

1-1-2018

# Type I General Exponential Class of Distributions

Gholamhossein G. Hamedani

*Marquette University*, gholamhoss.hamedani@marquette.edu

Haitham M. Yousof

*Benha University*

Mahdi Rasekhi

*Malayer University*

Morad Alizadeh

*Gulf University-Bushehr*

Seyed Morteza Najibi

*Shiraz University*

# Type I General Exponential Class of Distributions

G.G. Hamedani

Department of Mathematics, Statistics and Computer Science, Marquette University, USA  
g.hamedani@mu.edu

Haitham M. Yousof

Department of Statistics, Mathematics and Insurance, Benha University, Egypt  
haitham.yousof@fcom.bu.edu

Mahdi Rasekhi

Department of Statistics, Malayer University, Malayer, Iran  
rasekhimahdi@gmail.com

Morad Alizadeh

Department of Statistics, Persian Gulf University, Bushehr, Iran  
moradalizadeh78@gmail.com

Seyed Morteza Najibi

Department of Statistics, College of Sciences, Shiraz University, Shiraz, Iran  
mor.najibi@gmail.com

## Abstract

We introduce a new family of continuous distributions and study the mathematical properties of the new family. Some useful characterizations based on the ratio of two truncated moments and hazard function are also presented. We estimate the model parameters by the maximum likelihood method and assess its performance based on biases and mean squared errors in a simulation study framework.

**Keywords:** Maximum likelihood; Moment; Order Statistic; Quantile function; Hazard function; Characterization.

## 1. Introduction

Several continuous univariate models have been widely used for modeling real data sets in many areas such as life sciences, engineering, economics, biological studies and environmental sciences to name a few. Various families of distributions have been constructed by extending common families of continuous distributions. These generalized distributions give more flexibility by adding one "or more" parameters to the baseline model. For example, Gupta et al. (1998) proposed the exponentiated-G class, which consists of raising the cumulative distribution function (cdf) to a positive power parameter. Many other classes can be cited such as the Marshall-Olkin-G family by Marshall and Olkin (1997), beta generalized-G family by Eugene et al. (2002), a new method for generating families of continuous distributions by Alzaatreh et al. (2013), exponentiated T-X family of distributions by Alzaghal et al. (2013), transmuted exponentiated generalized-G family by Yousof et al. (2015), Kumaraswamy transmuted-G by Afify et al. (2016b), transmuted geometric-G by Afify et al. (2016a), Burr X-G by Yousof et al. (2016), exponentiated transmuted-G family by Merovci et al. (2016), odd-Burr generalized family by Alizadeh et al. (2016a) the complementary generalized

transmuted poisson family by Alizadeh et al. (2016b), transmuted Weibull G family by Alizadeh et al. (2016c), the Type I half-logistic family by Cordeiro et al. (2016a), the Zografos-Balakrishnan odd log-logistic family of distributions by Cordeiro et al. (2016b), generalized transmuted-G by Nofal et al. (2017), the exponentiated generalized-G Poisson family by Aryal and Yousof (2017) and beta transmuted-H by Afify et al. (2017), the beta Weibull-G family by Yousof et al. (2017), among others.

This paper is organized as follows. In Section 2, we define the Type I General Exponential (TIGE) class of distributions. Some of its special cases are presented in Section 3. In Section 4, we derive some of its mathematical properties such as the asymptotic, shapes of the density and hazard rate functions, mixture representation for the density, quantile function, moments, moment generating function (mgf) and mean deviations. Section 5 deals with some characterizations of the new family and estimation of the model parameters using maximum likelihood. In Section 6, a simulation study is performed to see the efficiency of Maximum Likelihood method. In Section 7, we illustrate the importance of the new family by means of two applications to real data sets. The paper is concluded in Section 8.

## 2. Type I General Exponential Class of distributions

The cdf of TIGE distributions is given by

$$F(x) = F(x; \lambda, \alpha, \xi) = \exp\{\lambda[1 - G(x; \xi)^{-\alpha}]\}, \quad x \in \mathbb{R}, \quad (1)$$

where  $\xi = (\xi_k) = (\xi_1, \xi_2, \dots)$  is a parameter vector, and  $\lambda$  and  $\alpha$  are positive parameters. The corresponding probability density function (pdf) is

$$f(x) = f(x; \lambda, \alpha, \xi) = \lambda \alpha g(x; \xi) G(x; \xi)^{-(\alpha+1)} \exp\{\lambda[1 - G(x; \xi)^{-\alpha}]\}. \quad (2)$$

The reliability function (RF)  $[R(X)]$ , hazard rate function (HRF)  $[h(X)]$ , reversed-hazard rate function (RHR)  $[r(x)]$  and cumulative hazard rate function (CHR)  $[H(X)]$  of the TIGE family are given, respectively, by

$$R(x) = 1 - \exp\{\lambda[1 - G(x; \xi)^{-\alpha}]\},$$

$$h(x) = \frac{\lambda \alpha g(x; \xi) G(x; \xi)^{-(\alpha+1)} \exp\{\lambda[1 - G(x; \xi)^{-\alpha}]\}}{1 - \exp\{\lambda[1 - G(x; \xi)^{-\alpha}]\}},$$

$$r(x) = \lambda \alpha G(x; \xi)^{-(\alpha+1)}$$

and

$$H(x) = -\log(1 - \exp\{\lambda[1 - G(x; \xi)^{-\alpha}]\}).$$

If  $U \sim U(0,1)$  and  $Q_G(\cdot)$  denote the quantile function of  $G$ , then

$$X_U = Q_G \left\{ \left[ 1 - \frac{1}{\lambda} \log(U) \right]^{-\frac{1}{\alpha}} \right\}$$

has cdf (1). Henceforth  $G(x; \xi) = G(x)$  and  $g(x; \xi) = g(x)$  and so on. Several structural properties of the extended distributions may be easily explored using mixture forms of Exp-G models. Therefore, we obtain mixture forms of exponentiated-G (“Exp-G”) for  $F(x)$  and  $f(x)$ . The cdf of the TIGE family in (1) can be expressed as

$$F(x) = e^\lambda e^{-\lambda G(x)^{-\alpha}} = e^\lambda \sum_{i=0}^{\infty} \frac{(-\lambda)^i}{i!} G(x)^{-\alpha i}. \quad (3)$$

After expanding  $G(x)^{-\alpha i}$  and performing some algebra, we obtain

$$F(x) = e^\lambda \sum_{i,k=0}^{\infty} \sum_{j=k}^{\infty} \frac{(-1)^{i+j+k} \lambda^i}{i!} \binom{-\alpha i}{j} \binom{j}{k} G(x)^k,$$

or

$$F(x) = \sum_{k=0}^{\infty} b_k \Pi_k(x), \tag{4}$$

where

$$b_k = e^\lambda \sum_{i=0}^{\infty} \sum_{j=k}^{\infty} \frac{(-1)^{i+j+k} \lambda^i}{i!} \binom{-\alpha i}{j} \binom{j}{k}$$

and  $\Pi_\delta(x) = G(x)^\delta$  is the cdf of the Exp-G distribution with power parameter  $\delta$ . Furthermore, the corresponding TIGE density function is obtained by differentiating (4)

$$f(x) = \sum_{k=0}^{\infty} b_k \pi_k(x), \tag{5}$$

where  $\pi_\delta(x) = \delta g(x) G(x)^{\delta-1}$  is the pdf of the Exp-G distribution with power parameter  $\delta$ . The properties of Exp-G distributions have been studied by many authors in recent years, e.g. see Nadarajah (2005) for exponentiated Gumbel (EGu), Shirke and Kakade (2006) for exponentiated log-normal (ELN) and Nadarajah and Gupta (2007) for exponentiated gamma distributions (EGa), among others.

