CORE
Metadata, citation and similar papers at core.ac.uk
Provided by KHALSA PUBLICATIONS

ISSN 2347-1921

Nonparametric Test for UBACT Class of Life Distribution Based on U-Statistic

S. E.Abu-Youssef[a] and B. I.Mohammed[a]*
Department of Mathematics, Faculty of Science, Al- Azhar University, Cairo, Egypt
M. E.Bakr[b]1
Egyptian General Authority for Meteorology, Ministry of Civil Aviation, Cairo, Egypt

Abstract.

Based on U-statistic, testing exponentially versus used better than aged in convex tail ordering (UBACT) class of life distribution is introduced for complete and censored data. Convergence of the proposed statistic to the normal distribution is proved. Selected critical values are tabulated for sample sizes 5(5)80 for complete data, and (61)(10)(201) for censored data. The Pitman asymptotic relative efficiency of the proposed tests to the other classes is studied. A numerical examples in medical science demonstrates practical application of the proposed test.

Key Words: UBAC class of life distribution; U-Statistic; hypothesis testing; asymptotic normality; efficiency.

Council for Innovative Research

Peer Review Research Publishing System

Journal: JOURNAL OF ADVANCES IN MATHEMATICS

Vol .11, No.2

www.cirjam.com , editorjam@gmail.com

INTRODUCTION

Let X be a non-negative continuous random variable representing equipment life with distribution function F and survival function $\bar{F}(x) = 1 - F(x)$; such that F(0-) = 0, given a unit which has survived up to time t; with distribution function $F_t(x)$ and survival function

$$\bar{F}_t(x) = \frac{\bar{F}(x+t)}{\bar{F}(t)}, \qquad x, t > 0,$$

and assume that X has a finite mean

$$u = E(X) = \int_0^\infty \bar{F}(u)du.$$

Some properties concerning the asymptotic behavior of Xt as $t \to \infty$ will be used.

Definition 1.1

If X is non-negative random variable, its distribution function $F(x)$ is said to be finitely and positively smooth if a number $\gamma \in (0,\infty)$ exists and,

$$\lim_{t\to\infty} \frac{\bar{F}(x+t)}{\bar{F}(t)} = e^{-\gamma x}, \qquad (1.1)$$

whereγ is called the asymptotic decay coefficient of X. See Bhattacharjee (1982).

Definition 1.2

The distribution function F is said to be Used better than aged (UBA) if it is finitely and positively smooth and satisfies

$$\bar{F}(x+t) \geq \bar{F}(t)e^{-\gamma x}. \qquad (1.2)$$

Definition 1.3

The distribution function F is said to be used better than aged in convex ordering (UBAC) if it is finitely and positively smooth and satisfies,

$$v(x+t) \geq \gamma^{-1}\bar{F}(t)e^{-\gamma x}, \qquad (1.3)$$

where

$$v(x+t) = \int_{x+t}^\infty F(z)dz.$$

Definition 1.4

The distribution $F(x)$ is called used better than aged in convex tail ordering (UBACT) if,

$$\Gamma(x+t) \geq \gamma^{-2}\bar{F}(t)e^{-\gamma x} \text{ for } t \geq 0, \qquad (1.4)$$

where,

$$\Gamma(x+t) = \int_{x+t}^\infty V(u)du.$$

The equality in (1.2) is achieved when $F(x)$ has an exponential distribution with mean μ equal to the coefficient of asymptotic decay γ , where the exponential distribution is the only one which has the no aging property.

We can see the details for these definitions in Abu-Youssef and Bakr (2014). Its dual class is used worse than used in convex tail order, denoted by UWACT, which is defined by reversing the above inequality. Then, it is clear that

IHR ⊂ DMRL ⊂ UBA ⊂ UBAC ⊂ UBACT.

See Willmot and Cai (2000).

Well known classes of life distributions include increasing failure rate (IFR), increasing failure rate in average (IFRA), new better than used (NBU), decreasing mean residual life (DMRL) and new better than used in expectation (NBUE). For definitions and properties of these criteria we refer Deshpande et al (1986), Barlow and Proschan (1981), Bryson and Siddique (1969).

