Aufgabe 1.1

Eine Umfrage ergab folgende Einkommensangaben (in TSD Euro):

i	1	2	3	4	5	6	7	8	9	10
x_i	15	36	24	12	75	63	19	21	43	16

Berechnen Sie:

1. $\sum_{i=1}^{10} x_i$
2. $\sum_{i=6}^{10}(5 + x_i)$
3. $\sum_{i=1}^{10} 2x_i$
4. $\sum_{i=2}^{5}(4 - 3x_i)^2$

Aufgabe 1.2

I.) Welches Skalenniveau weisen die folgenden Merkmale auf:
a) Tachostand, b) Schulnote, c) Uhrzeit, d) Gewicht, e) Tabellenplatz, f) Alter, g) Familienstand, h) Getreidesorte, i) Handygebühren.

II.) Welche Merkmale sind diskret, welche stetig:
a) Wasserstandsanzeige, b) Anzahl Vereinsmitglieder, c) Kraftstoffverbrauch, d) Ausschussstücke pro Tag, e) Weinkonsum (in l), f) Weinkonsum (in Fl.), g) Einwohner, h) Tage bis zur Prüfung.

III.) Ordnen Sie die folgenden Merkmale den Skalierungen nominal-, ordinalskaliert, metrisch und Merkmalstypen (stetig, diskret) zu:
a) Einkommen, b) Haarfarbe, c) Alter, d) Körpergröße, e) Geschlecht, f) Beruf, g) Schultypen, h) Anzahl an Kindern in Schulklassen, i) Raucher/Nichtraucher, j) Altersklassen.

Aufgabe 1.3

Die folgende Tabelle enthält die Sitzverteilung des 15. Deutschen Bundestages nach der Bundestagswahl 2002.

j	Partei	Anzahl der Sitze h_j	Anteile f_j
1	SPD	251	0.416
2	CDU	190	0.315
3	CSU	58	0.096
4	GRÜNE	55	0.091
5	FDP	47	0.078
6	PDS	2	0.003

Stellen Sie die absoluten und relativen Häufigkeiten des Merkmals „Parteizugehörigkeit" tabellarisch und graphisch dar.

Aufgabe 1.4

Eine Großbäckerei besitzt 20 Filialen unterschiedlicher Größe: 7 Filialen beschäftigen je 7 Mitarbeiter, 6 Filialen je 5 Mitarbeiter, 4 Filialen je 8 Mitarbeiter, 2 Filialen je 4 Mitarbeiter und 1 Filiale 11 Mitarbeiter.

a) Stellen Sie die Verteilung der relativen Häufigkeiten des Merkmals „Mitarbeiterzahl“ grafisch dar.

a) Ermitteln Sie die kumulierte Häufigkeitsverteilung und stellen Sie diese tabellarisch und grafisch dar.

c) Wie viel Prozent der Filialen beschäftigen zwischen 6 und 10 Mitarbeiter? Bestimmen Sie das Ergebnis mit Hilfe der empirischen Verteilungsfunktion.

Aufgabe 1.5

Der Besitzer eines kleinen Kinos macht sich Gedanken über die Wirtschaftlichkeit seines Hauses. An 100 Tagen zählt er die Zuschauer. Folgende Zahlen liegen ihm nun vor:

x_j	41	42	43	44	45	46	47	48	49	50	51
h_j	1	9	13	13	20	15	10	7	5	4	3

a) Bestimmen Sie die relativen Häufigkeiten und die kumulierte Häufigkeitsfunktion.

b) Zur Existenzerhaltung reicht es aus, wenn an 90% der Tage mindestens 48 Besucher kommen. Hat das Kino eine Überlebenschance?

c) Wie groß ist der Anteil der Tage, an denen der Kinobesitzer weniger als 270 Euro einnimmt (Eintrittspreis: 6 Euro)?

d) An wie viel Prozent der Tage kommen maximal 50, mindestens aber 45 Besucher in das Kino?

Aufgabe 1.6

Die 50 Beschäftigten eines Betriebes haben während eines bestimmten Zeitraums folgende Überstunden geleistet:

40, 32, 4, 18, 30, 46, 10, 10, 28, 20, 36, 44, 24, 42, 6, 2, 34, 34, 8, 26, 8, 14, 22, 40, 38, 34, 32, 24, 36, 18, 16, 26, 16, 34, 46, 20, 26, 36, 44, 26, 42, 4, 26, 44, 12, 42, 36, 14, 16, 32

a) Stellen Sie aus diesen Angaben eine Häufigkeitsverteilung in den Größenklassen $(0, 15]$, $(15, 25]$, $(25, 40]$ und $(40, 50]$ auf.

b) Stellen Sie die relativen Häufigkeiten sinnvoll graphisch dar.

c) Bestimmen Sie den Anteil der Beschäftigten, die

   c.1) höchstens 25
   c.2) über 25
   c.3) über 15 bis 40

   Überstunden gemacht haben.

Aufgabe 1.7

Bei 15 Haushalten wurde die Zahl der Haushaltsmitglieder ermittelt. Es ergaben sich folgende Werte:

3, 4, 1, 7, 2, 3, 4, 4, 5, 3, 2, 1, 2, 2, 10

Bestimmen Sie die absoluten und relative Häufigkeiten sowie die kumulierte Häufigkeitsfunktion.

a) Wie viele Haushalte haben mehr als 5 Mitglieder?
b) Wie groß ist der Anteil der Haushalte mit höchstens 4 Mitgliedern?
c) In wie viel Prozent der Haushalte leben zwischen 3 und 6 Menschen (jeweils einschließl.)?

Aufgabe 1.8

Aus der folgenden Tabelle ist die Religionszugehörigkeit (C: Christl. Religion, S: Sonstige) sowie die Zugehörigkeit zu den Berufsgruppen A: Arbeiter/Angestellte, B: Beamte, F: Freiberufler ersichtlich:

Person	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Beruf	A	F	B	B	A	A	B	F	F	F	A	B	A	A	A
Religion	C	S	S	C	C	C	C	S	C	S	C	C	S	C	C

Person	16	17	18	19	20
Beruf	A	F	B	B	A
Religion	C	S	C	C	S

Erstellen Sie eine Kontingenztabelle (Kreuztabelle) der Anteile.

