Statistik-Klausur am 24.09.2013

Bearbeitungszeit: 60 Minuten

Aufgabe 1
Es wird vermutet, dass die Auslieferungszeit (in Minuten) eines Lieferantenfahrers linear abhängt von der Anzahl der zu beliefernden Kunden. An den vergangenen vier Tagen ergaben sich folgende Werte:

Tag	Auslieferungszeit	Kunden
1	250	12
2	500	24
3	750	39
4	720	38

a) Wie lautet die lineare Regressionsgerade?

b) Interpretieren Sie den Wert der Steigung der Regressionsgeraden.

c) Mit welcher Auslieferungszeit ist gemäß der Regressionsgeraden zu rechnen, wenn 20 Kunden beliefert werden sollen?

d) Ist der unter Teilaufgabe c) berechnete Prognosewert aus statistischer Sicht verlässlich?

Aufgabe 2
In der nachfolgenden Tabelle sind für die Hauspreise in Deutschland die Veränderungen in Prozent gegenüber dem Vorjahr (Quelle: OECD) angegeben:

Jahr	Prozent
2004	−1,9
2005	−2,0
2006	0
2007	1,0
2008	0,6
2009	0,5
2010	2,6
2011	5,4

Um wie viel Prozent sind die Hauspreise

a) im Zeitraum 2003 bis 2011 insgesamt gestiegen?

b) im Zeitraum 2003 bis 2011 durchschnittlich pro Jahr gestiegen?

c) im Zeitraum 2006 bis 2011 durchschnittlich pro Jahr real gestiegen, wenn in diesem Zeitraum die Inflationsrate 9,5% insgesamt betrug?

Aufgabe 3

Ein Unternehmen produziert Holzpfeiler. Zur Qualitätssicherung wird vor dem Verkauf jeder Holzpfeiler vermessen und gewogen. Von besonderem Interesse sind dabei die Höhe und das Gewicht. Die Höhe (in m) wird als normalverteilt mit dem Erwartungswert 1,10 m und der Standardabweichung 0,05 m angenommen. Das Gewicht (in kg) ist ebenfalls normalverteilt mit dem Erwartungswert 5 kg und der Standardabweichung 0,1 kg.

a) Mit welcher Wahrscheinlichkeit ist ein zufällig ausgewählter Holzpfeiler schwerer als 4,8 kg?

b) Mit welcher Wahrscheinlichkeit liegt das Gewicht eines zufällig ausgewählten Holzpfeilers zwischen 4,8 kg und 5,2 kg?

c) Welches Gewicht wird mit einer Wahrscheinlichkeit von 95% nicht überschritten?

d) Sind die Ereignisse „ein zufällig ausgewählter Holzpfeiler ist größer als 1,10 m“ und „ein zufällig ausgewählter Holzpfeiler wiegt mehr als 5 kg“ stochastisch unabhängig, wenn 30% der Holzpfeiler größer als 1,10 m sind und mehr als 5 kg wiegen? Begründen Sie Ihre Antwort.

e) Eine Spedition soll eine Lieferung von 10 000 Holzpfeilern transportieren. Aufgrund der Beschaffenheit des verwendeten LKWs darf das Gesamtgewicht der Holzpfeiler maximal 50 020 kg betragen. Wie hoch ist näherungsweise die Wahrscheinlichkeit, dass diese Gewichtsgrenze nicht überschritten wird, wenn das Gewicht der einzelen Holzpfeiler stochastisch unabhängig voneinander ist?

Lösung zu Aufgabe 1:

a) X=Anzahl der Kunden
Y=Auslieferungszeit (in Min)

i	x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	12	250			
2	24	500			
3	39	750			
$n=4$	38	720			
$\sum$	113	2 220	71 610	3 685	1 393 400

$$b_1 = \frac{4 \cdot 71\,610 - 113 \cdot 2\,220}{4 \cdot 3\,685 - 113^2} = \frac{35\,580}{1\,971} = 18{,}05175$$

$$a_1 = \frac{2\,220 - b_1 \cdot 113}{4} = 45{,}03805$$

d.h. die Regressionsgerade lautet $f(x) = a_1 + b_1 \cdot x = 45{,}03805 + 18{,}05175x$

b) $b_1 = 18{,}05175$; d.h. ist ein Kunde mehr zu beliefern, so steigt die Auslieferungszeit um etwa 18 Minuten.

c) $a_1 + b_1 \cdot 20 = 406{,}0731$
d.h. es ist mit einer Auslieferungszeit von etwa 406 Minuten zu rechnen.

d) Der Prognosewert 406 ist ein interpolierter Wert, da $20 \in [12; 39]$ liegt.
$b_2 = \dfrac{35\,580}{4 \cdot 1\,393\,400 - 2\,220^2} = \dfrac{35\,580}{645\,200} = 0{,}05514569$
$r = \sqrt{b_1 \cdot b_2} = 0{,}9977356$
d.h. bei der Prognose handelt es sich um einen interpolierten Wert bei gleichzeitig starker Korrelation, insofern ist der Prognosewert 406 Min als zuverlässig anzusehen.

