XAIA INVESTMENT

# WHAT'S VaR?

Jan-Frederik Mai
XAIA Investment GmbH
Sonnenstraße 19, 80331 München, Germany
jan-frederik.mai@xaia.com

Date: August 1, 2013

**Abstract** The Value-at-Risk (VaR) of a portfolio is one of the most fundamental risk measures applied in the financial industry. Although sometimes being heavily criticized by theorists, in practice the VaR is an important number for many investors, both due to regulatory investment restrictions as well as for making investment decisions. The present article aims at presenting in simple terms how the VaR is computed in practice, and at collecting and discussing some of the critical aspects. Special attention is put on data availability, long holding periods, and implicit distributional assumptions underlying the common $\sqrt{t}$-rule.

## 1 Introduction

The Basel Accord was amended to add a charge for market risks in 1996. The primary goal was to keep track of the proprietary trading activities that have increased sharply during the 1990s, and for which no capital charges were assigned before. One of the major consequences is the obligation for banks (and similar investment vehicles) to report a so-called Value-at-Risk (VaR) for those parts of their trading and banking books that are exposed to market risks, such as over-the-counter derivatives. Loosely speaking, the idea of the VaR is that it is a single number – a risk measure – which is reported on a regular basis and increases with increasing riskiness of the portfolio (and vice versa). To accomplish this task banks in general are allowed to use the so-called *internal-models approach*, which is what's typically done. This means each and every bank comes up with its own, reasonable internal model for the computation of the VaR, but these internal models have to be accredited by the regulator before they can be used. Depending on the legislation and the business model, there exist strict regulatory rules about how big the VaR is allowed to become. As a consequence, the VaR is of paramount interest for investors and business leaders, and banks employ hundreds of high-qualified people in order to carry out a viable and reasonable computation of this number. Nowadays, more than 15 years later, several market standard approaches for the assignment of the VaR have been established and are commonly applied. Nevertheless, the academic literature does not get tired of indicating problematic issues surrounding the VaR – both conceptual problems as well as problems regarding the computational viability. The present article aims at explaining what the VaR is, how it is computed in practice, and discusses potential weaknesses of the VaR and its computation of which one should be aware.

## 2 Definition of VaR

Mathematically, the VaR is defined as a quantile of a random variable. For example, let $X_t$ denote the log-return of a portfolio in the period $[0, t]$. Denoting the value of the portfolio at time $t$ by $S_t$, where $t = 0$ is today, the log-return $X_t$ is defined as

$$X_t := \log\Big(\frac{S_t}{S_0}\Big) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\Big(\frac{S_t - S_0}{S_0}\Big)^k \approx \frac{S_t - S_0}{S_0},$$

where the second equality holds for future portfolio values $S_t < 2\,S_0$ that do not double on $[0, t]$, and the last equality is a first-order approximation showing that the log-return is actually similar to the discrete return. At least for small values of $t$ this approximation is justified in practice. In order to define the VaR one chooses a confidence level $1 - \alpha \in (0, 1)$, typically $\alpha$ is small so that $1 - \alpha = 95\%$ or $= 99\%$. The VaR with confidence level $1 - \alpha \in (0, 1)$ of a portfolio, denoted $VaR_t^{1-\alpha}$, is defined[1] as an $\alpha$-quantile of the random variable $X_t$, i.e. a value which satisfies the equality

$$\mathbb{P}\big(X_t \leq VaR_t^{1-\alpha}\big) = \alpha. \tag{1}$$

In words, the probability that the portfolio return $X_t$ is smaller or equal to $VaR_t^{1-\alpha}$ is precisely $\alpha$. In case the distribution function $x \mapsto F(x) := \mathbb{P}(X_t \leq x)$ of $X_t$ has jumps, e.g. if the distribution of $X_t$ is discrete, there might not exist any number $VaR_t^{1-\alpha}$ satisfying Equation (1). If this is the case, one determines the maximal value[2] $x_l$ satisfying $F(x_l) < \alpha$ and the minimal value $x_u$ satisfying $F(x_u) > \alpha$. The VaR can now be chosen as a number between $x_l$ and $x_u$, a common choice is

$$VaR_t^{1-\alpha} = x_l + \frac{\alpha - F(x_l)}{F(x_u) - F(x_l)}\,(x_u - x_l). \tag{2}$$

