Modul 00: Grundlagen

PI 6250 – Ökonometrie I

Max Heinze (mheinze@wu.ac.at)
Department für Volkswirtschaftslehre, WU Wien

6. März 2025

Einführung

Installation von R

Einführung in R mit RStudio

Summen

Einführung

Willkommen im Kurs **Ökonometrie I**!

In diesem Kurs beschäftigen wir uns, einfach gesagt, damit, wie wir Daten nutzen können, um Beweise für Hypothesen und Antworten auf Fragestellungen zu finden, die wir uns stellen.

Dafür benötigen wir solide **mathematische** und **statistische Grundkenntnisse**. Im Großen und Ganzen sind das Dinge, die schon im Maturastoff vorkommen und in den Statistik- und Mathematiklehrveranstaltungen im CBK wiederholt werden.

Dieser **Foliensatz zum Selbststudium** soll euch helfen, diese Grundlagen aufzufrischen. Wenn etwas nicht klar ist, besteht in der LV natürlich noch genug Zeit zum Nachfragen, aber ihr solltet über diesen Foliensatz grundlegend Bescheid wissen.

Die **englischsprachigen Übersetzungen** der wichtigsten Begriffe in diesem Foliensatz sollen euch dabei helfen, englischsprachige Lehrbücher (also die Kursliteratur) besser verstehen zu können und euch darauf vorbereiten, falls ihr einen englischsprachigen Kurs in Ökonometrie 2 oder 3 besuchen wollt.

Einführung

Installation von R

Einführung in R mit RStudio

Summen

Matrizen und Vektoren

Installation von RStudio

An verschiedenen Stellen in diesem Kurs, nicht zuletzt bei **Hausübungen**, benötigen wir ein Programm, mit dem wir statistische Berechnungen anstellen können. Welches Programm ihr dafür verwendet, ist euch freigestellt. Meine **Empfehlung** ist **R**. Auch alle Codebeispiele in diesem Kurs werden in R angeboten.

Eine komfotable Art und Weise, R zu verwenden, ist mit der *integrierten Entwicklungsumgebung* **RStudio**. RStudio ist dabei die Oberfläche, die wir verwenden, um Code in R zu schreiben; R selber ist ein separates Programm, das unseren Code ausführt und Ergebnisse liefert.

> **Installation von R und RStudio**
>
> Eine Installationsanleitung und der **Download für R** findet sich unter cran.r-project.org.
>
> Eine Installationsanleitung und der **Download von RStudio** findet sich unter posit.co/download/rstudio-desktop/

Installation von einer Statistiksoftware (z.B. R) wird im Kurs vorausgesetzt.

