Lecture 6: Risk and uncertainty

Prof. Dr. Svetlozar Rachev

Institute for Statistics and Mathematical Economics
University of Karlsruhe

Portfolio and Asset Liability Management

Summer Semester 2008

Copyright

The material is based on the text-book:

Svetlozar T. Rachev, Stoyan Stoyanov, and Frank J. Fabozzi
Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization: The Ideal Risk, Uncertainty, and Performance Measures
John Wiley, Finance, 2007

Prof. Svetlozar (Zari) T. Rachev
Chair of Econometrics, Statistics
and Mathematical Finance
School of Economics and Business Engineering
University of Karlsruhe
Kollegium am Schloss, Bau II, 20.12, R210
Postfach 6980, D-76128, Karlsruhe, Germany
Tel. +49-721-608-7535, +49-721-608-2042(s)
Fax: +49-721-608-3811
http://www.statistik.uni-karslruhe.de

Introduction

- There has been a major debate on the differences and common features of risk and uncertainty. Both notions are related but they do not coincide.
- Risk is often argued to be a subjective phenomenon involving exposure and uncertainty. That is, generally, risk may arise whenever there is uncertainty.
- In the context of investment management, exposure is identified with monetary loss. Thus, investment risk is related to the uncertain monetary loss to which a manager may expose a client.
- Subjectivity appears because two managers may define the same investment as having different risk.
- A major activity in many financial institutions is to recognize the sources of risk, then manage and control them. This is possible only if risk is quantified.

Introduction

- If we can measure the risk of a portfolio, then we can identify the financial assets which constitute the main risk contributors, reallocate the portfolio, and, in this way, minimize the potential loss by minimizing the portfolio risk.

Example 1

If an asset will surely lose 30% of its value tomorrow, then it is not risky even though money will be lost.
Uncertainty alone is not synonymous with risk either.
If the price of an asset will certainly increase between 5% and 10% tomorrow then there is uncertainty but no risk as there is no monetary loss.

$\Rightarrow$ Risk is qualified as an asymmetric phenomenon in the sense that it is related to loss only.

Introduction

- Uncertainty is an intrinsic feature of the future values of traded assets on the market.
- Investment managers do not know exactly the probability distribution of future prices or returns, but can infer it from the available data — they approximate the unknown law by assuming a parametric model and by calibrating its parameters.
- Uncertainty relates to the probable deviations from the expected price or return where the probable deviations are described by the unknown law.
- Therefore, a measure of uncertainty should be capable of quantifying the probable positive and negative deviations. Among uncertainty measures there are variance and standard deviation.

Introduction

Depending on the sources of risk, we can distinguish:

- **Market risk**
  It describes the portfolio exposure to the moves of certain market variables. There are four standard market risk variables — equities, interest rates, exchange rates, and commodities.
- **Credit risk**
  It arises due to a debtor's failure to satisfy the terms of a borrowing arrangement.
- **Operational risk**
  It is defined as the risk of loss resulting from inadequate or failed internal processes, people and systems.

Generally, a risk model consists of two parts:

1. Probabilistic models are constructed for the underlying sources of risk, such as market or credit risk factors, and the portfolio loss distribution is described by means of the probabilistic models.
2. Risk is quantified by means of a risk measure which associates a real number to the portfolio loss distribution.

$\Rightarrow$ Both steps are crucial. Non-realistic probabilistic models may compromise the risk estimate just as an inappropriate choice for the risk measure may do.

$\Rightarrow$ No perfect risk measure exists.

- A risk measure captures only some of the characteristics of risk and, in this sense, every risk measure is incomplete.
- We believe that it is reasonable to search for risk measures which are ideal for the particular problem under investigation.
- We start with several examples of widely used dispersion measures that quantify the notion of uncertainty.

Measures of dispersion

- Measures of dispersion calculate how observations in a dataset are distributed, whether there is high or low variability around the mean of the distribution.
- Intuitively, if we consider a non-random quantity, then it is equal to its mean with probability one and there is no fluctuation whatsoever around the mean.
- In this section, we provide several descriptive statistics widely used in practice and we give a generalization which axiomatically describes measures of dispersion.

Standard deviation

- Standard deviation is calculated as the square root of variance.
- It is usually denoted by σ_X, where X stands for the random variable we consider,

$$\sigma_X = \sqrt{E(X - EX)^2} \tag{1}$$

in which E stands for mathematical expectation.
- For a discrete distribution, equation (1) changes to

$$\sigma_X = \left(\sum_{k=1}^{n} (x_k - EX)^2 p_k \right)^{1/2},$$

where x_k, $k = 1, \ldots, n$ are the outcomes, p_k, $k = 1, \ldots, n$ are the probabilities of the outcomes,

$$EX = \sum_{k=1}^{n} x_k p_k$$

is the mathematical expectation.

- The standard deviation is always a non-negative number.
- If it is equal to zero, then the random variable (r.v.) is equal to its mean with probability one and, therefore, it is non-random. This conclusion holds for an arbitrary distribution.

Standard deviation

Why the standard deviation can measure uncertainty?

- Suppose that X describes the outcomes in a game in which one wins \$1 or \$3 with probabilities equal to 1/2.
- The mathematical expectation of X, the expected win, is \$2,

$$EX = 1(1/2) + 3(1/2) = 2.$$

  The standard deviation equals \$1,

$$\sigma_X = \left((1-2)^2\frac{1}{2} + (3-2)^2\frac{1}{2}\right)^{1/2} = 1.$$

- In this equation, both the positive and the negative deviations from the mean are taken into account. In fact, all possible values of the r.v. X are within the limits $EX \pm \sigma_X$.
- That is why it is also said that the standard deviation is a measure of statistical dispersion, i.e. how widely spread the values in a dataset are.

The interval $EX \pm \sigma_X$ covers all the possible values of X only in a few isolated examples.

Example 2

- Suppose that X has the normal distribution with mean equal to a, $X \in N(a, \sigma_X)$.
- The probability of the interval $a \pm \sigma_X$ is 0.683. That is, when sampling from the corresponding distribution, 68.3% of the simulations will be in the interval $(a - \sigma_X, a + \sigma_X)$.
- The probabilities of the intervals $a \pm 2\sigma_X$ and $a \pm 3\sigma_X$ are 0.955 and 0.997 respectively. (See the illustration on the next slide).

Standard deviation

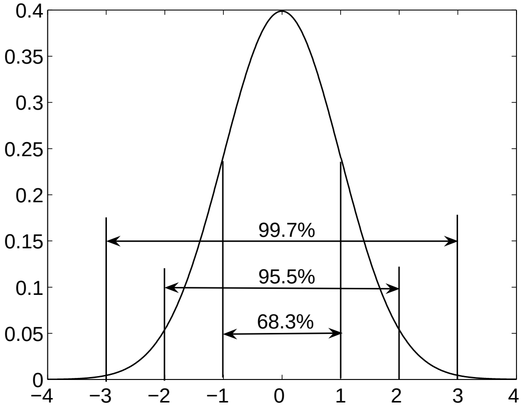

Figure: The standard normal density and the probabilities of the intervals $EX \pm \sigma_X$, $EX \pm 2\sigma_X$, and $EX \pm 3\sigma_X$, where $X \in N(0, 1)$, as a percentage of the total mass.

