LIVING ON THE EDGE

PAUL EMBRECHTS, SIDNEY RESNICK AND GENNADY SAMORODNITSKY*

ETH Zurich and Cornell University

Abstract

Risk management is currently a crucial topic in the world of finance. We argue that an important analytic tool in risk management should be *extreme value theory* and briefly review what this subject has to offer and illustrate the potential uses through an example.

1 Introduction.

Mainly due to the increase in volume and complexity of financial instruments traded, Risk Management (RM) has become a (if not *the*) key issue in any financial institution or corporation of some importance. Having grown from a relatively small set of technical standards set by banks internally or enforced upon banks by regulatory bodies, RM is now becoming an organizational force within the company touching each hierarchical level. Globally accepted rules are put into place (see for instance GARP (1996)) aimed at monitoring and managing the full diversity of risk (credit, market, liquidity, operational, ...). The final outcome is a better understanding and grasp of the way financial institutions not only assess their internal handling of risk but more importantly how their clients can benefit from a fully transparent risk profile of products and services. It is clear that RM is increasingly being used as a marketing tool. Within the institution, at the more technical level, a properly functioning RM system allows for a risk adjusted assessment of return on capital.

Glancing through past issues of RISK, one is struck by the rather sharp increase some three years ago in papers discussing the wider issues of risk. Examples include case studies like Barings, Orange County or Metallgesellschaft, and the publication and ongoing discussion of J.P. Morgan's RiskMetrics. RiskMetrics has been (and still is) important in so far that its publication has focused discussions on some of the underlying technical issues. No discussion has perhaps been more heated than the one on Value–at–Risk (VaR), or as some want to call it Capital–at–Risk. In McNew (1996), the author states that VaR has proved

*Sidney Resnick and Gennady Samorodnitsky were partially supported by NSF Grant DMS-9400535 at Cornell University. Paul Embrechts gratefully acknowledges the hospitality of the School of Operations Research and Industrial Engineering, Cornell University during Fall 1996.

to be a mixed blessing, and continues by saying that perhaps the biggest problem with VaR is the main assumption in the conventional models, i.e., that portfolio returns are normally distributed. Taking up the "nice" case of foreign exchange data, McNew states that it is unfortunate for VaR theorists that evidence on the distribution function of exchange rates indicates that there are too many extreme observations relative to what one would find if exchange rates were normally distributed. Brady (1996) stresses a similar point as he remarks that the past is not the future: all attention is on the edges, where we know that the models break down.

For the purpose of our paper, some of the problems related to VaR are best captured in Boudoukh et al. (1995). These authors stress that, from a RM perspective, managers care more about the size of the losses than the number of times they will face a loss. They continue by offering a worst case scenario (WCS) approach which they claim "is a concept most often associated with the analysis of RARE OR EXTREME events. WCS is concerned with the nature of an event which *by definition* is bound to happen". They continue by stressing that, since VaR incorporates a significant "ruin" probability, be it 5% or 1%, there is a need for an additional layer of prudence via a larger capital requirement. The approach commonly used in the context of the VaR measure is to use the VaR number as an indication and then simply to multiply this measure by some "hysteria factor". Often a factor of 3 is used. On the more fundamental issue on how to define coherent risk measures in finance and showing that VaR is not such a measure, see Artzner et al. (1997). Finally, a comprehensive review on VaR is Jorion (1997). See the latter text on p. 96 for a reasoning behind the use of the above factor 3.

The above discussion sets the scene for our contribution. Recall the main points:

- RM is interested in estimating tail probabilities and quantiles of profit–loss distributions, and indeed of general financial data;
- extremes matter;
- we want to have methods for estimating conditional probabilities concerning tail–events: given that we incur a loss beyond VaR, how far do we expect the excess to go;
- financial data show fat tails.

Extreme Value Theory (EVT) is a subject whose motivations match the four points highlighted above. It is our conviction that EVT has a very important role to play in some of the more technical discussions related to RM issues. At this point in time, nobody will be able to come up with the only true answer concerning VaR and related risk measures. What one can say however is that EVT will play an important part of the working methodology.

