Statistik-Klausur vom 29. September 2003

Aufgabe 1
Im Geschäftsbericht 2002 der Kali + Salz AG (= K + S) findet sich folgende Aufstellung:

Wertentwicklung eines K+S-Depots (Angaben in €)

Jahresende	1997	1998	1999	2000	2001	2002
Wert des Depots (in €)	5 000	7 620	9 424	12 734	16 388	13 559
Wertzuwachs (in % gegenüber Vorjahr)	•	+52,4	+23,7	+35,1	+28,7	−17,3

Quelle: Geschäftsbericht der Kali + Salz AG

Um wie viel Prozent ist der Wert des K+S-Depots von Ende 1997 bis Ende 2002 im Durchschnitt pro Jahr gewachsen?

Lösung zu Aufgabe 1
Geometrisches Mittel:

$$\sqrt[2002-1997]{\frac{13\,559}{5\,000}} = \sqrt[5]{2{,}7118} = 1{,}2208$$

d.h. im Zeitraum von 1997 bis 2002 betrug die durchschnittliche jährliche Steigerung 22,1 %.

zweiter Lösungsweg:

Jahr	Rate (in %)	Faktor
1997	-	-
1998	+52,4	1,524
1999	+23,7	1,237
2000	+35,1	1,351
2001	+28,7	1,287
2002	−17,3	0,827

Geometrisches Mittel:

$$\sqrt[2002-1997]{1{,}524 \cdot 1{,}237 \cdot 1{,}351 \cdot 1{,}287 \cdot 0{,}827} = \sqrt[5]{2{,}7108} = 1{,}2207 \widehat{=} 22{,}1\%$$

Aufgabe 2

Wirtschaftswachstum und Entwicklung der Arbeitslosigkeit
in sechs Industrieländern von 1991 bis 2001

Land	Wachstum des realen BIP (in %)	Veränderung der Arbeitslosenquote (in %)	
Deutschland	+16	+37	
Frankreich	+21	−9	
Italien	+17	+14	
Großbritannien	+30	−42	
Japan	+11	+131	
USA	+38	−32	

a) In dem Industrieland Spanien ist das BIP im Zeitraum von 1991 bis 2001 real um +30% gestiegen. Ermitteln Sie mit Hilfe der Methode der kleinsten Quadrate, wie sich die Arbeitslosenquote in Spanien in diesem Zeitraum verändert haben dürfte.

b) Analysieren und interpretieren Sie den Zusammenhang zwischen Wirtschaftswachstum einerseits und der Veränderung der Arbeitslosigkeit andererseits durch die Berechnung und Interpretation einer geeigneten statistischen Maßzahl!

Lösung zu Aufgabe 2

X = Wachstum des realen BIP (in %)

Y = Veränderung der Arbeitslosenquote (in %)

i	x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	16	37	592	256	1369
2	21	−9	−189	441	81
3	17	14	238	289	196
4	30	−42	−1260	900	1764
5	11	131	1441	121	17 161
6	38	−32	−1216	1444	1024
$\sum$	133	99	−394	3451	21 595

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

$$b_1 = \frac{6 \cdot (-394) - 133 \cdot 99}{6 \cdot 3451 - 133^2} = -\frac{15\,531}{3\,017} = -5{,}15$$

$$a_1 = \frac{99 - (-5{,}15) \cdot 133}{6} = 130{,}61$$

$f(30) = 130{,}61 - 5{,}15 \cdot 30 = -23{,}8$

d.h. gemäß der Methode der kleinsten Quadrate hat sich die Arbeitslosenquote in Spanien im Zeitraum von 1991 bis 2001 um $-23{,}8\%$ verändert.

b) $$b_2 = \frac{6 \cdot (-394) - 133 \cdot 99}{6 \cdot 21\,595 - 99^2} = -\frac{15\,531}{119\,769} = -0{,}13$$

Bestimmtheitsmaß $B = b_1 \cdot b_2 = (-5{,}15) \cdot (-0{,}13) = 0{,}67$

d.h. 67% der Streuung des Wachstums des BIP wird erklärt durch die Streuung der Regressionsgeraden.
Korrelationskoeffizient $r_{xy} = -\sqrt{B} = -\sqrt{0{,}67} = -0{,}82$
d.h. es liegt starker negativer linearer Zusammenhang vor. Je geringer das reale Wachstum des BIP ausfiel, desto stärker war der Anstieg der Arbeitslosigkeit. Und je höher das reale Wachstum des BIP ausfiel, desto stärker war der Rückgang der Arbeitslosigkeit.

