

Extreme Value Theory in Finance

Erik Brodin*

Claudia Klüppelberg[†]

Abstract

Extreme value theory is a practical and useful tool for modeling and quantifying risk. In this article, after introducing basic concepts, we indicate how to apply it within a financial framework.

1 Introduction

During the last decades the financial market has grown rapidly. As a result this has led to the demand and, consequently, to the development of risk management systems. The internationally recognised regulatory concept is called *Basel II accord*, which suggests international standards for measuring the adequacy of a bank's risk capital. Hence, when discussing risk in finance it is of importance to keep the Basel II accord in mind; see [35] for an interesting discussion of regulation of financial risk. Also, in the Basel II accord, the different aspects of financial risk are defined. We distinguish

- **Market risk:** the risk that the value of an investment will decrease due to movements in market factors.
- **Credit risk:** the risk of loss due to a debtor's non-payment of a loan (either the principal or interest (coupon) or both).
- **Operational risk:** the risk of losses resulting from inadequate or failed internal processes, people and systems, or external events.

In this paper we discuss extreme value theory in general and indicate how to use it to model, measure and assess financial risk. On balance, extreme value theory is a practical and useful tool for modeling and quantifying risk. However, it should, as with all model-based statistics, be used with care. For interesting reading of extreme value theory and risk management, see [15, 41].

The structure of the paper is as follows. In Section 2 we introduce basic concepts of extreme value theory for independent and identically distributed (iid) data, independent of financial applications. In Section 3 we introduce financial market risk and describe how to model and analyze it with tools from extreme value theory. Here we also present

*Department of Mathematical Sciences, Chalmers University of Technology, SE-412 96 Göteborg, Sweden, email: ebrodin@math.chalmers.se, www.chalmers.se/math/EN/

[†]Center for Mathematical Sciences, Munich University of Technology, D-85747 Garching, Germany, email: cklu@ma.tum.de, www.ma.tum.de/stat/

methods for temporal dependent financial time series and indicate multivariate modeling, which both are important for risk management to assess joint losses at subsequent times and of different instruments in a portfolio. As the role of extreme value theory in credit risk and operational risk has not yet been clarified and is under inspection, we refrain from a thorough presentation. We would like, however, to refer to the papers [9, 30, 33, 38] for an extreme value approach to credit risk and to [5, 6, 10, 37] for interesting developments in operational risk modeling. Further references can be found in these papers.

2 Extreme Value Theory

Extreme value theory (EVT) is the theory of modelling and measuring events which occur with very small probability. This implies its usefulness in risk modelling as risky events per definition happen with low probability. Textbooks on EVT include [3, 11, 16, 32, 39, 40]. The book [3] treats extreme value statistics from a mathematical statistician's point of view with interesting case studies, and contains also material on multivariate extreme value statistics, whereas [11] has more emphasis on applications. The book [16] combines theoretical treatments of maxima and sums with statistical issues, focusing on applications from insurance and finance. In [32] EVT for stationary time series is treated. The monograph [39] is an extreme value statistics book with various interesting case studies. Finally, [40] is a mathematically well balanced book aiming mainly at multivariate extreme value theory based on multivariate regular variation. We also want to mention the journal *Extremes* which specialises on all probabilistic and statistical issues in EVT and its applications.

In EVT there are two main approaches with their own strength and weakness. The first one is based on modelling the maximum of a sample (or a few largest values of a sample, called the upper order statistics) over a time period. In [16], Chapter 3, this approach is rigorously formulated based on the Fisher-Tippett theorem going back to 1928. The second approach is based on modelling excess values of a sample over a threshold within a time period. This approach is called *Peaks Over Threshold* or POT-method and has been suggested originally by hydrologists.

Both approaches are equivalent by the Pickands-Balkema-de Haan theorem presented in [16], Theorem 3.4.5. Statistics based on EVT has to use the largest (or smallest) values of a sample. They can be selected in different ways and we assume at the moment that we have iid data. The first method is based on a so-called block maxima method, where data are divided into blocks, whose maxima are assumed to follow an extreme value distribution. The second method uses the joint distribution function of upper order statistics and the third method uses the POT-method, i.e. it invokes excesses over a high threshold. In this paper we focus at the POT-method for two reasons. Firstly, there is an almost unison agreement that its performance is better than other EVT methods

when estimating quantiles; see e.g. [35]. Secondly, it can be extended easily to dependent data. This is also true for the block maxima method, there however the blocks have to be chosen so that the resulting block maxima are independent, so that maximum likelihood estimation (MLE) can be applied. However, then the block maxima method invokes considerably less data than the POT-method, which makes the POT-method more efficient.

