

Übungsblatt 1

Beschreibende Statistik

Stochastik@AIN2

Prof. Dr. Barbara Staehle

Sommersemester 2021

HTWG Konstanz

Einfache und mittelschwere Aufgaben

AUFGABE 1.1 TIKTOK-VIDEOS

Sie interessieren sich für die Frage, wie ein typisches TikTok-Video aussieht und führen eine statistische Stichprobenerhebung durch.

TEILAUFGABE 1.1.1 3 PUNKTE

Geben Sie **Beispiele** an, was für diese Untersuchung die folgenden Kenngrößen sein könnten:

- Statistische Einheiten
- Grundgesamtheit
- Stichprobe
- Merkmale (mindestens 3)
- **Beispielhafte** Merkmalsausprägungen (jeweils für Ihre Beispiele)
- Eigenschaften der Merkmale (qualitativ ordinal/nominal, quantitativ diskret/kontinuierlich) (jeweils für Ihre Beispiele)

TEILAUFGABE 1.1.2 1 PUNKT

Welche Eigenschaften von TikTok-Videos finden Sie besonders spannend und einer näheren Untersuchung wert? Begründen Sie Ihre Meinung.

AUFGABE 1.2 WOHNUNGSGRÖSSEN

10 Personen aus Konstanz wurden nach der Anzahl der Zimmer in ihrer Wohnung gefragt. Es ergab sich folgende Urliste:

3, 4, 5, 1, 5, 2, 1, 3, 1, 3

TEILAUFGABE 1.2.1 2 PUNKTE

Geben Sie die relative und absolute Häufigkeiten der genannten Zimmeranzahlen an.

TEILAUFGABE 1.2.2 2 PUNKTE

Zeichnen Sie ein Histogramm für die absoluten Häufigkeiten.

TEILAUFGABE 1.2.3 2 PUNKTE

Zeichnen Sie ein Diagramm für die empirische Verteilungsfunktion.

TEILAUFGABE 1.2.4 1 PUNKT

Geben Sie an, wie viele Prozent der Befragten

- a) maximal 4 Zimmer zur Verfügung haben.
- b) mindestens 2 und maximal 4 Zimmer zur Verfügung haben.

TEILAUFGABE 1.2.5 5 PUNKTE

Berechnen Sie für die Stichprobe (der Zimmergrößen):

- a) das arithmetische Mittel
- b) den Median
- c) den Modalwert
- d) das 25%-Quantil
- e) das 30%-Quantil
- f) das 66%-Quantil
- g) die Varianz
- h) die Standardabweichung
- i) den Interquartilabstand
- j) die Spannweite

TEILAUFGABE 1.2.6 1 PUNKT

In München werden 20 anderen Personen ebenfalls nach der Anzahl der Zimmer in ihrer Wohnung gefragt. Die Werte in dieser Stichprobe ergeben den gleichen Mittelwert wie die Ergebnisse aus Konstanz, allerdings ergeben diese eine **höhere Varianz**.

Was schließen Sie hieraus hinsichtlich des Unterschieds zwischen den beiden Stichproben?

AUFGABE 1.3 GELDSCHEINE, 2 PUNKTE

Angenommen, es wären von allen vorhandenen Werten (5 €, 10 €, 20 €, 50 €, 100 €, 200 €, 500 €) gleich viele Geldscheine im Umlauf.

- a) Bestimmen Sie den Median des Werts der Geldscheine im Umlauf.
- b) Berechnen Sie den Median, der sich nach der Einführung eines 1000 €-Scheins (und davon ebenfalls so viele in den Umlauf gebracht wie alle anderen) ergeben würde.

AUFGABE 1.4 AN DER FAHRRADBRÜCKE (KLAUSURAUFGABE WS 16/17), 4 PUNKTE

Alice steht an der Konstanzer Fahrradbrücke und zählt die Anzahl der Räder der Fahrzeuge (Einräder, Fahrräder, Dreiräder, Kinderwagen, ...) welche an ihr vorbei rollen.

Ihre Beobachtungen resultieren in folgender Urliste:

2, 4, 3, 1, 2, 4, 2, 2, 2, 3.

Berechnen Sie für die Beobachtungen von Alice:

- a) das arithmetische Mittel

- b) den Median
- c) den Modalwert
- d) das 10%-Quantil
- e) das 25%-Quantil
- f) das 75%-Quantil
- g) den Interquartilabstand
- h) die Spannweite

Geben Sie jeweils an, wie Sie die geforderten Charakteristika berechnet haben!

