Study of the New Finite Mixture of Weibull Extension Model: Identifiability, Properties and Estimation

Noura S. Mohamed*, Moshira A. Ismail, Sanaa A. Ismail

Department of Statistics, Faculty of Economics and Political Science, Cairo University, Giza, Egypt

Received September 28, 2021; Revised November 15, 2021; Accepted December 13, 2021

Cite This Paper in the following Citation Styles
(a): [1] Noura S. Mohamed, Moshira A. Ismail, Sanaa A. Ismail, "Study of the New Finite Mixture of Weibull Extension Model: Identifiability, Properties and Estimation," Mathematics and Statistics, Vol. 10, No. 1, pp. 35 - 60, 2022. DOI: 10.13189/ms.2022.100104.

(b): Noura S. Mohamed, Moshira A. Ismail, Sanaa A. Ismail (2022). Study of the New Finite Mixture of Weibull Extension Model: Identifiability, Properties and Estimation. Mathematics and Statistics, 10(1), 35 - 60. DOI: 10.13189/ms.2022.100104.

Copyright©2022 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract Finite mixture models have been used in many fields of statistical analysis such as pattern recognition, clustering and survival analysis, and have been extensively applied in different scientific areas such as marketing, economics, medicine, genetics and social sciences. Introducing mixtures of new generalized lifetime distributions that exhibit important hazard shapes is a major field of research aiming at fitting and analyzing a wider variety of data sets. The main objective of this article is to present a full mathematical study of the properties of the new finite mixture of the three-parameter Weibull extension model, considered as a generalization of the standard Weibull distribution. The new proposed mixture model exhibits a bathtub-shaped hazard rate among other important shapes in reliability applications. We analytically prove the identifiability of the new mixture and investigate its mathematical properties and hazard rate function. Maximum likelihood estimation of the model parameters is considered. The Kolmogrov-Smirnov test statistic is used to fit two famous data sets from mechanical engineering to the proposed model, the Aarset data and the Meeker and Escobar datasets. Results show that the two-component version of the proposed mixture is a superior fit compared to various lifetime distributions, either one-component or two-component lifetime distributions. The new proposed mixture is a significant statistical tool to study lifetime data sets in numerous fields of study.

Keywords Weibull Extension Model, Finite Mixture Models, Identifiability, Hazard Rate Function, Mean Residual Life, Skewness, Kurtosis, Maximum Likelihood Estimation, Goodness of Fit Tests

1. Introduction

Historically, survival analysis has been studied with the classical statistical distributions such as Gamma, exponential and Weibull distributions among others. Generalizations of such lifetime distributions have been introduced in the literature to fit a wider variety of data sets and to present multiple shapes of useful hazard functions. The Weibull distribution is perhaps the most widely used lifetime distribution model because of its flexibility and simple expressions for the density, survival, and hazard functions. However, it cannot capture the behavior of lifetime data that exhibit non-monotone hazard shapes (such as bathtub, inverse bathtub (unimodal)...etc.), often encountered in reliability engineering studies. A number of distributions were developed as generalizations or modifications of the Weibull distribution. Reference [1] introduced the Inverse generalized Weibull and generalized Inverse generalized Weibull (GIGW) distributions. They investigated the mixture model of two-component GIGW distributions. Reference [2] introduced the four-parameter new generalized Inverse Weibull distribution. They derived various structural properties and MLEs for the new modified distribution. Reference [3] studied a new
generalization of the flexible Weibull distribution with three parameters referred to as the exponentiated flexible Weibull extension. Their new distribution exhibits bathtub shape hazard rate function. Reference [4] proposed a generalized form of the Weibull distribution known as the Fréchet-Weibull distribution by using the T-X family. Reference [5] presented a detailed statistical analysis for their new proposed three-parameter Weibull Inverse Rayleigh distribution.

Our distribution of interest, the Weibull extension model (WEM), presented by [6], is an example of such generalizations that exhibits bathtub shape hazard rate function among other important shapes. The Weibull extension Model (WEM) is considered a generalization of the model studied by [7] by adding a scale parameter β in order to enhance its capability to fit diverse lifetime data. The model has been fitted to the failure times of Aarset data and the failure times of a sample of devices in a large system (see, [8]). Reference [9] investigated the estimation of the model parameters based on generalized order statistics.

The probability density function pdf, the cumulative distribution function cdf, survival and hazard functions of the three-parameter Weibull extension model are given by, (see, [5]).

\[ f(x) = \lambda \alpha \left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left\{\left(\frac{x}{\beta}\right)^{\alpha} + \lambda \beta \left(1 - e^{\frac{x}{\beta}^{\alpha}}\right)\right\}, \]

\[ x \geq 0, \quad \alpha, \beta, \lambda \geq 0. \quad (1.1) \]

\[ F(x) = 1 - \exp\left\{\lambda \beta \left(1 - e^{\frac{x}{\beta}^{\alpha}}\right)\right\}, \quad (1.2) \]

\[ S(x) = 1 - F(x) = \exp\left\{\lambda \beta \left(1 - e^{\frac{x}{\beta}^{\alpha}}\right)\right\}, \quad (1.3) \]

\[ h(x) = \lambda \alpha \left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left\{\left(\frac{x}{\beta}\right)^{\alpha}\right\} \quad (1.4) \]

where \(\alpha, \lambda\) are shape parameters and \(\beta\) is a scale parameter. For analysis purpose, the product \(\lambda \beta\) is referred to as the parameter \(\gamma\).

The shape of the hazard rate depends only on the shape parameter \(\alpha\). For \(\alpha \geq 1\),

(i). \(h(x)\) is increasing.
(ii). \(h(0) = 0\) if \(\alpha > 1\).
(iii). \(h(0) = \lambda\) if \(\alpha = 1\).
(iv). \(h(x) \to \infty\) as \(x \to \infty\).

For \(0 < \alpha < 1\) and the change point for \(h(x)\) defined by \(x^* = \beta \left(\frac{1}{\alpha} - 1\right)^{\frac{1}{\alpha}}\), we have that

(i). \(h(x)\) is decreasing for \(x < x^*\) and increasing for \(x > x^*\), this implies that the hazard rate function exhibits a bathtub shape.
(ii). \(h(x) \to \infty\) as \(x \to 0\) or \(x \to \infty\).
(iii). The change point \(x^*\) is directly proportional to \(\beta\) and inversely proportional to \(\alpha\).

We reproduced the proof of the hazard shapes of the WEM by applying Glaser’s theory (see, [10]) in Appendix A.

Survival mixture models have been widely applied as an efficient statistical tool for modeling and describing population heterogeneity in reliability, life testing applications, and in many fields of statistical analysis such as pattern recognition, clustering and survival analysis. Notably, no previous work in the literature presented the finite mixture of the Weibull extension model. This article introduces the new \(m\) -component mixture of Weibull extension distributions. The identifiability of the proposed mixture is proved. No statistical analysis is validated for a new proposed model without the latter being mathematically identifiable. Moreover, the usefulness of the model arises from the flexibility of its failure rate which accommodates increasing, bathtub and modified bathtub among other hazard patterns. These three patterns have been widely accepted in several fields, especially reliability and engineering fields. When fitting our model to two real datasets, it tends to be significantly superior to all competing distributions either one-component or two-components. It outstands the model by [7], this highlights the importance of adding scale parameters in the mixture case and that the two-component WEM mixture is a better representation for both datasets. In addition to the detailed mathematical presentation of the new mixture, identifiability and properties, we worked out a new general analytical form for the first derivative of the function \(\eta(x)\) in the mixture case. This form compares to the first derivative of the hazard rate function but is more tractable and useful to the analysis of the hazard shapes for more complex models. This proof in presented in Appendix B.

Following this introductory section, Section 2 presents the definition of the \(m\) -component mixture of WEM, its density, survival and hazard functions whereas the two-component mixture is presented in details with the shapes of the density function at various combinations of parameters. Section 3 presents the full analytical proof of the identifiability of the proposed mixture. Section 4 gives a detailed presentation of the mixture mathematical properties and a full study of its hazard function. Section 5 considers the maximum likelihood estimation of the model parameters. Finally, Section 6 gives an analysis and fit of two important datasets in reliability analysis, the Aarset data and the Meeker and Escobar data. A detailed comparison of fitting the data using the proposed model to several lifetime models is given as well.

