Aufgaben zur Einführung in die Statistik

Andreas Handl

Torben Kuhlenkasper

8. Januar 2016

Grundlage des vorliegenden Skripts sind Aufzeichnungen von Andreas Handl, die er bis zum Jahr 2007 an der Universität Bielefeld verfasst und für seine Lehrveranstaltungen verwendet hat. Seit 2012 werden die Skripten von Torben Kuhlenkasper weitergeführt sowie fortlaufend aktualisiert und erweitert.

Anmerkungen und Vorschläge zur Verbesserung und Ergänzung sind jederzeit willkommen und können an statistik@kuhlenkasper.de *gesendet werden. Weitere Skripten sind unter* www.skripten.kuhlenkasper.de *zu finden.*

Inhaltsverzeichnis

Kapitel 1

Grundbegriffe

Aufgabe 1
Sie wollen herausfinden, ob Masterstudenten an einem Fachbereich für Wirtschaftswissenschaften mit ihrem Studium zufrieden sind. Erstellen Sie einen geeigneten Fragebogen. Welches Skalenniveau besitzen die Merkmale?

Aufgabe 2
Es soll untersucht werden, ob sich der tägliche Alkoholkonsum von Frauen und Männern unterscheidet. Welche Probleme hat man, wenn man die Personen nach ihrem täglichen Alkoholkonsum befragt? Ist es möglich, die Daten durch Beobachtung zu gewinnen?

Aufgabe 3
Ein Student kann auf zwei Wegen zur Universität gelangen. Er will herausfinden, auf welchem Weg er schneller zur Universität kommt. Wodurch unterscheidet sich die Datengewinnung durch Beobachtung von der Datengewinnung durch Experiment?

Aufgabe 4
Geben Sie an, ob die folgenden Variablen diskret oder stetig sind und welches Skalenniveau sie besitzen:

Familienstand, Religion, Steuerklasse, Alter, Stückzahl in der Produktion, Reaktionszeit, Güteklasse von Restaurants, Fahrzeit, Körpergewicht, Geschlecht, Höhe der Miete, Gewichtsklasse von Eiern

Kapitel 2

Univariate Analyse

Aufgabe 5
Bei der Weltmeisterschaft 2002 im Fußball in Japan und Südkorea fanden 64 Spiele statt.
Im folgenden finden Sie eine Urliste der Anzahl der Tore je Spiel:

```
1 2 3 8 1 4 2 4 1 3 2 2 4 2 2 5
2 2 1 0 3 4 1 1 3 4 3 2 1 2 2 4
6 2 2 3 0 2 4 5 3 7 1 2 2 5 4 1
1 3 3 2 2 2 1 3 3 1 0 1 1 1 5 2
```

1. Erstellen Sie die Häufigkeitstabelle und interpretieren Sie diese.

2. Zeichnen und interpretieren Sie das Stabdiagramm!

3. Zeichnen Sie die empirische Verteilungsfunktion.

4. Wie groß ist der Anteil der Spiele, in denen genau 3 Tore geschossen wurden?

5. Wie groß ist der Anteil der Spiele, in denen mindestens ein Tor geschossen wurde?

Aufgabe 6
Im Wintersemester 2000/2001 beantworteten 270 Studierende die Frage nach ihrer Note in Mathematik im Abitur. In der folgenden Tabelle sind die absoluten Häufigkeiten der einzelnen Noten zu finden.

Note	Anzahl
1	27
2	89
3	93
4	45
5	16

1. Bestimmen Sie die relativen Häufigkeiten.
2. Zeichnen und interpretieren Sie das Stabdiagramm!
3. Zeichnen Sie die empirische Verteilungsfunktion.
4. Wie groß ist der Anteil der Studierenden, deren Abiturnote in Mathematik mindestens gut ist?

Aufgabe 7
Im Begleitheft der ersten Fünf Freunde CD-ROM finden sich Lösungshinweise, die aber kodiert sind. So findet sich unter *Wie komme ich auf die Insel?* folgender Text:

```
Yq hmi Mrwip fixvixir dy osirrir, fveyglwx hy jspkirhi
Kikirwxeirhi: Imri Ebx yrh hew Wimp eyw hiq Wglyttir yrh hmi
Vyhiv eyw hiq Zivwxigo zsr Kisvki. Eywwivhiq qeglx iwivwx Wmrr,
hmi Mrwip dy ivjsvwglir, airr hy hir Wglexdtper irxhigox lewx.
```

Jedem Buchstaben des Alphabets wurde ein anderer Buchstabe des Alphabets zugeordnet. Um den Text zu entschlüsseln, benötigen wir die Häufigkeit der Buchstaben in einem deutschsprachigen Text. Diese ist in der Tabelle 2.1 zu finden. Die Angaben sind in Prozent.

Tabelle 2.1: Häufigkeitsveretilung der Buchstaben in der deutschen Sprache(in Prozent)

a	6.51	e	17.4	i	7.55	m	2.53	q	0.09	u	4.35	y	0.04
b	1.89	f	1.66	j	0.27	n	9.78	r	7.00	v	0.67	z	1.13
c	3.06	g	3.01	k	1.21	o	2.51	s	7.27	w	1.89		
d	5.08	h	4.76	l	3.44	p	0.79	t	6.15	x	0.03		

In der Lösungshilfe steht, dass der Text mit Hilfe einer Cäsar-Verschlüsselung kodiert worden ist. Bei dieser wird jeder Buchstabe durch den Buchstaben verschlüsselt, der eine bestimmte Anzahl von Buchstaben hinter ihm steht. Wird also zum Beispiel das `a` durch das `h` verschlüsselt, so wird das `b` durch das `i` verschlüsselt, das `c` durch das `j`, u.s.w..

Entschlüsseln Sie mit Hilfe dieser Informationen den Text.

Aufgabe 8
Im Sommersemester 2000 fanden 8 Klausuren im Grundstudium statt. In der folgenden Tabelle finden sie die Notenverteilung der Klausuren beim ersten Termin:

Note Fach	1	2	3	4	5
BWL II	5	20	30	56	88
VWL II	8	21	37	19	26
Mathematik II	1	21	45	94	63
Statistik II	14	34	50	49	32
BWL IV	4	18	40	39	69
VWL IV	1	12	24	42	38
Informatik	9	14	29	27	87
Unternehmensforschung	0	1	61	54	69

1. Erstellen für jedes Fach ein Stabdiagramm.
2. Erstellen Sie die empirische Verteilungsfunktion der Noten im Fach VWL II.

Aufgabe 9
Im Wintersemester 1995/1996 hatten von den 250 Studienanfängern 139 den Leistungskurs Mathematik absolviert. Im Wintersemester 2000/2001 hatten von den 298 Studienanfängern 182 den Leistungskurs Mathematik absolviert. Erstellen Sie ein vergleichendes Paretodiagramm, wobei sie das Wintersemester 1995/1996 als Basis wählen sollten.

Aufgabe 10
Im Wintersemester 1995/1996 waren von den 250 Studienanfängern 69 weiblich, während im Wintersemester 2000/2001 von den 298 Studienanfängern 112 weiblich waren. Erstellen Sie ein vergleichendes Paretodiagramm, wobei sie das Wintersemester 1995/1996 als Basis wählen sollten.

Aufgabe 11
Im Sommersemester 2002 fanden 8 Klausuren im Grundstudium statt. In der folgenden Tabelle finden sie die Notenverteilung der Klausuren beim ersten Termin:

Note Fach	1	2	3	4	5
BWL II	23	60	60	63	103
VWL II	1	0	9	11	103
Mathematik II	2	31	49	86	78
Statistik II	13	19	24	26	63
BWL IV	3	21	41	42	75
VWL IV	4	48	45	45	56
Informatik	7	18	62	64	89
Unternehmensforschung	18	20	34	49	94

1. Erstellen für jedes Fach ein Stabdiagramm.
2. Betrachten Sie auch die Noten im Sommersemester 2000 in Aufgabe 8 auf Seite 4. Vergleichen Sie Verteilung der Noten im Fach VWL II in den beiden Semestern mit Hilfe eines vergleichenden Paretodiagramms.

Aufgabe 12
Von den in der ersten Vorlesung STATISTIK I im WS 1995/1996 abgegebenen Fragebögen wurden 20 der Männer und 20 der Frauen zufällig ausgewählt.
In der folgenden Liste ist das Alter angegeben.

Frauen:

```
20 21 21 21 21 21 22 22 19 19 23 21 22 23 19 23 22 23 19 22
```

Männer:

```
21 21 21 22 21 27 22 23 20 22 21 23 22 22 20 20 20 20 20 23
```

1. Erstellen Sie eine Häufigkeitsverteilung aller 40 Beobachtungen.
2. Zeichnen und interpretieren Sie das Stabdiagramm aller 40 Beobachtungen.
3. Zeichnen Sie die empirische Verteilungsfunktion aller 40 Beobachtungen.

4. Wie groß ist der Anteil der Personen, die höchstens 22 Jahre alt sind?

5. Erstellen Sie eine Häufigkeitstabelle für das Alter der Männer und eine für das Alter der Frauen.

Aufgabe 13
Wir betrachten die letzten drei Spieltage der Saison 2001/2002 in der Fußballbundesliga. Von den 27 Spielen wurden 25 ausgewählt und bei jedem Spiel die Zeit ermittelt, die man auf das erste Tor warten musste. Die Urliste sieht folgendermaßen aus:

```
67 54  5 21  4 48 47 33 20 25 26 36 5
23 50  9  6 30 10 15 17 26 35 42 39
```

Es werden folgende Klassen gebildet:

1.Klasse :	von 0	bis unter 15
2. Klasse :	von 15	bis unter 30
3. Klasse :	von 30	bis unter 45
4. Klasse :	von 45	bis unter 60
5. Klasse :	von 60	bis unter 75
6. Klasse :	von 75	bis unter 90

1. Erstellen Sie eine Häufigkeitstabelle.

2. Zeichnen und interpretieren Sie das Histogramm!

3. Zeichnen Sie die empirische Verteilungsfunktion!

4. Bestimmen Sie $\hat{F}(40)$!

5. Bestimmen Sie mit Hilfe der empirischen Verteilungsfunktion den Anteil der Spiele, bei denen man mehr als 30 aber höchstens 60 Minuten auf das erste Tor warten musste!

Aufgabe 14
Von den Fragebögen, die im Wintersemester 2000/2001 in der ersten Vorlesung zur Statistik I abgegeben wurden, wurden 50 zufällig ausgewählt und die Durchchnittsnote im Abitur festgehalten. Es ergibt sich folgende Urliste:

```
2.1 2.1 2.2 2.5 3.0 2.0 2.2 2.4 2.9 2.4
3.6 3.1 3.2 2.3 3.3 1.8 2.7 3.7 2.9 2.9
3.2 3.4 2.3 2.4 3.0 2.0 2.6 3.5 3.2 3.0
3.4 2.6 2.5 1.5 1.5 3.0 2.0 2.5 2.9 2.8
1.6 2.0 2.6 2.1 3.2 3.0 3.5 2.4 1.5 3.3
```

Es werden folgende Klassen gebildet:

1. Klasse :	von 1.0 bis unter 1.5
2. Klasse :	von 1.5 bis unter 2.0
3. Klasse :	von 2.0 bis unter 2.5
4. Klasse :	von 2.5 bis unter 3.0
5. Klasse :	von 3.0 bis unter 3.5
6. Klasse :	von 3.5 bis unter 4.0

1. Erstellen Sie eine Häufigkeitstabelle.

2. Zeichnen und interpretieren Sie das Histogramm!

3. Zeichnen Sie die empirische Verteilungsfunktion!

4. Bestimmen Sie $\hat{F}(2.2)$!

Aufgabe 15

In der Neuen Westfälischen vom 14.8.1996 erschienen 20 Anzeigen für Einzimmerwohnungen, bei denen die Kaltmiete (in DM) angegeben wurde. Diese und die Fläche der Wohnungen (in m^2) sind in den folgenden Tabellen angegeben:

Fläche	45	21	34	44	43	23	25	29	32	33
Preis	690	370	490	580	650	460	460	479	580	490

Fläche	65	44	35	25	42	30	43	22	34	37
Preis	660	560	500	383	580	400	650	580	522	469

1. Zunächst analysieren wir die Kaltmiete.

   Wählen Sie eine Klassenbreite von 70 und bilden Sie 5 Klassen, wobei die erste Klasse bei 350 beginnt.

   (a) Erstellen Sie die Häufigkeitstabelle!
   (b) Zeichnen und interpretieren Sie das Histogramm!
   (c) Zeichnen Sie die empirische Verteilungsfunktion!
   (d) Bestimmen Sie $\hat{F}(500)$!

2. Und nun zur Kaltmiete pro m^2.

   Bilden Sie 8 Klassen mit der Klassenbreite 2.50, wobei Sie bei 10.00 beginnen.

(a) Erstellen Sie die Häufigkeitstabelle!

(b) Zeichnen und interpretieren Sie das Histogramm!

(c) Zeichnen Sie die empirische Verteilungsfunktion!

(d) Wie groß ist $\hat{F}(15)$?

Aufgabe 16
Die Firma MFSL stellte bis Anfang des Jahres 2000 audiophile CDs mit Goldbeschichtung her. Nachdem die Firma Konkurs anmelden musste, wurden diese CDs zu begehrten Sammlerstücken. Ein Statistiker beschließt, seine MFSL GOLD-CD von DARK SIDE OF THE MOON von PINK FLOYD auf EBAY zu versteigern. Um eine Vorstellung vom realisierbaren Preis zu erhalten, beobachtet er den Markt. In der zweiten Februarwoche des Jahres 2001 wurden 9 CDs zu folgenden Höchstgeboten in Dollar ersteigert:

```
51  56  57  48  49  61  46  50  59
```

1. Bestimmen Sie den Mittelwert und den Median der Höchstgebote.
2. Bestimmen Sie die Spannweite und die Stichprobenvarianz der Höchstgebote.
3. Erstellen und interpretieren Sie den Boxplot der Daten.

Aufgabe 17
Ein Statistiker sieht sich in der Zeit vom 2.2.2001 bis zum 19.2.2001 alle Folgen der Show WER WIRD MILLIONÄR an und notiert sich am Ende der Sendung den realisierten Gesamtgewinn des Tages. Es ergaben sich folgende Werte (in Tausend DM):

```
34  17  96  33  189  282  33  66  64
```

1. Bestimmen Sie den Mittelwert und den Median.
2. Bestimmen Sie die Spannweite und die Stichprobenvarianz.
3. Erstellen und interpretieren Sie den Boxplot.

Aufgabe 18
Wir betrachten noch einmal die Daten aus Aufgabe 12 auf Seite 5.

1. Bestimmen Sie den Mittelwert und den Median des Alters der Frauen.
2. Bestimmen Sie den Mittelwert und den Median des Alters der Männer.

3. Bestimmen Sie den Mittelwert und den Median des Alters aller Personen.

4. Erstellen Sie den Boxplot des Alters der Frauen und den Boxplot des Alters der Männer.

   Auf welche Unterschiede zwischen den beiden Verteilungen deuten die Boxplots hin?

Aufgabe 19
Wir betrachten noch einmal den Datensatz und die Klasseneinteilung aus Aufgabe 13 auf Seite 6.

1. Welche Wartezeit bis zum ersten Tor wurde in 25 Prozent der Spiele nicht überschritten?

   Bestimmen Sie diesen Wert

   - mit Hilfe der empirischen Verteilungsfunktion,
   - aus den Rohdaten.

2. Bestimmen Sie aus den Rohdaten

   - das untere Quartil,
   - das obere Quartil,
   - den Median.

3. Zeichnen und interpretieren Sie den Boxplot.

4. Berechnen Sie den Mittelwert aus den Rohdaten.

5. Berücksichtigen Sie den Mittelwert in Ihrem Boxplot.

6. Berechnen Sie die Stichprobenvarianz s^2.

Aufgabe 20
Wir betrachten noch einmal den Datensatz und die Klasseneinteilung aus Aufgabe 14 auf Seite 6.

1. Welche Note wurde von 25 Prozent der Studierenden nicht überschritten?

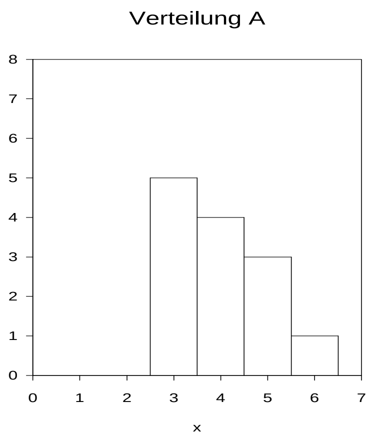
   Bestimmen Sie diesen Wert

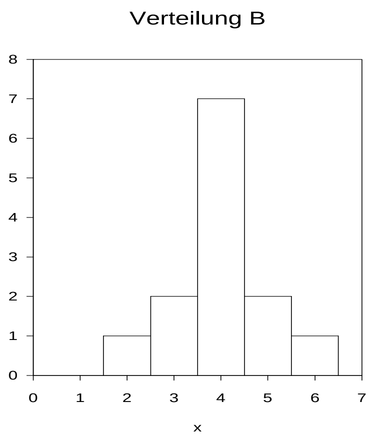
   - mit Hilfe der empirischen Verteilungsfunktion,
   - aus den Rohdaten.

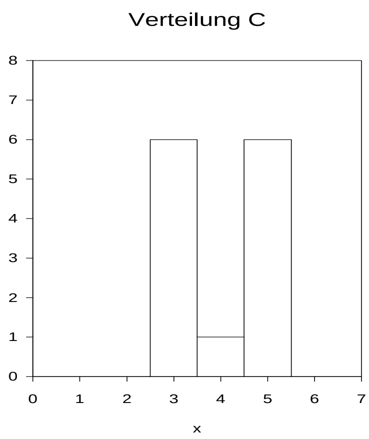
2. Bestimmen Sie aus den Rohdaten
   - das untere Quartil,
   - das obere Quartil,
   - den Median.
3. Zeichnen und interpretieren Sie den Boxplot.
4. Berechnen Sie den Mittelwert aus den Rohdaten.
5. Berücksichtigen Sie den Mittelwert in Ihrem Boxplot.
6. Berechnen Sie die Stichprobenvarianz s^2.

Aufgabe 21

Die folgenden Histogramme sind im Artikel *Variability- Does the standard deviation always measure it adequately* von *Louis A. Pingel* in der Zeitschrift *Teaching Statistics* 1993 auf der Seite 70 zu finden.







1. Berechnen Sie für jeden der drei Fälle den Mittelwert und den Median.

2. Begründen Sie, warum der Mittelwert und der Median in manchen Fällen identisch, in anderen hingegen unterschiedlich sind.

