

Article

On the Use of Lehmann's Alternative to Capture Extreme Losses in Actuarial Science

Emilio Gómez-Déniz ^{1,†}  and Enrique Calderín-Ojeda ^{2,*,†} 

¹ Department of Quantitative Methods in Economics and TiDES, University of Las Palmas de Gran Canaria, 35017 Las Palmas de Gran Canaria, Spain; emilio.gomez-denz@ulpgc.es

² Department of Economics, University of Melbourne, Melbourne, VIC 3010, Australia

* Correspondence: enrique.calderin@unimelb.edu.au

† These authors contributed equally to this work.

Abstract: This paper studies properties and applications related to the mixture of the class of distributions built by the Lehmann's alternative (also referred to in the statistical literature as max-stable or exponentiated distribution) of the form $[G(\cdot)]^\lambda$, where $\lambda > 0$ and $G(\cdot)$ is a continuous cumulative distribution function. This mixture can be useful in economics, financial, and actuarial fields, where extreme and long tails appear in the empirical data. The special case in which $G(\cdot)$ is the Stoppa cumulative distribution function, which is a good description of the random behaviour of large losses, is studied in detail. We provide properties of this mixture, mainly related to the analysis of the tail of the distribution that makes it a candidate for fitting actuarial data with extreme observations. Inference procedures are discussed and applications to three well-known datasets are shown.

Keywords: actuarial; Lehmann's alternative; mixture; right tail; Stoppa distribution

MSC: 62E10; 62F10; 62P05

1. Introduction

As (Boland 2007, p. IX) has pointed out, when a practitioner wants to model losses, there is usually considerable concern about the chances and sizes of large claims. More precisely, the study of the right tail of the distribution results is very important in order to not underestimate the size of large losses. The term long tail is useful here and is applied to rank-size distributions or rank-frequency distributions, which often form power laws. In this sense, the Pareto distribution and generalisations of this simple distribution, such as the Stoppa distribution (see Stoppa 1990; Kleiber and Kotz 2003), have been considered in this case. Other alternatives to the Pareto distribution have been proposed recently in the statistical and applied statistical literature. Some of them are provided by Sarabia and Prieto (2009), Gómez-Déniz and Calderín-Ojeda (2014, 2015) and Ghitany et al. (2018), among others.

One of the main advantages obtained when the Stoppa distribution is considered is the fact that this is a max-stable (maximum of a number of random variables) distribution, which is also related to extreme value theory. This extreme property can, in turn, be highlighted by making a mixture (compound) of it, gaining an even heavier tail.

In this paper, we study properties and applications related to the mixture of the class of probability distributions built by Lehmann's alternative (also known as max-stable or exponentiated distributions). The reader is recommended to see Lehmann (1953) and Sarabia and Castillo (2005), among other references. A positive random variable X is said to be built by the Lehmann's alternative method if its cumulative distribution function (cdf) can be represented as:

$$[G(x)]^\lambda, \quad \lambda > 0, \quad (1)$$



Citation: Gómez-Déniz, Emilio, and Enrique Calderín-Ojeda. 2024. On the Use of Lehmann's Alternative to Capture Extreme Losses in Actuarial Science. *Risks* 12: 6. <https://doi.org/10.3390/risks12010006>

Received: 30 October 2023

Revised: 12 December 2023

Accepted: 22 December 2023

Published: 28 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

where $G(\cdot)$ is a continuous cdf. This class of distributions is also known in the statistical literature as Lehmann's alternative and exponentiated distribution. In practise, if X represents, for instance, the loss of a given risk in a portfolio of insurances, a population of such policy-holders may be a ubiquitous variation in λ -values because of small fluctuations in the mean of losses, so that a policy-holder selected at random can be regarded as having a random value of λ , call it Λ . Here, Λ takes values in $(0, \infty)$. The distribution of Λ across the positive real numbers is referred to in the statistical literature as the mixing distribution (or density). These mixtures are useful in economics, financial, and actuarial fields, where extreme and long tails appear in the empirical data.

Although in the statistical literature, mixtures of discrete distributions for providing thick tails (for example, the well-known Poisson–Gamma distribution) are much more frequent, mixtures of continuous distributions have also been the subject of research in the field of applied statistics and, in particular, in actuarial statistics. One example of this is the exponential-inverse Gaussian model provided by [Frangos and Karlis \(2004\)](#). Other examples of continuous mixture distributions were studied by [Fung and Seneta \(2007\)](#) and [Gómez-Déniz et al. \(2013\)](#). Also, some mixtures of continuous distributions have recently been considered in the actuarial literature in [Tzougas and Karlis \(2020\)](#) and [Tzougas and Jeong \(2021\)](#).

The special case in which (1) corresponds to the Stoppa cdf (see for example, [Stoppa 1990](#); [Kleiber and Kotz 2003](#))—which is a good description of the random behaviour of large losses—is studied in some detail. The Stoppa distribution is a generalisation of the classical Pareto distribution and is a max-stable distribution with a nice and simple cdf. We provide properties of this mixture, mainly related to the analysis of the tail of the distribution that makes it a candidate for fitting actuarial data with extreme observations. Inference procedures are discussed and applications to two well-known datasets are shown. Furthermore, two regression models for this mixture are derived to fit bodily injury claims data.

The remainder of the paper is organised as follows. The mixture representation of a max-stable distribution is provided in Section 2. Here, the representation of the Stoppa distribution is also considered and some basic properties are derived. Section 3 discusses properties related to the right tail of this mixture distribution. Some special cases of the mixing distribution are considered in Section 4. In particular, special attention is paid to the generalised inverse Gaussian distribution, which includes as special cases the inverse Gaussian, Gamma, and exponential distributions, among other models. Next, the stochastic ordering of the resulting mixture models is examined. Finally, two regression models for this mixture of distributions are given. Then, three numerical applications are shown in Section 5 and conclusions are provided in the last section.

2. Lehmann's Alternative

The distribution function built by Lehmann's alternative method, $G(z)^\lambda$, assumed that $Z \sim G(z)$, can be interpreted in different ways (see for instance [Sarabia and Castillo 2005](#)); for example, it is the distribution of the last (maximum) order statistics in a random sample of size n . That is, consider the sample Z_1, \dots, Z_n of n independent and identically distributed random variables with common cdf $G(z)$ and define the ordered sample $Z_{1,n}, \dots, Z_{n,n}$. If the interest of the practitioner is the asymptotic distribution of the maxima M_n as $n \rightarrow \infty$, the distribution of M_n results in $G^n(z) = \Pr(M_n \leq z) = \Pr(Z_1 \leq z, \dots, Z_n \leq z)$. On the other hand, exponentiated distributions of the form $G(z)^\lambda$ also appear by applying the probability integral transformation or the quantile function of G to a Beta distribution with parameters α and 1.

We now consider extensions of the distributions built by the Lehmann's alternative method of the form

$$F(x|\lambda) = [G(x)]^\lambda, \quad \lambda > 0,$$

assuming that the parameter $\lambda > 0$ is not constant and varies according to some known probability distribution, say $m(\lambda; \boldsymbol{\omega})$, depending on a vector of parameters $\boldsymbol{\omega}$. This has the following stochastic representation,

$$\begin{aligned} X &\sim [G(x)]^\lambda = \exp[\lambda \log G(x)], \\ \lambda &\sim m(\lambda; \boldsymbol{\omega}). \end{aligned}$$

Thus, we study the mixture of the form,

$$F(x; \boldsymbol{\omega}) = E_{m(\lambda; \boldsymbol{\omega})} \{ \exp[\lambda \log G(x)] \}, \tag{2}$$

where $m(\lambda; \boldsymbol{\omega})$ is a genuine mixing probability density function of Λ , the parametric space in which the parameter λ moves, depending on a vector of parameters $\boldsymbol{\omega}$. Another possibility, which will not be studied here, is to consider that the parameter λ takes integer values $\{1, 2, \dots\}$ and allows λ to follow a discrete distribution such as the geometric discrete distribution. This is the family of geometric max-stable distribution considered by [Marshall and Olkin \(1997\)](#).

Observe that (2) can be written as

$$F(x; \boldsymbol{\omega}) = M_\Lambda(\log G(x)), \tag{3}$$

in which $M(\cdot)$ denotes the moment-generating function. Thus, attending to (3), any probability density function (pdf) with a closed-form for its moment generating function, such as the Gamma, the inverse Gaussian, and the generalised inverse Gaussian distributions, should be good candidates to introduce here.

The pdf, obtained from (2), is given by

$$f(x; \boldsymbol{\omega}) = \frac{g(x)}{G(x)} E_{m(\lambda; \boldsymbol{\omega})} \{ \lambda \exp[\lambda \log G(x)] \},$$

where $g(x) = dG(x)/dx$, and the hazard rate function results in

$$r(x; \boldsymbol{\omega}) = \frac{g(x)}{G(x)} \frac{E_{m(\lambda; \boldsymbol{\omega})} \{ \lambda \exp[\lambda \log G(x)] \}}{1 - E_{m(\lambda; \boldsymbol{\omega})} \{ \exp[\lambda \log G(x)] \}}.$$

The Lehmann's Alternative Stoppa Distribution

The natural generalisation of the Pareto distribution proposed by [Stoppa \(1990\)](#) (see also [Kleiber and Kotz 2003](#)) has a nice and simple cumulative distribution function (hereafter cdf) representation given by

$$F(x; \sigma, \theta, \lambda) = [\psi_{\sigma, \theta}(x)]^\lambda = \exp[\lambda \log \psi_{\sigma, \theta}(x)], \quad x > \sigma, \tag{4}$$

where $\sigma > 0$, $\theta > 0$, and $\lambda > 0$ is a shape parameter that allows for unimodality when $\lambda > 1$; furthermore, the mode is located at zero when $\lambda \leq 1$. Observe that

$$\psi_{\sigma, \theta}(x) = 1 - \left(\frac{\sigma}{x}\right)^\theta \tag{5}$$

is the cdf of the classical Pareto distribution with scale parameter $\sigma > 0$ and shape parameter $\theta > 0$ (see for instance [Arnold 1983](#)) with pdf

$$g(x) = \frac{\theta \sigma^\theta}{x^{\theta+1}}, \quad x > \sigma > 0, \quad \theta > 0$$

and survival function

$$\bar{G}(x) = \left(\frac{\sigma}{x}\right)^\theta, \quad x > \sigma > 0, \quad \theta > 0.$$

Note that (4) obeys the cdf of a Lehmann’s alternative (max-stable) distribution, such as the one given in (1). Thus, this cdf is simply a power transformation of the classical Pareto cdf obtained as a special case of (4) for $\lambda = 1$. Some additional properties, such as moments, Lorenz curves, the Gini index, and estimation of this distribution can be found in (Kleiber and Kotz 2003, chp. 3), where other generalisations of the classical Pareto distribution can also be viewed.

