Universität Regensburg
Mathematik

Modeling financial markets with extreme risk

Tobias Kusche

Preprint Nr. 04/2008

Modeling financial markets with extreme risk

Dr. Tobias Kusche

11. January 2008

1 Introduction

The Black-Scholes model (BS model) is a standard tool for modeling stock prices. If the initial value of the stock S is given by S_0 then the dynamic of S is described by the solution of the SDE

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t)$$

and the initial condition

$$S(0) = S_0.$$

Here, $\mu \in \mathbb{R}$ is called trend, $\sigma > 0$ volatility and $W(t)$ is a one dimensional standard Brownian motion. The solution is given by

$$S(t) = S(0)e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t)}.$$

An advantage of the BS model is that there are explicit formulas for a wide class of derivatives. Moreover, the methods for estimating the parameters μ and σ are easy to apply. But the BS-model has some disadvantages. For example, it does not explain the volatility smile. Moreover, the BS model underestimates the probability for large returns, i.e. it is light tailed, and jumps in the stock price are excluded. These jumps are produced by events with a strong impact on the underlying asset, for example natural disasters or international crisis.

In order to model the effect of such events, jump-diffusion models have been suggested, cf. [M]. These models are an extension of the BS model. At certain random times there occur jumps in the rate of return. Suppose the sock price is given by

$$S(t) = S(0)e^{\left(\mu - \beta\lambda - \frac{\sigma^2}{2}\right)t + \sigma W(t) + \sum_{i=1}^{N(t)} \ln(Y_i + 1)}, \tag{1}$$

where $N(t)$ is a poisson process with parameter $\lambda \geq 0$ and Y_i is an i.i.d sequence of integrable random variables (rvs) with mean β. It is shown in [S, Chapter 11] that (1) is the solution of the SDE

$$dS(t) = (\mu - \beta\lambda)S(t)dt + \sigma S(t)dW(t) + S(t-)dQ(t), \tag{2}$$

where

$$Q(t) = \sum_{i=1}^{N(t)} Y_i$$

is a compound poisson process. In the paper of [K, p. 1087] the random variable (rv) $\ln(Y_i + 1)$ has density

$$p \cdot \eta_1 \cdot e^{-\eta_1 \cdot y} 1_{\{y \geq 0\}} + q \cdot \eta_2 \cdot e^{\eta_2 \cdot y} 1_{\{y < 0\}},$$

where $\eta_i > 0$ and p, q are nonnegative such that $p + q = 1$. For further information on jump-diffusion models, we refer the reader to the work of Merton, [M], and [S].

After a time discretization, equation (1) delivers a model for the asset price of the shape

$$S_n = S_0 e^{\sum_{i=1}^n R_i}. \tag{3}$$

Roughly speaking, we apply extreme value theory to the sequence R_i of daily log-returns. For further information on extreme value theory, we refer the reader to [EKM]. The key assumption in this paper is

(A) The distribution function (df) of $\pm R_n$ belongs to the maximum domain of attraction of the Frechet distribution, i.e.

$$P\{\pm R_n > t\} \sim L_\pm(t) t^{-\alpha_\pm}, \quad t \to \infty,$$

for suitable $\alpha_\pm > 0$ and slowly varying functions $L_\pm$.

Section 2 starts with some basic results on extreme value theory. Then we show that real market data give rise to assumption (A). We use market data for the Dow Jones Industrial Average and the Nasdaq Composite.

In section 3, we prove that under suitable assumptions on the jumps $\ln(Y_i + 1)$, the log-returns of the process in (1) deliver a discrete-time model of the shape (3) such that (A) is fulfilled.

The work closes with section 4, where we suggest a model for the jumps that gives a qualitative explanation for a few properties of the market data from section 2.

2 Extremal events in log-returns

2.1 General setting

Suppose we have a discrete-time model for the price of an asset, given by

$$S_n = S_0 e^{\sum_{i=1}^{n} R_i}, \quad n \in \mathbb{N}_0. \tag{4}$$

We assume that the sequence

$$R_n = \ln\left(\frac{S_n}{S_{n-1}}\right), \quad n \in \mathbb{N},$$

is i.i.d and call it the daily log-returns. The index $n \in \mathbb{N}$ gives the number of days. Suppose $P\{R_n > 0\} \in (0,1)$ and that the distribution of R is absolutely continuous with respect to the Lebesgue measure. In order to apply extreme value theory, we require that

$$\pm R \in \mathrm{MDA}\left(\Phi_{\alpha_\pm}\right) \tag{5}$$

where $\alpha_\pm = \zeta_\pm^{-1}$ and $\zeta_\pm > 0$. Here, $\mathrm{MDA}\left(\Phi_{\alpha_\pm}\right)$ denotes the maximum domain of attraction of the Frechet distribution with parameter $\alpha_\pm$ and H_θ is the generalized extreme value distribution, i.e.

