STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, July 2025, pp 162–182.
Published online in International Academic Press (www.IAPress.org)

A New Mixed Gamma-Exponential Frailty Model under Heterogeneity Problem with Validation Testing for Emergency Care Data

Hafida Goual [1,*], Loubna Hamami [1], Mohamed S. Hamed [2,3]

[1] *Laboratory of probability and statistics LaPS, Badji Mokhtar-Annaba University, 12, P.O.Box, 23000 Annaba, Algeria*
[2] *Department of Business Administration, Gulf Colleges, KSA*
[3] *Department of Statistics, Mathematics and Insurance, Faculty of Commerce, Benha University, Egypt*

Abstract Frailty models play a crucial role in survival analysis as they account for unobserved differences among individuals, which may arise from various factors like genetics, environment, or lifestyle. These models help in identifying such factors and assessing their influence on survival outcomes. In this research, we introduce a new frailty model called the Mixed Gamma-Exponential (MxGEF) model for survival analysis. To evaluate its appropriateness, we apply the Rao-Robson-Nikulin (RR-Ni) and and the Bagdonaviμcius and Nikulin (B-Ni) goodness-of-fit tests, analyzing the distribution's characteristics and comparing its effectiveness against commonly used distributions in frailty modeling. Through simulation studies and real-world data applications, including a dataset collected from an emergency hospital in Algeria, we demonstrate how the MxGEF model effectively captures heterogeneity and improves model fitting. Our findings suggest that the MxGEF model is a promising alternative to existing frailty models, potentially enhancing the accuracy of survival analyses across various fields, including emergency care. Additionally, we explore the applicability of the MxGEF model in insurance through simulations and real data analysis, showcasing its versatility and potential impact in this domain.

Keywords Frailty model, Heterogeneity, Laplace transformation, Maximum likelihood, Regression Models, Statistical Testing.

AMS 2010 subject classifications 62N01, 62N02, 62E10

DOI: 10.19139/soic-2310-5070-2451

1. Introduction

Survival analysis is a key statistical methodology for modeling time-to-event data, with applications in medicine, engineering, economics, and insurance. A major challenge is accounting for unobserved heterogeneity among individuals, which can bias survival probability and hazard rate estimates (see Aalen (1988); Vaupel et al. (1979)). Frailty models address this by incorporating random effects ("frailties") to account for variability not explained by observed covariates. Traditional frailty models often rely on specific distributions like gamma or inverse Gaussian, which may not capture real-world complexity (see Wienke (2010) and Aalen (1988, 1992)). Alternatives include the generalized gamma frailty model (Balakrishnan and Peng (2006)) and the weighted Lindley frailty model (Mota et al. (2021)). This paper introduces the MxGEF model, combining gamma and exponential distributions for enhanced flexibility. Its performance is validated using the RR-Ni test (see Rao and Robson (1974) and Bagdonavičius and Nikulin (2011)). Goual et al. (2019) developed a modified goodness-of-fit test for censored and complete data, while Ibrahim et al. (2019) extended the Lindley distribution with a modified validation test, and Yadav et al. (2020) applied the RR-Ni test to the Topp-Leone-Lomax model. Abouelmagd et al. (2019)

*Correspondence to: Hafida Goual (Email: Goual.hafida@gmail.com). Laboratory of probability and statistics LaPS, Badji Mokhtar-Annaba University, 12, P.O.Box, 23000 Annaba, Algeria.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

proposed the zero-truncated Poisson Burr X family, and Mansour et al. (2020a) introduced a log-logistic lifetime model with copulas. Extensions of existing distributions, such as the exponentiated Weibull model (Mansour et al., 2020b) and reciprocal exponential model (Mansour et al., 2020c), enhance applicability. Bayesian and non-Bayesian approaches have improved validation techniques. Yousof et al. (2021) compared estimation methods for a new lifetime distribution, and Ibrahim et al. (2022a) explored censored and uncensored validation for the double Burr XII model. Recent studies focus on refining these techniques. Emam et al. (2023) addressed right-censored data, and Yousof et al. (2023a) introduced an alternative reliability model. Loubna et al. (2024) applied the quasi-xgamma frailty model to emergency care data, and Teghri et al. (2024) extended the Lindley-frailty model. Shehata et al. (2024) validated distributions using the RR-Ni test, and Hamedani et al. (2023) introduced a right-skewed one-parameter distribution for actuarial risk analysis.

Frailty models are essential in survival analysis for addressing unobserved heterogeneity caused by factors like genetics, environment, or lifestyle. These models enhance prediction accuracy by incorporating random effects beyond observed covariates (Loubna et al., 2024). In emergency care, where outcomes are influenced by diverse and unmeasured factors, frailty models provide a robust framework for analyzing survival patterns, assessing risks, and guiding decisions (Teghri et al. (2024)). The quasi-xgamma frailty model, introduced by Loubna et al. (2024), uses the RR-Ni test to validate its performance under censored and uncensored schemes, effectively capturing heterogeneity and improving model fit. Risk analysis is critical in emergency care, where accurate assessments impact patient outcomes. Salem et al. (2023) emphasize the importance of goodness-of-fit tests for validating distributions, while Hamedani et al. (2023) propose a right-skewed distribution for actuarial risk analysis. Practical applications include the reciprocal Weibull extension by Yousof et al. (2023) for modeling extreme values and the extended Gompertz model by Alizadeh et al. (2024) for reliability studies. Bayesian and non-Bayesian methodologies enhance risk quantification, as demonstrated by Ibrahim et al. (2023), Khedr et al. (2023) and Teghri et al. (2024). Frailty models thus offer powerful tools for analyzing heterogeneous data, with applications in healthcare, finance, and insurance. Future work should focus on refining these models for complex datasets and improving risk assessment methodologies.

Recently, Loubna et al. (2024) introduces the quasi-xgamma frailty model as a novel approach to survival analysis, particularly addressing heterogeneity problems in emergency care data. Following Loubna et al. (2024), this study presents a novel statistical model, referred to as the MxGEF model, which represents a significant advancement in the field of data analysis. The MxGEF model is designed to be versatile and adaptable, making it suitable for application to both simulated datasets and real-world data. Notably, it has been applied to emergency care data from Algeria, demonstrating its practical utility in analyzing complex, real-life scenarios. The MxGEF model exhibits superior performance compared to existing approaches in terms of both fit and interpretability. This claim is substantiated through rigorous validation using two robust statistical tests, the RR-Ni and the B-Ni. These tests confirm that the MxGEF model provides a more accurate and reliable fit to the data while maintaining clarity in interpretation, which is crucial for practical applications. Moreover, the model shows considerable potential in the domain of insurance risk analysis. By accurately modeling various types of risks, the MxGEF model can assist insurance companies in better assessing and managing uncertainties associated with their portfolios. This capability underscores the model's versatility and relevance across multiple fields, further highlighting its value as a cutting-edge tool for data-driven decision-making. Sections 2–6 cover the Cox-frailty model, estimation process, validation results, data analysis, and future directions. This paper introduces the MxGEF model , motivated by three key imperatives:

- The MxGEF model merges the strengths of gamma and exponential distributions, offering a novel parametric framework that balances mathematical tractability with the ability to model diverse hazard shapes (e.g., bathtub, monotonically increasing/decreasing). This hybrid structure addresses limitations of existing frailty models (e.g., quasi-xgamma, Lindley-frailty) that often lack closed-form expressions or sufficient adaptability for real data.
- The MxGEF model is rigorously validated using the RR-Ni and B-Ni goodness-of-fit tests, ensuring robustness under censored and uncensored data scenarios. Its application to a real-world emergency care

dataset from Algeria demonstrates superior performance in capturing heterogeneity and improving model fit compared to traditional approaches. This validation not only underscores its practical utility but also aligns with regulatory and ethical demands for transparent, reproducible risk assessment in high-stakes domains like healthcare and insurance.

- Beyond emergency care, the MxGEF model's modular design enables seamless integration into insurance risk analysis, actuarial science, and emerging fields like climate modeling. For example, its ability to handle extreme value scenarios (e.g., catastrophic insurance claims) and incorporate covariate effects positions it as a versatile tool for dynamic risk quantification. Furthermore, the model's compatibility with Bayesian frameworks and computational efficiency (via closed-form Laplace transforms) addresses critical gaps in handling high-dimensional or sparse datasets.

2. The Cox-frailty model

Consider the Cox proportional hazard (Cox-PH) model (see Cox (1972)) and an unexplained source of heterogeneity. The univariate frailty model's goal is to simulate unobserved factors in failure rates for unrelated items, demonstrating that the frailties are independent. Let $Z > 0$ and for an unobserved random variable that represents the frailty of the object. Then, the hazard-rate function for the i^{th} item is

$$\lambda\left(t_i|z_i, x_i\right) = z_i\lambda_0(t_i)\exp(x_i^T\underline{\beta}), \;\; i = 1, 2, ..., n, \tag{1}$$

where $\lambda_0(\cdot)$ refers to the hazard-rate function of the baseline model, $\underline{\beta} = \underline{\beta}_{(p\times 1)}$ is the vector of unknown regression coefficients for all $p < n$ (see Ibrahim et al. (2001)), where subject i has a unique frailty z_i, which is an unobserved non-negative number. Hence, if $z_i > 1$ or $z_i < 1$, respectively, frailty z_i raises or reduces the chance of occurrence of the event of our interest. The Cox-PH model is produced as a specific instance where $z_i = 1$ for every i. The following is how the model in (1) is used to determine the conditional survival function for the i^{th} subject:

$$S(t_i|z_i, x_i) = \exp(-z_i\Lambda_0(t_i)\exp(x_i^T\underline{\beta}))\,, i = 1, ..., n, \tag{2}$$

where the cumulative baseline hazard-rate function is

$$\Lambda_0(t_i) = \int_0^{t_i}\lambda_0(s)ds.$$

The conditional survival function (2) thus indicates the likelihood that the ith subject will live until time t_i given $Z = z_i$. We must integrate out the conditional survival function (2) on frailty in order to obtain the marginal survival (Mar-S) function, which does not depend on unseen variables. Keep in mind that this is equal to computing the frailty distribution's Laplace transform. In reality, if $f(z)$ is the frailty distribution, then we may get the following by integrating $S(t_i \mid z_i, x_i)$ from (2) on $Z = z_i$:

$$S(t_i|x_i) = \int_0^{\infty}\exp(-z_i\Lambda_0(t_i)\exp(x_i^T\underline{\beta}))f(z_i)dz_i = L_f(\Lambda_0(t_i)\exp(x_i^T\underline{\beta})), \tag{3}$$

where $L_f(\cdot)$ stands for the frailty distribution's Laplace transform, and the appropriate marginal probability density function (MPDF) is

$$f(t_i|x_i) = -\lambda_0(t_i)\exp(x_i^T\underline{\beta})L_f^{'}\left[\Lambda_0(t_i)\exp(x_i^T\underline{\beta})\right], i = 1, ..., n.$$

