International Journal of Statistics and Applications 2016, 6(3): 168-176
DOI: 10.5923/j.statistics.20160603.10

On a Generalization of Weibull Distribution and Its Applications

M. Girish Babu

Department of Statistics, Government Arts and Science College, Meenchanda, Kozhikode, India

Abstract The Weibull and related models have been used in many applications for solving a variety of problems from many disciplines. Here we introduce a new family of distributions, namely Weibull-truncated negative binomial distribution (WTNB) and study some properties of it. Exponential-truncated negative binomial (ETNB) and Marshall-Olkin Weibull (MOW) distribution are special cases of this distribution. We have analyzed a real data set of serum creatinine values and found that this new distribution is a good fit to model the data, compared to Weibull, Gamma, Exponentiated Weibull, ETNB and MOW models.

Keywords Exponential – truncated negative binomial distribution, Marshall-Olkin family of distributions, Shannon entropy, Weibull distribution

1. Introduction

The Weibull distribution is a very popular distribution for modeling lifetime data and for modeling phenomenon with monotone failure rates. The distribution is named after Waloddi Weibull who was the first to promote the usefulness of this to model the breaking strength of materials (Weibull, 1939). A similar model was proposed earlier by Rosin and Rammler (1933) in the context of modeling the variability in the diameter of powder particles being greater than a specific size. The earlier known publication dealing with the Weibull distribution is a work by Fisher and Tippet (1928) where this distribution is obtained as the limiting distribution of the smallest extremes in a sample. Gumbel (1958) refers to the Weibull distribution as the third asymptotic distribution of the smallest extremes. The Weibull, and related models have been used in many applications, and for solving a variety of problems from many disciplines. Jayakumar and Girish (2015) studied some generalizations of Weibull distribution and related time series models.

Marshall and Olkin (1997) proposed a new method of generating a family of distributions by introducing an additional parameter. Many authors have studied properties of various univariate distributions belonging to the family of Marshall-Olkin distributions, see, Alice and Jose (2003, 2005), Ghitany et al. (2005, 2007) and Jayakumar and Thomas (2008).

A generalization of the Marshall-Olkin distributions was introduced by Nadarajah et al. (2013) as follows: Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables with survival function $\bar{F}(x)$. Le N be a truncated negative binomial random variable with parameters $\alpha \in (0,1)$ and $\theta > 0$, that is,

$$P(N = n) = \frac{\alpha^\theta}{1-\alpha^\theta}\binom{\theta + n - 1}{\theta - 1}(1-\alpha)^n,\ n = 1,2, \ldots .$$

Consider $U_N = \min(X_1, X_2, \ldots, X_N)$.

Then $P(U_N > x) = \bar{G}(x; \alpha, \theta)$

$$= \frac{\alpha^\theta}{1-\alpha^\theta}\sum_{n=1}^{\infty}\binom{\theta + n - 1}{\theta - 1}\left((1-\alpha)\bar{F}(x)\right)^n$$

$$= \frac{\alpha^\theta}{1-\alpha^\theta}\left[\left(F(x) + \alpha\bar{F}(x)\right)^{-\theta} - 1\right] \tag{1.1}$$

Similarly, if $\alpha > 1$ and N is a truncated negative binomial random variable with parameters α^{-1} and $\theta > 0$, then $V_N = \max(X_1, X_2, \ldots, X_N)$ also has the survival function given in (1.1). Here note that in (1.1), if $\alpha \to 1$, then $\bar{G}(x; \alpha, \theta) \to \bar{F}(x)$. If $\theta = 1$, then this family reduces to the family of Marshall – Olkin distributions. Thus the family of distributions described in (1.1) is a generalization of the family of Marshall-Olkin distributions.

This family can be interpreted as follows. Suppose the failure times of a device are observed. Every time a failure occurs the device is repaired to resume function. Suppose also that the device is deemed no longer useable when a failure occurs that exceeds a certain level of severity. Let $X_1, X_2, \ldots$ denote the failure times and N denote the number of failures. Then U_N will represent the time to the first failure of the device and V_N will represent the life time of the device. Thus this family can be used to model both the time to the first failure and the life time.

The probability density function of (1.1) is

* Corresponding author:
giristat@gmail.com (M. Girish Babu)
Published online at http://journal.sapub.org/statistics

$$g(x;\alpha,\theta)=\frac{\theta\alpha^{\theta}(1-\alpha)f(x)}{(1-\alpha^{\theta})\left(F(x)+\alpha\bar{F}(x)\right)^{\theta+1}} \quad (1.2)$$

The hazard rate function is given by

$$h(x;\alpha,\theta)=\frac{\theta(1-\alpha)\bar{F}(x)h_F(x)}{\left(F(x)+\alpha\bar{F}(x)\right)\left[1-\left(F(x)+\alpha\bar{F}(x)\right)^{\theta}\right]} \quad (1.3)$$

where $h_F(x)=\frac{f(x)}{\bar{F}(x)}$ is the hazard rate corresponding to F.

Nadarajah et al. (2013) introduced and studied a new family of distribution as exponential-truncated negative binomial distribution with parameters α,θ and λ by substituting $\bar{F}(x)=e^{-\lambda x},\lambda>0,x>0$ in the survival function (1.1).

That is,

$$\bar{G}(x;\alpha,\theta,\lambda)=\frac{\alpha^{\theta}}{1-\alpha^{\theta}}\left[\left(1-e^{-\lambda x}+\alpha e^{-\lambda x}\right)^{-\theta}-1\right] \quad (1.4)$$

for $\alpha>0,\theta>0,\lambda>0$ and $x>0$.

