


Article

# On the Analysis of New COVID-19 Cases in Pakistan Using an Exponentiated Version of the M Family of Distributions

Rashad A. R. Bantan <sup>1</sup>, Christophe Chesneau <sup>2,\*</sup>, Farrukh Jamal <sup>3</sup>  and Mohammed Elgarhy <sup>4</sup>

<sup>1</sup> Department of Marine Geology, Faculty of Marine Science, King Abdulaziz University, Jeddah 21551, Saudi Arabia; rbantan@kau.edu.sa

<sup>2</sup> Department of Mathematics, Université de Caen, LMNO, Campus II, Science 3, 14032 Caen, France

<sup>3</sup> Department of Statistics, Govt. S.A Postgraduate College Dera Nawab Sahib, Bahawalpur 63100, Punjab, Pakistan; drfarrukh1982@gmail.com

<sup>4</sup> Valley High Institute for Management Finance and Information Systems, Obour 11828, Qaliubia, Egypt; m\_elgarhy85@sva.edu.eg

\* Correspondence: christophe.chesneau@unicaen.fr; Tel.: +33-02-3156-7424

Received: 12 May 2020; Accepted: 8 June 2020; Published: 11 June 2020



**Abstract:** This paper develops the exponentiated Mfamily of continuous distributions, aiming to provide new statistical models for data fitting purposes. It stands out from the other families, as it depends on two baseline distributions, with the use of ratio and power transforms in the definition of the main cumulative distribution function. Thanks to the joint action of the possibly different baseline distributions, flexible statistical models can be created, motivating a complete study in this regard. Thus, we discuss the theoretical properties of the new family, with emphasis on those of potential interest to the overall probability and statistics. Then, a new three-parameter lifetime distribution is derived, with the choices of the inverse exponential and exponential distributions as baselines. After pointing out the great flexibility of the related model, we apply it to analyze an actual dataset of current interest: the daily COVID-19 cases observed in Pakistan from 21 March to 29 May 2020 (inclusive). As notable results, we demonstrate that the proposed model is the best among the 15 top ranked models in the literature, including the inverse exponential and exponential models, several modern extensions of them depending on more parameters, and the “unexponentiated” version of the proposed model as well. As future perspectives, the proposed model can be of interest to analyze data on COVID-19 cases in other countries, for possible comparison studies.

**Keywords:** families of continuous distributions; exponentiated family of continuous distributions; entropy; parameter estimation; data analysis; COVID-19 epidemic

**MSC:** 62N05; 90B25

## 1. Introduction

The modeling and analysis of real-life data are essential to understand important features of random phenomena and to draw suitable conclusions as well. In particular, this requires the choice of statistical models based on probability distributions, whose adequateness against the observations will strongly influence the pertinence of the outputs. The analysis of recent data in applied sciences (environmental sciences, engineering, finance, etc.) has shown the limitations of the classical distributions, whose flexibility does not allow revealing some important details. To go further into these limitations, new distributions, often divided into specific families of distributions, have been created. A short list of the notorious families is the following: the

skew-normal family (see [1]), Marshall–Olkin-Gfamily (see [2]), exponentiated-G family (see [3]), beta-G family (see [4]), order statistics-G family (see [5]), sinh-arcsinh-G family (see [6]), transmuted-G (see [7]), gamma-G family (see [8]), Kumaraswamy-G (see [9]), Topp–Leone-G (see [10]), and ratio-exponentiated-G (see [11]). The global motivation behind them is to extend the modeling properties of a classical baseline distribution by adding one or more tuning parameters through the use of various flexible transformations (power, beta, gamma, ratio, etc.).

Among all the proposed families, the Mfamily of continuous distributions introduced by [12] stands out from the others due to its original construction; it is defined by a cumulative distribution function (cdf) based on a ratio involving two baseline cdfs, with possibly different characteristics. More specifically, the corresponding cdf is defined by:

$$F(x; \zeta_1, \zeta_2) = \frac{F_1(x; \zeta_1) + F_2(x; \zeta_2)}{1 + F_1(x; \zeta_1)}, \quad x \in \mathbb{R}, \tag{1}$$

where  $F_1(x; \zeta_1)$  and  $F_2(x; \zeta_2)$  are two cdfs of continuous distributions with sets of parameters represented by  $\zeta_1$  and  $\zeta_2$ , respectively. These two baseline cdfs can be chosen independently of each other, without a particular condition. However, for practical purposes, in order to avoid the over-parametrization phenomenon, it is recommended not to have too many parameters involved; one can reduce  $\zeta_1$  and  $\zeta_2$  to a unique parameter, or take  $\zeta_1$  and  $\zeta_2$  as two different parameters, or  $\zeta_1$  can be chosen as a subset of parameters of  $\zeta_2$ , or vice versa. Clearly, the M family contains a plethora of ratio distributions and models, since a multitude of choices for  $F_1(x; \zeta_1)$  and  $F_2(x; \zeta_2)$  is possible. However, to the best of our knowledge, this versatile aspect has not been fully explored yet. Indeed, in the former work of [12],  $F(x; \zeta_1, \zeta_2)$  was presented as (1), with the proof that it satisfied the properties of a valid cdf. Then, as a direct application, a new two-parameter lifetime distribution was defined by the cdf (1) under the following simple configuration:  $F_1(x; \zeta_1) = F_2(x; \zeta_2)$ , and  $F_1(x; \zeta_1)$  was chosen as the cdf of the Weibull distribution, i.e.,  $F_1(x; \zeta_1) = F_1(x; a, b) = 1 - e^{-(x/b)^a}$ ,  $a, b, x > 0$ . Thus, the corresponding cdf is given by:

$$F(x; a, b) = \frac{2 \left(1 - e^{-(x/b)^a}\right)}{2 - e^{-(x/b)^a}}, \quad a, b, x > 0. \tag{2}$$

As a main application, it was proven that the related model had a better fit to the exponentiated exponential, Weibull, and gamma models, for the failure times of the air conditioning system data from [13]. This nice result validated the entry of the M family on the short list. However, for the special configuration  $F_1(x; \zeta_1) = F_2(x; \zeta_2)$ , the M family loses its intrinsic originality for the following reasons: (i) it does not mix different features for the baseline cdfs; (ii) it is included in the well-known Marshall–Olkin family since we can express  $F(x; \zeta_1, \zeta_2)$  as:

$$F(x; \zeta_1, \zeta_2) = \frac{F_1(x; \zeta_1)}{1 - (1 - \theta)[1 - F_1(x; \zeta_1)]},$$

with  $\theta = 1/2$ . That is, the general form of  $F(x; \zeta_1, \zeta_2)$  is exploited at its minimum; the M family has not revealed all of its potential.

Based on the previous setting, the multiple contributions of the paper can be summarized as follows: (i) We introduce a simple and natural extension of the M family by the use of the power transform, called the EMfamily. (ii) We provide some mathematical results of this family, which are also new and applicable to the former M family. (iii) We consider  $F_1(x; \zeta_1)$  and  $F_2(x; \zeta_2)$  of different natures, i.e., exponential and inverse exponential, respectively, to create a new promising (three-parameter lifetime) distribution, which demonstrates a high modeling ability for data fitting; versatile shapes are observed for the main functions. (iv) We investigate the estimation of the model parameters by a top ranked method in terms of efficiency: the maximum likelihood method. (v) We apply this model to an actual dataset of COVID-19 cases observed in Pakistan during the year 2020. As a main result, for these

data of particular interest, the proposed model possesses an excellent fitting behavior, better than that of 15 other top ranked models in the literature, attesting to the importance of these findings.

The remainder of the works is outlined as follows. In Section 2, the EM family is introduced, and some of its mathematical results are proven. A special distribution of interest is presented in Section 3, with discussions. The estimation of the related model parameters is studied in Section 4. The application to a COVID-19 dataset is presented in Section 5. The conclusion is given in Section 6.

## 2. The EM Family

Here, the EM family is defined, with some of its important mathematical properties.

### 2.1. Definition

The EM family is the exponentiated version of the M family, that is a mix between the exponentiated-G and M families developed by [3,12], respectively. It is defined by the following cdf:

$$F(x; \zeta_1, \zeta_2, \gamma) = \left[ \frac{F_1(x; \zeta_1) + F_2(x; \zeta_2)}{1 + F_1(x; \zeta_1)} \right]^\gamma, \quad x \in \mathbb{R}, \tag{3}$$

where  $\gamma > 0$  is a shape parameter and  $F_1(x; \zeta_1)$  and  $F_2(x; \zeta_2)$  are two baseline cdfs of continuous distributions with sets of parameters represented by  $\zeta_1$  and  $\zeta_2$ , respectively. These two baseline cdfs can be chosen independently of each other. Naturally, by taking  $\gamma = 1$ , we rediscover the cdf of the former M family. The role of  $\gamma$  is to flexibilize the rigid ratio cdf given by (1), aiming to improve several of its characteristics (skewness, kurtosis, tails' heaviness, modes' properties, etc.). Furthermore, the following stochastic ordering result holds:

- if  $\gamma \leq 1$ , we have  $F(x; \zeta_1, \zeta_2) \leq F(x; \zeta_1, \zeta_2, \gamma)$ ,
- if  $\gamma > 1$ , we have  $F(x; \zeta_1, \zeta_2, \gamma) \leq F(x; \zeta_1, \zeta_2)$ ,

showing different perspectives of modeling for the EM family in comparison to the former M family. Among the notable studies employing the exponentiated technique, we may refer the reader to [14–16].

In addition to the cdf, the probability density function (pdf) of a continuous distribution plays a fundamental role in probability and statistics. The pdf of the EM family is given by differentiating  $F(x; \zeta_1, \zeta_2, \gamma)$  with respect to  $x$ , almost surely. After some developments, it is obtained as:

$$f(x; \zeta_1, \zeta_2, \gamma) = \gamma \frac{[1+F_1(x; \zeta_1)]f_2(x; \zeta_2) + [1-F_2(x; \zeta_2)]f_1(x; \zeta_1)}{[1+F_1(x; \zeta_1)]^2} \left[ \frac{F_1(x; \zeta_1) + F_2(x; \zeta_2)}{1 + F_1(x; \zeta_1)} \right]^{\gamma-1}, \tag{4}$$

$x \in \mathbb{R}$ ,

where  $f_1(x; \zeta_1)$  and  $f_2(x; \zeta_2)$  are the pdfs corresponding to  $F_1(x; \zeta_1)$  and  $F_2(x; \zeta_2)$ , respectively.

