sum_{r=0}^{\infty} \frac{\mu'_r}{r!} s^r, \tag{18}$$

where $\mu'_r = E(Y^r)$ is obtained from (15).

A second formula for $M(s)$ comes from (13) as

$$M(s) = \sum_{j=0}^{\infty} t_j^* M_{j+1}(s), \tag{19}$$

where $M_{j+1}(s)$ is the generating function of the $\text{Exp}^{j+1}(G)$ distribution. Hence, $M(s)$ can be determined from the Exp-G generating function.

A third formula for $M(s)$ can be derived from (12) as

$$M(s) = \alpha\beta \sum_{j=0}^{\infty} t_j \rho_j(s), \quad (20)$$

where the quantity $\rho_j(s) = \int_{-\infty}^{\infty} \exp(sx) G(x)^j g(x) dx$ follows from the baseline quantile function $Q_G(u) = G^{-1}(u)$ as

$$\rho_j(s) = \int_0^1 \exp[s Q_G(u)] u^j du. \quad (21)$$

We can derive the mgf of several EG distributions directly from (20) and (21). For example, the mgf's of the EGE (with parameter λ), EGL and EGP α (with parameter $\nu > 0$) distributions are given by

$$M(s) = \alpha\beta \sum_{j=0}^{\infty} t_j B(j+1, 1-\lambda s), \quad M(s) = \alpha\beta \sum_{j=0}^{\infty} t_j B(s+j+1, 1-s),$$

and

$$M(s) = \alpha\beta e^{-s} \sum_{j,r=0}^{\infty} \frac{t_j B(j+1, 1-r\nu^{-1})}{r!} s^r,$$

respectively.

Clearly, three representations for the characteristic function (chf) $\phi(s) = E[\exp(isX)]$ of the EG distributions are derived from (18)-(20) by $\phi(s) = M(is)$, where $i = \sqrt{-1}$.

6. Mean Deviations

The mean deviations about the mean ($\delta_1(Y) = E(|Y - \mu'_1|)$) and about the median ($\delta_2(Y) = E(|Y - M|)$) of Y can be expressed as

$$\delta_1(Y) = 2\mu'_1 F(\mu'_1) - 2m_1(\mu'_1) \quad \text{and} \quad \delta_2(Y) = \mu'_1 - 2m_1(M), \quad (22)$$

respectively, where $M = Q_G([1 - (1 - 2^{-\beta})^{1/\alpha}])$ is the median of Y , $\mu'_1 = E(Y)$ comes from (15), $F(\mu'_1)$ is easily calculated from (1) and $m_1(z) = \int_{-\infty}^z x f(x) dx$ is the first incomplete moment.

Now, we provide two alternative ways to compute $\delta_1(Y)$ and $\delta_2(Y)$. A general equation for $m_1(z)$ can be derived from (13) as

$$m_1(z) = \sum_{j=0}^{\infty} t_j^* J_{j+1}(z), \quad (23)$$

where

$$J_{j+1}(z) = \int_{-\infty}^z x h_{j+1}(x) dx. \tag{24}$$

(24) is the basic quantity to compute the mean deviations for the GE distributions. The mean deviations (22) depend only on the first incomplete moment of the Exp-G distributions. So, alternative representations for $\delta_1(Y)$ and $\delta_2(Y)$ are

$$\delta_1(Y) = 2\mu'_1 F(\mu'_1) - 2 \sum_{j=0}^{\infty} t_j^* J_{j+1}(\mu'_1) \quad \text{and} \quad \delta_2(Y) = \mu'_1 - 2 \sum_{j=0}^{\infty} t_j^* J_{j+1}(M).$$

A simple application of (23) and (24) refers to the exponentiated generalized Weibull (EGW) distribution. The exponentiated Weibull density function (for $x > 0$) with power parameter $j + 1$, shape parameter c and scale parameter β is

$$h_{j+1}(x) = c(j + 1) \beta^c x^{c-1} \exp\{-(\beta x)^c\} t[1 - \exp\{-(\beta x)^c\}]^j,$$

and then

$$J_{j+1}(z) = c(j + 1) \beta^c \sum_{r=0}^{\infty} (-1)^r \binom{j}{r} \int_0^z x^c \exp\{-(r + 1)(\beta x)^c\} dx.$$

The last integral reduces to the incomplete gamma function and then

$$J_{j+1}(z) = \beta^{-1} \sum_{r=0}^{\infty} \frac{(-1)^r (j + 1) \binom{j}{r}}{(r + 1)^{1+1/c}} \gamma(1 + c^{-1}, (r + 1)(\beta z)^c).$$

