OPEN ACCESS

International
Journal
Of Modern Engineering Research (IJMER)

The Exponentiated Marshall-Olkin Discrete Uniform Distribution With Application In Survival Analysis

[1]M-Gharib,[2]B-I- Mohammed,[1]W-E-R-Aghel

[1]Department of Mathematics, Faculty of Science, Ain Shams University, Abbassia, Cairo, Ejypt;
[2]Department of Mathematics, Faculty of Science, Al-Azhar University, Nasr City, Cairo, Ejypt;
Department of Mathematics, Faculty of Science, Ain-Shams University, Abbassia, Cairo, Ejypt, Ejypt.
Correspondence: W-E-R-Aghel,

ABSTRACT: *This paper introduces a new generalization of the Marshall-Olkin discrete uniform distribution introduced by Sandhya and Prasanth (2014). We refer to the new distribution as exponentiated Marshall-Olkin discrete uniform (E-MO-U) distribution. The new model contains the discrete uniform, the exponentiated discrete uniform and the Marshal Olkin discrete uniform distributions as special cases of the proposed model. The properties of the new model are discussed and the maximum likelihood estimation is used to estimate the parameters. While the properties of the mew model aren't in a closed form, then we numerically calculated the mean, the standard deviation, and Shannon's entropy of the given model at different values of the parameters. To examine the performance of our new model in fitting several data we use a real set of data to compare the fitting of the new model with some well-known models, which provides the best fit to all of the data. This model is capable of modeling various shapes of aging and failure criteria.*

Keywords*: Exponentiated; Reliability Function; Moment Generating Function; Inversion Method; Entropy; Maximum Likelihood Estimation.*

Mathematics Subject Classifications (2000) *33C90. 62E99.*

I. INTRODUCTION

Sometimes in real life it is difficult or inconvenient to get samples from a continuous distribution. Almost always the observed values are actually discrete because they are measured to only a finite number of decimal places and cannot really constitute all points in a continuum. Even if the measurements are taken on a continuous scale the observations may be recorded in a way making discrete model more appropriate. In some other situation because of precision of measuring instrument or to save space, the continuous variables are measured by the frequencies of non-overlapping class interval, whose union constitutes the whole range of random variable, and multinomial law is used to model the situation. In categorical data analysis with econometric approach existence of a continuous unobserved or latent variable underlying an observed categorical variable is presumed. Categorical variable is the observed as different discrete values when the unobserved continuous variable crosses a threshold value. Therefore, the inference is based on observed discrete values which are only indicative of the intervals to which unobserved continuous variable belongs but not its true values. Hence this is a case where one makes use of a discretization of the underlying continuous variable.

In survival analysis the survival function may be a function of count random variable that is a discrete version of underlying continuous random variable. For example the length of stay in an observation ward is counted by number of days or survival time of leukemia patients counted by number of weeks. From these examples it is clear that the continuous life time may not necessarily always be measured on a continuous scale but may often be counted as discrete random variables.

Moreover often the continuous failure timed at a generated from a complex system poses more derivational problem than that of a discrete version of the underlying continuous one. Despite these discrete life time distributions played only a marginal role in reliability analysis. Therefore, there is a need to focus on more realistic discrete life time distributions (RezaeiRoknabadi et al.(2009)). That is discretization of a continuous lifetime model is an interesting and intuitively appealing approach to derive a discrete lifetime model corresponding to the continuous one (Lai (2013)).

A continuous random variable may be characterized either by its probability density function (pdf), moment generating function (mgf), moments, hazard rate function etc. Basically construction of a discrete analogue from a continuous distribution is based on the principle of preserving one or more characteristic property of the continuous one. There are various methods by which discrete analogue Y of a continuous random variable X.

For any continuous distribution on $\Re^{+} = [0, \infty)$ with probability density function $g(x)$ (pdf) and a cumulative distribution function $G(x)$ (cdf), one can construct a discrete counterpart supported on the set of integers N_0={0,1,2, … }, whose probability mass function (PMF) is of the form

$$p_y = P(Y = y) = G(y + 1) - G(y), \qquad y = 0,1,2, \dots \dots \tag{1.1}$$

or

$$p_y = P(Y = y) = \bar{G}(y + 1) - \bar{G}(y),$$

where$\bar{G}(y)$ is the survival function (sf) of the random variable X. The resulting PMF will be in a compact form if the continuous (sf) is in compact form.

Numerous distributions are introduced in the literature based on this method as, discrete exponential or the geometric distribution, discrete Weibull, discrete geometric Weibull, the discrete normal, the discrete Maxwell and Discrete Burr distributions introduced by Bracquemond and Gaudoin (2003), Nakagawa and Osaki (1975), Bracquemond and Gaudoin (2003), Nakagawa and Osaki (1975), Krishna and Pundir (2007) and Krishna and Pundir (2009), respectively.

Another method to generating a discrete distribution is Marshall and Olkin generalization introduced by Marshall and Olkin (1997) using any discrete distribution. This generalization is generalized by adding an extra parameter $\theta > 0$ to the base distribution using

$$\bar{G}(y) = \frac{\theta . \bar{Q}(y)}{1 - (1 - \theta)\bar{Q}(y)},$$

$$\bar{F}(y) = \left(\bar{G}(y)\right)^{\gamma} = \left(\frac{\theta . \bar{Q}(y)}{1 - (1 - \theta)\bar{Q}(y)}\right)^{\gamma}.$$

Jayakumar and Mathew (2008) applied the new generalization to the semi-Burr distribution and introduce the exponentiated Marshall-Olkin semi-Burr distribution and derived different properties of the proposed distribution.

Jose and Alice (2005) discussed Marshall-Olkin family of distributions and their applications in time series modeling and reliability. Jose and Krishna (2011) have developed Marshall-Olkin extended uniform distribution. These works and most of the references there in, deal with continuous distribution. Not much work is seen in the discrete case. The reason behind this may be that it is difficult to obtain compact mathematical expressions for moments, reliability, and estimation in the discrete set up. Using discretizing technique defined in (1.1), the probability mass function (PMF) corresponding to the Marshal-Olkin family is given by

$$P(x) = G(x) - G(x - 1) = \frac{\theta q(x)}{[1 - (1 - \theta)\bar{Q}(x)][1 - (1 - \theta)\bar{Q}(x - 1)]}. \tag{1.2}$$

For a discrete uniform distribution with PMF $p(y) = 1/\alpha$; $y = 1,2,3, \dots, \alpha$, Sandhya and Prasanth (2014) introduced the Marshall-Olkin discrete uniform distribution with a survival function given by:

$$\bar{G}(y) = \frac{\theta(\alpha - y)}{\alpha\theta + (1 - \theta)y}; \quad y = 1,2,3, \dots, \alpha, \qquad \alpha, \theta > 0,$$

and it corresponding PMF and CDF are, respectively given by

$$g(y) = \frac{\alpha\theta}{[\alpha\theta + (1-\theta)y][\alpha\theta + (1-\theta)(y-1)]},$$

$$G(y) = 1 - \frac{\theta(\alpha - y)}{\alpha\theta + (1-\theta)y} = \frac{y}{\alpha\theta + (1-\theta)y}. \quad (1.3)$$

In this paper we will introduce the Exponentiated Marshall-Olkin discrete uniform (E-MO-U) distribution as a generalization of the Marshal-Olkin discrete uniform distribution, and illustrate its important features and properties.

