New methods to define heavy-tailed distributions with applications to insurance data

Zubair Ahmad\textsuperscript{a,*}, Eisa Mahmoudi \textsuperscript{a}, G. G. Hamedani \textsuperscript{b} and Omid Kharazmi \textsuperscript{c}

\textsuperscript{a}Department of Statistics, Yazd University, Yazd, Iran; \textsuperscript{b}Department of Mathematical and Statistical Sciences, Marquette University, Milwaukee, WI, USA; \textsuperscript{c}Department of Statistics, Faculty of Sciences, Vali-e-Asr university of Rafsanjan, Rafsanjan, Iran

\textbf{ABSTRACT}

Heavy-tailed distributions play an important role in modelling data in actuarial and financial sciences. In this article, nine new methods are suggested to define new distributions suitable for modelling data with an heavy right tail. For illustrative purposes, a special sub-model is considered in detail. Maximum likelihood estimators of the model parameters are obtained and a Monte Carlo simulation study is carried out to assess the behaviour of the estimators. Furthermore, some actuarial measures are calculated. A simulation study based on these actuarial measures is done. Finally, Bayesian analysis and performance of Gibbs sampling for the data sets are also carried out.

\section{1. Introduction}

In many applied areas, particularly in finance and actuarial sciences, data are usually positive, and their distribution is unimodal hump shaped and extreme values yield tails which are heavier than those of the standard well-known distributions. Classical distributions are not flexible enough to cater such heavy-tailed data sets. For example, (i) the Pareto distribution, which is widely used in modelling financial data sets, does not provide a reasonable fit for many applications and (ii) the Weibull model covers better the behaviour of small losses, but fails to cover the behaviour of large losses; see Bhati and Ravi [1]. In such situation, heavy-tailed distributions are reliable and accurate candidate models to employ. For positive data, heavy-tailed distributions are those whose right tail probabilities are greater than the exponential one (see Beirlant et al. [2]), that is

$$\lim_{x \to \infty} \frac{\exp(-\gamma x)}{1 - F(x)} = 0,$$

for any $\gamma > 0$,

where $F(x)$ is the cumulative distribution function (cdf).

For further detail, see McNeil [3] and Resnick [4]. Due to the importance of the heavy-tailed distributions in actuarial practice, the actuaries are motivated to introduce new flexible distributions. In this regard, serious attempts have been made and still growing rapidly.

The new developments have been made through many different approaches such as (i) transformation of variables, (ii) composition of two or more distributions, (iii) compounding of distributions and (iv) finite mixture of distributions.

Recent studies of Eling [5] and Adcock et al. [6] identify that skew-normal and skew student $t$ distributions are the best competitors as the skewed distributions adjust right-skewness and high kurtosis; for further detail see, Shushi [7] and Punzo [8]. However, insurance losses and financial risks take values on the positive real line and hence these skew class of distributions may not be appropriate as they are defined on $\mathbb{R}$. In such situations, the transformation of variables, particularly the exponential transformation, has proven to be substantial; see, for example, Azzalini et al. [9]. Bagnato and Punzo [10] showed that the transformation approach is simple to use but most often the inference as well as computation of the other distributional characteristics become complicated.

Another promising approach for obtaining new flexible heavy-tailed families of distributions, which gives reasonably good fit for heavy-tailed losses, is the method of composition; see Paula et al. [11], Klugman et al. [12], Nadarajah and Abu Bakar [13], and Bakar et al. [14]. However, it should be noted that the new
distributions obtained by the composition approach involve more than three parameters causing difficulties in the estimation process and computational efforts are required.

Another prominent approach is compounding of distributions to cater data modelling with unimodality, right-skewness and heavy tails [8,15,16]. However, the density obtained via this method may not have a closed form expression which makes the estimation more cumbersome as shown in Punzo et al. [15]. For a brief review about compounding of distributions, we refer to Tahir and Cordeiro [17].

Finite mixture models represent a further approach to define very flexible distributions which are also able to capture, for instance, multimodality of the underlying distribution [18–20]. The price to pay for this greater flexibility is a more complicated and computationally challenging inference.

Furthermore, Dutta and Perry [21] performed an empirical analysis of loss distributions and risk was estimated by different approaches such as Exploratory Data Analysis and other empirical approaches. These authors rejected the idea of using the exponential, gamma and Weibull models in modelling insurance losses due to the poor results. They concluded that “one would need to use a model that is flexible enough in its structure”, and this encourages researchers to search for more flexible probability distributions providing greater accuracy in fitting heavy-tailed insurance losses.

Carrying out this branch of distribution theory, Alzaatreh et al. [22] defined the T-X family method to introduce new families of distributions. Let \( \nu(t) \) be the probability density function (pdf) of a random variable, say \( T \), where \( T \in [m, n] \), \(-\infty \leq m < n < \infty \), and let \( W[F(x)] \) be a function of \( F(x) \) of a random variable, say \( X \), satisfying the conditions given below:

1. \( W[F(x)] \in [m, n] \),
2. \( W[F(x)] \) is differentiable and monotonically increasing, and
3. \( W[F(x)] \to m \) as \( x \to -\infty \) and \( W[F(x)] \to n \) as \( x \to \infty \).

The cdf of the T-X family of distributions is defined by

\[
G(x) = \int_m^{W[F(x)]} \nu(t) \, dt, \quad x \in \mathbb{R}, \tag{1}
\]

where \( W[F(x)] \) satisfies the conditions stated above. The pdf corresponding to (1) is

\[
g(x) = \frac{\partial}{\partial x} W[F(x)] \nu[W[F(x)]], \quad x \in \mathbb{R}.
\]

For the contributed work based on the idea of T-X approach, we refer to Ahmad et al. [23]. Using the approach of T-X method, one can introduce new members of survival family [24] via the cdf

\[
G(x) = 1 - \int_m^{W[F(x)]} \nu(t) \, dt, \quad x \in \mathbb{R}, \tag{2}
\]

where \( F(x) = 1 - F(x) \) is the survival function of the baseline distribution.

Recently, Mahdavi and Kundu [25] proposed a new prominent approach for introducing statistical distributions via the cdf given by

\[
G(x; \alpha, \xi) = \frac{\alpha^{\xi F(x)} - 1}{\alpha - 1}, \quad \alpha > 1, \alpha \neq 1, x \in \mathbb{R} \tag{3}
\]

with additional parameter \( \alpha \).

Under these premises, we are motivated to propose new families of distributions. Taking inspiration from (1)–(3), in this article, nine new families of distributions are proposed.

The paper is outlined as follows: new development to the heavy-tailed distributions is presented in Section 2. In Section 3, we define a special sub-model of the proposed family and provide plots of the density function. We derive some statistical properties in Section 4. Some characterizations results are provided in Section 5. Maximum likelihood estimation of the model parameters is addressed in Section 6. In the same section, the Monte Carlo simulation study is provided. Actuarial measures of the proposed method along with a simulation study are provided in Section 7. In Section 8, we provide two applications to real data to illustrate the importance of the new family. Bayesian analysis as well as a Gibbs sampling procedure for the considered data sets are discussed in Section 9. Finally, some concluding remarks are presented in Section 10.

2. New contribution to heavy-tailed distributions

As we mentioned in Section 1, the researchers are often in search of new heavy-tailed distributions suitable for modelling insurance losses. In this section, we propose some new useful methods to obtain heavy-tailed extended versions of the existing distributions.

2.1. New extended exponentiated-X family

Taking inspiration from (1), we introduce a new flexible class of distributions. The proposed class is called a new extended exponentiated-X (NEEx-X) family. Let \( T \sim \exp(1) \), then its cdf is given by

\[
V(t) = 1 - e^{-t}, \quad t \geq 0. \tag{4}
\]

The density function corresponding to (4) is

\[
\nu(t) = e^{-t}, \quad t > 0. \tag{5}
\]

If \( \nu(t) \) follows (5) and setting \( W[F(x)] = -\log \left( \frac{1 - \theta f(x)}{1 - \log[1 - F(x)]} \right) \) in (1), we define the
Taking inspiration from (2), we introduce a new method of proposing distributions, called type-I cosine exponentiated-X family. If \( v(t) \) follows (5) and setting
\[
W[F(x)] = -\log \left( \frac{\cos \left( \frac{\pi}{2} F(x; \xi)^a \right)}{1 - \log \left( 1 - F(x; \xi)^a \right)} \right)
\]
in (1), we define the pdf of the NTICEx-X family given by
\[
g(x) = \frac{af(x; \xi) F(x; \xi)^{a-1}}{\left( 1 - \log \left[ 1 - F(x; \xi)^a \right] \right)^2} \times \left\{ \theta \left[ 1 - \log \left[ A(x) \right] \right] + 1 - \theta F(x; \xi)^a \right\}, \quad x \in \mathbb{R},
\]
where \( A(x) = 1 - F(x; \xi)^a \). For \( a = 1 \), the NTICEx-X family gives the new extended-X(NE-X) family.

