A Goodness of Fit Approach to the Unknown Age (UBACT) Class of Life Distribution

S. E. Abu-Youssef
Department of Mathematics Faculty of Science
Al-Azhar University

M. E. Bakr
Department of Basic Science
Thebes Higher Institute for Engineering

ABSTRACT
Based on the goodness of fit approach, a new test is presented for testing exponentiality versus used better than aged in convex tail ordering UBACT class of life distribution. The percentiles of this test are tabulated for sample sizes \(n=1(5)100\). It is shown that the proposed test is simple and it has high relative efficiency for some commonly used alternatives. A numerical example in medical science demonstrates practical application of the proposed test.

Keywords
U-Statistics; Goodness of fit approach; UBACT class of life distribution; Hypothesis testing.

1. INTRODUCTION
A significant part of life testing problems is concerned with testing whether a life distribution belongs to a nonparametric family of aging.

In reliability theory, ageing life is usually characterized by a nonnegative continuous random variable \(X \geq 0\) representing equipment life with distribution function \(F\) and survival function \(1-F(x)\); given a unit which has survived up to time \(t\), its additional residual life \(X_t\), with distribution function \(F(x|t)\) and survival function, \(F(x|t) = \frac{F(x+t)}{F(t)}\), \(x, t > 0\)

Assume that \(X\) has a finite mean \(\mu = E(X) = \int_0^\infty F(u)du\).

Some properties concerning the asymptotic behavior of \(X_t\) as \(t \rightarrow \infty\) will be used.

Definition (1.1): age-smooth
If \(X\) is non-negative random variable, its distribution function \(F\) is said to be finitely and positively age-smooth if a number \(\gamma \in (0,1)\) exists such that: \(\lim_{t \rightarrow \infty} \frac{F(x+t)}{F(t)} = e^{-\gamma x}\)

Where \(\gamma\) is called the asymptotic decay coefficient of \(F\).

Definition (1.2): The distribution function \(F\) is said to be used better than aged (UBA) if it satisfies, (see Alzaid (1994))

\[
\int_t^\infty \frac{F(u)}{u} \geq \frac{F(t)}{t} \quad (1.2)
\]

which can be written as

\[
v(u) \geq \frac{F(t)}{u} \quad \text{Where} \quad v(u) = \int_t^\infty \frac{F(u)}{u} du
\]

Definition (1.4): The distribution function \(F\) is said to be used better than aged in convex order (UBAC) class of life distribution if we integrate the both sides of \((1.2)\) equation with respect to \(u\), (see Al-Nachawati and Alwasel (1997))

\[
\int_t^\infty \frac{F(u+t)}{u} du \geq \frac{F(t)}{u} e^{-\gamma x}, \quad t \geq 0 \quad (1.3)
\]

Many applications in reliability theory involve the modeling of life time data. In these applications, for the purchasing used items with unknown age, it may be realistic for the buyer to assume that those items have been used for a long period of time. Hence, it would be of a great importance to have some criteria to compute the remaining life of the purchased item with its performance under the true age. As a criterion for comparing ages of, for instance, electrical equipments, computers, radios or alike, Bhattacharjee (1986) discussed the tail behavior of age smooth failure distributions. Cline (1987) studied the connection between the class of age smooth distributions and the classes of life distributions with subexponential tails which have many applications in queuing theory, random walk and infinite divisibility.

Definition (1.5): The distribution \(F\) is called used better than aged in convex tail ordering UBACT if it is finitely and positively smooth and satisfies, (see Abu-Youssef and Bakr (2014))

\[
\int_x^\infty \int_z^\infty \frac{F(u+t)}{u} du dz \geq \gamma^{-2} F(x)e^{-\gamma x}, \quad t \geq 0 \quad (1.4)
\]

This can be written as,

\[
\Gamma(x + t) \geq \gamma^{-2} F(x) e^{-\gamma x}, \quad t \geq 0
\]

