Statistik-Klausur vom 2.10.2002

Bearbeitungszeit: 120 Minuten

Aufgabe 1
Statistik-Dozent K.R. lehrt an einer privaten FH in Köln, wohnt aber in Frankfurt am Main. Er hat - wegen möglicher saisonaler Schwankungen - zwei Semester lang (SS 01 und WS 01/02) ausprobiert, ob es für ihn - nur aus dem Blickwinkel der arbeitstäglichen Fahrzeit (berechnet vom Verlassen seiner Wohnung bis zum Betreten des FH-Gebäudes in Köln) - günstiger ist, mit dem privaten PKW oder mit dem Intercity-Express (ICE) der Deutschen Bahn AG zu fahren.

60 Fahrten mit dem PKW und 40 Fahrten mit dem ICE ergaben folgende Fahrzeiten:

Benötigte Fahrzeit von ... bis unter ... Minuten	Anzahl der Fahrten mit dem ...	
	... eigenen PKW	... ICE
110 - 120	5	-
120 - 130	15	1
130 - 150	20	9
150 - 180	10	20
180 - 240	10	10
$\sum$	60	40

Für welches der beiden Verkehrsmittel wird er sich auf Grund der von ihm gesammelten Daten entscheiden, wenn er

a) nach dem Kriterium <u>kürzere Fahrzeit</u> entscheidet?

b) nach dem Kriterium <u>gleichmäßigere Fahrzeit</u> entscheidet?

Begründen Sie Ihre Antwort durch Berechnung und Interpretation geeigneter statistischer Maßzahlen!

Lösung:
X = benötigte Fahrzeit mit PKW (in Minuten)
Y = benötigte Fahrzeit mit ICE (in Minuten)

a) Kriterium kürzere Fahrzeit

Benötigte Fahrzeit von ... bis unter ... Minuten	Klassen-mitte	Anzahl der Fahrten mit dem ... eigenen PKW	Dichte	Anzahl der Fahrten mit dem ... ICE	Dichte
110 - 120	115	5	0,0083	-	-
120 - 130	125	15	0,0250	1	0,0025
130 - 150	140	20	0,0167	9	0,0113
150 - 180	165	10	0,0056	20	0,0167
180 - 240	210	10	0,0028	10	0,0042
$\sum$		60		40	

arithmetisches Mittel:

$$\overline{x} \approx 115 \cdot \frac{5}{60} + 125 \cdot \frac{15}{60} + 140 \cdot \frac{20}{60} + 165 \cdot \frac{10}{60} + 210 \cdot \frac{10}{60} = \frac{9\,000}{60} = 150$$

$$\overline{y} \approx 125 \cdot \frac{1}{40} + 140 \cdot \frac{9}{40} + 165 \cdot \frac{20}{40} + 210 \cdot \frac{10}{40} = \frac{6\,785}{40} = 169{,}625 \approx 170$$

d.h. PKW-Fahrten sind im Durchschnitt kürzer.

50%-Punkte:

$x_{0,50} \approx 140$

$y_{0,50} \approx 165$

d.h. in 50% aller PKW-Fahrten lag die Fahrzeit unter 140 Minuten, während in 50% aller ICE-Fahrten die Fahrzeit weniger als 165 Minuten betrug; also sind die PKW-Fahrzeiten kürzer.

Modus:

$x_{\text{Modus}} \approx 125$

$y_{\text{Modus}} \approx 165$

d.h. PKW-Fahrzeiten sind kürzer.

b) Kriterium gleichmäßigere Fahrzeit

Da die Durchschnittswerte der PKW-Daten und der ICE-Daten „weit“ auseinander liegen, sollten hier der Variationskoeffizient oder die relative Quartilsabstand berechnet werden.