2. Finite Mixture of the Weibull Extension Model (FMWEM)

A random variable \(X\) follows a finite mixture of WEM (FMWEM) distribution if it has the following pdf.
\[
f(x|\Psi) = \sum_{j=1}^{m} p_j f_j(x|\theta_j), \tag{2.1}
\]
where, for any \( j \in \{1, \ldots, m\} \), \( f_j(x|\theta_j) \) is the pdf of the \( j \)-th component defined by (1.1) with \( \lambda, \alpha \) and \( \beta \) replaced by \( \lambda_j, \alpha_j \) and \( \beta_j \) respectively. \( p_j \) is the mixing proportion of the \( j \)-th component with \( p_j \geq 0 \) and \( \sum_{j=1}^{m} p_j = 1 \), \( \theta_j \) is the vector of parameters of \( f_j(x|\theta_j) \) and \( \Psi = (p_1, \ldots, p_m; \theta_1, \ldots, \theta_m) \) is the whole parameters’ vector. The cdf, reliability function, hazard function of the FMWEM are given by,

\[
F(x|\Psi) = \sum_{j=1}^{m} p_j F_j(x|\theta_j), \tag{2.2}
\]
\[
S(x|\Psi) = \sum_{j=1}^{m} p_j S_j(x|\theta_j), \tag{2.3}
\]
where \( F_j(x|\theta_j) \) and \( S_j(x|\theta_j) \) are the cdf and the reliability function of the \( j \)-th component defined by (1.2) and (1.3) respectively.

The hazard rate function is the ratio of the density function to its survival function (see, [11]). The failure rate function of the m-component mixture of a WEM is given by,

\[
h(x|\Psi) = \sum_{j=1}^{m} p_j f_j(x|\theta_j), \tag{2.4}
\]
where \( p_j(x) = \frac{p_j S_j(x|\theta_j)}{\sum_{j=1}^{m} p_j S_j(x|\theta_j)} \) with \( \sum_{j=1}^{m} p_j(x) = 1 \) and \( h_j(x) \) is the hazard rate function of the \( j \)-th component defined by (1.4) with \( \lambda, \alpha \) and \( \beta \) replaced by \( \lambda_j, \alpha_j \) and \( \beta_j \) respectively.

Two-component Mixture of WEM

The two-component mixture of WEM has the following pdf:

\[
f(x) = p f_1(x) + (1 - p) f_2(x)
\]
\[
= p \lambda_1 \alpha_1 \left( \frac{x}{\beta_1} \right)^{\alpha_1-1} \exp \left( \frac{-x}{\beta_1} \right) + \lambda_1 \beta_1 (1 - e^{\left( \frac{x}{\beta_1} \right)^{\alpha_1}}) + (1
\]
\[
- p \lambda_2 \alpha_2 \left( \frac{x}{\beta_2} \right)^{\alpha_2-1} \exp \left( \frac{-x}{\beta_2} \right) + \lambda_2 \beta_2 (1 - e^{\left( \frac{x}{\beta_2} \right)^{\alpha_2}}),
\]
where, \( x \geq 0, \ alpha_j, beta_j, lambda_j \geq 0, j = 1,2. \tag{2.5} \)

Figures 1-6 illustrate various shapes of the two-component mixture density function along with the densities of its components. Intervals containing parameter values and combinations of parameters are chosen based on the previous work of the one-component WEM considered in the literature; (see, [12]). Note that: Figures legend is described as follows: “pdf” refers to the mixture density, “pdf1” refers to the first component density and “pdf2” refers to the second component density.

Let \( \gamma_j = \lambda_j / \beta_j \), Figure 1 represents the density functions for \( \gamma_j \geq 1, 0 < \alpha_j < 1 \). Figure 2 represents the density functions for \( \gamma_j \geq 1, \alpha_j > 1 \). Figure 3 represents the density functions for \( \gamma_j < 1, \alpha_j > 1 \). Figure 4 represents the density functions for \( \alpha_j = 1, \gamma_j < 1 \) and \( \gamma_j \geq 1 \). Figure 5 represents the density functions for \( \gamma_j < 1, 0 < \alpha_j < 1 \). Finally, Figure 6 displays the density functions for the case where \( \alpha_j \) for one component is >1 whereas \( \alpha_j < 1 \) for the other component. From figures 1-6, the following is observed.

- As expected, the shape of the mixture pdf is closer to the shape of the pdf of the component associated with the larger mixing proportion.
- The pdf of the two-component mixture model is very flexible. By considering various choices of \( \alpha_j \) \((0 < \alpha_j < 1, \alpha_j = 1, \text{ or } \alpha_j > 1) \) and \( \gamma_j \) \((< 1 \text{ or } \geq 1) \), the mixture model can take strictly decreasing, unimodal, bimodal, and decreasing followed by unimodal shapes among other shapes.

In summary, the shapes of the new mixture density at a wide range of parameters’ values indicate its capability to represent multiple datasets in diverse scientific areas.
Fig 1 (a). \( (p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.7, 0.6, 0.8, 120, 100, 0.15, 0.2) \)

Fig 1 (b). \( (p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.3, 0.4, 0.6, 120, 100, 0.015, 0.02) \)
Figure 2 (a). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.3, 1.55, 2, 120, 100, 0.15, 0.2)\)

Figure 2 (b). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.7, 1.55, 2, 120, 100, 0.15, 2)\)
Study of the New Finite Mixture of Weibull Extension Model: Identifiability, Properties and Estimation

Figure 2 (c). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.6, 1.2, 1.84, 100, 120, 0.01, 0.0083)\)

Figure 3 (a). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.7, 1.2, 1.5, 10, 12, 0.04, 0.06)\)
Figure 3 (b). \( (p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.67, 1.2, 1.24, 5, 7, 0.15, 0.02) \)

Figure 4 (a). \( (p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.6, 1.1, 10, 12, 0.0876, 0.05) \)
Figure 4 (b). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.4, 1, 1, 12, 10, 0.15, 0.2)\)

Figure 5 (a). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.7, 0.4, 0.6, 8, 10, 0.02, 0.08)\)
Figure 5 (b). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.3, 0.84, 0.9, 100, 100, 0.002, 0.003)\)

Figure 5 (c). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.3, 0.84, 0.9, 10, 10, 0.002, 0.003)\)
3. Identifiability of the Model

For a new proposed mixture model, a researcher can meaningfully discuss the model mathematical properties, estimation of parameters, hypotheses testing about parameters only if the family of mixing distributions is said to be identifiable. Reference [13] stated and proved the following theorem:

Theorem:
Let $A$ be the set of all finite mixing distributions related to a class $\Phi$ of probability distributions $F_i$ for $i = 1, 2, \ldots, m$. Suppose for each $F_i \in \Phi$ there will be associated a transform $\varphi_i$ having domain of definition $D_{\varphi_i}$, and suppose that the mapping $M: F_i \to \varphi_i$ is linear. Suppose there exists a total ordering $(\leq)$ of $\Phi$ such that

(i). $F_1 \leq F_2 \ (F_1, F_2 \in \Phi)$ implies $D_{\varphi_1} \subseteq D_{\varphi_2}$.
(ii). For each $F_i \in \Phi$, there exists some $t_i$ in the closure of $T_i = \{t: \varphi_i(t) \neq 0\}$ such that $\lim_{t \to t_i} (\varphi_2(t)/\varphi_1(t)) = 0$ for $F_1 < F_2 \ (F_1, F_2 \in \Phi)$.

Then the class $A$ of all finite mixing distributions is identifiable relative to $\Phi$.

Applying Chandra’s approach, we proved the following proposition.

Proposition:

*The class of all finite mixing distributions relative to the WE distribution is identifiable.*

Proof:
Let $X$ be a random variable having pdf and cdf of the WEM given in (1.1) and (1.2), respectively. Then the $k^{th}$ moment of the WEM is given by

$$
\varphi(k) = E(X^k) = \int_0^\infty x^k f(x) dx = \int_0^\infty x^k \left(\frac{\lambda \alpha}{\beta} x^{\alpha-1} \exp\left(\frac{x}{\beta}\right) + \lambda \beta \left(1 - e^{\frac{x}{\beta}}\right)\right) dx.
$$

Let $y = \exp\left(\frac{x}{\beta}\right)$, then

$$
\varphi(k) = \lambda \beta^{k+1} e^{\lambda \beta} \int_1^\infty (\ln y)^k e^{-\lambda \beta y} dy = C \int_1^\infty (\ln y)^k e^{-\lambda \beta y} dy
$$

The lower limit 1 yields a divergent integral, we must have $\left\{\frac{k}{\alpha} > -1\right\}$ for the integral to converge; i.e. $\{k > -\alpha\}$. This implies that the domain of $\varphi(k)$ is $[-\alpha, \infty[.$

Proving condition (i) in Chandra’s theorem,
From (1.2), we have

$$F_1 \leq F_2$$

\[\text{Figure 6. (p. 2, 0.2, 0.8, 100, 0.6, 0.8)}\]
\[ 1 - \exp \left\{ \lambda_1 \beta_1 \left( 1 - e^{\left( \frac{x}{\beta_1} \right)^{\alpha_1}} \right) \right\} \leq 1 - \exp \left\{ \lambda_2 \beta_2 \left( 1 - e^{\left( \frac{x}{\beta_2} \right)^{\alpha_2}} \right) \right\} \]

This ordering is fulfilled under the following conditions on the model parameters,

(1) If \( \lambda_1 = \lambda_2 = \lambda \) and \( \beta_1 = \beta_2 = \beta \), then \( \alpha_1 > \alpha_2 \) for \( x < \beta \).

