

Bachelorprüfung: Mathematik 4 - Statistik (2 Stunden)

Bemerkungen:

- Es sind alle 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! Für die Note 6 brauchen nicht alle Aufgaben gelöst zu sein!
- Wenn nicht anders vermerkt, sind die Tests auf dem 5%-Niveau durchzuführen.
- Die nötigen Tabellen befinden sich auf den hintersten Seiten dieser Prüfung.
- Aufgaben 4 und 5 sind Multiple-Choice Aufgaben. Es ist jeweils genau eine Antwort korrekt. Eine korrekte Antwort gibt 1 **Plus**punkt und eine falsche Antwort $\frac{1}{2}$ **Minus**punkt. Minimal erhält man für eine ganze Multiple-Choice Aufgabe 0 Punkte. Tragen Sie die korrekten Antworten der Multiple Choice Aufgaben mit Kreuzchen in das separate Antwortblatt ein.

Viel Erfolg!

1. (7 Punkte)

Die Brauerei Wiesenburg verfügt über eine spezielle Maschine zur Abfüllung von Bierflaschen. Die Brauerei musste die Abfüllmaschine neulich auswechseln, da sie kaputt ging. Der Verein “lustige Biertrinker” hat nun den Verdacht, dass die Flaschen, welche mit der neuen Maschine abgefüllt werden, weniger Bier enthalten als die alten Flaschen. Daher messen die “lustigen Biertrinker” mit einem Präzisionsmessgerät die Biermengen in je 10 zufällig ausgewählten alten und neuen Flaschen und vergleichen die Werte (in Litern):

Alt	0.512	0.508	0.526	0.481	0.533	0.506	0.476	0.486	0.501	0.500
Neu	0.451	0.496	0.477	0.488	0.493	0.508	0.495	0.502	0.473	0.479
Differenz	0.061	0.012	0.049	-0.007	0.040	-0.002	-0.019	-0.016	0.028	0.021

Kennzahlen:

Alt	\bar{x}	=	0.5029	$\hat{\sigma}_x$	=	0.018
Neu	\bar{y}	=	0.4862	$\hat{\sigma}_y$	=	0.017
Differenz	$\bar{x} - \bar{y}$	=	0.0167	S_{pool}	=	0.018

Sie dürfen davon ausgehen, dass die Biermengen durch unabhängige $\mathcal{N}(\mu, \sigma^2)$ -verteilte Zufallsvariablen beschrieben werden können.

- Begründen Sie kurz, warum es sich um eine ungepaarte Stichprobe handelt.
- Bestimmen Sie die Null- und die Alternativhypothese H_0 und H_A .
- Führen Sie einen t-Test durch. Geben Sie den Wert der Teststatistik und den Verwerfungsbereich des Tests an. Wie entscheidet der Test?
- Die “lustigen Biertrinker” haben die Vermutung, dass die neue Maschine auf mindestens 0.02 Liter weniger geeicht ist als die alte. Sie haben glücklicherweise ein Statistikprogramm, welches ihnen das einseitige 95%-Vertrauensintervall für die Differenz (alt minus neu) liefert. Dieses beträgt $[0.0027, \infty)$ Liter. Würden die “lustigen Biertrinker” diese neue Hypothese auf dem 5%-Niveau verwerfen? Begründen Sie kurz.

2. (7 Punkte)

20% aller Kälber erkranken in den ersten sechs Lebensmonaten an einer bestimmten nicht ansteckenden Krankheit. Bauer Franz ist neu in der Kälberzucht, daher interessieren ihn gewisse Wahrscheinlichkeiten im Zusammenhang mit dieser Krankheit.

