PARETO TYPE DISTRIBUTIONS AND EXCESS-OF-LOSS REINSURANCE

Werner Hürlimann
Wolters Kluwer Financial Services Switzerland AG, Seefeldstrasse 69, CH-8008 Zürich

ABSTRACT
To be consistent with Extreme Value Theory the pricing of excess-of-loss reinsurance contracts should be based on Pareto type distributions. In this context, two recent Pareto type distributions are considered and compared. The state of a geometric Brownian motion after an exponentially distributed random time with log-normally distributed initial state generates a four-parameter Pareto type distribution, called double Pareto lognormal distribution. It has been introduced by Reed and Jorgensen and exhibits two-sided power-law behaviour. On the other hand, the mixture of a truncated exponential distribution in the lower-tail with a Pareto distribution in the upper-tail generates another analytical simple four-parameter Pareto type distribution, called exponential Pareto distribution. The statistical fitting of these two models is illustrated with a large claims data set and its application to the pricing of excess-of-loss reinsurance in actuarial science is illustrated. As a main result, it is shown that the selected best fitted right-tailed Pareto lognormal distribution with minimum Cramér-von Mises K-statistics yields the safest mean and standard deviation of excess-of-loss reinsurance layer risks. The corresponding method is useful for conservative excess-of-loss reinsurance pricing.

Keywords: *geometric Brownian motion, Pareto distribution, exponential distribution, log-normal distribution, excess-of-loss reinsurance, aggregate claims*

1. INTRODUCTION

The modeling and fitting of statistical distributions to loss or return data in actuarial science and finance is an important subject of wide interest. Well-known books from actuarial science include Hogg and Klugman [1], Panjer and Willmot [2] and Klugman et al. [3]. The usefulness in both actuarial science and finance of some three-parameter tractable analytical distributions like the normal inverted gamma mixture (Pearson type VII, generalized Student t) and the symmetric double Weibull distribution has been pointed in Hürlimann [4], [5].

This open subject has always been of interest and generates a lot of interesting statistical research. A main purpose of the present case study is further simplification and improvement. Whereas important applications are considered, the appearance of new phenomena leads to new questions. For example, it can sometimes be noticed that the preferred model among a class of fitted models leads to the highest mean and standard deviation of excess-of-loss reinsurance layer risks. In this situation, the selected model is consistent with the actuarial principle of safe or conservative pricing.

The desire to be consistent with Extreme Value Theory (EVT) in excess-of-loss reinsurance pricing has been recognized at an early time and various Pareto type claim size distributions have been used so far, among others the Benktander distributions (Benktander and Segerdahl [6], Benktander [7]), the log-gamma and the Burr distributions (Hogg and Klugman [1], Mack [8], p. 365).

In the present paper, two further Pareto type distributions are considered. In theory, several temporal stochastic phenomena can be modeled appropriately using a geometric Brownian motion (GBM), e.g. Black-Scholes option pricing, firm sizes, city sizes and individual incomes. In practice, the empirical data of such phenomena exhibits power-law behaviour, which contradicts the log-normal distribution underlying GBM. Recall the following reconciliation. A simple mechanism, which generates the power-law behaviour in the tails, consists to assume that the time of observation in a GBM is itself a random variable, whose distribution is close to an exponential distribution. For example, the state of a GBM after an exponentially distributed random time with fixed initial state generates the double Pareto distribution introduced in Reed [9]. A natural generalization consists to look at the state of a GBM after an exponentially distributed random time with log-normally distributed initial state. It generates a double Pareto lognormal distribution, which has been considered in Reed [10] and Reed and Jorgensen [11]. On the other hand, a practical and straightforward way to generalize the Pareto distribution consists to fit the lower tail using another two-parameter analytical distribution, for example a translated exponential distribution. The combination is the analytical four-parameter exponential Pareto distribution. The sketched analytical and tractable distributions are used as large claims distributions in the context of excess-of-loss reinsurance pricing. A more detailed content of the paper follows.

Section 2 describes the steps necessary to derive the double Pareto lognormal distribution from the geometric Brownian motion. Closed-form analytical expressions for the density, distribution and moments (in case they exist) of the distribution are given. Limiting cases include Pareto lognormal type distributions, which exhibit power-law behaviour in either the right-tail or the left-tail of the distribution. Section 3 discusses the application of Pareto type claim size distributions to excess-of-loss reinsurance. To fit the claim size distributions to data, we use three standard methods, namely maximum likelihood estimation (MLE), minimum chi-square estimation ($\min\chi 2$) and minimization of the Cramér-von Mises K-statistic (minK). Closed-form analytical expressions for the mean and standard deviation of excess-of-loss reinsurance layer losses are given in Section 4. A detailed case study based on a real-life sample of large claims is presented. We note that the selected best fitted right-tailed Pareto lognormal with minimum Cramér-von Mises K-statistics yields the safest mean and standard deviation.

