

Article

Scale Mixture of Gleser Distribution with an Application to Insurance Data

Neveka M. Olmos ¹, Emilio Gómez-Déniz ² and Osvaldo Venegas ^{3,*}

¹ Departamento de Estadística y Ciencias de Datos, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile; neveka.olmos@uantof.cl

² Department of Quantitative Methods in Economics, TIDES Institute, University of Las Palmas de Gran Canaria, 35017 Las Palmas de Gran Canaria, Spain; emilio.gomez-deniz@ulpgc.es

³ Departamento de Ciencias Matemáticas y Físicas, Facultad de Ingeniería, Universidad Católica de Temuco, Temuco 4780000, Chile

* Correspondence: ovenegas@uct.cl

Abstract: In this paper, the scale mixture of the Gleser (SMG) distribution is introduced. This new distribution is the product of a scale mixture between the Gleser (G) distribution and the Beta($a, 1$) distribution. The SMG distribution is an alternative to distributions with two parameters and a heavy right tail. We study its representation and some basic properties, maximum likelihood inference, and Fisher's information matrix. We present an application to a real dataset in which the SMG distribution shows a better fit than two other known distributions.

Keywords: Gleser distribution; heavy-tailed distribution; maximum likelihood; scale mixture

MSC: 62E15; 62E20



Citation: Olmos, N.M.; Gómez-Déniz, E.; Venegas, O. Scale Mixture of Gleser Distribution with an Application to Insurance Data. *Mathematics* **2024**, *12*, 1397. <https://doi.org/10.3390/math12091397>

Academic Editor: Jonas Šiaulyš

Received: 10 March 2024

Revised: 20 April 2024

Accepted: 22 April 2024

Published: 3 May 2024



Copyright: © 2024 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/>).

1. Introduction

The Pareto distribution (see Pareto [1]), named after Vilfredo Pareto, is a model widely used in various applied sciences, for example, actuarial sciences, economics, finance, life tests, climatology, and income distribution, in which the occurrence is generally described in extreme observations. Some investigators who have used the Pareto distribution are Beirlant et al. [2,3] and Resnick [4], among others. We say that a random variable X has a Pareto distribution if its probability density function (pdf) is given by

$$f_X(x; \alpha, \beta) = \frac{\beta \alpha^\beta}{x^{\beta+1}}, \quad x \geq \alpha \quad (1)$$

where $\alpha > 0$ is the scale parameter, and $\beta > 0$ is the shape parameter. We denote this as $X \sim \text{Pareto}(\alpha, \beta)$. The cumulative distribution function (cdf) corresponding to (1) is

$$F_X(x; \alpha, \beta) = 1 - \left(\frac{\alpha}{x}\right)^\beta, \quad x \geq \alpha \quad (2)$$

Johnson et al. [5] discuss various types of Pareto distributions, which differ from the Pareto density given in Equation (1). The density in Equation (1) is called Pareto Type I; however, in the present work we will refer to it simply as the Pareto distribution. Several extensions of the Pareto distribution can be found in the literature: Pickands [6] studied the generalized Pareto distribution and drew statistical inferences in the upper tail of a distribution function; Choulakian and Stephens [7] give tests of fit for the generalized Pareto distribution; additionally, Davison and Smith [8] discuss an application to hydrology, using river flow exceedances for a particular river over a period of 35 years. Aban et al. [9] list areas in which heavy-tailed distributions are considered applicable.

Gupta et al. [10], introduce a family of distributions which raises any cdf with positive support to a positive power. A particular case of this family of distributions is found when the cdf is represented by (2)—known as the exponentiated Pareto distribution. This distribution was first studied by Stoppa [11]; see also Kleiber and Kotz [12]. It is a particular case of the family studied by Gupta et al. [10]. Another generalization of the Pareto distribution is the beta-Pareto distribution introduced by Akinsete et al. [13]; Boumaraf et al. [14] applied various optimization methods to the beta-Pareto distribution. For further information on the Pareto distribution and its extensions, readers can consult the book by Arnold [15]. Gómez-Déniz and Calderín-Ojeda [16] study a distribution with one parameter which offers an alternative to the Pareto distribution; this distribution is considered a particular case of the log-gamma (LG) distribution, and in this paper, we denote it as LG2. We say that a random variable X has an LG2 distribution if its pdf is given by

$$f_X(x; \alpha, \beta) = \frac{\beta^2 \log(x + 1 - \alpha)}{(x + 1 - \alpha)^{\beta+1}}, \quad x \geq \alpha$$

where $\alpha \geq 0$ and $\beta \geq 0$. We denote this as $X \sim \text{LG2}(\alpha, \beta)$.

One function that is very necessary in this paper is the beta (B) function, which can be expressed as follows:

$$B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt = \int_0^\infty \frac{t^{a-1}}{(1+t)^{a+b}} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},$$

where $a > 0, b > 0$, and $\Gamma(\cdot)$ is the gamma function.

The beta function is the normalization constant of the beta distribution, i.e., we say that the random variable Y has a beta distribution with parameters a and b if its pdf is given by

$$f_Y(y; a, b) = \frac{1}{B(a, b)} y^{a-1}(1-y)^{b-1}, \quad 0 < y < 1$$

where $a > 0$ and $b > 0$.

The incomplete beta function is denoted as $B(y; a, b)$ and can be expressed as follows:

$$B(y; a, b) = \int_0^y t^{a-1}(1-t)^{b-1} dt, \quad 0 < y < 1$$

where $a > 0$ and $b > 0$. Another related function is the regularized incomplete beta function, denoted as $I_y(a, b)$, and expressed as $I_y(a, b) = B(y; a, b) / B(a, b)$.

