Truncated Inverted Kumaraswamy: Estimation and Predication Based on Maximum Likelihood Method

Mervat khalifa Abd Ellah Abd Elaal
Associate Professor, Department of Statistics, Faculty of Sciences, KAU

Truncated Inverted Kumaraswamy: Estimation and Predication Based on Maximum Likelihood Method

Mervat khalifa Abd Ellah Abd Elaal

Associate Professor,
Department of Statistics,
Faculty of Sciences, KAU

Truncation arises in many situations when failure of a unit is observed only if it fails before and/or after a certain period; since any unit, either industrial or biological cannot operate in the same condition forever. A random variable is said to be truncated if it is observed over part of its range where it can occur in various situations. For instance, in survival analysis, failures during the warranty period may not be counted. Items may also be replaced after certain time following the replacement policy, so that failures of the item are ignored.

Keywords: Truncation, survival analysis, survival analysis.

1. Introduction

Truncation arises in many situations when failure of a unit is observed only if it fails before and/or after a certain period; since any unit, either industrial or biological cannot operate in the same condition forever. A random variable is said to be truncated if it is observed over part of its range where it can occur in various situations. For instance, in survival analysis, failures during the warranty period may not be counted. Items may also be replaced after certain time following the replacement policy, so that failures of the item are ignored.

Double truncated distributions are used in cases where the occurrences are limited to values which lie above or below a given threshold or within a specified range. If occurrences are limited to values which lie below a given threshold, the lower (left) truncated distribution is obtained. Similarly, if occurrences are limited to values which lie above a given threshold, the upper (right) truncated distribution is obtained. [See Zhang and Xie (2011) and Isik*et al* (2017)]

Right truncation happens when studying of life testing and reliability of items such as an electronic component, light bulbs, etc.For instance,in industry, sometimes a minimum time; c, is required before which no failure occurs. This minimum time c is known as the guarantee time. Another use for right truncation at c is in epidemiological or biomedical applications where c may represent the latent period of some disease. For example, in cancer research problems, c is regarded as the time

elapsed between first exposure to carcinogen and theappearance of the tumor. [See Ateya and AL-Hussaini (2007)]

Left truncation takes place when failure of a unit is observed only if it fails after a certain period. Sometimes units may not be followed at the beginning of an experiment until all of them fail and the experimenter may have to start at a certain time and stop at a certain time when some of the units may still be working. [SeeOkasha and Alqanoo (2014)]

Many researchers who are interested in analyzing truncated data encountered in different fields, proposed the truncated versions of the usual statistical distributions.The truncated distributions iswide applicable e.g. to improve a forecasting actuarial model and particularly for modelling data from insurance payments that establish a deductible, to study the waiting times before service of the banks' customers,to the statistical analysis of masses of stars and of diameters of asteroids, to analyze the diameter data of trees truncate data specific threshold level, to predict the height distribution of small trees based on incomplete laser scanning data, to modelling the diameter distribution of forest, to characterize the observed Portuguese fire size distribution, to seismological data, on the development of the pit depths on a water pipe, etc.[See Singh *et al.*(2014)]

Some references in the field of the truncation distributions include AL-Yousef (2002), AL-Hussaini*et al*. (2006), Ateya and AL-Hussaini (2007), Nadarajah (2008),Zhang and Xie (2011), Raschke (2012), Okasha and Alqanoo (2014), Singh *et al.*(2014), Isik*et al* (2017) and Al-Omari (2018).

Kumaraswamy (1980) proposed a two-parameter Kumaraswamy distribution on (0, 1) denoted by Kum(*α, β*) with*probability density function* (pdf) and *cumulative distribution function* (cdf) as follows:

$$g(y; \alpha, \beta) = \alpha\beta y^{\alpha-1}(1 - y^\alpha)^{\beta-1}, \qquad 0 < y < 1, \tag{1}$$

and

$$G(y; \alpha, \beta) = 1 - (1 - y^\alpha)^\beta, \qquad 0 < y < 1. \tag{2}$$

Where α and β are the shape parameters.

The Kum distribution is applicable to many natural phenomena whose outcomes have lower and upper bounds, such as heights of individuals, scores obtained in a test, atmospheric temperatures and hydrological data. Also, it could be appropriate in situations where scientists use probability distributions which have infinite lower and or upper bounds to fit data, when in reality the bounds are finite. The Kumpdfhas the same basic properties as the beta distribution, but has the key advantage of a closed form cdf. [See Jones (2009)]

The *inverted Kumaraswamy*(IKum) distribution is important in a wide range of applications; engineering, medical research and lifetime problems.Abd AL-Fattah *et al.* (2017)discussedthe IKum distribution and derived the relation between the IKum

and other distributions. They obtained the *maximum likelihood* (ML), Bayes estimators and confidence intervals for the parameters, the *reliability function* (rf) and the *hazard rate function* (hrf) of the IKum distribution based on Type II censored samples. Also they illustrated the theoretical procedures by numerical study via simulation and real data. The IKum distribution can be derived from Kumdistribution using the transformation $X = \frac{1}{Y} - 1$, when Y has a Kum distribution.

The pdf and cdf of the IKum distribution are given, respectively, by

$$f(x; \alpha, \beta) = \alpha\beta(1 + x)^{-(\alpha+1)}(1 - (1 + x)^{-\alpha})^{(\beta-1)}, \quad x > 0, \qquad \alpha, \beta > 0, \tag{3}$$

and

$$F(x; \alpha, \beta) = (1 - (1 + x)^{-\alpha})^{\beta}, \qquad x > 0, \alpha, \beta > 0, \tag{4}$$

Where X~IKum with shape parameters α>0,β>0.

The rest of this paper is organized as follows: in Section 2, the distribution of the double truncated IKum (TIKum) distribution is derived. Some of its statistical propertiesare introduced in Section 3. In Section 4, ML estimators of the parameters, rf,hrfand reversed hazard rate function (rhrf)are obtained for the TIKum distribution based on complete samples. ML two-sample prediction for a future observation is considered in Section 5.A numerical study and an application to two real data sets are presented in Section 6; to discuss the various results developed.

