# Chapter 2

# Value at Risk and other risk measures

## 1   Motivation and basics

As we have already noted in the introduction, risk measurement based on proper risk measures is one of the fundamental pillars of the risk management. In this chapter, we will address in details the issue of such risk measures.

Let us start with a practical question. Imagine you have bought $1000. How risky is this investment? Putting differently, how large loss are you exposed to?

It seems that this question is not very well posed. In theory, at worst you lose everything – just imagine a gloomy scenario where the dollar collapses and becomes worthless. However, this information, even if absolutely rigorous, is not very useful for a risk manager.

To obtain a more practical answer, we can reformulate the question in a more specific way. In particular, we can estimate how much one can lose with probability e.g. 5 %. Alternatively, we can ask what is the expected loss given that financial markets enter into a hectic period which corresponds to a market drop scenario with a magnitude corresponding to an event which happens, e.g., once in 3 years. These questions lead us to a more precise definition of a risk measure.

The most widely used risk measure is Value at Risk (VaR). In the previous example, VaR refers to the loss that should not be exceeded with probability of 95 %. Generally, the definition of VaR is as follows:

**Definícia 1.** *Value at Risk refers to the maximum loss, which should not be exceeded during a specified period of time with a given probability level.*

We shall use the following notation: $\mathrm{VaR}_t^d(1 - \alpha)$ denotes the loss (taken with a positive sign) which should not be exceeded during the period of $d$ business days with

probability $1 - \alpha$. Saying differently,

$$P(X^d \leq -\text{VaR}_t^d(1-\alpha)) = \alpha,$$

where $X^d$ denotes the change in the portfolio value during the period of $d$ business days. In typical examples, we often consider $d = 10$ days and $\alpha = 5\%$ or $\alpha = 1\%$ ($\alpha$ is the probability of a loss exceeding VaR).

In addition, VaR can be expressed as the amount of loss, but is also often expressed as a relative loss (in %). To distinguish these two cases, let us use the notation $\overline{\text{VaR}}$ in the latter.

How to calculate VaR? Obviously, it depends on the volatility of value of the investment. More precisely, it is based on the distribution of its profit and loss (P& L).

Assume for now that the P& L distribution during a period of $d$ business days is known. Then $\text{VaR}_t^d(1 - \alpha)$ can be calculated as the $\alpha$ quantile of this distribution (taken with the opposite, i.e. positive sign). In particular,

$$\text{VaR}_t^d(1 - \alpha) = -\sup\{x \in \mathbb{R} \,|\, P(X^d \leq x) \leq \alpha\}.^1$$

In addition, if $f(x)$ is the probability density function of this distribution, one has

$$\int_{-\infty}^{-\text{VaR}_t^d(1-\alpha)} f(x)\,dx = \alpha.$$

Roughly speaking, 1-day VaR represents the maximum loss that your portfolio can suffer if the next day is a "bad day" for financial markets (which occurs in average once in each $1/\alpha$ business days). For example, if $\alpha = 5\%$, then 1-day VaR corresponds to the loss occurring once in $1/0.05 = 20$ business days (i.e. approximately once per month).

## 2   Methods of the VaR calculation

As we have already indicated, the most crucial part of the VaR calculation is the estimation of the distribution of the changes in the risk factors.

In general there are two possible categories of the methods for VaR calculation:

---

[1]This definition might seem slightly difficult to understand. To understand it better, consider an asset with the following distribution of its 1-day returns:

- -10 % with probability 2 % (i.e. $P(X^1 = -0.10) = 0.02$),

- - 5 % with probability 8 % (i.e. $P(X^1 = -0.05) = 0.08$),

- + 2 % with probability 90 % (i.e. $P(X^1 = 0.02) = 0.90$).

Hence $\text{VaR}_t^1(0.95) = 0.05$, since $P(X^1 < -0.05) = 0.02$ and $P(X^1 \leq -0.05) = 0.02 + 0.08 = 0.1$.

(i) **parametric method** – this method is based on a parametric class of distribution or, alternatively, on a model where some parameters need to be estimated,

(ii) **non-parametric method** – this method is based on the empirical distribution.

A combination of these two approaches (so called **semi-parametric methods**) can also be used.

These methods also differ in relation to the assumption of the stationarity of the distribution of the underlying risk factors. From the stylised facts about the financial data we already know that this distribution tends to be non-stationary. However, the calculation based on the assumption of stationarity may still be relevant, for example when we need a VaR estimation which is stable over a longer time horizon. Hence, we can distinguish two types of VaR calculations:

(i) unconditional VaR – based on the unconditional distribution of risk factors,

(ii) conditional VaR – based on the time-varying distribution.

Let us now introduce the most common calculation methods mentioned in the literature.

## 2.1 Historical simulation

**Calculation – simple case (a single equity)**

This is a non-parametric method. Let us illustrate this method on a simple example. Imagine that I have bought a share of Microsoft which is now worth $S_T$. Let us denote the relative returns by $r_t$, i.e.

$$r_t = \ln \frac{S_t}{S_{t-1}}.$$

Now assume that a sample of past $T$ realizations is available: $\{r_1, \ldots, r_T\}$. Define a window size $n$ business days (for one-year window, we can choose $n = 250$). We have a subsample of past $n$ observations $\{r_{T-n+1}, \ldots, r_T\}$. This subsample is assumed to approximate the cumulative distribution function. Hence, $\overline{\text{VaR}}_T^1(1-\alpha)$ is equal to the $\alpha$-quantile of this empirical distribution (taken with opposite sign).

