A large claim index

MARKUS AEBI, PAUL EMBRECHTS, and THOMAS MIKOSCH, Zürich

A Large Claim Index[1]

1 Introduction

An increasing demand is put on insurance mathematicians to adapt their, by now classical models so as to allow for extremal events or large claims. One of the key problems is the need for a dialogue between researchers and practitioners on the topic of 'what actually constitutes a large claim?' Many papers have already been written on this topic, see for instance the *Proceedings of the 4 Countries Astin-Symposium* (1984) for a series of papers on the subject, *Teugels* (1982) or *Embrechts/Veraverbeke* (1982). For a recent text on extreme value theory, see *Hahn* et al. (1991) and the references therein. In a recent discussion with an insurance practitioner, our statement concerning the Pareto distribution as a model for large claims was countered by: "the Pareto law, yes it means that 20 % of the individual claims are responsible for more than 80 % of the total claim amount!" Is there a way in which statistics can make the above statement more precise? The aim of the present paper is to show how an easy probabilistic argument yields a natural large claim index with which one can order classes of heavy-tailed distributions in a natural way. Based upon this index, statements like the one above can be made precise.

2 A large claim index

Let $(Y_1, \ldots, Y_n)$ be a (i.i.d) sample of positive random variables (r.v.'s) denoting the first n claims in a portfolio $Y = (Y_1, Y_2, \ldots)$, with distribution function (d.f.) F and finite mean $E(Y_1)$. Denote the associated ordered variables by $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$ where for instance $Y_{(1)} = \min(Y_1, \ldots, Y_n)$, the smallest claim, and $Y_{(n)} = \max(Y_1, \ldots, Y_n)$, the largest one. The total claim amount for the first n claims is $S_n = \sum_{i=1}^{n} Y_i$. If $\widehat{F}_n$ stands for the empirical

[1] Vorgetragen vom zweiten Autor an der Mitgliederversammlung der Schweizerischen Vereinigung der Versicherungsmathematiker vom 12. September 1992 in Winterthur.

distribution function (e.d.f), i.e.

$$\widehat{F}_n(y) = \frac{1}{n} \neq \{i \leq n : Y_i \leq y\}, \quad y \geq 0\,,$$

we denote by $\widehat{F}_n^{-1}$ the generalised inverse of $\widehat{F}_n$,

$$\widehat{F}_n^{-1}(x) = \inf\{y \geq 0 : \widehat{F}_n(y) \geq x\}\,, \quad 0 \leq x \leq 1\,.$$

$\widehat{F}_n^{-1}(y)$ is known as the empirical quantile function of F (or of $(Y_1, \ldots, Y_n)$). It follows immediately that for $1 \leq i \leq n$, $\widehat{F}_n^{-1}(\frac{i}{n}) = Y_{(i)}$.
With the above notation, the statement in the introduction regarding the Pareto law can be reformulated as the search for the d.f. of the ratio

$$T_n(\alpha) = \frac{Y_{([n\alpha])} + \cdots + Y_{(n)}}{S_n}, \quad 0 < \alpha < 1\,.$$

Here $[x]$ stands for the largest integer less than or equal to x. Hence $T_n(\alpha)$ is the proportion of the sum of the $(n - [n\alpha] + 1)$-st largest claims to the aggregate claim amount S_n in our portfolio. Exact distributional results for finite n are hard to obtain, hence we content ourselves with asymptotic estimates. Though variables of the form $T_n(\alpha)$ have been studied in the literature (see below), in order to keep the paper self-contained, we shall derive the basic results afresh. Already in *Aebi/Embrechts/Mikosch* (1992) (henceforth referred to as AEM) we indicated that the notion of Mallows metric has many applications to problems in insurance mathematics. Denote for $p > 0$:

$$\Gamma_p = \left\{ H \text{ d.f. on } \mathbb{R} : \int_{-\infty}^{+\infty} |x|^p H(dx) < \infty \right\},$$

the class of d.f.'s with finite p-th absolute moment.

