Wiederholung Wahrscheinlichkeitstheorie

Carsten Gips (HSBI)

Ereignisse und Wahrscheinlichkeit

- **Ereignisse** $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$: Ausgänge eines Zufallsexperiments
- **Elementarereignis**: Die $\omega_i \in \Omega$
  - decken *alle* möglichen Versuchsergebnisse ab, und
  - schließen sich gegenseitig aus
- **Wahrscheinlichkeit**:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der für A günstigen Fälle}}{\text{Anzahl der möglichen Fälle}}$$

Tafel: Würfeln

Rechenregeln: Kolmogorov Axiome

- $0 \leq P(A) \leq 1$
- $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$: $\sum_i P(\omega_i) = 1$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Daraus folgt (u.a.):

- $P(\Omega) = 1$
- $P(\emptyset) = 0$
- $P(A) = 1 - P(\neg A)$

Tafel: Würfeln: Elementar, gerade Zahl

Verbundwahrscheinlichkeiten

$$P(A, B) = P(B, A) = \text{ Wahrscheinlichkeit, dass A und B gleichzeitig auftreten}$$

	Halsschmerzen	$\neg$ Halsschmerzen
Schnupfen	0.04	0.06
$\neg$ Schnupfen	0.01	0.89

- $P(S, H) = 0.04$

Bedingte Wahrscheinlichkeit

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

	Halsschmerzen	¬ Halsschmerzen
Schnupfen	0.04	0.06
¬ Schnupfen	0.01	0.89

- $P(\text{Schnupfen} \mid \text{Halsschmerzen}) = \frac{P(S,H)}{P(H)} = \frac{0.04}{0.04+0.01} = 0.8$
- $P(\text{Halsschmerzen} \mid \text{Schnupfen}) = \frac{P(H,S)}{P(S)} = \frac{0.04}{0.04+0.06} = 0.4$

Marginalisierung

	Halsschmerzen	$\neg$ Halsschmerzen	$\sum$
Schnupfen	0.04	0.06	*0.1*
$\neg$ Schnupfen	0.01	0.89	*0.9*
$\sum$	*0.05*	*0.95*	*1*

$P(S) = P(S, H) + P(S, \neg H)$

Marginalisierung

	Halsschmerzen	$\neg$ Halsschmerzen	$\sum$
Schnupfen	0.04	0.06	*0.1*
$\neg$ Schnupfen	0.01	0.89	*0.9*
$\sum$	*0.05*	*0.95*	*1*

$P(S) = P(S, H) + P(S, \neg H)$

Seien $B_1, \ldots, B_n$ Elementarereignisse mit $\bigcup_i B_i = \Omega$. Dann ist

$$P(A) = \sum_i P(A, B_i) = \sum_i P(A|B_i)P(B_i)$$

Kettenregel

- **Produktregel**: Wegen $P(A|B) = \dfrac{P(A,B)}{P(B)}$ gilt $P(A,B) = P(A|B)P(B)$

- Verallgemeinerung (**Kettenregel**):

$$
\begin{aligned}
P(A_1, A_2, \ldots, A_n) &= P(A_n, \ldots, A_2, A_1) \\
&= P(A_n|A_{n-1}, \ldots, A_1)P(A_{n-1}, \ldots, A_1) \\
&= P(A_n|A_{n-1}, \ldots, A_1)P(A_{n-1}|A_{n-2}, \ldots, A_1)P(A_{n-2}, \ldots, A_1) \\
&= \ldots \\
&= P(A_n|A_{n-1}, \ldots, A_1)\ldots P(A_2|A_1)P(A_1) \\
&= \prod_i P(A_i|A_1, \ldots, A_{i-1})
\end{aligned}
$$

Bayes-Regel

Bedingte Wahrscheinlichkeit: $P(A, B) = P(A|B)P(B) = P(B|A)P(A)$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- $P(A)$ nennt man **"Prior"** oder **"A-priori-Wahrscheinlichkeit"**
- $P(B|A)$ nennt man **"Likelihood"**
- $P(A|B)$ nennt man **"Posterior"** oder **"A-posteriori-Wahrscheinlichkeit"**
- $P(B)$ ist ein Normierungsfaktor

Beispiel Bayes

- Bei Arthrose wird in 80 Prozent der Fälle ein steifes Gelenk beobachtet
- Eine von 10.000 Personen hat Arthrose
- Eine von 10 Personen hat ein steifes Gelenk

=> Ich habe ein steifes Gelenk. Habe ich Arthrose?