In the next section we summarize some of the main results of EVT and indicate where they can be used in the overall RM context. Clearly, we are not able to give all details here and the interested reader is referred to Resnick (1987), Leadbetter et al. (1983) and Embrechts et al. (1997) for comprehensive overviews, the latter in particular giving emphasis to applications to finance and insurance. See also Zangari (1997) and Longin (1997). Finally, the present paper can also be seen as a reaction to the following remark made by Alan Greenspan, Chairman of the Board of Governors of the FED at a Research Conference on Risk Management and Systemic Risk, Washington, D.C., 16 November 1996: "Work that characterizes the statistical distribution of extreme events would be useful, as well." EVT is offering precisely the methodology underlying such a characterization.

2 EVT: Basic Results.

In order to state the main results from EVT in their easiest form, we concentrate on a sample $X, X_1, X_2, \ldots X_n$ of independent, identically distributed (iid) random variables (rvs) with common distribution function (df) F. The iid assumption can be relaxed: this is indeed very important to realize. EVT has been worked out for processes both in discrete as well as continuous time, with or without independence and/or stationarity assumptions. It is fair to say that for most models encountered in finance, relevant EVT tools are available. See Embrechts et al. (1997) and the references therein for more details. For the purpose of this paper, one may think of X_i as the loss (or gain) of transaction (portfolio) i, the i–th absolute log–return of an underlying financial instrument or the i–th claim relating to an insurance loss. The latter example is specially important as such events can be contingent for so–called Act–of–God or catastrophe–linked bonds. Indeed, EVT has proved to be particularly useful in modeling catastrophic claims in reinsurance; see McNeil (1997) and Resnick (1997). A final example concerns the X_i's as credit losses. Here EVT will undoubtedly become useful in order to estimate the so–called unexpected loss and the stress loss. The latter nomenclature is taken from SBC's ACRA (Actuarial Credit Risk Accounting) and is indeed also to be found in JP Morgan's Credit Metrics. For instance, within ACRA, the stress loss is defined as the possible – although improbable – extreme scenario which the Bank must be able to survive.

Classical probability theory underlying most of the stochastic methods used in finance concerns sums of the individual X_i's: $S_n = X_1 + \cdots + X_n$. The relevant theorems relating to $\{S_n\}$ are the Laws of Large Numbers (LLN) describing that sample averages S_n/n approximate the mathematical expectation $\mu = E(X)$ and the Central Limit Theorem (CLT) which says that $\{S_n\}$ centered and scaled to have mean 0 and variance 1 has approximately a normal distribution. Indeed, it is the CLT which underlies the log–normality assumption in the Black–Scholes model, yields Brownian motion as the corner stone of most analytic models and leads to analytic VaR estimates based on normal quantiles.

Depending on the case at hand, a typical VaR or risk based capital estimate is calculated as:

$$\begin{array}{c}\text{Current}\\ \text{value of}\\ \text{portfolio}\end{array} \times \begin{array}{c}\text{Sensitivity of}\\ \text{portfolio to}\\ \text{underlying factors}\end{array} \times \begin{array}{c}\text{Potential change}\\ \text{in underlying}\\ \text{factors}\end{array} \tag{2.1}$$

where the last factor usually is of the form $k\sigma$ where $k \in \{1, 2, 3\}$ and σ stands for the standard deviation of the underlying P&L. The factor $k\sigma$ should give the risk manager the necessary statistical confidence with respect to the adequacy of the estimate produced. The outcome of (2.1) is then often further multiplied by some hysteria factor; see Jorion (1997). This all hints at the fact that standard methods are not catering enough for the fat tails in the loss data.

To show how EVT offers tools and techniques with potential use in finance, we summarize below some its main ideas in the case where iid loss data $X_1, \ldots, X_n$ with common, but unknown, df F are available. Of course, in practice VaR is calculated directly from a data driven model (the Black–Scholes log–normal model say) and not from a specific sample of losses. At the backtesting and calibration level however, the above set–up may be more relevant. Also, for the ease of exposition and indeed in order to highlight the main EVT procedures, we will stick to a somewhat idealized, sample–based VaR calculation. The reader should have no problem in translating our findings to other areas of insurance and finance. For the moment, it is the key ideas that matter. Refinements can be built in later.

