

Übungen zum Kapitel 3

Übung 3.1

In einem Fitnessstudio läuft bei 80 Kunden der Vertrag aus. Der Hälfte der Kunden wird ein Anschreiben geschickt, in dem man sich für die Mitgliedschaft bedankt und den Kunden eine Vertrags Verlängerung anbietet. Von den angeschriebenen Kunden verlängern 32 den Vertrag, von den nicht angeschriebenen 18.

Stellen Sie die Daten in einer Kontingenztabelle dar und geben Sie die Randhäufigkeiten an.

Übung 3.2

In einem Supermarkt wird die Aufenthaltsdauer Y [in Minuten] von 200 Kunden gemessen. Ergänzend wird dazu bei jedem Kunden der Ausgabenbetrag X [in €] registriert. Die Ergebnisse sind in der folgenden Tabelle zusammengestellt ():

	$Y=[0,5)$	[5,10)	[10,20)	[20,30)	[30,40)	[40,60)	[60,90)	<i>Summe</i>
$X=[0-10)$	5	6	8	6	-	-	-	25
[10-20)	3	5	12	10	-	-	-	30
[20-30)	-	1	5	10	6	5	5	32
[30-50)	-	-	3	5	15	15	15	53
[50-100)	-	-	-	5	5	20	30	60
<i>Summe</i>	8	12	28	36	26	40	50	200

- Wie hoch war der durchschnittliche Ausgabenbetrag der Kunden, deren Aufenthaltsdauer unter 5 Minuten war?
- Wie hoch war der gesamte Ausgabenbetrag aller Kunden und der durchschnittliche Ausgabenbetrag je Kunde?
- Wie hoch war der Umsatzanteil der Kunden mit einer Aufenthaltsdauer von einer Stunde und mehr am Gesamtumsatz?
- Wie hoch war die durchschnittliche Aufenthaltsdauer der Kunden, die zwischen 20 und 50 € ausgegeben haben?
- Wie hoch war der Anteil der Kunden, die 20 Minuten und länger und für 20 € und mehr eingekauft haben?
- Berechnen Sie die Regressionsgerade, den Korrelationskoeffizienten und das Bestimmtheitsmaß.

Übung 3.3

In der folgenden Tabelle sind für fünf aufeinanderfolgende Jahre die Marketingsausgaben und Umsätze eines Unternehmens abgedruckt:

Marketingsausgaben [in 10.000€]	Umsatz [in Millionen €]
50	6
20	4
10	3,5
30	4,7
80	5,0

- a) Berechnen Sie den Korrelationskoeffizienten.
- b) Interpretieren Sie das Ergebnis.

Übung 3.4

In der folgenden Tabelle sind die DAX-Stände und die geschätzten Nutzerzahlen des Internetdiensts Twitter für das Jahr 2010 abgedruckt. Beide Reihen steigen im Lauf des Jahres.

Monat (Ende)	Stand des DAX	Nutzer Twitter
Januar	5 608,79	194000
Februar	5 598,46	198000
März	6153,55	248000
April	6135,70	270 000
Mai	5 964,33	270 000
Juni	5 965,52	271 000
Juli	6147,97	269 000
August	5 925,22	268000
September	6 229,02	279 000
Oktober	6 601,37	344000
November	6 688,49	367000
Dezember	6914,19	425 000

- a) Berechnen Sie den Korrelationskoeffizienten zwischen den beiden Größen.
- b) Besteht ein Zusammenhang zwischen der Indexentwicklung und der Nutzerzahl von Twitter?
- c) Wie ist der Korrelationskoeffizient hier zu interpretieren?

Übung 3.5

Ein Industrieunternehmen möchte den Zusammenhang zwischen einem bestimmten Zusatz zu einem Werkstoff und der Bruchfestigkeit des Stoffes erforschen. Dazu führt der Werksleiter Experimente mit verschiedenen Anteilen am Zusatz durch. Die Bruchfestigkeit wird auf einer metrischen Skala von 1 bis 10 gemessen, wobei 10 einer maximalen Bruchfestigkeit entspricht.

