Shahid Bahonar University of Kerman
Journal of Mahani Mathematical Research
Print ISSN: 2251-7952 Online ISSN: 2645-4505
JMMR 2012

DISTRIBUTIONAL NIKULIN-RAO-ROBSON VALIDITY UNDER A NOVEL GAMMA EXTENSION WITH CHARACTERIZATIONS AND RISK ASSESSMENT

G.G. Hamedani ✉, Ahmad M. Aboalkhair, Khaoula Aidi, Ali S. Hadi, Haitham M. Yousof, and Mohamed Ibrahim

Special issue Dedicated to memory of professor Mahbanoo Tata

Article type: Research Article

(Received: 15 February 2024, Received in revised form 15 February 2024)

(Accepted: 16 March 2024, Published Online: 13 April 2024)

Abstract. In this work, a novel probability distribution is introduced and studied. Some characterizations are presented. Several financial risk indicators, such as the value-at-risk, tail-valueat-risk, tail variance, tail Mean-Variance, and mean excess loss function are considered under the maximum likelihood estimation, the ordinary least squares, the weighted least squares, and the Anderson Darling estimation methods. These four methods were applied for the actuarial evaluation under a simulation study and under an application to insurance claims data. For distributional validation under the complete data, the well-known Nikulin-Rao-Robson statistic is considered. The Nikulin-Rao-Robson test statistic is assessed under a simulation study and under three complete real data sets. For censored distributional validation, a new version of the Nikulin-Rao-Robson statistic is considered. The new Nikulin-Rao-Robson test statistic is assessed under a comprehensive simulation study and under three censored real data sets.

Keywords: Characterizations, Distributional Validition, Nikulin-Rao-Robson, Risk Assessment, Value-at-risk
2020 MSC: 62N01; 62N02; 62E10.

1. Introduction

In this paper we will present and study a new continuous probability distribution, but we will study the new distribution via new aspects that differ from those dealt with by most researchers. We will neglect many theoretical results and algebraic derivations, not because they are not important, but to allow for the opportunity to highlight more applied aspects in the field of risk assessment and analysis and in the field of distributive verification and its related practical applications on complete data and censored data. We will, however,

✉ gholamhoss.hamedani@marquette.edu, ORCID: 0000-0001-7976-1088
https://doi.org/10.22103/jmmr.2024.22954.1585

Publisher: Shahid Bahonar University of Kerman
How to cite: G.G. Hamedani, Ahmad M. Aboalkhair, Khaoula Aidi, Ali S. Hadi, Haitham M. Yousof and Mohamed Ibrahim, *Distributional Nikulin-Rao-Robson validity under a novel gamma extension with characterizations and risk assessment*, J. Mahani Math. Res. 2024; 13(5): 1 - 32.

cover some theoretical aspects of the new distribution by presenting and discussing some characterizations based on two truncated moments; in terms of the hazard function; and based on the basis of the conditional expectation of a function of the random variable (RV). However, as we have indicated, we will focus on practical and applied aspects of the proposed distribution in the following areas specifically:

(1) In the field of analyzing and evaluating the risks facing insurance companies by evaluating and analyzing insurance claims data by studying a set of commonly used financial indicators such as: the value-at-risk (VAR), tail-value-at-risk (TVAR) (also known as conditional tail expectation, conditional-value-at-risk (CVAR), tail variance (TV), tail Mean-Variance (TMV) and the mean excess loss (MEL) function. For the purpose of computing the main key risk indicators (KRIs), the following estimation techniques are discussed: the maximum likelihood estimation (MLE) method, the ordinary least squares (OLS) method, the weighted least squares estimation (WLSE) method, and the Anderson Darling estimation (ADE) method. These four aforementioned methods were used and employed in two different directions of financial and actuarial assessment, namely simulation under three confidence levels (CLs) and various sample sizes are considered for applications to insurance claims data.

(2) To complete the requirements of the actuarial analysis of risks, we provide a simulation study to compare the performance of the estimators of VaR based on insurance data

(3) In the framework of distributional validation and statistical hypothesis tests for the complete data, the well-known Nikulin-Rao-Robson (NRR) statistic (Y^2), which is based on the uncensored maximum likelihood estimators (UMLEs) on initial non-grouped data, is considered under a probability model called the Burr X exponentiated gamma (BXEG) model. The Y^2 statistic is assessed via a simulation study under three real data sets.

(4) In the framework of distributional validation and statistical hypothesis tests for the censored data, a modified NRR statistic (M^2), which is based on the censored maximum likelihood estimators (CMLEs) on initial non-grouped data, is considered under the BXEG model. The M^2 statistic is assessed via comprehensive simulation study under three real data sets.

Following Yousof et al. [37], the cumulative distribution function (CDF) of the BXEG model can be written as

$$F_{\underline{\mathbf{V}}}(x) = \left[1 - \varrho\left(x;\theta\right)\right]^{a} |_{x\geq 0}, \tag{1}$$

where $\underline{\mathbf{V}} = (a, \lambda, \theta)$, $a, \lambda, \theta > 0$, $\varsigma(x) = 1 - (1 + \lambda x)\exp(-\lambda x)$, $\varrho\ (x;\theta) = \exp\left\{-\left[\varsigma(x)^{-\theta} - 1\right]^{-2}\right\}$ and $[\varsigma(x)]^{\theta}$ refers to the CDF of the exponentiated gamma model proposed by Gupta et al. [12]. The exponentiated gamma model is flexible enough to accommodate both monotonic as well as nonmonotonic failure rates. The probability density function (PDF) correspondint to (1) can be expressed as

$$f_{\underline{\mathbf{V}}}(x) = 2a\theta\lambda^2 x \frac{\exp(-\lambda x)\,\varsigma(x)^{2\theta-1}\left[1 - \varsigma(x)^{\theta}\right]^{-3}}{\exp\left\{\left[\varsigma(x)^{-\theta} - 1\right]^{-2}\right\}[1 - \varrho(x;\theta)]^{-a+1}}|_{x>0}. \tag{2}$$

Generally, there are several criteria that may be applied to determine if a statistical model is legitimate. For the uncensored data, the most popular tests are those based on the empirical functions, such as the likelihood ratio test, Akaike information criteria, Bayesian information criteria, or chi-square tests. These tests include Kolmogorov-Smirnov, Anderson-Darling, and other statistics. The NRR statistic (Y^2), based on the MLEs on initial non-grouped data, is of particular importance among these goodness-of-fit evaluations. This Nikulin ([26], [27], [28]) and Rao and Robson [29] statistic restores information lost during data grouping and has a chi-square distribution. However, the existence of censorship renders all the conventional goodness-of-fit tests invalid and leads to several practical issues. As a result, several researchers offered various revisions of current goodness-of-fit tests. A modified NRR statistic was created by Bagdonavicius and Nikulin [4] for statistical distributions with unknown parameters and right censoring. This version of the NRR statistic may be used to fit data from domains like survival analysis, dependability, and others where data is often censored since it recovers all the information lost during data regrouping. In this study, we will provide modified NRR chi-square goodness-of-fit test statistics for fitting full and right-censored data to the suggested model, following Nikulin ([26], [27], [28] and Rao and Robson [29].

The NRR statistic is a well-known variant of the traditional chi-squared tests in the situation of full data. It is based on differences between two estimators of the probability for falling into grouping intervals. One estimate is based on the empirical distribution function, and the other on maximum likelihood estimates of the tested model's unobserved parameters using ungrouped initial data. (see Nikulin ([26], [27], [28]), and Rao and Robson [29] for more details and see Goual and Yousof [9], Goual et al. [10], Goual et al. [11] for more relevant applications under uncensored schemes). Generally, the statistical methods for testing hypotheses and the censored validity of parametric distributions are in increasing development, but the presence of censorship is considered as a big challenge. In the history of the statistical literature for verification tests in the case of controlled data, there are many contributions that are clear and cannot be ignored, and there are many who have contributed to the field of application.

In the statistical literature, there are not many studies that dealt with the test NRR, because of the scarcity of these studies, they can be counted, and here we will mention the recent ones of them: Goual et al. [10] for the odd Lindley exponentiated exponential validation by a modified NRR goodness of fit test with some applications to censored and uncensored data. Abouelmagd et al. [1] for distributional validity of the zero truncated Poisson-Burr-X G family of distributions. Ibrahim et al. [20] for new modified validation test under a new extension of Lindley distribution with characterizations and estimation different methods. Goual et al. [11] for validation of the Burr type XII inverse Rayleigh model via a modified NRR chi-squared goodness-of-fit test. Yadav et al. [33] for the distributional validation of the Topp-Leone-Lomax distribution via a modified NRR goodness-of-fit test with different classical estimation methods. Ibrahim et al. [18] for a modified NRR goodness-of-fit test for the censored distributional validaty using a new Burr type XII model with different classical methods of estimation and censored regression modeling. Yousof et al. [36] for Aa new inverted Rayleigh model with copulas, properties, various classical methods modified NRR right censored test for distributional validation. Finally, Yadav et al. [34] for distributional xgamma exponential validation via the NRR goodness-of- fit statistic test under censored and uncensored sample with estimation under different methods. For other applications, see Yousof et al. [38], Ibrahim et al. [17], Khalil et al. [22], Emam et al. [7], Emam et al. [8], Aidi et al. [2] and Yousof et al. [35].

In this study, the BXEG distribution is derived and used, the complete and right censored scenarios are used to validate a modified chi-squared goodness-of-fit test statistic based on the NRR test (Y^2) and the modified NRR test (M^2) respectively. First, the Y^2 statistic test is used for testing the null hypothesis H_0 according to which a certain complete sample belongs to a BXEG model. The NRR statistic test is evaluated using a simulation study via the Barzilai-Borwein (BB) algorithm (see Ravi and Gilbert [30]) in the case of complete data and a simulation study in the case of censored data. In the simulation studies, we have relied on the standard mean square error (MSEs) in the evaluation process, taking into account different sample sizes to help us evaluate the behavior of the test with an increase in the sample size. The Barzilai and Borwein gradient methodology has received a lot of interest from a variety of optimization areas. This is due to its practical usefulness, computer affordability, and simplicity. Using spectral analysis techniques, this paper proves root-linear global convergence for the Barzilai and Borwein method for strictly convex quadratic problems presented in infinite-dimensional Hilbert spaces. The application of these discoveries to two optimization problems controlled by partial differential equations is demonstrated.

Three uncensored real data sets are used to apply the NRR test (Y^2) for distributional validation. The uncensored real data that have been considered and included in the analysis are the times between failures for repairable items data, the reliability data and the strengths of glass fibers data. Three right-censored real data sets are used to apply the modified NRR test (M^2) for distributional validation. The censored real data that have been considered and included in the analysis are times to infection of kidney dialysis patients data, the bone marrow transplant data and the strength of a certain type of braided cord data.

The new NRR statistical test demonstrated that the new model is an effective substitute for examining two right censored data sets. In this regard, we will describe a few recent research results that extended the NRR in new or modified ways. Given that the NRR goodness-of-fit test has specific requirements, strict procedures, and demands censored data, it is important to note that the browser for statistical literature on this topic (NRR goodness-of-fit test) will not find many new NRR goodness-of-fit extensions but a few research that applied this test. As is generally known, obtaining fresh censored data to apply to and emphasise the significance of the new test is difficult. In the next few paragraphs, we will discuss a few recent research results that use this test on actual data that had been subject to right-wing censoring, along with a description of the findings from each study independently.