That is, for a random variable  $X$  defined on a generic probability set, say  $(\Omega, \mathcal{A}, P)$ , having the pdf of the EM family and any set  $A \subseteq \mathbb{R}$ , we have:

$$P(X \in A) = \int_A f(x; \zeta_1, \zeta_2, \gamma) dx.$$

The pdf is also central in the transfer theorem, which ensures that, for any function of  $X$ , say  $T(X)$ , the expectation of  $T(X)$  is given by:

$$E[T(X)] = \int_{-\infty}^{+\infty} T(x) f(x; \zeta_1, \zeta_2, \gamma) dx, \tag{5}$$

provided that it exists. From this formula, several types of moments, coefficients, probabilistic functions, and entropy can be defined (see [17]). Let us mention that, thanks to their integral expressions,  $P(X \in A)$  and  $E[T(X)]$  can be determined numerically with the help of any mathematical software.

### 2.2. Reliability Functions

The following functions of the EM family are central in various probability and statistics areas, with emphasis on reliability analysis. First of all, the survival function (sf) is specified as:

$$S(x; \xi_1, \xi_2, \gamma) = 1 - F(x; \xi_1, \xi_2, \gamma) = 1 - \left[ \frac{F_1(x; \xi_1) + F_2(x; \xi_2)}{1 + F_1(x; \xi_1)} \right]^\gamma, \quad x \in \mathbb{R}.$$

Furthermore, the hazard rate function (hrf), reversed hazard rate function (rhrf), and cumulative hazard rate function (chrf) are given by, respectively,

$$h(x; \xi_1, \xi_2, \gamma) = \frac{f(x; \xi_1, \xi_2, \gamma)}{S(x; \xi_1, \xi_2, \gamma)} = \gamma \frac{[1 + F_1(x; \xi_1)] f_2(x; \xi_2) + [1 - F_2(x; \xi_2)] f_1(x; \xi_1)}{[1 + F_1(x; \xi_1)] \{ [1 + F_1(x; \xi_1)]^\gamma - [F_1(x; \xi_1) + F_2(x; \xi_2)]^\gamma \}} [F_1(x; \xi_1) + F_2(x; \xi_2)]^{\gamma-1},$$

$$r(x; \xi_1, \xi_2, \gamma) = \frac{f(x; \xi_1, \xi_2, \gamma)}{F(x; \xi_1, \xi_2, \gamma)} = \gamma \frac{[1 + F_1(x; \xi_1)] f_2(x; \xi_2) + [1 - F_2(x; \xi_2)] f_1(x; \xi_1)}{[1 + F_1(x; \xi_1)] [F_1(x; \xi_1) + F_2(x; \xi_2)]},$$

and:

$$H(x; \xi_1, \xi_2, \gamma) = -\log[S(x; \xi_1, \xi_2, \gamma)] = -\log \{ [1 + F_1(x; \xi_1)]^\gamma - [F_1(x; \xi_1) + F_2(x; \xi_2)]^\gamma \} + \gamma \log[1 + F_1(x; \xi_1)], \quad x \in \mathbb{R}.$$

All the details on these functions, along with their applications in concrete settings, can be found in [18]. In the next part of the study, emphasis will be put on the pdf and hrf, due to their strong meaning in the fitting of data.

### 2.3. Properties

Stochastic ordering results: Now, we aim to compare the EM family with other existing families of distributions in the (usual) stochastic ordering sense (see [19]). That is, for two random variables  $X$  and  $Y$  for which at least one has the cdf of the EM family, we formalize the fact that  $X$  is less likely than  $Y$  to take any value lower than  $x$ , i.e.,  $P(X \leq x) \leq P(Y \leq x)$ . The main results are presented in the following proposition.

**Proposition 1.** *The following inequalities hold.*

- For  $\gamma_2 \geq \gamma_1$ , we have:

$$F(x; \xi_1, \xi_2, \gamma_2) \leq F(x; \xi_1, \xi_2, \gamma_1).$$

- In all cases, we have:

$$F(x; \xi_1, \xi_2, \gamma) \geq K(x; \xi_1, \xi_2, \gamma),$$

where  $K(x; \xi_1, \xi_2, \gamma)$  is the cdf of the exponentiated uniformly weighted two-component mixtures  $F_1(x; \xi_1)$  and  $F_2(x; \xi_2)$  with power parameter  $\gamma$ , i.e.,

$$K(x; \xi_1, \xi_2, \gamma) = [\lambda F_1(x; \xi_1) + (1 - \lambda) F_2(x; \xi_2)]^\gamma,$$

with  $\lambda = 1/2$ .

- If  $F_2(x; \xi_2) \geq F_1(x; \xi_1)$ , then we have:

$$F(x; \xi_1, \xi_2, \gamma) \geq Q(x; \xi_1, \xi_2, \gamma),$$

where  $Q(x; \xi_1, \xi_2, \gamma)$  is the cdf of the exponentiated “special cdf of the M family” introduced by (Section 2, [12]), i.e.,

$$Q(x; \xi_1, \xi_2, \gamma) = \left[ \frac{2F_1(x; \xi_1)}{1 + F_1(x; \xi_1)} \right]^\gamma.$$

**Proof.** For the first point, since  $[F_1(x; \xi_1) + F_2(x; \xi_2)]/[1 + F_1(x; \xi_1)] \in [0, 1]$ ,  $F(x; \xi_1, \xi_2, \gamma)$  is a decreasing function with respect to  $\gamma$ , giving the desired inequality. We prove the second point by noticing that  $1/[1 + F_1(x; \xi_1)] \geq 1/2$ , which gives:

$$F(x; \xi_1, \xi_2, \gamma) \geq \left\{ \frac{1}{2} [F_1(x; \xi_1) + F_2(x; \xi_2)] \right\}^\gamma = K(x; \xi_1, \xi_2, \gamma).$$

The third point follows immediately from the following inequality:  $F_1(x; \xi_1) + F_2(x; \xi_2) \geq 2F_1(x; \xi_1)$ . This ends the proof of Proposition 1.  $\square$

From the statistical point of view, Proposition 1 reveals a distributional hierarchy within the EM models and among other well-identified models of the literature. Thus, in comparison to these models, the EMIEE models offer new alternatives, depending on the repartition of the data.

Some series expansions: A series expansion of the cdf of the EM family in terms of simpler cdfs (defined as the product of exponentiated baseline cdfs) is described in the following proposition.

**Proposition 2.** *The cdf of the EM family can be expressed as:*

$$F(x; \xi_1, \xi_2, \gamma) = \sum_{k=0}^{+\infty} \sum_{\ell=0}^k \sum_{m=0}^{+\infty} a_{k,\ell,m} G_{\ell,m}(x; \xi_1, \xi_2),$$

where:

$$a_{k,\ell,m} = \binom{\gamma}{k} \binom{k}{\ell} \binom{-k}{m} (-1)^{k+\ell}$$

and  $G_{\ell,m}(x; \xi_1, \xi_2) = [F_1(x; \xi_1)]^m [F_2(x; \xi_2)]^\ell$ , which is a well-identified cdf.

**Proof.** The key of the proof is to notice that  $F(x; \xi_1, \xi_2, \gamma)$  can be written as:

$$F(x; \xi_1, \xi_2, \gamma) = \left[ 1 - \frac{1 - F_2(x; \xi_2)}{1 + F_1(x; \xi_1)} \right]^\gamma.$$

Now, from this expression, since  $[1 - F_2(x; \xi_2)]/[1 + F_1(x; \xi_1)] \in (0, 1)$  and  $F_1(x; \xi_1) \in (0, 1)$ , the general and standard binomial formulae give:

$$\begin{aligned} F(x; \xi_1, \xi_2, \gamma) &= \sum_{k=0}^{+\infty} \binom{\gamma}{k} (-1)^k [1 - F_2(x; \xi_2)]^k [1 + F_1(x; \xi_1)]^{-k} \\ &= \sum_{k=0}^{+\infty} \binom{\gamma}{k} (-1)^k \left\{ \sum_{\ell=0}^k \binom{k}{\ell} (-1)^\ell [F_2(x; \xi_2)]^\ell \right\} \left\{ \sum_{m=0}^{+\infty} \binom{-k}{m} [F_1(x; \xi_1)]^m \right\} \\ &= \sum_{k=0}^{+\infty} \sum_{\ell=0}^k \sum_{m=0}^{+\infty} a_{k,\ell,m} G_{\ell,m}(x; \xi_1, \xi_2). \end{aligned}$$

Then, we can remark that  $G_{\ell,m}(x; \xi_1, \xi_2)$  is a valid cdf, corresponding to the one of the random variable  $\max(X_1^{(1)}, \dots, X_1^{(m)}, X_2^{(1)}, \dots, X_2^{(\ell)})$ , where  $X_1^{(1)}, \dots, X_1^{(m)}$  are  $m$  independent and identically distributed (iid) random variables having the cdf  $F_1(x; \xi_1)$  and  $X_2^{(1)}, \dots, X_2^{(\ell)}$  are  $\ell$  iid random variables having the cdf  $F_2(x; \xi_2)$ , also independent of  $X_1^{(1)}, \dots, X_1^{(m)}$ . This ends the proof of Proposition 2.  $\square$

From Proposition 2, we can derive a useful sum expression of the pdf of the EM family, as presented below.

