Risk management

Introduction to the modeling of assets

Christian Groll

Interest rates and returns

General problem

Quantity of interest

$$\mathbf{Z} = \mathbf{g}(\mathbf{X}), \quad \mathbf{X} = (\mathbf{X_1}, \ldots, \mathbf{X_d})$$

- X_i are random variables
- X_i represent *uncertain risk factors*

Examples

portfolio return

- individual stocks $(X_1, \ldots, X_d)$
- g is aggregation function

option payoff

- single underlying X_1
- g is payoff function

Difference to regression setting

- X_i part of the model:
  - in regression analysis, all X_i are taken as given
  - here we need to specify a distribution for $(X_1, \ldots, X_d)$

Justification

- in regression analysis, explanatory variables with influence on first moment are observable upfront
- for financial variables, explanatory variables $(X_1, \ldots, X_d)$ sometimes only become observable simultaneously to Z
- many financial variables tend to exhibit almost constant mean over time: how they are distributed around their mean is important

Certain future payments

- in the simplest case, all risk factors $(X_1, \ldots, X_d)$ are perfectly known

Example

- *bank account* with given interest rate

Aggregation

- even without uncertainty, our quantity of interest commonly implies a *multi-dimensional* setting

Example

- *multi-period wealth* calculation with given annual interest rates

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

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

- multi-period compound interest:

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

Non-constant interest rates

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

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

Logarithmic interest rates

- logarithmic interest rates or *continuously compounded* interest rates are given by

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

Aggregation

- with logarithmic interest rates aggregation becomes a *sum* rather than a *product* of sub-period interest rates:

$$\begin{aligned} r_{1:t}^{log} &= \ln\left(1 + r_{1:t}\right) \\ &= \ln\left(\prod_{i=1}^{t}\left(1 + r_i\right)\right) \\ &= \ln\left(1 + r_1\right) + \ln\left(1 + r_2\right) + \ldots + \ln\left(1 + r_t\right) \\ &= r_1^{log} + r_2^{log} + \ldots + r_t^{log} \\ &= \sum_{i=1}^{t} r_i^{log} \end{aligned}$$

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:

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

Effective annual rate

- the *effective annual rate* R^{eff} is defined as the wealth after one year, given an initial wealth $W_0 = 1$

- with biannual compounding, we get

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

- it exceeds the simple annual rate:

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

m interest payments within a year

- *effective annual rate* after one year:

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

- for wealth after T years we get:

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

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

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

- compounding over T periods leads to

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

- under continuous 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*

T	$m = 1$	$m = 2$	$m = 3$	∞
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
5	1159.3	1160.5	1161.2	1161.8
...		...		...
9	1304.8	1307.3	1308.6	1310
10	1343.9	1346.9	1348.3	1349.9

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

Effective logarithmic rates

For logarithmic interest rates, a higher compounding frequency leads to

$$\begin{aligned} r^{log;eff} &= \ln(R^{eff}) \\ &= \ln(W_{m,1}) \\ &= \ln\left(\left(1+\frac{r}{m}\right)^m\right) \\ &\overset{m\to\infty}{\to} \ln(\exp(r)) \\ &= r \end{aligned}$$

Interpretation: If the bank were compounding interest rates continuously, the nominal interest rate r would equal the logarithmic effective rate.

Also:

- if $r^{log;eff} = r$ for continuous compounding,
- and continuous compounding leads to almost identical end of period wealth as simple compounding (see table above)
- the logarithmic transformation $r^{log} = \ln(1 + r)$ does change the value only marginally: $r^{log} \approx r$

Conclusion

In other words:

- we can interpret log-interest rates as roughly equal to simple rates
- still, log-interest rates are better to work with, as they increase linearly through aggregation over time

Conclusion

But: if interest rates get bigger, the approximation of simple compounding by continuous compounding gets worse!