$$H_{(\zeta,\mu,\psi)}(x) = \begin{cases} e^{-\left(1+\frac{\zeta}{\psi}(x-\mu)\right)^{-\frac{1}{\zeta}}}, & \zeta \neq 0 \\ \\ e^{-e^{-\frac{x-\mu}{\psi}}}, & \zeta = 0 \end{cases}$$

for $1 + \frac{\zeta}{\psi}(x-\mu) > 0$, $\zeta, \mu \in \mathbb{R}$, and $\psi > 0$, cf. [EKM, Chapter 6]. We have

$$\Phi_\alpha = H_{(\alpha^{-1},1,\alpha^{-1})}, \quad \alpha > 0.$$

Let us first consider the consequences of assumption (5). Define

$$M_n^\pm := \max(\pm R_1, ..., \pm R_n)$$

for $n \in \mathbb{N}$. Calculation yields

$$P(M_n^\pm \leq t) = F_\pm(t)^n$$

where $F_\pm$ is the df of $\pm R$. In view of [EKM, Definition 3.3.1], there exist norming constants $c_n^\pm > 0$ and $a_n^\pm \in \mathbb{R}$ such that

$$P\left(\frac{M_n^\pm - a_n^\pm}{c_n^\pm} \leq t\right) = F_\pm(c_n^\pm t + a_n^\pm)^n \to \Phi_{\alpha_\pm}(t), \quad n \to \infty. \tag{6}$$

2.2 Real market data

From yahoo.de we obtained $N = 19883$ daily quotations for the Dow Jones Industrial Average, starting at 01.10.1928 and ending at 06.12.2007. We denote these data by $\hat{S}_n$, $n = 0, 1, ..., N-1$, and assume that these data are a sample of the discrete-time process given in (4). Moreover, we set

$$\hat{R}_n := \ln\left(\frac{\hat{S}_n}{\hat{S}_{n-1}}\right), n = 1, 2, ..., N-1.$$

Here, $n = 0$ corresponds to the date 01.10.1928. In order to check for assumption (5), we define

$$\hat{R}_i^{\pm} := \pm\hat{R}_i$$

for $i = 1, ..., N-1$. Now we built maxima in blocks of length 20, i.e. set

$$K := \left[\frac{N-1}{20}\right],$$

and define

$$\hat{M}_i^{\pm} := \max\left(\hat{R}^{\pm}_{1+(i-1)\cdot 20}, ..., \hat{R}^{\pm}_{i\cdot 20}\right), \quad i = 1, ..., K.$$

Suppose (5) holds. Then by (6), we expect that there exists $\theta_{\pm} = (\zeta_{\pm}, \mu_{\pm}, \psi_{\pm})$ such that the distribution of $H_{\theta_{\pm}}$ is a good fit for the data $\hat{M}_i^{\pm}$, $i = 1, ..., K$.

2.3 Parameter estimation

The question is, how to obtain an estimate $\hat{\theta}_{\pm} = (\hat{\zeta}_{\pm}, \hat{\mu}_{\pm}, \hat{\psi}_{\pm})$ for $\theta_{\pm}$. We decided to apply the maximum likelihood estimator (MLE) described in [EKM, Section 6.3]. As an initial value for the MLE , we have chosen the parameter delivered by the probability-weighted moment estimator (PWM estimator) described in [HWW]. A description of the PWM estimator can also be found in [C]. The results are shown in table 1. We applied the same procedure to the data for the Nasdaq Composite with $N = 9295$ daily quotations starting at 05.02.1971 and ending on 06.12.2007. Source of the data is yahoo.de. The estimated parameters are shown in table 2. The tables 3 and 4 contain the initial values given by the PWM estimator. Figure 1 shows the QQ-plot for $\hat{M}_i^{\pm}$ and $H_{\hat{\theta}_{\pm}}$ (red line). The QQ-plot is almost the bisecting line and this indicates that the df $H_{\hat{\theta}_{\pm}}$ is a good fit for the data. For further information about the QQ-plot, see [EKM, Section 6] and the references therein. For each QQ-plot, we have computed a threshold for which at least 99 % of the plotted points have an x-coordinate bellow that threshold. The values of the thresholds are shown in table 5. The task now is to find a continuous time model that fulfills the following assumptions:

1. The price process of the asset has the shape of the process in (1).

2. After a discretization at times $i \cdot \tau = \frac{i}{250}$, $i \in \mathbb{N}_0$, the price process has the representation (4) and fulfills assumption (5).

