Slides for Risk Management

VaR and Expected Shortfall

Groll

Seminar für Finanzökonometrie

Prof. Mittnik, PhD

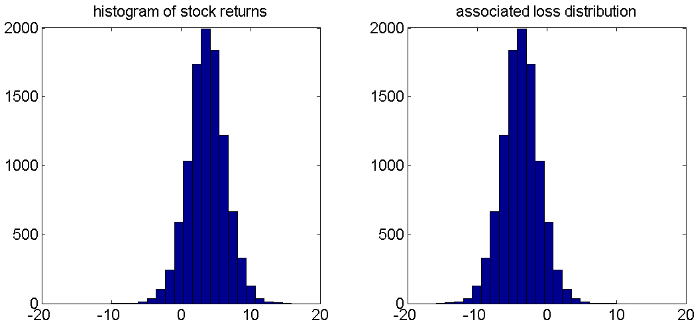
Notation

- risk often is defined as **negative deviation** of a given target payoff
- riskmanagement is mainly concerned with **downsiderisk**
- convention: focus on the **distribution of losses** instead of profits
- for prices denoted by P_t, the random variable quantifying losses is given by

$$L_{t+1} = -\left(P_{t+1} - P_t\right)$$

- distribution of losses equals distribution of profits flipped at x-axis

From profits to losses



Quantification of risk

- **decisions** concerned with managing, mitigating or hedging of risks have to be **based on quantification of risk** as basis of decision-making:
  - regulatory purposes: capital buffer proportional to exposure to risk
  - interior management decisions: freedom of daily traders restricted by capping allowed risk level
  - corporate management: identification of key risk factors (comparability)
- information contained in loss distribution is mapped to **scalar value**: information reduction

Decomposing risk

You are casino owner.

1. You only have one table of roulette, with one gambler, who bets 100 € on number 12. He only plays one game, and while the odds of winning are 1:36, his payment in case of success will be 3500 only. With expected positive payoff, what is your risk? $\Rightarrow$ **completely computable**
2. Now assume that you have multiple gamblers per day. Although you have a pretty good record of the number of gamblers over the last year, you still have to make an estimate about the number of visitors today. What is your risk? $\Rightarrow$ additional risk due to **estimation error**
3. You have been owner of The Mirage Casino in Las Vegas. What was your biggest loss within the last years?

Decomposing risk

- the closing of the show of Siegfried and Roy due to the attack of a tiger led to losses of hundreds of millions of dollars $\Rightarrow$ **model risk**

News | Weather | Traffic | Sports | Life & Style | Communities | KOMO 4 TV | KOMO NEWS

Roy Horn Of 'Siegfried And Roy' Critically Hurt In Tiger Attack

By KOMO Staff & News Services | Published: Oct 3, 2003 at 9:11 PM PDT | Last Updated: Aug 31, 2006 at 1:11 AM PDT

Recommend 0 | +1 0 | Tweet 0 | 0 Comments | Print | Email

komo 4 news

Roy Horn Attacked On Stage

LAS VEGAS - Illusionist Roy Horn of the duo Siegfried & Roy was in critical condition Saturday, a day after one of his own tigers mauled him during his popular Las Vegas show, biting him in the neck and dragging him off stage.

Authorities were uncertain about his chances for recovery.

Horn suffered a serious injury to the left side of his neck and underwent surgery late Friday.

Risk measurement frameworks

- **notional-amount approach:** weighted nominal value
    - nominal value as substitute for **outstanding** amount at risk
    - weighting factor representing riskiness of associated asset class as substitute for **riskiness** of individual asset
    - component of standardized approach of Basel capital adequacy framework
    - advantage: no individual risk assessment necessary - applicable even without empirical data
    - weakness: **diversification benefits** and **netting unconsidered**, strong simplification

Risk measurement frameworks

- **scenario analysis**:
  - **define** possible **future economic scenarios** (stock market crash of -20 percent in major economies, default of Greece government securities,...)
  - **derive** associated **losses**
  - determine risk as specified quantile of scenario losses (5*th* largest loss, worst loss, protection against at least 90 percent of scenarios,...)
  - since scenarios are not accompanied by statements about likelihood of occurrence, **probability dimension** is completely left **unconsidered**
  - scenario analysis **can be conducted** without any empirical data on the sole grounds of **expert knowledge**

Risk measurement frameworks

- **risk measures based on loss distribution:** statistical quantities of asset value distribution function
  - loss distribution
    - incorporates all information about **both probability and magnitude** of losses
    - includes **diversification and netting effects**
    - usually relies on empirical data
  - full information of distribution function reduced to **charateristics** of distribution for **better comprehensibility**
  - examples: standard deviation, Value-at-Risk, Expected Shortfall, Lower Partial Moments
  - **standard deviation**: **symmetrically** capturing positive and negative risks dilutes information about downsiderisk
  - overall loss distribution inpracticable: **approximate** risk measure of overall loss distribution by **aggregation of** asset **subgroup** risk **measures**

Value-at-Risk

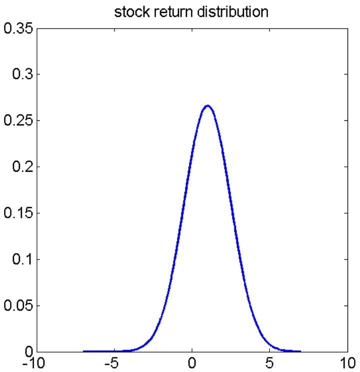
Value-at-Risk

The **Value-at-Risk (VaR) at the confidence level** α associated with a given loss distribution L is defined as the smallest value l that is not exceeded with probability higher than $(1-\alpha)$. That is,

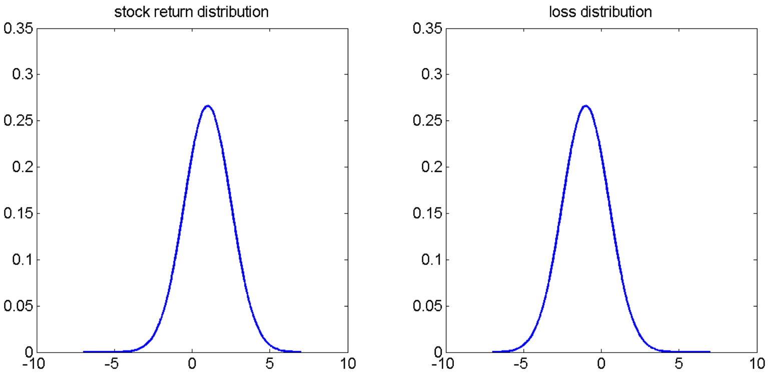
$$VaR_\alpha = \inf\left\{l \in \mathbb{R} : \mathbb{P}\left(L > l\right) \leq 1-\alpha\right\} = \inf\left\{l \in \mathbb{R} : F_L\left(l\right) \geq \alpha\right\}.$$

- typical values for α : $\alpha = 0.95$, $\alpha = 0.99$ or $\alpha = 0.999$
- as a measure of location, VaR does **not** provide any **information** about the nature of losses **beyond** the VaR
- the losses incurred by investments held on a daily basis exceed the value given by VaR_α only in $(1-\alpha)\cdot 100$ percent of days
- financial entity is **protected** in **at least** α-percent of days

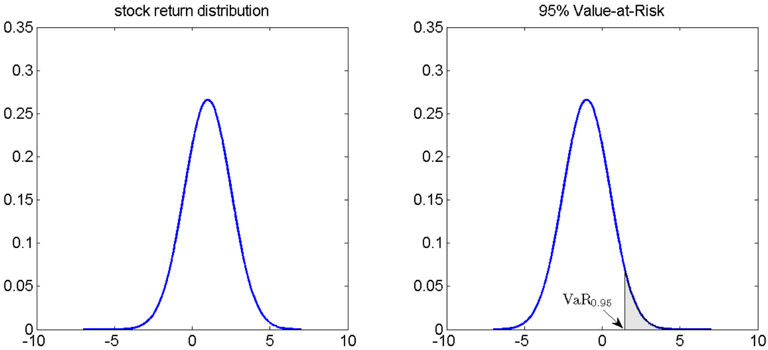
Loss distribution known



Loss distribution known



Loss distribution known

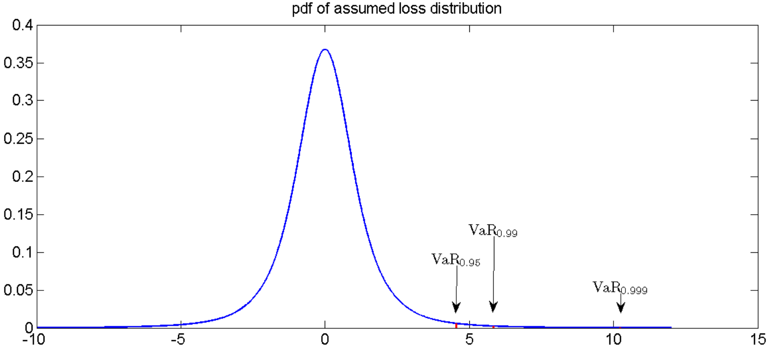


Estimation frameworks

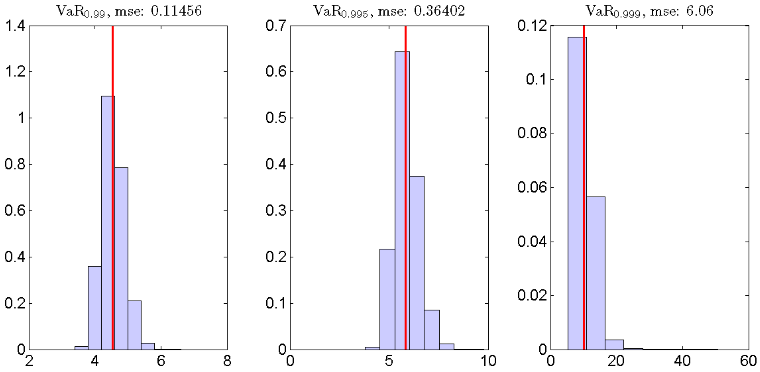
- in general: underlying loss distribution is **not known**
- two estimation methods for VaR:
  - **directly** estimate the associated **quantile** of **historical data**
  - estimate model for underlying **loss distribution**, and evaluate inverse cdf at required **quantile**
- derivation of VaR from a model for the loss distribution can be further decomposed:
  - **analytical** solution for quantile
  - **Monte Carlo Simulation** when analytic formulas are not available
- modelling the loss distribution inevitably entails **model risk**, which is concerned with possibly misleading results due to model **misspecifications**

Properties of historical simulation

- **simulation study**: examine properties of estimated **sample quantiles**
- assume *t*-**distributed** loss distribution with degrees-of-freedom parameter $\nu = 3$ and mean shifted by -0.004 :
  - $VaR_{0.99} = 4.54$
  - $VaR_{0.995} = 5.84$
  - $VaR_{0.999} = 10.22$
- estimate VaR for 100000 simulated samples of size 2500 (approximately 10 years in trading days)
- compare distribution of estimated VaR values with real value of applied underlying loss distribution
- even with sample size 2500, only 2.5 values occur above the 0.999 quantile on average $\Rightarrow$ high **mean squared errors (mse)**


pdf of assumed loss distribution
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-10
-5
0
5
10
15
VaR0.95
VaR0.99
VaR0.999

Distribution of estimated values



Modelling the loss distribution

- introductory model: assume **normally distributed** loss distribution

VaR normal distribution

For given parameters μ_L and σ VaR_α can be calculated **analytically** by

$$VaR_\alpha = \mu_L + \sigma\Phi^{-1}(\alpha) .$$

Proof.

