Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/110414

Modelling Reliability Data with Finite Weibull or Lognormal Mixture Distributions

*Emad E. Elmahdy**

Department of Mathematics, Science College, King Saud University, Riyadh 11451, P.O. 2455, Saudi Arabia.

Received: 24 Apr. 2017, Revised: 6 May 2017, Accepted: 9 May 2017
Published online: 1 Jul. 2017

Abstract: This paper presents a reliability analysis study of lifetime data based on Weibull and Lognormal distributions models. The main aim of this study is to compare two finite mixture distributions, Weibull mixture distribution (WMD) and Lognormal mixture distribution (LMD) for modelling heterogeneous survival data sets. This paper also provides the characterization of both WMD and LMD. A comparison of fitted cumulative distribution functions, probability density functions, hazard functions, reliability functions and the mean lifetime is obtained for different distribution models. The expectation-maximization (EM) and Levenberg-Marquardt algorithms are used for estimating the parameters of these mixture models. The goodness of fit is implemented by using different statistical methods such as the Kolmogorov-Smirnov (KS), Akaike's Information Criteria (AIC) tests and correlation coefficient to show the best fit for modelling survival data. This study give engineers some guidance for selecting the appropriate distribution model.

Keywords: Life Data Analysis; Weibull Distribution (WD); Lognormal Distribution (LND); Weibull Mixture Distribution (WMD); Lognormal Mixture Distribution (LMD); Maximum Likelihood Estimation (MLE) Method; Expectation-Maximization (EM) Algorithm; Non-linear median rank regression (NLMRR); Levenberg-Marquardt Algorithm; Goodness of Fit (GOF) Tests.

1 Introduction

Today's manufacturers exert much effort to design and produce highly reliable products to meet customers needs and respond to the increase in market competition. The reliability of a system, component, device, vehicle, and so on is the probability that it performs its function adequately for a specified period of time under the operation conditions intended [1]. An effective and widely used method of handling problems of reliability is that of accelerated life testing. This requires selecting a random sample of components of a certain product, putting on test under specified environmental conditions, and the times to failure of the individual components are observed. Reliability assessment depends on testing data at one or more levels from carefully designed experiments, statistical estimation and hypothesis tests, and model selection and validation [2]. Although life testing of components during the period of useful life is generally dependent on the exponential model, the failure rate of a component may not be constant throughout this period.

the period of initial failure may be so long that the component's main use is during this period, and in other instances the main purpose of life testing may be that of determining the time to wear-out failure rather than chance failure. In such cases, several statistical distributions can be used in analyzing lifetime data. For example, Gamma, Lognormal, and Weibull distributions are often used due to their significant usefulness in a wide range of situations [3].

The statistical analysis of lifetime data is a significant task in the discipline of reliability engineering. The Weibull distribution is the most useful distribution in modelling lifetime data, such as automobile components, electrical insulation, and ball bearings. The Weibull distribution is more popular in industrial applications than the Lognormal distribution. This may be partly because the lognormal distribution has been ignored by most reliability engineers for a long time. The lognormal distribution can be used for modelling physical phenomenon such as fatigue cracks or crack growth propagation, degradation failure modes, chemical reaction modes such as corrosion, material movement due to molecular diffusion or migration [4]. The lognormal distribution occurs whenever we encounter a random

* Corresponding author e-mail: eelmahdy@ksu.edu.sa

variable which is such that its logarithm has a normal distribution.

Many statistical methods such as moments, least squares estimation (LSE), non-linear median rank regression (NLMRR) with the Levenberg-Marquardt algorithm, maximum likelihood estimation (MLE), Bayes estimators, Monte Carlo simulation methods, and MLE using the expectation-maximization (EM) algorithm can be applied for estimating model parameters. These statistical methods may include graphical methods. In graphical methods, the estimates are obtained from plotting data where the plot depends on the selected model. They are useful for providing initial parameters estimates for different statistical estimation methods. The statistical estimation methods are more general and applicable to many kinds of models with different data types. In reliability analysis data types can be classified as complete data or incomplete data (censored data) [5,6,7,8].

Finite mixture models arise in problems of deciding between a finite number of probability distributions. They are important as probability models to describe some heterogeneous populations which can be regarded as being composed of a finite number of more homogeneous subpopulations [9,10]. In this study, two competing models are investigated, Weibull mixture distribution (WMD) and Lognormal mixture distribution (LMD) that are used to model lifetime data, it will be proved that both of them is a suitable and flexible model to analyse random durations in a possibly heterogeneous population.

The expectation-maximization (EM) algorithm is developed by Dempster, Laird and Rubin [11]. It gives us a general iteration procedure for computing MLE for parameters of mixture models. Each iteration consists of two steps: estimating expectation value for the likelihood function for data containing some missing values, these missing values may be a sub-population of a mixture distribution, and then maximizing it to get the parameter estimates of this mixture distribution. Weibull mixture distribution (WMD) modeling was studied by Elmahdy et al. [12,13]. In this study, the EM algorithm is implemented for complete and censored lifetime data to estimate the parameters of WMD.

