Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/120321

Modeling Survival Distributions For Certain Real-World Applications Using Nonparametric Hypothesis Testing

*M. E. Bakr[1,2], Arwa M. Alsahangiti[1] and A. M. Gadallah[3] **

[1]Department of Statistics and Operation Research, Faculty of Science, King Saud University, P.O. Box2455, Riyadh 11451, Saudi Arabia
[2]Department of Basic Sciences, Giza Higher Institute for Engineering, Cairo, Egypt
[3]Department of Basic Sciences, Thebes Higher Institute for Engineering, Thebes Academy, Cairo, Egypt

Received: 6 Sep. 2022, Revised: 20 Nov. 2022, Accepted: 24 Dec. 2016.
Published online: 1 Sep. 2023.

Abstract: By analyzing the failure behavior of the recorded survival data, we expect to analyze the different processing strategies or functionality of the tools or systems used in this nonparametric statistical test. It is anticipated that the test data would either behave exponentially or like the NBUE Property. If the survival results are NBUE, it is anticipated that the suggested treatment plan would be successful. Contrarily, As shown in the application section, if the data are exponential, the suggested treatment plan won't have any positive or negative consequences on patients. We evaluated the suggested test's efficacy and power for both complete and censored data, compared the outcomes with those of existing tests, and then applied the test to a variety of real-world data to demonstrate its validity.

Keywords: Censored data; Laplace transform; NBUE; Non-Parametric hypothesis testing; Pittman asymptotic efficiency; Reliability theory.

1 Introduction

Hypothesis estimation and hypothesis testing are the two primary statistical inference subfields. In most cases, we do not know the actual values of population parameters. Therefore, we have to evaluate them. But We do make some educated assumptions on the true values of the population parameters. The null hypothesis is frequently denoted by the abbreviationH_0, whereas the alternative hypothesis is denoted by the letter H_1. Before performing a significant statistical test, it is necessary to develop the null and alternative hypotheses.

For aging criteria, there are two different test kinds. You may model lifespan data using quantile functions (see Sreelakshm et al [1] and Sankaran and Midhu [2]). The other test evaluates if $\bar{F}(x)$ belongs to a specific aging class or has an exponential distribution.

Hypothesis estimation and testing are the two primary subfields of statistical inference. In the majority of cases, we are unsure of the population parameters' actual values. In order to estimate them, we must. The true values of the population parameters, however, are subject to some of our theories. Generally, H_0 is used to represent the null hypothesis, whereas H_1 is used to represent the alternative hypothesis. The null hypothesis and the alternative hypothesis need to be defined before you can run an insightful statistical test.

Definition 1.1 $F(x)$ has the NBU property iff,

$$\bar{F}(x)\bar{F}(t) \geq \bar{F}(x+t);\ t > 0. \tag{1}$$

Definition 1.2 $F(x)$ has the NBUE property iff,

$$\bar{\mathrm{V}}(x) \leq \mu\bar{F}(x);\ \mu, x > 0 \tag{2}$$

*Corresponding author e-mail: alaadean_mag@yahoo.com

Where $\overline{V}(x) = \int_x^{\infty} \overline{F}(u)du$

Hollander and Proschan [3], as well as Fernández-Ponce et al [4], are two authors who have suggested the probabilistic characteristics of the NBUE class of life distribution. Hassan [5], Navarro and Pellerey [6], EL-Sagheer et al [7], Navarro [8], Abu-Youssef et al [9, 10, 11], Atallah et al. [12], Al-Gashgari et al. [13], Gadallah et al [14], Gadallah [15], Abu-Youssef and Bakr [16] and Bakr et al [17, 18] are a few among them.

Our need to determine whether the data in this study is exponential or NBUE is causing an issue because we have some genuine data. We must first generate a test statistic before we can select or choose between H_0 and H_1. The test statistic measures how well a sample result corresponds with the hypothesis being tested using a random variable.

