Journal of
Risk and Financial Management

MDPI

Article

A New Generalization of the Pareto Distribution and Its Application to Insurance Data

Mohamed E. Ghitany [1,*], Emilio Gómez-Déniz [2] and Saralees Nadarajah [3]

[1] Department of Statistics and Operations Research, Faculty of Science, Kuwait University, Safat 13060, Kuwait
[2] Department of Quantitative Methods and T*i*DES Institute, University of Las Palmas de Gran Canaria, 35017 Gran Canaria, Spain; emilio.gomez-deniz@ulpgc.es
[3] School of Mathematics, University of Manchester, Manchester M13 9PL, UK; Saralees.Nadarajah@manchester.ac.uk
* Correspondence: meghitany@yahoo.com

Received: 26 November 2017; Accepted: 2 February 2018; Published: 7 February 2018

Abstract: The Pareto classical distribution is one of the most attractive in statistics and particularly in the scenario of actuarial statistics and finance. For example, it is widely used when calculating reinsurance premiums. In the last years, many alternative distributions have been proposed to obtain better adjustments especially when the tail of the empirical distribution of the data is very long. In this work, an alternative generalization of the Pareto distribution is proposed and its properties are studied. Finally, application of the proposed model to the earthquake insurance data set is presented.

Keywords: gamma distribution; estimation; financial risk; fit; pareto distribution

MSC: 62E10; 62F10; 62F15; 62P05

1. Introduction

In general insurance, only a few large claims arising in the portfolio represent the largest part of the payments made by the insurance company. Appropriate estimation of these extreme events is crucial for the practitioner to correctly assess insurance and reinsurance premiums. On this subject, the single parameter Pareto distribution (Arnold 1983; Brazauskas and Serfling 2003; Rytgaard 1990), among others has been traditionally considered as a suitable claim size distribution in relation to rating problems. Concerning this, the single parameter Pareto distribution, apart from its favourable properties, provides a good depiction of the random behaviour of large losses (e.g., the right tail of the distribution). Particularly, when calculating deductibles and excess–of–loss levels for reinsurance, the simple Pareto distribution has been demonstrated convenient, see for instance (Boyd 1988; Mata 2000; Klugman et al. 2008), among others.

In this work, an alternative to the Pareto distribution will be carried out. Properties and applications of this distribution will be studied here. As far as we know, these properties have not been studied for this distribution. In particular, we concentrate our attention to results connected with financial risk and insurance.

The paper is organized as follows. In Section 2, the new proposed distribution is shown, including some of its more relevant properties. Section 3 presents some interesting results connecting with financial risk and insurance. Next, Section 4 deals with parameter estimation, paying special attention to the maximum likelihood method. In Section 5, numerical application by using real insurance data is considered. Finally, some conclusions are given in the last section.

2. The Proposed Distribution

2.1. Probability Density Function

A continuous random variable X is said to have a generalized truncated log-gamma (GTLG) distribution if its probability density function (p.d.f.) is given by

$$f(x) = \frac{\theta^\lambda}{\alpha\Gamma(\lambda)} \left(\frac{x}{\alpha}\right)^{-\theta-1} \left(\log\frac{x}{\alpha}\right)^{\lambda-1}, \qquad x \geq \alpha, \quad \alpha,\theta,\lambda > 0, \tag{1}$$

where $\Gamma(z) = \int_0^\infty t^{z-1}\exp(-t)\,dt$ is the Euler gamma function. Note that, for all $\alpha,\theta > 0$, we have

$$f(\alpha) = \begin{cases} \infty, & \text{if } 0 < \lambda < 1, \\ \theta/\alpha, & \text{if } \lambda = 1, \\ 0, & \text{if } \lambda > 1, \end{cases} \qquad\qquad f(\infty) = 0.$$

As it can be easily seen, the parameter α marks a lower bound on the possible values that (1) can take on. When $\alpha = 1$, the GTLG distribution reduced to the log-gamma distribution proposed by Consul and Jain (1971) with p.d.f.

$$f_Z(z) = \frac{\theta^\lambda}{\Gamma(\lambda)} z^{-\theta-1}(\log z)^{\lambda-1}, \qquad z > 1, \quad \theta,\lambda > 0.$$

Note that Consul and Jain (1971) considered only the case $\lambda \geq 1$. For this case, they derived the raw moments and the distribution of the product of two independent log-gamma random variables. The p.d.f. (1) can now be obtained by the transformation $X = \alpha Z$.

Expression (1) is a particular case of the generalized truncated log–gamma distribution proposed in Amini et al. (2014) and related with the family proposed by Zografos and Balakrishnan (2009). When $\lambda = 1$, we obtain the famous Pareto distribution. In addition, when $\lambda = 2$, we obtain a distribution reminiscent of the distribution proposed in Gómez-Déniz and Calderín (2014). Properties and applications of this distribution will be studied here. In particular, we concentrate attention to results connecting with financial risk and insurance.

Theorem 1. *For all* $\alpha,\theta > 0$, $f(x)$ *is decreasing (increasing-decreasing) if* $0 < \lambda \leq 1$ $(\lambda > 1)$.

Proof. The first derivative of $f(x)$ given by

$$f'(x) = \left[-(\theta+1) + \frac{\lambda-1}{\log(x/\alpha)}\right]\frac{f(x)}{x},$$

which can be seen to be strictly negative if $0 < \lambda \leq 1$ and has a unique zero at $x_m = \alpha\exp\left[(\lambda-1)/(\theta+1)\right]$, if $\lambda > 1$. □

Note that the mode of $f(x)$ is given by α if $0 < \lambda \leq 1$ (x_m if $\lambda > 1$).

Figure 1 shows the p.d.f. (1) for selected values of λ and θ when $\alpha = 1$.