And \( \alpha_1 < \alpha_2 \) for \( x > \beta \).

(2) If \( \lambda_1 = \lambda_2 = \lambda \) and \( \alpha_1 = \alpha_2 = \alpha \), then \( \beta_1 \left( 1 - e^{\left( \frac{x}{\beta_1} \right)^{\alpha}} \right) \geq \beta_2 \left( 1 - e^{\left( \frac{x}{\beta_2} \right)^{\alpha}} \right) \).

For this inequality to be valid, we must have \( \beta_1 \geq \beta_2 \).

To fulfill \( D_{\varphi_1} \subseteq D_{\varphi_2} \), we must have \( (\alpha_1 < \alpha_2) \).

Proving condition (ii) in Chandra’s theorem:

For each \( F_1 \in \Phi \), there exists some \( t_1 \) in the closure of \( T_1 \{ \varphi_1(t) \neq 0 \} \) such that \( \lim_{t \to t_1} \frac{\varphi_2(t)}{\varphi_1(t)} = 0 \) for \( F_1 < F_2 \) \((F_1, F_2 \in \Phi)\).

Proof:

Consider the integral

\[ \int_1^a \left( \ln y \right)^m e^{-y} dy = \int_1^a \left( \ln y \right)^m e^{-y} dy + \int_1^\infty \left( \ln y \right)^m e^{-y} dy, a > 1. \]

The second integral in R.H.S converges \( \forall m \). As for \( \int_1^a \left( \ln y \right)^m e^{-y} dy \), the integrand is not bounded at 1 if \( m < 0 \). Consider \( G(m) = \int_1^a \left( \ln y \right)^m e^{-y} dy \); the integrand is a continuous function in the region \( \{(y, m), y \in [c, a], m \in [-d, b], 0 < d < b \text{ finite, } 1 < c < a \} \). Then we can deduce that \( G(m) \to G(m_0) \) as \( m \to m_0 \), \( m_0 \in [-d, b] \). Applying on \( \varphi(t) \),

\[ \lim_{t \to t_1} \frac{\varphi_2(t)}{\varphi_1(t)} \big|_{c}^{a} = \int_{c}^{a} \left( \ln y \right)^{-\alpha_1/\alpha_2} e^{-\lambda \beta y} dy < e^{-\lambda \beta c} \int_{c}^{a} \left( \ln y \right)^{-\alpha_1/\alpha_2} dy. \]

Consider the following transformation

\[ x = \ln y, y = e^x, dx = e^x dx, \text{ln} c < x < \ln a \]

\[ \lim_{t \to t_1} \frac{\varphi_2(t)}{\varphi_1(t)} \big|_{c}^{a} = \int_{c}^{a} \left( \ln y \right)^{-\alpha_1/\alpha_2} e^{-\lambda \beta y} dy < e^{-\lambda \beta c} \int_{\text{ln} c}^{\ln a} x^{-\alpha_1/\alpha_2} dx = e^{-\lambda \beta c} \frac{a e^{-\lambda \beta c}}{1 - \alpha_1/\alpha_2} \left[ \ln a \right]^{-1 - \alpha_1/\alpha_2} - \left( \ln c \right)^{1 - \alpha_1/\alpha_2}. \]

As \( c \to 1^+ \), \( R.H.S. = \frac{a e^{-\lambda \beta c}}{1 - \alpha_1/\alpha_2} \left( \ln a \right)^{1 - \alpha_1/\alpha_2} \) and tends to a constant. Hence;

\[ \lim_{t \to t_1} \frac{\varphi_2(t)}{\varphi_1(t)} \big|_{c}^{a} = \int_{\text{ln} c}^{\ln a} e^{-\lambda \beta y} dy = B, \quad \text{where} \quad c \to 1^+. \]

\( B \) is some constant.

On the other hand,

\[ \lim_{t \to t_1} \frac{\varphi_1(t)}{\varphi_1(t)} \big|_{c}^{a} = \int_{\text{ln} c}^{\ln a} \left( \ln y \right)^{-\alpha_1/\alpha_2} e^{-\lambda \beta y} dy > e^{-\lambda \beta c} \int_{\text{ln} c}^{\ln a} \left( \ln y \right)^{-1} dy. \]

Based on the previous transformation \( x = \ln y \), \( R.H.S. > ce^{-\lambda \beta a} (\ln a - \ln c) \) which tends to \( \infty \) as \( c \) tends to 1.

Thus \( \lim_{t \to t_1} \frac{\varphi_1(t)}{\varphi_1(t)} \big|_{c}^{a} \to 0. \)

And \( \lim_{t \to t_1} \frac{\varphi_2(t)}{\varphi_1(t)} = 0 \), which completes the proof.

4. Properties of the FMWEM

A detailed statistical analysis of the WEM was given in [12]. Reference [14] referred to the WEM as the modified Weibull distribution. He derived explicit algebraic formulas for the \( k \text{th} \) moment (about 0) of the distribution with \( 1/\alpha \) being a non-negative integer. The following presents the mathematical properties of our proposed FMWEM.

Moment Generating Function, Mean, Mode and Median, Variance, Skewness, Kurtosis

Let \( X \) be a random variable having pdf and cdf of the WEM given in (1.1) and (1.2), respectively.

The moment generating function of the FMWEM is given by

\[ M(t) = E(e^{tX}) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \varphi(k) \quad (4.1) \]

where \( \varphi(k) \), the \( k \text{th} \) moment of the FMWEM is given by

\[ \varphi(k) = \mu'_k = \sum_{j=1}^{m} p_j \varphi_j(k) \quad (4.2) \]

where \( \varphi_j(k) \), the \( k \text{th} \) moment of the WEM is given by (see, [12])

\[ \varphi_j(k) = E(X^k) = \int_{0}^{\infty} x^k f_j(x) dx = \int_{0}^{\infty} x^k \left( \frac{x^{a_j}}{\beta_j} \int_{\frac{x^{a_j}}{\beta_j}}^{\infty} \lambda_j \alpha_j \left( \frac{x^{a_j}}{\beta_j} \right)^{a_j-1} e^{(\frac{x^{a_j}}{\beta_j})} \right) \left( 1 - e^{(\frac{x^{a_j}}{\beta_j})} \right) \right) dx. \quad (4.3) \]

Using the transformation \( y = \exp \left( \frac{x^{a_j}}{\beta_j} \right) \), where \( n = \frac{k}{a_j} \)

\[ \varphi_j(k) = \int_{0}^{\infty} y^{k/n} \left( \frac{x^{a_j}}{\beta_j} \int_{\frac{x^{a_j}}{\beta_j}}^{\infty} \lambda_j \alpha_j \left( \frac{x^{a_j}}{\beta_j} \right)^{a_j-1} e^{(\frac{x^{a_j}}{\beta_j})} \right) \left( 1 - e^{(\frac{x^{a_j}}{\beta_j})} \right) \right) dx. \quad (4.3) \]

The mean of the FMWEM is given by,

\[ \mu'_1 = E(x) = \sum_{j=1}^{\infty} p_j \mu'_1. \quad (4.5) \]
where \( m = 2 \) in the two-component case and
\[
\mu'_{1j} = \frac{\beta_j \exp(\lambda_j \beta_j)}{\alpha_j} \int_1^\infty y^{-1} (\log y)^{1/\alpha_j - 1} \exp(-\lambda_j \beta_j y) \, dy.
\] (4.6)

The mode (modes) of the FMWEM is (are) obtained by differentiating the pdf of the WEM and by solving the following non-linear equation,
\[
\sum_{j=1}^{m} \frac{\lambda_j \alpha_j}{\beta_j} x^{\alpha_j - 2} \exp\left(\frac{x}{\beta_j}\right) + \lambda_j \beta_j \left(1 - e^{\left(\frac{x}{\beta_j}\right)}\right) \times \left\{ \frac{\alpha_j}{\beta_j} x^{\alpha_j} \left(1 - \lambda_j \beta_j e^{\left(\frac{x}{\beta_j}\right)}\right) + \alpha_j - 1 \right\} = 0 \tag{4.7}
\]
where \( m = 2 \) in the two-component case.

