Statistik-Klausur vom 11. Februar 2005

Bearbeitungszeit: 90 Minuten

Aufgabe 1

Ein Konzern erstellt einen Überblick über die Umsätze von drei Tochterunternehmen in der vergangenen Periode. Dazu werden die Umsätze pro Auftrag erfasst und in Klassen eingeteilt. Es ergibt sich folgendes Bild:

Anzahl der Aufträge gegliedert nach Tochterunternehmen und Umsatzhöhe

Umsätze in GE		Anzahl der Aufträge in Tochterunternehmen			Summe
über …	bis maximal …	A	B	C	
0	1 000	23	10	18	51
1 000	2 000	34	15	25	74
2 000	4 000	46	14	38	98
4 000	6 000	37	18	57	112
6 000	10 000	13	17	43	73
10 000	15 000	10	9	21	40
über 15 000		5	3	1	9
Summe		168	86	203	457

a) Die Tabelle basiert auf der statistischen Variablen X=„Umsatzhöhe eines Auftrags“. Ist diese Variable stetig oder diskret? Welche Skalierung hat diese statistische Variable?

b) Wie viel Prozent der Aufträge brachten dem Tochterunternehmen A einen Umsatz von mindestens 7 500 GE?

c) Wie viel Prozent der Aufträge brachten dem Tochterunternehmen B einen Umsatz von höchstens 3 000 GE?

d) Welches der drei Tochterunternehmen hat das höchste Umsatzniveau pro Auftrag? Beantworten Sie diese Frage durch Berechnung einer geeigneten Maßzahl.

Aufgabe 2

In einem Land wurde der Preisindex für die Lebenshaltung erhoben. Im Jahr 2000 wurde der Warenkorb neu den Verhältnissen angepasst. Es ergaben sich die folgenden Indexwerte:

Jahr	Indexwerte Basisjahr 1995	Basisjahr 2000
1998	102,3	
1999	104,1	
2000	106,2	100
2001		101,8
2002		103,9
2003		105,8
2004		107,5

a) Rechnen Sie die neue Indexreihe mit dem Basisjahr 2000 zurück bis 1998. Sie erhalten so eine durchgehende Reihe von 1998 bis 2004.

b) Um wie viel Prozent ist im Zeitraum von 1998 bis 2004 der Preisindex insgesamt gestiegen?

c) Geben Sie mit Hilfe der Methode der kleinsten Quadrate (lineare Regression) einen Prognosewert für 2005 an. Verwenden Sie hierzu die durchgehende Indexreihe aus a).

d) Wie verlässlich ist Ihre Prognose unter c)? Beantworten Sie diese Frage durch Berechnung und Interpretation einer geeigneten statistischen Maßzahl.

Aufgabe 3

Ein Versicherungsvertreter weiß aufgrund seiner langjährigen Erfahrung, dass 41% der neu abgeschlossenen Versicherungsverträge noch während der Rücktrittsfrist von den Kunden gekündigt werden.

1. An einem Tag hat der Versicherungsvertreter fünf Versicherungsverträge abgeschlossen.
   a) Wie groß ist die Wahrscheinlichkeit, dass alle Versicherungsverträge des Tages erfüllt, d.h. nicht gekündigt werden?
   b) Wie groß ist die Wahrscheinlichkeit, dass mindestens zwei der Versicherungsverträge erfüllt werden?
   c) Wie groß ist die Wahrscheinlichkeit, dass höchstens zwei der Versicherungsverträge erfüllt werden?

2. In einem Monat hat der Versicherungsvertreter 85 Versicherungsverträge abgeschlossen. Er erhält eine Sonderprämie, wenn er im Monat 50 Versicherungsverträge abschließt, die nicht während der Rücktrittsfrist gekündigt werden.
   a) Kann der Versicherungsvertreter damit rechnen, in diesem Monat die Sonderprämie zu bekommen?
   b) Wie groß ist die Wahrscheinlichkeit, dass der Versicherungsvertreter in diesem Monat die Sonderprämie erhält?

