

# Stochastik

## Grundbegriffe Wahrscheinlichkeit

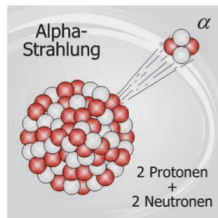
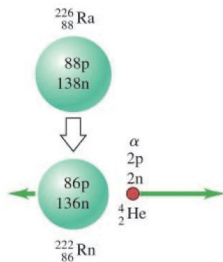
Mirko Birbaumer

Hochschule Luzern Technik & Architektur

- 1 Radioaktiver Alpha Zerfall
- 2 Wahrscheinlichkeitsmodell
- 3 Bedingte Wahrscheinlichkeit

# Radioaktiver Zerfall: Alpha Zerfall

- Nahezu 90% der bekannten Nuklide sind radioaktiv und zerfallen nach einer gewissen Zeit in neue Nuklide
- Beim Zerfall emittieren sie Alpha ( $\alpha$ ) oder Beta ( $\beta$ ) Teilchen oder Gamma ( $\gamma$ ) Strahlung

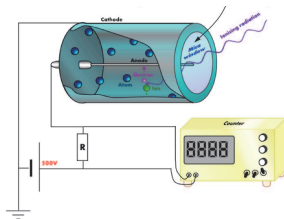


$^{226}_{88}\text{Ra}$  zerfällt in  $^{222}_{86}\text{Rn}$  unter Emission eines Alpha Teilchens

# Messgerät für radioaktiven Zerfall: Geigerzähler

- Geigerzähler: mit Argon (200hPa) gefülltes Stahlrohr.
- Radioaktive Strahlung ionisiert Atome des Edelgases.

Sehr dünnes Fenster aus Glimmer oder Mylar



- Anzahl Zerfälle in 10-Sekundenintervall zufällig: es können 16 oder 0 oder 120 Zerfälle gemessen werden

# Anzahl Zerfälle von americium 241 in 10 s

Anzahl Zerfälle	Beobachtet in Anzahl Experimente
0-2	18
3	28
4	56
5	105
6	126
7	146
8	164
9	161
10	123
11	101
12	74
13	53
14	23
15	15
16	9
17+	5

- Experiment wurde 1207 Mal wiederholt, jedes Mal wurde die Anzahl Zerfälle in 10 Sekunden gemessen
- In 28 der 1207 Experimente wurden 3 Alphateilchen gemessen

# Wahrscheinlichkeitsmodell

- **Zufallsexperimente:** Ausgang ist nicht exakt vorhersagbar, wie
  - Würfelwurf
  - Münzenwurf
  - Anzahl Zerfälle eines Alphastrahlers
- Ein **Wahrscheinlichkeitsmodell** beschreibt, welche Ereignisse in einem solchen Experiment möglich sind und welche Wahrscheinlichkeiten die verschiedenen Ergebnisse haben
- Beispiel: Würfel werfen
  - mögliche Ergebnisse: 1, 2, 3, 4, 5, 6
  - W'keit eine dieser Zahlen zu werfen, ist  $\frac{1}{6}$ , sofern Würfel fair

# Wahrscheinlichkeitsmodell

- Ein Wahrscheinlichkeitsmodell hat die folgenden Komponenten:
  - **Grundraum**  $\Omega$ : enthält alle möglichen **Elementarereignisse**  $\omega$
  - **Ereignis** „A“: Teilmenge des Grundraums
  - **Wahrscheinlichkeiten**  $P$
- **Elementarereignisse** sind mögliche Ergebnisse oder Ausgänge eines Experiments
- Elementarereignisse bilden den Grundraum:

$$\Omega = \underbrace{\{\text{mögliche Elementarereignisse } \omega\}}_{\text{mögliche Ausgänge/Resultate}}$$

# Beispiele Wahrscheinlichkeitsmodell

- Beim Würfelwurf ist der Grundraum

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

- Das Element  $\omega = 2$  ist ein Elementarereignis  $\rightarrow$  beim Würfeln wurde die Zahl 2 geworfen

