6 Value at Risk, Expected Shortfall and Risk Management

Ch 7 of the textbook
Extreme price movements in the financial markets are rare but important and crucial. The stock market crash on Wall Street in October 1987 and other big financial crises such as bankruptcy of Lehman brothers (2008) have attracted a great deal of attention among investors, practitioners and researchers. A market crash is characterized by increase in the market volatility (such as the VIX index) and big drops in the market index.

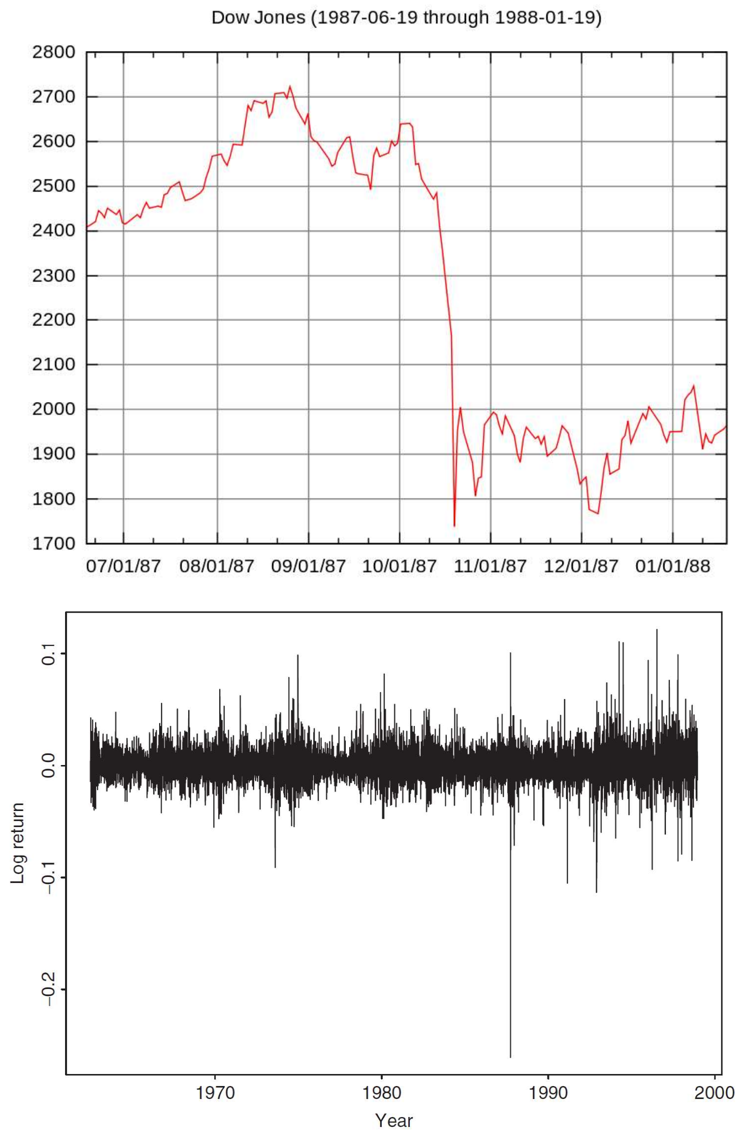

Figure 46: In finance, Black Monday refers to Monday, October 19, 1987, when stock markets around the world crashed. The crash began in Hong Kong and spread west to Europe, hitting the United States after other markets had already sustained significant declines. The Dow Jones Industrial Average (DJIA) fell exactly 508 points to 1,738.74 (22.61%). In Australia and New Zealand, the 1987 crash is also referred to as "Black Tuesday" because of the time zone difference.
Bottom panel: Time plot of daily log returns of IBM stock from July 3, 1962, to December 31, 1998.

Classification of Financial Risks

1. Market risk: risk due to changes in stock prices, interest rates, FX and commodity prices
2. Operational risk: includes legal and political risks
3. Credit risk: default risk

We start with the market risk, because

- easier to understand
- high-quality data are available
- the idea applicable to other types of risk

How can we measure the risk in the financial markets? How can we assess the risk present in the investments? To answer these questions, we will now look into what it means to have a "'coherent"' risk measure.

■ **Risk measure and coherence:** Let $L_t(\ell)$ be the loss variable of a financial position from time t to time $t+\ell$. Here ℓ denotes a holding period. Financial loss is concerned with the distribution of $L_t(\ell)$. All risk measures available in the literature are *summary statistics* of $L_t(\ell)$.

Coherence:

A risk measure η is *coherent* if for any two loss random variables X and Y:

1. Positive homogeneity: For $c > 0$, $\eta(cX) = c\eta(X)$
   The risk of a position is proportional to its size.
2. Translation invariance: For $c > 0$, $\eta(X + c) = \eta(X) + c$
3. Monotonicity: If $X \geq Y$ almost surely, then $\eta(X) \leq \eta(Y)$
   X always has better values than portfolio Y under almost all scenarios, then the risk of X should be less than the risk of Y.
4. Subadditivity: $\eta(X + Y) \leq \eta(X) + \eta(Y)$
   Related to diversification in finance.

Each of these properties have direct implications to how we view and quantify risk.

The first measure that we know is the variance or standard deviation, i.e., volatility of returns. If you have a portfolio with returns $\{r_t\}$ over time, you can measure the level of your risk by how much the return on your investment is deviating from the average returns. Standard deviation is an example of a coherent risk measure.

But when we care about extreme values, then variance might not be a sufficient measure. Sometimes our goal is to assess the negative extremes and their chances of occurring rather than the overall return. Variance on the other hand is symmetric and thus provides info regarding the entire data. In this chapter we look at risk measures that compare the tails, particularly in the lower tails, thus measuring the downside risk.

■ **What is Value at Risk (VaR)?**

We start with VaR, which is a measure of maximum loss of a financial position for a given probability. It is a point estimate of potential financial loss and contains a degree of uncertainty. VaR is an asymmetric measure of risk and is concerned with the maximum amount of loss for a given investment under "'normal"' conditions. Since the prices of assets are stochastic, we always relate VaR with a probability.

$VaR_{1-p}(L_t(\ell))$=an upper quantile of $L_t(\ell)$ for a small probability p.

VaR is defined for a given loss function $L_t(\ell)$, which is the loss from t to $t+\ell$. An example of a loss function can be the negative of the financial position i.e $L_t(\ell) = -r_{t+\ell}$ if you long the portfolio and $L_t(\ell) = r_{t+\ell}$ if you short the portfolio.

