The Value at Risk

Andreas de Vries*

FH Südwestfalen University of Applied Sciences, Haldener Straße 182, D-58095 Hagen, Germany

May 19, 2000

Contents

1 Introduction 1

2 The concept of Value at Risk 2

2.1 The case of $n$ subportfolios 3

3 Models to determine the VaR 4

3.1 Parametric models 4

3.2 Simulation models 5

4 How to calculate VaR 6

4.1 RiskMetrics 6

4.2 CreditMetrics 7

1 Introduction

The main business of banks and insurance companies is risk. Banks and financial institutions lend money, running the risk of losing the lended amount, and they borrow “short money” having less risk but higher expected rates of return. Insurance companies on the other hand earn a risk premium for guaranteeing indemnification for a negative outcome of a certain event.

The evaluation of risk is essential for both kinds of business. During the 1990’s there has been established a measure for risk in finance theory as well as in practice, the Value at Risk, VaR. It was mainly popularized by J.P. Morgan’s RiskMetrics, a database supplying the essential statistical data to calculate the VaR of derivatives.

In the context of finance Value at Risk is an estimate, with a given degree of confidence, of how much one can lose from a portfolio over a given time horizon. The portfolio can be that of a single trader, or it can be the portfolio of the whole bank. As a downside risk measure, Value at Risk concentrates on low probability events that occur in the lower tail of a distribution. In establishing a theoretical construct for VaR, Jorion [10] first defines the critical end of period portfolio value as the worst possible end-of-period portfolio value with a pre-determined confidence level “$1 - \alpha$” (e.g., 99%) These worst values should not be encountered more than $\alpha$ percent of the time.

For example, a Value at Risk estimate of 1 million dollars at the 99% level of confidence implies that portfolio losses should not exceed 1 million dollars more than 1% of the time over the given holding period [10].

Currently, Value at Risk is being embraced by corporate risk managers as an important tool in the overall risk management process. Initial interest in VaR, however, stemmed from its potential applications as a regulatory tool. In the wake of several financial disasters involving the trading of derivatives products, such as the Barings Bank collapse (see [10], regulatory agencies such as

*e-mail: de-vries@fh-swf.de*
the Securities and Exchange Commission or the BIS, in cooperation with several central banks, embraced VaR as a transparent measure of downside market risk that could be useful in reporting risks associated with portfolios of highly market sensitive assets such as derivatives. Since VaR focuses on downside risk and is usually reported in currency units, it is more intuitive than other statistical terms. It is commonly used for internal risk management purposes and is further being touted for use in risk management decision making by non-financial firms [10, 11].

However, an essential part of the balance sum of banks and financial institutions is made by the classical banking products such as credits and loans. Usually they are contracted over the counter (OTC) and are not traded on markets.

2 The concept of Value at Risk

The sample space $\Omega$ of the expected rates of return (or expected “relative returns”) $r$ on the investment $W$ in some arbitrary assets is mathematically represented by the set $\mathbb{R}$. We assume that the expected rates of return $r(t)$ with respect to the time horizon $t$ of the investment is a random variable determined by the distribution function $F: \Omega \rightarrow [0,1]$,

$$F(x) = \int_{-\infty}^{x} p(r) \, dr, \quad (1)$$

where $p$ is the corresponding probability density. This means in particular that the expected rate of return $r(t)$ will achieve a value less than $x\%$ ($x \in \mathbb{R}$) after time $t$ with probability $P(r(t) \leq x) = F(x)$. Let $\tilde{\Omega} = \mathbb{R}$ be the sample space of currency-valued returns $R = r(t)W$. The expected loss $L(t)$ with respect to the time horizon $t$ of the investment $W$ then is given as the negative difference between the return and the mean value, $L(t) = \mu W - R(t) = (\mu - r(t))W$. It is a quantity in currency units (cu). Note that any return less than the expected one means an effective loss, even if it is positive. Positive values of $L(t)$ mean a loss after time $t$, negative ones a gain. Its distribution function $\tilde{F}: \tilde{\Omega} \rightarrow [0,1]$ is simply given by $\tilde{F}(L) = 1 - F(L/W - \mu)$, or

