UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Einführung in die Wahrscheinlichkeitstheorie und Statistik
Prof. Dr. Rainer Dahlhaus
Nathawut Phandoidaen
Wintersemester 2019/2020

8. Übungsblatt – Lösungen

Aufgabe 29	Aufgabe 30	Aufgabe 31	Summe:

Übungsgruppe: Tutor(in):

Namen:

Aufgabe 29 (Stichprobenmittelwert, Stichprobenvarianz und empirische Verteilungsfunktion, 4 = 1 + 1.5 + 1.5 Punkte).

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, $n \in \mathbb{N}$ und $X_1, ..., X_n : \Omega \to \mathbb{R}$ unabhängig und identisch verteilte Zufallsvariablen mit $\mathbb{E}[X_1] = \mu$, $\mathrm{Var}(X_1) = \sigma^2$ und Verteilungsfunktion F. Seien für $t \in \mathbb{R}$,

$$\overline{X_n} := \frac{1}{n}\sum_{i=1}^n X_i, \qquad S_n^2 := \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X_n})^2, \qquad \hat{F}_n(t) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{X_i \le t\}}$$

der Stichprobenmittelwert, die Stichprobenvarianz und die empirische Verteilungsfunktion.

(a) Berechnen Sie $\mathbb{E}[\overline{X_n}]$ und $\mathrm{Var}(\overline{X_n})$ in Termen von n, μ, σ^2.

(b) Berechnen Sie $\mathbb{E}[S_n^2]$ in Termen von n, μ, σ^2.
Hinweis: Zeigen Sie zunächst, dass $S_n^2 = \frac{n}{n-1}\left(\left(\frac{1}{n}\sum_{i=1}^n X_i^2\right) - \overline{X_n}^2\right)$ *gilt.*

(c) Berechnen Sie die Verteilung von $n \cdot \hat{F}_n(t)$ und geben Sie dann $\mathbb{E}[\hat{F}_n(t)]$ und $\mathrm{Var}(\hat{F}_n(t))$ in Termen von $n, F(t)$ an.

Lösung:

Wir nutzen die folgenden bekannten Regeln aus der Vorlesung:

- Der Erwartungswert ist linear.
- Für die Varianz gilt $\mathrm{Var}(a \cdot X) = a^2 \cdot \mathrm{Var}(X)$ für deterministische Konstanten $a \in \mathbb{R}$ und Zufallsvariablen X.
- Für unabhängige Zufallsvariablen X, Y gilt $\mathbb{E}[XY] = \mathbb{E}X \cdot \mathbb{E}Y$ und $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

(a) Es gilt

$$\mathbb{E}[\overline{X_n}] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}X_i = \frac{1}{n}\sum_{i=1}^n \mu = \mu,$$

und

$$\mathrm{Var}(\overline{X_n}) = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(X_i) = \frac{1}{n^2}\sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n}.$$

(b) Es ist

$$\begin{aligned} S_n^2 &= \frac{1}{n-1}\sum_{i=1}^n \left(X_i^2 - 2X_i\overline{X_n} + \overline{X_n}^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^n X_i^2 - 2\overline{X_n}\cdot\sum_{i=1}^n X_i + \sum_{i=1}^n \overline{X_n}^2\right) \\ &= \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - 2\overline{X_n}^2 + \overline{X_n}^2\right) = \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - \overline{X_n}^2\right). \end{aligned}$$

Wir berechnen nun den Erwartungswert der einzelnen Teile. Es ist wegen $\mathbb{E}[X_i^2] = \mathrm{Var}(X_i) + \mathbb{E}[X_i]^2 = \sigma^2 + \mu^2$ gerade