MLE	$\hat{\zeta}$	$\hat{\mu}$	$\hat{\psi}$
$\hat{\theta}_+$	0.3202	0.0124	0.0059
$\hat{\theta}_-$	0.3020	0.0118	0.0066

Table 1: Parameter $\hat{\theta}_\pm$ for the Dow Jones index

MLE	$\hat{\zeta}$	$\hat{\mu}$	$\hat{\psi}$
$\hat{\theta}_+$	0.3518	0.0113	0.0060
$\hat{\theta}_-$	0.2630	0.0122	0.0074

Table 2: Parameter $\hat{\theta}_\pm$ for the Nasdaq Composite

PWM	$\hat{\zeta}$	$\hat{\mu}$	$\hat{\psi}$
$\hat{\theta}_+$	0.3330	0.0124	0.0058
$\hat{\theta}_-$	0.3261	0.0117	0.0064

Table 3: PWM estimate of $\theta_\pm$ for the Dow Jones index

3 Jump-diffusion models with fat tails

In order to price derivatives, we want to adapt a jump-diffusion model to the market data. Suppose the log-returns are given by

$$R(t) := \mu \cdot t + \sigma W(t) + \sum_{i=1}^{N(t)} \ln(Y_i + 1),$$

where

1. $W(t)$, $t \geq 0$, is a one dimensional standard brownian motion on the filtered probability space $(\Omega, \mathcal{F}, P, \mathbb{F})$ and the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ is complete.

2. $N(t)$, $t \geq 0$, is a poisson process on $(\Omega, \mathcal{F}, P, \mathbb{F})$ with parameter $\lambda > 0$.

PWM	$\hat{\zeta}$	$\hat{\mu}$	$\hat{\psi}$
$\hat{\theta}_+$	0.3347	0.0113	0.0060
$\hat{\theta}_-$	0.2340	0.0123	0.0076

Table 4: PWM estimate of $\theta_\pm$ for the Nasdaq Composite

	$\hat{\theta}_+$	$\hat{\theta}_-$
Dow Jones	0.0894	0.0742
Nasdaq	0.0764	0.0751

Table 5: Outliers in the QQ-plot

3. Y_i, $i \in \mathbb{N}$, is an i.i.d sequence of rvs such that $Y_i > -1$ and

$$\pm \ln(Y_i + 1) \in \mathrm{MDA}(\Phi_{\alpha_\pm}) \cap L^2(\Omega)$$

for some $\alpha_\pm > 0$.

4. The processes W, N and Y_i are independent. Moreover, $\mu \in \mathbb{R}$ and $\sigma > 0$ are constants.

We discretize the process

$$S(t) := e^{R(t)}, \quad t \geq 0,$$

at the time-points $\tau \cdot i$, $i \in \mathbb{N}_0$, where $\tau = \frac{1}{250}$ (daily quotations). Define

$$S_n := S(n \cdot \tau), \quad n \in \mathbb{N}_0,$$

and

$$R_n := \ln\left(\frac{S_n}{S_{n-1}}\right), \quad n \in \mathbb{N}.$$

It follows that

$$S_n = S_0 e^{\sum_{i=1}^n R_i},$$

where

$$R_i = \mu \cdot \tau + \sigma \left[W(i \cdot \tau) - W((i-1) \cdot \tau)\right] + \sum_{k=N((i-1)\cdot\tau)+1}^{N(i\cdot\tau)} \ln(Y_k + 1).$$

Especially, we have

$$R_n \sim \mu \cdot \tau + \sigma W(\tau) + \sum_{i=1}^{N(\tau)} \ln(Y_i + 1).$$

