Prof. Dr. A. Steland
N. Mause

RWTH Aachen, SS2023

Ausgabe: 04. Juli 2023 — Besprechung: 10. Juli 2023

Einführung in die angewandte Stochastik

Übungsblatt 10

Quantilstabellen mit den Werten zur Normalverteilung, der t-Verteilung und der χ^2 - Verteilung befinden sich am Ende dieses Übungsblattes.

Aufgabe 41

Auf 12 Versuchsflächen wurde eine neue Weizensorte angebaut. Die einzelnen Flächen erbrachten die folgenden Hektarerträge (in t):

$$3.56\ ,\ 3.37\ ,\ 3.78\ ,\ 3.12\ ,\ 3.72\ ,\ 3.41\ ,\ 3.56\ ,\ 3.66\ ,\ 3.71\ ,\ 3.49\ ,\ 3.56\ ,\ 3.40$$

Aus Erfahrung ist bekannt, dass diese Hektarerträge als Realisationen stochastisch unabhängiger, jeweils $N(\mu, \sigma^2)$-verteilter Zufallsvariablen mit $\mu \in \mathbb{R}$ und $\sigma > 0$ angesehen werden können.

(a) Ermitteln Sie ein zweiseitiges Konfidenzintervall für den Erwartungswert μ zum Konfidenzniveau 0.9 bei $1-\alpha = 0{,}9 \Leftrightarrow \alpha = 0{,}1$

(i) bekannter Varianz $\sigma^2 = 0.0324$,

(ii) unbekannter Varianz σ^2.

(b) Ermitteln Sie ein einseitiges unteres 99%–Konfidenzintervall für $1-\alpha = 0{,}99 \Leftrightarrow \alpha = 0{,}01$

(i) die Varianz σ^2,

(ii) die Standardabweichung σ.

Lösung: a) Zweiseitiges KI für μ zum Konf.niveau $1-\alpha = 0{,}9$ bei bekannter Varianz ist gegeben durch

$$KI_1 = \left[\bar{X} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} , \bar{X} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

$z_{1-\frac{\alpha}{2}}$: $(1-\frac{\alpha}{2})$-Quantil der $\mathcal{N}(0,1)$-Verteilung

Hier: $n = 12$, $1-\alpha = 0{,}9 \Leftrightarrow \alpha = 0{,}1 \Leftrightarrow 1-\frac{\alpha}{2} = 0{,}95$

$z_{1-\frac{\alpha}{2}} = z_{0{,}95} \overset{\text{Tabelle}}{\approx} 1{,}64$

$\sigma^2 = 0{,}0324 \Rightarrow \sigma = \sqrt{\sigma^2} = 0{,}18$, $\bar{X} = \frac{1}{12}\sum_{i=1}^{12} x_i \approx 3{,}528$

$$\Rightarrow KI_1 \approx \left[3{,}528 - 1{,}64 \cdot \frac{0{,}18}{\sqrt{12}} , 3{,}528 + 1{,}64 \cdot \frac{0{,}18}{\sqrt{12}}\right]$$

$$\approx [3{,}443 , 3{,}613]$$

(II) Varianz unbekannt

$$\hookrightarrow KI_2 = \left[\overline{X} - t_{1-\frac{\alpha}{2}}(n-1)\frac{\hat{\sigma}}{\sqrt{n}} , \overline{X} + t_{1-\frac{\alpha}{2}}(n-1)\frac{\hat{\sigma}}{\sqrt{n}}\right]$$

Es gilt $\hat{\sigma} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n \cdot \overline{X}^2\right)$

ist die erwartungstreue Punktschätzung für σ^2.

Hier ist $n = 12 \Rightarrow n-1 = 11$

$1 - \frac{\alpha}{2} = 0{,}95 \Rightarrow t_{1-\frac{\alpha}{2}}(n-1) = t_{0{,}95}(11) \overset{\text{Tabelle}}{\approx} 1{,}796$

