

A Modified Chi-square Type Test Statistic for the Double Burr X Model with Applications to Right Censored Medical and Reliability Data



Khaoula Aidi¹, Nadeem S. Butt², M. Masoom Ali³,
Mohamed Ibrahim⁴, Haitham M. Yousof⁵ and Wahid A. M. Shehata⁶

* Corresponding Author

¹Laboratory of probability and statistics LaPS, University Badji Mokhtar, Annaba, Algeria, khaoula.aidi@yahoo.fr

²Department of Family and Community Medicine, Faculty of Medicine Rabigh, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia, nshafique@kau.edu.sa

³Department of Mathematical Sciences, Ball State University, Muncie, Indiana 47306 USA, mali@bsu.edu

⁴Department of Applied, Mathematical and Actuarial Statistics, Faculty of Commerce, Damietta University, Damietta, Egypt, mohamed_ibrahim@du.edu.eg

⁵Department of Statistics, Mathematics and Insurance, Benha University, Benha 13518, Egypt, haitham.yousof@fcom.bu.edu.eg

⁶Department of Mathematics, Statistics and Insurance, Faculty of Business, Ain Shams University, Egypt, wahid75maher@yahoo.com

Abstract

A new modified version of the Bagdonavičius-Nikulín goodness-of-fit test statistic is presented for validity for the right censor case under the double Burr type X distribution. The maximum likelihood estimation method in censored data case is used and applied. Simulations via the algorithm of Barzilai-Borwein is performed for assessing the right censored estimation method. Another simulation study is presented for testing the null hypothesis under the modified version of the Bagdonavičius and Nikulín goodness-of-fit statistical test. Four right censored data sets are analyzed under the new modified test statistic for checking the distributional validation.

Key Words: Right Censored Validation; Barzilai-Borwein; Bagdonavičius-Nikulín; Double Burr type X; Goodness-of-fit Test; Simulation.

1. Introduction

Statistical methods for testing the validity of a parametric distribution are in increasing developments. Clearly, the presence of censored observations makes those methods unsuitable. For treating these problems there were many attempts. First, Habib and Thomas (1986) and Hollander and Pena (1992) proposed a modified version of the Chi-squared test for the randomly censored real data based on the well-known estimators of the Kaplan and Meyer. Then, Galanova (2012) considered some nonparametric modifications for the Anderson-Darling (AD) test statistic, Kolmogorov-Smirnov (KS) test statistic and the Cramer-Von-Mises (KM) test statistic for the accelerate failure models. Recently, Bagdonavičius and Nikulín (2011a,b) presented and applied a new type of Chi-squared goodness-of-fit test statistic for the censored data (right case), (see Bagdonavičius et al. (2013)). Then, the new type of Chi-squared goodness-of-fit test statistic of Bagdonavičius-Nikulín is applied for the distributional validation under the right censored schemes. Recently, the Bagdonavičius-Nikulín test is used for the distributional validation in case of the right censored data (see Mansour et al (2020a,b), Ibrahim et al. (2021) and Yousof et al. (2021a,b)).

In this work, a new modified type Chi-squared goodness-of-fit test statistic which is established based on the well-known Bagdonavičius-Nikulín test is presented and applied accordingly for distributional validation under the double

Burr type X (DBX) distribution model using the right censored case. A comprehensive simulation study under the right censored case via the well-known Barzilai-Borwein (BB) algorithm is performed for assessing and evaluating the right censored estimation method. Following Ravi and Gilbert (2009), we generated 10,000 samples with different size ($n = 15, 25, 50, 130, 350, 500, 1000$) from the DBX model using some very carefully selected initial values. The mean square errors (MSEs) are used for assessing the performance of the censored maximum likelihood. The new modified Bagdonavičius-Nikulin type test is applied using four right censored real data sets for distributional validity.

A random variable (RV) T is said to have the Burr type X (BX) distribution if its cumulative distribution function (CDF) is given by

$$F_{\theta_2}(t) = [1 - \exp(-t^2)]^{\theta_2} |_{t>0, \theta_2>0} \tag{1}$$

Yousof et al. (2017) proposed a new flexible family called the Burr type X family. Following Yousof et al. (2017), we construct the double Burr type X (DBX) model with CDF given by

$$F(x, \underline{\mathbb{V}}) = \left[1 - \exp\left(-\{[1 - \exp(-x^2)]^{-\theta_2} - 1\}^{-2}\right) \right]^{\theta_1}, \tag{2}$$

where $\underline{\mathbb{V}} = (\theta_1, \theta_2)$. The probability density function (PDF) of the DBX corresponding to (2) reduces to

$$f(x, \underline{\mathbb{V}}) = 4\theta_1\theta_2x \exp(-x^2) \exp\left(-\{[1 - \exp(-x^2)]^{-\theta_2} - 1\}^{-2}\right) \times \frac{[1 - \exp(-x^2)]^{2\theta_2 - 1}}{\{1 - [1 - \exp(-x^2)]^{\theta_2}\}^3} \left[1 - \exp\left(-\{[1 - \exp(-x^2)]^{-\theta_2} - 1\}^{-2}\right) \right]^{\theta_1 - 1}. \tag{3}$$

