Insurance: Mathematics and Economics 79 (2018) 247–259

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

On generalized log-Moyal distribution: A new heavy tailed size distribution

Deepesh Bhati *, Sreenivasan Ravi

Department of Studies in Statistics, University of Mysore, Manasagangotri, Mysuru 570006, India

ARTICLE INFO

Article history:
Received June 2017
Received in revised form February 2018
Accepted 12 February 2018
Available online 16 February 2018

Keywords:
Danish fire insurance losses
Heavy tailed distributions
Limited expected value
Log-Moyal distribution
Norwegian fire insurance losses
Regression modelling
Vehicle insurance losses

ABSTRACT

A new class of distributions, the generalized log-Moyal, suitable for modelling heavy tailed data is proposed in this article. This class exhibits desirable properties relevant to actuarial science and inference. The proposed distribution can be related to some well known distributions like Moyal, folded-normal and chi-square. Statistical inference of the model parameters is discussed using the method of quantiles and the method of maximum likelihood estimation. Three celebrated data sets, namely, Norwegian fire insurance losses, Danish fire insurance losses and vehicle insurance losses, are used to show the applicability of the new class of distributions. Parametric regression modelling is discussed assuming that the response variable follows the generalized log-Moyal distribution.

1. Introduction

Actuaries are often in search of distributions suitable for modelling insurance losses. Such data sets are skewed to the right, unimodal, hump-shaped and possess thick right tail. Realizing these, practitioners have proposed various probability models suitable for modelling these characteristics. However, probability distributions which accommodate some or all of these characteristics are few in number and hence deriving new models from the existing distributions attracts more researchers.

One of the important tasks in actuarial risk modelling is to have accurate prediction of losses occurring with very high monetary value. Underestimation of the probability of such losses leads to serious operational risk such as underestimating premium, bankruptcy, etc. In such circumstances, heavy tailed distributions give reasonably good fits for the right tail (McNeil, 1997; Resnick, 1997). Classical distributions, such as log-normal, Pareto, log–logistic, Fréchet, Lomax, gamma, inverse Gaussian and others (see Hogg and Klugman, 2009 and Klugman et al., 2012), are prominently used in modelling such losses. Distributions such as log-normal, gamma, Weibull and Inverse Gaussian are suitable for small size losses whereas Pareto, log-logistic, Fréchet and Lomax are suitable for large losses. In view of these, various generalizations have been proposed by adding new parameter(s) to these classical distributions. Mainly these generalizations are based on, but not limited to, the following four approaches (i) transformation method, (ii) composition of two or more distributions, (iii) compounding of distributions, and (iv) finite mixture of distributions.

Skewed distributions are popular models since they acclimatize right-skewness and high kurtosis. Recent works of Vernic (2006), Adcock et al. (2015), Kazemi and Noorizadeh (2015) and Eling (2012) identify that skew-normal and skew student t-distributions are two such promising models for insurance losses or financial risk modelling. However, insurance losses and financial risks take values on the positive real line and hence these skew classes of distributions may not be appropriate as they are real valued. In such cases, transformation of variables, especially exponential transformation, has proved to be significant (see for example, log-skew normal in Azzalini et al., 2003, and log-elliptical distributions in Landsman et al., 2016). Bagnato and Punzo (2013) point out that the transformation is simple to apply but often inference as well as computation of many distributional characteristics becomes very difficult. As we shall see later, though the proposed model in this article is obtained by an exponential transformation, closed form expressions are possible for distributional characteristics and inference, which are quite desirable. This is a key motivation for considering the proposed model.

Another popular method for obtaining a new family of distributions which gives reasonably good fit for light, moderate and heavy losses is the method of composition (Coorey and Ananda, 2005; Klugman et al., 2012). Distributions such as log-normal (Coorey and Ananda, 2005) or Weibull (Ciumara, 2006; Scollnik and Sun,

* Corresponding author.
E-mail addresses: deepesh.bhati@curaj.ac.in (D. Bhati), ravi@statistics.uni-mysore.ac.in (S. Ravi).

https://doi.org/10.1016/j.insmatheco.2018.02.002
0167-6687/

2012; Bakar et al., 2015) have been used for modelling the data up to a specified threshold whereas heavy tailed distributions such as Pareto (Coorey and Ananda, 2005), Lomax (Scollnik, 2007), Burr (Nadarajah and Bakar, 2014), log-logistic, para-logistic, generalized Pareto, inverse Burr, inverse Pareto, inverse para-logistic (Bakar et al., 2015) and Stoppa (Calderín-Ojeda and Kwok, 2016) have been used to model the data tail beyond the threshold. However, these models use fixed and a priori known mixing weights, and hence they can be very restrictive as observed in Calderín-Ojeda and Kwok (2016). To overcome this issue, Scollnik (2007) used unrestricted mixing weights. Recently, Reynkens et al. (2017) propose another splicing model by considering mixed Erlang distribution for the body and Pareto distribution for the tail and further extend this splicing approach for censored and/or truncated data. It should be noted that the models obtained by splicing usually involve more than three parameters.

Recently, Punzo et al. (2017) introduced a three-parameter compound distribution in order to take care of specifics such as uni-modality, hump-shaped, right-skewed and heavy tails. As indicated in Punzo et al. (2017), the resultant density obtained by this method may not always have closed form expressions which make the estimation more cumbersome. Also, not all moments are available in closed form and it has no skew parameter.

Another method which deals in modelling multi-modal insurance losses data sets is the k-component finite mixture models from parametric, non-Gaussian families of distributions. Bernardi et al. (2012) considered the skew normal mixture to model losses. Later, Lee and Lin (2010) and Verbelen et al. (2015) considered finite mixtures of Erlang distributions. Recently, this work has been extended by Miljkovic and Grün (2016) considering finite mixtures of Burr, Gamma, inverse Burr, inverse Gaussian, log-normal and Weibull.

Gómez-Déniz and Calderín-Ojeda (2015) proposed a different method to add a parameter to a family of distributions after making a change of variable in the truncated Cauchy distribution.

Though bringing flexibility to a model by introducing additional parameter(s) is a desirable feature, it makes the inference more complicated and computationally challenging. In particular, parameter estimation and model diagnostics become difficult (Brazauskas and Kleefeld, 2011). Often in the field of fire, motor, third party liability, catastrophe and other general insurance, the claim size distribution is modelled by heavy tailed distribution or by a distribution having tail heavier than exponential. Hence it is always of interest to propose a new class of distributions having tail heavier than exponential. The proposed two parameter model in this article is obtained by transformation and has uni-modality, right skewness and has tail heavier than exponential, which are desirable properties. This model also provides a better description of data with possibly heavy tails than the available two parameter distributions prevalent in the actuarial literature such as the classical Pareto, Burr, inverse Gamma, Fréchet, inverse para-logistic, log-logistic and possibly many others. Moreover, it possesses almost all the nice distributional properties such as moments of all orders, mean excess function, entropy, value-at-risk (VaR) and tail VaR (TVaR). Estimation of the model parameters using the method of quantiles and the method of maximum likelihood estimation have been carried out. More importantly, regression with the response variable assumed to follow the proposed model in the presence of covariates is studied, and the model seems to give better results than the existing models such as the Fréchet regression model.

The remainder of the article is structured as follows. The model is introduced in Section 2.1 and its relationships with other known distributions are shown in Section 2.2. Section 2.3 presents computation of some quantities like mode, quantile function, moments and entropy. Estimation of parameters is discussed in Section 3. Distribution fit with two celebrated insurance data sets is discussed in Section 4. Regression modelling is discussed in Section 5 with an illustration on vehicle insurance losses data set, followed by conclusions in Section 6.

2. A new heavy tailed size distribution

To begin with, some mathematical formulae used throughout are listed along with some relationships and series expansions.

(i) Incomplete gamma function:

$$\Gamma(a,z) = \int_z^\infty t^{a-1}e^{-t}dt, \quad z>0, \quad Re(a)>0, \\ = \Gamma(a) - \sum_{n=0}^{\infty}\frac{(-1)^n z^{a+n}}{n!(a+n)}. \tag{1}$$

(ii) Error function:

$$\begin{aligned}\texttt{erf}(z) &= \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt, \quad z>0;\\ &= \frac{1}{\sqrt{\pi}}\Gamma\left(\frac{1}{2},z^2\right), \\ &\qquad \text{(see 7.11.2 of Olver et al., 2010);}\\ &= \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n z^{2n+1}}{n!(2n+1)}.\end{aligned} \tag{2}$$

(iii) Complementary error function: $\texttt{erfc}(z) = \frac{2}{\sqrt{\pi}}\int_z^\infty e^{-t^2}dt = 1-\texttt{erf}(z); \quad z>0.$

2.1. The generalized log-Moyal distribution

Definition. Let

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma y}\left(\frac{\mu}{y}\right)^{1/2\sigma} e^{-\frac{1}{2}\left(\frac{\mu}{y}\right)^{1/\sigma}}, \quad y>0, \quad \mu>0, \sigma>0, \tag{3}$$

where μ and σ are, respectively, the scale and the shape parameters. It is easy to verify that f is a probability density function (pdf). Note that, if the random variable (rv) $X = \left(\frac{\mu}{Y}\right)^{1/2\sigma}$, with rv Y having pdf f, then X has a standard half-normal distribution. Also, if rv Z has the pdf $g(.)$ proposed by Moyal (1955): $g(z) = \frac{1}{\sqrt{2\pi}}\exp\left(-(z+\exp(-z))/2\right), \; z\in(-\infty,\infty)$, then $Y=\mu e^{\sigma Z}$. We call the pdf f in (3) as the **generalized log-Moyal** pdf and denote it by GlogM(μ,σ). The pdf (3) can also be represented in the form of a first order linear differential equation with boundary condition as stated in the following proposition whose proof is omitted since it is easy.

Proposition 1. *If $Y \sim$ GlogM(μ,σ), then its pdf f satisfies*

$$2y\sigma f'(y) + f(y)\left(2\sigma+1-\left(\frac{\mu}{y}\right)^{1/\sigma}\right) = 0,$$

with $f(1) = \frac{\mu^{1/2\sigma}}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\mu^{1/\sigma}}$.

To the best of our knowledge, Moyal distribution or any of its extensions have not been explored except the beta-Moyal distribution proposed in Cordeiro et al. (2012). This is one other motivation to study GlogM(μ,σ).

Distribution function: Using (2), the distribution function (df) of the generalized log-Moyal distribution is given by

$$F(y) = \frac{1}{\Gamma(\frac{1}{2})}\Gamma\left(\frac{1}{2},\frac{1}{2}\left(\frac{\mu}{y}\right)^{1/\sigma}\right) = \texttt{erfc}\left(\frac{1}{\sqrt{2}}\left(\frac{\mu}{y}\right)^{1/2\sigma}\right), \\ y>0. \tag{4}$$

Plots of the GlogM(μ,σ) pdf: Fig. 1 gives plots of the pdf f for different parameter values.

