Inference on a Four-parameter Generalized Weighted Exponential Distribution

Ramin Kazemi[a], Akram Kohansal[a], Afshin Fallah[a], Fariba Nasiri

[a]*Department of Statistics, Imam Khomeini International University, Qazvin, Iran*

Abstract. In this paper, we introduce a flexible extension of the generalized weighted exponential distribution which is called four-parameter generalized weighted exponential (FGWE) distribution. Theoretical properties of this model including the mode, median, moment generator function, moments, mean deviations, mean residual life, coefficients of skewness and kurtosis, hazard function, survival function, Bonferroni and Lorenz curves, reliability probability, entropy measures and stochasting ordering are derived and studied in details. We develop the maximum likelihood estimators of the unknown parameters and their corresponding asymptotic confidence intervals. A simulation study and a real world application are also worked out to assess the maximum likelihood estimators and to illustrate the applicability of the proposed distribution in practice.

1. Introduction

Exponential distribution plays an important role in statistical theory of reliability and lifetime analysis. Several extensions to the exponential distribution have been developed by many authors during the last decades. There is an increasing interest in using the weighted exponential distributions in modeling of the skewed positive data especially for lifetime analysis (see [1], [7], [12] and [13]). Some important probability distributions in statistics such as gamma belongs to the family of weighted exponential distributions. Many weighted exponential distribution have been developed so far via adding an extra shape parameter to a given symmetric distribution. Gupta and Kundu [8] and Al-Mutairi et al. [1], by using the idea of Azzalini [2], introduced two new classes of the univariate and bivariate weighted exponential distributions. They discussed distributional and inferential aspects of these families and show that the proposed distributions can be used effectively to analyze skewed positive data. In univariate case, as it has been noted by Gupta and Kundu [8], the univariate weighted exponential distribution, denoted by WE, can be used as an alternative to the gamma or generalized exponential distributions, because it has a probability density function whose shape is very close to the shape of these distributions. Also, this distribution can be obtained as hidden truncation model and may perform better than the well known gamma, log-normal, Weibull or generalized exponential distributions in many situations. As it has been considered in [8], a random variable X is

a weighted exponential (WE) distribution with shape parameter α and scale parameter λ, denoted by WE(α, λ), if the density function of X is given by

$$f_X(x|\alpha, \lambda) = \frac{1+\alpha}{\alpha}(1 - e^{-\lambda\alpha x})\lambda e^{-\lambda x}, \quad x, \alpha, \lambda > 0.$$

WE distribution was generalized to another two-parameter weighted exponential distributions TWE(α, λ) by [12] with the following pdf,

$$f_X(x|\alpha, \lambda) = \frac{(1+\alpha)(1+2\alpha)}{2\alpha^2}(1 - e^{-\lambda\alpha x})^2\lambda e^{-\lambda x}, \quad x, \alpha, \lambda > 0.$$

Shakhatreh [12] investigated the main properties of this distribution. He obtained and studied several statistical properties and statistical inferences. Two real data sets of which one is a right censored data set analyzed, and it showed that in both two cases our model fits much better than WE or some other existing models. Also, TWE distribution was generalized to the three-parameter generalized weighted exponential distributions GWE(α, λ, k) by Kharazmi et al. [10] with the following pdf,

$$f_X(x|\alpha, \lambda, k) = \frac{\alpha}{B(1/\alpha, k+1)}(1 - e^{-\lambda\alpha x})^k\lambda e^{-\lambda x}, \quad x, \alpha, \lambda > 0,$$

where $B(\cdot, \cdot)$ is the beta function and $k \in \mathbb{N} \cup \{0\}$. Kharazmi et al. [10] obtained several statistical and reliability properties of GWE distribution and investigated the estimation and inference procedure for distribution parameters. They showed that the proposed model can provide better fit than the recent class of weighted exponential by using two real data examples. We propose a simple extension of this distribution which is called four-parameter generalized weighted exponential (FGWE) distribution with adding a parameter γ. New properties not provided in [10] will be computed and the previous calculated properties will be extended.

Definition 1.1. *A random variable X is said to have four-parameter generalized weighted exponential distribution FGWE$(\alpha, \lambda, k, \gamma)$ with integer $k \in \mathbb{N} \cup \{0\}$ if the pdf of X is given as following:*

$$f_X(x|\alpha, \lambda, k, \gamma) = \frac{\alpha}{B(\frac{\gamma}{\alpha}, k+1)}(1 - e^{-\lambda\alpha x})^k\lambda e^{-\lambda\gamma x}, \quad x, \alpha, \lambda, \gamma > 0. \tag{1}$$

Theorem 1.2. *Suppose $X \sim$ FGWE$(\alpha, \lambda, k, \gamma)$. Then,*
1) For $k = 0, \gamma = 1$, the random variable X is distributed according to the exponential distribution, Exp(λ).
2) For $k = \gamma = 1$, the random variable X is distributed according to the weighted exponential distribution, WE(α, λ), introduced in [8].
3) For $\gamma = 1$, the random variable X is distributed according to the exponential distribution, GWE(α, λ, k).
4) For $\alpha = k = 1, \gamma = 1$, the random variable X is distributed according to the generalized exponential distribution, GE$(\lambda, 2)$, introduced in [6].
(5) The random variable X converges in distribution to the exponential distribution, Exp(λ), as $\alpha \longrightarrow +\infty$ and $\gamma = 1$.
6) The random variable $Y = cX$ for $c > 0$, is distributed according to FGWE$(\alpha, \lambda/c, k, \gamma)$.
7) The random variable $Y = e^{-\alpha\lambda X}$ with $\alpha, \lambda > 0$, has a weighted beta distribution, wbeta$\left(\frac{\gamma}{\alpha}, k+1\right)$.
8) The random variable $Y = 1 - e^{-\alpha\lambda X}$ with $\alpha, \lambda > 0$, has a weighted beta distribution, wbeta$\left(k+1, \frac{\gamma}{\alpha}\right)$.

Proof. Obvious. $\square$

The density of the FGWE distribution for different values of the parameters k, γ, α and λ, is plotted in Figure 1, indicates that the FGWE distribution can be a suitable candidate for a vast class of positively skewed data.

The rest of the paper is as follows. In Section 2, we give the mode, moment generator function, mean, variance, coefficients of skewness and kurtosis and higher moments of FGWE distribution. In Section 3, we compute the reliability probability for difeerent values of k and γ. In Section 4, we give the Shannon

and Reyni entropies. In Sections 5 and 6, we give the mean deviatins and mean residual life of FGWE distribution, respectively. Bonferroni and Lorenz curves are given in Section 7 and comparing WE and FGWE distributions with respet to stochastic ordering information are given in Section 8. Kharazmi et al. [10] claimed that due to the non-linearity of equations for the partial derivatives (with respect to α and λ), the MLEs of parameters can be obtained numerically for GWE distribution. But here and in Section 9, the maximum likelihood estimators of the all unknown parameters and their corresponding asymptotic confidence sets are developed. A simulation study and a real world application are also worked out to assess the maximum likelihood estimators and to illustrate the applicability of the proposed distribution, in Section 10. The conclusion of the paper is given in Section 11.
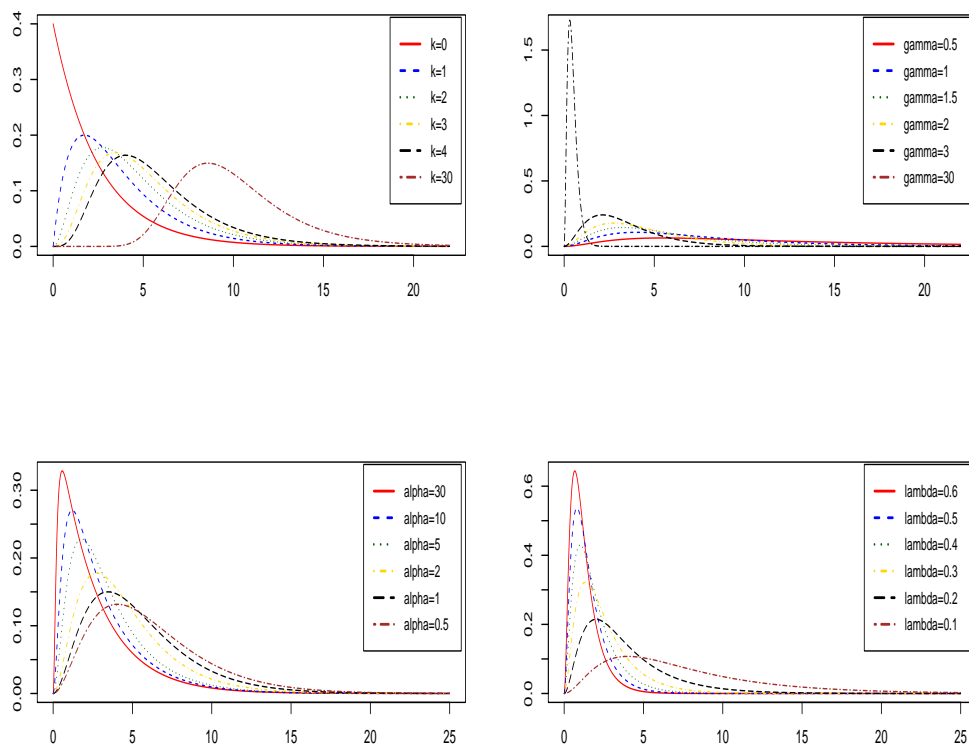

Figure 1: The density plot of the FGWE distribution for different values of the parameters k, α, λ and γ.