Where,

\[
\Gamma(x + t) = \int_x^\infty v(u) du, \quad \text{and} \quad v(x + t) = \int_x^\infty \frac{F(u)}{u} du
\]

Its dual class is used worse than used in convex tail order, denoted by UWACT, which is defined by reversing the above inequality.
Theorem (1.1):
If \( F \) is UBA class of life distribution, then \( F \) is UBAC class of life distribution, i.e.
\[
F \in \text{UBA} \Rightarrow F \in \text{UBAC}.
\]

Proof: \( \forall F \) is UBA, then by Integrate equation (1.1) as,
\[
\int_{x}^{\infty} F(u + t)du \geq \frac{F(t)}{\gamma} e^{-\gamma x}
\]
which gives,
\[
\int_{x}^{\infty} F(u + t)du \geq \frac{F(t)}{\gamma} e^{-\gamma x}
\]
Put \( y = u + t \), to get
\[
\int_{x+y}^{\infty} F(y)dy \geq \frac{F(t)}{\gamma} e^{-\gamma y},
\]
Taking \( v(x + t) = \int_{x+y}^{\infty} F(y)dy \), then
\[
v(x + t) \geq \frac{F(t)}{\gamma} e^{-\gamma y}.
\]
\( \therefore F \in \text{UBAC} \).

Theorem (1.2):
If \( F \) is UBC class of life distribution, then \( F \) is UBACT class of life distribution, i.e.
\[
F \in \text{UBAC} \Rightarrow F \in \text{UBACT}.
\]

Proof: \( \forall F \) is UBAC, then by Integrate equation (1.3) as,
\[
\int_{x}^{\infty} v(x + t)dx \geq \frac{F(t)}{\gamma} \int_{x}^{\infty} e^{-\gamma y}dx
\]
this gives,
\[
\int_{x}^{\infty} v(x + t)dx \geq \frac{F(t)}{\gamma^2} e^{-\gamma x}
\]
Put \( u = x + t \), to get
\[
\int_{x+t}^{\infty} v(u)du \geq \frac{F(t)}{\gamma^2} e^{-\gamma u},
\]
Taking \( \Gamma(z + t) = \int_{x+t}^{\infty} v(u)du \), then
\[
\Gamma(z + t) \geq \frac{F(t)}{\gamma^2} e^{-\gamma z}.
\]
\( \therefore F \in \text{UBACT} \).

\( F(x) \) has an exponential distribution if its mean \( \mu \) equal to the coefficient of the asymptotic decay \( \gamma \), and the exponential distribution is the only which has the lack of memory property.

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 definition and properties of these criteria we refer Deshpande et al (1986), Bryson and Siddique (1996), Alzaid (1994) discussed two classes of life distributions, namely used better than aged UBA and used better than aged in expectation UBAE. Willmot and Cai (2000) showed that the UBA class includes the decreasing mean residual life DMRL class. While Crescenzo (1996) has shown that

\[
\text{UBAE} \subset \text{HUBAE}
\]
and from theorem (1.1) and (1.2) we have the following

\[
\text{DMRL} \subset \text{UBA} \subset \text{UBAC} \subset \text{UBACT}
\]

\( \cup \)

\( \text{HUBAE} \subset \text{UBAE} \)

Testing exponentially against the classes of life distribution has seen a good deal of attention. For testing against IFR, we refer to Barlow and Proschan (1981), and Ahmad (1994), among others. While testing against DMRL see Ahmad (1992), and testing against UBA see Ahmad (2004). Abu-Youssef (2009) introduce a new test used better than aged in convex ordering of life distribution UBAC.

Recently Abu-Youssef, El-Bassiouny and El-Wasel and Ahmad et al. used a new methodology for testing by incorporating both HUBA and HUBAE in devising the test statistics for testing \( H_0 \) against the alternative; the life distribution is IFR, NBUC, HNBUE (harmonic new better than in expectation) and DMRL classes of life distributions. They obtained very simple statistics that are not only asymptotically equivalent in distribution and efficiency to classical procedures but also better in finite sample behaviors.