An explanation of this manuscript's structure is provided below: In order to create test statistics, we employ the Laplace transform method. In Section 2: for typical alternatives with a certain sample size, we calculate the test's power and Pitman's asymptotic efficiency, as well as the critical points of the Monte Carlo null distribution. We suggest a test to identify data that has been right-censored in Section 3. The importance of our test is finally discussed in Section 4 using a set of actual medical data.

2 A departure measures

To begin with, we establish a measure for the NBUE class's departure from exponentiality. Monte Carlo exponential distribution critical points, the powers for common alternatives and Pitman's asymptotic efficiencies are all used in this technique.

2.1 Complete Data Testing

For the exponential class, the distribution function is $F(x) = 1 - e^{-\beta x},\ for\ \beta, x > 0$. Our official objective is to pit $H_0 : F \in \xi$ vs $H_1 : F \in NBUE \setminus \xi$.

The next lemma provides a measure of departure.

Lemma 2.1.

$$\delta(s) = \frac{1}{s^2} E[(s+1) - (s+1)e^{-sx} - sx] \tag{3}$$

Proof:

According to equation (2), the measure of deviation from H_0 is defined as

$$\delta(s) = \int_0^{\infty} [\mu \overline{F}(x) - \overline{V}(x)]\ e^{-sx}\, dx,$$

Without sacrificing generality, use $\mu = 1$,

$$\delta(s) = \int_0^{\infty} e^{-sx}\overline{F}(x)\, dx - \int_0^{\infty}\int_x^{\infty} e^{-sx}\overline{F}(u)\, dudx,$$

$$= I_1 - I_2,$$

Where,

$$I_1 = \int_0^{\infty} \overline{F}(x)e^{-sx}\, dx$$

$$= \frac{1}{s} E[1 - e^{-sx}],$$

And

$$I_2 = \int_0^\infty \int_x^\infty \bar{F}(u)e^{-sx}dudx,$$

$$= \frac{1}{s^2}E[e^{-sx} + sx - 1].$$

This leads to the proof.

Set the empirical estimator of $\delta(s)$ in (3) as:

$$\hat{\delta}_n(s) = \frac{1}{s^2 n}\sum_i \{(s+1)(1-e^{-sX_i}) - sx_i\}.$$

Or

$$\hat{\delta}_n(s) = \frac{1}{n\,s^2}\sum_i \emptyset(X_i),$$

Where

$$\emptyset(X_i) = \{(1+s)(1-e^{-sX_i}) - sx_i\}.$$

To determine the limiting distribution of $\hat{\delta}_n(s)$, the U-statistic theory may be employed.

Set

$$\emptyset(X) = 1 - s - e^{-sx} - se^{-sx} - sx.$$

The following theorem can be used to show how the statistic given in (3) is asymptotic normal.

Theorem 2.1. The statistic, in accordance with U-statistics theory, has the following features.:

i- As $n \to \infty$, $(\hat{\delta}_n - \delta)$ is asymptotically normal, with zero mean and variance σ^2 such that:

$$\sigma^2 = E\left(\frac{s+1}{s^2} - \frac{s+1}{s^2}e^{-sx} - \frac{x}{s}\right)^2$$

ii- With H_0, we can minimize the variance to the

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

Proof:

i- Utilizing U-statistic theory, the results listed below are attained (Lee [19]):

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

$$\sigma^2 = \mathrm{Var}[\emptyset(X)] = E\left(\frac{s+1}{s^2} - \frac{s+1}{s^2}e^{-sx} - \frac{x}{s}\right)^2.$$

ii- With H_0, it is obvious that $\mu_0 = 0$, as well as the variance is:

$$\sigma_0^2(s) = \frac{2}{1+3s+2s^2}. \quad (4)$$

2.2 Critical Values

Using Mathematic 8 Programming, we determine the higher percentage points for our test.

Table 1 indicates a percentile point at $s = 0.5, 0.9$ that is significant statistically.