$$\begin{aligned} \mathbb{P}(L \leq VaR_\alpha) &= \mathbb{P}\left(L \leq \mu_L + \sigma\Phi^{-1}(\alpha)\right) \\ &= \mathbb{P}\left(\frac{L-\mu_L}{\sigma} \leq \Phi^{-1}(\alpha)\right) \\ &= \Phi\left(\Phi^{-1}(\alpha)\right) = \alpha \end{aligned}$$

□

Remarks

- **note:** μ_L in

$$VaR_\alpha = \mu_L + \sigma\Phi^{-1}(\alpha)$$

  is the expectation of the loss distribution

- if μ denotes the expectation of the asset return, i.e. the expectation of the **profit**, then the formula has to be modified to

$$VaR_\alpha = -\mu + \sigma\Phi^{-1}(\alpha)$$

- in practice, the assumption of **normally distributed** returns usually can be **rejected** both for loss distributions associated with credit and operational risk, as well as for loss distributions associated with market risk at high levels of confidence

Expected Shortfall

Definition

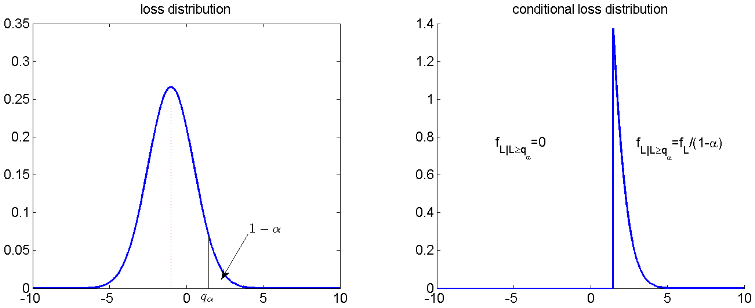
The **Expected Shortfall (ES)** with confidence level α denotes the **conditional expected loss**, given that the realized loss is equal to or exceeds the corresponding value of VaR_α :

$$ES_\alpha = \mathbb{E}\left[L|L \geq VaR_\alpha\right].$$

- given that we are in one of the $(1-\alpha) \cdot 100$ percent worst periods, how high is the loss that we have to expect?

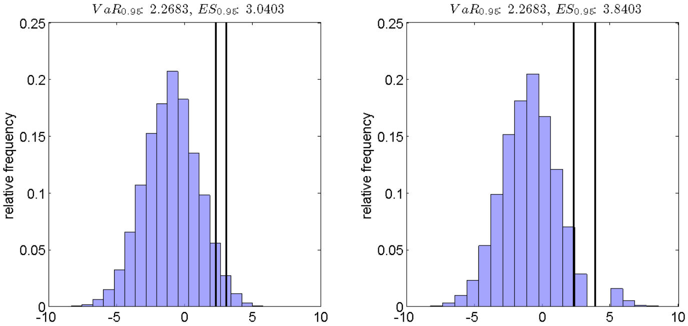
Expected Shortfall

- Expected Shortfall as expectation of **conditional loss distribution**:



Additional information of ES

- ES contains information about nature of losses **beyond** the *VaR* :

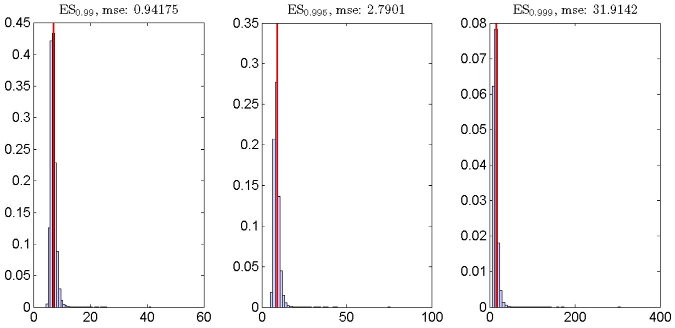


Estimation frameworks

- in general: underlying loss distribution is **not known**
- two estimation methods for ES:
  - **directly** estimate the **mean** of all values greater than the associated **quantile** of **historical data**
  - estimate model for underlying **loss distribution**, and calculate **expectation** of **conditional** loss distribution
- derivation of ES from a model for the loss distribution can be further decomposed:
  - **analytical** calculation of quantile and expectation: involves **integration**
  - **Monte Carlo Simulation** when analytic formulas are not available

Properties of historical simulation

- high **mean squared errors (mse)** for Expected Shortfall at high confidence levels:



ES under normal distribution

ES for normally distributed losses

Given that $L \sim \mathcal{N}\left(\mu_L, \sigma^2\right)$, the Expected Shortfall of L is given by

$$ES_\alpha = \mu_L + \sigma \frac{\phi\left(\Phi^{-1}\left(\alpha\right)\right)}{1-\alpha}.$$

Proof

$$
\begin{aligned}
ES_\alpha &= \mathbb{E}\left[L | L \geq VaR_\alpha\right] \\
&= \mathbb{E}\left[L | L \geq \mu_L + \sigma \Phi^{-1}(\alpha)\right] \\
&= \mathbb{E}\left[L | \frac{L-\mu_L}{\sigma} \geq \Phi^{-1}(\alpha)\right] \\
&= \mu_L - \mu_L + \mathbb{E}\left[L | \frac{L-\mu_L}{\sigma} \geq \Phi^{-1}(\alpha)\right] \\
&= \mu_L + \mathbb{E}\left[L - \mu_L | \frac{L-\mu_L}{\sigma} \geq \Phi^{-1}(\alpha)\right] \\
&= \mu_L + \sigma \mathbb{E}\left[\frac{L-\mu_L}{\sigma} | \frac{L-\mu_L}{\sigma} \geq \Phi^{-1}(\alpha)\right] \\
&= \mu_L + \sigma \mathbb{E}\left[Y | Y \geq \Phi^{-1}(\alpha)\right], \quad \text{with } Y \sim \mathcal{N}(0,1)
\end{aligned}
$$

Proof

Furthermore,

$$\mathbb{P}\left(Y \geq \Phi^{-1}(\alpha)\right) = 1 - \mathbb{P}\left(Y \leq \Phi^{-1}(\alpha)\right) = 1 - \Phi\left(\Phi^{-1}(\alpha)\right) = 1 - \alpha,$$

so that the conditional density as the scaled version of the standard normal density function is given by

$$\phi_{Y|Y\geq\Phi^{-1}(\alpha)}(y) = \frac{\phi(y)\,\mathbf{1}_{\{y\geq\Phi^{-1}(\alpha)\}}}{\mathbb{P}\left(Y \geq \Phi^{-1}(\alpha)\right)} = \frac{\phi(y)\,\mathbf{1}_{\{y\geq\Phi^{-1}(\alpha)\}}}{1-\alpha}.$$

Proof

Hence, the integral can be calculated as

$$\begin{aligned}
\mathbb{E}\left[Y|Y \geq \Phi^{-1}(\alpha)\right] &= \int_{\Phi^{-1}(\alpha)}^{\infty} y \cdot \phi_{Y|Y\geq\Phi^{-1}(\alpha)}(y)\,dy \\
&= \int_{\Phi^{-1}(\alpha)}^{\infty} y \cdot \frac{\phi(y)}{1-\alpha}\,dy \\
&= \frac{1}{1-\alpha}\int_{\Phi^{-1}(\alpha)}^{\infty} y \cdot \phi(y)\,dy \\
&\overset{(\star)}{=} \frac{1}{1-\alpha}\left[-\phi(y)\right]_{\Phi^{-1}(\alpha)}^{\infty} \\
&= \frac{1}{1-\alpha}\left(0+\phi\left(\Phi^{-1}(\alpha)\right)\right) \\
&= \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha},
\end{aligned}$$

with $(\star)$:

$$\left(-\phi(y)\right)' = -\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)\cdot\left(-\frac{2y}{2}\right) = y\cdot\phi(y)$$

Example: Meaning of VaR

You have invested 500,000 € in an investment fonds. The manager of the fonds tells you that the 99% Value-at-Risk for a time horizon of one year amounts to 5% of the portfolio value. Explain the information conveyed by this statement.

Solution

- for continuous loss distribution we have equality

$$\mathbb{P}\left(L \geq VaR_{\alpha}\right) = 1 - \alpha$$

- transform relative statement about losses into absolute quantity

$$VaR_{\alpha} = 0.05 \cdot 500,000 = 25,000$$

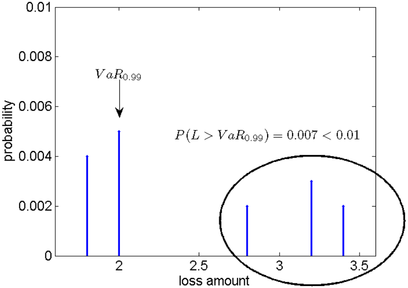
- pluggin into formula leads to

$$\mathbb{P}\left(L \geq 25,000\right) = 0.01,$$

  interpretable as "with probability 1% you will lose 25,000 € **or more**"
- a capital cushion of height $VaR_{0.99} = 25000$ is sufficient in exactly 99% of the times for continuous distributions

Example: discrete case

- example possible discrete loss distribution:



- the capital cushion provided by VaR_α would be sufficient even in 99.3% of the times
- interpretation of statement: "with **probability of maximal** 1% you will lose 25,000 € **or more**"

Example: Meaning of ES

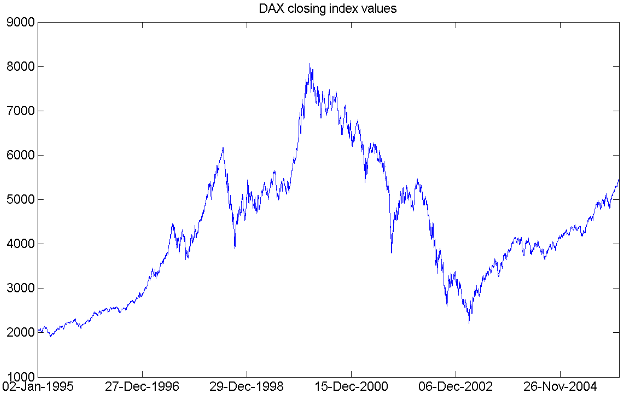
The fondsmanager corrects himself. Instead of the Value-at-Risk, it is the Expected Shortfall that amounts to 5% of the portfolio value. How does this statement have to be interpreted? Which of both cases does imply the riskier portfolio?