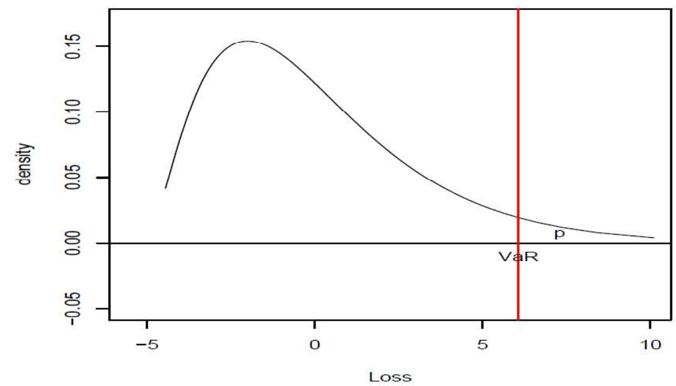

Figure 47: Definition of Value at Risk (VaR) for a continuous loss random variable based on the probability density function. Var is the upper quantile of the loss function over the holding period ℓ, $L_t(\ell)$, for a small tail probability, p.

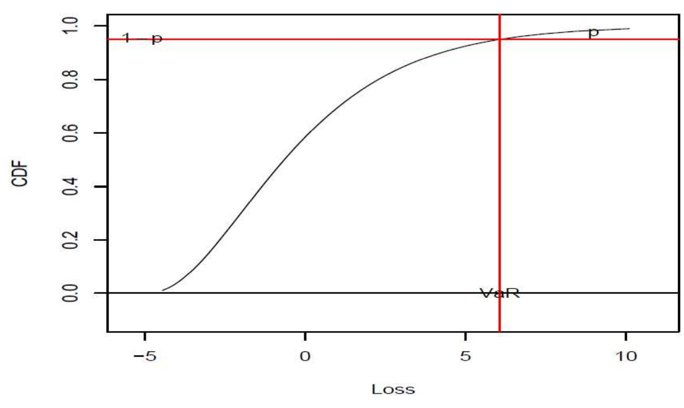

Figure 48: Definition of Value at Risk (VaR) for a continuous loss random variable based on the cumulative distribution function.

Thus, the probability of facing a loss greater than VaR_{1-p} is p. or, with probability $q = 1-p$, the potential loss is less than VaR_{1-p}.

Mathematically speaking, let $F_\ell(x)$ be the CDF of $x = L_t(\ell)$. The subscript of t is omitted from F to ease the notational burden. Then we have

1. $\text{VaR}_{1-p} = \inf\{x|F_\ell(x) \geq 1-p\}$ (Quantile definition)

2. $\Pr(L_t(\ell) \leq \text{VaR}_{1-p}) \geq 1-p$ (Investor perspective)
3. $\Pr(L_t(\ell) > \text{VaR}_{1-p}) \leq p$ (Regulatory perspective)

VaR can be used by:

1. Financial institutions- to asses their risk. They measure *the maximum loss of a financial position during a period with a given probability under normal conditions*. This measure of risk is viewed as a metric to judge financial positions by considering the maximum loss that can occur during normal conditions and is mathematically stated in (2). With probability $q = 1-p$, the potential loss is less than VaR_{1-p}.
2. Regulators- to set margin requirements. They measure *the minimum loss of a financial position within a certain period of time for a given (small) probability under extraordinary conditions*. They use this to find the amount of assets needed to cover possible losses (Margin setting). With probability p, the loss is greater than VaR_{1-p}. In other words, we need at least VaR_{1-p} as backup to sustain losses under extreme market conditions. Mathematically, we have the regulatory perspective in (3), which checks the probability of the loss being greater than VaR (If the loss function is continuous then we will have $\Pr(L_t(\ell) > \text{VaR}_{1-p}) = p$).

Calculating VaR boils down to calculating the quantile of the distribution of the loss function.

R packages: *VaRES* can be used to compute VaR and ES for many statistical distributions. *evir* (for extreme values), and *quantreg* (for quantile regression).

Recall the definition of quantiles: x_q is the $100q$-th quantile of the continuous distribution $F_\ell(x)$ if

$$q = F_\ell(x_q) \text{ i.e.,} \quad q = \Pr(L \leq x_q)$$

In general, we have $x_q = \min\{x | F_l(x_q) \geq q\}$.

We can calculate these quantiles using empirical techniques (non-parametric) or by making some assumptions about the loss distributions.

Discussion:

First, VaR is commonly used. It is simple, and some closed-form solutions are available:

1. Normal distribution: $L_t(\ell) \sim N(\mu_t(\ell), \sigma_t^2(\ell))$, then $Z_t = \frac{L_t - \mu_t(\ell)}{\sigma_t(\ell)}$ is standard normal and the VaR for $L_t(\ell)$ is

$$\text{VaR}_{1-p} = \mu_t(\ell) + z_{1-p}\sigma_t(\ell) \tag{6.1}$$

where z_{1-p} denotes the $(1-p)$th quantile of $N(0,1)$. Note that $z_{0.99} = 2.326$ and $z_{0.95} = 1.645$. In R, use *qnorm(0.95)* and *qnorm(0.99)*.

2. Student-t distribution: If $Y_t = (L_t(\ell) - \mu_t(\ell))/\sigma_t(\ell)$ follows a Student-t distribution with v degrees of freedom, then

$$\mathrm{VaR}_{1-p} = \mu_t(\ell) + t_{1-p,\nu}\sigma_t(\ell) \tag{6.2}$$

   where $t_{1-p,\nu}$ is the $(1-p)$th quantile of t_v. In R, use *qt(0.95,v)* and *qt(0.99,v)*.
3. Standardized t_v distribution: Assume $v > 2$. If $Y_t = (L_t(\ell) - \mu_t(\ell))/\sigma_t(\ell)$ follows a standardized Student-t distribution with v degrees of freedom, then

$$\mathrm{VaR}_{1-p} = \mu_t(\ell) + t^*_{1-p,\nu}\sigma_t(\ell) \tag{6.3}$$

   where $t^*_{1-p,\nu}$ is the $(1-p)$th quantile of standardized t_v. In R with package *fGarch*, use *qstd(0.95,nu=ν)* and *qstd(0.99,nu=ν)*.

