5. SOME COMPOSITE EXPONENTIAL-PARETO MODELS FOR ACTUARIAL PREDICTION

Sandra TEODORESCU*
RalucaVERNIC**

Abstract

Prediction is a very important and not so easy task for an actuary. An insurance company needs predictions of the future claims in order to evaluate premiums, to assess its financial situation, probabilities of ruin, etc. Therefore, modeling the claims distribution is of great importance, but since this distribution is usually different from the classical ones (e.g. skewed and heavy tailed), researchers are trying to find new models that can fit better to insurance data.

Such a composite model unifying a Lognormal and a Pareto distribution was introduced by Cooray and Ananda [1] and generalized by Scollnik [6]. In this paper we go even further and study a composite model obtained from two arbitrary distributions, then exemplify it with the Exponential and Pareto distributions. Some properties and statistical inference are also presented.

Keywords: composite models, mixture models, Exponential and Pareto distributions, composite Exponential-Pareto models, parameter estimation

JEL Classification: C14, G22

1. Introduction

Actuarial science has to work with a large set of statistical tools (for data collection, analysis, estimating, forecasting and valuation), in order to provide financial and underwriting data for risk management, and to assess marketing opportunities. Therefore, actuaries must forecast losses using models of random events. The choice of claims distributions becomes an important and sometimes quite difficult task.

A main classification of distributions with right-infinite domain is into “light” and “heavy” tailed, depending on how spread out they are in the right tail. In this sense, the

* Lecturer, Ph.D., Faculty of Economic Sciences, Ecological University of Bucharest, cezarina_teodorescu@yahoo.com
** Associate Professor,Ph.D., Faculty of Mathematics and Computer Science, Ovidius University of Constanta, rvernic@univ-ovidius.ro

Exponential distribution is used as a standard: if a distribution has a right tail less spread out than the Exponential one, then it is called "light-tailed", otherwise it is "heavy-tailed".

When studying the claims affecting a certain insurance portfolio, a frequent situation is that we have many small claims, but also a few large claims that generate a heavy-tailed distribution. Such situations are often encountered in, e.g., property insurance, auto insurance, etc. (see, for instance [4]). Then, the claims distribution can be modeled as a combination of two densities, consisting of a less heavy-tailed distribution up to a certain threshold value and of a heavy-tailed distribution from that threshold on. Such distributions are the composite ones, as suggested by Cooray and Ananda [1]. They constructed a composite model as:

$$f(x)=\begin{cases} cf_1(x) & if\ 0<x\le\theta \\ cf_2(x) & if\ \theta<x\le\infty \end{cases}, \tag{1.1}$$

where: f_1 and f_2 are probability density functions (pdf), while c is a normalizing constant.

In order to have a smooth pdf (1.1), continuity and differentiability conditions are imposed at θ, from where we obtain the constant c, and also a condition that reduces by one the number of unknown parameters of f_1 and f_2. As mentioned before, f_1 is usually taken to be a light-tailed distribution, while f_2 will be heavy-tailed.

In this paper, we chose the Exponential density for f_1 and a Pareto density for f_2. This choice is motivated by the fact that among the heavy-tailed distributions the Pareto one is preferred when it comes to model larger claims or reinsurance payments, while the Exponential distribution is easy to handle. Also, the Exponential distribution often appears in actuarial models for claims and, hence, plays an important role in models that allow for analytical computation of ruin probabilities (see, for instance [3]). For details on the Exponential and Pareto distributions see [2].

A composite Exponential-Pareto model in form (1.1) was already proposed by Teodorescu and Vernic [7]. The main characteristic of its pdf is that, even if its shape is similar to the Exponential, it has a larger tail that the corresponding Exponential, and a lighter tail than the corresponding Pareto. The composite Exponential-Pareto density obtained in [7] is

$$f(x)=\begin{cases} \dfrac{0.775}{\theta}e^{-\frac{1.35x}{\theta}}, & 0<x\le\theta \\ 0.2\dfrac{\theta^{0.35}}{x^{1.35}}, & \theta\le x<\infty \end{cases}. \tag{1.2}$$

Recently, Scollnik [6] proposed two different composite models based on Lognormal and Pareto distributions. In the same manner, we suggest two different models based on Exponential and Pareto distributions.

Section 2 is dedicated to the study of a mixture model equivalent to model (1.1), with focus on its pdf, cumulative distribution function (cdf), initial moments, characteristic function and a method for estimating the parameters. In Section 3 we reinterpret the composite Exponential-Pareto model from [7], which can be recognized as a two-component mixture model with fixed and *a priori* known mixing weights. In Sections 4 and 5 we introduce two new composite Exponential-Pareto models and discuss some of their properties based on the general theory of Section 2. Two numerical examples are presented in Section 6, based on two data samples considered in Teodorescu and Vernic [7]. We end the paper with some conclusions.

2. Some properties of a mixture model

Scollnik [6] noticed that the pdf of the mixture model (1.1) might also be written as:

$$f(x)=\begin{cases} rf_1^*(x), & -\infty < x \le \theta \\ (1-r)f_2^*(x), & \theta < x \le \infty \end{cases}, \tag{2.1}$$

where: $0 \le r \le 1$, while f_1^* and f_2^* are adequate truncations of the pdf-s f_1 and f_2. More precisely, if F_i denotes the cdf of f_i, we have:

$$\begin{aligned} f_1^*(x) &= \frac{f_1(x)}{\int_{-\infty}^{\theta} f_1(y)dy} = \frac{f_1(x)}{F_1(\theta)}, \quad -\infty < x \le \theta, \\ f_2^*(x) &= \frac{f_2(x)}{\int_{\theta}^{\infty} f_2(y)dy} = \frac{f_2(x)}{1-F_2(\theta)}, \quad \theta \le x < \infty. \end{aligned} \tag{2.2}$$

One should note that for more generality, we considered that the pdf f_1 and, hence, (2.1), may also be defined for negative values of *x*.

It is easy to see that pdf (2.1) can be interpreted as a two-component mixture model with mixing weights *r* and 1-*r* (i.e., a convex combination of two pdf-s),

$$f(x) = rf_1^*(x) + (1-r)f_2^*(x), \quad r \in [0,1]. \tag{2.3}$$

We will now deduce some characteristics of model (2.1) concerning its cdf, initial moments and the characteristic function. If *X* is a random variable (r.v.) with pdf *f*, we denote its initial *n*-th order moment by $E_n(f) = E(X^n)$, and its characteristic function by $\varphi_f = \varphi_X$.

