Revisiting the edge, ten years on

Valérie Chavez-Demoulin
EPF Lausanne, Switzerland

Paul Embrechts
ETH Zurich, Switzerland

October 14, 2008

Abstract

When these lines are written, it is January 21, 2008, a further "Black Monday" on the international markets. Stock indices have fallen between 5 and 10%. Which statistical tools help in describing such events and may help in understanding the consequences? In this paper we update our knowledge on the modelling of extremal events, in particular with a view towards applications to finance, insurance and risk management.

Keywords Aggregation, banking regulation, copula, dependence, extreme value theory, operational risk, quantile estimation, Value–at–Risk.

Mathematics Subject Classification 60G70, 62G32, 91B28

1 Looking back

Almost 10 years ago, Embrechts et al. [23] coined the phrase "Living on the edge" in order to stress the importance for risk managers within the banking industry to learn more about the relevant statistical tools for modelling extremal, i.e. “non–normal” events. The basic Extreme Value Theory (EVT) was summarized in Embrechts et al. [20]. The current subprime crisis, together with its consequences for international markets, shows that a deeper understanding of extreme events in statistical data from economics, insurance and finance is of high priority. The latter was already stressed in Embrechts [17]. In the present paper, we review some current EVT–based developments and exemplify these in the context of Quantitative Risk Management (QRM) where the latter is to be interpreted as presented in McNeil et al. [36]. The article is mainly written for a statistical audience with an interest in QRM–type of questions and less for the QRM specialist, though the latter, we hope, may learn from the material presented. Basic QRM terminology and notation is to be found in McNeil et al. [36] and will not be repeated here. An area of statistical research we will not touch upon is that of econometrics. Numerous scientific journals and books are solely devoted to that topic.

A main motivation for us to write this paper is our feeling that statistical research within finance is becoming more and more distant from some of the real issues out there in the world of risk management. As such, our contribution has a dual purpose: highlight some of the applied problems in risk management where EVT may play a role, and also point statisticians to some areas of considerable practical importance where more work is needed.

We start in the early 1990s. Let the random variable (rv) X denote the profit–and–loss position of the market risk of a bank. If we assume that $X \sim N(\mu, \sigma^2)$, then a commonly used (α–quantile) risk measure, Value–at–Risk at the level $\alpha \in (0,1)$, becomes

$$\mathrm{VaR}_\alpha(X) = \mu + \sigma\Phi^{-1}(\alpha) \tag{1}$$

where Φ denotes the distribution function (df) of a standard normal rv. The key input parameter in (1) is the so–called volatility parameter σ, to be estimated from data in function of the structure of the portfolio of assets underlying X. Of course, (1) is strongly based on the *assumption* that market returns are normally distributed, an assumption we know to hold only in very rare cases, if at all. It is at this point that EVT may enter, offering an alternative formula for α sufficiently close to 1, i.e. $F(u) < \alpha$:

$$\mathrm{VaR}_\alpha(X) = u + \frac{\beta}{\xi}\left(\left(\frac{1-\alpha}{\overline{F}(u)}\right)^{-\xi} - 1\right). \tag{2}$$

Here F is the df of X, $\overline{F} = 1-F$, u a threshold set by the modeller, and (ξ, β) are parameters to be estimated via the so–called Peaks–over–Threshold (POT) approach to EVT. Formula (2) is based on the much more realistic model assumption that large losses occur with considerable higher probability than predicted by the normal model underlying (1), i.e. it is assumed that F follows a power–tail (or Pareto–type) model

$$\overline{F}(x) = P(X > x) = x^{-1/\xi} L(x)\,, \tag{3}$$

where $0 < \xi < \infty$ and L is slowly varying in Karamata's sense:

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1\,, \quad t > 0\,. \tag{4}$$

Besides formula (2), EVT allows the quantitative risk manager to model the tail df $\overline{F}(x)$ for $x \geq u$, u sufficiently large, as

$$\overline{F}(x) \approx \overline{F}(u)\left(1 + \xi\frac{x-u}{\beta}\right)^{-1/\xi}. \tag{5}$$

Based on (5) and using for instance MLE, statistical estimates $\widehat{\overline{F}}(x)$ for $x \geq u$ can be obtained, inversion of (5) leads to formulas like (2). Furthermore, statistical uncertainty of these estimates can be quantified through the construction of confidence intervals using the profile likelihood approach. Other risk measures like Expected Shortfall

$$ES_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_s(X)\,ds$$

can easily be estimated in these models:

$$\widehat{ES}_\alpha(X) = \frac{\widehat{\mathrm{VaR}}_\alpha(X)}{1-\widehat{\xi}} + \frac{\widehat{\beta} - \widehat{\xi}u}{1-\widehat{\xi}}\,.$$

This was the situation when Embrechts et al. [23] was written. In Figure 1 we give an example including all the above, by now standard EVT ingredients. The upper panel contains the