The rest of the article is organized as follows. Section **II** introduced the CDF and the PMF of the derived exponentiated Marshall-Olkin uniform distribution. The reliability function of the subject model, hazard rate, cumulative hazard rate, the revered hazard rate and the cumulative reversed hazard rate functions are given in Section **III**. A useful expansion of the CDF and the PMF are given in Section **IV**. Section **V**, discusses the statistical properties including quantile functions, random numbers generation, central and non-central moments, the moment generating function, the incomplete moment, mean deviation and Shannon entropy. Order statistics from the distribution are given in Section **VI**. Section **VII**, gives the stress-strength model from the underlying distribution. In Section **VIII**, we demonstrate two methods of estimation the maximum likelihood estimates and the method of moments to estimate the unknown parameters. Finally, applications of the model using real data set are presented in Section **IX**.

II. EXPONENATED MARSHALL-OLKINDISCRETEUNIFORM (E-MO-U) DISTRIBUTION

By inserting (1.3) into the resilience parameter family of distributions, the CDF of the resulting discrete distribution is the exponentiated Marshall-Olkin discrete uniform distribution and given by:

$$F(y) = \left(\frac{y}{\alpha\theta + (1-\theta)y}\right)^{\gamma}, \qquad y = 0,1,2,\ldots,\alpha, \quad (2.1)$$

in which $\gamma > 0$ is the resilience parameter, and the corresponding PMF is given by

$$f(y) = \left(\frac{y+1}{\alpha\theta + (1-\theta)(y+1)}\right)^{\gamma} - \left(\frac{y}{\alpha\theta + (1-\theta)y}\right)^{\gamma}. \quad (2.2)$$

Figure (2.1) ***a, b, c*** **and** ***d*** bellow, illustrates some of the possible shapes of the PMF of the E-MO-U distribution for different values of the parameters.

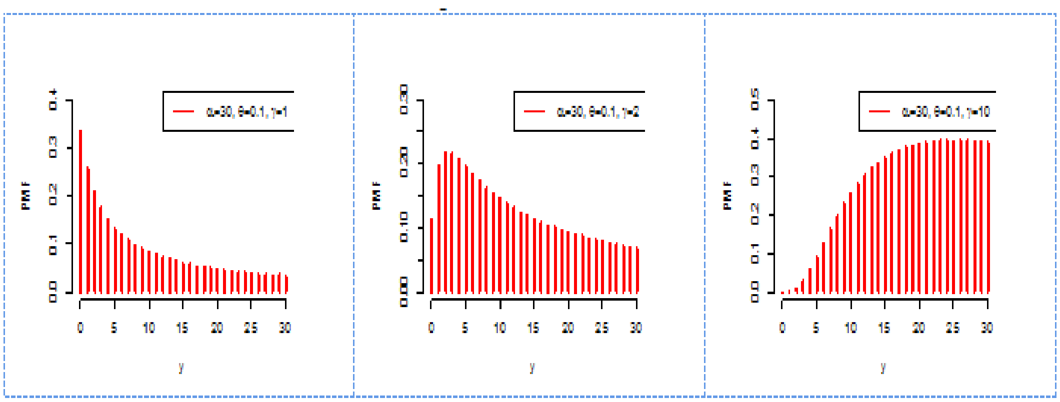

Figure 2.1(a). The PMF of the E-MO-U distribution for different values of γ at $\alpha = 30$ and$\theta = 0.1$.

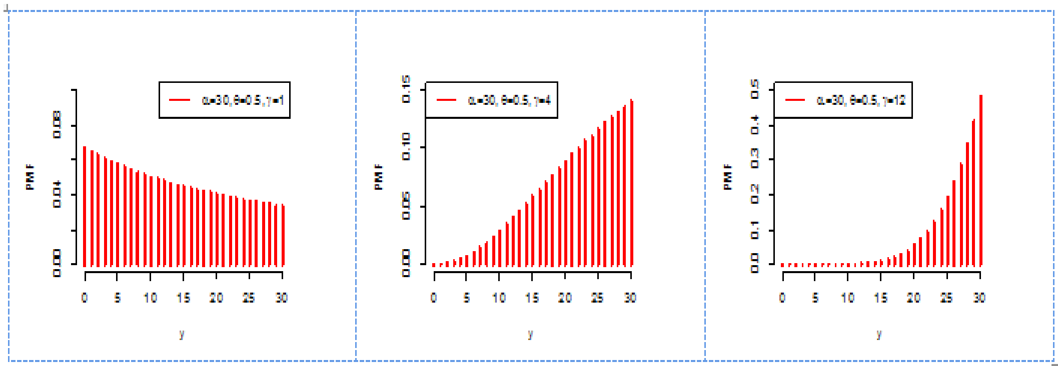

Figure 2.1 (b). The PMF of the E-MO-U distribution for different values of γ at $\alpha = 30$ and $\theta = 0.5$.

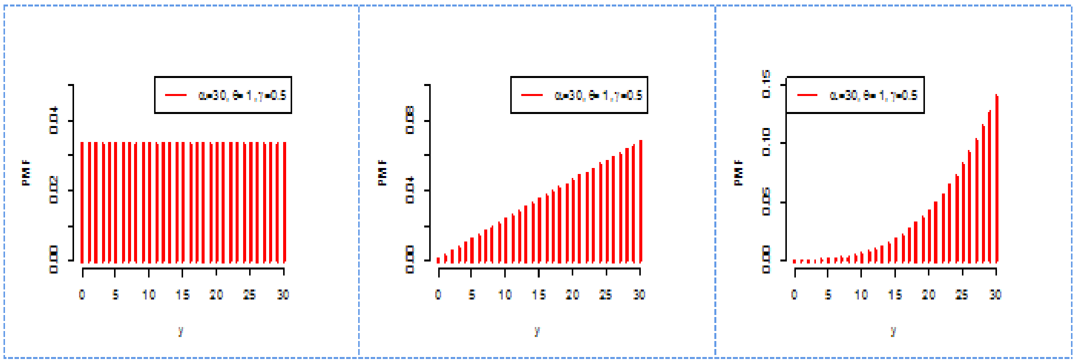

Figure 2.1(c). The PMF of the E-MO-U distribution for different values of γ at $\alpha = 30$ and $\theta = 1$.

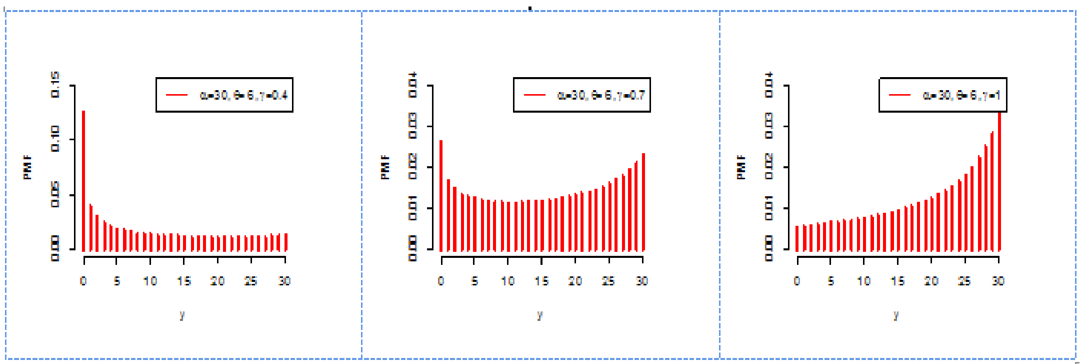

Figure 2.1(d). The PMF of the E-MO-U distribution different values of γ at $\alpha = 30$ and $\theta = 6$.

