

# ALGORITHMEN UND DATENSTRUKTUREN

## ÜBUNG 13: WAHRSCHEINLICHKEITEN

---

Eric Kunze

`eric.kunze@tu-dresden.de`

TU Dresden, 19. Januar 2022

letzte Änderung:  
19.01.2022, 11:44

# Aufgabe 1

---

Wir betrachten ein Zufallsexperiment  $(X, p)$  mit

- ▶ Ergebnismenge  $X$  und
- ▶ einer Funktion  $p: X \rightarrow [0, 1]$  mit  $\sum_{x \in X} p(x) = 1$   
(**Wahrscheinlichkeitsverteilung von  $X$** )

Wir betrachten ein Zufallsexperiment  $(X, p)$  mit

- ▶ Ergebnismenge  $X$  und
- ▶ einer Funktion  $p: X \rightarrow [0, 1]$  mit  $\sum_{x \in X} p(x) = 1$   
(**Wahrscheinlichkeitsverteilung von  $X$** )

Die Menge aller Wahrscheinlichkeitsverteilungen über  $X$  sei  $\mathcal{M}(X)$ . Jede Teilmenge  $\mathcal{M} \subseteq \mathcal{M}(X)$  heißt **Wahrscheinlichkeitsmodell**.

Wir betrachten ein Zufallsexperiment  $(X, p)$  mit

- ▶ Ergebnismenge  $X$  und
- ▶ einer Funktion  $p: X \rightarrow [0, 1]$  mit  $\sum_{x \in X} p(x) = 1$   
(**Wahrscheinlichkeitsverteilung von  $X$** )

Die Menge aller Wahrscheinlichkeitsverteilungen über  $X$  sei  $\mathcal{M}(X)$ . Jede Teilmenge  $\mathcal{M} \subseteq \mathcal{M}(X)$  heißt **Wahrscheinlichkeitsmodell**.

Ein Wahrscheinlichkeitsmodell  $\mathcal{M}$  heißt **beschränkt**, falls  $\mathcal{M} \neq \mathcal{M}(X)$ ; andernfalls unbeschränkt.

Wir betrachten ein Zufallsexperiment  $(X, p)$  mit

- ▶ Ergebnismenge  $X$  und
- ▶ einer Funktion  $p: X \rightarrow [0, 1]$  mit  $\sum_{x \in X} p(x) = 1$   
(**Wahrscheinlichkeitsverteilung von  $X$** )

Die Menge aller Wahrscheinlichkeitsverteilungen über  $X$  sei  $\mathcal{M}(X)$ . Jede Teilmenge  $\mathcal{M} \subseteq \mathcal{M}(X)$  heißt **Wahrscheinlichkeitsmodell**.

Ein Wahrscheinlichkeitsmodell  $\mathcal{M}$  heißt **beschränkt**, falls  $\mathcal{M} \neq \mathcal{M}(X)$ ; andernfalls unbeschränkt.

Führen wir nun zwei Zufallsexperimente nacheinander aus und nehmen dabei an, dass die beiden Experimente *unabhängig* voneinander sind. Folge das erste Experiment einer Verteilung  $p_1 \in \mathcal{M}(X_1)$  und das zweite Experiment einer Verteilung  $p_2 \in \mathcal{M}(X_2)$ , dann ist  $p_1 \times p_2 \in \mathcal{M}(X_1 \times X_2)$  eine Verteilung auf der Ergebnismenge  $X_1 \times X_2$  unseres zweistufigen Experiments:

$$(p_1 \times p_2)(a, b) = p_1(a) \cdot p_2(b).$$

„Einzelwahrscheinlichkeiten multiplizieren / erste Pfadregel“

## AUFGABE 1 – TEIL (A)

Beschreiben Sie die folgenden Zufallsexperimente formal durch Angabe einer Ergebnismenge und der zugehörigen Wahrscheinlichkeitsverteilung:

1. Das Werfen von zwei fairen Würfeln.
2. Das gleichzeitige Ziehen von jeweils einer Kugel aus zwei Urnen. Die erste Urne hat zwei weiße und drei schwarze Kugeln. Die zweite Urne hat eine rote, zwei blaue Kugeln und drei grüne Kugeln. Die Kugeln sind nicht unterscheidbar (abgesehen von ihrer Farbe).
3. Das sechsmalige Werfen einer Münze, die mit Wahrscheinlichkeit 0.3 Kopf ergibt.