Table 1. The Higher Percentile Values of $\hat{\delta}_n(s)$

	$\hat{\delta}_n(0.5)$			$\hat{\delta}_n(0.9)$		
n	90%	95%	99%	90%	95%	99%
5	0.280	0.304	0.338	0.237	0.256942	0.287
10	0.242	0.266	0.299	0.197	0.219304	0.249
15	0.214	0.240	0.277	0.174	0.199192	0.231
20	0.197	0.223	0.261	0.155	0.179	0.213
25	0.182	0.206	0.244	0.139	0.162	0.199
30	0.167	0.195	0.233	0.130	0.154	0.193
35	0.155	0.184	0.226	0.122	0.143	0.177
40	0.148	0.173	0.215	0.113	0.135	0.169
45	0.140	0.167	0.210	0.111	0.133	0.167
50	0.134	0.159	0.199	0.101	0.123	0.158
55	0.127	0.153	0.193	0.099	0.120	0.154
60	0.122	0.147	0.185	0.097	0.119	0.150
65	0.118	0.143	0.182	0.092	0.112	0.143
70	0.114	0.139	0.176	0.089	0.107	0.141
75	0.113	0.136	0.175	0.085	0.104	0.135
80	0.108	0.133	0.172	0.084	0.103	0.135
85	0.105	0.127	0.167	0.081	0.098	0.131
90	0.104	0.125	0.164	0.079	0.097	0.125
95	0.102	0.123	0.159	0.077	0.095	0.123
100	0.098	0.118	0.157	0.075	0.093	0.125

Table 1 demonstrates how our test's asymptotic normality increases.

2.2 Power Estimates

The results of the suggested test's efficacy are displayed in Table 2 at a significance level of $\alpha = 0.05$ (Mathematica 8 Programming).

Table 2. Powers estimates

family	n	$\theta = 2$	$\theta = 3$	$\theta = 4$

LFR	10	0.592	0.675	0.637
	20	0.946	0.987	0.987
	30	0.998	0.999	1
Makeham	10	1	1	1
	20	1	1	1
	30	1	1	1
Weibull	10	0.228	0.839	0.855
	20	0.930	1	1
	30	0.994	1	1
Gamma	10	0.832	0.900	0.925
	20	0.775	0.99	99
	30	995	1	1

Table 2 shows that the proposed test has perfect validity for Makeham distribution and growing validity for each of LFR, Weibull and Gamma distributions.

2.3 Pittman Efficiency

To assess the effectiveness of this process, we can contrast our exam with those from other courses.

The PAE of our test is assessed as follows in this section:

$$\mathrm{PAE}(\delta) = \frac{\left|\frac{\partial}{\partial\theta}\delta\right|_{\theta\to\theta_0}}{\sigma_0}$$

$$= \frac{1}{\sigma_0}\left|\frac{1}{s^2}\int_0^{\infty} e^{-sx}\overline{F}'_{\theta_0}(x)dx - \frac{1}{s(1-s)}\int_0^{\infty} x\overline{F}'_{\theta_0}(x)dx\right|$$

Where $\overline{F}'_{\theta_0}(x) = \frac{d}{d\theta}\overline{F}_{\theta}(u)\Big|_{\theta\to\theta_0}$.

In most cases, the Makeham, Weibull, LFR, and Gamm distributions are chosen in place of the exponential distribution.

(i) $\mathrm{PAE}(\delta, \mathrm{LFR}) = \frac{1}{\sigma_0}\left|\frac{2+s}{(s+1)^2}\right|$

(ii) $\mathrm{PAE}_{\mathrm{Makeham}}(\delta) = \frac{1}{\sigma_0}\left|\frac{3+s}{2(2+s)(1+s)}\right|$

(iii) $\mathrm{PAE}_{\mathrm{weibull}}(\delta,) = \frac{1}{\sigma_0}\left|\frac{\ln(1+s)-s(\gamma-1)}{s(1+s)}\right|, \quad \gamma(Euler\ constant)$

(iv) $\mathrm{PAE}_{\mathrm{Gamm}}(\delta) = \frac{1}{\sigma_0}\left|\frac{\ln(1+s)-s}{s^2}\right|.$

Comparisons are made between the efficacy of our suggested test $\hat{\delta}_n(s)$ and that of Ahmed et al. [20] $(\delta_n^{(1)})$ and Mugdadi and Ahmad [21] (δ_3) in Table 3.