Aufgabe 1.9

In einer Stadt wurde eine Mäuseart auf Krankheit untersucht. Die Mäuse wurden in drei Klassen eingeteilt, je nachdem, wo sie gefangen wurden: Zentrum, Peripherie und Land. Von den insgesamt 60 untersuchten Mäusen waren 15 krank. Im Zentrum hat man 8 Mäuse gefangen und bei 5 davon die Krankheit nachgewiesen; in der Peripherie waren hingegen von den 20 untersuchten Mäusen nur 2 krank. Stellen Sie die absoluten und relativen Anteile in einer Kreuztabelle dar.

Aufgabe 1.10

Erstellen Sie aus den folgenden Angaben einer Personalabteilung eine Kreuztabelle der absoluten und relativen Häufigkeiten.

Mitarbeiter	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Steuerklasse	1	2	5	5	4	3	2	4	1	2	3	2	1	3	4
Anz. Kinder	1	1	5	1	4	3	2	2	3	4	5	4	3	2	4

Mitarbeiter	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Steuerklasse	2	1	2	2	2	2	5	5	1	2	3	3	5	3	3
Anz. Kinder	4	1	2	3	4	5	3	2	3	4	3	4	1	2	3

Aufgabe 1.11

20 Haushalte (H) wurden nach ihrem monatlichen Einkommen (x) und den monatlichen Konsumausgaben (y) befragt. Das Ergebnis lautete:

H	1	2	3	4	5	6	7	8	9	10
x	800	1200	1100	1480	1300	900	1000	1200	800	925
y	700	820	930	1270	1160	840	620	970	680	750

H	11	12	13	14	15	16	17	18	19	20
x	1150	870	1420	950	1350	1280	1040	1470	1220	1120
y	870	870	1050	920	1250	1010	820	1310	1200	980

Stellen Sie die Beobachtungswerte in einer Kreuztabelle dar. Verwenden Sie folgende Klassen: über 500 bis 700, über 700 bis 900, über 900 bis 1100, über 1100 bis 1300, über 1300 bis 1500.

Aufgabe 2.1

Gegeben sei folgende Häufigkeitsverteilung eines metrischen Merkmals X:

Merkmalsausprägung	-2	-1	0	1	2	3	4	5
Abs. Häufigkeit	1	1	2	3	5	8	7	3

a) Stellen Sie die relative Häufigkeitsfunktion grafisch dar.

b) Berechnen Sie Modus, Median und arithmetisches Mittel. Fügen Sie die drei Größen in die Zeichnung ein.

Aufgabe 2.2

Bei der Befragung der Studierenden aller Statistik-Veranstaltungen ergaben sich für folgende Altersangaben die in der Tabelle angegebenen Häufigkeiten. Bestimmen Sie für diese Stichprobe die folgenden Maßzahlen: a) Arithmetisches Mittel; b) Modus; c) Median

Alter	h_j	H_j
19	3	3
20	7	10
21	18	28
22	20	48
23	29	77
24	23	100
25	19	119
26	11	130
27	7	137
28	8	145
29	2	147

Aufgabe 2.3

Bei einer Befragung der Studierenden zu ihrer Meinung bezüglich des Komforts der Hörsäle mittels einer Ratingskala (1=sehr bequem bis
7 = miserabel; Annahme: Merkmal ist intervallskaliert) ergab sich folgendes Bild:

Komfort	BWL	VWL	Mathe
1	1	7	2
2	2	6	4
3	3	5	6
4	4	4	8
5	5	3	6
6	6	2	4
7	7	1	2

Ermitteln Sie für jede Vorlesung Modus und arithmetisches Mittel.

Aufgabe 2.4

Eine psychologische Untersuchung beschäftigt sich mit der Nervosität und der Anpassung von Studenten vor einer anstehenden Klausur. Dazu wurde beobachtet, wie lange vor dem offiziellen Beginn der Klausur die Studenten vor den Hörsälen eintreffen. Die Beobachtungen ergaben folgendes Ergebnis:

Minuten vor der Klausur	Anzahl der Studenten	H_j
über 0 bis 1 Minute	2	2
über 1 bis 5 Minuten	10	12
über 5 bis 10 Minuten	55	67
über 10 bis 20 Minuten	38	105
über 20 bis 30 Minuten	15	120

a) Berechnen Sie den Median und das arithmetische Mittel.

b) Zeichnen Sie ein Histogramm.

Aufgaben 2.5, 2.6

2.5) In einem Unternehmen wurden bei 10 Mitarbeitern folgende Fehlzeiten, gemessen in Tagen pro Jahr, festgestellt: 9, 3, 13, 0, 62, 12, 4, 12, 7, 2. Wie groß ist der Mittelwert?

2.6) Der Bekannte Psychologe J. Ensen führt Intelligenztests durch und erhebt bei 2 Gruppen (a und b) gleichaltriger Studenten das metrische Merkmal „Intelligenzquotient“ (IQ). Er findet:

IQ (a)	90	90	97	99	98	145	114	80	85	102
IQ (b)	95	99	90	105	85	98	110	96	69	103

Berechnen Sie die durchschnittlichen IQs der beiden Gruppen, und geben Sie für jedes Gruppenmitglied die Abweichung zum jeweiligen Mittelwert an.