Definition 1
For $F, G \in \Gamma_p$, the Mallows metric of order p is defined as

$$d_p(F,G) = \inf\{\|X - Y\|_p : \mathcal{L}(X) = F, \quad \mathcal{L}(Y) = G\}$$

where $\|X\|_p = (E|X|^p)^{\min(1,1/p)}$ and $\mathcal{L}(X) = F$ denotes the fact that the r.v. X has d.f. F. □

So $d_p(F,G)$ is the smallest distance in p-th absolute moment between r.v.'s X and Y having the d.f.'s F, respectively G. When X and Y are r.v.'s so that $\mathfrak{L}(X) = F$, $\mathfrak{L}(Y) = G$, then we often write $d_p(X,Y)$ for $d_p(F,G)$.
For a summary of the basic results on d_p, see *AEM* (1992) and the references therein. In the following Proposition we summarize the results needed later.

Proposition 1 [*AEM* (1992, Lemma 2.1)]

(i) *Let H^{-1} denote the (generalized) inverse of the d.f. H and the* r.v. *U be uniformly distributed on $[0,1]$. Then, for $p \geq 1$, $F, G \in \Gamma_p$:*

$$d_p(F,G) = \|F^{-1}(U) - G^{-1}(U)\|_p \,.$$

In particular,

$$d_1(F,G) = \int_{-\infty}^{+\infty} |F(x) - G(x)|\, dx \,. \tag{2.1}$$

(ii) *Suppose $p > 0$ and (X_n) is a sequence of r.v.'s with d.f.'s (F_n) respectively. Then $d_p(F_n, F_0) \to 0$ as $n \to \infty$, if and only if $\|X_n\|_p \to \|X_0\|_p$ and $X_n \xrightarrow{d} X_0$ (convergence in distribution) as $n \to \infty$.*

Our main (theoretical) result is the following.

Theorem 1

Suppose (Y_i) are positive, i.i.d. r.v.'s, with distribution F and $0 < E(Y_1) < \infty$. Let $\alpha \in [0,1]$, then as $n \to \infty$,

$$T_n(\alpha) = \frac{Y_{([n\alpha])} + \cdots + Y_n}{S_n} \overset{\text{a.s.}}{\to} \frac{1}{E(Y_1)} \int_{\alpha}^{1} F^{-1}(x)\, dx \,. \tag{2.2}$$

Proof: Since the result is trivial for $\alpha = 0$ or $\alpha = 1$ we suppose $\alpha \in (0,1)$. According to the Strong Law of Large Numbers, we have $\frac{S_n}{n} \overset{\text{a.s.}}{\to} E(Y_1)$ for $n \to \infty$. Moreover, we can write

$$\begin{aligned}\frac{1}{n}(Y_{([n\alpha])} + \cdots + Y_{(n)}) &= \frac{1}{n} \sum_{i=[n\alpha]}^{n} \widehat{F}_n^{-1}\left(\frac{i}{n}\right) \\ &= \frac{1}{n} \int_{[n\alpha]}^{n} \widehat{F}_n^{-1}\left(\frac{x}{n}\right) dx = \int_{[n\alpha]/n}^{1} \widehat{F}_n^{-1}(x)\, dx \,.\end{aligned}$$

Thus we have

$$T_n(\alpha) = \frac{1}{E(Y_1)} \int\limits_{[n\alpha]/n}^{1} \widehat{F}_n^{-1}(x)\, dx\, (1 + o(1)), \text{ a.s..}$$

Note that

$$\int\limits_{0}^{1} \widehat{F}_n^{-1}(x)\, dx = \frac{1}{n} \sum_{i=1}^{n} Y_i \to E(Y_1), \quad \text{a.s.,}$$

so that for large n, $T_n(\alpha)$ will be less than or equal to 1. Moreover,

$$\left| \int\limits_{\alpha}^{1} \widehat{F}_n^{-1}(x)\, dx - \int\limits_{\alpha}^{1} F^{-1}(x)\, dx \right| \le \int\limits_{0}^{1} \left| \widehat{F}_n^{-1}(x) - F^{-1}(x) \right| dx = d_1(\widehat{F}_n, F)\,,$$

by Proposition 1 i). Now because of the SLLN $(E(Y_1) < \infty)$ and the Glivenko-Cantelli Theorem, [see *Chow/Teicher* (1978, p. 261)], Proposition 1 ii) implies that $d_1(\widehat{F}_n, F) \to 0$ as $n \to \infty$. Moreover,

$$\int\limits_{[n\alpha]/n}^{\alpha} \widehat{F}_n^{-1}(x)\, dx \le \frac{1}{n} Y_{(n)} = o(1), \text{ a.s.,} \quad \text{as } n \to \infty\,.$$

