



RESEARCH ARTICLE

# A new lifetime distribution by maximizing entropy: properties and applications

Ali Khosravi Tanak<sup>1,\*</sup> , Marziyeh Najafi<sup>2</sup> and G.R. Mohtashami Borzadaran<sup>3</sup> 

<sup>1</sup>Department of Statistics, Velayat University, Iranshahr, Iran.

<sup>2</sup>Department of Mathematics, Velayat University, Iranshahr, Iran.

<sup>3</sup>Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran.

\*Corresponding author: E-mail: [a.khosravi@velayat.ac.ir](mailto:a.khosravi@velayat.ac.ir)

**Keywords:** beta generalized Weibull distribution, differential equation, hazard rate function, lifetime distribution, maximum entropy principle

**MSC:** Primary 62B10, 94A17, 62N05; Secondary 46N30, 60E05, 65C20

## Abstract

The principle of maximum entropy is a well-known approach to produce a model for data-generating distributions. In this approach, if partial knowledge about the distribution is available in terms of a set of information constraints, then the model that maximizes entropy under these constraints is used for the inference. In this paper, we propose a new three-parameter lifetime distribution using the maximum entropy principle under the constraints on the mean and a general index. We then present some statistical properties of the new distribution, including hazard rate function, quantile function, moments, characterization, and stochastic ordering. We use the maximum likelihood estimation technique to estimate the model parameters. A Monte Carlo study is carried out to evaluate the performance of the estimation method. In order to illustrate the usefulness of the proposed model, we fit the model to three real data sets and compare its relative performance with respect to the beta generalized Weibull family.

## 1. Introduction

It is known that the approximation of probability distribution is a fundamental problem in statistical data analysis. The maximum entropy principle proposed by Jaynes [22] gives us a general method of approximating a probability distribution. It states that the least biased probability distribution that describes a partially known system is the probability distribution having maximum entropy in the class of all distributions compatible with the available information. Shannon entropy maximization under different constraints has been studied by many researchers. Kapur [24] and Kagan *et al.* [23] gave a maximum entropy characterization for various distributions under moment constraints. Some researchers obtained the distributions that maximize entropy under the constraints on economic inequality measures (for example, see [17, 26, 27, 35]).

The modeling and analysis of lifetime data play an important role in a wide variety of scientific and technological fields such as engineering, medicine, and biological sciences. In the last decades, a considerable amount of studies were devoted to the introduction of lifetime distributions that provide an adequate fit to real lifetime data. Many models have been introduced in the literature by extending the Weibull and exponential distributions (for example, see [8–10, 28, 30, 38]). The modeling of lifetime distributions using the maximum entropy principle has been of interest to several researchers. Ebrahimi [16] suggested some maximum entropy lifetime (MEL) models subject to constraints on the growth and curvature of the hazard function. Asadi *et al.* [4, 5] developed a maximum entropy procedure using differential inequality constraints for hazard rate function. Distributions with maximum entropy subject

to constraints on their L-moments were proposed by Hosking [21]. Asadi et al. [6] introduced two new maximum entropy methods for change point modeling of lifetime distribution. Recently, Du et al. [15] proposed a method based on the maximum entropy principle to estimate the hazard rate function.

One of the informative descriptive indices for a random variable  $X$  with cumulative distribution function (c.d.f.)  $F$  is

$$\eta = \int_0^\infty \bar{F}^\gamma(x) dx, \quad \gamma \geq 1, \tag{1}$$

where  $\bar{F}(x) = 1 - F(x)$  is called the survival function.  $\eta$  is widely used for measuring concentration, statistical heterogeneity, and randomness, so we name it the general index (GI). In the case  $\gamma = 1$ , GI equals the distribution's mean. In the case  $\gamma = n$ , where  $n$  is a natural number, GI coincides with the expected value of the first order statistic  $X_{1:n}$  from  $n$  independent and identically distributed nonnegative random variables  $X_1, \dots, X_n$  with survival function  $\bar{F}$ . There are many applications of order statistics in statistics and applied probability, especially in reliability theory. The extreme order statistic  $X_{1:n}$  represents the lifetimes of the series systems. Moreover, there are several applications of extreme values in meteorology (extremes of temperature and pressure), oceanography (waves and tides), aeronautics (gust loads), and hydrology (floods and droughts). For more details on the order statistics, we refer to [3] and [13]. GI also is used to gauge Gini index [20] and generalized Gini index [14] as measures of heterogeneity and inequality. In addition, some measures of uncertainty and randomness such as cumulative residual Rényi entropy [40] and cumulative residual Tsallis entropy [32] are obtained by GI. In this paper, we apply the maximum entropy principle under the constraints on the mean and GI to model lifetime distribution.

This paper is organized as follows. In Section 2, we first review some results on entropy maximization under some moment constraints. Then we derive the maximum entropy distribution subject to the constraints on the mean and GI. Some statistical properties of the obtained distribution and estimation of its parameters by the method of maximum likelihood are studied in Section 3. Also, we perform a simulation study to investigate the performance of the maximum likelihood estimators. Three empirical applications to real data are illustrated in Section 4 to elucidate the potentiality of the proposed model. Finally, the conclusion is provided in Section 5.

## 2. New maximum entropy distribution

Suppose that  $X$  is a random variable having a continuous c.d.f.  $F$  with probability density function (p.d.f.)  $f$ . Entropy, as a classical measure of information, is defined by

$$H(f) = - \int_{-\infty}^\infty f(x) \log f(x) dx,$$

provided the integral exists. It was originally introduced by Shannon [37]. In the literature,  $H(f)$  is often referred to as the entropy of  $X$  or Shannon's information about  $F$ . We refer the reader to [11] for more details and references.

The maximum entropy principle is a rational method for determining a consistent probability distribution. It states that among all possible distributions compatible with a given set of constraints, we should choose the one that has maximum entropy. The distribution so obtained is called maximum entropy distribution. Kapur and Kessevan [25] studied the problem of entropy maximization under the constraints on moments. In other words, by maximizing the Shannon entropy measure subject to the following moment constraints,

$$\int_{-\infty}^\infty g_j(x) f(x) dx = \mu_j, \quad j = 0, 1, \dots, m,$$

where  $g_0(x) = 1$  and  $g_j(x), j = 1, \dots, m$  are moment functions;  $\mu_0 = 1$  and  $\mu_j, j = 1, \dots, m$  are moment values corresponding to the  $g_j(x)$  moment function, the p.d.f. of maximum entropy distribution is obtained as follows:

$$f(x) = \exp\left(-\lambda_0 - \sum_{j=1}^m \lambda_j g_j(x)\right),$$

where  $\exp(\lambda_0) = \int_{-\infty}^{\infty} \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx$  and  $\lambda_1, \dots, \lambda_m$  are the Lagrange multipliers corresponding to parameters of the maximum entropy distribution. Some well-known distributions such as normal, beta, gamma, exponential, and Weibull can be expressed as maximum entropy distributions, which are obtained under certain moment constraints. For more information on maximum entropy distributions under moment constraints, we refer to [25, 41] and references therein.

In this section, we consider the problem of entropy maximization under the constraints on mean and GI. In other words, we intend to solve the problem: Within the class of distributions supported on the nonnegative half-line ( $x \geq 0$ ), and possessing a given mean and GI, which is the distribution that maximizes entropy? We answer this question in the next theorem.

**Theorem 2.1** *The maximum entropy distribution subject to constraints on mean and GI to be  $\mu$  and  $\eta$ , respectively, has the survival function*

$$\bar{F}(x) = \left[1 + \alpha \left(e^{\beta x} - 1\right)\right]^{-\frac{1}{\gamma-1}}, \quad x \geq 0, \tag{2}$$

where  $\alpha$  and  $\beta$  are positive parameters and depend on  $\mu$  and  $\eta$  from the constraints.

