THE MARSHALL – OLKIN TRANSMUTED - G FAMILY OF DISTRIBUTIONS: PROPERTIES AND APPLICATIONS

Roshini George

Department of Statistics, Loyola College (Autonomous), Chennai

S. Thobias

Department of Statistics, Loyola College (Autonomous), Chennai

Abstract: We introduce a new class of continuous distributions called the Marshall-Olkin Transmuted-G family which extends the Transmuted class by Shaw and Buckley (2007). We provide a comprehensive account of the mathematical properties of the new distribution. Various structural properties including ordinary and incomplete moments, generating function, Renyi and Mathai-Haubold entropy, order statistics are derived. Some special sub models of the new family are introduced involving Exponential, Kumaraswamy, Frechet and Weibull distributions. Different methods for estimating the model parameters of the family are discussed. A Monte Carlo simulation study is conducted to examine the bias, mean square error of the maximum likelihood estimators and width of the confidence interval for each parameter. We prove empirically the flexibility of the proposed model by means of two applications to real data sets to the two mentioned sub models.

Keywords: Incomplete Moments, Marshall-Olkin Transmuted-G Family, Monte Carlo Simulation Study, Renyi and Mathai-Haubold Enropy.

Introduction: In the last few years there has been a great interest in transmuted distributions and several of them have been investigated. The interest in developing more flexible statistical distributions remains strong nowadays. Shaw and Buckley (2007) pioneered an existing distribution that would offer more distributional flexibility. They used the Quadratic Rank Transmutation Map (QRTM) in order to generate a flexible family of distributions. The generated family is called the transmuted extended distribution, includes the parent distribution as a special case and gives more flexibility to model various types of data.

The transmutation map approach has been suggested by Shaw and Buckley is given as follows: To define a rank transmutation mapping in complete generality, consider two distributions with common sample space with cdf's F_1 and F_2.Then we can form $G_{R12}(u) = F_2(F_1^{-1}(u))$, $G_{R21}(u) = F_1(F_2^{-1}(u))$. This pair of maps takes unit interval I = [0,1] into itself and under suitable assumptions are mutual inverse and satisfy $G_{Rij}(0) = 0$ and $G_{Rij}(1) = 1$. It is also assumed that these rank transmutation maps are continuously differentiable. Under quadratic transmutation, the cdf of skewed distribution can be obtained by using the relation of cdf of base distribution as $F_2(x) = (1+\lambda)F_1(x) - \lambda(F_1(x))^2$ where $F_1(x)$ is the cdf of base distribution and $|\lambda| \leq 1$, known as shape parameter. According to QRTM approach the cdf satisfy the relationship $G(x) = (1+\lambda)F(x) - \lambda\big(F(x)\big)^2$, $|\lambda| \leq 1$, where F(x) is the cdf of the base distribution. It is observed that at $\lambda = 0$, the above cdf becomes the distribution function of the base random variable.

A random variable X has the Transmuted-G (TG) family if the probability density function(pdf) and cumulative distribution function(cdf) are defined through the QRTM method by

$$g(x) = h(x;\phi)[1+\lambda-2\lambda H(x;\phi)] \tag{1}$$

$$G(x) = (1+\lambda)H(x;\phi) - \lambda H(x;\phi)^2 \tag{2}$$

where $H(x;\phi)$ is the parent cdf and $h(x;\phi)$ is the parent pdf. Both functions depend on the parameter vector ϕ. Hence the random variable X following (1) with parameter λ and baseline vector of parameters ϕ is denoted by X~ $TG(\lambda,\phi)$.

The Marshall-Olkin -G family proposed by Marshall and Olkin (1997) by adding one parameter to the survival function $\bar{G}(x) = 1 - G(x)$, where G(x) is the baseline cdf. Then the Marshall-Olkin (MO) family has cdf defined by

$$F(x) = \frac{1-\bar{G}(x)}{1-\bar{\theta}\,\bar{G}(x)}\ ,\ -\infty < x < \infty,\ 0 < \theta < \infty \tag{3}$$

and the corresponding pdf is given by

$$f(x) = \frac{\theta g(x)}{\left[1-\bar{\theta}\,\bar{G}(x)\right]^2} \tag{4}$$

Then the hazard rate function is

$\tau(x) = \frac{h(x)}{1-\bar{\theta}\,\bar{G}(x)}$, where $h(x)$ is the hazard rate function(hrf) corresponding to $\bar{G}(x)$. It is clear that (3) provides a method to obtain a new distribution from an existing one which is more flexible.

The Marshall-Olkin Transmuted – G (MOTG) Family: Based on the TG family by Shaw and Buckley (2007) we construct a new distribution called the Marshall-Olkin Transmuted-G (MOTG) family of distribution and provide a comprehensive description of some of its mathematical properties. The cdf of the MOTG is

$$F(x) = \frac{(1+\lambda)H(x;\phi)-\lambda H(x;\phi)^2}{\theta+\bar{\theta}\ [(1+\lambda)H(x;\phi)-\lambda H(x;\phi)^2]} \tag{5}$$

The corresponding pdf is

$$f(x) = \frac{\theta h(x;\phi)[1+\lambda-2\lambda H(x;\phi)]}{\{\theta+\bar{\theta}\ [(1+\lambda)H(x;\phi)-\lambda H(x;\phi)^2]\}^2} \tag{6}$$

Hence the random variable X following (/6) is denoted by X~MOTG (θ, λ,ϕ).

The hazard rate function is given by

$$\tau(x) = \frac{h(x;\phi)[1+\lambda-2\lambda H(x;\phi)]}{\{1-(1+\lambda)H(x;\phi)-\lambda H(x;\phi)^2\}\{\theta+\bar{\theta}\ [(1+\lambda)H(x;\phi)-\lambda H(x;\phi)^2]\}}$$

The quantile function of X says Q(u) = $F^{-1}(u)$ is given by

$$Q(u) = \frac{1+\lambda-\sqrt{(1+\lambda)^2-4\lambda A}}{2\lambda}\ , \lambda \neq 0$$

where A= $\frac{u\theta}{1-u\bar{\theta}}$, u is a uniform variate on the unit interval [0 1].