**Corollary 1.** *Let us consider the notations of Proposition 2. The pdf of the EM family can be expressed as:*

$$f(x; \xi_1, \xi_2, \gamma) = \sum_{k=0}^{+\infty} \sum_{\ell=0}^k \sum_{m=0}^{+\infty} a_{k,\ell,m} g_{\ell,m}(x; \xi_1, \xi_2),$$

where  $g_{\ell,m}(x; \xi_1, \xi_2)$  is the pdf corresponding to  $G_{\ell,m}(x; \xi_1, \xi_2)$ , i.e.,

$$g_{\ell,m}(x; \xi_1, \xi_2) = m f_1(x; \xi_1) [F_1(x; \xi_1)]^{m-1} [F_2(x; \xi_2)]^\ell + \ell f_2(x; \xi_2) [F_1(x; \xi_1)]^m [F_2(x; \xi_2)]^{\ell-1}.$$

Owing to Corollary 1 and (5), one can provide the following sum expression for  $E[T(X)]$ :

$$E[T(X)] = \sum_{k=0}^{+\infty} \sum_{\ell=0}^k \sum_{m=0}^{+\infty} a_{k,\ell,m} \mathbb{I}_{k,\ell,m}, \tag{6}$$

where, by denoting  $Q_1(u; \xi_1)$  and  $Q_2(u; \xi_2)$  the inverse functions of  $F_1(x; \xi_1)$  and  $F_2(x; \xi_2)$ , respectively,  $\mathbb{I}_{k,\ell,m}$  is given by:

$$\begin{aligned} \mathbb{I}_{k,\ell,m} &= \int_{-\infty}^{+\infty} T(x) g_{\ell,m}(x; \xi_1, \xi_2) dx \\ &= m \int_{-\infty}^{+\infty} T(x) f_1(x; \xi_1) [F_1(x; \xi_1)]^{m-1} [F_2(x; \xi_2)]^\ell dx + \ell \int_{-\infty}^{+\infty} T(x) f_2(x; \xi_2) [F_1(x; \xi_1)]^m [F_2(x; \xi_2)]^{\ell-1} dx \\ &= m \int_0^1 u^{m-1} T[Q_1(u; \xi_1)] \{F_2[Q_1(u; \xi_1); \xi_2]\}^\ell dx + \ell \int_0^1 u^{\ell-1} T[Q_2(u; \xi_2)] \{F_1[Q_2(u; \xi_2); \xi_1]\}^m du. \end{aligned}$$

The involved integrals can have closed-forms, depending on the complexity of  $T[Q_1(u; \xi_1)]$ ,  $F_2[Q_1(u; \xi_1); \xi_2]$ ,  $T[Q_2(u; \xi_2)]$ , and  $F_1[Q_2(u; \xi_2); \xi_1]$ . Then, one can admit the following useful approximation:

$$E[T(X)] \approx \sum_{k=0}^K \sum_{\ell=0}^k \sum_{m=0}^M a_{k,\ell,m} \mathbb{I}_{k,\ell,m},$$

where  $K$  and  $M$  denote large integers, such that the residual term of the approximation is negligible. Hence, we approximate the complicated integral quantity  $E[T(X)]$  by a finite sum of computable coefficients, which can be more efficient than computing the integral directly.

### 3. On a Special EM Distribution

The EM family contains a myriad of new ratio distributions. Here, we focus on a new promising one, exploiting the mix of possibly baseline cdfs of a different nature.

#### 3.1. Definition and Shapes' Analysis

Here, we introduce the EM inverse exponential exponential (EMIEE) distribution with parameters  $\alpha > 0$ ,  $\beta > 0$ , and  $\gamma > 0$ , defined by the cdf given by (3), under the following configuration:

- $F_1(x; \alpha) = e^{-\alpha/x}$ ,  $x > 0$ , corresponding to the cdf of the inverse exponential distribution with parameter  $\alpha$  (see [20]),
- $F_2(x; \beta) = 1 - e^{-\beta x}$ ,  $x > 0$ , corresponding to the cdf of the standard exponential distribution with parameter  $\beta$ .

The choice of these functions is motivated by the following arguments: (i)  $F_1(x; \alpha)$  and  $F_2(x; \beta)$  are simple, both depending on only one parameter; (ii) the inverse exponential and exponential

distributions are complementary, showing different characteristics on the tails, with various polynomial-exponential decay, summarized in the following relation:

$$F_1(x; \alpha) = 1 - F_2\left[\frac{\alpha}{\beta x}; \beta\right].$$

We thus aim to mix the features of these two distributions following the scheme of the EM family. That is, the cdf of the EMIEE distribution is the following:

$$F(x; \alpha, \beta, \gamma) = \left[\frac{e^{-\alpha/x} + 1 - e^{-\beta x}}{1 + e^{-\alpha/x}}\right]^\gamma = \left[1 - \frac{e^{-\beta x}}{1 + e^{-\alpha/x}}\right]^\gamma, \quad x, \alpha, \beta, \gamma > 0. \tag{7}$$

It represents a three-parameter lifetime distribution, with remarkable flexible properties. This aspect is developed in the next part of the study. At first glance, note that, if  $\alpha \rightarrow +\infty$ , then  $F(x; \alpha, \beta, \gamma)$  becomes the cdf of the exponentiated exponential distribution with parameters  $\beta$  and  $\gamma$  (see [3]), and if  $\gamma = 1$ , we obtain the “unexponentiated” version of the distribution, naturally called the MIEE distribution.

By differentiating with respect to  $x$ , the pdf of the EMIEE distribution is given by:

$$f(x; \alpha, \beta, \gamma) = \gamma e^{\alpha/x - \beta x} \frac{\alpha + \beta x^2(1 + e^{\alpha/x})}{x^2(1 + e^{\alpha/x})^2} \left[1 - \frac{e^{-\beta x}}{1 + e^{-\alpha/x}}\right]^{\gamma-1}, \quad x, \alpha, \beta, \gamma > 0. \tag{8}$$

All the functions of Section 2.2 can be expressed in a similar manner. Here, we only mention the hrf, which remains of great interest for such a lifetime distribution (see [18]). Therefore, it is expressed as:

$$h(x; \alpha, \beta, \gamma) = \gamma e^{-\beta x} \frac{[\alpha + \beta x^2(1 + e^{\alpha/x})] [e^{-\alpha/x} + 1 - e^{-\beta x}]^{\gamma-1}}{x^2(1 + e^{\alpha/x}) \{ [1 + e^{-\alpha/x}]^\gamma - [e^{-\alpha/x} + 1 - e^{-\beta x}]^\gamma \}}, \quad x, \alpha, \beta, \gamma > 0. \tag{9}$$

The rest of the study is devoted to some properties of the EMIEE distribution, beginning with the shape properties of  $f(x; \alpha, \beta, \gamma)$  and  $h(x; \alpha, \beta, \gamma)$ .

The mode(s) analysis of the EMIEE distribution provides important information on the “tops of the bell shapes” of the related model. Mathematically, the mode(s) can be obtained by solving the following equation:  $df(x; \alpha, \beta, \gamma)/dx = 0$ , which is equivalent to:

$$-\frac{\alpha}{x^2} - \beta - \beta \frac{\alpha e^{\alpha/x} - 2x(1 + e^{\alpha/x})}{\alpha + \beta x^2(1 + e^{\alpha/x})} - \frac{2}{x} + 2\alpha \frac{e^{\alpha/x}}{x^2(1 + e^{\alpha/x})} + (\gamma - 1)e^{-\beta x} \frac{\alpha + \beta x^2(1 + e^{\alpha/x})}{x^2(1 + e^{\alpha/x})(1 + e^{-\alpha/x} - e^{-\beta x})} = 0.$$

Several solutions are possible, depending on the values of  $\alpha$ ,  $\beta$  and  $\gamma$ . After investigations, the EMIEE distribution is revealed to be unimodal or bimodal (including a “limiting mode” in zero in this last case). An analytical expression for a mode seems however not possible. The asymptotes of  $f(x; \alpha, \beta, \gamma)$  are studied below. After some developments, we get:

$$f(x; \alpha, \beta, \gamma) \sim \gamma \beta^\gamma x^{\gamma-1}, \quad x \rightarrow 0,$$

implying that  $f(x; \alpha, \beta, \gamma) \rightarrow +\infty$  if  $\gamma \in (0, 1)$ ,  $f(x; \alpha, \beta, \gamma) \rightarrow \beta$  if  $\gamma = 1$  and  $f(x; \alpha, \beta, \gamma) \rightarrow 0$  if  $\gamma > 1$ . This illustrates the importance of the power parameter  $\gamma$  in these asymptotes. Furthermore, we have:

$$f(x; \alpha, \beta, \gamma) \sim \frac{\gamma \beta}{2} e^{-\beta x} \rightarrow 0, \quad x \rightarrow +\infty.$$



One can remark that the parameter  $\alpha$  plays no role in these asymptotes. However, the fine variations of  $f(x; \alpha, \beta, \gamma)$  are complicated to handle analytically, due to a high level of complexity for the involved equations. For this reason, we propose a graphical approach in Figure 1.

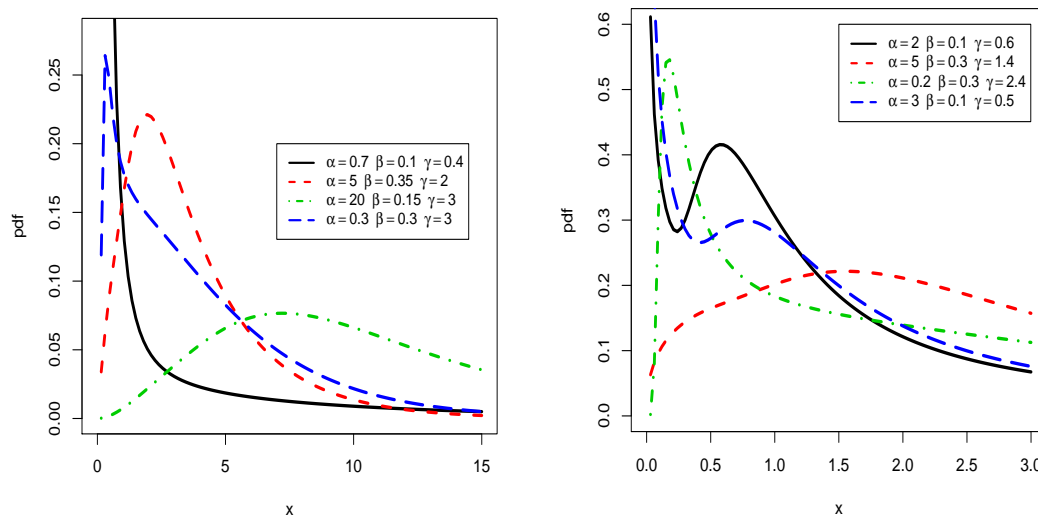


Figure 1. Plots for the pdf of the EMIEEdistribution, for several values of the parameters.

