

ON SUMS OF CLAIMS AND THEIR APPLICATIONS IN ANALYSIS OF PENSION FUNDS AND INSURANCE PRODUCTS

Rastislav Potocký, Helmut Waldl, Milan Stehlik*

Abstract:

The problem that assets of a fund are not sufficient to cover its liabilities is of extreme importance both for its members as well as for fund managers. We show that this problem can be solved via total claims distributions and give answers to the following questions: How much money will be needed in the first pillar in order to satisfy the requirements of pensioners in a time horizon and which groups of working people should join also the second pillar because their benefits from it will be greater than those from the first pillar? Though the paper concentrates primarily on the situation with Slovakian pension funds we believe that our findings are more general. We show that the alternative methods should be used for calculation of extremes. We discuss the so-called barrier strategy for treating the surplus of an insurance company and bring some new results concerning it.

Keywords: light- and heavy-tailed distributions, catastrophic events, claims, first and second pillars, robust approach, Johnson estimators, the 20–80 rule.

JEL Classification: C 19, G 22

1. Introduction

In pension fund theory the problem of solvency is of great importance. As widely known one recognizes the so-called first pillar which is a mandatory (unfunded) pension scheme operated under public (state) management and the second pillar which is a fully funded mandatory or voluntary scheme financed by employees contributions. Of course further pillars (third or even fourth) may exist, the two above-mentioned being the most important. Since the guaranteed pension may not be a fixed amount but may *e.g.* increase with national average earnings, in a system without capitalization contributions of the active members may not be sufficient and the system runs into difficulties when time of paying pensions comes. While in the first pillar based on pay-as-you-go principle we concentrate on the question of the limiting value of reserve value of the managers of the fund should put apart, another important problem arises for pensioners. It is the question which

* Rastislav Potocký, FMFI UK, Mlynská Dolina 842 48, Bratislava 4, Slovak Republic (potocky@fmph.uniba.sk); Helmut Waldl, JKU Linz, Altenbergerstraße 69, A-4040 Linz a. D., Austria; Milan Stehlik, JKU Linz, Austria and Universidad Tecnica Federico Santa Maria, Casilla 110-V, Valparaiso, Chile (Milan.Stehlik@jku.at, milan.stehlik@usm.cl).
The authors acknowledge the fruitful discussions with Zdeněk Fabián and Jozef Kiselák. Research was supported by ASO Project No. SK-0607-BA-018, by the VEGA Grant (Slovak Grant Agency) No. 1/3016/06 and by Aktion No. 50p14.

of the two pillars to join and with which preferences. This question, however, is not easy to answer. Here we analyse situation when employees are faced with the problem whether to contribute in the non-funded 1st pillar only or to join also the so-called second (funded) pillar. We recall that such a problem is very important in almost all Eastern European countries as increase of the members of the second pillar is expected to help solving problems with inadequate pensions in future but, on the other hand, it could cause deficit in the state pillar nowadays. Besides we compare the upper bounds based on normal approximation approach with those based on other distributions. The problem is illustrated on a real data set of Slovakian pensioners.

In classical risk theory the surplus of an insurance company (or a fund) is a function of the insurer's initial surplus (or free reserves), the premium received, the number of claims up to a certain time and the amount of claims. If in a time, say τ , the premium collected plus the initial surplus is less than the aggregate claim amount, the so-called ruin will occur. Such a situation has been studied for many years starting with the pioneer works of Cramér and Lundberg. On the other hand, if the business is successful increasing dividends can be expected. As models with a premium rate dependent on the surplus are too complicated, the so-called barrier strategy for treating the (possible) surplus of an insurance company is preferred. While in the majority of papers only exponentially distributed claims are mentioned we also discuss mixtures of exponentials.

On the other hand, the problem of total claim amount in a portfolio of policies covering catastrophic risks must be solved via heavy-tailed distributions. Among them Pareto, log-normal, log-gamma and mixtures of them are the most important (see Hewitt, Lefkowitz, 1979 and Hogg, Klugman, 1984). The aim of this paper is to demonstrate that for these distributions the behaviour of the total sum of claims is determined by the behaviour of their maximum. We present examples showing that total claims are highly sensitive to underlying individual distributions and to what extent parameter estimation influences the quality of the total sum.

Assessment of compound sums has many applications in insurance, auditing and operation risk capital assessment among others. In operation risk capital assessment there are many good reasons to use a lognormal distribution (see Alexander, 2003). Although this will not capture well the extremely high impact losses, these are by definition very rare indeed. But exact assessment of these upper quantiles of risk, for example calculation of operational risk capital at a percentile is of interest for Basel II related frameworks (see Alexander, 2005). This can be particularly interesting also for Solvency II. The Solvency II Directive 2009/138/EC is an EU Directive that codifies and harmonises the EU insurance regulation. Primarily this concerns the amount of capital that EU insurance companies must hold to reduce the risk of insolvency.

In the last part a specific problem is discussed. It is widely known that for extremal events the following situation occurs very often: 20% of the individual claims are responsible for more than 80% of the total (aggregated) claim amount in a portfolio of policies. The latter is called the 20–80 rule. Of course characterizing distributions which obey this rule is of great importance for insurance companies. A real data example illustrates the problem.

2. Material and Methods

We consider a homogeneous portfolio of independent, identically distributed positive claims X_k with the distribution function F and the finite expectation (mean) μ . The claims occur in random times T_n and their number in the time interval $[0, t]$ is counted by the process $N(t) = \sup\{n \geq 1, T_n \leq t\}$. If the inter-arrival times are exponentially distributed, $N(t)$ is a homogeneous Poisson process with intensity, say λ . This is the classical Cramér-Lundberg model. The corresponding process of aggregate (or total) claims is

$$S(t) = \sum_{i=1}^{N(t)} X_i \quad (\text{for a fixed } t \text{ we write } N \text{ and } S \text{ instead of } N(t) \text{ and } S(t), \text{ respectively}).$$

First we are interested in an asymptotic evaluation of probability $P\{S > x\}$. We suppose that N is either a random variable with the finite expectation λ or a constant (in the last mentioned case notation S_n is used instead of S).

The classical result is **Cramér's theorem**, which tells that if the moment generating function of X_i exists in a neighbourhood of zero and its variance is l , then

$$P\{S_n - nEX > x\} \approx \bar{\Phi}\left(x / n^{1/2}\right) \quad (1)$$

uniformly for $x = o(n^{1/6})$ where $\bar{\Phi}$ is the tail of the standard normal distribution.

While (1) gives only a rough estimation a more precise result can be obtained by using Petrov's theorem. (For the proof see Petrov, 1975.)

Petrov's theorem: Let the moment generating function $M(h) = Ee^{hx}$ of a random variable X exists in a neighbourhood of zero. Let $a = \sup\{h : M(h) < \infty\}$ and $h = h(x)$ be the unique solution of the equation $g(h) = \frac{M^{(1)}(h)}{M(h)} = x$. If we set $\sigma^2(h) = g^{(1)}(h)$, $a_0 = \lim_{h \rightarrow a} g(h)$ and suppose α_0 finite, then

$$P\{S_n - n\mu > x\} = \frac{\exp\{n(\ln M(h) - hx)\}}{h\sigma(h)\sqrt{2\pi n}}(1 + o(1)). \quad (2)$$

uniformly for $x \in \langle \varepsilon n, (a_0 - \varepsilon)n \rangle$. ($M^{(1)}(h)$ means the first derivative).