Table 3. PAE of $\delta_n^{(1)}$, $\delta_{F_n}^{(2)}$ and $\hat{\delta}(s)$

Distribution	$\delta_n^{(1)}$	δ_3	$\hat{\delta}_n(s)$				
			$s = 0.1$	$s = 0.5$	$s = 0.9$	$s = 1.5$	$s = 1.9$
LFR	1.3	0.4	1.41	1.4	1.3	1.25	1.22
Makeham	0.58	0.04	0.5	0.6	0.58	0.57	0.57
Weibull	0.97	0.2	1.0	1.0	0.98	0.9	0.89
Gamm	----	------	0.4	0.5	0.5	0.6	0.61

Table 3 indicates how well the chosen families performed in our exam.

3 Censored Data

Based on a characteristic of survival statistics, one of the most significant advancements in the life sciences has been made: The results will be skewed if some participants did not witness the important event at the conclusion of the study or during data processing. Because It's uncertain when the illness will stop or the time of survival is unclear, some patients might still be healthy or alive at the conclusion of the experiment. If someone cannot be contacted after a research session, they are known as incomplete observations or censored times.

3.1 Test Statistic

In this section, a statistical test is created to assess H_0 and H_1 using data that has been randomly right censored.

We may rephrase the measure of departure as follows using the Kaplan-Meier estimator and definition (2):

$$\delta_c(s) = \frac{1}{ns^2}\big((1+s)(1-v) - s\xi\big), \tag{5}$$

where

$$v(s) = \int_0^{\infty} e^{-sx} dF(x),$$

$$\hat{v}(s) = \sum_{m=1}^{l} e^{-sZ_{(m)}}\Big(\prod_{p=1}^{m-2} C_p^{\delta_p} - \prod_{p=1}^{m-1} C_p^{\delta_p}\Big)$$

$$\xi = \sum_{j=1}^{l}\prod_{k=1}^{j-1} C_m^{\delta_m}\big(Z_{(j)} - Z_{(j-1)}\big)$$

Table 4. The Higher Percentile Values of $\hat{\delta}_c$

$\hat{\delta}_c * 10^{-3}; s = 0.5$				$\hat{\delta}_c * 10^{-3}; s = 0.9$		
n	95%	98%	99%	95%	98%	99%
5	556	724	75	142	200	223
10	223	254	278	63	72	78
15	113	141	159	35	42	45
20	75	89	99	22	26	28

25	56	64	68	14	17	19
30	39	47	55	12	15	16
35	31	37	45	99	12	13
40	26	31	34	8	9	97
45	21	25	27	6	7	8
50	18	20	23	5	6	7
55	15	19	22	4	6	7
60	13	16	19	4	5	53
65	12	15	17	4	4	5
70	10	13	14	3	4	5
75	9	11	11	3	3	4
80	8	10	11	3	3	3
85	8	9	10	2	3	3
90	6	8	9	2	3	3
95	6	8	8	2	2	3
100	6	7	8	2	2	2

3.2 Power Estimates

The results of the suggested test's efficacy are displayed in Table 5.

Table 5. Powers estimates at $\alpha = 0.05$

n	θ	Distribution		
		Weibull	LFR	Gamma
10	1	0.9354	0.9983	0.9317
	2	0.9998	0.9999	0.8932
	3	1	1	0.8774
20	1	0.8945	0.9964	0.9971
	2	0.996	0.9998	0.9958
	3	1	1	0.9951
30	1	0.8783	0.9944	1
	2	0.99994	0.9999	1
	3	1	0.9999	1

Table 5 shows the impressive powers of our test for the Weibull, LFR, and Gamma families.

4 Applications

We show the applicability of the findings in this paper by applying them to real-world data sets.

Example 1: Applying the information from Abouammoh et al. [22], which shows the 40 leukemia patients' ages, expressed in years:

0.315	0.496	0.616	1.145	1.208	1.263	1.414	2.025	2.036
2.162	2.211	2.370	2.532	2.693	2.805	2.910	2.912	3.192
3.263	3.348	3.348	3.427	3.499	3.534	3.767	3.751	3.858
3.986	4.049	4.244	4.323	4.381	4.392	4.397	4.647	4.753
4.929	4.973	5.074	4.381					

We calculate the statistics in (3) for $\hat{\delta}(0.5) = 1.8$ and $\hat{\delta}(0.9) = 1.6$, and both of them are above the critical value of Table 1. Hence, the data set thus has the NBUE attribute.