From Figure 1, we see that  $f(x; \alpha, \beta, \gamma)$  has very versatile shape properties. In particular, one or two modes, reversed J shapes, several kinds of bathtub shapes, N shapes, abrupt spikes, plate shapes, and remarkable heaviness on the tails are observed, reaching some extreme situations in terms of modeling.

Let us now focus on the shape properties of  $h(x; \alpha, \beta, \gamma)$ . First of all, the critical points of  $h(x; \alpha, \beta, \gamma)$  can be obtained by solving the following equation:  $dh(x; \alpha, \beta, \gamma)/dx = 0$ , which is equivalent to:

$$\begin{aligned}
 & -\frac{\alpha}{x^2} - \beta - \beta \frac{\alpha e^{\alpha/x} - 2x(1 + e^{\alpha/x})}{\alpha + \beta x^2(1 + e^{\alpha/x})} - \frac{2}{x} + 2\alpha \frac{e^{\alpha/x}}{x^2(1 + e^{\alpha/x})} \\
 & + (\gamma - 1)e^{-\beta x} \frac{\alpha + \beta x^2(1 + e^{\alpha/x})}{x^2(1 + e^{\alpha/x})(1 + e^{-\alpha/x} - e^{-\beta x})} \\
 & + \gamma e^{-\beta x} \frac{[\alpha + \beta x^2(1 + e^{\alpha/x})] [e^{-\alpha/x} + 1 - e^{-\beta x}]^{\gamma-1}}{x^2(1 + e^{\alpha/x}) \{ [1 + e^{-\alpha/x}]^\gamma - [e^{-\alpha/x} + 1 - e^{-\beta x}]^\gamma \}} = 0.
 \end{aligned}$$

The complexity of this equation is an obstacle for providing exact analytical solutions; the number of solutions depends on the values of  $\alpha$ ,  $\beta$ , and  $\gamma$ , and no closed-form of them can be set, motivating the use of a graphical approach, as proposed later.

The asymptotes of  $h(x; \alpha, \beta, \gamma)$  are studied below. We have:

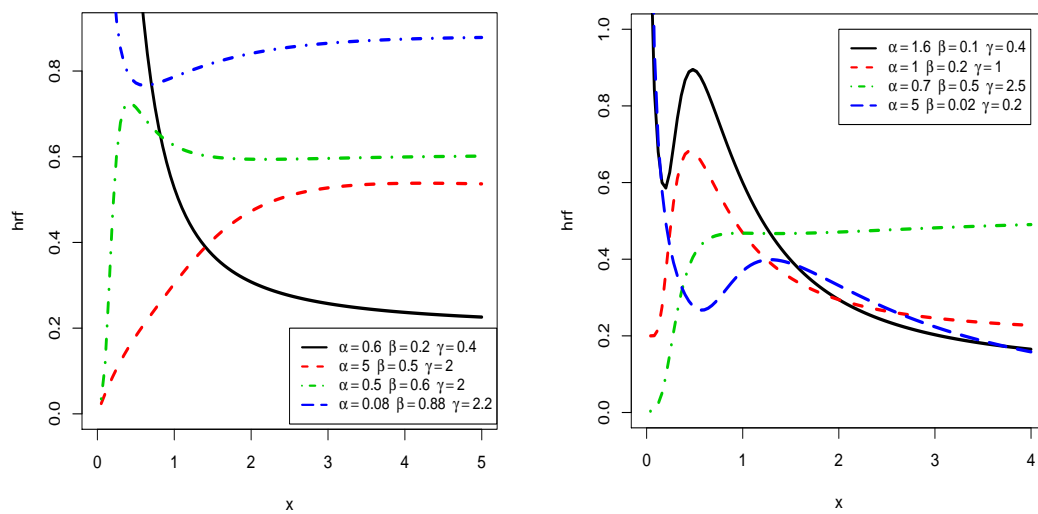
$$h(x; \alpha, \beta, \gamma) \sim \gamma \beta^\gamma x^{\gamma-1}, \quad x \rightarrow 0,$$

implying that  $h(x; \alpha, \beta, \gamma) \rightarrow +\infty$  if  $\gamma \in (0, 1)$ ,  $h(x; \alpha, \beta, \gamma) \rightarrow \beta$  if  $\gamma = 1$  and  $h(x; \alpha, \beta, \gamma) \rightarrow 0$  if  $\gamma > 1$ . Furthermore, we have:

$$h(x; \alpha, \beta, \gamma) \rightarrow \beta, \quad x \rightarrow +\infty.$$

As for  $f(x; \alpha, \beta, \gamma)$ , the parameter  $\alpha$  plays no role in the asymptotes of  $h(x; \alpha, \beta, \gamma)$ . Furthermore, the deep shape properties of  $h(x; \alpha, \beta, \gamma)$  are hard to present analytically. We thus complete our analysis by a graphical approach; some plots of  $h(x; \alpha, \beta, \gamma)$  are sketched in Figure 2.





**Figure 2.** Plots for the hazard rate function (hrf) of the EMIEE distribution, for several values of the parameters.

From Figure 2, we see that the hrf can be increasing, decreasing, with reserved J shapes, constant shapes, and N shapes. This wide panel of shapes indicates the great flexibility of the related distribution. In this regard, we may refer the reader to [21].

### 3.2. On Different Measures

The raw moments of the EMIEE distribution can be determined and computed by using (5) or (6), along with important probability and statistical measures. As an example, for a random variable  $X$  following the EMIEE distribution with parameters  $\alpha$ ,  $\beta$ , and  $\gamma$ , the  $k^{\text{th}}$  raw moment of  $X$  is given by  $\mu'_k = E(X^k)$ , corresponding to (5) with  $T(x) = x^k$ . Furthermore, from the raw moments, the following measures can be specified:

- the mean of  $X$  defined by  $\mu'_1$ , remaining the central parameter of the distribution,
- the variance of  $X$  given as  $\text{Var} = \mu'_2 - (\mu'_1)^2$ , providing a dispersion parameter,
- the standard deviation of  $X$  defined as  $\sigma = \text{Var}^{1/2}$ , corresponding to a dispersion parameter with the same unit as the mean,
- the skewness of  $X$  given by  $\text{SK} = [\mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3]/\sigma^3$ , measuring the lack of symmetry of tails of the EMIEE distribution (about the  $\mu'_1$ ),
- the kurtosis of  $X$  specified by  $\text{KU} = [\mu'_4 - 4\mu'_1\mu'_3 + 6(\mu'_1)^2\mu'_2 - 3(\mu'_1)^4]/\sigma^4$ , measuring how heavily the tails of the EMIEE distribution differ from those of a normal distribution,
- the coefficient of variation of  $X$  defined as  $\text{CV} = \sigma/\mu'_1$ , providing a dispersion parameter that can serve as a benchmark for comparison.

We refer the reader to the book of [17] for further details on these measures. A numerical treatment of these measures is proposed in Tables 1–3, for some selected values of the parameters.

Tables 1–3 show that the considered measures can take a wide range of values, with some increasing/decreasing tendencies depending on the increasing/decreasing tendencies of the values of the parameters. In particular, in Table 1, at  $\alpha = 0.5$  and  $\beta = 0.5$  and when the value of  $\gamma$  increases, then the values of SK, KU, and CV decrease, but the value of Var increases. Furthermore, from Table 2, at  $\gamma = 2.0$  and  $\beta = 0.5$  and when the value of  $\alpha$  increases, then the values of SK and KU increase, but the values of Var and CV decrease. Table 3 indicates that, at  $\gamma = 2.0$  and  $\alpha = 0.5$  when the value of  $\beta$  increases, the values of Var and CV decrease. The versatility of these measures is an additional quality of the EMIEE distribution.

**Table 1.** Numerical values of some moments, variance, skewness (SK), kurtosis (KU), and the coefficient of variation (CV) of the EMIEE distribution for some selected values of  $\gamma$  and at  $\alpha = 0.5$  and  $\beta = 0.5$ .

Measure	$\gamma = 1.5$	$\gamma = 2.0$	$\gamma = 2.5$	$\gamma = 3.0$	$\gamma = 3.5$	$\gamma = 4.0$	$\gamma = 4.5$
$\mu'_1$	1.709	1.921	2.051	2.344	2.601	2.830	3.037
$\mu'_2$	6.381	7.465	8.162	9.823	11.382	12.853	14.247
$\mu'_3$	37.486	44.483	49.068	60.272	71.130	81.674	91.927
$\mu'_4$	297.27	354.846	392.912	487.007	579.644	670.901	760.85
Var	3.459	3.775	3.956	4.329	4.616	4.842	5.024
SK	2.294	2.132	2.046	1.882	1.766	1.679	1.611
KU	10.633	9.646	9.155	8.281	7.711	7.315	7.026
CV	1.088	1.012	0.970	0.888	0.826	0.777	0.738

**Table 2.** Numerical values of some moments, variance, SK, KU, and CV of the EMIEE distribution for some selected values of  $\alpha$  and at  $\gamma = 2.0$  and  $\beta = 0.5$ .

Measure	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.75$
$\mu'_1$	1.846	1.910	1.963	2.009	2.088	2.123	2.139
$\mu'_2$	7.645	7.782	7.913	8.040	8.280	8.395	8.451
$\mu'_3$	47.023	47.542	48.055	48.564	49.567	50.061	50.306
$\mu'_4$	380.997	383.988	386.972	389.947	395.867	398.811	400.279
Var	4.238	4.134	4.060	4.002	3.919	3.887	3.874
SK	1.980	2.009	2.027	2.039	2.051	2.053	2.053
KU	8.646	8.846	8.982	9.081	9.209	9.249	9.264
CV	1.115	1.065	1.026	0.996	0.948	0.929	0.920

**Table 3.** Numerical values of some moments, variance, SK, KU, and CV of the EMIEE distribution for some selected values of  $\beta$  and at  $\alpha = 0.5$  and  $\gamma = 2.0$ .