Henceforward, we will write $X \sim S(\sigma, \theta, \lambda)$ to denote that a continuous random variable with support in (σ, ∞) follows the distribution given in (4).

To see how the Stoppa model works as compared to the classical Pareto distribution, we will consider data taken from (Hogg and Klugman 1984, p. 64) (see also Boyd 1988), which concern 40 losses recorded in 1977 related to wind catastrophes. This set of empirical data was recorded to the nearest USD 1,000,000, although the data include only losses of USD 2,000,000. Maximum likelihood estimates were computed for the classical Pareto distribution and the Stoppa distribution, assuming $\sigma = 1$, and the results appear in Table 1. As we can see, the Stoppa distribution provides a better fit than the one obtained with the classical Pareto distribution.

Table 1. Parameter estimators and log-likelihood for the different models considered.

Model	$\hat{\theta}$	$\hat{\lambda}$	NLL
Pareto	0.582	–	130.331
Stoppa	1.198	3.861	119.372

3. Extremal Properties

As was mentioned before, a random variable with non-negative support, such as the classical Pareto distribution, is commonly used in insurance and other financial scenarios to model the amount of claims (losses). In this sense, the size of the distribution tail is fundamental if it is desired that the chosen model allows us to capture amounts sufficiently far from the start of the distribution support, that is, extreme values. Due to this, the use of heavy right-tailed distributions such as the Pareto, lognormal, and Weibull (with shape parameter smaller than 1) distributions, among others, have been employed to model losses in motor third-party liability insurance, fire insurance, or catastrophe insurance, among others. In order to make the paper self-contained, in the following, we provide the formal definition of a heavy-tailed (heavy right-tailed) distribution (see Rolski et al. 1999).

Definition 1. Any probability distribution that is specified through its cdf $F(x)$ on the real line, is heavy right-tailed if $\limsup_{x \rightarrow \infty} (-\log \bar{F}(x)/x) = 0$.

Observe that $-\frac{d}{dx} \log \bar{F}(x)$ is the hazard function of a random variable with cdf $F(x)$.

An important issue in extreme value theory is the regular variation (see Bingham et al. 1989; Konstantinides 2017); that is, a flexible description of the variation of some function according to the polynomial form of the type $x^{-\delta} + o(x^{-\delta})$, $\delta > 0$. This concept is formalised in the following definition.

Definition 2. A distribution function (measurable function) is called regular varying at infinity with index $-\delta$ if it holds that

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(\tau x)}{\bar{F}(x)} = \tau^{-\delta},$$

where $\tau > 0$ and the parameter $\delta \geq 0$ is called the tail index.

The next proposition establishes that the survival function of the distribution given in (2) is a regular variation Lebesgue measure.

Proposition 1. *The survival function of any mixture with cdf as the one given in (2) when $G(x)$ is given in (5) is a survival function with regularly varying tails.*

Proof. This is straightforwardly verified by taking into account that

$$\begin{aligned} \limsup_{x \rightarrow \infty} \frac{\bar{F}(\tau x)}{\bar{F}(x)} &= \limsup_{x \rightarrow \infty} \frac{1 - E_{m(\lambda, \omega)}\{\exp[\lambda \log \psi_{\sigma, \theta}(\tau x)]\}}{1 - E_{m(\lambda, \omega)}\{\exp[\lambda \log \psi_{\sigma, \theta}(x)]\}} \\ &= \limsup_{x \rightarrow \infty} \frac{\tau \psi'_{\sigma, \theta}(\tau x) \psi_{\sigma, \theta}(x) E_{m(\lambda, \omega)}\{\lambda \exp[\lambda \log \psi_{\sigma, \theta}(\tau x)]\}}{\psi'_{\sigma, \theta}(x) \psi_{\sigma, \theta}(\tau x) E_{m(\lambda, \omega)}\{\lambda \exp[\lambda \log \psi_{\sigma, \theta}(x)]\}} \\ &= \tau^{-\theta}. \end{aligned}$$

Now, taking into account that $\theta > 0$, the statement of the result follows immediately. \square

Therefore, as a regular varying at infinity class is a long-tailed distribution, and the latter distribution is also heavy right-tailed, then any mixture of the Stoppa distribution is heavy right-tailed.

An immediate consequence of the previous result is the following (see [Jessen and Mikoshch 2006](#)):

Corollary 1. *If X, X_1, \dots, X_n is a sequence of iid random variables with a survival function provided by $\bar{F}(x) = 1 - F(x)$ and $S_n = \sum_{i=1}^n X_i, n \geq 1$, then*

$$\Pr(S_n > x) \sim \Pr(X > x) \text{ as } x \rightarrow \infty,$$

where the symbol \sim means it is asymptotically equivalent in probability. As a consequence of this result, if $P_n = \max_{i=1, \dots, n} X_i, n \geq 1$, then

$$\Pr(S_n > x) \sim n \Pr(X > x) \sim \Pr(P_n > x).$$

This result is important because, for large x , the event $\{S_n > x\}$ is due to the event $\{P_n > x\}$. Thus, exceedances of high thresholds by the sum S_n are explained by the exceedance of this cut-off by the largest value in the sample.

4. Some Candidates of Mixing Distributions

Consider the pdf of the generalised inverse Gaussian distribution given by

$$m(\lambda; \omega) = \frac{(\beta/\Gamma)^{\alpha/2}}{2K_\alpha(\sqrt{\beta\Gamma})} \lambda^{\alpha-1} \exp\left[-\frac{1}{2}\left(\beta\lambda + \frac{\Gamma}{\lambda}\right)\right], \quad \lambda > 0, \tag{6}$$

where $\omega = (\alpha, \beta, \Gamma)$ and $K_\nu(z)$ gives the modified Bessel function of the third kind and with index ν given by

$$K_\nu(z) = \sqrt{\pi} (2z)^\nu e^{-z} \mathcal{U}\left(\nu + \frac{1}{2}, 2\nu + 1, 2z\right),$$

$\mathcal{U}(\cdot, \cdot, \cdot)$ being the confluent hypergeometric function ([Johnson et al. 2005](#)) in which integral representation,

$$\mathcal{U}(a, b, z) = \frac{1}{\Gamma(a)} \int_0^\infty t^{a-1} (1+t)^{-a+b-1} \exp(-zt) dt.$$

A distribution with a pdf as the one given in (6) will be denoted as $GIG(\alpha, \beta, \Gamma)$.

The domain of the parameters $\omega = (\alpha, \beta, \Gamma)$ (see [Barndorff-Nielsen and Halgreen 1977](#); [Seshadri 1994](#); [Lemonte and Cordeiro 2011](#)) are given by $\alpha \in \mathbb{R}$ and $(\beta, \Gamma) \in Y_\alpha$, where

$$Y_\alpha = \begin{cases} \Gamma \geq 0, \beta > 0. & \text{if } \alpha > 0, \\ \Gamma > 0, \beta > 0, & \text{if } \alpha = 0, \\ \Gamma > 0, \beta \geq 0, & \text{if } \alpha < 0. \end{cases}$$

Its moment-generating function is given by

$$M(t) = \left(\frac{\beta}{\beta - 2t} \right)^{\alpha/2} \frac{K_\alpha(\sqrt{(\beta - 2t)\Gamma})}{K_\alpha(\sqrt{\beta\Gamma})}. \tag{7}$$

This distribution includes a lot of basic distributions such as the Gamma distribution ($\alpha > 0$ and $\Gamma \rightarrow 0$); the reciprocal Gamma distribution ($\alpha < 0$ and $\Gamma \rightarrow 0$); the inverse Gaussian distribution ($\alpha = -1/2$); the hyperbolic distribution ($\alpha = 0$); and others such as the exponential, chi-squared, half-normal, etc.

Now, by using (3) together with (7), we get the mixture of the Stoppa distributions with the generalised inverse Gaussian distribution (SGIG). In order to simplify the notation, we introduce the expression $\varphi_{\sigma,\beta,\theta}(x) = \beta - 2 \log \psi_{\sigma,\theta}(x)$. Then, its cdf results in

$$F(x) = \left(\frac{\beta}{\varphi_{\sigma,\beta,\theta}(x)} \right)^{\alpha/2} \frac{K_\alpha(\sqrt{\Gamma\varphi_{\sigma,\beta,\theta}(x)})}{K_\alpha(\sqrt{\Gamma\beta})}, \quad x > 0. \tag{8}$$

One special case obtained from the last is the mixture of the Stoppa distribution with the Gamma distribution (SG) with cdf,

$$F(x) = \left(\frac{\beta}{\varphi_{\sigma,\beta,\theta}(x)} \right)^{\alpha/2}, \quad x > 0. \tag{9}$$

Observe that, in this case, the mixture distribution is again a max-stable distribution. Another interesting sub-model obtained from the generalised inverse Gaussian distribution is the inverse Gaussian distribution, which is reached when $\alpha = -1/2$. In this case, (3) can be written as

$$F(x) = \exp\left\{ \sqrt{\Gamma\beta} - \sqrt{\Gamma\varphi_{\sigma,\beta,\theta}(x)} \right\}, \quad x > 0,$$

which corresponds to the mixture of the Stoppa model with the inverse Gaussian distribution (SIG).

Now, the pdf obtained from (8) can be written as

$$f(x) = \sqrt{\frac{\Gamma\beta^\alpha}{\varphi_{\sigma,\beta,\theta}(x)^{\alpha+1}}} \frac{\theta\sigma^\theta}{x^{\theta+1}\psi_{\sigma,\theta}(x)} \frac{K_{\alpha+1}(\sqrt{\Gamma\varphi_{\sigma,\beta,\theta}(x)})}{K_\alpha(\sqrt{\Gamma\beta})}, \quad x > 0. \tag{10}$$

From (9), the pdf results in

$$f(x) = \frac{\alpha\theta\sigma^\theta}{x^{\theta+1}\beta\psi_{\sigma,\theta}(x)} \left(\frac{\beta}{\varphi_{\sigma,\beta,\theta}(x)} \right)^{\alpha/2+1}, \quad x > 0.$$

It is known that $K_{-1/2}(z) = K_{1/2}(z) = \sqrt{\pi/(2z)} \exp(-z)$. Then, the pdf of the SIG distribution can be written as

$$f(x) = \sqrt{\frac{\Gamma}{\varphi_{\sigma,\beta,\theta}(x)}} \frac{\theta\sigma^\theta}{x^{\theta+1}\psi_{\sigma,\theta}(x)} \exp\left[\sqrt{\Gamma\beta} - \sqrt{\Gamma\varphi_{\sigma,\beta,\theta}(x)} \right], \quad x > 0.$$

In Figure 1, some graphs of the pdf (10) are shown for special cases of the parameters. As can be noted, all the cases shown in the graph include a large tail.

