Lösungshinweise zu den Aufgaben der deskriptiven Statistik

Lösungen zu Kapitel 2

Lösung 2.1

(1) diskret
(2) stetig
(3) diskret
(4) stetig

Lösung 2.2

(1) Bewegungsmasse
(2) Bestandsmasse
(3) Bestandsmasse
(4) Bewegungsmasse

Lösungen zu Kapitel 3

Lösung 3.1

(1) Verhältnisskala
(2) Verhältnisskala
(3) Nominalskala
(4) Intervallskala
(5) Ordinalskala
(6) Nominalskala
(7) Ordinalskala
(8) Intervallskala
(9) Verhältnisskala

Lösungen zu Kapitel 5

Lösung 5.1

a) $\sum_{i=1}^{5} x_i = x_1 + x_2 + x_3 + x_4 + x_5 = 100 + 120 + 140 + 150 + 180 = 690[kg]$

b) $\sum_{i=1}^{5} x_i + 5a = 690 + 5 \cdot 20 = 790[kg]$

c) $\sum_{i=1}^{5} (x_i + 0{,}2x_i) = \sum_{i=1}^{5} 1{,}2x_i = 1{,}2\sum_{i=1}^{5} x_i$

d) p_i sei das Gewicht der Verpackung in Prozent des Gewichts der Ware Nr. i. Dann gilt für das Gesamtgewicht aller Waren:

$$\sum_{i=1}^{5} (x_i + p_i x_i) = \sum_{i=1}^{5} (1 + p_i)\, x_i$$

Lösung 5.2

a) $\sum_{i=1}^{n} x_i = x_1 + x_2 + ... + x_n$

Falls $x_i = a$ für i = 1, 2, ... , n, erhält man:

$\sum_{i=1}^{n} x_i = a + a + ... + a = n \cdot a$

b) $\sum_{i=1}^{n} k x_i = k x_1 + k x_2 + ... k x_n = k(x_1 + x_2 + ... + x_n) = k \sum_{i=1}^{n} x_i$ (ausklammern)

c) $\sum_{i=1}^{n} (x_i + y_i) = (x_1 + y_1) + (x_2 + y_2) + ... + (x_n + y_n) =$

$= (x_1 + x_2 + ... + x_n) + (y_1 + y_2 + ... + y_n) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$

(Reihenfolge der Addition der Summanden beliebig)

Lösung 5.3

Ausgangspunkt ist die Matrix mit den Elementen

$$\begin{array}{cccc} a_{11} & a_{12} & .... & a_{1n} \\ a_{21} & a_{22} & .... & a_{2n} \\ . & . & . & . \\ . & . & . & . \\ a_{n1} & a_{n2} & .... & a_{nn} \end{array}$$

Die Hauptdiagonale wird gebildet durch die Elemente a_{11}, a_{22}, ... , a_{nn}. Oberhalb dieser liegen Werte a_{ij}, für die der erste Index i kleiner ist als der zweite Index j, z. B. a_{12}. Oberhalb der Diagonale liegen dann also die Werte:

$$\begin{array}{cccc} a_{12} & a_{13} & .... & a_{1n} \\ & a_{23} & .... & a_{2n} \\ & & .... & . \\ & & & a_{n-1,n} \end{array}$$

Die Addition dieser Werte ergibt, als Doppelsumme ausgedrückt, bei zeilenweiser Summation:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} a_{ij}$$

Lösung 5.4

Ausführlich geschrieben lautet die Frage: Wann ist
$(x_1 + x_2 + ... + x_n)(y_1 + y_2 + ... + y_n) = x_1y_1 + x_2y_2 + ... +x_ny_n$?

Oder in Matrixform ausgedrückt: Wann ist

$$\begin{array}{cccc} x_1y_1 + & x_1y_2 + & .... & x_1y_n \\ +x_2y_1 + & x_2y_2 + & .... & x_2y_n \\ . & . & . & . \\ . & . & . & . \\ +x_ny_1 + & x_ny_2 & .... & +x_ny_n \end{array} = \begin{array}{cccc} x_1y_1 & & & \\ & +x_2y_2 & & \\ & & .... & \\ & & & +x_ny_n \end{array} \quad ?$$

Man sieht, daß beide Seiten der Gleichung übereinstimmen, wenn die Summe der Werte außerhalb der Hauptdiagonalen gleich Null ist, d. h. wenn gilt:

$$\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} x_i y_j = 0$$

Lösung 5.5

B_{ig} sei die Bevölkerung des Alters i und des Geschlechts g. Die Bevölkerung des Alters 16-65 Jahre Geschlecht lautet dann:

g = 1 (Männer): $\sum_{i=16}^{65} B_{i1}$ (Personen)

g = 2 (Frauen): $\sum_{i=16}^{65} B_{i2}$ (Personen)

Männer und Frauen des Alters 16-65 Jahre insgesamt ergeben sich aus der Summe:

$$\sum_{i=16}^{65} \sum_{g=1}^{2} B_{ig} \text{ (Personen)}$$

Lösung 5.6

a) $\sum_{i=1}^{3} \sum_{j=1}^{4} x_i y_j n_{ij} =$

$$\begin{aligned} &= 10\cdot12\cdot3 \quad +10\cdot14\cdot2 \quad +10\cdot16\cdot0 \quad +10\cdot18\cdot0 \\ &\;\; +15\cdot12\cdot1 \quad +15\cdot14\cdot4 \quad +15\cdot16\cdot1 \quad +15\cdot18\cdot2 \\ &\;\; +20\cdot12\cdot0 \quad +20\cdot14\cdot1 \quad +20\cdot16\cdot3 \quad +20\cdot18\cdot4 \\ &= 5120 \end{aligned}$$

b1)
$$\begin{aligned} n_{1.} &= 3+2+0+0 &&= 5 \\ n_{2.} &= 1+4+1+2 &&= 8 \\ n_{3.} &= 0+1+3+4 &&= 8 \end{aligned}$$

b2)
$$\begin{aligned} n_{.1} &= 3+1+0 &&= 4 \\ n_{.2} &= 2+4+1 &&= 7 \\ n_{.3} &= 0+1+3 &&= 4 \\ n_{.4} &= 0+2+4 &&= 6 \end{aligned}$$

c1) $10 \cdot 5 + 15 \cdot 8 + 20 \cdot 8 = 330$

c2) $10^2 \cdot 5 + 15^2 \cdot 8 + 20^2 \cdot 8 = 5500$

c3) $330^2 = 108900$

d)
$$\begin{aligned} &\frac{3^2}{5 \cdot 4} + \frac{2^2}{5 \cdot 7} \\ &+ \frac{1^2}{8 \cdot 4} + \frac{4^2}{8 \cdot 7} + \frac{1^2}{8 \cdot 4} + \frac{2^2}{8 \cdot 6} \\ &+ \frac{1^2}{8 \cdot 7} + \frac{3^2}{8 \cdot 4} + \frac{4^2}{8 \cdot 6} \\ &= 1{,}6283 \end{aligned}$$

