Slides for Risk Management

Introduction to the modeling of assets

Groll

Seminar für Finanzökonometrie

Prof. Mittnik, PhD

1. Interest rates and returns
   - Fixed-income assets
   - Speculative assets

2. Probability theory
   - Probability space and random variables
   - Information reduction
   - Updating information
   - Functions of random variables
   - Monte Carlo Simulation
   - Measures under transformation

two broad types of investments:

fixed-income assets

- payments are known in advance
- only risk is risk of losses due to the failure of a counterparty to fulfill its contractual obligations: called **credit risk**

speculative assets

- characterized by random price movements
- modelled in a stochastic framework using random variables

Interest and Compounding

- given an interest rate of r per period and initial wealth W_t, the wealth one period ahead is calculated as

$$W_{t+1} = W_t\,(1+r)$$

Example

- $r = 0.05$ (annual rate), $W_0 = 500.000$, after one year:

$$500.000\left(1+\frac{5}{100}\right) = 500.000\,(1+0.05) = 525.000$$

- **compound interest** in general:

$$W_T(r, W_0) = W_0(1+r)^T$$

Compounding at higher frequency

- compounding can occur more frequently than at annual intervals
- m times per year: $W_{m,t}(r)$ denotes wealth in t for $W_0 = 1$

biannually

after six months:

$$W_{2,\frac{1}{2}}(r) = \left(1 + \frac{r}{2}\right)$$

after one year:

$$W_{2,1}(r) = \left(1 + \frac{r}{2}\right)\left(1 + \frac{r}{2}\right) = \left(1 + \frac{r}{2}\right)^2$$

- the **effective annual rate** exceeds the simple annual rate:

$$\left(1 + \frac{r}{2}\right)^2 > (1 + r) \Rightarrow W_{2,1}(r) > W_{1,1}(r)$$

Effective annual rate

m interest payments within a year

effective annual rate after one year:

$$W_{m,1}(r) := \left(1 + \frac{r}{m}\right)^m$$

after T years:

$$W_{m,T}(r) = \left(1 + \frac{r}{m}\right)^{mT}$$

- wealth is an increasing function of the interest payment frequency:

$$W_{m_1,t}(r) > W_{m_2,t}(r), \forall t \text{ and } m_1 > m_2$$

Continuous compounding

- the **continuously compounded rate** is given by the limit

$$W_{\infty,1}(r) = \lim_{m\rightarrow\infty}\left(1+\frac{r}{m}\right)^{m} = e^{r}$$

- compounding over T periods leads to

$$W_{\infty,T}(r) = \lim_{m\rightarrow\infty}\left(1+\frac{r}{m}\right)^{mT} = \left(\lim_{m\rightarrow\infty}\left(1+\frac{r}{m}\right)^{m}\right)^{T} = e^{rT}$$

- under continous compounding the value of an initial investment of W_0 grows **exponentially fast**
- comparatively simple for calculation of interest accrued in **between** dates of interest **payments**

Comparison of different interest rate frequencies

T	$m = 1$	$m = 2$	$m = 4$	∞
1	1030	1030.2	1030.3	1030.5
2	1060.9	1061.4	1061.6	1061.8
3	1092.7	1093.4	1093.8	1094.2
4	1125.5	1126.5	1127	1127.5
5	1159.3	1160.5	1161.2	1161.8
6	1194.1	1195.6	1196.4	1197.2
7	1229.9	1231.8	1232.7	1233.7
8	1266.8	1269	1270.1	1271.2
9	1304.8	1307.3	1308.6	1310
10	1343.9	1346.9	1348.3	1349.9

Table: Development of initial investment $W_0 = 1000$ over 10 years, subject to different interest rate frequencies, with annual interest rate $r = 0.03$

Non-constant interest rates

- for the case of **changing annual interest rates**, end-of-period wealth of **annually compounded** interest rates is given by

$$\begin{aligned} W_{1,t} &= (1+r_0)\cdot(1+r_1)\cdot\ldots\cdot(1+r_{t-1}) \\ &= \prod_{i=0}^{t-1}(1+r_i) \end{aligned}$$

- for **continuously compounded** interest rates, end-of-period wealth is given by

$$\begin{aligned} W_{\infty,t} &= \left(\lim_{m\rightarrow\infty}\left(1+\frac{r_0}{m}\right)^m\right)\cdot\ldots\cdot\left(\lim_{m\rightarrow\infty}\left(1+\frac{r_{t-1}}{m}\right)^m\right) \\ &= e^{r_0}\cdot e^{r_1}\cdot\ldots\cdot e^{r_{t-1}} \\ &= e^{r_0+\ldots+r_{t-1}} \\ &= exp\left(\sum_{i=0}^{t-1} r_i\right) \end{aligned}$$

Regarding continuous compounding

- **Why** bother with **continuous compounding**, as interest rates in the real world are always given at finite frequency?

→ the key to the answer of this question lies in the transformation of the **product** of returns **into a sum**

- as **interest rates** of fixed-income assets are assumed to be **perfectly known**, summation instead of multiplication only yields **minor advantages** in a world of computers
- however, as soon as **payments** are **uncertain** and have to be modelled as random variables, this transformation will make a **huge difference**

Returns on speculative assets

- let P_t denote the price of a speculative asset at time t
- **net return** during period t:

$$r_t := \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1$$

- **gross return** during period t:

$$R_t := (1 + r_t) = \frac{P_t}{P_{t-1}}$$

- returns calculated this way are called **discrete returns**
- returns on speculative assets vary from period to period

Calculating returns from prices

- while interest rates of fixed-income assets are usually known **prior** to the investment, returns of speculative assets have to be calculated **after** observation of prices

discrete case

$$P_T = P_0 (1 + r)^T \Leftrightarrow \sqrt[T]{\frac{P_T}{P_0}} = 1 + r$$

$$\Rightarrow r = \sqrt[T]{\frac{P_T}{P_0}} - 1$$

Continuously compounded returns

- defining the **log return**, or **continuously compounded return**, by

$$r_t^{log} := \ln R_t = \ln (1 + r_t) = \ln \frac{P_t}{P_{t-1}} = \ln P_t - \ln P_{t-1}$$

Exercise

Investor A and investor B both made one investment each. While investor A was able to increase his investment sum of 100 to 140 within 3 years, investor B increased his initial wealth of 230 to 340 within 5 years. Which investor did perform better?

Exercise: solution

- calculate mean annual interest rate for both investors
- investor A :

$$\begin{aligned} P_T &= P_0 \left(1+r\right)^T && \Leftrightarrow \\ 140 &= 100 \left(1+r\right)^3 && \Leftrightarrow \\ \sqrt[3]{\frac{140}{100}} &= (1+r) && \Leftrightarrow \\ r_A &= 0.1187 \end{aligned}$$

- investor B :

$$r_B = \left(\sqrt[5]{\frac{340}{230}} - 1\right) = 0.0813$$

- hence, investor A has achieved a higher return on his investment

Exercise: solution for continuous returns

- for comparison, solution of the exercise with respect to continous returns
- **continuously compounded returns** associated with an evolution of prices over a longer time period is given by

continuous case

$$P_T = P_0 e^{rT} \Leftrightarrow \frac{P_T}{P_0} = e^{rT} \Leftrightarrow \ln\left(\frac{P_T}{P_0}\right) = \ln\left(e^{rT}\right) = rT$$

$$r = \frac{(\ln P_T - \ln P_0)}{T}$$

Exercise: solution for continuous returns

- plugging in leads to

$$r_A = \frac{(\ln 140 - \ln 100)}{3} = 0.1121$$

$$r_B = \frac{(\ln 340 - \ln 230)}{5} = 0.0782$$

- conclusion: while the case of discrete returns involves calculation of the n-th root, the continuous case is computationally less demanding
- while continuous returns differ from their discrete counterparts, the ordering of both investors is unchanged
- keep in mind: **so far** we only treat returns retrospectively, that is, with given and **known realization of prices**, where any uncertainty involved in asset price evolutions already has been resolved

Aggregating returns

- compounded gross return over $n+1$ sub-periods:

$$
\begin{aligned}
R_{t,t+n} &:= R_t \cdot R_{t+1} \cdot R_{t+2} \cdot \ldots \cdot R_{t+n} \\
&= \frac{P_t}{P_{t-1}} \cdot \frac{P_{t+1}}{P_t} \cdot \ldots \cdot \frac{P_{t+n}}{P_{t+n-1}} \\
&= \frac{P_{t+n}}{P_{t-1}}
\end{aligned}
$$