Testing exponentially against the classes of life distribution has seen a good deal of attention. For testing against IHR, we refer to Barlow and Proschan (1981) and Ahmad (1994), among others. While testing against DMRL see Ahmad (1992), testing against UBA see Ahmad (2004) and tasting against UBAC see Abu-Youssef (2009), and Mohie El-Din et.al (2013). Finally tasting against UBACT see Abu-Youssef and Bakr (2014).

ISSN 2347-1921

The main object in this paper is to deal with the problem of testing $H_0: F$ is exponentialagainst H_1: Fis the largest class of life distribution UBACT. The paper is organized asfollows: in section 2, we give a test statistic based on U-statistic for complete data. Selectedcritical values are tabulated for sample sizes 5(5)80 is investigated in section 3. The Pitmanasymptotic efficiency for common alternatives is obtained in section 4. In section 5wepropose a test statistic based on U-statistic for censored data. Finally, A numerical examplesin medical science is demonstrated practical application for complete and censored data insection 6.

2 TESTING FOR COMPLETE DATA

The test presented on a sample $X_1, X_2, \ldots, X_n$ from a population with distribution $F(x)$. We wish to test the null hypothesis,

$$H_0 : \bar{F} \text{ is exponential distribution with mean 0, against,}$$

$$H_1 : \bar{F} \text{ is UBACT, and not exponential distribution.}$$

Let the measure of departure from H_0 in favor of H_1 is

$$\delta_{ut} = E[\boldsymbol{\Gamma}(\mathbf{x}+\mathbf{t}) - \gamma^{-2}\bar{F}(t)e^{-\gamma x})],$$

Which gives

$$\delta_{ut} = \int_0^\infty \int_0^\infty (\boldsymbol{\Gamma}(x+t) - \frac{1}{\gamma^2}\bar{F}e^{-\gamma x})dF(x)dF(t), \tag{2.1}$$

Remark that under $H_0: \delta_{ut} = 0$, while under $H_1: \delta_{ut} \geq 0$. Then to estimate δ_{ut} by $\hat{\delta}_{ut}$, let $X_1, X_2, \ldots, X_n$ be a random sample from F and let

$\hat{\boldsymbol{\Gamma}}_n(x) = \frac{1}{2n}\sum_{m=1}^{n}(X_m - x - t)^2 I(X_m > x+t)$ is the empirical distribution of $\Gamma(x)$,

$d\hat{F}_n(x) = \frac{1}{n}$ is the empirical distribution of $dF(x)$, Then,

$$\hat{\delta}_{ut_n} = \int_0^\infty \int_0^\infty (\hat{\boldsymbol{\Gamma}}_n(x+t) - \frac{1}{\gamma^2}\bar{F}_n e^{-\hat{\gamma}x})d\hat{F}_n(x)\, d\hat{F}_n(t),$$

i.e,

$$\hat{\delta}_{ut} = \frac{1}{2n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}(X^2{}_k + X^2{}_i + X^2{}_j + 2X_jX_i - -2X_kX_i - 2X_kX_j)I(X_k > X_i + X_j) - \frac{e^{-\gamma X_i}}{\gamma^2} \tag{2.2}$$

where,

$$I(y>t) = \begin{cases} 1 & \text{if, } y > t \\ 0, & \text{if, o.w.,} \end{cases}$$

let us rewrite (2.2) as the following,

$$\hat{\delta}_{ut} = \frac{1}{2n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\emptyset(X_i, X_j, X_k)$$

where,

$$\emptyset(X_i, X_j, X_k) = [(X_k - X_i - X_j)^2 I(X_k > X_i + X_j) - \frac{1}{\gamma^2}e^{-\hat{\gamma}X_i}].$$

To make the test scale invariant, we take,

$$\hat{\Delta}_{ut} = \hat{\delta}_{ut}/\bar{x}^2, \tag{2.3}$$

Set

$$\emptyset(X_1, X_2, X_3) = [(X_3 - X_1 - X_2)^2 I(X_3 > X_1 + X_2) - \frac{1}{\gamma^2}e^{-\hat{\gamma}X_1}]$$

Then $\hat{\Delta}_{ut}$ in (2.3) is equivalent to the U-statistic.

