STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, Month 2026, pp 1–24.
Published online in International Academic Press (www.IAPress.org)

A New Log-Exponential-Exponential Accelerated Failure Time Model for Censored and Uncensored Data with Applications and Rigorous NRR Validation

Mohamed Ibrahim[1,*], Hafida Goual[2], Abdullah H. Al-Nefaie[1], Sami A. Morsi[3], Haitham M. Yousof[4], Rehab Shehata Mahmoud[4]

[1] *Department of Quantitative Methods, School of Business, King Faisal University, Al Ahsa 31982, Saudi Arabia*
[2] *Laboratory of probabilities and statistics LaPS, Badji Mokhtar-Annaba University, 12, P.O. Box, 23000 Annaba Algeria*
[3] *Applied College, Abqaiq Branch, King Faisal University, P.O. Box 4000, Al-Ahsa 31982, Saudi Arabia*
[4] *Department of Statistics, Mathematics and Insurance, Faculty of Commerce, Benha University, Egypt*

Abstract The accelerated failure time (AFT) model is a cornerstone of survival analysis, yet traditional baseline distributions often lack the flexibility required to model complex, non-monotonic hazard rates. This paper proposes a novel AFT model based on the log-exponential-exponential (LEE) distribution, offering enhanced flexibility for reliability, medical, and actuarial data analysis. We derive the fundamental mathematical properties of the LEE-AFT model and develop the maximum likelihood estimators (MLEs) for both uncensored and right-censored data schemes. To ensure rigorous model validation, we adapt the modified Nikulin-Rao-Robson (NRR) chi-square statistic to assess the goodness-of-fit when parameters are estimated from complete and censored samples. Extensive Monte Carlo simulations are conducted to evaluate the consistency, bias, and efficiency of the MLEs, as well as to verify that the empirical rejection rates of the NRR test closely align with nominal theoretical levels. The practical applicability of the proposed framework is demonstrated using four real-world datasets: electric insulating fluid, body fat, motor insulation, and emergency care data. Furthermore, we present a comprehensive actuarial case study using Norwegian fire insurance claims to evaluate risk measures such as Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR). Fitted under log-linear, power-rule, and Arrhenius covariate assumptions, the AFT-LEE model consistently provides superior fits compared to classical baselines, confirming its value as a robust tool for complex survival analysis, reliability engineering, and financial risk management.

Keywords Accelerated Failure Time; Censored Data**;** Exponential Distribution; Maximum Likelihood Estimators; Nikulin-Rao-Robson; Actuarial Risk Measures.

AMS 2010 subject classifications 62N01; 62N02; 62E10, 60K10, 60N05.

DOI: 10.19139/soic-2310-5070-4838

1. Introduction

Survival analysis and reliability engineering rely heavily on the formulation of robust parametric models to understand time-to-event data and the impact of external covariates on system lifetimes. Among the various regression frameworks, the Accelerated Failure Time (AFT) model stands out for its direct and intuitive interpretation, wherein covariates act multiplicatively on the survival time, effectively accelerating or decelerating the life course of a system or individual (Bagdonavicius & Nikulin, 2002; Dupuy, 2014). While classical AFT models based on the Weibull, log-normal, or log-logistic distributions have been extensively utilized, they often fall short when confronted with real-world data exhibiting complex, non-monotonic hazard rate behaviors. To address

*Correspondence to: Mohamed Ibrahim (Email: miahmed@kfu.edu.sa). Department of Quantitative Methods, School of Business, King Faisal University, Al Ahsa 31982, Saudi Arabia.

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 2026 International Academic Press

this limitation, researchers have increasingly turned to generalized families of distributions that offer enhanced flexibility in modeling diverse hazard shapes (Afify et al., 2018; Yousof et al., 2018, 2026; Alizadeh et al., 2023; Alizadeh et al., 2025).

This paper introduces a novel AFT model built upon the LEE distribution. The proposed LEE-AFT model is designed to provide exceptional flexibility, seamlessly accommodating both monotonic and non-monotonic hazard rate functions. We comprehensively derive the mathematical properties of the model, including the cumulative distribution function (CDF), probability density function (PDF), survival function, and hazard rate functions. Parameter estimation is conducted using the Maximum Likelihood Estimation (MLE) method for both complete (uncensored) and right-censored data, with explicit derivations of the score functions and the Fisher Information Matrix (FIM). To solve the resulting non-linear systems, we employ advanced numerical optimization techniques, specifically the Barzilai-Borwein (BB) algorithm (see Ravi and Gilbert, 2009), which is highly efficient for high-dimensional objective functions.

Furthermore, a critical component of this study is the rigorous statistical validation of the proposed model. Recognizing that classical goodness-of-fit tests such as Pearson's chi-square, Kolmogorov-Smirnov (K–S), and Anderson-Darling (A–D), lose their theoretical validity when parameters are estimated from the observed data (Nikulin, 1973; Rao & Robson, 1974), we employ the modified Nikulin-Rao-Robson (NRR) chi-square statistic. The NRR statistic has been successfully adapted and applied in various contexts for model adequacy testing (Goual & Yousof, 2019; Goual et al., 2019; Goual et al., 2020; Ibrahim et al., 2022; Yousof et al., 2021; Yousof et al., 2022). In this paper, we tailor this modified test specifically for the LEE-AFT model, ensuring asymptotic correctness for both uncensored and right-censored samples (Bagdonavicius & Nikulin, 2011).

The paper is structured to first present the mathematical framework and MLE derivations for both data schemes, followed by extensive Monte Carlo simulation studies to evaluate the consistency and efficiency of the estimators and the empirical size of the NRR test. Finally, the practical utility of the LEE-AFT model is demonstrated through applications to diverse real-world datasets. These include the electric insulating fluid breakdown times under varying voltages (Lawless, 2003), body fat measurements (Neter et al., 1996; Johnson, 1996), motor insulation failure under thermal stress, and emergency care clinical data (Teghri et al., 2024; Loubna et al., 2024). Additionally, we extend the application to actuarial science using Norwegian fire insurance claims. The datasets are fitted under various covariate assumptions, including log-linear, power-rule, and Arrhenius models, demonstrating the broad applicability of the proposed framework in engineering, medical, reliability, and financial studies.

Despite the proliferation of generalized lifetime distributions in recent literature, a significant research gap persists in the simultaneous provision of flexible hazard modeling and statistically rigorous validation for censored data. Many existing AFT models are either too rigid to capture the non-monotonic hazard rates frequently observed in medical and engineering datasets, or they rely on classical goodness-of-fit tests (such as Pearson's chi-square or Kolmogorov-Smirnov) that are fundamentally invalid when model parameters are estimated from the observed data. This issue is particularly exacerbated in the presence of right-censoring, where the asymptotic distribution of traditional test statistics becomes heavily dependent on the underlying model and the estimation procedure. The scientific motivation behind this paper is twofold. First, we aim to bridge the flexibility gap by introducing the LEE-AFT model. Second, we are motivated to provide a mathematically sound validation framework by adapting the NRR statistic for the LEE-AFT model under both complete and right-censored schemes. By doing so, we offer practitioners a reliable, unified methodology for both modeling complex survival data and rigorously verifying model adequacy without resorting to ad-hoc or theoretically flawed diagnostic tools. According to Ibrahim et al. (2026), the CDF of the log-exponential-exponential (LEE) model can be expressed as

$$F_{\zeta,\alpha}(x) = \frac{1}{\exp(1)\log(2)} \exp\left[L_{\zeta,\alpha}(x)\right] \ln\left[1 + L_{\zeta,\alpha}(x)\right] | \alpha > 0, \zeta \in (-1,0) \cup (0,\infty), x > 0, \tag{1}$$

where

$$L_{\zeta,\alpha}(x) = \frac{1}{\ln(1+\zeta)} \ln\left\{1 + \zeta\left[1 - \exp(-\alpha x)\right]\right\}. \tag{2}$$

The corresponding PDF of (1) can then be expressed as:

$$f_{\zeta,\alpha}(x) = \frac{\zeta\alpha}{\exp(1)\ln(2)}\left\{\frac{1}{1+L_{\zeta,\alpha}(x)} + \ln\left[1+L_{\zeta,\alpha}(x)\right]\right\}\frac{\exp(-\alpha x)\exp\left[L_{\zeta,\alpha}(x)\right]}{\ln(1+\zeta)\ln\left\{1+\zeta\left[1-\exp(-\alpha x)\right]\right\}}. \quad (3)$$

Below, we introduce a novel AFT model based on the LEE distribution. This model is designed to provide greater flexibility in modeling survival data while accommodating both monotonic and non-monotonic hazard rate functions. To construct this model, we assume that n independent failure time variables are observed, and we consider the hypothesis H_0, which specifies the survival function given a vector of explanatory variables $z(t) = (z_0(t), z_1(t), ..., z_m(t)), \quad z_0(t) = 1$ represents the baseline covariate (e.g., temperature, stress, or other external factors). The survival function under this hypothesis takes the form:

$$S(t|z) = S_0\left(\int_0^t e^{-\beta^T z(u)}du;\zeta\right),$$

where $\beta = (\beta_0, \beta_1, ..., \beta_m)^T$ is a vector of unknown regression parameters, the function S_0 is a specified functional of time and does not depend on z_i. If explanatory variables are constant over time, the parametric AFT model has the form

$$S(t|z) = S_0\left[\exp\left(-\beta^T z\right)t;\zeta\right].$$

Consider the LEE distribution as baseline distribution where

$$\begin{aligned} S_0 = S_{\zeta,\alpha}(x) &= 1 - F_{\zeta,\alpha}(x) \\ &= 1 - \frac{1}{e\ln(2)}\exp\left[L_{\zeta,\alpha}(x)\right]\ln\left[1+L_{\zeta,\alpha}(x)\right] \\ &= 1 - \frac{1}{e\ln(2)}\exp\left[\frac{\ln\left(1+\zeta\left[1-e^{-\alpha x}\right]\right)}{\ln(1+\zeta)}\right]\ln\left[1+\frac{\ln\left(1+\zeta\left[1-e^{-\alpha x}\right]\right)}{\ln(1+\zeta)}\right]. \end{aligned}$$

So, the CDF of the LEE-AFT model can be expressed as

$$\begin{aligned} F_{\text{LEE-AFT}} &= \frac{1}{e\ln(2)}\exp\left[\frac{\ln\left(1+\zeta\left[1-e^{-\alpha t e^{-\beta^T z}}\right]\right)}{\ln(1+\zeta)}\right] \\ &\quad \times \ln\left[1+\frac{\ln\left(1+\zeta\left[1-e^{-\alpha t e^{-\beta^T z}}\right]\right)}{\ln(1+\zeta)}\right], \\ t &> 0, \alpha > 0, \zeta \in (-1,0)\cup(0,\infty), and \beta \in \mathbb{R}^{m+1}, \end{aligned}$$

For a given time $t > 0$, define the *scaled time*:

$$\mathcal{X}(t,\mathbf{z}) = t\exp\left(-\boldsymbol{\beta}^T\mathbf{z}\right).$$

The baseline link function $L_{\zeta,\alpha}(\cdot)$ evaluated at $\mathcal{X}(t,\mathbf{z})$ is:

$$L_{\zeta,\alpha}\left(\mathcal{X}(t,\mathbf{z})\right) = \frac{\log\left(1+\zeta\left[1-\exp\left(-\alpha t e^{-\boldsymbol{\beta}^T\mathbf{z}}\right)\right]\right)}{\log(1+\zeta)},$$

with $\alpha > 0$ and $\zeta \in (-1,0)\cup(0,\infty)$. For brevity in the following expressions, we denote:

$$L := L_{\zeta,\alpha}\left(\mathcal{X}(t,\mathbf{z})\right), \quad \mathcal{X} := te^{-\boldsymbol{\beta}^T\mathbf{z}}.$$

and then, the PDF of the LEE-AFT model can be re-expressed as

$$f_{\text{AFT-LEE}} = = e^{-\boldsymbol{\beta}^T\mathbf{z}} \cdot \frac{\exp(L)}{e\log(2)}\left[\log(1+L)+\frac{1}{1+L}\right]\frac{\zeta\alpha\exp\left(-\alpha\mathcal{X}\right)}{\log(1+\zeta)\left(1+\zeta\left[1-\exp\left(-\alpha\mathcal{X}\right)\right]\right)}$$
$$= \frac{\zeta\alpha e^{-\boldsymbol{\beta}^T\mathbf{z}}}{e\log(2)\log(1+\zeta)}\left[\log(1+L)+\frac{1}{1+L}\right]\frac{\exp\left(-\alpha te^{-\boldsymbol{\beta}^T\mathbf{z}}\right)\exp(L)}{\log\left(1+\zeta\left[1-\exp\left(-\alpha te^{-\boldsymbol{\beta}^T\mathbf{z}}\right)\right]\right)\left(1+\zeta\left[1-\exp\left(-\alpha te^{-\boldsymbol{\beta}^T\mathbf{z}}\right)\right]\right)}.$$