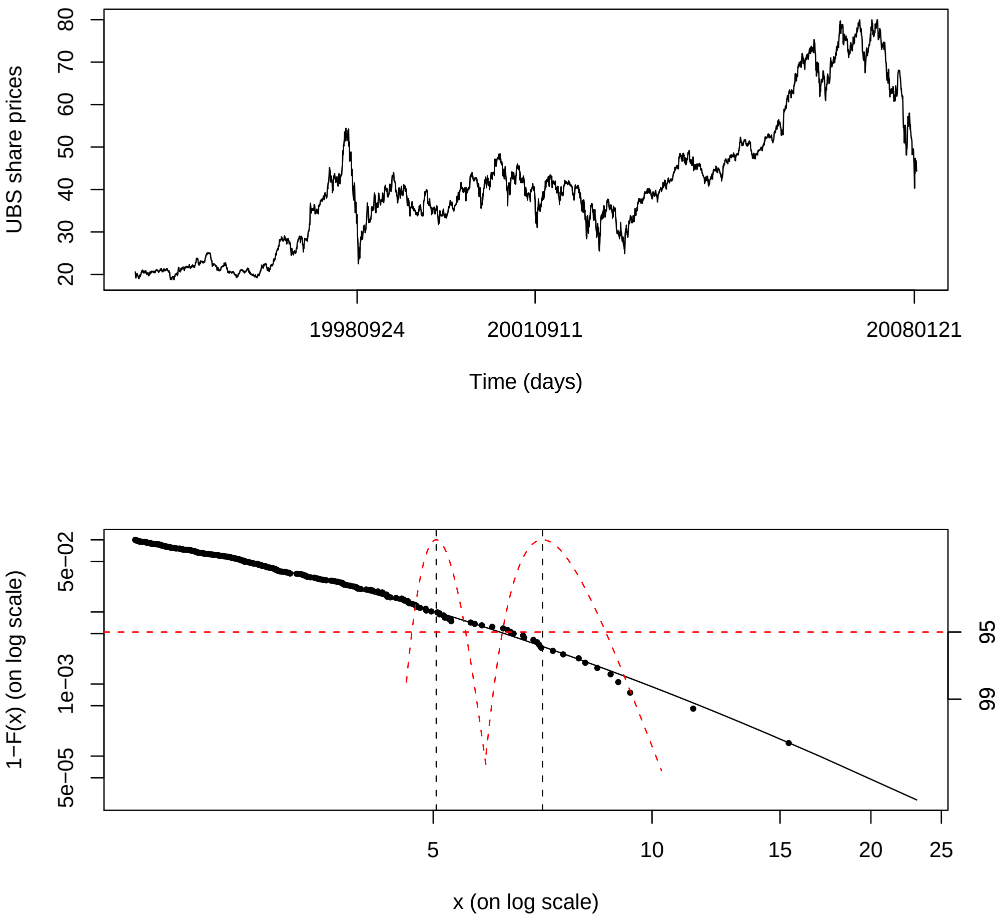

Figure 1: *UBS data for the period 1995.01.03 to 2008.02.01: daily share price (upper panel), loss returns in % together with a POT-based fit (lower panel). The vertical dotted lines show the estimates of* $\mathrm{VaR}_{99\%}$ *and* $\mathrm{ES}_{99\%}$. *The other curves are used to construct confidence intervals (horizontal line at 95%). The estimated* $\mathrm{VaR}_{99\%}$ *is 5.04 with 95% confidence interval* $[4.67; 5.54]$. *The estimated* $\mathrm{ES}_{99\%}$ *is 7.07 with 95% confidence interval* $[6.24; 8.62]$.

plot of the daily share price of UBS for the period 1995.01.03 to 2008.02.01. One clearly sees the drops in share price around the LTCM crisis in September 1998 and the more recent subprime crisis in 2008. The lower panel contains the loss returns (in %) of the data together with a POT-based fit of the tail above the threshold $u = 1.94\%$; see (5). In the analysis we assumed the return data to be iid and hence neglected stochastic volatility effects; see McNeil et al. [36] for details, and Mancini and Trojani [32] for a robust version of the former. The results obtained are $\mathrm{VaR}_{99\%} = 5.04[4.67; 5.54]$ and the $\mathrm{ES}_{99\%} = 7.07[6.24; 8.62]$. The numbers in brackets yield 95% confidence limits. If applied at the level of a portfolio, the

$\text{VaR}_{99\%}$ value, after scaling from 1–day returns to 10–day returns, would enter the capital charge formula for market risk of the bank; see McNeil et al. [36], formula (2.21). As a risk measure, Value–at–Risk enters fundamentally in the Regulatory Guidelines of the Basel Committee on Banking Supervision; see the various documents on the Committee's website www.bis.org/list/bcbs/index.htm. On the contrary, insurance regulators favour Expected Shortfall as this measure not only yields loss information *beyond* VaR, but also satisfies the fundamental axiom of *subadditivity* as championed by Artzner et al. [2]. For a related discussion on VaR and EVT, see Embrechts et al. [23]. SCOR [43] contains an industry view on the Solvency 2 guidelines for insurance within the so–called Swiss Solvency Test.

2 A brief discussion of some EVT caveats

Whereas from a theoretical point of view the use of EVT within QRM is compelling, one could ask the question why EVT has not yet advanced as *the* tool for rare event (high–quantile, say) estimation within QRM. A first reason is that in (1) the quantities μ (mean return), and more importantly σ (the volatility) have a deeply entrenched meaning for traders and risk managers. For instance, volatility is well–understood and can even be traded. On the other hand, though models based on (3) and hence (2) are statistically more realistic, the parameters (ξ, β) are much more difficult to interpret from a practitioner's point of view. The important shape parameter ξ parameterises the power–like tail behaviour. Typically $\xi \in (0.2, 0.5)$ for market risk data, however values $\xi \in (0.5, 1.5)$ are not uncommon for operational risk data. Note that values $\xi > 1$ in (3) correspond to *infinite mean* models. This leads to particular statistical estimation problems (see for instance Nešlehová et al. [38]) and to superadditivity of VaR, as discussed in Embrechts et al. [23]. Finally β is just a scale parameter. Though many models are in use which somehow "interpolate" between (1) and (2), industry often does resist new methodology, like EVT, which typically leads to higher capital charges. In the wake of the subprime crisis, this attitude will be more difficult to defend.