Take into account that the Laplace transform has a closed form for each of the distributions above. As a result, (3) may be used to derive the marginal hazard (Mar-H) function as follows:

$$\lambda(t_i|x_i) = -\frac{\lambda_0(t_i)\exp(x_i^T\underline{\beta})L_f^{'}\left[\Lambda_0(t_i)\exp(x_i^T\underline{\beta})\right]}{L_f(\Lambda_0(t_i)\exp(x_i^T\underline{\beta}))}, i = 1, 2, ..., n. \tag{4}$$

where $L_f^{'}(t) = \frac{\partial}{\partial t} L_f(t)$. The risk and survival of a person randomly selected from the research population are therefore calculated using the hazard and Mar-S functions (illustrated above) (see Wienke (2010)). As indicated earlier, using a frailty distribution with a Laplace transform on the closed-form facilitates parameter estimation and is necessary for estimating the Mar-S and hazard functions. However, when the frailty distribution lacks a Laplace transform on the closed-form, numerical integration or Markov Chain Monte Carlo approaches must be applied (see Balakrishnan and Peng (2006), Hougaard (2012), Robert and Casella (2013)). Frailty distribution must be taken into account while accounting for computational ease in univariate and multivariate survival data modelling (see Pickles and Crouchley (1995), Wienke (2010)). Hazard and Mar-S functions were created in this study.

2.1. The MxGEF model

Following Sen and Chandra (2017), the probability density function (PDF) of the NTPGEM model can be expressed as

$$f_{\underline{\mathbf{P}}}(x) = \frac{K^2}{1+K}\left[1 + \frac{K^{s-2}}{\Gamma(s)}x^{s-2}\right]\exp(-Kx); x \geq 0,$$

where $\underline{\mathbf{P}} = (s,k)., k > 0, s > 1$. Consider the frailty model in (1) where the frailty variable Z has a NTPGEM distribution (5) with mean one, i.e., $E[Z] = 1$. This assumption is required to identify the parameters of the derived model (see Elbers and Ridder (1982)). Hence, employing Mazucheli et al. (2016)'s alternate parameterization of the NTPGEM model in terms of mean, the PDF of the MxGEF model becomes

$$f(Z) = \frac{S}{1+\sqrt{S}}\left[1 + \frac{(\sqrt{S})^{S-2}}{\Gamma(S)}z^{S-1}\right]\exp(-\sqrt{S}z), \tag{5}$$

where $s > 0$ represents the (unknown) shape parameter. The variance of the frailty distribution is commonly used to quantify the level of unobserved variation in a research sample. If the PDF (5) is assumed to be a frailty distribution, the variance is

$$\sigma^2 = \frac{1}{(1+\sqrt{s})^2}\left(s^{\frac{3}{2}} - s^{\frac{1}{2}} + 1 + 2s^2\right).$$

The frailty PDF (5)'s Laplace transform, depending on its variance and $X \in \mathbb{R}$, is given by:

$$\begin{aligned} L_f(X) &= \frac{s}{(1+\sqrt{s})(X+\sqrt{s})} + \frac{s(\sqrt{s})^{s-2}}{(1+\sqrt{s})(X+\sqrt{s})^s} \\ &= \frac{s}{(1+\sqrt{s})}\left[\frac{1}{(X+\sqrt{s})} + \frac{(\sqrt{s})^{s-2}}{(X+\sqrt{s})^s}\right], \end{aligned} \tag{6}$$

where

$$\sqrt{S} = \frac{1+\sigma^2+\sqrt{\sigma^4+6\sigma^2-7}}{2},$$

and $\sigma^2 \geq 1$. In order to maintain simplicity, we analyze (6) at $X = \Lambda_0(t_i)\xi_i$, where $\xi_i = \exp(x_i^T\beta)$, and find that the marginal survival function (3) under MxGEF model is provided by

$$S(t_i|x_i) = \frac{s}{(1+\sqrt{s})}\left[\frac{1}{(X+\sqrt{s})} + \frac{(\sqrt{s})^{s-2}}{(X+\sqrt{s})^s}\right]. \tag{7}$$

The resulting marginal hazard function (4) is as follows:

$$\lambda(t_i|x_i) = \frac{\left[2^{\frac{(M(\sigma))^2}{4}} + (M(\sigma))^{\frac{(M(\sigma))^2}{4}}\right]}{2\left[\begin{array}{l} 2^{\frac{(M(\sigma))^2}{4}-2}\left(H_0(t_i)\exp(x_i^T\beta) + \frac{M(\sigma)}{2}\right)^{\left(\frac{M(\sigma)}{2}\right)^2} \\ + \left(H_0(t_i)\exp(x_i^T\beta) + \frac{M(\sigma)}{2}\right)(M(\sigma))^{\frac{(M(\sigma))^2}{4}-2} \end{array}\right]}, \tag{8}$$

where

$$M(\sigma) = \left(1 + \sigma^2 + \sqrt{\sigma^4 + 6\sigma^2 - 7}\right).$$

2.2. The MxGEF model under the Weibull baseline hazard function

The Weibull distribution's baseline hazard and cumulative hazard functions are provided by:

$$\lambda_0(t_i) = \frac{K}{\rho}\left(\frac{t_i}{\rho}\right)^{K-1} \quad \text{and} \quad \Lambda_0(t_i) = \left(\frac{t_i}{\rho}\right)^{K} \qquad t_i > 0 \tag{9}$$

where $K > 0$ represents the shape parameter and $\rho > 0$ represents the scale parameter. The hazard function of the Weibull distribution drops monotonously for $K < 1$; it is constant over time for $K = 1$ (exponential distribution); and it grows monoton when $K > 1$ (Wienke (2010)). As a result of plugging (10) into (9) and (8), the MxGEF model's marginal survival and hazard functions with Weibull baseline hazard function are, respectively,

$$\begin{aligned} S(t_i|x_i) &= \frac{M(\sigma)^2}{2(3+\sigma^2+\sqrt{\sigma^4+6\sigma^2-7})}\Phi\left(t_i;x_i,\sigma\right)^{-1} \\ &\quad + M(\sigma)^2\left[2(3+\sigma^2+\sqrt{\sigma^4+6\sigma^2-7})\right]^{-1}\left[M(\sigma)^{\left(\frac{M(\sigma)}{2}\right)^2-1}\right] \\ &\quad \times\left\{2^{\frac{M(\sigma)^2-8}{4}}\Phi\left(t_i;x_i,\sigma\right)^{\left(\frac{M(\sigma)}{2}\right)^2}\right\}^{-1}, \end{aligned} \tag{10}$$

where

$$\varrho\left(t_i;x_i\right) = \left(\frac{t_i}{\rho}\right)^{K}\exp(x_i^T\beta) \tag{11}$$

and

$$\begin{aligned} \lambda(t_i|x_i) &= \left[\frac{K}{\rho}\varrho\left(t_i;x_i\right)\right] \\ &\quad \times\begin{pmatrix} 2^{\frac{M(\sigma)^2-8}{4}}\left\{\Phi\left(t_i;x_i,\sigma\right)^{\left(\frac{M(\sigma)}{2}\right)^2}\right\}^2 \\ +\Phi\left(t_i;x_i,\sigma\right)^2\left(M(\sigma)\right)^{\frac{M(\sigma)^2-8}{4}}\left(\frac{M(\sigma)}{2}\right)^2\Phi\left(t_i;x_i,\sigma\right)^{\left(\frac{M(\sigma)}{2}\right)^2-1} \end{pmatrix} \\ &\quad \times\begin{pmatrix} \Phi\left(t_i;x_i,\sigma\right)\Phi\left(t_i;x_i,\sigma\right)^{\left(\frac{M(\sigma)}{2}\right)^2} \\ \times\left\{\begin{array}{c} 2^{\frac{M(\sigma)^2-8}{4}}\Phi\left(t_i;x_i,\sigma\right)^{\left(\frac{M(\sigma)}{2}\right)^2} \\ +\Phi\left(t_i;x_i,\sigma\right)M(\sigma)^{\frac{M(\sigma)^2-8}{4}} \end{array}\right\} \end{pmatrix}^{-1}. \end{aligned} \tag{12}$$

where $\Phi\left(t_i;x_i,\sigma\right) = \left[\varrho\left(t_i;x_i\right) + \frac{M(\sigma)}{2}\right]$.

