

Research Article

Actuarial Measures, Regression, and Applications of Exponentiated Fréchet Loss Distribution

Abdul Ghaniyyu Abubakari 

Department of Statistics, C. K. Tedam University of Technology and Applied Sciences, Navrongo, Ghana

Correspondence should be addressed to Abdul Ghaniyyu Abubakari; aabubakari@cktutas.edu.gh

Received 5 July 2022; Revised 2 August 2022; Accepted 4 August 2022; Published 5 September 2022

Academic Editor: Chin-Chia Wu

Copyright © 2022 Abdul Ghaniyyu Abubakari. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this study, a new loss distribution, called the exponentiated Fréchet loss distribution is developed and studied. The plots of the density function of the distribution show that the distribution can exhibit different shapes including right skewed and decreasing shapes, and various degrees of kurtosis. Several properties of the distribution are obtained including moments, mean excess function, limited expected value function, value at risk, tail value at risk, and tail variance. The estimators of the parameters of the distribution are obtained via maximum likelihood, maximum product spacing, ordinary least squares, and weighted least squares methods. The performances of the various estimators are investigated using simulation studies. The results show that the estimators are consistent. The new distribution is extended into a regression model. The usefulness and applicability of the new distribution and its regression model are demonstrated using actuarial data sets. The results show that the new loss distribution can be used as an alternative to modelling actuarial data.

1. Introduction

In actuarial practice, there is the need to appropriately model data sets. Achieving this can lead to optimal capital allocation as a result of accurate calculations of risk measures and insurance premiums. This is essential for risk management purposes. Due to this, probability distributions are very essential in actuarial practice. Several distributions have been used in actuarial practice including Pareto, gamma, beta, Fréchet, exponential and Weibull distributions. However, given the nature of actuarial data, specifically loss data, some of these distributions are not able to appropriately model such data. For instance, loss data are observed to be heavy-tailed in nature and require distributions that exhibit such properties to be able to model them [1, 2]. Thus, several new distributions have been developed and studied by researchers over the decades for modelling loss data.

Due to the nature of loss data, distributions that exhibit right skewness, such as extreme value distributions or their generalizations, are used to model them. The Fréchet distribution, also known as type II extreme value distribution, is

a special case of the generalized extreme value distribution. It has applications in several fields including actuarial science, finance, hydrology, and biological studies (see [3]). Due to its usefulness, several families of distributions have been developed for the generalization of the Fréchet distribution. Some of these include odd Fréchet family [4], extended odd Fréchet family [5, 6], transmuted odd Fréchet family [7], generalized odd Fréchet [8], and exponentiated Fréchet family [9]. Specifically, some generalizations of the Fréchet distribution include exponentiated Fréchet (EF) [10], beta Fréchet (BF) [11], gamma extended Fréchet [12], Kumaraswamy Fréchet (KF) [13], Weibull Fréchet [14], modified Fréchet [15], beta exponentiated Fréchet (BEF) [16], Burr X Fréchet [17], modified Kies-Fréchet [18], and extended Weibull Fréchet [19] distributions.

In this study, a new extension of EF distribution, known as the exponentiated Fréchet loss (EFL) distribution is developed and studied. The new distribution is developed using a family of loss distributions proposed by Ahmad et al. [20]. Regression models are very essential in relating a response variable to an independent variable(s). Letting the

response variable follow the EFL distribution, an EFL regression model can be developed. Thus, in this study, an EFL regression model is developed and its application demonstrated.

The rest of the article is organized as follows: Section 2 presents the EFL distribution. Some statistical properties, including moments and moment generating function, are presented in Section 3. Section 4 presents some actuarial properties including mean excess function, limited expected value function, value at risk, tail value at risk, and tail variance. Section 5 presents four parameter estimation methods for estimating the parameters of the distribution. Monte Carlo simulation studies to assess the performance of the estimators are carried out in Section 6. A new regression model based on the EFL distribution is given in Section 7. The usefulness of the new distribution and its regression model are demonstrated on real data sets in Section 8. Section 9 presents the conclusion of the study.

2. Exponentiated Fréchet Loss Distribution

Let the random variable X follow the family of loss distributions proposed by Ahmad et al. [20]. Then its cumulative distribution function (CDF) is given as follows:

$$F(x) = 1 - \frac{\alpha(1 - G(x))}{\alpha - \log(1 - G(x))}, \quad \alpha > 0, x \in \mathbb{R}, \quad (1)$$

where $G(x)$ is CDF of the baseline distribution. In this study, the EF distribution is used as the baseline distribution. The CDF of the EF distribution is given as follows:

$$G(x) = 1 - \left[1 - e^{-(\beta/x)^\lambda}\right]^\theta, \quad x > 0, \beta > 0, \theta > 0, \lambda > 0, \quad (2)$$

where β is a scale parameter and θ and λ are shape parameters. Substituting equation (2) into equation (1) gives the CDF of the EFL distribution as follows:

$$F(x) = 1 - \alpha \left[1 - e^{-(\beta/x)^\lambda}\right]^\theta \left\{ \alpha - \theta \log \left[1 - e^{-(\beta/x)^\lambda}\right] \right\}^{-1}, \quad x > 0, \alpha > 0, \beta > 0, \theta > 0, \lambda > 0. \quad (3)$$

The differential of equation (3) gives the probability density function (PDF) of the EFL distribution as follows:

$$f(x) = \alpha \theta \lambda \beta^\lambda x^{-(\lambda+1)} e^{-(\beta/x)^\lambda} \left[1 - e^{-(\beta/x)^\lambda}\right]^{\theta-1} \frac{\left\{1 + \alpha - \theta \log \left[1 - e^{-(\beta/x)^\lambda}\right]\right\}}{\left\{\alpha - \theta \log \left[1 - e^{-(\beta/x)^\lambda}\right]\right\}^2}, \quad x > 0. \quad (4)$$

The hazard rate function of ELF distribution is obtained as follows:

$$\tau(x) = \theta \lambda \beta^\lambda x^{-(\lambda+1)} e^{-(\beta/x)^\lambda} \frac{\left\{1 + \alpha - \theta \log \left[1 - e^{-(\beta/x)^\lambda}\right]\right\}}{\left[1 - e^{-(\beta/x)^\lambda}\right] \left\{\alpha - \theta \log \left[1 - e^{-(\beta/x)^\lambda}\right]\right\}}, \quad x > 0. \quad (5)$$

It should be noted that when $\theta = 1$ in the EFL distribution, Fréchet loss distribution is obtained. To show the flexibility of the EFL distribution, some plots of its PDF and hazard rate functions are obtained for some parameter values of the distribution and shown in Figure 1. It can be observed that the PDF can exhibit right-skewed, decreasing, approximately symmetric shapes, and various degrees of kurtosis. Also, the hazard rate function exhibit decreasing, increasing, and reverse J shapes.

2.1. Expansion of PDF. The expansion of the PDF of EFL distribution is obtained in this subsection. The expansion is

useful for obtaining some quantities, such as the moments, moment generating function, and other properties, of the distribution that involve integrals of the PDF or its functions. Using generalized binomial expansion defined as

$(1 - z)^a = \sum_{i=0}^{\infty} \binom{a}{i} (-1)^i z^i, |z| < 1$, we have

$$\left[1 - e^{-(\beta/x)^\lambda}\right]^{\theta-1} = \sum_{i=0}^{\infty} \binom{\theta-1}{i} (-1)^i e^{-i(\beta/x)^\lambda}. \quad (6)$$

Also, using the binomial expansion $(1 + z)^{-2} = \sum_{j=0}^{\infty} (-1)^{j-1} j z^{j-1}$, we have

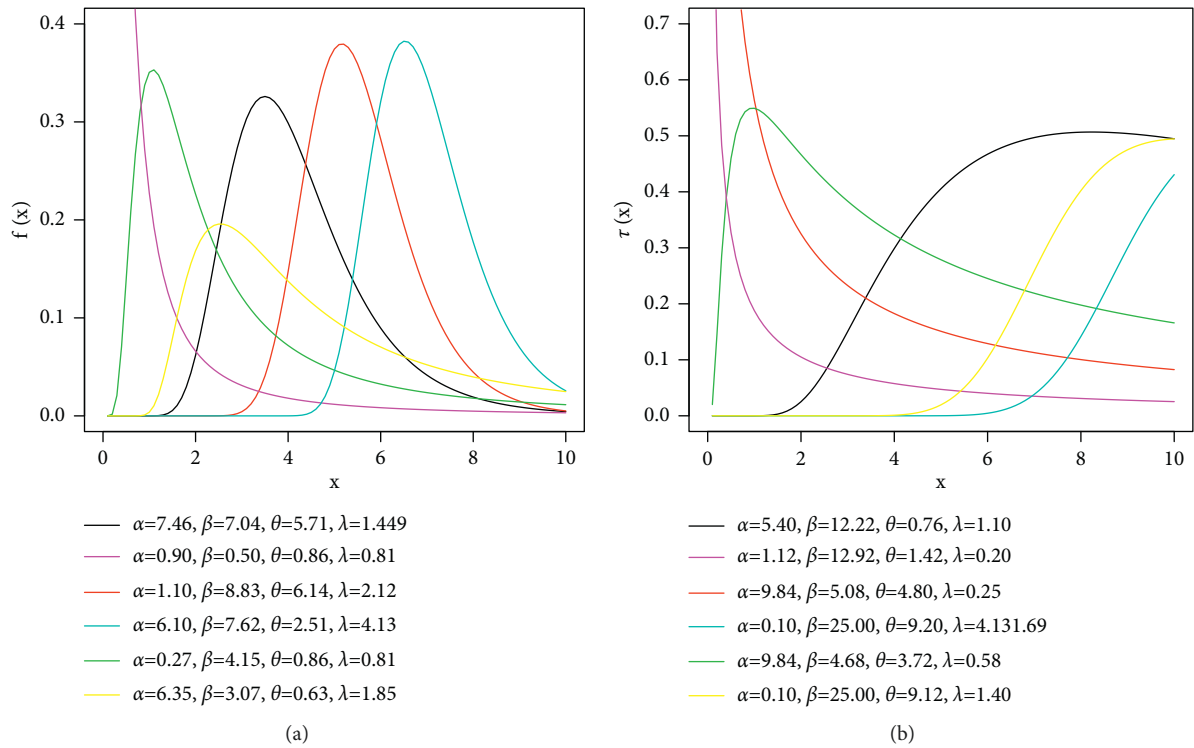


FIGURE 1: Plots of PDF (a) and hazard rate function (b) of the EFL distribution.