The existing software for the analysis of extreme events has been reviewed in [45], where a glossary of available packages can be found. We want to mention EVIS, which is written in SPlus, available at www.ma.h-w.ac.uk/~mcneil with corresponding package EVIR based on R. Also, we recommend the package EVD, also based on R, which includes programs for multivariate extreme value analysis at www.maths.lancs.ac.uk/~stephana. Finally, the MATLAB package EVIM is available at www.bilkent.edu.tr/~faruk.

The following introduction of the POT-method is based on [17]. The POT-method estimates a far out tail or a very high quantile, based on extreme observations of a sample and consists of three parts. Each part is based on a probabilistic principle which will be explained in the following paragraphs. Figure 2.1 serves as an illustration.

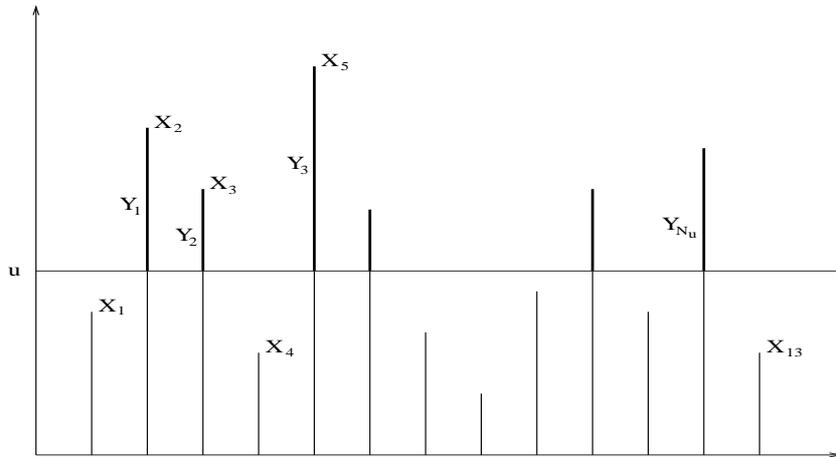


Figure 2.1: Data X_1, \dots, X_{13} with corresponding excesses Y_1, \dots, Y_{N_u} .

(1) Point process of exceedances

Given a high threshold u_n we index each observation of the sample X_1, \dots, X_n exceeding u_n . (In Figure 2.1 these are observations 2, 3, 5, 6, 10, 12.) To obtain a limit result, we let the sample size n tend to infinity and, simultaneously, the threshold u_n increase, and this in the correct proportion.

For iid data, each data point has the same chance to exceed the threshold u_n , the success probability being simply $P(X_i > u_n)$ for $i = 1, \dots, n$. Hence, the number of observations exceeding this threshold

$$\#\{i : X_i > u_n, i = 1, \dots, n\} = \sum_{i=1}^n I(X_i > u_n)$$

follows a binomial distribution with parameters n and $P(X_i > u_n)$. Here, $I(X_i > u_n) = 1$ or 0, according as $X_i > u_n$ or $\leq u_n$. If for some $\tau > 0$

$$nP(X_i > u_n) \rightarrow \tau, \quad n \rightarrow \infty, \quad (2.1)$$

then by the classical theorem of Poisson, the distribution of $\#\{i : X_i > u_n, i = 1, \dots, n\}$ converges to a Poisson distribution with parameter τ . If $X_i, i = 1, \dots, n$, come from an absolutely continuous distribution, (2.1) is a rather weak condition: for all known absolutely continuous distributions and every $\tau > 0$, a suitable series (u_n) can be found (see e.g. [16], Chapter 3). Indexing all points $\{i : X_i > u_n, i = 1, \dots, n\}$ in the interval $[0, n]$, the latter will become larger and larger whereas the indexed points will become sparser and sparser (as the threshold u_n rises with n). A more economic representation is gained by not plotting the points on the interval $[0, n]$ but rather on the interval $[0, 1]$. An observation X_i exceeding u_n will then be plotted not at i but at i/n . If for $n \in \mathbb{N}$ we define

$$N_n((a, b]) = \#\{i/n \in (a, b] : X_i > u_n, i = 1, \dots, n\}$$

for all intervals $(a, b] \subset [0, 1]$, then N_n defines a point process on the interval $[0, 1]$. This process is called the *time-normalized point process of exceedances*. Choosing u_n such that (2.1) holds, the series N_n of point processes converges (as $n \rightarrow \infty$) in distribution to a Poisson process with parameter τ . For the measure theoretical background on convergence of point processes see e.g. [16], Chapter 5.