Aufgabe 2.7

In zwei Klausuren erzielen 9 TeilnehmerInnen die folgenden Ergebnisse:

	Punkte	
i	Statistik	VWL
1	100	90
2	88	80
3	71	70
4	61	60
5	49	50
6	48	8
7	48	8
8	20	7
9	11	6

Notenschema

Punkte	Note
unter 50	5.0
ab 50	4.0
ab 63	3.0
ab 76	2.0
ab 89	1.0

a) Ermitteln Sie die Statistik und VWL-Noten der Studierenden.

b) Ermitteln Sie für beide Klausuren die absoluten und relativen Häufigkeiten der Punkte (Klassen in 20er-Schritten [0-20, 21-40, ..., 81-100]) und der Noten („Klassenspiegel“).

c) Zeichnen Sie diese Häufigkeitsverteilungen als Histogramme jeweils untereinander (Punkteverteilungen untereinander, Notenverteilungen ebenso) und vergleichen Sie.

d) Ermitteln Sie drei sinnvolle Lagemasse.

e) Was fällt besser aus - Statistik oder VWL?

Aufgabe 2.8

Die Altersverteilung der Abgeordneten des Deutschen Bundestages sah zum 31.12.1999 wie folgt aus:

Alter	Anzahl der Abgeordneten	f_k
19–33	28	0.0419
34–43	90	0.1345
44–48	103	0.15396
49–53	150	0.224
54–58	172	0.2571
59–63	103	0.15397
64–75	23	0.0344

a) Bestimmen Sie die relative Häufigkeitsverteilung des Merkmals “Alter”, und stellen Sie diese grafisch dar!

b) Bestimmen Sie den Median der Altersverteilung!

c) Welcher Anteil der Abgeordneten war zum Stichtag 50 Jahre oder älter?

Aufgabe 2.9

Für fünf Fahrten mit einem PKW werden Daten über die Durchschnittsgeschwindigkeit und den durchschnittlichen Benzinverbrauch ermittelt.

Fahrt	Strecke in Km	Durchschnittliche(r) Geschw. in Km/h	Durchschnittliche(r) Verbrauch in L/100Km
1	412	98	7.2
2	278	104	7.5
3	380	82	6.8
4	520	95	7.0
5	456	?	?
insg.	2046	96	7.1

a) Berechnen Sie die Durchschnittsgeschwindigkeit und den durchschnittlichen Benzinverbrauch für die Gesamtstrecke der ersten vier Fahrten!

b) Für alle fünf Fahrten betrug die Durchschnittsgeschwindigkeit 96 km/h und der durchschnittliche Benzinverbrauch 7,1 l/100km. Bestimmen Sie die Durchschnittsgeschwindigkeit und den durchschnittlichen Benzinverbrauch für die fünfte Fahrt!

Aufgabe 2.10

Ein Mitarbeiter am Institut für Bevölkerungswissenschaft untersucht insgesamt 151 Länder aus verschiedenen Erdteilen hinsichtlich des Alters ihres politischen Systems. Er erhält folgende Tabelle:

	Afrika	Amerika	Asien	Europa	Ozeanien
gering	29	8	18	2	43
mittel	3	4	2	7	2
hoch	1	17	5	10	0

a) Berechnen Sie die relative Häufigkeitsverteilung der Länder mit geringem und die relative Häufigkeitsverteilung der Länder mit hohem Alter ihres politischen Systems!

b) Charakterisieren Sie die Verteilungen von Teilaufgabe a) durch geeignete Lagemaße!

Aufgabe 2.11

In einem Betrieb werden während der Semesterferien sechs Studierende eingestellt.

Studierende	Eintritt (Kalenderwoche)	Austritt (Kalenderwoche)
A	31	35
B	33	42
C	38	43
D	32	35
E	34	43
F	39	42

Einstellungen erfolgen grundsätzlich an Montagen; entsprechend verlassen die Studierenden den Betrieb grundsätzlich an Freitagen.

a) Berechnen Sie die Lohnkosten für alle diese Studierenden, wenn pro Person und Woche 300.– Euro gezahlt werden!

b) Wie viele Studierende waren durchschnittlich (pro Woche) während der betrachteten Zeit im Betrieb?

c) Wie lange war ein Studierender durchschnittlich in diesem Betrieb beschäftigt?

Aufgabe 2.12

Gegeben sei die folgende Tabelle zur Betriebszugehörigkeit von 200 Beschäftigten eines Krankenhauses (ergänzt):

Betriebszugehörigkeit	h_j	f_j	F_j
bis ein Jahr	20	0.1	0.1
über 1 bis 3 Jahre	50	0.25	0.35
über 3 bis 5 Jahre	70	0.35	0.7
über 5 bis 10 Jahre	40	0.2	0.9
über 10	20	0.1	1

a) Wie viel Prozent der Beschäftigten haben eine Betriebszugehörigkeit von höchstens zwei Jahren? Welche Annahme muss für die Beantwortung dieser Frage getroffen werden?

b) Errechnen bzw. bestimmen Sie für die gegebenen Daten die folgenden Kennwerte: Modus, Median und arithmetisches Mittel. Interpretieren Sie die berechneten Lagemasse und gehen Sie dabei auf deren unterschiedliche Bedeutung ein. Welche zusätzlichen Annahmen mussten (ggf.) für die Bestimmung der einzelnen Kennwerte getroffen werden?

Aufgabe 2.13

In einem Landkreis wurden die jeweiligen Anzahlen an Mietwohnungen in bestimmten Preissegmenten (monatliche Nettomieten in Euro) erhoben. Es ergab sich folgende Verteilung:

Nettomiete in Euro	Anzahl an Wohnungen	Nettomiete in Euro	Anzahl an Wohnungen
über 0 bis 150	200	über 750 bis 900	100
über 150 bis 300	300	über 900 bis 1050	80
über 300 bis 450	250	über 1050 bis 1200	70
über 450 bis 600	100	über 1200 bis 1350	50
über 600 bis 750	150		

a) Stellen Sie die Verteilung der Nettomieten graphisch dar.

b) Berechnen Sie — wenn zulässig — das arithmetische Mittel und geben Sie Median und Modus an.

c) Beschreiben Sie die Verteilung der Nettomieten auf Basis der Ergebnisse aus a) und b) nicht nur anhand der statistischen Kenngrößen und Begriffe sondern auch inhaltlich.