- $\ln(1+x) = x$ for $x = 0$
- $\ln(1+x) \approx x$ for $x \neq 0$

Prices and returns

Returns on speculative assets

- 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
- returns on speculative assets usually vary from period to period

- 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*

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 :

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

- 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

Using continuous returns

- for comparison, using continuous returns

Continuous case

- *continuously compounded returns* associated with an evolution of prices over a longer time period is given by

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

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

Continuous case

- plugging in leads to

$$r_A = \frac{(\ln 130 - \ln 100)}{3} = 0.0875$$

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

- 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

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

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

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

Example

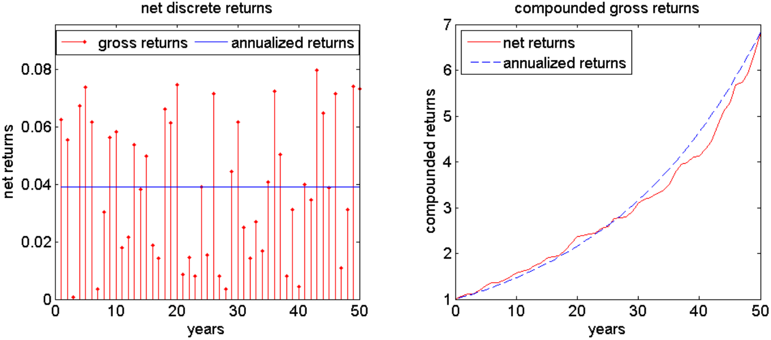

Figure 1

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

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

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*:

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

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

Example

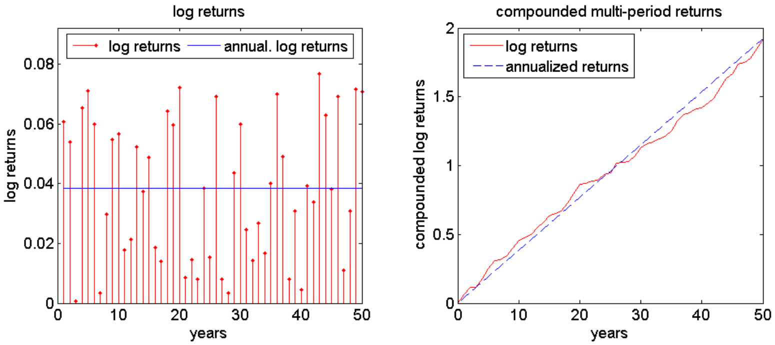

Figure 2

- convert result back to *normal world*:

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

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:

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

- 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 continuous returns overall?

- however: the *main advantage* only crops up in a setting of uncertain future returns, and their modelling as random variables!

Importance of returns

Why are *asset returns* so pervasive if *asset prices* are the real quantity of interest in many cases?

Non-stationarity

Most prices are not stationary:

- over long horizons stocks tend to exhibit a positive trend
- distribution changes over time

Consequence

- historic prices are not representative for future prices: mean past prices are a bad forecast for future prices

Returns

- returns are stationary in most cases

$\Rightarrow$ historic data can be used to estimate their current distribution

General problem

Quantity of interest

$$\mathbf{Z} = \mathbf{g}(\mathbf{X}), \quad \mathbf{X} = (\mathbf{X_1}, \ldots, \mathbf{X_d})$$

- as statistical requirements tend to force us to use returns instead of prices, almost always at least some X_i represent returns

Time horizon and aggregation

- lower frequency returns can be expressed as aggregation of higher frequency returns
- lack of data for lower frequency returns (as they need to be non-overlapping)

$\Rightarrow$ long horizons usually require aggregation of higher frequency returns: $X_t, X_{t+1}, \ldots$

Outlook: mathematical tractability

Only with log-returns we preserve a chance to end up with a linear function:

Quantity of interest

$$\begin{aligned}\mathbf{Z} &= g(\mathbf{X}) \\ &= g(Y_t, Y_{t+1}, \ldots, X_i) \\ &= \hat{g}(Y_t + Y_{t+1} + \ldots, X_i)\end{aligned}$$

Outlook: statistical fitting

The *central limit theorem* could justify modelling *logarithmic* returns as *normally distributed*:

- returns can be decomposed into *summation* over returns of *higher* frequency: e.g. annual returns are the sum of 12 monthly returns, 52 weakly returns, 365 daily returns,. . .

The central limit theorem states:

Independent of the distribution of high frequency returns, any sum of them follows a *normal distribution*, provided that the sum involves sufficiently many summands, and the following requirements are fulfilled:

- the high frequency returns are *independent* of each other
- the distribution of the low frequency returns allows finite second moments (variance)

- 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!