### 3. Properties

In this section, we investigate mathematical properties of the TIGE family of distributions including asymptotes, ordinary and incomplete moments, generating function, probability weighted moments and entropies. Established algebraic expansions to determine some structural properties of the TIGE family of distributions can be more efficient than computing them directly by numerical integration of its density function. The derivations derived throughout the article can be straightforwardly handled in most symbolic computation software platforms such as Mathematica, Maple and Matlab because of their ability to deal with such analytic expressions of enormous size and complexity.

Figure 1 shows different shapes of density and Hazard function for Type I General Exponential Weibull distribution(TIGEW).

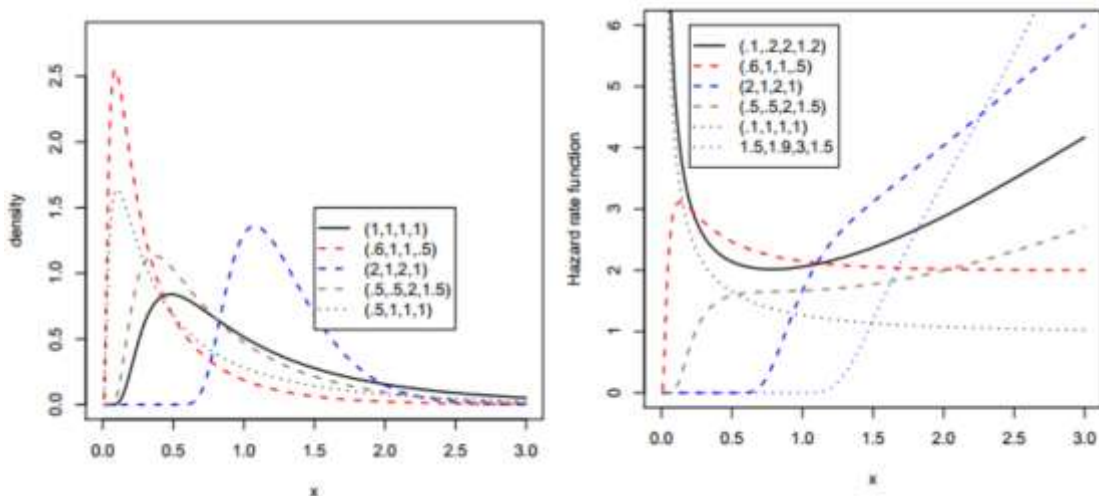


Figure 1: Different shapes of TIGEW pdf (left) and Hazard function (right)

### 3.1 Moments and generating function

The  $r$ th ordinary moment of  $X$  is given by

$$\mu'_r = E(X^r) = \int_{-\infty}^{\infty} x^r f(x)dx = \sum_{k=0}^{\infty} b_k E(Y_k^r), \tag{6}$$

where  $E(Y_k^r) = k \int_{-\infty}^{\infty} x^r g(x; \xi) G(x; \xi)^{k-1} dx$ , which can be computed numerically in terms of the baseline quantile function (qf)  $Q_G(u; \xi) = G^{-1}(u; \xi)$  such that  $E(Y_k^r) = k \int_0^1 Q_G(u; \xi)^r u^{k-1} du$ . For  $r = 1$  in (6), we obtain the mean of  $X$ . The last integration can be computed numerically for most parent distributions. The skewness and kurtosis measures can be calculated from the ordinary moments using well-known relationships. The  $r$ th central moment of  $X$ , say  $\mu_r$ , is given by

$$\mu_r = E(X - \mu)^r = \sum_{h=0}^r (-1)^h \binom{r}{h} (\mu'_1)^r \mu'_{r-h}.$$

The  $r$ th descending factorial moment of  $X$  (for  $r = 1, 2, \dots$ ) is given by

$$\mu'_{(r)} = E[X^{(r)}] = E[X(X - 1) \times \dots \times (X - r + 1)] = \sum_{k=0}^n s(r, k) \mu'_k,$$

where  $s(r, k) = (k!)^{-1} [d^k k^{(r)} / dx^k]_{x=0}$  is the Stirling number of the first kind. The cumulants ( $\kappa_n$ ) of  $X$  follow recursively from  $\kappa_n = \mu'_n - \sum_{r=0}^{n-1} \binom{n-1}{r-1} \kappa_r \mu'_{n-r}$ , where  $\kappa_1 = \mu'_1$ ,  $\kappa_2 = \mu'_2 - \mu_1'^2$ ,  $\kappa_3 = \mu'_3 - 3\mu_2' \mu_1' + \mu_1'^3$ , and so on. The skewness and kurtosis measures can also be calculated from the ordinary moments using well-known relationships. The main applications of the first incomplete moment refer to the mean deviations and the Bonferroni and Lorenz curves. These curves are very useful in economics, reliability, demography, insurance and medicine. The  $r$ th incomplete moment, say  $\varphi_r(t)$ , of  $X$  can be expressed from (5), as

$$\varphi_r(t) = \int_{-\infty}^t x^r f(x)dx = \sum_{k=0}^{\infty} b_k \int_{-\infty}^t x^r \pi_k(x)dx. \tag{7}$$

The mean deviations about the mean [ $\delta_1 = E(|X - \mu'_1|)$ ] and about the median [ $\delta_2 = E(|X - Me(X)|)$ ] of  $X$  are given by  $\delta_1 = 2\mu'_1 F(\mu'_1) - 2\varphi_1(\mu'_1)$  and  $\delta_2 = \mu'_1 - 2\varphi_1(M)$ , respectively, where  $\mu'_1 = E(X)$ ,  $Me(X) = Q(0.5)$  is the median,  $F(\mu'_1)$  is easily calculated from (1) and  $\varphi_1(t)$  is the first incomplete moment given by (7) with  $r = 1$ . A general equation for  $\varphi_1(t)$  can be derived from (7) as  $\varphi_1(t) = \sum_{k=0}^{\infty} b_k I_k(x)$ , where  $I_k(x) = \int_{-\infty}^t x \pi_k(x)dx$  is the first incomplete moment of the Exp-G model. The moment generating function (mgf)  $M_X(t) = E(e^{tX})$  of  $X$  can be derived from equation (5) as  $M_X(t) = \sum_{k=0}^{\infty} b_k M_k(t)$ , where  $M_\delta(t)$  is the mgf of  $Y_\delta$ . Hence,  $M_X(t)$  can be determined from the Exp-G generating function.