2. FROM THE GEOMETRIC BROWNIAN MOTION TO THE DOUBLE PARETO LOG-NORMAL

It is often assumed that the time evolution of a stochastic phenomenon X_t involves a variable but size independent proportional growth rate and can thus be modeled by a geometric Brownian motion (GBM) described by the stochastic differential equation

$$dX = \mu \cdot X \cdot dt + \sigma \cdot X \cdot dW\,, \tag{1}$$

where dW is the increment of a Wiener process. Since the proportional increment of a GBM in time dt has a systematic component $\mu \cdot dt$ and a white noise component $\sigma \cdot dW$, GBM can be viewed as a stochastic version of a simple exponential growth model. The GBM has long been used to model the evolution of stock prices (Black-Scholes option pricing model), firm sizes, city sizes and individual incomes. It is well-known that empirical studies on such phenomena often exhibit power-law behaviour. However, the state of a GBM after a fixed time T follows a lognormal distribution, which does not exhibit power-law behaviour.

Why does one observe power-law behaviour for phenomena apparently evolving like a GBM? A simple mechanism, which generates the power-law behaviour in the tails, consists to assume that the time of observation T is a random variable, whose distribution is close to an exponential distribution. This simple model has been proposed by Reed [9]. The distribution of X_T with fixed initial state X_0 is described by the double Pareto distribution $dP(X_0, \alpha, \beta)$ with density function

$$f_{X_T}(x) = \begin{cases} \left(\dfrac{\alpha\beta}{\alpha+\beta}\right)\dfrac{1}{X_0}\left(\dfrac{x}{X_0}\right)^{\beta-1}, & x \le X_0, \\ \left(\dfrac{\alpha\beta}{\alpha+\beta}\right)\dfrac{1}{X_0}\left(\dfrac{x}{X_0}\right)^{-\alpha-1}, & x \ge X_0, \end{cases} \tag{2}$$

where $\alpha, \beta > 0$, and $\alpha, -\beta$ are the positive roots of the characteristic equation

$$\left(\mu - \frac{1}{2}\sigma^2\right)z + \frac{1}{2}\sigma^2 z^2 = \lambda\,, \tag{3}$$

where λ is the parameter of the exponentially distributed random variable T.

A natural generalization of the double Pareto law has been introduced in Reed [10] and Reed and Jorgensen [11]. They consider the mixture of log-normal distributions arising from a GBM with log-normally distributed initial state $\ln X_0 \sim N(\nu, \tau^2)$ (instead of a fixed initial state as above), which is killed or stopped with the constant killing rate λ. Equivalently, this is the distribution of the state of a GBM after an exponentially distributed time of evolution. For completeness, let us derive shortly the form of this distribution. At fixed time t, the logarithm of this GBM is $Y_t = \ln X_t \sim N\left(\nu + \left(\mu - \frac{1}{2}\sigma^2\right)t, \tau^2 + \sigma^2 t\right)$, and has the moment generating function (mgf)

$$M_{Y_t}(s) = \exp\left\{\nu s + \tfrac{1}{2}\tau^2 s^2 + [\left(\mu - \tfrac{1}{2}\sigma^2\right)s + \tfrac{1}{2}\sigma^2 s^2]t\right\}. \tag{4}$$

Let now T be exponentially distributed with mgf

$$M_T(s) = \frac{\lambda}{\lambda - s}. \tag{5}$$

Then the state $Y_T = \ln X_T$ of an ordinary Brownian motion after the random time T has mgf

$$M_{Y_T}(s) = E_T[M_{Y_t}(s) | t = T] = e^{\nu s + \frac{1}{2}\tau^2 s^2} \cdot \frac{\alpha\beta}{(\alpha - s)(\beta + s)}, \tag{6}$$

where $\alpha, -\beta$ are the positive roots of (3). This shows that $Y_T = Z + W$ is the independent sum of a normally distributed random variable $Z \sim N(\nu, \tau^2)$ and a double exponential distribution or skewed Laplace distribution (Kotz et al. [12]) W with probability density function (pdf)

$$f_W(w) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} e^{\beta w}, & w \le 0, \\ \frac{\alpha\beta}{\alpha+\beta} e^{-\alpha w}, & w > 0. \end{cases} \tag{7}$$

Through convolution one obtains the pdf

$$g(y) = f_{Y_T}(y) = \frac{\alpha\beta}{\alpha+\beta} \varphi\left(\frac{y-\nu}{\tau}\right) [R\left(\alpha\tau - \frac{y-\nu}{\tau}\right) + R\left(\beta\tau + \frac{y-\nu}{\tau}\right)], \tag{8}$$

where $R(\cdot)$ denotes Mill's ratio of the complementary cumulative distribution function (cdf) to the pdf of a standard normal distribution given by

$$R(z) = \frac{\overline{\Phi}(z)}{\varphi(z)}, \quad \overline{\Phi}(z) = 1 - \Phi(z).. \tag{9}$$