- Alphazerfall: Grundraum

$$\Omega = \{0, 1, 2, 3, 4, \dots\}$$

da beliebig viele Zerfälle in 10-Sekundenintervall möglich

- Elementarereignis  $\omega = 6$   $\rightarrow$  in 10 Sek. 6 Zerfälle gemessen



# Beispiele Wahrscheinlichkeitsmodell

- 2-maliges Werfen einer Münze : Bezeichnungen  $K$  : „Kopf“ und  $Z$  : „Zahl“
  - Bezeichnungen  $K$  : „Kopf“ und  $Z$  : „Zahl“
- Alle möglichen Ergebnisse des Experimentes (Grundraum)

$$\Omega = \{KK, KZ, ZK, ZZ\}$$

- Elementarereignis ist z.B.  $\omega = KZ$

# Ereignis

- Unter einem **Ereignis**  $A$  versteht man eine Teilmenge von  $\Omega$ :

$$A \subset \Omega$$

„Ein Ereignis  $A$  tritt ein“ bedeutet, dass das Ergebnis  $\omega$  des Experiments zu  $A$  gehört.

- Beispiel: 2-maliges Werfen einer Münze
  - Ereignis  $A$ , wo genau einmal  $K$  geworfen wird
  - Ereignis besteht aus den Elementarereignissen  $KZ$  und  $ZK$
  - Das Ereignis  $A$  ist dann die Menge

$$A = \{KZ, ZK\}$$

- Würfeln  $ZZ$ , so trifft das Ereignis  $A$  *nicht* ein

# Beispiel: Würfeln

- Ereignis  $A$ : „eine ungerade Zahl würfeln“

- Dann ist

$$A = \{1, 3, 5\}$$

- Das Ereignis  $A$  tritt ein, wenn z.B. die Zahl 5 gewürfelt wird

- Ereignis  $B$ : eine Zahl kleiner als 7 würfeln

- Das ist natürlich immer der Fall und somit ist

$$B = \Omega$$

- $B$  heisst *sicheres* Ereignis

- $C$  das Ereignis „die Zahl sieben würfeln“

- Dies ist unmöglich

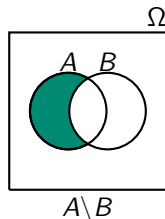
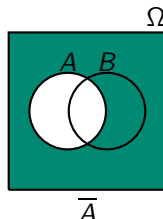
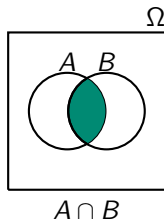
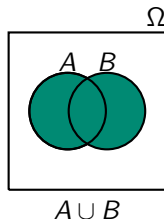
$$C = \emptyset$$

- $\emptyset$  ist die leere Menge, die kein Element enthält: *unmögliches* Ereignis

# Mengenlehre

- Die Operationen der Mengenlehre (Vereinigung, Durchschnitt, Komplement) werden für Ereignisse verwendet:

Name	Symbol	Bedeutung
Vereinigung	$A \cup B$	A <b>oder</b> B, nicht-exklusives "oder"
Durchschnitt	$A \cap B$	A <b>und</b> B
Komplement	$\bar{A}$	<b>nicht</b> A
Differenz	$A \setminus B = A \cap \bar{B}$	A <b>ohne</b> B



# Beispiel

- Ereignisse
  - $A$ : „morgen scheint die Sonne“
  - $B$ : „morgen regnet es“
- Dann bedeuten folgende Ereignisse:
  - $A \cup B$ :  
„morgen scheint die Sonne *oder* morgen regnet es“ (und dies kann auch bedeuten: „morgen scheint die Sonne *und* morgen regnet es“)
  - $A \cap B$ :  
„morgen scheint die Sonne *und* morgen regnet es“
  - $\bar{A}$ :  
„morgen scheint die Sonne nicht“

