Original Research Paper

The Uniform Truncated Negative Binomial Distribution and its Properties

[1]Bahady I. Kamel, [1]Shaban E. Abo Youssef and [2]Mohamed G. Sief

[1]*Department of Mathematics, Faculty of Science, Al-Azhar University, Nasr City, Cairo, Egypt*
[2]*Department of Mathematics, Faculty of Science, Fayoum University, Fayoum, Egypt*

Corresponding Author:
Mohamed G. Sief
Department of Mathematics,
Faculty of Science, Fayoum
University, Fayoum, Egypt
Email: mgr00@fayoum.edu.eg

Abstract: A new three parameters continuous distribution which is named the Uniform Truncated Negative Binomial (UTNB) distribution is introduced. It has the uniform and the Marshall-Olkin extended uniform distributions as sub-models. Some reliability and statistical properties of the new distribution are derived, including the shape behavior of the density and hazard rate functions, the mean residual life, moments and moment generating function, quantiles and related measures. The limiting distributions of the sample extremes, stochastic orderings, entropies and stress-strength reliability are derived. Maximum likelihood estimation is performed. Application to censored data illustrate the potentiality of the proposed distribution is provided.

Keywords: Uniform Distribution, Truncated Negative Binomial, Stress-Strength, Hazard Rate, Mean Residual Life, Maximum Likelihood

Introduction

The basic distributions (exponential, Weibull, gamma, log normal, Lomax and others) used widely in reliability and survival analysis, are described by Cox and Oakes (1984), have a limited range of behavior and cannot represent all the situations found in applications. For example, the exponential has a restricted hazard function, being constant. The Weibull is often described as flexible, but its hazard function is in fact restricted to being monotonically increasing or monotonically decreasing, or constant. The limitations of the standard distributions arouse the interest in developing extended or generalized distributions by adding further parameters. Marshall and Olkin (1997) introduced a new technique by adding a parameter to an existing distribution F to generate a family of distributions. The extended family taking the form:

$$\bar{G}(x) = \frac{\alpha \bar{F}(x)}{F(x) + \alpha \bar{F}(x)}, \alpha > 0 \qquad (1)$$

In the literature, many new parametric extensions of various univariate distributions were proposed using this method such as: Marshall-Olkin extended Lomax, Ghitany *et al.* (2007), Marshall-Olkin extended uniform (MOEU), Jose and Krishna (2011), Marshall-Olkin Pareto, Ghitany (2006), Marshall-

Olkin logistic, Alice and Jose (2005a), Marshall-Olkin semi-Weibull, Alice and Jose (2005b), Marshall-Olkin q-Weibull, Jose *et al.* (2010) and Marshall-Olkin Weibull, Ghitany *et al.* (2005).

Moreover, Jayakumar and Mathew (2008) introduced the exponentiated Marshall-Olkin scheme which has Marshall-Olkin family as a special case given by:

$$\bar{G}(x, \alpha, \gamma) = \left[\frac{\alpha \bar{F}(x)}{1 - (1 - \alpha) \bar{F}(x)} \right]^{\gamma} \qquad (2)$$

where, $\gamma, \alpha > 0, x \in \mathbb{R}$.

Also, a new family of distributions was proposed by Nadarajah *et al.* (2013) by adding an extra parameter to the Marshall-Olkin family of distributions whose survival function is given by:

$$\bar{G}(x) = \frac{\alpha^{\beta}}{1 - \alpha^{\beta}} \left[\left(F(x) + \alpha \bar{F}(x) \right)^{-\beta} - 1 \right] \qquad (3)$$
$$\alpha, \beta > 0, x \in \mathbb{R}, \alpha \neq 1$$

The family (3) is a generalization of the Marshall-Olkin family of distributions (1). Namely, if $\beta = 1$, then the introduced family reduces to the Marshall-Olkin family of distributions. Moreover, Nadarajah *et al.* (2013) introduced three-parameter generalization of the exponential distribution as a particular case of the new

family. Also, Jose and Sivadas (2015) used the family (3) to introduce the negative binomial Marshall-Olkin Rayleigh distribution.

The uniform distribution is regarded as the simplest probability model and is related to all distributions by the fact that the cumulative distribution function, taken as a random variable, follows uniform distribution over (0,1). This result is basic to the inverse method of the random variable generation. Some researchers utilized uniform distribution in connection with sample quasi-ranges. The distribution is also applied to determine power functions of tests of randomness. It is also applied in a power comparison of tests of non-random clustering. There are also numerous applications in nonparametric inference, such as Kolmogorov-Smirnov test for goodness of fit. The uses of uniform distribution to represent the distribution of round off errors and in connection with the probability integral transformations, are also well-known.

We proposed a new version of the family (3) by selecting the survival function of the uniform $(0,\theta)$ in Equation 3, where $\bar{F}(x) = 1 - \dfrac{x}{\theta}, 0 < x < \theta$, so we get:

$$\bar{G}(x) = \frac{\alpha^{\beta}}{1-\alpha^{\beta}}\left[\left(\frac{x}{\theta}+\alpha\left(1-\frac{x}{\theta}\right)\right)^{-\beta}-1\right] \qquad (4)$$

where, $0 < x < \theta$, $1 \neq \alpha$, $\beta > 0$. We refer to the distribution with the survival function (4), as the uniform truncated negative binomial distribution with parameters θ, α, β and will be denoted by UTNB (θ, α, β).

In this article, we introduce a new distribution which is named the Uniform Truncated Negative Binomial (UTNB) distribution. Some statistical characteristics of the proposed distribution are obtained and the maximum likelihood estimation of its parameters is discussed. The article is outlined as following: In section 2, the probability density function (pdf) and its behavior are studied. Section 3 discusses the hazard rate function and obtains the mean residual life function (mrlf). The moment generating function is expressed as an infinite sum and an explicit form for the *r*th moment is obtained in section 4. In section 5, we give an expression for the quantile function and also some related measures. Section 6 provides the limiting distributions of the sample extreme. The stochastic orderings are derived in section 7. In section 8, the stress-strength reliability $R = P(Y < X)$ of a system is discussed when both the strength X of the system and the stress Y, which act on it, are independent, non-identical UTNB distributed random variables. Rényi and Shannon entropies are determined in section 9. Section 10 discusses the maximum likelihood estimation of the model's parameters. In section 11, the

proposed model has been fitted to randomly censored data to motivate its usefulness. Finally, the concluding remarks are addressed in section 12.

The Shapes of the Density Function

The pdf of the UTNB distribution with is :

$$g(x) = \frac{(1-\alpha)\beta\alpha^{\beta}}{(1-\alpha^{\beta})\theta}\left(\alpha+\frac{1-\alpha}{\theta}x\right)^{-(\beta+1)} \qquad (5)$$

where, $\alpha, \beta > 0$, $0 < x < \theta$.

Remark 1

The UTNB distribution having the uniform and MOEU distributions as sub-models as:

- If $\alpha \to 1$, then the UTNB $(\theta, \alpha, \beta) \to U(0, \theta)$
- If $\beta \to 1$, we get the MOEU distribution

The following proposition gives simple sufficient condition for the shape behavior of the pdf (5)

Proposition 1

The pdf $g(x)$ of the UTNB (α, β, θ) is increasing (decreasing) if $\alpha > 1 (\alpha < 1)$, independent on β and θ.

