AIMS Mathematics, 9(8): 22513–22531.
DOI: 10.3934/math.20241095
Received: 14 March 2024
Revised: 06 July 2024
Accepted: 16 July 2024
Published: 19 July 2024

http://www.aimspress.com/journal/Math

Research article

Non-parametric hypothesis testing to address fundamental life testing issues in reliability analysis with some real applications

M. E. Bakr*

Department of Statistics and Operations Research, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

* **Correspondence:** Email: mmibrahim@ksu.edu.sa.

Abstract: Life categories and probability distributions are part of a new field in reliability that has emerged as a result of the daily generation of data that has become more complex across practical fields. This study demonstrated how well the U-statistics technique can be applied to real-world testing problems, producing more efficient processes that are on par with or even more successful than conventional approaches. Furthermore, there was room for improvement in the performance of these methods. An approach tending toward normalcy was supported by comparing a unique U-statistic test with the used better than age in moment generating ordering (UBAmgf) test statistic, In this manuscript, a novel nonparametric technique has been developed to test the belonging of a dataset to a distribution of a new statistical class survival function, the moment generating function for used better than aged (UBAmgf). This type of test was crucial in practical life, such as implementing a specific strategy of proposed therapy for a particular disease, deeming it futile if the survival data was exponential (accepting H_0) (the suggested therapeutic approach does not exhibit positive or negative effects on the patients). Once the survival data was UBAmgf, the treatment or device or system employed yields an expected overall current value better or higher than the older device governed by the asymptotic survival function (discussed in the Applications section). The appropriateness of the proposed statistical test's application range was properly determined by calculating its test efficiency and critical values and comparing them with other tests, whether in complete or censored data. Finally, we applied this proposed test technique in the manuscript to a different set of real data in both cases.

Keywords: reliability theory; hypothesis testing; UBA and UBAmgf classes of life distributions; medical data
Mathematics Subject Classification: 62G10, 62G20, 62N05

Notations and abbreviations

IFR	Increasing failure rate
IFRA	Increasing failure rate average
NBU	New better than used
NBUE	New better than used in expectation
NBRUL	New better than renewal used in laplace transform ordering
NBUCL	New better (worse) than used in a convex laplace ordering
DMRL	Decreasing mean residual life
UBA	Used better than age
UBAC	Used better than age in convex order
UBAC (2)	Used better than age in concave order
UBAL	Used better than age in Laplace transform
UBA_{mgf}	Used better than age in moment generating function

1. Introduction

In the field of probability, statistics, and related fields like reliability theory, survival analysis, and economics, the comparison of random probability distributions is considered a fundamental and highly effective principle. Among statisticians and reliability analysts, a primary focus lies in devising new models for various life distributions based on specific aging-related behaviors observed in survival data. Utilizing the U-statistic technique, a method developed by statisticians and reliability analysts, the nonparametric statistical test employed in this study assesses alternative treatment methodologies by examining failure patterns in collected survival data. In recent years, this approach has been increasingly utilized to evaluate exponentiality, particularly in monitoring patient survival times following the implementation of recommended strategies.

For the simulation of various aspects of aging, different categories of life distributions were considered. Increasing failure rate (IFR), new better than used (NBU), used better than age in expectation (UBAE), and Used better than age in moment generating function (UBAmgf) are the most common among these categories. See, for example, Esary et al. [1], Bryson and Siddiqui [2], Barlow and Proschan [3], Navarro J. [4], and Navarro and Pellerey [5]. Other aging categories, such as Fernandez-Ponce et al.'s NBU [6], have been considered by many researchers. Ahmad [7] introduced UBA and UBAE as well. The order of moment generating function for UBA was discussed by Abu Youssef et al. [8], Bakr et al. [9], Bakr and Al-Babtain [10], Ghosh and Mitra [11], EL-Sagheer et al. [12], Etman et al. [13], Bakr et al. [14], EL-Sagheer et al. [15], Mahmoud El-Morshedy et al. [16], Gadallah et al. [17], Abu-Youssef and El-Toony [18], Abu-Youssef and Gerges [19], Bakr [20], Mahmoud et al. [21], Gadallah [22], Bakr et al. [23], and Bakr [24].

Consequently, the following implications can be shown:

$$\text{IFR [1]} \Rightarrow \text{UBA [7]} \Rightarrow \text{UBAmgf [8]}$$
$$\Downarrow$$
$$\text{UBAE [7]}$$

A variety of events, including failure, death, relapse, reaction, the start of particular illnesses, divorce, or parole, are covered by examples of survival time data. Like any other stochastic variable, these intervals have a distribution and random changes. The probability density function, hazard

function, and survivorship function are frequently used to clarify or characterize the distribution of survival times. The other two functions can theoretically be inferred if any one of these functions is given. These three functions practically offer different features of the data. One of the main difficulties in analyzing survival data is estimating one or more of these functions from sampled data and making inferences about the survival pattern of the population.

For any random variable X, let $X_t = [X - t|X > t]$, $t \in \{x: F(x) < 1)$; denote a random variable whose distribution is the same as the conditional distribution of $X - t$, given that $X > t$ and it has a survival function which determines an individual's probability of surviving past a given time t, defined as follows,

$$\bar{F}_t(x) = \begin{cases} \dfrac{\bar{F}(x+t)}{\bar{F}(t)} & \bar{F}(t) > 0 \\ 0 & \bar{F}(t) = 0 \end{cases},$$

where X is the lifetime of a device which has a finite mean $\mu = E(X) = \int_0^\infty \bar{F}(u)du$, u ≥ 0, and the mean of X_t is called the mean residual life (MRL) and is given by

$$\mu(t) = E(X_t) = \frac{\int_t^\infty \bar{F}(u)du}{\bar{F}(t)}.$$

Ahmad [7] introduced UBA aging which we can say that, F is UBA $\forall t, \mathrm{x} \geq 0$, if

$$\mathrm{e}^{-\frac{\mathrm{x}}{\mu(\infty)}}\,\bar{\mathrm{F}}(\mathrm{t}) \leq \bar{\mathrm{F}}(\mathrm{x}+\mathrm{t}), \qquad \mathrm{x}, \mathrm{t} \geq 0,$$

along with UBAE which F is UBA if

$$\mu(\infty) \leq \mu(t),$$

where,

$$\mu(\infty) = \lim_{t\to\infty} \mu(t) = \frac{1}{h(\infty)}, h(t) = -\frac{d}{dt} ln\, \bar{F}(t) = \frac{f(t)}{\bar{F}(t)}, t \geq 0, \bar{F}(t) > 0.$$

Definition. The function F has the UBAmgf property if $0 < \mu(\infty) < \infty$ and $\forall(\mathrm{t}, \mathrm{x} \geq 0)$; (Abu Youssef et al. [8])

$$\int_0^\infty e^{sx}\bar{\mathrm{F}}(x+t)\,\mathrm{d}x \geq \frac{\mu(\infty)}{1-s\mu(\infty)}\bar{F}(t),\ \mathrm{s} \geq 0, \tag{1.1}$$

The purpose of this work is to assess the belonging of a dataset for the class of the UBAmgf life distribution by developing a new nonparametric statistical test based on the u-statistic technique, applicable to both complete and censored data. For both kinds of data, the power of the test and relative approximate efficiency of Pitman (PARE) will be used to evaluate the statistical test's performance. Often used substitutes for the asymptotic model are the Makeham, Weibull, and linear failure rate LFR distributions. Finally, to verify the efficacy and importance of the examination, this suggested methodology will be used on a collection of real-world data.

The structure of the article is as follows. Section 2 will cover the test statistic for complete data and PARE for our test statistic. Section 3 will offer the test statistic for censored data. We discuss several applications of real medical data of the suggested statistical test in Section 4. Finally, in

Section 5, we provide a conclusion.

2. Complete data testing

We create a metric to assess deviation from exponentiality within the UBAmgf category.

2.1. Testing technique

Assume that $X_1, X_2, \ldots, X_n$ is a random sample was extracted to apply this nonparametric testing. Various researchers have suggested nonparametric hypothesis testing for different life distributions, including Abu-Youssef et al. [25] and Mahmoud et al. [26].