We use different parametric approaches such as RiskMetrics or econometric modeling for the loss function. In each approach we find $\mu_t(\ell)$, $\sigma_t(\ell)$ and the underlying distribution and then calculate VaR_{1-p}.

Second, VaR is coherent for a normally distributed loss function. For instance, consider the subadditivity:

$$\begin{aligned}
\mathrm{Var}(X+Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y) \\
&= \sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y \\
&\le \sigma_x^2 + \sigma_y^2 + 2\sigma_x\sigma_y \\
&= (\sigma_x + \sigma_y)^2
\end{aligned}$$

Therefore, $\sigma_{x+y} \le \sigma_x + \sigma_y$. This implies that $z_{1-p}\sigma_{x+y} \le z_{1-p}\sigma_x + z_{1-p}\sigma_y$, or equivalently, VaR of $X+Y$ is less than or equal to the sum of VaR of X and VaR of Y.

VaR, however, is not coherent in general because the subadditivity property of VaR does not hold in general. VaR coherence exists if the distribution of the loss function is normal. A simple counterexample; see also Example 3.13 of Klugman, Panjer and Willmot (2008):

Suppose the CDF $F_\ell(x)$ of a continuous loss random variable X satisfies the following probabilities:

$$F_\ell(80) = 0.9215, \qquad F_\ell(90) = 0.95, \qquad F_\ell(100) = 0.97$$

For $p = 0.05$, the VaR of X is 90, because 90 is the 0.95th quantile of X. We denote this by $\mathrm{VaR}^x_{0.95} = 90$. Now, define two loss random variables X_1 and X_2 by

$$X_1 = \begin{cases} X & \text{if } X \le 100 \\ 0 & \text{if } X > 100 \end{cases} \quad \text{and } X_2 = \begin{cases} 0 & \text{if } X \le 100 \\ X & \text{if } X > 100 \end{cases}$$

These two loss variables are simply truncated versions of X and we have $X = X_1 + X_2$. Since the total probability must be 1, the CDF $F^1_\ell(X)$ of X_1 satisfies

$$F_\ell^1(80) = 0.9215/0.97 = 0.95, \quad F_\ell^1(90) = 0.95/0.97 = 0.9794, \quad F_\ell^1(100) = 0.97/0.97 = 1$$

The 0.95th quantile of X_1 is 80. Therefore, $\mathrm{VaR}^1_{0.95} = 80$, where the superscript 1 is used to denote X_1. On the other hand, $\Pr(X_2 \leq 0) = P(X \leq 100) = 0.97$. Therefore, the 0.95th quantile of X_2 is less than or equal to 0. We denote this by $\mathrm{VaR}^2_{0.95} \leq 0$. Taking the sum, we have $\mathrm{VaR}^1_{0.95}+\mathrm{VaR}^2_{0.95} \leq 80$.

In this particular instance, $X = X_1+X_2$, yet $\mathrm{VaR}^x_{0.95} > \mathrm{VaR}^1_{0.95}+\mathrm{VaR}^2_{0.95}$. Therefore, the subadditivity of VaR fails.

Remark 1. For a "long" position, use negative returns in your analyses as your loss function.

Remark 2. We use the *predictive density* of the loss (negative returns) to calculate VaR.

Remark 3. The $ amount of VaR is equal to

"Cash value of the financial position × VaR of negative log returns"

Remark 4. Finally VaR is just a quantile of the loss distribution. It does not describe the actual tail behavior of the loss random variable. It is not a perfect risk measure and does not fully describe the upper tail behavior of the loss function. In practice, two assets can have the same VaR yet face different losses when VaR is exceeded.

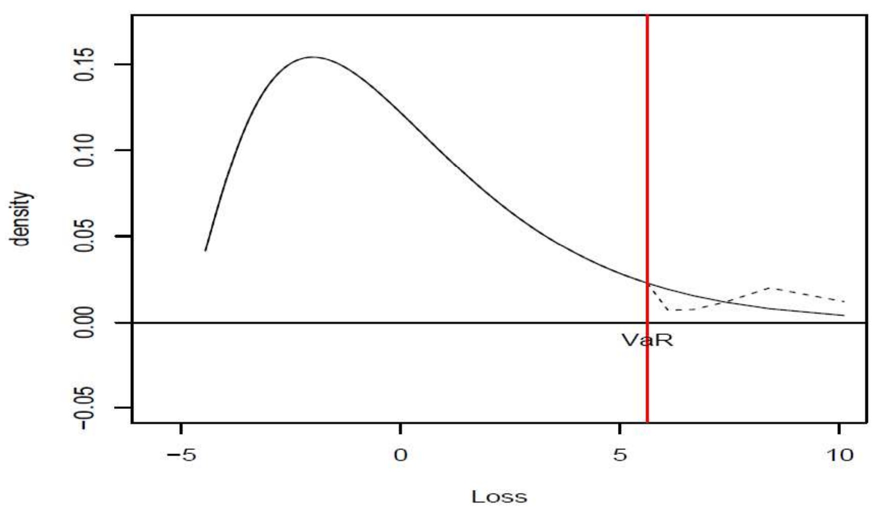

Figure 49: Density functions of two loss random variables that have the same VaR, but different loss implications.

■ **Expected Shortfall (ES)** Var is just a quantile of the loss distribution. It does not fully describe the upper tail behavior of the loss function. In practice, two assets can have the same VaR and yet face different losses if and when the VaR is exceeded. The actual loss, if it occurs, can obviously be greater than VaR. VaR may underestimate the actual loss. Given this it is important to realize that VaR is not a perfect risk measure.

To have a better assessment of the potential loss, we consider the expected value of the loss if the VaR is exceeded. This leads to the concept of Expected Shortfall also known as tail value at risk (TVaR) and conditional VaR (CVaR).

Simply put, ES is the expected loss of a financial position after a catastrophic event. ES of a loss variable X is defined as

$$ES_{1-p} = E[X|X > \text{VaR}_{1-p}] = \frac{\int_{\text{VaR}_{1-p}}^{\infty} xf(x)dx}{\Pr(X > \text{VaR}_{1-p})} = \frac{\int_{\text{VaR}_{1-p}}^{\infty} xf(x)dx}{p} \tag{6.4}$$

From the definition, ES is the expected loss of X given that X exceeds its VaR.