### 3.2 Probability weighted moments

The Probability weighted moments (PWMs) are expectations of certain functions of a random variable and they can be defined for any random variable whose ordinary moments exist. The PWMs method can generally be used for estimating parameters of a

distribution whose inverse form cannot be expressed explicitly. The  $(s, r)$ th PWMs of  $X$  following the TIGE family, say  $\rho_{s,r}$ , is given by

$$\rho_{s,r} = E\{X^s F(X)^r\} = \int_{-\infty}^{\infty} x^s F(x)^r f(x) dx.$$

Using equations (1) and (2), we can write

$$f(x)F(x)^r = \sum_{k=0}^{\infty} a_k \pi_k(x),$$

where

$$a_k = e^\lambda \sum_{i=0}^{\infty} \sum_{j=k}^{\infty} \frac{(-1)^{i+j+k} [\lambda(r+1)]^i}{i!} \binom{-\alpha i}{j} \binom{j}{k}.$$

Then, the  $(s, r)$ th PWMs of  $X$  can be expressed as

$$\rho_{s,r} = \sum_{k=0}^{\infty} a_k E(Y_k^s).$$

### 3.3 Order statistics and their moments

The order statistics plays an important role in real-life applications involving data relating to life testing studies. These statistics are required in many fields, such as climatology, engineering and industry to name a few. Suppose that  $X_1, \dots, X_n$  constitute a random sample from a TIGE distribution,  $X_{i:n}$  denotes the  $i$ th order statistic of this sample and  $f_{i:n}(x)$  denotes its pdf. We wish to find a linear expansion for  $f_{i:n}(x)$ . First, we note that

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

where  $K = \frac{1}{B(i, n-i+1)}$ . Now, consider the following equation:

$$\left( \sum_{i=0}^{\infty} a_i u^i \right)^n = \sum_{i=0}^{\infty} d_{n,i} u^i,$$

where  $n$  is a positive integer and the coefficients  $d_{n,i}$  (for  $i = 1, 2, \dots$ ) are determined from the recurrence equation (with  $d_{n,0} = a_0^n$ )

$$d_{n,i} = (i a_0)^{-1} \sum_{m=1}^i [m(n+1) - i] a_m d_{n,i-m}.$$

Then, the density function of the  $X_{i:n}$  can be expressed as

$$f_{i:n}(x) = \sum_{r,k=0}^{\infty} m_{r,k} \pi_{r+k+1}(x), \tag{8}$$

where  $\pi_{r+k+1}(x)$  stands for the pdf of the Exp-G distribution with power parameter  $r + k + 1$ .

$$m_{r,k} = \frac{n! (r+1) b_{r+1}}{(r+k+1) (i-1)!} \sum_{j=0}^{n-i} \frac{(-1)^j f_{j+i-1,k}^*}{(n-i-j)! j!}$$

in which,  $b_r$  is defined in Section 2 and the coefficients  $f_{j+i-1,k}^*$ 's can be determined such that  $f_{j+i-1,0}^* = b_0^{j+i-1}$  and for  $k \geq 1$

$$f_{j+i-1,k}^* = (k b_0)^{-1} \sum_{w=1}^k [w(j+i) - k] b_w f_{j+i-1,k-w}^*$$

Equation (8) reveals that the pdf of the TIGE order statistic can be expressed as a linear combination of the Exp-G densities. Therefore, some statistical and mathematical properties of these order statistics can be obtained by using this result. Analogous to the ordinary moments, we can derive the L-moments but it can be estimated by the linear combinations of order statistics in (8). They exist as long as the mean of the distribution exists, even if some higher moments may not exist, and are relatively robust to the effects of outliers. Based upon the moments in equation (8), we can derive explicit expressions for the L-moments of  $X$  as infinite weighted linear combinations of the means of suitable TIGE order statistics. They are linear functions of expected order statistics defined by

$$\lambda_r = \frac{1}{r} \sum_{d=0}^{r-1} (-1)^d \binom{r-1}{d} E(X_{r-d:r}), r \geq 1.$$

### 3.4 Moments of the Residual and Reversed Residual Lives

The  $n$ th moment of the residual life, say  $z_n(t) = E[(X - t)^n | X > t], n = 1, 2, \dots$ , uniquely determines  $F(x)$ . The  $n$ th moment of the residual life of  $X$  using equation (5) is given by

$$z_n(t) = \frac{1}{1 - F(t)} \int_t^\infty (x - t)^n dF(x) = \frac{1}{1 - F(t)} \sum_{k=0}^\infty b_k^* \int_t^\infty x^r \pi_k(x) dx,$$

where  $b_k^* = b_k \sum_{r=0}^n \binom{n}{r} (-t)^{n-r}$ . Another interesting function is the mean residual life (MRL) function or the life expectation at age  $t$  defined by  $z_1(t) = E[(X - t) | X > t]$ , which represents the expected additional life length for a unit which is alive at age  $t$ . The MRL has many applications in survival analysis in biomedical sciences, life insurance, maintenance and product quality control, economics and social studies, demography and product technology. The MRL of  $X$  can be obtained by setting  $n = 1$  in the last equation. The  $n$ th moment of the reversed residual life, say  $Z_n(t) = E[(t - X)^n | X \leq t]$  for  $t > 0$  and  $n = 1, 2, \dots$  uniquely determines  $F(x)$ . Therefore, The  $n$ th moment of the reversed residual life of  $X$  is given by

$$Z_n(t) = \frac{1}{F(t)} \int_0^t (t - x)^n dF(x) = \frac{1}{F(t)} \sum_{k=0}^\infty b_k^{**} \int_0^t x^r \pi_k(x) dx,$$

where  $b_k^{**} = b_k \sum_{r=0}^n (-1)^r \binom{n}{r} t^{n-r}$ . The mean inactivity time (MIT) or mean waiting time (MWT) also called the mean reversed residual life function is given by  $Z_1(t) = E[(t - X) | X \leq t]$ , and it represents the waiting time elapsed since the failure of an item on condition that this failure had occurred in  $(0, t)$ . The MIT of the TIGE family of distributions can be obtained easily by setting  $n = 1$  in the above equation.

### 3.5 Stress-strength model

The measure of dependability (reliability) of industrial components has many applications especially in the area of lifetime testing and engineering, to name a few. Stress-strength model is the most widely approach used for dependability estimation. This model is used in many applications in physics and engineering such as strength failure and system collapse. In stress-strength modeling,  $\mathbf{R} = \Pr(X_2 < X_1)$  is a measure of dependability of the system when it is subjected to random stress  $X_2$  and has strength  $X_1$  (e.g. see Kotz et al, 2003). The system fails if and only if the applied stress is greater than its strength and the component will function satisfactorily whenever  $X_1 > X_2$ .  $\mathbf{R}$  can be considered as a measure of system performance and naturally arise in electrical and electronic systems. Other interpretation is that, the reliability of the system is the probability of being the system strongly enough to overcome the stress imposed on it. Let  $X_1$  and  $X_2$  be two independent random variables with  $\text{TIGE}(\lambda_1, \alpha_1, \xi)$  and  $\text{TIGE}(\lambda_2, \alpha_2, \xi)$  distributions with the same parameter vector  $\xi$  for the baseline G. The reliability is defined by  $\mathbf{R} = \int_0^\infty f_1(x; \lambda_1, \alpha_1, \xi)F_2(x; \lambda_2, \alpha_2, \xi)dx$ . Then, We can write

$$\mathbf{R} = \sum_{k,h=0}^\infty b_{k,h} \int_0^\infty \pi_{k+h}(x)dx,$$

where

$$b_{k,h} = e^{\lambda_1 + \lambda_2} \sum_{i,w=0}^\infty \sum_{j=k}^\infty \sum_{m=h}^\infty \frac{(-1)^{i+j+k+w+m+h} \lambda_1^i \lambda_2^w}{i! w!} \binom{k}{k+h+1} \binom{-\alpha_1 i}{j} \binom{-\alpha_2 w}{m} \binom{j}{k} \binom{m}{h}.$$

Thus, the reliability can be reduced to

$$\mathbf{R} = \sum_{k,h=0}^\infty b_{k,h}.$$

### 4. Characterizations

In this section, we present characterizations of TIGE family based on a simple relationship between two truncated moments. Our characterization result employs a theorem due to Glänzel (1987), see Theorem 1 below. Note that the result holds also when the interval  $H$  is not closed. Moreover, it could be also applied when the *cdf*  $F$  does not have a closed form. As shown in Glänzel (1990), this characterization is stable in the sense of weak convergence.