$$\begin{aligned} \left\{ \alpha - \theta \log \left[1 - e^{-(\beta/x)^\lambda} \right] \right\}^{-2} &= \frac{1}{\alpha^2} \left\{ 1 + \frac{\theta}{\alpha} \left(-\log \left[1 - e^{-(\beta/x)^\lambda} \right] \right) \right\}^{-2} \\ &= \frac{1}{\alpha^2} \sum_{j=1}^{\infty} j (-1)^{j-1} \left(\frac{\theta}{\alpha} \right)^{j-1} \left(-\log \left[1 - e^{-(\beta/x)^\lambda} \right] \right)^{j-1}, \quad \alpha \geq \theta. \end{aligned} \tag{7}$$

Substituting equations (6) and (7) into the PDF of the EFL distribution in equation (4) gives

$$\begin{aligned} f(x) &= \lambda \beta^\lambda x^{-(\lambda+1)} \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \binom{\theta-1}{i} \left(\frac{\theta}{\alpha} \right)^i (-1)^{i+j-1} j e^{-(i+1)(\beta/x)^\lambda} \\ &\quad \times \left[(1+\alpha) \left(-\log \left[1 - e^{-(\beta/x)^\lambda} \right] \right)^{j-1} + \theta \left(-\log \left[1 - e^{-(\beta/x)^\lambda} \right] \right)^j \right]. \end{aligned} \tag{8}$$

Using the expansion $(-\log(1-z))^a = a \sum_{k=0}^{\infty} \sum_{m=0}^k \binom{k-a}{k} \binom{k}{m} (P_{m,k}/(a-m)) z^{a+k}$, where $P_{m,k} = (1/k) \sum_{n=1}^k ((mn+n-k)/(n+1)) (-1)^n P_{m,k-n}$, $k = 1, 2, \dots, P_{m,0} = 1$ [21, 22], we have

$$\left(-\log \left[1 - e^{-(\beta/x)^\lambda} \right] \right)^{j-1} = (j-1) \sum_{k=0}^{\infty} \sum_{m=0}^k \binom{k-j+1}{k} \binom{k}{m} \frac{P_{m,k}}{j-m-1} e^{-(j+k-1)(\beta/x)^\lambda}, \tag{9}$$

$$\left(-\log \left[1 - e^{-(\beta/x)^\lambda} \right] \right)^j = j \sum_{k=0}^{\infty} \sum_{m=0}^k \binom{k-j}{k} \binom{k}{m} \frac{P_{m,k}}{j-m} e^{-(j+k)(\beta/x)^\lambda}. \tag{10}$$

Substituting equations (9) and (10) into equation (8), and after some algebraic manipulations, gives the PDF of the EFL distribution as follows:

$$f(x) = \frac{\lambda\beta^\lambda}{x^{\lambda+1}} \left[(1 + \alpha)\Pi_{ijkm} e^{-(i+j+k)(\beta/x)^\lambda} + \theta\Pi_{ijkm}^* e^{-(i+j+k+1)(\beta/x)^\lambda} \right], \tag{11}$$

where $\Pi_{ijkm} = \sum_{i=0}^\infty \sum_{j=1}^\infty \sum_{k=0}^\infty \sum_{m=0}^k \binom{\theta-1}{i} \binom{k-j+1}{k} \binom{k}{m} (\theta/\alpha)^i (P_{m,k}/(j-m-1)) j(j-1)(-1)^{i+j-1}$, $\Pi_{ijkm}^* = \sum_{i=0}^\infty \sum_{j=1}^\infty \sum_{k=0}^\infty \sum_{m=0}^k \binom{\theta-1}{i} \binom{k-j}{k} \binom{k}{m} (\theta/\alpha)^i (P_{m,k}/(j-m)) j^2 (-1)^{i+j-1}$, and $P_{m,k} = (1/k) \sum_{n=1}^k ((mn+n-k)/(n+1)) (-1)^n P_{m,k-n}$, $k = 1, 2, \dots, P_{m,0} = 1$.

3. Statistical Properties

In this section, some statistical properties of the EFL distribution are obtained. These include the quantile function, ordinary and incomplete moments, and moment generating function.

3.1. Quantile Function. The quantile function of a distribution can be used to characterize the distribution. It can also be used to obtain random numbers from the distribution and obtain some quantile-based quantities, such as the skewness and kurtosis, of the distribution. The quantile function is obtained as the inverse function of the CDF of a distribution. That is, $Q_X(u) = F^{-1}(u)$. For the EFL distribution, the quantile function is obtained as follows:

$$Q_X(u) = \beta \left[-\log \left(1 - \left\{ \left(\frac{1-u}{\alpha} \right) W \left[\left(\frac{\alpha}{1-u} \right) e^\alpha \right] \right\}^{1/\theta} \right) \right]^{1/\lambda}, \tag{12}$$

$u \in (0, 1)$,

where $W(xe^x) = x$ is the Lambert function. Substituting $u = 0.5$ into the quantile function gives the median of the EFL distribution as follows:

$$\text{Median} = \beta \left[-\log \left(1 - \left\{ \frac{1}{2\alpha} W[2\alpha e^\alpha] \right\}^{1/\theta} \right) \right]^{1/\lambda}. \tag{13}$$

Moor's kurtosis (K) and Bowley's skewness (S) can be defined using the quantile function, respectively, as follows:

$$K = \frac{Q_x(7/8) - Q_x(5/8) + Q_x(3/8) - Q_x(1/8)}{Q_x(6/8) - Q_x(2/8)}, \tag{14}$$

$$S = \frac{Q_x(3/4) - 2Q_x(2/4) + Q_x(1/4)}{Q_x(3/4) - Q_x(1/4)}. \tag{15}$$

Plots of the kurtosis and skewness of the EFL distribution are obtained and shown in Figure 2 for $\alpha = 8.1$, $\beta = 1.8$, and a range of values for θ and λ . It can be observed that the EFL distribution can assume various degrees of kurtosis and also assume both negative and positive skewness.

3.2. Moments and Moment Generating Function. The ordinary moments, incomplete moments, and moment generating function (MGF) of the EFL distribution are given in this section. These properties are useful for characterizing the distribution and for obtaining some other properties of the distribution such as variance, coefficients of skewness and kurtosis, and mean excess function.

3.2.1. Ordinary Moments. The ordinary moment of a distribution is defined as $\mathbb{E}[X^r] = \int_0^\infty x^r f(x) dx$, $r = 1, 2, 3, \dots$. Thus, the ordinary moment of the EFL distribution is obtained by substituting the PDF of the distribution in equation (11) into the definition. This is given as follows:

$$\mathbb{E}[X^r] = \lambda\beta^\lambda \left[(1 + \alpha)\Pi_{ijkm} \int_0^\infty x^{r-(\lambda+1)} e^{-(i+j+k)(\beta/x)^\lambda} dx + \theta\Pi_{ijkm}^* \int_0^\infty x^{r-(\lambda+1)} e^{-(i+j+k+1)(\beta/x)^\lambda} dx \right]. \tag{16}$$

Letting $z = (i+j+k)(\beta/x)^\lambda$ and $y = (i+j+k+1)(\beta/x)^\lambda$ in the first and second integrals, respectively, in

equation (16), and after some algebraic manipulations, gives the ordinary moment of the EFL distribution as follows:

$$\mu_r' = \beta^r \left[(1 + \alpha)\Pi_{ijkm} (i+j+k) \frac{r}{\lambda} - 1 + \theta\Pi_{ijkm}^* (i+j+k+1) \frac{r}{\lambda} - 1 \right] \Gamma\left(1 - \frac{r}{\lambda}\right), \quad r < \lambda, r = 1, 2, \dots, \tag{17}$$

where $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$ is the gamma function and Π_{ijkm} and Π_{ijkm}^* are as defined in equation (11). The mean of

the EFL distribution is obtained by letting $r = 1$ in equation (17). Thus, the mean of EFL distribution is given as follows:

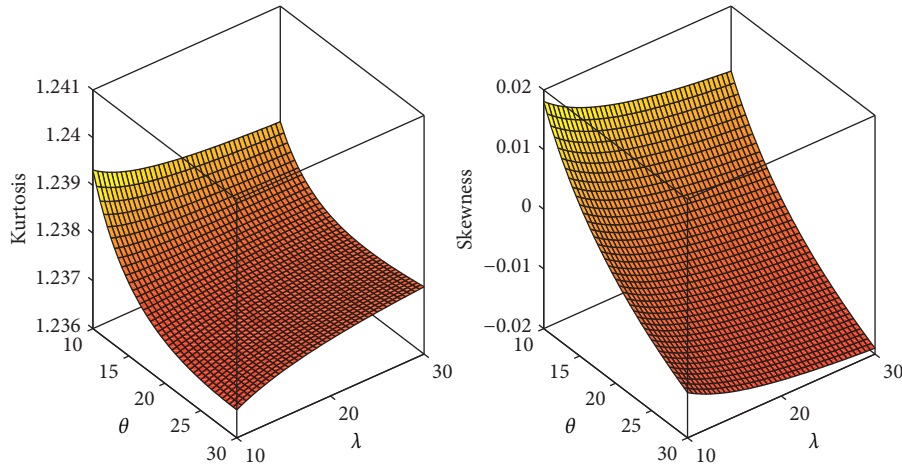


FIGURE 2: Kurtosis and skewness plots.

$$\mu = \beta \left[(1 + \alpha) \Pi_{ijkm} (i + j + k)^{1/\lambda - 1} + \theta \Pi_{ijkm}^* (i + j + k + 1)^{1/\lambda - 1} \right] \Gamma(1 - 1/\lambda), \lambda > 1. \tag{18}$$

Important measures such as standard deviation, coefficient of variation (CV), coefficients of skewness (CS), and kurtosis (CK) of the EFL distribution can be obtained via various ordinary moments of the distribution. The standard deviation and CV are measures of risk and are defined as $\sigma = \sqrt{\mu'_2 - \mu^2}$ and $CV = \sigma/\mu$, respectively. Also, CS and CK are defined, respectively, as $CS = (\mu'_3 - 3\mu'_2\mu + 2\mu^3)/\sigma^3$ and $CK = (\mu'_4 - 4\mu'_3\mu + 6\mu'_2\mu^2 - 3\mu^4)/\sigma^4$. Table 1 shows the first four moments, σ , CV, CS, and CK of the EFL distribution for three sets of parameter values: $I = (\alpha = 4.1, \beta = 4.2,$

$\theta = 4.4, \lambda = 1.5)$, $II = (\alpha = 1.9, \beta = 0.1, \theta = 2.3, \lambda = 1.9)$, and $III = (\alpha = 0.2, \beta = 1.1, \theta = 2.2, \lambda = 2.5)$. Again, it can be observed that the EFL distribution can exhibit various degrees of kurtosis and skewness, including negative skewness.