A second general formula for $m_1(z)$ can be derived by setting $u = G(x)$ in (13)

$$m_1(z) = \sum_{j=0}^{\infty} (j + 1) t_j^* T_j(z), \tag{25}$$

where $T_j(z)$ is given by

$$T_j(z) = \int_0^{G(z)} Q_G(u) u^j du. \tag{26}$$

For example, we can obtain the mean deviations of the EGE (with parameter λ), EGL and EGP α (with parameter $\nu > 0$) distributions from (25)-(26). The

quantities $T_j(z)$ can be derived from the following integrals (for $a > 0$) using Maple and Mathematica

$$\int_1^a x^j \log(x) dx = \frac{1 - a^{j+1}}{(j+1)^2} + \frac{a^{j+1} \log(a)}{(j+1)},$$

$$\int_0^a x^j \log(x) dx = \frac{a^{j+1}}{(j+1)^2} [(j+1) \log(a) - 1].$$

By changing variables and using the binomial expansion and these integrals, we obtain

$$T_j(z) = \lambda^{-1} \sum_{k=0}^j (-1)^k \binom{j}{k} \left[\frac{1 - \{1 - G(z)\}^{k+1}}{(k+1)^2} + \frac{\{1 - G(z)\}^{k+1} \log\{1 - G(z)\}}{(k+1)} \right],$$

$$T_j(z) = \frac{G(z)^{j+1}}{(j+1)^2} [(j+1) \log\{G(z)\} - 1]$$

$$- \sum_{k=0}^j (-1)^k \binom{j}{k} \left[\frac{1 - \{1 - G(z)\}^{k+1}}{(k+1)^2} + \frac{\{1 - G(z)\}^{k+1} \log\{1 - G(z)\}}{(k+1)} \right],$$

and

$$T_j(z) = \sum_{k=0}^{\infty} \frac{(-1)^k \binom{-\nu^{-1}}{k} G(z)^{k+j+1}}{(k+j+1)} - \frac{G(z)^{j+1}}{(j+1)},$$

for the EGE, EGL and EGP α distributions, respectively.

Applications of these equations can be conducted to obtain Bonferroni and Lorenz curves 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 $q = Q_G([1 - (1 - \pi^{1/\beta})^{1/\alpha}])$ is immediately calculated from the parent quantile function.

7. Order Statistics

The density $f_{i:n}(x)$ of the i th order statistic, for $i = 1, \dots, n$, from independent identically distributed random variables Y_1, \dots, Y_n is given by

$$f_{i:n}(x) = \frac{f(x)}{B(i, n-i+1)} F(x)^{i-1} \{1 - F(x)\}^{n-i}.$$

Substituting (1) and (2) in this equation, we can write

$$f_{i:n}(x) = \frac{\alpha\beta}{B(i, n-i)} g(x) \{1 - G(x)\}^{\alpha-1} [1 - \{1 - G(x)\}^\alpha]^{\beta i - 1}$$

$$\times \{1 - [1 - \{1 - G(x)\}^\alpha]^\beta\}^{n-i}.$$

Using the binomial expansion, $f_{i:n}(x)$ can be expressed as

$$f_{i:n}(x) = \frac{\alpha\beta}{B(i, n-i)} g(x) \{1 - G(x)\}^{\alpha-1} \sum_{k=0}^{n-i} (-1)^k \binom{n-i}{k} \times [1 - \{1 - G(x)\}^\alpha]^{\beta(i+k)-1}.$$

For β real non-integer, by applying (10) to the last term, we obtain

$$f_{i:n}(x) = \frac{\alpha\beta}{B(i, n-i)} g(x) \sum_{k=0}^{n-i} (-1)^k \binom{n-i}{k} \sum_{r=0}^{\infty} \frac{(-1)^r \Gamma(\beta(i+k))}{\Gamma(\beta(i+k)-r) r!} \times \{1 - G(x)\}^{\alpha(r+1)-1},$$

which can be rewritten as

$$f_{i:n}(x) = \frac{\alpha\beta}{B(i, n-i)} g(x) \sum_{l=0}^{\infty} s_l G(x)^l, \tag{27}$$

where the coefficients s_l can be calculated as

$$s_l = s_l(\alpha, \beta, i, n) = \sum_{k=0}^{n-i} \sum_{r=0}^{\infty} \frac{(-1)^{k+r+l} \Gamma(\beta(i+k)) \Gamma(\alpha(r+1)) \Gamma(n-i+1)}{\Gamma(\alpha(r+1)-l) \Gamma(\beta(i+k)-r) \Gamma(n-i-k+1) k! r! l!}.$$

(27) can be rewritten in terms of the Exp-G density functions as

$$f_{i:n}(x) = \frac{\alpha\beta}{B(i, n-i)} \sum_{l=0}^{\infty} s_l h_{l+1}(x). \tag{28}$$

(28) is the main result of this section. So, several mathematical properties of the EG order statistics (like ordinary and incomplete moments, generating function, mean deviations) can be obtained from those properties of the Exp-G distribution as shown previously for the EG distribution.