The Gleser (G) distribution was introduced by Gleser [17] and studied recently by Olmos et al. [18]; we say that a random variable X has a G distribution if its pdf is given by

$$f_X(x; \sigma, \alpha) = \frac{\sigma^\alpha x^{-\alpha}}{B(\bar{\alpha}, \alpha)(\sigma + x)}, \quad x > 0$$

where $\sigma > 0$ is the scale parameter, $0 < \alpha < 1$ is the shape parameter, and $\bar{\alpha} = 1 - \alpha$. We denote this as $X \sim G(\sigma, \alpha)$. Some properties of this pdf are the following:

- (a) The G distribution has unimodality (at 0).
- (b) Let $Y \sim \text{Beta}(1 - \alpha, \alpha)$. Then, $\frac{\sigma Y}{1-Y} \sim G(\sigma, \alpha)$.
- (c) The cdf of X is given by

$$F_X(x; \sigma, \alpha) = I_y(\bar{\alpha}, \alpha), \quad x > 0$$

where $y = \frac{x}{\sigma+x}$ and $I_y(\cdot, \cdot)$ is the regularized incomplete beta function defined above.

Andrews and Mallows [19] developed a study of mixtures of scales of the normal distribution, and their investigation provides a family of distributions with heavy tails (e.g., the well-known Student's t distribution). These families of distributions have been used

in the robust inference of symmetrical data. A scale mixture distribution can be obtained by “mixing” a base density with a scale distribution. The resulting distribution can be expressed as follows:

$$f(x) = \int g(x|\kappa(w))dH(w)$$

where $g(x|\kappa(w))$ is the distribution conditional on the random variable X given that $\kappa(w)$ is a positive function of a random variable W with distribution function $H(w)$. For example, in the case of the Student’s t distribution, $W \sim \text{Gamma}(\nu/2, \nu/2)$ and $\kappa(w) = 1/w$. Another case that has been studied several times in recent years is the case of the slash distribution, where $W \sim \text{Uniform}(0, 1)$ and $\kappa(w) = w^{1/q}$.

The object of this paper is to study the SMG distribution, which is related to the G distribution and is a product of a scale mixture between the G and Beta distributions. The SMG distribution has two parameters and a heavy right tail, making it a possible alternative to the Pareto and LG2 distributions, among others, for modelling data with atypical observations.

The paper is organized as follows. In Section 2, we study the principal properties of the SMG distribution. In Section 3, we estimate the parameters using the maximum likelihood (ML) method, a simulation study, and the Fisher’s information matrix, which will be useful for calculating the asymptotic standard errors of the parameter estimates. In Section 4, we show an application to real data. In Section 5, we offer some conclusions.

2. SMG Distribution and Properties

In this section, we introduce the SMG distribution and study some of its properties.

2.1. Density

Let $Z \sim \text{SMG}(\sigma, \alpha)$, and assume that the pdf of Z is given by

$$f_Z(z; \sigma, \alpha) = \frac{\alpha \sigma^\alpha}{B(\alpha, \alpha)} z^{-(\alpha+1)} \log\left(1 + \frac{z}{\sigma}\right) \tag{3}$$

where $z > 0, \sigma > 0$ is the scale parameter, $0 < \alpha < 1$ is a shape parameter, and $B(\cdot, \cdot)$ is the beta function.

Figure 1 shows some plots of the SMG distribution considering different values of its parameters.

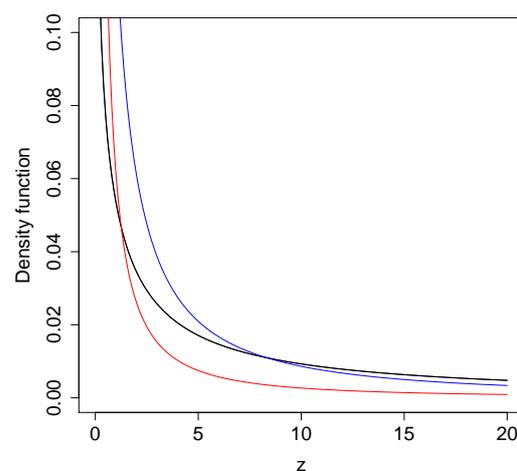


Figure 1. Examples of SMG(1, 0.3) (black), SMG(1, 0.7) (blue), and SMG(1, 0.9) (red) model.

Table 1 shows $P(Z > z)$ for different values of z in the SMG distribution, illustrating that the tails are heavier as the α parameter decreases.

Table 1. Tails comparison.

Distribution	$P(Z > 5)$	$P(Z > 10)$	$P(Z > 15)$
SMG(1, 0.9)	0.06493	0.04284	0.03310
SMG(1, 0.7)	0.26006	0.19377	0.16107
SMG(1, 0.3)	0.79235	0.73030	0.69215

2.2. Properties

The following propositions show one way of representing the SMG distribution. The first proposition shows that the SMG distribution is the product of a scale mixture between the G and Beta.

Proposition 1. Let $Z|U = u \sim G(\sigma u^{-1}, \alpha)$ and $U \sim \text{Beta}(\alpha, 1)$; then, $Z \sim \text{SMG}(\sigma, \alpha)$.

Proof. Calculating the integral directly, we have

$$f_Z(z; \sigma, \alpha) = \int_{\text{Rec}(u)} f(z|u) f_U(u) du = \int_0^1 \frac{(\sigma u^{-1})^\alpha z^{-\alpha}}{B(1 - \alpha, \alpha)(\sigma u^{-1} + z)} \alpha u^{\alpha-1} du,$$

the result of which is obtained with a change in variable $t = \sigma + uz$. □

Proposition 2. Let $X \sim G(\sigma, \alpha)$ and $Y \sim \text{Beta}(\alpha, 1)$ be independent random variables. Then, $Z = \frac{X}{Y} \sim \text{SMG}(\sigma, \alpha)$.