2. The Distribution of the Double Truncated Inverted Kumaraswamy

A random variable $X$is said to have double TIKum$(\underline{\theta})$ distributionif its pdf and cdf are given, respectively, by

$$f(x; \underline{\theta}) = k (1 + x)^{-(\alpha+1)}(1 - (1 + x)^{-\alpha})^{\beta-1}, \quad c < x < d, \tag{5}$$

where the constant k is

$$k = \frac{\alpha\beta}{(1 - (1 + d)^{-\alpha})^{\beta} - (1 - (1 + c)^{-\alpha})^{\beta}},$$

$\underline{\theta} = (\alpha, \beta, c, d)'$, $c$is the lower point truncation and $d$is the upper point truncation.

The corresponding cdf is

$$F(x; \underline{\theta}) = \frac{(1 - (1 + x)^{-\alpha})^{\beta} - (1 - (1 + c)^{-\alpha})^{\beta}}{(1 - (1 + d)^{-\alpha})^{\beta} - (1 - (1 + c)^{-\alpha})^{\beta}}, \quad c < x < d. \tag{6}$$

The rf, hrf and rhrfhave the following forms, respectively,

$$R(x) = \frac{(1 - (1+d)^{-\alpha})^{\beta} - (1 - (1+x)^{-\alpha})^{\beta}}{(1 - (1+d)^{-\alpha})^{\beta} - (1 - (1+c)^{-\alpha})^{\beta}}, \qquad c < x < d, \tag{7}$$

$$h(x) = \frac{\alpha\beta(1+x)^{-(\alpha+1)}(1 - (1+x)^{-\alpha})^{\beta-1}}{(1 - (1+d)^{-\alpha})^{\beta} - (1 - (1+x)^{-\alpha})^{\beta}}, \qquad c < x < d, \tag{8}$$

and

$$rh(x) = \frac{\alpha\beta(1+x)^{-(\alpha+1)}(1 - (1+x)^{-\alpha})^{\beta-1}}{(1 - (1+x)^{-\alpha})^{\beta} - (1 - (1+c)^{-\alpha})^{\beta}}, \qquad c < x < d. \tag{9}$$

2.1 Graphical description

Plots of pdf, hrf and rhrf of TIKum for different values of the parameter α, β, truncation points c and dare given, respectively, in Figures 1-3. From Figure1, one can observe that the curve of the TIkum density is monotonically decreasing and it has a unimodal curve. In Figure 2 the hrf curve represents most major hazard shapes such as decreasing, approximately symmetric and bathtub shaped respectively which depends on the values of its parameters. From Figure 3 one can notice that the curves of the rhrf are bathtub; decreasing, almost fixed and increasing.
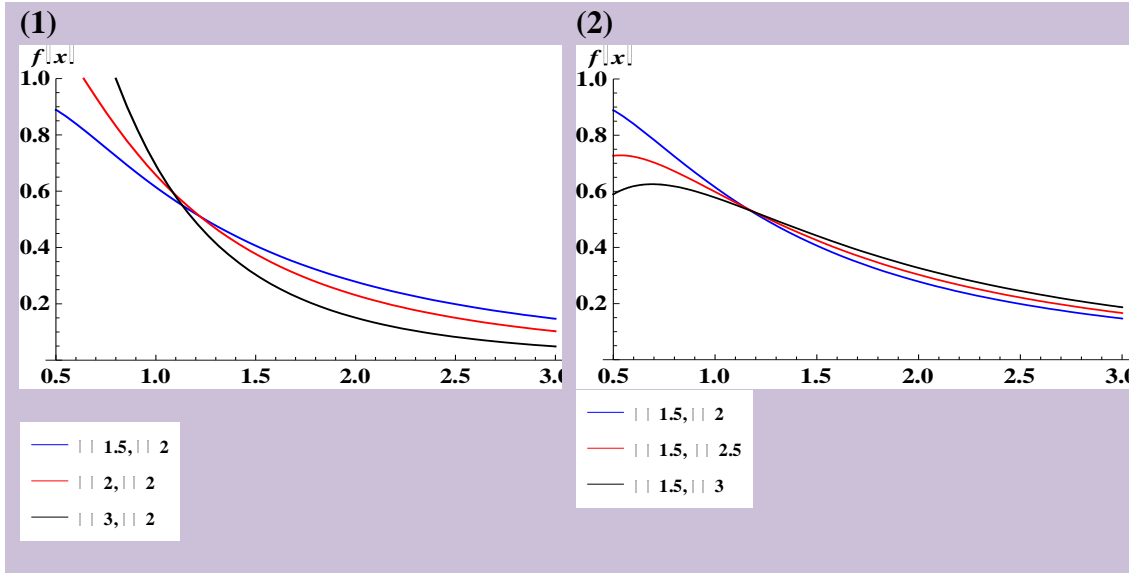

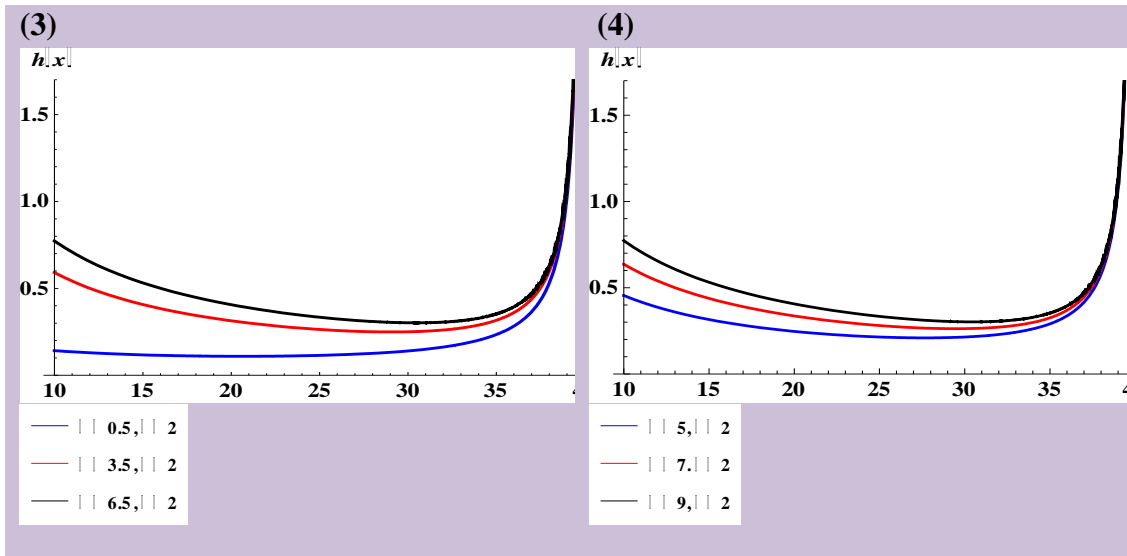**Figure 1:** Plots of the probability density function of double TIKum(c=0.5, d=3)