2. Characterization results

The characterizations of the BXEG distribution in the following ways are covered in this section: (i) on the basis of two truncated moments; (ii) in terms of the hazard function and (iii) on the basis of the conditional expectation of a function of the RV. It is not necessary for the CDF to have a closed form for characterisation (i). The characterizations will be presented in the following subsections.

2.1. Characterizations based on two truncated moments.

The characterizations of the BXEG distribution based on the connection between two truncated moments are covered in this subsection. The first characterisation makes use of the theorem of Glänzel [13]. It is obvious that the outcome is still true if interval H is not closed. Please refer to Glänzel [14] to see how stable this categorization is in terms of weak convergence.

Proposition 2.1.1. The RV $X : \Omega \to (0, \infty)$ is continuous, and assume

$$h_{(x)} = \frac{\left\{1 - \varsigma\,(x)^{\theta}\right\}^{3}}{\varrho\,(x;\theta)}\left[1 - \varrho\,(x;\theta)\right]^{1-a}$$

and $g_{(x)} = h_{(x)} \left[\varsigma(x)\right]^{2\theta}$ for $x > 0$. Then, the density of X is (2) if and only if the function $\zeta_{(x)}$ defined in the theorem of Glänzel [13] is

$$\zeta_{(x)} = \frac{1}{2}\left\{1 + \varsigma(x)^{2\theta}\right\}|_{x>0}.$$

Proof. If X has PDF (2), then

$$\overline{F}_{\underline{\mathbf{V}}}(x)\, E\left[h_{(X)} \mid {}_{X \geq x}\right] = a\left[1 - \varsigma(x)^{2\theta}\right]|_{x>0}$$

and

$$\overline{F}_{\underline{\mathbf{V}}}(x)\, E\left[g_{(X)} \mid {}_{X \geq x}\right] = \frac{a}{2}\left[1 + \varsigma(x)^{4\theta}\right]|_{x>0},$$

where $\overline{F}_{\underline{\mathbf{V}}}(x) = 1 - F_{\underline{\mathbf{V}}}(x)$ and finally

$$\zeta_{(x)} h_{(x)} - g_{(x)} = \frac{1}{2} h_{(x)}\left\{1 - \varsigma(x)^{2\theta}\right\} > 0|_{x>0}.$$

Conversely, if ζ has the above form, then

$$s'(x) = \frac{h_{(x)}\zeta'_{(x)}}{\zeta_{(x)} h_{(x)} - g_{(x)}} = 2\theta\lambda^2 \frac{x e^{-\lambda} \varsigma(x)^{2\theta-1}}{1 - \varsigma(x)^{2\theta}},$$

and hence $s(x) = -\log\left\{1 - \left[\varsigma(x)\right]^{2\theta}\right\}|_{x>0}$. In view of theorem of Glanzel (1987), X has PDF (2).

Corollary 2.1.1. If $X : \Omega \to (0, \infty)$ is a continuous RV and $h_{(x)}$ is as in Proposition 2.1.1., then, X has PDF (2) if and only if there exist functions g and ζ defined in the theorem of Glänzel (1987) satisfying the following first order differential equation

$$\frac{h_{(x)}\zeta'_{(x)}}{h_{(x)}\zeta_{(x)} - g_{(x)}} = 2\theta\lambda^2 \frac{x \exp(-\lambda)\, \varsigma(x)^{2\theta-1}}{1 - \varsigma(x)^{2\theta}}.$$

Corollary 2.1.2. The general solution of the above differential equation is

$$\zeta_{(x)} = -I_D(a, \lambda, \theta)\left[1 - \varsigma(x)^{2\theta}\right]^{-1},$$

where $I_D(a, \lambda, \theta) = \int 2\theta\lambda^2 x \exp(-\lambda)\, \varsigma(x)^{2\theta-1}\, g_{(x)}\left[h_{(x)}\right]^{-1} + D$ and D is a constant. A set of functions satisfying this differential equation is presented in Proposition 2.1.1 with $D = \frac{1}{2}$. Clearly, there are other triplets (h, g, ζ) satisfying the conditions of the theorem of Glänzel [13].

2.2. **Characterization based on hazard function.** The hazard function, $h_{F,\underline{\mathbf{V}}}$, of a twice differentiable distribution function, $F_{\underline{\mathbf{V}}}$ with density $f_{\underline{\mathbf{V}}}$, satisfies the first following trivial first differential equation $f\ '_{\underline{\mathbf{V}}}(x)/f\ _{\underline{\mathbf{V}}}(x) = h'_{F,\underline{\mathbf{V}}}(x)/h_{F,\underline{\mathbf{V}}}(x) - h_{F,\underline{\mathbf{V}}}(x)$. For many univariate continuous distributions, this is the only hazard function-based characterization that is currently available. The statement below gives a non-trivial characterization of the BXEG distribution for $a = 1$.

Proposition 2.2.1. Suppose $X : \Omega \to (0, \infty)$ is a continuous RV. The density of X is (2) if and only if the following differential equation holds

$$h'_{F,\underline{\mathbf{V}}}(x) + \lambda h_{F,\underline{\mathbf{V}}}(x) = 2\theta\lambda^2 \exp(-\lambda x) \frac{d}{dx} \left\{ \frac{x\varsigma(x)^{2\theta-1}}{\left[1-\varsigma(x)^{\theta}\right]^3} \right\} |_{x>0},$$

with the initial condition $\lim_{x\to 0} h_F(x) = 0$.

The proof. is straightforward and hence omitted.

2.3. **Characterizations based on conditional expectation.** Hamedani [15] makes the following claim, thus we will use it to characterize the BXEG distribution for $a = 1$.

Proposition 2.3.1. Suppose the RV $X : \Omega \to (a, b)$ is continuous with CDF $F_{\underline{\mathbf{V}}}$. If $\Upsilon_{(x)}$ is a differentiable function on (a, b) with $\lim_{x\to 0^+} \Upsilon_{(x)} = 1$, then for $\delta \neq 1$, $E\left[\Upsilon_{(X)} \mid {}_{X\geq x}\right] = \delta\Upsilon_{(X)}, x \in (a, b)$ if and only if $\Upsilon_{(x)} = \left[1 - F_{\underline{\mathbf{V}}}(x)\right]^{\frac{1}{\delta}-1} |_{x\in(a,b)}$.

Remark 2.3.1. Taking $(a, b) = (0, \infty)$, $\Upsilon_{(x)} = \exp\left\{-\frac{1}{2}\left[\varsigma(x)^{-\theta} - 1\right]^{-2}\right\}$ and $\delta = \frac{2}{3}$, A characterization of the BXEG distribution is provided in proposition 2.3.1. Obviously, there are further potential uses.

3. KRIs

The characterization of risk exposure that the probability-based distributions may offer is sufficient. One value, or at the very least a limited group of numbers, is frequently used to indicate the amount of risk exposure. These risk exposure statistics are obviously functions of a certain model and are frequently referred to as important KRIs. Such KRIs provide actuaries and risk managers with information on the degree to which a firm is exposed to specific types of risk. Numerous KRIs, including the VAR, the TVAR which also known as CVAR), the TV indicator, the TMV) and the MEL function, among others, can be taken into account and examined. The VaR is a quantile of the distribution of aggregate losses in particular. Actuaries and risk managers frequently focus on estimating the likelihood of a negative result, which may be conveyed using the VaR indicator at a certain probability/confidence level. This indicator is frequently used to calculate the amount of capital needed to deal with such probable negative situations. The VAR of the BXEG distribution at the $100q\%$ level, say VAR(X) or $\pi\left(q\right)$, is the $100q\%$ quantile (or percentile). Then, we can simply write

$$\text{VAR}\left(X\right)=\Pr\left(X>Q\left(U\right)\right)=\begin{cases}1\%|_{q=99\%}\\ 5\%|_{q=95\%}\\ \vdots\end{cases}, \tag{3}$$

where $Q\left(U\right)$ is from (3), for a one-year time when $q=99\%$, the interpretation is that there is only a very small chance (1%) that the insurance company will be bankrupted by an adverse outcome over the next year. Generally speaking, if the distribution of gains (or losses) is limited to the normal distribution, it is acknowledged that the number VAR(X) meets all coherence requirements. The data sets for insurance such as the insurance claims and reinsurance revenues are typically skewed whether to the right or to the left , though. Using the normal distribution to describe the revenues from reinsurance and insurance claims is not suitable. The TVAR of X at the $100q\%$ confidence level is the expected loss given that the loss exceeds the $100q\%$ of the distribution of X, then the TVAR of X can be expressed as

$$\text{TVAR}(X)=\mathbb{E}\left(X|X>\pi\left(q\right)\right)=\frac{1}{1-F_{\underline{\mathbf{V}}}\left(\pi\left(q\right)\right)}\int\limits_{\pi(q)}^{\infty}x\,f_{\underline{\mathbf{V}}}\left(x\right)dx,$$

then

$$\text{TVAR}(X)=\frac{1}{1-q}\int\limits_{\pi(q)}^{\infty}x\,f_{\underline{\mathbf{V}}}\left(x\right)dx. \tag{4}$$

The quantity TVAR(X), which gives further details about the tail of the BXEG distribution, is therefore the average of all the VaR values mentioned above at the confidence level q. Moreover, the TVAR(X) can also be expressed as TVAR$(X)=e\left(X\right)+$VAR(X) , where $e\left(X\right)$ is the mean excess loss (MEL) function evaluated at the $100q\%^{th}$ quantile (see Acerbi and Tasche 2002; Tasche, 2002; Wirch, 1990). When the $e\left(X\right)$ value vanishes, then TVAR(X) =VAR(X) and for the very small values of $e\left(X\right)$, the value of TVAR(X) will be very close to VAR(X) .The TV risk indicator, which Furman and Landsman (2006) developed, calculates the loss's deviation from the average along a tail. Explicit expressions for the TV risk indicator under the multivariate normal distribution were also developed by Furman and Landsman (2006). The TV risk indicator (TV(X)) can then be expressed as

$$\text{TV}(X)=\mathbb{E}\left(X^2|X>\pi\left(q\right)\right)-\left[\text{TVAR}(X)\right]^2. \tag{5}$$

As a statistic for the best portfolio choice, Landsman (2010) developed the TMV risk indicator, which is based on the TV risk indicator. Consequently, the TMV risk indicator may be written as

$$\mathrm{TMVq}(X) = \mathrm{TVAR}(X) + \pi \mathrm{TV}(X)|_{0<\pi<1}. \tag{6}$$

Then, for any RV, $\mathrm{TMVq}(X) > \mathrm{TV}(X)$ and, for $\pi = 1$, $\mathrm{TMVq}(X) = \mathrm{TVAR}(X)$. In view of the theoretical complexities and the fact that the quantile function is not known in a certain closed form, we will use the methods that provide numerical solutions. We will use ready-made programs such as "R" and "MATHCAD" to facilitate numerical operations. The use of the numerical methods has become popular recently for many reasons. The most important of which is the availability of ready-made statistical programs and the presence of lots of mathematically complex distributions and models. The fact that has become recognized and cannot be ignored in the field of statistical analysis and mathematical modeling is that the complexity of models is no longer the real problem facing researchers, because statistical programs and packages have, in fact, contributed a lot in simplifying these complexities by providing numerical solutions. In this paper, we have used numerical methods in the process of risk analysis and assessment (see Section 4), and numerical methods have also been used in the problem of distributional validation under the NRR and its new corresponding version (see Section 7).

