Hausaufgabenblatt 14

Aufgabe 1:

a) $\Omega = \{(\omega_1, \omega_2, \omega_3) \mid \omega_i \in \{0,1\} \ \forall i \in \{1,2,3\}\}$

$\mathcal{A} = \mathcal{P}(\Omega)$

mit $\bar{p} = (\bar{p}_1, \bar{p}_2, \bar{p}_3)$, $\bar{p}_i \in]0,1[$

und $\mathcal{P}_{\bar{p}}(\omega) = \bar{p}_1^{\omega_1} \cdot (1-\bar{p}_1)^{1-\omega_1} \cdot \bar{p}_2^{\omega_2} \cdot (1-\bar{p}_2)^{1-\omega_2} \cdot \bar{p}_3^{\omega_3} \cdot (1-\bar{p}_3)^{1-\omega_3}$

definieren das statistische Modell.

Es folgt für den Likelihood-Quotienten: $(q = (p_1, p_2, p_3))$

$$\varrho(\omega) = \frac{\mathcal{P}_{2q}(\omega)}{\mathcal{P}_q(\omega)} = \frac{0{,}20^{\omega_1} \cdot 0{,}40^{\omega_2} \cdot 0{,}80^{\omega_3} \cdot 0{,}80^{1-\omega_1} \cdot 0{,}60^{1-\omega_2} \cdot 0{,}2^{1-\omega_3}}{0{,}10^{\omega_1} \cdot 0{,}20^{\omega_2} \cdot 0{,}40^{\omega_3} \cdot 0{,}90^{1-\omega_1} \cdot 0{,}80^{1-\omega_2} \cdot 0{,}60^{1-\omega_3}} =$$

$$= 2^{\omega_1+\omega_2+\omega_3} \cdot \left(\frac{8}{9}\right)^{1-\omega_1} \cdot \left(\frac{3}{4}\right)^{1-\omega_2} \cdot \left(\frac{1}{3}\right)^{1-\omega_3}$$

Es folgt weiter: $\varrho((1,1,1)) > \varrho((0,1,1)) > \varrho((1,0,1)) > \varrho((1,1,0)) \geq \varrho((0,0,1)) >$

$> \varrho((0,1,0)) > \varrho((1,0,0)) > \varrho((0,0,0))$

Setze nun:
$$\varphi(\omega) = \begin{cases} 1 & \text{falls } \omega \in \{(1,1,1), (0,1,1)\} \\ c & \text{falls } \omega \in \{(1,0,1)\} \\ 0 & \text{sonst} \end{cases}$$

Es folgt: $\mathbb{E}_{\mathcal{P}_q}(\varphi) = 0{,}1 \cdot 0{,}2 \cdot 0{,}4 + 0{,}2 \cdot 0{,}4 \cdot 0{,}9 + c \cdot 0{,}10 \cdot 0{,}40 \cdot 0{,}80 \overset{!}{=} 0{,}10$

$$\Rightarrow \quad c = \frac{0{,}10 - 0{,}008 - 0{,}072}{0{,}032} = 0{,}625.$$

Dies ist nach dem Neyman-Pearson-Lemma ein Test mit maximaler Macht. (Entscheidungsregel φ (mit $c = 0{,}625$)).

Es folgt für die Macht:

$\mathbb{E}_{\mathcal{P}_{2q}}(\varphi) = 0{,}2 \cdot 0{,}4 \cdot 0{,}8 + 0{,}4 \cdot 0{,}8 \cdot 0{,}8 + 0{,}625 \cdot 0{,}2 \cdot 0{,}8 \cdot 0{,}6 = 0{,}38 = 38\%$

b) Sei jetzt:

$\Omega = \{0,1,2,3\}$

$\mathcal{A} = \mathcal{P}(\Omega)$

$$\mathcal{P}_q = 0{,}432\ \delta_0 + (0{,}048 + 0{,}108 + 0{,}288)\ \delta_1 + (0{,}012 + 0{,}032 + 0{,}072)\ \delta_2 + $$
$$+ 0{,}008\ \delta_3 .$$

und:

$$\mathcal{P}_{2q} = 0{,}096\ \delta_0 + (0{,}384 + 0{,}064 + 0{,}024)\ \delta_1 + (0{,}016 + 0{,}096 + 0{,}256)\ \delta_2 +$$
$$+ 0{,}064\ \delta_3 .$$