A RV X having PDF in (3) is denoted by $X \sim \text{DBX}(\underline{\mathbb{V}})$. The hazard rate function (HRF) and cumulative hazard rate function (CHRF) can be derived from the well known relationship. The new additional parameter θ_1 will allow us to study the tail behavior of the new density (3) with greater flexibility. Depending Yousof et al. (2017), many author presented and studied new lifetime distributions (see Jahanshahi et al. (2019) for the Burr X Fréchet distribution, Khalil et al. (2019) for the Burr X exponentiated Weibull distribution, Abouelmagd et al. (2019) for the Poisson Burr X Weibull distribution, Elsayed and Yousof (2019) for the Burr X Nadarajah Haghghi distribution, Ali et al. (2021a) for odd Burr Burr X distribution and Ali et al. (2021b) for the Marshall-Olkin Lehmann Burr X distribution).

2. Maximum likelihood estimation in censored data case

Suppose that X_1, X_2, \dots, X_n is a random sample with right censoring from the DBX distribution. The observed data $x_i = \min(X_i, C_i); i = 1, 2, \dots, n$ are the minimum of the survival time X_i and censoring time C_i for each subject in the sample. So, x_i can be written in the form $(x_i, \nabla_i)_{i=1, \dots, n}$ where, $\nabla_i = 1$ if X_i is the moment of failure (complete observation) and $\nabla_i = 0$ if X_i is the moment of censoring. The right censoring is assumed to be non informative, so the expression of the likelihood function is

$$l(x, \underline{\mathbb{V}}) = \prod_{i=1}^n f(x_i, \underline{\mathbb{V}})^{\nabla_i} S(x_i, \underline{\mathbb{V}})^{1 - \nabla_i} |_{\nabla_i=1, x_i < C_i}$$

The log-likelihood function of the DBX distribution is

$$L_n(\underline{\mathbb{V}}) = \sum_{i=1}^n \nabla_i \ln f(x_i, \underline{\mathbb{V}}) + \sum_{i=1}^n (1 - \nabla_i) \ln S(x_i, \underline{\mathbb{V}}) |_{\nabla_i=1, x_i < C_i}$$

where $S(x_i, \underline{\mathbb{V}})$ refers to the survival function of the DBX model, then

$$L_n(\underline{\mathbb{V}}) = \sum_{i=1}^n \nabla_i \left[\ln(4\theta_1\theta_2) + \ln(x_i) - x_i^2 + (2\theta_2 - 1)\ln(s_i) \right. \\ \left. - 3\ln(1 - s_i^{\theta_2}) - \varpi_i^{-2} + (\theta_1 - 1)\ln(1 - \exp(-\varpi_i^{-2})) \right] \\ + \sum_{i=1}^n (1 - \nabla_i) \ln \left[1 - (1 - \exp(-\varpi_i^{-2}))^{\theta_1} \right],$$

and the score functions are obtained as follows

$$\frac{\partial L_n(\mathbb{V})}{\partial \theta_1} = \sum_{i=1}^n \nabla_i \left[\frac{1}{\theta_1} + \ln(1 - \exp(-\varpi_i^{-2})) \right] - \sum_{i=1}^n \frac{(1 - \exp(-\varpi_i^{-2}))^{\theta_1} \ln(1 - \exp(-\varpi_i^{-2}))}{1 - (1 - \exp(-\varpi_i^{-2}))^{\theta_1}},$$

and

$$\frac{\partial L_n(\mathbb{V})}{\partial \theta_2} = \sum_{i=1}^n \nabla_i \left[\frac{1}{\theta_2} + 2\ln(s_i) + \frac{3s_i^{\theta_2} \ln s_i}{1 - s_i^{\theta_2}} - 2s_i^{\theta_2} \ln s_i \varpi_i^{-3} \right] - \sum_{i=1}^n \frac{2\theta_1 s_i^{\theta_2} \ln s_i \varpi_i^{-3} \varphi_i (1 - \varphi_i)^{\theta_1 - 1}}{1 - \varphi_i},$$

where $s_i = 1 - \exp(-x_i^2)$, $\varpi_i = s_i^{\theta_2} - 1$ and $\varphi_i = \exp(-\varpi_i^{-2})$. The maximum likelihood estimates of the unknown parameters can be obtained using various techniques, either software R, the EM algorithm or Newton Raphson method.