The quartile function of the FMWEM is derived by finding the value of \( q \) for which \( F(q) = \delta \) where
\[
F(q) = \sum_{j=1}^{m} p_j F_j(q) = \delta, \tag{4.8}
\]
where \( F_j(q) = 1 - \exp\left(\lambda_j \beta_j \left(1 - e^{\left(\frac{q}{\beta_j}\right)}\right)\right) \), \( k \) is the probability of the quartile \( q \). The value of the median is the value of \( q \) for which \( \delta = 0.5 \).

The variance of a two-component mixture is defined as follows, (see, [15])
\[
V(x) = p_1 \sigma_1^2 + p_2 \sigma_2^2 + p_1 p_2 (\mu'_{11} - \mu'_{12})^2. \tag{4.9}
\]

where \( \sigma_1^2, \sigma_2^2 \) are the variances of the two components respectively, \( \mu'_{11}, \mu'_{12} \) are the means of the two components respectively.

The Fisher skewness of the two-component mixture of the WEM is given by,
\[
S_k = \frac{\mu_3}{\sigma^3}, \tag{4.10}
\]
where \( \mu_3 \) is the third central moment of the two-component mixture and is given by
\[
\mu_3 = \mu'_{3} - 3 \mu'_{1} \mu'_{2} + 2(\mu'_{1})^3, \tag{4.11}
\]
where \( \mu'_{3} \) is the 3\textsuperscript{rd} non-central moment of the two-component mixture.

The Fisher kurtosis of the two-component mixture of the WEM is given by,
\[
Kurtosis = \frac{\mu_4}{\sigma^4} - 3, \tag{4.12}
\]
where \( \mu_4 \) is the fourth central moment of the two-component mixture and is given by
\[
\mu_4 = \mu'_{4} - 4 \mu'_{1} \mu'_{3} + 6(\mu'_{1})^2 \mu'_{2} - 3(\mu'_{1})^4, \tag{4.13}
\]
where \( \mu'_{4} \) is the 4\textsuperscript{th} non-central moment of the two-component mixture.

Table (1) shows some numerical values for the skewness and kurtosis of the two-component mixture at various combinations of the model parameters.

Table 1. Skewness and Kurtosis of the two-component mixture of WEM

| Parameters | p, \( \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2 \) | Skewness | Kurtosis |
|------------|-------------------------------------------------|----------|----------|
| \( a_1 = 1 \) | \( \gamma_1 < 1 \) | 0.6; 1; 1; 10; 12; 0.0876; 0.05 | 0.814 | 0.316 |
| \( a_1 > 1 \) | \( \gamma_1 > 1 \) | 0.4; 1; 1; 12; 10; 0.15; 0.2 | 1.088 | 1.086 |
| \( a_1 < 1 \) | \( \gamma_1 < 1 \) | 0.7; 0.4; 0.6; 8;10; 0.02; 0.08 | 1.767 | 3.288 |
| \( \lambda_1 \) | 0.3; 0.84; 0.9; 100; 100; 0.002; 0.003 | 0.463 | -0.32 |
| \( \gamma_1 > 1 \) | 0.3; 0.4; 0.6; 120;100; 0.015; 0.02 | 3.17 | 17.78 |
| | 0.7; 0.6; 0.8; 10; 10; 0.2; 0.2 | 1.87 | 4.5 |
| \( a_2 > 1 \) | \( \gamma_1 < 1 \) | 0.7; 1.2; 1.5; 10; 12; 0.04; 0.06 | 0.157 | -0.655 |
| | 0.6; 1.2; 1.24; 5; 7; 0.15; 0.02 | 0.81 | -0.32 |
| | 0.3; 1.5; 5; 2; 120; 100; 0.15; 0.2 | 0.607 | 0.114 |
| \( \gamma_1 > 1 \) | 0.6; 1.2; 1.84; 100; 120; 0.01; 0.0083 | 0.244 | -0.68 |
| \( a_1 > 1 \) | \( \gamma_1 > 1 \) | 0.7; 2; 0.8; 100; 100; 0.6; 0.8 | 0.61 | -0.35 |
Hazard Rate Function

Limits of the Hazard Rate Function

The hazard rate function of the two-component WEM satisfies the following limits.

\[
\begin{align*}
\text{Lemma (1): (a) } & \lim_{x \to 0} h(x) = 0 \quad \text{if } \alpha_j > 1, j = 1, 2. \\
p\lambda_1 & \alpha_1 = 1, \alpha_2 > 1. \\
(p - 1)\lambda_2 & \alpha_1 > 1, \alpha_2 = 1. \\
p\lambda_1 + (1 - p)\lambda_2 & \alpha_1 = \alpha_2 = 1. \\
\infty & \text{otherwise.}
\end{align*}
\]

(b) \( \lim_{x \to \infty} h(x) = \infty. \)

Proof: Part (a)

For the case \( \alpha_j > 1, \) note that \( \lim_{x \to 0} h_j(x) = 0. \)

Referring to equation (2.4) for the case of \( j = 2, \)

\[ h(x) = p(x)h_1(x) + (1 - p(x))h_2(x), \]

where

\[ p(x) = \frac{pS_1(x)}{pS_1(x) + (1 - p)S_2(x)}; \]

It is easily seen that \( p(0) = p. \)

Hence \( h(0) = 0. \)

Note that

\[ \lim_{x \to 0} h_j(x) = \lambda_j \quad \text{for } \alpha_j = 1; j = 1, 2. \]

Hence the hazard rate of the three cases (i) \( \alpha_1 = 1, \alpha_2 > 1, \) (ii) \( \alpha_1 > 1, \alpha_2 = 1, \) and (iii) \( \alpha_1 = \alpha_2 = 1 \) is easily obtained.

Now consider the following three cases:

1. \( 0 < \alpha_j < 1. \)
2. \( 0 < \alpha_1 < 1, \alpha_2 \geq 1. \)
3. \( 0 < \alpha_2 < 1, \alpha_1 \geq 1. \)

For case (i): \( \lim_{x \to 0} h(x) = \infty \)

\[ \therefore \lim_{x \to 0} h(x) = \infty. \]

For case (ii): \( \lim_{x \to 0} h(x) = \infty \)

\[ \lim_{x \to 0} h_2(x) = \left\{ \begin{array}{ll}
0 & \text{if } \alpha_2 > 1 \\
\lambda_2 & \text{if } \alpha_2 = 1
\end{array} \right. \\
\therefore \lim_{x \to 0} h(x) = \infty. \]

Case (iii) follows in a similar manner to case (ii).

Part (b)

Noting that \( 0 \leq p(x) \leq 1, \) and \( \lim_{x \to \infty} h_j(x) = \infty, \) it follows that

\[ \lim_{x \to \infty} h(x) = \infty. \]

Based on the above proved limits, we can deduce that \( h(x) \) cannot be decreasing neither inverse bathtub (unimodal).

Shapes of the Hazard Rate Function

It is well known that mixtures of decreasing failure rate (DFR) distributions are always DFR (see, [16]). On the other hand, mixtures of increasing failure rate (IFR) distributions can decrease, at least in some intervals of time. Note that IFR distributions are often used to model lifetimes governed by ageing processes, which means that the operation of mixing can dramatically change the pattern of population ageing, e.g., from positive ageing (IFR) to negative ageing (DFR), (see, [17])). Therefore, the study of the shape of the observed (mixture) failure rate in a heterogeneous setting is important in many applications (reliability, demography, risk analysis, etc.).

It is worth noting that the shape of the failure rate function (and the mixture failure rate as well) is often based on the Glaser’s approach ([10]), which states that the shape of the failure rate is mainly affected by the shape of the following function:

\[ \eta(x) = -\frac{f’(x)}{f(x)}, \]

which is easier to handle in many cases than the shape of the corresponding failure rate. For instance, when \( \eta(x) \) is increasing (decreasing) the failure rate is also increasing (decreasing).