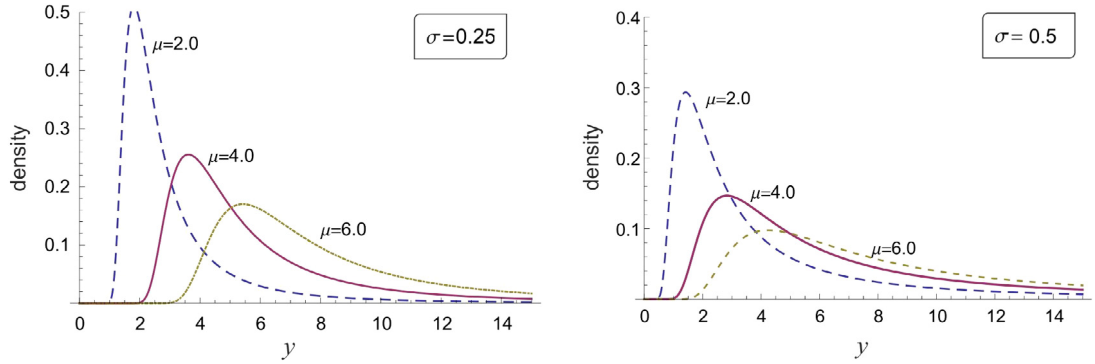

Fig. 1. Plots of the pdf of GlogM(μ, σ) for different parameter values.

2.2. Relationship of the generalized log-Moyal distribution with other distributions

From theoretical and practical perspectives, it is always of interest to know whether a class of distributions has some relation with other classes of distributions since such relationships help in understanding the distribution better and also in the estimation of parameters. The following proposition gives the relationships between GlogM(μ, σ) and some well known families of distributions and these relationships follow from definitions.

Proposition 2.

(i) *If the rv Z follows the standard Moyal distribution (see* Moyal, 1955*) then the rv $Y = \mu e^{\sigma Z} \sim$ GlogM(μ, σ).*
(ii) *If the rv X follows the half-normal $(0, \sigma^2)$ distribution, then the rv $Y = \mu\left(\frac{\sigma}{X}\right)^{2\sigma} \sim$GlogM($\mu, \sigma$).*
(iii) *If the rv X follows the generalized half-normal(θ, α) distribution,* Coorey and Ananda (2008)*, then the rv $Y = 1/X \sim$GlogM $(\mu = 1/\theta, \sigma = 1/2\alpha)$.*
(iv) *If the rv $X \sim \chi^2_{1/2}$, then the rv $Y = \mu X^{-\sigma} \sim$GlogM(μ, σ).*

2.3. Distributional properties of the generalized log-Moyal distribution

Some distributional properties of GlogM(μ, σ) are presented here and some of these will be used later.

The rth moment about the origin of $Y \sim$ GlogM(μ, σ) is given by

$$\mathbb{E}(Y^r) = \int_0^\infty y^r f(y) dy = \frac{\mu^r}{2^{r\sigma}\sqrt{\pi}} \Upsilon_\sigma(r), \qquad r < \frac{1}{2\sigma},$$

where $\Upsilon_\sigma(r) = \Gamma\left(\frac{1}{2} - r\sigma\right)$. Hence the mean and the variance of Y are given by

$$\mathbb{E}(Y) = \frac{\mu}{2^\sigma\sqrt{\pi}} \Upsilon_\sigma(1), \quad \sigma < 1/2,$$

$$\mathbb{V}(Y) = \frac{\mu^2}{2^{2\sigma}\sqrt{\pi}}\left(\Upsilon_\sigma(2) - \frac{1}{\sqrt{\pi}}\Upsilon_\sigma^2(1)\right), \quad \sigma < 1/4.$$

Proposition 3. *If $Y \sim$GlogM(μ, σ), then*

(i) $\mathbb{E}\left(\frac{1}{Y^r}\right) = \frac{2^{r\sigma}}{\sqrt{\pi}\mu^r}\Upsilon_\sigma(-r)$*, provided* $r > -\frac{1}{2\sigma}$;
(ii) $\mathbb{E}\left(\frac{\mu}{Y}\right)^{1/\sigma} = 1$;
(iii) $\mathbb{E}\left(\log\left(\frac{\mu}{Y}\right)\right) = -\sigma(\gamma + \log 2)$;
(iv) $\mathbb{E}\left(\left(\frac{\mu}{Y}\right)^{1/\sigma} \log\left(\frac{\mu}{Y}\right)\right) = \sigma(2 - \gamma - \log 2)$;
(v) $\mathbb{E}\left(\left(\frac{\mu}{Y}\right)^{1/\sigma} \log^2\left(\frac{\mu}{Y}\right)\right) = \frac{1}{2}\sigma^2\left(\pi^2 + (\log 2 - 4)\log 4 + 2\gamma(\gamma - 4 + \log 4)\right)$;

where $\gamma = 0.577216$ is Euler's gamma constant.

Proof. The proofs follow by straight forward computations using the integral representation of Euler's gamma constant, $\gamma = -\int_0^\infty e^{-t}\log t dt$. ■

The above proposition will be used later in determining the Shannon entropy and the Fisher information matrix.

2.3.1. Quantile function of the generalized log-Moyal distribution

The ϱth quantile $y_\varrho = \min\{y : F(y) \geq \varrho\}$, $0 < \varrho < 1$, of GlogM(μ, σ) is obtained by inverting the df (4) and is given by

$$y_\varrho = \mu\left(\sqrt{2}\,\texttt{erfc}^{\leftarrow}(\varrho)\right)^{-2\sigma},$$

where $\texttt{erfc}^{\leftarrow}$ is the inverse of $\texttt{erfc}$. The median of $Y \sim$ GlogM (μ, σ) is obtained by setting $\varrho = 1/2$, and is given by

$$y_{1/2} = \mu\left(\sqrt{2}\,\texttt{erfc}^{\leftarrow}(1/2)\right)^{-2\sigma} = \mu(0.6745)^{-2\sigma}.$$

2.3.2. Uni-modality of the generalized log-Moyal distribution

By taking the derivative of the logarithm of the pdf (3) and equating it to zero, we solve for the mode y_0 of $Y \sim$ GlogM(μ, σ), to get

$$y_0 = \frac{\mu}{(1 + 2\sigma)^\sigma}.$$

This shows that GlogM(μ, σ) is uni-modal. Fig. 2 gives more clarity about the skewness of $Y \sim$ GlogM(μ, σ). It can also be observed that mean > median > mode which indicates positive skewness.

2.3.3. The Shannon entropy of the generalized log-Moyal distribution

Proposition 4. *The Shannon entropy of $Y \sim$ GlogM(μ, σ) is given by*

$$\mathbb{H}(f) = \log(\sqrt{2\pi}\sigma) + \log\mu + \left(\frac{1}{2} + \sigma\right)(\gamma + \log 2) + \frac{1}{2}. \tag{5}$$

Proof. The Shannon entropy of the rv Y with pdf (3) is given by $\mathbb{H}(f) = -\int_0^\infty f(y)\log f(y) dy$. Using Proposition 3, we get (5). ■

2.3.4. Inverted bathtub shape of hazard rate of the generalized log-Moyal distribution

The hazard rate of GlogM(μ, σ) can be simply obtained by taking ratio of the pdf to the survival function and the limiting value of derivative of its reciprocal characterizes its tail, see Theorem 1.118, Ferreira and de Haan (2006). The hazard rate of Y having pdf (3) is

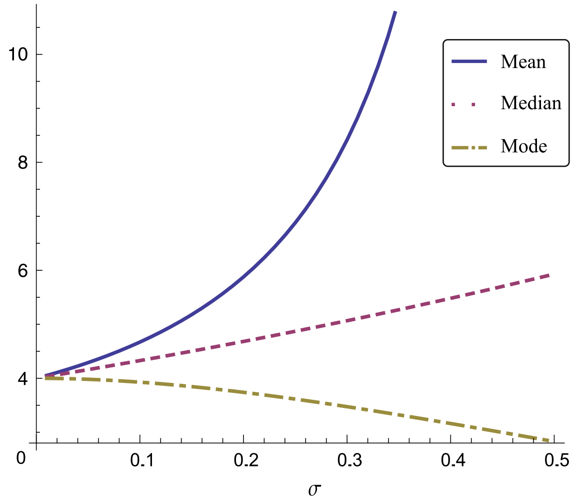

Fig. 2. Pattern of mean, median and mode of the generalized log-Moyal law for $\mu = 4$ and different values of σ.

given by

$$h(y) = \frac{1}{\sqrt{2\pi}\sigma y \cdot \mathtt{erf}\left(\frac{1}{\sqrt{2}}\left(\frac{\mu}{y}\right)^{\frac{1}{2\sigma}}\right)}\left(\frac{\mu}{y}\right)^{\frac{1}{2\sigma}} \times \exp\left(-\frac{1}{2}\left(\frac{\mu}{y}\right)^{1/\sigma}\right). \tag{6}$$

Theorem 1. *The hazard rate of $Y \sim$ GlogM(μ, σ) given in (6) has inverted bathtub shape.*

Proof. To obtain the shape of (6), we use a result in Glaser (1980). Consider the functions $\eta(.)$ and $\xi(.)$ defined as

$$\eta(y) = -\frac{f'(y)}{f(y)} = \frac{1 + 2\sigma - \left(\frac{\mu}{y}\right)^{1/\sigma}}{2y\sigma} \quad \text{and}$$

$$\xi(y) = \frac{\sqrt{2\pi}\sigma y \cdot \mathtt{erf}\left(\frac{1}{\sqrt{2}}\left(\frac{y}{\mu}\right)^{-\frac{1}{2\sigma}}\right)}{\left(\frac{\mu}{y}\right)^{\frac{1}{2\sigma}} \exp\left(-\frac{1}{2}\left(\frac{\mu}{y}\right)^{1/\sigma}\right)}.$$

The first derivatives of $\eta(.)$ and $\xi(.)$ are, respectively, given by

$$\eta'(y) = \frac{1}{2y^2\sigma^2}\left((1+\sigma)\left(\frac{\mu}{y}\right)^{1/\sigma} - \sigma(2\sigma + 1)\right),$$

$$\text{and} \quad \xi'(y) = -\sqrt{\frac{\pi}{2}}\frac{\left(-2\sigma + \left(\frac{\mu}{y}\right)^{1/\sigma} - 1\right)\mathtt{erf}\left(\frac{1}{\sqrt{2}}\left(\frac{y}{\mu}\right)^{-\frac{1}{2\sigma}}\right)}{e^{-\frac{1}{2}\left(\frac{y}{\mu}\right)^{-1/\sigma}}\left(\frac{y}{\mu}\right)^{-\frac{1}{2\sigma}}} - 1.$$

Further, $\eta'(y_0) \gtreqless 0$ for $y \lesseqgtr y_0$, where $y_0 = \mu\left(\frac{\sigma(2\sigma+1)}{1+\sigma}\right)^{-\sigma}$ and $\lim_{y\to 0}\xi'(y) = -\infty$. Also, $\lim_{y\to\infty}\xi'(y) = 2\sigma(>0)$, which implies that there exists $y_0 \in (0, \infty)$ such that $\xi'(y_0) = 0$. Hence from Theorem 1(d) of Glaser (1980), we conclude that the shape of the hazard rate is inverted bathtub. ■

Shapes of hazard rate of the generalized log-Moyal distribution for various parameter values are depicted in Fig. 3.