The expression for the departure measure can be derived from Eq (1.1), by performing the integration on both sides for Eq (1.1) as follows:

$$\begin{aligned}\delta(s) &= \int_0^\infty \left[\int_0^\infty e^{sx}\,\bar{F}(u+t)\,du - \frac{1}{1-s}\bar{F}(t)\right]dt \\ &= \int_0^\infty \int_0^\infty e^{sx}\,\bar{F}(u+t)\,du\,dt - \frac{1}{1-s}\int_0^\infty \bar{F}(t)\,dt, s \geq 0, \end{aligned} \tag{2.1}$$

which under $H_0: \delta(s) = 0$; see Theorem 2.

The theorem presented below is essential for the development of our test.

Theorem 1. Suppose X is a random variable from UBAmgf; according to Eq (2.1), we obtain

$$\delta(s) = \frac{1}{s^2}(\varphi(s) - 1) - \frac{\mu}{s(1-s)}, \tag{2.2}$$

which $\varphi(s) = \int_0^\infty e^{sx} dF(x), \mu = \int_0^\infty \bar{F}(t)\,dt,$ and $s \neq 1$.

Proof. Starting from (2.1), we have the following:

$$\begin{aligned}\delta(s) &= \int_0^\infty \int_0^\infty e^{sx}\bar{F}(u+t)dudt - \frac{1}{1-s}\mu, \\ &= I - \frac{1}{1-s}\mu, \end{aligned} \tag{2.3}$$

where,

$$\begin{aligned} I &= \int_0^\infty \int_0^\infty e^{su}\bar{F}(u+t)\,dudt \\ &= \frac{1}{s}\int_0^\infty e^{st}(1-e^{-st})\bar{F}(t)\,dt \\ &= \frac{1}{s}\int_0^\infty e^{st}\bar{F}(t)\,dt - \frac{1}{s}\int_0^\infty \bar{F}(t)\,dt \\ &= \frac{1}{s^2}(\varphi(s) - s\mu - 1). \end{aligned}$$

Consequently, the result follows.

The empirical estimate of $\delta(s)$ can be derived as follows:

$$\hat{\delta}(s) = \frac{1}{ns^2}\sum_i \left\{(e^{sX_i} - 1) - \frac{sX_i}{(1-s)}\right\}.$$

To make the test statistics scale invariant, we can get:

$$\hat{\Delta}(s) = \frac{1}{n\bar{X}s^2}\sum_i \left\{\frac{1}{s^2}\left(e^{sX_i} - \frac{sX_i}{(1-s)} - 1\right)\right\} = \frac{1}{n\bar{X}s^2}\,\emptyset(X), \tag{2.4}$$

where, $\emptyset(X) = \frac{1}{s^2}\left(e^{sx} - \frac{sx}{(1-s)} - 1\right)$.

The following theorem illustrates the asymptotic normality of $\hat{\Delta}(s)$.

Theorem 2. Using U-statistics theory (Lee [27]), the statistic $\delta(s)$ has the following properties: As $n \to \infty$, $\sqrt{n}(\hat{\Delta}(s) - \delta(s))$ is asymptotically normal with zero mean and $\sigma^2(s)$, where

$$\mu = E\left[\frac{1}{s^2}(\varphi(s) - 1) - \frac{\mu}{s(1-s)}\right].$$

$$\sigma^2(s) = var\left(\frac{1}{s^2}(\varphi(s) - 1) - \frac{\mu}{s(1-s)}\right),$$

and under H_0,

$$\mu = \int_0^\infty \left(\frac{1}{s^2}(e^{sx} - 1) - \frac{\mu}{s(1-s)}\right)e^{sx}dx$$

$$\sigma_0^2(s) = \frac{2}{(1 + 2s^2 - 3s)(s-1)^2}.$$

Proof. By direct calculation, we have the following

$$E[\emptyset(X)] = E\left(\frac{1}{s^2}(e^{sx} - 1) - \frac{x}{s(1-s)}\right),$$

$$= \int_0^\infty \left(\frac{1}{s^2}(e^{sx} - 1) - \frac{\mu}{s(1-s)}\right)e^{sx}dx = 0,$$

and

$$\sigma^2(s) = \int_0^\infty \left(\frac{1}{s^2}(e^{sx} - 1) - \frac{x}{s(1-s)}\right)^2 e^{-x}dx = \frac{2}{(1 + 2s^2 - 3s)(s-1)^2}.$$

2.2. Pitman Efficiency

In evaluating the effectiveness of this technique, the Pitman asymptotic efficiency (PAE) values are computed for additional option distributions and compared with various other tests. Specifically, we select $\delta^*(0.01)$ investigated by Abu-Youssef et al. [8] for the UBAmgf and $\xi(0.01,5)$ studied by El-Arishy et al. [28] for the renewal new better than used in the moment generating function (RNBUmgf). This method allows us to achieve the following assessments:

(i) LFR

$$\bar{F}_1(x) = e^{-x-\frac{x^2}{2}\theta},\ \theta, x \geq 0. \tag{2.5}$$

(ii) Weibull distribution (WD)

$$\bar{F}_2(x) = e^{-x^{\theta}}, \quad \theta \geq 1, , x \geq 0. \tag{2.6}$$

(iii) Makeham distribution (MD)

$$\bar{F}_3(x) = e^{-\theta(e^{-x}+x-1)-x}, \quad \theta, x \geq 0. \tag{2.7}$$

The PAE of $\delta(s)$ is equal to

$$\mathrm{PAE}(\delta(s)) = \frac{\left|\frac{\partial}{\partial\theta}\delta(s)\right|_{\theta\to\theta_0}}{\sigma_0(s)} = \frac{1}{\sigma_0(s)}\left|\frac{1}{s^2}\int_0^{\infty} e^{sx} d\bar{F}'_{\theta_0}(x) - \frac{1}{s(1-s)}\int_0^{\infty} x d\bar{F}'_{\theta_0}(x)\right|,$$

where $\bar{F}'_{\theta_0}(x) = \frac{d}{d\theta}\bar{F}_{\theta}(u)\Big|_{\theta\to\theta_0}$. This leads to:

(i) PAE in case of LFR:

$$\mathrm{PAE}(\delta(0.1)) = \frac{1}{\sigma_0}\left|\frac{1}{0.01}\int_0^{\infty} e^{0.1x} d\left(-\frac{x^2}{2}e^{-x}\right) - \frac{1}{0.09}\int_0^{\infty} x\, d\left(-\frac{x^2}{2}e^{-x}\right)\right| = 1.4.$$

$$\mathrm{PAE}(\delta(0.3)) = \frac{1}{\sigma_0}\left|\frac{1}{0.09}\int_0^{\infty} e^{0.3x} d\left(-\frac{x^2}{2}e^{-x}\right) - \frac{1}{0.21}\int_0^{\infty} x\, d\left(-\frac{x^2}{2}e^{-x}\right)\right| = 1.3.$$

(ii) PAE in case of WD:

$$\mathrm{PAE}(\delta(0.1)) = \frac{1}{\sigma_0}\left|\frac{-1}{0.01}\int_0^{\infty} e^{0.1x} d(-x\ln(x)e^{-x}) - \frac{1}{0.09}\int_0^{\infty} x d(-x\ln(x)e^{-x})\right| = 0.98.$$

$$\mathrm{PAE}(\delta(0.3)) = \frac{1}{\sigma_0}\left|\frac{-1}{0.09}\int_0^{\infty} e^{0.3x} d(-x\ln(x)e^{-x}) - \frac{1}{0.21}\int_0^{\infty} x d(-x\ln(x)e^{-x})\right| = 0.86.$$

(iii) PAE in case of MD:

$$\mathrm{PAE}(\delta(0.1)) = \frac{1}{\sigma_0}\left|\frac{-1}{0.01}\int_0^{\infty} e^{0.1x} d(-e^{-x}(e^{-x}+x-1)) - \frac{1}{0.09}\int_0^{\infty} x d(-e^{-x}(e^{-x}+x-1))\right| = 0.51.$$

$$\mathrm{PAE}(\delta(0.3)) = \frac{1}{\sigma_0}\left|\frac{-1}{0.09}\int_0^{\infty} e^{0.3x} d(-e^{-x}(e^{-x}+x-1)) - \frac{1}{0.21}\int_0^{\infty} x d(-e^{-x}(e^{-x}+x-1))\right| = 0.42.$$

Simplified PAE calculations for $\delta^*(0.01)$, $\xi(0.01,5)$ and our statistic test at $s = 0.1$ and $s = 0.3$ are summarized in Tables 1 and 2, that is, demonstrating that our tests for $F_1, F_2, \text{and } F_3$ are effective.