Einführung

Installation von R

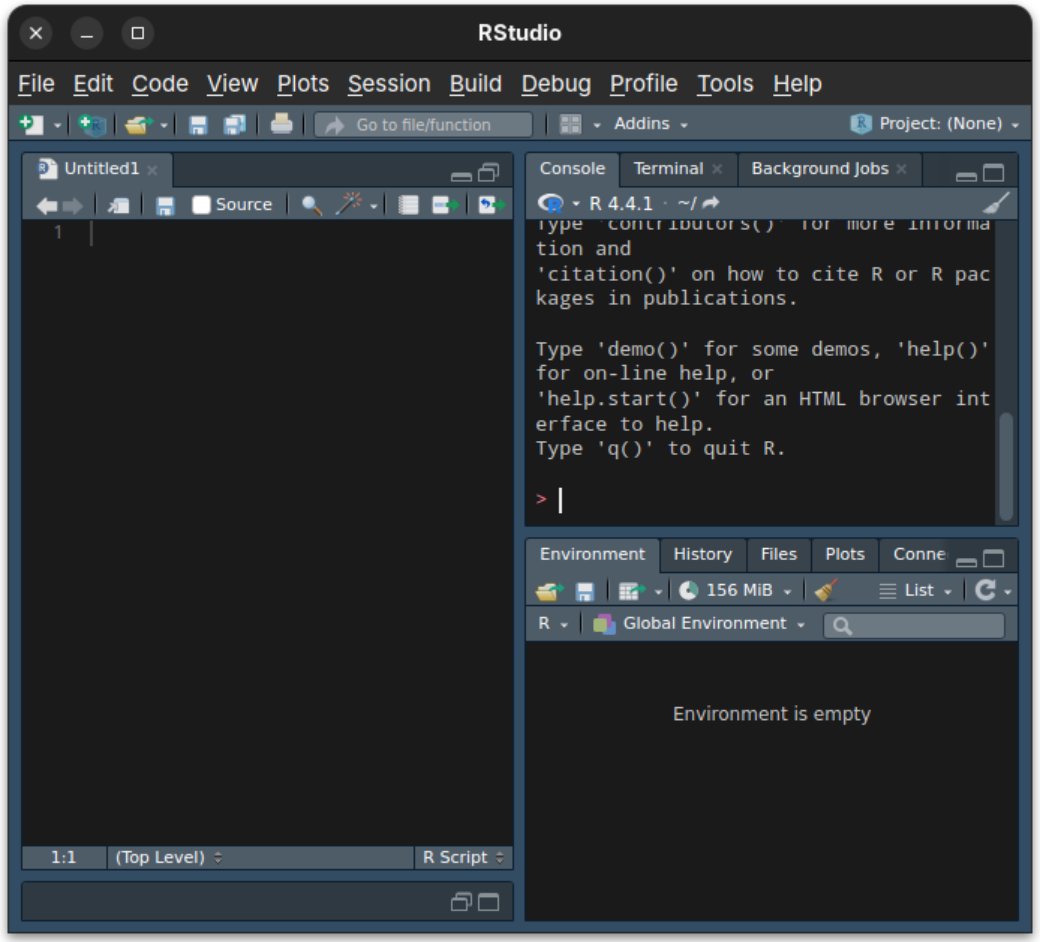
Einführung in R mit RStudio

Summen

Matrizen und Vektoren

Zufallsvariablen

RStudio



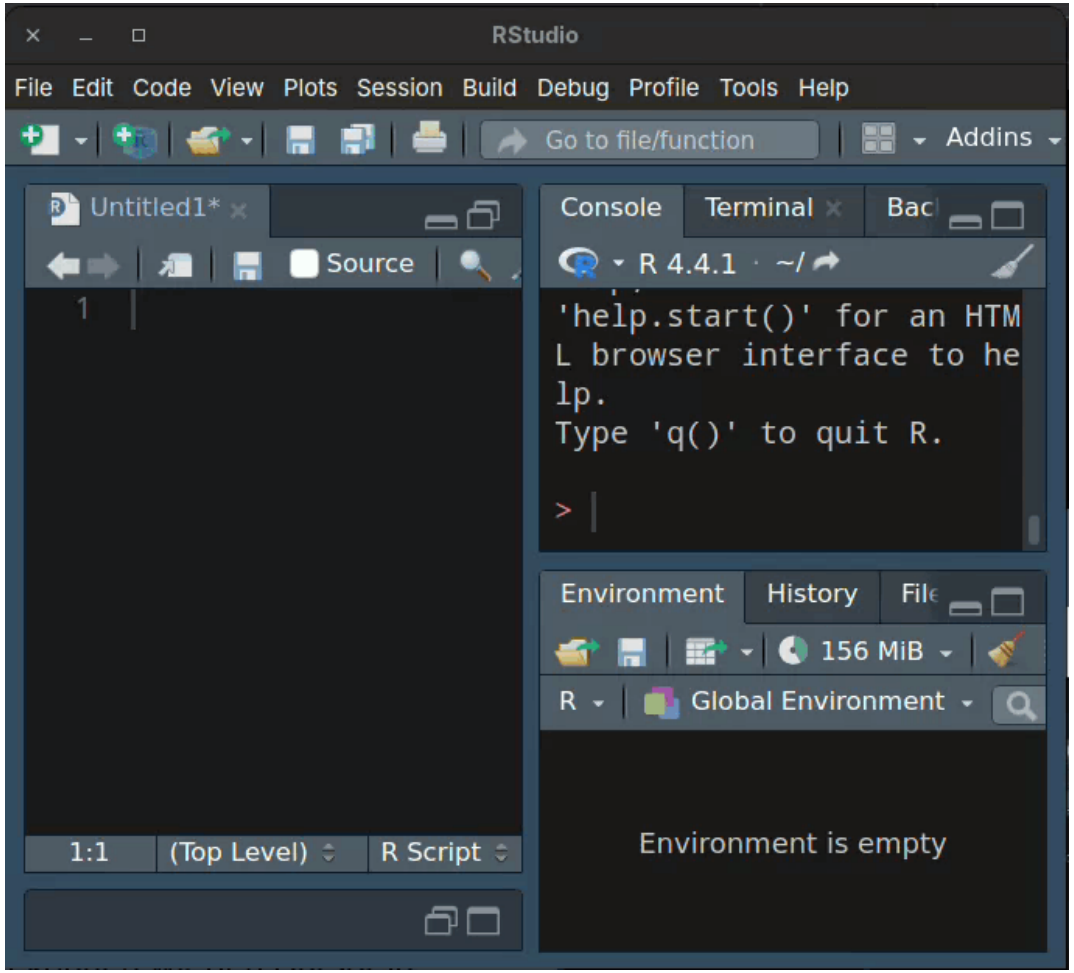
Die Standard-Anordnung ist etwas anders und das Standard-Thema ist hell. Beides kann in den Einstellungen geändert werden.

Etwa so sieht ein **RStudio**-Fenster aus:

- In der **Console** können wir **R**-Befehle ausführen. Beispielsweise können wir `1+1` eintippen und mit `Enter` bestätigen.
- Im **Environment** finden wir alle aktuell definierten Variablen und Datensätze.
- Unter **Plots** werden produzierte Grafiken angezeigt.
- Unter **Help** können wir die Dokumentation aufrufen.
- Im großen Feld links können wir ein **Skript** öffnen.

Skript vs. Console

- Mit einem **Skript** können wir verschiedene Zeilen Code **speichern** und immer dann ausführen, wenn wir sie benötigen.
- Wir können Code auch direkt aus der **Console** ausführen, dann können wir ihn aber nicht speichern.
- Im Skript können wir den Cursor in eine Zeile setzen und diese Zeile mit `Strg` + `Enter` (Linux, Windows) bzw. `cmd` + `Enter` (macOS) **ausführen**. Er wird dann sozusagen in die Console „eingefügt“.
- Alternativ können wir mehrere Zeilen markieren und ausführen, oder das ganze Skript auf einmal ausführen.



Grundlagen in R

Base R kommt mit vielen nützlichen **Funktionen**, manchmal werden wir für spezielle ökonometrische Zwecke aber Funktionen brauchen, die nicht in Base R enthalten sind. Oftmals sind diese in Paketen enthalten, die Entwickler:innen erstellt und dann (im Idealfall) im Comprehensive R Archive Network (CRAN) veröffentlicht haben. Diese zusätzlichen Pakete können wir wie folgt installieren:

```
1 install.packages("tidyverse")
2
3 library(tidyverse)
4
5 update.packages()
```

Die obigen Funktionen **installieren**, **laden**, und **aktualisieren** Pakete (Beachte: `install.packages()` braucht Anführungszeichen). Wir können die Dokumentation einer Funktion mit einem `?` aufrufen:

```
1 ?install.packages
```

Funktionen erkennen wir an den Klammern. Diese können Argumente enthalten, müssen aber nicht.

Variablen und Vektoren

Mithilfe von `<-` (oder `=`) können wir einem Variablennamen einen Wert **zuweisen**.

Das kann entweder ein Skalar sein, ein Vektor, oder etwas anderes (mehr später).

Mit `print()` können wir eine Variable ausgeben lassen. Es genügt aber auch, nur den Variablennamen zu schreiben (ohne `print()`).

Der Code rechts ist **interaktiv** und kann modifiziert und dann ausgeführt werden.

R Code | Start Over | Run Code

```r
# Zeilen, die mit # beginnen, sind Kommentare
# und werden ignoriert.

my_scalar <- 12

my_vector <- c(5, 2, 13, 15)

my_character_vector <- c("Wien", "Linz", "Graz")

# Geht auch ohne print()
print(my_vector)

print(my_character_vector)
```

Mathematik

Wir können R auch wie einen Taschenrechner verwenden. Der interaktive Code rechts zeigt verschiedene **mathematische Operationen**.