2.3. The MxGEF model under the Gompertz baseline hazard function

The Gompertz distribution's baseline hazard and cumulative hazard functions are provided by:

$$\lambda_0(t_i) = \rho_1\exp(K_1t_i) \quad \text{and} \quad \Lambda_0(t_i) = \frac{\rho_1}{K_1}(\exp(K_1t_i)-1) \quad , t_i > 0 \tag{13}$$

where $K_1 > 0$ and $\rho_1 > 0$ are the shape and scale parameters. If $K_1 < 0$, the Gompertz distribution is flawed (Calsavara et al. 2019a, b), since its cumulative hazard function converges to the constant $-\frac{\rho_1}{K_1}$ for $t \to \infty$, resulting in a cure or long-term survivors fraction $p_0 = exp(\frac{\rho_1}{K_1})$ in the research population. The exponential

distribution is derived as a special case for $K = 0$. As a result, the Gompertz distribution's hazard function might be decreasing ($K_1 < 0$), constant ($K_1 = 0$), or increasing ($K_1 > 0$). The marginal survival and hazard functions of the MxGEF model with Gompertz baseline hazard function are calculated by inserting (12) into (9) and (8),

$$S(t_i|x_i) = \frac{M(\sigma)^2}{2(3+\sigma^2+\sqrt{\sigma^4+6\sigma^2-7})} \left[\begin{array}{c} [T(t_i;x_i,\sigma)]^{-1} \\ + \left(\left[M^{-1}(\sigma)^{\left(\frac{M(\sigma)}{2}\right)^2-1} \right] \left\{ 2^{\frac{M(\sigma)^2-8}{4}} T(t_i;x_i,\sigma)^{\left(\frac{M(\sigma)}{2}\right)^2} \right\}^{-1} \right) \end{array} \right],$$

and

$$\lambda(t_i \mid x_i) = \left[\rho_1 \exp(K_1 t_i) \exp(x_i^T \beta)\right] \times \left[\begin{array}{c} \left[\begin{array}{c} 2^{\frac{M(\sigma)^2-8}{4}} \left\{ \frac{\rho_1}{K_1} \left(\left[\exp(K_1 t_i) - 1 \right] \exp(x_i^T \beta) + \frac{M(\sigma)}{2} \right]^{\left(\frac{M(\sigma)}{2}\right)^2} \right\}^2 + [T(t_i;x_i,\sigma)]^2 \\ \times (M(\sigma))^{\frac{M(\sigma)^2-8}{4}} \left(\frac{M(\sigma)}{2}\right)^2 [T(t_i;x_i,\sigma)]^{\left(\frac{M(\sigma)}{2}\right)^2-1} \end{array} \right] \\ \hline T(t_i;x_i,\sigma)\, T(t_i;x_i,\sigma)^{\left(\frac{M(\sigma)}{2}\right)^2} \\ \times \left[2^{\frac{M(\sigma)^2-8}{4}} [T(t_i;x_i,\sigma)]^{\left(\frac{M(\sigma)}{2}\right)^2} + T(t_i;x_i,\sigma)\, M(\sigma)^{\frac{M(\sigma)^2-8}{4}} \right] \end{array} \right],$$

where $T(t_i;x_i,\sigma) = \tau(t_i)\exp(x_i^T\beta) + \frac{M(\sigma)}{2}$, $\tau(t_i) = \frac{\rho_1}{K_1}[\exp(K_1 t_i) - 1]$. As previously stated, the marginal survival function is appropriate for $K > 0$ and inappropriate for $K < 0$. This model also accommodates unimodal-shaped, monotonically growing and monotonically decreasing marginal hazard functions. The MxGEF model with the Gompertz baseline hazard function is therefore more versatile than the MxGEF model with the Weibull baseline hazard function.

3. Estimation of MxGEF model's parameters

In an uncensored simulation study under the RR-Ni statistics, data are generated from a known distribution and then tested against a hypothesized distribution using one or more of the RR-Ni statistics. The performance of the statistics is evaluated based on their ability to correctly identify the underlying distribution, as well as their sensitivity to sample size, parameter values, and other factors. There are several motivations for conducting an uncensored simulation study under the RR-Ni statistics. One important motivation is to assess the statistical power of the RR-Ni tests under different scenarios. Statistical power is a measure of the ability of a test to detect a true effect or difference, and is influenced by factors such as sample size and effect size. By conducting a simulation study, researchers can determine the minimum sample size required to achieve a desired level of statistical power, and can assess the impact of other factors on test performance.

On the other hand, one important motivation for conducting a censored simulation study under the BG-NI statistics is to assess the statistical power of the tests under different types and levels of censoring. Censoring can lead to loss of information and reduced statistical power, so it is important to determine the minimum sample size required to achieve a desired level of statistical power under different types and levels of censoring. Another motivation for conducting a censored simulation study under the BG-NI statistics is to evaluate the accuracy and precision of the estimated distribution parameters, particularly when dealing with right-censored data. In many

cases, the goal of a goodness-of-fit test is not only to determine whether a particular distribution fits the data, but also to estimate the values of its parameters. Simulation studies can provide insights into the accuracy and precision of parameter estimates under different types and levels of censoring, and can inform decisions about which distribution to use for subsequent analyses.

3.1. Case of Weibull baseline hazard function

The ML approach for estimating parameters of the MxGEF model with Weibull and Gompertz baseline hazard functions is described in this section. ML estimators have appealing qualities under specific regularity constraints, such as consistency, efficiency, asymptotic normality, and others (Lehmann and Casella 2006). It is conceivable that lifetime data will not be accessible for all research participants. Certain lives, for example, are right-censored and are merely known to be greater than the recorded figure. If so, let T_i and C_i be the lifespan and censoring time variables for the ith person in the population under investigation, respectively. Assume T_i and C_i are independent random variables, and $\delta_i = \mathbf{1}_{(T_i \leq C_i)}$ is the censoring indicator (i.e., $\delta_i = 1$ if T_i is lifetime, and $\delta_i = 0$ otherwise). Then we see that $t_i = min\{T_i, C_i\}$. Let x_i represent a $p \times 1$ vector of variables observed in the i^{th} subject. The likelihood function for the model parameter vector $\underline{\mathbf{P}}$ under non-informative censoring is thus provided from a sample of n participants as $L(\underline{\mathbf{P}}) = \prod_{i=1}^{n} \lambda(t_i \mid x_i)^{\delta_i} S(t_i \mid x_i)$, where $S(\cdot \mid x_i)$ and $\lambda(\cdot \mid x_i)$ are the marginal survival and hazard functions given in Equations (8) and (9). As a result, the associated log-likelihood function is calculated using the natural logarithm of $L(\underline{\mathbf{P}})$. Then, the loglikelihood function for $\underline{\mathbf{P}} = (K, \rho, \sigma^2, \beta)$ is given by

$$\begin{aligned}
\log L(\underline{\mathbf{P}}) &= r\log(K) + (K-1)\sum_{i=1}^{n}\delta_i \log(t_i) - kr\log(\rho) + \sum_{i=1}^{n}\delta_i x_i^T\beta \\
&\quad - \sum_{i=1}^{n}\delta_i \log(U) - r\log(2) - \sum_{i=1}^{n}\delta_i \log(U_1) + 2n\log(M(\sigma)) - n\log(2) \\
&\quad - n\log(2 + M(\sigma)) + \sum_{i=1}^{n}\log(U_1) - \sum_{i=1}^{n}\log(U_2) \\
&\quad - n\left(\frac{M(\sigma)^2}{4} - 2\right)\log(2) - \frac{M(\sigma)^2}{4}\sum_{i=1}^{n}\log(U_2),
\end{aligned} \tag{14}$$

where: $r = \sum_{i=1}^{n}\delta_i$ is the failure number,

$$U = \left[2^{\frac{M(\sigma)^2}{4}} + (M(\sigma))^{\frac{M(\sigma)^2}{4}}\right],$$

$$\begin{aligned}
U_1 &= 2^{\frac{M(\sigma)^2}{4}-2}\Phi\left(t_i; x_i, \sigma\right)^{\left(\frac{M(\sigma)}{2}\right)^2} \\
&\quad + \Phi\left(t_i; x_i, \sigma\right) M(\sigma)^{\frac{\left(\sigma^2+\sqrt{\sigma^4+6\sigma^2-7}\right)^2-7}{4}},
\end{aligned}$$

and

$$U_2 = \varrho\left(t_i; x_i\right) + \frac{M(\sigma)}{2}.$$

Setting the nonlinear system of the score equations $\mathbf{I}_{(K)} = 0, \mathbf{I}_{(\rho)} = 0, \mathbf{I}_{(\sigma^2)} = 0$ and $\mathbf{I}_{(\beta)} = 0$ and solving them simultaneously yields the MLE $\widehat{\underline{\mathbf{P}}} = (\widehat{K}, \widehat{\rho}, \widehat{\sigma^2}, \widehat{\beta})^{\intercal}$. It is usually more convenient to use nonlinear optimization methods to solve these equations; such as the quasi-Newton algorithm to numerically maximize $\log L(\underline{\mathbf{P}})$.

3.2. Simulations: case of Weibull baseline hazard function

We consider the MxGEF model with Weibull baseline hazard function. The data were simulated $N = 13,000$ times; with parameter values $K = 0.7, \rho = 0.8, \sigma^2 = 0.7, \beta_1 = 0.9$ and sample sizes $n = 25, n = 40, n = 250$ and $n = 800$. Using the R software and the Barzilai-Borwein (BB) algorithm (Ravi, 2009) for calculating the averages of the simulated values of the maximum likelihood estimators $\widehat{K}, \widehat{\rho}, \widehat{\sigma^2}, \widehat{\beta_1}$ parameters and their mean squared errors (MSE). Table 1 presents the bias and mean squared error (MSE) of the maximum likelihood (ML) estimates for the parameters of the MxGEF model with a Weibull baseline hazard function, under varying sample sizes and censoring levels. The results provide valuable insights into the performance of the model under different conditions.

Firstly, as the sample size increases, both the bias and MSE decrease for all parameters, indicating improved estimation accuracy and efficiency. For instance, at the smallest sample size ($n = 25$), the bias and MSE values are relatively high, particularly for the parameter ρ (e.g., bias = 0.87548 and MSE = 0.04981 under 0% censoring). However, as the sample size grows to 800, the bias and MSE notably decrease (e.g., bias = 0.81003 and MSE = 0.03002 for ρ under 0% censoring). This trend is consistent across all parameters ($\widehat{K}, \widehat{\rho}, \widehat{\sigma^2}, \widehat{\beta_1}$), highlighting the importance of larger datasets for more reliable parameter estimation.

Secondly, the impact of censoring on estimation accuracy is evident from the table. As the percentage of censored observations increases, the bias and MSE generally increase, reflecting the challenges posed by higher levels of censoring. For example, at $n = 25$, the bias for ρ rises from 0.87548 (0% censoring) to 0.86000 (40% censoring), while the corresponding MSE increases from 0.04981 to 0.04587. Similarly, for $\widehat{\sigma^2}$, the MSE at $n = 40$ increases from 0.02347 (0% censoring) to 0.03641 (40% censoring). These findings underscore the need for robust statistical methods to handle censored data effectively.