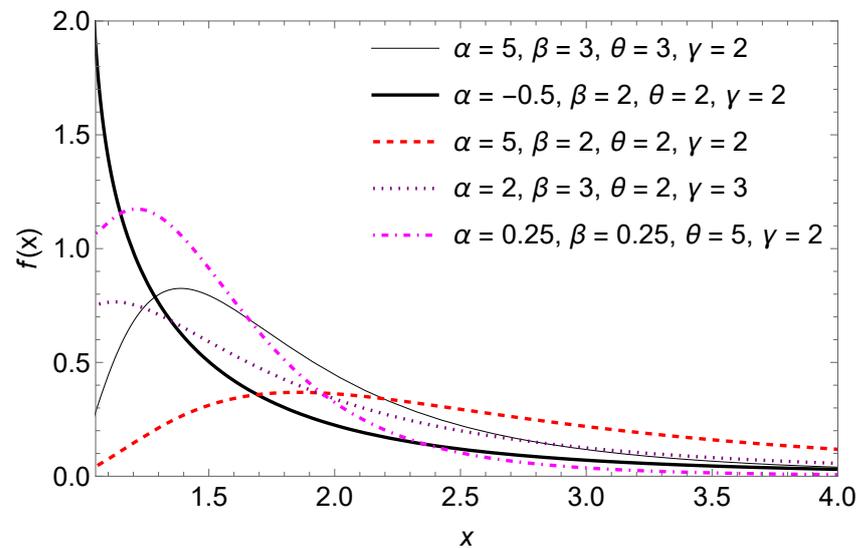


Figure 1. Different shapes of the pdf (10) for special values of the parameters.

In financial economics and risk theory, the accurate calculation of risk capital is an issue of crucial interest to researchers, regulators of financial institutions, and commercial vendors of financial products and services. In this sense, the Value-at-Risk (VaR), which is defined as the amount of capital required to ensure that the insurer does not become insolvent with a high degree of certainty, is an important risk measure. For a random variable X that follows the Stoppa and the different mixtures proposed here, i.e., the q -quantile, $0 < q < 1$ are given by

$$\begin{aligned}
 \text{S : } \quad & \text{VaR}[X; q] = \sigma(1 - q^{1/\lambda})^{-1/\theta}, \\
 \text{SG : } \quad & \text{VaR}[X; q] = \sigma \left\{ 1 - \exp \left[\frac{\beta}{2} (1 - q^{-2/\alpha}) \right] \right\}^{-1/\theta}, \\
 \text{SIG : } \quad & \text{VaR}[X; q] = \sigma \left\{ 1 - \exp \left[\frac{\beta}{2} - \frac{1}{2\Gamma} (\sqrt{\Gamma\beta} - \log q)^2 \right] \right\}^{-1/\theta}.
 \end{aligned}$$

4.1. Stochastic Ordering

It is our goal now to investigate the stochastic order of the distribution in (4) and its mixture. The analysis of stochastic ordering has been widely studied in many different areas such as economics, operations research, reliability, and statistics (e.g., survival analysis) among other fields. Also, comparing random variables is vital in risk theory. Numerous ordering concepts have been used in the statistical literature, e.g., the usual stochastic order, hazard rate order, reversed hazard rate order, and the likelihood ratio order among others. For a comprehensive understanding of stochastic ordering, the reader is encouraged to read [Shaned and Shanthikumar \(2007\)](#) and [Yu \(2009\)](#), among others. In this work, we use the stochastic dominance concept provided in [Dhaene et al. \(2006\)](#). We recall here the definition given in that book.

Definition 3. Let X_1 and X_2 be continuous random variables with cdf $F_{X_1}(x)$ and $F_{X_2}(x)$, respectively. Then, X_1 is said to be smaller than X_2 in the stochastic dominance sense (written as $X_1 \leq_{st} X_2$) if $F_{X_1}(x) \geq F_{X_2}(x)$ for all $x \in \mathbb{R}$.

In the following result, it is shown that the stochastic dominance can be expressed in terms of the Value-at-Risk.

Proposition 2. Let X_1 and X_2 be continuous random variables with cdf $F_{X_1}(x)$ and $F_{X_2}(x)$, respectively. We have that X_1 is smaller than X_2 in the stochastic dominance sense if and only if their respective quantiles are ordered:

$$X_1 \leq_{st} X_2 \iff VaR[X_1; q] \leq VaR[X_2; q], \text{ for all } q \in (0, 1).$$

Proof. See [Dhaene et al. \(2006\)](#). \square

Proposition 3. Let $X_1 \sim S(\sigma_1, \theta_1, \lambda_1)$ and $X_2 \sim S(\sigma_2, \theta_2, \lambda_2)$. Then, if $\lambda_1 < \lambda_2$, $\theta_1 > \theta_2$ and $\sigma_1 < \sigma_2$, then $X_1 \leq_{st} X_2$.

Proof. If $\lambda_1 < \lambda_2$, $\theta_1 > \theta_2$ and $\sigma_1 < \sigma_2$, we have that

$$VaR[X_1; q] = \sigma_1(1 - q^{1/\lambda_1})^{-1/\theta_1} \leq VaR[X_2; q] = \sigma_2(1 - q^{1/\lambda_2})^{-1/\theta_2},$$

and therefore by using the previous theorem we have that $X_1 \leq_{st} X_2$. \square

Proposition 4. Let $X_1 \sim SG(\sigma_1, \theta_1, \alpha_1, \beta_1)$ and $X_2 \sim SG(\sigma_2, \theta_2, \alpha_2, \beta_2)$. Then, if $\sigma_1 < \sigma_2$, $\theta_1 > \theta_2$, $\alpha_1 < \alpha_2$, and $\beta_1 > \beta_2$, then $X_1 \leq_{st} X_2$.

Proof. If $\sigma_1 < \sigma_2$, $\theta_1 > \theta_2$, $\alpha_1 < \alpha_2$, and $\beta_1 > \beta_2$, we have that

$$\begin{aligned} VaR[X_1; q] &= \sigma_1 \left\{ 1 - \exp \left[\frac{\beta_1}{2} \left(1 - q^{-2/\alpha_1} \right) \right] \right\}^{-1/\theta_1} \leq \\ VaR[X_2; q] &= \sigma_2 \left\{ 1 - \exp \left[\frac{\beta_2}{2} \left(1 - q^{-2/\alpha_2} \right) \right] \right\}^{-1/\theta_2}, \end{aligned}$$

and therefore by using the previous theorem, we have that $X_1 \leq_{st} X_2$. \square

It is easy to see that, if $X_1 \sim S(\sigma, \theta_1, \lambda)$ and $X_2 \sim S(\sigma, \theta_2, \lambda)$ with $\theta_1 < \theta_2$, then $\bar{F}(x_1) \leq_{st} \bar{F}(x_2)$. Furthermore, if $X_1 \sim S(\sigma, \theta, \lambda_1)$ and $X_2 \sim S(\sigma, \theta, \lambda_2)$ with $\lambda_1 < \lambda_2$, then $\bar{F}(x_2) \leq_{st} \bar{F}(x_1)$. Now, we have the following result.

Theorem 1. Let X_1 and X_2 be two mixture random variables with cdfs $F_1(x; \sigma, \theta_1, \omega)$ and $F_2(x; \sigma, \theta_2, \omega)$ obtained from (3). If $\theta_1 \leq \theta_2$, then $X_1 \leq_{st} X_2$.

Proof. The result follows by applying Theorem 1.C.17 in [Shaned and Shanthikumar \(2007\)](#) \square

More effort will be necessary to get this quantity for the SGIG distribution.

4.2. Estimation

In this section, we show how to estimate the parameters of the mixture distributions via maximum likelihood estimation. For that reason, let us assume that $\{x_1, x_2, \dots, x_n\}$ is a random sample selected from the distribution of interest (10) and also assume that $\sigma = \min\{x_i\}$, $i = 1, \dots, n$. The log-likelihood function, obtained from (10), is proportional to

$$\begin{aligned} \ell(\mathbf{\Omega}; \tilde{x}) &\propto n \left[\theta \log \sigma + \log \theta + \frac{\alpha}{2} \log \beta + \frac{1}{2} \log \Gamma - K_\alpha(\sqrt{\Gamma\beta}) \right] \\ &\quad - \frac{\alpha + 1}{2} \sum_{i=1}^n \log \varphi_{\sigma, \beta, \theta}(x_i) - \sum_{i=1}^n \psi_{\sigma, \theta}(x_i) \\ &\quad + \sum_{i=1}^n K_{\alpha+1} \sqrt{\Gamma \varphi_{\sigma, \beta, \theta}(x_i)}, \end{aligned}$$

where $\Omega = (\alpha, \beta, \theta, \Gamma)$ is the vector of parameters to be estimated.

Although in practice, the modified Bessel function of the third kind is implemented in most statistical packages (note that this is not the case of the econometric software WinRats), it is convenient to express this function in terms of the modified Bessel function of the first kind, $I_\nu(z)$. The relationship between both functions (see Johnson et al. 2005, p. 19) is given by

$$K_\nu(z) = \frac{\pi \csc(\pi\nu)}{2} [I_{-\nu}(z) - I_\nu(z)],$$

where

$$I_\nu(z) = \sum_{j=0}^{\infty} \frac{1}{j! \Gamma(j + \nu + 1)} \left(\frac{z}{2}\right)^{2j+\nu},$$

is the modified Bessel function of the first kind. The advantage of this function is that it is written as a series representation that facilitates (by truncation) the estimation procedure of the parameters of the model. More details about the maximum likelihood estimation are given in the Appendix A.

Conjugate Distribution

In this subsection, we will see that the GIG distribution is a conjugate distribution with respect to the Stoppa distribution.