# AUFGABE 1 – TEIL (A)

## 1. Werfen von zwei fairen Würfeln:

$$X = \{1, \dots, 6\} \times \{1, \dots, 6\} \quad \text{und} \quad p(a, b) = \frac{1}{36} \quad \forall (a, b) \in X$$

## 2. 2-Urnenmodell: Durch das *gleichzeitige* Ziehen können wir beide Versuche *unabhängig* voneinander betrachten.

$$\text{Urne 1: } X_1 = \{w, s\} \quad \text{mit} \quad p_1(w) = \frac{2}{5} \quad p_1(s) = \frac{3}{5}$$

$$\text{Urne 2: } X_2 = \{r, g, b\} \quad \text{mit} \quad p_2(r) = \frac{1}{6} \quad p_2(g) = \frac{1}{2} \quad p_2(b) = \frac{1}{3}$$

Der Gesamtversuch ist damit das unabhängige Produkt

$$X := X_1 \times X_2$$

$$p(a, b) := (p_1 \times p_2)(a, b) = p_1(a) \cdot p_2(b) \quad \forall (a, b) \in X_1 \times X_2$$

Genauer ist

$$\begin{array}{lll} p(w, r) = \frac{1}{15} & p(w, g) = \frac{2}{15} & p(w, b) = \frac{1}{5} \\ p(s, r) = \frac{1}{10} & p(s, g) = \frac{1}{5} & p(s, b) = \frac{3}{10} \end{array}$$



3. „Unfairer“ Münzwurf: Für *einen* Münzwurf gilt

$$X' = \{K, Z\}$$

$$p'(K) = 0.3$$

$$p'(Z) = 0.7.$$

Für den sechsfachen Münzwurf gilt dementsprechend

$$X := (X')^6 = \{K, Z\}^6$$

$$p(x) := p((x_1, x_2, \dots, x_6))$$

$$= (p' \times \dots \times p')(x_1, \dots, x_6)$$

$$= p'(x_1) \cdot p'(x_2) \cdot \dots \cdot p'(x_6)$$

Genauer kann man auch schreiben, dass

$$p(x) = 0.3^{K(x)} \cdot 0.7^{Z(x)}$$

wobei  $K(x)$  die Anzahl der Kopfwürfe in der Wurffolge  $x$  beschreibt und analog  $Z(x)$  die Anzahl der Zahlwürfe.

## AUFGABE 1 – TEIL (B)

Geben Sie für die folgenden Szenarien jeweils die Ergebnismenge und das Wahrscheinlichkeitsmodell an. Bestimmen Sie ob das Wahrscheinlichkeitsmodell beschränkt oder unbeschränkt ist. Wenn das Modell beschränkt ist, dann geben Sie eine Wahrscheinlichkeitsverteilung an, die sich nicht in dem Modell befindet.

1. Das Werfen eines Tetraeders mit beliebigen Eigenschaften.
2. Das Werfen eines Würfels, bei dem alle Seiten mit gerader Augenzahl die gleiche Wahrscheinlichkeit haben.
3. Das zweimalige Werfen der gleichen Münze, diese darf beliebige Eigenschaften haben.