- Wie gross ist die Wahrscheinlichkeit, dass von den nächsten drei Kälbern genau zwei an besagter Krankheit erkranken?
- Franz würde gerne berechnen, wie viele Kälber auf seinem Hof (mindestens) auf die Welt kommen müssen, damit mit mindestens 90% Wahrscheinlichkeit eines oder mehr die Krankheit besitzt.
- Ein neuer Impfstoff wurde getestet. In einer Studie wurden 100 Kälber damit geimpft. 15 Kälber sind trotzdem erkrankt. Wie gross ist die Wahrscheinlichkeit, dass 15 oder weniger als 15 Kälber erkranken, unter der Annahme, dass der Impfstoff keine Wirkung hat (d.h. 20% Wahrscheinlichkeit, dass die Krankheit auftritt)? Hinweis: Benutzen Sie eine geeignete Approximation.
- Berechnen Sie mit den Angaben aus Aufgabe c) ein approximatives zweiseitiges 90%-Vertrauensintervall für die Erkrankungswahrscheinlichkeit p mit dem neuen Impfstoff.

3. (8 Punkte)

Eine Firma unterzieht alle elektronischen Bauteile, die durch ein bestimmtes Produktionsverfahren hergestellt werden, einer ersten Qualitätskontrolle: im Mittel werden 92% der Bauteile akzeptiert. Die restlichen 8% werden einer strengeren Kontrolle unterzogen: von diesen werden im Mittel 25% verworfen und 75% akzeptiert.

Wenn Sie eine Teilaufgabe nicht lösen können, so verwenden Sie für die folgenden Teilaufgaben das angegebene Ersatzresultat, falls Sie dieses benötigen.

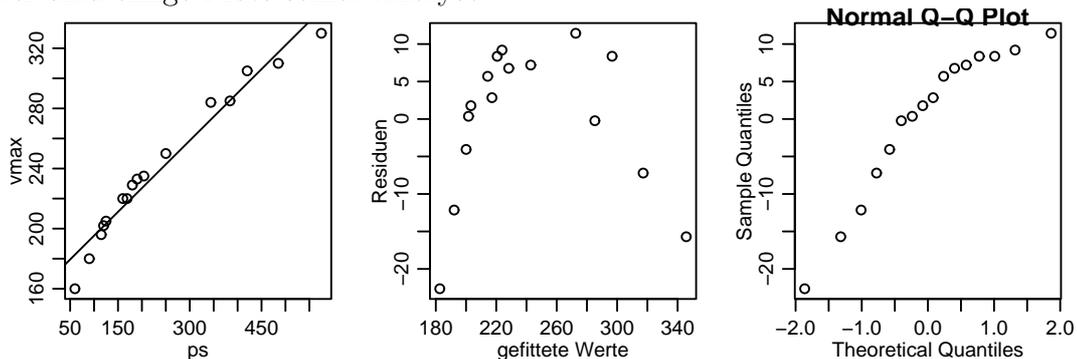
- Wie gross ist die Wahrscheinlichkeit, dass ein zufällig ausgewähltes Bauteil verworfen wird?
Ersatzresultat: 0.05
- Wie gross ist die Wahrscheinlichkeit, dass aus 20 zufällig ausgewählten Bauteilen höchstens 2 verworfen werden?
Ersatzresultat: 0.9534
- Im Mittel werden täglich 100 Bauteile verworfen: wie viele Bauteile werden täglich produziert?
Ersatzresultat: 5000
- Jede einzelne Anwendung der ersten Qualitätskontrolle kostet 10 Rappen, und jede einzelne Anwendung der strengeren Qualitätskontrolle kostet 25 Rappen. Wieviel wird im Mittel täglich gespart durch die Anwendung des obigen Verfahrens im Vergleich zum Verfahren, das alle Bauteile der strengeren Kontrolle unterzieht?