The pdf of WEM in (1.1) can be expressed in terms of exponential family of densities as follows,

\[ f(x) = C(\theta) \exp \left\{ \sum_{i=1}^{3} U_i(x, \theta) \right\}, \]

where \( \theta = \langle \lambda, \alpha, \beta \rangle \) is the vector of parameters, \( C(\theta) = \lambda \alpha \exp(\lambda \beta), \)

\[ U_1(x, \theta) = (\alpha - 1) \log(x/\beta), \quad U_2(x, \theta) = (x/\beta)^\alpha, \]

\[ U_3(x, \theta) = -\lambda \beta \exp(x/\beta)^\alpha. \]

\[ \therefore \eta(x, \theta) = -\sum_{i=1}^{3} U_i’(x, \theta). \]

\[ \eta’(x, \theta) = -\sum_{i=1}^{3} U_i''(x, \theta). \]

We reproduced the derivation of the shapes of the hazard function of the WEM in terms of the functions \( \eta(x, \theta) \) and \( \eta’(x, \theta) \) in Appendix (A). In the mixture case, reference [16] showed that

\[ \eta(x, \theta) = -[\varepsilon’(x, \theta)/\varepsilon(x, \theta) + \sum_{i=1}^{3} U_i’(x, \theta)]^{\frac{1}{\lambda}}. \]

where \( \varepsilon(x, \theta) = 1 + \frac{c_1(\alpha)}{c_2(\alpha)} \exp[\sum_{i=1}^{3} U_i(x, \theta) - \sum_{i=1}^{3} U_i’(x, \theta)], \)

\( c = (1 - \beta)/p. \]

Note that \( U_i(x, \theta), j = 1, 2, \) is defined as \( U_i(x, \theta) \) with \( \lambda, \alpha, \beta \) being replaced by \( \lambda_i, \alpha_j, \beta_j \) respectively. Using (4.16), it can be shown that

\[ \eta’(x, \theta) = a(x) \eta_1’(x, \theta) + (1 - a(x)) \eta_2’(x, \theta) - a \]

\[ (x)(1 - a(x))(\eta_2(x, \theta) - \eta_1(x, \theta))^2 \]

where \( a(x) = \frac{\varepsilon(x, \theta) - 1}{\varepsilon(x, \theta)}. \)
The proof of (4.18) is given in Appendix (B).

An analogous expression for the derivative of the hazard rate in terms of the derivatives of the components hazards was given, for example, by [18] and was used by several authors for studying and interpreting the shapes of the hazard rate. Equivalently, this can be done applying (4.18).

Figures 7-10 display multiple shapes of the two-component mixture hazard rate function along with the hazard function of the individual components at various combinations of the parameters. Figure 7 represents the hazard rate functions for \( \alpha_1 = \alpha_2 = 1 \). Figure 8 represents the hazard rate functions for \( \alpha_j > 1 \). Figure 9 represents the hazard rate functions for \( 0 < \alpha_j < 1 \). Finally, Figure 10 displays the hazard rate functions for the case where \( \alpha_j \) for one component is \( >1 \) whereas \( \alpha_j < 1 \) for the other component. In all figures of the hazard functions, the parameter \( \gamma_j \) is taken to be greater than 1 or less than 1.

Based on the limits proved of the two-component mixture hazard rate function and figures 7-10, the following comments are valid.

- Based on the proved limits of the hazard function, the latter cannot decrease or take an inverse bathtub (unimodal) shape. However, it can still decrease or have a unimodal shape for a certain interval of time. (see, figure (10))

- The mixture hazard rate tends with time to the hazard of the component associated with the smaller value for the shape parameter \( \alpha_j \). This can be explained by noting that the hazard rate of the mixture distribution always tends to the hazard rate of the stronger population (the component having the smaller hazard rate). It can be shown that the stronger population is the one having the smaller value of \( \alpha_j \) (when \( x \to \infty \)) holding the other parameter fixed.

- A mixture hazard function having one component with an increasing hazard rate (\( \alpha_j \geq 1 \)) and the second component having a bathtub shaped hazard rate (\( \alpha_j < 1 \)) may have a modified bathtub shape (increasing followed by bathtub shaped hazard) as in figure (10) for all intervals of \( \gamma_j \).

In short, the graphs of the hazard rate illustrate the flexibility of the proposed mixture to describe different failure patterns, namely increasing, bathtub, modified bathtub and decreasing-increasing-decreasing-increasing shapes.

Figures (7-10). Hazard Rate Function of the two-component mixture of WEM for different values of parameters \( \left( p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2 \right) \).

Note that: Figures legend id described as follows “h.r.m” refers to the mixture hazard function, “h.r.1” refers to the first component hazard function and “h.r.2” refers to the second component hazard function.

![Figure 7(a)](image-url)
Figure 7(b). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.4, 1, 1, 100, 120, 0.02, 0.05)\)

Figure 8(a). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.3, 1.55, 1.84, 120, 100, 0.15, 0.2)\)
Figure 8 (b). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.7, 1.7, 1.4, 10, 12, 0.2, 0.15)\)

Figure 8 (c). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.6, 1.4, 2, 10, 12, 0.04, 0.06)\)
Figure 9 (a). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.3, 0.6, 0.8, 100, 100, 2, 2)\)

Figure 9 (b). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.3, 0.84, 0.9, 100, 100, 0.002, 0.003)\)
Study of the New Finite Mixture of Weibull Extension Model: Identifiability, Properties and Estimation

Figure 9 (c). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.7, 0.9, 0.6, 100, 100, 0.002, 0.003)\)

Figure 9 (d). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.7, 0.6, 0.4, 8, 10, 0.02, 0.08)\)
Figure 10 (a). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.4, 1.4, 0.5, 5, 6, 0.05, 0.06)\)

Figure 10 (b). \((p, \alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.7, 2, 0.8, 100, 100, 0.05, 0.06)\)
Figure 11(a). Empirical TTT plot of Aarset Data

Figure 11(b). Empirical TTT plot of Meeker and Escobar Data

Figure 11. Empirical TTT plots of Aarset, Meeker and Escobar Datasets

Figure 12(a). Plot of the empirical versus effective cdfs of the two-component WEM mixture for Aarset Data
Mean Residual Life (MRL)

In industrial reliability studies of repair, replacement and other maintenance strategies, the MRL function may be more relevant than the hazard rate function. The hazard rate function relates only to the risk of immediate failure whereas, the MRL function summarizes the entire residual life distribution (see, [19]).

The MRL function of the $m$-component mixture of a WEM is given by,

$$m(x) = \sum_{j=1}^{m} p_j f_j(x) m_j(x), \quad (4.19)$$

where $m = 2$ in the two-component case, $m_j(x)$ is the MRL function of the $j^{th}$ component of the WEM (see, [20])

$$m_j(x) = \frac{1}{S_j(x)} \int_{x}^{\infty} S_j(x) \, dx = \exp(\lambda_j \beta_j e^{\frac{x}{\beta_j}}) \int_{x}^{\infty} \exp(\lambda_j \beta_j e^{\frac{x}{\beta_j}}) \, dx. \quad (4.20)$$

**5. Parameter Estimation**

Given a sample of $n$ independent observations from a mixture model defined by (2.1), the complete-data likelihood function is

$$L(\Psi) = \prod_{i=1}^{n} \prod_{j=1}^{m} p_j f_{j}(x_i | \theta_j)$$

$$= \prod_{i=1}^{n} \prod_{j=1}^{m} p_j \lambda_j \alpha_j \left(\frac{x_i}{\beta_j}\right)^{\alpha_j-1} \exp \left(\frac{x_i}{\beta_j}\right) + \lambda_j \beta_j \left(1 - e^{\frac{x_i}{\beta_j}}\right),$$

for $i = 1, \ldots, n; j = 1, \ldots, m. \quad (5.1)$

We want to get estimates that maximize $L(\Psi)$, i.e.,

$$\max L(\Psi), \quad (4.19)$$

where $\Psi = (p, \alpha_j, \beta_j, \lambda_j), \quad S = \{\Psi | 0 < p < 1; \alpha_j, \beta_j, \lambda_j > 0\}$. An estimate of $\Psi$ is said to be a feasible estimate if it belongs to the set $S$.