Note: For the two loss densities in Figure 49, their ES are different with the dash line corresponding to a higher value.

Assume that X is continuous. Let $u = F(x)$ for VaR$\leq X \leq \infty$. Then, we have $du = f(x)dx$, $F(\text{VaR}) = 1 - p$, $F(\infty) = 1$, and $x = F^{-1}(u) =$VaR$_u$. Equation 6.4 becomes

$$ES_{1-p} = \frac{\int_{1-p}^{1} \text{VaR}_u du}{p}$$

Thus, ES can be seen to average all VaR$_u$ for $1 - p \leq u \leq 1$. This averaging leads to *coherence* of ES.

Closed-form solutions for ES are also available for some loss distributions.

1. Normal distribution: If $L_t \sim N(\mu_t, \sigma_t^2)$, then

$$ES_{1-p} = \mu_t + \frac{f(z_{1-p})}{p}\sigma_t$$

where $f(z)$ is the pdf of $N(0, 1)$ and z_{1-p} is the $(1 - p)$th quantile of $f(z)$.

2. Student-t_v loss distribution: If $(L_t - \mu_t)/\sigma_t$ follows a Student-t distribution with v degrees of freedom, then

$$ES_{1-p} = \mu_t + \sigma_t \frac{f_v(x_{1-p})}{p}\left(\frac{v + x_{1-p}^2}{v - 1}\right)$$

where $f_v(x)$ denotes the pdf of t_v, and x_{1-p} is the $(1 - p)$th quantile of $f_v(x)$.

3. Standardized Student-t_v loss. Assume $v > 2$. If $(L_t - \mu_t)/\sigma_t$ follows directly a standardized Student-t distribution with v degrees of freedom, then

$$ES_{1-p} = \mu_t + \sigma_t \frac{f_v^*(x_{1-p}^*)}{p} \left(\frac{(v-2) + (x_{1-p}^*)^2}{v-1} \right)$$

where $f_v^*(x)$ denotes the pdf of t_v^*, and x_{1-p}^* is the $(1-p)$th quantile of $f_v^*(x)$.

Calculation of VaR involves several factors:

1. Tail probability p: $p = 0.01$ for risk management and $p = 0.001$ in stress testing.
2. The time horizon ℓ : 1 day or 10 days for market risk and 1 year or 5 years for credit risk.
3. The CDF $F_\ell(x)$ or the quantiles of the loss random variable.
4. The amount of the financial position or the mark-to-market value of the portfolio.

We define the loss random variable as $x_t = \begin{cases} r_t, & \text{if the position is short,} \\ -r_t, & \text{if the position is long.} \end{cases}$

The dollar amount of VaR is then the cash value of the financial position times the VaR of the loss variable. That is, VaR = Value×VaR(x_t). And, the dollar amount of ES is the cash value of the financial position times the ES of the loss variable. That is, ES = Value×ES(x_t).

■ Methods for calculating financial risk

1. RiskMetrics
2. Econometric modeling
3. Quantile and quantile regression
4. Extreme value theory: traditional and Peaks over Thresholds

Demonstration: To illustrate the various methods for assessing financial risk, we consider the daily log returns of IBM stock from January 2, 2001 to December 31, 2010 for 2515 observations. See Figure 50. We assume a long position of 1 million on the stock. The loss $x_t = -r_t$.

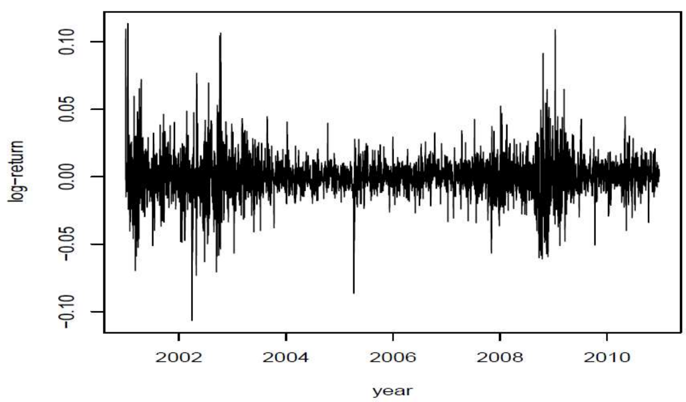

Figure 50: Daily log returns of IBM stock from January 2, 2001 to December 31, 2010.

RiskMetrics: J. P. Morgan developed the RiskMetrics methodology to VaR calculation; see Longerstaey and More (1995).

Let x_t denote the daily loss. RiskMetrics assumes that the loss, obtained from continuously compounded returns of the portfolio, follows a conditional normal distribution, i.e., $x_t|F_{t-1} \sim N(0, \sigma_t^2)$, where σ_t^2 follows the simple model:

$$\sigma_t^2 = \alpha\sigma_{t-1}^2 + (1-\alpha)x_{t-1}^2, \quad 0 < \alpha < 1. \tag{6.5}$$

Therefore, the log price $p_t = \ln(P_t)$ of a portfolio satisfies the difference equation $p_t - p_{t-1} = a_t$, $a_t = \sigma_t\epsilon_t$ where σ_t^2 is an IGARCH(1,1) process without drift. The value of α is often in the interval (0.9, 1) with a typical value of 0.94.

Riskmetrics assumes that

1. The loss function is symmetrically distribute, therefore VaR for long and short positions are the same.
2. $\mu_t = 0$.
3. σ_t^2 follows an IGARCH without drift.

Under this assumptions:

$$\begin{aligned} \mathrm{VaR}_{1-p}(1) &= z_{1-p}\sigma_{t+1}, \quad \text{For example } \mathrm{VaR}_{0.95}(1) = 1.65\sigma_{t+1} \\ \mathrm{ES}_{1-p}(1) &= \frac{f(z_{1-p})}{p}\sigma_{t+1}, \quad \text{For example } \mathrm{ES}_{0.95}(1) = \frac{f(1.65)}{0.05}\sigma_{t+1} \end{aligned}$$