2.4. Regularly varying tail behaviour of the generalized log-Moyal distribution

A df F with survival function (sf), $\bar{F}(.) = 1 - F(.)$, is said to be heavy tailed if for every $t > 0$, $\lim_{y\to\infty}\frac{\bar{F}(y)}{e^{-ty}} = \infty$. (see Section 2.5.1 of Rolski et al., 1999). A df F is said to belong to the regularly varying class if

$$\lim_{t\to\infty}\frac{\bar{F}(ty)}{\bar{F}(t)} = y^{-\gamma} \quad \forall\, y > 0,$$

denoted as $\bar{F} \in \mathtt{RV}_{-\gamma}$, where γ is called the tail index.

Further note that, the sf of GlogM(μ, σ) is $\bar{F}(y) = \frac{2}{\sqrt{\pi}}\sum_{n=1}^{\infty}\frac{(-1)^n}{n!(2n+1)2^{\frac{2n+1}{2}}}\left(\frac{\mu}{y}\right)^{\frac{2n+1}{2\sigma}}$, and has polynomially decaying tail which decays slower than that of the exponential sf e^{-ty}, $t > 0$. Hence, $\lim_{y\to\infty}\frac{\bar{F}(y)}{e^{-ty}} = \infty$ confirms that the GlogM(μ, σ) is heavy tailed.

Also, we have

$$\begin{aligned}\lim_{y\to\infty}\frac{\bar{F}(yt)}{\bar{F}(y)} &= \lim_{y\to\infty}\frac{\frac{2}{\sqrt{\pi}}\sum_{n=1}^{\infty}\frac{(-1)^n}{n!(2n+1)2^{\frac{2n+1}{2}}}\left(\frac{\mu}{yt}\right)^{\frac{2n+1}{2\sigma}}}{\frac{2}{\sqrt{\pi}}\sum_{n=1}^{\infty}\frac{(-1)^n}{n!(2n+1)2^{\frac{2n+1}{2}}}\left(\frac{\mu}{y}\right)^{\frac{2n+1}{2\sigma}}} \\ &= \lim_{y\to\infty}\frac{\frac{1}{t^{2\sigma}}\left(1 + o\left(\frac{1}{y^2}\right)\right)}{\left(1 + o\left(\frac{1}{y^2}\right)\right)} = t^{-2\sigma},\end{aligned} \tag{7}$$

and hence $\bar{F} \in \mathtt{RV}_{-2\sigma}$. Moreover, by Theorem 3.3.7, Embrechts et al. (2003), the df F belongs to the Fréchet max domain of attraction, which means that $c_n^{-1}\max(Y_1, \ldots, Y_n) \xrightarrow{d} W$, where $c_n = F^{\leftarrow}(1 - 1/n)$ and $Y_1, \ldots, Y_n$ are independent and identically distributed (iid) rvs having common df F, with rv W having the Fréchet distribution with parameter 2σ. ■

2.5. Actuarial measures

Some important actuarial measures for claim size distributions are discussed here.

(i) **LEV:** Most of the insurance companies transfer their financial risk over a threshold u by reinsurance. In deciding the reinsurance premium, one of the prime measures is the "expected value on or below the threshold u," known as the limited expected value (LEV). If the rv Y is the claim size with pdf f, then

$$\mathtt{LEV}_u(Y) = \mathbb{E}(Y \wedge u) = \int_0^u yf(y)dy + u\int_u^\infty f(y)dy.$$

If the rv $Y \sim$ GlogM(μ, σ), then

$$\mathtt{LEV}_u(Y) = \frac{2^{-\sigma}\mu}{\sqrt{\pi}}\Gamma\left(\frac{1}{2} - \sigma, \frac{1}{2}\left(\frac{\mu}{u}\right)^{1/\sigma}\right) + u \cdot \mathtt{erf}\left(\frac{1}{\sqrt{2}}\left(\frac{\mu}{u}\right)^{1/2\sigma}\right), \quad \sigma < 1/2.$$

(ii) **VaR:** To ensure the insolvency of the insurer with a specified degree of certainty, the actuarial measure Value-at-Risk (VaR) is widely used by practitioners. VaR of a rv Y is the qth quantile of its df (Artzner, 1999). If the rv $Y \sim$ GlogM(μ, σ), then

$$\mathtt{VaR}_q(Y) = \mu\left(\sqrt{2} \cdot \mathtt{erfc}^{\leftarrow}(q)\right)^{-2\sigma}.$$

(iii) **TVaR:** Artzner (1999) questioned the use of the VaR because it is not closed under sub-additivity which is a desirable property and hence suggested the use of the tail VaR (TVaR),

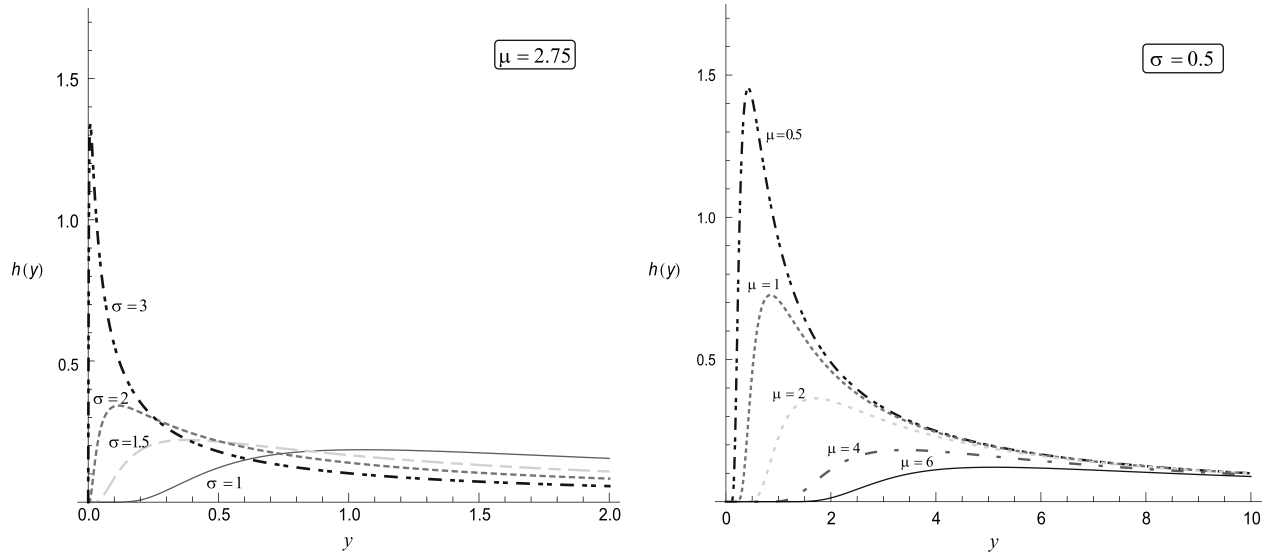

Fig. 3. Hazard rate of GlogM(μ, σ) for different values of σ with $\mu = 2.75$ on left hand side and for different values of μ with $\sigma = 0.5$ on the right side.

defined as $\texttt{TVaR} = \mathbb{E}(Y|Y > \texttt{VaR}_q(Y))$. If the rv $Y \sim$ GlogM(μ, σ), then

$$\begin{aligned}\texttt{TVaR}_q(Y) &= \frac{1}{1-q}\int_{\texttt{VaR}_q(Y)}^{\infty} y f_Y(y)dy,\\ &= \frac{2^{-\sigma}\mu}{1-q}\left(\Upsilon_\sigma(1) - \Gamma\left(\frac{1}{2}, (\texttt{erfc}^{\leftarrow}(1-q))^2\right)\right),\\ &\sigma < 1/2.\end{aligned}$$

(iv) **Excess of loss reinsurance:** To reduce the financial risk, a primary insurance company re-insures with another insurance company known as reinsurance company and the purpose of the reinsurance company is to pay the claim when it exceeds some pre-specified amount or when the claim size is within some reinsurance layers. Hence it is always of interest to the reinsurance company to compute the expected loss within reinsurance layers. Consider a single reinsurance layer where excess in loss is paid when the loss is within $(m, m+l)$ for some m, l positive. The actuarial measure $\mathbb{E}(\min(l, \max(0, Y-m)))$ is called the excess-of-loss reinsurance. If the loss rv $Y \sim$ GlogM(μ, σ), then the excess-of-loss reinsurance is given as

$$\begin{aligned}&\mathbb{E}(\min(l, \max(0, Y-m)))\\ &= \int_m^{m+l}(x-m)f_Y(y)dy + l\bar{F}(m+l),\\ &= (m+l)\cdot\texttt{erf}\left(\frac{1}{\sqrt{2}}\left(\frac{\mu}{l+m}\right)^{1/2\sigma}\right)\\ &\quad - m\cdot\texttt{erf}\left(\frac{1}{\sqrt{2}}\left(\frac{\mu}{m}\right)^{1/2\sigma}\right)\\ &\quad + \frac{\mu 2^{-\sigma}}{\sqrt{\pi}}\left(\Gamma\left(\frac{1}{2}-\sigma, \frac{1}{2}\left(\frac{\mu}{m+l}\right)^{1/\sigma}\right)\right.\\ &\quad \left.- \Gamma\left(\frac{1}{2}-\sigma, \frac{1}{2}\left(\frac{\mu}{m}\right)^{1/\sigma}\right)\right).\end{aligned}$$

(v) **The mean excess function:** The mean excess function $e(u)$ for a claim size rv Y is the expected value of excess claim on a policy with a fixed deductible amount u. If the rv $Y \sim$ GlogM(μ, σ), then

$$\begin{aligned}e(u) &= \mathbb{E}(Y-u|Y>u) = \frac{1}{\bar{F}(u)}\int_u^\infty tf(t)dt - u,\\ &= \frac{\mu 2^{-\sigma}\left(\Upsilon_\sigma(1) - \Gamma\left(\frac{1}{2}-\sigma, \frac{1}{2}\left(\frac{\mu}{u}\right)^{1/\sigma}\right)\right)}{\sqrt{\pi}\,\texttt{erf}\left(\frac{1}{\sqrt{2}}\left(\frac{\mu}{u}\right)^{1/2\sigma}\right)} - u,\\ &\sigma < 1/2.\end{aligned} \tag{8}$$

As in (7), since the GlogM(μ, σ) distribution has a regularly varying tail, from Pickands–Balkema–de Haan Theorem (Theorem 7.20, McNeil et al., 2005), the corresponding excess of loss rv $(Y-u|Y > u)$ converges in distribution to a rv having the generalized Pareto distribution as $u \to \infty$. Therefore the mean excess function in (8) tends to a linear function as $u \to \infty$. Also, by using the series expansions given in (1) and (2), the mean excess function in (8) can be represented as $e(u) = \frac{2\sigma u}{1-2\sigma}(1+o(1))$ which again shows that it tends to a linear function as $u \to \infty$. We use this property in Section 4.1 as a diagnostic tool to check if the GlogM(μ, σ) distribution is a good candidate for a data set.