By construction, it is natural to call this a Normal-Laplace distribution and write $Y_T \sim NL(\alpha, \beta, \tau, \nu^2)$. The cdf of the Normal-Laplace satisfies the analytical expression

$$G(y) = \Phi\left(\frac{y-\nu}{\tau}\right) - \varphi\left(\frac{y-\nu}{\tau}\right) \frac{\beta \cdot R\left(\alpha\tau - \frac{y-\nu}{\tau}\right) - \alpha \cdot R\left(\beta\tau + \frac{y-\nu}{\tau}\right)}{\alpha+\beta} \tag{10}$$

Since a Laplace random variable is an independent difference of two exponential random variables, one has the alternative representation (equality in distribution)

$$Y_T =_d \nu + \tau Z + \frac{1}{\alpha} E_1 - \frac{1}{\beta} E_2, \tag{11}$$

where E_1, E_2 are independent standard exponential random variables and Z is standard normal random variable, which is independent of E_1, E_2. It follows immediately that $X_T = \exp(Y_T)$ has the pdf

$$f(x) = f_{X_T}(x) = \frac{1}{x} g(\ln x), \quad x > 0. \tag{12}$$

Since $X_T = e^Z \cdot e^W$ is the independent product of a lognormal and a double Pareto distribution, it is natural to call this a double Pareto-lognormal distribution and write $X_T \sim dPLN(\alpha, \beta, \tau, \nu^2)$. Alternatively, the pdf (12) can be rewritten as

$$f(x) = \frac{\alpha\beta}{\alpha+\beta}\left\{ A(\alpha,\nu,\tau)x^{-\alpha-1}\Phi\left(\frac{\ln x-\nu}{\tau}-\alpha\tau\right) + A(-\beta,\nu,\tau)x^{\beta-1}\overline{\Phi}\left(\frac{\ln x-\nu}{\tau}+\beta\tau\right)\right\}, \tag{13}$$

with $A(\theta,\nu,\tau) = \exp\left(\theta\nu + \frac{1}{2}\theta^2\tau^2\right)$. The cdf of the double Pareto-lognormal satisfies the analytical expression

$$\begin{aligned} F(x) &= G(e^x) = \Phi\left(\frac{\ln x-\nu}{\tau}\right) \\ &-\frac{1}{\alpha+\beta}\left\{ \beta x^{-\alpha}A(\alpha,\nu,\tau)\Phi\left(\frac{\ln x-\nu}{\tau}-\alpha\tau\right) - \alpha x^{\beta-1}A(-\beta,\nu,\tau)\overline{\Phi}\left(\frac{\ln x-\nu}{\tau}+\beta\tau\right)\right\}. \end{aligned} \tag{14}$$

Some comments on this Pareto type distribution are in order. It is similar in shape to the log-hyperbolic distribution introduced by Barndorff-Nielsen [13] in connection with the “sand project” (investigation of the physics of wind-blown sand), which also exhibits Pareto behaviour asymptotically in both tails. Like the log-hyperbolic, it can be derived as a mixture of log-normal distributions. Indeed, it has been noticed that generalized hyperbolic distributions also arise as killed GBM's like the dPLN, however with more complicated killing rate functions (e.g. Eberlein [14]). In general, the "killing" property of stochastic processes has been studied in the book by Karlin and Taylor [15]. Compared to the log-hyperbolic, the dPLN has the advantage to be analytically more tractable, which makes it a very attractive from a practical viewpoint. Since the Normal-Laplace is infinitely divisible, it is possible to construct a Lévy process with increments following a Normal-Laplace distribution. The rich class of Lévy processes, extensively studied in Barndorff-Nielsen [16], can be used to model the logarithmic return of financial instruments. The attractive Normal-Laplace can thus be used to reflect the empirical fact that observed logarithmic returns for high frequency data have fatter tails than those of a normal distribution. In this setting, option pricing formulas can be derived applying the characteristic function approach (e.g. Schoutens [17], p. 20).

In the following, some useful properties of the Normal-Laplace and double Pareto-lognormal distributions are listed. The two special cases, obtained from the limiting cases $\alpha \to \infty$ and $\beta \to \infty$, are of main interest. They define two Normal-Exponential distributions, the right-tailed Normal-Exponential distribution $Y_1 \sim NEr(\alpha,\nu,\tau^2)$ and the left-tailed Normal-Exponential distribution $Y_2 \sim NEl(\beta,\nu,\tau^2)$ with pdf’s

$$g_1(y) = \alpha\varphi\left(\frac{y-\nu}{\tau}\right)R\left(\alpha\tau - \frac{y-\nu}{\tau}\right), \quad g_2(y) = \beta\varphi\left(\frac{y-\nu}{\tau}\right)R\left(\beta\tau + \frac{y-\nu}{\tau}\right). \tag{15}$$

With (8) these formulas show that the Normal-Laplace satisfies the mixture representation

$$g(y) = \frac{\beta}{\alpha+\beta}g_1(y) + \frac{\alpha}{\alpha+\beta}g_2(y). \tag{16}$$

