# Wahrscheinlichkeitstheorie und Statistik

Sommersemester 2020
FSU Jena

Prof. Schmalfuß
Stefan Engelhardt

## Hinweise und Musterlösung 5. Übungsblatt

### Hinweise

**Aufgabe 2.** (a) Erwartungswert ausrechnen oder Gegenbsp finden.

**Aufgabe 3.** (a) Tabelle und transformation

(b) in bekannte Formel einsetzen

**Aufgabe 4.** (a) Wie in 3.a) nur mit Studentverteilung und dann vergleichen.

(b) Nochmal das gleiche.

(c) $\beta$-Fehler mit Normalverteilung. Kann auch von Hand nachgerechnet werden (inverse Rechnung zum $\alpha$-Fehler)

(d) Verwende c) und Ausprobieren.

Musterlösungen

**Aufgabe 1.** Es ist bekannt, dass die Differentialgleichung

$$f'(t) = a(t) \cdot f(t) + g(t), \qquad f(0) = f_0,$$

die eindeutige Lösung

$$f(t) = f_0 \exp\left(\int_0^t a(r)\mathrm{d}r\right) + \int_0^t \exp\left(\int_r^t a(q)\mathrm{d}q\right) \cdot g(r)\ \mathrm{d}r$$

hat. Somit ergibt sich jeweils für

$$\begin{aligned}
P_3'(t) = -2P_3(t), P_3(0) = 1 \quad &\Rightarrow \quad P_3(t) = \exp\left(\int_0^t -2\mathrm{d}r\right) = e^{-2t} \\
P_2'(t) = -3P_2(t) + 2e^{-2t}, P_2(0) = 0 \quad &\Rightarrow \quad P_2(t) = \int_0^t \exp\big(-3(t-r)\big)2e^{-2r}\mathrm{d}r \\
&\qquad\qquad\;\; = 2e^{-3t}\left(e^t - 1\right) \\
P_1'(t) = -2P_1(t) + 3 \cdot 2e^{-3t}\left(e^t - 1\right), P_1(0) = 0 \quad &\Rightarrow \quad P_1(t) = \int_0^t \exp\big(-2(t-r)\big)6e^{-3r}\left(e^r - 1\right)\mathrm{d}r \\
&\qquad\qquad\;\; = 6e^{-2t}\left(t + e^{-t} - 1\right) \\
P_0'(t) = 2 \cdot 6e^{-2t}\left(t + e^{-t} - 1\right), P_0(0) = 0 \quad &\Rightarrow \quad P_0(t) = \int_0^t 12e^{-2r}\left(r - e^{-r} + 1\right)\mathrm{d}r \\
&\qquad\qquad\;\; = e^{-2t}(3 - 6t) - 4e^{-3t} + 1.
\end{aligned}$$

Probe:

$$\begin{aligned}
P_3(t) + P_2(t) + P_1(t) + P_0(t) &= e^{-2t} + 2e^{-3t}\left(e^t - 1\right) + 6e^{-2t}\left(t + e^{-t} - 1\right) + e^{-2t}(3 - 6t) - 4e^{-3t} + 1 \\
&= 1 + e^{-2t}(1 + 2 + 6t - 6 + 3 - 6t) + e^{-3t}(-2 + 6 - 4) \\
&= 1
\end{aligned}$$

**Aufgabe 2.** a) Nein, denn: $\mathbb{E}\frac{1}{n}\sum_{i=1}^n X_i + \frac{1}{n} = \frac{1}{n}\sum_{i=1}^n \mathbb{E}X_i + \frac{1}{n} = \mathbb{E}X + \frac{1}{n}$.

b) Ja, denn: $\mathbb{E}X_n = \mathbb{E}X$.

c) Ja, denn: $\mathbb{E}\frac{1}{n-1}\sum_{i=1}^{n-1} X_i = \frac{1}{n-1}\sum_{i=1}^{n-1}\mathbb{E}X_i = \frac{n-1}{n-1}\mathbb{E}X = \mathbb{E}X$.

d) Nein, denn: Betrachte $X$ mit $\mathbb{P}(X = -1) = \mathbb{P}(X = 1) = 0.5$. Dann
$\mathbb{E}\sqrt{\sum_{i=1}^n X_i^2} = \mathbb{E}\sqrt{\sum_{i=1}^n 1} = \sqrt{n} \neq 0 = \mathbb{E}X$.

Zur Erinnerung: Quantile sind für eine (hübsche) Verteilungsfunktion $F$ wie folg definiert: Sei $\alpha$ eine Wkt, also in $\alpha \in [0,1]$, dann ist das $\alpha$-Quantil der Wert $x \in \mathbb{R}$ für den gilt, dass $F(x) = \alpha$. Somit ist das $\alpha$-Quantil $q_\alpha$ (wird manchmal auch anders benannt) der Wert, für den gilt, dass $F(q_\alpha) = \alpha$.