Beispiel Bayes

- Bei Arthrose wird in 80 Prozent der Fälle ein steifes Gelenk beobachtet
- Eine von 10.000 Personen hat Arthrose
- Eine von 10 Personen hat ein steifes Gelenk

=> Ich habe ein steifes Gelenk. Habe ich Arthrose?

- Gegeben: $P(A) = 0.0001, P(S) = 0.1, P(S|A) = 0.8$
- Gesucht: $P(A|S)$

Beispiel Bayes

- Bei Arthrose wird in 80 Prozent der Fälle ein steifes Gelenk beobachtet
- Eine von 10.000 Personen hat Arthrose
- Eine von 10 Personen hat ein steifes Gelenk

=> Ich habe ein steifes Gelenk. Habe ich Arthrose?

- Gegeben: $P(A) = 0.0001, P(S) = 0.1, P(S|A) = 0.8$
- Gesucht: $P(A|S)$

$$P(A|S) = \frac{P(S|A)P(A)}{P(S)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008 = 0.08\%$$

Beispiel Bayes

- Bei Arthrose wird in 80 Prozent der Fälle ein steifes Gelenk beobachtet
- Eine von 10.000 Personen hat Arthrose
- Eine von 10 Personen hat ein steifes Gelenk

=> Ich habe ein steifes Gelenk. Habe ich Arthrose?

- Gegeben: $P(A) = 0.0001, P(S) = 0.1, P(S|A) = 0.8$
- Gesucht: $P(A|S)$

$$P(A|S) = \frac{P(S|A)P(A)}{P(S)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008 = 0.08\%$$

=> Wie wahrscheinlich ist ein steifes Gelenk ohne Arthrose, also $P(S|\neg A)$?

Beispiel Bayes

- Bei Arthrose wird in 80 Prozent der Fälle ein steifes Gelenk beobachtet
- Eine von 10.000 Personen hat Arthrose
- Eine von 10 Personen hat ein steifes Gelenk

=> Ich habe ein steifes Gelenk. Habe ich Arthrose?

- Gegeben: $P(A) = 0.0001, P(S) = 0.1, P(S|A) = 0.8$
- Gesucht: $P(A|S)$

$$P(A|S) = \frac{P(S|A)P(A)}{P(S)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008 = 0.08\%$$

=> Wie wahrscheinlich ist ein steifes Gelenk ohne Arthrose, also $P(S|\neg A)$?

Mit Marginalisierung: $P(S) = P(S|A)P(A) + P(S|\neg A)P(\neg A)$,
d.h. $0.1 = 0.8 \times 0.0001 + P(S|\neg A) \times (1 - 0.0001)$, d.h. $P(S|\neg A) = 0.0999$

Unabhängige Ereignisse

- $P(\text{Halsschmerzen},\ \text{Regen}) = P(\text{Regen} \mid \text{Halsschmerzen})P(\text{Halsschmerzen})$
- $P(\text{Regen} \mid \text{Halsschmerzen}) =\ ??$

Unabhängige Ereignisse

- $P(\text{Halsschmerzen},\ \text{Regen}) = P(\text{Regen} \mid \text{Halsschmerzen})P(\text{Halsschmerzen})$
- $P(\text{Regen} \mid \text{Halsschmerzen}) = \ ?? \ = P(\text{Regen})$

Unabhängige Ereignisse

- $P(\text{Halsschmerzen}, \ \text{Regen}) = P(\text{Regen} \mid \text{Halsschmerzen}) P(\text{Halsschmerzen})$
- $P(\text{Regen} \mid \text{Halsschmerzen}) = \ ?? \ = P(\text{Regen})$

- Zwei Ereignisse A und B sind **unabhängig**, wenn

$$P(A|B) = P(A)$$

$=> P(A, B) = P(A|B)P(B) = P(A)P(B)$

Wrap-Up

- Grundlagen der Wahrscheinlichkeitstheorie
    - Elementarereignisse und Wahrscheinlichkeit
    - Rechenregeln
    - Bedingte Wahrscheinlichkeit und Verbundwahrscheinlichkeit
    - Marginalisierung
    - (Bedingte) Unabhängigkeit
    - Bayes'sche Regel

LICENSE