(2) The generalized Pareto distribution

For the exceedances of a high threshold, we are not only interested in when and how often they occur, but also in how large the excess $X - u | X > u$ is. (In Figure 2.1 the excesses are labelled Y_1, \dots, Y_{N_u} and the number of exceedances is $N_u = 6$.) Under condition (2.1) it can be shown that for a measurable positive function a ,

$$\lim_{u \rightarrow \infty} P\left(\frac{X - u}{a(u)} > y \mid X > u\right) = (1 + \xi y)^{-1/\xi},$$

if the left-hand side converges at all. For $\xi = 0$ the right-hand side is interpreted as e^{-y} . For all $\xi \in \mathbb{R}$ the right-hand side is the tail of a distribution function, the so-called *generalized Pareto distribution*. If $\xi \geq 0$ the support of this distribution is $[0, \infty)$, for $\xi < 0$ the support is a compact interval. The case $\xi < 0$ is of no interest for our applications and therefore not considered.

(3) Independence

Finally, it can be shown that the point process of exceedances and the excesses, that is, the sizes of the exceedances, are in the limit independent.

How can these limit theorems be used to estimate tails and quantiles?

The following paragraph illustrates the POT-method for a given sample X_1, \dots, X_n . For a high threshold u we define

$$N_u = \#\{i : X_i > u, i = 1, \dots, n\}.$$

We will refer to the excesses of X_1, \dots, X_n as Y_1, \dots, Y_{N_u} , as indicated in Figure 2.1. The tail of F will be denoted by $\bar{F} = 1 - F$. Defining $\bar{F}_u(y) = P(Y_1 > y \mid X > u)$ yields

$$\bar{F}_u(y) = P(X - u > y \mid X > u) = \frac{\bar{F}(u + y)}{\bar{F}(u)}, \quad y \geq 0.$$

Consequently, we get

$$\bar{F}(u + y) = \bar{F}(u)\bar{F}_u(y), \quad y \geq 0. \quad (2.2)$$

An observation larger than $u + y$ is obtained if an observation exceeds u , i.e. an exceedance is required, and if, furthermore, such an observation has an excess over u that is also greater than y . An estimator of the tail (for values greater than u) can be obtained by estimating both tails on the right-hand side of (2.2).

This is done by exploiting (1)–(3).

We estimate $\bar{F}(u)$ by its empirical counterpart

$$\widehat{\bar{F}(u)} = \frac{1}{n} \sum_{i=1}^n I(X_i > u) = \frac{N_u}{n}.$$

Then, we approximate $\bar{F}_u(y)$ by the generalized Pareto distribution, where the scale function $a(u)$ has to be taken into account. The latter will be integrated into the limit distribution as a parameter. This gives

$$\bar{F}_u(y) \approx \left(1 + \xi \frac{y}{\beta}\right)^{-1/\xi},$$

where ξ and β have to be estimated (by $\widehat{\xi}$ and $\widehat{\beta}$). (Note that $\beta = \beta(u)$ is a function of u .) This results for given u in the following tail estimator:

$$\widehat{\bar{F}(u + y)} = \frac{N_u}{n} \left(1 + \widehat{\xi} \frac{y}{\widehat{\beta}}\right)^{-1/\widehat{\xi}}, \quad y \geq 0. \quad (2.3)$$

By inversion, one obtains for a given $\alpha \in (0, 1)$ an estimator of the α -quantile of the form

$$\widehat{x}_\alpha = u + \frac{\widehat{\beta}}{\widehat{\xi}} \left(\left(\frac{n}{N_u} (1 - \alpha) \right)^{-\widehat{\xi}} - 1 \right). \quad (2.4)$$

Finding the threshold u is important but by no means trivial. It is a classical variance-bias problem, made worse by the small sample of rare events data. The choice of the threshold is usually based on an extended explorative statistical analysis; cf. Section 6.2 in [16] or Chapter 4 in [11]. Once the threshold is found, parameter estimation can be performed by maximum likelihood. The seminal paper [43] shows that MLEs are asymptotically normal, provided that $\xi > -0.5$, and derives the asymptotic covariance matrix of the estimators; see [16], Section 6.5 or [11], Chapter 4. This allows also to calculate confidence intervals. Alternatively, they can also be obtained by the profile likelihood method; see [11]. There are algorithms on adaptive threshold selection, some asymptotically optimal, in settings similar to the POT-method, see e.g. [3], Section 4.7. However, to apply such black box methods can be dangerously misleading (see the Hill horror plot in Figure 4.1.13 of [16]) and we view such automatic methods as a complementary tool, which can be helpful in finding an appropriate threshold, but which should not be used stand alone.