Aufgabe 4
In der Fakultät einer Hochschule werden die drei Studiengänge BWL, Wirtschaftsrecht und Banking&Finance angeboten. Die Studierenden der Fakultät verteilen sich wie folgt auf die drei Studiengänge:

- 11% studieren Wirtschaftsrecht
- 17% studieren Banking&Finance
- der Rest studiert BWL

Der Frauenanteil in den drei Studiengängen beträgt:

- 48% in Wirtschaftsrecht
- 60% in Banking&Finance
- 53% in BWL

Wie viel Prozent der weiblichen Studierenden der Fakultät studieren

a) Wirtschaftsrecht?

b) Banking&Finance?

c) BWL?

Aufgabe 5
Die Fakultät einer Hochschule möchte den Zeitaufwand (gemessen in Stunden pro Woche) für Nebenbeschäftigungen ihrer 2 221 Studierenden ermitteln.

a) In einer früheren Umfrage zum gleichen Thema betrug der durchschnittliche Zeitaufwand pro Woche eines Studierenden zwölf Stunden und die Standardabweichung in der Stichprobe betrug 1,2 Stunden.
Wie viele Studierende sind mindestens zu befragen, damit das gesuchte 0,95-Konfidenzintervall höchstens die Breite von 0,4 hat, d.h. nur um $\pm 0{,}2$ Stunden vom wahren mittleren Zeitaufwand aller 2 221 Studierenden abweicht?

b) Bei einer Umfrage von 150 Studierenden ergab sich ein durchschnittlicher Zeitaufwand von 12,5 Stunden pro Woche, und die Standardabweichung in der Stichprobe betrug zwei Stunden.
Berechnen Sie ein 0,95-Konfidenzintervall für den mittleren Zeitaufwand eines Studierenden.

c) Wieso kann es passieren, dass die Breite des unter b) berechneten Konfidenzintervalls größer ist als die unter a) geforderte Breite von 0,4 Stunden?

Lösungen

Lösung zu Aufgabe 1

Zunächst werden die kumulierten absoluten Häufigkeiten berechnet, um die Einfallsklassen schneller zu erkennen.

Umsätze in GE		A		B		C	
0	1 000	23	23	10	10	18	18
1 000	2 000	34	57	15	25	25	43
2 000	4 000	46	103	14	39	38	81
4 000	6 000	37	140	18	57	57	138
6 000	10 000	13	153	17	74	43	181
10 000	15 000	10	163	9	83	21	202
über 15 000		5	168	3	86	1	203

a) Die Skalierung ist metrisch.
Da der Umsatz nur auf einen Cent (oder eine GE) genau erfasst werden kann, ist der Umsatz eine diskrete Variable. Da die Abstufungen, in denen der Umsatz erfasst wird, sehr fein sind, kann der Umsatz auch als stetige Variable bezeichnet werden.

b) $F_A(7\,500) = \dfrac{140}{168} + (7\,500 - 6\,000) \cdot \dfrac{13/168}{4\,000} = 0{,}833 + 0{,}029 = 0{,}862$
$1 - F(7\,500) = 0{,}138$
d.h. bei Tochterunternehmen A hatten etwa 14% aller Aufträge einen Mindestumsatz von 7 500 GE.

c) Da 3 000 die Mitte der dritten Klasse ist, können wir auch wie folgt rechnen:
$F_B(3\,000) = \dfrac{25}{86} + \dfrac{1}{2} \cdot \dfrac{14}{86} = \dfrac{32}{86} = 0{,}372$
d.h. bei Tochterunternehmen B hatten etwa 37% aller Aufträge einen Höchstumsatz von 3 000 GE.

d) $x^A_{0,50} \approx 2\,000 + \dfrac{0{,}5 - 57/168}{46/168} \cdot 2\,000 = 3\,173{,}91$

$x^B_{0,50} \approx 4\,000 + \dfrac{0{,}5 - 39/86}{18/86} \cdot 2\,000 = 4\,444{,}44$

$x^C_{0,50} \approx 4\,000 + \dfrac{0{,}5 - 81/203}{57/203} \cdot 2\,000 = 4\,719{,}30$

d.h. im Tochterunternehmen C ist das Umsatzniveau pro Auftrag am höchsten.