The theorem is valid even if the number N of claims is random, provided its expected value is large.

If X is exponentially distributed with parameter α , then $M(h) = \frac{\alpha}{\alpha - h}$ for $h < \alpha$ and hence

$$P\{S_n - n\mu > x\} = \frac{\exp\{n(\ln \alpha x - \alpha x + 1)\}}{\sqrt{2\pi n(\alpha x - 1)}}(1 + o(1)). \quad (3)$$

We use this result when analysing the pensions in the so-called state pillar in Slovakia. This approach also makes possible treating the case when a person is the member of the second (private) pillar.

Next we are interested in probabilities of large deviations of S under the assumption that F is heavy tailed. In particular, we assume that X_i has no finite exponential moments, *i.e.* standard large deviation theory leading to normal distribution does not apply (the existence of the moment generating function is crucial for proving Cramér's theorem, for more see *e.g.* Potocký, 2008). A random variable X (or its distribution function F) is said to be heavy tailed if $E(e^{rx}) = +\infty$ for any $r > 0$. Such a situation is typical for many distributions met in insurance, when one is interested in modelling large claims. In such a case distributions with exponentially decaying tail do not form adequate models. The most important heavy-tailed subclass is the subexponential class. The subexponentiality means that the tail of the sum of random variables becomes large by a dominating large random variable, *i.e.* F is subexponential if for every $n \geq 2$ and x tending to $+\infty$, $P(\sum_{i=1}^n X_i > x)$ is approximated with $P(\max_{1 \leq k \leq n} X_k > x)$, *i.e.* the distribution of the sum of random variables is determined by the distribution of their maximum.

Mikosch and Nagaev (2001) have proved that for a subexponential random variable X with a regularly varying tail (*e.g.* Pareto distributed claim) and under the assumption that the moment generating function of the counting variable N exists in a neighbourhood of zero, we have

$$\Delta(x) := \frac{P(S > x)}{P(X > x)} - E(N) \rightarrow 0, x \rightarrow +\infty. \quad (4)$$

(This result is known as Mikosch-Nagaev formula).

Without additional conditions the approximation of $P(S > x)$ by $E(N) P(X > x)$ may be very bad, since the rate of convergence in (4) can be arbitrarily slow as it is shown in Mikosch and Nagaev (2001). In the same paper they proved that $\Delta(x) = O(1/x), x \rightarrow +\infty$ under some regularity conditions and that the rate $O(1/x)$ cannot be improved even for such a regular distribution as Pareto. Christoph (2005), theorem 2, has proved that for regularly varying distributions with regularity index α , $1 < \alpha < 2$ and supposing $E(N^3) < \infty, \mu = E(X)$ we have

$$\Delta(x) = \frac{\alpha\mu(E(N^2) - E(N))}{x} + O(x^{-(r-\alpha)}), x \rightarrow \infty. \quad (5)$$

for $1 + \alpha < r \leq 2\alpha$

It follows that probability of S exceeding x is not negligible even for large x and this result should be taken into account by general insurance offices when dealing with large claims.

Moreover, especially when we analyze the sums of claims, robust estimates and procedures are needed. Here we empirically study the quality of the upper-quantile closeness drawn from classical parametric inference (method of moments and maximum likelihood estimation) and Johnson moments when the individual claim distribution is misspecified. Such studies are typical in various fields of statistics and actuarial sciences and are sometimes called sensitivity analysis.

The maximum likelihood estimation is very sensitive to deviations from the theoretical distributions, also in the class of heavy-tailed distributions. Not surprisingly (see *e.g.* Alexander, 2005) the maximum likelihood estimator of parameters failed to provide a reasonable estimation. Instead estimators based on Johnson moments should be used (see Stehlík, Potocký, Fabián, Waldl, 2010). More robust alternatives to MLE approach have been proposed by Juárez and Schucany (2004).

In this paper we show two examples. The first example in Section 4 shows, that the application of the Mikosch Nagaev formula for the estimation of upper extreme quantiles is very sensitive. In the second example the application of Gerd-Christoph's Formula (5) is discussed. In all these examples a significant difference in quantiles was observed and plotted by qq-plots. The qq-plot provides a somewhat informal but convenient way of graphical detection of such a difference. All simulations are made using Matlab 7.1.

In Section 5 we consider a corresponding process to $S(t) = \sum_{i=1}^{N(t)} X_i$, namely the process of the insurer's surplus at a time t . For a fixed t it is a random variable, since its value depends on the claims experience up to that time. It will be denoted by $U(t)$. So we have the model (the surplus or risk model)

$$U(t) = u + ct - S(t) \quad (6)$$

where c means the premium income rate in one time unit and $u \geq 0$ is called the initial surplus or free reserves. It follows easily that for $EU(t)/t \rightarrow c - \lambda\mu$ for $t \rightarrow \infty$, λ is intensity of homogeneous Poisson process $N(t)$. So the condition $c - \lambda\mu > 0$ is necessary for the solvency of the insurance company. However, it can happen that $U(t)$ falls below zero as a result of the last claim. In such a case we say that ruin has occurred. Therefore we define the probability of ultimate ruin as

$$\Psi(u) = P\{U(t) < 0 \text{ for some } t \in (0, \infty)\} \quad (7)$$

It is well known (see *e.g.* Embrechts, Kluppelberg, Mikosch, 2003) that if the solvency condition $c - \lambda\mu > 0$ holds and a constant $v > 0$ exists such that $\int_0^{\infty} e^{vx} \bar{F}(x) dx = c / \lambda$, then for all $u \geq 0$ $\Psi(u) \leq e^{-vu}$, where $\bar{F}(x) = 1 - F(x)$ is the tail distribution of claims.

It follows that the result may be only applied to such claims, whose moment generating function exists at least in a neighbourhood of zero, otherwise the constant cannot be found (of course, this is precisely the situation covered by Cramer's and Petrov's theorems). Such distributions are usually termed light-tailed distributions as the probability for exceeding large values goes to zero quickly. Among them the exponential distribution plays the key role.

A short inspection of the above-mentioned result says that the ruin probability decreases with u increasing. It follows from this that the surplus $U(t)$ tends to infinity

with probability 1. For this reason some economists suggested to reduce free reserves consistently to a predetermined value and to pay out the reduction as dividends (bonuses) to shareholders of the company.

In what follows we restrict our attention to the so-called dividend barrier strategy.

Replacing (6) by the modified surplus $Q(t) = U(t) - Z(t)$, where $Z(t)$ means the summed dividend up to time t , the dividend in time t is defined as follows: if $Q(t) < B$ no dividends are paid, if $Q(t) = B$ dividends are paid with rate c until the next claim occurs. B is called the barrier. We suppose that $u < B$. In what follows the expected discounted sum of dividends will be denoted by $V(u, B)$. The aim is to find B such that $V(u, B)$ is maximal.