2.6. Gamma and inverse Gaussian as conjugate priors for parameters of the generalized log-Moyal distribution

We re-parametrize the generalized log-Moyal pdf (3) by substituting $\tau = \mu^{1/\sigma}$ so that the transformed pdf is

$$f^*(y) = \frac{\sqrt{\tau}}{\sqrt{2\pi}\sigma}\left(\frac{1}{y}\right)^{\frac{1}{2\sigma}+1} e^{-\frac{\tau}{2}\left(\frac{1}{y}\right)^{\frac{1}{\sigma}}}, \quad y > 0. \tag{9}$$

The following results show that both the gamma and the inverse Gaussian distributions are conjugate priors for τ.

Theorem 1. *If $Y_i, i = 1, \ldots, n$, are iid rvs having pdf (9) and τ follows a prior gamma distribution $\zeta(.)$ with a shape parameter α and a scale parameter λ, i.e., $\zeta(u) \propto u^{\alpha-1}e^{-\lambda u}$, then the posterior distribution of τ given the sample information $\{y_1, \ldots, y_n\}$ is again a gamma distribution with shape parameter $n/2 + \alpha$ and scale parameter $\lambda + \frac{1}{2}\sum_{i=1}^n\left(\frac{1}{y_i}\right)^{\frac{1}{\sigma}}$.*

Theorem 2. *If $Y_i, i = 1, \ldots, n$ are iid rvs having pdf (9) and τ follows a prior inverse Gaussian distribution $\zeta(.)$ with parameters ϕ*

and ψ, *i.e.,* $\zeta(u) \propto u^{-3/2} e^{\frac{-1}{2}\left(\frac{\phi}{\psi^2}u+\frac{\psi}{u}\right)}$, *then the posterior distribution of* τ *given the sample information* $\{y_1,\ldots,y_n\}$ *is the generalized inverse Gaussian distribution GIG*$(\lambda^*, \psi^*, \phi^*)$ *where* $\lambda^* = (n-1)/2$, $\psi^* = \psi\sqrt{\phi}\left(\phi + \psi^2 \sum_{i=1}^{n}\left(\frac{1}{y_i}\right)^{\frac{1}{\sigma}}\right)^{-\frac{1}{2}}$ *and* $\phi^* = \phi$.

Proof. Both the above theorems follow from Bayes' theorem. ■

These results are used in credibility theory to decide the future premium charges given the past claim history. In order to obtain a legitimate and feasible formula for this premium, various methodologies have been proposed in the actuarial literature. Most of them are in the field of Bayesian decision theory. We propose to discuss the above results in detail in a future work.

3. Statistical inference for parameters of the generalized log-Moyal distribution

Two methods of estimation of parameters are discussed, viz., (a) the method of quantiles, and (b) the maximum likelihood method. Since the maximum likelihood estimators of parameters are not obtainable in closed form, numerical approximations are used with initial values of parameters chosen using the method of quantiles. The sample quantiles are compared with the population quantiles to obtain the quantile estimators for the parameters μ and σ. Let ϱ_i denote the ϱ_ith quantile of the random sample, $i = 1, 2$, so that $0 < \varrho_1 < \varrho_2 < 1$, and the corresponding population quantiles from a population following the generalized log-Moyal distribution are given by

$$y_{\varrho_1} = \mu\left(\sqrt{2}\,\mathtt{erfc}^{\leftarrow}(\varrho_1)\right)^{-2\sigma}, \quad \text{and}$$
$$y_{\varrho_2} = \mu\left(\sqrt{2}\,\mathtt{erfc}^{\leftarrow}(\varrho_2)\right)^{-2\sigma}.$$

Replacing the population quantiles y_{ϱ_i} by the empirical quantiles $y_{n_{\varrho_i}}$ and solving for the parameters μ and σ, we get

$$\tilde{\sigma} = \frac{1}{2}\cdot\frac{\log\left(y_{n_{\varrho_1}}\right) - \log\left(y_{n_{\varrho_2}}\right)}{\log\left(\mathtt{erfc}^{\leftarrow}(\varrho_2)\right) - \log\left(\mathtt{erfc}^{\leftarrow}(\varrho_1)\right)},$$
$$\tilde{\mu} = y_{n_{\varrho_1}}\left(\sqrt{2}\,\mathtt{erfc}^{\leftarrow}(\varrho_1)\right)^{2\tilde{\sigma}}.$$

If $\boldsymbol{y} = \{y_1,\ldots,y_n\}$ denotes a random sample of n observations from the GlogM(μ,σ), the log-likelihood function $l(\Theta)$, $\Theta = (\mu,\sigma)$, based on this sample $\boldsymbol{y}$ is given by

$$l(\Theta|\boldsymbol{y}) = -n\log(\sqrt{2\pi}) - n\log\sigma + \frac{n}{2\sigma}\log\mu - \left(\frac{1}{2\sigma}+1\right)\sum_{i=1}^{n}\log y_i - \frac{\mu^{1/\sigma}}{2}\sum_{i=1}^{n}\left(\frac{1}{y_i}\right)^{1/\sigma},$$

and the normal equations for μ and σ are

$$\frac{\partial l(\Theta|\boldsymbol{y})}{\partial\mu} = \frac{n}{2\sigma\mu} - \frac{1}{2\sigma\mu}\sum_{i=1}^{n}\left(\frac{\mu}{y_i}\right)^{1/\sigma},$$
$$\frac{\partial l(\Theta|\boldsymbol{y})}{\partial\sigma} = -\frac{n}{\sigma} + \frac{1}{2\sigma^2}\sum_{i=1}^{n}\log y_i - \frac{n}{2\sigma^2}\log\mu + \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\frac{\mu}{y_i}\right)^{1/\sigma}\log\left(\frac{\mu}{y_i}\right).$$

Equating these normal equations to zero, the maximum likelihood estimator $\hat{\mu}$ of μ is

$$\hat{\mu} = \left(\frac{1}{n}\sum_{i=1}^{n} y_i^{-\frac{1}{\hat{\sigma}}}\right)^{-\hat{\sigma}},$$

and that of σ, $\hat{\sigma}$ is obtained by solving the non-linear equation

$$2\hat{\sigma} - \frac{1}{n}\sum_{i=1}^{n}\log y_i + \frac{\sum_{i=1}^{n} y_i^{-1/\hat{\sigma}}\log y_i}{\sum_{i=1}^{n} y_i^{-1/\hat{\sigma}}} = 0.$$

The solution of the above equation is not in closed form and hence to determine the solution we use Newton–Raphson method with initial value of σ obtained from the method of percentiles. Further, to obtain confidence intervals for the parameters, we require the Fisher information matrix $\mathbf{I}(\Theta)$ with entries $\mathbf{I}_{ij}(\Theta) = \mathbb{E}\left(-\frac{\partial^2 l(\Theta)}{\partial\theta_i\partial\theta_j}\right)$, $i,j = 1, 2$, and $\Theta = (\mu,\sigma)$. The second order derivatives of the log-likelihood function are given by

$$\frac{\partial^2 l(\Theta|\boldsymbol{y})}{\partial\mu^2} = -\frac{n}{2\sigma\mu^2} - \left(\frac{1}{\sigma}-1\right)\frac{1}{2\sigma\mu^2}\sum_{i=1}^{n}\left(\frac{\mu}{y_i}\right)^{\frac{1}{\sigma}},$$
$$\frac{\partial^2 l(\Theta|\boldsymbol{y})}{\partial\mu\partial\sigma} = -\frac{n}{2\sigma^2\mu} + \frac{1}{2\sigma^3\mu}\sum_{i=1}^{n}\left(\frac{\mu}{y_i}\right)^{\frac{1}{\sigma}}\log\left(\frac{\mu}{y_i}\right) + \frac{1}{2\sigma^2\mu}\sum_{i=1}^{n}\left(\frac{\mu}{y_i}\right)^{\frac{1}{\sigma}},$$
$$\frac{\partial^2 l(\Theta|\boldsymbol{y})}{\partial\sigma^2} = \frac{n}{\sigma^2} + \frac{1}{\sigma^3}\sum_{i=1}^{n}\log\left(\frac{\mu}{y_i}\right) - \frac{1}{\sigma^3}\sum_{i=1}^{n}\left(\frac{\mu}{y_i}\right)^{\frac{1}{\sigma}}\log\left(\frac{\mu}{y_i}\right) - \frac{1}{2\sigma^4}\sum_{i=1}^{n}\left(\frac{\mu}{y_i}\right)^{\frac{1}{\sigma}}\log^2\left(\frac{\mu}{y_i}\right).$$

Hence, using Proposition 1, explicit expressions for the entries of the Fisher information matrix are

$$I_{11} = \mathbb{E}\left(-\frac{\partial^2 l(\Theta)}{\partial\mu^2}\right) = \frac{n}{2\sigma^2\mu^2},$$
$$I_{12} = I_{21} = \mathbb{E}\left(-\frac{\partial^2 l(\Theta)}{\partial\mu\partial\sigma}\right) = \frac{n}{2\sigma^2\mu}\left(\gamma + \log 2 - 2\right),$$
$$I_{22} = \mathbb{E}\left(-\frac{\partial^2 l(\Theta)}{\partial\sigma^2}\right) = \frac{n}{\sigma^2} + \frac{n}{4\sigma^2}\left(\pi^2 + (\log 2 - 4)\log 4 + 2\gamma(\gamma - 4 + \log 4)\right).$$

The Fisher information matrix is then given by

$$\mathbf{I}(\Theta) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}.$$

Under the usual regularity conditions, we have $\sqrt{n}\left(\hat{\Theta} - \Theta\right) \to N_2\left(\mathbf{0}, \mathbf{I}^{-1}(\Theta)\right)$, where $\mathbf{I}^{-1}(\Theta)$ is the inverse of the Fisher information matrix. The square root of the diagonal elements of $\mathbf{I}^{-1}$ gives the standard error of the estimates of the parameters μ and σ, respectively. The $100(1-\alpha)\%$ confidence intervals for μ and σ are, respectively, given by

$$\hat{\mu} \mp Z_{\alpha/2}\sqrt{\mathbf{I}_{11}^{-1}(\hat{\Theta})} \qquad \text{and} \qquad \hat{\sigma} \mp Z_{\alpha/2}\sqrt{\mathbf{I}_{22}^{-1}(\hat{\Theta})},$$

where $Z_{\alpha/2}$ is the upper $\alpha/2$th percentile of the standard normal distribution.

4. Data analysis using the generalized log-Moyal distribution

The applicability and superiority of the proposed GlogM(μ,σ) distribution and its comparison with other families of heavy tailed distributions in data analysis are discussed in this section using two celebrated actuarial data sets, viz., (a) the Norwegian fire losses data set `Norwegian` and (b) the Danish fire losses data set `Danish`.

Table 1
Empirical characteristics of `Norwegian` fire loss data set from 1990 to 1992.

Year	n	Min.	Q_1	Median	Mean	Q_3	Max.	SD	Skewness	Kurtosis
1990	628	500	788.80	1150	1973.50	1805.80	78 537	4257.28	11.73	182.52
1991	624	500	753.00	1098	1820.00	1909.00	49 692	3019.42	9.59	127.24
1992	615	500	734.50	1060	2184.20	1878.50	102 438	5529.40	12.24	192.85

Table 2
Empirical data analysis of the `Norwegian` fire loss data set.

