

# Tutorielle Tätigkeit im Modul *Statistik*

HENRY HAUSTEIN

## Beschreibung der Tätigkeit

Meine grundsätzliche Aufgabe ist das Halten von 2 Tutorien à 90 Minuten pro Woche. In den Tutorien werden jede Woche neue Übungsaufgaben besprochen; wie ich die Inhalte aufbereite, bleibt mir überlassen. Ich habe mich dafür entschieden meine Tutorien so zu halten, dass ich zuerst kurz die theoretischen Grundlagen aus der Vorlesung vorstelle und einige hilfreiche Formeln und Abbildungen an die Tafel schreibe. Dann lese ich die Aufgaben vor bzw. fasse kurz zusammen, was zu tun und wie man die Aufgabe in etwa lösen könnte, ohne hier einen vollständigen Lösungsweg zu erklären. Es geht mir bloß darum, einen möglichen Ansatz vorzustellen. Danach lasse ich den Studenten ein paar Minuten Zeit und beginne dann, zu jedem Studenten zu gehen und direkt nachzufragen, ob sie Probleme mit der Aufgabe oder andere Fragen haben. Das hat den Hintergrund, dass viele Studenten sich nicht trauen öffentlich im Tutorium ihre Fragen zu stellen, obwohl viele andere auch die selbe Frage haben. Aber wenn ich die Studenten dann direkt anspreche, kommen auch die Fragen. Je nach Länge der Aufgabe und nach restlicher Zeit im Tutorium schaffe ich es mindestens ein- oder zweimal bei jedem Studenten zu sein; natürlich gehe ich auch zwischen meinen Runden zu Studenten, wenn diese direkt eine Frage haben. Wenn die meisten Studenten fertig sind oder die Zeit weit fortgeschritten ist, stelle ich die Lösung, die ich vor dem Tutorium zu Hause erarbeitet habe, vor. Meistens kann ich die Vorstellung der Lösung recht kurz halten, da alle relevanten Fragen schon in den direkten Gesprächen mit den Studenten geklärt wurden.

Weiterhin betreue ich das Forum, in dem Studenten Fragen stellen können. Die meisten Fragen der Studenten kann ich selber beantworten, nur bei einigen organisatorischen Fragen, frage ich dann Martin Waltz, der dann eine Antwort im Forum verfasst.

Zudem habe ich während der Corona-Pandemie mich um die Erstellung der Online-Klausuren gekümmert. Aus den vorhandenen Aufgaben habe ich Varianten dieser Aufgaben erstellt, mit veränderten Zahlen oder leichten Veränderungen der Aufgabenstellung. Dabei mussten dann natürlich auch die Lösungen angepasst werden. Ich habe mir für diese Aufgabe ein Python-Programm geschrieben, was diese Anpassungen vollautomatisch machen kann und mehrere dieser Aufgaben zu einem Test zusammenbauen kann. Dieser Test kann dann zu OPAL Exam hochgeladen werden und als Klausur genutzt werden. Nach Corona wurden wieder Papierklausuren benötigt und ich habe mein Programm so angepasst, dass es nun auch Papierklausuren mit zufälligen Änderungen an den Aufgabenstellungen und zufälligen Zahlen erzeugen kann.

Während die Klausuren geschrieben werden, bin ich auch Teil der Aufsicht; das umfasst das Austeilen und

Einsammeln der Klausuren, das Vorbereiten des Raumes vor der Klausur und das Kontrollieren, dass keine Betrugsversuche stattfinden. Nachdem die Klausur geschrieben wurde, habe ich auch eine Aufgabe zugeteilt bekommen, die ich korrigiert habe.

## Syllabus des Kurses

Die Inhalte von Statistik 1 sind:

- Einführung in die deskriptive Statistik
- Empirische Verteilungen
- Streu- und Lagemaße
- Korrelation
- Kombinatorik
- Wahrscheinlichkeit von Ereignissen
- Bedingte Wahrscheinlichkeit, Satz von der totalen Wahrscheinlichkeit, Satz von Bayes
- Einführung in die diskreten und stetigen Verteilungen
- Eigenschaften von Erwartungswert und Varianz
- Diskrete Verteilungen (Binomial, Hypergeometrisch, Poisson, Geometrisch)
- Stetige Verteilungen (Gleich, Exponential, Normal, u.a.)