Wir können auch **Matrizen** definieren und mit ihnen rechnen.

R Code | Start Over | Run Code

```
my_matrix <- cbind(c(1, 2, 3), c(
2 * my_matrix
```

R Code | Start Over | Run Code

```
1 + 3

# Wir können das Ergebnis auch speichern:
my_sum <- 4 + 12

c(5, 2) + c(7, 3)

2^2
sqrt(2) # Wurzel
log(100) # Nat. Logarithmus
log(100, base = 10)
```

Deskriptive Statistik

Natürlich ist R vorwiegend eine **statistische Programmiersprache**. Simulieren wir einmal 100 Würfelwürfe:

R Code | Start Over | Run Code

```r
# Wir ziehen 100 mal
#aus den Werten 1 bis 6
throws <- sample(1:6,
         size = 100,
         replace = TRUE)
```

Führe diesen Code aus, damit wir weitermachen können.

Was ist der **Mittelwert** der Würfe?

R Code | Start Over | Run Code

```r
mean(throws)
```

Wir können auch andere Maßzahlen berechnen:

R Code | Start Over | Run Code

```r
summary(throws)
```

Weitere Funktionen sind `median()`, `min()`, `max()`, `length()`, `var()`, `sd()`, `sum()`, ...

Daten

Wir verwenden je nach Dateiformat eine andere Funktion, um Daten einzulesen: `read.csv()` für CSV, `readRDS()` für RDS, ...

Manche Datensätze sind auch in R bereits verfügbar, was besonders praktisch für Übungszwecke ist. Ein Beispiel ist `mtcars`. Mit `head()` können wir die ersten Zeilen ansehen.

R Code | Start Over | Run Code

```
1  head(mtcars)
```

Wenn wir Daten z.B. als CSV einlesen, müssen wir sie erst einem Namen zuweisen, z.B. durch `my_data <- read.csv("data.csv")`. Übrigens können wir Daten auf ähnliche Weise exportieren: `write.csv(my_data, "my_data.csv")`.

Dataframes

Die Struktur, in der die Daten gespeichert sind, heißt **Dataframe**. Die Zeilen eines Dataframe entsprechen einzelnen Beobachtungen, die Spalten entsprechen Variablen. Mit `View()` können wir den Datensatz in einem separaten Fenster ansehen. Wir können aber auch z.B. die Anzahl der Spalten und Zeilen herausfinden:

R Code | Start Over | Run Code

```
1  ncol(mtcars); nrow(mtcars)
```

Mit eckigen Klammern können wir bestimmte Zeilen und Spalten aufrufen. `mtcars[1,]` ist die erste Zeile von `mtcars`, `mtcars[,1]` ist die erste Spalte. Wir können einzelne Variablen auch mit folgender Notation aufrufen: `mtcars$mpg`.

Logik

Was passiert, wenn wir diesen Code ausführen?

R Code | Start Over | Run Code

```r
# Doppeltes == bedeutet "ist gleich", weil einfaches = für Zuweisungen
# gebraucht wird. Außerdem: <= für ≤, >= für ≥, != für ≠.
1 == 2
```

Wir können `TRUE` und `FALSE` auch verwenden, um Werte zu filtern:

R Code | Start Over | Run Code

```r
test_vector <- c(1, 2, 3)     # Vektor definieren
test_vector > 1               # Prüfen: Was ist >1?
test_vector[test_vector > 1] # Filtern: Nur Elemente >1 (also nur TRUE)
```

Weitere Operationen

Wir können auch mehrere Funktionen verbinden:

R Code | Start Over | Run Code

```
1 log(mean(mtcars$mpg))
2 # Wie muss dieser Code verändert werden, um d
3 # zu berechnen?
```

Oder Grafiken zeichnen:

R Code | Start Over | Run Code

```
1 hist(mtcars$mpg, breaks = 20)  # Histogramm m
2 # Probiere, einen Boxplot mit boxplot() zu er
```

Scatterplots

Oft wollen wir einen **Scatterplot** (ein **Punktwolkendiagramm**) zeichnen, um die Beziehnung zweier Variablen zueinander darzustellen.

Versuche, die Zeile `abline(lm(mpg~hp, data=mtcars), col="red")` hinzuzufügen, um eine rote **Regressionslinie** in den Plot zu zeichnen!

Wie stark sind `mpg` und `hp` korreliert?

R Code | Start Over | Run Code

```
1 cor(mtcars$hp, mtcars$mpg)
```

R Code | Start Over | Run Code

```
1 plot(mtcars$hp, mtcars$mpg)
```

Einführung

Installation von R

Einführung in R mit RStudio

Summen

Matrizen und Vektoren

Zufallsvariablen

Analyse einer Zufallsvariablen

Grundlagen des Summierens

Wenn wir eine Summe bilden, addieren wir verschiedene Dinge zusammen. Diese Dinge nennen wir **Summanden**. Diese Summanden können Zahlen sein, Funktionen, Vektoren, oder Matrizen. Das Bilden von Summen ist einfach und intuitiv, aber bei einer großen Anzahl von Summanden mühsam aufzuschreiben. Beispiel: Wir wollen die Summe aller natürlichen Zahlen von 1 bis 100 bilden:

$$
\begin{aligned}
&1+2+3+4+5+6+7+8+9+10+11+12+13+14+15+\\
&16+17+18+19+20+21+22+23+24+25+26+27+28+\\
&29+30+31+32+33+34+35+36+37+38+39+40+41+\\
&42+43+44+45+46+47+48+49+50+51+52+53+54+\\
&55+56+57+58+59+60+61+62+63+64+65+66+67+\\
&68+69+70+71+72+73+74+75+76+77+78+79+80+\\
&81+82+83+84+85+86+87+88+89+90+91+92+93+\\
&94+95+96+97+98+99+100 \qquad\qquad = 5050.
\end{aligned}
$$

Summenzeichen

Glücklicherweise können wir diese lange Summe auch einfacher aufschreiben:

$$\sum_{i=1}^{100} i = 5050.$$

Das **Summenzeichen** (engl. **summation operator**) wird oft gefürchtet (besonders wenn es in Gleichungen oft vorkommt, sehen die Gleichungen gerne kompliziert aus), ist aber sehr einfach:

- i unterhalb des Summenzeichens ist der **Summationsindex** (engl. **index of summation**).
- Die Zahl 1 in $i = 1$ ist der **Startwert** (engl. **lower bound**) der Summe.
- Die Zahl 100 ist der **Endwert** (engl. **upper bound**).
- Rechts von dem Summenzeichen steht der Ausdruck, den wir summieren, in diesem Fall i.