Year		Model						
		log gamma	Lomax	Fréchet	log logistic	para logistic	inverse para-logistic	GlogM
	Parameters	(β, λ)	(α, λ)	(α, β)	(γ, σ)	(β, σ)	(γ, σ)	(μ, σ)
1990	Estimates	(0.066, 108.955)	(5.616, 8592.01)	(1.956, 0.001)	(2.675, 1205.19)	(1.903, 1907.44)	(2.346, 0.001)	(863.199, 0.319)
	LL	−5157.500	−5344.870	−5089.040	−5149.440	−5200.830	−5114.430	−5081.650
	AIC	10 319.000	10 693.740	10 182.080	10 302.880	10 405.660	10 232.860	10 167.300
	BIC	10 333.342	10 708.082	10 196.422	10 317.222	10 420.002	10 247.202	10 181.642
	KS	0.756	0.534	0.749	0.480	0.508	0.507	0.875
	CvM	0.564	0.537	0.623	0.630	0.557	0.714	0.690
	AD	0.574	0.533	0.589	0.585	0.548	0.657	0.747
1991	Estimates	(0.065, 109.467)	(7.580, 11684)	(1.952, 0.001)	(2.624, 1176.41)	(1.897, 1874.48)	(2.326, 0.001)	(840.204, 0.318)
	LL	−5105.620	−5286.690	−5044.220	−5108.150	−5153.360	−5072.300	−5034.760
	AIC	10 215.240	10 577.380	10 092.440	10 220.300	10 310.720	10 148.600	10 073.520
	BIC	10 229.582	10 591.722	10 106.782	10 234.642	10 325.062	10 162.942	10 087.862
	KS	0.242	0.531	0.814	0.485	0.475	0.57	0.845
	CvM	0.069	0.555	0.6805	0.604	0.531	0.628	0.716
	AD	0.159	0.558	0.712	0.607	0.534	0.621	0.753
1992	Estimates	(0.079, 90.611)	(4.255, 6499.53)	(1.839, 0.0010)	(2.416, 1183.66)	(1.755, 1851.43)	(2.197, 0.001)	(837.877, 0.334)
	LL	−5108.970	−5268.320	−5033.290	−5107.440	−5154.280	−5068.690	−5017.820
	AIC	10 221.940	10 540.640	10 070.580	10 218.880	10 312.560	10 141.380	10 039.640
	BIC	10 236.282	10 554.982	10 084.922	10 233.222	10 326.902	10 155.722	10 053.982
	KS	0.682	0.527	0.710	0.474	0.460	0.635	0.946
	CvM	0.525	0.542	0.614	0.578	0.531	0.611	0.743
	AD	0.568	0.554	0.645	0.569	0.533	0.594	0.766

4.1. Analysis of the Norwegian fire losses data set

We begin with our first illustration by considering the Norwegian fire losses data set available at http://lstat.kuleuven.be/Wiley (in Chapter 1, file `norwegianfire.txt`). This data set describes year-wise Norwegian fire insurance portfolio from 1989 to 1992 and for the present study we consider data from year 1990 to 1992. The empirical characteristics of these three year fire losses data is shown in Table 1.

We employ the mean excess plot, a graphical diagnostic technique, to identify the possible class of distributions appropriate to model the `Norwegian` data set. Fig. 4 gives the mean excess plot which plots $(X_{(i)}, \hat{e}_n(X_{(i)}))$, $1 \le i \le n$, where $X_{(1)} \ge X_{(2)} \ge \cdots \ge X_{(n)}$ are in decreasing order, and sample mean excess function $\hat{e}_n(u) = \frac{\sum_{i=1}^{n}(X_i-u)I_{(X_i>u)}}{\sum_{i=1}^{n} I_{(X_i>u)}}$. The linearly increasing trend of $\hat{e}_n(u)$ for all the three years gives evidence to conclude that distributions belonging to the heavy tailed class are candidate fits for the `Norwegian` data set (see McNeil et al., 2005, Embrechts et al., 2003, Hogg and Klugman, 2009).

We now fit the GlogM(μ, σ) distribution and compare the fit with six other competitive heavy tailed distributions, namely, log gamma, Lomax, Fréchet, log-logistic, para-logistic and inverse para-logistic. We use the method of maximum likelihood to estimate the parameters in all these seven distributions. The parameter estimates, log-likelihood values, values of the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are computed and tabulated in Table 2.

We further compute the bootstrap p-values for the goodness-of-fit tests, namely, Kolmogorov–Smirnov (KS), Cramer–von Mises (CvM) and Anderson–Darling (AD) tests. It is clear from Table 2 that the GlogM(μ, σ) distribution is a better fit as it has the maximum log-likelihood value and the minimum AIC and BIC values. Moreover, the bootstrap p-values for KS, CvM and AD tests for the GlogM(μ, σ) distribution are the highest as compared to the other distributions. For these seven models, observed-versus-fitted quantile plots on the log–log scale, i.e. $\left(\log x_{(n-i+1)}, \log \hat{F}^{\leftarrow}\left(\frac{i}{n+1}\right)\right)$, $1 \le i \le n$, are presented in Fig. 5. Further, the plots in Fig. 5 also indicate that the GlogM(μ, σ) distribution gives a better fit than the other models in the sense that the points corresponding to the theoretical and empirical quantiles do not deviate much from the 45°straight line.

4.2. Analysis of the Danish fire insurance losses data set

The Danish fire insurance losses data set consists of 2492 losses in millions of Danish Kroner during the period from 1980 to 1990, arising from fire claims in Copenhagen. This data set is available in "`SMPractica1s`" add-on package of `R`, in the CRAN website http://cran.r-project.org/. We compare the GlogM(μ, σ) distribution with [nine] other competing two-parameter heavy tailed distributions including the generalized inverse Lindley distribution, recently introduced by Asgharzadeh et al. (2017). The log-likelihood values (LL) as well as the AIC and the BIC values along with the bootstrap p-values for the goodness-of-fit tests namely KS, CvM and AD are given in Table 3. Maximum value for LL and minimum values for AIC and BIC give evidence that the GlogM(μ, σ) distribution gives a better fit as compared to the other two-parameter models. Moreover, this is also ascertained by maximum p-values. Using estimates of the parameters, we have computed the empirical and estimated VaR for the GlogM(μ, σ) distribution which are presented in Table 4. In Table 5, we have presented the $\mathtt{LEV_u}$ values

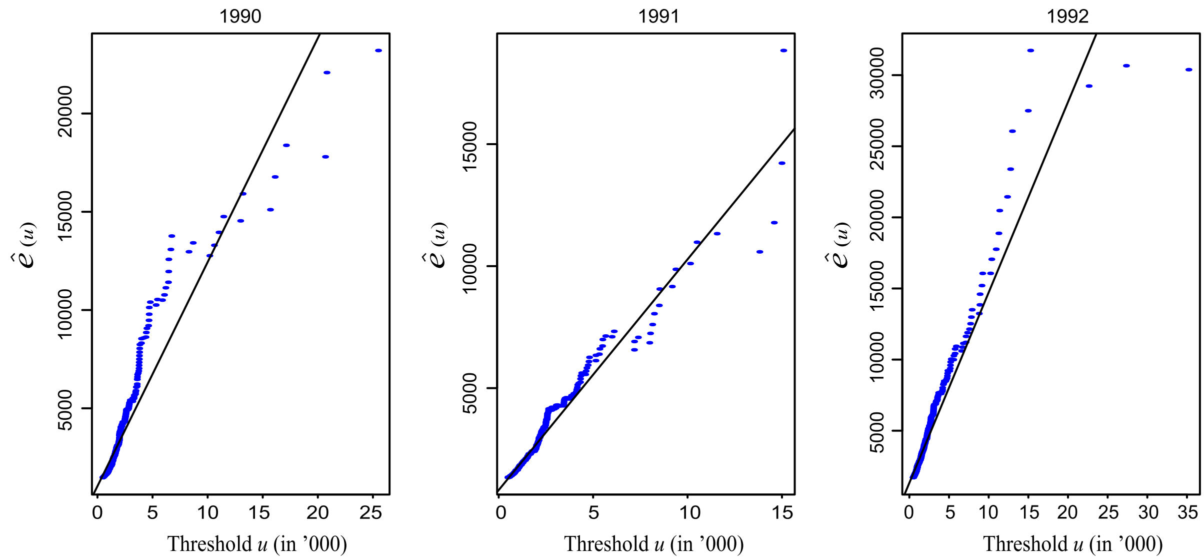

Fig. 4. Mean excess plot of losses in the years 1900, 1991 and 1992.

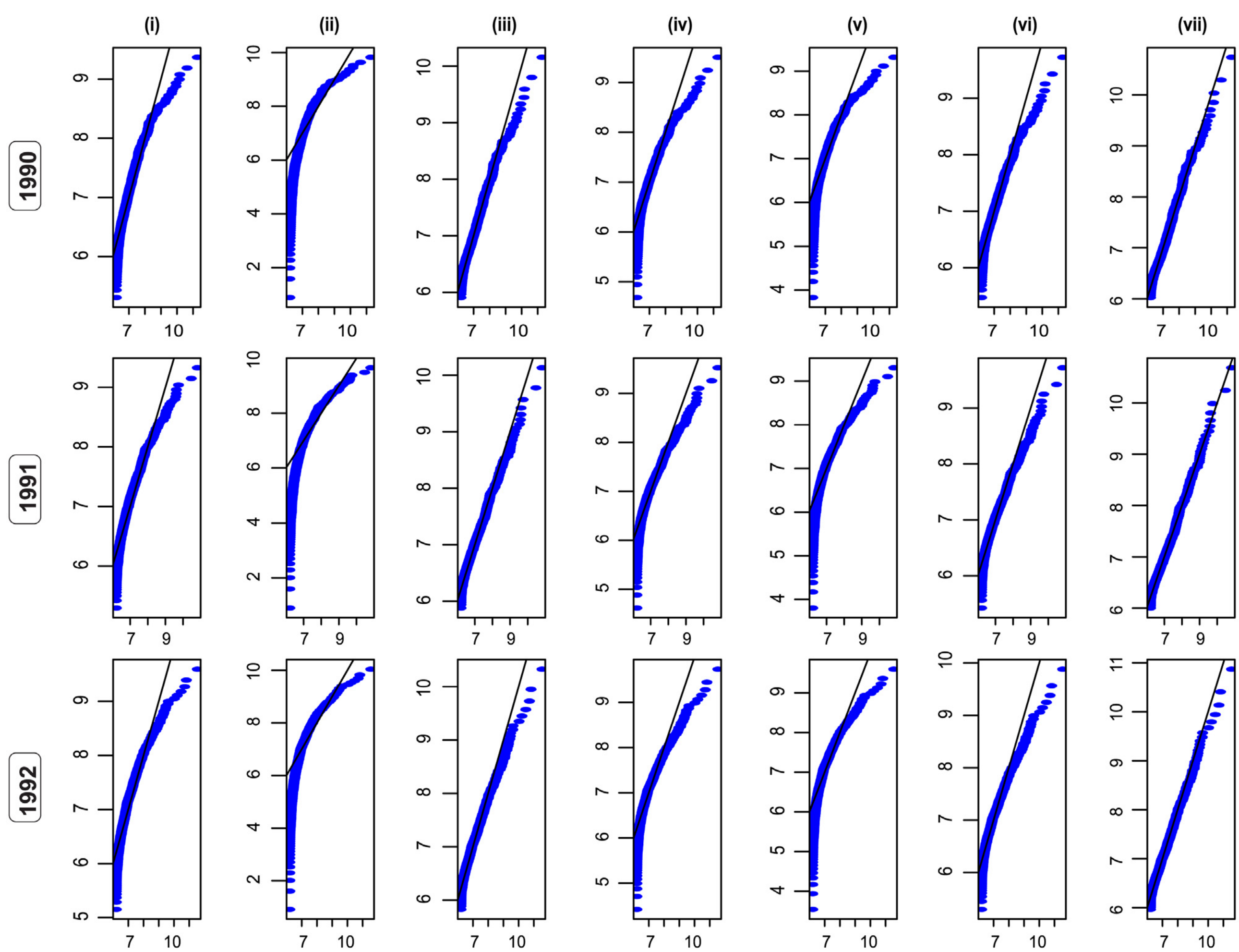

Fig. 5. Observed-versus-fitted quantile plots on log − log scale for 1990–1992 `Norwegian` data set for (i) log gamma distribution, (ii) Lomax, (iii) Fréchet, (iv) log-logistic, (v) Para-logistic, (vi) Inverse para-logistic and (vii) GlogM.

