

Einstieg

Wigand Rathmann

Sommersemester 2020

1. Thema: Daten und beschreibende Statistik

Willkommen zur C4

Wer begleitet Sie?

Dr. Wigand Rathmann

wigand.rathmann@fau.de, 09131/85-67129

www.math.fau.de/rathmann

Sprechstunde: Mi 11-12 Uhr

Dr. Marius Yamakou

marius.yamakou@fau.de, 09131/85-67127

<https://www.math.fau.de/angewandte-analysis/mitarbeitende/dr-marius-yamakou/>

Sprechstunde: Di 14-15 Uhr

0.1

Willkommen

Ziele

- Organisatorisches
- Einstieg ins Thema: Was sind Daten?
- Einstieg ins Thema: Erste Daten über Daten
- Was kommt im Semester.

0.2

Organisatorisches

Skript, Übungsaufgaben

Im StudON

<https://www.studon.fau.de/crs2897780.html>

Beitritt mit Kurspasswort

Kovarianzmatrix

Anmeldung

heute 13:45 Uhr bis Donnerstag, 23.04.2020 14 Uhr unter

Anmeldung zu den Übungsgruppen Mathe C4 (SoSe 2020)

Ausgabe und Abgabe der Übungen

Ausgabe der Übungsblätter: Do./Fr. **vor** der 1. Übung,
1. Blatt am 23.4.

Abgabe der Übungsblätter: bis Do. **bis** 12 Uhr (im Briefkasten oder online, wird auf dem Blatt bekanntgegeben),
1. Blatt am 5.5.

0.3

Literatur

Das Kurzsript ist folgenden Büchern und Skripten entnommen:

Hübner, G.: *Stochastik*, Vieweg+Teubner, Wiesbaden, 2009

Beyer, Hackel, Pieper, Tiedge: *Wahrscheinlichkeitsrechnung und mathematische Statistik*, Leipzig, Teubner, 1991

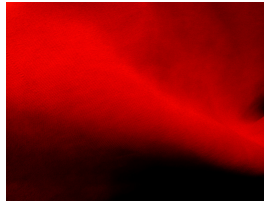
Groß, J.: *Grundlegende Statistik mit R*. Vieweg+Teubner, Wiesbaden, 2010

Graef, F.: *Skript zu Wahrscheinlichkeitsrechnung I*, FAU AM2, 2003

0.4

Stochastik

Stochastik =



Quelle: <https://www.publicdomainpictures.net/de/view-image.php?image=38242>, CC0

Warum?

0.5

Stochastik

Was brauchen wir aus der Analysis

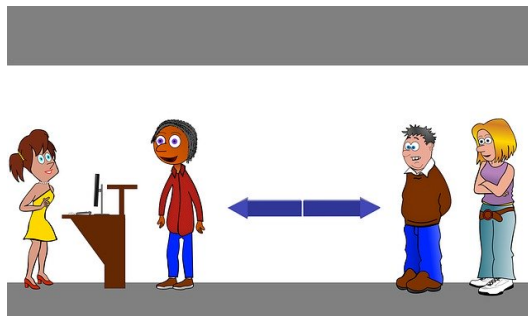
- Summen, Produkte
- Riemann-Integration
- Abbildungsbegriff

Was ist neu?

- Begriffe
- Modelle mathematisch zu formulieren

0.6

Warteschlange



Quelle: Bild von Christian Dorn auf Pixabay

0.7

Daten

Beispiel für Datenerhebung

Umfrage (<https://www.studon.fau.de/vote/CVHP>)



0.8

Daten

Datenarten

- nominal
- ordinal
- metrisch

Begriffe

- Modus
- Median
- Mittelwert
- Spannweite

0.9

Beispiel 1.1. Messung des Gewichts von zehnjährigen Kindern

Folgende Werte wurden gemessen:

34.6	34.5	28.6	37.8	24.8	24.9	27.7	34.0	24.9	30.4
32.0	24.9	35.6	38.1	32.4	33.1	41.5	29.6	37.2	26.6
33.0	36.8	34.1	29.9	29.3	27.6	29.9	35.0	24.1	36.0

Urliste

Die Liste aller ermittelten bzw. gemessenen Daten wird als **Urliste** oder **Datensatz** bezeichnet.

Was kann man mit diesen Zahlen anfangen?