As an example, a typical case in practice is that $X_t$ has a discrete distribution taking precisely $250$ different values $x_1, \ldots, x_{250}$ with equal probabilities $\mathbb{P}(X_t = x_k) = 1/250$, and $1-\alpha = 99\%$. In this case, there does not exist a value $VaR_t^{1-\alpha}$ satisfying Equation (1). The maximal value $x_l$ such that $F(x_l) < 1\%$ is precisely the second smallest value of $x_1, \ldots, x_{250}$, denoted $x_{(2)}$, and the minimal value $x_u$ such that $F(x_u) > 1\%$ is precisely the third smallest value of $x_1, \ldots, x_{250}$, denoted $x_{(3)}$. This is because

$$F(x_{(2)}) = \frac{2}{250} < 1\% < \frac{3}{250} = F(x_{(3)}).$$

In this case, the VaR equals the arithmetic mean of $x_{(2)}$ and $x_{(3)}$ since an application of Formula (2) yields

$$VaR_t^{99\%} = x_{(2)} + \frac{1\% - \frac{2}{250}}{\frac{3}{250} - \frac{2}{250}}\,(x_{(3)} - x_{(2)}) = \frac{x_{(2)} + x_{(3)}}{2}.$$

**Remark 2.1 (Loss or Gain?)**
In the present article, we define the VaR as a low quantile of the return distribution. Equivalently, the VaR can be defined as a

[1] Alternatively, the VaR may also be given in absolute terms instead of a percentage.

[2] Strictly speaking, the maximum might not exist, so we mean the supremum.

high quantile of the associated loss distribution, i.e. as a quantile of $-X_t$ instead of $X_t$. Sometimes this causes confusion and one must always be aware of what one is talking about when discussing the VaR.

The intuitive idea of the Value-at-Risk according to the definition above is a “worst-case“ portfolio return, where “worst-case“ is more precisely defined in terms of a small probability of the return taking values below the VaR. Given this definition, there are still challenges remaining, which we are going to discuss in the remaining sections:

- How do we determine the (required) function $F(x) = \mathbb{P}(X_t \leq x)$, i.e. the probability distribution function of the log-return? Section 3 discusses how this is done in practice.
- How reliable is the VaR if the holding period $t$ is large? And how to compute the VaR for long holding periods when not enough data is available? This is discussed in Section 4.
- Does the definition of the VaR suffer from fundamental shortcomings, e.g. the negligence of heavy tails? This is discussed in Section 5.

## 3 How is the VaR computed in practice?

Generally speaking, in order to compute the VaR two steps are required:

(a) A mathematical model for the probability distribution of the (random) return $X_t$ must be set up.

(b) The computational resources to compute the VaR, i.e. the respective quantile, from this modeled distribution must be available.

There are three methodologies that are commonly referred to as being standard in the market. The *historical method*, the *variance-covariance method*, and the *Monte Carlo method*. We find it very misleading that the Monte Carlo method is included in this list, because it is not really concerned with the modeling task (a) but rather a generic computational tool that can be applied to perform the task (b) for a wide family of models for $X_t$ (but this model has to be defined first, i.e. step (a) must be carried out first). Let us describe roughly what these methods propose:

- **Historical Method:** A model $X_t = f(R_1(t), R_2(t), \ldots, R_m(t))$ is set up which defines the log-return $X_t$ as a function of certain risk factors $R_1(t), \ldots, R_m(t)$, such as interest rates, FX rates, credit spreads and so on. Typically, the function $f$ is (the logarithm of) a sum over all portfolio constituents (swaps, bonds, CDS, stocks,...), and each summand is an instrument-specific pricing algorithm depending on certain market observables related to the risk factors (e.g. a bond pricer depending on interest rate data and credit spread data). A simplified approach could also be to set $m = 1$, $R_1(t) := S_t/S_0$ and $f(x) = \log(x)$, meaning that only the history of the overall portfolio is used as basis for the VaR computation. For each risk factor $i = 1, \ldots, m$, we observe a historical time series $R_i(-n\,t)$, $R_i(-(n-1)\,t)$, ..., $R_i(-2\,t)$, $R_i(-t)$, $R_i(0)$, where $R_i(0)$ is today's value of risk factor $i$. One computes the increments of the historical time series of all risk factors as $\Delta R_i(-k\,t) := R_i(-(k-1)\,t) - R_i(-k\,t)$, $k = 1, \ldots, n$, $i = 1, \ldots, m$. The probability distribution of $X_t$ is assumed to be discrete with equally-weighted outcomes. These finitely many possible values of $X_t$ are defined as

$$x_k := f\big(R_1(0) + \Delta R_1(-k\,t), \ldots, R_m(0) + \Delta R_m(-k\,t))\big),$$

$k = 1, \ldots, n$, leading to the probability distribution function

$$F(x) = \mathbb{P}(X_t \leq x) = \sum_{x_k \leq x} \underbrace{\mathbb{P}(X_t = x_k)}_{=1/n} = \frac{\#\{k \,:\, x_k \leq x\}}{n},$$

$x \in \mathbb{R}$. The distinct advantage of the historical method is that it comes with almost minimal modeling assumptions, except for the definition of the risk factors and the function $f$, which in practice boils down to specifying a standard pricing methodology for each and every instrument type in the portfolio. In particular – and this is probably the most convincing argument in favor of this method – there is no need to design a mathematical model for the dependence structure between the risk factors $R_1, \ldots, R_m$, which is a horrible mathematical exercise. Rather this dependence is taken into account implicitly by generating the possible outcomes $x_k$ via joint increments $(\Delta R_1(-k\,t), \ldots, \Delta R_m(-k\,t)))$ of the risk factors, i.e. the historical co- or counter-movements of risk factors are taken into account.

One major drawback of the historical method is that the number of scenarios is directly linked to the length of the time series used as input. For large portfolios – which intuitively have many more potential outcome scenarios than small portfolios – the obtained number of possible model scenarios is typically way too small. Consequently, in particular for applications in portfolio management, the historical method might be unreliable.

- **Variance-Covariance Method:** The first step is the same as in the historical method: one sets up a probabilistic model $X_t = f(R_1(t), R_2(t), \ldots, R_m(t))$ explaining the log-return in terms of risk factors. The difference is now that instead of using the observed increments $(\Delta R_1(-k\,t), \ldots, \Delta R_m(-k\,t))$, $k = 1, \ldots, n$, as equally-weighted scenarios, it is assumed that the required vector of risk factors $(R_1(t), \ldots, R_m(t))$ has a multivariate normal distribution with a certain mean vector $\boldsymbol{\mu}$ and covariance matrix $\Sigma$. Denoting by $n_{\boldsymbol{\mu},\Sigma}(\boldsymbol{x})$ the $m$-dimensional probability density function of $(R_1(t), \ldots, R_m(t))$, the required distribution function of $X_t$ is then given as

$$F(x) = \mathbb{P}(X_t \leq x) = \iint_{\mathbb{R}^m} f(\boldsymbol{x})\, n_{\boldsymbol{\mu},\Sigma}(\boldsymbol{x})\, d\boldsymbol{x}. \qquad (3)$$

Since high-dimensional integrals are difficult to compute, a further approximation is required. The so-called *Delta-Normal* *approximation* replaces the (in general highly non-linear) function $f$ by its first-order Taylor expansion, yielding

$$X_t = X_0 + \sum_{i=1}^{m} f_i \left(R_i(t) - R_i(0)\right),$$

where $f_i = f_i\big(R_1(0), \ldots, R_m(0)\big)$ denotes the partial derivative of $f$ with respect to the $i$-th risk factor $R_i$. With known partial derivatives (sensitivities) $f_i$ this approximation implies that the distribution of $X_t$ is normal, and its mean $\mu_X$ and variance $\sigma_X^2$ can be computed easily from $\boldsymbol{\mu}$ and $\Sigma$. This ultimately means that the VaR is merely a quantile of the univariate normal distribution with parameters $\mu_X, \sigma_X^2$, derived from the continuous distribution function

$$F(x) = \mathbb{P}(X_t \le x) = \int_{-\infty}^{x} \frac{1}{\sigma_X \sqrt{2\,\pi}}\, e^{-\frac{(t-\mu_X)^2}{2\,\sigma_X^2}}\, dt.$$