Let us note then despite this simple calculation method, the determination of the $\alpha$-quantile might not be implemented consistently. For example, one may treat the sorted values of changes as $0.5/n, 1.5/n, \ldots, (n-0.5)/n$ quantiles (observed empirical quantiles). Then the calculation of the $\alpha$-quantile (in case that it does not coincidences with one of these values) is based on the linear interpolation. This method is implemented e.g. in Matlab.
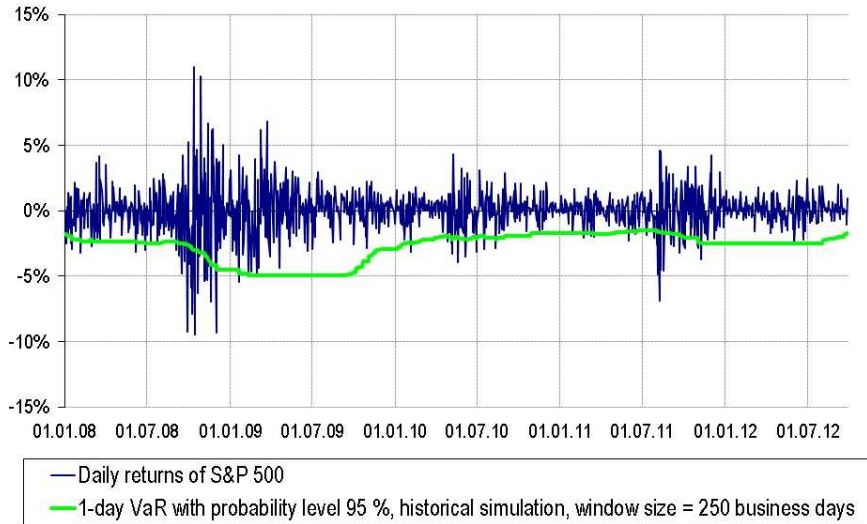
Figure 2.1: $\overline{\text{VaR}}_t^1(95\%)$ calculated with historical simulation.

To illustrate this calculation, let us take $\alpha = 2\%$ and the window size $n = 250$ business days. Then the observed empirical quantiles correspond to the level of probability of $0.002, 0.006, 0.010, 0.014, 0.018, 0.022 \ldots \ldots, 0.998$. Hence the $\alpha$-quantile is determined as the value in the middle between the $5^{th}$ and $6^{th}$ largest loss.

Alternatively, one may argue that the distribution between the observed empirical quantiles might not be linear which may in the worst case lead to an underestimation of the risk. Hence, we can choose to apply a more prudent method and take the closest worse value of the loss. In the previous example, this would correspond to the $5^{th}$ largest loss. Note that in this case, this approach yields the same result as simply taking the $(\alpha n)$-th largest loss which is a common approach as well.

**Advantages and disadvantages**

The main advantage of this method is the simple calculation which does not require any specific assumption on the class of the underlying distribution. However, the main issue is the assumption of stationarity (at least, we assume that the distribution on the next day will be the same as in the last $n$ business days). This assumption might turn to be very misleading, mainly in the case when a period of market turbulence appears after a long period of low volatility. In such a case, VaR calculated by a historical simulation only slowly adapts to the new conditions and tends to underestimate the risk for some period.

Care should be taken when determining the window size ($n$). If this window is too long, VaR is based on very old historical values which may not adequately represent the current situation on the market. In addition, the adjustment to any changes (e.g. in volatility) is very slow. On the hand, if the window is too short, the calculation might

not be sufficiently robust, especially for quantiles with a low probability level.

### Calculation – multiple and non-linear assets

The previous case of a single-equity portfolio was very simplistic since it only concerned one asset whose price directly coincided with the underlying risk factor. Let us now turn to a more complex case of the VaR calculation based on the historical simulation method.

Let us assume that we have a portfolio of $N$ financial instruments (equities, bonds, derivatives, ...). Let us denote the price of $i$-th financial instrument at day $t$ as $P_t^i$ ($i = 1, \ldots, N$, $t = T - n + 1, \ldots, T$, where $T = $ today). The price $P_t^i$ depends on risk factors (equity prices, FX rates, interest rates...) which may or may not be common to a subset of the financial instruments:

$$P_t^i = f^i(t, R_t^{i,1}, \ldots, R_t^{i,M_i})$$

where $f^i$ is the pricing function for the $i$-th asset.

The value of the portfolio at time $T$ is

$$P_T = \sum_{i=1}^{N} w_i P_T^i,$$

where the weight $w_i$ is the share of each instrument in the portfolio at $T$.

The general methodology of the historical simulation consists of the following steps:

1. Calculate daily relative changes for each risk factor:

$$r_t^{i,j} = \ln \frac{R_t^{i,j}}{R_{t-1}^{i,j}}.$$

2. Calculate "hypothetical" scenarios for tomorrow values of all risk factors based on the relative changes in the past:

$$\tilde{R}_{T+1,s}^{i,j} := R_T^{i,j} \, r_{T-s}^{i,j}, \qquad s = 0, \ldots, n - 2. \tag{2.1}$$

3. For each scenario, calculate "hypothetical" tomorrow prices of each financial instrument

$$\tilde{P}_{T+1,s}^i = f^i \left( T + 1, \tilde{R}_{T+1,s}^{i,1}, \ldots, \tilde{R}_{T+1,s}^{i,M_i} \right).$$

4. For each scenario, calculate "hypothetical" tomorrow value of whole portfolio as a weighted average of the prices of individual instruments:

$$P_{T+1,s} = \sum_{i=1}^{N} w_i P_{T+1,s}^i$$

5. For each scenario, calculate the 1-day change in the price of the portfolio:

$$\frac{P_{T+1,s}}{P_T}.$$

These values represent the P&L distribution which should be used to calculate the respective quantile.