*Proof.* The maximization problem is equivalent to minimize the convex functional

$$\int_0^{\infty} f(x) \log f(x) dx$$

subject to the constraints

$$\begin{cases} \int_0^{\infty} f(x) dx = 1, \\ \int_0^{\infty} xf(x) dx = \mu, \\ \int_0^{\infty} \bar{F}^{\gamma}(x) dx = \eta. \end{cases}$$

The corresponding Lagrangian is given by

$$\begin{aligned} L(f, \lambda) = & \int_0^{\infty} f(x) \log f(x) dx + \lambda_1 \left( \int_0^{\infty} f(x) dx - 1 \right) \\ & + \lambda_2 \left( \int_0^{\infty} xf(x) dx - \mu \right) + \lambda_3 \left( \int_0^{\infty} \bar{F}^{\gamma}(x) dx - \eta \right), \end{aligned}$$

where  $\lambda = (\lambda_1, \lambda_2, \lambda_3)$  is the vector of Lagrange multipliers. The first variation of the Lagrangian  $L(f, \lambda)$  is

$$\delta[L(f, \lambda)](\phi) = \lim_{\varepsilon \rightarrow 0} \frac{L(f + \varepsilon\phi, \lambda) - L(f)}{\varepsilon},$$

where  $\phi$  is an arbitrary test function. By some calculations, we have

$$\delta[L(f, \lambda)](\phi) = \int_0^\infty \left[ 1 + \log f(x) + \lambda_1 + \lambda_2 x + \lambda_3 \left( \gamma \int_0^x \bar{F}^{\gamma-1}(u) du \right) \right] \phi(x) dx.$$

Equating the first variation of the Lagrangian  $L(f, \lambda)$  to zero yields the following equation:

$$1 + \log f(x) + \lambda_1 + \lambda_2 x + \gamma \lambda_3 \int_0^x \bar{F}^{\gamma-1}(u) du = 0. \tag{3}$$

Differentiating both sides of Eq. (3) yields the differential equation

$$\frac{f'(x)}{f(x)} + \lambda_2 + \gamma \lambda_3 \bar{F}^{\gamma-1}(x) = 0.$$

Equivalently,

$$f'(x) + \lambda_2 f(x) + \gamma \lambda_3 f(x) \bar{F}^{\gamma-1}(x) = 0. \tag{4}$$

Using the fact that  $-\bar{F}'(x) = f(x)$ , we can write Eq. (4) in the following form,

$$-\bar{F}''(x) - \lambda_2 \bar{F}'(x) - \lambda_3 (\bar{F}^\gamma(x))' = 0. \tag{5}$$

Integrating both sides of Eq. (5) leads to the Riccati equation:

$$\bar{F}'(x) + \lambda_2 \bar{F}(x) + \lambda_3 \bar{F}^\gamma(x) = c,$$

where  $c$  is a constant. Using the fact that  $\bar{F}(x)$  is a survival probability function, we have  $\lim_{x \rightarrow \infty} \bar{F}(x) = 0$ . Consequently,  $c$  must be equal to zero. Thus, we arrive at the Bernoulli equation:

$$\bar{F}'(x) = c_1 \bar{F}(x) + c_2 \bar{F}^\gamma(x), \tag{6}$$

where  $c_1$  and  $c_2$  are arbitrary real coefficients. The solution of this equation is the survival function

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

where  $\alpha$  and  $\beta$  are positive-valued parameters. An outline of the derivation of Eq. (7) from the Bernoulli Eq. (6) is given in the Appendix. Since target and constraints functionals are convex, a global maximum is attained at the critical point  $\bar{F}(x)$ . □

### 3. Properties of the distribution

#### 3.1. Probability density function and hazard rate

We consider a reparameterization of the model (2) by putting  $\lambda = 1/(\gamma - 1)$  to achieve a parametric model for lifetime distribution with a simpler appearance. So, the c.d.f. of the reparameterized maximum entropy model is given by

$$F(x) = 1 - \left[ 1 + \alpha (e^{\beta x} - 1) \right]^{-\lambda}, \quad x > 0, \tag{8}$$

where the parameters  $\alpha > 0$  and  $\lambda > 0$  control the shapes of the distribution and  $\beta > 0$  is the scale parameter. We shall refer to the distribution given in (8) as the MEL distribution. If a random variable  $X$  has the MEL distribution, then we write  $X \sim \text{MEL}(\alpha, \beta, \lambda)$ .

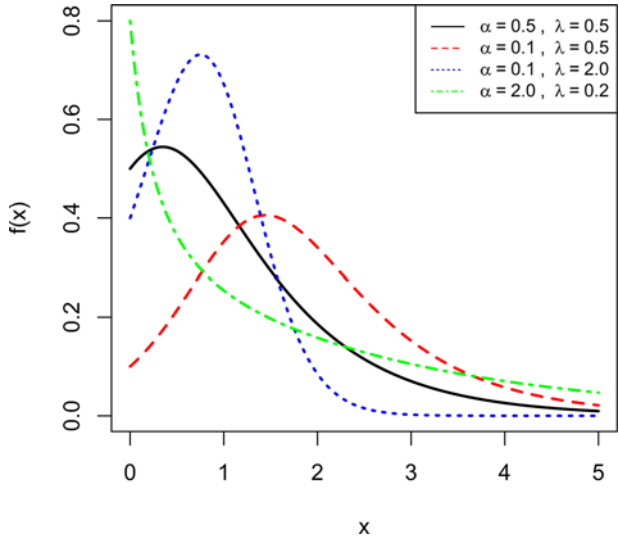


Figure 1. Probability density function of MEL distribution for  $\beta = 2$ .

The p.d.f. of the MEL distribution takes the form

$$f(x) = \alpha\beta\lambda e^{\beta x} [1 + \alpha(e^{\beta x} - 1)]^{-\lambda-1}, \quad x > 0. \tag{9}$$

The above p.d.f. is asymmetric and skewed. In the range  $\alpha \leq 1/(\lambda + 1)$ , the p.d.f. is unimodal and attains its global maximum at the value

$$x_{\text{mode}} = \frac{1}{\beta} \log\left(\frac{1 - \alpha}{\alpha\lambda}\right).$$

In the range  $\alpha > 1/(\lambda + 1)$ , the p.d.f. is monotone decreasing and attains its global maximum at the origin (i.e.,  $x_{\text{mode}} = 0$ ). At the parameter value  $\alpha = 1$ , the p.d.f. (9) coincides with the p.d.f. of exponential distribution with mean  $1/(\beta\lambda)$ . A schematic illustration of the p.d.f.s of MEL distribution with different parameters is depicted in Figure 1.

The inverse of the MEL distribution function yields a simple quantile function given by

$$Q(u) = \frac{1}{\beta} \log\left\{\frac{1}{\alpha} \left[(1 - u)^{-\frac{1}{\lambda}} - 1\right] + 1\right\}, \quad u \in (0, 1). \tag{10}$$

The above function facilitates ready quantile-based statistical modeling [19]. In addition,  $Q(u)$  gives a trivial random variable generation: if  $U \sim \mathcal{U}(0, 1)$ , then

$$X = \frac{1}{\beta} \log\left[\frac{1}{\alpha} \left(U^{-\frac{1}{\lambda}} - 1\right) + 1\right] \tag{11}$$

follows MEL distribution with parameters  $\alpha$ ,  $\beta$ , and  $\lambda$ . The median of MEL distribution can be derived from (10) by setting  $u = 0.5$ .

