STATISTICA, anno LXXVIII, n. 2, 2018

A NEW FAMILY OF DISTRIBUTIONS BASED ON THE HYPOEXPONENTIAL DISTRIBUTION WITH FITTING RELIABILITY DATA

Christophe Chesneau [1]
Université de Caen Basse-Normandie, LMNO, Département de Mathématiques, UFR de Sciences, F-14032, Caen, France

1. Introduction

Over the past decades, considerable efforts have been made to develop flexible distributions for modeling lifetime data. In many situations, the weighted exponential distributions reach this goal with a great success. They are characterized by probability density functions of the form: $f(x) = w(x)g(x)$, where $g(x)$ denotes a pdf of the exponential distribution with parameter $\lambda > 0$ and $w(x)$ denotes a weight function. A suitable choice for $w(x)$, based on practical or theoretical consideration, can lead a proper statistical model for analysis specific lifetime data. The literature covering weighted exponential distributions is vast and is growing fast. We refer to the review of Saghir *et al.* (2017), and the references therein.

In this paper, we introduce a new general family of distributions constructed from the so-called hypoexponential distribution and a transformation proposed by Al-Hussaini (2012). Let us recall that the hypoexponential distribution is characterized by the distribution of a sum of several independent exponential random variables with different parameters (see Amari and Misra, 1997; Akkouchi, 2008). It is used in many domains of application, including tele-traffic engineering and queuing systems. Several tuning parameters are thus necessary for a perfect model. Our new family of distributions profits of this flexibility to open the door of new applications.

By considering a specific configuration of this family, we exhibit a new weighted exponential distribution, called the EH($\mathbf{a}, \lambda$) distribution, having some interesting properties. The corresponding cumulative distribution function can also be viewed a special linear combination of several cumulative distribution functions associated to the well known exponentiated exponential distribution introduced by Gupta and Kundu (1999, 2001). "Special" in the sense that the coefficients of the combination depends only on

[1] Corresponding Author. E-mail: christophe.chesneau@unicaen.fr

the parameters of the exponentiated exponential distributions and can be negative. For this last reason, it also belongs to the family of finite generalized mixture distributions (see Bartholomew, 1969; Franco *et al.*, 2013; Bakouch and Popović, 2016). Mathematical properties of our distribution are studied, including moments, conditional moments, reversed conditional moments, moment generating function, Lorenz curve, Bonferroni curve, mean deviations and order statistics. Estimation of the parameters is determined using the method of maximum likelihood. Considering two real data sets of different natures, we show the superiority of our distribution in terms of some goodness-of-fit statistics ($-2\log(L)$, AIC, AICC, BIC, A^* and W^*) in comparison to the well-known distributions: the Lindley distribution (see Ghitany *et al.*, 2008), the exponential distribution and the exponentiated exponential distribution (see Gupta and Kundu, 1999, 2001).

The rest of the paper is organized as follows: Section 2 presents our new distribution. Some of its mathematical properties are described in Section 3. Distributions of order statistics following our distribution are investigated in Section 4. Estimation of the parameters via the maximum likelihood is studied in Section 5. Applications to real data sets are performed in Section 6. Some concluding remarks are outlined in Section 7.

2. THE EH($\mathbf{a},\lambda$) DISTRIBUTION

2.1. *On the hypoexponential distribution*

We now briefly present the so-called hypoexponential distribution. Let $m \geq 2$ be a positive integer and $\alpha_1, \alpha_2, \ldots, \alpha_m$ be m different real numbers. Let $X_1, X_2, \ldots, X_m$ be m independent random variables, with X_i following the exponential distribution of parameter $\alpha_i > 0$; the probability distribution function (pdf) of X_i is given by

$$f_{X_i}(x) = \alpha_i e^{-\alpha_i x}, \qquad x > 0.$$

Then the random variable $S = \sum_{i=1}^{m} X_i$ follows the hypoexponential distribution with a parameter vector $\mathbf{a} = (\alpha_1, \ldots, \alpha_m)$; a pdf of S is given by

$$f_S(x) = \sum_{i=1}^{m} \frac{\alpha_i}{P_i} e^{-\alpha_i x}, \qquad x > 0,$$

where

$$P_i = \prod_{\substack{j=1 \\ j \neq i}}^{m} \left(1 - \frac{\alpha_i}{\alpha_j}\right).$$

Note that P_i is completely determined by $\mathbf{a}$ and can be positive or negative according to the sign of $\prod_{\substack{j=1 \\ j \neq i}}^{m} \left(\alpha_j - \alpha_i\right)$, offering a great richness in the construction of $f_S(x)$. The corresponding cumulative distribution function (cdf) is given by

$$F_S(x) = 1 - \sum_{i=1}^{m} \frac{1}{P_i} e^{-\alpha_i x}. \tag{1}$$

Further details on the hypoexponential distribution can be found in Amari and Misra (1997), Akkouchi (2008) and Smaili *et al.* (2013). Before introducing the EH($\mathbf{a}$, λ) distribution, we present a more general family of distributions based on a general transformation of cdf and the hypoexponential distribution.

2.2. *A family of distributions based on the hypoexponential distribution*

Lemma 1 below presents an idea of construction of a new cdf using two existing cdfs.

LEMMA 1 (AL-HUSSAINI (2012)). *Let $G(x)$ and $H(x)$ be two cdfs of continuous distributions with $G(x)$ of support $(0, +\infty)$. Let $K(x)$ be the function defined by*

$$K(x) = 1 - G(-\ln(H(x))), \qquad x \in \mathbb{R}.$$

Then $K(x)$ is a cdf of a continuous distribution.

This result is useful for the construction of new families of distributions. Significant examples can be found in, e.g., Zografos and Balakrishnan (2009), Ristić and Balakrishnan (2012) and Pal and Tiensuwan (2015).

Let us now introduce a particular family of distributions which will be at the heart of this study. Let $H(x)$ be a cdf of a continuous distribution and $G(x)$ be the cdf associated to the hypoexponential distribution with a parameter vector $\mathbf{a} = (\alpha_1, \ldots, \alpha_m)$ (1). It follows from Lemma 1 that the following function $K(x)$ is a cdf

$$K(x) = 1 - G(-\ln(H(x))) = 1 - \left(1 - \sum_{i=1}^{m} \frac{1}{P_i} e^{-\alpha_i(-\ln(H(x)))}\right) = \sum_{i=1}^{m} \frac{1}{P_i} [H(x)]^{\alpha_i}.$$

From this expression, new distributions arise by taking specific cdf $H(x)$ (uniform distribution, normal distribution, Cauchy distribution, Pareto distribution...). We will call the associated distribution the GH(H; $\mathbf{a}$) distribution (for general hypoexponential distribution with a parameter vector $\mathbf{a} = (\alpha_1, \ldots, \alpha_m)$ using a cdf $H(x)$) for frequently used purpose in the present study or elsewhere.