Proposition 2.1.

Let F denote the cdf of the pdf given in (2.1). Then

$$F(x)=\begin{cases} r\dfrac{F_1(x)}{F_1(\theta)}, & -\infty < x \le \theta \\ r+(1-r)\dfrac{F_2(x)-F_2(\theta)}{1-F_2(\theta)}, & \theta \le x < \infty \end{cases},$$

Proof. We consider two cases and use formulas (2.2):

- when $-\infty < x \le \theta$,

$$F(x)=\int_{-\infty}^{x} f(y)dy = r\int_{-\infty}^{x} f_1^*(y)dy = r\int_{-\infty}^{x}\frac{f_1(y)}{F_1(\theta)}dy = r\frac{F_1(x)}{F_1(\theta)}.$$

- when $\theta \le x < \infty$,

$$F(x)=r\int_{-\infty}^{\theta} f_1^*(y)dy+(1-r)\int_{\theta}^{x} f_2^*(y)dy = r\frac{F_1(\theta)}{F_1(\theta)}+(1-r)\int_{\theta}^{x}\frac{f_2(y)}{1-F_2(\theta)}dy =$$

$$= r+(1-r)\frac{F_2(x)-F_2(\theta)}{1-F_2(\theta)}.$$

Therefore, we get the formula of *F*.

Proposition 2.2.

The initial n-th order moment of the pdf (2.3) is

$$E_n(f)=rE_n(f_1^*)+(1-r)E_n(f_2^*),$$

assuming that all the quantities involved exist.

Proof.

We have:

$$E_n(f)=\int_{-\infty}^{\infty} x^n f(x)dx = r\int_{-\infty}^{\infty} x^n f_1^*(x)dx+(1-r)\int_{-\infty}^{\infty} x^n f_2^*(x)dx =$$

$$= rE_n(f_1^*)+(1-r)E_n(f_2^*).$$

Proposition 2.3.

The characteristic function of the pdf (2.3) is

$$\varphi_f(t)=r\varphi_{f_1^*}(t)+(1-r)\varphi_{f_2^*}(t), \quad t\in \mathrm{R}.$$

Proof.

We have:

$$\varphi_f(t)=\int_{-\infty}^{\infty} e^{itx} f(x)dx = r\int_{-\infty}^{\infty} e^{itx} f_1^*(x)dx+(1-r)\int_{-\infty}^{\infty} e^{itx} f_2^*(x)dx =$$

$$= r\varphi_{f_1^*}(t)+(1-r)\varphi_{f_2^*}(t).$$

Statistical inference. Assuming that the pdf (2.1) depends on the real parameters $\delta_1,...,\delta_s,\theta,$ where $s\in \mathrm{N}$, we will now present an estimation algorithm based on the maximum likelihood (ML) method. Consider the random sample $(x_1,...,x_n)$. Without loss of generality, we assume that it is an ordered sample, i.e. $x_1 \le x_2 \le ... \le x_n$. In order to apply the ML method, we must know the integer value *m* such that the unknown parameter θ is between the m-th and (m+1)-th observations, i.e. $x_m \le \theta \le x_{m+1}$. Assuming that somehow we know this m, the likelihood function is

$$L(x_1,...,x_n;\delta_1,...,\delta_s,\theta) = \prod_{i=1}^{m} r f_1^*(x_i) \prod_{j=m+1}^{n} (1-r) f_2^*(x_j) = \\ = r^m (1-r)^{n-m} \prod_{i=1}^{m} f_1^*(x_i) \prod_{j=m+1}^{n} f_2^*(x_j). \quad (2.4)$$

Unfortunately, in general, we do not know the exact value of m; also, one should note that if m changes, the ML estimation also changes. Therefore, we suggest the following estimation algorithm that takes into consideration all possible values of m so that $x_m \le \theta \le x_{m+1}$:

Step 1. For each $m = 1,2,...,n-1,$ do evaluate $\hat{\delta}_1,...,\hat{\delta}_s,\hat{\theta}$ as solutions of the ML system

$$\begin{cases} \dfrac{\partial \ln L}{\partial \delta_i} = 0, \ i = 1,...s \\ \dfrac{\partial \ln L}{\partial \theta} = 0 \end{cases}, \quad (2.5)$$

If $x_m \le \hat{\theta} \le x_{m+1}$ then the ML estimations are

$$\hat{\delta}_i^{ML} = \hat{\delta}_i, i = 1,...,s, \ \hat{\theta}^{ML} = \hat{\theta}.$$

Step 2. If Step 1 does not give any solution for θ, then we are in one of two situations: m=n or m=0, hence we recommend using only f_1, and f_2, respectively, for the likelihood function.

Remark 2.1. With this algorithm, one has to check n-1 intervals, so that the computing time strongly depends on the magnitude of n.

3. The first composite Exponential-Pareto model-reinterpretation

It is not difficult to see that the composite Exponential–Pareto model (1.2), that we will call from now on the *first* composite Exponential–Pareto model, can be written in the form (2.1), with f_1 an Exponential pdf (hence f_1^* a right truncated Exponential pdf),

$$f_1(x) = \frac{1.35}{\theta} e^{-\frac{1.35x}{\theta}}, \quad x > 0 \Rightarrow f_1^*(x) = \frac{1.35}{\theta\left(1-e^{-1.35}\right)} e^{-\frac{1.35x}{\theta}} = \frac{1.82}{\theta} e^{-\frac{1.35x}{\theta}},$$

and f_2=f_2^* a Pareto pdf starting from θ, i.e.

$$f_2(x) = f_2^*(x) = \frac{0.35\theta^{0.35}}{x^{1.35}}, \quad x \ge \theta.$$

Therefore, the value of *r* results immediately as $r \cong 0.43$.
Thus, in our case, we can say that exactly 100*r*% ($\approx$43 %) of the observations are from an Exponential model truncated above at θ, and exactly 100($1-r$)% ($\approx$57%) of the observations are above θ, in accordance with a certain parameter restricted Pareto model. Hence, the model (1.2) can be interpreted as a two-component mixture model with *fixed* and *a priori* known mixing weights *r* and $1-r$. As $F_1(\theta)=1-e^{-1.35}$, it is easy to see that the truncation point θ is always the 0.74-th quintile of the original underlying Exponential model with density f_1.
Scollnik [6] observed that such a mixture model, with *fixed* and *a priori* known mixing weights *r* and $1-r$, is a very restrictive one. This is why in the following we extend this first Exponential–Pareto model to a general *r*.