Whereas S_n/n would correspond to an "average loss", the most extreme case within the range of the data concerns the largest loss $M_n = \max\{X_1, \ldots, X_n\}$. More generally, order the data

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

and we might be interested in the behavior of the k largest losses $X_{1,n}, X_{2,n}, \ldots, X_{k,n}$. Based on the P&L data $X_1, \ldots, X_n$ and given a confidence level α, an empirical VaR estimate would produce the k–th largest observation $X_{k,n}$ where k is approximately αn. By producing that estimate (or indeed any more sophisticated VaR measure), we would give management dollar value which, based on our data, will typically only be surpassed in $\alpha 100\%$ of cases. Now often, we would have insufficient data (especially when α is small, $\alpha = 0.001$ and $n = 100$ say) so that we have to extrapolate beyond the range of the data. So here is our first fundamental problem concerning tail estimation for P&L distributions. For given (small) α, calculate the level (α–quantile) u_α so that

$$P\left(X > u_\alpha\right) = 1 - F\left(u_\alpha\right) = \alpha\,. \tag{2.2}$$

Remember that F is not known. If X stands for monthly (log–)returns for a particular portfolio, u_α would correspond to a so–called $1/\alpha$–month return period in the language of insurance. If for instance $\alpha = 0.05$, then u_α is the 20–month return period, i.e. that value which on average is only surpassed once in 20 months.

Secondly, once the above level u_α is fixed, one would be interested in estimating the potential losses above u_α: "If we are hit (beyond VaR); by how much!" Therefore we need to be able to estimate the conditional probability df

$$P\left(X - u_\alpha \leq x \mid X > u_\alpha\right)\,, \tag{2.3}$$

i.e. the conditional probability that, given a loss beyond u_α, the excess loss $X - u_\alpha$ is no bigger than some level x. An estimate of this conditional probability will (in the case of sufficient data) involve the losses $X_{\ell,n}, X_{\ell-1,n}, \ldots, X_{1,n}$ above some (large) loss $X_{\ell+1,n}$. If insufficient data are available, we have to find a suitable model or approximation for (2.3). Though the conditional df in (2.3) has long been used as a standard measure in fields like insurance (excess–loss) and reliability/medical statistics (residual–life), its importance in finance is only now becoming clear. Names like "shortfall", "beyond VaR", etc... are being used for the quantity

$$e(u_\alpha) = E(X - u_\alpha \mid X > u_\alpha) \ , \tag{2.4}$$

the (conditional) mean excess loss, given that a loss above u_α (VaR say) has occurred. Within the actuarial literature, much is known about $e(u_\alpha)$; see Embrechts et al. (1997). The function $e(u_\alpha)$ is very useful in distinguishing between short–tailed and fat–tailed dfs. In the former case, $e(u_\alpha)$ typically decreases (in the normal case even to 0), whereas for fat–tailed dfs $e(u_\alpha)$ increases for α tending to 0, i.e. for u_α tending to infinity. One easily shows (Embrechts et al. (1997), p. 161), that for normally distributed data $e(u_\alpha) \approx u_\alpha^{-1}$ and $E(X \mid X > u_\alpha) = e(u_\alpha) + u_\alpha \approx u_\alpha$. Hence, if $u_\alpha = \mathrm{VaR}$, then $E(X \mid X > \mathrm{VaR}) \approx \mathrm{VaR}$. However, in the fat–tailed Pareto case with tail–parameter $\beta > 1$, i.e. $1 - F(x) = (1+x)^{-\beta}$, $x \geq 0$, say, one easily shows that $e(u_\alpha) = (1+u_\alpha)/(\beta-1)$ and consequently $E(X \mid X > u_\alpha) \approx \frac{\beta}{\beta-1} u_\alpha$. Once more, if $u_\alpha = \mathrm{VaR}$, then in the fat–tailed Pareto case $E(X \mid X > \mathrm{VaR}) \approx \frac{\beta}{\beta-1}\mathrm{VaR}$. From empirical studies, it follows that within insurance often $1 < \beta < 2$, whereas in finance a range $1.5 < \beta < 5$ is standard. The consequence of this for applications of the VaR methodology are obvious! For further discussions and examples, see Artzner et al. (1997) and Longin (1997).