The paper is organized as follows. In section 2, we propose a new family of distributions, namely Weibull-truncated negative binomial distribution (WTNB). We study some properties of WTNB, such as the behavior of hazard rate, moments, Shannon and Renyi entropies and distributions of order statistics. Also the maximum likelihood method of estimation is used to obtain the estimates of parameters of WTNB. In section 3, we analyze a real data set of serum creatinine (mg.dL) values and found that WTNB is the most appropriate model for the data. The performance of the WTNB is compared to the well known models such as Weibull, Gamma, exponentiated Weibull, ETNB and Marshall-Olkin Weibull using AIC, BIC, K-S statistic and P-values.

2. Weibull-Truncated Negative Binomial Distribution

We propose a new family of distribution named as Weibull-truncated negative binomial (WTNB) distribution with parameters $\alpha>0$, $\theta>0$, $c>0$ and $x>0$ by substituting $\bar{F}(x)=e^{-x^{c}},c>0,x>0$ in the survival function (1.1). Then,

$$\bar{G}(x;\alpha,\theta,c)=\frac{\alpha^{\theta}}{1-\alpha^{\theta}}\left[\left(1-e^{-x^{c}}+\alpha e^{-x^{c}}\right)^{-\theta}-1\right] \quad (2.1)$$

The probability density function of this distribution is

$$g(x;\alpha,\theta,c)=\frac{(1-\alpha)\theta\alpha^{\theta}cx^{c-1}e^{-x^{c}}}{(1-\alpha^{\theta})\left(1-e^{-x^{c}}+\alpha e^{-x^{c}}\right)^{\theta+1}} \quad (2.2)$$

The hazard rate function is given by

$$h(x;\alpha,\theta,c)=\frac{(1-\alpha)\theta cx^{c-1}e^{-x^{c}}}{\left(1-e^{-x^{c}}+\alpha e^{-x^{c}}\right)\left[1-\left(1-e^{-x^{c}}+\alpha e^{-x^{c}}\right)^{\theta}\right]} \quad (2.3)$$

The shape of density and hazard rate functions are shown in the following figures.

The cumulative probabilities at different choices of parameters are computed by using MATHCAD software and the results are shown in Table 2.1.

Some distributions arise as special cases of the WTNB(α,θ,c)

Case I : For $c=1$,

$$\bar{G}(x;\alpha,\theta)=\frac{\alpha^{\theta}}{1-\alpha^{\theta}}\left[(1-e^{-x}+\alpha e^{-x})^{-\theta}-1\right] \quad (2.4)$$

This is the survival function of two parameter Exponential-truncated negative binomial distribution.