4. (9 Punkte)

Christoph ist ein grosser Auto-Fan und analysiert die Autos, die in der neusten Ausgabe der Zeitschrift "Auto, Motor und Sport" vorgestellt werden. Dazu rechnet er eine lineare Regression mit der Anzahl PS (Pferdestärken) als erklärende Variable und der Höchstgeschwindigkeit (vmax, in km/h) als Zielgrösse. Das Modell ist:

$$vmax_i = \beta_0 + \beta_1 \cdot ps_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d.}$$

Hier sind einige Plots seiner Analyse:



- Welche der nachfolgenden Aussagen ist zutreffend?
 - Die Plots sind in Ordnung. Das Modell passt sehr gut.
 - Der Zusammenhang ist anscheinend nicht linear. Die Annahmen über die Fehler sind nicht verletzt.
 - Der Zusammenhang ist anscheinend nicht linear. Deshalb machen auch Aussagen über die Fehler keinen Sinn.
 - Der Zusammenhang ist anscheinend nicht linear. Dies stört aber nicht, da die Residuen sehr klein sind.

Christoph logarithmiert (zur Basis 10) sowohl die Anzahl PS als auch die Höchstgeschwindigkeit und rechnet nochmals eine lineare Regression. Das Modell ist nun:

$$\log(\text{vmax}_i) = \beta_0 + \beta_1 \cdot \log(\text{ps}_i) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d.}$$

Hier sind sein R-Output und einige Plots:

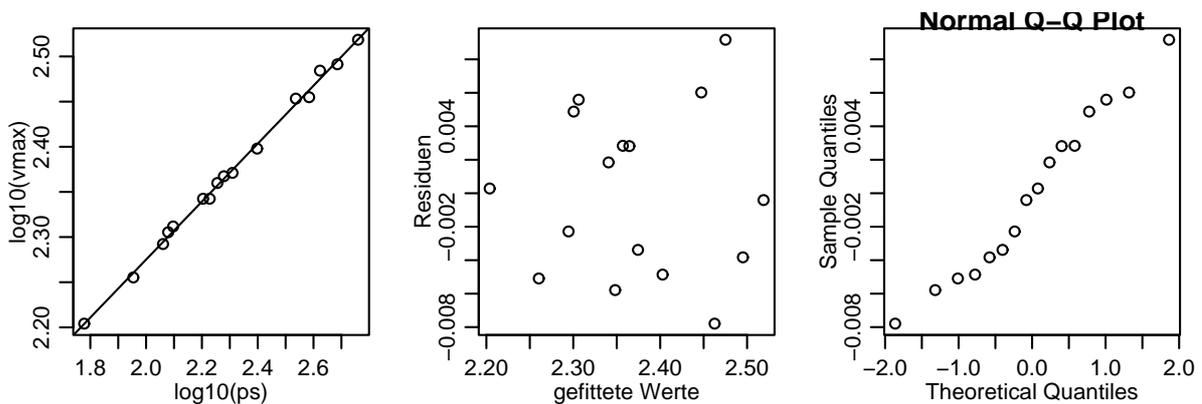
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.633	0.0110	??	<2e-16 ***
log10(ps)	0.321	0.0047	??	<2e-16 ***

Residual standard error: 0.005141 on 14 degrees of freedom

Multiple R-squared: ??, Adjusted R-squared: ??