for different levels u for the ten models and compare these with the empirical LEV values. The difference in values of the empirical LEV and the LEV obtained from the various models are presented in Fig. 6 and it is clear that the difference is the least for the GlogM(μ, σ) distribution in almost all the cases. The observed-versus-fitted QQ plots for all the ten models are given in Fig. 7.

Table 3
Data analysis for the Danish fire insurance losses data set.

Parameter	Models									
	log gamma	Lomax	inverse Gaussian	Fréchet	log-logistic	para-logistic	inverse para-logistic	Gen. inverse Lindley	inverse gamma	GlogM
$\hat{\alpha}$	–	5.169 (0.306)	–	2.01 (0.032)	–	–		1.944 (0.032)	2.753 (0.074)	
$\hat{\beta}$	0.19 (0.005)			0.695 (0.007)		1.846 (0.022)			4.447 (0.131)	
$\hat{\lambda}$	5.963 (0.167)	11.9 (0.802)	–		–		–	2.623 (0.045)		
$\hat{\mu}$			3.063 (0.058)							1.312 (0.013)
$\hat{\gamma}$			–	–	2.653 (0.045)	–	2.413 (0.032)			
$\hat{\sigma}$			0.541 (0.097)		1.77 (0.022)	2.807 (0.046)	0.908 (0.01)		–	0.321 (0.005)
l_{max}	−4308.56	−5051.91	−4516.31	−3966.83	−4280.58	−4514.88	−4093.32	−3954.30	−4097.90	−3932.99
AIC	8621.12	10 107.82	9036.62	7937.66	8565.17	9033.76	8190.64	7912.61	8199.80	7869.99
BIC	8632.764	10 110.613	9039.41	7940.45	8576.81	9045.41	8202.28	7924.30	8211.41	7872.78
KS	0.677	0.501	0.468	0.849	0.553	0.522	0.549	0.914	0.644	0.726
CvM	0.522	0.495	0.497	0.554	0.527	0.518	0.515	0.569	0.523	0.593
AD	0.520	0.494	0.312	0.562	0.536	0.519	0.513	0.594	0.532	0.649

Table 4
Empirical and estimated VaR for Danish fire insurance losses data set.

q	0.001	0.002	0.005	0.01	0.9	0.95	0.99	0.995	0.998	0.999
Empirical VaR	0.381	0.602	0.825	0.846	5.08	8.552	25.288	34.142	57.411	152.413
Estimate VaR	0.611	0.636	0.676	0.715	4.968	7.764	21.826	34.060	61.336	95.714

Table 5
Limited expected value LEV.

u	Empirical LEV	GlogM	gen. inverse Lindley	inverse gamma	Fréchet	Lomax	log gamma	log-logistic	para logistic	inverse Gaussian	inverse para-logistic
1	0.935	0.983	0.981	0.971	0.981	0.815	0.955	0.946	0.914	0.951	0.971
2	1.551	1.585	1.593	1.606	1.596	1.361	1.601	1.559	1.521	1.621	1.588
3	1.839	1.877	1.879	1.934	1.884	1.736	1.956	1.852	1.862	2.053	1.864
5	2.155	2.166	2.135	2.223	2.138	2.193	2.270	2.075	2.155	2.536	2.079
8	2.387	2.374	2.289	2.378	2.287	2.520	2.420	2.175	2.286	2.844	2.183
10	2.483	2.454	2.342	2.424	2.338	2.630	2.457	2.202	2.318	2.935	2.214
15	2.653	2.576	2.415	2.478	2.405	2.759	2.493	2.232	2.348	3.027	2.250
21	2.762	2.658	2.457	2.503	2.443	2.813	2.504	2.245	2.359	3.054	2.267
40	2.919	2.780	2.509	2.525	2.488	2.848	2.512	2.257	2.366	3.062	2.285
70	2.997	2.854	2.534	2.532	2.510	2.853	2.513	2.261	2.367	3.062	2.291
110	3.045	2.900	2.547	2.535	2.520	2.854	2.513	2.262	2.368	3.062	2.294
170	3.093	2.934	2.555	2.536	2.526	2.854	2.513	2.262	2.368	3.062	2.295
270	3.113	2.963	2.560	2.536	2.530	2.854	2.513	2.263	2.368	3.062	2.296

5. A regression application

In linear regression, the dependence of the response variable on the explanatory variable(s) is modelled via the conditional mean of the response variable. However, in situations where the response variable is assumed to follow a heavy tailed distribution, the mean of the distribution may not always exist. In such cases, the scale and/or shape parameters are treated as functions of the explanatory variable(s) in the distribution of the response variable. Beirlant et al. (1998) discussed the regression modelling assuming that the response variable follows the Burr distribution which was later generalized to distributions with Pareto-type tail in Beirlant and Goegebeur (2003). Recently, Gündüz et al. (2016) considered exponentiated Fréchet regression model for the response variable. Here we assume that the response variable follows the generalized log-Moyal distribution with transformed density (9) and assume that the shape parameter σ is a function of the explanatory variables. We consider an exponential link function which ensures the non-negativity of the shape parameter. We have

$$Y_i|\boldsymbol{x_i} \sim \mathrm{G\,log\text{-}M}(\tau, \sigma_i), \qquad \sigma_i = e^{\boldsymbol{\theta}^\top \boldsymbol{x_i}}, i = 1, \ldots, n.$$

where $\boldsymbol{x_i}^\top = (1, x_{i1}, \ldots, x_{is})$ and $\boldsymbol{\theta}^\top = (\theta_0, \theta_1, \ldots, \theta_s)$. The log-likelihood function is

$$l(\tau, \boldsymbol{\theta}|y, \boldsymbol{x}) = -\frac{n}{2}\log(2\pi) + \frac{n}{2}\log(\tau) - \sum_{i=1}^{n} \boldsymbol{\theta}^\top \boldsymbol{x_i} - \sum_{i=1}^{n}\left(\frac{e^{-\boldsymbol{\theta}^\top \boldsymbol{x_i}}}{2} + 1\right)\log y_i - \frac{\tau}{2}\sum_{i=1}^{n}\left(\frac{1}{y_i}\right)^{e^{-\boldsymbol{\theta}^\top \boldsymbol{x_i}}}. \tag{10}$$

The score function is obtained by differentiating the log-likelihood function (10) with respect to the unknown parameters $(\tau, \boldsymbol{\theta})$, and is given by $U^\top(\tau, \boldsymbol{\theta}) = \left(\frac{\partial l}{\partial \tau}, \frac{\partial l}{\partial \theta_k}\right)$, where

$$\frac{\partial l}{\partial \tau} = \frac{n}{2\tau} - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{1}{y_i}\right)^{e^{-\boldsymbol{\theta}^\top \boldsymbol{x_i}}},$$

$$\frac{\partial l}{\partial \theta_k} = -\sum_{i=1}^{n} x_{ik} - \frac{1}{2}\sum_{i=1}^{n} x_{ik} e^{-\boldsymbol{\theta}^\top \boldsymbol{x_i}} \log y_i + \frac{\tau}{2}\sum_{i=1}^{n} x_{ik} e^{-\boldsymbol{\theta}^\top \boldsymbol{x_i}} \log y_i \left(\frac{1}{y_i}\right)^{e^{-\boldsymbol{\theta}^\top \boldsymbol{x_i}}},$$