3 Financial Risk Management

We recall that market risk is the risk that the value of an investment will decrease due to movements in market factors. Examples include equity risk, interest rate risk, currency risk and commodity risk. The most common risk measure in the finance industry is the *Value-at-Risk* (VaR), which is also recommended in the Basel II accord. Consider the loss L of a portfolio over a given time period Δ , then VaR is a risk statistic that measures the risk of holding the portfolio for the time period Δ . Assume that L has distribution function F_L , then we define VaR at level $\alpha \in (0, 1)$ as

$$\text{VaR}_\alpha^\Delta(L) = \inf\{x \in \mathbb{R} : P(L > x) \leq 1 - \alpha\} = \inf\{x \in \mathbb{R} : F_L(x) \geq \alpha\}.$$

Typical values of α are 0.95 and 0.99, while Δ usually is 1 day or 10 days. For an overview on VaR in a more economic setting we refer to [24].

Although intuitively a good concept, VaR has met criticism from academia and practice because of some of its theoretical properties and its non-robustness in statistical estimation. Based on an axiomatic approach, [1] suggests so called coherent risk measures, which excludes VaR as it is in general not subadditive. An alternative set of axioms has been suggested in [21], leading to convex risk measures, motivated by the economic fact that risk may not increase linearly with the portfolios size.

A risk measure closely related to VaR, which is coherent, is the expected shortfall (ES), also known as Conditional VaR (CVaR), which is defined as the expected loss given that we have a loss larger than VaR. Assume that $E[|L|] < \infty$, then ES is defined as

$$\text{ES}_\alpha^\Delta(L) = E[L \mid L > \text{VaR}_\alpha^\Delta(L)].$$

Provided the loss distribution is continuous, then

$$\text{ES}_\alpha^\Delta(L) = \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_x^\Delta(L) dx.$$

Although ES is a more informative risk measure than VaR it suffers from even higher variability as we are using information very far out in the tail. For a discussion on the problematic description of complex risk with just one single number we refer to [41].

As VaR and ES are per definition based on a high quantile, EVT is a tailor made tool for estimating both. Instead of considering the whole distribution estimation is based on the tail of the loss distribution in the context of the acronym *letting the tail speak for itself*. This is the only way of obtaining an estimator based on the relevant data for an extreme event. We will here concentrate on the estimation procedure using the POT-method as described in Section 2.

Concretely, using the generalized Pareto distribution, $\text{VaR}_\alpha^\Delta(L)$ is estimated as the α -quantile of the log-differences (called logreturns) of the portfolio loss process $(L_t)_{t \in \mathbb{N}_0}$ in a time interval of length Δ . The logreturns are assumed to constitute a stationary time series, then the data used for estimation are given by

$$X_t^\Delta = \log(L_t) - \log(L_{t-\Delta}), \quad t = k\Delta, k \in \mathbb{N}_0.$$

Now, using equation (2.4) for a well-chosen threshold $u > 0$, we obtain an estimate for the quantile of the logreturn loss distribution as

$$\widehat{\text{VaR}}_\alpha(X^\Delta) = u + \frac{\hat{\beta}}{\hat{\xi}} \left(\left(\frac{n}{N_u} (1-\alpha) \right)^{-\hat{\xi}} - 1 \right). \quad (3.1)$$

Similarly, ES can be estimated, provided that the logreturns have finite expectation ($\xi < 1$), and we obtain

$$\widehat{\text{ES}}_\alpha(X^\Delta) = \frac{\widehat{\text{VaR}}_\alpha(X^\Delta)}{1-\hat{\xi}} + \frac{\hat{\beta} - \hat{\xi}u}{1-\hat{\xi}}. \quad (3.2)$$

In a bank, usually, a daily VaR is estimated, i.e. $\Delta = 1$ day based on daily logreturns. However, according to the Basel II accord $\Delta = 10$ days has to be estimated as well. This is statistically problematic as one needs a very long time series to get good estimates, which on the other hand puts stationarity to the question. To avoid this problem, and also to avoid multiple estimation, one often uses a scaling rule,

$$\widehat{\text{VaR}}_\alpha(X^\Delta) = \Delta^\kappa \widehat{\text{VaR}}_\alpha(X^1) \quad \text{and} \quad \widehat{\text{ES}}_\alpha(X^\Delta) = \Delta^\kappa \widehat{\text{ES}}_\alpha(X^1). \quad (3.3)$$

Here typically $\kappa = 0.5$ is chosen, the so called square root scaling rule, which is based on the central limit theorem; see [35] for a discussion on different values of κ . Applying such a scaling rule as in (3.3) can, however, be crossly misleading as realistic financial models usually do not allow for such naive scaling; cf. [13] for an example.