3.2.2. *Incomplete Moments.* The incomplete moment of a distribution with PDF $f(x)$ is defined as $m_r(y) = \int_0^y x^r f(x) dx, r = 1, 2, 3, \dots$. Substituting PDF of the EFL distribution in equation (11) into the definition gives

$$m_r(y) = \lambda \beta^\lambda \left[(1 + \alpha) \Pi_{ijkm} \int_0^y x^{r-(\lambda+1)} e^{-(i+j+k)(\beta/x)^\lambda} dx + \theta \Pi_{ijkm}^* \int_0^y x^{r-(\lambda+1)} e^{-(i+j+k+1)(\beta/x)^\lambda} dx \right]. \tag{19}$$

After substitution, similar to obtaining the r th ordinary moment, and some algebraic manipulations, the incomplete moment of the EFL distribution is obtained as follows:

$$m_r(y) = \frac{\beta^r (1 + \alpha) \Pi_{ijkm}}{(i + j + k)^{1-(r/\lambda)}} \Gamma\left(1 - \frac{r}{\lambda}, (i + j + k) \left(\frac{\beta}{y}\right)^\lambda\right) + \frac{\beta^r \theta \Pi_{ijkm}^*}{(i + j + k + 1)^{1-(r/\lambda)}} \Gamma\left(1 - \frac{r}{\lambda}, (i + j + k + 1) \left(\frac{\beta}{y}\right)^\lambda\right), \tag{20}$$

$$r < \lambda, r = 1, 2, 3, \dots,$$

where $\Gamma(a, y) = \int_y^\infty x^{a-1} e^{-x} dx$ is the upper incomplete gamma function. The first incomplete moment of the EFL distribution is given as follows:

$$m_1(y) = \frac{\beta(1 + \alpha) \Pi_{ijkm}}{(i + j + k)^{1-(r/\lambda)}} \Gamma\left(1 - 1/\lambda, (i + j + k) \left(\frac{\beta}{y}\right)^\lambda\right) + \frac{\beta \theta \Pi_{ijkm}^*}{(i + j + k + 1)^{1-(r/\lambda)}} \Gamma\left(1 - 1/\lambda, (i + j + k + 1) \left(\frac{\beta}{y}\right)^\lambda\right), \lambda > 1. \tag{21}$$

TABLE 1: First six moments, σ , CV, CS, and CK of the EFL distribution.

μ'_r	I	II	III
μ'_1	2.7653	0.0855	0.7762
μ'_2	8.8097	0.0087	0.6419
μ'_3	33.2061	0.0012	0.5793
μ'_4	154.8394	0.0003	0.5993
σ	1.0784	0.0356	0.1987
CV	0.3900	0.4238	0.2560
CS	1.9247	3.2234	2.5103
CK	12.0663	50.8141	20.7895

3.3. *Moment Generating Function.* The moment generating function (MGF) of a distribution is defined as $M_X(t) = \mathbb{E}[e^{tX}]$ and is useful in obtaining moments of the distribution. Using Taylor series expansion, MGF can be written

as $M_X(t) = \sum_{r=0}^{\infty} (t^r/r!) \mu'_r, r = 1, 2, 3, \dots$ The MGF of the EFL distribution is obtained by substituting the ordinary moment in equation (17) into the definition. This gives the MGF of the EFL distribution as follows:

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r \beta^r}{r!} \left[(1 + \alpha) \Pi_{ijkm} (i + j + k) \lambda^{\frac{r}{\lambda} - 1} + \theta \Pi_{ijkm}^* (i + j + k + 1) \lambda^{\frac{r}{\lambda} - 1} \right] \Gamma\left(1 - \frac{r}{\lambda}\right), r < \lambda, r = 1, 2, 3, \dots \quad (22)$$

4. Actuarial Properties

In this section, some actuarial properties of the EFL distribution are obtained. These include mean excess function, limited expected value function, value at risk, tail value at risk, and tail variance.

has its mean excess function defined as the expected payment, with losses less than t not paid. Also, in a mortality context, it can be defined as the remaining lifetime of an individual, given that the individual attained a particular age, say t . The mean excess function is defined as follows:

$$e(t) = \mathbb{E}[X - x | X > t] = \frac{1}{1 - F(x)} \int_t^{\infty} x f(x) dx - t. \quad (23)$$

4.1. *Mean Excess Function.* The mean excess function is useful in so many fields. It is also known as mean residual function or complete expectation of life. In an insurance context, an insurance policy with a fixed deductible, say t ,

Using the PDF of the EFL distribution given in equation (11), we have

$$\int_t^{\infty} x f(x) dx = \lambda \beta^\lambda \left[(1 + \alpha) \Pi_{ijkm} \int_t^{\infty} x^{-\lambda} e^{-(i+j+k)(\beta/x)^\lambda} dx + \theta \Pi_{ijkm}^* \int_t^{\infty} x^{-\lambda} e^{-(i+j+k+1)(\beta/x)^\lambda} dx \right]. \quad (24)$$

Letting $z_1 = (i + j + k)(\beta/x)^\lambda$ and $z_2 = (i + j + k + 1)(\beta/x)^\lambda$, and after some algebraic manipulations, we have

$$\int_t^{\infty} x f(x) dx = \frac{\beta(1 + \alpha) \Pi_{ijkm}}{(i + j + k)^{1-(1/\lambda)}} \gamma\left(1 - 1/\lambda, (i + j + k) \left(\frac{\beta}{t}\right)^\lambda\right) + \frac{\beta \theta \Pi_{ijkm}^*}{(i + j + k + 1)^{1-(1/\lambda)}} \gamma\left(1 - 1/\lambda, (i + j + k + 1) \left(\frac{\beta}{t}\right)^\lambda\right), \lambda > 1, \quad (25)$$

where $\gamma(a, y) = \int_0^y x^{a-1} e^{-x} dx$. Substituting equation (25) into equation (23) gives the mean excess function of the EFL distribution as follows:

$$e(t) = \frac{\beta}{1 - F(t)} \left[\frac{(1 + \alpha) \Pi_{ijkm}}{(i + j + k)^{1-(1/\lambda)}} \gamma\left(1 - 1/\lambda, (i + j + k) \left(\frac{\beta}{t}\right)^\lambda\right) + \frac{\theta \Pi_{ijkm}^*}{(i + j + k + 1)^{1-(1/\lambda)}} \gamma\left(1 - 1/\lambda, (i + j + k + 1) \left(\frac{\beta}{t}\right)^\lambda\right) \right] - t, \lambda > 1. \quad (26)$$

Figure 3 shows some plots of the mean excess function for three sets of parameter values of the EFL distribution. It can be observed that the mean excess function generally increases and can also assume both linear and nonlinear shapes.

4.2. Limited Expected Value Function. Given a policy limit or a deductible from a reinsurance perspective, say u , a limited loss random variable is defined as follows:

$$X\Lambda u = \min(X, u) = \begin{cases} X, & \text{if } X \leq u, \\ u, & \text{if } X > u. \end{cases} \quad (27)$$

The limited expected value function is defined as the expectation of the limited loss random variable given as follows:

$$E[X\Lambda u] = \int_0^u x f(x) dx + u(1 - F(u)) = m_1(u) + u(1 - F(u)), \quad (28)$$

where $m_1(u)$ is the first incomplete moment given in equation (21). Substituting equations (3) and (21) into the definition gives the limited expected value function of the EFL distribution as follows:

$$\begin{aligned} E[X\Lambda u] &= \frac{\beta(1 + \alpha)\Pi_{ijkm}}{(i + j + k)^{1-(1/\lambda)}} \Gamma\left(1 - 1/\lambda, (i + j + k)\left(\frac{\beta}{u}\right)^\lambda\right) + \frac{\beta\theta\Pi_{ijkm}^*}{(i + j + k + 1)^{1-(1/\lambda)}} \Gamma\left(1 - 1/\lambda, (i + j + k + 1)\left(\frac{\beta}{u}\right)^\lambda\right) \\ &+ \left(\frac{\alpha u \left[1 - e^{-(\beta/u)^\lambda}\right]^\theta}{\alpha - \theta \log\left[1 - e^{-(\beta/u)^\lambda}\right]}\right), \lambda > 1. \end{aligned} \quad (29)$$

4.3. Value at Risk. Value at risk (VaR) is a commonly used risk measure. VaR is defined as the loss that will not be exceeded with a given probability. Mathematically, given a probability p , $VaR = \inf\{x \in \mathbb{R} : P(X \leq x) \geq p\}$. Thus, VaR is also known as a quantile risk measure and is defined as $VaR = F^{-1}(p)$ for a continuous distribution. VaR of the EFL distribution with probability p is defined as follows:

$$VaR = \beta \left[-\log\left(1 - \left\{\left(\frac{1-p}{\alpha}\right) W\left[\left(\frac{\alpha}{1-p}\right) e^\alpha\right]\right\}^{1/\theta}\right) \right]^{1/\lambda}, p \in (0, 1). \quad (30)$$

4.4. Tail Value at Risk. Tail value at risk (TVaR) defines the expected value of the worst case of a loss. That is, TVaR measures the expectation of the losses beyond VaR. Given that $P(X \leq VaR) = p$, then TVaR is defined as follows:

$$TVaR = E[X|X > VaR] = \frac{1}{1-p} \int_{VaR}^\infty x f(x) dx. \quad (31)$$

Substituting equation (25), with $t = VaR$, into the definition gives the TVaR of EFL distribution as follows:

$$TVaR = \frac{\beta}{1-p} \left[\frac{(1 + \alpha)\Pi_{ijkm}}{(i + j + k)^{1-(1/\lambda)}} \gamma\left(1 - 1/\lambda, (i + j + k)\left(\frac{\beta}{VaR}\right)^\lambda\right) + \frac{\theta\Pi_{ijkm}^*}{(i + j + k + 1)^{1-(1/\lambda)}} \gamma\left(1 - 1/\lambda, (i + j + k + 1)\left(\frac{\beta}{VaR}\right)^\lambda\right) \right], \lambda > 1. \quad (32)$$

Figure 4 shows the plots of simulated values of VaR and TVaR for different parameter values and a range of confidence levels. It can be observed that increasing confidence levels are associated with increasing VaR and TVaR. This is consistent with practice, as more capital would have to be allocated for risk management purposes if a company wants to be safer at a higher probability.