The question most often asked by applied risk managers is no doubt "how to choose the threshold u". This gets to the core of EVT for which, it is our firm belief, no easy, automatic push-of-the-button answer can be given; see Chapter 7 in McNeil et al. [36] for a first discussion. A related issue, having attracted much less attention, concerns the rate of convergence of EVT based estimators. Theoretically (as $u \to \infty$, say) the rate of convergence underlying asymptotic formulas like (2) can be arbitrarily slow and crucially depends on the second–order properties of the slowly varying function L in (3). For operational risk data (as defined in Section 3 below), this point is very much highlighted in Degen et al. [13] and Degen and Embrechts [12]. In these papers, various applied consequences for the calculation of regulatory (or risk) capital are given. Just to give an idea, from Degen et al. [13] we learn that, in terms of the threshold $u(\to \infty)$, for the loggamma df the rate is $O(1/\log u)$ whereas for the popular g–and–h df with $g, h > 0$, the rate becomes the extremely slow $O(1/\sqrt{\log u})$. Recall that a rv X is g–and–h distributed, $g, h \in \mathbb{R}$, if for some constants $a, b \in \mathbb{R}$, and a rv $Z \sim N(0, 1)$,

$$X = a + b\frac{e^{gZ} - 1}{g} e^{\frac{1}{2}hZ^2}.$$

For a student t df, the rate is $O\left(1/u^2\right)$. In the case of a t, the slowly varying function L in (2) is asymptotically a constant, whereas for the g–and–h case with $g > 0$, $h(= \xi) > 0$, $L(x)$ is of the order $\exp\left(\sqrt{\log x}\right) / \sqrt{\log x}$. The rate of convergence for EVT–based estimation stands in stark contrast to the uniform $n^{-1/2}$ rate of convergence in the Central Limit Theorem for finite variance dfs, or $n^{-1/\alpha}$ in the α–stable case, $0 < \alpha < 2$.

The QRM literature contains numerous discussions on the use of EVT; the views range from very positive (see Moscadelli [37] for instance) to somewhat negative (Diebold et al. [15] and Dutta and Perry [16]), with some in between opinions (as in Jobst [28]). A key point however that often gets lost in such discussions is that using EVT, one not only bases estimation on a mathematically well understood theory (this is typically acknowledged) but that, for instance using the POT-method, one models conditional tail dfs above increasingly high thresholds by dfs (the Generalized Pareto dfs, GPDs) which satisfy an important *stability* property, i.e. the GPD–class is closed under conditioning over successive higher thresholds; see McNeil et al. [36], Lemma 7.22. When one uses a particular parametric class of dfs like the g–and–h, say, one may obtain a close fit over a restricted range of data values, but has no structural assumptions on the underlying data. We feel that more attention should be given to the latter; this will not be easy as economics in general, and finance more in particular to a high degree defy the existence of "fundamental laws" as for instance exist in physics. An excellent text stressing the importance of structural properties of dfs when fitting one-dimensional loss data is Marshall and Olkin [33].

Finally, regulators have to be aware of the statistical difficulties present in any estimation of high risk measures like $\mathrm{VaR}_\alpha(X)$ with $\alpha \in \{99, 99.9, 99.97\}$, say. The latter two confidence levels correspond to the calculation of risk capital for credit and operational risk, and economic capital, respectively (all on a yearly basis). The recent turmoils in the international banking system have already led to the introduction of a so–called Incremental Risk Charge (IRC) at the 99.9%, 1–year VaR level for market risk. It aims to cover the increasing amount of exposure in banks' trading books to credit risk related and often illiquid products whose risk is not reflected in the standard 1–to 10–day 99% VaR numbers. The IRC addresses the important cross border issue between market and credit risk so prominently exposed in the subprime crisis; see Basel Committee [6] for details. The statistical estimation of IRC and its backtesting is widely open.

Estimating a 1 in 1000 year event in any applied area is a daunting task; for credit and operational risk even more so because of the scarcity and often bad quality of the data. We do however prefer a careful use of EVT technology above the gormless guessing of some parametric model that may fit currently available data over a restricted range where no or only very few extreme observations are available. The fact that estimates based on EVT often yield unrealistically high estimates of risk capital is more a reflection on the unrealistic demand to base that capital on a VaR_α risk measure with $\alpha \geq 0.999$, say. As explained in Section 1, at least EVT yields (typically very wide) confidence intervals for such extreme risk measures. This often unsettles management of institutions who need to take business decisions partly based on these numbers. The only conclusion is that, beyond a certain level of confidence (the α above), a qualitative assessment has to kick in; one cannot solely rely on quantitative support. Once more, the subprime crisis yields a clear proof of this; investment–

quality ratings of CDO tranches are typically based on (very rare) default probabilities, in a range below 0.05%/year. Note here that the path of thinking went the other way: first model–based rare default probabilities for mortgage backed bonds are correlated and then, within the CDO structure, are mapped onto rating classes like a AAA from Standard & Poor's. These high (quality) ratings were then used to inform clients about the safety of their investments in products like CDOs based on subprime housing loans. History (by now) has taught the world of investment banking a very different lesson; hopefully, classical EVT helps some of these investment product factories to be a bit more humble when it comes to making statistical statements on very rare events and their interdependence. We briefly return to this issue in the next section.

3 Beyond classical EVT and a QRM example

The main methodological constraints on EVT applications, as discussed in Embrechts et al. [23], as well as in most publications on the subject, are the non–dynamic character of most of the models and the restriction to one dimension, $d = 1$. Of course, numerous theoretical developments have taken EVT from $d = 1$ to the multivariate case (MEVT, $d \geq 2$), as well as dynamic EVT models; the latter both considering extremes of stochastic processes, and (M)EVT where the model parameters are time dependent. Several of the recent textbooks treat MEVT; see for instance de Haan and Ferreira [14], Resnick [40] and Balkema and Embrechts [4]. For extremes of stochastic processes Berman [8] is a good start; see Hult and Lindskog [27] for more recent work. Chavez–Demoulin and Embrechts [10] contains a nice example within QRM for operational risk using time–dependent EVT parameter fitting. It is however our firm belief that the impact of all these developments on day–to–day QRM has been minimal so far. And yet, there are exciting statistical challenges out there, some of which we highlight below.