Lösung zu Aufgabe 2:
Zunächst werden die Faktoren der Veränderung berechnet:

Jahr	Prozent	Faktor
2004	−1,9	0,981
2005	−2,0	0,980
2006	0	1,000
2007	1,0	1,010
2008	0,6	1,006
2009	0,5	1,005
2010	2,6	1,026
2011	5,4	1,054

a) $0{,}981 \cdot 0{,}980 \cdot \ldots \cdot 1{,}054 = 1{,}061618$
d.h. im Zeitraum 2003 bis 2011 sind die Hauspreise um 6,2% insgesamt gestiegen.

b) $\sqrt[2011-2003]{1{,}061618} = \sqrt[8]{1{,}061618} = 1{,}007502$
d.h. im Zeitraum 2003 bis 2011 sind die Hauspreise um 0,8% durchschnittlich pro Jahr gestiegen.

c) $W = 1{,}010 \cdot 1{,}006 \cdot \ldots \cdot 1{,}054 = 1{,}104265$
$Q = \dfrac{W}{P} = \dfrac{1{,}104265}{1{,}095} = 1{,}008461$
$\sqrt[2011-2006]{1{,}008461} = \sqrt[5]{1{,}008461} = 1{,}001687$
d.h. im Zeitraum 2006 bis 2011 sind die Hauspreise um 0,2% durchschnittlich pro Jahr real gestiegen.

Lösung zu Aufgabe 3:
X=Höhe eines Holzpfeilers (in m) und $X \sim \mathsf{N}(\mu = 1{,}10; \sigma = 0{,}05)$
Y=Gewicht eines Holzpfeilers (in kg) und $Y \sim \mathsf{N}(\mu = 5; \sigma = 0{,}1)$

a) $P(Y > 4{,}8) = 1 - F_U\left(\dfrac{4{,}8 - 5}{0{,}1}\right) = 1 - FU(-2) = 1 - 0{,}023 = 0{,}977$
d.h. die Wahrscheinlichkeit beträgt 97,7%.

b) *1. Lösungsweg:*
$P(Y \leq 5{,}2) - P(Y \leq 4{,}8) = F_U\left(\dfrac{5{,}2 - 5}{0{,}1}\right) - 0{,}023 = F_U(2) - 0{,}023 = 0{,}977 - 0{,}023 = 0{,}954$
2. Lösungsweg:

$[4{,}8; 5{,}2] = [\mu - 2 \cdot \sigma; \mu + 2 \cdot \sigma]$ zweifaches zentrales Schwankungsintervall
d.h. die Wahrscheinlichkeit beträgt etwa 95%.

c) $0{,}95 = P(Y \leq y) \Leftrightarrow 1{,}6449 = \dfrac{y-5}{0{,}1} \Leftrightarrow y = 5 + 1{,}6449 \cdot 0{,}1 = 5{,}16449$
d.h. der gesuchte Wert beträgt etwa 5,16 kg.

d) $A =$ „Höhe $> 1{,}1$ m“ $= \{X > \mu_X\} \Rightarrow P(A) = 0{,}5$
$B =$ „Gewicht > 5 kg“$= \{Y > \mu_Y\} \Rightarrow P(B) = 0{,}5$
$P(A \cap B) = 0{,}30 \neq 0{,}25 = 0{,}5 \cdot 0{,}5 = P(A) \cdot P(B)$
d.h. die beiden Ereignisse sind stochastisch abhängig voneinander.

e) Faustregel: $n = 10\,000 \geq 30$ für ZGWS erfüllt.
$P(Y_1 + \ldots + Y_{10\,000} \leq 50\,020) \approx F_U\left(\dfrac{50\,020 - 10\,000 \cdot 5}{\sqrt{10\,000 \cdot 0{,}01}}\right) = F_U(2) = 0{,}977$
d.h. die gesuchte Wahrscheinlichkeit beträgt etwa 97,7%.