## AUFGABE 1 – TEIL (B)

1. **Werfen eines Tetraeders mit beliebigen Eigenschaften:** Sind die Seiten mit  $1, \dots, 4$  durchnummeriert, dann ist  $X = \{1, \dots, 4\}$ . Das Modell ist unbeschränkt, d.h. es gilt  $\mathcal{M} = \mathcal{M}(X)$ .
2. **Werfen eines Würfels, bei dem alle geraden Augenzahlen gleiche Wahrscheinlichkeit haben:** Es ist  $X = \{1, \dots, 6\}$  und das Modell ist gegeben durch

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

Klarerweise gibt es Verteilungen, die nicht in  $\mathcal{M}$  liegen, z.B.  $p$  mit

$$p(1) = p(2) = p(3) = \frac{1}{12} \quad \text{und} \quad p(4) = p(5) = p(6) = \frac{1}{4}.$$

Somit ist  $\mathcal{M}$  beschränkt.

3. **Zweimaliger Münzwurf:** Es ist

$$X = \{K, Z\}^2 = \{K, Z\} \times \{K, Z\} = \{(K, K), (K, Z), (Z, K), (Z, Z)\}.$$

Da zweimal die selbe Münze geworfen wird, gilt

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

Jede Verteilung  $p$  mit  $p(K, Z) \neq p(Z, K)$  liegt nicht in  $\mathcal{M}$  und  $\mathcal{M}$  ist somit beschränkt.

## Aufgabe 2

---

# KORPORA UND KORPUSWAHRSCHEINLICHKEITEN

Oftmals wissen wir aber die zugrundeliegende Verteilung nicht, sondern können lediglich die Ergebnisse des Experiments wahrnehmen. Zählen wir diese Beobachtungen, dann nennen wir das einen **X-Korpus** modelliert durch eine Funktion  $h: X \rightarrow \mathbb{R}^{\geq 0}$ . Man definiert die **Korpuswahrscheinlichkeit / Likelihood** von  $h$  unter einer Verteilung  $p$  als

$$L(h, p) = \prod_{x \in X} p(x)^{h(x)}.$$

Kennen wir nun aber die Verteilung  $p$  nicht, müssen wir sie aus den beobachteten Daten schätzen. Dies macht zum Beispiel der **Maximum-Likelihood-Schätzer** (MLE)

$$\text{mle}(h, \mathcal{M}) = \arg \max_{p \in \mathcal{M}} L(h, p).$$

Solange das Modell unbeschränkt gewählt wird, d.h. es werden alle Verteilungen über  $X$  zugelassen, dann ist der MLE genau die relative Häufigkeit von  $h$ .

## AUFGABE 2 – TEIL (A)

Sie haben zwei Würfel, davon ist einer fair und bei dem anderen ist jede der Seiten mit geraden Augenzahlen drei Mal so wahrscheinlich wie jede der Seiten mit ungeraden Augenzahlen. Sie wählen einen der beiden Würfel aus und erzeugen bei zehn Würfeln das folgende Korpus:

$$h(1) = 1, \quad h(2) = 2, \quad h(3) = 2, \quad h(4) = 1, \quad h(5) = 1, \quad h(6) = 3.$$

Welchen Würfel haben Sie wahrscheinlich gegriffen?

## AUFGABE 2 – TEIL (A)

Wir bestimmen für die Verteilungen der beiden Würfel die jeweilige Likelihood unter dem Korpus  $h$ .

**Verteilung des ersten Würfels:**

$$p_1(1) = p_1(2) = p_1(3) = p_1(4) = p_1(5) = p_1(6) = \frac{1}{6}$$

**Verteilung des zweiten Würfels:**

$$p_2(1) = p_2(3) = p_2(5) = \frac{1}{12} \quad \text{und} \quad p_2(2) = p_2(4) = p_2(6) = \frac{1}{4}$$