Probability theory

- *randomness*: the result is not known in advance
- *probability theory*: captures randomness in mathematical framework

Probability spaces and random variables

- *sample space* Ω: set of all possible outcomes or elementary events ω

Examples: *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: *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":

$$A =]10, \infty[\subset \Omega_4$$

Event space

- 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

- $\{\}$ denotes the *empty set*

Event space example

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

Events

- 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

A real-valued set function $\mathbb{P} : \mathcal{F} \rightarrow \mathbb{R}$, with properties

- $\mathbb{P}(A) > 0$ for all $A \subseteq \Omega$
- $\mathbb{P}(\Omega) = 1$
- For each finite or countably infinite collection of *disjoint* events (A_i) it holds:

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

$\Rightarrow$ quantifies for each event a probability of occurance

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*.

Example

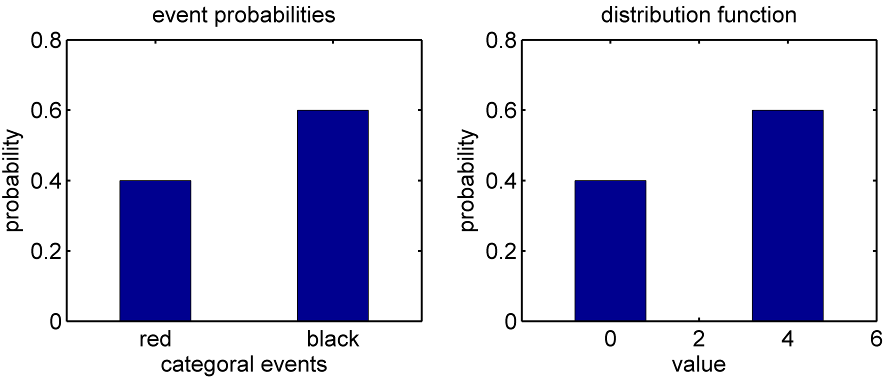

Figure 3:random variable with discrete values

Density function

- 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
- 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

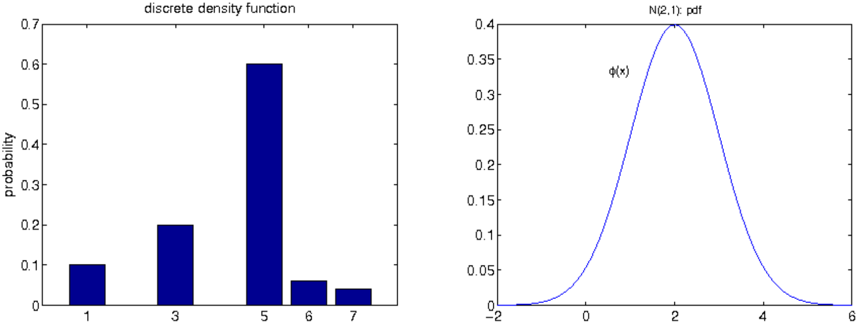

Figure 4

Cumulative distribution function The *cumulative distribution function* (cdf) of random variable X, denoted by $F(x)$,

indicates the probability that X takes on 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:

- $F(x)$ is a nondecreasing function of x;
- $\lim_{x\to\infty} F(x) = 1$;
- $\lim_{x\to-\infty} F(x) = 0$.
- furthermore:

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

Interrelation pdf and cdf: **discrete case**

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

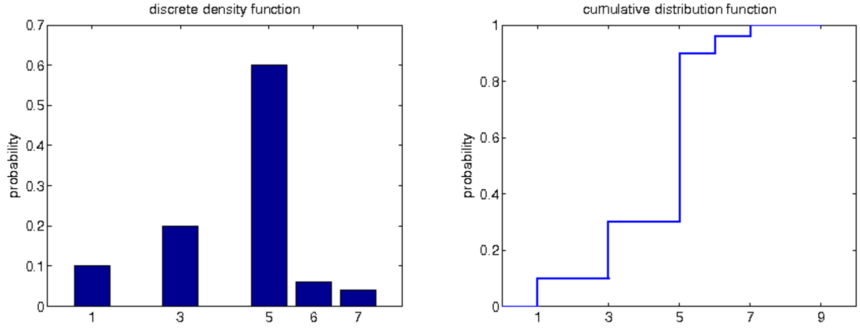

Figure 5

Interrelation pdf and cdf: **continuous case**

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

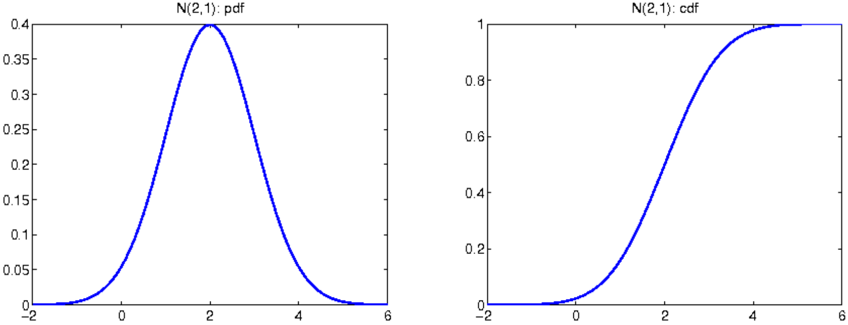

Figure 6

Information reduction

Modeling 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 $\Omega = \{1, 3, 5, 6, 7\}$
- modeling 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

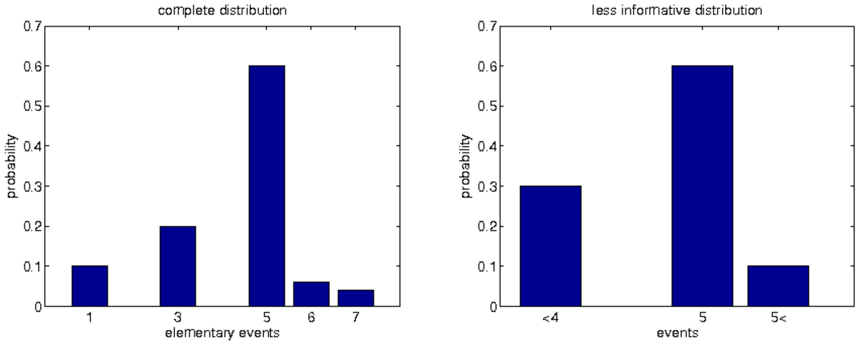

Figure 7

Information reduction discrete

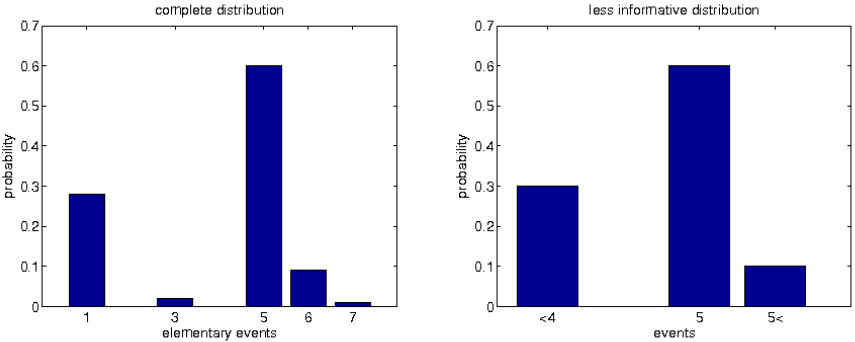

Figure 8

Information reduction continuous

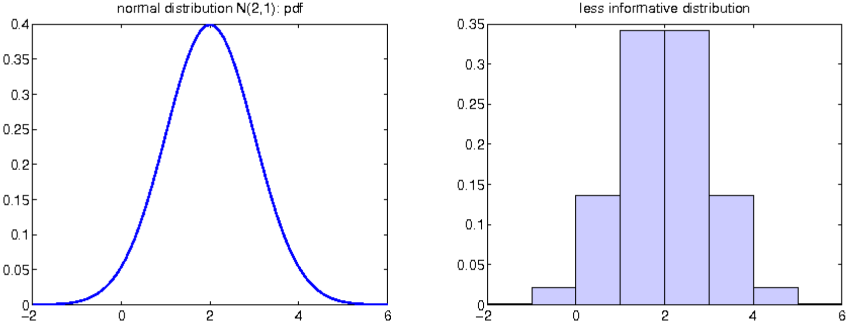

Figure 9

Measures of random variables

- complete distribution may not always be necessary
- 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