Measure	$\beta = 0.1$	$\beta = 0.4$	$\beta = 0.7$	$\beta = 1.0$	$\beta = 1.2$	$\beta = 1.5$	$\beta = 1.8$
$\mu'_1$	9.229	2.512	1.516	1.107	0.944	0.777	0.664
$\mu'_2$	191.116	12.562	4.283	2.180	1.549	1.023	0.730
$\mu'_3$	5878	94.852	18.244	6.439	3.794	1.993	1.182
$\mu'_4$	238100	952.018	103.814	25.473	12.458	5.209	2.562
Var	105.938	6.254	1.983	0.955	0.658	0.419	0.290
SK	1.980	2.039	2.053	2.049	2.040	2.023	2.001
KU	8.646	9.081	9.249	9.304	9.304	9.266	9.198
CV	1.115	0.996	0.929	0.883	0.860	0.832	0.811

We complete this part by discussing the entropy of the EMIEE distribution through the Rényi entropy defined by  $I_\delta = (1 - \delta)^{-1} \log [E\{[f(X; \alpha, \beta, \gamma)]^{\delta-1}\}]$ , with  $\delta > 0, \delta \neq 1$  and  $(\delta - 1)(1 - \gamma) < 1$  (this inequality is an additional condition to ensure the existence of  $I_\delta$ ). Therefore, one can express the main term via (5) or (6) by taking  $T(x) = [f(x; \alpha, \beta, \gamma)]^{\delta-1}$ . A numerical study of the Rényi entropy is performed in Table 4.

Table 4 shows the versatile nature of the Rényi entropy of  $X$ ; it can be positive or negative, with varying values. In some sense, this shows the flexibility of the amount of randomness of the EMIEE distribution. Further details about the Rényi entropy, and the general concept of entropy, can be found in [22].

**Table 4.** Numerical values of the Rényi entropy of the EMIEE distribution at different values of  $\alpha$ ,  $\beta$ , and  $\gamma$ .

Parameters			Rényi Entropy		
$\alpha$	$\beta$	$\gamma$	$\delta = 1.2$	$\delta = 2$	$\delta = 3$
0.5	0.5	1	0.451	0.285	0.182
0.5	0.5	1.8	0.650	0.531	0.449
0.5	0.5	3	0.782	0.705	0.656
0.5	0.5	4	0.834	0.770	0.730
0.5	0.5	5	0.865	0.805	0.768
1.5	0.5	3	0.788	0.714	0.666
2	0.5	3	0.793	0.720	0.674
2.5	0.5	3	0.799	0.727	0.682
3	0.5	3	0.805	0.734	0.690
0.5	1.5	3	0.311	0.236	0.190
0.5	2	3	0.191	0.118	0.072
0.5	2.5	3	0.100	0.028	−0.017
0.5	3	3	0.026	−0.044	−0.088

#### 4. Parameter Estimation

Here, we derive the maximum likelihood estimates (MLEs) of the EMIEE model parameters, along with a simulation study to illustrate their practical interest. We recall that the MLEs have the following desirable properties. They are (i) efficient, (ii) consistent, (iii) asymptotically normal, and (iv) easy to handle in practice. For these, we refer the reader to [23]. The mathematical basis of this method in the setting of the EMIEE distribution is given below.

Let  $x_1, x_2, \dots, x_n$  be  $n$  independent realizations of a random variable following the EMIEE distribution with parameters  $\alpha, \beta$ , and  $\gamma$ . Then, the MLEs of  $\alpha, \beta$ , and  $\gamma$  are defined as the “argmax of the likelihood function with respect to  $\alpha, \beta$ , and  $\gamma$ ”. Thus, by denoting them as  $\hat{\alpha}, \hat{\beta}$ , and  $\hat{\gamma}$ , they are defined by:

$$(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \operatorname{argmax}_{(\alpha, \beta, \gamma) \in (0, +\infty)^3} L(\alpha, \beta, \gamma),$$

where, based on (8),  $L(\alpha, \beta, \gamma)$  denotes the likelihood function given as:

$$L(\alpha, \beta, \gamma) = \prod_{i=1}^n f(x_i; \alpha, \beta, \gamma) = \prod_{i=1}^n \left\{ \gamma e^{\alpha/x_i - \beta x_i} \frac{\alpha + \beta x_i^2 (1 + e^{\alpha/x_i})}{x_i^2 (1 + e^{\alpha/x_i})^2} \left[ 1 - \frac{e^{-\beta x_i}}{1 + e^{-\alpha/x_i}} \right]^{\gamma-1} \right\}.$$

To avoid the complicated product form of the likelihood function, one can also define the MLEs as  $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \operatorname{argmax}_{(\alpha, \beta, \gamma) \in (0, +\infty)^3} \ell(\alpha, \beta, \gamma)$ , where  $\ell(\alpha, \beta, \gamma)$  denotes the log-likelihood function given by:

$$\begin{aligned} \ell(\alpha, \beta, \gamma) = \log[L(\alpha, \beta, \gamma)] &= n \log \gamma + \alpha \sum_{i=1}^n \frac{1}{x_i} - \beta \sum_{i=1}^n x_i + \sum_{i=1}^n \log \left[ \alpha + \beta x_i^2 (1 + e^{\alpha/x_i}) \right] \\ &- 2 \sum_{i=1}^n \log x_i - 2 \sum_{i=1}^n \log(1 + e^{\alpha/x_i}) + (\gamma - 1) \sum_{i=1}^n \log(1 + e^{-\alpha/x_i} - e^{-\beta x_i}) \\ &- (\gamma - 1) \sum_{i=1}^n \log(1 + e^{-\alpha/x_i}). \end{aligned}$$

Thus,  $\hat{\alpha}, \hat{\beta}$ , and  $\hat{\gamma}$  satisfy  $\partial \ell(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) / \partial \alpha = 0$ ,  $\partial \ell(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) / \partial \beta = 0$ , and  $\partial \ell(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) / \partial \gamma = 0$ , whose extended forms are the following:

$$\begin{aligned} & \sum_{i=1}^n \frac{1}{x_i} + \sum_{i=1}^n \frac{1 + \hat{\beta}x_i e^{\hat{\alpha}/x_i}}{\hat{\alpha} + \hat{\beta}x_i^2(1 + e^{\hat{\alpha}/x_i})} - 2 \sum_{i=1}^n \frac{e^{\hat{\alpha}/x_i}}{x_i(1 + e^{\hat{\alpha}/x_i})} - (\hat{\gamma} - 1) \sum_{i=1}^n \frac{e^{-\hat{\alpha}/x_i}}{x_i(1 + e^{-\hat{\alpha}/x_i} - e^{-\hat{\beta}x_i})} \\ & + (\hat{\gamma} - 1) \sum_{i=1}^n \frac{e^{-\hat{\alpha}/x_i}}{x_i(1 + e^{-\hat{\alpha}/x_i})} = 0, \\ & - \sum_{i=1}^n x_i + \sum_{i=1}^n \frac{x_i^2(1 + e^{\hat{\alpha}/x_i})}{\hat{\alpha} + \hat{\beta}x_i^2(1 + e^{\hat{\alpha}/x_i})} + (\hat{\gamma} - 1) \sum_{i=1}^n \frac{x_i e^{-\hat{\beta}x_i}}{1 + e^{-\hat{\alpha}/x_i} - e^{-\hat{\beta}x_i}} = 0 \end{aligned}$$

and:

$$\frac{n}{\hat{\gamma}} + \sum_{i=1}^n \log(1 + e^{-\hat{\alpha}/x_i} - e^{-\hat{\beta}x_i}) - \sum_{i=1}^n \log(1 + e^{-\hat{\alpha}/x_i}) = 0.$$

From this last equation, one can express  $\hat{\gamma}$  according to  $\hat{\alpha}$  and  $\hat{\beta}$  as:

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

(and plugging this expression into the first two equations, now depending on  $\hat{\alpha}$  and  $\hat{\beta}$  only). Since the above equations are complicated to solve analytically, the MLEs have no tractable expression. However, they can be approached by standard optimization algorithms, such as the Newton–Raphson or quasi-Newton Broyden–Fletcher–Goldfarb–Shannon (BFGS) algorithms, with the use of a statistical software. By expressing the second partial derivatives of the log-likelihood function with respect to  $\alpha$ ,  $\beta$ , and  $\gamma$ , we can determine the observed Fisher information matrix, allowing us to obtain the asymptotic variances, covariances, and standard errors (SEs) of the ML estimators for  $\alpha$ ,  $\beta$ , and  $\gamma$ , among others.

Now, let us illustrate the practical aspect of the MLEs by a simulation study, with the use of the R software (see [24]). The BFGS algorithm is considered. That is, we generated  $N = 10,000$  replications of samples  $(x_1, x_2, \dots, x_n)$  with  $n \in \{50, 100, 200, 500, 1000\}$  from a random variable following the EMIEE distribution defined with the six following sets of parameters, in turn: Y1( $\alpha = 1.2, \beta = 1.5, \gamma = 2.0$ ), Y2( $\alpha = 1.8, \beta = 1.5, \gamma = 2.0$ ), Y3( $\alpha = 2.5, \beta = 1.5, \gamma = 2.0$ ), Y4( $\alpha = 1.2, \beta = 2.0, \gamma = 2.0$ ), Y5( $\alpha = 1.8, \beta = 1.5, \gamma = 4.0$ ), and Y6( $\alpha = 3.0, \beta = 1.5, \gamma = 4.0$ ). Then, for each of these sets, we determined the average MLEs of the parameters defined as, for  $\omega = \alpha, \beta, \gamma$ ,

$$MLE_{\omega} = \frac{1}{N} \sum_{k=1}^N \hat{\omega}_k,$$

where  $\hat{\omega}_k$  denotes the MLE of  $\omega$  obtained at the  $k$ th replication, and the corresponding empirical mean squared errors (MSEs) defined as, for  $\omega = \alpha, \beta$  and  $\gamma$ ,

$$MSE_{\omega} = \frac{1}{N} \sum_{k=1}^N (\hat{\omega}_k - \omega)^2.$$

The results of this simulation study are in Tables 5 and 6.

The results in Tables 5 and 6 show the numerical efficiency of the maximum likelihood method for the EMIEE model. Indeed, we see that the MLEs were relatively close to the true values of the parameters, and globally, the MSEs decreased as  $n$  increased. This “numerical convergence” illustrated the well-known theoretical convergence properties of the MLEs.