Let us now consider a random variable Y following the GH(H; $\mathbf{a}$) distribution. Then the cdf of Y is given by

$$F_Y(x) = \sum_{i=1}^{m} \frac{1}{P_i} [H(x)]^{\alpha_i}, \qquad x \in \mathbb{R}. \tag{2}$$

Using the equality $\sum_{i=1}^{m} \frac{1}{P_i} = 1$, the survival function (sf) of Y is given by

$$S_Y(x) = 1 - F_Y(x) = 1 - \sum_{i=1}^{m} \frac{1}{P_i}[H(x)]^{\alpha_i} = \sum_{i=1}^{m} \frac{1}{P_i}[1 - [H(x)]^{\alpha_i}].$$

Denoting by $h(x)$ a pdf associated to $H(x)$, a pdf of Y is given by

$$f_Y(x) = h(x) \sum_{i=1}^{m} \frac{\alpha_i}{P_i}[H(x)]^{\alpha_i - 1}.$$

The hazard rate function (hrf) of Y is given by

$$h_Y(x) = \frac{f_Y(x)}{S_Y(x)} = h(x) \frac{\sum_{i=1}^{m} \frac{\alpha_i}{P_i}[H(x)]^{\alpha_i - 1}}{\sum_{i=1}^{m} \frac{1}{P_i}[1 - [H(x)]^{\alpha_i}]}.$$

REMARK 2. *The GH(H;* $\mathbf{a}$*) distribution belongs to the family of weighted distributions; we can express $f_Y(x)$ as: $f_Y(x) = w(x)g(x)$, where $g(x)$ denotes a pdf of the exponential distribution with parameter $\lambda > 0$ and $w(x) = \sum_{i=1}^{m} \frac{\alpha_i}{P_i}[H(x)]^{\alpha_i - 1}$.*

REMARK 3. *Since the coefficients $\frac{1}{P_1}, \frac{1}{P_2}, \ldots, \frac{1}{P_m}$ have possible negative and positive values with $\sum_{i=1}^{m} \frac{1}{P_i} = 1$, the GH($H$;* $\mathbf{a}$*) distribution also belongs to the family of finite generalized mixture of exponentiated baseline cdf $[H(x)]^{\alpha}$ (also known as Lehmann type-I distribution) (see Bartholomew, 1969; Franco* et al.*, 2013; Bakouch and Popović, 2016).*

REMARK 4. *In the particular case $m = 2$, $\mathbf{a} = (\alpha_1, \alpha_2)$, $\alpha_1, \alpha_2 > 0$, $\alpha_1 \neq \alpha_2$, we have $P_1 = \frac{\alpha_2 - \alpha_1}{\alpha_2}$ and $P_2 = \frac{\alpha_1 - \alpha_2}{\alpha_1}$,*

$$F_Y(x) = \frac{\alpha_2}{\alpha_2 - \alpha_1}[H(x)]^{\alpha_1} + \frac{\alpha_1}{\alpha_1 - \alpha_2}[H(x)]^{\alpha_2}$$

and

$$f_Y(x) = \frac{\alpha_1 \alpha_2}{\alpha_2 - \alpha_1} h(x) \left([H(x)]^{\alpha_1 - 1} - [H(x)]^{\alpha_2 - 1}\right).$$

Assuming that $\alpha_2 > \alpha_1$, note that $P_1 > 0$ and $P_2 < 0$.

The distribution of a sum of two independent GH(H;$\mathbf{a}$) distributions can be characterized. Let Y and Z be two independent random variables, Y follows the GH(H_1;$\mathbf{a}$) distribution with $\mathbf{a} = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ and Z follows the GH(H_2;$\mathbf{b}$) distribution with

$\mathbf{b} = (\beta_1, \beta_2, \ldots, \beta_m)$. Let $h_1(x)$ be a pdf associated to the cdf $H_1(x)$ and $h_2(x)$ be a pdf associated to the cdf $H_2(x)$. Set $Q_\ell = \prod_{\substack{j=1 \\ j\neq\ell}}^{m} \left(1 - \frac{\beta_\ell}{\beta_j}\right)$. Then a pdf for $Y + Z$ is given by

$$f_{Y+Z}(x) = \sum_{i=1}^{m}\sum_{\ell=1}^{m} \frac{\alpha_i \beta_\ell}{P_i Q_\ell} \int_{-\infty}^{+\infty} h_1(t) h_2(x-t)[H_1(t)]^{\alpha_i - 1}[H_2(x-t)]^{\beta_\ell - 1} dt.$$

This result is an immediate consequence of the continuous convolution formula.

The following result shows an ordering property between the GH($H;\mathbf{a}$) distribution and the exponentiated baseline cdf $[H(x)]^\alpha$.

PROPOSITION 5. *Let $F_Y(x)$ be (2). Then, for any $x \in \mathbb{R}$, we have*

$$F_Y(x) \geq 1 - \prod_{i=1}^{m}(1 - [H(x)]^{\alpha_i}).$$

Observe that the lower bound does not depend on $P_1, \ldots, P_m$ and it is a cdf.

PROOF. Let $X_1, X_2, \ldots, X_m$ be m independent random variables, with X_i following the exponential distribution of parameter $\alpha_i > 0$, $S = \sum_{i=1}^{m} X_i$ and $M = \sup(X_1, X_2, \ldots, X_m)$. Then, since the support of the random variables is positive, we have $M \leq S$ almost surely, implying that the corresponding cdfs of S and M, denoted by $F_S(x)$ and $F_M(x)$ respectively, satisfies $F_S(x) \leq F_M(x)$, with $F_M(x) = \prod_{i=1}^{m} \mathbb{P}(X_i \leq x) = \prod_{i=1}^{m}(1 - e^{-\alpha_i x})$. Therefore

$$\begin{aligned} F_Y(x) &= 1 - F_S(-\ln(H(x))) \geq 1 - F_M(-\ln(H(x))) \\ &= 1 - \prod_{i=1}^{m}\left(1 - e^{-\alpha_i(-\ln(H(x)))}\right) = 1 - \prod_{i=1}^{m}(1 - [H(x)]^{\alpha_i}). \end{aligned}$$

This ends the proof of Proposition 5. □

As an immediate consequence of Proposition 5, for any $\alpha_* \in \{\alpha_1, \ldots, \alpha_m\}$, we have $F_Y(x) \geq [H(x)]^{\alpha_*}$, showing that $F_Y(x)$ is always greater to any of its power cdfs components. This ordering result is of interest for the construction of statistical models where, after analysis, distributions with cdf of the form $[H(x)]^\alpha$ seems globally under the corresponding empirical cdf constructed from the data; the GH($H;\mathbf{a}$) distribution can be an alternative in this case.

2.3. *The EH($\mathbf{a}$, λ) distribution*

Let us now consider the GH($H;\mathbf{a}$) distribution defined with the cdf $H(x)$ associated to the exponential distribution of parameter $\lambda > 0$: $H(x) = 1 - e^{-\lambda x}$, $x > 0$. We will call the associated distribution the EH($\mathbf{a}, \lambda$) distribution (for exponential hypoexponential distribution with a parameter vector $\mathbf{a} = (\alpha_1, \ldots, \alpha_m)$ and λ) for frequently used purpose in the present study or elsewhere. Let us now consider a random variable Y following the EH($\mathbf{a}, \lambda$) distribution. Then the cdf of Y given by

$$F_Y(x) = \sum_{i=1}^{m} \frac{1}{P_i}(1 - e^{-\lambda x})^{\alpha_i}, \qquad x > 0.$$

This cdf can be viewed as a finite generalized mixture of exponentiated exponential distribution baseline cdf introduced by Gupta and Kundu (1999, 2001), with different parameters for the power and special weights depending only on the parameters (the standard cdf of the exponentiated exponential distribution is obtained by taking $m = 1$). Further details on finite mixtures can be found in Bartholomew (1969). For recent developments and applications of the exponentiated exponential distribution, we refer to Chaturvedi and Vyas (2017) and Bhattacharjee (2018). Thanks to its several tuning parameters, the EH($\mathbf{a}, \lambda$) distribution has a great flexibility which can be useful for determine a nice statistical model.