Summenzeichen

$$\sum_{i=1}^{100} i = 5050.$$

Wir verfahren wie folgt. Wir lassen i der Reihe nach einmal jeden Wert von Start- bis Endwert annehmen. Dann wird der Ausdruck rechts des Summenzeichens, i, zuerst 1, dann 2, ... und schließlich 100. Wir summieren dann alle diese Ausdrücke.

Der Summationsindex kann auch als Index einer Variable vorkommen:

$$\sum_{i=1}^{4} x_i = x_1 + x_2 + x_3 + x_4.$$

Summenzeichen

Wenn ein Term auf der rechten Seite des Summenzeichens den Summenindex nicht enthält, bleibt er in jedem Durchgang unverändert:

$$\sum_{i=1}^{100} c = 100c.$$

Für jede Konstante c gilt außerdem:

$$\sum_{i=1}^{n} cx_i = c \sum_{i=1}^{n} x_i$$

Wenn Start- und Endwert gleich sind:

$$\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i + y_i$$

Installation von R

Einführung in R mit RStudio

Summen

Matrizen und Vektoren

Zufallsvariablen

Analyse einer Zufallsvariablen

Analyse zweier Zufallsvariablen

Matrizen und Vektoren

Wir notieren Matrizen mit fettgedruckten Großbuchstaben ($\boldsymbol{X}$) und (Spalten)vektoren mit fettgedruckten Kleinbuchstaben ($\boldsymbol{x}$):

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & \cdots & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Der Fettdruck ist nicht unbedingt notwendig. Wenn wir handschriftlich trotzdem klarstellen wollen, dass es sich um eine Matrix oder einen Vektor handelt, können wir sie oder ihn stattdessen unterstreichen: $\underline{X}$ oder $\underline{x}$.

Dimensionen und Transposition

In der vorangegangenen Folie hatte $\boldsymbol{X}$ die **Dimensionen** $n \times k$ und $\boldsymbol{x}$ hatte die Dimension n. $\boldsymbol{x}$ war als Spaltenvektor angeschrieben, wir können $\boldsymbol{x}$ aber auch als Zeilenvektor anschreiben. Dafür müssen wir den Vektor **transponieren**: Vereinfacht gesagt werden aus Zeilen Spalten und aus Spalten Zeilen. Wir notieren die Transposition mit einem kleinen Strich (man kann aber auch ein hochgestelltes T verwenden):

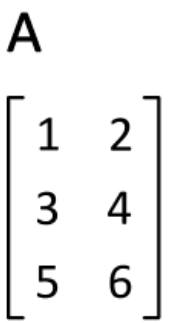
$$\boldsymbol{x}' = (x_1, x_2, \ldots, x_n)$$

Wir können auch **Matrizen transponieren**. Eine Matrix, die zuvor die Dimensionen $n \times k$ hatte, hat nach der Transposition die Dimensionsn $k \times n$.

Wenn wir eine transponierte Matrix wieder transponieren, erhalten wir:

$$(\boldsymbol{X}')' = \boldsymbol{X}$$

Darüber hinaus gilt: $(\boldsymbol{XZ})' = \boldsymbol{Z}'\boldsymbol{X}'$.



Animation von Wikipedia über Matrixtransposition.

Spezielle Matrizen

Wir begegnen in der Ökonometrie verschiedenen **speziellen Matrizen**:

- Eine $n \times n$-Matrix heißt **quadratische Matrix** (engl. **square matrix**).
- Wenn für diese quadratische Matrix gilt, dass $\boldsymbol{X}' = \boldsymbol{X}$, dann ist die Matrix **symmetrisch**.
- Eine quadratische ($n \times n$) Matrix, deren Elemente abseits der Hauptdiagonale alle 0 sind, heißt **diagonale Matrix** (engl. **diagonal matrix**) und kann mit $\text{diag}(x_1, x_2, \ldots, x_n)$ angeschrieben werden.
- Eine diagonale Matrix, deren Elemente der Hauptdiagonale alle 1 sind, heißt **Identitätsmatrix** oder **Einheitsmatrix** (engl. **identity matrix**) und wird mit $\boldsymbol{I}$ angeschrieben.
- Eine $n \times k$-Matrix, deren alle Elemente 0 sind, heißt **Nullmatrix** (engl. **zero matrix**) und wird mit $\mathbf{0}_{n \times k}$ angeschrieben.

Rang

Der **Rang** (engl. **rank**) einer Matrix ist definiert als die *Dimension des von den Spalten einer Matrix aufgespannten Vektorraumes* und wird als $\mathrm{rang}(\boldsymbol{X})$ oder $\mathrm{rank}(\boldsymbol{X})$ angeschrieben. Einfacher gesagt entspricht der Rang einer Matrix der **Anzahl ihrer linear unabhängigen Spalten** (engl. **linearly independent columns**). Eine Spalte ist dann linear unabhängig von den anderen, wenn sie nicht als lineare Kombination derselben ausgedrückt werden kann (also als eine Summe von Vielfachen der anderen Spalten). Wir betrachten die folgende Matrix:

$$\boldsymbol{X} = \begin{pmatrix} 12 & 2 & 10 \\ 3 & 1 & 2 \\ 7 & 4 & 3 \\ 8 & 6 & 2 \end{pmatrix}$$

Diese Matrix hat Rang 2. Sie hat zwar drei Spalten, aber die dritte Spalte ist eine lineare Kombination der ersten beiden: $x_{i3} = x_{i1} + (-1) \cdot x_{i2}$.

Wenn eine Matrix den maximal möglichen Rang für eine Matrix ihrer Dimensionen hat, hat sie **vollen Rang** (engl. **full rank**).