**Theorem 1.** Let  $(\Omega, \mathcal{F}, \mathbf{P})$  be a given probability space and let  $H = [d, e]$  be an interval for some  $d < e$  ( $d = -\infty, e = \infty$  might as well be allowed). Let  $X: \Omega \rightarrow H$  be a continuous random variable with the distribution function  $F$  and let  $q_1$  and  $q_2$  be two real functions defined on  $H$  such that

$$\mathbf{E}[q_2(X)|X \geq x] = \mathbf{E}[q_1(X)|X \geq x]\eta(x), \quad x \in H,$$

is defined with some real function  $\eta$ . Assume that  $q_1, q_2 \in C^1(H)$ ,  $\eta \in C^2(H)$  and  $F$  is twice continuously differentiable and strictly monotone function on the set  $H$ . Finally, assume that the Equation  $\eta q_1 = q_2$  has no real solution in the interior of  $H$ . Then  $F$  is uniquely determined by the functions  $q_1, q_2$  and  $\eta$ , particularly

$$F(x) = \int_a^x C \left| \frac{\eta'(u)}{\eta(u)q_1(u) - q_2(u)} \right| \exp(-s(u))du,$$



where the function  $s$  is a solution of the differential Equation  $s' = \frac{\eta' q_1}{\eta q_1 - q_2}$  and  $C$  is the normalization constant, such that  $\int_H dF = 1$ . Here is our first characterization.

**Proposition 1.** Let  $X: \Omega \rightarrow \mathbb{R}$  be a continuous random variable and let  $q_1(x) = \exp\{-\lambda(1 - (G(x))^{-\alpha})\}$  and  $q_2(x) = q_1(x)(G(x))^{-\alpha}$  for  $x \in \mathbb{R}$ . The random variable  $X$  belongs to the TIGE family if and only if the function  $\eta$  defined in Theorem 1 has the form

$$\eta(x) = \frac{1}{2}\{1 + (G(x))^{-\alpha}\}, \quad x \in \mathbb{R}.$$

Proof. Let  $X$  be a random variable with *pdf* of TIGE family, then

$$(1 - F(x))E[q_1(x)|X \geq x] = \lambda\{(G(x))^{-\alpha} - 1\}, \quad x \in \mathbb{R},$$

and

$$(1 - F(x))E[q_2(x)|X \geq x] = \frac{\lambda}{2}\{(G(x))^{-2\alpha} - 1\} \quad x \in \mathbb{R},$$

and finally

$$\eta(x)q_1(x) - q_2(x) = \frac{q_1(x)}{2}\{1 - (G(x))^{-\alpha}\} > 0 \quad \text{for } x \in \mathbb{R}.$$

Conversely, if  $\eta$  is given as above, then

$$s'(x) = \frac{\eta'(x)q_1(x)}{\eta(x)q_1(x) - q_2(x)} = -\frac{\alpha g(x)(G(x))^{-(\alpha+1)}}{1 - (G(x))^{-\alpha}}, \quad x \in \mathbb{R},$$

and hence

$$s(x) = -\log\{1 - (G(x))^{-\alpha}\}, \quad x \in \mathbb{R}.$$

Now, in view of Theorem 1,  $X$  has density of the TIGE family in (2).

**Corollary 1.** Let  $X: \Omega \rightarrow \mathbb{R}$  be a continuous random variable and let  $q_1(x)$  be as in Proposition 1. Then  $X$  has *pdf* (2) if and only if there exist functions  $q_2$  and  $\eta$  defined in Theorem 1 satisfying the differential equation

$$\frac{\eta'(x)q_1(x)}{\eta(x)q_1(x) - q_2(x)} = -\frac{\alpha g(x)}{(G(x))\{(G(x))^\alpha - 1\}}, \quad x \in \mathbb{R}. \tag{9}$$

The general solution of the differential equation in Corollary 1 is

$$\eta(x) = \{1 - (G(x))^{-\alpha}\}^{-1} \left[ \int \alpha g(x)(G(x))^{-(\alpha+1)}(q_1(x))^{-1} q_2(x) dx + D \right], \tag{10}$$

where  $D$  is a constant. Note that a set of functions satisfying the differential equation (9) is given in Proposition 1 with  $D = \frac{1}{2}$ . However, it should be also noted that there are other triplets  $(q_1, q_2, \eta)$  satisfying the conditions of Theorem 1.

### 5. Maximum likelihood estimation

Several approaches for parameter estimation were proposed in the literature but the maximum likelihood method is the most commonly employed. The maximum likelihood estimators (MLEs) enjoy desirable properties and can be used for constructing confidence intervals and regions, and the test statistics. The normal approximation for these estimators in large samples can be easily handled either analytically or numerically. Therefore, we consider the estimation of the unknown parameters of this family from complete samples only by maximum likelihood. Let  $x_1, \dots, x_n$  be a random sample from the TIGE distribution with parameters  $\lambda, \alpha$  and  $\xi$ . Let  $\Theta = (\lambda, \alpha, \xi)^u$  be the  $p \times 1$  parameter vector. For determining the MLE of  $\Theta$ , we have the log-likelihood function

$$\begin{aligned} \ell = \ell(\Theta) = n \log \lambda + n \log \alpha + \sum_{i=1}^n \log g(x_i; \xi) - (\alpha + 1) \sum_{i=1}^n \log G(x_i; \xi) \\ + \lambda \sum_{i=1}^n [1 - G(x_i; \xi)^{-\alpha}]. \end{aligned}$$

The components of the score vector,  $\mathbf{U}(\Theta) = \frac{\partial \ell}{\partial \Theta} = \left( \frac{\partial \ell}{\partial \lambda}, \frac{\partial \ell}{\partial \alpha}, \frac{\partial \ell}{\partial \xi} \right)^T$ , are given as

$$\begin{aligned} U_\lambda &= \frac{n}{\lambda} + \sum_{i=1}^n [1 - G(x_i; \xi)^{-\alpha}], \\ U_\alpha &= \frac{n}{\alpha} - \sum_{i=1}^n \log G(x_i; \xi) + \lambda \sum_{i=1}^n G(x_i; \xi)^{-\alpha} \log G(x_i; \xi) \end{aligned}$$

and (for  $r = 1, \dots, q$ )

$$U_{\xi_r} = \sum_{i=1}^n \frac{g'_r(x_i; \xi)}{g(x_i; \xi)} - (\alpha + 1) \sum_{i=1}^n \frac{G'_r(x_i; \xi)}{G(x_i; \xi)} + \lambda \alpha \sum_{i=1}^n G'_r(x_i; \xi) G(x_i; \xi)^{-\alpha-1},$$

where

$$g'_r(x_i; \xi) = \frac{\partial}{\partial \xi_r} [g(x_i; \xi)], G'_r(x_i; \xi) = \frac{\partial}{\partial \xi_r} [G(x_i; \xi)].$$