$$\tilde{F}(L) = 1 - \int_{-\infty}^{L} p(L'/W - \mu) \, dL', \quad (2)$$

with the probability density $-p(L'/W - \mu)$. The Value at Risk $\text{VaR}$ with respect to the time horizon $t$ of the investment then is defined as the maximal expected loss $L(t)$ not exceeded with probability $(1 - \alpha)$:

$$P(L(t) \leq \text{VaR}) = 1 - \alpha, \quad 0 \leq \alpha \leq 1. \quad (3)$$

In this paper the following convention is valid: $r$ denotes a relative quantity with respect to the invested capital $W$ of the asset (in %, say), whereas $R$ is the absolute return in currency units (cu).
\( \alpha \) is the default or downfall probability of the Value at Risk. For instance, according to the Basle Accord \([1]\) it should be be \( \alpha = 1\% \), and \( t = 10 \) days. The Value at Risk often is also called the “unexpected loss” of the investment, cf. \([12]\). We have \( P(L(t) \leq \text{VaR}) = \hat{F}(\text{VaR}) \), and the Value at Risk is nothing else than the \((1 - \alpha)\)-quantile of the random variable \( L(t) \):

\[
\text{VaR} = \hat{F}^{-1}(1 - \alpha),
\]

cf. \([10]\). To express it in terms of the expected rate of return \( r(t) \), the Value at Risk is the negative \( \alpha \)-quantile of the random variable \( r(t) \):

\[
\text{VaR} = (\mu - F^{-1}(\alpha)) W.
\]

Usually there only is interest on a positive Value at Risk. So we define the Value at Risk in the strict sense, \( \text{VaR}^+ \), by

\[
\text{VaR}^+ = \max(0, \text{VaR}).
\]

In general, the Value at Risk is depending on the confidence level \( \alpha \), the investment horizon \( t \), the investment \( W \), and the probability distribution \( F \). According to the famous “moment problem” the probability density \( \rho \), and thus the distribution \( F \), is determined by the moments \( m_k \) for \( k = 1, 2, \ldots \), if and only if \( F \) is of bounded variation on \( \Omega \), see the appendix and \([20]\) §1.4. Thus the Value at Risk depends on \( m_k \), i.e., \( \text{VaR} = f \cdot W \) with the function \( f = f_{\alpha,t}(m_1, m_2, \ldots) \) given as

\[
f = \mu - F^{-1}(\alpha).
\]

As a first approximation we restrict ourselves to the first two moments, \( m_1 \) and \( m_2 \). But because the mean value is exactly the first moment, \( \mu = m_1 \), and the volatility \( \sigma \) is related to the second moment by \( \sigma^2 = m_2 - \mu^2 \), we have approximately

\[
\text{VaR} = f_{\alpha,t}(\mu, \sigma) \cdot W.
\]

This approximation in fact is exact for a Gaussian normally distributed random variable, for then all other moments vanish. The approximation can be generalized straightforwardly to the case of an arbitrary distribution, besides the extension to the dependence on more moments, \( f = f(\mu, \sigma, m_3, m_4, \ldots) \): According to the “estimating function approach” \([8]\) one can transform the given random variable to a nearly standard distributed one.

### 2.1 The case of \( n \) subportfolios

Suppose a portfolio \( \Pi \) consisting of \( n \) subportfolios \( \Pi_1, \ldots, \Pi_n \) \((n \in \mathbb{N})\). For each subportfolio \( \Pi_i \) we denote the rate of return by \( r_i \), or relative return, on the invested capital \( W_i \) with time horizon \( t \). In the sequel we will often call \( W_i \) simply the investment of \( \Pi_i \). The total portfolio value is denoted by \( W_\Pi \), where \( W_\Pi = \sum_{i=1}^n W_i \). We suppose \( r_i \) to be a real-valued square-integrable random variable with a probability distribution \( F_i: \mathbb{R} \rightarrow [0,1] \).