Matrixaddition, Multiplikation von Matrix und Skalar

Matrixaddition passiert Element für Element:

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} + \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1k} \\ z_{21} & z_{22} & \cdots & z_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nk} \end{pmatrix} = \begin{pmatrix} x_{11} + z_{11} & x_{12} + z_{12} & \cdots & x_{1k} + z_{1k} \\ x_{21} + z_{21} & x_{22} + z_{22} & \cdots & x_{2k} + z_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} + z_{n1} & x_{n2} + z_{n2} & \cdots & x_{nk} + z_{nk} \end{pmatrix}$$

Multiplikation einer Matrix mit einem Skalar funktioniert auch Element für Element:

$$\alpha \cdot \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} = \begin{pmatrix} \alpha x_{11} & \alpha x_{12} & \cdots & \alpha x_{1k} \\ \alpha x_{21} & \alpha x_{22} & \cdots & \alpha x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha x_{n1} & \alpha x_{n2} & \cdots & \alpha x_{nk} \end{pmatrix}$$

Multiplikation zweier Matrizen

Die **Multiplikation zweier Matrizen** ist etwas komplizierter. Sei $\boldsymbol{X}$ eine 2×3-Matrix und $\boldsymbol{Z}$ eine 3×2-Matrix. Dann können wir die Matrizen wie folgt miteinander multiplizieren:

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix} \cdot \begin{pmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \end{pmatrix} = \begin{pmatrix} x_{11}z_{11} + x_{12}z_{21} & x_{11}z_{12} + x_{12}z_{22} & x_{11}z_{13} + x_{12}z_{23} \\ x_{21}z_{11} + x_{22}z_{21} & x_{21}z_{12} + x_{22}z_{22} & x_{21}z_{13} + x_{22}z_{23} \\ x_{31}z_{11} + x_{32}z_{21} & x_{31}z_{12} + x_{32}z_{22} & x_{31}z_{13} + x_{32}z_{23} \end{pmatrix}$$

Die folgende Tabelle hilft dabei, sich den Prozess zu visualisieren:

	z_{11}, z_{21}	z_{12}, z_{22}	z_{13}, z_{23}
x_{11}, x_{12}	$x_{11}z_{11} + x_{12}z_{21}$	$x_{11}z_{12} + x_{12}z_{22}$	$x_{11}z_{13} + x_{12}z_{23}$
x_{21}, x_{22}	$x_{21}z_{11} + x_{22}z_{21}$	$x_{21}z_{12} + x_{22}z_{22}$	$x_{21}z_{13} + x_{22}z_{23}$
x_{31}, x_{32}	$x_{31}z_{11} + x_{32}z_{21}$	$x_{31}z_{12} + x_{32}z_{22}$	$x_{31}z_{13} + x_{32}z_{23}$

Es ist leicht zu sehen, dass Matrizen **nur dann** miteinander **multipliziert werden können**, wenn die Anzahl der **Spalten der linken Matrix** der Anzahl der **Zeilen der rechten Matrix** entspricht.

Eigenschaften der Rechenoperationen

- Sowohl Matrixaddition als auch Skalarmultiplikation sind kommutativ: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ sowie $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.
- Anders als Skalarmultiplikation ist Matrixmultiplikation **nicht kommutativ**, das bedeutet, $\boldsymbol{X}\boldsymbol{Z}$ ist grundsätzlich nicht das gleiche wie $\boldsymbol{Z}\boldsymbol{X}$.
- Aufgrund der Bedingung für die Dimensionen der beiden Matrizen kann es sogar sein, dass $\boldsymbol{X}\boldsymbol{Z}$ existiert, $\boldsymbol{Z}\boldsymbol{X}$ aber nicht.
- Wir nennen eine Matrix $\boldsymbol{X}$ **idempotent**, wenn $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{X}$.
- Für Matrizen mit den entsprechenden Dimensionen gilt darüber hinaus:
  - $(\boldsymbol{X} + \boldsymbol{Z}) + \boldsymbol{A} = \boldsymbol{X} + (\boldsymbol{Z} + \boldsymbol{A})$
  - $\boldsymbol{X} + \boldsymbol{0} = \boldsymbol{X}$
  - $(\boldsymbol{X}\boldsymbol{Z})\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{Z}\boldsymbol{A})$
  - $\boldsymbol{A}(\boldsymbol{X} + \boldsymbol{Z}) = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{A}\boldsymbol{Z}$
  - $(\boldsymbol{X} + \boldsymbol{Z})\boldsymbol{A} = \boldsymbol{X}\boldsymbol{A} + \boldsymbol{Z}\boldsymbol{A}$
  - $\boldsymbol{I}\boldsymbol{X} = \boldsymbol{X}\boldsymbol{I} = \boldsymbol{X}$
  - $(\alpha \boldsymbol{X})\boldsymbol{Z} = \boldsymbol{X}(\alpha \boldsymbol{Z}) = \alpha(\boldsymbol{X}\boldsymbol{Z})$

Inverse

Eine **quadratische** Matrix $\boldsymbol{X}$ heißt **invertierbar** (engl. **invertible**), wenn eine Matrix $\boldsymbol{X}^{-1}$ existiert, sodass gilt:

$$\boldsymbol{X}\boldsymbol{X}^{-1} = \boldsymbol{X}^{-1}\boldsymbol{X} = \boldsymbol{I}.$$

In diesem Fall nennen wir $\boldsymbol{X}^{-1}$ die **Inverse** (engl. **inverse**) von $\boldsymbol{X}$. Wenn keine solche Matrix existiert, nennen wir $\boldsymbol{X}$ **singulär** (engl. **singular**) oder **nicht-invertierbar** (engl. **non-invertible**). Wenn eine Inverse existiert, ist sie **eindeutig** (engl. **unique**).