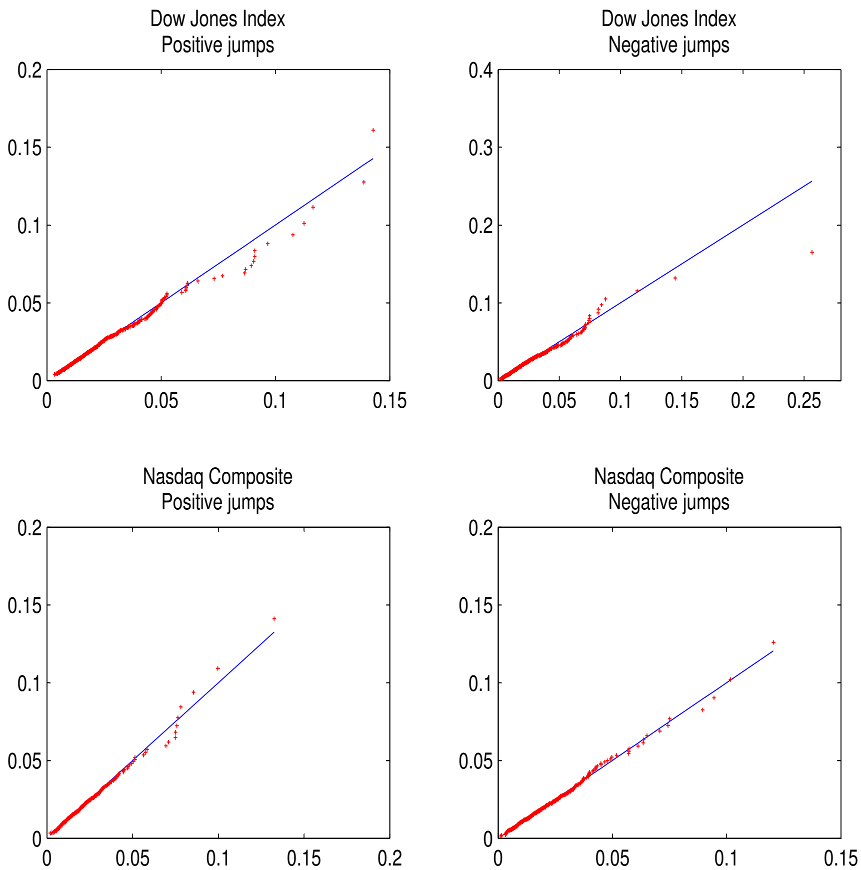

Figure 1: QQ-plot for $\hat{\theta}_{\pm}$ (red), bisecting line (blue)

In order to simplify the notation, we define

$$J := \sum_{i=1}^{N(\tau)} \ln(Y_i + 1).$$

Moreover, we denote the df of J by G.

3.1 Tail of the jump-part

Now we have to answer the question whether the df of R_n is in $\mathrm{MDA}(\Phi_{\alpha_\pm})$. Recall that

$$R_n = \ln\left(\frac{S_n}{S_{n-1}}\right), \quad n \in \mathbb{N}.$$

This is an i.i.d sequence and we have

$$R_n \sim \mu \cdot \tau + \sigma W(\tau) + J.$$

This leads to the question whether a random variable of the type

$$N(\mu, \sigma^2 \cdot \tau) + J$$

and $J \in \mathrm{MDA}(\Phi_{\alpha_\pm})$ is in $\mathrm{MDA}(\Phi_{\alpha_\pm})$ again. First, we clarify whether $\pm J \in \mathrm{MDA}(\Phi_{\alpha_\pm})$. The proof of the following lemma is a modification of the arguments given in the proof of [EKM, Lemma 1.3.1]. In contrast to [EKM, Lemma 1.3.1], we have to handle the case of a nonnegative rv.

Lemma 1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and X_1, X_2 independent rvs. Suppose that X_i is absolutely continuous with respect to the Lebesgue measure. Denote the df of X_i and $X_1 + X_2$ by F_i and F, respectively. Suppose there exist slowly varying functions $L_\pm$ that have a uniform positive lower bound for t large enough such that*

$$\overline{F_i}(t) = L_+(t) t^{-\alpha_+}$$

and

$$F_i(-t) = L_-(t) t^{-\alpha_-}$$

for $t \gg 1$. Moreover, assume

$$\forall n \in \mathbb{N} : \frac{L_\pm(t+n)}{L_\pm(t)} \to 1, \quad t \to \infty.$$

Then the following assertions hold:

1. *Positive tails:*
$$\frac{\bar{F}(t)}{\bar{F}_1(t) + \bar{F}_2(t)} \to 1, \quad t \to \infty.$$

2. *Negative tails:*
$$\frac{F(-t)}{F_1(-t) + F_2(-t)} \to 1, \quad t \to \infty.$$

Proof. Suppose for the moment, that X_1 and X_2 fulfill no other assumptions beside the independence. Fix $n \in \mathbb{N}$ and $t > 0$. Then we have

$$[\{X_1 > t + n\} \cap \{X_2 \geq -n\}] \cup [\{X_2 > t + n\} \cap \{X_1 \geq -n\}] \subset \{X_1 + X_2 > t\}.$$

As P is monotone and due to the formula

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

we have

$$\bar{F}_1(t+n)\bar{F}_2(-n) + \bar{F}_2(t+n)\bar{F}_1(-n) - \bar{F}_1(t+n)\bar{F}_2(t+n) \leq \bar{F}(t). \tag{7}$$