Example

investment $P_0 = 100$, net returns in percent $[3, -2, 4, 3, -1]$:

$$R_{0,4} = (1.03)\,(0.98)\,(1.04)\,(1.03)\,(0.99) = 1.075$$

$$P_4 = 100 \cdot 1.075 = 107.5$$

$$R_{0,4} = \frac{P_4}{P_0} = \frac{107.5}{100} = 1.075$$

Comparing different investments

- comparison of returns of alternative investment opportunities over different investment horizons requires computation of an **"average" gross return** $\bar{R}$ for each investment, fulfilling:

$$P_t \bar{R}^n \overset{!}{=} P_t R_t \cdot \ldots \cdot R_{t+n-1} = P_{t+n}$$

- in **net returns**:

$$P_t \left(1 + \bar{r}\right)^n \overset{!}{=} P_t \left(1 + r_t\right) \cdot \ldots \cdot \left(1 + r_{t+n-1}\right)$$

- solving for $\bar{r}$ leads to

$$\bar{r} = \left(\prod_{i=0}^{n-1} \left(1 + r_{t+i}\right) \right)^{1/n} - 1$$

- the **annualized gross return** is not an **arithmetic** mean, but a **geometric** mean

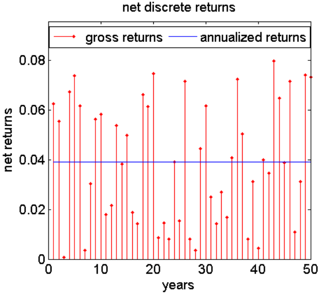
Aggregating continuous returns

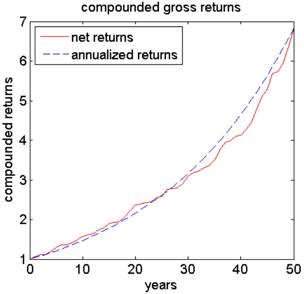
- when aggregating **log returns** instead of discrete returns, we are dealing with a **sum rather than a product** of sub-period returns:

$$\begin{aligned} r_{t,t+n}^{log} &:= \ln\left(1 + r_{t,t+n}\right) \\ &= \ln\left[\left(1 + r_t\right)\left(1 + r_{t+1}\right)\ldots\left(1 + r_{t+n}\right)\right] \\ &= \ln\left(1 + r_t\right) + \ln\left(1 + r_{t+1}\right) + \ldots + \ln\left(1 + r_{t+n}\right) \\ &= r_t^{log} + r_{t+1}^{log} + \ldots + r_{t+n}^{log} \end{aligned}$$

Example

The annualized return of 1.0392 is **unequal** to the simple arithmetic mean over the randomly generated interest rates of 1.0395!





Left: randomly generated returns between 0 and 8 percent, plotted against annualized net return rate. Right: comparison of associated compound interest rates.

Example

- two ways to calculate annualized net returns for previously generated random returns:

direct way

using gross returns, taking 50-th root:

$$\begin{aligned}\bar{r}^{ann}_{t,t+n-1} &= \left(\prod_{i=0}^{n-1}(1+r_{t+i})\right)^{1/n} - 1 \\ &= (1.0626 \cdot 1.0555 \cdot \ldots \cdot 1.0734)^{1/50} - 1 \\ &= (6.8269)^{1/50} - 1 \\ &= 0.0391\end{aligned}$$

via log returns

transfer the problem to the "logarithmic world":

- convert gross returns to log returns:

$$[1.0626, 1.0555, \ldots, 1.0734] \xrightarrow{log} [0.0607, 0.0540, \ldots, 0.0708]$$

- use arithmetic mean to calculate annualized return in the "logarithmic world":

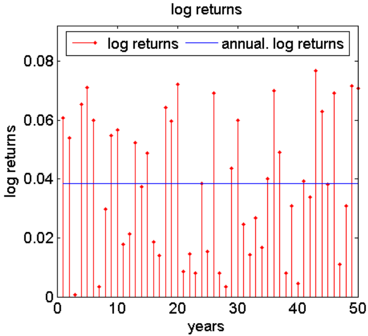
$$r_{t,t+n-1}^{log} = \sum_{i=0}^{n-1} r_{t+i}^{log} = (0.0607 + 0.0540 + ... + 0.0708) = 1.9226$$

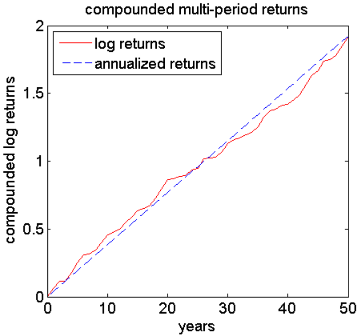
$$\bar{r}_{t,t+n-1}^{log} = \frac{1}{n} r_{t,t+n-1}^{log} = \frac{1}{50} 1.9226 = 0.0385$$

- convert result back to "normal world":

$$\bar{r}_{t,t+n-1}^{ann} = e^{\bar{r}_{t,t+n-1}^{log}} - 1 = e^{0.0385} - 1 = 0.0391$$

Example





Note: given a constant one-period return, the multi-period return **increases linearly** in the logarithmic world

Summary

- multi-period gross returns result from **multiplication** of one-period returns: hence, **exponentially increasing**
- multi-period logarithmic returns result from **summation** of one-period returns: hence, **linearly increasing**
- different calculation of returns from given portfolio values:

$$r_t = \frac{P_t - P_{t-1}}{P_t} \qquad r_t^{log} = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln P_t - \ln P_{t-1}$$

- however, because of

$$\ln(1+x) \approx x$$

discrete net returns and log returns are approximately equal:

$$r_t^{log} = \ln(R_t) = \ln(1+r_t) \approx r_t$$

Conclusions for known price evolutions

- given that prices / returns are already known, with **no uncertainty** left, **continuous** returns are computationally **more efficient**
- discrete returns can be calculated via a detour to continuous returns
- as the transformation of discrete to continuous returns does not change the ordering of investments, and as **logarithmic returns** are **still interpretable** since they are the limiting case of discrete compounding, why shouldn't we just **stick with continous returns overall**?
- however: the **main advantage** only crops up in a setting of **uncertain future returns**, and their modelling as random variables!

Outlook: returns under uncertainty

- **central limit theorem** could justify modelling **logarithmic** returns as **normally distributed**, since returns can be decomposed into **summation** over returns of **lower** frequency: e.g. annual returns are the sum of 12 monthly returns, 52 weakly returns, 365 daily returns,...
- independent of the distribution of low frequency returns, the **central limit theorem** states that any sum of these low frequency returns follows a **normal distribution**, provided that the sum involves sufficiently many summands, and the following requirements are fulfilled:
    - the low frequency returns are **independent** of each other
    - the distribution of the low frequency returns allows finite second moments (variance)

Outlook: returns under uncertainty

- this reasoning **does not apply to net / gross returns**, since they can not be decomposed into a sum of lower frequency returns
- keep in mind: these are **only hypothetical considerations**, since we have not seen any real world data so far!

Randomness

Probability theory

- **randomness**: the result is not known in advance
- **sample space** Ω: set of all possible **outcomes** or **elementary events** ω
- examples for **discrete** sample space:
  - roulette: $\Omega_1 = \{\text{red,black}\}$
  - performance: $\Omega_2 = \{\text{good,moderate,bad}\}$
  - die: $\Omega_3 = \{1, 2, 3, 4, 5, 6\}$
- examples for **continuous** sample space:
  - temperature: $\Omega_4 = [-40, 50]$
  - log-returns: $\Omega_5 =]-\infty, \infty[$

Events

- a subset $A \subset \Omega$ consisting of more than one **elementary event** ω is called **event**

examples

- "at least moderate performance": $A = \{\text{good,moderate}\} \subset \Omega_2$
- "even number": $A = \{2, 4, 6\} \subset \Omega_3$
- "warmer than 10 degrees": $A =]10, \infty[\subset \Omega_4$

- the set of all events of Ω is called **event space** $\mathcal{F}$
- usually it contains all possible subsets of Ω: it is the **power set** of $\mathcal{P}(\Omega)$

Events

event space example

$\mathcal{P}(\Omega_2) = \{\Omega, \{\}\} \cup \{\text{good}\} \cup \{\text{moderate}\} \cup \{\text{bad}\} \cup \{\text{good,moderate}\} \cup \{\text{good,bad}\} \cup \{\text{moderate,bad}\}$

- $\{\}$ denotes the **empty set**
- an event A is said to **occur** if any $\omega \in A$ occurs

example

If the performance happens to be $\omega = \{\text{good}\}$, then also the event $A =$"at least moderate performance" has occured, since $\omega \subset A$.