- Eine Matrix $\boldsymbol{X}$ ist genau dann **invertierbar**, wenn sie **vollen Rang** hat.
- Für invertierbare Matrizen gilt darüber hinaus:
  - $(\boldsymbol{X}^{-1})^{-1} = \boldsymbol{X}$
  - $(\boldsymbol{X}')^{-1} = (\boldsymbol{X}^{-1})'$
  - $(\boldsymbol{X}\boldsymbol{Z})^{-1} = \boldsymbol{Z}^{-1}\boldsymbol{X}^{-1}$

Einführung in R mit RStudio

Summen

Matrizen und Vektoren

Zufallsvariablen

Analyse einer Zufallsvariablen

Analyse zweier Zufallsvariablen

Zufallsvariablen

Angenommen, wir beobachten ein Zufallsereignis, wie z.B. einen Münzwurf oder das Werfen eines Würfels. Eine **Zufallsvariable** ist eine Veriable, die einen Wert annimmt, der von dem beobachteten Ereignis abhängt. Wir bezeichnen sie mit einem Großbuchstaben:

$$X$$

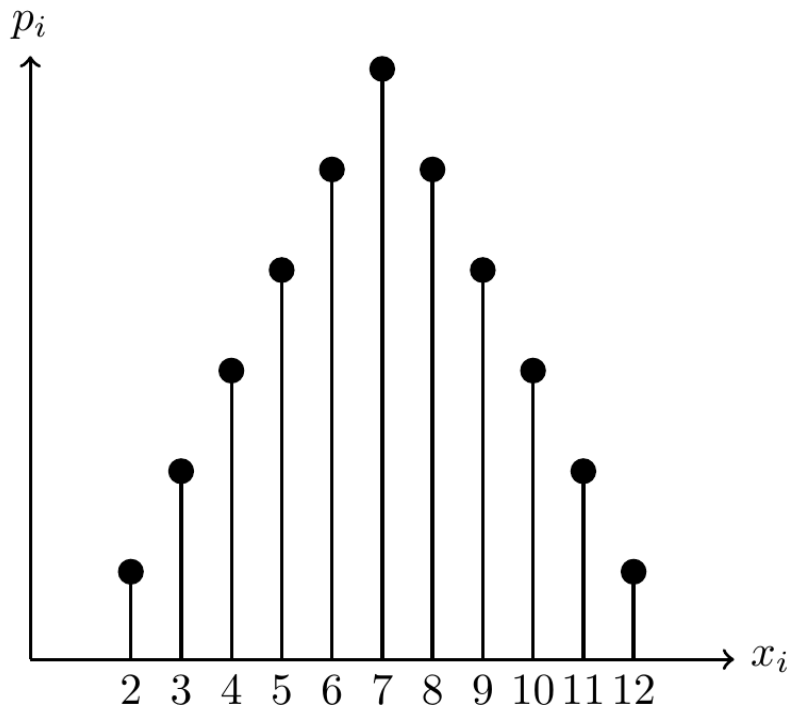
Wir bezeichnen alle möglichen Ergebnisse mit dem entsprechenden Kleinbuchstaben:

$$x_i$$

Diskrete Zufallsvariablen

Eine **diskrete Zufallsvariable** ist eine Zufallsvariable, die nur eine endliche oder abzählbar unendliche Anzahl möglicher Ergebnisse haben kann. Wenn die Variable X genannt wird, bezeichnen wir die Ergebnisse mit x_i und die zugehörigen Wahrscheinlichkeiten mit p_i. Beachte, dass die Summe aller Wahrscheinlichkeiten $\sum_i p_i$ gleich 1 sein muss.

Ein Beispiel für eine diskrete Zufallsvariable wäre das Werfen von zwei Würfeln. Die möglichen Ergebnisse sind $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$, und die zugehörigen Wahrscheinlichkeiten sind $\{\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36}\}$. Rechts ist die **Wahrscheinlichkeitsfunktion** (engl. **probability mass function**, **PMF**) dargestellt:



Eine **Bernoulli-Variable** ist eine diskrete Zufallsvariable, die nur zwei Ergebnisse annehmen kann, wie etwa ein Münzwurf.

Stetige Zufallsvariablen (1)

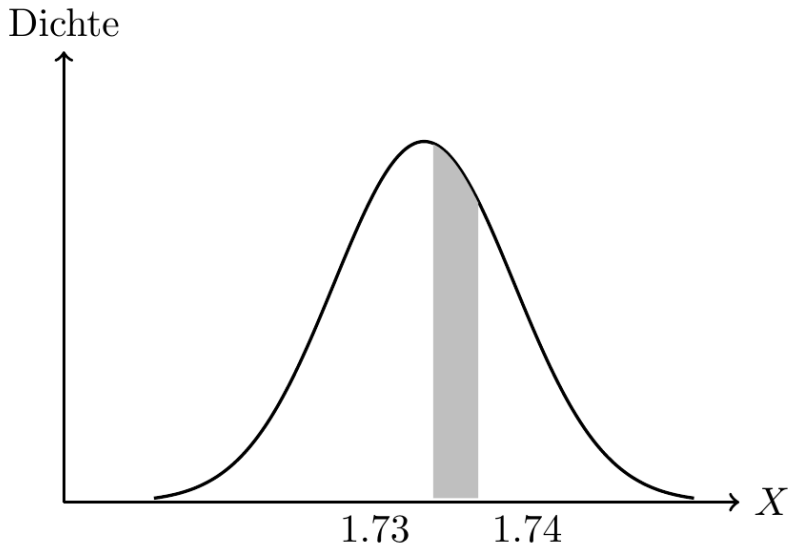
Eine **stetige** oder **kontinuierliche Zufallsvariable** (engl. **continuous random variable**) ist eine Zufallsvariable, die eine **überabzählbar unendliche** (engl. **uncountably infinite**) Anzahl unterschiedlicher Ergebnisse annehmen kann.

Wir wissen, dass es eine unendliche Anzahl von Ergebnissen gibt und dass die Summe all dieser 1 beträgt. Daraus folgt, dass die Wahrscheinlichkeit jedes einzelnen Ergebnisses gleich null ist. Daher gibt es auch keine Wahrscheinlichkeitsfunktion wie bei diskreten Zufallsvariablen.

Was wir jedoch tun können, ist, eine **Wahrscheinlichkeitsdichtefunktion** (engl. **probability density function**, **PDF**) zu zeichnen. Sie sagt uns die Wahrscheinlichkeit, dass das Ergebnis in ein bestimmtes Intervall fällt. Der Flächeninhalt unter der gesamten PDF entspricht 1.