4.5. Tail Variance. TVaR measures the expectation of losses exceeding VaR but does not measure the variability of these

losses. Tail variance (TV) measures the conditional variance of losses given that they exceed VaR at a given probability. TV at a probability of p is defined as follows:

$$\begin{aligned} TV &= E[X^2|X > VaR] - (TVaR)^2 \\ &= \frac{1}{1-p} \int_{VaR}^\infty x^2 f(x) dx - (TVaR)^2. \end{aligned} \quad (33)$$

Using the PDF of the EFL distribution in equation (1), we have

$$\int_{VaR}^\infty x^2 f(x) dx = \lambda\beta^\lambda \left[(1 + \alpha)\Pi_{ijkm} \int_{VaR}^\infty x^{1-\lambda} e^{-(i+j+k)(\beta/x)^\lambda} dx + \theta\Pi_{ijkm}^* \int_{VaR}^\infty x^{1-\lambda} e^{-(i+j+k+1)(\beta/x)^\lambda} dx \right]. \quad (34)$$

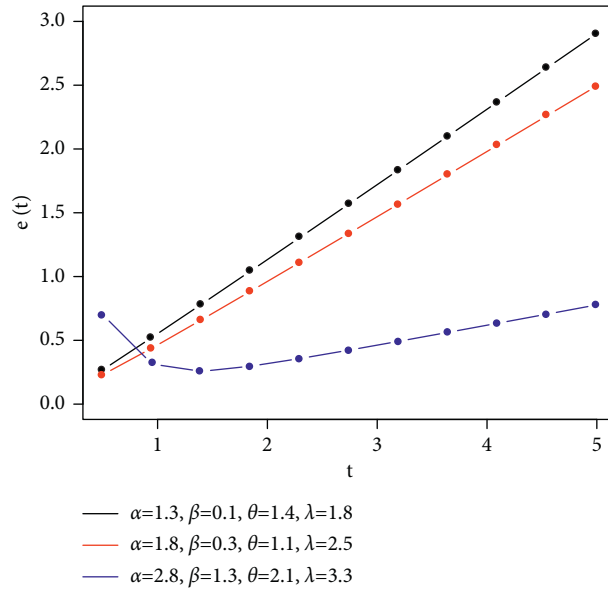


FIGURE 3: Mean excess function plots for some sets of parameter values.

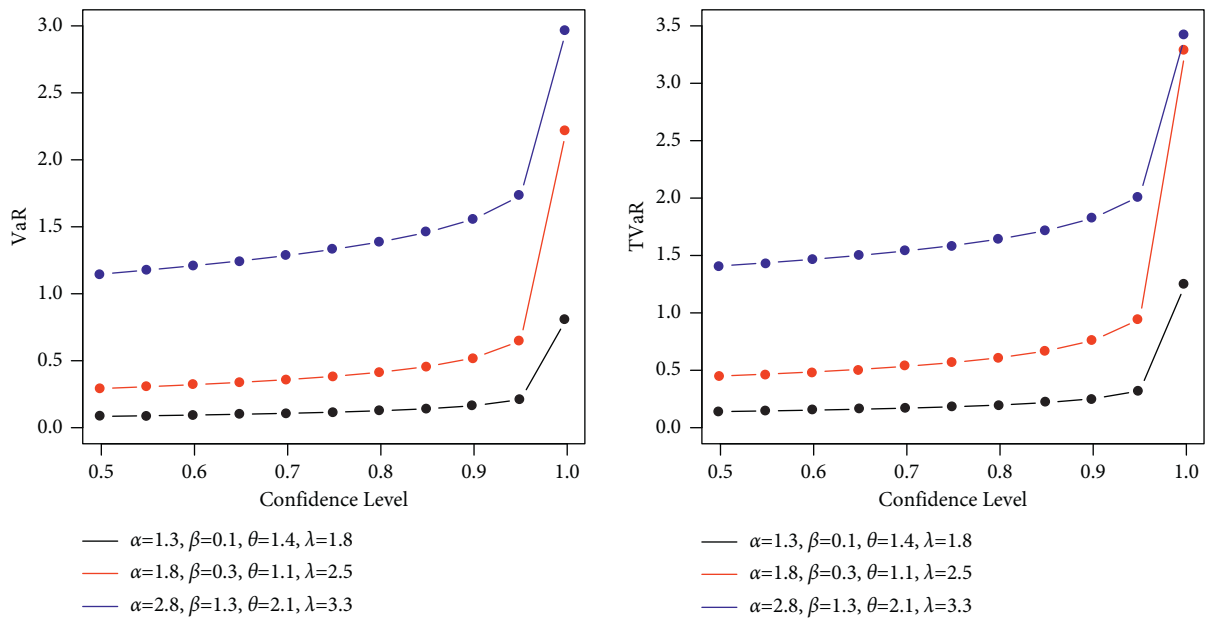


FIGURE 4: VaR and TVaR plots of the EFL distribution.

With the necessary substitutions and algebraic manipulations, we have

$$\int_{VaR}^{\infty} x^2 f(x) dx = \frac{\beta^2 (1 + \alpha) \Pi_{ijkm}}{(i + j + k)^{1-(2/\lambda)}} \gamma \left(1 - \frac{2}{\lambda}, (i + j + k) \left(\frac{\beta}{VaR} \right)^\lambda \right) + \frac{\beta^2 \theta \Pi_{ijkm}^*}{(i + j + k + 1)^{1-(2/\lambda)}} \gamma \left(1 - \frac{2}{\lambda}, (i + j + k + 1) \left(\frac{\beta}{VaR} \right)^\lambda \right), \lambda > 2, \quad (35)$$

where $\gamma(a, y) = \int_0^y x^{a-1} e^{-x} dx$. Substituting equation (35) into equation (33) gives the TV of the EFL distribution as follows:

$$TV = \frac{\beta^2}{1-p} \left[\frac{(1+\alpha)\Pi_{ijkm}}{(i+j+k)^{1-(2/\lambda)}} \gamma\left(1-\frac{2}{\lambda}, (i+j+k)\left(\frac{\beta}{VaR}\right)^\lambda\right) + \frac{\theta\Pi_{ijkm}^*}{(i+j+k+1)^{1-(2/\lambda)}} \gamma\left(1-\frac{2}{\lambda}, (i+j+k+1)\left(\frac{\beta}{VaR}\right)^\lambda\right) \right] - (TVaR)^2, \lambda > 2. \tag{36}$$

5. Parameter Estimation Methods

Estimators of the parameters of the EFL distribution are presented in this section. Four different estimation methods including maximum likelihood, maximum product spacing, least squares, and weighted least squares estimation methods are presented.

5.1. Maximum Likelihood Estimation. Let x_1, x_2, \dots, x_n be n independent and identically distributed random samples from the EFL distribution with a set of parameters $\phi = (\alpha, \beta, \theta, \lambda)'$. The total log-likelihood function of the density of the distribution given in equation (4) is obtained as follows:

$$\begin{aligned} \ell(\phi) = & n \log(\alpha\theta\lambda\beta^\lambda) - (\lambda + 1) \sum_{i=1}^n \log(x_i) - \sum_{i=1}^n \left(\frac{\beta}{x_i}\right)^\lambda + (\theta - 1) \sum_{i=1}^n \log\left[1 - e^{-\left(\frac{\beta}{x_i}\right)^\lambda}\right] \\ & + \sum_{i=1}^n \log\left\{1 + \alpha - \theta \log\left[1 - e^{-\left(\frac{\beta}{x_i}\right)^\lambda}\right]\right\} - 2 \sum_{i=1}^n \log\left\{\alpha - \theta \log\left[1 - e^{-\left(\frac{\beta}{x_i}\right)^\lambda}\right]\right\}. \end{aligned} \tag{37}$$

Equating the score functions, which are obtained by differentiating equation (37) with respect to each parameter, to zero and solving them simultaneously for the parameters give the maximum likelihood estimates (MLE) of the parameters of the EFL distribution. Numerical methods are employed to obtain the parameter estimates since the solution to the equations does not result in closed-form solutions.

with respect to each parameter.

5.2. Maximum Product Spacing Estimation. The maximum product spacing (MPS) method of obtaining parameters is an alternative to the maximum likelihood method. Let the ordered random samples of the EFL distribution be given as $x_{(1)}, x_{(2)}, x_{(3)}, \dots, x_{(n)}$ with CDF $F(x)$ given in equation (3). Define the uniform spacing as follows:

$$H_i(\phi) = F(x_{(i)}|\phi) - F(x_{(i-1)}|\phi), \tag{38}$$

where $F(x_{(0)}|\phi) = 0, F(x_{(n+1)}|\phi) = 1$ and $\sum_{i=0}^{n+1} H_i(\phi) = 1$. The MPS estimates of the parameters are obtained by maximizing the function as follows:

$$S(\phi) = \frac{1}{n+1} \sum_{i=1}^{n+1} \log H_i(\phi), \tag{39}$$

5.3. Ordinary and Weighted Least Squares Estimation. Let the ordered samples of the EFL distribution be given as $x_{(1)}, x_{(2)}, x_{(3)}, \dots, x_{(n)}$. The ordinary least squares (OLS) estimates of its parameters are obtained by minimizing the function as follows:

$$L(\phi) = \sum_{i=1}^n \left[F(x_{(i)}|\phi) - \frac{i}{n+1} \right]^2, \tag{40}$$

with respect to the parameters of the distribution. Also, the weighted least squares (WLS) estimates are obtained by minimizing the following function with respect to the parameters of the distribution:

$$W(\phi) = \sum_{i=1}^n \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F(x_{(i)}|\phi) - \frac{i}{n+1} \right]^2. \tag{41}$$

6. Simulation Studies

Simulation studies are carried out in this section to assess the performance of the parameter estimators. The R program with the nlmnb function is used for the simulation. The

function uses the L-BFGS-B optimization method. The simulation procedure is given as follows:

- (i) Generate $N = 3000$ samples of size $n = 20, 50, 100, 250, 500$ from the EFL distribution using its quantile function in equation (12).
- (ii) Compute the MLE, MPS, OLS, and WLS parameter estimates of the samples obtained in the previous step.
- (iii) For each parameter estimate, obtain the average estimate (AE), absolute bias (AB), and the root mean square error (RMSE) defined as follows:

$$AE = \frac{1}{N} \sum_{i=1}^N \hat{\phi}_i, AB = \frac{1}{N} \sum_{i=1}^N (\hat{\phi}_i - \phi_i) \text{ and RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{\phi}_i - \phi_i)^2}, \tag{42}$$

where $\phi = (\alpha, \beta, \theta, \lambda)$.