Probability measure

probability measure

- quantifies for each event a probability of occurance
- real-valued set function $\mathbb{P} : \mathcal{F} \rightarrow \mathbb{R}$, with $\mathbb{P}(A)$ denoting the probability of A, and properties
  1. $\mathbb{P}(A) > 0$ for all $A \subseteq \Omega$
  2. $\mathbb{P}(\Omega) = 1$
  3. For each finite or countably infinite collection of **disjoint** events (A_i) it holds:

$$\mathbb{P}(\cup_{i \in I} A_i) = \sum_{i \in I} \mathbb{P}(A_i)$$

Definition

The 3-tuple $\{\Omega, \mathcal{F}, \mathbb{P}\}$ is called **probability space**.

Random variable

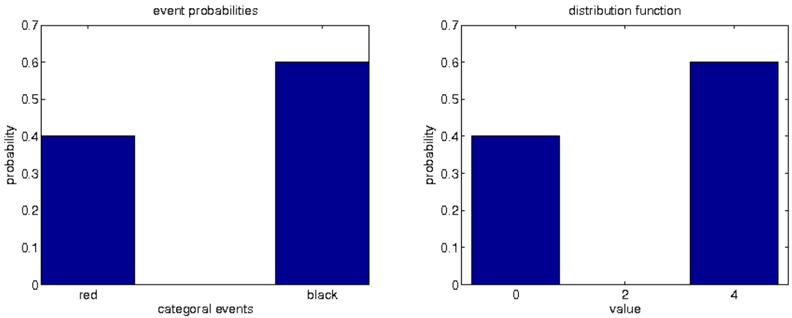
- instead of outcome ω itself, usually a mapping or function of ω is in the focus: when playing roulette, instead of outcome "red" it is more useful to consider associated gain or loss of a bet on "color"
- conversion of **categoral** outcomes to **real numbers** allows for further measurements / information extraction: expectation, dispersion,...

Definition

Let $\{\Omega, \mathcal{F}, \mathbb{P}\}$ be a probability space. If $X : \Omega \rightarrow \mathbb{R}$ is a real-valued function with the elements of Ω as its domain, then X is called **random variable**.

- a **discrete** random variable consists of a countable number of elements, while a **continuous** random variable can take any real value in a given interval

Example


event probabilities
probability
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
red
black
categoral events
distribution function
probability
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
value

Density function

- a **probability density function** determines the probability (possibly 0) for each event

discrete density function

For each $x_i \in X(\Omega) = \{x_i | x_i = X(\omega), \omega \in \Omega\}$, the function

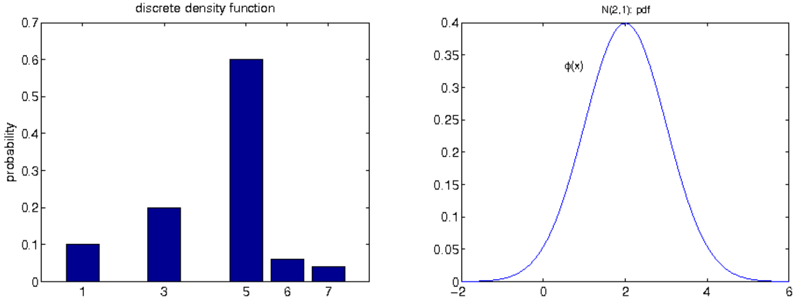
$$f(x_i) = \mathbb{P}(X = x_i)$$

assigns a value corresponding to the probability.

continuous density function

In contrast, the values of a continuous density function $f(x)$, $x \in \{x | x = X(\omega), \omega \in \Omega\}$ are not probabilities itself. However, they shed light on the relative probabilities of occurrence. Given $f(y) = 2 \cdot f(z)$, the occurrence of y is twice as probable as the occurrence of z.

Example


discrete density function
probability
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
3
5
6
7
N(2,1): pdf
φ(x)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−2
0
2
4
6

Cumulative distribution function

Definition

The **cumulative distribution function** (cdf) of random variable X, denoted by $F(x)$, indicates the probability that X assumes a value that is lower than or equal to x, where x is any real number. That is

$$F(x) = \mathbb{P}(X \leq x), \quad -\infty < x < \infty.$$

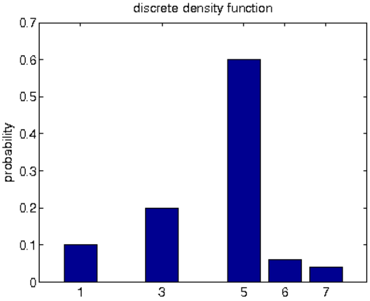
- a cdf has the following properties:
  1. $F(x)$ is a nondecreasing function of x;
  2. $\lim_{x\to\infty} F(x) = 1$;
  3. $\lim_{x\to-\infty} F(x) = 0$.
- furthermore:

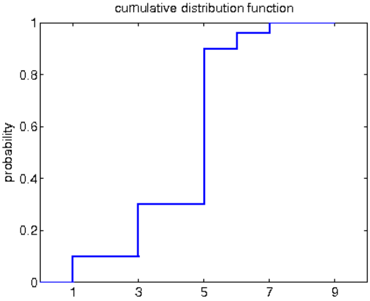
$$\mathbb{P}(a < X \leq b) = F(b) - F(a), \quad \text{for all } b > a$$

Interrelation pdf and cdf

Discrete case:

$$F(x) = \mathbb{P}(X \leq x) = \sum_{x_i \leq x} \mathbb{P}(X = x_i)$$

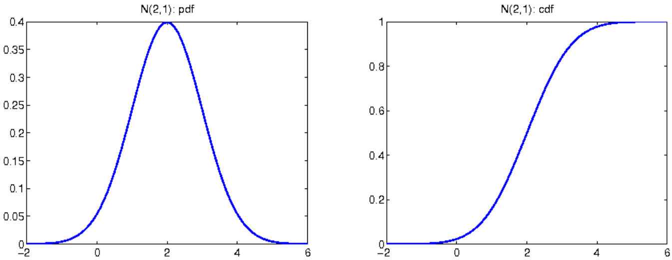




Interrelation pdf and cdf

Continuous case:

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(u)\, du$$



Modelling information

- both cdf as well as pdf, which is the derivative of the cdf, provide complete information about the distribution of the random variable
- may not always be necessary / possible to have complete distribution
- incomplete information modelled via **event space** $\mathcal{F}$

Example

- sample space given by $\Omega = \{1, 3, 5, 6, 7\}$
- modelling complete information about possible realizations:

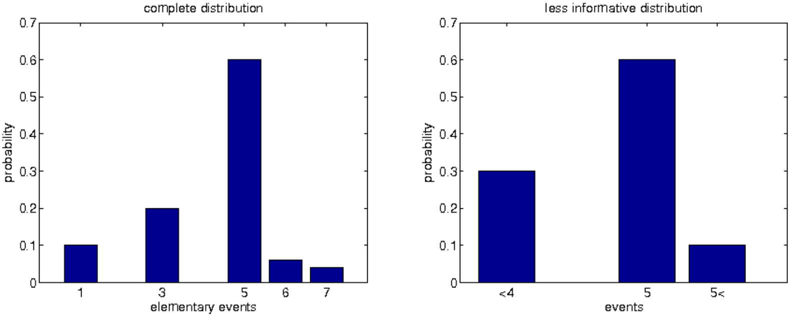
$$\begin{aligned}\mathcal{P}(\Omega) = & \{1\} \cup \{3\} \cup \{5\} \cup \{6\} \cup \{7\} \cup \\ & \cup \{1, 3\} \cup \{1, 5\} \cup \ldots \cup \{6, 7\} \cup \{1, 3, 5\} \cup \ldots \cup \{5, 6, 7\} \cup \\ & \cup \{1, 3, 5, 6\} \cup \ldots \cup \{3, 5, 6, 7\} \cup \{\Omega, \{\}\}\end{aligned}$$

- example of event space representing incomplete information could be

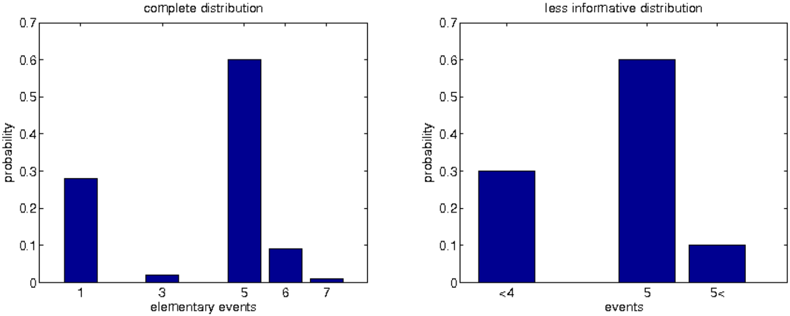
$$\mathcal{F} = \{\{1, 3\}, \{5\}, \{6, 7\}\} \cup \{\{1, 3, 5\}, \{1, 3, 6, 7\}, \{5, 6, 7\}\} \cup \{\Omega, \{\}\}$$