4. **Risk analysis using different estimation methods**

4.1. **Risk assessment under artificial data.** For the purpose of computing the above mentioned KRIs, the following estimation techniques are discussed in this section: the MLE method, the OLS method, the WLSE method, and the ADE method. Three CLs (q=70%, 90%, 99%) and N=1,000 with various sample sizes ($n = 50, 150, 300, 500$) are considered. All results are reported in Table 1 (n=50; $a = 2, \lambda = 0.7, \theta = 0.5$), Table 2 (n=150; $a = 2, \lambda = 0.7, \theta = 0.5$), Table 3 (n=300; $a = 2, \lambda = 0.7, \theta = 0.5$), and Table 4 (n=500; $a = 2, \lambda = 0.7, \theta = 0.5$). We have deliberately determined the parameter values in this simulation in a way that helps us in the risk analysis process using artificial values. The simulation aims mainly to evaluate the performance of the four methods of risk analysis in the hope of determining the most appropriate and best methods. from Table 1, Table 2, Table 3 and Table 4, we can show the following results:

(1) $\mathrm{VAR}(X)$, $\mathrm{TVAR}(X)$ and $\mathrm{TMVq}(X)$ increase when q increases for all estimation methods.
(2) $\mathrm{TV}(X)$ and $\mathrm{MEL}(X)$ decrease when q increases for all estimation methods.
(3) $\mathrm{VAR}(X)_{\mathrm{WLS}} < \mathrm{VAR}(X)_{\mathrm{ADE}} < \mathrm{VAR}(X)_{\mathrm{MLE}} < \mathrm{VAR}(X)_{\mathrm{OLSE}}$ for most values of q.

(4) $\mathrm{TVAR}(X)_{\mathrm{WLS}} < \mathrm{TVAR}(X)_{\mathrm{ADE}} < \mathrm{TVAR}(X)_{\mathrm{MLE}} < \mathrm{TVAR}(X)_{\mathrm{OLSE}}$ for most values of q.

(5) Through the results of the four tables, we can confirm that all methods perform in an acceptable manner, and one method cannot be preferred over another method decisively. Based on this basic result, we are compelled to present an application on actual data, in the hope that the application will help us in choosing one method over another and identifying the most appropriate and best methods.

Table 1. KRIs under artificial data for $n = 50$.

Method↓	$q\downarrow$	KRIs→	VAR(X)	TVAR(X)	TV(X)	TMVq(X)	MEL(X)
MLE	70%		1.67944	1.91420	0.03389	1.93115	0.23476
	90%		1.96297	2.12906	0.01955	2.13883	0.16608
	99%		2.32604	2.43409	0.00971	2.43895	0.10805
OLSE	70%		1.68364	1.91838	0.03390	1.93533	0.23474
	90%		1.96713	2.13326	0.01957	2.14304	0.16613
	99%		2.33033	2.43840	0.00988	2.44334	0.10807
WLSE	70%		1.68348	1.91916	0.03415	1.93623	0.23568
	90%		1.96812	2.13483	0.01970	2.14468	0.16671
	99%		2.33255	2.44113	0.00946	2.44586	0.10858
ADE	70%		1.68423	1.91931	0.03399	1.93631	0.23509
	90%		1.96813	2.13449	0.01962	2.14430	0.16636
	99%		2.33181	2.44021	0.00942	2.44492	0.10840

4.2. **Risk assessment under insurance claims data.** The historical growth of claims through time for each appropriate exposure (or origin) period is frequently shown in the historical insurance actual data in the form of a triangle presentation. The year the insurance policy was purchased or the time period during which the loss occurred may be regarded as the exposure period. It is obvious that the genesis period need not be annual. For instance, it may be monthly or quarterly origin periods. The development time of an origin period is known as the "claim age" or "claim lag." Data from separate insurance is frequently combined to represent uniform company lines, division levels, or risks. We examine the insurance claims payment triangle from a U.K. Motor Non-Comprehensive account in this article as a practical illustration. We choose a convenient origin period of 2007 to 2013. The insurance claims payment data frame displays the claims data in the manner in which a database would normally keep it. The origin year, which ranges from 2007 to 2013, the development year, and the incremental payments are all included in the first column. It's important to note that this data on insurance claims was initially

TABLE 2. KRIs under artificial data for $n = 150$.

Method↓	$q\downarrow$	KRIs→	VAR(X)	TVAR(X)	TV(X)	TMVq(X)	MEL(X)
MLE	70%		1.68221	1.91820	0.03423	1.93531	0.23599
	90%		1.96724	2.13415	0.01974	2.14402	0.16691
	99%		2.33208	2.44076	0.00947	2.44550	0.10868
OLSE	70%		1.68196	1.91813	0.03428	1.93527	0.23617
	90%		1.96722	2.13423	0.01976	2.14411	0.16701
	99%		2.33228	2.44101	0.00948	2.44575	0.10873
WLSE	70%		1.68041	1.91657	0.03427	1.93370	0.23616
	90%		1.96566	2.13263	0.01975	2.14251	0.16698
	99%		2.33064	2.43933	0.00946	2.44406	0.10869
ADE	70%		1.68208	1.91838	0.03431	1.93553	0.23630
	90%		1.96749	2.13458	0.01978	2.14447	0.16709
	99%		2.33273	2.44150	0.00949	2.44625	0.10877

TABLE 3. KRIs under artificial data for $n = 300$.

Method↓	$q\downarrow$	KRIs→	VAR(X)	TVAR(X)	TV(X)	TMVq(X)	MEL(X)
MLE	70%		1.68297	1.91915	0.03428	1.93629	0.23618
	90%		1.96823	2.13526	0.01977	2.14515	0.16703
	99%		2.33335	2.44211	0.00949	2.44685	0.10876
OLSE	70%		1.68179	1.91841	0.03440	1.93561	0.23663
	90%		1.96761	2.13490	0.01982	2.14481	0.16729
	99%		2.33327	2.44215	0.00951	2.44691	0.10888
WLSE	70%		1.68200	1.91814	0.03427	1.93527	0.23614
	90%		1.96722	2.13421	0.01976	2.14409	0.16699
	99%		2.33224	2.44096	0.00948	2.44570	0.10872
ADE	70%		1.68209	1.91939	0.03207	1.93543	0.23730
	90%		1.96799	2.13533	0.01984	2.14525	0.16734
	99%		2.33376	2.44267	0.00952	2.44743	0.10891

examined using a probability-based distribution. The capability of the insurance firm to handle such occurrences is of importance to actuaries, regulators, investors, and rating agencies. This work proposes certain KRI quantities for the left-skewed insurance claims data under the BXEG distribution, including VAR, TVAR, TV, and TMV. For more details, see et al. Artzner [3], Khedr et al. [21], Hamed et al. [16], Mohamed et al [24] and Ibrahim et al [19].

Table 4. KRIs under artificial data for $n = 500$.

Method↓	$q\downarrow$	KRIs→	VAR(X)	TVAR(X)	TV(X)	TMVq(X)	MEL(X)
MLE	70%		1.68164	1.91818	0.03438	1.93537	0.23654
	90%		1.96736	2.13460	0.01981	2.14450	0.16724
	99%		2.33291	2.44176	0.00951	2.44652	0.10885
OLSE	70%		1.68275	1.91921	0.03436	1.93639	0.23646
	90%		1.96836	2.13557	0.01981	2.14547	0.16721
	99%		2.33385	2.44269	0.00951	2.44745	0.10884
WLSE	70%		1.68217	1.91858	0.03434	1.93575	0.23640
	90%		1.96771	2.13487	0.01980	2.14477	0.16716
	99%		2.33310	2.44191	0.00950	2.44666	0.10881
ADE	70%		1.68289	1.91941	0.03438	1.93660	0.23652
	90%		1.96857	2.13582	0.01982	2.14573	0.16725
	99%		2.33415	2.44302	0.00952	2.44778	0.10887

Table 5 lists the KRIs under the insurance calims data and MLE method for the BXEG model where $\underline{\widehat{\mathbf{V}}} = (0.1306,\ 0.00057,\ 3.26359)$. Table 6 gives the KRIs under the insurance calims data and OLSE method for the BXEG model where $\underline{\widehat{\mathbf{V}}} = (0.16181,\ 0.000485,\ 2.166398)$. Table 7 shows the KRIs under the insurance calims data and WLSE method for the BXEG model where $\underline{\widehat{\mathbf{V}}} = (0.14933,\ 0.00054,\ 2.62303)$. Table 8 presents the KRIs under the insurance calims data and WLSE method for the BXEG model where $\underline{\widehat{\mathbf{V}}} = (0.15002,\ 0.00054,\ 2.65579)$. Based on these tables, the following results can be highlighted:

(1) For all risk assessment methods:

$$\begin{aligned} VaRq(X|_{1-q=30\%}) &< VaRq(X|_{1-q=25\%}) < ... \\ &< VaRq(X|_{1-q=10\%}) < VaRq(X|_{1-q=1\%}). \end{aligned}$$

(2) For all risk assessment methods:

$$\begin{aligned} TVaRq(X|_{1-q=30\%}) &< TVaRq(X|_{1-q=25\%}) < ... \\ &< TVaRq(X|_{1-q=10\%}) < TVaRq(X|_{1-q=1\%}). \end{aligned}$$

(3) For all risk assessment methods:

$$\begin{aligned} TV(X|_{1-q=30\%}) &> TV(X|_{1-q=25\%}) > ... \\ &> TV(X|_{1-q=10\%}) > TV(X|_{1-q=1\%}). \end{aligned}$$

(4) For all risk assessment methods:

$$\begin{aligned} TMVq(X|_{1-q=30\%}) &> TMVq(X|_{1-q=25\%}) > ... \\ &> TMVq(X|_{1-q=10\%}) > TMVq(X|_{1-q=1\%}). \end{aligned}$$

(5) For all risk assessment methods:

$$MEL(X|_{1-q=30\%}) > MEL(X|_{1-q=25\%}) > ... > MEL(X|_{1-q=10\%}) > MEL(X|_{1-q=1\%}).$$

(6) Under the MLE method: The VaRq(X) is monotonically increasing starts with 3673.32932 and ends with 6143.57495, the TVaRq(X) in monotonically incresing starts with 4715.02630 and ends with 6389.85619. However the TVq(X), the TMVq(X) and the MEL(X) are monotonically decreasing.

(7) Under the OLSE method: The VaRq(X) is monotonically increasing starts with 3672.18421 and ends with 6248.33526, the TVaRq(X) in monotonically incresing starts with 4743.90011 and ends with 6521.44231. However the TVq(X), the TMVq(X) and the MEL(X) are monotonically decreasing.

(8) Under the WLSE method: The VaRq(X) is monotonically increasing starts with 3598.25509 and ends with 6025.45856, the TVaRq(X) in monotonically incresing starts with 4614.29847 and ends with 6276.33491. However the TVq(X), the TMVq(X) and the MEL(X) are monotonically decreasing.

(9) Under the AE method: The VaRq(X) is monotonically increasing starts with 3619.36506 and ends with 6035.20922, the TVaRq(X) in monotonically incresing starts with 4630.7444 and ends with 6285.00127. However the TVq(X), the TMVq(X) and the MEL(X) are monotonically decreasing.

(10) Nearly for all q values, the OLSE method is recommended since it provides the most acceptable risk exposure analysis then the MLE method is recommended as a second one. However the other two methods are perform well.

Table 5. KRIs under the insurance calims data and MLE method for the BXEG model.