Proof. Using the Jacobian method, we have

$$\left. \begin{matrix} Z = XY^{-1} \\ V = Y \end{matrix} \right\} \Rightarrow \left. \begin{matrix} X = ZV \\ Y = V \end{matrix} \right\} \Rightarrow J = \begin{vmatrix} \frac{\partial X}{\partial Z} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial Z} & \frac{\partial Y}{\partial V} \end{vmatrix} = \begin{vmatrix} v & z \\ 0 & 1 \end{vmatrix} = v$$

$$\begin{aligned} f_{Z,V}(z, v) &= |J| f_{X,Y}(zv, v) \\ f_{Z,V}(z, v) &= v f_X(zv) f_Y(v), \quad z > 0, \quad 0 < v < 1 \end{aligned}$$

Then, marginalizing with respect to the variable V , we obtain the density function associated with Z :

$$\begin{aligned} f_Z(z; \sigma, \alpha) &= \frac{\alpha}{\sigma B(\bar{\alpha}, \alpha)} \int_0^1 \frac{\left(\frac{zv}{\sigma}\right)^{-\alpha} v^\alpha}{1 + \frac{zv}{\sigma}} dv \\ &= \frac{\alpha \sigma^{\alpha-1}}{B(\bar{\alpha}, \alpha) z^\alpha} \int_0^1 \frac{1}{1 + \frac{zv}{\sigma}} dv \end{aligned}$$

By making the change in variables $u = 1 + \frac{zv}{\sigma}$, the result is obtained. □

Two algorithms can be used to generate random numbers from the SMG model, as shown below. Algorithm 1 uses Proposition 1 and the composition method (see Tanner [20]); Algorithm 2 uses the representation given in Proposition 2.

Algorithm 1 for simulating $Z \sim \text{SMG}(\sigma, \alpha)$ can proceed as follows

- 1: Generate $U \sim \text{Beta}(\alpha, 1)$.
 - 2: Compute $Z|U = u \sim G(\sigma u^{-1}, \alpha)$.
-

Lemma 1. Let $Z \sim \text{SMG}(\sigma, \alpha)$. Then, we have the following:

1. $\mathbb{E}\left(\frac{Z}{(1+Z/\sigma)\log(1+Z/\sigma)}\right) = \sigma\alpha$
2. $\mathbb{E}\left(\frac{Z(2\sigma+Z)}{(1+Z/\sigma)^2\log(1+Z/\sigma)}\right) = \sigma^2\alpha$

Algorithm 2 for simulating $Z \sim SMG(\sigma, \alpha)$ can proceed as follows

- 1: Generate $V \sim Beta(1 - \alpha, \alpha)$.
- 2: Compute $X = \frac{\sigma V}{1-V}$.
- 3: Generate $Y \sim Beta(\alpha, 1)$.
- 4: Compute $Z = \frac{X}{Y}$.

Proof. Both results are obtained directly using the pdf given in (3). □

Proposition 3. Let $Z \sim SMG(\sigma, \alpha)$. Then, the cdf of Z is given by

$$F_Z(z; \sigma, \alpha) = I_y(\bar{\alpha}, \alpha) - \frac{\sigma^\alpha z^{-\alpha}}{B(\bar{\alpha}, \alpha)} \log\left(1 + \frac{z}{\sigma}\right), \quad z > 0 \tag{4}$$

where $\sigma > 0, 0 < \alpha < 1, y = \frac{z}{\sigma+z}$, and $I_y(\cdot, \cdot)$ is the regularized incomplete beta function.

Proof. Applying the cdf definition directly, we have

$$F_Z(z; \sigma, \alpha) = \frac{\sigma^\alpha}{B(\bar{\alpha}, \alpha)} \int_0^z t^{-(\alpha+1)} \log\left(1 + \frac{t}{\sigma}\right) dt$$

By applying the integration by parts method and considering $u = \log\left(1 + \frac{t}{\sigma}\right)$, the result is obtained. □

The survival function $s(t)$, which is the probability that an item will not fail before time t , is defined by $s(t) = 1 - F_X(t) = \bar{F}_X(t)$. The survival function for an SMG random variable is given by

$$\bar{F}_Z(t; \sigma, \alpha) = 1 - I_y(\bar{\alpha}, \alpha) + \frac{\sigma^\alpha t^{-\alpha}}{B(\bar{\alpha}, \alpha)} \log\left(1 + \frac{t}{\sigma}\right)$$

where $y = \frac{t}{\sigma+t}$. The hazards function $h(t)$, defined by $h(t) = \frac{f_X(t)}{\bar{F}_X(t)}$, for an SMG random variable is given by

$$h(t; \sigma, \alpha) = \frac{\alpha \sigma^\alpha t^{-(\alpha+1)} \log\left(1 + \frac{t}{\sigma}\right)}{B(\bar{\alpha}, \alpha) - B\left(\frac{t}{\sigma+t}; \bar{\alpha}, \alpha\right) + \sigma^\alpha t^{-\alpha} \log\left(1 + \frac{t}{\sigma}\right)}, \quad t > 0$$

Figure 2 shows the shape of the hazards function for different values of α , considering $\sigma = 1$.