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n X_i^2\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i^2] = \frac{1}{n}\sum_{i=1}^n(\sigma^2+\mu^2) = \sigma^2+\mu^2$$

und

$$\begin{aligned} \mathbb{E}[\overline{X_n}^2] &= \mathbb{E}\left[\frac{1}{n^2}\sum_{i,j=1}^n X_iX_j\right] = \frac{1}{n^2}\left[\sum_{i=1}^n \mathbb{E}[X_i^2] + \sum_{\substack{i,j=1\\ i\neq j}}^n \underbrace{\mathbb{E}[X_iX_j]}_{\overset{\text{unabh.}}{=}\mathbb{E}[X_i]\mathbb{E}[X_j]=\mu^2}\right] \\ &= \frac{1}{n^2}\Big[n\cdot(\sigma^2+\mu^2) + n\cdot(n-1)\cdot\mu^2\Big] \\ &= \frac{1}{n}\Big[(\sigma^2+\mu^2) + (n-1)\cdot\mu^2\Big]. \end{aligned}$$

Eingesetzt in die Darstellung für S_n^2 ergibt sich

$$\mathbb{E}[S_n^2] = \frac{n}{n-1}\Big((\sigma^2+\mu^2) - \frac{1}{n}\Big((\sigma^2+\mu^2)+(n-1)\mu^2\Big)\Big) = \frac{n}{n-1}\Big(\frac{n-1}{n}(\sigma^2+\mu^2) - \frac{n-1}{n}\mu^2\Big) = \sigma^2.$$

(c) Es gilt $\mathbb{P}(\mathbb{1}_{\{X_i\le t\}} = 1) = \mathbb{P}(X_i \le t) = F(t)$. Daher ist $\mathbb{1}_{\{X_i\le t\}} \sim \mathcal{B}(1, F(t))$. Da die X_i $(i = 1, ..., n)$ unabhängig sind, sind auch $\mathbb{1}_{\{X_i\le t\}}$ $(i = 1, ..., n)$ unabhängig. Es folgt nach Beispiel 8.13 der Vorlesung gerade

$$n\cdot\hat{F}_n(t) = \sum_{i=1}^n \mathbb{1}_{\{X_i\le t\}} \sim \mathcal{B}(n, F(t)).$$

Aus den bekannten Formeln für Erwartungswert und Varianz für eine Binomialverteilung erhalten wir $\mathbb{E}[n\cdot\hat{F}_n(t)] = n\cdot F(t)$ und $\mathrm{Var}(n\cdot\hat{F}_n(t)) = n\cdot F(t)\cdot(1-F(t))$, womit

$$\mathbb{E}[\hat{F}_n(t)] = F(t), \qquad \mathrm{Var}(\hat{F}_n(t)) = \frac{F(t)\cdot(1-F(t))}{n}$$

folgt.

Aufgabe 30 (Konstruktion von Konfidenzintervallen, 4 = 2 + 1 + 1 Punkte).

Seien $X_1, ..., X_n \overset{\text{iid}}{\sim} U[0,\theta]$ gleichverteilt auf $[0,\theta]$ mit Parameter $\theta > 0$ und $\mathbb{P}_\theta := \mathbb{P}^{X_1,...,X_n}$ die gemeinsame Verteilung dieser Zufallsvariablen.

(a) Sei $\theta_0 > 0$. Zeigen Sie, dass

$$\phi_{\theta_0} : \mathbb{R}^n \to \{0,1\}, \quad \phi_{\theta_0}(x_1, ..., x_n) = \begin{cases} 1, & \max(x_1, ..., x_n) \leq \theta_0 \cdot \alpha^{1/n}, \\ 0, & \text{sonst} \end{cases}$$

für die Hypothesen $H_0 : \theta \geq \theta_0$ gegen $H_1 : \theta < \theta_0$ ein Test zum Niveau $\alpha \in (0,1)$ ist.
Hinweis: Verwenden Sie zur Berechnung der Verteilungsfunktion Aufgabe 28 (a).

(b) Konstruieren Sie mittels Satz 9.3 aus ϕ_{θ_0} ein $(1-\alpha)$-Konfidenzintervall $S(X)$ für θ.