Figure 2.1 shows different shapes of the PMF while it gives, a monotonic increasing, decreasing, constant and uni-modal shapes, so we can conclude that the introduced distribution is a very flexible distribution in modeling various type of data.

III. RELIABILIY PROPERTIESOF THE E-MO-U DISTRIBUTION

In this section we present the survival, hazard rate, cumulative hazard rate, the reversed hazard rate and the cumulative reversed hazard rate for the Exponentiated Marshall-Olkin uniform distribution.

3.1The Survival Function

The E-MO-U distribution can be a useful characterization of life time data analysis of a given system. The survival function (SF), $\overline{F}(y)$, of the E-MO-U distribution is defined as:

$$\overline{F}(y) = 1 - \left(\frac{y}{\alpha\theta + (1-\theta)y}\right)^{\gamma}, \qquad y = 0,1,2,\dots,\alpha. \tag{3.1}$$

3.2 The Hazard Rate and the Cumulative Hazard Rate Functions

The other characteristic of interest of a random variable is the hazard rate function $h(y) = \frac{p_y}{\overline{F}(y)}$and defined by:

$$h(y) = \frac{\left(\frac{y+1}{\alpha\theta+(1-\theta)(y+1)}\right)^{\gamma} - \left(\frac{y}{\alpha\theta+(1-\theta)y}\right)^{\gamma}}{1 - \left(\frac{y}{\alpha\theta+(1-\theta)y}\right)^{\gamma}}, \tag{3.2}$$

We note that $h(x)$ might be constant, increasing, or decreasing depending or even bathtub on the values of the parameters involved.

For $\gamma = 1$, the hazard rate of the E-MO-U distribution reduced to hazard rate of the Marshall-Olkin uniform distribution. Sandhy and C. B. Prasanth 2014 shows that, for the Marshall-Olkin uniform distribution the distribution is with increasing failure rate (IFR) when $\theta > \frac{\alpha-2x}{2\alpha-2x}$, decreasing failure rate (DFR) when $\theta < \frac{\alpha-2x}{2\alpha-2x}$ and constant failure rate at $\theta = \frac{\alpha-2x}{2\alpha-2x}$. This results are valid for any given value of $\gamma > 0$.

Figure (3.1) a, b and c bellow, illustrates some of the possible shapes of the hazard rate function of the E-MO-U distribution for different values of the parameters.

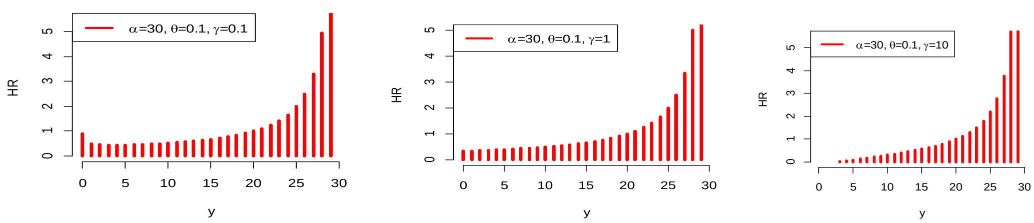

Figure 3.1 (a). The hazard of the E-MO-U distribution different values of γ at $\alpha = 30$ and$\theta = 0.1$.

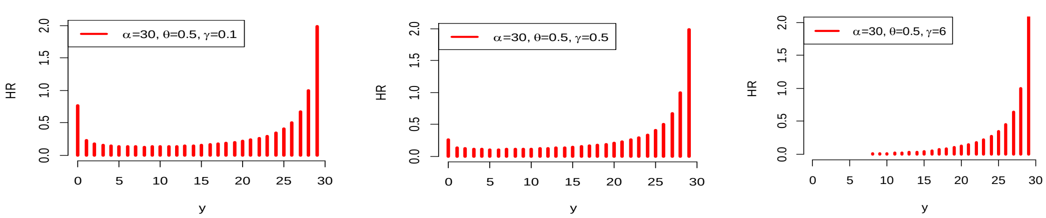

Figure 3.1 (b). The hazard of the E-MO-U distribution different values of γ at $\alpha = 30$ and$\theta = 0.5$.

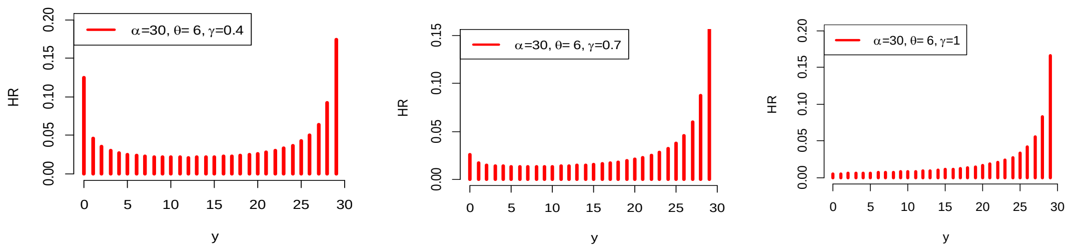

Figure 3.1 (c). The hazard of the E-MO-U distribution different values of γ at $\alpha = 30$ and$\theta = 6$.

Figure (3.1) shows different shapes of the hazard rate function while it gives, a monotonic decreasing, and bathtub shapes, so we can conclude that the introduced distribution is a very flexible distribution in modeling various types of data.
The cumulative hazard rate function, $H(y) = \sum_{i=0}^{y} h(y)$ of the E-MO-U distribution is given by:

$$H(y) = \sum_{i=0}^{y} \frac{\left(\frac{y+1}{\alpha\theta+(1-\theta)(y+1)}\right)^{\gamma} - \left(\frac{y}{\alpha\theta+(1-\theta)y}\right)^{\gamma}}{1 - \left(\frac{y}{\alpha\theta+(1-\theta)y}\right)^{\gamma}}, \tag{3.3}$$

where$H(y)$ is the total number of failures or deaths over an interval of time, which describes how the risk of a particular outcome changes with time for a E-MO-U distribution.