Aufgabe 2.14

Von Schulkindern wurden die Merkmale tägliche Fernsehdauer und gelesene Bücher pro Jahr erfasst. In der Kreuztabelle der beiden Merkmale sind hier lediglich die Randverteilungen dargestellt:

		Gelesene Bücher pro Jahr		
		0 bis 2	3 u. mehr	$\sum$
TV pro	0 bis 1 Std			40
Tag	mehr als 1 Std			60
	$\sum$	75	25	

a) Füllen Sie die offenen Zellen so aus, dass folgende Hypothese bestätigt wird: „Je länger die Kinder am Tag fernsehen, desto weniger Bücher lesen sie pro Jahr“

b) Füllen Sie die offenen Zellen so aus, dass garantiert kein Zusammenhang zwischen den beiden Merkmalen besteht.

Aufgabe 2.15

Die beiden Merkmale der Aufgabe 2.14 seien in folgenden Merkmalsklassen vorhanden:

	Code	TV pro Tag in Std	Bücher/Jahr
0	1	15	50
(bis) 1	2	25	20
(bis) 2	3	20	3
(bis) 4	4	17	1
(bis) 6	5	13	10
mehr als 6	6	10	16

a) Berechnen Sie den Interquartilsabstand für beide Merkmale.

b) Vergleichen Sie beide Merkmale bezüglich ihrer Streuung mit den berechneten Quartilsabständen und im Hinblick auf die Verteilungen insgesamt.

Aufgaben 3.1 – 3.4

3.1) Bestimmen Sie für die Stichprobe aus Aufgabe 2.2:
a) Standardabweichung $\tilde{s}$; b) Variationskoeffizient; c) 0.9-Quantil und 0.6-Quantil.

3.2) Bestimmen Sie für die Befragung aus Aufgabe 2.3 die Quartile.

3.3) Bestimmen Sie für die psychologische Untersuchung aus Aufgabe 2.4: a) das 0.1-Quantil; b) den Interquartilsabstand.

3.4) Die Marketingabteilung der PASTA AG möchte wissen, ob die Anzahl des verkauften Produktes (Basilikumpasta) in einem sehr engen oder breiten Bereich streuen. Sie befragen 9 von 88 Bioläden in Berlin, die folgende Verkaufszahlen pro Monat angeben: 24; 18; 12; 8; 32; 26; 10; 22; 28. Bestimmen Sie Spannweite und Standardabweichung der Stichprobe.

Aufgaben 3.5 – 3.6

3.5) In welcher der beiden Gruppen aus Aufgabe 2.6) streut der IQ stärker?

3.6) Bei der Produktion eines feinmechanischen Werkstücks müssen nacheinander 2 Arbeitsgänge ausgeführt werden. Der 1. Arbeitsgang stellt höhere Anforderungen an die Geschicklichkeit, der 2. Arbeitsgang dagegen höhere Anforderungen an die Aufmerksamkeit und Sorgfalt. Bei 6 Mechanikern werden folgende Zeiten gemessen (x: verbrauchte Arbeitszeit für den 1. Arbeitsgang; y: verbrauchte Arbeitszeit für den 2. Arbeitsgang (jeweils in Minuten)).

Mechniker: i	1	2	3	4	5	6
x_i	11	12	15	9	11	14
y_i	16	22	16	23	23	20

Welche der Arbeitszeiten variiert am stärksten?

Aufgabe 3.7

An einer bundesdeutschen Universität sind in den einzelnen Fakultäten im ersten Fachsemester folgende Anzahlen von Studierenden eingeschrieben:

Fakultät	Studierende im 1. Fachsemester	f_i
Ingenieurwesen	326	0.2399
Recht	224	0.1648
Theologie	60	0.0442
Wirtschaftswissenschaften	532	0.3915
Philosophie	89	0.0655
Pädagogik	128	0.0942

Charakterisieren Sie diese Häufigkeitsverteilung durch geeignete Lage- und Streuungsmaße!

Aufgaben 3.8 – 3.9

3.8) Gegeben seien vier Beobachtungspaare der gemeinsam auftretenden Merkmale x und y:

x_i	1	2	4	5
y_i	4	3	5	8

Berechnen Sie die Kovarianz.

3.9) Gegeben sind die folgenden Beobachtungspaare zweier metrisch messbarer Merkmale x und y: $(4;7)$, $(1;5)$, $(2;5)$, $(2;6)$, $(3;6)$, $(5;4)$, $(5;8)$, $(6;8)$, $(6;9)$, $(8;8)$. Zeichnen Sie ein Streudiagramm und bestimmen Sie die Kovarianz.

Aufgabe 3.10

Betrachten Sie die Urliste:
-7 -6 -2 2 3 4 6

Welche der folgenden Größen kann man nicht bestimmen:

a) das arithmetische Mittel
b) den Modalwert
c) den Median
d) das geometrische Mittel
e) den Variationskoeffizienten

Aufgabe 3.11

Der Produzent zweier Golfschläger x und y macht mit beiden Schlägern ein Experiment: 10 Spieler versuchen jeweils mit beiden Schlägern ein Ziel in 150 Metern Entfernung zu treffen. Das Ergebnis (alle Angaben in Meter) ist der folgenden Tabelle zu entnehmen (ein Spieler musste vorzeitig gehen, aus Gründen, die nichts mit der Fragestellung zu tun hatten):

	Meter	
Spieler	Schläger x	Schläger y
1	151	150
2	150	149
3	152	155
4	153	145
5	141	153
6	154	153
7	147	154
8	144.7	147
9	151	152
10		149

a) Berechnen Sie für beide Schläger das arithmetische Mittel und die empirische Standardabweichung.

b) Berechnen Sie für beide Schläger das 0.25-Quantil, das 0.75-Quantil und den Interquartilsabstand.

c) Welches Maß würden Sie anwenden, wenn Sie die Wahl zwischen Schläger x und y hätten und darauf achten müssten, welcher Schläger am genauesten schlägt?

