Musterlösung der Klausur Stochastik für Informatik*er vom 26.10.2020

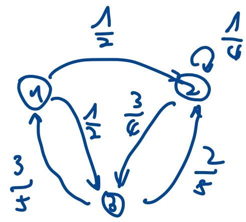
Exercise 1.1 (10 Punkte)
Sei $(X_n)_{n\in\mathbb{N}}$ eine Markovkette auf dem Zustandsraum $S = \{1, 2, 3\}$ mit Übergangsmatrix

$$P = \begin{pmatrix} a & 1/2 & 1/2 \\ 1/4 & 0 & b \\ 3/5 & c & 0 \end{pmatrix}$$

$a = 0$
$c = \frac{2}{5}$
$b = \frac{3}{4}$

(a) Bestimmen Sie a, b und c.

(b) Zeichnen Sie den Übergangsgraphen der Markov-Kette.



(c) Ist die Kette irreduzibel? ja

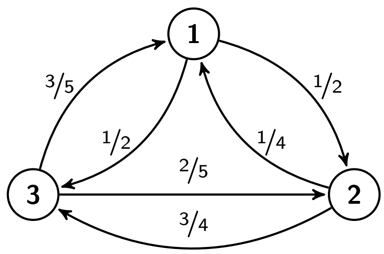
(d) Bestimmen Sie $\mathbb{P}(X_j = 1|X_0 = 1)$ für $j = 1, 2, 3$.

(e) Bestimmen Sie alle möglichen invarianten Verteilungen.

Lösungsskizze zu Exercise 1.1.

(a) Die Zeilensummen müssen 1 ergeben. Darum gilt: $a = 0, b = 3/4, c = 2/5$.

(b)



$(1,0,0)P = (1,0,0)\begin{pmatrix} 0 & \frac12 & \frac12 \\ \frac14 & 0 & \frac34 \\ \frac35 & \frac25 & 0 \end{pmatrix} = \left(0, \frac12, \frac12\right)$

$P^2 = \begin{pmatrix} 0 & \frac12 & \frac12 \\ \frac14 & 0 & \frac34 \\ \frac35 & \frac25 & 0 \end{pmatrix}\begin{pmatrix} 0 & \frac12 & \frac12 \\ \frac14 & 0 & \frac34 \\ \frac35 & \frac25 & 0 \end{pmatrix} = \begin{pmatrix} & & \\ & & \end{pmatrix}$

$(1,0,0)P^2 = \left(\frac{17}{40}, \frac15, \frac38\right)$

$40 + 28 + 5$

(c) Die Markov-Kette ist irreduzibel: Jeder Zustand kann von jedem Zustand aus erreicht werden.

$P^T - I = \begin{pmatrix} -1 & \frac14 & \frac35 \\ \frac12 & -1 & \frac25 \\ \frac12 & \frac34 & -1 \end{pmatrix} \Rightarrow \begin{array}{l} \frac14 v_2 + \frac35 v_3 = v_1 \\ v_2 = \frac12 v_1 + \frac25 v_3 \\ v_3 = \frac12 v_1 + \frac34 v_2 \end{array} \Rightarrow \left\{ \begin{array}{l} v_1 = \alpha \\ v_2 = \frac{4}{5}\alpha \\ v_3 = \frac12 \alpha \end{array} \right.$

$v_1 + v_2 + v_3 = \frac{43}{10}\alpha = 1$

(d) $\mathbb{P}(X_1 = 1|X_0 = 1) = 0$
Mit der Formel von der Gesamtwahrscheinlichkeit:
$\mathbb{P}(X_2 = 1|X_0 = 1) = \mathbb{P}(X_2 = 1, X_1 = 2|X_0 = 1) + \mathbb{P}(X_2 = 1, X_1 = 3|X_0 = 1) = 1/2 \cdot 1/4 + 1/2 \cdot 3/5 = 34/80 = 17/40 = 0.425$.