Jetzt werden also aus Einfachheitsgründen gleich die Verteilungen der Nullhypothese und Alternativhypothese angeschaut, da andere Verteilungen nicht benötigt werden.
Diese errechnen sich aus der Wahrscheinlichkeit, dass Fritz unter dem jeweiligen Maß $i \in \{0,1,2,3\}$ Aufgaben löst.
Es folgt für den Likelihood-Quotienten:

$$\frac{d\mathcal{P}_{2q}}{d\mathcal{P}_q} = 0{,}\overline{2}\ \mathbb{1}_{\{0\}} + \frac{0{,}472}{0{,}444}\ \mathbb{1}_{\{1\}} + \frac{0{,}368}{0{,}116}\ \mathbb{1}_{\{2\}} + 8 \cdot \mathbb{1}_{\{3\}}$$

Setze nun:
$$\varphi(\omega) = \begin{cases} 1 & \text{falls } \omega = 3 \\ c & \text{falls } \omega = 2 \\ 0 & \text{sonst.} \end{cases}$$

Es folgt: $\mathbb{E}_{\mathcal{P}_q}(\varphi) = 0{,}008 + c_2 \cdot 0{,}116 \overset{!}{=} 0{,}1$

$$c_2 = \frac{0{,}1 - 0{,}008}{0{,}116} = \frac{0{,}092}{0{,}116}$$

Dies ist nach dem Neyman-Pearson-Lemma ein Test mit maximaler Macht.
(Entscheidungsregel φ (mit $c = \frac{0{,}092}{0{,}116}$)).

Es folgt für die Macht:

$$\mathbb{E}_{\mathcal{P}_q}(\varphi) = 0{,}064 \cdot \frac{0{,}092}{0{,}116} \cdot 0{,}368 \approx 0{,}3558 \approx 35{,}6\,\%.$$

Die Macht diesen Tests ist kleiner, was klar scheint, da weniger Information vorhanden ist.

c) Bei Test a) wird die Alternativ-Hypothese, dass Fritz die Aufgaben mit doppelt so hoher W'keit richtig beantwortet, abgelehnt.

Das Ergebnis ist erstaunlich, da er die beiden schwierigsten Aufgaben richtig beantwortete, ist aber aufgrund der falschen Antwort in Aufgabe 3 zu erklären.

Beim Test in b) fällt die Entscheidung der Ablehnung der Nullhypothese in einem weiteren Zufallsexperiment. Mit W'keit $\frac{0{,}094}{0{,}116}$ soll sie abgelehnt werden mit W'keit $1-\frac{0{,}094}{0{,}116}$ nicht.

d.h. Mit hoher W'keit wird in diesem Test die Alternativhypothese angenommen und behauptet das Fritz doppelt so gut ist wie der Durchschnitt.

Dies liegt auch daran, dass dieser Test nicht unterscheiden kann, welche der 3 Aufgaben richtig beantwortet wurden.

Aufgabe 2:

a) statistisches Modell: $\Omega = [a,b]^{10}$

$\mathcal{A} = \mathcal{B}(\Omega)$

Mit $\omega = (\omega_1, \dots, \omega_{10})$ $\quad \mathcal{P}_{a,b}(\omega) = \prod_{i=1}^{10} \frac{1}{b-a} \mathbb{1}_{[a,b]}(\omega_i)$

Es folgt mit $\bar{\omega} = (10.55, 14.9, \dots 12.95)$ [wie in der Angabe]

$\mathcal{P}_{a,b}(\bar{\omega}) = 0$ falls $a > \min\{\bar{\omega}_i \mid i \in \{1 \dots 10\}\}$

oder $b < \max\{\bar{\omega}_i \mid i \in \{1, \dots 10\}\}$

Da das $\arg\max_{a,b} \mathcal{P}_{a,b}(\omega)$ bestimmt werden soll, muss außerdem $(b-a)$ minimiert werden.