8. Maximum Likelihood Estimation

We determine the maximum likelihood estimates (MLEs) of the parameters of the EG distribution from complete samples only. Let x_1, \dots, x_n be a random sample of size n from the $EG(\alpha, \beta, \gamma)$ distribution, where γ is a $p1$ vector of unknown parameters in the parent distribution $G(x; \gamma)$. The log-likelihood function

for the vector of parameters $\boldsymbol{\theta} = (\alpha, \beta, \gamma^T)^T$ can be expressed as

$$\begin{aligned} l(\boldsymbol{\theta}) = & n \log(\alpha) + n \log(\beta) + \sum_{i=1}^n \log [g(x_i; \gamma)] + (\alpha - 1) \sum_{i=1}^n \log [1 - G(x_i; \gamma)] \\ & + (\beta - 1) \sum_{i=1}^n \log \{1 - [1 - G(x_i; \gamma)]^\alpha\}, \end{aligned} \quad (29)$$

The log-likelihood can be maximized either directly by using the SAS (Proc NLMixed) or the MaxBFGS routine in the matrix programming language Ox (see, Doornik, 2007) or by solving the nonlinear likelihood equations obtained by differentiating (29). The components of the score vector $U(\boldsymbol{\theta})$ are

$$\begin{aligned} U_\alpha(\boldsymbol{\theta}) &= \frac{n}{\alpha} + \sum_{i=1}^n \log \left\{ 1 - G(x_i; \gamma) \right\} \left\{ 1 - \frac{(\beta - 1)[1 - G(x_i; \gamma)]^\alpha}{1 - [1 - G(x_i; \gamma)]^\alpha} \right\}, \\ U_\beta(\boldsymbol{\theta}) &= \frac{n}{\beta} + \sum_{i=1}^n \log \{1 - [1 - G(x_i; \gamma)]^\alpha\}, \\ U_{\gamma_j}(\boldsymbol{\theta}) &= \sum_{i=1}^n \left\{ \frac{[g(x_i; \gamma)]_{\gamma_j}}{g(x_i; \gamma)} - \frac{(\alpha - 1)[\dot{G}(x_i; \gamma)]_{\gamma_j}}{1 - G(x_i; \gamma)} \right. \\ &\quad \left. + \frac{\alpha(\beta - 1)[1 - G(x_i; \gamma)]^{\alpha-1}[\dot{G}(x_i; \gamma)]_{\gamma_j}}{1 - [1 - G(x_i; \gamma)]^\alpha} \right\}, \end{aligned}$$

where $[g(x_i; \gamma)]_{\gamma_j} = \partial g(x_i; \gamma) / \partial \gamma_j$, $[\dot{G}(x_i; \gamma)]_{\gamma_j} = \partial G(x_i; \gamma) / \partial \gamma_j$ for $j = 1, \dots, p$.

For interval estimation and hypothesis tests on the model parameters, we require the $(p + 2) \times (p + 2)$ observed information matrix $J = J(\boldsymbol{\theta})$ given in the Appendix. Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, the asymptotic distribution of $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is $\mathcal{N}_{(p+2)}(0, I(\boldsymbol{\theta})^{-1})$, where $I(\boldsymbol{\theta})$ is the expected information matrix. In practice, we can replace $I(\boldsymbol{\theta})$ by the observed information matrix evaluated at $\hat{\boldsymbol{\theta}}$ (say $J(\hat{\boldsymbol{\theta}})$). We can construct approximate confidence intervals and confidence regions for the parameters based on the multivariate normal $\mathcal{N}_{(p+2)}(0, J(\hat{\boldsymbol{\theta}})^{-1})$ distribution. The elements of the observed information matrix are in the Appendix.