3. Sind Mittelwert und Median in Fall C geeignete Maßzahlen der Lage?

4. Berechnen Sie für alle drei Fälle die Spannweite und die Stichprobenvarianz.

5. Zeichnen Sie für jeden Fall den zugehörigen Boxplot.

Aufgabe 22
Wir betrachten noch einmal die Aufgaben 8 und 11 auf den Seiten 4 und 4. Vergleichen Sie für beide Jahre den Mittelwert der Noten der Studierenden in den einzelnen Fächern.

Aufgabe 23
In der folgenden Tabelle ist die Regierungszeit (in Monaten) der ersten sechs Bundeskanzler der Bundesrepublik Deutschland zu finden.

Name	Regierungszeit
Konrad Adenauer	169
Ludwig Erhard	37
Kurt Georg Kiesinger	35
Willy Brandt	54
Helmut Schmidt	100
Helmut Kohl	193

1. Erstellen und interpretieren Sie den Boxplot der Daten.

2. Berechnen Sie den Mittelwert und den Median.

   Welche der beiden Maßzahlen halten Sie für geeigneter zur Beschreibung des Datensatzes?

Aufgabe 24
Vom 6.8.2002 bis zum 15.8.2002 wurde die Tageshöchsttemperatur (in Celsius) in Bielefeld bestimmt. Es ergaben sich folgende Werte:

```
23 21 19 24 19 23 22 20 23 27
```

1. Bestimmen Sie den Mittelwert der Temperatur.

2. Wir hätten die Temperatur auch in Fahrenheit angeben können.

   Zwischen der Temperatur in Celsius x und der Temperatur in Fahrenheit y besteht folgender Zusammenhang:

   $$y = 32 + 1.8 \cdot x$$

   Berechnen Sie den Mittelwert der Temperatur in Fahrenheit.

   Hinweis:

   Seien $x_1, \ldots, x_n$ reelle Zahlen.

   Sei

   $$y_i = a + b \cdot x_i$$

   für $i = 1, \ldots, n$.

   Zeigen Sie, dass dann gilt:

   $$\bar{y} = a + b \cdot \bar{x}$$

3. Bestimmen Sie den Median der Temperatur in Celsius und in Fahrenheit.

Aufgabe 25

Seien $x_1, \ldots, x_n$ reelle Zahlen.

Für welchen Wert a nimmt folgende Funktion $f(a)$ ihr Minimum an?

$$f(a) = \sum_{i=1}^{n} (x_i - a)^2$$

Interpretieren Sie das Ergebnis.

Hinweis:

Sie können die Differentialrechnung benutzen. Sie können aber auch folgende Beziehung benutzen:

$$\sum_{i=1}^{n} (x_i - a)^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 + n\,(\bar{x} - a)^2$$

Diese folgt nach einigen Umformungen aus

$$\sum_{i=1}^{n} (x_i - a)^2 = \sum_{i=1}^{n} (x_i - \bar{x} + \bar{x} - a)^2$$

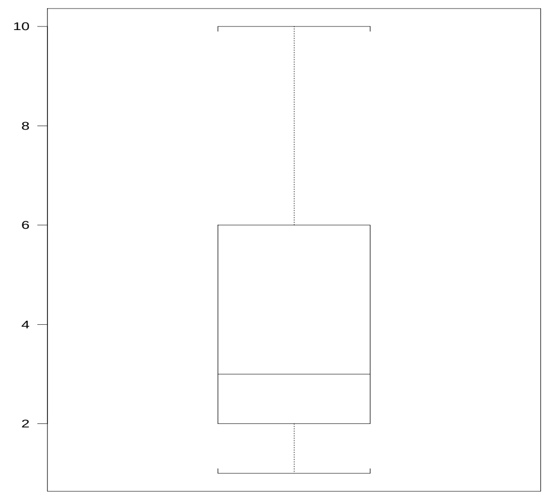
Führen Sie auch diese Umformungen durch.

Aufgabe 26
Gegeben sei folgender Datensatz:

`2 1 18 3 6`

1. Berechnen Sie für diesen Datensatz den Mittelwert und den Median.
2. Es stellt sich heraus, dass in obigem Datensatz keine 18, sondern 8 stehen sollte. Korrigieren Sie den Fehler und bestimmen Sie jetzt den Mittelwert und den Median.
3. Wie lauten der Mittelwert und Median, wenn Sie die 18 aus dem Datensatz herauslassen?
4. Man sagt, dass der Mittelwert ausreißerempfindlich ist. Das heißt, dass eine extreme Beobachtung einen großen Einfluss auf den Wert des Mittelwerts hat. Gilt dies auch für den Median?

Aufgabe 27
Gegeben sei folgender Boxplot eines Datensatzes.



1. Beschreiben Sie, wie sich der Boxplot ändert, wenn Sie zu jeder Beobachtung in der zugrundeliegenden Stichprobe den Wert 1 addieren.
2. Beschreiben Sie, wie sich der Boxplot ändert, wenn Sie jede Beobachtung in der zugrundeliegenden Stichprobe mit -1 multiplizieren.

Aufgabe 28
Gegeben sei ein Datensatz, der aus 20 Beobachtungen besteht. Der Mittelwert beträgt 10, der Median 9 und die Stichprobenvarianz 4. Wir multiplizieren jede Beobachtung in der Stichprobe mit 4. Wie groß ist der Mittelwert, der Median, die Stichprobenvarianz und die mittlere quadratische Abweichung der so gewonnenen Beobachtungen? Begründen Sie Ihre Aussage.

Kapitel 3

Bivariate Analyse

Aufgabe 29
Die Richterin Bettina Freimund analysierte die Gefangenenakten von 103 Wirtschaftsstraftätern und 113 alternden Dieben, um herauszufinden, 'wie es um den Resozialisierungsgedanken im Strafvollzug bestellt ist'. Sie hatte diese Gruppen ausgesucht, weil 'sie sich gleichermaßen Delikte gegen das Vermögen' hatten zuschulden kommen lassen und zugleich 'absolute Gegenpole bezüglich ihrer Stellung in der Gesellschaft' repräsentierten.
Ein Ergebnis der Untersuchung war, dass von den 216 Personen 109 Freigänger waren, wobei 68 aus der Gruppe der Wirtschaftskriminellen kamen. (Quelle: DER SPIEGEL 43/1991, S.85-92)

1. Stellen Sie eine Kontingenztabelle auf!
2. Bestimmen Sie die korrespondierenden bedingten relativen Häufigkeiten, interpretieren Sie diese und stellen Sie sie graphisch dar.
3. Erstellen Sie das vergleichende Paretodiagramm, wobei Sie die Wirtschaftsstraftäter als Basis wählen sollten.

Aufgabe 30
Der folgende Artikel war in der SÜDDEUTSCHEN ZEITUNG vom 8.8.1996 zu finden:

> Zahlenspiele um BSE-Prozente
>
> Zum Übertragungsrisiko von Kuh zu Kalb
>
> Von den 'bab's' gab es in Großbritannien bisher 28400 Fälle. Das Kürzel steht für 'born after ban'. Gemeint damit sind Rinder, die nach dem Verbot von Proteinfutter geboren sind, also nach 1988,

und trotzdem an der Seuche BSE erkrankten. Sie hatten, so die Vermutung britischer Experten, verseuchtes Futter bekommen. Doch nun ist sicher, dass manche Tiere sich bei ihren Müttern angesteckt haben.

Mit einem Risiko von einem Prozent erkrankt ein Kalb an der Seuche BSE, wenn seine Mutter dem Rinderwahnsinn erlegen ist. Das folgert das britische Landwirtschaftsministerium aus einer Studie, die 1989 begonnen wurde. Der Versuch selbst hat allerdings ein zehnfach höheres Risiko ergeben. Die englischen Experten argumentieren, dass ihre Experimente nicht die natürlichen Bedingungen widerspiegeln. Das Herunterrechnen des Risikos von zehn auf ein Prozent beruht jedoch auf Annahmen, die keineswegs gesichert sind, sondern vielmehr ein optimistisches Szenario darstellen.

Während der Studie hat man zwei Gruppen von jeweils 300 Kälbern beobachtet. Die eine bestand aus Tieren, deren Mütter an BSE erkrankt waren; die Kälber der anderen Kontroll-Gruppe stammten von gesunden Kühen ab. Von letzteren, 273 bislang geschlachteten Kälbern zeigten 13 Symptome der Krankheit. Dass es in dieser Gruppe überhaupt BSE-Fälle gibt, erklären die Experten damit, dass die Tiere vor Beginn der Studie verseuchtes Tiermehl gefressen hatten. Auch in der anderen Gruppe wurden 273 Tiere geschlachtet, von denen 42 an BSE erkrankt waren. Unter der Annahme, dass die Kälber der BSE-Mütter in gleichem Maße infiziertes Futter bekamen, errechneten die Forscher aus dem Vergleich der beiden Gruppen ein Risiko von 10 Prozent, dass der Erreger von Mutter zu Kalb überspringt.

Die Kälber der BSE-Gruppe sind alle im Verlauf von 13 Monaten vor dem Krankheitsausbruch bei der Mutter geboren, die meisten bis zu fünf Monate zuvor. Daraus schließen die englischen Experten, dass das Risiko einer Übertragung abnimmt, je größer die Zeitspanne zwischen Geburt des Nachwuchses und Krankheitsausbruch bei der Mutter ist.

Kritisch ist jedoch nach Meinung von Martin Groschup die Annahme, dass das Risiko einer Übertragung gleich Null ist für Kälber, die mehr als sechs Monate vor dem Krankheitsausbruch der Mutter geboren wurden. 'Es ist unwahrscheinlich, dass das Übertragungsrisiko plötzlich ansteigt', so der Experte von der Bundesanstalt für Viruskrankheiten der Tiere in Tübingen.

Trotzdem ziehen die britischen Fachleute folgendes Fazit: Weil die Inkubationszeit für den Rinderwahnsinn im Durchschnitt 60 Monate dauert, ist das Risiko unter natürlichen Bedingungen zehnmal geringer als in der Studie ermittelt, also ein Prozent. Sollten sich ihre Annahmen als falsch erweisen, könnte die Wahrscheinlichkeit einer Übertragung höher sein, auch weil die Ergebnisse der Studie mit einer statistischen Unsicherheit behaftet sind.

Unklar ist, wie die Jungtiere sich anstecken. In der Plazenta und im Blut sind Prionen, die potentiellen BSE-Erreger sind bisher nicht nachzuweisen. Eine Infektion über die Milch ist wenig wahrscheinlich, da die Kälber in der Regel nicht von der Mutter gesäugt werden, sondern nur die Vormilch erhalten. Auch diese gilt bislang nicht als infektiös. Möglicherweise stecken sich bereits Embryonen an. Zur Zeit läuft in England eine Studie, bei der man Embryonen von BSE-Tieren in neuseeländische Kühe eingepflanzt hat. Ergebnisse wird es wegen der langen Inkubationszeit aber erst in ein paar Jahren geben.

Während die Experten in Brüssel darüber debattieren, ob und wie viele Kälber man möglicherweise als Konsequenz der Studie zusätzlich schlachten muss, halten englische Wissenschaftler die Sorge für übertrieben. Die Zahl der BSE-Fälle sei seit 1993 von 1000 wöchentlich auf 200 gefallen, und sie werde weiter rapide abnehmen.

'Selbst wenn es ein Übertragungsrisiko von Kuh zu Kalb von 10 Prozent gäbe, würde nach zwei Generationen die Zahl der BSE-Tiere von 100 auf eins gesunken sein', sagte Kevin Taylor, Englands stellvetretender Chef-Tierarzt. Weniger positiv ausgedrückt heißt das aber auch, dass die Krankheit schwieriger als bislang vermutet auszurotten ist und über Generationen weitergereicht werden kann.

1. Erstellen Sie eine Kontingenztabelle mit den Variablen 'Mutter an BSE erkrankt' und 'Kalb an BSE erkrankt'.

2. Bestimmen Sie die interessierenden bedingten relativen Häufigkeiten.

3. Versuchen Sie nachzuvollziehen, wie das Risiko von 10 Prozent, dass der Erreger von Mutter zu Kalb überspringt, zustande kommt.

Aufgabe 31

Von den Passagieren auf der Titanic waren 337 in der ersten Klasse, 285 in der zweiten Klasse und 721 in der dritten Klasse. Es waren 885 Besatzungsmitglieder an Bord.

Von den Passagieren der ersten Klasse wurden nach dem Unglück 135 vermißt, von denen der zweiten Klasse 160, von denen der dritten Klasse 541 und von der Besatzung 674.

1. Erstellen Sie eine Kontingenztabelle.
2. Bestimmen Sie die korrespondierenden bedingten relativen Häufigkeiten, interpretieren Sie diese und stellen Sie sie graphisch dar.
3. Erstellen Sie das vergleichende Paretodiagramm, wobei Sie die erste Klasse als Basis wählen sollten.

Aufgabe 32

Es soll untersucht werden, ob der Zustand eines Patienten nach einer Operation vom Krankenhaus abhängt, in dem er operiert wird. Außerdem soll noch der gesundheitliche Zustand des Patienten bei der Einlieferung ins Krankenhaus berücksichtigt werden. Es wurden die Krankenhäuser A und B betrachtet. Es ergaben sich folgende Daten:

Für Patienten mit gutem Zustand bei der Einlieferung:

	Krankenhaus A	Krankenhaus B
gestorben	6	8
überlebt	594	592

Für Patienten mit schlechtem Zustand bei der Einlieferung:

	Krankenhaus A	Krankenhaus B
gestorben	57	8
überlebt	1443	192

1. Bestimmen Sie die relativen Überlebenshäufigkeiten der Patienten in den beiden Krankenhäusern für die beiden Zustände der Patienten bei der Einlieferung.
2. Bestimmen Sie die aggregierte Kontingenztabelle mit den Variablen 'Zustand nach der Operation' und 'Krankenhaus'. Bestimmen Sie die relativen Überlebenshäufigkeiten in den beiden Krankenhäusern.

3. Offensichtlich liegt das Simpsons Paradox vor. Kurze Erklärung.

Aufgabe 33
In einem Lehrgang wurden 200 Teilnehmern unter anderem nach ihrem Geschlecht befragt. Außerdem wurden sie gefragt, ob sie den Film Titanic gesehen haben. Die Personen wurden dann noch gebeten, den Satz 'Zu Risiken und Nebenwirkungen...' fortzusetzen. Wir nennen dieses Merkmal Satz. Wenn die Person den Satz richtig fortgesetzt hat, nimmt das Merkmal Satz den Wert richtig an, ansonsten den Wert falsch. Die folgende Tabelle zeigt das Ergebnis der Umfrage

		Satz	
Geschlecht	Titanic	richtig	falsch
w	ja	64	16
	nein	14	6
m	ja	28	32
	nein	14	26

1. Erstellen Sie die Kontingenztabelle der Merkmale Titanic und Satz und bestimmen Sie die bedingten relativen Häufigkeiten. Deuten diese auf einen Zusammenhang zwischen den beiden Merkmalen hin? Falls dies der Fall ist, interpretieren Sie diesen Zusammenhang.

2. Offenbar haben die Personen, die den Film gesehen haben, ein gutes Gedächtnis. Aber was hat der Film mit dem Gedächtnis zu tun? Die Lösung dieses Rätsels erhalten wir folgendermaßen:

   (a) Erstellen Sie die Kontingenztabelle der Merkmale Geschlecht und Titanic und bestimmen Sie die bedingten relativen Häufigkeiten. Deuten diese auf einen Zusammenhang zwischen den beiden Merkmalen hin? Falls dies der Fall ist, interpretieren Sie diesen Zusammenhang.

   (b) Erstellen Sie die Kontingenztabelle der Merkmale Geschlecht und Satz und bestimmen Sie die bedingten relativen Häufigkeiten. Deuten diese auf einen Zusammenhang zwischen den beiden Merkmalen hin? Falls dies der Fall ist, interpretieren Sie diesen Zusammenhang.

Aufgabe 34
Sie schlagen am 28. August 1999 die NW auf und suchen alle Einzimmerwohnungen heraus, die explizit in Uninähe liegen. Es sind acht. Die folgende Tabelle gibt neben der Fläche x_t auch die Kaltmiete y_t der acht Wohnungen an.

t	x_t	y_t
1	20	270
2	26	460
3	32	512
4	48	550
5	26	360
6	30	399
7	30	419
8	40	390

1. Erstellen Sie das Streudiagramm der Daten.
2. Bestimmen und interpretieren Sie den Wert des Korrelationskoeffizienten von Bravais-Pearson.
3. Bestimmen und interpretieren Sie den Wert des Korrelationskoeffizienten von Spearman.

Aufgabe 35
In der folgenden Tabelle ist das Alter x (in Jahren) und der Verkaufspreis y von Autos der Marke Golf angegeben. Wurden mehrere Autos eines Baujahrs annonciert, dann wurde der Mittelwert der Preise bestimmt)
(Quelle: SÜDDEUTSCHE ZEITUNG 21.7.1999)

t	x_t	y_t
1	2	20300
2	3	19200
3	4	17300
4	5	16100
5	6	13400
6	7	11600

1. Erstellen Sie das Streudiagramm der Daten.
2. Bestimmen und interpretieren Sie den Wert des Korrelationskoeffizienten von Bravais-Pearson.

3. Bestimmen und interpretieren Sie den Wert des Korrelationskoeffizienten von Spearman.

Aufgabe 36

1. Gegeben seien n Punktepaare (x_i, y_i), $i = 1, \ldots, n$. Für jedes i gilt

   $$y_i = a + b \cdot x_i$$

   mit

   $$b > 0.$$

   Zeigen Sie, dass gilt:

   $$r = 1.$$

2. Gegeben seien n Punktepaare (x_i, y_i), $i = 1, \ldots, n$. Für jedes i gilt

   $$y_i = a + b \cdot x_i$$

   mit

   $$b < 0.$$

   Zeigen Sie, dass gilt:

   $$r = -1.$$

Aufgabe 37

In der folgenden Tabelle sind die durchschnittlichen Verbraucherpreise (in EURO) für 1 kg Kalbsschnitzel und 1 kg Brathähnchen (Tiefkühlkost) für die Jahre 1987, 1988, 1989, 1990 und 1991 angegeben. (Quelle: Statistisches Jahrbuch 1994, S. 663)

Jahr	Kalbfleisch	Brathähnchen
1987	14.9	2.62
1988	14.8	2.54
1989	15.5	2.50
1990	16.2	2.56
1991	16.4	2.60

1. Erstellen Sie das Streudiagramm der Daten.

2. Bestimmen Sie den Korrelationskoeffizienten von Bravais-Pearson.

3. Wie ändert sich der Wert des Korrelationskoeffizienten von Bravais-Pearson, wenn die Preise in Pfennigen und nicht in DM angegeben werden?