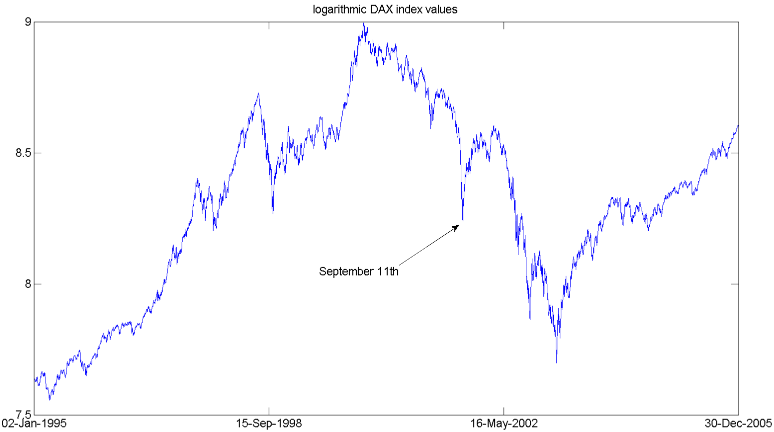
Solution

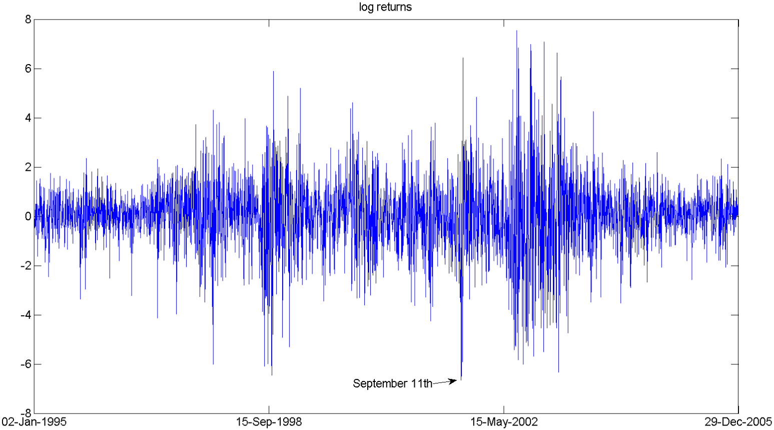
- given that one of the 1% worst years occurs, the expected loss in this year will amount to 25,000 €
- since always $VaR_\alpha \leq ES_\alpha$, the first statement implies $ES_\alpha \geq 25,000$ € $\Rightarrow$ the first statement implies the riskier portfolio

Example: market risk

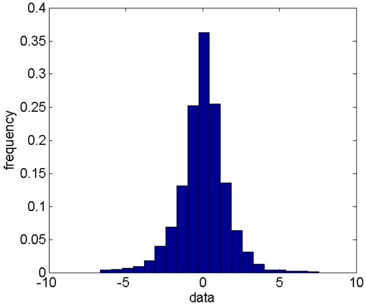
- estimating *VaR* for DAX

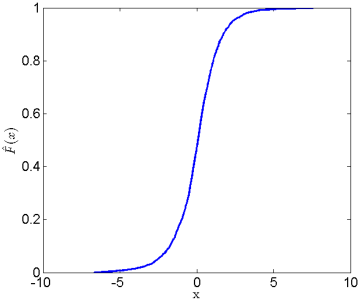



logarithmic DAX index values
9
8.5
8
7.5
02-Jan-1995
15-Sep-1998
16-May-2002
30-Dec-2005
September 11th

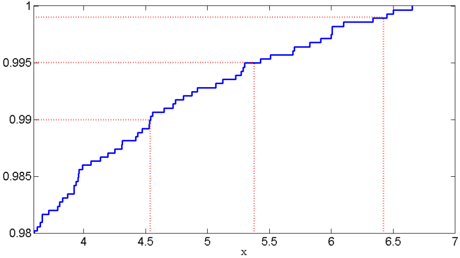

log returns
8
6
4
2
0
-2
-4
-6
-8
02-Jan-1995
15-Sep-1998
15-May-2002
29-Dec-2005
September 11th

Empirical distribution



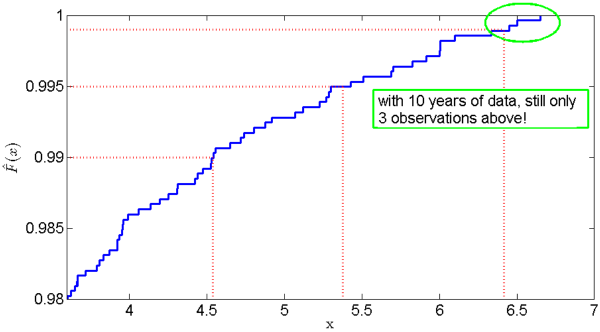


Historical simulation



- $VaR_{0.99} = 4.5380$, $VaR_{0.995} = 5.3771$, $VaR_{0.999} = 6.4180$
- $ES_{0.99} = 5.4711$, $ES_{0.995} = 6.0085$, $ES_{0.999} = 6.5761$

Historical simulation



Under normal distribution

- given estimated expectation for daily index returns, calculate estimated expected loss

$$\hat{\mu}_L = -\hat{\mu}$$

- plugging estimated parameter values of normally distributed losses into formula

$$VaR_\alpha = \mu_L + \sigma\Phi^{-1}(\alpha),$$

for $\alpha = 99\%$ we get

$$\begin{aligned}\widehat{VaR}_{0.99} &= \hat{\mu}_L + \hat{\sigma}\Phi^{-1}(0.99) \\ &= -0.0344 + 1.5403 \cdot 2.3263 = 3.5489\end{aligned}$$

- for $VaR_{0.995}$ we get

$$\widehat{VaR}_{0.995} = -0.0344 + 1.5403 \cdot 2.5758 = 3.9331$$

- for Expected Shortfall, using

$$ES_\alpha = \mu_L + \sigma \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha},$$

we get

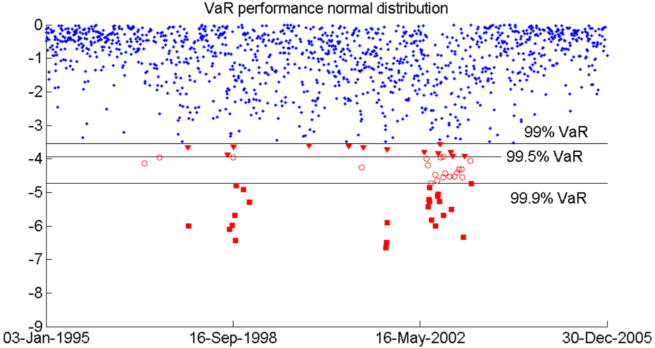
$$\begin{aligned}\widehat{ES}_{0.99} &= -0.0344 + 1.5403 \cdot \frac{\phi\left(\Phi^{-1}(0.99)\right)}{0.01} \\ &= -0.0344 + 1.5403 \cdot \frac{\phi(2.3263)}{0.01} \\ &= -0.0344 + 1.5403 \cdot \frac{0.0267}{0.01} \\ &= 4.0782,\end{aligned}$$

and

$$\begin{aligned}\widehat{ES}_{0.995} &= -0.0344 + 1.5403 \cdot \frac{\phi(2.5758)}{0.005} \\ &= -0.0344 + 1.5403 \cdot \frac{0.0145}{0.005} \\ &= 4.4325.\end{aligned}$$

Performance: backtesting

- how good did *VaR*-calculations with normally distributed returns perform?



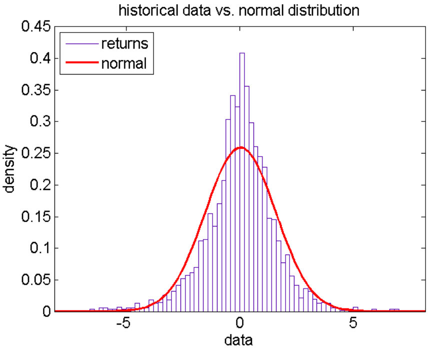
Backtesting: interpretation

- backtesting *VaR*-calculations based on assumption of **independent normally distributed losses** generally leads to two **patterns**:
  - percentage frequencies of *VaR*-**exceedances** are **higher** than the confidence levels specified: normal distribution assigns **too less probability to large losses**
  - *VaR*-exceedances occur in **clusters**: given an exceedance of *VaR* today, the likelihood of an additional exceedance in the days following is larger than average
  - clustered exceedances indicate **violation of independence** of losses over time
  - clusters have to be captured through **time series models**

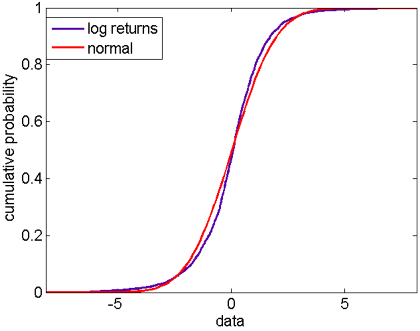
Model risk

- given that returns in the real world were indeed generated by an underlying normal distribution, we could determine the risk inherent to the investment up to a small error arising from **estimation errors**
- however, returns of the real world are **not normally distributed**
- in addition to the risk deduced from the model, the model itself could be significantly different to the processes of the real world that are under consideration
- the risk of deviations of the specified model from the real world is called **model risk**
- the results of the **backtesting procedure** indicate substantial **model risk** involved in the framework of **assumed normally distributed losses**

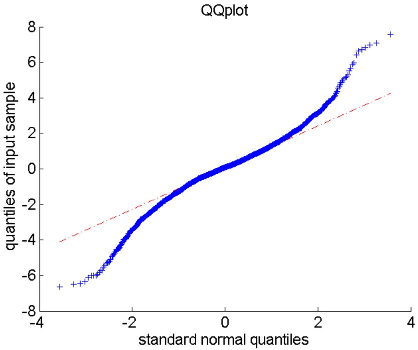
Appropriateness of normal distribution



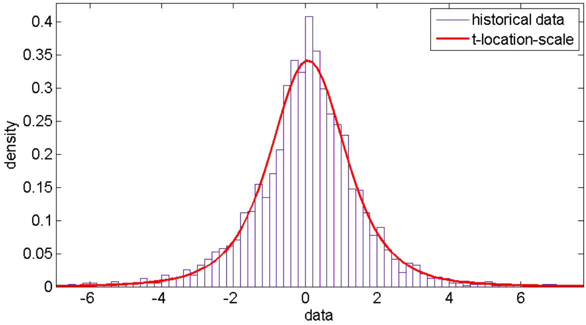
Appropriateness of normal distribution



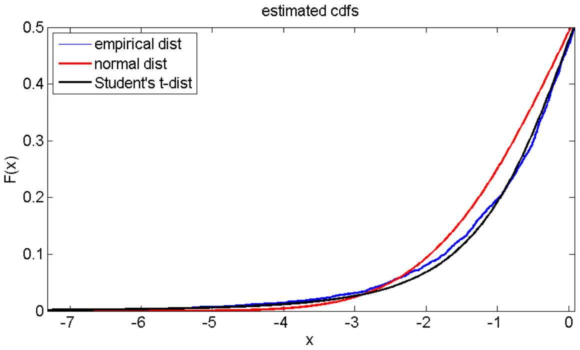
Appropriateness of normal distribution



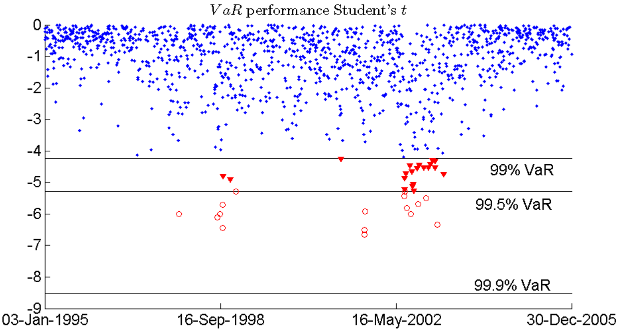
Student's t-distribution



Student’s t-distribution



Student's t-distribution



- note: clusters in *VaR*-exceedances remain

Comparing values

VaR	0.99	0.995	0.999
historical values	4.5380	5.3771	6.4180
normal assumption	3.5490	3.9333	4.7256
Student's t assumption	4.2302	5.2821	8.5283

ES	0.99	0.995	0.999
historical values	5.4711	6.0085	6.5761
normal assumption	4.0782	4.4325	5.1519
Student's t assumption	6.0866	7.4914	11.9183

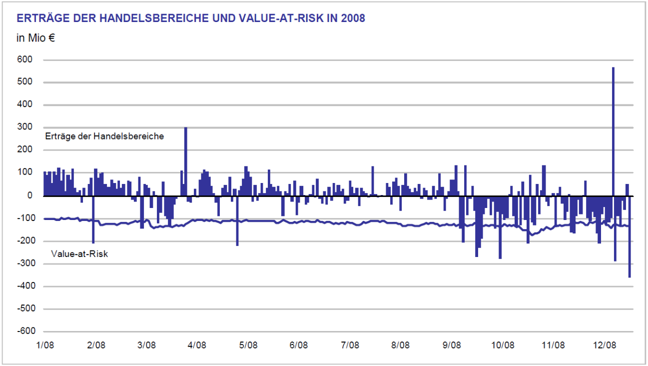
Comparing number of hits

sample size: 2779	$VaR_{0.99}$	$VaR_{0.995}$	$VaR_{0.999}$
historical values	28	14	3
frequency	0.01	0.005	0.001
normal assumption	57	44	24
frequency	0.0205	0.0158	0.0086
Student's t assumption	36	16	0
frequency	0.0130	0.0058	0