The natural logarithm of the likelihood function is given by,

$$\log L(\Psi) = \sum_{i=1}^{n} \log \sum_{j=1}^{m} p_j f_{j}(x_i | \theta_j) \quad (5.2)$$

The maximum likelihood estimators (MLEs) are obtained by differentiating (5.2) with respect to the parameters $p, \alpha_j, \beta_j, \lambda_j$ and equating them to 0. The normal equations are then given by,

$$\frac{\partial \log L}{\partial p_j} = \frac{1}{\sum_{i=1}^{n} \sum_{j=1}^{m} p_j f_{j}(x_i | \theta_j)} \left[ f_j(x_i, \theta) - f_m(x_i, \theta_m) \right] = 0,$$

$$\frac{\partial \log L}{\partial \alpha_j} = \frac{1}{\sum_{i=1}^{n} \sum_{j=1}^{m} p_j f_{j}(x_i | \theta_j)} \left[ \frac{\partial f_j}{\partial \alpha_j} \right] = 0,$$

$$\frac{\partial \log L}{\partial \beta_j} = \frac{1}{\sum_{i=1}^{n} \sum_{j=1}^{m} p_j f_{j}(x_i | \theta_j)} \left[ \frac{\partial f_j}{\partial \beta_j} \right] = 0,$$

$$\frac{\partial \log L}{\partial \lambda_j} = \frac{1}{\sum_{i=1}^{n} \sum_{j=1}^{m} p_j f_{j}(x_i | \theta_j)} \left[ \frac{\partial f_j}{\partial \lambda_j} \right] = 0.$$
\[
\frac{\partial \log L}{\partial \beta_j} = \sum_{i=1}^{n} p_j \frac{\partial f_j}{\partial \beta_j}(x_i, \theta_j) = 0, \quad j = 1, \ldots, m.
\]

The empirical TTT plot of both the Aarset and the MeekE datasets indicates that a bathtub shaped hazard rate function is more adequate to fit the data. Hence the two-component MWEM is an appropriate model in both cases. Table (3) summarizes the fit results of both data sets to our two-component WEM mixture, including parameters 'estimates, K-S values, p-values and estimated log-likelihood functions. Table (4) displays the results of fitting multiple lifetime distributions to the Aarset data. The competitive distributions are the additive Weibull (AddW) in [24], the WEM in [6], the modified Weibull in [25], the exponential flexible Weibull Extension (EFWE) distribution in [3], the two-component mixtures of three-parameter Weibull (M3pW) (see, [11]), exponentiated Weibull (MEW) (common scale parameters in both components) (see, [26]), Burr type III (MBurrIII) (see, [27]) and WEM are fitted as well. A reduced version of MWEM is the two-component mixture of the Chen model when the scale parameter \(\beta = 1\) for both components (MWEMR1).

Results in Table (3) assure the competence of the proposed two-component mixture of WEM to best fit both datasets in mechanical engineering. The descriptive statistics of the datasets (see, Table (2)) reveal heterogeneity and that it can be better represented by a mixture model with a hazard rate that exhibits bathtub shape. This seems evident when comparing the results of the one component distributions fitted with the results of our proposed mixture model. It is also noticed that the two-component mixture of Chen model, which is a reduced model of MWEM does not fit the data which shows the importance of including the scale parameters \(\beta_1\) and \(\beta_2\).

Table (5) shows the results of fitting multiple lifetime distributions to MeekE data. The one-component distributions fitted to the data are the additive Weibull (AddW), the WEM, the modified Weibull and the new modified Weibull (NMW) in [28]. We added the four-parameter Fréchet-Weibull (F-W) by [4] in our fit of the MeekE data. We tried to fit the latter to the Aarset data as well but estimation results were not good. Trying to fit the mixture model of F-W to the MeekE data was not a good choice based on the results of its one-component version.

The two-component mixtures of the Chen model and WEM are fitted as well. Although the six models fitted the data but the WEM mixture has the lowest K-S value and highest p-value.

Figures 12(a) and 12(b) compare the empirical cumulative distribution function computed in the K-S test with the cumulative distribution function of the two-component mixture of the WEM evaluated at the maximum likelihood estimates for Aarset and MeekE datasets respectively. The figures show minimum deviations and best representation of the data.

\[\omega(u) = \frac{1}{\mu} \int_0^{F^{-1}(u)} S(t)dt \quad 0 < u < 1. \quad (6.1)\]
### Table 2. Descriptive Statistics of Aarset, Meeker and Escobar Datasets

| Data             | N  | Min | 1st Quart. | Mean | Median | 3rd Quart. | Max |
|------------------|----|-----|------------|------|--------|------------|-----|
| Aarset           | 50 | 0.10| 13.50      | 45.69| 48.5   | 81.25      | 86  |
| Meeker & Escobar | 30 | 2   | 68.75      | 177.03| 196.5  | 298.25     | 300 |

IN: Data count

### Table 3. Results of fitting both data sets to the two-component mixture of WEM

| Data             | $\hat{p}$ | $\hat{a}_1$ | $\hat{a}_2$ | $\hat{b}_1$ | $\hat{b}_2$ | $\hat{\lambda}_1$ | $\hat{\lambda}_2$ | K-S (p-value) | Log-lik |
|------------------|-----------|-------------|-------------|-------------|-------------|---------------------|---------------------|--------------|---------|
| Aarset           | 0.598     | 0.495       | 1.4         | 13          | 14.97      | 0.025               | 11.6 \times 10^{-7} | 0.095 (0.964) | -214.55 |
| Meeker & Escobar | 0.614     | 0.575       | 0.36        | 2.993       | 2.504      | 3.9 \times 10^{-7}  | 0.0146             | 0.133 (0.9525) | -170.389 |

### Table 4. MLEs, K-S, p-values, AIC, BIC, AICC of models fitted to Aarset data

| Model            | MLEs                        | K-S | p-value | AIC   | AICC  | BIC   |
|------------------|-----------------------------|-----|---------|-------|-------|-------|
| AddW             | $\hat{a} = 1.133 \times 10^{-10}, \hat{b} = 0.086, \hat{p} = 0.477, \hat{\theta} = 4.214$ | 0.127 | 0.393   | 451   | 451.9 | 458.65 |
| ModW             | $\hat{b} = 0.062, \hat{r} = 0.356, \hat{\lambda} = 0.233$ | 0.129 | 0.38    | 460.3 | 460.86 | 466.08 |
| WEM              | $\hat{a} = 0.587, \hat{b} = 13.753, \hat{\lambda} = 0.00876$ | 0.14 | 0.281   | 469.3 | 469.81 | 475   |
| EFWE             | $\hat{a} = 0.015, \hat{b} = 0.381, \hat{\lambda} = 0.076$ | 0.158 | 0.171   | 455.664 | 456.19 | 461.4 |
| MWEM             | $\hat{\beta}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{b}_1, \hat{b}_2, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_1, \hat{\lambda}_2$ | 0.095 | 0.761   | 443.107 | 445.77 | 456.5 |
| MWEMR1           | $\hat{\beta}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{b}_1, \hat{b}_2, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_1, \hat{\lambda}_2$ | 0.44 | 0.000125 | 586.067 | 587.43 | 595.63 |
| MBurrIII         | $\hat{\beta} = 0.64, \hat{\lambda}_1 = 0.31, \hat{\lambda}_2 = 0.32, \hat{b}_1 = 1.09, \hat{b}_2 = 1.49$ | 0.78 | $1.23 \times 10^{-13}$ | 808.45 | 809.45 | 817.65 |
| M3pW             | $\hat{\beta}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{b}_1, \hat{b}_2, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_1, \hat{\lambda}_2$ | 1   | $< 2.2 \times 10^{-16}$ | 716.4 | 719.05 | 729.77 |

### Table 5. MLEs, K-S, p-values, AIC, BIC, AICC of models fitted to Meeker and Escobar data

| Model            | MLEs                        | K-S | p-value | AIC   | AICC  | BIC   |
|------------------|-----------------------------|-----|---------|-------|-------|-------|
| WEM              | $\hat{a} = 0.587, \hat{b} = 13.753, \hat{\lambda} = 0.00876$ | 0.167 | 0.38   | 364.516 | 365.44 | 368.72 |
| NMW              | $\hat{a} = 0.024, \hat{b} = 5.99 \times 10^{-8}, \hat{p} = 0.012, \hat{\theta} = 0.629, \hat{\lambda} = 0.056$ | 0.148 | 0.528 | 342.5 | 344.97 | 349.47 |
| AddW             | $\hat{a} = 1.32 \times 10^{-7}, \hat{\theta} = 0.019, \hat{\lambda} = 0.604, \hat{\beta} = 2.83$ | 0.191 | 0.23 | 364.22 | 365.8 | 369.83 |
| ModW             | $\hat{b} = 0.018, \hat{r} = 0.454, \hat{\lambda} = 7.133 \times 10^{-3}$ | 0.182 | 0.281 | 362.12 | 363.05 | 366.33 |
| F-W              | $\hat{a} = 1.424, \hat{b} = 5.8, \hat{\lambda} = 1.223, \hat{\beta} = 0.44$ | 0.996 | 1.872 \times 10^{-13} | 404.89 | 406.49 | 410.49 |
| MWEM             | $\hat{\beta} = 0.614, \hat{\lambda}_1 = 0.755, \hat{\lambda}_2 = 0.36, \hat{\beta}_1 = 2.995, \hat{\beta}_2 = 2.504, \hat{\lambda}_1 = 3.9 \times 10^{-7}, \hat{\lambda}_2 = 0.0146$ | 0.133 | 0.661 | 354.78 | 359.87 | 364.6 |
| MWEMR1           | $\hat{\beta}_1, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_2$ | 0.966 | 0.348, 3.61, 0.00163, 4.2522 | 0.233 | 0.076 | 377.74 | 380.2 | 384.7 |
## 7. Conclusion

This paper introduces a new finite mixture of a generalized form of the standard Weibull distribution, the Weibull Extension model introduced by [6]. The identifiability of the new mixture is analytically proved. For mathematical robustness of the study, the identifiability condition on the shape parameters is considered in all parameter choices and analysis. All the mathematical properties of the model are investigated. Maximum likelihood estimation of the new mixture is studied. Various combinations of the model parameters are applied to demonstrate diversity of the shapes of the density function and the hazard function. The new mixture exhibits increasing, bathtub and modified bathtub failure rates which makes the new lifetime model useful and flexible in representing many life data in practice. The two-component version of the new mixture is fitted to the Aarset and the Meeker and Escobar datasets. The model is the best fit, for both datasets, compared to many lifetime distributions, either one component or mixture distributions. In conclusion, the new FMWEM is identifiable and has tractable mathematical properties. The diversity of the density, hazard shapes and the results of fit assures the competence of our proposed mixture and usefulness in many scientific areas and applications.