4. Bestimmen Sie den Korrelationskoeffizienten von Spearman.

Aufgabe 38
Zeichnen Sie das Streudiagramm und berechnen Sie den Korrelationskoeffizienten von Bravais-Pearson!

1. Fall 1

i	x_i	y_i
1	-2	4
2	-1	1
3	0	0
4	1	1
5	2	4

2. Fall 2

i	x_i	y_i
1	1	3
2	2	5
3	3	7
4	4	9
5	5	11

3. Fall 3

i	x_i	y_i
1	1	5
2	2	4
3	3	3
4	4	2
5	5	1

4. Fall 3

i	x_i	y_i
1	1	2
2	2	1
3	2	2
4	2	3
5	3	2

Aufgabe 39
In der folgenden Tabelle finden Sie für das Jahr 1986 die Geschwindigkeitsbeschränkung x auf Landstraßen (in Meilen pro Stunde) und die Anzahl y der Toten pro 100 Millionen Autokilometern in 5 Ländern.

Land	Höchstgeschwindigkeit	Anzahl Tote
Dänemark	55	4.1
Japan	55	4.7
Kanada	60	4.3
Holland	60	5.1
Italien	75	6.1

1. Zeichnen Sie das Streudiagramm
2. Bestimmen Sie den Korrelationskoeffizienten von Bravais-Pearson.
3. Wie ändert sich der Wert des Korrelationskoeffizienten von Bravais-Pearson, wenn die Geschwindigkeitsbeschränkung nicht in Meilen, sondern in Kilometern bestimmt wird?
4. In England betrug die Geschwindigkeitsbeschränkung im Jahr 1986 70 Meilen pro Stunde. Die Todesrate lag bei 3.5.
   (a) Berücksichtigen Sie diesen Wert im Streudiagramm.
   (b) Wie ändert sich der Wert des Korrelationskoeffizienten von Bravais-Pearson, wenn Sie den Wert von England noch im Streudiagramm berücksichtigen?

   wird kleiner [ ]

   bleibt gleich [ ]

   wird größer [ ]

   kann man nicht sagen [ ]
5. Bestimmen Sie den Korrelationskoeffizienten von Spearman.

Kapitel 4

Wahrscheinlichkeitsrechnung

Aufgabe 40
Stellen Sie die folgenden Aussagen über Ereignisse in Venn-Diagrammen und in symbolischer Schreibweise dar:

1. Die Ereignisse A und B können nicht gleichzeitig eintreten.
2. Immer wenn A eintritt, tritt auch B ein.
3. B tritt nur ein, wenn auch A eintritt.
4. C tritt genau dann ein, wenn A und B eintreten.
5. C tritt genau dann ein, wenn zwar A, aber nicht B eintritt.

Aufgabe 41
A, B und C seien Ereignisse in E.
Zeichnen Sie jeweils ein Venn-Diagramm für die folgenden Ereignisse und geben Sie sie formal an:

1. A und B treten ein, C nicht.
2. Nur A tritt ein.
3. A tritt ein und B nicht.
4. A oder B treten ein, aber nicht beide.

Aufgabe 42
Eine Münze wird zweimal hintereinander geworfen.

1. Geben Sie die Ergebnismenge dieses Zufallsvorgangs an.
2. A sei das Ereignis, dass die Münze beim ersten Wurf *KOPF* zeigt.

   B sei das Ereignis, dass die Münze beim zweiten Wurf *KOPF* zeigt.

   Geben Sie $\bar{A}$, $A \cap B$, $A \cup B$, $B \cap \bar{A}$, $\bar{A} \cup \bar{B}$ an.
3. Wie läßt sich das Ereignis C, dass mindestens einmal *KOPF* eingetreten ist, aus den Ereignissen A und B darstellen ?

Aufgabe 43
Ein Würfel wird zweimal hintereinander geworfen.

1. Geben Sie die Ergebnismenge und die folgenden Ereignisse an:

A:	Die beiden Augenzahlen sind verschieden
B:	Die Augensumme ist gerade
C:	Die Augensumme ist höchstens 4
D:	Die zweite Augenzahl ist gerade
E:	Die zweite Augenzahl ist nicht größer als die erste

2. Um welche Ereignisse handelt es sich bei $A \cup B$, $A \cap C$, $B \cap D$, $\bar{D} \cap \bar{E}$?

Aufgabe 44
Bei einer Prüfung sind 25 Pozent der Prüflinge in Mathematik, 15 Prozent in Chemie und 10 Prozent in Chemie und Mathematik durchgefallen. Einer der Prüflinge wird zufällig ausgewählt.
Wie groß ist die Wahrscheinlichkeit, dass er

1. in mindestens einem der beiden Fächer durchgefallen ist?
2. nur in Mathematik durchgefallen ist?
3. in keinem Fach durchgefallen ist?
4. in genau einem Fach durchgefallen ist?

Aufgabe 45
Die Wahrscheinlichkeit, dass jemand die Zahnpasta *BLENDI* kauft, ist 0.5 und dass er die Seife *DUFTI* kauft, ist 0.6. Mindestens einen der beiden Artikel kauft er mit einer Wahrscheinlichkeit von 0.9.
Wie groß ist die Wahrscheinlichkeit, dass er

1. *BLENDI* und $DUFTI$ kauft?

2. keines von beiden kauft?

3. genau eines von beiden kauft?

4. nur *BLENDI* kauft?

Aufgabe 46
Ein elektronisches Gerät besteht aus zwei Komponenten A und B. Aus langjähriger Erfahrung weiß man, dass A mit Wahrscheinlichkeit 0.05 ausfällt. Die Wahrscheinlichkeit, dass A ausfällt, wenn B ausgefallen ist, beträgt 0.2. Außerdem ist bekannt, dass mit Wahrscheinlichkeit 0.02 beide ausfallen.

1. Wie groß ist die Wahrscheinlichkeit, dass B ausfällt, wenn A ausgefallen ist?

2. Wie groß ist die Wahrscheinlichkeit, dass B ausfällt?

3. Wie groß ist die Wahrscheinlichkeit, dass mindestens eins von beiden ausfällt?

4. Wie groß ist die Wahrscheinlichkeit, dass A und B nicht gleichzeitig ausfallen?

5. Wie groß ist die Wahrscheinlichkeit, dass B nicht ausfällt,wenn A ausgefallen ist?

Aufgabe 47
Ein Autokennzeichen besteht aus 5 Zeichen. Die ersten beiden Zeichen sind Buchstaben, die restlichen Ziffern.

1. Wie viele unterschiedliche Autokennzeichen gibt es?

2. Wie viele unterschiedliche Autokennzeichen gibt es, wenn keine Ziffer doppelt vorkommen darf?

3. Wie viele unterschiedliche Autokennzeichen gibt es, wenn nicht 4 Nullen gleichzeitig auftreten dürfen?

4. Wie viele unterschiedliche Autokennzeichen gibt es, bei denen eine aufsteigende Folge aufeinanderfolgender Ziffern vorkommt?

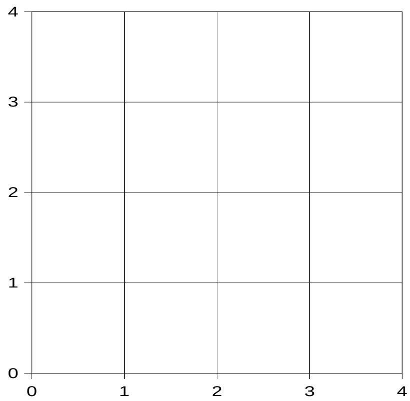
Aufgabe 48
1. 6 Würfel werden gleichzeitig geworfen. Mit welcher Wahrscheinlichkeit liegen alle 6 Augenzahlen auf dem Tisch?

2. In einem Zimmer sind r Personen versammelt.

   Wie wahrscheinlich ist es, dass mindestens zwei am selben Tag des Jahres Geburtstag haben?

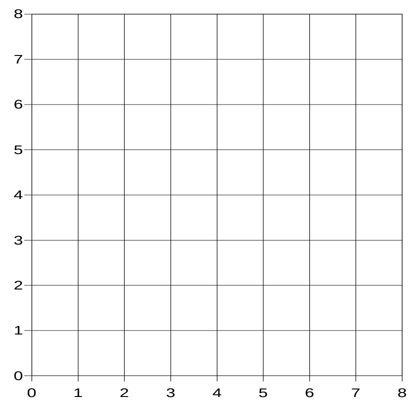
Aufgabe 49

1. Sepp wohnt in $(0,0)$ und arbeitet in $(4,4)$.

   Wie viele kürzeste Wege zur Arbeit gibt es für Sepp?

2. Abel wohnt in $(0,0)$ und arbeitet in $(8,8)$. Sein Arbeitskollege Kain wohnt in $(4,4)$. Abel fährt jeden Morgen zur Arbeit und nimmt unterwegs Kain mit.

   (a) Wie viele kürzeste Wege zur Arbeit gibt es für Abel?

(b) Wie viele gibt es, wenn er Kain nicht mitnimmt?

Aufgabe 50
Gegeben sei eine Gruppe, die aus 5 Männern und 7 Frauen besteht.

1. 5 dieser Personen sollen einen Ausschuss bilden. Wie viele unterschiedliche Ausschüsse gibt es?

2. Wie viele Auschüsse gibt es, die genau 2 Männer und 3 Frauen enthalten?

3. Es soll ein Ausschuss gebildet werden, der aus 3 Frauen besteht. Zwei der 7 Frauen wollen auf gar keinen Fall in diesem Ausschuss zusammenarbeiten.

   Wie viele Ausschüsse, die genau 3 Frauen enthalten, gibt es, in denen die beiden Frauen nicht zusammen enthalten sind?

Aufgabe 51
Der 'Deutsche Kaninchenzüchterverein e.V.' beschäftigt sich wieder mit sich selbst. Ein neuer Vorstand soll gewählt werden, bestehend aus dem 1. Vorsitzenden, 2. Vorsitzenden, Kassenwart, Revisor und Pressesprecher. Das bedeutet, dass jedes zweite Mitglied ein Amt erhält, denn der Verein umfasst nur 10 Mitglieder.
Die 'Igelzüchter 1902 e.V.' finden das sehr spießig. Sie haben selbstverständlich keinen Vorstand. Allerdings wird auch dort gewählt: Um effektiv arbeiten zu können, benötigen die 10 Stacheltierfreunde einen fünfköpfigen Geschäftsführenden Ausschuss.
Wie viele unterschiedliche Ergebnisse kann die Wahl zum

1. Geschäftsführenden Ausschuss der 'Igelzüchter 1902 e.V.'

2. Vorstand des 'Deutsche Kaninchenzüchterverein e.V.' erbringen?

Aufgabe 52

1. Auf wie viele unterschiedliche Arten können 4 Männer und 3 Frauen auf einer langen Bank nebeneinander sitzen?

2. Auf wie viele unterschiedliche Arten können 4 Männer und 3 Frauen auf einer langen Bank nebeneinandersitzen, wenn sowohl die Männer als auch die Frauen in Gruppen zusammensitzen sollen?

3. Auf wie viele unterschiedliche Arten können 4 Männer und 3 Frauen auf einer langen Bank nebeneinandersitzen, wenn nur die Männer in einer Gruppe zusammensitzen sollen?

Aufgabe 53
In einer Klausur gibt es drei Blöcke mit jeweils 4 Aufgaben.

1. Wie viele unterschiedliche Kombinationen von Aufgaben gibt es, wenn ein Student 9 Aufgaben bearbeiten muss und diese aus den 12 Aufgaben auswählen kann?

2. Wie viele unterschiedliche Kombinationen von Aufgaben gibt es, wenn ein Student aus jeder der Gruppen genau 3 Aufgaben bearbeiten muss.

3. Wie viele unterschiedliche Kombinationen von Aufgaben gibt es, wenn ein Student insgesamt 9 Aufgaben bearbeiten muss, wobei aus jeder Gruppe mindestens zwei Aufgaben bearbeitet werden müssen?

Aufgabe 54
Von einer Gruppe von Personen, die aus 3 Frauen und 3 Männern besteht, soll ein Gruppenfoto gemacht werden.

1. Wie viele unterschiedliche Fotos gibt es, wenn die 6 Personen nebeneinanderstehen?

2. Wie viele unterschiedliche Fotos gibt es, wenn links die drei Männer und rechts die drei Frauen nebeneinanderstehen sollen?

3. Bei den Personen handelt es sich um 3 Ehepaare.

   Wie viele unterschiedliche Fotos gibt es, auf denen die 6 Personen nebeneinanderstehen, wobei die Ehepartner aber nebeneinanderstehen?

4. Wie viele Möglichkeiten gibt es für ein Foto mit drei Personen?

Aufgabe 55

1. Geben Sie je ein Beispiel an für

$$\begin{aligned} P(A|B) &> P(A) \\ P(A|B) &= P(A) \\ P(A|B) &< P(A) \end{aligned}$$

2. Gilt für beliebige Ereignisse A und B

$$P(A|B) = 1 - P(A|\bar{B})$$

   oder

$$P(A|B) = 1 - P(\bar{A}|B)$$

Aufgabe 56

1. Seien A und B zwei Ereignisse mit $P(A) > 0$ und $P(B) > 0$.

   Zeigen Sie

   (a) Sind A und B disjunkt, so sind sie voneinander abhängig.

   (b) Sind A und B unabhängig, so sind sie nicht disjunkt.

2. Zeigen Sie:

   Sind die Ereignisse A und B unabhängig, so sind auch die Ereignisse $\bar{A}$ und $\bar{B}$ unabhängig.

Aufgabe 57

Eine faire Münze wird zweimal geworfen.
Wir betrachten die folgenden Ereignisse:

A: im 1. Wurf erscheint ZAHL

B: im 2. Wurf erscheint ZAHL

C: es erscheint genau einmal ZAHL

1. Welche Ereignispaare sind unabhängig?

2. Sind die Ereignisse A, B und C unabhängig?

Aufgabe 58

Ein meteorologisches Institut sagt mit Wahrscheinlichkeit 0.5 das Wetter richtig voraus, ein anderes trifft seine Voraussage unabhängig von ihm mit der Wahrscheinlichkeit 0.6.

1. Wie groß ist die Wahrscheinlichkeit, dass beide Institute das Wetter richtig voraussagen?

2. Wie groß ist die Wahrscheinlichkeit, dass mindestens eins der beiden Institute das Wetter richtig voraussagt?

3. Wie groß ist die Wahrscheinlichkeit, dass beide Institute das Wetter richtig voraussagen, wenn man weiß, dass mindestens eins der Institute das Wetter richtig vorausgesagt hat?

4. Wie groß ist die Wahrscheinlichkeit, dass genau eins der beiden Institute das Wetter richtig voraussagt?

Aufgabe 59

Ein System besteht aus den Komponenten `K1` und `K2`. Die Ausfallwahrscheinlichkeit von `K1` beträgt p und die Ausfallwahrscheinlichkeit von `K2` beträgt ebenfalls p. Die Ausfälle der Komponenten sind unabhängig voneinander.

1. Die beiden Komponenten sind parallelgeschaltet. Das System fällt also aus, wenn beide Komponenten ausfallen. Wie groß ist die Wahrscheinlichkeit hierfür?

2. Die beiden Komponenten sind hintereinandergeschaltet. Das System fällt also aus, wenn mindestens eine der beiden Komponenten ausfällt. Wie groß ist die Wahrscheinlichkeit hierfür?

3. Das System unter 2 wird nun so erweitert, dass zum Block der hintereinandergeschalteten Komponenten `K1` und `K2` ein Block aus den hintereinandergeschalteten Komponenten `K3` und `K4` parallelgeschaltet wird.

   Bestimmen Sie die Ausfallwahrscheinlichkeit des gesamten Systems. Unterstellen Sie dabei, dass die Ausfallwahrscheinlichkeit jeder der 4 Komponenten p beträgt, und dass die Ausfälle der einzelnen Komponenten voneinander unabhängig sind.

Aufgabe 60
Zwei Fußballspieler schießen einmal auf die Torwand. Vom Spieler A ist bekannt, dass er mit Wahrscheinlichkeit 0.4 ein Tor schießt, während diese Wahrscheinlichkeit beim Spieler B 0.5 beträgt. Die Schussergebnisse beider Spieler seien unabhängig voneinander.

1. Wie groß ist die Wahrscheinlichkeit, dass mindestens einer von beiden trifft?

2. Wie groß ist die Wahrscheinlichkeit, dass genau einer von beiden trifft?

3. Wie groß ist die Wahrscheinlichkeit, dass nur B trifft?

4. Die Spieler fordern den Moderator M auf, ebenfalls auf die Torwand zu schießen. Die Wahrscheinlichkeit, dass M trifft, beträgt 0.1. Wie groß ist die Wahrscheinlichkeit, dass genau einer der drei trifft, wenn die Schussergebnisse aller drei unabhängig voneinander sind?

Aufgabe 61
5 Prozent der Bevölkerung haben hohen Blutdruck. Von den Personen mit hohem Blutdruck trinken 75 Prozent Alkohol, während nur 50 Prozent der Personen ohne hohen Blutdruck Alkohol trinken.
Wieviel Prozent der Personen, die Alkohol trinken, haben hohen Blutdruck?

Aufgabe 62
Von den am Montag produzierten Autos einer Marke weisen 4 Prozent, von den am Freitag produzierten Autos weisen 3 Prozent und von den an den restliche Werktagen produzierten weisen jeweils 1 Prozent innerhalb des ersten Jahres erhebliche Mängel auf.
An jedem der Wochentage wird die gleiche Anzahl von Autos produziert.

1. Ein zufällig aus der Produktion einer Woche ausgewählter Wagen sei nicht in Ordung.

   Wie groß ist die Wahrscheinlichkeit, dass er am Montag produziert wurde?

2. Welches Ergebnis erhalten Sie, wenn am Freitag 12 Prozent der Autos produziert werden, und die restliche Produktion sich gleichmäßig auf die 4 übrigen Werktage verteilt?

Aufgabe 63
In einem Land gibt es die Parteien `A`, `B` und `C`, wobei `A` von 50 Prozent, `B` von 30 Prozent und `C` von 20 Prozent der Bürger gewählt wird.
Von den Wählern der Partei `A` sind 80 Prozent, von denen der Partei `B` 70 Prozent und von denen der Partei `C` 40 Prozent gegen eine Herabsetzung des Wahlalters auf 16 Jahre.

1. Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Bürger gegen die Herabsetzung des Wahlalters auf 16 Jahre ist?

2. Ein zufällig ausgewählter Bürger ist gegen die Herabsetzung des Wahlalters auf 16 Jahre. Welche Partei wählt er mit der größten Wahrscheinlichkeit?

Aufgabe 64
Die Ergebnismenge eines Zufallsvorgangs ist $E = \{a, b, c, d, e\}$.
Für die Ereignisse $A = \{a, b, c\}$ und $B = \{a\}$ gilt $P(A) = \frac{1}{2}$ und $P(B) = \frac{1}{3}$.
Außerdem seien die Ereignisse $C = \{b, c\}$, $D = \{d, e\}$ und $F = \{a, d, e\}$ gegeben.

1. Bestimmen Sie $P(C)$, $P(D)$ und $P(F)$.

2. Wie groß ist die Wahrscheinlichkeit für das Eintreten von A, wenn man weiß, dass B eingetreten ist?

3. Wie groß ist die Wahrscheinlichkeit für das Eintreten von B, wenn man weiß, dass A eingetreten ist?

4. Wie groß ist die Wahrscheinlichkeit, daß D und F eintreten, wenn man weiß, dass mindestens eins der beiden Ereignisse eingetreten ist?

Aufgabe 65
Der nachstehende Text stammt aus dem SPIEGEL 34/1991, S.212:

> Im Fernsehstudio steuert die Show *LET'S MAKE A DEAL* ihrem Höhepunkt entgegen. Der Kandidat hat die Endrunde erreicht, und damit beginnt ein Ritual, das ihm die Chance eröffnet, ein Auto zu gewinnen.
>
> Der Supergewinn verbirgt sich hinter einer von drei Holztüren; hinter den beiden anderen ist, gleichsam als lebende Niete, jeweils eine Ziege angepflockt. Der Kandidat trifft eine erste Wahl zwischen den drei Türen; beispielsweise entscheidet er sich für Tür Nummer eins. Seine Chancen, das Auto zu ergattern, stehen dabei eins zu drei.
>
> Dann folgt der zweite Schritt des Rituals. Der Showmaster öffnet eine der beiden übriggebliebenen Türen - im gewählten Beispiel die dritte. Naturgemäß (da der Spielleiter weiß, was sich hinter jeder Tür verbirgt) erscheint kein Auto, sondern es meckert eine Ziege.
>
> Einen weiteren Twist erfährt das Türenritual im dritten Teil der Prozedur: Der Showmaster stellt dem Kandidaten frei, seine Wahl ('Tür eins') noch einmal zu überdenken und vielleicht die zweite statt der ersten Tür zu öffnen.
>
> Soll er oder soll er nicht? Kann er durch einen Wechsel seines Tür-Tips die Gewinnaussichten verbessern? Oder bleibt die Gewinnchance, egal auf welche der beiden Resttüren seine Wahl fällt, halbe-halbe?

Versuchen Sie diese Fragen zu beantworten.

Wenn Sie gar nicht mehr vorankommen, schauen Sie doch einmal in die Zeitschrift CHANCE Vol.4 No.2, 1991, S.6-9.

Sehr zu empfehlen ist auch:

Gero von Randow: Das Ziegenproblem. rororo

Kapitel 5

Univariate Zufallsvariablen

Aufgabe 66
Eine faire Münze wird dreimal hintereinander geworfen. Fällt dreimal KOPF, so erhält Spieler `A` vom Spieler `B` 6 EURO. Fällt zweimal KOPF, so erhält Spieler `A` vom Spieler `B` 4 EURO. Der Spieler `A` zahlt an den Spieler `B` 6 EURO, wenn einmal KOPF fällt. Erscheint dreimal ZAHL, so braucht keiner der Spieler zu zahlen.
Sei X der Gewinn bzw. Verlust des Spielers `A`.

1. Welche Werte kann X annehmen?

2. Bestimmen Sie die Wahrscheinlichkeitsfunktion von X.

3. Zeichnen Sie die Verteilungsfunktion.

4. Bestimmen Sie folgende Wahrscheinlichkeiten:

   (a) $P(X = 0)$
   (b) $P(X \leq 0)$
   (c) $P(X > 4)$
   (d) $P(0 \leq X < 3)$

Aufgabe 67
Eine faire Münze werde dreimal geworfen. Die Zufallsvariable X ordnet jedem Ergebnis $e \in E$ die Differenz zwischen der Anzahl KOPF der Anzahl ZAHL zu.

1. Welche Werte kann X annehmen?

2. Bestimmen Sie die Wahrscheinlichkeitsfunktion von X.

3. Zeichnen Sie die Verteilungsfunktion.

4. Bestimmen Sie folgende Wahrscheinlichkeiten:

   (a) $P(X \leq 2)$
   (b) $P(X \leq 4)$
   (c) $P(X \geq 3.1)$
   (d) $P(X = 1)$

Aufgabe 68
Man nehme an, dass Knaben- und Mädchengeburten gleichwahrscheinlich sind, und Unabhängigkeit zwischen verschiedenen Geburten bestehe.
Sei X die Anzahl der Knaben bei drei Geburten.

1. Welche Werte kann X annehmen?

2. Bestimmen Sie die Wahrscheinlichkeitsfunktion von X.

3. Zeichnen Sie die Verteilungsfunktion.

4. Bestimmen Sie folgende Wahrscheinlichkeiten:

   (a) $P(X \leq 1)$
   (b) $P(X > 0)$
   (c) $P(1 < 3)$

Aufgabe 69
Eine faire Münze wird so oft geworfen, bis zum ersten Mal KOPF auftritt.
Sei X die Anzahl der dazu notwendigen Würfe.

1. Welche Werte kann X annehmen?

2. Bestimmen Sie die Wahrscheinlichkeitsfunktion von X.

3. Bestimmen Sie die Verteilungsfunktion.

4. Wie groß ist $P(X > 4)$?

5. Wie groß ist die Wahrscheinlichkeit, dass bei einer ungeraden Anzahl von Würfen zum ersten Mal KOPF fällt?

Aufgabe 70
Ein Mann kommt betrunken nach Hause. An seinem Schlüsselbund befinden sich 5 Schlüssel, von denen einer der Hausschlüssel ist. Da er nicht mehr weiß, welcher der richtige ist, wählt er einen zufällig aus und versucht mit diesem die Tür zu öffnen. Ist der Schlüssel falsch, so versucht er es noch einmal.
Sei X die Anzahl der Versuche, die er benötigt, bis es ihm gelingt die Tür zu öffnen.

1. Bestimmen Sie die Wahrscheinlichkeitsfunktion von X unter der Annahme, dass ihm vor jedem Versuch der Schlüsselbund zu Boden fällt, und er mit der Suche wieder von vorne beginnt.

   Wie groß ist in diesem Fall die Wahrscheinlichkeit, dass es ihm bei einer geraden Anzahl von Versuchen gelingt, die Tür zu öffnen?

2. Bestimmen Sie die Wahrscheinlichkeitsfunktion von X unter der Annahme, dass ihm vor jedem Versuch die Schlüssel bekannt sind, mit denen er schon vergeblich versucht hat, die Tür zu öffnen.

   Wie groß ist in diesem Fall die Wahrscheinlichkeit, dass es ihm bei einer geraden Anzahl von Versuchen gelingt, die Tür zu öffnen?

Aufgabe 71
Welche der folgenden Funktionen können wir als Dichtefunktion einer stetigen Zufallsvariablen X auffassen? Bestimmen Sie in diesen Fällen die Verteilungsfunktion.

1.
$$f(x) = \begin{cases} 1 & \text{für} \quad 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

2.
$$f(x) = \begin{cases} 2\,(1-x) & \text{für} \quad 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

3.
$$f(x) = \begin{cases} \dfrac{1}{x^2} & \text{für} \quad x \geq 1 \\ 0 & \text{sonst} \end{cases}$$

4.

$$f(x) = \begin{cases} \lambda\, e^{-\lambda\, x} & \text{für} \quad x \geq 0 \\ \\ 0 & \text{sonst} \end{cases}$$

5.

$$f(x) = \begin{cases} 0.5 & \text{für} \quad 0 \leq x < 1 \\ \\ 1 & \text{für} \quad 1 \leq x < 1.5 \\ \\ 0 & \text{sonst} \end{cases}$$

Aufgabe 72
Gegeben sei folgende Funktion

$$f(x) = \begin{cases} a\, x & \text{für} \quad 0 \leq x \leq 1 \\ \\ 0 & \text{sonst} \end{cases}$$

Für welchen Wert von a handelt es sich um die Dichtefunktion einer stetigen Zufallsvariablen X?
Gehen Sie im folgenden davon aus, dass a so gewählt wurde, dass es sich um eine Dichtefunktion handelt.

1. Bestimmen Sie die Verteilungsfunktion von X.
2. Zeichnen Sie die Dichtefunktion und die Verteilungsfunktion von X für $a = 2$.
3. Mit welcher Wahrscheinlichkeit nimmt X Werte größer als 0.5 an?
4. Für welche Zahl c nimmt X mit der Wahrscheinlichkeit 0.5 Werte an, die größer als c sind?

Aufgabe 73
Die Dichtefunktion einer Zufallsvariablen X, die gleichverteilt über dem Intervall $[a, b]$ ist, lautet:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{für} \quad a \leq x \leq b \\ \\ 0 & \text{sonst} \end{cases}$$

1. Zeichnen Sie die Dichtefunktion.

2. Bestimmen Sie die Verteilungsfunktion $F(x)$ und stellen Sie diese graphisch dar!

3. Es sei speziell $a = 0$ und $b = 1$.

   Berechnen Sie

   (a) $P(X \leq \frac{1}{2})$

   (b) $P(X \leq \frac{1}{3})$

   (c) $P(X \geq \frac{3}{4})$

Kapitel 6

Verteilungsparameter

Aufgabe 74
Tversky und Kahnemann fragten Personen, welche der beiden folgenden Alternativen sie vorzögen.

Alternative A

Man muss 750 EURO zahlen.

Alternative B

Eine faire Münze wird zweimal hintereinander geworfen. Fällt beide Male KOPF, so muss man nichts zahlen, ansonsten muss man 1000 EURO zahlen.
Vergleichen Sie die beiden Alternativen hinsichtlich des Erwartungswerts und der Varianz des Betrags, den man zahlen muss.

Aufgabe 75
Eine Urne enthält 5 Kugeln, von denen drei 10 g und zwei 20 g wiegen. Es werden zwei Kugeln *ohne Zurücklegen* gezogen. Die Zufallsvariable X ordnet jedem Ergebnis dieses Zufallsvorgangs das Gesamtgewicht der beiden Kugeln zu.

1. Leiten Sie die Wahrscheinlichkeitsfunktion von X her!

2. Berechnen Sie den Erwartungswert und die Varianz von X!

Was ändert sich unter 1. und 2., wenn *mit Zurücklegen* gezogen wird?

Aufgabe 76
Bei einem Spiel setzt jeder der beiden Mitspieler 1 EURO ein. Dann werden zwei Würfel geworfen und die Differenz aus der größeren Augenzahl und der kleineren Augenzahl bestimmt.

Nimmt die Differenz die Werte 0, 1 oder 2 an, so erhält Spieler `A` den gesamten Einsatz, ansonsten Spieler `B`.

1. Zeigen Sie, dass der erwartete Gewinn beider Spieler nicht gleich ist.
2. Wie hoch muss der Einsatz von Spieler `A` sein, damit der erwartete Gewinn beider Spieler gleich ist?

Aufgabe 77
Beim Sammelkartenspiel `POKEMON` kämpfen zwei Pokemons gegeneinander. Um einen Angriff durchführen zu können benötigt man genug Energie, die man seinem Pokenom in Form von Energiekarten gibt. Bei einem Angriff fügt man dem gegenerischen Pokemon Schadenspunkte zu. Das Basis-Pokemon `Nidoran` kann einen Angriff mit Namen *Kratzfurie* durchführen. Hierbei geht es folgendermaßen vor:

Wirf drei Münzen. Dieser Angriff fügt jedesmal, wenn die Münze KOPF zeigt, 10 Schadenspunkte zu.

Sei X die Gesamtzahl der Schadenspunkte, die dem angegriffenen Pokemon durch Nidoran mit dem Angriff *Kratzfurie* zugefügt werden.

1. Leiten Sie die Wahrscheinlichkeitsverteilung von X her.
2. Zeichnen Sie die Verteilungsfunktion von X.
3. Bestimmen Sie $E(X)$ und $Var(X)$.

Wenn Sie wissen wollen, wie ein Statistiker das Sammelkartenspiel POKEMON sieht, sollten Sie in Vol. 13 No.2 der Zeitschrift *Chance* schauen und den Artikel *Statistics and the Pokemon Craze: Trying to impress my kids* von Scott M. Berry lesen.

Aufgabe 78
Ein Spiel besitzt folgende Regeln:

Ein Spieler zahlt einen Betrag von 1 EURO als Einsatz. Es werden drei faire Würfel gleichzeitig geworfen. Erscheint die Zahl 5 auf einem Würfel, so erhält der Spieler 1 EURO. Erscheint die Zahl 5 auf zwei Würfeln, so erhält er 2 EURO. Erscheint die Zahl 5 auf drei Würfeln, so erhält er 3 EURO. Erscheint auf keinem Würfel die Zahl 5, so erhält er keine Auszahlung.

1. Man bestimme die Ergebnismenge des Zufallsexperiments.
2. X bezeichne den Gewinn (Auszahlung minus Einsatz) des Spielers. Welche Werte kann X annehmen und mit welchen Wahrscheinlichkeiten?
3. Man bestimme und zeichne die Verteilungsfunktion von X.
4. Man berechne den erwarteten Nettogewinn pro Spiel.
5. Ist das Spiel fair?
6. Man berechne die Varianz des Gewinns.

Aufgabe 79
Eine Tüte enthält 5 Gummibärchen, von denen 3 rot und 2 grün sind.

1. Der Tüte werden nacheinander zwei Gummibärchen zufällig entnommen, wobei das zuerst entnommene Gummibärchen nicht zurückgelegt wird.

   Sei X die Anzahl der grünen Gummibärchen, die entnommen wurden.

   (a) Bestimmen Sie die Wahrscheinlichkeitsfunktion von X.

   (b) Berechnen Sie $E(X)$ und $Var(X)$.

2. Beantworten Sie die Fragen unter 1.(a) und 1.(b), wenn das zuerst entnommene Gummibärchen in die Tüte zurückgelegt wird, bevor das zweite Gummibärchen gezogen wird.

Aufgabe 80
Gegeben seien die Dichtefunktionen

1.

$$f(x) = \begin{cases} 1 & \text{für} \quad 0 \leq x \leq 1 \\ \\ 0 & \text{sonst} \end{cases}$$

2.

$$f(x) = \begin{cases} 2\,(1-x) & \text{für} \quad 0 \leq x \leq 1 \\ \\ 0 & \text{sonst} \end{cases}$$

3.

$$f(x) = \begin{cases} 2\,x & \text{für} \quad 0 \leq x \leq 1 \\ \\ 0 & \text{sonst} \end{cases}$$

Bestimmen Sie jeweils $x_{0.5}$, $E(X)$ und $Var(X)$.

Aufgabe 81
Gegeben sei folgende Dichtefunktion einer stetigen Zufallsvariablen X:

$$f(x) = \begin{cases} \dfrac{1}{x^2} & \text{für} \quad x \geq 1 \\ \\ 0 & \text{sonst} \end{cases}$$

Zeigen Sie, dass der Erwartungswert von X nicht existiert.

Aufgabe 82
Gegeben sei folgende Dichtefunktion einer stetigen Zufallsvariablen X:

$$f(x) = \begin{cases} e^{-x} & \text{für} \quad x \geq 0 \\ \\ 0 & \text{sonst} \end{cases}$$

Bestimmen Sie $x_{0.5}$, $E(X)$ und $Var(X)$.

Aufgabe 83
Gegeben sei folgende Funktion $f(x)$ mit

$$f(x) = \begin{cases} 0.5 \cdot x - 0.5 & \text{für} \quad 1 \leq x \leq 3 \\ \\ 0 & \text{sonst} \end{cases}$$

1. Zeigen Sie, dass $f(x)$ eine Dichtefunktion ist.
2. Bestimmen Sie die zugehörige Verteilungsfunktion.
3. Bestimmen Sie den Median und den Erwartungswert.
4. Wie groß ist die Wahrscheinlichkeit, dass X Werte annimmt, die größer als 2 sind?

Kapitel 7

Multivariate Zufallsvariablen

Kapitel 8

Parameter multivariater Verteilungen

Aufgabe 84
Ein regelmäßiges Tetraeder ist ein von vier gleichseitigen Dreiecken begrenzter Körper, also so eine Art vierseitiger Würfel. Ein solches Tetraeder werde nun zweimal hintereinander geworfen, wobei bei den Würfen die unten liegende Seite von Interesse ist.

1. Geben Sie die Ergebnismenge dieses Zufallsvorgangs an.

2. Wir betrachten die beiden Zufallsvariablen:

   X: Summe der beiden Augenzahlen
   Y: Augenzahl beim ersten Wurf minus Augenzahl beim zweiten Wurf

   (a) Bestimmen Sie die gemeinsame Verteilung der Zufallvariablen X und Y und geben Sie die beiden Randverteilungen an.
   (b) Sind die Zufallsvariablen X und Y unabhängig?
   (c) Berechnen Sie die Kovarianz zwischen X und Y.

Aufgabe 85
Eine Urne enthalte drei Kugeln, die von 1 bis 3 durchnummeriert sind. Aus der Urne werden zwei Kugeln ohne Zurücklegen entnommen.
Sei X die größere der beiden Zahlen und Y die kleinere.

1. Bestimmen Sie die gemeinsame Wahrscheinlichkeitsfunktion von X und Y.

2. Geben Sie die beiden Randverteilungen an.

3. Sind die Zufallsvariablen X und Y unabhängig?

4. Berechnen Sie die Kovarianz zwischen X und Y.

Aufgabe 86
Eine Urne enthält 5 Kugeln, von denen zwei 20 g und drei 10 g wiegen. Es werden zwei Kugeln ohne Zurücklegen gezogen.
Sei X das Gesamtgewicht der beiden Kugeln und Y das Gewicht der leichteren Kugel.

1. Bestimmen Sie die gemeinsame Wahrscheinlichkeitsfunktion von X und Y.

2. Bestimmen Sie die Randverteilung von X und die Randverteilung von Y.

3. Bestimmen Sie die Kovarianz zwischen X und Y.

4. Sind die Zufallsvariablen X und Y unabhängig?

Aufgabe 87
Eine Urne enthält 3 Kugeln, die von 1 bis 3 durchnumeriert sind. Es werden zwei Kugeln ohne Zurücklegen gezogen.
Sei X die kleinere der gezogenen Zahlen und Y die Summe der beiden Zahlen.