Lösung 5.7

a1) $2 \cdot 4 + 9 \cdot 3 + 4 \cdot 8 + 3 \cdot 6 = 85$

a2) $(2+9+4+3)(4+3+8+6) = 378$

b1) $2^2 + 9^2 + 4^2 + 3^2 = 110$

b2) $(2+9+4+3)^2 = 324$

c1) $\frac{2}{4} + \frac{9}{3} + \frac{4}{8} + \frac{3}{6} = 4{,}5$

c2) $\frac{2+9+4+3}{4+3+8+6} = 0{,}857$

d1) $\ln 2 + \ln 9 + \ln 4 + \ln 3 = 5{,}37528$

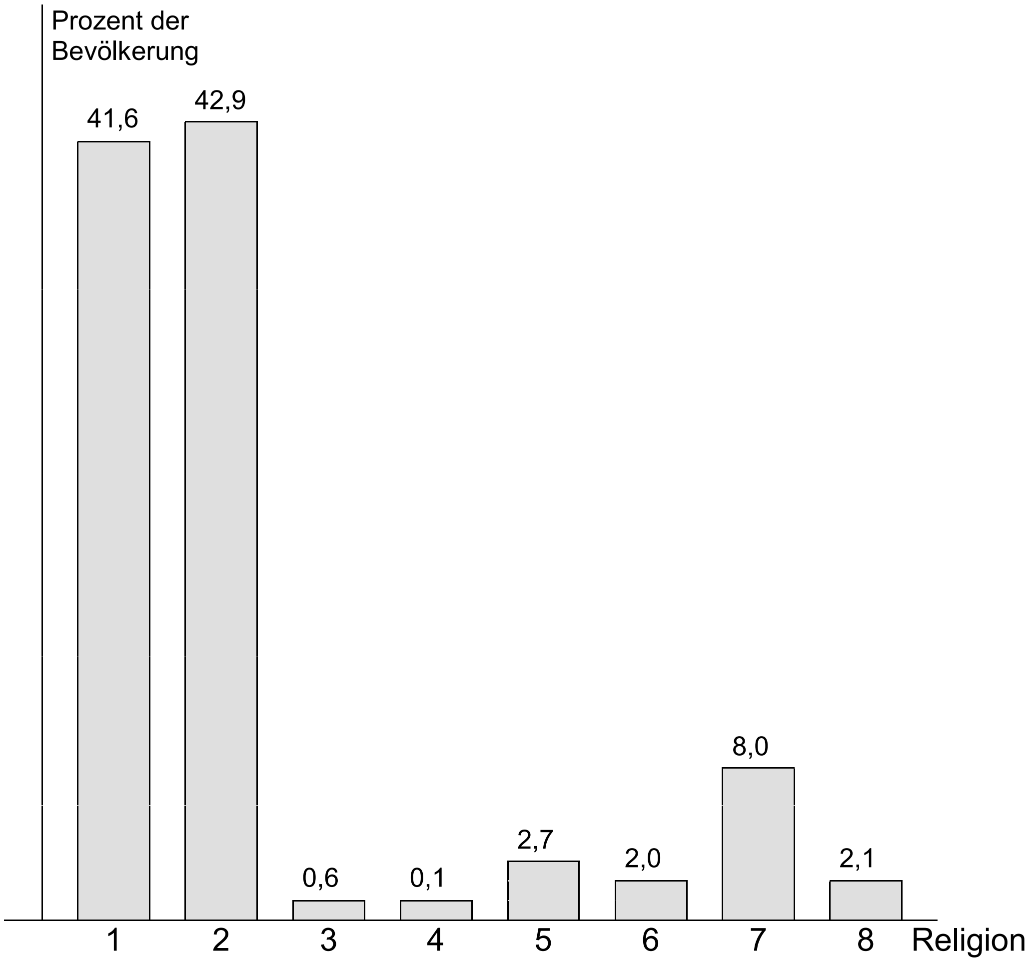
d2) $\ln(2+9+4+3) = \ln 18 = 2{,}89037$

Lösungen zu Kapitel 6

Lösung 6.1

Verteilung der Bevölkerung nach der Religionszugehörigkeit am 25.5.1987			
Nr.	Konfession	Anzahl der Personen (Mio. Personen)	Anteil der Personen (in %)
1	Evangelische Kirche	25,41	41,6
2	Römisch-katholische Kirche	26,23	42,9
3	Evangelische Freikirche	0,39	0,6
4	Jüdische Religionsgesellschaft	0,03	0,1
5	Islamische Religionsgemeinschaft	1,65	2,7
6	Andere Religionsgemeinschaften	1,20	2,0
7	Ohne Religionszugehörigkeit	4,91	8,0
8	Ohne Angabe der Zugehörigkeit	1,26	2,1
	Bevölkerung insgesamt	61,08	100,0
Quelle: Statistisches Jahrbuch für die Bundesrepublik Deutschland 2003, Seite 61			

Verteilung der Bevölkerung nach der Religionszugehörig-
keit am 25.7.1987
(Ergebnis der Volkszählung, früheres Bundesgebiet)
(in % der Bevölkerung)

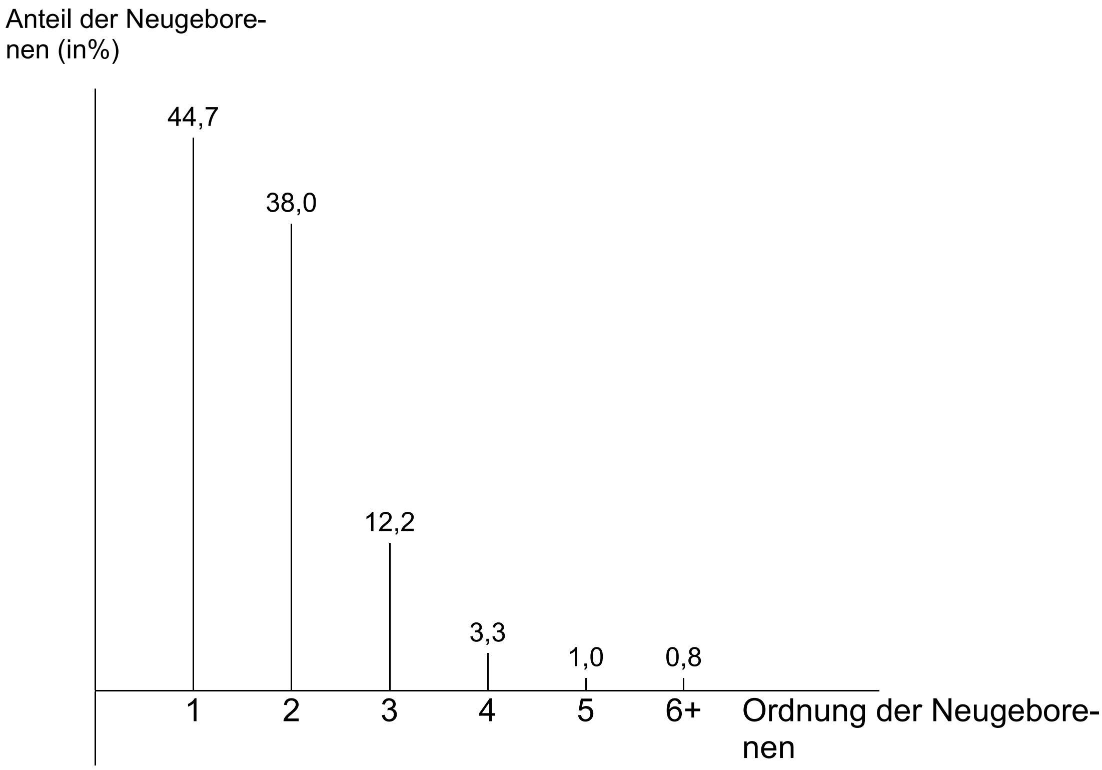


Lösung 6.2

a)

Verteilung der Neugeborenen nach der Geburtenfolge In Deutschland im Jahr 2000		
Ordnung des Kindes	Anzahl der Kinder (in 1.000)	Anteil der Kinder (in %)
1	262,3	44,7
2	223,0	38,0
3	71,6	12,2
4	19,6	3,3
5	5,9	1,0
6 oder höher	5,0	0,8
insgesamt	587,4	100,0

Verteilung der Neugeborenen nach der Geburtenfolge
In Deutschland im Jahr 2000



b) Der überwältigende Anteil der Neugeborenen entfällt auf Erstgeborene (44,7%) und Zweitgeborene (38,0%), insgesamt also 82,7% auf beide Kategorien zusammen. Lediglich 17,3% der Neugeborenen sind Kinder dritter oder höherer Ordnung. Die seh5r hohen Anteile der Erst- und Zweitgeborenen sind eine Folge der sehr niedrigen Fruchtbarkeit in Deutschland. Für eine weitere Beurteilung in dieser Hinsicht sind Vergleiche mit der Verteilung der Neugeborenen in Zeiten oder Ländern mit höherer Fruchtbarkeit als in Deutschland im Jahr 2000 sinnvoll.

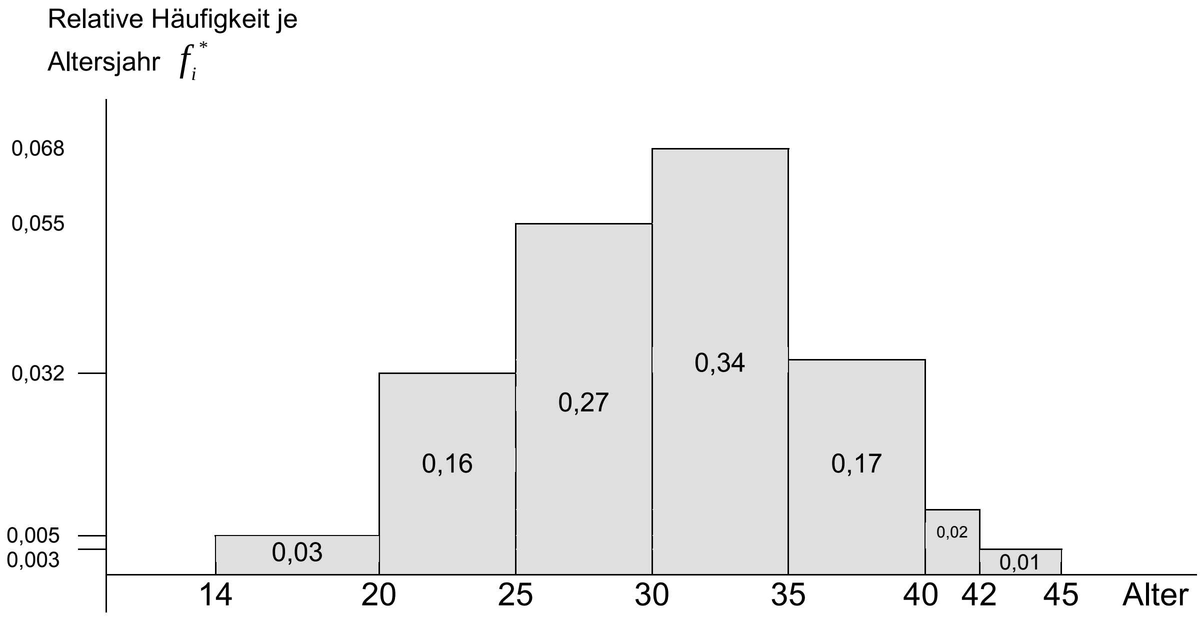
Lösung 6.3

a)