Die Inhalte von Statistik 2 sind:

- Diskrete und stetige Verteilungen
- Quantile
- Ungleichung von Tschebyscheff, Grenzwertsätze
- Parameterschätzung: ML- und KQ-Methoden, Eigenschaften
- Konfidenzintervalle
- Signifikanztests
- Zweistichprobentests
- Unabhängigkeitstests
- Anpassungstests
- Lineare Regression

# Auswahl an Übungsaufgaben

Aus Gründen der Übersichtlichkeit habe ich hier nur die erste Aufgabe eines jeden Übungsblattes angehängt.

## Aufgabe 1.1

Eine Gruppe Studentinnen der TU Dresden untersucht im Rahmen ihrer Seminararbeit das touristische Potenzial der sächsischen Landeshauptstadt. Die Befragung der Dresdner Touristen soll dabei Informationen zu folgenden Punkten liefern:

- |                               |   |
|-------------------------------|---|
| A) Wohnort des Befragten      | F) Nutzung des ÖPNV während des Aufenthalts (z.B. ja) |
| B) Verkehrsmittel der Anreise | G) Alter des Befragten                                |
| C) Aufenthaltsdauer           | H) Nettoeinkommen                                     |
| D) Zahl der Mitreisenden      | I) Tag der Ankunft                                    |
| E) Sternezahl der Unterkunft  |   |

Charakterisieren Sie die einzelnen Merkmale durch folgende Eigenschaften:

- nominal - ordinal - metrisch
- quantitativ - qualitativ
- diskret - stetig

## Aufgabe 2.1

Auf den Verkaufsportalen  $A$ ,  $B$ , und  $C$  wurden jeweils von verschiedenen Anbietern die Preise eines bestimmten Tablet Stiftes ermittelt. Auf Portal  $A$  ergab sich aufgrund von 11 Einzelwerten ein Durchschnittspreis von €15.28. Auf Portal  $B$  betrug der Durchschnittspreis basierend auf 6 Einzelwerten €16.44 und auf  $C$  wurde nach einer Auswahl von 13 Anbietern ein durchschnittlicher Preis von €14.86 ermittelt. Wie hoch ist der Durchschnittspreis des Tablet Stiftes unter Berücksichtigung sämtlicher Preisinformationen aus den Portalen  $A$ ,  $B$  und  $C$ ?

## Aufgabe 3.1

Ein Bausparkassenleiter möchte die Produktivität seiner Mitarbeiter in einem Monat anhand ihrer Vertragsabschlüsse ( $n = 60$ ) überprüfen. Zu diesem Zweck lässt er sich alle Bausparverträge des Monats unter Angabe des Mitarbeiternamens ( $A$ ,  $B$ ,  $C$ ,  $D$  oder  $E$ ) und der jeweiligen Abschlusshöhe (in Tausend €) auflisten:

In der folgenden Tabelle werden die Daten in einem sogenannten Stamm-Blätter-Diagramm dargestellt:

A		B		C		D		E	
1	35567789	1	55669	1	6799	1		1	35
2	113457	2	157	2	158	2		2	39
3	23357	3	03778	3	09	3		3	
4	135	4	02	4	0	4	055	4	
5	02	5		5		5	7	5	
6		6		6		6	005	6	
24		15		10		7		4	

### Aufgabe 4.1

Bestimmen Sie für die Daten

15	70	-3	13	-60	50	8	6	-3	-16
----	----	----	----	-----	----	---	---	----	-----

die Streuungs- und Lagemaße:

- (a) Mittelwert;
- (b) Oberes und unteres Quartil, Median;
- (c) Spannweite;
- (d) Quartilabstand;
- (e) Median der absoluten Abweichungen vom Median (MAD);
- (f) Empirische Varianz, empirische Standardabweichung, Stichprobenvarianz, Stichprobenstandardabweichung;
- (g) Stichprobenschiefe;
- (h) Zeichnen Sie den zugehörigen BOX-Plot.

