

Research Article

On Modeling the Earthquake Insurance Data via a New Member of the T-X Family

Zubair Ahmad,¹ Eisa Mahmoudi ,¹ and Omid Kharazmi²

¹Department of Statistics, Yazd University, P.O. Box 89175-741, Yazd, Iran

²Department of Statistics, Faculty of Sciences, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran

Correspondence should be addressed to Eisa Mahmoudi; emahmoudi@yazd.ac.ir

Received 27 December 2019; Revised 25 August 2020; Accepted 9 September 2020; Published 19 September 2020

Academic Editor: Friedhelm Schwenker

Copyright © 2020 Zubair Ahmad et al. 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.

Heavy-tailed distributions play an important role in modeling data in actuarial and financial sciences. In this article, a new method is suggested to define new distributions suitable for modeling data with a heavy right tail. The proposed method may be named as the Z-family of distributions. For illustrative purposes, a special submodel of the proposed family, called the Z-Weibull distribution, is considered in detail to model data with a heavy right tail. The method of maximum likelihood estimation is adopted to estimate the model parameters. A brief Monte Carlo simulation study for evaluating the maximum likelihood estimators is done. Furthermore, some actuarial measures such as value at risk and tail value at risk are calculated. A simulation study based on these actuarial measures is also done. An application of the Z-Weibull model to the earthquake insurance data is presented. Based on the analyses, we observed that the proposed distribution can be used quite effectively in modeling heavy-tailed data in insurance sciences and other related fields. Finally, Bayesian analysis and performance of Gibbs sampling for the earthquake data have also been carried out.

1. Introduction

In a number of applied areas such as finance and actuarial sciences, data sets are most often positive, and the respective distribution is unimodal hump-shaped and skewed to right having heavier tails as compared to the well-known classical distributions. These distributions are not much flexible to adequately model such types of heavy-tailed data sets. For example, (i) the Pareto distribution, which is frequently used to model financial data sets, does not provide a reasonable fit for many applications, for example, if we are interested in modeling only especially moderate-to-large losses altogether, then in such cases, the Pareto distribution may not be a suitable choice to use [1], and (ii) the Weibull model is capable of catering the behavior of small losses very closely, but, unfortunately, fails to provide an adequate fit to the large losses [2]. In such circumstances, the utilization of the heavy-tailed models may be a good choice to apply. For positive data, heavy-tailed distributions are those whose right-tail probabilities are greater than the exponential one [3], that is,

$$\lim_{x \rightarrow \infty} e^{px} (1 - F(x; \xi)) = \infty, \quad p > 0, \quad (1)$$

where $F(x; \xi)$ is the cumulative distribution function (cdf) depending on the parameter vector $\xi \in \mathbb{R}$.

Due to the usefulness and flexibility of the heavy-tailed models in financial and actuarial practice, actuaries are always intended to propose new statistical distributions. Therefore, serious attempts have been made to propose new statistical models and are still growing rapidly. The new contribution is made via 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, see Ahmad et al. [4].

Recent investigation of Eling [5] and Adcock et al. [6] determined that skew-normal and skew Student's t distributions are the most excellent competitors because the skewed distributions adjust right-skewness and high kurtosis; for the interested readers, one can refer to Shushi [7] and Punzo et al. [8]. However, insurance losses and monetary risks take values on the positive real line, and subsequently, these skew models may not be a suitable choice to

use as they are defined on \mathbb{R} . In such circumstances, the transformation of variable approach, especially, the exponential transformation, has demonstrated to be considerable; for details, see Azzalini et al. [9]. Bagnato and Punzo [10] showed that the transformation method of introducing new distributions is easy to use; however, most often, the inferences become complicated.

Another useful method of proposing new versatile heavy-tailed distributions, which provide the best fit to the heavy-tailed losses, is the methodology of composition, see Paula et al. [11], Klugman et al. [12], Nadarajah and Bakar [13] and Bakar et al. [14]. However, it ought to be noted that the new statistical models introduced via this approach involve more than three parameters inflicting difficulties in estimating the parameters, and more computational efforts are needed.

Another approach of introducing new distributions to cater data modeling adequately with unimodality is compounding of distributions, see Punzo et al. [15] and Mazza and Punzo [16]. Unfortunately, the density function of the distributions obtained by this approach might not have a closed-form expression that makes the estimation more complicated as shown in Punzo et al. [15].

The method of finite mixture models is another prominent approach to obtain new, very flexible models which are able to capture, for example, multimodality of the distribution under consideration, see Bernardi et al. [17], Miljkovic and Grün [18], and Punzo et al. [15]. No doubt, the distributions obtained via this approach are much flexible, but the inferences become more complicated and computationally challenging.

Furthermore, Dutta and Perry [19] 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 modeling insurance losses due to the poor results. They concluded that one would need to use a model that is flexible enough in its structure. This motivated the researchers to search for more flexible models offering greater accuracy in fitting the heavy-tailed data.

Hence, bringing flexibility to a model by introducing additional parameter(s) is a desirable feature [20–24]. In a number of recent papers, serious attempts have been made to introduce a variety of new heavy-tailed distributions, see Ahmad et al. [25] and Ahmad et al. [26]. Due to the importance of statistical distributions in financial science, a new family of distributions, called the Z-family, is introduced. The proposed family is introduced via the T-X family approach [27]. To illustrate the usefulness of the proposed method, a three-parameter submodel, called the Z-Weibull distribution, is taken and studied in detail. The proposed distribution provides a better description of the earthquake insurance data with possibly heavy tails than the available (i) two-parameter distributions such as Weibull, Burr-XII (B-XII), and generalized exponential (GE), (ii) three-parameter Weibull-claim (W-claim) and exponentiated Lomax (EL), and (iii) four-parameter beta-Weibull (BW) distributions, and possibly many others.

The rest of the article is structured in the following way. The proposed method is introduced in Section 2. The Z-Weibull model is considered in Section 3, and the shapes of its probability density function (pdf) are investigated in the same section. Estimation of parameters is discussed in Section 4. In the same section, a detailed Monte Carlo simulation study is conducted. Actuarial measures of the proposed method along with a simulation study are provided in Section 5. Distribution fit to the earthquake insurance data set is discussed in Section 6. Bayesian analysis as well as Gibbs sampling procedure for the real data set is discussed in Section 7. Future frame work is discussed in Section 8. Finally, some concluding remarks are presented in Section 9.

2. Development of the Z-Family

Let $p(t)$ be the pdf of a random variable, say T , where $T \in [n_1, n_2]$ for $-\infty \leq n_1, n_2 \leq \infty$, and let $W[F(x; \xi)]$ be a function of $F(x; \xi)$ of a random variable, say X , depending on the vector parameter ξ satisfying the conditions given in the following:

- (1) $W[F(x; \xi)] \in [n_1, n_2]$
- (2) $W[F(x; \xi)]$ is differentiable and monotonically increasing
- (3) $W[F(x; \xi)] \rightarrow n_1$ as $x \rightarrow -\infty$ and $W[F(x; \xi)] \rightarrow n_2$ as $x \rightarrow \infty$

Alzaatreh et al. [27] introduced a general method for generating new families of distributions called the T-X family which is defined by

$$G(x) = \int_{n_1}^{W[F(x; \xi)]} p(t) dt, \quad t \geq 0, \quad (2)$$

where $W[F(x; \xi)]$ satisfies the conditions mentioned above. The probability density function (pdf) corresponding to (2) is given by

$$g(x) = \left\{ \frac{d}{dx} W(F(x; \xi)) \right\} p(W[F(x; \xi)]). \quad (3)$$

Deploying the T-X proposal, several new classes of distributions have been introduced in the literature [28]. Let X have the exponential distribution with the pdf given by

$$p(t) = \theta e^{-\theta t}, \quad t, \theta > 0. \quad (4)$$

Using $\theta = 1$ in (4), we get

$$p(t) = e^{-t}, \quad t > 0. \quad (5)$$

On setting $W[F(x; \xi)] = -\log\{(1 - F(x; \xi))/(\beta^{F(x; \xi)})\}$ and $p(t) = e^{-t}$ in (2), we define the cdf of the Z-family of distributions by

$$G(x; \beta, \xi) = 1 - \left\{ \frac{1 - F(x; \xi)}{\beta^{F(x; \xi)}} \right\}, \quad \beta > 0, x, \xi \in \mathbb{R}, \quad (6)$$

where $F(x; \xi)$ is the baseline cdf. The expression in (6) represents a wide family of univariate continuous

distributions. Clearly, when $\beta = 1$, the cdf of the proposed family derived in (6) becomes identical to the baseline cdf. The pdf corresponding to (6) is given by

$$g(x; \beta, \xi) = f(x; \xi) \left(\frac{1 + (\log \beta)[1 - F(x; \xi)]}{\beta^{F(x; \xi)}} \right), \quad x \in \mathbb{R}. \quad (7)$$

The survival function (sf) and hazard rate function (hrf) corresponding to (6) are, respectively, given by

$$S(x; \beta, \xi) = \frac{1 - F(x; \xi)}{\beta^{F(x; \xi)}}, \quad x \in \mathbb{R},$$

$$h(x; \beta, \xi) = f(x; \xi) \left\{ \frac{1 + (\log \beta)[1 - F(x; \xi)]}{1 - F(x; \xi)} \right\}, \quad x \in \mathbb{R}. \quad (8)$$

Due to induction of the extra parameter, the Z-family provides greater distributional flexibility. The key motivations for using the Z-family in the practice are as follows:

- (i) A very useful and simple method of introducing an additional parameter to generalize the existing distributions
- (ii) To improve the characteristics and flexibility of the existing models
- (iii) To introduce new distributions having closed form of cdf and sf, as well as hrf
- (iv) To extend the existing distributions by introducing only one parameter, rather than adding two or more parameters
- (v) To provide the best fit to heavy-tailed insurance data sets
- (vi) To provide better fits than other modified models having same or higher number of parameters

3. The Z-Weibull Distribution

Most of the extended forms of distributions are introduced for one of the following aims: (i) an extension of the existing model to improve its characteristics, (ii) to obtain new distribution having a heavy right tail, and (iii) to introduce a model whose empirical fit is good to data. Here, we discuss the Z-Weibull distribution that can possess at least one of these aims. The Weibull random variable has the cdf and pdf given by $F(x; \xi) = 1 - e^{-\gamma x^\alpha}$ and $f(x; \xi) = \theta \gamma x^{\alpha-1} e^{-\gamma x^\alpha}$, respectively, where $\xi = (\alpha, \gamma)$. Then, the cdf of the Z-Weibull distribution has the following form:

$$G(x; \beta, \xi) = 1 - \frac{e^{-\gamma x^\alpha}}{\beta^{(1-e^{-\gamma x^\alpha})}}, \quad x \geq 0, \alpha, \gamma, \beta > 0. \quad (9)$$

The corresponding density is given by

$$g(x; \beta, \xi) = \frac{\alpha \gamma x^{\alpha-1} e^{-\gamma x^\alpha}}{\beta^{(1-e^{-\gamma x^\alpha})}} (1 + (\log \beta) e^{-\gamma x^\alpha}), \quad x > 0. \quad (10)$$

The sf, hrf, and reversed hazard rate function (rhfrf) of the proposed model are given by

$$S(x; \beta, \xi) = \frac{e^{-\gamma x^\alpha}}{\beta^{(1-e^{-\gamma x^\alpha})}}, \quad x > 0,$$

$$h(x; \beta, \xi) = \alpha \gamma x^{\alpha-1} (1 + (\log \beta) e^{-\gamma x^\alpha}), \quad x > 0,$$

$$r(x; \beta, \xi) = \frac{\alpha \gamma x^{\alpha-1} e^{-\gamma x^\alpha}}{\beta^{(1-e^{-\gamma x^\alpha})} - e^{-\gamma x^\alpha}} (1 + (\log \beta) e^{-\gamma x^\alpha}), \quad x > 0, \quad (11)$$

respectively.