Setting the nonlinear system of equations  $U_\lambda = U_\alpha = 0$  and  $U_\xi = \mathbf{0}$  and solving them simultaneously yields the MLE  $\hat{\Theta} = (\hat{\lambda}, \hat{\alpha}, \hat{\xi}^u)^u$ . To solve these equations, it is usually more convenient to use nonlinear optimization methods such as the quasi-Newton algorithm to numerically maximize  $\ell$ . For interval estimation of the parameters, we obtain the  $p \times p$  observed information matrix  $J(\Theta) = \left\{ \frac{\partial^2 \ell}{\partial r \partial s} \right\}$  (for  $r, s = \lambda, \alpha, \xi$ ), whose elements can be computed numerically. Under the standard regularity conditions when  $n \rightarrow \infty$ , the distribution of  $\hat{\Theta}$  can be approximated by a multivariate normal  $N_p(0, J(\hat{\Theta})^{-1})$  distribution to construct approximate confidence intervals for the parameters. Here,  $J(\hat{\Theta})$  is the total observed information matrix evaluated at  $\hat{\Theta}$ .

## 6. Simulation study

In this section, we survey the performance of the MLEs of the Type I General Exponential Weibull (TIGEW) distribution with respect to sample size  $n$ . This performance is done based on the following simulation study:

1. Generate 5000 samples of size  $n$  from TIGEW distribution. The inversion method was used to generate samples.
2. Compute the MLEs for 5000 thousand samples, say  $(\hat{\alpha}, \hat{\lambda}, \hat{\beta}, \hat{c})$  for  $i = 1, 2, \dots, 5000$ .
3. Compute the standard errors of the MLEs for the ten thousand samples, say  $(s_{\hat{\alpha}}, s_{\hat{\lambda}}, s_{\hat{\beta}}, s_{\hat{c}})$  for  $i = 1, 2, \dots, 5000$ . The standard errors were computed by inverting the observed information matrices.
4. Compute the biases, mean squared errors and coverage lengths given by

$$Bias(n) = \frac{1}{5000} \sum_{i=1}^{5000} (\hat{\epsilon}_i - \epsilon),$$

$$MSE(n) = \frac{1}{5000} \sum_{i=1}^{5000} (\hat{\epsilon}_i - \epsilon)^2$$

and

$$CL(n) = \frac{3.9199}{5000} \sum_{i=1}^{5000} s_{\hat{\epsilon}_i}$$

respectively, when  $\epsilon = (\alpha, \lambda, \beta, c)$ . Then these steps are repeated for  $n = 100, 105, 110, \dots, 300$  with  $\alpha = 1, \lambda = 1, \beta = 1$  and  $c = 1$ , so computing  $Bias(n), MSE(n), CL(n)$  for  $\epsilon = (\alpha, \lambda, \beta, c)$ . Figure 2 shows the variation of four biases with respect to  $n$ . The biases for each parameter either decrease or increase to zero as  $n$  goes to infinity. Figure 3 shows how the four mean squared errors vary with respect to  $n$ . The mean squared errors for each parameter decrease to zero as  $n \rightarrow \infty$ . In Figure 4, we show coverage lengths for each parameters with respect  $n$ . The coverage lengths for each parameter decrease to zero as  $n \rightarrow \infty$ .

The observations of this study are for only one choice for  $(\alpha, \lambda, \beta, c)$ , namely that  $(\alpha, \lambda, \beta, c) = (1, 1, 1, 1)$ . But the results were similar for other choices for  $(\alpha, \lambda, \beta, c)$ . In particular, 1) the biases for each parameter either decreased or increased to zero and appeared reasonably small at  $n = 300$ ; 2) the mean squared errors for each parameter decreased to zero and appeared reasonably small at  $n = 300$ ; 3) the coverage lengths for each parameter decreased to zero and appeared reasonably small at  $n = 300$ .

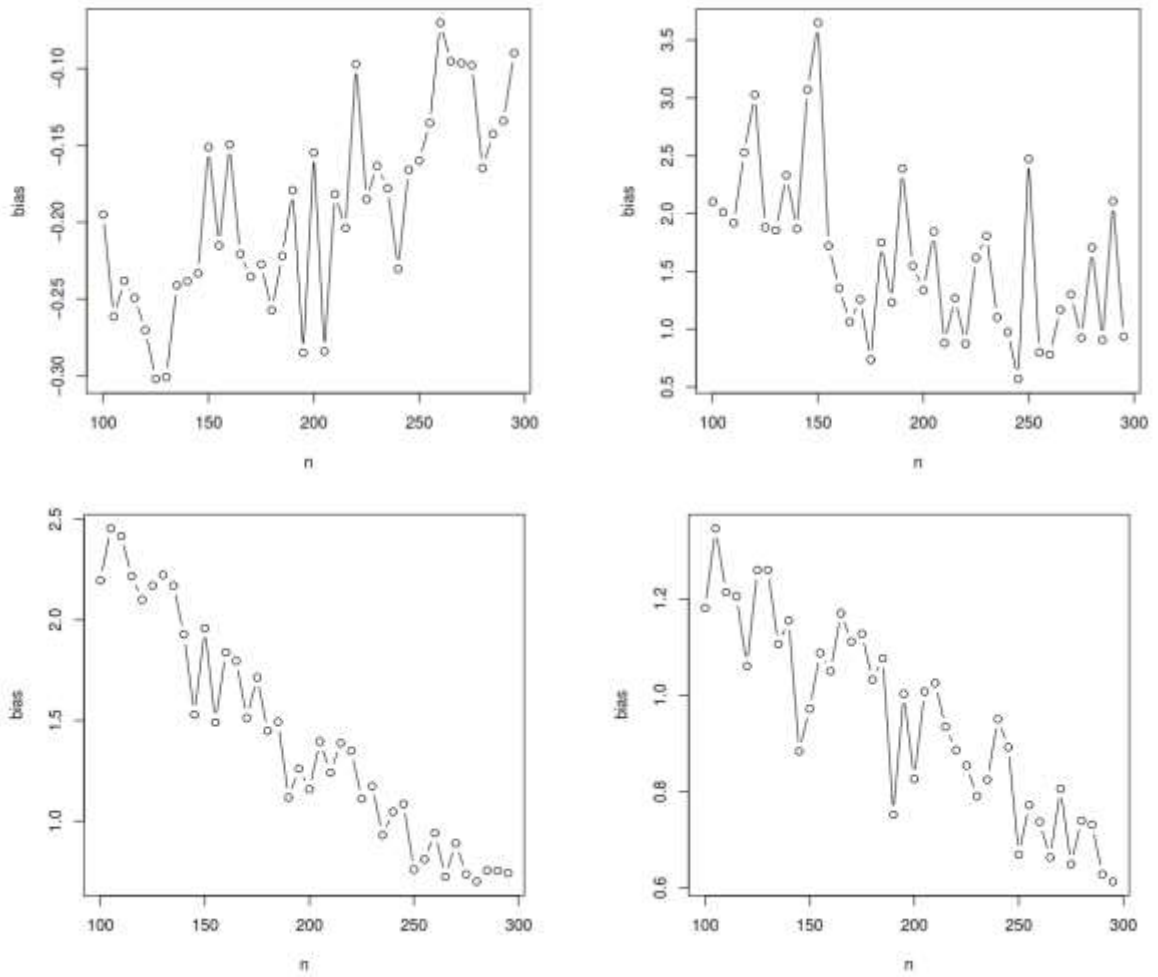
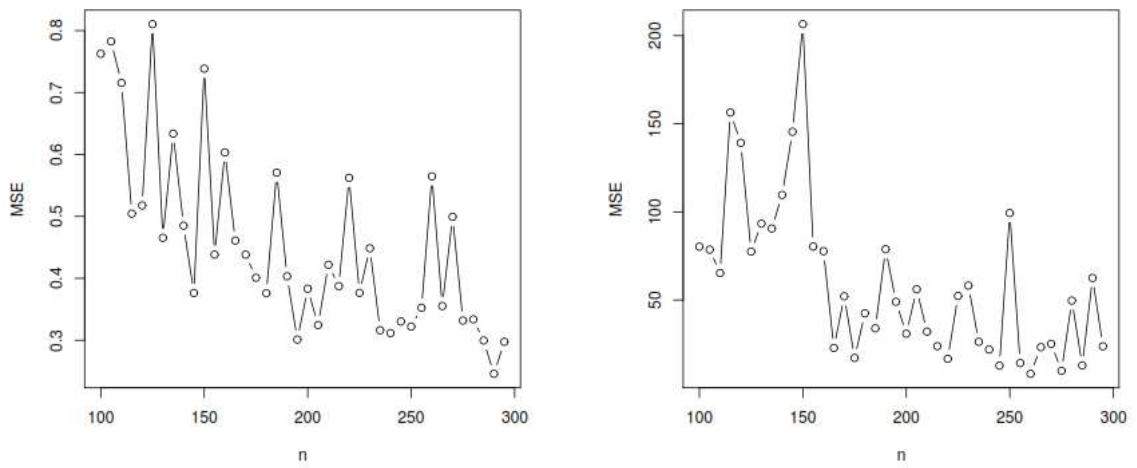