Method	VaRq(X)	TVaRq(X)	TVq(X)	TMVq(X)	MEL(X)
70%	3673.32932	4715.02630	500739.94226	255084.99743	1041.69698
75%	3968.42115	4894.08892	407070.53533	208429.35658	925.66777
80%	4280.10264	5086.96211	320774.76058	165474.34240	806.85947
85%	4616.32200	5300.73322	241765.88072	126183.67357	684.41122
90%	4995.10918	5550.62151	169527.7267	90314.48486	555.51234
95%	5470.41892	5881.34342	101407.57291	56585.129870	410.92450
99%	6143.57495	6389.85619	44408.50668	28594.109540	246.28124

Table 6. KRIs under the insurance calims data and OLSE method for the BXEG model.

Method	VaRq(X)	TVaRq(X)	TVq(X)	TMVq(X)	MEL(X)
70%	3672.18421	4743.90011	542265.56787	275876.68404	1071.7159
75%	3973.12544	4928.38033	445012.24229	227434.50148	955.25489
80%	4290.91416	5127.89496	355132.7358	182694.26287	836.98081
85%	4634.9389	5350.47828	272054.79375	141377.87516	715.53938
90%	5025.74843	5613.21676	194647.38881	102936.91117	587.46833
95%	5523.69632	5965.94666	120049.30962	65990.60147	442.25034
99%	6248.33526	6521.44231	52358.98199	32700.93330	273.10705

Table 7. KRIs under the insurance calims data and WLSE method for the BXEG model.

Method	VaRq(X)	TVaRq(X)	TVq(X)	TMVq(X)	MEL(X)
70%	3598.25509	4614.29847	482016.66052	245622.62874	1016.04338
75%	3885.06857	4789.04706	393813.20198	201695.64805	903.97848
80%	4187.59439	4977.61656	312579.62484	161267.42897	790.02216
85%	4514.30487	5187.35421	237867.78965	124121.24903	673.04934
90%	4883.94723	5433.90991	168797.62315	89832.72148	549.96268
95%	5351.87114	5762.93731	103230.38935	57378.13199	411.06617
99%	6025.45856	6276.33491	45414.59121	28983.63052	250.87635

Table 8. KRIs under the insurance calims data and WLSE method for the BXEG model.

Method	VaRq(X)	TVaRq(X)	TVq(X)	TMVq(X)	MEL(X)
70%	3619.36506	4630.7444	477422.91071	243342.19975	1011.37934
75%	3904.97787	4804.67882	390009.68636	199809.522	899.70095
80%	4206.12922	4992.34126	309572.24357	159778.46304	786.21204
85%	4531.27267	5201.06366	235588.3116	122995.21946	669.79099
90%	4899.09826	5446.42114	167251.13583	89071.98906	547.32289
95%	5364.73456	5773.89288	102359.34452	56953.56515	409.15832
99%	6035.20922	6285.00127	45164.32392	28867.16323	249.79205

5. Distributional validity

5.1. Distributional validity utilizing the UMLE method. Here, the UMLE method is used to estimate the BXEG distribution's parameters. Let $x_1, x_2, ...x_n$ be the observed values of the random sample from the BXEG model, the uncensored likelihood function is obtained by $L(\underline{\mathbf{V}}) =_{i=1}^{n} f_{\underline{\mathbf{V}}}(x_i)$. Then, the uncensored log-likelihood function is obtained as

$$l(\underline{\mathbf{V}}) = n\ln(2a\theta) + 2n\ln(\lambda) + \sum_{i=1}^{n}\ln(x_i) - \lambda\sum_{i=1}^{n}x_i + (2\theta-1)\sum_{i=1}^{n}\ln\varsigma_i - \sum_{i=1}^{n}s_i^2 + (a-1)\sum_{i=1}^{n}\ln(\varphi_i)$$

where $\varsigma_i = \varsigma(x_i) = 1-(1+\lambda x_i)\exp(-\lambda x_i)$, $s_i = \frac{\varsigma_i^{\theta}}{1-\varsigma_i^{\theta}}$, $\varphi_i = 1-\exp\left(-s_i^2\right)$.
The MLEs $\widehat{a}$, $\widehat{\lambda}$ and $\hat{\theta}$ of the unknown parameters a, λ and θ are derived from the following nonlinear score equations:

$$\frac{\partial l(\underline{\mathbf{V}})}{\partial a} = \frac{n}{a} + \sum_{i=1}^{n}\ln(\varphi_i)$$

$$\begin{aligned}\frac{\partial l(\underline{\mathbf{V}})}{\partial \lambda} &= \frac{2n}{\lambda} - \sum_{i=1}^{n}x_i + \lambda(2\theta-1)\sum_{i=1}^{n}\frac{x_i^2\exp(-\lambda x_i)}{\varsigma_i}\\ &\quad -2\lambda\theta\sum_{i=1}^{n}\frac{x_i^2\exp(-\lambda x_i)\,s_i^2}{\varsigma_i\left(1-\varsigma_i^{\theta}\right)}\\ &\quad +2\lambda\theta(a-1)\sum_{i=1}^{n}\frac{x_i^2\exp(-\lambda x_i)\,s_i^2\exp\left(-s_i^2\right)}{\varsigma_i(1-\varsigma_i^{\theta})\varphi_i},\end{aligned}$$

and

$$\begin{aligned}\frac{\partial l(\underline{\mathbf{V}})}{\partial \theta} &= \frac{n}{\theta} + 2\sum_{i=1}^{n}\ln(\varphi_i) - 2\sum_{i=1}^{n}s_i^2\ln\varsigma_i(1+s_i)\\ &\quad +2(a-1)\sum_{i=1}^{n}s_i^2\ln\varsigma_i(1+s_i)\exp\left(-s_i^2\right)\varphi_i^{-1}.\end{aligned}$$

To solve these equations simultaneously, we use ready-made statistical packages that are specially designed to solve this kind of equations. Hence, we employ numerical techniques like the Newton-Raphson method, the Monte Carlo method, or the BB-solve package to obyain the numerical solution.

5.2. Distributional validity utilizing the right CMLE method. Let us consider $X = (X_1, X_2, ..., X_n)^{\mathbf{T}}$ a sample from the BXEG with the parameter vector which can contain right censored data with fixed censoring time τ. Each X_i can be written as $X_i = (x_i, \Delta_i)$ where

$\Delta_i = \{.0 \quad \text{if } \ x_i \text{ is a censoring time } \ 1 \qquad \text{if } x_i \text{ is a failure time}$

The right censored likelihood function can be given by

$$l_n(\underline{\mathbf{V}}) =_{i=1}^n f_{\underline{\mathbf{V}}}^{\Delta_i}(x_i) S_{\underline{\mathbf{V}}}^{1-\Delta_i}(x_i).$$

where $S_{\underline{\mathbf{V}}}(x_i) = 1 - F_{\underline{\mathbf{V}}}(x_i)$ is the survival function of the BXEG model and then the right censored log-likelihood function $L_n(\underline{\mathbf{V}})$ is equivalent to

$$L_{n,\Delta_i}(\underline{\mathbf{V}}) = \sum_{i=1}^{n} \Delta_i \ln f_{\underline{\mathbf{V}}}(x_i) + \sum_{i=1}^{n} (1-\Delta_i) \ln S_{\underline{\mathbf{V}}}(x_i)$$

or

$$\begin{aligned} L_{n,\Delta_i}(\underline{\mathbf{V}}) &= \sum_{i=1}^{n} \Delta_i \begin{bmatrix} \ln(2a\theta) + 2\ln(\lambda) + \ln(x_i) - \lambda x_i \\ + (2\theta - 1)\ln \varsigma_i - s_i^2 + (a-1)\ln(\varphi_i) \end{bmatrix} \\ &+ \sum_{i=1}^{n} (1-\Delta_i) \ln(\varphi_i^a). \end{aligned}$$

The following nonlinear scoring equations must be solved in order to produce the right CMLEs:

$$\frac{\partial L_{n,\Delta_i}(\underline{\mathbf{V}})}{\partial a} = \sum_{i=1}^{n} \Delta_i \left[\frac{1}{a} + \ln(\varphi_i)\right] - \sum_{i=1}^{n} (1-\Delta_i) \frac{\varphi_i^a \ln(\varphi_i)}{1-\varphi_i^a}$$

$$\begin{aligned} \frac{\partial L_{n,\Delta_i}(\underline{\mathbf{V}})}{\partial \lambda} &= \sum_{i=1}^{n} \Delta_i \begin{bmatrix} \frac{2}{\lambda} - x_i + \frac{\lambda(2\theta-1)x_i^2 \exp(-\lambda x_i)}{\varsigma_i} \\ -\frac{2\lambda\theta x_i^2 \exp(-\lambda x_i) s_i^2}{\varsigma_i(1-\varsigma_i^\theta)} + \frac{2\lambda\theta(a-1)x_i^2 \exp(-\lambda x_i) s_i^2 \exp\left(-s_i^2\right)}{\varsigma_i(1-\varsigma_i^\theta)\varphi_i} \end{bmatrix} \\ &- \sum_{i=1}^{n} (1-\Delta_i) \frac{2\lambda\theta a x_i^2 \exp(-\lambda x_i) s_i^2 \exp\left(-s_i^2\right) \varphi_i^{a-1}}{\varsigma_i(1-\varsigma_i^\theta)(1-\varphi_i^a)} \end{aligned}$$

$$\begin{aligned} \frac{\partial L_{n,\Delta_i}(\underline{\mathbf{V}})}{\partial \theta} &= \sum_{i=1}^{n} \Delta_i \begin{bmatrix} \frac{1}{\theta} + 2\ln(\varphi_i) - 2s_i^2 \ln \varsigma_i (1+s_i) \\ +2(a-1)s_i^2 \ln \varsigma_i (1+s_i) \exp\left(-s_i^2\right) \varphi_i^{-1} \end{bmatrix} \\ &- \sum_{i=1}^{n} (1-\Delta_i) \frac{2a s_i^2 \ln \varsigma_i (1+s_i) \exp\left(-s_i^2\right) \varphi_i^{a-1}}{1-\varphi_i^a}. \end{aligned}$$

Similar, to the complete data scenario, we employ numerical techniques like the Newton-Raphson method, the Monte Carlo method, or the BB-solve package to compute the MLEs. Many authors does not prefer solving nonlinear systems of equations resulting from the setting the derivative of the likelihood function or its log to zero when the search space is of higher dimension than two. This is because of the existence of local maxima. Since the CDF of the GXED exists in closed form, you may wish to consider using the Elemental Percentile method presented in Castillo and Hadi [6].