Different plots for the pdf of the Z-Weibull distribution for selected parameter values are given in Figure 1.

4. Estimation and Monte Carlo Simulation Study

Several approaches to estimate the model parameter have been introduced in the literature, but the maximum likelihood estimation method is the most commonly employed. The maximum likelihood estimators (MLEs) enjoy several desirable properties and can be used for constructing confidence intervals and regions and also in test statistics. The normal approximation for MLEs in large samples can be easily handled either analytically or numerically. So, we estimate the parameters of the Z-family of distributions from complete samples via the maximum likelihood estimation method. Furthermore, we perform a comprehensive Monte Carlo simulation study to evaluate the performance of the MLEs.

4.1. Maximum Likelihood Estimation. In this section, we obtain the MLEs of the model parameters of the Z-family of distributions from complete samples only. Let x_1, x_2, \dots, x_n be the observed values from the Z-family of distributions with parameters β and ξ . The total log-likelihood function corresponding to (7) is given by

$$\ell(\beta, \xi) = \sum_{i=1}^n \log[f(x_i; \xi)] - \sum_{i=1}^n F(x_i; \xi) (\log \beta) + \sum_{i=1}^n \log[1 + (\log \beta)(1 - F(x_i; \xi))]. \quad (12)$$

The partial derivatives of (12) are

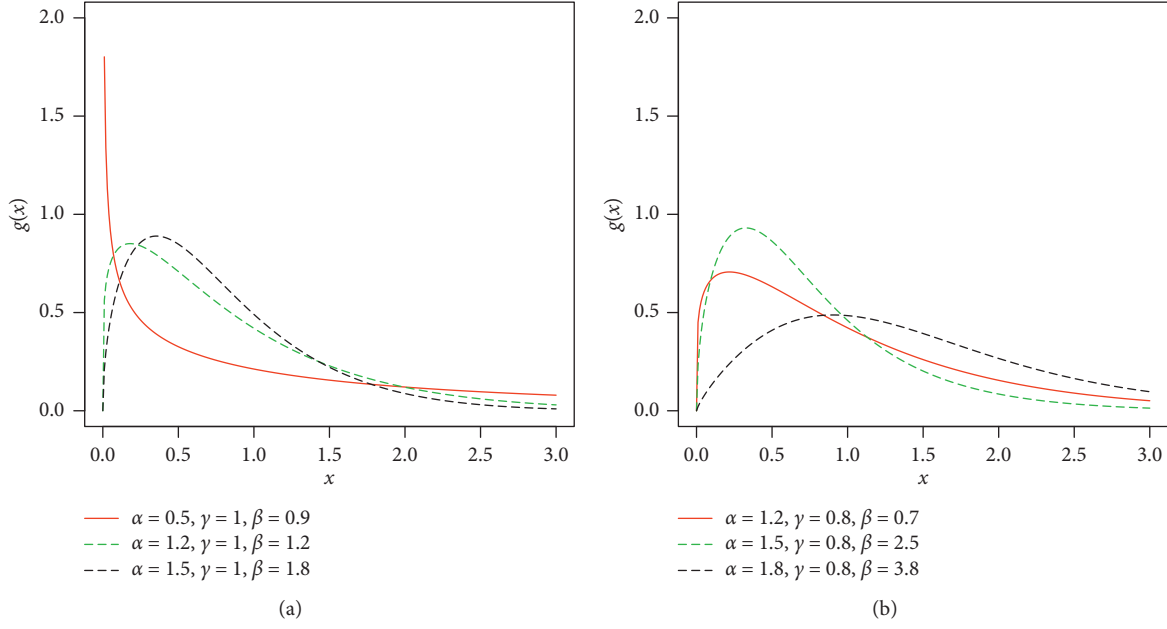


FIGURE 1: Plots for the density function of the Z-Weibull distribution.

$$\begin{aligned} \frac{\partial}{\partial \beta} \ell(\beta, \xi) &= - \sum_{i=1}^n \frac{F(x_i; \xi)}{\beta} + \frac{1}{\beta} \sum_{i=1}^n \frac{(1 - F(x_i; \xi))}{1 + (\log \beta)(1 - F(x_i; \xi))}, \\ \frac{\partial}{\partial \xi} \ell(\beta, \xi) &= \sum_{i=1}^n \frac{\partial f((x_i; \xi)/\partial \xi)}{f(x_i; \xi)} - \sum_{i=1}^n \frac{(\log \beta) \partial F((x_i; \xi)/\partial \xi)}{F(x_i; \xi)} - \sum_{i=1}^n \frac{\partial F((x_i; \xi)/\partial \xi)}{1 + (\log \beta)(1 - F(x_i; \xi))}. \end{aligned} \quad (13)$$

Setting $(\partial/\partial \beta)\ell(\beta, \xi)$ and $(\partial/\partial \xi)\ell(\beta, \xi)$ equal to zero and solving numerically these expressions simultaneously yield the MLEs of (β, ξ) .

4.2. Monte Carlo Simulation Study. This section offers a comprehensive simulation study to assess the behavior of the MLEs. The Z-family is easily simulated by inverting (6) as follows: if U has a uniform $U(0,1)$ distribution, then the nonlinear equation is

$$\log(1 - u) + \gamma x^\alpha + (1 - e^{-\gamma x^\alpha})(\log \beta) = 0. \quad (14)$$

Expression (14) can be used to simulate any special subcase of the Z-family. Here, we consider the Z-Weibull distribution to assess the behavior of the MLEs of the proposed method. We simulate the Z-Weibull distribution for two sets of parameters (set 1: $\alpha = 0.9, \beta = 0.3$, and $\gamma = 0.5$ and set 2: $\alpha = 1.3, \beta = 0.8$, and $\gamma = 1$). The simulation is performed via statistical software R through the library (rootSolve) command `mle`. The number of Monte Carlo replications made was 750 times. For maximizing

the log-likelihood function, we use the method = “L-BFGS-B” algorithm with `optim()`. The evaluation of the estimators was performed via the following quantities for each sample size: the empirical mean squared errors (MSEs) are calculated using the R package from the Monte Carlo replications. The MLEs are determined for each piece of simulated data, say, $(\hat{\alpha}_i, \hat{\gamma}_i, \hat{\beta}_i)$ for $i = 1, 2, \dots, 750$, and the biases and MSEs are computed, respectively, by

$$\text{bias}(w) = \left(\frac{1}{750} \right) \sum_{i=1}^{750} (\hat{w}_i - w), \quad (15)$$

$$\text{MSE}(w) = \left(\frac{1}{750} \right) \sum_{i=1}^{750} (\hat{w}_i - w)^2,$$

for $w = \alpha, \gamma, \beta$. We consider the sample sizes of $n = 25, 100, 200, 300, 400, 500, 600, 700, 750$. The empirical results are given in Tables 1 and 2. Corresponding to Tables 1 and 2, the simulation results are graphically displayed in

TABLE 1: Simulation results for the Z-Weibull distribution.

Set 1: $\alpha = 0.9$, $\beta = 0.3$, and $\gamma = 1.2$				
n	Parameters	MLEs	MSEs	Bias
25	α	1.1150	0.1937	0.2150
	β	1.2490	3.6872	0.9490
	γ	2.0431	4.7184	0.8431
50	α	1.0456	0.1526	0.1456
	β	0.9536	2.4614	0.6536
	γ	1.9282	4.3933	0.7281
100	α	0.9970	0.0783	0.0970
	β	0.7277	1.3741	0.4277
	γ	1.6722	2.3642	0.4722
200	α	0.9605	0.0472	0.0605
	β	0.5478	0.6878	0.2478
	γ	1.4691	1.1474	0.2691
400	α	0.9281	0.0163	0.0281
	β	0.3769	0.0901	0.0769
	γ	1.3115	0.4336	0.1115
600	α	0.9193	0.0085	0.0193
	β	0.3471	0.0383	0.0471
	γ	1.2424	0.1812	0.0424
800	α	0.9128	0.0067	0.0128
	β	0.3313	0.0279	0.0313
	γ	1.2405	0.1318	0.0405
1000	α	0.9006	0.0042	0.0076
	β	0.3184	0.0119	0.0184
	γ	1.2160	0.0928	0.0360

TABLE 2: Simulation results for the Z-Weibull distribution.

Set 2: $\alpha = 1.3$, $\beta = 0.8$, and $\gamma = 1.4$				
n	Parameters	MLEs	MSEs	Bias
25	α	1.3094	0.2002	0.2094
	β	1.9925	5.7399	1.2925
	γ	2.1937	4.3061	0.7937
50	α	1.2340	0.1333	0.1340
	β	1.6368	4.0681	0.9368
	γ	2.0939	3.3865	0.6939
100	α	1.1766	0.0697	0.0766
	β	1.2543	2.1560	0.5543
	γ	1.8715	2.1136	0.4715
200	α	1.1396	0.0368	0.0396
	β	0.9740	0.9313	0.2740
	γ	1.6771	1.0320	0.2771
400	α	1.1194	0.0153	0.0194
	β	0.8080	0.2293	0.1080
	γ	1.5204	0.3718	0.1204
600	α	1.1171	0.0094	0.0171
	β	0.7860	0.1229	0.0860
	γ	1.4516	0.1869	0.0516
800	α	1.1093	0.0065	0.0093
	β	0.7495	0.0833	0.0495
	γ	1.4443	0.1219	0.0443
1000	α	1.1050	0.0049	0.0050
	β	0.7807	0.0537	0.0307
	γ	1.4147	0.0935	0.0347

Figures 2 and 3. Based on Tables 1 and 2 and Figures 2 and 3, the following results are concluded:

- (i) Biases for all parameters are positive
- (ii) The parameters tend to be stable
- (iii) Estimated biases decrease when the sample size n increases
- (iv) Estimated MSEs decay toward zero when the sample size n increases

5. Actuarial Measures

One of the most important tasks of actuarial science 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 such as value at risk (VaR) and tail value at risk (TVaR) for the proposed distribution, which play a crucial role in portfolio optimization under uncertainty.