F-statistic: 4550 on 1 and 14 DF, p-value: < 2.2e-16



- 2) Mit wievielen Beobachtungen wurde die Regression berechnet?
 - a) 12
 - b) 13
 - c) 14
 - d) 15
 - e) 16
 - f) 4550
- 3) Wie gross ist die t-Teststatistik für den Test der Nullhypothese $H_0 : \beta_1 = 0$ (die Alternativhypothese ist $H_A : \beta_1 \neq 0$)?
 - a) 80.4
 - b) 68.3
 - c) 0.0047
 - d) 148.5
 - e) 4550
 - f) Keine Aussage möglich
- 4) Wird $H_0 : \beta_1 = 0$ auf dem 5% Niveau verworfen (die Alternative ist $H_A : \beta_1 \neq 0$)?
 - a) Ja
 - b) Nein
 - c) Keine Aussage möglich
- 5) Welches der folgenden Intervalle ist ein exaktes zweiseitiges 95% Vertrauensintervall für β_1 ?
 - a) $0.321 \pm 1.96 \cdot \frac{0.0047}{\sqrt{14}}$
 - b) $0.321 \pm 2.145 \cdot \frac{0.0047}{\sqrt{14}}$
 - c) $0.321 \pm 1.96 \cdot 0.0047$
 - d) $0.321 \pm 2.145 \cdot 0.0047$
- 6) Christoph kauft sich ein Auto mit 200 PS. Welche Höchstgeschwindigkeit (in km/h) kann er erwarten?
 - a) 66
 - b) 226
 - c) 232
 - d) 235
 - e) 240
 - f) 242
- 7) Wieviele PS sind in etwa nötig, damit ein Auto eine Höchstgeschwindigkeit von 300 km/h erreicht?
 - a) 268
 - b) 392
 - c) 408
 - d) 414
 - e) 426
 - f) 929
- 8) Wie gross ist das Bestimmtheitsmass R^2 ?
 - a) $R^2 < 2.2e-16$
 - b) $0.1 \leq R^2 < 0.3$
 - c) $0.3 \leq R^2 < 0.7$
 - d) $0.9 \leq R^2 < 1$
 - e) $R^2 = 1$
 - f) $R^2 > 1$

Christoph rechnet nun noch eine multiple Regression (wieder mit $\log(\text{vmax})$ als Zielvariable) mit der zusätzlichen Erklärenden Gewicht des Autos (in Tonnen). Hier ist der Output:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.627	0.011653	139.598	< 2e-16 ***
$\log_{10}(\text{ps})$	0.330	0.008327	39.687	5.93e-15 ***
gewicht	-0.012	0.008773	-1.364	0.196

9) Welche Aussage ist richtig?

- Bei der Schätzung des Koeffizienten von $\log_{10}(\text{ps})$ im multiplen linearen Modell wurde der lineare Einfluss von `gewicht` auf die Zielgröße berücksichtigt.
- Multiple Regression liefert ganz allgemein bis auf kleine numerische Abweichungen jeweils gleiche Parameterschätzungen wie die individuellen einfachen linearen Regressionen.
- Dass der Koeffizient von $\log_{10}(\text{ps})$ im multiplen Modell auch signifikant von Null verschieden ist, hätte man schon aus dem Output der einfachen linearen Regression vorhersagen können.
- Es stimmt überhaupt keine der obigen Aussagen.

5. (5 Punkte)

- Die Anzahl der Fehler, die bei einem bestimmten Computer während eines Tages auftreten, folgt einer Poisson-Verteilung. Die erwartete Anzahl Fehler ist 0.7. Wie gross ist die Standardabweichung?
 - 0.46
 - 0.21
 - 0.49
 - 0.84
 - 0.7
 - keine Aussage möglich
- Für die unabhängigen Zufallsvariablen X und Y seien $\mathbf{E}[X] = 2$, $\mathbf{E}[Y] = 1$, $\text{Var}(X) = 1$ und $\text{Var}(Y) = 2$. Wie gross ist die Varianz der Zufallsvariable $2X - Y$?
 - 2
 - 0
 - 2
 - 4
 - 6
 - keine Aussage möglich
- Die stetige Zufallsvariable X hat Dichte:

$$f(x) = \begin{cases} 2x & \text{falls } 0 \leq x \leq 1, \\ 0 & \text{sonst.} \end{cases}$$

Wie gross ist der Erwartungswert von X ?