# Axiome der Wahrscheinlichkeitsrechnung

- Eigenschaften von Wahrscheinlichkeiten

## Kolmogorov Axiome der Wahrscheinlichkeitsrechnung

Jedem Ereignis  $A$  wird eine **W'keit**  $P(A)$  zugeordnet, mit:

$$A1: P(A) \geq 0$$

$$A2: P(\Omega) = 1$$

$$A3: P(A \cup B) = P(A) + P(B) \quad \text{falls } A \cap B = \emptyset$$

- Bezeichnung  $P(A)$ : W'keit, dass das Ereignis  $A$  eintritt
- Ereignis  $A$ : „ungerade Zahl würfeln“ (bei fairem Würfel)

$$P(A) = \frac{1}{2}$$

- Der Buchstabe  $P$  steht für das englische *probability*
- $P(\Omega) = 1 \rightarrow$  W'keiten eines Ereignisses zwischen 0 und 1

# Beispiel: Münzwurf

- Wurf zweier Münzen: plausibel, dass alle 4 Elemente von

$$\Omega = \{KK, KZ, ZK, ZZ\}$$

gleich wahrscheinlich sind

- Wegen  $P(\Omega) = 1$  müssen sich die Wahrscheinlichkeiten zu eins aufaddieren:

$$P(KK) = P(KZ) = P(ZK) = P(ZZ) = \frac{1}{4}$$

# Rechenregeln

## Rechenregeln für W'keiten von Ereignissen

Sind  $A$ ,  $B$  und  $A_1, \dots, A_n$  Ereignisse, dann gilt

$$P(\overline{A}) = 1 - P(A) \quad \text{für jedes } A$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \text{für beliebige } A \text{ und } B$$

$$P(A_1 \cup \dots \cup A_n) \leq P(A_1) + \dots + P(A_n) \quad \text{für beliebige } A_1, \dots, A_n$$

$$P(B) \leq P(A) \quad \text{für beliebige } A \text{ und } B \text{ mit } B \subseteq A$$

$$P(A \setminus B) = P(A) - P(B) \quad \text{für beliebige } A \text{ und } B \text{ mit } B \subseteq A$$

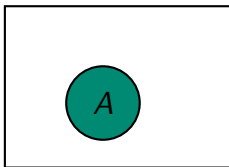


# Venn Diagramme: Wahrscheinlichkeit als Fläche

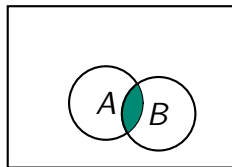
$$P(\Omega) = 1$$



$$P(A)$$

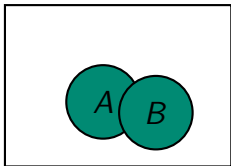


$$P(A \cap B)$$

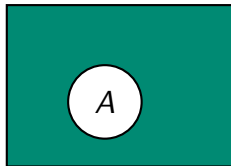


$$P(A \cup B)$$

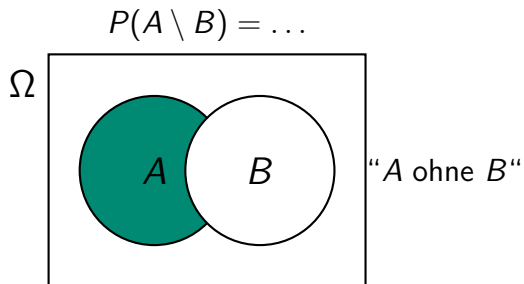
$$= P(A) + P(B) - P(A \cap B)$$



$$P(\bar{A}) = 1 - P(A)$$



# Knobelaufgabe



1.  $P(A) - P(B)$

2.  $P(A) + P(B)$

3.  $P(A) - P(A \cap B)$

4.  $P(A) + P(B) - P(A \cap B)$

# Diskrete Wahrscheinlichkeitsmodelle

- *Diskrete* Wahrscheinlichkeitsmodelle : Grundraum ist endlich oder unendlich und diskret