- given only incomplete information, originally distinct distributions can become indistinguishable

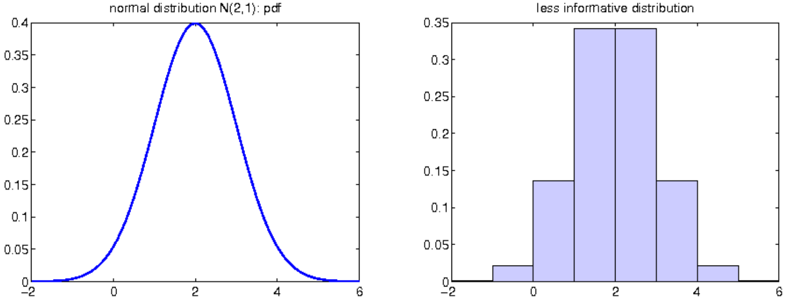
Information reduction discrete



Information reduction discrete



Information reduction continuous



Measures of random variables

- complete distribution may not always be necessary
- classification with respect to several measures can be sufficient:
  - probability of negative / positive return
  - return on average
  - worst case
- compress information of complete distribution for better comparability with other distributions
- compressed information is easier to interpret
- example: knowing **"central location"** together with an idea by how much X may **fluctuate** around the center may be sufficient
- measures of **location** and **dispersion**
- given only incomplete information conveyed by measures, distinct distributions can become indistinguishable

Expectation

The **expectation**, or **mean**, is defined as a weighted average of all possible realizations of a random variable.

discrete random variables

The **expected value** $\mathbb{E}[X]$ is defined as

$$\mathbb{E}[X] = \mu_X = \sum_{i=1}^{N} x_i \mathbb{P}(X = x_i).$$

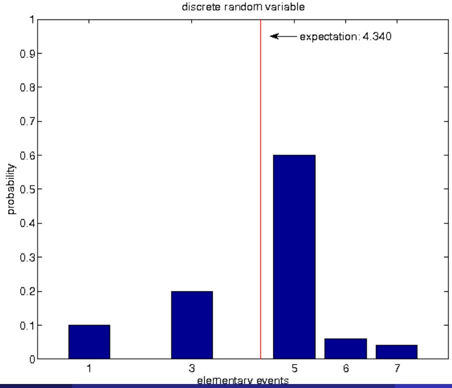
continuous random variables

For a continuous random variable with density function $f(x)$:

$$\mathbb{E}[X] = \mu_X = \int_{-\infty}^{\infty} x f(x)\, dx$$

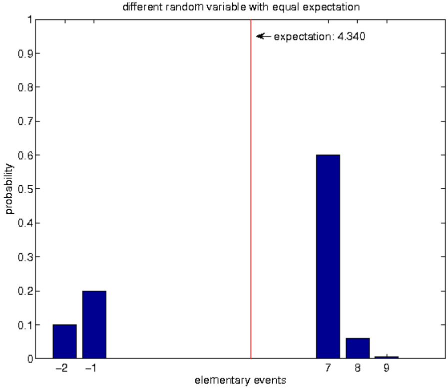
Example

$$\mathbb{E}[X] = \sum_{i=1}^{5} x_i \mathbb{P}(X = x_i)$$
$$= 1 \cdot 0.1 + 3 \cdot 0.2 + 5 \cdot 0.6 + 6 \cdot 0.06 + 7 \cdot 0.04 = 4.34$$



Example

$$\mathbb{E}[X] = -2 \cdot 0.1 - 1 \cdot 0.2 + 7 \cdot 0.6 + 8 \cdot 0.06 + 9 \cdot 0.0067 = 4.34$$



Variance

The **variance** provides a measure of **dispersion** around the mean.

discrete random variables

The **variance** is defined by

$$\mathbb{V}[X] = \sigma_X^2 = \sum_{i=1}^{N} (X_i - \mu_X)^2 \, \mathbb{P}(X = x_i),$$

where $\sigma_X = \sqrt{\mathbb{V}[X]}$ denotes the **standard deviation** of X.

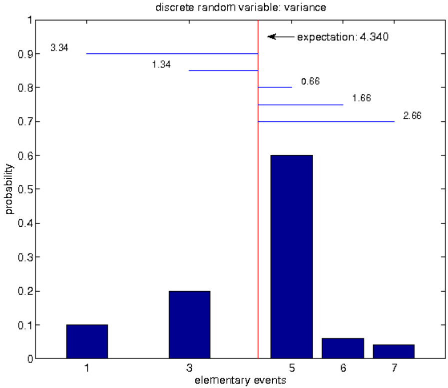
continuous random variables

For continuous variables, the **variance** is defined by

$$\mathbb{V}[X] = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x) \, dx$$

Example

$$\mathbb{V}[X] = \sum_{i=1}^{5} (x_i - \mu)^2 \, \mathbb{P}(X = x_i)$$
$$= 3.34^2 \cdot 0.1 + 1.34^2 \cdot 0.2 + 0.66^2 \cdot 0.6 + 1.66^2 \cdot 0.06 + 2.66^2 \cdot 0.04$$
$$= 2.1844 \neq 14.913$$



Quantiles

Quantile

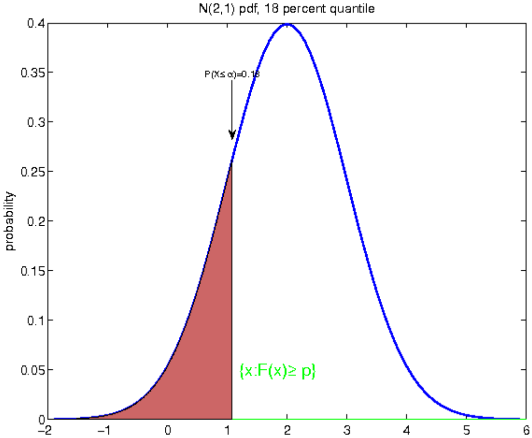
Let X be a random variable with cumulative distribution function F. For each $p \in (0,1)$, the p-**quantile** is defined as

$$F^{-1}(p) = \inf \{x | F(x) \geq p\}.$$

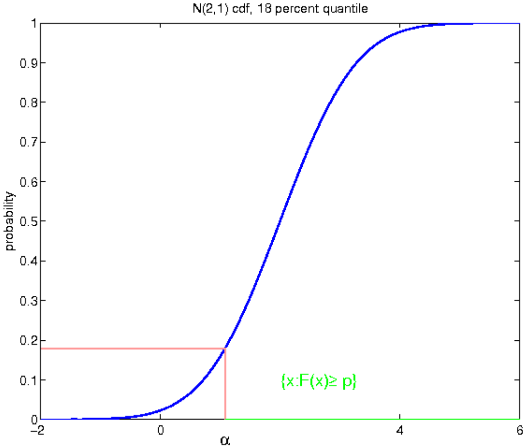
- **measure of location**
- divides distribution in two parts, with **exactly** $p * 100$ **percent** of the probability mass of the distribution to the left **in the continuous case:** random draws from the given distribution F would fall $p * 100$ percent of the time below the p-quantile
- for **discrete** distributions, the probability mass on the left has to be at least $p * 100$ percent:

$$F\left(F^{-1}(p)\right) = \mathbb{P}\left(X \leq F^{-1}(p)\right) \geq p$$

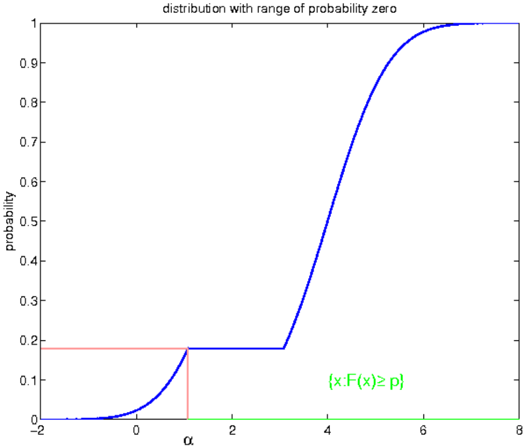
Example


N(2,1) pdf, 18 percent quantile
probability
P(X≤α)=0.18
{x:F(x)≥ p}
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−2
−1
0
1
2
3
4
5
6