The total volatility \( \sigma_\Pi \) of all subportfolios is given by

\[
\sigma_\Pi^2 = \frac{1}{W_\Pi^2} \sum_{i,j=1}^n \rho_{ij} \sigma_i \sigma_j W_i W_j.
\]

Here \( \rho_{ij} \) are the correlation coefficients, cf. \([2,7,15]\). They form the correlation matrix \( \hat{C} = (\rho_{ij})_{i,j} \) of \( r_i \) and \( r_j \) which has some important properties \([2,3,7]\): (i) it is symmetric, i.e. \( \rho_{ij} = \rho_{ji} \) for each \( i, j = 1, \ldots, n \), (ii) it is positive semi-definite, (iii) its diagonal entries are 1, \( \rho_{ii} = 1 \), and (iv) its entries each have an absolute value not greater than 1, \( |\rho_{ij}| \leq 1 \). The correlation matrix \( \hat{C} \) is related to the covariance matrix \( C = (c_{ij}) \) simply by the equation

\[
C = \sigma^T \hat{C} \sigma,
\]
or $c_{ij} = \rho_{ij}\sigma_i\sigma_j$. It is a measure of the linear dependence of $r_i$ and $r_j$. It induces a quadratic form $Q: \mathbb{R}^n \times \mathbb{R}^n \to [0,\infty)$ for two vectors $x, y \in \mathbb{R}^n$ by

$$Q(x, y) = x^T C y.$$  

The total invested portfolio capital $W_\Pi$ is the sum of all individual investments $W_i$ of the subportfolios $\Pi_i$: $W_\Pi = \sum_{i=1}^n W_i$. Accordingly, the total expected return of the portfolio $\Pi$ is given by $R = \sum_i r_i W_i$. If $v_i$ denotes the (isolated) Value at Risk of subportfolio $\Pi_i$, we have $v_i = f_i W_i$, according to (8). The distribution $F$ of the total portfolio is given implicitly by the consistency condition

$$\sum_{ij} f_i f_j W_i W_j = f^2 W_\Pi^2$$  

with $f = F^{-1}(\alpha)$.

Let now the distributions all be Gaussian, i.e. $f_i = \sigma_i$. Defining the vector $v = (v_1, \ldots, v_n)$ we conclude that $\text{VaR}^2 = v^T \hat{C} v$, or

$$\text{VaR}^2 = \sum_{ij} \rho_{ij} f_i f_j W_i W_j.$$  

Because $\text{VaR}(\lambda f, \lambda' W) = \lambda \lambda' \text{VaR}(f, W)$, the Value at Risk $\text{VaR}$ is a homogeneous function of degree 1 with respect to either $f$ and $W \in \mathbb{R}^n$, if the respectively other one is held fixed.

Straightforward calculation yields

$$\frac{f_i}{\partial f_i / \partial \sigma_i} \frac{\partial}{\partial \sigma_i} \text{VaR} = W_i \frac{\partial \text{VaR}}{\partial W_i} \text{VaR}$$  

(14)

3 Models to determine the VaR

I give a short survey about most popular the basic methods to determine the Value-at-Risk, parametric models and simulations.

3.1 Parametric models

In parametric models the changes of the portfolio value obey an assumed parametric probability distribution. Under certain conditions one receives a closed analytic solution depending on the probability parameters. The simplest ones are parametric models based on the normal distribution hypothesis. Here I mention four one of these:

**Portfolio-normal VaR.** Here the total portfolio return $r$ changes according to a normal distribution, i.e. $r \sim N(\mu, \sigma^2)$. Then the variable transformation $r \mapsto z = (r - \mu) / \sigma$ yields $F$ as the standard normal distribution, $F(x) = \Phi((x - \mu) / \sigma)$. Thus for an investment the Value at Risk on the confidence level $(1 - \alpha)$ is given by [3] as

$$\text{VaR} = (z_\alpha \sigma - \mu) W \quad \text{with} \quad z_\alpha = -\Phi^{-1}(\alpha),$$  

(15)

cf. [16]. Because losses are always on one side of the distribution, for risk considerations we only regard the one-sided confidence level.

