Statistik

Sommersemester 2017

Prof. Dr. Reimar Hofmann

Themengebiete

Stochastik

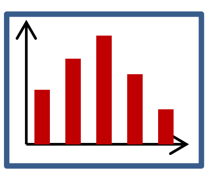
umfasst

beschreibende Statistik

Wahrscheinlichkeits-rechnung

schließende Statistik

Beobachtungsdaten aussagekräftig zusammenfassen.



Systeme mit Unsicherheit quantitativ beschreiben, Vorhersagen treffen

Rückschlüsse aus Beobachtungsdaten auf das zugrundeliegende System

Beschreibende Statistik

Definition

Beschreibende oder ***deskriptive*** Statistik:
Beschreibung und Darstellung von ***bekannten*** Daten, Zusammenfassen in Kennzahlen und Diagrammen.

Beispiele:

- Business Intelligence
- Arbeitslosenstatistik
- Server-Verfügbarkeitsstatistik
- Generell Visualisierung von Daten

wichtig

Wahrscheinlichkeitsrechnung

Definition
Wahrscheinlichkeitsrechnung ist ein Fachgebiet, das sich damit befasst, wie man Systeme, deren Verhalten **nicht exakt vorhersagbar** ist, trotzdem ***quantitativ beschreiben*** kann.

Meist bildet diese Beschreibung nicht die volle Komplexität des realen Systems ab, sondern ein **vereinfachtes Modell**.

Behandelt wird:

- Wie kann man ein solches Modell beschreiben?
- Wie kann man aus einem bekannten Modell Vorhersagen über das Verhalten des Systems ableiten?

NICHT betrachtet wird:

- Wie kann man das Modell zu einem vorliegenden System ermitteln? (-> schließende Statistik)

wichtig

Beispiel zur Wahrscheinlichkeitsrechnung

Untersucht werden Dinge wie:

Wenn bekannt ist, dass Max gegen ein bestimmtes Schachprogramm mit 70% Wahrscheinlichkeit gewinnt, wie wahrscheinlich ist es dann, dass er von 10 Spielen mehr als 8 gewinnt?

Nicht untersucht wird:

Wie findet man heraus, mit welcher Wahrscheinlichkeit Max gewinnt? Wie viele Testspiele sind nötig, um dies mit einer bestimmten Genauigkeit schätzen zu können?

(-> diese Fragen gehören zu *schließenden* Statistik)

Schließende Statistik

Definition

Schließende Statistik oder ***induktive*** S.:
Aus Daten Rückschlüsse auf das zugrundeliegende System (bzw.die Grundgesamtheit) ziehen.

Beispiele:

- Maschinelles Lernen (z.B. Bilderkennung, Robotik)
- Data Mining
- Lastprognose (Web-Server)
- Absatzprognose (Handel)
- Wahlanalyse

wichtig

Anwendungen in der Informatik

Aussagen über unsichere Sachverhalte treffen

- **Big Data / Predictive Analytics / Data Mining**:
  Z.B. Prognose von Kaufverhalten
- **Bild-/Sprachverarbeitung:**
  Klassifikation von Objekten/Lautfolgen, die teilweise zu *undeutlich sind, um für sich alleine eindeutig interpretierbar* zu sein.
- **Roboter** *in teilweise unbekannter Umgebung* **navigieren**.
- **Kompression**: Daten so codieren, dass häufige Bestandteile kurze, und selten vorkommende lange Codes haben. (Unsicher: Welche Daten werden zu codieren sein)
- **Kryptoanalysis**: Verfahren zum entschlüsseln *unbekannter Nachrichten* bei *unbekanntem Schlüssel* entwickeln.
- **Average-Case"-Komplexität** eines Algorithmus,
  z.B. optimale Dimensionierung einer Hash-Tabelle bestimmen.
  Unbekannt: Welcher Input wird zu verarbeiten sein.
- **Zuverlässigkeitsanalysen** (z.B. für Hochverfügbarkeitssysteme):
  Ausfallwahrscheinlichkeiten abschätzen.