**Table 5.** Simulation study for the EMIEE model: MLEs and MSEs with the following sets of parameters:  $Y1(\alpha = 1.2, \beta = 1.5, \gamma = 2.0)$ ,  $Y2(\alpha = 1.8, \beta = 1.5, \gamma = 2.0)$ , and  $Y3(\alpha = 2.5, \beta = 1.5, \gamma = 2.0)$ .

<i>n</i>	Y1		Y2		Y3	
	MLE	MLE	MLE	MLE	MLE	MLE
50	1.517	1.476	1.962	0.920	2.784	4.450
	1.712	0.176	1.541	0.057	1.622	0.112
	2.498	1.907	2.051	0.181	2.423	1.019
100	1.259	0.257	1.880	0.831	1.866	1.769
	1.568	0.069	1.628	0.045	1.483	0.019
	2.124	0.130	2.241	0.104	2.380	0.436
200	1.363	0.095	2.081	0.613	2.843	1.709
	1.583	0.014	1.550	0.026	1.497	0.012
	2.056	0.017	1.992	0.023	2.013	0.027
500	1.224	0.059	1.697	0.333	2.779	1.121
	1.532	0.013	1.463	0.008	1.504	0.010
	2.083	0.045	2.009	0.011	2.031	0.025
1000	1.131	0.029	1.825	0.262	2.552	0.461
	1.484	0.005	1.484	0.006	1.487	0.002
	2.003	0.016	2.007	0.010	1.978	0.003

**Table 6.** Simulation study for the EMIEE model: MLEs and MSEs with the following sets of parameters:  $Y4(\alpha = 1.2, \beta = 2.0, \gamma = 2.0)$ ,  $Y5(\alpha = 1.8, \beta = 1.5, \gamma = 4.0)$ , and  $Y6(\alpha = 3.0, \beta = 1.5, \gamma = 4.0)$ .

<i>n</i>	Y4		Y5		Y6	
	MLE	MLE	MLE	MLE	MLE	MLE
50	1.118	0.567	2.060	1.969	3.239	0.315
	2.016	0.134	1.774	0.165	2.082	0.521
	2.446	1.021	5.169	3.634	5.419	4.331
100	1.180	0.198	1.452	0.181	2.895	0.198
	1.931	0.127	1.517	0.034	1.800	0.172
	2.056	0.153	4.171	0.475	5.045	3.873
200	1.226	0.147	1.394	0.177	2.812	0.052
	1.990	0.042	1.550	0.020	1.722	0.086
	2.028	0.043	4.353	0.335	4.660	0.649
500	1.299	0.141	1.400	0.173	2.774	0.074
	2.018	0.014	1.509	0.012	1.736	0.069
	2.113	0.032	4.197	0.206	4.434	0.300
1000	1.177	0.050	1.669	0.093	2.781	0.056
	2.029	0.016	1.511	0.004	1.732	0.055
	2.044	0.009	4.200	0.077	4.289	0.222

### 5. Application to a COVID-19 Dataset

Here, we propose a concrete application with an actual dataset to assess the interest in the EMIEE model. The considered data, called the COVID-19 dataset, is presented below.

COVID-19, which can be renamed as the “the flu of 2020”, is due to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Sadly, it spread quickly in the beginning of the year 2020, taking thousands of victims, obliging governments to take exceptional measures to protect their people. The update as of 29 May 2020 situation of this pandemic tragedy can be found in [25–27]. Naturally, the overall comprehension of COVID-19 is a challenge for all scientists, but necessary for the sake of future generations. In this section, we modestly contribute to the subject by applying the EMIEE model to fit data of daily new COVID-19 confirmed cases in Pakistan from 21 March to 29 May 2020 (inclusive), showing that it is very efficient in this regard. We thus assumed that the new COVID-19

(confirmed) cases in Pakistan could be modeled by a continuous variable (since a discrete variable with a wide range of values could be considered as such) and provided a new statistical model that could be relevant for the following points:

- (i) Provide a precise estimation for some measures of interest related to COVID-19 cases in Pakistan (mean of cases, probability to have a certain number of cases, and so on),
- (ii) Compare the repartitions of the number of COVID-19 cases in Pakistan with those in other countries,
- (iii) Propose an efficient strategy for fitting data on COVID-19 cases in other countries,
- (iv) In a more challenging way, model the distribution of the number of cases for any pandemic with similar features and under a similar environment (with comparable populations, comparable climate, sanitary system, etc.).

The dataset was obtained from the following electronic address: <http://covid.gov.pk/stats/pakistan>. It is given as follows: {112, 157, 89, 108, 102, 133, 170, 121, 99, 236, 178, 250, 161, 258, 172, 407, 577, 210, 243, 281, 186, 254, 336, 342, 269, 543, 488, 463, 514, 427, 796, 555, 742, 642, 785, 783, 605, 751, 806, 942, 990, 1297, 989, 1083, 1315, 1049, 1523, 1764, 1637, 1991, 1476, 1140, 2255, 1452, 1430, 1581, 1352, 1974, 1841, 1932, 2193, 2603, 1743, 2164, 1748, 1356, 1446, 2241, 2636, 2429} corresponding to the dates {21 March 2020, 22 March 2020, . . . , 29 May 2020}, respectively.

Aiming to identify the possible shapes of the unknown hrf behind these data, we plot the total time on test (TTT) plot in Figure 3 (see [28] for further details on the use of TTT plots in data analysis).

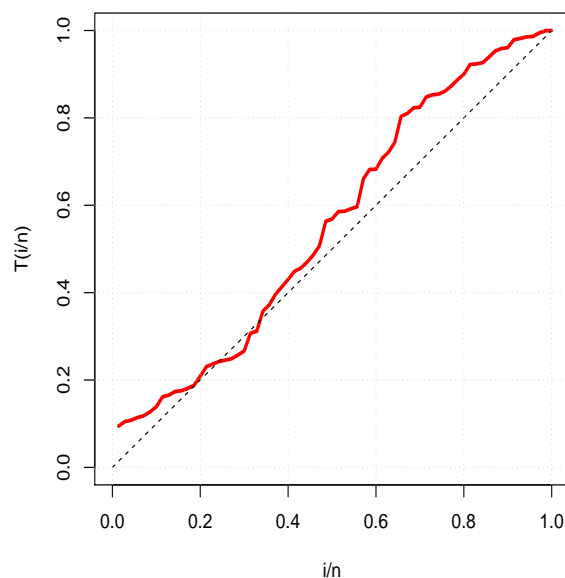


Figure 3. Total time on test (TTT) plot of the COVID-19 dataset.

In Figure 3, since the red line is convex, then concave, the unknown hrf probably presents a bathtub shape. Therefore, the EMIEE distribution is appropriate to fit the data.

Now, we aimed to compare the fitness of the EMIEE model with the one of 15 top ranked models in the literature: (i) the Weibull-exponential (WE) model by [29], (ii) the Lomax-exponential (LE) model by [30], (iii) the gamma-exponentiated exponential (GaE) model by [31], (iv) the beta Weibull (BW) model by [32], (v) the Kumaraswamy exponential (KE) model by [33], (vi) the Burr X-exponential (BXE) model by [34], (vii) the exponentiated exponential (EE) model by [35], (viii) the CSttransformation of exponential (CE) model by [36], (ix) the standard exponential (E) model (see [37], among others), (x) the alpha-power inverse Weibull (AIW) model by [38], (xi) the Gompertz inverse exponential (GomIE) model by [39], (xii) the Weibull-inverse exponential (WIE) model by [40], (xiii) the inverse Weibull-inverse exponential (IWIE) model by [41], (xiv) the inverse exponential (IE) model

by [20], and last, but not least, (xv) the “unexponentiated” version of the proposed EMIEE model, i.e., the MIEE model. We refer to the above references for the precise definitions of the related cdfs and pdfs, along with the Greek alphabet letters used for the parameters.

Then, the model parameters were estimated through the practice of the maximum likelihood method (with the BFGS algorithm). The R software was used in this regard. The calculations of the MLEs and SEs for all the model parameters are provided in Table 7.

**Table 7.** MLEs and standard errors (SEs) (under parentheses) of the model parameters for the COVID-19 dataset: Weibull-exponential (WE) model, the Lomax-exponential (LE) model, the gamma-exponentiated exponential (GaE) model, the beta Weibull (BW) model, the Kumaraswamy exponential (KE) model, the Burr X-exponential (BXE) model, the exponentiated exponential (EE) model, the CStransformation of exponential (CE) model, the standard exponential (E) model, the alpha-power inverse Weibull (AIW) model, the Gompertz inverse exponential (GomIE) model, the Weibull-inverse exponential (WIE) model, the inverse Weibull-inverse exponential (IWIE) model, the inverse exponential (IE) model.