Thirdly, the parameter $\widehat{\beta_1}$, which represents the effect of covariates in the model, shows a similar pattern of decreasing bias and MSE with increasing sample size and deteriorating performance with higher censoring levels. Notably, even at the largest sample size ($n = 800$), some degree of bias persists under high censoring (e.g., bias = 0.90210 and MSE = 0.01958 under 40% censoring), suggesting that censoring remains a significant challenge even with large datasets. Overall, the results demonstrate that the MxGEF model performs well under a variety of conditions, with improved accuracy and precision as sample size increases and censoring decreases. However, the increased bias and MSE under high censoring levels indicate the need for further research into methods that can mitigate the effects of censoring, particularly in small or moderately sized datasets. This analysis not only validates the robustness of the MxGEF model but also highlights areas where its application may require additional considerations. From Table 1, we observe that the maximum likelihood estimates for the MxGEF model with Weibull baseline hazard function are convergent, as evidenced by the decreasing trends in both bias and MSE with increasing sample sizes.

Specifically, the bias and MSE values for all parameters ($\widehat{K}, \widehat{\rho}, \widehat{\sigma^2}, \widehat{\beta_1}$) consistently decrease as the sample size grows from 25 to 800. This convergence indicates that the model's estimates become more accurate and efficient with larger datasets. Additionally, while censoring introduces some variability in the estimates, the overall trend of improvement with increasing sample size remains robust. These results suggest that the MxGEF model is well-suited for applications where large datasets are available, further validating its reliability and effectiveness in survival analysis. Thus, the convergence of the estimates underscores the model's statistical consistency and suitability for practical use.

Table 1: Bias and MSE of the ML estimates for the simulated data of the MxGEF model with Weibull baseline hazard function

n		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
		0%cens.		10%cens.		25%cens.		40%cens.	
25	ρ	0.87548	0.04981	0.86952	0.05002	0.85321	0.03996	0.86000	0.04587
	K	0.76958	0.04635	0.75316	0.04751	0.75558	0.04968	0.74875	0.04751
	σ^2	0.75481	0.04612	0.73625	0.04857	0.74986	0.04968	0.74266	0.04758
	β_1	0.96325	0.03999	0.94589	0.04657	0.97006	0.04587	0.95319	0.04578
40	ρ	0.86359	0.04003	0.85416	0.04787	0.849571	0.03225	0.85304	0.04120
	K	0.74317	0.04201	0.73167	0.03198	0.73410	0.03241	0.72348	0.03795
	σ^2	0.72954	0.02347	0.72015	0.03985	0.72968	0.04000	0.72406	0.03641
	β_1	0.95021	0.0300	0.93406	0.03958	0.96012	0.03698	0.91364	0.03121
250	ρ	0.82135	0.03845	0.81954	0.03652	0.82304	0.03164	0.82467	0.03775
	K	0.72214	0.02877	0.71845	0.02578	0.71864	0.02958	0.71005	0.02964
	σ^2	0.71067	0.01935	0.71384	0.02471	0.71357	0.02775	0.71111	0.02886
	β_1	0.92231	0.01247	0.92784	0.03012	0.9333	0.03124	0.91120	0.02657
800	ρ	0.81003	0.03002	0.80647	0.02425	0.81023	0.02996	0.81000	0.02746
	K	0.70085	0.02102	0.71130	0.02113	0.70064	0.01968	0.70301	0.02110
	σ^2	0.70200	0.01322	0.70604	0.02210	0.70651	0.01774	0.70009	0.02345
	β_1	0.9004	0.0215	0.9006	0.02335	0.90100	0.02005	0.90210	0.01958

3.3. Case of Gompertz baseline hazard function

Using the Gompertz baseline hazard function, the log-likelihood function for $\underline{\mathbf{P}} = (K_1, \rho_1, \sigma^2, \beta)$ is given as follows:

$$\begin{aligned}\log L(\underline{\mathbf{P}}) &= r\log(\rho_1) + \sum_{i=1}^{n}\delta_i(K_1 t_i + x_i^T\beta) + \sum_{i=1}^{n}\delta_i\log(U) - r\log(2)\\ &\quad - \sum_{i=1}^{n}\delta_i\log(U_1) + 2n\log(M(\sigma)) - n\log(2) - n\log(2+M(\sigma))\\ &\quad + \sum_{i=1}^{n}\log(U_1) - \sum_{i=1}^{n}\log(U_2) - n\left(\frac{M(\sigma)^2}{4} - 2\right)\log(2)\\ &\quad - \frac{M(\sigma)^2}{4}\sum_{i=1}^{n}\log(U_2).\end{aligned}$$

where

$$M(\sigma) = 1 + \sigma^2 + \sqrt{\sigma^4 + 6\sigma^2 - 7}$$

$$U = \left(2^{\frac{M(\sigma)^2}{4}} + (1M(\sigma))^{\frac{M(\sigma)^2-7}{4}}\right).$$

$$\begin{aligned}U_1 &= 2^{\frac{M(\sigma)^2-8}{4}}\left(\begin{array}{c}\frac{\rho}{K}(\exp(Kt_i)-1)\ \exp(x_i^T\beta)\\ +\frac{M(\sigma)}{2}\end{array}\right)^{\left(\frac{M(\sigma)}{2}\right)^2}\\ &\quad + \left(\begin{array}{c}\frac{\rho}{K}(\exp(Kt_i)-1)\ \exp(x_i^T\beta)\\ +\frac{M(\sigma)}{2}\end{array}\right)(M(\sigma))^{\frac{M(\sigma)^2-8}{4}}\end{aligned}$$

and

$$U_2 = \frac{\rho}{K}(\exp(Kt_i) - 1)\ \exp(x_i^T\beta) + \frac{M(\sigma)}{2}$$

Maximizing the log-likelihood functions (16) and (17), respectively, yields the appropriate ML estimators $\widehat{\underline{\mathbf{P}}}$ of parameter vectors $\underline{\mathbf{P}}$. It is worth noting that $\widehat{\underline{\mathbf{P}}}$ does not have a closed form. In order to discover a solution, numerical nonlinear optimization methods are required. These optimization approaches are implemented in BBsolve R software packages (see Ravi (2009)).

3.4. Simulations: case of Gompertz baseline hazard function

We consider the MxGEF model with Gompertz baseline hazard function. The data were simulated $N = 13,000$ times; with parameter values $K_1 = 0.9, \rho_1 = 0.6, \sigma^2 = 0.8, \beta_1 = 0.3$ and sample sizes $n = 25, n = 40, n = 250$ and $n = 800$. Using the R software and the Barzilai-Borwein (BB) algorithm (see Ravi, (2009)) for calculating the averages of the simulated values of the maximum likelihood estimators $\widehat{K_1}, \widehat{\rho_1}, \widehat{\sigma^2}, \widehat{\beta_1}$ parameters and their mean squared errors (MSE). Table 2 presents the bias and mean squared error (MSE) of the maximum likelihood estimates for the MxGEF model with a Gompertz baseline hazard function under varying sample sizes and censoring levels. The results reveal several key insights into the model's performance. First, as the sample size increases, both bias and MSE consistently decrease for all parameters ($\widehat{K_1}, \widehat{\rho_1}, \widehat{\sigma^2}, \widehat{\beta_1}$), indicating improved estimation accuracy and efficiency. For example, at the smallest sample size ($n = 25$), the bias for $\widehat{\rho_1}$ is 0.65247 (0% censoring), which reduces to 0.63025 at $n = 800$ under the same censoring level. Similarly, the MSE for $\widehat{\beta_1}$ decreases from 0.04751 at $n = 25$ to 0.03173 at $n = 800$. This trend highlights the importance of larger datasets in achieving more reliable parameter estimates. Second, the impact of censoring on estimation accuracy is evident, as higher levels of censoring generally lead to increased bias and MSE. For instance, at $n = 25$, the bias for $\widehat{\rho_1}$ rises from 0.65247 (0% censoring) to 0.67398 (40% censoring), while the corresponding MSE increases from 0.04652 to 0.04963. This pattern is consistent across all parameters, underscoring the challenges posed by higher censoring levels. Despite this, the model demonstrates robustness even under significant censoring, as the bias and MSE remain manageable with larger sample sizes. Finally, the parameter $\widehat{\beta_1}$, representing covariate effects, shows a similar trend of decreasing bias and MSE with increasing sample size, though it is slightly more sensitive to censoring compared to other parameters. For example, at $n = 800$, the MSE for $\widehat{\beta_1}$ increases from 0.02191 (10% censoring) to 0.01247 (40% censoring). Overall, these findings confirm the MxGEF model's ability to provide reliable estimates under various conditions, particularly when applied to sufficiently large datasets, while also highlighting the need for caution in scenarios with high censoring levels.

From Table 2, we observe that the maximum likelihood estimates for the MxGEF model with Gompertz baseline hazard function exhibit convergence as the sample size increases. The bias and mean squared error (MSE) values for all parameters $\widehat{K_1}, \widehat{\rho_1}, \widehat{\sigma^2}, \widehat{\beta_1}$ consistently decrease with larger sample sizes, indicating improved estimation accuracy and efficiency. For instance, at the smallest sample size ($n = 25$), the bias for $\widehat{\rho_1}$ is 0.65247 (0% censoring), which reduces to 0.63025 at $n = 800$ under the same censoring level. Similarly, the MSE for $\widehat{\beta_1}$ decreases from 0.04751 at $n = 25$ to 0.03173 at $n = 800$. Although censoring introduces some variability in the estimates, the overall trend of decreasing bias and MSE with increasing sample size remains robust. These findings confirm the model's ability to provide reliable parameter estimates when applied to sufficiently large datasets, thereby validating its suitability for survival analysis under the Gompertz baseline hazard function.