3.3The Reversed Hazard Rate and the Cumulative Reversed Hazard Rate Functions

The reversed hazard rate function$\tilde{h}(y) = \frac{p_y}{F(y)}$ and defined by:

$$\tilde{h}(y) = \frac{\left(\frac{y+1}{\alpha\theta+(1-\theta)(y+1)}\right)^{\gamma}}{\left(\frac{y}{\alpha\theta+(1-\theta)y}\right)^{\gamma}} - 1, \tag{3.4}$$

while the cumulative reversed hazed rate is given by:

$$\widetilde{H}(y) = \sum_{i=0}^{y} \frac{\left(\frac{y+1}{\alpha\theta+(1-\theta)(y+1)}\right)^{\gamma}}{\left(\frac{y}{\alpha\theta+(1-\theta)y}\right)^{\gamma}} - y, \tag{3.5}$$

IV. EXPANSIONFORTHE PMF AND THE CDFOF THE E-MO-U DISTRIBUTION

In this section we introduced another expression to the PMF and the CDF functions. The CDF of the E-MO-U distribution can be written as

$$F(y) = \frac{y^{\gamma}}{(\alpha\theta)^{\gamma}}\left[1 - \left(\frac{\theta-1}{\alpha\theta}\right)y\right]^{-\gamma}. \tag{4.1}$$

Now, let $k > 0$ be real non-integers. If $|z| < 1$, we have the series representations

$$(1-z)^{-k} = \sum_{i=0}^{\infty} \frac{\Gamma(k+i)}{\Gamma(k)i!} z^{i}.$$

From (4.1) and using the previous expansion, one can write the CDF of the E-MO-U distribution as

$$F(y)=\sum_{i=0}^{\infty}\frac{\Gamma(\gamma+i)}{\theta^{\gamma}\Gamma(\gamma)i!}.\left(\frac{\theta-1}{\theta}\right)^{i}\left(\frac{y}{\alpha}\right)^{\gamma+i}\equiv\sum_{i=0}^{\infty}w_{1i}(\theta,\gamma)\left(\frac{y}{\alpha}\right)^{\gamma+i}, \qquad (4.2)$$

where

$$w_{1i}(\theta,\gamma)=\frac{\Gamma(\gamma+i)}{\theta^{\gamma}\Gamma(\gamma)i!}\left(\frac{\theta-1}{\theta}\right)^{i}.$$

It is clear that (4.2) is a linear combination of the CDF of the exponentiated discrete uniform distributions. On the same manner, the PMF of the E-MO-U distribution can be written as

$$P(y)=\sum_{i=0}^{\infty}w_{2i}(\theta,\gamma)\left[(y+1)^{\gamma+i}-(y)^{\gamma+i}\right], \qquad (4.3)$$

where

$$\mathrm{w}_{2\mathrm{i}}(\theta,\gamma)=\frac{\Gamma(\gamma+\mathrm{i})}{(\alpha\theta)^{\gamma}\Gamma(\gamma)\ \mathrm{i}!}\left(\frac{\theta-1}{\theta}\right)^{\mathrm{i}}.$$

V. STATISTICAL PROPERTIES OF THE E-MO-U DISTRIBUTION

This section introduces some of the statistical properties of the E-MO-U distribution including the quantiles, generation of random numbers from the distribution, the central and non-central moments, the moment generating function and the Shannon's Entropy statistics.

5.1 Quintile of the E- MO- U Distribution

The q^{th} quantile is the solution of the Equation $F(y_q)=q$ using Equation (2.1), the q^{th} quantiles from the E-MO-U distribution after some simplifications is given by

$$y_q=\frac{\alpha\theta.\ q^{\frac{1}{\gamma}}}{1-(1-\theta)q^{\frac{1}{\gamma}}}. \qquad (5.1)$$

5.2 Random Numbers Generation from E-MO-U Distribution

Using the method of inversion [See Kelton and Law (2000)], one can generate random numbers from the E-MO-U distribution as

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

where$u$ is an observation from the uniform distribution on the unit interval. This yields

$$y=\frac{\alpha\theta.\ u^{\frac{1}{\gamma}}}{1-(1-\theta)u^{\frac{1}{\gamma}}}. \qquad (5.2)$$

5.3 The Moments of the E- MO- U Distribution

Depends on (4.3), one can get the r^{th} non-central moments, $E(x^r)=\mu^r=\sum_{y=0}^{\infty}y^r.p_y$, or (moments about the origin) of the E-MO-U distribution, as:

$$\mu^r=\sum_{y=0}^{\alpha}y^r.\sum_{i=0}^{\infty}w_{2i}(\theta,\gamma)\left[(y+1)^{\gamma+i}-(y)^{\gamma+i}\right]\quad=\sum_{i=0}^{\infty}w_{2i}(\theta,\gamma)\sum_{y=0}^{z}\left[(y+1)^{\gamma+i}-(y)^{\gamma+i}\right]y^r, \quad (5.3)$$

where

$$w_{2i}(\theta,\gamma) = \frac{\Gamma(\gamma+i)}{(\alpha\theta)^{\gamma}\Gamma(\gamma)i!}\left(\frac{\theta-1}{\theta}\right)^{i}.$$

Also the central moment, m_r, and the moment generating function, $E(e^{tY})$, can be obtained easily from the r^{th} non-central moments, μ^r, throw the relations:

$$m_r = E(Y-\mu)^r = \sum_{n=0}^{r}\binom{r}{n}(-\mu)^{r-n}.E(Y^r),$$

$$E(e^{tX}) = \sum_{r=0}^{\infty}\frac{t^r}{r!}E(X^r).$$

We numerically compute the expectation and standard deviation (Tables 1 and 2) of the E-MO-U distribution at different value of γ and θ at $\alpha = 30{,}100$ since compact expressions are not available for calculating the same.

Table 1: Expectation and standard deviation with different $\theta$and γat $\alpha = 30$.

γ \ θ	0.1		0.5		1		2		4	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
0.1	0.09329	1.21982	0.08985	1.20745	0.08742	1.19685	0.08439	1.18202	0.08088	1.16275
0.5	0.44431	2.67507	0.37619	2.55745	0.33442	2.46133	0.28799	2.33281	0.24031	2.17469
1	0.84259	3.70184	0.63023	3.41991	0.51667	3.19978	0.40309	2.91950	0.29981	2.59536
2	1.54394	5.04343	0.96946	4.42818	0.71759	3.98145	0.49969	3.45783	0.33097	2.91183
4	2.71557	6.72077	1.37957	5.52703	0.92311	4.74780	0.58791	3.9373	0.36857	3.20538

Table 2: Expectation and standard deviation with different $\theta$and γat $\alpha = 100$.

γ \ θ	0.1		0.5		1		2		4	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
0.1	0.09589	2.22649	0.09213	2.20279	0.08954	2.18264	0.08635	2.15455	0.08265	2.11813
0.5	0.44732	4.89643	0.37642	4.66447	0.33352	4.47845	0.28607	4.23174	0.23749	3.92969
1	0.83176	6.79541	0.61869	6.22478	0.505	5.79482	0.39131	5.25436	0.28783	4.6326
2	1.47999	9.29751	0.92744	8.00976	0.68175	7.13292	0.46801	6.12115	0.30158	5.06569
4	2.49049	12.4515	1.26704	9.85070	0.83557	8.32019	0.51501	6.74599	0.30227	5.29479

5.4Incomplete Moments of the E- MO- U Distribution

The incomplete moments of the income distribution form natural building blocks for measuring inequality. For example, the Lorenz and Bonferroni curves depend upon the incomplete moments of the income distribution. The r^{th} incomplete moment,
$\mu^r(z) = E(X^r|x<z) = \sum_{y=0}^{z} y^r.p_y$, of the new class is given by

$$\mu_r^{'}(z) = \sum_{y=0}^{z} y^r.\sum_{i=0}^{\infty} w_{2i}(\theta,\gamma)\left[(y+1)^{\gamma+i}-(y)^{\gamma+i}\right] = \sum_{i=0}^{\infty} w_{2i}(\theta,\gamma)\sum_{y=0}^{z}\left[(y+1)^{\gamma+i}-(y)^{\gamma+i}\right]y^r. \quad (5.4)$$

5.5Mean Absolute Deviations of the E- MO- U Distribution

The dispersion in a population is evidently measured to some extent by the totality of deviations from the mean or median. These are known as the mean deviation about the mean and the mean deviation about the median and are defined by

$$D_1(Y) = \sum_{y=0}^{\infty} |y-\mu| p_y \; and D_2(Y) = \sum_{y=0}^{\infty} |y-M| p_y,$$

respectively, where $\mu = E(X)$ is the mean of the E-MO-U distribution and M is its median.