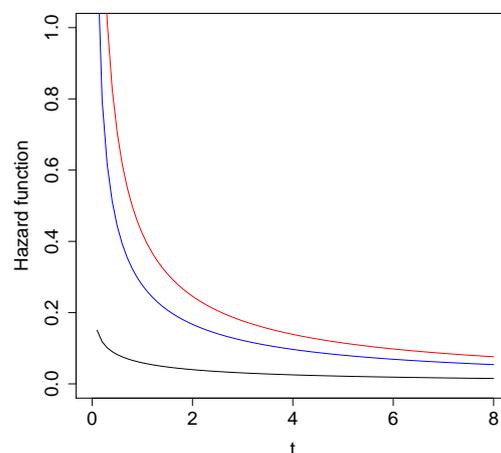


Figure 2. Examples of $h(t)$ for $\alpha = 0.3$ (black), $\alpha = 0.7$ (blue), and $\alpha = 0.9$ (red).

Proposition 4. Let $T \sim \text{SMG}(\sigma, \alpha)$. Then, the hazards function of T is decreasing for all $t > 0$.

Proof. Using the Theorem, item (b), given in Glaser [21], we have

$$\eta(t) = -\frac{f'(t)}{f(t)} = \frac{\alpha + 1}{t} - \frac{1}{(\sigma + t) \log(1 + \frac{t}{\sigma})}$$

where $f(t)$ is the pdf given in (3). Then, deriving the $\eta(t)$ function with respect to t , we have

$$\eta'(t) = -\left(\frac{\alpha + 1}{t^2} - \frac{1}{(\sigma + t)^2 \log(1 + \frac{t}{\sigma})} - \frac{1}{(\sigma + t)^2 \log^2(1 + \frac{t}{\sigma})} \right) < 0, \forall t > 0$$

and the result is concluded. \square

Figure 3 shows, by fixing $\sigma = 1$, that the $\eta'(t)$ function is always negative. For other different values of σ and α , the result is the same.

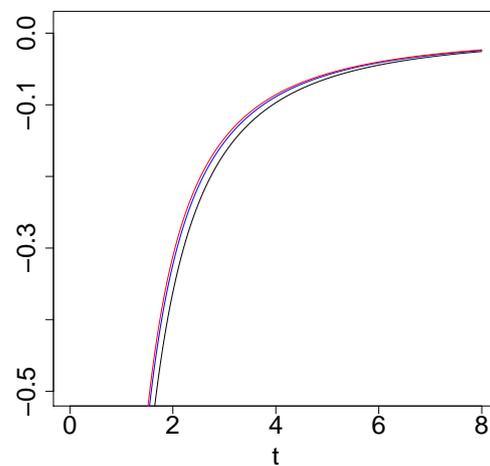


Figure 3. Examples of $\eta'(t)$ for $\alpha = 0.3$ (black), $\alpha = 0.7$ (blue), and $\alpha = 0.9$ (red).

Proposition 5. The r th moments of the random variable $Z \sim \text{SMG}(1, \alpha)$ do not exist for $r \geq \alpha$.

Proof. Using integration by parts, with $u = \log(1 + x)$, we have

$$\begin{aligned} \mathbb{E}(Z^r) &= \frac{\alpha \sigma^\alpha}{B(\alpha, \bar{\alpha})} \int_0^\infty z^{r-\alpha-1} \log(1+z) dz \\ &= \frac{\alpha \sigma^\alpha z^{r-\alpha} \log(1+z)}{(r-\alpha)B(\bar{\alpha}, \alpha)} \Big|_0^\infty - \frac{\alpha \sigma^\alpha}{(r-\alpha)B(\bar{\alpha}, \alpha)} \int_0^\infty \frac{z^{r-\alpha}}{1+z} dz \end{aligned}$$

The expression on the left side is equal to zero if $r < \alpha$ and does not exist if $r \geq \alpha$; the integral on the right side was analyzed in Proposition 1 (e) of the work of Olmos et al. [18] and does not converge. Therefore, the complete integral diverges, and this means that no r th moments exist for the random variable $Z \sim \text{SMG}(1, \alpha)$. \square

2.3. Tail of the Distribution

The Pareto model is a distribution with heavy right tail. Based on this characteristic, it is used in insurance to model the number of losses; hence, the size of the tail of the distribution is fundamental if the chosen model is to capture observations very distant from the start of the distribution support, i.e., extreme values. The concept of heavy tail is fundamental for this and other financial scenarios. The use of distributions with a heavy

right tail is of vital importance in economics, finance, and natural disasters. Pareto, EP, and log normal distributions, among others, have been used to model losses in insurance, reinsurance, and catastrophic insurance. It is known that any distribution of probability, specified by its cdf $F(x)$ on the real line, has a heavy right tail (see Rolski et al. [22]) if $\limsup_{x \rightarrow \infty} (-\log(s(x))/x) = 0$.

An important subject in the theory of extreme values is regular variation (see Bingham [23]). This concept is formalized in the following definition.

Definition 1. A distribution function is called regular varying at infinity with index $-\alpha$ if

$$\lim_{x \rightarrow \infty} \frac{s(tx)}{s(x)} = t^{-\alpha}$$

where the parameter $\alpha \geq 0$ is called the tail index.

The following proposition establishes that the survival function of the SMG distribution is a distribution with regular variation.

Proposition 6. The survival function of the random variable $X \sim \text{SMG}(\sigma, \alpha)$ is a survival function with regularly varying tails.

Proof. Applying the above definition and using L'Hospital's Rule, we have

$$\lim_{x \rightarrow \infty} \frac{s(tx)}{s(x)} = \lim_{x \rightarrow \infty} \frac{1 - F(tx)}{1 - F(x)} = t^{-\alpha} \lim_{x \rightarrow \infty} \frac{\log(1 + \frac{tx}{\sigma})}{\log(1 + \frac{x}{\sigma})} = t^{-\alpha+1} \lim_{x \rightarrow \infty} \frac{\sigma + x}{\sigma + tx}$$

and by calculating the latter limit, the result is obtained. \square

3. Inference

In this section, we estimate the parameters of the SMG model using the ML method; a simulation study and asymptotic estimation of the ML estimators are discussed.