Table 2: Bias and MSE of the ML estimates for the simulated data of the MxGEF model with Gompertz baseline hazard function

n		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
		0% cens.		10% cens.		25% cens.		40% cens.	
25	ρ_1	0.65247	0.04652	0.64328	0.04325	0.65003	0.04996	0.67398	0.04963
	K_1	0.94487	0.04621	0.95003	0.04652	0.94638	0.04357	0.95555	0.04581
	σ^2	0.87284	0.04751	0.84867	0.04756	0.86485	0.04158	0.87499	0.04751
	β_1	0.35953	0.04751	0.35254	0.04395	0.36358	0.04351	0.34968	0.04225
40	ρ_1	0.64439	0.04112	0.62369	0.03951	0.64381	0.03514	0.66634	0.04360
	K_1	0.94005	0.04125	0.92004	0.03617	0.92968	0.03968	0.93957	0.03625
	σ^2	0.86357	0.04003	0.83681	0.03698	0.84987	0.03138	0.83498	0.03251
	β_1	0.31254	0.03952	0.34615	0.03467	0.33296	0.03336	0.33650	0.03625
250	ρ_1	0.64005	0.03847	0.61541	0.02958	0.62013	0.02987	0.62198	0.03214
	K_1	0.92365	0.03421	0.91584	0.02473	0.91958	0.02985	0.92671	0.02854
	σ^2	0.81584	0.03216	0.82574	0.02618	0.83185	0.02758	0.81958	0.02361
	β_1	0.31000	0.03674	0.33578	0.02356	0.31520	0.02458	0.32655	0.02124
800	ρ_1	0.63025	0.02514	0.60688	0.02120	0.60008	0.02310	0.60307	0.02001
	K_1	0.90457	0.02317	0.91000	0.02220	0.90374	0.02132	0.90007	0.01254
	σ^2	0.81240	0.02316	0.81396	0.0230	0.81124	0.02008	0.80064	0.02300
	β_1	0.30327	0.03173	0.31002	0.02191	0.30024	0.02361	0.30217	0.01247

4. Validating the MxGEF model using the RR-Ni test

The degree to which a statistical model fits a given set of observations is quantified by the RR-Ni test statistic, making it a versatile and powerful tool in statistical analysis. Developed through contributions from Rao and Robson (1974) and further refined by Nikulin (1973a, 1973b, 1973c), this test statistic serves as a general goodness-of-fit measure applicable across various domains, including survival analysis, regression modeling, and time series analysis. Its broad applicability stems from its ability to evaluate the predictive accuracy of models while identifying potential issues that might otherwise go unnoticed. In essence, the RR-Ni test statistic plays a pivotal role in model selection, assessment of model fit, and diagnosis of model inadequacies.

One of the most notable features of the RR-Ni test statistic is its capacity to detect deviations from expected patterns that other statistical tests may overlook. Unlike some traditional goodness-of-fit tests, the RR-Ni test is robust to outliers, enabling it to identify and analyze datasets with extreme values effectively. This robustness makes the RR-Ni test particularly valuable in fields such as finance, where the detection and analysis of rare but significant events such as market crashes or large price fluctuations are critical. By providing insights into these extreme events, the RR-Ni test aids in understanding their underlying causes and developing strategies to mitigate their impacts.

The RR-Ni test statistic can be employed to compare the fit of different statistical models to the same dataset. This capability is essential for model selection, as it allows researchers to identify the model that best aligns with the observed data. By comparing the RR-Ni test statistics of competing models, analysts can determine which model provides the most accurate representation of the data, thereby enhancing the reliability of subsequent analyses. A fundamental use of the RR-Ni test statistic is its ability to assess the overall goodness of fit of a statistical model. When the RR-Ni test statistic is small, it indicates a strong alignment between the model and the data, suggesting that the model adequately captures the underlying patterns. Conversely, a large RR-Ni test statistic signals a poor fit, highlighting areas where the model may fail to represent the data accurately. This diagnostic capability ensures that models are appropriately specified and reliable for practical applications. Outliers, defined as data points that

deviate significantly from the general trend, can substantially affect the performance of statistical models. The RR-Ni test statistic is adept at identifying such outliers, allowing analysts to either adjust the model to account for these anomalies or remove them if they are deemed erroneous. By addressing outliers, the RR-Ni test helps improve the overall fit and predictive power of the model. Beyond evaluating model fit, the RR-Ni test statistic can diagnose underlying issues within a statistical model. For instance, a large RR-Ni test statistic may indicate that the model is misspecified or that certain assumptions underlying the model are violated. This diagnostic insight enables researchers to refine their models, ensuring they better reflect the complexities of the data and improving their predictive accuracy.

In addition to its technical applications, the RR-Ni test statistic holds significant importance in practical contexts. For example, in financial modeling, where the accuracy of predictions can have substantial economic implications, the RR-Ni test helps ensure that models are robust and reliable. Similarly, in survival analysis, where unobserved heterogeneity and censoring pose unique challenges, the RR-Ni test provides a rigorous framework for validating model assumptions and assessing their adequacy. Furthermore, in fields such as epidemiology and engineering, where understanding time-to-event data is crucial, the RR-Ni test facilitates the development of more precise and informative models. Under the RR-Ni statistic, we need to test the following null hypothesis

$$H_0 : \Pr\{z_i \leq z\} = F_{\underline{\mathbf{P}}}(z), \quad z \in \mathbb{R}, \quad \underline{\mathbf{P}} = (\underline{\mathbf{P}}_1, \underline{\mathbf{P}}_2, \cdots, \underline{\mathbf{P}}_s)^T,$$

Then, the RR-Ni statistic can be expressed as

$$Y^2(\widehat{\underline{\mathbf{P}}}_n) = X_n^2(\widehat{\underline{\mathbf{P}}}_n) + \frac{1}{n}\ell^T(\widehat{\underline{\mathbf{P}}}_n)(\mathbf{I}(\widehat{\underline{\mathbf{P}}}_n) - \mathbf{J}(\widehat{\underline{\mathbf{P}}}_n))^{-1}\ell(\widehat{\underline{\mathbf{P}}}_n),$$

where

$$X_n^2(\underline{\mathbf{P}}) = \left([np_1(\underline{\mathbf{P}})]^{-\frac{1}{2}}\left[-np_1(\underline{\mathbf{P}}) + \underline{\mathbf{P}}_1\right], \cdots, [np_b(\underline{\mathbf{P}})]^{-\frac{1}{2}}\left[-np_b(\underline{\mathbf{P}}) + \underline{\mathbf{P}}_b\right]\right)^T$$

and

$$\mathbf{J}(\underline{\mathbf{P}}) = B(\underline{\mathbf{P}})^T B(\underline{\mathbf{P}}),$$

refers to the information matrix where

$$B(\underline{\mathbf{P}}) = \left[\frac{1}{\sqrt{p}_i}\frac{\partial}{\partial\mu}(\underline{\mathbf{P}})\right]_{r\times s}|_{(i=1,2,\cdots,b \text{ and } \kappa=1,2,\cdots,s)},$$

and

$$\ell(\underline{\mathbf{P}}) = (\ell_1(\underline{\mathbf{P}}), ..., \ell_s(\underline{\mathbf{P}}))^T \text{ with } \ell_\kappa(\underline{\mathbf{P}}) = \sum_{i=1}^{r}\frac{\underline{\mathbf{P}}_i}{p_i}\frac{\partial p_i(\underline{\mathbf{P}})}{\partial\underline{\mathbf{P}}_\kappa},$$

The $Y^2(\widehat{\underline{\mathbf{P}}}_n)$ statistic has $(b-1)$ degrees of freedom (DF) and is accompanied by χ^2_{b-1} distribution, where the observations. $x_1, x_2, \cdots, x_n$ that are collected in $\mathbf{I}_1, \mathbf{I}_2, \cdots, \mathbf{I}_b$ (these b subintervals are mutually disjoint: $\mathbf{I}_j =]a_{j,b} - 1; a_{j,b}]$). The intervals $\mathbf{I}_j$'s limits for $a_{j,b}$ are determined as follows

$$p_j(\underline{\mathbf{P}}) = \int_{a_{j,b}-1}^{a_{j,b}} f_{\underline{\mathbf{P}}}(x)\,dx|_{(\ j=1,2,\cdots,b)},$$

and

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

4.0.1. Uncensored assessing for the RR-Ni statistic In numerous situations, the objective of a goodness-of-fit test extends beyond simply assessing whether a specific distribution adequately represents the data; it also involves estimating the parameters of that distribution. Simulation studies play a crucial role in evaluating the accuracy and precision of these parameter estimates under various conditions, thereby aiding in the selection of the most appropriate distribution for further analysis. In particular, simulation studies conducted in an uncensored setting using the RR-Ni statistic serve as a valuable tool for comparing and assessing different probability distributions in a controlled environment. These simulations offer insights into how well the RR-Ni test performs across diverse scenarios, helping to guide decisions regarding which distribution is best suited for subsequent analyses.

To validate the findings presented in this study, we performed a comprehensive numerical simulation analysis. To test the null hypothesis H_0, we generated RR-Ni statistics for the MxGEF model using simulated samples of varying sizes: n =20,40,250,350,600, and 1000, with a total sample size of 12,000. For different significance levels (ϵ=0.01,0.02,0.05,0.1), we calculated the average number of non-rejections under the null hypothesis, based on the condition $Y^2 \leq \chi^2_\epsilon\,(b-1)$ .The corresponding empirical and theoretical levels are summarized in Table 3. A close alignment between the empirical level values and their respective theoretical counterparts is evident, indicating strong agreement. Based on these results, we conclude that the proposed test demonstrates excellent performance for the MxGEF distribution, confirming its suitability for practical applications.

Table 3: Uncensored assessing for the RR-Ni statistic for $\epsilon = 0.01, 0.02, 0.05, 0.1$ and $N = 12000$.

$n \downarrow \epsilon \longrightarrow$	$\epsilon = 0.01$	$\epsilon = 0.02$	$\epsilon = 0.05$	$\epsilon = 0.1$
$n = 20$	0.9925	0.9830	0.9533	0.9023
$n = 40$	0.9921	0.9825	0.9527	0.9021
$n = 250$	0.9915	0.9818	0.9522	0.9014
$n = 350$	0.9910	0.9811	0.9515	0.9008
$n = 600$	0.9904	0.9808	0.9507	0.9007
$n = 1000$	0.9901	0.9805	0.9503	0.9004

4.1. Validating the MxGEF model using the B-Ni test

Due to Bagdonavicius and Nikulin (2011) and Bagdonavicius et al. (2013), we can verify the suitability of the MxGEF model when the parameters are unknown and the data are censored where null hypothesis can be expressed as

$$H_0 : F(x) \in F_0 = \left\{ F_0(x, \underline{\mathbf{P}}), x \in R^1,\ \underline{\mathbf{P}} \in \underline{\mathbf{P}} \subset R^s \right\},$$