Monte Carlo simulation can be applied in reliability analysis to implement simple relationships based on simulations. This type of simulation has different applications in risk analysis, design, quality control, etc. The Monte Carlo simulation is used for modeling today's complex systems that often involve order dependent failures [14]. The Monte Carlo simulation also presents one other important feature necessary for statistical inference of the distribution model. The Monte Carlo simulation method helps us to generate Weibull mixture distribution WMD lifetime data samples with different sizes, and studying the effects on analysis methods such as MLE using the EM algorithm and non-linear median rank regression (NLMRR) with the Levenberg-Marquardt algorithm for estimating mixing parameters of WMD [15].

The objective of this study is to compare the performance of different distributions such as three-parameterWeibull distribution (3p WD), WMD and LMD for fitting complete data random sample of a system of components by estimating the following:

- The parameters of each distribution using Maximum Likelihood Estimation (MLE) Method or Expectation-Maximization (EM) Algorithm or the Levenberg-Marquardt algorithm.
- The cumulative distribution function (CDF), probability density function (PDF), reliability function $R(t)$, hazard function $h(t)$, B-life, and mean time to failure (MTTF) of the system of components.
- The goodness of fit which is implemented by using different statistical methods such as the Kolmogorov-Smirnov (KS), Akaike's Information Criterion (AIC) tests and correlation coefficient to show the best fit for modeling survival data.

2 Modeling lifetime data

Modeling data that can be generated by some random process is a procedure for selecting a statistical distribution that best fits data. Modeling data is important for engineering design to avoid damage and errors in manufacture equipments. In general, in many industries, modeling is used for saving money and time through reducing the costs of manufacture and the ability to complete tasks and projects in specified time. In this section, some probability distributions are investigated for modeling lifetime data, such as Weibull distribution,Weibull mixture distribution (WMD), Lognormal distribution and Lognormal mixture distribution (LMD). There are some important features must be taken under consideration to model lifetime data by using Weibull mixture distributions as follows:

First, plotting cumulative distribution function (CDF) for data sample on Weibull plotting paper (WPP) and check it's fit by a smooth curve which may be concave, convex or likely S-shaped, which approaches to a straight line when data points become smaller [13]. It's suitable to select 3p WD model for concave downward curve, Weibull competing risk model for convex curve and WMD model for S-shaped curve.

Second, the possible shapes of the probability density function which depends on the parameter values, it may be decreasing, uni-modal, decreasing followed by unimodal or bimodal.

Third, The possible shapes of the hazard function which depends on the parameter values, it may be decreasing, increasing, uni-modal followed by increasing, decreasing followed by uni-modal, uni-modal followed by increasing or bi-modal followed by increasing, Glaser (1980) shows that a twofold Weibull mixture can never have a bathtub-shaped failure rate [8], Elmahdy et al. also show that the graph of the hazard function for a twofold Weibull mixture distribution can be decreasing failure rate followed by increasing [12].

2.1 Weibull Model

The three-parameterWeibull probability distribution (3p WD) is often used to model lifetime data. It has three parameters α, β and t_0. The Weibull probability density function (PDF) is defined as

$$f(t\ \beta,\alpha,t_0)=\frac{\beta}{\alpha}\left(\frac{t-t_0}{\alpha}\right)^{\beta-1}e^{-\left(\frac{t-t_0}{\alpha}\right)^{\beta}},\quad t>0 \quad (1)$$

for $\beta>0$, $\alpha>0$, and $-\infty<t_0<t$, where β is the shape parameter (determining the shape of the Weibull PDF), α (representing the characteristic life at which 63.2% of the population can be expected to have failed) and t_0 is called a threshold, location or shift parameter (sometimes called a failure-free time, minimum life or guarantee time).

If $t_0=0$ then the Weibull distribution is said to be two-parameter Weibull distribution or standard Weibull model. The PDF of the standard Weibull model is given by:

$$f(t|\beta,\alpha)=\frac{\beta}{\alpha}(\frac{t}{\alpha})^{\beta-1}\exp[-(\frac{t}{\alpha})^{\beta}],\quad t>0 \quad (2)$$

The mean life or mean time to failure ($MTTF$) is the average time that the units in a certain population are expected to operate before failure, the mean time to failure ($MTTF$) of a 3p WD is given by:

$$MTTF=t_0+\alpha\Gamma(1+\frac{1}{\beta}) \quad (3)$$

where $\Gamma(.)$ stands to gamma function.

2.2 Finite Weibull mixture model

Finite Weibull mixture models are univariate models, which can be used to model heterogeneous populations. The finite Weibull mixture model describes the density $f(t|\theta)$ as a combination of m weighted densities, which can be written as follows.

$$f(t|\theta)=\sum_{i=1}^{m}\omega_i f_i(t|\beta_i,\alpha_i) \quad (4)$$

$\theta=(\omega_1,\omega_2,...,\omega_m,\alpha_1,\alpha_2,..., \alpha_m,\beta_1,\beta_2,...,\beta_m)$ is the parameter vector for Weibull mixture distribution that including m sub-populations where $\omega_i>0$, $\alpha_i>0$, and $\beta_i>0$ denote the mixing weight, scale, and shape parameter of sub-population i respectively, $\sum_{i=1}^{m}\omega_i=1$. The probability density function of the standard Weibull model (two parameter Weibull distribution) for sub-population i is given by:

$$f_i(t|\beta_i,\alpha_i)=(\frac{\beta_i}{\alpha_i})(\frac{t}{\alpha_i})^{\beta_i-1}\exp[-(\frac{t}{\alpha_i})^{\beta_i}],\quad t>0; \quad (5)$$

therefore,

$$f(t|\theta)=\sum_{i=1}^{m}\omega_i(\frac{\beta_i}{\alpha_i})(\frac{t}{\alpha_i})^{\beta_i-1}\exp[-(\frac{t}{\alpha_i})^{\beta_i}]. \quad (6)$$