Expansion for the pdf and the cdf: By using the binomial expansion, we can show that the pdf (/6) has the expansion,

$$f(x) = g(x)\sum_j B_j \bar{G}(x)^j$$

$$= \sum_j B_j' \frac{d}{dx}\bar{G}\ (x)^{j+1}$$

where $B_j = \theta(1-\theta)^j(j+1)$

and $B_j' = \theta(1-\theta)^j$

Also, we can write

$$f(x) = g(x)\sum_j \sum_k u_{jk}\, G(x)^k$$

Where $u_{jk} = \frac{\theta(1-\theta)^j\Gamma(j+2)(-1)^k}{k!\Gamma(\text{j}-\text{k}+1)}$

Finally, the pdf can be expressed as a mixture of exp-G densities

$$f(x) = \sum_j \sum_k w_{jk}\, \pi_{k+1}(x) \tag{7}$$

where $w_{jk} = \frac{u_{jk}}{k+1}$

and $\pi_k(x) = kg(x)G(x)^{k-1}$ is the exp-G pdf with power parameter k > 0.

Also, we obtain the same mixture representation for the cdf.

$$F(x) = \sum_j \sum_k w_{jk}\, \Pi_{k+1}(x)$$

where $\Pi_{k+1}(x)$ is the cdf of exp-G.

Moments: Moments are necessary and important in any statistical analysis, especially in applications. They can be used to study the most important features and characteristics of a distribution such as tendency, dispersion, skewness and kurtosis. Let Y_{k+1} denotes the exp-G distribution with power parameter k+1. Then the r^{th} moment of X say μ_r' using (/7) as

$$\mu_r' = \sum_j \sum_k w_{jk} E(Y_{k+1}^r)$$

The $(r,k)^{th}$ probability weighted moment of X following MOTG distribution is defined by $\Gamma(r,k) = \int_{-\infty}^{\infty} x^r f(x) F(x)^k$. Then we have,

$$E(X^r) = \int_{-\infty}^{\infty} x^r f(x) dx$$
$$= \int_{-\infty}^{\infty} x^r h(x)[1+\lambda - 2\lambda H(x)] \sum_j \sum_k u_{jk} \left[(1+\lambda)H(x) - \lambda H(x)^2\right]^k dx$$
$$= \sum_j \sum_k u_{jk} \Gamma(r,k)$$

Especially, $E(X) = \sum_j \sum_k u_{jk} \Gamma(1,k)$.

Moment Generating Function:
We have $M_X(t) = E(e^{tX})$
$$= \sum_j \sum_k w_{jk} M_{k+1}(t),$$
where $M_{k+1}(t)$ is the mgf of Y_{k+1}.
Also, $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$
$$= \sum_j \sum_k u_{jk}' \int_{-\infty}^{\infty} e^{tx} \frac{d}{dx} G(x)^{k+1}$$
$$= \sum_j \sum_k u_{jk}' M_S(t)$$
where S follows exponentiated TG distribution.

Incomplete Moments: The s^{th}incomplete moment,
$$\phi_s(t) = \int_{-\infty}^{t} x^s f(x) dx$$
$$= \sum_j \sum_k w_{jk} \int_{-\infty}^{t} x^s \pi_{k+1}(x) dx$$
The mean deviation about mean and median are expressed by
$$\delta_1 = \int_0^{\infty} |X - \mu_1'| f(x) dx$$
$$= 2\mu_1 F(\mu_1') - 2\phi_1(\mu_1')$$
and $$\delta_2 = \int_0^{\infty} |X - M| f(x) dx$$
$$= \mu_1' - 2\phi_1(M)$$
where $\mu_1' = E(X)$, M = Median (X) = $\phi(1/2)$.
The first incomplete moment $\phi_1(t)$ can be obtained by setting s=1 in (8). $\phi_1(t)$ can be applied to construct Bonferroni and Lorenz curves defined for a given probability π by B(π) = $\phi_1(q)/(\pi\mu_1')$ and L(π) = $\phi_1(q)/\mu_1'$ where $q = Q(\pi)$ is the quantile function of X at π. Another application of the first incomplete moment is related to mean residual life and mean waiting time given by
$$m_1(t) = [1 - \phi_1(t)]/R(t) - t$$
and $$M_1(t) = t - [\phi_1(t)/F(t)]$$

Renyi and Mathai-Haubold Entropy: The entropy is interpreted as the expected uncertainty contained in f(x) about the predictability of an outcome of the random variable X. In physical situations when an appropriate density is selected, one procedure is the maximization of entropy.
The Renyi entropy of X is defined by
$$I_R(\delta) = (1-\delta)^{-1} log \int_{-\infty}^{\infty} f^{\delta}(x) dx, \ \ \delta > 0, \delta \neq 1.$$
We have,
$$f^{\delta}(x) = \alpha^{\delta} h(x)^{\delta} [1+\lambda - 2\lambda H(x)]^{\delta} \{1 - (1-\alpha)\left[1 - H(x)[1+\lambda - \lambda H(x)]\right]\}^{-2\delta}$$
$$= \sum_j \sum_k m_{jk} g(x)^{\delta} G(x)^k$$
where $m_{jk} = \alpha^{\delta}(-1)^k (1-\alpha)^j \Gamma(2\delta + j)/\Gamma(2\delta)\, k!\,(j-k)!$
Hence,
$$I_R(\delta) = (1-\delta)^{-1} log \sum_j \sum_k m_{jk} \int_{-\infty}^{\infty} g(x)^{\delta} G(x)^k dx$$
Note that the integral term depends only on the pdf and cdf of the base line distribution.
Shannon defined the basic measure of uncertainty associated with the random variable X is given by $H(x) = -\int_{-\infty}^{\infty} f(x) log f(x) dx$. The Renyi entropy is an extension of Shannon entropy.
Mathai – Haubold introduced the generalized information measure
$$M_a(X) = (\delta - 1)^{-1}\left[\int_{-\infty}^{\infty} f^{2-\delta}(x) dx - 1\right], \ \delta \neq 1, \ 0 < \delta < 2$$

$= (\delta - 1)^{-1}$
$\left[\sum_j \sum_k m_{jk} \int_{-\infty}^{\infty} g^{2-\delta}(x) G(x)^k dx - 1\right]$
where $m_{jk} = \alpha^{2-\delta}(-1)^k(1-\alpha)^j \Gamma(2(2-\delta)+j)/\Gamma(2\delta)\, k!\,(j-k)!$
As $\delta \to 1$, the above equation reduces to Shannon's information entropy.

