Einführung in die Wahrscheinlichkeitstheorie und Statistik
Wintersemester 2022/23
Dozent: Prof. Dr. Rainer Dahlhaus
Assistenz: Marilena Müller

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Lösungsskizzen zu Übungsblatt 8

Aufgabe 1

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

(i) Seien $X \sim \mathcal{P}(\lambda)$, $Y \sim \mathcal{P}(\mu)$ Poisson-verteilt mit Parametern $\lambda, \mu > 0$. Berechnen Sie die Verteilung von $Z := X + Y$.

(ii) Seien $X, Y \sim \mathcal{R}[0, 1]$ gleichverteilt auf $[0, 1]$. Berechnen Sie die Wahrscheinlichkeitsdichte f_Z von $Z := \max\{X, Y\}$.

(1+2 Punkte)

Beweis. (i) Es gilt

$$\begin{aligned}
\mathbb{P}(Z = z) = \mathbb{P}(X + Y = z) &= \sum_{k=0}^{\infty} \mathbb{P}(X = k, X + Y = z) \\
&= \sum_{k=0}^{z} \mathbb{P}(X = k, Y = z - k) = \sum_{k=0}^{z} \mathbb{P}(X = k) \cdot \mathbb{P}(Y = z - k) \\
&= \sum_{k=0}^{z} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{\mu^{z-k}}{(z-k)!} e^{-\mu} \\
&= e^{-(\lambda+\mu)} \cdot \frac{\mu^z}{z!} \cdot \sum_{k=0}^{z} \binom{z}{k} \cdot \left(\frac{\lambda}{\mu}\right)^k \cdot 1^{z-k} \\
&= e^{-(\lambda+\mu)} \cdot \frac{\mu^z}{z!} \cdot \left(\frac{\lambda}{\mu} + 1\right)^z \\
&= e^{-(\lambda+\mu)} \cdot \frac{(\lambda + \mu)^z}{z!},
\end{aligned}$$

d.h. $Z = X + Y$ ist Poisson-verteilt mit Parameter $\lambda + \mu$.

(ii) Wir berechnen zuerst die Verteilungsfunktion von Z für $z \in [0, 1]$. Für $X \sim \mathcal{R}[0, 1]$ gilt $\mathbb{P}(X \leq z) = z$.

$$\begin{aligned}
F_Z(z) = \mathbb{P}(Z \leq z) &= \mathbb{P}(\max(X, Y) \leq z) = \mathbb{P}(X \leq z, Y \leq z) \\
&= \mathbb{P}(X \leq z) \cdot \mathbb{P}(Y \leq z) = z^2.
\end{aligned}$$

Die Dichte für $z \in (0, 1)$ erhalten wir nun durch Differenzieren:

$$f_Z(z) = F_Z'(z) = 2z.$$

Da offensichtlich $Z \in [0, 1]$ gilt, muss $f_Z(z) = 0$ sein für $z \notin (0, 1)$. Wir erhalten insgesamt

$$f_Z(z) = 2z \cdot \mathbb{1}_{(0,1)}(z).$$

□

Aufgabe 2

Ein Hersteller behauptet von seinen Glühbirnen, dass diese 10 Kilostunden lang Licht spenden, ehe sie durchbrennen. Wir sind skeptisch und vermuten, dass die wahre Lebensdauer der Glühbirnen geringer ist. Für eine statistische Untersuchung nehmen wir an, dass die Lebensdauer (in Kilostunden) einer Glühbirne Exponential-verteilt ist mit unbekanntem Parameter $\lambda > 0$, wobei in dieser Modellierung (vgl. Erwartungswert einer Exponentialverteilung) dann $\frac{1}{\lambda}$ für die mittlere Lebensdauer steht. Wir kaufen $n = 10$ Glühbirnen, und erhalten folgende Lebensdauern (in 1000 Stunden):

8.2	4.4	7.3	5.1	10.2	12.6	5.5	9.3	7.1	3.8

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

$$H_0 : \lambda = \lambda_0 \quad \text{gegen} \quad H_1 : \lambda = \lambda_1$$

mit $\lambda_1 > \lambda_0$ zum Niveau $\alpha \in (0, 1)$ anzugeben.
Hinweis: Bestimmen Sie zunächst die Dichte f_λ der gemeinsamen Verteilung $\mathbb{P}_\lambda = \mathbb{P}^{X_1,...,X_n}$ von $X_1, ..., X_n$.