Standard deviation

$\Rightarrow$ The probabilities in the later example are specific for the normal distribution only.

- In the general case when the distribution of the r.v. X is unknown, we can obtain bounds on the probabilities by means of Chebyshev's inequality,

$$P(|X - EX| > x) \leq \frac{\sigma_X^2}{x^2}, \tag{2}$$

  provided that the r.v. X has a finite second moment, $E|X|^2 < \infty$.

- With the help of Chebyshev's inequality, we calculate that the probability of the interval $EX \pm k\sigma_X$, $k = 1, 2, \ldots$ exceeds $1 - 1/k^2$,

$$P(|X - EX| \leq k\sigma_X) \geq 1 - 1/k^2.$$

Standard deviation

- If we choose $k = 2$, we compute that $P(X \in EX \pm 2\sigma_X)$ is at least 0.75. Table 1 contains the corresponding bounds on the probabilities computed for several choices of k.

k	1.4	2	3	4	5	6	7
p_k	0.5	0.75	0.889	0.94	0.96	0.97	0.98

Table: The values $p_k = 1 - 1/k^2$ provide a lower bound for the probability $P(X \in EX \pm k\sigma_X)$ when the distribution of X is unknown.

Mean absolute deviation

- Even though the standard deviation is widely used, there are important cases where it is inappropriate — there are distributions for which the standard deviation is infinite.
- An example of an uncertainty measure also often used, which may be finite when the standard deviation does not exist, is the mean absolute deviation (MAD).

Mean absolute deviation

- MAD is defined as the average deviation in absolute terms around the mean of the distribution,

$$MAD_X = E|X - EX|, \tag{3}$$

  where X is a r.v. with finite mean.
- For a discrete distribution, equation (3) becomes

$$MAD_X = \sum_{k=1}^{n} |x_k - EX| p_k,$$

  where $x_k,\ k = 1, \ldots, n$, are the outcomes and $p_k,\ k = 1, \ldots, n$, are the corresponding probabilities.

Mean absolute deviation

- It is clear from the definition that both the positive and the negative deviations are taken into account in the MAD formula.
- The MAD is also a non-negative number and if it is equal to zero, then X is equal to its mean with probability one.
- The MAD and the standard deviation are two alternative measures estimating the uncertainty of a r.v. There are distributions, for which one of the quantities can be expressed from the other.

Example 3

If X has a normal distribution, $X \in N(a, \sigma_X^2)$, then

$$MAD_X = \sigma_X \sqrt{\frac{2}{\pi}}.$$

$\Rightarrow$ Thus, for the normal distribution case, the MAD is just a scaled standard deviation.

Semi-standard deviation

- The semi-standard deviation is a measure of dispersion which differs from the previous in that it takes into account only the positive or only the negative deviations from the mean. Therefore, it is not symmetric.
- The positive and the negative semi-standard deviations are:

$$\begin{aligned} \sigma_X^+ &= (E(X - EX)_+^2)^{1/2} \\ \sigma_X^- &= (E(X - EX)_-^2)^{1/2} \end{aligned} \tag{4}$$

where

$(x - EX)_+^2$	equals the squared difference between the outcome x and the mean EX if the difference is positive, $(x - EX)_+^2 = \max(x - EX, 0)^2$.
$(x - EX)_-^2$	equals the squared difference between the outcome x and the mean EX if the difference is negative, $(x - EX)_-^2 = \min(x - EX, 0)^2$.

Semi-standard deviation

- Thus, σ_X^+ takes into account only the positive deviations from the mean and it may be called an upside dispersion measure.
- Similarly, σ_X^- takes into account only the negative deviations from the mean and it may be called a downside dispersion measure.
- As with the standard deviation, both σ_X^- and σ_X^+ are non-negative numbers which are equal to zero if and only if the random variable equals its mean with probability one.

Semi-standard deviation

$\Rightarrow$ If the random variable is symmetric around the mean, then the upside and the downside semi-standard deviations are equal.

- For example, if X has a normal distribution, $X \in N(a, \sigma_X^2)$, then both quantities are equal and can be expressed by means of the standard deviation,

$$\sigma_X^- = \sigma_X^+ = \frac{\sigma_X}{\sqrt{2}}.$$

  If the distribution of X is skewed, then $\sigma_X^- \neq \sigma_X^+$.
- Positive skewness corresponds to larger positive semi-standard deviation, $\sigma_X^- < \sigma_X^+$.
- Similarly, negative skewness corresponds to larger negative semi-standard deviation, $\sigma_X^- > \sigma_X^+$.

- Measures of dispersion also include inter-quartile range and can be based on central absolute moments.
- The inter-quartile range is defined as the difference between the 75% and the 25% quantile.
- The central absolute moment of order k is defined as

$$m_k = E|X - EX|^k$$

and an example of a dispersion measure based on it is

$$(m_k)^{1/k} = (E|X - EX|^k)^{1/k}.$$

Axiomatic description

- The common properties of the dispersion measures we have considered can be synthesized into *axioms*.
- We denote the dispersion measure of a r.v. X by $D(X)$.

Positive shift	$D(X + C) \leq D(X)$ for all X and constants $C \geq 0$.
Positive homogeneity	$D(0) = 0$ and $D(\lambda X) = \lambda D(X)$ for all X and all $\lambda > 0$.
Positivity	$D(X) \geq 0$ for all X, with $D(X) > 0$ for non-constant X.

$\Rightarrow$ A dispersion measure is called any functional satisfying the axioms.

Axiomatic description

- According to the positive shift property, adding a positive constant does not increase the dispersion of a random variable.
- According to the positive homogeneity and the positivity properties, the dispersion measure D is equal to zero only if the random variable is a constant. Recall that it holds for the standard deviation, MAD, and semi-standard deviation.
- An example of a dispersion measure satisfying these properties is the colog measure defined by

$$\mathrm{colog}(X) = E(X \log X) - E(X)E(\log X).$$

where X is a positive random variable. The colog measure is sensitive to additive shifts and has applications in finance as it is consistent with the preference relations of risk-averse investors.

Deviation measures

- Rockafellar et al. (2006) provide an axiomatic description of dispersion measures. Their axioms define convex dispersion measures called deviation measures.
- Besides the axioms given above, the deviation measures satisfy the property

Sub-additivity	$D(X + Y) \leq D(X) + D(Y)$ for all X and Y.

and the positive shift property is replaced by

Translation invariance	$D(X + C) = D(X)$ for all X and constants $C \in \mathbb{R}$.

