Name:

Klausur zur Veranstaltung

Mathematisch-Naturwissenschaftliche Grundlagen

SoSe 2015, Erster Prüfungstermin

27 7.2015

Dozentin: Prof Dr Ulrike Grömping

Erlaubte Hilfsmittel

- Zwei selbstbeschriebene Blätter
- Taschenrechner
- Ein Tafelwerk Ihrer Wahl

Bitte schreiben Sie Ihre Ergebnisse auf leere Blätter,
wobei Sie den Lösungsweg wenigstens skizzieren sollten.

Kennzeichnen Sie jeweils deutlich die bearbeitete Aufgabe / Teilaufgabe,
und markieren Sie gegebenenfalls deutlich das Endergebnis.

Viel Erfolg!

Gesamtpunktzahl	76 Punkte
Sehr Gut-Garantie	ab 64 Punkte
Bestehensgarantie	ab 32 Punkte

40 Punkte

Klausur: Befriedigend (3,3)

Gesamt: Befriedigend (2,7)

Aufgabe	1	2	3	4	5	6	7
Max. Punktzahl	16	13	16	6	8	9	8
	13	10,5	14	1	1,5	–	–

Projekt: 1,7

Ich bin damit einverstanden,
dass meine Note mit Matrikel-Nr in Moodle veröffentlicht wird.

Unterschrift

Aufgabe 1 (Quantile, Boxplot, Lage- und Streuungsmaße) 16

Für ein Farbfeld sind auf 92 Farbkarten die folgenden Werte der Luminanz (L) gemessen worden (schon geordnet):

45,94	46,00	46,37	46,98	47,04	47,12	47,23	47,32	47,38	47,44
47,54	47,62	47,66	47,68	47,70	47,84	47,85	47,95	47,96	48,03
48,04	48,06	48,18	48,26	48,40	48,40	48,64	48,69	48,71	48,72
48,74	48,80	48,87	48,91	48,93	49,07	49,07	49,14	49,25	49,27
49,28	49,29	49,44	49,45	49,48	49,53	49,58	49,59	49,64	49,67
49,71	49,83	49,83	49,83	49,87	49,93	49,93	49,95	50,00	50,06
50,08	50,13	50,14	50,18	50,18	50,28	50,29	50,45	50,45	50,46
50,49	50,52	50,52	50,54	50,60	50,62	50,67	50,70	50,73	50,86
50,90	50,97	50,98	51,05	51,05	51,06	51,11	51,14	51,23	51,47
51,72	54,22								

(i) Berechnen Sie den Median, die beiden Quartile (25% und 75%), und das 10%-Quantil. 4

(ii) Erstellen Sie einen Boxplot *(volle Punktzahl nur, wenn Ihre Überlegungen zu etwaigen Ausreißern erkennbar sind).* 4

(iii) Berechnen Sie den Mittelwert und alle Ihnen bekannten Streuungsmaße für diese Daten. 6

Dabei dürfen Sie die folgenden Angaben verwenden:

$$\sum_{i=1}^{n} x_i = 4540.46$$

$$\sum_{i=1}^{n} (x_i - \overline{x})^2 = 188.131$$

$$\sum_{i=1}^{n} |x_i - x_{0.5}| = 101.338$$

(iv) Warum ist es nicht sinnvoll, für diese Daten den Modus x_{mod} anzugeben? 2

Aufgabe 2 (Beschreibung klassierter Daten) 13

Betrachtet werden klassierte Daten der Luminanz aus der vorherigen Aufgabe, zwecks einfacheren Rechnens betrachten wir nun eine leicht modifizierte Aufgabe mit 100 Karten.

Die folgende Tabelle liegt vor:

j	Luminanz Von über (x_j^u)	bis (x_j^o)	Anzahl (n_j)				
1	44	47	7				
2	47	48.5	22				
3	48.5	50	35				
4	50	52	36				
		Gesamt	100				

(i) Bestimmen Sie relative Häufigkeiten, Häufigkeitsdichten und die empirische Verteilungsfunktion. 6

(ii) Zeichnen Sie das Histogramm der Verteilung des Merkmals (nicht zu ordentlich!). 2

(iii) Berechnen Sie den Modus und den Median. 3

(iv) Zeichnen Sie diese ins Histogramm ein, und erläutern Sie jeweils den Bezug zum Histogramm. 2

Aufgabe 3 (Regression) 16

Fünf Unternehmen wurden nach der Anzahl monatlicher Aufträge (x) und dem Durchschnitts-volumen eines Auftrags (y; in Einheiten von 1000€) befragt.
Die folgenden Ergebnisse ergaben sich (fiktiv):

x_i	20	60	30	30	10
y_i	6	2	8	4	10

(i) Erstellen Sie ein Streudiagramm der Wertepaare, in das Sie später auch die Regressionsgerade einzeichnen können. 2

(ii) Berechnen Sie die Mittelwerte, Varianzen und die Kovarianz s_{xy}. 6

(iii) Berechnen Sie die Regressionsgerade. 2

(iv) Zeichnen Sie die Gerade in das Diagramm aus (i) ein. 2

(v) Berechnen Sie die Korrelation zwischen den *x*- und *y*-Werten, und geben Sie das Bestimmtheitsmaß der Regression an. 2

(vi) Interpretieren Sie die Steigung der Geraden in Bezug auf das zugrundeliegende (fiktive) Sachproblem. 2

Aufgabe 4 (Binomialwahrscheinlichkeiten) 6

Ein Druckprodukt wird mit einer Wahrscheinlichkeit von 2% beim Durchlauf durch die Falzmaschine beschädigt. Es werden 10 unbeschädigte Produkte benötigt. Mit welcher Wahrscheinlichkeit sind unter 11 gedruckten Produkten mindestens 10 unbeschädigt?

Hinweise:

Für $y \in \{0,1,\ldots,n\}$: $P(Y=y)=\binom{n}{y}p^y(1-p)^{n-y}$,

und $\binom{n}{0}=\binom{n}{n}=1$, $\binom{n}{1}=\binom{n}{n-1}=n$

Aufgabe 5 (Normalverteilung; Wahrscheinlichkeiten und Quantile s. Beiblatt) 8

Die Füllmenge in ml (X) eines Produkts mit Nennfüllmenge 100ml ist normalverteilt mit Erwartungswert μ=100 und Varianz σ^2=4.

(i) Die Fertigpackungsverordnung schreibt vor, dass bei Nenn-Füllmengen von 100 ml höchstens 2 Prozent der Packungen weniger als 95.5 ml enthalten dürfen. Prüfen Sie, ob die Fertigpackungsverordnung eingehalten wird. 2

(ii) Welches ist die schärfste Untergrenze, die gerade von nur zwei Prozent der Verpackungen unterschritten wird? 2

(iii) Es ist ebenfalls vorgeschrieben, dass das durchschnittliche Gewicht der Verpackungen die Nennfüllmenge nicht unterschreiten darf Zu diesem Zweck wird die durchschnittliche Füllmenge von 20 Packungen gemessen.

a. Wie ist diese verteilt? 2

b. Mit welcher Wahrscheinlichkeit unterschreitet sie die Nennfüllmenge von 100ml? 2

Aufgabe 6 (Konfidenzintervalle; Quantile s. Beiblatt) 9

Es wird ein Produkt abgefüllt, das eine Nennfüllmenge von 200g einhalten muss. Der Abnehmer hat eine Probe von 25 Packungen genommen und einen Mittelwert von 201g bei einer empirischen Varianz von s^2=16g erhalten.

(i) Geben Sie ein 95%-Konfidenzintervall für die erwartete Füllmenge an, wobei Sie von einer Normalverteilung der Füllmengen ausgehen dürfen. 3

(ii) Aufgrund des Ergebnisses von (i) beklagt sich der Abnehmer beim Lieferanten, dass nicht gewährleistet sei, dass der Erwartungswert mindestens der Nennfüllmenge entspricht. Der Lieferant hat zwar das Recht auf seiner Seite (da er nicht eine Konfidenzintervall-untergrenze von mindestens 200 g liefern muss), möchte aber trotzdem versuchen, den Abnehmer im Rahmen von dessen Argumentation zu überzeugen.

a. Zu diesem Zweck ermittelt er die durchschnittliche Füllmenge aus 400 Packungen, indem er 400 Packungen in ein großes Behältnis gibt und dessen Gesamtinhalt wiegt. Es ergibt sich ein Gewicht von 80.5 kg. Schätzen Sie auf dieser Basis die erwartete Füllmenge. 2

b. *Falls Sie a. nicht geschafft haben, arbeiten Sie hier bitte mit dem Mittelwert von oben weiter* 2

Aufgrund der Vorgehensweise in a. kann keine neue Schätzung für die Varianz bestimmt werden. Daher wird weiterhin die geschätzte Standardabweichung s des Abnehmers verwendet. Man kann mit der folgenden Formel ein 95%-Konfidenzintervall erhalten:

$$\text{Schätzung aus (ii) a.} \quad +/- \quad \frac{s}{\sqrt{n}}(\textit{quantil aus (i)}),$$

wobei das n in dieser Formel sich auf die neue größere Stichprobe bezieht.

Berechnen Sie dieses Konfidenzintervall

c. Mit Hilfe des Intervalls aus b. kann der Lieferant den Abnehmer überzeugen, dass alles in Ordnung ist, weil die Untergrenze größer als die Nennfüllmenge ist. Was hätte der Lieferant noch tun können, wenn auch die in b. berechnete Untergrenze noch unter 200g liegen würde? (Vergrößerung von α soll ausgeschlossen sein) 2

Aufgabe 7 (Testen und p-Werte; Quantile s. Beiblatt) 8

Der Altersdurchschnitt der Druck- und Medientechnik-Erstsemester bis zum Jahr 2011 lag bei etwa 24 Jahren. Es wird behauptet, dass die Druck- und Medientechnik-Erstsemester seit einiger Zeit im Durchschnitt jünger seien, und dass der Altersschnitt inzwischen bei nicht mehr als 23.5 Jahren liege.

(i) Stellen Sie ein Testproblem auf (also Nullhypothese H_0 und Alternativhypothese H_1), mit dem man diese Behauptung widerlegen kann. 2

(ii) Für das WS 2014/15 liegt eine Stichprobe von 43 Befragten vor, bei denen das durchschnittliche Alter bei 23,97 Jahren mit einer Varianz von s^2=14,69 (Quadratjahren) liegt. Führen Sie zum Niveau α=5% einen geeigneten Test für Testproblem aus (i) durch. 4

(iii) Es stehen zwei p-Werte im Raum, die angeblich zu dem Test aus (ii) gehören: 0.021 und 0.210. Welcher der beiden p-Werte ist der richtige, und warum? 2

MNG Formelsammlung für die Klausur SoSe 2015

Zusammenfassung zu den Konfidenzintervallen

Konfidenzintervalle für den Erwartungswert μ

- *Normalverteilung, σ bekannt*

Zweiseitiges $(1-\alpha)$-Konfidenzintervall für μ:

$$\left[\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}, \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}\right]$$

Einseitige $(1-\alpha)$-Konfidenzintervalle für μ:

$$\left[\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}, \infty\right[\quad \text{bzw} \quad \left]-\infty, \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha}\right]$$

- *Normalverteilung, σ unbekannt*

Zweiseitiges $(1-\alpha)$-Konfidenzintervall für μ:

$$\left[\bar{x} - \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}, \bar{x} + \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}\right]$$

Einseitige $(1-\alpha)$-Konfidenzintervalle für μ:

$$\left[\bar{x} - \frac{s}{\sqrt{n}} t_{n-1,1-\alpha}, \infty\right[\quad \text{bzw} \quad \left]-\infty, \bar{x} + \frac{s}{\sqrt{n}} t_{n-1,1-\alpha}\right]$$

- *beliebige Verteilung, n groß*

Zweiseitiges $(1-\alpha)$-Konfidenzintervall für μ:

$$\left[\bar{x} - \frac{s}{\sqrt{n}} z_{1-\alpha/2}, \bar{x} + \frac{s}{\sqrt{n}} z_{1-\alpha/2}\right]$$

Einseitige $(1-\alpha)$-Konfidenzintervalle für μ:

$$\left[\bar{x} - \frac{s}{\sqrt{n}} z_{1-\alpha}, \infty\right[\quad \text{bzw} \quad \left]-\infty, \bar{x} + \frac{s}{\sqrt{n}} z_{1-\alpha}\right]$$

Spezialfall: Konfidenzintervalle für die Erfolgswahrscheinlichkeit *p* bei Binomialverteilung und *n* groß

Faustregeln für n groß. $np(1-p) \geq 5$ oder (strenger) $np(1-p) \geq 9$

$$\hat{p} = \frac{Y}{n} = \frac{\text{Anzahl Erfolge}}{\text{Anzahl Versuche}}$$

Zweiseitiges $(1-\alpha)$-Konfidenzintervall für p:

$$\left[\hat{p} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

Einseitige $(1-\alpha)$-Konfidenzintervalle für p:

$$\left[\hat{p} - z_{1-\alpha} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, 1\right] \quad \text{bzw.} \quad \left[0, \hat{p} + z_{1-\alpha} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

MNG Formelsammlung für die Klausur SoSe 2015

Tests über den Erwartungswert μ einer Verteilung

Testprobleme

a) $H_0.\ \mu = \mu_0$ vs. $H_1\ \mu \neq \mu_0$
b) $H_0.\ \mu \leq \mu_0$ vs. $H_1\ \mu > \mu_0$
c) $H_0.\ \mu \geq \mu_0$ vs. $H_1\ \mu < \mu_0$

Normalverteilte Daten

Teststatistik: $T = \sqrt{n}\dfrac{\overline{x} - \mu_0}{s}$

Prüfverteilung: t(*n*-1)

a) Lehne H_0 ab, g.d w $|T| > t_{n-1;1-\frac{\alpha}{2}}$
b) Lehne H_0 ab, g.d w $T > t_{n-1;1-\alpha}$
c) Lehne H_0 ab, g.d w $T < t_{n-1;\alpha} = -t_{n-1;1-\alpha}$

Beliebig verteilte Daten, *n* groß

Teststatistik: $T = \sqrt{n}\dfrac{\overline{x} - \mu_0}{s}$

Prüfverteilung: N(0, 1)

a) Lehne H_0 ab, g.d w $|T| > z_{1-\frac{\alpha}{2}}$
b) Lehne H_0 ab, g.d.w $T > z_{1-\alpha}$
c) Lehne H_0 ab, g.d.w $T < z_{\alpha} = -z_{1-\alpha}$

Wahrscheinlichkeitstabelle für Aufgabe 5

Tabelliert ist $P(Z \leq z)$

Die Vorkommastellen von z finden Sie in den Zeilen

z	-0.0	-0.25	-0.5	-0.75
-3	0.00135	0.00058	0.00023	0.00009
-2	0.02275	0.01222	0.00621	0.00298
-1	0.15866	0.10565	0.06681	0.04006
0	0.50000	0.40129	0.30854	0.22663

z	0.0	0.25	0.5	0.75
0	0.50000	0.59871	0.69146	0.77337
1	0.84134	0.89435	0.93319	0.95994
2	0.97725	0.98778	0.99379	0.99702
3	0.99865	0.99942	0.99977	0.99991

Quantilstabelle für Aufgaben 5 bis 7

Quantile der Normalverteilung und der *t*-Verteilung

p	z_p	$t_{24.p}$	$t_{25.p}$	$t_{26.p}$	$t_{42.p}$	$t_{43.p}$	$t_{44.p}$
0 01	-2.3263	-2 4922	-2.4851	-2.4786	-2 4185	-2 4163	-2 4141
0.02	-2 0537	-2 1715	-2 1666	-2 1620	-2 1195	-2 1179	-2 1164
0 025	-1 9600	-2 0639	-2 0595	-2 0555	-2 0181	-2 0167	-2.0154
0 05	-1 6449	-1 7109	-1 7081	-1.7056	-1.6820	-1 6811	-1 6802
0 1	-1.2816	-1 3178	-1.3163	-1 3150	-1.3020	-1.3016	-1 3011
0 9	1 2816	1.3178	1 3163	1.3150	1 3020	1 3016	1 3011
0 95	1 6449	1 7109	1 7081	1 7056	1 6820	1 6811	1 6802
0 975	1.9600	2.0639	2.0595	2 0555	2 0181	2.0167	2.0154
0.98	2.0537	2.1715	2 1666	2 1620	2 1195	2 1179	2 1164
0 99	2 3263	2 4922	2 4851	2 4786	2 4185	2 4163	2 4141

① I

$X_{0,5} = \frac{X_{(46)} + X_{(47)}}{2} = \frac{49,53 + 49,58}{2} = \underline{\underline{49,555}}$

$X_{0,25} = [n \cdot p = 92 \cdot 0,25 =]\ X_{(23)} = \underline{\underline{48,18}}$

$X_{0,75} = [n \cdot p = 92 \cdot 0,75] = X_{(69)} = \underline{\underline{50,45}}$

Das wäre die andere Formel

$X_{0,1} = [n \cdot p = 92 \cdot 0,1] = 9,2 \to X_{(10)} = \underline{\underline{47,44}}$

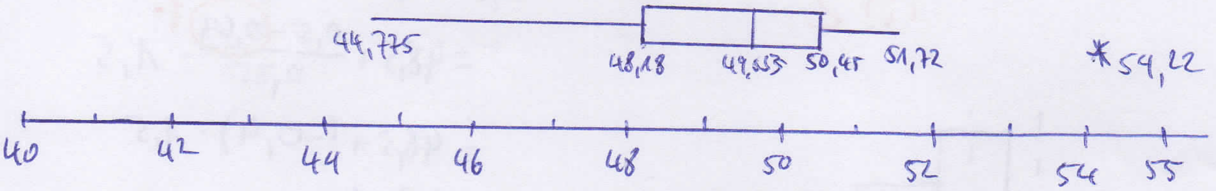
3

II

Boxbreite: $X_{0,75} - X_{0,25} = 50,45 - 48,18 = 2,27$

linkes Zaun $X_{0,25} - 1,5 \cdot 2,27 = 48,18 - 3,405 = 44,775$

rechter Zaun: $X_{0,75} + 1,5 \cdot 2,27 = 50,45 + 3,405 = 53,855$



4

III

$\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i = \frac{1}{92} \cdot 4540,46 = \underline{\underline{49,3528}}$

Spannweite: $X_n - X_1 = 54,22 - 45,94 = \underline{\underline{8,28}}$

IQA = Boxbreite = $\underline{\underline{2,27}}$ (s.o.)

$d = \frac{1}{n} \cdot \sum_{i=1}^{n} |x_i - x_{0,5}| = \frac{1}{92} \cdot 101,338 = \underline{\underline{1,1015}}$

$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{91} \cdot 188,131 = \underline{\underline{2,06737}}$

$s = \sqrt{s^2} = \sqrt{2,06737} = \underline{\underline{1,4378}}$

6

$VK = \frac{s}{|\bar{x}|} = \frac{1,4378}{49,3528} = \underline{\underline{0,0291}}$

IV

Weil die Farbfelder in keinem direkten Bezug zueinander stehen und somit der häufigste Merkmalswert irrelevant ist. ?

mee

13

(2)

I

j	x_j^u	x_j^o	n_j	h_j	d_j	$\hat{F}$
1	44	47	7	0,07	$0{,}02\overline{3}$	0,07
2	47	48,5	22	0,22	$0{,}14\overline{6}$	0,29
3	48,5	50	35	0,35	$0{,}2\overline{3}$	0,64
4	50	52	36	0,36	0,18	1,00
			100	1,0		

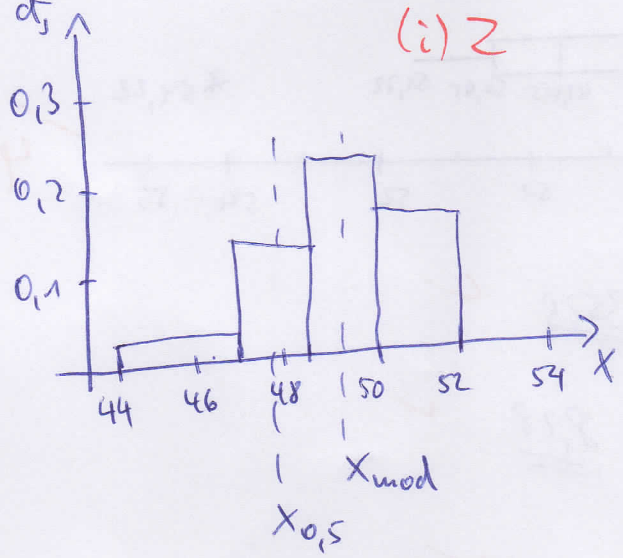
6

rel. Hfgkt $= h_j = \frac{n_j}{n}$

Hfgkt-dichte $= d_j = \frac{h_j}{\text{Klassenbreite}} = \frac{h_j}{x_j^o - x_j^u}$

emp Verteilungsf $= \hat{F}$

II
IV



(i) 2

(iv) 0,5

III $x_{0,5} = x_j^u + \frac{0{,}5 - \hat{F}(x_j^u)}{h_j} \cdot (x_j^o - x_j^u)$

f.

$= 48{,}5 + \frac{0{,}5 - 0{,}64}{0{,}35} \; 1{,}5$

$= 48{,}5 + (-0{,}4) \; 1{,}5$

$= \underline{\underline{47{,}9}}$

x_{mod} = zwischen 50 und 48,5

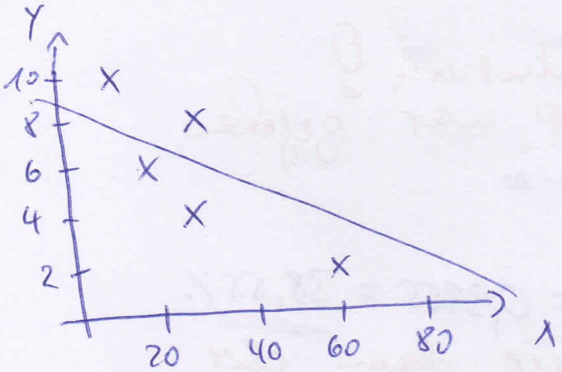
2

genau die Mitte

10,5

③

I



(i) 2

(ii) 2

II

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{5}\, 150 = 30$

$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{5} \cdot 30 = 6$

$S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \overset{f}{=} \frac{1}{4}\cdot\left((-10)^2 + 30^2 + 0^2 + 0^2 + (10^2)\right) = \frac{1}{4}\cdot(100 + 900 + 100)$ 20^2

$= \frac{1}{4}\cdot 1100 = 275$

$S_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{4}\cdot\left(0^2 + (-4)^2 + 2^2 + (-2)^2 + 4^2\right) = \frac{1}{4}\cdot(16 + 4 + 4 + 16)$

$= \frac{1}{4}\cdot 40 = 10$

$S_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})\cdot(y_i - \bar{y}) \overset{f}{=} \frac{1}{4}\cdot(-10 \cdot 0 + 30 \cdot (-4) + 0 \cdot 2 + 0 \cdot (-2) + 10 \cdot 4)$ $-20 \cdot 4$

$= \frac{1}{4}\cdot(-120 + 40) = \frac{1}{4}\cdot(-80) = -20$

5

III $\hat{y} = \hat{a} + \hat{b}x$

$\hat{b} = \frac{S_{xy}}{S_x^2} = \frac{-20}{275} = -0{,}0727$

$\hat{a} = \bar{y} - \hat{b}\cdot\bar{x} = 6 - (-0{,}0727)\cdot 30 = 6 - (-2{,}1818) = \cancel{8{,}1818}\ 8{,}1818$

$\hat{y} = \hat{a} + \hat{b}x = 8{,}1818 + (-0{,}0727)x$

2

V $r_{xy} = \frac{S_{xy}}{\sqrt{S_x^2 S_y^2}} = \frac{20}{\sqrt{275 \cdot 10}} = \frac{20}{52{,}44044} = -0{,}3814$

2

$R^2 = r_{xy}^2 = -0{,}3814^2 = 0{,}14545$

VI Je höher die Anzahl monatlicher Aufträge ist, desto geringer ist das Volumen. genauer?

14

1

④ Gesucht 2%-Quantil $x_{0,02}$ f.

~~P(Z 10/8 2)~~

$n = 11 \quad y = 10$

$\hat{p} = \frac{y}{n} = 0{,}909$ Unsinn! p war gegeben

$p = \binom{n}{y} p^y (1-p)^{n-y} = \binom{11}{10} 0{,}909^{10} (1-0{,}909)^{11-10}$

$= 11 \cdot 0{,}38516 \cdot 0{,}091 = 0{,}3855 = \underline{38{,}55\%}$

Das wäre denn die W'keit für Y = 10, gesucht war mind. 10

1

5

I $X \sim N(100, 4)$ ✓ $\mu = 100$ $\sigma^2 = 4$ $\sigma = 2$

$X_{0,02}$ gesucht

$P(X < 95{,}5)$

$P\left(Z_{0,02} \overset{\leq}{=} \underbrace{\frac{95{,}5 - 100}{2}}\right) = -2{,}25 \rightarrow 1{,}96\,\%$ f. $1{,}22\,\%$

Es wird mit $1{,}96\,\%$ gerade so eingehalten. (✓)

1,5