Large Loss Distributions

Yuriy Krvavych

Advanced Diploma in Actuarial Mathematics and Insurance
International Centre of Privatisation, Investment and Management
Kyiv, Ukraine

June 1, 1999

Contents

1 Introduction

The losses constitutes large loss are characterised by following conditions:

- grater than certain amount of money;
- have adverse financial effect;
- the event that such large loss occurs is rare;
- reside at right tail of a distribution function.

The lack of sufficient data concerning large claims is a serious problem for Actuary from both a practical and theoretical point of view. In many cases, in order to keep the variation of the aggregate claim amount reasonable, an insurer usually takes out reinsurance cover for his insurance portfolio, i.e. he protects himself against losses arising from large, excessively numerous or catastrophic claims by reinsuring large claim amounts or high claim frequency with one or more other insurance or reinsurance companies. In other cases, unless the larger claim size amounts are eliminated by reinsurance, the tail of the distribution function (d.f.) is of critical importance, whenever a heavy-tailed claim size d.f. F is under consideration.

The main properties of a frequently used analytic large claim size d.f. are considered in the unit 2 of this Diploma. The gamma, log-normal and Pareto distributions are typical candidates to be considered as an analytic d.f. of claim sizes which exceed some threshold.

The problem of coincidence between Burr distribution and General Pareto distribution (GPD) under testing the right tail of claim size models is raised in the unit 3.

Some of the basic tools from Extreme Value Theory (EVT) relevant for large losses are reviewed in the unit 4.

2 The main of the analytic claim size distribution functions.

A natural estimate for the claim size distribution is the observed size distribution, i.e. the claim size d.f. F is simply to be equal to the sample d.f.(tabular method)

$$F(Z) = \frac{number\ of\ claims\ of\ size \le Z}{total\ number\ of\ claims}. \tag{1}$$

The tabular method is appropriate only when there is a sufficiently large volume of claim data. This is rarely the case for the tail of the distribution, especially in situations where exceptionally large claims are possible. It is often desirable to try to find an explicit analytic expression for a claim size d.f. F . This is particularly the case if the claim statistics are too spars to use the tabular method. An analytic d.f. is often the easiest to handle in computations and programming, and it is often also used in those case where the sample distribution can be expected to provide a satisfactory model for the claim size. The suitable analytic d.f. may be found, for instance by curve-fitting, which is very close to the observed discrete d.f., and can be used instead.

(a) Three-parameter gamma distribution. One of the candidates for an analytic claim size d.f. F is the shifted gamma distribution, defined as $F(Z) = G(Z-d),\ Z > d$, where G is the $Gamma(\alpha, \beta)$ d.f. The d.f. F has three parameters, $\alpha,\ \beta,\ d,$ and is also referred to as the three-parameter gamma distribution.

For the given three main characteristics, the mean value m, the standard deviation $\sigma = \sigma_Z$, and the skewness $\gamma = \gamma_Z\ (>0)$ of the claim size distribution we may to find the values for the parameters $\alpha,\ \beta,\ d$ of the three-parameter gamma d.f. F so that these three characteristics match. So the following equations hold (see [1])

$$m = E(Z) = d + \frac{\alpha}{\beta}, \quad \sigma_Z^2 = \frac{\alpha}{\beta^2}, \quad \gamma_Z = \frac{2}{\sqrt{\alpha}}. \tag{2}$$

Solving for $\alpha,\ \beta,\ d$ gives

$$\alpha = \frac{4}{\gamma_Z^2}, \quad \beta = \frac{2}{\sigma_Z \cdot \gamma_Z}, \quad d = m - 2\frac{\sigma_Z}{\gamma_Z}. \tag{3}$$

We also may to calculate the limited expected value function L, which is the measure of goodness of fit of an analytic d.f. to the observed claim size d.f.

$$\begin{aligned} L(M) &\overset{def}{=} E(\min(M, Z)) = d + \frac{\alpha}{\beta} \cdot G_{\alpha+1,\beta}(M-d) - \\ &- (M-d)\cdot(1 - G_{\alpha,\beta}(M-d)), \quad M \ge d, \end{aligned} \tag{4}$$

where $G_{\alpha,\beta}$ denotes the $Gamma(\alpha, \beta)$ d.f.

The minimum d of a three-parameter gamma distribution is determined by the three basic characteristics of the claim size d.f. It is often the case that the initial point d of the d.f. fits the observed claim data badly. Then, if the range of small claims is important for the application, the three-parameter gamma d.f. has to rejected. It is even more fatal if the fit is bad for large claim. The three-parameter gamma d.f. is sometimes also used as an approximation for the d.f. F of the aggregate claim amount, if the skewness is small (see [1]).