4. The second composite Exponential-Pareto model

Using (2.1), we design a second composite Exponential-Pareto model as a truncated Exponential and Pareto mixture with threshold value θ, but *a priori* unrestricted mixing weights.

Proposition 4.1.
The second composite Exponential-Pareto density is given by

$$f(x)=\begin{cases} r\dfrac{\lambda e^{-\lambda x}}{1-e^{-\lambda\theta}}, & 0<x\le\theta \\ (1-r)\dfrac{\alpha\theta^{\alpha}}{x^{\alpha+1}}, & \theta<x\le\infty \end{cases}, \tag{4.1}$$

where the parameters $\lambda,\alpha,\theta>0$ *and* $0\le r\le 1$ *satisfy the continuity and differentiability conditions*

$$\lambda\theta=\alpha+1, \tag{4.2}$$

$$r=\frac{\alpha\left(1-e^{-\lambda\theta}\right)}{\alpha+e^{-\lambda\theta}}=\frac{\alpha\left(1-e^{-(\alpha+1)}\right)}{\alpha+e^{-(\alpha+1)}}. \tag{4.3}$$

Proof.
The pdf (4.1) results immediately from (2.1) and (2.2) by taking $f_1(x)=\lambda e^{-\lambda x}, x>0$ (i.e. an Exponential pdf) and $f_2(x)=\dfrac{\alpha\theta^{\alpha}}{x^{\alpha+1}}, x\ge\theta$ (i.e. a Pareto pdf), and noticing that $F_1(\theta)=1-e^{-\lambda\theta}, F_2(\theta)=0$. Hence,

$$f_1^*(x)=\frac{\lambda e^{-\lambda x}}{1-e^{-\lambda\theta}},\ \lambda>0 \text{ and } f_2^*(x)=\frac{\alpha\theta^{\alpha}}{x^{\alpha+1}},\ \alpha,\theta>0.$$

There are four unknown parameters: α, λ, θ and the mixing weight *r*. We would like the composite pdf to be smooth, so we will impose continuity and differentiability conditions at the threshold θ, i.e. $f(\theta - 0) = f(\theta + 0)$ and $f^{'}(\theta - 0) = f^{'}(\theta + 0)$, where $f^{'}$ is the first derivative of *f*.

From the first condition we obtain

$$r = \frac{\alpha\left(1 - e^{-\lambda\theta}\right)}{\lambda\theta e^{-\lambda\theta} + \alpha\left(1 - e^{-\lambda\theta}\right)}. \quad (4.4)$$

Observe that now the mixing weight *r* is not a fixed and known value as it was in the previous section. In this model, *r* is free to vary in the interval $[0, 1]$, its precise value depending on the particular values of α, λ and θ.

From the differentiability condition at θ we obtain

$$r = \frac{\alpha(\alpha + 1)\left(1 - e^{-\lambda\theta}\right)}{\lambda^2\theta^2 e^{-\lambda\theta} + \alpha(\alpha + 1)\left(1 - e^{-\lambda\theta}\right)}. \quad (4.5)$$

Condition (4.2) results easily by equating (4.4) and (4.5). Inserting (4.2) into the expression (4.4) of *r* gives

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

and, hence the condition (4.3) in both forms.

Remark 4.1. One may notice that because of the conditions (4.2)-(4.3), the number of unknown parameters can be reduced from four to only two (e.g., we can express *r* and *λ* in terms of *α* and *θ*). In order to reduce even more the number of free parameters of this model, we tried to impose a second derivative requirement $f^{''}(\theta - 0) = f^{''}(\theta + 0)$, which leads to the condition

$$\frac{r}{1 - e^{-\lambda\theta}}\lambda^3 e^{-\lambda\theta} = (1 - r)\alpha(\alpha + 1)(\alpha + 2)\frac{1}{\theta^3},$$

which together with (4.2) and (4.3) results in $\alpha + 1 = \alpha + 2 \Leftrightarrow 0 = 1$, which is impossible.

Remark 4.2. The second composite model is reduced to the first composite model for $\alpha \cong 0.35$ and $r \cong 0.43$.

In Figures 1 and 2, we plotted the pdf (4.1) for different values of the parameters *α* and *θ,* in order to see their effect on the tail of the distribution. One should notice that the tail of the second composite Exponential-Pareto distribution becomes heavier when *θ* increases (for fixed *α*) or when *α* decreases (for fixed *θ*).