- (iv) Steps (i) to (iii) are repeated for two parameter sets $I = (\alpha = 0.2, \beta = 1.2, \theta = 0.5, \lambda = 0.2)$ and $II = (\alpha = 2.1, \beta = 0.9, \theta = 0.7, \lambda = 0.8)$.

Tables 2 and 3 show the simulation results. It can be observed that all the estimation methods are consistent since their AE grows closer to the true parameter values, while AB and RMSE grow towards zero for all the estimation methods. However, generally, for smaller sample sizes, WLS performed better for α and λ , while MPS and MLE performed better for β and θ , respectively, for both sets of simulations. But, for larger sample sizes, MLE and MPS generally performed better for all the parameters in both simulations. Due to the desirable properties of MLE, it will be used to estimate the parameters of the distribution for application purposes.

7. EFL Regression Model

Regression analysis plays an important role in data analysis in most fields, including actuarial science. In this section, a new regression model with the response variable following the EFL distribution is given. Using the regression structure

$$h(\pi_i) = x'_i \delta, i = 1, 2, \dots, n, \tag{43}$$

where $x_i = (1, x_{i1}, x_{i2}, \dots, x_{ip})'$ is the i th vector of independent variables and $\delta = (\delta_0, \delta_1, \delta_2, \dots, \delta_p)'$ is the vector of parameters. $h(\pi_i)$ is known as a link function and links the response variable to the independent variables. Generally, the response variable is linked to the independent variables via the mean. But, also, the response variable can be linked to the independent variable via the quantile or a model parameter. In using a model parameter, a scale or shape parameter is used [2]. In this study, the shape parameter λ is used. Also, the log link function is used. This gives the response variable $Z|x_i$ following the EFL distribution with parameters $\phi = (\alpha, \beta, \theta, \lambda_i)$, where $\log(\lambda_i) = x'_i \delta$,

$i = 1, 2, \dots, n$. The PDF of the EFL regression model is given as follows:

$$f(z) = \alpha \theta \lambda \beta^\lambda z^{-(\lambda+1)} e^{-(\beta/z)^\lambda} \left[1 - e^{-(\beta/z)^\lambda} \right]^{\theta-1} \cdot \frac{\left\{ 1 + \alpha - \theta \log \left[1 - e^{-(\beta/z)^\lambda} \right] \right\}}{\left\{ \alpha - \theta \log \left[1 - e^{-(\beta/z)^\lambda} \right] \right\}^2}, z > 0. \tag{44}$$

The parameters of the EFL regression model can be obtained via the maximum likelihood method by maximizing the log-likelihood function given by

$$\begin{aligned} \ell(\phi) = & n \log(\alpha \theta) + \sum_{i=1}^n \log(\lambda_i) + \sum_{i=1}^n \lambda_i \log(\beta) \\ & - \sum_{i=1}^n (\lambda_i + 1) \log(z_i) - \sum_{i=1}^n \left(\frac{\beta}{z_i} \right)^{\lambda_i} \\ & + (\theta - 1) \sum_{i=1}^n \log \left[1 - e^{-(\beta/z)^\lambda} \right] \\ & + \sum_{i=1}^n \log \left\{ 1 + \alpha - \theta \log \left[1 - e^{-(\beta/z)^\lambda} \right] \right\} \\ & - 2 \sum_{i=1}^n \log \left\{ \alpha - \theta \log \left[1 - e^{-(\beta/z)^\lambda} \right] \right\}. \end{aligned} \tag{45}$$

For practical purposes, after fitting a model, residual analysis is used to diagnose the model and assess its adequacy. In this study, Cox-Snell [23] residual analysis is employed. Cox-Snell residuals are defined as $\hat{e}_i = -\log(1 - F(z_i; \hat{\phi}))$, $i = 1, 2, \dots, n$, where $\hat{\phi}$ is a vector of estimated parameters. The Cox-Snell residuals are standard exponentially distributed if the model fits the data. Checking the adequacy of a model using the Cox-Snell residuals can also be graphically investigated.

7.1. Simulation Studies. A Monte Carlo simulation study is carried out to assess the MLE estimators of the parameters of the EFL regression model. Three independent variables are considered in this simulation. Thus, the regression structure used is

$$\log(\lambda_i) = \delta_0 + \delta_1 x_{i1} + \delta_2 x_{i2} + \delta_3 x_{i3}, \quad i = 1, 2, \dots, n. \tag{46}$$

The process used for the simulation is as follows:

- (i) Generate 3,000 samples of sizes $n = 25, 50, 150, 300, 600$ from the EFL distribution using its quantile function and the independent variables, x_1, x_2 , and x_3 , from a uniform $U(0, 1)$ distribution
- (ii) Obtain the MLE estimates of the parameters $\alpha, \beta, \theta, \delta_0, \delta_1, \delta_2$ and δ_3

TABLE 2: Simulation results for $I = (\alpha = 0.2, \beta = 1.2, \theta = 0.5, \lambda = 0.2)$.

Parameter	n	AE				AB				RMSE			
		MLE	MPS	OLS	WLS	MLE	MPS	OLS	WLS	MLE	MPS	OLS	WLS
α	20	3.2791	4.0993	1.9601	1.902	3.1276	3.9164	1.8903	1.8327	3.8091	4.294	2.8471	2.8076
	50	2.4407	3.4328	1.5148	1.5036	2.2965	3.2551	1.4365	1.4189	3.2077	3.8727	2.3923	2.4244
	100	1.8339	2.9395	1.2619	1.2967	1.6927	2.7635	1.1575	1.1888	2.6758	3.5177	2.0413	2.1709
	250	1.0227	1.8845	0.9203	0.9756	0.8856	1.7126	0.8064	0.8546	1.7599	2.6189	1.5651	1.7263
	500	0.6038	1.0308	0.6002	0.6064	0.4686	0.8665	0.4807	0.4818	1.1109	1.6642	0.9883	1.1003
β	20	3.9546	3.1504	6.5333	6.9199	3.9658	3.3182	5.8841	6.2031	5.4102	4.7979	6.9779	7.1773
	50	3.1699	1.797	5.0604	5.2819	3.3674	2.3267	4.761	4.9327	4.841	3.632	6.0994	6.241
	100	3.0783	1.4832	4.4242	4.5742	3.2416	2.1145	4.2841	4.3904	4.7182	3.3239	5.7158	5.7927
	250	3.0658	1.477	4.158	3.9374	3.0776	1.985	4.0332	3.8157	4.557	3.1679	5.4782	5.2745
	500	3.1455	1.7671	3.9697	3.6965	2.9906	2.0082	3.7867	3.4836	4.4471	3.2061	5.2597	4.9558
θ	20	1.0928	1.3374	1.148	1.0989	0.7659	0.9689	0.9725	0.886	1.2382	1.544	1.4145	1.3158
	50	0.6415	0.7004	0.6976	0.6689	0.3247	0.3623	0.53	0.4374	0.5238	0.6101	0.7982	0.6847
	100	0.5464	0.566	0.5923	0.5652	0.2074	0.2143	0.345	0.2698	0.3066	0.3299	0.5133	0.4021
	250	0.5046	0.5097	0.5195	0.5097	0.128	0.1253	0.2073	0.1587	0.1671	0.1664	0.2852	0.2149
	500	0.4979	0.5002	0.4991	0.4979	0.0925	0.09	0.1401	0.1091	0.1172	0.1148	0.1793	0.14
λ	20	0.2904	0.3529	0.2204	0.2042	0.1198	0.1708	0.0773	0.063	0.191	0.279	0.1606	0.1326
	50	0.2885	0.347	0.2647	0.2428	0.1093	0.1575	0.1015	0.0807	0.1604	0.217	0.2013	0.1508
	100	0.2697	0.3233	0.2569	0.244	0.0894	0.1318	0.0877	0.075	0.1274	0.171	0.1471	0.1165
	250	0.242	0.2837	0.2424	0.2381	0.0617	0.0925	0.0696	0.0639	0.0844	0.1171	0.1047	0.0916
	500	0.226	0.2529	0.2296	0.2252	0.0465	0.0638	0.0553	0.0492	0.0643	0.0848	0.0774	0.0689

TABLE 3: Simulation results for $II = (\alpha = 2.1, \beta = 0.9, \theta = 0.7, \lambda = 0.8)$.

Parameter	n	AE				AB				RMSE			
		MLE	MPS	OLS	WLS	MLE	MPS	OLS	WLS	MLE	MPS	OLS	WLS
α	20	3.7972	4.4066	2.4961	2.2582	2.5633	2.7065	2.2488	2.2212	2.6358	2.7581	2.3539	2.3193
	50	3.4789	4.2889	2.3995	2.2666	2.4340	2.6591	2.2065	2.1678	2.5360	2.7230	2.3111	2.2800
	100	3.2977	4.1315	2.4378	2.4545	2.3295	2.5732	2.1869	2.2036	2.4582	2.6646	2.3009	2.3125
	250	3.0564	3.9016	2.5908	2.6516	2.1733	2.4051	2.1732	2.1253	2.3303	2.5528	2.2972	2.2719
	500	2.9788	3.7685	2.4422	2.5695	2.0164	2.2785	2.0835	2.0182	2.2216	2.4620	2.2249	2.1859
β	20	4.2828	3.9957	6.1518	6.7515	3.6117	3.3059	5.3834	5.9239	5.3204	5.0315	6.6637	7.0373
	50	2.9653	2.2402	4.7130	4.9302	2.3452	1.6507	4.0056	4.1863	4.0954	3.2000	5.5390	5.7511
	100	2.4394	1.6668	3.9783	3.9308	1.7897	1.0664	3.2788	3.2095	3.4669	2.3947	4.8195	4.9096
	250	1.9878	1.3874	2.9522	2.8671	1.2884	0.7392	2.2279	2.1350	2.7699	1.9259	3.6956	3.8618
	500	1.4846	1.0920	2.5043	2.2202	0.7513	0.4068	1.7556	1.4656	1.6941	1.0148	2.9840	2.8564
θ	20	1.6239	2.0225	1.7068	1.6268	1.1616	1.5209	1.3995	1.2951	1.8459	2.3794	2.0348	1.9323
	50	0.9928	1.0818	1.2033	1.0988	0.5447	0.6210	0.8431	0.6923	0.8965	1.0474	1.2886	1.0916
	100	0.8203	0.8516	1.0170	0.9211	0.3255	0.3547	0.5846	0.4405	0.4925	0.5600	0.9429	0.7055
	250	0.7534	0.7519	0.8538	0.8003	0.1830	0.1825	0.3218	0.2310	0.2534	0.2552	0.4783	0.3334
	500	0.7392	0.7324	0.7943	0.7687	0.1329	0.1301	0.2203	0.1661	0.1756	0.1714	0.2984	0.2237
λ	20	0.8258	0.9667	0.7894	0.5815	0.3902	0.4741	0.5319	0.3705	0.5337	0.7869	1.0189	0.5116
	50	0.8537	0.9726	0.7628	0.6804	0.3187	0.3692	0.4017	0.3269	0.4207	0.5205	0.6250	0.4235
	100	0.8160	0.9107	0.7229	0.7068	0.2411	0.2606	0.3138	0.2774	0.3161	0.3668	0.4117	0.3486
	250	0.7811	0.8482	0.7163	0.7244	0.1679	0.1590	0.2311	0.1975	0.2114	0.2081	0.2790	0.2401
	500	0.7772	0.8306	0.7087	0.7285	0.1264	0.1151	0.1907	0.1582	0.1582	0.1464	0.2257	0.1941