It should be noted that $Z = d + Q$, where Q is $Gamma(\alpha, \beta)$- distributed, and therefore some properties of $Gamma(\alpha, \beta)$ d.f. is very interested for us.

1) The sum of independent $Gamma(\alpha_1, \beta)$- and $Gamma(\alpha_2, \beta)$-distributed random variables is $Gamma(\alpha_1 + \alpha_2, \beta)$-distributed (see [1], [3]).

2) The $Gamma(1, \beta)$ d.f. F reduces to the exponential d.f. $F(Z) = 1 - e^{-\beta Z}, Z \geq 0$. Hence, the $Gamma(k, \beta)$ d.f. for positive integer values of k is obtained as the k-th convolution of the exponential d.f.

(b) Log-normal distribution. A frequently used claim size distribution is the log-normal. A claim size variable *Z* is log-normally distributed if it is of the form

$$Z = d + e^{W}, \tag{5}$$

where $W = W_{\mu,\sigma}$ is a normally distributed variable with mean value μ and standard deviation σ, and where *d* is the initial point of the range of the claim size variable *Z*.

Solving W from (5) gives $W = \ln(Z-d)$. Then $F(Z) = N\left(\frac{\ln(Z-d)-\mu}{\sigma}\right)$, and the density of the variable *Z* is

$$F'(Z) = \frac{1}{\sigma \cdot (Z-d) \cdot \sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(Ln(Z-d)-\mu)^2\right], \ Z > d, \tag{6}$$

where d, μ, σ are free parameters. For the given three main characteristics, the mean value *m*, the standard deviation $\sigma = \sigma_Z$, and the skewness $\gamma = \gamma_Z\ (>0)$ of the claim size d.f. *F* we may to find the values for the parameters d, μ, σ.

$$d = m - \frac{\sigma_Z}{\lambda}; \ \sigma^2 = \ln(1+\lambda^2); \ \ \mu = \ln(m-d) - \frac{1}{2}\sigma^2, \tag{7}$$

where λ is the auxiliary variable, which defined as the real root of the equation

$$\lambda^3 + 3\lambda - \gamma_Z = 0. \tag{8}$$

The limited expected value function of the log-normal distribution is

$$\begin{aligned} L(M) = & d + (m-d) \cdot N\left(\frac{\ln(M-d)-\mu-\sigma^2}{\sigma}\right) + \\ & + (M-d) \cdot \left[1 - N\left(\frac{\ln(M-d)-\mu}{\sigma}\right)\right] \end{aligned} \tag{9}$$

(c) Pareto distribution. An another frequently used analytic claim size distribution is the Pareto d.f. ($Pareto(\alpha, \beta, d)$ d.f.), which is defined by the formula

$$F(Z) = 1 - \left(\frac{d+\beta}{Z+\beta} \right)^{\alpha}, \quad Z \geq d, \tag{10}$$

where α, β, d are parameters which satisfy $\alpha > 0$ and $\beta > -d$.

The first parameter α, controls how heavy a tail the distribution has: the smaller the α, the heavier the tail. The second parameter β mainly affects the left-hand range of the distribution, and does not essentially change the tail of the distribution in the region where *Z* is significantly larger than β. The third parameter *d* constrains the range of Z. The parameter β can often be omitted as soon as the tail of the claim size distribution is to be fitted.

The main characteristics of the $Pareto(\alpha, \beta, d)$ d.f. are

$$\begin{aligned}
&m = \frac{\alpha \cdot d + \beta}{\alpha - 1}, \quad \alpha > 1; \\
&\sigma_Z^2 = \frac{\alpha \cdot (d+\beta)^2}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2; \\
&\gamma_Z = 2\frac{\alpha+1}{\alpha-3}\sqrt{\frac{\alpha-2}{\alpha}}, \quad \alpha > 3; \\
&L(M) = \frac{\alpha \cdot d + \beta \ + (M+\beta)(1-F(M))}{\alpha - 1}, \quad M \geq d.
\end{aligned} \tag{11}$$

Note that

$$L(M) \equiv L_\alpha(M) \to L_1(M) = d + (d+\beta)\ln\frac{M+\beta}{d+\beta}, \quad M \geq d. \tag{12}$$

Experience has shown that the Pareto formula (10) is often an appropriate model for the claim size distribution, particularly where exceptionally large claims may occur.