Alter der Mutter von ... bis unter ... Jahre	Anzahl der Kinder n_i (in 1.000)	Klassenbreite b_i	Anteil der Kinder $f_i = \frac{n_i}{\sum_{i=1}^{7} n_j}$	Relative Häufigkeitsdichte $f_i^* = \frac{f_i}{b_i}$
14 - 20	22,9	6	0,03	0,005
20 - 25	118,0	5	0,16	0,032
25 - 30	200,9	5	0,27	0,055
30 - 35	250,3	5	0,34	0,068
35 - 40	121,5	5	0,17	0,033
40 - 42	14,2	2	0,02	0,010
42 - 45	5,9	3	0,01	0,003
	733,7		1,00	

Der Anteil der Neugeborenen in der Klasse der 14 bis unter 20-jährigen Mütter beträgt 0,03 = 3%. Die relative Häufigkeitsdichte dieser Klasse beträgt 0,005, d. h. je Altersjahr dieser Klasse werden 0,5% der Kinder geboren, in der Klasse also insgesamt 6 mal 0,5% = 3%.

Verteilung der Neugeborenen nach dem Alter der Mutter
In Deutschland im Jahr 2001



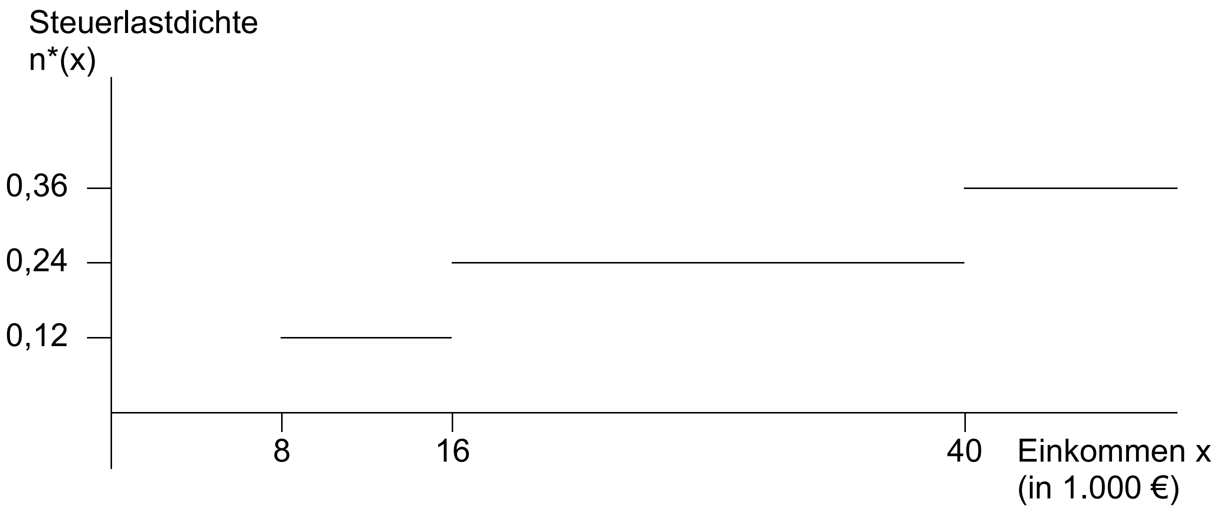
Die Angaben in den Stäben bezeichnen die relativen Häufigkeiten der Neugeborenen in den einzelnen Altersklassen der Mütter; der Maßstab auf der Ordinate gibt die Häufigkeitsdichte der Neugeborenen in den Altersklassen an.

b) Bei 3% der Neugeborenen war die Mutter jünger als 20 Jahre. Bei 20% der Neugeborenen war die mindestens 35 Jahre alt.
c) Die Geburten **verteilen** sich auf das Altersintervall von 14 bis 45 Jahre, also auf eine Spannweite von 31 Jahren. Zunächst steigt die Geburtenintensität an bis zur Altersklasse 30 bis unter 35 Jahre, und sinkt dann steil ab bis zur letzten Altersklasse 42 bis unter 45 Jahre.
Der **häufigste Wert** der Geburtenintensität liegt bei der Klasse 30-35, ein hohes Alter für das Maximum der Geburtenintensität. Auch dieses Ergebnis ist eine Begleiterscheinung der niedrigen Fruchtbarkeit in Deutschland. Ergänzende Vergleiche mit Perioden oder Ländern höherer Fruchtbarkeit erscheinen auch hier sinnvoll.

Lösung 6.4

a) Bei Häufigkeitsverteilungen wird den einzelnen Merkmalswerten die Anzahl der Fälle ihres Auftretens zugeordnet. In diesem Beispiel ist dies nicht der Fall. Hier wird jedem Merkmalswert, jedem Einkommenswert, die Anzahl der zu bezahlenden Euro an Steuern zugeordnet. So besagt die Aufgabenstellung, daß im Bereich zwischen 8.000 € und 16.000 € 0,12 Euro Steuern je Euro Einkommen zu bezahlen sind. Insgesamt hätte dann ein Einkommensbezieher mit 12.000 Euro für den Bereich zwischen 8.000 Euro und 12.000 Euro 4.000 mal 0,12 = 480 Euro Steuern zu bezahlen. Interpretiert man die Steuerlast von 0,12 Euro je Euro Einkommen als Häufigkeit, so kann man die vorliegenden Angaben als Häufigkeitsverteilung, die Steuerangaben als Werte einer absoluten Häufigkeitsdichtefunktion n*(x) interpretieren und entsprechend behandeln. Die Häufigkeitsdichtefunktion ("Steuerlastdichtefunktion") des Merkmals Einkommen (X) lautet dann gemäß den Angaben:

$$n^*(x) = \begin{cases} 0 & \text{für} \quad x < 8.000 \\ 0{,}12 & \text{für} \quad 8.000 \le x < 16.000 \\ 0{,}24 & \text{für} \quad 16.000 \le x < 40.000 \\ 0{,}36 & \text{für} \quad x \ge 40.000 \end{cases}$$



b) Die Häufigkeitssummenfunktion N(x) für einen Einkommenswert x ergibt sich anschaulich als Fläche unter der Steuerlastdichtefunktion bis zum Einkommenswert x. Daraus kann man einen einfachen Weg zu ihrer Berechnung finden. Für die jeden Abschnitt der Steuerlastdichte berechnet man zwei Punkte der Summenfunfktion und verbindet diese durch eine Gerade nach der Zwei-Punkte-Formel der Geradengleichung.

Erster Punkt (x_1, y_1):
An der Stelle x_1 = 8.000 ist die Steuerlast gleich Null, also gilt **(x_1, y_1) = (8000; 0)**

Zweiter Punkt (x_2,y_2):
An der Stelle x_2 = 16.000 ist die Steuerlast (16000-8000)·0,21 = 960 €. Somit ist **(x_2,y_2) = (16000; 960)**

Dritter Punkt (x_3, y_3):
An der Stelle x_3 = 40.000 beträgt die Steuerlast y_3= 960 + (40000-16000)·0,24 = 6720 €. Also ist **(x_3, y_3) = (40000; 6720)**.