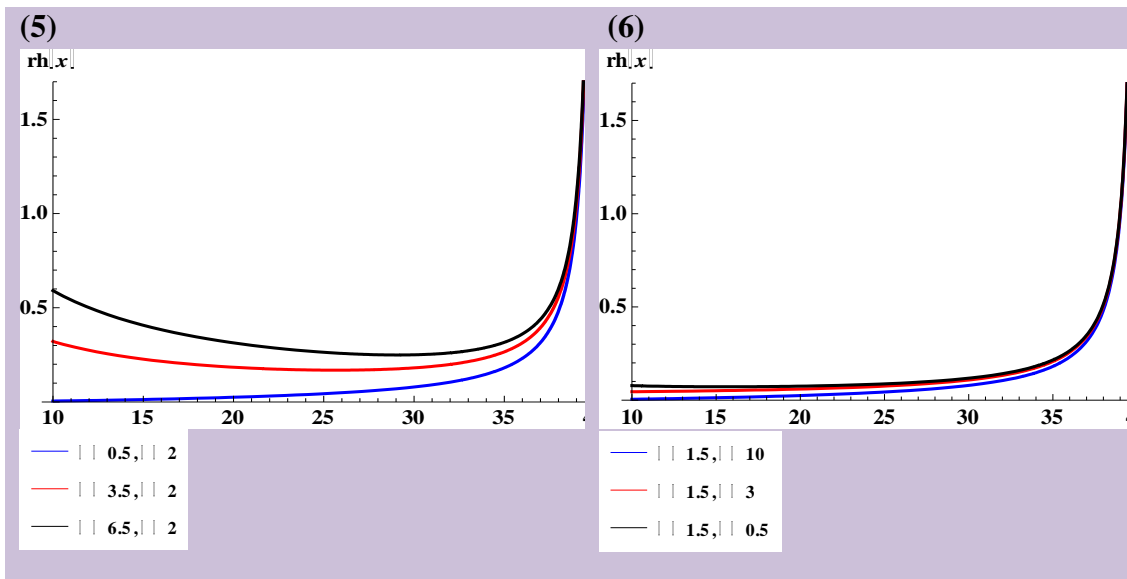**Figure 2:** Plots of the hazard rate function of double TIKum (c=10, d=40)

Figure 3: Plots of the reversed hazard rate function of double TIKum(c=10, d=40)

3. Statistical Properties

In this section, main properties of the TIKum (α,β,d,c) distribution are derived such as the quantile, median, the moments, moment generating function and order statistics.

3.1 Quantiles of the Truncated Inverted Kumaraswamy Distribution

The quantile function of the probability distribution is the inverse of its cdf andis used to describe the percentiles of the distribution. Hence, by inverting (6), an explicit expression for the quantile function of X is given below

$$Q(q) = F^{-1}(q), \qquad 0 < q < 1$$

$$= \left[1 - \left[q(1 - (1+d)^{-\alpha})^\beta + (1 - (1+c)^{-\alpha})^\beta(1-q)\right]^{\frac{1}{\beta}}\right]^{-\frac{1}{\alpha}} - 1, \qquad (10)$$

where Q(q) is the quantile function of the double TIKum and $q \in (0,1)$.

The median; the second quartile, can be obtained using (10) when q=0.5,

$$\text{Median} = \left[1 - \left[0.5(1 - (1+d)^{-\alpha})^\beta + (1 - (1+c)^{-\alpha})^\beta(1-0.5)\right]^{\frac{1}{\beta}}\right]^{-\frac{1}{\alpha}} - 1. \qquad (11)$$

3.2 The Mode of the Truncated Inverted Kumaraswamy Distribution

The mode of the double TIKum distribution is given by

$$\text{Mode} = \left(\frac{\alpha+1}{\alpha\beta+1}\right)^{-\frac{1}{\alpha}} - 1, \quad \alpha\beta \geq 1, \alpha > 0, \beta > 0. \qquad (12)$$

3.3 Moments and moment generating function

The rth non central moment of the double TIKum distribution is

$$\mu'_r = \frac{k}{\alpha}\sum_{j=0}^{r}\binom{r}{j}(-1)^{r-j}\left[B_l\left(1-\frac{r}{\alpha},\beta\right) - B_u\left(1-\frac{j}{\alpha},\beta\right)\right],$$

$$r = 1,2,\dots, \quad \alpha > j \ , j = 0,1,\dots,r, \qquad (13)$$

where $u = (1+d)^{-\alpha}$, $l = (1+c)^{-\alpha}$,

$\binom{r}{j} = \frac{r!}{j!(r-j)!}$is a binomial coefficient,$B_u(.,.)$and$B_l(.,.)$is the incomplete beta function. From (13) the mean, μ, and the variance,σ^2, of the double TIKum distribution are given, respectively, by

$$\mu = E(X) = \frac{k}{\alpha}\left[B_l\left(1 - \frac{1}{\alpha}, \beta\right) - B_u\left(1 - \frac{1}{\alpha}, \beta\right)\right], \alpha > 1, \tag{14}$$

and

$$\sigma^2 = Var(X) = \frac{k}{\alpha}\left[B_l\left(1 - \frac{2}{\alpha}, \beta\right) - 2B_l\left(1 - \frac{1}{\alpha}, \beta\right) + 2B_u\left(1 - \frac{1}{\alpha}, \beta\right) - B_u\left(1 - \frac{2}{\alpha}, \beta\right)\right]$$

$$-\left[\frac{k}{\alpha}\left[B_l\left(1 - \frac{1}{\alpha}, \beta\right) - B_u\left(1 - \frac{1}{\alpha}, \beta\right)\right]\right]^2. \tag{15}$$

