

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!
- 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 Schweizer Skinationalmannschaft möchte zwei verschiedene Wachse für sehr kalten Schnee miteinander vergleichen. Einerseits den “Ice Blizzard” und andererseits den “Freeze Rocket” Wachs. An einem eisig kalten Tag machen sie Tests. Alle 8 Leute des Kaders fahren einen ausgesteckten Kurs jeweils einmal mit einem Ski, der mit “Ice Blizzard” präpariert wurde und einmal mit einem Ski, der mit “Freeze Rocket” eingewachst wurde. Beide Male wird die Zeit gestoppt. Der Trainer notiert sich Folgendes (dabei bezeichnen X_i die Werte für “Ice Blizzard”, Y_i die Werte für “Freeze Rocket” und $D_i = X_i - Y_i$ in Sekunden):

	Marco	Carlo	Silvan	Didier	Marc	Ambrosi	Daniel	Tobias	Mittel
X_i	63.75	64.07	63.67	64.64	64.13	63.67	64.19	64.30	64.05
Y_i	64.63	64.28	64.90	64.56	64.05	63.51	64.75	64.38	64.38
D_i	-0.88	-0.21	-1.23	0.08	0.08	0.16	-0.56	-0.08	-0.33

Die geschätzten Standardabweichungen sind $\hat{\sigma}_x = 0.34$, $\hat{\sigma}_y = 0.44$ und $\hat{\sigma}_{x-y} = 0.51$.

Sie dürfen davon ausgehen, dass die zu Grunde liegenden Zufallsvariablen normalverteilt sind.

Den Trainer interessiert es nun, ob die Wahl des Wachses eine Auswirkung auf die Zeit hat.

- Handelt es sich hier um einen gepaarten oder einen ungepaarten Test? Begründen Sie kurz.
- Geben Sie die Null- und die Alternativhypothese an.
- Führen Sie einen t -Test auf dem 5%-Niveau durch. Geben Sie den Verwerfungsbereich an. Wie entscheidet dieser Test?
- Geben Sie ein zweiseitiges 95%-Vertrauensintervall an für die Differenz $\mu_x - \mu_y$.
- Angenommen die Standardabweichungen wären nicht geschätzt, sondern aus langjähriger Erfahrung bekannt. Würde sich der P-Wert vergrößern oder verkleinern, falls man anstatt eines t -Tests einen z -Test macht? Begründen Sie.

- 2. (10 Punkte)** Das Format eines Fernsehquizzes sieht vor, dass ein Kandidat 16 aufeinanderfolgende (und ganz verschiedene) Fragen gestellt bekommt und bei jeder Frage 4 Antworten zur Auswahl stehen, von denen genau eine richtig ist. Uns interessiert die Anzahl richtig beantworteter Fragen.

Wir nehmen zunächst an, dass der Kandidat die Antworten zu diesen Fragen gar nicht kennt und zufällig antwortet, sowie, dass seine Antworten zu verschiedenen Fragen unabhängig voneinander sind.

- a) Mit welcher Verteilung kann man dann die Anzahl richtiger Antworten beschreiben? (Verteilungsfamilie *und* Parameter angeben)
- b) Welches sind in diesem Fall Erwartungswert und Varianz der Anzahl richtiger Antworten?
- c) Nun wird wegen Ausfalls der nachfolgenden Sendung die Quizshow auf 50 Fragen verlängert (wiederum jede mit 4 Antworten zur Auswahl). Mit welcher stetigen Approximation kann man die Anzahl richtig beantworteter Fragen (aus den 50) nun einfacher modellieren? (Verteilungsfamilie *und* Parameter angeben)

Verwenden Sie bei den Teilaufgaben d) und e) die Approximation aus c).