(c) Sei nun $\alpha = 0.1$. Wie groß muss die Anzahl der Beobachtungen n gewählt werden, damit die Länge von $S(X)$ kleiner als 1 ist, wenn Sie wissen, dass die maximale Beobachtung den Wert 0.95 annimmt?

Lösung:

(a) Wir müssen zeigen, dass für alle $\theta \geq \theta_0$ gerade $\mathbb{P}_{\theta_0}(\phi_{\theta_0} = 1) \leq \alpha$ gilt.
Seien dazu $X_1, ..., X_n \overset{\text{iid}}{\sim} U[0,\theta]$. Dann ist bekannt (siehe 7. Übungsblatt, Aufgabe 28(a)), dass

$$Z_n := \max(X_1, ..., X_n)$$

die Verteilungsfunktion $F_{Z_n}(z) = \left(\frac{z}{\theta}\right)^n$ für $z \in [0,\theta]$ besitzt. Nun gilt

$$\mathbb{P}_\theta(\phi_{\theta_0} = 1) = \mathbb{P}(Z_n \leq \theta_0 \cdot \alpha^{1/n}) = F_{Z_n}(\theta_0 \cdot \alpha^{1/n}) \overset{\theta \geq \theta_0 \Rightarrow \theta_0 \cdot \alpha^{1/n} \leq \theta}{=} \left(\frac{\theta_0 \cdot \alpha^{1/n}}{\theta}\right)^n = \left(\frac{\theta_0}{\theta}\right)^n \alpha \overset{\theta \geq \theta_0}{\leq} \alpha.$$

(b) Der Annahmebereich des Tests ist für festes $\theta_0 \in \mathbb{R}$ gegeben durch

$$A(\theta_0) = \{x \in \mathbb{R}^n : \phi_{\theta_0}(x) = 0\} = \{x \in \mathbb{R}^n : \max(x_1, ..., x_n) > \theta_0 \cdot \alpha^{1/n}\}.$$

Damit lautet ein $(1-\alpha)$-Konfidenzintervall für θ laut Satz 9.3 der Vorlesung

$$\begin{aligned} S(X) = \{\theta > 0 : X \in A(\theta)\} &= \left\{\theta > 0 : \max\{X_1, ..., X_n\} > \theta \cdot \alpha^{1/n}\right\} \\ &= \left(0, \frac{\max\{X_1, ..., X_n\}}{\alpha^{1/n}}\right). \end{aligned}$$

(c) Die Länge von $S(X)$ ist

$$L(n) = \left|\frac{\max(X_1, ..., X_n)}{\alpha^{1/n}} - 0\right| = \frac{\max\{X_1, ..., X_n\}}{\alpha^{1/n}}$$

Für $\alpha = 0.1$ und $\max\{X_1, ..., X_n\} = 0.95$ erhalten wir

$$1 \geq L(n) = \frac{0.95}{0.1^{1/n}} \quad \Longleftrightarrow \quad n \geq \frac{\log(0.1)}{\log(0.95)} \approx 44.89,$$

d.h. es muss mindestens $n = 45$ gelten.

Aufgabe 31 (Konstruktion von gleichmäßig besten Tests und Konfidenzintervallen, 8 = 1 + 1 + 2 + 1 + 2 + 1 Punkte).

Wir untersuchen eine Maschine, die Schrauben herstellt. Uns ist bekannt (und das zweifeln wir nicht an), dass die Länge der Schrauben im Mittel $\mu = 5$ beträgt (Angabe in cm). Der Hersteller behauptet, dass die Standardabweichung der Länge der Schrauben höchstens $\sigma_0 = 0.3$ cm beträgt. Wir vermuten jedoch, dass die Standardabweichung tatsächlich größer ist. Für eine statistische Untersuchung nehmen wir an, dass die Länge der Schrauben normalverteilt ist mit Mittelwert μ und Standardabweichung σ. Für $n = 10$ Schrauben beobachten wir folgende Längen (in cm):