Our goal in this paper is to use similar methodology to obtain a very simple statistic for testing \( H_0 \) against \( H_1 \). The thread that connects most work mentioned here is a measure of departure from \( H_0 \), which is strictly positive under \( H_0 \) and is zero under \( H_1 \). Then, a sample version of this measure is used as test statistic, and its properties are studied. In section 1, we propose a test statistic, based on the goodness of fit approach, for testing \( H_0 \); \( F \) is exponential against \( H_1 \); \( F \) is the UBACT. Also, in section2, we present Monte Carlo null distribution critical points for sample size \( n = 1(5)100 \), in section 3. In section 4, we calculate the efficiency of the test statistic for some common alternatives and compared them to other procedures. Finally, an application in medical science is introduced in section 5.

2. TESTING UBACT CLASS OF LIFE DISTRIBUTION

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

\( H_0 : F \) is exponential distribution with mean \( \mu \), against,

\( H_1 : F \) is UBACT and not exponential distribution.

From (1.4) let the measure of departure from \( H_0 \) in favor of \( H_1 \) is given by
To make the test scale invariant, we take

\[ \Delta_0 = \frac{\delta_0}{\mu^2} \]

Theorem (2.1):

i) Let \( X \) be a non-negative random variable with distribution function \( F \), then

\[ \delta_0 = E \left[ \frac{1}{2} X^2 - 3e^{-X}Xe^{-X} - 2X + 3 - \frac{1 - e^{-X}}{y^2(y + 1)} \right] \]

ii) As \( n \to \infty, \sqrt{n}(\delta_0, \Delta_0) \) is asymptotically normal distribution with mean \( \mu = 0 \) and variance

\[ \sigma^2 = \text{var} \left( \frac{1}{2} X^2 - 3e^{-X}Xe^{-X} - 2X + 3 - \frac{1 - e^{-X}}{y^2(y + 1)} \right) \]

iii) Under \( H_0 : \delta_0 = 0 \) and \( \sigma^2 = 0.96296 \).

Proof: (i) We can write (2.1) as

\[ \delta_0 = E \left[ \int_0^\infty \int_0^\infty \frac{1}{2} (X - x - t)^2 I(X > x + t) \right] - \frac{1}{\mu^2} \int_0^\infty I(X > t)e^{-\mu t} \int e^{-(x+t)}dxdt \] (2.1)

Which it can be written as

\[ \delta_0 = E[I_1 - I_2] = E[I_1] - E[I_2] \] (2.2)

Which

\[ E[I_1] = E \left[ \int_0^\infty \int_0^\infty \frac{1}{2} (X - x - t)^2 I(X > x + t) e^{-(x+t)}dxdt \right] \]

And

\[ E[I_2] = E \left[ \int_0^\infty \int_0^\infty \frac{1}{2} I(X > t) e^{-\mu t} e^{-(x+t)}dxdt \right]. \]

Then, by direct integration

\[ E[I_1] = \int_0^X \int_0^X \frac{1}{2} (X - x - t)^2 e^{-(x+t)}dxdt \]

\[ = E \left[ \frac{1}{2} X^2 - 3e^{-X}Xe^{-X} - 2X + 3 \right] \] (2.3)

And

\[ E[I_2] = \int_0^\infty \int_0^\infty \frac{1}{\mu^2} e^{-\mu t} e^{-(x+t)} = E \left[ \frac{1 - e^{-X}}{y^2(y + 1)} \right] \] (2.4)

Substitute from (2.3) (2.4) in (2.2) we have

\[ \delta_0 = E \left[ \frac{1}{2} X^2 - 3e^{-X}Xe^{-X} - 2X + 3 - \frac{1 - e^{-X}}{y^2(y + 1)} \right] \]

\[ \sigma^2 = \text{var} \left( \frac{1}{2} X^2 - 3e^{-X}Xe^{-X} - 2X + 3 - \frac{1 - e^{-X}}{y^2(y + 1)} \right) \]