TABLE 4: Simulation results of EFL regression model.

Parameter	n	I			II		
		AE	AB	RMSE	AV	AB	RMSE
α	25	1.5750	1.3015	1.4018	1.6328	1.5141	1.9941
	50	1.8351	1.3302	1.4430	1.5665	1.4361	1.9328
	150	1.7534	1.2188	1.3764	1.2907	1.1592	1.6775
	300	1.7799	1.1097	1.3204	0.9486	0.8154	1.3465
	600	1.5831	0.7914	1.0999	0.4932	0.3511	0.7449

TABLE 4: Continued.

Parameter	n	I			II		
		AE	AB	RMSE	AV	AB	RMSE
β	25	0.8377	0.5794	0.9522	0.1858	0.1613	0.5030
	50	0.8812	0.6081	0.9992	0.1891	0.1633	0.4875
	150	0.9366	0.5873	1.0115	0.1429	0.1133	0.3203
	300	0.8054	0.4375	0.8552	0.1230	0.0807	0.2047
	600	0.7047	0.2937	0.6746	0.1117	0.0478	0.0908
θ	25	1.7392	1.1814	1.3640	0.8161	0.5502	0.8555
	50	1.9304	1.0784	1.2860	0.7391	0.4252	0.6870
	150	2.0814	0.7840	1.0541	0.5893	0.2040	0.3078
	300	1.9843	0.5641	0.8413	0.5707	0.1321	0.1822
	600	1.9311	0.3531	0.6311	0.5888	0.0792	0.1169
δ_0	25	0.1763	0.6405	0.7970	0.3600	0.7679	0.9953
	50	0.0387	0.4832	0.5792	0.2207	0.5229	0.6534
	150	-0.0842	0.3709	0.4597	0.1326	0.3392	0.4289
	300	-0.0579	0.2856	0.3903	0.0745	0.2712	0.3480
	600	-0.0036	0.2078	0.3231	0.0079	0.1944	0.2733
δ_1	25	-0.0218	0.5306	0.7228	0.0171	0.7157	0.9328
	50	0.0098	0.3110	0.4033	-0.0040	0.4494	0.5657
	150	0.0176	0.1435	0.1971	0.0361	0.2941	0.3551
	300	0.0211	0.1023	0.1431	0.0354	0.2684	0.3176
	600	0.0398	0.0673	0.1029	0.0884	0.2118	0.2596
δ_2	25	-0.0137	0.5601	0.7514	-0.0288	0.7691	0.9846
	50	-0.0026	0.3177	0.4193	0.0261	0.4868	0.6066
	150	0.0050	0.1478	0.2007	0.0220	0.3901	0.4568
	300	0.0246	0.1018	0.1415	0.0452	0.3558	0.4022
	600	0.0415	0.0646	0.0974	0.1150	0.2850	0.3421
δ_3	25	0.0863	0.9140	1.0952	0.0113	0.7855	1.0018
	50	0.0489	0.8593	0.9506	0.0180	0.5758	0.6966
	150	0.1009	0.7991	0.8608	0.0370	0.4694	0.5320
	300	0.1854	0.7146	0.8043	0.0607	0.4395	0.4844
	600	0.3569	0.5431	0.6982	0.1408	0.3592	0.4269

- (iii) Compute AE, AB, and RMSE of the parameter estimates
- (iv) For the parameters $(\alpha, \beta, \theta, \delta_0, \delta_1, \delta_2, \delta_3)$, repeat steps (i) to (iii) for the parameter values $I = (1.2, 0.5, 1.8, 0.2, 0.1, 0.1, 0.9)$ and $II = (0.2, 0.1, 0.6, 0.1, 0.3, 0.4, 0.5)$

The results of the simulation study are given in Table 4. It can be observed that the estimators of the parameters are consistent as the AE gets closer to the true parameter values, while AB and RMSE decrease with increasing sample size.

8. Applications to Real Data

In this section, the applications of the EFL distribution and EFL regression model to real data are demonstrated.

8.1. Application of EFL Distribution. In this subsection, the application of the EFL distribution is considered. The performance of the distribution is compared with several other distributions using Cramér von-Mises (CVM) and Anderson Darling (AD) goodness-of-fit measures. The distribution with the least of these measures is considered the best distribution. The distributions compared with the

EFL distribution include the Fréchet (F), exponentiated Fréchet (EF), beta exponentiated Fréchet (BEF), Kumaraswamy Fréchet (KF), Weibull (W), and Weibull loss (WL) [20] distributions.

8.1.1. Data 1: Catastrophe Data. The first data consist of the cost associated with natural catastrophic disasters in Australia from 1967 to 2014. The normed cost in millions of 2014 Australian dollars (AUD), computed as the inflated cost using the consumer price index, is used. The data can be found in the CASdatasets package [24] of the R program with the name auscathist. Table 5 shows the descriptive statistics of the data. The data has 206 observations with a wide range of values. Since the median is less than the mean, it suggests that the data is right-skewed.

Figure 5 shows the histogram and box plot of the data. Both figures confirm that the data is right-skewed. This suggests that the EFL distribution can be used to model the data.

The parameter estimates and their corresponding standard errors, in brackets, of the EFL distribution and the other competing distributions are shown in Table 6.

TABLE 5: Descriptive statistics of catastrophe data.

Minimum	Maximum	Range	Mean	Median	1 st Quartile	3 rd Quartile
2.144	4605.932	4603.788	287.830	77.195	32.164	216.305

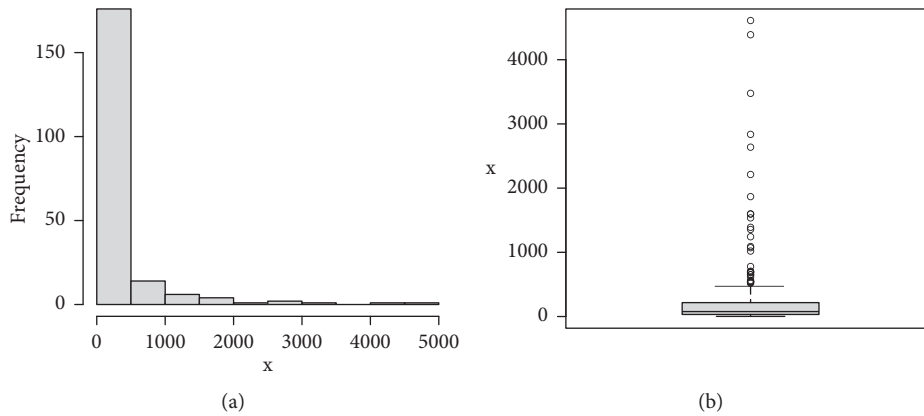


FIGURE 5: (a) Histogram and (b) box plot of catastrophe data.

TABLE 6: Parameter estimates and standard errors for catastrophe data.

Distribution	Parameter				
	a	α	β	θ	λ
EFL		12.9656 (43.7931)	498.2129 (0.8051)	4.2362 (0.9513)	0.3616 (0.0130)
F		40.0063 (4.2430)	0.6960 (0.0355)		
EF		2.6409 (0.6614)	0.4534 (0.0501)	166.6619 (72.9527)	
BEF	32.5139 (0.1268)	5.7122 (6.0689)	0.1245 (0.0215)	13.0977 (0.2957)	3.2091 (1.4834)
KF		4.4756 (0.4949)	5.6070 (3.0253)	8.8135 (0.0837)	0.3327 (0.0686)
W		0.6343 (0.0315)	0.0365 (0.0074)		
WL		0.6358 (0.0316)	0.0342 (0.0069)	16.4990 (<0.0001)	

TABLE 7: Goodness-of-fit measures of estimated distributions for catastrophe data.

Distribution	CVM	CVM p -value	AD	AD p -value
EFL	0.0623	0.8001	0.4372	0.8110
F	0.2708	0.1638	1.9174	0.1021
EF	0.1079	0.5481	0.7117	0.5490
BEF	0.0666	0.7734	0.4449	0.8032
KF	0.0672	0.7693	0.4440	0.8041
W	0.7541	0.0094	4.5831	0.0046
WL	0.7481	0.0097	4.5516	0.0047

Table 7 shows the goodness-of-fit measures of the distributions. It can be observed that the EFL distribution fits the data better than the competing distributions as it has the least of all the goodness-of-fit measures with large corresponding p -values.

Figure 6 shows the PDF plot superimposed on the histogram of the data, the CDF, and probability-probability (P-P) plots of the EFL distribution. It can be observed that the EFL distribution fits the data.

8.1.2. Data 2: Automobile Collision Data. The second data consist of severity, the average amount of claims (in pounds sterling) adjusted for inflation, of automobile

collisions in the United Kingdom. The data can be found in insuranceData package [25] of the R program with the name AutoCollision. Table 8 shows the descriptive statistics of the data. The data consists of 31 observations and indicates positive skewness, as its mean is greater than its median.

Figure 7 shows the histogram and box plot of the data. The data can be observed to be positively skewed; confirmation of the observation is made in Table 8.