Variationskoeffizient:

$$s_x^2 \approx (115-150)^2 \cdot \frac{5}{60} + (125-150)^2 \cdot \frac{15}{60} + (140-150)^2 \cdot \frac{20}{60} + (165-150)^2 \cdot \frac{10}{60} + (210-150)^2 \cdot \frac{10}{60} = 929{,}16$$

$$s_x = \sqrt{929{,}16} = 30{,}48$$

$$v_x = \frac{s_x}{\overline{x}} = \frac{30{,}48}{150} = 0{,}2032$$

$s_y^2 \approx (125-170)^2 \cdot \frac{1}{40} + (140-170)^2 \cdot \frac{9}{40} + (165-170)^2 \cdot \frac{20}{40} + (210-170)^2 \cdot \frac{10}{40} = 665{,}48$
$s_y = \sqrt{665{,}48} = 25{,}80$
$v_y = \frac{s_y}{\overline{y}} = \frac{25{,}80}{170} = 0{,}1518$
d.h. die Fahrzeiten mit dem ICE sind gleichmäßiger.

Relativer Quartilsabstand:
$\frac{x_{0,75} - x_{0,25}}{x_{0,50}} \approx 0{,}27$

$\frac{y_{0,75} - y_{0,25}}{y_{0,50}} \approx 0{,}18$
d.h. die Fahrzeiten mit dem ICE sind gleichmäßiger.

Aufgabe 2

Ausrüstungsinvestitionen[1)] in Preisen von 1995 in Deutschland von 1997 bis 2000 nach Quartalen (in Milliarden DM)

Jahr	Quartal			
	Jan.-März = I	April-Juni = II	Juli-Sept. = III	Okt.-Dez. = IV
1997	58	68	65	77
1998	64	72	72	86
1999	69	77	77	91
2000	77	83	84	97

1) Maschinen, Geräte, Fahrzeuge
Quelle: SVR-Jahresgutachten 2001/2002, Bonn, 21.11.2001, S. 401

Die Zeitreihe der Ausrüstungsinvestitionen wurde in ihre Komponenten zerlegt; die Ergebnisse sind in den Tabellen 2 bis 5 aufgeführt:

a) Tragen Sie die fehlenden Überschriften der Tabllen 2 bis 5 ein [mit anderen Worten: geben Sie an, welche Zeitreihenkomponente(n) in den Werten der jeweiligen Tabelle enthalten ist (bzw. sind)]!

b) Berechnen Sie, so weit dies möglich ist, die noch fehlenden Werte in den Tabellen 2 bis 5! (Hinweis: Runden Sie nach oben bzw. nach unten dabei jeweils auf volle Milliarden DM!)

Tabelle 2:

Jahr	Quartal			
	I	II	III	IV
1997				69
1998	70	72	74	75
1999	77	78	80	81
2000	83	85		

Tabelle 3:

Jahr	Quartale			
	I	II	III	IV
1997			-3	
1998	-6	0	-2	+11
1999	-8	-1	-3	+10
2000	-6	-2		

Tabelle 4:

Quartal	I	II	III	IV
	-7		-3	$+10$

Tabelle 5:

Jahr	Quartal			
	I	II	III	IV
1997	65	69	68	67
1998	71	73	75	76
1999	76	78	80	81
2000	84	84	87	

c) Was sagt der Wert „+10“ in Tabelle 4, Quartal IV, aus?

d) Was sagt der Wert „87“ in Tabelle 5, Jahr 2000, Quartal III, aus?