Now, we also make use of the additional assumptions. Divison by $\bar{F}_1(t) + \bar{F}_2(t)$ delivers

$$\frac{\bar{F}_1(t+n)}{\bar{F}_1(t)+\bar{F}_2(t)}\bar{F}_2(-n)+\frac{\bar{F}_2(t+n)}{\bar{F}_1(t)+\bar{F}_2(t)}\bar{F}_1(-n)-\frac{\bar{F}_1(t+n)}{\bar{F}_1(t)+\bar{F}_2(t)}\bar{F}_2(t+n) \leq \frac{\bar{F}(t)}{\bar{F}_1(t)+\bar{F}_2(t)}.$$

For sufficiently large t, the third addenda on the right side is smaller or equal to

$$\left(1+\frac{n}{t}\right)^{-\alpha_+}\frac{L_+(t+n)}{L_+(t)}\bar{F}_2(t+n)$$

and the latter one tends to zero as $t \to \infty$. Moreover,

$$\frac{\bar{F}_i(t+n)}{\bar{F}_1(t)+\bar{F}_2(t)} = \frac{1}{2}\left(1+\frac{n}{t}\right)^{-\alpha_+}\frac{L_+(t+n)}{L_+(t)} \to \frac{1}{2}, \quad t \to \infty.$$

It follows that

$$\frac{1}{2}\left(\bar{F}_1(-n)+\bar{F}_2(-n)\right) \leq \liminf_{t\to\infty}\frac{\bar{F}(t)}{\bar{F}_1(t)+\bar{F}_2(t)}.$$

Taking the limit $n \to \infty$, we obtain

$$1 \leq \liminf_{t\to\infty}\frac{\bar{F}(t)}{\bar{F}_1(t)+\bar{F}_2(t)}.$$

In order to finish the proof, we use the inclusion

$$\{X_1+X_2>t\} \subset \{X_1>(1-\delta)t\} \cup \{X_2>(1-\delta)t\} \cup \Big[\{X_1>\delta t\} \cap \{X_2>\delta t\}\Big]$$

for $\delta \in \left(0, \frac{1}{2}\right)$. This inclusion can be found in [EKM, Lemma 1.3.1] and holds for arbitrary rvs. As the rvs are independent, we obtain

$$\bar{F}(t) \leq \bar{F}_1((1-\delta)t) + \bar{F}_2((1-\delta)t) + \bar{F}_1(\delta t)\bar{F}_2(\delta t). \tag{8}$$

This delivers

$$\frac{\bar{F}(t)}{\bar{F}_1(t)+\bar{F}_2(t)} \leq (1-\delta)^{-\alpha_+}\frac{L_+((1-\delta)t)}{L_+(t)} + \delta^{-\alpha_+}\frac{L_+(\delta t)}{L_+(t)}\bar{F}_2(\delta t)$$

for t large enough. Sending $t \to \infty$, the slow variation of $L_\pm$ implies

$$\limsup_{t\to\infty}\frac{\bar{F}(t)}{\bar{F}_1(t)+\bar{F}_2(t)} \leq (1-\delta)^{-\alpha_+}.$$

Taking the limit $\delta \to 0$, we obtain the assertion. This yields the assertion for the positive tails. In order to prove the second part of the lemma, set

$$Y_i := -X_i$$

and

$$Y := Y_1 + Y_2.$$

Then

$$P\{-Y_i > t\} = F_i(-t)$$

and

$$P\{-Y > t\} = F(-t).$$

If we apply the first part of the lemma to these new rvs, we obtain the assertion on the negative tails. $\square$

We use this lemma to prove that at least the jump part fulfills the assumption (5) at the beginning of this article. For a subexponential df F, a version of the following lemma can be found in [EKM, Example 1.3.10]. Note that a subexponential df has support in $(0, \infty)$.

Lemma 2. *Suppose F is the df of $\ln(Y_i + 1)$. Then we have*

$$\bar{G}(t) \sim \lambda\tau\bar{F}(t)$$

and

$$G(-t) \sim \lambda\tau F(-t)$$

as $t \to \infty$.