Classification with respect to several measures can be sufficient:

- probability of negative / positive return
- return on average
- worst case
- 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 $f(x)$:

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

Examples

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

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

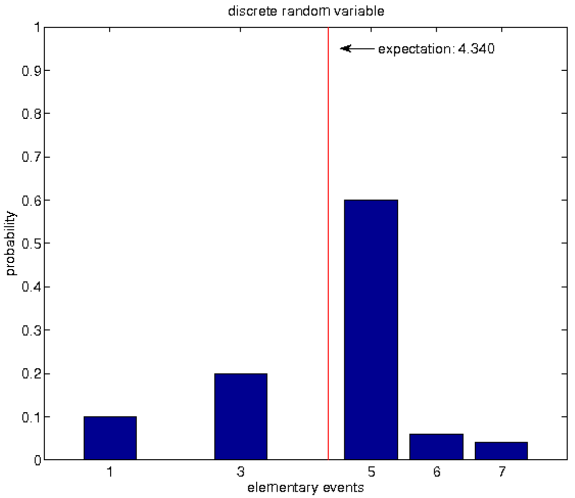
discrete random variable
expectation: 4.340
probability
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1
3
5
6
7
elementary events

Figure 10

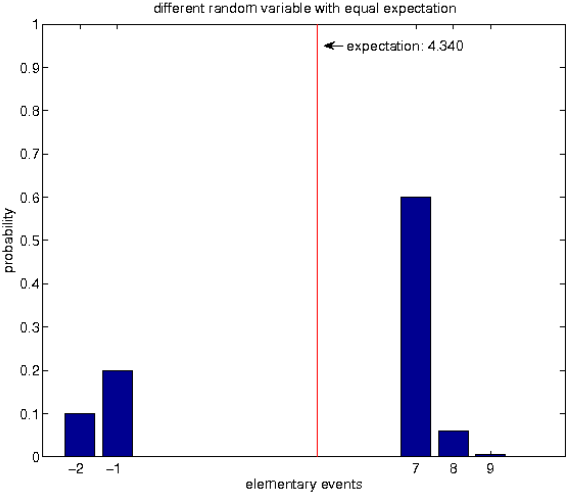
different random variable with equal expectation
expectation: 4.340
probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−2
−1
7
8
9
elementary events

Figure 11

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

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

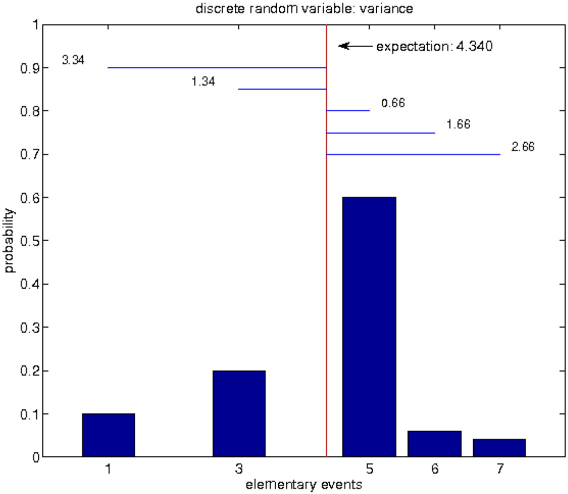
discrete random variable: variance
expectation: 4.340
3.34
1.34
0.66
1.66
2.66
probability
elementary events
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1
3
5
6
7