3.1. ML Estimation

For a random sample z_1, \dots, z_n derived from the $\text{SMG}(\sigma, \alpha)$ distribution, the log-likelihood function can be written as

$$\ell(\sigma, \alpha) = n \log(\alpha) + n\alpha \log(\sigma) - n \log(B(\bar{\alpha}, \alpha)) - (\alpha + 1) \sum_{i=1}^n \log(z_i) + \sum_{i=1}^n \log(\log(1 + z_i/\sigma)) \tag{5}$$

Setting the first derivatives equal to zero with respect to each parameter, we have

$$\sum_{i=1}^n \frac{z_i}{(\sigma + z_i) \log(1 + z_i/\sigma)} = n\alpha \tag{6}$$

$$\sum_{i=1}^n \log(z_i) + n\psi(\alpha) - n\psi(1 - \alpha) = \frac{n}{\alpha} + n \log(\sigma)$$

where $\psi(\cdot)$ is the digamma function. From (6), we obtain

$$\hat{\alpha}(\hat{\sigma}) = \frac{1}{n} \sum_{i=1}^n \frac{z_i}{(\hat{\sigma} + z_i) \log(1 + z_i/\hat{\sigma})} \tag{7}$$

and the maximum likelihood estimator for σ ($\hat{\sigma}$) is obtained by numerically resolving the following equation:

$$\frac{1}{n} \sum_{i=1}^n \log(z_i) - \psi(1 - \hat{\alpha}(\hat{\sigma})) + \psi(\hat{\alpha}(\hat{\sigma})) = \frac{n}{\hat{\alpha}(\hat{\sigma})} + n \log(\hat{\sigma}) \tag{8}$$

The estimator $\hat{\sigma}$ is the solution to Equation (8), and replacing it in (7), we obtain $\hat{\alpha}$. The solution for Equation (8) can be obtained by using numerical procedures such as the Newton–Raphson algorithm. Alternatively, these estimates can be found by directly maximizing the log-likelihood surface given by (5) and using the `optim` subroutine in the R software package [24].

3.2. Simulation Study

To evaluate the effectiveness of the proposed approach, we conducted a simulation study to assess the performance of the estimation procedure for the parameters α and σ in the SMG model. The study involved simulating 1000 samples from the SMG model with three different sample sizes: $n = 100, 200,$ and 300 . Table 2 shows the median of the bias estimate for each parameter (Bias), its standard errors (SE), and the estimated root of the mean squared error (RMSE); the empirical coverage probabilities (CPs) at 95% for the asymptotic intervals based on ML estimators are given. From Table 2, we conclude that the ML estimates are quite stable. The bias is reasonable and reduced when the sample size is increased. The empirical CP also comes closer to the nominal 95% as n increases.

Table 2. ML estimations for parameters α and σ of the SMG distribution.

True Value		Estimator	$n = 100$				$n = 200$				$n = 300$				
σ	α		Bias	SE	RMSE	CP	Bias	SE	RMSE	CP	Bias	SE	RMSE	CP	
1	0.1	α	0.0068	0.0093	0.0116	0.9190	0.0071	0.0066	0.0097	0.8460	0.0062	0.0053	0.0083	0.7990	
		σ	0.4032	2.2812	4.0426	0.9350	0.3730	1.2368	1.6858	0.9640	0.3860	0.8955	1.1496	0.9710	
	0.25	α	0.0071	0.0093	0.0120	0.9250	0.0065	0.0065	0.0093	0.8620	0.0065	0.0053	0.0083	0.7960	
		σ	4.0440	22.4201	34.7855	0.9390	3.5141	11.7503	15.3452	0.9600	3.5002	8.8431	11.3256	0.9780	
	0.5	α	0.0030	0.0265	0.0273	0.9630	0.0012	0.0185	0.0194	0.9470	−0.0002	0.0149	0.0153	0.9420	
		σ	0.0919	0.9833	1.2023	0.9110	0.0365	0.5801	0.6274	0.9300	0.0287	0.4429	0.4736	0.9320	
	0.75	α	0.0017	0.0266	0.0285	0.9580	0.0018	0.0184	0.0185	0.9570	0.0011	0.0150	0.0158	0.9550	
		σ	1.4353	10.1717	12.4871	0.9090	0.4046	5.7823	6.4739	0.9360	0.2705	4.5401	5.3019	0.9390	
	0.9	α	0.0000	0.0645	0.0663	0.9190	0.0006	0.0470	0.0486	0.9290	−0.0009	0.0389	0.0378	0.9390	
		σ	0.0090	1.3658	2.1311	0.8800	0.0027	0.7859	0.9812	0.8930	−0.0025	0.5698	0.6251	0.9090	
	10	0.1	α	0.0031	0.0643	0.0629	0.9210	0.0015	0.0471	0.0488	0.9270	0.0013	0.0388	0.0387	0.9390
			σ	0.2496	12.7439	24.3474	0.8950	0.4281	8.0812	10.7496	0.9060	0.3605	5.9643	6.8460	0.9230
0.25		α	−0.0026	0.0381	0.0423	0.9560	−0.0023	0.0262	0.0271	0.9540	−0.0017	0.0213	0.0224	0.9590	
		σ	−0.0281	0.6856	0.7240	0.8880	−0.0512	0.4548	0.4745	0.9130	−0.0301	0.3663	0.3837	0.9210	
0.5		α	−0.0033	0.0381	0.0420	0.9450	−0.0016	0.0261	0.0265	0.9470	−0.0011	0.0211	0.0217	0.9530	
		σ	−0.7347	6.8163	7.3233	0.8840	−0.3695	4.6192	4.8007	0.9290	−0.2962	3.6364	3.8096	0.9060	
0.75		α	0.0001	0.0116	0.0122	0.9550	−0.0005	0.0081	0.0082	0.9470	−0.0005	0.0066	0.0069	0.9470	
		σ	−0.0263	0.7099	0.7945	0.8810	−0.0256	0.4557	0.4643	0.9150	−0.0227	0.3672	0.3941	0.9230	
0.9		α	−0.0016	0.0117	0.0120	0.9600	−0.0002	0.0081	0.0083	0.9400	−0.0004	0.0066	0.0066	0.9480	
		σ	−0.7632	6.7365	7.6937	0.8640	−0.4858	4.5617	4.9562	0.9070	−0.1317	3.6848	3.8046	0.9260	