Theorem 2. *Let $X \sim S(\sigma, \theta, \lambda)$ and suppose that, given $\Lambda = \lambda$, $\Lambda \sim GIG(\alpha, \beta, \Gamma)$. Then we have that the posterior distribution of λ , given n years of individual independent experience, $\tilde{x} = (x_1, \dots, x_n)$ is again a $GIG(\alpha^*, \beta^*, \Gamma^*)$, where the updated parameters are given by,*

$$\alpha^* = \alpha + n, \tag{11}$$

$$\beta^* = \beta - 2 \sum_{i=1}^n \log \psi_{\sigma, \theta}(x_i)$$

$$\Gamma^* = \Gamma. \tag{12}$$

Proof. The pdf of the Stoppa distribution, obtained from (4), is given by

$$f(x) = \frac{\lambda \psi'_{\sigma, \theta}(x)}{\psi_{\sigma, \theta}(x)} \exp[\lambda \log \psi_{\sigma, \theta}(x)]. \tag{13}$$

Thus, given the sample information \tilde{x} , the likelihood, taken from (13), is proportional to

$$\lambda^n \exp \left\{ \lambda \sum_{i=1}^n \log \psi_{\sigma, \theta}(x_i) \right\}.$$

Therefore, by applying Bayes' theorem, we have that the posterior is proportional to

$$\lambda^{\alpha+n-1} \exp \left\{ -\frac{1}{2} \left[\left(\beta - 2 \sum_{i=1}^n \log \psi_{\sigma, \theta}(x_i) \right) \lambda + \frac{\Gamma}{\lambda} \right] \right\},$$

which is a GIG distribution with parameters (11) and (12). \square

Observe that, if $X \sim S(\sigma, \theta, \lambda)$ and Λ follow a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$, then using the previous result, we have that the posterior mean of the parameter λ given the sample information \tilde{x} results in

$$E(\Lambda | \tilde{x}) = \frac{\alpha + n}{\beta - \sum_{i=1}^n \log \psi_{\sigma, \theta}(x_i)}.$$

Consequently, this Bayesian expression takes the form

$$E(\Lambda|\tilde{x}) = (1 - z_n)r(\tilde{x}) + z_nE_g(\Lambda), \tag{14}$$

where $r(\tilde{x}) = -n / \sum_{i=1}^n \log \psi_{\sigma,\theta}(x_i)$ and

$$z_n = \frac{\beta}{\beta - \sum_{i=1}^n \log \psi_{\sigma,\theta}(x_i)}.$$

Since it is not verified that $E(X|\lambda) = \lambda$, then it is not guaranteed that the mean of the predictive distribution coincides with the posterior mean, i.e., $E(X_{n+1}|\tilde{x}) \neq E(\Lambda|\tilde{x})$. Therefore, although the expression (14) resembles the credibility formula widely introduced in actuarial statistics with credibility factor z_n , this is not a credibility premium since the purpose of that expression is to predict the expected aggregate claims size in the next period of time.

4.3. Novel Heavy-Tail Regression Models

The mean of the Stoppa distribution is given by the expression:

$$E(X) = \lambda\sigma B\left(1 - \frac{1}{\theta}, \lambda\right),$$

where $B(a, b)$ is the complete beta function defined as $\int_0^1 z^{a-1}(1-z)^{b-1}dz$ with $a > 0$ and $b > 0$. Therefore, the mean of the mixture Stoppa distribution can be obtained by compounding using the expression,

$$E(X) = E_\Lambda(E(X|\Lambda)) = \sigma E_\Lambda\left(\Lambda B\left(1 - \frac{1}{\theta}, \Lambda\right)\right).$$

Then, given a sample $\tilde{x} = (x_1, \dots, x_n)$, the new heavy-tailed regression models can be derived by writing the scale parameter of each model as

$$\sigma_i = \frac{\kappa \exp[\boldsymbol{\beta}^\top \mathbf{z}_i]}{1 + \exp[\boldsymbol{\beta}^\top \mathbf{z}_i]}, \quad i = 1, \dots, n.$$

Here, $\mathbf{z}_i = (z_{i1}, \dots, z_{ip})$ is a vector of explanatory variables and $\boldsymbol{\beta}^\top = (\beta_1, \dots, \beta_p)$ is a vector of regressors to be estimated. Note that, in the latter expression, κ does not depend on the subscript i . In the practical implementation, this parameter was chosen by using a grid search for manually specified values of this parameter in the interval $(0, \min\{x_1, \dots, x_n\})$.

4.3.1. SG Case

In this model, the log-likelihood function is given by

$$\begin{aligned} \ell(\boldsymbol{\Omega}; \tilde{x}) \propto & n \left[\theta \log \sigma_i + \log \theta + \log \alpha + \frac{\alpha}{2} \log \beta \right] - \sum_{i=1}^n \log \varphi_{\sigma_i, \beta, \theta}(x_i) \\ & - (\theta + 1) \sum_{i=1}^n \log x_i - \left(\frac{\alpha}{2} + 1 \right) \sum_{i=1}^n \log \psi_{\sigma_i, \theta}(x_i), \end{aligned} \tag{15}$$

where $\boldsymbol{\Omega} = (\alpha, \beta, \theta, \boldsymbol{\beta}^\top)$.

The maximum likelihood estimates of the SG distribution are obtained by solving the system of equations given by

$$\begin{aligned} \frac{n\alpha}{2\beta} - \left(\frac{\alpha}{2} + 1\right) \sum_{i=1}^n \frac{1}{\varphi_{\sigma_i, \beta, \theta}(x_i)} &= 0, \\ n\left(\frac{1}{\theta} + \log \sigma_i\right) + \sum_{i=1}^n \frac{(\sigma_i/x_i)^\theta}{\psi_{\sigma_i, \theta}(x_i)} \log\left(\frac{\sigma_i}{x_i}\right) \left[1 + \frac{\alpha}{2} - \frac{2}{\varphi_{\sigma_i, \beta, \theta}(x_i)}\right] - \sum_{i=1}^n \log x_i &= 0, \\ \frac{n}{\alpha} + \frac{\log \beta}{2} - \frac{1}{2} \sum_{i=1}^n \log \psi_{\sigma_i, \theta}(x_i) &= 0, \\ \frac{\beta_j}{1 + \exp[\boldsymbol{\beta}^\top \mathbf{z}_i]} \left(n\theta + \sum_{i=1}^n \frac{\theta}{\psi_{\sigma_i, \theta}(x_i)} \left(\frac{\sigma_i}{x_i}\right)^\theta \left[1 + \frac{\alpha}{2} - \frac{2}{\varphi_{\sigma_i, \beta, \theta}(x_i)}\right]\right) &= 0, \end{aligned}$$

where $j = 1, \dots, p$ and $i = 1, \dots, n$.

From these equations, the entries of the observed information matrix can be derived after tedious algebra (not reproduced here) by differentiating these equations with respect to the $p + 3$ parameters. The function (15) can be simply maximised by considering several values as seed points. Of course, the global maximum is not guaranteed by the difficulty to show that the log-likelihood function is concave. We have used different maximum search methods that are available in the MAXIMIZE in-built function in the WinRATS software package by using the BFGS algorithms.

4.3.2. SIG Case

When the mixing distribution is the inverse Gaussian, the log-likelihood function is given by

$$\begin{aligned} \ell(\boldsymbol{\Omega}; \tilde{x}) &\propto n \left[\theta \log \sigma_i + \log \theta + \frac{1}{2} \log \Gamma + \sqrt{\Gamma \beta} \right] \\ &\quad - (\theta + 1) \sum_{i=1}^n \log x_i - \frac{1}{2} \sum_{i=1}^n \log \varphi_{\sigma_i, \beta, \theta}(x_i) \\ &\quad - \sum_{i=1}^n \log \psi_{\sigma_i, \theta}(x_i) - \sqrt{\Gamma} \sum_{i=1}^n \sqrt{\varphi_{\sigma_i, \beta, \theta}(x_i)}, \end{aligned}$$

where $\boldsymbol{\Omega} = (\beta, \theta, \Gamma, \boldsymbol{\beta}^\top)$.

The normal equations are given, in this case, by

$$\begin{aligned} n\sqrt{\frac{\Gamma}{\beta}} - \sum_{i=1}^n \frac{1}{\varphi_{\sigma_i, \beta, \theta}(x_i)} - \sqrt{\Gamma} \sum_{i=1}^n (\varphi_{\sigma_i, \beta, \theta}(x_i))^{-1/2} &= 0, \\ n\left(\frac{1}{\theta} + \log \sigma_i\right) + \sum_{i=1}^n (\sigma_i/x_i)^\theta \log(\sigma_i/x_i) \left\{ \frac{1}{\psi_{\sigma_i, \theta}(x_i)} \left[1 - \frac{1}{\varphi_{\sigma_i, \beta, \theta}(x_i)} - \sqrt{\frac{\Gamma}{\varphi_{\sigma_i, \beta, \theta}(x_i)}}\right] \right\} - \sum_{i=1}^n \log x_i &= 0, \\ \frac{1}{\Gamma} + \sqrt{\frac{\beta}{\Gamma}} - \sum_{i=1}^n \sqrt{\frac{\varphi_{\sigma_i, \beta, \theta}(x_i)}{\Gamma}} &= 0, \\ \frac{\beta_j}{1 + \exp[\boldsymbol{\beta}^\top \mathbf{z}_i]} \left(n\theta + \sum_{i=1}^n \frac{\theta}{\psi_{\sigma_i, \theta}(x_i)} \left(\frac{\sigma_i}{x_i}\right)^\theta \left[1 - \frac{1}{\varphi_{\sigma_i, \beta, \theta}(x_i)} - \sqrt{\frac{\Gamma}{\varphi_{\sigma_i, \beta, \theta}(x_i)}}\right]\right) &= 0, \end{aligned}$$

where $j = 1, \dots, p$ and $i = 1, \dots, n$.

The standard errors of the estimates can be computed by following the same approach as above.

5. Numerical Experiments

In this section, the performance of the distributions introduced in this work is verified by employing three different sets of data. The first one is the `danishuni`, which can be downloaded from the *R* package `CASdatasets` and also from *Extreme Value Statistics* in *S-plus* libraries, collected at Copenhagen Reinsurance, comprising 2157 fire losses, adjusted for inflation to reflect 1985 values, over DKK 1,000,000 during the period 1980 to 1990, adjusted for inflation to reflect 1985. A detailed statistical analysis of this dataset can be found in [McNeil \(1988\)](#), in [Albrecher et al. \(2017\)](#), and also in [Embrechts et al. \(1997\)](#). The second dataset is `norfire`, which is also available in the *R* package `CASdatasets`, and includes 9181 fire losses over the period 1972 to 1992 from an unknown Norwegian insurer. A priority of NKR 500,000 (if this amount is exceeded, the reinsurer becomes liable to pay) was applied to derive this set of data.