Lösung:

Additives Modell:

$y_t = \underbrace{g_t}_{\approx x_t} + s_t + u_t$

Arbeitstabelle

Quartal	y_t	4er $\emptyset$	x_t	$d_t = y_t - x_t$	$\overline{d}_t$	$s_t = \overline{d}_t - \overline{d}$	$y_t - s_t$
I	58		–	–	$-6{,}\overline{6} \approx -7$	-7	65
II	68		–	–	-1	-1	69
		67					
III	65		68	-3	$-2{,}\overline{6} \approx -3$	-3	68
		68,5					
IV	77		69	$+8$	$9{,}\overline{6} \approx 10$	10	67
		69,5					
I	64		70	-6	-7	-7	71
		71,25					
II	72		72	0	-1	-1	73
		73,5					
III	72		74	-2	-3	-3	75
		74,75					
IV	86		75	11	10	10	76
		76					
I	69		77	-8	-7	-7	76
		77,25					
II	77		78	-1	-1	-1	78
		78,5					
III	77		80	-3	-3	-3	80
		80,5					
IV	91		81	10	10	10	81
		82					
I	77		83	-6	-7	-7	84
		83,75					
II	83		85	-2	-1	-1	84
		85,25					
III	84		–	–	-3	-3	87
IV	97		–	–	10	10	87

$$\overline{d} = \frac{1}{4}\left[\overline{d}_1 + \overline{d}_2 + \overline{d}_3 + \overline{d}_4\right] = \frac{1}{4}\left[-6{,}\overline{6} - 1 - 2{,}\overline{6} + 9{,}\overline{6}\right] = -0{,}1\overline{6} \approx 0$$

Tabelle 2: Geglättete Zeitreihe x_t

Jahr	Quartal			
	I	II	III	IV
1997	-	-	68	69
1998	70	72	74	75
1999	77	78	80	81
2000	83	85	-	-

Tabelle 3: Saisonkomponente plus Restkomponente $d_t = y_t - x_t \approx s_t + u_t$

Jahr	Quartale			
	I	II	III	IV
1997	-	-	-3	+8
1998	-6	0	-2	+11
1999	-8	-1	-3	+10
2000	-6	-2	-	-

Tabelle 4: Saisonphasen s_t

Quartal	I	II	III	IV
	-7	-1	-3	$+10$

Tabelle 5: Saisonbereinigte Zeitreihe $y_t - s_t$

Jahr	Quartal			
	I	II	III	IV
1997	65	69	68	67
1998	71	73	75	76
1999	76	78	80	81
2000	84	84	87	87

c) Im Winter von 1999 beträgt der Einfluss der Saisonkomponente auf die Ausrüstungsinvestition etwa 10 Milliarden DM und damit liegen Investitionen im Winter deutlich über dem Trend.

d) Der Wert „87“ gibt die Höhe der Ausrüstungsinvestition im 3. Quartal von 2000 an ohne die Einflüsse der Saison.

Aufgabe 3

a) Nennen Sie drei wichtige Unterschiede zwischen einem Zufallsprozess, der einer Binomialverteilung gehorcht, und einem Zufallsprozess, der einer Normalverteilung gehorcht!

b) Wie groß ist die Wahrscheinlichkeit dafür, beim Zahlenlotto „6 aus 49“ bei Abgabe einer einzigen Tipp-Reihe (z.B. 7; 14; 21; 28; 35; 42) sechs Treffer und damit den Hauptgewinn zu erzielen?

c) Angenommen, ein neues Zahlenlotto „5 aus 72“ wird eingeführt (komplett nach den Regeln des derzeitigen Zahlenlottos „6 aus 49“). Ist dann die Wahrscheinlichkeit für einen Hauptgewinn größer oder kleiner als beim derzeitigen Zahlenlotto „6 aus 49“? Begründen Sie Ihre Antwort durch geeignete Berechnungen!

Lösung:

a)
- Bei einer Binomialverteilung ist die zugrunde liegende Zufallsvariable eine diskrete Variable, während bei der Normalverteilung eine stetige Zufallsvariable unterstellt wird.
- Die Normalverteilung ist eine symmetrische Verteilung um den Erwartungswert, während die Binomialverteilung im Allgemeinen eine schiefe Verteilung besitzt (außer für $p = 0{,}5$).