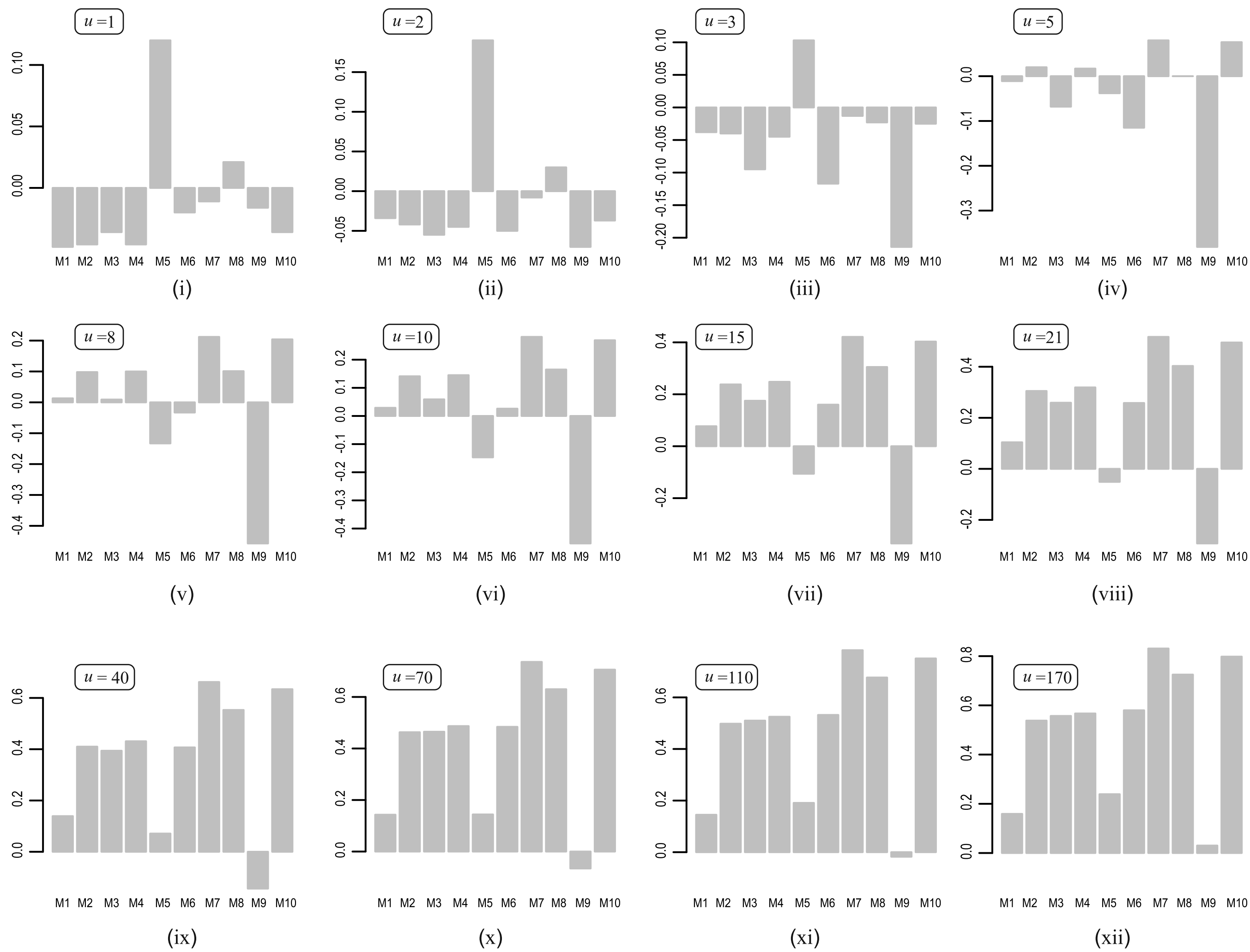

Fig. 6. Vertical bar in each figure represents the difference of empirical LEV for different u of Danish fire insurance losses data set with the fitted model (M_i), $i = 1, \ldots, 10$. Here M_1 = GlogM, M_2 = generalized inverse Lindley, M_3 = Inverse gamma, M_4 = Fréchet, M_5 = Lomax, M_6 = log gamma, M_7 = log-logistic, M_8 = para-logistic, M_9 = inverse Gaussian and M_{10} = Inverse para-logistic.

with $k = 0, 1, 2, \ldots, s$. Solving $U^\top(\tau, \boldsymbol{\theta}) = \mathbf{0}$ simultaneously, we obtain the estimates of the $s + 2$ parameters.

5.1. Residuals

For assessment of the regression model, we determine the residuals by using the transformation $R_i = Y_i^{\exp\left(-\boldsymbol{\theta}^\top \boldsymbol{x_i}\right)}$, $i = 1, \ldots, n$. This makes the residuals iid with pdf given by

$$f_{R_i}(r) = \frac{\sqrt{\tau}}{\sqrt{2\pi}\, r^{3/2}} e^{-\frac{\tau}{2r}}, \quad r > 0, \tau > 0. \tag{11}$$

Note that the pdf (11) does not depend on $\boldsymbol{x_i}$ and therefore R_i's have been called as the residuals. The quantile function of the distribution of the residuals is

$$Q(p) = \frac{\tau}{2}(\mathtt{erfc}^{\leftarrow}(p))^{-2}, \quad 0 < p < 1,$$

and the ordered pairs

$$\left(-\log(\mathtt{erfc}^{\leftarrow}(i/(n+1))), \log r_{(i)}\right),$$

with $r_{(1)} < r_{(2)} < \cdots < r_{(n)}$, denote the ordered observed residuals. We use the observed r_i's to construct a quantile plot for the distribution of the residuals. If the underlying distribution of the residuals follows the pdf (11), then the quantile plot will be a straight line with slope 2 and intercept $\log(\tau/2)$.

5.2. Application of the generalized log-Moyal regression to vehicle insurance losses data set

To illustrate the generalized log-Moyal regression model, we consider the vehicle insurance losses data set based on one-year vehicle insurance policies taken out in 2004 and 2005, available at http://www.businessandeconomics.mq.edu.au/our_departments/Applied_Finance_and_Actuarial_Studies/research/books/GLMsforInsuranceData. The original data set comprises of 67,856 policies and we consider 4624 policies having at least one claim. The details of each covariate considered for this study are presented in Table 6. The histogram of overall claim size along with the descriptive statistics are presented in Fig. 8.

We now fit the generalized log-Moyal regression model discussed above to this data set and estimate the parameters. We use the `optim()` function in R which uses the BFGS (Broyden–Fletcher–Goldfarb–Shanno) method. We compare our model with the Fréchet regression model, recently discussed in Gündüz et al. (2016). The log-likelihood value obtained for the generalized log-Moyal and the Fréchet distributions without covariates are −6626.23 and −6654.148, respectively, which gives an evidence that the generalized log-Moyal distribution shows better fit than the Fréchet distribution. We further assume that the shape parameter is a function of the covariates as discussed in Sections 5 and 5.1. To begin the numerical estimation procedure, we choose

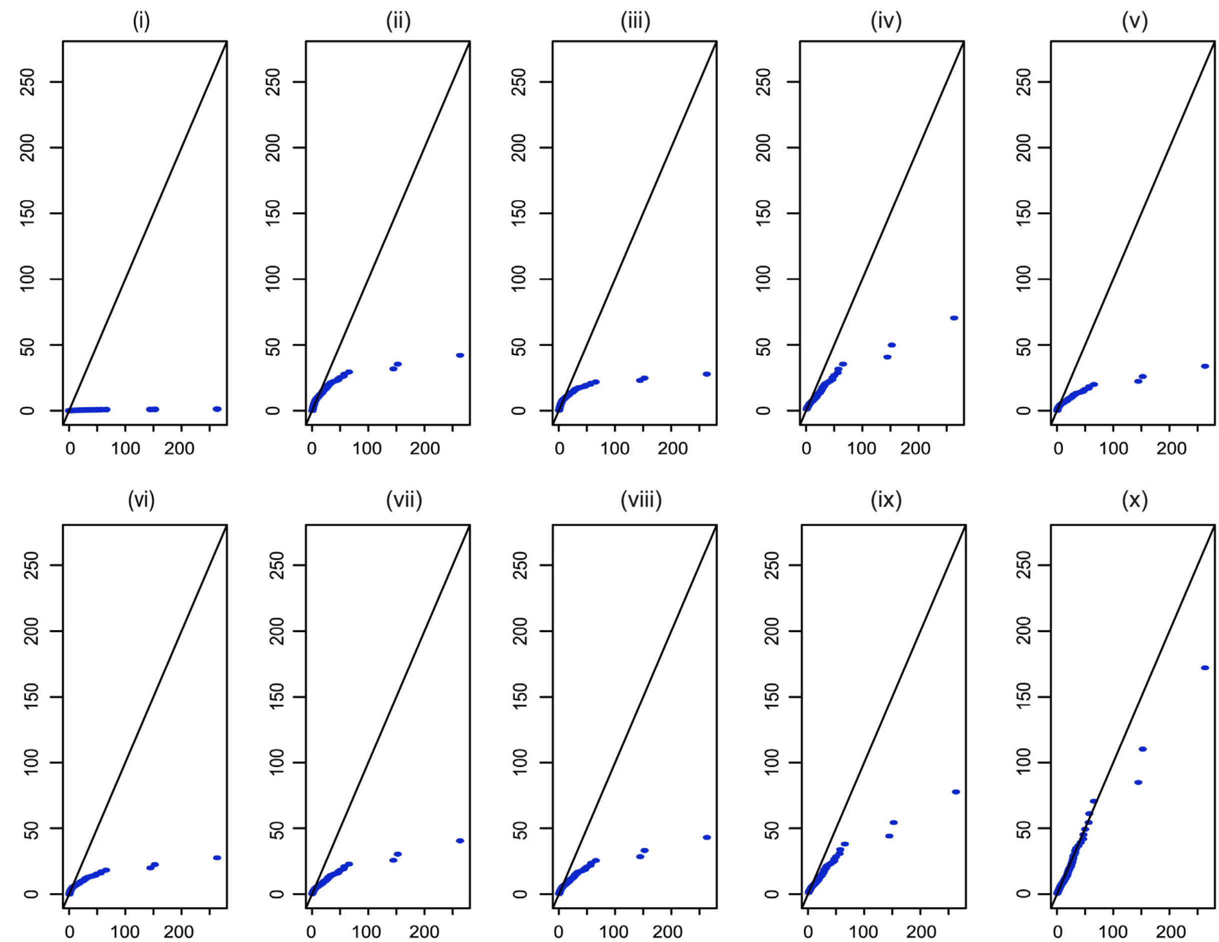

Fig. 7. QQ-plots of Danish fire losses data set for (i) log gamma, (ii) Lomax, (iii) inverse Gaussian, (iv) Fréchet, (v) log-logistic, (vi) para-logistic, (vii) inverse para-logistic, (viii) inverse Gamma, (ix) Generalized inverse Lindley and (x) GlogM.

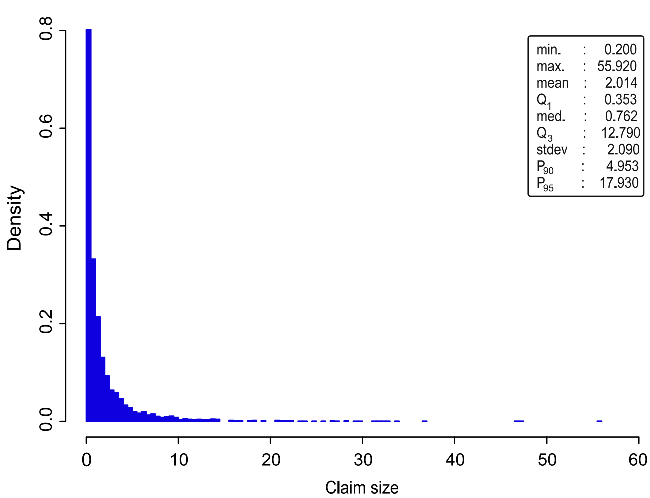

Fig. 8. Histogram of overall claim size.

the initial value for the scale parameter as 1 and parameters θ_k, $k = 0, 1, \ldots, 4$ as 0 which means that the shape parameter is not a function of the covariates. Similar selections are also made for the parameters of the Fréchet regression model. The estimates and their standard errors (SE) for both the regression models are presented in Table 7. Further, we test $\mathscr{H}_0 : \theta_k = 0$ against $\mathscr{H}_1 : \theta_k \neq 0$. We compute the t-ratio $= \hat{\theta}_k/\mathrm{SE}(\hat{\theta}_k)$, for testing the hypothesis and reject the null hypothesis if the absolute value of the t-ratio exceeds $t_{n-5,1-\alpha/2}$, where α is the level of significance. Also, the 95% confidence limits for the parameters are constructed by using the formula $\hat{\theta}_k \mp t_{n-5,1-\alpha/2} \cdot \mathrm{SE}(\hat{\theta}_k)$. All the covariates except the covariate VEHAGE are found to be significant in the generalized log-Moyal regression model which is not the case in the Fréchet regression model. Also, the values of the AIC and the BIC are lower for the generalized log-Moyal regression model, compared to those for the Fréchet regression model. Hence the generalized log-Moyal regression model is found more appropriate as compared to the Fréchet regression model for the vehicle insurance losses data set. To have a visual assessment of the fit of both the models, we give quantile plots of the residuals in Fig. 9 for both the models.