AUFGABE 1.5 TASCHENGELD (KLAUSURAUFGABE SS 18)

Frank schreibt für seine Schülerzeitung einen Artikel zum Thema „Taschengeld eines Achtklässlers“. Hierfür befragt er 6 zufällig ausgewählte Schüler der Klasse 8a nach ihrem monatlichen Taschengeld in Euro. Seine Umfrage ergibt die folgende Urliste:

25, 10, 30, 25, 35, 25

TEILAUFGABE 1.5.1 2 PUNKTE

Berechnen Sie für Frank's Urliste folgende Größen. Stellen Sie Ihren Rechenweg nachvollziehbar dar!

- a) Arithmetisches Mittel
- b) Median
- c) 75%-Quantil
- d) empirische Standardabweichung

TEILAUFGABE 1.5.2 3 PUNKTE

Beantworten Sie folgende Fragen möglichst allgemeinverständlich!

- a) Frank berechnet außerdem noch die Spannweite seiner Stichprobe zu $R = 5$ und den Interquartilabstand zu $I = 25$. Machen diese Berechnungen Sinn? Begründen Sie Ihre Meinung!
- b) Mittelwert und Median von Franks Daten sind gleich. Ist das immer so, oder ist das Zufall?
- c) Was hat der Wert des 75%-Quantils zu bedeuten? Fassen Sie Ihr Ergebnis in Worte.
- d) Frank führt seine Umfragen nach dem monatlichen Taschengeld noch in zwei anderen Klassen, den Klassen 8b und 8c, durch. In beiden Fällen erhält er das selbe arithmetische Mittel wie bei Klasse 8a. Im Fall von Klasse 8b erhält er allerdings eine Standardabweichung von 0, im Fall von Klasse 8c erhält er eine Standardabweichung, die deutlich größer ist als die der Klasse 8a. Was schließen Sie hieraus hinsichtlich der Stichproben des monatlichen Taschengelds aus den anderen beiden Klassen?
- e) Frank interessiert sich weiterhin für die Schuhgröße der Befragungsteilnehmer, sowie für ihre Mathenote. Er berechnet hierfür die Korrelationskoeffizienten von Schuhgröße und Taschengeld zu $r_{S,T} = 0.875$ und von Mathenote und Taschengeld zu $r_{M,T} = -0.038$. Was schließen Sie hieraus über den Zusammenhang der betrachteten Größen?
- f) Grace befragt 3 Schüler nach ihrem Taschengeld und berechnet den arithmetische Mittelwert dieser Stichprobe zu 300 €. Beschreibt diese Zahl die Stichprobe ausreichend gut, oder sollte sie besser noch andere Werte berechnen?

AUFGABE 1.6 SPORTWAGEN (KLAUSURAUFGABE WS 17/18)

Charlie besucht die IAA (Internationale Automobilausstellung). Er interessiert sich für Sportwagen und sucht daher Informationen zu den Höchstgeschwindigkeiten der vorgestellten Autos.

Am Stand von Hersteller 1 analysiert er die Datenblätter von 10 verschiedenen Fahrzeugen, denen er folgende Höchstgeschwindigkeiten entnimmt:

250, 330, 250, 350, 270, 300, 260, 330, 200, 400

TEILAUFGABE 1.6.1 2 PUNKTE

Berechnen Sie für die Stichprobe der von Charlie ermittelten Höchstgeschwindigkeiten folgende Größen. Stellen Sie Ihren Rechenweg nachvollziehbar dar!

- a) Arithmetisches Mittel
- b) Median
- c) 90%-Quantil
- d) Standardabweichung

TEILAUFGABE 1.6.2 3 PUNKTE

Gefragt sind Beschreibungen in Ihren eigenen Worten (Zielgruppe: nicht-technisches Publikum).

- a) Nennen Sie **zwei** Fakten, wie die von Ihnen berechneten Größen einzeln oder gemeinsam betrachtet die von Charlie erhobene Stichprobe der Höchstgeschwindigkeiten charakterisieren.
- b) Charlie entnimmt aus einem Datenblatt von Hersteller 2 ebenfalls die Höchstgeschwindigkeiten von 10 verschiedenen Fahrzeugen und berechnet den Mittelwert dieser Höchstgeschwindigkeiten zu 294 sowie die Standardabweichung zu 20. Was schließen Sie hieraus hinsichtlich des Unterschieds zwischen den Fahrzeugen der beiden Hersteller?
- c) Charlie betrachtet weiterhin die Preise der Reifen welche für die Fahrzeuge von Hersteller 1 zugelassen sind, sowie die Umweltfreundlichkeit des Fahrzeugs auf einer Skala von 0 (gar nicht) bis 10 (sehr).
Er berechnet hierfür die Korrelationskoeffizienten von Höchstgeschwindigkeit und Reifenpreis zu $r_{G,R} = 0.853$ und von Höchstgeschwindigkeit und Umweltfreundlichkeit zu $r_{G,U} = -1.137$. Was schließen Sie hieraus über den Zusammenhang der betrachteten Größen?