### 2.2. New type-I cosine exponentiated-X family

Again taking inspiration from (1), we introduce another new family of distributions, called new type-I cosine exponentiated-X(NTICEx-X) family. If \( v(t) \) follows (5) and setting
\[
W[F(x)] = -\log \left( \frac{\cos \left( \frac{\pi}{2} F(x; \xi)^a \right)}{1 - \log \left( 1 - F(x; \xi)^a \right)} \right)
\]
in (1), we define the cdf of the NTICEx-X family via
\[
G(x) = 1 - \frac{\cos \left( \frac{\pi}{2} F(x; \xi)^a \right)}{1 - \log \left( 1 - F(x; \xi)^a \right)}, \quad a > 0, x \in \mathbb{R},
\]
with pdf
\[
g(x) = \frac{af(x; \xi) F(x; \xi)^{a-1}}{\left( 1 - \log \left[ 1 - F(x; \xi)^a \right] \right)^2} \times \left\{ \theta \left[ 1 - \log \left[ A(x) \right] \right] + 1 - \theta F(x; \xi)^a \right\}, \quad x \in \mathbb{R}.
\]
For \( a = 1 \), the NTICEx-X family gives the new type-I cosine-X(NTIC-X) family.

### 2.3. Type-I cosine exponentiated-X family

Taking inspiration from (2), we introduce a new method of proposing distributions. The proposed method is called type-I cosine exponentiated-X(TICEx-X) family of distributions. If \( v(t) \) follows (5) and setting
\[
W[F(x)] = -\log \left( \frac{\cos \left( \frac{\pi}{2} F(x; \xi)^a \right)}{1 - \log \left( 1 - F(x; \xi)^a \right)} \right)
\]
in (2), we define the cdf of the TICEx-X family given by
\[
G(x) = \frac{\cos \left( \frac{\pi}{2} F(x; \xi)^a \right)}{1 - \log \left( 1 - F(x; \xi)^a \right)}, \quad a > 0, x \in \mathbb{R},
\]
The density function corresponding to (8) is
\[
g(x) = \frac{af(x; \xi) F(x; \xi)^{a-1}}{\left( 1 - \log \left[ 1 - F(x; \xi)^a \right] \right)^2} \times \left\{ \theta \left[ 1 - \log \left[ A(x) \right] \right] + 1 - \theta F(x; \xi)^a \right\}, \quad a > 0, x \in \mathbb{R}.
\]
For \( a = 1 \), the TICEx-X family gives the new Type-I cosine-X(TIC-X) family.

### 2.4. The exponent power exponentiated-X family

Take inspiration from (3), we introduce a new method of proposing probability distributions, called the exponent power exponentiated-X (EPEX-X) family of distributions. The cdf of the EPEX-X family is defined by the following expression
\[
G(x) = e^{\left[ \frac{1 - \log[1 - F(x; \xi)^a]}{e - 1} \right]} - 1, \quad a, \theta > 0, x \in \mathbb{R}.
\]
The pdf corresponding to (9) is
\[
g(x) = \frac{a\theta f(x; \xi) \left[ 1 - \log \left[ e - 1 - F(x; \xi)^a \right] \right]}{(e - 1) \left( 1 - F(x; \xi)^a \right)^2 \left[ \theta - \log \left( 1 - F(x; \xi)^a \right) \right]^2}, \quad x \in \mathbb{R}.
\]
For \( a = 1 \), in (9), we introduce a sub-case of the EPEX-X family, called the exponent power-X (EP-X) family by cdf
\[
G(x) = e^{\left[ \frac{1 - \log[1 - F(x; \xi)^a]}{e - 1} \right]} - 1, \quad \theta > 0, x \in \mathbb{R}.
\]
The pdf corresponding to (10) is
\[
g(x) = \frac{\theta f(x; \xi) e^{\left[ \frac{1 - \log[1 - F(x; \xi)^a]}{e - 1} \right]}}{(e - 1) \left( 1 - F(x; \xi)^a \right)^2 \left[ \theta - \log \left( 1 - F(x; \xi)^a \right) \right]^2}, \quad x \in \mathbb{R}.
\]
For \( \theta = 1 \), the EP-X family gives the new reduced exponent power-X (REP-X) family.

### 3. The exponent power-Weibull distribution: a new heavy-tailed distribution

In this section, we define a special sub-model of the exponent power-X family, called exponent power-Weibull (EP-W) distribution. Let \( F(x; \xi) \) be the cdf of the Weibull distribution given by \( F(x; \xi) = 1 - e^{-\gamma x^\alpha} \),
Some possible shapes for the pdf of the EP-W distribution, for selected values of the model parameters, are sketched in Figure 1.

4. Statistical properties

In this section, we study some statistical properties of the EP-W distribution such as quantile function, moments and moment generating function (mgf). Some key advantages of the EP-W distribution are:

- The mean and variance of the distribution are in explicit forms and simple to calculate.
- The quantile function of the distribution has a closed form which makes it easier to generate random numbers.

4.1. Quantile function

Let \( X \) be the random variable follow the EP-W distribution with cdf (12). Then, the quantile function (qf) of \( X \), say \( x_u = Q(u) = G^{-1}(u) \), can be obtained by inverting (12) as follows

\[
x_u = Q(u) = \left( \left[ \frac{\log (u (e - 1) + 1)}{\gamma} \right]^{\frac{1}{\theta}} - 1 \right)^{\frac{1}{\gamma}}.
\]

(14)

where \( u \in (0,1) \). The qf of the EP-W distribution has a closed form expression. In particular, the first quartile, second quartile (median) and third quartile are obtained by substituting \( u = 0.25, 0.5 \) and 0.75 in (14), respectively.

4.2. Moments

Some of the most important features and characteristics of a model can be obtain through its moments. The \( r \)th moment of a random variable \( X \), say \( \mu_r \), with pdf (13) is derived as

\[
\mu_r = \int_0^\infty \frac{a \gamma x^{r-1}}{(e - 1) (\theta + \gamma x^\alpha)} e^{\left(\frac{x^\alpha}{\gamma x^\alpha}\right)} dx.
\]

(15)

On solving, we have

\[
\mu_r = \sum_{i=0}^{\infty} \frac{1}{\eta^i (e - 1) i!} \int_0^\infty \frac{t^{(\frac{r}{\alpha}+1)+1}}{(1+t)^{(\frac{r}{\alpha}+1)+\left(\frac{\gamma}{\theta}\right)}} dt,
\]

where \( \eta = \gamma / \theta \). Using the definition of Beta type-II distribution, we have

\[
\mu_r = \frac{\sum_{i=0}^{\infty} B \left( \frac{r}{\alpha}+i+1, \frac{\gamma}{\theta} \right)}{\eta^i (e - 1) i!}.
\]

(16)

For \( r = 1 \), in (16), we get the first raw moment (mean) of the EP-W distribution given by

\[
\mu_1 = \frac{\sum_{i=0}^{\infty} B \left( \frac{1}{\alpha}+i+1, \frac{\gamma}{\theta} \right)}{\eta^i (e - 1) i!}.
\]

(17)

Similarly, the second raw moment of the EP-W distribution is given by

\[
\mu_2 = \frac{\sum_{i=0}^{\infty} B \left( \frac{2}{\alpha}+i+1, \frac{\gamma-2}{\theta} \right)}{\eta^i (e - 1) i!}.
\]

(18)

Now, the variance of the proposed distribution is obtained by using the following relation

\[
\mu_2 = \mu_1^2 - \left( \mu_1 \right)^2.
\]

(19)
distribution, \(MX(t)\), is given by
\[
MX(t) = \sum_{r, i=0}^{\infty} \frac{t^r}{\eta^r (e - 1)^i} \frac{1}{r!} \times B \left( \frac{r}{\alpha} + i + 1, \left( \frac{\alpha - r}{\alpha} \right) \right).
\]

The graphs of skewness and kurtosis of the EP-W distribution along with the corresponding contour plots are presented in Figures 2 and 3, respectively.

From Figures 2 and 3, it is clear that for fixed values of \(\theta\), the skewness is decreased when \(\alpha\) is increased. (ii) For fixed values of \(\alpha\), the skewness is increased when \(\theta\) is increased. (iii) For fixed values of \(\theta\), the kurtosis is decreased when \(\alpha\) is increased.

5. Characterization results

In designing a stochastic model for a particular data set, an investigator will be vitally interested to know if their selected model fits the necessary requirements. To this end, the investigator will rely on the characterizations of the selected distribution. Thus, the problem of characterizing a distribution is important in various fields and has recently attracted the attention of many researchers. Consequently, various characterization results have been reported in the literature. These characterizations have been established in different directions. This section is devoted to certain characterizations of the EP-X distribution based on a simple relationship between two truncated moments. Due to the nature of the cdf of this distribution, our characterizations may be the only possible ones. The first characterization result employs Theorem 1 of Glänzel [26], given in Appendix. As shown in Glänzel [27], this characterization is stable in the sense of weak convergence and can also be employed when the cdf does not have a closed form. We would like to mention that the goal of Theorem A.1 is to make \(\eta(x)\) as simple as possible.