In reliability analysis, the survivor (reliability) function $R(t|\theta)$, the hazard (failure rate) function $h(t|\theta)$ and the mean life or mean time to failure ($MTTF$) of a WMD can be defined respectively as follows.

$$R(t|\theta)=\sum_{i=1}^{m}\omega_i\exp[-(\frac{t}{\alpha_i})^{\beta_i}] \quad (7)$$

$$h(t|\theta)=\sum_{i=1}^{m}\omega_i(\frac{\beta_i}{\alpha_i})(\frac{t}{\alpha_i})^{\beta_i-1} \quad (8)$$

$$MTTF=\sum_{i=1}^{m}\alpha_i\Gamma(1+\frac{1}{\beta_i}) \quad (9)$$

There's another term used for estimating percentile life which is called B-life, it's the life by which a certain proportion (B%) of the population can be expected to be failure.

2.3 Lognormal model

The Lognormal distribution has many uses in engineering, biology and economy. This model has been used to model stress failure mechanisms, such as when a failure is caused by rupture. This model has also been used to represent the sizes of fragments from a breakage process, distribution of income, distribution of a variety of biological phenomena, and as a lifetime model for electronic and electromechanical components.

If the random variable T has a lognormal distribution, then $X=\ln t$ has a normal distribution. The probability density function (PDF) of thee-parameter lognormal model can be derived from this relationship and is given by:

$$f(t\ |\mu,\sigma^2,t_0)=\frac{1}{\sqrt{2\pi}\sigma(t-t_0)}\exp\left[-\frac{1}{2}\left(\frac{\ln(t-t_0)-\mu}{\sigma}\right)^2\right],\quad t>0 \quad (10)$$

for $\sigma>0$, $-\infty<\mu<\infty$ and $-\infty<t_0<t$, where the shape is determined by the parameter σ, the scalling by the parameter μ and the location by the parameter t_0.

The Lognormal distribution is called a two-parameter Lognormal distribution when $t_0 = 0$. The PDF of the two-parameter Lognormal model is given by:

$$f(t\,|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma t}\exp\left[-\frac{1}{2}\left(\frac{\ln t-\mu}{\sigma}\right)^2\right], \quad t>0 \tag{11}$$

and the $MTTF$ of 3p LND is given by:

$$MTTF = t_0 + \exp(\mu + \frac{\sigma^2}{2}) \tag{12}$$

2.4 Finite Lognormal mixture model

The finite Lognormal mixture model describes the density $f(t|\theta)$ as a combination of m weighted densities, which can be written as follows.

$$f(t|\theta) = \sum_{i=1}^{m}\omega_i f_i(t\,|\mu_i,\sigma_i^2) \tag{13}$$

$\theta = (\omega_1,\omega_2,\ldots,\omega_m,\mu_1,\mu_2,\ldots,\ \mu_m,\sigma_1^2,\sigma_2^2,\ldots,\sigma_m^2)$ is the parameter vector of an m-mixed Lognormal distribution where $\omega_i > 0$, $\mu_i > 0$, and $\sigma_i > 0$ denote the mixing weight, and the parameters that determine the scale, and the shape of the PDF of sub-population i respectively, $\sum_{i=1}^{m}\omega_i = 1$. The probability density function of the two-parameter Lognormal model for sub-population i is given by:

$$f_i(t\,|\mu_i,\sigma_i^2) = \frac{1}{\sqrt{2\pi}\sigma_i t}\exp\left[-\frac{1}{2}\left(\frac{\ln t-\mu_i}{\sigma_i}\right)^2\right], \quad t>0 \tag{14}$$

therefore,

$$f(t|\theta)=\sum_{i=1}^{m}\omega_i\frac{1}{\sqrt{2\pi}\sigma_i t}\exp\left[-\frac{1}{2}\left(\frac{\ln t-\mu_i}{\sigma_i}\right)^2\right]. \tag{15}$$

In reliability analysis, the survivor (reliability) function $R(t|\theta)$, the hazard (failure rate) function $h(t|\theta)$ and the mean time to failure ($MTTF$) of a LMD can be defined as follows.

$$R(t|\theta)=\sum_{i=1}^{m}\omega_i R_i(t\,|\mu_i,\sigma_i^2) \tag{16}$$

where, the reliability function of the two-parameter Lognormal model for sub-population i, is given by:

$$R_i(t|\mu_i,\sigma_i^2) = 1-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{1}{2}u_i^2}du_i, \quad z=\frac{\ln t-\mu_i}{\sigma_i} \tag{17}$$

or

$$R_i(t|\mu_i,\sigma_i^2) = 1-\Phi\left[\frac{\ln t-\mu_i}{\sigma_i}\right] \tag{18}$$

where, $\Phi(\cdot)$ is the CDF of the standard Normal distribution, $\Phi(\cdot)$ is tabulated in many publications [3].