6. Testing procedures

6.1. Testing procedures for the Y^2 statistic.

For testing the null hypothesis H_0 according to which a sample $X_1, X_2, ...X_n$ belongs to (1), where

$$H_0 = \Pr(X_i \leq x) = F_{\underline{\mathbf{V}}}(x)|||_{x\geq 0},$$

Consider r equiprobable grouping intervals $I_1, I_2, \cdots, I_r$ where $I_j =]b_{j-1}, b_j]$; $I_i \cap I_j = \phi \quad i \neq j$ and $\bigcup_{j=1}^{r} I_j = R^1$ such as

$$p_j = \int_{b_{j-1}}^{b_j} f_{\underline{\mathbf{V}}}(x)dx = \frac{1}{r}|_{\ j=1,2,..,r},$$

and $b_j = F^{-1}(j/r), \quad j = 1, 2, .., r$. If $\upsilon = (\upsilon_1, \upsilon_2, ..., \upsilon_r)^T$ represents the number of observed X_i grouping into these intervals I_j, and the vector $T_n(\underline{\mathbf{V}})$ is

$$T_n(\underline{\mathbf{V}}) = \left(\frac{\upsilon_1 - np_{1,(\underline{\mathbf{V}})}}{\sqrt{np_{1,(\underline{\mathbf{V}})}}}, \frac{\upsilon_2 - np_{2,(\underline{\mathbf{V}})}}{\sqrt{np_{2,(\underline{\mathbf{V}})}}}, ..., \frac{\upsilon_r - n}{\sqrt{np_{r,(\underline{\mathbf{V}})}}} \right)^T.$$

The NRR statistic Y^2 proposed by Nikulin [26], [27], [28] and Rao and Robson [29] is defined by

$$Y_n^2\left(\widehat{\underline{\mathbf{V}}}\right) = T_n^2\left(\widehat{\underline{\mathbf{V}}}\right) + \frac{1}{n} l^T\left(\widehat{\underline{\mathbf{V}}}\right)\left(I\left(\widehat{\underline{\mathbf{V}}}\right) - J\left(\widehat{\underline{\mathbf{V}}}\right)\right)^{-1} l\left(\widehat{\underline{\mathbf{V}}}\right),$$

where $I\left(\widehat{\underline{\mathbf{V}}}\right)$ and $J\left(\widehat{\underline{\mathbf{V}}}\right)$ are the estimated information matrices on non-grouped and grouped data respectively, and $\widehat{\underline{\mathbf{V}}}$ is the vector of the MLEs on initial data. The elements of the vector $l\left(\widehat{\underline{\mathbf{V}}}\right) = \left(l_k\left(\widehat{\underline{\mathbf{V}}}\right)\right)^T_{1\times s}$ are

$$l_k\left(\widehat{\underline{\mathbf{V}}}\right) = \sum_{i=1}^{r} \frac{\nu_j}{p_j} \frac{\partial}{\partial \widehat{\underline{\mathbf{V}}}_k} p_j\left(\widehat{\underline{\mathbf{V}}}\right) \text{ and } J\left(\widehat{\underline{\mathbf{V}}}\right) = B(\widehat{\underline{\mathbf{V}}})^T B(\widehat{\underline{\mathbf{V}}}),$$

where

$$B(\widehat{\underline{\mathbf{V}}}) = \left(\frac{1}{\sqrt{p_i}} \frac{\partial}{\partial \underline{\mathbf{V}}_i} p_i\right)_{r\times s} |_{k=1,2,...,s,\ i=1,2,...,r},$$

where s is the number of the model parameters. The distribution of $Y^2\left(\widehat{\underline{\mathbf{V}}}\right)$ is a chi-square with $r-1$ degrees of freedom. To construct the test statistic Y^2 corresponding to the BXEG with a parameter vector $\underline{\mathbf{V}} = (a, \lambda, \theta)^T$, first we calculate the MLEs $\underline{\widehat{\mathbf{V}}} = (\hat{a}, \hat{\lambda}, \hat{\theta})^T$ and the limit intervals b_j. Secondly, the derivatives $\frac{\partial}{\partial \underline{\widehat{\mathbf{V}}}_k} p_j \left(\underline{\widehat{\mathbf{V}}}\right)$ are deduced as follow

$$\frac{\partial}{\partial \hat{a}} p_j \left(\underline{\widehat{\mathbf{V}}}\right) = \varphi^a(b_j) \ln\left[\varphi(b_j)\right] - \ \varphi^a(b_{j-1}) \ln\left[\varphi(b_{j-1})\right],$$

$$\frac{\partial}{\partial \hat{\lambda}} p_j \left(\underline{\widehat{\mathbf{V}}}\right) = 2\lambda\theta a \left[\begin{array}{c} \frac{b_j^2 \exp(-\lambda b_j) s^2(b_j) \exp\left[-s^2(b_j)\right] \varphi^{a-1}(b_j)}{\varsigma(b_j)(1-\varsigma^{\theta}(b_j))} \\ -\frac{b_{j-1}^2 e^{-\lambda b_{j-1}} s^2(b_{j-1}) \exp\left[-s^2(b_{j-1})\right] \varphi^{a}-1(b_{j-1})}{\varsigma(b_{j-1})(1-\varsigma^{\theta}(b_{j-1}))} \end{array}\right],$$

and

$$\frac{\partial}{\partial \hat{\theta}} p_j \left(\underline{\widehat{\mathbf{V}}}\right) = 2a \left[\begin{array}{c} s^2(b_j) \ln\varsigma(b_j) \left[1 + s(b_j)\right] \exp\left[-s^2(b_j)\right] \varphi^{a-1}(b_j) \\ -s^2(b_{j-1}) \ln\varsigma(b_{j-1}) \left(1 + s(b_{j-1})\right) \exp\left[-s^2(b_{j-1})\right] \varphi^{a-1}(b_{j-1}) \end{array}\right].$$

Finally we obtain the statistic $Y^2 \left(\underline{\widehat{\mathbf{V}}}\right)$ which allows to verify if data belong to the BXEG distribution.

6.2. **Testing procedures for the M^2 test statistic with right censorship.** To verify if a right censored sample $X = (X_1, X_2, ..., X_n)^{\mathbf{T}}$ with fixed censored time τ, follows a parametric model $F_{0,\underline{\mathbf{V}}}(x)$, $\Pr(X_i \le x \mid H_0) = F_{0,\underline{\mathbf{V}}}(x)$, $x \ge 0$. The NRR statistic described above was adjustment by Bagdonavicius and Nikulin [4]. Generally, the NRR statistic is established based on the vector $\Lambda_j = \frac{1}{\sqrt{n}}(O_{j,X} - e_{j,X})|_{j=1,2,\ldots,r \text{ and } r \succ s}$, where $O_{j,X}$ and $e_{j,X}$ are the observed numbers of failures to fall and expected numbers of failures to fall into the grouping intervals I_j, the statistic M^2 is defined by

$$M^2 = \Lambda^T \widehat{\Sigma}^- \Lambda$$

where $\widehat{\Sigma}^-$ refers to the generalized inverse of the well-known covariance matrix $\widehat{\Sigma}$. For facilitating the calculation process, this novel NRR statistical test can be expressed as follows

$$M^2 = \sum_{j=1}^{r} \frac{1}{O_{j,X}} (O_{j,X} - e_{j,X})^2 + \Omega$$

with the quadratic form Ω obtained as

$$
\begin{aligned}
\Omega &= W^T\widehat{G}^{-}W, \widehat{A}_j = O_{j,X}/n, O_{j,X} = \sum_{i:X_i\in I_j}\Delta_i, \widehat{G} = [\widehat{g}_{ll'}]_{sxs}\ ,\\
\widehat{g}_{ll'} &= \widehat{i}_{ll'} - \sum_{j=1}^{r}\widehat{C}_{lj}\widehat{C}_{l'j}\hat{A}_j^{-1}, \quad \widehat{C}_{lj} = \frac{1}{n}\sum_{i:x_i\in I_j}\Delta_i\frac{\partial}{\partial\underline{\widehat{\mathbf{V}}}}\ln\tau_{\underline{\widehat{\mathbf{V}}}}(x_i),\\
\widehat{i}_{ll'} &= \frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{1}{\partial\underline{\widehat{\mathbf{V}}}_l}\partial\ln\tau_{\underline{\widehat{\mathbf{V}}}}(x_i)\frac{1}{\partial\underline{\widehat{\mathbf{V}}}_{l'}}\partial\ln\tau_{\underline{\widehat{\mathbf{V}}}}(x_i),\\
\widehat{W}_l &= \sum_{j=1}^{r}\widehat{C}_{lj}\widehat{A}_j^{-1}\Lambda_j|_{\ l,l'=1,\ldots,s},
\end{aligned}
$$

where $\tau_{\underline{\widehat{\mathbf{V}}}}(x_i) = f_{\underline{\widehat{\mathbf{V}}}}(x_i)/S_{\underline{\widehat{\mathbf{V}}}}(x_i)$ is the hazard rate function of the BXEG model. Under the null hypothesis H_0, the limit distribution of the statistic M^2 is a chi-square with $r = rank(\Sigma)$ degrees of freedom. For more details on modified chi-squatre tests, one can see the book by Voinov et al. [32]. For testing the null hypothesis that a right censored sample is described by the BXEG distribution, we develope M^2 corresponding to this distribution. At that end, we have to compute the MLEs $\underline{\widehat{\mathbf{V}}} = (\hat{a}, \hat{\lambda}, \widehat{\theta})^T$ on initial data (see section 3), the estimated information matrix $\widehat{i}_{ll'}$ which can be deduced from the score functions and the estimated limit intervals $\hat{b}_j$. To apply this test statistic, the expected failure times $e_{j,X}$ to fall into the grouping intervals I_j must be the same for any j, so the estimated interval limits $\hat{b}_j$ are equal to

$$
\hat{b}_j = H^{-1}\left[\frac{1}{n-i+1}\left(E_{j,X} - \sum_{l=1}^{i-1}H_{\underline{\widehat{\mathbf{V}}}}(x_l)\right), \underline{\widehat{\mathbf{V}}}\right],
$$

where $\hat{b}_r = \max\left(X_{(n)},\tau\right)$, $E_r = \sum_{i=1}^{n}H_{\underline{\widehat{\mathbf{V}}}}(x_i)$, $E_{j,X} = \frac{-j}{r-1}\sum_{i=1}^{n}\ln\left(1-\varphi_i^a\right)|_{\ j=1,..r-1}$, and $H_{\underline{\widehat{\mathbf{V}}}}(x_i) = -\ln\left[S_{\underline{\widehat{\mathbf{V}}}}(x_i)\right]$ is the cumulative rate function of the BXEG model. So, the numbers $e_{j,X}$ and $O_{j,X}$ can be obtained. Then, we can derive (and then calculate) the components of the estimated matrix $\widehat{K}$ as follows:

$$
\widehat{K}_{1j,\underline{\mathbf{V}}} = \frac{1}{n}\sum_{i:x_i\in I_j}^{n}\Delta_i\ \left[\frac{1}{a} + \frac{\ln(\varphi_i)}{1-\varphi_i^a}\right]
$$

$$
\widehat{K}_{2j,\underline{\mathbf{V}}} = \frac{1}{n}\sum_{i:x_i\in I_j}^{n}\Delta_i\left[\begin{array}{c}
\frac{2}{\lambda} - x_i + \frac{\lambda(2\theta-1)x_i^2\exp(-\lambda x_i)}{\varsigma_i}\\
-\frac{2\lambda\theta x_i^2\exp(-\lambda x_i)s_i^2}{\varsigma_i(1-\varsigma_i^\theta)}\\
+\frac{2\lambda\theta(a-1)x_i^2\exp(-\lambda x_i)s_i^2\exp\left(-s_i^2\right)}{\varsigma_i(1-\varsigma_i^\theta)\varphi_i}\\
+\frac{2\lambda\theta a x_i^2\exp(-\lambda x_i)s_i^2\exp\left(-s_i^2\right)\varphi_i^{a-1}}{\varsigma_i(1-\varsigma_i^\theta)(1-\varphi_i^a)}
\end{array}\right]
$$

$$\widehat{K}_{3j,\underline{\mathbf{V}}} = \frac{1}{n}\sum_{i:x_i\in I_j}^{n}\Delta_i\left[\begin{array}{c}\frac{1}{\theta}+2\ln(\varphi_i)+2(a-1)s_i^2\ln\varsigma_i(1+s_i)\exp\left(-s_i^2\right)\varphi_i^{-1}\\ -2s_i^2\ln\varsigma_i(1+s_i)+\frac{2as_i^2\ln\varsigma_i(1+s_i)\exp\left(-s_i^2\right)\varphi_i^{a-1}}{1-\varphi_i^a}\end{array}\right]$$

and the estimated matrix $\hat{W}$ is derived from the matrix $\widehat{K}$. Therefore the test statistic can be obtained easily:

$$M_n^2\left(\underline{\hat{\mathbf{V}}}\right) = \sum_{j=1}^{r}\frac{\left(O_{j,X}-e_{j,X}\right)^2}{O_{j,X}} + \hat{W}^T\left[\hat{\imath}_{ll'} - \sum_{j=1}^{r}\widehat{K}_{lj}\widehat{K}_{l'j}\hat{A}_j^{-1}\right]^{-1}\hat{W}$$

7. Assessing the $Y^2$and M^2 via some applications

We performed a significant investigation using numerical simulations in this section to demonstrate the flexibility and effectiveness of the tests suggested in this work. We then used actual data from reliability and survival analysis to run these tests.

