

Statistik-Klausur vom 14.7.2003

Bearbeitungszeit: 120 Minuten

Aufgabe 1

Bevölkerung der Bundesrepublik Deutschland 1950 und 2000
nach Altersgruppen

| Personen im Alter von ... bis unter ... Jahren | 1950 ¹⁾ | | 2000 ¹⁾ | |
|--|--------------------|------|--------------------|------|
| | Anzahl | % | Anzahl | % |
| Unter 14 | 14 870 911 | 21,4 | 11 841 487 | 14,4 |
| 14 - 21 | 7 233 151 | 10,4 | 6 511 593 | 7,9 |
| 21 - 40 | 17 287 640 | 24,9 | 22 536 777 | 27,4 |
| 40 - 65 | 23 205 210 | 33,5 | 27 675 669 | 33,6 |
| 65 und mehr | 6 749 385 | 9,7 | 13 694 014 | 16,6 |
| Insgesamt | 69 346 297 | 100 | 82 259 540 | 100 |

1) Stand: jeweils Jahresende

Quelle: Statistisches Bundesamt: Statistisches Jahrbuch für die BRD
2002, S. 58, Tab. 3.9

- Wie hat sich das Altersniveau in Deutschland von 1950 bis 2000 verändert?
- Wie haben sich die Altersunterschiede in Deutschland von 1950 bis 2000 verändert?

Aufgabe 2

Arbeitslose Akademiker in der BRD

| Jahr | Anzahl (in 1 000) | |
|------|----------------------|--|
| 2000 | 176 | |
| 2001 | 180 | |
| 2002 | 224 | |
| | | |

Quelle: Institut der deutschen Wirtschaft Köln

- Berechnen Sie für diesen Datensatz die lineare Regressionsgerade.
- Wie gut lassen sich die Daten durch eine Gerade beschreiben? Beantworten Sie die Frage durch Berechnung und Interpretation einer geeigneten statistischen Maßzahl.

- c) Geben Sie einen Prognosewert für die Anzahl der arbeitslosen Akademiker im Jahr 2003 an.

Aufgabe 3

Preisindex für die Lebenshaltung aller privaten Haushalte

| Jahr | Preisindex (1995 = 100) | Preisindex (2000 = 100) |
|------|-------------------------|-------------------------|
| 1991 | 87,2 | |
| 1995 | 100 | |
| 2000 | 106,9 | 100 |
| 2001 | | 102,0 |

- a) Um wie viel Prozent sind die Preise für die Lebenshaltung aller privaten Haushalte in Deutschland im Zeitraum von 1991 bis 2000 im Durchschnitt pro Jahr gestiegen?
- b) Um wie viel Prozent sind die Preise für die Lebenshaltung aller privaten Haushalte in Deutschland von 1995 bis 2001 gestiegen?

Aufgabe 4

Ein Versicherungsunternehmen bietet zwei Typen von Versicherungen an: Typ A und Typ B. Eine Analyse des Kundenstamms ergab, dass

- 35 % aller Kunden eine Versicherung vom Typ A abgeschlossen haben.
 - 60% der Kunden, die keine Versicherung vom Typ A abgeschlossen haben, eine Versicherung vom Typ B abgeschlossen haben.
 - jeder Kunde, der eine Versicherung vom Typ A hat, auch eine Versicherung vom Typ B abgeschlossen hat.
- a) Mit welcher Wahrscheinlichkeit hat ein zufällig ausgewählter Kunde eine Versicherung vom Typ B abgeschlossen?
- b) Mit welcher Wahrscheinlichkeit hat ein Kunde, der eine Versicherung vom Typ B hat, eine Versicherung vom Typ A abgeschlossen?

Aufgabe 5

Eine Unternehmung möchte eine Investition tätigen, die mit Risiken behaftet ist. Sie lässt das Investitionsvorhaben von einem Gutachterbüro prüfen. Dieses stellt Folgendes fest:

- In jeder der ersten zehn Perioden liegt das Risiko für einen Verlust bei 29%.
 - Die Verluste der einzelnen Perioden sind in den ersten zehn Perioden stochastisch unabhängig.
- a) Wie wahrscheinlich ist es, dass es in der ersten Periode zu einem Verlust kommt?