$\overline{X} \approx 3{,}528$, $\hat{\sigma} = \sqrt{\frac{1}{11}\sum_{i=1}^{12}(X_i - \overline{X})^2} \approx 0{,}185$

$$\Rightarrow KI_2 = \left[3{,}528 - 1{,}796 \cdot \frac{0{,}185}{\sqrt{12}} , 3{,}528 + 1{,}796 \cdot \frac{0{,}185}{\sqrt{12}}\right]$$

$$\approx [3{,}432 , 3{,}623]$$

b) (I) Einseitiges unteres KI für σ^2 bei unbekanntem μ zum Konf.niveau $1-\alpha$

$$KI_3 = \left[0 , \frac{n-1}{\chi^2_\alpha(n-1)} \cdot \hat{\sigma}^2\right]$$

$\hat{\sigma}^2 \approx 0{,}0342$ (siehe oben)

$n = 12$

$1 - \alpha = 0{,}99 \Rightarrow \alpha = 0{,}01 \Rightarrow \chi^2_\alpha(n-1) = \chi^2_{0{,}01}(11) \approx 3{,}05$

$$\Rightarrow KI_3 \approx [0 , 0{,}123]$$

(II) Ziehe die Wurzel

$$KI_4 = \left[\sqrt{0} , \sqrt{\frac{n-1}{\chi^2_\alpha(n-1)} \cdot \hat{\sigma}^2}\right] \approx [0 , \sqrt{0.123}]$$

$$\approx [0 , 0.351]$$

Aufgabe 42

Ein Mieter leistet monatliche Vorauszahlungen für die jährlichen Heizkosten seiner Wohnung an den Vermieter, der zur Festlegung der Abschläge einen Jahresverbrauch von 2 000 l Heizöl zugrunde legt. Der Mieter hält diese Angabe für überhöht und verweist auf die Verbrauchswerte der vergangenen sieben Jahre (in l):

1879, 2037, 1630, 2267, 1773, 2070, 1818

Die Streitfrage soll mittels einen geeigneten statistischen Tests entschieden werden. Nehmen Sie hierzu an, dass die angegebenen Verbrauchswerte als Realisationen stochastisch unabhängiger, jeweils $\mathcal{N}(\mu, \sigma^2)$ - verteilter Zufallsvariablen mit $\mu \in \mathbb{R}$ und (unbekanntem) $\sigma > 0$ angesehen werden können.

(a) Formulieren Sie die zugehörigen Hypothesen für den Fall, dass

(i) die Beweislast beim Mieter liegt,

(ii) Die Beweislast beim Vermieter liegt,

(iii) ein unabhängiger Sachverständiger entscheiden soll, ob der erwartete Jahresverbrauch von 2000 l abweicht.

(b) Geben Sie einen aus der Vorlesung bekannten statistischen Test an, mit dessen Hilfe die Streitfrage entschieden werden kann.

(c) Prüfen Sie jeweils mittels des in (b) angegebenen statistischen Tests, ob die Behauptungen des Mieters bzw. des Vermieters mit den gegebenen Daten zum Signifikanzniveau $\alpha = 5\%$ statistisch nachgewiesen werden können. Zu welcher Entscheidung gelangt der unabhängige Sachverständige?

Lösung: Test-Verfahren

1) Hypothesen aufstellen
(Das, was wir statistisch nachweisen möchten, muss immer als Alternative H_1 formuliert werden!)

2) Welche Test? (für μ)
- Gauß-Test bei bekannter Varianz
- t-Test bei unbekannter Varianz

3) Teststatistik und kritischen Wert berechnen

4) Vergleiche Teststatistik und kritischen Wert
↳ Testentscheidung: → Lehne H_0 ab
→ Lehne H_0 nicht ab

---

zur Aufgabe a) (i) Mieter will zeigen, dass der Verbrauch kleiner als 2000 ist.

Testproblem: $H_0: \mu \geq \mu_0 = 2000$ gegen $H_1: \mu < \underbrace{\mu_0}_{\text{Sollwert}} = 2000$

(ii)

$H_0: \mu \leq \mu_0$ gegen $H_1: \mu > \mu_0 = 2000$

(iii) $H_0: \mu = \mu_0$ gegen $H_1: \mu \neq \mu_0 = 2000$

b) Hier ist die Varianz <u>unbekannt</u>

$\Rightarrow$ Wir verwenden ein t-Test

(in (i) und (ii) einseitig, in (iii) zweiseitig)

c) Teststatistik: $T = \sqrt{n}\,\dfrac{\overline{X} - \mu_0}{S}$, $S = \sqrt{\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$

Hier ist

$n = 7$, $\mu_0 = 2000$, $\alpha = 0{,}05$

$\overline{x} = \frac{1}{7}\sum_{i=1}^{7} x_i \approx 1924.857$

$s \approx 213{,}786$

$\Rightarrow$ beobachtete Teststatistik

$$T \approx \sqrt{7}\,\frac{1924{,}857 - 2000}{213{,}786} \approx -0{,}930$$

Weiterhin ist

$$t_{1-\alpha}(n-1) = t_{0{,}95}(6) = 1{,}943$$

$$t_{1-\frac{\alpha}{2}}(n-1) = t_{0{,}975}(6) = 2{,}447$$

Nun gilt:

(i) $T \approx -0{,}930 \geq -1{,}943 = -t_{0{,}95}(6) = -t_{1-\alpha}(n-1)$

$\Rightarrow$ Nullhypothese $H_0: \mu \geq 2000$ wird <u>nicht</u> abgelehnt

(ii) $T \approx -0{,}930 \leq 1{,}943 = t_{0{,}95}(6) = t_{1-\alpha}(n-1)$

$\Rightarrow$ Nullhypothese $H_0: \mu \leq 2000$ wird <u>nicht</u> abgelehnt

(iii) $|T| \approx 0{,}930 \leq 2{,}447 = t_{0{,}975}(6) = t_{1-\frac{\alpha}{2}}(n-1)$

$\Rightarrow H_0: \mu = 2000$ wird <u>nicht</u> abgelehnt

Aufgabe 43

In einer Supermarktfiliale werden innerhalb eines festgelegten Zeitraumes jeweils die Angebotspreise $x_1, \dots, x_7$ und die zugehörigen Absatzmengen $y_1, \dots, y_7$ eines bestimmten Artikels notiert:

i	1	2	3	4	5	6	7
Preis x_i (in €/kg)	2.49	2.68	2.62	2.51	2.84	2.65	2.76
Absatzmenge y_i (in 100 kg)	6.7	6.2	5.8	6.2	5.4	6.5	5.9

Es sei das übliche Modell der linearen Regression zugrunde gelegt, d.h.

$$Y_i = a + bX_i + \epsilon_i, \quad i = 1, \dots, 7.$$

Hierbei seien die Fehler $\epsilon_1, \dots, \epsilon_i$ stochastisch unabhängig, jeweils $\mathcal{N}(0, \sigma^2)$ - verteilt mit unbekannter Varianz $\sigma^2 > 0$.

Bestimmen Sie die zugehörige geschätzte Regressionsgerade

$$\hat{f}(x) = \hat{a} + \hat{b}x, \quad x \in I,$$

und geben Sie deren Definitionsbereich I explizit an. Geben Sie weiter eine Schätzung für die Absatzmenge bei einem Angebotspreis von 2,70 €/kg an.

Lösung: Es gilt

$$\hat{b} = \frac{s_{xy}}{s_x^2} = \frac{cov(x,y)}{Var(x)} \quad \text{und} \quad \hat{a} = \bar{y} - \hat{b}\bar{x}$$

Hier haben wir

$$\bar{x} = \frac{1}{7}\sum_{i=1}^{7} x_i = 2{,}65$$

$$\bar{y} = \frac{1}{7}\sum_{i=1}^{7} y_i = 6{,}1$$

Weiterhin haben wir

$$\sum_{i=1}^{7} x_i y_i = 112{,}902$$

$$\sum_{i=1}^{7} x_i^2 = 49{,}2527$$

$$\Rightarrow \hat{b} = \frac{s_{xy}}{s_x^2} = \frac{cov(x,y)}{Var(x)} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

n

$$= \frac{\sum_{i=1} x_i \cdot y_i - n \cdot \bar{x} \cdot \bar{y}}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2}$$

$$= \frac{112.902 - 7 \cdot 2{,}65 \cdot 6{,}1}{49.2527 - 7 \cdot 2{,}65^2} \approx -2{,}658$$

$$\Rightarrow \hat{a} = \bar{y} - \hat{b}\bar{x} \approx 6{,}1 - (-2{,}658) \cdot 2{,}65$$

$$\approx 13.144$$

$\Rightarrow$ geschätzte Regressionsgerade

$$\hat{f}(x) = \hat{a} + \hat{b}x \approx 13{,}144 - 2.658 \cdot x,$$

$$x \in I := [x_{min}, x_{max}] = [2{,}49,\ 2{,}84]$$

Diese Funktion beschreibt die Absatzmenge y in Abhängigkeit von den Preis x.

Speziell für $x = 2{,}7$ gilt

$$\hat{f}(2{,}7) = 13{,}144 - 2{,}658 \cdot 2{,}7 = 5.9674$$

Standardnormalverteilung

Verteilungsfunktion $\Phi(x+h)$										
x	h									
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981

Beispiel: $X \sim \mathcal{N}(3{,}9)$,

$$P(X \leq 4.26) = P(\frac{X-3}{\sqrt{9}} \leq \frac{4.26-3}{3}) = P(X \leq 0.42) = 0.6628$$

t-Verteilung

q-Quantile der $t(df)$-Verteilung						
			q			
df	0.9	0.95	0.975	0.98	0.99	0.995
1	3.078	6.314	12.706	15.895	31.821	63.657
2	1.886	2.920	4.303	4.849	6.965	9.925
3	1.638	2.353	3.182	3.482	4.541	5.841
4	1.533	2.132	2.776	2.999	3.747	4.604
5	1.476	2.015	2.571	2.757	3.365	4.032
6	1.440	1.943	2.447	2.612	3.143	3.707
7	1.415	1.895	2.365	2.517	2.998	3.499
8	1.397	1.860	2.306	2.449	2.896	3.355
9	1.383	1.833	2.262	2.398	2.821	3.250
10	1.372	1.812	2.228	2.359	2.764	3.169
11	1.363	1.796	2.201	2.328	2.718	3.106
12	1.356	1.782	2.179	2.303	2.681	3.055
13	1.350	1.771	2.160	2.282	2.650	3.012
14	1.345	1.761	2.145	2.264	2.624	2.977
15	1.341	1.753	2.131	2.249	2.602	2.947
16	1.337	1.746	2.120	2.235	2.583	2.921
17	1.333	1.740	2.110	2.224	2.567	2.898
18	1.330	1.734	2.101	2.214	2.552	2.878
19	1.328	1.729	2.093	2.205	2.539	2.861
20	1.325	1.725	2.086	2.197	2.528	2.845
21	1.323	1.721	2.080	2.189	2.518	2.831
22	1.321	1.717	2.074	2.183	2.508	2.819
23	1.319	1.714	2.069	2.177	2.500	2.807
24	1.318	1.711	2.064	2.172	2.492	2.797
25	1.316	1.708	2.060	2.167	2.485	2.787
26	1.315	1.706	2.056	2.162	2.479	2.779
27	1.314	1.703	2.052	2.158	2.473	2.771
28	1.313	1.701	2.048	2.154	2.467	2.763
29	1.311	1.699	2.045	2.150	2.462	2.756
30	1.310	1.697	2.042	2.147	2.457	2.750
31	1.309	1.696	2.040	2.144	2.453	2.744
32	1.309	1.694	2.037	2.141	2.449	2.738

Beispiel: $X \sim t(8)$,

$$P(X \leq c) = 0.95 \Rightarrow c = 1.860$$

χ^2 - Verteilung

FG	0,01	0,025	0,05	0,1	0,5	0,9	0,95	0,975	0,99
1	0,0002	0,0010	0,0039	0,0158	0,4549	2,7055	3,8415	5,0239	6,6349
2	0,0201	0,0506	0,1026	0,2107	1,3863	4,6052	5,9915	7,3778	9,2103
3	0,1148	0,2158	0,3518	0,5844	2,3660	6,2514	7,8147	9,3484	11,345
4	0,2971	0,4844	0,7107	1,0636	3,3567	7,7794	9,4877	11,143	13,277
5	0,5543	0,8312	1,1455	1,6103	4,3515	9,2364	11,070	12,833	15,086
6	0,8721	1,2373	1,6354	2,2041	5,3481	10,645	12,592	14,449	16,812
7	1,2390	1,6899	2,1674	2,8331	6,3458	12,017	14,067	16,013	18,475
8	1,6465	2,1797	2,7326	3,4895	7,3441	13,362	15,507	17,535	20,090
9	2,0879	2,7004	3,3251	4,1682	8,3428	14,684	16,919	19,023	21,666
10	2,5582	3,2470	3,9403	4,8652	9,3418	15,987	18,307	20,483	23,209
11	3,0535	3,8157	4,5748	5,5778	10,341	17,275	19,675	21,920	24,725
12	3,5706	4,4038	5,2260	6,3038	11,340	18,549	21,026	23,337	26,217
13	4,1069	5,0088	5,8919	7,0415	12,340	19,812	22,362	24,736	27,688
14	4,6604	5,6287	6,5706	7,7895	13,339	21,064	23,685	26,119	29,141
15	5,2293	6,2621	7,2609	8,5468	14,339	22,307	24,996	27,488	30,578
16	5,8122	6,9077	7,9616	9,3122	15,338	23,542	26,296	28,845	32,000
17	6,4078	7,5642	8,6718	10,085	16,338	24,769	27,587	30,191	33,409
18	7,0149	8,2307	9,3905	10,865	17,338	25,989	28,869	31,526	34,805
19	7,6327	8,9065	10,117	11,651	18,338	27,204	30,144	32,852	36,191
20	8,2604	9,5908	10,851	12,443	19,337	28,412	31,410	34,170	37,566
21	8,8972	10,283	11,591	13,240	20,337	29,615	32,671	35,479	38,932
22	9,5425	10,982	12,338	14,041	21,337	30,813	33,924	36,781	40,289
23	10,196	11,689	13,091	14,848	22,337	32,007	35,172	38,076	41,638
24	10,856	12,401	13,848	15,659	23,337	33,196	36,415	39,364	42,980
25	11,524	13,120	14,611	16,473	24,337	34,382	37,652	40,646	44,314
26	12,198	13,844	15,379	17,292	25,336	35,563	38,885	41,923	45,642
27	12,879	14,573	16,151	18,114	26,336	36,741	40,113	43,195	46,963
28	13,565	15,308	16,928	18,939	27,336	37,916	41,337	44,461	48,278
29	14,256	16,047	17,708	19,768	28,336	39,087	42,557	45,722	49,588
30	14,953	16,791	18,493	20,599	29,336	40,256	43,773	46,979	50,892

χ^2 - Verteilung

q-Quantile der $\chi^2(df)$-Verteilung						
			q			
df	0.9	0.95	0.975	0.98	0.99	0.995
36	47.212	50.998	54.437	55.489	58.619	61.581
37	48.363	52.192	55.668	56.730	59.893	62.883
38	49.513	53.384	56.896	57.969	61.162	64.181
39	50.660	54.572	58.120	59.204	62.428	65.476
40	51.805	55.758	59.342	60.436	63.691	66.766
41	52.949	56.942	60.561	61.665	64.950	68.053
42	54.090	58.124	61.777	62.892	66.206	69.336
43	55.230	59.304	62.990	64.116	67.459	70.616
44	56.369	60.481	64.201	65.337	68.710	71.893
45	57.505	61.656	65.410	66.555	69.957	73.166
46	58.641	62.830	66.617	67.771	71.201	74.437
47	59.774	64.001	67.821	68.985	72.443	75.704
48	60.907	65.171	69.023	70.197	73.683	76.969
49	62.038	66.339	70.222	71.406	74.919	78.231
50	63.167	67.505	71.420	72.613	76.154	79.490
51	64.295	68.669	72.616	73.818	77.386	80.747
52	65.422	69.832	73.810	75.021	78.616	82.001
53	66.548	70.993	75.002	76.223	79.843	83.253
54	67.673	72.153	76.192	77.422	81.069	84.502
55	68.796	73.311	77.380	78.619	82.292	85.749
56	69.919	74.468	78.567	79.815	83.513	86.994
57	71.040	75.624	79.752	81.009	84.733	88.236
58	72.160	76.778	80.936	82.201	85.950	89.477
59	73.279	77.931	82.117	83.391	87.166	90.715
60	74.397	79.082	83.298	84.580	88.379	91.952
61	75.514	80.232	84.476	85.767	89.591	93.186
62	76.630	81.381	85.654	86.953	90.802	94.419
63	77.745	82.529	86.830	88.137	92.010	95.649
64	78.860	83.675	88.004	89.320	93.217	96.878
65	79.973	84.821	89.177	90.501	94.422	98.105
66	81.085	85.965	90.349	91.681	95.626	99.330
67	82.197	87.108	91.519	92.860	96.828	100.554
68	83.308	88.250	92.689	94.037	98.028	101.776
69	84.418	89.391	93.856	95.213	99.228	102.996
70	85.527	90.531	95.023	96.388	100.425	104.215