Es folgt: $\hat{b} = \max\{\bar{\omega}_i \mid i \in \{1, \dots, 10\}\} = 18.85$

$\hat{a} = \min\{\bar{\omega}_i \mid i \in \{1, \dots, 10\}\} = 9.2$

b) Sei X nun gleichverteilt auf dem Intervall $[a,b]$, $a<b$, $a,b\in\mathbb{R}$.

Es folgt:

$$\mathbb{E}[X] = \int_{\mathbb{R}} x \cdot \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)\, dx = \frac{1}{b-a}\left[\frac{1}{2}x^2\right]_a^b = \frac{b^2-a^2}{2\cdot(b-a)} = \frac{1}{2}(a+b)$$

$$\mathbb{E}[X^2] = \int_{\mathbb{R}} x^2 \cdot \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)\, dx = \frac{1}{b-a}\left[\frac{1}{3}x^3\right]_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ab+a^2}{3}$$

$$\mathbb{V}ar[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{b^2+ab+a^2}{3} - \frac{a^2+2ab+b^2}{4} = \frac{1}{12}(a-b)^2$$

Weiter gilt: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}\cdot(10.55 + 14.9 + \ldots + 12.95) = 12.94 =: c_1$

$$\bar{s} = \frac{1}{n-1}\sum_{i=1}^{n} x_i = \frac{1}{9}\left[(10.55-12.94)^2 + (14.9-12.94)^2 + \ldots + (12.95-12.94)^2\right] =$$

$$\approx 10.6499556 =: c_2$$

Jetzt muss gelten: (1) $\frac{1}{2}(a+b) = c_1$

(2) $\frac{1}{12}(a-b)^2 = c_2$

$\Rightarrow$ (1') $a = 2c_1 - b$ in (2): $\frac{1}{12}(2c_1-2b)^2 = c_2$

$$\Leftrightarrow (c_1-b)^2 = 3c_2$$

$$\Leftrightarrow |c_1-b| = \sqrt{3c_2}$$

$$\Leftrightarrow b_1 = c - \sqrt{3c_2} \vee b_2 = c_1 + \sqrt{3c_2}$$

$\Rightarrow$ aus (1'): $a_1 = c_1 + \sqrt{3c_2}$ und $a_2 = c_1 - \sqrt{3c_2}$

Da $a<b$ folgt: $\hat{a} = c_1 - \sqrt{3c_2} \approx 7{,}2876$

$\hat{b} = c_1 + \sqrt{3c_2} \approx 18{,}5924$

Die Schätzung in a) dürfte das Intervall zu klein schätzen, aber das Ergebnis ist plausibel. Die Schätzung in b) hingegen nicht, da $18.5924 < 18.85$!

Aufgabe 3:

a) $\Omega = \{0,1\}$

$\mathcal{A} = \mathcal{P}(\Omega)$

$\mathcal{P}_p(\omega) = p^{\omega} \cdot (1-p)^{1-\omega}$ mit $\omega \in \Omega$.

beschreibt das statistische Modell [Deutung: '1' Frage zum 'N-P'-Lemma wird gestellt.

'0' " " wird nicht gestellt]

b) Dies ist ein Binomialmodell für $n=1$.

Es folgt für den Maximum-Likelihoodschätzer nach Vorlesung:

$\hat{p}_{ML}(\omega) = \frac{\omega}{1} = \omega$

Aufgrund Annos Bericht gilt $\omega = 1$.

$\Rightarrow \hat{p}_{ML}(1) = 1$.

Es wird also $p=1$ geschätzt.

c) Es muss gelten: $\mathcal{P}_p(1: |1-p| \geq \varepsilon) \leq 1-0{,}95$

$\Leftrightarrow \mathcal{P}_p(1: 1-p \geq \varepsilon) \leq 0{,}05$

$\Leftrightarrow \mathcal{P}_p(1: p \leq 1-\varepsilon) \leq 0{,}05$

Nun gilt weiter: $\mathcal{P}_p(1: p \leq 1-\varepsilon) \leq 1-\varepsilon \leq 0{,}05$

$\Rightarrow \varepsilon \geq 0{,}95$

$\Rightarrow [0{,}05; 1]$ ist ein 95% Konfidenzintervall für p.