The coefficient of variation (CV), skewness(SK) and kurtosis (KU) of the TIKum distribution can be obtained as follows:

$$CV = \sqrt{\frac{\mu_2'}{\mu_1'^2}},$$

$$SK = \frac{\mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3}{(\mu_2' - \mu_1'^2)^{3/2}},$$

and

$$KU = \frac{\mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1' - 3\mu_1'^4}{(\mu_2' - \mu_1'^2)^2}.$$

The moment generating function $M_x(t) = E(e^{tx})$ is

$$M_x(t) = \sum_{\iota=0}^{\infty} \frac{t^\iota}{\iota}\left[\frac{k}{\alpha}\sum_{j=0}^{\iota}\binom{\iota}{j}(-1)^{\iota-j}\left[\beta_l\left(1 - \frac{j}{\alpha}, \beta\right) - \beta_u\left(1 - \frac{j}{\alpha}, \beta\right)\right]\right],$$

$$\iota = 1,2,\dots \quad ,\alpha > j, j = 0,1,\dots,\iota. \tag{16}$$

3.4 Order statistics of the truncated Inverted Kumaraswamy distribution

The pdf of the ith order statistic, $X_{(i)}$, from a random sample of size n from the double TIKum (α, β, c, d) distribution is

$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!\,(n-i)!}\left[\frac{\alpha\beta(1+x)^{-(\alpha+1)}(1-(1+x)^{-\alpha})^{\beta-1}}{(1-(1+d)^{-\alpha})^\beta - (1-(1+c)^{-\alpha})^\beta}\right]$$

$$\times\left[\frac{(1-(1+x)^{-\alpha})^\beta - (1-(1+c)^{-\alpha})^\beta}{(1-(1+d)^{-\alpha})^\beta - (1-(1+c)^{-\alpha})^\beta}\right]^{i-1}$$

8

$$\times \left[\frac{(1-(1+d)^{-\alpha})^{\beta} - (1-(1+x)^{-\alpha})^{\beta}}{(1-(1+d)^{-\alpha})^{\beta} - (1-(1+c)^{-\alpha})^{\beta}}\right]^{n-i}, c \le x \le d . \tag{17}$$

The pdf of the smallest order statistic can be obtained when $i=1$, in (17) as follows:

$$f_{X_{(1)}}(x) = n\left[\frac{\alpha\beta(1+x)^{-(\alpha+1)}(1-(1+x)^{-\alpha})^{\beta-1}}{(1-(1+d)^{-\alpha})^{\beta} - (1-(1+c)^{-\alpha})^{\beta}}\right]$$

$$\times \left[\frac{(1-(1+d)^{-\alpha})^{\beta} - (1-(1+x)^{-\alpha})^{\beta}}{(1-(1+d)^{-\alpha})^{\beta} - (1-(1+c)^{-\alpha})^{\beta}}\right]^{n-1}, \qquad c \le x \le d .$$

Similarly, the pdf of the largest order statistic can be obtained when $i=n$, in (17) as follows:

$$f_{X_{(n)}}(x) = n\left[\frac{\alpha\beta(1+x)^{-(\alpha+1)}(1-(1+x)^{-\alpha})^{\beta-1}}{(1-(1+d)^{-\alpha})^{\beta} - (1-(1+c)^{-\alpha})^{\beta}}\right]$$

$$\times \left[\frac{(1-(1+x)^{-\alpha})^{\beta} - (1-(1+c)^{-\alpha})^{\beta}}{(1-(1+d)^{-\alpha})^{\beta} - (1-(1+c)^{-\alpha})^{\beta}}\right]^{n-1}, \qquad c \le x \le d .$$

4. Maximum likelihood estimation

Let $X_1, X_2, \ldots, X_n$ be a random sample drawn from a population having a double TIKum($\underline{\theta}$) pdf given by (5). The *likelihood function* (LF) is given by

$$L(\underline{\theta}; \underline{x}) = (\alpha\beta)^n((1-(1+d)^{-\alpha})^{\beta} - (1-(1+c)^{-\alpha})^{\beta})^{-n} \prod_{i=1}^{n}(1+x_i)^{-(\alpha+1)}$$

$$\times \prod_{i=1}^{n}(1-(1+x_i)^{-\alpha})^{\beta-1}, \tag{18}$$

where $\underline{x} = (x_1, x_2 \ldots, x_n)$ is the vector of observations and $\underline{\theta} = (\alpha, \beta, c, d)'$.

The ML estimators of the parameters c and d, the truncated points c and d are, respectively, given by

$$\hat{c}_{ML} = \arg\max L(\underline{\theta}; \underline{x}) = X_{(1)} = min(X_{(i)}), \tag{19}$$

and

$$\hat{d}_{ML} = \arg\max L(\underline{\theta}; \underline{x}) = X_{(n)} = max(X_{(i)}). \tag{20}$$

The logarithm of the LF in (18) is given by

$$\ell \equiv ln\, L(\underline{\theta};\underline{x}) = n\, ln(\alpha) + n\, ln(\beta) - nln\big[(1-(1+d)^{-\alpha})^{\beta} - (1-(1+c)^{-\alpha})^{\beta}\big]$$

$$-(\alpha+1)\sum_{i=1}^{n} ln(1+x_i) + (\beta-1)\sum_{i=1}^{n} ln(1-(1+x_i)^{-\alpha}). \quad (21)$$

Replacing the parameters c and d by $\hat{c}$ and $\hat{d}$ in (21). Then equating the first partial derivatives of ℓ with respect to α and β to zero, one obtains