The hazard rate function characterizing the time-dependent nature of the aging process plays a key role in applied probability [33] and reliability engineering [7, 34]. The hazard rate function of a random

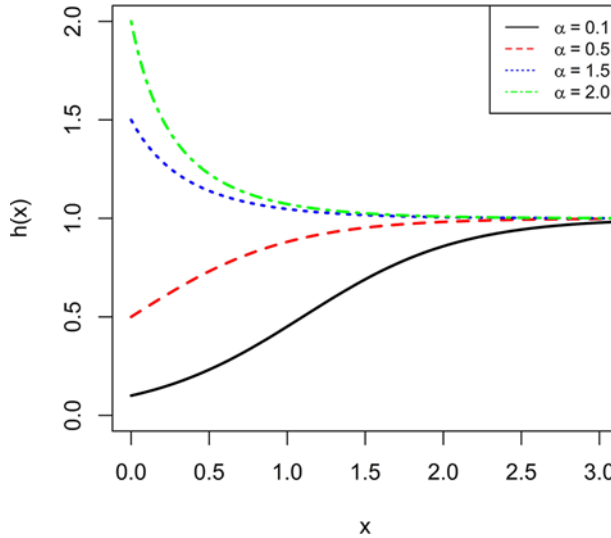


Figure 2. Hazard rate function of MEL distribution for  $\beta = 2, \lambda = 0.5$ .

variable  $X$  is defined as

$$h(x) = \frac{f(x)}{\bar{F}(x)}.$$

Considering the random variable  $X$  as a lifetime random variable, the hazard rate  $h(x)$  represents the likelihood that  $X$  be observed right after time  $x$ , given that it was not observed up to time  $x$ . The hazard rate function corresponding to (9) is

$$h(x) = \alpha\beta\lambda e^{\beta x} [1 + \alpha(e^{\beta x} - 1)]^{-1}, \quad x > 0. \tag{12}$$

Clearly, the shape of the hazard rate function of MEL distribution only depends on  $\alpha$ . It is decreasing for  $\alpha > 1$ , increasing for  $\alpha < 1$ , and constant for  $\alpha = 1$ . Plots of the hazard rate function for different values of the parameter  $\alpha$  are given in Figure 2. In the next theorem, we provide a characterization result for MEL distribution using an affine relation between the hazard rate function and the survival function.

**Theorem 3.1** *Let  $X$  be a random variable with p.d.f.  $f(x)$ , survival function  $\bar{F}(x)$ , and hazard rate function  $h(x)$ . Then  $X \sim MEL(\alpha, \beta, \lambda)$  if and only if the following relation holds:*

$$h(x) = a\bar{F}^c(x) + b, \tag{13}$$

where the transformations between the parameters  $(a, b, c)$  and  $(\alpha, \beta, \lambda)$  are given by

$$\begin{aligned} a &= (\alpha - 1)\beta\lambda, & b &= \beta\lambda, & c &= \frac{1}{\lambda}; \\ \alpha &= \frac{a}{b} + 1, & \beta &= bc, & \lambda &= \frac{1}{c}. \end{aligned}$$

*Proof.* If  $X \sim \text{MEL}(\alpha, \beta, \lambda)$ , then (13) follows from (8) and (12). Conversely, if (13) holds, then

$$\begin{aligned} \frac{f(x)}{\bar{F}(x)} &= a\bar{F}^c(x) + b, \\ \iff \bar{F}'(x) &= -a\bar{F}^{c+1}(x) - b\bar{F}(x). \end{aligned} \tag{14}$$

The solution of the Bernoulli Eq. (14) is given by

$$\bar{F}(x) = \left[ 1 + \left( \frac{a}{b} + 1 \right) \left( e^{bcx} - 1 \right) \right]^{-\frac{1}{c}}, \quad x \geq 0.$$

Therefore,  $X \sim \text{MEL}(\alpha, \beta, \lambda)$ . □

### 3.2. Moments

Some of the most important properties and characteristics of a distribution can be studied using its moments, such as the  $k$ th moment, moment generating function, and basic reliability features such as mean residual lifetime. Let  $X \sim \text{MEL}(\alpha, \beta, \lambda)$ . Using (9), the  $k$ th moment of  $X$  is given by

$$\mu'_k = E(X^k) = \frac{\lambda}{\beta^k \alpha^\lambda} \int_1^\infty (\log u)^k \left( u + \frac{1}{\alpha} - 1 \right)^{-\lambda-1} du. \tag{15}$$

The integral in Eq. (15) can be easily computed using most packages. The mean residual life (mrl) is the expected remaining lifetime, given survival up to time  $t$ . For a lifetime variable  $X$ , the mrl function is defined by

$$m(t) = E(X - t \mid X \geq t) = \frac{1}{\bar{F}(t)} \int_t^\infty \bar{F}(x) dx,$$

provided  $E(X)$  is finite. If  $X \sim \text{MEL}(\alpha, \beta, \lambda)$ , then the mrl function of  $X$  can be expressed as

$$m(t) = \frac{1}{\beta \alpha^\lambda} \left[ 1 + \alpha \left( e^{\beta t} - 1 \right) \right]^{-\lambda} \int_{e^{\beta t}}^\infty u^{-1} \left( u + \frac{1}{\alpha} - 1 \right)^{-\lambda} du.$$

Since increasing (decreasing) the hazard rate implies decreasing (increasing) the mrl [7], the mrl function of MEL distribution is decreasing for  $\alpha \leq 1$  and increasing for  $\alpha > 1$ .

### 3.3. Stochastic ordering

For nonnegative continuous random variables, stochastic ordering is a key tool for judging comparative behavior. There are different types of stochastic orderings, namely, the usual stochastic orders, the hazard rate order, the mrl order, and the likelihood ratio order. Now, we refer to a basic definition of stochastic order.

**Definition 3.2.** Let  $X$  and  $Y$  be random variables with p.d.f.s  $f_X$  and  $f_Y$ , respectively.  $X$  is smaller than  $Y$  in the likelihood ratio order, denoted by  $X \leq_{lr} Y$ , if  $f_X(x)/f_Y(x)$  is decreasing in  $x$ .

It is well-known that likelihood ratio ordering implies hazard rate ordering which, in turn, implies usual stochastic and mrl ordering. For more details on stochastic comparisons, the reader is referred to [36]. The MEL distributions are ordered with respect to the strongest likelihood ratio ordering, as shown in the following theorem.

**Theorem 3.3** Let  $X \sim MEL(\alpha_1, \beta, \lambda_1)$  and  $Y \sim MEL(\alpha_2, \beta, \lambda_2)$ . If  $\alpha_1 \geq \alpha_2$  and  $\lambda_1 \geq \lambda_2$ , then  $X \leq_{lr} Y$ .

*Proof.* First, note that

$$\frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)} = \beta e^{\beta x} \left\{ (\lambda_2 + 1) \frac{\alpha_2}{1 + \alpha_2 (e^{\beta x} - 1)} - (\lambda_1 + 1) \frac{\alpha_1}{1 + \alpha_1 (e^{\beta x} - 1)} \right\}.$$

Let  $\alpha_1 \geq \alpha_2$  and  $\lambda_1 \geq \lambda_2$ . Since  $t(u) = u/(1 + u(e^{\beta x} - 1))$  is an increasing function,  $t(\alpha_1) \geq t(\alpha_2)$ , so  $\frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)} \leq 0$ . Therefore, the likelihood ratio is decreasing, that is,  $X \leq_{lr} Y$ . Consequently,  $X$  is smaller than  $Y$  in the hazard rate, usual stochastic, and mrl ordering.  $\square$

### 3.4. Maximum likelihood estimation

In this section, we consider the estimation of the parameters of the MEL distribution by the method of maximum likelihood. Let  $x_1, \dots, x_n$  be a random sample of size  $n$  of the MEL distribution with unknown parameter vector  $\theta = (\alpha, \beta, \lambda)^T$ . The log-likelihood function for  $\theta$  based on the given random sample is

$$\ell(\theta) = n \log(\alpha\beta c) + \beta \sum_{i=1}^n x_i - (\lambda + 1) \sum_{i=1}^n \log \left[ 1 + \alpha (e^{\beta x_i} - 1) \right].$$

The maximum likelihood estimates (MLEs) of the unknown parameters are obtained by maximizing  $\ell(\theta)$  with respect to  $\theta$ . The first partial derivatives of  $\ell(\theta)$  with respect to the parameters are given by

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

The MLE  $\widehat{\theta} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda})^T$  of  $\theta = (\alpha, \beta, \lambda)^T$  can be obtained by solving the following equations simultaneously:

$$\frac{\partial \ell(\theta)}{\partial \alpha} = \frac{\partial \ell(\theta)}{\partial \beta} = \frac{\partial \ell(\theta)}{\partial \lambda} = 0.$$

There is no closed-form expression for the MLEs, so nonlinear optimization algorithms such as Newton–Raphson iterative technique can be applied to solve the equations and obtain the estimate  $\widehat{\theta}$  numerically.