Analogously, the corresponding survival function (SF), HRF and cumulative HRF of the LEE-AFT model are given by

$$S_{\text{AFTlee}} = 1-\frac{1}{e\log(2)}\exp\left(L\right)\log\left(1+L\right)$$
$$= 1-\frac{1}{e\log(2)}\exp\left[\frac{\log\left(1+\zeta\left[1-\exp\left(-\alpha te^{-\boldsymbol{\beta}^T\mathbf{z}}\right)\right]\right)}{\log(1+\zeta)}\right]\log\left[1+\frac{\log\left(1+\zeta\left[1-\exp\left(-\alpha te^{-\boldsymbol{\beta}^T\mathbf{z}}\right)\right]\right)}{\log(1+\zeta)}\right].$$

$$\lambda_{\text{AFT-LEE}} = e^{-\boldsymbol{\beta}^T\mathbf{z}} \cdot \frac{\exp(L)\left[\log(1+L)+\frac{1}{1+L}\right]\frac{\zeta\alpha\exp(-\alpha\mathcal{X})}{\log(1+\zeta)(1+\zeta[1-\exp(-\alpha\mathcal{X})])}}{e\log(2)-\exp(L)\log(1+L)}$$
$$= \frac{\zeta\alpha e^{-\boldsymbol{\beta}^T\mathbf{z}}}{\log(1+\zeta)} \cdot \frac{\exp(L)\left[\log(1+L)+\frac{1}{1+L}\right]\exp\left(-\alpha te^{-\boldsymbol{\beta}^T\mathbf{z}}\right)}{\left(e\log(2)-\exp(L)\log(1+L)\right)\left(1+\zeta\left[1-\exp\left(-\alpha te^{-\boldsymbol{\beta}^T\mathbf{z}}\right)\right]\right)},$$

and

$$\Lambda_{\text{AFT-LEE}} = -\log\left(1-\frac{1}{e\log(2)}\exp(L)\log(1+L)\right)$$
$$= -\log\left(1-\frac{1}{e\log(2)}\exp\left[\frac{\log\left(1+\zeta\left[1-\exp\left(-\alpha te^{-\boldsymbol{\beta}^T\mathbf{z}}\right)\right]\right)}{\log(1+\zeta)}\right]\log\left[1+\frac{\log\left(1+\zeta\left[1-\exp\left(-\alpha te^{-\boldsymbol{\beta}^T\mathbf{z}}\right)\right]\right)}{\log(1+\zeta)}\right]\right)$$

To contextualize the proposed LEE-AFT model within the broader landscape of modern survival analysis and reliability engineering, it is essential to acknowledge the significant strides made in developing flexible parametric families. Recent literature has demonstrated that classical distributions often lack the requisite flexibility to capture complex hazard rate behaviors, such as non-monotonic or bathtub-shaped curves, which are prevalent in medical and engineering datasets. To address this, numerous generalized families have been introduced, leveraging generator mechanisms to enhance modeling capacity. For instance, the work of Afify et al. (2018) and Cordeiro et al. (2018) highlighted the utility of composite structures like the Burr XII and Lindley-Weibull models in capturing diverse data shapes. Similarly, Alizadeh et al. (2018, 2023, 2024, 2025) and Yousof et al. (2018, 2019, 2021, 2022) have extensively explored transmuted and odd log-logistic generators, demonstrating their effectiveness in heteroscedastic regression models and risk analysis. These studies collectively underscore the trend toward using modular distributional frameworks to improve fit accuracy over traditional baselines.

Furthermore, the integration of covariates into these flexible distributions via Accelerated Failure Time (AFT) frameworks has gained considerable attention. While early foundational works by Crowder et al. (1991) and Murthy et al. (2004) established the theoretical underpinnings of reliability data analysis, recent applications have focused on extending these models to handle complex stress conditions. For example, Hashem et al. (2024, 2025) and Ibrahim et al. (2019, 2020, 2022, 2023, 2025) have developed various extended models, including Nadarajah-Haghighi and Burr Type XII extensions, specifically tailored for censored validation and Bayesian

estimation. These contributions illustrate the growing necessity for robust parameter estimation techniques that can accommodate both complete and right-censored data schemes without compromising statistical rigor. The development of such models is critical for accurately assessing the impact of external stresses, such as temperature or voltage, on system lifetimes.

In parallel with model development, the issue of rigorous statistical validation has emerged as a critical concern. Classical goodness-of-fit tests, such as Pearson's chi-square or Kolmogorov-Smirnov, are known to lose their theoretical validity when parameters are estimated from the data, particularly in the presence of censoring. To mitigate this, recent studies by Goual and Yousof (2019, 2020), Goual et al. (2019, 2020), and Yadav et al. (2020, 2022) have championed the use of the modified Nikulin-Rao-Robson (NRR) statistic. This approach, rooted in the foundational work of Nikulin (1973) and Rao and Robson (1974), provides an asymptotically correct chi-square test for model adequacy. Applications of the NRR test to various new distributions, such as the Odd Lindley Exponentiated Exponential and the Lomax Inverse Weibull, have demonstrated its superior performance in validating complex models against real-world data. Additionally, works by Mansour et al. (2020a-f) and Salem et al. (2023) have further refined these validation techniques for censored scenarios, emphasizing the need for unified methodologies that couple flexible modeling with statistically sound diagnostic tools.

2. The MLE for the LEE-AFT model

In this section, we apply the maximum likelihood method to estimate the parameters of the AFT for the LEE distribution. We give a detailed description of the method as well as the score functions and the elements of the FIM.

2.1. The MLE derivations

Let $x_1, \ldots, x_n$ be a RS from the AFT for the LEE model with parameters and α. Let $\boldsymbol{\theta} = (\lambda, \beta_0, \beta_1)^{\intercal}$ be the 4×1 parameter vector. For determining the MLE of $\boldsymbol{\theta}$, Using the auxiliary quantities:

$$\begin{aligned} A_i &= 1 + \zeta \left(1 - e^{-\lambda \mathcal{X}_i}\right), \\ L_i &= \frac{\log(A_i)}{\log(1+\zeta)}, \end{aligned}$$

we have the log-likelihood function the log-likelihood expressed compactly as:

$$\begin{aligned} \ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \Bigg\{ \delta_i \Bigg[& \log(\lambda) + \log(\zeta) - \boldsymbol{\beta}^{\top} \mathbf{z}_i - 1 - \log(\log 2) - \log(\log(1+\zeta)) \\ & + \log\left(\frac{1}{1+L_i} + \log(1+L_i)\right) - \lambda \mathcal{X}_i + L_i - \log\left(\log A_i\right) - \log A_i \Bigg] \\ & + (1-\delta_i) \log\left(1 - \frac{1}{e \log 2} e^{L_i} \log(1+L_i)\right) \Bigg\}, \end{aligned}$$

Setting the nonlinear system of score fucntions and solving them simultaneously yields the MLE $\widehat{\boldsymbol{\theta}} = (\widehat{\zeta}, \widehat{\alpha}, \widehat{\beta_0}, \widehat{\beta_1})^{\intercal}$. To solve these equations, it is usually more convenient to use nonlinear optimization methods such as the quasi-Newton algorithm to numerically maximize ℓ. Since, we can not find the explicit formulas for the MLEs of the parameters, we use numerical methods such as the Newton Raphson method, the Monte Carlo method, the BB algorithm or others.

Table 1. MLEs $\left(\widehat{\zeta}, \widehat{\alpha}, \widehat{\beta_0}, \widehat{\beta_1}\right)$ of the parameters, Bias, and mean squared errors for uncensored data.

$N = 13000$	$n = 15$	$n = 25$	$n = 50$	$n = 100$	$n = 250$	$n = 450$
$\widehat{\zeta}$	3.5642	3.5495	3.5387	3.5261	3.5137	3.5064
$Bias(\widehat{\zeta})$	0.0642	0.0495	0.0387	0.0261	0.0137	0.0064
$MSE(\widehat{\zeta})$	3.504×10^{-2}	2.812×10^{-2}	1.845×10^{-2}	7.821×10^{-3}	2.608×10^{-3}	1.186×10^{-3}
$\widehat{\alpha}$	2.5645	2.5512	2.5402	2.5288	2.5150	2.5023
$Bias(\widehat{\alpha})$	0.0645	0.0512	0.0402	0.0288	0.0150	0.0023
$MSE(\widehat{\alpha})$	3.571×10^{-2}	1.012×10^{-2}	0.837×10^{-2}	5.174×10^{-3}	1.446×10^{-3}	1.224×10^{-4}
$\widehat{\beta_0}$	1.9643	1.9606	1.9587	1.9561	1.9522	1.9504
$Bias(\widehat{\beta_0})$	0.0143	0.0106	0.0087	0.0061	0.0022	0.0004
$MSE(\widehat{\beta_0})$	8.187×10^{-2}	5.564×10^{-2}	3.053×10^{-2}	1.364×10^{-2}	0.195×10^{-2}	0.231×10^{-2}
$\widehat{\beta_1}$	2.5604	2.5584	2.5511	2.5506	2.5502	2.5501
$Bias(\widehat{\beta_1})$	0.0104	0.0084	0.0011	0.0006	0.0002	0.0001
$MSE(\widehat{\beta_1})$	4.251×10^{-2}	0.953×10^{-2}	6.251×10^{-3}	2.951×10^{-3}	1.845×10^{-3}	1.108×10^{-3}

2.2. Assessing the LEE-AFT model via a simulation study

We conduct a comprehensive simulation study using the R programming software to evaluate the performance of the maximum likelihood estimators (MLEs) for the parameters of the LEE-AFT model. This study is crucial for assessing the accuracy, consistency, and efficiency of the estimation methods under varying sample sizes and parameter settings. The results are obtained using a numerical optimization approach, specifically the Newton-Raphson method , which is widely recognized for its robustness and efficiency in solving nonlinear equations.

In this simulation setup, we assume that the data follows the LEE-AFT distribution. To ensure reliability and generalizability of the results, the data generation process is repeated N = 13000 times . The true parameter values used in the simulation are set as follows: $\zeta = 3.50, \alpha = 2.5, \beta_0 = 1.95, \beta_1 = 2.55$. These values represent the baseline hazard rate and the regression coefficients associated with the explanatory variables in the AFT model.

For computational purposes, we utilize the BB algorithm, as described in Ravi and Gilbert (2009), implemented in the R software. This algorithm is particularly well-suited for optimizing high-dimensional nonlinear objective functions and provides efficient computation of the MLEs. Using this approach, we calculate the averages of the simulated MLEs for the parameters $\zeta, \alpha, \beta_0, \beta_1$, along with their respective mean squared errors (MSE) and Bias. The MSE serves as a key metric to evaluate the precision and bias of the estimators across different sample sizes.

The simulation study considers six distinct sample sizes: $n = 15, n = 25, n = 50, n = 100, n = 250$, and $n = 450$. These sample sizes span a wide range, from small to large datasets, enabling us to examine how the performance of the estimators evolves as the sample size increases. For each sample size, the simulation generates synthetic datasets based on the specified parameter values and computes the MLEs iteratively over the 13000 replications. The results are then aggregated to compute the average estimates, Bias, and their corresponding MSEs. The outcomes of the simulation study are summarized in Table 1, which lists the MSE and Bias for the MLEs of the parameters $\zeta, \alpha, \beta_0, \beta_1$ across the different sample sizes. The table provides a clear depiction of how the MSE decreases as the sample size grows, reflecting the asymptotic properties of the MLEs. The results demonstrate that the proposed estimation method performs well under the specified simulation conditions, providing accurate and stable parameter estimates. Overall, this simulation study validates the practical applicability of the LEE-AFT model and highlights its potential for real-world survival analysis applications.

Table 1 presents the simulation results for the uncensored case. It is evident that the bias and MSE for all parameters decrease monotonically as the sample size increases from $n = 15$ to $n = 450$, confirming the consistency of the MLEs. For instance, the MSE of $\widehat{\zeta}$ drops from 3.504×10^{-2} at $n = 15$ to 1.186×10^{-3} at $n = 450$, demonstrating the estimator's rapid convergence to the true parameter value.

3. Validation of the LEE-AFT model

In the case of well-specified distributions, classical goodness-of-fit tests such as Pearson's chi-square, the Kolmogorov-Smirnov (K–S) statistic, and the Anderson-Darling (A–D) statistic are commonly employed to assess the adequacy of the proposed model. However, when the distribution's parameters are unknown and must be estimated from the observed data, these traditional tests lose their theoretical rigor. The resulting test statistics no longer follow their classical distributions, as their asymptotic behavior becomes dependent on both the underlying model and the estimation procedure used. When complete data are available, several techniques exist to assess the fit of probabilistic models to empirical data. Among these, Pearson's chi-square test is one of the most widely applied. Nevertheless, its applicability becomes limited in cases involving censored data or models with unknown parameters. To address these challenges, Nikulin (1973) and independently Rao and Robson (1974) developed a statistic, now commonly referred to as the NRR statistic, specifically designed for use with complete samples. This statistic is constructed based on the MLEs derived from the original data, and under regularity conditions, it converges in distribution to a chi-square distribution. For a detailed exposition of the construction and properties of this class of test statistics, readers are referred to Voinov et al. (2013) and Goual et al. (2019). The NRR-based methods have been successfully applied in various contexts, including model adequacy testing for the Lomax Inverse Weibull distribution (Goual et al., 2020), the Burr XII inverse Rayleigh model (Goual et al., 2019), and the Lindley Exponentiated distribution (Goual et al., 2019). Building upon this framework, the current section proposes a modified chi-square-type test based on the NRR statistic, tailored specifically to evaluate the fit of the newly introduced LEE model.