Stetige Zufallsvariablen (2)

Ein Beispiel für eine solche Variable wäre die **Körpergröße einer Person**. Es wäre sinnlos, nach der Wahrscheinlichkeit zu fragen, dass eine Person genau 1,734681092536 Meter groß ist. Diese Wahrscheinlichkeit ist null. Aber wir können uns die PDF ansehen und bestimmen, wie wahrscheinlich es ist, dass die Körpergröße der Person zwischen 1.73 und 1.74 Meter liegt:

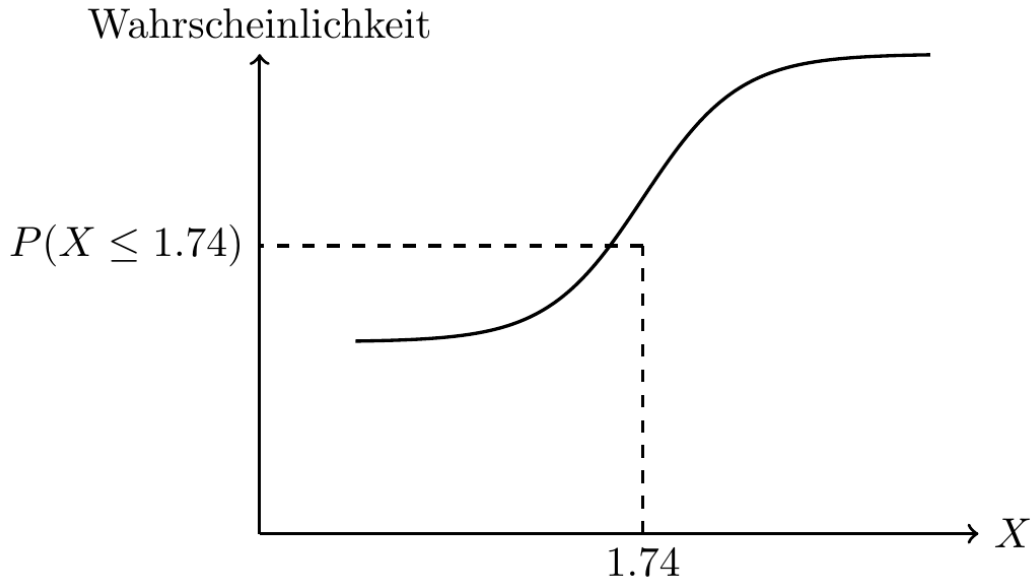


Ob eine unendliche Menge von Zahlen **abzählbar** oder **überabzählbar** unendlich ist, kann intuitiv beantwortet werden. Alle **natürlichen** Zahlen $\mathbb{N}$ sind abzählbar unendlich. Wir können klar einen Weg vorgeben, wie man sie zählt (beginne bei 0, dann 1, dann 2, dann 3, ...), wir wissen nur nicht, wo und wann der Weg *endet*. Schließlich ist er immer noch unendlich lang. Alle **reellen** Zahlen $\mathbb{R}$ sind jedoch überabzählbar unendlich. Wir können keinen eindeutigen Weg bestimmen, der alle Zahlen erreicht. Angenommen, wir beginnen bei 0, dann 0.001 was ist mit allen Zahlen dazwischen? Und allen Zahlen zwischen diesen Zahlen? Es gibt keinen Weg, sie alle zu zählen.

Wahrscheinlichkeitsverteilungsfunktion

Zusätzlich zur **Wahrscheinlichkeits-dichtefunktion** können wir die **Verteilungsfunktion** (engl. **cumulative distribution function**, **CDF**) zeichnen. Sie gibt die Wahrscheinlichkeit an, dass das Ergebnis gleich oder kleiner als ein bestimmter Wert ist. Die Funktion ist streng monoton steigend:

Die gestrichelte Linie zeigt, wie wir den Plot lesen: Der Wert der Dichtefunktion bei $X = 1.74$ stellt die Wahrscheinlichkeit dar, dass eine zufällig ausgewählte Person kleiner oder genau 1.74 Meter groß ist.



Summen

Matrizen und Vektoren

Zufallsvariablen

Analyse einer Zufallsvariablen

Analyse zweier Zufallsvariablen

Erwartungswert (1)

Kehren wir zu unserem Beispiel des Würfelwurfs zurück. Das Ergebnis ist eine diskrete Zufallsvariable mit den folgenden Ergebnissen und zugehörigen Wahrscheinlichkeiten:

Ergebnis	Wahrscheinlichkeit
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

Erwartungswert (2)

Der **Erwartungswert** (engl. **expected value** oder **expectation**) ist ein Konzept, das uns erlaubt – ganz einfach – zu analysieren, welchen Wert wir beim Würfeln *erwarten* können. Wir berechnen ihn als das arithmetische Mittel der Ergebnisse, gewichtet mit ihren jeweiligen Wahrscheinlichkeiten. Wir bezeichnen den Erwartungswert mit einem großen E:

$$\mathrm{E} := \sum_{i=1}^{n} x_i p_i$$

Der Erwartungswert für einen fairen Würfel beträgt 3.5. Wenn wir immer mehr Ergebnisse aus dieser Verteilung ziehen, d.h. den Würfel sehr oft werfen, wird sich der Durchschnitt aller Würfe immer mehr dem Erwartungswert annähern. Solange wir mit diskreten Variablen arbeiten, ist all dies recht einfach zu interpretieren. Bei kontinuierlichen Variablen wird es schwieriger, aber die allgemeine Intuition bleibt bestehen.

Regeln für Erwartungswerte

Wir beschäftigen uns in der Ökonometrie viel mit Erwartungswerten, daher ist es nützlich, einige Regeln im Umgang damit zu kennen.

