Stochastik - Musterlösung (BSc D-MAVT / BSc D-MATH / BSc D-MATL)

1. a) 1 b) 2 c) 3 d) 1 e) 1 f) 1 g) 3 h) 2 i) 1 j) 3

2. Wir definieren die folgenden Ereignisse:

$Z_A \ldots$ Zusage Anna,
$E_A \ldots$ Erscheinen Anna,

sowie analog

$Z_B \ldots$ Zusage Bernadett,
$E_B \ldots$ Erscheinen Bernadett.

Aus der Angabe ergibt sich

$P[Z_A] = \frac{6}{10} \quad \Rightarrow \quad P[Z_A^c] = \frac{4}{10}$

$P[E_A^c|Z_A] = \frac{1}{3} \quad \Rightarrow \quad P[E_A|Z_A] = \frac{2}{3}$

$P[Z_B^c] = \frac{6}{10} \quad \Rightarrow \quad P[Z_B] = \frac{4}{10}$

$P[E_B^c|Z_B] = \frac{2}{10}$

$P[E_B|Z_B^c] = \frac{p}{10} \quad \Rightarrow \quad P[E_B^c|Z_B^c] = 1 - \frac{p}{10}.$

a) Per Definition gilt $P[E_A^c|Z_A] := \frac{P[E_A^c \cap Z_A]}{P[Z_A]}$, die Wahrscheinlichkeit dass Anna zusagt und nicht erscheint errechnet sich daraus zu

$$P[Z_A \cap E_A^c] = P[E_A^c|Z_A]P[Z_A] = \frac{1}{3} \cdot \frac{6}{10} = \frac{1}{5}.$$

Analog ist die Wahrscheinlichkeit dass Anna zusagt und kommt

$$P[Z_A \cap E_A] = P[E_A|Z_A]P[Z_A] = \frac{2}{3} \cdot \frac{6}{10} = \frac{2}{5},$$

welche man auch durch

$$P[Z_A \cap E_A] = P[Z_A] - P[E_A^c \cap Z_A] = \frac{6}{10} - \frac{2}{10} = \frac{4}{10} = \frac{2}{5}$$

berechnen könnte.

Bemerkung betreffend Aufgabe c): Die Wahrscheinlichkeit $P[E_A] = P[E_A|Z_A]P[Z_A] + P[E_A|Z_A^c]P[Z_A^c]$ lässt sich aus den gegebenen Daten nicht berechnen, aber zumindest nach unten abschätzen, es ist $P[E_A] \geq P[E_A|Z_A]P[Z_A] = \frac{2}{3}\frac{6}{10} = \frac{4}{10}$.

Bitte wenden!

b) Mit dem Satz von der totalen Wahrscheinlichkeit erhalten wir

$$\begin{aligned}P[E_B^c] &= P[E_B^c|Z_B]P[Z_B] + P[E_B^c|Z_B^c]P[Z_B^c] = \frac{2}{10}\cdot\frac{4}{10} + \Big(1-\frac{p}{10}\Big)\cdot\frac{6}{10}\\ &\overset{p=2}{=} \frac{2}{10}\cdot\frac{4}{10} + \frac{8}{10}\cdot\frac{6}{10} = \frac{8}{100} + \frac{48}{100} = \frac{56}{100}.\end{aligned}$$

Die gesuchte Forderung an p ergibt sich aus der Ungleichung

$$P[E_B^c] = \frac{2}{10}\cdot\frac{4}{10} + (1-\frac{p}{10})\cdot\frac{6}{10} \leq \frac{1}{2},$$

woraus man $8 + 60 - 6p \leq 50$ erhält, zu $p \geq 3$.

c) Die Wahrscheinlichkeit, dass Bernadett damals zugesagt hat ergibt sich aus der Formel von Bayes zu

$$P[Z_B|E_B^c] = \frac{P[E_B^c|Z_B]P[Z_B]}{P[E_B^c]} = \frac{\frac{8}{100}}{\frac{56}{100}} = \frac{8}{56} = \frac{1}{7}.$$

d) Aus der Angabe folgt $P[E_A] = \frac{4}{10}$ und $P[E_B] = q$ und die Unabhängigkeit dieser Ereignisse. Die Wahrscheinlichkeit dass genau eines der beiden Mädchen erscheint ist gegeben durch

$$\begin{aligned}P[(E_A\cap E_B^c)\cup(E_A^c\cap E_B)] &= P[E_A\cap E_B^c] + P[E_A^c\cap E_B]\\ &= P[E_A]P[E_B^c] + P[E_A^c]P[E_b]\\ &= \frac{4}{10}(1-q) + \frac{6}{10}q = \frac{4}{10} + \frac{2}{10}q,\end{aligned}$$

wobei wir im zweiten Schritt die Unabhängigkeit von E_A und E_B (und damit jene von E_A und E_B^c usw.) benutzt haben, damit dieses Ereignis mehr als 50% Wahrscheinlichkeit hat, muss gelten

$$\frac{4}{10} + \frac{2}{10}q > \frac{5}{10},$$

woraus sich sofort $q > \frac{1}{2}$ ergibt.