EVT offers empirical finance in general and Risk Management in particular methods for estimating quantities like (2.2)–(2.4) and indeed many related quantities under flexible model assumptions. Such models include time dependent–parameter models (non–stationarity) and models involving exogenous variables. Note also that EVT–based solutions of (2.2)–(2.4) allow for a wide variety of shapes of the underlying dfs (P&L distribution, return df, credit–loss df, insurance claims df, ...).

Without going into excessive detail on how EVT works, we mention the main ingredients in the solution of the above problems.

Fact 1. Under widely applicable conditions, the df of the largest observation M_n of an iid sample $X_1, \ldots, X_n$ can be approximated by a member of the following class of extreme value distributions:

$$H_{\xi,\mu,\psi}(x) = \exp\left\{ - \left(1 + \xi \, \frac{x-\mu}{\psi}\right)_+^{-1/\xi} \right\} .$$

Here $y_+ = \max(y, 0)$. This three–parameter family of distributions has a *location* parameter $\mu \in \mathbb{R}$, a *scale* parameter $\psi > 0$ and (most importantly) a *shape* parameter $\xi \in \mathbb{R}$. The case $\xi = 0$ is to be interpreted as

$$H_{0,\mu,\psi}(x) = \exp\left\{ - \exp\left(- \frac{x-\mu}{\psi} \right) \right\} , \quad x \in \mathbb{R},$$

and is referred to as the *double exponential* or *Gumbel* distribution. For $\xi > 0$, $H_{\xi,\mu,\psi}$ is called the *Fréchet* distribution, for $\xi < 0$ the *Weibull*. An important distinction is that the Fréchet has unbounded support to the right and the Weibull has unbounded support to the left. *Figure 1* contains the density functions for the standard cases $H_{\xi,0,1}$ for $\xi = 0$, $\xi = 2/3$, $\xi = -2/3$.

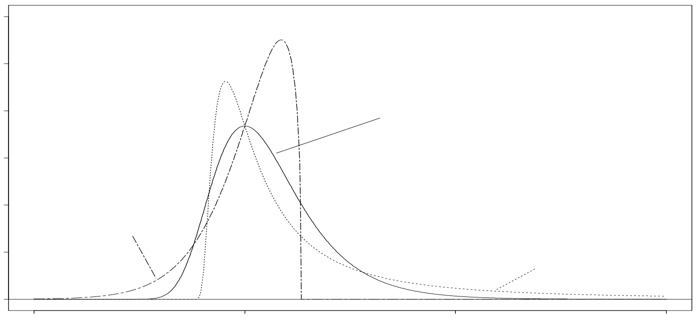

Figure 1: *Densities of the extreme value distributions.*

From Figure 1, we see the typical skew behaviour of extreme value distributions. Moreover, in the case $\xi > 0$ which is most important for finance, the tail $1 - H_{\xi,0,1}(x)$ behaves like $x^{-1/\xi}$, i.e. is fat–tailed. In order to be clear about the significance of the extreme value dfs and their link to the normal df, observe that for $X_1, \ldots, X_n$ iid, $N(\mu, \sigma^2)$, and $M_n = \max(X_1, \ldots, X_n)$,

$$P(M_n > x) \approx \Lambda\left(\frac{x - b_n}{a_n}\right) = H_{0,0,1}\left(\frac{x - b_n}{a_n}\right) \tag{2.5}$$

for suitable sequences (a_n) and (b_n) which can be calculated explicitly as functions of n, μ and σ. Hence the two–sided, skew Gumbel df Λ approximates the law governing the largest observation in a normal sample. A similar result, with different (a_n) and (b_n)'s holds for instance for exponential and lognormal data. In the case of fat–tailed data with $1 - F(x) \approx x^{-\alpha}$ say, the rhs in (2.5) has to be replaced by the Fréchet df $H_{1/\alpha,0,1}$. For details on these approximations, see Embrechts et al. (1997).

