

Statistik-Klausur vom 11. Juli 2006

Bearbeitungszeit: 90 Minuten

Aufgabe 1

Im Rahmen einer Befragung von 100 Studierenden wurde die wöchentliche Nachbereitungszeit in Stunden für die Vorlesungen Volkswirtschaftslehre und Statistik erfasst. Die Ergebnisse der Befragung sind der folgenden Tabelle zu entnehmen:

Wöchentliche Nachbereitungszeit in Stunden	Volkswirtschaftslehre (Anzahl Studierende)	Statistik (Anzahl Studierende)
0 bis unter 0,5	20	18
0,5 bis unter 1,5	25	32
1,5 bis unter 2,5	25	20
2,5 bis unter 4	15	20
4 und länger	15	10

- Welche Skalierung weist die Variable $X =$ „wöchentliche Nachbereitungszeit in Stunden“ auf? Ist diese Variable diskret oder stetig?
- Welche Vorlesung haben die Studierenden intensiver nachbereitet? Begründen Sie Ihre Antwort durch Berechnung einer geeigneten statistischen Maßzahl.
- Bei welcher Vorlesung waren die Unterschiede in der Nachbereitungszeit größer? Begründen Sie Ihre Antwort durch Berechnung einer geeigneten statistischen Maßzahl.
- Wie hoch ist der Anteil der Studierenden, die die Vorlesung Statistik mit mindestens drei Stunden pro Woche nachbereiten?

Aufgabe 2

Das reale Bruttoinlandsprodukt in der BRD wird seit der Revision 2005 nur noch als Index auf Basis 2000=100 veröffentlicht. In der folgenden Tabelle und Grafik sind die Angaben für das reale Bruttoinlandsprodukt abgebildet:

Reales Bruttoinlandsprodukt (2000=100)	
Jahr	Werte
1999	94,10
2000	100,00
2001	101,09
2002	99,47
2003	100,51
2004	104,95
2005	107,71



- Um wie viel Prozent ist die deutsche Wirtschaft von 1999 bis 2005 gewachsen? Wie hoch war die jährliche durchschnittliche Wachstumsrate in Prozent?
- Prognostizieren Sie mit Hilfe der Methode der kleinsten Quadrate (Lineare Regression) den Wert des Index für das Jahr 2006. Um wie viele Prozent würde laut Ihrem Ergebnis das reale Bruttoinlandsprodukt im Vorjahresvergleich wachsen?
- Das Bestimmtheitsmaß für die Regression aus b) beträgt 0,8. Wie lässt sich dieser Parameter interpretieren?
- Für 2006 prognostizieren die sechs führenden Forschungsinstitute in ihrem Frühjahrgutachten ein Wachstum des realen Bruttoinlandsprodukts von 1,8 % im Vorjahresvergleich. Wie erklären Sie sich den Unterschied zu Ihrem Ergebnis aus b)?

Aufgabe 3

„Die Älteren lieben neue Autos“. Fast zwei Drittel der so genannten Generation 60 plus, genau 63 Prozent, besitzen einen Neuwagen. Auf die Frage „Haben Sie Ihr Auto neu oder gebraucht gekauft?“ antworteten im März 2006 bei einer Umfrage in Deutschland die Fahrzeughalter wie folgt:

Alter	Gebrauchtwagen	Vorführgewagen	Neuwagen	Fahrzeughalter
18 - 29	73%	14%	13%	10 %
30 - 39	44%	19%	37%	20 %
40 - 49	44 %	17 %	39%	26 %
50 - 59	36%	15%	49%	20 %
60 plus	23%	14%	63%	24 %

Quellen: Aral-Marktforschung, KBA

- Wie viel Prozent der Halter eines Neuwagens sind
 - mindestens 60 Jahre alt?
 - 18 bis 29 Jahre alt?

3. 30 bis 59 Jahre alt?

- b) Sind die Ereignisse „zufällig befragter Fahrzeughalter kaufte einen Neuwagen“ und „zufällig befragter Fahrzeughalter ist mindestens 60 Jahre alt“ stochastisch unabhängig?

Aufgabe 4

Im Rahmen einer Qualitätskontrolle werden bei einem Unternehmen monatlich 30 Fertigungsstücke einer Prüfung unterzogen. Erfahrungsgemäß liegt die Ausschussrate bei 5 %. Die Qualität der einzelnen Fertigungsstücke sei stochastisch unabhängig voneinander.