Further, the likelihood ratio (LR) statistic can be used for comparing the EG distribution with some of its special models. We can compute the maximum values of the unrestricted and restricted log-likelihoods to construct LR statistics for testing some sub-models of the EG distribution. For example, the test of $H_0 : \alpha = 1$ versus $H : H_0 \text{ is not true}$ is equivalent to compare the EG and exponentiated type distributions and the LR statistic reduces to

$$w = 2\{\ell(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) - \ell(1, \hat{\beta}, \hat{\gamma})\},$$

where $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ are the MLEs under H and $\tilde{\beta}$ and $\tilde{\gamma}$ are the estimates under H_0 .

9. Applications

In this section, we use four real data sets to compare the fits of the EG distribution with those of three sub-models, i.e., the two exponentiated type distributions and the parent distribution itself. In each case, the parameters are estimated by maximum likelihood (Section 8) using the subroutine NLMixed in SAS. First, we describe the data sets and give the MLEs (and the corresponding standard errors in parentheses) of the parameters and the values of the Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC) and Bayesian Information Criterion (BIC) statistics. The lower the values of these criteria, the better the fit. Note that over-parameterization is penalized in these criteria, so that the two additional parameters in the EG model do not necessarily lead to smaller values of the AIC, BIC or CAIC statistics. Next, we perform LR tests (Section 8) for formal tests of the additional shape parameters. Finally, we provide histograms of the data sets to show a visual comparison of the fitted density functions.

(i) Ethylene data

These data were taken from a study by the University of São Paulo, ESALQ (Laboratory of Physiology and Post-colheita Biochemistry), which evaluated the effects of mechanical damage on banana fruits (genus *Musa* spp.); see Saavedra del Aguila *et al.* (2010) for more details. The major problem affecting bananas during and after harvest is the susceptibility of the mature fruit to physical damage caused during transport and marketing. Ethylene is a plant hormone important in post-harvest fruit. A high ethylene production can generate a fast senescence of fruit. We use 630 data points on ethylene and assume a normal parent distribution.

(ii) Wheaton River data

The data are the exceedances of flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada. The data consist of 72 exceedances for the years 1958-1984 rounded to one decimal place. These data were analyzed by Akinsete *et al.* (2008).

(iii) Stress level data

The following 101 data points represent the stress-rupture life of 49 kevlar epoxy strands, which were subjected to constant sustained pressure at the 90% stress level until all had failed, so that we have complete data with exact times of failure. The failure times in hours are given by Cooray and Ananda (2008).

(iv) Carbon data

These uncensored data on the breaking stress of carbon fibres (in Gba) are ob-

tained from Nichols and Padgett (2006).

Table 1 gives a descriptive summary of each sample. The carbon data has negative skewness. The ethylene, Wheaton River and stress level data have positive skewness and kurtosis, larger values of these sample moments being apparent in the ethylene data.

Table 1: Descriptive statistics

Data	Mean	Median	Mode	SD	Variance	Skewness	Kurtosis	Min.	Max.
Ethylene	8.29	1.17	0	16.18	261.66	3.85	20.03	0.0	141.05
Wheaton	12.20	9.5	1.7	12.3	151.22	1.5	3.19	0.1	64
Stress level	1.03	0.8	0.02	1.12	1.25	3.05	14.51	0.01	7.89
Carbon	2.76	2.84	1.61	0.89	0.80	-0.14	0.34	0.39	4.90

We compute the MLEs and the AIC, BIC and CAIC information criteria for each data set. For the ethylene data, we compare the fitted EGN model (with parameters α , β and $\gamma = (\mu, \sigma)^T$) with the fitted exponentiated normal (EN), Lehmann II normal (LIIN) and normal distributions. The MLEs of μ and σ for the normal distribution are taking as starting values for the numerical iterative procedure. For the wheaton River data, we compare the fitted EGGu model (with parameters α , β and $\gamma = (\mu, \sigma)^T$) with the fitted EG_u (exponentiated Gumbel), Lehmann II Gumbel (LIIG_u) and Gumbel distributions. The MLEs of μ and σ for the Gumbel distribution are taking as starting values for the iterative procedure. For the stress level data, we compare the EGF model (with parameters α , β and $\gamma = (\lambda, \sigma)^T$) with the fitted exponentiated Fréchet (EF), Lehmann II Fréchet (LIIF) and Fréchet distributions. The MLEs of λ and σ for the Fréchet distribution are taking as starting values for the iterative procedure. Further, for the carbon data, the fitted EGGa distribution (with parameters α , β and $\gamma = (a, b)^T$) is compared with the fitted EG_a, Lehmann II gamma (LIIG_a) and gamma models. Here, the MLEs of a and b for the gamma distribution are taking as starting values.

The results are reported in Table 2. Note that the three information criteria agree on the model's ranking in every case. For the ethylene, Wheaton River and stress level data, the lowest values of the information criteria correspond to the fitted EG distribution. Clearly, for the carbon data, based on the values of these statistics, we can conclude that the top two models are the EGG_a and LIIG_a distributions and the other distributions are far worse.

A formal test for the need of the extra parameters in the EG models can be performed using LR statistics (Section 8). Applying these LR tests to our four data sets yield the results in Table 3. For the carbon data, the additional parameters of the EGG_a distribution are not, in fact, necessary because the LR

Table 2: MLEs and information criteria

Ethylene	α	β	μ	σ	AIC	CAIC	BIC
EGN	0.8780 (0.0472)	455.79 (115.59)	-79.8710 (4.7022)	26.7977 (0.8233)	4846.2	4846.3	4864.0
EN	1 -	1.6688 (0.2667)	2.0018 (2.3450)	16.8665 (0.5433)	5230.1	5230.2	5243.5
LIIN	0.9367 (0.0844)	1 -	8.2365 (1.3425)	16.1102 (0.5336)	5296.0	5296.1	5309.3
Normal	1 -	1 -	8.2923 (0.6440)	16.1631 (0.4553)	5298.2	5298.3	5307.1
Wheaton River	α	β	μ	σ	AIC	CAIC	BIC
EGGu	0.0988 (0.0243)	0.4769 (0.1439)	2.6317 (0.7863)	1.6639 (0.0529)	521.8	522.4	530.9
EGu	1 -	1.0312 (14.99)	6.7165 (15.43)	8.1893 (0.8185)	544.9	545.2	551.7
LIIGu	0.1474 (0.0186)	1 -	0.5571 (0.5291)	1.7214 (0.0544)	524.5	524.8	531.3
Gumbel	1 -	1 -	6.9684 (1.0093)	8.1893 (0.8185)	542.9	543.0	547.4
Stress level	α	β	σ	λ	AIC	CAIC	BIC
EGF	606.36 (184.53)	0.1761 (0.0186)	1095.66 (375.16)	0.3104 (0.0105)	217.4	217.8	227.9
EF	1 -	0.3383 (0.2146)	1.4333 (1.4935)	0.6136 (0.0427)	271.9	272.2	279.8
LIIF	29.4149 (12.0856)	1 -	177.49 (16.90)	0.2289 (0.0209)	227.8	228.0	235.6
Fréchet	1 -	1 -	0.2451 (0.0423)	0.6136 (0.0422)	269.9	270.1	275.2
Carbon	α	β	a	b	AIC	CAIC	BIC
EGGa	181.07 (42.65)	0.6260 (0.2477)	5.4917 (1.5906)	0.3924 (0.3039)	180.1	180.7	188.8
EGa	1 -	0.2490 (0.0481)	24.5269 (1.7579)	6.5514 (1.0355)	183.7	184.0	190.2
LIIGa	44.7751 (9.5720)	1 -	4.2064 (1.1340)	0.3796 (0.4020)	179.5	179.9	186.0
Gamma	1 -	1 -	7.4880 (1.2755)	2.7135 (0.4781)	186.3	186.5	190.7

Table 3: LR tests

Ethylene	Hypotheses	LR Statistic w	p -value
EGN vs EN	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	385.9	< 0.00001
EGN vs LIIN	$H_0 : \beta = 1$ vs $H_1 : H_0$ is false	451.8	< 0.00001
EGN vs Normal	$H_0 : \alpha = \beta = 1$ vs $H_1 : H_0$ is false	456	< 0.00001
Wheaton River	Hypotheses	Statistic w	p -value
EGGu vs EGu	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	25.1	< 0.00001
EGGu vs LIIGu	$H_0 : \beta = 1$ vs $H_1 : H_0$ is false	4.7	0.0302
EGGu vs Gumbel	$H_0 : \alpha = \beta = 1$ vs $H_1 : H_0$ is false	26.1	< 0.00001
Stress level	Hypotheses	Statistic w	p -value
EGF vs EF	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	56.5	< 0.00001
EGF vs LIIF	$H_0 : \beta = 1$ vs $H_1 : H_0$ is false	12.4	0.0004
EGF vs Fréchet	$H_0 : \alpha = \beta = 1$ vs $H_1 : H_0$ is false	57.5	< 0.00001
Carbon	Hypotheses	Statistic w	p -value
EGGa vs EGa	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	5.6	0.0179
EGGa vs LIIGa	$H_0 : \beta = 1$ vs $H_1 : H_0$ is false	1.4	0.2367
EGGa vs Gamma	$H_0 : \alpha = \beta = 1$ vs $H_1 : H_0$ is false	10.2	0.0061

tests provide no indications against the LIIGa distribution. However, for the ethylene, Wheaton River and stress level data sets, we reject the null hypotheses in all three LR tests in favor of the new distributions. The rejection is extremely highly significant for the ethylene and stress level data, and highly or very highly significant for the Wheaton River data. This gives clear evidence of the potential need for the extra parameters in the proposed model when modelling real data.