Similarly, the limiting cases $\alpha \to \infty$ and $\beta \to \infty$ define two Pareto-lognormal distributions, the right-tailed Pareto-lognormal distribution $Y_1 \sim PLNr(\alpha,\nu,\tau^2)$ and the left-tailed Pareto-lognormal distribution $Y_2 \sim PLNl(\alpha,\nu,\tau^2)$ with pdf’s

$$f_1(x) = \alpha x^{-\alpha-1}A(\alpha,\nu,\tau)\Phi\left(\frac{\ln x-\nu}{\tau}-\alpha\tau\right), \quad f_2(x) = \beta x^{\beta-1}A(-\beta,\nu,\tau)\overline{\Phi}\left(\frac{\ln x-\nu}{\tau}+\beta\tau\right). \tag{17}$$

With (13) these formulas show that the double Pareto-lognormal satisfies the mixture representation

$$f(x) = \frac{\beta}{\alpha+\beta} f_1(x) + \frac{\alpha}{\alpha+\beta} f_2(x). \tag{18}$$

The mean, the variance and the third and fourth order cumulants of the Normal-Laplace are given by

$$\begin{aligned} &E[Y] = \nu + \alpha^{-1} - \beta^{-1}, \quad Var[Y] = \tau^2 + \alpha^{-2} + \beta^{-2}, \\ &\kappa_3 = 2\alpha^{-3} - 2\beta^{-3}, \quad \kappa_4 = 6\alpha^{-4} + 6\beta^{-4}. \end{aligned} \tag{19}$$

The r -th moment of the double Pareto-lognormal exists only if $r < \alpha$ and is given by

$$\mu_r = E[X^r] = \frac{\alpha\beta}{(\alpha - r)(\beta + r)} \exp\left(r\nu + \frac{1}{2} r^2 \tau^2 \right). \tag{20}$$

3. PARETO TYPE CLAIM SIZE DISTRIBUTIONS IN EXCESS-OF-LOSS REINSURANCE

In the framework of the classical *collective model* of risk theory, the *aggregate claims* of a portfolio of insurance risks are described by the random variable

$$X = \sum_{i=1}^{N} Y_i , \tag{21}$$

where the *claim sizes* Y_i are independent and identically distributed and independent from the random *claim number* N. It is assumed that the random variables Y_i are non-negative.

An *excess-of-loss* or *XL-reinsurance* contract with *deductible* d on a portfolio of risks covers for each claim Y_i the *excess claim size* $(Y_i - d)_+$, $i = 1,\dots,N$. In this setting, the *aggregate claims of the XL-reinsurance* are described by the random variable denoted by

$$X(d) = \sum_{i=1}^{N} (Y_i - d)_+ . \tag{22}$$

In practice, an XL-reinsurance contract will be limited to a maximum payment. If $L > d$ denotes the upper limit of the individual claims, which is covered by the contract, then the aggregate claims of the *XL-reinsurance layer with deductible d and limit L* is defined and denoted by

$$X(d, L) = X(d) - X(L) = \sum_{i=1}^{N} \left\{ (Y_i - d)_+ - (Y_i - L)_+ \right\}. \tag{23}$$

According to (23) the risk of the XL-reinsurance layer has a distribution from a collective model of risk theory. However, a risk-manager and/or a reinsurer, which does not know the number and the size of the original claims below the deductible d , will not be able to analyse this risk satisfactorily. Therefore, the collective model (23) is not appropriate to forecast the excess-of-loss risk of the layer. Fortunately, it is possible to construct a collective model for the layer on the basis of the collective model for the original claims such that the model contains only random variables which are observable for the reinsurer. This collective model is presented in Hess [18] and the related literature in Hess et al. [19], Franke and Macht [20], Mack [8] and Schmidt [21], [22].

Suppose that the analysis of the layer risk relies only on the data containing the claims above a limit OP, called *observation point*. This means that only forecasts for XL-reinsurance with a deductible $d \ge OP$ can be made. Denote by Y a random variable distributed as Y_i for all $i = 1,\dots,N$. It is assumed that $P(Y > OP) > 0$. In other words, the probability that a claim exceeds OP is strictly positive. The aggregate claims at disposal, which contains all claims exceeding OP, is described by the collective model of risk theory

$$X^{OP} = \sum_{i=1}^{N^{op}} Y_i^{OP} . \tag{24}$$

In this expression N^{OP} counts the number of claims above OP and is given by

$$N^{OP} = \sum_{i=1}^{N} B_i , \tag{25}$$

where the $B_i's$ are independent and identically distributed Bernoulli random variables, which are independent from N, such that $P(B_i = 1) = P(Y > OP)$. The claim sizes Y_i^{OP} are independent and identically distributed, and independent from N, and each is distributed like Y^{OP} with distribution $P(Y^{OP} \le x) = P(Y \le x | Y > OP), \quad x > OP$. This is the basic collective model of risk theory used in XL-reinsurance.