Asymptotic inference for  $\theta$  can be undertaken based on asymptotic normality of  $\widehat{\theta}$ . Under some regular conditions which are stated by Cox and Hinkley [12, Chapter 9] and are fulfilled for the new distribution, the asymptotic distribution of  $\sqrt{n}(\widehat{\theta} - \theta)$  is multivariate normal  $N_3(\mathbf{0}, \mathbf{K}_\theta^{-1})$ , where  $\mathbf{K}_\theta = \lim_{n \rightarrow \infty} n^{-1} \mathbf{J}_n(\theta)$  is the expected information matrix and  $\mathbf{J}_n(\theta)$  is the observed information matrix



defined by

$$J_n(\theta) = -\frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^T} = - \begin{bmatrix} J_{\alpha\alpha} & J_{\alpha\beta} & J_{\alpha\lambda} \\ J_{\alpha\beta} & J_{\beta\beta} & J_{\beta\lambda} \\ J_{\alpha\lambda} & J_{\beta\lambda} & J_{\lambda\lambda} \end{bmatrix}, \tag{16}$$

whose elements are

$$J_{\alpha\alpha} = -\frac{n}{\alpha^2} + (\lambda + 1) \sum_{i=1}^n \frac{(e^{\beta x_i} - 1)^2}{[1 + \alpha (e^{\beta x_i} - 1)]^2},$$

$$J_{\alpha\beta} = -(\lambda + 1) \sum_{i=1}^n \frac{x_i e^{\beta x_i}}{[1 + \alpha (e^{\beta x_i} - 1)]^2},$$

$$J_{\alpha\lambda} = -\sum_{i=1}^n \frac{e^{\beta x_i} - 1}{1 + \alpha (e^{\beta x_i} - 1)},$$

$$J_{\beta\beta} = -\frac{n}{\beta^2} - \alpha (\lambda + 1) \sum_{i=1}^n \frac{x_i e^{\beta x_i} + 1 - \alpha}{[1 + \alpha (e^{\beta x_i} - 1)]^2},$$

$$J_{\beta\lambda} = -\alpha \sum_{i=1}^n \frac{x_i e^{\beta x_i}}{1 + \alpha (e^{\beta x_i} - 1)}, \quad J_{\lambda\lambda} = -\frac{n}{\lambda^2}.$$

If  $K_\theta$  is replaced by the average sample information matrix evaluated at  $\hat{\theta}$ , that is,  $n^{-1}J_n(\hat{\theta})$ , then the asymptotic behavior remains valid. We can use asymptotic normality of  $\hat{\theta}$  for interval estimation and hypothesis tests on the model parameters. For example, approximate confidence intervals for  $\alpha, \beta$ , and  $\lambda$  are given, respectively, by  $\hat{\alpha} \pm z_{\tau/2} \widehat{se}(\hat{\alpha})$ ,  $\hat{\beta} \pm z_{\tau/2} \widehat{se}(\hat{\beta})$ , and  $\hat{\lambda} \pm z_{\tau/2} \widehat{se}(\hat{\lambda})$ , where  $\widehat{se}(\cdot)$  is the square root of the diagonal element of  $J_n^{-1}(\hat{\theta})$  corresponding to each parameter and  $z_{\tau/2}$  is the upper ( $\tau/2$ )th percentile of the standard normal distribution.

### 3.5. Simulation study

In this section, we conduct Monte Carlo simulation studies to evaluate the maximum likelihood estimation of the MEL distribution parameters. We consider the no censoring case for simplicity. The results are obtained from 5,000 Monte Carlo replications from simulations carried out using the software R. In each replication, a random sample of size  $n$  is drawn from the MEL( $\alpha, \beta, \lambda$ ) distribution, and the parameters are estimated by maximum likelihood method. The MEL random number generation was performed using (11). The true parameter values used in the data-generating processes are  $\alpha = 0.01, \beta = 0.02, \lambda = 0.03$  and  $\alpha = 0.3, \beta = 0.2, \lambda = 0.1$ . Tables 1 and 2 list the mean MLEs of the three model parameters along with the respective bias and mean squared errors (MSEs) for sample sizes  $n = 50, 100, 200$ , and 300.

From Tables 1 and 2, it is noted that the magnitude of bias and MSEs tend to zero as  $n \rightarrow \infty$ . Thus, the maximum likelihood technique can be used effectively for estimating the parameters of the MEL distribution.

## 4. Applications

In the following, we present three applications of the proposed three-parameter MEL distribution to real data. These applications will show the flexibility of the new distribution in modeling lifetime data.

In order to identify the shape of hazard rate function for a lifetime data, we shall consider a useful graphical tool based on the total time on test (TTT) plot, introduced by Aarset [1]. The TTT plot is

**Table 1.** Mean, bias, and MSE of the MLEs based on Monte Carlo simulation for  $\alpha = 0.01$ ,  $\beta = 0.02$ ,  $\lambda = 0.03$ .

$n$	Parameters	Mean	Bias	MSE
50	$\alpha$	0.12818	0.11818	0.08640
	$\beta$	0.01811	-0.00188	0.00008
	$\lambda$	0.04075	0.01075	0.00061
100	$\alpha$	0.08791	0.07791	0.05512
	$\beta$	0.01806	-0.00193	0.00005
	$\lambda$	0.03845	0.00845	0.00043
200	$\alpha$	0.04232	0.03232	0.02115
	$\beta$	0.01880	-0.00120	0.00003
	$\lambda$	0.03453	0.00453	0.00021
300	$\alpha$	0.02378	0.01379	0.00760
	$\beta$	0.01909	-0.00091	0.00002
	$\lambda$	0.03302	0.00302	0.00013

**Table 2.** Mean, bias, and MSE of the MLEs based on Monte Carlo simulation for  $\alpha = 0.3$ ,  $\beta = 0.2$ ,  $\lambda = 0.1$ .

$n$	Parameters	Mean	Bias	MSE
50	$\alpha$	0.13873	0.12873	0.29745
	$\beta$	0.23803	0.03803	0.01102
	$\lambda$	0.10683	0.00683	0.00826
100	$\alpha$	0.37610	0.07610	0.16856
	$\beta$	0.29060	0.09060	0.03901
	$\lambda$	0.09796	-0.00204	0.00518
200	$\alpha$	0.30011	0.00011	0.02739
	$\beta$	0.25310	0.05310	0.02003
	$\lambda$	0.09800	-0.00200	0.00240
300	$\alpha$	0.29082	-0.00918	0.01555
	$\beta$	0.23568	0.03568	0.01312
	$\lambda$	0.09980	-0.00020	0.00174

drawn by plotting  $T(r/n) = [(\sum_{i=1}^r x_{i:n}) + (n - r)x_{r:n}] / \sum_{i=1}^n x_{i:n}$  against  $r/n$ , where  $r = 1, \dots, n$  and  $x_{i:n} (i = 1, \dots, n)$  are the observed order statistics of the sample. It is a straight diagonal for constant hazard rates and convex (concave) for decreasing (increasing) hazard rates.

We also consider the beta generalized Weibull (BGW) distribution, introduced by Singla *et al.* [38] and its sub-models for the sake of comparison. Beta Generalized Exponential (BGE), Beta Weibull (BW), Generalized or Exponentiated Weibull (GW or EW), Generalized Rayleigh (GR), Beta Exponential (BE), Generalized Exponential (GE), Weibull (W), Rayleigh (R), and Exponential (E) are well-known sub-models of BGW. The BGW family is a rich class of generalized distributions that has captured considerable attention in modeling lifetime data over the last years (for example, see [8, 18, 28, 30, 31]). A detailed discussion of the BGW family of distributions is presented in the Appendix. In order to compare the proposed model with the BGW family, we use the following goodness-of-fit statistics:  $-\ell$  (the negative maximized log-likelihood), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Akaike second-order corrected Information Criterion (AICC). The best model is the one with the least values of these statistics. In case of small sample size

**Table 3.** Time to failure of the turbocharger of one type of engine.

1.6	2.0	2.6	3.0	3.5	3.9	4.5	4.6	4.8	5.0	5.1	5.3	5.4	5.6	5.8
6.0	6.0	6.1	6.3	6.5	6.5	6.7	7.0	7.1	7.3	7.3	7.3	7.7	7.7	7.8
7.9	8.0	8.1	8.3	8.4	8.4	8.5	8.7	8.8	9.0					

**Table 4.** MLEs of the parameters for the models fitted to the failure times data and the values of  $-\ell$ , AIC, BIC, and AICc.