- Für eine Konstante c: $\mathrm{E}(c) = c$
- Für Zufallsvariablen X, Y und Konstanten c, d:
  $\mathrm{E}(c \cdot X + d \cdot Y) = c \cdot \mathrm{E}(X) + d \cdot \mathrm{E}(Y)$
- Für Konstanten $c_1, \ldots, c_n$ und Zufallsvariablen $X_1, \ldots, X_n$:
  $\mathrm{E}\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i \mathrm{E}(X_i)$
- Für zwei *unabhängige* Zufallsvariablen X, Y: $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$

Varianz (1)

Oftmals reicht der Erwartungswert nicht aus, um eine Verteilung zu analysieren. Stellen Sie sich vor, Sie besitzen eine Firma, die Schrauben herstellt. Sie haben zwei Maschinen, die sie produzieren. Sie werben damit, dass Ihre Schrauben alle 35 Millimeter lang sind, aber in Wirklichkeit ist die Länge der Schrauben zufällig verteilt: Der Erwartungswert der Schraubenlänge beträgt für beide Maschinen 35mm. Allerdings produziert die Maschine A meist Schrauben, die sehr nah an der gewünschten Länge sind, während die Maschine B manchmal Schrauben ausgibt, die sogar 33 oder 37 Millimeter lang sind. Was ist der Unterschied zwischen diesen beiden Maschinen mit identischen Erwartungswerten?

Die Antwort lautet **Varianz**. Vereinhachend gesagt: Der Erwartungswert zeigt uns, wo das "Zentrum" einer Verteilung liegt. Die Varianz hingegen gibt an, wie weit die Ergebnisse tendenziell von dieser Erwartung abweichen. Wir bezeichnen sie als $\mathrm{Var}(X)$ und berechnen sie wie folgt:

$$\mathrm{Var}(X) := \mathrm{E}\left((X - \mu)^2\right),$$

wobei $\mu = \mathrm{E}(X)$.

Varianz (2)

Es leuchtet ein, dass die Varianz jeder Konstante null ist. Zu beachten ist außerdem die folgende Regel für eine Zufallsvariable X und Konstanten a, b:

$$\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X) + \mathrm{Var}(b) = a^2\mathrm{Var}(X)$$

Die Standardabweichung, bezeichnet als $\mathrm{sd}(X)$, ist einfach die Quadratwurzel der Varianz.

Analyse zweier Zufallsvariablen

Gemeinsame Wahrscheinlichkeitsverteilung

Angenommen, X und Y sind zwei diskrete Zufallsvariablen. Zusätzlich zu ihren individuellen Verteilungen können wir ihre **gemeinsame Verteilung** (engl. **joint distribution**) beschreiben. Dafür verwenden wir eine gemeinsame Wahrscheinlichkeitsfunktion:

$$f_{X,Y}(x, y) = P(X = x, Y = y)$$

Diese Funktion gibt einfach an, wie groß die Wahrscheinlichkeit für jede Kombination von X und Y ist. Wenn X und Y unabhängig sind, dann gilt:

$$f_{X,Y}(x, y) = f_X(x) f_Y(y),$$

wobei $f(x)$ und $f(y)$ die Wahrscheinlichkeitsfunktionen für X bzw. Y sind. Zwei Zufallsvariablen sind **unabhängig** (engl. **independent**), wenn das Ergebnis von X die Wahrscheinlichkeiten der möglichen Ergebnisse von Y nicht beeinflusst.

Bedingte Verteilung

Ein weiteres wichtiges Konzept ist die **bedingte Verteilung** (engl. **conditional distribution**). Die bedingte Wahrscheinlichkeitsdichtefunktion beschreibt, wie das Ergebnis von X das von Y beeinflusst:

$$f_{Y|X}(y|x) = P(Y = y|X = x) = \frac{f_{X,Y}(x, y)}{f_X(x)}, \text{ für alle } f_X(x) > 0$$

Wenn X und Y unabhängig sind, beeinflusst das Ergebnis von X nicht Y und somit gilt $f_{Y|X}(y|x) = f_Y(y)$.

Kovarianz

Die **Kovarianz** (engl. **covariance**) ähnelt einer “zwei-Variablen-Version” der Varianz. Wir könen damit zwei Verteilungen gemeinsam analysieren. Sie wird wie folgt definiert und mit $\mathrm{Cov}(X, Y)$ bezeichnet:

$$\mathrm{Cov}(X, Y) := \mathrm{E}\left((X - \mu_X)(Y - \mu_Y)\right),$$

wobei $\mu_X = \mathrm{E}_X(X)$ und $\mu_Y = \mathrm{E}_Y(Y)$.

Die Vorzeichen der Kovarianz können intuitiv interpretiert werden. Ist die Kovarianz positiv, erwarten wir, dass Y über seinem Mittelwert liegt, wenn X das ebenfalls tut. Ist die Kovarianz negativ, erwarten wir, dass Y unter seinem Mittelwert liegt, wenn X über seinem Mittelwert liegt. Einfach gesagt, zeigt eine positive Kovarianz, dass zwei Variablen positiv miteinander assoziiert sind, und umgekehrt. Eine Kovarianz von 0 bedeutet, dass keine Beziehung besteht. Wenn X und Y unabhängig sind, ist die Kovarianz immer 0.

Eine Assoziation in diesem Sinne bedeutet natürlich noch lange keinen kausalen Zusammenhang, aber mehr dazu im Kurs :)

Regeln für die Kovarianz

Folgende Regeln gelten für die Kovarianz:

$$\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$$

Für Konstanten a, b, c, d:

$$\mathrm{Cov}(aX + b, cY + d) = a \cdot c \cdot \mathrm{Cov}(X, Y)$$

Bedingte Erwartung

Angenommen, wir haben zwei Zufallsvariablen X und Y, die in irgendeiner Weise miteinander verbunden sind. Wir möchten wissen, was die Erwartung von Y ist, vorausgesetzt, X nimmt einen bestimmten Wert an. Dies wird als **bedingter Erwartungswert** bezeichnet und mit $\mathrm{E}(Y|X=x)$ notiert. Für bedingte Erwartungswerte gilt folgendes:

- Wenn $f(x)$ eine beliebige Funktion ist: $\mathrm{E}\left(f(X)|X\right) = f(X)$
- Für zwei Funktionen $f(X)$ und $g(x)$:
  $\mathrm{E}\left(f(X)Y + g(X) \mid X\right) = f(X)\mathrm{E}(Y|X) + g(X)$
- Wenn X und Y unabhängig sind: $\mathrm{E}(Y|X) = \mathrm{E}(Y)$
- Der **Satz der iterierten Erwartungswerte** (engl. **law of iterated expectations**):
  $\mathrm{E}\left(\mathrm{E}(Y|X)\right) = \mathrm{E}(Y)$