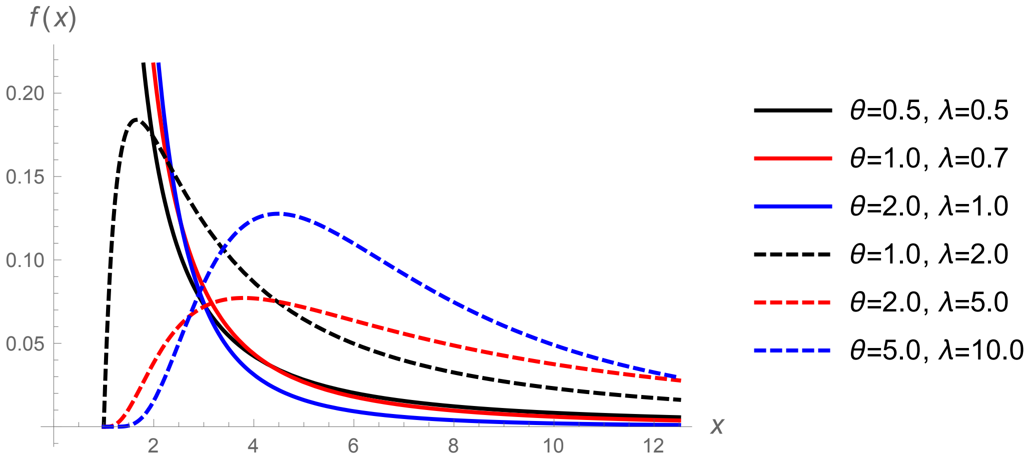

Figure 1. Probability density function of GTLG distribution for selected values of θ and λ when $\alpha = 1$.

2.2. Hazard Rate Function

The survival function (s.f.) of the GTLG distribution is given by

$$\overline{F}(x) = P(X > x) = \frac{\Gamma\left(\lambda, \theta \log\left(\frac{x}{\alpha}\right)\right)}{\Gamma(\lambda)}, \quad x \geq \alpha. \tag{2}$$

where $\Gamma(a, z) = \int_z^\infty t^{a-1} \exp(-t)\, dt$ is the incomplete gamma function. When λ is a positive integer, we have

$$\overline{F}(x) = (x/\alpha)^{-\theta} \sum_{k=0}^{\lambda-1} \frac{[\theta \log(x/\alpha)]^k}{k!}, \quad x \geq \alpha.$$

The hazard rate function (h.r.f.) of the GTLG distribution is given by

$$h(x) = \frac{f(x)}{\overline{F}(x)} = \frac{\theta^\lambda}{\alpha \Gamma\left(\lambda, \theta \log\left(\frac{x}{\alpha}\right)\right)} \left(\frac{x}{\alpha}\right)^{-\theta-1} \left(\log \frac{x}{\alpha}\right)^{\lambda-1}, \qquad x \geq \alpha, \quad \alpha, \theta, \lambda > 0. \tag{3}$$

Note that, $h(\alpha) = f(\alpha)$ and $h(\infty) = 0$.

Theorem 2. *For all $\alpha, \theta > 0$, $h(x)$ is decreasing (increasing-decreasing) if $0 < \lambda \leq 1$ ($\lambda > 1$).*

Proof. Let

$$\eta(x) = -\frac{f'(x)}{f(x)} = \left[(\theta + 1) - \frac{\lambda - 1}{\log(x/\alpha)}\right] \frac{1}{x}.$$

It is straightforward to show that $\eta(x)$ is decreasing if $0 < \lambda \leq 1$ and $\eta(x)$ is increasing-decreasing if $\lambda > 1$. Now by Glaser (1980), $h(x)$ is decreasing if $\lambda \leq 1$ and increasing-decreasing if $\lambda > 1$, since $f(\alpha) = h(\alpha) = 0$ when $\lambda > 1$. □

Figure 2 shows the h.r.f. (3) for selected values of λ and θ when $\alpha = 1$.

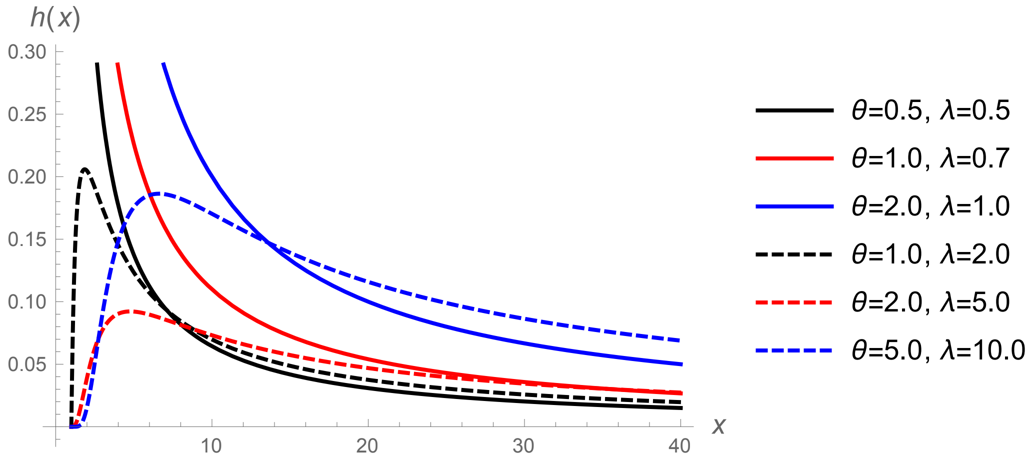

Figure 2. Hazard rate function of GTLG distribution for selected values of θ and λ when $\alpha = 1$.