Mittelschwere und schwere Aufgaben

AUFGABE 1.7 STICHPROBE MIT KLASSEN (ELEKTRONISCH LÖSEN!), 3 PUNKTE

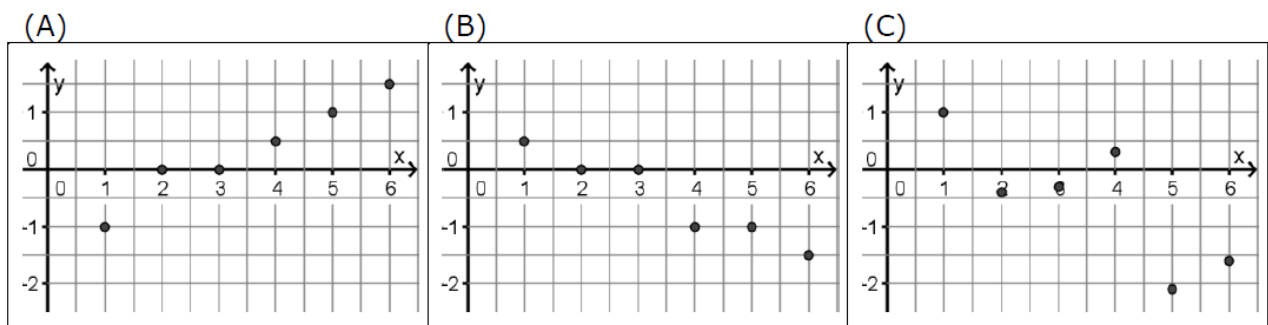
Gegeben ist folgende Urliste einer Stichprobe vom Umfang $n = 20$:

3, 7, 12, 18, 19, 20, 25, 25, 27, 28, 29, 31, 32, 34, 37, 38, 40, 41, 45, 47

- Gruppieren Sie die Stichprobenwerte in geeignete Klassen und bestimmen Sie die absoluten und die relativen Häufigkeiten der Klassen.
- Stellen Sie die für die Klassen erhaltenen absoluten Häufigkeiten als Histogramm dar.

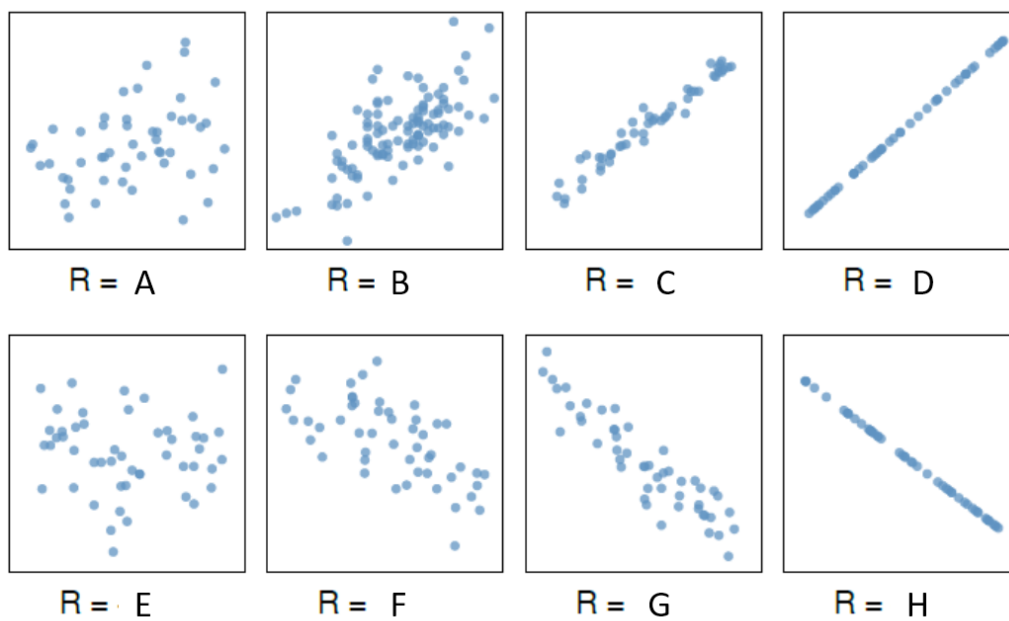
AUFGABE 1.8 PUNKTWOLKEN

TEILAUFGABE 1.8.1 3 PUNKTE



Bei welcher der dargestellten Punktwolken (A), (B) oder (C) erhält man als Korrelationskoeffizienten $r_{x,y} = -0,966$? Begründen Sie Ihre Meinung und nennen Sie für die beiden anderen Punktwolken jeweils einen Ablehnungsgrund.