Fahrplan der Statistik-Vorlesung

1. Wahrscheinlichkeitsrechnung von Ereignissen

- *Zufall, Wahrscheinlichkeit, Zufallsereignis*
- Wie bestimmt man Wahrscheinlichkeiten
- *bedingte* Wahrscheinlichkeit, stochastische *Unabhängigkeit*

2. Beschreibende Statistik von ***quantitativen Merkmalen*** und Wahrscheinlichkeitsrechnung von ***Zufallsvariablen***

– *Mittelwert/Erwartungswert, Quantile, Varianz, Standardabweichung*
– Häufigkeitsverteilung / *Wahrscheinlichkeitsverteilung*
– Wichtige Verteilungen (*Binomial-, Poisson-, Exponential-, Normalverteilung*)
– Multivariate Statistik: *Kovarianz* und *Korrelation*
– lineare Regression

3. Schließende Statistik

– Lernen von Modellparametern aus Beispielen (z.B. Maximum Likelihood Ansatz)
– Anzahl erforderlicher Beispiele abschätzen

Lernziel jeweils -*Wissen, Verständnis* und ***Anwendung***
- ***Notation*** lesen können.
(Aufgabentyp: Unbekannter Sachverhalt wird erklärt -> anwenden.)

1 Zufalls**ereignisse**

- **1.1 Grundbegriffe**
  - Zufall
  - relative/absolute Häufigkeiten
  - Wahrscheinlichkeiten

- 1.2 Wie bestimmt man Wahrscheinlichkeiten?
  - Durch Zählen: Laplace-Experimente
  - Aus bereits bekannten Wahrscheinlichkeiten: Rechengesetze für Wahrscheinlichkeiten

- 1.3 Bedingte Wahrscheinlichkeiten

1.1 Grundbegriffe

1.1 Kann man Systeme, deren Verhalten nicht sicher vorhersehbar ist, trotzdem ***quantitativ*** beschreiben?

Und wenn ja, was bringt das?

Einführende Beispiel-Problemstellung

Versicherung A bietet einem Suchmaschinenhersteller **1€ *pro Klick*** dafür, dass ihre Anzeige neben den Suchergebnissen zum Suchbegriff „Autoversicherung“ erscheint.

Versicherung B bietet für das gleiche **2€**.

Vereinfachende Annahmen:

- Nur *eine* Anzeige kann jeweils angezeigt werden.
- Google wird nur für die Fälle bezahlt, in denen die Anzeige auch anklickt wird.

Wovon würden Sie an Googles Stelle abhängig machen, welche Anzeige angezeigt wird?

Wir führen im Folgenden einige Grundbegriffe ein, die für die Beschreibung der Lösung erforderlich sind, und kommen dann auf dieses Beispiel zurück.

Zufall

(anschauliche Definition)

Von **Zufall** spricht man, wenn etwas nicht beeinflussbar oder vorhersehbar ist. Dies ist vor Allem dann klar, wenn ein *unter gleichen Umständen wiederholbares Experiment* unterschiedliche Ergebnisse liefern kann.

- **Reproduzierbarer** genauer **Ablauf**
- was wird als ***Ergebnis*** betrachtet, welche Aspekte ausgeblendet?
  → **Ω:** Menge der möglichen Ergebnisse, von denen immer *genau eines* eintritt.

Zufallsexperiment

Was *„gleiche Umstände“*, *„beeinflussbar“* oder *„vorhersehbar“* bedeutet, ist oft subjektiv, bzw. Gegenstand einer Abstraktion (=Vereinfachung) der Realität.

wichtig

Zufallsexperiment

Einige **Definitionen**:

Ein **Zufallsexperiment** ist ein Vorgang, bei dem nicht sicher vorhersehbar ist, welches von mehreren Ergebnissen eintreten wird. Die Menge aller ***möglichen Ergebnisse*** wird als Ω bezeichnet. Es muss sichergestellt sein, dass stets ***genau eines*** der Ergebnisse aus Ω eintritt.

Mit Bezug auf ein bestimmtes Zufallsexperiment verwendet man:

Jede Teilmenge $A \subseteq \Omega$ von wird als ***Ereignis*** des Zufallsexperiments bezeichnet. Nach Durchführung des Experimentes sagt man, ***A ist eingetreten,* wenn** für das tatsächliche eingetretene Ergebnis ω gilt: $\omega \in A$

Die Menge $\mathcal{A}$ aller bildbaren Ereignisse heißt ***Ereignismenge***.
(D.h. $\mathcal{A} = \mathcal{P}(\Omega)$ = Potenzmenge von Ω).