3. Modified chi-squared type test for right censored data

Methods for testing the validity of parametric models are in constant development, but the presence of censorship make them unsuitable. Habib and Thomas (1986), Hollander and Pena (1992) proposed modified chi-squared test based on Kaplan-Meyer estimators, Galanova et al. (2012) considered modifications of Kolmogorov-Smirnov statistic, Anderson-Darling statistic, Cramer Von-Mises statistic for accelerate failure models. In this work, we are interested in the modified chi-squared type test proposed by Bagdonavičius and Nikulin (2011a,b), Bagdonavičius et al.(2013), for parametric models with right censored data. Based on maximum likelihood estimates on non-grouped data, this test statistic is also based on the differences between the numbers of observed failures and the numbers of expected failures in the grouped intervals chosen. For this, random grouping intervals are considered as data functions. The description of the construction of this chi-squared type test is developed in Voinov *et al.*(2013). The test statistic is defined as follows. Suppose that X_1, X_2, \dots, X_n is a random sample with right censoring from a parametric model, and a finite time τ . The test statistic is defined as follows.

$$T_{r,\varepsilon}^2 = \sum_{j=1}^n \frac{1}{u(j, n)} (u(j, n) - e(j, n))^2 + QF$$

where $u(j, n)$ and $e(j, n)$ are the observed and the expected numbers of failure in grouping intervals, and QF is

$$QF = \mathbf{w}^T \hat{G}^{-1} \mathbf{w}, \hat{G} = [\hat{g}_{ll'}]_{s \times s} |_{l, l' = 1, \dots, s},$$

$$\hat{\mathbf{w}} = \hat{C} \hat{A}^{-1} Z = (\hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_s)^T,$$

$$Z_j = \frac{1}{\sqrt{n}} (u(j, n) - e(j, n)) |_{j=1, 2, \dots, r},$$

$$\mathbf{w}_l = \sum_{j=1}^r \hat{C}_{lj}(n) \hat{A}_j^{-1} Z_j |_{j=1, 2, \dots, r},$$

and

$$\hat{g}_{ll'} = \hat{t}_{ll'} - \sum_{j=1}^r \hat{C}_{lj}(n) \hat{C}_{l'j}(n) \hat{A}_j^{-1} |_{j=1, 2, \dots, r \text{ and } l, l' = 1, \dots, s}.$$

The limits $a(j, n, i)$ of r random grouping intervals $I_j = [(j - 1, n, i), (j, n, i)[$ are chosen such that the expected failure times falling into these intervals are the same for each $j = 1, \dots, r - 1$, $\hat{a}_r = \max(x_{(l)}, \tau)$. The estimated $\hat{a}_{\mathbb{V}}(j, n, i)$ is defined by

$$\hat{\alpha}_{\underline{V}}(j, n, i) = H^{-1} \left\{ \frac{1}{n-i+1} \left[E_j - \sum_{l=1}^{i-1} H(x_l, \underline{V}) \right], \underline{V} \right\},$$

where

$$\hat{\alpha}_r = \max(x_{(n)}, \tau),$$

and $H(x_l, \underline{V})$ is the CHRF of the DBX model distribution. This test statistic $\mathcal{T}_{r,\varepsilon}^2$ follows a chi-squared distribution.

3.1 Choice of random grouping intervals

Suppose that X_1, X_2, \dots, X_n is a random sample with right censoring from the double Burr type X (DBX) distribution and a finite time τ . In our case, the estimated $\hat{\alpha}_{\underline{V}}(j, n, i)$ is obtained as follows

$$\hat{\alpha}_{\underline{V}}(j, n, i) = \pm \sqrt{-\ln \left\{ 1 - [\mathcal{Q}(x_l, \underline{V}) + 1]^{-\frac{1}{\theta_2}} \right\}},$$

where $\underline{V} = (\hat{\theta}_1, \hat{\theta}_2)^T$ are the maximum likelihood estimates of the unknown parameters $\underline{V} = (\theta_1, \theta_2)^T$ on initial data, and

$$\mathcal{Q}(x_l, \underline{V}) = \left[-\ln \left(1 - \left\{ \frac{1}{n-i+1} [E_j - \sum_{l=1}^{i-1} H(x_l, \underline{V})] \right\}^{\frac{1}{\theta_1}} \right) \right]^{-\frac{1}{2}}.$$