Proof. As Y_i is i.i.d, calculation yields

$$\overline{F^{(n+1)*}}(t) = \int_{-\infty}^{+\infty} \overline{F^{n*}}(t-x)dF(x) = \int_{-\infty}^{+\infty} \overline{F^{n*}}(t-x)P_{\ln(Y_i+1)}(dx).$$

Division by $\bar{F}(t)$ and extending the fraction delivers

$$\frac{\overline{F^{(n+1)*}}(t)}{\bar{F}(t)} = \int_{-\infty}^{+\infty} \frac{\overline{F^{n*}}(t-x)}{\bar{F}(t-x)} \cdot \frac{\bar{F}(t-x)}{\bar{F}(t)} dF(x). \tag{9}$$

Define

$$c_n := \sup_{z\in\mathbb{R}} \frac{\overline{F^{n*}}(z)}{\bar{F}(z)}.$$

Equation (9) delivers

$$c_{n+1} \leq c_n \cdot \sup_{t\in\mathbb{R}} \int_{-\infty}^{+\infty} \frac{\bar{F}(t-x)}{\bar{F}(t)} dF(x).$$

we obtain

$$c_{n+1} \leq c_n \cdot c_2$$

and this implies

$$c_n \leq c_2^{n-1}, \quad n \geq 2.$$

Moreover, c_2 is finite because $F(t) < 1$ on $\mathbb{R}$ and in view of lemma 1 we have

$$\lim_{t\to\infty} \frac{\overline{F^{2*}}(t)}{\bar{F}(t)} = 2.$$

It follows that

$$\sum_{i=1}^{\infty} \frac{(\lambda\tau)^i}{i!} \cdot \frac{\overline{F^{i*}}(t)}{\bar{F}(t)}$$

is uniformly bounded in t. In view of the convergence

$$\frac{\overline{F^{i*}}(t)}{\bar{F}(t)} \to i, \quad t \to \infty,$$

the assertion follows. $\square$

It's left to prove that the diffusion process does not change the tail of the jump-part. More precisely, we want to prove that

$$N(\mu, \sigma^2) + \mathrm{MDA}(\Phi_{\alpha_\pm}) \in \mathrm{MDA}(\Phi_{\alpha_\pm}).$$

Lemma 3. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and X_i, $i = 1, 2$, independent rvs such that*

$$X_1 \sim N(\mu, \sigma^2)$$

and

$$P\{\pm X_2 > t\} = L_\pm(t) t^{-\alpha_\pm}$$

where $L_\pm$ is a slowly varying function, bounded from bellow by a positive constant for t sufficiently large. Assume further that

$$\forall n \in \mathbb{N} : \frac{L_\pm(t+n)}{L_\pm(t)} \to 1, \quad t \to \infty.$$

Then

$$P\{\pm(X_1 + X_2) > t\} \sim P\{\pm X_2 > t\}.$$

Proof. Proceed as in the proof of lemma 1 in order to obtain the equations (7) and (8). Divide (7) by $\bar{F}_2(t)$. We obtain

$$\frac{\bar{F}_1(t+n)}{\bar{F}_2(t)} \bar{F}_2(-n) + \frac{\bar{F}_2(t+n)}{\bar{F}_2(t)} \bar{F}_1(-n) - \frac{\bar{F}_2(t+n)}{\bar{F}_2(t)} \bar{F}_1(t+n) \le \frac{\bar{F}(t)}{\bar{F}_2(t)}.$$

Due to the assumption, we have

$$\bar{F}_1(t) \le c_1 e^{-p(t)}$$

for a constant $c_1 > 0$ and a polynomial p with degree one. This implies

$$\frac{\bar{F}_1(t+n)}{\bar{F}_2(t)} \leq \frac{c_1}{s} e^{\alpha_+ \ln(t) - p(t+n)} \to 0, \quad t \to \infty,$$

where $s > 0$ is the lower bound of $L_+(t)$. Taking the limit $t \to \infty$, we obtain

$$\bar{F}_1(-n) \leq \liminf_{t \to \infty} \frac{\bar{F}(t)}{\bar{F}_2(t)}.$$

As this is true for arbitrary $n \in \mathbb{N}$, it follows that

$$1 \leq \liminf_{t \to \infty} \frac{\bar{F}(t)}{\bar{F}_2(t)}.$$

Let us divide equation (8) by $\bar{F}_2(t)$ to obtain

$$\frac{\bar{F}(t)}{\bar{F}_2(t)} \leq \frac{\bar{F}_1((1-\delta)t)}{\bar{F}_2(t)} + \frac{\bar{F}_2((1-\delta)t)}{\bar{F}_2(t)} + \bar{F}_1(\delta t) \frac{\bar{F}_2(\delta t)}{\bar{F}_2(t)} \leq$$

$$\frac{c_1}{s} e^{\alpha_+ \ln(t) - p((1-\delta)t)} + (1-\delta)^{-\alpha_+} \frac{L_+((1-\delta)t)}{L_+(t)} + \bar{F}_1(\delta t) \delta^{-\alpha_+} \frac{L_+(\delta t)}{L_+(t)}.$$

It follows that

$$\limsup_{t \to \infty} \frac{\bar{F}(t)}{\bar{F}_2(t)} \leq (1-\delta)^{-\alpha_+}.$$

Taking the limit $\delta \to 0$, we obtain

$$\limsup_{t \to \infty} \frac{\bar{F}(t)}{\bar{F}_2(t)} \leq 1.$$

If we consider $-X_i$ instead of X_i, we obtain the assertion for the negative tails. $\square$

With lemma 1-3, we obtain that R_n fulfills assumption (5). So far, the model seems to be a good candidate for the real market data described in section 2.