Statt der Raten können Sie auch mit den Faktoren rechnen:

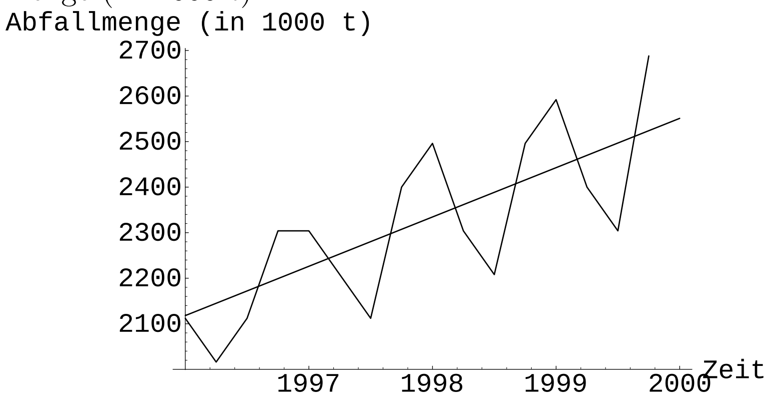
i	x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	1,16	1,37	1,5892	1,3456	1,8769
2	1,21	0,91			
3	1,17	1,14			
4	1,30	0,58			
5	1,11	2,31			
6	1,38	0,68			
$\sum$	7,33	6,99	8,2806	9,0051	10,1395

$b_1 = -5{,}147829$
$a_1 = 7{,}453931$
$f(1{,}30) = 7{,}453931 - 5{,}147829 \cdot 1{,}3 = 0{,}7617533 \mathrel{\widehat{=}} -23{,}8\%$
$r_{xy} = -0{,}8170329$
$B = 0{,}6675428$

Aufgabe 3
In den nachfolgenden zwei Grafiken ist im Zeitraum vom 1.1.1996 bis 31.12.1999 die quartalsmäßige Entwicklung der Höhe der Abfallmengen (in 1 000 t) einer Stadt dargestellt.

Grafik 1

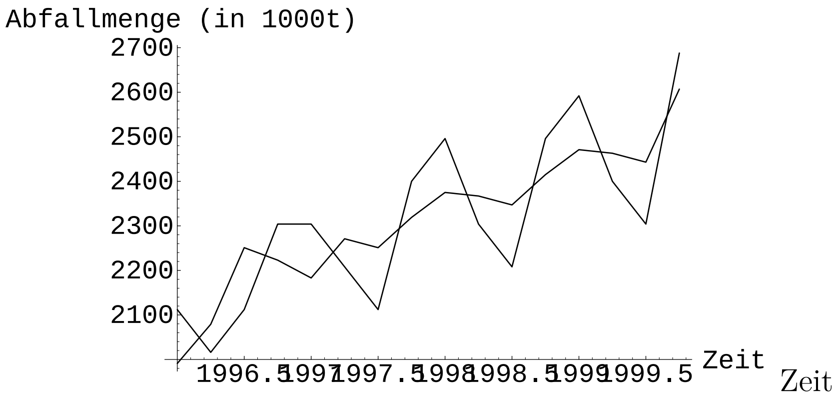
Abfallmenge (in 1 000 t)



Zeit

Grafik 2

Abfallmenge (in 1 000 t)



a) In einer Grafik wurde zusätzlich zur ursprünglichen Zeitreihe die lineare Trendfunktion eingezeichnet und in der anderen Grafik wurde zusätzlich zur ursprünglichen Zeitreihe die saisonbereinigte Zeitreihe eingetragen. In welcher der beiden Grafiken ist die saisonbereinigte Zeitreihe eingezeichnet?

b) Die Saisonkomponenten der vier Quartale betragen:

I	II	III	IV
121	-63	-139	81

Der lineare Trend ist $f(t) = -213\,952{,}46 + 108{,}25 \cdot t$. Geben Sie anhand dieser Daten einen Prognosewert über die Höhe der Abfallmenge im I. Quartal 2000 an. Das I. Quartal 2000 ist der Zeitpunkt $t = 2000$.