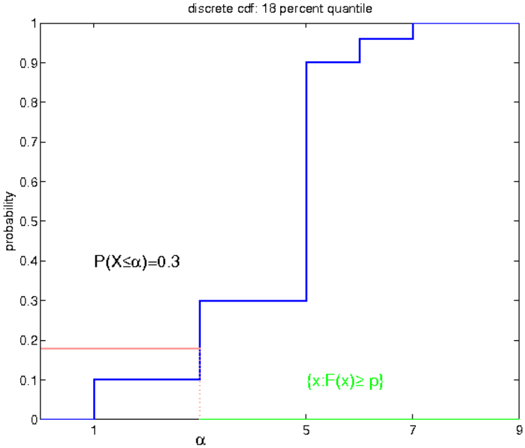
Example: cdf


N(2,1) cdf, 18 percent quantile
probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−2
0
4
6
α
{x:F(x)≥ p}

Example


distribution with range of probability zero
probability
α
{x:F(x)≥ p}
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−2
0
2
4
6
8

Example


discrete cdf: 18 percent quantile
probability
P(X≤α)=0.3
{x:F(x)≥ p}
α
1
5
7
9
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Information reduction / updating

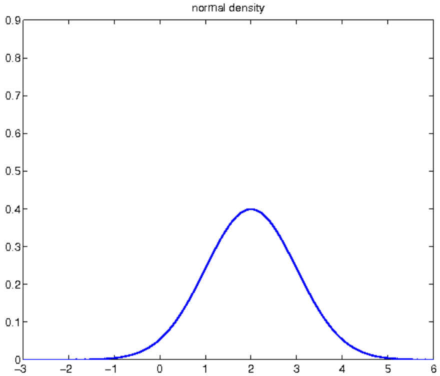
summary: information reduction

- **incomplete information** can occur in two ways:
  - a **coarse filtration**
  - only values of some **measures** of the underlying distribution are known (**mean**, **dispersion**, **quantiles**)
- any reduction of information implicitly induces that some formerly distinguishable distributions are **undistinguishable** on the basis of the limited information
- **tradeoff**: reducing information for better **comprehensibility** / **comparability**, or keeping as much information as possible

- opposite direction: **updating** information on the basis of new arriving information
- concept of **conditional probability**

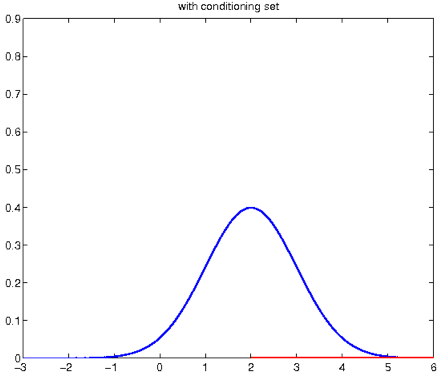
Example

- with knowledge of the underlying distribution, the information has to be updated, given that the occurrence of some event of the filtration is known



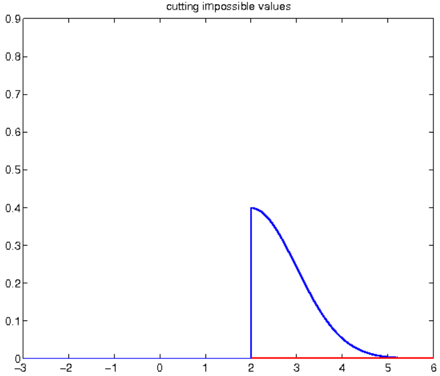
Conditional density

- normal distribution with mean 2
- incorporating the knowledge of a realization **greater than the mean**



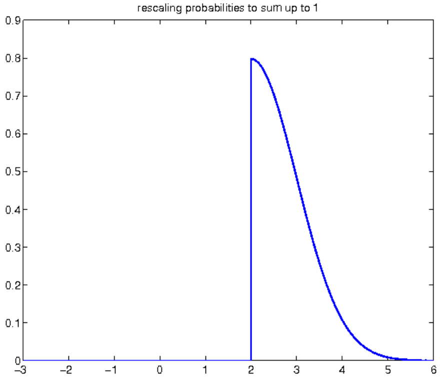
Conditional density

- given the knowledge of a realization higher than 2, probabilities of values below become zero



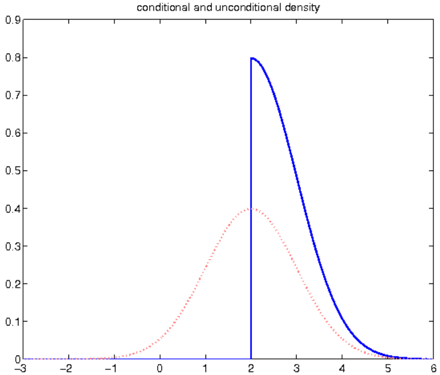
Conditional density

- without changing relative proportions, the density has to be rescaled in order to enclose an area of 1

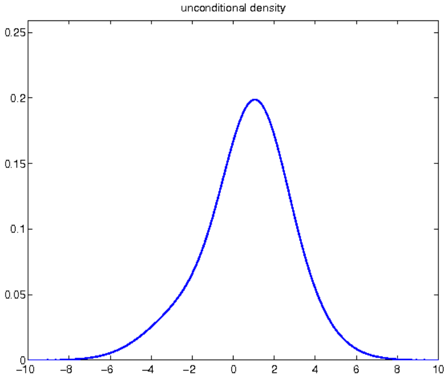


Conditional density

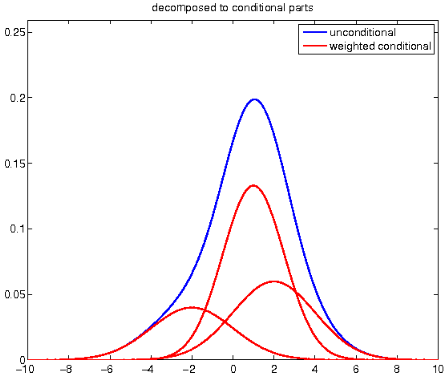
- original density function compared to updated conditional density