3.2 Shortfall distribution

A characteristic feature of a random variable with distribution function in MDA(Φ_α) is the shortfall distribution. If X is a random variable on the probability space $(\Omega, \mathcal{A}, P)$ with df F and $u \geq 0$ a certain threshold, then

$$F_u(t) := P(X - u \leq t | X > u), \quad t > 0,$$

is called shortfall df of X for the threshold u. According to [EKM, Theorem 3.4.13], there exists a function $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\lim_{u \to \infty} \sup_{t > 0} \left| F_u(t) - G_{\zeta, \beta(u)}(t) \right| = 0,$$

where $G_{\zeta,\beta}$ is the generalized Pareto distribution, cf. [EKM, p. 162]. If the given model fits the data $\hat{R}_n$ well, we expect to observe a Pareto distribution for the peaks over a sufficiently large threshold.

The real market data deliver the samples $\hat{R}_n^{\pm}$. In order to simplify the notation, we only treat the case $\hat{R}^+ = \hat{R}$. Let $r_1 \leq ... \leq r_{N-1}$ be the order statistics of $\hat{R}$. Let k be the smallest integer such that $r_k \geq 0.01$. We choose the threshold values

$$u_i = r_i$$

for $i = k, ..., N-1-100$. Define

$$I_i := \{j \geq i+1 : r_j > u_i\},$$

$$n_i := |I_i|$$

and denote the order statistics of $\{r_j : j \in I_i\}$ by

$$X_{n_i,n_i} \leq ... \leq X_{1,n_i}. \tag{10}$$

The QQ-plot of (10) is given by the points

$$\left(X_{j,n_i}, G^{-1}_{\zeta_+,\beta_i}(y^i_j)\right), \quad j = 1, ..., n_i,$$

where

$$\beta_i := \zeta_+ \cdot u_i, \quad i = k, ..., N-1-100,$$

and

$$y^i_j = \frac{n_i - j + 1}{n_i + 1}, \quad j = 1, ..., n_i.$$

In order to measure the deviation of the QQ-plot from the bisecting line, we use

$$e_i := \frac{1}{n_i} \sum_{j=1}^{n_i} \left| G^{-1}_{\zeta_+,\beta_i}(y^i_j) - X_{j,n_i} \right|.$$

For the negative tail, one only has to replace $\hat{R}^+$ by $\hat{R}^-$. We performed the calculations for both, the Dow Jones Industrial Average and the Nasdaq Composite. The results are given in figure 2. The plot shows the points (r_i, e_i), $i = k, ..., N-1-100$.

4 Modeling the jumps

In this section, we make a special choice for the rv $\ln(Y_1+1)$. For the distribution of $\ln(Y_1+1)$, we construct a denisity function similar to that in [K, p. 1087]. We replace the exponential density by one of Pareto-type.