The latter statement follows immediately from the Borel-Cantelli Lemma as shown in *Galambos* (1987, Corollary 4.3.1). Combining the above estimates, one obtains (2.2). □

Definition 2

Let F be a claimsize distribution with finite mean μ, $0 \le \alpha \le 1$. The Large Claim Index (l.c.i) *of F at α is defined as*

$$D_F(\alpha) = \frac{1}{\mu} \int\limits_{\alpha}^{1} F^{-1}(x)\, dx\,. \tag{2.3}$$

Its value indicates to what extend the $100(1-\alpha)\,\%$ largest claims in a claim portfolio contribute to the overall portfolio claim amount. □

In the next Paragraph we shall evaluate $D_F(\alpha)$ analytically for some F, compare our results with simulation data and 'solve' the Pareto law problem from the Introduction.

Remarks:

(i) The result in Theorem 1 is by no means new. Indeed functionals of the form D_F in (2.3) have been used in the economic literature for a long time. So is $L_F = 1 - D_F$ known as the Lorenz curve associated with F and is standardly used to model financial income data. For basic results and further references, see *Goldie* (1977) and *Villaseñor/Arnold* (1989) for instance. A good reference on the empirical version $D_{\widehat{F}_n}$ where $\widehat{F}_n$ is the e.d.f. of F, is *Csörgö* et al. (1986). Our Theorem 1, proved by different means, is to be found in the latter reference as Theorem 10.1.

(ii) Many more refined estimates on the convergence in (2.2) can be given. For instance, using the results in *Serfling* (1980, p. 283) it follows that for a general d.f. G on $\mathbb{R}$:

$$d_1(\widehat{G}_n, G) = O_p(n^{-1/2}), \quad \text{as } n \to \infty\,,$$

provided that

$$\int_{-\infty}^{+\infty} (G(x)(1 - G(x)))^{1/2}\, dx < \infty\,.$$

($O_p(\cdot)$ denotes of order $(\cdot)$ in probability). In our case, positive r.v.'s, this condition reduces to

$$\int_0^{\infty} (1 - F(x))^{1/2}\, dx < \infty\,,$$

which is fulfilled whenever $E(Y_1)^{2+\delta} < \infty$ for some $\delta > 0$. In the latter case, we have a $\sqrt{n}$-rate of convergence in Theorem 1. For further details on the asymptotic behaviour of empirical Lorenz curves, see *Csörgö* et al. (1986).

(iii) Theorem 1 can easily be adapted to cover the case where $(Y_1, \ldots, Y_n)$ is replaced by a sample $(Y_1, \ldots, Y_{N(t)})$ for some $t > 0$, where the claim arrival counting process $(N(t))_{t \geq 0}$ satisfies $N(t) \to \infty$ a.s., $t \to \infty$.

Theorem 1 gave the asymptotic behaviour of $T_n(\alpha)$ for claims with a finite first moment. In the case $E(Y_1) = \infty$, the following result yields useful information.

Theorem 2

Suppose $1 - F(y) \leq (1+y)^{-p}$ *for some* $p \in (0,1]$ *and* $y \geq y_0$, *and* $E(Y_1) = \infty$. *Then as* $n \to \infty$, $0 < \alpha < 1$, $T_n(\alpha) \xrightarrow{P} 1$.

Proof: Since $\frac{S_n}{n} \xrightarrow{\text{a.s.}} \infty$, it suffices to show that

$$\frac{Y_{(1)} + \cdots + Y_{([n\alpha])}}{n} = O_p(1)\,,$$

i.e. remains bounded in probability as $n \to \infty$.
Now $1 - F(y) \leq (1+y)^{-p}$ for some $y \geq y_0$, implies that for some x_1 and $x \geq x_1$:

$$F^{-1}(x) \leq (1-x)^{-1/p} - 1\,.$$

By Taylor expansion, the latter term remains bounded by Cx for some constant $C > 0$ and $x \leq x_2$, say. Thus

$$\frac{Y_{(1)} + \cdots + Y_{([\alpha n])}}{n} \leq \frac{y_0 \neq \{i \leq [\alpha n] : Y_{(i)} \leq y_0\}}{n} + \frac{1}{n} \sum_{i \leq [\alpha n] : Y_{(i)} \geq y_0} Y_{(i)}\,.$$