3.1 Time Series Approach

A stylized fact of financial logreturns is that they are seemingly uncorrelated, but the serial correlation of absolute or squared logreturns is significant. This is due to the empirical fact that a large absolute movement tends to be followed by a large absolute movement. One can view this as a varying temperature of the market introducing dependence into the data. In a naive EVT approach one ignores this effect and, therefore, produces estimators with non optimal performance.

To estimate VaR and/or ES for temporally dependent data two approaches have been suggested in the literature. The first approach is called *unconditional* VaR or ES estimation and is based on a POT-method, where possible dependence is taken into account. In such models, provided they are stationary, the POT-method presented by (1)-(3) in Section 2 has to be modified to account for possible clustering in the extremes. The point process of (2) can, by dependence, become a marked Poisson process. Exceedances over high thresholds appear in clusters, so that a single excess as in the Poisson process turns into a random number of excesses, whose distribution is for certain models known; see [32] or [16], Chapter 8, for the theoretical basis and [19, 25] for examples. As a consequence, the maximum value of a large sample of size n from such a dependent model behaves like the maximum value of a smaller sample size $n\theta$ for some $\theta \in (0, 1)$ of an iid sample with the same marginal distribution. The parameter θ is called extremal index and can under weak regularity conditions be interpreted as the inverse of the mean cluster size. This parameter enters into the tail and quantile estimators (2.3) and (2.4), which become for a high threshold $u > 0$

$$\widehat{F}(u + y) = \frac{N_u}{n\theta} \left(1 + \widehat{\xi} \frac{y}{\widehat{\beta}} \right)^{-1/\widehat{\xi}}, \quad y \geq 0, \quad (3.4)$$

respectively, for $\alpha \in (0, 1)$,

$$\widehat{x}_\alpha = u + \frac{\widehat{\beta}}{\widehat{\xi}} \left(\left(\frac{n\theta}{N_u} (1 - \alpha) \right)^{-\widehat{\xi}} - 1 \right). \quad (3.5)$$

The parameter θ has to be estimated from the data and is again, of course, highly dependent of the threshold u ; see [16], Section 8.1, for details and estimation procedures. New methods have been suggested in [20, 31].

The second approach is called *conditional* or *dynamic* VaR or ES estimation. The most realistic models introduce stochastic volatility and model a portfolio's daily logreturns by

$$X_t = \mu_t + \sigma_t Z_t, \quad t = 1, \dots, n$$

where σ_t is the stochastic volatility, μ_t is the expected return and Z_t is a random variable with mean zero and variance one. It starts by estimating μ_t and σ_t , usually by means of

quasi-maximum likelihood, and applies the classical POT-method from Section 2 to the residuals $\hat{Z}_t = (X_t - \hat{\mu}_t)/\hat{\sigma}_t$. The residuals \hat{Z}_t are sometimes called devolatilized logreturns. We obtain for a given time period Δ ,

$$\widehat{\text{VaR}}_\alpha(X^\Delta) = \hat{\mu}_{t+\Delta} + \hat{\sigma}_{t+\Delta}\widehat{\text{VaR}}_\alpha(Z^\Delta)$$

and

$$\widehat{\text{ES}}_\alpha(X^\Delta) = \hat{\mu}_{t+\Delta} + \hat{\sigma}_{t+\Delta}\widehat{\text{ES}}_\alpha(Z^\Delta).$$

The most prominent conditional mean and variance model for μ_t and σ_t is the ARMA-(G)ARCH model which considers the mean to be an ARMA process and the volatility to be a (G)ARCH process, see [7]. It is worth mentioning that Robert F. Engle received the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel in 2003 for his development of the (G)ARCH models. For a successful implementation of estimating VaR and ES using the ARMA-(G)GARCH model we refer to the explanatory paper [34]. It exemplifies that an AR-(G)ARCH model together with the generalized Pareto distribution for the residuals produces very accurate VaR and ES estimates.

Financial market data of liquid markets are tick-by-tick data, where each tick corresponds to a transaction. Such data are also termed high frequency data and have been in the focus of research in recent years. The accessibility of high frequency data has led to the development of continuous-time financial models, see e.g. [2, 28] for two different models and [29] for a comparison. VaR estimation based on high-frequency data has been considered in [4], estimating VaR on different time frames using high frequency data; see also [12] for a discussion on extreme returns of different time frames.

3.2 Multivariate Issues

Often when modelling financial risk one investigates the development of a portfolio's logreturns, aggregating different risk factors. This prevents the possibility of tracking joint large losses, which may indeed lead to complete disaster. Consequently, one should model a portfolio as a multivariate random vector. Unfortunately, this leads in financial applications easily to a very high-dimensional problem, and modeling of the dependence structure could be very difficult or even impossible. One way out is to use comparably few selected risk factors, or to group assets in different sectors or geographical regions.