Model	MLEs					$-\ell$	AIC	BIC	AICc
	$a$	$b$	$\lambda$	$\alpha$	$\beta$				
<i>5 parameters</i>									
BGW	0.011	70.46	0.115	22.021	10.28	77.781	165.561	174.005	167.325
<i>4 parameters</i>									
BGE	0.171	91.54	0.224	32.27	1	81.154	170.307	177.062	171.449
BW	0.907	0.380	0.177	1	4.282	82.129	172.258	179.014	173.401
<i>3 parameters</i>									
MEL	–	–	4.387	0.002	0.672	80.880	167.761	172.828	168.428
GW	1	1	0.481	11.93	1.010	90.429	186.859	191.926	187.526
BE	7.701	17.47	0.059	1	1	87.465	180.929	185.995	181.595
<i>2 parameters</i>									
GE	1	1	0.449	9.514	1	90.143	184.285	187.663	184.609

or a large number of parameters, AICc is preferred over AIC. In this section, the statistical packages are used by R 4.1 to obtain numerical results.

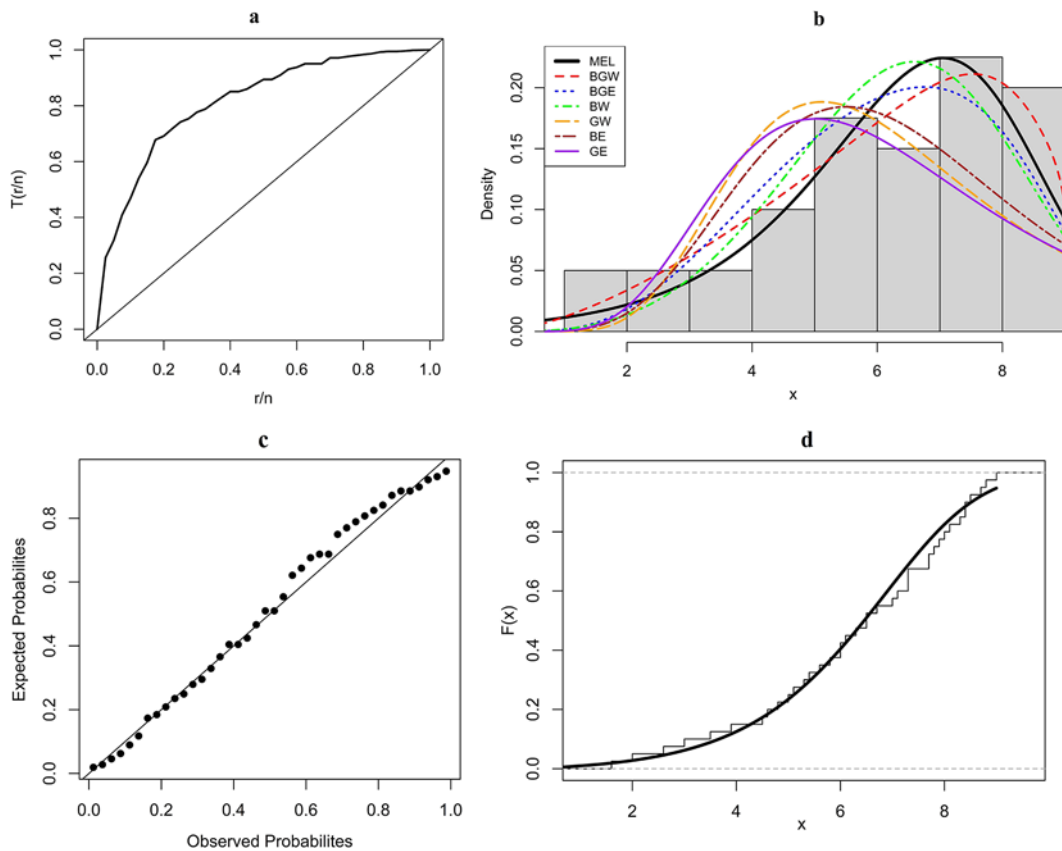
**4.1. Failure times of turbochargers**

The first data set is given by Xu *et al.* [42] on the time to failure ( $10^3$  h) of the turbocharger of one type of engine. Table 3 gives the measurements of the data set. Table 4 lists the MLEs of the parameters and the values of the goodness-of-fit statistics for the fitted models to the current data set.

From Table 3, we infer that the top two models for the first data set are the new three-parameter MEL and the five-parameter BGW distributions. However, the former has two fewer parameters to estimate and is, therefore, easier to use with less risk of overfitting. The TTT plot for failure times data presented in Figure 3(a) indicates an increasing hazard rate function. From Table 3, note that  $\hat{\alpha} < 1$  for the MEL model, which implies that the hazard rate function of this distribution is increasing, which is in accordance with Figure 3(a). The histogram of the data set with the plots of the estimated p.d.f.s of all fitted distributions are shown in Figure 3(b). This figure also depicts that the MEL and BGW distributions produce a better fit to the data set. Figure 3(c) shows the P–P plot for the MEL distribution and Figure 3(d) displays the empirical c.d.f. and estimated c.d.f. of the MEL distribution. From these figures, we also conclude that the proposed model is closely fitted to the failure times data.

The estimated variance–covariance matrix  $J_n^{-1}(\hat{\theta})$  of the MLEs under the MEL( $\alpha, \beta, \lambda$ ) distribution for the current data are given by

$$J_n^{-1}(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) = \begin{pmatrix} 1.35 \times 10^{-7} & -4.09 \times 10^{-6} & -0.00010 \\ -4.09 \times 10^{-6} & 0.00100 & -0.27173 \\ -0.00010 & -0.27173 & 8.19040 \end{pmatrix}.$$



**Figure 3.** Failure times data: (a) TTT plot; (b) histogram and p.d.f.s of the fitted models; (c) P–P plot of MEL model; and (d) empirical c.d.f. and estimated MEL c.d.f.

**Table 5.** Survival times of COVID-19 patients.

0.054	0.064	0.704	0.816	0.235	0.976	0.865	0.364	0.479
0.568	0.352	0.978	0.787	0.976	0.087	0.548	0.796	0.458
0.087	0.437	0.421	1.978	1.756	2.089	2.643	2.869	3.867
3.890	3.543	3.079	3.646	3.348	4.093	4.092	4.190	4.237
5.028	5.083	6.174	6.743	7.274	7.058	8.273	9.324	10.827
11.282	13.324	14.278	15.287	16.978	17.209	19.092	20.083	

Therefore, the approximate 95% confidence intervals for  $\alpha$ ,  $\beta$ , and  $\lambda$  are, respectively, [0.0022561, 0.0022567], [0.6527, 0.6921], and [−11.67, 20.44].