Figure 12

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

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

Quantile

- *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

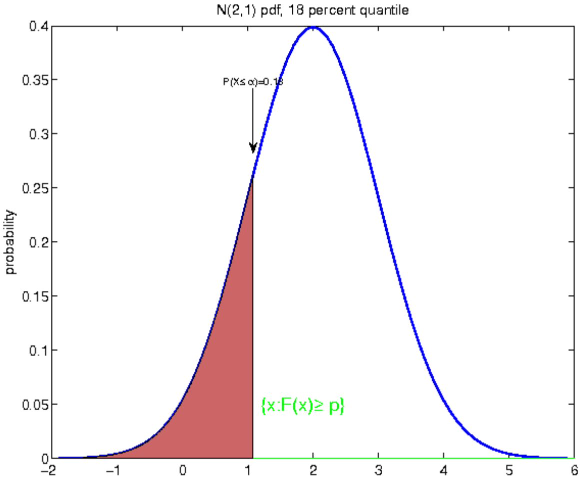

Figure 13

Example: cdf

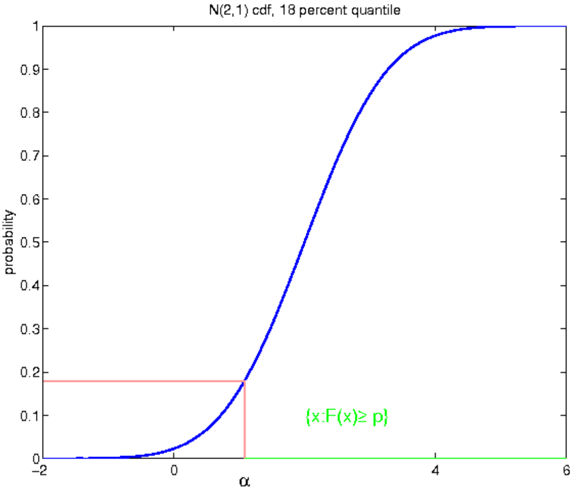

Figure 14

Example

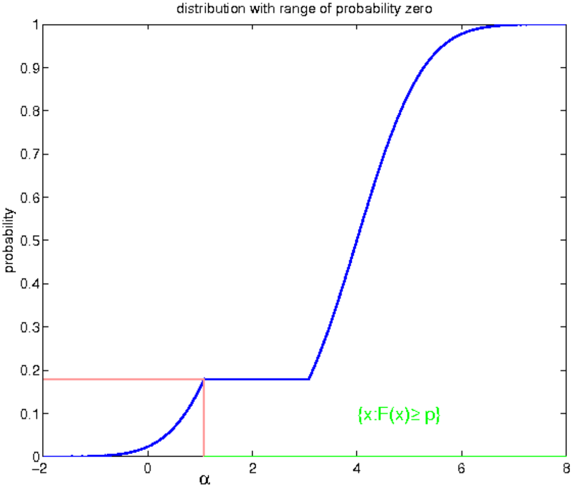

Figure 15

Example

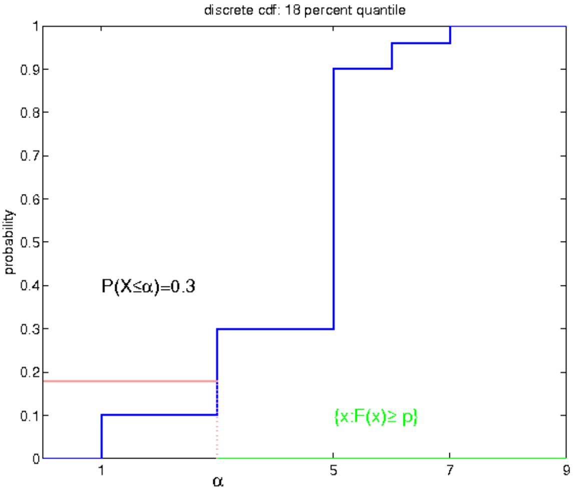

Figure 16

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

General problem

Quantity of interest

$$\varrho(\mathbf{Z}) = \mathbf{g}(\mathbf{X}), \quad \mathbf{X} = (\mathbf{X_1}, \ldots, \mathbf{X_d})$$

- instead of the complete distribution of Z, interest only lies in some measure ϱ (expectation, variance, ...)