Figure1

Pdf (4.1) for *θ*=10 and various values of *α*

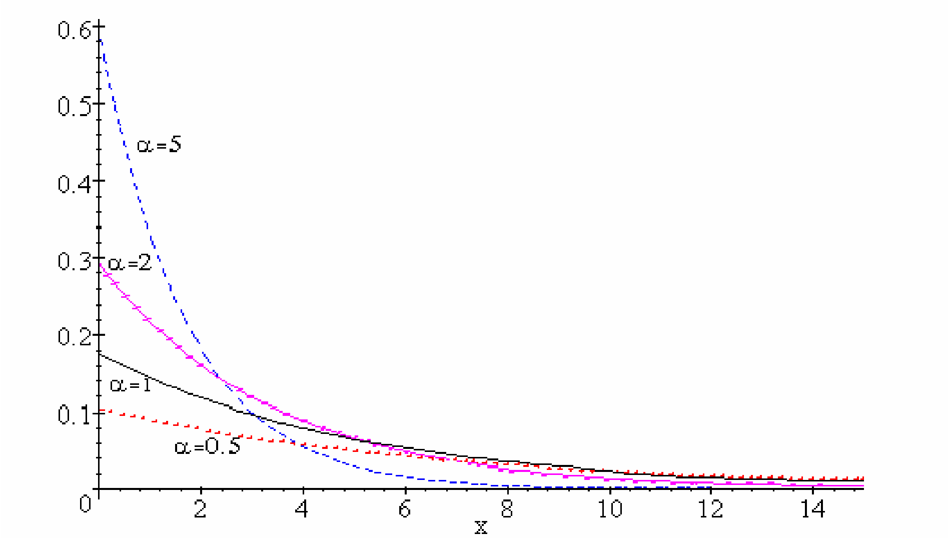
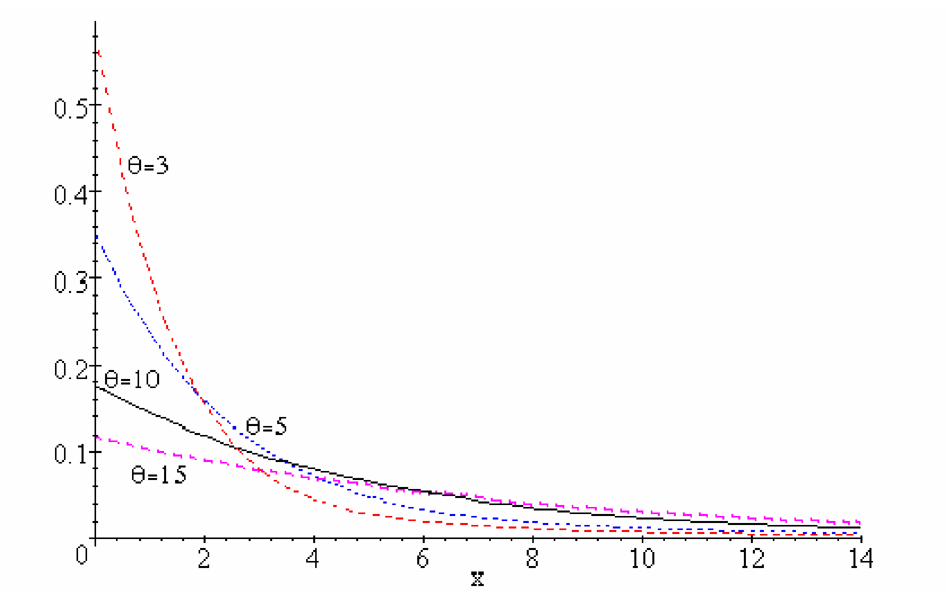


Figure 2

Pdf (4.1) for *α*=1 and various values of *θ*



Corollary 4.1.

The cdf of the second Exponential-Pareto model is given by

$$F(x)=\begin{cases} r\dfrac{1-e^{-\lambda x}}{1-e^{-\lambda\theta}}, & 0<x\le\theta \\ 1-(1-r)\left(\dfrac{\theta}{x}\right)^{\alpha}, & \theta\le x<\infty \end{cases}.$$

Proof. The formula results immediately from Proposition 2.1 and from

$$F_1(x)=1-e^{-\lambda x},\ F_2(x)=1-\left(\frac{\theta}{x}\right)^{\alpha}.$$

Corollary 4.2.

The initial n-th order moment of the second composite Exponential-Pareto distribution is given by

$$E_n(f)=r\frac{\Gamma(n+1,\lambda\theta)}{\lambda^n\left(1-e^{-\lambda\theta}\right)}+(1-r)\frac{\alpha\theta^n}{\alpha-n},\quad \alpha>n, \tag{4.6}$$

where: $\Gamma(\nu,t)=\int_0^t x^{\nu-1}e^{-x}dx,\ \nu,t>0,$ is the **lower incomplete gamma function**.

Proof.

We will use Proposition 2.2. We need the moments of the truncated Exponential distribution and of the Pareto distribution. We have

$$E_n(f_1^*)=\frac{\lambda}{1-e^{-\lambda\theta}}\int_0^{\theta}x^n e^{-\lambda x}dx=\frac{1}{\lambda^n\left(1-e^{-\lambda\theta}\right)}\int_0^{\lambda\theta}u^n e^{-u}du=\frac{\Gamma(n+1,\lambda\theta)}{\lambda^n\left(1-e^{-\lambda\theta}\right)}, \tag{4.7}$$

where we changed the variable $u=\lambda x$.

For $f_2^*(x)=f_2(x)=\dfrac{\alpha\theta^{\alpha}}{x^{\alpha+1}},\quad \alpha>0,\theta>0$, the initial *n-th* order moment is

$$E_n\left(f_2^*\right)=E_n\left(f_2\right)=\int_{\theta}^{\infty}x^n\cdot f_2(x)dx=\alpha\cdot\theta^{\alpha}\int_{\theta}^{\infty}x^{n-\alpha-1}dx=\frac{\alpha}{\alpha-n}\cdot\theta^n, \text{ for } \alpha>n.$$

Inserting them into Proposition 2.2 we get the stated result.

Remark 4.3. The values $\Gamma(n,\cdot)$ involved in formula (4.6), with *n* a positive integer, can be evaluated recursively as

$$\Gamma(n+1,x)=n\Gamma(n,x)-x^n e^{-x},\quad n\ge 1,\ x>0, \tag{4.8}$$

with starting value $\Gamma(1,x)=1-e^{-x},\ x>0$. Applying this recursion successively we also get

$$\Gamma(n+1,x)=n!-\sum_{k=0}^{n}\frac{n!}{(n-k)!}x^{n-k}e^{-x},\quad n\ge 1,\ x>0,$$

but using directly recursion (4.8) seems to be a more efficient method in this case.

Corollary 4.3.

The characteristic function of the second composite Exponential-Pareto distribution is given by

$$\varphi_f(t) = r\frac{\lambda\left(e^{(it-\lambda)\theta}-1\right)}{\left(1-e^{-\lambda\theta}\right)(it-\lambda)} + (1-r)\alpha\theta^{\alpha} I_{\alpha+1}, \quad t \in \mathrm{R},$$

where: $I_{\alpha+1} \overset{not}{=} \int_{\theta}^{\infty} x^{-(\alpha+1)} e^{itx} dx$.

Proof. We evaluate

$$\varphi_{f_1^*}(t) = \int_0^{\theta} e^{itx} \frac{\lambda e^{-\lambda x}}{1-e^{-\lambda\theta}} dx = \frac{\lambda\left(e^{(it-\lambda)\theta}-1\right)}{\left(1-e^{-\lambda\theta}\right)(it-\lambda)},$$

$$\varphi_{f_2^*}(t) = \int_{\theta}^{\infty} e^{itx} \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}} dx = \alpha\theta^{\alpha} \int_{\theta}^{\infty} e^{itx} x^{-(\alpha+1)} dx = \alpha\theta^{\alpha} I_{\alpha+1},$$

and insert them in Proposition 2.3.