The measures $D_1(Y)$ and $D_2(Y)$ can be expressed as

$$D_1(Y) = 2\mu.F(\mu) - 2T(\mu) \text{and} D_2(Y) = \mu - 2T(M),$$

where$T(z) = \sum_{y=0}^{z} y.p_y$ is the incomplete mean of Y and can be easily obtained from (5.2) by setting $r = 1$.

5.6 Shannon Entropiesof the E- MO- U Distribution

The entropy measure of a random variable Y with PMF ,P_y , is a measure of variation of the uncertainty. One of the popular entropy measures is the Shannon's Entropy given by

$$I_S = -\sum_{i=1}^{\infty} P_{y_i}.log\left(P_{y_i}\right).$$

For a E-MO-U distribution with pdf (2.2), then the Shannon's entropy is given by

$$I_S = -\sum_{i=1}^{\infty}\left(\frac{y+1}{\alpha\theta+(1-\theta)(y+1)}\right)^{\gamma} - \left(\frac{y}{\alpha\theta+(1-\theta)y}\right)^{\gamma} \times \left\{log\left[\left(\frac{y+1}{\alpha\theta+(1-\theta)(y+1)}\right)^{\gamma} - \left(\frac{y}{\alpha\theta+(1-\theta)y}\right)^{\gamma}\right]\right\}. \quad (5.5)$$

While compact expressions aren't available for calculating the entropy, so we numerically compute it at different values of the parameters (Tables 3 and 4).

Table 3: Entropy with different $\theta$and γat $\alpha = 30$.

γ \ θ	0.1	0.3	0.5	0.8	1	2	4	8
0.1	0.04505	0.04372	0.04290	0.04204	0.0416	0.04012	0.03851	0.0368
0.3	0.09578	0.08918	0.08507	0.08078	0.07859	0.07141	0.06388	0.05639
0.5	0.13031	0.11751	0.10945	0.10112	0.09694	0.08348	0.07003	0.05741
0.8	0.16807	0.14543	0.13117	0.11681	0.10978	0.08806	0.06792	0.05061
1	0.18768	0.15833	0.14004	0.12203	0.11337	0.08735	0.06431	0.04553
2	0.25159	0.19030	0.15639	0.12621	0.11278	0.07636	0.04906	0.03045
4	0.30523	0.20273	0.15524	0.11713	0.10137	0.06247	0.03731	0.02250
8	0.32202	0.19339	0.14010	0.10070	0.08537	0.05015	0.02964	0.01868

Table 4: Entropy with different $\theta$and γat $\alpha = 100$.

γ \ θ	0.1	0.3	0.5	0.8	1	2	4	8
0.1	0.02038	0.01962	0.01918	0.01873	0.01851	0.01777	0.01699	0.01619
0.3	0.04402	0.03999	0.03770	0.03543	0.03431	0.03072	0.02712	0.02364
0.5	0.06035	0.05222	0.04772	0.04336	0.04126	0.03473	0.02853	0.02294
0.8	0.07824	0.06357	0.05584	0.04863	0.04525	0.03526	0.02648	0.01923
1	0.08745	0.06852	0.05888	0.05009	0.04605	0.03442	0.02464	0.01696
2	0.11690	0.08058	0.06433	0.05075	0.04490	0.02949	0.01834	0.01091
4	0.14413	0.08730	0.06498	0.04789	0.04099	0.02949	0.01385	0.00772
8	0.16486	0.08809	0.06169	0.04312	0.03607	0.02023	0.01118	0.00630

VI. ORDER STATISTICSOF THE E-MO-U DISTRIBUTION

Order statistics are among the most fundamental tools in non-parametric statistics and inference. They enter the problems of estimation and hypothesis testing in a variety of ways. Let $F_i(y)$ and $f_i(y)$ be the CDF and PMF of the i-th order statistic of a random sample of size n from E-MO-U distribution.
Since, $F_i(y) = \sum_{k=i}^{n} \binom{n}{k} [F(y)]^k [1 - F(y)]^{n-k}$,

using the binomial expansion for $[1 - F(y)]^{n-k}$, we obtain the following result:

$$F_i(y) = \sum_{k=i}^{n} \sum_{j=0}^{n-k} \binom{n}{k} \binom{n-k}{j} (-1)^j [F(y)]^{k+j}.$$

Then, we can write $F_i(y)$ depends on (4) as

$$F_i(y) = \sum_{k=i}^{n} \sum_{j=0}^{n-k} \binom{n}{k} \binom{n-k}{j} (-1)^j \left(\frac{y}{\alpha\theta + (1-\theta)y}\right)^{\gamma(k+j)} \equiv \sum_{k=i}^{n} \sum_{j=0}^{n-k} \binom{n}{k} \binom{n-k}{j} (-1)^j F_1(y), \quad (6.1)$$

where $F_1(y)$ is the exponentiated Marshall-Olkin Uniform CDF with exponentiated parameter $\gamma(k + j)$.
The corresponding PMF of the i-th order statistic, $f_i(y) = F_i(y) - F_i(y - 1)$ for an integer value of y, then is given by

$$f_i(y) = \sum_{k=i}^{n} \sum_{j=0}^{n-k} \binom{n}{k} \binom{n-k}{j} (-1)^j f_1(y), \quad (6.2)$$

where $f_1(y)$ is the exponentiated Marshall-Olkin Uniform PMF with exponentiated parameter $\gamma(k + j)$.

VII. STRESS-STRENGTH PARAMETEROFTHE E-MO-U DISTRIBUTION

The stress-strength parameter $R = P(X > Y)$ is a measure of component reliability and its estimation problem when X and Y are independent and follow a specified common distribution has been discussed widely in the literature. Suppose that the random variable X is the strength of a component which is subjected to a random stress Y. Estimation of R when X and Y are independent and identically distributed following a well-known distribution has been considered in the literature. Many applications of the stress strength model, for its own nature, are related to engineering or military problems. There are also natural applications in Medicine or Psychology, which involves the comparison of two random variables, representing for example the effect of a specific drug or treatment administered to two groups, control and test.