Note that for $(\xi_i)_{i=1,\ldots,n+1}$ i.i.d. with an EXP(1)-distribution,

$$\left(F^{-1}\left(\frac{\xi_1 + \cdots + \xi_i}{\xi_1 + \cdots + \xi_{n+1}}\right)\right)_{i=1,\ldots,n} \stackrel{d}{=} (Y_{(i)})_{i=1,\ldots,n}$$

where $\stackrel{d}{=}$ denotes equality in distribution. See for instance *Bickel* and *Doksum* (1977, p. 44, 46). For this representation of the $Y_{(i)}$ and every fixed n

$$\frac{1}{n} \sum_{i \leq [\alpha n] : Y_{(i)} \geq y_0} Y_{(i)} \leq \frac{1}{n} C \sum_{i=1}^{[\alpha n]} \frac{\xi_1 + \cdots + \xi_i}{\xi_1 + \cdots + \xi_{n+1}} \leq C\,,$$

hence the result follows. □

The latter Theorem implies that for d.f.'s with infinite mean, the largest claims essentially determine the overall claim amount. This result should be compared and contrasted with the following known results in the case where $1 - F(x) \sim$

$x^{-\alpha}L(x)$ for $0 \leq \alpha < 1$ and L is slowly varying in Karamata's sense [see *Bingham/Goldie/Teugels* (1987), p. 419].

Proposition 2

Let $Y_1, \ldots, Y_n$ be positive i.i.d. r.v.'s with d.f. $F, S_n = \sum_{i=1}^{n} Y_i$ and $Y_{(n)} =$ $\max(Y_1, \ldots, Y_n)$.

(i) *(Maller-Resnick) Equivalent are:*
a) *$Y_{(n)}/S_n \to 1$ in Probability,*
b) *$1 - F$ is slowly varying.*

(ii) *(Breiman) Equivalent are:*
a) *$Y_{(n)}/S_n$ has a non-degenerate limit distribution,*
b) *$1 - F(x) \sim x^{-\alpha}L(x)$, L slowly varying and $0 < \alpha < 1$,*
c) *$E(S_n/Y_{(n)} - 1)$ has a positive finite limit.*

3 Numerical results

In Figure 1 (see Appendix) we have plotted the l.c.i. D_F for a wide range of d.f.'s rangeing from light-tailed distributions (gamma type) to heavy-tailed ones (Pareto). The diagonal would correspond to the Uniform distribution. Notice that the '20 – 80 % rule' or the so-called Pareto law from the Introduction corresponds to the point $(\alpha = 0.8,\ D_F(\alpha) = 0.8)$ which lies near curve 5 indicating that such data typically correspond to a Pareto $(p \approx 1.5)$ distribution having finite mean but infinite variance. The corresponding l.c.i. for instance for an exponential distribution is roughly 0.5, indicating that the 20 % largest claims account for 50 % of the total claim amount.
The parameterizations used are:

- Pareto $(p) : 1 - F(x) = (1 + x)^{-p}, \quad x \geq 0.$
- Loggamma $(a, \gamma; x_0)$:

$$f(x) = \frac{a^\gamma}{\Gamma(\gamma)x_0}(\log(x/x_0))^{\gamma-1}(x/x_0)^{-a-1}, \quad x > x_0\,.$$

  If deleted from the parameterization, $x_0 = 1$.
- Lognormal (μ, σ):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}x}e^{-\frac{1}{2}\sigma^2(\log x - \mu)^2}, \quad x > 0\,.$$

  The parameterization sdlog = 1 stands for $\sigma = 1$ and μ general.

- Gamma $(a, \nu) = \frac{1}{\Gamma(\nu)} a^\nu x^{\nu-1} e^{-ax}$, $x > 0$.
  The parameterization Gamma (shape $= k$) means $a = 1$ and $\nu = k$.
  Exponential $(\lambda) =$ Gamma $(\lambda, 1)$.

In all these cases the l.c.i. D_F can be calculated analytically.

