

## Article

# Survival and Reliability Analysis with an Epsilon-Positive Family of Distributions with Applications

Perla Celis <sup>1,\*</sup>, Rolando de la Cruz <sup>1</sup> , Claudio Fuentes <sup>2</sup> and Héctor W. Gómez <sup>3</sup> 

<sup>1</sup> Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Diagonal Las Torres 2640, Peñalolén, Santiago 7941169, Chile; rolando.delacruz@uai.cl

<sup>2</sup> Department of Statistics, Oregon State University, 217 Weniger Hall, Corvallis, OR 97331, USA; fuentes@oregonstate.edu

<sup>3</sup> Departamento de Matemática, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile; hector.gomez@uantof.cl

\* Correspondence: perla.celis@edu.uai.cl

**Abstract:** We introduce a new class of distributions called the epsilon-positive family, which can be viewed as generalization of the distributions with positive support. The construction of the epsilon-positive family is motivated by the ideas behind the generation of skew distributions using symmetric kernels. This new class of distributions has as special cases the exponential, Weibull, log-normal, log-logistic and gamma distributions, and it provides an alternative for analyzing reliability and survival data. An interesting feature of the epsilon-positive family is that it can be viewed as a finite scale mixture of positive distributions, facilitating the derivation and implementation of EM-type algorithms to obtain maximum likelihood estimates (MLE) with (un)censored data. We illustrate the flexibility of this family to analyze censored and uncensored data using two real examples. One of them was previously discussed in the literature; the second one consists of a new application to model recidivism data of a group of inmates released from the Chilean prisons during 2007. The results show that this new family of distributions has a better performance fitting the data than some common alternatives such as the exponential distribution.

**Keywords:** censored data; EM algorithm; epsilon-exponential distribution; exponential distribution; maximum likelihood; reliability analysis; survival analysis; stress-strength parameter



**Citation:** Celis, P.; de la Cruz, R.; Fuentes, C.; Gómez, H.W. Survival and Reliability Analysis with an Epsilon-Positive Family of Distributions with Applications. *Symmetry* **2021**, *13*, 908. <https://doi.org/10.3390/sym13050908>

Academic Editor: Wei-Shih Du

Received: 25 April 2021

Accepted: 18 May 2021

Published: 20 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The statistical analysis of reliability and survival data is an important topic in several areas, including medicine, epidemiology, biology, economics, engineering, and environmental sciences, to name a few. When using a parametric approach, one of the first steps for modeling the data is to choose a suitable distribution that can capture relevant features of the observations of interest. In this context, the gamma and Weibull distributions have become popular choices due to their flexibility that allows for a non-constant hazard rate function and to model skewed data. Although several alternatives were considered to accommodate different cases, researchers have continued to develop extensions and modifications of the standard distributions to increase the flexibility of the models see [1–3] for a few examples.

In this paper, we consider a generalization of the distributions with positive support and propose a new family of distributions, called the *epsilon-positive family*, whose construction is motivated by the ideas behind the generation of skew distributions using symmetric kernels. Specifically, we build upon the ideas from [4], where the authors start with a symmetric around zero distribution  $f$ , and define a family of distributions indexed by a parameter  $\gamma > 0$  as the set of densities of the form

$$h(x; \gamma) = \frac{2}{\gamma + \gamma^{-1}} \left( f(x/\gamma) \mathbf{1}_{\{x \geq 0\}} + f(\gamma x) \mathbf{1}_{\{x < 0\}} \right), \quad (1)$$

where  $1_A$  denotes the indicator function of the set  $A$ . Some extensions of this family include *epsilon-skew-normal* family introduced by [5] and the *epsilon-skew-symmetric* family introduced by [6], both discussed in some details in [7].

Here, starting with a probability density function  $g$  with positive support, we obtain a general class that extends the family of distributions with positive support, and that contains the Weibull, gamma and exponential distributions as special cases, depending on the choice of  $g$ . Furthermore, we discuss a stochastic representation and how to obtain maximum likelihood estimators for the members of this family. We also derive the corresponding survival and hazard functions and note that one interesting feature of this new class is that the hazard function is not necessarily constant.

The rest of the paper is organized as follows: in Section 2 we define the epsilon-positive family and obtain the hazard and survival functions, mean residual life and stress-strength parameters for this family. In addition, we discuss maximum likelihood estimation and how to obtain such estimates using an EM-type algorithm for the general case. In Section 3, we focus on one specific member of the family introduced in Section 2, namely *epsilon-exponential distribution*, and discuss its applicability in the analysis of survival data. In Section 4 we discuss two real data examples and we finish with a brief discussion in Section 5. We include Appendixes A and B with some of the technical details.

### 2. The Epsilon-Positive Family

Let  $g(\cdot) = g_Y(\cdot; \Psi)$  be a probability density function (pdf) with positive support and parameters  $\Psi \in \mathbb{R}^p$ . Then, for  $0 < \varepsilon < 1$ , the corresponding epsilon-positive (EP) family of distributions is defined as

$$f_X(x; \Psi, \varepsilon) = \frac{1}{2} \left[ g\left(\frac{x}{1+\varepsilon}\right) + g\left(\frac{x}{1-\varepsilon}\right) \right], \quad x > 0. \tag{2}$$

If a random variable  $X$  has the density given in (2), we say that  $X$  has an epsilon-positive distribution and write  $X \sim EP(\Psi, \varepsilon)$ .

Observe that as  $\varepsilon \downarrow 0$ ,  $f_X(x; \Psi, \varepsilon) \rightarrow g(x)1_{\{x>0\}}$  and therefore the distribution  $g_Y(\cdot; \Psi)$  can be seen as a particular member of the family.

The  $r$ th moment of  $X \sim EP(\Psi, \varepsilon)$ ,  $r = 1, 2, \dots$ , is given by

$$E(X^r) = \left( \frac{(1+\varepsilon)^{r+1}}{2} + \frac{(1-\varepsilon)^{r+1}}{2} \right) E(Y^r), \tag{3}$$

where  $E(Y^r)$  is the  $r$ th moment of  $Y \sim g_Y(\cdot; \Psi)$ . From (3) we obtain that the mean, variance, skewness (CS) and kurtosis (CK) coefficients are (respectively)

$$E(X) = (1 + \varepsilon^2)E(Y)$$

$$Var(X) = (1 + 3\varepsilon^2)E(Y^2) - (1 + \varepsilon^2)^2E^2(Y)$$

$$CS = \frac{(1 + 6\varepsilon^2 + \varepsilon^4)E(Y^3) - 3(1 + 4\varepsilon^2 + 3\varepsilon^4)E(Y)E(Y^2) + 2(1 + \varepsilon^2)^3E^3(Y)}{((1 + 3\varepsilon^2)E(Y^2) - (1 + \varepsilon^2)^2E^2(Y))^{3/2}}$$

and

$$CK = \frac{A}{((1 + 3\varepsilon^2)E(Y^2) - (1 + \varepsilon^2)^2E^2(Y))^2},$$

$$A = (1 + 10\varepsilon^2 + 5\varepsilon^4)E(Y^4) - 4(1 + 7\varepsilon^2 + 7\varepsilon^4 + \varepsilon^6)E(Y)E(Y^3) + 6(1 + 5\varepsilon^2 + 7\varepsilon^4 + 3\varepsilon^6)E^2(Y)E(Y^2) - 3(1 + 4\varepsilon^2 + 6\varepsilon^4 + 4\varepsilon^6 + \varepsilon^8)E^4(Y),$$

where  $E(Y^r)$ ,  $r = 1, 2, 3, 4$ , are the first four moments of the random variable  $Y \sim g_Y(\cdot; \Psi)$ .

To draw observations from an epsilon-positive distribution, we first notice that for  $0 < \epsilon < 1$ , if  $Y \sim g_Y(\cdot; \Psi)$  and  $U_\epsilon$  (independent from  $Y$ ) satisfies  $P(U_\epsilon = 1 + \epsilon) = 1 - P(U_\epsilon = 1 - \epsilon) = (1 + \epsilon)/2$ , then  $X = U_\epsilon Y \sim EP(\Psi, \epsilon)$ .

From this stochastic representation, it follows that we can generate EP random variables according to the Algorithm 1:

---

**Algorithm 1** Algorithm to generate observations from an epsilon-positive distribution.

---

**Require:** Initialize the algorithm fixing  $\Psi$  and  $\epsilon$

- 1: Generate  $Y$  from  $g_Y(\cdot)$  and  $U$  from  $Ber(p = \frac{1+\epsilon}{2})$
  - 2: **if**  $U = 1$  **then**
  - 3:      $U_\epsilon \leftarrow 1 + \epsilon$
  - 4: **else**
  - 5:      $U_\epsilon \leftarrow 1 - \epsilon$
  - 6: **end if**
  - 7: **return**  $X = U_\epsilon Y$ .
- 

Finally, observe that the definition in (2) can be easily extended so we can represent the epsilon-positive family as a finite scale mixture of positive distributions. In fact, for any  $0 < \epsilon < 1$  we can write

$$f_X(x; \Psi, \epsilon) = \sum_{\xi \in \mathcal{J}(\epsilon)} \pi_\xi(\epsilon) \frac{1}{\xi} g\left(\frac{x}{\xi}\right), \tag{4}$$

where  $x > 0$ ,  $\pi_\xi(\epsilon) > 0$  are mixing proportions satisfying  $\sum_{\xi \in \mathcal{J}(\epsilon)} \pi_\xi(\epsilon) = 1$ , and  $\mathcal{J}(\epsilon)$  is some finite subset of  $\Re$  that will typically depend on  $\epsilon$ . For instance, taking  $\mathcal{J}(\epsilon) = \{1 - \epsilon, 1 + \epsilon\}$  and  $\pi_\xi(\epsilon) = (1 \pm \epsilon)/2$  we recover the expression in (2). This representation will be particularly useful in order to obtain maximum likelihood estimates using EM-type algorithms, as we discuss in Section 2.5.