5.1. 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 represents the percentage loss in the portfolio value that will be equaled or exceeded only X percent of the time. VaR of a random variable X is the q^{th} quantile of its cdf, see Artzner [29]. If X follows the proposed method, then the VaR of X is

$$x_q = G^{-1}(u) = F^{-1}(t), \quad (16)$$

where t is the solution of $\beta^t(1-q) + t - 1 = 0$.

5.2. Tail Value at Risk. Another important measure is TVaR, also known as conditional tail expectation (CTE) or tail conditional expectation (TCE), which is used to quantify the expected value of the loss given that an event outside a given probability level has occurred. Let X follow the Z-Weibull distribution; then, TVaR of X is derived as

$$\text{TVaR}_q(x) = \left(\frac{1}{1-q} \right) \int_{\text{VaR}_q}^{\infty} xg(x)dx. \quad (17)$$

Using (10) in (17), we get

$$\begin{aligned} \text{TVaR}_q(x) &= \left(\frac{\alpha\gamma}{1-q} \right) \int_{\text{VaR}_q}^{\infty} \frac{x^{\alpha+1-1} e^{-\gamma x^\alpha}}{\beta(1-e^{-\gamma x^\alpha})} (1 + (\log \beta)e^{-\gamma x^\alpha}) dx, \\ \text{TVaR}_q(x) &= \left(\frac{\alpha\gamma}{1-q} \right) \sum_{i=0}^{\infty} \sum_{j=0}^i \frac{(\log \beta)^i (-1)^{(i+j)}}{(i-j)!j!} \int_{\text{VaR}_q}^{\infty} x^{\alpha+1-1} e^{-\gamma(j+1)x^\alpha} dx \\ &\quad + \left(\frac{\alpha\gamma}{1-q} \right) \sum_{i=0}^{\infty} \sum_{j=0}^i \frac{(\log \beta)^i (-1)^{i+j}}{(i-j)!j!} (\log \beta) \int_{\text{VaR}_q}^{\infty} x^{\alpha+1-1} e^{-\gamma(j+2)x^\alpha} dx. \end{aligned} \quad (18)$$

Recall the definition of incomplete gamma function in the form $\Gamma(\alpha, x) = \int_x^{\infty} t^{\alpha-1} e^{-t} dt$, so from (18), we get

$$\begin{aligned} \text{TVaR}_q(x) &= \left(\frac{1}{1-q} \right) \sum_{i=0}^{\infty} \sum_{j=0}^i \frac{(\log \beta)^i (-1)^{(i+j)}}{(i-j)!j! \gamma^{(1/\alpha)}} \frac{\Gamma((1/\alpha) + 1, \gamma(j+1)(\text{VaR}_q)^\alpha)}{(j+1)^{((1/\alpha)+1)}} \\ &\quad + \frac{1}{1-q} \sum_{i=0}^{\infty} \sum_{j=0}^i \frac{(\log \beta)^i (-1)^{(i+j)}}{(i-j)!j! \gamma^{(1/\alpha)}} \frac{\Gamma((1/\alpha) + 1, \gamma(j+2)(\text{VaR}_q)^\alpha)}{(j+2)^{((1/\alpha)+1)}}. \end{aligned} \quad (19)$$

5.3. Numerical Study of the Risk Measures. In this section, we provide numerical study of the VaR and TVaR for the Weibull and Z-Weibull distributions for different sets of parameters. The process is described as follows:

- (1) Random samples of size $n = 100$ are generated from the Weibull and Z-Weibull models
- (2) The parameters have been estimated via the maximum likelihood method

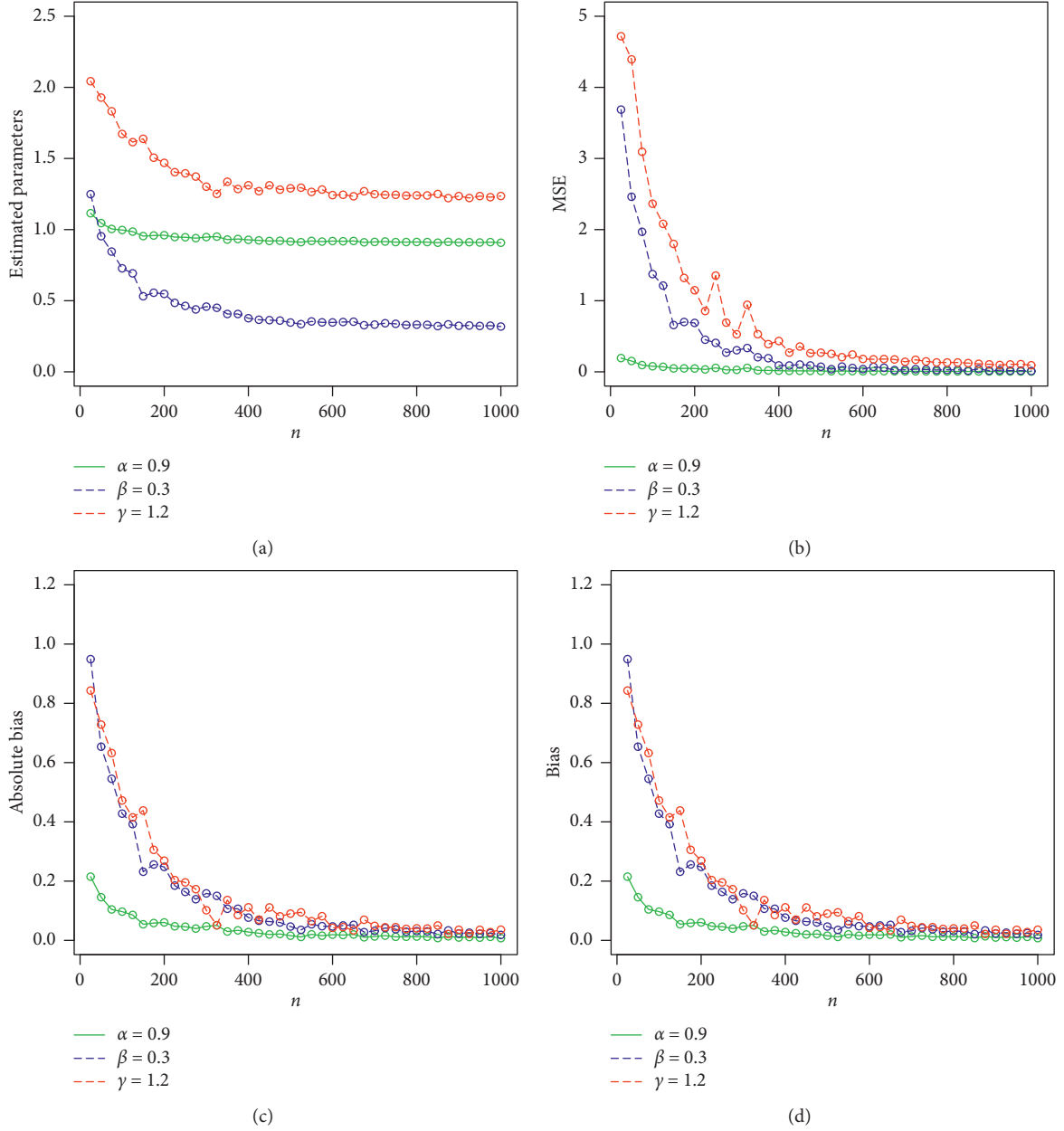


FIGURE 2: Graphical display of the results provided in Table 1. (a) Plot of estimated parameters vs. n . (b) Plot of MSE vs. n . (c) Plot of absolute bias vs. n . (d) Plot of bias vs. n .

- (3) 1000 repetitions are made to calculate the VaR and TVaR

The numerical results of the risk measures are provided in Tables 3 and 4 and displayed graphically in Figures 4 and 5 corresponding to each table.

The simulation is performed for Weibull and Z-Weibull 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 Z-Weibull model has higher values of the risk measures than the traditional Weibull distribution. The simulation results are graphically displayed in Figures 4 and 5, which show that the proposed model has a heavier tail than the Weibull distribution.

6. Practical Illustration via the Earthquake Insurance Data

The main applications of the heavy tail models are the so-called extreme value theory or insurance loss phenomena. In this section, we consider heavy-tailed earthquake insurance data to illustrate the usefulness of the Z-Weibull model. The data are reported by the “National Centers for Environmental Information” available at <https://ngdc.noaa.gov/hazard/earthqk.shtml>. We compare the goodness-of-fit results of the Z-Weibull distribution with the other well-known heavy-tailed distributions. The distribution functions of the competitive models are as follows:

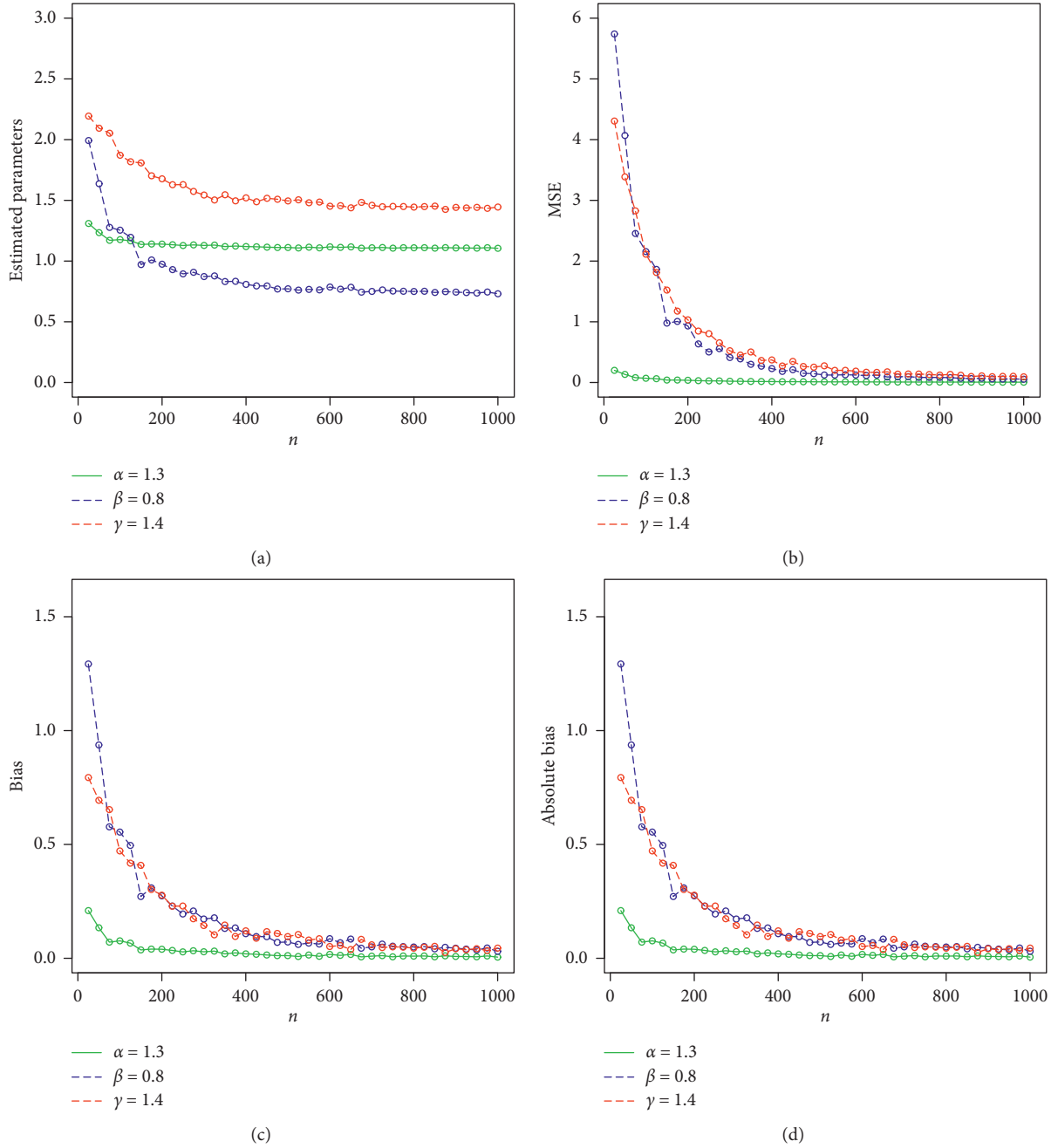


FIGURE 3: Graphical display of the results provided in Table 2. (a) Plot of estimated parameters vs. n . (b) Plot of MSE vs. n . (c) Plot of bias vs. n . (d) Plot of absolute bias vs. n .

(i) W-claim:

$$G(x) = \frac{\sigma(1 - e^{-\gamma x^\alpha})^2}{1 - (1 - \sigma)(1 - e^{-\gamma x^\alpha})}, \quad x \geq 0, \alpha, \gamma, \sigma > 0. \quad (20)$$

(ii) Weibull:

$$G(x) = 1 - e^{-\gamma x^\alpha}, \quad x \geq 0, \alpha, \gamma > 0. \quad (21)$$

(iii) B-XII:

$$G(x) = 1 - (1 + x^c)^{-k}, \quad x \geq 0, c, k > 0. \quad (22)$$

(iv) GE:

$$G(x) = (1 - e^{-\gamma x^\alpha})^a, \quad x \geq 0, a, \alpha, \gamma > 0. \quad (23)$$

(v) EL:

TABLE 3: Simulation results of the VaR and TVaR for the selected values of the parameters.

Dist.	Par	Level of significance	VaR	TVaR
Weibull	$\alpha = 1.2$ $\gamma = 0.85$	0.700	4.5373	7.3911
		0.750	5.0945	7.9076
		0.800	5.7601	8.5304
		0.850	6.5967	9.3205
		0.900	7.7437	10.4136
		0.950	9.6399	12.2384
		0.990	11.4765	14.0197
		0.999	19.5352	21.9150
Z-Weibull	$\alpha = 1.2$ $\beta = 1.4$ $\gamma = 0.85$	0.700	4.3778	13.2772
		0.750	5.5376	14.9465
		0.800	7.1015	17.1137
		0.850	9.3408	20.1026
		0.900	12.8999	24.6727
		0.950	20.0017	33.4043
		0.990	28.2956	43.2302
		0.999	80.4988	101.7184

TABLE 4: Simulation results of the VaR and TVaR for the selected values of the parameters.

Dist.	Par	Level of significance	VaR	TVaR
Weibull	$\alpha = 0.65$ $\gamma = 0.7$	0.700	1.5273	4.4325
		0.750	1.9174	4.9762
		0.800	2.4392	5.6792
		0.850	3.1801	6.6436
		0.900	4.3463	8.1092
		0.950	6.6444	10.8853
		0.990	9.2950	13.9806
		0.999	25.5667	32.0408
Z-Weibull	$\alpha = 0.65$ $\beta = 0.8$ $\gamma = 0.7$	0.700	1.7585	6.9318
		0.750	2.3222	7.9131
		0.800	3.1171	9.2175
		0.850	4.3112	11.0668
		0.900	6.3169	13.9926
		0.950	10.6148	19.8560
		0.990	16.0027	26.8019
		0.999	55.1566	73.1187

$$G(x) = (1 - (1 + \gamma x)^{-\alpha})^a, \quad x \geq 0, a, \alpha, \gamma > 0. \quad (24)$$

$$AIC = 2k - 2\ell. \quad (26)$$

(vi) BW:

$$G(x) = I_{(1-e^{-\gamma x^a})}(a, b), \quad x \geq 0, a, b, \alpha, \gamma > 0. \quad (25)$$

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), and consistent Akaike information criterion (CAIC), and (ii) two other goodness-of-fit measures including Anderson–Darling (AD) test statistic and -2ℓ . The discrimination measures are given as follows.

(vii) Akaike information criterion is given by

(viii) Consistent Akaike information criterion is given by

$$CAIC = \frac{2nk}{n-k-1} - 2\ell. \quad (27)$$

(ix) Bayesian information criterion is given by

$$BIC = k \log(n) - 2\ell. \quad (28)$$

(x) Hannan–Quinn information criterion is given by

$$HQIC = 2k \log(\log(n)) - 2\ell, \quad (29)$$

where ℓ denotes the log-likelihood function, k is the number of model parameters, and n is the sample size.

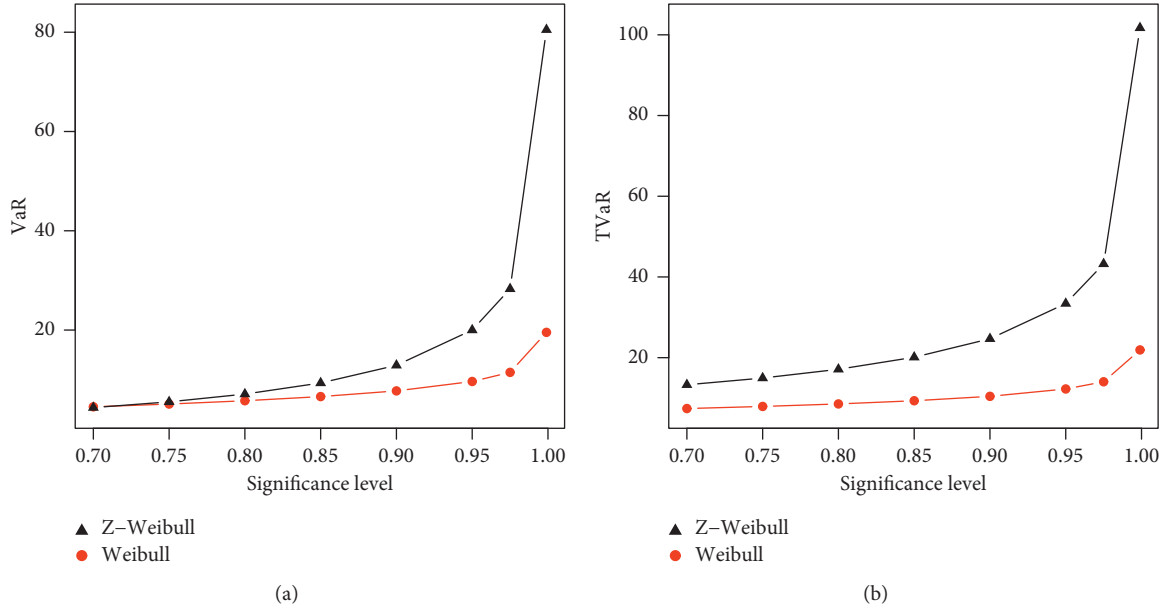


FIGURE 4: Graphical sketching of the results of the VaR and TVaR provided in Table 3.

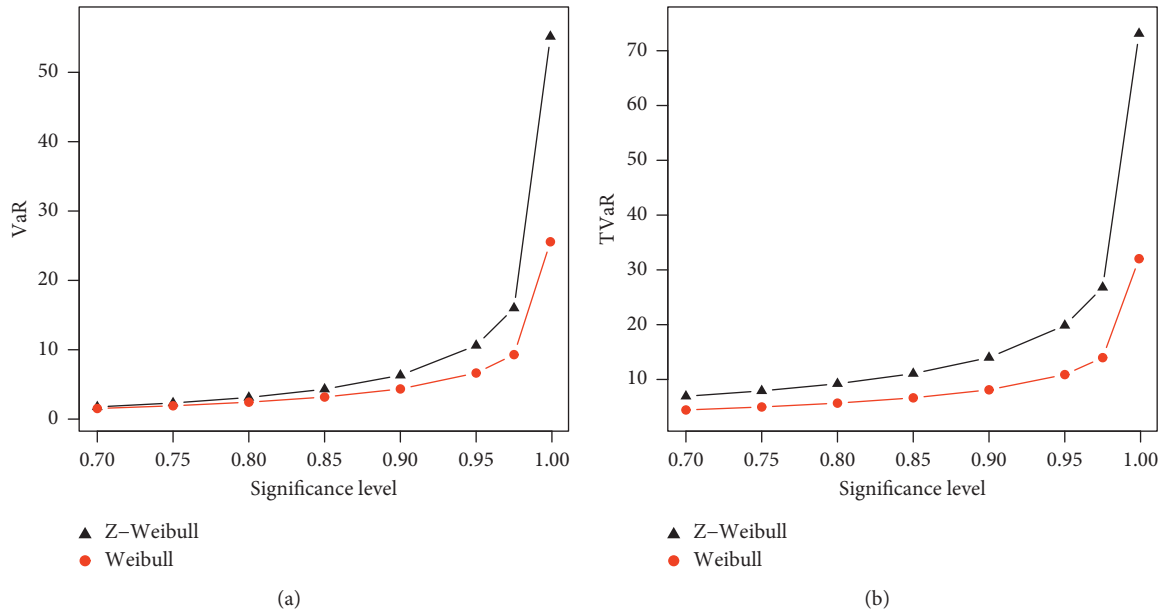


FIGURE 5: Graphical sketching of the results of the VaR and TVaR provided in Table 4.

(xi) The AD test statistic is given by

$$AD = -n - \left(\frac{1}{n}\right) \sum_{i=1}^n (2i-1) [\log G(x_i) + \log\{1 - G(x_{n-i+1})\}], \quad (30)$$

where n is the sample size and x_i is the i^{th} observation in the sample, calculated when the data are sorted in the ascending order.

All the computations are carried out using the `optim()` R-function with the argument `method="BFGS"` (see

Appendix). A model with lowest values for these measures could be chosen as the best model to fit the data. The values of MLEs of the parameters along with standard errors in parenthesis are presented in Table 5, whereas the discrimination measures are displayed in Table 6. The AD statistic and -2ℓ are provided in Table 7. Based on the considered data set, we have observed that the Z-Weibull distribution is the best fitted model among the above considered models.