**Asset-normal VaR.** This is a special case of the $n$-subportfolio case in section [2]. The portfolio is assumed to consist of $n$ assets $W_1, \ldots, W_n$ which have normally distributed relative changes $r_1, \ldots, r_n$, i.e. $r_i \sim N(\mu_i, \sigma_i)$. Then the portfolio Value-at-Risk is given by (15), with

$$\mu_\Pi = \frac{1}{W_\Pi} \sum_{i=1}^n \mu_i W_i, \quad \sigma^2_\Pi = \frac{1}{W_\Pi^2} \sum_{i,j=1}^n \rho_{ij} \sigma_i \sigma_j W_i W_j.$$  

(16)

\[A\] rigorous mathematical introduction can be found, e.g., in [17] §26.
Both the portfolio-normal and the asset-normal distributions have little practical utility. Despite the fact that the normal distribution hypothesis is not correct, correlation estimates of a realistic $n$-asset portfolio has to face permanent emergence and vanishing of financial instruments, due to their individually finite maturities or portfolio extension with new positions.

Delta-normal and Delta-Gamma-normal VaR. In practice, the distributions of daily changes in many market variables have essentially fatter tails than the normal distribution. In the light of the unappropriateness of asset-normal Values at Risk it appears to be more senseful to consider $m$ “risk factors” $S(\tau) = (S_1(\tau), \ldots, S_n(\tau))$ such as bond prices, interest rates, currency exchange rates, stock indices etc. The value $W_i$ of each asset in the portfolio $i$ then depends on these risk factors, $W_i = W_i(S(\tau))$, and the corresponding expected return $R$ after period $\tau$ is given by

$$R(\tau) = \sum_{i=1}^{n} W_i(S(\tau)) - W_i(0).$$

The expected rate of return $r$ is given by $r = R(\tau)/W_\Pi$, with $W_\Pi = \sum_{i=1}^{n} W_i(0)$. For example, a simple asset portfolio is a linear combination of $n$ stocks $S$, i.e. $W(t) = N S(t)$ with the $n \times m$-matrix $N$ with real entries. But if the assets includes options on stocks, the portfolio is not linear in the stock prices $S_i$.

For $i, j = 1, \ldots, m$ we define the parameters $\delta_i$ and $\gamma_{ij}$ by

$$\delta_i = \frac{\partial R}{\partial S_i}_{\tau=0}, \quad \gamma_{ij} = \frac{\partial^2 R}{\partial S_i \partial S_j}_{\tau=0},$$

(17)

Of course, $\gamma_{ij} = \gamma_{ji}$. By Taylor expansion we see that

$$R(\tau) = \sum_{i} \delta_i \Delta_i(\tau) + \frac{1}{2} \sum_{i,j} \gamma_{ij} \Delta_i(\tau) \Delta_j(\tau) + \ldots$$

(18)

where $\Delta_i(\tau) := S_i(\tau) - S_i(0)$. The determination of the Delta-normal VaR then starts with the approximation $\gamma_{ij} = 0$ and the assumption that the $\Delta_i$’s are independent and commonly normally distributed, $\Delta \sim N(0, tC)$. This means that the portfolio value change $R$ as a whole is normally distributed, but each single position needs not to be. We then get

$$\text{VaR} = z_\alpha \sqrt{\tau} \cdot \sqrt{d \cdot C \cdot d}$$

(19)

where $d = (\delta_1 S_1(0), \ldots, \delta_m S_m(0))$, and $z_\alpha = \Phi^{-1}(\alpha)$ is the $\alpha$-quantile of the standard normal distribution. For details see [9, 13].

For the Delta-Gamma normal distribution the $\gamma_{ij}$’s do not vanish. In this case the VaR can only be determined approximately or numerically. A broad introduction to the most commonly used methods is given in [13].

3.2 Simulation models

An alternative to parametric models are simulation models. The probability distribution of the relative returns of a portfolio emerges from a fictitious rate of return generated by scenarios.

Historical simulation or bootstrapping. The aim of this method is to generate a future distribution of possible future scenarios based on historical data. The data usually consist of daily returns for all possible assets, reaching back over a certain period. The simulation of the daily return for a day in the future then simply is done by choosing by chance (uniform distribution!) one of the historical returns.