Deviation measures

- As a consequence of the translation invariance axiom, the deviation measure is influenced only by the difference $X - EX$.
- If $X = EX$ in all states of the world, then the deviation measure is a constant and, therefore, it is equal to zero because of the positivity axiom.
- Conversely, if $D(X) = 0$, then $X = EX$ in all states of the world.
- The positive homegeneity and the sub-additivity axioms establish the convexity property of $D(X)$.
- Not all deviation measures are symmetric; that is, it is possible to have $D(X) \neq D(-X)$ if the r.v. X is not symmetric. This is an advantage because an investment manager is more attentive to the negative deviations from the mean.
- Examples of asymmetric deviation measures include the semi-standard deviation, σ_X^- defined in equation (4).

Deviation measures

- Deviation measures which depend only on the negative deviations from the mean are called downside deviation measures.
- The quantity $\tilde{D}(X)$ is a symmetric deviation measure if we define it as

$$\tilde{D}(X) := \frac{1}{2}(D(X) + D(-X)),$$

  where $D(X)$ is an arbitrary deviation measure.
- A downside deviation measure possesses several of the characteristics of a risk measure but it is not a risk measure. Consider the following example.

Deviation measures

Example 4

- Suppose that we have initially in our portfolio a common stock, X, with a current market value of \$95 and an expected return of 0.5% in a month.
- Let us choose one particular deviation measure, D_1, and compute $D_1(r_X) = 20\%$, where r_X stands for the portfolio return.
- Assume that we add to our portfolio a risk-free government bond, B, worth \$95 with a face value of \$100 and a one-month maturity. The return on the bond equals $r_B = \$5/\$95 = 5.26\%$ and is non-random. The return of our portfolio then equals $r_p = r_X/2 + r_B/2$.
- Using the axioms, $D_1(r_p) = D_1(r_X)/2 = 10\%$. The uncertainty of r_p decreases twice since the share of the risky stock decreases twice. Intuitively, the risk of r_p decreases more than twice if compared to r_X because half of the new portfolio earns a sure profit of 5.26%.
- This effect is due to the translation invariance which makes the deviation measure insensitive to non-random profit.

Probability metrics and dispersion measures

- Probability metrics are functionals which are constructed to measure distances between random quantities. Thus, every probability metric involves two random variables X and Y, and the distance between them is denoted by $\mu(X, Y)$ where μ stands for the probability metric.
- Suppose that μ is a compound probability metric. In this case, if $\mu(X, Y) = 0$, it follows that the two random variables are coincident in all states of the world. Therefore, the quantity $\mu(X, Y)$ can be interpreted as a measure of relative deviation between X and Y.
- A positive distance, $\mu(X, Y) > 0$, means that the two variables fluctuate with respect to each other and zero distance, $\mu(X, Y) = 0$, implies that there is no deviation of any of them relative to the other.

- The functional $\mu(X, EX)$ (the distance between X and the mean of X) provides a very general notion of a dispersion measure as it arises as a special case from a probability metric which represents the only general way of measuring distances between random quantities.

Measures of risk

- A risk measure may share some of the features of a dispersion measure but is, generally, a different object.
- Markowitz (1952) was the first to recognize the relationship between risk and reward and introduced standard deviation as a proxy for risk.
- The standard deviation is not a good choice for a risk measure because it penalizes symmetrically both the negative and the positive deviations from the mean. It is an uncertainty measure and cannot account for the asymmetric nature of risk, i.e. risk concerns losses only.
- Later, Markowitz (1959) suggested the semi-standard deviation as a substitute. But, as we've already stated, any deviation measure cannot be a true risk measures.

Value-at-risk

- A risk measure which has been widely accepted since 1990s is the value-at-risk (VaR).
- In the late 1980s, it was integrated by JP Morgan on a firmwide level into its risk-management system. In this system, they developed a service called RiskMetrics which was later spun off into a separate company called RiskMetrics Group.
- It is usually thought that JP Morgan invented the VaR measure. In fact, similar ideas had been used by large financial institutions in computing their exposure to market risk.

Value-at-risk

- In the mid 1990s, the VaR measure was approved by regulators as a valid approach to calculating capital reserves needed to cover market risk.
- The Basel Commettee on Banking Supervision released a package of amendments to the requirements for banking institutions allowing them to use their own internal systems for risk estimation.
- Capital reserves could be based on the VaR numbers computed internally by an in-house risk management system. Regulators demand that the capital reserve equal the VaR number multiplied by a factor between 3 and 4.
- Regulators link the capital reserves for market risk directly to the risk measure.

Value-at-risk

- VaR is defined as the minimum level of loss at a given, sufficiently high, confidence level for a predefined time horizon.
- The recommended confidence levels are 95% and 99%.

Example 5

Suppose that we hold a portfolio with a 1-day 99% VaR equal to $1 million. This means that over the horizon of 1 day, the portfolio may lose more than $1 million with probability equal to 1%.

Example 6

Suppose that the present value of a portfolio we hold is $10 million. If the 1-day 99% VaR of the return distribution is 2%, then over the time horizon of 1 day, we lose more than 2% ($200,000) of the portfolio present value with probability equal to 1%.

Value-at-risk

- Denote by $(1-\epsilon)100\%$ the confidence level parameter of the VaR. As we explained, losses larger than the VaR occur with probability ϵ. The probability ϵ, we call tail probability.
- Depending on the interpretation of the random variable, VaR can be defined in different ways. Formally, the VaR at confidence level $(1-\epsilon)100\%$ (tail probability ϵ) is defined as the negative of the lower ϵ-quantile of the return distribution,

$$VaR_{\epsilon}(X) = -\inf_{x}\{x|P(X \leq x) \geq \epsilon\} = -F_X^{-1}(\epsilon) \tag{5}$$

  where $\epsilon \in (0,1)$ and $F_X^{-1}(\epsilon)$ is the inverse of the distribution function.
- If the r.v. X describes random returns, then the VaR number is given in terms of a return figure. (See the illustration on the next slide).

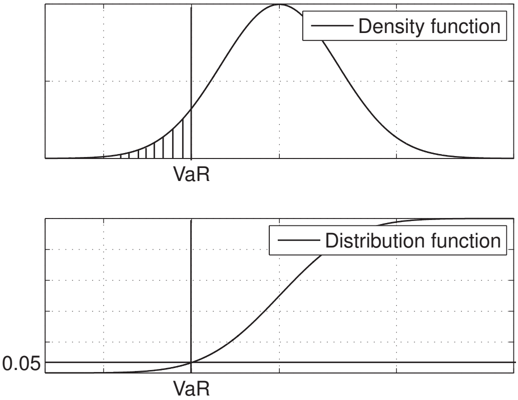

Figure: The VaR at 95% confidence level of a random variable X. The top plot shows the density of X, the marked area equals the tail probability, and the bottom plot shows the distribution function.