Table 9 shows the parameter estimates of the EFL distribution and the other competing distributions with their standard errors in brackets.

The goodness-of-fit measures of the fitted distributions are shown in Table 10 with their corresponding p -values. It

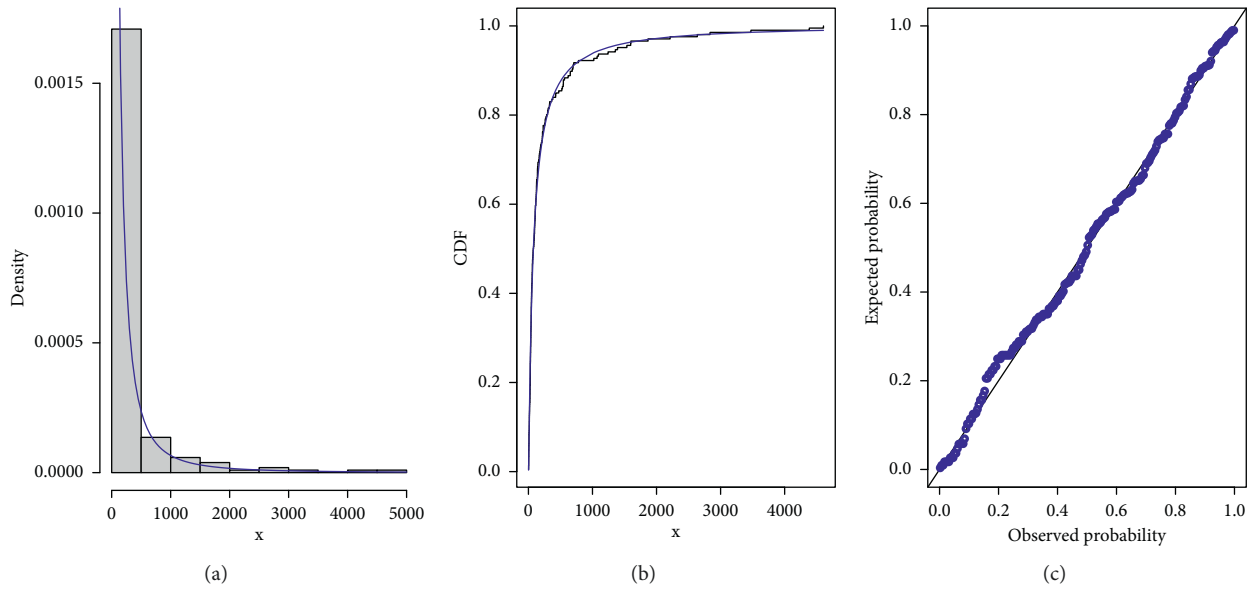


FIGURE 6: Histogram, CDF, and PP plots of EFL distribution for catastrophe data.

TABLE 8: Descriptive statistics of collision data.

No.	Minimum	Maximum	Range	Mean	Median	1 st Quartile	3 rd Quartile
31	153.6	797.8	644.2	277.2	250.6	211.2	298.4

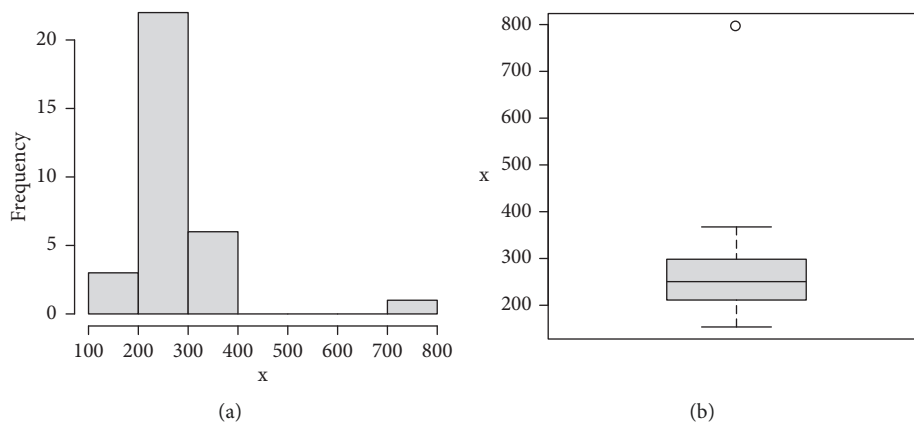


FIGURE 7: (a) Histogram and (b) box plot of collision data.

TABLE 9: Parameter estimates and standard errors for collision data.

Distribution	Parameter				
	a	α	β	θ	λ
EFL		0.0206 (0.0443)	600.2800 (0.0009)	1.2989 (2.3504)	1.6453 (0.1157)
F		231.4037 (9.7832)	4.4036 (0.5835)		
EF		1.2165 (0.7742)	3.998 (1.3859)	240.0675 (34.2864)	
BEF	68.7090 (8.1609)	1.5138 (1.3243)	3.2615 (3.4143)	78.8714 (24.0905)	1.0644 (1.2961)
KF		77.8484 (5.4353)	1.3758 (0.9322)	77.3238 (20.4437)	3.7389 (1.3732)
W		1.0693 (0.3072)	0.0025 (0.0045)		
WL		1.1387 (0.6357)	0.0017 (0.0063)	1090.2000 (<0.0001)	

TABLE 10: Goodness-of-fit measures of estimated distributions for collision data.

Distribution	CVM	CVM p -value	AD	AD p -value
EFL	0.0378	0.9466	0.2981	0.9391
F	0.0436	0.9169	0.3822	0.8653
EF	0.0414	0.9287	0.3653	0.8815
BEF	0.0419	0.9259	0.3647	0.8821
KF	0.0406	0.9329	0.3573	0.8890
W	1.6794	0.0000	7.9414	0.0001
WL	1.6015	0.0001	7.6057	0.0002

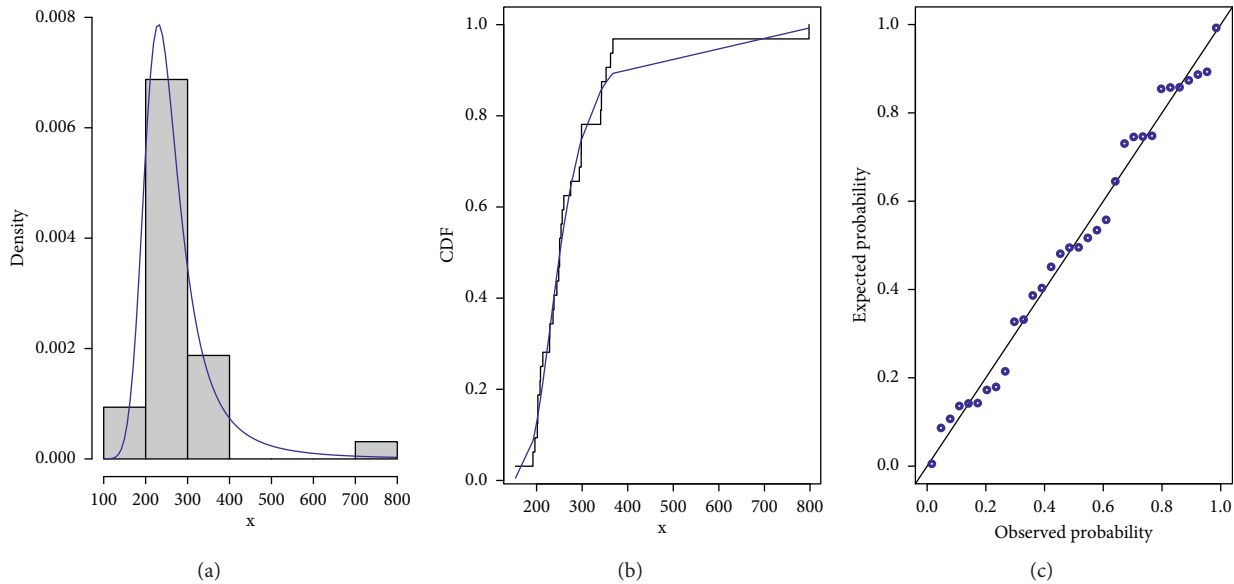


FIGURE 8: Histogram, CDF, and PP plots of EFL distribution for collision data.

can be observed that EFL distribution has the least of the measures and the largest p -value.

Figure 8 shows the histogram of the data with the fitted PDF, the CDF, and P-P plots of the EFL distribution. It can be observed that the EFL distribution can be used to model the automobile collision data.

8.2. Application of EFL Regression Model. This subsection presents an application of the EFL regression model to a real data set. The data used is obtained from insuranceData package [25] in the R program with the name dataOhlsson and comes from the former Swedish insurance company Wasa. The data contains aggregated data on all insurance policies and claims from 1994 to 1998. In this data set, the variables used are the claim cost (z) in 10,000 Swedish krona, vehicle age (x_1) and MC class, a classification by the so-called EV ratio, defined as $(\text{engine power in kW} \times 100) / (\text{vehicle weight in kg} + 75)$, rounded to the nearest lower integer. The 75 kg represents the average driver weight. The EV ratios are divided into seven classes. This data set was analyzed in a regression context by Gündüz and Genç [2]. The descriptive statistics of the claims and frequencies of the MC class are given in Table 11. It can be observed that there are 670 observations with more than zero claims. Also, MC class 6 can be observed to have the highest number of occurrences, with class 7 having the least.

The independent variable MC class is a categorical variable with seven levels and is coded using an indicator variable for the regression model. Given a categorical variable with p levels, then $p - 1$ new indicator variables are introduced. In such a case, one of the categories is chosen as a reference level. Usually, the level with the highest frequency is used as the reference level. Similar to Gündüz and Genç [2], level 6 of the MC class is chosen as the reference level because it has the highest number of occurrences as shown in Table 10. In this scenario, the following levels and their corresponding indicator variables are used: $(1, x_2)$, $(2, x_3)$, $(3, x_4)$, $(4, x_5)$, $(5, x_6)$, and $(7, x_7)$. Hence, the regression model considered is given as follows:

$$\log(\lambda_i) = \delta_0 + \sum_{i=1}^7 \delta_i x_{ij}, j = 1, \dots, n. \tag{47}$$

The performance of the EFL regression model is compared with the EF regression model with parametrization I as defined by Gündüz and Genç [2].

