Lösung zur 2. Aufgabe (AGS 10.1)

(a)

1. $X = \{1, \dots, 6\} \times \{1, \dots, 6\} = \{(i, j) \mid 1 \leq i, j \leq 6\}$, $p(i, j) = 1/36$
2. Dieses Zufallsexperiment ist das unabhängige Produkt der folgenden beiden Zufallsexperimente:

$$\begin{array}{llll} X_1 = \{W, S\} \ , & p_1(W) = 2/5 \ , & p_1(S) = 3/5 \ , & \\ X_2 = \{R, G, B\} \ , & p_2(R) = 1/6 \ , & p_2(B) = 1/3 \ , & p_2(G) = 1/2 \ . \end{array}$$

   Deshalb ist $X = \{W, S\} \times \{R, B, G\}$ und $p = p_1 \times p_2$, also:

$$\begin{array}{lll} p(W, R) = 1/15 \ , & p(W, G) = 2/15 & , p(W, B) = 1/5 \ , \\ p(S, R) = 1/10 \ , & p(S, G) = 1/5 \ , & p(S, B) = 3/10 \ , \end{array}$$

3. $X = \{K, Z\}^6$. Es gilt beispielsweise:

$$p(K, K, K, Z, K, K) = 0.3^5 \cdot 0.7 \ , \qquad p(K, Z, K, K, Z, Z) = 0.3^3 \cdot 0.7^3 \ .$$

   Allgemein: für jedes $w \in X$ ist $p(w) = 0.3^{K(w)} \cdot 0.7^{Z(w)}$, wobei $K(w)$ die Anzahl von K in w und $Z(w)$ die Anzahl an Z in w ist.

(b)

1. $X = \{1, 2, 3, 4\}$, wenn die Seiten von 1 bis 4 durchnummeriert sind. Das Modell ist unbeschränkt:

$$\mathcal{M} = \mathcal{M}(X)$$

2. $X = \{1, \dots, 6\}$.

$$\mathcal{M} = \{p \in \mathcal{M}(X) \mid p(2) = p(4) = p(6)\} \ .$$

   Das Modell ist beschränkt, da es z.B. die Verteilung p mit $p(i) = i/21$ nicht enthält.

3. $X = \{(K, K), (K, Z), (Z, K), (Z, Z)\}$.

$$\mathcal{M} = \big\{p \times p \mid p \in \mathcal{M}(\{K, Z\})\big\} \ .$$

   Das Modell ist beschränkt: jede Verteilung mit $p(K, Z) \neq p(Z, K)$ ist nicht enthalten.

(c)

1. Wir können auf zwei Arten vorgehen: (a) die Zufallsexperimente für beide Münzen aufstellen, jeweils die Wahrscheinlichkeit für das Ergebnis (K, Z, K, Z, K) bestimmen und die Münze mit der höheren Wahrscheinlichkeit wählen oder (b) die Maximum-Likelihood-Schätzung durchführen.

a) Beide Zufallsexperimente haben die Ergebnismenge $\{K, Z\}^5$. Bei der ersten und zweiten Münze ergibt sich:

$$p_1(K, Z, K, Z, K) = 0.7^3 \cdot 0.3^2 \approx 0.0309 \; ,$$
$$p_2(K, Z, K, Z, K) = 0.5^3 \cdot 0.5^2 \approx 0.0313 \; .$$

Die zweite Münze bietet also einen Vorteil.

b) Das Wahrscheinlichkeitsmodell ist $\mathcal{M} = \{p_1, p_2\}$ und der Korpus ist $h(K) = 3$, $h(Z) = 2$. Dann ist

$$L(h, p_1) = 0.7^3 \cdot 0.3^2 \approx 0.0309 \; ,$$
$$L(h, p_2) = 0.5^3 \cdot 0.5^2 \approx 0.0313 \; .$$

Dann ist

$$\mathrm{mle}(h, \mathcal{M}) =_{p \in \mathcal{M}} L(h, p) = p_2 \; .$$

2. Hierzu berechnen wir den Maximum-Likelihood-Schätzer. Das Modell ist das unbeschränkte Modell $M = \mathcal{M}(\{K, Z\})$. Der Korpus ist $h(K) = 3$ und $h(Z) = 2$. Dann ist

$$\hat{p} = \mathrm{mle}(h, M) =_{p \in \mathcal{M}} L(h, p) = \mathrm{rfe}(h) \; .$$

Also $\hat{p}(K) = 0.6$ und $\hat{p}(Z) = 0.4$.