Order Statistics: The order statistics and their moments have great importance in many statistical problems and they have many applications in reliability analysis and life testing. The order statistics arise in the study of reliability of a system. The order statistics can represent the lifetimes of units or components of a reliability system. Let $X_1, X_2, \dots, X_n$ be a random sample from MOTG family. Then the pdf of $X_{i:n}$ can be written as

$$f_{i:n}(x) = \frac{f(x)}{B(i,n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F^{j+i-1}(x)$$

where B(.) is the beta function.
We have

$$F^{j+i-1}(x) = \left\{\frac{(1+\lambda)H(x) - \lambda H(x)^2}{1-\bar{\alpha}\left[1-[(1+\lambda)H(x)-\lambda H(x)^2]\right]}\right\}^{j+i-1}$$

$$= \sum_l \sum_m (-1)^m (1-\alpha)^l \frac{\Gamma(j+i+l-1)\Gamma(l+1)}{\Gamma(j+i-1)\Gamma(l-m+1)l!m!} G(x)^{m+j+i-1}$$

Then

$$f(x)F^{j+i-1}(x) = \sum_k g(x)\, G(x)^{k+m+j+i-1} t_{jk}$$

where $t_{jk} =$

$$\sum_l \sum_m \frac{\alpha(1-\alpha)^{j+l}(-1)^{k+m}}{k!l!m!(k+m+j+i)} \frac{\Gamma(j+2)\Gamma(j+i+l-1)\Gamma(l+1)}{\Gamma(j-k+1)\Gamma(j+i-1)\Gamma(l-m+1)}$$

Therefore

$$f_{i:n}(x) = \frac{\sum_{k=0}^{\infty} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j}}{B(i,n-i+1)} t_{jk}\, \pi_{k+m+j+i}(x)$$

Thus, the density function of the MOTG order statistics is a mixture of exponential densities.
The moments of $X_{i:n}$ can be expressed as

$$E(X_{i:n}) = \frac{\sum_{k=0}^{\infty} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j}}{B(i,n-i+1)} t_{jk}\, E(Y_{k+m+j+i}).$$

The L-moments are expectations of certain linear combinations of order statistics. They exist whenever the mean of the distribution exists even though some higher moments may not exist, and are relatively robust to the effects of outliers. Based upon the moments of $X_{i:n}$ we can derive an explicit expression for the L-moments of the r^{th} order of the random variable X as infinite weighted linear combinations of the means of suitable MOTG order statistics and is defined by

$$\lambda_r = \frac{1}{r} \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} E\left(X_{r-j:r}\right);\ r \geq 1$$

Mixtures: Let $\bar{F}(x/\lambda), -\infty < x < \infty, -\infty < \lambda < \infty$ be the conditional survival function of a continuous random variable Λ. Let Λ follows a distribution with probability density function $\mathrm{m}(\lambda)$. A distribution with survival function

$$\bar{F}(x) = \int_{-\infty}^{\infty} \bar{F}(x/\lambda) m(\lambda) d\lambda,\ -\infty < x < \infty$$

is called a mixture distribution with mixing density $\mathrm{m}(\lambda)$. The following result shows that MOTG distribution can be expressed as a mixture.

Theorem 1: Let X be a continuous random variable with conditional survival function

$$\bar{F}(x/\lambda) = e^{-\left\{\left[1-H(x)[1+\lambda-\lambda H(x)]\right]^{-1}-1\right\}\lambda}$$

with probability density function $\mathrm{m}(\lambda) = \alpha e^{-\alpha\lambda}, \lambda > 0$. Then the random variable X has the MOTG distribution.
Proof: Under these assumptions, we obtain that the survival function of the random variable X is

$$\bar{F}(x) = \alpha \int_0^{\infty} e^{-\left\{\left[1-H(x)[1+\lambda-\lambda H(x)]\right]^{-1}-\bar{\alpha}\right\}\lambda} d\lambda$$

$$= \frac{\alpha[1-H(x)[1+\lambda-\lambda H(x)]]}{1-\bar{\alpha}[1-H(x)[1+\lambda-\lambda H(x)]]}$$

Thus, the random variable X has MOTG distribution.

Theorem 2: Let $\{X_i, i \geq 1\}$ be a sequence of independent and identically distributed random variables with TG(λ,ϕ) distribution. Let N be a geometric random variable which is independent of X_i with parameter θ such that P (N = n) = $\theta(1-\theta)^{n-1}$; $0 < \theta < 1, n = 0,1, \ldots$. Then U = Min $(X_1, X_2, \ldots, X_N)$ and V = Max $(X_1, X_2, \ldots, X_N)$ has MOTG family of distributions with parameter $(\theta$, λ, $\phi)$ and $(\frac{1}{\theta}$, λ, $\phi)$.

Proof: The survival function of U = Min $(X_1, X_2, \ldots, X_N)$ is given by

$$\bar{F}_U(x) = \sum_{n=1}^{\infty} P(X_1 \geq x, X_2 \geq x, \ldots, X_n \geq x) \; P(N = n)$$

$$= \sum_{n=1}^{\infty} [\bar{F}(x)]^n \, \theta(1-\theta)^{n-1}$$

$$= \frac{\theta \bar{F}(x)}{1-(1-\theta)\bar{F}(x)}$$

$$= \frac{\theta\{1-[(1+\lambda)H(x;\phi)-\lambda H(x;\phi)^2]\}}{1-(1-\theta)\{1-[(1+\lambda)H(x;\phi)-\lambda H(x;\phi)^2]\}}$$

which is the survival function of MOTG $(\theta$, λ, $\phi)$.

Consider the distribution function of V = Max $(X_1, X_2, \ldots, X_N)$ is given by

$$F_V(x) = \sum_{n=1}^{\infty} P(X_1 \leq x, X_2 \leq x, \ldots, X_n \leq x) \; P(N = n)$$

$$= \sum_{n=1}^{\infty} [F(x)]^n \, \theta(1-\theta)^{n-1}$$

$$= \frac{\theta F(x)}{1-(1-\theta)F(x)}$$

Then the survival function is

$$\bar{F}_V(x) = \frac{\bar{F}(x)}{\bar{F}(x)+\theta F(x)}$$

$$= \frac{\frac{1}{\theta}\bar{F}(x)}{1-\left(1-\frac{1}{\theta}\right)\bar{F}(x)}$$

$$= \frac{\frac{1}{\theta}\{1-[(1+\lambda)H(x;\phi)-\lambda H(x;\phi)^2]\}}{1-(1-\frac{1}{\theta})\{1-[(1+\lambda)H(x;\phi)-\lambda H(x;\phi)^2]\}}$$

which is the survival function of MOTG $(\frac{1}{\theta}$, λ, $\phi)$.