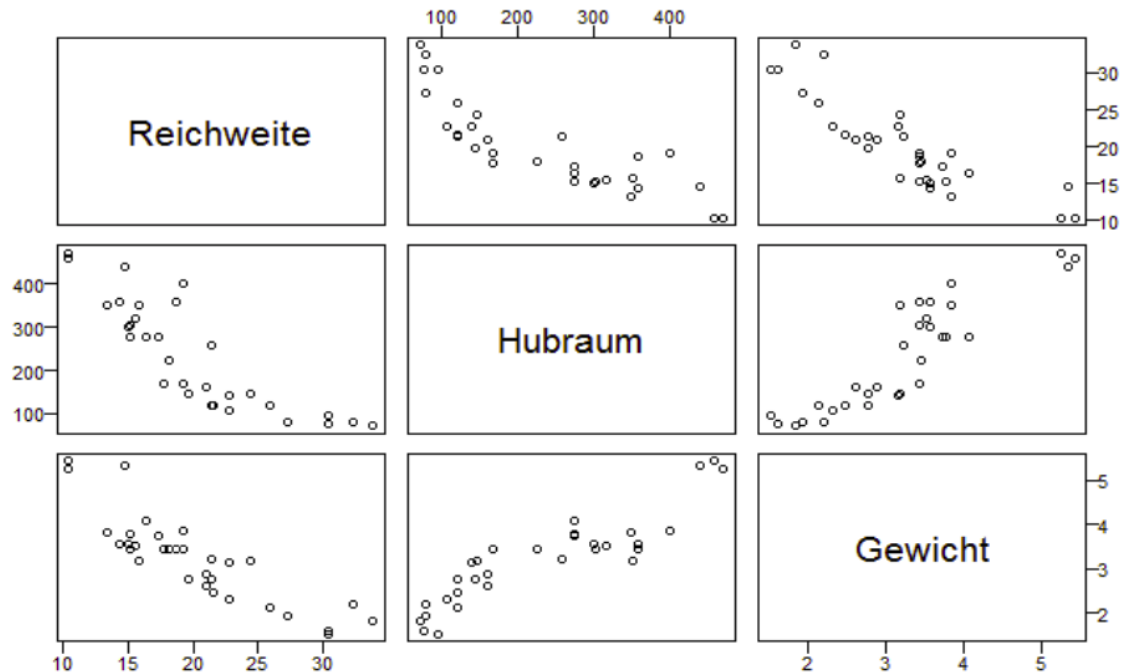
TEILAUFGABE 1.8.2 3 PUNKTE



Wählen Sie für jede der dargestellten Punktwolken aus den folgenden Werten denjenigen aus, der am besten für den Korrelationskoeffizienten passt. Achtung: Nicht alle Werte passen zu einem Bild.

−0.92 0.00 0.69 −0.08 −0.64 −0.20 0.33 1.00 −1.00 0.98

TEILAUFGABE 1.8.3 3 PUNKTE



Die abgebildete Matrix stellt Punktwolken für einige Eigenschaften einer Auswahl verschiedener Autotypen des gleichen Herstellers dar. Der Plot oben rechts z.B. zeigt auf der x-Achse den Faktor Gewicht, auf der y-Achse den Faktor Reichweite. Ein Punkt entspricht einem Auto, geeignete Einheiten fehlen, diese dürfen Sie sich selbst ausdenken.

Welche Aussagen über die Zusammenhänge zwischen diesen drei Faktoren lassen sich aus der Grafik ableiten?

AUFGABE 1.9 AUSBILDUNG UND GEHÄLTER (ELEKTRONISCH LÖSEN!), 3 PUNKTE

Fünf Mitglieder eines Sportclubs gehen der Frage nach, ob zwischen ihrer Schul- und Ausbildungszeit und ihrem Jahreseinkommen ein statistischer Zusammenhang besteht. Das Ergebnis der anonymen Befragung ist in folgender Tabelle wiedergegeben.