Moreover, the consideration of the EH($\mathbf{a}, \lambda$) distribution in comparison to other possible GH($H; \mathbf{a}$) distributions is motivated below by the consideration of two well-known distributions : the Weibull and the gamma distributions. Let $H_W(x; \alpha, \lambda)$ be the cdf of the Weibull distribution, with pdf : $f_W(x; \alpha, \lambda) = \alpha\lambda(\lambda x)^{\alpha-1} e^{-(\lambda x)^\alpha}$, $H_G(x; \alpha, \lambda)$ be the cdf of the gamma distribution, with pdf : $f_G(x; \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$ and $H_E(x; \alpha, \lambda)$ be the cdf of the exponentiated exponential distribution, with pdf : $f_E(x; \alpha, \lambda) = \alpha\lambda(1 - e^{-\lambda x})^{\alpha-1} e^{-\lambda x}$, $\alpha, \lambda, x > 0$. Ordering of lifetime distributions plays a crucial role in statistical literature, showing the most appropriated statistical model in some situation. Using ordering results among others, Gupta and Kundu (2001) proved that the exponentiated exponential family can be viewed as an alternative to gamma and Weibull distributions ; it may have a better fit compared to them. The result below shows that we can transpose this conclusion to the EH($\mathbf{a}, \lambda$) distribution in comparison to the GH($H_W; \mathbf{a}$) and GH($H_G; \mathbf{a}$) distributions.

PROPOSITION 6. *Let $F_E(x; \mathbf{a}, \lambda) = F_Y(x)$ be the cdf of the EH($\mathbf{a}, \lambda$) distribution, $F_W(x; \mathbf{a}, \lambda)$ be the cdf of the GH($H_W; \mathbf{a}$) distribution and $F_G(x; \mathbf{a}, \lambda)$ be the cdf of the GH($H_G; \mathbf{a}$) distribution. Then if, for any $i \in \{1, \ldots, m\}$, $\alpha_i - 1$ and P_i has the same sign, then we have*

$$F_W(x; \mathbf{a}, \lambda) \geq F_E(x; \mathbf{a}, \lambda) \geq F_G(x; \mathbf{a}, \lambda).$$

On the other hand, if, for any $i \in \{1, \ldots, m\}$, $\alpha_i - 1$ and P_i has the opposite sign, then we have

$$F_W(x; \mathbf{a}, \lambda) \leq F_E(x; \mathbf{a}, \lambda) \leq F_G(x; \mathbf{a}, \lambda).$$

PROOF. The result in Section 3 of Gupta and Kundu (2001) shows the following orderings:

$H_W(x;\alpha,\lambda) \geq H_E(x;\alpha,\lambda) \geq H_G(x;\alpha,\lambda)$ for $\alpha > 1$ and $H_W(x;\alpha,\lambda) \leq H_E(x;\alpha,\lambda) \leq H_G(x;\alpha,\lambda)$ for $\alpha < 1$. Indeed, since $F_E(x;\mathbf{a},\lambda) = \sum_{i=1}^m \frac{1}{P_i} H_E(x;\alpha_i,\lambda)$, $F_W(x;\mathbf{a},\lambda) = \sum_{i=1}^m \frac{1}{P_i} H_W(x;\alpha_i,\lambda)$ and $F_G(x;\mathbf{a},\lambda) = \sum_{i=1}^m \frac{1}{P_i} H_G(x;\alpha_i,\lambda)$, these orderings are respected with suitable assumption on α_i and P_i, which exactly corresponds to $\alpha_i - 1$ and P_i has the same sign for the first ordering, and $\alpha_i - 1$ and P_i has the opposite sign for the second ordering. This ends the proof of Proposition 6. □

Some elementary functions related to $F_Y(x)$ are given below. The sf of Y is given by

$$S_Y(x) = \sum_{i=1}^m \frac{1}{P_i}[1-(1-e^{-\lambda x})^{\alpha_i}].$$

A pdf of Y is given by

$$f_Y(x) = \lambda e^{-\lambda x} \sum_{i=1}^m \frac{\alpha_i}{P_i}(1-e^{-\lambda x})^{\alpha_i - 1}.$$

The hrf of Y is given by

$$h_Y(x) = \lambda e^{-\lambda x} \frac{\sum\limits_{i=1}^m \frac{\alpha_i}{P_i}(1-e^{-\lambda x})^{\alpha_i-1}}{\sum\limits_{i=1}^m \frac{1}{P_i}[1-(1-e^{-\lambda x})^{\alpha_i}]}.$$

Some asymptotic properties of the previous functions are given below: when $x \to 0$, we have

$$F_Y(x) \sim \sum_{i=1}^m \frac{1}{P_i}(\lambda x)^{\alpha_i}, \quad f_Y(x) \sim \lambda \sum_{i=1}^m \frac{\alpha_i}{P_i}(\lambda x)^{\alpha_i-1}, \quad h_Y(x) \sim \lambda \sum_{i=1}^m \frac{\alpha_i}{P_i}(\lambda x)^{\alpha_i-1}.$$

Note that, if $\inf(\alpha_1,\alpha_2,\ldots,\alpha_m) > 1$, we have $\lim\limits_{x\to 0} f_Y(x) = 0$.

If $\inf(\alpha_1,\alpha_2,\ldots,\alpha_m) < 1$, the limit depends on the signs of $P_1, P_2, \ldots, P_m$; we can have $\lim\limits_{x\to 0} f(x) = +\infty$ in some configurations. We have $\lim\limits_{x\to+\infty} f(x) = 0$. The limit for $h(x)$ when $x \to +\infty$ depends only on λ and $P_1, P_2, \ldots, P_m$.

REMARK 7. *As already observed in Remark 2, the EH($\mathbf{a}$, λ) distribution belongs to the family of weighted exponential distributions; we can express $f_Y(x)$ as: $f_Y(x) = w(x)g(x)$, where $g(x)$ denotes a pdf of the exponential distribution with parameter $\lambda > 0$ and* $w(x) = \sum\limits_{i=1}^m \frac{\alpha_i}{P_i}(1-e^{-\lambda x})^{\alpha_i-1}$.

REMARK 8. *Let us observe that, for a random variable U following the hypoexponential distribution with parameter a, the random variable $W = -\frac{1}{\lambda}\ln(1-e^{-U})$ follows the EH($\mathbf{a}$, λ) distribution.*

Remark 9. *In the particular case* $m = 2$, $\mathbf{a} = (\alpha_1, \alpha_2)$, $\alpha_1, \alpha_2 > 0$, $\alpha_1 \neq \alpha_2$, *we have* $P_1 = \frac{\alpha_2 - \alpha_1}{\alpha_2}$ *and* $P_2 = \frac{\alpha_1 - \alpha_2}{\alpha_1}$,

$$F_Y(x) = \frac{\alpha_2}{\alpha_2 - \alpha_1}[1 - e^{-\lambda x}]^{\alpha_1} + \frac{\alpha_1}{\alpha_1 - \alpha_2}[1 - e^{-\lambda x}]^{\alpha_2}$$

and

$$f_Y(x) = \frac{\lambda \alpha_1 \alpha_2}{\alpha_2 - \alpha_1} e^{-\lambda x}\left([1 - e^{-\lambda x}]^{\alpha_1 - 1} - [1 - e^{-\lambda x}]^{\alpha_2 - 1}\right).$$

Figures 1, 2 and 3 show the graphical features of pdfs, cdfs and hrfs of the EH($\mathbf{a}, \lambda$) distribution with $\mathbf{a} = (\alpha_1, \alpha_2)$, $\alpha_1 = \alpha$ and $\alpha_2 = \alpha + 0.1$, for several choices of parameters (α, λ). In this particular case, let us precise that

$$F_Y(x) = (1 + 10\alpha)(1 - e^{-\lambda x})^{\alpha} - 10\alpha(1 - e^{-\lambda x})^{\alpha + 0.1}, \tag{3}$$

a pdf is given by

$$f_Y(x) = \lambda\alpha(1 + 10\alpha)e^{-\lambda x}(1 - e^{-\lambda x})^{\alpha - 1}[1 - (1 - e^{-\lambda x})^{0.1}] \tag{4}$$

and the hrf is given by

$$h_Y(x) = \frac{\lambda\alpha(1 + 10\alpha)e^{-\lambda x}(1 - e^{-\lambda x})^{\alpha - 1}[1 - (1 - e^{-\lambda x})^{0.1}]}{1 - (1 + 10\alpha)(1 - e^{-\lambda x})^{\alpha} + 10\alpha(1 - e^{-\lambda x})^{\alpha + 0.1}}. \tag{5}$$