It is shown in Bühlmann (1970) and Gruber (1979) that $V(u, B)$ satisfies the integro-differential equation

$$V^{(1)}(u, B) = \frac{\lambda + \delta}{c} V(u, B) + \frac{\lambda}{c} \int_u^0 f(x) V(u - x, B) dx \quad (8)$$

with the boundary condition $V^{(1)}(B, B) = 1$ (in this formula δ means the force of interest and $f(x)$ means density). Solving equation (8) becomes easier if we set $V(u, B) = h(u)/h^{(1)}(B)$ for a suitable $h(u)$. Then we solve the equation

$$ch^{(1)}(u) = (\lambda + \delta)h(u) - \lambda \int_0^u h(u - x) f(x) dx \quad (9)$$

The exact solution in case of exponentially distributed claims can be found in the above mentioned books.

Finally, the so-called 20–80 rule is discussed. It means that 20% of the individual claims are responsible for more than 80% of the total claim amount in a portfolio of policies. A real data set shows how closely related this rule is to extremal events.

3. Pay-As-You-Go Pension Scheme vs. a Funded Pension Scheme

3.1 A brief description of pension funds

As widely known the social security systems in majority of European countries consist of the so-called first pillar which is a mandatory (unfunded) pension scheme operated under public (state) management (PAYG) and the second pillar which is a fully funded mandatory or voluntary scheme financed by employees contributions under private management.

Financial crisis together with demographic ageing are the main reasons for bad financial prospects for both PAYG latter pillar and funded second pillar. While the first has reduced the value of assets in the funded plans the former brings the threat of insufficient volume of money for the increasing number of pensioners.

The falling birth rates, diminishing work of older people and increasing life expectancy on the one side together with increasing unemployment and decreasing or stagnant

wages on the other one are the main reasons supporting the expectation that the number of the active members of the state pension scheme in many Eastern European countries will be smaller than the number of members who retire due to attainment of retirement age or for other reasons in the time horizon of 20 to 30 years. Since the guaranteed pension may not be a fixed amount, but may increase with national average earnings, contributions of the active members will be insufficient and the system runs into difficulties. For this reason people in Eastern European countries are recommended to join also the second pillar which is based on a different principle.

Briefly speaking the second pillar relies on funded pension systems which are realised through pension plans with defined benefits or with defined contributions. The main differences among them are as follows.

In a defined benefit pension plan the retirement benefits are promised in advance according to some predetermined formulae by the manager and contributions are chosen to maintain the fund in balance, while in a defined contribution plan contributions are fixed and benefits depend on the returns of the fund. So these plans treat the financial risk differently: in the first case the manager of the fund bears the risk of poor investment performance while in the second case this risk is left to pensioners.

In a defined benefit pension plan the main problem is to determine an ideal contribution rate and an ideal fund level (related with the actuarial liability) in such a way that the benefits promised will be guaranteed during the time. The plan is built with the contributions and investment earnings. The manager controls the stability and security of the plan by controlling the most important risks, namely the contribution rate risk and the solvency risk (or their combinations).

In plans of both types one recognizes an accumulation phase and a decumulation phase. While in the first phase the member contributes to the fund, in the second phase the fund converts his/her savings into a benefit. Three common forms of retirement benefits are:

- a) lump sum payment which is the most popular form as the most pensioners opt for this possibility;
- b) programmed withdrawals where remaining balances are inherited by relatives in the case of early death;
- c) a life annuity at the time of retirement. This form is becoming very popular in many countries, including *e.g.* Switzerland and majority of Eastern European countries are leaning to it.

However, also for funded pension plans the following holds: Apart from financial risk (asset, inflation, contribution rate and interest rate risks) longevity is becoming even more serious risk since life expectancy is increasing in almost all European countries. Hence longevity can make the benefit insufficient for the rest of pensioner's life also in the second pillar. In order to reduce this possibility a combination of both pillars is becoming a preferable solution.

3.2. The first (state) pillar in Slovakia

Table 1

Monthly Mean Brutto Salary in Slovakia, 1991-2007 (SKK)

Year	Max. Salaries	Min. Salaries	Year	Max. Salaries	Min. Salaries
1992	7,667	4,148	2000	30,021	6,785
1993	9,882	4,579	2001	31,825	7,262
1994	11,592	5,179	2002	34,041	8,533
1995	13,537	5,975	2003	34,443	8,840
1996	19,377	6,326	2004	39,452	9,446
1997	24,824	5,639	2005	42,544	10,199
1998	24,233	6,208	2006	45,349	10,947
1999	26,862	6,421	2007	51,154	12,945

Here we analyse the following situation: a person can be involved in two pillars – a state pension pillar (obligatory) and another based on a pension fund (voluntary). While conditions in the state pillar are given by law there is more freedom in the pillar based on a private pension fund. The recent development of pension funds in many Eastern European countries is well described (*e.g.* in Whitehouse, 2009). It should be mentioned that conditions for private funds are still under discussion in many countries (*e.g.* the Czech Republic) and the final solution may depend also on political situation.

We emphasize that we only analyse the combination: the state pillar + a private pillar as the other possibilities (*e.g.* being the member of a second pillar only) are not allowed in several countries.

We would like to find the answers to the next two questions: How much money will be needed in the first pillar in order to satisfy the requirements of pensioners in time horizon of twenty to thirty years? The second question sounds as follows: which groups of working people should join also the second pillar because their benefits from it will be greater than those from the first pillar alone?

Seeking the answer to the first question we are interested in the estimation of the

probabilities $P \left\{ \sum_{k=1}^N X_k > C \right\}$ where X_k are individual monthly claims of the members of the first pillar and C is a critical (limiting) value of the fund representing the amount the fund has gathered from the contributions of the active members or from other sources. We emphasize that it is possible to consider N as a constant or a random variable. In this paper we consider both cases. We use both Cramér's and Petrov's theorems and compare the upper bounds based on a normal approximation approach with those based on other distributions.

We provide an example based on a real data set of two closed groups of Slovakian people (opposite with respect to their incomes), all aged 45 in the year 1993 and are interested in the estimation of the total claim amount for these groups in the year 2010 when the members were supposed to retire. In the Statistical Yearbook (2010) of the Slovak Republic we have found the salaries evolution (see Labour Market, III. 3-10, Structure of Average Gross Nominal Monthly Wage of Employees in the Economy of the Slovak Republic), see Table 1. We suppose that the ratio pension/salary is approximately 0.5.

For the maximum salary group we obtained the mean of 27,930 with standard deviation 13,156.56, for the minimum salary group the mean was 7,464 and the standard deviation 2465.33.

If we put $N = 130,000$ (this number can be found in the above mentioned yearbook) and realize that the final salary was 51,154 in 2007 it follows that the needed value for the maximum salary group is $C = 3 \cdot 10^9$ with probability 0.1. On the other hand, the probability that for the maximum salary group total pensions will be greater than $1,9 \cdot 10^9$ is 0.4606213.

If N is random it is quite natural to choose a binomial model, namely $B(n,p)$ with $n = 130,000$ and $p = 0.8$ representing the probability that a 45-year person will survive to the age 62 years. A small simulation experiment with 10,000 simulations (in software R 2.12.2) showed that none of the aggregated sums exceeded $C = 3 \cdot 10^9$ if the aggregate claims (total pension) $\sum_{k=1}^N X_k$ follow the above mentioned model. In contrast for $C = 1,9 \cdot 10^9$ all simulated values were over C . Thus we can conclude $C = 1,9 \cdot 10^9$ that is not sufficient for this income group with probability 1 which is even higher than for constant N .