Since the Delta-Normal approximation is sometimes considered to be too coarse, there even exist refinements to second-order approximations for $f$. Summing up, however, one must conclude that the variance-covariance methodology bears numerous weaknesses, mostly stemming from the underlying assumption of the risk factors being multivariate normally distributed. For instance, the normal distribution might be too simplistic and ignore relevant features such as extreme scenarios or asymmetries. Moreover, reliable estimates for the required mean-, variance- and covariance-parameters $\boldsymbol{\mu}$ and $\Sigma$ might require much more historical data than available.

- **Monte Carlo Method:** By the term “Monte Carlo“ one always refers to stochastic simulation, i.e. the Monte Carlo method does not compute the VaR from given data directly, but instead uses the given data to generate a huge set of possible scenarios for $X_t$ from which the VaR is then estimated as the respective empirical quantile. In theory, given a certain probability distribution for the risk factors $R_1(t), \ldots, R_m(t)$ of the two aforementioned methods (or a probability distribution generated from a different, third method), one can always apply the Monte Carlo method in order to generate possible outcomes for $X_t$. This might sometimes be necessary when an exact computation is not possible. One example for such a case is if the variance-covariance method is applied and one does not want to apply the Delta-Normal approximation but rather wants to compute the VaR based on the distribution function (3). For large $m$ this is typically not possible due to the high-dimensional integration, but a Monte Carlo simulation is straightforward. Hence, as mentioned before, the Monte Carlo method does not really refer to a modeling approach for the random variable $X_t$ but merely serves as a computational tool added on top of a modeling approach.

Most standard is the historical method because it does not rely on too many modeling assumptions, is relatively straightforward to implement, and widely accepted by regulators.

## 4 How to compute a VaR with long holding period?

Clearly, $\alpha \mapsto VaR_t^{1-\alpha}$ is a non-decreasing function, i.e. the smaller the required confidence level, the smaller the uncertainty measured in terms of VaR. Concerning the dependence of the VaR on the holding period, however, the picture becomes more difficult. Intuitively one might think that $t \mapsto VaR_t^{1-\alpha}$ is a decreasing function, because the longer one is exposed to risk, the more risky one's position. However, what if the portfolio has a strong positive mean return and only minimal volatility? For instance, consider a fictitious investment strategy which is run for precisely three days. It is known in advance that the strategy gains one dollar on two of the three days and looses one dollar on the remaining day, only it is unknown on which of the three days the loss occurs. Clearly, with a holding period of three days the VaR (at any confidence level) equals $+1$ dollar, because this overall gain is certain. However, the $99\%$ confidence level VaR with one-day holding period equals $-1$ dollar, hence is more conservative than the VaR with longer holding period. Of course, one might argue that this example is pathological, but in fact there exist real-world portfolios which have a steady, positive growth rate and exhibit almost zero volatility - when this phenomenon actually appears. There is a second conclusion we can draw from this tiny example: The computation of a VaR at two different holding periods might necessitate two completely different calculations. In the example above, the three-day VaR is easy to compute because the three-day return is a constant. However, the one-day return is random and therefore the one-day VaR requires a more involved computation (although still easy due to the simplicity of the example). However, in real-world risk management applications this observation constitutes a severe problem. Quoting from BIS Consultative Document (2012) concerning the Basel III framework:

*“The Committee has agreed that the differentiation of market liquidity across the trading book will be based on the concept of liquidity horizons. It proposes that banks' trading book exposures be assigned to a small number of liquidity horizon categories. [10 days, 1 month, 3 months, 6 months, 1 year]. The shortest liquidity horizon (most liquid exposures) is in line with the current 10-day VaR treatment in the trading book. The longest liquidity horizon (least liquid exposures) matches the banking book horizon at one year. The Committee believes that such a framework will deliver a more graduated treatment of risks across the balance sheet.“*

This makes clear that a computation of a VaR with different holding periods is desirable. For the computation of a one-day VaR it is market standard to base the computation on historical data ranging back one year, or $250$ trading days. In order to avoid overlapping estimation intervals, the computation of a VaR with longer holding period requires a much longer data history. However, longer data history is often not available. For instance, think of a portfolio investing into newly issued bonds that have not existed two years ago.

Due to the lack of historical data, in practice it is not uncommon to compute a VaR with holding period one day, and then compute a VaR with longer holding period by transforming the one-day VaR by some formula, e.g. by the so-called $\sqrt{t}$-rule, see paragraph 4.1 below. However, it is important to point out that such a computation strongly relies on simplifying (and potentially unrealistic) assumptions. Mathematically speaking, assume we are given a distribution for the log-return $X_1$, and from it we need to derive the probability distribution of $X_t$ for a natural number $t \in \mathbb{N}$. Using $X_0 = 0$, we may rewrite

$$\begin{aligned} X_t &= X_t - (X_{t-1} - X_{t-1}) - \ldots - (X_1 - X_1) - X_0 \\ &= (X_t - X_{t-1}) + \ldots + (X_2 - X_1) + (X_1 - X_0) \\ &= \Delta X_t + \ldots + \Delta X_2 + \Delta X_1, \end{aligned}$$

i.e. the log-return over the period of length $t$ equals the sum of $t$ log-returns over a period of length $1$. Notice that the last equality uses the notation $\Delta X_k := X_k - X_{k-1}$. A simplifying assumption, which underlies all subsequently presented computation methods is the following:

(A) The (random) log-returns $\Delta X_1, \Delta X_2, \ldots, \Delta X_t$ over the next $t$ periods of length $1$ are stochastically independent and identically distributed.

Be aware that this is an assumption which is unrealistic, but not uncommon in Mathematical Finance. For instance, it is one of the crucial assumptions underlying the seminal Black-Scholes stock price model used for stock option pricing. In a certain sense, this assumption is conservative, since it completely ignores portfolio management decisions within the considered time horizon. For instance, if the sum of the first $20$ log returns $\Delta X_1, \ldots, \Delta X_{20}$ is strongly negative, in reality the portfolio management takes measures against a further downturn of the portfolio (e.g. shifting from risky into non-risky assets), which alters the probability distribution of the remaining log-returns $\Delta X_{21}, \ldots, \Delta X_t$ in a favorable way (e.g. introduce a floor to the overall loss). The effects of such management decisions are completely neglected under assumption (A).
In the sequel we present three possible ways to compute $VaR_t^{1-\alpha}$ from $VaR_1^{1-\alpha}$ under assumption (A). To this end, we assume that the distribution of $X_1$ is known and discrete, say

$$\mathbb{P}(X_1 = x_k) = \frac{1}{n}, \quad k = 1, \ldots, n.$$

This assumption is in accordance with the market practice of computing the VaR with the historical method. A typical size for $n$ encountered in practice is $n = 250$.

### 4.1 The $\sqrt{t}$-rule: assuming a normal law with mean zero

The VaR for the longer holding period is computed from the VaR with the shorter holding period according to the following formula:

$$VaR_t^{1-\alpha} = \sqrt{t}\, VaR_1^{1-\alpha}.$$

This rule is commonly applied in practice and widely accepted by regulators, even though it is not always justifiable. In addition to assumption (A) this rule implicitly assumes that the log-return $X_1$ has a normal distribution with mean zero. For a historically generated, discrete distribution of $X_1$ especially the assumption of zero mean can be highly questionable. If the mean of the distribution of $X_1$ is positive, then the $\sqrt{t}$-rule is (too) conservative. If it is negative, then the $\sqrt{t}$-rule sugarcoats the result, i.e. conceals riskiness.