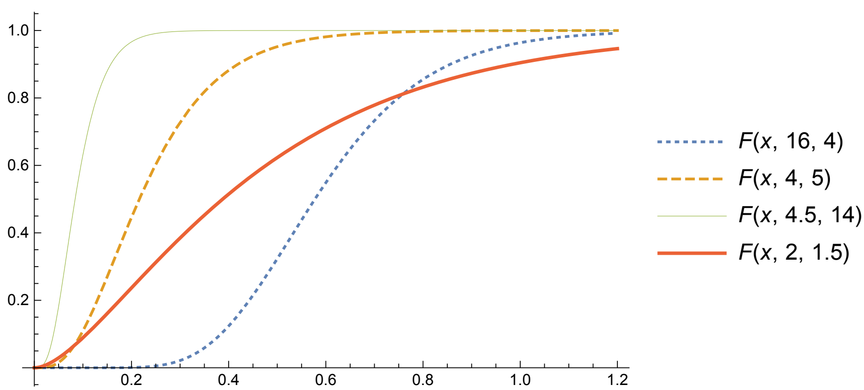

Figure 1 – Some cdfs $F(x) = F(x; (\alpha, \lambda))$ (3) with various values for α and λ.

We observe that the plots of the pdfs and hrfs of the EH($\mathbf{a}, \lambda$) distribution are very flexible; different shapes and curves can be of interest for modeling lifetime data.

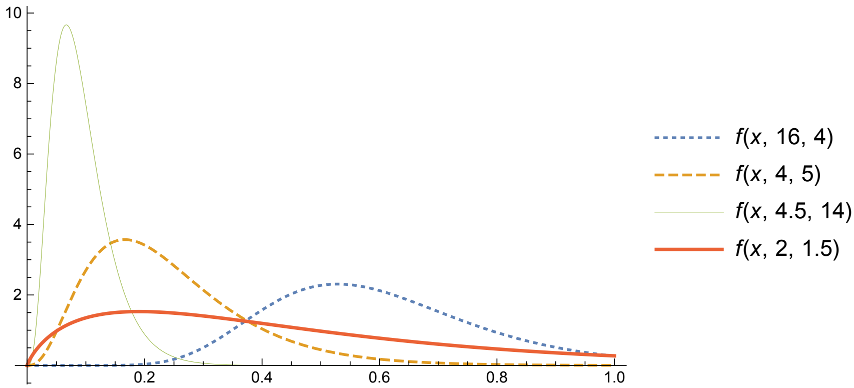

Figure 2 – Some pdfs $f(x) = f(x;(\alpha,\lambda))$ (4) with various values for α and λ.

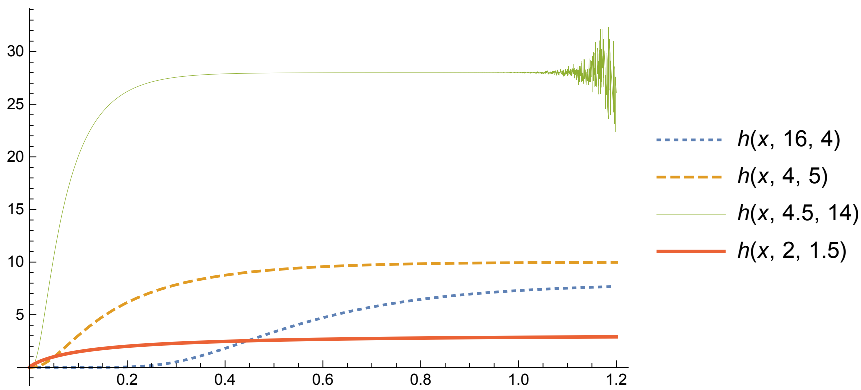

Figure 3 – Some hrfs $h(x) = h(x;(\alpha,\lambda))$ (5) with various values for α and λ.

REMARK 10. *The considered pdf $f_Y(x)$ (4) is a very special case of the EH($\mathbf{a},\lambda$) distribution and the value* 0.1 *is subjective. It can also be viewed as a particular two parameters pdf of the McDonald modified Weibull distribution introduced in Merovci and Elbatal (2013) or the beta exponentiated Weibull distribution developed by Cordeiro* et al. *(2013). Note that the main differences between $f_Y(x)$ (4) and similar looking pdfs as those associated to the weighted exponentiated exponential distribution introduced by Mahdavi (2015) or the transmuted exponentiated exponential distribution proposed by Khan* et al. *(2017), is the last term $[1-(1-e^{-\lambda x})^{0.1}]$ which only depends on λ, not α.*

3. Mathematical properties related to the EH($\mathbf{a},\lambda$) distribution

We now present some mathematical properties of our distribution.

The distribution of a sum of two independent EH($\mathbf{a},\lambda$) distributions can be characterized. Let Y and Z be two independent random variables, Y follows the EH($\mathbf{a},\lambda_1$) distribution with $\mathbf{a}=(\alpha_1,\alpha_2,\ldots,\alpha_m)$ and Z follows the EH($\mathbf{b},\lambda_2$) distribution with $\mathbf{b}=(\beta_1,\beta_2,\ldots,\beta_m)$. Set $Q_\ell=\prod_{\substack{j=1\\ j\neq \ell}}^{m}\left(1-\frac{\beta_\ell}{\beta_j}\right)$.

Then a pdf for $Y+Z$ is given by, for $x>0$,

$$f_{Y+Z}(x)=\lambda_1\lambda_2\sum_{i=1}^{m}\sum_{\ell=1}^{m}\frac{\alpha_i\beta_\ell}{P_iQ_\ell}\int_0^x e^{-\lambda_1 t}e^{-\lambda_1(x-t)}[1-e^{-\lambda_1 t}]^{\alpha_i-1}[1-e^{-\lambda_2(x-t)}]^{\beta_\ell-1}dt.$$

In the case $\lambda=\lambda_1=\lambda_2$, we can express $\mathbb{P}(Y\leq Z)$, which plays a central role in the context of reliability the stress-strength model : using the inequality $\int_0^{+\infty}\beta\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha+\beta-1}dx=\frac{\beta}{\alpha+\beta}$ (see Equation (2.1) Kundu and Gupta, 2005), we have

$$\begin{aligned}\mathbb{P}(Y\leq Z) &= \int_0^{+\infty}P(Y\leq z)\lambda e^{-\lambda z}\sum_{\ell=1}^{m}\frac{\beta_\ell}{Q_\ell}[1-e^{-\lambda z}]^{\beta_\ell-1}dz\\ &= \sum_{i=1}^{m}\sum_{\ell=1}^{m}\frac{1}{P_iQ_\ell}\int_0^{+\infty}\beta_\ell\lambda e^{-\lambda z}[1-e^{-\lambda z}]^{\alpha_i+\beta_\ell-1}dz\\ &= \sum_{i=1}^{m}\sum_{\ell=1}^{m}\frac{\beta_\ell}{P_iQ_\ell(\alpha_i+\beta_\ell)}.\end{aligned}$$

Let us work with only a random variable Y following the EH($\mathbf{a},\lambda$) distribution. Owing to the binomial series, we have the following expansion for $F_Y(x)$

$$F_Y(x)=\sum_{i=1}^{m}\frac{1}{P_i}(1-e^{-\lambda x})^{\alpha_i}=\sum_{i=1}^{m}\sum_{k=0}^{+\infty}\binom{\alpha_i}{k}(-1)^k\frac{1}{P_i}e^{-k\lambda x}.$$

Using again the binomial series, we have the following expansion for $f_Y(x)$

$$f_Y(x)=\lambda e^{-\lambda x}\sum_{i=1}^{m}\frac{\alpha_i}{P_i}(1-e^{-\lambda x})^{\alpha_i-1}=\sum_{i=1}^{m}\sum_{k=0}^{+\infty}\eta_{i,k}e^{-(k+1)\lambda x},$$

where

$$\eta_{i,k}=\lambda\binom{\alpha_i-1}{k}(-1)^k\frac{\alpha_i}{P_i}.$$

We can obtain the quantile function $Q_Y(x)$ via the nonlinear equation

$$F_Y(Q_Y(x))=x \quad\Leftrightarrow\quad \sum_{i=1}^{m}\frac{1}{P_i}(1-e^{-\lambda Q_Y(x)})^{\alpha_i}=x.$$

The EH($\mathbf{a}, \lambda$) distribution can be simulated by using $Y = Q_Y(U)$ where U is a random variable having the uniform distribution on $[0,1]$.