Note that this methodology implies one additional **advantage** of the historical simulation method when applied to a portfolio of several assets: We do not need to apply any explicit treatment of correlations between risk factors. These correlations are implicitly captured by using the empirical multivariate distribution of the changes in risk factors.

## 2.2   Parametric method – stationary distributions

**Calculation – simple case (a single equity)**

Parametric methods are based on a specific assumption about the underlying distribution of profits and losses. To start with a simple example, let us assume that the returns of equity prices for Microsoft follow a normal distribution. In addition, let us denote the estimated values of the mean and the standard deviation of this distribution by $\hat{\mu}$ and $\hat{\sigma}$, respectively. Then, VaR can be easily calculated as the $\alpha$-quantile of this distribution, i.e.

$$\text{VaR}_t^1(1 - \alpha) = -(\hat{\mu} + \phi^{-1}(\alpha)\hat{\sigma}), \tag{2.2}$$

where $\phi^{-1}(\alpha)$ is the $\alpha$-quantile of the standard normal distribution $N(0,1)$.

Again, the estimated values $\hat{\mu}$ and $\hat{\sigma}$ are computed from a window of size $n$, which is an important parameter for this method.

Another possibility is to use the Student t distribution. In this case, the formula for VaR calculation is

$$\overline{\text{VaR}}_t^1(1 - \alpha) = -\left(\hat{\mu} + t_\nu^{-1}(\alpha)\sqrt{\frac{\nu - 2}{\nu}}\,\hat{\sigma}\right),$$

where $\nu$ denotes degrees of freedom and $t_\nu^{-1}$ is the $\alpha$-quantile of the standard Student t distribution with $\nu$ degrees of freedom. Note that in this case, the estimated value of the standard deviation $\hat{\sigma}$ has to be adjusted. The reason is that the standard Student t distribution has standard deviation equal to $\sqrt{\frac{\nu}{\nu-2}}$ which is already "included" in the quantile.
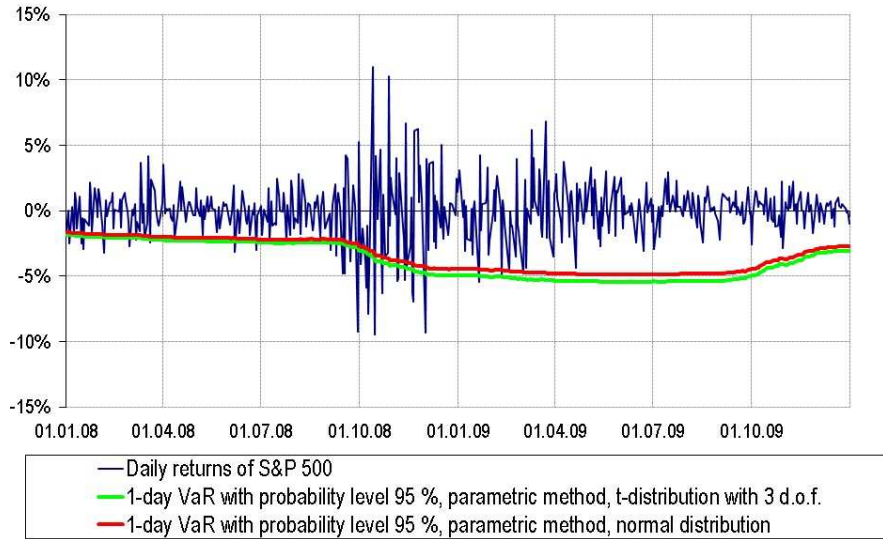
Figure 2.2: $\overline{\text{VaR}}_t^1(95\%)$ calculated with parametric method.

## Calculation – multiple linear assets

Let us now turn to the calculation of the VaR for portfolio of multiple assets. For the time being, assume that all the assets have linear pricing functions (or at least the pricing functions might be approximated by their linear approximations).

To illustrate the method, assume that the relative changes of risk factors for the next day follow a multivariate normal distribution, i.e.

$$\boldsymbol{r}_{T+1} \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

where $\boldsymbol{r}_{T+1}$ denotes the random vector of relative changes of risk factors.

Using the assumption of the linearity of the pricing function, the value of the portfolio can be written as

$$\frac{P_T}{P_{T-1}} = \boldsymbol{\beta}^\top \boldsymbol{r}_T,$$

where $\boldsymbol{\beta}$ is the vector of the coefficients. Hence one has

$$\frac{P_{T+1}}{P_T} \sim N(\boldsymbol{\beta}^\top \boldsymbol{\mu}, \boldsymbol{\beta}^\top \boldsymbol{\Sigma} \boldsymbol{\beta}).$$

This is univariate normal distribution, and VaR can be easily calculated as

$$\overline{\text{VaR}}_t^1(1-\alpha) = -\left(\boldsymbol{\beta}^\top \hat{\boldsymbol{\mu}} + \phi^{-1}(\alpha)\sqrt{\boldsymbol{\beta}^\top \hat{\boldsymbol{\Sigma}} \boldsymbol{\beta}}\right).$$

The method that has been described in this subsection is is often called **variance-covariance method**.