Value-at-risk

- If X describes random payoffs, then VaR is a threshold in dollar terms below which the portfolio value falls with probability ϵ,

$$VaR_\epsilon(X) = \inf_x \{x | P(X \leq x) \geq \epsilon\} = F_X^{-1}(\epsilon) \tag{6}$$

  where $\epsilon \in (0, 1)$ and $F_X^{-1}(\epsilon)$ is the inverse of the distribution function of the random payoff.
- VaR can also be expressed as a distance to the present value when considering the profit distribution.

- The random profit is defined as $X - P_0$ where X is the payoff and P_0 is the present value.
- The VaR of the random profit equals,

$$VaR_\epsilon(X - P_0) = -\inf_x \{x | P(X - P_0 \leq x) \geq \epsilon\} = P_0 - VaR_\epsilon(X)$$

  in which $VaR_\epsilon(X)$ is defined according to (6) since X is interpreted as a random payoff. In this case, the definition of VaR is essentially given by equation (5).
- If $VaR_\epsilon(X)$ in (5) is a negative number, then at tail probability ϵ we do not observe losses but profits. Losses happen with even smaller probability than ϵ, so the r.v. X bears no risk.

Value-at-risk

We illustrate one aspect in which VaR differs from the deviation measures and all uncertainty measures.

- As a consequence of the definition, if we add to the r.v. X a non-random profit C, the resulting VaR can be expressed by the VaR of the initial variable in the following way

$$VaR_\epsilon(X + C) = VaR_\epsilon(X) - C. \tag{7}$$

- Thus, adding a non-random profit decreases the risk of the portfolio. Furthermore, scaling the return distribution by a positive constant λ scales the VaR by the same constant,

$$VaR_\epsilon(\lambda X) = \lambda VaR_\epsilon(X). \tag{8}$$

- It turns out that these properties characterize not only VaR. They are identified as key features of a risk measure.

Recall the Example 4.

Example 7

- Initially, the portfolio we hold consists of a common stock with random monthly return r_X. We rebalance the portfolio so that it becomes an equally weighted portfolio of the stock and a bond with a non-random monthly return of 5.26%, $r_B = 5.26\%$. Thus, the portfolio return can be expressed as

$$r_p = r_X(1/2) + r_B(1/2) = r_X/2 + 0.0526/2.$$

- Using equations (7) and (8), we calculate that if $VaR_\epsilon(r_X) = 12\%$, then $VaR_\epsilon(r_p) \approx 3.365\%$ which is by far less than 6% — half of the initial risk.
- Recall that any deviation measure would indicate that the dispersion (or the uncertainty) of the portfolio return r_p would be twice as smaller than the uncertainty of r_X.

Value-at-risk

- The performance of VaR (as well as any other risk measure) is heavily dependent on the assumed probability distribution of the variable X.
- If we use VaR to build reserves in order to cover losses in times of crises, then underestimation may be fatal and overestimation may lead to inefficient use of capital.
- An inaccurate model is even more dangerous in an optimal portfolio problem in which we minimize risk subject to some constraints, as it may adversely influence the optimal weights and therefore not reduce the true risk.

Generally, VaR should be abandoned as a risk measure.
The most important drawback is:

- In some cases, the reasonable diversification effect that every portfolio manager should expect to see in a risk measure is not present; that is, the VaR of a portfolio may be greater than the sum of the VaRs of the constituents,

$$VaR_{\epsilon}(X+Y) > VaR_{\epsilon}(X) + VaR_{\epsilon}(Y), \tag{9}$$

  in which X and Y stand for the random payoff of the instruments in the portfolio.

$\Rightarrow$ This shows that VaR cannot be a true risk measure.

Value-at-risk

Let's show that VaR may satisfy (9).

Example 8

- Suppose that X denotes a bond which either defaults with probability 4.5% and we lose \$50 or it does not default (the loss is zero).
- Let Y be the same bond but assume that the defaults of the two bonds are independent events. The VaR of the two bonds 5% tail probability:

$$VaR_{0.05}(X) = VaR_{0.05}(Y) = 0.$$

- VaR fails to recognize losses occurring with probability smaller than 5%.
- A portfolio of the two bonds has the following payoff profile: it loses \$100 with probability of about 0.2%, loses \$50 with probability of about 8.6%, and the loss is zero with probability 91.2%. Thus, the corresponding 95% VaR of the portfolio equals \$50 and clearly,

$$\$50 = VaR_{0.05}(X+Y) > VaR_{0.05}(X) + VaR_{0.05}(Y) = 0.$$

What are the consequences of using a risk measure which may satisfy property (9)?

- It is going to mislead portfolio managers that there is no diversification effect in the portfolio and they may make the irrational decision to concentrate it only into a few positions. As a consequence, the portfolio risk actually increases.

- Another drawback is that VaR is not very informative about losses beyond the VaR level. It only reports that losses larger than the VaR level occur with probability equal to ϵ but it does not provide any information about the likely magnitude of such losses, for example.
- VaR is not a useless concept to be abandoned altogether. For example, it can be used in risk-reporting only as a characteristic of the portfolio return (payoff) distribution since it has a straightforward interpretation.
- The criticism of VaR is focused on its wide application by practitioners as a true risk measure which, in view of the deficiencies described above, is not well grounded and should be reconsidered.

Computing portfolio VaR in practice

- Suppose that a portfolio contains *n* common stocks and we are interested in calculating the daily VaR at 99% confidence level.
- Denote the random daily returns of the stocks by $X_1, \ldots, X_n$ and by $w_1, \ldots, w_n$ the weight of each stock in the portfolio. Thus, the portfolio return r_p can be calculated as

$$r_p = w_1 X_1 + w_2 X_2 + \ldots + w_n X_n.$$

- The portfolio VaR is derived from the distribution of r_p. The three approaches vary in the assumptions they make.

The approach of RiskMetrics

- The approach of RiskMetrics Group is centered on the assumption that the stock returns have a multivariate normal distribution. Under this assumption, the distribution of the portfolio return is also normal.
- In order to calculate the portfolio VaR, we only have to calculate the expected return of r_p and the standard deviation of r_p.
- The 99% VaR will appear as the negative of the 1% quantile of the $N(Er_p, \sigma^2_{r_p})$ distribution.
- The portfolio expected return can be directly expressed through the expected returns of the stocks,

$$Er_p = w_1 EX_1 + w_2 EX_2 + \ldots + w_n EX_n = \sum_{k=1}^{n} w_k EX_k, \tag{10}$$

where E denotes mathematical expectation.