### Aufgabe 5.1

Johannes fährt jeden Tag mit dem Fahrrad zur Uni. Eine Woche lang schreibt er seine Fahrtzeiten für Hin- und Rückweg in Minuten auf.

$i$	1	2	3	4	5
Hinweg $x$	28	29	25	30	28
Rückweg $y$	30	31	28	29	22

- Berechnen Sie Mittelwert und Varianz des Hinweges  $x$ .
- Berechnen Sie Mittelwert und Varianz des Rückweges  $y$ .
- Johannes möchte sich ein Rennrad kaufen. Er vermutet, dass seine Fahrtzeit auf dem Hinweg um den Faktor 0.95 kürzer wird und auf dem Rückweg sogar um 0.9. Allerdings wird er täglich 2 Minuten länger brauchen, um das Fahrrad abzuschließen. Berechnen Sie Mittelwert und Varianz der gesamten Fahrzeit  $z = 0.95 \cdot x + 0.9 \cdot y + 2$ . Nehmen Sie dabei an, dass kein Zusammenhang zwischen  $x$  und  $y$  besteht.
- Berechnen Sie Mittelwert und Varianz von  $x_{\text{zentriert}} = x - \bar{x}$ .
- Berechnen Sie Mittelwert und Varianz von  $x_{\text{standardisiert}} = \frac{x - \bar{x}}{s_x}$ .
- Berechnen Sie die Kovarianz zwischen  $x$  und  $y$ .
- Berechnen Sie die Varianz von  $z$  unter Berücksichtigung des Zusammenhangs zwischen  $x$  und  $y$ .

## Aufgabe 6.1

Um die Frage zu beantworten, ob es einen Zusammenhang zwischen der Haarfarbe eines Menschen und seiner Augenfarbe gibt, wurden für 200 Studenten, darunter 42 Blonde, die folgenden Daten erhoben:

Augenfarbe	Haarfarbe			
	blond	braun	rot	schwarz
blau		28	6	7
braun	5	58	14	29
grün	6	9		2

- Füllen Sie die fehlenden Stellen in der Tabelle aus und erstellen Sie die Kontingenztafel der relativen Häufigkeiten sowie die relativen Randhäufigkeiten!
- Überprüfen Sie, ob es Abhängigkeiten zwischen Haar- und Augenfarbe gibt. Besteht eine starke Abhängigkeit?
- Man untersucht die Augenfarbe der **blonden** Studenten. Bestimmen Sie die bedingten relativen Häufigkeiten für die Augenfarbe einer Person unter der Bedingung, dass die Person blond ist.

### Aufgabe 7.1

- (a) Wie viele Möglichkeiten gibt es, die Buchstaben des englischen Wortes TEA anzuordnen, wobei jeder einzelne nur einmal vorkommen darf?
- (b) Wie viele Möglichkeiten gibt es, die Buchstaben des deutschen Wortes TEE anzuordnen, wobei jeder einzelne nur einmal vorkommen darf?
- (c) Wie viele Möglichkeiten gibt es, die Buchstaben des Wortes STATISTIK anzuordnen, wobei jeder einzelne nur einmal vorkommen darf?

### Aufgabe 8.1

Geben Sie für die folgenden Ereignisoperationen die Ergebnisse an.

$\Omega$  - Grundgesamtheit,  $A, B \subset \Omega$ , und  $B \subset A$ :

$A \cup A, A \cap A, A \cup \emptyset, A \cap \emptyset, \emptyset \cap \Omega, A \cup \Omega, A \cap \Omega, A \cup \bar{A}, A \cap \bar{A}$

$A \cup B, A \cap B, A \cup \bar{B}, A \cap \bar{B}, \bar{A} \cap B, \bar{A} \cup \bar{B}, \bar{A} \cap \bar{B}$

### Aufgabe 9.1

Karl liebt den Alkohol. Die Wahrscheinlichkeit, dass er nach Büroschluss trinkt, ist 0.8. Karl ist vergesslich. Die Wahrscheinlichkeit, dass er seinen Schirm stehen lässt, ist 0.7. Wenn er getrunken hat, ist die Wahrscheinlichkeit sogar 0.8. Karl kommt ohne Schirm nach Hause.