Eine einelementige Teilmenge von Ω heißt ***Elementarereignis***.

Die leere Menge wird als das ***unmögliche Ereignis***, Ω selbst als das ***sichere Ereignis*** bezeichnet.

wichtig

Zufallsexperiment - Grundbegriffe

Beispiel:

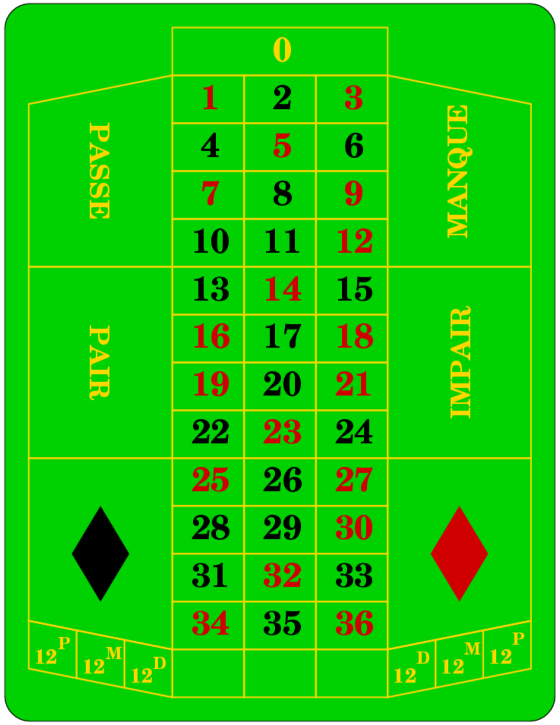
Beim Roulette betrachtet man üblicherweise
als Ergebnismenge: $\boldsymbol{\Omega} = \{0; 1; 2; 3; \ldots ; 36\}$

Beispiele für zugehörige **Ereignisse**:

P := {2; 4; 6; …; 36} = *„Pair“* = *„gerade Zahl“*

M := {1; 2; 3; …; 18} = „Manque“ = „Zahl aus der unteren Hälfte“

E := {1} = *„die Eins fällt“* (dies ist ein Elementarereignis)



Bildquelle: Von Lars H. Rohwedder de:Benutzer:RokerHRO - Eigenes Werk, Gezeichnet mit XFig, Gemeinfrei, https://commons.wikimedia.org/w/index.php?curid=267028

Wenn das Zufallsexperiment als Ergebnis $\omega = 3$ hat, würde gelten:

- $3 \notin P$, also ist das Ereignis P nicht eingetreten
- $3 \in M$, also ist das Ereignis M eingetreten

Egal, welches Ergebnis $\omega \in \Omega$ das Zufallsexperiment hat, gilt:

- Das unmögliche Ereignis {} tritt nicht ein, weil $\omega \notin \{\}$.
- Das sichere Ereignis Ω tritt ein, weil $\omega \in \Omega$.

Verknüpfung von Ereignissen
(anschauliche Definition)

Ereignis: Irgendetwas, von dem interessiert, ob es bei einem Zufallsexperiment eintritt. Formal: **Menge von Ergebnissen**.

Anders als *Ergebnisse* müssen sich *Ereignisse* nicht gegenseitig ausschließen.

Gibt es zu einem Zufallsexperiment mehrere interessante Ereignisse, z.B. *A* und *B, so* lassen sich weitere ableiten:

A ***und*** B : $A \cap B$

A ***oder*** *B* : $A \cup B$

nicht A : $\overline{A} = \Omega \setminus A$

Tritt ein, wenn bei einer Durchführung des Experimentes etwas passiert, das sowohl zum Ereignis A gehört als auch zum Ereignis B.

Definition: A, B heißen „***unvereinbar***“ oder „***disjunkt***“, wenn sie sich gegenseitig ausschließen, d.h. wenn $A \cap B = \{\}$. Das Gegenteil ist „***vereinbar***“.