Vierter Punkt (x_4, y_4):
Beispielsweise an der Stelle x_4 = 50.000 beträgt die Steuerlast y_4 = 6720 + (50000 - 40000)·0,36 = 10320 €. Also ist **(x_4, y_4) = (50000; 10320)**

Zwischen diesen vier Punkten ist die Summenfunktion geradlinig, da die Steuerlastdichte eine Parallele zur Abszisse ist. Nach der Zwei-Punkte-Formel der Geradengleichung bestimmt man die Summenfunktion. Die Zwei-Punkte-Formel lautet:

$$\frac{y-y_1}{x-x_1}=\frac{y_2-y_1}{x_2-x_1}$$

Angewendet auf die beiden ersten Punkte erhält man die Gleichung:

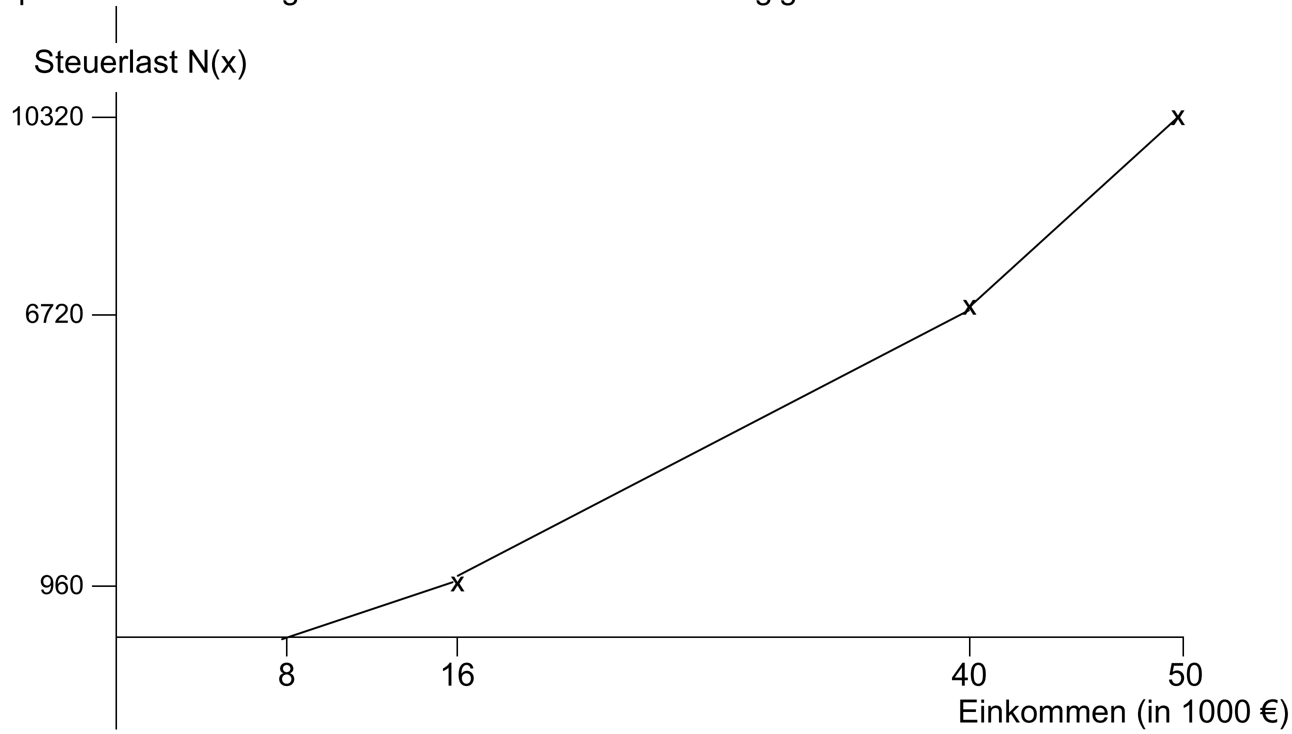
$$\frac{y-0}{x-8000}=\frac{960-0}{16000-8000}$$

Die Auflösung der Gleichung nach y ergibt y = 0,12x - 960
Entsprechend verwendet man für die beiden weiteren Abschnitte den zweiten und dritten Punkt bzw. den dritten und vierten Punkt. Als Ergebnis erhält man die Summenfunktion N(x) wie folgt:

$$N(x)=\begin{cases} 0 & \textit{für} \quad x < 8.000 \\ 0{,}12x-960 & \textit{für} \quad 8.000 \le x < 16.000 \\ 0{,}24x-2880 & \textit{für} \quad 16.000 \le x < 40.000 \\ 0{,}36x-7680 & \textit{für} \quad x \ge 40.000 \end{cases}$$

Hinweis: Zu dieser Funktion kommt man auch durch abschnittweises Integrieren über die Steuerlastdichtefunktion (vgl. hierzu die Ausführungen im Skript.

Graphische Darstellung der Steuerlastfunktion in Abhängigkeit vom Einkommen:



c) Dividiert man die Steuerlast durch da Einkommen, so erhält man den Prozentsatz des Einkommens, der an Steuern abgeführt werden muß, also den Steuersatz in Abhängigkeit vom Einkommen. Diese Funktion t(x) = N(x)/x läßt sich unmittelbar angeben:

$$t(x) = \frac{N(x)}{x} = \begin{cases} 0 & \text{für} \quad x < \ 8.000 \\ 0{,}12 - \dfrac{960}{x} & \text{für} \ 8.000 \le x < \ 16.000 \\ 0{,}24 - \dfrac{2880}{x} & \text{für} \ 16.000 \le x < 40.000 \\ 0{,}36 - \dfrac{7680}{x} & \text{für} \quad x \ge 40.000 \end{cases}$$

Der Steuersatz ist also Null für x = 8.000, steigt bis zu einem Einkommen von **16.000 €** auf 0,12 - 960/16.000 = 0,06 = **6%** an und steigt dann weiter bis zum Einkommen von **40.000 €** auf 0,24 - 2.880/40.000 = 0,168 = **16,8%**. Der Steuersatz steigt dann weiter bis zur Obergrenze von 0,36 = 36% bei unendlich hohem Einkommen an. Für ein Einkommen von **50.000 €** beträgt der Steuersatz 0,36 - 7680/50.000 = 0,2064 = **20,64%.**

Lösung 6.5

Tabelle ..... Wertpapierbestand des Unternehmens U 31.03.1999		
Emittierende Firma	Art des Wertpapiers	
	Aktien (in Stück)	Anleihen (in €)
BMW Bayer Siemens VW	15 20 40 20	- 1.000,- - 500,-
Insgesamt	95	1.500,-
Quelle: Geschäftsbericht des Unternehmens U vom 31.03.1999		

Lösung 6.6

a) Im Histogramm wird auf der Abszisse das betrachtete Merkmal und die zugehörigen Merkmalsklassen abgetragen. Auf der Ordinate wird die Häufigkeitsdichte einer jeden Merkmalsklasse, d. h. die Häufigkeiten je Merkmalseinheit in der Klasse abgetragen.
Im Fall des Altersaufbaus der Bevölkerung haben wir das Merkmal Alter, das im vorliegenden Beispiel in Einjahresklassen aufgeteilt ist. Dargestellt sind also die Anzahl der Personen je Altersjahr (Alter von 0 bis unter 1Jahr, 1 bis unter 2 Jahre usw.). Im Unterschied zur "normalen" Darstellung ist hier allerdings das Merkmal nicht auf der Abszisse sondern auf der Ordinate abgetragen und die Häufigkeitsdichte auf der Abszisse. Dadurch kann man die Histogramme für Männer und Frauen in einer Graphik darstellen.

b) Die Ursachen der **Einkerbungen** sind in der Graphik angegeben.
Der **Bevölkerungsaufbau** nimmt immer stärker die Form einer Urne an: die niedrige Fruchtbarkeit in Deutschland führt dazu, daß die Besetzung der jungen Altersklassen immer geringer wird, weil immer weniger Kinder geboren werden. Der Männerüberschuß in den jüngeren Altersjahren und Frauenüberschuß in den höheren Altersjahren läßt sich wie folgt erklären: Unter den Neugeborenen ist der Knabenanteil höher (in Deutschland etwa 51,4%) als der Mädchenanteil (in Deutschland demnach etwa 48,6%). Andererseits ist in allen Altersklassen die Sterblichkeit der Männer höher als die der Frauen. Daher wird der durch die Neugeborenen erzeugte Männerüberschuß im Laufe des Alterns durch die höhere Sterblichkeit der Männer "aufgezehrt" und wird zu einem Frauenüberschuß. Daneben gibt es weitere Ursachen, die das Verhältnis der Anzahl von Männern und Frauen beeinflussen (z. B. Wanderungen, Kriege usw.).

Lösungen zu Kapitel 7

Lösung 7.1

Gesucht: $\tilde{x}_{0,10}$

$\tilde{x}_{0,10}$ liegt in der Klasse Nr. 2, denn:

$$F(x_1^0) = \frac{22{,}9}{733{,}7} = 0{,}03 \leq 0{,}10 < 0{,}19 = \frac{22{,}9 + 118{,}0}{733{,}7} = F(x_2^0)$$

Dann ist:

$$\tilde{x}_{0,10} = x_1^0 + b_2 \cdot \frac{0{,}10 - F(x_1^0)}{f_2} = 20 + 5 \cdot \frac{0{,}10 - 0{,}03}{0{,}16} = 22{,}19\,[Jahre]$$

Bis zum Alter der Müttervon 22,2 Jahren werden 10% der Kinder zur Welt gebracht.

b) *Gesucht*: $\tilde{x}_{0,90}$

$\tilde{x}_{0,90}$ liegt in der Klasse Nr. 5, denn:

$$F(x_4^0) = \frac{22{,}9 + 118{,}0 + 200{,}9 + 250{,}5}{733{,}7} = 0{,}80 \leq 0{,}90 < 0{,}0{,}97 = \frac{22{,}9 + \ldots 121{,}5}{733{,}7} = F(x_5^0)$$

Dann ist:

$$\tilde{x}_{0,90} = x_4^0 + b_5 \cdot \frac{0{,}90 - F(x_4^0)}{f_5} = 35 + 5 \cdot \frac{0{,}90 - 0{,}80}{0{,}17} = 37{,}94\,[Jahre]$$

D. h. bis zur Vollendung des 38. Lebensjahres brachten die Mütter 90% der Kinder zur Welt.

c) In das Altersintervall [22,2; 38,0] der Mütter fallen 80% der Neugeborenen.