3.3. Fisher’s Information Matrix

Let us now consider $Z \sim SMG(\sigma, \alpha)$. For a single observation z of Z , the log-likelihood function for $\theta = (\sigma, \alpha)$ is given by

$$\ell(\theta) = \log(\alpha) + \alpha \log(\sigma) - \log(B(\bar{\alpha}, \alpha)) - (\alpha + 1) \log(z) + \log(\log(1 + z/\sigma))$$

The corresponding first and second partial derivatives of the log-likelihood function are derived in Appendix A. It can be shown that the Fisher’s information matrix, denoted by $I_F(\cdot)$, for the G distribution is provided by

$$I_F(\theta) = \begin{pmatrix} \frac{\alpha \alpha_\alpha}{\sigma^2 B(\bar{\alpha}, \alpha)} & -\frac{1}{\sigma} \\ -\frac{1}{\sigma} & \psi'(\alpha) + \psi'(\bar{\alpha}) + \frac{1}{\alpha^2} \end{pmatrix}$$

where $\psi'(\cdot)$ is the trigamma function and $a_\alpha = \int_0^\infty \frac{u^{1-\alpha}}{(1+u)^2 \log(1+u)} du$.

Hence, for large samples, the ML estimator, $\hat{\theta}$, of θ is asymptotically normal bivariate, that is, the following is true;

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{\mathcal{L}} N_2(\mathbf{0}, I_F(\theta)^{-1})$$

the asymptotic variance of the ML estimator $\hat{\theta}$ is, therefore, the inverse of Fisher’s information matrix $I_F(\theta)$. Since the parameters are unknown, the observed information matrix is usually considered, where the unknown parameters are estimated by ML.

4. Application with Real Data

In this section, we analyze an application to a real dataset, comparing the Pareto and LG2 distributions with the fit of the SMG distribution. To compare the models, we use the Akaike information criterion, AIC (see Akaike [25]), and the Bayesian information criterion, BIC (see Schwarz [26]). The dataset represents accident indices (annual data in billions of USD) for earthquake insurance in California from 1971 to 1993 for values greater than zero. The data are given in Embrechts et al. [27]. The descriptive statistics of these data are shown in Table 3, where CS is the coefficient of asymmetry of the sample, and CK is the coefficient of kurtosis of the sample. The CK is quite high, and this shows the presence of atypical observations. The boxplot in Figure 4 shows the existence of two very extreme data. These atypical data make the right tail heavier.

Table 3. Descriptive statistics for loss ratios data.

n	Median	Mean	Variance	CS	CK
19	5.0	15.72	889.582	3.204	12.648

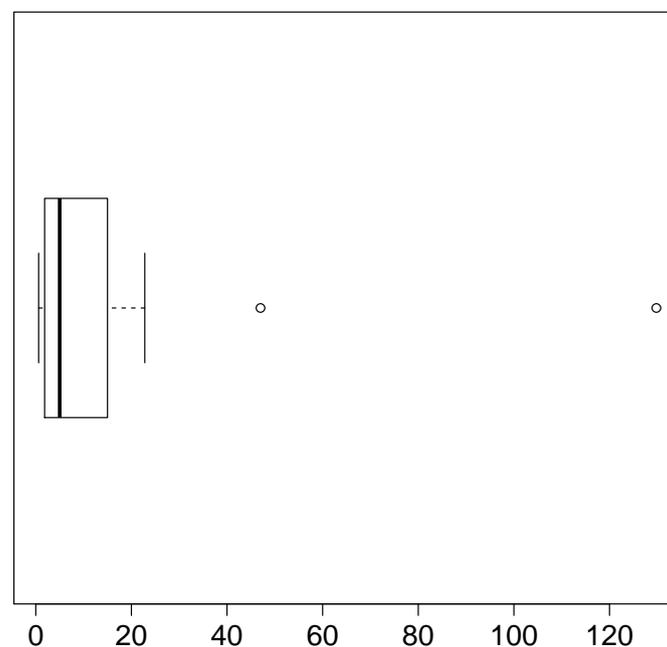


Figure 4. Boxplot for of loss ratios data.

Table 4 shows the ML estimates together with standard errors (in brackets) for the parameters of the SMG, Pareto, and LG2 models, as well as the values of the AIC and BIC criteria for each model. For the Pareto and LG2 models, the value $\alpha = 0.1$ was used.

Table 4. Loss ratios data: model, ML estimates, and AIC and BIC values.