Ausbildungsdauer X	9	13	15	18	20
Jahresgehalt Y [1000 €]	18	37	61	125	59

Berechnen Sie den empirischen Korrelationskoeffizient der beiden untersuchten Merkmale. Interpretieren Sie diesen Wert!

AUFGABE 1.10 FERNSEHEN UND AGRESSIONEN (ELEKTRONISCH LÖSEN!), 3 PUNKTE

In einer Studie zur Auswirkung von YouTube-Filmen mit gewalttätigen Szenen auf das Sozialverhalten von Kindern wurden ein Aggressivitätsmaß Y (auf einer Skala von 0-10), die Zeitdauer X (in Minuten), während der das Kind gewöhnlich solche Filme sieht, und das Geschlecht Z des Kindes (0 männlich, 1 weiblich) erfasst.

Die Ergebnisse in einer zufällig gewählten Testgruppe waren wie folgt:

i	1	2	3	4	5	6	7	8	9	10	11	12	13
x_i	10	50	30	70	80	60	90	40	10	20	30	50	60
y_i	4	5	2	6	6	8	7	2	7	3	5	1	3
z_i	0	0	0	0	0	0	0	1	1	1	1	1	1

Verwenden Sie diese Testergebnisse, für eine (empirisch begründete) Aussage zum Thema "macht YouTube aggressiv"?

Empfohlene Vorgehensweise:

- Zeichnen Sie ein Streudiagramm für die 13 Kinder, und berechnen Sie den Korrelationskoeffizienten zwischen Dauer des Medienkonsums und Aggressionslevel ohne Berücksichtigung des Geschlechts.
- Zeichnen Sie nun für Jungen und Mädchen getrennt jeweils ein Streudiagramm und berechnen Sie für beide Geschlechter den Korrelationskoeffizienten.
- Vergleichen Sie alle Ergebnisse und ziehen Sie Ihre Folgerungen.

AUFGABE 1.11 KLAUSURNOTEN (KLAUSURAUFGABE SS19)

Prof. Schmidt liest für alle Studierenden des Studiengangs „Bachelor Super-Informatik“ die Vorlesungen „Mathematik 1“ und „Mathematik 2“.

Um die Noten der beiden Vorlesungen vergleichen, analysiert sie die Noten, welche 6 zufällig ausgewählte Studierende (s_1, s_2, \dots, s_6) in den beiden Klausuren jeweils erreichten. Sie erhält die folgende Tabelle:

Studierende(r)	s_1	s_2	s_3	s_4	s_5	s_6
Note Mathematik 1	1.3	5.0	1.3	2.7	4.0	3.7
Note Mathematik 2	2.0	3.0	1.0	2.0	1.0	2.7

TEILAUFGABE 1.11.1 4 PUNKTE

Berechnen Sie **nur für die Noten der Vorlesung Mathematik 1** folgende Größen. Stellen Sie Ihren Rechenweg nachvollziehbar dar, bzw. geben Sie die Verwendung des Taschenrechners oder eines anderen elektronischen Tools an!

- Arithmetisches Mittel
- Alle Quartile
- 90%-Quantil
- empirische Standardabweichung

TEILAUFGABE 1.11.2 4 PUNKTE

Beantworten Sie folgende Fragen in ganzen Sätzen und möglichst allgemeinverständlich!

- Was bedeutet der Wert des 90%-Quantils (Ihr Ergebnis aus Aufgabe 1.11.1c) für die Noten von Mathematik 1?
- Prof. Schmidt berechnet für die Noten der Vorlesung Mathematik 2 folgende Größen:

$$\bar{x} = 2.0, \quad (Q1, Q2, Q3) = (1.0, 2.0, 2.7), \quad \tilde{x}_{0.9} = 3.0, \quad s = 0.8462$$

Verwenden Sie diese Zahlen, sowie Ihre eigenen Berechnungen aus Aufgabe 1.11.1 um **zwei** Gemeinsamkeiten oder Unterschiede für die Noten in Mathematik 1 und 2 abzuleiten. Begründen Sie Ihre Meinung!

- Welcher der folgenden Werte ist der korrekte empirische Korrelationskoeffizienten für die Noten in Mathematik 1 und Mathematik 2 $r_{1,2}$ (für die obige Stichprobe)? Begründen Sie Ihre Meinung.

$$r_{1,2} = \quad 0 \quad -0.571 \quad 0.423 \quad 1.1 \quad 0.952$$