Table 1. PAE of $\xi(0.01,5), \delta^*(0.01)$ and $\delta(0.1)$ & $\delta(0.3)$.

Distribution	$\xi(0.1,5)$	$\delta^*(0.01)$	$\delta(0.1)$	$\delta(0.3)$
LFR	0.943	1.4	1.4	1.3
WD	0.695	0.597	0.98	0.86
MD	0.192	0.102	0.51	0.42

Table 2. PARE of $\delta(0.1)$ and $\delta(0.3)$ with respect to $\xi(0.01,5)$ and $\delta^*(0.01)$.

Distribution	$e(\delta(0.1),\xi)$	$e(\delta(0.3),\xi)$	$e(\delta(0.1),\delta^*)$	$e(\delta(0.3),\delta^*)$
LFR	1.48	1.38	1	0.93
Weibull	1.41	1.24	1.64	1.44
Makeham	2.57	2.19	5	4.12

Our statistics clearly perform well, as seen in Tables 1 and 2.

Critical values

The upper percentile of $\hat{\Delta}_{un}(s)$ is calculated in this section using Mathematica programming based on a randomly chosen sample size of 10,000.

Upon examining Tables 3 and 4, as well as Figure 1, it's noticeable that the behavior of critical values tends to approach a normal distribution as the sample size increases.

Table 3. Upper percentile points of $\hat{\Delta}(s)$.

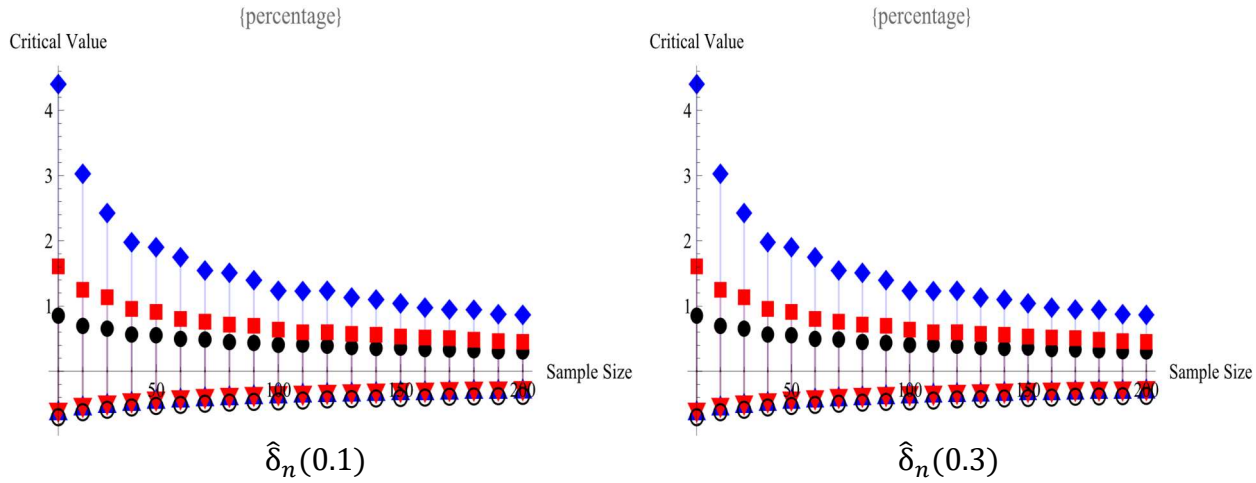
	$\hat{\Delta}_{un}(0.1)$			$\hat{\Delta}_{un}(0.3)$		
n	90%	95%	99%	90%	95%	99%
10	0.589599	1.0548	2.38187	0.90344	1.70724	4.36234
20	0.478266	0.789781	1.55303	0.769239	1.28583	3.28312
30	0.405417	0.610084	1.19421	0.617771	1.0489	2.5286
40	0.347127	0.53271	1.02942	0.603084	0.996328	2.05001
50	0.314902	0.483235	0.921412	0.53828	0.870271	1.91975
60	0.300995	0.449961	0.831487	0.518077	0.825426	1.72713
70	0.282376	0.418966	0.710617	0.470304	0.771146	1.5557
80	0.263594	0.372777	0.64752	0.444529	0.710233	1.50348
90	0.245934	0.349949	0.582578	0.41206	0.659672	1.38659
100	0.234955	0.344106	0.603165	0.410608	0.639758	1.31727
110	0.220522	0.315423	0.548694	0.386204	0.616867	1.26791
120	0.219413	0.319273	0.533454	0.382229	0.590118	1.25532
130	0.206842	0.292671	0.490819	0.380101	0.588795	1.13755
140	0.205538	0.288783	0.486047	0.36453	0.563253	1.10836
150	0.197803	0.277123	0.471442	0.34982	0.536596	1.02944

Continued on next page

	$\hat{\Delta}_{un}(0.1)$			$\hat{\Delta}_{un}(0.3)$		
n	90%	95%	99%	90%	95%	99%
160	0.186517	0.268745	0.441818	0.34347	0.503481	0.99703
170	0.182598	0.260694	0.434232	0.340778	0.500624	0.974983
180	0.174062	0.249311	0.398776	0.321098	0.489979	0.969483
190	0.173023	0.243926	0.383918	0.304935	0.465992	0.883525
200	0.172233	0.242028	0.380304	0.302076	0.464217	0.879726

Table 4. Lower percentile points of $\hat{\Delta}$(s).

	$\hat{\Delta}_{un}(0.1)$			$\hat{\Delta}_{un}(0.3)$		
n	90%	95%	99%	90%	95%	99%
10	-0.471067	-0.426453	-0.395363	-0.592297	-0.635645	-0.7004
20	-0.419978	-0.372406	-0.340227	-0.526397	-0.564187	-0.625656
30	-0.38235	-0.339507	-0.302180	-0.479946	-0.525115	-0.587352
40	-0.364249	-0.312609	-0.278526	-0.450476	-0.496252	-0.560209
50	-0.348252	-0.294385	-0.258681	-0.418322	-0.468409	-0.537834
60	-0.333397	-0.2797	-0.246042	-0.392849	-0.443084	-0.511545
70	-0.31856	-0.266462	-0.228035	-0.380477	-0.430749	-0.497798
80	-0.303186	-0.249586	-0.215193	-0.360892	-0.410798	-0.481773
90	-0.298034	-0.244458	-0.208772	-0.348319	-0.398453	-0.473025
100	-0.286063	-0.232463	-0.198667	-0.341951	-0.387858	-0.458414
110	-0.280916	-0.226196	-0.191198	-0.328458	-0.372738	-0.452141
120	-0.269198	-0.218881	-0.185565	-0.316119	-0.364621	-0.432687
130	-0.263995	-0.211496	-0.179830	-0.307513	-0.353976	-0.424693
140	-0.252446	-0.203445	-0.171397	-0.295639	-0.343957	-0.415618
150	-0.249648	-0.199637	-0.167331	-0.292632	-0.341521	-0.411077
160	-0.241313	-0.191188	-0.160843	-0.287568	-0.331935	-0.40222
170	-0.247363	-0.191327	-0.160827	-0.27644	-0.323556	-0.397487
180	-0.235357	-0.185306	-0.155513	-0.27222	-0.318282	-0.390303
190	-0.233028	-0.182623	-0.150336	-0.268424	-0.31458	-0.379604
200	-0.228591	-0.178437	-0.151272	-0.26155	-0.307621	-0.372109



$\hat{\delta}_n(0.1)$ $\hat{\delta}_n(0.3)$

Figure 1. The Relation of critical values with sample size.