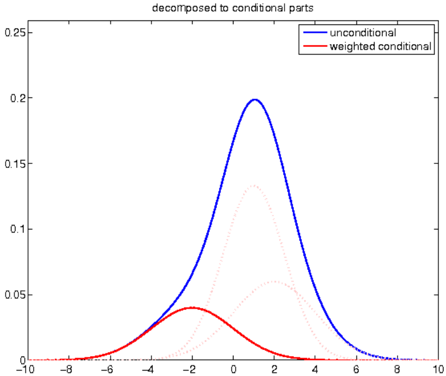
Decompose density



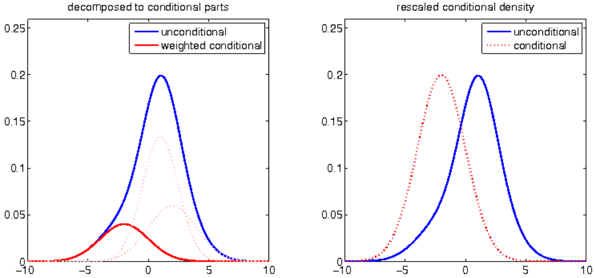
Decompose density



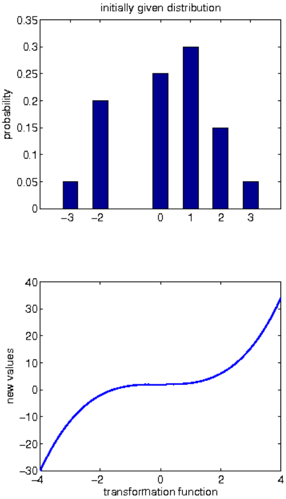
Decompose density

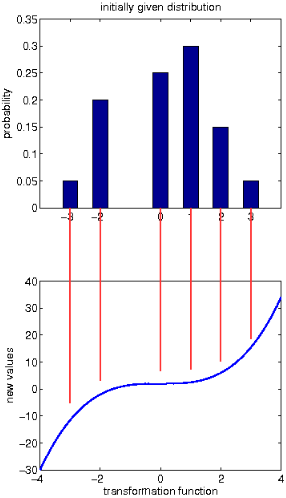


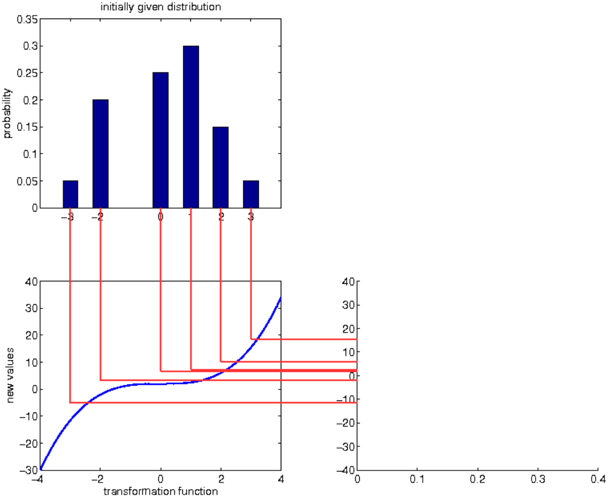
Decompose density

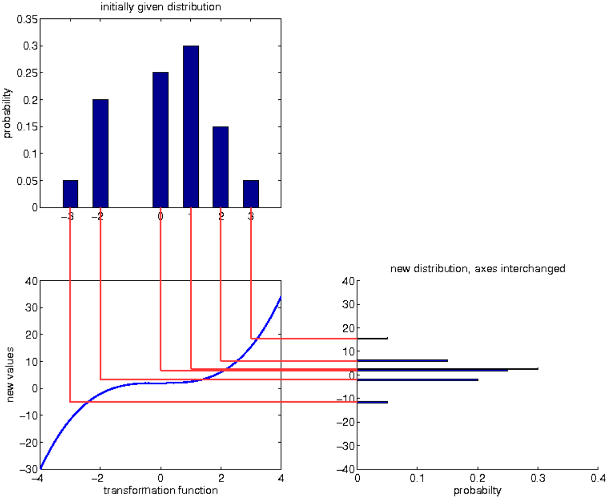


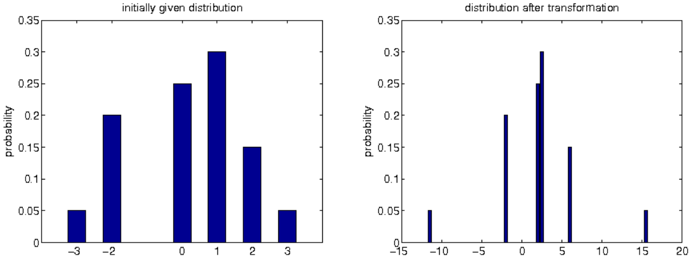
Functions of random variables: example



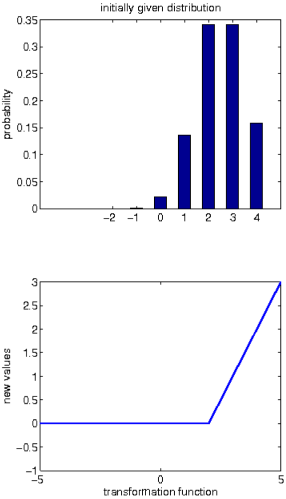

initially given distribution
probability
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−3
−2
0
1
2
3
new values
40
30
20
10
0
−10
−20
−30
−4
−2
0
2
4
transformation function

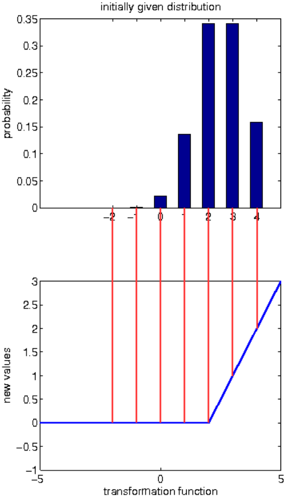

initially given distribution
probability
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−3
−2
0
1
2
3
new values
40
30
20
10
0
−10
−20
−30
−4
−2
0
2
4
transformation function
40
30
20
10
0
−10
−20
−30
−40
0
0.1
0.2
0.3
0.4

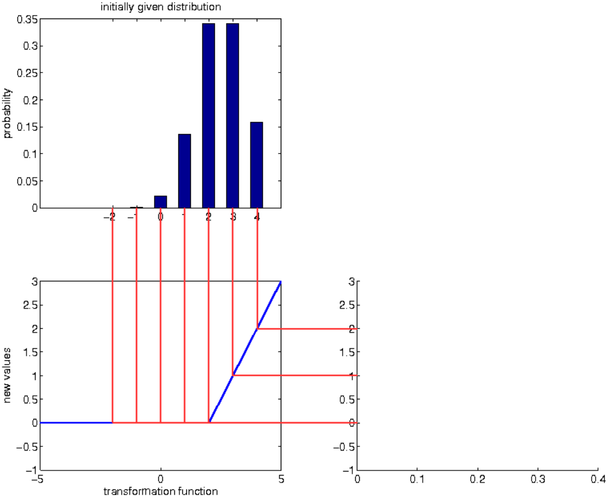

initially given distribution
probability
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−3
−2
0
1
2
3
new values
40
30
20
10
0
−10
−20
−30
−4
−2
0
2
4
transformation function
new distribution, axes interchanged
40
30
20
10
0
−10
−20
−30
−40
0
0.1
0.2
0.3
0.4
probabilty

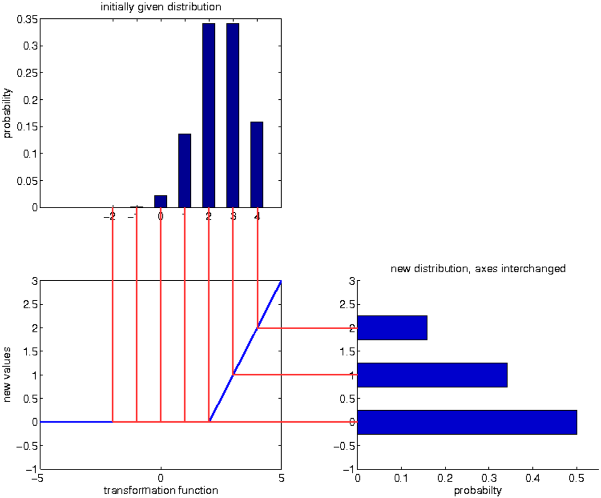

initially given distribution
probability
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−3
−2
0
1
2
3
distribution after transformation
probability
−15
−10
−5
0
5
10
15
20

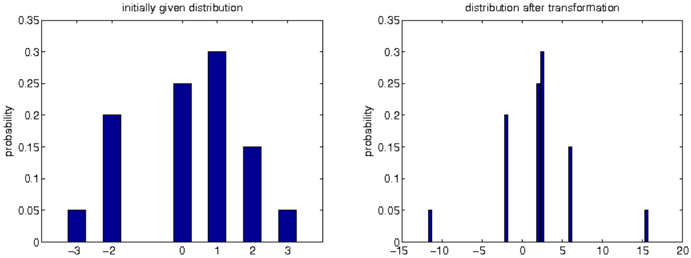
Example: call option




initially given distribution
probability
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−2
−1
0
1
2
3
4
new values
3
2.5
2
1.5
1
0.5
0
−0.5
−1
−5
0
5
transformation function


initially given distribution
probability
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−2
−1
0
1
2
3
4
new values
3
2.5
2
1.5
1
0.5
0
−0.5
−1
−5
0
5
transformation function
0
0.1
0.2
0.3
0.4


initially given distribution
probability
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−2
−1
0
1
2
3
4
new values
3
2.5
2
1.5
1
0.5
0
−0.5
−1
−5
0
5
transformation function
new distribution, axes interchanged
0
0.1
0.2
0.3
0.4
0.5
probabilty


initially given distribution
probability
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−3
−2
0
1
2
3
distribution after transformation
probability
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−15
−10
−5
0
5
10
15
20

Analytical formula

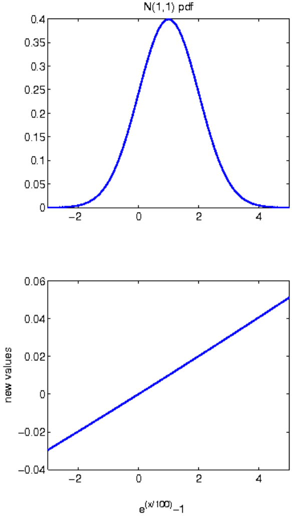
transformation theorem

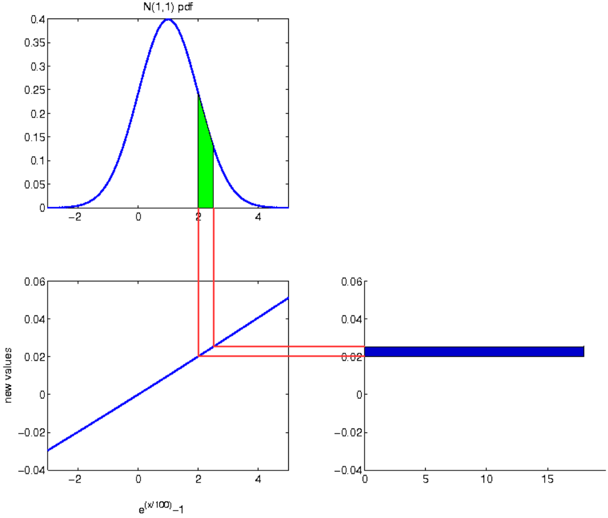
Let X be a random variable with density function $f(x)$, and $g(x)$ be an **invertible bijective** function. Then the density function of the **transformed random variable** $Y = g(X)$ in any point z is given by