**Calculation – multiple non-linear assets**

In general, the assumption about the linear pricing functions needs to be relaxed. In this case, we can still model the changes in risk factors as a random vector from the multivariate normal distribution. However, the analytical formula similar to the one given in the previous subsection might not exist. Hence, we employ the procedure which has been described in the case of historical simulation. However, the scenarios are generated from the multivariate normal distribution instead of past realizations. Equation (2.1) has to be replaced by

$$\tilde{R}^{i,j}_{T+1,s} := R^{i,j}_T \, r^{i,j}_s, \qquad s = 0, \dots, n-2,$$

where the $k$ random vectors $(r^1_s, \dots, r^M_s)$, $s = 0, \dots, k-1$ represents independent draws from the multivariate normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $M$ denotes the total number of risk factors.

How to simulate random vectors from the multivariate normal distribution? We can apply the following procedure:

1. Generate a random vector $\boldsymbol{x}$ from $N(\boldsymbol{0}, \boldsymbol{I})$.

2. Compute the Cholesky decomposition of the variance-covariance matrix $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^\top,$$

   where $\boldsymbol{A}$ is a lower triangular matrix.

3. Vector $\boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{x}$ has then the distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Indeed, one has

$$\mathrm{Cov}(\boldsymbol{y}) = \boldsymbol{A}\mathrm{Cov}(\boldsymbol{x})\boldsymbol{A}^\top = \boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{\Sigma}.$$

   This method is an example of a so-called Monte Carlo method.

**Advantages and disadvantages**

The main advantage of this methods is apparently that the computation is very simple. On the other hand, the assumption of normal distribution is unrealistic. In addition, the slow response in case that the market enters into a turbulent period is again an issue. Furthermore, the calculation of the mean return is not robust to the choice of a time window, since it only depends on the first and last observations.

## 2.3   Parametric method – non-stationary distributions

**Calculation – simple case (a single equity)**

As it has been noted in the introductory lectures, a stationary distribution with a constant volatility does not correspond to the empirical characteristics of the data. In particular,

the real data tend to exhibit a significant level of heteroskedasticity, mainly in the form of so-called volatility clustering.

Hence, one can propose to model the time-varying volatility using so-called Exponentially weighted moving average (EWMA) which is a method proposed by Riskmetrics. In this case, the volatility is assumed to follow an AR(1) process:

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1-\lambda)(r_t - \mu)^2,$$

where $\mu$ is the unconditional mean and $\lambda$ is an arbitrary constant from $(0, 1)$.

Let us note that this formula can be derived as follows:

The historical volatility is estimated as

$$\sigma_t^2 = \frac{1}{n} \sum_{i=t-n+1}^{t} (r_i - \mu)^2.$$

This formula imposes equal weights $(1/n)$ for the last $n$ observations. In order to have the volatility more time-varying, we may increase the weight of the latest observation and use exponentially decreasing weights for other observations. Hence, we modify the calculation as follows:

$$\sigma_t^2 = (1-\lambda) \sum_{i=-\infty}^{t} \lambda^{t-i}(r_i - \mu)^2 = (1-\lambda)(r_t - \mu)^2 + (1-\lambda) \sum_{i=-\infty}^{t-1} \lambda^{t-i}(r_i - \mu)^2 =$$

$$= (1-\lambda)(r_t - \mu)^2 + \lambda \left[ (1-\lambda) \sum_{i=-\infty}^{t-1} \lambda^{t-i-1}(r_i - \mu)^2 \right] = (1-\lambda)(r_t - \mu)^2 + \lambda \sigma_{t-1}^2,$$

which is the formula for EWMA-based volatility. Note that the scaling factor $(1-\lambda)$ is needed in order to have the weights summing up to 1.

In practical application, the value of $\lambda$ is set to 0.94, which is the value proposed by Riskmetrics as well. However, it would be clearly preferential to estimate this value directly from the data (e.g. using maximum likelihood method). This is the basis for so-called GARCH (=generalized autoregressive conditional heteroskedasticity) models. These models have been proposed by Robert F. Engle, a winner of the 2003 Nobel Memorial Prize in Economic Sciences.

The GARCH(1,1) model has the following form:

$$r_t = c + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \alpha \varepsilon_{t-1}^2.$$