$\alpha = 2v_2 - \frac45 v_3 = 2v_3 + \frac32 v_2$

$\frac12 v_2 = \frac{14}{5} v_3$

Mit der Formel von der Gesamtwahrscheinlichkeit:
$\mathbb{P}(X_3 = 1|X_0 = 1) = \mathbb{P}(X_3 = 1, X_2 = 3, X_1 = 2|X_0 = 1) + \mathbb{P}(X_3 = 1, X_2 = 2, X_1 = 3|X_0 = 1) = 1/2 \cdot 3/4 \cdot 3/5 + 1/2 \cdot 2/5 \cdot 1/4 = 9/40 + 2/40 = 11/40 = 0.275$.

$\frac75 v_3 + \frac35 v_3 = 2v_3 = \alpha$

(e) Die invarianten Verteilungen sind die Lösungen von $\pi^T P = \pi^T$ oder $(P - I)^T \pi = 0$ (beide sind korrekt). Berechnung:

$$\begin{aligned} (I) &\quad -\pi_1 + 1/4\pi_2 + 3/5\pi_3 = 0 \\ (II) &\quad 1/2\pi_1 - \pi_2 + 2/5\pi_3 = 0 \\ (III) &\quad 1/2\pi_1 + 3/4\pi_2 - \pi_3 = 0 \\ (IV) &\quad \pi_1 + \pi_2 + \pi_3 = 1 \end{aligned}$$

$-\pi_1 + \frac45 \pi_3 = 0$

$\frac74 \pi_2 - \frac75 \pi_3 = 0$

$\pi_2 = \frac45 \pi_3$

$\left(\frac45 + \frac45 + 1\right)\pi_3 = 1$

Durch, z.B., $(II - III) : -7/4\pi_2 + 7/5\pi_3 = 0 \;\Rightarrow\; \pi_2 = 4/5\pi_3$ und $(3 \cdot I - III) : -7/2\pi_1 + 14/5\pi_3 = 0 \;\Rightarrow\; \pi_1 = 4/5\pi_3$. Zusammen mit (IV) ist also:

$\pi_3 = \frac{5}{13}$

$\pi_1 = \pi_2 = \frac{4}{13}$

$$1 = \frac45\pi_3 + \frac45\pi_3 + \pi_3 = \frac{13}{5}\pi_3 \;\Rightarrow\; \pi_3 = \frac{5}{13} \;\Rightarrow\; \pi_1 = \pi_2 = \frac45\,\frac{5}{13} = \frac{4}{13}.$$

(Das ist die eindeutige Lösung des Gleichungssystems, also gibt es nur eine invariante Verteilung, aber das muss man nicht erwähnen.)

Exercise 1.2 (10 Punkte)
In einer Großstadt gibt es insgesamt $n \in \mathbb{N}$ studentische Mensen, die im Sommersemester 2020 Corona-bedingt alle geschlossen blieben. Im Wintersemester wird jede Mensa unabhängig voneinander mit Wahrscheinlichkeit $\frac{1}{2n+10}$ wiedereröffnet.
Sei X_n die Anzahl an Mensen, die wiedereröffnet werden (in Abhängigkeit von n).

$\mathrm{Bin} \sim (n, \frac{1}{2n+10})$

(a) Welchen Namen hat die Verteilung von X_n und welche Parameter hat sie?

(b) Bestimmen Sie $\mathbb{E}[X_n]$ und $\mathbb{V}[X_n]$ (in Abhängigkeit von n). $\mathbb{E}[X_n] = \frac{n}{2n+10}$ $\mathbb{V}[X_n] = \frac{n \cdot (2n+9)}{(2n+10)^2}$

(c) Sei nun $n = 20$. Berechnen Sie genau die Wahrscheinlichkeit, dass mindestens eine Mensa wiedereröffnet wird. $\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X=0) = 1 - \binom{20}{0}\left(\frac{2n+9}{2n+10}\right)^{20} = 1 - \left(\frac{49}{50}\right)^{20}$