### 4.2 Improved $\sqrt{t}$-rule: assuming a normal law (with arbitrary mean)

In addition to assumption (A) one makes the following assumption:

(A+) The log-returns $\Delta X_1, \ldots, \Delta X_t$ are normally distributed with a mean parameter $\mu$ and a variance parameter $\sigma^2$.

To set this assumption into action in practice one has to approximate the discrete distribution of $X_1$ by a normal distribution, which for large portfolios can sometimes be justified pretty well. The obvious and best-possible fit of a normal law to the given discrete probability distribution is the one with parameters $\mu$ and $\sigma^2$ as follows:

$$\mu = \frac{1}{n}\sum_{k=1}^{n} x_k, \quad \sigma^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \mu)^2.$$

Under this choice and assumptions (A) and (A+), the log-return $X_t = \Delta X_1 + \ldots + \Delta X_t$ has a normal distribution with mean $t\,\mu$ and variance $t\,\sigma^2$. Hence, the VaR can be computed as follows:

$$\begin{aligned} VaR_t^{1-\alpha} &= \alpha\text{-quantile of the normal distribution} \\ &\qquad\qquad \text{with mean } t\,\mu \text{ and variance } t\,\sigma^2 \\ &= t\,\mu + \sqrt{t}\,\sigma\,\Phi^{-1}(\alpha), \end{aligned}$$

where $\Phi^{-1}$ denotes the inverse of the standard normal distribution function, e.g. $\Phi^{-1}(0.01) \approx -2.32635$. From this formula one observes that the $\sqrt{t}$-rule is obtained if and only if $\mu = 0$, which is not true in general. Figure 1 visualizes an exemplary fit of a normal distribution to a given discrete distribution of the log-return $X_1$; the latter being represented as a histogram.

### 4.3 Only assumption (A)

Dropping the assumption (A+) of normality in the previous paragraph, the required distribution of $X_t = \Delta X_1 + \ldots + \Delta X_t$ is the $t$-fold convolution of the given discrete distribution of $X_1$. In general, this distribution is very difficult to compute, because its support consists of up to $n^t$ different values. However, it is not too difficult to simulate this distribution efficiently, i.e. compute $VaR_t^{1-\alpha}$ using Monte Carlo simulation. In (Mai, Scherer, 2012a, Algorithm 6.13, p. 247) it is described how to simulate random variables taking only $n$ values with computational effort in $\mathcal{O}(\log(n))$, general information concerning the Monte Carlo method can be found in (Mai, Scherer, 2012a, Chapter 7). One generates a huge number of scenarios for $X_t$ and the VaR is finally computed as the empirical quantile of the simulated samples.

## 5 What are fundamental pitfalls of VaR?

The Value-at-Risk is a quantile, i.e. a functional, of the distribution of the portfolio log-return $X_t$. Generally speaking, the information content of this number suffers from two major problems one has to keep in mind:

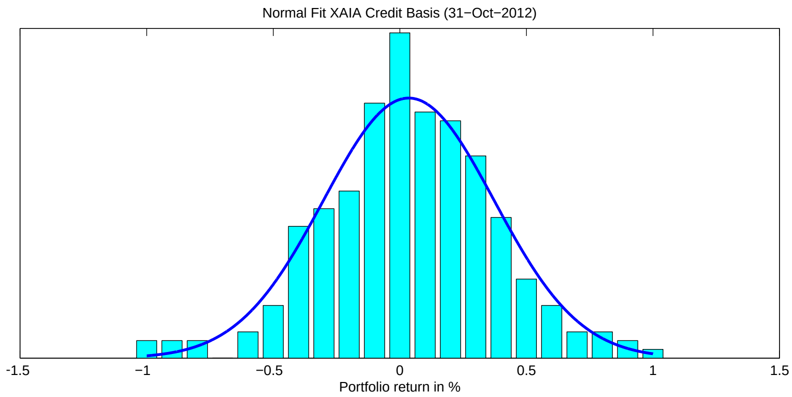


Fig. 1: Best fit of a normal distribution to a given discrete distribution with $n = 250$ values. The dark blue line shows the density of the fitted normal law, and the light blue histogram shows the discrete distribution. The data correspond to the $250$ scenarios generated for the daily log-return of the fund XAIA Credit Basis on October 31, 2012.