Aufgabe 3.12

Bei einer Befragung von $n = 2542$ erwerbstätigen Personen wird nach der Anzahl der im Haushalt lebenden Kindern („Anzahl Kinder“: 0, 1, 2, 3 oder mehr) und der Sorge um die Sicherheit des Arbeitsplatzes gefragt (0: „Mache mir große Sorgen“, 1: „Mache mir einge Sorgen“, 2: „mache mir keine Sorgen“). Die zweidimensionale Häufigkeitsverteilung ist in folgender Tabelle gegeben.

Sorge	Anzahl Kinder				
	0	1	2	≥ 3	Summe
0	113	72	54	25	264
1	400	198	128	79	805
2	843	306	242	82	1473
Summe	1356	576	424	186	2542

a) Sind die beiden Merkmale in der Stichprobe unabhängig (mit Begründung)?

b) Wählen Sie einen geeigneten Kennwert, berechnen Sie die Stärke des Zusammenhangs und interpretieren Sie das Ergebnis.

c) Wie groß ist (in der Stichprobe) die Chance, dass jemand der/die in einem Haushalt ohne Kinder lebt, angibt sich große Sorgen um die Sicherheit des Arbeitsplatzes zu machen, im Verhältnis zu jemandem, der/die angibt, sich keine Sorgen zu machen?

d) Wie ist diese Chance relativ zur entsprechenden Chance bei Personen, die angeben mit drei oder mehr Kindern in einem Haushalt zu leben?

Aufgabe 3.13

Die folgende Tabelle gibt absolute Häufigkeiten von Wahlabsicht (CDU, SPD, FDP, Grüne, PDS) und Konfession (katholisch evangelisch, keine) in einer Stichprobe vom Umfang von $n = 2168$ wider.

Wahlabsicht	Konfession			
	katholisch	evangelisch	keine	
CDU	327	306	141	774
SPD	198	300	216	714
FDP	49	109	41	199
Grüne	92	129	134	355
PDS	10	16	100	126
	676	860	632	2168

Wie stark ist der Zusammenhang in dieser Stichprobe ? Begründen Sie die Wahl des von Ihnen gewählten Maßes.

Aufgabe 4.1

Fortsetzung der Übungsaufgabe, Abschnitt 7, „Regressionsanalyse“:
Es wird vermutet, dass es zwischen den Variablen „Ersparnisse“ und „Jahreseinkommen“ einen positiven linearen Zusammenhang gibt. Für 10 befragte Haushalte ergaben sich die folgenden Werte:

i	Einkommen	Ersparnisse
1	34.2	2.8
2	40.8	4.1
3	76.7	8.8
4	52.6	5.8
5	50.1	4.9
6	42.5	4.5
7	73.5	8.1
8	61.4	7.7
9	56.9	7.0
10	47.3	4.3

Berechnen Sie R^2.

Aufgabe 4.2

Es soll die Beziehung zwischen Werbeausgaben (X) in 10 000 Euro einerseits und Verkaufseinnahmen in 100 000 Euro andererseits untersucht werden. In einer Stichprobe wurden die entsprechenden Beträge für $n = 28$ Produkte erhoben. Ein mit den Vorarbeiten beauftragter Praktikant berechnete folgende Kennwerte:
$r_{xy} = 0.5371$, $s_x^2 = 1.229$, $s_y^2 = 205.04$, $\bar{x} = 9.9836$, und $\bar{y} = 81.824$.
Gehen Sie bei den folgende Berechnungen davon aus, dass alle notwendigen Annahmen erfüllt sind.

a) Geben Sie die Stichproben-Regressionsgerade für die Regression von Y (abhängige Variable) auf X (unabhängige Variable) an.

b) In welchem Ausmaß können die Verkaufseinnahmen durch die Werbeausgaben „erklärt“ werden ?

c) Zur Vorhersage von Verkaufseinnahmen aufgrund von Werbeausgaben bietet sich einerseits die Stichproben-Regressionsgerade, andererseits das arithmetische Mittel der Verkaufseinnahmen an. Welches dieser beiden Vorhersagemodelle würden Sie im vorliegenden Fall wählen ? (Beziehen Sie in Ihre Begründung auch die Werte relevanter quantitativer Größen mit ein).

Aufgabe 4.3

In einem wissenschaftlichen Projekt soll der Einfluss der Berufserfahrung auf das logarithmierte Einkommen untersucht werden. Dazu wurden $n = 120$ zufällig ausgewählte Personen nach ihrem Einkommen und ihrer Berufserfahrung gefragt. Die ersten Berechnungen für die Berufserfahrung in Jahren (X) und das logarithmierte Einkommen (Y) liefern $\bar{x} = 10.2$, $\bar{y} = 8.8$, $\sum_{i=1}^{n} x_i^2 = 24844$, $\sum_{i=1}^{n} y_i^2 = 26459$, $\sum_{i=1}^{n} x_i y_i = 19854$ und $r_{xy} = 0.624$.

a) Geben Sie die Stichproben-Regressionsgerade für die Regression von „logarithmiertes Einkommen“ (abhängige Variable) auf „Berufserfahrung“ (unabhängige Variable) an.

b) Welches logarithmierte Einkommen würden Sie für eine Person mit einer Berufserfahrung von 20 Jahren vorhersagen ?

Aufgabe 4.4

Die Handelskette Poldi plant die Erweiterung der Verkaufsflächen einer Filiale um 1000 m^2. Zur Entscheidungsfindung wurden die Umsatzzahlen y_i (in Millionen Geldeinheiten) und die Verkaufsflächen x_i (in 1000 m^2) von acht Filialen erhoben. Folgende Daten wurden beobachtet:

x_i	0.48	0.31	1.29	0.98	0.94	1.21	1.49	1.12
y_i	2.21	2.05	5.37	3.68	4.03	4.79	5.84	4.91

Erste Analysen ergaben:

$r_{xy} = 0.984$, $\sum_{i=1}^{8} x_i^2 = 8.7732$, $\sum_{i=1}^{8} y_i^2 = 148.8646$, $\bar{x} = 0.9775$ und $\bar{y} = 4.11$.

a) Berechnen Sie die Regressionsgerade.

b) Mit welchem Umsatz könnte die Handelskette bei einer Verkaufsfläche von 1700 m^2 rechnen?

c) Berechnen Sie das Bestimmtheitsmaß für die Regression von Umsatzzahlen auf Verkaufsfläche.