$$h(t|\theta)=\sum_{i=1}^{m}\omega_i h_i(t\,|\mu_i,\sigma_i^2) \tag{19}$$

where, the hazard function of the two-parameter Lognormal model for sub-population i, is given by:

$$h_i(t|\mu_i,\sigma_i^2) = \frac{f_i(t)}{R_i(t)} = \frac{\frac{1}{\sqrt{2\pi}\sigma_i t}\exp\left[-\frac{1}{2}\left(\frac{\ln t-\mu_i}{\sigma_i}\right)^2\right]}{1-\Phi\left[\frac{\ln t-\mu_i}{\sigma_i}\right]} \tag{20}$$

$$MTTF = \sum_{i=1}^{m}\omega_i\exp(\mu_i+\frac{\sigma_i^2}{2}) \tag{21}$$

3 Parameter estimation for mixture models

In this section, MLE using the EM algorithm and NLMRR method are introduced to estimate the paramete vector $\theta = (\theta_1,\theta_2,\ldots,\theta_i,\ldots,\theta_m)$ of an m-mixed distribution which is formed by identical distributions. Given a grouped ordered time-to-failure data $t_1,t_2,\ldots,t_n$ of identical units of a random sample of size n of a certain product are obtained from a reliability life testing experiment.

3.1 MLE using the EM Algorithm for estimating WMD model parameters

The EM algorithm for estimating parameters is an optimizing method for a log-likelihood function [11]. Given a current estimate $\theta^{(r)}$, we define the expectation of a log-likelihood function for grouped and complete ordered time-to-failure data as follows.

$$Q(\theta,\theta^{(r)})=\sum_{j=1}^{n}n_j P_i(t_j,\theta^{(r)})\ln[f(t_j|\theta)] \tag{22}$$

$P_i(t_j,\theta^{(r)})$ is the posterior probability of a unit that belongs to the i th sub-population $(i=1,2,\ldots,m)$, which failed at time t_j, where n_j denotes the number of units that failed in the jth group of the exact failure data, consequently, $n=\sum_{j=1}^{n}n_j$ is the sample size in the life test experiment. The EM algorithm is based on two main steps. The E step estimates $Q(\theta,\theta^{(r)})$ and the M step selects $\theta^{(r+1)} = Arg\ \max_{\theta}\left\{Q(\theta,\theta^{(r)})\right\}$ by equating the first derivatives of $Q(\theta,\theta^{(r)})$ with respect to each parameter $\theta_i \in \theta$ with zero. These two steps are repeated alternately until $\left|\theta^{(r+1)} - \theta^{(r)}\right| \to 0$.

By Applying EM algorithm to estimate the parameter vector of an m-mixed Weibull distribution, thereby obtaining the following recurrence relations, for censored extensions, see also E. E. Elmahdy [13].

$$\omega_i^{(r+1)} = \frac{1}{n}\left[\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)})\right] \tag{23}$$

where $P_i(t_j, \theta^{(r)}) = \frac{\omega_i^{(r)} f_i(t_j|\beta_i^{(r)}, \alpha_i^{(r)})}{\sum_{i=1}^{m} \omega_i^{(r)} f_i(t_j|\beta_i^{(r)}, \alpha_i^{(r)})}$ is the posterior probability.

$$\alpha_i^{(r+1)} = \left[\frac{\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)})(t_j)^{\beta_i^{(r+1)}}}{\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)})}\right]^{1/\beta_i^{(r+1)}} \tag{24}$$

$$g(\beta_i^{(r+1)}) = \frac{1}{\beta_i^{(r+1)}} + \frac{\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)}) \ln(t_j)}{\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)})} - \frac{\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)})(t_j)^{\beta_i^{(r+1)}} \ln(t_j)}{\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)})(t_j)^{\beta_i^{(r+1)}}} = 0 \tag{25}$$

By taking a good initial estimate of $\theta^{(r)}$ and then solving Eqs. (23), (24) and (25) numerically, we can find MLE estimates of $\omega_i^{(r+1)}$, $\beta_i^{(r+1)}$ and $\alpha_i^{(r+1)}$ for sub-population i . For further illustration, see the proposed algorithm for estimating the parameters of the WMD for modelling complete failure data [12].

3.2 MLE using the EM Algorithm for estimating LMD model parameters

Given a current estimate $\theta^{(r)}$, the expectation of a log-likelihood function for Lognormal mixture distribution (LMD) can be defined as

$$Q(\theta, \theta^{(r)}) = \sum_{j=1}^{n}\sum_{i=1}^{m} n_j P_i(t_j, \theta^{(r)}) \ln[\omega_i f_i(t_j|\mu_i, \sigma_i^2)] \tag{26}$$

$$Q(\theta, \theta^{(r)}) = \sum_{j=1}^{n}\sum_{i=1}^{m} n_j P_i(t_j, \theta^{(r)}) \ln(\omega_i) + \sum_{j=1}^{n}\sum_{i=1}^{m} n_j P_i(t_j, \theta^{(r)}) \ln(f_i(t_j|\mu_i, \sigma_i^2)) + \lambda(\sum_{i=1}^{m} \omega_i - 1) \tag{27}$$

where λ is the lagrange multiplier, on condution that $\sum_{i=1}^{m} \omega_i = 1$. EM algorithm is based on two steps, first step is finding this expectation which is called the E step and the second one is finding the value $\theta^{(r+1)}$ of θ that maximizes $Q(\theta, \theta^{(r)})$ which is called the M-step (the maximization step).