- a) Wie hoch ist die Wahrscheinlichkeit, dass im Monat August höchstens zwei Fertigungsstücke als Ausschuss aussortiert werden?
- b) Im Monat September wird die Anzahl der geprüften Fertigungsstücke im Rahmen einer Sonderaktion auf 240 erhöht. Wie hoch ist jetzt näherungsweise die Wahrscheinlichkeit, dass höchstens 16 Fertigungsstücke als Ausschuss aussortiert werden?

Aufgabe 5

Ein Marktforschungsinstitut soll Aufschluss darüber geben, wie viel Geld für Urlaubsreisen ausgegeben wird. Dazu möchte das Institut ein 0,95-Konfidenzintervall für die mittlere Höhe der Ausgaben für Urlaubsreisen berechnen. Bei einer Umfrage: „Wie viel Euro planen Sie für Ihre nächste Urlaubsreise auszugeben?“ erhielt das Institut 100 Antworten. Es ergaben sich folgende Daten x_1, \dots, x_{100} :

i	x_i	$(x_i - \bar{x})^2$
1	300	239 121
2	600	35 721
\vdots		
100	800	121
Σ	78 900	9 137 900

- a) Berechnen Sie aus den Daten das gesuchte Konfidenzintervall.
- b) Interpretieren Sie das berechnete Konfidenzintervall.

Lösung zu Aufgabe 1

- a) X ist metrisch skaliert und stetig.
- b) X = wöchentliche Nachbereitungszeit (in h) eines Studierenden für VWL
 Y = wöchentliche Nachbereitungszeit (in h) eines Studierenden für Statistik

$$x_{0,50} \approx 1,5 + \frac{0,5 - 0,45}{0,25} \cdot 1 = 1,7$$

$$y_{0,50} \approx 1,5$$

d.h. für VWL wird länger gearbeitet.

$$\begin{aligned}
\text{c) } x_{0,25} &\approx 0,5 + \frac{0,25 - 0,2}{0,25} \cdot 1 = 0,7 \\
x_{0,75} &\approx 2,5 + \frac{0,75 - 0,7}{0,15} \cdot 1,5 = 3 \\
\frac{x_{0,75} - x_{0,25}}{x_{0,5}} &= \frac{3 - 0,7}{1,7} = 1,3529 \\
y_{0,25} &\approx 0,5 + \frac{0,25 - 0,18}{0,32} \cdot 1 = 0,7188 \\
y_{0,75} &\approx 2,5 + \frac{0,75 - 0,7}{0,2} \cdot 1,5 = 2,875 \\
\frac{y_{0,75} - y_{0,25}}{y_{0,5}} &= \frac{2,875 - 0,7188}{1,5} = 1,4375
\end{aligned}$$

d.h. gemessen am relativen Quartilsabstand sind die Unterschiede in der Nachbereitungszeit für Statistik stärker.

$$\begin{aligned}
\text{d) } F_y(3) &\approx 0,7 + \frac{0,2}{1,5}(3 - 2,5) = 0,7667 \\
&\text{d.h. } 23 \% \text{ bereiten } 3 \text{ h oder länger nach.}
\end{aligned}$$

Lösung zu Aufgabe 2

Die Indexwerte in der Tabelle sind die Werte jeweils am Jahresende.

$$\text{a) } \sqrt[6]{\frac{107,71}{94,10}} = \sqrt[6]{1,1446} = 1,0228$$

d.h. im Zeitraum von 1999 bis 2005 ist das BIP um insgesamt 14,46 % und pro Jahr um durchschnittlich 2,28 % gestiegen.

$$\text{b) }$$

x_i	y_i	$x_i \cdot y_i$	x_i^2
1	94,10		
2	100,00		
3	101,09		
4	99,47		
5	100,51		
6	104,95		
7	107,71		
28	707,83	2 881,47	140

$$b_1 = \frac{7 \cdot 2 881,47 - 28 \cdot 707,83}{7 \cdot 140 - 28^2} = 1,791071$$

$$a_1 = \frac{707,83 - 1,791071 \cdot 28}{7} = 93,954286$$

$$a_1 + b_1 \cdot 8 = 108,2829$$

$$\frac{108,2829}{107,71} = 1,0053$$

d.h. der Prognosewert für 2006 ist 108,28, was einem Wachstum gegenüber dem Vorjahr von 0,53% entspricht.

c) $r = \sqrt{0,8} = 0,89$

d.h. starker positiver linearer Zusammenhang.

d) Es handelt sich bei der Prognose um eine Extrapolation. Ferner gehen die wirtschaftlich schwachen Jahre 2001, 2002, 2003 in die Prognose mit ein. Außerdem hat im letzten Jahr ein Regierungswechsel stattgefunden (von SPD zu CDU/SPD), bei dem erfahrungsgemäß die Wirtschaft wieder mehr investiert.