2. Mode, median and moments

Notice that

$$\frac{d}{dx} \log f_X(x|\alpha, \lambda, k, \gamma) = -\lambda\gamma + \frac{k\alpha\lambda e^{-\alpha\lambda x}}{1 - e^{-\alpha\lambda x}},$$
$$\frac{d^2}{dx^2} \log f_X(x|\alpha, \lambda, k, \gamma) = \frac{-k\alpha^2\lambda^2 e^{-\alpha\lambda x}}{(1 - e^{-\alpha\lambda x})^2} < 0,$$

thus the density function (1) is a log-concave function of x for all $x > 0$. Therefore, the FGWE distribution is strongly unimodal and it's unique mode is located at

$$x = \frac{1}{\lambda\alpha}\log\left(\frac{k}{\gamma}\alpha + 1\right).$$

We have

$$(1 - e^{-\lambda\alpha x})^k = \sum_{i=0}^{k}\binom{k}{i}(-1)^i e^{-i\lambda\alpha x}.$$

Thus we can show the probability density function of X as follows:

$$f_X(x|\alpha, \lambda, k, \gamma) = \frac{\alpha}{B(\frac{\gamma}{\alpha}, k+1)}\lambda\sum_{i=0}^{k}\binom{k}{i}(-1)^i e^{-\lambda(\alpha i+\gamma)x}.$$

The cumulative distribution function and the hazard function of the FGWE$(\alpha, \lambda, k, \gamma)$ are given, respectively, by

$$F_X(x|\alpha, \lambda, k, \gamma) = \frac{\alpha}{B(\frac{\gamma}{\alpha}, k+1)}\sum_{i=0}^{k}\binom{k}{i}\frac{(-1)^i}{\alpha i + \gamma}[1 - e^{-\lambda(\alpha i+\gamma)x}] \tag{2}$$

and

$$H_X(x|\alpha, \lambda, k) = \frac{\frac{\alpha}{B(\frac{\gamma}{\alpha}, k+1)}\lambda e^{-\lambda\gamma x}(1 - e^{-\lambda\alpha x})^k}{1 - \frac{\alpha}{B(\frac{\gamma}{\alpha}, k+1)}\sum_{i=0}^{k}\binom{k}{i}\frac{(-1)^i}{\alpha i+\gamma}[1 - e^{-\lambda(\alpha i+\gamma)x}]},$$

for $x > 0$ and 0 otherwise. There is no explicit solution for the equation $F_X(m|\alpha, \lambda, k, \gamma) = 1/2$, so the median m can not be given explicitly.

Corollary 2.1. *Suppose that $X \sim FGWE(\alpha, \lambda, k, \gamma)$. If we use the conditional notations $f_Z(\cdot|\theta)$ and $F_Z(\cdot|\theta)$ to show the pdf and cdf of an exponential random variable Z with mean equal to $1/\theta$, then*

$$f_X(x|\alpha, \lambda, k, \gamma) = \frac{\alpha}{\gamma B(\frac{\gamma}{\alpha}, k+1)}f_Z(x|\lambda\gamma)(F_Z(x|\alpha\lambda))^k,$$

$$F_X(x|\alpha, \lambda, k, \gamma) = \frac{\alpha}{B(\frac{\gamma}{\alpha}, k+1)}\sum_{i=0}^{k}\binom{k}{i}\frac{(-1)^i}{\alpha i + \gamma}F_Z(x|\lambda(\alpha i + \gamma))$$

In the next theorem we derive the moment generator function and the first two moments of the FGWE distribution.

Theorem 2.2. *Suppose $X \sim FGWE(\alpha, \lambda, k, \gamma)$. Then*
i) The moment generator function (mgf) of the random variable X is

$$M_X(t) = \frac{\Gamma\left(1 + k + \frac{\gamma}{\alpha}\right)\Gamma\left(\frac{\lambda\gamma - t}{\alpha\lambda}\right)}{\Gamma\left(\frac{\gamma}{\alpha}\right)\Gamma\left(k + \frac{(\gamma+\alpha)\lambda - t}{\lambda\alpha}\right)}, \quad t - \lambda\gamma < 0 \tag{3}$$

where $\Gamma(\cdot)$ is the gamma function.
ii) The first two moments are given, respectively, by

$$\mathbb{E}(X) = \frac{1}{\lambda\alpha}\left(\psi\left(1 + k + \frac{\gamma}{\alpha}\right) - \psi\left(\frac{\gamma}{\alpha}\right)\right), \tag{4}$$

where $\psi(\cdot)$ is the digamma function and

$$\mathbb{V}ar(X) = \frac{1}{\lambda^2\alpha^2}\left(\psi'\left(\frac{\gamma}{\alpha}\right) - \psi'\left(1 + k + \frac{\gamma}{\alpha}\right)\right).$$

Proof. We have

$$\int_0^\infty (1 - e^{-\lambda \alpha x})^k e^{-(\lambda \gamma - t)x} dx = \frac{B(\gamma/\alpha - t/(\lambda \alpha), k+1)}{\lambda \alpha}.$$

Due to the definition of the moment generator function and this fact that $\frac{d^n M_X(t)}{dt^n}\Big|_{t=0} = \mathbb{E}(X^n)$ proof is completed. $\square$

Since the exponential distribution ($k = 0, \gamma = 1$) is an especial case of the FGWE distribution, the main distributional characteristics of the GWE distribution, i.e., the probability density function (pdf), cdf, mgf, moments, etc., can be formulated as an explicit expression of their corresponding counterparts in the exponential distribution. For example, since

$$\psi^{(m)}(z+1) = \psi^{(m)}(z) + \frac{(-1)^m m!}{z^{m+1}}, \quad m \geq 1$$

we have

$$\mathbb{E}(X) = \frac{1}{\lambda}, \quad \mathbb{V}ar(X) = \frac{1}{\lambda^2}.$$

With simple but long calculations,

$$
\begin{aligned}
\mathbb{E}(X^3) &= \frac{1}{\lambda^3 \alpha^3}\Big(\psi(1+k+\gamma/\alpha)^3 - 3\psi(1+k+\gamma/\alpha)^2 \psi(\gamma/\alpha) - \psi(\gamma/\alpha)^3 \\
&+ 3\psi(\gamma/\alpha)(\psi'(1+k+\gamma/\alpha) - \psi'(\gamma/\alpha)) \\
&+ 3\psi(1+k+\gamma/\alpha)(\psi(\gamma/\alpha)^2 - \psi'(1+k+\gamma/\alpha) + \psi'(\gamma/\alpha)) \\
&+ \psi''(1+k+\gamma/\alpha) - \psi''(\gamma/\alpha)\Big), \\
\mathbb{E}(X^4) &= \frac{1}{\lambda^4 \alpha^4}\Big(\psi(1+k+\gamma/\alpha)^4 - 4\psi(1+k+\gamma/\alpha)^3 \psi(\gamma/\alpha) \\
&+ \psi(\gamma/\alpha)^4 + 3\psi'(1+k+\gamma/\alpha)^2 - 6\psi(\gamma/\alpha)^2(\psi'(1+k+\gamma/\alpha) - \psi'(\gamma/\alpha)) \\
&- 6\psi'(1+k+\gamma/\alpha)\psi'(\gamma/\alpha) + 3\psi'(\gamma/\alpha)^2 \\
&+ 6\psi(1+k+\gamma/\alpha)^2(\psi(\gamma/\alpha)^2 - \psi'(1+k+\gamma/\alpha) + \psi'(\gamma/\alpha)) \\
&- 4\psi(\gamma/\alpha)(\psi''(1+k+\gamma/\alpha) - \psi''(\gamma/\alpha)) \\
&- 4\psi(1+k+\gamma/\alpha)(\psi(\gamma/\alpha)^3 - 3\psi(\gamma/\alpha)(\psi'(1+k+\gamma/\alpha) - \psi'(\gamma/\alpha)) \\
&- \psi''(1+k+\gamma/\alpha) + \psi''(\gamma/\alpha) - \psi^{(3)}(1+k+\gamma/\alpha) + \psi^{(3)}(\gamma/\alpha)\Big).
\end{aligned}
$$