**Proposition 5.1:** Suppose \(X\) is a continuous random variable. Let \(q_1(x) = \exp \left\{ \frac{\theta}{e - 1 - F(x; \xi)} - 1 \right\} \) and \(q_2(x) = q_1(x) \left( \theta - \log (1 - F(x; \xi)) \right)^{-1} \) for \(x \in \mathbb{R} \). Then \(X\) has density function (11) if and only if the function \(\eta\), defined in Theorem A.1, is given by
\[
\eta(x) = \frac{1}{2} \left( \theta - \log \left(1 - F(x; \xi)\right) \right)^{-1}, \quad x \in \mathbb{R}.
\]
Proof: Suppose $X$ is a random variable with density function (11), then we have

$$(1 - G(x))E\left[q_1(X) \mid X \geq x\right]$$

$$= \frac{\theta}{e - 1} [\theta - \log (1 - F(x; \xi))]^{-1}, \quad x \in \mathbb{R},$$

$$(1 - G(x))E\left[q_2(X) \mid X \geq x\right]$$

$$= \frac{\theta}{2(e - 1)} [\theta - \log (1 - F(x; \xi))]^{-2}, \quad x \in \mathbb{R},$$

and

$$\eta(x) q_1(x) - q_2(x)$$

$$= -\frac{1}{2} q_1(x) [\theta - \log (1 - F(x; \xi))]^{-1} < 0$$

for $x \in \mathbb{R}$.

Conversely, if $\eta$ is of the above form, then

$$s'(x) = \frac{\eta'(x) q_1(x)}{\eta(x) q_1(x) - q_2(x)}$$

$$= f(x; \xi) [\theta - \log (1 - F(x; \xi))]^{-1}, \quad x \in \mathbb{R},$$

and consequently

$$s(x) = \log ([\theta - \log (1 - F(x; \xi))]), \quad x \in \mathbb{R}.$$

Now, according to Theorem A.1, $X$ has pdf (11).

Corollary 5.1: Suppose $X : \Omega \rightarrow \mathbb{R}$ is a continuous random variable and $q_1(x)$ is as in Proposition 5.1. Then $X$ has density function (11) if and only if there exist functions $q_2$ and $\eta$ defined in Theorem A.1 satisfying the following first-order differential equation

$$\eta'(x) q_1(x) = \frac{f(x; \xi) [\theta - \log (1 - F(x; \xi))]^{-1}}{1 - F(x; \xi)}, \quad x \in \mathbb{R}.$$

Corollary 5.2: The differential equation in Corollary 5.1 has the following general solution

$$\eta(x) = \left[\int \frac{f(x; \xi)}{1 - F(x; \xi)} (q_1(x))^{-1} q_2(x) \, dx + D\right]^{-1}$$

in which $D$ is a constant. We like to mention that a set of functions satisfying the above first-order differential equation is given in Proposition 5.1 with $D = 0$. Clearly, there are other triplets $(q_1, q_2, \eta)$ satisfying the conditions of Theorem A.1.

6. Maximum likelihood estimation and Monte Carlo simulation study

In this section, we use the maximum likelihood method to estimate model parameters and also provide a Monte Carlo simulations to assess the behaviour of these estimators.

6.1. Maximum likelihood estimation

In this sub-section, we obtain the maximum likelihood estimators (MLEs) of the model parameters of the EP-W distribution from complete samples only. Let $x_1, x_2, \ldots, x_n$ be an observed sample of size $n$ obtained from (13). The corresponding log-likelihood function can be expressed as

$$\log L(x; \alpha, \theta, \gamma) = -n \log (e - 1) + n \log (\alpha) + n \log (\theta) + \sum_{i=1}^{n} \log (x_i)$$

and

$$-2 \sum_{i=1}^{n} \left(\theta + \gamma x_i^2\right) + \sum_{i=1}^{n} \frac{\gamma x_i^2}{\theta + \gamma x_i^2}. $$

The partial derivatives of (21) with respect to the parameters $(\alpha, \theta, \gamma)$ are given, respectively, by

$$\frac{\partial \log L(x; \alpha, \theta, \gamma)}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} \log (x_i)$$

$$+ n \sum_{i=1}^{n} \frac{(\theta + \gamma x_i^2)(\log x_i)}{(\theta + \gamma x_i^2)^2}$$

$$- 2 \gamma \sum_{i=1}^{n} \frac{(\log x_i) x_i^2}{(\theta + \gamma x_i^2)},$$

$$\frac{\partial \log L(x; \alpha, \theta, \gamma)}{\partial \theta} = -2 \sum_{i=1}^{n} \frac{1}{\theta + \gamma x_i^2},$$

and

$$\frac{\partial \log L(x; \alpha, \theta, \gamma)}{\partial \gamma} = \frac{n}{\gamma} - 2 \sum_{i=1}^{n} \frac{x_i^2}{(\theta + \gamma x_i^2)}$$

$$+ \sum_{i=1}^{n} \frac{\theta x_i^2}{(\theta + \gamma x_i^2)^2}. $$

Solving

$$\frac{\partial \log L(x; \alpha, \theta, \gamma)}{\partial \alpha} = 0,$$

$$\frac{\partial \log L(x; \alpha, \theta, \gamma)}{\partial \theta} = 0$$

and

$$\frac{\partial \log L(x; \alpha, \theta, \gamma)}{\partial \gamma} = 0,$$

simultaneously yields the MLEs $(\hat{\alpha}, \hat{\theta}, \hat{\gamma})$ of $(\alpha, \theta, \gamma)$.

6.2. Monte Carlo simulation study

In this sub-section, we perform a Monte Carlo (MC) simulation study with the objective to assess the behaviour
Table 1. Simulation results for the EP-W distribution.

| n | parameters | MLE | MSEs | Biases | MLE | MSEs | Biases |
|---|------------|-----|------|--------|-----|------|--------|
| 25 | \(\hat{\alpha}\) | 1.45781 | 0.07176 | 0.05781 | 1.68430 | 0.09778 | 0.08430 |
|    | \(\hat{\theta}\) | 0.56975 | 0.03086 | 0.06975 | 1.29588 | 0.08168 | 0.09588 |
|    | \(\hat{\gamma}\) | 1.11005 | 0.05338 | 0.11005 | 1.06545 | 0.07545 | 0.06545 |
| 50 | \(\hat{\alpha}\) | 1.44211 | 0.03493 | 0.04211 | 1.64965 | 0.04159 | 0.04965 |
|    | \(\hat{\theta}\) | 0.54253 | 0.01148 | 0.04253 | 1.27387 | 0.03730 | 0.07387 |
|    | \(\hat{\gamma}\) | 1.09308 | 0.02294 | 0.09308 | 1.03817 | 0.02547 | 0.03817 |
| 100 | \(\hat{\alpha}\) | 1.41918 | 0.01554 | 0.01918 | 1.61664 | 0.02004 | 0.01664 |
|    | \(\hat{\theta}\) | 0.52919 | 0.00501 | 0.02919 | 1.25112 | 0.01628 | 0.05112 |
|    | \(\hat{\gamma}\) | 1.06569 | 0.01336 | 0.06569 | 1.03636 | 0.01294 | 0.03636 |
| 300 | \(\hat{\alpha}\) | 1.40481 | 0.00480 | 0.00481 | 1.60327 | 0.00360 | 0.00237 |
|    | \(\hat{\theta}\) | 0.51642 | 0.00127 | 0.01642 | 1.23301 | 0.00566 | 0.03301 |
|    | \(\hat{\gamma}\) | 1.03382 | 0.00366 | 0.03382 | 1.01953 | 0.00385 | 0.01953 |
| 500 | \(\hat{\alpha}\) | 1.40375 | 0.00279 | 0.00279 | 1.60237 | 0.00360 | 0.00237 |
|    | \(\hat{\theta}\) | 0.51185 | 0.00078 | 0.01185 | 1.22082 | 0.00278 | 0.02082 |
|    | \(\hat{\gamma}\) | 1.02895 | 0.00242 | 0.02895 | 1.01908 | 0.00246 | 0.01908 |
| 700 | \(\hat{\alpha}\) | 1.40467 | 0.00211 | 0.00211 | 1.60493 | 0.00254 | 0.00493 |
|    | \(\hat{\theta}\) | 0.51105 | 0.00057 | 0.01105 | 1.21818 | 0.00206 | 0.01818 |
|    | \(\hat{\gamma}\) | 1.02214 | 0.00160 | 0.02214 | 1.01624 | 0.00176 | 0.01623 |
| 750 | \(\hat{\alpha}\) | 1.40089 | 0.00192 | 0.00192 | 1.60026 | 0.00237 | 0.0026 |
|    | \(\hat{\theta}\) | 0.51094 | 0.00053 | 0.01094 | 1.21943 | 0.00203 | 0.01943 |
|    | \(\hat{\gamma}\) | 1.02156 | 0.00139 | 0.02156 | 1.01371 | 0.00153 | 0.01371 |

Figure 4. Plots of MLEs and MSEs of the EP-W model for \(\alpha = 1.4, \theta = 0.5\) and \(\gamma = 1\).