2.3. Mean Residual Life Function

For the GTLG$(\alpha, \lambda, \theta)$, we have

$$\int_x^\infty y f(y) dy = \mu \int_x^\infty f_{\theta-1}(y) dy = \mu\, \overline{F}_{\theta-1}(x), \qquad x > \alpha, \quad \theta > 1.$$

where, for $\theta > 1$, $\mu = E(X) = \alpha \left(1 - \frac{1}{\theta}\right)^{-\lambda}$ is the mean of the GTLG distribution, and $f_{\theta-1}(x)$ ($\overline{F}_{\theta-1}(x)$) is the p.d.f. (1) (s.f. (2)) when θ is replaced by $\theta - 1$.

The mean residual life function (m.r.l.f.) of the GTLG distribution is given by

$$\begin{aligned} e(x) &= E(X - x | X > x) \\ &= \frac{1}{\overline{F}(x)} \int_x^\infty y f(y)\, dy \; - \; x \\ &= \mu\, \frac{\overline{F}_{\theta-1}(x)}{\overline{F}_{\theta}(x)} \; - \; x, \qquad x > \alpha, \quad \theta > 1. \end{aligned} \tag{4}$$

Theorem 3. *For all $\alpha > 0, \theta > 1$, the m.r.l.f. $e(x)$ is increasing (decreasing-increasing) if $0 < \lambda \leq 1$ ($\lambda > 1$).*

Proof. Since $h(x)$ is decreasing for $0 < \lambda \leq 1$, it follows that, in this case, $e(x)$ is increasing. In addition, since $h(x)$ is increasing-decreasing for $\lambda > 1$ and $f(\alpha)e(\alpha) = 0$, it follows that, in this case, $e(x)$ is decreasing-increasing, by Gupta and Akman (1995). □

From the point of view of a risk manager, the expression $e(x) + x = E(X|X > x)$ is the so-called *Expected Shortfall*, that is the conditional mean of X given X exceeds a given quantile value x. This is a risk measurement appropriate to evaluate the market risk or credit risk of a portfolio.

Figure 3 shows the m.r.l.f. (4) for selected values of λ and θ when $\alpha = 1$.

It is noted that, unlike the classical Pareto distribution, this expression is not a linear function of x.

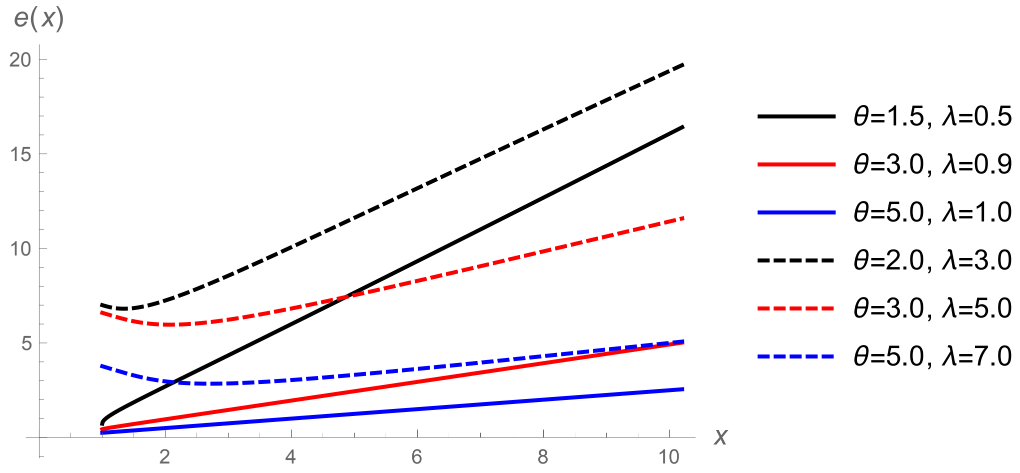

Figure 3. Mean residual life function of GTLG distribution for selected values of θ and λ when $\alpha = 1$.

2.4. Moments

The GTLG distribution with p.d.f. (1) can be obtained from a monotonic transformation of the gamma distribution, as it can be seen in the next result.

Theorem 4. *Let us assume that Y follows a Gamma(λ, θ) distribution with p.d.f. $f(y) \propto y^{\lambda-1}\exp(-\theta y)$, where $\lambda > 0$ and $\theta > 0$. Then the random variable*

$$X = \alpha\, e^{Y} \tag{5}$$

has p.d.f. (1).

Proof. The proof follows after a simple change of variable. □

Note that $Z = e^{Y}$ has a log-gamma distribution over $(1, \infty)$. That is $X = \alpha Z$ as indicated before.

Now, by using representation (5) and the moments of the Gamma distribution, the expression for the r-th moment about zero of distribution (1) is easily obtained,

$$\mu'_r = E(X^r) = \alpha^r\, M_Y(r) = \alpha^r \left(1 - \frac{r}{\theta}\right)^{-\lambda}, \qquad r = 1, 2, \ldots,$$

provided $\theta > r$ and $\lambda > 0$.

In particular, the mean is given by

$$\mu = \alpha \left(1 - \frac{1}{\theta}\right)^{-\lambda}, \qquad \theta > 1, \tag{6}$$

and the variance is given by

$$\sigma^2 = \alpha^2 \left[\left(1 - \frac{2}{\theta}\right)^{-\lambda} - \left(1 - \frac{1}{\theta}\right)^{-2\lambda}\right], \qquad \theta > 2, \tag{7}$$

Furthermore, by using the representation given by (5) the following result is obtained

$$E\left[\log\left(\frac{X}{\alpha}\right)\right]^r = E(Y^r) = \frac{\Gamma(\lambda + r)}{\theta^r \Gamma(\lambda)}, \qquad r = 1, 2, \ldots. \tag{8}$$

Solving the equation

$$\mu = \alpha \left(1 - \frac{1}{\theta}\right)^{-\lambda},$$

in θ, we obtain

$$\theta = \frac{(\mu/\alpha)^{1/\lambda}}{(\mu/\alpha)^{1/\lambda} - 1}, \qquad \mu > \alpha.$$

This implies that the covariates can be introduced into the model in a simple way.

