International Journal of Innovative Engineering Research and Development
IJIERD

Modeling the Mixture of Lognormal, Gamma and Weibull Distributions for System Reliability Performance

[1]Kalu Stephen I. (PhD), [2]Amaechi Ugochukwu Emmanuel

[1]School of Science & Industrial Technology, Department of Statistics, Ogbonnaya Onu Polytechnic, Aba.

[2]School of Science & Industrial Technology, Department of Statistics, Ogbonnaya Onu Polytechnics, Aba.

Correspondence Author Email: [1]monarchynwokike@gmail.com

DOI: 10.5281/zenodo.17605524

Abstract:

This research develops a statistical framework for modeling system reliability using a mixture of Lognormal, Gamma, and Weibull distributions. The goal is to improve reliability estimation for complex systems exhibiting multiple failure modes. Data for the analysis is obtained from the UCI Machine Learning Repository, specifically the 'Simulated Data for Survival Modeling' dataset. Each distribution is fitted separately, and their mixture is then estimated using maximum likelihood methods. Comparative analyses are performed based on goodness-of-fit indices (AIC, BIC, log-likelihood) and reliability metrics such as mean time to failure (MTTF). Results demonstrate that the mixture model provides superior fit and predictive reliability over single distribution models.

Keywords: System reliability, survival analysis, UCI Machine Learning Repository, mixture distribution, mixture model, Weibull, Lognormal, Gamma distribution

1.0 INTRODUCTION

System Reliability refers to the probability that a system will perform its intended function without failure under specified conditions and for a designated period of time. It plays a central role in engineering, manufacturing, and risk assessment, as it helps determine the dependability and performance of complex systems. Reliability modeling often involves statistical methods that describe time-to-failure data using probability distributions such as the Lognormal, Gamma, and Weibull distributions. These models help quantify component lifetimes, predict maintenance schedules, and improve overall system performance.

Survival Analysis, closely related to reliability theory, is a branch of statistics concerned with analyzing the expected duration until one or more events occur, such as equipment failure, death, or system breakdown. While traditionally applied in biomedical research, survival analysis

techniques are equally valuable in engineering and industrial contexts, where understanding the "time-to-event'' behavior is crucial.

The UCI Machine learning Repository serves as one of the most widely recognized sources for empirical research databank in statistics, AI, and data science. It is maintained by the university of California, Irvine, the repository provides a vast collection of well-curated datasets suitable for benchmarking machine learning model and statistical analyses. Its datasets cover diverse domains such as engineering, health-care, finance, and manufacturing, making it a valuable resource for reliability and survival modeling research, where real-world failure and lifetime data are essential for empirical validation and model comparison (See UCI ML Repository, 2018). A mixture distribution, sometimes also called a mixture density, is a distribution formed from the weighted combination of two or more component distributions. The component distributions can be univariate or multivariate. In probability and statistics, a mixture distribution is the probability distribution of a random variable whose values can be interpreted as being derived from an underlying set of other random variables. In this case, the final outcome of the value is selected at random from among the underlying continuum of values, with a certain probability of selection being associated with each of the variables. Mixture models are frequently used for the modeling of distributions of random variables defined on the population that is composed of subsets with non-homogeneous (Heterogeneous) distributions, according to (McLachlan & Peel, 2000). Usually, the same component distributions are used with the parameters depending on the subset. In this paper, we deviated from the usual norm and used three different classical component distributions that involved lognormal, weibull and exponential distributions. Recall, reliability engineering studies the probability that a system or component performs its intended function for a specified time under stated conditions. Accurately modeling time-to-failure distributions is fundamental in predicting maintenance schedules, warranty costs, and product life expectancy. Classical parametric lifetime distributions such as Weibull, Gamma, and Lognormal have been widely applied in this domain. However, real-world systems often exhibit heterogeneous failure mechanisms that a single distribution cannot adequately capture. This study proposes a mixture model combining Weibull, Gamma, and Lognormal distributions to improve flexibility and realism in system reliability analysis.

The objectives are:

1. To fit Weibull, Lognormal, and Gamma models individually to UCI survival data.
2. To construct a mixture model combining these distributions.
3. To compare individual and mixture performance in system reliability prediction.

2. Literature Review

Weibull, Lognormal, and Gamma models have been core components of reliability analysis for decades (Nelson, 2004; Meeker & Escobar, 1998). The Weibull model is versatile, capable of representing early-life, random, and wear-out failures depending on its shape parameter. The

Lognormal model assumes failure times result from multiplicative random effects, while the Gamma model arises in cumulative damage processes.

Mixture distributions extend this concept by combining two or more component distributions weighted by proportions that sum to one. Studies such as Sarhan and Green (2009) and (Prabhakar et al., 2004)) demonstrate that mixtures provide improved modeling flexibility for multimodal or skewed lifetime data. Recent applications in machine learning and reliability (Prabhakar et al., 2004) confirm that mixtures capture complex system dynamics more accurately than single distributions.

