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

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

2. Bezeichne S, (bzw. G) das Ereignis, dass Maja es mit einem schlechten (bzw. guten) Bewerber zu tun hat, und A das Ereignis, dass der Bewerber von der Verwaltung angenommen wird.

a) Wir betrachten die Ereignisse S und A von oben. Dann ist

$$\mathbb{P}[S \cap A] = \mathbb{P}[A|S] \cdot \mathbb{P}[S] = 0.1 \cdot 0.7 = 0.07.$$

b) Da $\mathbb{P}[A] = \mathbb{P}[S \cap A] + \mathbb{P}[G \cap A] = 0.07 + 0.6 \cdot 0.3 = 0.25$ ist, erhalten wir mit der Bayesschen Formel

$$\mathbb{P}[S|A] = \frac{\mathbb{P}[S \cap A]}{\mathbb{P}[S \cap A] + \mathbb{P}[G \cap A]} = \frac{\mathbb{P}[S \cap A]}{\mathbb{P}[A]} = \frac{0.07}{0.25} = 0.28$$

für die gefragte Wahrscheinlichkeit.

c) Bei jedem Kandidaten, den Maja vorschlägt, beträgt die Wahrscheinlichkeit, dass die Verwaltung ihn akzeptiert $\mathbb{P}[A] = 0.25$, und $1 - \mathbb{P}[A] = 0.75$ ist die Wahrscheinlichkeit, dass ein Kandidat abgelehnt wird. Die Wahrscheinlichkeit, dass n Kandidaten abgelehnt werden ist also $(1 - \mathbb{P}[A])^n = 0.75^n$. Wir suchen daher das kleinste ganzzahlige $n \in \mathbb{N}$, sodass $1 - (1 - \mathbb{P}[A])^n \geq 0.95$, oder äquivalent $(1 - \mathbb{P}[A])^n \leq 0.05$:

$$\begin{aligned} & 0.75^n \leq 0.05 \\ \Leftrightarrow \quad & n \geq \frac{\log(0.05)}{\log(0.75)} = \log_{0.75}(0.05) \approx 10.41 \end{aligned}$$

liefert $n \geq 10.41,\ n \in \mathbb{N}$. Maja muss deshalb mindestens 11 Kandidaten vorschlagen.

d) Sei B das Ereignis, dass der jeweilige Bewerber bereit ist das Sofa zu kaufen. Dann fordern wir

$$\begin{aligned} P[G|B] &= \frac{P[G \cap B]}{P[B]} = \frac{0.3(1-p)}{0.3(1-p)+0.7(1-p/7)} \\ &= \frac{3-3p}{10-4p} \overset{!}{\geq} \frac{1}{8}. \end{aligned}$$

Auflösen nach p ergibt $3 - 3p \geq \frac{10}{8} - \frac{p}{2}$, oder $p \leq 0.7$. Also ist der grösstmögliche Wert von p den Maja verlangen kann $p = 0.7$.

3. a) X ist $\text{Bin}(60, p)$-verteilt.
Die Wahrscheinlichkeit $\mathbb{P}[X = x]$ für ein beliebiges $x \in \{0, \ldots, 60\}$ beträgt

$$\mathbb{P}[X = x] = \binom{60}{x} p^x (1-p)^{60-x}$$

und der Erwartungswert ist $\mathbb{E}[X] = 60p$.

b) Es gilt $\mu_1 = \mathbb{E}[X] = 60p$, und aus dem arithmetischen Mittel der Stichprobe erhält man

$$\frac{1}{5}\sum_{i=1}^{5} x_i = \frac{2+2+4+1+5}{5} = 60\hat{p},$$

und daraus $\hat{p} = \frac{14}{300} \approx 0.047$.

c) Jeder der sechs Kollegen kontrolliert 10 Fahrgäste. Die Wahrscheinlichkeit, dass keiner von diesen Schwarzfahrer ist, beträgt $(1-p)^{10}$. Deshalb ist die Anzahl Y der Kontrolleure, die bei einem Kontrollgang keinen Schwarzfahrer finden, $\text{Bin}\left(6, (1-p)^{10}\right)$-verteilt. Die Likelihood-Funktion in Abhängigkeit von y ist

$$L(p) = \mathbb{P}_p[Y = y] = \binom{6}{y} \left((1-p)^{10}\right)^y \left(1-(1-p)^{10}\right)^{6-y},$$

für die log-Likelihoodfunktion erhält man

$$l(p) = \log L(p) = \log\binom{6}{y} + y \log\left((1-p)^{10}\right) + (6-y)\log\left(1-(1-p)^{10}\right),$$

und $\frac{\partial l(p)}{\partial p} = 0$ liefert

$$\begin{aligned} & \frac{10y(1-\hat{p})^9}{(1-\hat{p})^{10}} = \frac{(6-y)10(1-\hat{p})^9}{1-(1-\hat{p})^{10}} \\ \Leftrightarrow \quad & y\left((1-(1-\hat{p})^{10}\right) = (6-y)(1-\hat{p})^{10} \\ \Leftrightarrow \quad & y = 6(1-\hat{p})^{10}. \end{aligned}$$

Siehe nächstes Blatt!