Returning to the crucial question of estimating beyond VaR (or shortfall), under reasonable conditions on F, there exists a canonical class of dfs approximating the conditional df in (2.3) for u_α large, i.e. α small. For finance and insurance applications, this is a crucial point. These *generalized Pareto distributions* (GPD) are defined as

$$G_{\xi,\nu,\beta}(x) = 1 - \left(1 + \xi\,\frac{x - \nu}{\beta}\right)_+^{-1/\xi} \tag{2.6}$$

where ξ is the shape parameter corresponding to the extreme value distribution, ν and β are again location and scale parameters. In the case $\xi > 0$ (the most important case for finance), the GPD has a heavy–tailed Pareto distribution. We have therefore reached

Fact 2: The GPD dfs (2.6) are natural approximations to the excess df (conditional VaR df) in (2.3).
Remark: Facts 1 and 2 are linked because the conditions needed in order to decide on the GPD fit depend on the approximation coming out of Fact 1.

We have sketched some of the main problems EVT can solve. A key task is now to work out the theory in such a way than an end–user can safely apply the methodology. The following is therefore of crucial importance for this end–user (risk manager, quant, actuary, ...).
Fact 3: The main tools and techniques from EVT have been worked out to be used on a large variety of data and models. Standard software is made available. A sample of S–plus routines can be downloaded from http://www.math.ethz.ch/m̃cneil/software.html.

The next section contains an example from finance.

3 An example

In order to illustrate some of the above techniques, we briefly discuss an example based on daily equity (BMW) return data over the period (January 2, 1973 till July 23, 1996). We concentrate on the left tail (i.e. negative daily return values). In order to keep in line with the positive sign for losses as used in the previous section, we take absolute values and denote the df of these values by F. The resulting series has $n = 2770$ observations. In Figure 2, we have plotted the empirical estimate $e_n(u)$, $u \geq 0$ of the mean excess function $e(u)$ in (2.4). This means that

$$e_n(u) = \frac{1}{N_u} \sum_{i=1}^{n} (X_i - u)_+$$

where $N_u = \# \{i = 1, \ldots, n \mid X_i > u\}$, i.e. N_u is the number of exceedances of u. Note the increasing behaviour from $u = 0.02$ onwards clearly indicating fat (even Pareto type) tails. In Figure 3, we have fitted the generalised Pareto distribution to the excess df in (2.3). All the data above the threshold $u = 0.02$ were used for this plot. The crucial shape parameter ξ has the value 0.223 which corresponds to a Pareto tail with value $1/\xi = 4.484$. From this plot, one can read off the conditional probability of high excesses, given that indeed we have an exceedance of $u = 0.02$. Of course, we can change the latter value as desired: for each u, a new model has to be fitted. On the other hand, one might be interested in the tail probabilities $1 - F(x)$ or high quantiles like in (2.2). A plot of this is given in Figure 4. Note that we use for both axes a log scale. The reason for this is that it magnifies the tail of F. An exact Pareto tail would be linear on this scale. On Figure 4, we have also plotted some quantiles u_α (see (2.2)) together with 95% confidence intervals. These intervals are the sections cut off by the two parabolas of the horizontal line through 95. The resulting values are:

$$\begin{aligned}
&\alpha = 0.01\,, \qquad u_\alpha = 0.040\,, \qquad \mathrm{CI}(95\%) = (0.038, 0.043)\,,\\
&\alpha = 0.001\,, \qquad u_\alpha = 0.081\,, \qquad \mathrm{CI}(95\%) = (0.070, 0.101)\,.
\end{aligned}$$

Hence a 95% confidence interval for a 0.1% event in these daily return data is (7.0%,10.1%). There is a lot more we could examine at this point: dependency in the data, sensitivity of estimates to the threshold $u = 0.02$, more details on the statistical tools used, ... All these, and indeed many more points can be addressed through careful use of EVT. We refer the interested reader to the references for more information and recommend experimenting with the software posted on the web.