Let's divide the limited amount of time $[0, \tau]$ into $\kappa | \kappa = 1, 2, \cdots, s$ shorter time periods. Where is the maximum runtime of the research and $\mathbf{I}_j = (a_{j-1}, a_{j,b}]$; $0 =< a_{0,b} < a_{1,b} ... < a_{\kappa-1,b} < a_{\kappa,b} = +\infty$. The anticipated worth of $\widehat{a_{j,b}}$ can be said the following if $x_{(i)}$ is the i^{th} element in the ordered statistics $(x_{(1)}, , , x_{(n)})$ and if $\mathbf{\Lambda}^{-1}$ refers to the cumulative hazard-rate function and

$$\widehat{a_{j,b}} = \mathbf{\Lambda}^{-1} \left((E_{j,X} - \sum_{l=1}^{i-1} \mathbf{\Lambda}(x_{(l)}, \widehat{\underline{\mathbf{P}}}))/(n-i+1), \widehat{\underline{\mathbf{P}}} \right), \quad \widehat{a_\kappa} = x_{(n)}|_{(j=1,\ldots,\kappa)},$$

where

$$e_{j,Z} = E_\kappa/\kappa \text{ for every } j.$$

and

$$\begin{aligned} E_{j,Z} &= (n-i+1)\mathbf{\Lambda}(\widehat{a_{j,b}},\widehat{\underline{\mathbf{P}}}) + \sum_{l=1}^{i-1}\mathbf{\Lambda}(x_{(l)},\widehat{\underline{\mathbf{P}}}) \\ &= \sum_{i:z_i>a_{j,b}} (\mathbf{\Lambda}(a_{j,b}\wedge z_i,\widehat{\underline{\mathbf{P}}}) - \mathbf{\Lambda}(a_{j-1},\widehat{\underline{\mathbf{P}}}), E_\kappa = \sum_{i=1}^{n}\mathbf{\Lambda}(z_i,\widehat{\underline{\mathbf{P}}}). \end{aligned}$$

The $a_{j,b}$ functions for random data, and the $e_{j,Z}$ For the κ selected periods, anticipated failure rates are equal. Statistical data $Y_n^2 = \mathbf{Z}^T\widehat{\mathbf{S}}^{-1}\mathbf{Z}$, where

$$\mathbf{Z} = (Z_1,...,Z_\kappa)^x, Z_j = \frac{1}{\sqrt{n}}(\mathbf{W}_{j,Z} - e_{j,Z})|_{(\ j=1,2,...,\kappa)}$$

and $\mathbf{W}_{j,Z}$ can be used to test a hypothesis since it reflects the total number of failures that have been recorded throughout these time-shared. The elements of the B-Ni test statistic

$$Y_n^2 = \sum_{j=1}^{\kappa}\frac{1}{\mathbf{W}_{j,Z}}(\mathbf{W}_{j,Z} - e_{j,Z})^2 + \mathbf{D}_{W,G}$$

where

$$\begin{aligned} \mathbf{D}_{W,G} &= \widehat{\mathbf{V}}^T\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{V}}, \widehat{\mathbf{S}}^{-1} = \widehat{\mathbf{B}}^{-1} + \widehat{\mathbf{M}}^{-1}\widehat{\mathbf{B}}^T\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{M}}\widehat{\mathbf{B}}^{-1}, \\ \widehat{\mathbf{G}} &= [\widehat{g}_{ll'}]_{s\times s} = \widehat{i} - \widehat{\mathbf{M}}\widehat{\mathbf{B}}^{-1}\widehat{\mathbf{M}}^x, \\ \widehat{\mathbf{M}}_{lj} &= \frac{1}{n}\sum_{i:z_i\in\mathbf{I}_j}\rho_i\frac{\partial}{\partial\underline{\mathbf{P}}}\ln\left[\lambda_{i,\underline{\mathbf{P}}}(z_i)\right], \mathbf{W}_{j,Z} = \sum_{i:z_i\in\mathbf{I}_j}\rho_i,\ \widehat{\mathbf{B}}_j = n^{-1}\mathbf{W}_{j,Z}, \\ \widehat{\mathbf{V}}_l &= \sum_{j=1}^{\kappa}\widehat{\mathbf{M}}_{lj}\widehat{\mathbf{B}}_j^{-1}\mathbf{Z}_j, \quad l,l'=1,...,s, \\ \widehat{i}_{ll'} &= n^{-1}\sum_{i=1}^{n}\rho_i\frac{\partial}{\partial\underline{\mathbf{P}}_l}\ln\left[\lambda_{i,\underline{\mathbf{P}}}(z_i)\right]\frac{\partial}{\partial\underline{\mathbf{P}}_{l'}}\ln\left[\lambda_{i,\widehat{\underline{\mathbf{P}}}}(z_i)\right] \end{aligned}$$

and

$$\widehat{g}_{ll'} = \widehat{i}_{ll'} - \sum_{j=1}^{\kappa}\widehat{\mathbf{M}}_{lj}\widehat{\mathbf{M}}_{l'j}\hat{A}_j^{-1},$$

and

$$\widehat{\mathbf{M}}_{lj} = \frac{1}{n}\sum_{i:z_i\in\mathbf{I}_j}\rho_i\frac{\partial}{\partial\underline{\mathbf{P}}}\ln\left[\lambda_{i,\widehat{\underline{\mathbf{P}}}}(z_i)\right].$$

4.1.1. Censored assessing for the B-Ni statistic Censored simulation studies under the B-Ni statistics are an important tool for evaluating and comparing different probability distributions when dealing with censored data. These studies can provide valuable insights into the performance of the B-Ni tests under different types and levels of censoring, and can inform decisions about which distribution to use for subsequent analyses. It is intended that the sample that was produced ($N = 12000$) will be censored at 20% and that DF$= 5$ To check if the sample agrees with the MxGEF model's null hypothesis, grouping intervals will be used. For various theoretical levels, we determine the average value of the non-rejection numbers of the null hypothesis. ($\epsilon = 0.01, 0.02, 0.05, 0.1$), where $Y^2 \leq \chi_\epsilon^2(r-1)$. The theoretical and empirical levels are compared in Table 4, which demonstrates how closely the determined empirical level matches the value of the relevant theoretical level. We conclude that the customized test is ideally suited to the MxGEF model as a consequence.

Table 4: Censored assessing for the B-Ni statistic for $\epsilon = 0.01; 0.02; 0.05; 0.1$ and $N = 12000$.

$n \downarrow \& \epsilon \longrightarrow$	$\epsilon = 0.01$	$\epsilon = 0.02$	$\epsilon = 0.05$	$\epsilon = 0.1$
$n = 20$	0.9932	0.9828	0.9532	0.9021
$n = 40$	0.9925	0.9815	0.9524	0.9016
$n = 250$	0.9916	0.9810	0.9514	0.9011
$n = 350$	0.9910	0.9806	0.9510	0.9009
$n = 600$	0.9906	0.9803	0.9508	0.9003
$n = 1000$	0.9903	0.9801	0.9506	0.9001

We conclude from these findings that the empirical significance level of the Y_n^2 The theoretical level of the chi-square distribution on degrees of freedom corresponds to the statistical level at which it is statistically significant. The censored data acquired from the MxGEF distribution may thus be satisfactorily fitted using the suggested test, according to this evidence.

5. The emergency care data

In the field of medicine, frailty models serve as valuable tools for analyzing risk factors and predicting the prognosis of various diseases. These models enable researchers to investigate how individual-level characteristics, such as age, gender, and genetic predispositions, affect patient outcomes. Additionally, they account for unmeasured or unobserved factors that may contribute to the risk of disease progression or mortality. Frailty models are widely applied in epidemiological research, clinical trials, and cohort studies to evaluate the impact of different treatments or interventions on patient health, providing a comprehensive framework for understanding complex health dynamics.

The real dataset utilized in this study was provided by the emergency department of the public proximity health institution (Echatt, El Tarf, Algeria) and encompasses observations collected throughout March 2023. The aim of this research was to investigate the relationship between various clinical variables and outcomes in patients seeking care at the emergency department. Ethical guidelines were strictly followed, and the necessary approvals were obtained prior to data collection. The dataset consists of 30 unique individuals, each representing a distinct observation. Six key clinical variables were recorded for each individual: age (in years), minimum and maximum blood pressure (in mmHg), blood glucose level (in mg/dL), heart rate (in beats per minute, BPM), and oxygen saturation (SaO2 %). To ensure high-quality data, stringent measures were implemented during the collection process. These included meticulous documentation of patient information, adherence to standardized measurement protocols, and regular quality checks to identify and address any missing or inconsistent data. Such rigorous procedures enhance the reliability and accuracy of the dataset, making it particularly valuable for analyzing the relationships between clinical factors and emergency room outcomes. This dataset provides an opportunity to evaluate the goodness-of-fit of the MxGEF model distribution and its ability to accurately represent the observed patterns and variability in emergency care data. Specifically, we present point estimates for two fitted models: the MxGEF model with a Weibull baseline hazard-rate function and the MxGEF model with a Gompertz baseline hazard-rate function. To determine the most appropriate model among those fitted to the data, we employ the modified chi-squared test proposed by Bagdonavičius and Nikulin (2011). This statistical approach allows us to assess the validity and applicability of the MxGEF distribution in the context of emergency care data, thereby contributing to a deeper understanding of the underlying survival dynamics and heterogeneity in this critical healthcare setting. For more new data sets see Abiad et al. (2025), Alizadeh et al. (2025), Das et al. (2025) and Ibrahim et al. (2025).