Proof

Taking the first derivative of $\log(g)$ then we have:

$$(\log g(x))' = -\frac{(\beta+1)(1-\alpha)}{\alpha\theta+(1-\alpha)x}, 0 < x < \theta$$

If $\alpha \in (0,1)$ then $(\log(g))' < 0$ for all $x \in (0,\theta)$ and hence, $g(x)$ is decreasing with $g(0) = (\alpha-1)\beta/(\alpha^{\beta}-1)\alpha\theta$ and $g(\theta) = 0$.

If $\alpha > 1$ then $(\log(g))' > 0$ for all $x \in (0,\theta), \theta > 0$, it follows that $g(x)$ is increasing.

Figure 1 shows the behavior of the pdf for selected values of the parameters α, θ, β.

The Hazard rate and Mean Residual Life Functions

The hazard rate function (hrf) of the UTNB distribution is given by:

$$h(x) = \frac{(\alpha-1)\beta}{((\alpha-1)x-\alpha\theta)\left(\left(\alpha+\frac{(1-\alpha)x}{\theta}\right)^{\beta}-1\right)} \qquad (6)$$

where, $1 \neq \alpha, \beta > 0, 0 < x < \theta$.

The following proposition gives simple condition under which the hrf (6) is increasing or bathtub.

Proposition 2

The hrf $h(x)$ of the UTNB (α,β,θ) is increasing (bathtub) if $\alpha^\beta(1+\beta)-1\geq0(<0)$.

Proof

By taking the first derivative of $h(x)$ we get:

$$h'(x) = \frac{(\alpha-1)^2\beta}{\left(\alpha\theta+(1-\alpha)x\right)^2\left(\left(\alpha+\frac{(1-\alpha)x}{\theta}\right)^\beta-1\right)^2}\phi(x)$$

where, $\phi(x)=(\beta+1)\left(\alpha+\frac{(1-\alpha)x}{\theta}\right)^\beta-1, 0<x<\theta$.

If $\phi(0) = (\beta+1)\alpha^\beta-1\geq0$, then $h'(x)\geq0$ and hence, $h(x)$ is increasing for all $0<x<\theta$.

If $\phi(0) <0$, thus the function $\phi(x)$ has one root say,

$x_o = \frac{\theta\left(\alpha-(\beta+1)^{-1/\beta}\right)}{\alpha-1}$ and it is negative for $x<x_o$ and positive for $x>x_o$. This implies that the function $h'(x)$ is negative for $x<x_o$ and positive for $x>x_o$. Thus, we conclude that $h(x)$ is bathtub shaped.

Figure 2 shows the graph of the hrf for the UTNB distribution for selected values of the parameters α,β,θ.
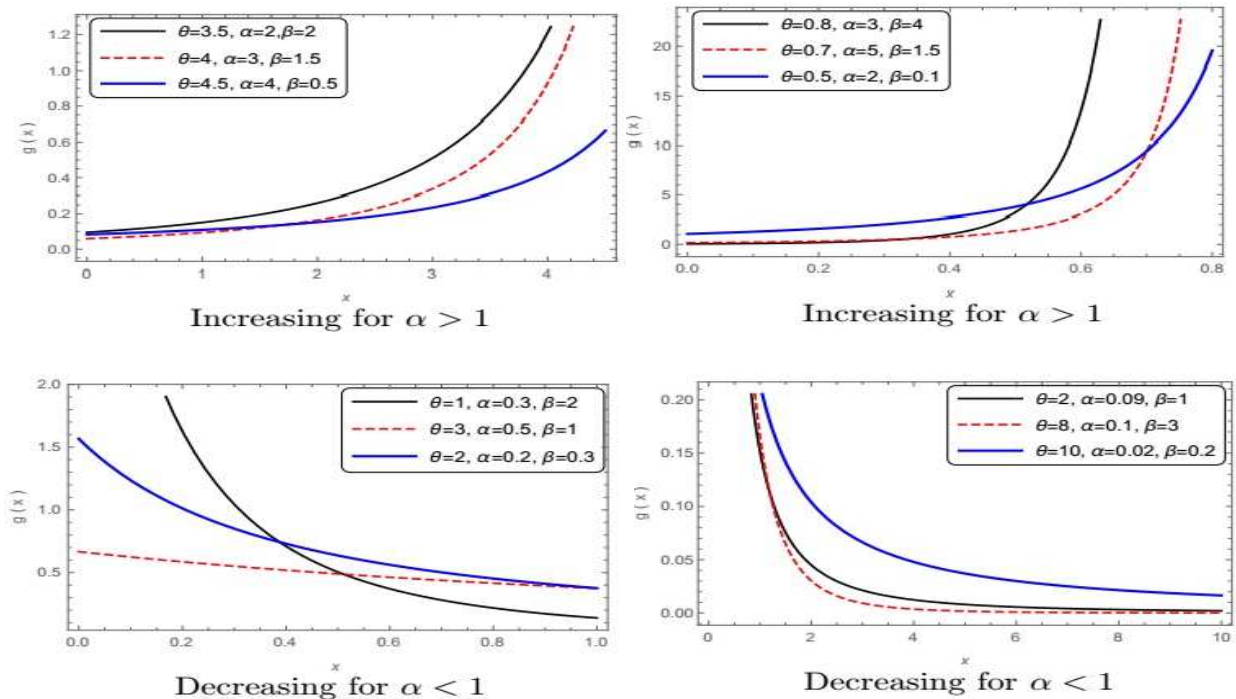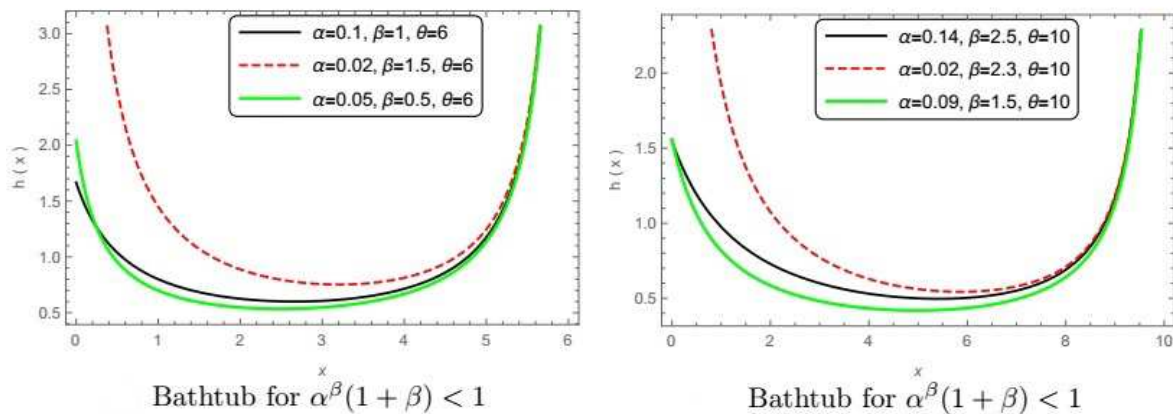