This method is very simple to implement. Its advantages are that it naturally incorporates any correlation between assets and any non-normal distribution of asset prices. The main disadvantage is that it requires a lot of historical data that may correspond to completely different economic circumstances than those that currently apply.
Monte Carlo simulation. Monte Carlo simulation is the generation of returns (and often of asset price paths) by the use of random numbers, drawn from an assumed model distribution, to build up a distribution of future scenarios.

The VaR is calculated as the appropriate percentile of the probability distribution of $r$. Suppose, for example, that we calculate 5,000 different sample values of $r$. The 1-day 99% VaR is the value of $r$ for the 50th worst outcome; the 1-day 95% VaR is the value for the 250th worst outcome.

The main advantages of Monte Carlo simulation are: It is the most powerful and theoretically most flexible method, because it is not restricted to a given risk term distribution; the grade of exactness can be improved by raising the sample number.

The main drawback is that it tends to be slow because a company’s complete portfolio has to be revalued many times. One way of speeding things up is to assume that equation (18) describes the relationship between $r$ and the $S_i$’s. We then avoid the need for a complete revaluation of the portfolio. This is sometimes referred to as the partial simulation approach.

4 How to calculate VaR

Typically the data required for the calculations of VaR are statistical parameters for the ‘underlyings’ and measures of a portfolio’s current exposure to these underlyings. The parameters include volatilities and correlations of the assets and, for longer time horizons, drift rates. In October 1994 the American bank JP Morgan introduced the system RiskMetrics as a publicly accessible service for the estimation of VaR parameters for tradable assets such as stocks, bonds, and equities. In April 1997, JP Morgan also proposed a similar approach, together with a data service, for the estimation of risks associated with the risk of default of loans, CreditMetrics.

4.1 RiskMetrics

A detailed technical description of the method for estimating financial parameters can be found at the web site [www.jpmorgan.com](http://www.jpmorgan.com). Here I shortly describe how volatilities and correlations are estimated from historical data.

Estimating volatility. The volatility of an asset is measured as the annualized standard deviation of returns. There are many ways of taking this measurement. In RiskMetrics the volatility $\sigma_i$ on day $i$ is measured as the square root of a variance that is an exponentially weighted moving average (EMWA) of the square of returns,

$$\sigma_i^2 = \frac{1 - \lambda}{\Delta t} \sum_{j = -\infty}^{i} \lambda^{i-j} (R_j - \langle R \rangle)^2,$$

where $\Delta t$ is the timestep (usually one day), $R_j$ is the return on day $j$, and $\langle R \rangle$ is the mean value over the period from day $i$ to $j$ (it is usually neglected, assuming that the time horizon is sufficiently small). The parameter $0 < \lambda \leq 1$ represents the weighting attached to the past volatility versus the present return. This difference in weighting is more easily seen if we simply calculate $\sigma_i^2 - \lambda \sigma_{i-1}^2$ yielding (for $\langle R \rangle \approx 0$)

$$\sigma_i^2 = \lambda \sigma_{i-1}^2 + (1 - \lambda) R_i / \Delta t.$$

JP Morgan has chosen the parameter $\lambda$ as either 0.94 for a horizon of one day and 0.97 for a horizon of one month.

Estimating correlation. Similarly to the estimating of volatility, RiskMetrics uses an exponentially weighted estimate

$$\sigma_{12,i} = \lambda \sigma_{12,i-1} (1 - \lambda) R_{1,i} R_{2,i} / \Delta t.$$
4.2 CreditMetrics

The CreditMetrics methodology is described in great detail at the web site [www.jpmorgan.com](http://www.jpmorgan.com). The corresponding dataset consists of four data kinds: transition matrices, yield curves, spread, and correlations. They depend on a model for change of credit rating: A credit rating or grade is assigned to firms as an estimate of their creditworthiness. This usually is done by rating agencies, the famous of which are Standard & Poor’s and Moody’s. Standard & Poor’s rate businesses as one of AAA, AA, A, BBB, BB, B, CCC, or Default. Moody’s use Aaa, Aa, A, Baa, Ba, B, Caa, Ca, C. The credit rating agencies continually gather data on individual firms and will, depending on the information, grade or regrade a company according to well-specified criteria. Migration to a higher rating will increase the value of a bond and decrease its yield, since it is less likely to default.