Theorem 2.1

i) When $n \to \infty$, then $\sqrt{n}(\Delta_{ut} - \hat{\Delta}_{ut})$ is convergence asymptotically normal distribution with mean 0 and variance,

$$\sigma^2 = \mathrm{var}(\int_0^\infty \int_{X+u}^\infty (v - X - u)^2 f(v)f(u)dvdu - e^{-X} + \int_0^\infty \int_{X+u}^\infty (v - u - X)^2 f(v)f(u)dvdu - \int_0^\infty e^{-u} f(u)du + \int_0^X \int_0^{X-u} (X - v - u)^2 f(v)f(u)dvdu - \int_0^\infty e^{-v} f(v)dv) \quad (2.4)$$

ii) Under H_0: $\Delta_{ut_n} = 0$, and $\sigma^2 = \frac{104}{27}$.

iii)If $F(x)$ is continuous UBACT, then the test is consistent.

Proof

(i) and (ii) follow from the standard theory of U-statistics cf. Lee (1990) by direct calculation. To prove part (iii),let $D(x,t) = \Gamma(x+t) - \frac{1}{\gamma^2}\bar{F}e^{-\gamma x}$. Since $F(x)$ UBACT, then $D(x,t) > 0$for at least one value of x, tcall x_0, t_0. Set

$(x_1, t_1) = \mathrm{Inf}\{(x,t): x \geq x_0, t \geq t_0, \bar{F}(x) = \bar{F}(x_0)\}$, thus,

$$D(x_1, t_1) = \Gamma(x_1 + t_1) - \frac{1}{\gamma^2}\bar{F}(t_1)e^{-\gamma x_1} > \Gamma(x_0 + t_0) - \frac{1}{\gamma^2}\bar{F}(t_0)e^{-\gamma x_0} = D(x_0, t_0),$$

and $F(x_1 + \delta) - F(x_1) > 0$.

Since x_1 and t_1 are points of increasing of F, thus $\Delta_{ut} > 0$.

Remark:

The statistic T_n is consistent for the parameter $g(\theta)$ if

$$\lim_{n\to\infty} p[|T_n - g(\theta)|] = 0\,, \varepsilon > 0.$$

i.e.,

$$T_n - g(\theta) as\ n \to \infty.$$

This complete the proof.

MONTE CARLO NULL DISTRIBUTION CRITICAL POINTS

In practice, simulated percentiles for small samples are commonly used by applied statisticians and reliability analyst. We have simulated the upper percentile points for 95%, 98%, 99%. Table 1 gives these percentile points of statistic $\hat{\Delta}_{ut}$ in (2.3) and the calculations are based on 10000 simulated samples of sizes n = 5(5)80. The percentiles values change slowly as n increase.

ISSN 2347-1921

Table 1. Critical values of $\hat{\Delta}_{ut}$

n	95%	98%	99%
5	1.345	1.672	1.891
10	0.855	1.084	1.239
15	0.623	0.810	0.937
20	0.492	0.655	0.764
25	0.401	0.546	0.644
30	0.334	0.466	0.556
35	0.284	0.406	0.489
40	0.242	0.357	0.435
45	0.209	0.318	0.391
50	0.180	0.283	0.352
55	0.148	0.246	0.312
60	0.132	0.226	0.289
65	0.105	0.195	0.256
70	0.087	0.173	0.232
75	0.071	0.155	0.211
80	0.062	0.143	0.198

It is clear from Table 1 that, the percentiles values decreases slowly as the sample size increases where is shown in Figure 1.