3. Censored data

People are frequently susceptible to right censoring in survival analysis or analysis of time-to-event data. The Kaplan-Meier estimator provides a consistent estimator of the common marginal distribution function F within the widely recognized framework of random (or independent) censoring (one minus). So, in order to extend U-statistics to right-censored data (under independent or random censoring), one method is to integrate (average) the kernel regard to the product of Kaplan-Meier estimators of the failure time distribution. Such statistics are generally referred to as the Kaplan-Meier U-statistics (Datta et al. [29]). Thus far, asymptotic normality of a Kaplan-Meier U-statistic of order two has been established in the literature (Bose and Sen, [30]), which is given by

$$\widehat{U} = \frac{\sum_{1\le i<j\le n} h(T_i, T_j) W_i W_j}{\sum_{1\le i<j\le n} W_i W_j},$$

where W_i is the mass assigned to the ith right-censored failure time T_i by a Kaplan-Meier estimator. As can be seen from Bose and Sen [30], one faces extensive asymptotic calculations with this form of censored U-statistic.

3.1. Testing exponentiality

A proposed test statistic is suggested in this section to test the H_0 and H_1 hypotheses, where randomly right-censored data are used. We can express the measure of departure as follows:

$$\delta_c(s) = \frac{1}{s^2}\sum_{j=1}^{n}\prod_{k=1}^{j-1}\left(\left(\sum_{m=1}^{n}\tau - 1\right) - \frac{s}{(1-s)}(\xi)\right), \tag{3.1}$$

where $\tau = \sum_{m=1}^{n} e^{sZ_{(m)}}[\prod_{p=1}^{m-2} C_p^{I_p} - \prod_{p=1}^{m-1} C_p^{I_p}]$, and $\xi = \sum_{m=1}^{n}\left[\prod_{p=1}^{m-2} C_p^{I_p} - \prod_{p=1}^{m-1} C_p^{I_p}\right]$.

3.2. Critical values

Once again, Mathematica programming is used in this section to get the upper percentile of $\delta_c(s)$ using a randomly selected sample size of 10,000.

After examining Table 5, as well as Figure 2, it's noticeable that the behavior of critical values tends to approach a normal distribution in all cases as the sample size increases.

Table 5. Upper percentile points of $\delta_c(s)$.

	$\hat{\delta}_c(0.01)$			$\hat{\delta}_c(0.03)$		
n	90%	95%	99%	90%	95%	99%
5	1571.41	1958.13	1978.57	169.641	211.715	215.147
10	566.416	653.487	845.747	58.8706	69.2857	91.4355
15	307.293	360.621	457.593	31.3036	37.5648	47.879
20	194.979	226.324	282.967	20.7336	24.9836	32.5422
25	140.75	165.07	211.233	14.7634	17.5481	22.6889
30	110.104	128.472	170.311	10.9404	13.1516	16.7708
35	84.0231	98.327	134.255	8.38991	9.86629	13.3953
40	66.9122	77.9284	103.379	7.09847	8.496	10.475
45	58.2185	67.1532	86.8666	5.74715	6.9643	9.68581
50	50.0569	58.0822	79.4455	4.65359	5.80076	7.84431
55	42.7518	49.6379	66.847	4.24271	5.49049	7.38808
60	38.2681	45.7686	58.2368	3.4532	4.30971	5.82472
65	32.8402	39.0926	49.166	3.24946	3.84523	5.05027
70	29.1869	34.5325	43.7522	2.68896	3.25801	4.43777
75	26.5801	31.3683	41.9513	2.44625	3.21558	4.34292
80	22.7444	27.8858	35.9219	2.32005	2.85058	3.87856
85	21.812	25.7467	34.335	2.07935	2.61302	3.70679
90	19.9409	23.8876	31.1644	1.76254	2.26448	3.05033
95	16.9682	21.0559	26.9644	1.6898	2.06181	2.8819
100	17.983	21.1462	26.4482	1.47302	1.79136	2.54551

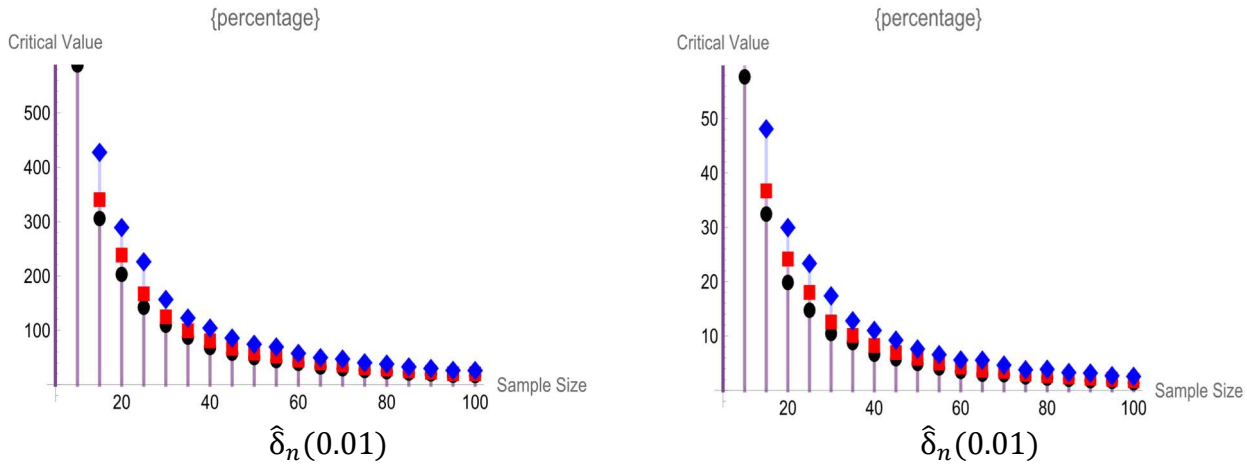


Figure 2. The correlation between sample size and critical values.

4. Applications in practical scenarios

To display the importance of the results in this research, we applied these results to some sets of real data.

4.1. Case of uncensored data

Dataset# 1: Reviewing the data in Almetwally et al. [31] (see Figure 3), which show the ages of 36 patients (in years) with COVID-19, have the following ordered observations:

3.1091, 3.3825, 3.1444, 3.2135, 2.4946, 3.5146, 4.9274, 3.3769, 6.8686, 3.0914, 4.9378, 3.1091, 3.2823, 3.8594, 4.0480, 4.1685, 3.6426, 3.2110, 2.8636, 3.2218, 2.9078, 3.6346, 2.7957, 4.2781, 4.2202, 1.5157, 2.6029, 3.3592, 2.8349, 3.1348, 2.5261, 1.5806, 2.7704, 2.1901, 2.4141, 1.9048