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\hat{\beta}} + \sum_{i=1}^{n} ln\big(1-(1+x_i)^{-\hat{\alpha}}\big) -$$

$$n\frac{\left(1-(1+\hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}} ln\left(1-(1+\hat{d})^{-\hat{\alpha}}\right) - \left(1-(1+\hat{c})^{-\hat{\alpha}}\right)^{\hat{\beta}} ln\left(1-(1+\hat{c})^{-\hat{\alpha}}\right)}{\left(1-(1+\hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}} - \left(1-(1+\hat{c})^{-\hat{\alpha}}\right)^{\hat{\beta}}} = 0, \quad (22)$$

and

$$\frac{\partial \ell}{\partial \alpha}$$

$$= \frac{n}{\hat{\alpha}} - \sum_{i=1}^{n} ln(1+x_i) + (\hat{\beta}-1)\sum_{i=1}^{n} \frac{(1+x_i)^{-\hat{\alpha}} ln(1+x_i)}{(1-(1+x_i)^{-\hat{\alpha}})}$$

$$- n\hat{\beta}\frac{\left(1-(1+\hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}-1}(1+\hat{d})^{-\hat{\alpha}} ln(1+\hat{d}) - \left(1-(1+\hat{c})^{-\hat{\alpha}}\right)^{\hat{\beta}-1}(1+\hat{c})^{-\hat{\alpha}} ln(1+\hat{c})}{\left(1-(1+\hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}} - (1-(1+\hat{c})^{-\hat{\alpha}})^{\hat{\beta}}}$$

$$= 0 \;. \quad (23)$$

Equations (22) and (23) are two nonlinear equations which can be solved using Newton- Raphson method to obtain the ML estimators$\hat{\alpha}$and$\hat{\beta}$. The invariance property of the ML estimators can be used to obtain the ML estimator for the rf and the hrf; $\hat{R}(x)$ and $\hat{h}(x)$, as given below

$$\hat{R}(x) = \frac{\left(1-(1+\hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}} - \left(1-(1+x)^{-\hat{\alpha}}\right)^{\hat{\beta}}}{\left(1-(1+\hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}} - (1-(1+\hat{c})^{-\hat{\alpha}})^{\hat{\beta}}}, \qquad \hat{c} < x < \hat{d}, \quad (24)$$

and

$$\hat{h}(x) = \frac{\hat{\alpha}\hat{\beta}(1+x)^{-(\hat{\alpha}+1)}\left(1-(1+x)^{-\hat{\alpha}}\right)^{\hat{\beta}-1}}{\left(1-(1+\hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}} - (1-(1+x)^{-\hat{\alpha}})^{\hat{\beta}}}, \qquad \hat{c} < x < \hat{d}. \quad (25)$$

The asymptotic normality of the ML estimates can be used to compute the asymptotic confidence intervals for the parameters, rfandhrf.

Hence, a two sided approximate 100 $(1\text{-}\tau)\%$ confidence intervals for ω is given by

$$L_\omega = \hat{\omega} - Z_{(1-\tau/2)}\sqrt{var(\hat{\omega})} \text{ and } U_\omega = \hat{\omega} + Z_{(1-\tau/2)}\sqrt{var(\hat{\omega})}, \quad (26)$$

whereL_ω and U_ω are the *lower limit* (LL) and *upper limit* (UL), $\hat{\omega}$ is $\hat{\alpha}, \hat{\beta}, \hat{R}(x)$, or $\hat{h}(x)$, respectively, and $Z_{(1-\tau/2)}$ is a standard normalvariate and τ is the confidence coefficient.

5. Maximum Likelihood Prediction

This section considered the point and interval ML prediction for a future observation from the double TIKum $(\underline{\theta})$ distribution based on two-sample prediction.

Suppose that $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$are the first nordered life times in a random sample from the double TIKum $(\underline{\theta})$ distributionand assume$Y_{(1)} < Y_{(2)} < \cdots < Y_{(m)}$is a future independent random sample (of size m) from the same distribution. For the future sample of size m, let $Y_{(s)}$denotes the sth order statistic $1 \leq s \leq m$.

The pdf of $Y_{(s)}$can be obtained from (17) just by replacing $x_{(i)}$ by $y_{(s)}$

$$h\big(y_{(s)}|\underline{\theta}\big) = D(s)f\big(y_{(s)}|\underline{\theta}\big)\big[F\big(y_{(s)}|\underline{\theta}\big)\big]^{s-1}\big[1 - F\big(y_{(s)}|\underline{\theta}\big)\big]^{m-s}, c < y_{(s)} < d, \quad (27)$$

where

$$D(s) = s\binom{m}{s} = \frac{m!}{(s-1)!\,(m-s)!}, \qquad s = 1,2,\dots,m. \quad (28)$$

Substituting (5) and (6) in (17),thenusing (27) and the binomial expansion. Hence, the pdf of$y_{(s)}$is as follows:

$$h\big(y_{(s)}|\underline{\theta}\big) = D(s)\, Z\, \alpha\, \beta\big(1 + y_{(s)}\big)^{-(\alpha+1)}\big(1 - 1 + y_{(s)}^{-\alpha}\big)^{((s-j+i)\beta)-1}$$

$$\times (1 - (1+d)^{-\alpha})^{(m-s-i)\beta}(1 - (1+c)^{-\alpha})^{j\beta}$$

$$\times \big[(1 - (1+d)^{-\alpha})^\beta - (1 - (1+c)^{-\alpha})^\beta\big]^{-m}, \qquad c \leq y_{(s)} \leq d; \ \underline{\theta} > 0, \quad (29)$$

Where

$$Z = \sum_{j=0}^{s-1}(-1)^j\binom{s-1}{j}\sum_{i=0}^{m-s}(-1)^i\binom{m-s}{i}, \quad (30)$$

and $D(s)$is defined in (28).

Assuming that the parameters $\underline{\theta}$ are unknown and independent, then the *maximum
likelihood prediction density* (MLPD) of $Y_{(s)}$ given $\underline{x}$ can be obtained using the
conditional pdf of the order statistic which is given by (29) after replacing the vector
of parameters $\underline{\theta}$ by their ML estimators $\hat{\underline{\theta}}_{ML}$ as given below

$$h_s\left(y_{(s)}|\hat{\underline{\theta}}_{ML}\right) = D(s)\, Z\, \hat{\alpha}\hat{\beta}\left(1 + y_{(s)}\right)^{-(\hat{\alpha}+1)}\left(1 - 1 + y_{(s)}^{-\hat{\alpha}}\right)^{((s-j+i)\hat{\beta})-1}$$

$$\times \left(1 - (1 + \hat{d})^{-\alpha}\right)^{(n-s-i)\hat{\beta}}\left(1 - (1 + \hat{c})^{-\hat{\alpha}}\right)^{j\hat{\beta}}$$

$$\times \left[\left(1 - (1 + \hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}} - \left(1 - (1 + \hat{c})^{-\hat{\alpha}}\right)^{\hat{\beta}}\right]^{-n}, \qquad c \le y_{(s)} \le d, \;\; \underline{\hat{\theta}} > 0. \tag{31}$$

where $D(s)$ is given by (28) and Z is given by (30).