## APPENDIX A:

Derivation of the shapes of the hazard rate of the WEM in terms of $\eta(x, \theta)$ and $\eta'(x, \theta)$

We applied Glaser’s theorem ([10]) to the one-component Weibull extension model using the exponential family density to prove the shapes of the hazard function.

$$f(x, \theta) = \lambda x^{\alpha-1} \exp \left\{ \frac{x}{\beta} \right\} \exp \left\{ \lambda \beta \left( 1 - e^{(x/\beta)^\alpha} \right) \right\}, \ x \geq 0; \ \alpha, \lambda, \beta \geq 0.$$

$f(x, \theta)$ can be expressed in terms of the exponential family of densities as follows

$$f(x, \theta) = C(\theta) \exp \left\{ \sum_{i=1}^k U_i(x, \theta) \right\}$$

where $\theta = \{\alpha, \lambda, \beta\}$ is the vector of parameters or constants.

$$U_1(x, \theta) = (\alpha - 1) \log \left( \frac{x}{\beta} \right)$$

$$U_2(x, \theta) = \left( \frac{x}{\beta} \right) \lambda$$

$$U_3(x, \theta) = -\lambda \beta x^{(x/\beta)^\alpha}$$

Then,

$$\eta(x, \theta) = -\sum_{i=1}^3 U'_i(x, \theta)$$

$$= - \left\{ \frac{\alpha - 1}{x} + \alpha \frac{x}{\beta^a} x^{a-1} - \lambda e^{(x/\beta)^\alpha} \left( \frac{x}{\beta} \right)^{a-1} \right\}$$

$$= \lambda e^{(x/\beta)^\alpha} \left( \frac{x}{\beta} \right)^{a-1} - \alpha \frac{x}{\beta^a} x^{a-1} + 1 - \alpha \frac{1}{x}.$$

$$\eta'(x, \theta) = -\sum_{i=1}^3 U''_i(x, \theta)$$

$$= \lambda e^{(x/\beta)^\alpha} \frac{\alpha - 1}{\beta^a-1} x^{a-2} + \lambda \left( \frac{x}{\beta} \right)^{a-1} e^{(x/\beta)^\alpha} \frac{\alpha}{\beta^a} x^{a-1}$$

$$- \alpha \frac{\alpha - 1}{\beta^a} x^{a-2} \frac{1 - \alpha}{x^2}$$

$$= \frac{1}{x^2} \left( \frac{\lambda (\alpha - 1)}{\beta^{a-1}} e^{(x/\beta)^\alpha} \lambda x^a + \frac{\alpha}{\beta^{a-1}} e^{(x/\beta)^\alpha} x^{2a} - (1 - \alpha) \right)$$

$$= \frac{1}{x^2 \beta^a x^a} \left( e^{(x/\beta)^\alpha} \left[ \beta + x^a \frac{\beta-\alpha}{\beta-a} \right] + \frac{\beta^a}{\lambda} x^{-a} \left[ 1 - \frac{1}{\alpha} \right] \right)$$

Evidently, if $\alpha \geq 1$, the derivative $\eta'(x, \theta)$ will be greater than or equal to 0; as the first two positive terms in the previous equation dominate the $3^{rd}$ negative term. This implies an increasing failure rate (IFR) for the one component WEM.

For $0 < \alpha < 1$,

Let $y = \left( \frac{x}{\beta} \right) \alpha$; \ \ $x = \beta y^{\frac{1}{\alpha}}$; \ \ $x, y \geq 0$.

$$\eta'(x, \theta)|_y = \frac{1}{x^2} \left( \lambda y^{\alpha}(\alpha - 1) ye^y + \alpha \lambda y^2 e^y + (\alpha - 1) \right)$$

$$- x(\alpha - 1) y.$$
\[
\eta'(x, \theta) = \begin{cases} 
  g_1(y) + g_2(y) & \text{if } y > \frac{1}{\alpha} - 1 \text{ or } y = \frac{1}{\alpha} \\
  0 & \text{if } y < \frac{1}{\alpha} 
\end{cases}
\]

Hence, at a certain value \( y_0 \), such that \( \frac{1}{\alpha} - 1 < y_0 < \frac{1}{\alpha} \), \( \eta'(x, \theta) = 0 \).

For \( y < y_0 \), \( \eta'(x, \theta) < 0 \).

And for \( y > y_0 \), \( \eta'(x, \theta) > 0 \).

Noting that \( \lim_{x \to 0} f(x) = \infty \), the hazard rate function for the one component WEM has a bathtub shape (BT) for \( 0 < \alpha < 1 \).

**APPENDIX B:**

Derivation of (4.18)

In the mixture case, recall that

\[ \eta(x, \theta) = -[\varepsilon'(x, \theta)/\varepsilon(x, \theta) + \sum_{i=1}^{3} U'_i(x, \theta)], \quad (4.16) \]

where

\[ \varepsilon(x, \theta) = 1 + \frac{C_i(\theta)}{C_{g_2}(\theta)} \exp \left\{ \sum_{i=1}^{3} U_{i1}(x, \theta) - \sum_{i=1}^{3} U_{i2}(x, \theta) \right\}, \quad c = \frac{(1-p)}{p} \quad (4.17) \]

Let \( C = \frac{C_i(\theta)}{C_{g_2}(\theta)} \), then

\[ \frac{\varepsilon'(x, \theta)}{\varepsilon(x, \theta)} = \frac{1}{\varepsilon(x, \theta)} C \exp \left\{ \sum_{i=1}^{3} U_{i1}(x, \theta) - \sum_{i=1}^{3} U_{i2}(x, \theta) \right\} \]

\[ \times \left\{ \sum_{i=1}^{3} U'_i(x, \theta) - \sum_{i=1}^{3} U'_{i2}(x, \theta) \right\} \]

\[ = \frac{\varepsilon(x, \theta) - 1}{\varepsilon(x, \theta)} \left\{ \sum_{i=1}^{3} U'_{i1}(x, \theta) - \sum_{i=1}^{3} U'_{i2}(x, \theta) \right\} \]

\[ = \alpha(x) \left( \eta_2(x, \theta) - \eta_1(x, \theta) \right) \]

where \( 0 < \alpha(x) = \frac{\varepsilon(x, \theta) - 1}{\varepsilon(x, \theta)} < 1 \).

\[ \frac{\varepsilon'(x, \theta)}{\varepsilon(x, \theta)} = \frac{1}{\varepsilon(x, \theta)} \left\{ C \exp \left\{ \sum_{i=1}^{3} U_{i1}(x, \theta) \right. \right. \]

\[ - \sum_{i=1}^{3} U_{i2}(x, \theta) \left. \left. \right\} \left( \eta_2(x, \theta) - \eta_1(x, \theta) \right) \right\} \]

\[ + C \exp \left\{ \sum_{i=1}^{3} U_{i1}(x, \theta) - \sum_{i=1}^{3} U_{i2}(x, \theta) \left( \eta_2'(x, \theta) \right. \right. \]

\[ - \eta_1'(x, \theta) \left. \left. \right\} \right\} \]

\[ = \frac{\varepsilon(x, \theta) - 1}{\varepsilon(x, \theta)} \left\{ \left[ \eta'_2(x, \theta) - \eta'_1(x, \theta) \right] \right. \]

\[ + \left( \eta_2(x, \theta) - \eta_1(x, \theta) \right) \right\} \]

\[ = \alpha(x) \left[ \left( \eta'_2(x, \theta) - \eta'_1(x, \theta) \right) \right. \]

\[ + \left. \left( \eta_2(x, \theta) - \eta_1(x, \theta) \right) \right\} \]

Then the first derivative of the function \( \eta(x, \theta) \) defined in (4.16) can be expressed as follows,

\[ \eta'(x, \theta) = \eta'_2(x, \theta) - \eta'_1(x, \theta) \]

\[ = a(x)\left[ \left( \eta'_2(x, \theta) - \eta'_1(x, \theta) \right) \right. \]

\[ + \left. \left( \eta_2(x, \theta) - \eta_1(x, \theta) \right) \right\} \]

(4.18)

The first derivative of the function \( \eta(x, \theta) \) in the mixture case, given in (4.18), represents an alternative to the first derivative of the hazard function \( h(x) \). This assures the ability of the function \( \eta(x, \theta) \) to interpret shapes of the failure rate function.