Aufgaben 5.1, 5.2

5.1) Welche Aussagen zu Wahrscheinlichkeiten beliebiger Ereignisse sind richtig?

a) $P(A \cup B) = P(A) + P(B)$

b) $0 < P(A) < 1$

c) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

d) $P(A \cap B) = P(A)\, P(B)$

e) $0 \leq P(A) \leq 1$

5.2) Wie groß ist die Wahrscheinlichkeit, mit viermaligem Werfen eines Würfels

a) P(„viermal 6“)

b) P(„viermal keine 6“)

c) P(„mindestens eine 6“)

d) P(„6,6,6,5“)

e) dreimal 6 und einmal 5.

Aufgaben 5.3, 5.4

5.3) Aus dem Bestand der Familien mit drei Kindern wird eine Familie per Zufall ausgewählt. Wie groß ist – unter der Annahme einer Geburtenwahrscheinlichkeit für Jungen und Mädchen von 0.5 – die Wahrscheinlichkeit dafür, dass zwei Kinder Mädchen sind, wenn

a) als Vorinformation bekannt ist, dass unter den drei Kindern mindestens ein Mädchen ist.

b) als Vorinformation bekannt ist, dass das älteste Kind ein Mädchen ist.

c) keine Vorinformation genutzt werden kann.

5.4) Aus einer Urne, die drei weiße und zwei schwarze Kugeln enthält, wurden auf gut Glück zwei Kugeln entnommen und in eine zweite Urne gelegt, in der sich bereits vier weiße und vier schwarze Kugeln befanden. Wie groß ist die Wahrscheinlichkeit dafür, dass eine hierauf zufällig aus der zweiten Urne entnommene Kugel weiß ist?

Aufgabe 5.5

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

a) Bestimmen Sie $P(C)$, $P(D)$ und $P(F)$

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

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

d) Wie groß ist die Wahrscheinlichkeit, dass bei einem Zufallsexperiment D bzw. F eintritt, wenn man weiß, dass das jeweils andere Ereignis eingetreten ist?

Aufgabe 5.6

5.6) Zur Auswahl von Studienbewerbern will eine Universität zwei Aufnahmetests einsetzen. Die Tests sollen gleichwertig sein, d.h. die jeweils einander entsprechenden Aufgaben sollen den gleichen Schwierigkeitsgrad besitzen. In einer Untersuchung bearbeiteten 600 Versuchspersonen die Testaufgaben. Für zwei einander entsprechende Aufgaben in beiden Tests wurde folgendes Untersuchungsergebnis vorgelegt:

Anzahl der Personen	Aufgabe 1 gelöst	Aufgabe 1 nicht gelöst
Aufgabe 2 gelöst	116	174
Aufgabe 2 nicht gelöst	211	99

Werten Sie das Untersuchungsergebnis grob, indem Sie die relativen Häufigkeiten der Ereignisse A („Aufgabe 1 gelöst") und B („Aufgabe 2 gelöst") miteinander vergleichen. Zeigen Sie, dass A und B voneinander abhängige Ereignisse sind.

Aufgabe 5.7

5.7) Ein Kraftfahrzeughändler weiß aus Erfahrung, dass bei den in Zahlung genommenen Wagen 50% Mängel am Motor, 70% an der Karosserie und 30% an Motor und Karosserie aufweisen. Wie groß ist die Wahrscheinlichkeit dafür, dass ein in Zahlung genommener Wagen

a) weder am Motor noch an der Karosserie Mängel aufweist?

b) auch einen Mangel am Motor besitzt, wenn bekannt ist, dass die Karosserie beschädigt ist?

Aufgabe 5.8

5.8) Ein elektronisches Gerät besteht aus 2 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.

a) Wie groß ist die Wahrscheinlichkeit, dass B ausfällt, wenn A ausgefallen ist?

b) Wie groß ist die Wahrscheinlichkeit, dass B ausfällt?

c) Wie groß ist die Wahrscheinlichkeit, dass mindestens eins von beiden ausfällt?

d) Wie groß ist die Wahrscheinlichkeit, dass A und B nicht gleichzeitig ausfallen?

Aufgabe 5.9

Für vier gleichrangige Stellen bei einer Behörde bewerben sich vier Frauen und sechs Männer, die den Qualitätsanforderungen genügen. Die Besetzung der Stellen wird durch eine Zufallsauswahl nach dem Modell ohne Zurücklegen vorgenommen.

1. Wie groß ist die Wahrscheinlichkeit, dass keine bzw. mindestens eine Frau eingestellt wird?
2. Wie groß ist die Wahrscheinlichkeit, dass gleich viele Männer und Frauen eingestellt werden?
3. Es werden tatsächlich drei Männer und eine Frau eingestellt. Für eine weitere freiwerdende Stelle wird aus den restlichen Bewerbern/Bewerberinnen eine(r) zufällig ausgewählt. Wie groß ist die Wahrscheinlichkeit, dass diese zusätzliche Stelle mit einem männlichen Bewerber besetzt wird?

Aufgabe 5.10

Aus langjähriger Erfahrung weiß man, dass 70% der Versuchspersonen (Vpn) einen bestimmten Eignungstest A bestehen, 60% einen anderen Eignungstest B bestehen und 50% beide Tests bestehen. Wählt man also zufällig eine Vpn aus der entsprechenden Grundgesamtheit aus und definiert die Ereignisse

A: Vp besteht Test A

B: Vp besteht Test B

so gilt: $P(A) = 0.7$, $P(B) = 0.6$ und $P(A \cap B) = 0.5$

a) Sind A und B stochastisch abhängig?

b) Berechnen Sie $P(\overline{A})$, $P(\overline{B})$ und $P(A \cup B)$.

c) Berechnen Sie $P(A|B)$ und $P(B|A)$.