Updating information

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

Unconditional density

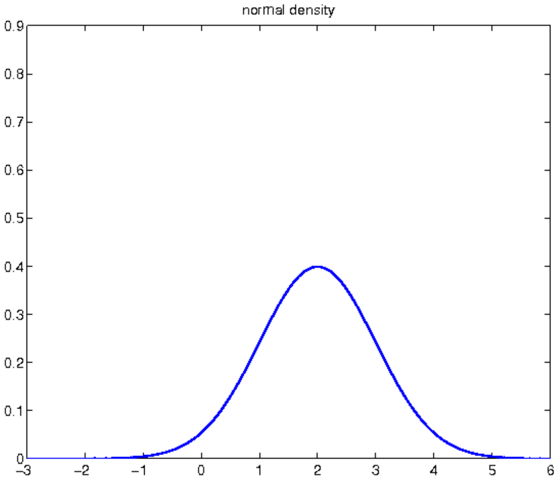

Figure 17

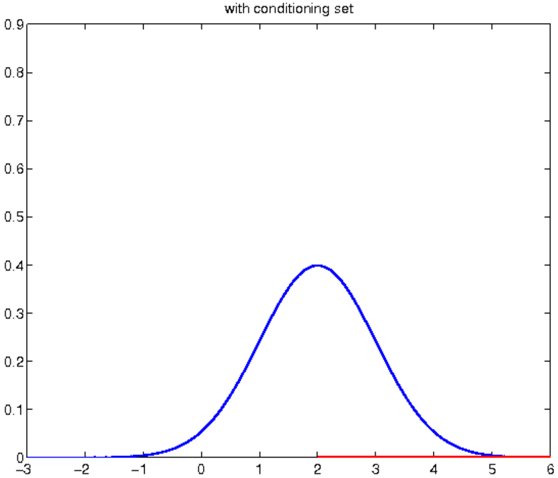

Figure 18

Given the knowledge of a realization higher than 2, probabilities of values below become zero:

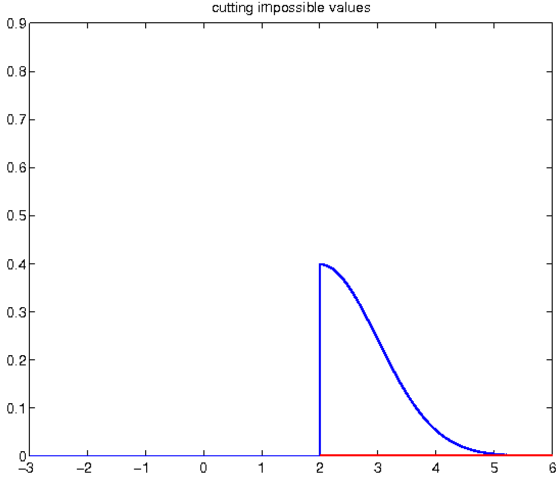

Figure 19

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

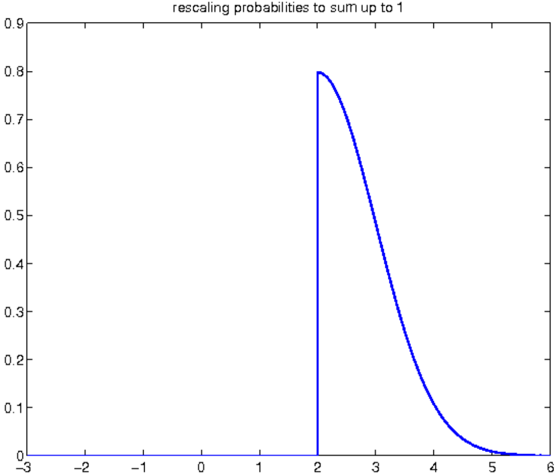

Figure 20

- original density function compared to updated conditional density

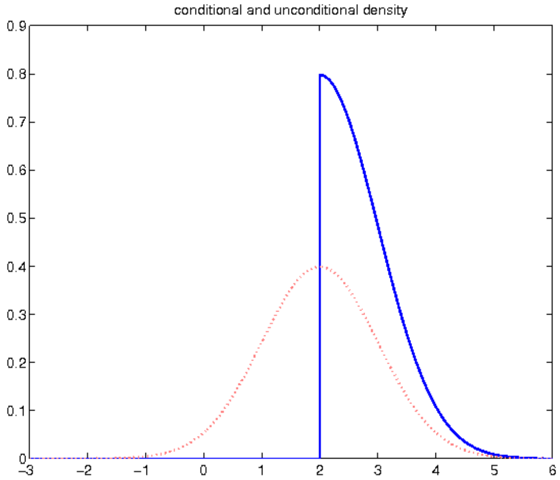

Figure 21

Decomposing density

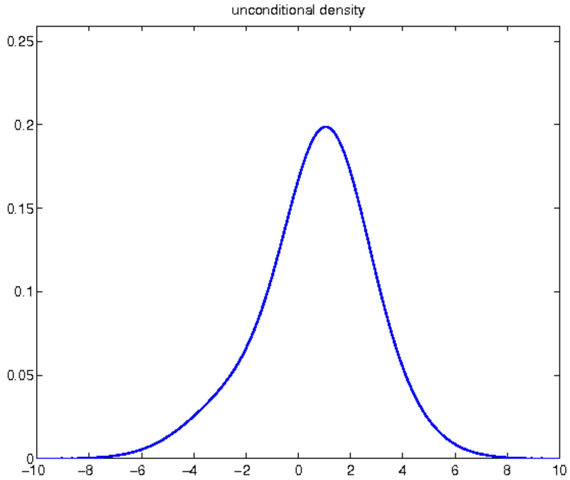

Figure 22

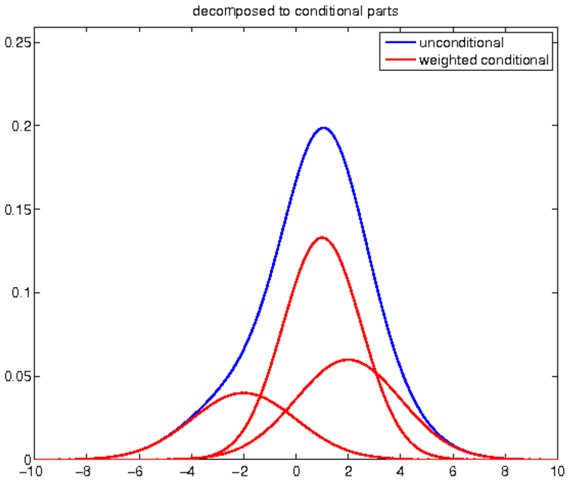
decomposed to conditional parts
unconditional
weighted conditional
0.25
0.2
0.15
0.1
0.05
0
−10
−8
−6
−4
−2
0
2
4
6
8
10

Figure 23

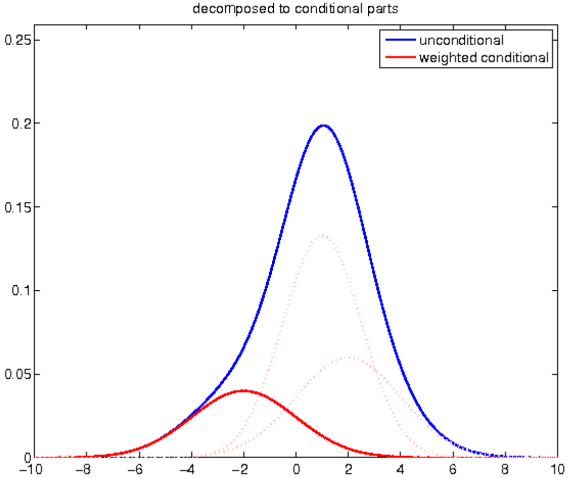
decomposed to conditional parts
unconditional
weighted conditional
0.25
0.2
0.15
0.1
0.05
0
−10
−8
−6
−4
−2
0
2
4
6
8
10

Figure 24

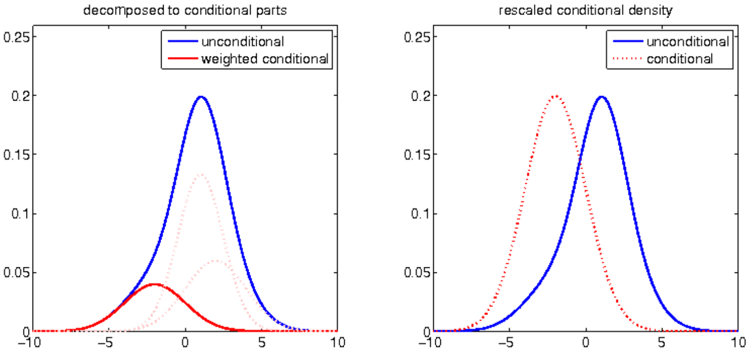
decomposed to conditional parts
unconditional
weighted conditional
0.25
0.2
0.15
0.1
0.05
0
−10
−5
0
5
10
rescaled conditional density
unconditional
conditional
0.25
0.2
0.15
0.1
0.05
0
−10
−5
0
5
10

Figure 25