Figure 5. Plots of Biases and absolute biases of the EP-W model for \(\alpha = 1.4, \theta = 0.5\) and \(\gamma = 1\).
of the MLEs of EP-W model via the \texttt{optim()} R-function with the argument \texttt{method = "L-BFGS-B"}; see Appendix. It is used for maximizing the log-likelihood function of a probabilistic model. We consider 750 MC-replicates under different sample sizes $n = 25, 50, \ldots, 750$. For each sample size, we compute the average MLEs, mean square errors (MSE), Biases and Absolute biases. The results obtained after performing the MC simulation are provided in Table 1 and displayed graphically in Figures 4–7.

7. Actuarial measures

One of the most important tasks of financial and actuarial sciences institutions is to evaluate the exposure to market risk in a portfolio of instruments, which arise from changes in underlying variables such as prices of equity, interest rates or exchange rates. In this section, we calculate some important risk measures including value at risk (VaR), tail value at risk (TVaR), tail variance (TV) and tail variance premium (TVP) for the proposed distribution, which play a crucial role in portfolio optimization under uncertainty.

7.1. Value at risk

In the context of actuarial sciences, the VaR is widely used by practitioners as a standard financial market risk measure. It is also known as the quantile risk measure or quantile premium principle. The VaR is always specified with a given degree of confidence say $q$ (typically 90%, 95% or 99%), and represent the percentage loss in portfolio value that will be equaled or exceeded only $X\%$ of the time. VaR of a random variable $X$ is the $q$th quantile of its cdf, see Artzner [28]. If $X$ follows the EP-W distribution, then its VaR is

$$x_q = \left\{ \left( \frac{\gamma}{\theta} \left[ \log (q (\epsilon - 1))^{-1} - 1 \right] \right)^{-\frac{1}{\alpha}} \right\} \frac{1}{\alpha}. \quad (25)$$
7.2. Tail value at risk

Another important measure is TVaR, also known as conditional tail expectation (CTE) or tail conditional expectation (TCE), used to quantify the expected value of the loss given that an event outside a given probability level has occurred. If \( X \) follows the EP-W distribution, then its TVaR is derived as

\[
\text{TVaR}_q (X) = \frac{1}{1 - q} \int_{\text{VaR}_q}^{\infty} x \, g(x; \theta, \xi) \, dx. \tag{26}
\]

Using (13) in (26), we get

\[
\text{TVaR}_q (X) = \frac{\alpha \theta^2}{(1 - q) (e - 1)} \sum_{i = 0}^{\infty} \frac{\gamma^{i + 1}}{i!} \times \int_{\text{VaR}_q}^{\infty} x^{\alpha i + 1 + \alpha - 1} \, dx.
\]

Using the series \( e^x = \sum_{i = 0}^{\infty} \frac{x^i}{i!} \), we have

\[
\text{TVaR}_q (X) = \frac{\alpha \theta^2}{(1 - q) (e - 1)} \sum_{i = 0}^{\infty} \frac{\gamma^{i + 1}}{i!} \times \int_{\text{VaR}_q}^{\infty} x^{\alpha i + 1 + \alpha - 1} \, dx.
\]

Let \( \eta = \frac{\gamma}{\theta} \), then from (27), we have

\[
\text{TVaR}_q (X) = \frac{\alpha}{(1 - q) (e - 1)} \sum_{i = 0}^{\infty} \frac{\eta^{i + 1}}{i!} \times \int_{\text{VaR}_q}^{\infty} x^{\alpha i + 1 + \alpha - 1} \, dx,
\]

\[
\text{TVaR}_q (X) = \frac{1}{(1 - q) (e - 1)} \sum_{i = 0}^{\infty} \frac{\eta^i}{(\eta^{\frac{i + 1}{\alpha}})^{\frac{i + 1}{\alpha}}} \times \int_{\text{VaR}_q}^{\infty} \frac{t^{\frac{i + 1}{\alpha}}}{(1 + t)^{\frac{i + 1}{\alpha}}} \, dt,
\]

\[
\text{TVaR}_q (X) = \frac{1}{(1 - q) (e - 1)} \sum_{i = 0}^{\infty} \frac{1}{(\eta^{\frac{i}{\alpha}})^{\frac{i}{\alpha}}} \times \int_{\text{VaR}_q}^{\infty} \frac{t^{\frac{i + 1}{\alpha} - 1}}{(1 + t)^{\frac{i + 1}{\alpha} + (\frac{i + 1}{\alpha})}} \, dt.
\]

Finally, we get

\[
\text{TVaR}_q (X) = \frac{1}{(1 - q) (e - 1)} \sum_{i = 0}^{\infty} \frac{1}{(\eta^{\frac{i}{\alpha}})^{\frac{i}{\alpha}}} \times B \left( \frac{1}{\alpha} + i + 1, \frac{\alpha - 2}{\alpha} \right) - \text{TVaR}_q ^2. \tag{28}
\]

7.3. Tail variance

The TV is one of the most important actuarial measures which considers the tail variance beyond the VaR. The TV of a EP-W distributed random variable is

\[
\text{TV}_q (X) = E (X^2 | X > x_q) - (\text{TVaR}_q)^2. \tag{29}
\]

Consider

\[
E (X^2 | X > x_q) = \frac{\alpha \theta^2}{(1 - q) (e - 1)} \times \int_{\text{VaR}_q}^{\infty} x^{\alpha^2 + 1 + \alpha - 1} \, dx.
\]

On solving, we get

\[
E (X^2 | X > x_q) = \frac{1}{(1 - q) (e - 1)} \sum_{i = 0}^{\infty} \frac{1}{(\eta^\frac{i}{\alpha})^\alpha} \times \int_{\text{VaR}_q}^{\infty} \frac{t^{\frac{i + 1}{\alpha} - 1}}{(1 + t)^{\frac{i + 1}{\alpha} + (\frac{i + 1}{\alpha})}} \, dt.
\]

Finally, we have

\[
E (X^2 | X > x_q) = \frac{1}{(1 - q) (e - 1)} \sum_{i = 0}^{\infty} \frac{1}{(\eta^\frac{i}{\alpha})^\alpha} \times B \left( \frac{2}{\alpha} + i + 1, \frac{\alpha - 2}{\alpha} \right) \times \left\{ 1 - F \left( \frac{\gamma (\text{VaR}_q)^\alpha}{1 + \gamma (\text{VaR}_q)^\alpha}, \frac{2}{\alpha} + i + 1, \frac{\alpha - 2}{\alpha} \right) \right\}. \tag{30}
\]

Using (28) and (30) in (29), we get the expression for the TV of the EP-W distribution.

7.4. Tail variance premium

The TVP is another important measure playing an essential role in insurance sciences. The TVP of the EP-W distributed random variable is

\[
\text{TVP}_q (X) = \text{TVaR}_q + \delta \text{TVaR}_q. \tag{31}
\]

where \( 0 < \delta < 1 \). Using the expression (28) and (29) in (31), we get the tail variance premium of the proposed distribution.
7.5. Numerical study of the risk measures

In this sub-section, we provide numerical study of the VaR, TVaR, TV and TVP for the Weibull and EP-W distributions for different sets of parameters. The process is described below:

- Random sample of size \( n = 100 \) are generated from the Weibull and EP-W distributions and parameters have been estimated via the maximum likelihood method.
- 1000 repetitions are made to calculate the VaR, TVaR, TV and TVP for these distributions.

The numerical results of the risk measures are provided in Tables 2 and 3 and displayed graphically in Figures 8–11 corresponding to each table.

| Table 2. Simulation results of the actuarial measures for the selected values of the parameters. |
| Dist. Parameters | Level of significance | VaR | TVaR | TV | TVP |
|-------------------|-----------------------|------|------|---|-----|
| Weibull           |                       |      |      |   |     |
| \( \alpha = 2.1 \) | 0.700                 | 1.5200| 1.9853| 0.1513| 2.0232 |
| \( \gamma = 0.5 \) | 0.800                 | 1.7452| 2.1641| 0.1290| 2.1964 |
| \( \alpha = 2.1 \) | 0.850                 | 1.8874| 2.2808| 0.1170| 2.3101 |
| \( \gamma = 0.5 \) | 0.900                 | 2.0698| 2.4342| 0.1035| 2.4601 |
| \( \alpha = 2.1 \) | 0.950                 | 2.3461| 2.6732| 0.0867| 2.6948 |
| \( \gamma = 0.5 \) | 0.975                 | 2.5905| 2.8898| 0.0746| 2.9084 |
| \( \alpha = 2.1 \) | 0.999                 | 3.4924| 3.7179| 0.0453| 3.7293 |
| EP-W              |                       |      |      |   |     |
| \( \alpha = 2.1 \) | 0.700                 | 2.3150| 4.8127| 78.0599| 24.3277 |
| \( \gamma = 0.5 \) | 0.800                 | 2.9649| 5.9132| 113.4390| 34.2730 |
| \( \alpha = 2.1 \) | 0.850                 | 3.4806| 6.8156| 147.9874| 43.8125 |
| \( \gamma = 0.5 \) | 0.900                 | 4.3091| 8.2954| 215.3835| 62.1413 |
| \( \alpha = 2.1 \) | 0.950                 | 6.0887| 11.5370| 409.4976| 113.9114 |
| \( \gamma = 0.5 \) | 0.975                 | 8.4974| 15.9801| 779.0482| 210.7422 |
| \( \alpha = 2.1 \) | 0.999                 | 38.2903| 71.3062| 15,474.6648| 3940.1724 |