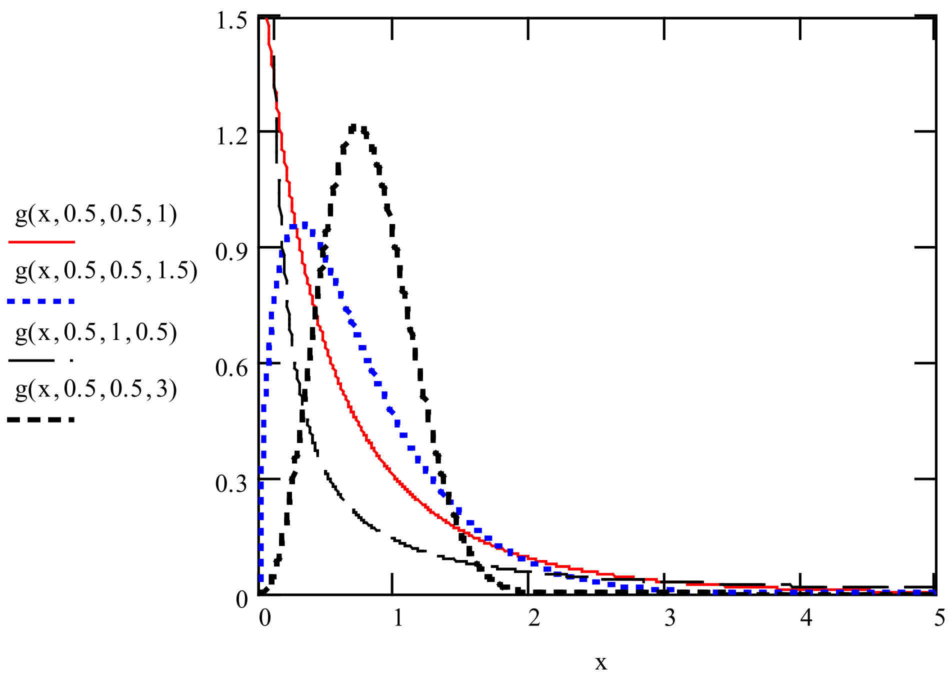


Figure 2.1. Probability density function of WTNB distribution

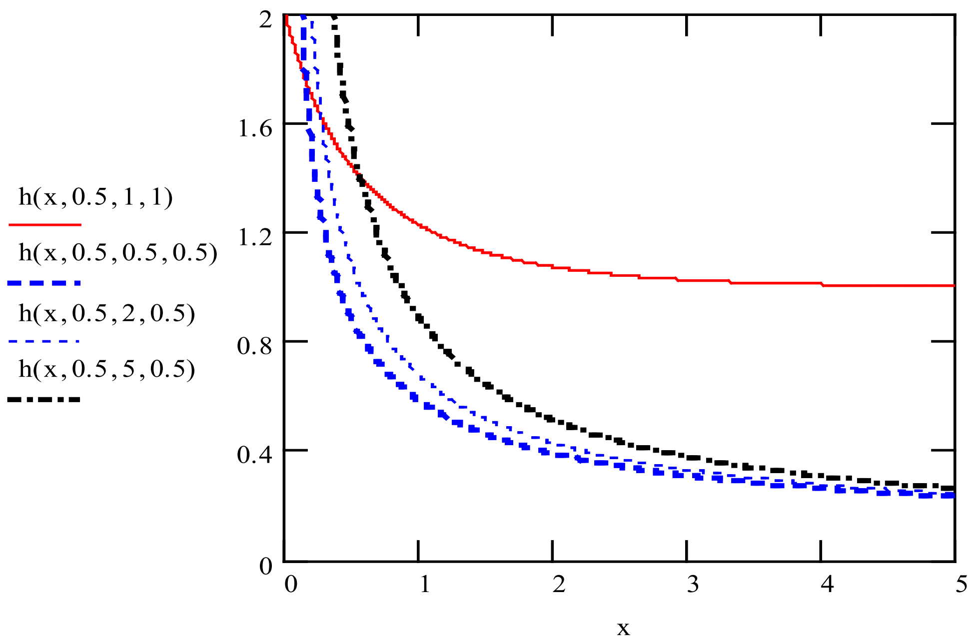


Figure 2.2. Hazard rate function of WTNB distribution

Table 2.1. Cumulative Probabilities of WTNB distribution

X	$\theta = 1$, $\alpha = 0.5$, $c = 0.5$	$\theta = 0.5$, $\alpha = 0.5$, $c = 0.5$	$\theta = 1.5$, $\alpha = 0.5$, $c = 0.5$	$\theta = 2$, $\alpha = 0.5$, $c = 0.5$
0	0.775	0.996	0.805	0.833
1	0.862	0.997	0.883	0.901
2	0.903	0.997	0.918	0.932
3	0.927	0.997	0.939	0.95
4	0.944	0.997	0.953	0.961
5	0.955	0.998	0.963	0.969
6	0.963	0.998	0.97	0.975
7	0.97	0.998	0.975	0.979
8	0.974	0.998	0.979	0.983
9	0.978	0.998	0.982	0.985
10	0.982	0.999	0.985	0.988
11	0.984	0.999	0.987	0.989
12	0.986	0.999	0.989	0.991
13	0.988	0.999	0.99	0.992
14	0.989	0.999	0.991	0.993
15	0.991	0.999	0.992	0.994
16	0.992	0.999	0.993	0.995
17	0.993	0.999	0.994	0.995
18	0.994	0.999	0.995	0.996
19	0.994	0.999	0.995	0.996
20	0.995	0.999	0.996	0.997
21	0.995	0.999	0.996	0.997
22	0.996	0.999	0.997	0.997
23	0.996	0.999	0.997	0.998
24	0.997	0.999	0.997	0.998
25	0.997	1	0.997	0.998
26	0.997	1	0.998	0.998
27	0.997	1	0.998	0.998
28	0.998	1	0.998	0.998
29	0.998	1	0.998	0.999
30	0.998	1	0.998	0.999
31	0.998	1	0.999	0.999
32	0.998	1	0.999	0.999
33	0.999	1	0.999	0.999

Case II: For $\theta = 1$,

$$\bar{G}(x,\alpha,c) = \frac{\alpha\, e^{-x^c}}{1-(1-\alpha)e^{-x^c}} \tag{2.5}$$

This is the survival function of Marshall-Olkin extended Weibull distribution with parameters α and c.

Case III: For $\theta = 1, \alpha = 1$

$$\bar{G}(x,c) = e^{-x^c} \tag{2.6}$$

This is the survival function of one parameter Weibull distribution.

Case IV: For $\theta = 1,\ \alpha = 2$

$$\bar{G}(x,c) = \frac{2}{1+e^{x^c}} \tag{2.7}$$

This is the survival function of the generalized half logistic distribution.

Case V: When $\alpha \to 1$, WTNB(α,θ,c) reduces to the one parameter Weibull distribution.

A random sample from WTNB(α,θ,c) distribution can be simulated as

$$X = \left[-\ln\left[\frac{1-\alpha\left[\alpha^{\theta}+Y(1-\alpha^{\theta})\right]^{\frac{-1}{\theta}}}{1-\alpha}\right]\right]^{\frac{1}{c}} \quad \text{for } Y\sim U(0,1). \tag{2.8}$$

2.1. Moments

Suppose that X has the WTNB(α,θ,c) distribution. The nth moment can be written as

$$E(X^n) = \frac{(1-\alpha)\theta\alpha^\theta c}{(1-\alpha^\theta)}\int_0^\infty \frac{x^{n+c-1}e^{-x^c}}{[1-(1-\alpha)e^{-x^c}]^{\theta+1}}dx \tag{2.9}$$

Taking $u = e^{-x^c}$, (2.9) reduces to

$$E(X^n) = \frac{(1-\alpha)\theta\alpha^\theta}{(1-\alpha^\theta)}\int_0^1 \frac{(-\log(u))^{\frac{n}{c}}}{[1-(1-\alpha)u]^{\theta+1}}du \tag{2.10}$$

If $|1-\alpha| < 1$, then by the series expansion $(1-x)^{-m} = \sum_{k=0}^{\infty}\binom{m+k-1}{m-1}x^k$, equation (2.10) can be written as

$$E(X^n) = \frac{(1-\alpha)\theta\alpha^\theta}{(1-\alpha^\theta)}\sum_{k=0}^{\infty}\binom{\theta+k}{\theta}(1-\alpha)^k\int_0^1 u^k(-\log u)^{\frac{n}{c}}du$$

where $\int_0^1 u^k(-\log u)^{\frac{n}{c}}du = \frac{\left(\frac{n}{c}\right)\left(\frac{n}{c}-1\right)\ldots\left(\frac{n}{c}-(n-1)\right)}{(k+1)^{n+1}}$

Therefore,

$$E(X^n) = \frac{(1-\alpha)\theta\alpha^\theta\left(\frac{n}{c}\right)\left(\frac{n}{c}-1\right)\ldots\left(\frac{n}{c}-(n-1)\right)}{(1-\alpha^\theta)}\sum_{k=0}^{\infty}\binom{\theta+k}{\theta}\frac{(1-\alpha)^k}{(k+1)^{n+1}}. \tag{2.11}$$