(b) Vereinfachen Sie ϕ^*, indem Sie zeigen, dass der Likelihood-Quotient monoton ist. Folgern Sie, dass ϕ^* sogar gleichmäßig bester Test ist für

$$H_0 : \lambda = \lambda_0 \quad \text{gegen} \quad H_1' : \lambda > \lambda_0.$$

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

$$H_0' : \lambda \leq \lambda_0 \quad \text{gegen} \quad H_1' : \lambda > \lambda_0.$$

Hinweis: Definieren Sie $Z_n := \lambda \cdot \sum_{i=1}^n X_i$. Überlegen Sie sich, dass die Verteilung von Z_n nicht mehr von λ abhängt. Zeigen Sie dann, dass $\mathbb{P}_\sigma(\phi^* = 1) \leq \alpha$ gilt für $\lambda \leq \lambda_0$.

(d) Es ist bekannt, dass $Z_n \sim \Gamma(n, 1)$, wobei $\Gamma(n, 1)$ die Gamma-Verteilung bezeichnet. Es bezeichne $\Gamma_{n,1,\alpha}$ das α-Quantil dieser Verteilung. Berechnen Sie c^{**} in Termen von λ_0 und $\Gamma_{n,1,\alpha}$.

(e) Wir drücken nun die Abhängigkeit des Tests ϕ^* von λ_0 explizit aus, indem wir ihn mit $\phi^*_{\lambda_0}$ bezeichnen. Bestimmen Sie den Annahmebereich $A(\lambda_0) := \{x \in \mathbb{R}^n : \phi^*_{\lambda_0}(x) = 0\}$ des Tests und damit ein gleichmäßig bestes $(1 - \alpha)$-Konfidenzintervall $S(X)$ für λ. Für welche falschen Parameter wurde dieses Intervall konstruiert?

(f) Sei nun $\alpha = 0.05$. Werden Sie den Hersteller der Glühbirnen 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 λ aus (e) an.
Hinweis: Hier sind einige Quantile der $\Gamma(n, 1)$-Verteilung: $\Gamma_{10,1,0.05} = 5.426$, $\Gamma_{10,1,0.95} = 15.706$.

(2+1+2+1+2+1 Punkte)

Beweis. (a) Die Dichte f_λ der gemeinsamen Verteilung $\mathbb{P}_\lambda = \mathbb{P}^{X_1,\dots,X_n}$ von $X_1, ..., X_n$ ist wegen der Unabhängigkeit von $X_1, ..., X_n$ gegeben durch

$$\begin{aligned} f_\lambda(x) &= f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n) \\ &= \prod_{i=1}^n \left(\lambda \cdot e^{-\lambda x_i} \mathbb{1}_{\{x_i \geq 0\}}(x_i)\right) \\ &= \lambda^n \cdot e^{-\lambda \cdot \sum_{i=1}^n x_i} \cdot \mathbb{1}_{\{x_1,\dots,x_n \geq 0\}}(x_1, \dots, x_n). \end{aligned}$$

Setze $x = (x_1, ..., x_n)$, wir nehmen im Folgenden $x_1, ..., x_n \geq 0$ an, da dies unabhängig vom Parameter λ für die Realisierungen $X_1, ..., X_n$ erfüllt ist. Für $\lambda_0 < \lambda_1$ ist der Likelihood-Quotient daher gegeben durch

$$L(x) = \frac{f_{\lambda_1}(x)}{f_{\lambda_0}(x)} = \left(\frac{\lambda_1}{\lambda_0}\right)^n \cdot e^{(\lambda_0 - \lambda_1) \cdot \sum_{i=1}^n x_i}. \tag{1}$$

Der beste Test zum Niveau $\alpha \in (0, 1)$ für die Hypothesen H_0, H_1 lautet daher nach dem (verallgemeinerten) Neyman-Pearson-Lemma 6.18:

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

wobei c^* bestimmt wird aus $\mathbb{P}_{\lambda_0}(\phi^* = 1) = \mathbb{P}_{\lambda_0}(L(x) > c^*)$.