3.2 Quadratic form of the statistic $\mathcal{T}_{r,\varepsilon}^2$

To calculate the quadratic form (QF) of the statistic $\mathcal{T}_{r,\varepsilon}^2$, and as its distribution does'nt depend on the parameters, so we can use the estimated matrices $\hat{\mathcal{W}}$, \hat{C} and the estimated information matrix \hat{I} . The elements of \hat{C} defined by

$$\hat{C}_{lj}(n) = \frac{1}{n} \sum_{i: x_i \in I_j} \nabla_i \frac{\partial}{\partial \underline{V}_l} \ln h(x_i; \underline{V}),$$

are obtained as below

$$\hat{C}_{1j}(n) = \frac{1}{n} \sum_{i: x_i \in I_j} \nabla_i \left[\frac{1}{\theta_1} + \frac{\ln(1 - \exp(-\varpi_i^{-2}))}{1 - (1 - \exp(-\varpi_i^{-2}))^{\theta_1}} \right],$$

and

$$\hat{C}_{2j}(n) = \frac{1}{n} \sum_{i: x_i \in I_j} \nabla_i \left[\begin{array}{l} \frac{1}{\theta_2} + 2\ln(s_i) + \frac{3s_i^{\theta_2} \ln s_i}{1 - s_i^{\theta_2}} - 2s_i^{\theta_2} \ln s_i \varpi_i^{-3} \\ - \frac{2(\theta_1 - 1)s_i^{\theta_2} \ln s_i \varpi_i^{-3} \varphi_i}{1 - \varphi_i} + \frac{2\theta_1 s_i^{\theta_2} \ln s_i \varpi_i^{-3} \varphi_i (1 - \varphi_i)^{\theta_1 - 1}}{1 - \varphi_i} \end{array} \right].$$

Therefore, the estimated matrix $\hat{\mathcal{W}}$ can be deduced from \hat{C} .

3.3 Estimated information matrix $\hat{I}(n)$

We need also the estimated information matrix $\hat{I}(n)$ of the DBX distribution with right censoring. After some calculations and some simplifications, we have obtained the elements of the matrix as follows:

$$\hat{i}_{11}(n) = \frac{1}{n} \sum_{i=1}^n \nabla_i \left(\frac{1}{\theta_1} + \frac{\ln(1 - \exp(-\varpi_i^{-2}))}{1 - (1 - \exp(-\varpi_i^{-2}))^{\theta_1}} \right)^2,$$

$$\hat{i}_{22}(n) = \frac{1}{n} \sum_{i=1}^n \nabla_i \left(\frac{\frac{1}{\theta_2} + 2\ln(s_i) + \frac{3s_i^{\theta_2} \ln s_i}{1 - s_i^{\theta_2}} - 2s_i^{\theta_2} \ln s_i \varpi_i^{-3}}{-\frac{2(\theta_1 - 1)s_i^{\theta_2} \ln s_i \varpi_i^{-3} \varphi_i}{1 - \varphi_i} + \frac{2\theta_1 s_i^{\theta_2} \ln s_i \varpi_i^{-3} \varphi_i (1 - \varphi_i)^{\theta_1 - 1}}{1 - \varphi_i}} \right)^2,$$

and

$$\begin{aligned} \hat{i}_{12}(n) &= \frac{1}{n} \sum_{i=1}^n \nabla_i \left(\frac{1}{\theta_1} + \frac{\ln(1 - \exp(-\varpi_i^{-2}))}{1 - (1 - \exp(-\varpi_i^{-2}))^{\theta_1}} \right) \\ &\times \left(\frac{\frac{1}{\theta_2} + 2\ln(s_i) + \frac{3s_i^{\theta_2} \ln s_i}{1 - s_i^{\theta_2}} - 2s_i^{\theta_2} \ln s_i \varpi_i^{-3}}{-\frac{2(\theta_1 - 1)s_i^{\theta_2} \ln s_i \varpi_i^{-3} \varphi_i}{1 - \varphi_i} + \frac{2\theta_1 s_i^{\theta_2} \ln s_i \varpi_i^{-3} \varphi_i (1 - \varphi_i)^{\theta_1 - 1}}{1 - \varphi_i}} \right). \end{aligned}$$

As all the components of the statistic are given explicitly, then we obtain the test statistic for the DBX distribution with unknown parameters and right censored data. This statistic follows a chi-squared distribution with r degrees of freedom.