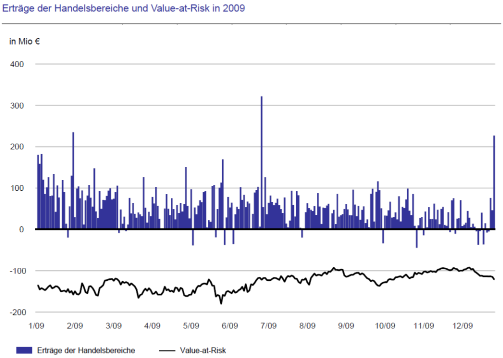
- note: exceedance frequencies for historical simulation equal predefined confidence level per definition →**overfitting**

Model risk

- besides sophisticated modelling approaches, even Deutsche Bank seems to fail at *VaR*-estimation: $VaR_{0.99}$



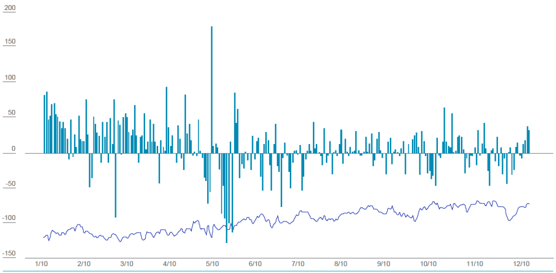
Model risk


Erträge der Handelsbereiche und Value-at-Risk in 2009
in Mio €
400
300
200
100
0
-100
-200
1/09
2/09
3/09
4/09
5/09
6/09
7/09
8/09
9/09
10/09
11/09
12/09
Erträge der Handelsbereiche
Value-at-Risk

Model risk

Buy-and-hold Erträge der Handelsbereiche und Value-at-Risk in 2010

in Mio €



— Buy-and-Hold Erträge der Handelsbereiche
— Value-at-Risk

- given only information about VaR_α^A of random variable A and VaR_α^B of random variable B, there is in general **no sufficient information** to calculate VaR for a function of both:

$$VaR_\alpha^{f(A,B)} \neq g\left(VaR_\alpha^A, VaR_\alpha^B\right)$$

- in such cases, in order to calculate $VaR_\alpha^{f(A,B)}$, we have to derive the distribution of $f(A,B)$ first
- despite the **marginal distributions** of the constituting parts, the transformed distribution under f is affected by the way that the margins are related with each other: the **dependence structure** between individual assets is **crucial** to the determination of $VaR_\alpha^{f(A,B)}$

Multi-period case

- as **multi-period returns** can be calculated as simple **sum** of sub-period returns in the **logarithmic case**, we aim to model

$$VaR_{\alpha}^{f(A,B)} = VaR_{\alpha}^{A+B}$$

- even though our object of interest relates to a simple sum of random variables, easy **analytical solutions** apply only in the very **restricted cases** where summation preserves the distribution: A, B and $A + B$ have to be of the same distribution
- this property is **fulfilled** for the case of **jointly normally** distributed random variables

Multi-asset case

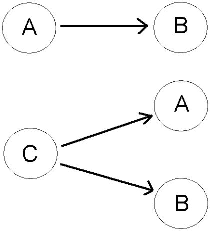
- while portfolio returns can be calculated as **weighted sum** of individual assets for **discrete returns**, such an easy relation does **not** exist for the case of **logarithmic returns**
- discrete case:

$$r_P = w_1 r_1 + w_2 r_2$$

- logarithmic case:

$$\begin{aligned} r_P^{log} &= \ln\left(1 + r_P\right) \\ &= \ln\left(1 + w_1 r_1 + w_2 r_2\right) \\ &= \ln\left(1 + w_1\left[\exp\left(\ln\left(1 + r_1\right)\right) - 1\right] + w_2\left[\exp\left(\ln\left(1 + r_2\right)\right) - 1\right]\right) \\ &= \ln\left(w_1 \exp\left(r_1^{log}\right) + w_2 \exp\left(r_2^{log}\right)\right) \end{aligned}$$

Origins of dependency



- **direct influence:**
  - smoking → health
  - recession → probabilities of default
  - both directions: wealth ↔ education
- **common underlying influence:**
  - gender → income, shoe size
  - economic fundamentals → BMW, Daimler

Independence

- two random variables X and Y are called **stochastically independent** if

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)$$

  for all x, y

- the occurrence of one event makes it neither more nor less probable that the other occurs:

$$\begin{aligned}\mathbb{P}(X = x | Y = y) &= \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} \\ &\overset{indep.}{=} \frac{\mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)}{\mathbb{P}(Y = y)} \\ &= \mathbb{P}(X = x)\end{aligned}$$

- knowledge of the realization of Y does **not** provide **additional information** about the distribution of X

Example: independent dice

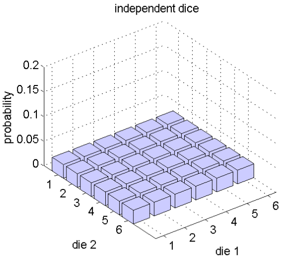
- because of **independency**, joint probability is given by product:

$$\mathbb{P}(X=5, Y=4) \overset{indep.}{=} \mathbb{P}(X=5) \cdot \mathbb{P}(Y=4) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

- joint distribution given by

$$\mathbb{P}(X=i, Y=j) \overset{indep.}{=} \mathbb{P}(X=i) \cdot \mathbb{P}(Y=j) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36},$$

for all $i, j \in \{1, \ldots, 6\}$

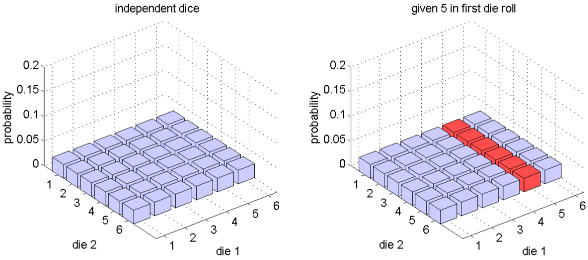


Example: independent dice

- because of **independency**, realization of die 1 does not provide additional information about occurrence of die 2:

$$\mathbb{P}\left(X = x | Y = 5\right) \overset{indep.}{=} \mathbb{P}\left(X = x\right)$$

- **conditional distribution:** relative distribution of probabilities (red row) has to be **scaled up**
- unconditional distribution of die 2 is equal to the conditional distribution given die 1

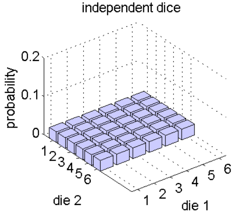


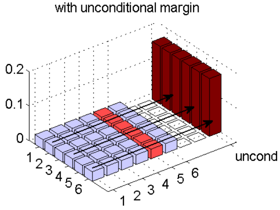
Example: unconditional distribution

- given the joint distribution, the **unconditional marginal** probabilities are given by

$$\mathbb{P}(X = x) = \sum_{i=1}^{6} \mathbb{P}(X = x, Y = i)$$

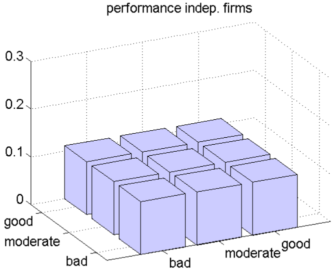
- **marginal distributions** hence are obtained by **summation** along the appropriate direction:





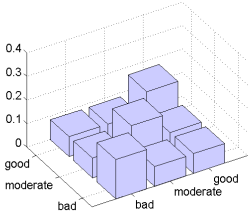
Example: independent firms

- firms A and B, both with possible performances **good**, **moderate** and **bad**
- each performance occurs with **equal probability** $\frac{1}{3}$
- **joint distribution** for the case of **independency**:



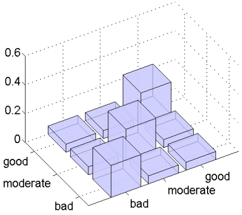
Example: dependent firms

- firm A **costumer** of firm B: demand for good of firm B depending on financial condition of A
- good financial condition $A \rightarrow$ high demand $\rightarrow$ high income for $B \rightarrow$ increased likelihood of good financial condition of firm B:
  - $f_A = \frac{1}{3}\delta_{\{a=good\}} + \frac{1}{3}\delta_{\{a=mod.\}} + \frac{1}{3}\delta_{\{a=bad\}}$
  - $f_B = \frac{1}{2}\delta_{\{b=a\}} + \frac{1}{4}\delta_{\{b=good\}}\mathbf{1}_{\{a\neq good\}} + \frac{1}{4}\delta_{\{b=mod.\}}\mathbf{1}_{\{a\neq mod.\}} + \frac{1}{4}\delta_{\{b=bad\}}\mathbf{1}_{\{a\neq bad\}}$



Example: common influence

- firm A and firm B **supplier** of firm C: demand for goods of firm A and B depending on financial condition of C :
  - $f_C = \frac{1}{3}\delta_{\{c=good\}} + \frac{1}{3}\delta_{\{c=mod.\}} + \frac{1}{3}\delta_{\{c=bad\}}$
  - $f_A = \frac{3}{4}\delta_{\{a=c\}} + \frac{1}{8}\delta_{\{a=good\}}\mathbf{1}_{\{c\neq good\}} + \frac{1}{8}\delta_{\{a=mod.\}}\mathbf{1}_{\{c\neq mod.\}} + \frac{1}{8}\delta_{\{a=bad\}}\mathbf{1}_{\{c\neq bad\}}$
  - $f_B = \frac{3}{4}\delta_{\{b=c\}} + \frac{1}{8}\delta_{\{b=good\}}\mathbf{1}_{\{c\neq good\}} + \frac{1}{8}\delta_{\{b=mod.\}}\mathbf{1}_{\{c\neq mod.\}} + \frac{1}{8}\delta_{\{b=bad\}}\mathbf{1}_{\{c\neq bad\}}$



Example: common influence

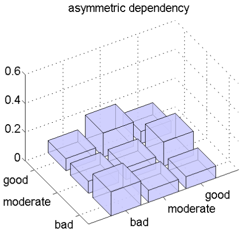
- get common distribution of firm A and B:

$$\begin{aligned}\mathbb{P}(A=g,B=g) &= \mathbb{P}(A=g,B=g|C=g)\\ &\quad+\mathbb{P}(A=g,B=g|C=m)+\mathbb{P}(A=g,B=g|C=b)\\ &= \frac{3}{4}\cdot\frac{3}{4}\cdot\frac{1}{3}+\frac{1}{8}\cdot\frac{1}{8}\cdot\frac{1}{3}+\frac{1}{8}\cdot\frac{1}{8}\cdot\frac{1}{3}\\ &= \frac{19}{68}\end{aligned}$$