The estimates of the parameters and their corresponding *p*-values (given in brackets), the negative of the maximum of the log likelihood function, and the AIC for the two aforementioned sets of data are shown in Table 2 for the Stoppa distribution and for the three mixture models previously considered, i.e., SG, SIG, and SGIG. For all the models and the two datasets, the parameter σ was chosen by using a grid search for manually specified values of this parameter in the interval $(0, \min\{x_1, \dots, x_n\})$. The validation of the models is carried out using the following information criteria: the negative log likelihood (NLL), computed by taking the negative of the value of the log-likelihood evaluated at the maximum likelihood estimates, and the Akaike's information criterion (AIC), computed as twice the NLL, evaluated at the ML estimates plus twice the number of estimated parameters. Moreover, we also incorporate the Kolmogorov–Smirnov test (KS) and the Anderson–Darling test (AD) to show the fit of the model to the empirical data in terms of the cdf. For these test statistics, smaller values of these tests indicate a better fit of the model to the empirical data. Note that they do not only provide a way to measure the fit in terms of the cdfs, but also allow us to perform hypothesis testing for model validation purposes. An extremely small *p*-value might lead to a confident rejection of the null hypothesis that the data come from the given model. It can be seen that the SG distribution provides the best fit for the Danish dataset, whereas the SGIG returns the lowest values for NLL and AIC for the Norwegian set of data. For the former dataset, and using KS and AD tests, none of the models are rejected; however, it is noted that, for both tests, the Stoppa distribution is rejected for the Norwegian dataset.

For comparison reasons, we have also fitted shifted versions of lognormal, Weibull, and Burr distributions for the Danish and Norwegian datasets with the following densities:

$$f(x) = \frac{1}{\lambda(x-\sigma)\sqrt{2\pi}} \exp\left[-\frac{1}{2\lambda^2}(\log(x-\sigma)-\theta)^2\right], \quad x > \sigma, \theta \in \mathbb{R}, \lambda > 0,$$

$$f(x) = \frac{\theta}{\lambda} \left(\frac{x-\sigma}{\lambda}\right)^{\theta-1} \exp\left[-\left(\frac{x-\sigma}{\lambda}\right)^\theta\right], \quad x > \sigma, \lambda > 0, \theta > 0,$$

$$f(x) = \frac{\lambda\theta(x-\sigma)^{\theta-1}}{(1+(x-\sigma)^\theta)^{\lambda+1}}, \quad x > \sigma, \lambda > 0, \theta > 0,$$

respectively. For all these distributions, the parameter σ was estimated by using the method explained above. A comparison of these models with the other heavy-tailed distributions used to model the Danish and the Norwegian datasets can be found in [Gómez-Déniz et al. \(2022\)](#) and [Gómez-Déniz and Calderín-Ojeda \(2015\)](#), respectively. This catalogue of heavy-tailed distributions includes shifted versions of the lognormal, inverse Gaussian, and generalised Gamma, and also inverse Gamma, log Gamma, and Fréchet and Pareto ArcTan, among other models.

The estimation of the parameters for all the distributions used in this work has been completed using the method of maximum likelihood by using `Mathematica`® v. 12.0 and has also been verified via `WinRATS` v. 7.0. The codes are available from the authors upon

request. Standard errors of the estimates were obtained by finite differentiation. The WinRATS v. 7.0 software package also gives the option to directly calculate the maximum of the log-likelihood returning the entries of the Fisher information matrix. The parameters can also be estimated via an EM algorithm as shown in the Appendix A.

Table 2. Parameter estimates and their corresponding *p*-values (in brackets), negative of the maximum of the log likelihood function (NLL), AIC, Kolmogorov–Smirnov and Anderson–Darling tests for the Weibull, Lognormal, Burr, and Stoppa distributions and mixture models: SG, SIG, and SGIG.

Danish Data									
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\Gamma}$	NLL	AIC	KS	AD
lognormal		1.557 (<0.001)		−0.290 (<0.001)		3404.340	6812.68	0.2549 (<0.001)	365.672 (<0.001)
Weibull		1.584 (<0.001)		0.660 (<0.001)		3510.610	7025.22	0.0724 (<0.001)	31.752 (<0.001)
Burr		1.223 (<0.001)		1.115 (<0.001)		3339.440	6682.89	>1 (<0.001)	>100 (<0.001)
Stoppa		1.163 (<0.001)		1.395 (<0.001)		3342.420	6688.84	0.0277 (0.3772)	0.00055 (0.4930)
SG	19.550 (0.002)		14.241 (0.008)	1.512 (<0.001)	≈0	3334.788	6675.580	0.022 (0.6368)	0.00031 (0.7955)
SIG	−0.5		5.883 (0.012)	1.517 (<0.001)	11.323 (<0.001)	3335.000	6676.000	0.022 (0.6368)	0.000328 (0.7705)
SGIG	−4.546 (<0.001)		2.989 (<0.001)	1.517 (<0.001)	16.370 (<0.001)	3335.150	6678.300	0.022 (0.6368)	0.00034 (0.767)
Norwegian Data									
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\Gamma}$	NLL	AIC	KS	AD
lognormal		6.313 (<0.001)		1.537 (<0.001)		21,097.400	42,198.80	0.0682 (<0.001)	23.082 (<0.001)
Weibull		0.689 (<0.001)		1134.39 (<0.001)		21,150.500	42,305.00	0.0765 (<0.001)	37.756 (<0.001)
Burr		0.0078 (<0.001)		20.295 (<0.001)		23,685.620	47,375.20	0.419 (<0.001)	704.395 (<0.001)
Stoppa		1.143 (<0.001)		1.124 (<0.001)		21,045.290	42,094.600	0.0854 (<0.001)	0.00627 (<0.001)
SG	6.404 (<0.001)		2.686 (<0.001)	1.593 (<0.001)	≈0	20,931.033	41,868.100	0.0158 (0.9011)	0.00016 (0.8685)
SIG	−0.5		0.559 (<0.001)	1.673 (<0.001)	4.472 (<0.001)	20,931.400	41,868.700	0.0162 (0.8846)	0.00021 (0.7905)
SGIG	1.559 (<0.001)		1.588 (<0.001)	1.632 (<0.001)	1.626 (<0.001)	20,929.900	41,867.900	0.0154 (0.9163)	0.00016 (0.8685)

These results are confirmed by Figure 2. In this figure, the graphs of the histogram of both datasets (Danish left, Norwegian right) are shown. We have also superimposed the fitted densities. Also in Figure 3, we have plotted the smooth cdf of the empirical data, and the fitted cdf of all the distributions previously considered have been superimposed. It is observable that the SG (dashed line) and SGIG (dotted-dashed line) distributions adhere closely to Danish and Norwegian datasets, respectively.

Now, we select a distribution that yields a feasible characterisation of the loss process for both sets of data, we should check that the theoretical limited expected values, calculated numerically by

$$L(x) = E[\min(X, x)] = \int_0^x y dF(y) + x\bar{F}(x), \tag{16}$$

adhere closely to the empirical ones. As is already known, (16) is the expected quantity per claim retained by the insured on a policy with a fixed amount deductible of x . Here, the empirical limited expected value function was computed based on the expression $E_n(x) = \frac{1}{n} \sum_{i=1}^n \min(x_i, x)$. Obviously, when x tends to infinity, $L(x)$ and $E_n(x)$ converge to $E(X)$ and the sample mean, respectively.

Tables 3 and 4 display the limited expected value for several values of the policy limit x considered for the Danish and Norwegian datasets, respectively. It is observed that the values obtained from the mixing distributions adhere closely to the observed empirical limited expected values obtained, as compared to the Stoppa distribution for both datasets.

The absolute errors between the empirical values and the fitted values shown in Tables 3 and 4 are shown in the graphs that appear in Figure 4. As can be seen, at least for the Danish data, the fit is improved to the extent that the value of the deductible increases. The pattern for the Norwegian data is challenging to observe.

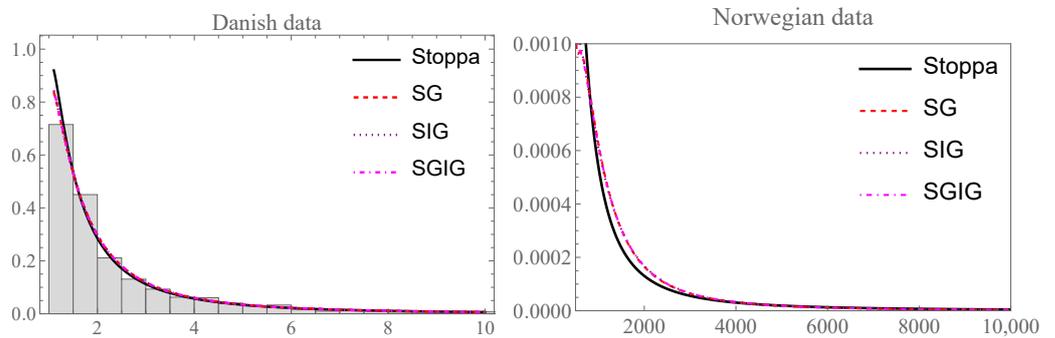


Figure 2. Empirical (smoothed) distribution (histogram) and theoretical distribution model of the different distributions considered for the Danish (left panel) and Norwegian (right panel) datasets.

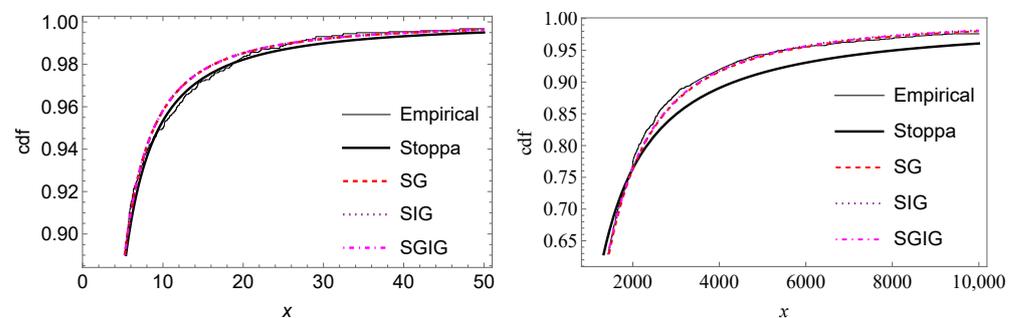


Figure 3. Smooth cdf of the empirical Danish (left panel) and Norwegian (right panel) claims data as compared to the theoretical models.