Using the above results, the measures of skewness and kurtosis of the FGWE distribution can be obtained as

$$
\begin{aligned}
Skewness(X) &= \frac{\mathbb{E}(X^3) - 3\mathbb{E}(X^2)\mathbb{E}(X) + 2\mathbb{E}^3(X)}{\mathbb{V}ar(X)^{3/2}}, \\
Kurtosis(X) &= \frac{\mathbb{E}(X^4) - 4\mathbb{E}(X^3)\mathbb{E}(X) + 6\mathbb{E}(X^3)\mathbb{E}^2(X) - 3\mathbb{E}^4(X)}{\mathbb{V}ar(X)},
\end{aligned}
$$

respectively. Since

$$\int_0^\infty x^n e^{-\lambda(\alpha i + \gamma)x} dx = \frac{\Gamma(n+1)}{((\gamma + \alpha i)\lambda)^{n+1}}, \quad n \geq 1$$

we have

$$\mu'_n = \mathbb{E}(X^n) = \frac{\alpha \lambda \Gamma(n+1)}{B(\frac{\gamma}{\alpha}, k+1)} \sum_{i=0}^k \binom{k}{i} \frac{(-1)^i}{((\gamma + \alpha i)\lambda)^{n+1}}.$$

We can also find the n-th central moment of the FGWE distribution as ($\mu = \mu'_1 = \mathbb{E}(X)$):

$$\mathbb{E}(X - \mu)^n = \sum_{j=0}^{n} \binom{n}{j} \mu'_j (-\mu)^{n-j}.$$

3. Reliability probability

In the following theorem we compute the probability that one of two independently distributed FGWE random variables exceeds another. In the context of the reliability, this probability is in core of interest and is known as the reliability probability, $R = P(X > Y)$.

Theorem 3.1. *Suppose that two independent random variables X and Y are distributed according to the distributions FGWE($\alpha, \lambda, k_1, \gamma$) and FGWE($\alpha, \lambda, k_2, \gamma$), respectively. Then the reliability probability is given by*

$$R = P(X > Y) = \frac{\alpha \sum_{i=0}^{k_1} \binom{k_1}{i} \frac{(-1)^i}{\alpha i + \gamma} B(\frac{2\gamma + \alpha i}{\alpha}, k_2 + 1)}{B(\frac{\gamma}{\alpha}, k_1 + 1) B(\frac{\gamma}{\alpha}, k_2 + 1)}.$$

Proof. Set $C(\alpha, k, \gamma) := \frac{1}{\alpha} B(\frac{\gamma}{\alpha}, k + 1)$. Thus

$$\sum_{i=0}^{k} \binom{k}{i} \frac{(-1)^i}{\alpha i + \gamma} = \frac{B(\frac{\gamma}{\alpha}, k + 1)}{\alpha} = C(\alpha, k, \gamma). \tag{5}$$

Consider the cdf of the FGWE distribution given in (2). Then

$$\begin{aligned}
P(X > Y) &= 1 - \int_0^{+\infty} F_X(y) f_Y(y) dy \\
&= 1 - \int_0^{+\infty} (C(\alpha, k_1, \gamma))^{-1} \sum_{i=0}^{k_1} \binom{k_1}{i} \frac{(-1)^i}{\alpha i + k_2} \left(1 - e^{-\lambda(\alpha i + k_2)y}\right) f_Y(y) dy \\
&= 1 - (C(\alpha, k_1, \gamma))^{-1} \sum_{i=0}^{k_1} \binom{k_1}{i} \frac{(-1)^i}{\alpha i + \gamma} \int_0^{+\infty} f_Y(y) dy \\
&+ (C(\alpha, k_1, \gamma))^{-1} \sum_{i=0}^{k_1} \binom{k_1}{i} \frac{(-1)^i}{\alpha i + \gamma} \int_0^{+\infty} e^{-\lambda(\alpha i + \gamma)y} (C(\alpha, k_2, \gamma))^{-1} e^{-\lambda \gamma y} (1 - e^{-\lambda \alpha y})^{k_2} \\
&= (C(\alpha, k_1, \gamma))^{-1} (C(\alpha, k_2, \gamma))^{-1} \sum_{i=0}^{k_1} \binom{k_1}{i} \frac{(-1)^i}{\alpha i + \gamma} C(\alpha, k_2, 2\gamma + \alpha i).
\end{aligned}$$

□

For the special case of $k_1 = k_2 = 0$ and $\gamma = 1$ where X and Y are independent random variables with exponential distribution, Exp(λ), we have $P(X > Y) = 0.5$.

Theorem 3.2. *Suppose two random variables X and Y are distributed according to the distributions FGWE($\alpha, \lambda, k, \gamma_1$) and FGWE($\alpha, \lambda, k, \gamma_2$), respectively. Then the reliability probability is given by*

$$R = P(X > Y) = \frac{\alpha \sum_{i=0}^{k} \binom{k}{i} \frac{(-1)^i}{\alpha i + \gamma_2} B(\frac{2\gamma_2 + \alpha i}{\alpha}, k + 1)}{B(\frac{\gamma_1}{\alpha}, k + 1) B(\frac{\gamma_2}{\alpha}, k + 1)}.$$

Proof. The proof is quite similar to the proof of Theorem 3.1. □

Theorem 3.3. *Suppose $D_{KL}(X, Y)$ denotes the Kullback-Leibler (KL) divergence between two random variables $X \sim$ GWE(α, λ, k) and $Y \sim$ FGWE$(\alpha, \lambda, k, \gamma)$. Then*

$$D_{KL}(X, Y) = \log \frac{B(\frac{1}{\alpha}, k+1)}{B(\frac{\gamma}{\alpha}, k+1)} + \frac{\gamma - 1}{\alpha} \Big(\psi\big(\frac{\gamma}{\alpha}\big) - \psi\big(1 + k + \frac{\gamma}{\alpha}\big) \Big).$$

Proof. We have

$$
\begin{aligned}
D_{KL}(X, Y) &= \int_0^{+\infty} f_Y(x|\alpha, \lambda, k) \log \frac{f_Y(x|\alpha, \lambda, k, \gamma)}{f_X(x|\alpha, \lambda, k)} \, dx \\
&= \int_0^{+\infty} f_Y(x|\alpha, \lambda, k) \log \left(\frac{B(\frac{1}{\alpha}, k+1)}{B(\frac{\gamma}{\alpha}, k+1)} e^{-\lambda(\gamma - 1)x} \right) dx \\
&= \log \frac{B(\frac{1}{\alpha}, k+1)}{B(\frac{\gamma}{\alpha}, k+1)} - \lambda(\gamma - 1) \mathbb{E}_{FGWE}(X),
\end{aligned}
$$

where $\mathbb{E}_{FGWE}(\cdot)$ denotes the expectation with respect to the density $f_Y(x|\alpha, \lambda, k, \gamma)$. The proof completes due to the following fact:

$$\mathbb{E}_{FGWE}(X) = \frac{1}{\lambda \alpha} \Big(\psi\big(1 + k + \frac{\gamma}{\alpha}\big) - \psi\big(\frac{\gamma}{\alpha}\big) \Big).$$

$\square$

We compare GWE and FGWE distributions via Kullback-Leibler divergence in Figure 2.
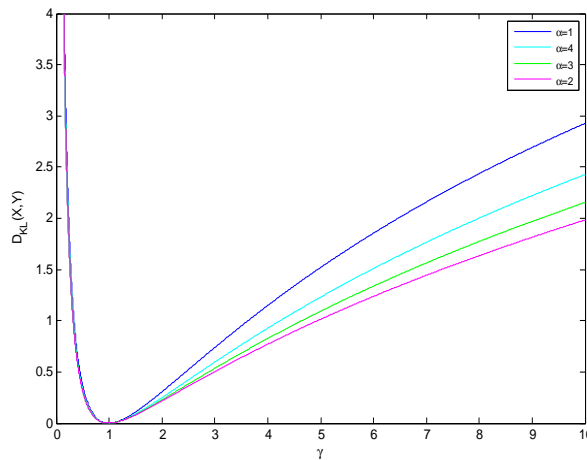
Figure 2: Kullback-Leibler divergence between GWE and FGWE for some selected parameter values of α and γ with $k = 2$.