Statistical inference. In order to apply the algorithm described at the end of Section 2, if $x_1 \le x_2 \le \ldots \le x_n$ is an ordered sample and $x_m \le \theta \le x_{m+1}$, the likelihood function (2.4) becomes in this case

$$L(x_1,\ldots,x_n;\alpha,\theta) = r^m (1-r)^{n-m} \prod_{i=1}^{m} \frac{\lambda e^{-\lambda x_i}}{1-e^{-\lambda\theta}} \prod_{j=m+1}^{n} \frac{\alpha\theta^{\alpha}}{x_j^{\alpha+1}}.$$

Replacing now *r* and λ in terms of *α* and *θ* as given in (4.2)-(4.3), and making some calculation we get

$$L(x_1,\ldots,x_n;\alpha,\theta) = \frac{\alpha^n(\alpha+1)^n \theta^{\alpha(n-m)-m} e^{-(\alpha+1)(n-m)}}{\left(\alpha+e^{-(\alpha+1)}\right)^n} \cdot \frac{e^{-\frac{\alpha+1}{\theta}S_m}}{(P_{n-m})^{\alpha+1}},$$

where: $S_m = \sum_{i=1}^{m} x_i$, $P_{n-m} = \prod_{j=m+1}^{n} x_j$. Hence, the likelihood system (2.5) is in this case

$$\begin{cases} \dfrac{\partial \ln L}{\partial\theta} = \dfrac{\alpha(n-m)-m}{\theta} + \dfrac{\alpha+1}{\theta^2} S_m = 0 \\ \dfrac{\partial \ln L}{\partial\alpha} = \dfrac{n(2\alpha+1)}{\alpha(\alpha+1)} - \dfrac{n(\alpha+1)}{\alpha+e^{-(\alpha+1)}} + (n-m)\ln\theta - \dfrac{S_m}{\theta} + m - \ln P_{n-m} = 0 \end{cases}.$$

From the first equation we obtain

$$\theta = \frac{(\alpha+1)S_m}{m-\alpha(n-m)}. \tag{4.9}$$

One may notice that in order to have the compulsory condition $\theta > 0$, the following condition results from (4.9)

$$m - \alpha(n-m) > 0 \Leftrightarrow \alpha < \frac{m}{n-m}. \qquad (4.10)$$

Inserting now (4.9) into the second equation of the likelihood system leads to the following equation in α

$$\frac{n(2\alpha+1)}{\alpha(\alpha+1)} - \frac{m-\alpha(n-m)}{\alpha+1} - \frac{n(\alpha+1)}{\alpha+e^{-(\alpha+1)}} + (n-m)\ln\frac{\alpha+1}{m-\alpha(n-m)} + \ln\frac{S_m^{n-m}}{P_{n-m}} + m = 0 \Leftrightarrow$$

$$\frac{n}{\alpha} - \frac{n(\alpha+1)}{\alpha+e^{-(\alpha+1)}} + (n-m)\ln\frac{\alpha+1}{m-\alpha(n-m)} + \ln\frac{S_m^{n-m}}{P_{n-m}} + n = 0. \qquad (4.11)$$

This equation must be numerically solved for α. Once we obtained a positive solution, we check if it satisfies the condition (4.10). If so, then we insert it in (4.9) and, hence, the value of θ results.

One should note that if in this case $\hat{\theta}^{ML}$ is closer to x_1 or x_n, it is better to choose the Pareto or Exponential models, respectively.

5. The third composite Exponential-Pareto model (The composite Exponential-type II Pareto model)

The third composite Exponential-Pareto model will also be developed in terms of the mixture model (2.1). This time we will make use of a version of the generalized Pareto distribution (GPD) above the threshold value θ, i.e. instead of $f_2^*(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}$, we consider $g(x) = \frac{\alpha(\alpha\beta)^\alpha}{(\alpha\beta-\theta+x)^{\alpha+1}}$, where $\theta > 0, \alpha > 0, \beta > 0, x > \theta$. This is also known as the Lomax or type II Pareto distribution, see e.g. Johnson *et al*. [2]. Now, letting $\gamma = \alpha\beta - \theta,\ \gamma > -\theta$, this density function may be written as

$$g(x) = \frac{\alpha(\gamma+\theta)^\alpha}{(\gamma+x)^{\alpha+1}}, \quad x > \theta. \qquad (5.1)$$

Proposition 5.1.

The third composite Exponential-Pareto density is given by

$$f(x) = \begin{cases} r\dfrac{\lambda e^{-\lambda x}}{1-e^{-\lambda\theta}}, & 0 < x \le \theta \\ (1-r)\dfrac{\alpha(\gamma+\theta)^\alpha}{(\gamma+x)^{\alpha+1}}, & \theta < x \le \infty \end{cases}, \qquad (5.2)$$

where the parameters $\lambda, \alpha, \theta > 0$, $\gamma > -\theta$ *and* $0 \le r \le 1$ *satisfy the continuity and differentiability conditions*

$$\lambda(\gamma + \theta) = \alpha + 1, \tag{5.3}$$

$$r = \frac{\alpha\left(1 - e^{-\lambda\theta}\right)}{\alpha + e^{-\lambda\theta}}. \tag{5.4}$$

Proof.

Like in the proof of Proposition 4.1, we immediately obtain the pdf (5.2) from (2.1) by replacing f_1^* with the truncated Exponential pdf, and f_2^* with the type II Pareto pdf given in (5.1). This pdf has five unknown parameters: $\alpha, \lambda, \theta, \gamma$ and the mixing weight *r*.

We now impose the continuity condition at the threshold point θ, i.e. $f(\theta - 0) = f(\theta + 0)$. This gives

$$r = \frac{\alpha\left(1 - e^{-\lambda\theta}\right)}{\lambda(\gamma + \theta)e^{-\lambda\theta} + \alpha\left(1 - e^{-\lambda\theta}\right)}. \tag{5.5}$$

Again the mixing weight *r* is not fixed and depends on the other four parameters.

We may ensure that the resulting density function is smooth if we also impose a differentiability condition at θ, so that $f'(\theta - 0) = f'(\theta + 0)$. This restriction gives

$$r = \frac{\alpha(\alpha + 1)\left(1 - e^{-\lambda\theta}\right)}{\lambda^2(\gamma + \theta)^2 e^{-\lambda\theta} + \alpha(\alpha + 1)\left(1 - e^{-\lambda\theta}\right)}. \tag{5.6}$$

From (5.5) and (5.6) we easily obtain the first condition, $\lambda(\gamma + \theta) = \alpha + 1$. Inserting this into (5.5) gives the expression (5.4) of *r*.