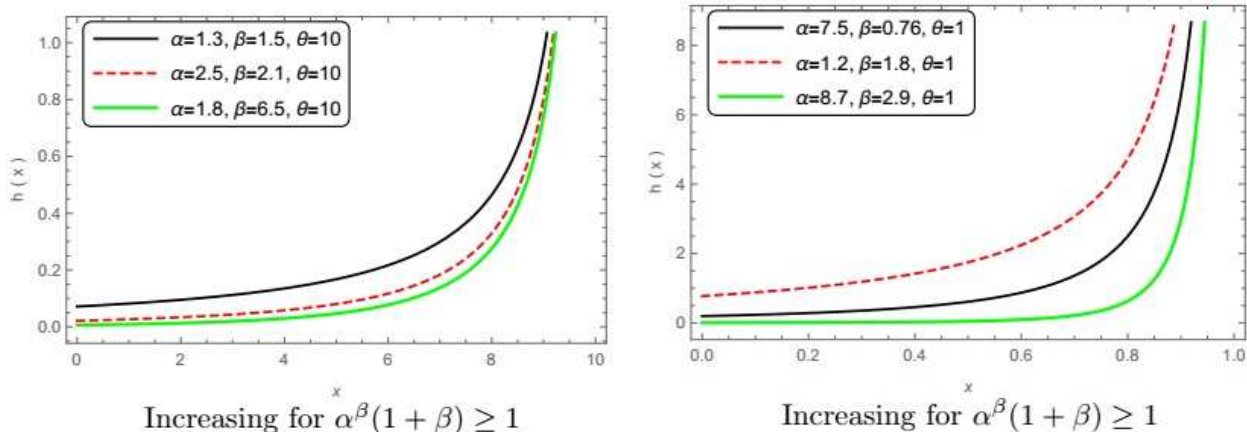Fig. 1. Some possible shapes of the UTNB pdf

Fig. 1. The hrf of the UTNB distribution for selected values of the parameters α, β, θ

An important ageing property for the UTNB distribution is the mean residual life function (mrlf), which was introduced by Watson and Wells (1961) to analyse burn-in problems. It has been studied by reliabilists, statisticians, survival analysts and others. It is defined simply as the expected value of the remaining lifetime beyond an age *t*. For $X\sim$ UTNB(α, β, θ), the residual life random variable at age *t*, is denoted by:

$$X_t = X - t \mid X > t$$

The mrlf is defined formally as:

$$\mu(t) = E(X - t | X > t)$$
$$= \frac{1}{\overline{G}(t)} \int_t^\infty \overline{G}(x) dx = \frac{1}{\left(\alpha + \frac{1-\alpha}{\theta}t\right)^{-\beta} - 1} \int_t^\theta \left[\left(\alpha + \frac{1-\alpha}{\theta}x\right)^{-\beta} - 1 \right] dx$$
$$= \frac{(t-\theta) + \frac{\theta}{(1-\alpha)(1-\beta)} \left(1 - \left(\alpha + \frac{1-\alpha}{\theta}t\right)^{1-\beta}\right)}{\left(\left(\alpha + \frac{1-\alpha}{\theta}t\right)^{-\beta} - 1\right)}$$

where, $1 \neq \alpha$, $\beta > 0, 0 \le t \le \theta$.

Moment Generating Function and Moments

In this section, we present an infinite sum for the moment generating function (mgf) of the UTNB distribution. Also, we derive an explicit form of the *r*th moment of the UTNB distribution. In the sequel, we use the substitution $u = \alpha + \frac{1-\alpha}{\theta}x$.

The mgf is defined as:

$$M(t, \alpha, \beta, \theta) = E(e^{tX}) = \int_0^\infty e^{tx} g(x) dx$$
$$= \int_0^\theta e^{tx} \frac{(1-\alpha)\beta\alpha^\beta}{(1-\alpha^\beta)\theta} \left(\alpha + \frac{1-\alpha}{\theta}x\right)^{-(\beta+1)} dx$$
$$= \frac{\beta\alpha^\beta e^{-\frac{\alpha\theta}{(1-\alpha)}t}}{(1-\alpha^\beta)} \int_\alpha^1 e^{\frac{\theta t}{(1-\alpha)}u} u^{-\beta-1} du$$
$$M(t, \alpha, \beta, \theta) = \frac{\beta\alpha^\beta e^{-\frac{\alpha\theta}{(1-\alpha)}t}}{(1-\alpha^\beta)} \sum_{n=0}^\infty \frac{\left(\frac{\theta t}{1-\alpha}\right)^n (1-\alpha^{n-\beta})}{n!(n-\beta)}$$

$n \neq \beta, \alpha \neq 1$

The *r*th moment for the UTNB random variable *X* is given by:

$$E(X^r) = \int_0^\infty x^r g(x) dx$$
$$= \int_0^\theta x^r \frac{(1-\alpha)\beta\alpha^\beta}{(1-\alpha^\beta)\theta} \left(\alpha + \frac{1-\alpha}{\theta}x\right)^{-(\beta+1)} dx$$
$$= \frac{\beta\alpha^\beta \theta^r}{(1-\alpha^\beta)(1-\alpha)^r} \int_\alpha^1 u^{-\beta-1}(u-\alpha)^r du$$
$$= \frac{\left(\frac{\alpha\theta}{1-\alpha}\right)^r}{(\alpha^\beta-1)\Gamma(\beta)} \left[\beta\Gamma(\beta) B_\alpha(\beta-r, r+1) - \Gamma(r+1)\Gamma(\beta-r) \right]$$

where, $B_z(a,b) = \int_0^z t^{a-1}(1-t)^{b-1} dt$ is the incomplete beta function.

Quantiles, Kurtosis, Skewness

The *q*th quantile of the UTNB distribution is given by:

$$Q(q) = G^{-1}(q) = \frac{\theta}{\alpha-1}\left(\alpha - \left(q - (q-1)\alpha^{-\beta}\right)^{-\frac{1}{\beta}}\right) \qquad (7)$$

where, $0 \leq q \leq 1$ and $G^{-1}(\cdot)$ is the inverse function of the distribution function. As special cases, we have:

- The qth quantile of the uniform distribution ($\alpha \to 1$) is given as $x_q = q^\theta$
- The qth quantile of the MOEU distribution ($\beta = 1$) is given by $x_q = \dfrac{\alpha\theta q}{(\alpha-1)q+1}$

In particular, the median of UTNB distribution is given by:

$$X_{\text{med}} = \frac{\theta}{\alpha-1}\left(\alpha - 2^{1/\beta}\left(\alpha^{-\beta}+1\right)^{-1/\beta}\right)$$

Also, we can get the median of the uniform and that of the MOEU as special cases respectively as $X_{\text{med}} = \dfrac{\theta}{2}$ and $X_{\text{med}} = \dfrac{\alpha\theta}{\alpha+1}$.

A new quantile measures for the dispersion of a distribution around the values $\mu \pm \sigma$ will be presented. The Bowley skewness, Keeping (1962) is given by:

$$SK = \frac{Q\left(\frac{1}{4}\right) - 2Q\left(\frac{2}{4}\right) + Q\left(\frac{3}{4}\right)}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)}$$

Also, the Moors kurtosis, Moors (1988) is given by:

$$K = \frac{Q\left(\frac{1}{8}\right) + Q\left(\frac{3}{8}\right) - Q\left(\frac{5}{8}\right) + Q\left(\frac{7}{8}\right)}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)}$$

Figure 3 displays the Bowley skewness and Moors kurtosis for selected values of the parameters which shows how the skewness and kurtosis are independent on θ and depend only on α and β.

