Extreme Value Theory in Finance and Insurance

Paul Embrechts

Department of Mathematics, ETH Zürich, Switzerland

The year 1998 had a long (financially) hot summer and autumn. During the LTCM–crisis and the following instabilities in financial markets worldwide, numerous critical statements on how these markets were behaving were filling newspapers and magazine columns. From the point of view of Extreme Value Theory (EVT), relevant examples include:

— The quants may have placed too much faith in their exquisitely tuned computer models... Sources say, Long–Term–Capital's worst–case scenario was only about 60% as bad as the one that actually happened. (*Business Week*, September 21, 1998).

— In a certain sense, maybe the problem wasn't too much rocket science, but too little. Extreme, synchronized rises and falls in financial markets occur infrequently but they do occur. The problem with the models is that they did not assign a high enough chance of occurrences to the scenario in which many things go wrong at the same time – the "perfect storm" scenario (Ibid.).

— These kind of moves are unprecedented in the floating rate era since the 1970s (after the 18% drop of the dollar against the Yen; S. Hannah in the *Financial Times*, October 9, 1998).

— Numerous critical remarks on existing risk management systems and techniques were voiced, here is an example: "What went wrong was the firm's risk–management model – which is similar to those used by the best and brightest banks." (*The Economist*, October 17, 1998)

However, already in November 1995, A. Greenspan stated:

— Work that characterizes the statistical distribution of extreme events would be useful as well.

Finally, in the already quoted October 17 Economist, the following two statements are pointing to the necessary work ahead:

— Regulators have criticised LTCM and banks for not "stress–testing" risk models against extreme market movements .... The markets have been through the financial equivalent of several Hurricane Andrews hitting Florida all at once. Is the appropriate response to accept that it was mere bad luck to run into such a rare event – or to get new forecasting models that assume more storms in future?

— VaR purports to give a bank an idea of the amount of money it is likely to lose over a certain period. Unfortunately, there are some big flaws in the VaR approach. The first is the word "likely" .... it seems that extreme outcomes happen more frequently than the models allow.

These events and comments definitely set the scene for a course correction (change of paradigms) in some of our RM tools. I am not claiming that EVT will give that new paradigm, but as we stated in the Reader Guidelines to Embrechts, Klüppelberg and Mikosch (1997), referred to as EKM below: "Though not providing a risk manager in a bank with the financial product he or she can use for monitoring financial risk on a global scale, we will provide that manager with stochastic methodology needed for the construction of various components of such a global tool."

It clearly is impossible to summarise the 650 pages of EKM in these few pages, I will try to summarise EVT's main potential input in insurance and banking. Applications in these fields which have already benefited from EVT methodology include:

— modelling claim–size distributions in non–life insurance;

— stress testing of catastrophe insurance portfolios;

— pricing of multi–line, multi–year, rare event (high layer) reinsurance products;

— pricing of catastrophe triggered bonds, and

— enhancement of integrated RM tools allowing to go beyond VaR.

I refrain from giving references on the above, beyond EKM, the interested reader is advised to check references either in EKM, or consult our Web–site http://www.math.ethz.ch/finance/ where examples of the above can be found in papers of myself and my colleagues at the ETHZ: Delbaen, Frey, McNeil and Schmock. Software programs, within S–Plus are freely available through the above web–site (look for EVIS under McNeil).

Before we state the main EVT results, it is imperative to understand where these methods deviate from classical models in finance and insurance, but also to understand which techniques remain similar. In order to do so, let us start from the easiest model: suppose we have data $X_1, \dots, X_n$ independent and identically distributed (iid) with distribution function F, i.e. $F(x) = P(X_1 \le x)$. The random variables (rvs) X_i may denote for instance:

— successive (log–) returns in a financial model (here, admittedly the iid assumption would not be very appropriate; this can however be corrected for by the fitting of a stochastic volatility model and then the X_i's would be the residuals);

— successive insurance claims, data on catastrophic events;

— losses in a credit portfolio, or

— historically or Monte Carlo simulated portfolio values in a VaR calculation.

"Classical" theory is all about sums $S_n = \sum_{i=1}^n X_i$ (total returns, accumulated claim amount), and through the general Central Limit Theorem (see EKM, Theorem 2.2.15, approximations in distributions to S_n lead to the class of α–stable distributions $(0 < \alpha \le 2)$ including the well–known normal distribution (Black–Scholes model) for $\alpha = 2$. Other "known" cases are the Cauchy ($\alpha = 1$) and the inverse Gaussian ($\alpha = 1/2$). By embedding the partial sums into a stochastic process (linear interpolation, say) we find as limiting processes Brownian motion ($\alpha = 2$) and the α–stable processes ($\alpha < 2$). All these processes have seen applications throughout finance, especially Brownian motion is the main building block of nearly all models in modern mathematical finance. The famous Basle Committee $\sqrt{t}$–rule for scaling a 1–day VaR to a t–day VaR is a consequence of its properties (and indeed of the above Central Limit Theory). I would like to stress however that $\sqrt{t}$ scaling really restricts to Brownian motion ($\alpha = 2$).