Aufgabe 6.1

In einer Urne befinden sich $n = 4$ Kugeln, welche die Zahlen 2,4,8 und 16 tragen. Es werden nach dem Modell mit Zurücklegen $n = 2$ Kugeln entnommen. Man definiert die Zufallsvariable X als Durchschnitt der Zahlen, die die beiden entnommenen Kugeln tragen.

a) Zählen Sie die 16 möglichen Resultate des Zufallsvorgangs in Form von Zahlenpaaren auf und errechnen Sie jeweils den zugehörigen Wert für X.

b) Ermitteln Sie die Wahrscheinlichkeitsfunktion der Zufallsvariablen X.

c) Zeichnen Sie die Wahrscheinlichkeits- und die Verteilungsfunktion von X.

Aufgabe 6.2

Für eine stetige Zufallsvariable gilt:

$$f(x) = \begin{cases} 4bx & \text{für } 0 \leq x < 1 \\ 0.5 - bx & \text{für } 1 \leq x \leq 5 \\ 0 & \text{sonst.} \end{cases}$$

Bestimmen Sie den Parameter b so, dass $f(x)$ eine Dichtefunktion von X ist. Ermitteln Sie die zugehörige Verteilungsfunktion und skizzieren Sie deren Verlauf. Wie groß ist die Wahrscheinlichkeit einen Wert von höchstens 1, 2, 4 zu beobachten?

Aufgabe 7.1

Aus einer Urne mit 6 roten und 4 schwarzen Kugeln werden zwei Kugeln ohne Zurücklegen gezogen. Es ist

$$X = \begin{cases} 1 & \text{falls rote Kugel im ersten Zug} \\ 0 & \text{sonst} \end{cases}$$
$$Y = \begin{cases} 1 & \text{falls rote Kugel im zweiten Zug} \\ 0 & \text{sonst.} \end{cases}$$

a) Geben Sie die möglichen Wertepaare (x_j, y_j) dieser Zufallsvariablen an und erstellen Sie eine Kreuztabelle der bivariaten Wahrscheinlichkeitsfunktion mit den entsprechenden Randverteilungen.

b) Stellen Sie alle mögliche bedingten Verteilungen in Tabellenform dar.

c) Sind die beiden Variablen X und Y unabhängig (Begründung)?

Aufgaben 7.2 – 7.4

7.2) Die diskrete Zufallsvariable X hat die folgende Verteilung:

x	0	1	2	4
$f(x)$	0.2	0.4	0.3	0.1

Berechnen Sie Erwartungswert und Varianz.

7.3) Sie haben sich nach Ihrem Lottogewinn 100 Aktien der SeeleKomm AG (SK) und 200 Aktien der Nord-Eis (NE) zugelegt. SK notiert zu 480 Euro und NE zu 240 Euro. Ihr Freund, der Börsenmakler Richard Risiko hat für den nächsten Börsentag vorhergesagt, dass SK mit 1/3 Wahrscheinlichkeit gleich bleibt und mit einer Wahrscheinlichkeit von 2/3 um 10% anzieht. Dagegen befürchtet er, dass NE angesichts des kühlen Wetters 10% seines Wertes verliert oder höchstens gleich bleibt. Die Chancen dafür ständen 50:50, meint er. Welchen Wert Ihres Aktienpaketes erwarten Sie auf dieses Basis?

7.4) Eine Versicherungsagentur erzielt bei guter Konjunktur einen Monatsgewinn von 5000 Euro, bei fallender Konjunktur von 2500 Euro und in einer Rezession macht sie 3500 Euro Verlust. Die Wahrscheinlichkeit für eine gute Konjunktur sei $P(S) = 0.5$ und für eine fallende Konjunktur $P(R) = 0.3$. Wie hoch ist der Erwartungswert des monatlichen Gewinns für die Agentur?

Aufgaben 7.5 – 7.8

7.5) Auf einem Jahrmarkt gibt es folgendes Glücksspiel. In einem Topf sind 9 Kugeln, davon 4 weiße (w) und 5 schwarze (s). Es dürfen mit einem Griff 3 Kugeln herausgenommen werden. Für jede gezogene weiße Kugel wird ein Gewinn von 1 Euro ausgezahlt, jede gezogene schwarze Kugel bedeutet 1 Euro Verlust für den Spieler. Mit welchem durchschnittlichen Gewinn kann der Veranstalter für den Spielbetrieb rechnen?

7.6) Z sei $N(0,1)$-verteilt. Bestimmen Sie a, b, c und d aus:
a)$P(Z \leq a) = 0.6$, b) $P(Z > b) = 0.8$, c) $P(|Z| \leq c) = 0.6$, d) $P(|Z| > d) = 0.3$.

7.7) Die Zufallsvariable Z sei $N(0,1)$-verteilt. Bestimmen Sie:
a) $P(0 < Z \leq 2.4)$, b) $P(-1.3 < Z \leq 0)$, c) $P(-0.8 < Z \leq 0.8)$, d) $P(Z \leq 2.1)$, e) $P(Z > -0.1)$ f) $P(0.2 < Z \leq 1.6)$.

7.8) Die Zufallsvariable X sei $N(100,100)$-verteilt. Bestimmen Sie a, b und c aus:
a) $P(X \leq a) = 0.7$, b) $P(X > b) = 0.65$, c) $P(|X - 100| \leq c) = 0.5$.

Aufgaben 7.9 – 7.10

7.9) Eine Maschine füllt Kaffeepakete so, dass die Masse des eingefüllten Kaffees normalverteilt ist mit $\mu = 520g$ und $\sigma = 15g$. Für die Pakete werden 500g als Füllgewicht angegeben.

a) Wie viel % der Pakete sind untergewichtig?

b) Wie viel % der Pakete wiegen mehr als 550g?

c) Wie groß müsste μ (bei gleichem σ) sein, damit nur 2% der Pakete untergewichtig sind?