(d) Sei nun $n = 500$. Berechnen Sie die Wahrscheinlichkeit, dass genau 2 Mensen wiedereröffnet werden, mit einer (für großes n) geeigneten Approximation. $\lambda = 500 \cdot \frac{1}{1010} = \frac{100}{202} = \frac{50}{101} \approx \frac{1}{2}$

(e) Sei X die Zufallsvariable, die in Aufgabe d) die Anzahl der Mensen approximiert, die wiedereröffnet werden. Zeigen Sie, dass

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X].$$

$\lambda = n \cdot \frac{1}{2n+1}$

$\lim_{n\to\infty} \mathbb{E}[X_n] = \lim_{n\to\infty} \frac{n}{2n+1} = \frac{1}{2}$

Lösungsskizze zu Exercise 1.2.

(a) $X_n \sim \mathrm{Bin}(n, \frac{1}{2n+10})$.

(b) $\mathbb{E}[X_n] = \frac{n}{2n+10}$, $\mathbb{V}[X_n] = n \cdot \frac{1}{2n+10} \cdot \left(1 - \frac{1}{2n+10}\right) = \frac{n(2n+9)}{(2n+10)^2} = \frac{2n^2+9n}{4n^2+20n+100}$. (Der erste Schritt o.ä. reicht.)

(c) $\mathbb{P}(X_n \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \left(\frac{2n+9}{2n+10}\right)^n = 1 - \left(\frac{49}{50}\right)^{20} \approx 0.3324$.

(d) Wir nehmen die Poissonapproximation mit Parameter $\lambda = \lim_{n\to\infty} n \cdot \frac{1}{2n+10} = \frac{1}{2}$. Damit erhalten wir [mit X aus Aufgabe e)]

$$\mathbb{P}(\text{genau 2 Mensen werden wiedereröffnet}) \approx \mathbb{P}(X = 2) = \frac{0.5^2}{2!} \mathrm{e}^{-1/2} \approx 0.0758.$$

(Falls man eine Normalapproximation verwendet, kann man maximal 1 Punkt bekommen.)

(e) $\mathbb{E}[X] = \lambda = 0.5$, da X poissonverteilt ist, und laut Aufgabe b) gilt $\lim_{n\to\infty} \mathbb{E}[X_n] = \lim_{n\to\infty} \frac{n}{2n+10} = 0.5$.

Exercise 1.3 (10 Punkte)
In einem Tierheim wohnen 2 schwarze und 2 getigerte Katzen. 2 von diesen 4 Katzen werden zufällig und gleichzeitig ausgewählt und zum Tierarzt gebracht, um eine Impfung gegen Tollwut zu bekommen. Sei X die Anzahl an ausgewählten schwarzen Katzen und Y die Anzahl an ausgewählten getigerten Katzen.
In den folgenden Aufgaben a) und b) erwarten wir ausnahmsweise keine Begründung.

(a) Geben Sie die gemeinsame Verteilung von X und Y mit Randverteilung in der folgende Tabelle an.

$\binom{2}{1}\binom{2}{1}$ $\quad \frac{\binom{2}{2}}{\binom{4}{2}} = \frac{1}{\frac{4!}{2!\,2!}}$

$X \backslash Y$	0	1	2	p_X
0	0	0	$\frac{1}{6}$	$\frac{1}{6}$
1	0	$\frac{2}{3}$	0	$\frac{2}{3}$
2	$\frac{1}{6}$	0	0	$\frac{1}{6}$
p_Y	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$	1

$(X-Y)(\Omega) = \{-2, 2, 0\}$

(b) Bestimmen Sie den minimalen Wertebereich der Zufallsvariablen $X - Y$.
Hinweis: der minimale Wertebereich enthält genau die Werte, die $X - Y$ mit positiver Wahrscheinlichkeit annimmt.