Table 12 shows the parameter estimates of the EFL and EF regression models with their corresponding standard errors (SE) and p -values. Also, the average marginal estimates (AME), which measures the average contribution, of each independent variable is presented in Table 12. Again, the negative log likelihood $(-\ell)$, Akaike information criteria (AIC), and Bayesian information criteria (BIC) are also presented.

TABLE 11: Descriptive statistics of data.

Variable	No.	Minimum	Maximum	Mean	Median	1 st Quartile	3 rd Quartile
Claims	670	0.0016	36.5347	2.5436	0.9015	0.3031	2.9304
MC class	1	2	3	4	5	6	7
Frequency	45	57	158	93	144	168	5

TABLE 12: Regression parameter estimates.

	EFL				EF			
	Estimate	SE	<i>p</i> -value	AME	Estimate	SE	<i>p</i> -value	AME
α	113.0807	0.0063	0.0000					
β	1290.2940	0.0213	0.0000		5.2535	0.0451	0.0000	
θ	35.3300	2.3085	0.0000		134.0171	0.0003	0.0000	
δ_0	-1.5562	0.0221	0.0000		-1.9628	0.0613	0.0000	
Vehicle age	-0.0084	0.0013	0.0000	-0.0013	0.0077	0.0039	0.0468	0.0009
MC class								
1	-0.0518	0.0348	0.1367	-0.0080	-0.1111	0.1216	0.3607	-0.0134
2	-0.0333	0.0317	0.2934	-0.0051	0.1543	0.1093	0.1582	0.0186
3	-0.0565	0.0231	0.0144	-0.0087	-0.1934	0.0776	0.0127	-0.0233
4	-0.0516	0.0268	0.0541	-0.0080	-0.1042	0.0912	0.2532	-0.0125
5	-0.0382	0.0234	0.1027	-0.0059	-0.0790	0.0798	0.3226	-0.0095
7	0.0770	0.0936	0.4109	0.0119	0.1813	0.2939	0.5373	0.0218
$-\ell$			1171.068				1184.263	
AIC			2364.136				2388.525	
BIC			2413.716				2433.598	

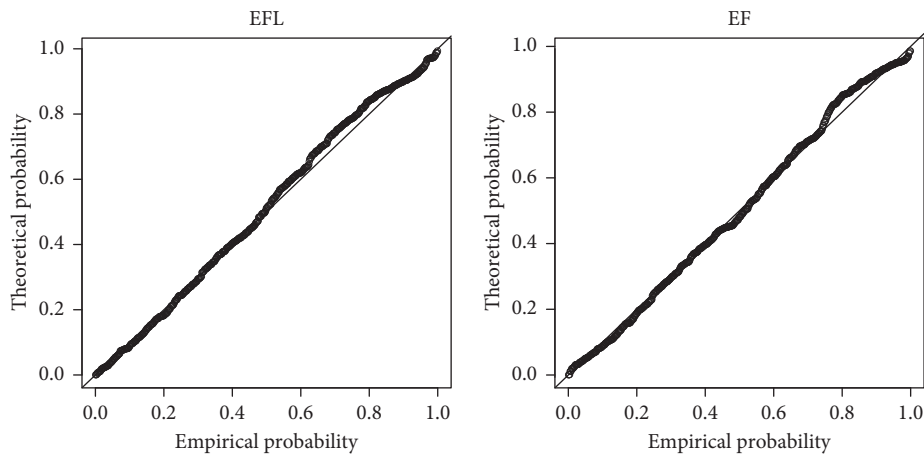


FIGURE 9: Cox-Snell residuals P-P plot.

It can be observed from Table 12 that the vehicle age and MC class 3 are significant and significantly different from MC class 6 at a 5% significance level for both regression models. Both of these variables have a negative impact on the claims, as can be observed from their AME for the EFL regression. However, vehicle age contributes positively to the claims in the EF regression model, while MC class 3 contributes negatively. Finally, EFL regression performs better in modelling the data as compared to the EF model, as the EFL model has the least values in terms of $-\ell$, AIC and BIC measures.

Cox-Snell residuals analysis is performed on the fitted models to evaluate their fit. Figure 9 shows the P-P plots of the empirical probabilities of the residuals against the theoretical probabilities from the standard exponential distribution. It can be observed that the EFL regression has more

plotted points closer to the diagonal as compared to the EF regression model. This confirms that the EFL regression model performed better than the EF regression model in modelling the data.

9. Conclusion

A new loss distribution, called the exponentiated Fréchet loss distribution, is developed and studied. Various statistical properties including the quantile function, moments, and moment generating function are obtained. Also, some actuarial properties including value at risk, tail value at risk, and tail variance of the distribution are obtained. Four estimation methods are used to obtain the estimators of the loss distribution. Simulations studies are performed to assess

the performance of the estimators. The new loss distribution is extended into a regression model. The usefulness of the new loss distribution and its regression model are demonstrated using real data sets. The results show that the exponentiated Fréchet loss distribution and its regression model can serve as an alternative to modelling loss data.

Data Availability

The data used for the analysis are openly available, and the sources are stated in the manuscript.

Conflicts of Interest

The author declares that there are no conflicts of interest with respect to this research.

References

- [1] H. H. Panjer, *Operational Risk, Modeling Analytics*, Wiley, Hoboken (NJ), 2006.
- [2] F. F. Gündüz and A. İ Genç, “The exponentiated Fréchet regression: an alternative model for actuarial modelling purposes,” *Journal of Statistical Computation and Simulation*, vol. 86, no. 17, pp. 3456–3481, 2016.
- [3] S. Kotz and S. Nadarajah, *Extreme Value Distributions: Theory and Applications*, Imperial College Press, London, 2000.
- [4] M. A. Haq and M. Elgarhy, “The odd Fréchet-G family of probability distributions,” *J. Stat. Appl. Probab.*, vol. 7, pp. 189–203, 2018.
- [5] S. Nasiru, “Extended odd Fréchet-G family of distributions,” *Journal of Probability and Statistics*, vol. 2018, pp. 1–12, 2018.
- [6] H. M. Yousof, M. Rasekhi, E. Altun, and M. Alizadeh, “The extended odd Fréchet family of distributions: properties, applications and regression modeling,” *International Journal of Applied Mathematics & Statistics*, vol. 30, no. 1, pp. 1–30, 2018.
- [7] M. M. Badr, I. Elbatal, F. Jamal, C. Chesneau, and M. Elgarhy, “The Transmuted odd Fréchet-G family of distributions: Theory and applications,” *Mathematics*, vol. 8, no. 6, p. 958, 2020.
- [8] S. Marganpoor, V. Ranjbar, M. Alizadeh, and K. Abdollahnezhad, “Generalised odd Fréchet family of distributions: properties and applications,” *Statistics in Transition New Series*, vol. 21, no. 3, pp. 109–128, 2020.
- [9] L. A. Baharith and H. H. Alamoudi, “The exponentiated Fréchet generator of distributions with applications,” *Symmetry*, vol. 13, no. 4, p. 572, 2021.
- [10] S. Nadarajah and S. Kotz, “The exponentiated Fréchet distribution,” *Interstat Electronic Journal*, vol. 14, pp. 01–07, 2003.
- [11] S. Nadarajah and A. K. Gupta, “The beta Fréchet distribution,” *Far East Journal of Theoretical Statistics*, vol. 14, pp. 15–24, 2004.
- [12] R. V. d. Silva, T. A. d. Andrade, D. B. Maciel, R. P. S. Campos, and G. M. Cordeiro, “A new lifetime model: The gamma extended Fréchet distribution,” *Journal of Statistical Theory and Applications*, vol. 12, no. 1, p. 39, 2013.
- [13] M. E. Mead and A. R. Abd-Eltawab, “A Note on Kumaraswamy-Fréchet distribution,” *Australian Journal of Basic and Applied Sciences*, vol. 8, pp. 294–300, 2014.
- [14] A. Z. Afify, H. M. Yousof, G. M. Cordeiro, E. M. Ortega, and Z. M. Nofal, “The Weibull Fréchet distribution and its applications,” *Journal of Applied Statistics*, vol. 43, no. 14, pp. 2608–2626, 2016.
- [15] C. J. Tablada and G. M. Cordeiro, “The modified Fréchet distribution and its properties,” *Communications in Statistics - Theory and Methods*, vol. 46, no. 21, pp. 10617–10639, 2017.
- [16] M. Mead, A. Z. Afify, G. Hamedani, and I. Ghosh, “The beta exponential Fréchet distribution with applications,” *Austrian Journal of Statistics*, vol. 46, no. 1, pp. 41–63, 2017.
- [17] T. Abouelmagd, M. Hamed, A. Afify, H. Al-Mofleh, and Z. Iqbal, “The Burr X Fréchet distribution with its properties and applications,” *Journal of Applied Probability and Statistics*, vol. 29, pp. 23–51, 2018.
- [18] M. M. A. Sobhi, “The modified Kies-Fréchet distribution: properties, inference and application,” *AIMS Mathematics*, vol. 6, no. 5, pp. 44691–44714, 2021.
- [19] E. A. Hussein, H. M. Aljohani, and A. Z. Afify, “The extended Weibull-Fréchet distribution: properties, inference, and applications in medicine and engineering,” *AIMS Mathematics*, vol. 7, no. 1, pp. 225–246, 2021.
- [20] Z. A. Ahmad, E. M. Mahmoudi, and G. G. Hamedani, “A family of loss distributions with an application to the vehicle insurance loss data,” *Pakistan Journal of Statistics and Operation Research*, vol. 15, no. 3, pp. 731–744, 2019.
- [21] S. Nadarajah, G. M. Cordeiro, and E. M. M. Ortega, “The Zografos-Balakrishnan-G family of distributions: Mathematical properties and applications,” *Communications in Statistics - Theory and Methods*, vol. 44, no. 1, pp. 186–215, 2015.
- [22] G. M. Cordeiro, E. M. Ortega, and T. G. Ramires, “A new generalized Weibull family of distributions: Mathematical properties and applications,” *Journal of Statistical Distributions and Applications*, vol. 2, no. 1, p. 13, 2015.
- [23] D. R. Cox and E. J. Snell, “A general definition of residuals,” *Journal of the Royal Statistical Society: Series B*, vol. 30, no. 2, pp. 248–265, 1968.
- [24] C. Dutang and A. Charpentier, “Package ‘CASdatasets,’” 2020, <http://cas.uqam.ca/pub/>.
- [25] A. Wolny-Dominiak and M. Trzęsiok, “Package ‘insuranceData,’” 2014, <https://cran.r-project.org/package=insuranceData>.