4. Entropy measures

Shannon entropy is a central concept of information theory for expressing the uncertainty about a random variable. Renyi defined a generalization of Shannon entropy which depends on a parameter v. Renyi entropy defined by

$$H_v(f(x)) = \frac{1}{1 - v} \log \Big(\int f^v(x) dx \Big),$$

where $v > 0$ and $v \neq 1$. Renyi entropy tends to Shannon entropy as $v \to 1$. We have

$$\int_0^\infty (1 - e^{-\lambda \alpha x})^{kv} e^{-\lambda v \gamma x} dx = \frac{B((\gamma v)/\alpha, kv + 1)}{\alpha \lambda}.$$

Thus if $X \sim FGWE(\alpha, \lambda, k, \gamma)$, then

$$H_v(f(x)) = \frac{1}{1-v} \log \left(\frac{(\alpha \lambda)^{v-1} B((\gamma v)/\alpha, kv + 1)}{(B(\frac{\gamma}{\alpha}, k + 1))^v} \right).$$

Also, as $v \to 1$, the Shannon entropy is given by

$$H_1(f(x)) = -\log \left(\frac{\alpha \lambda}{B(\gamma/\alpha, k + 1)} \right) - k\psi(1 + k) + (\gamma/\alpha + k)\psi(1 + k + \gamma/\alpha) - \frac{\gamma}{\alpha}\psi(\gamma/\alpha).$$

For the exponential distribution ($k = 0$, $\gamma = 1$), since $\psi(z + 1) = \psi(z) + \frac{1}{z}$, we have

$$H_1(f(x)) = 1 - \log \lambda.$$

5. Mean deviations

The mean deviations can be used as a measure of spread in a population. The mean deviations about the mean and about the median are given by

$$\delta_1(X) = \int_0^\infty |x - \mu| f(x) dx$$

and

$$\delta_2(X) = \int_0^\infty |x - m| f(x) dx,$$

respectively, where $\mu = \mathbb{E}(X)$ and $m = median(X)$. These quantities can be calculated as

$$\delta_1(X) = 2\mu F_X(\mu) - 2 \int_0^\mu x f(x) dx$$

and

$$\delta_2(X) = \mu - 2 \int_0^m x f(x) dx.$$

Since

$$\int_0^t x e^{-\lambda(\alpha i + \gamma)x} dx = \frac{e^{-\lambda t(\alpha i + \gamma)}(e^{\lambda t(\alpha i + \gamma)} - \lambda t(\alpha i + \gamma) - 1)}{\lambda^2(\alpha i + \gamma)^2},$$

we have

$$\delta_1(X) = 2\mu F_X(\mu) - \frac{2\alpha \lambda}{B(\frac{\gamma}{\alpha}, k + 1)} \sum_{i=0}^k \binom{k}{i} \frac{(-1)^i e^{-\lambda \mu(\alpha i + \gamma)}(e^{\lambda \mu(\alpha i + \gamma)} - \lambda \mu(\alpha i + \gamma) - 1)}{\lambda^2(\alpha i + \gamma)^2}$$

and

$$\delta_2(X) = \mu - \frac{2\alpha \lambda}{B(\frac{\gamma}{\alpha}, k + 1)} \sum_{i=0}^k \binom{k}{i} \frac{(-1)^i e^{-\lambda m(\alpha i + \gamma)}(e^{\lambda m(\alpha i + \gamma)} - \lambda m(\alpha i + \gamma) - 1)}{\lambda^2(\alpha i + \gamma)^2}.$$

6. Mean residual life

In life testing situations, the expected additional lifetime given that a component has survived until time x is a function of x, called the mean residual life. More specifically, if the random variable X represents the life of a component, then the mean residual life is given by $m(x) = \mathbb{E}(X - x | X > x)$ and can be expressed as

$$m(x) = \frac{1}{\bar{F}(x)} \int_x^\infty \bar{F}(u)du,$$

where $\bar{F} = 1 - F$. From (5),

$$F_X(x|\alpha, \lambda, k, \gamma) = 1 - \frac{\alpha}{B(\frac{\gamma}{\alpha}, k+1)} \sum_{i=0}^k \binom{k}{i} \frac{(-1)^i}{\alpha i + \gamma} e^{-\lambda(\alpha i + \gamma)x}.$$

Hence,

$$m(x) = \frac{\sum_{i=0}^k \binom{k}{i} \frac{(-1)^i}{(\alpha i + \gamma)^2 \lambda} e^{-\lambda(\alpha i + \gamma)x}}{\sum_{i=0}^k \binom{k}{i} \frac{(-1)^i}{\alpha i + \gamma} e^{-\lambda(\alpha i + \gamma)x}}.$$

7. Bonferroni and Lorenz curves

We can construct Bonferroni and Lorenz curves, which are important in several fields such as economics, reliability, demography, insurance and medicine. They are defined as the following:

$$B(F(x)) = \frac{1}{\mu F(x)} \int_0^x t f(t) dt$$

and

$$L(F(x)) = \frac{1}{\mu} \int_0^x t f(t) dt.$$

If $X \sim FGWE(\alpha, \lambda, k, \gamma)$, then

$$B(F(x)) = \frac{\alpha \lambda}{\mu F(x) B(\frac{\gamma}{\alpha}, k+1)} \sum_{i=0}^k \binom{k}{i} \frac{(-1)^i e^{-\lambda x(\alpha i + \gamma)} (e^{\lambda x(\alpha i + \gamma)} - \lambda x(\alpha i + \gamma) - 1)}{\lambda^2 (\alpha i + \gamma)^2}$$

and

$$L(F(x)) = \frac{\alpha \lambda}{\mu B(\frac{\gamma}{\alpha}, k+1)} \sum_{i=0}^k \binom{k}{i} \frac{(-1)^i e^{-\lambda x(\alpha i + \gamma)} (e^{\lambda x(\alpha i + \gamma)} - \lambda x(\alpha i + \gamma) - 1)}{\lambda^2 (\alpha i + \gamma)^2}.$$

Figure 3 shows Bonferroni and Lorenz curve plots for some selected paparmeter values.
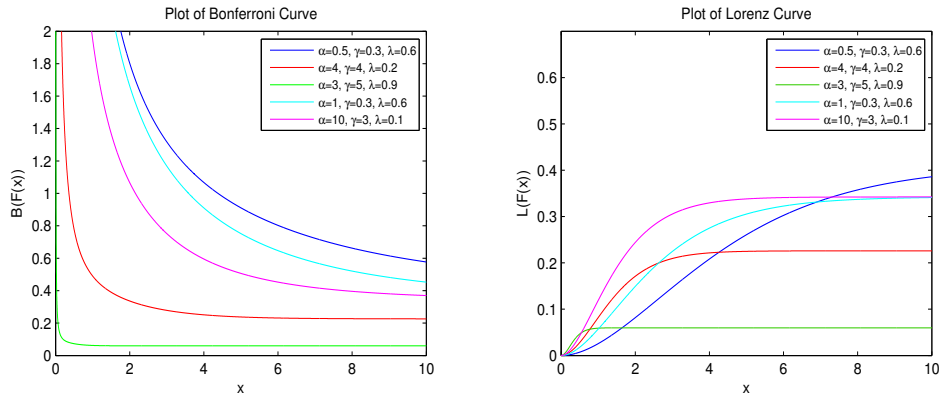
Figure 3: Bonferroni and Lorenz curve plots for some selected paparmeter values of α, γ and λ with $k = 2$.

8. Stochastic ordering

In this section, we are interested in comparing WE and FGWE distributions with respect to stochastic ordering information. The comparing two random variables are very important in reliability theory, risk analysis, and other disciplines. There are many possibilities to compare random variables or their distributions, respectively, with each other. One of the most important orderings among stochastic orderings is the likelihood ratio ordering which compares lifetimes of systems with respect to their likelihood information. In this section, we give a basic theorem for comparing $FGWE(\alpha, \lambda, k, \gamma)$ and $WE(\alpha k, \lambda \gamma)$ distributions. Now, let us to give a quick review of required definitions of stochastic orders and notation which are used in the following theorem. Let X and Y be two random variables with distribution functions F and G, survival functions $\bar{F} = 1 - F$ and $\bar{G} = 1 - G$ and density functions f and g. It is said that X is smaller than Y in the following expression:

1. Likelihood ratio order ($X \leq_{lr} Y$) if $g(x)/f(x)$ is increasing in x.
2. Hazard rate order ($X \leq_{hr} Y$) if $\bar{G}(x)/\bar{F}(x)$ is increasing in x.
3. Usual stochastic order ($X \leq_{st} Y$) if $\bar{F}(x) \leq \bar{G}(x)$.
4. Mean residual life order ($X \leq_{mrl} Y$) if $\mathbb{E}(X - x | X > x) \leq \mathbb{E}(Y - x | Y > x)$.