(c) Sind X und Y unabhängig? Sind sie identisch verteilt? Warum? Nein   ja, da $\mathbb{P}(X=k) = \mathbb{P}(Y=k)$ $k = \{0,1,2\}$

(d) Berechnen Sie $\mathbb{E}[X]$ und $\mathbb{E}[Y]$. $\mathbb{E}[X] = \mathbb{E}[Y] = 1 \cdot \frac{2}{3} + \frac{1}{3} = 1$

(e) Berechnen Sie die Kovarianz von X und Y. $\mathbb{E}[XY] = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$ $\mathrm{Cov}(X,Y) = \frac{2}{3} - 1 = -\frac{1}{3}$

Ersatzlösung für maximal 7 Punkte: falls Sie die gemeinsame Verteilung nicht ausrechnen konnten, nehmen Sie an, dass $\mathbb{P}(X = k, Y = l) = \frac{1}{9}$ für alle $k, l \in \{0, 1, 2\}$ gilt, und beantworten Sie alle weitere Fragen im Hinblick auf diese gemeinsame Verteilung.

Lösungsskizze zu Exercise 1.3.

(a)

$X \backslash Y$	0	1	2	p_X
0	0	0	$\frac{1}{6}$	$\frac{1}{6}$
1	0	$\frac{2}{3}$	0	$\frac{2}{3}$
2	$\frac{1}{6}$	0	0	$\frac{1}{6}$
p_Y	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$	

(Bei falscher gemeinsamen Verteilung geben wir Folgefehler.)

(b) $(X - Y)(\Omega) = \{-2, 0, 2\}$.

(c) X und Y sind nicht unabhängig, da z.B. $0 = \mathbb{P}(X = Y = 0) \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 0)$ (oder andere Werte in der Tabelle).
Andere Begründungen sind auch möglich. Z.B. könnte man laut Aufgabe e) sagen, dass X und Y sogar nicht unkorreliert sind. Oder man könnte auch mit dem Wertebereich bzw. Verteilung von $X + Y$ oder $X - Y$ argumentieren.
X und Y sind identisch verteilt, da laut der Tabelle $\mathbb{P}(X = k) = \mathbb{P}(Y = k)$ für alle $k \in \{0, 1, 2\}$ gilt. (Man kann auch einfach die Symmetrie der Rollen der schwarzen und getigerten Katzen erwähnen.)

(d) **1. Lösung**: $\mathbb{E}X = \frac{1}{6} \cdot 0 + \frac{2}{3} \cdot 1 + \frac{1}{6} \cdot 2 = 1$. $\mathbb{E}Y = 1$, wegen einer ähnlichen Rechnung, oder weil X, Y identisch verteilt sind.
2. Lösung: Wir beobachten, dass $\mathbb{E}[X + Y] = 2$, da $\mathbb{P}(X + Y = 2) = 1$. Wegen Linearität des Erwartungswerts und Symmetrie der Rollen von X und Y muss $\mathbb{E}X = \mathbb{E}Y = 1$ gelten.

(e) $\mathrm{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y = \mathbb{E}(XY) - 1$.

Es gelten $X = 2, Y = 0 \Rightarrow XY = 0$, $X = Y = 1 \Rightarrow XY = 1$ und $X = 0, Y = 2 \Rightarrow XY = 0$.

$$\mathbb{E}(XY) = \sum_{k \in XY(\Omega)} k\mathbb{P}(XY = k) = \frac{1}{6} \cdot 0 + \frac{2}{3} \cdot 1 + \frac{1}{6} \cdot 0 = \frac{2}{3},$$

(Falls die Verteilung von XY von der Rechnung nachvollziehbar ist, muss man sie nicht explizit erwähnen, z.B. das Display alleine reicht schon), es folgt $\operatorname{cov}(X, Y) = -\frac{1}{3}$.
Man kann $\operatorname{cov}(X, Y)$ natürlich auch als $\mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]$ berechnen.