- Begriff „diskret“ z.B. Menge

$$\Omega = \{0, 1, \dots, 10\}$$

die *endlich und deshalb diskret* ist

- *unendliche*, aber trotzdem *diskrete* Menge

$$\Omega = \mathbb{N}_0 = \{0, 1, 2, \dots\}$$

- Die Menge  $\Omega = \mathbb{R}$  (die Menge aller Zahlen, die man als Dezimalbrüche schreiben kann) ist *nicht* diskret

- $\Omega = \mathbb{R}$  wird später für Messdaten eine sehr wichtige Rolle spielen

# Berechnung von W'keiten für diskrete Modelle

Im diskreten Fall ist die Wahrscheinlichkeit eines Ereignisses

$$A = \{\omega_1, \omega_2, \dots, \omega_n\}$$

durch die Wahrscheinlichkeiten der zugehörigen Elementarereignisse  $P(\omega)$  festgelegt:

$$P(A) = P(\omega_1) + P(\omega_2) + \dots + P(\omega_n) = \sum_{\omega_i \in A} P(\omega_i)$$

# Beispiel: 2-maliges Münzenwerfen

- Es ist  $A$  : „genau einmal  $K$  werfen“, also

$$A = \{KZ, ZK\}$$

- W'keit  $P(A)$ , dass das Ereignis  $A$  eintritt

$$P(A) = P(KZ) + P(ZK) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

- Ereignis  $B$ : „mindestens einmal Kopf werfen“

- Wegen  $B = \{KZ, ZK, KK\}$  gilt

$$P(B) = P(KZ) + P(ZK) + P(KK) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

# Beispiel: 2-maliges Münzenwerfen

- $P(B)$  mit  $B = \{KZ, ZK, ZZ\}$  : einfacher mit der sogenannten *Gegenw'keit* berechnen
- Das Komplement  $\overline{B}$  von  $B$  ist

$$\overline{B} = \{ZZ\}$$

- Aus erster Rechenregel (siehe oben):

$$P(B) = 1 - P(\overline{B}) = 1 - \frac{1}{4} = \frac{3}{4}$$

# Modell von Laplace

- Annahme: jedes Elementarereignis hat die gleiche W'keit (**Modell von Laplace**)
- Ereignis  $E = \{\omega_1, \omega_2, \dots, \omega_g\}$ ; Grundraum  $m$  Elemente
- Wahrscheinlichkeiten addieren sich zu 1, also

$$P(\omega_k) = \frac{1}{|\Omega|} = \frac{1}{m}$$

Für ein Ereignis  $E$  im Laplace Modell gilt also

$$P(E) = \sum_{k: \omega_k \in E} P(\{\omega_k\}) = \frac{g}{m}$$

- Man teilt die Anzahl der „**günstigen**“ Elementarereignisse durch die Anzahl der „**möglichen**“ Elementarereignisse

# Beispiel: Laplace Modell

- Es werden zwei Würfel geworfen
- Wie gross ist die W'keit, dass die Augensumme 7 ergibt?
- Elementarereignis beschreibt die Augenzahlen auf beiden Würfeln
- Dieses Ergebnis in der Form  $(1, 4)$  schreiben, wenn der eine Würfel eine 1 und der andere eine 4 zeigt
- Es sind insgesamt 36 Elementarereignisse möglich:

$$\Omega = \{(1, 1), (1, 2), \dots, (6, 5), (6, 6)\}$$



# Beispiel: Laplace Modell

- Ereignis  $E$ : die Augensumme 7 wird gewürfelt
- Es gibt davon 6 Elementarereignisse:

$$E = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$$

- Da alle Elementarereignisse gleich wahrscheinlich , ist W'keit für das Ereignis  $E$ :

$$P(E) = \frac{|E|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}$$

# Stochastische Unabhängigkeit

- Falls  $P(A)$  und  $P(B)$  bekannt, lässt sich daraus  $P(A \cap B)$  i.A. nicht berechnen
- Spezialfall liegt vor, wenn die Berechnung von  $P(A \cap B)$  aus  $P(A)$  und  $P(B)$  mit Hilfe folgender Produktformel möglich ist