Daraus erhalten wir den Maximul-Likelihood Schätzer durch auflösen nach $\hat{p}$

$$\hat{p} = 1 - \sqrt[10]{\frac{y}{6}} \overset{y=4}{\approx} 0.04.$$

Alternativ, mit der Substitution $q := (1-p)^{10}$ und $Y \sim \mathrm{Bin}(6, q)$ ist die Likelihood-Funktion

$$L(q) = \mathbb{P}_q[Y = y] = \binom{6}{y} q^y (1-q)^{6-y},$$

für die log-Likelihoodfunktion erhält man

$$l(q) = \log L(q) = \log \binom{6}{y} + y \log(q) + (6-y)\log(1-q),$$

woraus man mit $\frac{\partial l(q)}{\partial q} = 0$ die Gleichung

$$\frac{y}{\hat{q}} = \frac{6-y}{1-\hat{q}},$$

und durch Auflösen nach $\hat{q}$ den Maximum-Likelihood Schätzer von q bekommt:

$$\hat{q} = \frac{y}{6} \overset{y=4}{=} \frac{2}{3}.$$

Daraus erhält man durch Zurücksubstituieren von $\hat{q} = (1-\hat{p})^{10}$

$$\hat{p} = 1 - \sqrt[10]{\frac{2}{3}} \approx 0.04.$$

d) K ist Geom(p)-verteilt. Gegeben sind Angaben der sechs Kontrolleure (d.h. $n = 6$), die als unabhängig angenommen werden können. Deshalb gilt

$$\begin{aligned} L(p) &= \mathbb{P}_p[K_1 = k_1, K_2 = k_2, K_3 = k_3, K_4 = k_4, K_5 = k_5, K_6 = k_6] \\ &= \prod_{i=1}^{6} (1-p)^{k_i - 1} p = (1-p)^{\sum_{i=1}^{6} k_i - 6} p^6, \end{aligned}$$

und für die log-Likelihoodfunktion erhalten wir

$$l(p) = \log L(p) = \left(\sum_{i=1}^{6} k_i - 6 \right) \log(1-p) + 6 \log p.$$

Aus $\frac{\partial l(p)}{\partial p} = 0$ und $\sum_{i=1}^{6} k_i = 118$ folgt

$$\frac{\sum_{i=1}^{6} k_i - 6}{1-\hat{p}} = \frac{6}{\hat{p}} \quad \Rightarrow \quad \frac{112}{1-\hat{p}} = \frac{6}{\hat{p}} \quad \Rightarrow \quad \hat{p} = \frac{6}{118} \approx 0.05.$$

Bitte wenden!

4. **a)** Bei den Angaben aus der Aufgabenstellung ist der z-Test am besten geeignet, denn M ist normalverteilt mit bekannter Streuung $\sigma = 120 \cdot 0.01 = 1.2$. Der Test ist zweiseitig.

b) Getestet wird auf den Mittelwert μ von $M \sim \mathcal{N}\left(\mu, \sigma^2\right)$. Die Null- und Alternativhypothese lauten

$$H_0 : \mu = \mu_0 := 100 \qquad\qquad H_A : \mu \neq \mu_0.$$

c) Die Teststatistik für Charles' Test lautet $\overline{M}_n$, wobei $n = 10$ ist.
Die Teststatistik für den z-Test aus der Vorlesung lautet

$$Z = \sqrt{n}\frac{\overline{M}_n - \mu_0}{\sigma},$$

i.e. die Teststatistik Z entsteht aus $\overline{M}_n$ durch Transformation auf die Standardnormalverteilung.
Der Verwerfungsbereich von Charles' Test ist $V\!B^{\overline{M}_n} = (-\infty, 99.5] \cup [100.5, \infty)$.
Der zum z-Test gehörige Verwerfungsbereich ergibt sich wie folgt aus dem Verwerfungsbereich von Charles' Test:

$$\overline{M}_n \notin (99.5, 100.5) \quad \Leftrightarrow \quad Z = \sqrt{n}\frac{\overline{M}_n - \mu_0}{\sigma} \notin \left(\sqrt{n}\frac{99.5 - \mu_0}{\sigma}, \sqrt{n}\frac{100.5 - \mu_0}{\sigma}\right),$$

also

$$\begin{aligned} V\!B^Z &= \left(-\infty, \sqrt{10}\frac{99.5 - 100}{1.2}\right] \cup \left[\sqrt{10}\frac{100.5 - 100}{1.2}, \infty\right) \\ &= (-\infty, -1.31] \cup [1.31, \infty). \end{aligned}$$

d) Unter der Nullhypothese ist $\overline{M}_n \sim \mathcal{N}\left(\mu_0, \frac{\sigma^2}{n}\right)$, also

$$\overline{M}_{10} = \mathcal{N}\left(100, \frac{(120 \cdot 0.01)^2}{10}\right) = \mathcal{N}\left(100, 0.144\right).$$

Das Niveau α bzw. der Fehler erster Art ist bei beiden Tests gleich, und es beträgt

$$\begin{aligned} \mathbb{P}_{\mu=100}\left(\overline{M}_{10} \in V\!B^{\overline{M}_{10}}\right) &= \mathbb{P}_{\mu=100}\left(\overline{M}_{10} \in (-\infty, 99.5] \cup [100.5, \infty)\right) \\ &= 2\,\mathbb{P}_{\mu=100}\left(\overline{M}_{10} \leq 99.5\right) = 2\,F_{\mathcal{N}(100,0.144)}(99.5) \\ &= 2\,\Phi\left(\frac{99.5 - 100}{\sqrt{0.144}}\right) = 2\,\Phi(-1.32) == 2 \cdot 0.0934 = 0.1868, \end{aligned}$$

wobei die zweite Gleichheit aus Symmetriegründen folgt. Dies ist folglich auch der Wert für $\mathbb{P}_{\mu=100}\left(Z \in V\!B^Z\right)$ (Transformation auf die Standardnormalverteilung).