3. a)

$$\begin{aligned}P(\min(T_1,T_2)\leq 5) &= 1 - P(\min(T_1,T_2) > 5) = 1 - P(T_1 > 5, T_2 > 5)\\ &= 1 - P(T_1>5)P(T_2>5) = 1-(1-F(5))^2\\ &= 1-(1-(1-e^{-\frac{1}{2}}))^2 = 1-e^{-1} = 0.63.\end{aligned}$$

Verallgemeinert auf beliebiges n und Zeit t ergibt das

$$\begin{aligned}P(\min(T_1,\ldots,T_n)\leq t) &= 1 - P(\min(T_1,\ldots,T_n) > t) = 1 - P(T_1>t,\ldots,T_n>t)\\ &= 1 - P(T_1>t)\cdots P(T_n>t) = 1-(1-F(t))^n\\ &= 1-(1-(1-e^{-\frac{1}{10}t}))^n = 1-e^{-\frac{nt}{10}}.\end{aligned}$$

Siehe nächstes Blatt!

b)

$$P(\max(T_1, T_2) \le 5) = P(T_1 \le 5, T_2 \le 5) = P(T_1 \le 5)P(T_2 \le 5)$$
$$= F(5)^2 = (1 - e^{-\frac{1}{2}})^2 = 0.15.$$

Verallgemeinert auf beliebiges n und Zeit t ergibt das

$$P(\max(T_1, \dots, T_n) \le t) = P(T_1 \le t, \dots, T_n \le t) = P(T_1 \le t) \cdots P(T_n \le t)$$
$$= F(t)^n = (1 - e^{-\frac{t}{10}})^n.$$

c) Wir haben:

$$f(z) = F'(z) = \frac{1}{10} \exp\left(-\exp\left(-\left(\frac{z}{10} - \mu\right)\right)\right) \exp\left(-\left(\frac{z}{10} - \mu\right)\right).$$

d) Die Log-likelihood ist:

$$\begin{aligned} L(\mu) &= \sum_{i=1}^{k} \log\left(\frac{1}{10} e^{-\left(\frac{m_i}{10} - \mu\right)} \exp\left(-e^{-\left(\frac{m_i}{10} - \mu\right)}\right)\right) \\ &= \sum_{i=1}^{k} \left(-\log(10) - \left(\frac{m_i}{10} - \mu\right) - e^{-\left(\frac{m_i}{10} - \mu\right)}\right) \\ &= -\log(10)k - \frac{1}{10}\sum_{i=1}^{k} m_i + k\mu - e^{\mu} \sum_{i=1}^{k} e^{-\frac{m_i}{10}}. \end{aligned}$$

Diese Funktion gilt es zu maximieren, also leiten wir zuerst ab,

$$\frac{\partial L}{\partial \mu} = k - e^{\mu} \sum_{i=1}^{k} e^{-\frac{m_i}{10}}.$$

Dann setzten wir $\frac{\partial L}{\partial \mu} = 0$ und lösen nach μ auf, d.h.

$$\hat{\mu} = -\log\left(\frac{1}{k}\sum_{i=1}^{k} e^{-\frac{m_i}{10}}\right).$$

4. **a)** Die Stichprobe ist gepaart.

Da keine Normalverteilung der Daten gegeben ist, werden wir den Wilcoxon-Test wählen, da dieser die vorhandenen Informationen besser ausnützt als der noch zur Wahl stehende Vorzeichentest.

b) Wir wollen herausfinden ob der Mittelwert μ der Differenzen von 0 abweicht, das heißt wir werden einen zweiseitigen Test zum Niveau $\alpha = 5\%$ mit den Hypothesen

$H_0 : \mu = 0$

$H_A : \mu \neq 0$

durchführen.

Bitte wenden!

c) Wenn man die Differenzen betragsmäßig ordnet, dann haben die positiven Differenzen 0.2, 0.6 und 3.2 die Ränge 1, 3 und 9, die Summe ergibt 13, gemäß der Tabelle für den 5%-Verwerfungsbereich beim Wilcoxon-Test (für eine Stichprobengröße von 12), liegt 13 (gerade noch) im Verwerfungsbereich, die Nullhypothese wird demnach also verworfen und der Test bestätigt, dass das Taxiunternehmen A im Schnitt schneller ist.

d) Für den Erfolgsparameter p lautet die Nullhypothese $p = p_0 = 0.5$ und die Alternativhypothese $p \neq p_0$. Empirisch ergibt sich aus 3 Erfolgen von 12 Versuchen $\hat{p} = \frac{3}{12} = 0.25$.

Aus dem Nomogramm erhalten wir ein 5%-Vertrauensintervall von ungefähr $[0.05, 0.57]$, und da $p_0 = 0.5$ in diesem Intervall liegt, wird die Nullhypothese nicht verworfen, mit diesem Test kann also nicht bestätigt werden, dass eines der beiden Unternehmen im Schnitt schneller ist.

Verwendet man hingegen die Normalapproximation zur Berechnung des Vertrauensintervalls für die Binomialverteilung, erhält man ein Vertrauensintervall von

$$\begin{aligned}\left[\frac{x}{n} \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sqrt{\frac{x}{n}\left(1 - \frac{x}{n}\right)\frac{1}{n}}\right] &= \left[0.25 \pm 1.96\sqrt{0.25\,(1 - 0.25)\,\frac{1}{12}}\right] \\ &= [0.25 \pm 0.245] = [0.005, 0.495]\,.\end{aligned}$$

Demnach wird die Nullhypothese verworfen, da 0.5 nicht in diesem Intervall liegt. Sie sollte aber nicht verworfen werden, da 3 (oder 9) Erfolge noch nicht signifikant genug sind um auf $p \neq 0.5$ zu schließen, d.h. wenn man 12 Mal eine Münze mit Erfolgsparameter $p = 0.5$ wirft sind 3 Erfolge nichts Ungewöhnliches. Tatsächlich ist $n = 12$ noch etwas zu klein um die Normalapproximation zu verwenden.