If $|1-\alpha| < \alpha$, then we have

$$E(X^n) = \frac{(1-\alpha)\theta}{(1-\alpha^\theta)\alpha}\int_0^1 \frac{(-\log(1-u))^{\frac{n}{c}}}{[1+\left(\frac{1-\alpha}{\alpha}\right)u]^{\theta+1}}du \tag{2.12}$$

Taking $u = 1-v$ in equation (2.12) and using equation (2.6.5.3) of Prudnikov et al. (1986), we have

$$\begin{aligned}
E(X^n) &= \frac{(1-\alpha)\theta}{(1-\alpha^\theta)\alpha}\int_0^1 \frac{(-\log v)^{\frac{n}{c}}}{[1+\left(\frac{1-\alpha}{\alpha}\right)(1-v)]^{\theta+1}}dv \\
&= \frac{(1-\alpha)\theta}{(1-\alpha^\theta)\alpha}\int_0^1 (-\log v)^{\frac{n}{c}} \\
&= \frac{\theta}{1-\alpha^\theta}\sum_{k=0}^{\infty}\binom{\theta+k}{\theta}(-1)^k\left(\frac{1-\alpha}{\alpha}\right)^{k+1}\int_0^1(-\log v)^{\frac{n}{c}}(1-v)^k\,dv \\
&\sum_{k=0}^{\infty}(-1)^k\binom{\theta+k}{\theta}\left(\frac{1-\alpha}{\alpha}\right)^k(1-v)^k\;dv \\
&= \frac{\theta\left(\frac{n}{c}\right)\left(\frac{n}{c}-1\right)\ldots\left(\frac{n}{c}-(n-1)\right)}{1-\alpha^\theta}\sum_{k=0}^{\infty}\binom{\theta+k}{\theta}(-1)^k\left(\frac{1-\alpha}{\alpha}\right)^{k+1}\sum_{j=0}^{k}\frac{(-1)^j(k+1-j)_j}{j!(j+1)^{\frac{n}{c}+1}}
\end{aligned} \tag{2.13}$$

2.2. Shannon Entropy

Entropy is a measure of variation or uncertainty. The Renyi entropy of a random variable with probability density function $g(.)$ is defined as

$$I_R(\gamma) = \frac{1}{1-\gamma}\log\int_0^\infty g^\gamma(x)\,dx, \gamma > 0, \gamma \neq 1. \tag{2.14}$$

The Renyi entropy of WTNB(α,θ,c) is

$$I_R(\gamma) = \frac{1}{1-\gamma}\log\int_0^\infty\left[\frac{(1-\alpha)\theta\alpha^\theta cx^{c-1}e^{-x^c}}{(1-\alpha^\theta)(1-(1-\alpha)e^{-x^c})^{\theta+1}}\right]^\gamma dx$$

Letting $u = e^{-x^c}$, we have

$$I_R(\gamma) = \frac{1}{1-\gamma}\log\left\{\left(\frac{(1-\alpha)\theta\alpha^\theta c}{(1-\alpha^\theta)}\right)^\gamma\int_0^1 \frac{(-\log u)^{\gamma\left(1-\frac{1}{c}\right)}u^\gamma}{(1-(1-\alpha)e^{-x^c})^{\gamma(\theta+1)}}\,du\right\} \tag{2.15}$$

The Shannon entropy is given by

$$E[-\log g(X)] = \log\left(\frac{1-\alpha^{\theta}}{(1-\alpha)\theta\alpha^{\theta}c}\right) - (c-1)E(\log X) + E(X^c) + (\theta+1)\,E\left(\log\left(1-(1-\alpha)e^{-X^c}\right)\right) \quad (2.16)$$

2.3. Order Statistics

Let $X_1, X_2, \ldots, X_n$ are independent random variables having the WTNB (α, θ, c) distribution. Let $X_{i:n}$ denote the i^{th} order statistic. The probability density function of $X_{i:n}$ is

$$g_{i:n}(x;\alpha,\theta,c) = \frac{n!}{(i-1)!\,(n-i)!}\, g(x;\alpha,\theta,c)\, G^{i-1}(x;\alpha,\theta,c)\bar{G}^{n-i}(x;\alpha,\theta,c)$$

$$= \frac{(-1)^{n-i}n!(1-\alpha)\theta\alpha^{\theta(n+1-i)}c.x^{c-1}e^{-x^c}}{(i-1)!(n-i)!\left(1-\alpha^{\theta}\right)^n\left(1-(1-\alpha)e^{-x^c}\right)^{\theta+1}}\left[1-\frac{\alpha^{\theta}}{\left(1-(1-\alpha)e^{-x^c}\right)^{\theta}}\right]^{i-1}\left[1-\frac{1}{\left(1-(1-\alpha)e^{-x^c}\right)^{\theta}}\right]^{n-i}. \quad (2.17)$$

Using the binomial series expansion, the probability density function can be written as

$$g_{i:n}(x;\alpha,\theta,c) = \frac{(-1)^{n-i}n!\;\alpha^{\theta(n-i)}}{(i-1)!\,(n-i)!\,(1-\alpha^{\theta})^n}\sum_{k=0}^{i-1}\sum_{l=0}^{n-i}\binom{i-1}{k}\binom{n-i}{l}\frac{(-1)^{k+l}\left(1-\alpha^{\theta(k+l+1)}\right)}{\alpha^{\theta l}(k+l+1)}$$

$$g(x;\alpha,\theta(k+l+1),c). \quad (2.18)$$

This shows that $X_{i:n}$ is a finite mixture of WTNB random variables.