We can find $\omega_i^{(r+1)}$ of ω_i which maximizes $Q(\theta, \theta^{(r)})$, by taking the derivative of Eq. (27) with respect to ω_i equal to zero, :

$$\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)}) + \lambda\omega_i = 0 \tag{28}$$

Summing both sides over i and using the fact that $\sum_{i=1}^{m} P_i(t_j, \theta^{(r)}) = 1$, we get $\lambda = -n$, consequently

$$\omega_i^{r+1} = \frac{1}{n}\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)}) \tag{29}$$

To obtain the value $\mu_i^{(r+1)}$ of μ_i which maximizes $Q(\theta, \theta^{(r)})$, taking the derivative of Eq. (27) with respect to μ_i equal to zero, we get:

$$\frac{\partial Q(\theta, \theta^{(r)})}{\partial \mu_i} = 0 \tag{30}$$

$$\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)}) \frac{\partial \ln(f_i(t_j|\mu_i, \sigma_i^2))}{\partial \mu_i} = 0 \tag{31}$$

$$\sum_{j=1}^{n} \frac{n_j P_i(t_j, \theta^{(r)})[\ln t_j - \mu_i^{(r+1)}]}{(\sigma_i^2)^{(r+1)}} = 0 \tag{32}$$

$$\mu_i^{(r+1)} = \frac{\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)}) \ln t_j}{\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)})} \tag{33}$$

Also, we can find the value $(\sigma_i^2)^{(r+1)}$ of σ_i^2 which maximizes $Q(\theta, \theta^{(r)})$, taking the derivative of Eq. (27) with respect to σ_i^2 equal to zero, we get:

$$\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)}) \frac{\partial \ln(f_i(t_j|\mu_i, \sigma_i^2))}{\partial \sigma_i^2} = 0 \tag{34}$$

$$\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)}) . \frac{-1}{2(\sigma_i^2)^{(r+1)}} [1 - \frac{1}{(\sigma_i^2)^{(r+1)}} (\ln t_j - \mu_i^{(r+1)})^2] = 0 \tag{35}$$

$$(\sigma_i^2)^{(r+1)} = \frac{\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)})(\ln t_j - \mu_i^{(r+1)})^2}{\sum_{j=1}^{n} n_j P_i(t_j, \theta^{(r)})} \tag{36}$$

Taking a good initial guess of $\theta^{(r)}$, consequently knowing $P_i(t_j, \theta^{(r)})$, and by updating Eqs. (29), (33) and (36) until $\left|\theta^{(r+1)} - \theta^{(r)}\right| \to 0$ we can solve these Eqs. numerically to find MLE estimates of $\omega_i^{(r+1)}$, $\mu_i^{(r+1)}$and $(\sigma_i^2)^{(r+1)}$ of subpopulation i.

3.3 NLMRR for estimating the parameters of mixture models

Here, the parameter optimization can be implemented for LMD and WMD models using the Levenberg-Marquardt algorithm. The Levenberg-Marquardt algorithm is a non-linear iterative optimization method that can be used to minimize the sum of squares for the residuals due to error, *SSR* [15]. When regression analysis is applied to the WMD model to estimate its parameters, **MATLAB** program can be used for non-linear median rank regression (NLMRR), which is based on the modified Levenberg-Marquardt algorithm and median rank method. Estimates of the parameters in Eq. (7) and Eq. (16) are required to fit lifetime data with the WMD or LMD models. These parameters can be evaluated by using *SSR*, which can be defined as:

$$SSR = \sum_{i=1}^{n} (R_i - \hat{R}_i)^2 \quad (37)$$

where $\hat{R}_i$ denotes the approximated value of the reliability function which can be calculated using Eqs. (7), (16), (17) and (18) and R_i is the actual value of the reliability function at t_i, which can be determined by plotting a probability graph for the given lifetime data on WPP using various methods, such as the median rank method, Kaplan–Meier, or Benard's median rank [6,13]. The required parameter estimates θ are the values that minimize *SSR*.