Figure 2:  $Bias_{\alpha}(n)$  (top left),  $Bias_{\lambda}(n)$  (top right),  $Bias_{\beta}(n)$  (bottom left) and  $Bias_c(n)$  (bottom right) versus  $n = 100, 105, 110, \dots, 300$ .



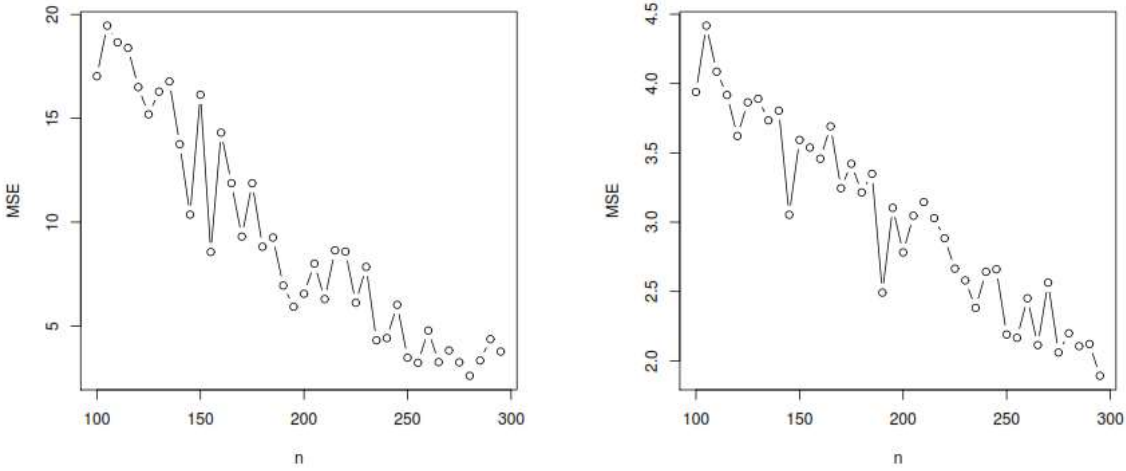


Figure 3:  $MSE_{\alpha}(n)$  (top left),  $MSE_{\lambda}(n)$  (top right),  $MSE_{\beta}(n)$  (bottom left) and  $MSE_c(n)$  (bottom right) versus  $n = 100, 105, 110, \dots, 300$ .