3.1. The NRR statistic test for the LEE-AFT model

To test the hypothesis H_0 according to which $T_1, T_2, \cdots, T_n$, an n-sample comes from a parametric family $F_{\boldsymbol{\theta}}(t)$

$$H_0 : \Pr\{T_i \leq t\} = F_{\boldsymbol{\theta}}(t), \quad t \in \mathbb{R},$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \underline{\mathbf{V}}_2, \cdots, \boldsymbol{\theta}_s)^T$ represents the vector of unknown parameters, Nikulin (1973) and Rao and Robson (1974) proposed K^2 the NRR statistic defined as below. Observations $T_1, T_2, \cdots, T_n$ are grouped in r subintervals $\mathbf{I}_1, \mathbf{I}_2, \cdots, \mathbf{I}_r$ mutually disjoint $\mathbf{I}_j =]a_j\text{-}1; a_j]$; where $j = \overline{1;r}$. The limits a_j of the intervals $\mathbf{I}_j$ are obtained such that

$$p_j(\boldsymbol{\theta}) = p_j(\boldsymbol{\theta}; a_{j-1}, a_j) = \int_{a_{j-1}}^{a_j} f_{\boldsymbol{\theta}}(t)dt|_{(\ j=1,2,\cdots,r)},$$

so

$$a_j = F^{-1}\left(\frac{j}{r}\right)|_{(j=1,\cdots,r-1)}.$$

If $\nu_j = (\nu_1, \nu_2, \cdots, \nu_r)^T$ is the vector of frequencies obtained by the grouping of data in these $\mathbf{I}_j$ intervals

$$\nu_j = \sum_{i=1}^{n} 1_{\{t_i \in \mathbf{I}_j\}}\ |_{(j=1,\ldots,r)}.$$

The NRR statistic is given by

$$K^2(\widehat{\boldsymbol{\theta}}_n) = X_n^2(\widehat{\underline{\mathbf{V}}}_n) + \frac{1}{n}\mathbf{L}^T(\widehat{\boldsymbol{\theta}}_n)(\mathbf{I}(\widehat{\boldsymbol{\theta}}_n) - \mathbf{J}(\widehat{\boldsymbol{\theta}}_n))^{-1}\mathbf{L}(\widehat{\boldsymbol{\theta}}_n),$$

where

$$X_n^2(\boldsymbol{\theta}) = \left(\frac{\nu_1 - np_1(\underline{\mathbf{V}})}{\sqrt{np_1(\boldsymbol{\theta})}}, \frac{\nu_2 - np_2(\boldsymbol{\theta})}{\sqrt{np_2(\boldsymbol{\theta})}}, \cdots, \frac{\nu_r - np_r(\boldsymbol{\theta})}{\sqrt{np_r(\underline{\mathbf{V}})}}\right)^T$$

and $\mathbf{J}(\boldsymbol{\theta})$ is the information matrix for the grouped data defined by

$$\mathbf{J}(\boldsymbol{\theta}) = B(\boldsymbol{\theta})^T B(\underline{\mathbf{V}}),$$

Table 2. Empirical levels K^2 and corresponding theoretical levels for uncensored data.

$N = 5000$	$n = 17$	$n = 25$	$n = 50$	$n = 100$	$n = 250$	$n = 450$
$\alpha = 0.01$	0.019	0.017	0.015	0.014	0.012	0.011
$\alpha = 0.05$	0.055	0.053	0.052	0.051	0.0508	0.0506
$\alpha = 0.1$	0.164	0.144	0.138	0.127	0.117	0.0999

with

$$B(\boldsymbol{\theta}) = \left[\frac{1}{\sqrt{p}_i}\frac{\partial p_i(\boldsymbol{\theta})}{\partial \mu}\right]_{r\times s}|_{(i=1,2,\cdots,r \text{ and } k=1,\cdots,s)},$$

then

$$\mathbf{L}(\boldsymbol{\theta}) = (\mathbf{L}_1(\boldsymbol{\theta}), ..., \mathbf{L}_s(\boldsymbol{\theta}))^T \text{ with } \mathbf{L}_k(\boldsymbol{\theta}) = \sum_{i=1}^{r}\frac{\nu_i}{p_i}\frac{\partial}{\partial \boldsymbol{\theta}_k}p_i(\underline{\mathbf{V}}),$$

where $\mathbf{I}_n(\widehat{\boldsymbol{\theta}_n})$ represents the estimated FIM and $\widehat{\boldsymbol{\theta}_n}$ is the maximum likelihood estimator of the parameter vector. The K^2 statistic follows a distribution of chi-square χ^2_{r-1} with $(r-1)$ degrees of freedom.

3.2. Simulation studies under the NRR statistic K^2

Consider a sample $T_{1:n}$ where $T = T_{1:n} = (T_1, T_2, \cdots, T_n)^T$. If these data are distributed in accordance with the LEE-AFT model, then $P\{T_{1:n} \leq t\} = F_{\underline{\mathbf{V}}}(t)$; with unknown parameters $\boldsymbol{\theta} = (\zeta, \alpha, \beta_0, \beta_1)^T$, by fitting the NRR statistic created in the preceding section, a chi-square goodness-of-fit test is created. The MLEs $\widehat{\boldsymbol{\theta}_n}$ of the unknown parameters of the LEE-AFT model are computed on the initial data. Since, the statistic K^2 not dependent on the parameters, we can therefore use the estimated Fisher information matrix (FIM) $I_n(\widehat{\boldsymbol{\theta}_n})$. All the components of the statistic K^2, for the LEE-AFT distribution are provided, therefore K^2 can be deduced easily.

In order to support the results obtained in this work, a numerical simulation is performed. Therefore, in order to test the null hypothesis H_0 of the LEE-AFT model, we calculated 8000 sample data simulations ($n = 17$, $n = 25$, $n = 50$, $n = 100$, $n = 250$ and $n = 450$) from LEE-AFT distribution, after calculating the value of the criterion statistic K^2, we count the number of rejected cases of the null hypothesis H_0. When $K^2 > \chi^2\,(k = r-1)$, the significance is different level $\alpha\,(1\%, 5\%, 10\%)$. The simulation results of the rejected level of K^2 and its theoretical value are shown in Table 2 below.

Table 2 demonstrates the validity of the NRR test for uncensored data. The empirical rejection rates converge to the nominal levels ($\alpha = 0.01, 0.05, 0.10$) as sample size increases. At $n = 450$, the empirical level for $\alpha = 0.05$ is 0.0506, indicating excellent size control and validating the asymptotic chi-square approximation.

3.3. Applications to real data

We take into account the following real data sets and confirm the presumption that their distribution is consistent with the LEE-AFT model in order to demonstrate the applicability of the proposed modified chi-square goodness-of-fit test.

3.3.1. Electric insulating fluid data The failure times of 76 electrical insulating fluids, which were tested under varying voltages ranging from 26 to 38 kilovolts, are documented in the work of Lawless (2003). This dataset has been widely utilized in reliability studies due to its relevance in understanding the impact of voltage stress on the durability and performance of insulating materials. Bagdonavicius and Nikulin (2011) further analyzed this dataset to assess its compatibility with the exponential and Weibull accelerated failure time power-rule models. Their study aimed to determine whether these traditional parametric models could adequately describe the relationship between voltage levels and the failure times of the insulating fluids. In this section, we extend their analysis by evaluating how well the data fits our proposed LEE-AFT model. The LEE-AFT model offers a more flexible framework

Table 3. Electric insulating fluid data breakdown times.

Voltage level (z_i)	n_i	Breakdown time x_i
26	3	5.79,1579.52,2323.7
28	5	68.85,426.07,110.29,108.29,1067.6
30	11	17.05,22.66,21.01,175.88,139.07,144.12,
		20.46,43.40,194.90,47.30,7.74
32	15	0.40,82.85,9.88,89.29,215.10,2.75,0.79,
		15.93,3.91,0.27,0.69,100.58,27.80,13.95,53.24
34	19	0.96,4.15,0.19,0.78,8.01,31.75,7.35,6.50,8.27,33.91,
		32.52,3.16,4.85,2.78,4.67,1.31,12.06,36.71,72.89
36	15	1.97,0.59,2.58,1.69,2.71,25.50,0.35,0.99,
		3.99,3.67,2.07,0.96,5.35,2.90,13.77
38	8	0.47,0.73,1.40,0.74,0.39,1.13,0.09,2.38

compared to the conventional exponential and Weibull models, as it can accommodate both monotonic and non-monotonic hazard rate functions. This flexibility makes it particularly suitable for modeling complex datasets like the one under consideration, where the influence of voltage stress on failure times may exhibit nonlinear patterns. By applying the LEE-AFT model to this dataset, we aim to demonstrate its ability to provide a better fit and capture the underlying failure mechanisms more accurately. The results of this evaluation will not only validate the practical utility of the LEE-AFT model but also highlight its advantages over existing models in analyzing real-world reliability data.The data observations are given in Table 3:

Table 3 summarizes the raw data used for the electric insulating fluid analysis, showing the high variability in breakdown times across different voltage levels, which necessitates a flexible AFT model.

1- In case of $\varphi(z) = z$ log-linear assumption: Using R statistical software (the BB package) we find the values of the MLEs of LEE-AFT distribution parameters :

$$\widehat{\zeta} = 1.5487, \widehat{\alpha} = 2.6185, \widehat{\beta}_0 = 3.005, \widehat{\beta}_1 = 0.6487,$$

we choose $r = 8$ intervals and the estimated FIM can be expressed as :

$$I\left(\widehat{\boldsymbol{\theta}}\right) = \begin{pmatrix} 4.0215 & -2.5124 & -12.0217 & 3.6254 \\ & 3.5124 & -1.9538 & 10.9032 \\ & & 28.9014 & -11.8474 \\ & & & 5.3816 \end{pmatrix},$$

and then the NRR statistic : $K^2 = 19.56842$. For the critical value : $\alpha = 0.01$, we find $K^2 > \chi^2_{0.01}(7) = 18.4753$.

2- In case of $\varphi(z) = \log(z)$ power-rule assumption: We find the values of the MLEs of the LEE-AFT distribution parameters:

$$\widehat{\zeta} = 1.5003, \widehat{\alpha} = 1.9658, \widehat{\beta}_0 = 5.06474, \widehat{\beta}_1 = -1.20544,$$

we take $r = 8$ intervals and the estimated FIM can be:

$$I\left(\widehat{\boldsymbol{\theta}}\right) = \begin{pmatrix} 2.6254 & -24.0324 & 4.91570 & -5.9157 \\ & 9.0607 & 0.9935 & -28.9514 \\ & & 8.6352 & 2.3174 \\ & & & 1.0885 \end{pmatrix},$$

the NRR statistic is $K^2 = 10.10241$. For the critical values : $\alpha = 0.01$, $\alpha = 0.05$ and $\alpha = 0.1$, we find

$$K^2 < \chi^2_{0.01}(7) = 18.4753,$$

$$K^2 < \chi^2_{0.05}(7) = 14.0671,$$

Table 4. Body fat dataset observations.

z_{i1} (triceps skinfold measurement); z_{i2} (thigh circumference); x_i(body-fat)
$19.5, 24.7, 30.7; 43.1, 49.8, 51.9; 11.9, 22.8, 18.7$
$29.8, 19.1, 25.6; 54.3, 42.2, 53.9; 20.1, 12.9, 21.7$
$31.4, 27.9, 22.1; 58.5, 52.1, 49.9; 27.1, 25.4, 21.3$
$25.5, 31.1, 30.4; 53.5, 56.6, 56.7; 19.3, 25.4, 27.2$
$18.7, 19.7, 14.6, 29.5; 46.5, 44.2, 42.7, 54.4; 11.7, 17.8, 12.8, 23.9$
$27.7, 30.2, 22.7, 25.2; 55.3, 58.6, 48.2, 51.0; 22.6, 25.4, 14.8, 21.1$

and

$$K^2 < \chi^2_{0.1}(7) = 12.0170.$$

respectively. So, we can assume that electric insulating fluid data of Lawless (2003) correspond appropriately to the LEE-AFT model.

3- In case of $\varphi(z) = 1/z$ arrehnius model: We fit these data by the LEE-AFT model. Using R statistical software (the BB package) we find the values of the MLEs of the LEE-AFT distribution parameters :

$$\widehat{\zeta} = 3.257, \widehat{\alpha} = 1.849, \widehat{\beta}_0 = -4.0067, \widehat{\beta}_1 = 5.6175,$$

we take $r = 8$ intervals and the estimated FIM expressed as :

$$I\left(\widehat{\boldsymbol{\theta}}\right) = \begin{pmatrix} 5.6184 & 7.6295 & -20.0064 & -4.9173 \\ & 0.9996 & -2.3022 & -6.3578 \\ & & 3.3698 & 1.2457 \\ & & & 3.7006 \end{pmatrix},$$

the NRR statistic is: $K^2 = 19.9485$. For the critical value : $\alpha = 0.01$, we find $K^2 > \chi^2_{0.01}(7) = 18.4753$.

3.3.2. Body fat data set The data of Neter et al. (1996) provides information on ($n = 20$) body fat, triceps skinfold thickness, thigh circumference, and mid-arm circumference for twenty healthy females aged 20 to 34. The data are presented in Table 4.

Table 4 displays the anthropometric measurements used as covariates. Note that while body fat is not a time-to-event variable, it is treated here as a positive continuous response in an AFT framework for demonstration purposes.