Die Ereignisse  $A$  und  $B$  heissen **stochastisch unabhängig**, wenn

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

# Beispiel: 2-maliges Werfen einer Münze

- Es sei  $A$ : „ $K$  im 1. Wurf“ und  $B$ : „ $K$  im 2. Wurf“
- Dann gilt

$$P(A) = P(KK) + P(KZ) = \frac{1}{2}$$

und analog

$$P(B) = P(KK) + P(ZK) = \frac{1}{2}$$

- Mit  $A \cap B$  wird das Ereignis „ $K$  im 1. *und* im 2. Wurf“ beschrieben

# Beispiel: 2-maliges Werfen einer Münze

- Es gilt dann für die W'keit dieses Ereignisses

$$P(A \cap B) = P(KK) = \frac{1}{4}$$

- Da

$$P(A)P(B) = \frac{1}{4}$$

gilt zudem

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

- D.h.: Ereignisse  $A$  und  $B$  sind (stochastisch) unabhängig

# Stochastische Unabhängigkeit: Umkehrung

- Viel wichtiger ist jedoch der umgekehrte Schluss
- Wenn zwischen den Ereignissen  $A$  und  $B$  *kein kausaler Zusammenhang* besteht (d.h. es gibt keine gemeinsamen Ursachen oder Ausschlüsse), dann nehmen wir stochastische Unabhängigkeit an und damit *muss* gelten:

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

- In diesem Fall kann also  $P(A \cap B)$  aus  $P(A)$  und  $P(B)$  berechnet werden

# Beispiel: Würfel und Karte

- Werfen zuerst einen Würfel und ziehen nachher eine Karte aus einem vollständigen Stapel Jasskarten
- Wie gross ist die W'keit zuerst eine 3 zu werfen und nachher einen König zu ziehen?
- Bezeichnung dieses Elementarereignis mit  $3K$
- Die W'keit eine 3 zu werfen ist  $\frac{1}{6}$ , also

$$P(3) = \frac{1}{6}$$

# Beispiel: Würfel und Karte

- Die Wahrscheinlichkeit einen König zu ziehen ist  $\frac{1}{9}$ :

$$P(K) = \frac{1}{9}$$

- Würfelwurf hat auf das Ziehen einer Jasskarte keinen Einfluss

$$P(3K) = P(3) \cdot P(K) = \frac{1}{6} \cdot \frac{1}{9} = \frac{1}{54}$$

# Beispiel: Münzenwerfen

- Sei wieder  $A$ : „ $K$  im 1. Wurf“ und  $B$ : „ $K$  im 2. Wurf“.
- Plausible Annahme:, gibt keinen kausalen Zusammenhang zwischen dem Ergebnis des ersten und des zweiten Wurfs gibt
- Mit anderen Worten hat das Ergebnis des 1. Wurfs (Ereignis  $A$ ) keinen Einfluss auf das Ergebnis des 2. Wurfs (Ereignis  $B$ )
- Die Ereignisse  $A$  und  $B$  sind also unabhängig
- Deshalb kann man  $P(A \cap B)$  wie folgt berechnen:

$$P(A \cap B) = P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$



# Beispiel: Abhängigkeit beim Würfelwurf

- Zwei Würfel werden geworfen
- Sei  $E$  das Ereignis, dass der erste Würfel die Augenzahl 4 zeigt und  $F$  das Ereignis, dass die Augensumme 6 ist
- Sind die Ereignisse  $E$  und  $F$  unabhängig voneinander?
- Um dies festzustellen  $\rightarrow$  gilt Produktformel oder nicht?