Consider the case of operational risk (OR): "Operational risk is defined as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition includes legal risk, but excludes strategic and reputational risk". For a large international bank, the yearly, total OR loss rv has the representation

$$L = \sum_{i=1}^{b} \sum_{j=1}^{r} \sum_{k=1}^{N_{ij}} X_k^{ij} I_{\left\{X_k^{ij} > c_{ij}\right\}} \,. \tag{6}$$

Here b stands for the number of relevant business lines, r for the number of risk types and X_k^{ij} is the k–th loss over the past year for business line i (BL_i) and loss type j, N_{ij} being the total (random) number of such losses. The constants c_{ij} correspond to a minimal value below which no losses are reported. Standard values within the Basel II guidelines are $b = 8$, $r = 7$ and $c_{ij} = 10\,000$ US\$ for all i, j. The key statistical task, imposed by Basel II, now concerns the estimation of

$$F_L^{-1}(0.999) = \mathrm{VaR}_{99.9\%}^{\mathrm{OR}} \,. \tag{7}$$

As already mentioned above, this is a daunting task! In Figure 2 we plot some operational risk data; the data are aggregated at the level of three risk types. We also have included the mean excess plots as discussed in McNeil et al. [36], p. 279, clearly indicating the heavy–tailedness (power–tail behaviour) of the data. For a detailed analysis of these data using EVT, see Chavez–Demoulin and Embrechts [10]. Some of the modelling issues related to (6) and (7) are:

- (MI1) Top–down (F_L) modelling versus a bottom–up $\left(F_{X_k^{ij}}, F_{N_{ij}}, i, j, k\right)$ approach.
- (MI2) (Often used in practice) Aggregate data business line wise yielding estimates $\widehat{\mathrm{VaR}}\,(\mathrm{BL}_1)\,, \ldots, \widehat{\mathrm{VaR}}\,(\mathrm{BL}_b)$. Add up to obtain $\widehat{\mathrm{VaR}}(\mathrm{Total}) = \sum_{i=1}^{b} \widehat{\mathrm{VaR}}\,(\mathrm{BL}_i)$ and use diversification arguments to arrive at the reported regulatory capital value of OR: $\mathrm{VaR}(\mathrm{Reported}) = (1-\delta)\widehat{\mathrm{VaR}}(\mathrm{Total})$ where in practice δ–values may be in the range of 10–30%.
- (MI3) Compare the influence on VaR(Reported) for different aggregation schemes.
- (MI4) Estimate e.g. $\mathrm{VaR}_\alpha\,(\mathrm{BL}_i)$ at a level $\alpha < 99.9\%$ ($\alpha = 90\%$, say) and scale upwards.

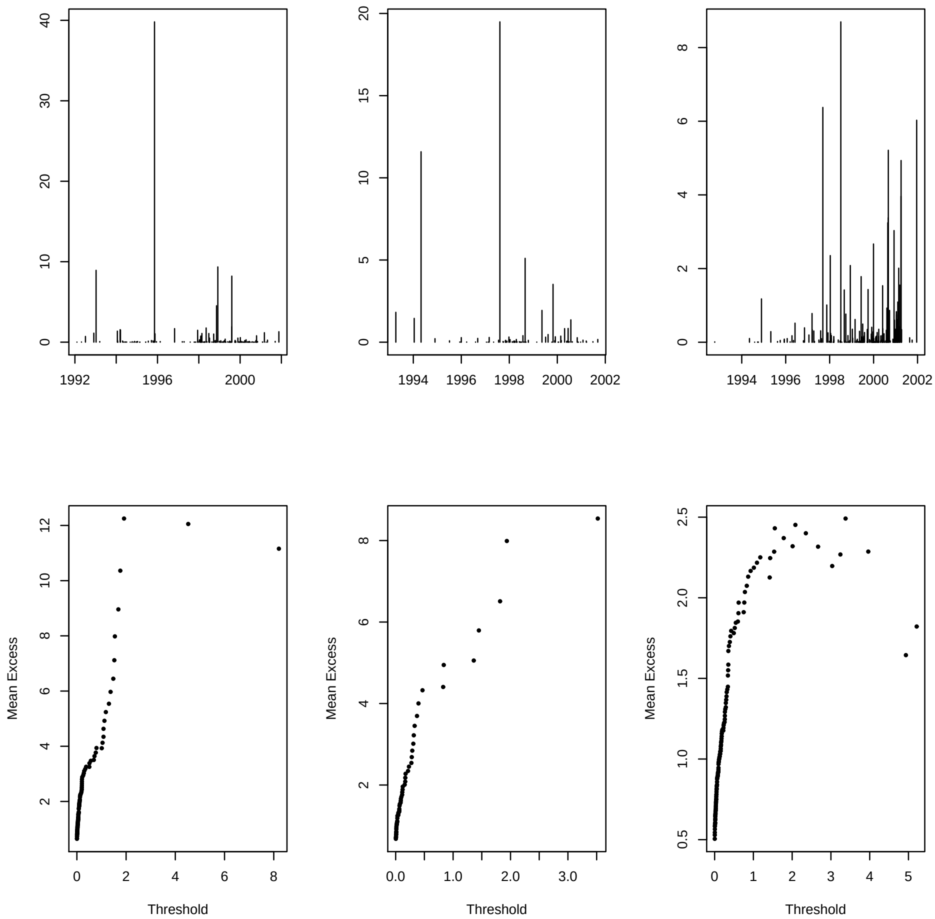

Figure 2: *Operational Risk data (top) together with their corresponding mean excess plots (bottom).*

(MI5) According to Basel II, the final VaR–based capital charge has to be based on a combination of internal, external and expert opinion data. This immediately leads to interesting statistical questions:

Q1: how to calibrate individual institution data to industry–wide data;

Q2: how to quantify expert opinions (Bayesian methodology, inter–rater agreement, ...), and

Q3: data–type combination (as for instance done in credibility theory; see Bühlmann and Gisler [9] for the general theory and Lambrigger et al. [30] for an application to OR.).

These are just some of the, for practice important issues and scientific guidance on how to solve the underlying statistical problems are badly needed. For instance, what could be the difference when in (MI2) and (MI3) the data were first aggregated risk type wise. Under the current regulatory guidelines banks would be free to choose. Clearly the answer depends on the dependence structure of the underlying ($b \times r$–)matrix of loss data; a first attempt to solve this problem is to be found in Embrechts and Puccetti [22]. An important issue underlying the δ–diversification in (MI2) is the already mentioned subadditivity property for VaR, i.e.

$$\mathrm{VaR}_{\alpha}\left(L_1 + L_2\right) \leq \mathrm{VaR}_{\alpha}\left(L_1\right) + \mathrm{VaR}_{\alpha}\left(L_2\right) \,, \tag{8}$$

which one hopes to hold for loss random vectors (L_1, L_2). See Embrechts et al. [21] for a discussion of (8).