When one is interested in extremes in data, one is less concerned about S_n (or averages S_n/n), but much more in the largest observation $M_n = \max(X_1, \ldots, X_n)$, say. More generally, denote the ordered data as follows:

$$X_{n,n} = \min(X_1, \ldots, X_n) \leq X_{n-1,n} \leq \cdots \leq X_{2,n} \leq X_{1,n} = M_n \,.$$

Consequently, $X_{k,n}$ is the k–th largest observation among our data $X_1, \ldots, X_n$. In order to avoid difficulties in this notation for discrete data, assume that our df F has a density. The rvs $X_{n,n}, \ldots, X_{1,n}$ are called the order statistics. Insurance for a long time (especially reinsurance) has been interested, not only in S_n, but also in sums of few upper order statistics $X_{1,n}, \ldots, X_{k,n}$ say. Increasingly finance is (has to be) getting interested: the extreme observations are hurting (on the loss side) or pleasing (on the gains side). There now are two fundamental theorems, like the Central Limit Theorem, but now for M_n, or more generally for the $X_{k,n}$'s. I formulate these results in a mathematically non-fully–precise form, stressing the main ideas rather than giving the final details; see EKM for the latter.

Result 1 (Fisher–Tippett, Gnedenko)

The asymptotic distribution of normalised maxima M_n is of the following type:

$$H_\xi(x) = \begin{cases} \exp\left\{-(1+\xi x)_+^{-1/\xi}\right\} & \text{if } \xi \neq 0\,, \\ \exp\{-\exp(-x)\} & \text{if } \xi = 0\,. \end{cases}$$

Here the notation $x_+ = \max(x, 0)$. The case $\xi > 0$ is referred to as the Fréchet case, $\xi < 0$ the Weibull and $\xi = 0$ for obvious reasons the double exponential (or Gumbel). For applications to insurance and finance, mostly $\xi \geq 0$ is relevant. A consequence of the derivation of H_ξ as an appropriate limit distribution for normalised maxima is that H_ξ can be used directly in order to model data which themselves arise as maxima: e.g. X_i is the largest loss (written positively say) in a credit portfolio over a specific period of time. One can directly fit H_ξ (or more generally $H_{\xi;\mu,\nu}(x) = H_\xi\left(\frac{x-\mu}{\nu}\right)$) to the data using for instance maximum likelihood theory. However, for the applications we have in mind (excess–of–loss reinsurance treaties, measuring beyond VaR), the following result is more important. In deriving it, the theory underlying Result 1 is used extensively.

Result 2 (Pickands, Balkema-de Haan)

Again, suppose that our data are $(X,)\ X_1, \ldots, X_n$ iid with distribution function F. In many applications, for instance VaR or catastrophy excess–of–loss treaties, one is interested in estimating two quantities.

(a) High quantiles u_p of F, where $u_p = F^{-1}(p)$, $0 < p < 1$. Here F^{-1} is the properly define inverse (or quantile) function of F. This means that $P(X > u_p) = 1 - p$, which for p close to 1 (0.95 or 0.99 say) yields u_p as a high level which is only exceeded with small probability. VaR typically is such a number, also the so–called attachment point in an excess–of–loss reinsurance treaty. The problem in practice is of course that one doesn't know F, hence we have to estimate F from the available data (typically either simulated or observed). Also, for p very close to 1 (i.e. $(X > u_p)$ a very rare event), we may be at or even outside the range of available data.

(b) Once we have estimated u_p above (again think of VaR) one should be even more interested in the excess risk still available above u_p. For this, one needs to know the so–called conditional excess distribution:

$$F_u(x) = P(X - u \leq x \mid X > u)\,, \quad x \geq 0\,.$$

When $u = \text{VaR}$, I propose to call F_{VaR} the CDVaR: the Conditional Distribution above VaR. Result 2 now gives a canonical approximation for F_u (u large):

$$F_u(x) \approx G_{\xi,\beta(u)}(x)\,, \quad u \to \infty\,, \quad x \geq 0\,,$$

for some positive function $\beta(\cdot)$ and $G_{\xi,\beta}$ the so–called generalised Pareto distribution,

$$G_{\xi,\beta}(x) = \begin{cases} 1 - \left(1 + \xi \frac{x}{\beta}\right)_+^{-1/\xi}\,, & \xi \neq 0\,, \\ 1 - e^{-x/\beta}\,, & \xi = 0\,. \end{cases}$$

The meaning of "$\approx$" can be made precise; see for instance EKM, Theorem 3.4.13.