| Table 3. Simulation results of the actuarial measures for the selected values of the parameters. |
| Dist. Parameters | Level of significance | VaR | TVaR | TV | TVP |
|-------------------|-----------------------|------|------|---|-----|
| Weibull           |                       |      |      |   |     |
| \( \alpha = 2.5 \) | 0.700                 | 0.9985| 1.2467| 0.0410| 1.2569 |
| \( \gamma = 1.2 \) | 0.800                 | 1.1214| 1.3412| 0.0341| 1.3497 |
| \( \alpha = 2.5 \) | 0.900                 | 1.2942| 1.4814| 0.0264| 1.4880 |
| \( \gamma = 1.2 \) | 0.950                 | 1.4378| 1.6032| 0.0215| 1.6085 |
| \( \alpha = 2.5 \) | 0.975                 | 1.5626| 1.7120| 0.0181| 1.7165 |
| \( \gamma = 1.2 \) | 0.999                 | 2.0083| 2.1163| 0.0102| 2.1188 |
| EP-W              |                       |      |      |   |     |
| \( \alpha = 2.5 \) | 0.700                 | 1.8512| 3.3312| 9.0836| 5.6021 |
| \( \theta = 1.5 \) | 0.800                 | 2.0411| 3.6089| 10.4370| 6.2182 |
| \( \gamma = 1.2 \) | 0.900                 | 2.3805| 3.9720| 12.359 | 7.0685 |
| \( \theta = 1.5 \) | 0.950                 | 4.2013| 7.0116| 36.4278| 16.1186 |
| \( \gamma = 1.2 \) | 0.975                 | 5.5728| 9.2439| 62.7387| 24.9286 |
| \( \alpha = 2.5 \) | 0.999                 | 19.9616| 32.9239| 789.0310| 230.1817 |

Figure 8. Graphical sketching of the results of VaR and TVaR provided in Table 3.
The simulation is performed for the Weibull and EP-W for the selected values of parameters. A model with higher values of the Risk measures is said to have a heavier tail. The simulated results provided in Tables 3 and 4 show that the proposed EP-W model has higher values of the risk measures than the traditional Weibull distribution. The simulation results are graphically displayed in Figures 8–11, which show that the proposed model has heavier tail than the Weibull distribution.
8. An application to a medical care insurance data set

In this section, we use two insurance data sets to illustrate the importance and flexibility of the proposed distribution. The comparison of the proposed distribution is made with a nested model, the Weibull distribution, and with some other non-nested models such as the generalized exponential (GE), Lomax, Burr, exponentiated Weibull (EW), generalized odd Burr III-Weibull (GOBXIII-W) and generalized log-Moyal (GLM) distributions. It is important to emphasize that the EW distribution is a popular model for analysing data in the applied areas; see Mudholkar and Srivastava [29]. The GE is another non-nested model and offers the characteristics of the Weibull and gamma distributions; see Gupta and Kundu [30]. The Lomax and Burr-XII (B-XII) distributions are prominent competitors for modelling large losses and offer a wide range of applications in financial sciences. The proposed EP-W model is also compared with the GLM distribution of Bhati and Ravi [1]. We also considered the five-parameter non-nested GOBXIII-W distribution of Haq et al. [31]. The cdfs of the competing distributions are:

- The EW distribution
  \[ G(x; \alpha, \gamma) = \left(1 - e^{-x^\gamma}\right)^\alpha, \quad x \geq 0, \alpha, \gamma > 0. \]

- The GE distribution
  \[ G(x; \alpha, \gamma) = \left(1 - e^{-x^\gamma}\right)^\alpha, \quad x \geq 0, \alpha, \gamma > 0. \]

- The Lomax distribution
  \[ G(x; \alpha, \gamma) = 1 - (1 + x^\gamma)^{-\alpha}, \quad x \geq 0, \alpha, \gamma > 0. \]

- The B-XII distribution
  \[ G(x; c, k) = 1 - (1 + x^c)^{-k}, \quad x \geq 0, c, k > 0. \]

- The GOBXIII-W distribution
  \[ G(x; \alpha, \gamma, \beta, c, k) = \left[ 1 + \left( \frac{1 - (1 - e^{-x^\gamma})^\beta}{1 - e^{-x^\gamma}} \right)^c \right]^{-k}, \quad x \geq 0, \alpha, \gamma, \beta, c, k > 0. \]

- The GLM distribution
  \[ G(x; \alpha, \gamma) = \text{erf}\left( \frac{1}{\sqrt{2}} \left( \frac{\gamma}{x} \right)^{1/2} \right), \quad x \geq 0, \alpha, \gamma > 0. \]

Next, we consider certain analytical measures in order to verify which distribution fits better the considered data. These analytical measures include (i) discrimination measures such as Akaike information criterion (AIC), Bayesian information criterion (BIC), Hannan-Quinn information criterion (HQIC), Consistent Akaike information criterion (CAIC), and (ii) two other goodness of fit measures including Anderson–Darling (AD) test statistic and Cramer-von Mises test statistic. The discrimination measures are given by

- The AIC is
  \[ AIC = 2k - 2l, \]

- The BIC is
  \[ BIC = k \log(n) - 2l, \]

- The HQIC
  \[ HQIC = 2k \log(\log(n)) - 2l, \]

- The CAIC is
  \[ CAIC = \frac{2nk}{n - k - 1} - 2l, \]

where \( l \) denotes the log-likelihood function evaluated at the MLEs, \( k \) is the number of model parameters and \( n \) is the sample size. The considered goodness-of-fit measures are given by

| Dist.    | \( \hat{\alpha} \) | \( \hat{\gamma} \) | \( \hat{\beta} \) | \( \hat{\alpha} \) | \( \hat{\beta} \) | \( \hat{\gamma} \) | \( \hat{k} \) |
|----------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| EP-W     | 2.353           | 0.466           | 3.488           | (0.0870)        | (0.2007)        | (0.7362)        |
| Weibull  | 1.436           | 0.132           |                 | (0.0458)        | (0.0119)        |                 |
| EW       | 0.924           | 0.721           | 4.813           | (0.2287)        | (0.0888)        | (1.4897)        |
| GE       |                 |                 | 1.247           |                 | 0.908           | (1.6334)        |
| Lomax    | 2.548           | 3.927           | 3.927           | (1.1160)        |                 | (1.5244)        |
| B-XII    |                 |                 |                 |                 |                 | 0.261           | 3.462           |
| GOBXIII-W| 2.905           | 0.975           | 1.736           | (1.7563)        | (0.8936)        | (2.0955)        |
| GLM      | 1.487           | 0.869           | 1.084           | (1.487)         | (0.9064)        | (2.1094)        |
The AD test statistic is given by
\[
AD = -n - \frac{1}{n} \sum_{i=1}^{n} (2i - 1) \times \left[ \log G(x_i) + \log \{1 - G(x_{n-i+1})\} \right],
\]
where \( n \) is the sample size and \( x_i \) is the \( i \)th observation in the sample, calculated when the data is sorted in ascending order.

The CM test statistic is given by
\[
CM = \frac{1}{12n} + \sum_{i=1}^{n} \left[ \frac{2i - 1}{2n} - G(x_i) \right]^2.
\]

Table 5. Discrimination and goodness of fit measures of the competing models.

| Dist.       | AIC       | BIC       | CAIC      | HQIC      | CM       | AD       |
|------------|-----------|-----------|-----------|-----------|----------|----------|
| EP-W       | 2288.154  | 2300.950  | 2288.200  | 2293.164  | 0.280    | 2.103    |
| Weibull    | 2321.130  | 2329.660  | 2321.160  | 2324.477  | 0.943    | 6.249    |
| EW         | 2291.085  | 2304.880  | 2292.231  | 2297.095  | 0.359    | 2.520    |
| GE         | 2312.160  | 2320.956  | 2312.206  | 2315.170  | 0.393    | 2.985    |
| Lomax      | 2447.169  | 2455.699  | 2447.192  | 2450.509  | 0.437    | 3.451    |
| B-XII      | 2504.136  | 2512.667  | 2504.159  | 2507.477  | 2.744    | 6.786    |
| GOBXIII-W  | 2302.647  | 2315.709  | 2303.634  | 2311.896  | 2.037    | 3.624    |
| GLM        | 2298.098  | 2310.309  | 2299.091  | 2303.065  | 0.385    | 2.704    |

For the optimization and calculation of the analytical measures, we use the `optim()` R-function with the argument `method = "BFGS"`; see Appendix.