Example 2: 51 liver cancer patients age from the Elminia Cancer Center of the Egyptian Ministry of Health who started receiving medical evaluations in 2000 is shown using data from Hassan [23] and Mahmoud et al. [24]. Only 39 patients (right-censored) are included in the study, while the remaining 12 are not (missing from the study). A list of ordering details (in days) is provided below.

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

The Censored data are:

30	30	30	30	30	60	150	150	150	150	150	185

In the two situations of $\hat{\delta}_c(0.5)$ and $\hat{\delta}_c(0.9)$ as n=51, we calculate the statistics in (5) $\hat{\delta}_c(0.5) = 0.00012$ and $\hat{\delta}_c(0.9) = 0.00365264$, both of which are under the critical value of Table 4. Consequently, the data collection is devoid of the NBUE characteristic.

Example 3: Think about the information in Lee and Wolfe [25] and Kamran Abbas et al. [26] regarding the survival periods of 61 patients with incurable lung cancer treated with cyclophosphamide. The patients whose treatment was discontinued due to a deteriorating condition are represented by the 33 uncensored observations and the 28 censored observations.

In the two situations of $\hat{\delta}_c(0.4)$ and $\hat{\delta}_c(0.6)$, we calculate the statistics in (5), both of which are above the critical value of Table 4. Consequently, the data set have the property NBUE.

5. Conclusions

In this paper, we provided a non-parametric NBUE testing method that is based on the Laplace transform and operates in both complete and censored data. The recommended statistics' percentages were simulations. Using Pitman's asymptotic relative efficiency and well-known life distributions like the LFR and Weibull families, the efficacy of our recommended tests was contrasted to the efficacy of many other classes. The results of the research were then applied to real data sets from clinical settings.

Conflicts of Interest Statement

The authors declare there is no conflict of interest.

Ethics Statement

This research did not require ethical approval. Data Availability Statement Data associated with the manuscript is public and has been referenced appropriately.

References

[1] Sreelakshmi, N., Sudheesh K. Kattumannil & Asha G., Quantile based tests for exponentiality against DMRQ and NBUE alternatives. Journal of the Korean Statistical Society, 47(2), 180-200 (2018).

[2] Sankaran, P. G., & Midhu, N. N., Testing exponentiality using mean residual quantile function, Statistical Papers, 57, 235–247 (2016).

[3] Hollander, M. & Prochan, F., Test for mean residual life. Biometrika, 62, 585-593 (1975).

[4] Ferna´ndez-Ponce JM, Pellerey F. & Rodrı´guez-Grin˜olo MR., On a new NBUE property in multivariate sense: an application, Computational Statistics & Data Analysis,55(12) 3283–3294 (2011).

[5] Hassan, M. KH., A generalized Hollander-Proschan test for NBUE alternative based on U-statistics approach, International Journal of Reliability and Applications 16(2), 113-122 (2015).

[6] Navarro, J. & Pellerey, F., Preservation of ILR and IFR aging classes in sums of dependent random variables, Appl Stochastic Models Business and Industry, 38(2), 240-261 (2021).

[7] EL-Sagheer, R. M., Abu-Youssef, S. E., Sadek, A., Omar, K. M. & Etman, W. B. H., Characterizations and testing NBRUL class of life distributions based on Laplace transform technique, Journal of Statistics Applications & Probability, 11(1), 75-88 (2022).

[8] Navarro, J., Preservation of DMRL & IMRL aging classes under the formation of order statistics and coherent systems, Statistics and Probability Letters, 137 (1), 264-268 (2018).

[9] Abu-Youssef S. E., Ali N. S. A. & Bakr M. E., "Used better than aged in mgf ordering class of life distribution with application of hypothesis testing" Journal of Statistics Applications & Probability Letters, vol. 7, no. 1, 23-32, (2020).