Special Models: In this section we provide four special models of MOTG family. The pdf (6) will be most tractable when h(x) and H(x) have simple analytic expressions. These special models extend several widely known distributions in the literature.

The Marshall-Olkin Transmuted Exponential (MOT-E) Distribution: The exponential distribution has pdf and cdf given by $h(x) = pe^{-px}$ $(x > 0)$ and $H(x) = 1 - e^{-px}$, then the pdf and cdf of MOTE distribution is given by

$$f(x) = \frac{\theta p e^{-px}[1+\lambda-2\lambda(1-e^{-px})]}{\{\theta+\bar{\theta}\;[(1+\lambda)(1-e^{-px})-\lambda(1-e^{-px})^2]\}^2}$$

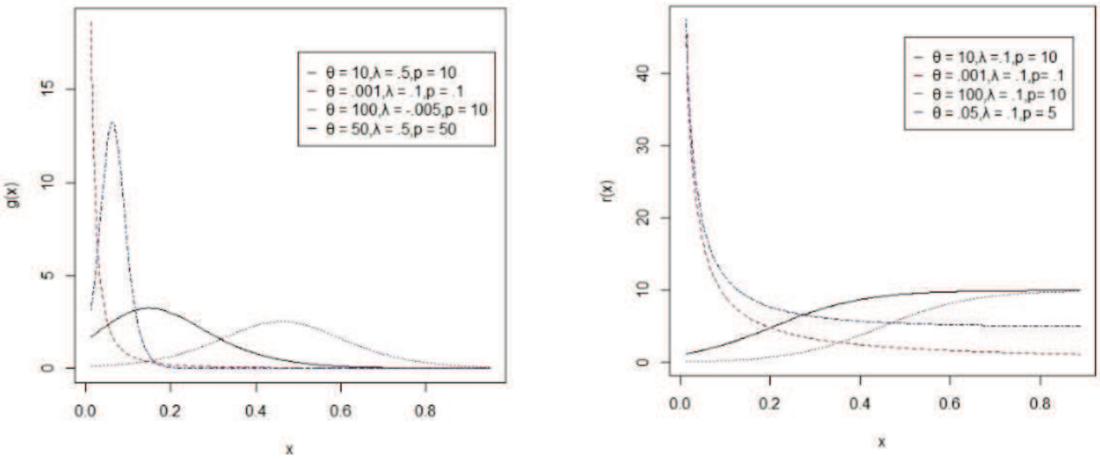
and $F(x) = \frac{(1+\lambda)(1-e^{-px})-\lambda(1-e^{-px})^2}{\theta+\bar{\theta}\;[(1+\lambda)(1-e^{-px})-\lambda(1-e^{-px})^2]}$

The hazard rate function (hrf) is

$$r(x) = \frac{pe^{-px}[1+\lambda-2\lambda(1-e^{-px})]}{\{1-[(1+\lambda)(1-e^{-px})-\lambda(1-e^{-px})^2]\}}$$

$$\frac{1}{\{\theta+\bar{\theta}\;[(1+\lambda)(1-e^{-px})-\lambda(1-e^{-px})^2]\}}$$

The MOTE distribution reduces to the Transmuted exponential(TE) distribution when $\theta = 1$. Also when $\lambda = 0$, it reduces to Marshall-Olkin Exponential distribution. Fig. 1 displays some possible shapes of the density and hazard rate function of the distribution. It reveals that the pdf of MOTE distribution can be reversed J shape, right skewed or unimodal. The hazard rate function can be decreasing or increasing.



(a) The MOT-E density plots (b) The MOT-E hrf plots

Fig. 1

The Marshall-Olkin Transmuted Kumaraswamy(MOT-Kw) Distribution: The Kumaraswamy (Kw) distribution has pdf and cdf for $x \in [0\ 1]$ and a, b > 0 given by
$h(x) = abx^{a-1}(1-x^a)^{b-1}$ and $H(x) = 1-(1-x^a)^b$ respectively. Then the MOTKw density function is

$$f(x) = \frac{\theta a b x^{a-1}(1-x^a)^{b-1}[1+\lambda-2\lambda(1-(1-x^a)^b)]}{\{\theta+\bar{\theta}\ [(1+\lambda)(1-(1-x^a)^b)-\lambda(1-(1-x^a)^b)^2]\}^2}$$

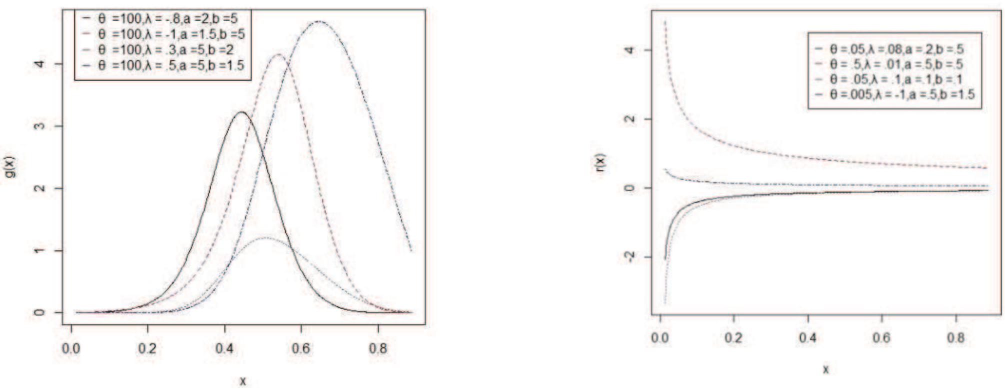
The corresponding distribution function is

$$F(x) = \frac{[1-(1-x^a)^b][1+\lambda-\lambda(1-(1-x^a)^b)]}{\theta+\bar{\theta}\ [1-(1-x^a)^b][1+\lambda-\lambda(1-(1-x^a)^b)]}$$

The q^{th} quantile is given by

$$x_q = \left\{1-\left[1-\frac{(1+\lambda)-\sqrt{(1+\lambda)^2-4\lambda A}}{2\lambda}\right]^{1/b}\right\}^{1/a}$$

This distribution reduces to Transmuted Kumaraswamy(TKw) when $\theta = 1$. For $\lambda = 0$, we obtain the Marshall-Olkin Kumaraswamy(MOKw) distribution. Also, when $\theta = 1$ and $\lambda = 0$ it follows a Kumaraswamy(Kw) distribution. Fig. 2 displays the plots of density and hazard rate function for the MOTKw distribution.