8.1. Data 1: medical care insurances

Here, we illustrate the EP-W distribution by analysing a heavy-tailed data set representing medical care insurances. The data set is available at https://data.world/datasets/insurance. The maximum likelihood estimates with the standard error (in parentheses) of the fitted models for the analysed data are presented in Table 4. The analytical measures of the EP-W and other considered models are provided in Table 5. Based on the above-mentioned measures, the proposed model fits the heavy-tailed medical care insurance data set better than the other models. In support of the results provided in Table 5, the estimated pdf and cdf of the proposed distribution are plotted in Figure 12. The PP and Kaplan–Meier survival plots are sketched in Figure 13. The proposed model may be an interesting alternative to the other existing models in the literature for modelling positively skewed heavy tail data. The practical application to the medical care insurance data shows...
Table 6. Estimated values with standard error (in parentheses) of the competing models for data 2.

| Dist.   | $\hat{\alpha}$ | $\hat{\gamma}$ | $\hat{\theta}$ | $\hat{\beta}$ | $\hat{\xi}$ | $\hat{k}$ |
|---------|----------------|----------------|----------------|--------------|----------|---------|
| EP-W    | 1.390          | 0.002          | 2.517          |              |          |         |
| Weibull | 1.019          | 0.003          |                |              |          |         |
| EW      | 0.889          | 0.008          | 1.476          | (0.4897)     |          |         |
| GE      | 1.234          | 0.604          |                |              |          |         |
| Lomax   | 0.428          | 19.267         | 0.056          | 3.459        |         |         |
| Burr-XII|                |               | (1.1479)       |             | (7.6205) |         |
| GOBXIII-W| 1.209         | 0.042          | 0.054          | 503.460      |         |         |
| GLM     | 0.854          | 1.238          | 0.049          | 0.107        |         |         |
|         | (0.7498)       |               | (0.3848)       | (0.09048)    |         |         |

Table 7. Discrimination and goodness of fit measures of the competing models for data 2.

| Dist.  | AIC  | BIC  | CAIC | HQAIC | CM  | AD  |
|--------|------|------|------|-------|-----|-----|
| EP-W   | 425.490 | 429.887 | 426.347 | 426.947 | 0.044 | 0.115 |
| Weibull| 432.353 | 439.256 | 434.089 | 435.746 | 0.054 | 0.447 |
| EW     | 430.226 | 435.623 | 431.083 | 432.683 | 0.049 | 0.183 |
| GE     | 433.924 | 440.855 | 435.338 | 438.895 | 0.056 | 0.462 |
| Lomax  | 466.167 | 469.099 | 466.581 | 467.139 | 0.204 | 1.215 |
| Burr-XII| 503.460 | 506.391 | 503.873 | 504.431 | 0.228 | 1.362 |
| GOBXIII-W| 431.220 | 438.618 | 433.897 | 434.678 | 0.054 | 0.359 |
| GLM    | 429.908 | 433.097 | 430.654 | 430.324 | 0.051 | 0.259 |

that the proposed distribution should be taken into account among other possible distributions for insurance data in which the properties of a heavy-tailed distribution are present.

8.2. Data 2: vehicle insurance losses

The second data set, available at https://data.world/datasets/insurance, refers to vehicle insurance losses. The competing models are also applied to these data. The estimates with the standard error (in parentheses)

Figure 14. Estimated pdf and cdf plots of the EP-W distribution for the vehicle insurance loss data.

Figure 15. PP and Kaplan–Meier survival plots of the EP-W distribution for the vehicle insurance loss data.
Table 8. Bayes estimator and posterior risk under different loss functions.

| Loss function | Bayes estimator | Posterior risk |
|---------------|----------------|---------------|
| $L_1$ = SELF = $(\theta - d)^2$ | $E(\theta | x)$ | $Var(\theta | x)$ |
| $L_2$ = WSELF = $\frac{\mu - d^2}{\sigma^2}$ | $(E(\theta^{-1} | x))^{-1}$ | $E(\theta | x) - (E(\theta^{-1} | x))^{-1}$ |
| $L_3$ = MSELF = $(1 - \frac{d}{\theta})^2$ | $\frac{E(\theta^{-1} | x)}{\left(E(\theta^{-1} | x)\right)^2}$ | $1 - \frac{E(\theta^{-1} | x)^2}{\left(E(\theta^{-1} | x)\right)^2}$ |
| $L_4$ = PLF = $\frac{\mu - d^2}{\sigma^2}$ | $\sqrt{E(\theta^2 | x)}$ | $2 \left(\sqrt{E(\theta^2 | x)} - E(\theta | x)\right)$ |
| $L_5$ = KLF = $\left(\frac{\sigma^2 - \mu^2}{\sigma^2 | x}\right)$ | $\sqrt{E(\theta^2 | x)}$ | $2 \left(\sqrt{E(\theta^2 | x)E(\theta^{-1} | x)} - 1\right)$ |

Table 9. Bayesian estimates and their posterior risks of the parameters under different loss functions based on the medical care insurance data.

| Bayes loss functions | $\alpha$ Estimate | $\alpha$ Risk | $\gamma$ Estimate | $\gamma$ Risk | $\theta$ Estimate | $\theta$ Risk |
|----------------------|------------------|--------------|------------------|--------------|------------------|--------------|
| SELF                 | 2.3404           | 0.0065       | 0.4666           | 0.0174       | 3.4543           | 0.9733       |
| WSELF                | 2.3376           | 0.0028       | 0.4226           | 0.0440       | 3.1250           | 0.3293       |
| MSELF                | 2.3348           | 0.0012       | 0.3695           | 0.1257       | 2.7428           | 0.1223       |
| PLF                  | 2.3418           | 0.0028       | 0.4849           | 0.0366       | 3.5924           | 0.2762       |
| KLF                  | 2.3390           | 0.0012       | 0.4440           | 0.1015       | 3.2855           | 0.1028       |

Table 10. Credible and HPD intervals of the parameters $\alpha$, $\gamma$ and $\theta$ for the medical care insurance data.

| Parameters | Credible interval | HPD interval |
|------------|------------------|--------------|
| $\alpha$   | (2.284, 2.395)   | (2.201, 2.516) |
| $\gamma$   | (0.364, 0.566)   | (0.2507, 0.7267) |
| $\theta$   | (2.681, 4.222)   | (1.649, 5.197) |

Figure 16. Plots of Bayesian analysis and performance of Gibbs sampling for Insurance data set. Trace plots of each parameter of EP-W distribution.

9. Bayesian estimation

Bayesian inference procedure has been taken into consideration by many statistical researchers, especially those in the field of the survival analysis and reliability engineering. In this section, a complete sample data is analysed through Bayesian point of view. We assume that the parameters $\alpha$, $\gamma$ and $\theta$ of EP-W distribution have independent prior distributions as

$$\alpha \sim \text{Gamma}(a, b), \quad \gamma \sim \text{Gamma}(c, d), \quad \theta \sim \text{Gamma}(e, f).$$

The analytical measures of the EP-W and other considered models are provided in Table 6. We can see that the EP-W model again outperforms all the fitted competitive models under these statistics. For the second data set, the fitted density and distribution functions of the EP-W model are plotted in Figure 14. We note that the EP-W distribution best captures the fitted pdf and cdf. The PP and Kaplan–Meier survival plots of the EP-W model are shown in Figure 15. From Figure 15, we can see that the proposed model is closely followed by the PP and Kaplan–Meier survival plots.
Figure 17. Plots of Bayesian analysis and performance of Gibbs sampling for Insurance data set. Autocorrelation plots of each parameter of EP-W distribution.

Figure 18. Plots of Bayesian analysis and performance of Gibbs sampling for Insurance data set. Histogram plots of each parameter of EP-W distribution.
where \(a, b, c, d, e\) and \(f\) are positive. Hence, the joint prior density function is formulated as follow:

\[
\pi(\alpha, \gamma, \theta) = \frac{b^d d^e}{\Gamma(a) \Gamma(c) \Gamma(e)} \alpha^{a-1} \gamma^{c-1} \theta^{e-1} e^{-(ba + dy + e\theta)},
\]

(32)

In the Bayesian estimation the actual value of the parameter, may be adversely affected by the loss when choosing an estimator. This loss can be measured by a function of the parameter and the corresponding estimator. Five well-known loss functions and associated Bayesian estimators and corresponding posterior risk are presented in Table 8.