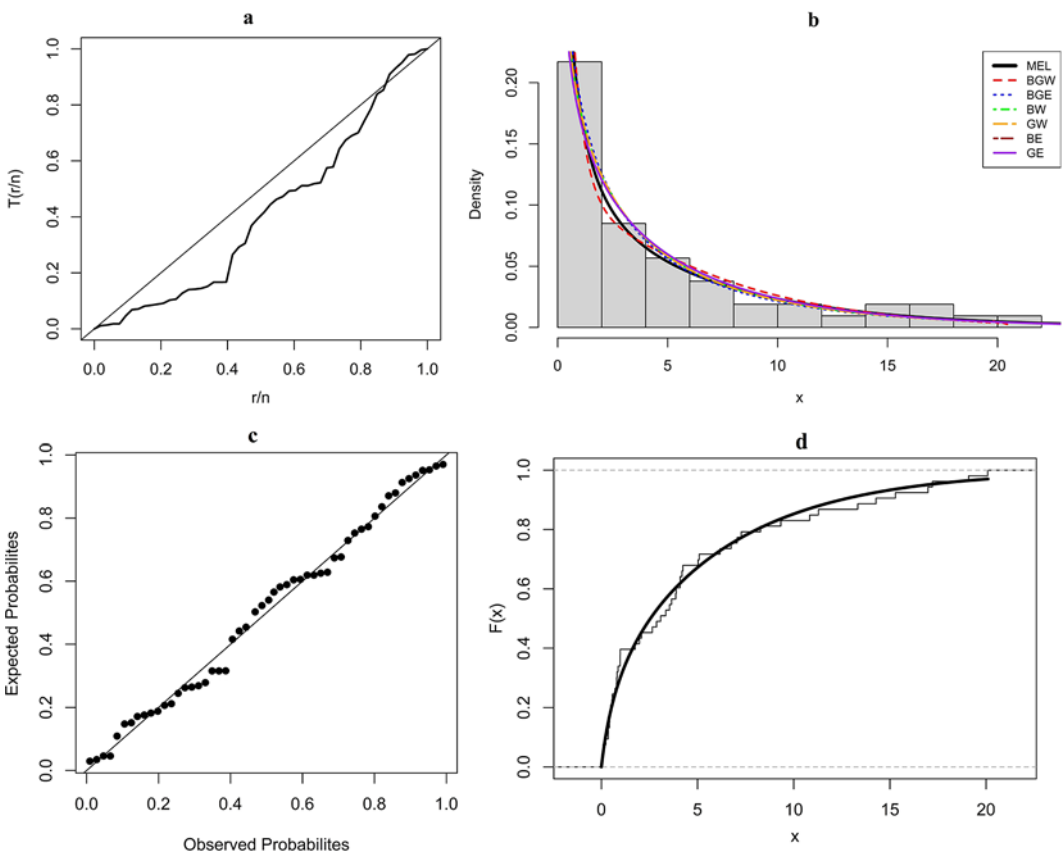
#### 4.2. Survival times of COVID-19 patients

The second real data set presented in Table 5 corresponds to the survival times (in days) of 53 COVID-19 patients from January to February 2020 in China [29]. The value of MLEs of the parameters and the goodness-of-fit statistics for the fitted models to the current data are given in Table 6.

From this table, we can conclude that the MEL distribution presents the best performance when compared with the other fitted models to these data. We have  $\hat{\alpha} > 1$  for the MEL model in Table 6,

**Table 6.** MLEs of the parameters for the models fitted to COVID-19 data and the values of  $-\ell$ , AIC, BIC, and AICc.

Model	MLEs					$-\ell$	AIC	BIC	AICc
	$a$	$b$	$\lambda$	$\alpha$	$\beta$				
<i>5 parameters</i>									
BGW	0.144	0.100	1.208	6.700	1.127	131.245	272.490	282.341	273.766
<i>4 parameters</i>									
BGE	4.783	3.156	0.038	0.207	1	133.100	274.199	282.080	275.032
BW	0.879	0.803	0.270	1	0.853	133.385	274.770	282.651	275.603
<i>3 parameters</i>									
MEL	–	–	0.259	3.628	0.610	132.237	270.475	276.386	270.965
GW	1	1	0.207	0.879	0.854	133.386	272.772	278.683	273.262
BE	0.691	0.960	0.170	1	1	133.420	272.840	278.751	273.330
<i>2 parameters</i>									
GE	1	1	0.163	0.692	1	133.420	270.840	274.780	271.080



**Figure 4.** COVID-19 data: (a) TTT plot; (b) histogram and p.d.f.s of the fitted models; (c) P–P plot of MEL model; and (d) empirical c.d.f. and estimated MEL c.d.f.

implying that the hazard rate function is decreasing. This fact is also in accordance with Figure 4(a). We provide the histogram of the data set with the plots of the estimated p.d.f.s in Figure 4(b), the P–P plot for the new model in Figure 4(c), and the empirical c.d.f. and estimated c.d.f. of

**Table 7.** Life of fatigue fracture of Kevlar 373 epoxy.

0.0251	0.0886	0.0891	0.2501	0.3113	0.3451	0.4763	0.5650	0.5671
0.6566	0.6748	0.6751	0.6753	0.7696	0.8375	0.8391	0.8425	0.8645
0.8851	0.9113	0.9120	0.9836	1.0483	1.0596	1.0773	1.1733	1.2570
1.2766	1.2985	1.3211	1.3503	1.3551	1.4595	1.4880	1.5728	1.5733
1.7083	1.7263	1.7460	1.7630	1.7746	1.8275	1.8375	1.8503	1.8808
1.8878	1.8881	1.9316	1.9558	2.0048	2.0408	2.0903	2.1093	2.1330
2.2100	2.2460	2.2878	2.3203	2.3470	2.3513	2.4951	2.5260	2.9911
3.0256	3.2678	3.4045	3.4846	3.7433	3.7455	3.9143	4.8073	5.4005
5.4435	5.5295	6.5541	9.0960					

**Table 8.** MLEs of the parameters for the models fitted to the fatigue life data and the values of  $-\ell$ , AIC, BIC, and AICc.

Model	MLEs					$-\ell$	AIC	BIC	AICc
	$a$	$b$	$\lambda$	$\alpha$	$\beta$				
<i>5 parameters</i>									
BGW	6.131	0.434	1.103	0.190	1.174	122.019	254.039	265.692	254.896
<i>4 parameters</i>									
BGE	0.630	1.788	0.445	2.443	1	122.050	252.100	261.423	252.663
BW	1.415	0.744	0.760	1	1.122	122.156	252.313	261.636	252.876
<i>3 parameters</i>									
MEL	–	–	0.300	0.182	2.432	120.403	246.808	253.800	247.141
GW	1	1	0.609	1.443	1.101	122.163	250.327	257.319	250.661
BE	1.680	1.501	0.487	1	1	122.227	250.455	257.447	250.788
<i>2 parameters</i>									
GE	1	1	0.703	1.709	1	122.244	248.487	253.149	248.652

the MEL distribution in Figure 4(d). These figures show that the new distribution fits these data adequately.

The estimated variance–covariance matrix  $J_n^{-1}(\hat{\theta})$  of the MLEs under the MEL( $\alpha, \beta, \lambda$ ) distribution for the current data are given by

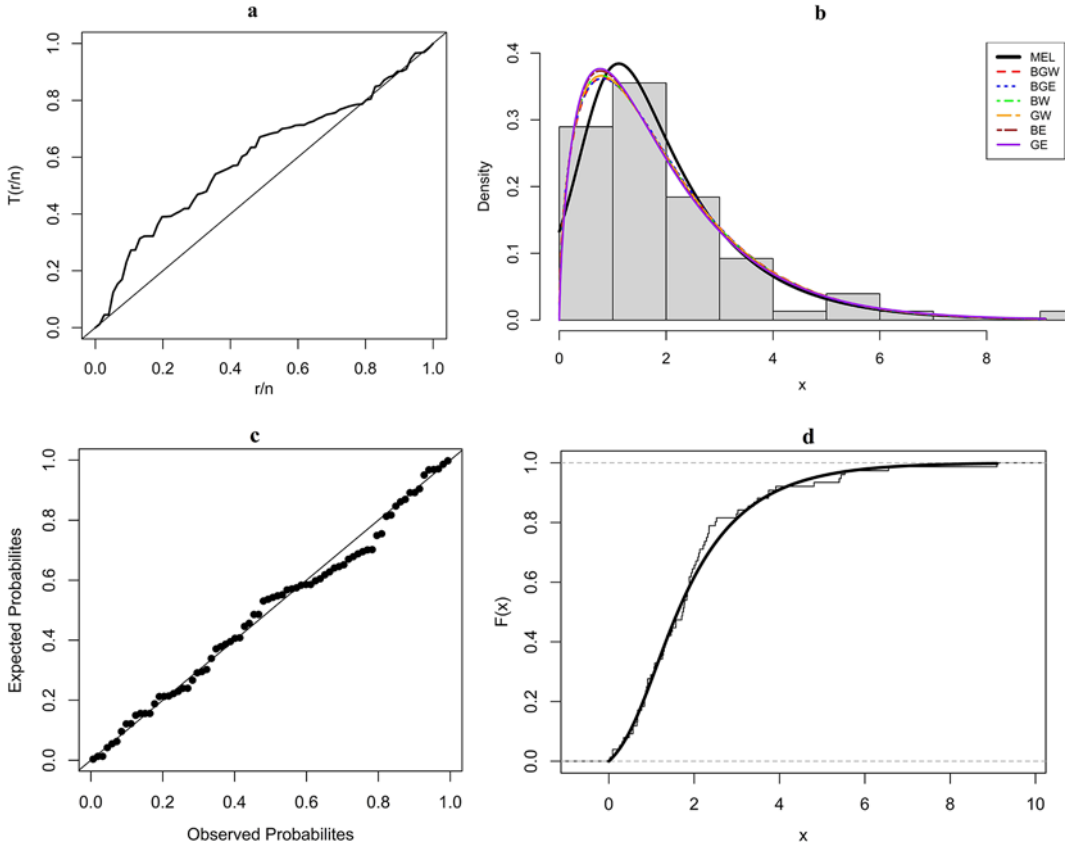
$$J_n^{-1}(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) = \begin{pmatrix} 3.24042 & 0.10985 & -0.08521 \\ 0.10985 & 0.24189 & -0.08444 \\ -0.08521 & -0.08444 & 0.03143 \end{pmatrix}.$$