Almost all of these studies consider continuous distributions for $X$and Y, because many practical applications of the stress-strength model in engineering fields presuppose continuous quantitative data. A complete review is available in Kotz et al. (2003). However, in this regard, a relatively small amount of work is devoted to discrete or categorical data. Data may be discrete by nature. For example, the stress pattern in a step-stress accelerated life test can be treated as a discrete random variable of which the possible values can be obtained from all stress levels, and the corresponding probabilities can be obtained from the acting times of each stress levels. Moreover, the stress state of a component can be categorized based on the characteristic of external loads. For instance, the stress state of a mechanical component can be simply classified as state 1, state 2 and state 3, which correspond to low load, moderate load and heavy load, respectively. More generally, according to the change of external loads, the stress of a component can be categorized into arbitrary finite state: state 1, state 2, ... , state m.
The stress-strength parameter, in discrete case, is defined as

$$R = P(X > Y) = \sum_{X=0}^{\infty} f_X(x).F_Y(x),$$

wheref_X and F_Y denote the PMF and CDF of the independent discrete random variables X and Y, respectively. Now, let $X \sim E - MO - U(\vartheta_1)$and $Y \sim E - MO - U(\vartheta_2)$, where
$\vartheta_1 = (\alpha_1, \theta_1, \gamma_1)^T$and $\vartheta_2 = (\alpha_2, \theta_2, \gamma_2)^T$. Using Equations (4.2) and (4.3), we obtain

$$R = P(X > Y) = \sum_{X=0}^{\infty} \left[\left(\frac{y+1}{\alpha_1\theta_1 + (1-\theta_1)(y+1)}\right)^{\gamma_1} - \left(\frac{y}{\alpha_1\theta_1 + (1-\theta_1)y}\right)^{\gamma_1}\right] \times \left(\frac{y}{\alpha_2\theta_2 + (1-\theta_2)y}\right)^{\gamma_2}. \qquad (7.1)$$

Depending on (4.2) and (4.3), $\boldsymbol{R}$ can be written as

$$R = \sum_{X=0}^{\infty} \sum_{i=0}^{\infty} \sum_{i=0}^{\infty} w_{2i}(\theta, \gamma)\, w_{1i}(\theta_2, \gamma_2) \left(\frac{y}{\alpha_2}\right)^{\gamma_2 + i} [(y+1)^{\gamma+i} - (y)^{\gamma+i}], \quad (7.2)$$

where

$$w_{1i}(\theta_2, \gamma_2) = \frac{\Gamma(\gamma_2 + i)}{\theta_2^{\gamma_2}\Gamma(\gamma_2)i!}\left(\frac{\theta_2 - 1}{\theta_2}\right)^i,$$

and

$$w_{2i}(\theta_1, \gamma_1) = \frac{\Gamma(\gamma_1 + i)}{(\alpha_1\theta_1)^{\gamma_1}\Gamma(\gamma_1)i!}\left(\frac{\theta_1 - 1}{\theta_1}\right)^i.$$

Now, assume that $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_m$ are independent observations from $X \sim E - MO - U(\vartheta_1)$and $Y \sim E - MO - U(\vartheta_2)$, respectively. The total likelihood function is $\ell_R(\vartheta^*) = \ell_n(\vartheta_1).\ell_m(\vartheta_2)$, where $\vartheta^* = (\vartheta_1, \vartheta_2)$. The score vector is given by:

$$U_R(\vartheta^*) = \left(\frac{\partial \ell_R}{\partial \alpha_1}, \frac{\partial \ell_R}{\partial \theta_1}, \frac{\partial \ell_R}{\partial \gamma_1}, \frac{\partial \ell_R}{\partial \alpha_2}, \frac{\partial \ell_R}{\partial \theta_2}, \frac{\partial \ell_R}{\partial \gamma_2}\right),$$

and the MLE of ϑ^*, say $\hat{\vartheta}^*$, may be attained from the nonlinear equation $U_R(\vartheta^*) = 0$. Thus, by inserting the MLEs in equation (7.2) the stress-strength parameter R will be estimated.

VIII. ESTIMATIONOF THE E-MO-U DISTRIBUTION

In this section, the maximum likelihood estimation is used to estimate the unknown parameters. An equation is presented to estimate the parameters using the method of moments.

8.1 The Maximum likelihood Method

To apply the method of maximum likelihood for estimating the parameter vector
$\vartheta = (\alpha, \theta, \gamma)^T$of E-MO-U distribution, assume that $x = (x_1, x_2, \ldots, x_n)^T$ is a random sample of size n from an E-MO-U distribution. The log-likelihood function becomes

$$\mathcal{L} = \sum_{i=1}^{n} \log\left[\left(\frac{y+1}{\alpha\theta + (1-\theta)(y+1)}\right)^{\gamma} - \left(\frac{y}{\alpha\theta + (1-\theta)y}\right)^{\gamma}\right]. \qquad (8.1)$$

Then, the first derivatives of $\mathcal{L}$ with respect to the vector of the parameters are

$$\frac{\partial \mathcal{L}}{\partial \gamma} = \sum_{i=1}^{n} \frac{B_{\alpha,\theta,\gamma}(y_i+1) - B_{\alpha,\theta,\gamma}(y_i)}{A_{\alpha,\theta,\gamma}(x_i)} = 0\,, \qquad (8.2)$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = \sum_{i=1}^{n} \frac{C_{\alpha,\theta,\gamma}(y_i+1) - C_{\alpha,\theta,\gamma}(y_i)}{A_{\alpha,\theta,\gamma}(x_i)} = 0, \qquad (8.3)$$

$$\frac{\partial \mathcal{L}}{\partial \alpha} = \sum_{i=1}^{n} \frac{D_{\alpha,\theta,\gamma}(y_i+1) - D_{\alpha,\theta,\gamma}(y_i)}{A_{\alpha,\theta,\gamma}(x_i)} = 0\,. \tag{8.4}$$

where,

$$A_{\alpha,\theta,\gamma}(y_i) = \left(\frac{y+1}{\alpha\theta + (1-\theta)(y+1)}\right)^{\gamma} - \left(\frac{y}{\alpha\theta + (1-\theta)y}\right)^{\gamma},$$

$$B_{\alpha,\theta,\gamma}(y_i) = \left(\frac{y}{\alpha\theta + (1-\theta)y}\right)^{\gamma} log\left(\frac{y}{\alpha\theta + (1-\theta)y}\right),$$

$$C_{\alpha,\theta,\gamma}(y_i) = \frac{-\gamma(\alpha - y)y^{\gamma}}{(\alpha\theta + (1-\theta)y)^{\gamma+1}}\ , \quad \text{and}$$

$$D_{\alpha,\theta,\gamma}(y_i) = \frac{-\gamma\,\theta\,y^{\gamma}}{(\alpha\theta + (1-\theta)y)^{\gamma+1}}.$$

The solutions of likelihood equations (8.2) to (8.4) provide the maximum likelihood estimators (MLEs) of $\vartheta = (\alpha, \theta, \gamma)^T$, say $\hat{\vartheta} = (\hat{\alpha}, \hat{\theta}, \hat{\gamma})^T$, which can be obtained by a numerical method such as the three variable Newton-Raphson type procedure.
For interval estimation and hypothesis tests on the model parameters, we require the information matrix. The 3×3 observed information matrix is

$$I_n(\hat{\vartheta}) = \begin{bmatrix} -\frac{\partial^2 \mathcal{L}}{\partial \alpha^2} & -\frac{\partial^2 \mathcal{L}}{\partial \alpha \partial \theta} & -\frac{\partial^2 \mathcal{L}}{\partial \alpha \partial \gamma} \\ -\frac{\partial^2 \mathcal{L}}{\partial \theta \partial \alpha} & -\frac{\partial^2 \mathcal{L}}{\partial \theta^2} & -\frac{\partial^2 \mathcal{L}}{\partial \theta \partial \gamma} \\ -\frac{\partial^2 \mathcal{L}}{\partial \gamma \partial \alpha} & -\frac{\partial^2 \mathcal{L}}{\partial \gamma \partial \theta} & -\frac{\partial^2 \mathcal{L}}{\partial \gamma^2} \end{bmatrix}.$$

One can show that the E-MO-U distribution satisfies the regularity conditions which are fulfilled for parameters in the interior of the parameter space but not on the boundary (see, e.g., Cox and Hinkley, 1974). Hence, the MLE vector $\hat{\vartheta}$ is consistent and asymptotically normal. That is, $I_n^{1/2}(\vartheta)(\hat{\vartheta} - \theta)$ converges in distribution to trivariate normal with the (vector) mean zero and the identity covariance matrix.
One can use the normal distribution of $\hat{\vartheta}$ to construct approximate confidence regions for some parameters. Indeed, an asymptotic $100(1-\xi)$ confidence interval for each parameter ϑ_i, is given by

$$\left(\hat{\vartheta}_i - Z_{\xi/2}\sqrt{\hat{J}_{ii}}\,,\ \hat{\vartheta}_i + Z_{\xi/2}\sqrt{\hat{J}_{ii}}\right), \qquad i = 1,2,3,$$

where $\hat{J}_{ii}$ denotes the (i,i) diagonal element of $I_n^{-1}(\hat{\vartheta})$ and $Z_{\xi/2}$ is the $(1-\xi/2)$ -th quantile of the standard normal distribution.