(a) The MOT-Kw density plots (b) The MOT-Kw hrf plots

Fig. 2

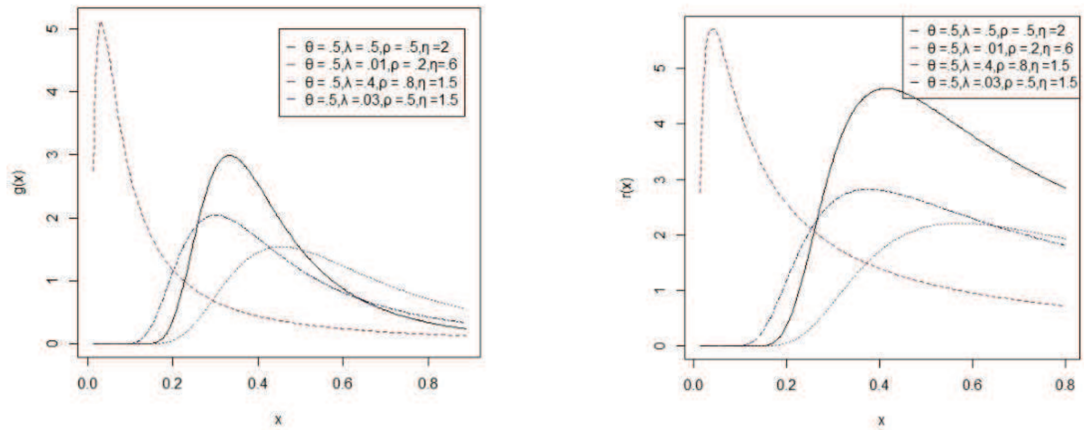
The Marshall-Olkin Transmuted Frechet(MOT-Fr) Distribution: The pdf and cdf of Frechet distribution (for x > 0) is given by $h(x) = \frac{\eta}{\rho}\left(\frac{\rho}{x}\right)^{\eta+1} e^{-\left(\frac{\rho}{x}\right)^{\eta}}$ and $H(x) = e^{-\left(\frac{\rho}{x}\right)^{\eta}}$ respectively. The Frechet distribution is one of the important distributions in extreme value theory. The pdf of the MOTFr distribution is given by

$$f(x) = \frac{\theta\eta\rho^{\eta}x^{-(\eta+1)}e^{-\left(\frac{\rho}{x}\right)^{\eta}}\left[1+\lambda-2\lambda e^{-\left(\frac{\rho}{x}\right)^{\eta}}\right]}{\left\{\theta+(1-\theta)e^{-\left(\frac{\rho}{x}\right)^{\eta}}\left[1+\lambda-\lambda e^{-\left(\frac{\rho}{x}\right)^{\eta}}\right]\right\}^{2}}$$

and the corresponding cdf is

$$F(x) = \frac{e^{-\left(\frac{\rho}{x}\right)^{\eta}}\left[1+\lambda-\lambda e^{-\left(\frac{\rho}{x}\right)^{\eta}}\right]}{\theta+(1-\theta)e^{-\left(\frac{\rho}{x}\right)^{\eta}}\left[1+\lambda-\lambda e^{-\left(\frac{\rho}{x}\right)^{\eta}}\right]}$$

The MOTFr distribution reduces to Transmuted Frechet(TFr) distribution when $\theta = 1$. For $\lambda = 0$, it reduces to Marshall-Olkin Frechet distribution. Also, when $\theta = 1$ and $\lambda = 0$ we get the Frechet distribution. The plots in Fig. 3 show some possible shapes of the density and hazard rate functions of the MOTFr distribution. The Fig. shows that the pdf of MOTFr model is very flexible. It can be left skewed or right skewed.



(a) Plots of the MOT-Fr pdf. (b) Plots of the MOT-Fr hrf.

Fig. 3

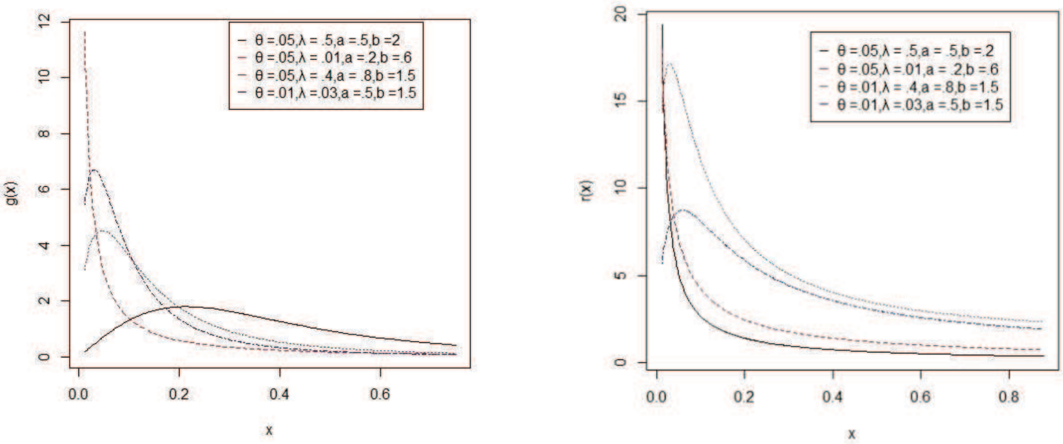
The Marshall-Olkin Transmuted Weibull(MOT-W) Distribution: The Weibull distribution (for x > 0) has pdf and cdf given by $h(x) = ba^{\eta}x^{b-1}e^{-(ax)^{b}}$ and $H(x) = 1 - e^{-(ax)^{b}}$ respectively. Then the density function of the MOTW distribution is given by

$$f(x) = \frac{\theta ba^{b}x^{b-1}e^{-(ax)^{b}}\left[1+\lambda-2\lambda\left(1-e^{-(ax)^{b}}\right)\right]}{\left\{\theta+\bar{\theta}\left(1-e^{-(ax)^{b}}\right)\left[1+\lambda-\lambda\left(1-e^{-(ax)^{b}}\right)\right]\right\}^{2}}$$

The corresponding cdf is

$$F(x) = \frac{\left(1-e^{-(ax)^{b}}\right)\left[1+\lambda-\lambda\left(1-e^{-(ax)^{b}}\right)\right]}{\theta+\bar{\theta}\left(1-e^{-(ax)^{b}}\right)\left[1+\lambda-\lambda\left(1-e^{-(ax)^{b}}\right)\right]}$$

The MOTW distribution reduces to Transmuted Weibull(TW) when $\theta = 1$. Also if $\lambda = 0$, we get the Marshall-Olkin Weibull(MOW) distribution. The density and hazard rate plots of MOTW distribution is given in Fig. 4.