Lösung zu Aufgabe 3

a) Grafik 2 stellt zusätzlich zur ursprünglichen Zeitreihe die saisonbereinigte Zeitreihe dar.

b) $f(2000) = -213\,952{,}46 + 108{,}25 \cdot 2000 = 2547{,}54$
$f(2000) + s_1 = 2547{,}54 + 121 = 2668{,}54$
d.h. der Prognosewert für das I. Quartal 2000 beträgt 2 669 000 Tonnen.

Aufgabe 4

Ein Zulieferer eines Automobilkonzerns führt vor der Auslieferung seiner Produkte eine umfassende Qualitätskontrolle durch. Dabei werden seit Jahren die drei Fehlertypen A, B und C festgestellt. Fehlertyp A tritt bei 2%, Fehlertyp B bei 3% und Fehlertyp C bei 5% aller Produktionsstücke auf. Bei 1% der Produktionsstücke werden Fehlertyp A und Fehlertyp B festgestellt.

Die Ereignisse „Ein zufällig ausgewähltes Produktionsstück hat einen Fehler vom Typ A“ und „Ein zufällig ausgewähltes Produktionsstück hat einen Fehler vom Typ C“ sind stochastisch unabhängig.

a) Berechnen Sie die Wahrscheinlichkeit, dass bei einem zufällig ausgewählten Produktionsstück mindestens einer der beiden Fehlertypen A oder B festgestellt werden.

b) Berechnen Sie die Wahrscheinlichkeit, dass bei einem zufällig ausgewählten Produktionsstück vom Fehlertyp B auch der Fehlertyp A festgestellt wird.

c) Sind die Ereignisse „Ein zufällig ausgewähltes Produktionsstück hat einen Fehler vom Typ A“ und „Ein zufällig ausgewähltes Produktionsstück hat einen Fehler vom Typ B“ stochastisch unabhängig? Begründen Sie Ihre Antwort.

d) Berechnen Sie die Wahrscheinlichkeit, dass bei einem zufällig ausgewählten Produktionsstück mindestens einer der beiden Fehlertypen A oder C festgestellt werden.

e) Berechnen Sie die Wahrscheinlichkeit, dass bei einem zufällig ausgewählten Produktionsstück vom Fehlertyp C auch Fehlertyp A festgestellt wird.

Lösung zu Aufgabe 4
A=zufällig ausgewähltes Produktionsstück hat Fehler A
B=zufällig ausgewähltes Produktionsstück hat Fehler B
C=zufällig ausgewähltes Produktionsstück hat Fehler C
$P(A) = 0{,}02$
$P(B) = 0{,}03$
$P(C) = 0{,}05$
$P(A \cap B) = 0{,}01$
A, C sind stoch. unabh., d.h. insb. $P(A \cap C) = P(A) \cdot P(C) = 0{,}02 \cdot 0{,}05 = 0{,}001$

Wir tragen die Wahrscheinlichkeiten in Arbeitstabellen für A, B und für A, C ein:

	A	$\overline{A}$	
B	0,01		0,03
$\overline{B}$			
	0,02		1

	A	$\overline{A}$	
C	0,001		0,05
$\overline{C}$			
	0,02		1

Jetzt lassen sich die Arbeitstabellen vervollständigen:

	A	$\overline{A}$	
B	0,01	0,02	0,03
$\overline{B}$	0,01	0,96	0,97
	0,02	0,98	1

	A	$\overline{A}$	
C	0,001	0,049	0,05
$\overline{C}$	0,019	0,931	0,95
	0,02	0,98	1

a) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0{,}02 + 0{,}03 - 0{,}01 = 0{,}04$
d.h. die Wahrscheinlichkeit beträgt 0,04.

b) $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0{,}01}{0{,}03} = 0{,}33$
d.h. die Wahrscheinlichkeit beträgt 0,33.

c) $P(A \mid B) = 0{,}33 \neq 0{,}02 = P(A)$
d.h. die Ereignisse A, B sind stochastisch abhängig.

d) $P(A \cup C) = P(A) + P(C) - P(A \cap C) = P(A) + P(C) - P(A) \cdot P(C) = 0{,}02 + 0{,}05 - 0{,}02 \cdot 0{,}05 = 0{,}069$
d.h. die Wahrscheinlichkeit beträgt 0,069.

e) $P(A \mid C) = P(A) = 0{,}02$
d.h. die Wahrscheinlichkeit beträgt 0,02.
oder $P(A \mid C) = \dfrac{P(A \cap C)}{P(C)} = \dfrac{0{,}001}{0{,}05} = 0{,}02$

Aufgabe 5

Ein Unternehmen produziert Fliesen in den Farben gelb und weiß. Da beide Farben bei den Kunden gleich beliebt sind, besteht 50% der Produktion aus gelben und 50% der Produktion aus weißen Fliesen.