Lösung zu Aufgabe 2

a) Verkettung:

Jahr	Basisjahr 2000
1998	$\frac{100}{106{,}2} \cdot 102{,}3 = 96{,}3$
1999	$\frac{100}{106{,}2} \cdot 104{,}1 = 98{,}0$
2000	100
2001	101,8
2002	103,9
2003	105,8
2004	107,5

b) $\dfrac{\text{Wert}_{2004}}{\text{Wert}_{1998}} = \dfrac{107{,}5}{96{,}3} = 1{,}116303$

d.h. im Zeitraum von 1998 bis 2004 ist der Preisindex um insgesamt 11,6% gestiegen.

c) Lineare Regression:

1. Lösungsweg

i	x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
	1	96,3	96,3	1	9 273,69
	2	98,0	196,0	4	9 604,00
	3	100	300,0	9	10 000,00
	4	101,8	407,2	16	10 363,24
	5	103,9	519,5	25	10 795,21
	6	105,8	634,8	36	11 193,64
	7	107,5	752,5	49	11 556,25
$\sum$	28	713,3	2 906,3	140	72 786,03

Gesucht: $f(8) = a_1 + b_1 \cdot 8 =?$

$$b_1 = \frac{7 \cdot 2\,906{,}3 - 28 \cdot 713{,}3}{7 \cdot 140 - 28^2} = \frac{371{,}7}{196} = 1{,}896429$$

$$a_1 = \frac{713{,}3 - 1{,}896429 \cdot 28}{7} = 94{,}314286$$

$f(8) = 94{,}314286 + 1{,}896429 \cdot 8 = 109{,}4857$

d.h. der vorhergesagte Preisindex für das Jahr 2005 beträgt 109,5.

2. Lösungsweg

Weniger ist zu rechnen, wenn die sieben Zeitpunkte mit 0 bis 6 durchnumeriert werden:

i	x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
	0	96,3	0	0	9 273,69
	1	98,0	98,0	1	9 604,00
	2	100	200,0	4	10 000,00
	3	101,8	305,4	9	10 363,24
	4	103,9	415,6	16	10 795,21
	5	105,8	529,0	25	11 193,64
	6	107,5	645,0	36	11 556,25
$\sum$	21	713,3	2 193	91	72 786,03

Gesucht: $f(7) = a_1 + b_1 \cdot 7 =?$

$b_1 = 1{,}896429$
$a_1 = 96{,}210714$
$f(7) = 96{,}210714 + 1{,}896429 \cdot 7 = 109{,}5$

d) $b_2 = \dfrac{371{,}7}{7 \cdot 72\,786{,}03 - 713{,}3^2} = \dfrac{371{,}7}{705{,}32} = 0{,}5269948$
Bestimmtheitsmaß $B = b_1 \cdot b_2 = 1{,}896429 \cdot 0{,}5269948 = 0{,}999408$
Korrelationskoeffizient $r = +\sqrt{0{,}999408} = 0{,}999704$
d.h. es liegt ein starker positiver linearer Zusammenhang vor.
d.h. auf die Berechnung unter c) ist jedoch kein Verlass, da es sich um eine Extrapolation handelt.

Lösung zu Aufgabe 3
1. Lösungsweg
X = Anzahl der Versicherungsverträge, die erfüllt werden
$X \sim \mathsf{B}(n; p = 0{,}59)$

1. $X \sim \mathsf{B}(n = 5; p = 0{,}59)$
   a) $P(X = 5) = \binom{5}{5} \cdot 0{,}59^5 \cdot 0{,}41^0 = 0{,}0715$
   d.h. die Wahrscheinlichkeit beträgt 0,0715.
   b) $P(X \geq 2) = 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)] = 1 - 0{,}0116 - 0{,}0834 = 0{,}9050$
   $P(X = 0) = \binom{5}{0} \cdot 0{,}59^0 \cdot 0{,}41^5 = 0{,}0116$
   $P(X = 1) = \binom{5}{1} \cdot 0{,}59^1 \cdot 0{,}41^4 = 0{,}0834$
   d.h. die Wahrscheinlichkeit beträgt 0,9050.
   c) $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0{,}0116 + 0{,}0834 + 0{,}2399 = 0{,}3349$
   $P(X = 2) = \binom{5}{2} \cdot 0{,}59^2 \cdot 0{,}41^3 = 0{,}2399$
   d.h. die Wahrscheinlichkeit beträgt 0,3349.