The generalisation of the Pareto d.f. is often represented by the Burr distribution

$$F(Z) = 1 - \left(\frac{v}{(Z-d)^r + v} \right)^{\alpha}, \quad Z \geq d, \tag{13}$$

where α, r, v are positive parameters. Indeed, with $r = 1$ and $v = d + \beta$ we obtain the formula (10). In actual fact there is generalised Pareto distribution (GPD)

$$GPD(x,s,k) = \begin{cases} 1 - (1 - kx/s)^{1/k}, & k \neq 0 \\ 1 - \exp(-x/s), & k = 0 \end{cases}, \tag{14}$$

where $x > 0$, *k* is the Shape parameter, $s > 0$ – the Scale parameter. Note for $k \leq 0$: $0 \leq x < \infty$, and for $k > 0$: $0 \leq x < s/k$. There are some special cases of GPD:

c1) $k = 0$: Exponential with mean s ;

c2) $k = 1/2$: Triangular distribution ;

c3) $k = 1$: Uniform distribution on $[0, s]$; (15)

c4) $k < 0$: Pareto distribution,

and there are many interested properties, some of which is :

p1) if Y is Exponentially distributed with mean 1, then $X = s\cdot(1-\exp(-kY)/k$ is GPD;

p2) If X is GPD(. , s, k), then the excess $\{X-u \mid X \geq u\}$ is GPD(. , $s`(u)$, k) ; (16)

p3) The mean excess function of a GPD, i.e., the conditional mean of excesses, is a <u>linear</u> function of the threshold u :

$$E(X-u \mid X \geq u) = b_0 + b_1 u \quad , \tag{17}$$

where $b_0 = \dfrac{s}{1+k}$, $b_1 = -\dfrac{k}{1+k}$.

3 Burr distribution function vs. generalised Pareto distribution function and large claims.

It turn out that the right tail of the ISO US (Insurance Service Office in the USA) liability claim severity model fitted to claims generally up to a limit of \$1 000 000 is a Burr distribution with x-exponent equal 1 or ISO Burr distribution (see [2]):

$$\Pr ob.(X \leq x) = 1-\left(\beta/(x+\beta)\right)^{\alpha} \tag{18}$$

for $x \geq$ some positive lower-truncation point t .

Note that the ISO Burr distribution is a GPD with

$$k = -\alpha^{-1}, \text{ and } s = \beta/\alpha .$$

For most kinds of liability exposures, the Shape parameter is approximately $\alpha \cong 1$.

For instance:

1) Commercial auto (13 tables)	$0.95 \leq \alpha \leq 1.45$;
2) Premises (3 tables)	$1.10 \leq \alpha \leq 1.70$;
3) Products (3 tables)	$1.10 \leq \alpha \leq 1.20$;
4) Medical Prof. Liability (6 tables)	$0.75 \leq \alpha \leq 1.40$.

On the other hand, an analysis of very, very large claims by a large, global reinsurance company Swiss Re has shown that their distribution could be approximated with a GPD with $k \cong -1$.

Hence, there is an interested question: *does this mean that from a very low truncation point, heavy tailed liability claims distributions are approximately GPD with* $k \cong -1$ *???*

In order to answer for this question we consider the tail approximation problem in the following unit.

4 Probabilistic laws applicable to insurance.

4.1 Central Limit Type Laws. Suppose that $X_1, X_2, ..., X_n$ are independent identically distributed, i.i.d., random variables with distribution function F ("underlying distribution"). Form partial sums:

$$S_n = X_1 + X_2 + ... + X_n, \quad n \ge 1 \tag{19}$$

Select normalising constants a_n and $b_n > 0$ to normalise (standardise) the partial sums

$$\frac{S_n - a_n}{b_n}. \tag{20}$$

Then the following questions are take place.

Question 1:

Do the normalised partial sums converge in distribution to a random variable Y with distribution function G ("limiting distribution")?

i.e., $\frac{S_n - a_n}{b_n} \to Y, \quad \text{as} \quad n \to \infty$?

Question 2:

What is the form of the "limiting distribution" G? That is, how do we characterise G?

The partial answer of the Question 2 is:

For many F's, G is standard normal: Y is standard normal distributed with d.f. Φ, where $a_n = n\mu$ and $b_n = \sigma\sqrt{n}$ (Central Limit Theorem).

The class of F distributions whose normalised partial sums converge in distribution to a standard normal Φ is called the Domain of Attraction of Φ, written as DA(Φ).

Question 3:

Is the standard normal the only "limiting distribution"? What is the Domain of Attraction of Φ ? That is, which F's belongs to DA(Φ) ?