5.6	5.2	4.1	3.7	6.5	3.6	6.0	6.1	5.2	4.2

(a) Sei $n \in \mathbb{N}$ und seien $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ mit $\sigma > 0$. Nutzen Sie das auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ verallgemeinerte Neyman-Pearson-Lemma 6.19, um einen besten Test $\phi^* : \mathbb{R}^n \to \{0, 1\}$ für die einfachen Hypothesen

$$H_0 : \sigma = \sigma_0 \quad \text{gegen} \quad H_1 : \sigma = \sigma_1$$

mit $\sigma_1 > \sigma_0$ zum Niveau $\alpha \in (0, 1)$ anzugeben.
Hinweis: Bestimmen Sie zunächst die Dichte f_σ der gemeinsamen Verteilung $\mathbb{P}_\sigma = \mathbb{P}^{X_1,...,X_n}$ von $X_1, ..., X_n$ mit Hilfe von Satz 8.11.

(b) Zeigen Sie mittels der Technik der monotonen Likelihood-Quotienten, dass ϕ^* die Form

$$\phi^*(x) = \begin{cases} 1, & T(x) > c^{**}, \\ 0, & T(x) \leq c^{**} \end{cases}$$

besitzt, wobei $T(x) = \sum_{i=1}^n (x_i - \mu)^2$. Folgern Sie, dass ϕ^* sogar gleichmäßig bester Test ist für $H_0 : \sigma = \sigma_0$ gegen $H_1' : \sigma > \sigma_0$.

(c) Zeigen Sie, dass ϕ^* sogar gleichmäßig bester Test ist für $H_0' : \sigma \leq \sigma_0$ gegen $H_1' : \sigma > \sigma_0$.
Hinweis: Definieren Sie $Z_n := \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2$. Überlegen Sie sich, dass die Verteilung von Z_n nicht mehr von σ abhängt. Zeigen Sie dann, dass $\mathbb{P}_\sigma(\phi^* = 1) \leq \alpha$ gilt für $\sigma \leq \sigma_0$.

(d) Es ist bekannt, dass $Z_n \sim \chi_n^2$, wobei χ_n^2 die Chi-Quadrat-Verteilung mit n Freiheitsgraden bezeichnet. Zeigen Sie, dass $c^{**} = \sigma_0^2 \cdot \chi_{n,1-\alpha}^2$, wobei $\chi_{n,1-\alpha}^2$ das $(1-\alpha)$-Quantil dieser Verteilung bezeichnet.

(e) Wir drücken nun die Abhängigkeit des Tests ϕ^* von σ_0 explizit aus, indem wir ihn mit $\phi_{\sigma_0}^*$ bezeichnen. Bestimmen Sie den Annahmebereich $A(\sigma_0) := \{x \in \mathbb{R}^n : \phi_{\sigma_0}^*(x) = 0\}$ des Tests und damit ein gleichmäßig bestes $(1-\alpha)$-Konfidenzintervall $S(X)$ für σ mittels Satz 9.3. Für welche falschen Parameter wurde dieses Intervall konstruiert?

(f) Sei nun $\alpha = 0.05$. Werden Sie den Hersteller der Maschine auf Basis unserer Beobachtungen und des Tests ϕ^* aus (b) vorwerfen, dass die Angabe der Standardabweichung falsch ist? Geben Sie auf Basis unserer Beobachtungen das 95%-Konfidenzintervall für σ aus (e) an.
Hinweis: Hier sind einige Quantile der χ_n^2-Verteilung: $\chi_{10,0.05}^2 = 3.94$, $\chi_{10,0.95}^2 = 18.31$.

Lösung:

(a) Die Dichte f_σ der gemeinsamen Verteilung $\mathbb{P}_\sigma = \mathbb{P}^{X_1,...,X_n}$ von $X_1, ..., X_n$ ist wegen der Unabhängigkeit von $X_1, ..., X_n$ gegeben durch (setze $x = (x_1, ..., x_n)$)

$$\begin{aligned} f_\sigma(x) &= f_{X_1}(x_1) \cdot ... \cdot f_{X_n}(x_n) = \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu)^2 \right) \right) \\ &= \frac{1}{(2\pi)^{n/2}} \cdot \frac{1}{\sigma^n} \cdot \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right). \end{aligned}$$

Für $\sigma_1 > \sigma_0$ ist der Likelihood-Quotient daher gegeben durch

$$L(x) = \frac{f_{\sigma_1}(x)}{f_{\sigma_0}(x)} = \left(\frac{\sigma_0}{\sigma_1} \right)^n \cdot \exp\left(\frac{1}{2} \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} \right) \sum_{i=1}^n (x_i - \mu)^2 \right). \qquad (*)$$

Der beste Test zum Niveau $\alpha \in (0,1)$ für die Hypothesen H_0 und H_1 lautet daher nach dem (verallgemeinerten) Neyman-Pearson-Lemma 6.19 gerade

$$\phi^*(x) = \begin{cases} 1, & L(x) > c^*, \\ 0, & L(x) \leq c^*, \end{cases}$$

wobei c^* aus $\mathbb{P}_{\sigma_0}(\phi^* = 1) = \mathbb{P}_{\sigma_0}(L(x) > c^*)$ bestimmt wird.

(b) An der Darstellung $(*)$ sehen wir, dass $L(x)$ für $\sigma_1 > \sigma_0$ (d.h. $\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} > 0$) monoton wachsend in $T(x) := \sum_{i=1}^n (x_i - \mu)^2$ ist. Daher vereinfacht sich ϕ^* zu

$$\phi^*(x) = \begin{cases} 1, & T(x) > c^{**}, \\ 0, & T(x) \leq c^{**}, \end{cases}$$

wobei nun c^{**} bestimmt wird aus $\mathbb{P}_{\sigma_0}(\phi^* = 1) = \mathbb{P}_{\sigma_0}(T(x) > c^{**})$. Da der gesamte Test ϕ^* und insbesondere die Bestimmungsgleichung für c^{**} nicht mehr von σ_1 abhängt, ist der Test ϕ^* auch gleichmäßig bester Test für die Hypothesen H_0 gegen H_1'.

(c) Zu zeigen ist, dass für alle $\sigma \leq \sigma_0$ gilt:

$$\mathbb{P}_\sigma(\phi^* = 1) = \mathbb{P}_\sigma(T(x) > c^{**}) \leq \alpha.$$

Seien nun wieder $X_1, ..., X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$. Dann sind $\frac{X_i - \mu}{\sigma} \sim N(0,1)$, i = 1,...,n, standardnormalverteilte und unabhängige Zufallsvariablen. Entsprechend ist die Verteilung von $Z_n = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2$ unabhängig von σ (und μ). (Das genaue Argument lautet: Die gemeinsame Verteilung von $\frac{X_i - \mu}{\sigma}$, $i = 1, ..., n$, hängt nicht mehr von σ ab, und Z_n ist nur eine Funktion dieser n Zufallsvariablen). Wir erhalten (offensichtlich muss $c^{**} \geq 0$ gelten)

$$\begin{aligned} \mathbb{P}_\sigma(\phi^* = 1) = \mathbb{P}(\sigma^2 \cdot Z_n > c^{**}) &= \mathbb{P}(Z_n > \frac{c^{**}}{\sigma^2}) \\ &\overset{\sigma \leq \sigma_0}{\leq} \mathbb{P}(Z_n > \frac{c^{**}}{\sigma_0^2}) = \mathbb{P}(\sigma_0^2 \cdot Z_n > c^{**}) \overset{(*)}{=} \mathbb{P}_{\sigma_0}(\phi^* = 1). \quad (1) \end{aligned}$$

Anmerkung zu $(*)$:
Im Grunde definieren wir hier eine weitere Zufallsvariable $W_n = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0}\right)^2$, welche unabhängig von σ_0 (und μ) ist und die selbe Verteilung wie Z_n hat. Dadurch erhalten wir die Gleichungskette

$$\mathbb{P}(\sigma_0^2 \cdot Z_n > c^{**}) = \mathbb{P}(\sigma_0^2 \cdot W_n > c^{**}) = \mathbb{P}_{\sigma_0}(\phi^* = 1).$$

Beachte nun, dass ein gleichmäßig bester Test ϕ^{**} für die Hypothesen H_0' gegen H_1' das folgende Optimierungsproblem Nr. 1 lösen muss:
Für alle Tests $\phi : \mathbb{R}^n \to \{0,1\}$ mit

$$\forall\, \sigma \leq \sigma_0 : \mathbb{P}_\sigma(\phi = 1) \leq \alpha \quad \text{(Fehler 1. Art } \leq \alpha) \tag{2}$$

gilt

$$\forall\, \sigma > \sigma_0 : \mathbb{P}_\sigma(\phi^{**} = 0) \leq \mathbb{P}_\sigma(\phi = 0) \quad \text{(Fehler 2. Art wird minimiert).}$$

Die Bedingung (2) ist stärker als die entsprechende Bedingung für einen gleichmäßigen Test von H_0 gegen H_1', wo die Tests ϕ nur $\mathbb{P}_{\sigma_0}(\phi = 1) \leq \alpha$ erfüllen müssen. Das bedeutet für das Optimierungsproblem Nr. 1 kommen weniger Tests in Frage als für das Optimierungsproblem

Nr. 2 von H_0 gegen H_1'. Naturgemäß muss dann eine Lösung ϕ^{**} aus Optimierungsproblem Nr. 1 "schlechter" sein als unsere Lösung ϕ^* von Optimierungsproblem Nr. 2, d.h. es gilt

$$\forall\, \sigma > \sigma_0 : \quad \mathbb{P}_\sigma(\phi^{**} = 0) \geq \mathbb{P}_\sigma(\phi^* = 0).$$

Wir haben aber soeben in (1) gezeigt, dass ϕ^* sogar in der Menge der Tests für das Optimierungsproblem Nr. 1 enthalten ist. Das bedeutet (da ϕ^{**} Lösung des Optimierungsproblems Nr. 1), dass

$$\forall\, \sigma > \sigma_0 : \quad \mathbb{P}_\sigma(\phi^{**} = 0) \leq \mathbb{P}_\sigma(\phi^* = 0)$$

gilt. Entsprechend erhalten wir insgesamt

$$\forall\, \sigma > \sigma_0 : \quad \mathbb{P}_\sigma(\phi^{**} = 0) = \mathbb{P}_\sigma(\phi^* = 0).$$

und damit die Tatsache, dass ϕ^* eine Lösung des Optimierungsproblems Nr. 1 ist, d.h. ϕ^* ist gleichmäßig bester Test für H_0' gegen H_1'.

(d) Seien $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma_0^2)$. Die Bestimmungsgleichung für c^{**} lautet

$$\alpha = \mathbb{P}_{\sigma_0}(\phi^* = 1) = \mathbb{P}_{\sigma_0}(T(x) > c^{**}) = \mathbb{P}(\sigma_0^2 Z_n > c^{**}) = 1 - \mathbb{P}(Z_n \leq \frac{c^{**}}{\sigma_0^2}).$$

Damit folgt

$$\mathbb{P}(Z_n \leq \frac{c^{**}}{\sigma_0^2}) = 1 - \alpha \quad \Longleftrightarrow \quad \frac{c^{**}}{\sigma_0^2} = \chi^2_{n,1-\alpha} \quad \Longleftrightarrow \quad c^{**} = \sigma_0^2 \cdot \chi^2_{n,1-\alpha}.$$

(e) In (a) bis (d) haben wir den gleichmäßig besten Test

$$\phi^*_{\sigma_0}(x) = \begin{cases} 1, & T(x) > \sigma_0^2 \chi^2_{n,1-\alpha}, \\ 0, & T(x) \leq \sigma_0^2 \chi^2_{n,1-\alpha} \end{cases}$$

für die Hypothesen $H_0' : \sigma \leq \sigma_0$ gegen $H_1' : \sigma > \sigma_0$ konstruiert. Der Annahmebereich des Tests lautet demzufolge

$$A(\sigma_0) = \{x \in \mathbb{R}^n : \phi^*_{\sigma_0}(x) = 0\} = \{x \in \mathbb{R}^n : T(x) \leq \sigma_0^2 \chi^2_{n,1-\alpha}\}.$$

Für das Konfidenzintervall für σ aus Satz 9.3 erhalten wir

$$\begin{aligned} S(X) &= \{\sigma > 0 : X \in A(\sigma)\} = \{\sigma > 0 : T(X) \leq \sigma^2 \chi^2_{n,1-\alpha}\} = \left[\sqrt{\frac{T(X)}{\chi^2_{n,1-\alpha}}}, \infty\right) \\ &= \left[\sqrt{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{n,1-\alpha}}}, \infty\right). \end{aligned}$$

Um die falschen Parameter, die bei der Konstruktion dieses Konfidenzintervalls zugrunde lagen, zu ermitteln, müssen wir die Bereiche der Hypothesen genauer studieren. Bei unserem Test haben wir Hypothesen der Form

$$H_0' : \sigma \in H(\sigma_0) \quad \text{gegen} \quad H_1' : \sigma \in K(\sigma_0)$$

mit $H(\sigma_0) = (0, \sigma_0]$ und $K(\sigma_0) = (\sigma_0, \infty)$. Die falschen Parameter berechnen sich durch

$$\overline{K}(\sigma) = \{\overline{\sigma} > 0 : \sigma \in K(\overline{\sigma})\} = \{\overline{\sigma} > 0 : \sigma \in (\overline{\sigma}, \infty)\} = \{\overline{\sigma} > 0 : \sigma > \overline{\sigma}\} = (0, \sigma)$$

(Dies folgt auch aus Bemerkung 9.5, erste Zeile). Das heißt, bei der Konstruktion des Konfidenzintervalls waren wir daran interessiert, eine untere Grenze für die Standardabweichung zu erhalten, und haben daher kleine σ als falsche Parameter deklariert.

(f) Für die konkreten Beobachtungen aus der Aufgabenstellung erhalten wir (beachte $\mu = 5$, $n = 10$):

$$T(X) = \sum_{i=1}^{n} (X_i - \mu)^2 = 10.0,$$

und mit $\sigma_0^2 = 0.3^2$ und $\chi^2_{n,1-\alpha} = \chi^2_{10,0.95} = 18.31$ ergibt sich

$$\sigma_0^2 \chi^2_{n,1-\alpha} = 1.65$$

Damit gilt für unsere Beobachtungen $T(X) > \sigma_0^2 \chi^2_{n,1-\alpha}$, d.h. $\phi^*(X) = 1$. Wir verwerfen also die Nullhypothese und können statistisch gesichert davon ausgehen, dass die angegebene Standardabweichung von $\sigma_0 = 0.3$ zu klein ist. Wir sollten den Hersteller deswegen kontaktieren.
Für das 95%-Konfidenzintervall für σ aus (e) erhalten wir

$$S(X) = \left[\sqrt{\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\chi^2_{n,1-\alpha}}}, \infty \right) = \left[\sqrt{\frac{10}{18.31}}, \infty \right) \approx [0.74, \infty).$$

Abgabe:
In Zweiergruppen, bis spätestens Donnerstag, den **19. Dezember 2019, 11:00 Uhr**.
(Die Zettelkästen für das Abgabeblatt sind im 1. OG, INF 205, vor dem Dekanat.)

Homepage der Vorlesung:
`https://ssp.math.uni-heidelberg.de/ews-WS1920/index.html`