For $\varphi(z) = z$ as a log-linear assumption: We fit these data by the LEE-AFT model. Using R statistical software (the BB package) we find the values of the MLEs of LEE-AFT distribution parameters :

$$\widehat{\zeta} = 4.5127, \widehat{\alpha} = 1.0483, \widehat{\beta}_0 = -2.0047, \widehat{\beta}_1 = -0.9937, \widehat{\beta}_2 = 0.0098.$$

we take $r = 4$ intervals and the estimated FIM expressed as

$$I\left(\widehat{\boldsymbol{\theta}}\right) = \begin{pmatrix} 8.0025 & -12.8457 & -22.0547 & -17.1653 & 1.9538 \\ & 1.06253 & 8.6137 & -7.6473 & 3.1428 \\ & & 5.8091 & -4.613457 & -11.6620 \\ & & & 3.7006 & -8.9436 \\ & & & & 7.0729 \end{pmatrix},$$

and then the NRR statistic : $K^2 = 5.678660$. For different critical values : $\alpha = 1\%, \alpha = 5\%$ and $\alpha = 10\%$, we find

$$K^2 < \chi^2_{0.01}(3) = 11.3448, K^2 < \chi^2_{0.05}(3) = 7.8147$$
$$K^2 < \chi^2_{0.1}(3) = 6.2513,$$

respectively.

Table 5. Variables in Johnson's data set.

Variable	Name	Details	Variable	Name	Details
z_1	age	Age (years)	z_8	thigh	Circumference (cm)
z_2	weight	Weight (lb)	z_9	knee	Circumference (cm)
z_3	height	Height (in)	z_{10}	ankle	Circumference (cm)
z_4	neck	Circumference (cm)	z_{11}	bicepes	Circumference (cm)
z_5	chest	Circumference (cm)	z_{12}	forearm	Circumference (cm)
z_6	ab	Circumference (cm)	z_{13}	wrist	Circumference (cm)
z_7	hip	Circumference (cm)	x	pcfat	Body fat $(\%)$

3.3.3. Johnson's data set Johnson (1996) utilized a comprehensive dataset to explore challenges associated with multiple regression analysis. The dataset consisted of a response variable, which was the estimated percentage of body fat, alongside 13 continuous covariates measured for $n = 252$ males . These covariates included demographic and anthropometric measurements such as age, weight, height, and 10 distinct body circumference measurements. Below is a detailed breakdown of the variables included in the dataset in Table 5:

Table 5 outlines the rich set of predictors available in Johnson's dataset, allowing for robust multivariate analysis.

In our case, we used two covariates density (Density determined from underwater weighing gm/cm^3) and age (years). We consider the log linear assumption $(\varphi(z) = z)$ and we fit this data by the LEE-AFT model. The values of the MLEs parameters:

$$\widehat{\zeta} = 2.1348, \widehat{\alpha} = 2.9683, \widehat{\beta}_0 = -0.9968, \widehat{\beta}_1 = 06748, \widehat{\beta}_2 - 3.1363,$$

We take $r = 15$ intervals and the estimated FIM $I\left(\widehat{\underline{\mathbf{V}}}\right)$ expressed as

$$I\left(\widehat{\boldsymbol{\theta}}\right) = \begin{pmatrix} 1.6253 & 0.6195 & 7.6859 & -1.6653 & 11.9008 \\ & 2.0257 & -18.6247 & -17.0003 & -0.18488 \\ & & 1.1095 & 6.60427 & -10.6060 \\ & & & 3.7876 & 0.9963 \\ & & & & 1.4019 \end{pmatrix},$$

The NRR statistic test: $K^2 = 19.33462$. For different critical values : $\alpha = 0.01, \alpha = 0.05$ and $\alpha = 0.1$, we find

$$\begin{aligned} K^2 &< \chi^2_{0.01}(14) = 29.1412, \\ K^2 &< \chi^2_{0.05}(14) = 23.6847, \\ K^2 &< \chi^2_{0.1}(14) = 21.06414, \end{aligned}$$

respectively. This data can be fitted by our proposed LEE-AFT model with the log-linear assumption $(\varphi(z) = z)$. One can affirm that our proposed LEE-AFT model can be an appropriate distribution of this data.

4. Censored case

4.1. Maximum likelihood estimation

The likelihood function of the right-censored LEE-AFT is

$$\begin{aligned} L(x;\vartheta) &= \prod_{i=1}^{n} f(x_i;\vartheta) = \prod_{i=1}^{n} \lambda^{\delta_i}(x_i;\vartheta) S(x_i;\vartheta), \delta_i = 1_{\{T_i \le C_i\}} \\ &= \prod_{i=1}^{n} \left\{ \frac{\lambda e^{-\beta^T z} \left[\theta\left(1 + \lambda x_i \exp\left(-\beta^T z_i\right)\right) + 1\right]}{1 + \lambda x_i \exp\left(-\beta^T z_i\right)} \right\}^{\delta_i} \times \frac{\exp\left(-\lambda x_i \exp\left(-\beta^T z_i\right)\right)}{1 + \lambda x_i \exp\left(-\beta^T z_i\right)}. \end{aligned}$$

The log-likelihood function is

$$\ell(\boldsymbol{\theta}) = \sum_{i\in F}\Bigg[\log(\lambda) + \log(\zeta) - \boldsymbol{\beta}^\top \mathbf{z}_i - 1 - \log(\log 2) - \log(\log(1+\zeta)) + \log\left(\frac{1}{1+L_i} + \log(1+L_i)\right) - \lambda\mathcal{X}_i + L_i - \log\left(\log A_i\right) - \log A_i\Bigg] + \sum_{i\in C}\log\left(1 - \frac{1}{e\log 2}e^{L_i}\log(1+L_i)\right),$$

where $\mathcal{X}_i = t_i e^{-\boldsymbol{\beta}^\top \mathbf{z}_i}$ is defined explicitly. F and C denote the sets of uncensored and censored observations, respectively.

To present the score functions explicitly without auxiliary shorthands for derivatives, define the following core quantities for each individual i:

$$\mathcal{X}_i = t_i \exp\left(-\boldsymbol{\beta}^\top \mathbf{z}_i\right), \qquad \mathcal{M}_i = 1 + \zeta\left(1 - e^{-\alpha\mathcal{X}_i}\right),$$

$$\mathcal{L}_i = \frac{\log(\mathcal{M}_i)}{\log(1+\zeta)}, \qquad \mathcal{Q}_i = \frac{1}{1+\mathcal{L}_i} + \log(1+\mathcal{L}_i),$$

$$\mathcal{P}_i = e^{\mathcal{L}_i}\log(1+\mathcal{L}_i), \qquad \mathcal{D}_i = 1 - \frac{1}{e\log 2}\mathcal{P}_i.$$

The score vector is $\mathbf{U}(\boldsymbol{\theta}) = (U_\alpha, U_\zeta, U_{\beta_0}, U_{\beta_1})^\top = \partial\ell/\partial\boldsymbol{\theta}$.

$$U_\alpha = \sum_{i\in F}\Bigg[\frac{1}{\alpha} + \frac{1}{\mathcal{Q}_i}\frac{\mathcal{L}_i}{(1+\mathcal{L}_i)^2}\left(\frac{\zeta\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\,\mathcal{M}_i}\right) - \mathcal{X}_i + \left(\frac{\zeta\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\,\mathcal{M}_i}\right) - \frac{\zeta\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i\log(\mathcal{M}_i)} - \frac{\zeta\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i}\Bigg] + \sum_{i\in C}\left[-\frac{\frac{1}{e\log 2}\,\mathcal{P}_i\,\mathcal{Q}_i\,\zeta\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{D}_i\,\log(1+\zeta)\,\mathcal{M}_i}\right].$$

$$U_\zeta = \sum_{i\in F}\Bigg[\frac{1}{\zeta} - \frac{1}{(1+\zeta)\log(1+\zeta)} + \frac{1}{\mathcal{Q}_i}\frac{\mathcal{L}_i}{(1+\mathcal{L}_i)^2}\left(\frac{\partial\mathcal{L}_i}{\partial\zeta}\right) + \left(\frac{\partial\mathcal{L}_i}{\partial\zeta}\right) - \frac{\log(1+\zeta)}{\log(\mathcal{M}_i)}\left(\frac{\partial\mathcal{L}_i}{\partial\zeta}\right) - \frac{1-e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i}\Bigg] + \sum_{i\in C}\left[-\frac{\frac{1}{e\log 2}\,\mathcal{P}_i\,\mathcal{Q}_i}{\mathcal{D}_i}\left(\frac{\partial\mathcal{L}_i}{\partial\zeta}\right)\right],$$

where $\partial\mathcal{L}_i/\partial\zeta$ is explicitly given by:

$$\frac{\partial\mathcal{L}_i}{\partial\zeta} = \frac{1-e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\,\mathcal{M}_i} - \frac{\log(\mathcal{M}_i)}{(1+\zeta)\log^2(1+\zeta)}.$$

$$U_{\beta_k} = \sum_{i\in F}\Bigg[- z_{ik} + \frac{1}{\mathcal{Q}_i}\frac{\mathcal{L}_i}{(1+\mathcal{L}_i)^2}\left(\frac{\zeta\alpha z_{ik}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\,\mathcal{M}_i}\right) + \alpha z_{ik}\mathcal{X}_i + \left(\frac{\zeta\alpha z_{ik}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\,\mathcal{M}_i}\right) - \frac{\zeta\alpha z_{ik}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i\log(\mathcal{M}_i)} - \frac{\zeta\alpha z_{ik}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i}\Bigg] + \sum_{i\in C}\left[-\frac{\frac{1}{e\log 2}\,\mathcal{P}_i\,\mathcal{Q}_i\,\zeta\alpha z_{ik}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{D}_i\,\log(1+\zeta)\,\mathcal{M}_i}\right].$$

where

$$\frac{\partial\mathcal{L}_i}{\partial\beta_k} = \frac{\zeta\alpha z_{ik}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\,\mathcal{M}_i}, k = 0, 1$$

4.1.1. The elements of the matrix $\widehat{C}$ The elements of the $s \times k$ matrix $\widehat{\mathbf{C}}$ (where $s = 4$ is the number of parameters, and k is the number of intervals) are defined as:

$$\widehat{C}_{lj} = \frac{1}{n}\sum_{i:t_i\in I_j}\delta_i\frac{\partial\ell_i}{\partial\theta_l}, \qquad l = 1,\ldots,4, \quad j = 1,\ldots,k,$$

where $\partial\ell_i/\partial\theta_l$ are the components of the score vector for the LEE-AFT model.

Specifically, for the LEE-AFT model, the elements of $\widehat{\mathbf{C}}$ are:

$$\widehat{C}_{1j} = \frac{1}{n}\sum_{i:t_i\in I_j}\delta_i\left[\frac{1}{\alpha} + \frac{1}{\mathcal{Q}_i}\frac{L_i}{(1+L_i)^2}\left(\frac{\zeta\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i}\right) - \mathcal{X}_i + \frac{\zeta\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\zeta\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i\log(\mathcal{M}_i)} - \frac{\zeta\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i}\right],$$

$$\widehat{C}_{2j} = \frac{1}{n}\sum_{i:t_i\in I_j}\delta_i\Bigg[\frac{1}{\zeta} - \frac{1}{(1+\zeta)\log(1+\zeta)} + \frac{1}{\mathcal{Q}_i}\frac{L_i}{(1+L_i)^2}\left(\frac{1-e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\log(\mathcal{M}_i)}{(1+\zeta)\log^2(1+\zeta)}\right) + \frac{1-e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\log(\mathcal{M}_i)}{(1+\zeta)\log^2(1+\zeta)} - \frac{\log(1+\zeta)}{\log(\mathcal{M}_i)}\left(\frac{1-e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\log(\mathcal{M}_i)}{(1+\zeta)\log^2(1+\zeta)}\right) - \frac{1-e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i}\Bigg],$$

$$\widehat{C}_{3j} = \frac{1}{n}\sum_{i:t_i\in I_j}\delta_i\begin{bmatrix} -1 + \frac{1}{\mathcal{Q}_i}\frac{L_i}{(1+L_i)^2}\left(\frac{\zeta\alpha z_{i0}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i}\right) \\ +\alpha z_{i0}\mathcal{X}_i + \frac{\zeta\alpha z_{i0}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\zeta\alpha z_{i0}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i\log(\mathcal{M}_i)} - \frac{\zeta\alpha z_{i0}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i} \end{bmatrix},$$

$$\widehat{C}_{4j} = \frac{1}{n}\sum_{i:t_i\in I_j}\delta_i\begin{bmatrix} -z_{i1} + \frac{1}{\mathcal{Q}_i}\frac{L_i}{(1+L_i)^2}\left(\frac{\zeta\alpha z_{i1}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i}\right) \\ +\alpha z_{i1}\mathcal{X}_i + \frac{\zeta\alpha z_{i1}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\zeta\alpha z_{i1}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i\log(\mathcal{M}_i)} - \frac{\zeta\alpha z_{i1}\mathcal{X}_i e^{-\alpha\mathcal{X}_i}}{\mathcal{M}_i} \end{bmatrix},$$

where

$$\mathcal{Q}_i = \frac{1}{1+L_i} + \log(1+L_i), \qquad \mathcal{P}_i = e^{L_i}\log(1+L_i), \qquad \mathcal{D}_i = 1 - \frac{1}{e\log 2}\mathcal{P}_i,$$

and $z_{i0} = 1$ for all i.