(iii) \( \delta_0 \) Under \( H_0 \) is

\[ \delta_0 = \int_0^\infty \left( \frac{1}{2} X^2 - 3e^{-X}Xe^{-X} - 2X + 3 - \frac{1 - e^{-X}}{2} \right) e^{-X}dX = 0 \]

And the variance of \( \delta_0 \) is

\[ \sigma^2 = E(\delta_0)^2 - (E(\delta_0))^2. \]

So

\[ \sigma^2 = \int_0^\infty \left( \frac{1}{2} X^2 - 3e^{-X}Xe^{-X} - 2X + 3 - \frac{1 - e^{-X}}{2} \right)^2 e^{-X}dX = 0.96 \]

3. MONTE CARLO NULL DISTRIBUTION CRITICAL POINTS FOR \( \Delta_0 \) TEST

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%, 97%, 99%. Table (3.1) gives these percentile points of statistic \( \Delta_0 \) in (2.1) and the calculations are based on 10000 simulated samples of sizes \( n = 1(5)100 \). The percentiles values change slowly as \( n \) increase. To use the above test, calculate \( \sqrt{n}(\delta_0, \Delta_0) \) and reject \( H_0 \) if this exceeds the normal variety value \( Z_{n-1} \)

<table>
<thead>
<tr>
<th>n</th>
<th>95%</th>
<th>98%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.829</td>
<td>2.208</td>
<td>2.465</td>
</tr>
<tr>
<td>6</td>
<td>0.638</td>
<td>0.793</td>
<td>0.898</td>
</tr>
<tr>
<td>11</td>
<td>0.423</td>
<td>0.537</td>
<td>0.615</td>
</tr>
<tr>
<td>16</td>
<td>0.322</td>
<td>0.417</td>
<td>0.481</td>
</tr>
<tr>
<td>21</td>
<td>0.257</td>
<td>0.340</td>
<td>0.396</td>
</tr>
<tr>
<td>26</td>
<td>0.231</td>
<td>0.288</td>
<td>0.338</td>
</tr>
<tr>
<td>31</td>
<td>0.177</td>
<td>0.245</td>
<td>0.291</td>
</tr>
<tr>
<td>36</td>
<td>0.152</td>
<td>0.215</td>
<td>0.258</td>
</tr>
<tr>
<td>41</td>
<td>0.129</td>
<td>0.188</td>
<td>0.228</td>
</tr>
<tr>
<td>46</td>
<td>0.112</td>
<td>0.168</td>
<td>0.205</td>
</tr>
<tr>
<td>51</td>
<td>0.096</td>
<td>0.149</td>
<td>0.185</td>
</tr>
<tr>
<td>56</td>
<td>0.085</td>
<td>0.136</td>
<td>0.170</td>
</tr>
<tr>
<td>61</td>
<td>0.074</td>
<td>0.122</td>
<td>0.155</td>
</tr>
</tbody>
</table>
It is clear from Table (3.1) that, the percentiles values decreases slowly as the sample size increases, which is shown in Figure (3.1) the relation between the sample size and critical values in the case of censored data \( \tilde{\delta} \).

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>66</th>
<th>71</th>
<th>76</th>
<th>81</th>
<th>86</th>
<th>91</th>
<th>96</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentile</td>
<td>0.063</td>
<td>0.053</td>
<td>0.046</td>
<td>0.039</td>
<td>0.030</td>
<td>0.027</td>
<td>0.022</td>
</tr>
<tr>
<td>Critical Value</td>
<td>0.110</td>
<td>0.098</td>
<td>0.090</td>
<td>0.081</td>
<td>0.070</td>
<td>0.067</td>
<td>0.060</td>
</tr>
<tr>
<td>Percentage</td>
<td>0.142</td>
<td>0.128</td>
<td>0.119</td>
<td>0.110</td>
<td>0.098</td>
<td>0.094</td>
<td>0.087</td>
</tr>
</tbody>
</table>

Table (3.1): Critical Values of \( \tilde{\delta} \)