Table 3. Limited expected value for the different distributions considered, and different values of the fixed amount deductible x for the Danish dataset.

Deductible	Limited Expected Value				
	Empirical	Stoppa	SG	SIG	SGIG
6.00	2.42371	2.42879	2.44939	2.43608	2.43616
6.15	2.43641	2.44278	2.46265	2.44931	2.44940
6.30	2.44871	2.45630	2.47544	2.46206	2.46217
6.45	2.46052	2.46938	2.48778	2.47437	2.47450
6.60	2.47209	2.48205	2.49971	2.48627	2.48640
6.75	2.48345	2.49433	2.51123	2.49776	2.49791
6.90	2.49459	2.50623	2.52238	2.50888	2.50904
7.05	2.50547	2.51778	2.53317	2.51964	2.51982
7.20	2.51607	2.52899	2.54363	2.53007	2.53025
7.35	2.52637	2.53989	2.55376	2.54017	2.54037
7.50	2.53644	2.55047	2.56359	2.54997	2.55019
7.65	2.54620	2.56077	2.57313	2.55949	2.55971
7.80	2.55563	2.57080	2.58240	2.56872	2.56896
7.95	2.56491	2.58056	2.59140	2.57770	2.57795
8.10	2.57401	2.59006	2.60015	2.58642	2.58668
8.25	2.58294	2.59933	2.60866	2.59491	2.59518
8.40	2.59165	2.60836	2.61694	2.60317	2.60345
8.55	2.60026	2.61717	2.62501	2.61121	2.61150
8.70	2.60876	2.62577	2.63286	2.61904	2.61934
8.85	2.61703	2.63417	2.64052	2.62667	2.62698
9.00	2.62514	2.64237	2.64798	2.63410	2.63443
9.15	2.63324	2.65039	2.65526	2.64136	2.64169
9.30	2.64117	2.65822	2.66236	2.64844	2.64878
9.45	2.64891	2.66588	2.66929	2.65534	2.65570
9.60	2.65653	2.67337	2.67606	2.66209	2.66245
9.75	2.66414	2.68070	2.68267	2.66868	2.66905
9.90	2.67175	2.68788	2.68913	2.67512	2.67550

Table 4. Limited expected value for the different distributions considered, and different values of the fixed amount deductible x for the Norwegian dataset.

Deductible	Limited Expected Value				
	Empirical	Stoppa	SG	SIG	SGIG
600	594.306	592.565	594.307	594.14	594.189
1100	936.322	900.662	935.339	934.021	935.392
1600	1139.38	1084.90	1136.21	1136.55	1137.46
2100	1269.48	1214.54	1267.37	1269.27	1269.35
2600	1357.43	1313.79	1360.36	1363.11	1362.55
3100	1421.89	1393.80	1430.21	1433.22	1432.30
3600	1473.58	1460.58	1484.96	1487.80	1486.72
4100	1516.11	1517.74	1529.23	1531.64	1530.56
4600	1551.11	1567.61	1565.91	1567.72	1566.75
5100	1580.55	1611.75	1596.91	1598.02	1597.23
5600	1606.31	1651.30	1623.52	1623.88	1623.31
6100	1629.25	1687.09	1646.68	1646.24	1645.93
6600	1650.39	1719.74	1667.04	1665.8	1665.77
7100	1669.51	1749.72	1685.13	1683.08	1683.33
7600	1686.98	1777.43	1701.32	1698.47	1699.02
8100	1703.07	1803.17	1715.92	1712.29	1713.13
8600	1717.58	1827.19	1729.17	1724.77	1725.91
9100	1730.79	1849.69	1741.26	1736.11	1737.54
9600	1743.01	1870.85	1752.34	1746.46	1748.19

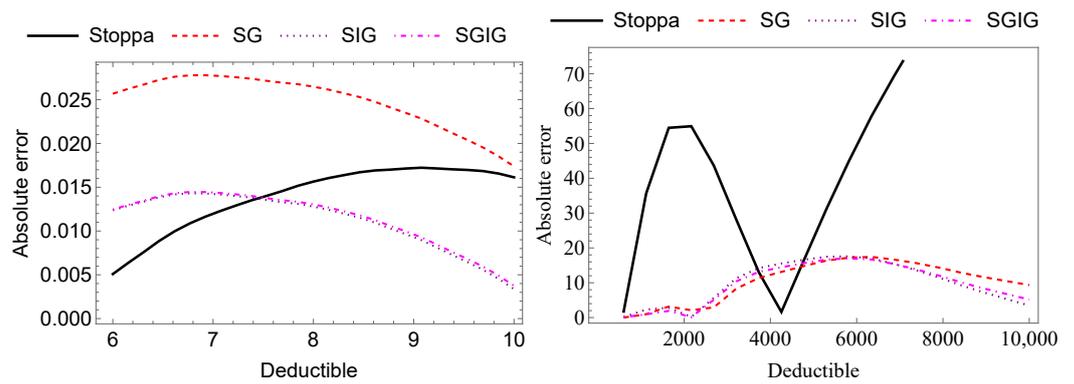


Figure 4. Absolute errors of the limited expected values for Danish dataset (left) and Norwegian dataset (right).

The third dataset deals with automobile bodily injury claims using data from the Insurance Research Council (IRC), a division of the American Institute for Chartered Property Casualty Underwriters and the Insurance Institute of America. The data, collected in 2002, include demographic information about the claimants, attorney involvement, and economic losses (in thousands of USD), among other variables. As several of these covariates include missing observations, we will only use a sample of 1091 losses. This dataset can also be downloaded from the R package *CASdatasets*, see also [Frees \(2010\)](#). We consider as a response variable the claimant’s total economic loss. The empirical distribution of this variable combines losses of small, moderate, and large sizes, which makes it suitable for fitting heavy-tailed distributions. Other remarkable features of this set of data are unimodality, skewness, and a long right tail, showing a high likelihood of extremely expensive losses. Below, in Figure 5, we have plotted the histogram of this dataset. We have also superimposed the densities of the S, SG, and SIG distributions. The SIG distribution adheres closely to empirical data.

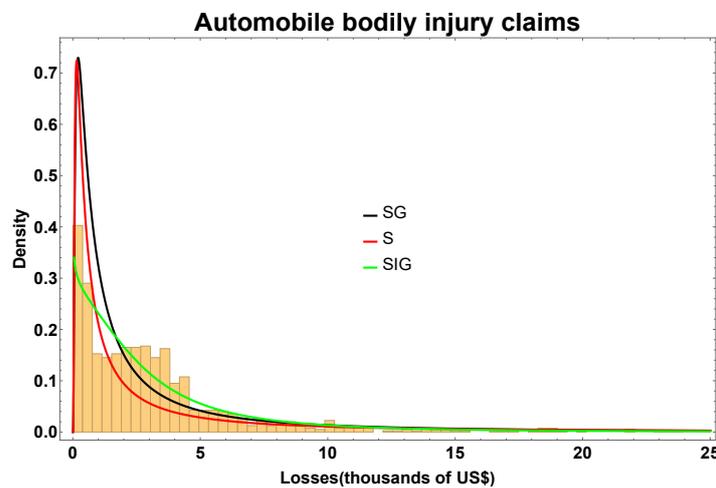


Figure 5. Histogram and densities for the S, SG, and SIG distributions for the automobile bodily injury claims dataset.

This dataset also includes the following covariates:

- *ATTORNEY* takes the value 1 if the claimant is represented by an attorney and 0 otherwise;
- *CLMSEX* takes the value 1 if the claimant is male and 0 otherwise;
- *MARRIED* takes the value 1 if the claimant is married and 0 otherwise;
- *SINGLE* takes the value 1 if the claimant is single and 0 otherwise;
- *WIDOWED* takes the value 1 if the claimant is widowed and 0 otherwise;
- *CLMINSUR*, whether or not the claimant’s vehicle was uninsured (=1 if yes and 0 otherwise);

- *SEATBELT*, whether or not the claimant was wearing the seatbelt/child restraint (=1 if yes and 0 otherwise);
- *CLMAGE*, claimant’s age.

Now, by using these covariates, we will explain the total losses in terms of the set of explanatory variables by using the Stoppa and SG and SIG regression models. From left to right, in Table 5, the parameter estimates, standard errors (S.E.), and the corresponding *p*-values calculated based on the *t*-Wald statistics for the three regression models are illustrated. Furthermore, NLL and AIC values for each model are provided in the last two rows of the table. For the *i*th policyholder, the total amount y_i follows the specified model whose scale parameter σ_i depends on the above set of covariates through the aforementioned link function. In view of its relatively low *p*-value, the estimates associated with the explanatory variables *INTERCEPT*, *ATTORNEY*, *SEATBELT*, and *CLMAGE* are statistically significant at the 5% significance levels for the three models considered. In addition, the shape parameter θ is also statistically significant at the same nominal level for all the models discussed here. It is noted that the parameters of the mixing distribution are also significant at this nominal level. Finally, the SIG provides the best fit to this dataset in terms of the two measures of the model selection considered.

Table 5. Parameter estimates, standard errors (S.E.), and *p*-values for automobile bodily injury claims dataset under S, SG, and SIG regression models. NLL and AIC are included in the last two rows. The response variable is total losses.