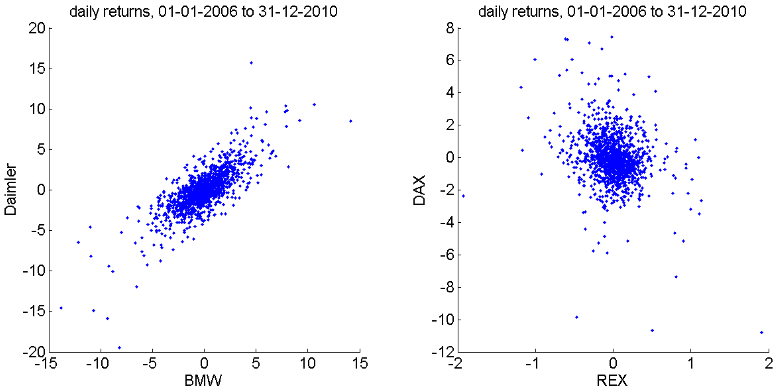
$$\begin{aligned}\mathbb{P}(A=g,B=m) &= \mathbb{P}(A=g,B=m|C=g)\\ &\quad+\mathbb{P}(A=g,B=m|C=m)+\mathbb{P}(A=g,B=m|C=b)\\ &= \frac{3}{4}\cdot\frac{1}{8}\cdot\frac{1}{3}+\frac{1}{8}\cdot\frac{3}{4}\cdot\frac{1}{3}+\frac{1}{8}\cdot\frac{1}{8}\cdot\frac{1}{3}\\ &= \frac{13}{192}\end{aligned}$$

Example: asymmetric dependency

- two **competitive** firms with common **economic fundamentals:**
  - in times of **bad** economic conditions: both firms tend to perform bad
  - in times of **good** economic conditions: due to competition, a prospering competitor most likely comes at the expense of other firms in the sector
- **dependence** during **bad** economic conditions **stronger** than in times of booming market



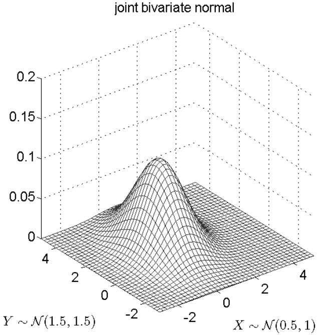
Empirical

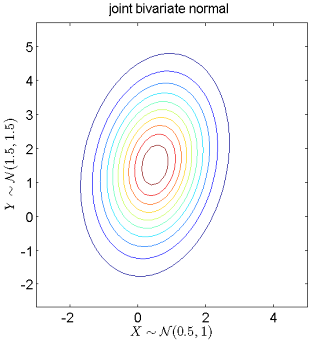

daily returns, 01-01-2006 to 31-12-2010
Daimler
BMW
20
15
10
5
0
-5
-10
-15
-20
-15
-10
-5
0
5
10
15
daily returns, 01-01-2006 to 31-12-2010
DAX
REX
8
6
4
2
0
-2
-4
-6
-8
-10
-12
-2
-1
0
1
2

Bivariate normal distribution

$$f(x,y) = \frac{1}{2\pi\sigma_{\mathbf{X}}\sigma_{\mathbf{Y}}\sqrt{1-\rho^2}} \cdot exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_{\mathbf{X}})^2}{\sigma_{\mathbf{X}}^2} + \frac{(y-\mu_{\mathbf{Y}})^2}{\sigma_{\mathbf{Y}}^2} - \frac{2\rho(x-\mu_{\mathbf{X}})(y-\mu_{\mathbf{Y}})}{\sigma_{\mathbf{X}}\sigma_{\mathbf{Y}}}\right]\right)$$

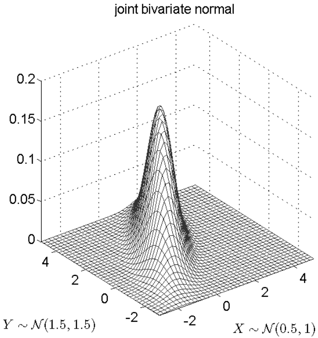
- $\rho = 0.2$:

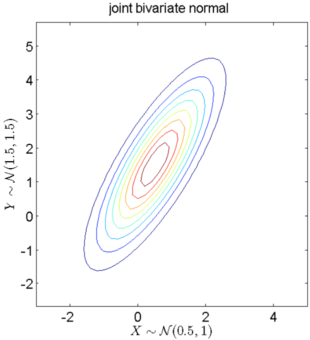




Bivariate normal distribution

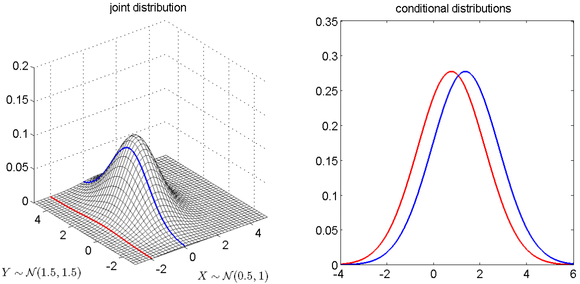
- $\rho = 0.8$:





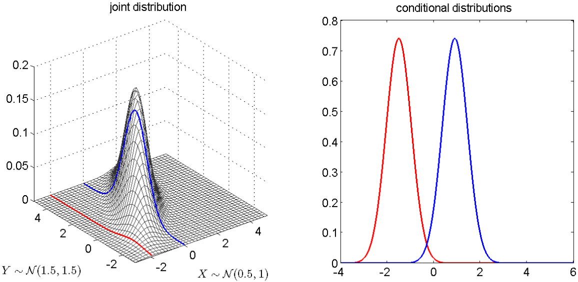
Conditional distributions

- distribution of Y conditional on $X = 0$ compared with distribution conditional on $X = -2$:



Conditional distributions

- $\rho = 0.8$: the **information** conveyed by the known realization of X **increases** with increasing dependency between the variables



Interpretation

- given **jointly normally distributed** variables X and Y, you can think about X as being a **linear transformation** of Y, up to some **normally distributed noise** term ϵ :

$$X = c_1 (Y - c_2) + c_3 \epsilon,$$

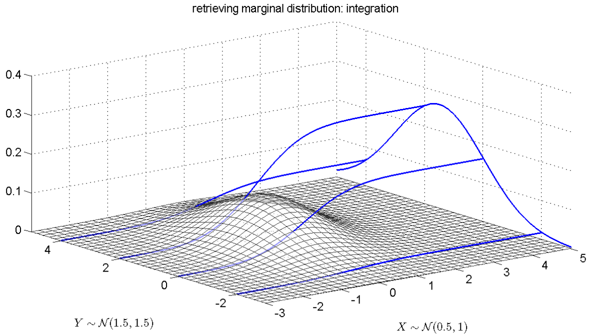
with

$$c_1 = \frac{\sigma_X}{\sigma_Y} \rho$$
$$c_2 = \mu_Y$$
$$c_3 = \sigma_X \sqrt{1 - \rho^2}$$

- proof will follow further down

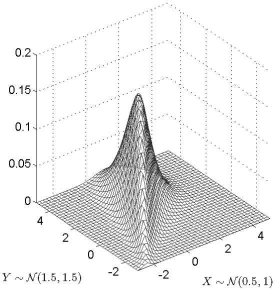
Marginal distributions

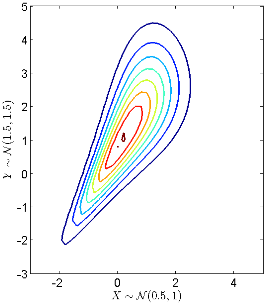
- marginal distributions are obtained by integrating out with respect to the other dimension



Asymmetric dependency

- there exist joint bivariate distributions with normally distributed margins that can not be generated from a bivariate normal distribution





Covariance

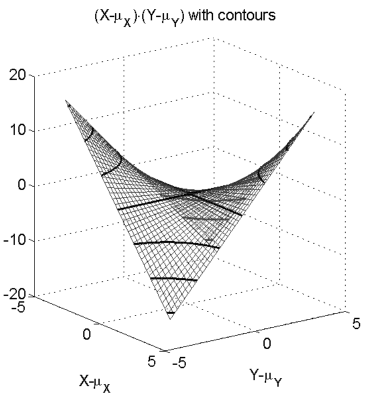
Covariance

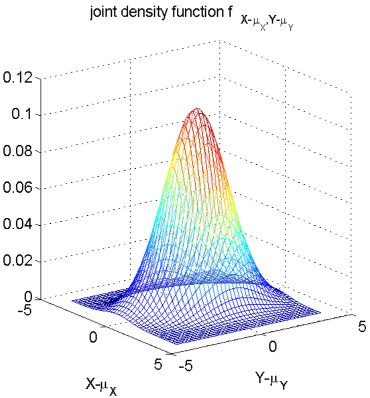
The **covariance** of two random variables X and Y is defined as

$$\mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y].$$

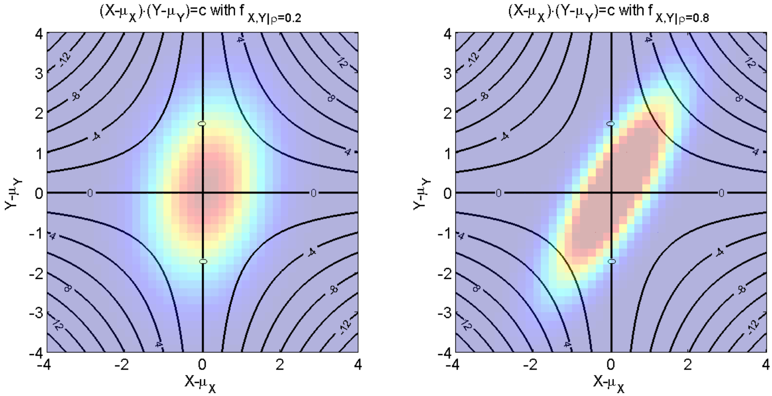
- captures tendency of variables X and Y to jointly take on values above the expectation
- given that $Cov(X, Y) = 0$, the random variables X and Y are called **uncorrelated**
- $Cov(X,X) = \mathbb{E}\left[(X - \mathbb{E}[X]) \cdot (X - \mathbb{E}[X])\right] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{V}[X]$

Covariance


(X-μ_X)·(Y-μ_Y) with contours
20
10
0
-10
-20
-5
0
5
X-μ_X
-5
0
5
Y-μ_Y


joint density function f X-μ_X,Y-μ_Y
0.12
0.1
0.08
0.06
0.04
0.02
0
-5
0
5
X-μ_X
-5
0
5
Y-μ_Y

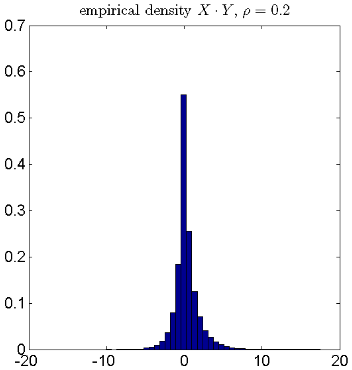
Covariance



```
% initialize parameters
mu1 = 0;
mu2 = 0;
sigma1 = 1;
sigma2 = 1.5;
rho = 0.2;
n = 10000;

% simulate data
data = mvnrnd([mu1 mu2],[sigma1^2 rho*sigma1*sigma2; rho*
  sigma1*sigma2 sigma2^2],n);
data = data(:,1).*data(:,2);
hist(data,40)
```

Covariance


empirical density $X \cdot Y$, $\rho = 0.2$
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-20
-10
0
10
20
empirical density $X \cdot Y$, $\rho = 0.8$
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-10
0
10
20
30

Covariance under linear transformation

$$
\begin{aligned}
Cov\left(aX+b, cY+d\right) &= \mathbb{E}\left[\left(aX+b-\mathbb{E}\left[aX+b\right]\right)\left(cY+d-\mathbb{E}\left[cY+d\right]\right)\right] \\
&= \mathbb{E}\left[\left(aX-\mathbb{E}\left[aX\right]+b-\mathbb{E}\left[b\right]\right)\left(cY-\mathbb{E}\left[cY\right]+d-\mathbb{E}\left[d\right]\right)\right] \\
&= \mathbb{E}\left[a\left(X-\mathbb{E}\left[X\right]\right)\cdot c\left(Y-\mathbb{E}\left[Y\right]\right)\right] \\
&= ac\cdot\mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]\right)\cdot\left(Y-\mathbb{E}\left[Y\right]\right)\right] \\
&= ac\cdot Cov\left(X, Y\right)
\end{aligned}
$$

Linear Correlation

Linear correlation

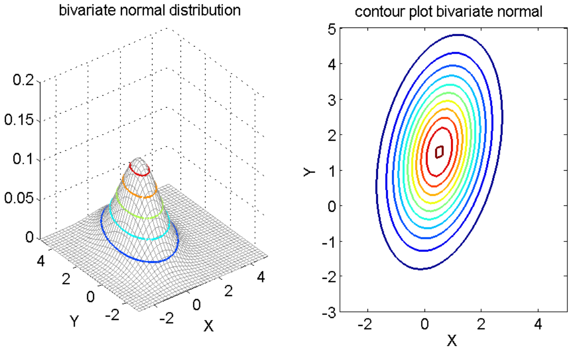
The linear correlation coefficient between two random variables X and Y is defined as

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}},$$

where $Cov(X, Y)$ denotes the covariance between X and Y, and σ_X^2 , σ_Y^2 denote the variances of X and Y .