2.4. Estimation

For a given sample$(x_1, x_2, \ldots, x_n)$, the log-likelihood function is given by

$$\log L(\alpha,\theta,c) = n\log\left[\frac{(1-\alpha)\theta\alpha^{\theta}c}{1-\alpha^{\theta}}\right] + (c-1)\sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n}x_i^c - (\theta+1)\sum_{i=1}^{n}\log\left[1-(1-\alpha)e^{-x_i^c}\right] \quad (2.19)$$

The partial derivatives of the log-likelihood function with respect to the parameters are

$$\frac{\partial \log L}{\partial\alpha} = \frac{n\left(-1+(1-\theta)\alpha^{\theta}+\theta\alpha^{\theta-1}\right)}{(1-\alpha)\left(1-\alpha^{\theta}\right)} + \frac{n\theta}{\alpha} - (\theta+1)\sum_{i=1}^{n}\frac{e^{-x_i^c}}{1-(1-\alpha)e^{-x_i^c}} \quad (2.20)$$

$$\frac{\partial \log L}{\partial\theta} = \frac{n\log\alpha}{1-\alpha^{\theta}} + \frac{n}{\theta} - \sum_{i=1}^{n}\log\left(1-(1-\alpha)e^{-x_i^c}\right) \quad (2.21)$$

$$\frac{\partial \log L}{\partial c} = \frac{n}{c} + \sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n}x_i^c\log x_i - (\theta+1)(1-\alpha)\sum_{i=1}^{n}\frac{x_i^c\log x_i e^{-x_i^c}}{1-(1-\alpha)e^{-x_i^c}} \quad (2.22)$$

Equating these partial derivatives equal to zero and solving these equations numerically, we get the maximum likelihood estimates of α, θ and c.

3. Application

In this section, we analyze a real data set and found that WTNB distribution gives the best fit for the data. We consider a real data set of serum creatinine (mg/dL) value of 300 samples reported on February 2nd, 2016 at the Biochemistry Laboratory of Govt. Medical College, Calicut, Kerala. The data are follows:

0.83 1.05 1.37 1.11 0.26 0.39 0.30 0.66 0.65 0.74 0.71 0.64 1.06 0.38 0.88 0.82 0.68 1.51 0.68 1.33 1.05 0.53 1.15 0.77 0.86 1.03 1.21 1.22 1.69 1.70 1.02 0.17 0.79 0.34 0.40 0.60 0.69 0.63 0.76 0.49 0.55 1.42 0.62 0.42 0.50 0.72 0.43 0.46 0.88 1.35 0.48 1.43 0.57 0.58 0.39 0.99 0.85 1.00 0.85 0.73 0.66 0.81 0.43 0.47 0.24 0.46 0.27 1.32 0.42 1.13 0.51 0.59 0.72 1.29 0.58 0.35 0.80 0.93 1.13 0.90 0.67 0.98 0.73 0.35 0.89 1.12 1.35 0.94 0.33 0.59 0.43 1.29 1.14 0.77 0.68 0.38 0.73 0.48 0.33 1.03 0.68 0.47 0.61 0.84 1.00 0.71 1.15 0.58 0.70 0.28 0.23 0.85 0.96 1.02 0.36 0.73 0.53 0.75 0.71 0.62 0.71 0.69 1.25 1.09 0.44 0.86 0.67 0.81 1.30 0.60 0.56 1.02 0.53 0.27 1.64 0.64 0.47 1.29 0.57 0.94 0.36 1.16 1.12 1.07 0.75 1.13 1.34 0.43 1.09 0.79 1.58 0.56 0.83 0.99 1.05 1.04 0.57 1.22 0.52 1.25 0.57 0.75 1.03 0.58 1.87 0.65 1.50 0.57 0.16 1.27 0.81 0.64 1.33 0.94 1.44 0.83 1.02 0.82 1.27 0.83 1.07 0.78 0.41 0.66 1.08 0.21 0.33 0.54 0.55 0.67 1.16 0.93 0.81 0.49 0.68 0.49 0.79 1.26 0.58 0.43 0.88 0.51 1.58 0.71 1.04 0.91 0.73 0.89 0.57 1.04 0.65 0.31 0.72 0.92 0.72 0.53 0.92 0.54 1.08 1.20 0.37 0.88 1.33 0.56 1.15 0.71 0.39 0.87 0.17 0.68 0.77 0.66 0.47 0.51 0.41 0.38 0.73 0.26 0.82 0.84 0.85 0.40 1.05 0.35 0.69 0.71 1.14 0.79 0.94 0.53 0.34 1.19 1.25 0.26 0.79 0.96 0.43 0.46 0.52 0.84 0.19 0.64 0.96 0.97 0.72 1.62 1.12 1.11 0.77 0.79 0.63 0.49 1.04 1.42 0.79 0.78 0.75 0.97 0.72 1.02 0.45 0.56 1.30 0.50 0.56 1.14 0.94 1.37 1.39 1.07 1.02 0.99 0.76 1.43 0.33 1.32 0.78 1.44 0.51 0.77

Creatine is a chemical made by the body and is used to supply energy mainly to muscle. Serum creatinine is a chemical waste product of creatine. Creatinine is removed from the body entirely by the kidneys. If kidney function is not normal, the creatinine level will increase. The normal range of creatinine is about 0.7 to 1.3 mg/dL.

The descriptive measures of the 300 samples are as follows:

Initially a histogram of the data is plotted and a normal curve is embedded in it and we observe that normal distribution is not suitable to model the data since the data is highly positively skewed.