Transition matrices. The probability that a given company migrates to another rating class in one year’s time is given by a transition matrix \( M = (p_{ij}) \), where \((p_{ij})\) denotes the transition probability from rating class \(i\) to \(j\) over a given finite horizon. The matrix is time-dependent, forming a Markov process as time passes. For instance, a typical S&P transition matrix looks like:

<table>
<thead>
<tr>
<th></th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>.90829</td>
<td>.08272</td>
<td>.00736</td>
<td>.00065</td>
<td>.00066</td>
<td>.00014</td>
<td>.00006</td>
<td>.00012</td>
</tr>
<tr>
<td>AA</td>
<td>.00665</td>
<td>.90890</td>
<td>.07692</td>
<td>.00583</td>
<td>.00064</td>
<td>.00066</td>
<td>.00029</td>
<td>.00011</td>
</tr>
<tr>
<td>A</td>
<td>.00092</td>
<td>.02420</td>
<td>.91305</td>
<td>.05228</td>
<td>.00678</td>
<td>.00227</td>
<td>.00009</td>
<td>.00041</td>
</tr>
<tr>
<td>BBB</td>
<td>.00042</td>
<td>.00320</td>
<td>.05878</td>
<td>.87459</td>
<td>.04964</td>
<td>.01078</td>
<td>.00110</td>
<td>.00149</td>
</tr>
<tr>
<td>BB</td>
<td>.00039</td>
<td>.00126</td>
<td>.00644</td>
<td>.07710</td>
<td>.81159</td>
<td>.08397</td>
<td>.00970</td>
<td>.00955</td>
</tr>
<tr>
<td>B</td>
<td>.00044</td>
<td>.00211</td>
<td>.00361</td>
<td>.00718</td>
<td>.07961</td>
<td>.80767</td>
<td>.04992</td>
<td>.04946</td>
</tr>
<tr>
<td>CCC</td>
<td>.00127</td>
<td>.00122</td>
<td>.00423</td>
<td>.01195</td>
<td>.02690</td>
<td>.11711</td>
<td>.64479</td>
<td>.19253</td>
</tr>
<tr>
<td>Default</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Unless the time horizon is very long, the largest probability is typically for the bond to remain at its initial rating.

In the CreditMetrics framework, the time horizon is one year.

Yield curves and spreads. The CreditMetrics dataset consists of the risk-free yield to maturity for several currencies. It contains yields for maturity of 1, 2, 3, 5, 7, 10, and 30 years. Additionally, for each credit rating the dataset gives the spread above the riskless yield for each maturity. (Thus the spreads denote the differences between riskless bonds and a rated credit.) Evidently, the riskier the credit the higher the yield: Higher yield for risky credits is compensation for the possibility of not receiving future coupons or the principal. Often one also speaks of the “credit spread premium.”

![Figure 2: Yield curve and spreads](image-url)
Let us examine, as a simple example, a 5-year fixed-rate loan of €100 million at 6% annual interest. The market value of the loan is equal to its present value $V$,

$$
V = \frac{6}{1 + r_1 + s_1} + \frac{6}{(1 + r_2 + s_2)^2} + \frac{6}{(1 + r_3 + s_3)^3} + \frac{106}{(1 + r_4 + s_4)^4},
$$

measured in million €. Here $r_i$ denotes the risk-free rate (derived from the yield to maturity curve) expected after year $i$, and $s_i$ is the corresponding annual credit spread. If the loan is downgraded, the spread $s_i$ will rise, and thus the present value will fall. A credit rating upgrade has the opposite effect. Suppose that, during the first year, the borrower gets upgraded from BBB to A. Then $V$ is given by, say,

$$
V = 6 + \frac{6}{1.0372} + \frac{6}{1.0432^2} + \frac{6}{1.0493^3} + \frac{106}{1.0532^4} = €108.66.
$$

Hence at the end of the first year, if the loan borrower is upgraded from BBB to A, the €100 million (book value) loan has a market value of €108.66 million.