Also Petrov's theorem confirmed that the probability for aggregate claims being greater than $C = 3 \cdot 10^9$ is zero (the parameter of the exponential distribution was fitted with the maximum likelihood method).

When considering the minimum salary group with $N = 2 \cdot 130,000$ we obtained approximately $1,5 \cdot 10^9$ for the value of C with probability 0.1 (using Cramér's theorem). A simulation experiment (10,000 simulations) with a final salary of 12,945 and $N \sim B(2 \cdot 130,000; 0.8)$ confirmed the non-exceedance of C by the aggregated sum.

To conclude, these observations also support the diversification of pensions in the high salary group (this fact can be easily explained if one realises that the standard deviation in the high salary group is approximately 5 times higher than in low salary groups).

It is worth mentioning that boosting nominal values of salaries, *e.g.* for max. group, bring us new challenges of their modelling, since standard assumptions may be violated (see *e.g.* Uherek, Stehlík, Strelec, 2011 or Stehlík, Strelec, 2009).

Therefore caution should be taken before automatic application of methods working well for pensions modelling in Western Europe, since the application of automatic methods hoping that the data will enforce its true structure is deceptive.

3.3 The first vs. second pillar

The second aim of this part is to show that a person should be very careful when deciding whether to choose the first pillar only or to prefer a mix of both pillars. It should be mentioned that private pillars are not a remedy for everything and their yields were decreasing in the last years.

In what follows ${}_n p_x$ means the probability of a person aged x years to survive the next n years. If a person belongs to the state pillar only his/her contribution is 18% of his/her salary. If, however, he/she chooses both pillars, the contributions become 9% for each pillar. As the first pillar has already been analysed we concentrate on the second pillar. We suppose that after reaching the retirement age (62 years given by law) the member of the second pillar buys a life annuity (the conditions are also given by law but there are some possibilities concerning the length of annuity or whether annuity is level or not). We compute the monthly pension for various life distributions. Our computations are based on the well-known equation of equivalence

$$\sum_{t=0}^k u X_t (1+r_t)^{k-t} {}_t p_{t_0} = \sum_{t=0}^n P (1+r_t)^{-t} {}_t p_T \quad (10)$$

where u is the contribution rate (usually given by law), X_t means employee's salary in the year t , t_0 means the employee's age when he/she entered the scheme, P stands for monthly pension and T for retirement age. We suppose that interest rate r_t is deterministic and does not depend on t . Of course, the case of stochastic interest rates is much more complicated and will be treated in a separate paper.

The decision rule is as follows: If pension P from the second pillar is greater than half of the state pillar pension the pensioner is recommended to join also the second pillar. In the opposite case he/she should be only in the first pillar.

We suppose that people obey mortality published in Statistical Yearbook of the Slovak Republic (2010). It is well known that Weibull, Gompertz-Makeham and even gamma fit the life data in Slovakia well. In order to find the parameters of these distributions the method of moments was used. The results are in Table 2.

For the maximum salary group we obtained the mean of 27,930 with standard deviation 13156.56, for the minimum salary group the mean was 7,464 and the standard deviation 2465.33.

Table 2

Pensions Evaluated for Various Life Times and Interest Rate $r=0.01$

Weibull	P_max=9148.5	P_min=2483
Gamma	P_max=11261.7	P_min=3065.8
Gompertz-Makeham	P_max=20950.4	P_min=5688.2

Realizing that the minimum pension is approximately 7,500 Slovakian korunas (250 euro) the results obtained support our intuitive feeling that joining the second pillar is not good for people from the low salary group as they cannot rely on large gains from private pillar and hence being the members of the state fund is safer. On the other hand, high salary group should join also a private pillar as gains from it outperform losses from obtaining only the half of the first pillar pension. It should be mentioned that the results obtained support also the fact that the accumulation phase should not be shorter than 20 years otherwise the pensions will be low because of the possible long decumulation phase.

So, also from this simple example we can see the relevance of distinguishing between salary groups. It is clear that higher salary groups should be involved, at least partially, in the 2nd pillar. The study also supports the well-known fact that postponing the retirement age will ease the problem of liabilities of the first pillar.

4. Estimation of the Extreme Upper Quantiles for Pareto and Similar Distributions

4.1. Mikosh-Nagaev formula, its importance and drawbacks

Unlike exponential and similar distributions the solution of the problem of large deviations in the case of heavy-tailed distributions brings completely different results. It follows from Mikosh-Nagaev formula (4) that probabilities of large deviations of total claims can be approximated by means of distribution of their maximum. However, it means that such probabilities can be quite significant even for relatively large claims. Moreover, the convergence of $P(S > x)$ to $E(N) P(X > x)$ can be very slow as shown above. Here we demonstrate drawbacks of both Mikosh-Nagaev and Christoph formulas for insurance practice. It follows that averaging techniques are insufficient when dealing with heavy-tailed claims. In other words knowing mean and variation of a portfolio of such claims tells us almost nothing about behaviour of the largest claim(s).

Here we demonstrate that the application of Mikosh-Nagaev results to the estimation of the extreme upper quantiles is rather non-robust and sensitive.

We made 1,000,000 simulations of a Poisson variate N with parameter $\lambda = 10,000$ and of the sum of N log-gamma variables with parameters $\alpha = 10$ and $\gamma = 2.5$. For a comparison we also simulated the sum of $\lambda = 10,000$ log-gamma variables with

parameters $\alpha = 10$ and $\gamma = 2.5$. The two distributions of the sums (with N Poisson distributed and fixed at $N = \lambda$ respectively) seem to be the same.

We studied the speed of convergence of the Mikosch-Nagaev theorem with the empirical distribution of the sum of N iid log-gamma variates with N Poisson distributed. First we compared $P(S > x)$ with $E(N) \cdot P(X > x)$ which is shown in Figure 1. It seems as if $E(N) \cdot P(X > x) - P(S > x) \rightarrow 0$ with $x \rightarrow \infty$. Convergence does not look that fast if we plot $\frac{P(S > x)}{P(X > x)}$ against $E(N)$ (see Figure 2) but it still seems to be acceptable. The applicability of the theorem can be seen if we plot $\frac{P(S > x)}{P(X > x)}$ for small values of $P(S > x)$ (see Figure 2 and 3). It seems as if $\frac{P(S > x)}{P(X > x)}$ would converge to $E(N)$ but the theorem is not applicable for probability values used in actuarial practice as the convergence is too slow.