Multivariate EVT provides the theoretical background to model and analyze joint extreme events by concentrating on dependence in extreme observations. An indication of extreme dependence in financial data is the well-known fact that market values tend to fall together in a market crash. Multivariate EVT makes a contribution to this problem by offering a tool to model, analyze and understand the extreme dependence structures.

In multivariate EVT there are mainly two approaches. The first one utilizes multivariate regular variation. This approach is based on the fact that marginal models for high

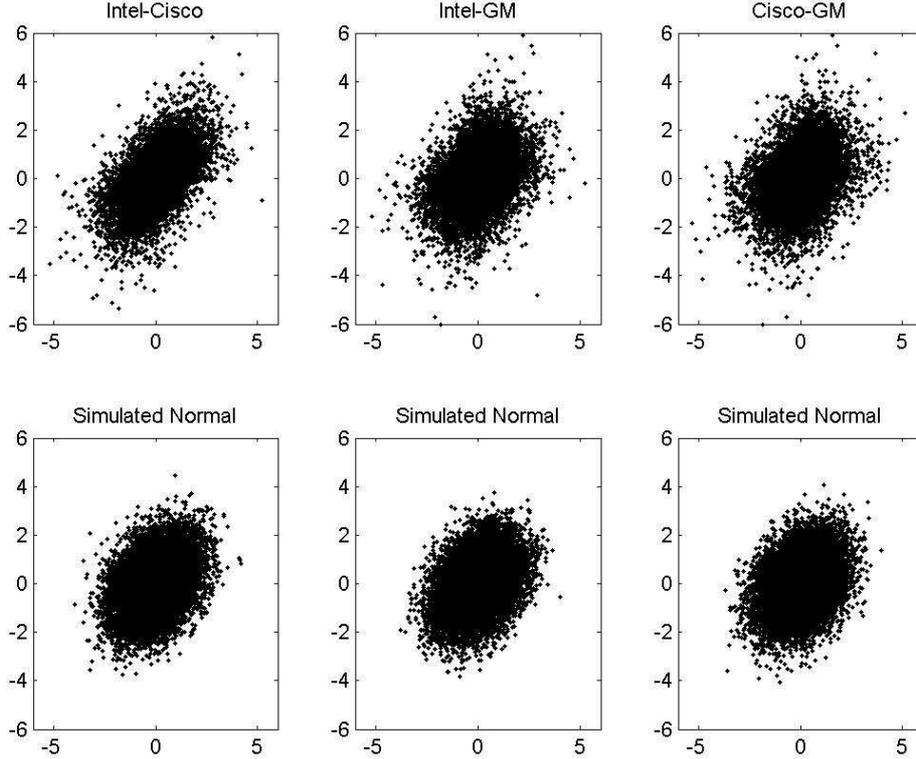


Figure 3.1: Bivariate stock data versus simulated normal random numbers with same (estimated) means and variances.

risk data are very well modelled by Pareto-like models, whose tails decrease like $P(X > x) = x^{-\alpha} \ell(x)(1 + o(1))$ as $x \rightarrow \infty$ for some function ℓ satisfying $\lim_{t \rightarrow \infty} \ell(xt)/\ell(t) = 1$ for all $x > 0$. Given that all marginal distributions in the portfolio have such tails with the same α extreme dependence can be modelled by the so-called *spectral measure*. Indeed, a multivariate regularly varying random vector has univariate regularly varying distribution tails along each direction from the origin and a measure, the spectral measure, which describes the extreme dependence. More precisely, a random vector X in \mathbb{R}^d is *regularly varying with index* $\alpha \in [0, \infty)$ and spectral measure σ on the Borel sets of the unit sphere $\mathbb{S}^{d-1} := \{z \in \mathbb{R}^d : \|z\| = 1\}$, if for all $x > 0$

$$\frac{P(\|X\| > xt, X/\|X\| \in \cdot)}{P(\|X\| > t)} \xrightarrow{v} x^{-\alpha} \sigma(\cdot) \quad t \rightarrow \infty,$$

where $\|\cdot\|$ denotes any norm in \mathbb{R}^d and \xrightarrow{v} denotes vague convergence; for details see [36, 40] and references therein. This approach has been applied for instance in [22, 44] to analyze high-frequency foreign exchange (FX) data.