Estimate (S.E.)	Regression Model		
	S	SG	SIG
<i>INTERCEPT</i>	−3.704 (1.041)	2.512 (0.019)	17.659 (2.713)
<i>p</i> -value	0.0004	<0.0001	<0.0001
<i>ATTORNEY</i>	1.651 (0.095)	2.374 (0.244)	20.014 (2.565)
<i>p</i> -value	<0.0001	<0.0001	<0.0001
<i>CLMSEX</i>	0.023 (0.087)	2.544 (0.293)	−0.092 (0.177)
<i>p</i> -value	0.7941	<0.0001	0.6022
<i>MARRIED</i>	−0.352 (0.272)	7.009 (0.233)	−0.054 (0.582)
<i>p</i> -value	0.1951	0.3210	0.9260
<i>SINGLE</i>	−0.434 (0.281)	2.363 (0.183)	−0.065 (0.615)
<i>p</i> -value	0.1222	<0.0001	0.9158
<i>WIDOWED</i>	−1.596 (0.488)	10.005 (14.472)	14.631 (2.536)
<i>p</i> -value	0.0011	0.4802	<0.0001
<i>CLMINSUR</i>	0.079 (0.146)	3.086 (0.714)	0.271 (0.365)
<i>p</i> -value	0.5898	<0.0001	0.4575
<i>SEATBELT</i>	−1.258 (0.361)	2.508 (0.043)	−19.1757 (2.7936)
<i>p</i> -value	0.0005	<0.0001	<0.0001
<i>CLMAGE</i>	0.019 (0.003)	0.393 (0.003)	0.0144 (0.0066)
<i>p</i> -value	<0.0001	<0.0001	0.0279
θ	0.736 (0.015)	1.647 (0.008)	1.061 (0.034)
<i>p</i> -value	<0.0001	<0.0001	<0.0001
λ, α, β	831.609 (528.599)	1.093 (0.039)	0.001 (0.000)
<i>p</i> -value	0.1160	<0.0001	0.0236
$-, \beta, \Gamma$	–	0.000 (0.000)	741.355 (104.096)
<i>p</i> -value	–	<0.0001	<0.0001
NLL	2608.38	2566.17	2558.82
AIC	5238.76	5156.34	5141.65

As the estimates of the parameter θ significantly differ in these models, we now estimate the value of the tail index, i.e., $1/\theta$ estimated by the commonly used Hill's estimator. For this, we consider the set of observations representing the n losses y_1, \dots, y_n and let $y_{(n)}, \dots, y_{(1)}$ be reordered (reversed order statistics) in such a way that $y_{(1)}$ is the highest value in the sample. Then, we construct the sets of numbers H_k and θ_k for $k > 2$. The set H_k is defined by $H_k = \frac{1}{k} \sum_{i=1}^k \log y_{(i)} - \log y_{(k)}$ and $\theta_k = 1/H_k$. Then, θ_k is an estimate of θ when k is increased until it seems inappropriate to proceed. Below, Figure 6 presents the values of θ_k when the value of k takes values in $\{2, \dots, n\}$. Note that the values of the estimator stabilise in the neighbourhood of 1.1. These values can be compared with the estimate of the tail indexes $1/\theta$ for each one of the three models given in Table 5. It would be interesting to study how the parameters' estimates of these three models vary if we fix the the estimate of the tail index based on the Hill's estimator and then estimate the other parameters via maximum likelihood estimation.

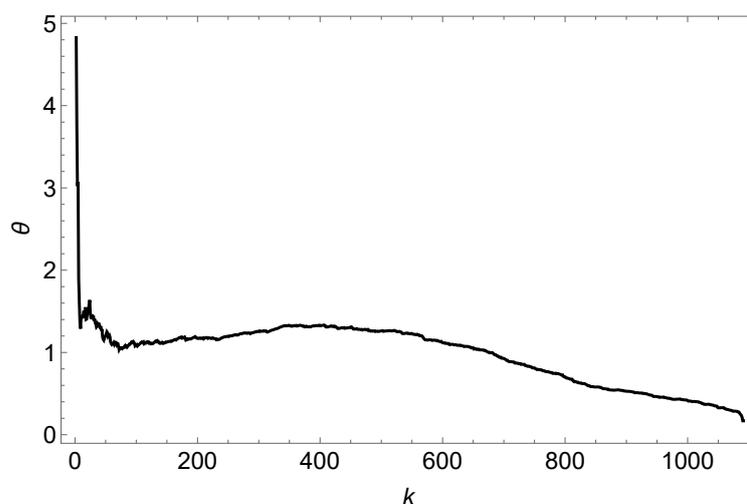


Figure 6. Hill's estimator θ_k values varying k , calculated for the automobile bodily injury claims dataset.

6. Conclusions

Although mixtures of discrete distributions have been largely considered in the statistical literature, the mixture of continuous distributions have been explored to a lesser extent in actuarial statistics. This paper considered properties and applications related to the mixture of probability distributions built by Lehmann's alternative method (the class of max-stable continuous distributions). In particular, a mixture involving the shape parameter of the Stoppa distribution was introduced in this article. The special case of the generalised inverse Gaussian was chosen as the mixing distribution and was thoroughly investigated in this work. Additionally, we provided properties related to the right-tail of the distribution that is of vital interest in risk theory. In addition, we examined the stochastic ordering of this new class of probability distributions. Also, a Bayesian analysis of this family was carried out, in which it was shown that the generalised inverse Gaussian distribution conjugates with the Stoppa distribution. Furthermore, for the case where the mixing distribution is Gamma, the mean of the posterior distribution can be written as a convex sum of a functional expression of the data and the mean of the prior distribution.

The parameters of these mixture models were estimated by the method of maximum likelihood via maximisation of the log-likelihood surface, and the models were empirically validated by using two well-known datasets in the actuarial literature. Moreover, two regression models based on this mixture were implemented for two particular cases of the mixing distribution: Gamma and inverse Gaussian distributions. These models were used to fit bodily injury claims data. Some covariates associated with insurance claimants were considered when building a modification of the logit link function for the scale parameter σ . Because of the results obtained, we have found that the proposed models in this paper are a valid alternative to other parametric fat-tailed models in the literature.

Author Contributions: Conceptualisation, E.G.-D. and E.C.-O.; methodology, E.G.-D. and E.C.-O.; software, E.G.-D. and E.C.-O.; validation, E.G.-D. and E.C.-O.; formal analysis, E.G.-D. and E.C.-O.; investigation, E.G.-D. and E.C.-O.; writing—original draft preparation, E.G.-D. and E.C.-O.; writing—review and editing, E.G.-D. and E.C.-O.; visualisation, E.G.-D. and E.C.-O.; supervision, E.G.-D. and E.C.-O.; project administration, E.G.-D. and E.C.-O.; funding acquisition, E.G.-D. and E.C.-O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by grant PID2021-127989OB-I00 (Agencia Estatal de Investigación, Ministerio de Ciencia e Innovación, Spain).

Data Availability Statement: Publicly available datasets were analyzed in this study.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Parameters Estimation for SG and SIG Distributions

We provide here the log-likelihood function and the normal equations from which the parameters of the SG and SIG distributions can be estimated.

Appendix A.1.1. SG Case

In this case the log-likelihood function is proportional to

$$\begin{aligned} \ell(\mathbf{\Omega}; \tilde{x}) \propto & n \left[\theta \log \sigma + \log \theta + \log \alpha + \frac{\alpha}{2} \log \beta \right] - \sum_{i=1}^n \log \varphi_{\sigma, \beta, \theta}(x_i) \\ & - (\theta + 1) \sum_{i=1}^n \log x_i - \left(\frac{\alpha}{2} + 1 \right) \sum_{i=1}^n \log \psi_{\sigma, \theta}(x_i), \end{aligned}$$

where $\mathbf{\Omega} = (\alpha, \beta, \theta)$. The mle of the SG distribution are obtained by solving the system of equations given by

$$\begin{aligned} \frac{n\alpha}{2\beta} - \left(\frac{\alpha}{2} + 1 \right) \sum_{i=1}^n \frac{1}{\varphi_{\sigma, \beta, \theta}(x_i)} &= 0, \\ n \left(\frac{1}{\theta} + \log \sigma \right) - \sum_{i=1}^n \log x_i + \sum_{i=1}^n \frac{(\sigma/x_i)^\theta}{\psi_{\sigma, \theta}(x_i)} \log \left(\frac{\sigma}{x_i} \right) \left[1 - \frac{\alpha + 2}{\varphi_{\sigma, \beta, \theta}(x_i)} \right] &= 0, \end{aligned}$$

where

$$\alpha = \left\{ \frac{1}{2} \left[\frac{1}{n} \sum_{i=1}^n \log \varphi_{\sigma, \beta, \theta}(x_i) - \log \beta \right] \right\}^{-1}.$$

Appendix A.1.2. SIG Case

$$\begin{aligned} \ell(\mathbf{\Omega}; \tilde{x}) \propto & n \left[\theta \log \sigma + \log \theta + \frac{1}{2} \log \Gamma + \sqrt{\Gamma \beta} \right] \\ & - (\theta + 1) \sum_{i=1}^n \log x_i - \frac{1}{2} \sum_{i=1}^n \log \varphi_{\sigma, \beta, \theta}(x_i) \\ & - \sum_{i=1}^n \log \psi_{\sigma, \theta}(x_i) - \sqrt{\Gamma} \sum_{i=1}^n \sqrt{\varphi_{\sigma, \beta, \theta}(x_i)}, \end{aligned}$$

where $\mathbf{\Omega} = (\beta, \theta, \Gamma)$.

The normal equations are given, in this case, by,

$$n\sqrt{\frac{\Gamma}{\beta}} - \sum_{i=1}^n \frac{1}{\varphi_{\sigma,\beta,\theta}(x_i)} - \sqrt{\Gamma} \sum_{i=1}^n (\varphi_{\sigma,\beta,\theta}(x_i))^{-1/2} = 0,$$

$$n\left(\frac{1}{\theta} + \log \sigma\right) + \sum_{i=1}^n (\sigma/x_i)^\theta \log(\sigma/x_i) \left\{ \frac{1}{\psi_{\sigma,\theta}(x_i)} \left[1 - \frac{1}{\varphi_{\sigma,\beta,\theta}(x_i)} \right. \right.$$

$$\left. \left. - \sqrt{\frac{\Gamma}{\varphi_{\sigma,\beta,\theta}(x_i)}} \right] \right\} - \sum_{i=1}^n \log x_i = 0,$$

where

$$\Gamma = \left\{ \frac{1}{n} \sum_{i=1}^n \sqrt{\varphi_{\sigma,\beta,\theta}(x_i)} - \sqrt{\beta} \right\}^{-2}.$$

Appendix A.2. EM Algorithm for the SG and SIG Regression Models

The estimation of the parameters in the mixture of Stoppa regression models can be carried out via an EM algorithm. When the model contains explanatory variables the E-step computes the posterior expectations of the sufficient statistics, i.e., $\sum \lambda_i$ and $\sum \log \lambda_i$, while at the M-step, firstly the posterior expectation is included in the first term of the right hand side of (A1) to fit an Stoppa regression model to update the regression coefficients and later, the parameter of the mixing distribution, i.e., Gamma or inverse gaussian, is updated.