(a) The MOT-W density plots. (b) The MOT-W hrf plots.

Fig. 4

Methods of Estimation: In this section we describe the six estimation methods for estimating the parameters θ, λ and ϕ of the MOTG family. For all the methods we consider the case when the parameters are unknown.

Method of Maximum Likelihood (MLE): Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from the MOTG family of distribution. Then the loglikelihood function is given by

$$l = n\log\theta + \sum_{i=1}^{n} \log h(x_i;\phi) + \sum_{i=1}^{n}[1+\lambda-2\lambda H(x_i;\phi)] - 2\sum_{i=1}^{n} \log\{\theta+\bar{\theta}\,[(1+\lambda)H(x_i;\phi) - \lambda H(x_i;\phi)^2]\}$$

Maximizing the loglikelihood, we have the following system of non-linear equations:

$$\frac{\partial l}{\partial\theta} = \frac{n}{\theta} - 2\sum_{i=1}^{n}\frac{\{1-[(1+\lambda)H(x_i;\phi)-\lambda H(x_i;\phi)^2]\}}{\{\theta+\bar{\theta}\,[(1+\lambda)H(x_i;\phi)-\lambda H(x_i;\phi)^2]\}}$$

$$\frac{\partial l}{\partial\lambda} = \sum_{i=1}^{n}\frac{1-2H(x_i;\phi)}{[1+\lambda-2\lambda H(x_i;\phi)]} - 2\sum_{i=1}^{n}\frac{\bar{\theta}[H(x_i;\phi)-H(x_i;\phi)^2]}{\{\theta+\bar{\theta}\,[(1+\lambda)H(x_i;\phi)-\lambda H(x_i;\phi)^2]\}}$$

$$\frac{\partial l}{\partial\phi} = \sum_{i=1}^{n}\frac{h'(x_i;\phi)}{h(x_i;\phi)} - 2\lambda\sum_{i=1}^{n}\frac{H'(x_i;\phi)}{1+\lambda-2\lambda H(x_i;\phi)} - 2\sum_{i=1}^{n}\frac{\bar{\theta}[(1+\lambda)H'(x_i;\phi)-2\lambda H(x_i;\phi)H'(x_i;\phi)]}{\{\theta+\bar{\theta}\,[(1+\lambda)H(x_i;\phi)-\lambda H(x_i;\phi)^2]\}}$$

This system of non-linear equations can be solved numerically by any software to obtained the estimates $\hat{\theta}_{MLE}$, $\hat{\lambda}_{MLE}$ and $\hat{\phi}_{MLE}$.

Method of Maximum Product of Spacings (MPS): This method is based on an idea that the differences (spacings) of the consecutive points should be identically distributed. Define the uniform spacings of a random sample from the MOTG distribution as:

$$D_i = F(x_i) - F(x_{i-1}); \quad i = 1,2,\ldots,n$$

where $F(x_i) = 0$ and $F(x_{n+1}) = 1$. Clearly $\sum_{i=1}^{n+1} D_i = 1$. The maximum product of spacings estimates are obtained by maximizing the geometric mean of the spacings:

$$G = [\prod_{i=1}^{n+1} D_i]^{\frac{1}{n+1}}$$

Then we have, $\log G = \frac{1}{n+1}\sum_{i=1}^{n+1}\log D_i$

The non-linear equations are given by,

$$\frac{\partial \log G}{\partial\theta} = \frac{1}{n+1}\sum_{i=1}^{n+1}[F_1'(x_i) - F_1'(x_{i-1})] = 0$$

$$\frac{\partial \log G}{\partial\lambda} = \frac{1}{n+1}\sum_{i=1}^{n+1}[F_2'(x_i) - F_2'(x_{i-1})] = 0$$

$\frac{\partial \log G}{\partial \phi} = \frac{1}{n+1} \sum_{i=1}^{n+1} [F_3'(x_i) - F_3'(x_{i-1})] = 0$
where
$F_1'(x_i) = \frac{-[(1+\lambda)H(x_i) - \lambda H(x_i)^2]\{1-[(1+\lambda)H(x_i) - \lambda H(x_i)^2]\}}{\{\{\theta + \bar{\theta}\,[(1+\lambda)H(x_i) - \lambda H(x_i)^2]\}\}^2}$

$F_2'(x_i) = \frac{\theta[H(x_i) - H(x_i)^2]}{\{\{\theta + \bar{\theta}\,[(1+\lambda)H(x_i) - \lambda H(x_i)^2]\}\}^2}$

$F_3'(x_i) = \frac{\theta[(1+\lambda)H'(x_i) - 2\lambda H(x_i)H'(x_i)]}{\{\{\theta + \bar{\theta}\,[(1+\lambda)H(x_i) - \lambda H(x_i)^2]\}\}^2}$
This system of non-linear equations can be solved by any software to obtained the estimates.

Method of Least Square (LSE): Let $x_1, x_2, \ldots, x_n$ be the observed values from the MOTG distribution. By considering the associated order statistics $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ and it is known that $[F(X_{i:n})] = \frac{i}{n+1}$.
The least square estimates are obtained by minimizing the function:
$S(\theta, \lambda, \phi) = \sum_{i=1}^{n} \left[F(x_i) - \frac{i}{n+1}\right]^2$
Then the estimates are obtained by solving the following non-linear equations:
$\sum_{i=0}^{n} \left[F(x_i) - \frac{i}{n+1}\right] F_1'(x_i) = 0,$
$\sum_{i=0}^{n} \left[F(x_i) - \frac{i}{n+1}\right] F_2'(x_i) = 0,$
$\sum_{i=0}^{n} \left[F(x_i) - \frac{i}{n+1}\right] F_3'(x_i) = 0.$

Method of Weighted Least Square (WLS): The weighted least square estimates are obtained by minimizing the function:
$W(\theta, \lambda, \phi) = \sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F(x_i) - \frac{i}{n+1}\right]^2$
The estimates are obtained by solving the non-linear equations:
$\sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F(x_i) - \frac{i}{n+1}\right] F_1'(x_i) = 0,$
$\sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F(x_i) - \frac{i}{n+1}\right] F_2'(x_i) = 0,$
$\sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F(x_i) - \frac{i}{n+1}\right] F_3'(x_i) = 0.$