$$\mathcal{J}_{r,\varepsilon}^2(\hat{\mathbb{V}}) = \sum_{j=1}^r \frac{1}{u(j, n)} [u(j, n) - e(j, n)]^2 + \hat{\mathbf{w}}^T \left[\hat{i}_{ll'} - \sum_{j=1}^r \hat{C}_{ij} \hat{C}_{l'j} \hat{A}_j^{-1} \right]^{-1} \hat{\mathbf{w}}.$$

4. Simulations

An important simulation study is carried out to show the performance of the techniques used and the feasibility of the goodness-of-fit test developed in this work. At this end, we generated $N = 10,000$ right censored samples with different sizes $n_1 = 15, n_2 = 25, n_3 = 50, n_4 = 130, n_5 = 350, n_6 = 500, n_7 = 1000$ from the DBX model with parameter values $(\theta_1 = 2, \theta_2 = 1.2)$. Firstly, we compute the MLEs of the unknown parameters and then the criteria $\mathcal{J}_{r,\varepsilon}^2$ of the corresponding samples are provided.

4.1 Maximum likelihood estimation for DBX model

Using statistical software *R* and the BB algorithm, we calculate the maximum likelihood estimates of the unknown parameters and the corresponding MSEs (see Ravi (2009) for more details). The results of simulating the $\mathcal{J}_{r,\varepsilon}^2$ test under the BB algorithm and the DBX model are given in Table 1.

Table 1: Simulating the $\mathcal{J}_{r,\varepsilon}^2$ test under the BB algorithm.

$n \rightarrow$	n_1	n_2	n_3	n_4	n_5	n_6	n_7
<i>AVs&MSE</i> ↓	15	25	50	130	350	500	1000
$\hat{\theta}_1$	1.8763	1.8969	1.9236	1.9424	1.9759	1.9826	1.9952
<i>MSE</i>	0.0157	0.0139	0.0111	0.0079	0.0068	0.0042	0.0023
$\hat{\theta}_2$	1.346	1.3162	1.2863	1.2452	1.2331	1.2283	1.2093
<i>MSE</i>	0.0196	0.0182	0.0169	0.0134	0.0092	0.0078	0.0048

4.2 Testing the null hypothesis under the test statistic $\mathcal{J}_{r,\varepsilon}^2$

For testing the null hypothesis H_0 that right censored data arose from the DBX model, we compute the criteria statistic $\mathcal{J}_{r,\varepsilon}^2(\hat{\mathbb{V}})$ as defined above for 10,000 simulated samples from the hypothesized distribution with different sizes ($n = 15, 25, 50, 130, 350, 500, 1000$). Then, we calculate empirical levels of significance, when $\mathcal{J}_{r,\varepsilon}^2 > \chi_{r,\varepsilon}^2$, corresponding to theoretical levels of significance ($\varepsilon = 1\%, 5\%, 10\%$). We choose $r = 5$. The results are reported in Table 2.

Table 2: Simulated levels of significance for $\mathcal{J}_{r,\epsilon}^2(\hat{\mathbb{V}})$ test for DBX model

$n \rightarrow$	n_1	n_2	n_3	n_4	n_5	n_6	n_7
$\epsilon \downarrow$	15	25	50	130	350	500	1000
$\epsilon = 1\%$	0.0042	0.0045	0.0055	0.0067	0.008	0.0089	0.0094
$\epsilon = 5\%$	0.0446	0.0457	0.0462	0.0475	0.0483	0.0491	0.0503
$\epsilon = 10\%$	0.0921	0.0933	0.0944	0.0959	0.0964	0.0978	0.0989

The null hypothesis H_0 for which simulated samples are fitted by DBX distribution, is widely validated for the different levels of significance. Therefore, the test proposed in this work, can be used to fit data from this new distribution.

5. Applications

In this section, we introduce four applications of the DBX distribution to four real-life right censored data sets. The first right censored data set is from a study conducted on the effects of ploidy on the prognosis of patients with cancers of the mouth. Patients were selected who had a paraffin-embedded sample of the cancerous tissue taken at the time of surgery. Follow-up survival data was obtained on each patient. The tissue samples were examined using a flow cytometer to determine if the tumor had an aneuploidy (abnormal) or diploid (normal) DNA profile using a technique discussed in Sickle-Santanello et al. (1988). The data below relates to patients with tongue cancer. The death times in weeks under aneuploid tumors are: (1, 3, 3, 4, 10, 13, 13, 16, 16, 24, 26, 27, 28, 30, 30, 32, 41, 51, 65, 67, 70, 72, 73, 77, 91, 93, 96, 100, 104, 157, 167, 61*, 74*, 79*, 80*, 81*, 87*, 87*, 88*, 89*, 93*, 97*, 101*, 104*, 108*, 109*, 120*, 131*, 150*, 231*, 240*, 400*).