In order to assess if the model is appropriate, the histograms of the data, the plots of the fitted EGN, EN, LIIN, EGGu, EGu, LIIGu, EGF, EF, LIIF, EGGa, EGa, LIIGa, normal, Gumbel, Fréchet and gamma density functions are displayed in Figure 5. We can conclude that the EG distributions are very suitable to these data.

10. Concluding Remarks

We propose a new class of exponentiated generalized (EG) distributions which includes as special cases the two classes of Lehmann's (1953) alternatives. The EG class extends several common distributions studied recently such as the exponentiated exponential, exponentiated Weibull, exponentiated gamma, exponentiated Fréchet and exponentiated Gumbel distributions (see, for example, Mudholkar and Srivastava, 1993; Gupta *et al.*, 1998; Gupta and Kundu, 2001; Nadarajah and Kotz, 2006), among several others. Indeed, for any baseline distribution, we

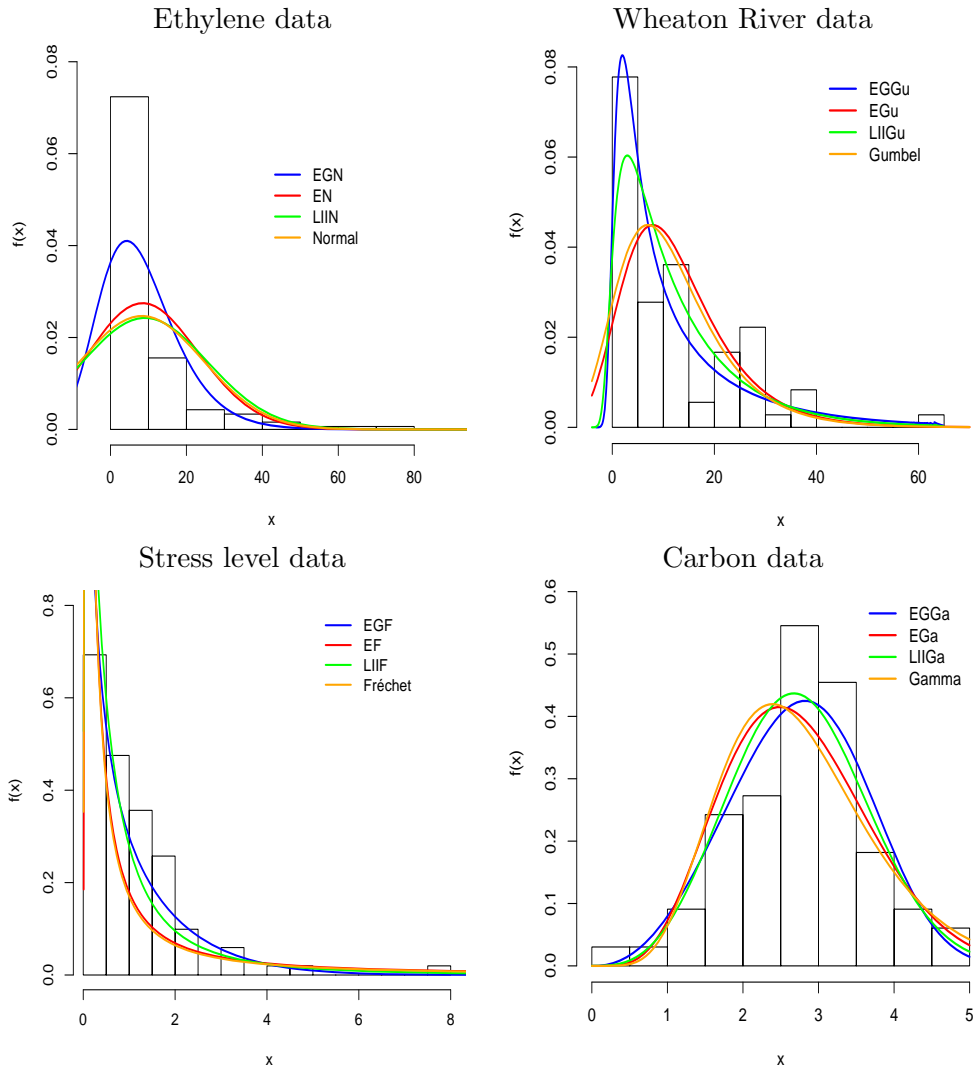


Figure 5: Estimated densities of the EG models for the analysed data sets

can easily define the corresponding EG distribution. The quantile, moments, generating function and mean deviations of some generated distributions have tractable mathematical properties. Some of these properties are readily obtained from those of the exponentiated and baseline distributions. For example, the moments of the EG distribution can be expressed explicitly in terms of an infinite sum of probability weighted moments of the baseline G distribution. The same happens for the moments of their order statistics.

We discuss maximum likelihood estimation and inference on the parameters based on likelihood ratio statistics for testing nested models. Four applications of the new class of distributions to real data are given to show the feasibility

of our proposal. We hope this generalization may attract wider applications in statistics. Some suggestions and directions for future research on the new class of models include simulation studies, asymptotic properties of the maximum likelihood estimates and performance and comparison of the Bayesian, bootstrap and Jackknife methods for estimation of the model parameters. Finally, we can also define regression models for the logarithmic of the random variable of the EG class of distributions.