**Aufgabe 3.** a) $\alpha = \Phi_{1,4}(x) = \mathbb{P}(N_{1,4} \leq x) = \mathbb{P}(N_{0,1} \leq \frac{x-1}{\sqrt{4}}) = \Phi(\frac{x-1}{2}) = 1 - \Phi(-\frac{x-1}{2})$. Für $\alpha = 0.05$ erhalten wir mit der Tabelle, dass $-\frac{x-1}{2} = 1.64$ und somit $x = -2.28$ also $q_{0.05} = -2.28$. Ebenso erhalten wir für $\alpha = 0.975$, dass $\frac{x-1}{2} = 1.96$ und somit $x = 4.92$ also $q_{0.975} = 4.92$.

b) Aus der Vorlesung ist bekannt, dass für ein Zweiseitiges Konfidenzintervall mit bekannter Varianz gilt:

$$\mathbb{P}\left(\mu \in \left(\bar{x} - q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)\right) = 1 - \alpha,$$

wobei $\bar{x} = \frac{1}{n}\sum_{i=1}^n X_i$ und $q_{1-\frac{\alpha}{2}}$ das $(1 - \frac{\alpha}{2})$-Quantil der Normalverteilung ist. Da $\alpha = 1 - 0.95 = 0.05$ erhalten wir $q_{1-\frac{\alpha}{2}} = 1.96$ und damit das Konfidenzinterval $\left(\bar{x} - q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = \left(50.1 - 1.96 \cdot \frac{4}{5}, 50.1 + 1.96 \cdot \frac{4}{5}\right) = (48.532, 51.668)$

**Aufgabe 4.** In der Aufgabe war ein Schreibfehler, dass die Stichprobe aus 20 Elementen besteht. Es waren 25 gemeint, zum einfacherern Rechnen und deswegen waren auch nur hierfür die Werte aus der Tabelle angegeben. Im Folgenden rechne ich nun beiden durch.

a) Da ein einseitiges Konfidenzinterval gesucht ist, ist dieses gegeben durch $\left(\mu_0 - t_{1-\alpha,n-1}\frac{s}{\sqrt{n}}, \infty\right)$. Somit erhalten wir für $\alpha = 0.05$ und $n = 20$ das Intervall $\left(15 - 1.729 \cdot \frac{3}{\sqrt{20}}, \infty\right) \approx (13.84, \infty)$. Für $\alpha = 0.05$ und $n = 25$ ergibt sich das Intervall $\left(15 - 1.711 \cdot \frac{3}{5}, \infty\right) \approx (13.97, \infty)$. Da $\bar{x} = 13.5$ in beiden Fällen nicht dem Konfidenzintervall liegt, würde $H_0$ abgelehnt werden.

b) Für $\alpha = 0.005$ erhalten wir für $n = 20$: $\left(15 - 2.861 \cdot \frac{3}{\sqrt{20}}, \infty\right) \approx (13.08, \infty)$ und für $n = 25$: $\left(15 - 2.797 \cdot \frac{3}{5}, \infty\right) \approx (13.3218, \infty)$, wesegen hier in beiden Fällen $H_0$ nicht abgewiesen werden würde.

c) Im Falle einer bekannten Varianz, ist der $\beta$-Fehler gegeben durch

$$\begin{aligned}
\beta &= \mathbb{P}_{\mu_a}\left(\bar{x} \in K^C\right) \\
&= \mathbb{P}_{\mu_a}\left(\bar{x} \in \left(\mu_0 - q_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}, \infty\right)\right) \\
&= \mathbb{P}_{\mu_a}\left(\frac{\frac{1}{n}\sum_{i=1}^n X_i - \mu_a}{\sigma}\sqrt{n} > \frac{\mu_0 - \mu_a}{\sigma}\sqrt{n} - q_{1-\alpha}\right) \\
&= 1 - \Phi\left(\frac{\mu_0 - \mu_a}{\sigma}\sqrt{n} - q_{1-\alpha}\right).
\end{aligned}$$

Für $n = 20$ ergibt sich somit $\beta = 1 - \Phi\left(\frac{15-12.5}{3}\sqrt{20} - 2.58\right) \approx \Phi(1,14678) \approx 0.12508$ und für $n = 25$: $\beta = 1 - \Phi\left(\frac{15-12.5}{3}5 - 1.64\right) \approx \Phi(1.5867) \approx 0.05592$.

Wenn wir die Voraussetzung der bekannten Varianz nicht hätten, ergäbe sich

$$\begin{aligned}
\beta &= \mathbb{P}_{\mu_a}\left(\bar{x} \in K^C\right) \\
&= \mathbb{P}_{\mu_a}\left(\bar{x} \in \left(\mu_0 - t_{1-\alpha,n-1} \cdot \frac{s}{\sqrt{n}}, \infty\right)\right) \\
&= \mathbb{P}_{\mu_a}\left(\frac{\frac{1}{n}\sum_{i=1}^n X_i - \mu_a}{s}\sqrt{n} > \frac{\mu_0 - \mu_a}{s}\sqrt{n} - t_{1-\alpha,n-1}\right) \\
&= 1 - T\left(n - 1, \frac{\mu_0 - \mu_a}{s}\sqrt{n} - t_{1-\alpha,n-1}\right).
\end{aligned}$$

und des Weiteren für $n = 20$: $\beta = 1 - T\,(19, 3.7268 - 2.861) = 1 - T\,(19, 0.8658) \approx 0.1987$ und für $n = 25$: $\beta = 1 - T\,(24, 3.7268 - 2.797) = 1 - T\,(24, 0.9298) \approx 0.1809$.

d) Wir stellen die Gleichung für den $\beta$-Fehler nach $n$ um und erhalten so:

$$n = \left(\frac{q_{1-\beta} + q_{1-\alpha}}{\mu_0 - \mu_a}\right)^2 \sigma^2.$$

Einsetzen ergibt $\left(\frac{1.64+2.58}{15-12.5}\right)^2 9 = 25,64$. Für $n = 26$ wäre der $\beta$-Fehler also kleiner als 0.05.

**Aufgabe 5.** Wie in der Vorlesung definieren wir

$$X := \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \qquad y := \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 2 \end{pmatrix}.$$

Hierfür ist bekannt, dass die die optimale Gerade gegeben ist durch $g(x) = \beta_0 + \beta_1 x$, wobei

$$\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \left(X^T X\right)^{-1} X^T y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Somit ist die im Sinne der kleinsten Quadrate am besten passende Gerade $g(x) = x$.