The partial answer of the Question 2 is the following example:

Normalised partial sums from Cauchy distribution with density $f(x) = \frac{1}{\pi}\frac{\theta}{\theta^2 + x^2}$ do not have a normal limiting distribution. That is Cauchy distribution does not belong to DA(Φ) !

Question 4:

What is the form of the "limiting distribution" G?

Full answer:

The characteristic function of G is usually of the form:

$\varphi_Y(t) = \exp(-c|t|^{\alpha})$, where $0 < \alpha \le 2$. Note that the following Stable

Laws hold (see [3]):

1) if $\alpha = 2$, then G is standard normal;
2) if $\alpha = 1$, then G is Cauchy.

Since, in this work we are interested in the large losses, i.e. random variables $M_n = \max(X_1, X_2, ..., X_n)$, it should be noted that for $\frac{M_n - a_n}{b_n}$ there are some Limit Laws from *Extreme Value Theory* (EVT). Some of the basic tools from EVT relevant for large losses are reviewed in the following subunit.

4.1 Limiting distributions of excess losses in Extreme Value Theory.

As its name suggests, the Extreme Value Theory (EVT) is concerned with the modelling of extreme events and in the last few years various authors (Embrechts & Klüppelberg 1993, McNeil 1996) have noted that the theory is relevant to the modelling of extreme insurance losses. The key result in EVT is the Pickands-Balkema-de Haan theorem (Balkema & de Haan 1974, Pickands 1975) which essentially says that, for a wide class of distributions, losses, which exceed high enough thresholds, follow the GPD.

Before consider the main results from EVT relevant to large losses we consider the following easy exercise:

Exercise: It is stated that within a given portfolio, claims follow an exponential d.f. with mean 10 (thousands dollars, say). We have now observed 100 such claims with largest loss 50. What if the largest loss would have been 100?

Solution:

The basic assumption yields that

$X_1, X_2, ..., X_{100}$ are i.i.d. with d.f. $P(X_1 \le x) = 1 - e^{-x/10}$, $x \ge 0$.

Therefore, for $M_n = \max(X_1, X_2, ..., X_n)$,

$$P(M_{100} > x) = 1 - (P(X_1 \le x))^{100} = 1 - (1 - e^{-x/10})^{100}. \tag{21}$$

From this, we immediately obtain

$$\begin{aligned} &P(M_{100} \ge 50) = 0.4914 \\ &P(M_{100} \ge 100) = 0.00453. \end{aligned} \tag{22}$$

However, rather than doing the exact calculations above, consider the following asymptotic argument. First, for all $n \ge 1$ and $x \in R$,

$$\begin{aligned} &P\left(\frac{M_n}{10} - \ln n \le x\right) = P(M_n \le 10(x + \ln n)) = \\ &= \left(1 - \frac{e^{-x}}{n}\right)^n, \end{aligned} \tag{23}$$

so that

$$\lim_{n\to\infty} P\left(\frac{M_n}{10} - \ln n \le x\right) = e^{-e^{-x}} = \Lambda(x). \tag{24}$$

Therefore, use the approximation

$$P(M_n \le x) \approx \Lambda(x/10 - \ln n) \tag{25}$$

to obtain

$$\begin{aligned} P(M_{100} \ge 50) &\approx 0.4902 \\ P(M_{100} \ge 100) &\approx 0.00453, \end{aligned} \tag{26}$$

very much in agreement with the exact calculations above.

Suppose we were asked the same question, but now we would have much less specific information on $F(x) = P(X_1 \le x)$, could we still proceed? This is exactly the point where classical EVT enters. In the above exercise, we have proved the following result:

Theorem 1 *Suppose $X_1, X_2, \ldots, X_n$ are i.i.d random variables with distribution function $F \sim EXP(\lambda)$, then for $x \in R$:*

$$\lim_{n\to\infty} P(\lambda\, M_n - \ln n \le x) = \Lambda(x). \tag{27}$$

Here are the key questions:

Q1: What is special about Λ? Can we get other limiting distributions, possibly for other d.fs F?

Q2: How do we find the normalising constants λ and $\ln n$ in general? That is, find a_n and b_n so that

$$\lim_{n\to\infty} P\left(\frac{M_n - a_n}{b_n} \le x\right) \tag{28}$$

exists.

Q3: Given a limit coming out of Q1, for which d.fs F and normalising constants from Q2, do we have convergence to that limit? Can one say something about second order behaviour, i.e. speed of convergence?

The solution to Q1 forms part of the famous Fisher-Tippett theorem (see [4]).