For more details see Calabria and Pulcini [32]. Next, we provide the posterior probability distribution for a complete data set. We define the function \(\psi\) as

\[
\psi(\alpha, \gamma, \theta) = \alpha^{a-1} \gamma^{c-1} \theta^{e-1} e^{-(ba + dy + e\theta)},
\]

\[
\alpha > 0, \quad \gamma > 0, \quad \theta > 0.
\]

The joint posterior distribution in terms of a given likelihood function \(L(\text{data})\) and joint prior distribution \(\pi(\alpha, \gamma, \theta)\) is defined as

\[
\pi^\ast(\alpha, \gamma, \theta | \text{data}) \propto \pi(\alpha, \gamma, \theta)L(\text{data}).
\]

(33)

Hence, we obtain the joint posterior density of the parameters \(\alpha, \gamma\) and \(\theta\) for the complete sample data by combining the likelihood function and the joint prior density (33). Therefore, the joint posterior density function is given by

\[
\pi^\ast(\alpha, \gamma, \theta | x) = K \psi(\alpha, \gamma, \theta) \prod_{i=1}^{n} \frac{\alpha \theta \gamma x_i^{a-1} \exp \left( \frac{\gamma x_i^{a}}{\alpha \theta + \gamma x_i^{a}} \right)}{(\theta + \gamma x_i^{a})^2},
\]

(34)

**Table 11.** Bayesian estimates and their posterior risks of the parameters under different loss functions based on the vehicle insurance losses data.

| Bayes Loss functions | \(\hat{\alpha}\) | \(\hat{\gamma}\) | \(\hat{\theta}\) | Estimate | Risk | Estimate | Risk | Estimate | Risk |
|----------------------|-----------------|-----------------|-----------------|----------|------|----------|------|----------|------|
| SELF                 | 1.3047          | 0.0307          | 0.0080          | 3.2091   | 1.5498 |
| WSELF                | 1.2826          | 0.0220          | 0.0026          | 0.0054   | 0.4389 |
| MSELF                | 1.2618          | 0.0162          | 0.0020          | 0.9001   | 2.2058 |
| PLF                  | 1.3164          | 0.0235          | 0.0104          | 0.0048   | 3.4422 |
| KLF                  | 1.2936          | 0.0171          | 0.0045          | 2.9456   | 0.1789 |

**Table 12.** Credible and HPD intervals of the parameters \(\alpha\), \(\gamma\) and \(\theta\) for the vehicle insurance losses data.

| Parameters | Credible interval | HPD interval |
|------------|-------------------|--------------|
| \(\alpha\) | (1.190, 1.392)    | (0.971, 1.669) |
| \(\gamma\) | (0.003, 0.011)    | (0.007, 0.021) |
| \(\theta\) | (2.264, 3.978)    | (0.958, 5.579) |

**Figure 19.** Plots of Bayesian analysis and performance of Gibbs sampling for Insurance data set. Trace plots of each parameter of EP-W distribution for the vehicle insurance losses data.
Figure 20. Plots of Bayesian analysis and performance of Gibbs sampling for Insurance data set. Autocorrelation plots of each parameter of EP-W distribution.

Figure 21. Plots of Bayesian analysis and performance of Gibbs sampling for Insurance data set. Histogram plots of each parameter of EP-W distribution.
where $K$ is given by

$$K^{-1} = \int_0^\infty \int_0^\infty \int_0^\infty \varphi(\alpha, \gamma, \eta) \times \prod_{i=1}^n \alpha^\gamma \gamma x_i^{\alpha-1} \exp \left( \frac{\gamma x_i^\alpha}{\theta + \gamma x_i^\alpha} \right) \, d\alpha \, d\gamma \, d\theta.$$ 

It is clear from Equation (34) that there is no closed form for the Bayesian estimators under the Five loss functions described in Table 8, so we suggest using a MCMC procedure based on 10,000 replicates to compute Bayesian estimators. The corresponding Bayesian estimates and posterior risk are provided in Table 9. Table 10 provides 95% credible and HPD intervals for each parameter of the EP-W distribution. The posterior samples extracted by using Gibbs sampling technique. Moreover, we provide the posterior summary plots in Figures 19–21. These plots confirm that the sampling process is of the prime quality and the convergence does occur.

For the vehicle insurance losses, the Bayesian estimates along with corresponding the posterior risk are provided in Table 11. Table 12 provides 95% credible and HPD intervals for each parameter of the EP-W distribution. The posterior samples extracted by using Gibbs sampling technique. Corresponding to the vehicle insurance data set, the posterior summary plots are provided in Figures 19–21. These plots confirm that the sampling process is of the prime quality and the convergence does occur.

10. Concluding remarks

To cater data in financial and actuarial sciences, a number of methods to define heavy-tailed distributions have been proposed. In this regard, we further carried this branch of distribution theory and proposed nine new families of distributions. A three-parameter special model, called EP-W distribution, is studied in detail. Some mathematical properties along with certain characterizations are derived. Actuarial measures of the proposed model are also calculated and a simulation study has been conducted to show the usefulness of the proposed method in actuarial sciences. To prove the potential of the EP-W distribution, we considered two insurance data sets and compared its goodness of fit with other well-known distributions. The proposed distribution outclassed the competitive models by considering certain analytical measures. We hope that the new development will attract wider applications in the financial and insurance sciences and would be quite helpful for the new comers in the field of distribution theory.

Acknowledgments

The authors are grateful to the Editor-in-Chief, the Associate Editor and the anonymous referees for many of their valuable comments and suggestions which lead to this improved version of the paper. The first two authors also acknowledge the support of the Yazd University, Iran.

Disclosure statement

This article is drafted from the Ph.D work of Mr. Zubair Ahmad.

Data availability statement

This work is mainly a methodological development and has been applied on secondary data related to financial and insurance sciences, but if required, data will be provided.

ORCID

Eisa Mahmoudi http://orcid.org/0000-0001-6109-7342
Omid Kharazmi http://orcid.org/0000-0001-6638-2185

References

[1] Bhati D, Ravi S. On generalized log-Moyal distribution: a new heavy tailed size distribution. Insur Math Econ. 2018;79:247–259.
[2] Beirlant J, Matthys G, Dierckx G. Heavy-tailed distributions and rating. ASTIN Bull. 2001;31(1):37–58.
[3] McNeil AJ. Estimating the tails of loss severity distributions using extreme value theory. ASTIN Bull. 1999;27(1):117–137.
[4] Resnick SI. Discussion of the Danish data on large fire insurance losses. ASTIN Bull. 1999;27(1):139–151.
[5] Eling M. Fitting insurance claims to skewed distributions: are the skew-normal and skew-student good models. Insur Math Econ. 2012;51(2):239–248.
[6] Adcock C, Eling M, Loperfi do N. Loperfi distributions in finance and actuarial science: a review. Eur J Financ. 2015;21(13-14):1253–1281.
[7] Shushi T. Skew-elliptical distributions with applications in risk theory. Eur Actuar J. 2017;7:277–296.
[8] Punzo A. A new look at the inverse Gaussian distribution with applications to insurance and economic data. J Appl Stat. 2019;46(7):1260–1287.
[9] Azzalini A, Del Cappello T, Kotz S. Log-skew-normal and log-skew-t distributions as models for family income data. J Inco Dist. 2002;11:12–20.
[10] Bagnato L, Punzo A. Finite mixtures of unimodal beta and gamma densities and the $k$-bumps algorithm. Comput Stat. 2013;28(4):1571–1597.
[11] Paula GA, Leiva V, Barros M, et al. Robust statistical modeling using the Birnbaum Saunders $t$ distribution applied to insurance. Appl Stoch Models Bus Ind. 2012;28(1):16–34.
[12] Klugman SA, Panjer HH, Willmot GE. Loss models: from data to decisions. 4th ed. John Wiley and Sons, Inc; 2012. (Wiley series in probability and statistics).
[13] Nadarajah S, Abu Bakar SA. New composite models for the Danish fire insurance data. Scand Actuar J. 2014;2014(2):180–187.
[14] Bakar SA, Hamzah N, Maghsoudi M, et al. Modeling loss data using composite models. Insur Math Econ. 2015;61:146–154.
[15] Punzo A, Bagnato L, Maruotti A. Compound unimodal distributions for insurance losses. Insur Math Econ. 2018;81:95–107.
as well be allowed). Let \( X : \Omega \rightarrow H \) be a continuous random variable with the distribution function \( F \) and let \( q_1 \) and \( q_2 \) be two real functions defined on \( H \) such that

\[
E \left[ q_2 (X) \mid X \geq x \right] = E \left[ q_1 (X) \mid X \geq x \right] \eta (x), \quad x \in H,
\]

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

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

where the function \( s \) is a solution of the differential equation

\[
s' = \frac{\eta'}{\eta} q_1 / \eta q_1 - q_2 \quad \text{and} \quad C \text{ is the normalization constant, such that } \int_a^x dF = 1.
\]

We like to mention that this kind of characterization based on the ratio of truncated moments is stable in the sense of weak convergence (see, Glänzel [27]), in particular, let us assume that there is a sequence \( \{X_n\} \) of random variables with distribution functions \( \{F_n\} \) such that the functions \( q_1 n , q_2 n \) and \( \eta n \) \((n \in N)\) satisfy the conditions of Theorem A.1 and let \( X_1 n \rightarrow q_1 , q_2 n \rightarrow q_2 \) for some continuously differentiable real functions \( q_1 \) and \( q_2 \). Let, finally, \( X \) be a random variable with distribution \( F \). Under the condition that \( q_1 n (X) \) and \( q_2 n (X) \) are uniformly integrable and the family \( \{F_n\} \) is relatively compact, the sequence \( X_n \) converges to \( X \) in distribution if and only if \( \eta n \) converges to \( \eta \), where

\[
\eta (x) = \frac{E \left[ q_2 (X) \mid X \geq x \right]}{E \left[ q_1 (X) \mid X \geq x \right]}.
\]