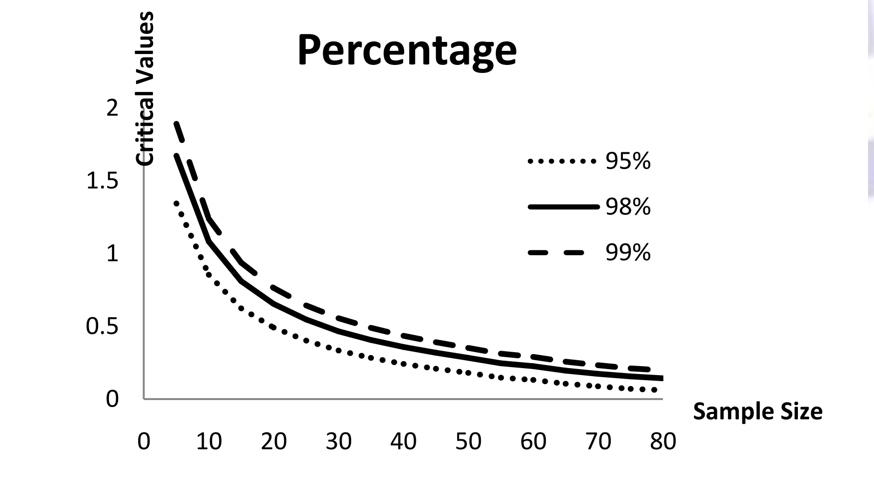

Figure 1. The Relation between sample size and critical values of $\hat{\Delta}_{ut}$

ISSN 2347-1921

ASYMPTOTIC RELATIVE EFFICIENCY (ARE)

Since the above test statistic $\Delta_{ut_n} = \frac{\delta_{ut}}{x^2}$ is new and no other tests are known for these class UBACT. We may compare this to those of the other classes classes. Here we choose the $\hat{\delta}_2$ presented by Ahmad (2004) for (UBAE) class of life distribution and $\hat{\Delta}_k$ presented Mohie El-Din et al (2014) for (UBAC) class of life distribution. The comparisons are achived by using Pitman asymptotic relative efficiency (PARE), which is defined as follows:

Let T_{1n} and T_{2n} be two statistics for testing $H_0 : F_{\theta_x} \in \{F_x\}, \theta_n = \theta + \frac{C}{2\sqrt{n}}$ with C an arbitrary constant, then PARE of T_{1n} relative to T_{2n} is defined by

$$e(T_{1n}, T_{2n}) = \frac{\grave{\mu}_1(\theta_0)}{\sigma_1(\theta_0)} \Big/ \frac{\grave{\mu}_2(\theta_0)}{\sigma_2(\theta_0)} \tag{4.1}$$

where $\grave{\mu}_i = \lim_{n\to\infty} \frac{\partial}{\partial\theta} E(T_{in})_{\theta\to\theta_0}$ and $\sigma_i^2(\theta_0) = \lim_{n\to\infty} var(T_{in}), i = 1, 2.$

Two of the most commonly used alternatives (cf. Hollander and Proschan (1972)) they are:

(i) Linear failure rate family

$$\bar{F}_1(x) = e^{-x-\frac{x^2}{2}\theta}, \quad x, \theta \geqslant 0 \tag{4.2}$$

(ii) Makeham family:

$$\bar{F}_2(x) = e^{-x-\theta(x+e^{-x}-1)}, \quad x, \theta \geqslant 0 \tag{4.3}$$

Note that H_0 (the exponential disribution) is attained at θ = 0 in (i) and (ii). The Pitman's asymptotic efficiency (PAE) of $\hat{\Delta}_{ut}$ is equal to

$$\text{eff}_F = \frac{\left|\frac{\partial}{\partial\theta}\Delta\ |_{\theta=\theta_0}\right|}{\sigma_0} \tag{4.4}$$

Direct calculations of PAE of $\hat{\delta}_2$ and $\hat{\Delta}_k$are summarized in table (2), the efficienciesin table (2) shows clearly our U-statistic $\hat{\Delta}_{ut}$perform well for $F_1$and F_2