Ersatzlösung: alle Einträge in den Randverteilungen sind $\frac{1}{3}$; $(X - Y)(\Omega) = \{-2, 1, 0, 1, 2\}$; X und Y sind unabhängig, da $\mathbb{P}(X = k, Y = l) = \mathbb{P}(X = k)\mathbb{P}(Y = l)$ für alle $k, l \in \{0, 1, 2\}$; X und Y sind identisch verteilt, da $\mathbb{P}(X = k) = \mathbb{P}(Y = k)$ für alle $k \in \{0, 1, 2\}$ oder wegen Symmetrie der Ersatztabelle; $\mathbb{E}X = \frac{1}{3} \cdot (0 + 1 + 2) = 1$; $\mathbb{E}Y = 1$ wegen ähnlicher Rechnung oder da X, Y identisch verteilt sind; $\operatorname{cov}(X, Y) = 0$, da X, Y unabhängig sind.

Exercise 1.4 (10 Punkte)

Sei X eine Zufallsvariable mit stetiger Verteilungsfunktion

$$F_X(x) = \begin{cases} 0, & \text{falls } x \leq -3/2 \\ a(x + 3/2), & \text{falls } -3/2 < x \leq 0 \\ 3/4 + bx - x^2, & \text{falls } 0 < x \leq b/2 \\ 1, & \text{falls } b/2 < x. \end{cases}$$

$\lim_{x \to \frac{b}{2}} F_X(x) = \frac{3}{4} + \frac{b^2}{2} - \frac{b^2}{4}$

$= \frac{3}{4} + \frac{b^2}{4} = 1$

$\Rightarrow b^2 = 1 \quad b = 1$

$\frac{3}{2}a = \frac{3}{4} \Rightarrow a = \frac{1}{2}$

(a) Bestimmen Sie die Werte für $a, b > 0$.

(b) Bestimmen Sie den Wert für c, so dass $\mathbb{P}(X \leq c) = 15/16$.

$F_X(c) = \mathbb{P}(X \leq c) = \frac{3}{4} + c - c^2 = \frac{15}{16}$

(c) Bestimmen und skizzieren Sie die Dichte $f(x)$ von X.

$c^2 - c + \frac{3}{16} = 0$

(d) Berechnen Sie $\mathbb{E}[X]$.

$(c - \frac{1}{2})^2 = -\frac{3}{16} + \frac{1}{4}$

$= \frac{1}{16}$

Lösungsskizze zu Exercise 1.4.

(a) Da F_X stetig ist, muss für $x = 0$ gelten: $3/2a = 3/4 \;\Rightarrow\; a = 1/2$.
Ebenso gilt in $x = b/2$: $3/4 + b^2/2 - b^2/4 = 1 \;\Rightarrow\; b^2/4 = 1/4 \;\Rightarrow\; b = 1$.

$c_{1,2} = \pm \frac{1}{4} + \frac{1}{2}$

$c_1 = \frac{3}{4} \quad c_2 = \frac{1}{4}$

(b) Es gilt: $15/16 = \mathbb{P}(X \leq c) = F_X(c) = 3/4 + c - c^2$. Daraus folgt: $c^2 - c + 3/16 = 0$. Mit pq-Formel: $c_{1,2} = 1/2 \pm \sqrt{1/4 - 3/16} = 1/2 \pm 1/4$.
Da $0 < c \leq 1/2$, ist also $c = 1/4$.

$c \in]0, \frac{b}{2} = \frac{1}{2}]$

$c_2 = \frac{1}{4}$

(c) Wegen $f(x) = F'_X(x)$ gilt:

$$f(x) = \begin{cases} 0, & \text{falls } x \leq -3/2 \\ 1/2, & \text{falls } -3/2 < x \leq 0 \\ 1 - 2x, & \text{falls } 0 < x \leq 1/2 \\ 0, & \text{falls } 1/2 < x. \end{cases}$$

$x - 2x^2$

$\mathbb{E}[X] = \int_{-\frac{3}{2}}^{0} x\frac{1}{2}dx + \int_{0}^{\frac{1}{2}} (1-2x)x\,dx$