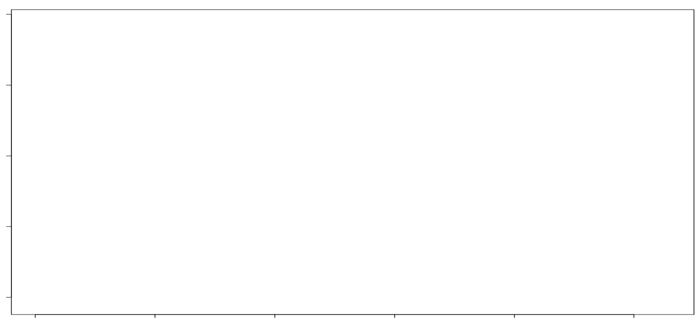

Figure 2: *Sample mean excess function $e_n(u)$ for the BMW return data.*

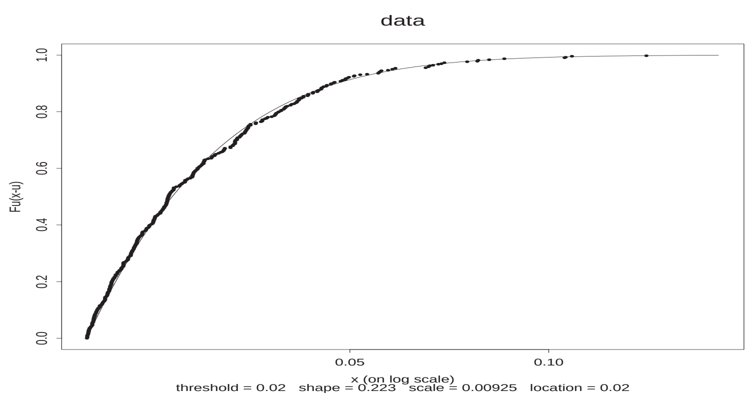

Figure 3: *Fit of the excess df F_u for the BMW return data.*

4 Final comments

The above discussion scratched the surface of what EVT offers. Its applicability to finance will be examined in much greater detail. The fact that the method already has proven important in fields like reliability, reinsurance, hydrology and environmental science enhances our belief that relevant applications in the realm of finance in general and RM in particular will be found. There is no alternative: if risk managers want to look at the edge, the proper

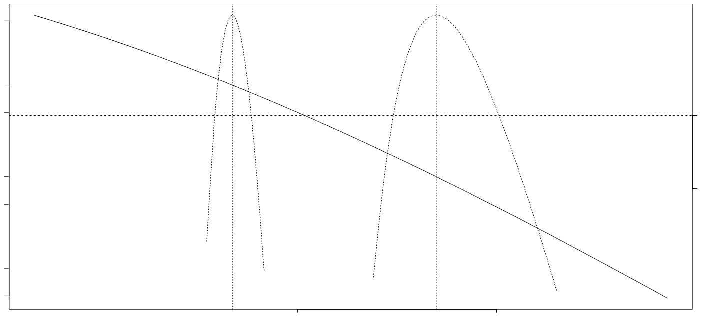

Figure 4: *Quantile estimates (99% and 99.9%) for the BMW return data together with 95% confidence intervals. Doubly logarithmic scale.*

tools must be used which implies reliance on classical EVT and its numerous extensions to dependent and multivariate data, as well as to stochastic processes (including some of the standard stochastic volatility models).

References

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1996) A characterization of measures of risk. Preprint. Université de Strasbourg.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1997) Coherent measures of risk. RISK, to appear.

Boudoukh, J., Richardson, M. and Whitelaw, R. (1995) Expect the worst. RISK, 8 (9), 100–101.

Brady, R. (1996) Value at Risk. RISK special supplement, June 1996, p. 32.

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

GARP (1996) Generally accepted risk principles. Coopers & Lybrand.

Jorion, P. (1997) *Value at Risk*. Irwin, Chicago.

Leadbetter, M., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, New York.

Longin, F.M. (1997) Beyond the VaR. Preprint. ESSEC, Paris.

Matthews, R. (1996) Far out forecasting. In: The New Scientist 2051, October 12, 1996, 36–40.

McNeil, A. (1997) Estimating the tails of loss severity distributions using Extreme Value Theory. ASTIN Bulletin, 27, 117–137.

McNew, L. (1996) So near, so VaR. RISK, 9 (10), 54–56.

Resnick, S. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer–Verlag, New York.

Resnick, S. (1997) Discussion of the Danish data on large fire insurance losses. ASTIN Bulletin, 27, 139–151.

Zangari, P. (1997) Catering for an event. RISK, 10(7), 34–36.