Limiting Distribution of Sample Extremes

Let X_1, X_2, …, X_n be a random sample of size n from the UTNB distribution, then the sample minima and the sample maxima which are a long stand area in application of probability and statistics, are respectively:

$$X_{1:n} = \min(X_1, X_2, ..., X_n), \; X_{n:n} = \max(X_1, X_2, ..., X_n)$$
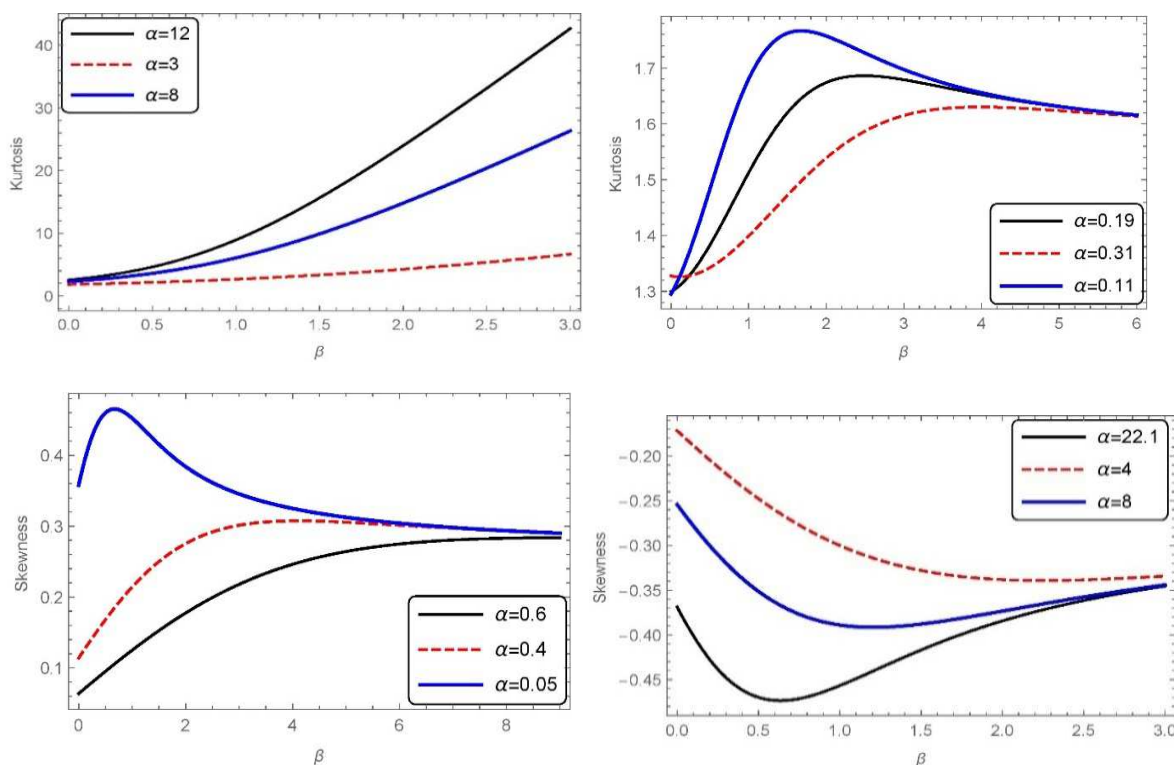

Fig. 3. Some possible shapes of the Bowley skewness and Moors kurtosis for the UTNB distribution

Using the asymptotic result of $X_{1:n}$ and $X_{n:n}$, Arnold *et al.* (1992), as following:

- For the minimal order statistic $X_{1:n}$, we have:

$$\lim_{n \to \infty} P\{X_{1:n} \le a_n^* + b_n^* t\} = 1 - \exp\left[-t^c\right], \ t > 0, c > 0 \qquad (8)$$

(of the Weibull type) where $a_n^* = F^{-1}(0)$ and $b_n^* = F^{-1}(1/n) - F^{-1}(0)$ if and only if:

- $F^{-1}(0)$ is finite
- For all $t > 0$, $c > 0$, we have:

$$\lim_{\varepsilon \to 0^+} \frac{F\left(F^{-1}(0) + \varepsilon t\right)}{F\left(F^{-1}(0) + \varepsilon\right)} = t^c$$

- For the maximal order statistic $X_{n:n}$, we have:

$$\lim_{n \to \infty} P\{X_{1:n} \le a_n + b_n t\} = \exp\left[-t^k\right]$$
$$t > 0, k > 0 \qquad (9)$$

(of the Weibull type) where $a_n = F^{-1}(1)$ and:

$$b_n = F^{-1}(1) - F^{-1}\left(1 - \frac{1}{n}\right)$$

if and only if:

- $F^{-1}(1)$ is finite.
- For all $t > 0$, $k > 0$, we have:

$$\lim_{\varepsilon \to 0^+} \frac{\overline{F}\left(F^{-1}(1) - \varepsilon t\right)}{\overline{F}\left(F^{-1}(1) - \varepsilon\right)} = t^k$$

In the following proposition we derived the limiting distributions of $X_{1:n}$ and $X_{n:n}$ from the UTNB distribution.

Proposition 3

Let $X_{1:n}$ and $X_{n:n}$ be the smallest and the largest order statistic from the UTNB distribution, then we have:

- $\lim_{n \to \infty} P\{X_{1:n} \le b_n^* t\} = 1 - e^{-t}, t > 0$

where:

$$b_n^* = \frac{\alpha \theta}{\alpha - 1}\left(1 - \left(\frac{n}{\alpha^\beta + n - 1}\right)^{1/\beta}\right)$$

- $\lim_{n \to \infty} P\{X_{n:n} \le a_n + b_n t\} = e^{-t}, t > 0$

where:

$$a_n = \theta, b_n = \frac{\theta}{1-\alpha}\left(1 - \left(\frac{n}{\alpha^{-\beta} + n - 1}\right)^{1/\beta}\right)$$

Proof

We have $G^{-1}(0) = 0$ is finite and:

$$\lim_{\varepsilon \to 0^+} \frac{G(\varepsilon t)}{G(\varepsilon)} = t \lim_{\varepsilon \to 0^+} \frac{\left(\dfrac{\varepsilon - \alpha\varepsilon}{\theta} + \alpha\right)^{\beta+1}}{\left(\alpha - \dfrac{(\alpha-1)t\varepsilon}{\theta}\right)^{\beta+1}} = t$$

Using Equation 8 and 7 we have $a_n^* = 0$ and:

$$b_n^* = \frac{\alpha\theta}{\alpha - 1}\left(1 - \left(\frac{n}{\alpha^\beta + n - 1}\right)^{1/\beta}\right)$$

We have $G^{-1}(1) = \theta$ is finite and:

$$\lim_{\varepsilon \to 0^+} \frac{\overline{G}(\theta - \varepsilon t)}{G(\theta - \varepsilon)} = t \left(\lim_{\varepsilon \to 0^+} \frac{\left(\dfrac{(\alpha-1)\varepsilon}{\theta} + 1\right)^{\beta+1}}{\left(\dfrac{(\alpha-1)t\varepsilon}{\theta} + 1\right)^{\beta+1}}\right) = t$$

Thus, we obtain that $k = 1$, also using (7) we get $a_n = \theta$ and $b_n = \dfrac{\theta}{1-\alpha}\left(1 - \left(\dfrac{n}{\alpha^{-\beta} + n - 1}\right)^{1/\beta}\right)$, therefore statement (2) flows from (9).

Remark 2

For $\beta = 1$, i.e., for the MOEU distribution, the norming constants b_n^*, b_n are given respectively by:

$$\frac{\alpha\theta}{\alpha + n - 1}, \frac{\alpha\theta(n-1)}{\alpha(n-1) + 1}$$

(see, e.g., Jose and Krishna (2011)).

If $\alpha \to 1$, i.e., the uniform distribution, the norming constants will be $b_n^* = \dfrac{\theta}{n}, b_n = \dfrac{n-1}{n\theta}$.