- d) Der Kandidat beantwortet genau 15 der 50 Fragen richtig. Kann man auf dem 5%-Signifikanzniveau ausschliessen (verwerfen), dass er geraten hat statt tatsächlich etwas zu wissen? Schreiben Sie dazu in Formeln Null- und Alternativhypothese, die Teststatistik und den Verwerfungsbereich auf, und geben Sie die Entscheidung des Tests an.
- e) Nehmen wir nun an, der Kandidat habe in Wahrheit die Erfolgswahrscheinlichkeit 40% bei jeder Aufgabe. Wie gross ist die Wahrscheinlichkeit, dass im vorigen Test (bei Teilaufgabe d)) trotzdem die Nullhypothese beibehalten wird?

3. (10 Punkte)

Markus möchte sich mit Corinne verabreden und ruft sie deshalb mit seinem Natel an. Leider ist die Natel-Verbindung sehr schlecht. Wir nehmen an, dass ein von Markus am Natel gesprochenes Wort entweder richtig übermittelt oder gar nicht übermittelt wird, und dies unabhängig von den anderen gesprochenen Wörtern. Die Wahrscheinlichkeit, dass ein Wort richtig übermittelt wird, sei $3/4$.

a) Markus teilt Corinne folgenden Satz mit: “Wir treffen uns nicht um fünf Uhr am Hauptbahnhof”. Wie gross ist die Wahrscheinlichkeit, dass Corinne den Satz richtig empfängt? Wie gross ist die Wahrscheinlichkeit, dass Corinne fälschlicherweise folgenden Satz empfängt: “Wir treffen uns um fünf Uhr am Hauptbahnhof”?

b) Wie gross ist die Wahrscheinlichkeit, dass vom ursprünglichen Satz mindestens 2 Wörter richtig übermittelt werden?

Hinweis: Betrachten Sie das Komplementär-Ereignis.

c) Durch mehrmaliges Aussprechen von jedem Wort möchte Markus erreichen, dass seine Mitteilung richtig verstanden wird. Wie gross ist die Wahrscheinlichkeit, dass Markus irgendein Wort k -mal aussprechen muss bis es richtig übermittelt wird? Wie gross ist die Wahrscheinlichkeit, dass Markus ein Wort mehr als 3-mal aussprechen muss bis es richtig übermittelt wird?

Hinweis: Markus muss ein Wort genau k -mal aussprechen, wenn die ersten $k-1$ Versuche das Wort zu übermitteln scheitern und der k -te Versuch gelingt.

d) Mit wievielen Wiederholungen eines Wortes bis zur ersten erfolgreichen Übermittlung rechnet Markus im Durchschnitt?

Hinweis: Berechnen Sie den Erwartungswert. Verwenden sie dazu:

$$\sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2} \quad \text{falls} \quad 0 < x < 1$$

e) Markus spricht jedes Wort aus seinem ursprünglichen Satz 4-mal aus. Wie gross ist die Wahrscheinlichkeit, dass Corinne den Satz richtig versteht?

4. (8 Punkte)

Wir betrachten hier Daten, welche den Zusammenhang des Gewichts in Kilogramm (Y), des Alters in Monaten (x_1) und der Grösse in Zentimetern (x_2) von Kleinkindern untersuchen.

An diese Daten wurde folgendes Regressionsmodell angepasst:

$$Y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d.}$$

Hier sehen Sie einen Teil des R-Outputs:

Coefficients:

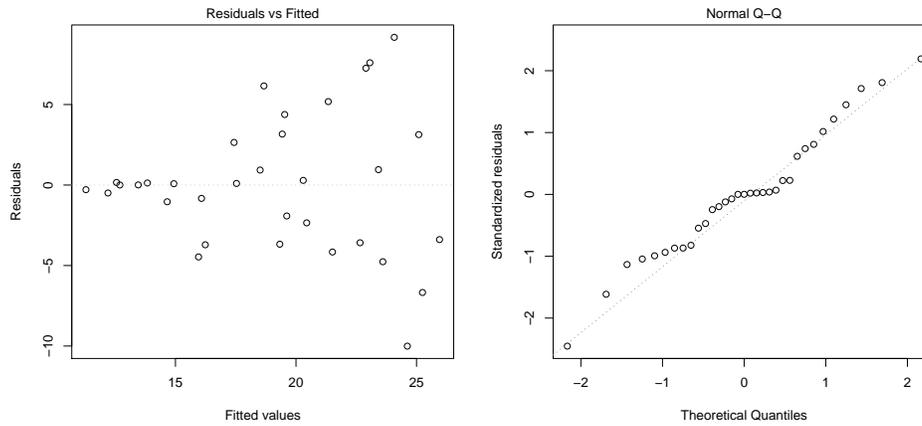
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.645	???	0.049	0.962
Alter	0.070	0.110	???	0.528
Groesse	0.132	???	???	???

Residual standard error: 4.377 on 30 degrees of freedom

Multiple R-squared: 0.5084, Adjusted R-squared: 0.4757

F-statistic: 15.51 on 2 and 30 DF, p-value: 2.364e-05

- 1) Wie gross ist der Wert der t-Teststatistik von $\hat{\beta}_1$ (Alter)?
a) 0.008 b) 0.636 c) 0.058 d) 4.8
- 2) Mit wievielen Beobachtungen wurde die Regression berechnet?
a) 33 b) 32 c) 31 d) 30
- 3) Hat mindestens eine der beiden erklärenden Variablen **Alter** oder **Groesse** einen signifikanten Einfluss auf die Zielgrösse?
a) Ja.
b) Nein.
c) Man müsste zwei einfache Regressionen machen, um dies zu entscheiden.
d) Keine Aussage möglich.
- 4) Wird die Nullhypothese $H_0 : \beta_1 = 0$ auf dem 5%-Niveau verworfen (die Alternative ist $H_A : \beta_1 \neq 0$)? β_1 ist der Koeffizient des Alters im Regressionsmodell.
a) Ja b) Nein c) Keine Aussage möglich
- 5) Wird $H_0 : \beta_2 = 0$ auf dem 5%-Niveau verworfen (die Alternative ist $H_A : \beta_2 \neq 0$)?
a) Ja b) Nein c) Keine Aussage möglich.
- 6) Welches der folgenden Intervalle ist ein exaktes zweiseitiges 95%-Vertrauensintervall für β_1 (Alter)?
a) $0.07 \pm 1.96 \cdot 0.11$ b) $0.07 \pm 1.96 \cdot \frac{0.11}{\sqrt{30}}$
c) $0.07 \pm 2.04 \cdot \frac{0.11}{\sqrt{30}}$ d) $0.07 \pm 2.04 \cdot 0.11$
- 7) Max und sein jüngerer Bruder Moritz unterscheiden sich um genau 20cm. Wie gross ist der erwartete Gewichtsunterschied?
a) Der erwartete Gewichtsunterschied der beiden Kinder ist 3.29kg.
b) Der erwartete Gewichtsunterschied der beiden Kinder ist 2.64kg.
c) Keine Aussage möglich.



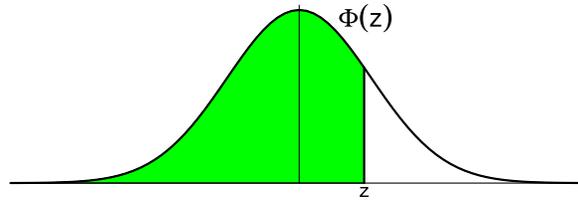
8) Betrachten Sie die gezeigten Plots. Welche der nachfolgenden Aussagen ist zutreffend?

- a) Die Modellannahmen über die Fehler scheinen plausibel.
- b) Die Normalitätsannahme der Fehler ϵ_i scheint grob verletzt zu sein, aber die Annahme der konstanten Varianz scheint plausibel.
- c) Die Annahme der konstanten Varianz der Fehler ϵ_i scheint verletzt zu sein, aber Normalitätsannahme der Fehler ϵ_i scheint plausibel.
- d) Die Normalitätsannahme der Fehler ϵ_i und die Annahme der konstanten Varianz der Fehler ϵ_i scheinen grob verletzt zu sein.