7.1. **Simulating the Y^2 under the UMLE method.** For simulating the M^2 under the UMLE, the data were simulated $N = 10,000$ times under the sample sizes $n_1 = 25$, $n_2 = 50$, $n_3 = 130$, $n_4 = 350$, $n_5 = 500$, $n_6 = 1000$. Using the BB algorithm and the R software, the MLEs and their mean square errors (MSEs) are calculated and presented in Table 9. For testing the null hypothesis H_0 according to which the data follows the BXEG distribution, we calculate the Y^2 statistical test, then it is compared with the different empirical levels of rejection of the null hypothesis H_0 when $Y^2 > \chi^2_\alpha(r-1)$ where empirical levels of rejection of the null hypothesis are $\alpha_1 = 0.01$, $\alpha_2 = 0.05$ and $\alpha_3 = 0.10$. Table 10 gives the theoretical risk and empirical risk for complete case. The levels simulated for the statistic Y^2 agree with those corresponding to the theoretical levels of the chi-square distribution with $(r-1)$ degrees of freedom, it is noticed, after accounting for simulation errors. In light of this, we can state that the test suggested in this study can appropriately adapt the data obtained from a BXEG model. For example:

(1) For α_1, it is observed that: $Y^2_{0.01} = (0.0048,\ 0.0053,\ 0.0061,\ 0.0076,\ 0.0083,\ 0.0091) < \chi^2_{0.01}(r-1)$. So, we can accept the null hypothesis that the data follows the BXEG distribution.
(2) For α_2, it is seen that: $Y^2_{0.05} = (0.0274,\ 0.0286,\ 0.0312,\ 0.0346,\ 0.0389,\ 0.0409) < \chi^2_{0.05}(r-1)$. Therefore, we can accept the null hypothesis that the data follows the BXEG distribution.
(3) For α_3, it is noted that: $Y^2_{0.10} = (0.0867,\ 0.0894,\ 0.0923,\ 0.0963,\ 0.0997,\ 0.1013) < \chi^2_{0.10}(r-1)$. Hence, we can accept the null hypothesis that the data follows the BXEG distribution.

TABLE 9. MLEs and their MSEs for the complete case.

MLEs & MSEs$\downarrow$ $n \rightarrow$	25	50	130	350	500	1000
$\hat{a}\vert(a_0 = 2)$	1.9623	1.9698	1.9743	1.9876	1.9952	1.9994
MSE	0.0078	0.0064	0.0052	0.0039	0.0024	0.0016
$\hat{\lambda}\vert(\lambda_0 = 0.7)$	0.7312	0.7282	0.7221	0.7164	0.7094	0.7010
MSE	0.0067	0.0058	0.0048	0.0034	0.0019	0.0012
$\hat{\theta}\vert(\theta_0 = 1.5)$	1.4731	1.4764	1.4816	1.4865	1.4921	1.4998
MSE	0.0075	0.0061	0.0042	0.0029	0.0015	0.0008

TABLE 10. Theoretical risk and empirical risk for complete case.

$\alpha_{i,1=1,2,3} \downarrow n \rightarrow$	25	50	130	350	500	1000
$\alpha_1 = 0.01$	0.0048	0.0053	0.0061	0.0076	0.0083	0.0091
$\alpha_2 = 0.05$	0.0274	0.0286	0.0312	0.0346	0.0389	0.0409
$\alpha_3 = 0.10$	0.0867	0.0894	0.0923	0.0963	0.0997	0.1013

7.2. **Simulating the M^2 under the CMLE method.** For simulating the M^2 under the uncensored maximum likelihood method, the data were simulated $N = 10,000$ times under the sample sizes $n_1 = 25, n_2 = 50, n_3 = 130, n_4 = 350, n_5 = 500, n_6 = 1000$. Using the BB algorithm and the R software, the MLEs and their mean square errors (MSEs) are calculated and presented in Table 11. For testing the null hypothesis H_0 according to which the data follows the BXEG distribution, we calculate the Y^2 statistical test, then it is compared with the different empirical levels of rejection of the null hypothesis H_0 when $M^2 > \chi^2_\alpha(r)$ where empirical levels of rejection of the null hypothesis are $\alpha_1 = 0.01$, $\alpha_2 = 0.05$ and $\alpha_3 = 0.10$. Table 12 gives the theoretical risk and empirical risk for complete case. The levels simulated for the statistic Y^2 agree with those corresponding to the theoretical levels of the chi-square distribution with (r) degrees of freedom, it is noticed, after accounting for simulation errors. In light of this, we can state that the test suggested in this study can appropriately adapt the data obtained from a BXEG model. For example:

(1) For α_1, it is noted that:
$$M^2_{0.01}(r) = (0.0062,\ 0.0068,\ 0.0076,\ 0.0087,\ 0.0098,\ 0.0106) < \chi^2_{0.01}(r).$$
So, we can accept the null hypothesis that the data follows the BXEG distribution.

(2) For α_2, it is seen that:

$M^2_{0.05}(r) = (0.0384,\ 0.0397,\ 0.0422, 0.0461,\ 0.0479,\ 0.0483) < \chi^2_{0.05}(r)$.
Therefore, we can accept the null hypothesis that the data follows the BXEG distribution.

(3) For α_3, it is observed that:
$M^2_{0.10}(r) = (0.0949,\ 0.0953,\ 0.0966,\ 0.0972,\ 0.0986,\ 0.0994) < \chi^2_{0.10}(r)$.
Hence, we can accept the null hypothesis that the data follows the BXEG distribution.

Table 11. MLEs and their MSEs for the censored case.

MLEs & MSEs$\downarrow$ $n \rightarrow$	25	50	130	350	500	1000
$\hat{a}\lvert(a_0 = 1.5)$	1.5224	1.5189	1.5143	1.5102	1.5078	1.5029
MSE	0.0088	0.0079	0.0058	0.0046	0.0034	0.0023
$\hat{\lambda}\lvert(\lambda_0 = 1.2)$	1.2233	1.2191	1.2111	1.2094	0.5067	0.5012
MSE	0.0079	0.0067	0.0049	0.0037	0.0028	0.0018
$\hat{\theta}\lvert(\theta_0 = 2.5)$	2.4719	2.4778	2.4836	2.4897	2.4921	2.4978
MSE	0.0081	0.0059	0.0043	0.0033	0.0023	0.0013

Table 12. Theoretical risk and empirical risk for censored case.

$\alpha_{i,1=1,2,3} \downarrow n \rightarrow$	$n_1 = 25$	$n_2 = 50$	$n_3 = 130$	$n_4 = 350$	$n_5 = 500$	$n_6 = 1000$
$\alpha_1 = 0.01$	0.0062	0.0068	0.0076	0.0087	0.0098	0.0106
$\alpha_2 = 0.05$	0.0384	0.0397	0.0422	0.0461	0.0479	0.0483
$\alpha_3 = 0.10$	0.0949	0.0953	0.0966	0.0972	0.0986	0.0994

8. **Data analysis**

Three examples from various fields are used to demonstrate the applicability of the proposed paradigm. We utilize M^2 to fit the first one's censored data from a survival analysis to predicted distributions. For the whole data scenario, Y^2 is built to see if the suggested model can accurately represent the two more occurrences.

8.1. **Real applications for uncensored data.**

8.1.1. *Times between failures for repairable items data.* The first data set is given by Murthy et al. [25]. These data have had a great deal of analysis and study, as many researchers and scholars have modeled and analyzed them and drawn many conclusions about them. The data refers to the time between failures for repairable items. The data are: 1.43, 0.11, 0.71, 0.77, 2.63, 1.49, 3.46, 2.46, 0.59, 0.74, 1.23, 0.94, 4.36, 0.40, 1.74, 4.73, 2.23, 0.45, 0.70, 1.06, 1.46, 0.30, 1.82, 2.37, 0.63, 1.23, 1.24, 1.97, 1.86, 1.17. Using R software and BB algorithms, we have $\hat{a} = 2.4196$, $\hat{\lambda} = 0.8956$ and $\hat{\theta} = 3.7482$. Then, taking for example 6 intervals, then $r = 6$, and calculate the Fisher information matrix (FIMx) on the initial data, we have

$$I(\underline{\mathbf{V}}) = \begin{pmatrix} 8.6352 & 5.0254 & 6.6235 \\ & 9.8455 & 7.6235 \\ & & 10.4632 \end{pmatrix}$$

Then, by calculating the NRR test statistic $Y^2_{0.05}(5)$ test statistics we have $Y^2_{0.05}(5) = 8.6247$. Since

$$Y^2_{0.05}(5) = 8.6247 < \chi^2_{0.05}(5) = 11.0705,$$

therefore, we can accept the null hypothesis that the times between failures for repairable items data follows the BXEG distribution.

8.1.2. *Reliability data.* The second data set is the reliability data which given by Cabarbaye and Faure [5]. These data have undergone extensive examination and investigation, as several researchers and academics have modelled, examined, and concluded on them. The data are: 0,313, 360, 231, 286, 340, 212, 287, 243, 170, 141, 150, 593, 328, 234, 206, 108, 134, 231, 218, 281, 192, 457, 269, 201, 181, 277, 479, 272, 223, 272, 163, 370, 217, 182, 202, 451, 303. Using R software and BB algorithms, we have $\hat{a} = 3.1526$, $\hat{\lambda} = 0.9432$ and $\hat{\theta} = 2.1635$. Then, taking for example 6 intervals, then $r = 6$, and calculate the FIMx on the initial data, we have

$$I(\underline{\mathbf{V}}) = \begin{pmatrix} 15.6325 & 23.5162 & 27.9523 \\ & 38.6243 & 31.2652 \\ & & 42.6153 \end{pmatrix}$$

Then, by calculating the NRR test statistic $Y^2_{0.05}(5)$ test statistics we have $Y^2_{0.05}(6) = 10.6781$. Since

$$Y^2_{0.05}(6) = 10.6781 < \chi^2_{0.05}(6) = 12.59159,$$

therefore, we can accept the null hypothesis that the reliability data follows the BXEG distribution.