Aufgabe 4:

a) Sei $\Omega = (\mathbb{R}^+)^{10}$

$\mathcal{A} = \mathcal{B}(\Omega)$

$P_\mu, \mu \in \mathbb{R}^+$ das Maß mit Lebesgue-Dichte:

$$f_\mu(x_1, \dots, x_{10}) = \left(\frac{1}{\mu}\right)^{10} \exp\left(-\frac{1}{\mu}(x_1 + \dots + x_{10})\right)$$

Dies gilt, da wegen der Unabhängigkeit und der gleichen Verteilung von $T_1, T_2 - T_1, T_3 - T_2, \dots, T_{10} - T_9$ für die gemeinsame Dichte von $(T_1, T_2 - T_1, \dots, T_{10} - T_9)$ folgt:

$$f_\mu(x_1, \dots, x_n) = f_\mu(x_1) \cdot f_\mu(x_2) \cdot \ldots \cdot f_\mu(x_n)$$

Weiter wurde $f_\mu(x_i) = \frac{1}{\mu} \exp\left(-\frac{1}{\mu} x_i\right) \; \forall i \in \{1, \dots 10\}$ gesetzt (also $a = \frac{1}{\mu}$), da im weiteren Verlauf immer der Erwartungswert von T_1 benötigt wird.

b) Es gilt: $T_{10} = \sum_{i=2}^{10} (T_i - T_{i-1}) + T_1 = x_1 + \dots + x_{10}$

Es folgt:
$$\frac{dP_1}{dP_0} = \frac{f_{\mu_1}(x_1, \dots, x_n)}{f_{\mu_0}(x_1, \dots, x_n)} = \frac{\left(\frac{1}{\mu_1}\right)^{10} \exp\left(-\frac{1}{\mu_1}(x_1 + \dots + x_{10})\right)}{\left(\frac{1}{\mu_0}\right)^{10} \exp\left(-\frac{1}{\mu_0}(x_1 + \dots + x_{10})\right)} = \frac{\mu_0^{10}}{\mu_1^{10}} \exp\left(-\left(\frac{1}{\mu_1} - \frac{1}{\mu_0}\right)(x_1 + \dots + x_{10})\right) =$$

$$= \frac{\mu_0^{10}}{\mu_1^{10}} \exp\left(-\left(\frac{1}{\mu_1} - \frac{1}{\mu_0}\right) T_{10}\right) = g(T_{10})$$

wobei: $g(x) = \frac{\mu_0^{10}}{\mu_1^{10}} \exp\left(-\left(\frac{1}{\mu_1} - \frac{1}{\mu_0}\right) x\right)$ offensichtlich messbar ist.

Es folgt: T_{10} ist eine suffiziente Statistik nach Vorlesung.

c) Da $0 < \mu_0 < \mu_1$ gilt: $g(x)$ ist streng monoton fallend.

Weiter gilt, da $T_{10} = \sum_{i=2}^{10} (T_i - T_{i-1}) + T_1$ und $T_1, T_2 - T_1, \dots T_{10} - T_9 \sim \text{Exponential}(\mu)$

$T_{10} \sim \text{Gamma}\left(10, \frac{1}{\mu}\right)$.

Wegen b) folgt nun weiter:

Setze: $\Omega = (\mathbb{R}^+)$

$\mathcal{A} = \mathcal{B}(\Omega)$

$\mathcal{P}_\mu,\ \mu \in \mathbb{R}^+$ ist das Maß mit Lebesgue-Dichte:

$$f_\mu(x) = \frac{\left(\frac{1}{\mu}\right)^{10}}{\Gamma(10)} x^9 e^{-\left(\frac{1}{\mu}\right)^{10} x}$$

Es genügt also dieses Modell für den Test zu verwenden.