Let us consider the gamma function: $\Gamma(\nu) = \int_0^{+\infty} x^{\nu-1} e^{-x} dx$, $\nu > 0$. The r-th moment of Y is given by

$$\begin{aligned} E(Y^r) &= \int_{-\infty}^{+\infty} x^r f_Y(x) dx = \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \int_0^{+\infty} x^r e^{-(k+1)\lambda x} dx \\ &= \frac{\Gamma(r+1)}{\lambda^{r+1}} \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \frac{1}{(k+1)^{r+1}}. \end{aligned}$$

The moment generating function of Y is given by, for $t < \lambda$,

$$\begin{aligned} M_Y(t) &= E(e^{tY}) = \int_{-\infty}^{+\infty} e^{tx} f_Y(x) dx = \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \int_0^{+\infty} e^{(t-(k+1)\lambda)x} dx \\ &= \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \frac{1}{(k+1)\lambda - t}. \end{aligned}$$

Proceeding as in Equation (2.3) of Gupta and Kundu (2001), with the change of variable $y = e^{-\lambda x}$, we also have, for $t < \lambda$,

$$M_Y(t) = \sum_{i=1}^{m} \frac{1}{P_i} \frac{\Gamma(\alpha_i + 1)\Gamma\left(1 - \frac{t}{\lambda}\right)}{\Gamma\left(\alpha_i - \frac{t}{\lambda} + 1\right)}.$$

Let us consider the lower incomplete gamma function: $\Gamma(t,\nu) = \int_0^t x^{\nu-1} e^{-x} dx$. Then we have

$$\begin{aligned} \int_0^t x^r f_Y(x) dx &= \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \int_0^t x^r e^{-(k+1)\lambda x} dx \\ &= \frac{1}{\lambda^{r+1}} \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \frac{\Gamma((k+1)\lambda t, r+1)}{(k+1)^{r+1}}. \end{aligned}$$

This equality will be useful in the next. For $t > 0$, the conditional r-th moment of Y is given by

$$\begin{aligned} E(Y^r \mid Y > t) &= \frac{1}{1 - F_Y(t)} \int_t^{+\infty} x^r f_Y(x) dx \\ &= \frac{1}{1 - F_Y(t)} \left(E(Y^r) - \int_0^t x^r f_Y(x) dx \right) \\ &= \frac{1}{\lambda^{r+1}(1 - F_Y(t))} \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \frac{\Gamma(r+1) - \Gamma((k+1)\lambda t, r+1)}{(k+1)^{r+1}}. \end{aligned}$$

For $t > 0$, the reversed conditional r-th moment of Y is given by

$$\begin{aligned} E(Y^r \mid Y \leq t) &= \frac{1}{F_Y(t)} \int_0^t x^r f_Y(x) dx \\ &= \frac{1}{F_Y(t)\lambda^{r+1}} \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \frac{\Gamma((k+1)\lambda t, r+1)}{(k+1)^{r+1}}. \end{aligned}$$

Let $\mu = E(Y)$. The Lorenz curve $L(F_Y(t))$ is given by

$$L(F_Y(t)) = \frac{1}{\mu} \int_0^t x f_Y(x) dx = \frac{1}{\mu \lambda^{r+1}} \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \frac{\Gamma((k+1)\lambda t, 2)}{(k+1)^{r+1}}.$$

The Bonferroni curve $B(F_Y(t))$ is given by

$$B(F_Y(t)) = \frac{1}{\mu F_Y(t)} \int_0^t x f_Y(x) dx = \frac{1}{\mu F_Y(x) \lambda^{r+1}} \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \frac{\Gamma((k+1)\lambda t, 2)}{(k+1)^{r+1}}.$$

The mean deviation of Y about the mean μ can be expressed as

$$\begin{aligned} \delta_1(Y) &= E(|Y-\mu|) = 2\mu F_Y(\mu) - 2\int_0^{\mu} x f_Y(x) dx \\ &= 2\mu F_Y(\mu) - \frac{2}{\lambda^{r+1}} \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \frac{\Gamma((k+1)\lambda\mu, 2)}{(k+1)^{r+1}}. \end{aligned}$$

The mean deviation of Y about the median M can be expressed as

$$\begin{aligned} \delta_2(Y) &= E(|Y-M|) = \mu - 2\int_0^{M} x f_Y(x) dx \\ &= \mu - \frac{2}{\lambda^{r+1}} \sum_{i=1}^{m} \sum_{k=0}^{+\infty} \eta_{i,k} \frac{\Gamma((k+1)\lambda M, 2)}{(k+1)^{r+1}}. \end{aligned}$$

4. ORDER STATISTICS

Order statistics are essential in many areas of statistics. They naturally appear in the probabilistic analysis of reliability of a system. We explicit here the distributions of these order statistics in the context of our new distribution. Let $Y_1, Y_2, \ldots, Y_n$ be n independent and identically distributed random variables following the EH($\mathbf{a}, \lambda$) distribution. Let us consider its order statistics : $Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}$. A pdf of the i-th order statistic $Y_{i:n}$ is given by, for $x > 0$,

$$\begin{aligned} f_{i:n}(x) &= \frac{n!}{(i-1)!(n-i)!}[F_Y(x)]^{i-1}[1-F_Y(x)]^{n-i} f_Y(x) \\ &= \frac{n!}{(i-1)!(n-i)!}\left[\sum_{i=1}^{m}\frac{1}{P_i}(1-e^{-\lambda x})^{\alpha_i}\right]^{i-1} \times \\ &\left[1-\sum_{i=1}^{m}\frac{1}{P_i}(1-e^{-\lambda x})^{\alpha_i}\right]^{n-i} \lambda e^{-\lambda x}\sum_{i=1}^{m}\frac{\alpha_i}{P_i}(1-e^{-\lambda x})^{\alpha_i-1}. \end{aligned}$$

In particular, a pdf of $Y_{1:n} = \inf(Y_1, Y_2, \ldots, Y_n)$ is given by

$$f_{1:n}(x) = \lambda n e^{-\lambda x}\left[1-\sum_{i=1}^{m}\frac{1}{P_i}(1-e^{-\lambda x})^{\alpha_i}\right]^{n-1}\sum_{i=1}^{m}\frac{\alpha_i}{P_i}(1-e^{-\lambda x})^{\alpha_i-1}$$

and a pdf of $Y_{n:n} = \sup(Y_1, Y_2, \ldots, Y_n)$ is given by

$$f_{n:n}(x) = \lambda n e^{-\lambda x}\left[\sum_{i=1}^{m}\frac{1}{P_i}(1-e^{-\lambda x})^{\alpha_i}\right]^{n-1}\sum_{i=1}^{m}\frac{\alpha_i}{P_i}(1-e^{-\lambda x})^{\alpha_i-1}.$$