$$P(E \cap F) \stackrel{?}{=} P(E) \cdot P(F)$$

- Ereignis  $E \cap F$  besteht nur aus Elementarereignis  $(4, 2)$

- Damit gilt für dessen Wahrscheinlichkeit

$$P(E \cap F) = P((4, 2)) = \frac{1}{36}$$

- Es gilt natürlich  $P(E) = \frac{1}{6}$
- Das Ereignis  $F$  besteht aus den Elementen

$$F = \{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}$$

- Es gilt

$$P(F) = \frac{5}{36}$$

- Produktformel überprüfen:

$$P(E) \cdot P(F) = \frac{1}{6} \cdot \frac{5}{36} = \frac{5}{216} \neq \frac{1}{36} = P(E \cap F)$$

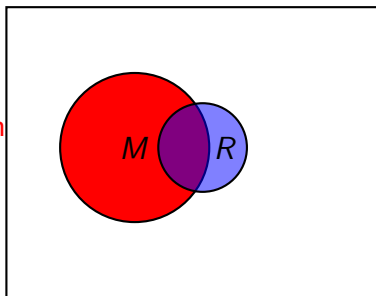
- Deshalb sind  $E$  und  $F$  nicht unabhängig.

- Abhängigkeit der beiden Ereignisse  $E$  und  $F$  auch ohne Rechnung leicht einsehen
- Werfen den ersten Würfel und erhalten eine 6
- Keine Möglichkeit mehr in zwei Würfeln die Augensumme 6 zu erreichen (die W'keit ist 0)
- Zuerst irgendeine andere Zahl werfen, so Möglichkeit (W'keit ungleich 0) auf die Augensumme 6 zu kommen
- W'keit die Augensumme 6 zu erzielen hängt von der Augenzahl des ersten Wurfes ab
- Also können  $E$  und  $F$  nicht unabhängig voneinander sein

# Bedingte Wahrscheinlichkeit

$\Omega$ : Studenten dieser VL

M: Männlich  
 $P(M)$



R: Raucher  
 $P(R)$

$$P(R|M)$$

Wa. für Raucher

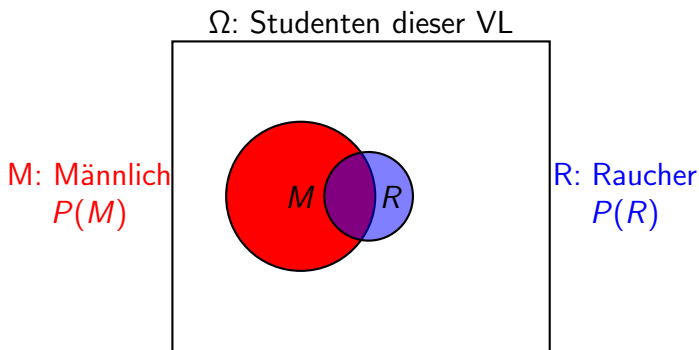
wenn ein Mann gewählt wurde

$$P(M|R)$$

Wa. für Mann

wenn ein Raucher gewählt wurde

# Bedingte Wahrscheinlichkeit



Welche Aussagen sind korrekt?

1.  $P(M|R) = P(R|M)$
2.  $P(M|R) > P(R|M)$
3.  $P(M|R) < P(R|M)$

# Bedingte Wahrscheinlichkeit

- Die **bedingte W'keit** ist die W'keit, dass das Ereignis  $A$  eintritt, wenn wir schon wissen, dass  $B$  eingetreten ist
- Bezeichnung:

$$P(A|B)$$

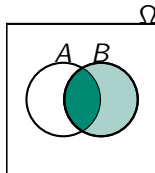
- Der Längsstrich wird als „unter der Bedingung“ gelesen
- Die bedingte Wahrscheinlichkeit  $P(A|B)$  wird definiert durch