### 2.1. Reliability Properties

From the definition, we obtain that the survival function  $S_X(x; \Psi, \epsilon) = P(X > x)$  for this family is given by

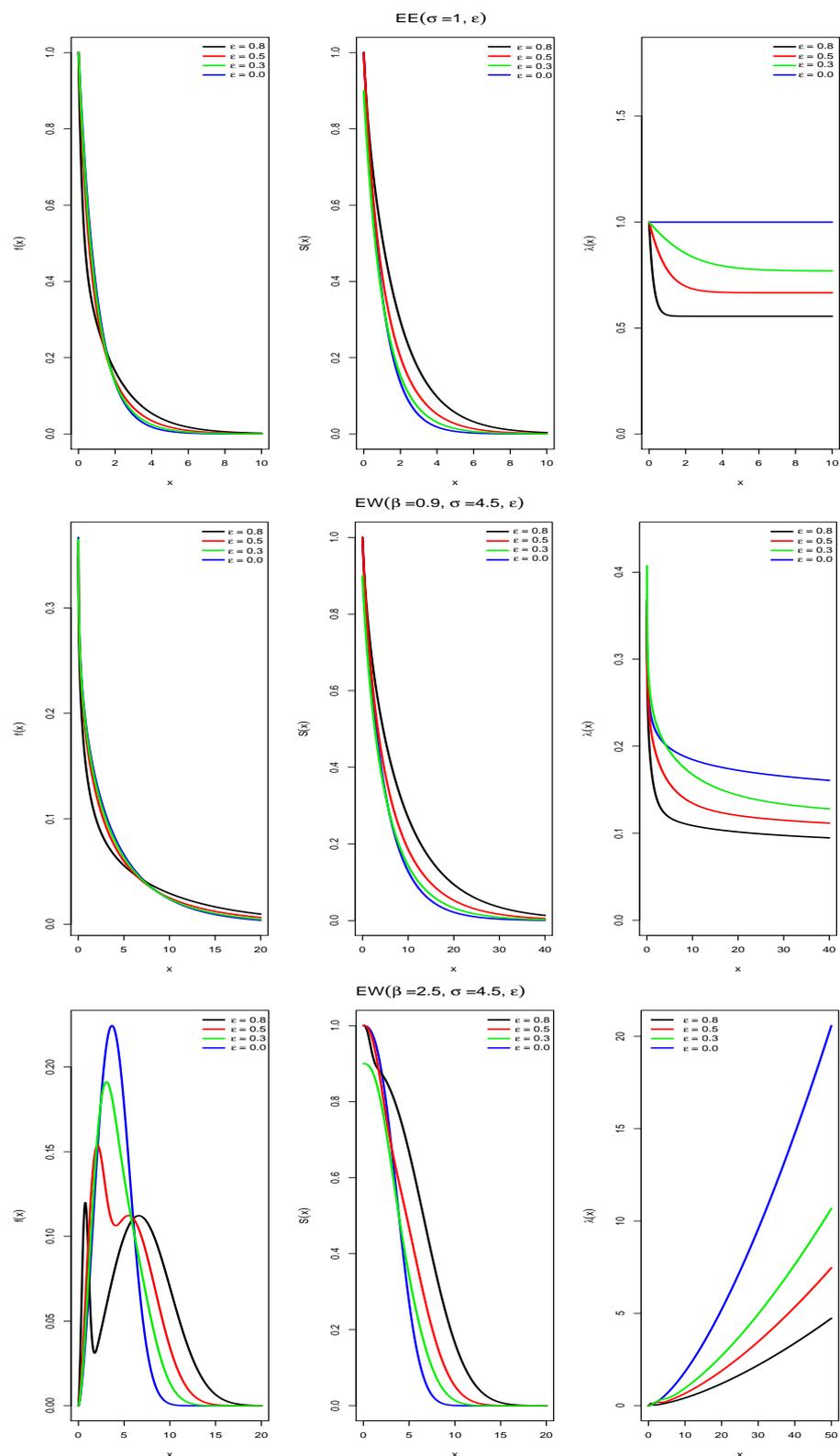
$$S_X(x; \Psi, \epsilon) = \left(\frac{1 + \epsilon}{2}\right) S_Y\left(\frac{x}{1 + \epsilon}\right) + \left(\frac{1 - \epsilon}{2}\right) S_Y\left(\frac{x}{1 - \epsilon}\right), \tag{5}$$

where  $S_Y(\cdot)$  is the survival function associated with the density  $g_Y(\cdot; \Psi)$ . Similarly, the hazard function  $\lambda_X(x; \Psi, \epsilon) = f_X(x; \Psi, \epsilon) / S_X(x; \Psi, \epsilon)$  is given by

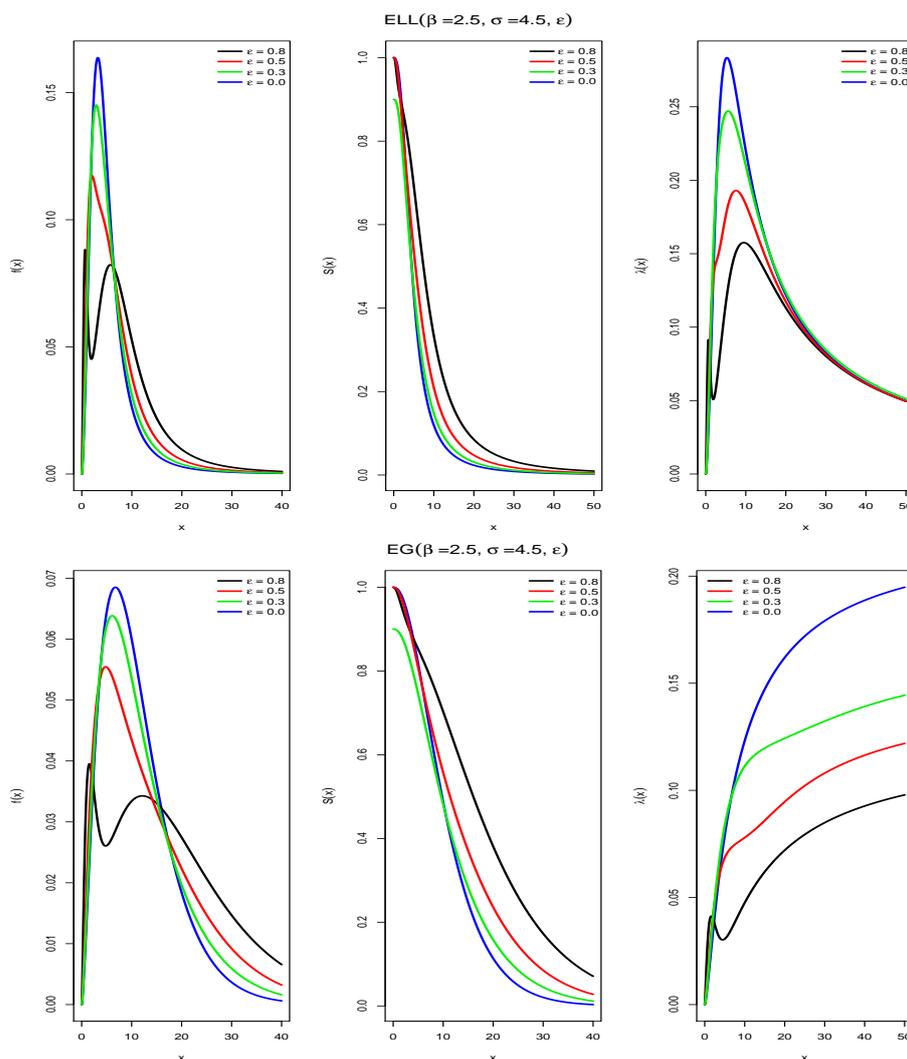
$$\lambda_X(x; \Psi, \epsilon) = \frac{1 + r(x)}{(1 + \epsilon)R\left(\frac{x}{1 + \epsilon}\right) + (1 - \epsilon)R\left(\frac{x}{1 - \epsilon}\right)r(x)}, \tag{6}$$

where  $r(x) = g_Y\left(\frac{x}{1 - \epsilon}\right) / g_Y\left(\frac{x}{1 + \epsilon}\right)$ , and  $R(\cdot) = S_Y(\cdot) / g_Y(\cdot)$  is the Mills ratio.

Table 1 shows some examples of the densities that can be extended using the definition of the epsilon-positive family, with the corresponding densities, survival and hazard functions. Figures 1 and 2 show the pdf, survival and hazard functions of the epsilon-exponential, epsilon-Weibull, epsilon-log-logistic and epsilon-gamma distributions. We can see that in the case of the epsilon-Weibull, epsilon-log-logistic and epsilon-gamma distributions a bimodal shape is obtained when the value of the parameter  $\epsilon$  is 0.9.



**Figure 1.** Examples of the probability density  $f(x)$ , survival  $S(x)$  and hazard  $\lambda(x)$  functions of epsilon-exponential distribution,  $EE(\sigma, \epsilon)$ , and epsilon-Weibull distribution,  $EW(\alpha, \sigma, \epsilon)$ . Please note that the exponential and Weibull distributions correspond to the case  $\epsilon = 0$ .