- b) Wie wahrscheinlich ist es, dass in der ersten Periode ein Verlust, in der zweiten Periode kein Verlust und in der dritten Periode wieder ein Verlust erwirtschaftet werden?
- c) Wie wahrscheinlich ist es, dass es in genau drei der ersten zehn Perioden zu einem Verlust kommt?
- d) Wie wahrscheinlich ist es, dass es in höchstens drei der ersten zehn Perioden zu einem Verlust kommt?

Aufgabe 6

Bei einem Versandhandel entstehen 6 Euro Zusatzkosten, falls der Kunde seine Ware retourniert und mit der Retoure eine Neubestellung aufgibt. Die Zusatzkosten betragen 4 Euro, falls der Kunde lediglich seine Ware retourniert und keine Neubestellung aufgibt. Aus Erfahrung weiß man, dass

- 60% aller Bestellungen nicht retourniert werden
- 30% aller Bestellungen zwar retourniert werden, jedoch mit der Retoure keine Neubestellung aufgegeben wird
- 10% aller Bestellungen retourniert werden und gleichzeitig eine Neubestellung aufgegeben wird

- a) Wie hoch ist die erwartete Höhe an Zusatzkosten pro Bestellung?
- b) Berechnen Sie die Varianz der Variablen „Zusatzkosten (in €) pro Bestellung“.
- c) Die Unternehmensleitung interessiert sich für die Wahrscheinlichkeit, dass bei 100 Bestellungen die Zusatzkosten insgesamt höchstens 200 Euro betragen. Drücken Sie dieses Ereignis entweder über die Variable „Summe der Zusatzkosten (in €) bei 100 Bestellungen“ oder über die Variable „Durchschnittliche Zusatzkosten (in €) bei 100 Bestellungen“ aus und berechnen Sie näherungsweise die Wahrscheinlichkeit für dieses Ereignis.

Lösungen

Lösung von Aufgabe 1

X = Alter (1950)

Y = Alter (2000)

| Klasse | 1950 | | 2000 | |
|-------------|---------|-------|---------|-------|
| | n_j/n | F_X | n_j/n | F_Y |
| unter 14 | 0,214 | 0,214 | 0,144 | 0,144 |
| 14 - 21 | 0,104 | 0,318 | 0,079 | 0,223 |
| 21 - 40 | 0,249 | 0,567 | 0,274 | 0,497 |
| 40 - 65 | 0,335 | 0,902 | 0,336 | 0,833 |
| 65 und mehr | 0,097 | 0,999 | 0,166 | 0,999 |

a) 50%-Punkt:

$$x_{0,50} \approx 21 + \frac{0,50 - 0,318}{0,249} \cdot 19 = 34,8876 \approx 35$$

$$y_{0,50} \approx 40 + \frac{0,50 - 0,497}{0,336} \cdot 25 = 40,2232 \approx 40$$

d.h. 1950 war etwa die Hälfte der Bevölkerung mindestens 35 Jahre alt, während 2000 etwa die Hälfte der Bevölkerung mindestens 40 Jahre alt war.

b) $x_{0,25} \approx 14 + \frac{0,25 - 0,214}{0,104} \cdot 7 = 16,4231$

$$x_{0,75} \approx 40 + \frac{0,75 - 0,567}{0,335} \cdot 25 = 53,6567$$

$$y_{0,25} \approx 21 + \frac{0,25 - 0,223}{0,274} \cdot 19 = 22,8723$$

$$y_{0,75} \approx 40 + \frac{0,75 - 0,497}{0,336} \cdot 25 = 58,8244$$

relativer Quartilsabstand:

$$\frac{x_{0,75} - x_{0,25}}{x_{0,50}} = \frac{53,6567 - 16,4231}{34,8876} = 1,0672$$

$$\frac{y_{0,75} - y_{0,25}}{y_{0,50}} = \frac{58,8244 - 22,8723}{40,2232} = 0,8938$$

d.h. gemessen mit dem relativen Quartilsabstand waren 1950 die Altersunterschiede stärker.