8.1.3. *Strengths of glass fibers data.* The third data set is the strengths of glass fibers data which given by Smith and Naylor [31]. Additionally, as many researchers and academics have modelled, examined, and drawn several inferences from the strengths of fiber glass data, they have received considerable attention in statistical modelling. The data are: 1.014, 1.081, 1.082, 1.185, 1.223, 1.248, 1.267, 1.271, 1.272, 1.275, 1.276, 1.278, 1.286, 1.288, 1.292, 1.304, 1.306, 1.355, 1.361, 1.364, 1.379, 1.409, 1.426, 1.459, 1.46, 1.476, 1.481, 1.484, 1.501, 1.506, 1.524, 1.526, 1.535, 1.541, 1.568, 1.579, 1.581, 1.591, 1.593, 1.602, 1.666, 1.67, 1.684, 1.691, 1.704, 1.731, 1.735, 1.747, 1.748, 1.757, 1.800, 1.806, 1.867, 1.876, 1.878, 1.91, 1.916, 1.972, 2.012, 2.456, 2.592, 3.197, 4.121. Using R software and BB algorithms, we have $\hat{a} = 2.0631$, $\hat{\lambda} = 0.5326$ and $\hat{\theta} = 1.6352$. Then, taking for example 6 intervals, then $r = 7$, and calculate the FIMx on the initial data, we have

$$I(\underline{\mathbf{V}}) = \begin{pmatrix} 45.6325 & 23.6132 & 15.6236 \\ & 35.6214 & 16.5236 \\ & & 13.5264 \end{pmatrix}$$

Then, by calculating the NRR test statistic $Y^2_{0.05}(5)$ test statistics we have $Y^2_{0.05}(6) = 9.6325$. Since

$$Y^2_{0.05}(6) = 9.6325 < \chi^2_{0.05}(6) = 12.59159,$$

therefore, we can accept the null hypothesis that the glass fibers data follows the BXEG distribution.

8.2. **Real applications for censored data.**

8.2.1. *Times to infection of kidney dialysis patients.* Consider data of times to infection of kidney dialysis patients (see Klein and Moeschberger [23]). Infection times: 1.5, 3.5, 4.5, 4.5, 5.5, 8.5, 8.5, 9.5, 10.5, 11.5, 15.5, 16.5, 18.5, 23.5 26.5. Censored observations: 2.5, 2.5, 3.5, 3.5, 3.5, 4.5, 5.5, 6.5, 6.5, 7.5, 7.5, 7.5, 7.5, 8.5, 9.5, 10.5, 11.5, 12.5, 12.5, 13.5, 14.5, 14.5, 21.5, 21.5, 22.5, 22.5, 25.5, 27.5. Using R software and BB algorithms, we have $\hat{a} = 2.634$, $\hat{\lambda} = 0.7487$ and $\hat{\theta} = 1.415$. Table 5 gives the values of $\widehat{a}_j, e_{j,X}, O_{j,X}, \widehat{K}_{1j,\underline{\mathbf{V}}}, \widehat{K}_{2j,\underline{\mathbf{V}}}$ and $\widehat{K}_{3j,\underline{\mathbf{V}}}$ under $r = 5$. Table 13 gives the values of $\widehat{a}_j, e_{j,X}, O_{j,X}, \widehat{K}_{1j,\underline{\mathbf{V}}}, \widehat{K}_{2j,\underline{\mathbf{V}}}, \widehat{K}_{3j,\underline{\mathbf{V}}}$ for data of times to infection of kidney dialysis patients.

TABLE 13. Values of $\widehat{a}_j, e_{j,X}, O_{j,X}, \widehat{K}_{1j,\underline{\mathbf{V}}}, \widehat{K}_{2j,\underline{\mathbf{V}}}, \widehat{K}_{3j,\underline{\mathbf{V}}}$ for data of times to infection of kidney dialysis patients.

$\hat{a}_j$	5.2	8.1	12.3	17.5	27.5
$O_{j,X}$	10	8	9	7	9
$\widehat{K}_{1j,\underline{V}}$	1.1284	0.9317	1.2746	1.3742	0.9764
$\widehat{K}_{2j,\underline{V}}$	0.8463	0.9468	0.4869	0.3748	0.7417
$\widehat{K}_{3j,\underline{V}}$	0.7486	0.8974	0.4637	0.9713	0.8264
$e_{j,X}$	5.5134	5.5134	5.5134	5.5134	5.5134

Then, by calculating the modified NRR test statistic $M^2_{0.05}(5)$ test statistics we have $M^2_{0.05}\,(5) = 4.9845 = 7.940\,7$. Since

$$M^2_{0.05}(5) = 7.940\,7 < \chi^2_{0.05}(6) = 11.0705,$$

therefore, we can accept the null hypothesis that the data of times to infection of kidney dialysis patients follows the BXEG distribution.

8.2.2. *The bone marrow transplant data.* Consider data of times to infection of kidney dialysis patients for 38 patients (see Klein and Moeschberger [23]). Time to death: 1, 86, 107, 110, 122, 156, 162, 172, 243, 262, 262, 269, 276, 371, 417, 418, 466, 487, 526, 716, 781, 1111, 1182, 1199, 1279, 1377, 1433, 1496**.** Censored observations: 350, 1330, 194,226, 1167, 1462, 1602, 2081, 530, 996, 1330**.** Using R software and BB algorithms, we have $\hat{a} = 2.1635$, $\hat{\lambda} = 0.9635$ and $\hat{\theta} = 1.4362$. Table 6 gives the values of $\widehat{a}_j, e_{j,X}, O_{j,X}, \widehat{K}_{1j,\underline{\mathbf{V}}}, \widehat{K}_{2j,\underline{\mathbf{V}}}$ and $\widehat{K}_{3j,\underline{\mathbf{V}}}$ under $r = 4$. Table 14 gives the values of $\widehat{a}_j, e_{j,X}, O_{j,X}, \widehat{K}_{1j,\underline{\mathbf{V}}}, \widehat{K}_{2j,\underline{\mathbf{V}}}, \widehat{K}_{3j,\underline{\mathbf{V}}}$ for the bone marrow transplant data.

TABLE 14. Values of $\widehat{a}_j, e_{j,X}, O_{j,X}, \widehat{K}_{1j,\underline{\mathbf{V}}}, \widehat{K}_{2j,\underline{\mathbf{V}}}, \widehat{K}_{3j,\underline{\mathbf{V}}}$ for the bone marrow transplant data.

$\hat{a}_j$	260	480	1180	2081
$O_{j,X}$	11	9	8	10
$\widehat{K}_{1j,\underline{V}}$	2.3162	1.2356	0.9856	1.8564
$\widehat{K}_{2j,\underline{V}}$	1.5236	0.8456	0.7452	1.4266
$\widehat{K}_{3j,\underline{V}}$	3.1526	2.2513	1.7456	2.6351
$e_{j,X}$	4.6523	4.6523	4.6523	4.6523

Then, by calculating the modified NRR test statistic $M^2_{0.05}(4)$ test statistics we have $M^2_{0.05}\,(4) = 8.3265$. Since

$$M^2_{0.05}(4) = 8.3265 < \chi^2_{0.05}(4) = 9.4877,$$

therefore, we can accept the null hypothesis that the bone marrow transplant data follows the BXEG distribution.

8.2.3. *Strength of a certain type of braided cord data.* We apply the findings from this analysis to real data derived from reliable sources for the third data set. The forces of 48 pieces of cord that had withstood for a certain amount of time were examined as part of an experiment to learn more about the strength of a specific type of braided cord after the weather. Strength: 41.7, 43.9, 49.9, 50.1, 50.8,51.9, 52.1, 52.3, 52.3, 52.4, 52.6, 52.7, 53.1, 53.6, 53.6, 53.9, 53.9, 54.1, 54.6, 54.8, 54.8,55.1, 55.4, 55.9, 56, 56.1, 56.5, 56.9, 57.1, 57.1, 57.3, 57.7, 57.8, 58.1, 58.9, 59, 59.1, 59.6, 60.4, 60.7. Censored observations:

26.8, 29.6, 33.4, 35, 36.3, 40, 41.9, 42.5. Using R software and BB algorithms, we have $\hat{a} = 2.9365$, $\hat{\lambda} = 0.9635$ and $\hat{\theta} = 1.4362$. Table 7 gives the values of $\widehat{a}_j, e_{j,X}, O_{j,X}, \widehat{K}_{1j,\underline{\mathbf{V}}}, \widehat{K}_{2j,\underline{\mathbf{V}}}$ and $\widehat{K}_{3j,\underline{\mathbf{V}}}$ under $r = 4$. Table 15 list the values of $\widehat{a}_j, e_{j,X}, O_{j,X}, \widehat{K}_{1j,\underline{\mathbf{V}}}, \widehat{K}_{2j,\underline{\mathbf{V}}}, \widehat{K}_{3j,\underline{\mathbf{V}}}$ for strength of certain type of braided cord data.

Table 15. Values of $\widehat{a}_j, e_{j,X}, O_{j,X}, \widehat{K}_{1j,\underline{\mathbf{V}}}, \widehat{K}_{2j,\underline{\mathbf{V}}}, \widehat{K}_{3j,\underline{\mathbf{V}}}$ for strength of certain type of braided cord data.

$\hat{a}_j$	43	53	57	60.7
$O_{j,X}$	9	11	16	12
$\widehat{K}_{1j,\underline{V}}$	0.9356	1.5236	1.9642	1.4263
$\widehat{K}_{2j,\underline{V}}$	0.7152	2.0135	1.8236	0.9563
$\widehat{K}_{3j,\underline{V}}$	0.8623	1.9632	1.4326	2.2365
$e_{j,X}$	4.2153	4.2153	4.2153	4.2153

Then, by calculating the modified NRR test statistic $M^2_{0.05}(4)$ test statistics we have $M^2_{0.05}(4) = 6.9536$. Since

$$M^2_{0.05}(4) = 6.9536 < \chi^2_{0.05}(4) = 9.4877,$$

therefore, we can accept the null hypothesis that the strength of certain type of braided cord data follows the BXEG distribution.

9. Concluding remarks

A novel continuous probability distribution called the Burr X exponentiated gamma (BXEG) distribtion is introduced and studied in this work, but we will approach it from fresh angles that depart from those often covered in the literature. The BXEG model is basically derived based on the the Burr X family of Yousof elt al. (2017). In order to highlight more practical aspects in the areas of risk assessment and analysis, distributive verification, and its related practical applications on complete data and censored data, we chose to ignore many theoretical results and algebraic derivations, this is not to say that they are not important. However, by presenting and discussing some novel characterizations based on some related theories, such as characterizations based on the two truncated moments, characterizations in terms of the hazard function, and characterizations based on the conditional expectation of a function of the random variable, we were able to cover some theoretical aspects of the BXEG distribution. By analyzing a collection of commonly used financial indicators, such as the value-at-risk (VAR), tail-value-at-risk (TVAR), tail variance (TV), tail Mean-Variance (TMV), and mean excess loss (MEL) function, it is possible to analyse and evaluate the risks that insurance firms face. The maximum likelihood estimation approach, the ordinary least squares method, the weighted least squares estimation method, and the Anderson Darling estimation method are all described as estimate strategies for the major important risk indicators. These four methods were used and applied for the actuarial evaluation and a comparison is presented for determining the best method under a simulation study (for artificial assessment) and under an application to insurance claims data. The simulation is performed under three degrees of confidence, consideration of various sample sizes. With regard to the application to insurance claims data, the following results can be highlighted:

(1) For all risk assessment methods:

$$VaRq(X|_{1-q=30\%}) < ... < VaRq(X|_{1-q=1\%}), TVaRq(X|_{1-q=30\%}) < ... < TVaRq(X|_{1-q=1\%}),$$

$$TV(X|_{1-q=30\%}) > ... > TV(X|_{1-q=1\%}), TMVq(X|_{1-q=30\%}) > ... > TMVq(X|_{1-q=1\%}),$$

and

$$MEL(X|_{1-q=30\%}) > ... > MEL(X|_{1-q=1\%}).$$

(2) The VaRq(X) under the MLE method is monotonically increasing starts with 3673.32932 and ends with 6143.57495, the TVaRq(X) in monotonically incresing starts with 4715.02630 and ends with 6389.85619. However the TVq(X), the TMVq(X) and the MEL(X) are monotonically decreasing.