Lösung zu Aufgabe 3

N =Neuwagen

A_1 =18 bis 29 Jahre	$P(N A_1) = 0,13$	$P(A_1) = 0,10$
A_2 =30 bis 39 Jahre	$P(N A_2) = 0,37$	$P(A_2) = 0,20$
A_3 =40 bis 49 Jahre	$P(N A_3) = 0,39$	$P(A_3) = 0,26$
A_4 =50 bis 59 Jahre	$P(N A_4) = 0,49$	$P(A_4) = 0,20$
A_5 =60 Jahre oder älter	$P(N A_5) = 0,63$	$P(A_5) = 0,24$

Daraus ergeben sich folgende Anteile:

$$\begin{aligned}
 P(N \cap A_1) &= 0,13 \cdot 0,10 = 0,0130 \\
 P(N \cap A_2) &= 0,37 \cdot 0,20 = 0,0740 \\
 P(N \cap A_3) &= 0,39 \cdot 0,26 = 0,1014 \\
 P(N \cap A_4) &= 0,49 \cdot 0,20 = 0,0980 \\
 P(N \cap A_5) &= 0,63 \cdot 0,24 = 0,1512
 \end{aligned}$$

	A_1	A_2	A_3	A_4	A_5	Σ
N	0,0130	0,0740	0,1014	0,0980	0,1512	0,4376
Σ	0,10	0,20	0,26	0,20	0,24	1

a) 1. $P(A_5 | N) = \frac{P(A_5 \cap N)}{P(N)} = \frac{0,1512}{0,4376} = 0,3455$

d.h. 35 % der Halter eines Neuwagens sind mindestens 60 Jahre alt.

2. $P(A_1 | N) = \frac{P(A_1 \cap N)}{P(N)} = \frac{0,0130}{0,4376} = 0,0297$

d.h. 3 % der Halter eines Neuwagens sind 18 bis 29 Jahre alt.

3. $1 - 0,0297 - 0,3455 = 0,6248$

d.h. 63% der Halter eines Neuwagens sind 30 bis 59 Jahre alt.

2. Lösungsweg

$$P(A_2 \cup A_3 \cup A_4 | N) = \frac{0,0740 + 0,1014 + 0,0980}{0,4376} = 0,6247715$$

b) $P(N | A_5) = 0,63 \neq 0,4376 = P(N)$

d.h. die Ereignisse A_5 und N sind nicht stochastisch unabhängig.

d.h. die Ereignisse A_5 und N sind stochastisch abhängig.

Lösung zu Aufgabe 4:

X =Anzahl der Ausschusstücke unter den n geprüften Stücken

a) $X \sim B(n = 30; p = 0,05)$

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$P(X = 0) = \binom{30}{0} \cdot 0,05^0 \cdot 0,95^{30} = 0,2146$$

$$P(X = 1) = \binom{30}{1} \cdot 0,05^1 \cdot 0,95^{29} = 0,3389$$

$$P(X = 2) = \binom{30}{2} \cdot 0,05^2 \cdot 0,95^{28} = 0,2586$$

$$0,2146 + 0,3389 + 0,2586 = 0,8122$$

d.h. die Wahrscheinlichkeit beträgt 0,8122.

b) $X \sim \mathbf{B}(n = 240; p = 0,05)$

Faustregel für ZGWS: $np = 12 \geq 10$ und $n(1 - p) = 228 \geq 10$

$$P(X \leq 16) \approx F_U \left(\frac{16 + 0,5 - 12}{\sqrt{12 \cdot 0,95}} \right) = F_U(1,3328) = 0,909$$

d.h. die Wahrscheinlichkeit beträgt näherungsweise 0,909 (exakt 0,9042767).

Lösung zu Aufgabe 5:

$$\text{a) } \bar{x} = \frac{78\,900}{100} = 789$$

$$s_x^2 = \frac{9\,137\,900}{100} = 91\,379$$

$$s_x = \sqrt{91\,379} = 302,2896$$

Faustregel $n \geq 30$ erfüllt

$$\bar{x} \pm 1,96 \cdot \frac{s}{\sqrt{n}} = 789 \pm 1,96 \cdot \frac{302,2896}{10} = 789 \pm 59,2488 = [729,7512; 848,2488]$$

b) Das Intervall [730; 848] ist ein geschätzter Bereich für das Intervall, in dem die mittlere Ausgabenhöhe mit einer Wahrscheinlichkeit von 0,95 liegt.