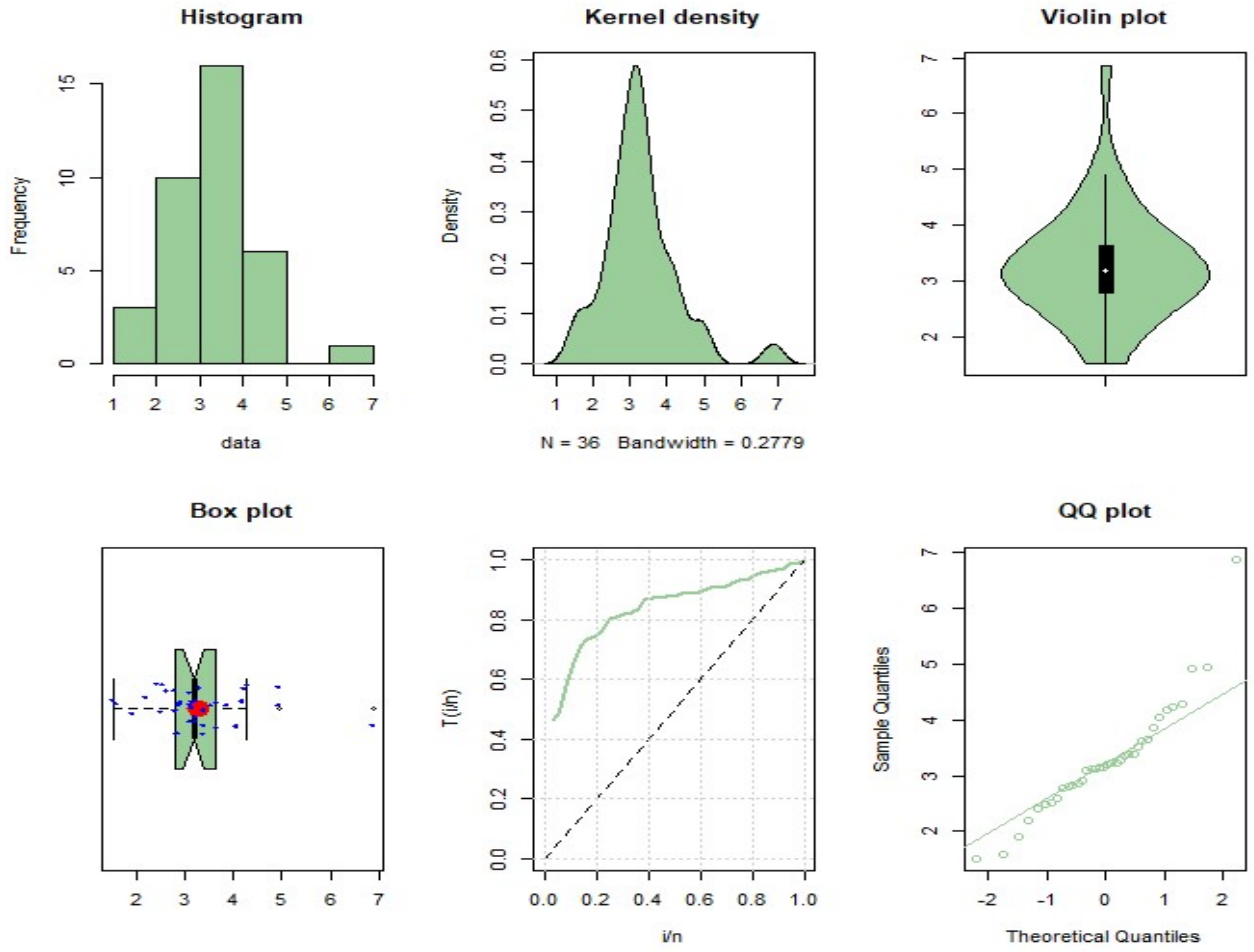


Figure 3. Representation of (COVID-19) dataset.

In the two situations of $\hat{\Delta}(0.1)$ and $\hat{\Delta}(0.3)$ as $n = 36$, we calculate the statistic in (2.4) $\hat{\Delta}(0.1) = 2.6$ and $\hat{\Delta}(0.3) = 2.7$, compared with the Table 3's critical value. The exponentiality null hypothesis is rejected by our test for any value of s, we conclude.

Dataset# 2: Survival times (in years) after diagnosis of 43 patients with a specific type of leukemia are represented in this dataset from Johnson and Kotz [32], (see Figure 4).

0.019	0.129	0.159	0.203	0.485	0.636	0.748	0.781
0.869	1.175	1.206	1.219	1.219	1.282	1.356	1.362
1.458	1.564	1.586	1.592	1.781	1.923	1.959	2.134
2.413	2.466	2.548	2.652	2.951	3.038	3.600	3.655
3.745	4.203	4.690	4.888	5.143	5.167	5.603	5.633
6.192	6.655	6.874					

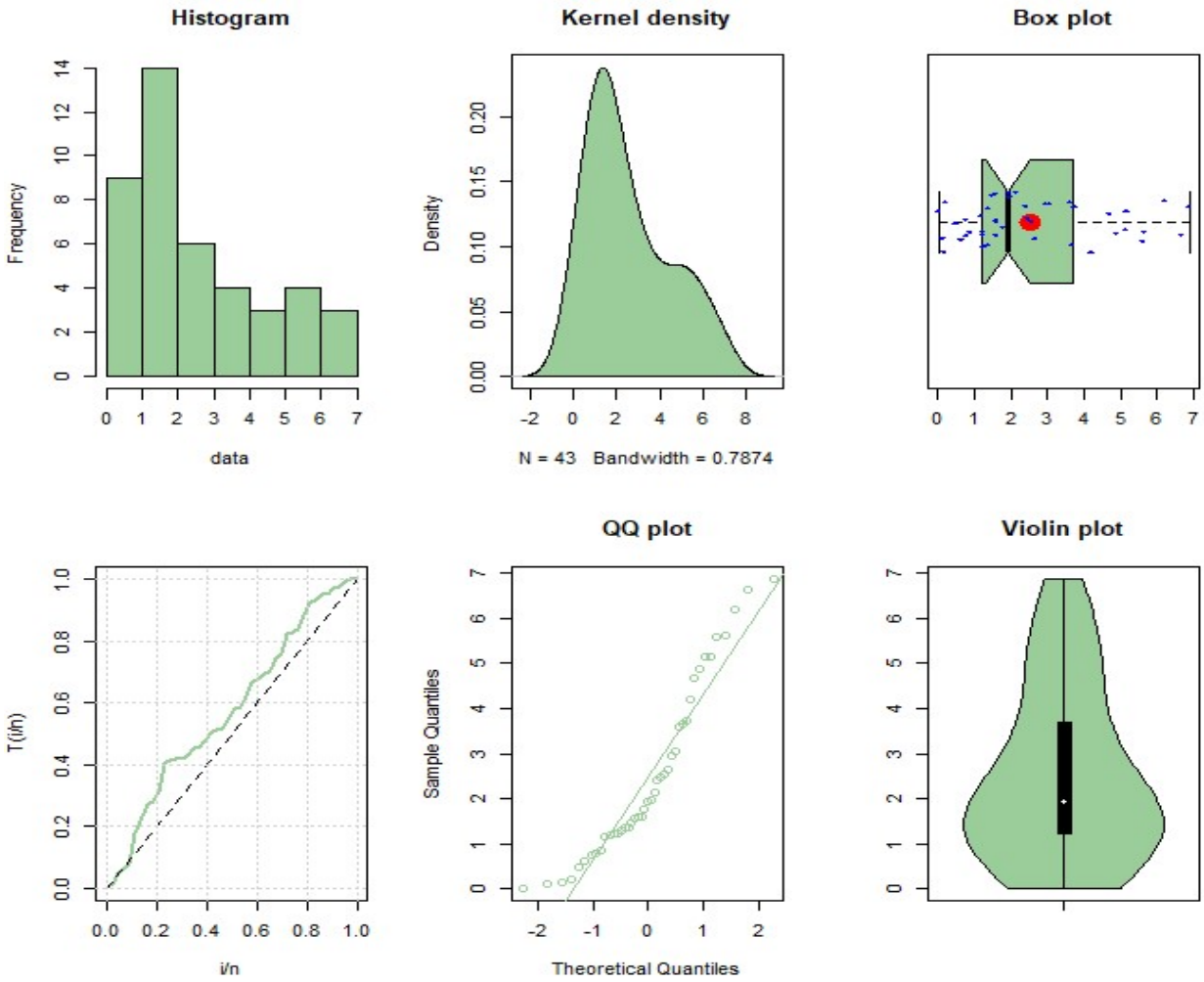


Figure 4. Representation of (leukemia) dataset.

We calculate the statistic in (2.4) $\hat{\Delta}(0.1) = 3.2$ and $\hat{\Delta}(0.3) = 4.3$, compared with the Table 3's critical value. The exponentiality null hypothesis is rejected by our test for any value of s, we conclude.

4.2. Case of censored data

Dataset# 3: We derive $\hat{\delta}_c(s)$ which is smaller than the crucial value of Table 4 in all cases by

applying the statistic in (3.1) to the data in Hassan [33] which reflects the ages of 51 patients (in days) with liver tumors and taken from Egypt and only 39 patients are observed (see Figure 5). The data was as follows:

10	14	14	14	14	14	15	17	18	20
20	20	20	20	23	23	24	26	30	30
+30	+30	+30	+30	+30	31	40	49	51	52
60	+60	61	67	71	74	75	87	96	105
107	107	107	116	150	+150	+150	+150	+150	+150
+185									

The exponentiality null hypothesis fails to be rejected by our test for any value of s, we conclude.