- Die Dichte der Normalverteilung ist an der Stelle $x = \mu$ am größten, während bei der Binomialverteilung die Wahrscheinlichkeitsfunktion an Stelle np sogar null betragen kann.

b) X = Anzahl der Treffer
$P(X = 6) = \frac{1}{\binom{49}{6}} = \frac{1}{13\,983\,816}$
d.h. die Wahrscheinlichkeit ist sehr gering.

c) X = Anzahl der Treffer
$P(X = 5) = \frac{1}{\binom{72}{5}} = \frac{1}{13\,991\,544}$
d.h. die Wahrscheinlichkeit für einen Hauptgewinn ist kleiner.

Aufgabe 4
Im Rahmen einer VWL-Klausur wird nach dem Verfahren „multiple choice“ geprüft. Die Klausur besteht insgesamt aus sechs Fragen. Zu jeder Frage stehen vier Antworten zur Auswahl, von denen genau eine Antwort richtig ist und drei Antworten falsch sind. Ein Prüfling muss sich für genau eine Antwort entscheiden. Lässt der Prüfling eine Frage unbeantwortet, so gilt sie als falsch beantwortet. Alle Prüflinge, die mindestens drei der sechs Fragen richtig beantworten, haben die Prüfung bestanden.

a) Wie viele unterschiedliche Möglichkeiten gibt es, die sechs Fragen zu beantworten?

b) Wie wahrscheinlich ist es, dass ein Prüfling, der überhaupt nichts weiß und nur rät, d.h. alles sechs Antworten zufällig auswählt und ankreuzt, trotzdem alle sechs Fragen richtig beantwortet?

c) Wie viele Ankreuz-Möglichkeiten aus a) führen dazu, dass ein Prüfling die Prüfung besteht?

d) Ein Prüfling hat schon zwei der sechs Fragen durch selbst erarbeitetes Wissen richtig beantwortet. Bei den restlichen Fragen hat dieser Prüfling keinerlei Wissen und kreuzt zufällig an. Wie groß ist die Wahrscheinlichkeit dafür, dass dieser Prüfling die Prüfung nicht besteht?

Lösung:

a) Für alle sechs Fragen muss eine der vier möglichen Antworten angekreuzt werden.
6 aus 4 mit Zurücklegen mit Berücksichtigung der Reihenfolge $\widehat{=} 4^6 = 4\,096$
d.h. es gibt 4 096 verschiedene Ankreuz-Möglichkeiten.

b) X = Anzahl der richtig erratenen Antworten
$X \sim \mathsf{B}(n = 6; p = 0{,}25)$
$P(X = 6) = \binom{6}{6} \cdot 0{,}25^6 \cdot 0{,}75^0 = \frac{1}{4\,096} = 0{,}0002$
d.h. die Wahrscheinlichkeit beträgt 0,0002.

c) X = Anzahl der richtig erratenen Antworten
Wenn X=3 beträgt, weiß man noch nicht, welche der 6 Fragen richtig beantwortet

wurden. Um die Anzahl der Möglichkeiten zu ermitteln, müssen wir erstens aus den 6 Fragen die drei richtig beantworteten Fragen festlegen, .d.h. 3 aus 6 ohne Zurücklegen ohne Berücksichtigung der Reihenfolge ziehen. Und zweitens müssen wir für die 3 falsch beantworteten Fragen das falsche Kreuz festlegen; denn es gibt pro falsche Antwort jeweils drei falsche Kreuze; d.h. es gibt 3^3 Möglichkeiten.
Insgesamt also:
$X = 3$ $\quad \binom{6}{3} \cdot 3^3 = 540$
Und weiter:
$X = 4$ $\quad \binom{6}{4} \cdot 3^2 = 135$
$X = 5$ $\quad \binom{6}{5} \cdot 3 = 18$
$X = 6$ $\quad 1$
Insgesamt sind das 540+135+18+1=694 Möglichkeiten.