Method of Cramer- Von- Mises (CVM): Cramer Von Mises is a type of minimum distance estimators. The likelihood function is given by
$C(\theta, \lambda, \phi) = \frac{1}{12n} + \sum_{i=1}^{n} \left[F(x_i) - \frac{2i-1}{2n}\right]^2$
The estimates can be obtained by solving these non-linear equations:
$\sum_{i=1}^{n} \left[F(x_i) - \frac{2i-1}{2n}\right] F_1'(x_i) = 0,$

$\sum_{i=1}^{n} \left[F(x_i) - \frac{2i-1}{2n}\right] F_2'(x_i) = 0,$
$\sum_{i=1}^{n} \left[F(x_i) - \frac{2i-1}{2n}\right] F_3'(x_i) = 0.$

Method of Anderson-Darling (ADE): The Anderson Darling estimates are obtained by minimizing the function:
$A(\theta, \lambda, \phi) = -n - \frac{1}{n} \sum_{i=1}^{n} \{(2i-1) log F(x_i) \quad + \quad log \bar{F}(x_i)(2n+1-2i)\}$
The estimates are obtained by solving these non-linear equations:
$\sum_{i=1}^{n} (2i-1) \frac{F_1'(x_i)}{F(x_i)} -$

$$\sum_{i=1}^{n}(2n+1-2i)\frac{F_1'(x_i)}{\bar{F}\,(x_i)}=0,$$

$$\sum_{i=1}^{n}(2i-1)\frac{F_2'(x_i)}{F(x_i)}-\sum_{i=1}^{n}(2n+1-2i)\frac{F_2'(x_i)}{\bar{F}\,(x_i)}=0,$$

$$\sum_{i=1}^{n}(2i-1)\frac{F_3'(x_i)}{F(x_i)}-\sum_{i=1}^{n}(2n+1-2i)\frac{F_3'(x_i)}{\bar{F}\,(x_i)}=0.$$

Monte Carlo Simulation Study: In this section, a simulation study is conducted to assess the performance of MOTKw distribution. We study the performance by conducting various simulation for different sample sizes. It is repeated for N = 5000 times each with sample size n = 25,50,75,100,200,500 and parameter values $\theta = 0.05, \lambda = 0.3, a = 0.5, b = 0.1$. The average bias, root mean square error and the average width are computed.

a) Average bias of the MLE $\hat{\vartheta}$ of the parameter $\vartheta = \theta, \lambda, a, b$:

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\vartheta}-\vartheta)$$

b) Root mean squared error (RMSE):

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\vartheta}-\vartheta)^2}$$

c) Average Width(AW) of 95% confidence intervals of the parameter ϑ.

The numerical values of the parameters for different sample sizes are listed below in Table I.

Table I: Average Bias, RMSE and AW θ = 0.05, λ = 0.= 0.3, a=0.5 and b = 0.1

Parameter	n	Average Bias	RMSE	AW
θ	25	0.32289	0.68898	3.15262
	50	0.22277	0.41176	1.79424
	75	0.16923	0.29305	1.25495
	100	0.15022	0.22752	1.01051
	200	0.12286	0.15269	0.65896
	500	0.10654	0.11708	0.39156
λ	25	-0.27153	0.29366	5.75235
	50	-0.25532	0.29344	4.79329
	75	-0.25291	0.29235	4.31187
	100	-0.25656	0.29169	3.91449
	200	-0.27311	0.29137	3.05036
	500	-0.29197	0.29059	2.02911
a	25	-0.02998	0.09983	0.48824
	50	-0.02348	0.07834	0.34293
	75	-0.01836	0.06611	0.28233
	100	-0.01642	0.05861	0.24380
	200	-0.01579	0.04340	0.17033
	500	-0.01511	0.03043	0.10667
b	25	4.5818	5.60629	13.8461
	50	4.09197	4.57659	9.32185
	75	3.82779	4.09917	7.51958
	100	3.75488	3.96208	6.49111
	200	3.67529	3.76936	4.52624
	500	3.62198	3.65697	2.81105

From the results, we can verify that as the sample size n increases, the RMSEs decay toward zero. We also observe that for all the parametric values, the biases decrease as the sample size n increases. Also, the table shows that the average confidence widths decrease as the sample size increases. Thus, our simulation study results show that the maximum likelihood estimation procedure works very well.

Application: In this section, we provide two applications to real data to illustrate the importance of the MOTE and MOTFr models. The first data represents the life of fatigue fracture of Kevlar 373/epoxy subjected to constant pressure at 90% stress level until all had failed. The data was extracted from Abdul-Moniem and Seham 2015[10].

Data set I: 0.0251,0.0886,0.0891,0.2501,0.3113,0.3451,
0.4763, 0.5650,0.5671,0.6566,0.6748,0.6751,
0.6753, 0.7696,0.8375,0.8391,0.8425,0.8645,
0.8851, 0.9113,0.9120,0.9836,1.0483,1.0596,
1.0773,1.1733,1.2570,1.2766,1.2985,1.3211,
1.3503, 1.3551,1.4595, 1.4880,1.5728,1.5733,
1.7083, 1.7263,1.7460,1.7630, 1.7746,1.8275,
1.8375,1.8503,1.8808, 1.8878,1.8881,1.9316,
1.9558, 2.0048,2.0408,2.0903,2.1093, 2.1330,
2.2100, 2.2460,2.2878,2.3203,2.3470,2.3513,
2.4951, 2.5260,2.9911,3.0256,3.2678,3.4045,
3.4846, 3.7433,3.7455,3.9143, 4.8073,5.4005,
5.4435, 5.5295, 6.5541, 9.0960.