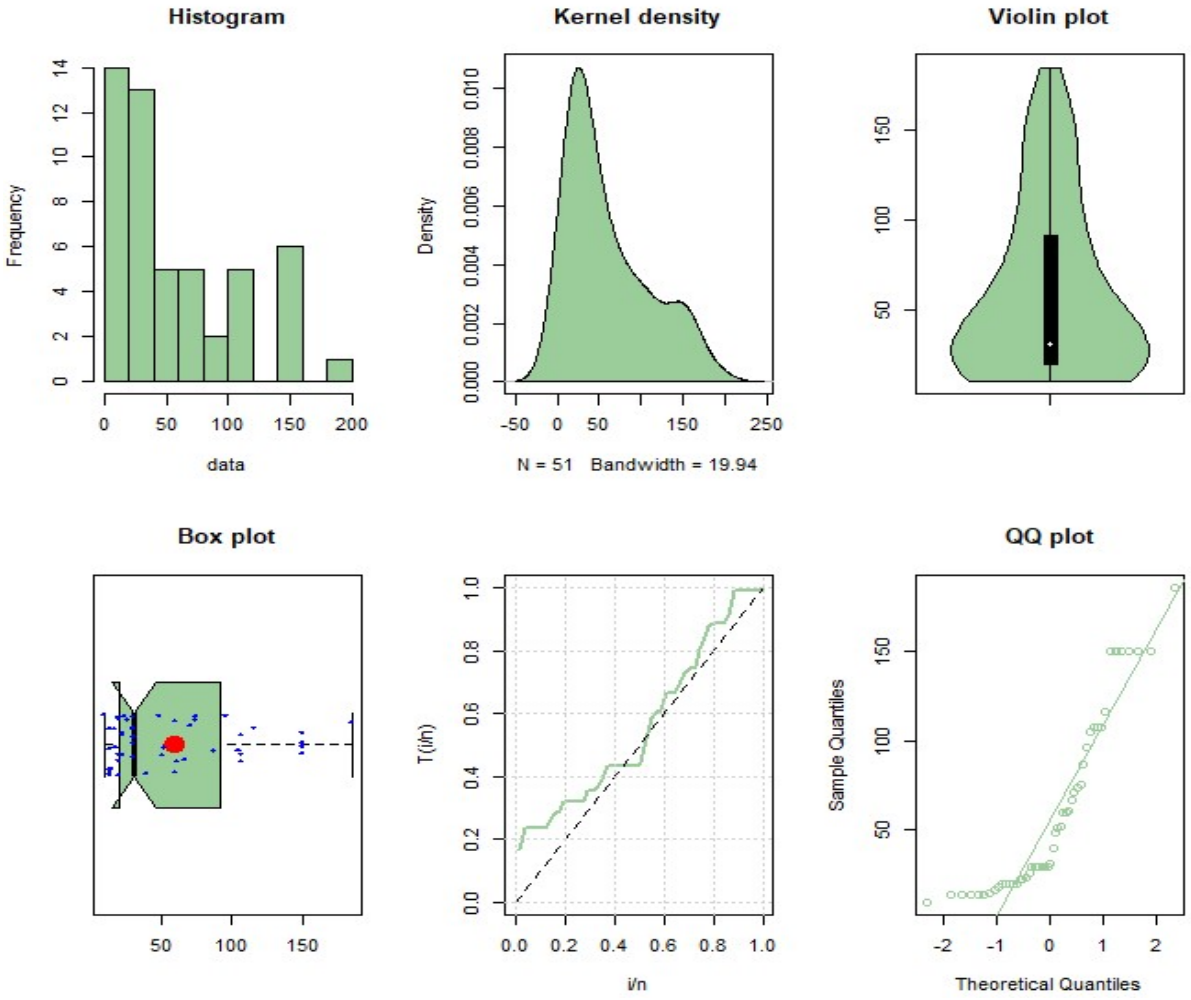


Figure 5. Representation of (liver tumors) dataset.

Dataset# 4: Comparison of two treatments and two diets:

Thirty melanoma patients were evaluated to compare the ability of the immunotherapies BCG (Bacillus Calmette-Guerin) and Corynebacterium parvum to extend survival time. Table 6 shows the

age, gender, treatment received, and time spent alive (Lee and Wolfe [34]). All individuals were resected before starting treatment, thus there was no evidence of melanoma when they started. Calculating the length of survival and comparing the survival time distributions in each group is the most common purpose with this type of data.

Table 6. Age, gender, treatment received, and time spent alive in months.

N. Patient	Age	Gender	Treatment Received	Survival
1	59	F	BCG	33.7+
2	50	F	BCG	3.9
3	76	M	BCG	10.5
4	66	F	BCG	5.4
5	33	M	BCG	19.5
6	23	F	BCG	23.8+
7	40	F	BCG	7.9
8	34	M	BCG	16.9+
9	34	M	BCG	16.6+
10	38	F	BCG	33.7+
11	54	F	BCG	17.1+
12	49	M	C. parvum	8
13	35	M	C. parvum	26.9+
14	22	M	C. parvum	21.4+
15	30	M	C. parvum	18.1+
16	26	F	C. parvum	16+
17	27	M	C. parvum	6.9
18	45	F	C. parvum	11+
19	76	F	C. parvum	24.8+
20	48	M	C. parvum	23+
21	91	M	C. parvum	8.3
22	82	F	C. parvum	10.8+
23	50	F	C. parvum	12.2+
24	40	M	C. parvum	12.5+
25	34	M	C. parvum	24.4
26	38	M	C. parvum	7.7
27	50	M	C. parvum	14.8+
28	53	F	C. parvum	8.2+
29	48	F	C. parvum	8.2+
30	40	F	C. parvum	7.8+

The question is whether or not the two therapy groups have a statistically significant difference in survival. Is the difference shown by the data significant or just a result of the sample's randomness? A statistical significance test is required. A statistical test that ignores patient characteristics, on the other hand, makes sense only if the prognostic indications of the two groups of patients are similar. So far, it's been considered that the patients in the two groups are similar and that the only difference is treatment. As a result, it is necessary to assess the homogeneity of the two groups prior to executing a statistical test.

Although there are no well-established prognostic indicators for melanoma patients, it has been observed that women and the young have a greater survival rate than men and the old (UBA class of life distribution). Furthermore, the stage of the disease has a big influence on survival.

Now we have a dataset that claims to compare the two survival distributions in two ways: First, the data is exponential (which could reflect one of the immunotherapies), and second, the data is UBAmgf. There are various parametric and nonparametric tests to support one of the two hypotheses; however, we will proceed to compare the two survival distributions using nonparametric techniques because we do not know the survival distribution that the data follows. Sections 2 and 3 explain two tests that are appropriate.

We derive $\hat{\delta}_c(0.01) = 6.9 \times 10^{23}$ and $\hat{\delta}_c(0.03) = 2.3 \times 10^{23}$ which is greater than the crucial value of Table 5 in all cases by applying the statistic in (3.1). The exponentiality null hypothesis is rejected by our test for any value of s, we conclude, (see Figure 6).