1. Leiten Sie die gemeinsame Wahrscheinlichkeitsfunktion von X und Y her.

2. Geben Sie die Wahrscheinlichkeitsfunktion von X und die Wahrscheinlichkeitsfunktion von Y an.

3. Bestimmen Sie $E(X)$.

4. Sind X und Y unabhängig?

5. Was wissen Sie aufgrund von 4. über $Cov(X,Y)$?

Kapitel 9

Verteilungsmodelle

Aufgabe 88
Ein Testbogen setzt sich aus 10 Fragen zusammen. Bei jeder Frage gibt es 5 Antwortalternativen, von denen genau eine richtig ist. Ein Prüfling versucht, die richtigen Antworten zu erraten.
Sei X die Anzahl der richtig geratenen Antworten.

1. Wie ist X verteilt?
2. Bestimmen Sie den Erwartungswert und die Varianz von X!
3. Wie groß ist die Wahrscheinlichkeit,
   (a) den Test zu bestehen, wenn dazu mehr als die Hälfte aller Fragen richtig beantwortet werden muss?
   (b) genau 2 Antworten richtig zu erraten?
   (c) mehr als 2 Antworten richtig zu erraten?
   (d) mindestens 2 Antworten richtig zu erraten?
   (e) höchstens 2 Antworten richtig zu erraten?
   (f) weniger als 2 Antworten richtig zu erraten?

Aufgabe 89
Eine automatisch gesteuerte Ampelanlage zeigt 10 Sekunden lang grün, 40 Sekunden lang rot, 10 Sekunden lang grün, u.s.w. . Ein Zyklus dauert also 50 Sekunden, da Fußgängerampeln kein gelbes Licht haben. Herr Walker muss diesen Übergang täglich benutzen. Seine Ankunftszeiten am Fußgängerüberweg sind rein zufällig verteilt. Er geht also bis zum Straßenrand, schaut dann erst, ob die Ampel rot oder grün zeigt und verhält sich als braver Verkehrsteilnehmer entsprechend der jeweiligen Anzeige.

1. Gestern musste Herr Walker zehnmal über diesen Übergang gehen.

   Sei X die Anzahl der Fälle, in denen Herr Walker die Ampel ohne Verzögerung überqueren konnte.

   (a) Wie ist X verteilt?

   (b) Bestimmen Sie den Erwartungswert und die Varianz von X!

   (c) Wie groß ist die Wahrscheinlichkeit, dass er genau fünfmal ohne Verzögerung weitergehen konnte?

   (d) Wie groß ist die Wahrscheinlichkeit, dass er jedesmal warten musste?

2. Die Ampelphase wurde so geändert, dass 20 Sekunden lang grün, 30 Sekunden lang rot, u.s.w. angezeigt wird.

   Beantworten Sie die Fragen unter 1. für diesen Fall.

Aufgabe 90

Eine Urne enthält 4 schwarze und 6 weiße Kugeln. Aus der Urne wurden 100 mal hintereinander jeweils drei Kugeln gezogen, wobei aber nicht bekannt ist, ob mit oder ohne Zurücklegen gezogen wurde. Es ergab sich folgende Häufigkeitsverteilung für die Anzahl der schwarzen Kugeln.

x_i	n_i
0	15
1	49
2	32
3	4

Für welche Art der Ziehung spricht dies Ergebnis?

Aufgabe 91

Eine Warenlieferung enthält 20 Stücke, von denen 5 ohne Zurücklegen entnommen werden. Ist unter den 5 kein oder ein defektes, so wird die Lieferung angenommen, ansonsten wird sie zurückgewiesen. Wir nehmen an, dass die Lieferung 3 defekte Stücke enthält.
Sei X die Anzahl der defekten Stücke in der Stichprobe.

1. Wie ist X verteilt?

2. Wie groß ist die Wahrscheinlichkeit, dass die Lieferung zurückgewiesen wird?

Aufgabe 92
Ein Obsthändler erhält eine Lieferung von 200 Orangen, von denen 20 nicht in Ordnung sind.

1. Es werden 3 Orangen ohne Zurücklegen entnommen.

   Sei X die Anzahl der Orangen, die nicht in Ordnung sind.

   (a) Geben Sie die Wahrscheinlichkeitsfunktion von X an. Begründung!

   (b) Wie groß ist die Wahrscheinlichkeit, dass genau 2 Orangen nicht in Ordnung sind?

   (c) Geben Sie $E(X)$ an.

2. Es werden drei Orangen mit Zurücklegen entnommen.

   Sei X die Anzahl der Orangen, die nicht in Ordnung sind.

   (a) Geben Sie die Wahrscheinlichkeitsfunktion von X an. Begründung!

   (b) Wie groß ist die Wahrscheinlichkeit, dass genau 2 Orangen nicht in Ordnung sind?

   (c) Geben Sie $E(X)$ an.

Aufgabe 93
Von einer Maschine wird Joghurt in Becher abgefüllt. Jeder abgefüllte Becher wird von einem Arbeiter vom Band genommen und in einen Korb gestellt. Für den Deckel eines Bechers sind die Zustände defekt und intakt möglich. Die Wahrscheinlichkeit, dass ein Deckel defekt ist, beträgt 0.1. Wir unterstellen, dass diese Wahrscheinlichkeit für jeden Becher konstant ist. Außerdem unterstellen wir, dass die Zustände der Becher voneinander unabhängig sind.

1. Der Arbeiter soll die Maschine anhalten, sobald er einen defekten Becher vom Band genommen hat. Sei X die Anzahl der intakten Becher, die vor dem defekten Becher vom Band genommen wurden.

   (a) Geben Sie die Wahrscheinlichkeitsfunktion von X an.

   (b) Wie groß ist die Wahrscheinlichkeit, dass X Werte annimmt, die größer als 2 sind?

2. Um sich ein Bild vom Zustand der Becher zu machen, werden 10 Becher vom Band genommen. Sei Z die Anzahl der defekten Becher unter den 10.

(a) Geben Sie die Wahrscheinlichkeitsfunktion von Z an.

(b) Bestimmen Sie
$$P(Z > 2).$$

Aufgabe 94

Bei einer Münze beträgt die Wahrscheinlichkeit für das Auftreten von ZAHL gleich p. Die Münze werde nun so oft hintereinander geworfen, bis zum ersten Mal ZAHL auftritt.

Sei X die Anzahl der hierzu benötigten Würfe.

1. Bestimmen Sie die Wahrscheinlichkeitsfunktion von X.

2. Zeigen Sie, dass die Wahrscheinlichkeit, dass bei einer ungeraden Anzahl von Würfen ZAHL zum ersten Mal auftritt,
$$\frac{p}{1-(1-p)^2}$$
beträgt.

3. Bestimmen Sie den Wert von p, für den die unter 2. angegebene Wahrscheinlichkeit 0.6 beträgt.

Aufgabe 95

Ein Würfel wird so oft geworfen, bis zum ersten Mal die 6 fällt.

Sei X die Anzahl der benötigten Würfe.

1. Welche Werte kann X annehmen?

2. Bestimmen Sie die Wahrscheinlichkeitsfunktion von X.

3. Bestimmen Sie die Verteilungsfunktion von X.

Aufgabe 96

In den 432 Jahren von 1500 bis 1931 brach 299 mal ein Krieg aus. Dabei wird eine militärische Aktion Krieg genannt, wenn sie entweder offiziell erklärt wurde und mindestens 50000 Mann daran beteiligt waren, oder sie zu bedeutenden Grenzverschiebungen führte. Die folgende Tabelle gibt die Anzahl der Jahre wieder, in denen 0, 1, 2, 3 oder 4 Kriege ausbrachen:

x_i	n_i
0	223
1	142
2	48
3	15
4	4

1. Stellen Sie dieser empirischen Häufigkeitsverteilung die Stabdiagramme der Poissonverteilung mit den Parametern $\lambda = 0.6$, $\lambda = 0.7$ bzw. $\lambda = 0.8$ gegenüber. Welche dieser Poissonverteilungen passt sich dem Beobachtungsmaterial besser an?

2. Gehen Sie im folgenden von der Poissonverteilung aus, die sich in 1. dem Beobachtungsbefund besser anpasst.

   (a) Wie groß ist die Wahrscheinlichkeit, dass

      i. in einem Jahr höchstens ein Krieg ausbricht

      ii. mehr als 2 Kriege in einem Jahr ausbrechen?

   (b) Welche Anzahl von ausgebrochenen Kriegen würden Sie in einem Jahr erwarten? Vergleichen Sie diesen Wert mit dem arithmetischen Mittel.

Aufgabe 97
An einem Bankschalter werden die Kundenankünfte (Anzahl der pro 10-Minuten-Intervall ankommenden Kunden) beachtet. Für 50 derartige Zeitintervalle erhält man folgende Ergebnisse:

```
1  0  0  0  0  1  2  2  0  2
1  1  0  4  1  0  0  3  0  1
1  0  2  1  1  3  0  2  2  2
0  0  1  2  1  2  1  6  2  0
0  0  1  1  2  1  2  0  1  1.
```

1. Berechnen Sie die relativen Häufigkeiten und stellen Sie diese graphisch in einem Stabdiagramm dar.

2. Stellen Sie dieser empirischen Häufigkeitsverteilung die Stabdiagramme der Poissonverteilung mit den Parametern $\lambda = 1$, $\lambda = 0.5$ bzw. $\lambda = 2$ gegenüber. Welche dieser Poissonverteilungen passt sich dem Beobachtungsmaterial besser an?

3. Gehen Sie im folgenden von der Poissonverteilung aus, die sich in 2. dem Beobachtungsbefund besser anpasst.

   (a) Wie groß ist die Wahrscheinlichkeit, dass in einem 10-Minuten-Intervall

      i. nicht mehr als 2 Personen,

      ii. mehr als 1 Person,

      iii. mindestens 1, aber höchstens 4 Personen ankommen?

(b) Wie viele ankommende Personen würden Sie in einem 10-Minuten-Intervall erwarten? Vergleichen Sie diesen Wert mit dem arithmetischen Mittel.

Aufgabe 98
Neben der Binomialverteilung ist die Normalverteilung für Statistik II sehr wichtig.

1. Welche Bedeutung haben die Parameter μ und σ^2?

2. Geben Sie die Dichtefunktion an!

3. Wie lauten Erwartungswert und Varianz?

4. Welche Verteilung hat
$$Z = \frac{X - \mu}{\sigma},$$
falls X normalverteilt ist mit den Parametern μ und σ^2?

Aufgabe 99
Eine Reifenfirma untersucht die Lebensdauer eines neu entwickelten Reifens. Dabei zeigt sich, dass die ermittelte Lebensdauer der Reifen gut durch eine Normalverteilung mit den Parametern $\mu = 36000$ und $\sigma = 4000$ angenähert werden kann.

1. Wie groß ist die Wahrscheinlichkeit dafür, dass ein zufällig ausgewählter Reifen höchstens 48000 km hält?

2. Welche Lebensdauer wird von 95 Prozent der Reifen nicht überschritten?

3. Wie groß ist die Wahrscheinlichkeit dafür, dass ein zufällig ausgewählter Reifen mehr als 28000 km hält?

4. Welche Lebensdauer wird von 90 Prozent der Reifen überschritten?

5. Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Reifen länger als 28000 km und weniger als 44000 km hält?

6. Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Reifen länger als 30000 km und weniger als 46000 km hält?

7. Welches sind die Grenzen für das 2-fache zentrale Schwankungsintervall?

8. Errechnen Sie das zentrale Schwankungsintervall, in das 95 Prozent der Reifen fallen.

Aufgabe 100

Nehmen wir an, dass das Papiergewicht X (in Gramm) einer Klausur mit $\mu = 90$ und $\sigma = 16$ normalverteilt ist.

1. Wie groß ist die Wahrscheinlichkeit, dass die Klausur mehr als 95 g wiegt?

2. Berechnen Sie $x_{0.75}$!

   Wie kann man diesen Wert interpretieren?

3. Berechnen Sie das zentrale Schwankungsintervall, in das das Gewicht von 80 Prozent der Klausuren fällt!

4. Wie groß ist die Wahrscheinlichkeit, dass die Klausur mehr als 80 g und weniger als 94 g wiegt?

5. Welches Gewicht wird von 99 Prozent der Klausuren nicht überschritten?

Aufgabe 101

Ein Zwischenprodukt muss vom Produktionsbetrieb zum Endmontagebetrieb mit dem LKW transportiert werden. Die Fahrtdauer ist dabei normalverteilt mit Erwartungswert 5 Stunden und Standardabweichung 0.5 Stunden.

1. Berechnen Sie den Median, das untere und das obere Quartil der Fahrtdauer.

2. Berechnen Sie die Wahrscheinlichkeit, dass die Fahrtdauer mindestens 6 Stunden beträgt.

3. Berechnen Sie die Wahrscheinlichkeit, dass die Fahrtdauer zwischen dem Median und dem oberen Quartil liegt.

4. Geben Sie die Grenzen des zweifachen Schwankungsintervalls an.

Aufgabe 102

Eine Berliner Autowerkstatt für Auspuffschnellreparaturen hat über einen längeren Zeitraum anhand betrieblicher Aufzeichnungen ermittelt, dass die Reparaturzeiten in guter Näherung exponentialverteilt sind. Die Abfertigungsrate beträgt $\lambda = 0.75$ [PKW/Stunde], d.h. durchschnittlich benötigt die Werkstatt 80 Minuten für eine Auspuffreparatur.

1. Wie groß ist die Wahrscheinlichkeit, dass die Reparaturzeit unter 80 Minuten bleibt?

2. Welche Reparaturzeit wird in 50 Prozent aller Fälle nicht überschritten?

3. Die Werkstatt wirbt mit dem Slogan: 'In einer Stunde haben Sie für Jahre Ruhe!'.

   Wie groß ist die Wahrscheinlichkeit, dass die Reparaturzeit länger als eine Stunde dauert?

Kapitel 10

Stichproben

Kapitel 11

Stichprobenfunktionen

Aufgabe 103
Das Gewicht X des in '50 g-Packungen' abgefüllten Gewürzpulvers sei normalverteilt mit Erwartungswert 50 und Varianz 4. Wenn bei einer Kontrolle das Gewicht unter 45 g liegt, muss der Hersteller mit einer Strafe rechnen.

1. Wie groß ist die Wahrscheinlichkeit, dass der Hersteller mit einer Strafe rechnen muss?

2. Wie groß ist die Wahrscheinlichkeit, dass das Füllgewicht mehr als 47 aber höchstens 53 g beträgt?

3. Welches Füllgewicht wird von 80 Prozent der Packungen überschritten?

4. Geben Sie die Grenzen des zweifachen Schwankungsintervalls an. Wieviel Prozent der Beobachtungen fallen in dieses Intervall?

5. In welchem zentralen Schwankungsintervall liegt das Füllgewicht mit Wahrscheinlichkeit 0.9?

6. Bei einer neuen Maschine beträgt die Standardabweichung für das Füllgewicht nur noch 1. Welches Gewicht kann der Hersteller für den Erwartungswert einstellen, wenn mit der gleichen Wahrscheinlichkeit wie oben das Füllgewicht über 45 g liegen soll?

Aufgabe 104
Ein Student weiß aus Erfahrung, dass seine Fahrzeit zur Universität normalverteilt ist mit Erwartungswert 30 und Varianz 4.

1. Wie groß ist die Wahrscheinlichkeit, dass die Fahrzeit höchstens 26 Minuten beträgt?

2. Bestimmen Sie die Grenzen des zweifachen Schwankungsintervalls.

3. Welche Fahrzeit wird an 70 Prozent der Tage überschritten?

4. Würde sich der Wert in 3. erhöhen oder verringern, wenn sich die Standardabweichung erhöht?

Aufgabe 105
Eine Urne enthält 10 Kugeln, von denen vier 10 g und sechs 20 g wiegen.

1. Es wird eine Kugel aus der Urne gezogen. Sei X das Gewicht der gezogenen Kugel.

   (a) Bestimmen Sie die Wahrscheinlichkeitsfunktion von X.

   (b) Berechnen Sie $E(X)$ und $Var(X)$.

2. Die Zusammensetzung der Urne sei unbekannt. Um sich ein Bild vom Inhalt der Urne zu machen, werden zwei Kugeln **mit** Zurücklegen aus der Urne gezogen.

   Sei X_1 das Gewicht der ersten gezogenen Kugel und X_2 das Gewicht der zweiten gezogenen Kugel.

   (a) Geben Sie die Wahrscheinlichkeitsfunktion von X_1 und die Wahrscheinlichkeitsfunktion von X_2 an.

   (b) Bestimmen Sie $E(X_1)$, $E(X_2)$, $Var(X_1)$ und $Var(X_2)$.

   (c) Bestimmen Sie die gemeinsame Wahrscheinlichkeitsfunktion von X_1 und X_2.

   (d) Bestimmen Sie mit Hilfe von 2. (c) die Wahrscheinlichkeitsfunktion der Zufallsvariablen $\bar{X} = 0.5 \cdot (X_1 + X_2)$.

   (e) Bestimmen Sie mit Hilfe von 2.(b) $E(\bar{X})$ und $Var(\bar{X})$.
   Hätten Sie diese Werte auch anders bestimmen können?

3. Beantworten Sie die Fragen 2.(a) bis 2.(e) beim Ziehen **ohne** Zurücklegen.

4. An welchen Stellen bestehen Unterschiede zwischen dem Ziehen mit und dem Ziehen ohne Zurücklegen?

Aufgabe 106
Das Gewicht X einer leeren kleinen Coca-Cola Flasche sei normalverteilt mit Erwartungswert 420 g und Standardabweichung 4. Das Gewicht Y des Inhalts einer vollen kleinen Coca-Cola Flasche ist normalverteilt mit Erwartungswert

330 g und Standardabweichung 3. Sei Z das Gewicht einer vollen kleinen Coca-Cola Flasche.

1. Wie ist Z verteilt?

2. Wie groß ist der Erwartungswert und die Varianz des Gewichts einer vollen Coca-Cola Flasche?

3. Wie groß ist die Wahrscheinlichkeit, dass eine volle Coca-Cola Flasche mehr als 760 g wiegt?

4. In einer Kiste befinden sich 5 volle Coca-Cola Flaschen.

   (a) Welches Gesamtgewicht erwarten Sie?

   (b) Wie groß ist die Wahrscheinlichkeit, dass dieses Gesamtgewicht weniger als 3500 g beträgt?