Let us consider now the vector of complete data $\tilde{x} = (x_1, \dots, x_n)$ and the missing data $\tilde{\lambda} = (\lambda_1, \dots, \lambda_n)$ that contains the observed data and the missing data. By writing $\theta_i = \exp[\beta^\top \mathbf{z}_i]$, $i = 1, \dots, n$, the complete data log-likelihood takes the form

$$\ell_c(\beta^\top, \alpha; \tilde{x}, \tilde{\lambda}) = \sum_{i=1}^n \log f(x_i | \lambda_i, \beta^\top) + \sum_{i=1}^n \log f(\lambda_i | \omega), \tag{A1}$$

where $\mathbf{z}_i = (z_{i1}, \dots, z_{ip})$ is a vector of explanatory variables and $\beta^\top = (\beta_1, \dots, \beta_p)$ is a vector of regressors.

Appendix A.2.1. SG Regression Model

Here we consider the Gamma as mixing distribution. We will assume that $E(\Lambda) = 1$, i.e., $\alpha = \beta$. In this case, the EM scheme is given as follows. From the current estimates after the j -th iteration, $(\hat{\beta}^\top, \hat{\alpha}^{(j)})$, the new estimates are calculated as follows:

- at the E-step, calculate the pseudo-values

$$t_i = E\left(\Lambda_i | x_i, \hat{\theta}_i, \hat{\alpha}^{(j)}\right) = \frac{\alpha^{(j)} + 1}{\alpha^{(j)} - \log \psi_{\sigma, \hat{\theta}_i^{(j)}}(x_i)}$$

$$s_i = E\left(\log \Lambda_i | x_i, \hat{\theta}_i, \hat{\alpha}^{(j)}\right) = \Psi(1 + \hat{\alpha}^{(j)}) - \log \hat{\alpha}^{(j)} - \log(1 - \log \psi_{\sigma, \hat{\theta}_i^{(j)}}(x_i)),$$

where $\Psi(\cdot)$ is the diGamma function.

- at the M-step, first update the regressors $\hat{\beta}^\top$ by fitting an Stoppa regression model by including the covariates in the shape parameter as described above, by using the pseudo-values t_i and s_i . Then update the estimate of the parameter α by letting

$$\hat{\alpha}^{(j+1)} = \exp\left[\frac{\sum_{i=1}^n t_i - \sum_{i=1}^n s_i - n}{n} + \Psi(\hat{\alpha}^{(j)})\right].$$

- If some convergence condition is satisfied then stop iterating, otherwise move back to the E-step for another iteration.

Appendix A.2.2. SIG Regression Model

In this case, the EM scheme is given as follows. From the current estimates after the j -th iteration, $(\hat{\beta}^{\top(j)}, \hat{\Gamma}^{(j)})$, the new estimates are calculated as follows:

- at the E-step, calculate numerically the pseudo-values

$$\begin{aligned} m_i &= E\left(1/\Lambda_i|x_i, \hat{\theta}_i, \hat{\Gamma}^{(j)}\right); t_i = E\left(\Lambda_i|x_i, \hat{\theta}_i, \hat{\Gamma}^{(j)}\right) \text{ and} \\ s_i &= E\left(\log \Lambda_i|x_i, \hat{\theta}_i, \hat{\Gamma}^{(j)}\right). \end{aligned}$$

- at the M-step, first update the regressors $\hat{\beta}^{\top(j+1)}$ by fitting an Stoppa regression model by including the covariates in the shape parameter as described above, by using the pseudo-values t_i and s_i . Then update the estimate of the parameter Γ by letting

$$\hat{\Gamma}^{(j+1)} = \sqrt{\frac{n}{\sum_{i=1}^n t_i + \sum_{i=1}^n m_i - 2n}}.$$

- If some convergence condition is satisfied then stop iterating, otherwise move back to the E-step for another iteration.

The standard errors of the estimates $(\hat{\beta}^{\top(j)}, \hat{\Gamma}^{(j)})$ can be computed in the last iteration of the EM algorithm for these two models following the methodology in [Louis \(1982\)](#).

References

- Albrecher, Hansjörg, Jan Beirlant, and Jozef L. Teugels. 2017. *Reinsurance: Actuarial and Statistical Aspects*. Hoboken: John Wiley & Sons.
- Arnold, Barry. 1983. *Pareto Distributions*. Hoboken: Taylor & Francis Group.
- Barndorff-Nielsen, Ole, and Christian Halgreen. 1977. Infinite divisibility of the hyperbolic and generalised inverse gaussian distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 38: 309–11. [\[CrossRef\]](#)
- Bingham, Nicholas, Charles M. Goldie, and Jozef L. Teugels. 1989. *Regular Variation, Volume 27 of Encyclopedia of Mathematics and its Applications*. Cambridge: Cambridge University Press.
- Boland, Philip J. 2007. *Statistical and Probabilistic Methods in Actuarial Science*. Boca Raton: Chapman & Hall.
- Boyd, Albert V. 1988. Fitting the truncated Pareto distribution to loss distributions. *JSS* 31: 151–58. [\[CrossRef\]](#)
- Dhaene, John, Steven Vanduffel, Marc J. Goovaerts, Rob Kaas, Qihe Tang, and David Vyncke. 2006. Risk measures and comonotonicity: A review. *Stochastic Models* 22: 573–606. [\[CrossRef\]](#)
- Embrechts, Paul, Claudia Klüppelberg, and Thomas Mikosch. 1997. *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer.
- Frangos, Nikolaos, and Dimitris Karlis. 2004. Modelling losses using an exponential-inverse gaussian distribution. *Insurance: Mathematics and Economics* 35: 53–67.
- Frees, Edward W. 2010. *Regression Modelling with Actuarial and Financial Applications*. Cambridge: Cambridge University Press.
- Fung, Thomas, and Eugene Seneta. 2007. Tailweight, quantiles and kurtosis: A study of competing distributions. *Operations Research Letters* 35: 448–54. [\[CrossRef\]](#)
- Ghitany, Mohamed E., Emilio Gómez-Déniz, and Saralees Nadarajah. 2018. A new generalization of the Pareto distribution and its application to insurance data. *Journal of Risk and Financial Management* 11: 10. [\[CrossRef\]](#)
- Gómez-Déniz, Emilio, and Enrique Calderín-Ojeda. 2014. A suitable alternative to the Pareto distribution. *Hacettepe Journal of Mathematics and Statistics* 43: 843–60.
- Gómez-Déniz, Emilio, and Enrique Calderín-Ojeda. 2015. Modelling insurance data with the Pareto ArcTan distribution. *ASTIN Bulletin* 45: 639–60. [\[CrossRef\]](#)
- Gómez-Déniz, Emilio, Enrique Calderín-Ojeda, and José M. Sarabia. 2013. Gamma-generalized inverse Gaussian class of distributions with applications. *Communications in Statistics-Theory and Methods* 42: 919–33. [\[CrossRef\]](#)
- Gómez-Déniz, Emilio, Víctor Leiva, Enrique Calderín-Ojeda, and Christophe Chesneau. 2022. A novel claim size distribution based on a Birnbaum-Saunders and Gamma mixture capturing extreme values in insurance: Estimation, regression, and applications. *Computational and Applied Mathematics* 41: 171. [\[CrossRef\]](#)
- Hogg, Robert V., and Stuart A. Klugman. 1984. *Loss Distributions*. New York: John Wiley & Sons.
- Jessen, Anders H., and Thomas Mikosch. 2006. Regularly varying functions. *Publications de L'institut Mathématique* 80: 171–92. [\[CrossRef\]](#)

- Johnson, Norman, Adrienne W. Kemp, and Samuel Kotz. 2005. *Univariate Discrete Distributions*. Hoboken: John Wiley, Inc.
- Kleiber, Christian, and Samuel Kotz. 2003. *Statistical Size Distributions in Economics and Actuarial Sciences*. Hoboken: John Wiley & Sons, Inc.
- Konstantinides, Dimitrios G. 2017. *Risk Theory. A Heavy Tail Approach*. Singapore: World Scientific Publishing.
- Lehmann, Erich L. 1953. The power of rank test. *Annals of Mathematical Statistics* 24: 23–43. [[CrossRef](#)]
- Lemonte, Artur J., and Gauss M. Cordeiro. 2011. The exponentiated generalized inverse Gaussian distribution. *Statistics & Probability Letters* 81: 506–17.
- Louis, Thomas A. 1982. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society, Series B* 44: 226–33. [[CrossRef](#)]
- Marshall, Albert W., and Ingram Olkin. 1997. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* 84: 641–52. [[CrossRef](#)]
- McNeil, Alexander J. 1988. Estimating the tails of loss severity distribution using extreme value theory. *ASTIN Bulletin* 27: 117–37. [[CrossRef](#)]
- Rolski, Tomasz, Hanspeter Schmidli, Volker Schmidt, and Jozef L. Teugels. 1999. *Stochastic Processes for Insurance and Finance*. Hoboken: John Wiley & Sons.
- Sarabia, José M., and Enrique Castillo. 2005. About a class of max-stable families with applications to income distributions. *Metron* 63: 505–27.
- Sarabia, José M., and Faustino Prieto. 2009. The Pareto-positive stable distribution: A new descriptive model for city size data. *Physica A* 388: 4179–91. [[CrossRef](#)]
- Seshadri, Vanamamalai. 1994. *The Inverse Gaussian Distribution: A Case Study in Exponential Families*. Oxford: Science Publications.
- Shaked, Moshe, and J. George Shanthikumar. 2007. *Stochastic Orders*. Springer Series in Statistics. New York: Springer.
- Stoppa, Gabriele. 1990. *Proprietà Campionarie di un Nuovo Modello Pareto Generalizzato*. Padova: Cedam, pp. 137–44.
- Tzougas, George, and Dimitris Karlis. 2020. An EM algorithm for fitting a new class of mixed exponential regression models with varying dispersion. *ASTIN Bulletin* 50: 555–83. [[CrossRef](#)]
- Tzougas, George, and Himchan Jeong. 2021. An expectation-maximization algorithm for the exponential-generalized inverse gaussian regression model with varying dispersion and shape for modelling the aggregate claim amount. *Risks* 9: 19. [[CrossRef](#)]
- Yu, Yaming. 2009. Stochastic ordering of exponential family of distributions and their mixtures. *Journal of Applied Probability* 46: 244–54. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.