UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Lösungsskizzen zum Extrablatt

Aufgabe 1 (i) Sei $\Omega = \mathbb{N}_0$, $\mathcal{A} = \mathcal{P}(\Omega)$ und $p : \Omega \to [0,1]$ für $p \in (0,1)$, $r \in \mathbb{N}$, $\omega \in \Omega$ gegeben durch

$$p(\omega) = \binom{\omega + r - 1}{\omega} p^r (1-p)^\omega.$$

Zeigen Sie, dass p eine Zähldichte eines Wahrscheinlichkeitsmaßes auf $(\Omega, \mathcal{A})$ ist. Was kann mit Hilfe dieser Zähldichte modelliert werden?
Zeigen Sie die Regel $\binom{\alpha+k-1}{k} = (-1)^k \binom{-\alpha}{k}$ *für den verallgemeinerten Binomialkoeffizienten* $\binom{\alpha}{k} := \frac{\alpha\cdot(\alpha-1)\cdot\ldots\cdot(\alpha-(k-1))}{k!}$*, der für* $\alpha \in \mathbb{Z}$ *und* $k \in \mathbb{N}_0$ *definiert ist. Beweisen Sie dann, dass für die binomische Reihe* $(1+x)^\alpha = \sum_{k=0}^\infty \binom{\alpha}{k} x^k$ *gilt, wobei* $\alpha \in \mathbb{Z}$*,* $x \in (-1,1)$*.*

(ii) Sei wieder $\Omega = \mathbb{N}_0$ und $\mathbb{P}$ als Wahrscheinlichkeitsverteilung auf $(\Omega, \mathcal{P}(\Omega))$ definiert durch

$$\mathbb{P}(\{n\}) := 2^{-n-1}$$

für $n \in \mathbb{N}_0$. Weiter sei eine Abbildung gegeben durch

$$X : \Omega \to \mathbb{R}, \quad n \mapsto n \mod 3.$$

Bestimmen Sie die induzierte Verteilung $\mathbb{P}^X$ auf $(\Omega^X, \mathcal{P}(\Omega^X)$.

(iii) Nach langjähriger Erfahrung wissen Sie, dass Sie bei einer Runde Skat mit Wahrscheinlichkeit $p = 0.2$ gewinnen. Sie sind nun zu einem Spieleabend eingeladen worden, bei dem ausschließlich Skat gespielt wird. Wie hoch ist die Wahrscheinlichkeit, dass Sie genau beim 30. Spiel zum sechsten Mal gewinnen?

Beweis. (i) Zunächst zur Formel für den verallgemeinerten Binomialkoeffizienten: Es gilt

$$\begin{aligned}\binom{\alpha+k-1}{k} &= \frac{(\alpha+k-1)(\alpha+k-2)\ldots(\alpha+k-k)}{k!} \\ &= \frac{(-1)^k(-\alpha-k+1)(-\alpha-k+2)\ldots(-\alpha)}{k!} = (-1)^k\binom{-\alpha}{k}.\end{aligned}$$

Nun zur binomischen Reihe. Wir zeigen zuerst die Konvergenz der Reihe für $|x| < 1$ mit Hilfe des Quotientenkriteriums; sei dazu $a_k := \binom{\alpha}{k} x^k$, dann gilt

$$\left|\frac{a_{k+1}}{a_k}\right| = \left|\frac{\binom{\alpha}{k+1}x^{k+1}}{\binom{\alpha}{k}x^k}\right| = \left|\frac{\alpha-k}{k+1}\right| |x|$$

und damit $\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = |x| < 1$. Somit konvergiert die Reihe $\sum_{k=0}^\infty a_k$. Die Darstellung erhalten wir, in dem wir $f(x) = (1+x)^\alpha$ im Punkt 0 Taylor-Entwickeln. Dazu bestimmen wir die k-ten Ableitungen:

$$f^{(k)}(x) = \alpha\cdots(\alpha-(k-1))(1+x)^{\alpha-k}$$

und wir erhalten sofort

$$(1+x)^\alpha = f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = \sum_{k=0}^{\infty} \frac{(\alpha \cdots (\alpha-(k-1))}{k!} x^k = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$$

Bemerkung: Da $f(x) = (1+x)^\alpha$ eine analytische Funktion ist (für $\alpha \geq 0$ handelt es sich um ein Polynom, für $\alpha < 0$ um einen Quotienten von analytischen Funktionen), wird f innerhalb des Konvergenzbereiches tatsächlich durch die Taylorreihe dargestellt.