In February 2009, the US Office of the Comptroller of the Currency (OCC) and the National Institute of Statistical Sciences (NISS) start a two–year joint research effort with a conference on "Exploring Statistical Issues in Financial Risk Modeling and Banking Regulation" clearly stressing the need for more statistical research within QRM. The statistical modelling of operational risk data will be one of the key research areas to focus on; see NISS's website www.niss.org for an update on this effort.

A research topic on which very little is known so far concerns the joint modelling of the regulatory risk classes market (MR), credit (CR) and operational (OR) risk. The recent introduction of the IRC touched upon in Section 2 very much highlights the importance of this. In particular, how can one make sense of a global VaR–based risk measure through

$$\mathrm{VaR}_{\mathrm{Total}} = \mathrm{VaR}^{\mathrm{MR}}_{\alpha_1,T_1} + \mathrm{VaR}^{\mathrm{CR}}_{\alpha_2,T_2} + \mathrm{VaR}^{\mathrm{OR}}_{\alpha_3,T_3} \quad (+\ ?) \tag{9}$$

where the (+ ?) is included to indicate that addition only makes sense if we have rescaled $(\alpha_1, \alpha_2, \alpha_3)$ and (T_1, T_2, T_3) to a common (α, T). For a discussion on time scaling within QRM, see for instance Embrechts et al. [19] and the references therein. Some early attempts for solving the risk aggregation problem (9) for banks are Aas et al. [1] and Rosenberg and Schuermann [41].

One of the key tools used in practice in discussions of the above type ((9) and the determination of δ in (MI2)) is linear correlation. This brings us to another important aspect of the modelling of extremes in a context very relevant for QRM: the analysis of high risk scenarios or stress testing. Besides the standard VaR–based analyses discussed above, banks (and insurance companies) have to stress test their portfolios by sufficiently extreme scenarios. The development of probabilistic and statistical theory necessary for the formulation of such scenarios is at this moment of crisis of very high priority. The following quote from Myron Scholes in Scholes [42] already highlighted this point in the year 2000, and this in the wake of the LTCM crisis: "Now is the time to encourage the BIS and other regulatory bodies to support studies on stress testing and concentration methodologies. Planning for crises is more important than VaR analysis." The subprime crisis and the resulting systemic risk for the global financial system stresses the need for this to the extreme; see for instance the section on "Stress testing process" in Basel Committee [7].

For the development of such crisis tools, EVT–based statistical methodology will have to play an important role. In the next section we briefly discuss one of such techniques, the copula concept. It plays a crucial, though not always positive role in the turmoils around the current subprime crisis. We decided to include a short section on copulas as we feel that this concept also has been, and still is, widely misused within the financial community.

4 Copulas and stress testing

Though extreme events for financial and insurance data manifest themselves through high gains and/or losses, very often it is the *comovement* of underlying instruments in times of crises that triggers such events; the so–called *perfect storm scenario*. Therefore, the modelling of the joint occurrence of extremes within QRM ought to be of prime concern. A tool which over the last ten years has gained enormous interest in this context is that of *copula*. The concept is a rather trivial one; in our view, for QRM, its virtues are mainly pedagogical with some usefulness towards the development of stress scenarios. Researchers as well as practitioners not only ought to be aware of the concept's potential, but increasingly so of its limitations.

Suppose $L_1, \ldots, L_d$ are rvs with marginal dfs $F_1, \ldots, F_d$. Then for any df C on the unit hypercube $[0,1]^d$ with standard uniform marginals,

$$F_{\mathbf{L}}(\boldsymbol{\ell}) = C\left(F_1\left(\ell_1\right), \ldots, F_d\left(\ell_d\right)\right) \tag{10}$$

is a df with marginal dfs $F_1, \ldots, F_d$. The copula C yields *a* dependence scenario between the rvs $L_1, \ldots, L_d$. The copula concept was introduced in the 1950s but gained wider prominence through the world of finance only in the late 1990s, and this mainly through QRM–type questions posed by industry. For a historical review of the latter and a discussion on the pros and cons of copula–based modelling in QRM, see Embrechts [18]. A bibliographical review highlighting applications to finance is Genest et al. [25]. From the latter paper we learn that more than 40% of the papers written on the subject belong to the field of finance, and this despite the fact that copula applications to finance only appeared after 2000 in the academic literature. A Google search (copula, risk) on October 9, 2008 lists about 140000 entries whereas (copula, finance) gets 111000! So where does the copula concept actually enter into the extended (M)EVT scheme?

First of all, classical QRM is strongly concerned with (linear) correlation. How may a change in dependence, measured through linear correlation, imply changes in the riskiness of a financial position? Such changes can be dramatic. McNeil et al. [36], Figure 8.1, discusses an example of a bond portfolio in which a substantial change in the default loss at portfolio level occurs when the assumption of uncorrelated loans ($\rho = 0$) is replaced by dependent loans with a low ($\rho = 0.005$) default correlation. Changing default correlation only mildly may have a considerable effect on the default risk at portfolio level. This is particularly worrying as standard statistical estimators for ρ are non–robust. As a consequence (default) correlation estimation (also at matrix level) is of great importance. A further worrying fact comes from the copula world and concerns the rating of the underlying credit risk: marginal dfs (of firm value, say) and (default) correlation do not uniquely determine a stochastic default model. An example should make the consequences clear; we use the set up of the standard Merton–KMV model as discussed in McNeil et al. [36], Section 8.2.3, in which the firm values of d companies are modelled through a d–dimensional geometric Brownian motion. As a rather immediate consequence, default calculations at year–end can be based on a multivariate normal distribution with correlation matrix $\mathcal{P}$ which, as an example, we assume of the equi–correlation type, $\rho_i = \rho = 0.038$. The individual default probabilities are to be 0.005 (medium credit quality). By way of example, both a Gaussian copula as well as t–copula model are calibrated, with 10 degrees of freedom for the latter, say. Both models are fully in line with the given model constraints (given marginal dfs and correlation). One then shows that, at the 95% VaR level, the t_{10}–copula model more than doubles the required risk capital as compared to the Gauss–copula model. Formulated differently, the rating of the credit model very much depends on the dependence model (copula) used. Note that this difference cannot be explained by marginal loss dfs and linear correlation. This example nicely shows that copula modelling may be useful for the construction of stress scenarios for dependence. The question however "which is the right copula to use" is as difficult as "which is the right joint model to use"; this point over and over needs stressing when talking to practitioners. Embrechts [18] has more on this.