c) $X =$ Anzahl der richtig erratenen Antworten
$X \sim \mathsf{B}(n = 6; p = 0{,}25)$
$P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)$
Jetzt ist es schneller, die Wkt. für das Gegenereignis zu berechnen:
$P(X \geq 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) = 1 - \binom{6}{0} \cdot 0{,}25^0 \cdot 0{,}75^6 - \binom{6}{1} \cdot 0{,}25^1 \cdot 0{,}75^5 - \binom{6}{2} \cdot 0{,}25^2 \cdot 0{,}75^4 = 1 - 0{,}8306 = 0{,}1694$
d.h. die Wahrscheinlichkeit beträgt 0,17.

d) Jetzt sind nur noch vier Fragen übrig, die angekreuzt werden müssen.
$Y =$ Anzahl der richtig erratenen Antworten
$Y \sim \mathsf{B}(n = 4; p = 0{,}25)$
$P(Y = 0) = \binom{4}{0} \cdot 0{,}25^0 \cdot 0{,}75^4 = 0{,}3164$
d.h. die Wahrscheinlichkeit beträgt 0,32.

Aufgabe 5

Ein börsengehandeltes Wertpapier hat am 30. September 2002 einen Kurs von 100 € . Binnen eines Monats bleibt der Kurs konstant mit der Wahrscheinlichkeit null. Binnen eines Monats steigt der Kurs mit der Wahrscheinlichkeit 0,4 anhand des Vervielfachungsfaktors 1,25 auf 125 € . Binnen eines Monats fällt der Kurs und mit der Wahrscheinlichkeit 0,6 anhand des Faktors $\frac{1}{1{,}25} = 0{,}80$ auf 80 € .
Daraus folgt unter anderen auch, dass der Kurs nach zwei Monaten, in denen zuerst eine Kurssteigerung erfolgt und anschließend ein Kursrückgang (oder auch in umgekehrter Reihenfolge), wieder genau den Ausgangswert 100 € aufweist (100 € $\cdot 1{,}25 \cdot 0{,}80 = 100$ € $\cdot 0{,}80 \cdot 1{,}25 = 100$ €).
Die Anzahl der Aufwärtsbewegungen des Kurses binnen n Monaten kann als binomialverteilt angesehen werden.

a) Wie viele Aufwärtsbewegungen des Kurses sind im Zeitraum vom 1. Oktober 2002 bis 28. Februar 2003 zu erwarten?

b) Wie groß ist die Wahrscheinlichkeit, dass der Kurs vom 1. Oktober 2002 bis zum 31. März 2003 nur Aufwärtsbewegungen macht?

c) Wie groß ist die Wahrscheinlichkeit, dass der Kurs des Wertpapiers nach vier Monaten (gerechnet vom 1. Oktober 2002 an) wieder den Ausgangswert 100 € hat?

Lösung:
X = Anzahl der Aufwärtsbewegungen binnen n Monaten
$X \sim \mathsf{B}(n; p = 0{,}4)$

a) $X \sim \mathsf{B}(n = 5; p = 0{,}4)$
$E[X] = np = 5 \cdot 0{,}4 = 2$
d.h. in dem Zeitraum ist mit zwei Aufwärtsbewegungen zu rechnen.

b) $X \sim \mathsf{B}(n = 6; p = 0{,}4)$
$P(X = 6) = \binom{6}{6} \cdot 0{,}4^6 \cdot 0{,}6^0 = 0{,}0041$
d.h. die Wahrscheinlichkeit für sechs Aufwärtsbewegungen in sechs Monaten beträgt 0,0041.

c) $X \sim \mathsf{B}(n = 4; p = 0{,}4)$
$P(X = 2) = \binom{4}{2} \cdot 0{,}4^2 \cdot 0{,}6^2 = 0{,}3456$
d.h. die Wahrscheinlichkeit dafür, dass nach vier Monaten wieder der Anfangskurs erreicht wird, beträgt 0,3456.