Table II: Descriptive Statistics of Data I

Min	Q_1	Q_2	Mean	Q_3	Max	Variance	Skewness	Kurtosis
0.0251	0.9048	1.7362	1.9592	2.2959	9.096	2.4774	1.9797	8.1608

Table III: Goodness of Fit Statistics for Data I

Distribution	Estimates	AIC	CAIC	BIC	HQIC
MOTE (θ, λ, p)	3.644 -0.631 1.842	-174.0	-173.7	-167.0	-171.2
TE (λ, p)	-0.849 0.727	247.0	247.2	251.7	248.9
BE (p, a, b)	0.485 1.679 1.509	250.5	250.8	253.2	257.4
KwTE (λ, p, a, b)	-1.012 4.781 0.456 1.559	-142.2	-141.6	-132.8	-138.4

We compare the MOTE distribution with Transmuted Exponential(TE), Beta Exponential(BE) and Kumaraswamy Transmuted Exponential(KwTE) distributions. The Akaike information criterion(AIC), Bayesian information criterion (BIC), consistent Akaike information criterion(CAIC) and Hannan-Quinn information criterion(HQIC) are computed. The numerical values are listed in the Table. Based on these values it is clear that MOTE distribution is the best fit for the first data compared to the other models that were considered.

The second data consists of 100 observations from Nichols and Padgett (2006) on breaking stress of carbon fibres in (Gba) [14].
Data set II: 3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11, 4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09,1.87, 3.15, 4.90, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.53, 2.67, 2.93, 3.22,
3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31,2.85,
2.56, 3.56, 3.15, 2.35, 2.55, 2.59,2.38, 2.81, 2.77, 2.17, 2.83, 1.92, 1.41, 3.68, 2.97, 1.36, 0.98, 2.76,4.91, 3.68, 1.84, 1.59,3.19, 1.57, 0.81, 5.56, 1.73, 1.59, 2.00, 1.22, 1.12, 1.71,
2.17, 1.17, 5.08, 2.48, 1.18,3.51, 2.17,1.69,
1.25, 4.38, 1.84, 0.39, 3.68, 2.48,0.85, 1.61, 2.79, 4.70, 2.03, 1.80, 1.57, 1.08, 2.03, 1.61, 2.12, 1.89, 2.88, 2.82, 2.05, 3.65.

Table IV: Descriptive Statistics of Data II

Min	Q_1	Q_2	Mean	Q_3	Max	Variance	Skewness	Kurtosis
0.390	1.840	2.700	2.61	3.220	5.560	1.0279	0.3682	3.1049

Table V: Goodness of Fit Statistics for Data II

Distribution	Estimates	AIC	CAIC	BIC	HQIC
MOTFr$(\theta, \lambda, \rho, \eta)$	10.666 2.109 0.990 1.734	-384.1	-383.5	-373.7	-380.3
TFr (λ, ρ, η)	0.253 1.159 0.705	-272.6	-272.8	-264.8	-269.8
Fr (ρ, η)	0.872 1.381	16.9	17.0	22.1	18.7

Here we use the data to compare the MOTFr model with Transmuted Frechet(TFr) and Frechet(Fr) models. It is shown that the MOTFr model has the lowest values for all goodness-of-fit statistics among all fitted models. So, the MOTFr model can be chosen as the best model.

It is shown, by means of two real data sets, that special cases of the MOTG family can give better fits than other models generated by well-known families. From the discussions given above, it is clear that Marshall -Olkin transmuted family of distributions provide a flexible class of distributions having applications in various areas like distribution theory, reliability theory etc.

Acknowledgements: The authors would like to thank the editor and the anonymous referees for valuable comments which greatly improved the paper.

References:

1. A. Marshall and I. A. Olkin. “A new method for adding parameter to a family of distributions with applications to the exponential and weibull families.” Biometrika, vol.84 (1997): 641-652.
2. A. M. Mathai and J. Haubold. “Pathway models, Tsallis statistics, superstatistics and a generalized measure of entropy.” Physics A, vol.375 (2006): 110-122.
3. A. Z. Afify, G. G. Hamedani, I. Ghosh and M. E. Mead. “The transmuted Marshall-Olkin Frechet distribution: properties and applications”. International journal of statistics and probability, vol.4 (2015): 132-148.
4. A. Z. Afify, G. M. Cordeiro, H. M. Yousof, Z. M. Nofal and A. Alzaatreh. “The Kumaraswamy transmuted G family of distributions: properties and applications.” Journal of data science, vol.14 (2016):245-270.
5. B. C. Arnold, N. Balakrishnan and H. N. Nagaraja. “A first course in order statistics.” New York Wiley (1992).
6. C. E. Shannon. “A mathematical theory of communication.” Bell system technical journal, vol.27 (1948): 379-423, 623-656.
7. E. Krishna, K. K. Jose, T. Alice and M. M. Ristic. “The Marshall-Olkin Fréchet distribution.” Communications in statistics, Theory and methods, vol.42, no.22 (2013): 4091-4107.
8. G. R. Aryal and C. P. Tsokos. “Transmuted weibull distribution: a generalization of the weibull probability distribution.” European journal of pure and applied mathematics, vol.4, no.2 (2011): 89-102.
9. G. R. Aryal. “Transmuted log-logistic distribution.” Journal of statistics applications and probability, Vol.2, no.1 (2013): 11-20.
10. I. B. Abdul-Moniem and M. Seham. “Transmuted Gompertz distribution.” Computation and Applied Mathematics, vol.1, no.3 (2015): 88-96.
11. I. Elbatal, L. S. Diab and N. A. Abdul Alim. “Transmuted generalized linear exponential distribution.” International journal of computer applications, vol.83, no.17 (2013): 29-37.
12. K. Jayakumar and R. N. Pillai. “A first order autoregressive Mittag-Leffler process.” Journal of applied probability, vol.30 (1993):462-466.
13. M. Bourgnignon, I. Ghosh and G. M. Cordeiro. “General results for the transmuted family of distributions and new models.” Hindawi publishing corporation: Journal of probability and statistics, vol.2016(2016).

14. M. D. Nochols and W. J. Padgett. "A bootstrap control chart for Weibull percentiles." Quality and reliability Engineering International, vol.22 (2006): 141-151.
15. O. M. Bdair. "Different methods of estimation for Marshall-Olkin exponential distribution." Journal of applied statistical science, vol.19 (2012): 13-29.
16. R. C. Gupta, M. E. Ghitany and D. K. Mutairi. "Estimation of reliability from Marshall-Olkin extended lomax distributions." Journal of statistical computation and simulation, vol.80 (2010): 937-947.
17. W. Shaw and I. Buckley. "The alchemy of probability distributions: beyond Gram- Charlier expansions and a skew-kurtotic normal distribution from a rank transmutation map." Research report. (2007).
18. Z. M. Nofal, A. Z. Afify, H. M. Yousof and G. M. Cordeiro. "The generalized transmuted-G family of distributions." Communications in statistics, Theory and methods, vol.46 (2017): 4119-4136.

***