This model can be estimated using the maximum likelihood method. However, when estimating this model, we have to ensure that the estimated volatility stays positive. Clearly a sufficient condition for this is that the parameters in the volatility equation are non-negative. In addition, it is required that the estimation of the unconditional
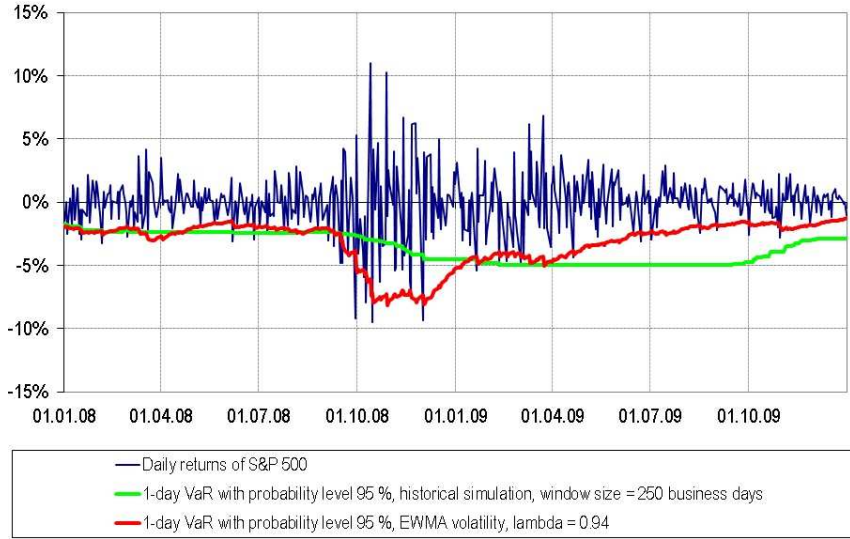
Figure 2.3: $\overline{\mathrm{VaR}}_t^1(95\%)$ calculation based on EWMA / GARCH model.

volatility

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta}$$

is finite.

VaR is then easily calculated as

$$\overline{\mathrm{VaR}}_t^1(1 - \alpha) = -(\hat{c} + \phi^{-1}(\alpha)\hat{\sigma}_t).$$

Note that the GARCH(1,1) model can be generalized to the GARCH($q$,$p$) model as follows:

$$r_t = c + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2.$$

**Multivariate GARCH models (MGARCH)**

In case of a portfolio with several risk factors, we have to apply so called multivariate GARCH model.

Clearly, the most general approach to the multivariate extension of the GARCH model is that each element of the covariance matrix is modeled separately as an AR (1) process. However, this approach has two major drawbacks. First, there are too many

parameters. Second (and even more importantly), the covariance matrix might not be positively semi-definite.

Hence, another form of the model has been proposed. It is so-called BEKK-GARCH(1,1) model:

$$\Sigma_t = C^\top C + A^\top \Sigma_{t-1} A + B^\top \varepsilon_{t-1} \varepsilon_{t-1}^\top B,$$

where $A$, $B$ and $C$ are $N \times N$ square matrices of parameters and $C$ is an upper triangular matrix. By the construction, the model ensures that the covariance matrix is always positively semi-definite.

The VaR calculation is a straightforward analogy to the stationary case (either by an explicit formula in case of linear pricing functions or via Monte Carlo simulations otherwise).

There is also one other type of multivariate GARCH models which is often used. It is so-called CCC-GARCH model (=constant conditional correlation). In this model, it is assumed that volatilities are time-dependent but correlations between factors remain constant.

CCC-GARCH model has the following form:

$$r_t = c + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \Sigma_t),$$

$$\Sigma_t = \mathrm{diag}(\sigma_t) R \, \mathrm{diag}(\sigma_t),$$

$$(\sigma_t^i)^2 = \omega_i + \beta_i (\sigma_{t-1}^i)^2 + \alpha_i (\varepsilon_{t-1}^i)^2,$$

where $R$ is the (stationary) correlation matrix. This matrix is computed as the (unconditional) variance-covariance matrix of standardised residuals.

Compared to the BEKK-GARCH model, CCC-GARCH has considerably lower number of parameters, whereas it still preserves the positively semi-definiteness of the covariance matrix.

## 2.4 Implied volatility

Implied volatilities might be used as an alternative method for the model-based time-varying volatility. The main advantage of the implied volatility is that it is based on market data and represents market-based estimates of future risk. It does not suffer from the inherent inertia unlike any metric based on historical data.

Implied volatility of an asset is calculated from derivative prices (e.g. from options using an inverse of the Black-Scholes formula).

However, it should be noted that this method has some drawbacks as well:

- The value of an implied volatility is not uniform across different strike prices and maturities (due to the existence of volatility smile / smirk).

- Implied volatilities might not be sufficiently reliable in case of a low liquidity in derivative market.

- In case of a multiple-asset portfolio, an estimation of correlation is needed as well. However, unlike implied volatilities, the existence of derivative instruments for calculating implied correlations is scarce.

## 2.5   Transforming 1-day VaR to $d$-day VaR

So far, we have focused on calculating of VaR with 1-day holding period (i.e. $\text{VaR}^1$). However, in many practical application longer holding period is required. The main reason is that it is assumed that in downturn markets where the liquidity is hampered, it might take several days to liquidate a risky position. Therefore a practical question arise – how to calculate $d$-day VaR?

One possibility is to take $d$-day returns instead of 1-day returns. However, one needs to work only with non-overlapping returns, otherwise a problem of autocorrelation arises. However, such an approach significantly reduces the number of data. For instance, if we calculate 10-day VaR based on data from 250 business days, we would be left with 25 data only.

Therefore, it is a common practice to calculate 1-day VaR and to scale it up to $d$-day VaR with a scaling factor. Most often, this scaling factor is set to $\sqrt{d}$. However it should be noted, that despite a wide usage of this scaling factors, it is only correct when the returns follow geometric random walk (i.e. based on normal distribution) with zero mean. We will address this issue later when discussing the extreme value theory.

# 3   Usage of VaR

VaR is mainly used for the following purposes:

1. **Determination of minimum capital requirement.** VaR can be interpreted as the minimum amount of capital to cover losses with a pre-defined probability level. The requirement to hold this capital is set by the internal management of the banks, by supervisory authorities as well as by expectations of investors and credit rating agencies.