4. ASYMPTOTIC RELATIVE EFFICIENCY (ARE)

Since the above test statistic \( \Delta_G = \frac{\delta_G}{\mu} \) is new and no other tests are known for this class (UBACT), We may compare our test to the other classes. Here we choose the test \( \delta_G \) presented by Ahmad (2004) for (UBAE) class of life distribution and \( \Delta_G \) presented Mohie El-Din et al for (UBAC) class. Then comparisons are achived by using Pitman asymptotic relative efficiency (PAE) which is defined as follows:

Let \( T_{1n} \) and \( T_{2n} \) be two statistics for testing

\[
H_0: F_{\theta_0} \in \{ F_x \}, \theta_n = \theta + \frac{C}{\sqrt{n}}
\]

With \( C \) an arbitrary constant, then PAE of \( T_{1n} \) relative to \( T_{2n} \) is defined by

\[
e(T_{1n}, T_{2n}) = \frac{\mu_1(\theta_0)}{\mu_2(\theta_0)} \frac{\sigma_1(\theta_0)}{\sigma_2(\theta_0)}
\]

Where

\[
\mu_1(\theta_0) = \lim_{n \to \infty} \frac{\partial}{\partial \theta} \mathbb{E} (T_{1n}) \bigg|_{\theta = \theta_0}.
\]

and

\[
\sigma_2^2(\theta_0) = \lim_{n \to \infty} \operatorname{var}(T_{1n}).
\]

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

(i) Linear failure rate family

\[
\tilde{F}_1(x) = e^{-x^{-\theta}}, \quad \theta, x \geq 0. \tag{4.1}
\]

(ii) Makeham family:

\[
\tilde{F}_2(x) = e^{-x-\theta(x+e^{-\theta})}, \quad \theta, x \geq 0. \tag{4.1}
\]

Note that \( H_0 \) (the exponential distribution) is attained at \( \theta = 0 \) in (i) and (ii). The Pitman's asymptotic efficiency (PAE) of \( \Delta_G \) is equal to

\[
\text{eff}_F = \frac{\left| \frac{\partial}{\partial \theta} \right|_{\theta = \theta_0}}{\sigma_0}
\]

By substituting of \( \Delta_G \) we get

\[
\delta_G(\theta) = \left. \frac{\partial}{\partial \theta} \Delta \right|_{\theta = \theta_0}
\]

And by differentiation

\[
\delta_G(\theta) = \left. \int_{x_0}^{x} \left( \tilde{F}_0(x+u) - \frac{\tilde{F}_0(t)}{y} e^{-uy} \right) e^{-x+u} du \right|_{x = x_0} \tag{4.4}
\]

Below is a necessary calculations to get the Pitman asymptotic relative efficiency (PAE) of the test \( \delta_G \) for the distributions (i), (ii) which has been mentioned frequently used in reliability and life tests.

(i) linear failure rate distribution.

From equations (4.1), we get at \( \theta_0 = 0 \)

\[
\tilde{F}_1(x) = e^{-x^2} \quad \Rightarrow \quad \tilde{F}_0(x) = \frac{-x^2}{2} e^{-x}. \tag{4.5}
\]

And

\[
\Gamma_{\theta}(x+u) = \int_{-\infty}^{x+u} \tilde{F}_0(t) dt \quad u \geq 0\]

So

\[
\Gamma_{\theta}(x+u) = \int_{-\infty}^{x} \int_{x+u}^{x+u} \frac{-u^2}{2} e^{-u^2} du \quad u \geq 0.
\]
When $\theta_0 = 0$ we get
\[
\gamma_0'(x + t) = \int_x^\infty \int_{x+t}^\infty \frac{u^2}{2} e^{-u} du dz
= \frac{1}{2} e^{-t^2}(t^2 + 2tx + 4t + x^2 + 4x + 6).
\] (4.6)