2. $X \sim \mathsf{B}(n = 85; p = 0{,}59)$

   a) $E[X] = 85 \cdot 0{,}59 = 50{,}15$
      d.h. er kann erwarten, die Prämie zu bekommen.

   b) ZGWS
      Faustregel $np = 50{,}15 \geq 10$ und $n(1-p) = 34{,}85 \geq 10$ ist erfüllt
      $P(X \geq 50) = 1 - P(X < 50) = 1 - P(X \leq 49) \approx 1 - F_U\left(\frac{49+0{,}5-50{,}15}{\sqrt{50{,}15 \cdot 0{,}41}}\right) =$
      $1 - F_U(-0{,}1434) = 1 - 0{,}443 = 0{,}557.$
      d.h. die Wahrscheinlichkeit ist etwas höher als 50%.

2. Lösungsweg zu Aufgabe 3:
Y = Anzahl der Versicherungsverträge, die nicht erfüllt werden
$Y \sim \mathsf{B}(n; p = 0{,}41)$

1. $Y \sim \mathsf{B}(n = 5; p = 0{,}41)$

   a) $P(Y = 0) = 0{,}0715$

   b) $P(Y \leq 3) = 0{,}9050$

   c) $P(Y \geq 3) = 0{,}3349$

2. $Y \sim \mathsf{B}(n = 85; p = 0{,}41)$

   a) $E[Y] = n \cdot p = 85 \cdot 0{,}41 = 34{,}85$
      d.h. erwartungsgemäß werden 34,85 Verträge von 85 Verträgen nicht erfüllt.
      d.h. erwartungsgemäß werden $85 - 34{,}85 = 50{,}15$ Verträge erfüllt.

   b) Faustregel $np \geq 10$ und $n(1-p) \geq 10$ ok
      $$P(Y \leq 35) \approx F_U\left(\frac{35 + 0{,}5 - 34{,}85}{20{,}5615}\right) = F_U(0{,}1433) = 0{,}557$$

Lösung zu Aufgabe 4

$$\begin{array}{lcl@{\qquad\qquad}lcl}
0{,}48 & = & P(\text{Frau}|\text{W-Recht}) & P(\text{W-Recht}) & = & 0{,}11 \\
0{,}60 & = & P(\text{Frau}|\text{B \& FI}) & P(\text{B \&FI}) & = & 0{,}17 \\
0{,}53 & = & P(\text{Frau}|\text{BWL}) & P(\text{BWL}) & = & 0{,}72
\end{array}$$

$P(\text{Frau} \cap \text{W-Recht}) = 0{,}48 \cdot 0{,}11 = 0{,}0528$
$P(\text{Frau} \cap \text{B \& FI}) = 0{,}60 \cdot 0{,}17 = 0{,}1020$
$P(\text{Frau} \cap \text{BWL}) = 0{,}53 \cdot 0{,}72 = 0{,}3816$

Arbeitstabelle:

	W-Recht	B &FI	BWL	
F	0,0528	0,1020	0,3816	0,5364
M				
	0,11	0,17	0,72	1

a) $P(\text{W-Recht}|\text{F}) = \dfrac{0{,}0528}{0{,}5364} = 0{,}0984$
   d.h. der Anteil beträgt 10%.

b) $P(\text{B\&FI}|\text{F}) = \dfrac{0{,}1020}{0{,}5364} = 0{,}1902$
   d.h. der Anteil beträgt 19%.

c) $P(\text{BWL}|\text{F}) = \dfrac{0{,}3816}{0{,}5364} = 0{,}7114$
   d.h. der Anteil beträgt 71%.

Lösung zu Aufgabe 5

a) $n \geq \dfrac{1{,}96^2 \cdot 1{,}2^2}{0{,}2^2} = 138{,}2976$
   d.h. es sind mindestens 139 Studierende zu befragen.

b) Faustregel $n \geq 30$ ok
   0,95-KI für $E[X]$:
   $[12{,}5 \pm 1{,}96 \cdot \dfrac{2}{\sqrt{150}}] = [12{,}5 \pm 0{,}32] = [12{,}18\,;\ 12{,}82] \approx [12{,}2\,;\ 12{,}8]$

c) Bei der Berechnung des Mindest-Stichprobenumfangs unter a) wurde die Standardabweichung aus einer früheren Stichprobe verwendet. Sobald dieser alte Wert nicht mehr zutrifft, ist auch der berechnete Mindest-Stichprobenumfang hinfällig.