(b) An der Darstellung (1) sehen wir, dass $L(x)$ für $\lambda_0 < \lambda_1$ (d.h. $\lambda_0 - \lambda_1 < 0$) monoton fallend in $T(x) := \sum_{i=1}^n x_i$ ist. Daher vereinfacht sich ϕ^* zu

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

wobei nun c^{**} bestimmt wird aus $\mathbb{P}_{\lambda_0}(\phi^* = 1) = \mathbb{P}_{\lambda_0}(T(x) < c^{**})$. Da der gesamte Test ϕ^* und insbesondere die Bestimmungsgleichung für c^{**} nicht mehr von λ_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 $\lambda \leq \lambda_0$ gilt:

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

Seien nun wieder $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathcal{E}(\lambda)$. Dann sind $\lambda X_i \sim \mathcal{E}(1)$ exponentialverteilte und unabhängige Zufallsvariablen. Das kann einfach mit der lineare Transformation nachgerechnet werden. Ist $X \sim p_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{\{x \geq 0\}}$, dann gilt für $Y := \lambda X$: $p_Y(y) = \frac{1}{\lambda} p_Y(\frac{y}{\lambda}) = e^{-y} \mathbb{1}_{\{y \geq 0\}}(y)$. Das ist genau die Dichte einer $\mathcal{E}(1)$-Verteilung. Entsprechend ist die Verteilung von $Z_n = \sum_{i=1}^n (\lambda X_i)$ unabhängig von λ. Denn die gemeinsame Verteilung von λX_i, $i = 1, ..., n$ hängt nicht mehr von λ ab und Z_n ist nur eine Funktion dieser n Zufallsvariablen. Wir erhalten mit $c^{**} \geq 0$, dass

$$\begin{aligned} \mathbb{P}_\sigma(\phi^* = 1) &= \mathbb{P}\left(\frac{1}{\lambda} \cdot Z_n < c^{**}\right) = \mathbb{P}(Z_n < \lambda c^{**}) \leq \mathbb{P}(Z_n < \lambda_0 c^{**}) \\ &= \mathbb{P}(\sigma_0^2 \cdot Z_n < c^{**}) = \mathbb{P}_{\sigma_0}(\phi^* = 1). \end{aligned}$$

(d) Wir müssen nun zunächst c^{**} explizit bestimmen. Dazu betrachten wir die Bestimmungsgleichung und versuchen, die angegebene Information über Z_n zu verwerten. Für $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathcal{E}(\lambda_0)$ soll folgendes erfüllt sein.

$$\alpha = \mathbb{P}_{\lambda_0}(T(x) < c^{**}) = \mathbb{P}\left(\frac{1}{\lambda_0} Z_n < c^{**}\right) = \mathbb{P}(Z_n < \lambda_0 c^{**}),$$

d.h. es muss gelten $\lambda_0 c^{**} = \Gamma_{n,1,\alpha}$ bzw. $c^{**} = \frac{\Gamma_{n,1,\alpha}}{\lambda_0}$.

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

$$\phi^*_{\lambda_0}(x) = \begin{cases} 1, & T(x) < \frac{\Gamma_{n,1,\alpha}}{\lambda_0}, \\ 0, & T(x) \geq \frac{\Gamma_{n,1,\alpha}}{\lambda_0} \end{cases}$$

für die Hypothesen $H_0' : \lambda \leq \lambda_0$ gegen $H_1' : \lambda > \lambda_0$ konstruiert. Der Annahmebereich des Tests lautet demzufolge

$$A(\lambda_0) = \{x \in \mathbb{R}^n : \phi^*_{\lambda_0}(x) = 0\} = \left\{x \in \mathbb{R}^n : T(x) \geq \frac{\Gamma_{n,1,\alpha}}{\lambda_0}\right\}.$$