Figure 1

Approximation for a Large x

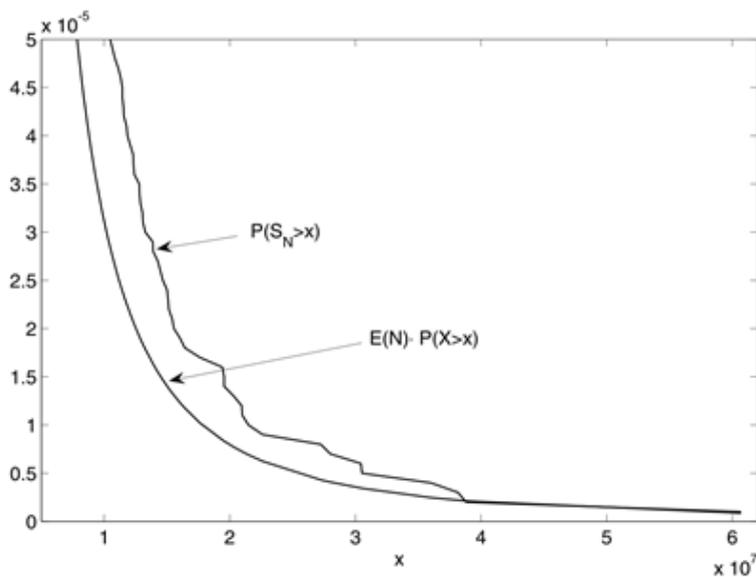


Figure 2

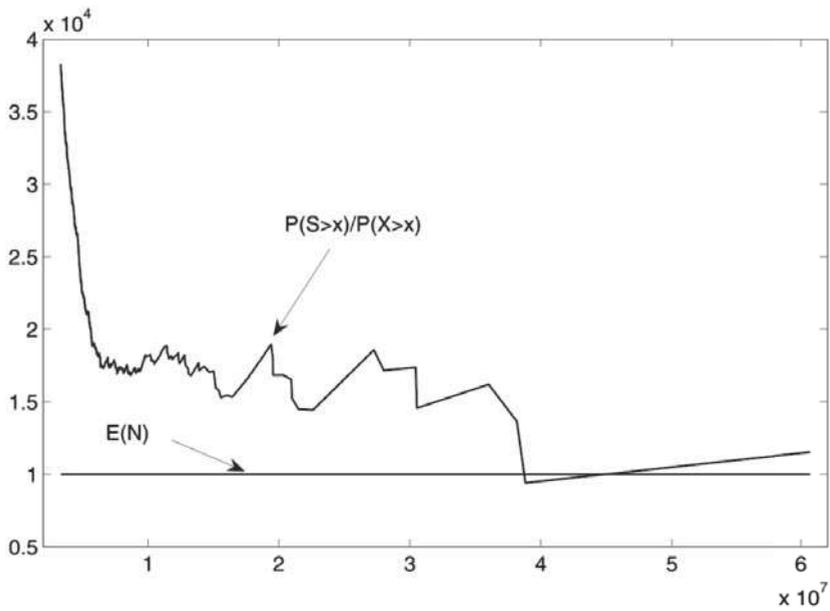
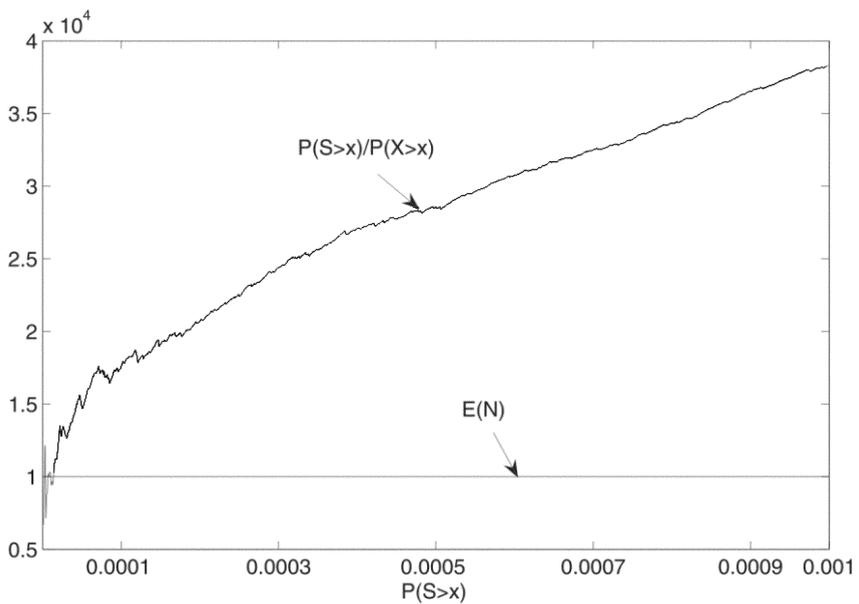
Convergence in the Mikosch-Nagaev Theorem with loggamma Data

Figure 3

Fit of Upper Quantiles by Mikosch Nagaev Formula

4.2 An application of Christoph's formula

Here it is demonstrated that the application of Christoph's Formula (5) to the estimation of the extreme upper quantiles is rather non-robust and sensitive.

We made 1,000,000 simulations of the sum of N Pareto distributed random variates with $\alpha = 1.5$, $\lambda = 1$ where N is Poisson distributed with parameter $\lambda = 10,000$. In Figure 4 we can see $\Delta(x)$ and $\frac{\alpha\mu(E(N^2)-E(N))}{x}$. It looks as if Christoph's theorem based asymptotics

does not work very robust in this setup. This can be observed also in Figure 5, where we plotted $\varepsilon(x) := \Delta(x) - \frac{\alpha\mu(E(N^2)-E(N))}{x}$ and $x^{\alpha-r}$ for $1 + \alpha < r \leq 2\alpha$ against $P(S > x)$.

This asymptotic fit can be explained by the fact that we know $\Delta(x)$ only empirically and that the variation of the empirical $\Delta(x)$ is very high for high x -values. Nevertheless, we can see that Christoph's theorem is unfeasible in our example especially for probability values used in actuarial practice (see Figure 5).

Figure 4

Fitting in Christoph's Theorem

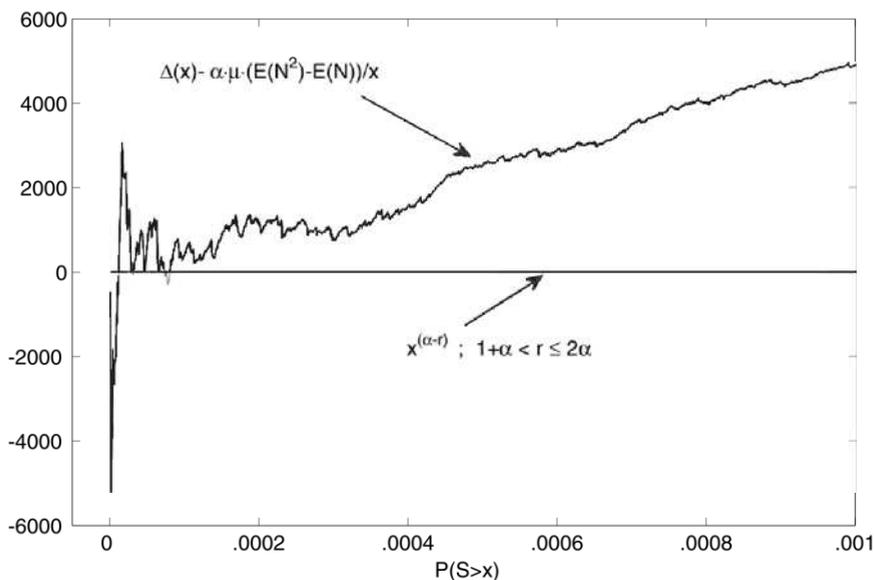
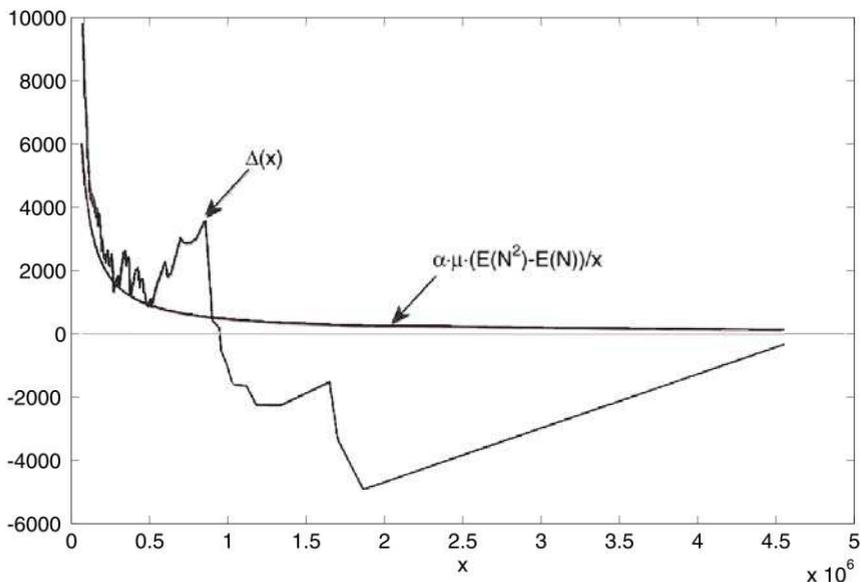