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

- Interpretation:  $P(A|B)$  ist die W'keit für das Ereignis  $A$ , wenn wir wissen, dass das Ereignis  $B$  schon eingetroffen ist

# Verdeutlichung der Formel mit Flächen

- Graphisch



- Es ist  $|\Omega| = 1$
- $P(A|B)$  Flächeninhalt der dunkel gefärbten Flächen
- $P(B)$  Flächeninhalt der gesamten gefärbten Fläche  $B$
- Der Anteil der dunkelgefärbten Fläche zur gefärbten Fläche ist dann

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

# Beispiel: Medizinischer Test

- Ein medizinischer Test soll für eine Krankheit feststellen, ob eine Person an dieser Krankheit erkrankt ist oder nicht
- Natürlich ist dieser Test nicht ganz genau:
  - Zeigt manchmal die Krankheit an, obwohl die Person gesund ist
  - Zeigt die Krankheit nicht an, obwohl die Person krank ist
- Sie gehen zum Arzt und machen diesen Test auf eine tödliche Krankheit: Test ist positiv
- **Frage:** Wie gross ist die W'keit, dass Sie wirklich krank sind?



- Bezeichnungen:
  - $D$ : Krankheit ist vorhanden;  $\overline{D}$ : Krankheit ist nicht vorhanden
  - $+$ : Test zeigt Krankheit an;  $-$ : Test zeigt Krankheit nicht an
- W'keiten in Tabelle sind durch Versuche bekannt

	$D$	$\overline{D}$
$+$	0.009	0.099
$-$	0.001	0.891

- Z.B.: W'keit, dass die Krankheit vorhanden ist *und* der Test positiv ausfällt

$$P(D \cap +) = 0.009$$

- Grund, warum diese W'keit recht klein ist: nur kleiner Prozentsatz der Bevölkerung hat Krankheit

- Verschiedene bedingte W'keiten:
  - $P(+|D)$ : W'keit, dass ein Kranker auch wirklich positiv getestet wird
  - $P(-|\overline{D})$ : W'keit, dass ein Gesunder richtigerweise negativ getestet wird
  - $P(D|+)$ : W'keit, dass positiv Getesteter auch wirklich krank ist
- Berechnen zuerst die W'keit  $P(+|D)$ :

$$P(+|D) = \frac{P(+ \cap D)}{P(D)} = \frac{0.009}{0.009 + 0.001} = 0.9$$

- Dabei haben wir für  $P(D)$  folgende Tatsache benützt

$$P(D) = P(D \cap +) + P(D \cap -) = 0.009 + 0.001$$

- $P(D)$  : Summe der Einträge in der Tabelle in der Spalte unter  $D$
- Die Kranken sind entweder positiv oder negativ getestet

- Bedingte W'keit  $P(-|\overline{D})$ :

$$P(-|\overline{D}) = \frac{P(- \cap \overline{D})}{P(\overline{D})} = \frac{0.891}{0.891 + 0.099} = 0.9$$

- Anscheinend ist dieser Test recht genau
- Kranke Personen werden zu 90 % als positiv eingestuft, und gesunde Personen werden zu 90 % als negativ eingestuft
- Angenommen, Sie gehen zu einem Test und dieser wird als positiv eingestuft
- Wie gross ist die W'keit, dass Sie die Krankheit wirklich haben?

- Die meisten Leute würden 0.9 antworten
- Müssen Sie sich also grosse Sorgen machen und das Testament schreiben oder einer Sterbehilfeorganisation beitreten?
- Die *richtige* Antwort ist die bedingte W'keit  $P(D|+)$ :

$$P(D|+) = \frac{P(+ \cap D)}{P(+)} = \frac{0.009}{0.009 + 0.099} = 0.08$$

- Was bedeutet nun dieses Resultat?
- Die bedingte W'keit  $P(D|+)$  ist die W'keit, dass man bei einem positiven Test auch wirklich krank ist
- Diese beträgt aber nur 8 %
- Man hat bei einem positiven Test also nur zu 8 % auch wirklich die Krankheit
- Ein positiver Test sagt hier also sehr wenig darüber aus, ob man die Krankheit hat oder nicht