**Correlations.** To examine the behavior of a portfolio of risky bonds we must consider whether there is any correlation between rating or default of one bond to the rating or default of another. In other words: Are bonds issued by different companies or governments correlated? The CreditMetrics dataset gives the correlations between major indices in many countries.

Each company issuing bonds has the return on its stock decomposed into parts correlated with these indices and a part which is specific to the company. By relating all bond issuers to these indices we can determine correlations between the companies in our portfolio.

CreditMetrics is, above all, a way of measuring risk associated with default issues. From the CreditMetrics methodology one can calculate the risk, measured by standard deviation, of the risky portfolio over the required horizon. Because of the risk of default the distribution of returns is highly skewed, it is far from being normal. For details see [14 §§4 & 10], [19 §45].

**Appendix: Moments and cumulants**

Let $X$ be a real random variable and $p$ a probability distribution on the phase space $\Omega$ such that for a given $k \in \mathbb{N}$ the expected value of $|X|^k$ exists:

$$
\mathbb{E}[|X|^k] = \int_{\Omega} |x|^k p(x) \, dx < \infty.
$$

Then the $k$-th moment, or moment of order $n$, is defined by

$$
m_k := \mathbb{E}[x^k] = \int_{\Omega} x^k p(x) \, dx.
$$

Figure 3: The cash flow and the single present values resulting from each payment.
Especially the cases $k = 1$ and $k = 2$ are important: The first moment is called mean value, $\mu = m_1$, and the second moment is related to the variance $\sigma^2 := \langle (X - \mu)^2 \rangle$ simply by $\sigma^2 = m_2 - \mu^2$. For details see, e.g., [4] §1.2, [5] §3, or [7] §1.

From a theoretical point of view the moments are interesting. In the famous “moment problem” one asks: Does the knowledge of all moments $m_n$ suffice to determine the probability distribution? The problem is solved by application of the fundamental Hahn-Banach theorem under quite general assumptions, eg. [20] §1.4. However, the log-normal distribution for instance has the remarkable property that the knowledge of all its moments is not sufficient to characterize the corresponding distribution, see [4] §1.3.2.

For important computational purposes it is convenient to introduce the characteristic function of the probability density $p(x)$, defined as its Fourier transform $\hat{p}(z)$:

$$\hat{p}(z) := \int_{\mathbb{R}} e^{ixz} p(x) \, dx. \quad (25)$$

Since $p(x)$ is a probability distribution, we always have $\hat{p}(0) = 1$. The moments of $p(x)$ are easily obtained by the corresponding derivatives of $\hat{p}(x)$ at $z = 0$,

$$m_k = (-i)^k \left. \frac{d^k}{dz^k} \hat{p}(z) \right|_{z=0}. \quad (26)$$

We define the $k$-th normalized cumulant $\hat{c}_k$ of the probability distribution $p(x)$ as the $k$-th derivative of the logarithm of its characteristic function:

$$\hat{c}_k := \left. \frac{1}{(i\sigma)^k} \frac{d^k}{dz^k} \log \hat{p}(z) \right|_{z=0}. \quad (27)$$

We have simply $\hat{c}_2 = 1$. One often uses the third and fourth normalized cumulants, called the skewness $\hat{c}_3$ and the kurtosis $\kappa = \hat{c}_4$:

$$\hat{c}_3 = \left( \frac{(x - \mu)^3}{\sigma^3} \right), \quad \hat{c}_4 = \left( \frac{(x - \mu)^4}{\sigma^4} \right) - 3. \quad (28)$$

The definitions of the cumulants may look a bit arbitrary, but these quantities have remarkable properties. For instance, they simply add by summing independent random variables. Moreover for a Gaussian distribution all cumulants of order larger than two are vanishing, $\hat{c}_k = 0$ for $k \geq 3$. Hence the cumulants, in particular the kurtosis $\kappa$, are a measure of the distance between the given probability distribution $p(x)$ and the Gaussian distribution. For details see [4].

References

[1] Basler Ausschuss für Bankenaufsicht: Richtlinien für das Risikomanagement im Derivativgeschäft (1994); Änderungen der Eigenkapitalvereinbarung zur Einbeziehung der Marktrisiken (1996);