The *ML predictive estimator* (MLPE) of the future observation $Y_{(s)}$ can be derived
using (31) as follows:

$$\hat{y}_{(s)ML} = E\left(y_{(s)}|\hat{\underline{\theta}}_{ML}\right) = \int_{y_{(s)}} y_{(s)}\, h_s\left(y_{(s)}|\hat{\underline{\theta}}_{ML}\right) dy_{(s)}$$

$$= \int_{y_{(s)}} y_{(s)}\, D(s)\, Z\, \hat{\alpha}\hat{\beta}\left(1 + y_{(s)}\right)^{-(\hat{\alpha}+1)}\left(1 - 1 + y_{(s)}^{-\hat{\alpha}}\right)^{((s-j+i)\hat{\beta})-1}$$

$$\times \left(1 - (1 + \hat{d})^{-\alpha}\right)^{(n-s-i)\hat{\beta}}\left(1 - (1 + \hat{c})^{-\hat{\alpha}}\right)^{j\hat{\beta}}$$

$$\times \left[\left(1 - (1 + \hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}} - \left(1 - (1 + \hat{c})^{-\hat{\alpha}}\right)^{\hat{\beta}}\right]^{-n} dy_{(s)}, \qquad c \le y_{(s)} \le d, \;\; \underline{\hat{\theta}} > 0. \tag{32}$$

where $D(s)$ and Z are given, respectively, by (28) and (30).

The ML predictive bounds for $Y_{(s)}$

A 100(1-τ) % *maximum likelihood predictive bounds* (MLPB) for the future
observation $Y_{(s)}$, such that $P\left(L_{(s)}(\underline{x}) < Y_{(s)} < U_{(s)}(\underline{x})|\underline{x}\right) = 1 - \tau$, are as follows:

$$P\left(Y_{(s)} > L_{(s)}(\underline{x})|\underline{x}\right) = \int_{L_{(s)}(\underline{x})}^{\infty} h_1\left(y_{(s)}|\hat{\underline{\theta}}_{ML}\right)dy_{(s)} = 1 - \frac{\tau}{2}, \tag{33}$$

and

$$P\left(Y_{(s)} > U_{(s)}(\underline{x})|\underline{x}\right) = \int_{U_{(s)}(\underline{x})}^{\infty} h_1\left(y_{(s)}|\hat{\underline{\theta}}_{ML}\right)dy_{(s)} = \frac{\tau}{2}. \tag{34}$$

Substituting (31) in (33) and (34), the MLPB are obtained as follows:

$$P(Y_{(s)} > L_{(s)}(\underline{x})|\underline{\hat{\theta}}_{ML}) = \int_{L_{(s)}(\underline{x})}^{\infty} D(s)\, Z\, \hat{\alpha}\hat{\beta}(1 + y_{(s)})^{-(\hat{\alpha}+1)}(1 - 1 + y_{(s)}^{-\hat{\alpha}})^{((s-j+i)\hat{\beta})-1}$$

$$\times \left(1 - (1 + \hat{d})^{-\alpha}\right)^{(n-s-i)\hat{\beta}}\left(1 - (1 + \hat{c})^{-\hat{\alpha}}\right)^{j\hat{\beta}}$$

$$\times \left[\left(1 - (1 + \hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}} - \left(1 - (1 + \hat{c})^{-\hat{\alpha}}\right)^{\hat{\beta}}\right]^{-n} dy_{(s)} = 1 - \frac{\tau}{2}, \qquad (35)$$

and

$$P(Y_{(s)} > U_{(s)}(\underline{x})|\underline{\hat{\theta}}_{ML}) = \int_{U_{(s)}(\underline{x})}^{\infty} D(s)\, Z\, \hat{\alpha}\hat{\beta}(1 + y_{(s)})^{-(\hat{\alpha}+1)}(1 - 1 + y_{(s)}^{-\hat{\alpha}})^{((s-j+i)\hat{\beta})-1}$$

$$\times \left(1 - (1 + \hat{d})^{-\alpha}\right)^{(n-s-i)\hat{\beta}}\left(1 - (1 + \hat{c})^{-\hat{\alpha}}\right)^{j\hat{\beta}}$$

$$\times \left[\left(1 - (1 + \hat{d})^{-\hat{\alpha}}\right)^{\hat{\beta}} - \left(1 - (1 + \hat{c})^{-\hat{\alpha}}\right)^{\hat{\beta}}\right]^{-n} dy_{(s)}dy_{(s)} = \frac{\tau}{2}, \qquad (36)$$

where $s = 1, 2, 3, \ldots, m$.

5.1 Special cases:

• All results obtained in this paper for the TIKum distribution give corresponding results for the left TIKum distribution when $d = \infty$ in (5).

• Results can also be obtained for right TIKum distribution when $c = 0$ in (5).

6. Numerical Illustration

This section aims to investigate the precision of the theoretical results of estimation and prediction on basis of simulated and real data.

6.1 Simulation study

In this subsection, a simulation study is conducted to illustrate the performance of the presented ML estimates on the basis of generated data from the TIKum($\underline{\theta}$)distribution. The ML averages of the parameters, rf and hrf based on complete samples are computed.Moreover, confidence intervals of the parameters, rf and hrf are calculated. Also, the ML two-sample predictors (point and interval) for a

future observation are obtained. All simulation studies are performed using Mathematica 9.