Theorem 2 *Suppose $X_1, X_2, \ldots, X_n$ are i.i.d random variables with distribution function F and a_n, b_n are constants so that for some non-generate limiting distribution G,*

$$\lim_{n\to\infty} P\left(\frac{M_n - a_n}{b_n} \le x\right) = G(x)\ , \qquad x \in R. \tag{29}$$

Then G is of one of the following types:

- *Type 1 (Fréchet):*

$$\Phi_\alpha(x) = \begin{cases} 0, & x \le 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases} \qquad \alpha > 0 \tag{30}$$

- *Type 2 (Weibull):*

$$\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha), & x \le 0 \\ 1, & x > 0 \end{cases} \qquad \alpha > 0 \tag{31}$$

- *Type 3 (Gumbel):*

$$\Lambda(x) = \exp(-e^{-x})\ , \qquad x \in R\ . \tag{32}$$

The distributions of one of the above three types are called *extreme value distributions*. Alternatively, any extreme value distribution can be represented by *generalised extreme value distribution* (GEVD)

$$H_{\xi;\mu,\sigma}(x) = \exp\left\{-\left(1 + \xi\frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right\}, \qquad x \in R\ . \tag{33}$$

Here $\xi \in R, \mu \in R$ and $\sigma > 0$. The case $\xi > 0 (\xi < 0)$ corresponds to the Freshet (Weibull)-type d.f. with $\xi = 1/\alpha\,(\xi = -1/\alpha)$, whereas by continuity $\xi = 0$ corresponds to the Gumbel or double exponential-type d.f. Some examples of extreme value distributions are given in figure 1. Note that the Freshet case (the Weibull case) corresponds to a model with finite lower (upper) bound; the Gumbel is two-sided unbounded.

Answering Q2 and Q3 is much more complicated. Below a complete answer (due to Gnedenko) for the Freshet case is formulated. This case is the most important for applications to (re) insurance. For a general d.f. *F*, the inverse of *F* is defined as follows:

$$F^{\leftarrow}(t) = \inf\{x \in R : F(x) \ge t\},\ \ 0 < t < 1. \tag{34}$$

Using this notation, the *p*-quantile of *F* is defined as

$$x_p = F^{\leftarrow}(p), \quad 0 < p < 1. \tag{35}$$

Theorem 3 *Suppose* $X_1, X_2, ..., X_n$ *are i.i.d random variables with distribution function F satisfying*

$$\lim_{t\to\infty} \frac{1 - F(tx)}{1 - F(x)} = x^{-\alpha}, \qquad x > 0,\ \alpha > 0. \tag{36}$$

Then for x > 0,

$$\lim_{n\to\infty} P\left(\frac{M_n - a_n}{b_n} \le x\right) = \Phi_\alpha(x) \tag{37}$$

where $a_n = 0$, $b_n = F^{\leftarrow}(1 - 1/n)$. *The converse of this result also holds true.*

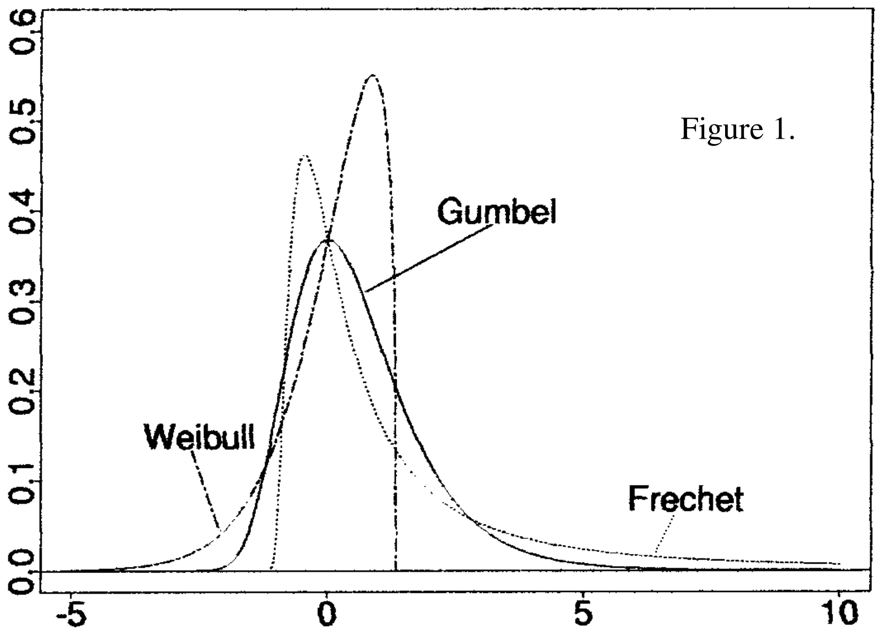

Figure 1.