Model	ML Estimates	AIC	BIC
SMG(σ, α)	$\hat{\sigma} = 2.0190(0.793)$, $\hat{\alpha} = 0.573(0.069)$	151.815	153.704
Pareto(α, β)	$\hat{\beta} = 0.249(0.057)$	157.878	158.822
LG2(α, β)	$\hat{\beta} = 0.688(0.112)$	176.332	177.277

We observe that the smallest values of the AIC and BIC criteria correspond to the SMG model, meaning that the SMG model fits the data better than the Pareto and LG2 models. The SEs of the ML of the SMG model were calculated using the method described in Section 3.3.

Table 5 shows three goodness-of-fit tests for all the distributions considered; this shows that the classic Pareto distribution is rejected for this dataset.

Table 5. Test statistics (*p*-values) of goodness-of-fit tests of the considered models.

Model	Kolmogorov–Smirnov	Anderson–Darling	Cramér–Von Misses
SMG(σ, α)	0.238(0.197)	0.215(0.823)	0.033(0.783)
Pareto(α, β)	0.360(0.010)	3.574 (0.014)	0.706(0.012)
LG2(α, β)	0.255(0.142)	0.239(0.746)	0.039(0.675)

We observe that the goodness-of-fit tests indicate that the SMG and LG2 distributions are appropriate for fitting these data; the selection criteria of the AIC and BIC models indicate that the SMG distribution provides a better fit than the other two. For this reason, the SMG distribution may be considered as an alternative to the classic Pareto distribution. Figure 5 shows the Q–Q plot of the quantile residuals (QRs) of the SMG distribution (see Dunn and Smyth [28]).

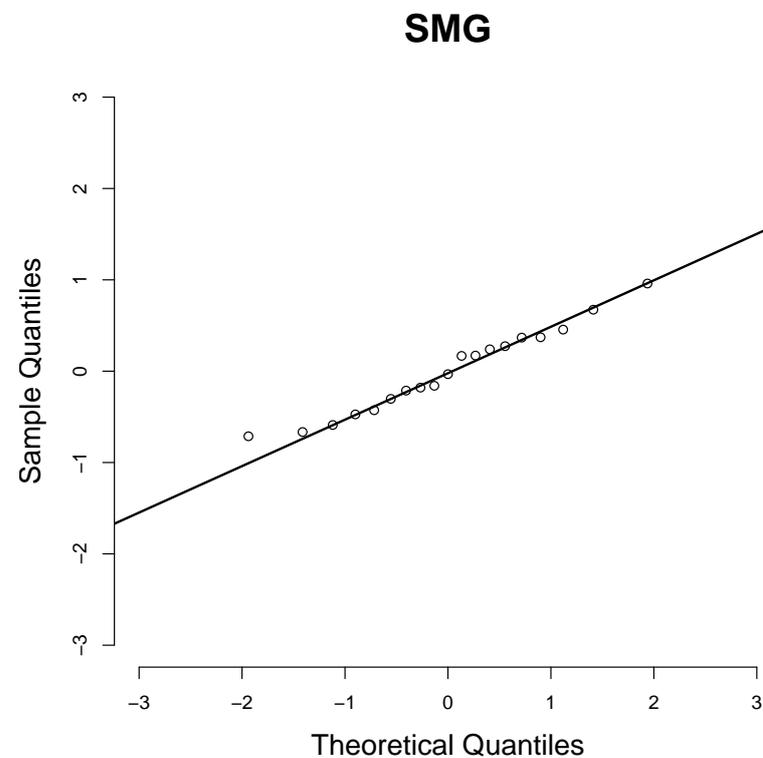


Figure 5. Q–Q plots of the QRs for SMG distribution.

5. Conclusions

This paper presents a study of the SMG distribution; we study some of its properties, estimate the parameters using the ML method, and analyze an application to real data. The SMG distribution is the product of a scale mixture of the G and Beta distributions. This distribution has two parameters, making it interesting as a competitor with several two-parameter models used, for example, in actuarial statistics. The SMG model appears to be a viable alternative for fitting data with atypical observations. Some other characteristics of the SMG model are as follows:

- The right tail of the SMG model is heavier for small values of parameter α .
- There are representations of the SMG model, given in Propositions 1 and 2.
- The SMG distribution has a heavy right tail, see Proposition 6
- The pdf, cdf, and hazard function are explicit and are represented by known functions.
- The analysis of the application shows that the SMG distribution is a good candidate for modeling loss ratios data, performing better than two well-known distributions, the Pareto and LG2 distributions.

Author Contributions: Conceptualization, N.M.O. and E.G.-D.; methodology, N.M.O. and E.G.-D.; software, N.M.O. and E.G.-D.; validation, N.M.O., E.G.-D., and O.V.; formal analysis, N.M.O. and O.V.; investigation, E.G.-D. and N.M.O.; writing—original draft preparation, N.M.O. and E.G.-D.; writing—review and editing, O.V. and E.G.-D.; funding acquisition, O.V. All authors have read and agreed to the published version of the manuscript.

Funding: The research work by N.M. Olmos was supported by Semillero UA-2024.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data set is available in the text.

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

Appendix A

The first derivatives of $\ell(\theta)$ are given by

$$\begin{aligned} \frac{\partial \ell(\theta)}{\partial \sigma} &= \frac{\alpha}{\sigma} - \frac{z}{\sigma^2(1+z/\sigma) \log(1+z/\sigma)} \\ \frac{\partial \ell(\theta)}{\partial \alpha} &= \frac{1}{\alpha} + \log(\sigma) + \psi(\bar{\alpha}) - \psi(\alpha) - \log(z) \end{aligned}$$

The second derivatives of $l(\theta)$ are

$$\begin{aligned} \frac{\partial^2 \ell(\theta)}{\partial \sigma^2} &= -\frac{\alpha}{\sigma^2} + \frac{z(2\sigma+z)}{\sigma^2(\sigma+z)^2 \log(1+z/\sigma)} - \frac{z^2}{\sigma^2(\sigma+z)^2 \log^2(1+z/\sigma)} \\ \frac{\partial^2 \ell(\theta)}{\partial \sigma \partial \alpha} &= \frac{1}{\sigma} \\ \frac{\partial^2 \ell(\theta)}{\partial \alpha^2} &= -\frac{1}{\alpha^2} - \psi'(\bar{\alpha}) - \psi'(\alpha) \end{aligned}$$

where $\psi(\cdot)$ and $\psi'(\cdot)$ are the digamma and trigamma functions, respectively.