The simulation algorithm based on complete sample data is

- Several data sets are generated from double TIKum distribution for a combination of the population parameter values of α, β, c and d and for samples ofsize 30, 50 and 100 using R=400 replications for each sample size.
- The transformation between the uniform distribution and double TIKum distribution is given as follows:

$$x_i = \left[1 - \left[u_i(1 - (1 + d)^{-\alpha})^\beta + (1 - (1 + c)^{-\alpha})^\beta(1 - u_i)\right]^{\frac{1}{\beta}}\right]^{-\frac{1}{\alpha}} - 1. \tag{37}$$

- The population parameter values of α, β, c and $d$used in this simulation study are $(1.5,\ 2, 0.6$ and $\ 2)$.
- A computer program; depending on Mathematica 9, is derivedusing the iterative technique of Newton Raphsonmethod to solve the nonlinear likelihood equations simultaneously. Hence the ML averages $\hat\alpha, \hat\beta, \hat c$ and $\hat d$ are computed.
- Also, the rf and the hrf are estimated for different values of time; $t_0 = (0.5,\ 0.8,\ 1)$.
- Evaluating the performance of the estimates of$\alpha, \beta, c, d, R(t_0)$ and $h(t_0)$is considered through some measurements of accuracy. The precision and variation of the MLestimates, are studied using the*relative absolute bias* (RAB) and the *relative error* (RE), where

$$RAB = \frac{\left|\bar{\hat\theta} - \theta\right|}{\theta}, \quad RE = \frac{\sqrt{ER(\hat\theta)}}{\theta}\text{where the estimated risk; } ER(\hat\theta) = \frac{1}{M}\sum_{i=1}^{M}(\hat\theta_i - \theta)^2.$$

- Table 1 presented the ML averages, estimated risks, relative absolute biases, variances and 95% Confidence intervals of the parameters; from the double TIKum distribution for different samplesof size n=30, 50, 100 and number of replications R=400.

- Table 2 displayed the ML averages, estimated risks, relative absolute biases of ML estimates and 95% confidence intervals of the reliability and hazard rate at $t_0 = (0.5,\ 0.8,\ 1)$; from TIKum distribution for different samples of size n=30, 50, 100 andnumber of replications R=400.

- The ML two-sample predictors and bounds are presented in Table 3.

6.2 Concluding Remarks

• One can observe from Table 1 and 2 that the ML averages are very close to the initial values of the parameters as the sample size increases. Also, RABs and REs are decreasing when the sample size is increasing. Also, the estimates are consistent and approaches the true parameter values as the sample size increases.

• The previous remark is expected since increasing the sample size means that more information is provided by the sample and hence increases the accuracy of the estimates.

• The ML averages intervals include the estimates.

• The lengths of the confidence intervals of the parameters become narrower as the sample size increases.

6.3 Applications

In this subsection, how the proposed methods can be used in practice is demonstrated. Two real lifetime data sets are analyzed.

Application 1:

The first application is given by Hinkley (1977). The data refers to thirty successive values of March precipitation (in inches) in Minneapolis/St Paul: 0.77, 1.74, 0.81, 1.20, 1.95, 1.20, 0.47, 1.43, 3.37, 2.20, 3.00, 3.09, 1.51, 2.10, 0.52, 1.62, 1.31, 0.32, 0.59, 0.81, 2.81, 1.87, 1.18, 1.35, 4.75, 2.48, 0.96, 1.89, 0.90, 2.05.

Application 2:

The second application is the time between failures for repairable items used by Murthy et al. (2004).The data is 1.43, 0.11, 0.71, 0.77, 2.63, 1.49, 3.46, 2.46, 0.59, 0.74, 1.23, 0.94, 4.36, 0.40, 1.74, 4.73, 2.23, 0.45, 0.70, 1.06, 1.46, 0.30, 1.82, 2.37, 0.63, 1.23, 1.24, 1.97, 1.86, 1.17.

To check the validity of the fitted model, the Kolmogorov–Smirnov and chi–squared goodness of fit tests are performed through the R programming language. The p values are given, respectively, by 0.925 and 0.6649. The p value given in each case showed that the model fits the data very well.

The ML estimates of the parameters, rf, hrf and their ERs and RABs, for the real data based on complete sample are displayed in Table 4. The ML two-sample predictors (point and interval) for a future observation are presented in Table 5.

From Tables 4 and 5, one can observe that:

• The estimates are very close to each other for each data set.
• The ML intervals include the estimates.

- The results based on the real data ensure the simulation results.

Table 1. ML averages, estimated risks,relative absolute biases, variances and 95% confidence intervals of the parameters from the double TIKum distribution for different samples ofsize n and replication R=400(α=1.5, β=2, c=0.6 and d=2)

n	$\widehat{\underline{\theta}}$	Averages	ER	RAB	Variance	UL	LL	Length
30	$\hat{\alpha}$	1.5377	0.3548	0.0252	0.3534	1.5755	0.4546	1.1209
	$\hat{\beta}$	2.0883	0.3422	0.0441	0.3344	2.2125	1.2662	0.9462
	$\hat{c}$	0.6282	0.0016	0.0469	0.0008	0.6842	0.5722	0.1119
	$\hat{d}$	1.9649	0.0132	0.0426	0.0059	2.0666	1.7631	0.3035
50	$\hat{\alpha}$	1.4201	0.2530	0.0233	0.2466	2.0244	1.3208	0.7036
	$\hat{\beta}$	2.0985	0.2995	0.0299	0.2961	2.4194	2.0013	0.4181
	$\hat{c}$	0.6157	0.0005	0.0262	0.0002	0.6454	0.5859	0.0594
	$\hat{d}$	1.9508	0.0047	0.0246	0.0022	2.0435	1.8581	0.1854
100	$\hat{\alpha}$	1.4193	0.1493	0.0213	0.1476	1.5238	1.2964	0.2274
	$\hat{\beta}$	2.0597	0.2301	0.0293	0.2265	2.3953	2.0506	0.3447
	$\hat{c}$	0.6145	0.0001	0.0141	0.0001	0.6242	0.5927	0.0315
	$\hat{d}$	1.9762	0.0012	0.0119	0.0006	2.0251	1.9273	0.0978

Table 2. ML averages, estimated risks, relative absolute biases and 95% confidence intervals of the reliability and hazard rate at t_0 = (0.5, 0.8, 1) from TIKum distribution for different samples of size n andreplication R=400 (α=1.5, β=2, c=0.6 and d=2)

n	t_0	Estimators	Averages	ER	RAB	UL	LL	Length
30	0.5	$\widehat{R}(t_0)$	0.8512	0.0025	0.0168	0.9457	0.5568	0.3889
		$\widehat{h}(t_0)$	0.1728	0.0812	0.0986	0.2837	0.0222	0.2614
	0.8	$\widehat{R}(t_0)$	0.5988	0.0076	0.0099	0.7694	0.4283	0.3411