The following implications hold among these stochastic orders:

$$Y \leq_{lr} X \implies Y \leq_{hr} X \implies Y \leq_{st(mrl)} X.$$

Theorem 8.1. *Suppose $Y \sim FGWE(\alpha, \lambda, k, \gamma)$ and $X \sim WE(\alpha k, \lambda \gamma)$. Then $X \leq_{lr} Y$.*

Proof. It is sufficient to see that the ratio $f_Y(x)/f_X(x)$ is increasing in x. We have

$$\frac{d}{dx}\left(\frac{f_Y(x)}{f_X(x)}\right) = \frac{e^{k\alpha\lambda x}(1 - e^{-\alpha\lambda x})^k(-e^{\alpha\lambda x} + e^{k\alpha\lambda x})k\alpha\lambda}{(-1 + e^{\alpha\lambda x})(-1 + e^{k\alpha\lambda x})^2} \geq 0.$$

$\square$

9. Maximum likelihood estimation

In this section we consider the maximum likelihood (ML) estimation of parameters of the FGWE distribution. Consider a random sample $X_1, \ldots, X_n$ from the $FGWE(\alpha, \lambda, k, \gamma)$ distribution. Then, the likelihood function is given by

$$L(\alpha, \lambda, k, \gamma | x_1, \ldots, x_n) = \left(\frac{\alpha\lambda}{B(\frac{\gamma}{\alpha}, k+1)}\right)^n \left\{\prod_{i=1}^{n}(1 - e^{-\lambda\alpha x_i})\right\}^k e^{-\lambda\gamma \sum_{i=1}^{n} x_i}. \tag{6}$$

We first explain the maximum likelihood estimation of parameter k assuming the values of λ, γ and α are known. This parameter plays an important role as different values of k are correspond to the different families of the distributions. Maximizing the likelihood function with respect to k can not be done by differentiating the likelihood function because of the factorials and because k must be an integer. For known values of the parameters λ, γ and α, the profile likelihood function of the parameter k is given by

$$L(k) = L(k|\mathbf{x}) \propto \left(\frac{1}{B(\frac{\gamma}{\alpha}, k+1)} \right)^n \left\{ \prod_{i=1}^{n} (1 - e^{-\lambda \alpha x_i}) \right\}^k.$$

Suppose $\hat{k}$ denotes the maximum likelihood estimator of the parameter k. Then, it must simultaneously satisfies the following two inequalities

$$L(\hat{k}) \geq L(\hat{k} - 1) \tag{7}$$
$$L(\hat{k}) > L(\hat{k} + 1). \tag{8}$$

From the equation (7), we have

$$\frac{L(\hat{k})}{L(\hat{k} - 1)} \geq 1 \Leftrightarrow \left\{ \frac{B(\frac{\gamma}{\alpha}, \hat{k})}{B(\frac{\gamma}{\alpha}, \hat{k} + 1)} \right\}^n T(\mathbf{x}; \alpha, \lambda) > 1,$$

where $T(\mathbf{x}; \alpha, \lambda) = \prod_{i=1}^{n} (1 - e^{-\lambda \alpha x_i})$. Since $T(\mathbf{x}; \alpha, \lambda)$ is a positive quantity, we have

$$\frac{L(\hat{k})}{L(\hat{k} - 1)} \geq 1 \quad \Leftrightarrow \quad \frac{B(\frac{\gamma}{\alpha}, \hat{k})}{B(\frac{\gamma}{\alpha}, \hat{k} + 1)} \geq (T(\mathbf{x}; \alpha, \lambda))^{-\frac{1}{n}}$$

$$\Leftrightarrow \quad 1 + \frac{\gamma}{\alpha \hat{k}} \geq (T(\mathbf{x}; \alpha, \lambda))^{-\frac{1}{n}}.$$

Therefore the inequality (7) holds if and only if

$$\hat{k} \leq \gamma \left\{ \alpha \left(T(\mathbf{x}; \alpha, \lambda)^{-\frac{1}{n}} - 1 \right) \right\}^{-1}. \tag{9}$$

Similarly, it can be shown that the inequality (8) holds if and only if

$$\hat{k} > \gamma \left\{ \alpha \left(T(\mathbf{x}; \alpha, \lambda)^{-\frac{1}{n}} - 1 \right) \right\}^{-1} - 1. \tag{10}$$

Combining the inequalities (9) and (10) it is resulted that $\hat{k}$ must satisfies the inequality

$$\hat{k} \leq \gamma \left\{ \alpha \left(T(\mathbf{x}; \alpha, \lambda)^{-\frac{1}{n}} - 1 \right) \right\}^{-1} < \hat{k} + 1.$$

Therefore the explicit form of the maximum likelihood estimator of parameter k is obtain to be

$$\hat{k} = \left[\frac{\gamma}{\alpha \left(T(\mathbf{x}; \alpha, \lambda)^{-\frac{1}{n}} - 1 \right)} \right], \tag{11}$$

where $[b]$ denotes the largest integer less than or equal to b. It can be seen that, despite of the work of Feldman and Fox [5], for the success probability in binomial distribution, the maximum likelihood estimator of parameter k has a closed form.

Now, we can use the profile likelihood of the parameters $(\alpha, \lambda, \gamma)$ to find their maximum likelihood estimators. According to the equation (6), the profile log-likelihood function of α, λ and γ, assuming k are known, is obtained to be

$$\ell(\alpha, \lambda, \gamma) = -n \log C(\alpha, k, \gamma) + n \log \lambda - \lambda \gamma \sum_{i=1}^{n} x_i + k \sum_{i=1}^{n} \log(1 - \exp\{-\lambda \alpha x_i\}). \tag{12}$$

We also note that by definition of $C(\alpha, k, \gamma)$ in relation (5),

$$C_\alpha^{(n)}(\alpha, \lambda, \gamma) := \frac{\partial^n C(\alpha, k, \gamma)}{\partial \alpha^n} = n! \sum_{i=0}^{k} \binom{k}{i} \frac{i^n (-1)^{n+i}}{(\alpha i + \gamma)^{n+1}}$$

and

$$C_\gamma^{(n)}(\alpha, \lambda, \gamma) := \frac{\partial^n C(\alpha, k, \gamma)}{\partial \gamma^n} = n! \sum_{i=0}^{k} \binom{k}{i} \frac{(-1)^{n+i}}{(\alpha i + \gamma)^{n+1}}, \quad n = 1, 2, \dots .$$

Therefore equating the first order derivations of (12) with respect to α, λ and γ to zero leads to the following system of equations:

$$\frac{\partial \ell(\alpha, \lambda, \gamma)}{\partial \alpha} = -n \frac{C_\alpha^{(1)}(\alpha, \lambda, \gamma)}{C(\alpha, \lambda, \gamma)} + k\lambda \sum_{i=1}^{n} \frac{x_i e^{-\lambda \alpha x_i}}{1 - e^{-\lambda \alpha x_i}} = 0$$

$$\frac{\partial \ell(\alpha, \lambda, \gamma)}{\partial \lambda} = \frac{n}{\lambda} - \gamma \sum_{i=1}^{n} x_i + k\alpha \sum_{i=1}^{n} \frac{x_i e^{-\lambda \alpha x_i}}{1 - e^{-\lambda \alpha x_i}} = 0,$$

$$\frac{\partial \ell(\alpha, \lambda, \gamma)}{\partial \gamma} = -n \frac{C_\gamma^{(1)}(\alpha, \lambda, \gamma)}{C(\alpha, \lambda, \gamma)} - \lambda \sum_{i=1}^{n} x_i = 0. \tag{13}$$

One can use an iterative maximization procedure to find the solutions of equations (13).

9.1. Confidence interval for interested parameters

We are interesting to drive the exact distribution of $\hat{k}$. It should be noted that it is not possible to use the asymptotic normality of the ML estimators for $\hat{k}$ because the log-likelihood function is not differentiable with respect to the parameter k and due to discreteness of the distribution of this estimator. According to the equation (11), $\hat{k}$ is a nonnegative integer random variable. One can write

$$P(\hat{k} = b) = P\left(\left[\frac{\gamma}{\alpha\left(T(\mathbf{x}; \alpha, \lambda)^{-\frac{1}{n}} - 1\right)}\right] = b\right)$$

$$= P\left(\left(\frac{\alpha b}{\alpha b + \gamma}\right)^n \leq T(\mathbf{x}; \alpha, \lambda) < \left(\frac{\alpha(b+1)}{\alpha(b+1) + \gamma}\right)^n\right),$$

where $T(\mathbf{x}; \alpha, \lambda) = \prod_{i=1}^{n}(1 - e^{-\lambda \alpha x_i})$. Recalling the part (8) of Theorem 1.2, when α and λ are known, the statistics $T(\mathbf{x}; \alpha, \lambda)$ is a product of n independent weighted beta random variables. The product of the independent beta random variables belongs to a family whose densities are particular Meijier G-functions. This distribution is too complex to be used for inferential proposes. Therefore, we use the central limit theorem to derive the asymptotic distribution of the logarithm of this random variable. We have

$$P(\hat{k} = b) = P\left(\log\left(\frac{\alpha b}{\alpha b + \gamma}\right) \leq \frac{1}{n} \sum_{i=1}^{n} \log(Y_i) < \log\left(\frac{\alpha(b+1)}{\alpha(b+1) + \gamma}\right)\right)$$