5. (8 Punkte)

- 1) Ein elektrisches Bauteil habe eine Lebensdauer T (in Stunden), die mit einer Exponentialverteilung mit Parameter $\lambda = 1/500$ gut modellierbar ist. Bestimme den Zeitpunkt t , der mit 90% Wahrscheinlichkeit “überlebt” wird.
a) 25.6 b) 1151.3 c) 500
d) 2302.6 e) 52.7 f) Aussagen a) bis e) sind falsch
- 2) Es sei $X \sim \mathcal{N}(\mu, \sigma^2)$ mit $\mu = 32$ und $\sigma^2 = 10$. Ferner sei $Y \sim \text{Poisson}(\lambda)$ mit $\lambda = 18$ und unabhängig von X . Wie gross ist $\mathbf{E}[X - Y]$?
a) 50 b) 41 c) 14
d) 28 e) 23 f) Aussagen a) bis e) sind falsch
- 3) Seien X und Y unabhängige Zufallsvariablen wie in Teilaufgabe 3.): Wie gross ist $\text{Var}(X - \frac{1}{3}Y)$?
a) 8 b) 12 c) 4
d) 16 e) 50 f) Aussagen a) bis e) sind falsch
- 4) In einem Glückspiel wird pro Runde mit Wahrscheinlichkeit 0.5 das Guthaben verdoppelt respektive halbiert. Was ist das erwartete Guthaben nach einer Runde dieses Spiels, wenn mit einem Vermögen von 100 gestartet wird?
a) 100 b) 50 c) 10
d) 125 e) 75 f) Aussagen a) bis e) sind falsch
- 5) Max hat in seinem Sparschwein 500 5-Rappen Münzen angesammelt. Er wirft nun sein Sparschwein auf den Boden, so dass es zerbricht. Wir nehmen an, dass die Münzen unabhängig voneinander Kopf oder Zahl zeigen. Mit welcher Verteilung lässt sich approximativ die Anzahl Münzen modellieren, die Kopf zeigen?
a) Poissonverteilung b) Uniforme Verteilung
c) Normalverteilung d) Exponentialverteilung
- 6) Das 95%-Vertrauensintervall bei einem 1-Stichproben t-Test für den Erwartungswert μ sei $[0.125, 3.375]$. Wird die Nullhypothese $H_0 : \mu = 0.13$ auf dem 1%-Niveau verworfen?
a) Ja b) Nein c) Keine Aussage möglich
- 7) Wenn $X \sim \text{Poisson}(100)$ verteilt ist, so gilt für $Y = \frac{1}{10}X$:
a) $Y \sim \text{Poisson}(1)$ b) $Y \sim \text{Poisson}(10)$
c) $Y \sim \mathcal{N}(1, \frac{1}{10})$ d) Aussagen a) bis c) sind falsch
- 8) Das Signifikanzniveau sorgt bei einem statistischen Test dafür, dass:
a) die Wahrscheinlichkeit eines Fehlers 2. Art kontrolliert wird
b) die Macht kontrolliert wird
c) die Wahrscheinlichkeit eines Fehlers 1. Art kontrolliert wird
d) Aussagen a) bis c) sind falsch

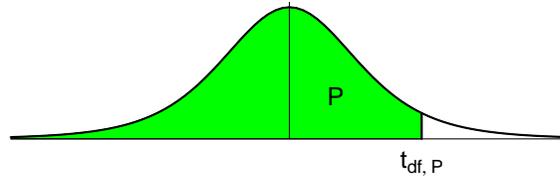
Tabelle der Kumulativen Normalverteilung $\Phi(z) = P[Z \leq z]$, $Z \sim \mathcal{N}(0, 1)$



Bsp.: $P[Z \leq 1.96] = 0.975$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

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