Let the limiting distribution of $\left(X_{1:n} - a_n^*\right)/b_n^*$ and $(X_{n:n} - a_n)/b_n$ are, respectively, denoted as $G^*(t)$ and $G(t)$. Then, it is well-known from Arnold *et al.* (1992) that for any finite $i > 1$, the limiting distributions of the random variables $(X_{i:n} - a_n)/b_n$ and $(X_{n-i+1:n} - a_n)/b_n$ are, respectively, given by:

$$\lim_{n \to \infty} P\{X_{i:n} \le a_n^* + b_n^* t\} = 1 - \sum_{j=0}^{i-1}\left(1 - G^*(t)\right)\frac{\{-\log[1 - G^*(t)]\}^j}{j!}$$

$$\lim_{n \to \infty} P\{X_{n-i+1:n} \le a_n + b_n t\} = \sum_{j=0}^{i-1} G(t)\frac{\{-\log G(t)\}^j}{j!}$$

From proposition (3) we conclude that, for any finite $i>1$, the limiting distribution of the *i*th and the $(n-i+1)$th order statistic from the UTNB distribution, respectively, are given by:

$$\lim_{n\to\infty} P\left\{X_{i:n} \le b_n^* t\right\} = 1 - \sum_{j=0}^{i-1} e^{-t} \frac{t^j}{j!} = 1 - P(Z < i), t > 0$$

$$\lim_{n\to\infty} P\left\{X_{n-i+1:n} \le a_n + b_n t\right\} = \sum_{j=0}^{i-1} e^{-t} \frac{t^j}{j!} = P(Z < i) \ t > 0$$

where, b_n^*, a_n and b_n are given in proposition (3) and Z is discrete random variable follows Poisson distribution with mean t.

Stochastic Orderings

A very important tool for judging the comparative behavior is the stochastic ordering. If X and Y are two continuous random variables, then we can say that X is smaller than Y in the:

- Stochastic ordering ($X \le_{st} Y$) if $\bar{F}_X(t) \le \bar{F}_Y(t)$, for all $t \ge 0$
- Hazard rate ordering ($X \le_{hr} Y$) if $h_x(t) \le h_Y(t)$, for all $t \ge 0$
- Mean residual life ordering ($X \le_{mrl} Y$) if $m_X(t) \le m_Y(t)$, for all $t \ge 0$
- Likelihood ratio ordering ($X \le_{lr} Y$) if $\dfrac{f_X(t)}{f_Y(t)}$

  decreasing, for all $t \ge 0$

According to the well-known result of Shaked and Shanthikumar (1994) for establishing stochastic ordering of distribution we have:

$$X \le_{lr} Y \Rightarrow X \le_{hr} Y \Rightarrow X \le_{mrl} Y$$
$$\Downarrow$$
$$X \le_{st} Y \tag{10}$$

In the following proposition, we show that the UTNB distribution is ordered with respect to the strongest "likelihood ratio" ordering.

Proposition 4

Let $X \sim UTNB(\alpha_1, \beta_1, \theta_1)$ and $Y \sim UTNB(\alpha_2, \beta_2, \theta_2)$, then $(X \le_{lr} Y)$ and hence $(X \le_{hr} Y)$, $(X \le_{mrl} Y)$ and $(X \le_{st} Y)$ if and only if one of the following cases is satisfied:

- If $(\alpha_1 = \alpha_2, \theta_1, \theta_2)$ and $(\alpha > 1, \beta_1 < \beta_2)$ or $(\alpha < 1, \beta_1 > \beta_2)$
- If $(\alpha_1 = \alpha_2, \beta_1 = \beta_2)$ and $(\alpha > 1, \theta_1, \theta_2)$ or $(\alpha < 1, \theta_1, \theta_2)$
- If $(\theta_1 = \theta_2, \beta_1 = \beta_2)$ and $(\alpha_1 < \alpha_2)$

Proof

It is sufficient to show that $(X \le_{lr} Y)$ if $\dfrac{g_X(t)}{g_Y(t)}$

decreasing, for all $t \ge 0$ in this cases. Then we get the required results.

Stress-Strength Reliability

If a system subject to a stress, the reliability of the system is defined as the probability that it accomplishes its required function adequately without failure. For stress-strength models both the strength of the system, X and the stress, Y, imposed on it by its operating environments are considered to be random variables. The reliability, R, of the system can be stated as, $R = P(Y<X)$. Now, Let X and Y are two independent but not identically distributed UTNB random variables represent the strength of a system and the stress acting on it respectively. Suppose that $X \sim UTNB(\alpha, \beta_1, \theta)$ and $Y \sim UTNB(\alpha, \beta_2, \theta)$. That is the survival functions of X and Y are respectively given by:

$$\bar{G}_1(x) = \frac{\alpha^{\beta_1}}{1-\alpha^{\beta_1}}\left[\left(\alpha + \frac{1-\alpha}{\theta}x\right)^{-\beta_1} - 1\right]$$

and:

$$\bar{G}_2(y) = \frac{\alpha^{\beta_2}}{1-\alpha^{\beta_2}}\left[\left(\alpha + \frac{1-\alpha}{\theta}y\right)^{-\beta_2} - 1\right]$$

Then the reliability of the system will be:

$$R = P(Y < X) = \int_0^\theta P(X > Y \mid Y = y).g_2(y)dy$$

$$= \int_0^\theta \bar{G}_1(y).g_2(y)dy$$

$$= \frac{(1-\alpha)\beta_2\alpha^{\beta_1+\beta_2}}{\theta(1-\alpha^{\beta_2})(1-\alpha^{\beta_1})}\int_0^\theta \left\{ \begin{matrix} \left(\left(\alpha + \frac{(1-\alpha)y}{\theta}\right)^{-\beta_1} - 1\right) \\ \times\left(\alpha + \frac{(1-\alpha)y}{\theta}\right)^{-\beta_2-1} \end{matrix} \right\} dy$$

$$= \frac{\beta_2 - \alpha^{\beta_1}\left(-\beta_1\alpha^{\beta_2} + \beta_2 + \beta_1\right)}{\left(\alpha^{\beta_1} - 1\right)\left(\alpha^{\beta_2} - 1\right)\left(\beta_2 + \beta_1\right)}$$

Entropies

The entropy of a random variable X is a measure of uncertainty or randomness of dynamical system modeled by this random variable. We derive two popular entropy measures that are Rényi entropy, Song (2001) and Shannon's entropy, Shannon (1951) of the UTNB random variable. The Rényi entropy is defined as:

$$I_R(\gamma) = (1-\gamma)^{-1}\log\left(\int_0^\infty g^\gamma(x)dx\right) \tag{11}$$

where, $\gamma > 0$ and $\gamma \ne 1$. If $X \sim UTNB(\alpha, \beta, \theta)$ then we have:

$$\int_0^\theta g^\gamma(x)\,dx = \int_0^\theta \left(\frac{(1-\alpha)\beta\alpha^\beta}{(1-\alpha^\beta)\theta}\right)^\gamma \left(\alpha + \frac{1-\alpha}{\theta}x\right)^{-\gamma(\beta+1)} dx$$

$$= \frac{(1-\alpha)^\gamma\beta^\gamma\alpha^{\gamma\beta}}{(1-\alpha^\beta)^\gamma\theta^\gamma}\int_0^\theta \left(\alpha + \frac{1-\alpha}{\theta}x\right)^{-\gamma(\beta+1)} dx$$