Copulas allow examples like the above to be taken one step further: within QRM it is by now very well established that (linear) correlation only allows for a rather one–sided view on dependence. The subprime crisis made this point very clear. At the danger of gross oversimplification, it is fair to say that within the overall chain of events involving excessive brokerage of subprime loans, new CDO product constructions through investment banks, and finally over–optimistic credit ratings, mathematics had a small, but crucial role to play. The industry models used to price CDOs were too simplistic with respect to the dynamic aspects of widening credit spreads in periods of crisis, as well as too optimistic when it came to predicting joint extremes. It is here that MEVT, in the guise of copulas, entered. A further example will make this clear. One of the standard tools used for the pricing of structured products like CDOs is the so-called Li–model; see Li [31]. The basic idea is to model time–to–default by exponential rvs and introduce dependence through the copula construction (10). In Figure 3 we have given two stylised examples with equal marginal dfs, EXP(1), and linear correlation 50%. In the left panel we have put a Gaussian copula on these marginals,

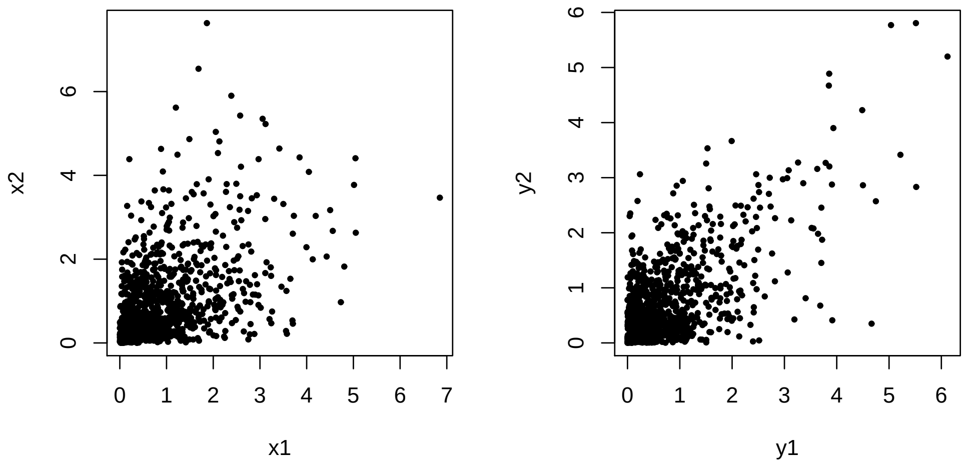

Figure 3: *Bivariate unit-exponential survival data with equal linear correlation* $\rho = 0.5$ *and different copulas: Gaussian copula left and Gumbel–copula right calibrated so as to achieve 50% correlation. For the definition of the Gumbel–copula, see McNeil et al. [36], p. 192.*

in the right one a Gumbel–copula with the parameters chosen so as to calibrate to the 50% correlation. Note that the Gaussian copula case contains fewer joint extreme events than the Gumbel one, an observation that can easily be proved analytically; see McNeil et al. [36], Section 5.2.3. The main point is that using Gauss–copula technology for the pricing of CDOs will typically underestimate the number of joint defaults and as a consequence may lead to a higher credit rating. An early warning on this, with considerable practical consequences, was reported by M. Whitehouse in a front page article in *The Wall Street Journal* of September 12, 2005 with the telling title: “How a formula ignited market that burned some big business.” The “formula” referred to was the above Li–model. Few realised then that similar formulas were ticking away like time bombs in some of the structured products now exploding. And yet, when used properly, copulas yield an excellent stress testing tool. We were told by a risk manager of a larger insurance company how he declined a request for an investment of his company in some CDO tranches; his risk assessment was based on a simple copula stress testing of the type exemplified in Figure 3. I.e., using some straightforward copula modelling, he questioned the AAA–rating quality of some CDO tranches built on subprime loans.

5 Some statistical challenges related to MEVT

In Section 3, we very briefly mentioned some generalisations of classical, one–dimensional EVT. For applications to finance (and indeed beyond) there is a need for a truly workable multivariate theory. Though there are several approaches to MEVT available in the literature, more work is needed on the statistical theory in arbitrary dimensions.

Suppose $\mathbf{X} = (X_1, \ldots, X_d)'$ is a vector of rvs representing the value of d financial instruments, where the notion of instrument is to be interpreted broadly. For expository purposes, assume $\mathbf{X}$ has a density $f_{\mathbf{X}}$ on $\mathbb{R}^d$. The statistical estimation of rare events based on $f_{\mathbf{X}}$ is of considerable importance within QRM. The notion "rare" will become clear in the examples below.