Remark 1. A nice property of such a special IGARCH model is that the conditional distribution of a multi-period return is easily available. Specifically, for a k-period horizon, the log-return from time $t+1$ to time $t+k$ is $r_t[k] = r_{t+1} + ... + r_{t+k-1} + r_{t+k}$. Under the special IGARCH(1,1) model in Equation 6.5, the conditional distribution $r_t[k]|F_t$ is normal with mean zero and variance $\sigma_t^2[k] = Var(r_t[k]|F_t)$. Using the independence assumption of ϵ_t and Equation 6.5, we have

$$\sigma_t^2[k] = Var(r_t[k]|F_t) = \sum_{i=1}^{k} Var(r_{t+i}|F_t) = k \times Var(r_{t+1}|F_t) = k \times [\alpha\sigma_{t-1}^2 + (1-\alpha)x_{t-1}^2]$$

The last equality is because in IGARCH model without drift, all volatility forecasts are equal to $Var(r_{t+1}|F_t)$. Therefor, $r_t[k] \sim N(0, k\sigma_{t+1}^2)$, and the risk measures from time $t+1$ to time $t+k$ can be calculated as

$$\begin{aligned} \mathrm{VaR}_{1-p}[k] &= z_{1-p}\sqrt{k}\sigma_{t+1} \\ \mathrm{ES}_{1-p}[k] &= \frac{f(z_{1-p})}{p}\sqrt{k}\sigma_{t+1} \end{aligned}$$

Note that if σ_t^2 follows an IGARCH with drift, or $\mu \neq 0$, then $\mathrm{VaR}_{1-p}[k] \neq \sqrt{k}\mathrm{VaR}_{1-p}(1)$.

Example 2: Demonstration with IBM stock returns. We fit the special IGARCH(1,1) model of Equation (6.5) to obtain an estimate of the parameter α and obtain

$\hat{\alpha} = 0.943(0.007)$. In addition, using $x_{2515} = -0.00061$ and $\sigma_{2515} = 0.00734$, we have $\sigma_{2516} = 0.007133$. Consequently, using RiskMetrics, we have

$$\mathrm{VaR}_{0.95} = 0.01173, \quad \mathrm{VaR}_{0.99} = 0.01659, \quad \mathrm{ES}_{0.95} = 0.01471, \quad \mathrm{ES}_{0.99} = 0.01901,$$

Therefore,

$$\mathrm{VaR}_{0.95} = 0.01173 \times 1,000,000 = \$11,730$$

$$\mathrm{ES}_{0.95} = 0.01471 \times 1,000,000 = \$14,710$$

and

$$\mathrm{VaR}_{0.99} = 0.01659 \times 1,000,000 = \$16,590$$

$$\mathrm{ES}_{0.99} = 0.01901 \times 1,000,000 = \$19,010$$

Finally, for 15 days holding period, we have (square-root of time rule)

$$\mathrm{VaR}_{0.95}(15) = \sqrt{15} \times \$11,730 = \$45,430$$

$$\mathrm{ES}_{0.95}(15) = \sqrt{15} \times \$14,710 = \$56.972.$$

Remark 2. Consider multiple positions: Under RiskMetrics assumptions, for 2 assets, we have

$$\mathrm{VaR} = \sqrt{\mathrm{VaR}_1^2 + \mathrm{VaR}_2^2 + 2\rho_{12}\mathrm{VaR}_1\mathrm{VaR}_2}$$

This formula is obtained using the assumption that the joint distribution of the log returns of assets involved in the portfolio is multivariate normal ($(X, Y)' \sim N_2(\mu, \Sigma)$).

The generalization of VaR to a position consisting of m instruments is straightforward as

$$\mathrm{VaR} = \sqrt{\sum_{i=1}^{m} \mathrm{VaR}_i^2 + 2\sum_{i<j}^{m} \rho_{ij}\mathrm{VaR}_i\mathrm{VaR}_j}$$

where ρ_{ij} is the cross-correlation coefficient between returns of the ith and jth instruments and VaR_i is the VaR of the ith instrument.

A simple derivation for the prior formula for two assets

Basic setup: Two assets with log returns r_{1t} and r_{2t}. The portfolio consists of w_1 and w_2 amounts invested in asset 1 and asset 2, respectively.

Under RiskMetrics, we have

$$r_{1t}|F_{t-1} \sim N(0, \sigma_{1t}^2), \quad \sigma_{1t}^2 = \beta\sigma_{1t-1}^2 + (1-\beta)r_{1t-1}^2$$

$$r_{2t}|F_{t-1} \sim N(0, \sigma_{2t}^2), \quad \sigma_{2t}^2 = \beta\sigma_{2t-1}^2 + (1-\beta)r_{2t-1}^2$$

VaR for the two assets are $w_1\text{VaR}_1$ and $w_2\text{VaR}_2$, respectively. For instance, for tail probability 0.05, VaR for the first asset is $1.645 w_1\sigma_{1t}$. Let p_t be the log return of the portfolio. Then, we have

$$p_t \approx w r_{1t} + (1-w) r_{2t}$$

with $w = \frac{w_1}{w_1+w_2}$. The prior approximation becomes equality for simple returns.

Remark. $(w_1 + w_2)w = w_1$ and $(w_1 + w_2)(1 - w) = w_2$.

Under RiskMetrics, we have

$$p_t|F_{t-1} \sim N(0, \sigma_{pt}^2)$$

where

$$\sigma_{pt}^2 = \text{Var}(r_{pt}|F_{t-1}) = w^2\sigma_{1t}^2 + (1-w)^2\sigma_{2t}^2 + 2\rho w(1-w)\sigma_{1t}\sigma_{2t}$$

For tail probability of 0.05, we have $\text{VaR}_p = 1.645(w_1 + w_2)\sigma_{pt}$. Therefore, the square of VaR for the portfolio with tail probability 0.05 is

$$\begin{aligned}
(\text{VaR}_{pt})^2 &= 1.645^2(w_1 + w_2)^2\sigma_{pt}^2 \\
&= 1.645^2(w_1 + w_2)^2 \times [w^2\sigma_{1t}^2 + (1-w)^2\sigma_{2t}^2 + 2\rho w(1-w)\sigma_{1t}\sigma_{2t}] \\
&= 1.645^2[w_1^2\sigma_{1t}^2 + w_2^2\sigma_{2t}^2 + 2\rho w_1 w_2\sigma_{1t}\sigma_{2t}] \\
&= \text{VaR}_1^2 + \text{VaR}_2^2 + 2\rho\text{VaR}_1\text{VaR}_2
\end{aligned}$$

This is exactly similar to $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\rho \times \text{Std}(X)\text{Std}(Y)$.