Nicht immer ist die Ergebnismenge so klar wie beim Roulette.
Beispiel:
A:=„morgen ist es wolkenlos“ B:=„morgen ist es windig“ C:=„morgen regnet es“

wichtig

Häufigkeit

Angenommen, man hat ein Zufallsexperiment n Mal durchgeführt und die Ergebnisse festgehalten.

Definition

Für ein beliebiges Ereignis A heißt die absolute Anzahl an Durchführungen des Experimentes, bei denen A eingetreten ist, ***absolute Häufigkeit von A*** (bei uns mit h_A bezeichnet).

Der relative **Anteil** der Durchführungen, bei denen A eingetreten ist, heißt ***relative Häufigkeit von A*** (bei uns mit f_A bezeichnet).

Hat man das Experiment n Mal durchgeführt, gilt also $f_A = \frac{h_A}{n}$

Im Unterschied zur Begriff „Wahrscheinlichkeit“ bezieht sich „Häufigkeit“ also auf bereits **bekannte Ergebnisse**.

Bemerkung:

Häufigkeiten lassen sich nicht nur für Ergebnisse von Zufallsexperimenten angeben, sondern für beliebige Daten, die n gleichartige Objekte beschreiben. Der Begriff gehört also zur **beschreibenden Statistik**, nicht zur Wahrscheinlichkeitsrechnung.

Beispiel: *Relative Häufigkeit von Frauen unter den Kursteilnehmern.*

wichtig

Häufigkeit (Beispiel)

Von den letzten 500 Websitenbesuchern haben 100 die Anzeige angeklickt.

Dann ist die ***relative Häufigkeit von*** $\underbrace{\textbf{\textit{Klicks}}}_{\text{betrachtetes Ereignis}}$ 20% $\left(= \frac{100}{500}\right)$.

Wahrscheinlichkeit (1a): Statistischer W-Begriff

Empirische Beobachtung

Führt man *n* unabhängige Wiederholungen eines Zufallsexperimentes durch, und betrachtet dabei ein bestimmtes Ereignis A, so scheint dessen relative Häufigkeit *mit* zunehmendem *n* immer gegen einen Grenzwert zu konvergieren.

n	k	f_A
100	0	0,0%
1000	15	1,5%
10000	105	1,1%
100000	1012	1,0%

Mit anderen Worten:

Man beobachtet, dass das Ergebnis *einer* Durchführung eines Zufallsexperimentes zwar unvorhersehbar ist, die relativen Häufigkeiten **bei einer großen Anzahl an Wiederholungen** aber **vorhersagbar** sind: Diesen ***Grenzwert der relativen Häufigkeit*** von *A* für wachsenden Stichprobenumfang bezeichnen wir als **Wahrscheinlichkeit** des Ereignisses A. **(Statistischer Wahrscheinlichkeitsbegriff)**

Für beliebig oft wiederholbare Zufallsexperimente sind Wahrscheinlichkeiten damit objektiv beliebig genau bestimmbar.

Wahrscheinlichkeit (1a): Statistischer W-Begriff

Problem: Die „Definition“ des Begriffs „Wahrscheinlichkeit“ auf der vorigen Folie ist im mathematischen Sinn keine, weil dazu bewiesen sein müsste, dass dieser Grenzwert immer existiert.

Deswegen gehen die Mathematiker anders vor. Sie überlegen, welche Eigenschaften Wahrscheinlichkeiten haben müssten, wenn es die Grenzwerte der relativen Häufigkeiten geben sollte, und sagen: Wir nennen alles, was diese Eigenschaften erfüllt „Wahrscheinlichkeit“. Diese Eigenschaften sind die Kolmogorov-Axiome der nächsten Folie.

Wahrscheinlichkeit (axiomatische Definition)

Definition: Sei Ω die Ergebnismenge und $\mathcal{A}$ die Ereignismenge eines Zufallsexperimentes, und $\boldsymbol{P : \mathcal{A} \to [0; 1]}$ **eine Funktion** mit folgenden Eigenschaften (Kolmogorov-Axiome):

- $P(\Omega) = 1$ (sicheres Ereignis)
- Für beliebige *unvereinbare* Ereignisse (d.h. Ereignisse mit $A \cap B = \{\}$), gilt:
  $P(A \cup B) = P(A) + P(B)$

Dann heißt das Tripel $(\Omega, \mathcal{A}, P)$ *Wahrscheinlichkeitsraum*, und P *Wahrscheinlichkeitsmaß*, ***Wahrscheinlichkeitsverteilung*** oder kurz ***W-Verteilung***.