Für das Konfidenzintervall für λ aus Satz 9.3 erhalten wir:

$$\begin{aligned} S(X) &= \{\lambda > 0 : X \in A(\lambda)\} = \left\{\lambda > 0 : T(X) \geq \frac{\Gamma_{n,1,\alpha}}{\lambda}\right\} \\ &= \left[\frac{\Gamma_{n,1,\alpha}}{T(X)}, \infty\right) = \left[\frac{\Gamma_{n,1,\alpha}}{\sum_{i=1}^n X_i}, \infty\right). \end{aligned}$$

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

$$H_0' : \lambda \in H(\lambda_0) \quad \text{gegen} \quad H_1' : \lambda \in K(\lambda_0),$$

mit $H(\lambda_0) = (0, \lambda_0]$ und $K(\lambda_0) = (\lambda_0, \infty)$. Die falschen Parameter berechnen sich durch

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

Dies folgt auch aus Bemerkung 9.5. Das heißt, bei der Konstruktion des Konfidenzintervalls waren wir daran interessiert, eine Untergrenze für den Parameter λ zu erhalten, und haben daher kleine λ als falsche Parameter deklariert.

(f) Für die konkreten Beobachtungen aus der Aufgabenstellung erhalten wir für $n = 10$:

$$T(X) = \sum_{i=1}^n X_i = 73.5,$$

und mit $\lambda_0 = \frac{1}{10}$ (Man beachte, dass die mittlere Lebensdauer hier $\frac{1}{\lambda_0} = 10$ ist.), $\alpha = 0.05$, sowie $\Gamma_{n,1,\alpha} = \Gamma_{10,1,0.05} = 5.426$ ergibt sich

$$\frac{\Gamma_{10,1,0.05}}{\lambda_0} = 54.26.$$

Damit gilt für unsere Beobachtungen $T(X) > \frac{\Gamma_{10,1,0.05}}{\lambda_0}$, d.h. $\phi^*(X) = 0$. Wir verwerfen die Nullhypothese also nicht. Das bedeutet, unsere Beobachtungen genügen nicht, um die Behauptungen des Herstellers als falsch zu brandmarken.
Für das 95%-Konfidenzintervall für λ aus (e) erhalten wir:

$$S(X) = \left[\frac{\Gamma_{n,1,\alpha}}{\sum_{i=1}^n X_i}, \infty\right) = [0.074, \infty).$$

□

Aufgabe 3

Seien $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma_0^2)$ normalverteilt mit Parameter $\mu \in \mathbb{R}$ und festem $\sigma_0^2 = 5$, und $\mathbb{P}_\mu := \mathbb{P}^{X_1,...,X_n}$ die gemeinsame Verteilung dieser Zufallsvariablen.

(a) Sei $\mu_0 \in \mathbb{R}$. Zeigen Sie, dass

$$\phi : \mathbb{R}^n \to \{0,1\}, \quad \phi(x_1, ..., x_n) = \begin{cases} 1, & |\mu_0 - \frac{1}{n}\sum_{i=1}^n x_i| > \frac{\sigma_0}{\sqrt{n}} \cdot q_{1-\frac{\alpha}{2}}, \\ 0, & \text{sonst} \end{cases}$$

ein Test zum Niveau $\alpha \in (0,1)$ für die Hypothesen

$$H_0 : \mu = \mu_0 \quad \text{gegen} \quad H_1 : \mu \neq \mu_0$$

ist, wobei $q_{1-\frac{\alpha}{2}}$ das $(1-\frac{\alpha}{2})$-Quantil der Standardnormalverteilung bezeichnet.

(b) Konstruieren Sie mittels Satz 9.3 aus ϕ ein $(1-\alpha)$-Konfidenzintervall $S(X)$ für μ.

(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?