Definiere nun $\varphi(x) := \begin{cases} 1 & \text{für } x \geq c \\ 0 & \text{sonst.} \end{cases}$

Es folgt: $$\mathbb{E}_{\mathcal{P}_{\mu_0}}[\varphi] = \underbrace{\int_c^\infty \frac{\left(\frac{1}{\mu_0}\right)^{10}}{\Gamma(10)} x^9 e^{-\left(\frac{1}{\mu_0}\right)x} dx}_{(*)} \overset{!}{=} \alpha$$

$$(*) = \left[-\frac{\left(\frac{1}{\mu_0}\right)^9}{9!} x^9 e^{-\frac{1}{\mu_0}x}\right]_c^\infty + \int_c^\infty \frac{\left(\frac{1}{\mu_0}\right)^9}{\Gamma(9)} x^8 e^{-\frac{1}{\mu_0}x} =$$

$$= \left[0 + \frac{\left(\frac{1}{\mu_0}\right)^9}{9!} c^9 e^{-\frac{1}{\mu_0}c}\right] + \int_c^\infty \frac{\left(\frac{1}{\mu_0}\right)^9}{9!} x^8 e^{-\frac{1}{\mu_0}x} =$$

Schritt iterieren!

$$\overset{\downarrow}{=} \sum_{k=0}^{9} \frac{\left(\frac{1}{\mu_0}\right)^k}{k!} c^k e^{-\frac{1}{\mu_0}c} \underset{(**)}{\overset{!}{=}} \alpha$$

Nun wird c so bestimmt, dass $(**)$ gilt. [wohl nicht allgemein auflösbar!]

Dies ist nach dem Neyman-Pearson-Lemma ein Test mit maximaler Macht.

d.h. Falls $T_{10} \geq c$ wird die Nullhypothese abgelehnt.

Falls $T_{10} < c$ wird sie nicht abgelehnt.

d) $(**)$ zeigt sofort, dass der Test nicht von der Wahl von μ_1 abhängt, da μ_1 in der Gleichung nicht auftaucht.

e) Es gilt nach Vorlesung:

$$p(\omega) = P_{\mu_0}\left[T_{10} \geq T_{10}(\omega)\right] = P_5\left[T_{10} \geq 10\right] = \int_{10}^{\infty} \frac{\left(\frac{1}{5}\right)^{10}}{\Gamma(10)} x^9 e^{-\left(\frac{1}{5}\right)x} dx =$$

$$= \sum_{k=0}^{9} \frac{\left(\frac{1}{5}\right)^k}{k!} 10^k e^{-\frac{1}{5}\cdot 10} \approx 1{,}00$$

Es wird folglich die Nullhypothese nicht abgelehnt. (auf dem 5% Niveau).

f) Es gilt nach Vorlesung für das Konfidenzintervall:

$$P_{\mu_0}\left(T_{10} \in \Omega : \mu_0 \in C(T_{10})\right) \geq 1 - 0{,}05 = 0{,}95$$

$$P_{\mu_0}\left(T_{10} \in \Omega : \mu_0 \in \left]0, a(T_{10})\right[\right) \geq 0{,}95$$

$$\Rightarrow \int_{T_{10}}^{\infty} \left(\frac{1}{a(T_{10})}\right)^k \cdot \frac{1}{k!} x^k \cdot e^{-\frac{1}{a(T_{10})}x} dx \geq 0{,}95$$

$$\sum_{k=0}^{9} \frac{\left(\frac{1}{a(T_{10})}\right)^k}{k!} T_{10}^k e^{-\frac{1}{a(T_{10})}\cdot T_{10}} \geq 0{,}95$$

Mit: $z = \frac{T_{10}}{a(T_{10})}$

$$\sum_{k=0}^{9} \frac{z^k}{k!} e^{-z} \geq 0{,}95$$

Außerdem soll z möglichst groß gewählt werden, dass $a(T_{10}) = \frac{T_{10}}{z}$ möglichst klein wird.

Mit dem Computer kann man bestimmen: $\sum_{k=0}^{9} \frac{\bar{z}^k}{k!} e^{-\bar{z}} \approx 0{,}95$ für $\bar{z} = 5{,}426$.

$\Rightarrow$ Konfidenzintervall $= \left]0, \frac{T_{10}}{5{,}426}\right[$

g) $\mu_0 = 5$ liegt nicht in diesem Konfidenzintervall, da $1 - p(\omega) < 0{,}05$!