The approach of RiskMetrics

- Similarly, the variance of the portfolio return $\sigma^2_{r_p}$ can be computed through the variances of the stock returns and their covariances,

$$\sigma^2_{r_p} = w_1^2\sigma^2_{X_1} + w_2^2\sigma^2_{X_2} + \ldots + w_n^2\sigma^2_{X_n} + \sum_{i \neq j} w_i w_j \operatorname{cov}(X_i, X_j),$$

  in which the last term appears because we have to sum up the covariances between all pairs of stock returns.

- There is a more compact way of writing down the expression for $\sigma^2_{r_p}$ using matrix notation,

$$\sigma^2_{r_p} = w'\Sigma w, \tag{11}$$

  in which $w = (w_1, \ldots, w_n)$ is the vector of portfolio weights

The approach of RiskMetrics

- Σ is the covariance matrix of stock returns,

$$\Sigma = \begin{pmatrix} \sigma^2_{X_1} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma^2_{X_2} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma^2_{X_n} \end{pmatrix},$$

  in which σ_{ij}, $i \neq j$, is the covariance between X_i and X_j, $\sigma_{ij} = \mathrm{cov}(X_i, X_j)$.

- As a result, we obtain that the portfolio return has a normal distribution with mean given by equation (10) and variance given by equation (11).

The approach of RiskMetrics

- The standard deviation is the scale parameter of the normal distribution and the mean is the location parameter.
- Due to the normal distribution properties, if $r_p \in N(Er_p, \sigma^2_{r_p})$, then

$$\frac{r_p - Er_p}{\sigma_{r_p}} \in N(0,1).$$

- Because of the properties (7) and (8) of the VaR, the 99% portfolio VaR can be represented as,

$$VaR_{0.01}(r_p) = q_{0.99}\sigma_{r_p} - Er_p \tag{12}$$

where the standard deviation of the portfolio return σ_{r_p} is computed from equation (11), the expected portfolio return Er_p is given in (10), and $q_{0.99}$ is the 99% quantile of the standard normal distribution. Note that $q_{0.99}$ is a quantity independent of the portfolio composition

The approach of RiskMetrics

- The parameters which depend on the portfolio weights are the standard deviation of portfolio returns σ_{r_p} and the expected portfolio return.
- VaR under the assumption of normality is symmetric even though, by definition, VaR is centered on the left tail of the distribution; that is, VaR is asymmetric by construction. This result appears because the normal distribution is symmetric around the mean.
- The approach of RiskMetrics can be extended for other types of distributions: Student's t and stable distributions.

The Historical Method

- The historical method does not impose any distributional assumptions; the distribution of portfolio returns is constructed from historical data. Hence, sometimes the historical simulation method is called a non-parametric method.
- For example, the 99% daily VaR of the portfolio return is computed as the negative of the empirical 1% quantile of the observed daily portfolio returns. The observations are collected from a predetermined time window such as the most recent business year.

The Historical Method

While the historical method seems to be more general as it is free of any distributional hypotheses, it has a number of major drawbacks:

1. It assumes that the past trends will continue in the future. This is not a realistic assumption because we may experience extreme events in the future, which have not happened in the past.
2. It treats the observations as independent and identically distributed (i.i.d.) which is not realistic. The daily returns data exhibits clustering of the volatility phenomenon, autocorrelations and so on, which are sometimes a significant deviation from the i.i.d. assumption.
3. It is not reliable for estimation of VaR at very high confidence levels. A sample of one year of daily data contains 250 observations which is a rather small sample for the purpose of the 99% VaR estimation.

The Hybrid Method

- The hybrid method is a modification of the historical method in which the observations are not regarded as i.i.d. but certain weights are assigned to them depending on how close they are to the present.
- The weights are determined using the exponential smoothing algorithm. The exponential smoothing accentuates the most recent observations and seeks to take into account time-varying volatility phenomenon.

The Hybrid Method

The algorithm of the hybrid approach consists of the following steps.

1. Exponentially declining weights are attached to historical returns, starting from the current time and going back in time. Let $r_{t-k+1}, \ldots, r_{t-1}, r_t$ be a sequence of k observed returns on a given asset, where t is the current time. The i-th observation is assigned a weight
$$\theta_i = c^* \lambda^{t-i},$$
where $0 < \lambda < 1$, and $c = \frac{1-\lambda}{1-\lambda^k}$ is a constant chosen such that the sum of all weights is equal to one, $\sum \theta_i = 1$.
2. Similarly to the historical simulation method, the hypothetical future returns are obtained from the past returns and sorted in increasing order.
3. The VaR measure is computed from the empirical c.d.f. in which each observation has probability equal to the weight θ_i.

The Hybrid Method

- Generally, the hybrid approach is appropriate for VaR estimation of heavy-tailed time series.
- It overcomes, to some degree, the first and the second deficiency of the historical method but it is also not reliable for VaR estimation of very high confidence levels.

The Monte Carlo Method

- The Monte Carlo method requires specification of a statistical model for the stocks returns. The statistical model is multivariate, hypothesizing both the behavior of the stock returns on a stand-alone basis and their dependence.
- For instance, the multivariate normal distribution assumes normal distributions for the stock returns viewed on a stand-alone basis and describes the dependencies by means of the covariance matrix.
- The multivariate model can also be constructed by specifying explicitly the one-dimensional distributions of the stock returns, and their dependence through a copula function.

The Monte Carlo Method

The Monte Carlo method consists of the following basic steps:

Step 1. *Selection of a statistical model.* The statistical model should be capable of explaining a number of observed phenomena in the data, e.g. heavy-tails, clustering of the volatility, etc., which we think influence the portfolio risk.

Step 2. *Estimation of the statistical model parameters.* A sample of observed stocks returns is used from a predetermined time window, for instance the most recent 250 daily returns.

Step 3. *Generation of scenarios from the fitted model.* Independent scenarios are drawn from the fitted statistical model. Each scenario is a vector of stock returns.

Step 4. *Calculation of portfolio risk.* Compute portfolio risk on the basis of the portfolio return scenarios obtained from the previous step.

The advantage of Monte Carlo method:

- It does not require any closed-form expressions and, by choosing a flexible statistical model, accurate risk numbers can be obtained.

The disadvantage of Monte Carlo method:

- The computed portfolio VaR is dependent on the generated sample of scenarios and will fluctuate a little if we regenerate the sample. This side effect can be reduced by generating a larger sample. An illustration is provided in the following example.

The Monte Carlo Method

- Suppose that the daily portfolio return distribution is standard normal and, therefore, at Step 4 of the algorithm we have scenarios from the standard normal distribution. So that, we can compute the 99% daily VaR directly from formula (12).
- Nevertheless, we will use the Monte Carlo method to gain more insight into the deviations of the VaR based on scenarios from the VaR computed according to formula (12).
- In order to investigate how the fluctuations of the 99% VaR change about the theoretical value, we generate samples of different sizes: 500, 1,000, 5,000, 10,000, 20,000, and 100,000 scenarios.
- The 99% VaR is computed from these samples and the numbers are stored. We repeat the experiment 100 times. In the end, we have 100 VaR numbers for each sample size.
- We expect that as the sample size increases, the VaR values will fluctuate less about the theoretical value which is $VaR_{0.01}(X) = 2.326$, $X \in N(0, 1)$.