By substitution for the function (4.5) and (4.6) in (4.4) we find,
\[
\delta'(\theta) = \int_0^\infty \int_0^x e^{-2(t+x)} \left( t^2 + 2tx + 4t + x^2 + 4x + \frac{t^2}{2} \right) dx dt = \frac{47}{32}.
\]

And we get the asymptotic relative efficiency from equation (4.3)
\[
eff_F = \frac{47}{26} = 1.4967 \quad (4.7)
\]

(ii) Makeham distribution.
From equation (4.2) we get at $\theta_0 = 0$
\[
F_0(x) = (1 - x - e^{-x})e^{-x-\theta(x+x^{-1})} \Rightarrow F_0(x) = (1 - x - e^{-x})e^{-x}.
\] (4.8)

And
\[
\gamma_0'(x + t) = \int_x^\infty \int_{x+t}^\infty (1 - u - e^{-u}) e^{-u-\theta(u+e^{-u}-1)} du dz.
\]

When $\theta_0 = 0$ we get
\[
\gamma_0'(x + t) = \int_x^\infty \int_{x+t}^\infty (1 - u - e^{-u}) e^{-u} du dz
= e^{-t-x} \left( -1 - \frac{1}{4} e^{-t-x} - t - x \right).
\] (4.9)

By substitution for the function (4.8) and (4.9) in equation (4.3) we find
\[
\delta'(\theta) = \int_0^\infty \int_0^x e^{-2(t+x)} \left( -1 - \frac{1}{4} e^{-t-x} - t - x - (1 - t - e^{-t}) \right) dx dt = \frac{35}{72}.
\]

From equation (3.3)
\[
eff_F = \frac{35}{72} = 1.4967. \quad (4.10)
\]

The null hypothesis is at $\theta = 0$ for linear failure rate and Makeham distributions respectively. Direct calculations of PAE of $\Delta_2$, $\Delta_k$ and $\Delta_G$ are summarized in table (4.1), the efficiencies in table (4.2) shows clearly our U-statistic $\Delta_G$ perform well for $F_1$ and $F_2$.

\[
\begin{array}{|c|c|c|}
\hline
\text{Distribution} & \Delta_2 & \Delta_k & \Delta_G \\
\hline
F_1 & 0.630 & 0.565 & 1.4967 \\
F_2 & 0.385 & 0.245 & 0.49537 \\
\hline
\end{array}
\]

Table (4.1) PAE of $\Delta_2$, $\Delta_k$ and $\Delta_G$.

In table (4.2), we give PARE's of $\Delta_G$ with respect to $\Delta_2$ and $\Delta_k$ whose PAE are mentioned in table (4.1).

\[
\begin{array}{|c|c|c|}
\hline
\text{Distribution} & e(\Delta_2, \Delta_2) & e(\Delta_2, \Delta_k) \\
\hline
F_1 & 2.38 & 2.56 \\
F_2 & 1.29 & 2.02 \\
\hline
\end{array}
\]

Table (4.2): PARE of $\Delta_G$ with respect to $\Delta_2$ and $\Delta_k$.

It is clear from table (4.2) that the statistic $\Delta_G$ perform well for $F_1$ and $F_2$ and it is more efficient than both $\Delta_2$ and $\Delta_k$ for all cases mentioned above. Hence our test, which deals the much larger UBAC is better and also simpler.

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

10 14 14 14 14 14 14 15 17 18 20 20 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 (2.1) is $H_1$, which is greater than the cross ponding critical value of the table (3.1) is $(0.129)$. Then we accept $H_1$ which states that the set of data have UBACT property under significant level $\alpha = 0.05$. Therefore the data has UBACT Property.

5. CONCLUSION
Testing exponentiality versus the unknown age (UBACT) is proposed. The percentiles of proposed test are tabulated Table (3.1), which clearly show the critical values of the estimate of departure from measure is decreasing as the sample size increasing. This test is more efficient than the test for (UBAC). It gives a very good power for most alternatives.

6. ACKNOWLEDGEMENTS
The authors would like to thank the editors and the referees for their valuable comments and suggestions.

7. REFERENCES