		$\hat{h}(t_0)$	0.3165	0.0999	0.1074	0.3754	0.2459	0.1295
	1	$\hat{R}(t_0)$	0.4761	0.0102	0.0145	0.6736	0.3387	0.3349
		$\hat{h}(t_0)$	0.4611	0.0936	0.0775	0.5625	0.3659	0.1967
50	0.5	$\hat{R}(t_0)$	0.8482	0.0015	0.0132	0.9219	0.7745	0.1474
		$\hat{h}(t_0)$	0.1826	0.0543	0.0369	0.2039	0.1013	0.1027
	0.8	$\hat{R}(t_0)$	0.5145	0.0073	0.0159	0.7805	0.4485	0.3319
		$\hat{h}(t_0)$	0.2196	0.0633	0.0259	0.3090	0.1801	0.1289
	1	$\hat{R}(t_0)$	0.4665	0.0047	0.0060	0.6006	0.3324	0.2682
		$\hat{h}(t_0)$	0.4818	0.0416	0.0559	0.5028	0.3609	0.1419
100	0.5	$\hat{R}(t_0)$	0.8422	0.0010	0.0059	0.9037	0.7806	0.1231
		$\hat{h}(t_0)$	0.2195	0.0314	0.0221	0.3230	0.2113	0.1017
	0.8	$\hat{R}(t_0)$	0.5075	0.0021	0.0044	0.6967	0.5183	0.1784
		$\hat{h}(t_0)$	0.3715	0.0211	0.0258	0.4480	0.3309	0.1171
	1	$\hat{R}(t_0)$	0.4421	0.0029	0.0172	0.5809	0.3738	0.2072
		$\hat{h}(t_0)$	0.5982	0.0191	0.0114	0.6404	0.4027	0.1377

Table 3. ML predictive and bounds of the future observation under two-sample prediction ($n_1 = 100$, $n_2 = 30$, $\alpha = 1.5$, $\beta = 2$, $c = 0.6$ and $d = 2$)

s	$\hat{y}_{(s)ML}$	UL	LL	Length
3	0.6865	0.9645	0.6183	0.0762
18	1.2874	1.4985	1.0100	0.4886
30	3.6865	2.1871	1.6183	0.5687

Table 4. ML estimates of the parameters, rf, hrf, their estimated risks and relative absolute biases for the real data

Real data	n	Estimators	ML	ER	RAB
Application I	30	$\hat{\alpha}$	1.5376	0.1139	0.2814
		$\hat{\beta}$	2.8995	0.0101	0.0335
		$\hat{c}$	0.3200	0.0784	0.4667
		$\hat{d}$	4.7500	0.5625	0.1875
		$\hat{R}(t_0)$	0.9204	0.0148	0.1167
		$\hat{h}(t_0)$	0.5234	0.0190	0.3575
Application II	30	$\hat{\alpha}$	1.4378	0.0189	0.1060
		$\hat{\beta}$	2.4002	0.1602	0.2001
		$\hat{c}$	0.1100	0.2401	0.8167
		$\hat{d}$	4.7300	0.5329	0.1825
		$\hat{R}(t_0)$	0.8366	0.0562	0.2208
		$\hat{h}(t_0)$	0.6059	0.0085	0.1318

Table 5. ML predictive and bounds of the future observation for real data under two-sample prediction

s	Application I				Application II			
	$\hat{y}_{(s)ML}$	UL	LL	Length	$\hat{y}_{(s)ML}$	UL	LL	Length
1	0.0437	0.0472	0.0431	0.0040	0.0697	0.0683	0.0650	0.0032
9	0.0937	0.0985	0.0925	0.0062	0.1891	1.1928	0.1873	0.0045
17	1.2469	1.2532	1.2451	0.0081	1.0827	1.0966	1.0792	0.0063

References

Abd AL-Fattah, A. M.,AL-Dayian, G. R. and EL-Helbawy, A. A. (2017). Inverted Kumaraswamy distribution: properties and estimation. Pakistan Journal of Statistics, 33(1):37-61.

Al-Omari, A. I., (2018). Acceptance sampling plans based on truncated life tests for Sushiladistribution. Journal of Mathematical Fundamental Science, 50 (1):72-83.

AL-Hussaini, E.K., AL-Dayian, G.R. and AL-Angary, A. M. (2006). Bayesian prediction bounds under the truncated Type1 generalized logistic model. Journal of the Egyptian Mathematical Society, 14(1):55-67.

AL-Yousef, M.H. (2002).Estimation in a doubly truncated Burr distribution. Journal King Saudi University Admin Sciences, 1(1):1-9.

Ateya, S.F. and AL-Hussaini, E. K. (2007). Estimation under truncated generalized Cauchy lifetime model. Journal of the Egyptian Mathematical Society, 1:181-193.

Gupta, R. C., Gupta, R. D. and Gupta, P. L. (1998). Modeling failure time data by Lehman alternatives, Communications in Statistics - Theory and Methods, 27(4): 887-904.

Isik, H., Turfan, D. and Ozel, G. (2017). The double truncated Dagum distribution with applications. Journal of mathematics and Statistics an International Journal,18(2): 61-78.

Hinkley, D. (1977). On quick choice of power transformations, Journal of the Royal Statistical Series C (Applied Statistics), 26(1), 67-69.

Jones, M. C. (2009). Kumaraswamy's distribution: A beta-Type distribution with some tractability advantages. Journal of Statistical Methodology, 6 (1): 70-81.

Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. Journal of Hydrology, 46, 79-88.

Murthy, D.N.P., Xie, M. and Jiang, R. (2004). Weibull Models, Wiley series in probability and statistics, John Wiley and Sons, New York.

Okasha,M.K. and Alqanoo, I.M.A. (2014). Inference on the doubly truncated Gamma distribution for lifetime data. International Journal of Mathematics and Statistics Invention, 2 (11):1-17.

Nadarajah, S.(2008). A truncated inverted beta distribution with application to air pollution data. Stochastic Environmental Research and RiskAssessment,22:285-289.

Raschke, M. (2012). Inference for the truncated exponential distribution.Stochastic Environmental Research and Risk Assessment, 26: 127-138.

Singh, S.K., Singh, U. and Sharma, K.(2014). The truncated Lindley distribution: inference and application. Journal of Statistics Application Probability an International Journal,3(2): 219-228.

Zhang,T. and Xie, M. (2011). On the upper truncated Weibull distribution and its reliability implications. Reliability Engineering and System Safety, 96: 194-200.