(EX1) Suppose $A \in \mathcal{B}\left(\mathbb{R}^d\right)$ such that $P(\mathbf{X} \in A)$ is sufficiently small, estimate

$$P(\mathbf{X} \in A) = \int_A f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}\,. \tag{11}$$

(EX2) Suppose $\Psi : \mathbb{R}^d \to \mathbb{R}$, measurable, $B \in \mathcal{B}(\mathbb{R})$ rare, and $A \in \mathcal{B}\left(\mathbb{R}^d\right)$ general, estimate

$$P\left(\mathbf{X} \in A \mid \Psi(\mathbf{X}) \in B\right)\,. \tag{12}$$

(EX3) Suppose $I_1, I_2, J_1, J_2 \subset \{1, 2, \ldots, d\}$ with $I = I_1 \cup I_2$ and denote the subvectors $\mathbf{X}_1$ and $\mathbf{X}_2$, making up $\mathbf{X}$, in an obvious way. Let $\Psi_1 : \mathbb{R}^{I_1} \to \mathbb{R}^{J_1}$, $\Psi_2 : \mathbb{R}^{I_2} \to \mathbb{R}^{J_2}$ be measurable and $A_i \in \mathcal{B}\left(\mathbb{R}^{J_i}\right)$, $i = 1, 2$, be rare, i.e. $P\left(\Psi_i\left(\mathbf{X}_i\right) \in A_i\right)$, is small, $i = 1, 2$. Estimate the so–called *spillover* or *contagion* probabilities

$$P\left(\Psi_1\left(\mathbf{X}_1\right) \in A_1 \mid \Psi_2\left(\mathbf{X}_2\right) \in A_2\right)\,. \tag{13}$$

Practical questions of the above type abound in QRM. For instance, if $d = 2$ and in (13) "rare" is interpreted as "exceeding VaR_α" for α close to 1, we obtain the spillover probability:

$$P\left(X_1 > \mathrm{VaR}_\alpha\left(X_1\right) \mid X_2 > \mathrm{VaR}_\alpha\left(X_2\right)\right)$$

which can be estimated based on standard componentwise MEVT as discussed in Coles [11]. For a more recent reference with QRM applications in mind, see Ramos and Ledford [39]. We do not know of good statistical techniques for handling the estimation of (13) for $d \gg 2$. Questions of this type are relevant for analysing systemic risk; hence the "spillover" of risk/crisis from one market (or bank) to another.

In Balkema and Embrechts [4], a multivariate generalisation of the POT–method is given motivated by questions of the type (12) coming from portfolio theory. In (12), take $\Psi(\mathbf{x}) = \mathbf{w}'\mathbf{x}$ for some portfolio weights $\mathbf{w} \in \mathbb{R}^d$. As rare event we take $B = (q_\alpha, \infty)$ for some high quantile q_α (VaR_α, say) and denote $H_\alpha = \left\{\mathbf{x} \in \mathbb{R}^d : \mathbf{w}'\mathbf{x} > q_\alpha\right\}$, the corresponding remote half–space in $\mathbb{R}^d$. The above reference contains an asymptotic theory for the affinely scaled vector $\mathbf{X}$:

$$\beta_{H_\alpha}^{-1}(\mathbf{X}) \xrightarrow{P_{H_\alpha}} \mathbf{Z}\,, \quad \alpha \uparrow 1\,, \quad \mathbf{Z} \text{ non–degenerate}\,, \tag{14}$$

where $P_{H_\alpha}(\cdot) = P\left(\cdot \mid \mathbf{X} \in H_\alpha\right)$ and (β_{H_α}) is a family of affine transformation. In the isotropic case, i.e. the limit in (14) exists for all $\mathbf{w} \in \mathbb{R}^d \backslash \{\mathbf{0}\}$, the limit random vectors $\mathbf{Z}$ are characterised and their (maximum) domains of attraction determined. The statistical theory based on (14) needs to be developed and numerous concrete examples worked out. For some first statistical results in this direction, see Fougères and Soulier [24]. Balkema et al. [5] contains an application of the theory based on (14) to some of the credit models discussed in Section 4.

Questions of the type (11) have received considerable attention in the statistics as well as in the finance literature. They occur under the names of *rare event simulation* (Asmussen and Glynn [3]) and *importance sampling* (Glasserman [26] and McLeish [34]). A nice paper applying importance sampling and EVT to questions of the type (12) is McLeish [35].

6 Conclusion

Statisticians with an interest in finance and insurance in general and QRM more in particular have to become aware of the broader picture. Statistics offers a powerful and key theory for use in QRM. Indeed in very few fields of application do we have statistical quantities hard-wired in the law, like VaR–based capital requirements within the Basel II regulatory framework for large international banks; see also Jorion [29]. In the coming years it will be crucial to see whether statisticians take up the challenge and become concerned with some of the real issues underlying QRM going forward. We do hope that, from the point of view of the modelling of extremes, our paper will have contributed a bit towards this potentially fruitful discussion. Examples like the OCC–NISS project mentioned in Section 3 give some hope. Going back to Embrechts et al. [23], statistics has outgrown "playing on the edge" and has potentially entered the much more demanding task of "mastering living on the edge".

Acknowledgment. The authors would like to thank the referees for some critical comments on an earlier version of the paper.

Note added in proof. Since the first version of this paper, markets worldwide have weakened considerably. Pure investment banks have ceased to exist, several financial institutions have either defaulted or needed saving. Even countries are under risk of default. Systemic risk has spread from Wall Street to Main Street so much so that the cover of *The Economist* of October 4th–10th 2008 read: "World on the edge". Fear of a worldwide recession is real. Internationally, countries work out plans to restabilise the economy. More than ever, scientists in general, and statisticians in particular are called upon to help restoring the quantitative, regulatory foundations of the financial system and develop methods and techniques which help society at large to navigate more safely close to the occasional economic edge.

References

[1] Aas, K., Dimiakos, X.K., Øksendal, A. (2007). Risk capital aggregation. *Risk Management* 9, 82–107.

[2] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D. (1999). Coherent Measures of Risk. *Math. Finance* 9, 203-228.

[3] Asmussen, S., Glynn, P.W. (2007). *Stochastic Simulation: Algorithms and Analysis.* Berlin: Springer.

[4] Balkema, G., Embrechts, P. (2007). *High Risk Scenarios and Extremes: A Geometric Approach.* Zurich Lectures in Advanced Mathematics. Zurich: European Mathematical Society Publishing House.