Table below contains the result of the experiment.

Number of Scenarios	99% VaR	95% confidence interval
500	2.067	[1.7515, 2.3825]
1,000	2.406	[2.1455, 2.6665]
5,000	2.286	[2.1875, 2.3845]
10,000	2.297	[2.2261, 2.3682]
20,000	2.282	[2.2305, 2.3335]
50,000	2.342	[2.3085, 2.3755]
100,000	2.314	[2.2925, 2.3355]

Table: The 99% VaR of the standard normal distribution computed from a sample of scenarios. The 95% confidence interval is calculated from 100 repetitions of the experiment. The true value is $VaR_{0.01}(X) = 2.326$.

The Monte Carlo Method

- From the 100 VaR numbers, we calculate the 95% confidence interval for the true value given in the third column.
- The confidence intervals cover the theoretical value 2.326 and also we notice that the length of the confidence interval decreases as the sample size increases. This effect is best illustrated with the help of the boxplot diagrams[1] shown on the next slide.
- A sample of 100,000 scenarios results in VaR numbers which are tightly packed around the true value while a sample of only 500 scenarios may give a very inaccurate estimate.

[1] A boxplot, or a box-and-whiskers diagram, is a convenient way of depicting several statistical characteristics of the sample. The size of the box equals the difference between the third and the first *quartile* (75% quantile – 25% quantile), also known as the *interquartile range*. The line in the box corresponds to the median of the data (50% quantile). The lines extending out of the box are called *whiskers* and each of them is long up to 1.5 times the interquartile range. All observations outside the whiskers are labeled outliers and are depicted by a plus sign.

The Monte Carlo Method

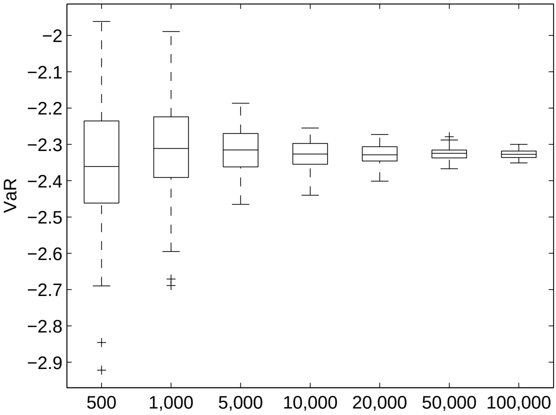

Figure: Boxplot diagrams of the fluctuation of the 99% VaR of the standard normal distribution based on scenarios. The horizontal axis shows the number of scenarios and the boxplots are computed from 100 independent samples.

The Monte Carlo Method

- This simple experiment shows that the number of scenarios in the Monte Carlo method has to be carefully chosen.
- The approach we used to determine the fluctuations of the VaR based on scenarios is a statistical method called parametric bootstrap.
- The bootstrap methods in general are powerful statistical methods which are used to compute confidence intervals when the problem is not analytically tractable but the calculations may be quite computationally intensive.
- When the portfolio contains complicated instruments such as derivatives, it is no longer possible to use a closed-form expression for the portfolio VaR (and any risk measure in general) because the distribution of portfolio return (or payoff) becomes quite arbitrary.

The Monte Carlo Method

- The practical implementation of Monte Carlo method is a very challenging endeavor from both software development and financial modeling point of view.
- The portfolios of big financial institutions often contain products which require yield curve modeling, development of fundamental and statistical factor models, and, on top of that, a probabilistic model capable of describing the heavy tails of the risk-driving factor returns, the autocorrelation, clustering of the volatility, and the dependence between these factors.
- Processing large portfolios is related to manipulation of colossal data structures which requires excellent skills of software developers in order to be efficiently performed.

Back-testing of VaR

- Suppose that we calculate the 99% daily portfolio VaR. This means that according to our assumption for the portfolio return (payoff) distribution, the portfolio loses more than the 99% daily VaR with 1% probability.
- The question is whether this estimate is correct; that is, *does the portfolio really lose more than this amount with 1% probability?*
- This question can be answered by back-testing of VaR.

Back-testing of VaR

Generally, the procedure consists of the following steps:

Step 1. Choose a time window for the back-testing. Usually the time window is the most recent one or two years.

Step 2. For each day in the time window, calculate the VaR number.

Step 3. Check if the loss on a given day is below or above the VaR number computed the day before. If the observed loss is larger, then we say that there is a case of an *exceedance*. (See the illustration on the next slide).

Step 4. Count the number of exceedances. Check if there are too many or too few of them by verifying if the number of exceedances belong to the corresponding 95% confidence interval.

Back-testing of VaR

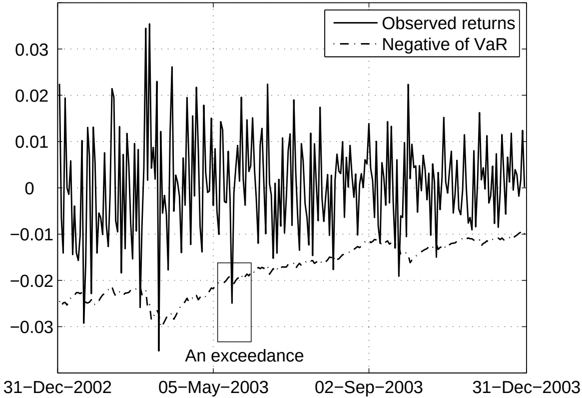

Figure: The observed daily returns of S&P 500 index between December 31, 2002 and December 31, 2003 and the negative of VaR. The marked observation is an example of an exceedance.

Back-testing of VaR

- If in Step 4 we find out that there are too many number of exceedances, so that losses exceeding the corresponding VaR happen too frequently. If capital reserves are determined on the basis of VaR, then there is a risk of being incapable of covering large losses.
- Conversely, if the we find out that there are too few number of exceedances, then the VaR numbers are too pessimistic. This is also an undesirable situation. Note that the actual size of the exceedances is immaterial, we only count them.

Back-testing of VaR

- The confidence interval for the number of exceedances is constructed using the indicator-type events “we observe an exceedance”, “we do not observe an exceedance” on a given day.
- Let us associate a number with each of the events similar to a coin tossing experiment. If we observe an exceedance on a given day, then we say that the number 1 has occurred, otherwise 0 has occurred.
- If the back-testing time window is two years, then we have a sequence of 500 zeros and ones and the expected number of exceedances is 5. Thus, finding the 95% confidence interval for the number of exceedances reduces to finding an interval around 5 such that the probability of the number of ones belonging to this interval is 95%.