Table 3.1. Descriptive statistics of Creatinine (mg/dL)

Mean	Median	SD	Skewness	Kurtosis	Min	Max
0.798	0.755	0.337	0.476	-0.221	0.16	1.87

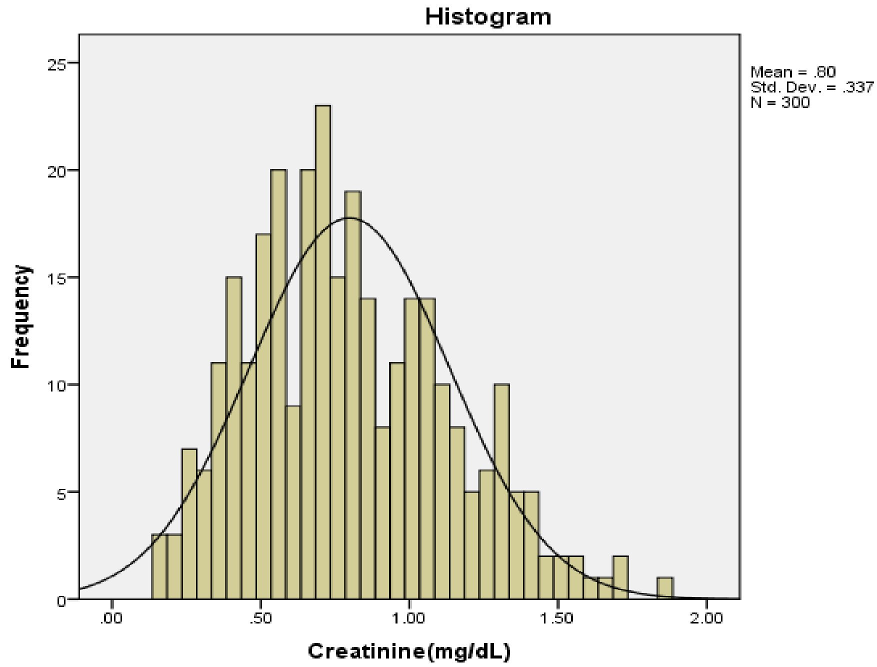
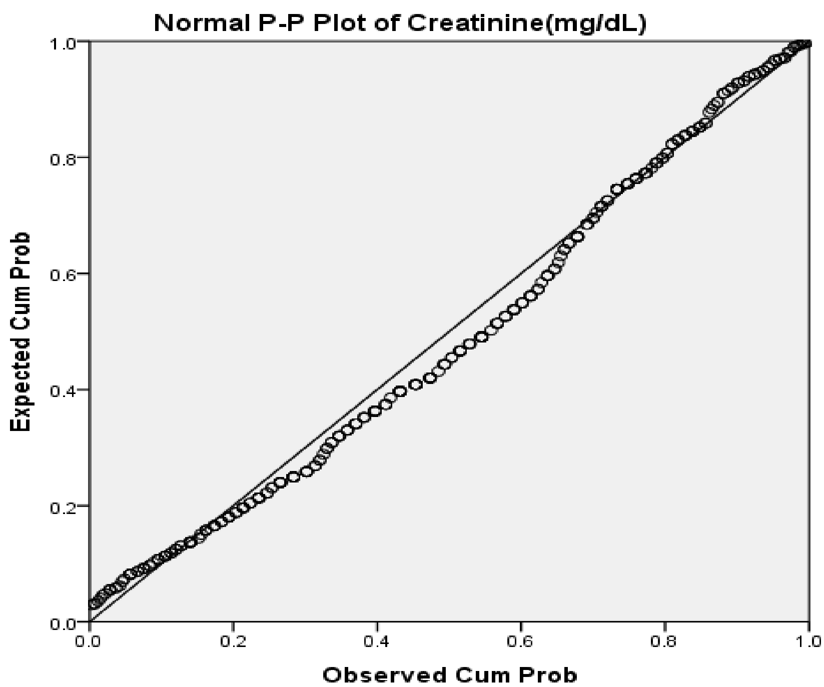


Figure 3.1. Empirical structure of the serum creatinine data

The P-P plot and Q-Q plot of the creatinine data is as follows



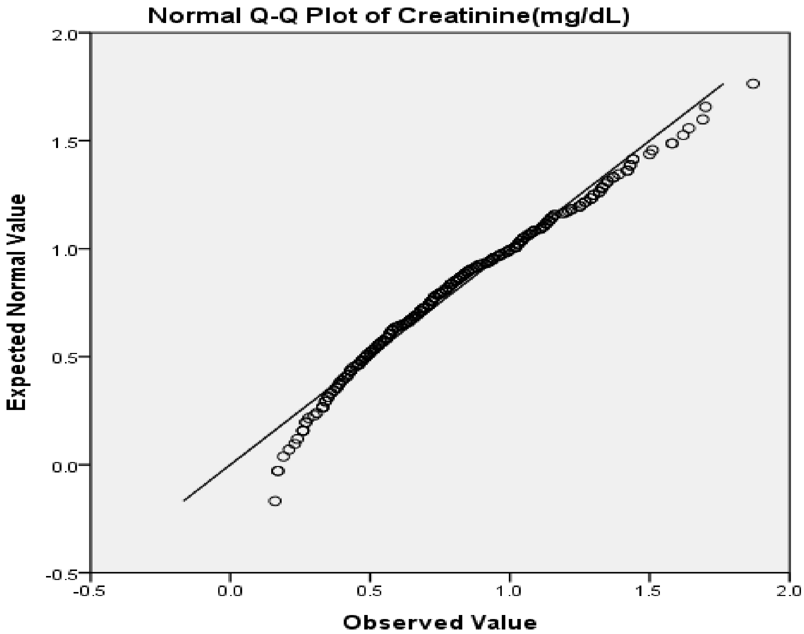


Figure 3.2. P-P plot and Q-Q plot of the serum creatinine data

Since the data shows a positively skewed nature, the symmetrical distributions will not be a suitable choice. So we fitted the data using WTNB with three parameters and compared with the well known distributions like Weibull, Gamma, Exponentiated Weibull, Marshall-Olkin Weibull and ETNB. For comparing the goodness of fit of the model we used the information criteria, Akaike Information Criterion (AIC = -2 log L+2*k*), Bayesian Information Criterion (BIC = -2log L+*k* log *n*) and the Kolmogorov-Smirnov statistic, where *k* is the number of unknown parameters, *log L* is the log-likelihood function value and *n* is the sample size. The results are presented in Table 3.2.