   The minimum required capital (MRC) for regulatory purposes is calculated in accordance with Article 364 of the Regulation (EU) No 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms as follows:

$$MRC = \max \left\{ \text{VaR}_t^{10}(99\%), \frac{m_c}{60} \sum_{i=1}^{60} \text{VaR}_{t-i}^{10}(99\%) \right\} +$$

$$+ \max \left\{ \text{sVaR}_t^{10}(99\%), \frac{m_s}{60} \sum_{i=1}^{60} \text{sVaR}_{t-i}^{10}(99\%) \right\}, \tag{2.3}$$

where $m_c$ and $m_s$ are bank-specific multiplicators prescribed by the regulator based on the model quality (between 3 and 4) and sVaR is so-called stressed VaR. Stressed VaR is VaR calculated for the current portfolio with inputs calibrated to historical data from a continuous 1-year period of significant financial stress relevant to the institution´s own portfolio.

2. **Management tool.** VaR is used in internal limit systems in the asset management, e.g. limits set for the individual dealers. When deciding on the asset allocation, institutions can decide on the basis of comparing the marginal increase in VaR.

3. **Insurance premia** can be viewed as measure of riskiness of insured. It might be based on VaR. claims.

4. **Margin calls** used by central clearing counterparties in derivative transactions are also based on VaR.

# 4 Setting of parameters

As it is clear from the previous discussion, there are several parameters of the VaR model which need to be set ex-ante. This section provides some more insights into this issue.

1. **Confidence level** $(1 - \alpha)$ depends mainly on how VaR is going to be used. For a daily monitoring of the portfolio and setting limits for traders, the level of probability of 95 % is often used. On the other hand, when VaR is used for calculating the capital buffer, supervisory authorities set a higher level, usually 99 %. Rating authorities require an even higher level (which might exceed 99.9 %) in order to obtain a high credit rating. These differences in values are caused by the trade-off between robustness and safety. Apparently, it would be safer for the institution to set the confidence level very high and hence take into account also losses with a very low probability. On the other hand, the calculation of VaR with such high confidence levels might not be sufficiently robust, because the calculation is then based only on a handful of extreme observations. This is one of the reason why regulators require banks to use the probability level of 99 %, but then scale-up the result in order to calculate the minimum capital requirement.

2. **Holding period** (number of days) depends mainly on the liquidity of the market. The standard requirement of 10 business days is based on the assumption that it might take up to 10 business days (i.e. two weeks) for a bank to liquidate its position when it starts incurring losses. However, there are several other factors which might influence the setting of this parameter, e.g.:

   - Portfolio and risk type – the number of days is usually lower for market risk and higher (even up to one year) when credit risk prevails,
   - VaR usage – regulators require 10 days, but for setting internal limits for dealers shorter horizon might be considered in order to have the calculation more robust,
   - The extend of using linear approximations in pricing functions – where the portfolio contains a high share of non-linear instruments, it is better to keep the horizon shorter, otherwise the calculation would yield incorrect results.

3. **Positions** of individual financial instruments should contain all positions; short positions are taken with negative sign. It is important to include all contingent positions (e.g. derivatives or other contractual commitments) as well.

4. **Risk factors** needs to be properly identified, including all fundamental risk factors (interest rates, FX rates, equity prices) as well as the risk factors for specific risk (e.g. credit spreads).

5. **Correlations** between risk factors need not to be neglected. It should be taken into account that these correlations are not stationary and they are rising sharply in the period of high stress. The treatment of correlations might differ in individual calculation methods: In case of the historical simulation, correlations are included implicitly (non-parametric method). However, modelling of correlations might be challenging in other methods. In case that the portfolio only includes a low number of risk factors, M-GARCH models might be used. In case of high number of risk factors, some dimension-reduction techniques need to be applied (e.g. principal component analysis, which will be explained later in this course).

# 5  Critique of VaR

There are several major shortcomings which need to be taken into account when using VaR:

- **Model risk** should be taken into account in a proper manner. VaR calculation is in particular exposed to an incorrect model of statistical distribution of risk factor changes as well as to incorrect pricing functions (using invalid assumptions, or neglecting risk factors). In addition, missing, unreliable and improper data might also pose a problem.

- **Market liquidity risk** is neglected. One of the major assumption of VaR models is that a risky position will be fully liquidated during a period of several days at the prevailing market prices. However, market turmoils are often interlinked with low market liquidity. This is mainly caused by herding behavior of market participants and the fact that the general mood on the market changes significantly (from bullish to bearish). In addition, a negative impact of signal risk should be taken into account – when one major participant starts closing his positions, others assume that there is a problem with that asset and start to close their positions as well.

- **VaR is only a quantile.** Actually, VaR represents the **minimum** loss given we are in a market turmoil. However, we do not have any information about the magnitude of losses exceeding VaR. Hence, market participants which are only limited by VaR have incentives to be exposed to a high-risk portfolio structure with huge losses with probability beyond the confidential level.[2]

- **VaR is not a subaditive risk measure.** It would be reasonable to expect that a proper risk measure satisfies a condition that the sum of risks of two individual parts of a portfolio does not exceed the risk of the whole portfolio (effect of the diversification). However, this might not be true for VaR.