8.2 Method of Moments

To apply the method of moments for estimating the parameters α, θ and γ of E-MO-U distribution, we need to equate the population moments to the corresponding sample moments and subsequently solve the two equations simultaneously. Since the moments of the new distribution cannot be obtained in closed forms, the equations can't be solved via ordinary techniques. So, we resort to a method of pseudo moment by minimizing

$$S(\alpha, \theta, \gamma) = \left(M_1 - E(Y)\right)^2 + \left(M_2 - E(Y^2)\right)^2 + \left(M_3 - E(Y^3)\right)^2,$$

with respect to α, θ and γ, where $M_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$, $M_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ and $M_3 = \frac{1}{n}\sum_{i=1}^{n} x_i^3$ are the first, second and the third sample moments, respectively.

IX. APPLICATIONOF THE E-MO-U DISTRIBUTION

In this section, the E-MO-U model will be examined for two real data sets.

9.1 Application

The first data is given by Chakraborty (2010) on the number of European red mites on apple leaves. This data set consists of 123 observation and are presented in Table 5

Table 5: Number of European Red Mites on Apple Leaves.

Number of European red mites	0	1	2	3	4	5	6	7
Observed frequency	70	38	17	10	9	3	2	1

We compare the fitting of the E-MO-U model with 7 well known discrete models. In each case, the parameters are estimated by maximum likelihood as described in Section 8. We have fitted the E-MO-U distribution to the data, and compared this model with discrete exponentiatedWeibull (EW), discrete transmuted geometric (TG), discrete exponentiated generalized geometric (EGG), discrete modified Weibull type I (MW-I), discrete modified Weibull type II (MW-II), discrete additive Weibull (AW), discrete modified Weibull extension (MWE) and discrete Burr-III (B-III) distributions introduced by Nekoukhou and Bidram (2015), Chakraborty and Bhati (2016), Bidram et al. (2016), reported in Chakraborty (2015), reported in Chakraborty (2015), reported in Chakraborty (2015), reported in Chakraborty (2015) and AL-Huniti and AL-Dayian(2012).

The model selection is carried out using the likelihood (L), Akaike information criterion (AIC), the Bayesian information criterion (BIC) and the Kolmogorov-Smirnov (KS) test.

Table 6 lists the MLEs of the parameters and the values of the AIC and BIC statistics for the tested models. Based on the criterion, we conclude that the E-MO-U distribution provides a superior fit to these data than the other models.

Table 6. MLEs (standard errors in parentheses) and the measures AIC, BIC and KS test to Number of European red mites on apple leaves.

Model	**Estimated Parameters**				***L***	**AIC**	**BIC**	**KS**
E-MO-U	**θ =0.0233 (0.0381)**	**γ =5.8130 (8.9752)**	**$\alpha = 7$**		**-219.435**	**442.8699**	**448.8912**	**0.4567**
EW	γ =0.4545 (0.3470)	α =1.5235 (0.6730)	p =0.8098 (0.2329)		-222.0388	450.0777	459.1096	0.4702
TG	α=-0.0242 (0.2912)	q=0.5303 (0.0539)			-222.438	448.876	454.8972	0.4622
EGG	α=1.6986 (2.1612)	θ=0.5064 (0.0823)	γ=0.7526 (0.4958)		-222.3655	450.7311	459.763	0.4752
MW-I	α=0.2682 (13.3768)	β=1.0176 (0.2252)	q=0.6140 (3.1619)		-222.4288	450.8575	459.8894	0.4613
MW-II	α=0.0785 (0.0770)	β=0.8099 (0.2123)	q= 0.5556 (0.0428)		0.4704145	449.7673	458.7992	0.4704
AW	θ=0.7613 (0.4525)	γ=1.5256 (0.7521)	q_1=0.6031 (0.1882)	q_2=0.8827 (0.2662)	-222.1147	452.2295	464.272	0.4787
MWE	θ=0.0142 (0.007)	β=0.2606 (0.0176)	q=0.9679 (0.0098)		-222.1292	450.2584	450.2584	0.4727
B-III	c=1.7821 (0.1654)	θ=0.3241 (0.0417)			-227.4882	458.9765	464.9977	0.4880

Figure 3 below, shows the empirical CDF and the theoretical versus the empirical PMF and CDF for the underlying distribution which shows a great fitting to the data.

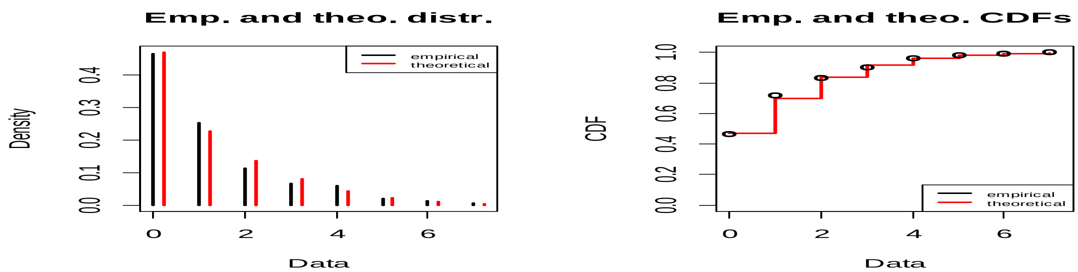

Figure 3.The theoretical versus the empirical PMF and CDF for the E-MO-U distribution.

9.2 Numbers of Ten Shots Fired From a Rifle

The data were given by Nikora and Lyu (1996) on the number of failures of software observed over 62 weeks. The data are presents in table 5 below.

Table 7: Numbers of failures of software observed over 62 weeks.

Failure	0	1	2	3	4	5	6	7	8	9	10	11
Frequency	20	10	11	10	2	3	3	0	0	1	1	1

We compare the fitting of the E-MO-U model with discrete exponentiatedWeibull (EW), discrete transmuted geometric (TG), discrete modified Weibull type I (MW-I), discrete modified Weibull type II (MW-II) and discrete modified Weibull extension (MWE) distributions.

In each case, the parameters are estimated by maximum likelihood. In each case the expected frequency form each distribution has been calculated and perform the chi-square test. Table 8 below summarizes the results

Table 8. MLEs (standard errors in parentheses) and the measures AIC and BIC to numbers of failures of software observed over 62 weeks.