It is well-known that the two-parameter Pareto distribution is an appropriate distribution often used to fit large claims distributions in (re)insurance. This has been a first choice in the practice of reinsurance for a long time (see e.g. Schmitter [23], Schmitter and Bütikofer [24], Doerr [25], Schmutz and Doerr [26]) and it is consistent with the theoretical results from Extreme Value Theory (e.g. Embrechts et al. [27]).

Once the large claims Pareto distribution has been fitted in an adequate way, one often observes a rather poor fit in the lower tail of the distribution. To remedy for this disadvantage several Pareto type distributions can be used. As argued in the previous Section, the double Pareto lognormal distribution, or the simpler right-tailed Pareto lognormal might be an attractive and appropriate model. For comparison and decision purposes, other models have to be taken into account. A practical and straightforward way to generalize the Pareto distribution consists to fit the lower tail using another simple two-parameter analytical distribution, for example a translated exponential distribution.

The considered combined four-parameter *exponential Pareto* distribution, written $EP(\alpha, \beta, T, \gamma)$ takes the form

$$F(x) = \begin{cases} 1 - \exp\left(-\dfrac{x-\alpha}{\beta}\right), & \alpha \le x \le T, \\ 1 - \exp\left(-\dfrac{T-\alpha}{\beta}\right) \cdot \left(\dfrac{x}{T}\right)^{-\gamma}, & x \ge T. \end{cases} \tag{26}$$

To fit claim size distributions to data, we use in the following case study three standard methods, namely maximum likelihood estimation (MLE), minimum chi-square estimation ($\min \chi 2$) and minimization of the Cramér-von Mises *K*-statistic (minK). The latter method is especially recommended to judge the goodness-of-fit of the tails of distributions (e.g D'Agostino and Stephens [28]). A main advantage of the proposed Pareto type claim size distributions and fitting methods are their full analytical tractability and the fast and numerical stable evaluation of all required quantities using computer implemented minimization algorithms.

Since the claim size densities have explicit analytical expressions, the required log-likelihood functions for MLE are explicit. The used chi-square statistic and *K*-statistic for $\min \chi 2$ and minK are defined as follows. Given a random sample $x_1 \le x_2 \le \ldots \le x_n$ of claims data (order statistics of the claims data), let $y_i = \frac{i-1}{n-1} \in [0,1], \quad i = 1, \ldots, n$, be the corresponding percentile ranks. The fitted values of the distribution function are denoted by $\hat{y}_i = F(x_i), \quad i = 1, \ldots, n$. The chi-square value is defined by

$$\chi_2 = \sum_{i=2}^{n} \frac{\left[y_i - y_{i-1} - (\hat{y}_i - \hat{y}_{i-1})\right]^2}{\hat{y}_i - \hat{y}_{i-1}} , \tag{27}$$

and the Cramér-von Mises *K*-statistic by

$$K = n \cdot \sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{\hat{y}_i \cdot (1 - \hat{y}_i)} . \tag{28}$$

4. A CASE STUDY IN EXCESS-OF-LOSS REINSURANCE

According to the collective model for XL-reinsurance introduced in Section 3, we assume for ease of notation that the aggregate claims of a portfolio of insurance risks are described by the random variable $X = \sum_{i=1}^{N_T} Y_i$, where the *claim sizes* Y_i are independent and identically distributed like Y and independent from the random *claim number* N_T. The variable N_T counts the number of claims above a fixed threshold T and its mean is denoted here $\lambda = E[N_T]$. For simplicity, we assume that N_T is Poisson distributed with parameter λ, which is a reasonable assumption as soon as the expected number of claims is relatively small.

The mean and standard deviation of an XL-reinsurance layer with deductible $d > T$ and upper limit $L > d$ will be evaluated according to (23) with N replaced by N_T. These quantities are formally given by

$$\begin{aligned} &E[X(d,L)] = \lambda \cdot (m(d) - m(L)), \\ &Var[X(d,L)] = \lambda \cdot (m2(d) - m2(L) - 2 \cdot (L-d) \cdot m(L)), \\ &m(x) = \int_x^\infty \overline{F}(t)dt, \quad m2(x) = 2 \cdot \int_x^\infty m(t)dt. \end{aligned} \tag{29}$$

For the exponential Pareto claim size distribution (26), the required formal expressions are obtained in a straightforward way and in case $x > T$ they are given by

$$\begin{aligned} &m(x) = \exp[-\left(\frac{T-\alpha}{\beta}\right)] \cdot \frac{T}{\gamma - 1} \cdot \left(\frac{x}{T}\right)^{-(\gamma-1)}, \\ &m2(x) = \exp[-\left(\frac{T-\alpha}{\beta}\right)] \cdot \frac{2T^2}{(\gamma-1)(\gamma-2)} \cdot \left(\frac{x}{T}\right)^{-(\gamma-2)}. \end{aligned} \tag{30}$$

In the present case study, the three-parameter right-tailed Pareto lognormal claim size distribution $Y \sim PLNr(\alpha, \nu, \tau^2)$ introduced in (17) will be compared with the exponential Pareto claim size. Its distribution function, obtained from (14) with $\beta \to \infty$, is given by

$$F(x) = \Phi\left(\frac{\ln x - \nu}{\tau}\right) - A(\alpha, \nu, \tau) x^{-\alpha} \Phi\left(\frac{\ln x - \nu}{\tau} - \alpha\tau\right). \tag{31}$$