A d.f. F satisfying (36) is called regularly varying with index $-\alpha$, denoted by $\overline{F} = 1 - F \in \Re_{-\alpha}$. An important consequence of the condition $\overline{F} \in \Re_{-\alpha}$ is that for a r.v. with d.f. F

$$EX^{\beta} = \begin{cases} < \infty, & \beta < \alpha, \\ = \infty, & \beta > \alpha. \end{cases} \tag{38}$$

In insurance applications, one often finds α-values in the range (1, 2). Theorem 3 is also formulated as: the maximal domain of attraction (MDA) of Φ_α is $\Re_{-\alpha}$, i.e.

$$MDA(\Phi_\alpha) = \Re_{-\alpha}. \tag{39}$$

D.fs belonging to $\Re_{-\alpha}$ are for obvious reasons also called of Pareto type. For instance, if F is Pareto(one parameter) d.f., i.e.

$$F(x) = 1 - kx^{-\alpha}, \qquad \alpha > 0,\ k > 0,\ x \ge k^{1/\alpha}$$

then

$$a_n = 0, \quad b_n = (kn)^{1/\alpha}. \tag{40}$$

Though we can calculate the normalising constants, the calculation of b_n depends on the tail of *F*, which in practice is unknown.
Let $x_F \equiv \sup\{x \in R : F(x) < 1\}$ $(\leq \infty)$ - "right endpoint" of a d.f. *F*. Then for two other cases (Weibull and Gumbel cases), the necessary and sufficient conditions for $F \in MDA(GEVD)$ are as follows:

Theorem 4

1) ***(Weibull case)*** $F \in MDA(\Psi_\alpha)$ *if and only if*

$$x_F < \infty \quad and \quad \lim_{h \to 0} \frac{1 - F(x_F - xh)}{1 - F(x_F - h)} = x^{\alpha}, \qquad (41)$$

$$\alpha > 0, for \;\; each \;\; x > 0$$

2) ***(Gumbel case)*** $F \in MDA(\Lambda)$ *if and only if*

for some srictly positive function $g(t)$

$$\lim_{t \to \infty} \frac{1 - F(t + xg(t))}{1 - F(t)} = e^{-x}, \quad for \;\; x \in R. \qquad (42)$$

Note that $MDA(\Phi_\alpha)$ includes such heavy-tailed distributions as Pareto, Burr, loggamma, Cauchy and *t*-distributions, and for $F \in MDA(\Phi_\alpha)$ we have $x_F = \infty$. This class is interested in insurance applications because insurance loss data are generally heavy-tailed.

The $MDA(\Lambda)$ contains such diverse d.fs as the exponential, normal, log-normal, gamma. These distributions are called medium-tailed distributions and the distribution of excess over sufficiently high thresholds may in this case approximated by the exponential distribution. Note that, the log-normal distribution has a moderately heavy tail and has historically been a popular model for loss severity distributions, however it is not as heavy-tailed as the distributions in $MDA(\Phi_\alpha)$.

Distributions in $MDA(\Psi_\alpha)$ are short-tailed distributions such as the Uniform and Beta distributions. This class is of lesser interest in insurance applications.

Theorem 3 and Theorem 4 are the basis of EVT. In order to show how this theory can be put into practice, consider for instance the pricing of an XL-treaty. Typically, the priority (or attachment point) *u* is determined as a *t*-year event corresponding to a specific claim event with claim size d.f. *F*, say. This means that

$$u = u_t = F^{\leftarrow}(1 - 1/t). \qquad (43)$$

In our notation (35) used before, $u_t = x_{1-1/t}$. Whenever *t* is large (typically the case in the catastrophic (i.e. rare) event situation), the following result due to Balkema, de Haan, (see [4]) is very useful.

<u>Theorem 5 (Balkema-de Haan)</u> *Suppose* $X_1, X_2, ..., X_n$ *are i.i.d random variables with distribution function F. Equivalent are:*

1) $F \in MDA(H_\xi)$, $\xi \in R$,
2) *for some function* $\beta : R^+ \to R^+$,

$$\lim_{u \to x_F} \sup_{0 < x < x_F - u} \left| F_u(x) - GPD(x, \beta(u), -\xi) \right| = 0 , \tag{44}$$

where $F_u(x) = P(X - u \le x \mid X > u)$.