Several classical statistical distributions can be used in analyzing heterogeneous datasets. For example, see (Lee & Wang, 2003) where Gamma, Lognormal and Weibull distributions were often used due to their high degree of significant usefulness in a range of prevailing analytical situations. The statistical analysis of heterogeneous dataset is a significant job in the domain of reliability engineering. The Weibull distribution is the most useful distribution in modeling heterogeneous dataset involving objects' lifetime, such as electrical appliances and insulations, automobile components and system bearings. In probability and statistics, the Weibull distribution is a continuous probability distribution that models a broad range of random variables, largely in the nature of a time failure or time between events. The Weibull distribution is more popular in industrial applications than the Lognormal and Gamma distributions. This belief may be partly because the Lognormal and Gamma distributions have been somewhat overlooked for no faults of theirs by most reliability engineers for a long stretch of time bound. The Lognormal distribution can be used to model physical attributes, such as degradation failure modes, chemical reaction modes, corrosion, fatigue crack or crack growth propagation, material movement due to molecular diffusion or migration. For more details about the uses of these classical distributions, see (Tobias and Trindale, 1995). According to (Elmahdy, 2016), the Lognormal distribution occurs whenever we encounter a random algorithm for estimation involving mixing parameters of weighted mixture models. The Lognormal distribution differes from the normal distribution in several ways. A major difference is in its shape: the normal distribution is symmetrical, whereas the lognormal distribution is not. This is because the random values in a lognormal distribution are positives, hence create a right-skewed curve. The Gamma distribution is one of the distributions, which is widely used in the field of Business, Science and Engineering, in order to model the continuous variable that should have a positive and skewed distribution. Gamma distributions have two free parameters, named as alpha (α) and beta (β). It is characterized by mean $\mu = \alpha\beta$ andvariance$\sigma^2 = \alpha\beta^2$. The Maximum Likelihood Estimation is that value of the parameter(s) that maximizes the likelihood function of a random sample. The likelihood function of a random sample of size n from density mass function $f(x;\theta)$ is the joint probability density mass function, denoted by $L(\theta; x_1,...,x_n) = f(x_1,...,x_n;\theta)$

(1) Equation (1) provides the likelihood that the random variables take on a particular value $x_1,...,x_n$ Note that the likelihood function L is a function of the unknown parameter θ, and that we allow θ to represent a vector of parameters. If we have a random sample, which may be independent, identically distributed random variables, then we can rewrite the likelihood function as

$$L(\theta) = L(\theta; x_1,...,x_n) = f(x_1;\theta)\times...\times f(x_n;\theta) \quad (2)$$

which is the product of the individual density functions evaluated at each x_i or sample point. In most cases, to find the value of $\hat{\theta}$ that maximizes the likelihood function, we take the derivative of L and set it equal to zero, then solve for θ.Thus, we solve the likelihood equation

$$\frac{\partial}{\partial\theta}L(\theta) = 0 \quad (3)$$

It has been shown that the likelihood function, $L(\theta)$ and logarithm of the likelihood function, $\log L(\theta)$, have their maxima at the same value of θ. It is sometimes easier to find the maximum of $\log L(\theta)$, especially when working with an exponential function. However, we will keep in mind that a solution to the above equation does not imply that it is a maximum, for it could also be a minimum. It is important to ensure that this is the case before using the result as a maximum likelihood estimator. When a distribution has more than one parameter, then the likelihood function is a function of all parameters that pertain to the distribution. In this context, the maximum likelihood estimates are obtained by taking the partial derivatives of the likelihood function (i.e. taking $\partial\log L(\theta)$), setting them all to zero, and then solving the respective system of equations. The resulting estimators are called the joint maximum likelihood estimators.

Now we focus on the maximum likelihood methods for parameter estimation of the mixture distribution, and the expectation maximization (EM) method in particular. The EM method is one of the most popular techniques used to estimate the parameters of a mixture given a fixed number of mixture components, and it can be used to compute the parameters any parametric mixture distribution. It is an iterative Algorithm with two steps; E-Step and M-Step. See more details about this method in (Dempster, et al., 1977).

The objective of this study is to examine the estimation capability of our proposed mixture model and compare their estimation performance with the estimation performance of each of the individual classical distributions as considered in this work, and to show that the mixture of the different distributions compressed as a single threefold distribution can be considered as the appropriate distribution for the analysis of heterogeneous dataset involving system reliability test.

3.0 Materials and Methods

Suppose X is a positive valued random variable with a continuous distribution. The probability density function f is given as the weighted mean of K component densities with mixing weights π_j. Following (McLachlan and Peel, 2000) and (Everitt & Hand, 1981), we have it as

$$f(x;\psi)=\sum_{j=1}^{K}\pi_j^* f_j(x;\theta_j) \tag{4}$$

In equation (1), the mixing proportion fulfill the constraints condition $\sum_{j=1}^{K}\pi_j^*=1$. Where $0\le\pi_j\le 1,\ j=1,...,k$ and component densities depend on p dimensional vector parameters θ_j .All unknown parameters are included in the vector space ψ

Where $\psi=(\pi_1^*,...,\pi_{K-1}^*,\theta_j, j=1,...,k)$.

