Vorlesung 3 Wahrscheinlichkeitstheorie

Mit Hilfe der Wahrscheinlichkeitstheorie können wir "Zufall" in Daten modellieren. Einem Ereignis A wollen wir dafür eine "Wahrscheinlichkeit" $P(A) \in [0,1]$ zuordnen.

Es gibt zwei grundlegende Interpretationen, wie Zufall durch Angabe von $P(A)$ zu verstehen.

1) $P(A) \approx$ relative Häufigkeit des Ereignisses A, wobei die relative Häufigkeit aus n Zufallsexperimenten berechnet.
   D.h. $P(A) \approx \frac{k}{n}$, $k =$ Anzahl der Experimente mit Ausgang.
   Außerdem soll $\approx$ zu $=$ werden, wenn $n \to \infty$.
   Dieser Ansatz wird frequentistischer Wahrscheinlichkeitsbegriff genannt.

2) $P(A)$ ist ein Erfahrungswert, der anhand von beobachteten Daten generiert wird. Insbesondere ist $P(A)$ keine vom Beobachter unabhängige Größe und kann sich durch neue Daten ändern. D.h., unvollständige Information über deterministische Prozesse lässt sich durch $P(A)$ modellieren. Dieser Ansatz wird Bayes'scher Wahrscheinlichkeitsbegriff genannt.

Beide Ansätze haben ihre Berechtigung. Die mathematische Definition von Wahrscheinlichkeit ist unabhängig davon und wie folgt:

Definition 2.1

Sei Ω eine nichtleere Menge und $\mathcal{A} \subset 2^{\Omega} = \{ A \mid A \subset \Omega \}$.

Wir nennen $\mathcal{A}$ eine σ-Algebra, falls

1) $\Omega \in \mathcal{A}$

2) Falls $A \in \mathcal{A}$, dann $\Omega \setminus A \in \mathcal{A}$.

3) Falls $A_n, n \in \mathbb{N}$, dann auch $\bigcup_{n \in \mathbb{N}} A_n$.

Definition 2.2

Ein Wahrscheinlichkeitsraum ist ein Tripel $(\Omega, \mathcal{A}, P)$, wobei

1) Ω eine nicht-leere Menge ist

2) $\mathcal{A}$ ist eine σ-Algebra in 2^{Ω}

3) $P: \mathcal{A} \to [0,1]$ ein Wahrscheinlichkeitsmaß.

D.h: $P(\Omega) = 1$ und

$$\mathbb{P}\Big(\bigcup_{i \in \mathbb{N}} A_n\Big) = \sum_{i=1}^{\infty} P(A_i) \quad , A_i \cap A_j = \emptyset \text{ für } i \neq j$$

Ω heißt Ereignisraum, $A \in \mathcal{A}$ heißt Ereignis und $P(A)$ heißt die Wahrscheinlichkeit von A.

Beispiel 2.3.

$\Omega = \{0,1\}$, $\mathcal{A} = \{\emptyset, \{0\}, \{1\}, \Omega\}$

$P(\{0\}) = 1/2$, $P(\{0\}) + P(\{1\}) = P(\{0,1\}) = P(\Omega) = 1$

$\to P(\{1\}) = 1/2$.

Oftmals für uns ist $\Omega = \mathbb{R}$ oder $\Omega = \mathbb{R}^n$. Hier wählen wir immer immer die sog. Borel-σ-Algebra, die σ-Algebra, die durch Intervalle / Boxen erzeugt wird.

Definition 2.4

Eine Zufallsvariable X ist eine Abbildung zwischen Wahrscheinlichkeitsräumen

$$X: (\Omega', \mathcal{A}', P') \to (\Omega, \mathcal{A}, P)$$

so dass für alle $A \in \mathcal{A}$ gilt

$$X^{-1}(A) \in \mathcal{A}' \qquad = \{a \in \Omega' \mid X(a) \in A\}$$

und

$$P(A) = P'(X^{-1}(A))$$

Falls $\Omega = \mathbb{R}$ oder $\mathbb{R}^n$ nennen wir X eine reelle Zufallsvariable

Wir schreiben $P(A) =: P(X \in A)$.

Falls der Wertebereich von X diskret ist, nennen wir X diskret

— " — kontinuierlich, — " — kontinuierlich.

Beispiel 2.5

Ω' = Menge aller Münzwürfe

$X : \Omega' \to [0,1]$, $X(\omega) = 0$, falls der Wurf ω auf Kopf landet

$X(\omega) = 1$, — " — Zahl landet.

Wir wollen eine Wahrscheinlichkeit definieren, im Fall das ein Ereignis bereits eingetreten ist.

Definition 2.6 (Bedingte Whr.)

Die bedingte Wahrscheinlichkeit von A gegeben B (für $P(B) > 0$) ist

$$P[A \mid B) := \frac{P(A \cap B)}{P(B)}$$

Beispiel 2.7

$\Omega = \{1, \dots, 6\}$, $A = \{2\}$, $B = \{2, 4, 6\}$ und $P(\{k\}) = 1/6$ für $k = 1, \dots, 6$.

$$P(A) = 1/6, \qquad P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2} = 1/3$$

Theorem 2.8 (Satz von Bayes').

Seien A und B Ereignisse mit $P(A), P(B) > 0$. Dann gilt:

$$P(A \mid B) = P(B \mid A) \, \frac{P(A)}{P(B)}$$

Beweis

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B \mid A) = \frac{P(A \cap B)}{P(A)} \qquad \square$$

Oft wollen wir $P(A \mid B)$ verstehen und können $P(B \mid A)$ berechnen. Der Satz von Bayes hilft uns beide in Zusammenhang zu bringen.