The second approach is based on Pickands' representation of a multivariate extreme value distribution, which aims at the asymptotic frequency for joint large values in a portfolio. For two risk factors or assets X and Y with distribution functions F and G , respectively, the risk manager is interested in $P(X > x \text{ or } Y > y)$ for x, y large. This is

measured by the quantity

$$\begin{aligned}\Lambda(x, y) &= \lim_{n \rightarrow \infty} n P(F(X) > 1 - \frac{x}{n} \text{ or } G(Y) > 1 - \frac{y}{n}) \\ &= \int_0^{\pi/2} \left(\frac{x}{1 \vee \cot \theta} \vee \frac{y}{1 \vee \tan \theta} \right) \Phi(d\theta),\end{aligned}$$

and the limit exists under weak regularity conditions with finite measure Φ on $(0, \pi/2)$ satisfying

$$\int_0^{\pi/2} \frac{1}{1 \wedge \cot \theta} \Phi(d\theta) = \int_0^{\pi/2} \frac{1}{1 \wedge \tan \theta} \Phi(d\theta) = 1$$

The measure Φ has to be estimated from the data, which can be done parametrically and non-parametrically. As it is more intuitive to estimate (dependence) functions instead of measures: *Pickands' dependence function* and a so-called *tail-dependence function* have been introduced. They both aim at assessing joint extreme events. For Pickands' dependence function we refer to [11], Chapter 8, and [3], Section 8.2.5, where parametric and non-parametric estimation procedures have been suggested and investigated. The tail dependence function has been proposed and estimated non-parametrically in [23]. Non-parametrical approaches are based on fundamental papers like [14]; for further references see this paper.

Estimates as above can be improved for elliptical distributions or even for distributions with elliptical copula; cf. [26, 27]. The tail dependence function is also invoked in [8], which includes an in depth analysis of multivariate high-frequency equity data.

Multivariate analogues to the POT-method of Section 2 have been developed. We refer to [3, 11] for further details. Multivariate generalized Pareto distributions are natural models for multivariate POT data; see [42] and further references there.

Acknowledgement

The paper was partly written while the first author visited the Center for Mathematical Sciences of the Munich University of Technology. He takes pleasure in thanking the colleagues there for their hospitality. Financial support by the Chalmerska Forskningsfonden through a travelling grant is gratefully acknowledged.

References

- [1] Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999) Coherent measures of risk. *Mathematical Finance* **9**, 203-228.
- [2] Barndorff-Nielsen, O.E. and Shephard, N. (2001) Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion). *J. R. Statist. Soc. Ser. B* **63**, 167-241.

- [3] Beirlant, J. Goegebeur, Y. Segers, J. and Teugels, J. (2004), *Statistics of Extremes: Theory and Applications*. Wiley, Chichester.
- [4] Beltratti, A. and Morana, C. (1999), Computing value at risk with high frequency data. *Journal of Empirical Finance* **6**(5), 431-455.
- [5] Böcker, K. and Klüppelberg, C. (2005) Operational VaR: a closed-form approximation. *RISK*, December 2005, 90-93.
- [6] Böcker, K. and Klüppelberg, C. (2006) Multivariate models for operational risk. Submitted for publication.
- [7] Bollerslev, T, Engle, R.F. and Nelson, D. (1994) ARCH Models. In: R. Engle and D. McFadden (Eds.) *Handbook of Econometrics, Volume IV*, pp. 2959-3038. Elsevier, Amsterdam.
- [8] Brodin, E. and Klüppelberg, C. (2006) Modeling, estimation and visualization of multivariate dependence for high-frequency data. *Preprint* Munich University of Technology. www.ma.tum.de/stat/.
- [9] Campbell, R. A. and Huisman, R. (2003) Measuring credit spread risk: Incorporating the tails. *J. Portfolio Management* Summer 2003, 121-127.
- [10] Chavez-Demoulin, V., Embrechts, P. and Nešlehová (2005) Quantitative models for operational risk: extremes, dependence and aggregation. *Journal of Banking and Finance* **30**(10), 2635-2658.
- [11] Coles, S. G. (2001) *An Introduction to Statistical Modeling of Extreme Values*. Springer, London.
- [12] Dacorogna, M., Müller, U., Pictet, O. and de Vries, C. (2001) Extremal forex returns in extremely large data sets. *Extremes* **4**(2), 105-127.
- [13] Drost, F.C. and Nijman, T.E. (1993) Temporal aggregation of GARCH processes. *Econometrica* **61**, 909-927.
- [14] Einmahl, J., de Haan, L. and Piterbarg, V.I. (2001) Nonparametric estimation of the spectral measure of an extreme value distribution. *Annals of Statistics* **29**, 1401-1423.
- [15] Embrechts, P. (Ed.) (2000) *Extremes and Integrated Risk Management*. UBS Warburg and Risk Books.
- [16] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.
- [17] Emmer, S., Klüppelberg, C. and Trüstedt, M. (1998) VaR - ein Maß für das extreme Risiko. *Solutions* **2**, 53-63.