It is exactly the so-called *excess* d.f. F_u which both risk managers as well as reinsurers should be interested in. Theorem 5 states that for large *u*, F_u has a generalised Pareto d.f. Now in order to estimate the tail $\overline{F}(u+x)$ for a fixed large value of *u* and all $x \ge 0$ consider the trivial identity

$$\overline{F}(u+x) = \overline{F}(u)\,\overline{F}_u(x), \qquad u, x \ge 0. \tag{45}$$

In order to estimate $\overline{F}(u+x)$, one first estimates $\overline{F}(u)$ by the empirical estimator

$$(\overline{F}(u))^\wedge = \frac{\#\{1 \le i \le n : X_i > u\}}{n}. \tag{46}$$

In order to have a "good" estimator for $\overline{F}(u)$, we need *u* not too large: the level *u* has to be well within the data. Given such a *u*-value, we approximate $\overline{F}_u(x)$ via (44) by

$$(\overline{F}_u(x))^\wedge = \overline{GPD}(x, \hat{\beta}(u), -\hat{\xi}) \tag{47}$$

for some estimators $\hat{\beta}(u)$, $\hat{\xi}$ depending on *u*.

There are some methods of estimation of GPD parameters:

a) <u>Based on mean excess function</u>

Using property (17) of GPD, if *X* is GPD (., s, k), then $e(u) = E(X - u \mid X \ge u) = b_0 + b_1 u$ denotes a "Population Line", where

$b_0 = \dfrac{s}{1+k}$, $b_1 = -\dfrac{k}{1+k}$, i.e. $e(u) = E(X - u \mid X \ge u) = \dfrac{s - ku}{1+k}$.

Sample mean excess function, defined by

$$e_n(u) = \frac{\sum_{i=1}^{n} (X_i - u)^+}{\sum_{i=1}^{n} 1_{\{X_i > 1\}}}, \tag{48}$$

is an empirical estimate of the mean excess function *e*(*u*).

From properties p1) and p2) we can state that $\{X - u \mid X \ge u\}$ is GPD(. , *s*`(*u*), *k*) and $e(u) = \dfrac{s`(u)}{k+1}$, i.e. $e(u) = \dfrac{s`(u)}{k+1} = \dfrac{s - ku}{k+1}$.

Hence we have

$$s`(u) = s - ku\,. \tag{49}$$

b) <u>Based on the method of moments</u>

If X is GPD, then using property p1) of GPD, we have

$$E\left(1-\frac{k}{s}X\right)^r = \frac{1}{1+kr}\text{, provided } 1+kr > 0$$

$$\hat{s} = \frac{1}{2}\overline{X}\left(\frac{(\overline{X})^2}{(\overline{X^2})-(\overline{X})^2}+1\right),\quad \hat{k} = \frac{1}{2}\left(\frac{(\overline{X})^2}{(\overline{X^2})-(\overline{X})^2}-1\right). \tag{50}$$

c) <u>Based on the method of probability-weighted</u>

Moments

$$a'_r = E\{X[1-F(X)]^r\} = \frac{s}{(r+1)(r+1+k)}\text{, provided } k > -1$$

$$\hat{s} = \frac{2a'_0\,a'_1}{a'_0 - 2a'_1},\quad \hat{k} = \frac{a'_0}{a'_0 - 2a'_1} - 2\,. \tag{51}$$

d) <u>Based on maximum likelihood estimation</u>

In end of this work we apply the method a) to 5724 flood insurance claims in Ukrainian property insurance (from Oranta Insurance Company).

The sample mean excess plot for this data is shown in figure 2.