Count	Observed frequency	Expected frequency					
		E-MO-U	EW	TG	MW-I	MW-II	MWE
0	20	20.623	19.556	15.727	18.897	19.443	20.926
1	10	10.757	12.783	17.854	13.684	13.011	11.080
2	11	9.125	9.172	12.353	9.487	9.167	8.542
3	10	8.896	6.538	8.416	6.494	6.467	6.544
> 3	11	12.599	13.951	7.649	13.438	13.912	14.907
sum	62	62	62	62	62	62	62
Parameter estimates		θ =0.1133 (0.1058)	γ = 0.664 (0.5448)	α =-0.152 (0.3811)	α =0.3440 (8.8921)	α =0.0234 (0.0601)	θ=0.0902 (0.4039)
		γ =1.5202 (1.0963)	α = 1.2796 (0.5851)	q =0.6628 (0.0572)	β=1.0482453 (0.3646)	β=0.9443 (0.2597)	β=0.3129 (0.1867)
		α =11 (0.00)	p =0.8234 (0.2346)		q=0.7630 (1.3706)	q=0.6923 (0.0546)	q=0.9455 (0.1251)
L_{max}		-121.0313	-121.7864	-121.8662	-121.8926	-121.8926	-122.0297
AIC		246.0625	249.5729	247.7325	249.7851	249.7851	250.0594
BIC		250.3168	255.9543	251.9868	256.1665	256.1665	256.4408
χ^2		2.6524	3.4375	6.5302	3.6329	3.61816	3.7033
p-value		0.1241	0.0637	0.0382	0.0566	0.0572	0.0543

Based on the criterion in table 8, we conclude that the E-MO-U distribution provides a better fit to these data than the other models based on the χ^2 test. Figure 4 below, shows the empirical CDF and the theoretical versus the empirical PMF and CDF for the underlying distribution which shows a great fitting to the data.

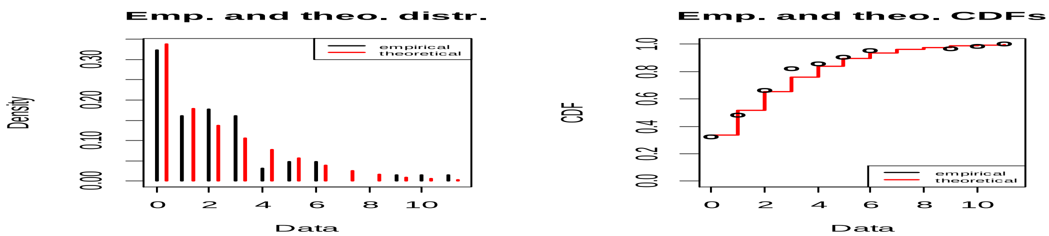

Figure 4. The theoretical versus the empirical PMF and CDF for the U-MO-U distribution.

BIBLIOGRAPHY

[1]. A. P. Nikora and M. R. Lyu (1996), "Software reliability measurement experience," in Handbook of Software Reliability Engineering, M. R. Lyu, Ed., McGraw-Hill, New York, NY, USA, 255-301.

[2]. AL-Huniti, A. A., & AL-Dayian, G. R. (2012). Discrete Burr type III distribution. American Journal of Mathematics and Statistics, 2(5), 145-152.

[3]. Barbiero, A. (2010). A discretizing method for reliability computation in complex stress-strength models.World. Acad. Sci. Eng. Technol, 47, 75-81.

[4]. Bidram, H., Roozegar, R., &Nekoukhou, V. (2016). Exponentiated generalized geometric distribution:A new discrete distribution. HacettepeJournal of Mathematics and Statistics, 1767.

[5]. Bidram, H., Roozegar, R., &Nekoukhou, V. (2016). Exponentiated generalized geometric distribution:A new discrete distribution. HacettepeJournal of Mathematics and Statistics, 1767.

[6]. Chakraborty, S. (2010). On some distributional properties of the family of weighted generalized Poisson distribution. Commun.Statist.Theor.Meth., 39, 2767-2788.

[7]. Chakraborty, S. (2015). Generating discrete analogues of continuous probability distributions-A survey of methods and constructions. Journal of Statistical Distributions and Applications, 2(1), 6.

[8]. Chakraborty, S., &Bhati, D. (2016). Transmuted geometric distribution with applications in modeling and regression analysis of count data. SORT-Statistics and Operations Research Transactions, 1(1), 153-176.

[9]. Cox, D. R., &Hinkley, D. V. (1979). Theoretical statistics. CRC Press.

[10]. Good, I. J. (1953). The population frequencies of species and the estimation of population parameters.Biometrika, 237-264.

[11]. Jayakumar, K., & Mathew, T. (2008). On a generalization to Marshall–Olkin scheme and its application to Burr type XII distribution. Statistical Papers, 49(3), 421-439.

[12]. Jose, K. K., & Alice, T. (2005). Marshall-Olkin Family of Distributions-Applications in Time Series Modeling and Reliability ,JC publications.

[13]. Jose, K. K., & Krishna, E. (2011). Marshall-Olkin extended uniform distribution. In ProbStat Forum (Vol. 4, No. October, pp. 78-88).

[14]. Kelton, W., and Law, A. (2000). Simulation Modeling and Analysis. MA: McGraw Hill, Boston.

[15]. Kotz, S., &Pensky, M. (2003). The stress-strength model and its generalizations: theory and applications. World Scientific.

[16]. Krishna, H., &Pundir, P. S. (2007). Discrete maxwell distribution. InterStat, 3.

[17]. Krishna, H., &Pundir, P. S. (2009). Discrete Burr and discrete Pareto distributions. Statistical Methodology, 6(2), 177-188.

[18]. Lai, C. D., & Wang, D. Q. (1995). A finite range discrete life distribution. International Journal of Reliability, Quality and Safety Engineering, 2(02), 147-160.

[19]. Lai, C. D., Xie, M., & Murthy, D. N. P. (2003). A modified Weibull distribution. IEEE Transactions on reliability, 52(1), 33-37.

[20]. Lisman, J. H. C., & Van Zuylen, M. C. A. (1972). Note on the generation of most probable frequency distributions. StatisticaNeerlandica, 26(1), 19-23.

[21]. Marshall, A. W., &Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibullfamilies.Biometrika, 84(3), 641-652.

[22]. Nakagawa, T., & Osaki, S. (1975). The discrete Weibull distribution. IEEE Transactions on Reliability, 24(5), 300-301.

[23]. Nekoukhou, V., &Bidram, H. (2015). The exponentiated discrete Weibull distribution. SORT, 39, 127-146.

[24]. Roknabadi, A. H., Borzadaran, G. R., &Khorashadizadeh, M. (2009). Some aspects of discrete hazard rate function in telescopic families. Economic Quality Control, 24(1), 35-42.

[25]. Sandhya, E., &Prasanth, C. B. (2014). Marshall-Olkin Discrete Uniform Distribution. Journal of probability,10.

[26]. Sato, H., Ikota, M., Sugimoto, A., & Masuda, H. (1999). A new defect distribution metrology with a consistent discrete exponential formula and its applications. IEEE Transactions on Semiconductor Manufacturing, 12(4), 409-418.

[27]. Yiokari, N., Xekalaki, E., &Karlis, D. (2001). On some discrete valued time series models based on mixtures and thinning. In Proceedings of the 5th Hellenic-European conference on computer mathematics and its applications, Athens, Greece (pp. 872-877).