Lösung zur 3. Aufgabe (AGS 10.4)

(a) Wir müssen drei Möglichkeiten betrachten: entweder haben wir die Würfel 1 & 2, 1 & 3 oder 2 & 3 gewählt. Das durchgeführte Zufallsexperiment unterliegt also entweder der Wahrscheinlichkeitsverteilung $p_1 \times p_2$, $p_1 \times p_3$ oder $p_2 \times p_3$ (aus Gründen der Symmetrie müssen wir die Fälle $p_2 \times p_1$, $p_3 \times p_1$ sowie $p_3 \times p_2$ nicht zusätzlich betrachten). Es gilt:

$$L(h, p_1 \times p_2) = (0.5 \cdot 0.3)^2 \cdot (0.5 \cdot 0.7)^4 \cdot (0.5 \cdot 0.3)^4 \cdot (0.5 \cdot 0.7)^6 \; ,$$
$$L(h, p_1 \times p_3) = (0.5 \cdot 0.6)^2 \cdot (0.5 \cdot 0.4)^4 \cdot (0.5 \cdot 0.6)^4 \cdot (0.5 \cdot 0.4)^6 \; ,$$
$$L(h, p_2 \times p_3) = (0.3 \cdot 0.6)^2 \cdot (0.3 \cdot 0.4)^4 \cdot (0.7 \cdot 0.6)^4 \cdot (0.7 \cdot 0.4)^6 \; .$$

Mit einem Taschenrechner kann man diese Werte ausrechnen:

$$L(h, p_1 \times p_2) \approx 3.14 \cdot 10^{-10} \; ,$$
$$L(h, p_1 \times p_3) \approx 7.46 \cdot 10^{-11} \; ,$$
$$L(h, p_2 \times p_3) \approx 1.0 \cdot 10^{-10} \; .$$

Sie haben also schätzungsweise die ersten beiden Münzen gewählt.

(b) Wir betrachten das Wahrscheinlichkeitsmodell

$$\mathcal{M} = \left\{ p_1 \times p_2 \mid p_1, p_2 \in \mathcal{M}(\{K, Z\}) \right\} \; .$$

Für dieses Modell gilt $\hat{p} = \text{mle}(h, \mathcal{M}) = \text{rfe}(h_1) \times \text{rfe}(h_2)$, wobei h_1 und h_2 jeweils $\{K, Z\}$-Korpora sind mit

$$h_1(K) = h(K,K) + h(K,Z) = 6 \,, \qquad h_1(Z) = h(Z,K) + h(Z,Z) = 10 \,,$$
$$h_2(K) = h(K,K) + h(Z,K) = 6 \,, \qquad h_2(Z) = h(K,Z) + h(Z,Z) = 10 \,.$$

Dann ist

$$\text{rfe}(h_1)(K) = \frac{3}{8} \,, \qquad \text{rfe}(h_1)(Z) = \frac{5}{8} \,,$$
$$\text{rfe}(h_2)(K) = \frac{3}{8} \,, \qquad \text{rfe}(h_2)(Z) = \frac{5}{8} \,.$$

Das sind die Wahrscheinlichkeitsverteilungen der beiden Münzen. Übrigens: da $\hat{p} = \text{rfe}(h_1) \times \text{rfe}(h_2)$, erhält man

$$\hat{p}(K,K) = \frac{9}{64} \,, \qquad \hat{p}(K,Z) = \frac{15}{64} \,, \qquad \hat{p}(Z,K) = \frac{15}{64} \,, \qquad \hat{p}(Z,Z) = \frac{25}{64} \,.$$

Lösung zur 4. Aufgabe (AGS 10.6)

(a) Das Experiment kann mit einer der Verteilungen $p_1 \times p_2$, $p_1 \times p_3$ oder $p_2 \times p_3$ durchführt werden, wobei:

$$(p_1 \times p_2)(K,K) = 0.5 \cdot 0.3 = 0.15 \,, \qquad (p_1 \times p_2)(K,Z) = 0.5 \cdot 0.7 = 0.35 \,,$$
$$(p_1 \times p_2)(Z,K) = 0.5 \cdot 0.3 = 0.15 \,, \qquad (p_1 \times p_2)(Z,Z) = 0.5 \cdot 0.7 = 0.35 \,,$$

$$(p_1 \times p_3)(K,K) = 0.5 \cdot 0.6 = 0.3 \,, \qquad (p_1 \times p_3)(K,Z) = 0.5 \cdot 0.4 = 0.2 \,,$$
$$(p_1 \times p_3)(Z,K) = 0.5 \cdot 0.6 = 0.3 \,, \qquad (p_1 \times p_3)(Z,Z) = 0.5 \cdot 0.4 = 0.2 \,,$$

$$(p_2 \times p_3)(K,K) = 0.3 \cdot 0.6 = 0.18 \,, \qquad (p_2 \times p_3)(K,Z) = 0.3 \cdot 0.4 = 0.12 \,,$$
$$(p_2 \times p_3)(Z,K) = 0.7 \cdot 0.6 = 0.42 \,, \qquad (p_2 \times p_3)(Z,Z) = 0.7 \cdot 0.4 = 0.28 \,.$$