Aufgabe 107
Das Gewicht von Gummibärchen ist normalverteilt mit Erwartungswert 1.8 g und Varianz 0.01. In einer Tüte sind 36 Gummibärchen.

1. Wie groß ist die Wahrscheinlichkeit, dass das Gesamtgewicht aller Gummibärchen in der Tüte mehr als 66 g beträgt?

2. Wie groß sind der Erwartungswert und der Median des Gesamtgewichts der Gummibärchen?

3. Welches Nettogewicht wird von 95 Prozent aller Tüten nicht überschritten?

Aufgabe 108

1. Ein Student wirft zehnmal eine Münze.

   (a) Wie groß ist die erwartete Anzahl von `Kopf`, wenn die Münze fair ist?

   (b) Wie groß ist die Wahrscheinlichkeit, dass bei 10 Würfen einer fairen Münze kein einziges Mal `Kopf` auftritt?

2. Sie werfen eine Münze zehnmal hintereinander. Jedesmal tritt `Kopf` auf. Spricht dieses Ergebnis dafür, dass die Münze fair ist?

3. Sie werfen eine faire Münze hundertmal hintereinander. Wie groß ist approximativ die Wahrscheinlichkeit, dass der Anteil von `Kopf` zwischen 0.4 und 0.6 liegt?

Aufgabe 109
Ein Autohändler weiß aus Erfahrung, dass ein interessierter Kunde mit Wahrscheinlichkeit 0.2 ein Auto kauft. Außerdem weiß er, dass die Kaufentscheidungen unterschiedlicher Kunden voneinander unabhängig sind.

1. An einem Tag waren genau 5 interessierte Kunden in seinem Geschäft.
   (a) Wie groß ist die Wahrscheinlichkeit, dass alle 5 ein Auto gekauft haben?
   (b) Wie groß ist die erwartete Anzahl von Käufern?
2. In einem Monat waren genau 100 interessierte Kunden in seinem Geschäft.
   (a) Wie groß ist die erwartete Anzahl von Käufern unter den 100?
   (b) Wie groß ist approximativ die Wahrscheinlichkeit, dass mindestens 10 ein Auto gekauft haben?

Aufgabe 110
Ein Ladenbesitzer weiß aus langer Erfahrung, dass sein täglicher Umsatz annähernd normalverteilt ist mit Erwartungswert 800 und Standardabweichung 100.
Wie groß ist die Wahrscheinlichkeit, dass der durchschnittliche Umsatz

1. an 25 zufällig ausgewählten Tagen des Jahres
2. an 100 zufällig ausgewählten Tagen des Jahres zwischen 740 und 860 liegt?

Aufgabe 111
Die Anzahl X der Seiten, die mit einer Patrone eines Tintenstrahldruckers gedruckt werden können, sei normalverteilt mit Erwartungswert 6000 und Varianz 250000.

1. Geben Sie das k-fache zentrale Schwankungsintervall an, in das 99 Prozent der Seitenzahlen fallen.
2. Wie groß ist die Wahrscheinlichkeit, dass man mindestens 6600 Seiten drucken kann?
3. Wie groß ist die Wahrscheinlichkeit, dass die durchschnittliche Seitenzahl $\bar{X}$ von 25 Druckpatronen mehr als 5850 Seiten aber höchstens 6220 Seiten beträgt?

4. Wird die Wahrscheinlichkeit unter 3. größer oder kleiner, wenn man die durchschnittliche Seitenzahl $\bar{X}$ von 20 Druckpatronen betrachtet?

   Geben Sie eine kurze verbale Begründung.

Aufgabe 112

1. Welche Bedingungen müssen die Zufallsvariablen $X_1, X_2, \ldots, X_n$ erfüllen, um Zufallsstichprobe genannt zu werden?
2. Stellen Sie alles zusammen, was Ihnen über die Verteilung des arithmetischen Mittels $\bar{X}$ einer Zufallsstichprobe bekannt ist.

Aufgabe 113

Eine Urne enthält 100 Kugeln. Die Hälfte der Kugeln wiegt 30 Gramm und die andere Hälfte 60 Gramm.

Nehmen wir an, ein Statistiker, dem der Inhalt der Urne unbekannt sei, wolle das Durchschnittsgewicht der Kugeln in der Urne schätzen. Dazu entnimmt er der Urne drei Kugeln mit Zurücklegen. Es gibt 8 mögliche Stichproben.

1. Erstellen Sie eine Liste der Stichproben.
2. Wir wollen den Mittelwert $\bar{X}$ und den Median $X_{0.5}$ als Schätzfunktionen für $E(X)$ miteinander vergleichen:
   (a) Berechnen Sie für jede Stichprobe den Wert, den $\bar{X}$ und den Wert, den $X_{0.5}$ annimmt.
   (b) Bestimmen Sie die Wahrscheinlichkeitsfunktion der Schätzfunktion $\bar{X}$ und die Wahrscheinlichkeitsfunktion der Schätzfunktion $X_{0.5}$
   (c) Bestimmen Sie den $E(\bar{X})$ und $Var(\bar{X})$.
   (d) Bestimmen Sie den $E(X_{0.5})$ und $Var(X_{0.5})$.
   (e) Welche der beiden Schätzfunktionen würden Sie vorziehen?

Aufgabe 114

Eine Urne enthält 100 Kugeln. Die Hälfte der Kugeln wiegt 10 Gramm und die andere Hälfte 30 Gramm.

Nehmen wir an, ein Statistiker, dem der Inhalt der Urne unbekannt sei, wolle die Varianz des Gewichts der Kugeln in der Urne schätzen. Dazu entnimmt er der Urne zwei Kugeln mit Zurücklegen.

Es gibt 4 mögliche Stichproben.

1. Erstellen Sie eine Liste der Stichproben.

2. Wir wollen die mittlere quadratische Abweichung D^2 und die Stichprobenvarianz S^2 als Schätzfunktionen für $Var(X)$ miteinander vergleichen:

   (a) Berechnen Sie für jede Stichprobe den Wert, den D^2 und den Wert, den S^2 annimmt.

   (b) Bestimmen Sie die Wahrscheinlichkeitsfunktion der Schätzfunktion D^2 und die Wahrscheinlichkeitsfunktion der Schätzfunktion S^2

   (c) Bestimmen Sie den $E(D^2)$ und $Var(D^2)$.

   (d) Bestimmen Sie den $E(S^2)$ und $Var(S^2)$.

   (e) Welche der beiden Schätzfunktionen würden Sie vorziehen?

Kapitel 12

Schätzung von Parametern

Aufgabe 115
Für den Erwartungswert μ einer Zufallsvariablen werden folgende Schätzfunktionen vorgeschlagen.

$$
\begin{aligned}
T_1 &= \bar{X} \\
T_2 &= X_n \\
T_3 &= \frac{X_1 + X_n}{2} \\
T_4 &= \frac{1}{n-1} \sum_{i=1}^{n} X_i \\
T_5 &= \frac{1}{n^2} \sum_{i=1}^{n} X_i
\end{aligned}
$$

Dabei werde angenommen, dass $X_1, \ldots, X_n$ unabhängig sind. Außerdem soll gelten $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$ für $i = 1, \ldots, n$.

1. Welche der Schätzfunktionen sind erwartungstreu?
2. Welche der erwartungstreuen Schätzfunktionen hat die kleinste Varianz?
3. Welche der Schätzfunktionen sind konsistent?

Aufgabe 116
Zwei Statistiker entwickeln unabhängig voneinander jeweils eine Schätzfunktion für den unbekannten Parameter θ. Für die Schätzfunktion T_1 des ersten Statistikers gilt $E(T_1) = 3 \cdot \theta$ und $Var(T_1) = 1$ und für die Schätzfunktion

T_2 des zweiten Statistikers gilt $E(T_2) = 2 \cdot \theta$ und $Var(T_2) = 9$. Die beiden Statistiker beschließen, eine Linearkombination $T = a \cdot T_1 + b \cdot T_2$ ihrer Schätzfunktionen zu bilden.

1. Welche Bedingungen müssen a und b erfüllen, damit T erwartungstreu für θ ist?

2. Bestimmen Sie a und b so, dass T erwartungstreu und varianzminimal für θ ist.

Aufgabe 117
$X_1, \ldots, X_n$ seien unabhängige, identisch verteilte Zufallsvariablen mit Erwartungswert μ und Varianz σ^2. Der Parameter μ soll geschätzt werden. Hierzu stehen folgende Schätzfunktionen zur Verfügung:

$$\begin{aligned} T_1 &= \frac{1}{2}\bar{X}_1 + \frac{1}{2}\bar{X}_2 \\ T_2 &= \frac{1}{2(n-1)}\sum_{i=1}^{n-1} X_i + 0.5 \cdot X_n \end{aligned}$$

1. Zeigen Sie, dass die beiden Schätzfunktionen erwartungstreu für μ sind.

2. Bestimmen Sie die Varianzen der beiden Schätzfunktionen.

3. Untersuchen Sie die beiden Schätzfunktionen hinsichtlich der Konsistenz.

Aufgabe 118
Sei T_1 eine erwartungstreue Schätzfunktion für den Parameter θ zum Stichprobenumfang n, wobei gilt

$$Var(T_1) = \frac{2}{n}$$

Außerdem sei noch die Schätzfunktion $T_2 = T_1 + \frac{1}{n}$ gegeben.
Kreuzen Sie die wahren Aussagen an.

T_1 ist asymptotisch erwartungstreu für θ	[ ]
T_1 ist konsistent für θ	[ ]
T_2 ist erwartungstreu für θ	[ ]
T_2 ist asymptotisch erwartungstreu für θ	[ ]
T_2 ist konsistent für θ	[ ]

Aufgabe 119
$X_1, \ldots, X_n$ seien unabhängige, identisch mit Erwartungswert μ verteilte Zufallsvariablen. Dann gilt für die Schätzfunktion $\hat{\mu} = X_1$:

$\hat{\mu}$ ist asymptotisch erwartungstreu für μ	[ ]
$\hat{\mu}$ ist konsistent für μ	[ ]
$\hat{\mu}$ ist nicht erwartungstreu für μ	[ ]
Der Bias von $\hat{\mu}$ ist 0	[ ]

Aufgabe 120
Für die Schätzfunktion T_1 gilt $E(T_1) = \theta$ und $Var(T_1) = \dfrac{2}{n}$.

Für die Schätzfunktion T_2 gilt $E(T_2) = \dfrac{n-1}{n}\theta$ und $Var(T_2) = \dfrac{1}{n}$.

Kreuzen Sie die wahren Aussagen an.

T_2 ist asymptotisch erwartungstreu für θ	[ ]
Der Bias von T_2 ist gleich $\frac{\theta}{n}$	[ ]
T_1 ist erwartungstreu für θ	[ ]
T_2 ist effizienter als T_1	[ ]
T_1 ist konsistent für θ	[ ]

Aufgabe 121
T sei eine Schätzfunktion für den Parameter θ. Welche Aussage kann man mit Hilfe der Varianz $Var(T)$ von T machen?

Wie sehr T um $E(T)$ streut	[ ]
Wie sehr T um θ streut	[ ]

Aufgabe 122
T sei eine Schätzfunktion für den Parameter θ. Welche Aussage kann man mit Hilfe des mittleren quadratischen Fehlers $MSE(T, \theta)$ von T_n machen?

Wie sehr T um $E(T)$ streut	[ ]
Wie sehr T um θ streut	[ ]

Aufgabe 123

1. Was versteht man unter dem Selektionsbias? Geben Sie Beispiele an.

2. Was versteht man unter dem Nonresponse Bias? Geben Sie Beispiele an.

Aufgabe 124

Eine Urne enthält 10 Kugeln. Sie kann zwei unterschiedliche Zusammensetzungen aufweisen:

Zustand I: 7 Kugeln wiegen 20 g , 3 Kugeln wiegen 10 g

Zustand II: 4 Kugeln wiegen 20 g , 6 Kugeln wiegen 10 g

1. Bestimmen Sie das jeweilige Durchschnittsgewicht θ der Kugeln in der Urne.

2. Der Zustand der Urne und damit auch der Wert von θ seien unbekannt. Um θ zu schätzen, entnehmen wir der Urne zwei Kugeln mit Zurücklegen.

   (a) Stellen Sie die Wahrscheinlichkeiten der einzelnen Stichproben in Abhängigkeit von den Werten von θ in einer Tabelle zusammen.

   (b) Geben Sie die Maximum-Likelihood-Schätzfunktion für θ an.

   (c) Bei einer Ziehung wiegen beide Kugeln 20 g. Wie lautet der M-L-Schätzwert für θ?

   (d) Ist der M-L-Schätzer für θ erwartungstreu?

Aufgabe 125

Eine Urne enthält 5 Kugeln. Sie kann zwei unterschiedliche Zusammensetzungen aufweisen:

Zustand I: 1 Kugel ist weiß , 4 Kugeln sind schwarz

Zustand II: 3 Kugeln sind weiß , 2 Kugeln sind schwarz

Sei p der Anteil der weißen Kugeln in der Urne.
Der Zustand der Urne und damit auch der Wert von p seien unbekannt. Um p zu schätzen, entnehmen wir der Urne zwei Kugeln mit Zurücklegen.

1. Stellen Sie die Wahrscheinlichkeiten der einzelnen Stichproben in Abhängigkeit von den Werten von p in einer Tabelle zusammen.

2. Geben Sie die Maximum-Likelihood-Schätzfunktion für p an.

3. Ist der M-L-Schätzer für p erwartungstreu?
4. Bei einer Ziehung sind beide Kugeln weiß.
   (a) Stellen Sie die Likelihood-Funktion graphisch dar.
   (b) Wie lautet der M-L-Schätzwert für p?

Aufgabe 126
In einer Schachtel befinden sich 5 Kugeln, die entweder weiß oder rot sind. Die Anzahl θ der roten Kugeln ist unbekannt (θ=0, 1, 2, 3, 4 oder 5).
Ein Statistiker zieht mit Zurücklegen 3 Kugeln aus der Urne und erhält folgendes Ergebnis:

```
rot  weiss  rot
```

1. Wie groß ist die Wahrscheinlichkeit dieses Ereignisses, wenn $\theta = 2$ ist?
2. Wie lautet die Likelihoodfunktion für θ aufgrund dieses Ergebnisses?
3. Was gibt diese Likelihoodfunktion an?
4. Skizzieren Sie $L(\theta)$.
5. Bestimmen Sie den Maximum-Likelihood-Schätzwert für θ.

Aufgabe 127
Maximum-Likelihood-Schätzer sind

manchmal erwartungstreu	[ ]
immer erwartungstreu	[ ]
nie erwartungstreu	[ ]

Aufgabe 128
Eine Urne enthält 10 Geldstücke. Sie kann zwei unterschiedliche Zusammensetzungen aufweisen:

Zustand I: 3 mal 1 EURO und 7 mal 2 EURO

Zustand II: 6 mal 1 EURO und 4 mal 2 EURO

Von Interesse ist der Gesamtbetrag θ des Geldes in der Urne.
Der Zustand der Urne und damit auch der Wert von θ seien unbekannt. Um θ zu schätzen, entnehmen wir der Urne zwei Geldstücke mit Zurücklegen.

1. Stellen Sie die Wahrscheinlichkeiten der einzelnen Stichproben in Abhängigkeit von den Werten von θ in einer Tabelle zusammen.

2. Geben Sie die Maximum-Likelihood-Schätzfunktion für θ an.

3. Bei einer Ziehung werden zwei 1 EURO Stücke gezogen. Wie lautet der M-L-Schätzwert für θ?

4. Ist der M-L-Schätzer für θ erwartungstreu?

Aufgabe 129
Bestimmen Sie den M-L-Schätzer $\hat{\lambda}$ des Parameters λ für eine Zufallsstichprobe vom Umfang n aus einer mit dem Parameter λ poissonverteilten Grundgesamtheit. Ist der M-L-Schätzer $\hat{\lambda}$ erwartungstreu für λ? Ist der M-L-Schätzer $\hat{\lambda}$ konsistent für λ?

Aufgabe 130
Bestimmen Sie den M-L-Schätzer $\hat{p}$ des Parameters p für eine Zufallsstichprobe vom Umfang n aus einer mit dem Parameter p bernoulliverteilten Grundgesamtheit. Ist der M-L-Schätzer $\hat{p}$ erwartungstreu für p? Ist der M-L-Schätzer $\hat{p}$ konsistent für p?

Aufgabe 131
Die Firma MFSL stellte bis Anfang des Jahres 2000 audiophile CDs mit Goldbeschichtung her. Nachdem die Firma Konkurs anmelden musste, wurden diese CDs zu begehrten Sammlerstücken. Ein Statistiker beschließt, seine MFSL GOLD-CD von DARK SIDE OF THE MOON von PINK FLOYD auf EBAY zu versteigern. Um eine Vorstellung vom realisierbaren Preis zu erhalten, beobachtet er den Markt. In der zweiten Februarwoche des Jahres 2001 wurden 9 CDs zu folgenden Höchstgeboten in Dollar ersteigert:

```
51  56  57  48  49  61  46  50  59
```

1. Erstellen Sie einen Boxplot der Daten.

2. Wir gehen davon aus, dass die Höchstgebote normalverteilt sind. Bestimmen Sie für den unbekannten Parameter μ dieser Normalverteilung ein Konfidenzintervall, das μ mit der Wahrscheinlichkeit 0.95 überdeckt.

Aufgabe 132
Ein Statistiker sieht sich in der Zeit vom 2.2.2001 bis zum 19.2.2001 alle Folgen der Show WER WIRD MILLIONÄR an und notiert sich am Ende der Sendung den realisierten Gesamtgewinn des Tages. Es ergaben sich folgende Werte (in Tausend DM):

`34 17 96 33 189 282 33 66 64`

1. Erstellen Sie einen Boxplot der Daten.
2. Wir gehen davon aus, dass die Gesamtgewinne normalverteilt sind. Bestimmen Sie für den unbekannten Parameter μ dieser Normalverteilung ein Konfidenzintervall, das μ mit der Wahrscheinlichkeit 0.95 überdeckt.

Aufgabe 133
Von den Fragebögen, die in der Vorlesung abgegeben wurden, wurden 9 zufällig ausgewählt und die Körpergröße der Studenten ermittelt. Der Mittelwert betrug 178 cm und die Varianz betrug 88.889.

Gehen Sie im folgenden davon aus, dass die Körpergröße der Studenten normalverteilt ist.