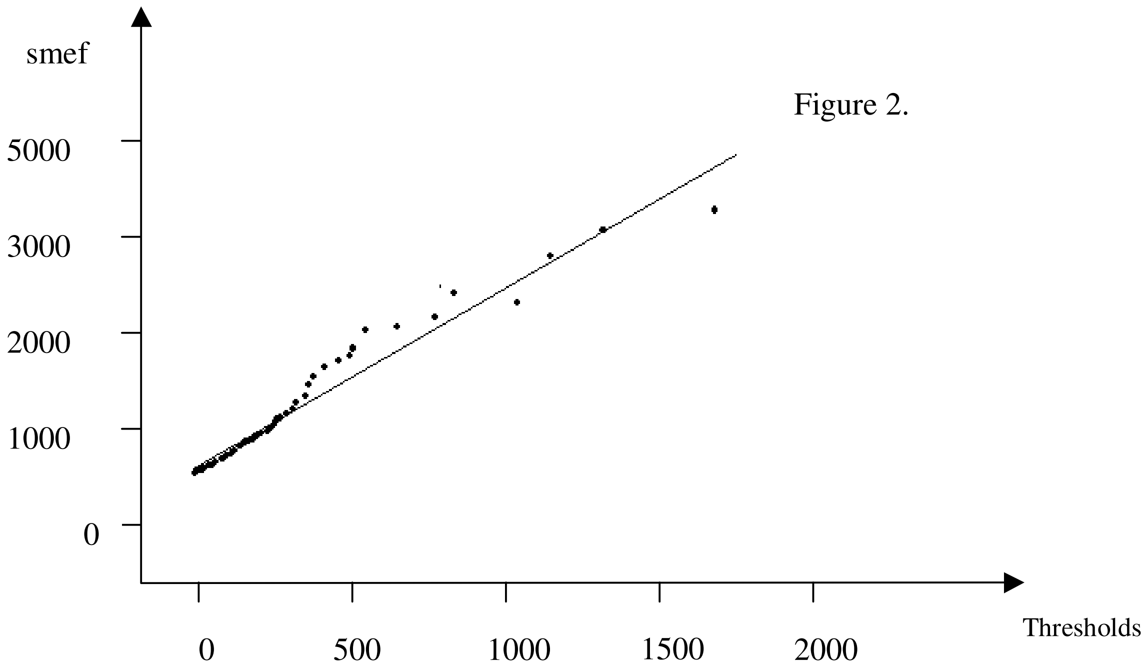

Figure 2.

In fact, apart from the extreme right hand points that are averages of only a small number of large excesses, the whole plot is approximately linear. In order to find a fit of the underlying excess d.f. F_u by GPD we have chosen a very high threshold (in our case it equals 380). This threshold exceeded by 78 losses and represents a threshold at about the 96th percentile of the data. In figure 3 GPD is fitted to 78 losses that exceed the threshold 380. The parameter estimates are $\widehat{k} = -0.677$ and $\widehat{s} = 63.33$, well then $\widehat{s}`(380) = 316.79$.

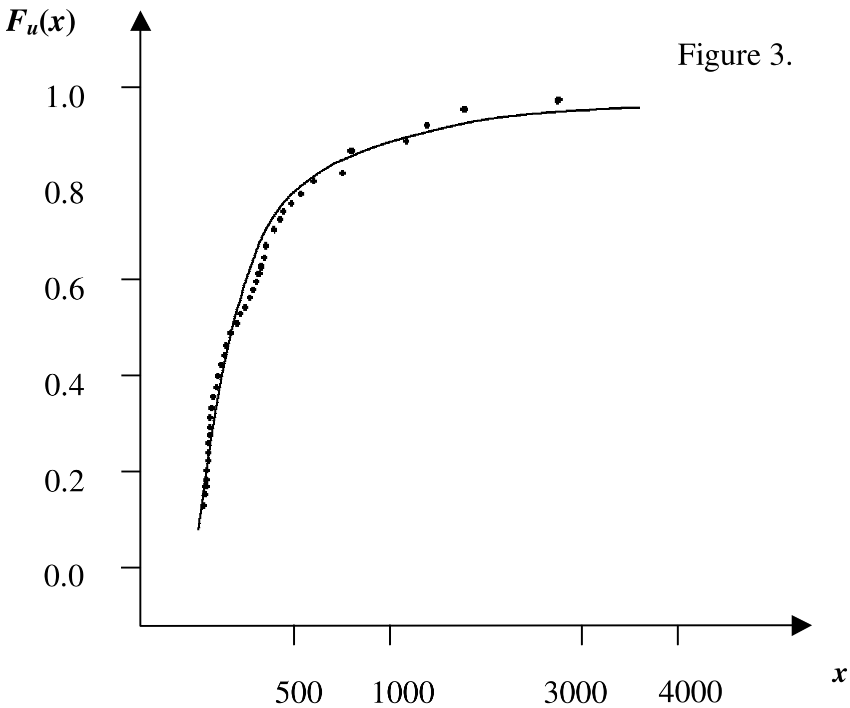

Figure 3.

5 References

1. Currie, I. D. (1992) *Loss Distributions.* Heriot-Watt University, Edinburg.
2. Gary Patrik and Farrokh Guahi (1998 CAS Spring Meeting) *An extremely important application of Extreme Value Theory to reinsurance pricing.* Swiss Re America.
3. Loeve, M. (1955) *Probability Theory.* Van Nostrand, New York.
4. Embrechts, P., Klüppelberg, C. and T. Mikosh (1997) *Modelling extremal events for insurance and finance.* Springer-Verlag, Berlin.