Um die Maximum-Likelihood-Schätzung für den Korpus h mit unvollständigen Daten durchführen zu können, muss man zunächst die Wahrscheinlichkeitsverteilungen $p_1 \times p_2$, $p_1 \times p_3$ und $p_2 \times p_3$ auf das Niveau der Beobachtungen heben:

$$(p_1 \times p_2)'(=) = (p_1 \times p_2)(K,K) + (p_1 \times p_2)(Z,Z) = 0.15 + 0.35 = 0.5 \,,$$
$$(p_1 \times p_2)'(\neq) = (p_1 \times p_2)(K,Z) + (p_1 \times p_2)(Z,K) = 0.35 + 0.15 = 0.5 \,,$$

$$(p_1 \times p_3)'(=) = (p_1 \times p_3)(K,K) + (p_1 \times p_3)(Z,Z) = 0.3 + 0.2 = 0.5 \,,$$
$$(p_1 \times p_3)'(\neq) = (p_1 \times p_3)(K,Z) + (p_2 \times p_3)(Z,K) = 0.2 + 0.3 = 0.5 \,,$$

$$(p_2 \times p_3)'(=) = (p_2 \times p_3)(K,K) + (p_1 \times p_3)(Z,Z) = 0.18 + 0.28 = 0.46 \,,$$
$$(p_2 \times p_3)'(\neq) = (p_2 \times p_3)(K,Z) + (p_2 \times p_3)(Z,K) = 0.12 + 0.42 = 0.54 \,.$$

Dann ist

$$L(h, (p_1 \times p_2)') = \big((p_1 \times p_2)'(=)\big)^2 \cdot \big((p_1 \times p_2)'(\neq)\big)^3 = 0.5^2 \cdot 0.5^3 \,,$$
$$L(h, (p_1 \times p_3)') = \big((p_1 \times p_3)'(=)\big)^2 \cdot \big((p_1 \times p_3)'(\neq)\big)^3 = 0.5^2 \cdot 0.5^3 \,,$$
$$L(h, (p_2 \times p_3)') = \big((p_2 \times p_3)'(=)\big)^2 \cdot \big((p_2 \times p_3)'(\neq)\big)^3 = 0.46^2 \cdot 0.54^3 \,.$$

Offensichtlich ist $L(h,(p_2 \times p_3)')$ größer als $L(h,(p_1 \times p_2)')$ und $L(h,(p_1 \times p_3)')$. Daher haben Sie mutmaßlich die Münzen 2 & 3 gegriffen.

(b) Wir berechnen zunächst für jeden der statistischen Analysatoren den vervollständigten Korpus (E-Schritt):

$$
\begin{aligned}
h_{d_1}(K,K) &= 2 \cdot \frac{1}{2} = 1\,, & h_{d_1}(K,Z) &= 3 \cdot \frac{1}{2} = \frac{3}{2}\,,\\
h_{d_1}(Z,K) &= 3 \cdot \frac{1}{2} = \frac{3}{2}\,, & h_{d_1}(Z,Z) &= 2 \cdot \frac{1}{2} = 1\,,\\[1ex]
h_{d_2}(K,K) &= 2 \cdot \frac{1}{3} = \frac{2}{3}\,, & h_{d_2}(K,Z) &= 3 \cdot \frac{1}{9} = \frac{1}{3}\,,\\
h_{d_2}(Z,K) &= 3 \cdot \frac{8}{9} = \frac{8}{3}\,, & h_{d_2}(Z,Z) &= 2 \cdot \frac{2}{3} = \frac{4}{3}\,,\\[1ex]
h_{d_3}(K,K) &= 2 \cdot \frac{1}{4} = \frac{2}{4}\,, & h_{d_3}(K,Z) &= 3 \cdot \frac{3}{4} = \frac{9}{4}\,,\\
h_{d_3}(Z,K) &= 3 \cdot \frac{1}{4} = \frac{3}{4}\,, & h_{d_3}(Z,Z) &= 2 \cdot \frac{3}{4} = \frac{6}{4}\,.
\end{aligned}
$$

Zu diesen Korpora bestimmen wir den Maximum-Likelihood-Schätzer (M-Schritt). Dabei geht man wie in Aufgabe 3(b) vor. Wir erhalten die folgenden geschätzten Wahrscheinlichkeitsverteilungen für die Münzen:

	Wahrscheinlichkeit für Kopf von Münze 1	Zahl Münze 1	Kopf Münze 2	Zahl Münze 2
d_1	1/2	1/2	1/2	1/2
d_2	1/5	4/5	2/3	1/3
d_3	11/20	9/20	1/4	3/4

Das führt zu den folgenden Wahrscheinlichkeitsverteilungen auf der Ergebnismenge $\{(K,K),(K,Z),(Z,K),(Z,Z)\}$:

	$p(K,K)$	$p(K,Z)$	$p(Z,K)$	$p(Z,Z)$
d_1	1/4	1/4	1/4	1/4
d_2	2/15	1/15	8/15	4/15
d_3	11/80	33/80	9/80	27/80

Wir leiten die folgenden bedingten Wahrscheinlichkeiten ab:

	$p\big((K,K)\mid=\big)$	$p\big((K,Z)\mid\neq\big)$	$p\big((Z,K)\mid\neq\big)$	$p\big((Z,Z)\mid=\big)$
d_1	1/2	1/2	1/2	1/2
d_2	1/3	1/9	8/9	2/3
d_3	11/38	33/42	9/42	27/38

Nur die statistischen Analysatoren d_1 und d_2 sind konsistent!