**Figure 2.** Examples of the probability density  $f(x)$ , survival  $S(x)$  and hazard  $\lambda(x)$  functions of epsilon-log-logistic distribution,  $ELL(\sigma, \epsilon)$ , and epsilon-gamma distribution,  $EG(\alpha, \sigma, \epsilon)$ . Please note that the log-logistic and gamma distributions correspond to the case  $\epsilon = 0$ .

**Table 1.** Hazard rate,  $\lambda(\cdot)$ , Survival,  $S(\cdot)$ , and density,  $f(\cdot)$ , functions of some probability models that can be generalized using the definition in (2) In the table  $I(a, \beta) = \int_0^a \Gamma(\beta)^{-1} u^{\beta-1} e^{-u} du$ .

	$\lambda(y)$	$S(y)$	$f(y)$
Exponential	$\frac{1}{\sigma} (> 0)$	$\exp(-\frac{y}{\sigma})$	$\frac{1}{\sigma} \exp(-\frac{y}{\sigma})$
Weibull	$(\frac{\beta}{\sigma^\beta}) y^{(\beta-1)} (\beta, \sigma > 0)$	$\exp(-[\frac{y}{\sigma}]^\beta)$	$\frac{\beta y^{(\beta-1)}}{\sigma^\beta} \exp(-[\frac{y}{\sigma}]^\beta)$
Log-logistic	$\frac{(\frac{\beta}{\sigma})(\frac{y}{\sigma})^{\beta-1}}{1+(\frac{y}{\sigma})^\beta} (\beta, \sigma > 0)$	$(1 + [\frac{y}{\sigma}]^\beta)^{-1}$	$\frac{(\frac{\beta}{\sigma})(\frac{y}{\sigma})^{\beta-1}}{(1+(\frac{y}{\sigma})^\beta)^2}$
Gamma	$\frac{f(y)}{S(y)}$	$1 - I(y/\sigma, \beta)$	$\frac{y^{(\beta-1)} \exp(-\frac{y}{\sigma})}{\sigma^\beta \Gamma(\beta)}$

2.2. Mean Residual Life

The mean residual life or life expectancy is an important characteristic of the model. It gives the expected additional lifetime given that a component has survived until time  $t$ . For a non-negative continuous random variable  $X \sim EP(\Psi, \epsilon)$  the mean residual life ( $mrl$ ) function is defined as

$$mrl(t) = E(X - t | X > t) = E(X | X > t) - t \tag{7}$$

where  $t > 0$ . The above conditional expectation is given by

$$E(X|X > t) = \int_t^\infty \frac{xf_X(x)}{P(X > t)}dx = \int_t^\infty \frac{xf_X(x)}{1 - F_X(t)}dx = \frac{1}{S_X(t)} \int_t^\infty xf_X(x)dx. \tag{8}$$

Calculation of the integral in (8) is done in the same way as the calculation of the mean. Thus,

$$I = \int_t^\infty xf_X(x)dx = \int_t^\infty x \frac{1}{2} \left[ g_Y\left(\frac{x}{1+\epsilon}\right) + g_Y\left(\frac{x}{1-\epsilon}\right) \right] dx.$$

Making the changes of variables  $z = \frac{x}{1+\epsilon}$  and  $u = \frac{x}{1-\epsilon}$  we have

$$\begin{aligned} I &= \frac{(1+\epsilon)^2}{2} \int_{t/(1+\epsilon)}^\infty zg_Y(z)dz + \frac{(1-\epsilon)^2}{2} \int_{t/(1-\epsilon)}^\infty ug_Y(u)du \\ &= \frac{(1+\epsilon)^2}{2} S_Y(t_1) \int_{t_1}^\infty \frac{zg_Y(z)}{S_Y(t_1)}dz + \frac{(1-\epsilon)^2}{2} S_Y(t_2) \int_{t_2}^\infty \frac{ug_Y(u)}{S_Y(t_2)}du \\ &= \frac{(1+\epsilon)^2}{2} S_Y(t_1)E(Y|Y > t_1) + \frac{(1-\epsilon)^2}{2} S_Y(t_2)E(Y|Y > t_2) \\ &= \frac{(1+\epsilon)^2}{2} S_Y(t_1)(E(Y - t_1|Y > t_1) + t_1) + \frac{(1-\epsilon)^2}{2} S_Y(t_2)(E(Y - t_2|Y > t_2) + t_2) \\ &= \frac{(1+\epsilon)^2}{2} S_Y(t_1)(mrl_Y(t_1) + t_1) + \frac{(1-\epsilon)^2}{2} S_Y(t_2)(mrl_Y(t_2) + t_2), \end{aligned}$$

where  $t_1 = \frac{t}{1+\epsilon}$ ,  $t_2 = \frac{t}{1-\epsilon}$ , and  $mrl_Y(t_i) = E(Y - t_i|Y > t_i)$ ,  $i = 1, 2$  corresponds to the mean residual life of the random variable  $Y \sim g_Y(\cdot)$ . Finally, Equation (7) can be written as

$$mrl(t) = \frac{\frac{(1+\epsilon)^2}{2} S_Y\left(\frac{t}{1+\epsilon}\right)(mrl_Y(t_1) + t_1) + \frac{(1-\epsilon)^2}{2} S_Y\left(\frac{t}{1-\epsilon}\right)(mrl_Y(t_2) + t_2)}{\left(\frac{1+\epsilon}{2}\right) S_Y\left(\frac{t}{1+\epsilon}\right) + \left(\frac{1-\epsilon}{2}\right) S_Y\left(\frac{t}{1-\epsilon}\right)} - t. \tag{9}$$

### 2.3. Stress-Strength Parameter

An important concept in reliability theory is the stress-strength parameter. Let  $X_1$  denote the strength of a system or component with a stress  $X_2$ . Then, the stress-strength parameter is defined as  $R = P(X_2 < X_1)$ , which can be viewed as a measure of the system performance. In the next theorem, we look at this quantity when  $X_1$  and  $X_2$  are independent random variables with epsilon-positive distributions.

**Theorem 1.** Suppose  $X_1$  and  $X_2$  are random variables independently distributed as  $X_1 \sim EP(\Psi_1, \epsilon_1)$  and  $X_2 \sim EP(\Psi_2, \epsilon_2)$ , the reliability of the system with stress variable ( $X_2$ ) and strength variable ( $X_1$ ) is given by

$$\begin{aligned} R = P(X_2 < X_1) &= \frac{(1+\epsilon_2)^2}{4} (aP(Y_2 < aY_1) + bP(Y_2 < bY_1)) \\ &\quad + \frac{(1-\epsilon_2)^2}{4} (cP(Y_2 < cY_1) + dP(Y_2 < dY_1)), \end{aligned} \tag{10}$$

where  $a = \frac{1+\epsilon_1}{1+\epsilon_2}$ ,  $b = \frac{1-\epsilon_1}{1+\epsilon_2}$ ,  $c = \frac{1+\epsilon_1}{1-\epsilon_2}$ ,  $d = \frac{1-\epsilon_1}{1-\epsilon_2}$ , and  $Y_i \sim g_{Y_i}(\cdot; \Psi_i)$ ,  $i = 1, 2$ , with  $Y_1$  independent of  $Y_2$ .

**Proof of Theorem 1.** Making the changes of variables  $z = \frac{x_1}{1+\varepsilon}$  and  $u = \frac{x_1}{1-\varepsilon}$  we have

$$\begin{aligned}
 P(X_2 < X_1) &= \int_0^\infty F_{X_2}(x_1)f_{X_1}(x_1)dx_1 \\
 &= \int_0^\infty \left[ \left( \frac{1+\varepsilon_2}{2} \right) G_{Y_2} \left( \frac{x_1}{1+\varepsilon_2} \right) + \left( \frac{1-\varepsilon_2}{2} \right) G_{Y_2} \left( \frac{x_1}{1-\varepsilon_2} \right) \right] \\
 &\quad \times \left[ \frac{1}{2} g_{Y_1} \left( \frac{x_1}{1+\varepsilon_1} \right) + \frac{1}{2} g_{Y_1} \left( \frac{x_1}{1-\varepsilon_1} \right) \right] dx_1 \\
 &= \frac{(1+\varepsilon_2)^2}{4} a \int_0^\infty G_{Y_2}(az)g_{Y_1}(z)dz + \frac{(1+\varepsilon_2)^2}{4} b \int_0^\infty G_{Y_2}(bu)g_{Y_1}(u)du \\
 &\quad \times \frac{(1-\varepsilon_2)^2}{4} c \int_0^\infty G_{Y_2}(cz)g_{Y_1}(z)dz + \frac{(1-\varepsilon_2)^2}{4} d \int_0^\infty G_{Y_2}(du)g_{Y_1}(u)du \\
 &= \frac{(1+\varepsilon_2)^2}{4} aP(Y_2 < aY_1) + \frac{(1+\varepsilon_2)^2}{4} bP(Y_2 < bY_1) + \\
 &\quad \frac{(1-\varepsilon_2)^2}{4} cP(Y_2 < cY_1) + \frac{(1-\varepsilon_2)^2}{4} dP(Y_2 < dY_1).
 \end{aligned}$$

□

Observe that the same concept can be used to make comparisons between two systems. For example, if  $X_1$  and  $X_2$  denote instead the lifetimes of systems  $S_1$  and  $S_2$  respectively, then, a probability  $P(X_1 < X_2) > 0.5$  would indicate that the system  $S_2$  is better than the system  $S_1$  in a stochastic sense.