Using partial integration and somewhat tedious but straightforward calculations, one obtains the following formulas

$$m(x) = E[Y] \cdot \overline{\Phi}\left(\frac{\ln x - \nu}{\tau} - \tau\right) - x \cdot \overline{\Phi}\left(\frac{\ln x - \nu}{\tau}\right) + \frac{A(\alpha, \nu, \tau)}{\alpha - 1} x^{-(\alpha-1)} \cdot \Phi\left(\frac{\ln x - \nu}{\tau} - \alpha\tau\right). \tag{32}$$

$$\begin{aligned} m2(x) = {} & E[Y^2] \cdot \overline{\Phi}\left(\frac{\ln x - \nu}{\tau} - 2\tau\right) - 2E[Y]x \cdot \overline{\Phi}\left(\frac{\ln x - \nu}{\tau} - \tau\right) \\ & + x^2 \cdot \overline{\Phi}\left(\frac{\ln x - \nu}{\tau}\right) + \frac{2A(\alpha, \nu, \tau)}{(\alpha-1)(\alpha-2)} x^{-(\alpha-2)} \cdot \Phi\left(\frac{\ln x - \nu}{\tau} - \alpha\tau\right). \end{aligned} \tag{33}$$

Samples of large claims of bigger size are scarce in practice. Therefore, a case study based on a real-life sample of large claims of size 49 has been undertaken. Table 1 summarizes the results of the chosen fitting procedure.

Table 1. Exponential Pareto and Pareto lognormal fit to data

Exponential Pareto

Method	Scale	Location	Observation Point	Pareto index	Chi-square	K statistic
MLE	1'210'512	525'000	1'000'000	1.50763	0.573	7.74
Min K	1'354'622	525'000	1'000'000	1.58931	0.588	5.36
Min Chi^2	1'188'357	525'000	1'000'000	1.43927	0.571	9.17

Pareto lognormal

Method	Scale - tau	Location - vu	Pareto index	Chi-square	K statistic
MLE	0.06832	13.54312	1.60671	114.835	11'859.23
Min K	0.31052	13.54432	1.60671	0.608	7.89
Min Chi^2	0.34106	13.50428	1.60671	0.595	15.21

The chi-square values and K-statistics for the exponential Pareto are excellent for all three fitting methods. For the Pareto lognormal the MLE do not lead to low chi-square values and K-statistics albeit the graph of the fit looks very good. The Figures 1 and 2 show graphs of these best fitting claim size distributions. No significant differences between the Exponential Pareto and Pareto lognormal best fits are observed. Therefore, we conclude that the more parsimonious three-parameter Pareto lognormal should be preferred.

