§64: BEDINGTE WAHRSCHEINLICHKEITEN

64.1. Motivation

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

64.2. Beispiel

Aufteilung der 1500 Angehörigen eines Betriebs nach Geschlecht und Rauchergewohnheiten:

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

Ω: Menge der Betriebsangehörigen

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

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

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

$$P(A) = \frac{|A|}{|\Omega|} = \frac{800}{1500} = \frac{8}{15} \quad \text{Raucheranteil}$$

$$P(B) = \frac{|B|}{|\Omega|} = \frac{900}{1500} = \frac{3}{5} \quad \text{Frauenanteil}$$

$$P(A \cap B) = \frac{|A \cap B|}{|\Omega|} = \frac{600}{1500} = \frac{2}{5}$$ Anteil der rauchenden Frauen

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

$$P(A|B) = \frac{|A\cap B|}{|B|} = \frac{\frac{|A\cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{P(A\cap B)}{P(B)} \qquad (*)$$

$$= \frac{2/5}{3/5} = \frac{2}{3}.$$

Man nennt $P(A|B)$ die <u>bedingte Wahrscheinlichkeit</u> von A unter der Bedingung (Hypothese) B.

Es gilt stets (vgl. $(*)$):

$$\boxed{P(A\cap B) = P(A|B)\cdot P(B)}$$

Zwei Ereignisse A, B heißen <u>unabhängig</u>, wenn

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

64.3. <u>Verallgemeinerung</u>

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

Dann gilt:

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

<u>Beweis:</u> Die rechte Seite läßt sich schreiben als

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

$$= P(A_1\cap\dots\cap A_n) \qquad \square.$$

64.4. Beispiel

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

A_1: irgendein Ergebnis für 1. Würfel

A_2: ein vom 1. Ergebnis versch. Ergebnis für Würfel 2

⋮

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

$$P(A_1 \cap \ldots \cap A_6) = P(A_1) \cdot P(A_2 | A_1) \cdot \ldots \cdot P(A_6 | A_1 \cap \ldots \cap A_5)$$

$$= 1 \cdot \frac{5}{6} \cdot \ldots \cdot \frac{1}{6} = \frac{6!}{6^6} \approx 0{,}015$$

64.5. Beispiel

Ein Krebstest ist mit 96%iger Sicherheit positiv, falls der Patient Krebs hat, mit 94%iger Sicherheit 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 Ws., 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.

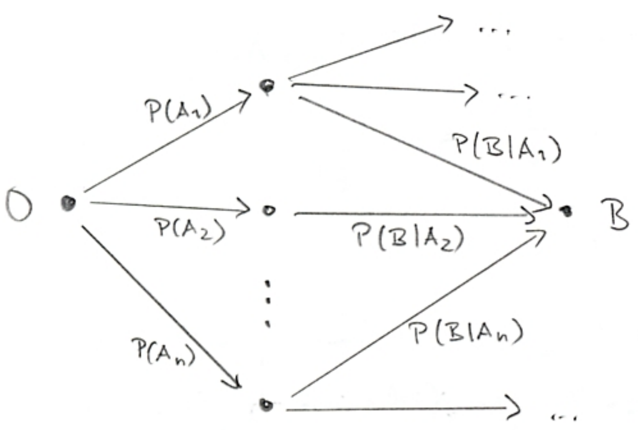
Fazit: Um eine seltene Krankheit zuverlässig zu erkennen, darf ein Test nur sehr wenige "false positives" haben.

64.6 Satz (Satz von der totalen Wahrscheinlichkeit)

Sei $\Omega = \dot{\bigcup}_{i=1}^{n} A_i$ eine Partition 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:

$$\boxed{P(B) = \sum_{i=1}^{n} P(A_i)\, P(B \mid A_i)}$$

Veranschaulichung



Fahrer startet bei O und fährt mit Ws. $P(A_1), \dots, P(A_n)$ zu $A_1, \dots, A_n$. Die Ws. von dort nach B zu fahren, beträgt $P(B \mid A_1), \dots, P(B \mid A_n)$.
Die Gesamtws., dass der Fahrer nach B gelangt, ist die Summe der Ws. alle Pfade von O nach B:

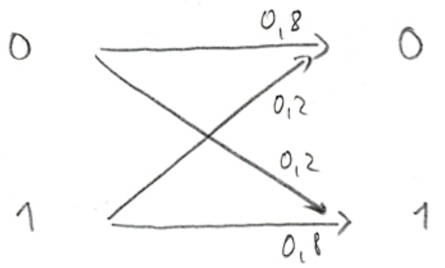
$$P(B) = \sum_{i=1}^{n} P(A_i)\, P(B \mid A_i)$$

Bem.: Im Beispiel 64.5 haben wir im Nenner bereits den Satz von der totalen Ws. 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 Ws. für 0000000 sei 0,1.

Es wird 0100110 empfangen. Wie groß ist die Ws., dass 0000000 gesendet wurde?



$$P(0000000 \mid 0100110)$$

$$= \frac{0{,}1 \cdot 0{,}2^3 \cdot 0{,}8^4}{0{,}1 \cdot 0{,}2^3 \cdot 0{,}8^4 + 0{,}9 \cdot 0{,}2^4 \cdot 0{,}8^3} \approx 0{,}308$$

Man wird den Block also als 1 lesen, obwohl die Mehrzahl der Bits Nullen sind!

Eng verwandt mit dem Satz von der totalen Ws. ist

64.8. Satz (Formel von Bayes)

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

$$\boxed{P(A_k \mid B) = \frac{P(A_k)\, P(B \mid A_k)}{\sum_{i=1}^{n} P(A_i)\, P(B \mid A_i)}}$$

Beweis: Folgt aus $P(A_k|B) = \dfrac{P(A_k \cap B)}{P(B)}$, indem man

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

$$P(A_k \cap B) = P(A_k)\, P(B|A_k) \qquad (64.2)$$

einsetzt. $\square$

Bem.: In 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 eine Wirkung misst und nach der wahrscheinlichsten Ursache fragt (inverses Problem).

64.5. Anwendungsbeispiele

a) Ein Arzt beobachtet bei einem Patienten ein Symptom B. Es kann von n verschiedenen Ursachen A_k, $k=1,\dots,n$ 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. auch 64.7).

c) In der Computertomographie schickt man Röntgenstrahlung 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.