Table 2.PAE of $\hat{\delta}_2$&$\hat{\Delta}_{UK}$and $\hat{\Delta}_K$

Distribution	$\hat{\delta}_2$	$\hat{\Delta}_{UK}$	$\hat{\Delta}_{ut}$
F_1 Linear failure rate	0.630	0.565	0.748
F_2 Makeham	0.385	0.245	0.248

In Table 3, we give PARE.s of $\hat{\Delta}_{ut}$ with respect to $\hat{\delta}_2$ and $\hat{\Delta}_{UK}$ whose PAE are mentioned in Table 2

Table 3.PARE of $\hat{\Delta}_K$ with respect to $\hat{\delta}_2$ and $\hat{\Delta}_{UK}$

Distribution	$eff_i(\hat{\Delta}_{ut}, \hat{\delta}_2)$	$eff_i(\hat{\Delta}_{ut}, \hat{\Delta}_{UK}$
F_1 Linear failure rate	1.2	1.3
F_2 Makeham	0.7	1.01

It is clear from Table 3 that the statistic $\hat{\Delta}_{ut}$perform well for $\bar{F}_1$ and $\bar{F}_2$ and it is more efficient than both $\hat{\delta}_2$ and $\hat{\Delta}_{UK}$ for all cases mentioned above. Hence our test, which deals the much larger UBAC is better and also simpler.

TESTING FOR CENSORED DATA

In this section, a test statistic is proposed to test

H_0 ($\bar{F}$is exponential distribution with mean μ) versus

H_1 ($\bar{F}$is UBACT and not exponential distribution); with randomly right-censored data.

Such a censored data is usually the only information available in a life-testing model or in a clinical study where patients may be lost (censored) before the completion of a study. This experimental situation can formally be modeled as follows:

Suppose n objects are put on test, and $X_1, X_2, \ldots, X_n$ denote their true life time. We assume that $X_1, X_2, \ldots, X_n$ be independent, identically distributed (i.i.d.) according to a continuous life distribution F.

Let$Y_1, Y_2, \ldots, Y_n$ be (i.i.d.) according to a continuous life distribution G and assume that X′s and Y's are independent. In the randomly right-censored model, we observe the pairs (Z_i, δ_i), $i = 1, \ldots, n$, where $Z_i = \min(X_i, Y_i)$ and

$$\delta_i = \begin{cases} 1 & \text{if } \; Z_i = X_i \text{ (i th observation is uncensored)} \\ 0 & \text{if } \; Z_i = Y_i \text{(i th observation is censored).} \end{cases}$$

Let $Z_{(0)} < Z_{(1)} < \ldots < Z_{(n)}$ denoted the ordered of Z's and δ_i is the δ corresponding to $Z_{(i)}$, respectively. Using the Kaplan and Meier estimator in the case of censored data

(Z_i, δ_i), $i = 1, \ldots, n$, then the proposed test statistic is given by (2.3) can be written using right censored data as

$$\hat{\delta}_K^c = \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{f}(x)\left[\widehat{\Gamma}_n(x+t) - \bar{F}_n(t)e^{-\gamma Z_{(j)}}\right]\left[\prod_{p=1}^{i-2} C_i^{\delta_i} - \prod_{p=1}^{i-1} C_i^{\delta_i}\right]\left[\prod_{q=1}^{j-2} C_i^{\delta_i} - \prod_{q=1}^{j-1} C_i^{\delta_i}\right] \quad (5.1)$$

Where

$$\widehat{\Gamma}_n(x+t) = \int_x^\infty \int_z^\infty \bar{F}_n(u+t)dudt = \int_x^\infty \int_{z+t}^\infty \bar{F}_n(u)dudt \quad (5.2)$$

$$= \int_x^\infty \left[\hat{\mu} - \sum_{k=1}^{l}\prod_{m=1}^{k-1} C_m^{\delta_m}(Z_k - Z_{k-1})\right]$$

$$l = i + j \quad if \;\; Z_i + Z_j < Z_n$$

$$l = n \quad if \;\; Z_i + Z_j > Z_n$$

$$\hat{\mu} = \sum_{j=1}^{l}\prod_{k=1}^{j-1} C_k^{\delta_k}(Z_{(j)} - Z_{(j-1)}),$$