- 0.5
 - 0.67
 - 1
 - 1.33
 - 2
 - keine Aussage möglich
- Die diskrete Zufallsvariable X ist definiert durch $P[X = 1] = 0.3$, $P[X = 2] = 0.2$, und $P[X = 4] = 0.5$. Wie gross ist die Varianz von X ?
 - 2.45
 - 1.41
 - 2.3
 - 1.81
 - 2.7
 - keine Aussage möglich

5) Betrachten Sie die folgenden Darstellungen. Welche der aufgeführten Zuordnungen von Histogramm und Boxplot ist richtig?

a) A2, B1, C3, D4, E5

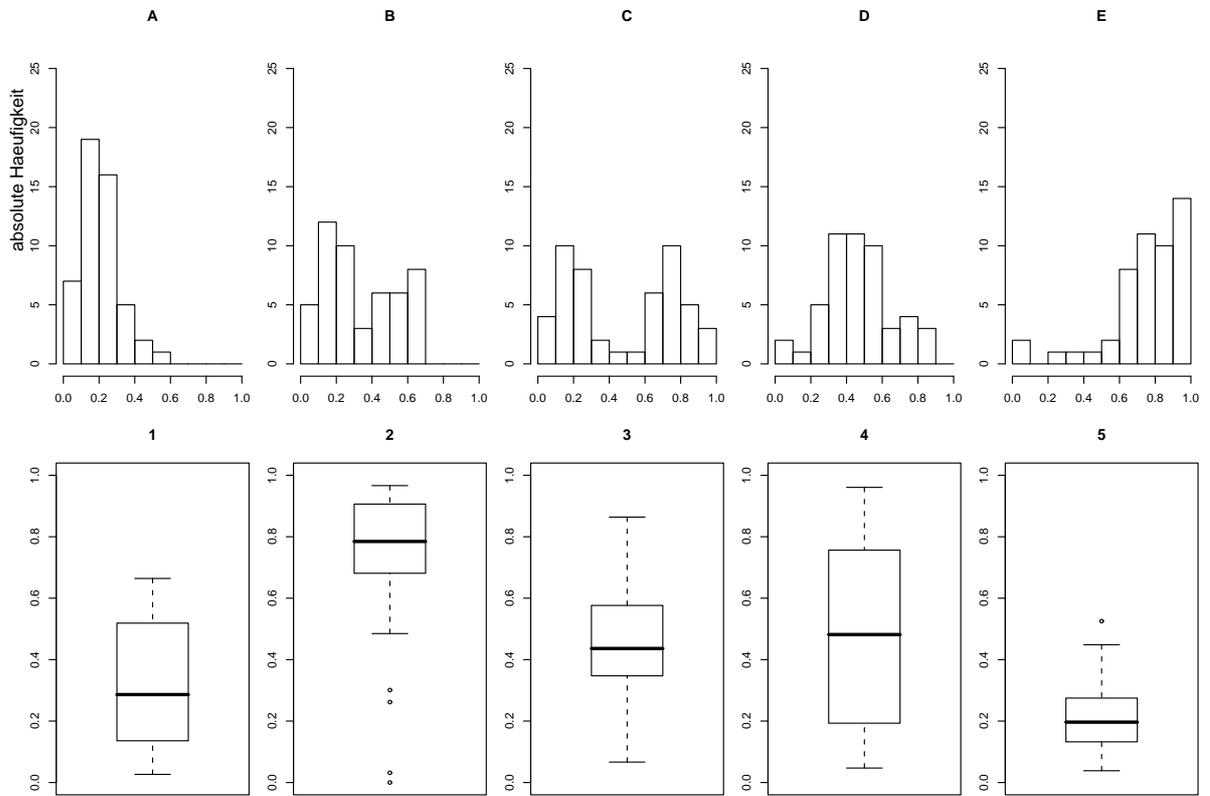
b) A5, B4, C1, D3, E2

c) A5, B1, C4, D3, E2

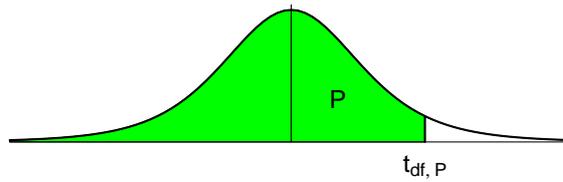
d) A1, B5, C3, D4, E2

e) A1, B5, C4, D3, E2

f) A5, B1, C3, D4, E2



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
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