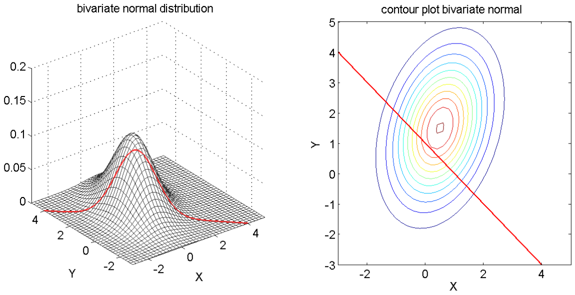
- in the elliptical world, any given distribution can be completely described by its margins and its correlation coefficient.

Jointly normally distributed variables



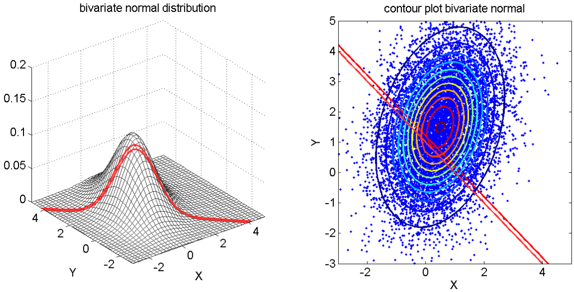
Jointly normally distributed variables

- all two-dimensional points on red line result in the value 4 after summation
- for example: $4 + 0$, $3 + 1$, $2.5 + 1.5$, or $-5 + 1$



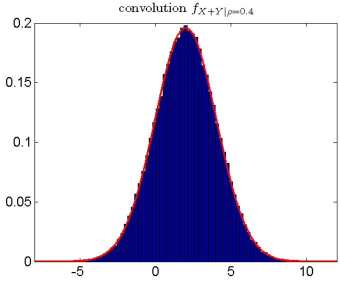
Jointly normally distributed variables

- approximate distribution of $X+Y$: counting the number of simulated values between the lines gives estimator for relative frequency of a summation value between 4 and 4.2



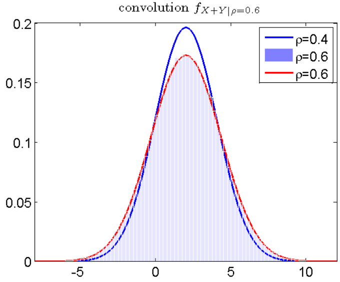
Distribution of sum of variables

- dividing two-dimensional space into series of line segments leads to approximation of distribution of new random variable $X + Y$



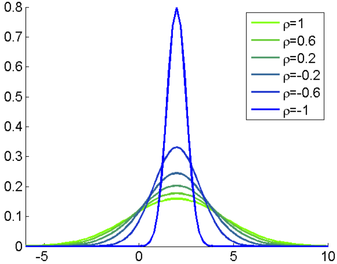
Effects of correlation

- **increasing correlation** leads to higher probability of joint large positive or large negative realizations
- joint large realizations of same sign lead to high absolute values after summation: **increasing probability in** the **tails**



Effects of correlation

- small variances in case of negative correlations display **benefits of diversification**



Moments of sums of variables

Theorem

Given random variables X and Y of ***arbitrary distribution*** *with existing first and second moments, the first and second moment of the summed up random variable $Z = X + Y$ are given by*

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y],$$

and

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2Cov(X,Y).$$

In general, for more than 2 variables, it holds:

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i],$$

$$\mathbb{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{V}(X_i) + \sum_{i \neq j}^{n} Cov(X_i, X_j).$$

Proof: linearity in joint normal distribution

$$
\begin{aligned}
\mathbb{V}(X+Y) &= \mathbb{E}\left[(X+Y) - \mathbb{E}\left[(X+Y)^2\right]\right] \\
&= \mathbb{E}\left[(X+Y)^2\right] - (\mathbb{E}[X+Y])^2 \\
&= \mathbb{E}\left[X^2 + 2XY + Y^2\right] - (\mathbb{E}[X] + \mathbb{E}[Y])^2 \\
&= \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 + \mathbb{E}\left[Y^2\right] - \mathbb{E}[Y]^2 + 2\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y] \\
&= \mathbb{V}(X) + \mathbb{V}(Y) + 2(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]) \\
&= \mathbb{V}(X) + \mathbb{V}(Y) + 2Cov(X,Y)
\end{aligned}
$$

Moments of sums of variables

- calculation of $\mathbb{V}(X+Y)$ **requires** knowledge of the **covariance** of X and Y
- however, more detailed information about the dependence structure of X and Y is not required
- for linear functions in general:

$$\mathbb{E}[aX+bY+c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$$

$$\mathbb{V}(aX+bY+c) = a^2\mathbb{V}(X) + b^2\mathbb{V}(Y) + 2ab\cdot Cov(X,Y)$$

- also:

$$Cov(X+Y,K+L) = Cov(X,K)+Cov(X,L)+Cov(Y,K)+Cov(Y,L)$$

Proof: linearity underlying joint normal distribution

- define random variable Z as

$$Z := \frac{\sigma_X}{\sigma_Y}\rho\left(Y - \mu_Y\right) + \mu_X + \sigma_X\sqrt{1-\rho^2}\epsilon,$$

$$\epsilon \sim \mathcal{N}\left(0,1\right)$$

- then the expectation is given by

$$\begin{aligned}\mathbb{E}\left[Z\right] &= \frac{\sigma_X}{\sigma_Y}\rho\mathbb{E}\left[Y\right] - \frac{\sigma_X}{\sigma_Y}\rho\mu_Y + \mu_X + \sigma_X\sqrt{1-\rho^2}\mathbb{E}\left[\epsilon\right] \\ &= \mu_Y\left(\frac{\sigma_X}{\sigma_Y}\rho - \frac{\sigma_X}{\sigma_Y}\rho\right) + \mu_X \\ &= \mu_X\end{aligned}$$

- the variance is given by

$$\begin{aligned}\mathbb{V}(Z) &\overset{Y\perp\epsilon}{=} \mathbb{V}\left(\frac{\sigma_X}{\sigma_Y}\rho Y\right) + \mathbb{V}\left(\sigma_X\sqrt{1-\rho^2}\epsilon\right)\\ &= \frac{\sigma_X^2}{\sigma_Y^2}\sigma_Y^2\rho^2 + \sigma_X^2\left(1-\rho^2\right)\cdot 1\\ &= \sigma_X^2\left(\rho^2+1-\rho^2\right)\\ &= \sigma_X^2\end{aligned}$$

- being the sum of two independent normally distributed random variables, Z is normally distributed itself, so that we get

$$Z \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right) \Leftrightarrow Z \sim X$$

- for the conditional distribution of Z given $Y = y$ we get

$$Z|_{Y=y} \sim \mathcal{N}\left(\mu_X + \frac{\sigma_X}{\sigma_Y}\rho\left(y - \mu_Y\right), \left(\sigma_X\sqrt{1-\rho^2}\right)^2\right)$$

- it remains to show

$$Z|_{Y=y} \sim X|_{Y=y}$$

to get

$$(Z, Y) \sim (X, Y)$$

Conditional normal distribution I

- for jointly normally distributed random variables X and Y, conditional marginal distributions remain normally distributed:

$$
\begin{aligned}
f(x|y) &= \frac{f(x,y)}{f(y)} \\
&= \frac{\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}}{\frac{1}{\sqrt{2\pi}\sigma_Y}} \\
&\quad \cdot \frac{exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}-\frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right)}{exp\left(-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right)} \\
&\overset{!}{=} \frac{1}{\sqrt{2\pi}\sigma}\cdot exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)
\end{aligned}
$$

Conditional normal distribution II

$$
\begin{aligned}
\frac{1}{\sqrt{2\pi}\sigma} &\stackrel{!}{=} \frac{\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}}{\frac{1}{\sqrt{2\pi}\sigma_Y}} \\
&= \frac{\sqrt{2\pi}}{2\pi\sigma_X\sqrt{1-\rho^2}} \\
&= \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}} \\
\Leftrightarrow \sigma &= \sigma_X\sqrt{1-\rho^2}
\end{aligned}
$$

Conditional normal distribution III

$$exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \overset{!}{=} \frac{exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}-\frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right)}{exp\left(-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right)}$$

$$\Leftrightarrow \frac{-(x-\mu)^2}{\sigma^2} = -\frac{1}{(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}-\frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right] + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$$

$$\Leftrightarrow \frac{(x-\mu)^2}{\sigma^2} = \frac{1}{(1-\rho^2)\sigma_X^2}\left((x-\mu_X)^2+\frac{\sigma_X^2}{\sigma_Y^2}(y-\mu_Y)^2\right) + \frac{1}{(1-\rho^2)\sigma_X^2}\left(-2\rho(x-\mu_X)(y-\mu_Y)\frac{\sigma_X}{\sigma_Y}-\frac{\sigma_X^2}{\sigma_Y^2}(y-\mu_Y)^2(1-\rho^2)\right)$$

Conditional normal distribution IV

$$\Leftrightarrow (x-\mu)^2 = (x-\mu_X)^2 + \left(1-\left(1-\rho^2\right)\right)\frac{\sigma_X^2}{\sigma_Y^2}\left(y-\mu_Y\right)^2 - 2\rho\left(x-\mu_X\right)\left(y-\mu_Y\right)\frac{\sigma_X}{\sigma_Y}$$

$$\Leftrightarrow (x-\mu)^2 = \left(\left(x-\mu_X\right) - \frac{\sigma_X}{\sigma_Y}\rho\left(y-\mu_Y\right)\right)^2$$

$$\Leftrightarrow \mu = \mu_X + \frac{\sigma_X}{\sigma_Y}\rho\left(y-\mu_Y\right)$$

- for $(X, Y) \sim \mathcal{N}_2\left(\begin{bmatrix}\mu_X \\ \mu_Y\end{bmatrix}, \begin{bmatrix}\sigma_X^2 & \rho\cdot\sigma_X\sigma_Y \\ \rho\cdot\sigma_X\sigma_Y & \sigma_Y^2\end{bmatrix}\right)$, given the realization of Y, X is distributed according to

$$X \sim \mathcal{N}\left(\mu_X + \frac{\sigma_X}{\sigma_Y}\rho\left(y-\mu_Y\right), \left(\sigma_X\sqrt{1-\rho^2}\right)^2\right)$$