$$dF_n(Z_i) = \prod_{q=1}^{j-2} C_i^{\delta_i} - \prod_{q=1}^{j-1} C_i^{\delta_i},$$

$$\hat{f}(x) = \sum_{k=1}^{l} \delta_k K(X - Z_{(k)})$$

$$\bar{F}_n(t) = \prod_{m<t} C_m^{\delta_m},$$

$$C_m = \frac{n-m}{n-m+1}, \quad t \in [0, z_{(m)}].$$

Table 4 shows the critical values percentiles $\hat{\delta}_{ut}^c$ for sample size n= (61)(10)(201) and Figure 2 shows the relation between the sample size and critical values in the case of censored data.

Table 4. Critical values of $\hat{\Delta}_K^c$

n	95%	98%	99%
61	3.894	3.988	4.050
71	1.990	2.077	2.135
81	1.313	1.394	1.448
91	0.983	1.059	1.110
101	0.792	0.864	0.913
11	0.670	0.739	0.786
121	0.586	0.652	0.697
131	0.524	0.588	0.631
141	0.478	0.539	0.580
151	0.441	0.500	0.540
161	0.411	0.468	0.507
171	0.387	0.442	0.480
181	0.366	0.420	0.456
191	0.348	0.401	0.436
201	0.333	0.384	0.419

It is clear from Table 1 that, the percentiles values decreases slowly as the sample size increases where is shown in Figure 2.

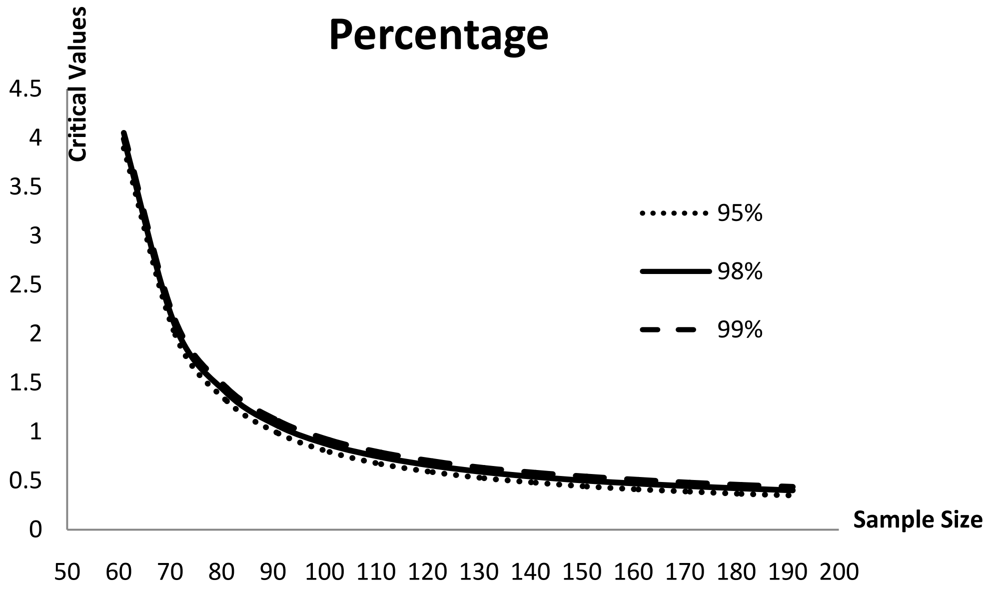

Figure 2. The relation between sample size and critical values

ISSN 2347-1921

Applications

1. **Applications for complete data**

Example 1

The following data represent 39 liver cancers patients taken from El Minia Cancer Center Ministry of Health Egypt (Attia (2004). The ordered life times (in days) are:

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

It is found that the test statistics for the set data by using equation (2.3) is $\hat{\Delta}_{ut} = 1.71113 * 10^7$, which is greater than the crossposting critical value of the table (1) is (0.242). Thenwe accept H_1which states that the set of data have UBACT property under significant level$\alpha = 0.05$.