Figure 5

Fit of Christoph's Theorem for Large Claims**5. The 20-80 Rule and Its Meaning for Insurance Business**

The 20-80 rule is often used by applied actuaries. It states that 20% of the individual claims are responsible for more than 80% of the total claim amount. The link to heavy-tailed distributions seems to be obvious as exactly for these distributions the behaviour of the sum is determined by the behaviour of its largest component(s). The following example supports this statement.

We consider a set consisting of 75 claims reported by a non-life insurance company (see Table 3, also Figure 6). As they strongly differ in magnitude and obey the 20-80 rule we suspect they could come from a heavy-tailed distribution.

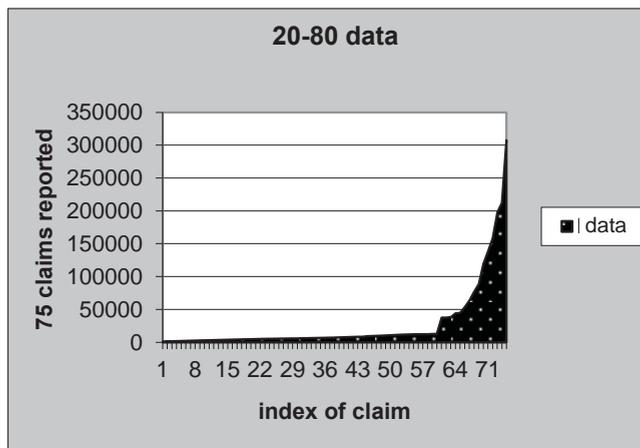
Table 3

20-80 Data

1295	3700	5262	6457	9271	12250	63840
1480	3899	5493	6645	9344	12515	76598
1603	3956	5747	6918	9934	12571	88779
1924	4134	5832	7279	10262	12616	119018
1945	4173	5962	7502	10730	12928	138629
2050	4258	6079	7611	11064	37871	158029
2132	4429	6142	7821	11250	37882	198075
2521	4721	6181	8015	11457	38571	212252
3119	4848	6229	8239	11857	44245	307556
3156	5000	6350	8545	12178	44979	
3488	5024	6446	8665	12228	53560	

Source: own calculation

Figure 6

20-80 Data

Recalling the density of Pareto distribution $f(x) = \alpha \lambda^\alpha (\lambda + x)^{-(\alpha+1)}$ we have the following results for testing Pareto distribution for the new data estimators:

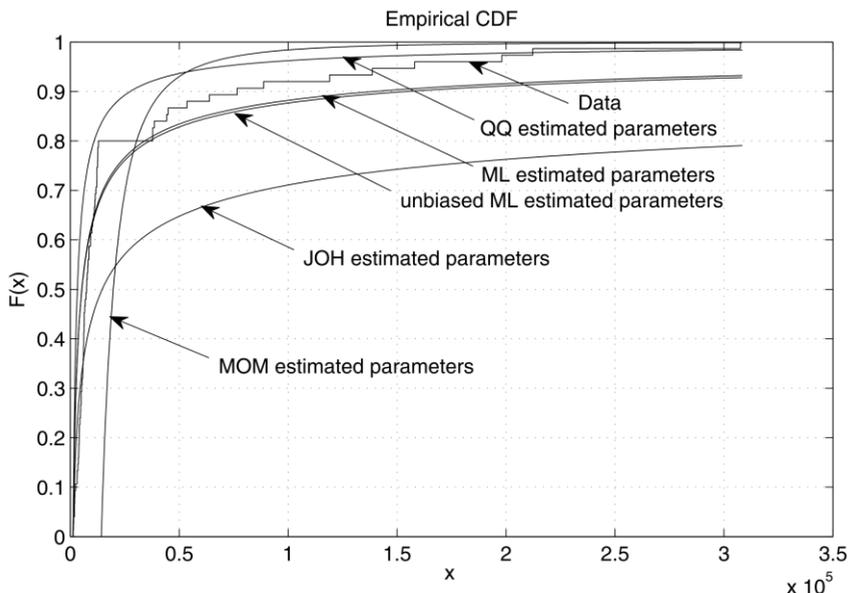
$$\lambda_{ML} = 1295 \quad \alpha_{ML} = 0.4916; \quad \lambda_{unbML} = 1258.9 \quad \alpha_{unbML} = 0.4784;$$

$$\lambda_{MOM} = 14261 \quad \alpha_{MOM} = 2.1199; \quad \alpha_{QQ} = 0.7536; \quad \alpha_{JOH} = 0.2855$$

(ML stands for maximum likelihood, unbML for unbiased likelihood, MOM for method of moments, JOH for Johnson and QQ for QQ estimator). See also Figure 7.

Figure 7

Non Fit of Pareto



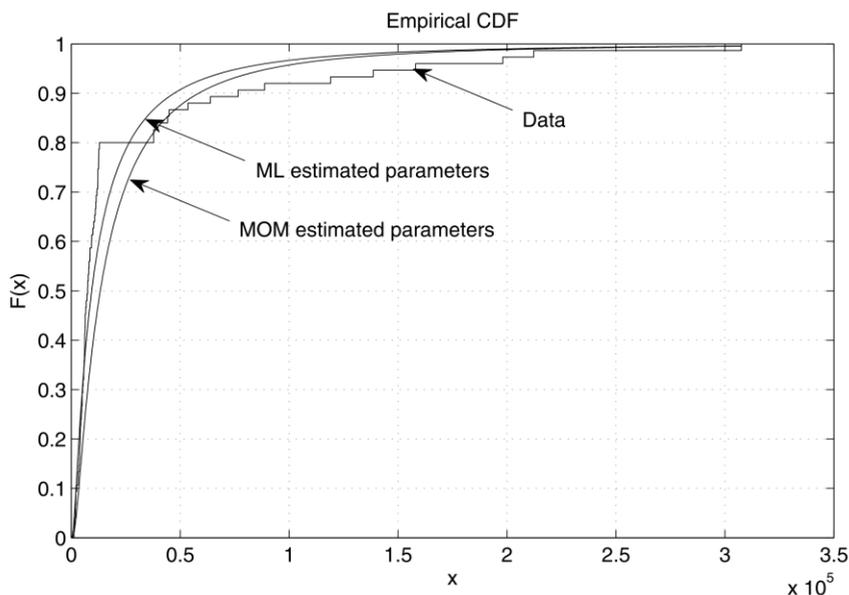
We have p-values for testing Pareto distribution with the according parameters: ML: $6.68e - 5$ unbiased ML: $7.01e - 5$ Johnson estimator: $4.84e - 7$ QQ: $7.01e - 12$, moment estimator: $2.62e - 43$. Testing is based on the exact likelihood ratio test of the homogeneity of Pareto distributions, which is derived in Stehlík, Potocký, Fabián and Waldl (2010) and used here as an omnibus test for Pareto distributions. From this test p-values are simulated in Matlab. For more on properties of such a test see Stehlík and Wagner (2009). We conclude that the data do not look like they would come from