Lösung 7.2

Gesucht: $\tilde{x}_{0,75}$

Das dritte Quartil entfällt auf die Ordnung Nr. 2 der Geburten, denn für Zweitgeborene gilt:

$$F(x=1) = \frac{262{,}3}{587{,}4} = 0{,}45 < 0{,}75 < 0{,}83 = \frac{262{,}3 + 223{,}0}{587{,}4} = F(x=2)$$

D. h. etwa 75% der Neugeborenen weisen die Geburtenfolge 1 oder 2 auf.

Lösung 7.3

Gesucht: $\tilde{x}_{0,25}$
n·α = 8 · 0,25 = 2 (ganzzahlig); daher istzu wählen:
i = n·α+1 = 3; Folglich ist:

$$\tilde{x}_{0,25} = \frac{1}{2}(x_{e_2} + x_{e_3}) = \frac{1}{2}(770 + 785) = 777{,}5 \ [€]$$

D. h. etwa 25% der Preise betragen 777,5 [€] oder weniger.

Lösungen zu Kapitel 8

Lösung 8.1

a) $$\bar{x} = \frac{15 \cdot 450 + 16 \cdot 1400 + 17 \cdot 3300 + ... + 44 \cdot 900}{733700} = \frac{21747750}{733700} = 29{,}64\,[Jahre]$$

Das Durchschnittsalter beträgt ca. 30 Jahre.

b) Das mittlere Alter, gemessen durch den Median, beträgt 30 Jahre, denn:

$$F(x = 29) = \frac{450 + 1400 + ... + 47750}{73370} = 0{,}4659$$

$$F(x = 30) = \frac{450 + 1400 + ... + 47750 + 52550}{73370} = 0{,}5375$$

Folglich gilt:

$$F(x = 29) = 0{,}45 < 0{,}50 < 0{,}5375 \ = F(x = 30)$$

Also ist der Median

$$\tilde{x} = 30\,[Jahre]$$

c) Graphik, Stabdiagramm

d) $$\bar{x}_H = \frac{733700}{\frac{1}{1} \cdot 450 + \frac{1}{3{,}1} \cdot 1400 + ... + \frac{1}{1{,}5} \cdot 900} = 44{,}98\,[Jahre]$$

Begründung für die Wahl des harmonischen Mittels (**Faustregel**): Es werden Beziehungszahlen gemittelt (Geborene je 1000 Frauen) und die Größe im Zähler der Beziehungszahl als Gewicht verwendet.

Begründung **inhaltlich**: Im Zähler von $\bar{x}_H$ stehen die Geborenenzahlen, im Nenner die Gesamtzahl der Frauen, denn z. B 1400/3,1 ergibt die Anzahl der 16-jährigen Frauen. Anzahl der Geborenen dividiert durch die Anzahl der Frauen (in 1000) ergibt dann die Anzahl der Geborenen je 1000 Frauen im Durchschnitt aller Altersjahre.

e) TFR = 1,0+3,1+7,5+14,1+24,6+ ....+1,5 = 1347,3 [Kinder]
Die Summe der altersspezifischen Fruchtbarkeitsraten heißt **Totale Fruchtbarkeitsrate** (**TFR**). Sie gibt an, wieviele Kinder von 1000 Frauen im Laufe ihres Lebens zur Welt gebracht werden, wenn man von der Sterblichkeit der Mütter absieht. Im vorliegenden Beispiel für das Jahr 2001 waren es 1347 Kinder. Knapp die Hälfte davon (48,6%) sind Mädchen. Das bedeutet, 1000 Frauen bringen etwa 655 Mädchen im Laufe ihres Lebens zur Welt. Wenn man die Sterblichkeit der Mütter im Laufe des gebärfähigen Alters berücksichtigt, sind es noch etwas weniger Mädchen. Die Zahl der Neugeborenen Mädchen ist also nicht ausreichend, um die 1000 Frauen der Müttergeneration zu ersetzen. Die Bevölkerung wird langfristig sinken, wenn nicht Wanderungen das Geburtendefizit ausgleichen.

Lösung 8.2

a) Bei Verwendung der Klassenmitten zur Berechnung erhält man den Mittelwert:

$$\bar{x} = \frac{30 \cdot 2 + 50 \cdot 16 + ... + 150 \cdot 3}{100} = 83{,}40 \text{ [€]}$$

Der durchschnittliche Tageslohn je Beschäftigten beträgt 83,40 [€]

b) $\bar{x}$ sei der durchschnittliche Tageslohn **vor** der Lohnerhöhung, $\bar{y}$ der durchschnittliche Tageslohn **nach** der Lohnerhöhung. Dann ergibt sich:
b1) $\bar{y}$ = $\bar{x}$ + 10 = 93,40 [€]
b2) $\bar{y}$ = $1{,}10\,\bar{x}$ = 91,74 [€]
b3) $\bar{y}$ = $1{,}08\,\bar{x} + 5$ = 95,07 [€]

Lösung 8.3

Betrieb A:

$$\bar{w}_A = \sqrt[4]{\frac{256}{120}} - 1 = 1{,}2086 - 1 = 0{,}2086 = 20{,}86\,\%$$

Betrieb B:

$$\bar{w}_B = \sqrt[4]{1{,}24 \cdot 1{,}08 \cdot 0{,}97 \cdot 1{,}29} - 1 = 1{,}1378 - 1 = 0{,}1378 = 13{,}78\,\%$$

Die durchschnittliche jährliche Wachstumsrate ist also in Betrieb A mit 20,86% höher als in Betrieb B mit 13,78%.

Lösungen zu Kapitel 9

Lösung 9.1

a) R = x_{max} - x_{min} = 45 - 14 = 31 [Jahre]

b) $\tilde{x}_{0,10} = x_1^0 + b_2 \cdot \dfrac{\alpha - F(x_1^0)}{f_2} = 20 + 5 \cdot \dfrac{0{,}10 - 0{,}03}{0{,}16} = 22{,}19\,[Jahre]$

$\tilde{x}_{0,90} = x_4^0 + b_5 \cdot \dfrac{\alpha - F(x_4^0)}{f_5} = 35 + 5 \cdot \dfrac{0{,}90 - 0{,}81}{0{,}16} = 37{,}81\,[Jahre]$

$R_{80} = \tilde{x}_{0,9} - \tilde{x}_{0,10} = 37{,}81 - 22{,}19 = 15{,}62\,[Jahre]$

Lösung 9.2

a) Berechnung der Rundenzeiten für Fahrer II (1 Runde = 50 km) in min/50km aus den Durchschnittsgeschwindigkeiten.

1. Runde:

$$198\,[\frac{km}{h}] \rightarrow \frac{1}{198{,}3}[\frac{h}{km}] = \frac{50}{198{,}3}[\frac{h}{50km}] = \frac{50 \cdot 60}{198{,}3}[\frac{\min}{50km}] = 15{,}13\,[\frac{\min}{50km}]$$

2. Runde

$$199{,}9\,[\frac{km}{h}] \rightarrow \frac{50 \cdot 60}{199{,}9}[\frac{\min}{50km}] = 15{,}01[\frac{\min}{50km}]$$

3. Runde

$$205{,}7\,[\frac{km}{h}] \rightarrow \frac{50 \cdot 60}{205{,}7}[\frac{\min}{50km}] = 14{,}58\,[\frac{\min}{50km}]$$

4. Runde

$$203{,}6\,[\frac{km}{h}] \rightarrow \frac{50 \cdot 60}{203{,}6}[\frac{\min}{50km}] = 14{,}73\,[\frac{\min}{50km}]$$

5. Runde

$$202{,}5\,[\frac{km}{h}] \rightarrow \frac{50 \cdot 60}{202{,}5}[\frac{\min}{50km}] = 14{,}81[\frac{\min}{50km}]$$

b) $\bar{x}_I = \dfrac{15{,}3 + 15{,}2 + 14{,}9 + 14{,}6 + 15{,}0}{5} = 15{,}0\ [\dfrac{\min}{50\,km}]$

$\bar{x}_{II} = \dfrac{15{,}13 + 15{,}01 + 14{,}58 + 14{,}73 + 14{,}81}{5} = 14{,}85\ [\dfrac{\min}{50\,km}]$

Fahrer II hat die niedrige durchschnittliche Rundenzeit.

c) $s_I^2 = \frac{1}{5}\sum_{i=1}^{5} x_{Ii}^2 - \overline{x}_I^2 = \frac{1}{5}\cdot 225 - 15^2 = 0{,}06\,[(\frac{\min}{50km})^2]$

$s_{II}^2 = \frac{1}{5}\sum_{i=1}^{5} x_{IIi}^2 - \overline{x}_{II}^2 = \frac{1}{5}\cdot 1103{,}1 - 14{,}85^2 = 0{,}0975\,[(\frac{\min}{50km})^2]$

d) $v_s(I) = \frac{s_I}{\overline{x}_I} = \frac{\sqrt{0{,}06}}{15} = 0{,}016 = 1{,}6\%$

$v_s(II) = \frac{s_{II}}{\overline{x}_{II}} = \frac{\sqrt{0{,}0975}}{14{,}85} = 0{,}021 = 2{,}1\%$

Die Standardabweichung der Rundenzeiten des Fahrers I beträgt 1,6% seiner durchschnittlichen Rundenzeit. Der entsprechende Wert von Fahrer II beträgt 2,1%. Fahrer I erzielte also gleichmäßigere Rundenzeiten als Fahrer II.

e) Gleichmäßigere Rundenzeiten drücken eine höhere Konzentrationsfähigkeit über die längere Zeit aus. Sie ist bei Fahrer I stärker ausgeprägt als bei Fahrer II. Die schnellste Rundenzeit ist bei beiden Fahrern gleich.

