

## Schriftliche Prüfung (90 Minuten)

### Bemerkungen:

- Alle schriftlichen Hilfsmittel und ein Taschenrechner sind erlaubt.
- Mobiltelefone 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.
- Der Lösungsweg muss (ausser bei den Multiple-Choice-Aufgaben) immer ersichtlich sein.
- Bei den Multiple-Choice-Aufgaben 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 zugehörige Antwortblatt ein.
- Die nötigen Tabellen befinden sich auf den hintersten Seiten dieser Prüfung.

**Viel Erfolg!**

1. (10 Punkte) Ein Wissenschaftler behauptet, dass er ein einzelnes Gen gefunden hat, das einen kausalen Einfluss darauf hat, wie lange die Pflanzenart *Arabidopsis* blüht. Um dies zu überprüfen werden ihm fünf Paare von jeweils zwei genetisch identischen Pflanzen (Replikate) zur Verfügung gestellt. Jeweils bei einer der beiden schaltet er das besagte Gen aus ("Gen-Knockout") und vergleicht die Blütezeiten der Pflanzen (in Tagen).

Pflanze	1	2	3	4	5
Replikat 1 ohne Knockout	4.1	4.8	4.0	4.5	4.0
Replikat 2 mit Knockout	3.1	4.3	4.5	3.0	3.5

Nehmen Sie an, dass die Differenzen Blütezeit Replikat 1 minus Blütezeit Replikat 2 normalverteilt mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$  sind.

- Es handelt sich um einen gepaarten Test. Warum?
- Geben Sie die Null- und die Alternativhypothese an und begründen Sie kurz Ihre Wahl.
- Geben Sie eine Schätzung  $\hat{\sigma}^2$  für die Varianz  $\sigma^2$  der Differenz an (mit Lösungsweg).
- Führen Sie den geeigneten  $t$ -Test zum Niveau 0.05 durch: Bestimmen Sie den Wert der Teststatistik  $T$  und dessen Verteilung unter der Nullhypothese, den Verwerfungsbereich für  $T$  und den Testentscheid. (Wenn Sie obige Aufgabe c) nicht lösen konnten, benutzen Sie im Folgenden  $\hat{\sigma}^2 = 0.55$ .)
- Bestimmen Sie ein zweiseitiges 95%-Vertrauensintervall für  $\mu$ .
- Kann der Wissenschaftler das Testergebnis durch mehrfaches Kopieren der Daten verändern, d.h. dadurch, dass er vorgibt insgesamt  $5 \cdot k$  unabhängige Pflanzenpaare untersucht zu haben und jeden Datenpunkt  $k$ -fach notiert?  
(Geben Sie eine kurze Erklärung für Ihre Antwort, in der Sie ohne Beweis erwähnen, was mit dem Mittelwert und der Standardabweichung geschieht. Sie dürfen benutzen, dass  $t_{f;1-\alpha/2}$  mit steigendem  $f$  gegen  $z_{1-\alpha/2}$  konvergiert.)

- 2. (8 Punkte)** Bei der online Musikfirma MeTunes können die Kunden zuerst einen Teil eines Liedes hören und dann entscheiden, ob sie das Lied kaufen wollen. Für ein Lied des Rappers 20Rappen nimmt MeTunes auf Grund langjähriger Erfahrung an, dass etwa (unabhängig von der Anzahl Hörer und unabhängig vom Tag) 40% der täglichen Hörer das Lied schlussendlich kaufen. Sei weiterhin  $X$  die Anzahl der Kunden, die ein Lied des Rappers 20Rappen an einem Tag hören und kaufen.

An einem Tag wurde ein Lied des Rappers 20Rappen von 10 Kunden gehört.

- a) Geben Sie die Verteilung von  $X$  inklusive Parameter. Wie gross ist die Wahrscheinlichkeit, dass das Lied von keinem der ersten fünf Kunden gekauft wurde?
- b) Wie gross ist die Wahrscheinlichkeit, dass das Lied genau von 3 Kunden gekauft wurde? Wie gross ist die Wahrscheinlichkeit, dass das Lied von höchstens 3 Kunden gekauft wurde?

Der Rapper 20Rappen vermutet, dass eine Kollaboration mit der Sängerin Pihanna die Wahrscheinlichkeit, dass dieses neue Lied von einem Hörer gekauft wurde, erhöhen wird. Als Experiment wird das neue Lied von 20Rappen und Pihanna an einem Tag von 10 Leuten gehört, von welchen 6 Leuten es kaufen. Ist die Vermutung des Rappers richtig? Mit obigen Daten soll das nun getestet werden (auf dem Signifikanzniveau 0.05).

- c) Was sind die Null- und die Alternativhypothese?
- d) Führen Sie einen geeigneten exakten Test durch. Geben Sie den Verwerfungsbereich für  $X$  und den Testentscheid an.
- e) Wie gross ist die Macht des Tests bei konkreter Alternative  $\pi = 0.5$ ?
- f) MeTunes ist ebenfalls an dem Experiment interessiert und lädt 250 Kunden ein, das neue Lied von 20Rappen und Pihanna zu hören. Das Lied wurde von 108 Kunden gekauft. Führen Sie erneut einen Hypothesentest durch und berechnen Sie anschliessend das zweiseitige 95%-Intervall für  $\pi$ .

3. (7 Punkte) In Zürich wird untersucht, inwiefern die Anzahl Liter verkauften Glühweins (wine) durch die Aussentemperatur (temp) und den Preis (price) beeinflusst wird. Es wurden 30 Messungen gemacht und folgendes Modell angepasst:

$$\text{wine}_i = \beta_0 + \beta_1 \cdot \text{temp}_i + \beta_2 \cdot \text{price}_i + \varepsilon_i, \quad \varepsilon_i \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2), \quad i = 1, \dots, 30.$$

Der (unvollständige) Regressionsoutput sieht wie folgt aus:

Residuals:

Min	1Q	Median	3Q	Max
-1.6759	-0.1504	-0.0064	0.1704	1.3397

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.0122	???	-0.11	0.91
temp	-1.9600	0.1012	-19.37	<2e-16 ***
price	9.5989	11.0057	???	???

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

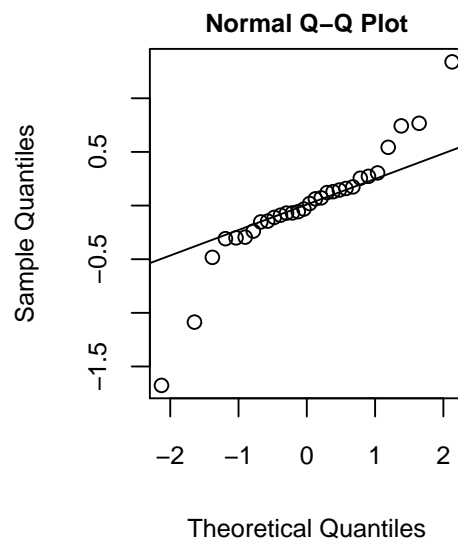
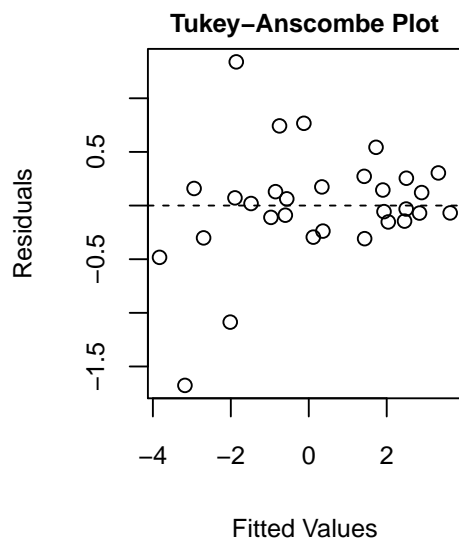
Residual standard error: 0.551 on ?? degrees of freedom

Multiple R-squared: 0.938, Adjusted R-squared: 0.934

F-statistic: 205 on 2 and ?? DF, p-value: <2e-16

- 1) Was ist der Standardfehler von  $\hat{\beta}_0$ ? (auf zwei signifikante Stellen genau)
  - a) -0.013
  - b) 0.33
  - c) 0.11
  - d) 0.87
- 2) Wie gross ist die Schätzung von  $\sigma^2$ ?
  - a) 0.551
  - b) 0.3
  - c) 0.74
  - d) 0.88
- 3) Mit wievielen Freiheitsgraden wurde der "residual standard error" berechnet?
  - a) 28
  - b) 3
  - c) 2
  - d) 27
- 4) Berechnen Sie das exakte zweiseitige 95%-Konfidenzintervall für  $\beta_1$ . (auf zwei signifikante Stellen genau)
  - a) [-2.13, -1.78]
  - b) [-12.98, 32.18]
  - c) [-2.17, 1.75]
  - d) [-2.17, -1.75]
- 5) Wird die Nullhypothese  $H_0 : \beta_2 = 0$  auf dem 1%-Niveau verworfen?

- a) Ja.
  - b) Nein.
  - c) Keine Angabe möglich.
- 6) Eine Person verkauft Glühwein zum Preis von 3 CHF. Wie soll die Aussentemperatur nach obigem Modell sein, wenn sie 18 Liter Glühwein zu verkaufen erwartet?
- a)  $-6.5$
  - b)  $4.3$
  - c)  $5.5$
  - d)  $2.5$
- 7) Betrachten Sie die nachfolgenden Plots. Welche der folgenden Aussagen trifft zu?
- a) Alle Modellannahmen sind erfüllt.
  - b) Die Fehlervarianz ist nicht konstant, aber die Normalverteilungsannahme ist plausibel.
  - c) Die Fehlervarianz ist konstant, aber die Normalverteilungsannahme trifft nicht zu.
  - d) Sowohl konstante Fehlervarianz als auch Normalverteilungsannahme treffen nicht zu.



4. (8 Punkte) Die folgenden Aufgaben sind zufällig angeordnet und insbesondere nicht nach Schwierigkeitsgrad sortiert.

1) Es gilt für ein Ereignis  $A$ , dass  $\text{odds}(A) = 3$ . Was ist die Wahrscheinlichkeit für  $A$ ?

- a)  $1/3$
- b)  $0.5$
- c)  $0.25$
- d)  $0.75$

2) Die geometrische Verteilung beschreibt die Anzahl Versuche  $x$ , die nötig sind, um einen Erfolg zu haben bei unabhängigen Bernoulliversuchen mit Erfolgswahrscheinlichkeit  $p$  (z.B. bei einem Münzwurf). Dies ist eine diskrete Verteilung mit Wahrscheinlichkeitsfunktion

$$P(X = x) = p(1 - p)^{x-1}, \quad x = 1, 2, 3, \dots$$

wobei  $0 < p < 1$  die Erfolgswahrscheinlichkeit ist.

Zwei Personen ziehen unabhängig voneinander an einer Losbude so lange ein Los, bis sie den ersten Treffer haben. Dabei ist  $x_1 = 5$  und  $x_2 = 7$ . Was ist der Maximum Likelihood Schätzer für  $p$ ?

- a)  $0.083$
- b)  $0.833$
- c)  $0.171$
- d)  $0.166$

3) Es gelten für zwei Ereignisse  $A$  und  $B$ , dass  $P(A|B) = 0.2$ ,  $P(A) = 0.5$  und  $P(B) = 0.4$ . Wie gross ist  $P(B|A)$ ?

- a)  $0.25$
- b)  $0.16$
- c)  $0.08$
- d)  $0.04$

4) Für zwei unabhängige Zufallsvariablen  $X$  und  $Y$  gelte  $\text{Var}(X) = 3$  und  $\text{Var}(Y) = 2$ . Welchen Wert hat  $\text{Var}(3X - 2Y + 5)$ ?

- a)  $35$
- b)  $60$
- c)  $17$
- d)  $22$

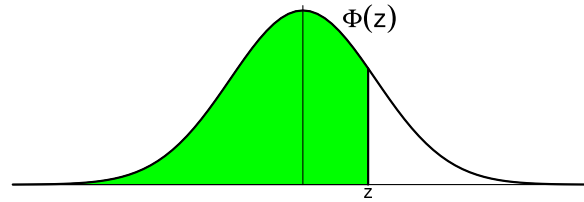
5) Ein Casino kauft 100 neue Spielautomaten. Pro Abend werden ca. 100 Spiele pro Automat gemacht. Sowohl einzelne Spiele als auch verschiedene Automaten sind unabhängig. Pro Spiel und Automat folgt der Gewinn folgender Verteilung:

Gewinn	-1	0	1
Wahrscheinlichkeit	0.3	0.2	0.5

Gib ein approximatives 95% Vertrauensintervall für den Gewinn, der pro Abend durch die neuen Maschinen entsteht, an.

- a) [2829.1, 3170.9]
  - b) [1829.1, 2170.9]
  - c) [1856.6, 2143.4]
  - d) [1998.3, 2001.7]
- 6) Bei fixem  $\alpha$  hat ein einseitiger Test verglichen mit dem entsprechenden zweiseitigen Test
- a) einen kleineren Fehler 1. Art
  - b) ein kleineres Signifikanzniveau
  - c) einen grösseren Fehler 2. Art
  - d) mehr Macht
- 7) Es sei  $X \sim N(5, 3^2)$ . Was ist  $P(X \leq 8)$ ?
- a) 0.54
  - b) 0.63
  - c) 0.84
  - d) 1
- 8) Ein Test mit einer Nullhypothese  $H_0$  ergibt einen  $p$ -Wert von 0.01. Welche der folgenden Aussagen trifft zu:
- a) Die Nullhypothese  $H_0$  ist mit 1% Wahrscheinlichkeit richtig.
  - b) Die Alternative ist mit 99% Wahrscheinlichkeit richtig.
  - c) Die Alternative wird mit 1% Wahrscheinlichkeit fälschlicherweise angenommen.
  - d) Keine der obigen Aussagen ist richtig.

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

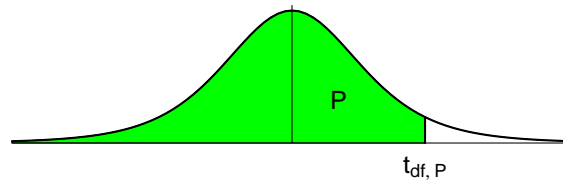


Bsp.:  $P[Z \leq 1.96] = 0.975$

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
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
$\infty$	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576