6. Conclusions

In the present work, we have proposed a versatile two parameter heavy tailed distribution allowing closed-form expressions for a lot of insurance measures and have studied some of its properties. The applicability of this new family of distributions has been illustrated using two well known data sets from insurance and the model performs reasonably well as compared to other popular heavy tailed distributions. Assuming that the response variable follows the proposed model, non-linear regression has been discussed along with an illustration comparing the proposed regression model with a Fréchet regression model.

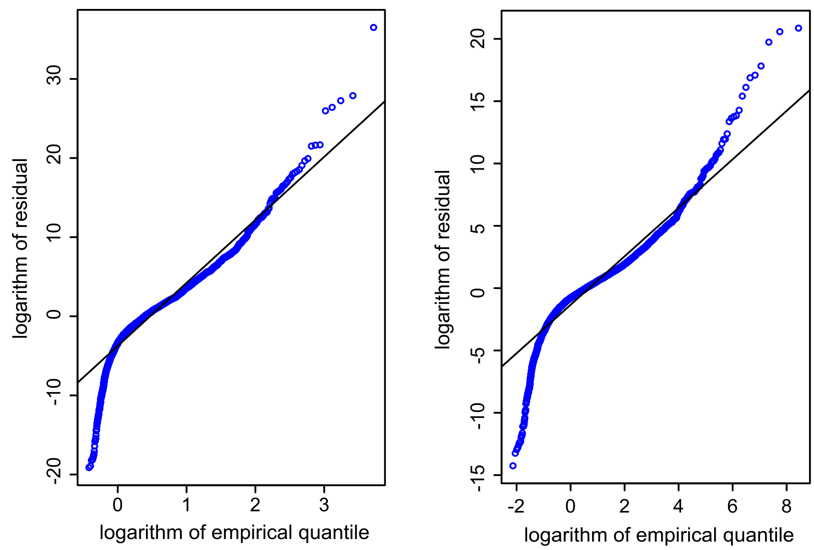

Fig. 9. Generalized log-Moyal quantile plot of the residuals (left) and Fréchet quantile plot of the residuals (right).

Table 6
Details of response and explanatory variables of vehicle insurance data set.

Response variable	
CLMSIZE	Total claim (in '000)
Explanatory variables	
EXPSR	Exposure over the range (0–1).
GENDR	Gender of the policy holder (0 for Male, 1 for female).
VEHAGE	Age of vehicle: 1 (youngest), 2, 3, 4.
AGECAT	Driver's age category: 1 (youngest), 2, 3, 4, 5, 6.

Table 7
Results for the generalized log-Moyal and the Fréchet regression.

Explanatory variables	Estimate	SE	95% Confidence limits	t-ratio	p-value
	Generalized log-Moyal regression				
Intercept($\hat{\theta}_0$)	0.7763	0.1781	(0.427, 1.125)	4.359	0.000
EXPSR($\hat{\theta}_1$)	−0.8621	0.0825	(−1.024, −0.700)	−10.452	0.000
VEHAGE($\hat{\theta}_2$)	−0.0596	0.0410	(−0.139, 0.021)	−1.454	0.146
GENDR($\hat{\theta}_3$)	0.6050	0.0194	(0.567, 0.643)	31.168	0.0000
AGECAT($\hat{\theta}_4$)	−0.3785	0.0099	(−0.398, −0.359)	−38.304	0.0000
$\hat{\tau}$	0.0514	0.0044	(0.042, 0.060)	–	–
log-likelihood	**−716.417**				
AIC	**1444.833**				
BIC	**1483.467**				
	Fréchet regression				
Intercept($\hat{\theta}_0$)	−2.190	0.2045	(−2.592, −1.789)	−10.712	0.000
EXPSR($\hat{\theta}_1$)	1.3945	0.0945	(1.209, 1.579)	14.758	0.000
VEHAGE($\hat{\theta}_2$)	0.2451	0.0469	(0.153, 0.337)	5.2274	0.000
GENDR($\hat{\theta}_3$)	−0.4735	0.0226	(−0.517, −0.429)	−20.962	0.000
AGECAT($\hat{\theta}_4$)	0.3807	0.0119	(0.357, 0.404)	32.091	0.000
$\hat{\tau}$	0.2134	0.0096	(0.194, 0.232)	–	–
log-likelihood	−735.404				
AIC	1482.807				
BIC	1521.441				

Acknowledgements

The authors acknowledge suggestions of two anonymous referees which helped in improving the presentation of the paper.

Thanks first authors's employer Central University of Rajasthan and the University Grants Commission (UGC) for awarding fellowship under UGC Faculty Development Programme and the University of Mysore for support to carry out this work. The second author's research work was supported by UGC Major Research Project Stat-2013-19168 and the University of Mysore.

References

Adcock, C.J., Eling, M., Loperfido, N., 2015. Skewed distributions in finance and actuarial science: a review. Eur. J. Finance 21 (13–14), 1253–1281.
Artzner, P., 1999. Application of coherent risk measures to capital requirements in insurance. N. Am. Actuar. J. 3 (2), 11–25.
Asgharzadeh, A., Nadarajah, S., Sharafi, F., 2017. Generalized inverse Lindley distribution with application to Danish fire insurance data. Comm. Statist. Theory Methods 46 (10), 5000–5021.
Azzalini, A., Dal Cappello, T., Kotz, S., 2003. Log-skew-normal and log-skew-t distributions as models for family income data. J. Income Distrib. 11, 12–20.

Bagnato, L., Punzo, A., 2013. Finite mixtures of uni-modal beta and gamma densities and the k−bumps algorithm. Comput. Statist. 28 (4), 1571–1597.
Bakar, S.A.A., Hamzah, N.A., Maghsoudi, M., Nadarajah, S., 2015. Modelling loss data using composite models. Insurance Math. Econom. 61, 1146–1154.
Beirlant, J., Goegebeur, Y., 2003. Regression with response distributions of Pareto-type. Comput. Statist. Data Anal. 42 (4), 595–619.
Beirlant, J., Goegebeur, Y., Verlaak, R., Vynckier, P., 1998. Burr regression and portfolio segmentation. Insurance Math. Econom. 23 (3), 231–250.
Bernardi, M., Maruotti, A., Petrella, M., 2012. Skew mixture models for loss distributions: A Bayesian approach. Insurance Math. Econom. 51 (3), 617–623.
Brazauskas, V., Kleefeld, A., 2011. Folded and log-folded-t distributions as models for insurance loss data. Scand. Actuar. J. (1), 59–74.
Calderín-Ojeda, E., Kwok, C.F., 2016. Modelling claims data with composite Stoppa models. Scand. Actuar. J. 9, 817–836.
Ciumara, R., 2006. An actuarial model based on the composite Weibull–Pareto distribution. Math. Rep.–Buchar. 8 (4), 401–414.
Coorey, K., Ananda, M.M., 2005. Modelling actuarial data with a composite lognormal-Pareto model. Scand. Actuar. J. 5, 321–334.
Coorey, K., Ananda, M.M., 2008. A generalization of the half-Normal distribution with applications to lifetime data. Comm. Statist. Theory Methods 37 (9), 1323–1337.
Cordeiro, G.M., Nobre, J.S., Pescim, R.R., Ortega, E.M.M., 2012. The Beta-Moyal: a useful-skew distribution. Int. J. Res. Rev. Appl. Sci. 10 (2).
Eling, M., 2012. Fitting insurance claims to skewed distributions: Are the skew-normal and skew-student good models? Insurance Math. Econom. 51, 239–248.
Embrechts, P., Klüppelberg, C., Mikosch, T., 2003. Modelling Extremal Events. Springer.
Ferreira, A., de Haan, L., 2006. Extreme Value Theory: An Introduction. Springer.
Glaser, R.E., 1980. Bathtub and related failure rate characterizations. J. Amer. Statist. Assoc. 75 (371), 667–672.
Gómez-Déniz, E., Calderín-Ojeda, E., 2015. Modelling insurance data with the Pareto arctan distribution. Astin Bull. 45 (3), 639–660.
Gündüz, F.F., Ali, I., Genç, A.I., 2016. The exponentiated Fréchet regression: an alternative model for actuarial modelling purposes. J. Stat. Comput. Simul. 86 (17), 3456–3481.
Hogg, R.V., Klugman, S.A., 2009. Loss distributions. In: Wiley Series in Probability and Statistics, vol. 249, Wiley, New York.
Kazemi, R., Noorizadeh, M., 2015. A comparison between skew-logistic and skew-normal distributions. Matematika 31 (1), 15–24.
Klugman, S.A., Panjer, H.H., Willmot, G.E., 2012. Loss Models: From Data to Decisions, fourth ed. In: Wiley Series in Probability and Statistics, John Wiley and Sons, Inc., Hoboken NJ.
Landsman, Z., Makov, U., Shushi, T., 2016. Tail conditional moments for elliptical and log-elliptical distributions. Insurance Math. Econom. 71, 179–188.
Lee, S.C.K., Lin, X.S., 2010. Modeling and evaluating insurance losses via mixtures of Erlang distributions. N. Am. Actuar. J. 14 (1), 107–130.
McNeil, A.J., 1997. Estimating the tails of loss severity distributions using extreme value theory. Astin Bull. 27 (1), 117–137.
McNeil, A.J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, Princeton NJ.
Miljkovic, T., Grün, B., 2016. Modeling loss data using mixtures of distributions. Insurance Math. Econom. 70, 387–396.
Moyal, J.E., 1955. Theory of ionization fluctuations. Lond. Edinb. Dublin Philos. Mag. J. Sci. 46 (374), 263–280.
Nadarajah, S., Bakar, S.A.A., 2014. New composite models for the Danish fire insurance data. Scand. Actuar. J. 2, 180–187.
Olver, F.W.J., Lozier, D.W., Boisvert, R.F., Clark, C.W., 2010. NIST Handbook of Mathematical Functions. Cambridge University Press, New York.
Punzo, A., Bagnato, L., Maruotti, A., 2017. Compound unimodal distributions for insurance losses. Insurance Math. Econom. Accepted Manuscript.
Resnick, S.I., 1997. Discussion of the danish data on large fire insurance losses. Astin Bull. 27 (1), 139–151.
Reynkens, T., Verbelen, R., Beirlant, J., Antonio, K., 2017. Modelling censored losses using splicing: A global fit strategy with mixed Erlang and extreme value distributions. Insurance Math. Econom. 77, 65–77.
Rolski, T., Schmidli, H., Schmidt, V., Teugels, J., 1999. Stochastic Processes for Insurance and Finance. John Wiley and sons.
Scollnik, D.P., 2007. On composite log-normal-pareto models. Scand. Actuar. J. 1, 20–33.
Scollnik, D.P., Sun, C., 2012. Modelling with weibull-pareto models. N. Am. Actuar. J. 16 (2), 260–272.
Verbelen, R., Gong, L., Antonio, K., Badescu, A., Lin, S., 2015. Fitting mixtures of Erlangs to censored and truncated data using the EM algorithm. Astin Bull. 45 (3), 729–758.
Vernic, R., 2006. Multivariate skew-normal distributions with applications in insurance. Insurance Math. Econom. 38 (2), 413–426.