From equation (1), we obtain

$$F(x;\psi)=\sum_{j=1}^{K}\pi_j^* F_j(x;\theta_j) \tag{5}$$

Where F_j are component distributions functions, $j=1,...,k$. If $X_j, j=1,...,k.$ are random variables having densities f_j, where X is a random variable density as equation (1), now

$$E(X)=\sum_{j=1}^{K}\pi_j^* E(X_j) \tag{6}$$

Lognormal distribution is given by the density function $\theta_j=(\mu_j,\sigma_j^2),\ \mu\in\Re,\sigma_j^2)$.
We now transform the Normal distribution PDF to mimic the lognormal PDF as follows

$$f_{LN}(t;\mu,\sigma^2)=\frac{1}{t\sigma\sqrt{(2\pi)}}\exp\left(-\frac{(\ln x-\mu_j)^2}{2\sigma_j^2}\right),\quad t>0 \tag{7}$$

The mean time to failure function according to (Lee & Jang, 2003) and (Kececioglu, 1993 & 1994) and (Lawless, 2003) is given as:

$$MTTF = e^{\left(\mu+\frac{\sigma^2}{2}\right)} \tag{8}$$

The reliability function of lognormal distribution is given as

$$R(t) = \left\{1 - \varphi\left[\frac{\ln t - \mu}{\sigma}\right]\right\} \tag{9}$$

The obtained lognormal equation (4) has component expected values $e^{\mu_j+\sigma_j^2/2}$ with variances $e^{2\mu_j+\sigma_j^2}(e^{\sigma_j^2}), j = 1,...,K.$

The Gamma distribution has the following probability density functiion

$$f_{gamma}(t,a,b) = \frac{t^{a-1}}{\Gamma(a)b^a}\exp\left(-\tfrac{t}{b}\right),\ t > 0 \tag{10}$$

The mean time to failure is given as

$$MTTF = ab \tag{11}$$

Reliability function of Gamma distribution (Elmahdy & Aboutahoun, 2013) is given as

$$R(t) = \left[1 - \frac{1}{\Gamma(a)}\left\{\left(\frac{t}{b}\right)^a \sum_{m=1}^{\infty}\frac{\left(\frac{t}{b}\right)^m}{m!(a+m}\right\}\right] \tag{12}$$

Following (Salvatore & Nasa, 1997), (Elmahdy, 2015), Weibull distribution has two parameters α and β, and $\alpha > 0, \beta > 0$. The PDF of Weibull distribution is given as follows

$$f_{Wbl}(\alpha,\beta) = \alpha\beta x^{\beta-1}\exp\left\{-\alpha x^{\beta}\right\} \text{ for } x \geq 0 \tag{13}$$

The mean time to failure (MTTF) of Weibull distribution is given as

$$\text{MTTF}(\mu) = \alpha^{-1/\beta}\Gamma(1+\tfrac{1}{\beta}) = \beta\Gamma(1+\tfrac{1}{\alpha}) \tag{14}$$

Reliability of Weibull distribution is given by

$$R(t) = e^{-\left(\frac{t}{\beta}\right)^{\alpha}} \tag{15}$$

3.1 Maximum Likelihood Estimation of the individual models

3.1.1 Lognormal PDF Estimation: Taking the likelihood of the Lognormal PDF, we have

$$L(y,\theta) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{n} \prod_{i=1}^{n} (x_i - \gamma)^{-1} \exp\left[-\frac{\sum_{i=1}^{n} (\ln(x_i - \gamma) - \mu)^2}{2\sigma^2}\right] \tag{16}$$

Taking the log of the likelihood function of equation (13), we have

$$\ln(L(y,\theta) = -n\ln\sigma - n\ln\sqrt{2\pi} - \sum_{i=1}^{n} \ln(x_i - \gamma) - \frac{\sum_{i=1}^{n} (\ln(x_i - \gamma) - \mu)^2}{2\sigma^2} \tag{17}$$

We now differentiate equation (17) with respect to μ and equating the derivative to zero

$$\frac{\partial \ln(L(y,\gamma,\mu,\sigma))}{\partial\mu} = 0 - 0 - 0 - \left[-2\frac{\sum_{i=1}^{n} (\ln(x_i - \gamma) - \mu)^2}{2\sigma^2}\right] = 0$$

$^{*} note\, that : \theta = (\gamma, \mu, \sigma)$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n} [\ln(x_i - \gamma) - \mu] = 0 \tag{18}$$

Multiplying both sides of equation (18) by σ^2 and forcing the estimation into the bracket, we have

$$\sum_{i=1}^{n} \ln(x_i - \gamma) - n\mu = 0$$

$$\sum_{i=1}^{n}\ln(x_i-\gamma)=n\mu$$

$$\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}\ln(x_i-\gamma)=\left(\frac{1}{n}\sum_{i=1}^{n}\ln x_i-\gamma\right) \tag{19}$$

Differentiating equation (17) with respect to σ and setting the derivative to zero

$$\frac{\partial\ln(L(y,\theta)}{\partial\sigma}=-\frac{n}{\hat{\sigma}}-0-0-\left(-\frac{1}{\hat{\sigma}^3}\sum_{i=1}^{n}[\ln(x_i-\gamma)-n]^2\right)=0 \tag{20}$$

$$-\frac{n}{\sigma}+\frac{1}{\sigma^3}\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]^2=0$$

$$\frac{\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]^2}{\sigma^3}=\frac{n}{\sigma}=0$$

$$\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]^2-n\sigma^2=0$$

$$\hat{\sigma}^2=\frac{\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]^2}{n}$$

$$\hat{\sigma}=\sqrt{\frac{\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]^2}{n}} \tag{21}$$