a Pareto distribution. However, log gamma (its density is $f(x) = \frac{a^b (\ln x)^{b-1}}{x^{a+1} \Gamma(b)}$) fits

better. We have used the matlab parametrization for the log gamma distribution. We

got the estimators $\hat{a}_{ML} = 59.1685$, $\hat{b}_{ML} = 0.1555$, $\hat{a}_{MOM} = 75.1256$, $\hat{b}_{MOM} = 0.127$ and p-values ML: 0.0054 and MOM: $6.36e - 7$. See also Figure 8.

Figure 8

Fit of log gamma Distribution**6. Dividend Strategy and Its Role in Risk Management of Insurance Companies**

In this example we present the solution of dividend strategy for a mixture of two exponential distributions with density functions $f_1(x) = \alpha e^{-\alpha x}$ and $f_2(x) = \beta e^{-\beta x}$ respectively where $0 < \alpha < \beta$, i.e. the density function of the mixture will be $f(x) = pf_1(x) + (1-p)f_2(x)$.

To be comparable with Bühlmann (1970), p. 174, we have chosen the following values for α , β , p and λ (see Table 4).

Table 4

Barrier for Mixtures of Two Exponentials

α	β	p	λ	c	B
0.001	0.002	0.9	1000	1.1 million	85016.5
0.002	0.004	0.9	200	1.1 million	306100

Having in mind the result for exponential distributions we seek the solution in the form $h(u) = \sum_{i=0}^2 C_i e^{r_i u}$ with suitable C_i , r_i , $i = 0, 1, 2$. Substituting in (9) and comparing the coefficients of $e^{r_i u}$, $i = 0, 1, 2$ we obtain that they satisfy the equation

$$cr - (\lambda + \delta) + \lambda \left(\frac{p\alpha}{\alpha + r} + \frac{(1-p)\beta}{\beta + r} \right) = 0 \quad (11)$$

which may be rewritten as

$$cr^3 + ((\alpha + \beta)c - \lambda - \delta)r^2 + (\alpha\beta c - \delta(\alpha + \beta) - \lambda(\alpha(1-p) + \beta p))r - \delta\alpha\beta = 0$$

(11) has three roots, one nonnegative r_0 and two negative r_1 and r_2 , in descending order. Moreover $-\beta < r_2 < -\alpha < r_1 < 0 < r_0$ holds.

Comparing the coefficients e^{-au} of and $e^{-\beta u}$ respectively, we obtain that

$$C_0 = \frac{(\alpha + r_0)(\beta + r_0)}{(r_0 - r_1)(r_0 - r_2)} > 0, C_1 = -\frac{(\alpha + r_1)(\beta + r_1)}{(r_1 - r_2)(r_0 - r_1)} < 0, C_2 = -\frac{(\alpha + r_2)(\beta + r_2)}{(r_2 - r_1)(r_0 - r_2)} < 0. \quad (12)$$

which coincides with the result in Sheldon, Willmot and Drekcic (2003). From this we obtain the form of $V(u, B)$ and find the value of B for which it attains its maximum. We know that it is the value for which $h^{(2)}(B) = 0$. This leads to the equation

$$e^{r_0 B} \frac{r_1^2}{r_0^2} K_1 e^{r_1 B} + \frac{r_2^2}{r_0^2} K_2 e^{r_2 B} \quad (13)$$

$$\text{with } K_1 = \frac{C_1}{C_0} \text{ and } K_2 = \frac{C_2}{C_0}.$$

The value of B for some mixtures of two exponentials can be found in Table 4. Setting $p = 1$ and comparing it with the exponential distribution discussed in Bühlmann (1970), on p. 174 we see that the results obtained here agree well. Recall that the ruin probability for a mixture of exponentials is given, e.g. in Potocký and Stehlík (2009). The case where the risk model is influenced by a constant interest rate is treated, e.g. in Potocký (2008).

The goal of this example was to demonstrate the case when the parametric mixture is involved. The solution of such a problem may ask for subtle statistical techniques, especially when information on the so-called lower/upper contamination is not available. The problem is treated in general, e.g. in Stehlík, Wagner (2011).

7. Conclusions

In recent decades the field of financial and insurance risk management has undergone an explosive development. Risk evaluation plays an important role also in pension fund theory and practice. This paper shows that assessing liabilities of a fund can be solved via the theory of total claims distributions and brings a solution based on large deviation results. In the second part it discusses the so-called barrier strategy for light-tailed distributions as well as the favourable estimation for fitting individual heavy-tailed data or the aggregated claims of such heavy tailed individuals. Some special topics on extremal data are discussed, too.

The main novelty of the paper is that

(1) it enables the application of results on large deviations to find liabilities of a pension fund and the analysis of the case of the so-called mixed preferences, *i.e.* when a person is a member of both the first (non-funded) pillar and the second (or third) pillar with equal or unequal preferences. We have shown the importance of pensioner's proper choice between state and private pillars. Moreover, we discuss some drawbacks of pension funds.

(2) We present a solution of the barrier strategy for treating the surplus of an insurance company also for mixtures of exponentials, namely a formula for the optimal barrier is given. (It means that when this barrier is reached insurance company starts paying dividends in a certain rate until the next claim occurs).

(3) We discuss the problem of large deviations in the case of heavy-tailed distributions and demonstrate drawbacks of both Mikosh-Nagaev and Christoph formulas for insurance practice. It follows that averaging techniques are insufficient when dealing with large claims. In other words knowing mean and variation of a portfolio of such claims tells us almost nothing about behaviour of the largest claim(s).

(4) We demonstrate that the distribution of aggregate claims is highly sensitive to the distribution of individual ones. This is in agreement with Brazauskas and Serfling, 2000 showing that small errors in the estimation of the tail index can already produce large errors in the estimation of quantiles based on the tail index. Hence robust operators and procedures have to be implemented. To our best knowledge, for collective claims similar computations/simulations have not yet been made, especially novelty is the usage of the Christoph and Mikosch-Nagaev formula and their fit. Discussing aggregated claims extends the results obtained in Stehlík, Potocký, Fabián, Waldl (2010).

(5) We discuss the 20/80 rule which is very popular among applied actuaries. Our findings confirm that distributions obeying this rule should be sought among heavy-tailed ones.