As we can see, the results show that the Z-Weibull distribution provides a better fit than the other competitors. Hence, the proposed model can be used as a best

TABLE 5: Estimated values of the proposed and other competitive models.

Dist.	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\beta}$	\hat{a}	\hat{b}	\hat{c}	\hat{k}	$\hat{\sigma}$
Z-Weibull	2.346 (0.020)	0.564 (0.029)	2.295 (0.248)					
W-claim	2.023 (0.076)	0.709 (0.056)						1.783 (0.197)
Weibull	2.110 (0.091)	0.889 (0.086)						
B-XII						1.480 (0.030)	2.851 (0.107)	
GE		2.283 (0.015)		4.308 (0.053)				
EL	4.804 (4.910)	3.879 (5.400)		4.335 (0.054)				
BW	1.609 (0.055)	0.644 (0.304)		1.616 (0.087)	2.237 (1.228)			

TABLE 6: Discrimination measures of the proposed and other competitive models.

Dist.	AIC	BIC	CAIC	HQIC
Z-Weibull	23286.95	23310.57	23289.92	23294.69
W-claim	23345.86	23364.09	23355.04	23354.73
Weibull	23379.64	23395.38	23386.84	23384.80
B-XII	24197.63	24213.37	24197.63	24202.78
GE	23846.71	23862.45	23853.57	23851.86
EL	23863.99	23887.60	23871.99	23871.72
BW	23338.24	23369.73	23343.17	23348.56

TABLE 7: Analytical measures of the fitted models.

Dist.	-2ℓ	AD
Z-Weibull	11640.48	0.840
W-claim	11667.98	0.986
Weibull	11687.82	1.071
B-XII	12096.81	5.139
GE	11921.35	4.965
EL	11928.99	5.012
BW	11665.12	0.952

candidate model for modeling insurance data sets. Furthermore, in support of Tables 6 and 7, the estimated cdf and pdf are plotted in Figure 6. The Kaplan–Meier survival plot and PP plot are sketched in Figure 7, whereas the QQ plot of the proposed distribution and box plot of the earthquake data are presented in Figure 8. From the estimated pdf in Figure 6, it is clear that the proposed distribution provides an adequate fit to the heavy-tailed earthquake data. From Figures 6 and 7, we can easily detect that the proposed distribution fits the estimated cdf and Kaplan–Meier survival plots very closely. The PP and QQ plots which serve as a tool for graphical display of the analytical measures show that the Z-Weibull distribution provides the best fit to real data. Finally, the box plot (Figure 8) of the data is graphical evidence that the data possess tail skewed to the right.

Furthermore, using the earthquake insurance data, we obtained the values of the Kolmogorov–Smirnov (KS) statistic of the proposed and other competing models. Then, we applied the parametric bootstrap technique [30] and bootstrapped the p value for all the competing models. The KS statistic and the corresponding bootstrapped p value are provided in Table 8. Based on the results provided in Table 8, we conclude that the proposed model is the best candidate model among the competing distributions for modeling the insurance claim data.

7. Bayesian Estimation

Bayesian inference procedure has been taken into consideration by many statistical researchers, especially those in the field of survival analysis and reliability engineering. In this section, complete sample data are analyzed through the Bayesian point of view. We assume that the parameters α , γ , and β of the Z-Weibull distribution have independent prior distributions as follows:

$$\begin{aligned}\alpha &\sim \text{Gamma}(a, b), \\ \gamma &\sim \text{Gamma}(c, d), \\ \beta &\sim \text{Gamma}(e, f),\end{aligned}\tag{31}$$

where a, b, c, d, e , and f are positive. More about choosing gamma priors, refer Kundu and Howlader [31], S. Dey and T. Dey [32], Dey et al. [33], and Dey et al. [34]. Hence, the joint prior density function is formulated as follows:

$$\pi(\alpha, \gamma, \beta) = \frac{b^a d^c f^e}{(\Gamma a)(\Gamma c)(\Gamma e)} \alpha^{a-1} \gamma^{c-1} \beta^{e-1} \exp\{-(b\alpha + d\gamma + f\beta)\}.\tag{32}$$

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-

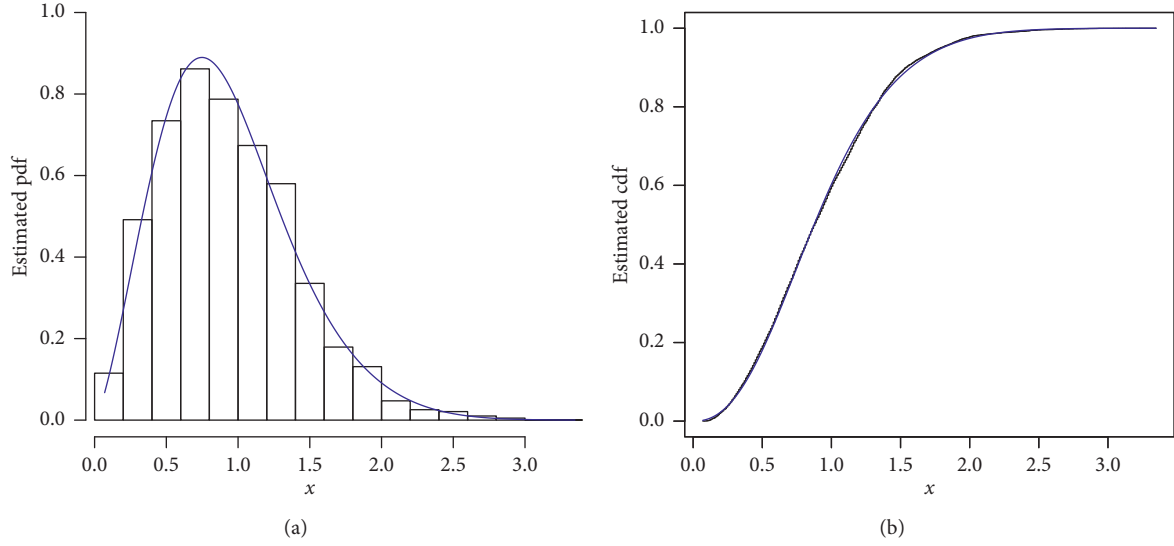


FIGURE 6: Estimated pdf and cdf of the Z-Weibull distribution.

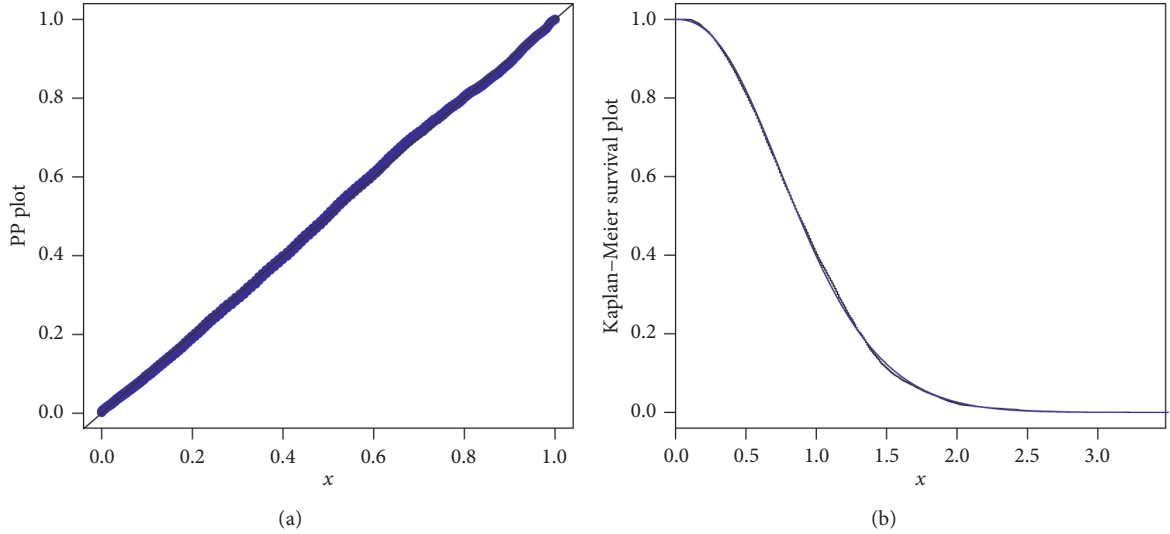


FIGURE 7: PP and Kaplan-Meier survival plots of the Z-Weibull distribution.

known loss functions and associated Bayesian estimators and corresponding posterior risk are presented in Table 9.

For more details, see Calabria and Pulcini [35] and Dey et al. [36]. Next, we provide the posterior probability distribution for a complete data set. We define the function φ as

$$\varphi(\alpha, \gamma, \beta) = \alpha^{a-1} \gamma^{c-1} \beta^{e-1} \exp\{-(b\alpha + d\gamma + f\beta)\}, \quad (33)$$

$$\cdot \alpha > 0, \gamma > 0, \beta > 0, \beta \neq 1.$$

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

$$\pi^*(\alpha, \gamma, \beta | \text{data}) \propto \pi(\alpha, \gamma, \beta) L(\text{data}). \quad (34)$$

Hence, the joint posterior density of parameters α , γ , and β for complete sample data is obtained by combining the

likelihood function and joint prior density (32). Therefore, the joint posterior density function is given by

$$\pi^*(\alpha, \gamma, \beta | \underline{x}) = K \varphi(\alpha, \gamma, \beta | \underline{x}) \prod_{i=1}^n \frac{\alpha \gamma x_i^{\alpha-1} e^{-\gamma x_i^\alpha} (1 + \log(\beta)) e^{-\gamma x_i^\alpha}}{\beta^{1-e^{-\gamma x_i^\alpha}}}, \quad (35)$$

where

$$K^{-1} = \int_0^\infty \int_0^\infty \int_0^\infty \varphi(\alpha, \gamma, \beta) \prod_{i=1}^n \frac{\alpha \gamma x_i^{\alpha-1} e^{-\gamma x_i^\alpha} (1 + \log(\beta)) e^{-\gamma x_i^\alpha}}{\beta^{1-e^{-\gamma x_i^\alpha}}} d\alpha d\gamma d\beta. \quad (36)$$

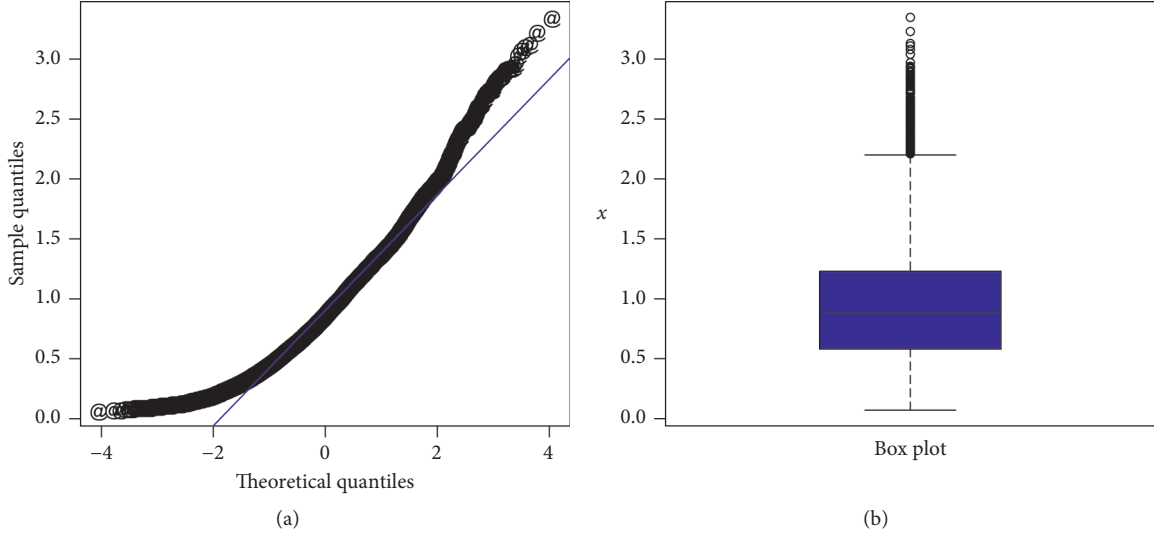


FIGURE 8: QQ plot of the Z-Weibull distribution and box plot of the earthquake insurance data.