Therefore, the approximate 95% confidence intervals for  $\alpha, \beta$ , and  $\lambda$  are, respectively,  $[-2.72, 9.98]$ ,  $[0.1362, 1.0844]$ , and  $[0.1974, 0.3206]$ .

### 4.3. Life of fatigue fracture of Kevlar 373 epoxy

The third data set reported by Andrews and Herzberg [2] represents the life of fatigue fracture of Kevlar 373 epoxy that is subjected to constant pressure at the 90 stress level until all have failed. The measurements of this data set are presented in Table 7.

Table 8 and Figure 5(b) clearly prove the superiority of the new model over other competitive models for the current data. The TTT plot for the fatigue life data displayed in Figure 5(a) indicates increasing



**Figure 5.** Fatigue life data: (a) TTT plot; (b) histogram and p.d.f.s of the fitted models; (c) P–P plot of MEL model; and (d) empirical c.d.f. and estimated MEL c.d.f.

hazard rate function, which is confirmed by  $\hat{\alpha}$  for the MEL model in Table 8. Figure 5(c) and (d) also depict that the MEL distribution produces an adequate fit to the third data set.

The estimated variance–covariance matrix  $J_n^{-1}(\theta)$  of the MLEs under the MEL( $\alpha, \beta, \lambda$ ) distribution for the current data are given by

$$J_n^{-1}(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) = \begin{pmatrix} 0.00923 & -0.06351 & 0.00692 \\ -0.06351 & 0.23383 & -0.16097 \\ 0.00692 & -0.16097 & 0.02502 \end{pmatrix}.$$

Thus, the approximate 95% confidence intervals for  $\alpha, \beta,$  and  $\lambda$  are, respectively,  $[-0.0067, 0.3700], [0.3520, 4.5121],$  and  $[-0.0096, 0.6103].$

**5. Conclusion**

In this paper, we have proposed a new three-parameter distribution, so-called the MEL distribution, by using the maximum entropy principle subject to the constraints on the mean and a GI. This GI is widely used for measuring concentration, variability, and uncertainty. The proposed MEL model has two shape parameters and one scale parameter. It includes exponential distribution as a special sub-model. The MEL p.d.f. can take various forms depending on its shape parameters. The new model can present both increasing and decreasing hazard rate functions. We have provided some of the mathematical

properties of the new distribution, including quantiles, moments, mrl, characterization, and stochastic ordering. The model parameters are estimated using the maximum likelihood approach. The simulation results show that the method of maximum likelihood performs well in estimating the model parameters. Finally, three real data sets are analyzed to illustrate the importance and flexibility of the new proposed distribution in modeling lifetime data. In conclusion, the MEL distribution may provide a flexible mechanism for fitting a wide spectrum of positive real-world data sets. We hope the proposed model might serve as an alternative model to other models available in the literature for modeling real data in areas such as engineering, survival analysis, economics, and so on.

**Acknowledgments.** The authors thank the editor and anonymous reviewers for their insightful suggestions and positive remarks, which have considerably improved this paper.

**Competing Interest.** The authors declare no conflict of interest.

## References

- [1] Aarset, M.V. (1987). How to identify bathtub hazard rate. *IEEE Transactions on Reliability* 36(1): 106–108.
- [2] Andrews, D.F. & Herzberg, A.M. (2012) *Data: A collection of problems from many fields for the student and research worker*. Berlin: Springer Science and Business Media.
- [3] Arnold, B.C., Balakrishnan, N., & Nagaraja, H.N. (2008). *A first course in order statistics*. New York: John Wiley and Sons.
- [4] Asadi, M., Ebrahimi, N., Hamedani, G.G., & Soofi, E.S. (2004). Maximum dynamic entropy models. *Journal of Applied Probability* 41(2): 379–390.
- [5] Asadi, M., Ebrahimi, N., & Soofi, E.S. (2005). Dynamic generalized information measures. *Statistics and Probability Letters* 71(1): 85–98.
- [6] Asadi, M., Ebrahimi, N., Soofi, E.S., & Zarezadeh, S. (2014). New maximum entropy methods for modeling lifetime distributions. *Naval Research Logistics* 61(6): 427–434.
- [7] Barlow, R.E. & Proschan, F. (1996). *Mathematical theory of reliability*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- [8] Barreto-Souza, W., Santos, A.H., & Cordeiro, G.M. (2010). The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation* 80(2): 159–172.
- [9] Block, H.W., Langberg, N.A., & Savits, T.H. (2012). A mixture of exponential and IFR gamma distributions having an upsidedown bathtub-shaped failure rate. *Probability in the Engineering and Informational Sciences* 26(4): 573–580.
- [10] Carrasco, J.M.F., Ortega, E.M.M., & Cordeiro, G.M. (2008). A generalized modified Weibull distribution for lifetime modeling. *Computational Statistics and Data Analysis* 53(2): 450–462.
- [11] Cover, T.M. & Thomas, J.A. (2006). *Elements of information theory*. New York: Wiley.
- [12] Cox, D.R. & Hinkley, D.V. (1974). *Theoretical statistics*. London: Chapman and Hall.
- [13] David, H.A. & Nagaraja, H.N. (2003). *Order statistics*. New York: Wiley.
- [14] Donaldson, D. & Weymark, J.A. (1980). A single-parameter generalization of the Gini index of inequality. *Journal of Economic Theory* 22(1): 67–86.
- [15] Du, Y.M., Ma, Y.H., Wei, Y.F., Guan, X., & Sun, C.P. (2020). Maximum entropy approach to reliability. *Physical Review E* 101(1): p.012106.
- [16] Ebrahimi, N. (2000). The maximum entropy method for lifetime distributions. *Sankhya (A)* 62: 236–243.
- [17] Eliazar, I. & Sokolov, M. (2010). Maximization of statistical heterogeneity: from Shannon's entropy to Gini's index. *Physica A* 389(16): 3023–3038.
- [18] Famoye, F., Lee, C., & Olumolade, O. (2005). The beta-Weibull distribution. *Journal of Statistical Theory and Applications* 4(2): 121–136.
- [19] Gilchrist, W.G. (2001). *Statistical modelling with quantile functions*. Boca Raton: Chapman and Hall/CRC.
- [20] Gini, C. (1914). Sulla misura della concentrazione e della variabilit'a dei caratteri. *Atti del Reale Istituto Veneto di Scienze* 73: 1203–1248.
- [21] Hosking, J.R. (2007). Distributions with maximum entropy subject to constraints on their L-moments or expected order statistics. *Journal of Statistical Planning and Inference* 137(9): 2870–2891.
- [22] Jaynes, E.T. (1957). Information theory and statistical mechanics. *Physical Review* 106(4): 620–630.
- [23] Kagan, A.M., Y.M. Linnik, & Rao, C.R. (1973). *Characterization problems in mathematical statistics*. New York: Wiley.
- [24] Kapur, J.N. (1989). *Maximum entropy models in a science and engineering*. New York: Wiley.
- [25] Kapur, J.N. & Kessevan, H.K. (1992). *Entropy optimization principle with applications*. San Diego: Academic Press.
- [26] Khosravi Tanak, A., Mohtashami Borzadaran, G.R., & Ahmadi, J. (2015). Entropy maximization under the constraints on the generalized Gini index and its application in modeling income distributions. *Physica A* 438: 657–666.