- *Imprecise probability law:* The underlying probability distribution of $X_t$ is unknown and must therefore be estimated somehow, typically from historical data. This estimation procedure might bear essential errors.
- *Information loss by quantile transformation:* Projecting the complete probability distribution of $X_t$ onto one single number, relevant information about the future return distribution is ignored.

Regarding the estimation of the probability distribution of $X_t$, several points have to be mentioned. Firstly, typical methods applied in practice rely solely on historical data. This might ignore relevant future developments that have never occurred before (or at least not in the considered time series). Providing an example: assume that an investment strategy heavily applies a certain type of derivative for which the regulatory setup changes essentially within the upcoming month. The future returns of the portfolio might suffer significantly from this regulatory change. Since the information about the upcoming change is known already today, ideally this information should be incorporated into the VaR calculation – of course, this might not be an easy task. However, every model relying on historical data doesn't do that. Secondly, since $X_t$ depends on numerous portfolio constituents and their interdependence, the estimation of the probability distribution is a very hard task in general. The aforementioned methods tackle this issue either by resorting to the empirical joint distribution of risk factors (historical method) or to normality assumptions for risk factors (variance-covariance method). Like every model, both approaches provide a trade-off between realism and viability, and one might argue whether the trade-off is acceptable or not. By the way, already the modeling step of defining the risk factors and how they influence the return distribution certainly simplifies reality. One of the most critical aspects in the estimation routine is definitely the dependence structure between the constituents. A mathematical description of dependence between numerous portfolio constituents is way beyond trivial, and the common treatment of dependence via estimated correlation matrices is often not appropriate, see Mai, Scherer (2012b) for more issues related with dependence modeling and related references.

Regarding the information loss by the quantile transformation, the seminal article Artzner et al. (1999) axiomatically defines “reasonable“ risk measures and shows that Value-at-Risk in general is not of such kind. Loosely speaking, the Value-at-Risk provides no information about how bad things can go, given we already know that things will go bad. For instance, if $X_t$ has a distribution which is positive with $99.1\%$ probability, but in $0.9\%$ of all cases equals a total loss, then the VaR at $99\%$ confidence level is positive. This example shows how the VaR tends to neglect heavy tails of a distribution, because they are truncated at a quantile. Such a drawback can be overcome by resorting to the more conservative risk measure which computes the conditional expectation over all scenarios below the VaR, called[3] *Tail-VaR*. It can be shown that this risk measure is reasonable in the sense of Artzner et al. (1999), and it is safe to say that most experts in academia and practice – and also regulators – nowadays agree on the fact that the VaR should be replaced by the Tail-VaR in the future regulatory framework.

## 6 Conclusion

It has been explained what the Value-at-Risk is and how it is computed in practice. Moreover, it has been illustrated how a VaR with long holding period could be computed. Finally, some fundamental pitfalls of VaR have been pointed out.

## Acknowledgements

Helpful comments by Christian Hering and Christofer Vogt on earlier versions of the manuscript are highly appreciated.

## References

P. Artzner, F. Delbaen, J. Eber, D. Heath, Coherent measures of risk, *Mathematical Finance* **9:3** (1999) pp. 203–228.

BIS Consultative Document, Fundamental review of the trading book, Annex 4, available at BIS.org (2012).

J.-F. Mai, M. Scherer, Simulating Copulas, *Series in Quantitative Finance, Imperial College Press, Singapore* (2012).

J.-F. Mai, M. Scherer, What makes dependence modeling challenging? Pitfalls and how to circumvent them, *XAIA homepage article* (2012).

[3]The Tail-VaR is sometimes also called *Expected Shortfall*, *Conditional Tail Expectation*, *CVaR*, or *Average VaR*.