Eq. (37) can be written as:

$$SSR = E^{'}.E \quad (38)$$

where $E_i = R_i - \hat{R}_i$, $E = \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_i \\ \vdots \\ E_n \end{bmatrix}$ and $E^{'}$ is the transpose of E

Marquardt techniques depends on finding the gradient of *SSR* with respect to the set of parameters θ as follows:

$$\frac{1}{2}\frac{\partial (E^{'}.E)}{\partial \theta} = -X^{'}.R + X^{'}.R(\theta) = -X^{'}.E \quad (39)$$

where X is $m \times n$ matrix including the partial derivatives of R with respect to the parameters, $X = \frac{\partial R}{\partial \theta}$, and E is $n \times 1$ matrix including the error at each data point. The gradient method can be used to determine the best direction of moving in the θ_i space to obtain the smallest sum of squares for the residuals due to error as follows:

$$\theta_{i+1} = \theta_i + \kappa.X^{'}.E \quad (40)$$

where κ is a control variable adjusts how far to move in the direction opposite to the gradient for updating the parameter values, but this method can not specify how far to move for finding the optimal solution. Levenberg-Marquardt technique treats this problem by using Gauss-Newton method. This method assumes that $R(\theta)$ can be expanded in θ_i space by using Taylor series about θ_0 as follows:

$$R(\theta) = R(\theta_0) + X.(\theta - \theta_0) + \ldots \quad (41)$$

By taking under consideration only the linear terms in the above equation and assuming that θ are the exact parameter values i.e there's no error, therfore by the aids of Eq. (39), one can deduce that:

$$X^{'}.[R(\theta_0) + X.(\theta - \theta_0)] = X^{'}.R \quad (42)$$

Consequently, the updating formula of Gauss-Newton method can be written as:

$$\theta_{i+1} = \theta_i + (X^{'}.X)^{-1}.X^{'}.E \quad (43)$$

The Levenberg-Marquardt algorithm combines these two methods through the following general formula:

$$\theta_{i+1} = \theta_i + (X^{'}.X + \lambda.I)^{-1}.X^{'}.E \quad (44)$$

where λ is the scaling parameter which balances the gradient-steepest-decent and Gauss-Newton methods. The optimal solution is obtained by adjusting λ and taking a good initial values for the parameters. The Levenberg-Marquardt algorithm is a stable, efficient and easily programmable.

4 Goodness of fit (GOF) tests

The goodness of fit is implemented by using different statistical methods such as the Kolmogorov-Smirnov (KS), Akaike's Information Criterion (AIC) tests and correlation coefficient to show the best fit for modeling survival data. When inferences are to be based on a statistical model, it is of course important to be satisfied as to the appropriateness of the model. As a minimum the model should be consonant with data in regard to goodness of fit tests and other assessment procedures [2]. Inference for the selected model also depends on sample sizes and in what the model is used. Goodness of fit tests are necessary for making a decision for selecting the best model but not sufficient. Studying physical failure analysis and prior engineering experience are required.

4.1 The Kolmogorov-Smirnov (KS) test

Consider X be a random variable with distribution function $F(x)$, let $x_1, x_2, \ldots, x_i, \ldots, x_n$ be an ordered random sample then the empirical distribution function can be defined as

$$\bar{F_n}(x_i) = \frac{i}{n}, 1 \leq i \leq n \tag{45}$$

and the Kolmogorov-Smirnov (KS) statistics can be defined as

$$D_n = \max_{1 \leq i \leq n} [\frac{i}{n} - F_0(x_i)] \tag{46}$$

where $F_0(x)$ stands to a specified family of distribution models that can fit complete (uncensored) data, i.e., the parameters of $F_0(x)$ are determined. It's clear that KS statistics measures the distance between $\bar{F_n}(x_i)$ and $F_0(x_i)$, the best fit has the smallest distance.

4.2 Akaike's Information Criterion (AIC)

Akaike's Information Criterion (AIC) is defined as

$$AIC = -2\ln(l(\hat{\theta})) + 2K \tag{47}$$

where $\ln(l(\hat{\theta}))$ is the natural logarithm of the maximum likelihood for the proposed model and k is the number of independently adjusted parameters for the model. Since, AIC is dependent on the maximum likelihood function, it's effective and unbiased for large sample data of size more than 30. To select the best fitting distribution among competing models, AIC is calculated for each one, the best fit has the minimum AIC [12].

4.3 The least squares fit criterion

The objective function $J(\theta)$ can be defined as

$$J(\theta) = \sum_{i=1}^{n} [y(x_i;\theta) - y_i]^2 \tag{48}$$

where x_i and y_i , $1 \leq i \leq n_i$ are the transformed values of the data set and $y(x_i;\theta)$ are the Weibull transformed values for the model with parameter vector θ

The squared value for the correlation coefficient r can be defined as

$$r^2 = 1 - \frac{J}{S} \tag{49}$$

where S represents the sum of squares of the deviation of y values from their mean $\bar{y}$, for the best fit, $J \approx 0$ and consequently $r \approx 1$.