2.5. Conjugate Distributions

The following results show that both the inverse Gaussian distribution and the gamma distribution are conjugate with respect to the distribution proposed in this work.

Theorem 5. *Let X_i, $i = 1, 2, \ldots, n$ independent and identically distributed random variables following the p.d.f. (1). Let us suppose that θ follows a prior inverse Gaussian distribution $\pi(\theta)$ with parameters τ and ϕ, i.e., $\pi(\theta) \propto \theta^{-3/2} \exp\left[-\frac{1}{2}\left(\frac{\phi}{\tau^2}\theta + \frac{\phi}{\theta}\right)\right]$. Then the posterior distribution of θ given the sample information $(X_1, \ldots, X_n)$ is a generalized inverse Gaussian distribution $GIG(\lambda^*, \tau^*, \phi^*)$, where*

$$\begin{aligned} \lambda^* &= n\lambda - \frac{1}{2}, \\ \tau^* &= \tau\sqrt{1 + \frac{2\tau^2}{\phi}\sum_{i=1}^{n}\log(x_i/\alpha)}, \\ \phi^* &= \phi. \end{aligned}$$

Proof. The result follows after some computations by applying Bayes' Theorem and arranging parameters. □

Theorem 6. *Let X_i, $i = 1, 2, \ldots, n$ independent and identically distributed random variables following the p.d.f. (1). Let us suppose that θ follows a prior gamma distribution $\pi(\theta)$ with a shape parameter $\tau > 0$ and a scale parameter $\sigma > 0$, i.e., $\pi(\theta) \propto \theta^{\tau-1}\exp(-\sigma\theta)$. Then the posterior distribution of θ given the sample information $(X_1, \ldots, X_n)$ is again a gamma distribution with shape parameter $\tau + n\lambda$ and scale parameter $\sigma + \log(x_i/\alpha)$.*

Proof. Again, the result follows after some algebra by using Bayes' Theorem and arranging parameters. □

2.6. Stochastic Ordering

Stochastic ordering of positive continuous random variables is an important tool for judging the comparative behavior. We will recall some basic definitions, see (Shaked and Shanthikumar 2007).

Let X and Y be random variables with p.d.f.s $f(x)$ and $g(y)$ (s.f.s $\overline{F}(x)$ and $\overline{G}(y)$) (h.r.f.s $h(x)$ and $r(y)$), respectively.

A random variable X is said to be smaller than a random variable Y in the

(i) stochastic order (denoted by $X \preceq_{ST} Y$) if $\overline{F}(x) \leq \overline{G}(x)$ for all x,
(ii) hazard rate order (denoted by $X \preceq_{HR} Y$) if $h(x) \geq r(x)$ for all x,
(iii) likelihood ratio order (denoted by $X \preceq_{LR} Y$) if $\frac{f(x)}{g(x)}$ decreases for all x.

The following implications are well known:

$$X \preceq_{LR} Y \Rightarrow X \preceq_{HR} Y \Rightarrow X \preceq_{ST} Y.$$

Members of the family of distributions with p.d.f. (1) are ordered with respect to the strongest "likelihood ratio" ordering, as shown in the following theorem.

Theorem 7. *Let X and Y be two continuous random variables distributed according to (1) with p.d.f.'s given by $f(x) = f(x;\alpha,\theta_1,\lambda)$ and $g(x) = f(x;\alpha,\theta_2,\lambda)$, respectively. If $\theta_1 \geq \theta_2 > 0$, then $X \preceq_{LR} Y$ ($X \preceq_{HR} Y$) ($X \preceq_{ST} Y$)*

Proof. Firstly, let us observe that the ratio

$$\frac{f(x)}{g(x)} = \left(\frac{\theta_1}{\theta_2}\right)^{\lambda} \alpha^{\theta_1-\theta_2} x^{\theta_2-\theta_1}$$

wih derivative

$$\left(\frac{f(x)}{g(x)}\right)' = \left(\frac{\theta_1}{\theta_2}\right)^{\lambda} \alpha^{\theta_1-\theta_2}(\theta_2-\theta_1)x^{\theta_2-\theta_1-1} \leq 0,$$

for all $\theta_1 \geq \theta_2 > 0$, proving the theorem. □

Properties for higher-order stochastic dominance in financial economics can be obtained following the line of the work of (Guo and Wong 2016). In this regard, let X and Y be random variables defined on $[a,b]$ with p.d.f.'s $f(x), g(y)$ and s.f.'s $\overline{F}(x), \overline{G}(y)$, respectively, satisfying

$$F_j^D(x) = \int_x^b F_{j-1}^D(y)dy, \qquad G_j^D(x) = \int_x^b G_{j-1}^D(y)dy, \qquad j \geq 1,$$

where $F_0^D(x) = f(x)$, $G_0^D(x) = g(x)$, $F_1^D(x) = \overline{F}(x)$, and $G_1^D(x) = \overline{G}(x)$.

A random variable X is said to be smaller than a random variable Y

(i) in the *first-order* descending stochastic dominance (denoted by $X \preceq^1 Y$) iff $F_1^D(x) \leq G_1^D(x)$ for each $x \in [a,b]$.
(ii) in the *second-order* descending stochastic dominance (denoted by $X \preceq^2 Y$) iff $F_2^D(x) \leq G_2^D(x)$ for each $x \in [a,b]$.
(iii) in the *N-order* descending stochastic dominance (denoted by $X \preceq^N Y$) iff $F_N^D(x) \leq G_N^D(x)$ for each $x \in [a,b]$ and $F_k^D(a) \leq G_k^D(a)$ for $2 \leq k \leq N-1$, $N \geq 3$.