Remark 5.1. One may notice that because of the conditions (5.3)-(5.4), the number of unknown parameters was reduced from five to three (e.g., by expressing *r* and *λ* in terms of *α,γ* and *θ*). We tried to reduce further the number of free parameters of this model, by imposing a second derivative requirement $f''(\theta - 0) = f''(\theta + 0)$, which leads to the condition

$$\frac{r}{1 - e^{-\lambda\theta}}\lambda^3 e^{-\lambda\theta} = (1 - r)\alpha(\alpha + 1)(\alpha + 2)\frac{1}{(\gamma + \theta)^3},$$

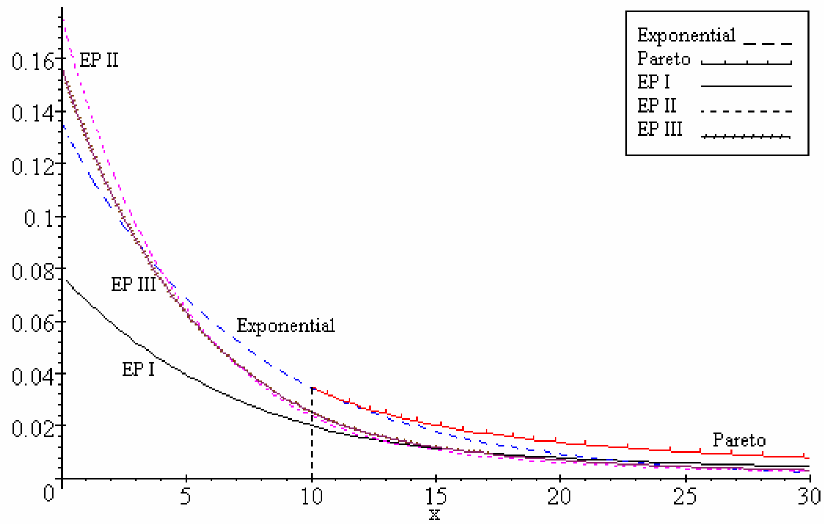
and using (5.3)-(5.4) we get $\alpha + 1 = \alpha + 2 \Leftrightarrow 0 = 1$, which is impossible.

Remark 5.2. The third composite model is reduced to the second one for $\gamma = 0$.

For comparison, in Figure 3 we plotted all three composite Exponential-Pareto pdf-s, and also the corresponding Exponential and Pareto pdf-s. One may notice that the second and the third composite Exponential-Pareto models are more flexible, because they have more parameters.

Figure 3

Exponential, Pareto and all three composite Exponential-Pareto (EP I, II, III) pdf curves for θ=10, α=γ=1



Corollary 5.1.

The cdf of the third Exponential-Pareto model is given by:

$$F(x)=\begin{cases} r\dfrac{1-e^{-\lambda x}}{1-e^{-\lambda\theta}}, & 0<x\le\theta \\ 1-(1-r)\left(\dfrac{\gamma+\theta}{\gamma+x}\right)^{\alpha}, & \theta\le x<\infty \end{cases} \tag{5.7}$$

Proof. We apply Proposition 2.1 with $F_1(x)=1-e^{-\lambda x}$ and

$$F_2(x)=\int_\theta^x \frac{\alpha(\gamma+\theta)^\alpha}{(\gamma+y)^{\alpha+1}}dy=(\gamma+\theta)^\alpha\left(\frac{1}{(\gamma+\theta)^\alpha}-\frac{1}{(\gamma+x)^\alpha}\right)=1-\left(\frac{\gamma+\theta}{\gamma+x}\right)^\alpha .$$

Thus, the cdf of the third composite model results as (5.7).

Corollary 5.2.

The initial n-th order moment of the third composite Exponential-Pareto distribution is given by

$$E_n(f)=r\frac{\Gamma(n+1,\lambda\theta)}{\lambda^n\left(1-e^{-\lambda\theta}\right)}+(1-r)\alpha\sum_{k=0}^{n}C_n^k\frac{(-\gamma)^{n-k}(\gamma+\theta)^k}{\alpha-k},\quad \alpha>n,$$

where Γ *is the lower incomplete gamma function as introduced in Corollary 4.2.*

Proof. We will use again Proposition 2.2. Like in the proof of Corollary 4.2, we get from (4.7): $E_n(f_1^*) = \dfrac{\Gamma(n+1,\lambda\theta)}{\lambda^n\left(1-e^{-\lambda\theta}\right)}$.

For $f_2^*(x) = g(x) = \dfrac{\alpha(\gamma+\theta)^\alpha}{(\gamma+x)^{\alpha+1}}$, $\alpha>0, \theta>0, \gamma>-\theta$, the initial *n-th* order moment is

$$E_n\left(f_2^*\right) = E_n(g) = \int_\theta^\infty x^n g(x)\,dx = \alpha(\gamma+\theta)^\alpha \int_\theta^\infty x^n (x+\gamma)^{-\alpha-1}\,dx. \qquad (5.8)$$

Denoting $I = \int_\theta^\infty x^n (x+\gamma)^{-\alpha-1}\,dx$ and changing the variable $y = x+\gamma$, we get

$$I = \int_{\gamma+\theta}^\infty \frac{(y-\gamma)^n}{y^{\alpha+1}}dy = \sum_{k=0}^n C_n^k(-\gamma)^{n-k}\int_{\gamma+\theta}^\infty y^{k-\alpha-1}dy = \sum_{k=0}^n C_n^k(-\gamma)^{n-k}\left.\frac{y^{k-\alpha}}{k-\alpha}\right|_{\gamma+\theta}^\infty =$$

$$= \sum_{k=0}^n C_n^k(-\gamma)^{n-k}\left(-\frac{(\gamma+\theta)^{k-\alpha}}{k-\alpha}\right) = \sum_{k=0}^n C_n^k \frac{(-\gamma)^{n-k}}{(\alpha-k)(\gamma+\theta)^{\alpha-k}}, \quad \alpha>n.$$

One should notice that the condition $\alpha > n$ is essential for the existence of this integral. Inserting this result into (5.8) we get

$$E_n\left(f_2^*\right) = \alpha(\gamma+\theta)^\alpha \sum_{k=0}^n C_n^k \frac{(-\gamma)^{n-k}}{(\alpha-k)(\gamma+\theta)^{\alpha-k}} = \alpha\sum_{k=0}^n C_n^k \frac{(-\gamma)^{n-k}(\gamma+\theta)^k}{\alpha-k}, \quad \alpha>n.$$

Using this and (4.7) in Proposition 2.2, the stated result is immediate.