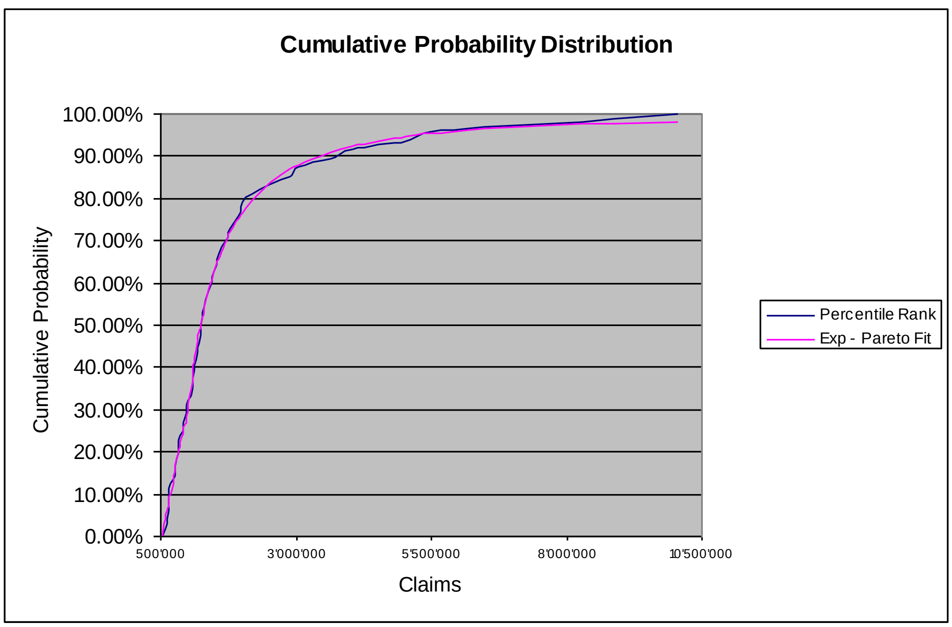

Figure 1. Graph of the Exponential Pareto fit by the minK method

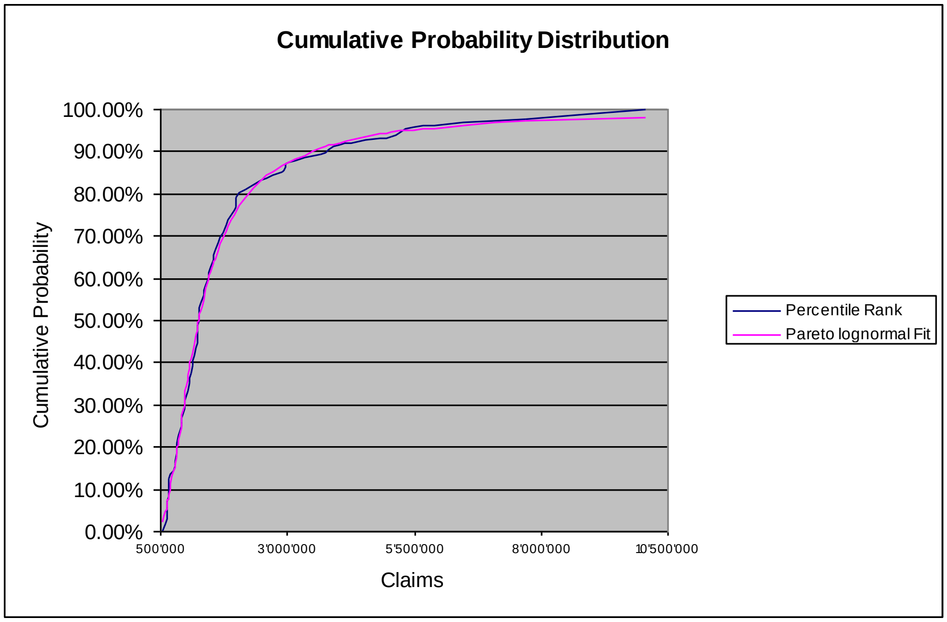

Figure 2. Graph of the Pareto lognormal fit by the minK method

Next, the Tables 2 and 3 display the impact of the chosen fitting methods on the evaluation of the mean and standard deviations of XL-reinsurance layers. We note that the preferred Pareto lognormal with minK estimators yields the safest mean and standard deviations, a choice that we recommend for conservative pricing of excess-of-loss reinsurance.

Table 2. Mean of XL-reinsurance layers for fitted models

XL mean						
Layer 1	**MLE**	**Exponential Pareto Min K**	**Min Chi^2**	**MLE**	**Pareto lognormal Min K**	**Min Chi^2**
23.5 xs 1.5	7'627'205	7'056'029	8'394'244	11'827'930	13'332'047	12'821'232
23 xs 2	6'264'061	5'696'715	6'989'569	9'513'770	10'730'176	10'322'800
22.5 xs 2.5	5'335'627	4'790'116	6'015'911	7'977'220	8'997'517	8'656'354
22 xs 3	4'651'348	4'132'944	5'288'255	6'867'351	7'745'715	7'452'065
21 xs 4	3'692'544	3'229'461	4'252'296	5'347'659	6'031'650	5'802'988
20 xs 5	3'039'504	2'626'881	3'534'216	4'338'618	4'893'548	4'708'032
Layer 2						
5 xs 25	212'619	169'188	264'641	274'512	309'624	297'886
10 xs 25	377'655	298'714	472'436	484'057	545'971	525'273
15 xs 25	510'540	401'788	641'407	650'389	733'577	705'767
20 xs 25	620'505	486'211	782'442	786'326	886'901	853'279
25 xs 25	713'452	556'921	902'565	899'963	1'015'072	976'591
30 xs 25	793'357	617'212	1'006'543	996'685	1'124'166	1'081'549
40 xs 25	924'413	715'052	1'178'609	1'153'290	1'300'802	1'251'488
50 xs 25	1'028'170	791'535	1'316'290	1'275'380	1'438'507	1'383'973

Table 3. Standard deviation of XL-reinsurance layers for fitted models

XL standard deviation						
		Exponential Pareto			**Pareto lognormal**	
Layer 1	**MLE**	**Min K**	**Min Chi^2**	**MLE**	**Min K**	**Min Chi^2**
23.5 xs 1.5	8'371'259	7'775'739	9'038'279	9'992'214	10'611'879	10'408'668
23 xs 2	7'948'681	7'357'815	8'605'055	9'447'881	10'033'912	9'841'867
22.5 xs 2.5	7'576'786	6'994'340	8'220'113	8'975'679	9'532'427	9'349'992
22 xs 3	7'240'846	6'668'967	7'869'821	8'553'807	9'084'387	8'910'527
21 xs 4	6'644'934	6'097'658	7'243'266	7'814'695	8'299'429	8'140'592
20 xs 5	6'120'432	5'600'059	6'687'061	7'172'405	7'617'300	7'471'518
Layer 2						
5 xs 25	1'007'188	897'282	1'124'887	1'142'628	1'213'504	1'190'280
10 xs 25	1'859'573	1'649'723	2'084'203	2'098'940	2'229'135	2'186'473
15 xs 25	2'599'586	2'297'868	2'922'521	2'921'321	3'102'527	3'043'150
20 xs 25	3'254'282	2'867'374	3'668'436	3'642'869	3'868'831	3'794'788
25 xs 25	3'841'966	3'375'491	4'341'404	4'285'802	4'551'644	4'464'534
30 xs 25	4'375'601	3'834'363	4'955'267	4'865'747	5'167'563	5'068'665
40 xs 25	5'316'466	4'637'656	6'044'042	5'879'442	6'244'135	6'124'633
50 xs 25	6'129'083	5'325'629	6'991'025	6'746'048	7'164'496	7'027'380