Theorem 8. *Let X and Y be two continuous random variables distributed according to (1) with p.d.f.'s given by $f(x) = f(x;\alpha,\theta_1,\lambda)$ and $g(y) = f(y;\alpha,\theta_2,\lambda)$, respectively.*

(i) If $\theta_1 \geq \theta_2 > 0$, then $X \preceq^1 Y$.

(ii) If $\theta_1 \geq \theta_2 > 0$, then $X \preceq^2 Y$.

(iii) If $\theta_1 \geq \theta_2 > 1$, then $X \preceq^N Y$ for $N \geq 3$.

Proof. (i) For $\theta_1 \geq \theta_2 > 0$, we have

$$F_1^D(x) = \overline{F}_{\theta_1}(x) = \frac{\Gamma(\lambda, \theta_1 \log(x/\alpha))}{\Gamma(\lambda)} \leq \frac{\Gamma(\lambda, \theta_2 \log(x/\alpha))}{\Gamma(\lambda)} = \overline{G}_{\theta_2}(x) = G_1^D(x).$$

Therefore, for $\theta_1 \geq \theta_2 > 0$, $X \preceq^1 Y$.

(ii) For $\theta_1 \geq \theta_2 > 0$, we have

$$F_2^D(x) = \int_x^{\infty} F_1^D(y)dy \leq \int_x^{\infty} G_1^D(y)dy = G_2^D(x).$$

Therefore, for $\theta_1 \geq \theta_2 > 0$, $X \preceq^2 Y$.

(iii) For $\theta_1 \geq \theta_2 > 1$, we have

$$F_3^D(x) = \int_x^\infty F_2^D(y)dy \leq \int_x^\infty G_2^D(y)dy = G_3^D(x).$$

Also, for $\theta_1 \geq \theta_2 > 1$, we have

$$F_2^D(\alpha) = \int_\alpha^\infty \overline{F}_{\theta_1}(y)dy = \mu_{\theta_1} \leq \mu_{\theta_2} = \int_\alpha^\infty \overline{G}_{\theta_2}(y)dy = G_2^D(\alpha).$$

Therefore, for $\theta_1 \geq \theta_2 > 1$, $X \preceq^3 Y$.

Now assume that, for $\theta_1 \geq \theta_2 > 1$, $X \preceq^N Y$ for some $N \geq 3$, i.e., $F_N^D(x) \leq G_N^D(x)$ for each $x \in [a,b]$ and $F_k^D(a) \leq G_k^D(a)$ for $2 \leq k \leq N-1$, $N \geq 3$.

Now for $\theta_1 \geq \theta_2 > 1$, we have

$$F_{N+1}^D(x) = \int_x^\infty F_N^D(y)dy \leq \int_x^\infty G_N^D(y)dy = G_{N+1}^D(x).$$

Also, for $\theta_1 \geq \theta_2 > 1$, we have

$$F_N^D(\alpha) = \int_\alpha^\infty F_{N-1}^D(y)dy \leq \mu_{\theta_2} = \int_\alpha^\infty G_{N-1}^D(y)dy = G_N^D(\alpha).$$

Therefore, for $\theta_1 \geq \theta_2 > 1$, $X \preceq^N Y$ for all $N \geq 3$. □

3. Some Theoretical Financial Results

The integrated tail distribution function (also known as equilibrium distribution function):

$$F_I(x) = \frac{1}{E(X)} \int_\alpha^x \overline{F}(y)\, dy, \qquad x > \alpha.$$

is an important probability model that often appears in insurance and many other applied fields (see for example Yang 2004).

For the GTLG$(\alpha, \lambda, \theta)$, we have

$$\int_\alpha^x yf(y)dy = \mu \int_\alpha^x f_{\theta-1}(y)dy = \mu\, F_{\theta-1}(x), \qquad x > \alpha, \quad \theta > 1.$$

The integrated tail distribution of the GTLG $(\alpha, \lambda, \theta)$ is given by

$$\begin{aligned} F_I(x) &= \frac{1}{\mu} \int_\alpha^x \overline{F}_\theta(y)dy \\ &= \frac{1}{\mu} \left\{ x\, \overline{F}_\theta(x) - \alpha + \mu \int_\alpha^x f_{\theta-1}(y)dy \right\} \\ &= \frac{1}{\mu} \left\{ x\, \overline{F}_\theta(x) - \alpha + \mu\, F_{\theta-1}(x) \right\}, \qquad x > \alpha, \quad \theta > 1. \end{aligned} \tag{9}$$

Under the classical model (see Yang 2004) and assuming a positive security loading, ρ, for the claim size distributions with regularly varying tails we have that, by using (3), it is possible to obtain an approximation of the probability of ruin, $\Psi(u)$, when $u \to \infty$. In this case the asymptotic approximations of the ruin function is given by

$$\Psi(u) \quad \sim \quad \frac{1}{\rho}\overline{F}_I(u), \quad u \to \infty.$$

where $\overline{F}_I(u) = 1 - F_I(u)$.

The use of heavy right-tailed distribution is of vital importance in general insurance. In this regard, Pareto and log-normal distributions have been employed to model losses in motor third liability insurance, fire insurance or catastrophe insurance. It is already known that any probability distribution, that is specified through its cumulative distribution function $F(x)$ on the real line, is heavy right-tailed if and only if for every $t > 0$, $e^{tx}\overline{F}(x)$ has an infinite limit as x tends to infinity. On this particular subject, (1) decays to zero slower than any exponential distribution and it is long-tailed since for any fixed $t > 0$ (see Rytgaard 1990) it is verified that

$$\overline{F}(x+t) \sim \overline{F}(x), \quad x \to \infty.$$

Therefore, as a long-tailed distribution is also heavy right-tailed, the distribution introduced in this manuscript is also heavy right–tailed.