In summary: The distribution functions $\{H_\xi : \xi \in \mathbb{R}\}$ are used to model the stochastic behaviour of maxima of iid rvs. On the other hand $\{G_{\xi,\beta} : \xi \in \mathbb{R},\, \beta > 0\}$ is called for when one wants to model conditional excesses over high thresholds (going beyond VaR).

Of course there are various important conditions underlying the above. Indeed applications of EVT should be carefully checked against these conditions; see EKM for details. As already mentioned above, standard software (EVIS) is made available for doing the necessary computations in a typical risk management environment.

An Example: Rather than going more into the technical details of how EVT works, I present the numbers EVT produces coming out of a real example. The data underlying the model are fire insurance losses (similar examples typically would occur in credit risk management). The data base consists of historical loss data and the aim is to estimate the so–called 1 in 100 event, i.e. $u_{0.99}$ in our notation above (again, one could think of similar examples where $u_{0.99}$ is replaced by the 1–day VaR at 99%). Using EVT for this insurance portfolio the estimate for $u_{0.99}$ amounted to $\hat{u}_{0.99} = 27.3$ (Mio US$, say). Hence the probability of a loss exceeding 27.3 is 0.01. Now clearly one wants to learn about the uncertainty in this estimate. A full application of EVT (using maximum likelihood methodology) yields a 95% confidence interval for $u_{0.99}$ of (23.3, 33.1). Hence we are 95% sure that the latter interval contains the unknown quantity $u_{0.99}$. Further, management would (should) be interested in the next question: "What happens if we have a loss exceeding $\hat{u}_{0.99} = 27.3$? What is the excess risk still available in the data?". For this, we turn to $F_{27.3}$, or as an example we estimate the conditional mean loss above 27.3, i.e.

$$e(27.3) = E(X \mid X > 27.3)\,.$$

Again using EVT in its full strength, this conditional mean can be estimated as $\hat{e}(27.3) = 58.2$. Hence, given that there will be a loss above 27.3 Mio US$ (our estimate for the 1 in 100 event), we expect this loss to be 58.2 Mio US$. Once more, a 95% confidence interval can be calculated yielding (41.6, 154.7)! The latter interval may seem really wide: indeed this is the whole point of EVT, it produces best estimates on the extreme events. Hence the resulting confidence intervals turn out to be very skew to the right (as one would expect). Such (upper) boundaries of confidence intervals can then be used as a value for stress testing. Clearly, as they stand, the above results can easily be communicated to management and regulators. Translated into a VaR context (though more work is needed on getting the basic data) EVT not only produces overall a better VaR estimate (see for instance McNeil and Frey (1998)) but also allows for the construction of appropriate(!) confidence intervals around VaR and for the estimation (with confidence limits) of the excess risk above VaR. See Embrechts, Resnick and Samorodnitsky (1998) for a more detailed discussion on applications to insurance, and Schmock (1998) for an example in the realm of catastrophy risk securitisation

Needless to say that I have only scanned the tip of the EVT–iceberg. Various generalisations can be considered:

- non–iid data (think of stochastic volatility models in finance, or catastrophic event models containing trend and seasonality in insurance);
- regression type VaR models where the EVT parameters depend on some exogenous economic data;
- multivariate extremes, and

— the analysis of dependencies in more–dimensional, non–normal data (see for instance Embrechts, McNeil and Straumann (1999) for a start on this).

I am convinced that EVT will become a standard tool in integrated risk management. For the simple reason that, by definition, IRM is all about (hopefully) rare events.

References

P. Embrechts, C. Klüppelberg, T. Mikosch (1997) *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.

P. Embrechts, A. McNeil, D. Straumann (1999) Correlation and dependency in risk management: properties and pitfalls. ETH preprint.

P. Embrechts, S. Resnick, G. Samorodnitsky (1998) Extreme value theory as a risk management tool. *North American Actuarial Journal*. To appear.

A. McNeil, R. Frey (1998) Estimation of tail–related risk measures for heteroskedastic financial time series: an extreme value approach. ETH preprint.

U. Schmock (1998) Estimating the value of the Wincat coupons of the Winterthur insurance convertible bond: A study of the model risk. *ASTIN Bulletin*. To appear.

Paul Embrechts
Department of Mathematics
ETHZ
CH–8092 Zurich, Switzerland
Tel. +411 632 3419
FAX +411 632 1523
email: embrechts@math.ethz.ch