Model	$\alpha$	$\beta$	$\gamma$	$\lambda$	$\theta$	$\eta$
EMIEE	1.5577 (0.5860)	0.1184 (0.0188)	2.4646 (0.5511)	- -	- -	- -
WE	- -	- -	- -	1.2850 (1.3391)	0.9664 (0.1893)	0.0528 (0.0387)
LE	44.7999 (13.9974)	0.0477 (0.0199)	67.3943 (1.2461)	- -	- -	- -
GaE	- -	- -	- -	0.0822 (0.0450)	0.7317 (0.3352)	1.2936 (0.5946)
BW	- -	- -	1.6505 (0.0083)	0.9444 (0.0069)	3.4276 (1.2777)	0.0867 (0.0109)
KE	- -	- -	- -	1.1891 (0.0191)	2.9329 (0.1555)	0.0982 (0.0744)
BXE	0.0391 (0.0027)	0.3782 (0.0489)	- -	- -	- -	- -
EE	1.3633 (0.2239)	0.1284 (0.0184)	- -	- -	- -	- -
CE	16.9399 (2.27821)	16.6595 (3.1211)	8.1926 (2.116)	- -	- -	- -
E	- -	0.1061 (0.0126)	- -	- -	- -	- -
AIW	7.3531 (1.3953)	2.6650 (0.8931)	1.1892 (0.1124)	- -	- -	- -
GomIE	0.2035 (0.1355)	1.0786 (0.1925)	1.9361 (0.8576)	- -	- -	- -
WIE	0.2472 (0.2197)	0.9974 (0.1553)	2.0841 (1.2271)	- -	- -	- -
IWIE	0.3554 (0.5157)	0.9506 (0.1172)	9.7728 (5.2750)	- -	- -	- -
IE	3.7629 (0.4497)	- -	- -	- -	- -	- -
MIEE	4.3601 (1.11734)	0.0741 (0.0123)	- -	- -	- -	- -



Among the information provided by Table 7, the parameters of the EMIEE model, i.e.,  $\alpha$ ,  $\beta$ , and  $\gamma$ , are estimated as:

$$\hat{\alpha} = 1.5577, \quad \hat{\beta} = 0.1184, \quad \hat{\gamma} = 2.4646.$$

Therefore, based on (8), the corresponding estimated pdf is given by:

$$\hat{f}(x) = f(x; \hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \hat{\gamma} e^{\hat{\alpha}/x - \hat{\beta}x} \frac{\hat{\alpha} + \hat{\beta}x^2(1 + e^{\hat{\alpha}/x})}{x^2(1 + e^{\hat{\alpha}/x})^2} \left[ 1 - \frac{e^{-\hat{\beta}x}}{1 + e^{-\hat{\alpha}/x}} \right]^{\hat{\gamma}-1}. \tag{10}$$

Thus,  $\hat{f}(x)$  is an estimated function of the unobservable underlying pdf of the number of COVID-19 cases in Pakistan. By the use of this function, one can estimate the quantities of interest. Some basics of them are presented below. By denoting  $X$  the random variable modeling the daily COVID-19 confirmed cases in Pakistan during the epidemic, the probability that  $X$  belongs to a chosen interval, say  $[a, b]$ , can be estimated by  $\hat{p}_{a,b} = \int_a^b \hat{f}(x)dx$ . For instance, the probability that the COVID-19 cases in Pakistan are less than a certain values  $c$  is given by  $\hat{p}_{0,c}$ . More generally, an estimation of the mean of a certain transformation of  $X$ , say  $T(X)$ , can be estimated by  $\hat{\mu}_* = \int_0^{+\infty} T(x)\hat{f}(x)dx$ . For instance, the average number of COVID-19 cases in Pakistan can be approximated with precision by  $\hat{\mu}_*$  by taking  $T(x) = x$ , and so on.

As planned, a comparison of the models in terms of fitting was performed. We decided which was the best model by determining the values of the following statistical measures: minus complete log-likelihood function ( $-\hat{\ell}$ ), Akaike information criterion (AIC), Bayesian information criterion (BIC), Cramer–von Mises (W) criterion, and Anderson–Darling (A) criterion. Furthermore, we considered the value of the Kolmogorov–Smirnov (KS) statistic and its  $p$ -value. The best model was the one having the smallest  $-\hat{\ell}$ , AIC, BIC, W, A, and KS and the largest KS  $p$ -value. For the considered data, the obtained values are shown in Table 8.

**Table 8.** Values of  $-\hat{\ell}$ , AIC, BIC, W, Anderson–Darling (A) criterion, KS, and KS  $p$ -value of the 16 considered models for the COVID-19 dataset.

Model	$-\hat{\ell}$	AIC	BIC	W	A	KS	KS $p$ -Value
EMIEE	221.3346	448.6692	455.4147	0.1228	0.8148	0.0991	0.5679
WE	223.8891	453.7781	460.5236	0.1393	0.9449	0.1036	0.4122
LE	223.9463	453.8925	460.6380	0.1399	0.9494	0.1060	0.3838
GaE	223.8191	453.6381	460.3836	0.1396	0.9447	0.1042	0.4050
BW	223.8808	455.7615	464.7555	0.1629	1.0745	0.1172	0.2695
KE	224.0199	454.0399	460.7854	0.1577	1.0497	0.1130	0.3088
BXE	224.5807	453.1613	457.6583	0.1552	1.0425	0.1040	0.4074
EE	225.2817	454.5633	459.0603	0.1632	1.1044	0.1085	0.3563
CE	228.1471	462.2942	469.0396	0.1240	0.8424	0.1392	0.1203
E	226.9804	455.9608	458.2093	0.1608	1.0887	0.1048	0.3979
AIW	231.7256	469.4511	476.1966	0.3406	2.0919	0.12049	0.2413
GomIE	223.1276	452.2553	459.0007	0.1454	0.9659	0.1107	0.3327
WIE	223.2084	452.4168	459.1623	0.1518	1.0024	0.1134	0.3049
IWIE	233.9508	473.9016	480.6471	0.4037	2.4377	0.1367	0.1327
IE	233.4393	468.8786	471.1271	0.3844	2.3318	0.1330	0.1532
MIEE	228.6196	461.2392	465.7362	0.2013	1.2916	0.1577	0.0548

From Table 8, we see that the EMIEE model was the best among all the considered models, with the following numerical criteria:  $-\hat{\ell} = 221.3346$ ,  $AIC = 448.6692$ ,  $BIC = 455.4147$ ,  $W = 0.1228$ ,

$A = 0.8148$ ,  $KS = 0.0991$ , and  $KS p$ -value = 0.5679. One can notice that the EMIEE model outperformed the baseline E and IE models, and also, the MIEE model was derived from the former M family, validating the use of the exponentiated transform for fitting purposes.

Figure 4 shows the estimated pdf as described in (10) over the histogram of the data. Figure 5 presents the estimated cdf, i.e., based on (7),  $\hat{F}(x) = F(x; \hat{\alpha}, \hat{\beta}, \hat{\gamma})$ , over the empirical cdf of the data.

The probability-probability (P-P) plot in Figure 6 shows how closely the estimated and empirical cdfs agreed.

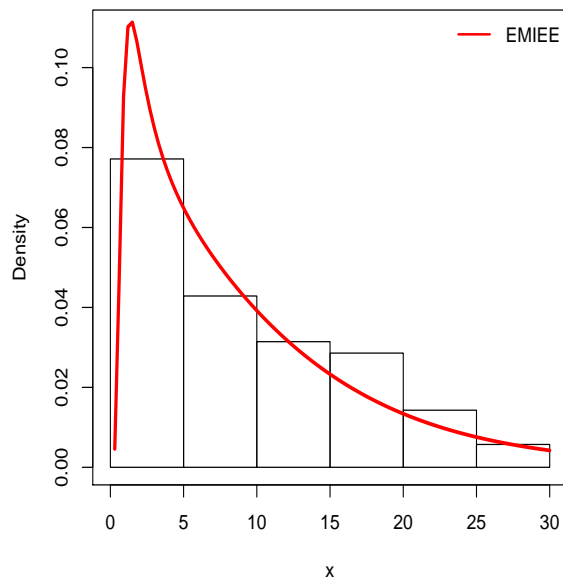


Figure 4. Estimated pdf of the EMIEE model over the histogram for the COVID-19 dataset.

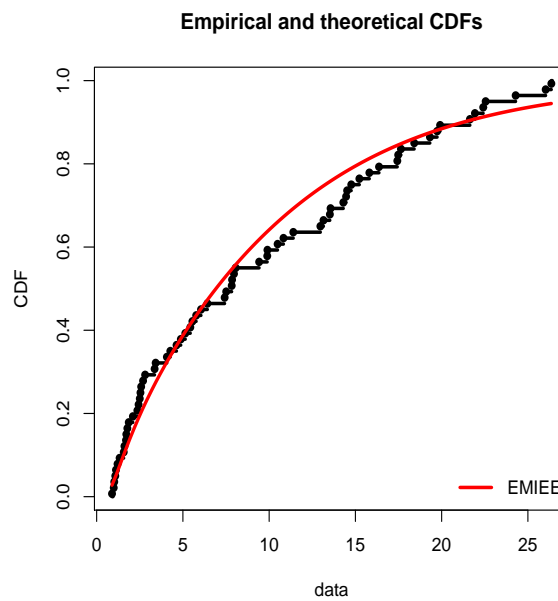


Figure 5. Estimated cdf of the EMIEE model over the histogram for the COVID-19 dataset.

In all the graphics, we see that the red curves fit perfectly the black data objects, motivating the importance of the EMIEE model in the analysis of the COVID-19 dataset. We end this application by displaying the estimated hrf of the EMIEE model in Figure 7.

We see that the estimated hrf has a bathtub shape, which was in coherence with what was interpreted in Figure 3.

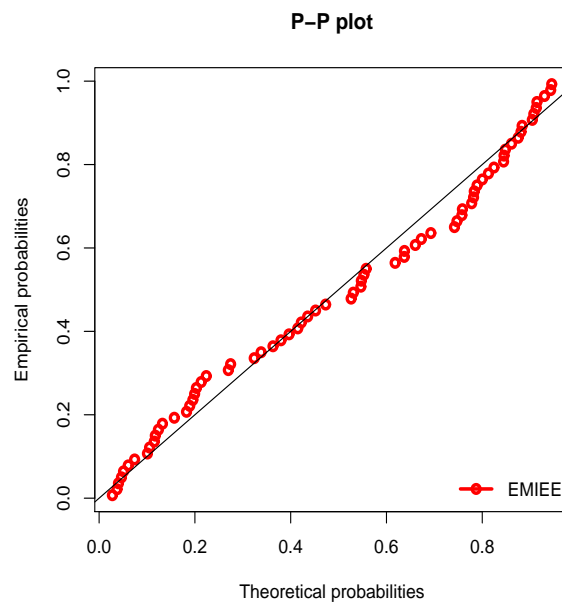


Figure 6. Probability-probability (P-P) plot of the EMIEE model for the COVID-19 dataset.

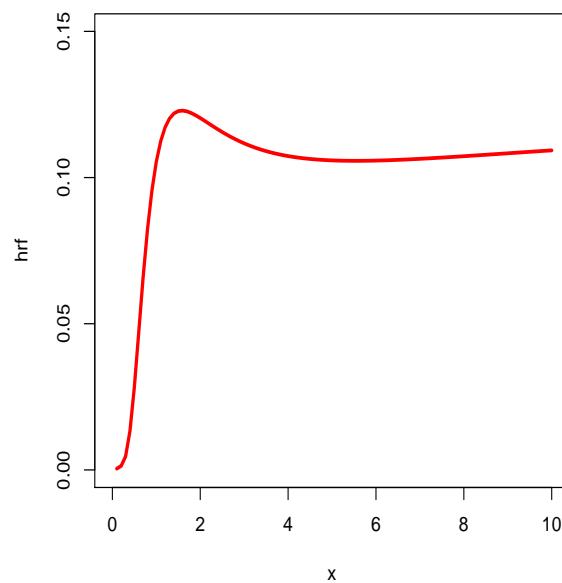


Figure 7. Estimated hrf of the EMIEE model for the COVID-19 dataset.