$= \frac{1}{4}x^2 \Big|_{-\frac{3}{2}}^{0} + \left(\frac{1}{2}x^2 - \frac{2}{3}x^3\right)\Big|_0^{\frac{1}{2}}$

$= -\frac{1}{4} \cdot \frac{9}{4} + \frac{1}{8} - \frac{2}{3} \cdot \frac{1}{8}$

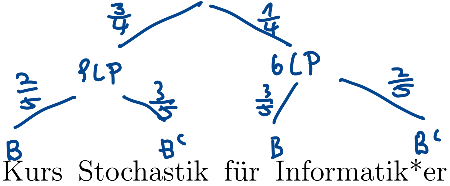
$= -\frac{8}{16} - \frac{1}{8} = -\frac{11}{16}$

(d)

$$\mathbb{E}X = \int_{\mathbb{R}} xf(x)dx = \int_{-3/2}^{0} \frac{x}{2}dx + \int_0^{1/2} (x - 2x^2)dx$$

$$= \frac{x^2}{4}\Big|_{-3/2}^{0} + \frac{x^2}{2} - \frac{2x^3}{3}\Big|_0^{1/2} = \frac{9}{4} + \frac{1}{8} - \frac{1}{12} = \frac{-54 + 12 - 8}{96} = -\frac{25}{48} \approx -0.5208.$$

$\frac{1}{4} \cdot \frac{8}{4}$



Exercise 1.5 (10 Punkte, nur 6LP)
Eine zufällig ausgewählte Studierende beim Kurs Stochastik für Informatik*er schreibt mit Wahrscheinlichkeit $\frac{3}{4}$ die 9LP-Klausur und mit Wahrscheinlichkeit $\frac{1}{4}$ die 6LP-Klausur. Die 6LP-Klausur enthält mit Wahrscheinlichkeit $\frac{3}{5}$ eine Aufgabe über bedingte Wahrscheinlichkeiten, wobei die 9LP-Klausur mit Wahrscheinlichkeit $\frac{2}{5}$ eine Aufgabe über bedingte Wahrscheinlichkeiten enthält.

(a) Stellen Sie die Situation als Baum dar. Schreiben Sie die im Aufgabentext gegebenen Wahrscheinlichkeiten an die richtigen Stellen im Baum.

(b) Mit welcher Wahrscheinlichkeit hat die (zufällig ausgewählte) Studierende eine Aufgabe über bedingte Wahrscheinlichkeiten in ihrer Klausur? $\mathbb{P}(B) = \mathbb{P}(B|9LP)\mathbb{P}(9LP) + \mathbb{P}(B|6LP)\mathbb{P}(6LP)$
$= \frac{3}{10} + \frac{3}{20} = \frac{9}{20}$

(c) Angenommen, die Studierende hatte eine Aufgabe über bedingte Wahrscheinlichkeiten in ihrer Klausur. Was ist wahrscheinlicher: sie hat die 9LP-Klausur geschrieben oder sie hat die 6LP-Klausur geschrieben? $\mathbb{P}(9LP|B) = \frac{\mathbb{P}(9LP \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|9LP)\mathbb{P}(9LP)}{\mathbb{P}(B)} = \frac{\frac{2}{5}\cdot\frac{3}{4}}{\frac{9}{20}} = \frac{2}{3}$

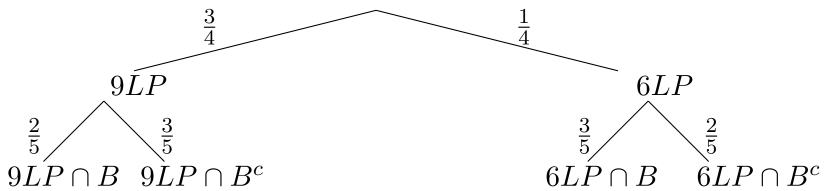
(d) Für beide Klausurvarianten (sowohl 9LP als auch 6LP) gilt: falls die Klausur eine Aufgabe über bedingte Wahrscheinlichkeiten enthält, muss man in dieser Aufgabe mit Wahrscheinlichkeit $\frac{2}{3}$ einen Baum skizzieren. Wie groß ist die Wahrscheinlichkeit, dass die Studierende **keinen** Baum in ihrer Klausur skizzieren muss?
(Besonders hier: begründen Sie Ihre Aussage.)