#### 2.4. Maximum Likelihood Estimation

Let  $\tilde{X}_n = (X_1, \dots, X_n)$  be a random sample from an  $EP(\Psi, \varepsilon)$  distribution. Then, the maximum likelihood estimator (MLE) of  $\theta = (\Psi, \varepsilon)'$  is given by

$$\theta_{MLE} = (\hat{\Psi}, \hat{\varepsilon})_{MLE} = \arg \max_{\Psi, \varepsilon} \ell(\Psi, \varepsilon; \tilde{X}_n), \tag{11}$$

where  $\ell(\Psi, \varepsilon; \tilde{X}_n) = \sum_{i=1}^n \log f_{X_i}(x_i; \Psi, \varepsilon)$  is the log-likelihood.

Although the MLE for the  $EP$  family is conceptually straightforward, typically closed form solutions are not available and the MLE need to be obtained numerically. One possibility is the Newton–Raphson algorithm, with iteration equation

$$\hat{\theta}^{(k+1)} = \hat{\theta}^{(k)} - [H(\hat{\theta}^{(k)})]^{-1}u(\hat{\theta}^{(k)}), \tag{12}$$

where  $\theta^{(k)}$  be the current estimate of  $\theta$ ,  $u(\theta)$  denote the vector of first derivatives of  $\ell(\theta; \tilde{X}_n)$ , and  $H(\theta)$ .

A disadvantage of this approach is that it requires redthe calculation of the second derivatives of the likelihood function and repeated inversion of potentially large matrices, which can be computationally intensive. Instead, we can consider an expectation–maximization (EM) approach see [8] as a general iterative method for data sets with missing (or incomplete) data.

The mixture representation proposed in (4) is particularly useful in order to use an EM–type algorithm to estimate the model parameters, since it provides a hierarchical scheme for the  $EP$  family. Next, we show how to implement maximum likelihood estimation using an EM–type algorithm for the  $EP$  family.

#### 2.5. MLE via the EM Algorithm

From (4), the log-likelihood takes the form,

$$\ell(\Psi, \varepsilon; \tilde{X}_n) = \sum_{i=1}^n \log \left( \sum_{\xi \in \mathcal{J}(\varepsilon)} \pi_\xi(\varepsilon) \frac{1}{\xi} g \left( \frac{x_i}{\xi} \right) \right),$$

where the derivatives with respect to  $\Psi$  and  $\varepsilon$  typically lead to a system of equations with no closed form solution. To address this problem, we can “augment” the data  $\tilde{X}_n$  using an unobservable matrix  $W = (w_{ij}), i = 1, \dots, n; j = 1, \dots, m = |\mathcal{J}(\varepsilon)|$ , with elements  $w_{ij}$  defined as

$$w_{ij} = \begin{cases} 1, & \text{if observation } x_i \text{ comes from the distribution } \frac{1}{\zeta_j} g\left(\frac{x_i}{\zeta_j}\right), \\ 0, & \text{otherwise} \end{cases}$$

where  $\zeta_1, \dots, \zeta_m$  denote the distinct elements of  $\mathcal{J}(\varepsilon)$ . This way, each row of  $W$  contains only one 1 and zero 0, and the (complete) log-likelihood for the augmented data  $Y = (\tilde{X}_n, W)$  is given by

$$\ell_c(\Psi, \varepsilon; Y) = \sum_{i=1}^n \sum_{j=1}^m w_{ij} \left( \log \left( \frac{\pi_{\zeta_j}(\varepsilon)}{\zeta_j} \right) + \log g \left( \frac{x_i}{\zeta_j} \right) \right).$$

Then, if we denote by  $\hat{\theta}^{(s)} = (\Psi^{(s)}, \varepsilon^{(s)})$  the estimate of  $\theta = (\Psi, \varepsilon)$  at iteration  $s$ , and by  $\mathcal{Q}(\theta, \hat{\theta}^{(s)})$ , the conditional expectation of  $\ell_c(\theta; Y)$  given  $\tilde{X}_n$  and  $\hat{\theta}^{(s)}$ , we obtain

$$\begin{aligned} \mathcal{Q}(\theta, \hat{\theta}^{(s)}) &= E(\ell_c(\theta; Y) | \tilde{X}_n, \hat{\theta}^{(s)}) \\ &= \sum_{i=1}^n \sum_{j=1}^{m^{(s)}} w_{ij}^{(s)} \left( \log \left( \frac{\pi_{\zeta_j}(\varepsilon)}{\zeta_j} \right) + \log g \left( \frac{x_i}{\zeta_j} \right) \right), \end{aligned}$$

where  $m^{(s)} = |\mathcal{J}(\varepsilon^{(s)})|$ , and

$$w_{ij}^{(s)} = \frac{\pi(\varepsilon^{(s)})_j \frac{1}{\zeta_j} g\left(\frac{x_i}{\zeta_j}\right)}{\sum_{j=1}^{m^{(s)}} \pi_{\zeta_j}(\varepsilon^{(s)}) \frac{1}{\zeta_j} g\left(\frac{x_i}{\zeta_j}\right)}.$$

From here, it follows that for  $\mathcal{J}(\varepsilon) = \{1 - \varepsilon, 1 + \varepsilon\}$ , the iteration  $s$  of the EM algorithm takes the form:

- **E-step:** For  $i = 1, \dots, n$ , compute

$$w_{ij}^{(s)} = \frac{g\left(\frac{x_i}{1+\varepsilon^{(s)}}\right)}{g\left(\frac{x_i}{1+\varepsilon^{(s)}}\right) + g\left(\frac{x_i}{1-\varepsilon^{(s)}}\right)}, \quad \zeta_j = 1 + \varepsilon.$$

- **M-step:** Given  $\varepsilon^{(s)}$  and  $\Psi^{(s)}$ , compute

$$\begin{aligned} \varepsilon^{(s+1)} &= \frac{2}{n} \sum_{i=1}^n w_{ij}^{(s)} - 1, \quad j = 1 + \varepsilon, \\ \Psi^{(s+1)} &= \arg \max_{\Psi} \mathcal{Q}(\theta, \hat{\theta}^{(s)}). \end{aligned}$$

The E and M steps are alternated repeatedly until a convergence criteria is satisfied. For the variance estimation of the MLEs we consider the bootstrapping method suggested in [9].

### 3. The Epsilon-Exponential Distribution

If we take  $g_Y(y; \Psi = \sigma) = \frac{1}{\sigma} e^{-y/\sigma} 1_{\{y>0\}}$ , the pdf of an exponential distribution, the expression in (2) becomes

$$f_X(x; \sigma, \varepsilon) = \frac{1}{2\sigma} \left[ e^{-x/(1+\varepsilon)\sigma} + e^{-x/(1-\varepsilon)\sigma} \right], \quad x > 0, \tag{13}$$

where  $\sigma > 0$  and  $0 < \varepsilon < 1$ . We say that a random variable  $X$  has epsilon–exponential (EE) distribution with scale parameter  $\sigma$  and shape parameter  $\varepsilon$  if its density has the form in (13), and we write  $X \sim EE(\sigma, \varepsilon)$ .

Recall that the  $r$ th moment,  $r = 1, 2, \dots$ , of  $Y \sim Exp(\sigma)$  is  $E(Y^r) = r!\sigma^r$ . From (3), when  $X \sim EE(\sigma, \varepsilon)$ , we obtain that the mean, variance, skewness (CS) and kurtosis (CK) coefficients are, respectively,

$$\begin{aligned}
 E(X) &= (1 + \varepsilon^2)\sigma \\
 Var(X) &= \sigma^2(1 + 4\varepsilon^2 - \varepsilon^4) \\
 CS &= \frac{(2 + 18\varepsilon^2 - 6\varepsilon^4 + 2\varepsilon^6)}{(1 + 4\varepsilon^2 - \varepsilon^4)^{3/2}} \\
 CK &= \frac{(9 + 120\varepsilon^2 + 18\varepsilon^4 - 3\varepsilon^8)}{(1 + 4\varepsilon^2 - \varepsilon^4)^2}.
 \end{aligned}$$

Please note that for any value of  $\varepsilon \in (0, 1)$ ,  $CS > 0$  and  $CK > 0$ . It can be seen that  $2 < CS < 2.3$  and  $9 < CK < 11.023$ . Figure 3 depicts the behavior of the skewness (CS) and kurtosis (CK) coefficients as a function of  $\varepsilon$ . In the figures we observe that the maximum skewness is attained at  $\varepsilon = 0.4$ , while the maximum kurtosis coefficient is obtained when  $\varepsilon = 0.37$ .

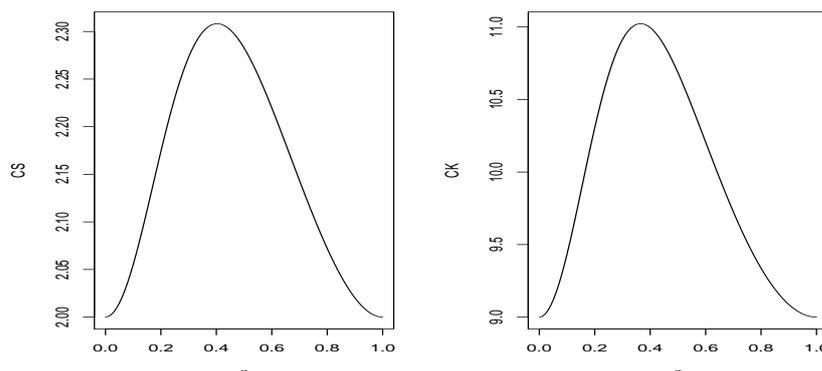


Figure 3. Skewness (CS) and kurtosis (CK) coefficients for  $X \sim EE(\sigma, \varepsilon)$ .