  To illustrate this issue, consider the following example: We have 2 assets ($X$ and $Y$) with return 0 % with probability of 99,1 % and Ű 10 % with probability of 0,9 %, respectively. Hence, VaR(99 %) for each asset taken separately is 0. The combined portfolio ($X + Y$) exhibits the following loss distribution:

  – 20 % (probability of 0,0081 %)

  – 10 % (probability of 1,7838 % = 2*0,009*0,991)

  – 0 % (probability of 98,2081 %)

  Hence, VaR(99 %) for this combined portfolio is 10.

  In practical situation, the subaditivity problems matters mainly in the following cases:

  - The marginal univariate distributions are skewed or have fat tails.

  - Portfolio of assets exhibits high tail dependence, even in case of symmetric univariate distributions.

It is interesting to note that despite this important drawback, VaR has become the major risk measure in the worldwide financial industry. The massive use of the VaR started in early nineties and its development was most extensive at J. P. Morgan. In 1999, the Basel Committee on Banking Supervision included VaR in the Basel Capital

---

[2]For more details, see e.g. the dissertation thesis by Csilla Krommerová.

Accord and hence it became a standard regulatory measure of market risk. In 2012 the same Committee issued a consultative document where it identified major shortcomings of VaR. Nevertheless, VaR is still the main pillar of the market risk regulation and is widely used in the industry.

# 6   Expected shortfall as an alternative risk measure

## 6.1   Definition

Some of the shortcomings of VaR identified in the previous section might be overcome when using another type of risk measure – **the expected shortfall**. Expected shortfall is defined as the value of the expected loss *given that* the loss is worse than VaR, i.e.

$$ES_t^d(1-\alpha) = -E\left[r_t^d \mid r_t^d \leq -\text{VaR}_t^d(1-\alpha)\right]$$

In addition, if $f(x)$ is the probability density function of the distribution of $r^d$, one has

$$ES_t^d(1-\alpha) = \int_{-\infty}^{-\text{VaR}_t^d(1-\alpha)} x f(x) \, dx.$$

The name of this measure is not unique across literature. The other names that are used are conditional value at risk (CVaR), average value at risk (AVaR), and expected tail loss (ETL). However, I would advise you to always check the exact definition when working with a risk measure.

## 6.2   Calculation

The expected shortfall might be easily calculated by the methods introduced above where only their slight adjustment is needed.

1. **Historical simulation.** When calculating VaR using the historical simulation, it is equal to the $\alpha$-quantile of the empirical distribution. As the straightforward extension, the expect shorftall is then equal to the average of the losses not exceeding this quantile (taken with positive sign).

2. **Linear parametric method** – the case of a univariate normal distribution: By denoting 1-day returns by $r_t$, expected loss (omitting the upper index) is given by

$$ES_t(1-\alpha) = -E[r_t \mid r_t \leq -\text{VaR}_t(1-\alpha)] =$$

$$= -\sigma E\left[\frac{r_t - \mu}{\sigma} \,\Big|\, \frac{r_t - \mu}{\sigma} \le \Phi^{-1}(\alpha)\right] - \mu = -\mu - \sigma \int_{-\infty}^{\Phi^{-1}(\alpha)} x\, \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx =$$

$$= -\mu + \sigma \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right]_{-\infty}^{\Phi^{-1}(\alpha)} = -\mu + \sigma \left[\phi(x)\right]_{-\infty}^{\Phi^{-1}(\alpha)} = -\mu + \sigma\phi(\Phi^{-1}(\alpha))$$

It is worth noting that expected shortfall can be equivalently calculated as follows:

$$ES_t(1-\alpha) = -\frac{1}{\alpha}\int_0^\alpha VaR_t(1-u)du.$$

*Sketch of the proof:*

$$ES_t(1-\alpha) = -E[r_t \,|\, r_t \le -\text{VaR}_t(1-\alpha)] = \frac{1}{\alpha}E\left[r_t \chi_{r_t \le -\text{VaR}_t(1-\alpha)}\right] =$$

$$= \frac{1}{\alpha}E\left[F^{-1}(u)\chi_{u\le\alpha}\right] = \frac{1}{\alpha}\int_0^\alpha F^{-1}(u)\,du = \frac{1}{\alpha}\int_0^\alpha -VaR_t(1-u)\,du.$$

$\chi$ is an indicator function.

The third equality is based on the following result:

Let $F$ be the distribution function of $r_t$. Assume for now that it is strictly increasing. If $u$ is random variable with a standard uniform distribution, then the random variable $y := F^{-1}(u)$ has distribution function as $r_t$. Indeed, one has

$$P(y \le r_t) = P(F^{-1}(u) \le r_t) = P(u \le F(r_t)) = F(r_t).$$

In addition, the last equality is based on the fact that VaR is the value of the loss corresponding to the $\alpha$-quantile.

# 7   Backtesting

Backtesting is the assessment of the quality of the VaR model. This assessment is based on the analysis of cases when the actual loss in the past observations was higher than VaR (exceedances or failures).

There are two important properties which should be met by an adequate VaR model:

1. **Unconditional coverage property:** The number of exceedances should correspond to the probability level of the VaR model. For example, if $\alpha = 5\%$, the

share of exceedances recorded in the past observations should be approximately 5 %. If this share is higher, the VaR model tends to underestimate the risk. In the opposite case, the VaR model is too conservative.

2. **Independence property:** Besides the correct number of exceedances, it is also important that they should be distributed randomly accross the sample of past observations. Any clustering of exceedances (e.g. in case of a market turmoil) points to deficiensies in the VaR model. In other words, the occurances of exceedances should not be autocorrelated.