5.1. Validation of the MxGEF model under the Weibull baseline hazard-rate function

Assuming that these data are distributed according to the MxGEF model with Weibull baseline hazard-rate function. Then, using R statistical software (the BB package), the maximum likelihood estimates of the parameter vector $\underline{\mathbf{P}}$ are obtained as

$$\begin{aligned}\widehat{\kappa} &= (0.822547, \widehat{\rho} = 0.63951, \widehat{\sigma^2} = 1.03591,\\ \widehat{\beta_1} &= 0.015748, \widehat{\beta_2} = 0.50024, \widehat{\beta_3} = 0.20368,\\ \beta_4 &= -0.39517, \widehat{\beta_5} = 0.27845, \widehat{\beta_6} = 0.83695).\end{aligned}$$

According to Bagdonavičius and Nikulin (2011) for censored data, we take for example 5 intervals $(r = 5)$ as number of classes. The elements of the estimated Fisher information matrix $I\left(\widehat{\underline{\mathbf{P}}}\right)$ are presented as follows:

$$I\left(\widehat{\underline{\mathbf{P}}}\right) = \begin{pmatrix} 1.35 & -3.27 & 0.363 & 0.001 & 1.111 & 1.006 & 0.0236 & 0.9026 & 2.6255 \\ & 0.626 & 0.325 & -2.966 & -0.002 & 0.363 & -9.3025 & 1.8547 & 0.0001 \\ & & 1.025 & -7.263 & 0.965 & 0.0003 & -8.3262 & 0.9681 & 0.1057 \\ & & & 1.954 & 2.125 & -0.252 & 0.00215 & 0.0002 & 0.1547 \\ & & & & 2.152 & 0.097 & 1.0025 & 1.6685 & 0.6326 \\ & & & & & 0.952 & 1.0255 & -6.000 & 2.31125 \\ & & & & & & 0.6685 & -3.2626 & 4.0216 \\ & & & & & & & 0.3155 & 1.2515 \\ & & & & & & & & 1.9658 \end{pmatrix}$$

The calculated value of the test statistic for the proposed MxGEF model with a Weibull baseline hazard-rate function is $Y_n^2 = 9.0025741$. Comparing this value with the critical value from the chi-squared distribution, $\chi^2_{0.05}(4) = 9.488$, we observe that Y_n^2 is less than the critical value. This result implies that we fail to reject the null hypothesis at the 5% significance level, indicating that the data is consistent with the proposed model. Therefore, the emergency care data can be adequately fitted using the MxGEF model with a Weibull baseline hazard-rate function. The model demonstrates a proper fit, effectively capturing the underlying survival patterns and accounting for unobserved heterogeneity in the dataset. This validation underscores the robustness and applicability of the MxGEF model in analyzing complex survival data.

5.2. Validation of the MxGEF model under the Gompertz baseline hazard-rate function

Assuming that these data are distributed according to the MxGEF model with Gompertz baseline hazard-rate function. Then, using R statistical software (the BB package), the maximum likelihood estimator of the parameter vector $\underline{\mathbf{P}}$ can be obtained as

$$\begin{aligned}\widehat{\kappa_1} &= 1.005428, \widehat{\rho_1} = 0.90035, \widehat{\sigma^2} = 1.01584,\\ \widehat{\beta_1} &= -0.248331, \widehat{\beta_2} = 0.215862, \widehat{\beta_3} = 0.02913,\\ \widehat{\beta_4} &= 0.723841, \widehat{\beta_5} = 0.0700561, \widehat{\beta_6} = 0.68597.\end{aligned}$$

We take $r = 5$ intervals and the estimated Fisher matrix expressed as

$$I\left(\widehat{\underline{\mathbf{P}}}\right) = \begin{pmatrix} 0.96 & -6.24 & 0.216 & 0.952 & -4.686 & 1.037 & 0.962 & 1.03325 & 1.002 \\ & 1.025 & 2.00 & 1.003 & 2.003 & 0.002 & 0.951 & 0.001 & 0.209 \\ & & 1.033 & 1.957 & -9.33 & 1.026 & -8.620 & 0.633 & 1.625 \\ & & & 0.966 & 0.022 & 1.025 & 1.855 & 0.001 & 1.251 \\ & & & & 0.326 & 1.856 & 0.097 & -3.36 & 1.003 \\ & & & & & 1.220 & 1.954 & 0.125 & 0.004 \\ & & & & & & 1.204 & 1.025 & 2.032 \\ & & & & & & & 1.965 & 1.006 \\ & & & & & & & & 0.549 \end{pmatrix}.$$

To evaluate the compatibility of the emergency care data with the proposed MxGEF model with a Gompertz baseline hazard-rate function, we calculate the value of the Bagdonavičius and Nikulin (2011) statistic, denoted as Y_n^2. The computed value of this statistic is Y_n^2 =8065984. This test statistic plays a crucial role in assessing the goodness-of-fit of the model to the observed data. Next, we compare Y_n^2 with the critical value from the chi-squared distribution $\chi^2_{0.05}$ at a significance level of α =5%. For this test, the degrees of freedom are determined by the number of parameters estimated in the model. In our case, the degrees of freedom are 4, where $r = 5$ represents the total number of parameters in the MxGEF model with the Gompertz baseline hazard-rate function. From the chi-squared distribution table, the critical value for α=5% and 4 degrees of freedom is $\chi^2_{0.05}(5-1) = 9.488$. Upon comparing the calculated test statistic $Y^2 < \chi^2_{0.05}(5-1) = 9.488$ with the critical value $\chi^2_{0.05}(5-1) = 9.488$, it becomes evident that $Y^2 < \chi^2_{0.05}(5-1)$. This result implies that the null hypothesis, which assumes that the data follows the proposed MxGEF model with a Gompertz baseline hazard-rate function, cannot be rejected at the 5% significance level. Consequently, we conclude that the emergency care data is indeed compatible with the proposed model. This compatibility indicates that the MxGEF model with a Gompertz baseline hazard-rate function provides an adequate fit to the emergency care data, effectively capturing the underlying survival patterns and accounting for unobserved heterogeneity. The robustness of the model is further validated through its ability to align with the observed data under rigorous statistical testing. Thus, the MxGEF model can be confidently applied for analyzing and predicting survival outcomes in emergency care settings, offering valuable insights into risk assessment and decision-making processes.

6. Conclusion

Frailty models are essential in survival analysis for addressing unobserved heterogeneity among individuals, which can stem from factors such as genetics, environmental influences, or lifestyle choices. These models enable the identification of such factors and their impact on survival outcomes, thereby improving the accuracy of predictions. In this paper, we introduced the MxGEF model as a novel approach to survival analysis. To assess its suitability, we utilized the RR-Ni goodness-of-fit test, examining the model's characteristics and comparing its performance with commonly used distributions in frailty modeling. Through extensive simulation studies and real-world applications, including data from an emergency hospital in Algeria, we demonstrated that the MxGEF model effectively captures heterogeneity and provides superior model fit compared to existing alternatives. Our results indicate that the MxGEF model is a promising advancement in frailty modeling, with the potential to enhance the precision of survival analyses in diverse fields, particularly in emergency care. Furthermore, we explored the applicability of the MxGEF model in the insurance sector through simulations and real data analysis, highlighting its versatility and potential to address risk assessment challenges in this domain. The MxGEF model represents a significant contribution to the field of survival analysis, offering improved accuracy and broader applicability across multiple disciplines. Future research could further investigate its performance under varying conditions and expand its use in other domains where heterogeneity plays a critical role.

While the MxGEF model represents a significant advancement in addressing unobserved heterogeneity in survival analysis, several avenues remain open for future research. One key area involves extending the model to handle more complex datasets, such as those with competing risks or recurrent events, which are common in medical and engineering applications. Additionally, further exploration of Bayesian methodologies could enhance the model's flexibility and ability to incorporate prior information, particularly in scenarios with limited data. Another promising direction is the development of computationally efficient algorithms for estimating parameters in high-dimensional settings, ensuring scalability for large datasets. The integration of machine learning techniques, such as neural networks, could also improve the model's predictive power by capturing nonlinear relationships in the data. Furthermore, validating the MxGEF model under various censoring mechanisms, including interval and informative censoring, would broaden its applicability across diverse fields. Investigating the performance of the model in real-time data streams, such as those from wearable health devices, could provide new insights into dynamic risk assessment. Lastly, expanding the model's application to interdisciplinary domains, such as climate science and social sciences, would demonstrate its versatility and relevance beyond traditional fields like medicine

and insurance. These future studies promise to refine and extend the capabilities of frailty models, enhancing their utility in both theoretical and practical contexts.

Apendix: Description of R Code

```
# Clear the workspace
rm(list=ls())

# Create an empty matrix with 0 rows and 4 columns (number of parameters)
matrix_data <- matrix(nrow = 0, ncol = 4)
n <- 20  # Sample size
S <- 5000   # Number of simulations
iteration <- 1

# Covariates
x1 <- 0.01
x2 <- 0.1

# Generate Weibull random variable
t <- rweibull(n, 0.8, 0.9)

# Loop for simulations
while (iteration <= S) {
  # Parameter estimates
  p0 <- c(0.9, 0.8, 0.5, 0.7)

  # Define the function for parameter estimation
  DQXg <- function(p) {
    # Log-likelihood function (implementation goes here)
  }

  # Load BB library for optimization
  library(BB)

  # Check for convergence and store results
  result <- BBsolve(par = p, fn = DQXg, quiet = TRUE)
  print(result$par)
  matrix_data <- rbind(matrix_data, result$par)

  iteration <- iteration + 1
}
```

REFERENCES

1. Aalen, O .O. (1988). Heterogeneity in Survival Analysis. Statistics in Medicine ,7, 1121-1137.
2. Aalen, O.O, Tretli, S. (1999). Analysing the incidence of testis cancer by means of a frailty model. Cancer Causes and Control ,10, 285-292.
3. Aalen, O.O. (1992). Modelling Heterogeneity in Survival Analysis by the Compound Poisson Distribution. Annals of Applied Probability, 4 (2), 951-972.