Putting $w = \alpha + \dfrac{1-\alpha}{\theta}x$, then we have:

$$\int_0^\theta g^\gamma(x)\,dx = \frac{(1-\alpha)^{\gamma-1}\beta^\gamma\alpha^{\gamma\beta}}{(1-\alpha^\beta)^\gamma\theta^{\gamma-1}}\int_\alpha^1 w^{-\gamma(\beta+1)}\,dw$$

$$= \frac{(1-\alpha)^{\gamma-1}\beta^\gamma\alpha^{\gamma\beta}\left(\alpha^{1-\gamma(\beta+1)}-1\right)}{(1-\alpha^\beta)^\gamma\theta^{\gamma-1}\left(1-\gamma(\beta+1)\right)}$$

Therefore, Rényi entropy is given by:

$$I_R(\gamma) = (1-\gamma)^{-1}\log\left(\frac{(1-\alpha)^{\gamma-1}\beta^\gamma\alpha^{\gamma\beta}\left(\alpha^{1-\gamma(\beta+1)}-1\right)}{(1-\alpha^\beta)^\gamma\theta^{\gamma-1}(1-\gamma(\beta+1))}\right) \qquad (12)$$

Now we consider the Shannon entropy which is defined as $E[-\log(g(X))]$. It is the special case of the Rényi entropy for $\gamma\uparrow 1$. For the UTNB random variable we have:

$$E\left[-\log\left(g(X)\right)\right] = -\log\left((1-\alpha)\beta\alpha^\beta\right)$$
$$+\log\left((1-\alpha^\beta)\theta\right) - E\left[(\beta+1)\log\left(\alpha + \frac{1-\alpha}{\theta}X\right)\right] \qquad (13)$$

First, we derive the expectation:

$$E\left[\log\left(\alpha + \frac{1-\alpha}{\theta}X\right)\right] = \frac{\log(\alpha)}{1-\alpha^\beta} - \frac{1}{\beta}$$

Therefore, we get:

$$E\left[-\log\left(g(X)\right)\right] = \log\left((1-\alpha^\beta)\theta\right)$$
$$-\log\left((1-\alpha)\beta\alpha^\beta\right) + (\beta+1)\left(\frac{\log(\alpha)}{1-\alpha^\beta} - \frac{1}{\beta}\right) \qquad (14)$$

Maximum Likelihood Estimation

Suppose that $X_1, X_2, \ldots, X_n$ be the failure time of individual i and let $Y_1, Y_2, \ldots, Y_n$ be the corresponding censoring time. Assuming that X_i's and Y_i's are independent. If X_i's ~UTNB and Y_i are assumed to have a non-informative distribution. One observes the pairs (T_i, Δ_i) where $T_i = \min(X_i, Y_i)$ and $\Delta_I = I(X_i \le Y_i)$ is the censoring indicator. Given the data (t_1, δ_1), (t_2, δ_2), …, (t_n, δ_n) then the likelihood function is given by:

$$L_n(\alpha, \beta, \theta) = \prod_{i=1}^n [g(t_i)]^{\delta_i}\left[\bar{G}(t_i, \alpha, \beta, \theta)\right]^{1-\delta_i}$$

$$= \prod_{i=1}^n \left\{ \left(\frac{(1-\alpha)\beta\alpha^\beta}{(1-\alpha^\beta)\theta}\left(\alpha + \frac{1-\alpha}{\theta}t_i\right)^{-(\beta+1)}\right)^{\delta_i} \times \left(\frac{\alpha^\beta}{1-\alpha^\beta}\left[\left(\alpha + \frac{1-\alpha}{\theta}t_i\right)^{-\beta} - 1\right]\right)^{1-\delta_i} \right\}$$

and the log-likelihood function is then given by:

$$l_n = \sum_{i=1}^n \left\{ \begin{array}{l} \delta_i\log\left(\frac{(1-\alpha)\beta}{\theta}\right) + \beta\log(\alpha) \\ -\log(1-\alpha^\beta) - \delta_i(\beta+1)\log\left(\alpha + \frac{1-\alpha}{\theta}t_i\right) \\ +(1-\delta_i)\log\left(\left(\alpha + \frac{1-\alpha}{\theta}t_i\right)^{-\beta} - 1\right) \end{array} \right\}$$

Taking the first partial derivative of the log-likelihood function with respect to α, β, θ we have:

$$\frac{\partial l_n}{\partial \alpha} = \sum_{i=1}^n \left\{ \begin{array}{l} \frac{\beta\alpha^{\beta-1}}{1-\alpha^\beta} + \frac{\beta}{\alpha} - \frac{\delta_i}{1-\alpha} - \frac{(\beta+1)\delta_i\left(1-\frac{t_i}{\theta}\right)}{\alpha + \frac{\alpha-1}{\theta}t_i} \\ -\frac{\beta(1-\delta_i)\left(1-\frac{t_i}{\theta}\right)\left(\alpha + \frac{\alpha-1}{\theta}t_i\right)^{-\beta-1}}{\left(\alpha + \frac{\alpha-1}{\theta}t_i\right)^{-\beta} - 1} \end{array} \right\}$$

$$\frac{\partial l_n}{\partial \beta} = \sum_{i=1}^n \left\{ \begin{array}{l} \frac{\log(\alpha)}{1-\alpha^\beta} + \frac{\delta_i}{\beta} - \delta_i\log\left(\alpha + \frac{\alpha-1}{\theta}t_i\right) \\ -\frac{(1-\delta_i)\left(\alpha + \frac{\alpha-1}{\theta}t_i\right)^{-\beta}\log\left(\alpha + \frac{\alpha-1}{\theta}t_i\right)}{\left(\alpha + \frac{\alpha-1}{\theta}t_i\right)^{-\beta} - 1} \end{array} \right\}$$

and:

$$\frac{\partial l_n}{\partial \theta} = \sum_{i=1}^n \left\{ \begin{array}{l} -\frac{\delta_i}{\theta} - \frac{\beta(1-\delta_i)\left(\frac{\alpha t_i}{\theta^2} - \frac{t_i}{\theta^2}\right)\left(\alpha + \frac{\alpha-1}{\theta}t_i\right)^{-\beta-1}}{\left(\alpha + \frac{\alpha-1}{\theta}t_i\right)^{-\beta} - 1} \\ -\frac{(\beta+1)\delta_i\left(\frac{\alpha t_i}{\theta^2} - \frac{t_i}{\theta^2}\right)}{\alpha + \frac{\alpha-1}{\theta}t_i} \end{array} \right\}$$

Maximum likelihood estimates (mles) $\hat{\alpha}, \hat{\beta}, \hat{\theta}$ are obtained numerically as the solution of the non-linear equations:

$$\frac{\partial l_n}{\partial \alpha} = 0, \frac{\partial l_n}{\partial \beta} = 0, \frac{\partial l_n}{\partial \theta} = 0$$

Application to Censored Data

In this section, we fit the proposed UTNB model to randomly censored data representing the serum reversal times (in years) of a random sample of 148 children contaminated with HIV from vertical transmission at the university hospital of Ribeirão Preto school of medicine from 1986 to 2001, Silva (2004):