$$= P\left(\log\left(\frac{\alpha b}{\alpha b + \gamma}\right) \leq \bar{W} < \log\left(\frac{\alpha(b+1)}{\alpha(b+1) + \gamma}\right)\right), \tag{14}$$

where $Y_i = 1 - e^{-\lambda \alpha X_i}$ has the weighted beta distribution ($wbeta(k+1, \frac{\gamma}{\alpha})$), $\bar{W} = \frac{1}{n} \sum_{i=1}^{n} W_i$ and $W_i = \log(Y_i)$. It is well known that the mean and the variance of the natural logarithm of a beta random variable can be expressed in terms of the digamma and trigamma functions. If $Z \sim beta(\alpha, \beta)$, then $\mathbb{E}(\log(Z)) = \psi(\alpha) - \psi(\alpha + \beta)$ and $\mathbb{E}(\log(Z))^2 = (\psi(\alpha) - \psi(\alpha + \beta))^2 + \psi_1(\alpha) - \psi_1(\alpha + \beta)$, where $\psi_1(\cdot)$ denote the trigamma functions. Thus, the mean and variance of the random variable $W_i = \log(Y_i)$ is obtained to be

$$\mu_w = \mathbb{E}(W_i) = \mathbb{E}(\log(Y_i)) = \frac{1}{\alpha} \{\psi(k+1) - \psi(k + 1 + \gamma/\alpha)\}$$

and

$$
\begin{aligned}
\sigma_w^2 &= \mathbb{V}ar(W_i) = \mathbb{E}(\log(Y_i))^2 - \mathbb{E}^2(\log(Y_i)) \\
&= \frac{1}{\alpha}\left\{\mathbb{E}(\log(Z))^2 - \mathbb{E}^2(\log(Z))\right\} \\
&= \frac{1}{\alpha}\left\{\psi_1(k+1) - \psi_1(k+1+\gamma/\alpha)\right\}.
\end{aligned}
$$

For enough large sample sizes the central limit theorem implies that $\bar{W} \sim N(\mu_w, \frac{\sigma_w^2}{n})$. Therefore due to the equation (14) the sampling distribution of the $\hat{k}$, in the form of a probability mass function, is obtained to be

$$
P(\hat{k} = b) \approx \Phi\left(\frac{\sqrt{n}(\log(\frac{\alpha(b+1)}{\alpha(b+1)+\gamma}) - \mu_w)}{\sigma_w}\right) - \Phi\left(\frac{\sqrt{n}(\log(\frac{\alpha b}{\alpha b + \gamma}) - \mu_w)}{\sigma_w}\right),
$$

where $b = 0, 1, 2, \ldots$ and $\Phi(\cdot)$ denotes the cdf of the standard normal distribution. Once the parameters k, γ and α are estimated, the quantiles of this distribution can be used to compute the moments of $\hat{k}$, to construct the confidence intervals and to test the hypothesis concerning k. The moments of a given real valued function of $\hat{k}$, $g(\hat{k})$, is obtain to be $\mathbb{E}(g(\hat{k})) = \sum_{b=1}^{\infty} g(b)P(\hat{k} = b)$. The sampling distribution of the maximum likelihood estimator of parameter k for $n = 100$, $k = 2$ and different values of parameters α, γ is plotted in Figure 4. It can be seen that this distribution is a right skewed discrete distribution that assigns small probability to the large values. Therefore by ignoring the values with zero or too small probabilities from the sampling distribution of $\hat{k}$, the interested moments of this distribution can be calculated as $\mathbb{E}(g(\hat{k})) = \sum_{b=1}^{M} g(b)P(\hat{k} = b)$, where M denotes a large natural value, e.g. 100. We provide the graphs of the root mean square error (RMSE) and the standard deviation (SD) of $\hat{k}$ for various values of the sample size and parameters α, γ in Figure 5. It can be seen that the RMSE and SD of the proposed maximum likelihood estimator are both decreasing function in terms of sample size. Of course it should be noted that, as a drawback of the proposed estimator, after a strick decreasing in the RMSE of this estimator, the rate of the decreasing is very small for large values of the sample size. Since the variance is exponentially decreasing, this facts indicates that some biases remains even for large sample sizes. From the Figure 5 it is also resulted that the grater values of the parameters α, γ are correspond to the smaller values of the RMSE and variance of the proposed maximum likelihood estimator for k. The cdf and quantile function of the asymptotic sampling distribution of estimator $\hat{k}$ are given by

$$
\begin{aligned}
F_{\hat{k}}(t) &= \sum_{b \le [t]} P(\hat{k} = b), \\
F_{\hat{k}}^{-1}(p) &= \inf\left\{b \in Z; F_{\hat{k}}(b) \ge p\right\},
\end{aligned}
$$

respectively. The quantile function can be used to provide a $(1 - \alpha)\%$ confidence interval for parameter k. For this it suffice to find two quantiles of sampling distribution of $\hat{k}$, say $q_L(\hat{k})$ and $q_U(\hat{k})$, such that

$$
P(q_L(\hat{k}) < k < q_U(\hat{k})) = 1 - \alpha.
$$

When all parameters are unknown the maximum likelihood estimates of the unknown parameters for the FGWE distribution necessarily obtain via an iterative procedure. Therefore, derivation of the exact sampling distribution of the maximum likelihood estimators of λ, α and γ is not possible. Hence, the asymptotic distribution of these estimators could be used instead. Despite of the maximum likelihood estimator of parameter k, for parameters α, γ and λ the estimators share the fundamental optimality properties of the maximum likelihood estimators, i.e. they are unbiased, efficient and asymptotically normal [3]. We provide the asymptotic distribution of the maximum likelihood estimators that are necessary for the inferential proposes. It suffices to compute the Fisher information matrix for obtaining the asymptotic variance of the estimators.
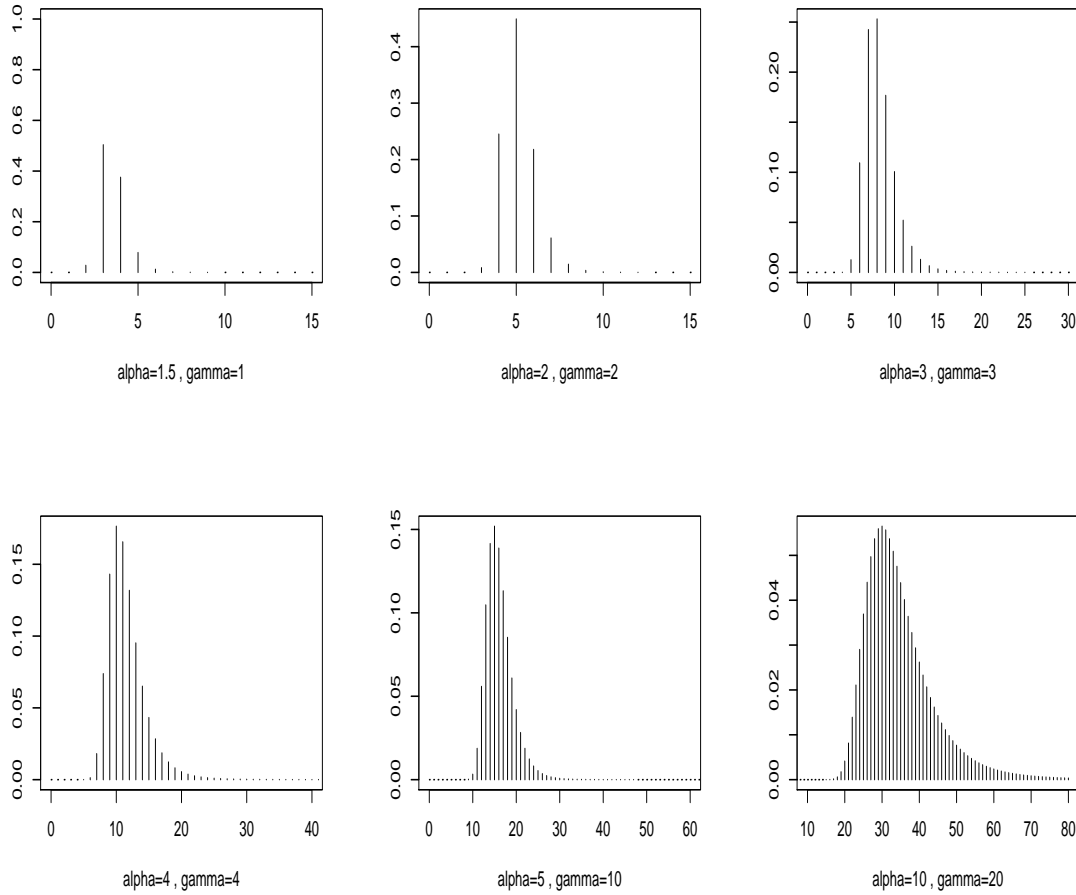
Figure 4: The sampling distribution of the maximum likelihood estimator of parameter k for $n = 100$, $k = 2$ and different values of parameters α, γ.