1. The result can be generalized to more than two assets.
2. The formula continues to hold for expected shortfall provided that the mean returns of the two assets are zero. [Which is assumed under RiskMetrics.] If the means are not zero, then some adjustments are needed for the portfolio.

Remarks:

1. RiskMetrics assumes that log returns are normally distributed. Thus if the data does not have normal-like characteristics, the method would not work up to our expectations.
2. The loss function is assumed to be symmetrically distributed. However security returns tend to have fat tails.
3. The VaR obtained for long and short positions are the same.
4. We always assume $\mu = 0$ when using RiskMetrics. If not, then the time rule for VaR does not exist i.e $VaR(r_t\{k\}) = k\mu + 1.65\sqrt{k}\sigma_{t+1} \neq \sqrt{k}VaR(r_t)$.

Example 3: Consider a simple portfolio consisting of 40% in AAA bonds and 60% on IBM stock. The market value of the portfolio is U.S. \$1 million. To measure the bond returns, we employ the daily log-returns of the Bank of America Merrill Lynch U.S. Corp AAA total return index from January 2, 2001 to December 31, 2010. The data of bond index are obtained from the Federal Reserve Bank at St. Louis.

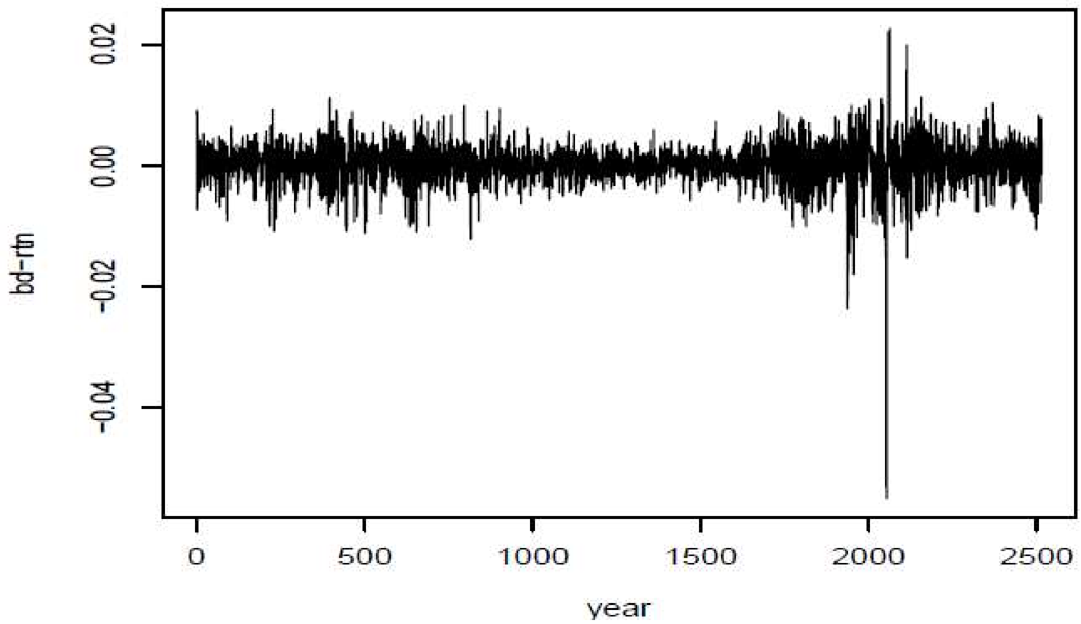

Figure 51: Daily log returns of bond index from January 2, 2001 to December 31, 2010. The bond index is the Bank of America Merrill Lynch U.S. Corp AAA total return index.

Figure 51 shows the log-returns of the bond index. Like stock returns, bond returns also exhibit the pattern of volatility clustering and weak stationarity. For bond returns,

$$\sigma_t^2 = 0.9577\sigma_{t-1}^2 + (1 - 0.9577)r_{t-1}^2$$

from which we have

$$\mathrm{VaR}_{0.95} = 0.00705, \quad \text{and} \quad \mathrm{VaR}_{0.99} = 0.00997$$

Recall, from Example 2 of the lecture notes, that for the daily log returns of IBM stock, we have

$$\mathrm{VaR}_{0.95} = 0.01173, \quad \text{and} \quad \mathrm{VaR}_{0.99} = 0.01659$$

The sample correlation coefficient of the log returns between IBM stock and AAA bond index is -0.2215. Consequently, for the portfolio we have

$$\mathrm{VaR}^e_{0.95} = 0.01173 \times 0.6 = 0.00704, \quad \text{and} \quad \mathrm{VaR}^b_{0.95} = 0.00705 \times 0.4 = 0.00282$$

where the superscripts e and b denote equity and bond returns, respectively. The $\mathrm{VaR}_{0.95}$ for the portfolio is then

$$\mathrm{VaR}_p = \sqrt{(\mathrm{VaR}^e)^2 + (\mathrm{VaR}^b)^2 + 2(-0.2215)\mathrm{VaR}^e\mathrm{VaR}^b} = 0.006978$$

For this particular instance, we see that with tail probability $p = 0.05$, the VaR of the portfolio is less than the VaR of each component. More specifically, with \$1 million investment, we have

1. Equity market only: $\mathrm{VaR}_{0.95} = \$11{,}730$.

2. Bond market only: $\text{VaR}_{0.95}$= \$7,050.
3. Portfolio (60-40): $\text{VaR}_{0.95}$= \$6,978.

This result is expected because VaR is a coherent risk measure under the normality assumption. The example, thus, demonstrates the value of diversification.

Discussion: RiskMetrics has several advantages

1. Simplicity: Normal distribution, square-root of time rule, and multiple assets (portfolio)
2. Transparency

It also has some serious weaknesses:

1. Assumed model is rejected by empirical data
2. The square-root of time rule fails if either of the model assumptions is rejected.

■ **Econometric Modeling:**

A general approach to VaR calculations is to use time series models that we went through in chapters 2 and 3. Upon building such a model, we can then forecast the parameters required and estimate for VaR.