Lösung von Aufgabe 2

| | x_i | y_i | $x_i \cdot y_i$ | x_i^2 | y_i^2 |
|----------|-------|-------|-----------------|---------|---------|
| | 0 | 176 | 0 | 0 | 30 976 |
| | 1 | 180 | 180 | 1 | 32 400 |
| | 2 | 224 | 448 | 4 | 50 176 |
| Σ | 3 | 580 | 628 | 5 | 113 552 |

a) Die unabhängig Variable ist die Zeit x

Regressionsgerade : $f(x) = a_1 + b_1 \cdot x$

$$b_1 = \frac{3 \cdot 628 - 3 \cdot 580}{3 \cdot 5 - 3^2} = \frac{144}{6} = 24$$

$$a_1 = \frac{580 - 24 \cdot 3}{3} = \frac{508}{3} = 169,\bar{3}$$

d.h. die Regressionsgerade lautet: $f(x) = 169,\bar{3} + 24 \cdot x$

b) $b_2 = \frac{144}{3 \cdot 113 552 - 580^2} = \frac{144}{4 256} = 0,0338$

Bestimmtheitsmaß $B = b_1 \cdot b_2 = 24 \cdot 0,0338 = 0,8120$

Korrelationskoeffizient $r_{xy} = \sqrt{B} = \sqrt{0,8120} = 0,9011$

d.h. es liegt starke Korrelation vor.

- c) $f(3) = 169,3 + 24 \cdot 3 = 241,3333$
d.h. gemäß der Methode der kleinsten Quadrate ist im Jahr 2003 mit 241 000 arbeitslosen Akademikern zu rechnen.

Lösung von Aufgabe 3

a) $\sqrt[9]{\frac{106,9}{87,2}} = \sqrt[9]{1,2259} = 1,02289$

d.h. die Preise für die Lebenshaltung aller privaten Haushalte sind in Deutschland von 1991 bis 2000 um durchschnittlich 2,3% pro Jahr gestiegen.

b)

| Zeitraum | Steigerung |
|-------------|------------|
| 1995 - 2000 | + 6,9 % |
| 2000 - 2001 | + 2,0 % |

$$1,069 \cdot 1,02 = 1,0904$$

d.h. die Preise für die Lebenshaltung aller privaten Haushalte in Deutschland sind im Zeitraum von 1995 bis 2001 um rund 9% gestiegen.

Lösung von Aufgabe 4

A = Kunde schließt eine Versicherung vom Typ A ab

B = Kunde schließt eine Versicherung vom Typ B ab

Gegeben sind folgende Information:

- $P(A) = 0,35$
- $P(B | \bar{A}) = 0,60$
- $A \subset B \Rightarrow A \cap B = A$ und $A \cap \bar{B} = \emptyset$

Aus der 1. Information folgt, dass gilt:

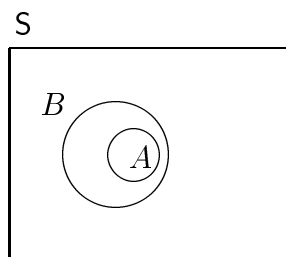
$$P(\bar{A}) = 1 - 0,35 = 0,65$$

Mit der 2. Information haben wir:

$$P(B | \bar{A}) = \frac{P(\bar{A} \cap B)}{P(\bar{A})} \Rightarrow P(\bar{A} \cap B) = P(B | \bar{A}) \cdot P(\bar{A}) = 0,60 \cdot 0,65 = 0,39$$

Aus der 3. Information folgt, dass gilt:

$$P(A \cap \bar{B}) = 0$$



Jetzt können wir die Arbeitstabelle aufstellen:

| | | | |
|-----------|------|-----------|------|
| | A | \bar{A} | |
| B | 0,35 | 0,39 | 0,74 |
| \bar{B} | 0 | 0,26 | 0,26 |
| | 0,35 | 0,65 | 1 |

a) $P(B) = 0,74$

d.h. 74 % aller Kunden haben eine Versicherung vom Typ B abgeschlossen.

oder $P(B) = P(A \cap B) + P(\bar{A} \cap B) = 0,35 + 0,39 = 0,74$

b) $P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{0,35}{0,74} = 0,4730$

d.h. unter denjenigen Kunden, die eine Versicherung vom Typ B abgeschlossen haben, beträgt der Anteil der Kunden, die auch eine Versicherung vom Typ A abgeschlossen haben, genau 47,3%.