4.1.2. Information matrix $\widehat{I}$ For right-censored data, the estimated Fisher information matrix $\widehat{\mathbf{I}} = [\widehat{I}_{ll'}]$ in our LEE-AFT model is defined as the outer product of the score functions, summed over all uncensored observations:

$$\begin{aligned} \widehat{I}_{ll'} &= \frac{1}{n}\sum_{i=1}^{n}\delta_i\, s_{i,l}\, s_{i,l'}, \\ l, l' &= 1, 2, 3, 4, \end{aligned}$$

where $(s_{i,1}, s_{i,2}, s_{i,3}, s_{i,4}) = (A_i, B_i, C_i, D_i)$. Here the log-hazard is $\ell_i = \log \lambda(t_i|\mathbf{z}_i, \boldsymbol{\theta})$. Its derivatives with respect to the parameters are defined as:

$$A_i = \frac{\partial \ell_i}{\partial \alpha}, B_i = \frac{\partial \ell_i}{\partial \zeta}, C_i = \frac{\partial \ell_i}{\partial \beta_0}, D_i = \frac{\partial \ell_i}{\partial \beta_1}.$$

These functions are explicitly given by:

$$A_i = \frac{1}{\alpha} + \frac{1}{Q_i}\frac{L_i}{(1+L_i)^2}\left(\frac{\zeta \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i}\right) - \mathcal{X}_i + \frac{\zeta \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\zeta \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\mathcal{M}_i \log(\mathcal{M}_i)} - \frac{\zeta \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\mathcal{M}_i},$$

$$\begin{aligned} B_i = {} & \frac{1}{\zeta} - \frac{1}{(1+\zeta)\log(1+\zeta)} + \frac{1}{Q_i}\frac{L_i}{(1+L_i)^2}\left(\frac{1-e^{-\alpha \mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\log(\mathcal{M}_i)}{(1+\zeta)\log^2(1+\zeta)}\right) \\ & + \frac{1-e^{-\alpha \mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\log(\mathcal{M}_i)}{(1+\zeta)\log^2(1+\zeta)} - \frac{\log(1+\zeta)}{\log(\mathcal{M}_i)}\left(\frac{1-e^{-\alpha \mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\log(\mathcal{M}_i)}{(1+\zeta)\log^2(1+\zeta)}\right) - \frac{1-e^{-\alpha \mathcal{X}_i}}{\mathcal{M}_i}, \end{aligned}$$

$$C_i = -1 + \frac{1}{Q_i}\frac{L_i}{(1+L_i)^2}\left(\frac{\zeta \alpha \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i}\right) + \alpha \mathcal{X}_i + \frac{\zeta \alpha \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\zeta \alpha \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\mathcal{M}_i \log(\mathcal{M}_i)} - \frac{\zeta \alpha \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\mathcal{M}_i},$$

$$D_i = -z_{i1} + \frac{1}{Q_i}\frac{L_i}{(1+L_i)^2}\left(\frac{\zeta \alpha z_{i1} \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i}\right) + \alpha z_{i1}\mathcal{X}_i + \frac{\zeta \alpha z_{i1} \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\log(1+\zeta)\mathcal{M}_i} - \frac{\zeta \alpha z_{i1} \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\mathcal{M}_i \log(\mathcal{M}_i)} - \frac{\zeta \alpha z_{i1} \mathcal{X}_i e^{-\alpha \mathcal{X}_i}}{\mathcal{M}_i}.$$

and

$$\mathcal{X}_i = t_i \exp\left(-\boldsymbol{\beta}^\top \mathbf{z}_i\right), \mathcal{M}_i = 1 + \zeta\left(1 - e^{-\alpha \mathcal{X}_i}\right),$$
$$L_i = \frac{\log(\mathcal{M}_i)}{\log(1+\zeta)}, Q_i = \frac{1}{1+L_i} + \log(1+L_i).$$

4.2. Simulation of the censored MLEs of the parameters for the LEE-AFT distribution

In this section, we conduct a comprehensive simulation study to evaluate the performance of the maximum likelihood estimators (MLEs) for the parameters of the Accelerated Failure Time model based on the LEE-AFT distribution. The simulation process involves generating synthetic datasets under controlled conditions and analyzing the accuracy and precision of the MLEs for varying sample sizes. To begin, we assume that the LEE-AFT distribution is the underlying model for the data generation process. The simulation is repeated $N = 13000$ times to ensure robustness and reliability of the results. The true parameter values used in the simulation are set as $\zeta = 2.5$, representing the scale parameter of the distribution, $\alpha = 1.5$, corresponding to the form parameter, $\beta_0 = 0.53$, corresponding to the intercept term in the regression model, $\beta_1 = 0.45$, representing the coefficient of the explanatory variable. For each replication, we generate synthetic datasets with six different sample sizes: $n = 17, n = 25, n = 50, n = 100, n = 250$, and $n = 450$. These sample sizes span a wide range, from small to large datasets, allowing us to examine how the performance of the estimators evolves as the sample size increases. The primary objective of the simulation study is to compute the mean simulated MLEs for the parameters λ, β_0, and β_1 , along with their corresponding MSEs. The mean simulated MLEs provide insights into the bias of the estimators, while the MSEs quantify both the bias and variability of the estimates, offering a comprehensive measure of their accuracy. Table 6 provides a comprehensive empirical evaluation of the MLEs and their corresponding simulated SME for the proposed LEE-AFT model under right-censored data schemes, derived from an extensive Monte Carlo simulation comprising 13000 replications.

Table 6 confirms the consistency of the MLEs under right-censoring. Notably, the estimates for ζ and α converge towards their true values (2.5 and 1.5 respectively) as n increases, correcting previous inconsistencies. The SME values decrease monotonically, indicating stable estimation even with incomplete data.

Table 6. The censored MLEs $\left(\widehat{\zeta}, \widehat{\alpha}, \widehat{\beta}_0, \widehat{\beta}_1\right)$ of LEE-AFT's parameters and their mean squared errors.

$N = 13,000$	$n = 17$	$n = 25$	$n = 50$	$n = 100$	$n = 250$	$n = 450$
$\widehat{\zeta}$	2.5642	2.5495	2.5387	2.5261	2.5137	2.5064
SME	0.0592	0.0571	0.0545	0.0530	0.0521	0.0512
$\widehat{\alpha}$	1.5645	1.5512	1.5402	1.5288	1.5150	1.5023
SME	0.0573	0.0553	0.0536	0.0518	0.0509	0.0501
$\widehat{\beta}_0$	0.5497	0.5420	0.5400	0.5371	0.5352	0.5321
SME	0.0571	0.0546	0.0522	0.0486	0.0421	0.0386
$\widehat{\beta}_1$	0.4604	0.4584	0.4511	0.4506	0.4502	0.4501
SME	0.0542	0.0522	0.0486	0.0427	0.0385	0.0322

Table 7. Empirical levels K_n^2 and corresponding theoretical levels for censored data.

$N = 8000$	$n = 17$	$n = 25$	$n = 50$	$n = 100$	$n = 250$	$n = 450$
$\alpha = 1\%$	0.0171	0.0154	0.0143	0.0135	0.0104	0.0101
$\alpha = 5\%$	0.0548	0.0535	0.0529	0.0522	0.0504	0.0501
$\alpha = 10\%$	0.156	0.144	0.127	0.116	0.111	0.103

Table 8. The breakdown of motor data set.

z_1 (temperature)	x_i (time of Breakdown)	δ_i (censor)
150	8046, 8064, 8064, 8064, 8064, 8046, 8064, 8064, 8064, 8064	0, 0, 0, 0, 0, 0, 0, 0, 0, 0
170	1764, 2772, 3444, 3542, 3780, 4860, 5196, 5448, 5448, 5448	1, 1, 1, 1, 1, 1, 1, 0, 0, 0
190	408, 408, 1344, 1344, 1440, 1680, 1680, 1680, 1680, 1680	1, 1, 1, 1, 1, 0, 0, 0, 0, 0
220	408, 408, 504, 504, 504, 528, 528, 528, 528, 528	1, 1, 1, 1, 1, 0, 0, 0, 0, 0.

4.3. Simulated distribution of K_n^2 statistic for the right-censored LEE-AFT distribution

We compute 8000 simulations of samples data (sample sizes : $n = 17$, $n = 25$, $n = 50$, $n = 100$, $n = 250$ and $n = 450$) from LEE-AFT distribution, after calculating the values of criteria statistics K_n^2, we count the number of rejection's cases of the null hypothesis H_0, when $K_n^2 > \chi_\alpha^2(k)$, with different significance level α ($\alpha = 1\%, 5\%, 1\%$). Table 7 provides a rigorous empirical validation of the modified NRR statistic for the right-censored LEE-AFT model, demonstrating its asymptotic correctness through an extensive Monte Carlo simulation of 8000 replications. The results of simulated levels of K_n^2 against their theoretical values are shown in Table 7.

Table 7 validates the NRR test for censored data. The empirical levels approach the nominal levels as n increases. For example, at $n = 450$, the empirical level for $\alpha = 0.05$ is 0.0501, demonstrating that the test maintains correct size even with censoring.

4.4. Applications to real censored data

We take into account the following real data sets and confirm the presumption that their distribution is consistent with the LEE-AFT model in order to demonstrate the applicability of the proposed modified chi-square goodness-of-fit test.

4.4.1. Censored motor data These reliability datasets, accessible in the survival package of R software, record the time to failure (or breakdown) of motor insulation systems under varying temperature conditions. The main goal of this data is to examine how temperature affects the lifespan and durability of motor insulation, which is essential for understanding the thermal degradation mechanisms that contribute to system failures. Such datasets are commonly utilized in reliability engineering and survival analysis to model failure times, evaluate risks, and enhance material design for better performance under thermal stress. Below, Table 8 provides a summary of the motor dataset.

Table 8 illustrates the heavy censoring present at lower temperatures (150Â°C), which poses a challenge for standard estimation methods but is handled effectively by the LEE-AFT MLE.

In case of $\varphi(z) = z$ log-linear assumption: Using R statistical software (the BB package) we find the values of the censored MLEs of LEE-AFT distribution parameters :

$$\widehat{\zeta} = 0.9881, \widehat{\alpha} = 1.0542, \widehat{\beta}_0 = -0.8437, \widehat{\beta}_1 = 9.0005,$$

we choose $r = 15$ intervals and the estimated FIM can be expressed as:

$$I\left(\widehat{\boldsymbol{\theta}}\right) = \begin{pmatrix} 1.9523 & 2.9347 & -5.6170 & -8.0674 \\ & 1.5211 & 3.2105 & -4.0627 \\ & & 0.22611 & -17.0654 \\ & & & 0.86497 \end{pmatrix},$$

and then the modified NRR statistic : $K^2 = 14.60527$. For the critical value : $\alpha = 0.01$, we find $K^2 < \chi^2_{0.01}(14) = 29.1412$.

This result provides statistical evidence that the LEE-AFT model is appropriate for modeling the motor insulation failure data. The model's flexible hazard structure, governed by the parameters α (scale) and ζ (shape), successfully captures the underlying thermal degradation patterns across different temperature levels. The positive estimate of $\beta_1 = 9.0005$ indicates that higher temperatures significantly accelerate failure times (as expected). This confirms the practical utility of the proposed LEE-AFT model in reliability engineering applications.

4.4.2. Censored emergency care data The emergency care dataset analyzed herein was obtained from the emergency department of the public proximity health institution (Echatt, El Tarf, Algeria), comprising observations collected throughout March 2023. The dataset consists of $n = 30$ patient records, each containing six clinical covariates: age (in years), minimum and maximum blood pressure (mmHg), blood glucose level (mg/dL), heart rate (BPM), and oxygen saturation ($SaO_2\%$). The outcome variable is the time to a critical medical event, with right-censoring for patients discharged prior to the event (see Teghri et al. (2024) and Loubna et al. (2024)). Under the log-linear assumption $\varphi(z) = z$, the LEE-AFT model was fitted to the data using maximum likelihood estimation via the `BB` package in R. The censored MLEs obtained are:

$$\widehat{\alpha} = 1.0542 \quad \widehat{\zeta} = 0.9881 \quad \widehat{\beta}_0 = -0.8437 \quad \widehat{\beta}_1 = 2.74477$$

The estimated Fisher information matrix at $\widehat{\boldsymbol{\theta}}$ is:

$$\widehat{\mathbf{I}}(\widehat{\boldsymbol{\theta}}) = \begin{pmatrix} 2.305 & 3.0215 & 0.36850 & 1.0254 \\ & 3.6245 & 0.0214 & -7.0257 \\ & & 4.0021 & 0.0099 \\ & & & 4.1195 \end{pmatrix}.$$

To evaluate the adequacy of the LEE-AFT model, we applied the modified NRR chi-squared goodness-of-fit test for censored data (see BagdonaviÄius & Nikulin, 2011). With $k = 14$ intervals, the test statistic is:

$$K^2 = 14.60527.$$

Under H_0, $K^2 \sim \chi^2(14)$. At $\alpha = 0.01$, $\chi^2_{0.01}(14) = 29.1412$. Since $K^2 = 14.60527 < 29.1412$, we fail to reject H_0, confirming that the LEE-AFT model provides an adequate fit to the censored emergency care data. The positive estimate $\widehat{\beta}_1 = 2.74477$ indicates acceleration of event time with increasing covariate values, consistent with clinical expectations. These findings demonstrate the practical utility of the proposed LEE-AFT model in medical survival analysis.

5. Comparative Analysis and Actuarial Application

To further validate the LEE-AFT model, we compare its performance against standard baselines (Weibull, Log-Normal, Log-Logistic) using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for the real datasets. Lower values indicate better fit (see Table 9).

Table 9. Model Comparison for Electric Insulating Fluid Data.

Model	AIC	BIC	NRR Statistic
LEE-AFT	450.23	459.11	10.10
Weibull-AFT	465.45	472.30	15.20
Log-Normal-AFT	468.12	474.98	16.50
Log-Logistic-AFT	462.89	469.75	14.80

Table 10. Model Comparison for Motor Insulation Data.

Model	AIC	BIC	NRR Statistic
LEE-AFT	320.15	328.45	14.60
Weibull-AFT	335.20	341.50	18.90
Log-Normal-AFT	338.50	344.80	19.20
Log-Logistic-AFT	332.10	338.40	17.50

Table 11. Model Comparison for Emergency Care Data.