TABLE 8: KS and the corresponding bootstrapped p value of the proposed and other models.

Dist.	KS	Bootstrapped p value
Z-Weibull	0.205	0.976
W-claim	0.367	0.806
Weibull	0.476	0.704
B-XII	0.954	0.408
GE	0.590	0.605
EL	0.869	0.502
BW	0.406	0.775

TABLE 9: Bayes estimator and posterior risk under different loss functions.

Loss function	Bayes estimator	Posterior risk
$L_1 = \text{SELF} = (\theta - d)^2$	$E(\theta/x)$	$V(\theta/x)$
$L_2 = \text{WSELF} = ((\theta - d)^2)/\theta$	$(E(\theta^{-1}/x))^{-1}$	$E(\theta/x) - (E(\theta^{-1}/x))^{-1}$
$L_3 = \text{MSELF} = (1 - (d/\theta))^2$	$(E(\theta^{-1}/x))/(E(\theta^{-2}/x))$	$1 - ((E(\theta^{-1}/x)^2)/(E(\theta^{-2}/x)^2))$
$L_4 = \text{PLF} = ((\theta - d)^2)/d$	$\sqrt{E(\theta^2/x)}$	$2(\sqrt{E(\theta^2/x)} - E(\theta/x))$
$L_5 = \text{KLF} = (\sqrt{(d/\theta)} - \sqrt{(d/\theta)})$	$\sqrt{(E(\theta/x))/(E(\theta^{-1}/x))}$	$2(\sqrt{E(\theta/x)E(\theta^{-1}/x)} - 1)$

Moreover, the marginal posterior density of α , γ , and β , assuming that $\Theta = (\alpha, \gamma, \beta)$, is given by

$$\pi(\Theta_i | \underline{x}) = \int_0^\infty \int_0^\infty \pi^*(\Theta_i | \underline{x}) d\Theta_j d\Theta_k, \quad (37)$$

where $i, j, k = 1, 2, 3, i \neq j \neq k$, and also Θ_i is the i -th member of vector Θ .

From (35) and (37), it is clear that there is no closed form for the Bayesian estimators under the five loss functions described in Table 8. Therefore, we use the MCMC procedure based on 10,000 replicates to compute Bayesian estimators.

Because of intractable integrals associated with joint posterior and marginal posterior distributions, one needs to

use numerical software to solve the integral equations numerically. The two most popular MCMC methods are the Metropolis-Hastings algorithm [37, 38] and the Gibbs sampling [39]. Gibbs sampling is a special case of the Metropolis-Hastings algorithm which generates a Markov chain by sampling from the full set of conditional distributions. Often, Bayesian inference requires computing intractable integrals to generate posterior samples. In practice, simulations related to Gibbs sampling are conducted through special software WinBUGS. WinBUGS software was developed in 1997 to simulate data of complex posterior distributions, where analytical or numerical integration techniques cannot be applied. One may also use OpenBUGS software, which is an open-source version of WinBUGS.

TABLE 10: Bayesian estimates and their posterior risks of the parameters under different loss functions based on the earthquake insurance data.

Bayes	$\hat{\alpha}$		$\hat{\gamma}$		$\hat{\beta}$	
Loss functions	Estimate	Risk	Estimate	Risk	Estimate	Risk
SELF	2.34808	0.00048	0.56078	0.00113	2.35742	0.08661
WSELF	2.34787	0.00020	0.55876	0.00201	2.32162	0.03580
MSELF	2.34766	$8.90e-05$	0.55675	0.00359	2.28686	0.01497
PLF	2.34819	0.00020	0.56179	0.00202	2.37572	0.03659
KLF	2.34798	$8.879e-05$	0.55977	0.00360	2.33945	0.01536

Using Gibbs sampling, we obtain samples from the joint posterior distribution and then use OpenBUGS software to carry out the Bayesian analysis.

The process is described as follows:

- (i) Gibbs sampling technique is used to generate posterior samples
- (ii) 10,000 replicates are made to compute the Bayesian estimators via OpenBUGS software
- (iii) The idea of Congdon [40] was to implement and choose $a = b = c = d = e = f = 0.0001$ as we do not have any prior information about hyperparameters

The corresponding Bayesian estimates and posterior risk are provided in Table 10. Table 11 provides 95% credible and HPD intervals for each parameter of the Z-Weibull distribution. Moreover, we provide the posterior summary plots in Figures 9–11. These plots confirm that the sampling process is of the prime quality, and the convergence does occur.

8. Discussion and Future Framework

Statistical decision theory addresses the state of uncertainty and provides a rational framework for dealing with problems of actuarial and financial decision-making. The insurance data sets are generally skewed to the right and heavy-tailed. The traditional distributions are not flexible enough to counter complex forms of data such as insurance science data.

Due to the importance of statistical distributions in actuarial sciences, a number of papers have been appeared in the literature aiming to improve the characteristics of the existing distributions. Although this has been achieved, unfortunately, the numbers of parameters have been increased, and the estimation of parameters and derivation of mathematical problems become complicated.

To provide a better description of the insurance science data, therefore, in this study, an attempt has been made to introduce a new family of statistical distributions aiming to increase the flexibility of the existing distributions. A special submodel of the proposed family offers the best fitting to the heavy-tailed insurance science data. The maximum likelihood method is adopted to estimate the model parameters, and a comprehensive Monte Carlo simulation study is done to evaluate the behavior of the estimators.

To show the usefulness of the proposed method in insurance sciences, a real-life application of the earthquake insurance data is discussed. Analyzing the data set, it showed that the proposed model performs much better than the other competitive distributions.

From the above discussion, it is obvious that the researchers are always in search of new flexible distributions. Therefore, to bring further flexibility in the proposed model, we suggest to introduce its extended versions. The proposed method can further be extended by introducing a shape parameter to the model.

- (i) A random variable X is said to follow the extended version of the Z-family if its cdf is given by

$$G(x; a, \beta, \xi) = 1 - \left\{ \frac{1 - F(x; \xi)^a}{\beta^{F(x; \xi)}} \right\}, \quad a, \beta > 0, x, \xi \in \mathbb{R}, \quad (38)$$

where a is the additional shape parameter. For $a = 1$, expression (38) reduces to (6). The new proposal may be named as the exponentiated Z-family. For the illustrative purposes, one may consider its special sub-case, called the exponentiated Z-Weibull (EZ-Weibull) distribution, defined by the cdf

$$G(x; a, \beta, \xi) = 1 - \left\{ \frac{1 - (1 - e^{-\gamma x^a})^a}{\beta^{(1 - e^{-\gamma x^a})}} \right\}, \quad a, \alpha, \gamma, \beta > 0, x \in \mathbb{R}. \quad (39)$$

Due to the introduction of the additional shape parameter, the suggested extension may be much flexible in modeling data in insurance sciences and other related fields.

- (ii) Another extension of the Z-family is given by

$$G(x; \theta, \beta, \xi) = 1 - \left\{ \frac{1 - F(x; \xi)}{\beta^{F(x; \xi)}} \right\}^\theta, \quad \theta, \beta > 0, x, \xi \in \mathbb{R}, \quad (40)$$

where θ is the additional shape parameter. For $\theta = 1$, expression (40) reduces to (6). The model defined in (40) may be named as the extended Z-family.

TABLE 11: Credible and HPD intervals of the parameters for the earthquake insurance data.

Parameters	Credible interval	HPD interval
α	(2.334, 2.364)	(2.303, 2.388)
γ	(0.536, 0.585)	(0.410, 0.625)
β	(2.139, 2.562)	(1.827, 2.883)

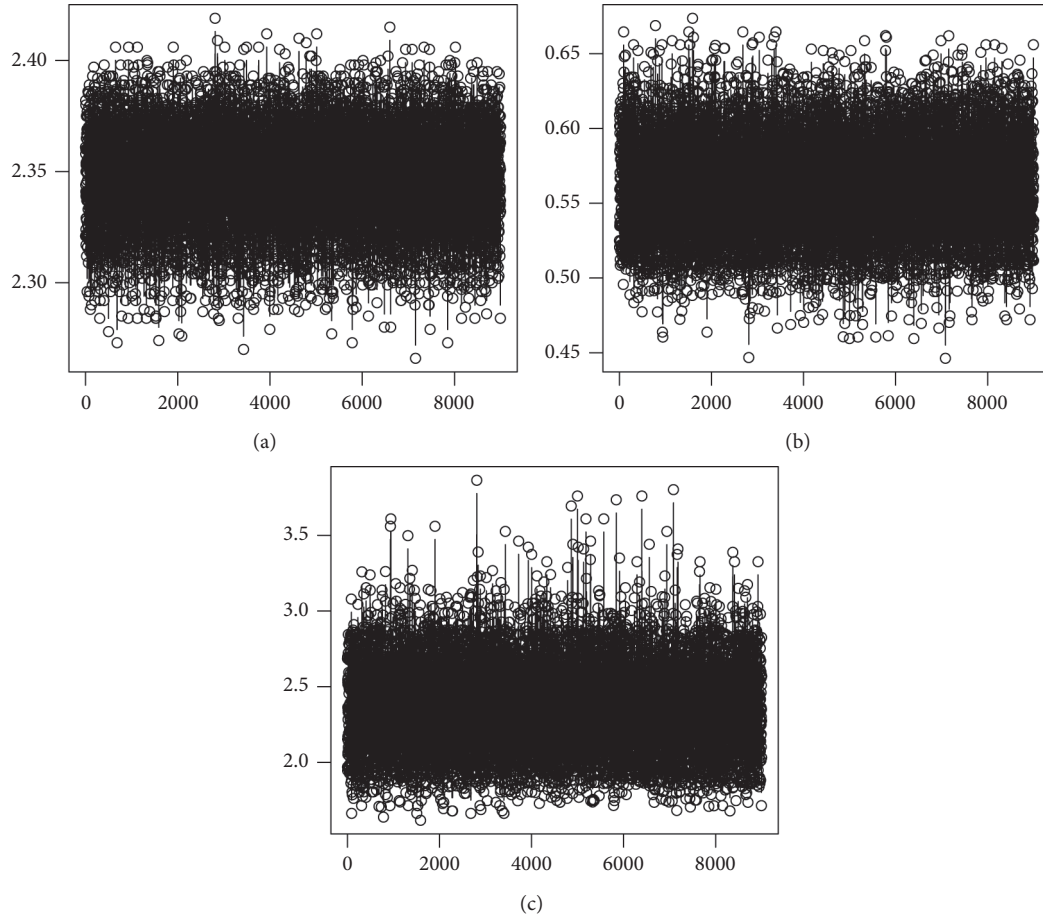


FIGURE 9: Trace plots of each parameter of the Z-Weibull distribution. (a) Alpha. (b) Gamma. (c) Beta.