Table 3.2. Parameter estimates and goodness of fit statistics for various models fitted to the serum creatinine data

Distribution	Parameters	AIC	BIC	K-S	P value
Weibull $f(x;\lambda,c)=c\lambda^c x^{c-1}e^{-(\lambda x)^c},\ \lambda,c>0.$	$\hat{\lambda}$ =1.11, $\hat{c}$ =2.55.	183.8	184.7	0.0457	0.558
Gamma $f(x;\beta,\lambda)=\frac{\lambda^\beta}{\Gamma(\beta)}x^{\beta-1}e^{-\lambda x},\ \beta,\lambda>0.$	$\hat{\beta}$ =5.22, $\hat{\lambda}$ =6.54.	184.3	185.3	0.0432	0.6287
Exponentiated Weibull $f(x;\lambda,\theta,c)=c\lambda^c\theta x^{c-1}e^{-(\lambda x)^c}\left(1-e^{-(\lambda x)^c}\right)^{\theta-1},$ $\lambda,\theta,c>0.$	$\hat{\lambda}$ =1.36, $\hat{\theta}$ =1.66, $\hat{c}$ =1.95.	183.1	184.6	0.0393	0.7428
ETNB $f(x;\alpha,\theta,\lambda)=\frac{(1-\alpha)\theta\alpha^\theta\lambda e^{-\lambda x}}{(1-\alpha^\theta)(1-(1-\alpha)e^{-\lambda x})^{\theta+1}},$ $\alpha,\theta,\lambda>0$	$\hat{\alpha}$ =2.53, $\hat{\theta}$ =7.71, $\hat{\lambda}$ =3.73.	191.4	192.8	0.0482	0.4877
Marshall-Olkin Weibull $f(x;\alpha,\lambda,c)=\frac{\alpha c\lambda^c x^{c-1}e^{-(\lambda x)^c}}{[1-(1-\alpha)e^{-(\lambda x)^c}]^2},$ $\alpha,c,\lambda>0$	$\hat{\alpha}$=0.39, $\hat{\lambda}$=0.91, $\hat{c}$=3.06.	183.6	185.1	0.0358	0.8364
WTNB $f(x;\alpha,\theta,c)=\frac{(1-\alpha)\theta\alpha^\theta c\,x^{c-1}e^{-x^c}}{(1-\alpha^\theta)(1-(1-\alpha)e^{-x^c})^{\theta+1}},$ $\alpha,\theta,c>0$	$\hat{\alpha}$ =0.51, $\hat{\theta}$ =0.52, $\hat{c}$ =2.84.	182.9	184.3	0.0338	0.8832

From the above Table 3.2., we can see that the WTNB distribution is the suitable model for the given data. The probability plots for the fitted models are presented in figure 3.3.