Statistical inference. We want to apply again the algorithm described in Section 2. If $x_1 \le x_2 \le \ldots \le x_n$ is an ordered sample and $x_m \le \theta \le x_{m+1}$, for (5.2) the likelihood function (2.4) becomes

$$L(x_1,\ldots,x_n;\alpha,\gamma,\theta) = \prod_{i=1}^m r\frac{\lambda e^{-\lambda x_i}}{1-e^{-\lambda\theta}}\prod_{j=m+1}^n (1-r)\frac{\alpha(\gamma+\theta)^\alpha}{\left(\gamma+x_j\right)^{\alpha+1}}.$$

Replacing now *r* and λ in terms of *α,* γ and *θ* as given in (5.3)-(5.4), after some calculation we obtain

$$L(x_1,\ldots,x_n;\alpha,\gamma,\theta) = \frac{\alpha^n(\alpha+1)^n(\gamma+\theta)^{\alpha(n-m)-m}\, e^{-(\alpha+1)(n-m)\frac{\theta}{\gamma+\theta}}}{\left(\alpha+e^{-(\alpha+1)\frac{\theta}{\gamma+\theta}}\right)^n}\cdot\frac{e^{-\frac{\alpha+1}{\gamma+\theta}S_m}}{\left(P_{n-m}\right)^{\alpha+1}},$$

where: $S_m = \sum_{i=1}^{m} x_i$, $P_{n-m} = \prod_{j=m+1}^{n} \left(\gamma + x_j\right)$. Hence, the likelihood system is

$$\begin{cases} \dfrac{\partial \ln L}{\partial \theta} = \dfrac{\alpha(n-m)-m}{\gamma+\theta} + (S_m + m\gamma)\dfrac{\alpha+1}{(\gamma+\theta)^2} - \dfrac{\alpha(\alpha+1)n\dfrac{\gamma}{(\gamma+\theta)^2}}{\alpha + e^{-(\alpha+1)\frac{\theta}{\gamma+\theta}}} = 0 \\ \dfrac{\partial \ln L}{\partial \alpha} = \dfrac{n(2\alpha+1)}{\alpha(\alpha+1)} - \dfrac{n\left(\dfrac{\theta}{\gamma+\theta}\alpha+1\right)}{\alpha + e^{-(\alpha+1)\frac{\theta}{\gamma+\theta}}} + (n-m)\ln(\gamma+\theta) - \dfrac{S_m}{\gamma+\theta} + \dfrac{m\theta}{\gamma+\theta} - \ln P_{n-m} = 0 \\ \dfrac{\partial \ln L}{\partial \gamma} = \dfrac{\alpha(n-m)-m}{\gamma+\theta} + (S_m - m\theta)\dfrac{\alpha+1}{(\gamma+\theta)^2} + \dfrac{\alpha(\alpha+1)n\dfrac{\theta}{(\gamma+\theta)^2}}{\alpha + e^{-(\alpha+1)\frac{\theta}{\gamma+\theta}}} - (\alpha+1)\sum_{i=m+1}^{n}\dfrac{1}{\gamma+x_i} = 0 \end{cases}$$

This system looks very complicated. After some calculation, from the first and the third equations we get

$$\theta = \frac{S_m}{\sum_{i=m+1}^{n} \frac{\gamma}{\gamma + x_i} + \frac{n}{\alpha+1} + m - n} - \gamma, \tag{5.9}$$

while from the first and second equations we obtain

$$\alpha = \frac{n\gamma + m\theta - S_m}{\gamma \ln P_{n-m} + S_m + (n-m)(\theta - \gamma\ln(\gamma+\theta))}. \tag{5.10}$$

We then suggest the following approach: replace the ML system with the optimization problem

$$\begin{cases} \max\limits_{\alpha,\gamma,\theta} L(x_1,\ldots,x_n;\alpha,\gamma,\theta) \\ \alpha,\theta > 0 \\ \gamma > -\theta \end{cases}$$

with constraints (5.9), (5.10) and $x_m \le \theta \le x_{m+1}$. Such an optimization problem can be solved by metaheuristic approaches (we suggest, for instance, a Variable Neighborhood Search VNS algorithm, see Mladenovic and Hansen [5]).

6. Numerical examples – Model comparison

Let us now illustrate the estimation procedure described in Section 2. In Teodorescu and Vernic [7], we conducted a statistical study on two data samples generated from the Exponential-Pareto model (1.2), i.e. from the first composite Exponential-Pareto model, as called in this paper. In the following, we consider again these two data samples and try to fit also the newly introduced second composite Exponential-Pareto model (4.1). Since the first composite model was already fitted to these data in [7], we also present the corresponding results under the name of EP I (i.e., first Exponential-Pareto model), in order to compare the fitting of both models. The second composite Exponential-Pareto model is shortly denoted by EP II.

We would expect that, taking into consideration the randomness of the generating process, the second composite model fits better that the first one, because it allows for more variability (having more parameters).

The data analysis was realized using Excel and Mathcad software. In Mathcad, we solved equation (4.11) using the existing function Root that implements the secant method.

6.1 First example

The first data set was sampled from an Exponential-Pareto population (1.2) with parameter θ = 5, and has only n = 100 values. From Teodorescu and Vernic [7], the estimated value of the first composite Exponential-Pareto parameter by a MLE algorithm (described in [7]) is $\widetilde{\theta} = 5.427$.

Using the algorithm indicated for the second composite Exponential-Pareto model, we obtained the following estimated values:

$$m = 42; \quad \hat{\alpha} = 0.3983, \hat{\theta} = 6.5092, \hat{\lambda} = 0.2148, \hat{r} = 0.4647.$$

We also noticed that for a few values of m close to 42, the equation (4.11) has solution, but the corresponding solution of θ does not satisfy the condition $x_m \le \theta \le x_{m+1}$. Also, as soon as we get some distance of the good value of m (in both directions), the equation (4.11) has no real solution. Hence, at least for these data, we found only one corresponding solution.