The second right censored data set we consider is the bone marrow transplant data (Klein and Moeschberger (2003)) for patients suffering from acute lymphoblastic leukemia. This data consist of time (in days) to death or on study time after an allogenic bone marrow transplant for 38 patients. The bone marrow transplant is a standard treatment for acute leukemia. Recovery following bone marrow transplant is a complex process. Immediately following transplantation, patients have depressed platelet counts and have higher hazard rate for the development of infections but as the time passes the hazard decreases. Data are given as (1, 86, 107, 110, 122, 156, 162, 172, 194, 226*, 243, 262, 262, 269, 276, 350*, 371, 417, 418, 466, 487, 526, 530*, 716, 781, 996*, 1111, 1167*, 1182, 1199, 1279, 1330*, 1377, 1433, 1462*, 1496, 1602*, 2081*).

The third right censored data set we apply are the results obtained from Crowder et al. (1991). In an experiment to obtain information on the strength of a certain type of braided cord after the weather the forces of 48 pieces of cord having resisted for a determined time were studied. The right censored force values observed are given as follows: (26.8*, 29.6*, 33.4*, 35*, 36.3, 40*, 41.7, 41.9*, 42.5*, 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).

Fourth right censored data set we consider are times to infection of kidney dialysis patients (see Nahman et al. (1992)). The times to infection data are: (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, 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*).

For the all right censored data, we use the Bagdonavičius-Nikulin test statistic provided above to verify if these data are modeled by the DBX distribution. For the first right censored data, we calculate the maximum likelihood estimates of the unknown parameters

$$\hat{\mathbb{V}} = (\hat{\theta}_1, \hat{\theta}_2)^T = (2.566, 1.346)^T,$$

where the data are grouped into $r = 5$ intervals. We give the necessary calculations of the $\hat{a}_{\underline{v}}(j, n, i)$, $e(j, n)$, $u(j, n)$, $\hat{C}_{1j}(n)$ and $\hat{C}_{2j}(n)$ for the first right censored data in the following Table 3.

Table 3: Values of $\hat{a}_{\underline{v}}(j, n, i), e(j, n), u(j, n), \hat{C}_{1j}(n), \hat{C}_{2j}(n)$

$\hat{a}_{\underline{v}}(j, n, i)$	29.654	71.968	90.536	129.25	400
$u(j, n)$	12	9	10	12	7
$\hat{C}_{1j}(n)$	1.1963	0.9865	1.1236	1.3621	0.7469
$\hat{C}_{2j}(n)$	1.3642	0.6186	0.8263	1.2233	0.5632
$e(j, n)$	4.1896	4.1896	4.1896	4.1896	4.1896

Then we obtain the value of the test statistic $\mathcal{J}_{r,\varepsilon}^2: \mathcal{J}_{5,0.05}^2 = \chi_{5,0.05}^2 + QF = 5.3926 + 2.8691 = 8.2617$. For significance level $\varepsilon = 0.05$, the critical value $\chi_{5,0.05}^2 = 11.0705$ is greater than the value of $\mathcal{J}_{r,\varepsilon}^2 = 8.2617$, so we can say that the proposed DBX model fits these data.

For the second right censor example, we have

$$\hat{\underline{V}} = (\hat{\theta}_1, \hat{\theta}_2)^T = (1.7203, 0.9833)^T.$$

where the data are grouped into $r = 4$ intervals. The results of the $\hat{a}_{\underline{v}}(j, n, i), e(j, n), u(j, n), \hat{C}_{1j}(n)$ and $\hat{C}_{2j}(n)$ for the second right censored data are given in the following Table 4.

Table 4: Values of $\hat{a}_{\underline{v}}(j, n, i), e(j, n), u(j, n), \hat{C}_{1j}(n), \hat{C}_{2j}(n)$

$\hat{a}_{\underline{v}}(j, n, i)$	250.5	475.6	1175.7	2081
$u(j, n)$	11	9	8	10
$\hat{C}_{1j}(n)$	0.9834	1.2342	0.8933	1.2191
$\hat{C}_{2j}(n)$	1.0347	1.0733	0.9933	1.1349
$e(j, n)$	4.6325	4.6325	4.6325	4.6325

Then we obtain the value of the test statistic $\mathcal{J}_{r,\varepsilon}^2: \mathcal{J}_{4,0.05}^2 = \chi_{4,0.05}^2 + QF = 7.6325$. For significance level $\varepsilon = 0.05$, the critical value $\chi_{4,0.05}^2 = 9.4877$ is greater than the value of $\mathcal{J}_{r,\varepsilon}^2 = 7.6325$, so we can say that the proposed model DBX fit these data.