This stability theorem makes sure that the convergence of distribution functions is reflected by corresponding convergence of the functions \( q_1 , q_2 \) and \( \eta \), respectively. It guarantees, for instance, the “convergence” of characterization of the Wald distribution to that of the Lévy–Smirnov distribution if \( \alpha \rightarrow \infty \). A further consequence of the stability property of Theorem A.1 is the application of this theorem to special tasks in statistical practice such as the estimation of the parameters of discrete distributions. For such purpose, the functions \( q_1 , q_2 \) and, specially, \( \eta \) should be as simple as possible. Since the function triplet is not uniquely determined it is often possible to choose \( \eta \) as a linear function. Therefore, it is worth analysing some special cases which helps to find new characterizations reflecting the relationship between individual continuous univariate distributions and appropriate in other areas of statistics. In some cases, one can take \( q_1 (x) \equiv 1 \), which reduces the condition of Theorem A.1 to \( E[q_2 (X) \mid X \geq x] = \eta (x) \), \( x \in H \). We, however, believe that employing three functions \( q_1 , q_2 \) and \( \eta \) will enhance the domain of applicability of Theorem A.1.

### Appendix

**Theorem A.1:** Let \((\Omega , F, P)\) be a given probability space and let \(H = [d , e]\) be an interval for some \(d < e\) \((d = -\infty , e = \infty \) might

### R Code for analysis

Note: In the following R-code, \(a\) is used for \(\alpha\), \(\varepsilon\) is used for \(\theta\), \(g\) is used for \(\gamma\) and \(\Omega\) is used for proposed model.
```r
library(AdequacyModel)
data <- read.csv(file.choose(), header=TRUE)
data <- data[,1]
data <- data[!is.na(data)]
data <- data/1000
data

df_pm <- function(par, x)
{
a  <- par[1]
t  <- par[2]
g  <- par[3]
a*t*g*(x^(a-1))*(exp((g*x^a)/(t+g*x^a)))*(1/(t+g*x^a)^2)*(1/(exp(1)-1))
}
cdf_pm <- function(par, x)
{
a  <- par[1]
t  <- par[2]
g  <- par[3]
(exp((g*x^a)/(t+g*x^a))-1)/(exp(1)-1)
}
set.seed(0)
goodness.fit(pdf=pdf_pm, cdf=cdf_pm,
starts = c(0.1,0.1,0.1), data = data,
method = "B", domain = c(0,Inf), mle = NULL)

###############################################################
# R Code for Simulation
###############################################################
###############################################################
# Generating a Random Sample of Size n from the EP-Weibull
###############################################################

rEP_Weibull <- function(par, n)
{
a  <- par[1]; t  <- par[2]; g  <- par[3]
u <- runif(n)
x <- c()

for (i in 1:n)
{
x[i] = ((g/t)*(((log(u[i]*(exp(1)-1)+1))^(-1)-1)))^(-1/a)
}
return(x)
}

###############################################################
The pdf of the EP-Weibull
###############################################################

dEP_Weibull <- function(par, x)
{
a  <- par[1]
t  <- par[2]
g  <- par[3]
a*t*g*(x^(a-1))*(exp((g*x^a)/(t+g*x^a)))*(1/(t+g*x^a)^2)*(1/(exp(1)-1))
}

###############################################################
The cdf of the EP-Weibull
###############################################################
pEP_Weibull <- function(par, x)
{
a  <- par[1]
t  <- par[2]
g  <- par[3]
(exp((g*x^a)/(t+g*x^a))-1)/(exp(1)-1)
```

The loglikelihood function of the EP-Weibull

\[
\text{loglikelihoodEP_Weibull} \left( \text{par} \right) = \begin{cases} 
\log(a) + \log(t) + \log(g) + (a-1) \log(x) + \left( \frac{g \cdot x^a}{t+g \cdot x^a} \right) + \log\left( \frac{1}{\exp(1)-1} \right) 
& \text{if } a > 0 \land t > 0 \land g > 0 \land \min(x) > 0 \land \min\left( \frac{1}{t+g \cdot x^a} \right) > 0 \\
-9999999.9 
& \text{else}
\end{cases}
\]

par <- c(1.4, 0.5, 1); a = 1.4; t = 0.5; g = 1
n_replicas = 750
matriz_par <- matrix(0, 30, 3)
matriz_bias <- matrix(0, 30, 3)
matriz_MSE <- matrix(0, 30, 3)
matriz_std <- matrix(0, 30, 3)
colnames(matriz_par) <- c("a", "t", "g")
colnames(matriz_bias) <- c("a", "t", "g")
colnames(matriz_MSE) <- c("a", "t", "g")
colnames(matriz_std) <- c("a", "t", "g")

cont = 1
n = 25
while(n <= 750){
  par_mean <- c(0, 0, 0)
  std_mean <- c(0, 0, 0)
  bias <- c(0, 0, 0)
  MSE <- c(0, 0, 0)
  replica = 1
  while(replica <= n_replicas)
    {
      x <- rEP_Weibull(par, n)
      Data <- x
      result = optim(c(a, t, g), loglikelihoodEP_Weibull, hessian = F, control = list(fnscale = -1), method = "L-BFGS-B", lower = c(0.001, 0.001, 0.001), upper = c(5, 5, 5))
      if (class(result) != "try-error" && result$convergence == 0)
        {
          par_mean <- par_mean + result$par
          bias = bias + (result$par - par)
          MSE = MSE + (result$par - par)^2
          replica = replica + 1
        }
    }
  par_mean = par_mean/n_replicas
  bias = bias/n_replicas
  MSE = MSE/n_replicas
}
matriz_par[cont,] = par_mean
matriz_std[cont,] = std_mean
matriz_bias[cont,] = bias
matriz_MSE[cont,] = MSE

print("mean = ")
print( par_mean )
print("bias = ")
print( bias )
print("MSE = ")
print(MSE)
n = n + 25
cont = cont + 1
}

print(matriz_par)
print(matriz_MSE)
print(matriz_bias)

n=seq(25, 750, 25)

###############################################################
########## Plots of Parameters
###############################################################

plot(n,matriz_par[,1], type="o", col="green", lty=1, lwd=2,xlab="n", ylab="Estimated Parameters",ylim=c(0.4,2))
lines(n,matriz_par[,2], col="blue", lty=5, lwd=2,type="o")
lines(n,matriz_par[,3], col="red", lty=8,lwd=2,type="o")
title("Plot of Estimated Parameters vs n")
legend(525,1.9, legend=c(expression(paste(alpha,"=","1.4")),
expression(paste(theta,"=","0.5")),
expression(paste(gamma,"=","1"))),
   lty =c(1,5,8),cex=1,
   col=c('green','blue','red'))

###############################################################
########## Plots of MSEs
###############################################################

plot(n,matriz_MSE[,1], col="green", lty=1, lwd=2,type="o", xlab="n", ylab="MSE", ylim=c(0,0.1))
lines(n,matriz_MSE[,2], col="blue", lty=5, lwd=2,type="o")
lines(n,matriz_MSE[,3], col="red", lty=8,lwd=2,type="o")
title("Plot of MSE vs n")
legend(550,0.09, legend = c(expression(paste(alpha)),
expression(paste(theta)) ),
   expression(paste(gamma))),
   lty =c(1,5,8),cex=1,
   col=c('green','blue','red'))

###############################################################
########## Plots of Absolute Biases
###############################################################

plot(n,abs(matriz_bias[,1]), type="o", col="green", lty=1, lwd=2,xlab="n", ylab="Absolute Bias",ylim=c(0,0.15))
lines(n,abs(matriz_bias[,2]), col="blue", lty=2, lwd=2,type="o")
lines(n,abs(matriz_bias[,3]), col="red", lty=3,lwd=2,type="o")
title("Plot of Absolute Bias vs n")
legend(550,0.14, legend = c(expression(paste(alpha)),
expression(paste(theta))),
   expression(paste(gamma))),
   lty =c(1,5,8),cex=1,
col=c(‘green’, ‘blue’, ‘red’))

# Plots of Biases

plot(n, matriz_bias[,1], type="o", col="green", lty=1, lwd=2, xlab="n", ylab="Bias", ylim=c(0,0.15))
lines(n, matriz_bias[,2], col="blue", lty=5, lwd=2, type="o")
lines(n, matriz_bias[,3], col="red", lty=8, lwd=2, type="o")
title("Plot of Bias vs n")
legend(550, 0.14, legend = c(expression(paste(alpha)),
expression(paste(theta)),
expression(paste(gamma))),
lty =c(1,5,8), cex=1,
col=c(‘green’,‘blue’,‘red’))

#