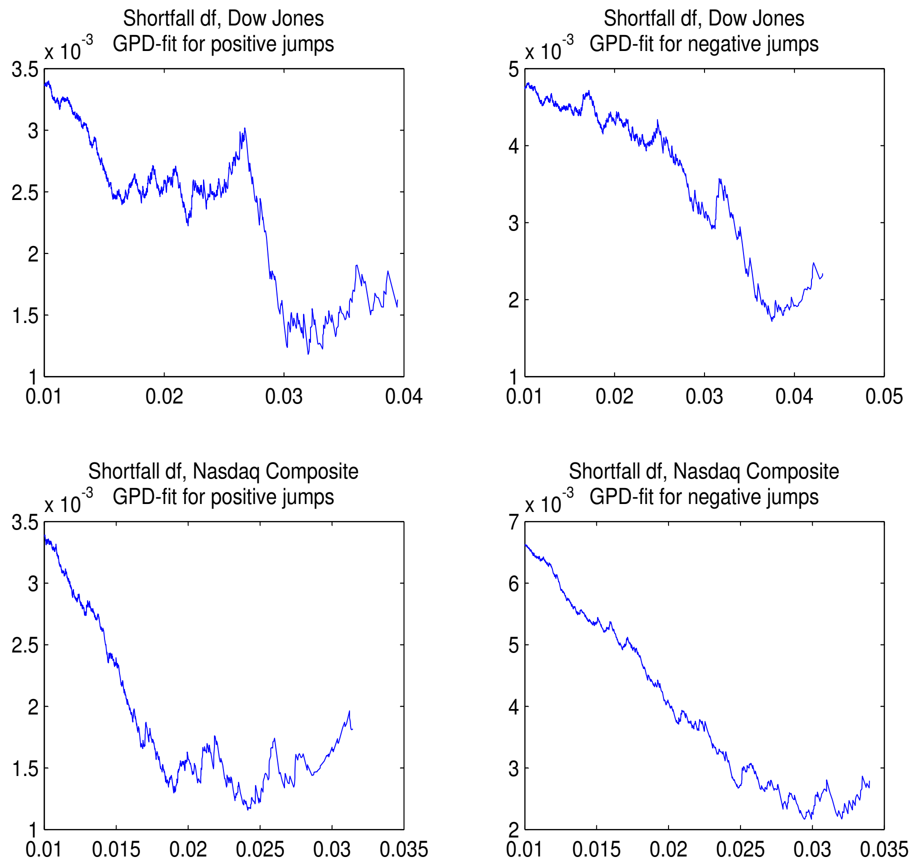

Figure 2: Error in the QQ-plot for the shortfall df with parameter $\hat{\zeta}_{\pm}$ and $\hat{\zeta}_{\pm} \cdot u_i$

Definition 1. *Suppose $\alpha_{\pm} > 2$, $u_{\pm} > 0$ and $p \in [0, 1]$. Define*

$$f(x) := p \cdot \frac{\alpha_+}{x} \left(\frac{u_+}{x}\right)^{\alpha_+} \cdot 1_{\{z \geq 0\}}(x) + (-1)^{\alpha_- + 1}(1-p) \cdot \frac{\alpha_-}{x} \left(\frac{u_-}{x}\right)^{\alpha_-} \cdot 1_{\{z<0\}}(x)$$

for $x \in \mathbb{R}$. Moreover, set

$$F_{\pm}(t) := 1 - \left(\frac{u_{\pm}}{t}\right)^{\alpha_{\pm}}$$

for $t \geq u_{\pm}$ and zero otherwise.

Suppose now that the distribution of $\ln(Y_1 + 1)$ has density f with respect to the Lebesgue measure. Calculation yields

$$\bar{F}(t) = p \cdot \bar{F}_+(t)$$

and

$$F(-t) = (1-p) \cdot \bar{F}_-(t)$$

for $t \geq 0$. Therefore,

$$\pm \ln(Y_1 + 1) \in \mathrm{MDA}(\Phi_{\alpha_\pm}).$$

Note further that

$$P\{\ln(Y_1 + 1) > 0\} = p,$$

i.e. with probability p an observed jump is positive. With lemma 1 - 3, we obtain

$$\frac{P\{R_n > t\}}{\bar{F}_+(t)} \rightarrow p \cdot \lambda \cdot \tau$$

and

$$\frac{P\{R_n \leq -t\}}{\bar{F}_-(t)} \rightarrow (1-p) \cdot \lambda \cdot \tau$$

as $t \rightarrow \infty$. Moreover, if $u, t > 0$ then

$$F_u(t) = 1 - \left(1 + \frac{t}{u}\right)^{-\alpha} = G_{\zeta_+, u \cdot \zeta_+}(t).$$

In view of the results of section 3.2, the density f seems to be a good candidate for the density of $\ln(Y_1 + 1)$, because

1. it explains the observed tails,
2. the shortfall df is a Pareto distribution.

References

[C] Marcelo G. Cruz: Modeling, measuring and hedging operational risk. Wiley Finance Series 2002

[EKM] P. Embrechts, C. Klüppelberg, T. Mikosch: Modelling Extremal Events for Insurance and Finance. Springer-Verlag Berlin Heidelberg 1997.

[HWW] Hosking, J. R. M.; Wallis, J. R.; Wood, E. F. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics 27 (1985), no. 3, 251–261.

[K] S. G. Kou: A jump diffusion model for option pricing. Manag. Sci. 48 (2002), 1086–1101.

[M] Robert C. Merton: Option Pricing when underliying stock returns are discontinuous, J. Fin. Economics 3 (1976), 125-144.

[S] Steven E. Shreve: Stochastic Calculus for Finance II - Continuous-Time Models. Springer-Verlag New York 2004.