Für jedes Ereignis $A \in \mathcal{A}$ heißt dann
$P(A)$ die „***Wahrscheinlichkeit des Ereignisses A***" .

Man kann leicht zeigen, dass daraus folgt:

Satz:

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

$P(\{\}) = 0$ (Wahrscheinlichkeit des unmöglichen Ereignisses)

wichtig

Wahrscheinlichkeit (axiomatische Definition)

Die (wenigen) Kolmogorov-Axiome genügen, um daraus die gesamte Wahrscheinlichkeitstheorie zu folgern, darunter auch

Satz: Gesetz der großen Zahlen

Mit wachsender Anzahl *n* von Wiederholungen eines Zufallsexperimentes nähert sich die relative Häufigkeit jedes Ereignisses A mit an Sicherheit grenzender Wahrscheinlichkeit beliebig nahe der Wahrscheinlichkeit von A.

(präzisere Formulierung siehe nächste Folie)

wichtig

Wahrscheinlichkeit (axiomatische Definition)

Die (wenigen) Kolmogorov-Axiome genügen, um daraus die gesamte Wahrscheinlichkeitstheorie zu folgern, darunter auch

Satz (**Gesetz der großen Zahlen**)
Die Wahrscheinlichkeit dafür, dass
beim Ziehen einer Zufallsstichprobe vom Umfang n
der Unterschied zwischen der beobachteten relativen Häufigkeit f_A und der Wahrscheinlichkeit P(A) eines Ereignisses A
eine beliebig vorgebbare Genauigkeit $\varepsilon > 0$ einhält,
konvergiert für n gegen Unendlich gegen Eins:

$$\lim_{n \to \infty} P\left(| P(A) - f_A | < \varepsilon \right) = 1$$

Praktisch bedeutet das, dass mit an Sicherheit grenzender Wahrscheinlichkeit sich für hinreichend große Stichproben die relative Häufigkeit eines Ereignisses beliebig nahe der Wahrscheinlichkeit annähert.

Lösung der einführenden Beispiel-Problemstellung

Relevant sind die **Klickwahrscheinlichkeiten p_A und p_B**, mit denen Anzeige *A* bzw. *B* angeklickt werden, wenn Sie neben den Suchergebnissen zu „Autoversicherung“ erscheinen.

Bei einer sehr großen Zahl *N* an Versuchen wird die relative Häufigkeit von Klicks sehr nahe an diesen Wahrscheinlichkeiten liegen. Falls *N* Mal Anzeige A eingeblendet wird, ist also mit $N \cdot p_A \cdot 1€$ Einnahmen zu rechnen, bei B wären es $N \cdot p_B \cdot 2€$.

Google sollte also genau dann Anzeige A anzeigen, wenn $N \cdot p_A \cdot 1€ > N \cdot p_B \cdot 2€$, also wenn $p_A > 2p_B$, sonst B .

Wie passt das zu den formalen Definitionen der vorigen Folien?

Wir betrachten folgendes **Zufallsexperiment** für Anzeige A

- Wir warten auf die nächste Suchanfrage nach „Autoversicherung“, und blenden Anzeige A ein.
- Wir ignorieren Aspekte wie Herkunftsregion der Suchanfrage (**Abstraktion**)
- Wir beobachten, ob die Anzeige angeklickt wird, d.h. Ω = {Click, NoClick}

- Dieses Experiment kann als Zufallsexperiment betrachtet werden, weil es *unter gleichen Umständen reproduzierbar* ist, aber trotzdem unterschiedliche Ergebnisse liefern kann.
- Bezüglich dieses Zufallsexperimentes definieren wir $p_A := P(\text{Click})$

Für B betrachten wir ein anderes, analoges Zufallsexperiment und ermitteln daraus die Klickwahrscheinlichkeit p_B für Anzeige B.