First, we will have to consider a model for the loss, $\{x_t\}$, and a model for the conditional variance. We stick to GARCH for simplicity but other volatility models are possible choices to explore. We will thus estimate the parameters of the model given below:

$$x_t = \begin{cases} r_t, & \text{if short position} \\ -r_t, & \text{if long position} \end{cases}$$

where the model is an ARMA(p,q)-GARCH(m,s) model with innovations from distribution $D(0,1)$:

$$x_t = \phi_0 + \sum_{i=1}^{p} \phi_i x_{t-i} + a_t - \sum_{j=1}^{q} \theta_j a_{t-j}$$
$$a_t = \epsilon_t \sigma_t, \epsilon_t \sim D(0,1)$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

1. Estimate the model parameters.
2. Use one step ahead forecasts for x_t and σ_t^2 for calculating VaR_{t+1}.

Hence, depending on the density of the error terms, we can derive the one-period-ahead VaR and ES.For example given that the one step ahead forecasts for expected loss and volatility is:

$$\hat{x}_t(1) = E[x_{t+1}|F_t] = \phi_0 + \sum_{i=1}^{p} \phi_i x_{t-i}$$
$$\hat{\sigma}_t^2(1) = E[\sigma_{t+1}^2|F_t] = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

We have:

- If $\epsilon_t \sim N(0,1)$, then $x_{t+1}|F_t \ N(\hat{x}_t(1), \hat{\sigma}_t^2(1))$, and we will get:

$$VaR_{0.95} = \hat{x}_t(1) + \hat{\sigma}_t(1) \times 1.65$$
$$ES_{0.95} = \hat{x}_t(1) + \hat{\sigma}_t(1) \times \frac{f(1.65)}{0.05}$$

- If $\epsilon_t \sim t_\nu^*$:

$$VaR_{0.95} = \hat{x}_t(1) + \hat{\sigma}_t(1) \times t_\nu^*(0.95)$$
$$ES_{0.95} = \hat{x}_t(1) + \frac{f_\nu^*(t_\nu^*(0.95))}{0.05} \frac{\nu - 2 + (t_\nu^*(0.95))^2}{\nu - 1} \hat{\sigma}_t(1)$$

The approach considers modeling seriously, but it requires human intervention. Also, multiple-period risk measures become tedious.

Example: Consider again the daily log returns of IBM stock.

Model 1: A Gaussian GARCH(1,1) model. The fitted model is

$$x_t = -6.01 \times 10^{-4} + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0,1)$$

$$\sigma_t^2 = 4.378 \times 10^{-6} + 0.884\sigma_{t-1}^2 + 0.101a_{t-1}^2$$

The 1-step-ahead predictions at $T = 2515$ are -6.01×10^{-4} and 7.82×10^{-3}, respectively, for mean and volatility. Consequently, we have

$$\text{VaR}_{0.95} = 0.01227, \quad \text{and} \quad \text{ES}_{0.95} = 0.01554,$$

$$\text{VaR}_{0.99} = 0.01760, \quad \text{and} \quad \text{ES}_{0.99} = 0.02025.$$

These results imply $\text{VaR}_{0.95} = \$12,270$ and $\text{ES}_{0.95} = \$15,540$ for the next trading day.

Model 2: A GARCH(1,1) model with standardized Student-t innovations. The fitted model is

$$x_t = -4.113 \times 10^{-4} + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t^*_{5.751}$$

$$\sigma_t^2 = 1.992 \times 10^{-6} + 0.928\sigma_{t-1}^2 + 0.0645a_{t-1}^2$$

The 1-step ahead predictions at $T = 2515$ are -4.113×10^{-4} and 0.00801, respectively, for mean and volatility. Therefore, the risk measures for the financial position are $\text{VaR}_{0.95} = \$12,399.8$ and $\text{ES}_{0.95} = \$17,564$.

Using econometric approaches makes multiperiod risk measures tedious to find:

Suppose at time h (forecast horizon), we want to forecast the k-horizon VaR of an asset with log returns r_t, we have:

$$r_h[k] = r_{h+1} + r_{h+2} + \ldots + r_{h+k}$$

We need to get the conditional mean and variance of $r_h[k]$. Then, given the information F_h, we use the forecasting techniques in chapter 2 and chapter 3 for the mean and volatility respectively.

For $E[r_{h+\ell}|F_h]$, we can use the MA representation of our ARMA model i.e $r_t = \mu + a_t + \psi_1 a_{t-1} + \ldots$, to get:

$$\hat{r}_h(l) = E[r_{h+l}|F_h] = \mu + \psi_l a_h + \psi_{l+1} a_{h-1} + \ldots$$
$$e_h(l) = r_{h+l} - \hat{r}_h(l) = a_{h+l} + \psi_1 a_{h+l-1} + \ldots + \psi_{l-1} a_{h+1}$$

with $\psi_0 = 1$. Thus,

$$e_h[k] = e_h(1) + e_h(2) + \ldots + e_h(k)$$
$$\Rightarrow e_h[k] = a_{h+k} + (\psi_1 + 1)a_{h+k-1} + \ldots + (\sum_{i=0}^{k-1} \psi_i)a_{h+1} \ (\because \psi_0 = 1)$$
$$\Rightarrow Var(e_h[k]) = Var_h(a_{h+k}) + (\psi_1^2 + 1)Var_h(a_{h+k-1}) + \ldots + (\sum_{i=0}^{k-1} \psi_i^2)Var_h(a_{h+1})$$
$$= \hat{\sigma}_h^2(k) + (\psi_1^2 + 1)\hat{\sigma}_h^2(k-1) + \ldots + (\sum_{i=0}^{k-1} \psi_i^2)\hat{\sigma}_h^2(1)$$

We can use the methods in Chapter 3 to calculate each $\hat{\sigma}_h^2(l)$ for $l = 1, ...k$.

- **Empirical Quantiles** A non-parametric approach to VaR calculations. No specific distributional assumption on the loss function of a portfolio is imposed. The only assumption is that the distribution of the returns in the prediction period is the same as that in the sample period.

  Arrange $x_1, x_2, ..., x_n$ such that $x(1) \le x(2) \le ... \le x(n)$. Note that $x(1)$ is the minimum loss in the sample (minimum order statistic), and $x(n)$ is the maximum loss in the sample (maximum order statistic). Then qth empirical quantile is $x(nq)$. In practice, q may not satisfy that nq is a positive integer. In this case we can use the simple interpolation to obtain the empirical quantiles:

  If $\ell_1 < nq < \ell_2$, then define $q_1 = \frac{\ell_1}{n}$ and $q_2 = \frac{\ell_2}{n}$. The qth empirical quantile is

$$\frac{q_2 - q}{q_2 - q_1} x_(\ell_1) + \frac{q - q_1}{q_2 - q_1} x_(\ell_2)$$

  For instance, if $n = 825$, we would obtain $nq = 825 \times 0.95 = 783.75$. By using the above form, (and assuming $l_1 = 783$ and $l_2 = 784$, the qth empirical quantile will be $0.75x^{(784)} + 0.25x^{(783)}$.

