

2. Vordiplom: Statistik (2 Stunden)

Benötigte Tabellen befinden sich auf den letzten zwei Seiten der Aufgabenstellungen. Wenn nicht anders vermerkt, führen Sie die Tests auf dem 5%-Niveau durch. Geben Sie jeweils Null- und Alternativhypothese, Verwerfungsbereich und Testentscheidung an.

Berechnungen sind jeweils auf 10% genau verlangt.
 Aufgaben 4 und 5 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 erhalte 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 Glück!

1. (5 Punkte) Ein Hersteller möchte untersuchen, wie viele der von ihm verkauften Messgeräte bereits im ersten Jahr ausfallen.

a) Wie gross ist die Wahrscheinlichkeit, dass genau 99 von 100 Geräten im ersten Jahr ausfallen, wenn die Wahrscheinlichkeit für einen Ausfall 10% beträgt? (Wir nehmen natürlich an, dass die Ereignisse "Ausfall von Gerät i " ($i = 1, \dots, 100$) unabhängig voneinander sind.)
 Geben Sie das exakte Resultat an.

b) Wie gross ist die Wahrscheinlichkeit, dass maximal 7 Geräte von 100 im ersten Jahr ausfallen, wenn die Wahrscheinlichkeit für einen Ausfall 10% beträgt? Approximieren Sie die Verteilung der Anzahl Geräteausfälle mit einer geeigneten Normalverteilung und berechnen Sie die gesuchte Wahrscheinlichkeit mit Hilfe dieser Approximation.
 c) Um zu beweisen, dass tatsächlich weniger als 10% der Geräte im ersten Jahr ausfallen, beträgt der Hersteller 100 zufällig ausgewählte Kunden und stellt fest, dass das Gerät nur bei 7 Kunden im ersten Jahr ausfiel. Kann der Hersteller seine Aussage statistisch beweisen?
 (Statt dem Verwerfungsbereich können Sie auch den P-Wert angeben.)

2. (7 Punkte) Die unten angegebenen Daten zeigen die Grösse (in cm) von 18 unabhängigen Messungen an Pflanzen, die entweder mit dem herkömmlichen Dünger oder mit einem neuen Dünger behandelt wurden. (Aus Gründen der Übersichtlichkeit wurden die Daten in jeder Gruppe sortiert.)

Daten:	alter Dünger	36	38	39	40	43	51	52	53	92
	neuer Dünger	43	43	49	53	59	63	69	81	83
Kennzahlen:	alter Dünger	$\bar{x} = 49.33, s_x^2 = \hat{\sigma}_x^2 = 298$								
	neuer Dünger	$\bar{y} = 60.33, s_y^2 = \hat{\sigma}_y^2 = 226$								

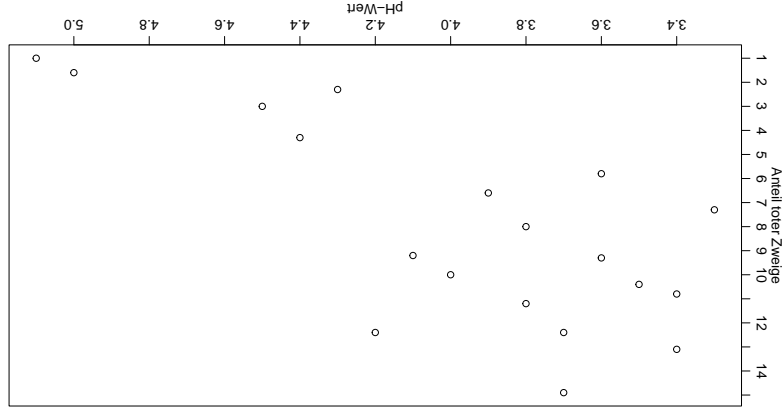
a) Geben Sie ein exaktes (zweiseitiges) 95%-Vertrauensintervall für die erwartete Grösse in der Gruppe mit dem neuen Dünger an. (Nehmen Sie an, die Daten seien normalverteilt.)
 b) Lässt sich zeigen, dass der neue Dünger zu grösseren Pflanzen führt? Führen Sie einen geeigneten t -Test auf dem 5%-Niveau durch, um diese Frage zu beantworten.
 c) Um die Frage aus b) zu beantworten wurde ebenfalls ein Wilcoxon-Test durchgeführt. Der P-Wert war 0.035. Welche Antwort gibt der Wilcoxon-Test auf die Frage in b)?
 d) Welcher der beiden Tests aus b) und c) ist zu bevorzugen? Begründen Sie kurz, aber konkret anhand des oben angegebenen Datensatzes.