Another important issue in extreme value theory is the regular variation (see Bingham 1987; Rytgaard 1990). A distribution function is called regular varying at infinity with index $-\beta$ if

$$\lim_{x\to\infty} \frac{\overline{F}(tx)}{\overline{F}(x)} = t^{-\beta},$$

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

Theorem 9. *The GTLG distribution is regularly varying at infinity with index $-\theta$.*

Proof. Using L'Hospital rule, we have

$$\begin{aligned}
\lim_{x\to\infty} \frac{\overline{F}(tx)}{\overline{F}(x)} &= \lim_{x\to\infty} \frac{\Gamma(\lambda, \theta \log \frac{tx}{\alpha})}{\Gamma(\lambda, \theta \log \frac{x}{\alpha})} = \lim_{x\to\infty} \frac{-\left(\theta \log \frac{tx}{\alpha}\right)^{\lambda-1} e^{-\theta \log \frac{tx}{\alpha}} \left(\frac{\theta}{x}\right)}{-\left(\theta \log \frac{x}{\alpha}\right)^{\lambda-1} e^{-\theta \log \frac{x}{\alpha}} \left(\frac{\theta}{x}\right)} \\
&= \lim_{x\to\infty} \left(1 + \frac{\log t}{\log \frac{x}{\alpha}}\right)^{\lambda-1} t^{-\theta} = t^{-\theta},
\end{aligned}$$

for all $\alpha, \theta, \lambda > 0$. □

As a consequence of this result we have that if $X, X_1, \ldots, X_n$ are i.i.d. random variables with common s.f. (2) and $S_n = \sum_{i=1}^{n} X_i$, $n \geq 1$, then

$$\Pr(S_n > x) \sim \Pr(X > x) \quad \text{as} \quad x \to \infty.$$

Therefore, if $P_n = \max_{i=1,\ldots,n} X_i$, $n \geq 1$, we have that

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

This means that for large x the event $\{S_n > x\}$ is due to the event $\{P_n > x\}$. Therefore, exceedance of high thresholds by the sum S_n are due to the exceedance of this threshold by the largest value in the sample.

On the other hand, let the random variable X represent either a policy limit or reinsurance deductible (from an insurer's perspective); then the limited expected value function L of X with cdf $F(x)$, is defined by

$$\begin{aligned}
L(x) &= E[\min(X, x)] \\
&= \int_{\alpha}^{x} y\, f_\theta(y)\, dy + x\, \overline{F}_\theta(x) \qquad (10)\\
&= \mu\, F_{\theta-1}(x) + x\, \overline{F}_\theta(x), \qquad x > \alpha, \quad \theta > 1.
\end{aligned}$$

Note that $L(x)$ represents the expected amount per claim retained by the insured on a policy with a fixed amount deductible of x.

Figure 4 shows the limited expected value function (10) for selected values of λ and θ when $\alpha = 1$.

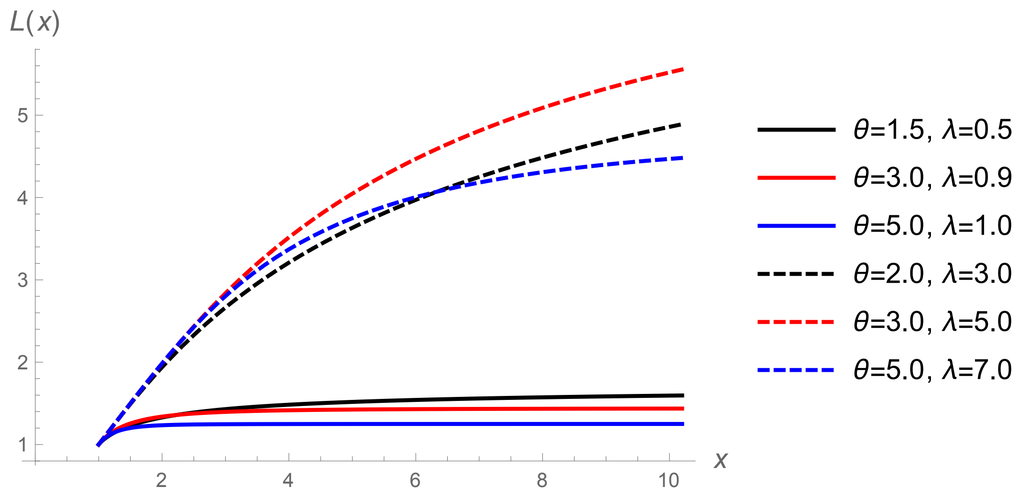

Figure 4. Limited expected value function of GTLG distribution for selected values of θ and λ when $\alpha = 1$.

On the other hand, the new distribution can also be applied in rating excess–of–loss reinsurance as it can be seen in the next result.

Theorem 10. *Let X be a random variable denoting the individual claim size taking values only for individual claims greater than d. Let us also assumed that X follows the pdf (1), then the expected cost per claim to the reinsurance layer when the losses excess of m subject to a maximum of l is given by*

$$\begin{aligned} E[\min(l, \max(0, X-m))] &= \frac{\theta^\lambda}{\Gamma(\lambda)} \left[m \left(R(\lambda, \theta, m+l) - R(\lambda, \theta, m) \right) \right. \\ &\quad \left. + \alpha \left(R(\lambda, \theta-1, m) - R(\lambda, \theta-1, m+1) \right) \right] \\ &\quad + l\overline{F}(m+l), \end{aligned}$$

where $R(a, b, z) = \log^\lambda(z/\alpha) E_{1-a}(b \log(z/\alpha))$, being $E_n(z) = \int_1^\infty t^{-n} \exp(-zt)\, dt$ the exponential integral function.