(3) The VaRq(X) under the OLSE method is monotonically increasing starts with 3672.18421 and ends with 6248.33526, the TVaRq(X) in monotonically incresing starts with 4743.90011 and ends with 6521.44231. However the TVq(X), the TMVq(X) and the MEL(X) are monotonically decreasing.

(4) The VaRq(X) under the WLSE method is monotonically increasing starts with 3598.25509 and ends with 6025.45856, the TVaRq(X) in monotonically incresing starts with 4614.29847 and ends with 6276.33491. However the TVq(X), the TMVq(X) and the MEL(X) are monotonically decreasing.

(5) The VaRq(X) under the AE method is monotonically increasing starts with 3619.36506 and ends with 6035.20922, the TVaRq(X) in monotonically incresing starts with 4630.7444 and ends with 6285.00127. However the TVq(X), the TMVq(X) and the MEL(X) are monotonically decreasing.

(6) Nearly for all q values, the OLSE method is recommended since it provides the most acceptable risk exposure analysis then the MLE method is recommended as a second one. However the other two methods are perform well.

In the framework of distributional validation and statistical hypothesis tests for the complete data, the well-known Nikulin-Rao-Robson statistic (Y^2), which is based on the uncensored maximum likelihood estimators on initial non-grouped data, is of considered under the BXEG model. The Y^2 statistic is assessed under a simulation study and under three real data sets as well and the following results can be highlighted:

- For the times between failures for repairable items data: $Y^2_{0.05}(5) = 8.6247 < \chi^2_{0.05}(5) = 11.0705$, therefore, we can accept the null hypothesis that the times between failures for repairable items data follows the BXEG distribution.
- For the reliability data: $Y^2_{0.05}(6) = 10.6781 < \chi^2_{0.05}(6) = 12.59159$, therefore, we can accept the null hypothesis that the reliability data follows the BXEG distribution.
- For the strengths data: $Y^2_{0.05}(6) = 10.6781 < \chi^2_{0.05}(6) = 12.59159$, therefore, we can accept the null hypothesis that the strengths data follows the BXEG distribution.

In the framework of distributional validation and statistical hypothesis tests for the censored data, a modified NRR statistic (M^2), which is based on the censored maximum likelihood estimators on initial non-grouped data, is of considered under the BXEG model. The M^2 statistic is assessed under comprehensive simulation study and under three real data sets and the following results can be highlighted:

- For the times to infection of kidney dialysis patients data, $M^2_{0.05}(5) = 7.9407 < \chi^2_{0.05}(6) = 11.0705$, therefore, we can accept the null hypothesis that the data of times to infection of kidney dialysis patients follows the BXEG distribution.
- For the bone marrow transplant data, $M^2_{0.05}(4) = 8.3265 < \chi^2_{0.05}(4) = 9.4877$, therefore, we can accept the null hypothesis that the bone marrow transplant data follows the BXEG distribution.
- For the strength of certain type of braided cord data, $M^2_{0.05}(4) = 6.9536 < \chi^2_{0.05}(4) = 9.4877$, therefore, we can accept the null hypothesis that thestrength of certain type of braided cord data follows the BXEG distribution.

10. Author Contributions

G. Hamedani: original draft preparation, conceptualization, methodology, supervision. Ahmad Aboalkhair: writing, review and editing. Khaoula Aidi: resources, validation, software, formal analysis, data curation. Ali Hadi: original draft preparation, supervision, writing, review and editing. Haitham Yousof: original draft preparation, methodology, project administration, data curation. Mohamed Ibrahim: original draft preparation, validation, software, data curation.

11. Data Availability Statement

The data sets are provided in the paper.

12. Aknowledgement

We would like to thank the reviewers for their thoughtful comments and efforts towards improving our manuscript.

13. Ethical considerations

The authors avoided from data fabrication and falsification.

14. Funding

No funding received.

15. Conflict of interest

The authors declare no conflict of interest.

References

[1] Abouelmagd, T. H. M., Hamed, M. S., Hamedani, G. G., Ali, M. M., Goual, H., Korkmaz, M. C. and Yousof, H. M. (2019a). The zero truncated Poisson Burr X family of distributions with properties, characterizations, applications, and validation test. Journal of Nonlinear Sciences and Applications (JNSA), 12(5), 314-336.

[2] Aidi, K., Butt, N. S., Ali, M. M., Ibrahim, M., Yousof, H. M. and Shehata, W. A. M. (2021). A Modified Chi-square Type Test Statistic for the Double Burr X Model with Applications to Right Censored Medical and Reliability Data. Pakistan Journal of Statistics and Operation Research, 17(3), 615-623.

[3] Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent measures of risk. Mathematical finance, 9(3), 203-228.

[4] Bagdonavicius, V., & Nikulin, M. (2011). Chi-squared goodness-of-fit test for right censored data. International Journal of Applied Mathematics and Statistics, 24(1), 1-11.

[5] Cabarbaye, A., Faure, J., Laulheret, R., & Cabarbaye, A. (2009). Interest of a global optimization tool for reliability models adjustment and systems optimization. In Reliability, Risk, and Safety, Three Volume Set (pp. 617-624). CRC Press.

[6] Castillo, E. and Hadi, A. S. (1995). A Method for Estimating Parameters and Quantiles of Continuous Distributions of Random Variables, Computational Statistics and Data Analysis, 20, 421–439.

[7] Emam, W.; Tashkandy, Y.; Hamedani, G.G.; Shehab, M.A.; Ibrahim, M.; Yousof, H.M. (2023). A Novel Discrete Generator with Modeling Engineering, Agricultural and Medical Count and Zero-Inflated Real Data with Bayesian, and Non-Bayesian Inference. Mathematics 2023, 11, 1125. https://doi.org/10.3390/math11051125

[8] 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 2023, 11, 897. https://doi.org/10.3390/math11040897

[9] Goual, H. and 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.

[10] Goual, H., Yousof, H. M. and 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 & Applications (JNSA), 13(6), 330-353.

[11] Goual, H., Yousof, H. M. and 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.

[12] Gupta, R. C., Gupta, P. L. and Gupta, R. D. (1998). Modeling failure time data by Lehman alternatives. Communications in Statistics-Theory and methods, 27(4), 887-904.

[13] Glänzel, W. (1987). A characterization theorem based on truncated moments and its application to some distribution families, Mathematical Statistics and Probability Theory (Bad Tatzmannsdorf, 1986), Vol. B, Reidel, Dordrecht, 1987, 75–84.

[14] Glänzel, W. (1990). Some consequences of a characterization theorem based on truncated moments, *Statistics:* A Journal of Theoretical and Applied Statistics, 21(4), 613–618.

[15] Hamedani, G.G. (2013). On certain generalized gamma convolution distributions **II**, Technical Report, No. 484, MSCS, Marquette University, 2013.

[16] 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

[17] Ibrahim, M., Aidi, K., Ali, M. M. and Yousof, H. M. (2023). A Novel Test Statistic for Right Censored Validity under a new Chen extension with Applications in Reliability and Medicine. Annals of Data Science, 10(5), 1285–1299.

[18] 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.

[19] Ibrahim, M.; Emam, W.; Tashkandy, Y.; Ali, M.M.; Yousof, H.M. 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

[20] Ibrahim, M., Yadav, A. S., Yousof, H. M., Goual, H. and 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.

[21] Khedr, A. M., Nofal, Z. M., El Gebaly, Y. M., & Yousof, H. M. A Novel Family of Compound Probability Distributions: Properties, Copulas, Risk Analysis and Assessment under a Reinsurance Revenues Data Set. Thailand Statistician, forthcoming.

[22] Khalil, M. G., Yousof, H. M., Aidi, K., Ali, M. M., Butt, N. S. and Ibrahim. M. (2023). Modified Bagdonavicius-Nikulin Goodness-of-fit Test Statistic for the Compound Topp Leone Burr XII Model with Various Censored Applications. Statistics, Optimization & Information Computing, forthcoming.

[23] Klein, J. P. and Moeschberger, M. L. (2003). Survival analysis: techniques for censored and truncated data (Vol. 1230). New York: Springer.

[24] 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.

[25] Murthy, D. P., Xie, M., & Jiang, R. (2004). Weibull models. John Wiley & Sons.

[26] Nikulin, M. S., (1973a). Chi-square test for normality. In: Proceedings of the International Vilnius Conference on Probability Theory and Mathematical Statistics, Vol. 2, 199-122.

[27] Nikulin, M. S., (1973b), chi-squared test for continuous distribution with shift and scale parameters, Theory of Probability and its Applications, 18(3), 559-568.

[28] Nikulin. M.S, (1973c). On a Chi-squared test for continuous distributions, Theory of Probability and its Applications. 19, 638-639.

[29] 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.

[30] Ravi, V. and 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, J. Statist. Software, 32(4), 1-26.

[31] Smith, R. L. and Naylor, J. (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. Journal of the Royal Statistical Society: Series C (Applied Statistics), 36(3), 358-369.

[32] Voinov, V., Nikulin, M., and Balakrishnan, N. (2013), chi-squared Goodness of Fit Tests with Applications, Academic Press, Elsevier.

[33] Yadav, A. S., Goual, H., Alotaibi, R. M., Rezk, H., Ali, M. M. and 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.

[34] 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.

[35] Yousof, H. M., Ali, M. M., Aidi, K., Ibrahim, M. (2023a). 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.

[36] 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.

[37] Yousof, H. M., Afify, A. Z., Hamedani, G. G. and Aryal, G. (2017). The Burr X generator of distributions for lifetime data. Journal of Statistical Theory and Applications, 16(3), 288-305.

[38] Yousof, H. M., Aidi, K., Hamedani, G. G., & Ibrahim, M. (2023b). Chi-squared type test for distributional censored and uncensored validity with numerical assessments and real data applications. Japanese Journal of Statistics and Data Science, 6(2), 729-758.

G.G. Hamedani
Orcid number: 0000-0001-7976-1088
Department of Mathematical and Statistical Sciences
Marquette University
USA
Email address: gholamhoss.hamedani@marquette.edu

Ahmad M. Aboalkhair[1,2]
Orcid number: 0000-0002-6782-6961
[1] Department of Quantitative Methods, School of Business
King Faisal University
Al-Ahsa 31982, Saudi Arabia
[2] Department of Applied Statistics and Insurance, Faculty of Commerce
Mansoura University, Mansoura 35516
Mansoura, Egypt
Email address: aaboalkhair@kfu.edu.sa

Khaoula Aidi
Orcid number: 0000-0002-4756-3690
Laboratory of probability and statistics LaPS
University Badji Mokhtar
Annaba, Algeria
Email address: khaoula.aidi@yahoo.fr

Ali S. Hadi[1,2]
Orcid number: 0000-0002-1533-8333
[1] Department of Statistics
American University
Cairo, Egypt
[2] Department of Statistics
Cornell University
Ithaca, NY, USA
Email address: ahadi@aucegypt.edu

Haitham M. Yousof
Orcid number: 0000-0003-4589-4944
Department of Statistics, Mathematics and Insurance,
Benha University
Benha, Egypt
Email address: haitham.yousof@fcom.bu.edu.eg

Mohamed Ibrahim[1,2]
Orcid number: 0000-0003-4893-9669
[1] Department of Quantitative Methods, School of Business
King Faisal University
Al-Ahsa 31982, Saudi Arabia
[2] Department of Applied, Mathematical and Actuarial Statistics
Faculty of Commerce
Damietta University
Damietta, Egypt
Email address: miahmed@kfu.edu.sa