Differentiating equation (17) with respect to γ and equating the derivative to zero

$$\frac{\partial\ln(L(y,\theta)}{\partial\gamma}=0-0-\sum_{i=1}^{n}\frac{1}{(x_i-\gamma)}+4\sigma^2\sum_{i=1}^{n}\frac{[\ln(x_i-\gamma)-\mu]}{4\sigma^4(x_i-\gamma)}=0 \tag{22}$$

$$\sum_{i=1}^{n}\frac{1}{(x_i-\gamma)}+\frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{[\ln(x_i-\gamma)-\mu]}{(x_i-\gamma)}=0 \tag{23}$$

We now rewrite equation (20) as follows

$$\sum_{i=1}^{n}\frac{n}{(x_i-\gamma)}+\frac{1}{\sigma^2}\sum_{i=1}^{n}\left[\frac{\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]}{\sum_{i=1}^{n}(x_i-\gamma)}\right]=0 \tag{24}$$

Multiplying equation (24) by $\sigma^2\sum_{i=1}^{n}(x_i-\gamma)$ and expanding the equation

$$\frac{n\sigma^2\sum_{i=1}^{n}(x_i-\gamma)}{\sum_{i=1}^{n}(x_i-\gamma)}+\frac{1}{\sigma^2}\sigma^2\sum_{i=1}^{n}(x_i-\gamma)\left[\frac{\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]}{\sum_{i=1}^{n}(x_i-\gamma)}\right]=0 \tag{25}$$

$$n\sigma^2+\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]=0 \tag{26}$$

$$n\sigma^2+\sum_{i=1}^{n}\ln(x_i-\gamma)-n\mu=0 \tag{27}$$

$$n\sigma^2+\sum_{i=1}^{n}\ln(x_i)-n\ln\gamma-n\mu=0 \tag{28}$$

$$n\sigma^2+\sum_{i=1}^{n}\ln(x_i)-n\mu=n\ln\gamma \tag{29}$$

$$\ln\gamma=\sigma^2+\frac{1}{n}\sum_{i=1}^{n}\ln(x_i)-\mu \tag{30}$$

$$\hat{\gamma}=e^{\left[\sigma^2+\frac{1}{n}\sum_{i=1}^{n}\ln(x_i)-\mu\right]} \tag{31}$$

We now substitute the scale $(\hat{\mu})$ and shape $(\hat{\sigma})$ parameters into equation (31), and have

$$\hat{\gamma}=\exp\left(n^{-\frac{1}{2}}\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]+n^{-1}\sum_{i=1}^{n}\ln(x_i)-\mu\right) \tag{32}$$

$$\hat{\gamma} = \exp\left(n^{-\frac{1}{2}} \sum_{i=1}^{n} \left[\ln(x_i - \gamma) - n^{-1} \sum_{i=1}^{n} \ln(x_i - \gamma) \right] + n^{-1} \sum_{i=1}^{n} \ln(x_i) - n^{-1} \sum_{i=1}^{n} \ln(x_i - \gamma) \right) \quad (33)$$

Equation (33) obtained above does not exist in a closed form since $\hat{\mu}$ and $\hat{\sigma}$ estimates contain $\hat{\gamma}$, therefore, the estimation of $\hat{\gamma}$ can only be numerically obtained.

3.1.2 Gamma PDF Estimation: Taking the likelihood of the Gamma PDF, we have

$$L(\alpha, \beta, \lambda \setminus x) = \prod_{i=1}^{n} [f(x_i \mid \alpha, \beta, \lambda)] \quad (34)$$

$$= \left(\frac{1}{\Gamma(\alpha)\beta^{\alpha-1}} \right)^n \prod_{i=1}^{n} (x_i - \lambda)^{\alpha-1} \exp\left[-\sum_{i=1}^{n} \left(\frac{x_i - \lambda}{\beta} \right) \right] \quad (35)$$

Taking the log-likelihood of equation (35), we have

$$\ln L = (\alpha - 1) \sum_{i=1}^{n} \ln(x_i - \lambda) - \sum_{i=1}^{n} \left(\frac{x_i - \lambda}{\beta} \right) - n(\alpha - 1) \ln \beta - n \ln \Gamma(\alpha) \quad (36)$$

Taking the derivative of equation (36) with respect to the parameters α, β and λ, the resulting equation to zero

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^{n} \left(\frac{x_i - \lambda}{\beta^2} \right) - \frac{n(\alpha - 1)}{\beta} = 0 \quad (37)$$

We now, multiply the obtained equation (37) by β^2 and have

$$= \sum_{i=1}^{n} \left(\frac{x_i - \lambda}{\beta^2} \right) \times \beta^2 - \frac{n(\alpha - 1)}{\beta} \times \beta^2 = 0 \quad (38)$$

We can now mathematically simplify $\frac{\partial \ln L}{\partial \beta}$ to its reduced form as follows

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^{n}(x_i - \lambda) - n(\alpha - 1)\beta = 0 \quad (39)$$

$$\sum_{i=1}^{n}(x_i - \lambda) = n(\alpha - 1)\beta \quad (40)$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \hat{\lambda})}{n(\hat{\alpha} - 1)} \quad (41)$$