Damit können wir nun nachrechnen, dass es sich bei $\mathbb{p}$ um eine Zähldichte handelt:

$$\begin{aligned}
\sum_{\omega\in\Omega} \mathbb{p}(\omega) &= \sum_{\omega\in\Omega} \binom{\omega+r-1}{\omega} p^r (1-p)^\omega = p^r \sum_{\omega\in\Omega} \binom{\omega+r-1}{\omega} (1-p)^\omega \\
&= p^r \sum_{\omega\in\Omega} \binom{-r}{\omega} (-1)^\omega (1-p)^\omega = p^r \sum_{\omega\in\Omega} \binom{-r}{\omega} (p-1)^\omega \\
&= p^r (1+(p-1))^{-r} = p^r p^{-r} = 1
\end{aligned}$$

Durch diese Zähldichte kann die Anzahl der Misserfolge vor dem r-ten Erfolg bei einem Bernoulli-Schema mit Erfolgswahrscheinlichkeit p modelliert werden, d.h. $\mathbb{p}(\omega)$ gibt die Wahrscheinlichkeit an genau ω Misserfolge vor dem r-ten Erfolg zu haben.

(ii) Es ist $\mathrm{Bild}(X) = \{0,1,2\}$. Die Wahrscheinlichkeitsverteilung $\mathbb{P}^X$ ist durch die Zuordnung $x \mapsto \mathbb{P}^X(\{x\})$ bereits eindeutig angeben. Das heißt, wir müssen folgende Terme bestimmen.

$$\mathbb{P}^X(\{x\}) = \mathbb{P}(X^{-1}(\{x\})), \qquad x \in \{0,1,2\}.$$

Es gilt $X^{-1}(\{x\}) = x + 3\mathbb{N}_0 = \{x + 3n : n \in \mathbb{N}_0\}$. Daher folgt

$$\begin{aligned}
\mathbb{P}^X(\{x\}) &= \mathbb{P}(X^{-1}(\{x\})) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(\{3n+x\}) = \sum_{n=0}^{\infty} 2^{-(3n+x)-1} \\
&= 2^{-x-1} \cdot \sum_{n=0}^{\infty} \left(\frac{1}{8}\right)^n = 2^{-x-1} \cdot \frac{1}{1-\frac{1}{8}} = \frac{8}{7} \cdot 2^{-x-1},
\end{aligned}$$

konkret

$$\mathbb{P}^X(\{0\}) = \frac{4}{7}, \quad \mathbb{P}^X(\{1\}) = \frac{2}{7}, \quad \mathbb{P}^X(\{2\}) = \frac{1}{7}.$$

(iii) Um beim 30. Spiel zum 6. Mal zu gewinnen, müssen wir vor dem 6. Erfolg genau 24 Misserfolge haben. Die Wahrscheinlichkeit hierfür ist nach (i) mit $p = 0.2$, $r = 6$, $\omega = 24$:

$$\mathbb{p}(24) = \binom{29}{24} 0.2^6 0.8^{24} \approx 0.036$$

$\square$

Aufgabe 2

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \mathbb{N}$ eine Zufallsvariable mit Werten in den natürlichen Zahlen.