1. Stellen Sie ein Konfidenzintervall zum Konfidenzniveau 0.95 für den Erwartungswert der Körpergröße der Studenten auf.
2. Wie ändert sich das Ergebnis unter 1., wenn die Varianz der Grundgesamtheit bekannt ist und 96.04 beträgt?
3. Wie groß muss unter 2. das Konfidenzniveau sein, damit das Intervall nur halb so lang ist?

Aufgabe 134
Ein Statistiker berechnet aus dem gleichen Datensatz 3 Konfidenzintervalle für μ bei Normalverteilung mit bekanntem $\sigma^2 = 4$. Das erste Konfidenzintervall zum Konfidenzniveau $1 - \alpha = 0.95$ hat die Länge l_1, das zweite zum Konfidenzniveau $1 - \alpha = 0.9$ hat die Länge l_2 und das dritte zum Konfidenzniveau $1 - \alpha = 0.99$ hat die Länge l_3.

Ordnen Sie die drei Konfidenzintervalle nach ihrer Länge. Beginnen Sie mit dem kürzesten.

Aufgabe 135
Ein Statistiker konstruiert 20 unabhängige Konfidenzintervalle zum Konfidenzniveau 0.95. Wie groß ist die Wahrscheinlichkeit, dass

1. alle
2. keins
3. die Hälfte

den wahren Wert des Parameters überdecken?

Aufgabe 136
Bei einer Befragung in der Statistik I Vorlesung gaben 13 von den 250 Befragten an, dass sie FDP wählen würden, wenn am nächsten Sonntag Bundestagswahl wäre. Unterstellen Sie im folgenden, dass diese Umfrage repräsentativ für die Studienanfänger in Wirtschaftswissenschaften in Deutschland ist.

1. Stellen Sie ein Konfidenzintervall zum Konfidenzniveau 0.95 für den unbekannten Anteil p der FDP-Wähler auf.

2. In einer Statistikklausur sollten Studenten mit den obigen Zahlenangaben und Annahmen ein Konfidenzintervall zum Konfidenzniveau 0.99 für den unbekannten Anteil p der FDP-Wähler aufstellen.

   Es gab folgende Lösungen

   (a) $[0.016, 0.088]$
   (b) $[0.032, 0.072]$
   (c) $[0.020, 0.080]$
   (d) $[0.017, 0.087]$

   Geben Sie mit Begründung an, bei welchen es sich nicht um ein Konfidenzintervall zum Konfidenzniveau 0.99 handeln kann.

   Sie sollten nicht rechnen, sondern 1. beachten und denken.

3. Nehmen wir an, Sie wollen ein Konfidenzintervall zum Konfidenzniveau 0.95 für den unbekannten Anteil p der FDP Wähler aufstellen, dessen Länge höchstens 0.02 beträgt.

   Wieviele Leute müssen Sie mindestens befragen,

   (a) wenn Sie keinerlei Vorinformation besitzen,
   (b) wenn Sie wissen, dass der Anteil der FDP Wähler höchstens 7 Prozent beträgt?

Aufgabe 137

1. Eine Partei will ihren Wähleranteil schätzen. Dazu werden 100 Personen befragt, von denen 25 die Partei wählen würden.

   (a) Stellen Sie ein Konfidenzintervall für den Anteil der Wähler der Partei in der Bevölkerung zum Konfidenzniveau 0.95 auf.
   (b) Wieviele Personen müssten befragt werden, um ein Konfidenzintervall zum Konfidenzniveau 0.95 zu erhalten, dessen Länge halb so groß ist?

2. Es soll ein Konfidenzintervall für p zum Konfidenzniveau 0.99 bei $n = 500$ aufgestellt werden. Ist dieses länger, wenn $\hat{p}$ gleich 0.3 oder wenn $\hat{p}$ gleich 0.7 ist?

Aufgabe 138
Bei einer Befragung in einer Statistik-I-Vorlesung gaben von den 225 Studenten 90 an, bei der nächsten Bundestagswahl die SPD zu wählen.

1. Stellen Sie ein Konfidenzintervall zum Konfidenzniveau 0.99 für den unbekannten Anteil p der SPD-Wähler auf.

2. Wie groß muss der Stichprobenumfang n mindestens gewählt werden, damit die Länge des Konfidenzintervalls den Wert 0.02 bei einem Konfidenzniveau 0.99 nicht überschreitet, wenn vermutet wird, dass die SPD höchstens 40 Prozent der Stimmen erhält?

3. Wie groß kann die Länge des Konfidenzintervalls bei einem Niveau von 0.95 und einem Stichprobenumfang von $n = 5000$ höchstens werden, wenn die Vorinformation unter 2. berücksichtigt wird?

Aufgabe 139
In Berlin gab es Mitte der Siebziger Jahre ein Papier mit dem Titel: 'Wie man beim Aufstellen von Konfidenzintervallen das Denken vermeidet.' Das Papier bestand aus einem Flussdiagramm, in dem die einzelnen Möglichkeiten strukturiert dargestellt wurden.

Versuchen Sie ein solches Flußdiagramm zu erstellen.

Die erste Entscheidung könnte sein: p-Problem oder μ-Problem.

Kapitel 13

Grundbegriffe statistischer Tests

Aufgabe 140
Erklären Sie die Begriffe Fehler 1.Art und Fehler 2.Art

1. an den Begriffen Produzentenrisiko und Konsumentenrisiko,
2. am Beispiel eines Feuerwehralarms (blinder Alarm und notwendiger, aber unterlassener Alarm).

Aufgabe 141
In einem Gerichtssaal wartet ein Angeklagter auf sein Urteil. Das Gericht hat sich zur Beratung zurückgezogen und muss über schuldig und nicht schuldig entscheiden. Welche beiden Fehler können die Richter bei ihrer Entscheidung begehen?
In der Justiz gilt der Grundsatz: 'Im Zweifel für den Angeklagten'. Wie müssten die Wahrscheinlichkeiten für die möglichen Fehler des Gerichts aussehen, wenn der obige Grundsatz gültig wäre?
Vergleichen Sie die Situation mit einem statistischen Test.

Aufgabe 142
Ein Student der Statistik II hat einen Test konstruiert, bei dem $\alpha > 1 - \beta$ gilt mit:

P(H ablehnen|H richtig) $= \alpha$

P(H nicht ablehnen|H falsch)$=\beta$

Ist ein solcher Test sinnvoll?

Aufgabe 143
Ein Nichtstatistiker schlägt für jedes Testproblem vor, einer Urne mit 5 weißen und 95 roten Kugeln eine Kugel zufällig zu entnehmen und H ablehnen,

wenn die gezogene Kugel weiß ist. Bestimmen Sie die Wahrscheinlichkeit für den Fehler 1.Art und die Güte des Tests.

Aufgabe 144
Um einen Test durchzuführen, wirft ein fauler Statistiker vier faire Münzen und entscheidet sich für die Gegenhypothese, wenn bei allen Münzen `Kopf` erscheint. In allen anderen Fällen entscheidet er sich für die Hypothese. Geben Sie die Wahrscheinlichkeit für den Fehler 1. Art an.

Aufgabe 145

1. Eine Hypothese wird zum Niveau $\alpha = 0.01$ abgelehnt. Wird sie dann auch zum Niveau $\alpha = 0.05$ abgelehnt?
2. Eine Hypothese wird zum Niveau $\alpha = 0.05$ abgelehnt. Wird sie dann auch zum Niveau $\alpha = 0.01$ abgelehnt?
3. Eine Hypothese wird zum Niveau $\alpha = 0.01$ nicht abgelehnt. Wird Sie dann auch zum Niveau 0.05 nicht abgelehnt?
4. Eine Hypothese wird zum Niveau $\alpha = 0.05$ nicht abgelehnt. Wird Sie dann auch zum Niveau 0.01 nicht abgelehnt?

Aufgabe 146
Die beiden möglichen Zustände einer Grundgesamtheit seien chararkterisiert durch den Parameter θ, wobei θ die Werte 0.1 oder 0.2 annehmen kann. Für eine Stichprobe vom Umfang 1 sind folgende Wahrscheinlichkeitsverteilungen bekannt:

x \ θ	0.1	0.2
0	0.10	0.95
1	0.90	0.05

Es soll getestet werden $H : \theta = 0.1$ gegen $G : \theta = 0.2$. Die Entscheidungsregel lautet:

Entscheidung für H, wenn gilt $x = 1$.

Entscheidung für G, wenn gilt $x = 0$.

Sei α die Wahrscheinlichkeit des Fehlers 1.Art und β die Wahrscheinlichkeit des Fehlers 2.Art.
Welche der folgenden Aussagen sind dann richtig?

$\alpha + \beta = 0.15$ [ ]

$\alpha = 0.05$ [ ]

$\beta = 1 - \alpha$ [ ]

$\beta < \alpha$ [ ]

Kapitel 14

Das Einstichprobenproblem

Aufgabe 147
Wir kommen noch einmal auf Aufgabe 131 zurück. Der Statistiker will seine CD nur ins Ebay stellen, wenn der erwartete Höchstpreis mehr als 50 Dollar beträgt. Führen Sie einen geeigneten Test zum Niveau $\alpha = 0.05$ durch.

Aufgabe 148
Wir kommen noch einmal auf Aufgabe 132 zurück. Der Statistiker will überprüfen, ob der erwartete Gewinn eines Tages mehr als 100000 DM beträgt. Führen Sie einen geeigneten Test zum Niveau $\alpha = 0.05$ durch.

Aufgabe 149
Tabelle 14.1 gibt die Häufigkeitsverteilung der Anzahl der Knaben unter den ersten vier Kindern in 3343 schwedischen Familien wieder:

Tabelle 14.1: Anzahl der Knaben unter den ersten vier Kindern in 3343 schwedischen Familien

Anzahl Kinder	abs. Häufigkeit
0	183
1	789
2	1250
3	875
4	246

(Quelle: A.W.F. Edwards, M. Fraccaro: Distribution and Sequences of Sexes in a selected Sample of Swedish Families. Annals of Human Genetics. Cambridge University Press, 1960, vol. 24, S.246)

Sei X die Anzahl der Knaben unter den vier Kindern.
Testen Sie zum Niveau $\alpha = 0.05$, ob X binomialverteilt ist.

Aufgabe 150
Bei der Fußball WM 98 in Frankreich fanden 64 Spiele statt. In der folgenden Liste finden Sie die Anzahl der pro Spiel erzielten Tore.

3	4	4	2	0	1	3	5	4	0	1	1	4	2	1	2
2	3	2	3	2	4	1	0	1	4	5	4	5	3	1	3
3	2	3	3	3	4	7	4	4	2	2	1	1	3	2	2
1	5	0	5	3	3	1	4	0	5	3	3	2	3	3	3

1. Erstellen Sie die Häufigkeitstabelle.
2. Zeichnen Sie das Stabdiagramm.
3. Testen Sie zum Niveau $\alpha = 0.05$, ob die Anzahl der pro Spiel erzielten Tore poissonverteilt ist.

Aufgabe 151
Warum sind die Statistiken

$$U = \sum_{i=1}^{k} (n_i - n \cdot p_i)$$

und

$$V = \sum_{i=1}^{k} (n_i - n \cdot p_i)^2$$

für einen Anpassungstest nicht geeignet?

Aufgabe 152
Ein Statistiker führt einen Chiquadratanpassungstest auf Normalverteilung mit bekanntem μ und unbekanntem σ^2 durch. Er bildet 6 Klassen. Geben Sie den kritischen Wert an, wenn das Niveau $\alpha = 0.05$ beträgt.

Kapitel 15

Das Zweistichprobenproblem

Aufgabe 153
Im Rahmen der PISA-Studie wurden die Leistungen in den Bereichen `Lesekompetenz` und `Mathematische Grundbildung` bestimmt. Tabelle 15.1 zeigt die Punkte von 10 Ländern.

Tabelle 15.1: Punkte in den Bereichen `Lesekompetenz` und `Mathematische Grundbildung`

Land	Lesekompetenz	Mathematische Grundbildung
BR	396	334
D	484	490
F	505	517
I	487	457
FL	483	514
L	441	446
A	507	515
PL	479	470
P	470	454
E	493	476

Quelle: Deutsches PISA-Konsortium: PISA 2000

1. Wir wollen testen, ob die erwartete Punktezahl in beiden Bereichen gleich hoch ist.

Die Differenzen aus den Punkten im Bereich `Mathematische Grundbildung` und den Punkten im Bereich `Lesekompetenz` betragen:

```
> d
 BR   D   F   I  FL   L   A  PL   P   E
-62   6  12 -30  31   5   8  -9 -16 -17
```

2. Erstellen Sie den Boxplot der Differenzen.
3. Wir betrachten zunächst den Vorzeichentest.
   (a) Bestimmen Sie den Wert der Teststatistik des Vorzeichentests.
   (b) Spricht der Wert der Teststatistik des Vorzeichentests für oder gegen H_0 ($\alpha = 0.05$)?
   (c) Bestimmen Sie den Wert der Überschreitungswahrscheinlichkeit des zweiseitigen Tests.
4. Wir betrachten nun den Vorzeichen-Rang-Test.
   (a) Bestimmen Sie den Wert der Teststatistik des Vorzeichen-Rang-Test.
   (b) Spricht der Wert der Teststatistik des Vorzeichen-Rang-Test. für oder gegen H_0 ($\alpha = 0.05$)?
5. Schauen wir uns nun den t-Test an.
   (a) Bestimmen Sie den Wert der Teststatistik des t-Tests.
   (b) Spricht der Wert der Teststatistik des t-Tests für oder gegen H_0 ($\alpha = 0.05$)?
6. Welchen der beiden Tests würden Sie wählen?

Aufgabe 154
Es soll untersucht werden, ob die erwartete Anzahl geschossener Tore in beiden Halbzeiten eines Bundesligaspiels gleich groß ist. Zur Beantwortung dieser Frage wurden 10 Bundesligaspiele der Saison 2002/2003 zufällig ausgewählt. Die folgende Tabelle 15.2 gibt für jedes Spiel die Anzahl der in beiden Halbzeiten geschossenen Tore wieder.

Tabelle 15.2: Anzahl Tore in erster und zweiter Halbzeit

Spiel	1	2	3	4	5	6	7	8	9	10
Anzahl Tore 1. Halbzeit	1	2	4	0	0	1	0	3	1	0
Anzahl Tore 2. Halbzeit	3	3	2	3	4	6	2	2	3	1

1. Bestimmen Sie den Wert der Teststatistik des Vorzeichentests, wenn man die Differenz aus 2. Halbzeit und 1. Halbzeit betrachtet.
2. Bestimmen Sie den Wert der Überschreitungswahrscheinlichkeit des zweiseitigen Tests.
3. Wird zum Niveau 0.05 die Hypothese abgelehnt, dass die erwartete Anzahl von geschossenen Toren in beiden Halbzeiten identisch ist?

Aufgabe 155
Im Grundfach *Schwimmen* legten in einem Wintersemester in Theorie und Praxis bei einem Dozenten 19 Studenten die praktische und theoretische Prüfung ab. Für jeden Studenten wurde die Note X in Theorie und die Note Y in Praxis bestimmt. Es soll überprüft werden, ob die Studenten in Theorie und Praxis gleich gut sind.
Es ergaben sich folgende Werte für die Differenz $D = X - Y$:

```
 0.6  0.3  0.3  0.3  0.6 -0.2 0.0  1.0  0.3  0.0
 0.2 -0.6  0.0 -0.3  0.2  0.4  1.8  1.3 -0.8
```

1. Es soll der t-Test durchgeführt werden.
   (a) Auf welchen Annahmen beruht der t-Test?
   (b) Geben Sie die Hypothesen an.
   (c) Führen Sie den t-Test zum Niveau $\alpha = 0.05$ durch.
   (d) Angenommen, der Student hätte überprüfen wollen, ob die Note Y in Theorie im Mittel schlechter ist als die Note X in Praxis.
   Zu welcher Entscheidung wäre er in diesem Fall gekommen?
2. Ein anderer Student schlägt den Vorzeichentest vor, wobei er alle Nullen aus der Stichprobe streicht. Zu welcher Entscheidung kommt der Vorzeichentest zum Niveau 0.05?

3. Ein weiterer Student schlägt den Vorzeichen-Rang-Test vor. Zu welcher Entscheidung kommt der Vorzeichen-Rang-Test zum Niveau 0.05?

Aufgabe 156

Im Rahmen der PISA-Studie wurde in den einzelnen Ländern die durchschnittliche Klassengröße ermittelt. Wir bilden zwei Klassen, wobei die Länder der ersten Klasse eine kleine und die Länder der zweiten Klasse eine hohe Klassengröße besitzen. Aus jeder Klasse wurden jeweils 4 Länder ausgewählt. Die Tabelle 15.3 zeigt die Ergebnisse im Bereich `Naturwissenschaftliche Grundbildung` dieser Länder.

Tabelle 15.3: Ergebnisse im Bereich Naturwissenschaftliche Grundbildung

klein	hoch
496	487
481	513
476	532
500	499

Quelle: Deutsches PISA-Konsortium: PISA 2000

1. Wir überprüfen, ob die erwartete Punktezahl in den beiden Gruppen identisch ist.

   (a) Wir betrachten zuerst den Wilcoxon-Rangsummentest.

      i. Bestimmen Sie den Wert der Teststatistik.

      ii. Zeigen Sie, dass die Überschreitungswahrscheinlichkeit beim zweiseitigen Test 0.2 beträgt.

      iii. Zu welcher Entscheidung kommt der Wilcoxon-Test zum Signifikanzniveau $\alpha = 0.05$? Begründen Sie Ihre Aussage.

   (b) Schauen wir uns nun den t-Test und den F-Test an. Bei den Ländern mit geringer Klassengröße beträgt der Mittelwert der Punkte 488.25 und bei den Ländern mit hoher Klassengröße beträgt er 507.75. Bei den Ländern mit geringer Klassengröße beträgt die Stichprobenvarianz der Punkte 133.583 und bei den Ländern mit hoher Klassengröße beträgt sie 374.25.

      i. Auf welchen Annahmen beruht der t-Test?

      ii. Bestimmen Sie den Wert der Teststatistik des t-Tests.

iii. Zu welcher Entscheidung kommt der zweiseitige t-Test zum Niveau 0.05?

iv. Auf welchen Annahmen beruht der F-Test?

v. Bestimmen Sie den Wert der Teststatistik des F-Tests.

vi. Zu welcher Entscheidung kommt der zweiseitige F-Test zum Niveau 0.05?