**Likelihoods:**

$$\begin{aligned} L(h, p_1) &= \prod_{x=1}^6 p_1(x)^{h(x)} = \left(\frac{1}{6}\right)^1 \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^1 \cdot \left(\frac{1}{6}\right)^1 \cdot \left(\frac{1}{6}\right)^3 \\ &= \left(\frac{1}{6}\right)^{|h|} = \left(\frac{1}{6}\right)^{10} \end{aligned}$$

$$\begin{aligned} L(h, p_2) &= \prod_{x=1}^6 p_2(x)^{h(x)} = \left(\frac{1}{12}\right)^1 \cdot \left(\frac{1}{4}\right)^2 \cdot \left(\frac{1}{12}\right)^2 \cdot \left(\frac{1}{4}\right)^1 \cdot \left(\frac{1}{12}\right)^1 \cdot \left(\frac{1}{4}\right)^3 \\ &= \left(\frac{1}{12}\right)^4 \cdot \left(\frac{1}{4}\right)^6 \end{aligned}$$

Wir müssen nun noch entscheiden, welche der beiden Likelihoods größer ist – vorzugsweise ohne Taschenrechner:

$$L(h, p_1) > L(h, p_2)$$

$$\Leftrightarrow \left(\frac{1}{6}\right)^{10} > \left(\frac{1}{12}\right)^4 \cdot \left(\frac{1}{4}\right)^6 = \left(\frac{1}{3} \cdot \frac{1}{4}\right)^4 \cdot \left(\frac{1}{4}\right)^6 = \left(\frac{1}{3}\right)^4 \cdot \left(\frac{1}{4}\right)^{10}$$

$$\Leftrightarrow \left(\frac{1}{6} \cdot \frac{4}{1}\right)^{10} = \left(\frac{2}{3}\right)^{10} > \left(\frac{1}{3}\right)^4$$

$$\Leftrightarrow 3^4 > \left(\frac{3}{2}\right)^{10} = \frac{3^{10}}{2^{10}}$$

$$\Leftrightarrow 3^{-6} > \frac{1}{2^{10}}$$

$$\Leftrightarrow 2^{10} > 3^6$$

$$\Leftrightarrow 2^5 > 3^3$$

$$\Leftrightarrow 32 > 27$$

Somit ist also  $L(h, p_1) > L(h, p_2)$  und wir haben sicherlich den ersten Würfel genutzt.



## AUFGABE 2 – TEIL (B)

Sie haben zwei Würfel, davon ist einer fair und bei dem anderen ist jede der Seiten mit geraden Augenzahlen drei Mal so wahrscheinlich wie jede der Seiten mit ungeraden Augenzahlen. Sie wählen einen der beiden Würfel aus und erzeugen bei zehn Würfeln das folgende Korpus:

$$h(1) = 1, \quad h(2) = 2, \quad h(3) = 2, \quad h(4) = 1, \quad h(5) = 1, \quad h(6) = 3.$$

Nehmen Sie an, dass Sie das Korpus  $h$  mit einem beliebigen Würfel erzeugt haben. Schätzen Sie die Wahrscheinlichkeitsverteilung des Würfels ab.

## AUFGABE 2 – TEIL (B)

Der Würfel wurde beliebig gewählt, d.h. ohne besondere Eigenschaften oder Anforderungen an die Verteilung. Damit betrachten wir das unbeschränkte Wahrscheinlichkeitsmodell  $\mathcal{M} = \mathcal{M}(X)$ . Wir erhalten den Maximum Likelihood Schätzer  $\hat{p}$  von  $h$  und  $\mathcal{M}$  als relative Häufigkeiten:

$$\hat{p} = \text{mle}(h, \mathcal{M}(X)) = \text{rfe}(h).$$

Für die relativen Häufigkeiten gilt die Formel

$$\text{rfe}(h)(x) := \frac{h(x)}{|h|} \quad \text{mit} \quad |h| := \sum_{x \in X} h(x).$$

Es gilt also

$$\begin{aligned} \hat{p}(1) &= \frac{1}{10} & \hat{p}(4) &= \frac{1}{10} \\ \hat{p}(2) &= \frac{2}{10} & \hat{p}(5) &= \frac{1}{10} \\ \hat{p}(3) &= \frac{2}{10} & \hat{p}(6) &= \frac{3}{10} \end{aligned}$$