  Most statistical software provides empirical quantiles for a given data set.

  Example: Consider the daily log returns of IBM stock. Since $2515 \times 0.95 = 2389.25$, we let $\ell_1 = 2389$ and $\ell_2 = 2390$. The empirical 95% quantile of the negative log returns can be obtained as

$$\hat{x}_{0.95} = (\ell_2 - 2389.25) \times x(\ell_1) + (2389.25 - \ell_1) \times x(\ell_2) = 0.75 \times x(2389) + 0.25 \times x(2390) = 0.02654$$

  $x(i)$ is the ith order statistic of the loss variable x_t. In this particular instance, $x(2389) = 0.02652$ and $x(2390) = 0.02657$. Finally, with $p = 0.05$, the sample expected shortfall is d $\text{ES}_{0.95} = \$39,949$ for the next trading day:

$$ES_{0.95} = \frac{1}{N_{0.95}} \sum_{i=1}^{T} x(i) \times \mathrm{I}[x(i) > \hat{x}_{0.95}]$$

  where $N_{0.95}$ is the numbers of $x(i)$ greater than $\hat{x}_{0.95}$, and $\mathrm{I}[x(i) > \hat{x}_{0.95}] = 1$ if $x(i) > \hat{x}_{0.95}$, otherwise $\mathrm{I}[x(i) > \hat{x}_{0.95}] = 0$.

Advantage: 1- Simple. 2- No restrictive assumption on the distribution except that it stays the same.

Disadvantages: 1- Fails to consider the effect of explanatory variables.2- When q is small, the empirical quantile is not an efficient estimate of the theoretical quantile.

- **Quantile Regression:** see Koenker and Bassett (1978).

In Ordinary Least Squares technique for estimating the parameters of linear regression models, we minimize the following function for β:

$$\sum_{t=1}^{T}(x_t - \beta z_t)^2$$

where $z_t = (z_{1t}, z_{2,t}, \ldots, z_{kt})$ is the vector of explanatory variables, $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$ is the coefficient vector and $x_t = \beta z_t + \epsilon_t$, where $\epsilon_t \sim N(0, \sigma_e^2)$.

After finding $\hat{\beta}$ that minimizes the equation, then our forecast or expected value of $x_{t+1}|z_{t+1}$ is

$$E[x_{T+1}|x_{1:T}, z_{1:T}, z_{T+1}] = \hat{\beta} z_{T+1}$$

If we are interested in finding other quantiles in general, we can still find a β^o such that:

$$x_q^{T+1} = \beta^0 z_{T+1}$$

Let's assume that $z_t \in F_{t-1}$ has some impact on x_t:

$$x_t = \beta' z_t + a_t$$

We can estimate the conditional quantile $x_q|F_{t-1}$ of x_t given F_{t-1} as

$$\hat{x}_q|F_{t-1} \equiv \inf\{\beta_0' z_t | R_q(\beta_0) = \min\} \tag{6.6}$$

where $R_q(\beta_0) = \min$ means that β_0 is obtained by

$$\beta_0 = arg\min_{\beta} \sum_{t=1}^{T} w_q(x_t - \beta' z_t)$$

and

$$w_q(y) = \begin{cases} qy & \text{if } y \geq 0 \\ \\ (q-1)y & \text{if } y < 0 \end{cases} \tag{6.7}$$

Quantile regression directly models a specific quantile rather than the whole return distribution.

Example: Again, consider the daily log returns of IBM stock. We employ a quantile regression with two predictors. The first predictor is the lag-1 daily volatility of the IBM stock and the second predictor is the lag-1 VIX index of Chicago Board Options Exchange (CBOE). More specifically, we consider the quantile regression

$$x_t = \beta_0 + \beta_1 s_{t-1} + \beta_2 \nu_{t-1} + e_t,\ e_t \sim N(0, \sigma_e^2) \tag{6.8}$$

where $x_t = -r_t$ with r_t being the daily log return of IBM stock, s_{t-1} is the lag-1 daily IBM stock volatility obtained from fitting a Gaussian GARCH(1,1) model to x_t, and ν_{t-1} is the lag-1 VIX index obtained from CBOE. Here we use the VIX index, not the percentage VIX.

Applying the quantile regression with $q = 0.95$, we obtain

$$\hat{\beta}_0 = -0.001(0.003), \qquad \hat{\beta}_1 = 1.17724(0.22268), \qquad \hat{\beta}_2 = 0.02809(0.01615)$$

where the number in parentheses denotes standard error. As expected the 95th quantile of the IBM negative daily log returns depends critically on the lag-1 IBM daily volatility and marginally on the lag-1 VIX index. Based on the model, we have $\hat{Q}(0.95|z_{2514}) = 0.013574$. This implies that $\mathrm{VaR}_{0.95} = \$13,574$ for the financial position. Figure 52 shows the negative IBM log returns $x_t = -r_t$ and the fitted values of the quantile regression with probability $q = 0.95$. The plot also shows that VaR is time-varying and highlights the fact that the actual loss may vary when the loss exceeds VaR. R package: Quantreg.

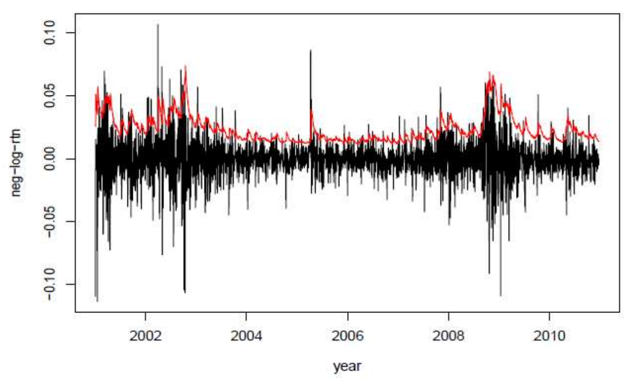

Figure 52: Time plot of the negative daily log-returns of IBM stock from January 3, 2001 to December 31, 2010. The upper line shows the 0.95th quantiles obtained by the quantile regression of Equation (6.8)