In order to check the distribution fitting, we applied the χ^2 test using the empirical and theoretical frequencies given in Table 1 (we approximated the values at four decimals by rounding). The χ^2 distances, calculated as

$$d^2 = \sum \frac{(sample\ abs.freq. - n \times theor.freq.)^2}{n \times theor.freq.},$$

for both Exponential-Pareto distributions are:

First Exponential-Pareto model (from [7]): $d^2_{EP\,I} = 11.054$

Second Exponential-Pareto model: $d^2_{EP\,II} = 12.499$.

The χ^2 test accepts both models as expected, but in this case, the first Exponential-Pareto model gives the best fit, since it has the smallest χ^2 distance.

Table 1

Grouped data and theoretical frequencies

(column 2 results from the data sample, while columns 3 and 5 are calculated using the first and second Exponential-Pareto distribution functions, respectively)

Classes	Sample absolute freq., f_i	EP I freq., p_i	$\frac{(f_i - np_i)^2}{np_i}$	EP II freq., $\hat{p}_i$	$\frac{(f_i - n\hat{p}_i)^2}{n\hat{p}_i}$
[0, 1)	15	0.1263	0.4409	0.1193	0.7887
[1, 4)	18	0.2353	1.3019	0.2365	1.3515
[4, 8)	13	0.1371	0.0369	0.1511	0.2952
[8, 15)	10	0.0989	0.0010	0.1092	0.0776
[15, 30)	13	0.0866	2.1709	0.0926	1.5106
[30, 100)	9	0.1085	0.3154	0.1109	0.3952
[100, 300)	9	0.0660	0.8651	0.0639	1.0668
[300, 500)	6	0.0230	5.9089	0.0214	6.9443
[500, 7930)	7	0.0730	0.0128	0.0634	0.0693
Σ	*n* = 100		11.0543		12.4993

6.2 Second example

The second data sample, of *n* = 500, was taken from the first Exponential-Pareto model (1.2) with θ = 10. From Teodorescu and Vernic [7], the estimated value of the first composite Exponential-Pareto parameter by a MLE algorithm is $\widetilde{\theta} = 9.1069$. The estimation algorithm for the second composite Exponential-Pareto model gives the estimated values

$$m = 219; \quad \hat{\alpha} = 0.3426, \hat{\theta} = 8.9042, \hat{\lambda} = 0.1508, \hat{r} = 0.4192.$$

When running the algorithm, we noticed again that equation (4.11) together with condition $x_m \le \theta \le x_{m+1}$ has a unique solution. The χ^2 test (see Table 2 for involved frequencies, approximated by rounding) gives the distances:

First Exponential-Pareto model (from [7]): $d^2_{EP\,I} = 24.1939$

Second Exponential-Pareto model: $d^2_{EP\,II} = 23.7611$.

The χ^2 test accepts both models as expected, but this time, the second Exponential-Pareto model fits best, having the smallest χ^2 distance.

Table 2

Grouped data and theoretical frequencies

(the columns significance is the same as in Table 1)

Classes	Sample absolute freq., f_i	EP I freq. p_i	$\frac{(f_i - np_i)^2}{np_i}$	EP II freq. $\hat{p}_i$	$\frac{(f_i - n\hat{p}_i)^2}{n\hat{p}_i}$
[0, 3)	93	0.2060	0.9766	0.2064	1.0148
[3, 6)	78	0.1320	2.1645	0.1313	2.3143
[6, 9)	48	0.0846	0.7582	0.0835	0.9272
[9, 14)	47	0.0833	0.6755	0.0813	0.9949
[14, 25)	33	0.0906	3.3621	0.0896	3.1087
[25, 40)	34	0.0611	0.3844	0.0606	0.4462
[40, 100)	54	0.0938	1.0716	0.0935	1.1213
[100, 300)	24	0.0792	6.1486	0.0795	6.2534
[300, 600)	16	0.0363	0.2647	0.0368	0.3118
[600, 1500)	14	0.0363	0.9617	0.0370	1.0903
[1500, 2500)	11	0.0157	1.2422	0.0161	1.0813
[2500, 10000)	15	0.0309	0.0136	0.0318	0.0523
[10000, 30000)	11	0.0158	1.2143	0.0164	0.9479
[30000, 10^5)	11	0.0115	4.6748	0.0121	3.9992
[10^5, $21806 \cdot 10^3$)	11	0.0187	0.2804	0.0200	0.0972
Σ	*n* = 500		24.1939		23.7611

7. Conclusions

In order to minimize financial losses associated with uncertain undesirable insurance events, actuaries must forecast losses using models of random events. Therefore, modeling the claim distribution is of great importance and researchers are trying to find new models that fit better to insurance data.

In this paper, we proposed three composite Exponential-Pareto models as alternatives to the composite Lognormal-Pareto models studied by Scollnick [6]. These composite Exponential-Pareto models have attractive properties concerning the pdf, cdf, moments, likelihood function, etc. The second and the third models are more flexible than the first one because they have more parameters. On the other hand, as compared to the third model, the second one can be estimated easier, while its shape is similar.

As future plans, considering the fact that in Section 2 we presented the general framework for creating composite models, we aim at building other composite models as well, such as Gamma-Pareto or Weibull-Pareto models. These models could be used in situations where other models underestimate the tail probability or, in other words, in situations where other models underestimate the premium to be paid in case of large losses.

Acknowledgment. Raluca Vernic acknowledges the support given by the research grant GAR 24/2007-2008.

References

[1] Cooray, K, Ananda, M.A. (2005), "Modeling actuarial data with a composite Lognormal-Pareto model". *Scandinavian Actuarial Journal,* 5: 321-334.

[2] Johnson, N. L., Kotz, S., Balakrishnan, N. (1994), *Continuous Univariate Distributions*. Wiley, New York.

[3] Kaas, R., Goovaerts, M., Denuit, M., Dhaene, J. (2001), *Modern Actuarial Risk Theory.* Kluwer Academic Publishers, Boston.

[4] Klugman, S.A., Panjer, H.H., Willmot, G.E. (2004), *Loss models: from data to decisions* (2nd edition). New York, John Wiley & Sons, Inc.

[5] Mladenovic, N., Hansen, P. (1997), "Variable Neighborhood Search". *Computers and Operations Research,* 24: 1097-1100.

[6] Scollnik, D.P.M. (2007), "On composite Lognormal-Pareto models". *Scandinavian Actuarial Journal,* 1: 20-33.

[7] Teodorescu, S., Vernic, R. (2006), "A composite Exponential–Pareto distribution". *The Annals of the "Ovidius" University of Constanta*, Mathematics Series, vol.XIV (1): 99-108.