Die Fliesen werden unsortiert verkauft, d.h. die Kunden können nur eine Anzahl von Fliesen bestellen, wissen aber nicht, wie viele gelbe bzw. weiße Fliesen sie bekommen.

a) Kunde A bestellt 10 Fliesen. Wie wahrscheinlich ist es, dass darunter weniger als 3 gelbe Fliesen sind?

b) Kunde B bestellt 2 500 Fliesen. Wie wahrscheinlich ist es, dass darunter 1 200 oder weniger gelbe Fliesen sind?

Lösung zu Aufgabe 5

a) Y = Anzahl der gelben Fliesen unter den 10 gekauften
$Y \sim \mathsf{B}(n = 10; p = 0{,}5)$
$P(Y < 3) = P(Y \leq 2) = P(Y = 0) + P(Y = 1) + P(Y = 2) = 0{,}0547$
d.h. die Wahrscheinlichkeit beträgt 0,0547.

b) Y = Anzahl der gelben Fliesen unter den 2 500 gekauften
$Y \sim \mathsf{B}(n = 2\,500; p = 0{,}5)$
$n \cdot p = 1\,250 \geq 10$ und $n(1-p) = 1\,250 \geq 10$ d.h. Faustregel für ZGWS erfüllt

$$\begin{aligned} P(Y \leq 1\,200) &\approx F_U\left(\frac{1\,200 + 0{,}5 - np}{\sqrt{np(1-p)}}\right) \\ &= F_U\left(\frac{1\,200 + 0{,}5 - 1\,250}{\sqrt{625}}\right) \\ &= F_U(-1{,}98) \\ &= 0{,}024 \end{aligned}$$

d.h. die Wahrscheinlichkeit beträgt 0,024.
Ohne Korrektur-Term:

$$\begin{aligned} P(Y \leq 1\,200) &\approx F_U\left(\frac{1\,200 - np}{\sqrt{np(1-p)}}\right) \\ &= F_U\left(\frac{1\,200 - 1\,250}{\sqrt{625}}\right) \\ &= F_U(-2) \\ &= 0{,}023 \end{aligned}$$

Aufgabe 6

Die KVB möchte Aufschluss über den Anteil der Kunden erhalten, die über die Haltestelle „Friesenplatz“ fahren. Ein beauftragtes Marktforschungsinstitut schlägt vor, ein 0,95-Konfidenzintervall für diesen Anteilswert zu berechnen.

a) Wie viele Kunden sind mindestens zu befragen, damit das gesuchte 0,95-Konfidenzintervall höchstens die Breite von 0,10 hat, d.h. eine Abweichung um $\pm 5\%$-Punkte vom wahren Wert hat?

b) Von 400 befragten Kunden fuhren 83 über die Haltestelle „Friesenplatz“.

1. Berechnen Sie anhand dieser Stichprobe das gesuchte 0,95-Konfidenzintervall.
2. Interpretieren Sie das erhaltene Konfidenzintervall.

Lösung zu Aufgabe 6:

a) $n \geq \dfrac{1{,}96^2 \cdot 0{,}25}{0{,}05^2} = 384{,}16$

d.h. es sind mindestens 385 Kunden zu befragen.

b) $\dfrac{83}{400} \pm 1{,}96 \cdot \sqrt{\dfrac{\frac{83}{400} \cdot \frac{317}{400}}{400}} = 0{,}2075 \pm 0{,}0397 = [0{,}1678; 0{,}2472]$

d.h. $[16{,}8\%; 24{,}7\%]$ ist ein geschätzter Bereich für das Intervall, in dem der Anteilswert mit der Wahrscheinlichkeit 0,95 liegt.