For the third right censored data, $\hat{\underline{V}} = (\hat{\theta}_1, \hat{\theta}_2)^T = (1.8363, 1.0634)^T$. Data are grouped into $r = 4$ intervals. Table 5 gives the results of $\hat{a}_{\underline{v}}(j, n, i), e(j, n), u(j, n), \hat{C}_{1j}(n)$ and $\hat{C}_{2j}(n)$ for the third right censored data.

Table 5: Values of $\hat{a}_{\underline{v}}(j, n, i), e(j, n), u(j, n), \hat{C}_{1j}(n), \hat{C}_{2j}(n)$

$\hat{a}_{\underline{v}}(j, n, i)$	42.8	53.5	57.2	60.7
$u(j, n)$	9	12	17	10
$\hat{C}_{1j}(n)$	1.3323	1.6436	1.0638	1.2782
$\hat{C}_{2j}(n)$	0.8886	0.9589	0.9736	0.9833
$e(j, n)$	4.3215	4.3215	4.3215	4.3215

Then we obtain the value of the test statistic $\mathcal{J}_{r,\varepsilon}^2: \mathcal{J}_{4,0.05}^2 = \chi_{4,0.05}^2 + QF = 7.0356$. For significance level $\varepsilon = 0.05$, the critical value $\chi_{4,0.05}^2 = 9.4877$ is greater than the value of $\mathcal{J}_{r,\varepsilon}^2 = 7.0356$, so we can say that the proposed model DBX fit these data.

For the fourth right censored data, $\hat{\mathbb{V}} = (\hat{\theta}_1, \hat{\theta}_2)^T = (1.9346, 1.2102)^T$, where the data are grouped into $r = 4$ intervals. Table 6 lists the results of $\hat{a}_{\hat{\mathbb{V}}}(j, n, i)$, $e(j, n)$, $u(j, n)$, $\hat{C}_{1j}(n)$ and $\hat{C}_{2j}(n)$ for the third fourth censored data.

Table 6: Values of $\hat{a}_{\hat{\mathbb{V}}}(j, n, i)$, $e(j, n)$, $u(j, n)$, $\hat{C}_{1j}(n)$, $\hat{C}_{2j}(n)$

$\hat{a}_{\hat{\mathbb{V}}}(j, n, i)$	5.8	10.2	18.2	27.5
$u(j, n)$	12	11	10	9
$\hat{C}_{1j}(n)$	0.7342	0.6534	0.3342	0.8212
$\hat{C}_{2j}(n)$	1.0833	1.1078	1.0393	1.0567
$e(j, n)$	2.6354	2.6354	2.6354	2.6354

Then we obtain the value of the test statistic $\mathcal{T}_{r,\varepsilon}^2: \mathcal{T}_{4,0.05}^2 = \chi_{4,0.05}^2 + QF = 6.956$. For significance level $\varepsilon = 0.05$, the critical value $\chi_{4,0.05}^2 = 9.4877$ is greater than the value of $\mathcal{T}_{r,\varepsilon}^2 = 6.956$, so we can say that the proposed model DBX fit these data.

6. Concluding remarks

In this article, a modified Bagdonavičius-Nikulin goodness-of-fit test statistic is presented and applied for distributional validation under the right censor case and for the double Burr type X distribution. The modified test statistic ($\mathcal{T}_{r,\varepsilon}^2$) is given along with all its relevant components. For this purpose, the maximum likelihood estimation method in censored data case is used and applied along with a simulation study via the algorithm of Barzilai-Borwein performed for assessing the right censored estimation method. Another simulation study is presented for testing the null hypothesis under the modified version of the Bagdonavičius and Nikulin goodness-of-fit statistical test. Four right censored data sets are analyzed under the new modified test statistic for checking the distributional validation. According to the modified Bagdonavičius-Nikulin goodness-of-fit test statistic, the double Burr type X distribution can be used in modeling the censored medicine and reliability real data sets. Below are results of the modified test statistic under the right censored data sets:

- For the death times in weeks under aneuploid tumors data: $\mathcal{T}_{5,0.05}^2 = 8.2617 < \chi_{5,0.05}^2$.
- For the bone marrow transplant data: $\mathcal{T}_{4,0.05}^2 = 7.6325 < \chi_{4,0.05}^2$.
- For strength data of a certain type of braided cord: $\mathcal{T}_{4,0.05}^2 = 7.0356 < \chi_{5,0.05}^2$.
- For the times to infection data: $\mathcal{T}_{4,0.05}^2 = 6.956 < \chi_{4,0.05}^2$.