7.10) Die Lebensdauer X (in km) eines Automotors einer bestimmten Marke sei näherungsweise normalverteilt mit dem Erwartungswert $\mu = 105000$ und der Standardabweichung $\sigma = 10000$.

a) Bei wie viel Prozent der Motoren übersteigt die Lebensdauer 120000 km?

b) Bei wie viel Prozent der Motoren weicht die Lebensdauer um mehr als 12000 km vom Erwartungswert ab?

c) Geben Sie die Grenzen für ein um den Erwartungswert symmetrisches 90%-iges Intervall an.

d) Berechnen Sie die Grenzen des um den Erwartungswert symmetrischen Intervalls, in das 99.73% der Lebensdauern fallen.

Aufgaben 7.11 – 7.12

7.11) Der Apotheker eines Krankenhauses hat eine Abfüllmaschine für Abführ-Tabletten. Die Abfüllung unterliege einer Normalverteilung mit einem Mittel von $\mu = 100$ und einer Standardabweichung von 5.

a) Wie groß ist die Wahrscheinlichkeit dafür, dass

  a.1) in einem Glas zwischen 90 und 110 Tabletten sind?
  a.2) mehr als 105 Tabletten in einem Glas sind?
  a.3) höchstens 90 Tabletten in einem Glas sind?

b) Welche Mindest-Anzahl an Tabletten ist mit 90-prozentiger Sicherheit in einem Glas? In welchem Bereich liegen die zentralen 95% der Anzahlen abgefüllter Tabletten?

7.12) Ein Biowaren-Handel füllt sein Spitzenprodukt - das Bio-Müsli „tropic" - zu je 500g selbst maschinell ab (μ). Die Abfüllung unterliege einer Normalverteilung mit einer Standardabweichung von 5g.

a) Wie groß ist die Wahrscheinlichkeit dafür, dass in einem Paket

  a.1) zwischen 490 und 510 Gramm sind?
  a.2) mehr als 510 Gramm sind?
  a.3) höchstens 495 Gramm sind?

b) Welches Mindest-Füllgewicht bietet ein Paket mit 90-prozentiger Sicherheit?

c) In welchem Gewichts-Bereich liegen die zentralen 95% der abgefüllten Pakete?

Aufgaben 7.13 – 7.14

7.13) Die erwartete Schifffahrtdauer zwischen Frederikshavn und Göteborg ist 3 Stunden (180 Minuten) mit einer Standardabweichung von 15 Minuten. Es wird angenommen, dass die Schifffahrtdauer normalverteilt ist.

a) Berechnen Sie die Wahrscheinlichkeit dafür, dass die Schiffahrtdauer mehr als 215 Minuten ist.

b) Ein Vertreter fährt 10 Mal pro Vierteljahr mit der Fähre. Berechnen Sie die Wahrscheinlichkeit dafür, dass die durchschnittliche Fahrdauer der 10 Fahrten größer als 215 Minuten ist.

7.14) In einer Nahrungsmittelfabrik stehen drei Maschinen, die Haferflockenpakete abfüllen. Der verantwortliche Betriebsingenieur weiß aus langen Messreihen, dass die Füllgewichte der Pakete (approximativ) normalverteilt sind: $X_1 \sim N(201, 0.81^2)$, $X_2 \sim N(203, 3.24^2)$, $X_3 \sim N(202, 1.69^2)$. Bei einem Abfüllgewicht von 199g oder weniger, wird das Paket von der Endkontrolle zurückgewiesen.

a) Welche Maschine liefert den größten Anteil an ordnungsgemäß gefüllten Paketen?

b) Der Betriebsingenieur kann durch Änderung der Standardabweichung die Produktqualität der einzelnen Maschinen variieren. Wie groß muss er das neue σ_i, $(i = 1, 2, 3)$ wählen, damit bei der i-ten Maschine die Wahrscheinlichkeit, dass die Endkontrolle ein Paket zurückweist, genau 0.05 beträgt?

Aufgaben 7.15 – 7.17

7.15) Das Füllgewicht von Kartoffelsäcken sei normalverteilt mit $\sigma = 5$. Der Erwartungswert μ sei durch Änderungen an der Füllmaschine zu beeinflussen, dabei wird σ nicht verändert. Wie groß ist μ mindestens zu wählen, damit höchstens 3% der Säcke ein Gewicht von 50 kg oder weniger haben?

7.16) Bei einer Lieferung von Präzisionsteilen sei deren Durchmesser normalverteilt mit $\mu = 0.614$ mm und $\sigma = 0.007$ mm. Wie viel Prozent Ausschuss sind zu erwarten, wenn der Durchmesser der Teile a) über 0.600 mm, b) höchstens 0.620 mm betragen soll?

7.17) Die Länge von Gehwegplatten sei normalverteilt mit $\mu = 400$ mm und $\sigma = 5$ mm.

a) Wie groß ist der Ausschussanteil, wenn die Länge der Platten über 390 mm betragen soll?

b) Wie groß ist die Wahrscheinlichkeit, dass eine Platte nicht länger als 407.5 cm ist?

Aufgabe 7.18

Bei einer Klausur mit einer maximalen Punktzahl von 100 seien die Ergebnisse (näherungsweise) normalverteilt mit $\mu = 60$ und $\sigma = 10$. Bestimmen Sie den Anteil der Studenten,

a) die durchgefallen sind, wenn zum Bestehen der Klausur über 50 Punkte erforderlich sind,

b) die die Note „gut“ erhalten, wenn diese für Punktzahlen von über 80 bis einschließlich 95 vergeben wird.

c) Auf welchen Wert muss die Mindestpunktzahl festgelegt werden, wenn nicht mehr als 10% der Studenten durchfallen sollen?