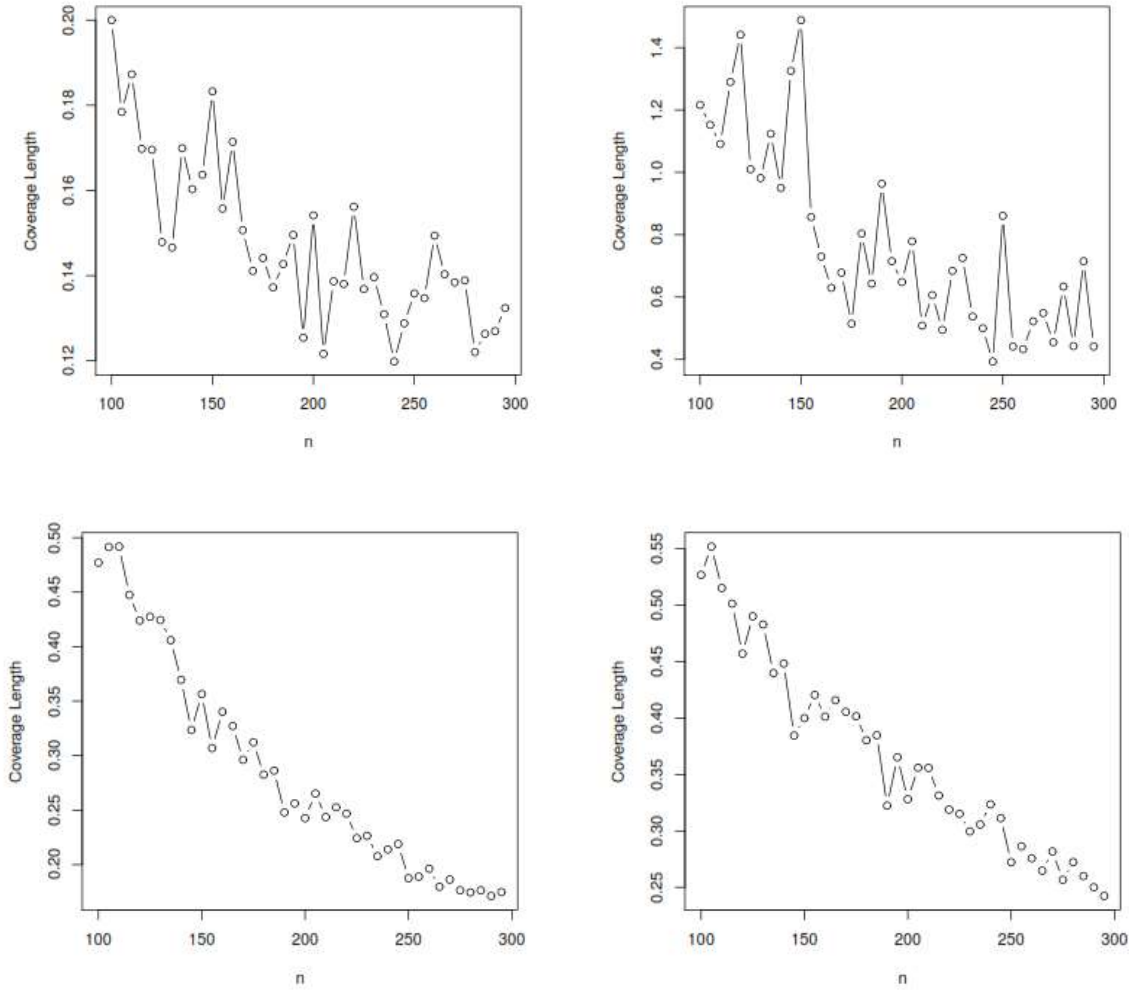


Figure 4:  $CL_{\alpha}(n)$  (top left),  $CL_{\lambda}(n)$  (top right),  $CL_{\beta}(n)$  (bottom left) and  $CL_c(n)$  (bottom right) versus  $n = 100, 105, 110, \dots, 300$ .

### 7. Applications

We demonstrate the flexibility of the TIGEW distribution in the application by help of two data sets. We compare TIGEW with Kw-Weibull (Cordero et al., 2010), Beta-Weibull (Lee et al., 2007) and Beta- Exponentiated Weibull (Cordeiro et al., 2013) distributions. The first data set is given by Ghitany et al. (2008) on the waiting times (in minutes) before service of 100 Bank customers. The data set consists of 202 observations and was also analyzed by Gupta and Singh (2013), Al-Zahrani and Gindwan (2014) and Shanker (2015). The second data set is the body mass index from 202 elite Australian athletes who trained at the Australian Institute of Sport from Cook and Weisberg (1994). These data are existed in "DPpackage" of R statistical program.

The MLE of parameters, –maximized log-likelihood function, Akaike information criterion (AIC), Bayesian information criterion (BIC), Hannan-Quinn information criterion (HQIC), Consistent Akaike information criterion (CAIC) (Burnham and Anderson, 2002) statistics are determined by fitting mentioned distributions using the two data sets. In general, the smaller values of these statistics show the better fit to the data sets. The MLEs are computed using the Limited-Memory Quasi-Newton Code for Bound-Constrained Optimization (L-BFGS-B). The estimated parameters based on MLE procedure are ginen in Tables 1 and 2, whereas the values of goodness-of-fit statistics are given in Tables 3 and 4. In the applications, the information about the hazard shape can help in selecting a particular model. To do so, a device called the total time on test (TTT) plot (Aarset, 1987) is useful. The TTT plot is obtained by plotting

$$G(r/n) = \left[ \left( \sum_{i=1}^r y_{i,n} \right) + (n - r)y_{(r)} \right] / \sum_{i=1}^n y_{i,n}$$

where  $r = 1, \dots, n$  and  $y_{i,n}$  ( $i = 1, \dots, n$ ) are the order statistics of the sample, against  $r/n$ . If the shape is a straight diagonal the hazard is constant. It is convex shape for decreasing hazards and concave shape for increasing hazards.

The TTT plot for both data sets presented in Figure 5, from this Figure we note that first and second data sets have increasing failure rate functions. In both real data sets, the results show that the TIGEW distribution yields a better fit than other distributions.

**Table 1: Parameters estimates and standard deviation in parenthesis for first dataset**

Model	Estimates	-Log Likelihood
TIGEW( $\alpha, \beta, \lambda, c$ )	0.20(0.01), 7.39(0.73), 1.09(0.07), 7.83(0.56)	316.955
Kw-W( $a, b, \beta, c$ )	2.33(0.40), 0.30(0.03), 1.13(0.04), 0.33(0.02)	316.975
BW( $a, b, \beta, c$ )	3.56(0.22), 4.94(0.31), 0.73(0.04), 0.05( $3e^{-3}$ )	317.023
BEW( $a, b, \alpha, c, \lambda$ )	26.98(1.07), 7.88(0.30), 0.14( $5e^{-3}$ ), 0.89(0.03), 0.02( $1e^{-3}$ )	316.994

**Table 2: Parameters estimates and standard deviation in parenthesis for second dataset**

Model	Estimates	-Log Likelihood
TIGEW( $\alpha, \beta, \lambda, c$ )	0.90(0.06), 60.54(4.25), 2.16(0.02), 11.44(0.09)	488.987
kw-W( $a, b, \beta, c$ )	33.82(2.34), 1.03(0.07), 2.41(0.03), 0.07( $6e^{-4}$ )	489.108
BW( $a, b, \beta, c$ )	76.09(2.59), 4.45(0.14), 1.33(0.01), 0.09( $8e^{-3}$ )	490.102

**Table 3: Formal goodness of fit statistics for first dataset**

Model	Goodness of fit criteria			
	<i>AIC</i>	<i>BIC</i>	<i>HQIC</i>	<i>CAIC</i>
TIGEW	641.91	652.33	640.01	642.33
Kw-W	641.94	652.37	640.05	642.37
BW	642.04	652.46	640.15	642.46
BEW	643.98	657.01	641.62	644.62

**Table 4: Formal goodness of fit statistics for second dataset**

Model	Goodness of fit criteria			
	<i>AIC</i>	<i>BIC</i>	<i>HQIC</i>	<i>CAIC</i>
TIGEW	985.97	999.20	984.65	986.17
Kw-W	986.217	999.45	984.89	986.42
BW	988.205	1001.43	986.88	988.40

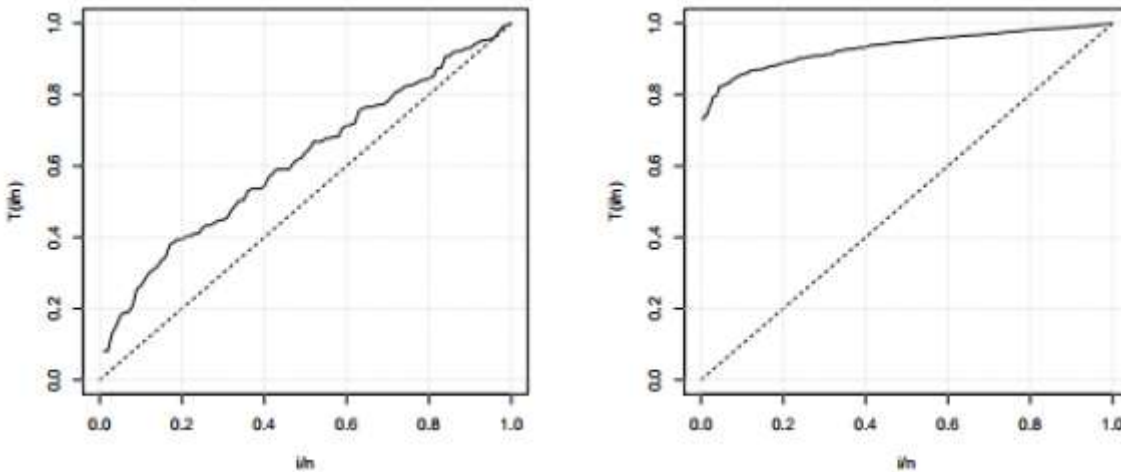


Figure 5 : TTT-plot for the first dataset (left Fig.) and for the second dataset (right Fig.)