Jointly normally distributed variables

Theorem

Given ***jointly normally distributed*** *univariate random variables* $X \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right)$ *and* $Y \sim \mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$*, the deduced random vector* $Z := X + Y$ *is also* ***normally distributed****, with parameters*

$$\mu_Z = \mu_X + \mu_Y$$

and

$$\sigma_Z = \sqrt{\mathbb{V}(X) + \mathbb{V}(Y) + 2Cov(X, Y)}.$$

That is,

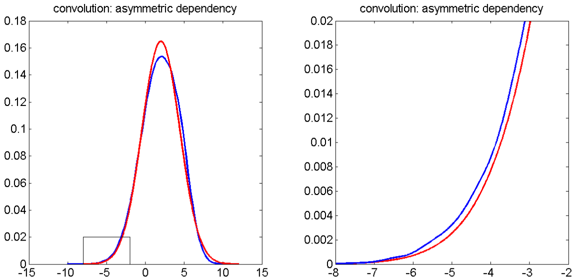
$$X + Y \sim \mathcal{N}\left(\mu_X + \mu_Y, \mathbb{V}(X) + \mathbb{V}(Y) + 2Cov(X, Y)\right).$$

Remarks

- note: the bivariate random vector (X, Y) has to be distributed according to a **bivariate normal** distribution, i.e. $(X, Y) \sim \mathcal{N}_2(\mu, \Sigma)$
- given that $X \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right)$ and $Y \sim \mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$, with **dependence** structure **different** to the one implicitly given by a bivariate normal distribution, the **requirements** of the theorem are **not fulfilled**
- in general, with deviating dependence structure we can only infer knowledge about **first** and **second moments** of the distribution of $Z = X + Y$, but we are **not able** to deduce the **shape** of the distribution

Convolution with asymmetric dependence

- univariate normally distributed random vectors $X \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right)$ and $Y \sim \mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$, linked by **asymmetric dependence** structure with stronger dependence for negative results than for positive results
- approximation for distribution of $X + Y$:



- note: $X + Y$ does **not** follow a **normal** distribution!
  $X + Y \nsim \mathcal{N}\left(\mu, \sigma^2\right)$

Independence over time

Assumption

The return of any given period shall be **independent** of the returns of previous periods:

$$\mathbb{P}\left(r_t^{log} \in [a, b], r_{t+k}^{log} \in [c, d]\right) = \mathbb{P}\left(r_t^{log} \in [a, b]\right) \cdot \mathbb{P}\left(r_t^{log} \in [c, d]\right),$$

for all $k \in \mathbb{Z}$, $a, b, c, d \in \mathbb{R}$.

Consequences

- consequences of assumption of independence over time combined with
  - case 1: **arbitrary** return **distribution**
    - moments of multi-period returns can be derived from moments of one-period returns: **square-root-of-time** scaling for **standard deviation**
    - **multi-period** return **distribution** is **unknown**: for some important risk measures like *VaR* or *ES* no analytical solution exists
  - case 2: **normally distributed** returns
    - moments of multi-period returns can be derived from moments of one-period returns: square-root-of-time scaling for standard deviation
    - **multi-period** returns follow **normal distribution**: *VaR* and *ES* can be derived according to **square-root-of-time scaling**

Multi-period moments

- **expectation:** (independence unnecessary)

$$\mathbb{E}\left[r^{log}_{t,t+n-1}\right] = \mathbb{E}\left[\sum_{i=0}^{n-1} r^{log}_{t+i}\right] = \sum_{i=0}^{n-1} \mathbb{E}\left[r^{log}_{t+i}\right] = \sum_{i=0}^{n-1} \mu = n\mu$$

- **variance:**

$$\mathbb{V}\left(r^{log}_{t,t+n-1}\right) = \mathbb{V}\left(\sum_{i=0}^{n-1} r^{log}_{t+i}\right) = \sum_{i=0}^{n-1} \mathbb{V}\left(r^{log}_{t+i}\right) + \sum_{i\neq j}^{n-1} Cov\left(r^{log}_{t+i}, r^{log}_{t+j}\right)$$

$$= \sum_{i=0}^{n-1} \mathbb{V}\left(r^{log}_{t+i}\right) + 0 = n\sigma^2$$

- **standard deviation:**

$$\sigma_{t,t+n-1} = \sqrt{\mathbb{V}\left(r^{log}_{t,t+n-1}\right)} = \sqrt{n\sigma^2} = \sqrt{n}\sigma$$

Distribution of multi-period returns

- assumption: $r_t^{log} \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- consequences:
  - random vector $\left(r_t^{log}, r_{t+k}^{log}\right)$ follows a **bivariate normal distribution** with zero correlation because of **assumed independence**

$$\left(r_t^{log}, r_{t+k}^{log}\right) \sim \mathcal{N}_2\left(\left[\begin{array}{c}\mu \\ \mu\end{array}\right], \left[\begin{array}{cc}\sigma^2 & 0 \\ 0 & \sigma^2\end{array}\right]\right)$$

  - as a **sum** of components of a multi-dimensional **normally distributed** random vector, **multi-period returns** are **normally** distributed themselves
- using formulas for multi-period moments we get

$$r_{t,t+n-1}^{log} \sim \mathcal{N}\left(n\mu, n\sigma^2\right)$$

Multi-period VaR

- notation:
  - $\mu_n := \mathbb{E}\left[r^{log}_{t,t+n-1}\right] = n\mu$
  - $\sigma_n := \sigma_{t,t+n-1} = \sqrt{n}\sigma$
  - $VaR^{(n)}_\alpha := VaR_\alpha\left(r^{log}_{t,t+n-1}\right)$
- rewriting VaR_α for multi-period returns as function of one-period VaR_α :

$$
\begin{aligned}
VaR^{(n)}_\alpha &= -\mu_n + \sigma_n \Phi^{-1}(\alpha) \\
&= -n\mu + \sqrt{n}\sigma\Phi^{-1}(\alpha) \\
&= -n\mu + \sqrt{n}\mu - \sqrt{n}\mu + \sqrt{n}\sigma\Phi^{-1}(\alpha) \\
&= \left(\sqrt{n} - n\right)\mu + \sqrt{n}\left(-\mu + \sigma\Phi^{-1}(\alpha)\right) \\
&= \left(\sqrt{n} - n\right)\mu + \sqrt{n}VaR_\alpha\left(r^{log}_t\right)
\end{aligned}
$$

Multi-period VaR

- furthermore, for the case of $\mu = 0$ we get

$$VaR_{\alpha}^{(n)} = \sqrt{n}\sigma\Phi^{-1}(\alpha) = \sqrt{n}VaR_{\alpha}\left(r_t^{log}\right)$$

- this is known as the **square-root-of-time scaling**
- requirements:
  - returns are independent through time: no autocorrelation
  - returns are normally distributed with zero mean: $r_t^{log} \sim \mathcal{N}(0, \sigma^2)$

Multi-period ES

$$
\begin{aligned}
ES_\alpha^{(n)} &= -\mu_n + \sigma_n \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha} \\
&= -n\mu + \sqrt{n}\sigma \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha} \\
&= \left(\sqrt{n} - n\right)\mu + \sqrt{n}\left(-\mu + \sigma \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha}\right) \\
&= \left(\sqrt{n} - n\right)\mu + \sqrt{n} ES_\alpha
\end{aligned}
$$

- again, for $\mu = 0$ the **square-root-of-time scaling** applies:

$$
ES_\alpha^{(n)} = \sqrt{n} ES_\alpha
$$

Example: market risk

- extending DAX example, with parameters of normal distribution fitted to real world data given by $\hat{\mu} = 0.0344$ and $\hat{\sigma} = 1.5403$
- calculate multi-period *VaR* and *ES* for 5 and 10 periods
- using multi-period formulas for *VaR* and *ES*:

$$\begin{aligned} VaR_{\alpha}^{(n)} &= \left(\sqrt{n} - n\right)\mu + \sqrt{n} VaR_{\alpha}\left(r_t^{log}\right) \\ &= \left(\sqrt{5} - 5\right) \cdot 0.0344 + \sqrt{5} VaR_{\alpha} \end{aligned}$$

$$\begin{aligned} ES_{\alpha}^{(n)} &= \left(\sqrt{n} - n\right)\mu + \sqrt{n} ES_{\alpha} \\ &= \left(\sqrt{5} - 5\right) \cdot 0.0344 + \sqrt{5} ES_{\alpha} \end{aligned}$$

- using previously calculated values, for 5-day returns we get :

$$VaR_{0.99}^{(5)} = \left(\sqrt{5} - 5\right) \cdot 0.0344 + \sqrt{5} \cdot 3.5489$$
$$= -2.7639 + 7.9356$$
$$= 5.1716$$

$$ES_{0.99}^{(5)} = -2.7639 + 9.1191 = 6.3552$$

- for 10-day returns we get:

$$VaR_{0.99}^{(10)} = 10.9874$$

$$ES_{0.99}^{(10)} = -0.2352 + \sqrt{10} \cdot 4.0782 = 12.6612$$

Example: multi-period portfolio loss

- let $S_{t,i}$ denote the price of stock i at time t
- given λ_i shares of stock i, the portfolio value in t is given by

$$P_t = \sum_{i=1}^{d} \lambda_i S_{t,i}$$

- one-day portfolio loss:

$$L_{t+1} = -\left(P_{t+1} - P_t\right)$$

Model setup

- target variable: n-day cumulated portfolio loss for periods $\{t, t+1, \dots, t+n\}$:

$$L_{t,t+n} = -\left(P_{t+n} - P_t\right)$$

- capture uncertainty by **modelling logarithmic returns** $r_t^{log} = \log S_{t+1} - \log S_t$ **as random variables**
- consequence: instead of directly modelling the distribution of our target variable, our model treats it as function of stochastic risk factors, and tries to model the distribution of the risk factors

$$L_{t,t+n} = f\left(r_t^{log}, \ldots, r_{t+n-1}^{log}\right)$$

- **flexibility**: changes in target variable (portfolio changes) do not require re-modelling of the stochastic part at the core of the model

Function of risk factors

$$
\begin{aligned}
L_{t,t+n} &= -\left(P_{t+n} - P_t\right) \\
&= -\left(\sum_{i=1}^{d} \lambda_i S_{t+n,i} - \sum_{i=1}^{d} \lambda_i S_{t,i}\right) \\
&= -\sum_{i=1}^{d} \lambda_i \left(S_{t+n,i} - S_{t,i}\right) \\
&= -\sum_{i=1}^{d} \lambda_i S_{t,i} \left(\frac{S_{t+n,i}}{S_{t,i}} - 1\right) \\
&= -\sum_{i=1}^{d} \lambda_i S_{t,i} \left(exp\left(\log\left(\frac{S_{t+n,i}}{S_{t,i}}\right)\right) - 1\right)
\end{aligned}
$$

Function of risk factors

$$= -\sum_{i=1}^{d} \lambda_i S_{t,i} \left(exp \left(\log \left(\frac{S_{t+n,i}}{S_{t,i}} \right) \right) - 1 \right)$$

$$= -\sum_{i=1}^{d} \lambda_i S_{t,i} \left(exp \left(r^{log}_{(t,t+n),i} \right) - 1 \right)$$

$$= -\sum_{i=1}^{d} \lambda_i S_{t,i} \left(exp \left(\sum_{k=0}^{n} r^{log}_{(t+k),i} \right) - 1 \right)$$

$$= g \left(r^{log}_{t,1}, r^{log}_{t+1,1}, \ldots, r^{log}_{t+n,1}, r^{log}_{t,2}, \ldots, r^{log}_{t,d}, \ldots r^{log}_{t+n,d} \right)$$

- target variable is **non-linear function of risk factors**: non-linearity arises from non-linear portfolio aggregation in logarithmic world