5 Application

In Wire Fatigue Experiment [6], forty-eight stranded stainless steel wire was ruptured by clamping the wire in needle nose pliers and hanging a 1.65 pound weight on it, using $3/4$ liter of water, followed by a 2.2 pound weight on it, using one liter of water. The pliers were rotated through 180 degrees, alternating clockwise and counterclockwise. The number of half twists to total rupture (failure) was recorded. The wire data in Table 1 represent the number of half twists to total rupture for Wire Fatigue Experiment. Table 2 presents the GOF tests: KS, AIC and r estimates obtained for different competing distribution models: 2p WD, 2p LND, 3p WD, 3p LND, WMD and LMD respectively. It's found in Table 2 that the smallest AIC for WMD, the smallest KS for LMD and the largest r for LMD, thus LMD is the most closely fit these data. Figures 1-4 show clearly that 3p WD or WMD and LMD are reasonable to model the life data. Figure 1 also shows that B-0.1 or 10% life for all illustrated models is about 12.6. Table 2 also shows MTTF estimations for the competing distribution models: 2p WD, 2p LND, 3p WD, 3p LND, WMD and LMD respectively, clearly, the estimated values are very closed. Figure 2 shows that the shape of probability density function is bimodal for both WMD and LMD but it's uni-modal for 3p WD. Figure 3 shows a comparison of estimated reliability functions obtained for 3p LND, WMD and LMD models. Figure 4 shows the different possible shapes for the graph of the hazard function (failure rate), the shape of hazard function is uni-modal followed by increasing for both WMD and LMD where it's increasing rapidly for WMD and It's continuous increasing and concave downward for 3p WD where it has a horizontal asymptote as $x \to \infty$.

Table 1. Wire data set:

37	30	27	51	10	24	15	14
34	34	42	25	15	13	16	12
27	21	37	35	18	17	17	13
35	27	41	41	14	17	20	16
28	24	32	24	17	19	15	20
45	39	27	33	18	13	11	13

Table 2. GOF tests and MTTF results for different distribution models:

Model	GOF tests			MTTF
	KS	AIC	r	
2p WD	0.1458	364.1615	0.9456	24.0789
2p LND	0.0978	356.7795	0.9762	24.5654
3p WD	0.0746	354.2064	0.9909	24.5114
3p LND	0.0895	357.7758	0.9873	24.7562
WMD	0.0653	**353.8896**	0.9912	24.1536
LMD	**0.0486**	354.1050	**0.9939**	24.4801

6 Conclusion

This paper presented stable and efficient approaches that are easily programmable for modeling lifetime data such

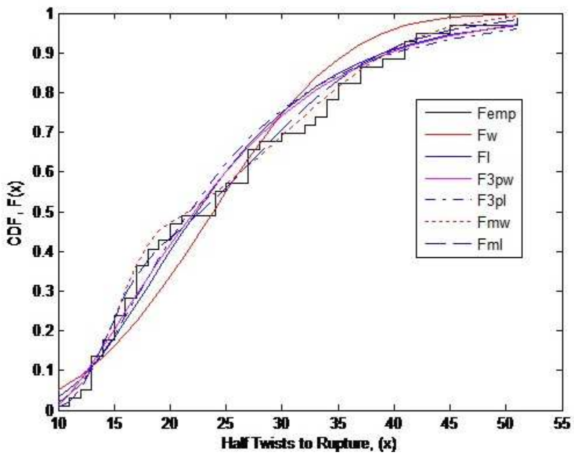

Fig. 1: A comparison of fitted CDFs obtained for different life data distribution models.

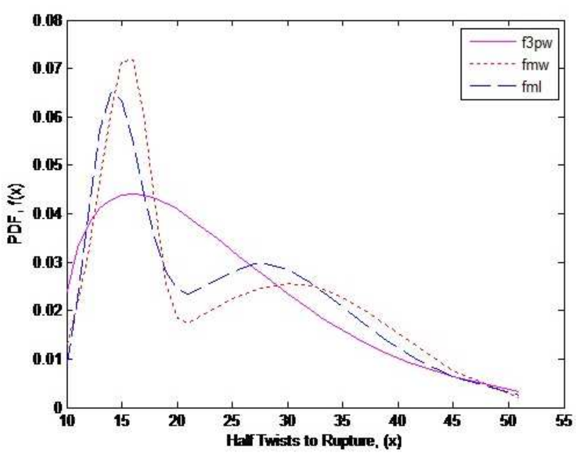

Fig. 2: A comparison of fitted PDFs obtained for different life data distribution models.

as EM Algorithm and Levenberg-Marquardt algorithm. A numerical application is implemented through the proposed algorithms, accurate parameters estimates are obtained. This paper also presents a comparison of the fitted CDFs, PDFs, reliability functions and hazard functions of Weibull, Lognormal, Weibull mixture (WMD) and Lognormal mixture (LMD) models. Goodness of fit (GOF) based on different statistical methods such as the Kolmogorov-Smirnov (KS), Akaike's Information Criteria (AIC) tests and correlation coefficient are used to select the best distribution for modeling life data. The WMD and LMD are considered as competing distribution models, both of them are reasonable to fit life data with accuracy. Also, for selecting the best model to fit life data, the physical

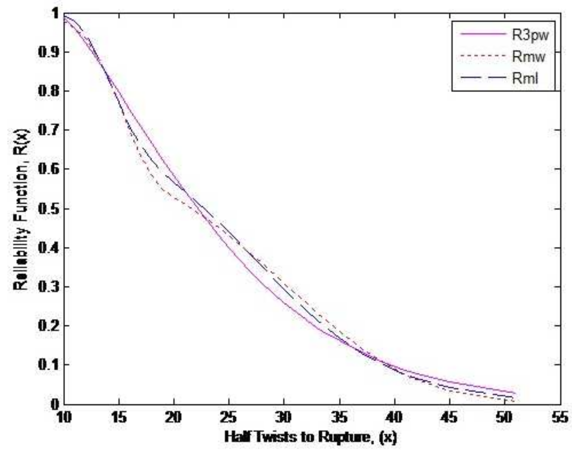

Fig. 3: A comparison of fitted reliability functions obtained for different life data distribution models.