[27] Khosravi Tanak, A., Mohtashami Borzadaran, G.R., & Ahmadi, J. (2018). New functional forms of Lorenz curves by maximizing Tsallis entropy of income share function under the constraint on generalized Gini index. *Physica A* 511: 280–288.

[28] Lee, C., Famoye, F., & Olumolade, O. (2007). Beta-Weibull distribution: some properties and applications to censored data. *Journal of Modern Applied Statistical Methods* 6(1): 173–186.

[29] Liu, X., Ahmad, Z., Gemeay, A.M., Abdulrahman, A.T., Hafez, E.H., & Khalil, N. (2021). Modeling the survival times of the COVID-19 patients with a new statistical model: A case study from China. *PLoS One* 16(7): e0254999.

[30] Nadarajah, S. & Kotz, S. (2006). The beta exponential distribution. *Reliability Engineering and System Safety* 91(6): 689–697.

[31] Nadarajah, S., Cordeiro, G.M., & Ortega, E.M. (2013). The exponentiated Weibull distribution: A survey. *Statistical Papers* 54(3): 839–877.

[32] Rajesh, G. & Sunoj, S.M. (2019). Some properties of cumulative Tsallis entropy of order  $\alpha$ . *Statistical Papers* 60(3): 933–943.

[33] Ross, S.M. (2007). *Introduction to Probability Models*. Boston: Academic press.

[34] Rausand, M. & Hoyland, A. (2004). *System Reliability Theory: Models, Statistical Methods, and Applications*. New York: Wiley.

[35] Ryu, H.K. (2008). Maximum entropy estimation of income distributions from Bonferroni indices, In D. Chotikapanich (ed), *Modeling Income Distributions and Lorenz Curves*. New York: Springer.

[36] Shaked, M. & Shanthikumar, J.G. (2007). *Stochastic orders and their applications*. New York: Springer.

[37] Shannon, C.E. (1948). A mathematical theory of communication. *The Bell System Technical Journal* 27(3): 379–423.

[38] Singla, N., Jain, K., & Sharma, S.K. (2012). The beta generalized Weibull distribution: Properties and applications. *Reliability Engineering and System Safety* 102: 5–15.

[39] K.P. Singh, C.M. Lee, & E.O. George (1988). On generalized log-logistic model for censored survival data. *Biometrical Journal* 30(7): 843–850.

[40] Sunoj, S.M. & Linu, M.N. (2012). Dynamic cumulative residual Rényi’s entropy. *Statistics* 46(1): 41–56.

[41] Usta, I. & Kantar, Y.M. (2011). On the performance of the flexible maximum entropy distributions within partially adaptive estimation. *Computational Statistics and Data Analysis* 55(6): 2172–2182.

[42] Xu, K., Xie, M., Tang, L.C., & Ho, S.L. (2003). Application of neural networks in forecasting engine systems reliability. *Applied Soft Computing* 2(4): 255–268.

**Appendix A. Appendix**

**A.1. Solution of the Bernoulli equation**

The substitution  $y = \bar{F}^{1-\gamma}$  reduces the Bernoulli Eq. (6) to the linear differential equation

$$\frac{dy}{dx} = c_1y + c_2. \tag{A.1}$$

By considering the boundary condition  $y(0) = 1$ , the solution of differential equation (A.1) is given by

$$y(x) = \frac{c_1 + c_2}{c_1} e^{c_1x} - \frac{c_2}{c_1}.$$

Putting  $\alpha = \frac{c_1+c_2}{c_1}$  and  $\beta = c_1$ , we arrive at Eq. (7).

**A.2. The BGW family**

Singla *et al.* [38] introduced a flexible five-parameter model called BGW distribution by considering the baseline distribution of beta generalized distribution [39] to be EW. The c.d.f. and p.d.f. of BGW distribution, respectively, are given by

$$F(x) = \frac{1}{B(a, b)} \int_0^{(1-e^{-(\lambda x)^\beta})^\alpha} w^{a-1} (1-w)^{b-1} dw, \quad x > 0$$

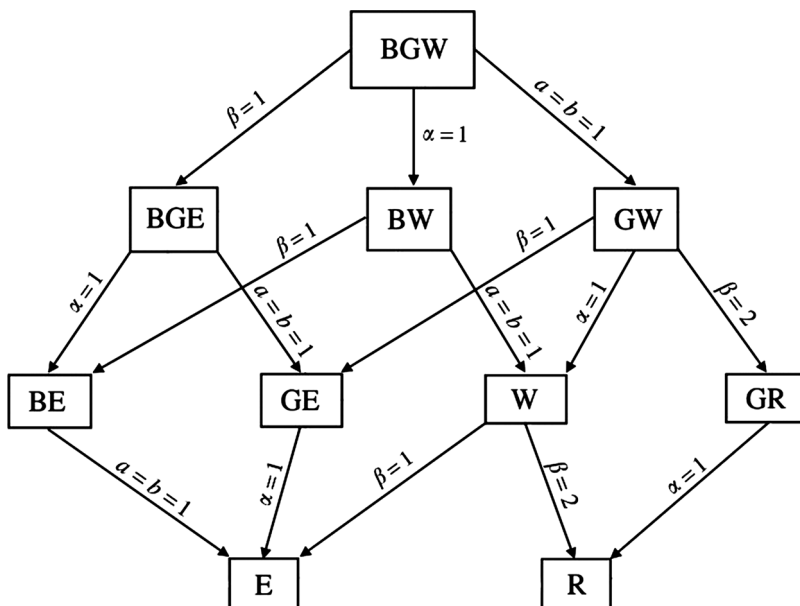


Figure A1. Sub-models of BGW.

and

$$f(x) = \frac{\alpha\beta\lambda^\beta x^{\beta-1}}{B(a,b)} \left(1 - e^{-(\lambda x)^\beta}\right)^{\alpha a-1} \left\{1 - \left(1 - e^{-(\lambda x)^\beta}\right)^\alpha\right\}^{b-1} e^{-(\lambda x)^\beta}, \quad x > 0,$$

where  $B(a, b)$  is the beta function,  $\lambda > 0$  is a scale parameter and  $a, b, \alpha, \beta > 0$  are shape parameters. This distribution exhibits increasing, decreasing, bathtub, and upside-down bathtub-shaped failure rate functions for different parametric combinations. Singla *et al.* [38] expressed the p.d.f. and c.d.f. of BGW distribution as mixtures of EW distribution. The BGW distribution unifies many existing distributions as its sub-models with applications in modeling a wide spectrum of real data sets in reliability and engineering. The sub-models of BGW distribution are as follows:

1. For  $\beta = 1$ , we get the BGE distribution.
2. If  $a = b = 1$ , then BGW distribution reduces to the GW distribution. If in addition  $\beta = 2$ , we obtain GR distribution.
3. BW distribution arises as a special case of BGW by taking  $\alpha = 1$ .
4. Assuming  $a = b = \beta = 1$ , we get GE distribution.
5. With  $\alpha = \beta = 1$ , BE distribution can be obtained.
6. For  $a = b = \alpha = 1$ , we obtain the Weibull (W) distribution with parameters  $\lambda$  and  $\beta$ . If in addition  $\beta = 2$ , the BGW distribution becomes Rayleigh (R) distribution.

The above relationships are also depicted in pictorial form in Figure A1.