(6) As the simulations and a real-data example shows, the favourable estimation is highly sensitive on the underlying parametric model of the heavy tailed data. Our findings for the Pareto model are in accord with conclusions of Brazauskas and Serfling (2003) that the maximum likelihood (ML) estimator is efficient but not robust and should be replaced by a competitor. It is not robust also for the case of misspecification of the heavy-tailed distribution. One of the possible substitute is also an estimator based on robustified Johnson score. However, more research should be conducted to characterize the cases when such an estimation gives favourable trade-offs between efficiency and robustness. It follows that the classical methods of moments and the maximum likelihood method do not reflect the heavy-tailed character of the data satisfactorily. In this paper therefore we have presented some alternative methods how to treat this problem.

References

- Alexander, C.** (2003), "Technical Document for OpRisk2021 Module 1." [cit. 2009-09-02], available from www.2021solutions.com.
- Alexander, C.** (2005), "Assessment of Operational Risk Capital." In Frenkel, M., Hommel, U., Rudolf, M., eds., *Risk Management, Challenge and Opportunity*. 2nd revised and Enlarged Edition. Berlin: Springer.
- Bühlmann, H.** (1970), *Mathematical Methods in Risk Theory*. Springer-Verlag.
- Brazauskas, V., Serfling, R.** (2000), "Robust and Efficient Estimation of the Tail Index of a Single-Parameter Pareto Distribution." *North American Actuarial Journal*, Vol. 4, No. 4, pp. 12–27.
- Brazauskas, V., Serfling, R.** (2003), "Favorable Estimators for Fitting Pareto Models: A Study Using Goodness-of-fit Measures with Actual Data." *Astin Bulletin*, Vol. 33, No. 2, pp. 365–381.
- Christoph, G.** (2005), "Exact Rates of Convergence to Ruin Probabilities for Regularly Varying Random Variables." In Frenkel, I. et al., eds., *Proceedings of the International Symposium on Stochastic Models in Reliability, Safety, Security and Logistics*. Beer Sheva, Israel, pp. 75–78.
- Christoph, G.** (2004), "Exact Rates of Convergence for Compound Sums of Random Variables with Common Regularly Varying Distribution Functions." In Antonov, V. et al., *Longevity, Aging and Degradation Models in Reliability, Public Health, Medicine and Biology*. Vol. 2. St. Petersburg, pp. 56–67.
- Embrechts, P., Kluppelberg, C., Mikosch, T.** (2003), *Modeling Extremal Events. Applications of Mathematics*. 4th Edition. Springer.
- Fabián, Z.** (2001), "Induced Cores and Their Use in Robust Parametric Estimation." *Communications in Statistics - Theory and Methods*, Vol. 30, No. 3, pp. 537–556.
- Fabián, Z.** (2006), "Johnson Point and Johnson Variance." In Hušková, M., Janžura, M. eds., *Proc. Prague Stochastics 2006*. Prague: Matfyzpress, pp. 354–363.
- Fabián, Z.** (2007), "Parametric Estimates by Generalized Moment Method." Research Report No. 1014, Institute of Computer Sciences, AS CR.
- Fabián, Z.** (2008), "New Measures of Central Tendency and Variability of Continuous Distributions." *Communications in Statistics – Theory and Methods*, Vol. 37, No. 2, pp. 159–174.
- Fabián, Z., Stehlík, M.** (2008), "A Note on Favorable Estimation when Data Is Contaminated." *Communications in Dependability and Quality Management*, Vol. 11, No. 4, pp. 36–43.
- Gruber, H. U.** (1979), *An Introduction to Mathematical Risk Theory*. Philadelphia, PA.: S. S. Huebner Foundation.
- Hewitt, Ch. C., Lefkowitz, B.** (1979), "Methods for Fitting Distributions to Insurance Loss Data." *Proceedings of the Casualty Actuarial Society*, Vol. 66, pp. 139–160.
- Hogg, R. V., Klugman, S. A.** (1984), *Loss Distributions*. New York: Wiley.
- Hommel, U., Ritter, M.** (2005), "New Approaches to Managing Catastrophic Insurance Risk." In Frenkel, M., Hommel, U., Rudolf, M. eds., *Risk Management, Challenge and Opportunity*. 2nd Revised and Enlarged Edition. Springer.
- Juárez, S. F., Schucany, W. R.** (2004), "Robust and Efficient Estimation of the Generalized Pareto Distribution." *Extremes*, Vol. 7, No. 3, pp. 237–251.
- Mikosch, T., Nagaev, A.** (2001), "Rates in Approximations to Ruin Probabilities for Heavy-Tailed Distributions." *Extremes*, Vol. 4, No. 1, pp. 67–78.
- Pacáková, V., Linda, B.** (2009), "Simulations of Extreme Losses in Non-Life Insurance." *Economics and Management*, Vol. 12, No. 4, pp. 97–102.
- Petrov, V. V.** (1975), *Sums of Independent Random Variables*. Springer.
- Philbrick, S. W.** (1985), "Practical Guide to Single Parameter Pareto Distribution." *Proceedings of the Casualty Actuarial Society*, Vol.72, pp.44–84.
- Potocký, R.** (2008), "On a Dividend Strategy of Insurance Companies." *Economics and Management*, Vol. 11, No. 4, pp. 103–110.

- Potocký, R., Stehlík, M.** (2007), "Stochastic Models in Insurance and Finance with Respect to Basel II." *Journal of Applied Mathematics, Statistics and Informatics*, Vol. 3, No. 2, pp. 237–245.
- Potocký, R., Stehlík, M.** (2009), "Statistical Analysis of Mixtures Underlying Probability of Ruin." *Acta universitatis agriculturae et silviculturae Mendelianae Brunensis*, Vol. 62, No. 6, pp. 209–214. ISSN 1211-8516.
- Sheldon, L. X. E., Willmot, G., Drekić, S.** (2003), "The Classical Risk Model with a Constant Dividend Barrier: Analysis of the Gerber-Shiu Discounted Penalty Function." *Insurance: Mathematics and Economics*, Vol. 33, No. 3, pp. 551–566.
- Stehlík, M., Potocký, R., Fabián, Z., Waldl, H.** (2010), "On the Favourable Estimation of Fitting Heavy Tailed Data." *Computational Statistics*, Vol. 25, No. 3, pp. 485–503.
- Stehlík, M., Wagner, H.** (2011), "Exact Likelihood Ratio Testing for Homogeneity of Exponential Distribution." *Communications in Statistics - Simulation and Computation*, Vol. 40, pp. 663–684.
- Stehlík, M., Strelec, L.** (2009), "On Normality Assumptions for Claims in Insurance." *Acta universitatis agriculturae et silviculturae Mendelianae Brunensis*, 2009, Vol. 62, No. 3, pp.141–146. ISSN 1211-8516.
- Uherek, M., Stehlík, M., Strelec, L.** (2011), "On Robust Analysis of Paycheck: Case Study." *Acta universitatis agriculturae et silviculturae Mendelianae Brunensis*, Vol. 64, No. 4, pp. 371–378. ISSN 1211-8516.
- Vandewalle, B., Beirlant, J., Christmann, A., Hubert, M.** (2007), "A Robust Estimator for the Tail Index of Pareto-Type Distributions." *Computational Statistics & Data Analysis*, Vol. 51, No. 12, pp. 6252–6268.
- Whitehouse, E.** (2009), "Pensions during the Crisis: Impact on Retirement Income Systems and Policy Responses." *The Geneva Papers*, Vol. 34, No. 4, pp. 536–547.