$p\backslash 1-\alpha$	0.5	0.4	0.3	0.2	0.1	0.05	0.01	0.005	0.001
1.01	0.998	0.997	0.995	0.992	0.986	0.980	0.965	0.958	0.943
1.05	0.991	0.985	0.976	0.963	0.936	0.908	0.843	0.816	0.756
1.1	0.983	0.972	0.956	0.930	0.882	0.833	0.723	0.679	0.587
1.2	0.969	0.950	0.922	0.878	0.798	0.718	0.555	0.495	0.379
1.4	0.948	0.918	0.873	0.804	0.685	0.575	0.372	0.306	0.194
1.7	0.928	0.886	0.825	0.736	0.589	0.460	0.248	0.188	0.098
2	0.914	0.865	0.795	0.694	0.532	0.397	0.190	0.136	0.062
3	0.890	0.829	0.744	0.626	0.446	0.307	0.119	0.078	0.028
5	0.872	0.802	0.708	0.580	0.392	0.255	0.086	0.052	0.016
10	0.859	0.784	0.684	0.549	0.359	0.225	0.068	0.040	0.011
1000	0.847	0.767	0.661	0.522	0.331	0.200	0.056	0.032	0.008
Exp(.)	0.847	0.767	0.661	0.522	0.330	0.200	0.056	0.031	0.008

Table 1: Table of large claim index D_F for different p-Pareto distributions and the exponential distribution.

Figure 2 plots D_F for the Pareto family and the exponential distribution. These plots are summarised in Table 1 above where we see that indeed for $p = 1.4$, $D_F(0.8) = 0.804$ (corresponding to $1 - \alpha = 0.2$). We also observe that on an l.c.i.-basis, the results for Pareto $(p \geq 5)$ are very close to the exponential case. The dramatic situation for Pareto $(p \leq 1.2)$ is exemplified in the last column where, for instance for $p = 1.01$, nearly 95 % of the total claim amount is due to 0.1 % of the individual claims! In the extreme case of very heavy-tailed distributions, like those in Proposition 2, the corresponding l.c.i.-curve converges more and more to a curve with value 1 on (0,1).
The final two Figures 3 and 4, highlight the convergence of the empirical l.c.i. $D_{\widehat{F}_n}$ to D_F as discussed in Remark i) of the previous paragraph, in the case of an exponential distribution (Figure 3) and the Pareto $(p = 2)$ case in Figure 4. Even in this latter case, which violates the condition $E(Y_1^{2+\delta}) < \infty$ from Remark ii) in Paragraph 2, we still have a reasonable rate of convergence.

Conclusion

The large claim index $D_F(\alpha)$ may be useful in highlighting the appropriateness of certain claim-size distributions used in the modelling of catastrophic events. Its interpretation is close to how 'the actuary in the field' would describe what large claims are.

Acknowledgement

The authors would like to thank Herold Dehling for useful discussions related to this paper.

Markus Aebi, Paul Embrechts and Thomas Mikosch
Department of Mathematics
ETH-Zentrum
8092 Zürich

References

Aebi, M./Embrechts, P./Mikosch, T. (1992): Stochastic discounting and Mallows metric. ETH preprint.

Bickel, P.J./Doksum, K.A. (1977): Mathematical Statistics. Basic Ideas and Selected Topics. Holden-Day, San-Franciso.

Bingham, N.H./Goldie, C.M./Teugels, J.L. (1987): Regular Variation. Cambridge University Press, Cambridge.

Chow, Y.S./Teicher, H. (1978): Probability Theory. Springer-Verlag, New York.

Csörgö, M./Csörgö, S./Horváth, L. (1986): An Asymptotic Theory for Empirical Reliability and Concentration Processes. Lecture Notes in Statistics *33*, Springer Verlag, Berlin.

Embrechts, P./Veraverbeke, N. (1982): Estimates for the probability of ruin with special emphasis on the possibility of large claims. Insurance: Mathematics and Economics *1*, 55 – 72.

Galambos, J. (1987): The Asymptotic Theory of Extreme Order Statistics. Krieger Publ. Co., Malabar, Florida.

Goldie, C.M. (1977): Convergence theorems for empirical Lorenz curves and their inverses. Adv. Appl. Prob. *9*, 765 – 791.

Hahn, M.G./Mason, D.M./Weiner, D.C. (Eds.) (1991): Sums, Trimmed Sums and Extremes. Birkhäuser, Basel.