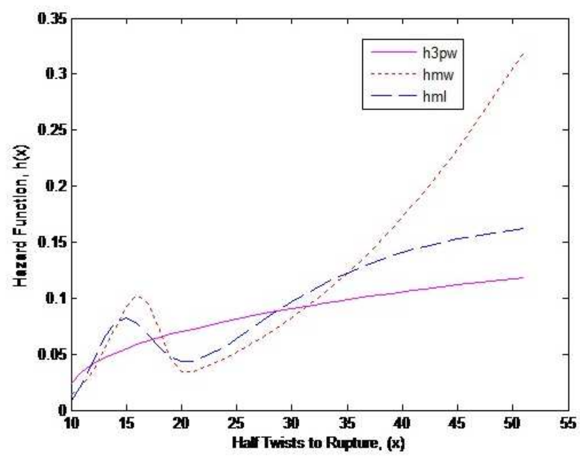

Fig. 4: A comparison of fitted hazard functions obtained for different life data distribution models.

failure analysis must be considered besides GOF tests. In this paper also, some important features and evidences are considered to justify the choice of a twofold mixture distribution model for modeling lifetime data by using WMD or LMD such as the shape of PDF which can be appeared as bi-modal and the graph of the hazard function for a twofold mixture distribution WMD or LMD which can be uni-modal followed by increasing, WMD or LMD can never have a bathtub-shaped failure rate.

Acknowledgement

The author thanks the editor and the anonymous referees for their valuable comments and suggestions which have helped to greatly improve this paper. The author extends his appreciation to King Saud University, Deanship of Scientific Research, College of Science Research Center for supporting this project.

References

[1] W. Q. Meeker and l. A. Escobar, Statistical methods for reliability data, New York Wiley, 1998.
[2] J. F. Lawless, Statistical Models and Methods for Lifetime Data, second ed., Wiley, 2002.
[3] R. A. Johnson, I. Miller, and J. E. Freund, Probability and Statistics for Engineers, eighth ed., Pearson Education Limited Co., UK, 2014.
[4] P. A. Tobias, and D. C. Trindade, Applied Reliability, second ed., Van Nostrand Reinhold Co., New York, 1995.
[5] W. Nelson, Applied Life Data Analysis, John Wiley & Sons, Inc., New York, 1982.
[6] R. B. Abernethy, The New Weibull Handbook: Reliability & Statistical Analysis for Predicting Life, Safety, Survivability, Risk, Cost, and Warranty Claims (Fifth ed.), Florida, 2010**.**
[7] D. Kececioglu, Reliability & Life Testing Handbook, Prentice Hall, Inc., Englewood Cliffs, New Jersey,Vol. 1 and 2, 1993 and 1994.
[8] D. N. Prabhakar Murthy, M. Xie and R. Jiang, Weibull Models , John Wiley & Sons, Inc., 2004.
[9] B. S. Everitt and D. J. Hand, Finite Mixture Distribution, Chapman & Hall, London,1981.
[10] J. S. Maritz and T. Lwin, Emperical Bayes Methods, second ed., Chapman & Hall, London,1989.
[11] A. P. Dempster, N. M. Laird and D. B. Rubin, Maximum likelihood estimation from incomplete data via the EM algorithm, J. Roy. Statist. Soc. Ser. B 39 (1977) 1–38.
[12] E. E. Elmahdy and A. W. Aboutahoun, A new Approach for Parameter Estimation of Finite Weibull Mixture Distributions for Reliability Modelling, Applied Math. Modelling 37(4) Feb. (2013) 1800–1810.
[13] E. E. Elmahdy, A new approach for Weibull modeling for reliability life data analysis, Applied Mathematics and Computation 250 (2015) 708–720.
[14] Salvatore J. Bavuso; NASA; Hampton, Aerospace Applications of Weibull and Monte Carlo Simulation with Importance Sampling, IEEE, Annual Reliability and Maintainability Symposium, Proc. 1997.
[15] E. E. Elmahdy, Reliability analysis for Monte Carlo simulation using the expectation-maximization algorithm for a Weibull mixture distribution model, J Appl Computat Math, 5(3) June (2016).
[16] P. M. Lee, Bayesian Statistics: An Introduction, Third ed., Oxford University Press Inc., New York, 2004.
[17] A. Gelman, J. B. Carlin, H. S. Stern and D. B. Rubin, Bayesian Data Analysis, second ed., Chapman & Hall, CRC, 2003.
[18] E. G. Tsionas, Bayesian analysis of finite mixtures of Weibull distributions, Commun. Statist.Theory Methods 31 (1) (2002) 37–48.
[19] E. T. Lee and J. W. Wang, Statistical Methods for Survival Data Analysis, John Wiley & Sons Inc; (3rd Edition), 2003.

Emad E Elmahdy received the PHD degree in applied mathematics at Tanta University (Egypt). He's assistant professor of applied mathematics, science college at King Saud University. He's interested in Weibull modelling and mathematical physics subjects. He has many papers in reliability analysis. He is referee of several international scientific journals.