$\mathbb{P}(S) = \mathbb{P}(B)\cdot\frac{2}{3} = \frac{9}{20}\cdot\frac{2}{3} = \frac{3}{10}$

$\mathbb{P}(S^c) = \frac{7}{10}$

Lösungsskizze zu Exercise 1.5.

(a) Notation: $9LP, 6LP$: jeweilige Klausurvariante wurde geschrieben, B: Aufgabe über bedingte Wahrscheinlichkeit in der Klausur.



(b) Formel von der Gesamtwahrscheinlichkeit (kann auch implizit verwendet werden) liefert:

$$\mathbb{P}(B) = \mathbb{P}(B|9LP)\mathbb{P}(9LP) + \mathbb{P}(B|6LP)\mathbb{P}(6LP) = \frac{2}{5}\cdot\frac{3}{4} + \frac{3}{5}\cdot\frac{1}{4} = \frac{9}{20}\Big(= 0.45\Big).$$

(c) Bayes-Formel (kann auch implizit verwendet werden) liefert:

$$\mathbb{P}(9LP|B) = \frac{\mathbb{P}(B|9LP)\mathbb{P}(9LP)}{\mathbb{P}(B)} = \frac{\frac{2}{5}\cdot\frac{3}{4}}{\frac{9}{20}} = \frac{6}{9} = \frac{2}{3}.$$

(Hier wurde $\mathbb{P}(B)$ aus der Lösung von b) verwendet, aber man kann auch die Rechnung von b) wiederholen.)
Da $\mathbb{P}(9LP|B) + \mathbb{P}(6LP|B) = 1$ und $\mathbb{P}(9LP|B) > \frac{1}{2}$, ist es wahrscheinlicher, dass die Studierende die 9LP-Klausur geschrieben hat, falls sie eine Aufgabe über bedingte Wahrscheinlichkeiten in ihrer Klausur hatte. (Alternativ kann man $\mathbb{P}(6LP|B)$ wieder mit der Bayes-Formel berechnen und feststellen, dass das gleich $\frac{1}{3} < \mathbb{P}(9LP|B)$ ist.)

(d) Sei B' das Ereignis, dass man in der Klausur einen Baum skizzieren muss, gesucht wird.
1. Lösung: Es gilt

$$\begin{aligned}\mathbb{P}(B') &= \mathbb{P}(B' \cap 9LP) + \mathbb{P}(B' \cap 6LP) = \mathbb{P}(B' \cap B \cap 9LP) + \mathbb{P}(B' \cap B \cap 6LP) \\ &= \mathbb{P}(B'|B \cap 9LP)\mathbb{P}(B|9LP)\mathbb{P}(9LP) + \mathbb{P}(B'|B \cap 6LP)\mathbb{P}(B|6LP)\mathbb{P}(6LP) \\ &= \frac{2}{3} \cdot \frac{2}{5} \cdot \frac{3}{4} + \frac{2}{3} \cdot \frac{3}{5} \cdot \frac{1}{4} = \frac{12}{60} + \frac{6}{60} = \frac{18}{60} = \frac{3}{10}.\end{aligned}$$

(Die zweite Zeile des Displays muss nicht explizit da stehen, wenn sie von der Rechnung nachvollziehbar ist.)
Deshalb ist die gesuchte Wahrscheinlichkeit $\mathbb{P}(B'^c) = 1 - \mathbb{P}(B') = \frac{7}{10}(= 0.7)$.
2. Lösung: laut Aufgabenstellung gilt $\mathbb{P}(B'|B) = \frac{2}{3}$ (die Unabhängigkeit von B' und $9LP$ bzw. $6LP$ muss nicht explizit erwähnt werden). Deswegen gilt $\mathbb{P}(B') = \mathbb{P}(B'|B)\mathbb{P}(B) = \frac{9}{20} \cdot \frac{2}{3} = \frac{3}{10}$ und somit ist $\mathbb{P}(B'^c) = 1 - \mathbb{P}(B') = \frac{7}{10}(= 0.7)$ die gesuchte Wahrscheinlichkeit.