Proceedings of the 4 Countries ASTIN-Symposium (1984): Akersloot, September 1984, Astin-Group, The Netherlands.

Serfling, R.J. (1980): Approximation Theorems of Mathematical Statistics. John Wiley, New York.

Teugels, J.L. (1982): Large claims in insurance mathematics. Astin Bulletin *13* (2), 81 –88.

Villaseñor, J.A./Arnold, B.C. (1989): Elliptical Lorenz curves. J. Econometrics *40*, 327 – 338.

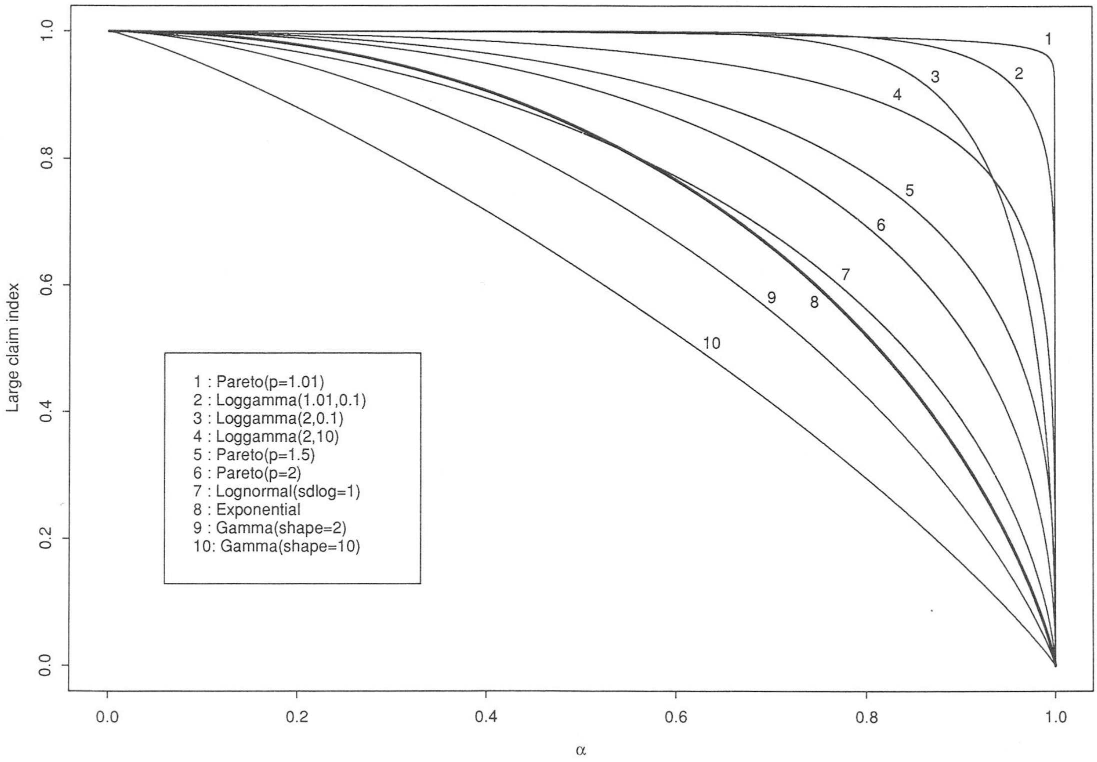

Figure 1: Large claim index D_F for several distribution functions

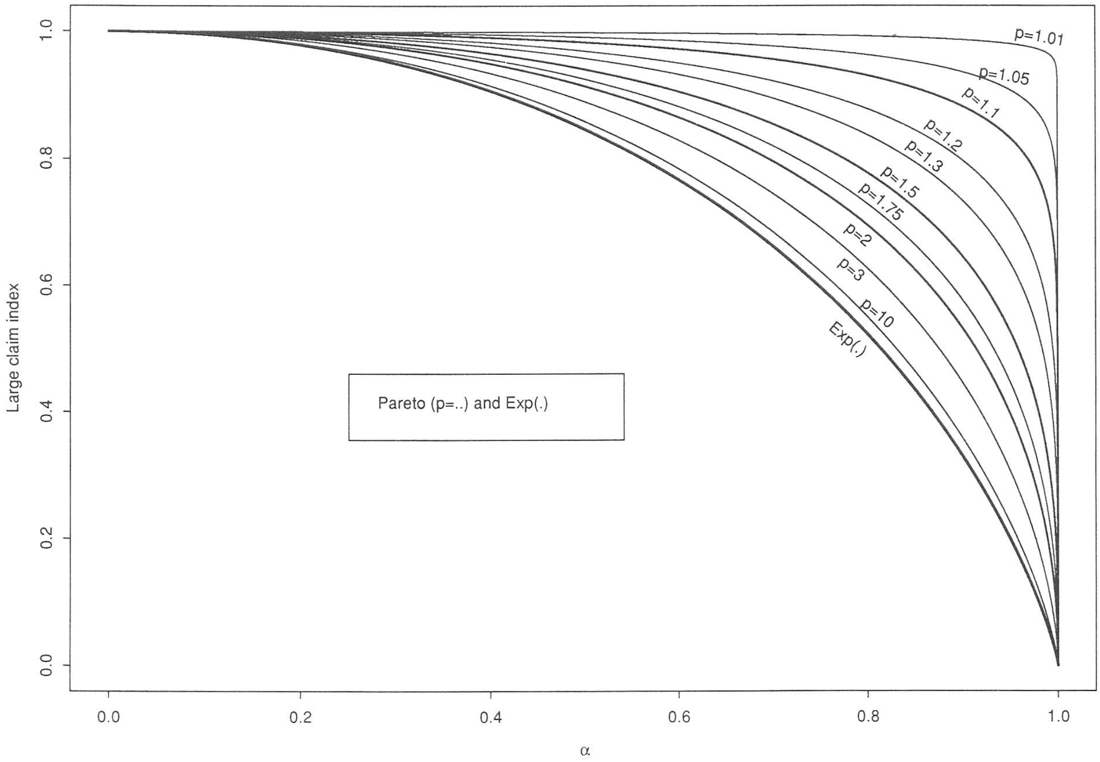

Figure 2: Large claim index D_F for the Pareto family and exponential distributions