Proof. The result follows by having into account that

$$E[\min(l, \max(0, X-m))] = \int_m^{m+l} (x-m) f(x)\, dx + l\overline{F}(m+l),$$

from which we get the result after some tedious algebra. □

4. Maximum Likelihood Estimation

In the following it will be assumed that $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is a random sample selected from the GTLG distribution with known parameter α and unknown parameters $\boldsymbol{\nu} = (\theta, \lambda)$ from the p.d.f. (1). Then, the log–likelihood function is given by

$$\ell(\boldsymbol{\nu}; \mathbf{x}) = n \left[\lambda \log \theta - \log \alpha \Gamma(\lambda) \right] + \sum_{i=1}^n \left[-(\theta+1) \log \left(\frac{x_i}{\alpha} \right) + (\lambda - 1) \log \log \left(\frac{x_i}{\alpha} \right) \right]. \qquad (11)$$

The maximum likelihood estimates (MLEs) $\widehat{\boldsymbol{\nu}} = (\widehat{\theta}, \widehat{\lambda})$, of the parameters $\boldsymbol{\nu} = (\theta, \lambda)$ are obtained by solving the score equations:

$$\begin{aligned} \frac{\partial \ell}{\partial \theta} &= \frac{n\lambda}{\theta} - \sum_{i=1}^{n} \log(x_i/\alpha) = 0, \\ \frac{\partial \ell}{\partial \lambda} &= n\left[\log\theta - \psi(\lambda)\right] + \sum_{i=1}^{n} \log\log(x_i/\alpha) = 0, \end{aligned}$$

where $\psi(\cdot)$ is the digamma function. Therefore,

$$\widehat{\theta} = \frac{n\widehat{\lambda}}{\sum_{i=1}^{n} \log(x_i/\alpha)},$$

where $\widehat{\lambda}$ is the solution of the equation:

$$\log(\widehat{\lambda}) - \psi(\widehat{\lambda}) - \log\left[\frac{1}{n}\sum_{i=1}^{n} \log(x_i/\alpha)\right] + \frac{1}{n}\sum_{i=1}^{n} \log\log(x_i/\alpha) = 0.$$

The second partial derivatives are given by

$$\begin{aligned} \frac{\partial^2 \ell}{\partial \theta^2} &= -\frac{n\lambda}{\theta^2}, \\ \frac{\partial^2 \ell}{\partial \theta \partial \lambda} &= \frac{n}{\theta}, \\ \frac{\partial^2 \ell}{\partial \lambda^2} &= -n\psi'(\lambda). \end{aligned}$$

The expected Fisher's information matrix is given by

$$\mathcal{I}(\boldsymbol{\nu}) = \begin{bmatrix} \frac{n\lambda}{\theta^2} & -\frac{n}{\theta} \\ & \\ -\frac{n}{\theta} & n\psi'(\lambda) \end{bmatrix}. \qquad (12)$$

Now the estimated variance-covariance matrix of the MLEs $\widehat{\boldsymbol{\nu}}$ is given by the inverse matrix $\mathcal{I}^{-1}(\widehat{\boldsymbol{\nu}})$.

It is known that under certain regularity conditions, the maximum likelihood estimator $\widehat{\boldsymbol{\nu}}$ converges in distribution to a bivariate normal distribution with mean equal to the true parameter value and variance-covariance matrix given by the inverse of the information matrix. That is, $\widehat{\boldsymbol{\nu}} \xrightarrow{D} \mathcal{N}(\boldsymbol{\nu}, \mathcal{I}^{-1}(\boldsymbol{\nu}))$, which provides a basis for constructing tests of hypotheses and confidence regions. The regularity conditions are verified by taking into account that the Fisher's information matrix exists and is non-singular and that the parameter space is a subset of the real line and the range of x is independent of ν. Furthermore, additional computations provides that $E\left(\frac{\partial f(x)}{\partial \nu}\right) = \mathbf{0}$ and that $\frac{\partial^3 f(x)}{\partial \nu^3}$ is bounded.

5. Numerical Application

Because the main application of the heavy tail distributions is the so-called extreme value theory, we consider a data set coming from catastrophic events. The data set represents loss ratios (yearly data in billion of dollars) for earthquake insurance in California from 1971 through 1993 for values larger than zero. The data are given in Embrechts et al. (1999).

For comparison with other heavy tail distributions, we consider the following models:

(1) Pareto distribution:

$$f(x) = \frac{\theta}{x}\left(\frac{\alpha}{x}\right)^{\theta}, \qquad x \geq \alpha, \quad \alpha, \theta > 0.$$

(2) Shifted log-normal (SLN):

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

(3) Burr distribution:

$$f(x) = \frac{\theta\lambda(x-\alpha)^{\theta-1}}{\left[1-(x-\alpha)^{\theta}\right]^{\lambda+1}}, \qquad x \geq \alpha, \quad \alpha, \theta, \lambda > 0.$$

(4) Stoppa distribution:

$$f(x) = \frac{\lambda\theta}{x}\left(\frac{\alpha}{x}\right)^{\theta}\left[1-\left(\frac{\alpha}{x}\right)^{\theta}\right]^{\lambda-1}, \qquad x \geq \alpha, \quad \alpha, \theta, \lambda > 0.$$

(5) Log-gamma distribution (LG):

$$f(x) = \frac{(1+x-\alpha)^{-1-1/\theta}}{\theta^{\lambda}\Gamma(\lambda)} \log^{\lambda-1}(1+x-\alpha), \qquad x \geq \alpha, \quad \alpha, \theta, \lambda > 0.$$

Table 1 provides parameter estimates together with standard errors (in brackets) using the maximum likelihood estimation method of the parameters θ and λ when $\alpha = 0.1$. This table also gives the negative log-likelihood (NLL), Akaike's Information Criteria (AIC), Bayesian information criterion (BIC), and Consistent Akaike's Information Criteria (CAIC).