Consequently, the cdf of the i-th order statistic $Y_{i:n}$ is given by

$$\begin{aligned} F_{i:n}(x) &= \int_0^x f_{i:n}(t)dt = \frac{n!}{(i-1)!(n-i)!}\sum_{k=0}^{n-i}\binom{n-i}{k}\frac{(-1)^k}{i+k}[F_Y(x)]^{i+k} \\ &= \frac{n!}{(i-1)!(n-i)!}\sum_{k=0}^{n-i}\binom{n-i}{k}\frac{(-1)^k}{i+k}\left[\sum_{i=1}^{m}\frac{1}{P_i}(1-e^{-\lambda x})^{\alpha_i}\right]^{i+k}. \end{aligned}$$

Further, for $j<k$, a joint pdf of $(Y_{j:n}, Y_{k:n})$ is given by, for $0<x_j<x_k$,

$$\begin{aligned} & f_{(j:n,k:n)}(x_j,x_k) \\ = & \frac{n!}{(j-1)!(n-k)!(k-j-1)}[F_Y(x_j)]^{j-1}[F_Y(x_k)-F_Y(x_j)]^{k-j-1}\times \\ & [1-F_Y(x_k)]^{n-k} f_Y(x_j) f_Y(x_k) \\ = & \frac{n!}{(j-1)!(n-k)!(k-j-1)}\left[\sum_{i=1}^{m}\frac{1}{P_i}(1-e^{-\lambda x_j})^{\alpha_i}\right]^{j-1}\times \\ & \left[\sum_{i=1}^{m}\frac{1}{P_i}[(1-e^{-\lambda x_k})^{\alpha_i}-(1-e^{-\lambda x_j})^{\alpha_i}]\right]^{k-j-1} \\ \times & \left[1-\sum_{i=1}^{m}\frac{1}{P_i}(1-e^{-\lambda x_k})^{\alpha_i}\right]^{n-k}\lambda^2 e^{-\lambda(x_j+x_k)}\sum_{i=1}^{m}\frac{\alpha_i}{P_i}(1-e^{-\lambda x_j})^{\alpha_i-1}\times \\ & \sum_{i=1}^{m}\frac{\alpha_i}{P_i}(1-e^{-\lambda x_k})^{\alpha_i-1}. \end{aligned}$$

5. Maximum likelihood estimation

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from the EH($\mathbf{a}$, λ) distribution with unknown parameters $\alpha_1, \ldots, \alpha_m, \lambda$. We consider the maximum likelihood estimation providing the maximum likelihood estimators (MLEs) $\hat{\alpha}_1, \ldots, \hat{\alpha}_m, \hat{\lambda}$ for the parameters $\alpha_1, \ldots, \alpha_m, \lambda$. Let us recall that the MLEs have some statistical desirable properties (under regularity conditions) as the sufficiency, invariance, consistency, efficiency and asymptotic normality. Using the observed information matrix, asymptotic confidence interval for $\alpha_1, \ldots, \alpha_m, \lambda$ can be constructed. Details can be found in Larsen and Marx (2000).

Let $\mathbf{\Theta} = (\alpha_1, \ldots, \alpha_m, \lambda)$ and $y_1, y_2, \ldots, y_n$ be the observed values. The likelihood function is given by

$$L(\mathbf{\Theta}) = \lambda^n e^{-\lambda \sum_{u=1}^{n} y_u} \prod_{u=1}^{n}\left(\sum_{i=1}^{m}\frac{\alpha_i}{P_i}(1-e^{-\lambda y_u})^{\alpha_i-1}\right).$$

The log-likelihood function is given by

$$\ell(\mathbf{\Theta}) = \log(L(\mathbf{\Theta})) = n\log(\lambda) - \lambda\sum_{u=1}^{n} y_u + \sum_{u=1}^{n}\log\left(\sum_{i=1}^{m}\frac{\alpha_i}{P_i}(1-e^{-\lambda y_u})^{\alpha_i-1}\right).$$

The nonlinear log-likelihood equations given by $\frac{\partial \ell(\Theta)}{\partial \Theta} = 0$ are listed below

$$\frac{\partial \ell(\Theta)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{u=1}^{n} y_u + \sum_{u=1}^{n} y_u e^{-\lambda y_u} \frac{\sum_{i=1}^{m} \frac{\alpha_i(\alpha_i - 1)}{P_i}(1-e^{-\lambda y_u})^{\alpha_i - 2}}{\sum_{i=1}^{m} \frac{\alpha_i}{P_i}(1-e^{-\lambda y_u})^{\alpha_i - 1}} = 0 \tag{6}$$

and, for any $q \in \{1, \dots, m\}$,

$$\frac{\partial \ell(\Theta)}{\partial \alpha_q} = \sum_{u=1}^{n} \frac{\frac{\partial}{\partial \alpha_q}\left(\sum_{i=1}^{m} \frac{\alpha_i}{P_i}(1-e^{-\lambda y_u})^{\alpha_i - 1}\right)}{\sum_{i=1}^{m} \frac{\alpha_i}{P_i}(1-e^{-\lambda y_u})^{\alpha_i - 1}} = 0. \tag{7}$$

Let us now investigate the numerator. We have

$$\sum_{i=1}^{m} \frac{\alpha_i}{P_i}(1-e^{-\lambda y_u})^{\alpha_i - 1} = \frac{\alpha_q}{P_q}(1-e^{-\lambda y_u})^{\alpha_q - 1} + \sum_{\substack{i=1 \\ i \neq q}}^{m} \frac{\alpha_i}{P_i}(1-e^{-\lambda y_u})^{\alpha_i - 1}.$$

Hence

$$\begin{aligned}
&\frac{\partial}{\partial \alpha_q}\left(\sum_{i=1}^{m} \frac{\alpha_i}{P_i}(1-e^{-\lambda y_u})^{\alpha_i - 1}\right) \\
&= \frac{\partial}{\partial \alpha_q}\left(\frac{\alpha_q}{P_q}\right)(1-e^{-\lambda y_u})^{\alpha_q - 1} + \frac{\alpha_q}{P_q}(1-e^{-\lambda y_u})^{\alpha_q - 1}\log(1-e^{-\lambda y_u}) \\
&+ \sum_{\substack{i=1 \\ i \neq q}}^{m} \alpha_i (1-e^{-\lambda y_u})^{\alpha_i - 1} \frac{\partial}{\partial \alpha_q}\left(\frac{1}{P_i}\right).
\end{aligned}$$

Observe that

$$\begin{aligned}
\frac{\partial}{\partial \alpha_q}\left(\frac{\alpha_q}{P_q}\right) &= \frac{\partial}{\partial \alpha_q}\left(\alpha_q \prod_{\substack{j=1 \\ j \neq q}}^{m} \frac{\alpha_j}{\alpha_j - \alpha_q}\right) \\
&= \prod_{\substack{j=1 \\ j \neq q}}^{m} \frac{\alpha_j}{\alpha_j - \alpha_q} + \alpha_q \left(\prod_{\substack{j=1 \\ j \neq q}}^{m} \frac{\alpha_j}{\alpha_j - \alpha_q}\right)\left(\sum_{\substack{v=1 \\ v \neq q}}^{m} \frac{1}{\alpha_v - \alpha_q}\right).
\end{aligned}$$

On the other hand, for $i \neq q$, we have

$$\frac{\partial}{\partial \alpha_q}\left(\frac{1}{P_i}\right) = \frac{\partial}{\partial \alpha_q}\left(\prod_{\substack{j=1\\ j\neq i}}^{m} \frac{\alpha_j}{\alpha_j-\alpha_i}\right) = \left(\prod_{\substack{j=1\\ j\neq\{i,q\}}}^{m} \frac{\alpha_j}{\alpha_j-\alpha_i}\right)\frac{\partial}{\partial \alpha_q}\left(\frac{\alpha_q}{\alpha_q-\alpha_i}\right)$$
$$= -\left(\prod_{\substack{j=1\\ j\neq\{i,q\}}}^{m} \frac{\alpha_j}{\alpha_j-\alpha_i}\right)\frac{\alpha_i}{(\alpha_q-\alpha_i)^2}.$$

Putting these equalities together, we obtain an unified equation for (7). The MLEs are solutions of (6) and (7). These equations are not solvable analytically, but some numerical iterative methods, as Newton-Raphson method, can be used. The solutions can be approximate numerically by using software such as MATHEMATICA, MAPLE and R. Here we work with MATHEMATICA, see Wolfram (1999).