Recall that the survival function of the exponential distribution is of the form  $S_Y(y) = e^{-y/\sigma}1_{\{y>0\}}$  and the mean residual life is of the form  $mrl_Y(t) = \sigma$ . Then, it follows from (5), (6) and (9) that if  $X \sim EE(\sigma, \varepsilon)$ , then the survival and hazard functions and mean residual life are (respectively)

$$\begin{aligned}
 S(x; \sigma, \varepsilon) &= \frac{(1 + \varepsilon)}{2}e^{-x/(1+\varepsilon)\sigma} + \frac{(1 - \varepsilon)}{2}e^{-x/(1-\varepsilon)\sigma}, \\
 \lambda(x; \sigma, \varepsilon) &= \frac{1}{\sigma} \left[ \frac{e^{-x/(1+\varepsilon)\sigma} + e^{-x/(1-\varepsilon)\sigma}}{(1 + \varepsilon)e^{-x/(1+\varepsilon)\sigma} + (1 - \varepsilon)e^{-x/(1-\varepsilon)\sigma}} \right] \text{ and} \\
 mrl(t) &= \sigma \left[ \frac{(1 + \varepsilon)^2e^{-t/(1+\varepsilon)\sigma} + (1 - \varepsilon)^2e^{-t/(1-\varepsilon)\sigma}}{(1 + \varepsilon)e^{-t/(1+\varepsilon)\sigma} + (1 - \varepsilon)e^{-t/(1-\varepsilon)\sigma}} \right].
 \end{aligned}$$

Interestingly, in contrast to the exponential distribution, it can be shown that the hazard function  $\lambda(x; \sigma, \varepsilon)$  of the EE distribution is not constant, but decreasing in  $x$ . This feature can be easily observed in Figure 1 (top panel), where we show the pdf, survival and hazard functions of the EE distribution for different values of the parameter  $\varepsilon$  when  $\sigma = 1$ . Please note that  $\lambda(x; \sigma, \varepsilon) \rightarrow \lambda_Y(x; \sigma) = 1/\sigma$  as  $\varepsilon \rightarrow 0$ . Additionally,  $mrl(t) \rightarrow mrl_Y(t) = \sigma$  as  $\varepsilon \rightarrow 0$ .

Suppose  $X_1$  and  $X_2$  are random variables independently distributed as  $X_1 \sim EE(\sigma_1, \varepsilon_1)$  and  $X_2 \sim EE(\sigma_2, \varepsilon_2)$ , using Theorem 1 the reliability of the system with stress variable ( $X_2$ ) and strength variable ( $X_1$ ) is given by

$$R = \frac{(1 + \varepsilon_2)^2}{4} \left( \frac{a^2\sigma_1}{a\sigma_1 + \sigma_2} + \frac{b^2\sigma_1}{b\sigma_1 + \sigma_2} \right) + \frac{(1 - \varepsilon_2)^2}{4} \left( \frac{c^2\sigma_1}{c\sigma_1 + \sigma_2} + \frac{d^2\sigma_1}{d\sigma_1 + \sigma_2} \right).$$

Please note that when  $(\varepsilon_1, \varepsilon_2) \rightarrow (0, 0)$ ,  $R \rightarrow \frac{\sigma_1}{\sigma_1 + \sigma_2}$  which corresponds to the stress-strength reliability model of the exponential distributions with parameter  $\sigma_1$  for  $X_1$  (strength), and with parameter  $\sigma_2$  for  $X_2$  (stress), respectively.

Maximum likelihood estimations for the parameters  $\sigma$  and  $\varepsilon$  of the epsilon-exponential distribution, can be obtained following the strategy described in Sections 2.4 and 2.5 (see the Appendix B for details). Let  $X_1, X_2, \dots, X_n$  and  $Y_1, Y_2, \dots, Y_m$  be random samples from  $EE(\sigma_1, \varepsilon_1)$  and  $EE(\sigma_2, \varepsilon_2)$ , respectively. Having estimates of  $(\sigma_1, \varepsilon_1, \sigma_2, \varepsilon_2)$ , say  $(\hat{\sigma}_1, \hat{\varepsilon}_1, \hat{\sigma}_2, \hat{\varepsilon}_2)$ , by the invariance property of the MLE, the MLE of  $R$  becomes

$$\hat{R} = \frac{(1 + \hat{\varepsilon}_2)^2}{4} \left( \frac{\hat{a}^2\hat{\sigma}_1}{\hat{a}\hat{\sigma}_1 + \hat{\sigma}_2} + \frac{\hat{b}^2\hat{\sigma}_1}{\hat{b}\hat{\sigma}_1 + \hat{\sigma}_2} \right) + \frac{(1 - \hat{\varepsilon}_2)^2}{4} \left( \frac{\hat{c}^2\hat{\sigma}_1}{\hat{c}\hat{\sigma}_1 + \hat{\sigma}_2} + \frac{\hat{d}^2\hat{\sigma}_1}{\hat{d}\hat{\sigma}_1 + \hat{\sigma}_2} \right).$$

*Numerical Experiments*

To illustrate the properties of the estimators we performed a small simulation study considering 5000 simulated datasets for different pair of values of  $\sigma$  and  $\varepsilon$  using Algorithm 1. The goal of the study is to observe the behavior of the MLEs for the model parameters using our proposed EM algorithm considering different sample sizes.

Table 2 summarizes the simulation results. In the table, the columns labeled as “estimate” show the average of the estimators obtained in the simulations, and the columns labeled “SD” show the sample standard deviation of the corresponding estimators. To obtain the standard errors we used the bootstrap method with  $B = 150$  samples. We observe that the estimates are quite stable and fairly accurate, reporting (on average) numbers close to the true values of the parameters in all cases. Please note that as expected, the precision of the estimates improves as the sample size increase.

**Table 2.** Summary of simulation results.

True Value		n	$\sigma$		$\varepsilon$	
$\sigma$	$\varepsilon$		Estimate	SD	Estimate	SD
0.3	0.3	n = 50	0.286	0.047	0.376	0.126
		n = 100	0.292	0.037	0.350	0.127
		n = 200	0.292	0.030	0.335	0.118
		n = 500	0.295	0.022	0.317	0.106
		n = 1000	0.299	0.017	0.302	0.089
	0.5	n = 50	0.287	0.054	0.558	0.133
		n = 100	0.290	0.043	0.542	0.129
		n = 200	0.294	0.037	0.527	0.123
		n = 500	0.297	0.030	0.512	0.111
		n = 1000	0.298	0.023	0.506	0.091
	0.8	n = 50	0.299	0.059	0.818	0.126
		n = 100	0.299	0.047	0.809	0.122
		n = 200	0.301	0.038	0.801	0.108
		n = 500	0.303	0.029	0.794	0.088
		n = 1000	0.302	0.022	0.794	0.069

Table 2. Cont.

True Value		n	$\sigma$		$\varepsilon$	
$\sigma$	$\varepsilon$		Estimate	SD	Estimate	SD
0.5	0.3	n = 50	0.476	0.079	0.379	0.125
		n = 100	0.484	0.062	0.355	0.125
		n = 200	0.486	0.049	0.337	0.121
		n = 500	0.494	0.037	0.316	0.107
		n = 1000	0.497	0.029	0.303	0.088
	0.5	n = 50	0.477	0.088	0.558	0.133
		n = 100	0.484	0.074	0.542	0.131
		n = 200	0.488	0.059	0.528	0.124
		n = 500	0.494	0.048	0.512	0.109
		n = 1000	0.498	0.039	0.502	0.091
	0.8	n = 50	0.484	0.104	0.838	0.129
		n = 100	0.499	0.076	0.809	0.119
		n = 200	0.503	0.064	0.799	0.110
		n = 500	0.505	0.049	0.793	0.089
		n = 1000	0.504	0.038	0.793	0.070
0.8	0.3	n = 50	0.762	0.125	0.374	0.124
		n = 100	0.771	0.099	0.356	0.125
		n = 200	0.779	0.080	0.336	0.120
		n = 500	0.789	0.058	0.314	0.105
		n = 1000	0.796	0.046	0.301	0.089
	0.5	n = 50	0.765	0.139	0.555	0.134
		n = 100	0.771	0.116	0.543	0.134
		n = 200	0.782	0.096	0.528	0.123
		n = 500	0.789	0.076	0.515	0.110
		n = 1000	0.797	0.062	0.503	0.091
	0.8	n = 50	0.791	0.153	0.815	0.130
		n = 100	0.800	0.126	0.804	0.122
		n = 200	0.798	0.103	0.798	0.106
		n = 500	0.806	0.079	0.794	0.089
		n = 1000	0.807	0.060	0.793	0.070

#### 4. Survival and Reliability Analysis

Let  $T \geq 0$  represent the survival time until the occurrence of a “death” event. In this context, suppose we have  $n$  subjects with lifetimes determined by a survival function  $S(t)$ , and that the  $i$ th subject is observed for a time  $t_i$ . If the individual dies at time  $t_i$ , its contribution to the likelihood function is given by  $\mathcal{L}_i = f(t_i)$ , where  $f(t) = -S'(t)$  is the event density associated with  $S(t)$ , or equivalently,  $\mathcal{L}_i = S(t_i)\lambda(t_i)$ , where  $\lambda(t)$  is the corresponding hazard function. On the other hand, if the  $i$ th individual is still alive at time  $t_i$  and therefore, under non-informative censoring, all we can say is that their lifetime exceeds  $t_i$ . It follows that the contribution of a censored observation to the likelihood is simply given by  $\mathcal{L}_i = S(t_i)$ . Notice that in either case we evaluate the survival function at time  $t_i$ , because in both cases the  $i$ th subject was alive until (at least) time  $t_i$ . A death will multiply this contribution by the hazard  $\lambda(t_i)$ , but a censored observation will not.