[5] Balkema, G., Embrechts, P., Lysenko, N. (2008). Meta densities and the shape of their sample clouds. ETH Zurich. Preprint.

[6] Basel Committee (2008a). Guidelines for computing capital for incremental risk in the trading book. BCBS Consultative Document, July 2008.

[7] Basel Committee (2008b). Principles for sound liquidity risk management and supervision. BCBS, September 2008.

[8] Berman, S.M. (1992). *Sojourns and Extremes of Stochastic Processes.* Pacific Grove, Ca: Wadsworth–Brooks/Cole.

[9] Bühlmann, H., Gisler, A. (2005). *A Course in Credibility Theory and Its Applications.*. Berlin: Springer.

[10] Chavez–Demoulin, V., Embrechts, P. (2004). Smooth extremal models in finance and insurance. *J. Risk and Insurance* 71, 183–199.

[11] Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values.* London: Springer.

[12] Degen, M., Embrechts, P. (2008). EVT-based estimation of risk capital and convergence of high quantiles. *Adv. Appl. Probab.* 40(3), 696–715.

[13] Degen, M., Embrechts, P., Lambrigger, D.D. (2007). The quantitative modelling of operational risk: between g–and–h and EVT. *ASTIN Bulletin* 37, 265–291.

[14] de Haan, L., Ferreira, A. (2006). *Extreme Value Theory. An Introduction.* New York: Springer.

[15] Diebold, F.X., Schuermann, T. Stroughair, J. (1998). Pitfalls and opportunities in the use of extreme value theory in risk management. In A.–P. N. Refenes, J.D. Moody and A.N. Burgess (eds.) *Advances in Computational Finance*, 3–12. Reprinted in *J. Risk Finance*, 1 (Winter 2000), 30–36.

[16] Dutta, K., Perry, J. (2006). A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital. Federal Reserve Bank of Boston, Working Paper No. 6-13.

[17] Embrechts, P. (2004). Extremes in economics and the economics of extremes. In *Extreme Values in Finance, Telecommunications, and the Environment*, (Eds. Finkenstädt and Rootzén). Boca Raton, FL: Chapman & Hall/CRC, pp. 169–183.

[18] Embrechts, P. (2009). Copulas: A personal view. *J. Risk and Insurance.* To appear.

[19] Embrechts, P., Kaufmann, R., Patie, P. (2005). Strategic long–term financial risks: single risk factors. *Computational Optimization and Applications* 32(1/2), 61–90.

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

[21] Embrechts, P., Lambrigger, D.D., Wüthrich, M.V. (2008). Multivariate extremes and the aggregation of dependent risks: examples and counter-examples. *Extremes.* To appear.

[22] Embrechts, P., Puccetti, G. (2008). Aggregating risk across matrix structured loss data: the case of operational risk. *J. Operational Risk* 3(2), 29–44.

[23] Embrechts, P., Resnick, S., Samorodnitsky, G. (1998). Living on the edge. *RISK*, January 1998, 96-100. Also published in: *Hedging with Trees: Advances in Pricing and Risk Managing Derivatives*, M. Broadie and P. Glasserman (eds.). New York: Risk Books, pp. 239-243.

[24] Fougères, A.–L., Soulier, P. (2008). Modeling and estimation of conditional excesses. Université Paris, Ouest–Nanterre. Preprint.

[25] Genest, C., Gendron, M., Baudeau–Brien, M. (2008). The advent of copulas in finance. *European J. Finance.* To appear.

[26] Glasserman, P. (2004), *Monte Carlo Methods in Financial Engineering.* New York: Springer.

[27] Hult, H., Lindskog, F. (2005). Extremal behavior for regularly varying stochastic processes. *Stoch. Proc. Appl.* 115 (2), 249–274.

[28] Jobst, A.A. (2007). Operational risk: the sting is still in the tail but the poison depends on the dose. *J. Operational Risk* 2(2), 3–59.

[29] Jorion, P. (2006). *Value–at–Risk*. 3rd Edition. New York: McGraw-Hill.

[30] Lambrigger, D.D., Shevchenko, P.V., Wüthrich, M.V. (2007). The quantification of operational risk using internal data, relevant external data and expert opinion. *J. Operational Risk* 2(3), 3–27.

[31] Li, D. (2001). On default correlation: a copula function approach. *J. Fixed Income* 9, 43–54.

[32] Mancini, L., Trojani, F. (2007). Robust Value–at–Risk prediction. Swiss Finance Institute Research Paper No. 07–31.

[33] Marshall, A.W., Olkin, I. (2007). *Life Distributions.* New York: Springer.

[34] McLeish, D.L. (2005). *Monte Carlo Simulation & Finance.* New York: John Wiley.

[35] McLeish, D.L. (2008). Bounded relative error importance sampling and rare event simulation. University of Waterloo. Preprint.

[36] McNeil, A.J., Frey, R., Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools.* Princeton: Princeton University Press.

[37] Moscadelli, M. (2004). The modelling of operational risk: experience with the analysis of the data collected by the Basel Committee. Bank of Italy, working Paper No 517.

[38] Nešlehová, J., Embrechts, P., Chavez–Demoulin, V. (2006). Infinite–mean models and the LDA for operational risk. *J. Operational Risk* 1(1), 3–25.

[39] Ramos, A., Ledford, A. (2008). A new class of models for bivariate joint tails. Oxford–Man Institute of Quantitative Finance, University of Oxford. Preprint.

[40] Resnick, S.I. (2006). *Heavy–Tail Phenomena: Probabilistic and Statistical Modeling.* New York: Springer.

[41] Rosenberg, J., Schuermann, T. (2006). A general approach to integrated risk management with skewed, fat–tailed risks. *J. Financial Economics* 79, 569–614.

[42] Scholes M.S. (2000). Crisis and risk management. *American Economic Review* 90, 17–22.

[43] SCOR (2008). *From Principle–Based Risk Management to Solvency Requirements. Analytical Framework for the Swiss Solvency Test.* SCOR Switzerland AG, Zurich.