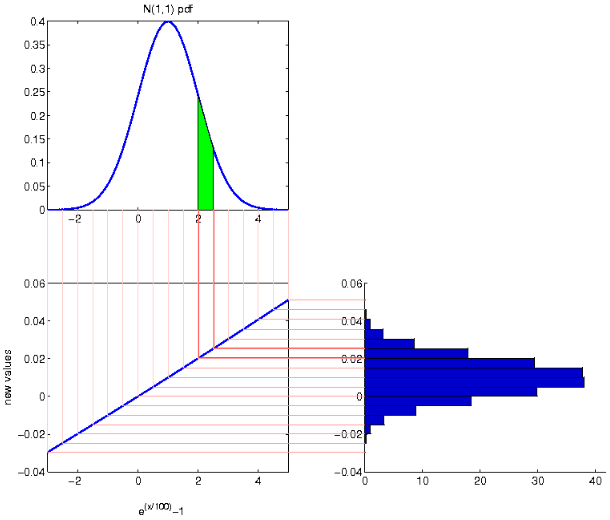
$$f_Y(z) = f_X\left(g^{-1}(z)\right) \cdot \left|\left(g^{-1}\right)'(z)\right|.$$

problems:

- given that we can calculate a measure ϱ_X of the random variable X, it is not ensured that ϱ_Y can be calculated for the new random variable Y, too: e.g. if ϱ envolves integration
- what about non-invertible functions?


N(1,1) pdf
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−2
0
2
4
0.06
0.04
0.02
0
−0.02
−0.04
new values
−2
0
2
4
e^(x/100)−1


N(1,1) pdf
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−2
0
2
4
0.06
0.04
0.02
0
−0.02
−0.04
new values
$e^{(x/100)}-1$
0
5
10
15


N(1,1) pdf
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−2
0
2
4
0.06
0.04
0.02
0
−0.02
−0.04
new values
e^(x/100)−1
0
10
20
30
40

Analytical solution

- Traditional financial modelling assumes **logarithmic returns** to be distributed according to a normal distribution, so that, for example, $100 \cdot r^{log}$ is modelled by $R^{log} := 100 \cdot r^{log} \sim \mathcal{N}(1, 1)$.
- given a **percentage logarithmic return** R^{log}, the **net return** we observe in the real world can be calculated as a function of R^{log} by

$$r = e^{R^{log}/100} - 1$$

- hence, the associated distribution of the **net return** has to be calculated according to the transformation theorem:

$$f_r(z) = f_{R^{log}}\left(g^{-1}(z)\right) \cdot \left|\left(g^{-1}\right)'(z)\right|$$

with transformation function $g(x) = e^{x/100} - 1$

- calculate each part

- calculation of g^{-1} :

$$\begin{aligned} x = e^{y/100} - 1 &\Leftrightarrow \\ x + 1 = e^{y/100} &\Leftrightarrow \\ \log(x+1) = y/100 &\Leftrightarrow \\ 100 \cdot \log(x+1) = y & \end{aligned}$$

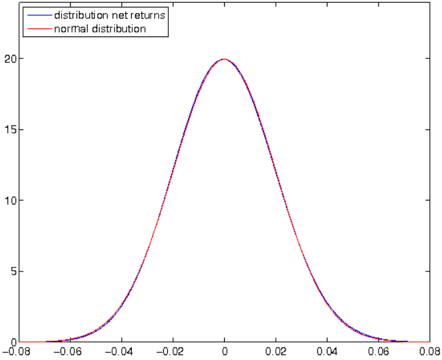
- calculation of the derivative $\left(g^{-1}\right)'$ of the inverse of g^{-1} :

$$(100 \cdot \log(x+1))' = 100 \cdot \frac{1}{x+1}$$

- plugging in leads to:

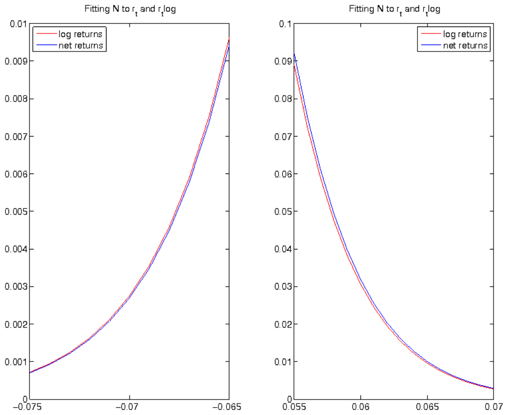
$$f_r(z) = f_{R^{log}}(100 \cdot \log(z+1)) \cdot \left|\frac{100}{z+1}\right|$$

- although only visable under some magnification, there is a difference between a normal distribution which is directly fitted to the net returns and the distribution which arises for the net returns by fitting a normal distribution to the logarithmic returns

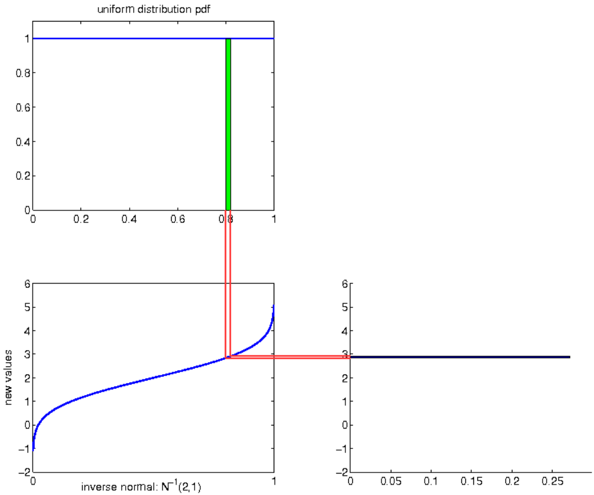


Comparison of tails

- magnification of the tail behavior shows that the resulting distribution from fitting a normal distribution to the logarithmic returns assigns more probability to extreme negative returns as well as less probability to extreme positive returns

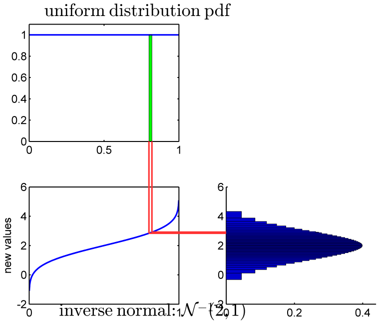


- example: application of an inverse normal cumulative distribution as transformation function to a uniformly distributed random variable



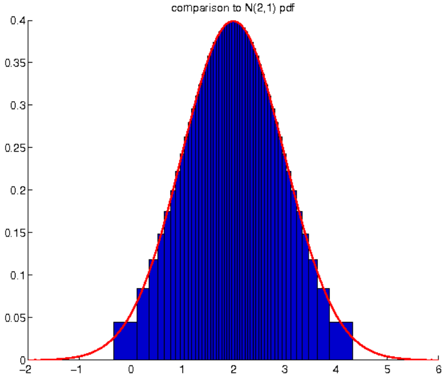
Monte Carlo Simulation

- the resulting density function of the transformed random variable seems to resemble a normal distribution



Monte Carlo Simulation

- a more detailed comparison shows: the resulting approximation has the shape of the normal distribution with the exact same parameters that have been used for the inverse cdf as transformation function



Monte Carlo Simulation

Proposition

Let X be a univariate random variable with distribution function F_X. Let F_X^{-1} be the quantile function of F_X, i.e.

$$F_X^{-1}(p) = \inf\{x | F_X(x) \geq p\}, \quad p \in (0, 1).$$

Then for any standard-uniformly distributed $U \sim \mathbb{U}[0, 1]$ we have $F_X^{-1}(U) \sim F_X$. This gives a simple method for simulating random variables with arbitrary distribution function F.

Proof

Proof.