Aufgabe 157

Ein Statistiker will untersuchen, ob sich das erwartete Höchstgebot für die MFSL GOLD CD DARK SIDE OF THE MOON von Februar 2001 zu März 2001 geändert hat. Hierzu wählt er die ersten 5 CDs des Monats Februar und die ersten 5 CDs des Monats März aus und ermittelt deren Höchstgebot. Es ergaben sich folgende Werte:

Februar:	51 56 57 48 49
März:	46 50 58 79 82

1. Begründen Sie, warum es sich um ein unverbundenes Zweistichprobenproblem handelt.

2. Erstellen Sie die Boxplots der beiden Stichproben.

3. Führen Sie den Wilcoxon Rangsummentest durch ($\alpha = 0.05$).

4. Führen Sie den t-Test durch ($\alpha = 0.05$).

5. Für welchen Test entscheiden Sie sich?

Aufgabe 158

Von den Studenten, die im WS 91/92 die Statistik I Klausur geschrieben haben, wurden 5 zufällig ausgewählt und von den Studenten, die im SS 92 die Statistik II Klausur geschrieben haben, wurden 5 zufällig ausgewählt. Bei allen 10 Studenten wurde die Punktezahl in der jeweiligen Klausur bestimmt. Es ergaben sich folgende Werte:

Statistik I:	59 37 45 80 30
Statistik II:	66 61 56 76 96

Es soll untersucht werden, ob sich die Punktezahlen zwischen der Statistik I und der Statistik II Klausur unterscheiden.

1. Begründen Sie, warum es sich um ein unverbundenes Zweistichprobenproblem handelt.

2. Erstellen Sie die Boxplots der beiden Stichproben.
3. Begründen Sie, warum der Wilcoxon Rangsummentest besser geeignet ist als der t-Test.
4. Führen Sie den Wilcoxon Rangsummentest durch ($\alpha = 0.05$).
5. Führen Sie auch den t-Test durch ($\alpha = 0.05$).

Aufgabe 159
Im Rahmen der PISA-Studie wurde in den einzelnen Ländern ermittelt, wie hoch der Zeitaufwand für Hausaufgaben beträgt. Dabei wurde unterschieden zwischen `geringem` und `hohem` Zeitaufwand. Tabelle 15.4 zeigt die Ergebnisse im Bereich `Lesekompetenz` von jeweils 5 Ländern.

Tabelle 15.4: Ergebnisse im Bereich `Lesekompetenz` von jeweils 5 Ländern im Rahmen der PISA-Studie

gering	492	494	507	516	522
hoch	474	479	480	487	493

Quelle: Deutsches PISA-Konsortium: PISA 2000

1. Wir wollen überprüfen, ob die erwartete Punktezahl in den beiden Gruppen identisch ist.
   (a) Wir betrachten zuerst den Wilcoxon-Rangsummentest.
      i. Bestimmen Sie den Wert der Teststatistik.
      ii. Bestimmen Sie die Überschreitungswahrscheinlichkeit beim zweiseitigen Test.
      iii. Zu welcher Entscheidung kommt der Wilcoxon-Test zum Signifikanzniveau $\alpha = 0.05$? Begründen Sie Ihre Aussage.
   (b) Schauen wir uns nun den t-Test an.
      i. Auf welchen Annahmen beruht der t-Test?
      ii. Bestimmen Sie den Wert der Teststatistik des t-Tests.
      iii. Führen Sie den t-Test zum Niveau $\alpha = 0.05$ durch.
   (c) Schauen wir uns nun den F-Test an.
      i. Auf welchen Annahmen beruht der F-Test?

ii. Bestimmen Sie den Wert der Teststatistik des F-Tests.

iii. Führen Sie den zweiseitigen F-Test zum Niveau $\alpha = 0.05$ durch.

Aufgabe 160
Was ist der Unterschied zwischen dem verbundenen und dem unverbundenen Zweistichprobenproblem? Geben Sie für beide jeweils ein Anwendungsbeispiel an! Geben Sie für beide jeweils zwei Tests an. Gehen Sie dabei auf die Annahmen ein, die den Tests zugrundeliegen!

Aufgabe 161
Wiederholung, Blockbildung und Randomisieren sind drei Prinzipien der Versuchsplanung.

1. Stellen Sie dar, inwieweit diese Prinzipien zur Verbesserung der Auswertung beitragen.

2. Stellen Sie dar, inwieweit diese Prinzipien im unverbundenen und verbundenen Zweistichprobenproblem relevant sind.

Aufgabe 162
Gebe Sie bei jedem der folgenden Fälle an, ob es sich um ein verbundenes oder um ein unverbundenes Zweistichprobenproblem handelt.

1. Um zu überprüfen, ob sich die Reaktionszeit vor dem Essen von der Reaktionszeit nach dem Essen unterscheidet, werden 10 Personen zufällig ausgewählt und ihre Reaktionszeit vor und nach dem Essen verglichen.

2. Es soll untersucht werden, ob sich Krankenhäuser hinsichtlich der Verweildauer nach einer bestimmten Operation unterscheiden. Hierzu werden in jedem Krankenhaus die Verweilzeiten von jeweils 5 Patienten bestimmt.

3. Es soll untersucht werden, ob BWL-Studenten in der Statistik II Klausur besser abschneiden als VWL-Studenten. Dazu werden jeweils 10 Klausuren zufällig ausgewählt und die Punktezahl notiert.

Kapitel 16

Einfaktorielle Varianzanalyse

Aufgabe 163
Im Rahmen der PISA-Studie wurde in den einzelnen Ländern die durchschnittliche Klassengröße ermittelt. Wir bilden drei Klassen, wobei die Länder der ersten Klasse eine kleine, die Länder der zweiten Klasse eine mittlere und die Länder der dritten Klasse eine hohe Klassengröße besitzen. Aus jeder Klasse wurden jeweils 4 Länder ausgewählt. Tabelle 16.1 zeigt die Ergebnisse im Bereich `Naturwissenschaftliche Grundbildung` dieser Länder.

Tabelle 16.1: Punkte in im Bereich Naturwissenschaftlicher Grundbildung in drei Gruppen

klein	496	481	476	500
mittel	375	550	552	422
hoch	487	513	532	499

Quelle: Deutsches PISA-Konsortium: PISA 2000

1. Stellen Sie die ANOVA-Tabelle auf und interpretieren Sie sie.
2. Testen Sie zum Niveau $\alpha = 0.05$, ob sich die drei Gruppen hinsichtlich der Lage unterscheiden.

Aufgabe 164
Tabelle 16.2 enthält die Arbeitslosenquote im Jahr 2000 in 15 Ländern.

Tabelle 16.2: Arbeitslosenquote in ausgewählten Ländern.

Land	Arbeitslo-senquote	Land	Arbeitslo-senquote	Land	Arbeitslo senquote
Deutschland	7.9	Bulgarien	16.9	USA	4.0
Frankreich	9.3	Kroatien	16.1	Kanada	6.8
Italien	10.4	Rumänien	7.2	Japan	4.7
Österreich	3.7	Russland	13.4	Korea	4.1
England	5.4	Tschechien	8.8	Australien	6.6

Die Daten stammen aus 3 Gruppen. Die ersten 5 Länder sind Mitglied der EU und bilden die erste Gruppe. Die mittleren 5 Länder gehören zu Europa und nicht zur EU und bilden die zweite Gruppe. Die letzten 5 Länder gehören nicht zu Europa und bilden die dritte Gruppe. Wir gehen im Folgenden davon aus, dass es sich um eine Stichprobe handelt.

Es soll untersucht werden, ob sich die Verteilung des Merkmals `Arbeitslosenquote` in den 3 Gruppen unterscheidet. Die folgende Tabelle enthält die Ergebnisse der Varianzanalyse. Es fehlen aber einige Zahlen.

Quelle der Variation	Quadrat-summen	Freiheits-grade	Mittlere Quadratsummen	F
zwischen den Gruppen	138.745			
innerhalb der Gruppen	112.852			
Gesamt				

1. Ergänzen Sie die Tabelle.
2. Vollziehen Sie nach, wie die Quadratsumme zwischen den Gruppen bestimmt wurde.
3. Überprüfen Sie zum Niveau $\alpha = 0.05$, ob sich die Erwartungswerte der drei Gruppen unterscheiden.

Aufgabe 165
Im Rahmen der PISA-Studie wurde in den einzelnen Ländern ermittelt, wie hoch der Zeitaufwand für Hausaufgaben beträgt. Dabei wurde unterschieden

zwischen `geringem`, `mittlerem` und `hohem` Zeitaufwand. Tabelle 16.3 zeigt die Ergebnisse im Bereich `Lesekompetenz` von jeweils 5 Ländern.

Tabelle 16.3: im Bereich Lesekompetenz von jeweils 5 Ländern im Rahmen der PISA-Studie

gering	492	494	507	516	522
mittel	396	497	525	529	470
hoch	474	479	480	487	493

Quelle: Deutsches PISA-Konsortium: PISA 2000

1. Stellen Sie die ANOVA-Tabelle auf und interpretieren Sie sie.
2. Testen Sie zum Niveau $\alpha = 0.05$, ob sich die drei Gruppen hinsichtlich der Lage unterscheiden.

Kapitel 17

Unabhängigkeit und Homogenität

Aufgabe 166
Studienanfänger wurden gefragt, ob sie noch bei den Eltern wohnen, und welche Partei sie wählen würden. Tabelle 17.1 zeigt die Ergebnisse der Umfrage.

Tabelle 17.1: Ergebnis einer Umfrage

Partei Wohnen bei den Eltern	CDU	SPD	FDP	Grüne
ja	35	25	5	10
nein	40	20	15	25

1. Bestimmen Sie die Anteile der Parteien in den beiden Gruppen und berücksichtigen Sie diese bedingten relativen Häufigkeiten in der Tabelle. Bei welchen Parteien besteht ein beträchtlicher Unterschied zwischen den Studenten, die bei den Eltern wohnen, und den Studenten, die nicht bei den Eltern wohnen?

2. Erstellen Sie ein vergleichendes Paretodiagramm der beiden Gruppen, wobei Sie die Gruppe der Studenten, die bei den Eltern wohnen, als Basis wählen sollten.

3. Testen Sie, ob sich das Wahlverhalten der Studenten, die bei den Eltern wohnen, von dem Wahlverhalten der Studenten, die nicht bei den Eltern wohnen, unterscheidet ($\alpha = 0.05$).

4. Testen Sie, ob der Anteil der Studenten, die zu Hause wohnen, gleich dem Anteil der Studenten ist, die nicht zu Hause wohnen.

Aufgabe 167
Von den Passagieren auf der Titanic waren 337 in der ersten Klasse, 285 in der zweiten Klasse und 721 in der dritten Klasse. Von den Passagieren der ersten Klasse wurden nach dem Unglück 135 vermißt, von denen der zweiten Klasse 160 und von denen der dritten Klasse 541.

1. Erstellen Sie die Kontingenztabelle.

2. Erstellen Sie ein vergleichendes Paretodiagramm der beiden Gruppen, wobei Sie die Gruppe der Personen, die nicht verunglückt sind, als Basis wählen sollten.

3. Testen Sie zum Niveau $\alpha = 0.05$, ob die beiden Merkmale unabhängig sind.

Aufgabe 168
Bei einer Befragung in der Statistik I Vorlesung wurde unter anderem nach dem Geschlecht und der Anzahl der Geschwister gefragt. In Tabelle 17.2 sind die Ergebnisse für die Studenten zusammengestellt, die höchstens zwei Geschwister haben.

Tabelle 17.2: Ergebnis einer Umfrage

Geschlecht Anzahl Geschwister	weiblich	männlich
0	7	44
1	39	77
2	17	26

1. Testen Sie, ob die Merkmale Geschlecht und Anzahl Geschwister unabhängig voneinander sind. ($\alpha = 0.05$)

2. Testen Sie, ob die Anzahl der Geschwister bei weiblichen und männlichen Studenten zusammen binomialverteilt ist mit den Parametern n=2 und p=0.5.($\alpha = 0.05$)

Aufgabe 169
In der Statistik I Vorlesung wurden 250 Fragebögen abgegeben. Von den Teilnehmern sind 69 weiblich. Außerdem haben 111 Teilnehmer Mathematik als Leistungskurs belegt. Von den Frauen haben 23 den Mathematik-Leistungskurs besucht.

1. Testen Sie, ob 40 Prozent der Studenten der Wirtschaftswissenschaften den Mathematik-Leistungskurs besucht haben ($\alpha = 0.05$).

2. Ist der Besuch des Mathematik-Leistungskurses unabhängig vom Geschlecht ($\alpha = 0.05$)?

Aufgabe 170
An der Steuerberater-Prüfung 1999/2000 nahmen 5118 Personen teil. Von den 2636 Personen, die ein Universitätsstudium absolviert haben, bestanden 1502 die Prüfung, von den 1240, Personen, die ein Fachhochschulstudium absolviert haben, bestanden 655 die Prüfung und von den 1242, Personen, die eine Lehrzeit mit Gesellenprüfung absolviert haben, bestanden 593 die Prüfung

1. Erstellen Sie die Kontingenztabelle.

2. Erstellen Sie ein vergleichendes Paretodiagramm der drei Gruppen, wobei Sie die Gruppe der Personen, die ein Universitätsstudium absolviert haben, als Basis wählen sollten.

3. Testen Sie zum Niveau $\alpha = 0.05$, ob die beiden Merkmale unabhängig sind.

Aufgabe 171
Bei einer Befragung wurden 25 Personen nach ihrem Geschlecht befragt. Außerdem mussten die Personen den folgenden Satz ergänzen

Zu Risiken und Nebenwirkungen ...

Von den 13 Frauen haben 7 und von den Männern 3 den Satz richtig ergänzt.

1. Erstellen Sie die Kontingenztabelle.

2. Erstellen Sie ein vergleichendes Paretodiagramm der beiden Gruppen, wobei Sie die Gruppe der weiblichen Personen als Basis wählen sollten.

3. Testen Sie zum Niveau $\alpha = 0.05$, ob die beiden Merkmale unabhängig sind.

Kapitel 18

Das lineare Modell

Aufgabe 172
Über 5 Perioden wurden folgende Preise x_t (in DM) und Absatzmengen y_t (in Stück) für ein Produkt ermittelt. Diese sind in Tabelle 18.1 zu finden.

Tabelle 18.1: Preis und Absatzmengen eines Produktes

t	1	2	3	4	5
x_t	1	2	3	4	5
y_t	8	9	5	1	2

1. Erstellen Sie das Streudiagramm der Daten.
2. Schätzen Sie die Regressionsgerade nach der Methode der Kleinsten Quadrate.
3. Interpretieren Sie $\hat{a}$ und $\hat{b}$.
4. Zeichnen Sie die Regressionsgerade in das Streudiagramm.
5. Berechnen Sie das Bestimmtheitsmaß R^2.
6. Stellen Sie das Konfidenzintervall für β_0 auf.
7. Testen Sie, ob der Preis einen Einfluss auf die Absatzmenge hat. ($\alpha = 0.05$)

Aufgabe 173
Tabelle 18.2 ist der jährliche Ertrag y_t eines Agrarproduktes für die Jahre 1985 bis 1989 zu finden.

Tabelle 18.2: jährlicher Ertrag eines Agrarproduktes für die Jahre 1985 bis 1989

t	1985	1986	1987	1988	1989
y_t	40	80	170	360	750

Wir suchen ein geeignetes Modell für den Zusammenhang zwischen t und y_t.
Erstellen Sie zunächst das Streudiagramm.
Deutet dieses Streudiagramm auf lineares Wachstum hin?
Ein Student unterstellt lineares Wachstum:

$$y_t = a + b\,t + u_t$$

1. Zeigen Sie, dass die Kleinst-Quadrate-Schätzer $\hat{a}$ und $\hat{b}$ von a und b gegeben sind durch $\hat{a} = -337510$ und $\hat{b} = 170$.
2. Bestimmen Sie die Schätzer $\hat{a}$ und $\hat{b}$, falls wir die Zeit messen durch $t = 1, \ldots, 5$.
3. Welcher Zusammenhang besteht zwischen $\hat{y}_t$ und $\hat{y}_{t-1}$ bei linearem Wachstum?

Ein anderer Student meint, dass der Zusammenhang zwischen y_t und t durch exponentielles Wachstum viel besser beschrieben werden könnte, da der Ertrag sich jedes Jahr annähernd verdoppelt.
Er unterstellt also

$$y_t = e^{a+b\,t}\,u_t.$$

1. Zeigen Sie, dass die Kleinst-Quadrate-Schätzer $\hat{a}$ und $\hat{b}$ von a und b gegeben sind durch $\hat{a} = -1458.574$ und $\hat{b} = 0.7366$.
2. Ergibt sich auch aus der Schätzung, dass der Ertrag sich jedes Jahr annähernd verdoppelt?
3. Welcher Zusammenhang besteht zwischen $\hat{y}_t$ und $\hat{y}_{t-1}$ bei exponentiellem Wachstum?

Aufgabe 174
In Tabelle 18.3 ist das wöchentliche Einkommen x_t und die wöchentlichen Ausgaben y_t (in 100 DM) von 8 Personen zu finden.

Tabelle 18.3: wöchentliches Einkommen und die wöchentlichen Ausgaben (in 100 DM) von 8 Personen

t	1	2	3	4	5	6	7	8
x_t	1.7	2.7	3.6	4.6	5.7	6.7	8.1	12.0
y_t	0.8	1.2	1.5	1.8	2.2	2.3	2.6	3.1

Wir suchen ein geeignetes Modell für den Zusammenhang zwischen x_t und y_t.
Erstellen Sie zunächst das Streudiagramm.
Deutet dieses Streudiagramm auf einen linearen Zusammenhang zwischen x_t und y_t hin?
Ein Student unterstellt einen linearen Zusammenhang:

$$y_t = a + b\,x_t + u_t$$

1. Zeigen Sie, dass die Kleinst-Quadrate-Schätzer $\hat{a}$ und $\hat{b}$ von a und b gegeben sind durch

$$\begin{aligned} \hat{a} &= 0.6855 \\ \hat{b} &= 0.2221 \end{aligned}$$

2. Halten Sie das Modell für geeignet?

3. Welcher Zusammenhang besteht zwischen $\hat{y}_t$ und $\hat{y}_{t-1}$ bei einem linearem Zusammenhang?

Ein anderer Student schlägt einen semilogarithmischen Zusammenhang der folgenden Form vor:

$$y_t = a + b\,\ln x_t + u_t$$

1. Welche Idee steckt hinter dem Vorschlag des Studenten?

2. Zeigen Sie, dass die Kleinst-Quadrate-Schätzer $\hat{a}$ und $\hat{b}$ von a und b gegeben sind durch

$$\begin{aligned} \hat{a} &= 0.0477 \\ \hat{b} &= 1.2066 \end{aligned}$$

3. Interpretieren Sie den Wert des Parameters $\hat{a}$.