The backtesting of VaR model is important first of all for the internal management of the institutions. They should ensure that their VaR figures are calculted properly. In addition, the backtesting is also interesting for regulators, e.g. as a part of the assessment of a bankŠs application for internal model approval when the bank intends to apply the advanced approach for the calculation of regulatory capital requirement for market risk.

## 7.1   Regulatory test

The regulatory test is actually not a proper statistical test. Instead, it is a very simple rule which provides regulators with a guidance on the quality of the VaR model. The regulatory test is specified in the Article 366 of the Regulation (EU) No 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms.

The idea of this test is that the multipliers $m_c$ and $m_s$ in the calculation of the minimum required capital 2.3 are increased if the number of exceedances is too high. In particular, these multipliers are calculated as $3 +$ add-on factor, where the add-on factor depends on the number of exceedances in the most recent 250 business days and is determined in accordance with the following table:

| Number of exceedance | Add-on factor |
|:---:|:---:|
| less than 5 | 0 |
| 5 | 0.4 |
| 6 | 0.5 |
| 7 | 0.65 |
| 8 | 0.75 |
| 9 | 0.85 |
| 10 or more | 1 |

## 7.2 Kupiec´s POF test

This statistical test is based on Kupiec (1995)[3] and examines whether the proportion of failures (POF) calculated from the data corresponds to its theoretical value. Hence, it is focused on the unconditional coverage property only.

The null hypothesis is that the actual relative frequency of exceedances ($q$) corresponds to the chosen probability level ($p$), which is equal to $\alpha$.

The test is a likelihood ratio test and is based on the following statistics:

$$LR = -2(\ln L(K, N, p) - \ln L(K, N, q)),$$

where $K$ denotes the number of exceedances in $N$ days (hence $q = K/N$) and $L$ denotes the likelihood function of the binomial distribution, i.e.

$$L(K, N, \theta) = \binom{N}{K} \theta^K (1 - \theta)^{N-K}.$$

The *LR* statistics is asymptotically $\chi^2$ distributed with 1 degree of freedom.

If $p \approx q$, then *LR* is close to zero and we do not reject the null hypothesis. On the other hand, if $q$ is significantly different from $p$, the likelihood function of the binomial distribution calculated at $\theta = p$ is significantly below its maximum value, which is attained at $\theta = q$. Hence, *LR* is significantly positive and the null hypothesis is rejected.

## 7.3 Christoffersen´s test

Unlike the previous test, this test focuses on both the properties (the correct number of exceedances as well as their independence). It is based on Christoffersen (1998)[4].

The statistics is as follows:

$$LR_{ch} = LR_{pof} + LR_{ind},$$

where

$LR_{pof}$ is the statistics from the Kupiec´s POF test,

$LR_{ind}$ is the statistics for testing the null hypothesis that exceptions are independent.

Statistics $LR_{ind}$ is calculated based on the following formula:

$$LR_{ind} = -2\ln\left(\frac{(1-q)^{N_{00}+N_{10}}q^{N_{01}+N_{11}}}{(1-q_{01})^{N_{00}}(1-q_{11})^{N_{10}}(q_{01})^{N_{01}}(q_{11})^{N_{11}}}\right),$$

---

[3]KUPIEC, N. H. : Techniques for Verifying the Accuracy of Risk Measurement Models, Journal of Derivatives, Vol. 3, No. 2, 1995

[4]CHRISTOFFERSEN, P. F. : Evaluating Interval Forecasts, Internal Economic Review, Vol. 39, No. 4, November 1998

where

$N_{01}$ – the number of days with an exceedance following after no exceedance

$N_{11}$ – the number of days with an exceedance following after an exceedance

$N_{00}$ – the number of days with no exceedance following after no exceedance

$N_{10}$ – the number of days with no exceedance following after an exceedance

$q_{01}$ – relative frequency of exceedances given no exceedance occurred the day before

$q_{11}$ – relative frequency of exceedances given an exceedance occurred the day before

The conditional relative frequencies are calculated as follows:

$$q_{01} = \frac{N_{01}}{N_{00} + N_{01}}, \qquad q_{11} = \frac{N_{11}}{N_{10} + N_{11}}, \qquad q = \frac{N_{01} + N_{11}}{N_{00} + N_{01} + N_{10} + N_{11}}.$$

It is ease to observe that under the null hypothesis, one has $q_{01} = q_{11} = q$ (the relative frequency of exceedance should not depend on whether an exceedance occurred on the previous day or not) and the value of $LR_{ind}$ is zero.

$LR_{ch}$ has asymptotically $\chi^2$ distribution with 2 degrees of freedom.

## 7.4   Other backtesting methods

There are several important properties of VaR models which are not covered by the test mentioned above. For example, the Christoffersen´s test is only focused on the situation where an exceedance follows just on the next day after the previous exceedance. Other types of autocorrelation are not captured. This issue is addressed by e.g. Kupiec´s mixed test[5].

The previous tests also neglect the actual size of an exceedance. A test taking into account this size has been proposed e.g. by Lopez[6].

These additional tests are out of the scope of this document. However, the interested readers and / or well-motivated risk managers might follow the references given in the footnotes.

---

[5]HAAS, M. : New Methods in Backtesting, Financial Engineering Research Center, Working Paper, 2001

[6]LOPEZ, CH.A. : Methods for Evaluating Value-at-Risk Estimates, Economic Policy Review, Vol. 4, No. 3, 1998