Let X be a continuous random variable with cumulative distribution function F_X, and let Y denote the transformed random variable $Y := F_X^{-1}(U)$. Then

$$F_Y(x) = \mathbb{P}(Y \leq x) = \mathbb{P}\left(F_X^{-1}(U) \leq x\right) = \mathbb{P}(U \leq F_X(x)) = F_X(x)$$

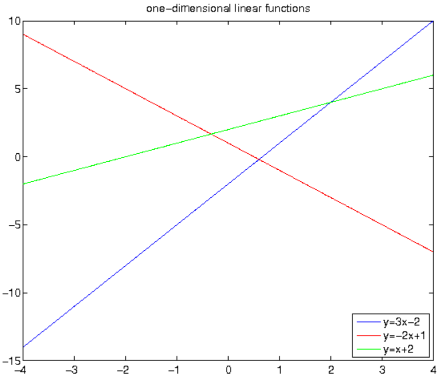
so that Y has the same distribution function as X. □

Linear transformation functions

- a one-dimensional **linear** transformation function is given by

$$g(x) = ax + b$$

- examples of linear functions:



Effect on measures

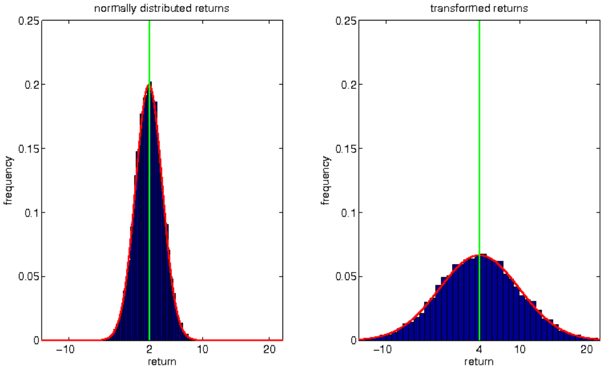
- determine effects of **linear transformation** on measures derived from the distribution function
- **example:** given $X \sim \mathcal{N}(2,4)$, calculate mean and variance of $Y := g(X) = 3X - 2$ via **Monte Carlo Simulation**
  - simulate 10,000 uniformly distributed random numbers
  - transform uniformly distributed numbers via inverse of $\mathcal{N}(2,4)$ into $\mathcal{N}(2,4)$-distributed random numbers
  - apply linear function $g(x) = 3x - 2$ on each number
  - calculate sample mean and sample variance

Matlab code

```
U = rand(10000,1);
returns = norminv(U,2,2);
transformedReturns = 3*returns-2;
sampleMean = mean(transformedReturns);
sampleVariance = var(transformedReturns);
```

Solution

$$\hat{\mu} = 4.0253, \quad \hat{\sigma}^2 = 36.1843 \Leftrightarrow \hat{\sigma} = 6.0153$$



Analytical solution: general case

- calculate inverse g^{-1} :

$$x = ay + b \Leftrightarrow x - b = ay \Leftrightarrow \frac{x}{a} - \frac{b}{a} = y$$

- calculate derivative $\left(g^{-1}\right)'$:

$$\left(\frac{x}{a} - \frac{b}{a}\right)' = \frac{1}{a}$$

- putting together gives:

$$f_{g(X)}(z) = f_X\left(g^{-1}(z)\right) \cdot \left|\left(g^{-1}\right)'\right| = f_X\left(\frac{z}{a} - \frac{b}{a}\right) \cdot \left|\frac{1}{a}\right|$$

- interpretation: **stretching** by factor a, **shifting** b units to the right

Effect on expectation

- **stretching** and **shifting** the distribution also directly translates into the formula for the expectation of a **linearly transformed random variable** $Y := aX + b$:

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

- **possible application**: given expectation $\mathbb{E}[X]$ of stock return, find expected wealth when investing initial wealth W_0 and subtracting the fixed transaction costs c
- hence, focus on linearly transformed random variable

$$\mathbb{E}[Y] = \mathbb{E}[W_0 \cdot X - c],$$

calculated by

$$W_0\mathbb{E}[X] - c$$

Effect on variance

- using the formula for the expectation, the effect of a linear transformation on the variance

$$\mathbb{V}[Y] = \mathbb{E}\left[(Y - \mathbb{E}[Y])^2\right]$$

of the random variable can be calculated by

$$\begin{aligned}
\mathbb{V}[aX + b] &= \mathbb{E}\left[(aX + b - \mathbb{E}[aX + b])^2\right] \\
&= \mathbb{E}\left[(aX + b - a\mathbb{E}[X] - b)^2\right] \\
&= \mathbb{E}\left[(a(X - \mathbb{E}[X]) + b - b)^2\right] \\
&= a^2\mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \\
&= a^2\mathbb{V}[X]
\end{aligned}$$

- note: calculation of mean and variance of a linearly transformed variable neither requires detailed information about the distribution of the original random variable, nor about the distribution of the transformed random variable
- knowledge of the respective values of the original distribution is sufficient
- the analytically computated values for expectation and variance of the example amount to

$$\mathbb{E}[3X-2] = 3\mathbb{E}[X] - 2 = 3 \cdot 2 - 2 = 4$$

$$\mathbb{V}[3X-2] = 3^2\mathbb{V}[X] = 9 \cdot \sigma_X^2 = 9 \cdot 2^2 = 36$$

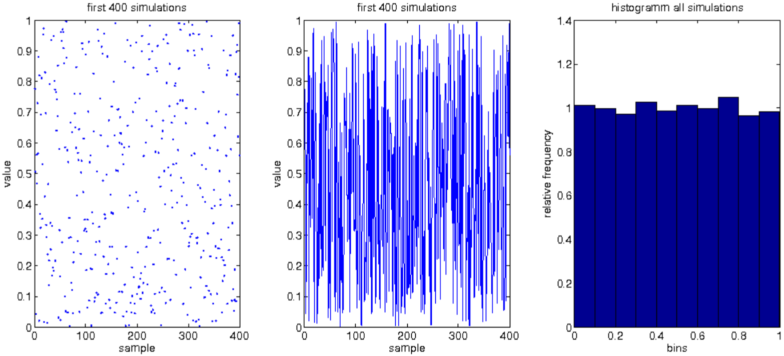
- for **non-linear** transformations, such simple formulas do **not exist**
- most situations require simulation of the transformed random variable and subsequent calculation of the sample value of a given measure

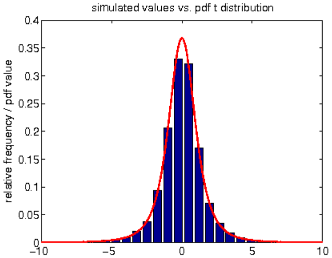
Summary / outlook

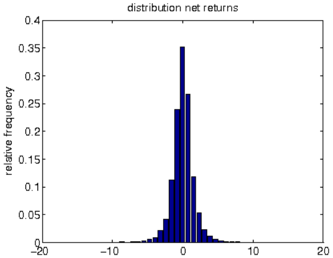
- given random variable X of **arbitrary distribution** F_X, associated values $\mathbb{E}[X]$ and $\mathbb{V}(X)$, and a **linear** transformation $Y = f(X)$, we can also get $\mathbb{E}[Y]$ and $\mathbb{V}(X)$ very simple
- **modelling practices**: taking hypothetical considerations as given, **continuous returns** are modelled as **normally distributed**
- consequences:
  - $\mathbb{E}[X]$ and $\mathbb{V}(X)$ are easily obtainable
  - since discrete real world returns are non-linear transformation of log-returns, $\mathbb{E}$ and $\mathbb{V}$ are not trivially obtained here

```
U = rand(10000,1); % generate uniformly distributed RV
t = tinv(U,3);    % transform to t-distributed values

% transform to net returns
netRets = (exp(t/100)-1)*100;

% transform net returns via butterfly option payoff
  function:
payoff = subplus(netRets+2)-2*subplus(netRets)+subplus(
  netRets-2);

% calculate 95 percent quantile:
value = quantile(payoff,0.95)
```

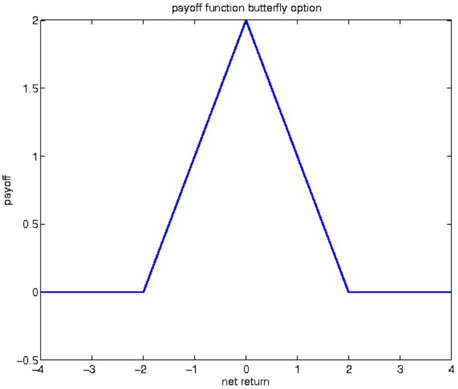
Example


first 400 simulations
value
sample
first 400 simulations
value
sample
histogramm all simulations
relative frequency
bins


simulated values vs. pdf t distribution
relative frequency / pdf value
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−10
−5
0
5
10


distribution net returns
relative frequency
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−20
−10
0
10
20

- payoff profile butterfly option



- expected payoff approximated via Monte Carlo simulation: 1.9305