4. Abiad, M., Alsadat, N., Abd El-Raouf, M. M., Yousof, H. M., & Kumar, A. (2025). Different copula types and reliability applications for a new fisk probability model. Alexandria Engineering Journal, 110, 512-526.
5. Abouelmagd, T. H. M., Hamed, M. S., Hamedani, G. G., Ali, M. M., Goual, H., Korkmaz, M. C., & Yousof, H. M. (2019). The zero truncated Poisson Burr X family of distributions with properties, characterizations, applications, and validation test. Journal of Nonlinear Sciences and Applications, 12(5), 314-336.
6. Ahmed, B., Ali, M. M. and Yousof, H. M. (2022). A Novel G Family for Single Acceptance Sampling Plan with Application in Quality and Risk Decisions, Annals of Data Science, 10.1007/s40745-022-00451-3
7. 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.
8. 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.
9. Alizadeh, M., Afshari, M., Ranjbar, V., Merovci, F. and 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.
10. Aljadani, A., Mansour, M. M., & Yousof, H. M. (2024). A Novel Model for Finance and Reliability Applications: Theory, Practices and Financial Peaks Over a Random Threshold Value-at-Risk Analysis. Pakistan Journal of Statistics and Operation Research, 20(3), 489-515. https://doi.org/10.18187/pjsor.v20i3.4439
11. Bagdonavičius, V., and Nikulin, M. (2011). Chi-squared goodness-of-fit test for right censored data.International Journal of Applied Mathematics and Statistics, 24, 30-50.
12. Bagdonavičius, V., Levuliene, R., J., and Nikulin, M. (2013). Chi-squared goodness-of-fit tests for parametric accelerated failure time models. Communications in Statistics-Theory and Methods, 42(15), 2768-2785.
13. Balakrishnan, N.Peng., Y. (2006). Generalized gamma frailty model. Stat Med, 25(16), 2797-2816.
14. Cox DR. (1972). Regression models and life-tables. J Roy Stat Soc: Ser B (Methodol), 34(2), 187-202.
15. 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 Laplace Distribution for House Prices. Mathematical and Computational Applications, 30(1), 4.
16. Elbatal, I., Diab, L. S., Ghorbal, A. B., Yousof, H. M., Elgarhy, M. and 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.
17. Elbers, C., Ridder, G. (1982). True and spurious duration dependence: The identifiability of the proportional hazard model. Rev Econ Stud, 49(3), 403-409.
18. Emam, W.; Tashkandy, Y.; Goual, H.; Hamida, T.; Hiba, A.; Ali, M.M.; Yousof, H.M.; Ibrahim, M. A New One-Parameter Distribution for Right Censored Bayesian and Non-Bayesian Distributional Validation under Various Estimation Methods. Mathematics 2023, 11, 897. https://doi.org/10.3390/math11040897
19. Goual, H., and Yousof, H. M. (2020a). Validation of Burr XII inverse Rayleigh model via a modified chi-squared goodness-of-fit test. Journal of Applied Statistics, 47(3), 393-423.
20. Goual, H., Yousof, H. M. and Ali, M. M. (2020b). Lomax inverse Weibull model: properties, applications, and a modified Chi-squared goodness-of-fit test for validation. Journal of Nonlinear Sciences & Applications (JNSA), 13(6), 330-353.
21. 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.
22. Hamed, M. S., Cordeiro, G. M. and 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
23. Hamedani, G. G., Goual, H., Emam, W., Tashkandy, Y., Ahmad Bhatti, F., Ibrahim, M., & Yousof, H. M. (2023). A new right-skewed one-parameter distribution with mathematical characterizations, distributional validation, and actuarial risk analysis, with applications. Symmetry, 15(7), 1297.
24. 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.
25. Hashempour, M., Alizadeh, M. and Yousof, H. M. (2023). A New Lindley Extension: Estimation, Risk Assessment and Analysis Under Bimodal Right Skewed Precipitation Data. Annals of Data Science, 1-40.
26. Hougaard, P. (2012). Analysis of multivariate survival data. Springer Science & Business Media, Berlin.
27. Ibrahim, J., Chen, M., Sinha, D. (2001). Bayesian survival analysis springer series in statistics. Springer, New York, 978-981.
28. Ibrahim, M., Ali, M. M., Goual, H. and Yousof, H. M. (2022). Censored and uncensored validation for the double Burr XII model using a new Nikulin-Rao-Robson goodness-of-fit test with Bayesian and non-Bayesian estimation, Pakistan Journal of Statistics and Operation Research, 18(4), 2022.
29. Ibrahim, M., Altun, E., Goual, H., and 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.
30. 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
31. Ibrahim, M., Hamedani, G. G., Butt, N. S. and Yousof, H. M. (2022). Expanding the Nadarajah Haghighi Model: Copula, Censored and Uncensored Validation, Characterizations and Applications. Pakistan Journal of Statistics and Operation Research, 18(3), 537-553. https://doi.org/10.18187/pjsor.v18i3.3420
32. Ibrahim, M., Yadav, A. S., Yousof, H. M., Goual, H., & Hamedani, G. G. (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.
33. Ibrahim, M.; Emam, W.; Tashkandy, Y.; Ali, M.M.; Yousof, H.M. (2023). Bayesian and Non-Bayesian Risk Analysis and Assessment under Left-Skewed Insurance Data and a Novel Compound Reciprocal Rayleigh Extension. Mathematics 2023, 11, 1593. https://doi.org/10.3390/ math11071593
34. Ibrahim. M., Aidi, K., Ali, M. M. and Yousof, H. M. (2021). The Exponential Generalized Log-Logistic Model: Bagdonavičius-Nikulin test for Validation and Non-Bayesian Estimation Methods. Communications for Statistical Applications and Methods, 29(1), 681–705.
35. Khedr, A. M., Nofal, Z. M., El Gebaly, Y. M. and Yousof, H. M. (2023). A Novel Family of Compound Probability Distributions: Properties, Copulas, Risk Analysis and Assessment under a Reinsurance Revenues Data Set. Thailand Statistician, forthcoming.
36. Korkmaz, M. Ç., Altun, E., Yousof, H. M., Afify, A. Z. and Nadarajah, S. (2018). The Burr X Pareto Distribution: Properties, Applications and VaR Estimation. Journal of Risk and Financial Management, 11(1), 1.
37. Lehmann, EL., Casella, G. (2006) .Theory of point estimation. Springer Science & Business Media, Berlin.
38. 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.
39. Loprinzi, C. L., Laurie, J. A., Wieand, H. S., Krook, J. E., Novotny, P. J., Kugler, J. W. and Klatt, N. E. (1994). Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group. Journal of Clinical Oncology, 12(3), 601-607.
40. 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.
41. 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 (2020) 57–66. DOI: 10.47443/cm.2020.0018
42. 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.
43. Mansour, M. M., Ibrahim, 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.
44. Mansour, M., Korkmaz, M. Ç., Ali, M. M., Yousof, H. M., Ansari, S. I., & Ibrahim, M. (2020). A generalization of the exponentiated Weibull model with properties, Copula and application. Eurasian Bulletin of Mathematics, 3(2), 84-102.
45. 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.
46. Mansour, M., Yousof, H. M., Shehata, W. A. M., & Ibrahim, M. (2020). A new two parameter Burr XII distribution: properties, copula, different estimation methods and modeling acute bone cancer data. Journal of Nonlinear Science and Applications, 13(5), 223-238.
47. Mazucheli, J., Coelho-Barros, EA., Achcar, JA. (2016). An alternative reparametrization for the weighted lindley distribution. Pesquisa Operacional , 36(2), 345-353.
48. McGilchrist, C.A., Aisbett, C.W. (1991). Regression with Frailty in Survival Analysis. Biometrics, 47, 461-466.
49. Mota, A., Milani, E.A., Calsavara, V.F., Tomazella, V.L., Leão, J., Ramos, P.L., Ferreira, P.H. and Louzada, F. (2021). Weighted Lindley frailty model: estimation and application to lung cancer data. Lifetime Data Analysis, 27(4), 561-587.
50. Nikulin, M.S, (1973b). Chi-squared test for continuous distributions with shift and scale parameters, Theory of Probability and its Aplications. 18, 559-568.
51. Nikulin, M.S, (1973c). On a Chi-squared test for continuous distributions, Theory of Probability and its Aplications. 19, 638-639.
52. Nikulin, M.S. (1973a). Chi-squared test for normality. in proceedings of the International Vilnius Conference on Probability Theory and Mathematical Statistics, 2, 119-122.
53. Pickles, A., Crouchley, R. (1995). A comparison of frailty models for multivariate survival data. Stat Med , 14(13), 1447-1461.
54. Rao, K. C., Robson, D. S. (1974). A Chi-square statistic for goodness-of-fit tests within the exponential family. Communication in Statistics, 3, 1139-1153.
55. Rasekhi, M., Altun, E., Alizadeh, M. and 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.
56. Robert, C., Casella, G.(2013). Monte Carlo statistical methods. Springer Science & Business Media, Berlin.
57. Salah, M. M., El-Morshedy, M., Eliwa, M. S. and Yousof, H. M. (2020). Expanded Fréchet Model: Mathematical Properties, Copula, Different Estimation Methods, Applications and Validation Testing. Mathematics, 8(11), 1949.
58. 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.
59. Shehata, W. A., Goual, H., Hamida, T., Hiba, A., Hamedani, G. G., Al-Nefaie, A. H., Ibrahim, M., Butt, N. S., Osman, R. M. A., and 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.
60. Shrahili, M.; Elbatal, I. and Yousof, H. M. Asymmetric Density for Risk Claim-Size Data: Prediction and Bimodal Data Applications. Symmetry 2021, 13, 2357. https://doi.org/10.3390/sym13122357
61. Tasche, D. (2002), Expected Shortfall and Beyond, Journal of Banking and Finance, 26, 1519-1533.

62. Teghri, S., Goual, H., Loubna, H., Butt, N. S., Khedr, A. M., Yousof, H. M., Ibrahim, 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.
63. Vaupel, J.W., Manton, K.G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16(3), 439-454.
64. Wienke, A. (2010). Frailty models in survival analysis. CRC Press, Boca Raton.
65. Yadav, A. S., Goual, H., Alotaibi, R. M., Rezk, H., 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.
66. Yadav, A. S., Shukla, S., Goual, H., Saha, M. and 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.
67. Yousof, H. M., Aidi, K., Hamedani, G. G and 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.
68. 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.
69. Yousof, H. M., Ali, M. M., Goual, H. and 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.
70. 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.
71. 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
72. 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. 2023; 11(9):2179. https://doi.org/10.3390/math11092179
73. 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.
74. Yousof, H. M., Goual, H., Hamida, T., Hiba, A., Hamedani, G.G. and Ibrahim, M. (2022). Censored and Uncensored Nikulin-Rao-Robson Distributional Validation: Characterizations, Classical and Bayesian estimation with Applications.
75. 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.
76. Yousof, H. M., Khaoula, M. K., Goual, H., Hamedani, G.G. and Ibrahim, M. (2022). Accelerated failure time estimation for a novel exponential model with characterizations, validations and different methods of estimation.
77. Yousof, H. M., Saber, M. M., Al-Nefaie, A. H., Butt, N. S., Ibrahim, M. and Alkhayyat, S. L. (2024). A discrete claims-model for the inflated and over-dispersed automobile claims frequencies data: Applications and actuarial risk analysis. Pakistan Journal of Statistics and Operation Research, 261-284.
78. 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 2023, 11, 1284. https://doi.org/10.3390/math11061284
79. Yousof, H.M.; Tashkandy, Y.; Emam, W.; Ali, M.M.; Ibrahim, M. (2023). A New Reciprocal Weibull Extension for Modeling Extreme Values with Risk Analysis under Insurance Data. Mathematics 2023, 11, 966. https://doi.org/10.3390/ math11040966