Consider the profile log-likelihood function of α, λ and γ in (12). The Fisher information matrix is obtain to be

$$\mathbf{I}(\alpha, \lambda, \gamma) = \begin{pmatrix} I_{11}(\alpha, \lambda, \gamma) & I_{12}(\alpha, \lambda, \gamma) & I_{13}(\alpha, \lambda, \gamma) \\ I_{21}(\alpha, \lambda, \gamma) & I_{22}(\alpha, \lambda, \gamma) & I_{23}(\alpha, \lambda, \gamma) \\ I_{31}(\alpha, \lambda, \gamma) & I_{32}(\alpha, \lambda, \gamma) & I_{33}(\alpha, \lambda, \gamma) \end{pmatrix}, \tag{15}$$

where

$$
\begin{aligned}
I_{11}(\alpha, \lambda, \gamma) &= -\mathbb{E}\left(\frac{\partial^2 \ell(\alpha, \lambda, \gamma)}{\partial \alpha^2}\right) = n\frac{C_\alpha^{(2)}(\alpha, \lambda, \gamma)}{C(\alpha, \lambda, \gamma)} - n\left(\frac{C_\alpha^{(1)}(\alpha, \lambda, \gamma)}{C(\alpha, \lambda, \gamma)}\right)^2 + nk\lambda^2 \mathbb{E}\left(\frac{X^2 e^{-\lambda\alpha X}}{1 - e^{-\lambda\alpha X}}\right) \\
&+ nk\lambda^2 \mathbb{E}\left(\frac{X^2 e^{-2\lambda\alpha X}}{(1 - e^{-\lambda\alpha X})^2}\right), \\
I_{22}(\alpha, \lambda, \gamma) &= -\mathbb{E}\left(\frac{\partial^2 \ell(\alpha, \lambda, \gamma)}{\partial \lambda^2}\right) = \frac{n}{\lambda^2} + nk\alpha^2 \mathbb{E}\left(\frac{X^2 e^{-\lambda\alpha X}}{1 - e^{-\lambda\alpha X}}\right) + nk\alpha^2 \mathbb{E}\left(\frac{X^2 e^{-2\lambda\alpha X}}{1 - e^{-\lambda\alpha X}}\right),
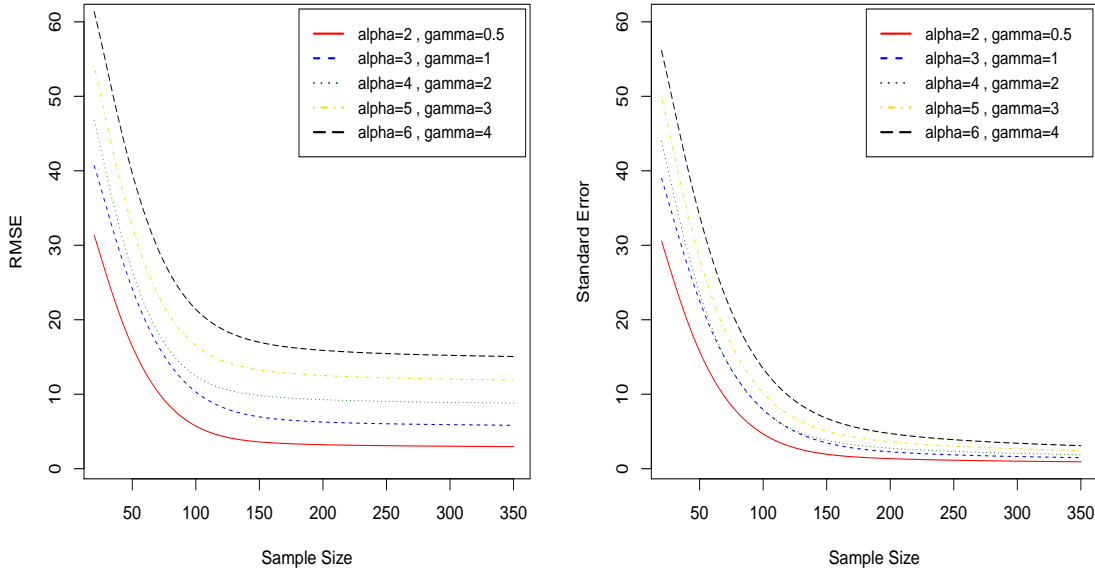\end{aligned}
$$

Figure 5: The RMSE and SD of the maximum likelihood estimator of parameter k for different values of parameters α, γ and sample size.

$$I_{33}(\alpha, \lambda, \gamma) = -\mathbb{E}\left(\frac{\partial^2 \ell(\alpha, \lambda, \gamma)}{\partial \gamma^2}\right) = n\frac{C_\gamma^{(2)}(\alpha, \lambda, \gamma)}{C(\alpha, \lambda, \gamma)} - n\left(\frac{C_\gamma^{(1)}(\alpha, \lambda, \gamma)}{C(\alpha, \lambda, \gamma)}\right)^2,$$

$$I_{12}(\alpha, \lambda, \gamma) = I_{21}(\alpha, \lambda, \gamma) = -\mathbb{E}\left(\frac{\partial^2 \ell(\alpha, \lambda, \gamma)}{\partial \lambda \partial \alpha}\right) = nk\lambda\alpha\mathbb{E}\left(\frac{X^2 e^{-\lambda\alpha X}}{1 - e^{-\lambda\alpha X}}\right) + nk\lambda\alpha\mathbb{E}\left(\frac{X^2 e^{-2\lambda\alpha X}}{(1 - e^{-\lambda\alpha X})^2}\right),$$

$$I_{13}(\alpha, \lambda, \gamma) = I_{31}(\alpha, \lambda, \gamma) = -\mathbb{E}\left(\frac{\partial^2 \ell(\alpha, \lambda, \gamma)}{\partial \gamma \partial \alpha}\right) = n\frac{C_{\alpha,\gamma}^{(1)}(\alpha, \lambda, \gamma)}{C(\alpha, \lambda, \gamma)} - n\frac{C_\alpha^{(1)}(\alpha, \lambda, \gamma)C_\gamma^{(1)}(\alpha, \lambda, \gamma)}{C^2(\alpha, \lambda, \gamma)},$$

$$I_{23}(\alpha, \lambda, \gamma) = I_{32}(\alpha, \lambda, \gamma) = -\mathbb{E}\left(\frac{\partial^2 \ell(\alpha, \lambda, \gamma)}{\partial \gamma \partial \lambda}\right) = n\mathbb{E}(X),$$

where

$$C_{\alpha,\gamma}^{(1)}(\alpha, \lambda, \gamma) = \frac{\partial^2 C(\alpha, k, \gamma)}{\partial \alpha \partial \gamma}.$$

It is easy to show that the expectation terms in the above expressions can be computed via the following general formula

$$\mathbb{E}\left(\frac{X^r e^{-p\lambda\alpha X}}{(1 - e^{-\lambda\alpha X})^s}\right) = \frac{\alpha\Gamma(r+1)}{\lambda^r B(\frac{\gamma}{\alpha}, k+1)} \sum_{i=0}^{k-s} \binom{k-s}{i} \frac{(-1)^i}{(\alpha(i+p) + \gamma)^{r+1}},$$

for $0 \le s \le k$ and $r, p \in \mathbb{N}$. Now, based on normality and optimality of the maximum likelihood estimators, we have

$$(\hat{\alpha}, \hat{\lambda}, \hat{\gamma}) \sim AN_3((\alpha, \lambda, \gamma), \mathbf{I}(\alpha, \lambda, \gamma)^{-1}).$$

This asymptotic distribution can be used to construct confidence interval and to test hypotheses concerning α, λ and γ. The %95 asymptotic confidence interval of the parameters α, λ and γ are obtained, respectively,

to be

$$\alpha \pm 1.96 \sqrt{I_{11}(\hat{\alpha}, \hat{\lambda}, \hat{\gamma})^{-1}},$$

$$\lambda \pm 1.96 \sqrt{I_{22}(\hat{\alpha}, \hat{\lambda}, \hat{\gamma})^{-1}},$$

and

$$\gamma \pm 1.96 \sqrt{I_{33}(\hat{\alpha}, \hat{\lambda}, \hat{\gamma})^{-1}}.$$

10. Data Illustration

In this section, we used a simulation study and a real example are used to assess the proposed maximum likelihood estimators of the FGWE distribution parameters.