Anteil am Zusatz	Bruchfestigkeit
1 %	0,5
2%	3,3
3%	2,7
4%	6,5
5%	5,4

- Zeichnen Sie ein Streudiagramm der Daten.
- Erklären Sie die Bruchfestigkeit durch den Anteil am Zusatz.
- Wie gut erklärt der Anteil des Zusatzes die Bruchfestigkeit? Bestimmen Sie dazu das Bestimmtheitsmaß.

Übung 3.6

In der folgenden Tabelle sind die DAX-Stände und die Kurse der BMW-Aktie für 2010 abgedruckt:

Monat (Ende)	Stand des DAX	Stand BMW
Januar	5 608,79	30,94
Februar	5 598,46	29,78
März	6153,55	34,23
April	6135,70	37,09
Mai	5 964,33	37,97
Juni	5 965,52	40,15
Juli	6147,97	41,33
August	5 925,22	41,51
September	6 229,02	51,43
Oktober	6 601,37	51,41
November	6 688,49	58,04
Dezember	6914,19	58,71

Als stetige Renditen bezeichnet man $R_t = \ln(Kurs_t) - \ln(Kurs_{t-1})$. Dabei bezeichnet t in obiger Tabelle den Monat, $t - 1$ bezeichnet entsprechend den Vormonat. Mit $\ln()$ ist der natürliche Logarithmus bezeichnet.

- Berechnen Sie die stetigen Monatsrenditen für beide Kursreihen für Februar bis Dezember.
- Berechnen Sie den Korrelationskoeffizienten nach Pearson zwischen den Renditen.
- Erklären Sie die Rendite der BMW-Aktie durch die Renditen des DAX.

Übung 3.7

In der Personalabteilung eines Unternehmens hat man festgestellt, dass ein Zusammenhang zwischen dem Alter eines Mitarbeiters und der Anzahl der Krankheitstage besteht. Die Korrelation ist 0,3. Das durchschnittliche Alter der Belegschaft beträgt 40 Jahre, die Standardabweichung ist 10. Die durchschnittliche Anzahl der Krankheitstage in einem Jahr beträgt 4 bei einer Standardabweichung von 3.

- Berechnen Sie aus den gegebenen Daten eine Regression, die die Anzahl der Krankheitstage durch das Alter des Mitarbeiters erklärt.
- Mit wie vielen Krankheitstagen ist bei einem 60-jährigen Mitarbeiter zu rechnen?
- Wie hoch ist das Bestimmtheitsmaß?

Übung 3.8

X und Y seien zwei metrisch skalierte Merkmale. Folgende Urliste sei gegeben:

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	2	2	6	6	3	3	5	5	3	3	5	5
y_i	0	4	0	4	1	3	1	3	1	3	1	3

Welche der folgenden Aussagen ist bzw. sind richtig?

- Der Korrelationskoeffizient zwischen X und Y nach Pearson ist 0.
- Die nach der KQ-Methode geschätzte Steigung der Regressionsgeraden von Y auf X ist 0.
- Der nach der KQ-Methode geschätzte Achsenabschnitt der Regressionsgeraden von Y auf X ist 0.
- Das Bestimmtheitsmaß ist 0.
- X und Y sind linear unabhängig.
- Alle Aussagen sind falsch.

Übung 3.9

Aus einer Befragung von Einzelhändlern liegen für die beiden Merkmale X : Anzahl der Mitarbeiter und Y : Höhe des Umsatzes folgende statistischen Kennzahlen vor:

- die Varianz der y -Werte ist 1,25,
- das arithmetische Mittel der x -Werte ist 1,5,
- das arithmetische Mittel der y -Werte ist 2,5,
- die Kovarianz zwischen Y und X ist -1,25,
- der Korrelationskoeffizient nach Bravais-Pearson ist -1.

Welche der folgenden Aussagen ist bzw. sind richtig?

- Das Bestimmtheitsmaß ist 1.
- Die Steigung der Regressionsgerade ist -1.
- Der Achsenabschnitt der Regressionsgerade ist 4.
- Die Varianz der x -Werte ist so groß wie die Varianz von y -Werten.

e) Der Variationskoeffizient von X ist kleiner als der Variationskoeffizient von Y .

f) Alle Aussagen sind falsch.

de, so ergibt sich: