

Bachelorprüfung: Statistik (1 Stunde)

Bemerkungen:

- Es sind alle mitgebrachten schriftlichen Hilfsmittel und der Taschenrechner erlaubt.
- Natels sind auszuschalten!
- Lesen Sie zuerst alle Aufgaben durch! Verweilen Sie nicht zu lange bei einem Aufgabenteil, der Ihnen grosse Schwierigkeiten bereitet! Es wird nicht erwartet, dass Sie alle Aufgaben vollständig lösen.
- Die nötigen Tabellen befinden sich auf den hintersten Seiten dieser Prüfung.
- Bei den Aufgaben 1 und 2 muss der Lösungsweg ersichtlich sein, sonst gibt es keine Punkte.
- Für die Antworten zu den Aufgaben 1 und 2 sind die zugehörigen Antwortblätter zu benutzen. Bitte schreiben Sie alle Antworten, welche Sie abgeben möchten, nur auf die vorbereiteten Antwortblätter. Antworten auf eigens mitgebrachten Blättern werden nicht korrigiert. Es sind zwei Reserve-Antwortblätter vorhanden.
- Aufgaben 3 und 4 sind Multiple-Choice-Aufgaben. Es ist jeweils genau eine Antwort korrekt. Eine korrekte Antwort gibt 1 **Pluspunkt** und eine falsche Antwort $\frac{1}{2}$ **Minuspunkt**. Minimal erhält man für eine ganze Multiple-Choice Aufgabe 0 Punkte. Tragen Sie die korrekten Antworten der Multiple-Choice-Aufgaben mit Kreuzchen in das zugehörige Antwortblatt ein.
- Rechnen Sie immer mit den exakten Zwischenresultaten weiter. Runden sie alle auf dem Papier angegebenen Endergebnisse sinnvoll.

Viel Erfolg!

1. (8 Punkte)

Homer S. kauft seine Lebensmittel entweder im Kwik-E-Mart oder im Wull-Mart. Bis jetzt ist er davon ausgegangen, dass beide Läden im Mittel gleich teuer sind. Er möchte nun überprüfen, ob dies wirklich so ist. Deshalb geht er an einem Tag in beide Geschäfte, und kauft von seinen Lieblingsprodukten je eines. Die Produkte sind identisch in beiden Geschäften. Er stellt folgende Preise (in USD) fest:

Produkt	Donuts	Burgers	Hot-Dogs	Speck	Duff Bier	Pizza
Kwik-E-Mart	4	6	3	1	3	3
Wull-Mart	5	5	4	2	5	2

- a) Handelt es sich hier um einen gepaarten oder einen ungepaarten Test? Begründen Sie.
- b) Welche Annahme über die Verteilung der Differenzen der Preise macht man typischerweise, um einen t -Test durchzuführen? Wie kann man diese Annahme überprüfen?
- c) Geben Sie die Null- und die Alternativhypothese an.
- d) Wir betrachten die Differenzen (Preis Kwik-E-Mart) $-$ (Preis Wull-Mart). Schätzen Sie die Varianz σ^2 dieser Differenzen.
- e) Führen Sie einen t -Test auf dem 10%-Niveau durch. Was ist der Testentscheid?
- f) Geben Sie ein zweiseitiges 90%-Vertrauensintervall für den Mittelwert der Differenzen μ an. Geben Sie die Antwort auf drei Nachkommastellen genau an.

2. (7 Punkte)

Marco ist ein Spargelliebhaber. Deshalb konsumiert er während der Spargelsaison jeden Mittag einen Spargelsalat in der Kantine, der jeweils 10 Spargelstücke enthält. Diese werden in der Küche durchs Schneiden ganzer Spargeln in 6 Stücke hergestellt.

- a) Marco mag besonders die Spargelspitzen, deshalb zählt er jeweils deren Anzahl (Variable X) auf seinem Teller. Wie ist X verteilt unter der Annahme, dass bei der Herstellung des Salats zufällig 10 Stücke aus einem grossen, gut durchmischten Topf genommen werden? (Verteilungsfamilie sowie Parameter angeben!)
- b) Wie gross ist die Wahrscheinlichkeit $p = P(X \geq 3)$, mindestens 3 Spargelspitzen auf seinem Teller zu finden? (Numerisches Resultat auf 4 Stellen.)
- c) Marco verdächtigt das Kantinenpersonal, beim Zubereiten der Salate gezielt Spargelspitzen herauszusuchen und zu naschen. Deshalb zählt er die Zahl der Spitzen, die er über die ganze Spargelsaison verspeist; er nennt diese Zahl Y . Während einer Spargelsaison isst er 48 Spargelsalate, verspeist also insgesamt $48 \cdot 10 = 480$ Spargelstücke. Wie ist Y *approximativ* verteilt unter der Annahme, dass das Personal *keine* Spitzen nascht? (Verteilungsfamilie und Parameter angeben.)
- d) Während einer Saison zählt Marco $Y = 72$ Spargelspitzen. Stimmt seine Vermutung, dass das Personal Spitzen nascht? Führen Sie mit der Verteilung aus ?? einen Test auf dem 5%-Niveau durch. Geben Sie explizit die Null- und Alternativhypothese an; muss ein- oder zweiseitig getestet werden?

3. (8 Punkte)

Es liegen uns Daten über die Entwicklung der Geburtenrate von 20 lateinamerikanischen Ländern vor. Die Daten beinhalten die prozentuale Abnahme der Geburtenrate (y) zwischen den Jahren 1965 und 1975 in den Ländern und einen Index (x) der den sozialen Hintergrund der Länder beschreibt. Folgendes lineare Modell wurde an diese Daten angepasst:

$$y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i, \quad \varepsilon_i \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2), \quad i = 1, \dots, 20.$$

Der Regressionsoutput sieht wie folgt aus:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	???	9.6416	-2.295	0.03398
x	0.5052	0.1308	???	???

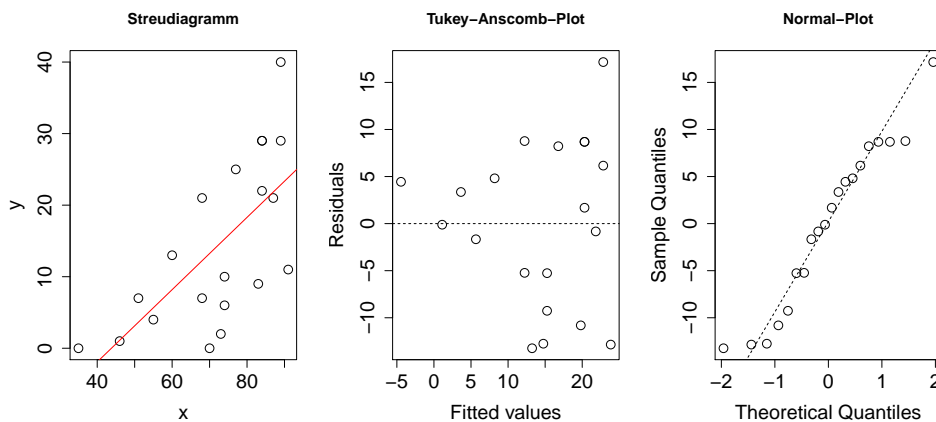
Residual standard error: 8.973 on ??? degrees of freedom

Multiple R-squared: 0.4532, Adjusted R-squared: 0.4228

F-statistic: 14.92 on 1 and 18 DF, p-value: 0.001141

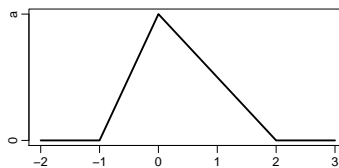
- Wie gross ist der Intercept $\hat{\beta}_0$?
 - 9.6416
 - 30.863
 - 3.863
 - 24.679
 - 23.245
 - 22.125
- Wieviele Freiheitsgrade hat der "Residual standard error"?
 - 20
 - 19
 - 16
 - 17
 - 18
 - 21
- Wie gross ist der t -Wert der Schätzung von β_1 ?
 - 1.863
 - 0.52
 - 3.863
 - 2.458
 - 10.253
 - 3.121
- Welches der folgenden Intervalle ist ein exaktes zweiseitiges 95% Vertrauensintervall für β_1 ?
 - $0.5052 \pm 2.101 \cdot \frac{0.1308}{\sqrt{18}}$
 - $0.5052 \pm 1.960 \cdot \frac{0.1308}{\sqrt{18}}$
 - $0.5052 \pm 2.101 \cdot 0.1308$
 - $0.5052 \pm 1.960 \cdot 0.1308$
- Wird die Nullhypothese $H_0 : \beta_1 = 0$ auf dem 1%-Niveau verworfen?
 - Ja
 - Nein
 - Keine Aussage möglich
- Wie gross ist die Schätzung von σ^2 ?
 - 9.23
 - 50.34
 - 80.51
 - 8.973
 - 40.86
 - Keine Aussage möglich

7) Betrachten Sie die Plots. Welche der folgenden Aussagen ist zutreffend?



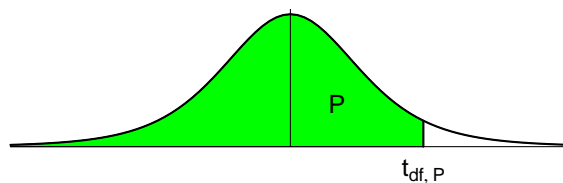
- Die Modellannahmen über die Fehler sind erfüllt.
 - Die Modellannahmen über die Fehler scheinen plausibel. Es gibt aber Ausreisser.
 - Die Annahme der konstanten Varianz scheint grob verletzt zu sein, aber die Normalitätsannahme der Fehler ist plausibel.
 - Es gibt systematische Fehler im Modell.
- 8) Was passiert mit der Regression, wenn an der Stelle (40, 30) noch eine Beobachtung hinzugefügt wird?
- $\hat{\beta}_1$ wird kleiner, $\hat{\sigma}$ wird kleiner
 - $\hat{\beta}_1$ wird kleiner, $\hat{\sigma}$ wird grösser
 - $\hat{\beta}_1$ wird grösser, $\hat{\sigma}$ wird kleiner
 - $\hat{\beta}_1$ wird grösser, $\hat{\sigma}$ wird grösser

Die folgende Abbildung zeigt die Wahrscheinlichkeitsdichte f_X einer Zufallsvariablen X :



- 7) Welchen Wert hat das Maximum a der Dichte f_X ?
- a) $\frac{1}{2}$ b) $\frac{4}{5}$ c) 1
d) $\frac{3}{2}$ e) 4 f) keinen der genannten Werte
- 8) Ist der Median μ der Zufallsvariablen X grösser oder kleiner als ihr Erwartungswert $\mathcal{E}(X)$?
- a) $\mu < \mathcal{E}(X)$
b) $\mu = \mathcal{E}(X)$
c) $\mu > \mathcal{E}(X)$
d) das hängt von der Anzahl Beobachtungen von X ab

Perzentile der t-Verteilung



Bsp.: $t_{9; 0.975} = 2.262$

df	$t_{0.60}$	$t_{0.70}$	$t_{0.80}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750
31	0.255	0.530	0.853	1.309	1.696	2.040	2.452	2.744
32	0.255	0.530	0.853	1.309	1.694	2.037	2.449	2.738
33	0.255	0.530	0.853	1.308	1.693	2.035	2.445	2.733
34	0.255	0.529	0.852	1.307	1.691	2.032	2.441	2.728
35	0.255	0.529	0.852	1.306	1.690	2.030	2.438	2.724
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660
90	0.254	0.526	0.846	1.291	1.662	1.987	2.368	2.632
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358	2.617
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576