References

1. Abouelmagd, T. H. M., Hamed, M. S. and Yousof, H. M. (2019). Poisson Burr X Weibull distribution. Journal of Nonlinear Sciences & Applications (JNSA), 12(3), 173-183.
1. Ali, M. M., Yousof, H. M. and Ibrahim, M. (2021a). A new version of the generalized Rayleigh distribution with copula, properties, applications and different methods of estimation. Optimal Decision Making in Operations Research & Statistics: Methodologies and Applications, VOL 1, 1-25.
2. Ali, M. M., Ibrahim, M. and Yousof, H. M. (2021b). Expanding the Burr X model: properties, copula, real data modeling and different methods of estimation. Optimal Decision Making in Operations Research & Statistics: Methodologies and Applications, VOL 1, 26-49.
3. Habib, M.G., Thomas, D.R. (1986). Chi-squared Goodness-of-Fit Tests For Randomly Censored Data, Annals of Statistics, 14(2), 759-765.
1. Hollander, M., Pena, E.(1992). Chi-square goodness-of-fit test for randomly censored data. JASA 87(417), 458–463.
2. Galanova, N. S., Lemeshko, B. Y. and Chimitova, E. V. (2012). Using nonparametric goodness-of-fit tests to validate accelerated failure time models. Optoelectronics, Instrumentation and Data Processing, 48(6), 580-592.
3. Bagdonavičius, V., & Nikulin, M. (2011a). Chi-squared tests for general composite hypotheses from censored samples. Comptes Rendus Mathematique, 349(3-4), 219-223.
4. Bagdonavičius, V., Nikulin, M.(2011b). Chi-squared Goodness-of-fit Test for Right Censored Data. International Journal of Applied Mathematics and Statistics, 24, 30-50.

5. Bagdonavičius, V., Levulienė, 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.
6. Crowder M. J., Kimber A.C., Smith R. L and Sweeting, T. J. (1991): *Statistical analysis of reliability data*, CHAPMAN & HALL/CRC.
7. Elsayed, H. A. H. and Yousof, H. M. (2019). The Burr X Nadarajah Haghghi distribution: statistical properties and application to the exceedances of flood peaks data. *Journal of Mathematics and Statistics*, 15, 146-157.
8. Klein J. P. and Moeschberger M. L. (2003). *Survival Analysis : Techniques for Censored and Truncated Data*, Springer, New York.
9. 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*, forthcoming.
10. Jahanshahi, S.M.A., Yousof, H. M. and Sharma, V. K. (2019). The Burr X Fréchet Model for Extreme Values: Mathematical Properties, Classical Inference and Bayesian Analysis. *Pak. J. Stat. Oper. Res.*, 15(3), 797-818.
11. Khalil, M. G., Hamedani, G. G. and Yousof, H. M. (2019). The Burr X exponentiated Weibull model: Characterizations, mathematical properties and applications to failure and survival times data. *Pakistan Journal of Statistics and Operation Research*, 141-160.
12. Mansour, M. M., Ibrahim, M., Aidi, K., Shafique Butt, N., Ali, M. M., Yousof, H. M. and Hamed, M. S. (2020a). 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.
13. Mansour, M., Rasekhi, M., Ibrahim, M., Aidi, K., Yousof, H. M. and Elrazik, E. A. (2020b). 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.
14. Sickle-Santanello, B. J., Farrar, W. B., Decenzo, J. F., Keyhani-Rofagha, S., Klein, J. P., Laufman, H. and Toole, R. V. (1988). Technical and statistical improvements for flow cytometric DNA analysis of paraffin-embedded tissue. *Cytometry*, 9(6), 954-599.
15. Voinov, V., Nikulin, M., Balakrishnan, N. (2013). *Chi-Squared Goodness of Fit Tests with Applications*, Academic Press, Elsevier.
16. 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.
17. Yousof, H. M., Aidi, K., Hamedani, G. G. and Ibrahim, M. (2021a). A new parametric lifetime model 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.
18. Yousof, H. M., Ali, M. M., Aidi, K., Hamedani, G. G. and Ibrahim, M. (2021b). A new lifetime distribution with properties, characterizations, validation testing, different estimation methods. *Statistics, Optimization & Information Computing*, forthcoming.