Back-testing of VaR

- If we assume that the corresponding events are independent, then there is a complete analogue of this problem in terms of coin tossing.
- We toss independently 500 times an unfair coin with probability of success equal to 1%. *What is the range of the number of success events with 95% probability?*
- In order to find the 95% confidence interval, we can resort to the normal approximation to the binomial distribution. The formula is,

$$\text{left bound} = N\epsilon - F^{-1}(1 - 0.05/2)\sqrt{N\epsilon(1-\epsilon)}$$
$$\text{right bound} = N\epsilon + F^{-1}(1 - 0.05/2)\sqrt{N\epsilon(1-\epsilon)}$$

  where N is the number of indicator-type events, ϵ is the tail probability of the VaR, and $F^{-1}(t)$ is the inverse distribution function of the standard normal distribution.
- In the example, $N = 500$, $\epsilon = 0.01$, and the 95% confidence interval for the number of exceedances is [0, 9]. If we are back-testing the 95% VaR, under the same circumstances the confidence interval is [15, 34].

Back-testing of VaR

- Note that the statistical test based on the back-testing of VaR at a certain tail probability cannot answer the question if the distributional assumptions for the risk-driving factors are correct in general.
- For instance, if the portfolio contains only common stocks, then we presume a probabilistic model for stocks returns.
- By back-testing the 99% daily VaR of portfolio return, we verify if the probabilistic model is adequate for the 1% quantile of the portfolio return distribution; that is, we are back-testing if a certain point in the left tail of the portfolio return distribution is sufficiently accurately modeled.

Coherent risk measures

Can we find a set of desirable properties that a risk measure should satisfy?

- An answer is given by *Artzner et. al.* (1998). They provide an axiomatic definition of a functional which they call a coherent risk measure.
- We denote the risk measure by the functional $\rho(X)$ assigning a real-valued number to a random variable. Usually, the r.v. X is interpreted as a random payoff.
- In the remarks below each axiom, we provide an alternative interpretation which holds if X is interpreted as random return.

Monotonicity $\qquad \rho(Y) \leq \rho(X)$, if $Y \geq X$ in almost sure sense.

Monotonicity states that if investment A has random return (payoff) Y which is not less than the return (payoff) X of investment B at a given horizon in all states of the world, then the risk of A is not greater than the risk of B.

This is quite intuitive but it really does matter whether the random variables represent random return or profit because an inequality in almost sure sense between random returns may not translate into the same inequality between the corresponding random profits and vice versa.

- Suppose that X and Y describe the random percentage returns on two investments A and B and let $Y = X + 3\%$. Apparently, $Y > X$ in all states of the world. The corresponding payoffs are obtained according to the equations

$$\begin{aligned}\text{Payoff}(X) &= I_A(1 + X)\\ \text{Payoff}(Y) &= I_B(1 + Y) = I_B(1 + X + 3\%)\end{aligned}$$

  where I_A is the initial investment in opportunity A and I_B is the initial investment in opportunity B.
- If the initial investment I_A is much larger than I_B, then $\text{Payoff}(X) > \text{Payoff}(Y)$ irrespective of the inequality $Y > X$.
- In effect, investment A may seem less risky than investment B it terms of payoff but in terms of return, the converse may hold.

Positive Homogeneity — $\rho(0) = 0$, $\rho(\lambda X) = \lambda\rho(X)$, for all X and all $\lambda > 0$.

The positive homogeneity property states that scaling the return (payoff) of the portfolio by a positive factor scales the risk by the same factor.

The interpretation for payoffs is obvious — if the investment in a position doubles, so does the risk of the position. We give a simple example illustrating this property when X stands for a random percentage return.

Coherent risk measures

The positive homogeneity property

- Suppose that today the value of a portfolio is I_0 and we add a certain amount of cash C. Then the value of our portfolio is $I_0 + C$.
- The value tomorrow is random and equals $I_1 + C$ in which I_1 is the random payoff. The return of the portfolio equals

$$X = \frac{I_1 + C - I_0 - C}{I_0 + C} = \frac{I_1 - I_0}{I_0}\left(\frac{I_0}{I_0 + C}\right)$$
$$= h\frac{I_1 - I_0}{I_0} = hY$$

  where $h = I_0/(I_0 + C)$ is a positive constant.
- The axiom positive homogeneity property implies that $\rho(X) = h\rho(Y)$; that is, the risk of the new portfolio will be the risk of the portfolio without the cash but scaled by h.

Sub-additivity $\rho(X+Y) \leq \rho(X) + \rho(Y)$, for all X and Y.

If X and Y describe random payoffs, then the sub-additivity property states that the risk of the portfolio is not greater than the sum of the risks of the two random payoffs.

- The positive homogeneity property and the sub-additivity property imply that the functional is convex

$$\begin{aligned}\rho(\lambda X + (1-\lambda)Y) &\leq \rho(\lambda X) + \rho((1-\lambda)Y)\\ &= \lambda\rho(X) + (1-\lambda)\rho(Y)\end{aligned}$$

  where $\lambda \in [0, 1]$.
- If X and Y describe random returns, then the random quantity $\lambda X + (1-\lambda)Y$ stands for the return of a portfolio composed of two financial instruments with returns X and Y having weights λ and $1-\lambda$ respectively.
- The convexity property states that the risk of a portfolio is not greater than the sum of the risks of its constituents, meaning that it is the convexity property which is behind the diversification effect that we expect in the case of X and Y denoting random returns.

Invariance $\rho(X + C) = \rho(X) - C$, for all X and $C \in \mathbb{R}$.

The invariance property has various labels. Originally, it was called *translation invariance* while in other texts it is called *cash invariance*.

- If X describes a random payoff, then the invariance property suggests that adding cash to a position reduces its risk by the amount of cash added.
- This is motivated by the idea that the risk measure can be used to determine capital requirements.
- As a consequence, the risk measure $\rho(X)$ can be interpreted as the minimal amount of cash necessary to make the position free of any capital requirements,

$$\rho(X + \rho(X)) = 0.$$

The invariance property has a different interpretation when X describes random return.

- Suppose that the r.v. X describes the return of a common stock and we build a long-only portfolio by adding a government bond yielding a risk-free rate r_B.
- The portfolio return equals $wX + (1-w)r_B$, where $w \in [0,1]$ is the weight of the common stock in the portfolio. Note that the quantity $(1-w)r_B$ is non-random by assumption.
- The invariance property states that the risk of the portfolio can be decomposed as

$$\begin{aligned} \rho(wX + (1-w)r_B) &= \rho(wX) - (1-w)r_B \\ &= w\rho(X) - (1-w)r_B \end{aligned} \tag{13}$$

where the second equality appears because of the positive homogeneity property.

The risk measure admits the following interpretation:

- Assume that the constructed portfolio is equally weighted, i.e. $w = 1/2$, then the risk measure equals the level of the risk-free rate such that the risk of the equally weighted portfolio consisting of the risky asset and the risk-free asset is zero.
- The investment in the risk-free asset will be, effectively, the reserve investment.