- Sortieren
- Gruppieren ($< 30, \geq 30$)
- In Klassen zusammenfassen $[23, 25]$ bis $(41, 43]$.

Definition 1.2. Die aus einer Beobachtung oder Messung aufgezeichneten Daten heißen **Datensatz**. Es wird auch von der **Urliste** gesprochen. Ist der Datensatz aufgrund einer gezielten Teilerhebung entstanden, heißt er **Stichprobe**.

https://www.studon.fau.de/pg636985_2897784.html

0.10

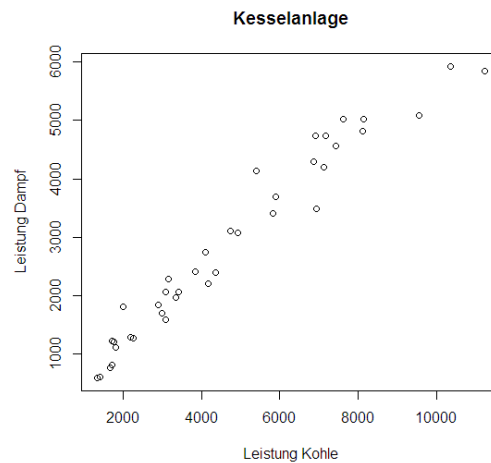
Daten

Was können wir mit Daten tun?

- Sortieren
- Gruppieren
- Kennzahlen berechnen:
 - Median, Mittelwerte, ...
 - Quantile, ...
- empirische Verteilungsfunktion bauen
- Visualisieren (Histogramme, Boxplots)

0.11

Bivariate Daten



https://www.studon.fau.de/pg636987_2897784.html

0.12

Inhalte

- Zufallsexperimente und Wahrscheinlichkeitsräume,
- Laplace-Experimente, Kombinatorik,
- Bedingte Wahrscheinlichkeiten,
- Wahrscheinlichkeiten auf diskreten Ergebnismengen,
- Verteilungen und Zufallsvariable,
- Funktionen von Zufallsvariablen,
- Erwartungswert und Varianz,
- Normalverteilung,
- Grundbegriffe der mathematischen Statistik,
- Grenzwertsätze.

0.13

Selbststudium

Ihre Aufgaben

- (i) Eignen Sie sich die Begriffe aus der beschreibenden Statistik an. Ordnungsstatistik, Rangwert, Mittelwerte, Median, Perzentile, Klassen, empirische Varianz, Standardabweichung, Kovarianz etc.
- (ii) Machen Sie sich mit den graphischen Möglichkeiten vertraut (Histogramm, Boxplot, empirische Verteilungsfunktion)

- (iii) Wiederholen Sie bitte den Begriff **Lineares Ausgleichsproblem** aus dem 1. Semester

Weiterführende Fragen

- (i) Finden Sie ein Faustregel, wieviele Klassen für einen Datensatz günstig sind?
- (ii) Gibt es Möglichkeiten, die Güte einer Regression zu messen?

0.14

Selbststudium

Quellen

- Skript Kapitel 2 (https://www.studon.fau.de/file2897817_download.html)
- Kopien Buch: Hübner, G. Stochastik. Vieweg. Kapitel 1
- StudOn Lernmodul https://www.studon.fau.de/pg636985_2897784.html
- Seite im Lernmodul zum Ausprobieren https://www.studon.fau.de/pg705490_2897784.html (kleiner Datensatz).

Zum Anschauen

Hier finden Sie Visualisierungen der Coronadaten. Dabei werden auch Hilfsmittel der linearen Regression genutzt.

<https://netmath.vcrp.de/downloads/Einheiten/Rathmann/CORONA-Datenanzeigen.html>

<https://netmath.vcrp.de/downloads/Einheiten/Rathmann/CORONA-LinWachstum.html>

<https://netmath.vcrp.de/downloads/Einheiten/Rathmann/CORONA-ExpWachstum-1.html>

<https://netmath.vcrp.de/downloads/Einheiten/Rathmann/CORONA-ExpWachstum-2.html>

0.15

Ihre Fragen

... stellen

- Im Forum <https://www.studon.fau.de/fm2897793.html>
- Die Fragen, die bis Donnerstag gestellt wurden, werden am Freitag in der Online-Runde diskutiert.

0.16

Guten Start

Guten Start ins neue Semester!

0.17