Example 2.

In an experiment at Florida state university to study the effect of methylmercury poisoning on the life lengths of fish goldfish were subjected to various dosages ofmethyl mercury (Kochar (1985)). At one dosage level the ordered times to death in weekare:

6	6.143	7.286	8.714	9.429	9.857	10.143	11.571	11.714	11.714

It is found that the test statistics for the set data by using equation (2.3) is $\hat{\Delta}_{ut} = 0.648704$,which is smaller than the crossposting critical value of the table (3 1) (0 855). Then weacceptH_0which states that the set of data do not have UBACT property under significantlevel $\alpha = 0.05$.

2. Applications for censored data

Example 3

On the basis of right-censored data for lung cancer patients from Pena (2002). These data consists of 86 survival times (in month) with 22 right censored. The whole life times

i) Non-censored data

0.99	1.28	1.77	1.97	2.17	2.63	2.66	2.76	2.79	2.86
2.99	3.06	3.15	3.45	3.71	3.75	3.81	4.11	4.27	4.34
4.40	4.63	4.73	4.93	4.93	5.03	5.16	5.17	5.49	5.68
5.72	5.85	5.98	8.15	8.62	8.48	8.61	9.46	9.53	10.05
10.15	10.94	10.94	11.24	11.63	12.26	12.65	12.78	13.18	13.47
13.96	14.88	15.05	15.31	16.13	16.46	17.45	17.61	18.20	18.37
19.06	20.70	22.54	23.36						

ii) Censored data

11.04	13.53	14.23	14.65	14.91	15.47	15.47	17.05
17.28	17.88	17.97	18.83	19.55	19.55	19.75	19.78
19.95	20.04	20.24	20.73	21.55	21.98		

It is found that the test statistics for the set of data $\hat{\Delta}^c_{ut} = 0.682989$. Then we accept$H_0$ which states that the set of data do not have UBACT property under significant level $\alpha = 0.05$.

REFERENCES

1. Abu-Youssef, S.E. (2009). Nonparametric Test for Used Better Than Aged in ConvexOrdering Class (UBAC) of Life Distributions with Hypothesis Testing,*Intentional Journal ofReliability and Applications*, **10**, 81-88.
2. Abu-Youssef, S. E. and Bakr, M.E. (2014). Some properties of UBACT class of lifedistribution,*Journal of Advanced Research in Applied Mathematics*, **6**, 1-9.
3. Abu-Youssef, S. E. and Bakr, M.E. (2014). A goodness of fit approach to class of lifedistribution with unknown age,International Journal of Computer Applications, 30-35
4. Abu-Youssef, S. E., B. I. MohammedandBakr, M.E. (2014). Nonparametric test for unknown age class of life distributions, International Journal of Reliability and Applications, 99-110.
5. Ahmed, I.A. (2004). Some Properties of Classes of Life Distributions with UnknownAge, *Statistics and Probability Letters*, 333-342.
6. Barlow ,R. E. and Doksum, K. A. Isoomic (1972). Tests for Covex Ordering,*Proceeding of 6th Berkeley Symposium*, 93-323.
7. Barlow, R. E. andProschan, F. (1981).*Statistical Theory of Reliability and Life TestingProbability Models*, To Begin With, Silver-Spring, MD.
8. Deshpand, J.V., Kocher, S.C.and Singh, H. (1986).Aspects of Positive aging, *Journal of Applied Probability*, 1472-1483.
9. Kanjo, A. J. (1993). Testing for new is better than used, *Communications in Statistics-Theory and Methods*,311-321.
10. Kochar, S.C. (1985). Testing exponentiality against monotone failure rate average, *Communications in Statistics-Theory and Methods*, 381-392.
11. Mohie El-Din, M.M., Abu-Youssef, S.E. and Hassan, M.KH. (2014). On testing ofunknown age classes of life distributions, submitted.
12. Willmot, G.E. and Cai, J. (2000). On Classes of Lifetime Distributions with unknownage, *Probability in the Engineering and Informational Sciences*.473-484.