# Bayes' Theorem

## Bayes' Theorem

Das **Bayes Theorem** liefert einen oft nützlichen Zusammenhang zwischen  $P(A|B)$  und  $P(B|A)$ :

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

**Beispiel:** Das Bayes Theorem liefert die gleiche Lösung wie unsere obige Rechnung:

$$P(D|+) = \frac{P(+|D)P(D)}{P(+)} = \frac{0.9 \cdot (0.009 + 0.001)}{0.009 + 0.099} = \frac{0.009}{0.009 + 0.099} = 0.08$$

# Gesetz der totalen Wahrscheinlichkeit

- Ein weiterer nützlicher Begriff ist die *totale Wahrscheinlichkeit*
- Menge  $A$  in Mengen  $A_1, \dots, A_k$  unterteilt, die miteinander keine Schnittmenge haben und zusammen (Vereinigung) die ganze Menge  $A$  bilden
- Eine solche Aufteilung nennen wir eine **Partitionierung**
- Für den Würfelwurf ist folgende Partitionierung möglich:

$$A_1 = \{1\}, \quad A_2 = \{2, 4\}, \quad A_3 = \{3, 5, 6\}$$

- Es gilt also

$$A_1 \cap A_2 = \emptyset; \quad A_1 \cap A_3 = \emptyset; \quad A_2 \cap A_3 = \emptyset$$

und

$$A_1 \cup A_2 \cup A_3 = A$$

# Gesetz der totalen Wahrscheinlichkeit

## Gesetz der totalen Wahrscheinlichkeit

Für die Partitionierung  $A_1, \dots, A_k$  und für jedes beliebige Ereignis  $B$  gilt:

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \dots + P(B|A_k)P(A_k) \\ &= \sum_{i=1}^k P(B|A_i)P(A_i) \end{aligned}$$

# Beispiel: Spam-Mail

- Teilen Emails in drei Kategorien ein:

$A_1$  : „spam“,    $A_2$  : „niedrige Priorität“,    $A_3$  : „hohe Priorität“

- Aus früheren Beobachtungen bekannt:

$$P(A_1) = 0.7, \quad P(A_2) = 0.2, \quad \text{und} \quad P(A_3) = 0.1$$

- Es gilt

$$P(A_1) + P(A_2) + P(A_3) = 1$$

wie es bei einer Partitionierung auch sein sollte

- Ereignis  $B$ : das Wort „free“ taucht in der Email auf - kommt sehr oft in Spam-Mails vor



# Beispiel: Spam-Mail

- Von früheren Beobachtungen bekannt

$$P(B|A_1) = 0.9, \quad P(B|A_2) = 0.01, \quad \text{und} \quad P(B|A_3) = 0.01$$

- Hier ergibt die Summe nicht 1
- Dies sind die W'keiten, mit der das Wort „free“ in den drei Mailkategorien vorkommt
- Angenommen, es kommt eine Email an, die das Wort „free“ enthält: Wie gross ist die W'keit, dass es sich um Spam handelt?

- Das Bayes Theorem zusammen mit dem Gesetz der totalen Wahrscheinlichkeit liefert die Lösung:

$$\begin{aligned}
 P(A_1|B) &= \frac{P(B|A_1)P(A_1)}{P(B)} \\
 &= \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)} \\
 &= \frac{0.9 \cdot 0.7}{(0.9 \cdot 0.7) + (0.01 \cdot 0.2) + (0.01 \cdot 0.1)} \\
 &= 0.995
 \end{aligned}$$

- Viele Spamfilter basieren tatsächlich auf diesem Prinzip
- Die Mails werden nach Worten wie „free“, „credit“, etc. durchsucht, die häufig in Spam-Mails vorkommen, in den anderen aber eher nicht