Lösung von Aufgabe 5

$$X_i = \begin{cases} 0 & ; \text{ falls in der } i\text{-ten Periode kein Verlust gemacht wird} \\ 1 & ; \text{ falls in der } i\text{-ten Periode Verlust gemacht wird} \end{cases} ; i = 1, \dots, 10$$

$$P(X_i = 1) = 0,29$$

X_1, \dots, X_{10} stochastisch unabhängig.

a) $P(X_1 = 1) = 0,29$

d.h. die Wahrscheinlichkeit beträgt 0,29.

$$\begin{aligned} \text{b) } P(X_1 = 1 \cap X_2 = 0 \cap X_3 = 1) &= P(X_1 = 1) \cdot P(X_2 = 0) \cdot P(X_3 = 1) \\ &= 0,29 \cdot 0,71 \cdot 0,29 \\ &= 0,0597 \end{aligned}$$

d.h. die Wahrscheinlichkeit beträgt 0,0597.

c) $Y =$ Anzahl der Perioden, in denen ein Verlust gemacht wird.

$$Y \sim \mathbf{B}(n = 10; p = 0,29)$$

$$P(Y = 3) = \binom{10}{3} \cdot 0,29^3 \cdot 0,71^7 = 0,2662$$

d.h. die Wahrscheinlichkeit beträgt 0,2662.

$$\begin{aligned} \text{d) } P(Y \leq 3) &= P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3) \\ &= \binom{10}{0} \cdot 0,29^0 \cdot 0,71^{10} + \binom{10}{1} \cdot 0,29^1 \cdot 0,71^9 \\ &\quad + \binom{10}{2} \cdot 0,29^2 \cdot 0,71^8 + 0,2662 \\ &= 0,0326 + 0,1330 + 0,2444 + 0,2662 \\ &= 0,6761 \end{aligned}$$

d.h. die Wahrscheinlichkeit beträgt 0,6761.

Lösung von Aufgabe 6

$X =$ Zusatzkosten (in Euro) pro Bestellung

| | | | |
|------------|-----|-----|-----|
| x | 0 | 4 | 6 |
| $P(X = x)$ | 0,6 | 0,3 | 0,1 |

a) $E[X] = 0 \cdot 0,6 + 4 \cdot 0,3 + 6 \cdot 0,1 = 1,8$

d.h. pro Bestellung ist mit Zusatzkosten in Höhe von 1,80 Euro zu rechnen.

b) $Var[X] = (0 - 1,8)^2 \cdot 0,6 + (4 - 1,8)^2 \cdot 0,3 + (6 - 1,8)^2 \cdot 0,1 = 5,16$
d.h. die Varianz beträgt 5,16.

c) Zentraler Grenzwertsatz (Faustregel $n \geq 30$ erfüllt)

$$P(X_1 + \dots + X_{100} \leq 200) \approx F_U \left(\frac{200 - 100 \cdot 1,8}{\sqrt{100 \cdot 5,16}} \right) = F_U \left(\frac{200 - 180}{\sqrt{516}} \right) =$$

$$F_U(0,8805) = 0,811$$

d.h. die Wahrscheinlichkeit beträgt annähernd 0,811.

oder

$$P(\bar{X} \leq 2) \approx F_U \left(\frac{2 - 1,8}{\sqrt{5,16/100}} \right) = F_U \left(\frac{2 - 1,8}{\sqrt{0,0516}} \right) = F_U(0,8805) = 0,811$$