[10] Abu-Youssef S. E., Ali N. S. A. & Bakr M. E., "Non-parametric testing for unknown age (UBAL) class of life distribution", Journal of testing and evaluation, vol. 48, no. 6, P. 4576, (2019).

[11] Abu-Youssef S. E., Mohammed B. I. & Bakr M. E., "Nonparametric test for unknown age class of life distributions", International Journal of Reliability and Applications, vol. 15, no. 2, 99-110, (2014).

[12] Atallah,M.A., Mahmoud,M.A.W. & Alzahrani,B.M., A new test for exponentiality versus NBUmgf life distribution based on Laplace transform, Quality and Reliability Engineering International, 30(8), 1353-1359 (2014).

[13] Al-Gashgari, F. H., Shawky, A. I. & Mahmoud, M. A. W, A non-parametric test for exponentiality against NBUCA class of life distribution based on Laplace transform, Quality and Reliability Engineering International, 32(1), 29-36 (2016).

[14] Gadallah A. M., Mohammed B. I., Abdulhakim A. Al-Babtain, Saima Khan Khosa, Mutua Kilai, Yusuf M. & Bakr M. E., " Modeling various Survival Distributions using A Non-Parametric Hypothesis Testing based on Laplace Transform Approach with some Real Applications ", Computational and Mathematical Methods in Medicine, vol. 2022, Article ID 5075716, 8 pages, (2022).

[15] Gadallah, A. M., Testing EBUmgf Class of Life Distributions based on Laplace Transform Technique, Journal of Statistics Applications & Probability. 6(3), 471-477 (2017).

[16] Abu-Youssef S. E. & Bakr M. E., "Laplace transform order for unknown age class of life distribution with some applications," Journal of Statistics Applications & Probability, vol. 7, no. 1, 105-114, (2018).

[17] Bakr M. E., Abdulhakim A. Al-Babtain, & Saima Khan Khosa, "Statistical Modeling of Some Cancerous Diseases Using the Laplace Transform Approach of Basic Life Testing Issues", Computational and Mathematical Methods in Medicine, vol. 2022, Article ID 8964869, 8 pages, (2022).

[18] Bakr M. E., Nagy M. & Abdulhakim A. Al-Babtain, "Non-parametric hypothesis testing to model some cancers based on goodness of fit", AIMS Mathematics, 7(8): 13733–13745, (2022).

[19] Lee A. J., U-Statistics Theory and Practice, Marcel Dekker, Inc. New York, NY(1990).

[20] Ahmad, I.A., Alwasel, I.A. & Mugdadi, A.R., A goodness of fit approach to major life testing problems, International Journal of Reliability and Applications, 2, 81-97 (2001).

[21] Mugdadi, A.R. & Ahmad, I.A., Moment inequalities derived from comparing life with its equlibrium form, Journal of Statistical Planning and Inference, 134, 303-317 (2005).

[22] Abouammoh, A.M., Abdulghani, S.A. & Qamber I.S., On partial orderings and testing of new better than renewal used classes, Reliability Engineering System Safety, 43(1), 37–41 (1994).

[23] Hassan, M.KH., On testing exponentiality ageing UBA class of life distribution based on Laplace transform, Sankhya B, 79, 142-155 (2017).

[24] Mahmoud, M.A.W., Hassan, E.M.A. & Gadallah A.M., On NRBUL class of life distributions, Journal of the Egyptian Mathematical Society, 26(3), 483-490 (2018).

[25] Lee E. T. & Wolfe R. A., A simple test for independent censoring under the proportional hazards model, Biometrics, 54, 1176-1182 (1998).

[26] Kamran Abbas, Zamir Hussain, Noreen Rashid, Amjad Ali, Muhammad Taj, Sajjad Ahmad Khan, Sadaf Manzoor, Umair Khalil, & Dost Muhammad Khan., Bayesian Estimation of Gumbel Type-II Distribution under Type-II Censoring with Medical Applications, Journal of Computational and Mathematical Methods in Medicine, 7, 1-11(2020).