Appendix

The elements of the observed information matrix $J(\boldsymbol{\theta})$ for the parameters (α, β, γ) are:

$$\begin{aligned}
J_{\alpha,\alpha} &= -\frac{n}{\alpha^2} + (\beta - 1) \sum_{i=1}^n \frac{[1 - G(x_i; \gamma)]^\alpha \{\log[1 - G(x_i; \gamma)]\}^2}{\{1 - [1 - G(x_i; \gamma)]^\alpha\}^2}, \\
J_{\alpha,\beta} &= -\sum_{i=1}^n \frac{[1 - G(x_i; \gamma)]^\alpha \log[1 - G(x_i; \gamma)]}{1 - [1 - G(x_i; \gamma)]}, \\
J_{\alpha,\gamma_j} &= -\sum_{i=1}^n \frac{[\dot{G}(x_i; \gamma)]_{\gamma_j}}{1 - G(x_i; \gamma)} + (\beta - 1) \sum_{i=1}^n \frac{[\dot{G}(x_i; \gamma)]_{\gamma_j} [1 - G(x_i; \gamma)]^{\alpha-1}}{\{1 - [1 - G(x_i; \gamma)]^\alpha\}^2} \\
&\quad \times \{[1 + \alpha \log[1 - G(x_i; \gamma)]]\{1 - [1 - G(x_i; \gamma)]\} + \alpha[1 - G(x_i; \gamma)]^\alpha\}, \\
J_{\beta,\beta} &= -\frac{n}{\beta^2}, \\
J_{\beta,\gamma_j} &= \alpha \sum_{i=1}^n \frac{[\dot{G}(x_i; \gamma)]_{\gamma_j} [1 - G(x_i; \gamma)]^{\alpha-1}}{\{1 - [1 - G(x_i; \gamma)]^\alpha\}^2}, \\
J_{\gamma_j,\gamma_s} &= \sum_{i=1}^n \frac{[\ddot{g}(x_i; \gamma)]_{\gamma_j\gamma_s} g(x_i; \gamma) - [\dot{g}(x_i; \gamma)]_{\gamma_j} [\dot{g}(x_i; \gamma)]_{\gamma_s}}{[g(x_i; \gamma)]^2} \\
&\quad + \sum_{i=1}^n \frac{[\ddot{G}(x_i; \gamma)]_{\gamma_j\gamma_s} [1 - G(x_i; \gamma)] + [\dot{G}(x_i; \gamma)]_{\gamma_j} [\dot{G}(x_i; \gamma)]_{\gamma_s}}{[1 - G(x_i; \gamma)]^2} \\
&\quad - \sum_{i=1}^n \frac{[1 - G(x_i; \gamma)]^{\alpha-2}}{1 - [1 - G(x_i; \gamma)]} \{(\alpha - 1)[\dot{G}(x_i; \gamma)]_{\gamma_j} [\dot{G}(x_i; \gamma)]_{\gamma_s} \\
&\quad - [1 - G(x_i; \gamma)][\ddot{G}(x_i; \gamma)]_{\gamma_j\gamma_s}\} \\
&\quad + \alpha \sum_{i=1}^n \frac{[\dot{G}(x_i; \gamma)]_{\gamma_j} [\dot{G}(x_i; \gamma)]_{\gamma_s} [1 - G(x_i; \gamma)]^{2(\alpha-1)}}{\{1 - [1 - G(x_i; \gamma)]^\alpha\}^2},
\end{aligned}$$

where $[\ddot{g}(x_i; \gamma)]_{\gamma_j\gamma_s} = \partial^2 g(x_i; \gamma) / \partial \gamma_j \partial \gamma_s$, $[\ddot{G}(x_i; \gamma)]_{\gamma_j\gamma_s} = \partial^2 G(x_i; \gamma) / \partial \gamma_j \partial \gamma_s$, and $j, s = 1, \dots, p$.