A lower value of these measures is desirable. These results show that the proposed GTLG distribution provides better fit than the considered competing distributions. Table 2 shows three goodness-of-fit tests for all considered models and that the classical Pareto model is rejected for this data set.

Table 1. Estimated values of the considered models when $\alpha = 0.1$.

Distribution	Estimates (S.E.)	NLL	AIC	BIC	CAIC
Pareto	$\theta = 0.249\,(0.057)$	77.939	157.878	158.822	159.822
SLN	$\theta = 1.477\,(0.239)$ $\lambda = 1.668\,(0.339)$	66.080	136.161	138.05	140.05
Burr	$\theta = 2.287\,(0.895)$ $\lambda = 0.243\,(0.106)$	67.352	138.703	140.592	142.592
Stoppa	$\theta = 0.768\,(0.159)$ $\lambda = 12.013\,(6.065)$	66.321	136.643	138.532	140.532
LG	$\theta = 0.802\,(0.271)$ $\lambda = 2.474\,(0.755)$	66.273	136.547	138.435	140.435
GTLG	$\theta = 1.845\,(0.606)$ $\lambda = 7.401\,(2.352)$	65.987	135.974	137.863	139.863

Table 2. Test statistics (p-values) of goodness-of-fit tests of the considered models when $\alpha = 0.1$.

Distribution	Kolmogorov-Smirnov	Cramér-Von Misses	Anderson-Darling
Pareto	0.360 (0.010)	0.706 (0.012)	3.574 (0.014)
SLN	0.116 (0.933)	0.031 (0.970)	0.205 (0.989)
Burr	0.182 (0.500)	0.090 (0.633)	0.483 (0.761)
Stoppa	0.148 (0.746)	0.040 (0.933)	0.242 (0.974)
LN	0.149 (0.731)	0.043 (0.918)	0.257 (0.966)
GTLG	0.148 (0.745)	0.040 (0.932)	0.242 (0.974)

6. Conclusions

In this paper, a continuous probability distribution function with positive support suitable for fitting insurance data has been introduced. The distribution, that arises from a monotonic transformation of the classical Gamma distribution, can be considered as a generalization of the log-gamma distribution. This new development, which has a promising approach for data modeling in the actuarial field, may be very useful for practitioners who handle large claims. For that reason, it can be deemed as an alternative to the classical Pareto distribution. Besides, an extensive analysis of its mathematical properties has been provided.

Acknowledgments: The authors would like to express their gratitude to two anonymous referees for their relevant and useful comments. Emilio Gómez-Déniz was partially funded by grant ECO2013-47092 (Ministerio de Economía y Competitividad, Spain and ECO2017-85577-P (Ministerio de Economía, Industria y Competitividad, Agencia Estatal de Investigación)).

Author Contributions: Mohamed E. Ghitany, Emilio Gómez-Déniz and Saralees Nadarajah have contributed jointly to all of the sections of the paper.

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

References

Amini, Morteza, S. M. T. K. MirMostafaee, and Jafar Ahmadi. 2014. Log-gamma-generated families of distributions. *Statistics* 48: 913–32.

Arnold, Barry C. 1983. *Pareto Distributions*. Silver Spring: International Cooperative Publishing House.

Bingham, Nicholas H. 1987. *Regular Variation*. Cambridge: Cambridge University Press.

Boyd, Albert V. 1988. Fitting the truncated Pareto distribution to loss distributions. *Journal of the Staple Inn Actuarial Society* 31: 151–58.

Brazauskas, Vytaras, and Robert Serfling. 2003. Favorable estimator for fitting Pareto models: A study using goodness-of-fit measures with actual data. *ASTIN Bulletin* 33: 365–81.

Consul, Prem C., and Gaurav C. Jain. 1971. On the log-Gamma distribution and its properties. *Statistische Hefte* 12: 100–6.

Embrechts, Paul, Sidney I. Resnick, and Gennady Samorodnitsky. 1999. Extreme value theory as a risk management tool. *North American Actuarial Journal* 3: 30–41.

Glaser, Ronald E. 1980. Bathtub and related failure rate characterizations. *Journal of the American Statistical Association* 75: 667–72.

Gómez-Déniz, Emilio, and Enrique Calderín. 2014. A suitable alternative to the Pareto distribution. *Hacettepe Journal of Mathematics and Statistics* 43: 843–60.

Guo, Xu, and Wing-Keung Wong. 2016. Multivariate stochastic dominance for risk averters and risk seekers. *RAIRO-Operations Research* 50: 575–86.

Gupta, Ramesh C., and Akman, H. Olcay. 1995. Mean residual life function for certain types of non-monotonic ageing. *Communications in Statistics-Stochastic Models* 11: 219–25.

Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot. 2008. *Loss Models: From Data to Decisions*, 3rd ed. Hoboken: Wiley.

Mata, Ana J. 2000. Princing excess of loss reinsurance with reinstatements. *ASTIN Bulletin* 30: 349–68.

Rytgaard, Mette. 1990. Estimation in the Pareto distribution. *ASTIN Bulletin* 20: 201–16.

Shaked, Moshe, and Jeyaveerasingam G. Shanthikumar. 2007. *Stochastic Orders*. Series: Springer Series in Statistics. New York: Springer.

Yang, Hailiang. 2004. Crámer-Lundberg asymptotics. In *Encyclopedia of Actuarial Science*. New York: Wiley, pp. 1–6.

Zografos, Konstantinos, and Narayanaswamy Balakrishnan. 2009. On families of beta and generalized gamma-generated distributions and associated inference. *Statistical Methodology* 6: 344–62.