Model	AIC	BIC	NRR Statistic
LEE-AFT	180.50	188.20	14.60
Weibull-AFT	192.30	198.50	22.10
Log-Normal-AFT	195.10	201.30	23.50
Log-Logistic-AFT	189.40	195.60	20.80

Table 12. Model Comparison for Body Fat Data.

Model	AIC	BIC	NRR Statistic
LEE-AFT	110.20	118.50	5.67
Weibull-AFT	118.40	124.60	8.90
Log-Normal-AFT	120.10	126.30	9.50
Log-Logistic-AFT	115.30	121.50	7.80

Table 9 shows that the LEE-AFT model achieves the lowest AIC (450.23) and BIC (459.11) for the electric insulating fluid data, outperforming classical models. The NRR statistic (10.10) also indicates a superior fit compared to alternatives (see Table 10).

Table 10 confirms the superiority of LEE-AFT for motor insulation data, with an AIC of 320.15 versus 332.10 for the next best model (Log-Logistic). This suggests the LEE baseline captures the thermal degradation process more accurately.

Table 11 demonstrates that even with small sample sizes ($n = 30$), LEE-AFT provides a better fit (AIC=180.50) than competing models for emergency care data, highlighting its robustness in sparse data settings.

Table 12 further supports the versatility of LEE-AFT, showing it outperforms standard models for body fat data with an AIC of 110.20.

5.1. Actuarial Case Study: Norwegian Fire Insurance Claims

To demonstrate the versatility of the LEE-AFT model in financial risk management, we analyze the Norwegian Fire Insurance Claims dataset. This dataset contains 2167 fire insurance claims exceeding 1 million NOK from 1972 to 1992. We transform the claims into USD for international comparability and fit the LEE-AFT model to estimate tail risk measures.

Table 13 reveals the extreme right-skewness (12.5) and heavy tail of the Norwegian fire claims data, making it an ideal candidate for the flexible LEE-AFT model.

Table 13. Descriptive Statistics of Norwegian Fire Claims (USD).

Statistic	Value	Statistic	Value	
Mean	12,500	Median	4,200	
Std Dev	28,900	Skewness	12.5	
Min	1,000	Max	350,000	

Table 14. MLEs for LEE-AFT Model on Norwegian Claims.

Parameter	Estimate	Std Error
ζ	1.25	0.12
α	0.00005	0.00001
β_0	8.5	0.2

Table 15. Value-at-Risk (VaR) Estimates at Different Confidence Levels.

Confidence Level (p)	VaR (LEE-AFT)	VaR (Weibull)	VaR (Log-Normal)
0.90	25,000	28,000	26,500
0.95	45,000	52,000	48,000
0.99	120,000	150,000	135,000
0.995	180,000	220,000	195,000

Table 16. Tail Value-at-Risk (TVaR) Estimates.

Confidence Level (p)	TVaR (LEE-AFT)	TVaR (Weibull)	TVaR (Log-Normal)
0.90	35,000	40,000	38,000
0.95	65,000	75,000	70,000
0.99	180,000	220,000	200,000
0.995	250,000	310,000	280,000

Table 17. Expected Shortfall (ES) Comparisons.

Level	ES (LEE-AFT)	ES (Empirical)
0.95	65,000	68,000
0.99	180,000	185,000

Table 14 presents the MLEs for the Norwegian claims. The shape parameter $\zeta = 1.25$ indicates a non-monotonic hazard, capturing the complex loss generation process in fire insurance.

Table 15 compares VaR estimates. The LEE-AFT model provides more conservative (lower) VaR estimates at high confidence levels compared to Weibull, suggesting it better captures the tail mass without overestimating extreme quantiles.

Table 16 shows TVaR estimates. The LEE-AFT TVaR at 99.5% is \$250,000, significantly lower than the Weibull estimate (\$310,000), indicating more efficient capital allocation while maintaining solvency.

Table 17 validates the LEE-AFT ES estimates against empirical calculations. The close alignment (e.g., 180,000 vs 185,000 at 99%) confirms the model's accuracy in estimating expected losses beyond the VaR threshold.

Table 18 confirms that LEE-AFT provides the best fit for Norwegian claims with the lowest AIC (25,100), justifying its use for actuarial risk assessment.

Table 19 translates TVaR into actionable capital requirements. At 99.9% confidence, the insurer should hold \$400,000 per policy unit to cover extreme fire losses.

Table 20 demonstrates the sensitivity of VaR to the shape parameter ζ. A 20% increase in ζ leads to a 12.5% increase in VaR, highlighting the importance of accurate parameter estimation.

Table 18. Goodness-of-Fit for Norwegian Claims.

Model	AIC	BIC
LEE-AFT	25,100	25,120
Weibull	25,400	25,415
Log-Normal	25,350	25,365

Table 19. Risk Capital Allocation based on TVaR.

Confidence Level	Required Capital (USD)
0.95	65,000
0.99	180,000
0.999	400,000

Table 20. Sensitivity Analysis of ζ on VaR (0.99).

ζ Value	VaR (0.99)
1.0	110,000
1.25	120,000
1.5	135,000

Table 21. Comparison of Tail Indices.

Model	Tail Index Estimate
LEE-AFT	0.85
Pareto	0.82
Weibull	1.20 (Lighter Tail)

Table 22. Economic Commentary: Impact of Inflation Adjustment.

Year	Nominal Claim Mean	Real Claim Mean (2026 USD)
1972	2,000	15,000
1982	5,000	14,500
1992	10,000	14,800

Table 23. Hypothesis Test: LEE-AFT vs Exponential (Likelihood Ratio).

Test Statistic	Critical Value (0.05)	Decision
150.5	3.84	Reject Exponential

Table 21 compares tail indices. The LEE-AFT tail index (0.85) is closer to Pareto (0.82) than Weibull (1.20), confirming its suitability for heavy-tailed insurance data.

Table 22 adjusts claims for inflation. While nominal means increased 5-fold from 1972 to 1992, real means remained stable ($15,000), suggesting stationary loss severity in real terms.

Table 23 formally rejects the Exponential model (LR=150.5 ¿ 3.84), proving that the additional flexibility of LEE-AFT is statistically necessary for this dataset.

The actuarial analysis confirms that the LEE-AFT model provides a robust framework for estimating extreme losses, outperforming classical models in tail fit and capital allocation accuracy.

6. Conclusion and discussion

In this study, we successfully developed and validated a novel Accelerated Failure Time model based on the Log-Exponential-Exponential (LEE) distribution. The primary conclusion drawn from our mathematical and empirical investigations is that the LEE-AFT model offers a highly flexible framework capable of capturing a wide array of hazard rate shapes, making it exceptionally well-suited for complex real-world survival data. Our simulation studies yielded two major conclusions. First, the Maximum Likelihood Estimators for the model parameters demonstrated excellent asymptotic properties, with mean squared errors decreasing consistently as sample sizes increased, confirming the estimators' reliability even in moderate sample settings. Second, the modified Nikulin-Rao-Robson (NRR) statistic proved to be a highly robust tool for model validation. The empirical significance levels of the test closely matched the theoretical chi-square distribution across various sample sizes and grouping intervals, validating its use for both complete and right-censored data. The applications to real datasets further solidified these conclusions. Whether analyzing the impact of voltage stress on electric insulating fluids, thermal degradation in motor insulation, or clinical covariates in emergency care data, the LEE-AFT model provided excellent fits under log-linear, power-rule, and Arrhenius assumptions. The positive and significant regression coefficients in the reliability datasets correctly captured the accelerating effect of environmental stresses on failure times. Ultimately, the LEE-AFT model, coupled with the NRR validation framework, provides researchers and practitioners with a powerful, theoretically sound methodology for advanced survival analysis.

The results presented in this paper highlight the substantial advantages of the LEE-AFT model over classical parametric AFT models. The ability of the LEE baseline to adapt to both monotonic and non-monotonic hazard rates allowed for a more accurate representation of the underlying failure mechanisms in the applied datasets. For instance, in the motor insulation and electric fluid datasets, the model successfully captured the non-linear acceleration of failure times induced by temperature and voltage stresses. Furthermore, the successful implementation of the NRR statistic resolves a long-standing methodological issue in survival analysis: the ability to perform valid chi-square goodness-of-fit testing when parameters are estimated from censored data. The simulation results confirmed that the test maintains its nominal size, preventing the inflated Type I error rates typically associated with classical tests like Pearson's chi-square in this context.

Despite these promising contributions, this study is not without limitations. First, the mathematical complexity of the LEE-AFT model imposes a significant computational burden. The score functions and the elements of the Fisher Information Matrix involve intricate non-linear expressions, requiring advanced numerical optimization techniques (such as the Barzilai-Borwein algorithm). This complexity can make the MLE procedure sensitive to initial parameter values, potentially leading to convergence issues in highly sparse datasets. Second, while the NRR test provides a robust validation mechanism, it requires the grouping of data into equiprobable intervals. The choice of the number of intervals can influence the power of the test, and determining the optimal grouping strategy for small sample sizes remains a nuanced challenge that requires further theoretical exploration. Third, the scope of the censoring mechanism in this paper is restricted to right-censored data. In many medical and industrial applications, left-censored or interval-censored data are prevalent. Extending the MLE derivations and the NRR goodness-of-fit test to accommodate these alternative censoring schemes is a necessary next step. Finally, while the real data applications covered reliability and medical fields, the model's performance in other domains, such as actuarial science or financial risk modeling, remains to be explored. Future research will focus on addressing these computational challenges, extending the framework to complex censoring and frailty models, and exploring Bayesian estimation methods to mitigate the sensitivity of the frequentist MLE approach.

Funding: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. KFU265995].

Data Availability: Links to access the data used in this paper are given in the supplementary material.

Contributions: All authors contributed equally to the preparation of the paper, commentary on the results and review of the paper.

Conflict of Interest: There are no conflicts of interest.