(a) Es sei $X \sim \mathcal{G}(p)$ geometrisch verteilt mit Parameter $p \in (0,1)$. Zeigen Sie, dass für alle $k, n \in \mathbb{N}$ gilt, dass

$$\mathbb{P}(X = k) = \mathbb{P}(X = n - 1 + k \mid X \geq n). \tag{1}$$

Diese Eigenschaft wird als "Gedächtnislosigkeit" der geometrischen Verteilung bezeichnet.

(b) Zeigen Sie, dass eine Zufallsvariable $X : \Omega \to \mathbb{N}$, die (1) erfüllt, geometrisch verteilt ist mit einem geeigneten Parameter p.
Hinweis: Definieren Sie $p := \mathbb{P}(X = 1)$, nutzen Sie (1) mit $k = 1$ und leiten Sie so eine rekursive Beziehung der Form $\mathbb{P}(X \geq n+1) = c \cdot \mathbb{P}(X \geq n)$ her mit geeignetem $c \in \mathbb{R}$.

Beweis. Wir halten zunächst für beide Aufgabenteile fest, dass $X = n - 1 + k \geq n$, da $k \in \mathbb{N}$. D.h. $\{X = n - 1 + k\} \subset \{X \geq n\}$, also

$$\{X = n - 1 + k, X \geq n\} = \{X = n - 1 + k\} \cap \{X \geq n\} = \{X = n - 1 + k\}.$$

Damit ist

$$\mathbb{P}(X = n - 1 + k | X \geq n) = \frac{\mathbb{P}(X = n - 1 + k, X \geq n)}{\mathbb{P}(X \geq n)} = \frac{\mathbb{P}(X = n - 1 + k)}{\mathbb{P}(X \geq n)}. \tag{2}$$

(a) Da X geometrisch verteilt ist mit Parameter p, gilt:

$$\mathbb{P}(X = k) = (1-p)^{k-1} \cdot p$$

für $k \in \mathbb{N}$. Damit folgt

$$\begin{aligned} \mathbb{P}(X < n) &= \sum_{k=1}^{n-1} \mathbb{P}(X = k) = \sum_{k=1}^{n-1} (1-p)^{k-1} \cdot p \\ &= p \cdot \sum_{k=0}^{n-2} (1-p)^k = p \cdot \frac{1 - (1-p)^{n-1}}{1 - (1-p)} = 1 - (1-p)^{n-1}, \end{aligned}$$

also $\mathbb{P}(X \geq n) = 1 - \mathbb{P}(X < n) = (1-p)^{n-1}$. Hiermit erhalten wir:

$$\begin{aligned} \mathbb{P}(X = n - 1 + k | X \geq n) &= \frac{\mathbb{P}(X = n - 1 + k)}{\mathbb{P}(X \geq n)} = \frac{(1-p)^{(n-1+k)-1} \cdot p}{(1-p)^{n-1}} \\ &= (1-p)^{k-1} \cdot p = \mathbb{P}(X = k). \end{aligned}$$

Erklärung der Gedächtnislosigkeit:

- $\mathbb{P}(X = k)$ ist die Wahrscheinlichkeit, dass nach $k - 1$ Misserfolgen beim k-ten Versuch ein Erfolg eingetreten ist.
- $\{X \geq n\}$ bezeichnet also das Ereignis, dass die ersten $n-1$ Versuche Misserfolge gewesen sind und Erfolg erst ab dem n-ten Versuch (oder später!) eintritt.
- (1) bedeutet also: Die Wahrscheinlichkeit, dass der erste Erfolg beim k-ten Versuch eintritt, ist gleich der Wahrscheinlichkeit, dass der erste Erfolg beim $(n-1+k)-$ten Versuch auftritt, wenn bis zum $(n-1)$-ten Versuch kein Erfolg eingetreten ist.
- D.h. es ist für den k-ten Versuch uninteressant, wie viele $(n-1)$ Misserfolge vorher stattgefunden haben. Die Unabhängigkeit des Ausgangs des k-ten Versuchs von den Ausgängen der vorherigen Versuche ist die Gedächtnislosigkeit der geometrischen Verteilung.