5. REFERENCES

[1]. R. Hogg and S. Klugman, Loss Distributions, John Wiley. New York (1984).

[2]. H. Panjer and G. Willmot, Insurance Risk Models, Society of Actuaries. Schaumburg (1992).

[3]. S. Klugman, H. Panjer and G. Willmot, Loss Models, From Data to Decisions, John Wiley, New York (1998).

[4]. W. Hürlimann, Financial data analysis with two symmetric distributions, *ASTIN Bulletin* **31(2)**, 187-211 (2001).

[5]. W. Hürlimann, Fitting bivariate cumulative returns with copulas, *Computational Statistics and Data Analysis* **45(2)**, 355-372 (2004).

[6]. G. Benktander and C.-O. Segerdahl, On the analytical representation of claim distribution with special reference to excess of loss reinsurance, In: Transactions of the International Congress of Actuaries, 626-46 (1960).

[7]. G. Benktander, Schadenverteilung nach Grösse in der Nicht-Lebensversicherung, *Bulletin of the Swiss Association of Actuaries*, 263-84 (1970).

[8]. Th. Mack, Schadenversicherungsmathematik, Verlag Versicherungswirtschaft, Karlsruhe (1997).

[9]. W.J. Reed, The Pareto, Zipf and other power laws, *Economics Letters* **74**, 15-19 (2001).

[10]. W.J. Reed, The Pareto law of incomes – an explanation and an extension, *Physica* **A319**, 469-486 (2003).

[11]. W.J. Reed and M. Jorgensen, The double Pareto-lognormal distribution – A new parametric model for size distribution, *Communications in Statistics – Theory & Methods* **33(8)**, 1733-1753 (2004).

[12]. S. Kotz, T.J. Kozubowski and K. Podgorski, The Laplace Distribution and Generalizations, Birkhäuser (2001).

[13]. O.E. Barndorff-Nielsen, Exponentially decreasing distributions for the logarithm of particle size, *Proceedings of the Royal Society of London, Ser.* **A 353**, 401-419 (1977).

[14]. E. Eberlein, Application of generalized hyperbolic Lévy motions to finance, In: O.E. Barndorff-Nielsen, T. Mikosch and S. Resnick (Eds.), Lévy Processes: Theory and Applications, Birkhäuser Boston, 319-336 (2001).

[15]. S. Karlin and H.M. Taylor, An Introduction to Stochastic Processes, vol. II, Academic Press, New York (1981).

[16]. O.E. Barndorff-Nielsen, T. Mikosch and S. Resnick (Eds.), Lévy Processes: Theory and Applications, Birkhäuser, Boston (2001).

[17]. W. Schoutens, Lévy Process in Finance. J. Wiley & Sons, Chichester (2003).
[18]. K.Th. Hess, W. Macht and K.D. Schmidt, Thinning of risk processes, Dresdner Schriften zur Versicherungsmathematik **1** (1995).
[19]. K.Th. Hess, Das kollektive Modell der Risikotheorie in der Schadenexzedenten-Rückversicherung, *Allgemeines Statistisches Archiv* **87**, 309-320 (2003).
[20]. T. Franke and W. Macht, Decomposition of risk processes, Dresdner Schriften zur Versicherungsmathematik **2** (1995).
[21]. K.D. Schmidt, Lectures on Risk Theory. Teubner, Stuttgart (1996).
[22]. K.D. Schmidt, Versicherungsmathematik, Springer, Berlin (2002).
[23]. H. Schmitter, Quotierung von Sach-Schadenexzedenten mit Hilfe des Paretomodells, Swiss Re publications (1978).
[24]. H. Schmitter and P. Bütikofer, Abschätzung von Risikoprämien für Sach-Schadenexzedenten mit Hilfe des Paretomodells, Swiss Re publications (1997).
[25]. R. Doerr, Property Excess of Loss: Pareto Rating, Swiss Re publications (1980).
[26]. M. Schmutz and R. Doerr, Das Paretomodell in der Sach-Rückversicherung, Swiss Re publications (1998).
[27]. P. Embrechts, C. Klüppelberg and Th. Mikosch, Modelling Extremal Events for Insurance and Finance. Springer-Verlag, Berlin (1997).
[28]. R.B. D'Agostino and M.A. Stephens, Goodness-of-Fit Techniques, Marcel Dekker (1986).