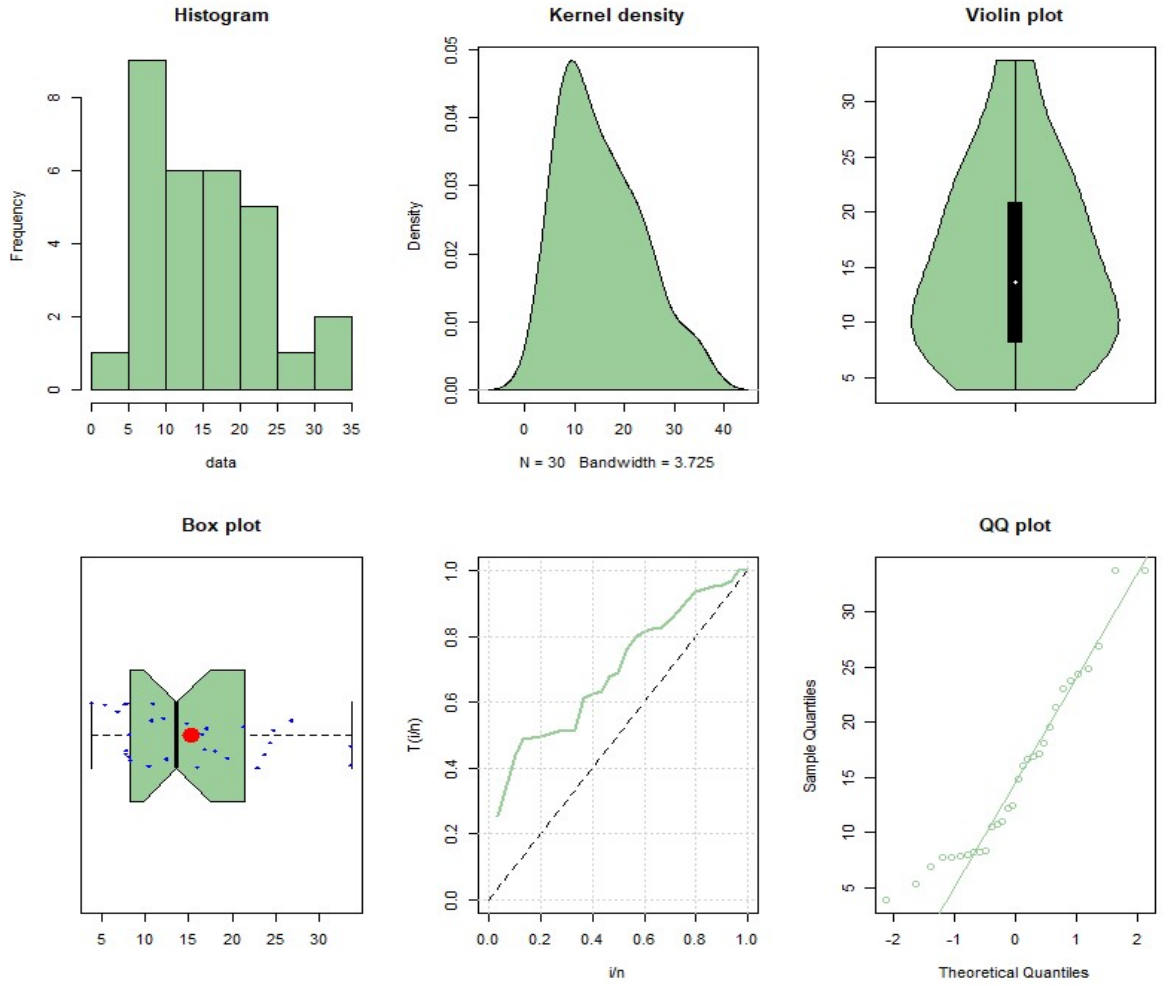


Figure 6. Representation of (tumors) dataset.

Dataset# 5: We derive $\hat{\delta}_c(0.01) = 1528$ and $\hat{\delta}_c(0.03) = 6.7 \times 10^{10}$ which is greater than the crucial value of Table 5 in all cases by applying the statistic in (3.1) to the data in Kamran Abbas et al. [35] and in Lee and Wolfe [36], which reflects the ages of 61 patients (in weeks) (see Figure 7) with inoperable lung cancer, and only 33 patients are observed. The data was as follows:

+0.14	+0.14	+0.29	+0.43	0.43	+0.57	+0.57	+1.86	2.86	+3.00
+3.00	3.14	3.14	+3.29	+3.29	3.43	3.43	3.71	3.86	+6.00
+6.00	+6.14	6.14	6.86	+8.71	9.00	9.43	+10.57	10.71	10.86
11.14	+11.86	13.00	14.43	+15.57	15.71	+16.57	+16.57	+17.29	18.43
18.57	+18.71	20.71	+21.29	23.86	+26.00	+27.57	29.14	29.71	+32.14
+33.14	40.57	48.57	49.43	53.86	61.86	66.57	68.71	68.96	72.86
72.86									

The exponentiality null hypothesis is rejected by our test for any value of s, we conclude.

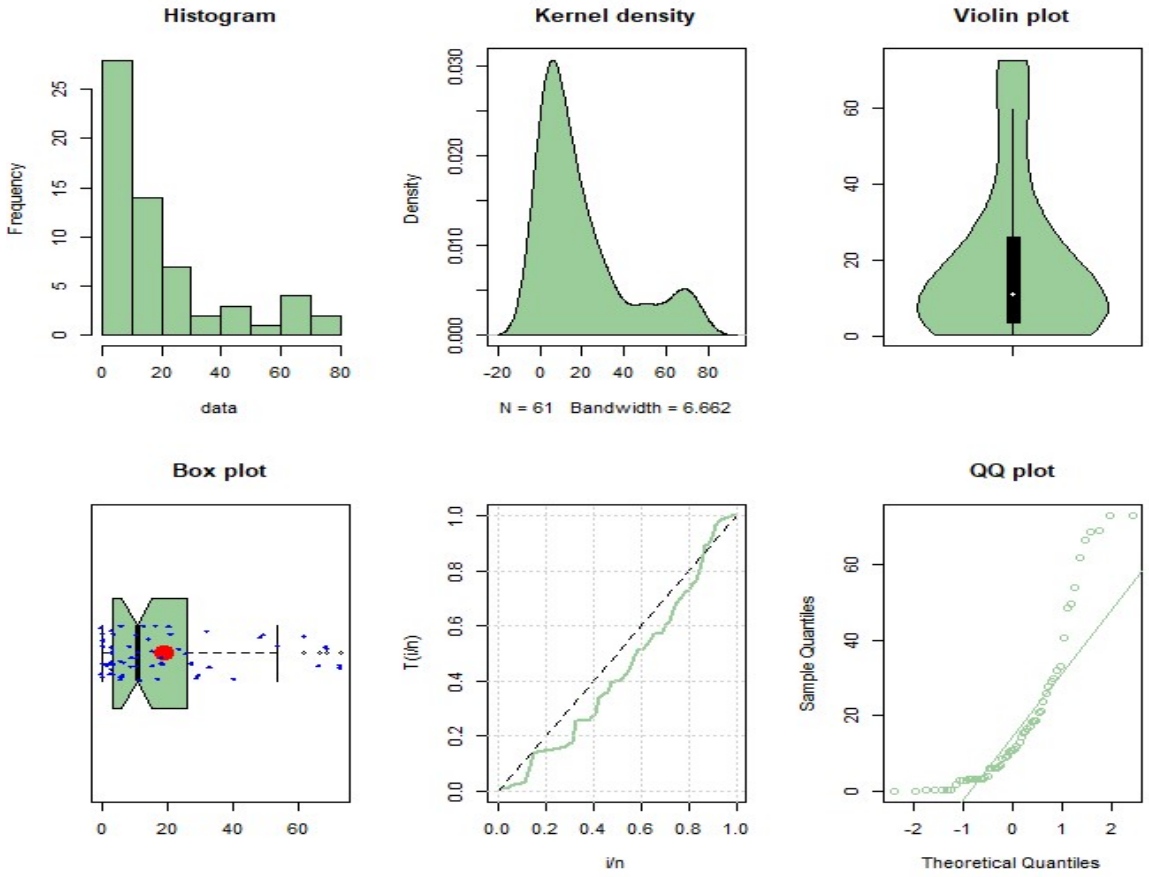


Figure 7. Representation of (lung cancer) dataset.

5. Conclusions

Recent advancements in life testing methodologies significantly depart from traditional approaches due to the increasing complexity of data in various practical domains. This shift has expanded the realm of reliability assessment, with a focus on life categories and specific probability distributions. Our research emphasizes the efficacy of incorporating the U-statistics technique into life testing scenarios, resulting in more streamlined processes compared to conventional methods, particularly advantageous when dealing with smaller sample sizes. The exponential distribution holds a crucial position in statistics as a continuous distribution, widely employed for its memoryless

property, especially in reliability theory, life testing, and stochastic systems. It is vital to ascertain whether a random sample conforms to this distribution. This study introduced nonparametric tests based on the U-statistics approach to assess both conformity to exponential distribution and the presence of exponential distribution in datasets, considering diverse data compositions. Our investigation validated the asymptotic normality of the proposed test and determined critical percentile values through Monte Carlo simulations. Additionally, we computed power estimates for alternative distributions, offering insights for both censored and uncensored data, and discussed strategies for appropriately handling truncated data. Our findings underscored the practical utility of these tests in analyzing reliability across bioscience sustainability data, encompassing both censored and uncensored datasets.

Use of AI tools declaration

The author declares he has not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This study was funded by Researchers Supporting Project number (RSPD2024R1004), King Saud University, Riyadh, Saudi Arabia.

Conflict of interest

The author declares no conflict of interest.