Simplification for dimension of time

- **assuming normally** distributed daily returns r_t^{log} as well as **independence** of daily returns **over time**, we know that multi-period returns

$$r_{t,t+n}^{log} = \sum_{i=0}^{n} r_{t+i}^{log}$$

have to be **normally distributed** with parameters

$$\mu_n = n\mu$$

and

$$\sigma_n = \sqrt{n}\sigma$$

Simplification for dimension of time

- the input parameters can be reduced to

$$L_{t,t+n} = -\sum_{i=1}^{d} \lambda_i S_{t,i} \left(exp \left(r^{log}_{(t,t+n),i} \right) - 1 \right)$$
$$= h \left(r^{log}_{(t,t+n),1}, \ldots, r^{log}_{(t,t+n),d} \right)$$

- non-linearity still holds because of non-linear portfolio aggregation

Application: real world data

- estimating parameters of a normal distribution for historical daily returns of *BMW* and *Daimler* for the period from 01.01.2006 to 31.12.2010 we get

$$\mu^B = -0.0353, \quad \sigma^B = 2.3242$$

and

$$\mu^D = -0.0113 \quad \sigma^D = 2.6003$$

- assuming independence over time, the parameters of 3-day returns are given by

$$\mu_3^B = -0.1058, \quad \sigma_3^B = 4.0256$$

and

$$\mu_3^D = -0.0339, \quad \sigma_3^D = 4.5039$$

Asset dependency

- so far, the marginal distribution of individual one-period returns has been specified, as well as the distribution of multi-period returns through the assumption of independence over time
- however, besides the marginal distributions, in order to make derivations of the model, we also have to **specify the dependence structure** between different assets
- once the dependence structure has been specified, simulating from the complete two-dimensional distribution and plugging into function h gives **Monte Carlo solution** of the target variable

Linearization

- in order to eliminate non-linearity, **approximate function** by linear function

$$f\left(x+\Delta t\right)=f\left(x\right)+f^{'}\left(x\right)\cdot\Delta t$$

- denoting $Z_t := \log S_t$, makes Δt expressable with risk factors:

$$\begin{aligned}
P_{t+1} &= \sum_{i=1}^{d} \lambda_i S_{t+1,i} \\
&= \sum_{i=1}^{d} \lambda_i exp\left(\log\left(S_{t+1,i}\right)\right) \\
&= \sum_{i=1}^{d} \lambda_i exp\left(Z_{t,i} + r_{t+1,i}^{log}\right) \\
&= f\left(Z_t + r_{t+1}^{log}\right)
\end{aligned}$$

Linearization

- function $f(u) = \sum_{i=1}^{d} \lambda_i exp(u_i)$ has to be approximated by differentiation
- differentiating with respect to the single coordinate i:

$$\frac{\partial f(u)}{\partial u_i} = \frac{\partial \left(\sum_{i=1}^{d} \lambda_i exp(u_i) \right)}{\partial u_i} = \lambda_i exp(u_i)$$

$$\begin{aligned}
\Rightarrow f\left(Z_t + r_{t+1}^{log}\right) &\approx f(Z_t) + f'(Z_t)\, r_{t+1}^{log} \\
&= f(Z_t) + \sum_{i=1}^{d} \frac{\partial f(Z_t)}{\partial Z_{t,i}} r_{t+1,i}^{log} \\
&= \sum_{i=1}^{d} \lambda_i exp(Z_{t,i}) + \sum_{i=1}^{d} \lambda_i exp(Z_{t,i})\, r_{t+1,i}^{log} \\
&= \sum_{i=1}^{d} \lambda_i S_{t,i} + \sum_{i=1}^{d} \lambda_i S_{t,i} r_{t+1,i}^{log}
\end{aligned}$$

Linearization

- linearization of one-period portfolio loss:

$$\begin{aligned} L_{t+1} &= -\left(P_{t+1} - P_t\right) \\ &\approx -\left(\sum_{i=1}^{d} \lambda_i S_{t,i} + \sum_{i=1}^{d} \lambda_i S_{t,i} r_{t+1,i}^{log} - \sum_{i=1}^{d} \lambda_i S_{t,i}\right) \\ &= -\left(\sum_{i=1}^{d} \lambda_i S_{t,i} r_{t+1,i}^{log}\right) \\ &= a_1 r_{t+1,1}^{log} + a_2 r_{t+1,2}^{log} + \ldots + a_d r_{t+1,d}^{log} \end{aligned}$$

- linearization of 3-period portfolio loss:

$$L_{t,t+2} \approx -\left(\sum_{i=1}^{d} \lambda_i S_{t,i} r_{(t,t+2),i}^{log}\right)$$

- now that non-linearities have been removed, make use of fact that linear function of normally distributed returns is still normally distributed
- **assuming** the dependence structure between daily returns of *BMW* and *Daimler* to be symmetric, joint returns will follow a **bivariate normal distribution**, and 3-day returns of *BMW* and *Daimler* also follow a joint normal distribution
- given the covariance of daily returns, the covariance of 3-day returns can be calculated according to

$$
\begin{aligned}
Cov\left(r^B_{t,t+2}, r^D_{t,t+2}\right) &= Cov\left(r^B_t + r^B_{t+1} + r^B_{t+2}, r^D_t + r^D_{t+1} + r^D_{t+2}\right) \\
&= \sum_{i,j=0}^{2} Cov\left(r^B_{t+i}, r^D_{t+j}\right) \\
&= Cov\left(r^B_t, r^D_t\right) + Cov\left(r^B_{t+1}, r^D_{t+1}\right) + Cov\left(r^B_{t+2}, r^D_{t+2}\right) \\
&= 3Cov\left(r^B_t, r^D_t\right)
\end{aligned}
$$

Application

- with estimated correlation $\hat{\rho} = 0.7768$, the covariance becomes

$$\widehat{Cov}\left(r_t^B, r_t^D\right) = \hat{\rho} \cdot \left(\sigma^B\right) \cdot \left(\sigma^D\right) = 0.7768 \cdot 2.3242 \cdot 2.6003 = 4.6947$$

and

$$\widehat{Cov}\left(r_{t,t+2}^B, r_{t,t+2}^D\right) = 14.0841$$

- simulating from two-dimensional normal distribution, and plugging into function h will give a simple and fast approximation of the distribution of the target variable
- as the target variable is a **linear function of jointly normally** distributed risk factors, it has to be normally distributed itself: hence, an **analytical solution** is possible

Recapturing involved assumptions

- individual **daily logarithmic returns** follow **normal** distribution
- returns are **independent over time**
- **non-linear function** for target variable has been **approximated by linearization**
- **dependence structure** according to **joint normal** distribution

Example: coherence

Consider a portfolio consisting of $d = 100$ corporate bonds. The probability of default shall be 0.5% for each firm, with occurrence of **default independently of each other**. Given no default occurs, the value of the associated bond increases from $x_t = 100$ € this year to $x_{t+1} = 102$ € next year, while the value decreases to 0 in the event of default.
Calculate $VaR_{0.99}$ for a portfolio A consisting of 100 shares of **one single given corporate**, as well as for a portfolio B, which consists of one **share of each** of the 100 different corporate bonds. Interpret the results. What does that mean for VaR as a risk measure, and what can be said about Expected Shortfall with regard to this feature?

Example

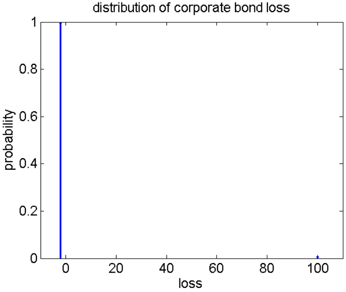
- **setting**: $d = 100$ different corporate bonds, each with values given by

	t	$t+1$	
value	100	102	0
probability		0.995	0.005

- defaults are independent of each other

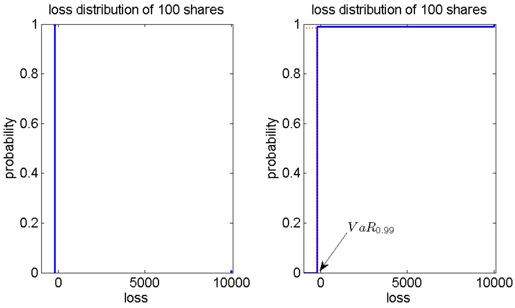
Example

- associated loss distribution:



Example

- portfolio A : 100 bonds of one given firm
- $VaR^A_{0.99} = \inf\{l \in \mathbb{R} : F_L(l) \geq 0.99\} = -200$



- portfolio B: 100 bonds, one of each firm
- number of defaults are distributed according to Binomial distribution:

$$\mathbb{P}\left(\textit{no defaults}\right) = 0.995^{100} = 0.6058$$

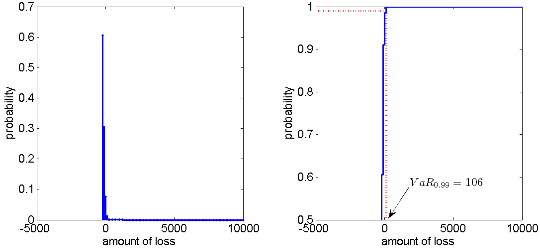
$$\mathbb{P}\left(\textit{one default}\right) = 0.995^{99} \cdot 0.005 \cdot \begin{pmatrix} 100 \\ 1 \end{pmatrix} = 0.3044$$

$$\mathbb{P}\left(\textit{two defaults}\right) = 0.995^{98} \cdot 0.005^{2} \cdot \begin{pmatrix} 100 \\ 2 \end{pmatrix} = 0.0757$$

$$\mathbb{P}\left(\textit{three defaults}\right) = 0.995^{97} \cdot 0.005^{3} \cdot \begin{pmatrix} 100 \\ 3 \end{pmatrix} = 0.0124$$

- hence, because of $\mathbb{P}\left(\textit{defaults} \leq 2\right) = 0.9859$ and $\mathbb{P}\left(\textit{defaults} \leq 3\right) = 0.9983$, to be protected with probability of at least 99%, the capital cushion has to be high enough to offset the losses associated with 3 defaults
- losses for 3 defaults: $-2 \cdot 97 + 100 \cdot 3 = 106$
- hence, $VaR^{B}_{0.99} = 106$

- resulting loss distribution:



- note: due to **diversification** effects, the risk inherent to portfolio B should be less than the risk inherent to portfolio A
- *VaR* as a measure of risk fails to account for this reduction of risk: it is **not subadditiv**
- *ES* does fulfill this property: it is **subadditiv**

Coherence of risk measures

Definition

Let $\mathcal{L}$ denote the set of all possible loss distributions which are almost surely finite. A risk measure ϱ is called **coherent** if it satisfies the following axioms:

Translation Invariance: For all $L \in \mathcal{L}$ and every $c \in \mathbb{R}$ we have

$$\varrho(L + c) = \varrho(L) + c.$$

It follows that $\varrho(L - \varrho(L)) = \varrho(L) - \varrho(L) = 0$: the portfolio hedged by an amount equal to the measured risk does not entail risk anymore

Subadditivity: For all $L_1, L_2 \in \mathcal{L}$ we have

$$\varrho(L_1 + L_2) \leq \varrho(L_1) + \varrho(L_2).$$

Due to diversification, the joint risk cannot be higher than the individual ones.

Coherence of risk measures

Definition

Positive Homogeneity: For all $L \in \mathcal{L}$ and every $\lambda > 0$ we have

$$\varrho(\lambda L) = \lambda \varrho(L).$$

Increasing the exposure by a factor of λ also increases the risk measured by the same factor.

Monotonicity: For all $L_1, L_2 \in \mathcal{L}$ such that $L_1 \leq L_2$ almost surely we have

$$\varrho(L_1) \leq \varrho(L_2).$$