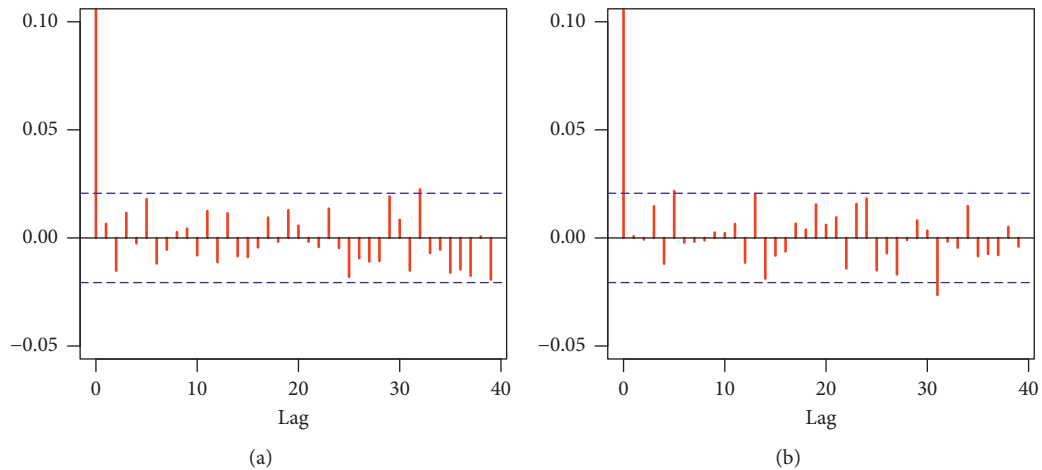
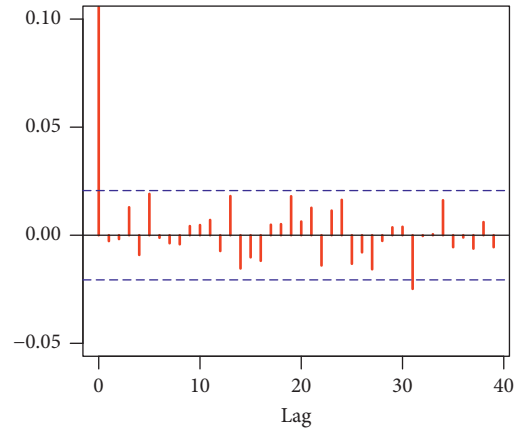
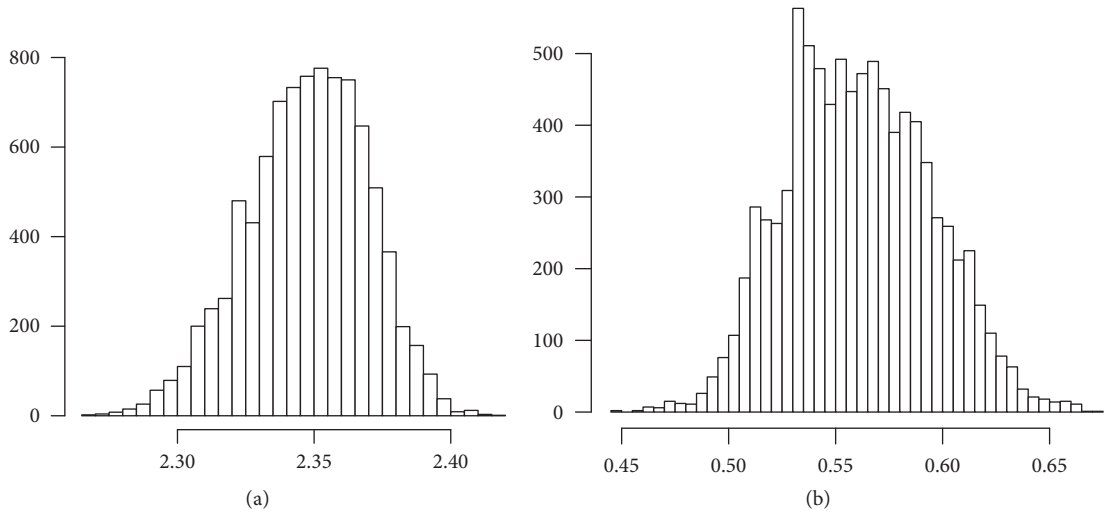


FIGURE 10: Continued.



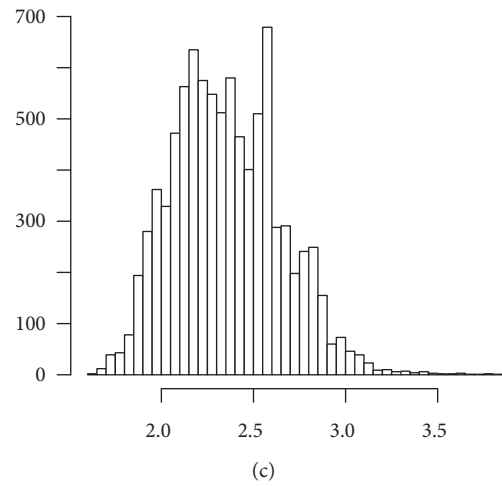
(c)

FIGURE 10: Autocorrelation plots of each parameter of the Z-Weibull distribution. (a) Alpha. (b) Gamma. (c) Beta.



(a)

(b)



(c)

FIGURE 11: Histogram plots of each parameter of the Z-Weibull distribution. (a) Alpha. (b) Gamma. (c) Beta.

- (iii) Another generalized version of the Z-family can be introduced via

$$G(x; a, \theta, \beta, \xi) = 1 - \left\{ \frac{1 - F(x; \xi)^a}{\beta^{F(x; \xi)}} \right\}^\theta, \quad (41)$$

$$a, \theta, \beta > 0, x, \xi \in \mathbb{R},$$

where a and θ are the additional shape parameters. Clearly, for $a = 1$, expression (41) reduces to (40), and for $\theta = 1$, expression (41) reduces to (38). However, for $a = \theta = 1$, expression (41) reduces to (6). The model introduced in (41) may be named as the extended exponentiated Z-family.

- (iv) Another generalized version of the new extended Z-family can be introduced via

$$G(x; a, \theta, \beta, \xi) = 1 - \left\{ \frac{1 - F(x; \xi)^a}{\beta^{F(x; \xi)}} \right\}^\theta, \quad a, \theta, \beta > 0, x, \xi \in \mathbb{R}. \quad (42)$$

9. Concluding Remarks

A variety of methods for proposing new heavy-tailed distributions have been developed to model data related to financial and actuarial sciences. We carried out this area of research further and introduced a new heavy-tailed distribution family. Some distributional properties are derived, and the method of maximum likelihood estimation is discussed to estimate the model parameters. In addition to distributional properties, some actuarial properties are also derived. Based on the actuarial measures, a comprehensive simulation study is conducted. We focused our concentration on a three-parameter special model called the Z-Weibull distribution. To prove the potential and usefulness of the Z-Weibull distribution, earthquake insurance data are analyzed, and its comparison is made with the other well-known distributions. While analyzing the earthquake insurance data, it is observed that the proposed model performs better than the other competitive models. Bayesian analysis using the earthquake data is also provided. Finally, some new extensions are also suggested which may further improve the characteristics of the proposed family.

Appendix

R-Code for Analysis

Note: in the following R-code, a is used for α , g is used for γ , b is used for β , and pm is used for the proposed model.

```
data ← read.csv(file.choose(), header = TRUE)
data = data[1]
data = data[!is.na(data)]
data = data/100
data
hist(data)
```

```
###—Probability density function
pdf_pm ← function(par,x)
{
  a = par[1]
  g = par[2]
  b = par[3]
  (a * g * (x^(a-1)) * (exp(-g * x*a)) * (b*(exp(-g * x*a) - 1))) * (1 + log(b) * exp(-g * x*a))
}
###—Cumulative distribution function.
cdf_pm ← function(par,x)
{
  a = par[1]
  g = par[2]
  b = par[3]
  1 - (exp(-g * x*a)/(b*(1 - exp(-g * x*a))))
}
set.seed(0)
goodness.fit(pdf = pdf_pm,
cdf = cdf_pm,
starts = c(0.5,0.5,0.5), data = data,
method = "BFGS," domain = c(0,Inf),mle = NULL)
%%%%-----
% Estimated pdf
%%%%-----
x ← read.csv(file.choose(), header = TRUE)
x = x[,1]
x = x[!is.na(x)]
x = x/100
x
###—Parameter values
a = 2.346
g = 0.564
b = 2.295
pdf = (a * g
* (x^(a-1)) * (exp(-g * x*a)) * (b*(exp(-g * x*a) - 1))) * (1 + log(b) * exp(-g * x*a))
f = pdf
x = sort(x)
yrange = c(0,1)
xrange = c(min(x),max(x))
hist(x, freq = FALSE, breaks = 15, xlim = xrange,
ylim = yrange,
ylab = "Estimated pdf," xlab = "x," main = "")
par(new = TRUE)
lines(x,f, xlim = xrange, lty = 1, ylab = "y" ylim = yrange,
lwd = 3,
col = "blue," xlab = "")
par(new = TRUE)
```

```

#####
%Estimated cdf
#####
x ← read.csv(file.choose(), header = TRUE)
x = x[,1]
x = x[!is.na(x)]
x = x/100
x
###—Parameter values
a = 2.346
g = 0.564
b = 2.295
x ← sort(x)
F1 ← ecdf(x)
ecdf ← F1(c(x))
zweibullcdf ← 1 - ((exp(-g * x̂a)/
(b(1 - exp(-g * x̂a))))))
plot(x, ecdf, lty = 1, lwd = 2.5, type = "s," xlab = "x,"
ylab = "Estimated cdf," ylim = c(0,1), xlim = c(min(x),
max(x)), col = "black")
par(new = TRUE)
plot(x, zweibullcdf, lty = 1, lwd = 3, type = "l,"
xlab = "x,"
ylab = "Estimated cdf," ylim = c(0,1), xlim = c(min(x),
max(x)), col = "blue")
par(new = TRUE)
#####
%Kaplan-Meier survival plot
#####
x ← read.csv(file.choose(), header = TRUE)
x = x[,1]
x = x[!is.na(x)]
x = x/100
x
###—Parameter values
library(survival)
delta = rep(1,length(x))
x ← sort(x)
km = survfit(Surv(x,delta)~1)
plot(km,conf.int = FALSE, ylab = "Kaplan-Meier survival plot,"xlab = "x")
a = 2.346
g = 0.564
b = 2.295
ss ← function(x)
{
(exp(-g * x̂a)/(b(1 - exp(-g * x̂a))))