Acknowledgements

The financial support from CAPES and CNPq is gratefully acknowledged. We thank the associate editor and the reviewer for constructive comments.

References

- Aarts, R. M. (2000). Lauricella functions. From MathWorld, A Wolfram Web Resource, created by Eric W. Weisstein. (<http://mathworld.wolfram.com/LauricellaFunctions.html>).
- Akinsete, A., Famoye, F. and Lee, C. (2008). The beta-Pareto distribution. *Statistics* **42**, 547-563.
- Chaudhry, M. A. and Zubair, S. M. (2002). *On a Class of Incomplete Gamma Functions with Applications*. Chapman and Hall, Boca Raton, Florida.
- Cooray, K. and Ananda, M. M. A. (2008). A generalization of the half-normal distribution with applications to lifetime data. *Communications in Statistics - Theory and Methods* **37**, 1323-1337.
- Cordeiro, G. M. and Nadarajah, S. (2011). Closed-form expressions for moments of a class of beta generalized distributions. *Brazilian Journal of Probability and Statistics* **25**, 14-33.
- Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods* **31**, 497-512.
- Exton, H. (1978). *Handbook of Hypergeometric Integrals: Theory, Applications, Tables, Computer Programs*. Halsted Press, New York.
- Gupta, R. C., Gupta, P. L. and Gupta, R. D. (1998). Modeling failure time data by Lehmann alternatives. *Communications in Statistics - Theory and Methods* **27**, 887-904.
- Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions. *Australian and New Zealand Journal of Statistics* **41**, 173-188.
- Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and Weibull. *Biometrical Journal* **43**, 117-130.
- Lehmann, E. L. (1953). The power of rank tests. *Annals of Mathematical Statistics* **24**, 23-43.

- Mathai, A. M. (1993). *Hypergeometric Functions of Several Matrix Arguments: A Preliminary Report*. Centre for Mathematical Sciences, Trivandrum.
- Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability* **42**, 299-302.
- Mudholkar, G. S., Srivastava, D. K. and Kollia, G. D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of American Statistical Association* **91**, 1575-1583.
- Nadarajah, S. (2005). The exponentiated Gumbel distribution with climate application. *Environmetrics* **17**, 13-23.
- Nadarajah, S. and Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematicae* **92**, 97-111.
- Nadarajah, S. and Gupta, A. K. (2007). The exponentiated gamma distribution with application to drought data. *Calcutta Statistical Association Bulletin* **59**, 29-54.
- Nichols, M. D. and Padgett, W. J. (2006). A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International* **22**, 141-151.
- Prudnikov, A. P., Brychkov, Y. A. and Marichev, O. I. (1986). *Integrals and Series (vol. 1-3)*. Breach Science, Amsterdam, Netherlands.
- Saavedra del Aguila, J., Heiffig-del Aguila, L. S., Sasaki, F. F., Tsumanuma, G. M., Ongarelli, M. G., Spoto, M. H. F., Jacomino, A. P., Ortega, E. M. M. and Kluge, R. A. (2010). Postharvest modifications of mechanically injured bananas. *Revista Iberoamericana de Tecnologia Postcosecha* **10**, 73-85.
- Shirke, D. T. and Kakade, C. S. (2006). On exponentiated lognormal distribution. *International Journal of Agricultural and Statistical Sciences* **2**, 319-326.

Received February 24, 2012; accepted June 22, 2012.

Gauss M. Cordeiro
Departamento de Estatística
Universidade Federal de Pernambuco
52171-900, Recife, PE, Brazil
gausscordeiro@uol.com.br

Edwin M. M. Ortega
Departamento de Ciências Exatas
Universidade de São Paulo
13418-900, Piracicaba, SP, Brazil
edwin@usp.br

Daniel C. C. da Cunha
Departamento de Matemática,
Universidade Federal de Pernambuco
52171-900, Recife, PE, Brazil
daniel.csmr@gmail.com