(2+1+1 Punkte)

Beweis. (a) Da die Nullhypothese H_0 nur aus einem Wert, nämlich μ_0, besteht, ist nur zu zeigen, dass $\mathbb{P}_{\mu_0}(\phi^* = 1) \leq \alpha$. Seien dazu $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu_0, \sigma_0^2)$ und $\overline{X_n} := \frac{1}{n}\sum_{i=1}^n X_i$ der Mittelwert. Dann ist bekannt (siehe Vorlesung: Anwendung 8.19), dass

$$Z_n := \sqrt{n} \cdot \frac{\overline{X_n} - \mu_0}{\sigma_0} \sim \mathcal{N}\sqsupseteq(0,1).$$

Beachte, dass $q_{1-\frac{\alpha}{2}} \geq 0$, da $\alpha \in (0,1)$ und somit $1 - \frac{\alpha}{2} > 0.5$, und $\Phi(0) = 0.5$ gilt. Damit erhalten wir

$$\begin{aligned}
\mathbb{P}_{\mu_0}(\phi^* = 1) &= \mathbb{P}(|\mu_0 - \overline{X_n}| \ > \frac{\sigma_0}{\sqrt{n}} \cdot q_{1-\frac{\alpha}{2}}) \\
&= \mathbb{P}(|Z_n| > q_{1-\frac{\alpha}{2}}) = \mathbb{P}(Z_n > q_{1-\frac{\alpha}{2}}) + \mathbb{P}(Z_n < -q_{1-\frac{\alpha}{2}}) \\
&= (1 - \Phi(q_{1-\frac{\alpha}{2}})) + \Phi(-q_{1-\frac{\alpha}{2}}) \\
&= 2 - 2 \cdot \Phi(q_{1-\frac{\alpha}{2}}) = 2 - 2 \cdot (1 - \frac{\alpha}{2}) = \alpha.
\end{aligned}$$

Die Beziehung $\Phi(-z) = 1 - \Phi(z)$ gilt, da die Dichte einer Standardnormalverteilung $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ symmetrisch um 0 ist, und daher ist deren Stammfunktion punktsymmetrisch um $(0, \Phi(0)) = (0, \frac{1}{2})$.

(b) Der Annahmebereich des Tests ist für festes $\mu_0 \in \mathbb{R}$

$$A(\mu_0) = \left\{ x \in \mathbb{R}^n : \left|\mu_0 - \frac{1}{n}\sum_{i=1}^n x_i\right| \leq \frac{\sigma_0}{\sqrt{n}} \cdot q_{1-\frac{\alpha}{2}} \right\},$$

damit lautet ein $(1-\alpha)$-Konfidenzintervall für μ laut Satz 9.3 der Vorlesung:

$$\begin{aligned}
S(X) &= \{\mu \in \mathbb{R} : X \in A(\mu)\} = \left\{ \mu \in \mathbb{R} : \left|\mu - \overline{X_n}\right| \leq \frac{\sigma_0}{\sqrt{n}} \cdot q_{1-\frac{\alpha}{2}} \right\} \\
&= \left[\overline{X_n} - \frac{\sigma_0}{\sqrt{n}} q_{1-\frac{\alpha}{2}}, \overline{X_n} + \frac{\sigma_0}{\sqrt{n}} \cdot q_{1-\frac{\alpha}{2}} \right].
\end{aligned}$$

Sind $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma_0^2)$, so enthält $S(X)$ den Parameter μ also mit einer Wahrscheinlichkeit von mindestens $1 - \alpha$.

(c) Die Länge von $S(X)$ ist

$$L(n) = \left| \left(\overline{X_n} + \frac{\sigma_0}{\sqrt{n}} \cdot q_{1-\frac{\alpha}{2}} \right) - \left(\overline{X_n} - \frac{\sigma_0}{\sqrt{n}} q_{1-\frac{\alpha}{2}} \right) \right| = \frac{2\sigma_0}{\sqrt{n}} \cdot q_{1-\frac{\alpha}{2}}$$

Für $\alpha = 0.1$ gilt nach der Quantiltabelle $q_{1-\frac{\alpha}{2}} = q_{0.95} = 1.645$, und mit $\sigma_0^2 = 5$ erhalten wir

$$1 \geq L(n) = \frac{7.357}{\sqrt{n}} \quad \Leftrightarrow \quad n \geq 7.357^2 = 54.13,$$

d.h. es muss mindestens $n = 55$ gelten.

$\square$