## **8. Conclusions**

There has been a great interest among the statisticians and practitioners to generate new extended families. In this paper, we present a new class of distributions called the Type I General Exponential (TIGE) family of distributions, The mathematical properties of this new family including explicit expansions for the ordinary and incomplete moments, generating function, mean deviations, order statistics, probability weighted moments are provided. Characterizations based on two truncated moments are presented. The model parameters are estimated by the maximum likelihood estimation method and the observed information matrix is determined. Simulation results to assess the performance of the maximum likelihood estimators are discussed. It is shown that a special case of the TIGE class (TIGEW) can provide a better fit than other models generated by well-known families in two real data sets.

## **Acknowledgements**

The authors are grateful to Professor Nadeem Shafique Butt and the reviewers for their comments and suggestions which certainly improved the presentation of the content of this work.

## **References**

1. Afify A. Z., Alizadeh, M., Yousof, H.M., Aryal, G. and Ahmad, M. (2016a). The transmuted geometric-G family of distributions: theory and applications. Pak. J. Statist., forthcoming.
2. Afify A. Z., Cordeiro, G.M., Yousof, H.M., Alzaatreh, A. and Nofal, Z.M. (2016b). The Kumaraswamy transmuted-G family of distributions: properties and applications. J. Data Sci., forthcoming.
3. Afify, A.Z., Yousof, H.M. and Nadarajah, S. (2017). The beta transmuted-H family of distributions: properties and applications. *Statistics and its Inference*, 10, 505–520.
4. Alizadeh, M., Cordeiro, G.M., Nascimento, A.D.C. Lima M.D.S. and Ortega, E.M.M. (2016a). Odd-Burr generalized family of distributions with some applications. *Journal of Statistical Computation and Simulation*, 83, 326-339.
5. Alizadeh, M., Yousof, H.M., Afify A.Z., Cordeiro, G.M., and Mansoor, M. (2016b). The complementary generalized transmuted poisson-G family. *Austrian Journal of Statistics*, forthcoming.
6. Alizadeh, M., Rasekhi, M., Yousof, H. M. and Hamedani G. G. (2016c). The transmuted Weibull G family of distributions. *Hacettepe Journal of Mathematics and Statistics*, forthcoming.
7. Aryal, G. R. and Yousof, H. M. (2017). The exponentiated generalized-G Poisson family of distributions. *Economic Quality Control*, forthcoming.
8. Alzaatreh, A., Lee, C. and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71, 63-79.



9. Alzaghal, A., Famoye, F., Lee, C. (2013). Exponentiated T-X family of distributions with some applications. *International Journal of Probability and Statistics*, 2, 31-49.
10. Al-Zahrani, B., Gindwan, M. (2014). Parameter estimation of a two-parameter Lindley distribution under hybrid censoring. *International Journal of System Assurance Engineering and Management*, 5: 628-636.
11. Cook, R. D., Weisberg, S. (1994), *An Introduction to Regression Graphics*. John Wiley.
12. Cordeiro, G. M., Alizadeh, M. and Diniz Marinho, P. R. (2016a). The Type I half-logistic family of distributions. *Journal of Statistical Computation and Simulation*, 86, 707-728.
13. Cordeiro, G. M., Alizadeh, M., Ortega, E. M. and Serrano, L. H. V. (2016b). The Zografos-Balakrishnan odd log-logistic family of distributions: Properties and Applications. *Hacetatepe Journal of Mathematics and Statistics*, forthcoming.
14. Cordeiro, G.M., Gomes, A.E., da-Silva, C.Q. and Ortega, E.M. (2013). The beta exponentiated Weibull distribution. *Journal of Statistical Computation and Simulation*, 83, 114-138.
15. Cordeiro, G.M., Ortega, E.M. and Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute*, 347, 1317-1336.
16. Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Commun. Stat. Theory Methods*, 31, 497-512.
17. Famoye, F., Lee, C. and Olumolade, O. (2005). The Beta-Weibull Distribution, *Journal of Statistical Theory and Applications* 4, 121-136.
18. Ghitany, M.E., Atieh, B. and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78, 493-506.
19. Glänzel, W. (1987). A characterization theorem based on truncated moments and its application to some distribution families, *Mathematical Statistics and Probability Theory (Bad Tatzmannsdorf, 1986)*, Vol. B, Reidel, Dordrecht, 75-84.
20. Glänzel, W. (1990). Some consequences of a characterization theorem based on truncated moments, *Statistics: A Journal of Theoretical and Applied Statistics*, 21, 613-618.
21. Gupta, R.C., Gupta, P. L. and Gupta, R. D. (1998). Modeling failure time data by Lehmann alternatives. *Commun. Stat. Theory Methods*, 27, 887-904.
22. Gupta, P.K., Singh, B. (2013). *International Journal of System Assurance Engineering and Management*. *International Journal of System Assurance Engineering and Management*, 4: 378-385.
23. Kotz, S., Lai, C.D., and Xie M. (2003). On the effect of redundancy for systems with dependent components. *IIE Trans*, 35, 1103–1110.

*Type I General Exponential Class of Distributions*

24. Lee, C., Famoye, F. and Olumolade, O. (2007). Beta-Weibull distribution: some properties and applications to censored data. *Journal of Modern Applied Statistical Methods*, 6, 173-186.
25. Marshall, A. W. and Olkin, I. (1997). A new methods for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84, 641-652.
26. Merovci, F., Alizadeh, M., Yousof, H. M. and Hamedani G. G. (2016). The exponentiated transmuted-G family of distributions: theory and applications. *Commun. Stat. Theory Methods*, forthcoming.
27. Nadarajah, S. (2005). The exponentiated Gumbel distribution with climate application. *Environmetrics*, 17, 13–23.
28. Nadarajah, S. and Gupta, A.K. (2007). The exponentiated gamma distribution with application to drought data. *Calcutta Statistical Association Bulletin*, 59, 29–54.
29. Nofal, Z. M., Afify, A. Z., Yousof, H. M. and Cordeiro, G. M. (2017). The generalized transmuted-G family of distributions. *Commun. Stat. Theory Methods*, 46, 4119-4136.
30. Shanker, R. (2015). Shanker Distribution and Its Applications. *International Journal of Statistics and Applications*, 5, 338-348.
31. Shirke, D.T and Kakade, C.S. (2006). On exponentiated lognormal distribution. *International Journal of Agricultural and Statistical Sciences*, 2, 319–326.
32. Yousof, H. M., Afify, A. Z., Alizadeh, M., Butt, N. S., Hamedani, G. G. and Ali, M. M. (2015). The transmuted exponentiated generalized-G family of distributions, *Pak. J. Stat. Oper. Res.*, 11, 441-464.
33. Yousof, H. M., Afify, A. Z., Hamedani, G. G. and Aryal, G. (2016). the Burr X generator of distributions for lifetime data. *Journal of Statistical Theory and Applications*, forthcoming.
34. Yousof, H. M., Rasekhi, M., Afify, A. Z., Alizadeh, M., Ghosh, I. and Hamedani G.G. (2017). The beta Weibull-G family of distributions: theory, characterizations and applications, *Pakistan Journal of Statistics*, forthcoming.