Alternative interpretations are also possible.

- Suppose that the present value of the position with random percentage return X is I_0. Assume that we can find a government security earning return r_B^* at the horizon of interest.
- *How much should we reallocate from* I_0 *and invest in the government security in order to hedge the risk* $\rho(X)$*?*
- The needed capital C should satisfy the equation

$$\frac{I_0 - C}{I_0}\rho(X) - \frac{C}{I_0} r_B^* = 0$$

which is merely a re-statement of equation (13) with the additional requirement that the risk of the resulting portfolio should be zero. The solution is

$$C = I_0 \frac{\rho(X)}{\rho(X) + r_B^*}.$$

- Note that if in the invariance property the constant is non-negative, $C \geq 0$, then it follows that $\rho(X + C) \leq \rho(X)$. This result is in agreement with the monotonicity property as $X + C \geq X$.
- The invariance property can be regarded as an extension of the monotonicity property when the only difference between X and Y is in their means.
- According to the discussion in the previous section, VaR is not a coherent risk measure because it may violate the sub-additivity property.

- An example of a coherent risk measure is the AVaR defined as the average of the VaRs which are larger than the VaR at a given tail probability ϵ.
- The accepted notation is $AVaR_\epsilon(X)$ in which ϵ stands for the tail probability level.
- A larger family of coherent risk measures is the family of spectral risk measures which includes the AVaR as a representative.
- The spectral risk measures are defined as weighted averages of VaRs.

Risk measures and dispersion measures

- Both classes, the deviation measures and the coherent risk measures, are not the only classes capable of quantifying statistical dispersion and risk respectively.[2]
- They describe basic features of uncertainty and risk and, in effect, we may expect that a relationship between them exists.

[2] See the appendix to the lecture, describing the convex risk measures.

Risk measures and dispersion measures

- The common properties are the sub-additivity property and the positive homogeneity property.
- The specific features are the monotonicity property and the invariance property of the coherent risk measures and the translation invariance and positivity of deviation measures.
- The link between them concerns a sub-class of the coherent risk measures called strictly expectation bounded risk measures and a sub-class of the deviation measures called lower range dominated deviation measures.
- This link has an interesting implication for constructing optimal portfolios[3].

[3] Will be discussed in Lecture 8

Risk measures and dispersion measures

- A coherent risk measure $\rho(X)$ is called strictly expectation bounded if it satisfies the condition

$$\rho(X) > -EX \quad (14)$$

for all non-constant X, in which EX is the mathematical expectation of X.

- If X describes the portfolio return distribution, then the inequality in (14) means that the risk of the portfolio is always greater than the negative of the expected portfolio return.
- A coherent risk measure satisfying this condition is the AVaR, for example.

Risk measures and dispersion measures

- A deviation measure $D(X)$ is called lower-range dominated if it satisfies the condition

$$D(X) \leq EX \tag{15}$$

  for all non-negative random variables, $X \geq 0$.
- A deviation measure which is lower range dominated is, for example, the downside semi-standard deviation σ_X^- defined in (4).

Risk measures and dispersion measures

- The relationship between the two sub-classes is a one-to-one correspondence between them established through the equations

$$D(X) = \rho(X - EX) \tag{16}$$

and

$$\rho(X) = D(X) - E(X). \tag{17}$$

- That is, if $\rho(X)$ is a strictly expectation bounded coherent risk measure, then through the formula in (16) we obtain the corresponding lower range dominated deviation measure and, conversely, through the formula in (17), we obtain the corresponding strictly expectation bounded coherent risk measure.

Risk measures and dispersion measures

- There is a deviation measure behind each strictly expectation bounded coherent risk measure.
- Consider the AVaR for instance. Since it satisfies the property in (14), according to the relationship discussed above, the quantity

$$D_\epsilon(X) = AVaR_\epsilon(X - EX)$$

  represents the deviation measure underlying the AVaR risk measure at tail probability ϵ.
- The quantity $D_\epsilon(X)$, as well as any other lower range dominated deviation measure, is obtained by computing the risk of the centered random variable.

Risk measures and stochastic orders

- Suppose that we estimate the risk of X and Y through a risk measure ρ.
- *If all risk-averse investors prefer X to Y, then does it follow that $\rho(X) \leq \rho(Y)$?*
- This question describes the issue of consistency of a risk measure with the SSD order.
- Intuitively, a realistic risk measure should be consistent with the SSD order since there is no reason to assume that an investment with higher risk as estimated by the risk measure will be preferred by all risk-averse investors.

- The monotonicity property of the coherent risk measures implies consistency with first-order stochastic dominance (FSD).
- The condition that $X \geq Y$ in all states of the world translates into the following inequality in terms of the c.d.f.s,

$$F_X(x) \leq F_Y(x), \quad \forall x \in \mathbb{R},$$

  which, in fact, characterizes the FSD order.
- If all non-satiable investors prefer X to Y, then any coherent risk measure will indicate that the risk of X is below the risk of Y.

Risk measures and stochastic orders

- The defining axioms of the coherent risk measures cannot guarantee consistency with the SSD order. Therefore, if we want to use a coherent risk measure in practice, we have to verify separately the consistency with the SSD order.
- DeGiorgi(2005) shows that the AVaR, and spectral risk measures in general, are consistent with the SSD order.
- Note that if the AVaR, for example, is used to measure the risk of portfolio return distributions, then the corresponding SSD order concerns random variables describing returns.
- Similarly, if the AVaR is applied to random variables describing payoff, then the SSD order concerns random payoffs.

$\Rightarrow$ SSD orders involving returns do not coincide with SSD orders involving payoffs.

Summary

We discussed in detail the following dispersion measures:

1. the standard deviation
2. the mean absolute deviation
3. the upside and downside semi-standard deviations
4. an axiomatic description of dispersion measures
5. the family of deviation measures

We also discussed in detail the following risk measures:

1. the value-at-risk
2. the family of coherent risk measures

Summary

We emphasized that a realistic statistical model for risk estimation includes two essential components:

- a realistic statistical model for the financial assets return distributions and their dependence, capable of accounting for empirical phenomena, and
- a true risk measure capable of describing the essential characteristics of risk.

Summary

- We explored a link between risk measures and dispersion measures through two sub-classes of coherent risk measures and deviation measures.
- Behind every such coherent risk measure, we can find a corresponding deviation measure and vice versa. The intuitive connection between risk and uncertainty materializes quantitatively in a particular form.
- Finally, we emphasized the importance of consistency of risk measures with the SSD order.

Svetlozar T. Rachev, Stoyan Stoyanov, and Frank J. Fabozzi
Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization: The Ideal Risk, Uncertainty, and Performance Measures
John Wiley, Finance, 2007.

Chapter 6.