References

1. Pareto, V. *Cours d'Économie Politique*; Librairie Droz: Lausanne, Switzerland, 1897.
2. Beirlant, J.; Teugels, J.L.; Vynckier, P. *Practical Analysis of Extreme Values*; Leuven University Press: Leuven, Belgium, 1996.
3. Beirlant, J.; Joossens, E.; Segers, J. Generalized Pareto fit to the society of actuaries' large claims database. *N. Am. Actuar. J.* **2004**, *8*, 108–111. [[CrossRef](#)]

4. Resnick, S.I. Discussion of the Danish data on large fire insurance losses. *ASTIN Bull.* **1997**, *27*, 139–151. [[CrossRef](#)]
5. Johnson, N.L.; Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions*, 2nd ed.; Wiley Series in Probability and Statistics; Wiley: New York, NY, USA, 1995.
6. Pickands, J. Statistical inference using extreme order statistics. *Ann. Stat.* **1975**, *3*, 119–131.
7. Choulakian, V.; Stephens, M.A. Goodness-of-fit for the generalized Pareto distribution. *Technometrics* **2001**, *43*, 478–484. [[CrossRef](#)]
8. Davison, A.C.; Smith, R.L. Models for Exceedances Over High Thresholds (with comments). *J. R. Stat. Soc. Ser. B* **1990**, *52*, 393–442. [[CrossRef](#)]
9. Aban, I.B.; Meerschaert, M.M.; Panorska, A.K. Parameter estimation for the truncated Pareto distribution. *J. Am. Statist. Assoc.* **2006**, *101*, 270–277. [[CrossRef](#)]
10. Gupta, R.C.; Gupta, R.D.; Gupta, P.L. Modeling failure time data by Lehman alternatives. *Commun. Stat. Theory Methods* **1998**, *27*, 887–904. [[CrossRef](#)]
11. Stoppa, G. Proprieta campionarie di un nuovo modello Pareto generalizzato. In Proceedings of the Atti XXXV Riunione Scientifica della Societa Italiana di Statistica, Padova, Italy, 18–21 April 1990; Cedam: Padova, Italy, 1990; pp. 137–144.
12. Kleiber, C.; Kotz, S. *Statistical Size Distributions in Economics and Actuarial Sciences*; John Wiley & Sons: Hoboken, NJ, USA, 2003.
13. Akinsete, A.; Famoye, F.; Lee, C. The Beta-Pareto distribution. *Statistics* **2008**, *42*, 547–563. [[CrossRef](#)]
14. Boumaraf, B.; Seddik-Ameur, N.; Barbu, V.S. Estimation of Beta-Pareto Distribution Based on Several Optimization Methods. *Mathematics* **2020**, *8*, 1055. [[CrossRef](#)]
15. Arnold, B.C. Pareto Distributions. In *Monographs on Statistics & Applied Probability*, 2nd ed.; Chapman & Hall: Boca Raton, FL, USA, 2015.
16. Gómez-Déniz, E.; Calderín-Ojeda, E. A suitable alternative to the Pareto distribution. *Hacet. J. Math. Stat.* **2014**, *43*, 843–860.
17. Gleser, L.J. The Gamma Distribution as a Mixture of Exponential Distributions. *Am. Stat.* **1989**, *43*, 115–117. [[CrossRef](#)]
18. Olmos, N.M.; Gómez-Déniz, E.; Venegas, O. The Heavy-Tailed Gleser Model: Properties, Estimation, and Applications. *Mathematics* **2022**, *10*, 4577. [[CrossRef](#)]
19. Andrews, D.F.; Mallows, C.L. Scale mixtures of normal distributions. *J. R. Stat Soc. Ser. B* **1974**, *36*, 99–102. [[CrossRef](#)]
20. Tanner, M.A. Tools for statistical inference. In *Methods for the Exploration of Posterior Distributions and Likelihood Functions*, 3rd ed.; Springer: New York, NY, USA, 1996.
21. Glaser, R.E. Bathtub and Related Failure Rate Characterizations. *J. Am. Stat. Assoc.* **1980**, *75*, 667–672. [[CrossRef](#)]
22. Rolski, T.; Schmidli, H.; Schmidt, V.; Teugel, J. *Stochastic Processes for Insurance and Finance*; John Wiley & Sons: Hoboken, NJ, USA, 1999.
23. Bingham, N. *Regular Variation*; Cambridge University Press: Cambridge, UK, 1987.
24. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021. Available online: <https://www.R-project.org/> (accessed on 15 January 2024).
25. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Automat. Cont.* **1974**, *19*, 716–723. [[CrossRef](#)]
26. Schwarz, G. Estimating the dimension of a model. *Ann. Statist.* **1978**, *6*, 461–464. [[CrossRef](#)]
27. Embrechts, P.; Sidney, I.R.; Gennady, S. Extreme value theory as a risk management tool. *N. Am. Actuar. J.* **1999**, *3*, 30–41. [[CrossRef](#)]
28. Dunn, P.K.; Smyth, G.K. Randomized Quantile Residuals. *J. Comput. Graph. Stat.* **1996**, *5*, 236–244. [[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.