Lösung 9.3

a) $\overline{x} = \frac{60\cdot 18 + 84\cdot 61 + 108\cdot 88 + 132\cdot 49 + 156\cdot 27 + 190\cdot 7}{250} = 110{,}872 \approx 111[h]$

(Verwendung der Klassenmitten zur Berechnung des Mittelwertes)

b) R = 192 - 48 = 144 [h]

c) $s_{X,n}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 n_i - \overline{x}^2 = \frac{3.285.196}{250} - 111^2 = 848{,}18[h^2]$

$s_{X,n} = \sqrt{848{,}18} = 29{,}12[h]$

d) Je niedriger die Streuung ist, desto gleichmäßiger und störungsfreier ist der Durchlauf.

Lösung 9.4

a) $\overline{x}_1 = 12[\min]$

$\overline{x}_2 = 20[\min]$

$\overline{x}_{1+2} = 12 + 20 = 32\ [\min]$

Die Daten können z. B. zur Festlegung von Akkordzeiten genutzt werden.

b) Zum Vergleich der Streuung von Verteilungen verwendet man sinnvollerweise den Variationskoeffizienten.

$$s_1^2 = \frac{1}{6}\sum_{i=1}^{6} x_{1i}^2 - \bar{x}_1^2 = \frac{88}{6} - 12^2 = 4[\min^2]$$

$$s_2^2 = \frac{1}{6}\sum_{i=1}^{6} x_{2i}^2 - \bar{x}_2^2 = \frac{2.454}{6} - 20^2 = 9[\min^2]$$

$$s_{1+2}^2 = \frac{1}{6}\sum_{i=1}^{6} x_{(1+2)i}^2 - \bar{x}_{(1+2)}^2 = \frac{6.182}{6} - 32^2 = 6,\bar{3}[\min^2]$$

$$v_s(1) = \frac{\sqrt{4}}{12} = 0,17$$

$$v_s(2) = \frac{9}{20} = 0,0,15$$

$$v_s(1+2) = \frac{\sqrt{6,\bar{3}}}{32} = 0,08$$

Arbeitsgang Nr. 1 weist die größte relative Streuung auf.

c) Die Streuung der Arbeitszeiten die zeigt die Ungleichheit in der Geschicklichkeit (Vorgang 1) bzw. der Aufmerksamkeit (Vorgang 2) der Arbeitnehmer. Betriebliche Ausbildung ist eine Möglichkeit zur Reduktion der Ungleichheit.

Lösung 9.5

a) Es ist eine allgemeine Erfahrungstatsache, daß in Verteilungen mit steigendem Mittelwert die Streuung, gemessen durch ein absolutes Streuungsmaß wie die Varianz oder die Standardabweichung ansteigt. steigt. Der Mittelwert des Kurses der Porsche-Aktie ist demnach größer als derjenige der Daimler-Chrysler-Aktie Ein Vergleich der Streuung von Verteilungen mit unterschiedlichem Mittelwert durch ein absolutes Streuungsmaß ist daher wenig sinnvoll. Man verwendet statt dessen ein relatives Streuungsmaß, bei dem Die absolute Streuung auf einen Mittelwert bezogen wird. Dies geschieht beispielsweise durch die Verwendung des **Variationskoeffizienten**, zu dessen Berechnung die Standardabweichung auf das arithmetische Mittel bezogen wird.

b) $$v_s(DC) = \frac{36,18}{50,59} = 0,72 = 72\%$$

$$v_s(Po) = \frac{182,96}{396,10} = 0,46 = 46\%$$

Die Berechnung zeigt, daß die relative Streuung der Porsche-Aktie kleiner ist als die der DC-Aktie, obwohl es sich bei den absoluten Streuungen umgekehrt verhält. Beide relative Streuungen sind sehr hoch. Der Kurs der DC-Aktie schwankt durchschnittlich 72% um den Kursmittelwert, bei der Porsche-Aktie sind es durchschnittlich 46%.

c) Hohe Streuung bedeutet hohe Schwankungsintensität (hohe "**Volatilität**")der Kurse und damit je nach Kurslage die Möglichkeit hohe Gewinne zu erzielen oder hohe Verluste zu erleiden, also hohes Risiko.

Lösungen zu Kapitel 10

Lösung 10.1

a)

$f_{y_j \mid x_i}$	Raucherverhalten Y		Summe
	y_1 = Raucher	y_2 = Nichtraucher	
$f_{y_j \mid x_1 = m}$	$\frac{36}{80} = 0{,}45$	$\frac{44}{80} = 0{,}55$	1,00
$f_{y_j \mid x_2 = w}$	$\frac{27}{60} = 0{,}45$	$\frac{33}{60} = 0{,}55$	1,00

b)
Da die bedingten Verteilungen des Merkmals Raucherverhalten (Y) für Männer und Frauen gleich sind, ist für die vorliegende Gesamtheit das Raucherverhalten unabhängig vom Geschlecht.

c)

	Raucherverhalten Y		Summe
	y_1 = Raucher	y_2 = Nichtraucher	
$n(y)$	63	77	140
$f(y)$	0,45	0,55	1,00

Die relative Randverteilung ist gleich den beiden bedingten relativen Häufigkeitsverteilungen.

Lösung 10.2

a)

$f(y_j \mid x_i)$	Herstellkosten					
	$y_1 = 15$	$y_2 = 20$	$y_3 = 25$	$y_4 = 25$	$y_5 = 30$	$\overline{y} \mid x_i$
$f(y_j \mid x_1 = 40)$	0,40	0,60	0	0	0	18
$f(y_j \mid x_2 = 50)$	0,30	0,40	0,30	0	0	20
$f(y_j \mid x_3 = 60)$	0,20	0,40	0,30	0,10	0	21,5
$f(y_j \mid x_4 = 70)$	0	0,40	0,45	0,10	0,05	24
$f(y_j \mid x_5 = 80)$	0	0,20	0,20	0,40	0,20	28

b) Die Herstellkosten steigen mit steigender Produktionsmenge. Dies ist aus der Entwicklung der bedingten Verteilungen des Merkmals Y und aus den bedingten Mittelwerten zu ersehen.

Lösungen zu Kapitel 11

Lösung 11.1

a) Die Kinderzahl ist das unabhängige Merkmal; von ihr ist die Höhe der Konsumausgaben abhängig.

b) Bedingte Häufigkeitsfunktionen und bedingte Mittelwerte

$f(y_j \mid x_i)$	Monatliche Ausgaben Y						$\bar{y} \mid x_i$
	y_1=1500	y_1=1600	y_1=1700	y_1=1800	y_1=1900	y_1=2000	
$f(y_j \mid x_1)$	0,625	0,3125	0,0625	0	0	0	1544
$f(y_j \mid x_2)$	0,1111	0,6667	0,2222	0	0	0	1611
$f(y_j \mid x_3)$	0,0667	0,6667	0,1667	0,1000	0	0	1630
$f(y_j \mid x_4)$	0	0,2778	0,5556	0,1667	0	0	1689
$f(y_j \mid x_5)$	0	0	0,2500	0,5000	0,2500	0	1800
$f(y_j \mid x_6)$	0	0	0,2500	0,4375	0,3125	0	1806
$f(y_j \mid x_7)$	0	0	0	0,0385	0,1923	0,7692	1973

Die allgemeine Regressionsfunktion ist die geradlinige Verbindung der (x_i, $\bar{y} \mid x_i$)-Wertepaare.