### 6. Conclusions

In this paper, we derived a natural extension of the M family, called the exponentiated M (EM) family. We investigated its main mathematical properties and discussed its ability in terms of statistical modeling. Light was shed on a new promising distribution of the EM family, based on the inverse exponential and exponential distributions. It was called the EM inverse exponential exponential (EMIEE) distribution. We investigated the estimation of the EMIEE model parameters by a reputed method: the maximum likelihood method. We applied it to analyze new COVID-19 cases in Pakistan during 21 March to 29 May 2020 (inclusive), with fair comparisons with 15 other solid models. The fitting results were quite favorable to the EMIEE model. That is, the EMIEE model could be used for similar analyses in other countries, allowing comparisons in this regard and, consequently, a better understanding of the COVID-19 pandemic.

**Author Contributions:** Investigation, R.A.R.B., C.C., F.J., and M.E. All authors contributed equally to this work. All authors read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under Grant No. (RG-2-150-37).

**Acknowledgments:** We warmly thank the three reviewers for their thorough and constructive comments. This project was funded by the Deanship of Scientific Research (DSR), at King Abdulaziz University, Jeddah, under Grant No. (RG-2-150-37). The authors, therefore, acknowledge with thanks the DSR's technical and financial support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Azzalini, A. A class of distributions which includes the normal ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
2. Marshall, A.; Olkin, I. A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika* **1997**, *84*, 641–652. [[CrossRef](#)]
3. Gupta, R.D.; Kundu, D. Exponentiated exponential family: An alternative to Gamma and Weibull distributions. *Biom. J.* **2001**, *43*, 117–130. [[CrossRef](#)]
4. Eugene, N.; Lee, C.; Famoye, F. Beta-normal distribution and its applications. *Commun. Statist. Theory Methods* **2002**, *31*, 497–512. [[CrossRef](#)]
5. Jones, M. Families of distributions arising from distributions of order statistics. *Test* **2004**, *13*, 1–43. [[CrossRef](#)]
6. Jones, M.; Pewsey, A. Sinh-arcsinh distributions. *Biometrika* **2009**, *96*, 761–780. [[CrossRef](#)]
7. Shaw, W.T.; Buckley, I.R. The Alchemy of Probability Distributions: Beyond Gram-Charlier Expansions, and a Skew-kurtotic-normal Distribution from a Rank Transmutation Map. *arXiv* **2009**, arXiv:0901.0434.
8. Zografos, K.; Balakrishnan, N. On families of beta- and generalized gamma-generated distributions and associated inference. *Stat. Methodol.* **2009**, *6*, 344–362. [[CrossRef](#)]
9. Cordeiro, G.M.; de Castro, M. A new family of generalized distributions. *J. Stat. Comput. Simul.* **2011**, *81*, 883–893. [[CrossRef](#)]
10. Al-Shomrani, A.; Arif, O.; Shawky, A.; Hanif, S.; Shahbaz M.Q. Topp-Leone family of distributions: Some properties and application. *Pak. J. Stat. Oper. Res.* **2016**, *12*, 443–451. [[CrossRef](#)]
11. Bantan, R.A.R.; Jamal, F.; Chesneau, C.; Elgarhy, M. On a new result on the ratio exponentiated general family of distributions with applications. *Mathematics* **2020**, *8*, 598. [[CrossRef](#)]
12. Kumar, D.; Singh, U.; Singh, S.K.; Mukherjee, S. The new probability distribution: An aspect to a Life time distribution. *Math. Sci. Lett.* **2017**, *6*, 35–42. [[CrossRef](#)]
13. Linhart, H.; Zucchini, W. *Model Selection*; John Wiley and Sons: New York, NY, USA, 1986.
14. Tahir, M.H.; Cordeiro, G.M.; Alizadeh, M.; Mansoor, M.; Zubair, M.; Hamedani, G.G. The odd generalized exponential family of distributions with applications. *J. Stat. Distrib. Appl.* **2015**, *2*, 1–28. [[CrossRef](#)]
15. Merovci, F.; Alizadeh, M.; Yousof, H.M.; Hamedani, G.G. The exponentiated transmuted-G family of distributions: Theory and applications. *Commun. Stat. Theory Methods* **2017**, *46*, 10800–10822. [[CrossRef](#)]
16. Aldahlan, M.A.; Jamal, F.; Chesneau, C.; Elbatal, I.; Elgarhy, M. Exponentiated power generalized Weibull power series family of distributions: Properties, estimation and applications. *PLoS ONE* **2020**, *15*, 1–25. [[CrossRef](#)]
17. Cordeiro, G.M.; Silva, R.B.; Nascimento, A.D.C. *Recent Advances in Lifetime and Reliability Models*; Bentham Books: Sharjah, UAE, 2020; doi:10.2174/97816810834521200101. [[CrossRef](#)]
18. Klein, J.P.; Moeschberger, M.L. *Survival Analysis: Techniques for Censored and Truncated Data*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2003.
19. Shaked, M.; Shanthikumar, J.G. *Stochastic Orders*; Springer: New York, NY, USA, 2007.
20. Keller, A.Z.; Kamath, A.R. Reliability analysis of CNC Machine Tools. *Reliab. Eng.* **1982**, *3*, 449–473. [[CrossRef](#)]
21. Glaser, R.E. Bathtub and related failure rate characterization. *J. Am. Stat. Assoc.* **1980**, *75*, 667–672. [[CrossRef](#)]
22. Amigo, J.M.; Balogh, S.G.; Hernandez, S. A brief review of generalized entropies. *Entropy* **2018**, *20*, 813. [[CrossRef](#)]
23. Casella, G.; Berger, R.L. *Statistical Inference*; Duxbury Advanced Series; Thomson Learning: Pacific Grove, CA, USA, 2002.

24. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2013. Available online: <http://www.R-project.org/> (accessed on 1 May 2020).
25. CDCP. Centers for Disease Control and Prevention. Coronavirus Symptoms and Diagnosis. 2020. Available online: <https://www.cdc.gov/coronavirus/2019-ncov/symptomstesting/symptoms.html> (accessed on 1 May 2020).
26. WHO. WHO Director-General's Opening Remarks at the Media Briefing on COVID-19—11 March 2020. 2020. Available online: <https://www.who.int/dg/speeches/detail/who-director-generals-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> (accessed on 1 May 2020).
27. WHO. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). 2020. Available online: <https://www.who.int/docs/default-source/coronaviruse/who-chinajoint-mission-on-covid-19-final-report.pdf> (accessed on 1 May 2020).
28. Aarset, M.V. How to identify bathtub hazard rate. *IEEE Trans. Reliab.* **1987**, *36*, 106–108. [[CrossRef](#)]
29. Oguntunde, P.E.; Balogun, O.S.; Okagbue, H.I.; Bishop, S.A. The Weibull-exponential distribution: Its properties and applications. *J. Appl. Sci.* **2015**, *15*, 1305–1311. [[CrossRef](#)]
30. Golzar, N.H.; Ganji, M.; Bevrani, H. The Lomax-exponential distribution, some properties and applications. *J. Statist. Res. Iran* **2016**, *13*, 131–153. [[CrossRef](#)]
31. Ristic, M.M.; Balakrishnan, N. The gamma-exponentiated exponential distribution. *J. Stat. Comput. Simul.* **2012**, *82*, 1191–1206. [[CrossRef](#)]
32. Lee, C.; Famoye, F.; Olumolade, O. Beta-Weibull distribution: Some properties and applications to censored data. *J. Mod. Appl. Stat. Methods* **2007**, *6*, 173–186. [[CrossRef](#)]
33. Rodrigues, J.A.; Silva, A.P. The exponentiated Kumaraswamy-exponential distribution. *Br. J. Appl. Sci. Technol.* **2015**, *10*, 1–12. [[CrossRef](#)]
34. Oguntunde, P.E.; Adejumo, A.O.; Owoloko, E.A.; Rastogi, M.K.; Odetunmbi, O.A. The Burr X-exponential distribution: Theory and applications. In Proceedings of the World Congress on Engineering 2017 Vol I WCE 2017, London, UK, 5–7 July 2017.
35. Nadarajah, S. The exponentiated exponential distribution: A survey. *AStA Adv. Stat. Anal.* **2011**, *95*, 219–251. [[CrossRef](#)]
36. Chesneau, C.; Bakouch, H.; Hussain, T. A new class of probability distributions via cosine and sine functions with applications. *Commun. Stat. Simul. Comput.* **2019**, *48*, 2287–2300. [[CrossRef](#)]
37. Balakrishnan, N.; Basu, A.P. *The Exponential Distribution: Theory, Methods and Applications*; Taylor and Francis: Philadelphia, PA, USA, 1995.
38. Ramadan, D.A.; Walaa, M.A. On the alpha-power inverse Weibull distribution. *Int. J. Comput. Appl.* **2018**, *181*, 6–12.
39. Oguntunde, P.E.; Khaleel, M.A.; Adejumo, A.O.; Okagbue, H.I.; Opanuga, A.A.; Owolabi, F.O. The Gompertz inverse exponential (GoIE) distribution with applications. *Cogent Math. Stat.* **2018**, *5*, 1507122. [[CrossRef](#)]
40. Oguntunde, P.E.; Adejumo, A.O.; Owoloko, E.A. The Weibull-inverted exponential distribution: A generalization of the inverse exponential distribution. In Proceedings of the World Congress on Engineering, London, UK, 5–7 July 2017; pp. 16–19.
41. Aldahlan, M.A. The inverse Weibull inverse exponential distribution with application. *Int. J. Contemp. Math. Sci.* **2019**, *14*, 17–30. [[CrossRef](#)]

