

§64: BEDINGTE WAHRSCHEINLICHKEITEN

64.1 Motivation

Im §61 haben wir den Begriff der Wahrscheinlichkeit (WS) kennengelernt. Oft hat man Zusatzinformationen, mit denen sich präzisere Wahrscheinlichkeitsaussagen machen lassen. Dies führt zu den so genannten bedingten Wahrscheinlichkeiten.

64.2 Beispiel:

Aufteilung der 1500 Angehörigen eines Betriebs nach Geschlecht und Raucherwohnheiten:

	Frauen (B)	Männer ($\bar{B}$)
Raucher (A)	600	200
Nichtraucher ($\bar{A}$)	300	400

Ω : Menge der Betriebsangehörigen

A : Menge der Raucher ($\bar{A} = \Omega \setminus A$: Nichtraucher)

B : Menge der Frauen ($\bar{B} = \Omega \setminus B$: Männer)

Wir lösen eine Person zufällig aus: Dabei treffen folgende Wahrscheinlichkeiten zu: ($|A|$: Mächtigkeit von A)

$$\begin{aligned}P(A) &= \frac{|A|}{|\Omega|} = \frac{800}{1500} = \frac{8}{15} && \text{Raucheranteil} \\P(B) &= \frac{|B|}{|\Omega|} = \frac{900}{1500} = \frac{3}{5} && \text{Frauenanteil} \\P(A \cap B) &= \frac{|A \cap B|}{|\Omega|} = \frac{600}{1500} = \frac{2}{5} && \text{Anteil der rauchenden Frauen}\end{aligned}$$

Wie groß ist die Wahrscheinlichkeit $P(A|B)$, dass eine Person raucht, wenn es sich um eine Frau handelt?

$$P(A|B) = \frac{|A \cap B|}{|B|} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{P(A \cap B)}{P(B)} = \frac{\frac{2}{5}}{\frac{3}{5}} = \frac{2}{3} \quad (*)$$

Man nennt $P(A|B)$ die bedingte Wahrscheinlichkeit von A unter der Bedingung (Hypothese) B . Es gilt stets (vgl. $(*)$)

$$P(A \cap B) = P(A|B) \cdot P(B)$$

Zwei Ereignisse A und B heißen unabhängig, wenn gilt:

$$P(A \cap B) = P(A) \cdot P(B)$$

64.3 Verallgemeinerung:

Seien $A_1, \dots, A_n \subset \Omega$ und $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$. Dann gilt:

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \dots \cdot P(A_n|A_1 \cap \dots \cap A_{n-1})$$

Beweis:

Die rechte Seite lässt sich schreiben als

$$P(A_1) \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \cdot \dots \cdot \frac{P(A_1 \cap \dots \cap A_n)}{P(A_1 \cap \dots \cap A_{n-1})} = P(A_1 \cap \dots \cap A_n)$$

64.4 Beispiel:

Wie groß ist die Wahrscheinlichkeit, mit 6 Würfeln 6 verschiedene Zahlen zu würfeln?

A_1 : irgendein Ergebnis für Würfel 1

A_2 : ein vom ersten Ergebnis verschiedenes Ergebnis für Würfel 2

$\vdots$

A_6 : ein von $A_1, \dots, A_5$ verschiedenes Ergebnis für Würfel 6

Mit 64.3 gilt:

$$\begin{aligned} P(A_1 \cap \dots \cap A_6) &= P(A_1) \cdot P(A_2|A_1) \cdot \dots \cdot P(A_6|A_1 \cap \dots \cap A_5) \\ &= 1 \cdot \frac{5}{6} \cdot \frac{4}{6} \cdot \dots \cdot \frac{1}{6} = \frac{6!}{6^6} \approx 0.015 \end{aligned}$$

64.5 Beispiel:

Ein Krebstest ist mit 96%iger Wahrscheinlichkeit positiv, falls der Patient Krebs hat, mit 94%iger Wahrscheinlichkeit negativ, falls er keinen Krebs hat. Bei einem Patienten, in dessen Altersgruppe 0.5% aller Personen Krebs haben, verläuft der Test positiv. Wie groß ist die Wahrscheinlichkeit, dass er tatsächlich Krebs hat?