Study of a particular case. Let us now consider a simple two parameters EH($\mathbf{a}, \lambda$) distribution where $\mathbf{a} = (\alpha_1, \ldots, \alpha_m)$, for any $i \in \{1, \ldots, m\}$, $\alpha_i = \alpha + \epsilon_i$, with $\epsilon_1, \ldots, \epsilon_m$ denote different fixed positive real numbers. Then $\alpha > 0$ and $\lambda > 0$ can be estimated via the maximum likelihood method. It is enough to set $\mathbf{\Theta} = (\alpha, \lambda)$, using (6) with $\alpha_i = \alpha + \epsilon_i$ and consider the new equation

$$\frac{\partial \ell(\mathbf{\Theta})}{\partial \alpha} = \sum_{u=1}^{n} \frac{\frac{\partial}{\partial \alpha}\left(\sum\limits_{i=1}^{m} \frac{\alpha+\epsilon_i}{P_i}(1-e^{-\lambda y_u})^{\alpha+\epsilon_i-1}\right)}{\sum\limits_{i=1}^{m} \frac{\alpha+\epsilon_i}{P_i}(1-e^{-\lambda y_u})^{\alpha+\epsilon_i-1}} = 0. \tag{8}$$

We can explicit the numerator by observing that

$$\frac{\partial}{\partial \alpha}\left(\sum_{i=1}^{m} \frac{\alpha+\epsilon_i}{P_i}(1-e^{-\lambda y_u})^{\alpha+\epsilon_i-1}\right)$$
$$= \sum_{i=1}^{m} \frac{\partial}{\partial \alpha}\left(\frac{\alpha+\epsilon_i}{P_i}\right)(1-e^{-\lambda y_u})^{\alpha+\epsilon_i-1}$$
$$+ \sum_{i=1}^{m} \frac{\alpha+\epsilon_i}{P_i}(1-e^{-\lambda y_u})^{\alpha+\epsilon_i-1}\ln(1-e^{-\lambda y_u}),$$

with

$$\frac{\partial}{\partial \alpha}\left(\frac{\alpha+\epsilon_i}{P_i}\right) = \left(\prod_{\substack{j=1\\ j\neq i}}^{m} \frac{1}{\epsilon_j-\epsilon_i}\right)\frac{\partial}{\partial \alpha}\left(\prod_{j=1}^{m}(\alpha+\epsilon_j)\right)$$
$$= \left(\prod_{\substack{j=1\\ j\neq i}}^{m} \frac{1}{\epsilon_j-\epsilon_i}\right)\left(\prod_{j=1}^{m}(\alpha+\epsilon_j)\right)\left(\sum_{j=1}^{m} \frac{1}{\alpha+\epsilon_j}\right).$$

Again, these equations can be solve numerically. Thanks to its simplicity, this particular EH($\mathbf{a}, \lambda$) distribution will be considered in the applications presented in the next section.

6. ILLUSTRATIVE REAL DATA EXAMPLES

In this section, we analysis real data sets to show that the EH($\mathbf{a}, \lambda$) distribution can be a better model than other existing distributions. We consider the following well-known weighted exponential distributions for comparison purpose.

Lindley distribution. A pdf associated to the Lindley distribution of parameter $\theta > 0$ is given by

$$f(x) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}, \qquad x > 0.$$

Theory and applications related to this distribution can be found in Ghitany *et al.* (2008).

Exponential distribution. A pdf associated to the exponential distribution of parameter $\lambda > 0$ is given by

$$f(x) = \lambda e^{-\lambda x}, \qquad x > 0.$$

Exponentiated exponential distribution (E Exponential). A pdf associated to the Exponentiated exponential distribution of parameters $\alpha, \lambda > 0$ is given by

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

Details about this distribution can be found in Gupta and Kundu (1999).

Exponential hypoexponential distribution (Exp Hypo). To simplify the situation, we consider a particular simple two parameters EH($\mathbf{a}, \lambda$) distribution with $m = 2$, $\mathbf{a} = (\alpha_1, \alpha_2)$, $\alpha_1 = \alpha$ and $\alpha_2 = \alpha + 0.1$, $\alpha > 0$ and $\lambda > 0$ are the parameters to be estimated; the value 0.1 is subjective. The corresponding pdf is given by

$$f(x) = \lambda\alpha(1+10\alpha)e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1}[1-(1-e^{-\lambda x})^{0.1}], \qquad x > 0.$$

REMARK 11. *Another two parameters EH($\mathbf{a}, \lambda$) distribution can be used, as the one defined with $m = 3$, say $\mathbf{a} = (\alpha_1, \alpha_2, \alpha_3)$, $\alpha_1 = \alpha$, $\alpha_2 = \alpha + \epsilon$ and $\alpha_3 = \alpha + \upsilon$, for fixed ϵ and υ with $\epsilon \neq \upsilon$, the parameters to be estimated are $\alpha > 0$ and $\lambda > 0$.*

The two considered real data sets are described as follows:

Data set 1. The data set contains $n = 63$ measures related to the strength of 1.5cm glass fibers. It is reported in Smith and Naylor (1987):

0.55, 0.93, 1.25, 1.36, 1.49, 1.52, 1.58, 1.61, 1.64, 1.68, 1.73, 1.81, 2, 0.74, 1.04, 1.27, 1.39, 1.49, 1.53, 1.59, 1.61, 1.66, 1.68, 1.76, 1.82, 2.01, 0.77, 1.11, 1.28, 1.42, 1.5, 1.54, 1.6, 1.62, 1.66, 1.69, 1.76, 1.84, 2.24, 0.81, 1.13, 1.29, 1.48, 1.5, 1.55, 1.61, 1.62, 1.66, 1.7, 1.77, 1.84, 0.84, 1.24, 1.3, 1.48, 1.51, 1.55, 1.61, 1.63, 1.67, 1.7, 1.78, 1.89

Data set 2. The data set contains $n = 128$ measures on the remission times in months of bladder cancer patients. It is extracted from Lee and Wang (2003):

0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69, 5.49

For each data set, we compare the fitted distributions using the criteria: $-2\log(L)$, AIC (Akaike information criterion), AICC (Akaike information criterion corrected), BIC (Bayesian information criterion), Anderson-Darling (A^*) and Cramer-von Mises (W^*). Let us precise that $\log(L)$ is the log-likelihood taking with the estimate values, AIC$=-2\log(L)+2k$, AICC$=$AIC$+\frac{2k(k+1)}{n-k-1}$ and BIC$=-2\log(L)+k\log(n)$, where k denotes the number of estimated parameters and n denotes the sample size. The best fitted distribution corresponds to lower $-2\log(L)$, AIC, AICC, BIC, A^* and W^*. We see in Tables 1 and 2 that the EH($\mathbf{a}, \lambda$) distribution has the smallest $-2\log(L)$, AIC, AICC, BIC, A^* and W^* for the two data sets (excepted for the second data set where the exponential distribution has a smallest BIC), indicating that it is a serious competitor to the other considered distributions. To show graphically the behaviors of our model, Figure 4 shows the fitted densities superimposed on the histogram of Data sets 1 and 2.

TABLE 1
The values of $-2\log(L)$, AIC, AICC and BIC of the fitted distributions of Data sets 1 and 2.

	Model	Estimate of parameters	$-2\log(L)$	AIC	AICC	BIC
Data set 1	Lindley	$\hat{\theta} = 0.996116$	162.56	164.56	164.62	166.70
	Exponential	$\hat{\lambda} = 0.663647$	177.66	179.66	179.73	181.80
	E Exponential	$(\hat{\alpha}, \hat{\lambda}) = (31.3489, 2.61157)$	62.76	66.76	66.96	71.05
	Exp Hypo	$(\hat{\alpha}, \hat{\lambda}) = (24.0816, 1.83894)$	55.67	59.67	59.87	63.96
Data set 2	Lindley	$\hat{\theta} = 0.196$	839.05	841.06	841.09	843.89
	Exponential	$\hat{\lambda} = 0.106773$	828.68	830.68	830.71	833.53
	E Exponential	$(\hat{\alpha}, \hat{\lambda}) = (1.21795, 0.121167)$	826.15	830.15	830.25	835.85
	Exp Hypo	$(\hat{\alpha}, \hat{\lambda}) = (1.44399, 0.057924)$	825.49	829.49	829.59	835.20

TABLE 2
The values of A^* and W^* of the fitted distributions of Data sets 1 and 2.

	Model	A^*	W^*
Data set 1	Lindley	16.44785	3.35845
	Exponential	18.65557	3.89287
	E Exponential	4.39118	0.80479
	Exp Hypo	3.80978	0.69927
Data set 2	Lindley	2.80755	0.52238
	Exponential	1.18068	0.17951
	E Exponential	0.71817	0.12839
	Exp Hypo	0.66252	0.11474

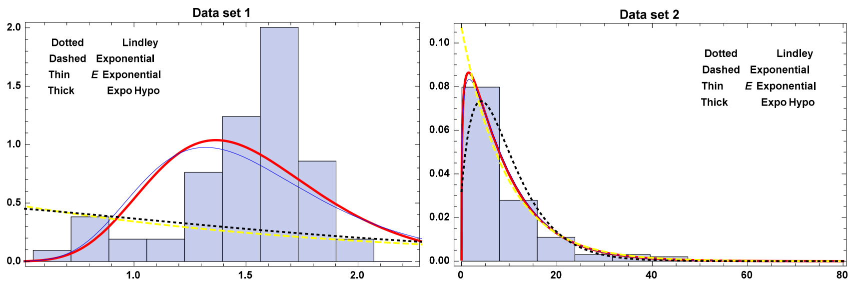

Figure 4 – The fitted densities superimposed on the histogram of Data sets 1 and 2.

7. Conclusion

A new weighted exponential distribution based on the hypoexponential distribution is introduced. Some of its structural properties are studied. The analysis of two real data sets shows that the fit of the model related to our new distribution can be superior to other models. As future work, we plan to study the GH(H; $\mathbf{a}$) distribution with another cdf $H(x)$ and also to provide more applications for the EH($\mathbf{a}, \lambda$) distribution with $m \geq 3$. Inverted or transmuted transformations of EH($\mathbf{a}, \lambda$) can be of interest too for some applications.

Acknowledgements

The authors would like to thank the referees for their valuable comments which helped to improve the manuscript.

References

M. Akkouchi (2008). *On the convolution of exponential distributions*. Journal of the Chungcheong Mathematical Society, 21, no. 4, pp. 501–510.

E. K. Al-Hussaini (2012). *Composition of cumulative distribution functions*. Journal of Statistical Theory and Applications, 11, pp. 323–336.

S. V. AMARI, R. B. MISRA (1997). *Closed-form expression for distribution of the sum of independent exponential random variables*. IEEE Transactions on Reliability, 46, no. 4, pp. 519–522.

H. S. BAKOUCH, B. V. POPOVIĆ (2016). *Lindley first-order autoregressive model with applications*. Communications in Statistics - Theory and Methods, 45, no. 17, pp. 4988–5006.

D. J. BARTHOLOMEW (1969). *Sufficient conditions for a mixture of exponentials to be a probability density function*. Annals of Mathematical Statistics, 40, pp. 2183–2188.

S. BHATTACHARJEE (2018). *A counting process with generalized exponential inter-arrival times*. Statistica, 78, no. 1, pp. 3–14.

A. CHATURVEDI, S. VYAS (2017). *Estimation and testing procedures for the reliability functions of exponentiated distributions under censorings*. Statistica, 77, no. 1, pp. 13–31.

G. M. CORDEIRO, A. E. GOMES, C. Q. DA SILVA, E. M. M. ORTEGA (2013). *The beta exponentiated Weibull distribution*. Journal of Statistical Computation and Simulation, 83, no. 1, pp. 114–138.

M. FRANCO, N. BALAKRISHNAN, D. KUNDU, J. M. VIVO (2013). *Generalized mixtures of Weibull components*. Test, 23, no. 3, pp. 515–535.

M. E. GHITANY, B. ATIEH, S. NADARAJAH (2008). *Lindley distribution and its applications*. Mathematics and Computers in Simulation, 78, no. 4, pp. 493–506.

R. D. GUPTA, D. KUNDU (1999). *Generalized exponential distributions*. Australian and New Zealand Journal of Statistics, 41, no. 2, pp. 173–188.

R. D. GUPTA, D. KUNDU (2001). *Exponentiated exponential distribution: an alternative to gamma and Weibull distributions*. Biometrical Journal, 43, no. 1, pp. 117–130.

M. S. KHAN, R. KING, I. L. HUDSON (2017). *Transmuted generalized exponential distribution: A generalization of the exponential distribution with applications to survival data*. Communications In Statistics - Simulation And Computation, 46, no. 6, pp. 4377–4398.

D. KUNDU, R. D. GUPTA (2005). *Estimation of $p(y < x)$ for generalized exponential distribution*. Metrika, 61, no. 3, pp. 291–308.

R. J. LARSEN, M. L. MARX (2000). *An introduction to mathematical statistics and its applications*. Pearson, Boston.

E. T. LEE, J. W. WANG (2003). *Statistical Methods for Survival Data Analysis*. Wiley, New York.

A. MAHDAVI (2015). *Two weighted distributions generated by exponential distribution*. Journal of Mathematical Extension, 9, no. 3, pp. 1–12.

F. MEROVCI, I. ELBATAL (2013). *The McDonald modified Weibull distribution: Properties and applications*. Eprint, arXiv:1309.2961.

M. PAL, M. TIENSUWAN (2015). *Some properties of gamma generated distributions*. Statistica, 75, no. 4, pp. 391–403.

M. M. RISTIĆ, N. BALAKRISHNAN (2012). *The gamma-exponentiated exponential distribution*. Journal of Statistical Computation and Simulation, 82, no. 8, pp. 1191–1206.

A. SAGHIR, G. G. HAMEDANI, S. TAZEEM, A. KHADIM (2017). *Weighted distributions: A brief review, perspective and characterizations*. International Journal of Statistics and Probability, 6, no. 3, pp. 109–131.

K. SMAILI, T. KADRI, S. KADRY (2013). *Hypoexponential distribution with different parameters*. Applied Mathematics, 4, no. 4, pp. 624–631.

R. L. SMITH, J. C. NAYLOR (1987). *A comparison of maximum likelihood and Bayesian estimators for the three parameter Weibull distribution*. Applied Mathematics, 36, no. 3, pp. 358–369.

S. WOLFRAM (1999). *The Mathematica Book*. Cambridge University Press, Cambridge.

K. ZOGRAFOS, N. BALAKRISHNAN (2009). *On families of beta- and generalized gamma-generated distributions and associated inference*. Statistical Methodology, 6, no. 4, pp. 344–362.

SUMMARY

In this paper, a new general family of distributions using the hypoexponential distribution is introduced and studied. A special case of this family is explored in detail, corresponding to a new finite generalized mixture of generalized exponential distributions. Some of their mathematical properties are provided. We investigate maximum likelihood estimation of the model parameters. Two real data sets are used to prove the potential of this distribution among some recent extensions of the exponential distribution.

Keywords: Hypoexponential distribution; Generalized mixture; Moments; Maximum likelihood estimation.