**REFERENCES**

[1] Jain, K., Singla, N. and Sharma, S.K., “The Generalized Inverse Generalized Weibull distribution and its properties”, Journal of Probability, Hindawi Publishing Corporation, vol. 2014, pp. 1-11, 2014. https://doi.org/10.1155/2014/736101.

[2] Khan, M.S. and King, R., “New generalized Inverse Weibull distribution for lifetime modeling”, Communications for Statistical Applications and Methods, vol. 23, no. 2, pp. 147-161, 2016. https://doi.org/10.5351/CISAM.2016.23.2.147.

[3] El-Desouky, B.S., Mustafa, A. and Al-Garash, S., “The exponentiated Flexible Weibull Extension distribution”, Open Journal of Modelling and Simulation, vol. 5, pp. 83-97, 2017. http://www.scirp.org/journal/PaperInformatio n.aspx?PaperID=73438& - abstract.

[4] Deka, D., Das B., Baruah, B. K. and Baruah, B., “Some Properties on Fréchet-Weibull Distribution with Application to Real Life Data”, Mathematics and Statistics, vol. 9, no. 1, pp. 8-15, 2021. https://doi.org/10.13189/ms.2021.090102.

[5] Ogunsanya, A. S., Yahya, W. B., Adegoke, T. M., Iluno, C., Adeere, O. R. and Ekum, M. I., "A New Three-Parameter Weibull Inverse Rayleigh Distribution: Theoretical Development and Applications," Mathematics and Statistics, vol. 9, no. 3, pp. 249-272, 2021. https://doi.org/10.13189/ms.2021.090306.

[6] Xie, M., Tang, Y. and Goh, T. N., “A modified Weibull extension with bathtub-shaped failure rate function”, Reliability Engineering & Systems Safety, vol. 73, no. 2, pp. 279-285, 2002. http://dx.doi.org/10.1016/S0951-8320(02) 00022-4.
[7] Chen, Z., “A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function”, Statistics and Probability Letters, vol. 49, pp. 155-161, 2000. https://doi.org/10.1016/S0167-7152(00)00444-4.

[8] Gupta, A., Mukherjee, B., and Upadhyay, S. K., “Weibull extension: A Bayes study using Markov chain Monte Carlo simulation”, Reliability Engineering and System Safety, vol. 93, pp. 1434–1443, 2008. https://doi.org/10.1016/j.ress.2007.10.008.

[9] Abu El Fotouh, S. and Nassar, M.M.A., “Estimation for the parameters of the Weibull Extension model based on generalized order statistics”, Int. J. Contemp. Math. Sciences, vol. 6, no. 36, pp. 1749-1760, 2011. http://www.m-hikari.com/ijcms-2011/33-36-2011/nassarIJCMS33-36-2011.pdf.

[10] Glaser, R.E., “Bathtub and related failure rate characteristics”, Journal of the American Statistical Association, vol. 75, pp. 667-672, 1980. https://www.tandfonline.com/doi/abs/10.1080/01621459.1980.10477530.

[11] Lai, C. D., “Generalized Weibull Distributions”, Springer Briefs in Statistics, Springer, 2014. https://www.springerprofessional.de/en/generalized-weibull-distributions/4124868.

[12] Tang, Y., Xie, M. and Goh, T. N., “Statistical analysis of a Weibull extension model”, Communications in Statistics, Theory and Methods, vol. 32, no. 5, pp. 913-928, 2003. DOI: 10.1081/STA-120019952.

[13] Chandra, S., “On the mixtures of probability distributions”, Scandinavian Journal of Statistics, vol. 4, pp. 105-112, 1977. https://www.jstor.org/stable/4615560.

[14] Nadarajah, S., “On the moments of the modified Weibull distribution”, Reliability Engineering and System Safety, vol. 90, pp. 114–117, 2005. http://dx.doi.org/10.1016/j.ress.2004.09.002.

[15] Frühwirth-Schnatter, S., “Finite Mixture and Markov Switching Models”. New York: Springer, 2006. https://link.springer.com/article/10.1007/s11336-009-9121-4.

[16] Barlow, R.E. and Proschan, F., “Statistical Theory of Reliability and Life Testing Probability Models”. Holt, Rinehart and Winston, New York, 1975.

[17] Finkelstein, M., “Understanding the shape of the mixture failure rate (with engineering and demographic applications)”, Appl. Stoch. Models Busin. Ind., vol. 25, pp. 643–663, 2009. https://doi.org/10.1002/asmb.817.

[18] Al-Hussaini, E.K. and Sultan, K.S., “Reliability and hazard rate functions based on finite mixture models”. In: Balakrishnan, N. and Rao, C.R., in Handbook in Statistics: Advances in Reliability, (20th ed.) Elsevier Science, Holland, 2001, pp. 139-183. https://doi.org/10.1016/S0169-7161(01)20007-8.

[19] Lai, C.D., Zhang, L. Y., and Xie, M., “Mean residual life and other properties of Weibull related bathtub shape failure rate distributions”, International Journal of Reliability, Quality and Safety Engineering, vol. 11, no. 2, pp. 113–132, 2004. https://doi.org/10.1142/S0218539304001397.

[20] Xie, M., Goh, T. N. and Tang, Y., “On changing points of mean residual life and failure rate function for some generalized Weibull distributions”. Reliability Engineering & Systems Safety, vol. 84, pp. 293–299, 2004. https://doi.org/10.1011/j.l.1467-842X.2010.00569.

[21] Aarset, M.V., “How to identify a bathtub hazard rate”, IEEE Transactions on Reliability, vol. 36, pp. 106-108, 1987. https://doi.org/10.1109/TR.1987.5222310.

[22] Meeker, W.Q. and Escobar, L.A., “Statistical Methods for Reliability Data”, 2nd ed.), New York: John Wiley, 1998. ISBN: 978-0-471-4328-4.

[23] Barlow, R.E. and Campo, R., “Total time on test processes and applications to failure data analysis”, in Reliability and Fault Tree Analysis, eds Barlow, R. E. et al., SIAM, Philadelphia, 1975, pp. 451–481. https://scholar.google.com/scholar_lookup?title=Reliability+and+Fault+Tree+Analysis&author=Barlow+R.&E.&author=Campo+R.&author=Barlow+R.+E.&publication+year=1975&pages=451-481.

[24] Xie, M. and Lai, C.D., “Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function”. Reliability Engineering and System Safety, vol. 52, pp. 87-93, 1996. https://doi.org/10.1016/0951-8320(95)00149-2

[25] Lai, C.D., Xie, M. and Murthy, D.N.P., “A modified Weibull distribution”. IEEE. Transactions on Reliability, vol. 52, no. 1, pp. 33-37, 2003. http://dx.doi.org/10.1109/TR.2002.805788

[26] Abd-Elrahman, A.M. and Mohamad, M.A., “Estimation of Mixed Exponentiated Weibull Parameters in life testing”, International Mathematical Forum, vol. 11, no. 16, pp. 769-781, 2016. http://dx.doi.org/10.12988/imf.2016.6445.

[27] Al-Moheissen, A.S., “A mixture of two Burr type III distributions: Identifiability and Estimation under type II censoring”, Mathematical Problems in Engineering, vol. 2016, no. 3, pp. 1-12, 2016. http://dx.doi.org/10.1155/2016/7035279.

[28] Almalki, S.J. and Yuan, J., “A new modified Weibull distribution”, Reliability Engineering and System Safety, vol. 111, pp. 164-170, 2013. https://econpapers.repec.org/scripts/redir.plf?u=https%3A%2F%2Fdoi.org%2F10.1016%2F0213-2521%2Fj.ress.2012.10.018&h=repec:eee:reensy:v:111:y:2013:i:6:p:164-170.