- [18] Fasen, V. and Klüppelberg, C. (2005) Extremes of supOU processes. Proceedings to the Abel Symposium 2005. Springer. To appear.
- [19] Fasen, V., Klüppelberg, C. and Lindner, A. (2006) Extremal behavior of stochastic volatility models. In: A.N. Shiryayev, M.d.R. Grossinho, P.E. Oliveira, and M.L. Esquivel (Eds.) *Stochastic Finance*, pp. 107-155. Springer, New York.
- [20] Ferro, T. A and Segers, J. (2003) Inference for clusters of extreme values. *J. R. Statist. Soc. Ser. B* **65**, 545-556.
- [21] Föllmer, H. and Schied, A. (2002) Convex measures of risk and trading constraints. *Finance and Stochastics* **6**(4), 429-447.
- [22] Hauksson, H., Dacorogna, M., Domenig, T., Müller, U. and Samorodnitsky, G. (2001) Multivariate extremes, aggregation and risk estimation. *Quantitative Finance* **1** (1), 79-95.
- [23] Hsing, T., Klüppelberg, C. and Kuhn, G. (2004) Dependence estimation and visualization in multivariate extremes with applications to financial data. *Extremes* **7**, 99-121.
- [24] Jorion, P. (2001) *Value at Risk: The New Benchmark for Measuring Financial Risk*. McGraw-Hill, New York.
- [25] Klüppelberg, C. (2004) Risk management with extreme value theory. In: Finkenstädt, B. and Rootzén, H. (Eds) *Extreme Values in Finance, Telecommunication and the Environment*, pp.101-168. Chapman and Hall/CRC, Boca Raton.
- [26] Klüppelberg, C., Kuhn, G. and Peng, L. (2005a) Semi-parametric models for the multivariate tail dependence function. *Bernoulli*. To appear.
- [27] Klüppelberg, C., Kuhn, G. and Peng, L. (2005b) Estimating the tail dependence of an elliptical distribution. Submitted for publication.
- [28] Klüppelberg, C., Lindner, A. and Maller, R. (2004) A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour. *J. Appl. Prob.* **41**(3), 601-622.
- [29] Klüppelberg, C., Lindner, A. and Maller, R. (2006) Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models. In: Kabanov, Y., Lipster, R. and Stoyanov, J. (Eds.) *From Stochastic Calculus to Mathematical Finance. The Shiryayev Festschrift*, pp. 393-419. Springer, Berlin.
- [30] Kuhn, G. (2004) Tails of credit default portfolios. Technical Report. Munich University of Technology. Available at <http://www-m4.ma.tum.de/Papers/>
- [31] Laurini F. and Tawn, J.A. (2003) New estimators for the extremal index and other cluster characteristics. *Extremes* **6**(3), 189-211.

- [32] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.
- [33] Lukas, A., Klaasen, P., Spreij, P. and Straetmans S. (2003) Tail behavior of credit loss distributions for general latent factor models. *Appl. Math. Finance* **10**(4), 337-357.
- [34] McNeil, A. and Frey, R. (2000) Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance* **7**, 271-300.
- [35] McNeil, A., Frey, R. and Embrechts, P. (2005) *Quantitative Risk Management*. Princeton University Press, Princeton.
- [36] Mikosch, T. (2005) How to model multivariate extremes if one must? *Statistica Neerlandica* **59**(3), 324-338.
- [37] Moscadelli, M. (2004) The modelling of operational risk: experience with the analysis of the data collected by the Basel committee. *Technical Report* **517**, Banca d'Italia.
- [38] Phoa, W. (1999) Estimating credit spread risk using extreme value theory. *J. Portfolio Management* Spring 1999, 69-73.
- [39] Reiss, R.-D. and Thomas, M. (2001) *Statistical Analysis of Extreme Values. With Applications to Insurance, Finance, Hydrology and Other Fields. 2nd ed.* Birkhäuser, Basel.
- [40] Resnick, S. I. (1987) *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
- [41] Rootzén, H. and Klüppelberg, C. (1999) A single number can't hedge against economic catastrophes. *Ambio* **28**, 550-555.
- [42] Rootzén, H. and Tajvidi, N. (2006) Multivariate generalized Pareto distribution. *Bernoulli* **12**(5), 917-930.
- [43] Smith, R.L. (1987) Estimating tails of the probability distributions. *Ann. Stat.* **15**, 1174-1207.
- [44] Stărică, C. (1999) Multivariate extremes for models with constant conditional correlations. *Journal of Empirical Finance* **6**, 515-553.
- [45] Stephenson, A. and Gilleland, E. (2006) Software for the analysis of extreme events: The current state and future directions. *Extremes* **8**(3), 87-109.