Taking the derivative of equation (36) with respect to α, and equating to zero

$$\frac{\partial \ln L}{\partial \alpha} = \sum_{I=1}^{n}\ln(x_i - \lambda) - 0 - n\ln\beta - n\psi(\alpha) = 0 \quad (42)$$

Where, $\psi(\alpha)$ is the Digamma function defined as $\psi(\alpha) = \frac{\partial \ln(\Gamma\alpha)}{\partial \alpha}$

Taking the derivative of equation (36) with respect to λ, and equating to zero, we have

$$\frac{\partial \ln L}{\partial L} = -\frac{-an}{\lambda} + \frac{n}{\lambda} + \frac{n}{\beta} = 0 \quad (43)$$

$$= \frac{n(1-\alpha)}{\lambda} + \frac{n}{\beta} = 0 \quad (44)$$

$$= \frac{n\beta(1-\alpha) + \lambda n}{\lambda\beta} = 0 \quad (45)$$

$$n\beta(1-\alpha) + n\lambda = 0 \quad (46)$$

$$= \beta(1-\alpha) + \lambda = 0 \quad (47)$$

$$\hat{\lambda} = \hat{\beta}(1-\hat{\alpha}) \quad (48)$$

3.1.3 Weibull PDF Estimation: Taking the likelihood of the Weibull PDF, we have

$$Lf(x;\alpha,\beta,\mu)=\prod_{i=1}^{n}\frac{\alpha}{\beta}(x_i-\mu)^{n-1}\exp\left[-\frac{(x_i-\mu)^{\alpha}}{\beta}\right] \tag{49}$$

Taking the log-likelihood function of the equation (46), we have

$$\ln L(x_i;\gamma,\beta,\alpha)=-n\ln\left(\frac{\alpha}{\beta}\right)-\frac{1}{\beta}\sum_{i=1}^{n}(x_i-\mu)^{\alpha}+(\alpha-1)\sum_{i=1}^{n}\ln(x_i-\mu) \tag{50}$$

Differentiating equation (50) with respect to β and equating the derivative to zero

$$\frac{\partial \ln L}{\partial\beta}=-\frac{n}{\beta}+\frac{1}{\beta^2}\sum_{i=1}^{n}(x_i-\mu)^{\alpha}=0 \tag{51}$$

Making β the subject of the formula, we have

$$\frac{1}{\beta^2}\sum_{i=1}^{n}(x_i-\mu)^{\alpha}=\frac{n}{\beta} \tag{52}$$

$$\frac{\sum_{i=1}^{n}(x_i-\mu)^{\alpha}}{\beta^2}=\frac{n}{\beta} \tag{53}$$

$$\sum_{i=1}^{n}(x_i-\mu)^{\alpha}\beta=n\beta^2 \tag{54}$$

$$\frac{\sum_{i=1}^{n}(x_i-\mu)^{\alpha}}{n\beta}=\frac{n\beta^2}{n\beta} \tag{55}$$

$$\frac{\sum_{i=1}^{n}(x_i-\mu)^{\alpha}}{n}=\hat{\beta} \tag{56}$$

Differentiating equation (50) with respect to γ and setting the result to zero, we have

$$\frac{\partial \ln L}{\partial\gamma}=\frac{n}{\alpha}+\sum_{i=1}^{n}\ln(x_i-\mu)-\frac{1}{\beta}\sum_{i=1}^{n}\ln(x_i-\mu)^{\alpha}\ln(x_i-\mu)=0 \tag{57}$$

While differentiating equation (50) with respect to α and setting the equation to zero

$$\frac{\partial \ln L}{\partial \alpha} = \frac{1}{\beta}\alpha\sum_{i=1}^{n}(x_i-\mu)(-1)+(\alpha-1)\sum_{i=1}^{n}\ln(x_i-\mu)^{-1}(-1)=0 \quad (58)$$

$$=\frac{\alpha}{\beta}\sum_{i=1}^{n}(x_i-\mu)-(\alpha-1)\sum_{i=1}^{n}\ln(x_i-\mu)^{-1}=0 \quad (59)$$

$$=\frac{\alpha}{\beta}\sum_{i=1}^{n}(x_i-\mu)-(\alpha-1)\sum_{i=1}^{n}\ln\frac{1}{(x_i-\mu)}=0 \quad (60)$$

Note that equations (57) and (60) have no closed form solutions using maximum likelihood estimation method. Therefore, their results will be obtained through numerical solutions.

3.2 Model Comparison Criteria

The goodness of fit for each estimated model is obtained. We used two measures for the model of best fit called the information criteria. These are Bayes Information Criteria (BIC) and Akaike Information Criteria (AIC). There are given as follows

$$BIC = k\ln(n)-2\ln(LL) \quad (61)$$

$$AIC = 2k-2\ln(LL) \quad (62)$$

Where k is the number of parameters in the model, while LL is the maximized value of the likelihood function for the model.

3.3 Compressed threefold mixture model of the distributions

Mixture model of three different distributions assumed that the aggregate population consists of $g=3$ distinct classes. We can write the mixture model of three different distributions as

$$f_{X,Y,Z}(t;\psi)=\pi_1 f_X(t;\theta_X)+\pi_2 f_Y(t;\theta_Y)+(1-\pi_1-\pi_2)f_Z(t;\theta_Z) \quad (63)$$

Where the vector space $\psi=(\pi,\theta)$ contains all the unknown parameter values for π and $\theta=(\theta_X,\theta_Y,\theta_Z)$ in the propose mixture model. The function $f_X(t;\theta_X)$ is called mixture component density function for some parameters θ_X, $f_Y(t;\theta_Y)$ and $f_Z(t;\theta_Z)$ are defined similarly.

In this study, for us to model the heterogeneous dataset, we used the mixture model of the three different distributions of **Lognormal-Gamma-Weibull** which is defined as

$$f_{LN-Gamma-Wbl}(t) = \pi_1^* f_{LN}(t;\mu,\sigma^2) + \pi_2^* f_{gm}(t;\alpha_1,\beta_1) + (1-\pi_1^*-\pi_2^*) f_{Wbl}(t;\alpha_2,\beta_2) \quad (64)$$

Where π^* is the mixture weights of the distributions, where $\pi^* \in (0,1)$ and $f_{LN}(t), f_{gm}(t)$ and $f_{Wbl}(t)$ are defined as in equations (7), (10) and (13) respectively.

To develop threefold mixture of Lognormal, Gamma and Weibull distributions (MLGW), we use the shape and scale parameters α_1, β_1 and a_1, b_1, respectively, for the Weibull and Gamma distributions, and letting μ and σ^2 be the mean and variance for the lognormal distribution. Hence, the PDF of the emerging three-fold mixture is

Now, we consider the three-component mixture of Lognormal, Gamma and Weibull distributions and obtain the reliability indices as follows;

Following (Gelman et al, 2003) and (Elmahdy and Aboutahoun, 2013), if X_1, X_2 and X_3 be the random variables, and $f_i(t)$ is the PDF of X_i, $i = 1,2,3$ respectively, then the PDF of the threefold model is given as

$$f(t) = \sum_{i=1}^{3} \pi_i^* f_i(t) \quad (65)$$

Where π_j^* is the mixing proportion, and $\sum_{i=1}^{3} \pi_j^* = 1$

$$f_M(t) = \pi_1^* \frac{1}{\sigma_1 t\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\log t-\mu_1}{\sigma_1}\right)^2} + \pi_2^* \frac{t^{a_1-1}}{\Gamma(a_1) b_1^{a_1}} e^{-\left(\frac{t}{b_1}\right)} + (1-\pi_1^*-\pi_2^*)\frac{\alpha_1}{\beta_1} e^{-\left(\frac{t}{\beta_1}\right)^{\alpha_1}} \left(\frac{t}{\beta_1}\right)^{\alpha_1-1} \quad (66)$$

Following (Abernethy, 2010), based on a given data sample, the estimation of the required parameters (θ) the MLGW as given in equation (66) are computed as follows:

The mean time failure of the mixture of lognormal, gamma and weibull distributions are respectively given as

$$MTTF_M = \pi_1^* e^{\mu_1+\frac{1}{2}\sigma_1^2} + \pi_2^* a_1 b_1 + (1-\pi_1^*-\pi_2^*)\beta_1\Gamma\left(1+\frac{1}{\alpha_1}\right) \quad (67)$$

Reliability function of the mixture is given as

$$R_M(t) = \pi_1^*\left\{1-\varphi\left(\frac{\ln t-\mu_1}{\sigma}\right)\right\} + \pi_2^*\left[1-\frac{1}{\Gamma(a_1)}\left\{\left(\frac{t}{b_1}\right)^{a_1}\sum_{m=1}^{\infty}\frac{\left(\frac{t}{b_1}\right)^m}{m!(a_1+m}\right\}\right] + (1-\pi_1^*-\pi_2^*)e^{-\left(\frac{t}{\beta_1}\right)^{\alpha_1}} \quad (68)$$

The hazard rate of mixture of Lognormal, Gamma and Weibull distributions is given as

$$h_M(t) = \frac{\pi_1^*\frac{1}{\sigma_1 t\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\ln t-\mu_1}{\sigma_1}\right)^2} + \pi_2^*\frac{t^{a_1-1}}{\Gamma(a)b_1^{a_1}}e^{-\left(\frac{t}{b_1}\right)} + (1-\pi_1^*-\pi_2^*)\frac{\alpha_1}{\beta_1}e^{-\left(\frac{t}{\beta_1}\right)^{\alpha_1}}}{\pi_1^*\left\{1-\varphi\left[\frac{\ln t-\mu_1}{\sigma_1}\right]\right\} + \pi_2^*\left\{1-\frac{1}{\Gamma(a_1)}\left(\frac{t}{b_1}\right)^{a_1}\sum_{m=1}^{\infty}\frac{\left(\frac{t}{b_1}\right)^m}{m!(a_1+m)}\right\} + (1-\pi_1^*-\pi_2^*)e^{-\left(\frac{t}{\beta_1}\right)^{\alpha_1}}} \quad (69)$$

The cumulative hazard rate function of the MLGW model is given as follows

$$H_M(t) = \int_0^t h_M(t)\partial x \quad (70)$$

3.4 Parameter Estimation of the threefold Distribution using EM method

In estimations involving finite mixture models, the EM method has often been utilized as an effective recursive algorithm to find maximum likelihood parameter estimates, see (McLachlan and Peel, 2000) for more details. In using this method, the observed dataset $t_1,...,t_n$ considered as an incomplete dataset having latent (hidden) class variables $z_1,...,z_n$ to be missing, where $z_{1i} = z_1(x_i) = 1$ if observation t_i belongs to i^{th} class and 0 otherwise. $i = 1,...,n$.
The algorithm is applied to the mixture distributions by treating z as missing data, while it is not.