Exercise 1.6 (10 Punkte, nur 9LP)
In verschiedenen Berliner Sportvereinen wird das Raucherverhalten (Raucher vs. Nichtraucher) der Sportlerinnen und Sportler in Mannschafts- und Individualsportarten untersucht. Dabei wurden die unten angegebenen Antworten gegeben. Es soll die Frage beantwortet werden, ob das Raucherverhalten einer sporttreibenden Person unabhängig davon ist, ob sie die Sportart in einer Mannschaft oder im Einzel betreibt. Das Signifikanzniveau beträgt $\alpha = 5\%$.

	Raucher	Nichtraucher
Fußball	23	15
Handball	17	12
Volleyball	12	17
Triathlon	1	3
Crossfit	2	5
Tennis	4	7

(a) Mittels welchem Test können Sie die Forschungsfrage beantworten?
Wie lautet die Nullhypothese?

(b) Tragen Sie die empirischen Häufigkeiten in folgende Tabelle ein:

	Raucher	Nichtraucher	Total
Mannschaft			
Individual			
Total			

(c) Was ist der Freiheitsgrad f?

(d) Was ist der Vergleichswert? Nutzen Sie dafür folgende Tabelle:

f \ α	0.025	0.05	0.1	0.9	0.95	0.975
1	0.001	0.004	0.016	2.706	3.841	5.024
2	0.051	0.103	0.211	4.605	5.991	7.378
3	0.216	0.352	0.584	6.251	7.815	9.348
4	0.484	0.711	1.064	7.779	9.488	11.143

(e) Bestimmen Sie die theoretischen Häufigkeiten und tragen Sie sie in folgende Tabelle ein:

	Raucher	Nichtraucher	Total
Mannschaft			
Individual			
Total			

(f) Bestimmten Sie den Testwert.

(g) Wird die Nullhypothese angenommen? Begründen Sie Ihre Antwort.

Lösungsskizze zu Exercise 1.6.

(a) χ^2-Test auf Unabhängigkeit. Die Nullhypothese lautet: Ob eine sporttreibende Person raucht oder nicht, ist unabhängig davon, ob sie die Sportart in einer Mannschaft oder im Einzel betreibt.

(b)

	Raucher	Nichtraucher	Total
Mannschaft	52	44	96
Individual	7	15	22
Total	59	59	118

(c) Freiheitsgrad $f = (m-1)(k-1) = (2-1)(2-1) = 1$.

(d) $\chi^2_{1-\alpha,f} = \chi^2_{0.95,1} = 3.841$.

(e) Theoretischen Häufigkeiten werden berechnet mittels $F_{i,j} = \frac{N_{i,*} \cdot N_{*,j}}{N_{\text{total}}}$

	Raucher	Nichtraucher	Total
Mannschaft	$96 \cdot 56/118 = 48$	$96 \cdot 56/118 = 48$	96
Individual	$22 \cdot 56/118 = 11$	$22 \cdot 56/118 = 11$	22
Total	59	59	118

(f) Testwert $\chi^2 = \sum_{i,j} \frac{(F_{i,j}-N_{i,j})^2}{F_{i,j}} = \frac{4^2}{48} + \frac{4^2}{48} + \frac{4^2}{11} + \frac{4^2}{11} = \frac{32}{48} + \frac{32}{11} \approx 3.58 < 3\frac{2}{3}$.

(g) Da der Testwert kleiner als der Vergleichswert ist: $3.58 = \chi^2 < \chi^2_{0.95,1} = 3.84$, wird die Nullhypothese nicht verworfen.