0.0027+	0.0027+	0.0027+	0.0027+	0.0055+	0.0055+
0.0055+	0.0055+	0.0055+	0.0055+	.0082+	0.0082+
0.0082+	0.0082+	0.0082+	0.0082+	0.0082+	0.011+
0.0137	0.0164+	0.0164+	0.0219+	0.0247+	0.0247
0.0274+	0.0301+	0.0274+	0.0384+	0.0384+	0.0438+
0.0521+	0.0521+	0.0521+	0.0548+	0.0575+	0.0603+
0.0658+	0.0658+	0.0822+	0.0849+	0.0877+	0.0877+
0.0904+	0.0932+	0.1068+	0.1068+	0.1068+	0.126 +
0.1288+	0.137+	0.1534+	0.1534+	0.1644	0.1671+
0.1753+	0.189+	0.2164+	0.2493+	0.2493+	0.2548+
0.2603	0.2658+	0.2712+	0.274+	0.2767+	0.2849+
0.2904+	0.2932 +	0.3068+	0.3534	0.3534	0.3589
0.4055	0.4082	0.4247+	0.4274	0.4356+	0.4712
0.4795+	0.4822	0.4822	0.4877+	0.4877+	0.4904+
0.5233+	0.526+	0.537 +	0.5534+	0.5589	0.5589
0.5726	0.5781	0.5863	0.6164+	0.6274	0.6301+
0.6521	0.6904+	0.6932	0.6959	0.7151	0.7342+
0.7342	0.737	0.7397	0.7425	0.7479+	0.7507
0.7562	0.7918	0.7945	0.7973	0.7973	0.800
0.8055	0.8137+	0.8137	0.8137	0.8274	0.8466+
0.8795+	0.8822	0.8822	0.8877	0.9151	0.9151
0.9151	0.9205	0.9288+	0.9288	0.9342	0.937+
0.9425	0.9479	0.9507+	0.9507+	0.9534	0.9589+
0.9616	0.9616	0.9671	0.9671	0.9726	0.9736
0.9808	0.9836	0.9918+	1.000		

+: Censored data

The data set contains 84 right censored observations which constitute 56.7% of the sample size (heavy censoring). We have fitted the uniform and the MOEU distributions to the data set using MLE and then compared it with the UTNB distribution. In addition, for model selection, we use the Akiake Information Criterion (AIC), the Bayesian Information Criterion (BIC), the Consistent Akiake Information Criterion (CAIC) and the Hannan-Quinn Information Criterion (HQIC) defined as:

$$AIC = -2l_n\left(\hat{\alpha}, \hat{\beta}, \hat{\theta}\right) + 2k,$$

$$BIC = -2l_n\left(\hat{\alpha}, \hat{\beta}, \hat{\theta}\right) - k\log(n),$$

$$CAIC = -2l_n\left(\hat{\alpha}, \hat{\beta}, \hat{\theta}\right) + \frac{2kn}{n-k-1},$$

$$HQIC = -2l_n\left(\hat{\alpha}, \hat{\beta}, \hat{\theta}\right) + 2k\log(\log(n))$$

where, k is the number of parameters in the model and n is the sample size. The model with lowest AIC or (BIC, CAIC and HQIC) value is the one that better fits the data. Table 1 show that the fitted UTNB distribution should be selected based on either the AIC or (BIC, CAIC and HQIC) procedure for the given data.

To test the null hypothesis H_0: The data follows the uniform distribution versus H_a: The data follows the UTNB distribution, we used the Likelihood Ratio Test statistic (LRT) which has an asymptotic chi-square distribution with 2 degrees of freedom. H_0 is rejected at a significance level of 0.05% if $LRT > \chi^2_{(2,0.05\%)}$. The uniform distribution is rejected in favor of the UTNB distribution for this data set ($LRT = 77.5203 > \chi^2_{(2,0.05)} = 5.9915$). This result shows that UTNB is a great improvement over the uniform distribution for this censored data set.

Let n be the total number of breaking stress of carbon fibers whose survival times, uncensored data, are available. Retable the n survival times in order of increasing magnitude such that $t_1 \leq t_2 \leq \ldots \leq t_n$. The Kaplan and Meier (1958), estimator (KME), also known as the product limit estimator, gave a non-parametric estimator of the survival function of the underlying as follows:

$$\bar{S}_n(t) = \prod_{t:t_i \leq t} 1 - \frac{\delta(i)}{n-i+1} \quad t > 0$$

Table 1. A comparison between the Log-likelihood, AIC, BIC, CAIC and HQIC for the fitted UTNB, MOEU and uniform distributions

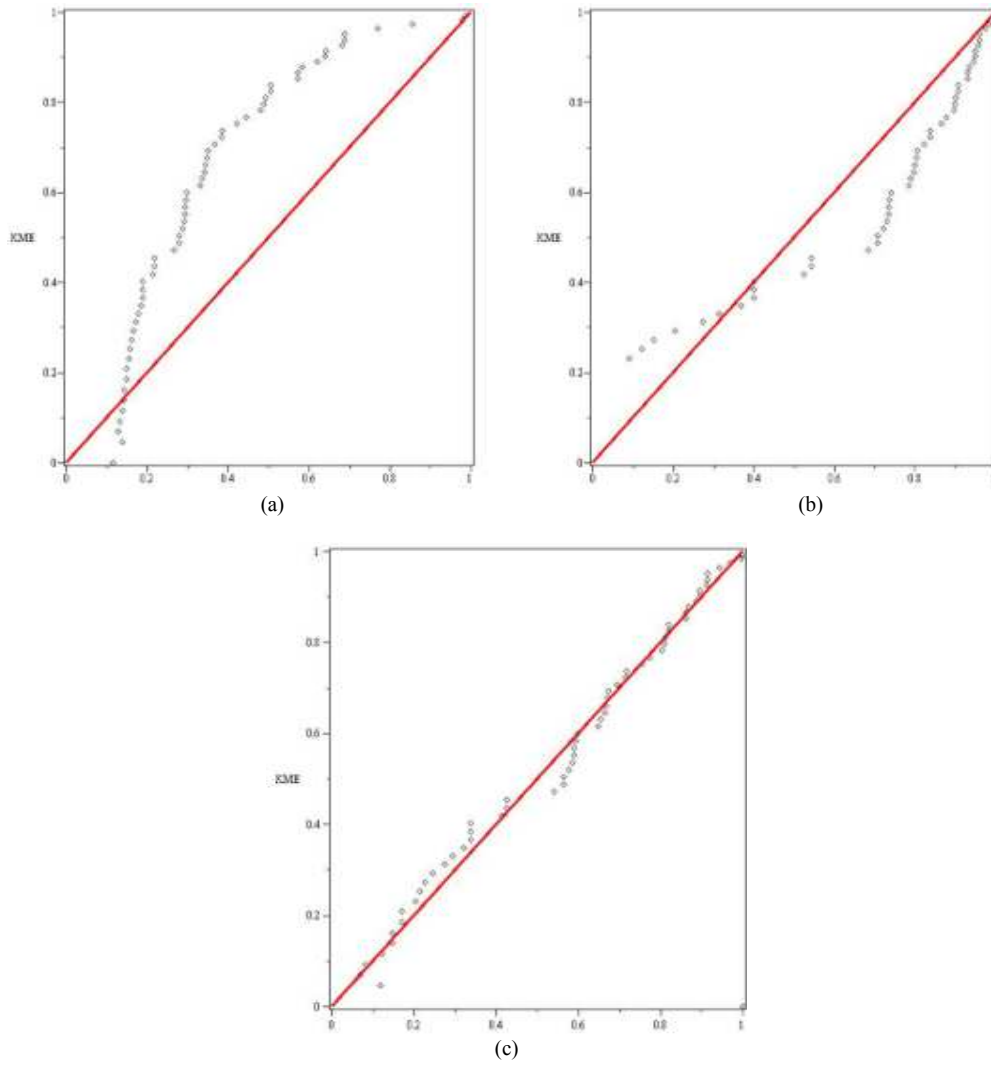| Models	Parameters	MLE	Log-likelihood	AIC	BIC	CAIC	HQIC
UTNB	$\hat{\alpha}$	5.44	0.87	4.26	-16.732	4.43	7.91
	$\hat{\beta}$	0.21					
	$\hat{\theta}$	0.99					
MOEU	$\hat{\alpha}$	13.42	-8.99	21.98	7.99	22.07	24.42
	$\hat{\theta}$	0.96					
Uniform	$\hat{\theta}$	1.12	-37.89	77.79	70.79	75.81	79.01