We can combine these contributions into a single expression using a death indicator  $d_i$ , taking the value one if individual  $i$  died and the value zero otherwise. The resulting likelihood function  $\mathcal{L}$  is of the form

$$\mathcal{L} = \prod_{i=1}^n \lambda(t_i)^{d_i} S(t_i) = \prod_{i=1}^n f(t)^{d_i} S(t)^{1-d_i}.$$

In the next section we will assume that the random variable  $T$  follows an epsilon-positive family, and show how estimate the model parameters using the EM algorithm.

#### 4.1. Estimation Using the EM Algorithm

Let  $T_1, \dots, T_n$  denote the survival times,  $T_i \sim EP(\Psi, \varepsilon)$ . Using the notation introduced in the previous section, the observed data are a collection of pairs  $\tilde{X}_n = \{(T_1, d_1), \dots, (T_n, d_n)\}$ , where the  $d_i, i = 1, \dots, n$ , are the censoring indicators.

In order to implement the EM algorithm, we augment the observed data  $\tilde{X}_n$  with the unobservable matrix  $W$  defined in Section 2.5, and obtain the (complete) likelihood is given by

$$\mathcal{L}_c(\theta; \tilde{X}_n, W) = \prod_{i=1}^n \prod_{j=1}^m \left\{ \pi_{\xi_j}(\varepsilon) \left[ \frac{1}{\xi_j} g\left(\frac{t_i}{\xi_j}\right) \right]^{d_i} \left[ S\left(\frac{t_i}{\xi_j}\right) \right]^{1-d_i} \right\}^{w_{ij}}$$

with corresponding (complete) log-likelihood  $\ell_c = \log \mathcal{L}_c$ .

Then, if  $\hat{\theta}^{(s)} = (\Psi^{(s)}, \varepsilon^{(s)})'$  be the estimate of  $\theta = (\Psi, \varepsilon)'$  at iteration  $s$ , and we denote by  $\mathcal{Q}(\theta, \hat{\theta}^{(s)})$  the conditional expectation of  $\ell_c(\theta; \tilde{X}_n, W)$  given the observed data  $\tilde{X}_n$  and  $\hat{\theta}^{(s)}$ , we obtain

$$\begin{aligned} \mathcal{Q}(\theta, \hat{\theta}^{(s)}) &= E(\ell_c(\theta; \tilde{X}_n, W) | \tilde{X}_n, \hat{\theta}^{(s)}) \\ &= \sum_{i=1}^n \sum_{j=1}^{m^{(s)}} w_{ij}^{(s)} \left\{ \log \pi_{\xi_j}(\varepsilon) + d_i \log \left[ \frac{1}{\xi_j} g\left(\frac{t_i}{\xi_j}\right) \right] + (1 - d_i) \log S\left(\frac{t_i}{\xi_j}\right) \right\}, \end{aligned}$$

where

$$w_{ij}^{(s)} = \frac{\pi_{\xi_j}(\varepsilon) \left[ \frac{1}{\xi_j} g\left(\frac{t_i}{\xi_j}\right) \right]^{d_i} \left[ S\left(\frac{t_i}{\xi_j}\right) \right]^{1-d_i}}{\sum_{j=1}^{m^{(s)}} \pi_{\xi_j}(\varepsilon) \left[ \frac{1}{\xi_j} g\left(\frac{t_i}{\xi_j}\right) \right]^{d_i} \left[ S\left(\frac{t_i}{\xi_j}\right) \right]^{1-d_i}}.$$

Then, for  $\mathcal{J}(\uparrow) = \{1 - \varepsilon, 1 + \varepsilon\}$ , the iteration  $s$  of the algorithm takes the form:

- **E-step:** For  $i = 1, \dots, n$ , compute

$$w_{ij}^{(s)} = \frac{w_{i,+}}{w_{i,+} + w_{i,-}}, \quad \text{for } \xi_j = 1 + \varepsilon,$$

where

$$w_{i,+} = \left[ g\left(\frac{t_i}{1 + \varepsilon^{(s)}}\right) \right]^{d_i} \left[ (1 + \varepsilon^{(s)}) S\left(\frac{t_i}{1 + \varepsilon^{(s)}}\right) \right]^{1-d_i}$$

and

$$w_{i,-} = \left[ g\left(\frac{t_i}{1 - \varepsilon^{(s)}}\right) \right]^{d_i} \left[ (1 - \varepsilon^{(s)}) S\left(\frac{t_i}{1 - \varepsilon^{(s)}}\right) \right]^{1-d_i}.$$

- **M-step:** Given  $\varepsilon^{(s)}$  and  $\Psi^{(s)}$ , compute

$$\varepsilon^{(s+1)} = \frac{2}{n} \sum_{i=1}^n w_{ij}^{(s)} - 1, \quad \xi_j = 1 + \varepsilon,$$

$$\Psi^{(s+1)} = \arg \max_{\Psi} \mathcal{Q}(\theta, \hat{\theta}^{(s)}).$$

##### 4.1.1. EM Algorithm for the Epsilon-Exponential Distribution

Suppose that the survival times  $T_i \sim EE(\sigma, \varepsilon)$ . Then the EM algorithm takes the form:

- **E-step:** For  $i = 1, \dots, n$ , compute

$$w_{ij}^{(s)} = \frac{w_{i,+}}{w_{i,+} + w_{i,-}}, \quad \text{for } \xi_j = 1 + \varepsilon,$$

where

$$w_{i,+} = \left[ \frac{1}{\sigma} e^{-t_i / (1 + \varepsilon^{(s)}) \sigma^{(s)}} \right]^{d_i} \left[ (1 + \varepsilon^{(s)}) e^{-t_i / (1 + \varepsilon^{(s)}) \sigma^{(s)}} \right]^{1 - d_i}$$

and

$$w_{i,-} = \left[ \frac{1}{\sigma} e^{-t_i / (1 - \varepsilon^{(s)}) \sigma^{(s)}} \right]^{d_i} \left[ (1 - \varepsilon^{(s)}) e^{-t_i / (1 - \varepsilon^{(s)}) \sigma^{(s)}} \right]^{1 - d_i}.$$

- **M-step:** Given  $\varepsilon^{(s)}$  and  $\sigma^{(s)}$ , compute

$$\varepsilon^{(s+1)} = \frac{2}{n} \sum_{i=1}^n w_{ij}^{(s)} - 1, \quad \zeta_j = 1 + \varepsilon,$$

$$\sigma^{(s+1)} = \frac{\sum_{i=1}^n \sum_{j \in \mathcal{J}(\varepsilon^{(s)})} w_{ij}^{(s)} \left( \frac{t_i}{j} \right)}{\sum_{i=1}^n \sum_{j \in \mathcal{J}(\varepsilon^{(s)})} w_{ij}^{(s)} d_i}.$$

### 5. Real Data Examples

In this section, we use two examples to illustrate the proposed distributions using (un)censored data sets.

#### 5.1. Example 1: Maintenance Data

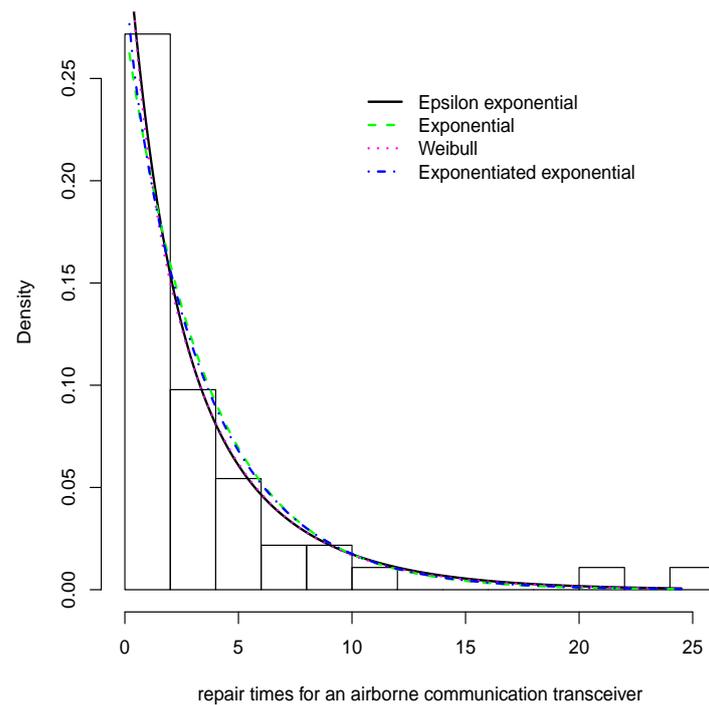
First, we consider a real data set originally analyzed by [10]. The complete data set correspond active repair times (in hours) for an airborne communication transceiver, and can be found in Table 3.