(b) Sei $X : \Omega \to \mathbb{N}$ eine $\mathbb{N}$-wertige Zufallsvariable. Definiere $p := \mathbb{P}(X = 1)$. Dann gilt für alle $n \in \mathbb{N}$

$$p = \mathbb{P}(X = 1) = \frac{\mathbb{P}(X = n)}{\mathbb{P}(X \geq n)} = \frac{\mathbb{P}(X \geq n) - \mathbb{P}(X \geq n+1)}{\mathbb{P}(X \geq n)} = 1 - \frac{\mathbb{P}(X \geq n+1)}{\mathbb{P}(X \geq n)},$$

woraus folgt $\mathbb{P}(X \geq n+1) = (1-p) \cdot \mathbb{P}(X \geq n)$. Damit folgt für beliebiges $n \in \mathbb{N}$, dass

$$\begin{aligned}\mathbb{P}(X \geq n) &= (1-p) \cdot \mathbb{P}(X \geq n-1) = (1-p)^2 \cdot \mathbb{P}(X \geq n-2) \\ &= ... = (1-p)^{n-1} \cdot \mathbb{P}(X \geq 1) = (1-p)^{n-1}.\end{aligned}$$

Es ist $\mathbb{P}(X \geq 1) = 1$, da X eine $\mathbb{N}$-wertige Zufallsvariable. Insgesamt erhalten wir

$$\mathbb{P}(X = n) = \mathbb{P}(X \geq n) - \mathbb{P}(X \geq n+1) = (1-p)^{n-1} - (1-p)^n = (1-p)^{n-1} \cdot p.$$

Das heißt, X ist geometrisch verteilt.

□

Aufgabe 3

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \mathbb{R}$ eine steig verteilte Zufallsvariable, die einer Pareto-Verteilung folgt, d.h. $X \sim \text{Pareto}(\alpha, x_m)$ mit Parametern $\alpha > 0$ und $x_m > 0$. Die Dichte ist gegeben durch

$$f_X(x) = C_{\alpha,x_m} \cdot x^{-(\alpha+1)} \mathbb{1}_{\{x \geq x_m\}}(x) = \begin{cases} C_{\alpha,x_m} \cdot x^{-(\alpha+1)}, & x \geq x_m, \\ 0, & x < x_m \end{cases}.$$

(a) Bestimmen Sie die Konstante C_{α,x_m} so, dass $f_X(x)$ eine Wahrscheinlichkeitsdichte ist.

(b) Bestimmen Sie die Verteilungsfunktion $F_X(x)$ von X.

(c) Welche bekannte Verteilung besitzt die Zufallsvariable Y mit $Y = \log\left(\frac{X}{x_m}\right)$?

(d) Sei nun $\alpha = x_m = 1$. Berechnen Sie die Wahrscheinlichkeiten $\mathbb{P}(1 \leq X \leq 2)$ und $\mathbb{P}(X > 2)$.

Beweis. (a) Damit f_X eine Wahrscheinlichkeitsdichte ist, müssen zwei Dinge erfüllt sein: $f_X \geq 0$ überall und $\int_{\mathbb{R}} f_X(x)\mathrm{d}x = 1$. Die erste Bedingung ist offensichtlich erfüllt, solange $C_{\alpha,x_m} \geq 0$ ist (man muss für diese fehlende Begründung nicht unbedingt Punkte abziehen).
Aus der zweiten Bedingung ermitteln wir C_{α,x_m} wie folgt.

$$1 = \int_{\mathbb{R}} f_X(x)\mathrm{d}x = C_{\alpha,x_m} \int_{x_m}^{\infty} x^{-(\alpha+1)}\mathrm{d}x = C_{\alpha,x_m} \cdot \Big[-\frac{1}{\alpha} \cdot x^{-\alpha} \Big]_{x_m}^{\infty} = C_{\alpha,x_m} \cdot \frac{1}{\alpha} \cdot x_m^{-\alpha},$$

die Gleichung ist genau mit $C_{\alpha,x_m} = \alpha x_m^{\alpha}$ erfüllt.

(b) Es gilt für $x \geq x_m$, dass

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(y)\mathrm{d}y \\ &= C_{\alpha,x_m} \cdot \int_{x_m}^{x} y^{-(\alpha+1)}\mathrm{d}y = C_{\alpha,x_m} \cdot \Big[-\frac{1}{\alpha} \cdot y^{-\alpha} \Big]_{x_m}^{x} \\ &= -\frac{C_{\alpha,x_m}}{\alpha} \Big[x^{-\alpha} - x_m^{-\alpha} \Big] = 1 - \left(\frac{x_m}{x}\right)^{\alpha}. \end{aligned}$$

und andererseits für $x < x_m$

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(y)\ \mathrm{d}y = 0,$$

da die Dichte f_X erst für Werte größer als x_m nicht Null ist. Insgesamt erhalten wir

$$F_X(x) = \begin{cases} 0, & x < x_m, \\ 1 - \left(\frac{x_m}{x}\right)^{\alpha}, & x \geq x_m \end{cases}$$

(c) Mit der Funktion $g(x) := \log(\frac{x}{x_m})$ (die stetig differenzierbar und streng monoton wachsend ist) gilt nach dem Transformationssatz für stetige Zufallsvariablen 6.16 für die Dichte f_Y von $Y = g(X)$:

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left|\frac{d}{dy} g^{-1}(y)\right|.$$

Hier ist

$$\begin{aligned} y = g(x) &\Leftrightarrow y = \log(\frac{x}{x_m}) \\ &\Leftrightarrow e^y = \frac{x}{x_m} \\ &\Leftrightarrow x = x_m \cdot e^y =: g^{-1}(y) \end{aligned}$$

und $\frac{d}{dy} g^{-1}(y) = x_m \cdot e^y$. Daher erhalten wir

$$\begin{aligned} f_Y(y) &= f_X(x_m \cdot e^y) \cdot x_m \cdot e^y \\ &= \Big(C_{\alpha,x_m} \cdot x_m^{-(\alpha+1)} e^{-y(\alpha+1)} \cdot \mathbb{1}_{\{x_m e^y \geq x_m\}}\Big) \cdot \Big(x_m \cdot e^y\Big) \\ &= \alpha \cdot e^{-\alpha y} \cdot \mathbb{1}_{\{y \geq 0\}}, \end{aligned}$$

da $x_m e^y \geq x_m \Leftrightarrow y \geq 0$. Das ist die Dichte einer $\mathcal{E}(\alpha)$-Verteilung, d.h. $Y \sim \mathcal{E}(\alpha)$.

(d) Es gilt mit $\alpha = x_m = 1$, dass $C_{\alpha,x_m} = 1$ und $f_X(x) = x^{-2}\mathbb{1}_{\{x \geq 1\}}$.

$$\mathbb{P}(1 \leq X \leq 2) = \int_1^2 f_X(x)\mathrm{d}x = \Big[-x^{-1}\Big]_1^2 = -\frac{1}{2} + 1 = \frac{1}{2}.$$

Alternativ kann diese Wahrscheinlichkeit auch mittels der Verteilungsfunktion $F_X(x) = \left(1 - \frac{1}{x}\right)\mathbb{1}_{\{x \geq 1\}}(x)$ berechnet werden:

$$\mathbb{P}(1 \leq X \leq 2) = \mathbb{P}(X \leq 2) - \mathbb{P}(X \leq 1) = F_X(2) - F_X(1) = \Big(1 - \frac{1}{2}\Big) - \Big(1 - 1\Big) = \frac{1}{2}.$$

Aufgrund Feststellung $\mathbb{P}(X \leq 1) = F_X(1) = 0$ (sieht man auch an der Wahrscheinlichkeitsdichte f_X selbst) gilt

$$\mathbb{P}(X > 2) = 1 - \mathbb{P}(X \leq 1) - \mathbb{P}(1 \leq X \leq 2) = \frac{1}{2}.$$

Bemerkung: Wir haben in (d) ohne weiteren Kommentar benutzt, dass $\mathbb{P}(X = x) = 0$ bei stetigen Verteilungen. Die Wahrscheinlichkeit, dass eine stetig verteilte Zufallsvariable genau einen bestimmten Wert $x \in \mathbb{R}$ annimmt, ist Null.

□

Aufgabe 4

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 $X_1, ..., X_n$.

(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

$$H_0 : \sigma = \sigma_0 \quad \text{gegen} \quad H_1' : \sigma > \sigma_0.$$

(c) Zeigen Sie, dass ϕ^* sogar gleichmäßig bester Test ist für

$$H_0' : \sigma \leq \sigma_0 \quad \text{gegen} \quad H_1' : \sigma > \sigma_0.$$

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 $Z_n \sim \chi^2_n$, wobei χ^2_n die Chi-Quadrat-Verteilung mit n Freiheitsgraden bezeichnet. Es bezeichne $\chi^2_{n,1-\alpha}$ das $(1-\alpha)$-Quantil dieser Verteilung. Zeigen Sie, dass $c^{**} = \sigma_0^2 \cdot \chi^2_{n,1-\alpha}$.

(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. 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^2_n-Verteilung: $\chi^2_{10,0.05} = 3.94$, $\chi^2_{10,0.95} = 18.31$.

Beweis. (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

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

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

$$L(x) = \frac{f_{\sigma_1}(x)}{f_{\sigma_0}(x)} = \left(\frac{\sigma_0}{\sigma_1} \right)^n \cdot \exp\left(\frac{1}{2} \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} \right) \sum_{i=1}^{n} (x_i - \mu)^2 \right). \tag{3}$$

mit $x = (x_1, ..., x_n)$. Der beste Test zum Niveau $\alpha \in (0, 1)$ für die Hypothesen H_0, H_1 lautet daher nach dem (verallgemeinerten) Neyman-Pearson-Lemma 6.19:

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

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

(b) An der Darstellung (3) sehen wir, dass $L(x)$ für $\sigma_1 > \sigma_0$, also $\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} > 0$, monoton wachsend in $T(x) := \sum_{i=1}^{n} (x_i - \mu)^2$ ist. Daher vereinfacht sich ϕ^* zu

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

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 $\sigma \le \sigma_0$ gilt, dass

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

Seien nun wieder $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$. Dann sind $\frac{X_i - \mu}{\sigma} \sim \mathcal{N}(0, 1)$ 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 eigentliche 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 mit $c^{**} \ge 0$ gelten, dass

$$\begin{aligned} \mathbb{P}_\sigma(\phi^* = 1) &= \mathbb{P}(\sigma^2 \cdot Z_n > c^{**}) && (4) \\ &= \mathbb{P}(Z_n > \frac{c^{**}}{\sigma^2}) \overset{\sigma \le \sigma_0}{\le} \mathbb{P}(Z_n > \frac{c^{**}}{\sigma_0^2}) && (5) \\ &= \mathbb{P}(\sigma_0^2 \cdot Z_n > c^{**}) = \mathbb{P}_{\sigma_0}(\phi^* = 1). && (6) \end{aligned}$$

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

$$\forall \sigma \le \sigma_0 : \mathbb{P}_\sigma(\phi = 1) \le \alpha \tag{7}$$

gilt

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

Die Bedingung (7) 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 (4)-(6) 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

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

(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 zu ermitteln, die bei der Konstruktion dieses Konfidenzintervalls zugrunde lagen, 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 Untergrenze 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:

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

□

Aufgabe 5

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und seien $X, Y : \Omega \to \mathbb{R}$ zwei Zufallsvariablen.

(a) Seien X, Y stochastisch unabhängig. Zeigen Sie, dass gilt

$$\begin{aligned} \operatorname{Kov}(XY, Y) &= \mathbb{E}[X] \cdot \operatorname{Var}(Y), \\ \operatorname{Var}(XY) &= \operatorname{Var}(X)\operatorname{Var}(Y) + \mathbb{E}[X]^2 \operatorname{Var}(Y) + \mathbb{E}[Y]^2 \operatorname{Var}(X). \end{aligned}$$

(b) Gelte $\operatorname{Var}(X) = \operatorname{Var}(Y)$. Zeigen Sie, dass

$$\operatorname{Kov}(X + Y, X - Y) = 0.$$

(c) Seien $X, Y \sim \operatorname{Ber}(1, p)$ unabhängig und Bernoulli-verteilt mit Parameter $p \in (0, 1)$. Zeigen Sie, dass $X + Y$ und $X - Y$ nicht stochastisch unabhängig sind.

Beweis. (a) Es ist

$$\begin{aligned} \operatorname{Kov}(XY, Y) &= \mathbb{E}[XY^2] - \mathbb{E}[XY]\mathbb{E}[Y] \\ &= \mathbb{E}[X]\mathbb{E}[Y^2] - \mathbb{E}[X]\mathbb{E}[Y]^2 \\ &= \mathbb{E}[X](\mathbb{E}[Y^2] - \mathbb{E}[Y]^2) = \mathbb{E}[X]\operatorname{Var}(Y) \end{aligned}$$

und

$$\begin{aligned}
&\mathrm{Var}(XY) \\
&= \mathbb{E}[(XY)^2] - \mathbb{E}[XY]^2 = \mathbb{E}[X^2]\mathbb{E}[Y^2] - \mathbb{E}[X]^2\mathbb{E}[Y]^2 \\
&= \Big(\mathrm{Var}(X)\mathrm{Var}(Y) + \mathbb{E}[X^2]\mathbb{E}[Y]^2 + \mathbb{E}[X]^2\mathbb{E}[Y^2] - \mathbb{E}[X]^2\mathbb{E}[Y]^2\Big) - \mathbb{E}[X]^2\mathbb{E}[Y]^2 \\
&= \mathrm{Var}(X)\mathrm{Var}(Y) + \mathbb{E}[Y]^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2) + \mathbb{E}[X]^2(\mathbb{E}[Y^2] - \mathbb{E}[Y]^2) \\
&= \mathrm{Var}(X)\mathrm{Var}(Y) + \mathbb{E}[Y]^2\mathrm{Var}(X) + \mathbb{E}[X]^2\mathrm{Var}(Y).
\end{aligned}$$

(b) Hier ist mit der Bilinearität/Symmetrie der Kovarianz

$$\begin{aligned}
\mathrm{Kov}(X+Y, X-Y) &= \mathrm{Kov}(X,X) \underbrace{-\mathrm{Kov}(X,Y) + \mathrm{Kov}(Y,X)}_{=0} -\mathrm{Kov}(Y,Y) \\
&= \mathrm{Var}(X) - \mathrm{Var}(Y) = 0.
\end{aligned}$$

(c) Es ist

$$\mathbb{P}(X+Y=0, X-Y=1) = \mathbb{P}(\emptyset) = 0,$$

da $X+Y=0$ bei $X,Y \in \{0,1\}$ schon $X=Y=0$ impliziert, aber dann $X-Y = 0 \neq 1$ gilt. Aber

$$\begin{aligned}
\mathbb{P}(X+Y=0)\cdot\mathbb{P}(X-Y=1) &= \mathbb{P}(X=0, Y=0)\cdot\mathbb{P}(X=1, Y=0) \\
&= (1-p)^2 \cdot p(1-p) \neq 0.
\end{aligned}$$

Damit ist $\mathbb{P}(X+Y=0, X-Y=1) \neq \mathbb{P}(X+Y=0)\cdot\mathbb{P}(X-Y=1)$, bei Unabhängigkeit müsste hier aber Gleichheit gelten.

□