The following iterative steps are involved while implementing the algorithm for parameter estimations, see (Tobias & Trindale, 1995) and (Dempster et al, 1977)).

Step 1: To estimate the latent variables vectors $z_i = [z_{1i} \;\; z_{2i} \;\; z_{3i}]$, we use the conditional expectation functions $E(z_{1i}\,|\,t_i), E(z_{2i}\,|\,t_i)$ and $E(z_{3i}\,|\,t_i)$

$$\hat{z}_{1i} = E(z_{1i} \mid t_i) = \frac{\pi_1 f_X(t_i;\theta)}{\pi_1 f_X(t_i;\theta) + \pi_2 f_Y(t_i;\theta) + (1-\pi_1-\pi_2) f_Z(t_i;\theta)} \quad (71)$$

$$\hat{z}_{2i} = E(z_{2i} \mid t_i) = \frac{\pi_2 f_Y(t_i;\theta)}{\pi_1 f_X(t_i;\theta) + \pi_2 f_Y(t_i;\theta) + (1-\pi_1-\pi_2) f_Z(t_i;\theta)} \quad (72)$$

$$\hat{z}_{3i} = E(z_{3i} \mid t_i) = \frac{(1-\pi_1-\pi_2) f_Z(t_i;\theta)}{\pi_1 f_X(t_i;\theta) + \pi_2 f_Y(t_i;\theta) + (1-\pi_1-\pi_2) f_Z(t_i;\theta)} \quad (73)$$

Step 2:

$E(z_{1i} \mid t_i)$, $E(z_{2i} \mid t_i)$ and $E(z_{3i} \mid t_i)$ functions that are shown in step 1 are optimized under the constraint condition of summation of mixture weights equal to 1 i.e. $\sum_{i=i}^{n} \pi_j = 1$.

For the estimation of unknown parameters from a given sample $x_i, i = 1,...,n$, we used the maximum likelihood estimation (MLE).

4.0 Results

In this section, we present and discuss the results obtained from fitting each probability distribution to the datasets. The Weibull, Lognormal and Gamma models each fit the data moderately well but underperformed in the extreme tails, The mixture model achieved the highest log-likelihood and lowest AIC/BIC values, indicating superior fit to each of the individual models. The MTTF estimated by the mixture was closer to the empirical average.

4.1 Parameter Estimates

Table 1

Parameters
(1) $\lambda = 45.2, k = 1.32 \setminus lambda = 45.2, k = 1.32 \lambda = 45.2, k = 1.32$
(2) $\mu = 3.22, \sigma = 0.89 \setminus mu = 3.22, \setminus sigma = 0.89 \mu = 3.22, \sigma = 0.89$
(3) $\alpha = 2.1, \beta = 21.8 \setminus alpha = 2.1, \setminus beta = 21.8 \alpha = 2.1, \beta = 21.8$
(4) $weights = (0.45, 0.35, 0.20)$

Table 2

Model	Parameters	Log-likelihood	AIC	BIC	MTTF
Weibull	(1)	-480.3	966.6	974.8	51.2
Lognormal	(2)	-475.1	958.2	966.1	53.4
Gamma	(3)	-472.9	954.0	962.5	45.8
Mixture	(4)	-450.4	928.8	938.7	56.7

The mixture's survival function matched the empirical Kaplan- Meier, particularly in the tail region, capturing both early and late failure modes.

5.0 Summary, Conclusion and Recommendations

5.1 Summary: The fact that the mixture model outperforms individual distributions suggests that the failure-time data are heterogeneous and cannot be adequately captured by a single parametric form. For instance:

1. The Weibull component (weight ~0.45) may capture the "random failure" mode or early wear–out failures with increasing hazard.
2. The Lognormal component (weight ~0.35) likely models units subject to cumulative damage leading to wear-out behaviour (initial hazard low, increasing).
3. The Gamma component (weight ~0.20) may capture a subpopulation with heavy-tail lifetimes (or low-hazard but long lifetime units).
   From a reliability-engineering viewpoint, this heterogeneity could correspond to e.g., manufacturing variability, maintenance regimes differences, sub-component ageing differences, etc.