3. (7 Punkte) Eine Rohrleitung besteht aus 5 Segmenten. Da die Ausflussmenge kleiner ist als die Zufuhsmenge, muss irgendwo ein Leck bestehen. Wir nehmen an, dass es genau ein Leck gibt. Die Wahrscheinlichkeit, dass sich das Leck in Segment i befindet, sei für alle Segmente gleich, nämlich $1/5$. Ziel ist es, das defekte Segment mittels Inspektionen der Durchflussmenge bei den 4 Zwischenstücken zwischen den Segmenten ausfindig zu machen. (Beispiel: Wenn die Durchflussmenge beim Zwischenstück zwischen dem 2. und 3. Segment kleiner ist als die Zufuhsmenge, so ist entweder das 1. oder das 2. Segment defekt.)
 Im folgenden werden zwei Strategien zum Auffinden des Lecks verglichen.

a) Es wird sukzessive bei allen 4 Zwischenstücken die Durchflussmenge gemessen, bis man das Leck eindeutig identifizieren kann. Sei X die Anzahl benötigter Inspektionen. Mit welchen Wahrscheinlichkeiten nimmt X die Werte $1, \dots, 4$ an? Berechnen Sie zudem $E[X]$ und $\text{Var}(X)$.
 b) Eine etwas raffiniere Strategie als die in a) beschriebene besteht darin, die Inspektion jeweils in der "Mitte" des Gebiets durchzuführen, in dem sich das Leck noch befinden kann, wobei man das Gebiet in jedem Schritt so gut wie möglich "halbiert". (Man misst also zuerst die Durchflussmenge direkt nach dem 2. Segment, und untersucht dann je nach Resultat entweder die ersten 2 oder die letzten 3 Segmente analog.)
 Es sei Y die Anzahl mit dieser Strategie benötigten Untersuchungen. Welche Werte nimmt Y mit welcher Wahrscheinlichkeit an? Berechnen sie zudem $E[Y]$ und $\text{Var}(Y)$.
 c) Wieviele Inspektionen spart Strategie b) "im Schnitt" gegenüber Strategie a)?

4. (5 Punkte) Der untenstehende Scatterplot stellt 19 Messungen an Zuckerbüschen

von x_i = pH-Wert des Bodens und y_i = Anteil toter Zweige in der Krone (in %)



graphisch dar.

An diese Daten wurde ein lineares Regressionsmodell der Form $Y_i = \beta_0 + \beta_1 x_i + E_i$ mit $E_i \sim \mathcal{N}(0, \sigma^2)$ i.i.d. angepasst (Y_i sind die Anteile toter Zweige, x_i die pH-Werte).

Hier sehen Sie einen Teil des R-Outputs:

Coefficients:	
Estimate	Std. Error
(Intercept)	31.04
pH	-5.79
...	1.36
...	...

Multiple R-squared: 0.5151, Adjusted R-squared: 0.4866

1) Wie gross ist ungefähr die Korrelation ρ zwischen x_i und y_i ?

- a) $-1.0 \leq \rho \leq -0.8$
- b) $-0.8 < \rho \leq -0.3$
- c) $-0.3 < \rho \leq +0.3$
- d) $+0.3 < \rho \leq +0.8$
- e) $+0.8 < \rho \leq +1.0$

2) Wie entscheidet der Test der Nullhypothese $H_0: \beta_1 = 0$ gegen $H_A: \beta_1 \neq 0$?

- a) H_0 beibehalten
- b) H_0 verwerfen

3) Welches der folgenden Intervalle ist ein exaktes 95%-Vertrauensintervall für β_0 ?

- a) $31.04 \pm 2.11 \cdot 5.45$
- b) $31.04 \pm 2.09 \cdot 5.45$
- c) $31.04 \pm \frac{2.09}{\sqrt{19}} \cdot 5.45$
- d) keines der Intervalle

4) Eine Messung hat die Koordinaten $(4.0, 10.0)$. Wie gross sind der angepasste Wert und das Residuum dieser Beobachtung (auf 10% genau)?

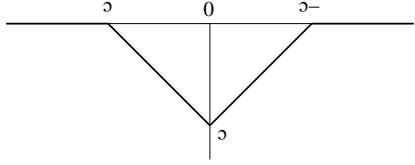
- a) 4.0; -2.1
- b) 7.9; 10.0
- c) 7.9; 2.1
- d) 4.0; 6.0
- e) 10.5; -0.5

5) Wie gross ist die Schätzung $\hat{\sigma}$ von σ ungefähr?

- a) $0 < \hat{\sigma} \leq 1$
- b) $1 < \hat{\sigma} \leq 6$
- c) $6 < \hat{\sigma} \leq 12$
- d) grösser als 12

5. (5 Punkte)

1) Betrachten Sie die folgende Dichtefunktion:



Der Wert von c beträgt

- a) 1/2
- b) 1
- c) 3/2
- d) 2

2) Die Verteilungsfunktion einer Zufallsvariablen X ist gegeben durch

$$F(x) = \begin{cases} 0 & \text{für } x < 0 \\ x_{10} & \text{für } 0 \leq x \leq 1 \\ 1 & \text{für } x > 1 \end{cases}$$

Wie gross ist dann $E[X]$?

- a) 1/13
- b) 1/12
- c) 2/3
- d) 10/12
- e) 10/11

3) Für eine Zufallsvariable Y sei $E[Y] = 2$ und $\text{Var}(Y) = 10$. Wie gross ist dann $E[3Y + Y^2]$?

- a) 8
- b) 10
- c) 16
- d) 20
- e) 24

4) Die Wetterprognose eines Lokalsenders sagt schönes Wetter mit einer Wahrscheinlichkeit von 0.8 korrekt an (bedingte Wahrscheinlichkeit, gegeben dass das Wetter tatsächlich schön wird). Mit Wahrscheinlichkeit 0.2 sagt sie in diesem Fall fälschlicherweise schlechtes Wetter voraus.

Schlechtes Wetter sagt die Prognose mit einer Wahrscheinlichkeit von 0.6 korrekt voraus.

Die Wahrscheinlichkeit von schönem Wetter sei 0.7, diejenige von schlechtem 0.3.

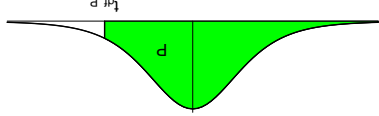
Wie gross ist die Wahrscheinlichkeit für eine korrekte Vorhersage?

- a) 0.64
- b) 0.70
- c) 0.74
- d) 0.78
- e) 0.83

5) Gleiche Ausgangslage wie in Teilaufgabe 4. Wie gross ist die bedingte Wahrscheinlichkeit, dass das Wetter tatsächlich schön wird, wenn die Wetterprognose dies behauptet?

- a) 56/83
- b) 8/10
- c) 56/68
- d) 24/26

Perzentile der t-Verteilung



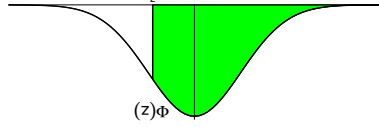
Bsp.: $t_{0.975} = 2.262$

df	$t_{0.60}$	$t_{0.70}$	$t_{0.80}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326

\sqrt{x}	\sqrt{z}	x
1	1.41	1
2	1.73	4
3	1.73	9
4	2.00	16
5	2.24	25
6	2.45	36
7	2.65	49
8	2.83	64
9	3.00	81
10	3.16	100
11	3.32	121
12	3.46	144
13	3.61	169
14	3.74	196
15	3.87	225
16	4.00	256
17	4.12	289
18	4.24	324
19	4.36	361
20	4.47	400
25	5.48	625
30	5.92	900
35	6.32	1225
40	6.32	1600
45	6.71	2025
50	7.07	2500
55	7.42	3025
60	7.75	3600
65	8.07	4225
70	8.37	4900
75	8.64	5625
80	8.94	6400
85	9.19	7225
90	9.49	8100
100	10.00	10000

Tabelle der Quadratwurzeln und Quadratzahlen

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



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

z	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	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	0.8868	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9047	0.9066	0.9082	0.9098	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	0.9663	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	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974	0.9974
2.8	0.9975	0.9976	0.9977	0.9977	0.9977	0.9978	0.9979	0.9980	0.9981	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	0.9988	0.9989	0.9989	0.9989	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	0.9993	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

Antwortblatt für die Multiple Choice Aufgaben:

Kreuzen Sie bitte die zutreffenden Kästchen deutlich an.

Es ist jeweils nur eine Antwort korrekt.

Aufgabe 4.: a) 1) 2) 3) 4) 5)

b) 1) 2) 3) 4) 5)

c) 1) 2) 3) 4) 5)

d) 1) 2) 3) 4) 5)

e) 1) 2) 3) 4) 5)

Aufgabe 5.: a) 1) 2) 3) 4) 5)

b) 1) 2) 3) 4) 5)

c) 1) 2) 3) 4) 5)

d) 1) 2) 3) 4) 5)

e) 1) 2) 3) 4) 5)