Wie groß ist die Wahrscheinlichkeit dafür, dass er dieses Mal nicht getrunken hat?

### Aufgabe 10.1

Gegeben ist folgende Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & \text{für } x < 1; \\ 0.2 & \text{für } 1 \leq x < 3; \\ 0.5 & \text{für } 3 \leq x < 7; \\ 1 & \text{für } 7 \leq x. \end{cases}$$

- (a) Wie lautet die zugehörige Wahrscheinlichkeitsfunktion? Bitte graphisch darstellen.
- (b) Bestimmen Sie die Wahrscheinlichkeiten  $\mathbb{P}(1 \leq X < 3)$ ,  $\mathbb{P}(0 \leq X < 3)$ ,  $\mathbb{P}(1 \leq X < 4)$  und  $\mathbb{P}(1 < X < 7)$ !
- (c) Berechnen Sie den Erwartungswert und die Varianz!

### Aufgabe 11.1

Gegeben ist folgende Dichtefunktion einer Zufallsvariablen  $Y$ ?

$$f(y) = \begin{cases} \frac{1}{12} y & \text{für } 1 \leq y \leq 5; \\ 0 & \text{sonst.} \end{cases}$$

- (a) Bestimmen Sie die Verteilungsfunktion dieser Zufallsvariablen!
- (b) Bestimmen Sie die Wahrscheinlichkeit des Ereignisses  $(2 < Y < 3)$ !

### Aufgabe 12.1

Aus langjähriger Erfahrung ist der Abteilung für Familie und Soziales einer Stadtverwaltung bekannt, wie sich die gemeinsame Verteilung der Anzahl der Kinder pro Familie  $X_1$  und der Anzahl der PKW pro Familie  $X_2$  zusammensetzt. Die gemeinsame Wahrscheinlichkeitsfunktion von  $X_1$  und  $X_2$  ist gegeben durch:

Anzahl der Kinder, $X_1$	Anzahl der PKW, $X_2$		
	1	2	3
0	0.08	0.28	0.01
1	0.10	0.14	0.06
2	0.07	0.05	0.06
3	0.06	0.03	0.01
4	0.04	0.01	0

Bestimmen Sie

- (a) die Randwahrscheinlichkeiten von  $X_1$  und  $X_2$ ;
- (b) die Randverteilungsfunktion von  $X_1$ ;
- (c) den Erwartungswert für die Anzahl der Kinder bzw. PKW pro Familie;
- (d) die Wahrscheinlichkeit, dass es pro Familie genau zwei PKW und höchstens zwei Kinder gibt;
- (e) die (bedingte) Wahrscheinlichkeitsfunktion der Anzahl der PKW für die Familien ohne Kinder;
- (f) die (bedingte) erwartete Anzahl der PKW pro Familie ohne Kinder.
- (g) die Kovarianz und die Korrelation zwischen  $X_1$  und  $X_2$ .

### Aufgabe 13.1

Nur 3 % der kontrollierten Fahrgäste der Dresdner Straßenbahnen besitzen keinen gültigen Fahrschein. Die Ergebnisse der Kontrolle seien unabhängig, d.h. es wird vereinfacht angenommen, dass nur Einzelpersonen kontrolliert werden.

- (a) Mit welcher Wahrscheinlichkeit wird bei der Kontrolle von 10 Personen wenigstens ein "Schwarzfahrer" erwischt?
- (b) Wie groß ist die Wahrscheinlichkeit, dass der Kontrolleur erst bei der zehnten Kontrolle den ersten "Schwarzfahrer" aufspürt?