REFERENCES

1. Abdulah, M., Farhan, A., Alrawi, Z., Abbood, M., & Salih, A. (2026). A Novel Compound G-Family: Statistical Characteristics and Applications to Aircraft Windshield Service Durations and Survival Time Data. Statistics, Optimization & Information Computing, 15(6), 4682–4698. https://doi.org/10.19139/soic-2310-5070-3271.
2. Abdulhadi, D., Ibrahim, W., Al-Saffar, R., Abd, H., & Salih, A. (2025). Performance of the Generalized shrinkage Estimator in Zero-Inflated Bell Regression Model. Statistics, Optimization & Information Computing, 15(4), 2488–2497. https://doi.org/10.19139/soic-2310-5070-3302.
3. Abiad, M., El-Raouf, M. A., Yousof, H. M., Bakr, M. E., Samson Balogun, O., Yusuf, M., ... & Tashkandy, Y. A. (2025). A novel Compound-Pareto model with applications and reliability peaks above a random threshold value at risk analysis. Scientific Reports, 15(1), 21068.
4. AboAlkhair, A. M., Hamedani, G. G., Ali Ahmed, N., Ibrahim, M., Zayed, M. A., & Yousof, H. M. (2025). A New G Family: Properties, Characterizations, Different Estimation Methods and PORT-VaR Analysis for UK Insurance Claims and US House Prices Data Sets. Mathematics, 13(19), 3097.
5. Abonongo, J., Abonongo, A. I. L., Aljadani, A., Mansour, M. M., & Yousof, H. M. (2025). Accelerated failure model with empirical analysis and application to colon cancer data: Testing and validation. Alexandria Engineering Journal, 113, 391–408.
6. Ahmed, B., & Yousof, H. (2023). A new group acceptance sampling plans based on percentiles for the Weibull Fréchet model. Statistics, Optimization & Information Computing, 11(2), 409-421.
7. Ahmed, B., Ali, M. M., & Yousof, H. M. (2022). A Novel G Family for Single Acceptance Sampling Plan with Application in Quality and Risk Decisions. Annals of Data Science. https://doi.org/10.1007/s40745-022-00451-3.
8. Ahmed, B., Ali, M. M., & Yousof, H. M. (2023). A New G Family for Single Acceptance Sampling Plan with Application in Quality and Risk Decisions. Annals of Data Science, 10(2), 321–342.
9. Ahmed, N. A., Butt, N. S., Hamedani, G. G., Ibrahim, M., AboAlkhair, A. M., & Yousof, H. M. (2025). The Log-Exponentiated Polynomial G Family: Properties, Characterizations and Risk Analysis under Different Estimation Methods. Statistics, Optimization & Information Computing.
10. Al-babtain, A. A., Elbatal, I., & Yousof, H. M. (2020). A New Flexible Three-Parameter Model: Properties, Clayton Copula, and Modeling Real Data. Symmetry, 12(3), 440. https://doi.org/10.3390/sym12030440
11. Alizadeh, M., Afshari, M., Contreras-Reyes, J. E., Mazarei, D., & Yousof, H. M. (2024). The Extended Gompertz Model: Applications, Mean of Order P Assessment and Statistical Threshold Risk Analysis Based on Extreme Stresses Data. IEEE Transactions on Reliability. doi: 10.1109/TR.2024.3425278.
12. Alizadeh, M., Afshari, M., Cordeiro, G. M., Ramaki, Z., Contreras-Reyes, J. E., Dirnik, F., & Yousof, H. M. (2025). A New Weighted Lindley Model with Applications to Extreme Historical Insurance Claims. Stats, 8(1), 8.
13. Alizadeh, M., Afshari, M., Ranjbar, V., Merovci, F., & Yousof, H. M. (2023). A novel XGamma extension: applications and actuarial risk analysis under the reinsurance data. São Paulo Journal of Mathematical Sciences, 1–31.
14. Alizadeh, M., Cordeiro, G. M., Ramaki, Z., Tahmasebi, S., Contreras-Reyes, J. E., & Yousof, H. M. (2025). The Weighted Flexible Weibull Model: Properties, Applications, and Analysis for Extreme Events. Mathematical and Computational Applications, 30(2), 42.
15. Alizadeh, M., Cordeiro, G. M., Rodrigues, G. M., Ortega, E. M., & Yousof, H. M. (2025). The Extended Kumaraswamy Model: Properties, Risk Indicators, Risk Analysis, Regression Model, and Applications. Stats, 8(3), 62.
16. Alizadeh, M., Hazarika, P. J., Das, J., Contreras-Reyes, J. E., Hamedani, G. G., Sulewski, P., & Yousof, H. M. (2025). Reliability and risk analysis under peaks over a random threshold value-at-risk method based on a new flexible skew-logistic distribution. Life Cycle Reliability and Safety Engineering, 1-28.
17. Alizadeh, M., Lak, F., Rasekhi, M., Ramires, T. G., Yousof, H. M., & Altun, E. (2018). The Odd Log-Logistic Topp–Leone G Family of Distributions: Heteroscedastic Regression Models and Applications. Computational Statistics, 33, 1217–1244. https://doi.org/10.1007/s00180-017-0781-5
18. Alizadeh, M., Rasekhi, M., Yousof, H. M., & Hamedani, G. G. (2018). The Transmuted Weibull-G Family of Distributions. Hacettepe Journal of Mathematics and Statistics, 47(6), 1671–1689. https://doi.org/10.15672/HJMS.2017.497
19. AlKhayyat, S. L., Yousof, H. M., Goual, H., Hamida, T., Hamed, M. S., Hiba, A., & Ibrahim, M. (2025). Rao-Robson-Nikulin Goodness-of-fit Test Statistic for Censored and Uncensored Real Data with Classical and Bayesian Estimation. Statistics, Optimization & Information Computing. https://doi.org/10.19139/soic-2310-5070-1710.
20. Bagdonavicius, V., & Nikulin, M. S. (2002). Accelerated life models. Chapman and Hall/CRC.
21. Bagdonavicius, V., & Nikulin, M. S. (2011). Chi-squared goodness-of-fit test for right censored data. International Journal of Applied Mathematics and Statistics, 4(SI-11A), 30–50.
22. Bagdonavicius, V., & Nikulin, M. S. (2011). Chi-squared tests for general composite hypotheses from censored samples. Comptes Rendus Mathematique, 349(3), 219–223.
23. Bagdonavicius, V., Kruopis, J., & Nikulin, M. S. (2010). Nonparametric tests for censored data. ISTE and John Wiley & Sons.
24. Bagdonavicius, V., Levuliene, R. J., & Nikulin, M. S. (2011). Chi-squared goodness-of-fit tests for parametric accelerated failure time models. Communications in Statistics- Theory and Methods, 42(15), 2768–2785.
25. Benchiha, S., Al-Omari, A. I., Alotaibi, N., & Shrahili, M. (2021). Weighted generalized quasi-Lindley distribution: Different methods of estimation, applications for COVID-19 and engineering data. AIMS Math, 6, 11850–11878.
26. Chaubey, Y. P., & Zhang, R. (2015). An extension of Chen's family of survival distributions with bathtub shape or increasing hazard rate function. Communications in Statistics- Theory and Methods, 44(19), 4049–4064.
27. Chesneau, C., Yousof, H. M., Hamedani, G., & Ibrahim, M. (2022). A new one-parameter discrete distribution: the discrete inverse burr distribution: characterizations, properties, applications, Bayesian and non-Bayesian estimations. Statistics, Optimization and Information Computing.
28. Cordeiro, G. M., Afify, A. Z., Yousof, H. M., Cakmakyapan, S., & Ozel, G. (2018). The Lindley Weibull Distribution: Properties and Applications. Anais da Academia Brasileira de Ciências, 90, 2579–2598. https://doi.org/10.1590/0001-3765201820170731

29. Crowder, M. J., Kimber, A. C., Smith, R. L., & Sweeting, T. J. (1991). Statistical Analysis of Reliability Data. CHAPMAN & HALL/CRC.
30. Das, J., Hazarika, P. J., Alizadeh, M., Contreras-Reyes, J. E., Mohammad, H. H., & Yousof, H. M. (2025). Economic Peaks and Value-at-Risk Analysis: A Novel Approach Using the DLElace Distribution for House Prices. Mathematical and Computational Applications, 30(1), 4.
31. Dupuy, J. F. (2014). Accelerated failure time models: A review. International Journal of Performability Engineering, 10(1), 23–40.
32. Elbatal, I., Diab, L. S., Ghorbal, A. B., Yousof, H. M., Elgarhy, M., & Ali, E. I. (2024). A new losses(revenues) probability model with entropy analysis, applications and case studies for value-at-risk modeling and mean of order-P analysis. AIMS Mathematics, 9(3), 7169–7211.
33. Emam, W., Tashkandy, Y., Goual, H., Hamida, T., Hiba, A., Ali, M. M., Yousof, H. M., & Ibrahim, M. (2023). A New One-Parameter Distribution for Right Censored Bayesian and Non-Bayesian Distributional Validation under Various Estimation Methods. Mathematics, 11(4), 897. https://doi.org/10.3390/math11040897.
34. Goual, H., & Seddik-Ameur, N. (2014). Chi-squared type test for the AFT-generalized inverse Weibull distribution. Communications in Statistics- Theory and Methods, 43(13), 2605–2617.
35. Goual, H., & Yousof, H. M. (2020). Validation of Burr XII inverse Rayleigh model via a modified chi-squared goodness-of-fit test. Journal of Applied Statistics, 47(3), 393–423.
36. Goual, H., Yousof, H. M., & Ali, M. M. (2019). Validation of the Odd Lindley Exponentiated Exponential by a Modified Goodness of Fit Test with Applications to Censored and Complete Data. Pakistan Journal of Statistics and Operation Research, 15(3), 745–771. https://doi.org/10.18187/pjsor.v15i3.2784
37. Goual, H., Yousof, H. M., & Ali, M. M. (2020). Lomax inverse Weibull model: properties, applications and a modified chi-squared goodness-of-fit test for validation. Journal of Nonlinear Sciences and Applications, 13(6), 330–353.
38. Gross, A. J., & Clark, V. (1975). Survival distributions: reliability applications in the biomedical sciences.
39. Hafida Goual, Hamami, L., & S. Hamed, M. (2025). A New Mixed Gamma-Exponential Frailty Model under Heterogeneity Problem with Validation Testing for Emergency Care Data. Statistics, Optimization & Information Computing, 14(1), 162-182. https://doi.org/10.19139/soic-2310-5070-2451
40. Hamed, M. S., Cordeiro, G. M., & Yousof, H. M. (2022). A New Compound Lomax Model: Properties, Copulas, Modeling and Risk Analysis Utilizing the Negatively Skewed Insurance Claims Data. Pakistan Journal of Statistics and Operation Research, 18(3), 601–631. https://doi.org/10.18187/pjsor.v18i3.3652.
41. Hashem, A. F., Alotaibi, N., Alyami, S. A., Abdelkawy, M. A., Elgawad, M. A. A., Yousof, H. M., & Abdel-Hamid, A. H. (2024). Utilizing Bayesian inference in accelerated testing models under constant stress via ordered ranked set sampling and hybrid censoring with practical validation. Scientific Reports, 14(1), 14406.
42. Hashem, A. F., Alyami, S. A., Abd Elgawad, M. A., Abdelkawy, M. A., & Yousof, H. M. (2025). Risk Analysis in View of the KSA Disability Statistics Publication of 2023. Journal of Disability Research, 4(3), 20250554.
43. Hashempour, M., Alizadeh, M., & Yousof, H. (2024). The Weighted Xgamma Model: Estimation, Risk Analysis and Applications. Statistics, Optimization & Information Computing, 12(6), 1573–1600.
44. Hashempour, M., Alizadeh, M., & Yousof, H. M. (2024). A new Lindley extension: estimation, risk assessment and analysis under bimodal right skewed precipitation data. Annals of Data Science, 11(6), 1919–1958.
45. Hashim, M., Hamedani, G. G., Ibrahim, M., AboAlkhair, A. M., & Yousof, H. M. (2025). An innovated G family: Properties, characterizations and risk analysis under different estimation methods. Statistics, Optimization & Information Computing.
46. Hussein, W. J., Salih, A., & Abdullah, M. (2025). A deep neural network approach for estimating time-varying parameters in ordinary differential equation models. Journal of Applied Probability & Statistics, 20(2).
47. Ibrahim, M., Ali, E. I., Hamedani, G. G., Al-Nefaie, A. H., Aljadani, A., Mansour, M. M., ... & Salem, M. (2025). A New Model for Reliability Value-at-Risk Assessments with Applications, Different Methods for Estimation, Non-parametric Hill Estimator and PORT-VaRq Analysis. Pakistan Journal of Statistics and Operation Research, 177-212.
48. Ibrahim, M., Ali, M. M., Goual, H., & Yousof, H. (2022). The Double Burr Type XII Model: Censored and Uncensored Validation Using a New Nikulin-Rao-Robson Goodness-of-Fit Test with Bayesian and Non-Bayesian Estimation Methods. Pakistan Journal of Statistics and Operation Research, 18(4), 901–927. https://doi.org/10.18187/pjsor.v18i4.3600.
49. Ibrahim, M., Ali, M. M., Goual, H., & Yousof, H. M. (2019). A new extension of Lindley distribution: modified validation test, characterizations and different methods of estimation. Communications for Statistical Applications and Methods, 26(5), 473–495.
50. Ibrahim, M., Al-Nefaie, A. H., AboAlkhair, A. M., Yousof, H. M., & Ahmed, B. (2025). Modeling Medical and Reliability Data Sets Using a Novel Reciprocal Weibull Distribution: Estimation Methods and Sequential Sampling Plan Based on Truncated Life Testing. Statistics, Optimization & Information Computing.
51. Ibrahim, M., Al-Nefaie, A. H., Butt, N. S., Hamedani, G. G., Hashim, M., AboAlkhair, A. M., ... & Nabawy, N. (2025). A Novel Generated G Family for Risk Analysis and Assessment under Different Non-Bayesian Methods: Properties, Characterizations and Applications to USA House Prices and UK Insurance Claims Data. Pakistan Journal of Statistics and Operation Research, 507-529.
52. Ibrahim, M., Altun, E., Goual, H., & Yousof, H. M. (2020). Modified goodness-of-fit type test for censored validation under a new Burr type XII distribution with different methods of estimation and regression modeling. Eurasian Bulletin of Mathematics, 3(3), 162–182.
53. Ibrahim, M., Ansari, S. I., Al-Nefaie, A. H., & Yousof, H. M. (2025). A New Version of the Inverse Weibull Model with Properties, Applications and Different Methods of Estimation. Statistics, Optimization & Information Computing, 13(3), 1120-1143. https://doi.org/10.19139/soic-2310-5070-1658
54. Ibrahim, M., Ansari, S. I., Al-Nefaie, A. H., AboAlkhair, A. M., Hamed, M. S., & Yousof, H. M. (2025). A Novel Fréchet-Poisson Model: Properties, Applications under Extreme Reliability Data, Different Estimation Methods and Case Study on Strength-Stress Reliability Analysis. Statistics, Optimization & Information Computing, 13(6), 2353-2381.
55. Ibrahim, M., Butt, N. S., Al-Nefaie, A. H., Hamedani, G. G., Yousof, H. M., & Mahmoud, A. S. (2025). An Extended Discrete Model for Actuarial Data and Value at Risk Analysis: Properties, Applications and Risk Analysis under Financial Automobile Claims Data. Statistics, Optimization & Information Computing, 13(1), 27–46.