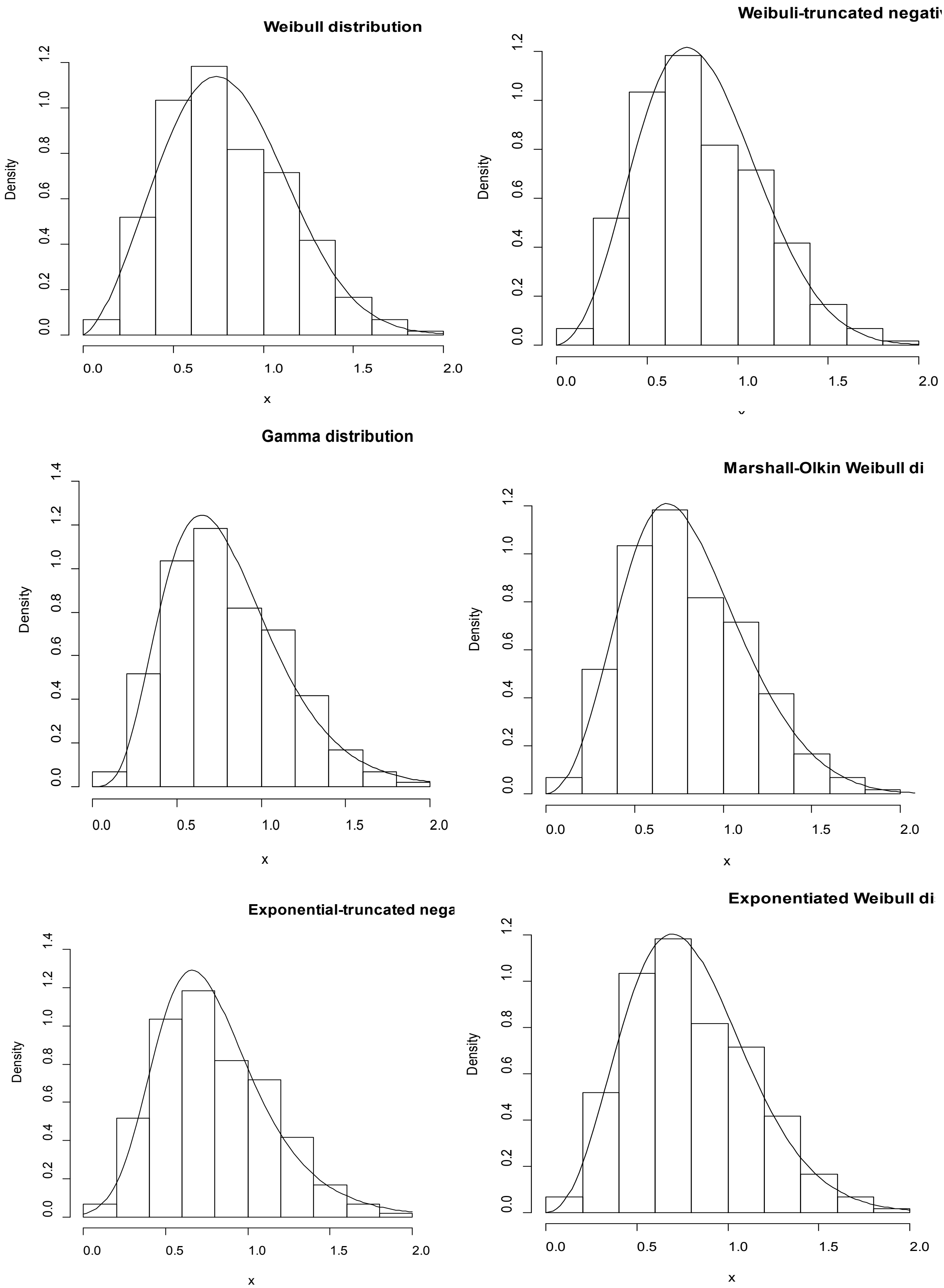


Figure 3.3. Fitted probability density function for the serum creatinine (mg/dL) data

We have used the *nlm* function of **R** software to compare the empirical distribution and the theoretical distribution of all the six distributions mentioned above. From the results out of the six fitted models, WTNB is closer to the empirical distribution of the serum creatinine data. The following figure shows the closeness of the distribution functions.

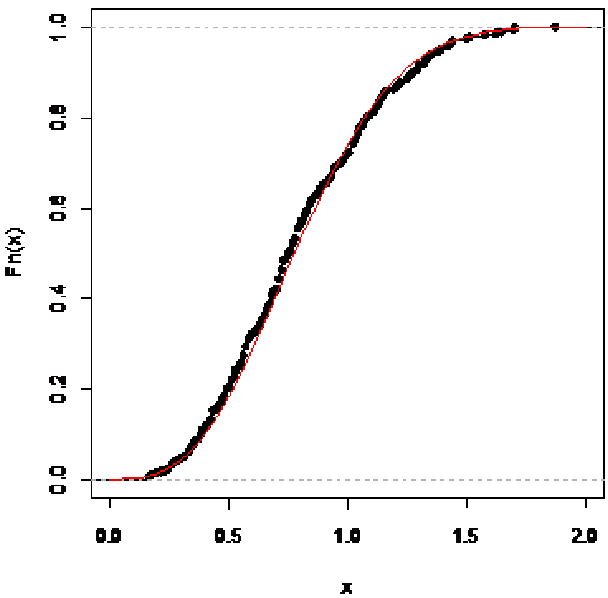


Figure 3.4. WTNB distribution function and empirical distribution

4. Conclusions

In the present paper, we have studied a new family of distributions, namely Weibull truncated negative binomial distributions. Some properties of this distribution such as hazard rate, moments, Shannon and Renyi entropies, distributions of order statistics are studied. This distribution is found to be the most appropriate model to fit the serum creatinine data compared to the Weibull, Gamma, Exponentiated Weibull, ETNB and MOW models. We hope that the new family will attract wider application in life time modeling.

REFERENCES

[1] Alice, T. and K.K. Jose (2003): Marshall-Olkin Pareto process, Far East Journal of Theoretical Statistics, 9, 117-132.

[2] Alice, T. and Jose, K.K. (2005): Marshall-Olkin semi-Weibull minification processes, Recent Advances in Statistical Theory and Applications, 1, 6-17.

[3] Fisher, R.A. and Tippett, L.H.C. (1928): Limiting forms of the frequency distribution of the largest and smallest member of a sample, Proceedings of the Cambridge Philosophical Society, 24, 180-190.

[4] Ghitany, M.E., Al-Hussaini, E.K. and Al-Jaralla, R.A. (2005): Marshall-Olkin extended Weibull distribution and its application to censored data, Journal of Applied Statistics, 32, 1025-1034.

[5] Ghitany, M.E., Al-Awadgi, F.A. and Alkhalfan, I.A. (2007): Marshall-Olkin extended Lomax distribution and its application to censored data, Communication in Statistical Theory and Methods, 36, 1855-1866.

[6] Gumbel, E.J. (1958): Statistics of Extremes, Columbia University Press, New York.

[7] Jayakumar, K. and Girish, B.M. (2015): Some generalizations of Weibull distribution and related processes, Journal of Statistical Theory and Applications, 14, 425-434.

[8] Jayakumar, K. and Thomas, M. (2008): On a generalization to Marshall-Olkin scheme and its application to Burr type XII distribution, Statistical Papers, 49, 421-439.

[9] Marshall, A.W. and Olkin, I. (1997): A new method for adding parameter to a family of distributions with application to exponential and Weibull families, Biometrika, 84, 641-652.

[10] Nadarajah, S., Jayakumar, K. and Ristic, M.M. (2013): A new family of lifetime models, Journal of Statistical Computation and Simulation, 83, 1389-1404.

[11] Prudnikov, A.P., Brychkov, Y.A., and Marichev, O.I. (1986): Integrals and Series Vol.1, Gordon and Breach Sciences, Amsterdam, Netherlands.

[12] Rosin, P. and Rammler, E. (1933): The laws governing the fineness of powdered coal, Journal of the Institute of Fuels, 6, 29-36.

[13] Weibull, W. (1939): A statistical theory of the strength of material, Ingeniors Vetenskaps Akademiens Handligar Report No.151, Stockholm.