**Table 3.** Repair times (in hours) of 46 transceivers.

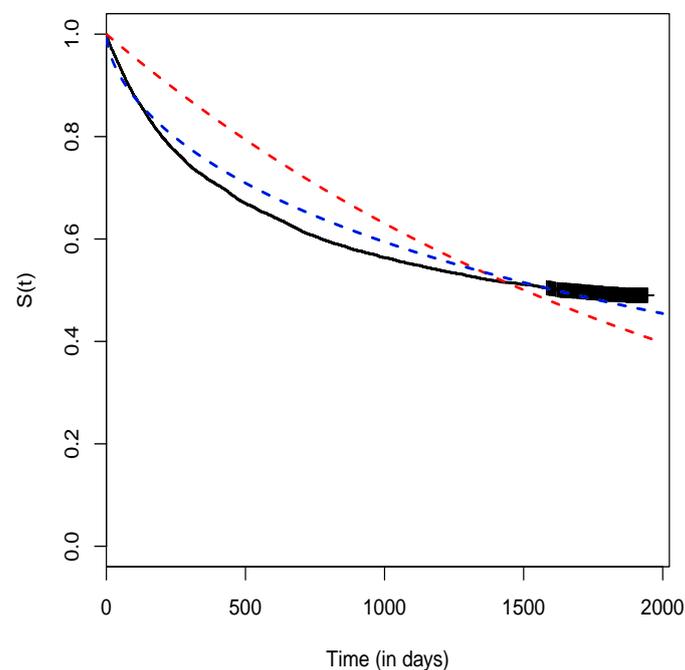
0.2	0.3	0.5	0.5	0.5	0.5	0.6	0.6	0.7	0.7
0.7	0.8	0.8	1.0	1.0	1.0	1.0	1.1	1.3	1.5
1.5	1.5	1.5	2.0	2.0	2.2	2.5	2.7	3.0	3.0
3.3	3.3	4.0	4.0	4.5	4.7	5.0	5.4	5.4	7.0
7.5	8.8	9.0	10.3	22.0	24.5				

Using the EM algorithm described in Section 4.1.1 we fit an epsilon–exponential (EE) distribution to the active repair times. We obtain that the maximized log–likelihood value  $-103.806$ . Alternatively, we also fit an exponential (Exp), exponentiated–exponential (EExp) and Weibull (Wei) distribution, obtaining the maximized log–likelihood values of  $-105.006$ ,  $-104.983$  and  $-104.470$ , respectively. For model comparison, we use the Akaike information criterion (AIC) introduced in [11], given by  $AIC = -2\hat{l} + 2k$  where  $\hat{l}$  is the maximized log–likelihood and  $k$  is the number of parameters of the distribution under consideration.

The best model is deemed to be the one with the smallest AIC. For the data set in the example, we obtain  $AIC_{EE} = 211.611$ ,  $AIC_{Exp} = 212.012$ ,  $AIC_{EExp} = 213.966$ , and  $AIC_{Wei} = 212.939$ . It follows that in terms of the AIC criteria, the epsilon–exponential shows the best performance when fitting these data. Figure 4 shows the fit of the different models used in the example. Figure 5 displays the three estimated survival functions for this data set.



**Figure 4.** The density functions of the fitted epsilon exponential, exponential, Weibull and exponentiated exponential distributions.



**Figure 5.** Fit of the survival functions: Kaplan–Meier estimator (solid line), exponential (dashed line red) and epsilon–exponential (dashed line blue) distributions.

### 5.2. Example 2: Recidivism Data

For the second example we use real data obtained from the official records of Gendarmerie of Chile on all inmates released from the Chilean prisons during 2007 after serving a sentence of imprisonment by robbery.

The data set contains records of 9477 inmates and the follow-up period from release until 30 April 2012. In this study, recidivism is defined to occur when a released prisoner goes back to prison for the original or any other offense.

Overall, 52.2% of the inmates in the cohort were convicted of one or more offenses and returned to prison within 64 months of release. About 11.8% of the cohort returned to prison within three months of release, and 30% returned within a year of release.

Table 4 shows the observed proportion of the cohort returning to prison within 1, 3, 6, 12, 18, 24, 36, 48 and 64 months of release. We observe that the cumulative proportion of recidivism grew quickly over the first 12 months after release, increasing by more than 7% every 3 months. After that, the percent increases were smaller and over longer periods.

**Table 4.** Observed recidivism by time after release.

Time after Release	% Observed Recidivists
1 month	4.7%
3 month	11.8%
6 month	19.9%
12 month	30.0%
18 month	35.9%
24 month	40.8%
36 month	46.6%
48 month	50.4%
64 month	52.2%

To analyze the time to recidivism, we determined the number of days between an inmate's release and his return to prison. Because some inmates did not reoffend, we have censored data and we used the EM algorithm described in Section 4.1.1 to fit an epsilon-exponential distribution to the time to recidivism. The maximized log-likelihood value for an assumed epsilon-exponential distribution is easily calculated to be  $-42,067.84$ . In comparison, we also fit an exponential distribution yielding a maximized log-likelihood value of  $-42,632.81$ .

Looking at the AIC values, we obtain  $AIC_{EE} = 84,139.68$  and  $AIC_{Exp} = 85,267.62$  for the epsilon-exponential and exponential model respectively, and therefore we conclude, the epsilon-exponential is a better model for these data, based on this criteria.

Finally, we also analyzed the survival time using the Kaplan-Meier estimator. Figure 5 displays the three estimated survival functions for this data set. We observe a close agreement between the Kaplan-Meier survival curve and the epsilon-exponential distribution.

## 6. Discussion

We introduced a new class of distributions with positive support called epsilon-positive which are generated from any distributions with positive support. This new class of distributions includes the exponential, Weibull, log-normal, etc. ones as special cases. We discussed a stochastic representation for this family, as well as parameter estimation, using the maximum likelihood approach via the Newton-Raphson. In addition, we showed that the elements of this new family can be expressed as a finite scale mixture of positive distributions, which facilitates the implementation of EM-type of algorithms.

We then focused on particular member of this family, called epsilon-exponential distribution, and discuss its applicability in the analysis of survival and reliability data. In this context, we considered the censored data case, and we show how we can use this new family to analyze this type of data sets. For the new class of distributions and, in particular, for the epsilon-exponential distribution we estimate the model parameters using the EM algorithm. An interesting feature of the hazard function of the epsilon-exponential distribution is that is not constant at difference of the exponential one. This feature increases the flexibility of the models allowing its use in a broader range of scenarios.

This greater flexibility is corroborated in the two examples considered in this paper where the AIC criteria shows a better performance our proposed epsilon-exponential model when compared to commonly used alternatives such as the exponential one. The results suggests that the epsilon-exponential distribution should be considered to be a

legitimate alternative for the analysis of survival and reliability data in both the censored and uncensored case.

**Author Contributions:** Conceptualization, R.d.I.C., H.W.G.; methodology, P.C., R.d.I.C., C.F. and H.W.G.; software, P.C.; validation, R.d.I.C. and H.W.G.; formal analysis, P.C.; investigation, P.C., R.d.I.C., C.F. and H.W.G.; resources, R.d.I.C. and H.W.G.; data curation, P.C. and R.d.I.C.; writing—original draft preparation, P.C., R.d.I.C., C.F. and H.W.G.; writing—review and editing, P.C., R.d.I.C., C.F. and H.W.G.; visualization, P.C.; supervision, R.d.I.C. and H.W.G.; project administration, R.d.I.C. and H.W.G.; funding acquisition, R.d.I.C. and H.W.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by ANID FONDECYT grant number 1181662 & Anillo ACT-87 project, and grant SEMILLERO UA-2021 (Chile).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** The authors would like to thank Gendarmerie of Chile for sharing the data on recidivism used in the paper.

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

## Appendix A. The EE-MLE

For a sample of  $n$  independent identically distributed (i.i.d.) observations  $\tilde{X}_n = (X_1, \dots, X_n)$  from  $EE(\sigma, \varepsilon)$ ,  $\theta = (\sigma, \varepsilon)'$ , has to be estimated from the data. The log-likelihood function is given by

$$\ell = \ell(\sigma, \varepsilon; \tilde{X}_n) = -n \log(2\sigma) + \sum_{i=1}^n \log\left(e^{-x_i/(1+\varepsilon)\sigma} + e^{-x_i/(1-\varepsilon)\sigma}\right). \quad (A1)$$

Differentiating (A1) with respect to  $\sigma$  and  $\varepsilon$  and equating to 0 respectively, we obtain the likelihood equations

$$\begin{aligned} \frac{\partial \ell}{\partial \sigma} &= \frac{-n}{\sigma} + \frac{1}{\sigma^2(1+\varepsilon)(1-\varepsilon)} \sum_{i=1}^n x_i a_i(\sigma, \varepsilon) = 0 \\ \frac{\partial \ell}{\partial \varepsilon} &= \frac{1}{\sigma(1+\varepsilon)^2(1-\varepsilon)^2} \sum_{i=1}^n x_i b_i(\sigma, \varepsilon) = 0, \end{aligned}$$

where

$$a_i(\sigma, \varepsilon) = \frac{(1-\varepsilon)e^{-x_i/(1+\varepsilon)\sigma} + (1+\varepsilon)e^{-x_i/(1-\varepsilon)\sigma}}{e^{-x_i/(1+\varepsilon)\sigma} + e^{-x_i/(1-\varepsilon)\sigma}} \quad (A2)$$

and

$$b_i(\sigma, \varepsilon) = \frac{(1-\varepsilon)^2 e^{-x_i/(1+\varepsilon)\sigma} + (1+\varepsilon)^2 e^{-x_i/(1-\varepsilon)\sigma}}{e^{-x_i/(1+\varepsilon)\sigma} + e^{-x_i/(1-\varepsilon)\sigma}}. \quad (A3)$$

Please note that no closed form solutions are available to obtain the MLEs of  $\sigma$  and  $\varepsilon$ , respectively. Therefore, the Newton–Raphson algorithm can be implemented. Let  $u(\theta) = \frac{\partial \ell}{\partial \theta}$  and  $H_{2 \times 2}(\theta) = \frac{\partial^2 \ell}{\partial \theta \partial \theta'}$  for the epsilon–exponential distribution. Let  $u(\theta) = (\partial \ell / \partial \sigma, \partial \ell / \partial \varepsilon)'$  the vector of first derivatives. Define

$$H_{11} = \frac{\partial^2 \ell}{\partial \sigma^2}, H_{12} = \frac{\partial^2 \ell}{\partial \sigma \partial \varepsilon} \text{ and } H_{22} = \frac{\partial^2 \ell}{\partial \varepsilon^2}.$$

The entries  $H_{rs}$ ,  $r, s = 1, 2$ , of the symmetric matrix of second partial derivatives for the epsilon–exponential distribution are

$$H_{11} = \frac{\partial^2 \ell}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{2}{\sigma^3(1+\varepsilon)(1-\varepsilon)} \sum_{i=1}^n x_i a_i(\sigma, \varepsilon) - \frac{1}{\sigma^4(1+\varepsilon)^2(1-\varepsilon)^2} \sum_{i=1}^n x_i^2 a_i^2(\sigma, \varepsilon) + \sum_{i=1}^n x_i b_i(\sigma, \varepsilon).$$

$$H_{22} = \frac{\partial^2 \ell}{\partial \varepsilon^2} = \frac{4\varepsilon}{\sigma(1+\varepsilon)^3(1-\varepsilon)^3} \sum_{i=1}^n x_i b_i(\sigma, \varepsilon) + \frac{1}{\sigma^2(1+\varepsilon)^4(1-\varepsilon)^4} \sum_{i=1}^n x_i^2 h_i(\sigma, \varepsilon) - \frac{1}{\sigma(1+\varepsilon)(1-\varepsilon)} \sum_{i=1}^n x_i^2 b_i^2(\sigma, \varepsilon).$$

$$H_{12} = \frac{\partial^2 \ell}{\partial \varepsilon \partial \sigma} = -\frac{1}{\sigma^2(1+\varepsilon)^2(1-\varepsilon)^2} \sum_{i=1}^n x_i b_i(\sigma, \varepsilon) + \frac{1}{\sigma^2(1+\varepsilon)^4(1-\varepsilon)^4} \sum_{i=1}^n x_i c_i(\sigma, \varepsilon) - \frac{1}{\sigma^3(1+\varepsilon)^6(1-\varepsilon)^6} \sum_{i=1}^n x_i^2 b_i^2(\sigma, \varepsilon),$$

where

$$c_i(\sigma, \varepsilon) = \frac{(1-\varepsilon)^3 e^{-x_i/(1+\varepsilon)\sigma} + (1+\varepsilon)^3 e^{-x_i/(1-\varepsilon)\sigma}}{e^{-x_i/(1+\varepsilon)\sigma} + e^{-x_i/(1-\varepsilon)\sigma}},$$

and

$$h_i(\sigma, \varepsilon) = \frac{(1-\varepsilon)^4 e^{-x_i/(1+\varepsilon)\sigma} + (1+\varepsilon)^4 e^{-x_i/(1-\varepsilon)\sigma}}{e^{-x_i/(1+\varepsilon)\sigma} + e^{-x_i/(1-\varepsilon)\sigma}},$$

and the quantities  $a_i(\sigma, \varepsilon)$  and  $b_i(\sigma, \varepsilon)$  are defined in Equations (A2) and (A3), respectively.

The functions  $u(\theta)$  and  $H_{2 \times 2}(\theta)$  define the terms of the Newton–Raphson iteration equation given in (12). To implement the Newton–Raphson algorithm, we can use the moments estimates for  $\sigma$  and  $\varepsilon$  as starting values.

Next, we show that the EM–type algorithm describe in Section 2.5 can be implemented to find the MLEs of the parameters of the epsilon–exponential distribution.

## Appendix B. An EM–Type Algorithm for the EE–MLE

In order to implement the EM algorithm to estimate the model parameters of the epsilon–exponential distribution we need to choose  $g(u) = \frac{1}{\sigma} e^{-u/\sigma}$  in the EM algorithm described in Section 2.5. Let  $\tilde{X}_n = (X_1, \dots, X_n)$  be a random sample from  $EE(\sigma, \varepsilon)$ . The complete log–likelihood is

$$\ell_c(\sigma, \varepsilon; Y) = \sum_{i=1}^n \sum_{j=1}^m w_{ij} \left( \log \left( \frac{\pi_{\xi_j}(\varepsilon)}{\xi_j} \right) - \frac{x_i}{\xi_j \sigma} - \log \sigma \right).$$

Let  $\hat{\theta}^{(s)} = (\sigma^{(s)}, \varepsilon^{(s)})'$  be the estimate of  $\theta = (\sigma, \varepsilon)'$  at iteration  $s$ , and denote by  $\mathcal{Q}(\theta, \hat{\theta}^{(s)})$  the conditional expectation of  $\ell_c(\theta; Y)$  given the observed data  $\tilde{X}_n$  and  $\hat{\theta}^{(s)}$ . We obtain

$$\begin{aligned}\mathcal{Q}(\theta, \hat{\theta}^{(s)}) &= E(\ell_c(\theta; Y) | \tilde{X}_n, \hat{\theta}^{(s)}) \\ &= \sum_{i=1}^n \sum_{j=1}^{m^{(s)}} w_{ij}^{(s)} \left( \log \left( \frac{\pi_{\xi_j}(\varepsilon)}{\xi_j} \right) - \frac{x_i}{\xi_j \sigma} \right) - n \log \sigma,\end{aligned}$$

where  $w_{ij}^{(s)} = \frac{\pi_{\xi_j}(\varepsilon^{(s)}) \frac{1}{\xi_j} e^{-x_i/\xi_j \sigma^{(s)}}}{\sum_{j=1}^{m^{(s)}} \pi_{\xi_j}(\varepsilon^{(s)}) \frac{1}{\xi_j} e^{-x_i/\xi_j \sigma^{(s)}}}$ . Therefore, the iteration  $s$  of the algorithm takes

the form:

- **E-step:** For  $i = 1, \dots, n$ , compute

$$w_{ij}^{(s)} = \frac{e^{-x_i/(1+\varepsilon^{(s)})\sigma^{(s)}}}{e^{-x_i/(1+\varepsilon^{(s)})\sigma^{(s)}} + e^{-x_i/(1-\varepsilon^{(s)})\sigma^{(s)}}}, \quad \xi_j = 1 + \varepsilon.$$

- **M-step:** Given  $\varepsilon^{(s)}$  and  $\sigma^{(s)}$ , compute

$$\varepsilon^{(s+1)} = \frac{2}{n} \sum_{i=1}^n w_{ij}^{(s)} - 1, \quad \xi_j = 1 + \varepsilon,$$

$$\sigma^{(s+1)} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{m^{(s)}} w_{ij}^{(s)} \left( \frac{x_i}{\xi_j} \right).$$

## References

1. Mudholkar, G.S.; Srivastava, D.K.; Kollia, G.D. A generalization of the weibull distribution with application to the analysis of survival data. *J. Am. Stat. Assoc.* **1996**, *91*, 1575–1583. [\[CrossRef\]](#)
2. Gupta, R.D.; Kundu, D. Exponentiated exponential family: An alternative to gamma and weibull distributions. *Biom. J.* **2001**, *43*, 117–130. [\[CrossRef\]](#)
3. Cooray, K. Analyzing lifetime data with long-tailed skewed distribution: the logistic-sinh family. *Stat. Model.* **2005**, *5*, 343–358. [\[CrossRef\]](#)
4. Fernández, C.; Steel, M.F. On Bayesian modeling of fat tails and skewness. *J. Am. Stat. Assoc.* **1998**, *93*, 359–371.
5. Mudholkar, G.S.; Hutson, A.D. The epsilon skew normal distribution for analyzing near normal data. *J. Statist. Plann. Inference* **2000**, *83*, 291–309. [\[CrossRef\]](#)
6. Arellano-Valle, R.B.; Gómez, H.W.; Quintana, F.A. Statistical inference for a general class of asymmetric distributions. *J. Statist. Plann. Inference* **2005**, *128*, 427–443. [\[CrossRef\]](#)
7. Jones, M. A note on rescalings, reparametrizations and classes of distributions. *J. Statist. Plann. Inference* **2006**, *136*, 3730–3733. [\[CrossRef\]](#)
8. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the em algorithm. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1977**, *39*, 1–38.
9. Givens, G.H.; Hoeting, J.A. *Computational Statistics*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 2012.
10. Chhikara, R.S.; Folks, J.L. The inverse gaussian distribution as a lifetime model. *Technometrics* **1977**, *19*, 461–468. [\[CrossRef\]](#)
11. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **1974**, *19*, 716–723. [\[CrossRef\]](#)