References

1. J. D. Esary, A. W. Marshall, F. Proschan, Some reliability applications of the hazard transform, *SIAM J. Appl. Math*., **18** (1994), 849–860. https://doi.org/10.1137/0118077
2. M. C. Bryson, M. M. Siddiqui, Some criteria for aging, *J. Am. Stat. Assoc*., **64** (1969), 1472–1483. https://doi.org/10.1080/01621459.1969.10501072
3. R. E. Barlow, F. Proschan, *Statistical theory of reliability and life testing. Probability models*, To Begin With, Silver Spring, MD, 1981.
4. J. Navarro, Preservation of DMRL and IMRL aging classes under the formation of order statistics and coherent systems, *Stat. Prob. Lett.,* **137** (2018), 264–268. https://doi.org/10.1016/j.spl.2018.02.005
5. J. Navarro, F. Pellerey, Preservation of ILR and IFR aging classes in sums of dependent random variables, *Appl. Stoch. Model. Bus.*, 2021, 1–22. https://doi.org/10.1002/asmb.2657
6. J. M. Fernandez-Ponce, F. Pellerey, M. R. Rodrıguez-Grinolo, On a new NBUE property in multivariate sense: An application, *Comput. Stat. Data Anal*., **55** (2011), 3283–3294. https://doi.org/10.1016/j.csda.2011.06.010
7. I. A. Ahmad, Some properties of classes of life distributions with unknown age, *Stat. Prob. Lett.,* **9** (2004), 333–342. https://doi.org/10.1016/j.spl.2004.06.029
8. S. E. Abu-Youssef, N. S. A. Ali, M. E. Bakr, Used better than aged in mgf ordering class of life distribution with application of hypothesis testing, *J. Stat. Appl. Prob. Lett.,* 2020, 1–10.

9. M. E. Bakr, A. A. Al-Babtain, S. K. Khosa, Statistical modeling of some cancerous diseases using the Laplace transform approach of basic life testing Issues, *Comput. Math. Method. Med.*, 2022, 1–8. https://doi.org/10.1155/2022/8964869
10. E. B. Mahmoud, A. A. Al-Babtain, Non-parametric hypothesis testing for unknown aged class of life distribution using real medical data, *Axioms*, **14** (2022), 2353.
11. S. Ghosh, M. Mitra, A new test for exponentiality against HNBUE alternatives, *Commun. Stat. Theor Meth.,* **49** (2020)*,* 27–43. https://doi.org/10.1080/03610926.2018.1528370
12. R. M. EL-Sagheer, M. S. Eliwa, K. M. Alqahtani, M. EL-Morshedy, Asymmetric randomly censored mortality distribution: Bayesian framework and parametric bootstrap with application to COVID-19 data, *J. Math.*, 2022, 8300753. https://doi.org/10.1155/2022/8300753
13. W. B. H. Etman, R. M. EL-Sagheer, S. E. Abu-Youssef, A. Sadek, On some characterizations to NBRULC class with hypotheses testing application, *Appl. Math. Inf. Sci.*, **16** (2022), 139–148. https://doi.org/10.18576/amis/160201
14. M. E. Bakr, M. Nagy, A. A. Al-Babtain, Non-parametric hypothesis testing to model some cancers based on goodness of fit, *AIMS Math.,* **7** (2022), 13733–13745. https://doi.org/10.3934/math.2022756
15. R. M. EL-Sagheer, M. A. W. Mahmoud, W. B. H. Etman, Characterizations and testing hypotheses for NBRUL-to class of life distributions, *J. Stat. Theory Pract*., **16** (2022), 31–44. https://doi.org/10.1007/s42519-022-00258-8
16. M. El-Morshedy, A. Al-Bossly, R. M. EL-Sagheer, B. Almohaimeed, W. B. H. Etman, M. S. Eliwa, A moment inequality for the NBRULC class: Statistical properties with applications to model asymmetric data, *Symmetry*, **14** (2022), 2353. https://doi.org/10.3390/sym14112353
17. A. M. Gadallah, B. I. Mohammed, A. A. Al-Babtain, S. K. Khosa, M. Kilai, M. Yusuf, et al., Modeling various survival distributions using a nonparametric hypothesis testing based on Laplace transform approach with some real applications, *Comput. Math. Method. Med*., 2022, 5075716. https://doi.org/10.1155/2022/5075716
18. S. E. Abu-Youssef, A. A. El-Toony, A new class of life distribution based on Laplace transform and it's applications, *Inform. Sci. Lett*., **11** (2022), 355–362. https://doi.org/10.18576/isl/110206
19. S. E. Abu-Youssef, S. T. Gerges, Based on the goodness of fit approach, a new test statistic for testing NBUCmgf class of life distributions, *Pak. J. Statist*., **38** (2022), 129–144.
20. M. E. Bakr, Statistical modeling for some real applications in reliability analysis using non-parametric hypothesis testing, *Symmetry*, 2023, 1735. https://doi.org/10.3390/sym15091735
21. M. A. W. Mahmoud, R. M. EL-Sagheer, W. B. H. Etman, Moments inequalities for NBRUL distributions with hypotheses testing applications, *Austrian J. Stat*., **47** (2018), 95–104. https://doi.org/10.17713/ajs.v47i1.579
22. A. M. Gadallah, B. I. Mohammed, A. A. Al-Babtain, S. K. Khosa, M. Kilai, M. Yusuf, et al., Modeling various survival distributions using a nonparametric hypothesis testing based on Laplace transform approach with some real applications, *Comput. Math. Method. Med.,* 2022, 5075716. https://doi.org/10.1155/2022/5075716
23. M. E. Bakr, B. G. Kibria, A. M. Gadallah, A new non-parametric hypothesis testing with reliability analysis applications to model some real data, *J. Radiat. Res. Appl. Sci.*, **16** (2023), 100724. https://doi.org/10.1016/j.jrras.2023.100724
24. M. E. Bakr, On basic life testing issues in medical research using non-parametric hypothesis testing, *Qual. Reliab. Engng. Int.*, 2023, 1–12.
25. S. E. Abu-Youssef, N. S. A. Ali, M. E. Bakr, Non-parametric testing for unknown age (UBAL) class of life distribution, *J. Test. Eval.,* **48** (2019), 1–13. https://doi.org/10.1520/JTE20170726

26. M. A. W. Mahmoud, L. S. Diab, D. M. Radi, Testing exponentiality against renewal new better than used in Laplace transform order: applications in reliability, *J. Egypt. Math. Soc.,* **29** (2019), 49–60. https://doi.org/10.1186/s42787-019-0044-7
27. A. J. Lee, *U-statistics theory and practice*, Marcel Dekker, New York, 1990.
28. S. M. El-Arishy, L. S. Diab, E. S. El-Atfy, Testing exponentially against RNBUmgf based on Laplace transform technique, *J. Stat. Appl. Prob.*, **8** (2019), 229–239. https://doi.org/10.18576/jsap/080307
29. S. Datta, D. Bandyopadhyay, G. A. Satten, Inverse probability of censoring weighted U-statistics for right-censored data with an application to testing hypotheses, *Scand. J. Stat.*, **37** (2010), 680–700. https://doi.org/10.1111/j.1467-9469.2010.00697.x
30. A. Bose, A. Sen, Asymptotic distribution of the Kaplan--Meier U-statistics, *J. Multivariate Anal*., **83** (2002), 84–123. https://doi.org/10.1006/jmva.2001.2039
31. E. M. Almetwally, R. Alharbi, D. Alnagar, E. H. Hafez, A new inverted toppleone distribution: Applications to the COVID-19 mortality rate in two different countries, *Axioms*, **10** (2021), 25. https://doi.org/10.3390/axioms10010025
32. N. L. Johnson, S. Kotz, *Encyclopedia of statistical sciences*, New York: Wiley., **3** (1983)**.**
33. M. K. Hassan, On testing exponentiality ageing UBA class of life distribution based on Laplace transform, *Sankhya B*, **79** (2017), 142–155. https://doi.org/10.1007/s13571-015-0112-4
34. E. T. Lee, J. W. Wang, *Statistical methods for survival data analysis*, 3 Eds., John Wiley & Sons, Inc., Hoboken, New Jersey, 2003.
35. K. Abbas, Z. Hussain, N. Rashid, A. Ali, M. Taj, S. A. Khan, et al., Bayesian estimation of gumbel type-II distribution under type-II censoring with medical applications, *J. Comput. Math. Method. Med.,* **7** (2020), 1–11. https://doi.org/10.1155/2020/1876073
36. E. T. Lee, R. A. Wolfe, A simple test for independent censoring under the proportional hazards model, *Biometrics*, **54** (1998), 1176–1182. https://doi.org/10.2307/2533867

AIMS Press

© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)