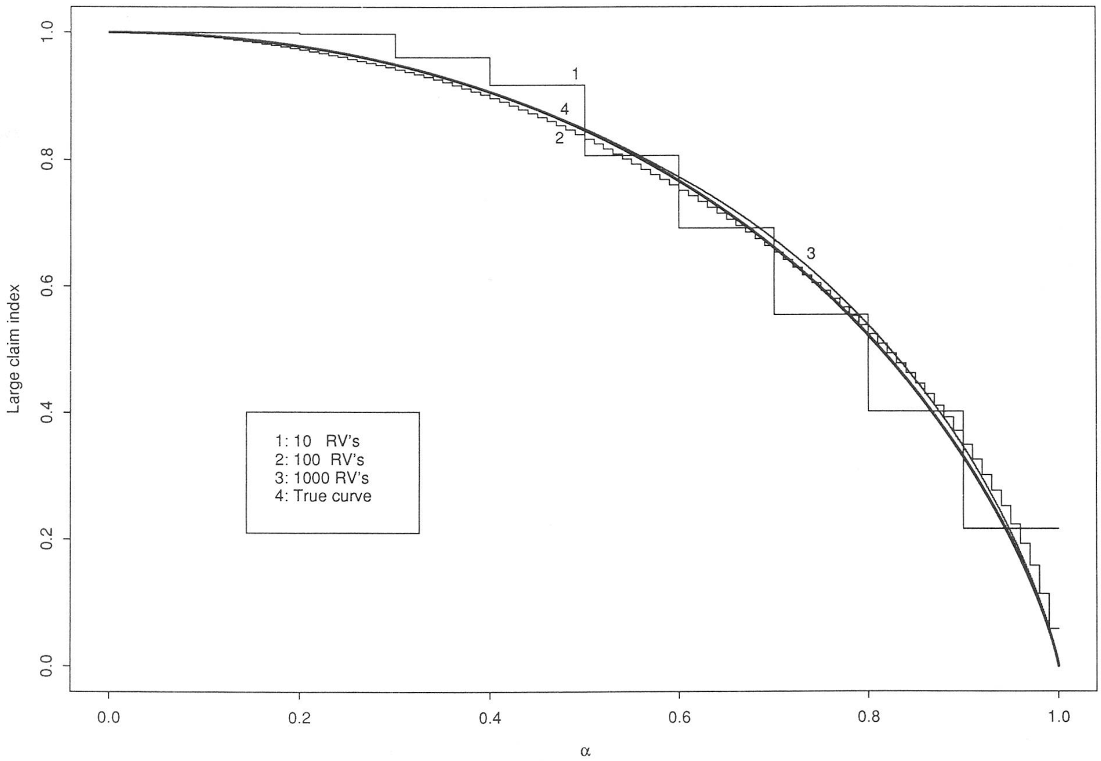

Figure 3: Convergence of the empirical large claim index $D_{\widehat{F}_n}$ to D_F for the exponential distribution

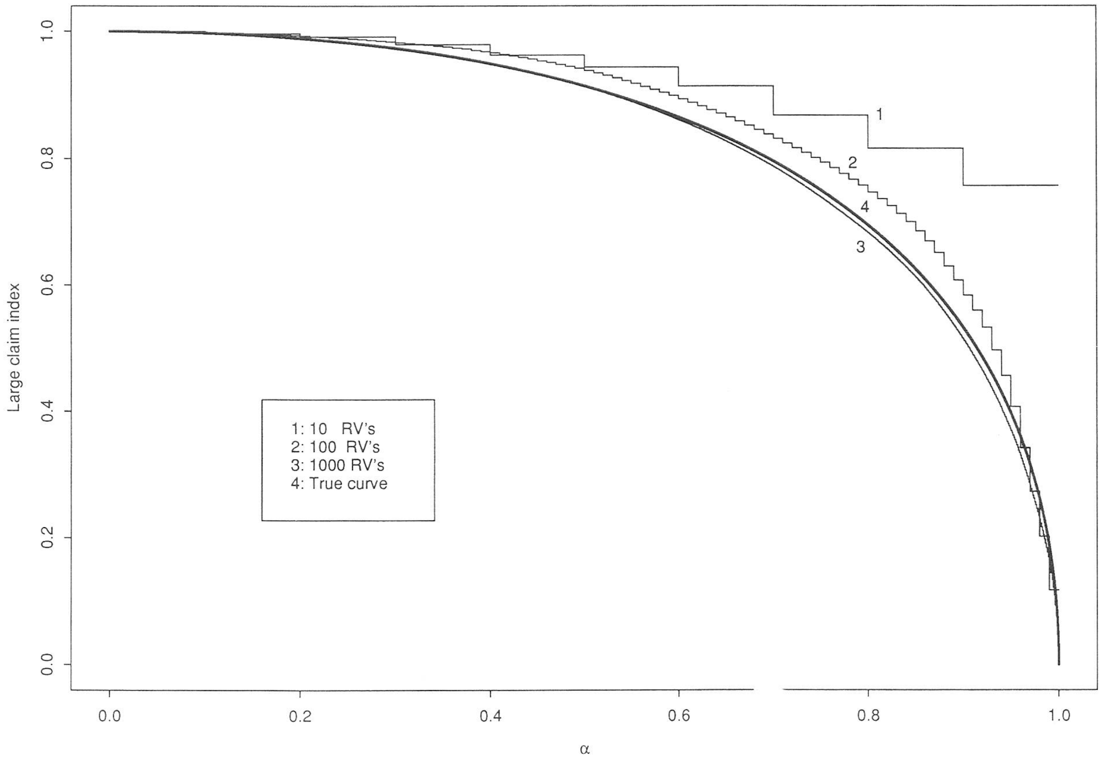

Figure 4: Convergence of the empiricial large claim index $D_{\widehat{F}_n}$ to D_F for the Pareto $(p = 2)$ distribution

Summary

A notion of large claim index is introduced allowing to distinguish between various families of claim-size distributions on the basis of their appropriateness to model catastrophic events.

Zusammenfassung

Es wird der Begriff des Grossschaden-Index eingeführt, welcher es erlaubt, zwischen verschiedenen Familien von Schadenverteilungen zur Modellierung von Katastrophen auszuwählen.

Résumé

On definit un index de grand sinistre. Cet index permet de choisir entre différentes familles de répartition de sinistres afin d'établir des modèles de catastrophes.