Graphik zeichnen (vgl. Skript S. 114): die (x_i, $\bar{y} \mid x_i$)-Wertepaare in die (x,y)-Ebene eintragen und geradlinig verbinden.

c)

$\sum_{i=1}^{7} x_i n_{i.}$ = 555; $\sum_{i=1}^{7} x_i^2 n_{i.}$ = 2383; $\sum_{j=1}^{6} y_j n_{.j}$ = 346300

$\sum_{i=1}^{7}\sum_{j=1}^{6} x_i y_j n_{ij}$ = 1002300; $\sum_{j=1}^{6} y_j^2 n_{.j} = 595390000$

$$a_1 = \frac{203\sum_{i=1}^{7}\sum_{j=1}^{6} x_i y_j n_{ij} - \sum_{i=1}^{7} x_i n_{i.} - \sum_{j=1}^{6} y_j n_{.j}}{203\sum_{i=1}^{7} x_i^2 n_{i.} - (\sum_{i=1}^{7} x_i n_{i.})^2} = \frac{203 \cdot 1002300 - 555 \cdot 346300}{203 \cdot 2383 - 555^2} =$$

$$64{,}1369\,[€/\,Kind]$$

d. h. die monatlichen Ausgaben nehmen mit einem zusätzlichen Kind durchschnittlich um 64,14 [€] zu.

$$a_0 = \frac{\sum_{j=1}^{6} y_j n_{.j} - a_1 \sum_{i=1}^{7} x_i n_{i.}}{203} = \frac{346300 - 64{,}1369 \cdot 555}{203} = 1530{,}5617\,[€]$$

d. h. bei Null Kindern betragen die durchschnittlichen Ausgaben 1530,56 [€].

Die Regressionsgleichung lautet somit:

$$\hat{y} = 1530{,}56 + 64{,}14$$

(Die Gerade in die Graphik einzeichnen!)

Lösung 11.2

a) Graphik zeichnen; schildern, wie sie sich im Zeitablauf entwickelt und mögliche gesellschaftliche oder ökonomische Ursachen dafür nennen.

b)

Jahr	x_{e_i}	y_{e_i}	$x_{e_i} y_{e_i}$	$x_{e_i}^2$
1970	0	24,3	0	0
1975	5	24,8	124	25
1980	10	25,2	252	100
1985	15	26,2	393	225
1990	20	26,9	538	400
1995	25	28,2	705	625
Summe	75	155,6	2012	1375

$$a_1 = \frac{6\sum_{i=1}^{6} x_{e_i} y_{e_i} - \sum_{i=1}^{6} x_{e_i} \sum_{i=1}^{6} y_{e_i}}{6\sum_{i=1}^{6} x_{e_i}^2 - (\sum_{i=1}^{6} x_{e_i})^2} = 0{,}1531\,\text{[Altersjahre/Jahrfünft]}$$

$$a_0 = \frac{\sum_{i=1}^{6} y_{e_i} - a_1 \sum_{i=1}^{6} x_{e_i}}{6} = 24{,}02\,\text{[Altersjahre]}$$

a_0 formal: Achsenabschnitt; inhaltlich: Trendwert 1970
a_1 formal: Steigung der Trendgeraden; inhaltlich: durchschnittliche Veränderung des Erstgeburtsalters je Jahrfünft

c) $\hat{y}(1999) = \hat{y}(29) = 24{,}02 + 0{,}1531 \cdot 29 = 28{,}5$ [Altersjahre
tatsächliches Ergebnis: 28,9 [Altersjahre]

d) Ungeeignet, weil:
formal: Wertepaare legen eher eine gekrümmte Regressionsgerade (Trendgerade) nahe. Man sagt, ein linearer Trend stellt eine **Fehlspezifikation** der Regressionsfunktion dar.
inhaltlich: weitere starke Erhöhung stößt an biologische Grenze. Dieses Argument widerspricht dem formalen, der linearen oder gekrümmten Trendfunktion

e) Per Augenschein paßt sich die Gerade sehr gut an die Wertepaare an, sie besitzt also hohe Anpassungsqualität (dies kann auch durch Berechnung des Determinationskoeffizienten gezeigt werden; dieser beträgt $r_{xy}^2 = 0{,}96$, er wird jedoch erst in Kap. 12 eingeführt). Wegen der Bedenken in d) besitzt der lineare Trend jedoch wenig Prognosequalität.

Lösung 11.3

a1)

Jahr	$\frac{BIP_r(t)}{BIP_r(1996)}$	$\frac{ET(t)}{ET(1996)}$	$\frac{ETSt(t)}{ETSt(1996)}$
1996	100,0	100,0	100,0
1997	95,9	99,7	99,5
1998	103,4	100,8	100,0
1999	105,5	102,1	100,5
2000	108,5	103,8	101,3
2001	109,5	104,6	100,7
1002	109,6	103,8	99,6

Graphische Darstellung!

a2)

1. Variante: Produktionssteigerung je Erwerbstätigen

$$\sqrt[6]{\frac{\frac{BIP_r(2002)}{BIP_r(1996)}}{\frac{ET(2002)}{ET(1996)}}} - 1 \quad = 0{,}009 = 0{,}9\% \; je\ Jahr$$

2. Variante

$$\sqrt[6]{\frac{\frac{BIP_r(2002)}{BIP_r(1996)}}{\frac{ETSt(2002)}{ETSt(1996)}}} - 1 \quad = 0{,}016 = 1{,}6\% \; je\ Jahr$$

Variante 2 ist ein genaues Maß für die geleistete Arbeitszeit. Variante 1 ist dafür ungenau, weil es nicht zwischen Ganztagsbeschäftigten und Teilzeitbeschäftigten unterscheidet.

a3) Das Wachstum des realen Bruttoinlandsprodukts betrug im betrachteten Zeitraum von 6 Jahren insgesamt 9,6%, also etwa 1,5% je Jahr. Die Erwerbstätigenzahl stieg um 3,8%, die Anzahl der geleisteten Erwerbstätigenstunden blieb in etwa konstant. Die Erhöhung der Beschäftigtenzahl im betrachteten Zeitraum erfolgte demnach vor allem durch Erhöhung der Teilzeitbeschäftigung. Die Produktionssteigerung war also im wesentlichen auf Grund des Produktivitätsfortschritts möglich. Die Produktion je Erwerbstätigenstunde erhöhte sich um durchschnittlich 1,6% je Jahr und damit in gleicher Weise wie das Produktionsergebnis, so daß für die Erhöhung der Erwerbstätigkeit kein Raum blieb.

b1) Berechnung der Regressionsgeraden

b2) Graphische Darstellung

b3) Bei Steigerung der realen Produktion um 1% steigt im Untersuchungszeitraum die Erwerbstätigkeit um durchschnittlich 0,88%

b4) $w_{ET} = 0$ -> $0 = -0{,}74 + 0{,}88\, w_{BIPr}$, d.h.

$w_{BIPr} = 0{,}74/0{,}88 = 0{,}84$

d. h. das BIP muß real um mindestens 0,84% wachsen, um ein Wachstum der Erwerbstätigkeit zu erzeugen.

Lösung 11.4

a) Der Anteil der Ausgaben für Wohnungsmiete (y) am ausgabefähigen Einkommen (x) nimmt mit steigendem Einkommen ab (Schwabe´sches Gesetz).

b) Berechnung der Regressionsgeraden

c) Der Anteil der Ausgaben für Wohnungsmiete am ausgabefhigen Einkommen beträgt y/x; unter Verwendung der Regressionsgeraden erhält man:

$$\frac{\hat{y}}{x} = \frac{0{,}1013x + 445{,}59}{x} = 0{,}1013 + \frac{445{,}59}{x}$$

Der Anteil nähert sich also mit steigendem Einkommen immer mehr dem Wert 0,1013 = 10,13% von oben her an.

d) Die Wertepaare legen eher eine gekrümmte Regressionsfunktion nahe. Die Gerade ist eine Fehlspezifikation der Regressionsfunktion, obwohl sie eine gute Anpassung an die Wertepaare liefert. Ihre Prognosequalität ist schlecht; bei Verwendung einer gekrümmten Regressionsfunktion würde sich der Anteil der Wohnungsmieten am ausgabefähigen Einkommen einem Wert unter 0,1013 annähern.

Lösungen zu Kapitel 12

Lösung 12.1

a) Graphische Darstellung; Angaben zur Person Nr. 8 können z. B. weggelassen werden.
b) Beurteilung
c) $\sum_{i=1}^{9} x_{e_i} = 1550$ $\quad \sum_{i=1}^{9} y_{e_i} = 628$ $\quad \sum_{i=1}^{9} x_{e_i} y_{e_i} = 108309$

$\sum_{i=1}^{9} x_{e_i}^2 = 267460$ $\quad \sum_{i=1}^{9} y_{e_i}^2 = 44286$

$$r_{xy} = \frac{n\sum_{i=1}^{n} x_{e_i} y_{e_i} - \sum_{i=1}^{n} x_{e_i} \sum_{i=1}^{n} y_{e_i}}{\sqrt{[n\sum_{i=1}^{n} x_{e_i}^2 - (\sum_{i=1}^{n} x_{e_i})^2][n\sum_{i=1}^{n} y_{e_i}^2 - (\sum_{i=1}^{n} y_{e_i})^2]}} = 0{,}313$$

$$r_{xy}^2 = 0{,}313^2 = 0{,}098$$

r_{xy} = 0,313, d. h. es liegt ein schwacher positiver linearer Zusammenhang vor. $r_{xy}^2 = 0{,}313^2 = 0{,}098$ bedeutet, daß nur etwa 10% der Varianz der y-Werte durch den linearen Zusammenhang erklärt werden, die übrigen 90% sind unerkärte Streuung.