10.1. Simulation Study

We simulate n observation from a FGWE distribution with $k = 2$, $\lambda = \alpha = 0.1$ and $\gamma = 2$. As we pointed out earlier, the greater values of the parameter α are correspond to the smaller values of the RMSE and variance of the proposed maximum likelihood estimator for k. Therefore this choice of α is adopted to evaluate the worst situation in estimation of the parameter k. The bias and the Root Mean Squares Error (RMSE) of the maximum likelihood estimators of the parameters are presented in Table 1. It can be seen that the RMSE values of the estimators decrease as the sample size increases. This is due to the asymptotic optimality properties of the maximum likelihood estimators. Following Lachos et al. [11] and Cancho et al. [4], we also proposed selecting the best fit by inspecting the Akaike information criterion ($AIC = 2r - 2\ell(\hat{\theta})$) and the Schwarz Bayesian information criterion ($BIC = -\ell(\hat{\theta}) + 0.5r \log n$), where θ, r and n denote the vector of parameters, the number of estimated parameter and the sample size, respectively. These criteria are computed for the FGWE distribution along with the correspond values for the GWE, WE and exponential distribution in order to provide a comprehensive assessment and comparison of different candidate distributions for a set of right skewed positive observations and the results are presented in Table 2. Throughout the number of iterations is 5000 which seems large enough to take into account the uncertainty in sampling procedure.

Table 1: The bias and Root Mean Squares Error (RMSE) of the maximum likelihood estimators of the parameters.

Estimator	Criterion	Sample Size			
		10	25	50	100
$\hat{\alpha}$	Bias	-0.0100	-0.0100	-0.0100	-0.0100
	RMSE	0.0100	0.0100	0.0100	0.0100
$\hat{\lambda}$	Bias	0.0100	0.0100	0.0100	0.0100
	RMSE	0.0100	0.0100	0.0401	0.0100
$\hat{\gamma}$	Bias	0.0100	0.0200	0.0100	0.0100
	RMSE	0.0100	0.0200	0.0301	0.0100
$\hat{k}$	Bias	0.3500	0.3014	0.2090	0.1262
	RMSE	0.7644	0.5788	0.4580	0.3558

The values of the -log-likelihood, AIC and BIC indicate that the FGWE distribution provides a better fit to data in all cases. An important point is that distinguishing between GWE, WE and exponential distribution may by difficult due to the similar values of the AIC and BIC for these three distributions, for instance consider the sample size 10, the AIC values for the GWE, WE and exponential distributions are, respectively, 44.50 and 44.94. Whereas, there is a considerable difference between the values of these criteria for the FGWE distribution and other distributions. Obviously, as the sample size increases, the asymptotic properties of the maximum likelihood estimators implies that the quality of the fit increases for the FGWE distribution and decreases for the GWE, WE and exponential distributions.

Table 2: The values of the minus log-likelihood (-loglike), AIC, BIC criteria and K-S statistic for the FGWE distribution along with their correspond values for the GWE, WE and exponential distributions.

		Sample Size			
Criterion	Distribution	10	25	50	100
	FGWE	40.27	101.35	203.09	406.15
	GWE	42.47	105.35	212.11	417.32
-loglike	WE	42.50	106.30	212.61	425.13
	Exponential	42.94	107.64	215.45	431.04
	FGWE	43.27	104.36	206.09	409.16
	GWE	44.45	108.12	213.61	425.13
AIC	WE	44.50	108.30	214.61	427.13
	Exponential	44.94	108.64	216.45	432.05
	FGWE	43.72	106.19	208.96	413.06
	GWE	44.12	107.51	215.19	426.81
BIC	WE	44.81	109.52	216.52	429.74
	Exponential	44.10	109.24	217.41	433.35
	FGWE	0.125	0.136	0.165	0.187
	GWE	0.129	0.145	0.174	0.192
K-S	WE	0.131	0.151	0.182	0.193
	Exponential	0.140	0.159	0.184	0.199

10.2. Real Example

Consider the following real data set that shows the lifetimes of 52 industrial productions [9]:

7.62 10.58 1.10 6.89 3.33 8.69 3.22 2.27 4.59 4.91 3.78 2.88 3.49 1.77 6.50 1.73 3.47 5.73 3.03 3.74 5.51 0.68 4.71 4.29 2.43 5.95 5.88 1.51 7.51 3.39 4.23 1.27 2.25 0.97 2.63 8.21 1.96 7.18 1.35 1.90 1.60 7.31 11.79 1.61 1.39 2.19 1.74 1.58 4.09 1.96 1.28 0.17.

The values of the minus log-likelihood, AIC and BIC criteria, K-S and corresponding p-values resulted from fitting the FGWE distribution to this dataset are presented in Table 3. We also provide the corresponding values for the GWE, WE and exponential distribution as another possible candidates for right skewed positive data that both are special cases of the FGWE distribution. The values of these criteria show that the

Table 3: Comparison between the FGWE, GWE, WE and exponential distributions by using different criteria.

	Distribution			
Criterion	FGWE	GWE	WE	Exponential
- loglike	112.32	114.55	119.73	698.63
AIC	114.54	117.55	121.73	699.63
BIC	116.98	119.38	122.95	700.23
K-S	0.098	0.176	0.187	0.198
p-value	0.845	0.433	0.532	0.476

FGWE distribution has better fit to the data than those of the other distributions. Notice that although the GWE distribution improves upon the corresponding WE and exponential model but it does not provides a good fit as much as the FGWE distribution does. Also, the profile log likelihood plots are added in Figure 6.
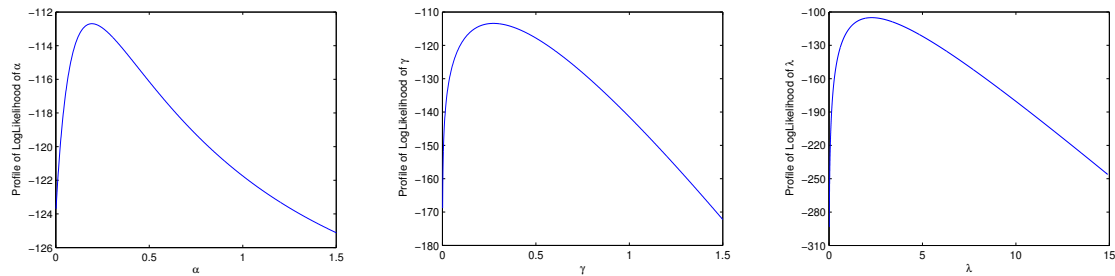
Figure 6: Profile of log likelihood functions for data set in Subsection 10.2.

11. Conclusions

We discussed the main inferential aspects of four-parameter generalized weighted exponential distribution containing the maximum likelihood estimation, interval estimation and some reliability properties. We also developed some new results about the FGWE distribution from the distribution theory point of view and provided some interesting closed form expressions for the this distribution and its related characteristics. Our simulation study indicates to the good efficiency of the proposed estimators. We showed that distinguishing between the GWE, WE and exponential distribution may be difficult for small sample sizes. Also, the real data analysis provides evidence indicating there are some data for them the FGWE distribution provides a considerably better fit in comparison to the GWE, WE and exponential distributions.

References

[1] D. K. Al-Mutairi, M. E. Ghitany, D. Kundu, A new bivariate distribution with weighted exponential marginals and its multivariate generalization, Stat. Pap. 52 (2011) 921–936.
[2] A. Azzalini, A class of distributions which includes the normal ones, Scand. J. Stat. 12 (1985) 171–178.
[3] D. R. Cox, D. V. Hinkley, Theoretical Statistics, Chapman and Hall, London, 1974.
[4] V. G. Cancho, V. H. Lachos, E. M. M. Ortega, A nonlinear regression model with skew-normal errors, Stat. Pap. 51(3) (2010) 547–558.
[5] D. Feldman, M. Fox, Estimation of the parameter n in the binomial distribution, J. Am. Stat. Assoc. 69 (1968) 990–996.
[6] R. D. Gupta, D. Kundu, Generalized exponential distributions, Aust. NZ. J. Stat. 41(2) (1999) 173–188.
[7] R. D. Gupta, D. Kundu, Generalized exponential distributions; different method of estimations. J. Stat. Comput. Sim. 69 (2000) 315–338.
[8] R. D. Gupta, D. Kundu, A new class of weighted exponential distributions, Statistics. 43 (2009) 621–643.
[9] R. Kazemi, Fundamentals of Reliability, Jahade Daneshgahi Press, Qazvin, 2017.
[10] O. Kharazmi, A. Mahdavi, M. Fathizadeh, Generalized weighted exponential distribution, Commun. Stat-Simul. C. 44 (2015) 1557–1569.
[11] V. H. Lachos, H. Bolfarine, R. B. Arellano-Valle, L. C. Montenegro, Likelihood based inference for multivariate skew-normal regression models, Commun. Stat-Theor. Meth. 36 (2007) 1769–1786
[12] M. K. Shakhatreh, A two-parameter of weighted exponential distributions, Stat. Prob. Lett. 82(2) (2012) 52–261.
[13] R. Shongkour, M. A. S. Adnan, Wrapped weighted exponential distributions, Stat. Prob. Lett. 82(1) (2012) 77–83.