Ereignis K : Patient hat Krebs

Ereignis T : Test ist positiv

$$P(K|T) = \frac{P(K \cap T)}{P(T)} = \frac{0.005 \cdot 0.96}{0.005 \cdot 0.96 + 0.995 \cdot 0.06} \approx 0.074$$

Der Patient kann also noch relativ beruhigt sein.

Fazit: Um eine seltene Krankheit zuverlässig zu erkennen, darf ein Test nur sehr wenige „false positives“ haben.

64.6 Satz: (Satz von der totalen Wahrscheinlichkeit)

Sei $\Omega = \bigcup_{i=1}^n A_i$ (disjunkt vereinigt) eine Partition (disjunkte Vereinigung) von Ω in mögliche Ereignisse $A_i, i = 1, \dots, n$. Ferner sei $P(A_i) > 0$ für alle i . Dann gilt für jedes Ereignis B :

$$P(B) = \sum_{i=1}^n P(A_i) \cdot P(B|A_i)$$

Ein Fahrer startet bei 0 und fährt mit Wahrscheinlichkeit $P(A_1), \dots, P(A_n)$ zu $A_1, \dots, A_n$. Die Wahrscheinlichkeit, von dort nach B zu fahren, beträgt $P(B|A_1), \dots, P(B|A_n)$. Die Gesamtwahrscheinlichkeit, dass der Fahrer nach B gelangt, ist die Summe der Wahrscheinlichkeiten aller Pfade von 0 nach B .

$$P(B) = \sum_{i=1}^n P(A_i) \cdot P(B|A_i)$$

Bemerkung: Im Beispiel 64.5 haben wir im Nenner bereits den Satz von der totalen Wahrscheinlichkeit verwendet.

64.7 Beispiel:

Um einen binären Nachrichtenkanal robuster gegenüber Störungen zu machen, sendet man die Bitfolge 0000000 statt 0 und 1111111 statt 1. Störungen treten in 20% aller Fälle auf und die Wahrscheinlichkeit für 0 sei 0,1.

Es wird 0100110 empfangen: Wie groß ist die Wahrscheinlichkeit, dass 0000000 gesendet wurde?

$$P(0000000|0100110) = \frac{0.1 \cdot 0.2^3 \cdot 0.8^4}{0.1 \cdot 0.2^3 \cdot 0.8^4 + 0.9 \cdot 0.2^4 \cdot 0.8^3} \approx 0.308$$

Man wird den Block also als 1 lesen, obwohl die Mehrzahl der Bits Nullen sind! Eng verwandt mit dem Satz von der totalen Wahrscheinlichkeit ist

64.8 Satz: (Formel von Bayes)

Sei $P(B) > 0$ und seien die Voraussetzungen von Satz 64.6 erfüllt. Dann gilt:

$$P(A_k|B) = \frac{P(A_k) \cdot P(B|A_k)}{\sum_{i=1}^n P(A_i) \cdot P(B|A_i)}$$

Beweis:

Folgt aus $P(A_k|B) = \frac{P(A_k|B)}{P(B)}$, in dem man

$$P(B) = \sum_{i=1}^n P(A_i) \cdot P(B|A_i) \quad (64.6)$$

$$P(A_k \cap B) = P(A_k) \cdot P(B|A_k) \quad (64.2)$$

einsetzt.

Bemerkung:

Im Satz 64.8 wird $P(A_k|B)$ aus $P(B|A_i), i = 1, \dots, n$ berechnet, d.h. Ursache und Wirkung kehren sich um. Eine typische Anwendung besteht darin, dass man weiß, welche Ursachen welche Auswirkungen haben. Hat man eine Auswirkung gemessen, fragt man sich nach der wahrscheinlichsten Ursache (inverses Problem).

64.9 Anwendungsbeispiele

- (a) Ein Arzt beobachtet bei einem Patienten ein Symptom B . Es kann von n verschiedenen Krankheiten $A_1, \dots, A_n$ herrühren. Um die wahrscheinlichste Ursache zu finden, muss man also $P(A_k|B)$ abschätzen.
- (b) Aus einem verrauschten Bild will man das wahrscheinlichste unverrauschte Bild rekonstruieren. (vgl. Beispiel 64.7)
- (c) In der Computertomographie schickt man Röntgenstrahlen in verschiedenen Richtungen durch den Patienten und misst die durchgedrungene Intensität. Aus diesen Auswirkungen versucht man, Rückschlüsse auf die Ursache (Gewebe, Knochen, Tumor,...) zu ziehen.