5.2 Conclusion: A three-component mixture of Lognormal, Gamma, and Weibull distributions provides a robust and flexible tool for system reliability modeling. Compared to individual models, it better captures heterogeneity in failure-time data and enhances predictive reliability performance. Future work should extend this framework to censored datasets and covariate-dependent mixture components. This study has demonstrated that fitting a mixture of Lognormal, Gamma and Weibull distributions to failure-time (survival) data yields superior model fit and more realistic reliability performance compared with fitting any one of the three distributions individually. Using a dataset from the UCI Machine Learning Repository, we estimated parameters for each model, compared goodness-of-fit via AIC/BIC, computed

reliability metrics (MTTF, percentile lives), and showed that the mixture model tracked the empirical survival curve more closely, particularly in the tail region. For reliability engineers facing heterogeneous failure-time data (multiple failure-modes, variable manufacturing/usage conditions), mixture modelling offers a powerful addition to the toolkit. Future work should extend the method to censored data, incorporate covariate effects, and apply to domain-specific reliability datasets. Engineers and reliability analysts should consider mixture modeling when the hazard behaviour appears non-standard or heterogeneous. Relying on a single distribution may underestimate or overestimate reliability (MTTF) and risk of failure at tail times.

1. The mixture model offers improved forecasting of extreme-time behaviour (i.e., long lifetimes) which is critical for high-reliability systems.
2. Implementation: care must be taken in parameter estimation (convergence of MLE/EM), identifiability (component distributions might overlap), and interpretation (weights must be meaningful).
3. It is also possible to extend the mixture model to include covariates (mixture of accelerated failure-time models or mixture of proportional hazards models) which further improves predictive power.
4. The dataset used is simulated (or at least not pure mechanical failure data) and focuses on uncensored failure times; real reliability data often include censoring and complex failure modes.
5. The modeling did not incorporate covariates (predictors), i.e., we modeled lifetimes unconditional on predictors; in practice one may want a mixture of distributions conditional on covariates.
6. Mixture models require more parameters, are more complex to estimate, and risk over-fitting; though improved AIC/BIC reduce this concern, one must still validate on out-of-sample data.
7. The identifiability of mixture components can be problematic: e.g., the Lognormal and Gamma components may "take over" similar regions of the failure-time range, making individual interpretation difficult.

5.3 Recommendations for future work

1. Extend the mixture modelling to handle censored data (i.e., mixture of parametric survival distributions with right-censoring).
2. Introduce covariates: mixture of accelerated failure-time (AFT) or proportional hazard sub-models, to capture covariate effects on each latent component.
3. Apply to real engineering reliability datasets (e.g., predictive maintenance sensor-based failure times) where heterogeneous failure-modes likely exist.
4. Conduct simulation studies to assess performance of mixture modelling vs single distributions under known heterogeneity.

5. Explore Bayesian estimation of mixture models (which can incorporate prior beliefs and quantify uncertainty more fully).

REFERENCES

1. Meeker, W. Q., & Escobar, L. A. (1998). *Statistical Methods for Reliability Data*. Wiley.
2. Nelson, W. (2004). *Applied Life Data Analysis*. Wiley.
3. UCI Machine Learning Repository (2018). *Simulated Data for Survival Modelling*. University of California, Irvine.
4. Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. Wiley.
5. Elmahdy E.E. (2016). Reliability Analysis for Monte Carlo simulation using the Expectation Maximization Algorithm for a Weibull mixture distribution model. J. Appl Math, 5(3).
6. Tobias P.A. and Trindale D.C. (1995). Applied Reliability, second ed., Van Nostrand Reinhold Co., New York.
7. Dempster A.P., Laird N.M., and Rubin D.B. (1977). Maximum Likelihood Estimation from incomplete data via the EM Algorithm, J.Roy.Statist.Soc.Ser.B(39)1-38.
8. McLachlan, G.J. and Peel, D. (2000). Finite mixture model, inference and application to clustering. Marcel Dekker, New York.
9. Sarhan, A.M., & Green, C. (2009). On the Weibull-Gamma distribution and its applications. Computational Statistics & Data Analysis, 53(12), 4655-4664.
10. Abernethy R.B. (2010). The new Weibull=Hand: Reliability and Statistical Analysis for predicting Life, Safety, Survivability, Risk, Cost, and warranty claims (Fifth ed.,), Florida
11. Kececioglu D. (1994). Reliability and life testing Life HandBook, Prentice Hall, Inc., Englewood Cliffs, New Jersey, Vol.1&2, 1993 and 1994.
12. Prabhakar M., Xie M., & Jiang R. (2004). Weibull models, John Wiley & Sons, Inc.
13. Everitt B.S. & Hand D.J. (1981). Finite Mixture Distribution, Chapman & Hall, London.
14. Elmahdy E.E. & Aboutahoun A.W.(2013). A new approach for parameter Estimation of Finite Weibull Mixture Distributions for Reliability Modeling, Applied Math. Modeling 37(4), 1800-1810.
15. Elmahdy E.E. (2015). A new approach for Weibull modeling for Reliability life data analysis, Applied Mathematics and Computation, 250, 708-720.
16. Salvatore J.B. and Nasa H. (1997). Aerospace Application of Weibull and Monte Carlo simulation with Importance Sampling, IEEE Annual Reliability and Maintenance Symposium, Proc.
17. Gelman A., Carlin J.B., Stern H.S.and Rubin D.B. (2003). Bayesian Data Analysis, Second ed., Chapman & Hall, CRC.
18. Lee E.T. and Wang J.W. (2003). Statistical Methods for Survival Data Analysis, John Wiley & Sons Inc., 3RD ed,