```

```

}
lines(seq(0, 3.5, length.out = 100),ss(seq(0, 3.5,
length.out = 100)),
col = "blue," lwd = 3)
#####
%PP plot
#####
x ← read.csv(file.choose(), header = TRUE)
x = x[,1]
x = x[!is.na(x)]
x = x/100
x
cdfLD1 = function(x, a, g, b)
{
1 - (exp(-g * x̂a)/(b(1 - exp(-g * x̂a))))
}
x = sort(x)
n = length(x)
###—Empirical distribution function
Fn = seq(1,n)/n
plot(Fn, cdfLD1(x, 2.346, 0.564, 2.295), xlab = "x,"
ylab = "PP Plot," pch = 21, col = "blue," bg = "blue")
abline(0,1)
#####
%QQ plot
#####
x ← read.csv(file.choose(), header = TRUE)
x = x[,1]
x = x[!is.na(x)]
x = x/100
x
###—Parameter values
a = 2.346
g = 0.564
b = 2.295
x ← sort(x)
F1 ← ecdf(x)
ecdf ← F1(c(x))
zweibullcdf ← 1 - (exp(-g * x̂a)/(b(1 - exp(-g * x̂a))))
qqnorm(x, pch = "@," col = "black," main = "")
qqline(x, pch = "@," col = "blue," ylab = "X," lty = 1,
lwd = 3)
#####
%Box plot
#####
x < read.csv(file.choose(), header = TRUE)
x = x[,1]

```

```
x = x[!is.na(x)]
x = x/100
x
boxplot(x, main = , col = "blue," ylab = "x," xlab = "Box
Plot")
%%%%-----
```

Data Availability

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

Disclosure

This article is drafted from the PhD work of the first author (Zubair Ahmad).

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] M. Guillen, F. Prieto, and J. M. Sarabia, "Modelling losses and locating the tail with the Pareto positive stable distribution," *Insurance: Mathematics and Economics*, vol. 49, no. 3, pp. 454–461, 2011.
- [2] D. Bhati and S. Ravi, "On generalized log-Moyal distribution: a new heavy tailed size distribution," *Insurance: Mathematics and Economics*, vol. 79, pp. 247–259, 2018.
- [3] J. Beirlant, G. Matthys, and G. Dierckx, "Heavy-tailed distributions and rating," *ASTIN Bulletin*, vol. 31, no. 1, pp. 37–58, 2001.
- [4] Z. Ahmad, E. Mahmoudi, G. G. Hamedani, and O. Kharazmi, "New methods to define heavy-tailed distributions with applications to insurance data," *Journal of Taibah University for Science*, vol. 14, no. 1, pp. 359–382, 2020.
- [5] M. Eling, "Fitting insurance claims to skewed distributions: are the skew-normal and skew-student good models?" *Insurance: Mathematics and Economics*, vol. 51, no. 2, pp. 239–248, 2012.
- [6] C. Adcock, M. Eling, and N. Loperfido, "Skewed distributions in finance and actuarial science: a review," *The European Journal of Finance*, vol. 21, no. 13–14, pp. 1253–1281, 2015.
- [7] T. Shushi, "Skew-elliptical distributions with applications in risk theory," *European Actuarial Journal*, vol. 7, no. 1, pp. 277–296, 2017.
- [8] A. Punzo, A. Mazza, and A. Maruotti, "Fitting insurance and economic data with outliers: a flexible approach based on finite mixtures of contaminated gamma distributions," *Journal of Applied Statistics*, vol. 45, no. 14, pp. 2563–2584, 2018.
- [9] A. Azzalini, T. Del Cappello, and S. Kotz, "Log-skew-normal and log-skew- t distributions as models for family income data," *Journal of Income Distribution*, vol. 11, no. 3–4, p. 2, 2002.
- [10] L. Bagnato and A. Punzo, "Finite mixtures of unimodal beta and gamma densities and the k -bumps algorithm," *Computational Statistics*, vol. 28, no. 4, pp. 1571–1597, 2013.
- [11] G. A. Paula, V. Leiva, M. Barros, and S. Liu, "Robust statistical modeling using the Birnbaum-Saunders- t distribution applied to insurance," *Applied Stochastic Models in Business and Industry*, vol. 28, no. 1, pp. 16–34, 2012.
- [12] S. A. Klugman, H. H. Panjer, and G. E. Willmot, *Loss Models: from Data to Decisions*, Vol. 715, John Wiley & Sons, Hoboken, NJ, USA, 2012.
- [13] S. Nadarajah and S. A. A. Bakar, "New composite models for the Danish fire insurance data," *Scandinavian Actuarial Journal*, vol. 2014, no. 2, pp. 180–187, 2014.
- [14] S. A. Bakar, N. A. Hamzah, M. Maghsoudi, and S. Nadarajah, "Modeling loss data using composite models," *Insurance: Mathematics and Economics*, vol. 61, pp. 146–154, 2015.
- [15] A. Punzo, L. Bagnato, and A. Maruotti, "Compound unimodal distributions for insurance losses," *Insurance: Mathematics and Economics*, vol. 81, pp. 95–107, 2018.
- [16] A. Mazza and A. Punzo, "Modeling household income with contaminated unimodal distributions," in *New Statistical Developments in Data Science*, pp. 373–391, Springer, Cham, Switzerland, 2019.
- [17] M. Bernardi, A. Maruotti, and L. Petrella, "Skew mixture models for loss distributions: a Bayesian approach," *Insurance: Mathematics and Economics*, vol. 51, no. 3, pp. 617–623, 2012.
- [18] T. Miljkovic and B. Grün, "Modeling loss data using mixtures of distributions," *Insurance: Mathematics and Economics*, vol. 70, pp. 387–396, 2016.
- [19] K. Dutta and J. Perry, "A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital," Federal Reserve Bank of Boston, Boston, MA, USA, Working Papers, No. 06-13, 2006.
- [20] S. Nasiru, "Extended odd Fréchet-G family of distributions," *Journal of Probability and Statistics*, vol. 2018, Article ID 2931326, 12 pages, 2018.
- [21] M. A. Cortés, D. Elal-Olivero, and J. F. Olivares-Pacheco, "A new class of distributions generated by the extended bimodal-normal distribution," *Journal of Probability and Statistics*, vol. 2018, Article ID 9753439, 10 pages, 2018.
- [22] H. Al-Mofleh, "On generating a new family of distributions using the tangent function," *Pakistan Journal of Statistics and Operation Research*, vol. 14, no. 3, pp. 471–499, 2018.
- [23] M. E. Mead, G. M. Cordeiro, A. Z. Afify, and H. A. Mofleh, "The alpha power transformation family: properties and applications," *Pakistan Journal of Statistics and Operation Research*, vol. 15, pp. 525–545, 2019.
- [24] W. He, Z. Ahmad, A. Z. Afify, and H. Goual, "The arcsine exponentiated-X family: validation and insurance application," *Complexity*, vol. 2020, Article ID 8394815, 18 pages, 2020.
- [25] Z. Ahmad, E. Mahmoudi, S. Dey, and S. K. Khosa, "Modeling vehicle insurance loss data using a new member of T-X family of distribution," *Journal of Statistical Theory & Practice*, vol. 19, no. 2, pp. 133–147, 2020.
- [26] Z. Ahmad, E. Mahmoudi, and S. Dey, "A new family of heavy-tailed distribution with an application to the heavy-tailed insurance loss data," *Communication in Statistics: Simulation & Computation*, vol. 49, 2020.
- [27] A. Alzaatreh, C. Lee, and F. Famoye, "A new method for generating families of continuous distributions," *Metron*, vol. 71, no. 1, pp. 63–79, 2013.
- [28] Z. Ahmad, G. G. Hamedani, and N. S. Butt, "Recent developments in distribution theory: a brief survey and some new generalized classes of distributions," *Pakistan Journal of*

- Statistics and Operation Research*, vol. 15, no. 1, pp. 87–110, 2019.
- [29] P. Artzner, “Application of coherent risk measures to capital requirements in insurance,” *North American Actuarial Journal*, vol. 3, no. 2, pp. 11–25, 1999.
 - [30] W. Stute, W. G. Manteiga, and M. P. Quindimil, “Bootstrap based goodness-of-fit-tests,” *Metrika*, vol. 40, no. 1, pp. 243–256, 1993.
 - [31] D. Kundu and H. Howlader, “Bayesian inference and prediction of the inverse Weibull distribution for type-II censored data,” *Computational Statistics & Data Analysis*, vol. 54, no. 6, pp. 1547–1558, 2010.
 - [32] S. Dey and T. Dey, “On progressively censored generalized inverted exponential distribution,” *Journal of Applied Statistics*, vol. 41, no. 12, pp. 2557–2576, 2014.
 - [33] S. Dey, S. Ali, and C. Park, “Weighted exponential distribution: properties and different methods of estimation,” *Journal of Statistical Computation and Simulation*, vol. 85, no. 18, pp. 3641–3661, 2015.
 - [34] S. Dey, T. Dey, S. Ali, and M. S. Mulekar, “Two-parameter Maxwell distribution: properties and different methods of estimation,” *Journal of Statistical Theory and Practice*, vol. 10, no. 2, pp. 291–310, 2016.
 - [35] R. Calabria and G. Pulcini, “Point estimation under asymmetric loss functions for left truncated exponential samples,” *Communications in Statistics-Theory and Methods*, vol. 25, no. 3, pp. 585–600, 1996.
 - [36] S. Dey, S. Singh, Y. M. Tripathi, and A. Asgharzadeh, “Estimation and prediction for a progressively censored generalized inverted exponential distribution,” *Statistical Methodology*, vol. 32, pp. 185–202, 2016.
 - [37] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines,” *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.
 - [38] W. K. Hastings, “Monte Carlo sampling methods using Markov chains and their applications,” *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.
 - [39] S. Geman and D. Geman, “Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.
 - [40] P. Congdon, *Bayesian Statistical Modelling*, John Wiley & Sons, New York, NY, USA, 2001.