Lösung 12.2

a)
$\sum_{i=1}^{8} x_{e_i} = 42.845$ $\quad \sum_{i=1}^{8} y_{e_i} = 7.903$ $\quad \sum_{i=1}^{8} x_{e_i} y_{e_i} = 54.381.081$

$\sum_{i=1}^{8} x_{e_i}^2 = 348.520.000$ $\quad \sum_{i=1}^{8} y_{e_i}^2 = 9.082.067$

Daraus ergibt sich:

$r_{xy} = 0{,}9785$, d. h. zwischen dem ausgabefähigen Einkommen und den Ausgaben für Wohnungsmieten besteht ein starker, positiver, linearer Zusammenhang.

$r_{xy}^2 = 0{,}9575$, d. h. etwa 96% der Varianz der Ausgaben für Wohnungsmieten werden durch den linearen Zusammenhang erklärt.

b)
Augenschein und $r_{xy}^2 = 0{,}9575$ zeigen, daß die Gerade eine ausgezeichnete **Anpassungsgüte** aufweist. Der Augenschein zeigt jedoch auch, daß die Wertepaare eher eine gekrümmte Regressionsfunktion nahelegen und Prognosen mit der Geraden in die Irre führen werden. Die Gerade hat daher gleichzeitig eine geringe **Prognosequalität**.

Lösung 12.3

Die erforderlichen Teilsummen zur Berechnung des Korrelationskoeffizienten finden sich in Lösung 11.1.c. Man erhält dann:

$r_{xy} = 0{,}8767$, d. h. zwischen der Kinderzahl und den monatlichen Ausgaben eines Haushalts besteht ein starker, positiver, linearer Zusammenhang.
$r_{xy}^2 = 0{,}7686$, d. h. etwa 77% der Varianz der Konsumausgaben eines Haushalts werden durch den Zusammenhang mit der Kinderzahl erklärt.

Lösungen zu Kapitel 13

Lösung 13.1

a)

$$V_{2000,2003} = \frac{\sum_{i=1}^{3} p_i^{2003} q_i^{2003}}{\sum_{i=1}^{3} p_i^{2000} q_i^{2000}} = \frac{162}{200} = 0{,}81$$

b)

$$P_{2000,2003}^{La} = \sum_{i=1}^{3} \frac{p_i^{2003}}{p_i^{2000}} \cdot \frac{p_i^{2000} q_i^{2000}}{\sum_{i=1}^{3} p_j^{2000} q_j^{2000}} = 1{,}82$$

$$P_{2000,2003}^{Pa} = \frac{1}{\sum_{i=1}^{3} \frac{1}{\frac{p_i^{2003}}{p_i^{2000}}} \cdot \frac{p_i^{2003} q_i^{2003}}{\sum_{j=1}^{3} p_j^{2003} q_j^{2003}}} = 1{,}62$$

c)

$$Q_{2000,2003}^{La} = \frac{V_{2000,2003}}{P_{2000,2003}^{Pa}} = 0{,}5$$

$$Q_{2000,2003}^{Pa} = \frac{V_{2000,2003}}{P_{2000,2003}^{La}} = 0{,}445$$

Lösung 13.2

a)

$$V_{2000,2003} = \frac{\sum_{i=1}^{3} p_i^{2003} q_i^{2003}}{\sum_{i=1}^{3} p_i^{2000} q_i^{2000}} = \frac{32420}{46640} = 0{,}695$$

b)

$$P_{2000,2003}^{La} = \frac{\sum_{i=1}^{3} p_i^{2003} q_i^{2000}}{\sum_{i=1}^{3} p_i^{2000} q_i^{2000}} = \frac{49980}{46640} = 1{,}072$$

$$P_{2000,2003}^{Pa} = \frac{\sum_{i=1}^{3} p_i^{2003} q_i^{2003}}{\sum_{i=1}^{3} p_i^{2000} q_i^{2003}} = \frac{32420}{30260} = 1{,}071$$

c)

$$Q_{2000,2003}^{La} = \frac{V_{2000,2003}}{P_{2000,2003}^{Pa}} = 0{,}649$$

$$Q_{2000,2003}^{Pa} = \frac{V_{2000,2003}}{P_{2000,2003}^{La}} = 0{,}649$$

Lösung 13.3

a)

x = 19,87 (Promille)

b)

y = (16,86·103,5 + 19,87·117,9)/36,73 = 111,29

c)

Vergleich

d)

$$Q_{2000,2003} = \frac{1{,}12}{1{,}1129} = 1{,}006$$

Das Ergebnis ist ein Mengenindex nach Paasche, da der Preisindex vom Typ Laspeyres ist.

Lösung 13.4

a)

Jahr	Verbraucherprerisindex	
	1995 = 100	2000 = 100
1995	100,0	**93,5**
1996	101,4	**94,9**
1997	103,3	**96,6**
1998	104,3	**97,6**
1999	104,9	**98,1**
2000	106,9	100,0
2001	**109,0**	102,0
2002	**110,5**	103,4
2003	**111,7**	104,5

Die fettgedruckten Zahlen stellen die Verlängerung des Index zur Basis 1995 bzw. die Rückrechnung des Index zur Basis 2000 dar.

b)
Problem: der Laspeyres-Index ist nicht verkettbar (vgl. Text im Skript)

Lösung 13.5.

a)

(1) Jahr	(2) Arbeitnehmerentgelt (Mrd. €)	(3) Verbraucherpreisindex 2000 = 100	(4) Arbeitnehmerentgelt Preisbereinigt Sp. (2)/ Sp. (3)
1995	996,20	93,9	1060,92
1996	1005,30	95,3	1054,88
1997	1009,22	97,1	1039,36
1998	1030,56	98,0	1051,59
1999	1057,78	98,6	1072,80
2000	1099,09	100,0	1099,09
2001	1121,28	102,0	1099,29
2002	1130,46	103,4	1093,29
2003	1132,37	104,5	1083,61

Die preisbereinigte Reihe gibt das reale Arbeitnehmerentgelt an, also das Einkommen nach Ausschaltung des Preiseinflusses. Sie kann als Mengenreihe interpretiert werden. Die Entwicklung des Realeinkommens zeigt, wie sich die Gütermenge entwickelt hat, die man für das Einkommen kaufen kann. Es zeigt sich, daß das reale Arbeitnehmerentgelt seit 1995 praktisch unverändert geblieben ist; im Gegensatz dazu ist das nominale Arbeitnehmerentgelt leicht angestiegen, genauer um $\sqrt[8]{\frac{1132{,}37}{996{,}20}} - 1 = 1{,}6\%$ pro Jahr.

b)
Nachteil: Der Preisindex ist ein Index vom Typ Laspeyres. Zur Preisbereinigung müßte ein Preisindex nach Paasche verwendet werden, denn:
Mengenentwicklung = Wertentwicklung/Preisindex Laspeyres (vgl. Text im Skript 13.5f).
Vorteil: Zur Preisbereinigung (Deflationierung) des Arbeitnehmerentgelts ist ein Index erforderlich, der die Preisentwicklung der Güter angibt, für die das Arbeitnehmerentgelt verwendet wird. Es wird im wesentlichen für Güter der Lebenshaltung ausgegeben, bis auf die Ersparnis, die der Zukunftsvorsorge dient. Insgesamt ist der Preisindex der Lebenshaltung daher zur Deflationierung geeignet, der Preisindex für das Bruttoinlandsprodukt wäre auch geeignet.

Lösung 13.6

a1)
$KKP_{N,Lo}^{La} = 0{,}86$ [£/€]
3.000 [€] · 0,86 [£/€] = 2.580[£]

a2)
3.000 [€] · 0,62 [£/€] = 1.860[£]

a3) 1.860/2.580 = 0,72 = 72%
Je umgetauschte n € erhält er Waren im Wert von 72 cts.

b)
$$KKP_{N,Lo}^{La} = \sum_{i=1}^{4} \frac{p_i^{Lo}}{p_i^N} \cdot g_i^N = 0{,}785 \text{ [£/€]}$$
$$KKP_{N,Lo}^{Pa} = \frac{1}{\sum_{i=1}^{4} \frac{1}{\frac{p_i^{Lo}}{p_i^N}} \cdot g_i^{Lo}} = 0{,}76 \text{ [£/€]}$$

Die Kaufkraftparität nach Paasche, d. h. in diesem Beispiel bei Verwendung des britischen Warenkorbs, ist niedriger als die nach Laspeyres (bei Verwendung des deutschen Warenkorbs). Die Ausgaben für die Lebenshaltung eines Deutschen in England sind also niedriger, wenn er sich dem Verbrauchsschema in England anpasst. Fraglich ist dabei, ob er durch das englische Verbrauchsschema das gleiche Bedürfnisbefriedigungsniveau erreicht wie durch das deutsche.