(a)

(b)

(c)

Fig. 4. P-P plot of KME versus fitted UTNB, MOEU and uniform survival functions (a) fitted uniform (b) fitted MOEU (c) fitted UTNB
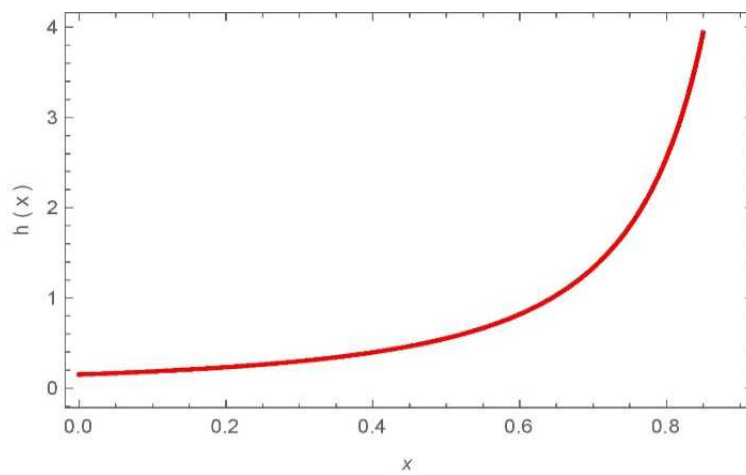

Fig. 5. The estimated hrf of the UTNB distribution based on the given data

Figure 4 shows respectively the P-P plot of the KME of the fitted UTNB, MOEU and uniform distributions for the given data. The figures indicate the goodness-of-fit of the fitted UTNB as compared with the fitted uniform and MOEU survival functions.

Also, whereas $\hat{\alpha}^{\hat{\beta}}\left(1+\hat{\beta}\right)=1.73\geq1$ according to proposition (2) the estimated hrf is increasing as shown in Fig. 5.

Conclusion

For the first time, we propose a three-parameter continuous distribution which generalizes the uniform and the Marshall-Olkin extended uniform distributions. We refer to the new model as the Uniform Truncated Negative Binomial (UTNB) distribution. Provide several statistical properties for the proposed model including reliability measures: The density, the hazard rate, mean residual life, moment generating function, moments, quantiles, kurtosis, skewness, limiting distribution of sample extremes, stochastic orderings, stress-strength reliability and entropies. Estimation via maximum likelihood is straightforward. Censored data application of the UTNB distribution shows that it could provide a better fit than the uniform and the MOEU distributions. We hope that the proposed extended model may attract wider applications in survival analysis.

Acknowledgement

The authors are grateful to the referees for their comments and suggestions, which certainly improved the presentation of the content of the paper.

Author's Contributions

Bahady I. Kamel: Offering ideas, encouragement, proof read text and equations.

Shaban E. Abo Youssef: Revise, improve and editing thefinal drafts of the paper.

Mohamed G. Sief: Analyzed and interpreted the result, prepared the manuscript and responsible for the manuscript correction, proof reading and paper submission.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Alice, T. and K. Jose, 2005a. Marshall-Olkin logistic processes. STARS Int. J., 6: 1-11.

Alice, T. and K. Jose, 2005b. Marshall-Olkin semi-Weibull minification processes. Recent Adv. Stat. Theory Applied.

Arnold, B.C., N. Balakrishnan and H.N. Nagaraja, 1992. A First Course in Order Statistics. 1st Edn., SIAM, Philadelphia, ISBN-10: 0898719062, pp: 279.

Cox, D.R. and D. Oakes, 1984. Analysis of Survival Data. CRC Press, ISBN-10: 041224490X, pp: 208.

Ghitany, M., E. Al-Hussaini and R. Al-Jarallah, 2005. Marshall-Olkin extended Weibull distribution and its application to censored data. J. Applied Stat., 32: 1025-1034. DOI: 10.1080/02664760500165008

Ghitany, M., 2006. Marshall-Olkin extended Pareto distribution and its application .Proceedings of the 30th International Conference of the Forum for Interdisciplinary Mathematics on Interdisciplinary Mathematical and Statistical Techniques, Sept. 1-4, New University of Lisbon-Tomar Polytechnic Institute, Portugal.

Ghitany, M.E., F. Al-Awadhi and L. Alkhalfan, 2007. Marshall-Olkin extended Lomax distribution and its application to censored data. Commun. Stat. Theory Meth., 36: 1855-1866.
DOI: 10.1080/03610920601126571

Jayakumar, K. and T. Mathew, 2008. On a generalization to Marshall-Olkin scheme and its application to Burr type XII distribution. Stat. Papers, 49: 421-439.
DOI: 10.1007/s00362-006-0024-5

Jose, K., S.R. Naik and M.M. Ristić, 2010. Marshall-Olkin q-Weibull distribution and max-min processes. Stat. Papers, 51: 837-851.
DOI: 10.1007/s00362-008-0173-9

Jose, K. and E. Krishna, 2011. Marshall-Olkin extended uniform distribution. ProbStat Forum, 4: 78-88.

Jose, K. and R. Sivadas, 2015. Negative binomial Marshall-Olkin Rayleigh distribution and its applications. Economic Quality Control, 30: 89-98.
DOI: 10.1515/eqc-2015-0009

Kaplan, E.L. and P. Meier, 1958. Nonparametric estimation from incomplete observations. Am. Stat. Assoc., 53: 457-481.
DOI: 10.1080/01621459.1958.10501452

Keeping, E.S., 1962. Introduction to Statistical Inference. 1st Edn., Courier Corporation, New York, ISBN-10: 0486685020, pp: 451.

Marshall, A. and I. Olkin, 1997. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika, 84: 641-652.
DOI: 10.1093/biomet/84.3.641

Moors, J.J.A., 1988. A quantile alternative for kurtosis. Statistician, 37: 25-32. DOI: 10.2307/2348376

Nadarajah, S., K. Jayakumar and M.M. Ristić, 2013. A new family of lifetime models. J. Stat. Comput. Simulat., 83: 1389-1404.
DOI: 10.1080/00949655.2012.660488

Shaked, M. and J.G. Shanthikumar, 1994. Stochastic Orders and their Applications. Academic Press, New York, ISBN-10: 0126386102, pp: 545.

Shannon, C.E., 1951. Prediction and entropy of printed english. Bell Syst. Tech. J., 30: 50-64. DOI: 10.1002/j.1538-7305.1951.tb01366.x

Silva, A.N., 2004. Estudo evolutivo das crianças expostas ao HIV e notificadas pelo ncleo de vigilância epidemiolǵica do HCFMRP-USP [dissertation]. Universidade de Sﻣo Paulo.

Song, K.S., 2001. Rényi information, loglikelihood and an intrinsic distribution measure. J. Stat. Plann. Inference, 93: 51-69. DOI: 10.1016/S0378-3758(00)00169-5

Watson, G.S. and W.T. Wells, 1961. On the possibility of improving the mean useful life of items by eliminating those with short lives. Technometrics, 3: 281-298. DOI: 10.1080/00401706.1961.10489946