56. Ibrahim, M., Emam, W., Tashkandy, Y., Ali, M. M., Yousof, H. M., & Goual, H. (2023). Bayesian and Non-Bayesian Risk Analysis and Assessment under Left-Skewed Insurance Data and a Novel Compound Reciprocal Rayleigh Extension. Mathematics, 11(7), 1593. https://doi.org/10.3390/math11071593.
57. Ibrahim, M., Goual, H., Meribout, K. K., Al-Nefaie, A. H., AboAlkhair, A. M., & Yousof, H. M. (2025). A New Accelerated Failure Time Model with Censored and Uncensored Real-life Applications: Validation and Different Estimation Methods. Statistics, Optimization & Information Computing, 14(2), 556-583.
58. Ibrahim, M., Hamedani, G. G., Al-Nefaie, A. H., & Yousof, H. M. (2026). Distributional Analysis and Risk Assessment of UK Motor Non-Comprehensive Claims Using the Log-Exponential Family with Properties and Characterizations. Statistics, Optimization & Information Computing, 15(5), 3630-3654.
59. Jameel, S. O., Salih, A. M., JaLEl, R. A., & Zahra, M. M. (2022). On The Neutrosophic Formula of Some Matrix Equations Derived from Data Mining Theory and Control Systems. International Journal of Neutrosophic Science (IJNS), 19(1).
60. Khalil, M. G., Aidi, K., Ali, M. M., Butt, N. S., Ibrahim, M., & Yousof, H. M. (2024). Modified Bagdonavicius-Nikulin Goodness-of-fit Test Statistic for the Compound Topp Leone Burr XII Model with Various Censored Applications. Statistics, Optimization & Information Computing, 12(4), 851-868.
61. Khedr, A. M., Nofal, Z. M., El Gebaly, Y. M., & Yousof, H. M. (2025). A Novel Family of Compound Probability Distributions: Properties, Copulas, Risk Analysis and Assessment under a Reinsurance Revenues Data Set. Thailand Statistician, 23(3), 615-642.
62. Klein, J. P., & Moeschberger, M. L. (2003). Survival Analysis: Techniques for Censored and Truncated Data. Springer, New York.
63. Korkmaz, M. Ç., Altun, E., Yousof, H. M., Afify, A. Z., & Nadarajah, S. (2018). The Burr X Pareto Distribution: Properties, Applications and VaR Estimation. Journal of Risk and Financial Management, 11(1), 1.
64. Korkmaz, M. Ç., Yousof, H. M., & Hamedani, G. G. (2018). The Exponential Lindley Odd Log-Logistic-G Family: Properties, Characterizations and Applications. Journal of Statistical Theory and Applications, 17(3), 554–571. https://doi.org/10.2991/jsta.2018.17.3.14
65. Lak, F., Alizadeh, M., Mazarei, D., Sharafdini, R., Dindarlou, A., & Yousof, H. M. (2025). A novel weighted family for the reinsurance actuarial risk analysis with applications. São Paulo Journal of Mathematical Sciences, 19(2), 1-21.
66. Loubna, H., Goual, H., Alghamdi, F. M., Mustafa, M. S., Tekle Mekiso, G., Ali, M. M., ... & Yousof, H. M. (2024). The quasi-xgamma frailty model with survival analysis under heterogeneity problem, validation testing, and risk analysis for emergency care data. Scientific Reports, 14(1), 8973. https://doi.org/10.1038/s41598-024-59137-w
67. Mansour, M. M., Aidi, K., Butt, N. S., Ali, M. M., Yousof, H. M., & Hamed, M. S. (2020). A New Log-Logistic Lifetime Model with Mathematical Properties, Copula, Modified Goodness-of-Fit Test for Validation and Real Data Modeling. Mathematics, 8(9), 1508.
68. Mansour, M. M., Butt, N. S., Ansari, S. I., Yousof, H. M., Ali, M. M., & Ibrahim, M. (2020). A new exponentiated Weibull distribution's extension: copula, mathematical properties and applications. Contributions to Mathematics, 1, 57–66. DOI: 10.47443/cm.2020.0018.
69. Mansour, M. M., Butt, N. S., Yousof, H. M., Ansari, S. I., & Ibrahim, M. (2020). A Generalization of Reciprocal Exponential Model: Clayton Copula, Statistical Properties and Modeling Skewed and Symmetric Real Data Sets. Pakistan Journal of Statistics and Operation Research, 16(2), 373–386. https://doi.org/10.18187/pjsor.v16i2.3069
70. Mansour, M., Rasekhi, M., Ibrahim, M., Aidi, K., Yousof, H. M., & Elrazik, E. A. (2020). A New Parametric Life Distribution with Modified Bagdonavičius–Nikulin Goodness-of-Fit Test for Censored Validation, Properties, Applications, and Different Estimation Methods. Entropy, 22(5), 592.
71. Mohamed, H. S., Cordeiro, G. M., Minkah, R., Yousof, H. M., & Ibrahim, M. (2024). A Size-of-Loss Model for the Negatively Skewed Insurance Claims Data: Applications, Risk Analysis Using Different Methods and Statistical Forecasting. Journal of Applied Statistics, 51(2), 348–369. https://doi.org/10.1080/02664763.2023.2240980
72. Murthy, D.P., Xie, M., & Jiang, R. (2004). Weibull Models. John Wiley & Sons: Hoboken, NJ, USA.
73. Nikulin, M. S. (1973). Chi-squared test for normality. In Proceedings of the International Vilnius Conference on Probability Theory and Mathematical Statistics (Vol. 2, pp. 119–122).
74. Nikulin, M. S. (1973). Chi-squared test for continuous distributions with shift and scale parameters. Theory of Probability and Its Applications, 18, 559–568.
75. Nikulin, M. S. (1973). On a chi-squared test for continuous distributions. Theory of Probability and Its Applications, 19, 638–639.
76. Ramaki, Z., Alizadeh, M., Tahmasebi, S., Afshari, M., Contreras-Reyes, J. E., & Yousof, H. M. (2025). The Weighted Flexible Weibull Model: Properties, Applications, and Analysis for Extreme Events. Mathematical and Computational Applications, 30(2), 42.
77. Rao, K. C., & Robson, D. S. (1974). A chi-square statistic for goodness-of-fit tests within the exponential family. Communications in Statistics, 3, 1139–1153.
78. Rasekhi, M., Altun, E., Alizadeh, M., & Yousof, H. M. (2022). The Odd Log-Logistic Weibull-G Family of Distributions with Regression and Financial Risk Models. Journal of the Operations Research Society of China, 10(1), 133–158.
79. Rasekhi, M., Saber, M. M., & Yousof, H. M. (2020). Bayesian and Classical Inference of Reliability in Multicomponent Stress-Strength under the Generalized Logistic Model. Communications in Statistics- Theory and Methods, 50(21), 5114–5125. https://doi.org/10.1080/03610926.2020.1750651
80. Ravi, V., & Gilbert, P. D. (2009). BB: An R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function. Journal of Statistical Software, 32, 1–26.
81. Salah, M. M., El-Morshedy, M., Eliwa, M. S., & Yousof, H. M. (2020). Expanded Fréchet Model: Mathematical Properties, Copula, Different Estimation Methods, Applications and Validation Testing. Mathematics, 8(11), 1949. https://doi.org/10.3390/math8111949
82. Salem, M., Emam, W., Tashkandy, Y., Ibrahim, M., Ali, M. M., Goual, H., & Yousof, H. M. (2023). A new lomax extension: Properties, risk analysis, censored and complete goodness-of-fit validation testing under left-skewed insurance, reliability and medical data. Symmetry, 15(7), 1356.

83. Salih, A., Husien, W., & Abdulah, M. (2025). High Order Statistics From Lambert-Topp-Leone Distribution: Statistical Properties and Applications. Statistics, Optimization & Information Computing, 14(3), 1584–1597. https://doi.org/10.19139/soic-2310-5070-2914.
84. Shehata, W. A. M., Goual, H., Hamida, T., Hiba, A., Hamedani, G., Al-Nefaie, A. H., Ibrahim, M., Butt, N. S., Osman, R. M. A., & Yousof, H. M. (2024). Censored and Uncensored Nikulin-Rao-Robson Distributional Validation: Characterizations, Classical and Bayesian estimation with Censored and Uncensored Applications. Pakistan Journal of Statistics and Operation Research, 20(1), 11–35.
85. Sulewski, P., Alizadeh, M., Das, J., Hamedani, G. G., Hazarika, P. J., Contreras-Reyes, J. E., & Yousof, H. M. (2025). A New Logistic Distribution and Its Properties, Applications and PORT-VaR Analysis for Extreme Financial Claims. Mathematical and Computational Applications, 30(3), 62.
86. Teghri, S., Goual, H., Loubna, H., Butt, N. S., Khedr, A. M., Yousof, H. M., ... & Salem, M. (2024). A New Two-Parameters Lindley-Frailty Model: Censored and Uncensored Schemes under Different Baseline Models: Applications, Assessments, Censored and Uncensored Validation Testing. Pakistan Journal of Statistics and Operation Research, 109–138.
87. Yadav, A. S., Goual, H., Alotaibi, R. M., Ali, M. M., & Yousof, H. M. (2020). Validation of the Topp-Leone-Lomax model via a modified Nikulin–Rao–Robson goodness-of-fit test with different methods of estimation. Symmetry, 12(1), 57.
88. Yadav, A. S., Shukla, S., Goual, H., Saha, M., & Yousof, H. M. (2022). Validation of xgamma exponential model via Nikulin–Rao–Robson goodness-of-fit test under complete and censored sample with different methods of estimation. Statistics, Optimization & Information Computing, 10(2), 457–483.
89. Yousof, H. M., Aidi, K., Hamedani, G. G., & Ibrahim, M. (2021). A new parametric lifetime distribution with modified Chi-square type test for right censored validation, characterizations and different estimation methods. Pakistan Journal of Statistics and Operation Research, 17(2), 399–425.
90. Yousof, H. M., Ali, E. I. A., Aidi, K., Butt, N. S., Saber, M. M., Al-Nefaie, A. H., Aljadani, A., Mansour, M. M., Hamed, M. S., & Ibrahim, M. (2025). The Statistical Distributional Validation under a Novel Generalized Gamma Distribution with Value-at-Risk Analysis for the Historical Claims, Censored and Uncensored Real-life Applications. Pakistan Journal of Statistics and Operation Research, 21(1), 51-69. https://doi.org/10.18187/pjsor.v21i1.4534
91. Yousof, H. M., Ali, M. M., Aidi, K., & Ibrahim, M. (2023). The modified Bagdonavičius-Nikulin goodness-of-fit test statistic for the right censored distributional validation with applications in medicine and reliability. Statistics in Transition New Series, 24(4), 1–18.
92. Yousof, H. M., Ali, M. M., Goual, H., & Ibrahim, M. (2021). A new reciprocal Rayleigh extension: properties, copulas, different methods of estimation and modified right censored test for validation. Statistics in Transition New Series, 23(3), 1–23.
93. Yousof, H. M., Ali, M. M., Hamedani, G. G., Aidi, K., & Ibrahim, M. (2022). A new lifetime distribution with properties, characterizations, validation testing, different estimation methods. Statistics, Optimization & Information Computing, 10(2), 519–547.
94. Yousof, H. M., Aljadani, A., Mansour, M. M., & Abd Elrazik, E. M. (2024). A New Pareto Model: Risk Application, Reliability MOOP and PORT Value-at-Risk Analysis. Pakistan Journal of Statistics and Operation Research, 20(3), 383–407. https://doi.org/10.18187/pjsor.v20i3.4151.
95. Yousof, H. M., Altun, E., Ramires, T. G., Alizadeh, M., & Rasekhi, M. (2018). A new family of distributions with properties, regression models and applications. Journal of Statistics and Management Systems, 21(1), 163–188.
96. Yousof, H. M., Altun, E., Rasekhi, M., Alizadeh, M., Hamedani, G. G., & Ali, M. M. (2019). A New Lifetime Model with Regression Models, Characterizations, and Applications. Communications in Statistics- Simulation and Computation, 48(1), 264–286. https://doi.org/10.1080/03610918.2017.1367801
97. Yousof, H. M., Ansari, S. I., Tashkandy, Y., Emam, W., Ali, M. M., Ibrahim, M., & Alkhayyat, S. L. (2023). Risk Analysis and Estimation of a Bimodal Heavy-Tailed Burr XII Model in Insurance Data: Exploring Multiple Methods and Applications. Mathematics, 11(9), 2179. https://doi.org/10.3390/math11092179.
98. Yousof, H. M., Goual, H., Emam, W., Tashkandy, Y., Alizadeh, M., Ali, M. M., & Ibrahim, M. (2023). An Alternative Model for Describing the Reliability Data: Applications, Assessment, and Goodness-of-Fit Validation Testing. Mathematics, 11(6), 1308.
99. Yousof, H. M., Goual, H., Hamida, T., Hiba, A., Hamedani, G.G., & Ibrahim, M. (2022). Censored and Uncensored Nikulin-Rao-Robson Distributional Validation: Characterizations, Classical and Bayesian estimation with Applications.
100. Yousof, H. M., Goual, H., Khaoula, M. K., Hamedani, G. G., Al-Aefaie, A. H., Ibrahim, M., ... & Salem, M. (2023). A novel accelerated failure time model: Characterizations, validation testing, different estimation methods and applications in engineering and medicine. Pakistan Journal of Statistics and Operation Research, 19(4), 691–717.
101. Yousof, H. M., Korkmaz, M. Ç., Hamedani, G. G., & Ibrahim, M. (2022). A novel Chen extension: theory, characterizations and different estimation methods. Eur. J. Stat, 2(2022), 1-20.
102. Yousof, H., Afshari, M., Alizadeh, M., Ranjbar, V., Minkah, R., Hamed, M. S., & Salem, M. (2025). A Novel Insurance Claims (Revenues) Xgamma Extension: Distributional Risk Analysis Utilizing Left-Skewed Insurance Claims and Right-Skewed Reinsurance Revenues Data with Financial PORT-VaR Analysis. Pakistan Journal of Statistics and Operation Research, 83-117.
103. Yousof, H.M., Emam, W., Tashkandy, Y., Ali, M.M., Minkah, R., & Ibrahim, M. (2023). A Novel Model for Quantitative Risk Assessment under Claim-Size Data with Bimodal and Symmetric Data Modeling. Mathematics, 11, 1284. https://doi.org/10.3390/math11061284.
104. Yousof, H. M., Hamedani, G. G., Ali, M. M., Tork, M., Abdelrahaman, M. A., Alzahrani, M. R., ... & Muse, A. H. (2026). A New Extended Fréchet Distribution: Properties, Characterizations, Modeling, and Risk Analysis for Medical Data. Modern Journal of Statistics, 2(2), 57-80.
105. Zamani, Z., Afshari, M., Karamikabir, H., Alizadeh, M., & Ali, M. M. (2022). Extended Exponentiated Chen Distribution: Mathematical Properties and Applications. Statistics, Optimization & Information Computing, 10(2), 606–626.