1.1 Was Sie gelernt haben sollten

- Womit befasst sich ***Statistik***?
- Unter welchen Voraussetzungen spricht man von ***Zufallsexperiment***?
- Wann spricht man von **relativen Häufigkeiten**, wann von **Wahrscheinlichkeiten**? Welche Beziehung besteht zwischen beiden?
- Das Ergebnis eines Zufallsexperimentes ist zwar nicht sicher vorhersehbar, die relativen Häufigkeiten der Ergebnisse bei vielen Wiederholungen aber schon.
- Wofür stehen die zusammengesetzten Ereignisse $A \cap B$, $A \cup B$ bzw. $\bar{A}$?
- Wann nennt man Ereignisse „***unvereinbar***“?

Übungsaufgaben 1.A.1 – 1.A.5

1.2 Wie bestimmt man Wahrscheinlichkeiten?

- 1.1 Grundbegriffe
  - Zufall
  - relative/absolute Häufigkeiten
  - Wahrscheinlichkeiten
- 1.2 Wie bestimmt man Wahrscheinlichkeiten?
  - Empirisch
  - Durch Zählen: Laplace-Experimente
  - Aus bereits bekannten Wahrscheinlichkeiten: Rechengesetze für Wahrscheinlichkeiten
- 1.3 Bedingte Wahrscheinlichkeiten

In einer VL: Formalisierung des Google-Beispiels mit ZWEI Zufallsexperimenten, Laplace (Urnenbeispiel)

Wie bestimmt man Wahrscheinlichkeiten? (1)

-> Empirisch: Durch häufiges Wiederholen des Experimentes.

Um die Wahrscheinlichkeit eines Ereignisses A zu schätzen:

- Das Experiment sehr oft wiederholen
- Die relative Häufigkeit von A ermitteln.
- Prüfen, ob die Schätzgenauigkeit ausreicht. (-> schließende Statistik)

Beispiel 1

Google will schätzen, mit welcher Wahrscheinlichkeit eine bestimmte Anzeige angeklickt wird, wenn sie neben den Suchergebnissen zum Suchbegriff „Autoversicherung“ als oberstes angezeigt wird. Dazu hat Google die Anzeige bei 10 000 passenden Suchen ausprobiert, davon wurden 100 angeklickt. Schätzen Sie die Klickwahrscheinlichkeit.

Lösung: 100/10000 = **1%**

Beispiel 2

Sie haben drei Mal mit einem Würfel aus einem Zauberkasten gewürfelt und kein einziges Mal Sechs gewürfelt. Schätzen Sie die Wahrscheinlichkeit, mit der Sechs kommt.

Lösung: Keine sinnvolle Schätzung möglich, da zu wenig Wiederholungen.

wichtig

Wie bestimmt man Wahrscheinlichkeiten? (2)

-> **Durch Zählen**, wenn man weiß, dass alle Ergebnisse gleich wahrscheinlich sind.

Definition 26.7:
Ein **Laplace-Experiment** ist ein Zufallsexperiment mit folgenden Eigenschaften:

- Es gibt nur endlich viele mögliche Ergebnisse (zu unterscheidende Fälle).
- Jedes dieser Ergebnisse ist gleich wahrscheinlich.

Bemerkung: Oft ergibt sich die Laplace-Eigenschaft aus einer Symmetrie-Überlegung, so z.B. bei fast allen Glücksspielen, bei denen alle Karten, Lose, oder Seiten eines Würfels austauschbar sind.

Satz 26.8: Bei einem Laplace-Experiment mit *n* möglichen Ergebnissen hat jedes die Wahrscheinlichkeit $\frac{1}{n}$. Ein Ereignis *A*, das *k* davon umfasst, hat die Wahrscheinlichkeit

$$P(A) = \frac{k}{n} = \frac{Anzahl\ zu\ A\ gehörender\ Ergebnisse}{Anzahl\ möglicher\ Ergebnisse}.$$

Beispiel 26.9:

- Wie wahrscheinlich ist es, mit einem fairen Würfel eine gerade Zahl zu würfeln? $\frac{3}{6} = 0.5$
- In einer Charge von 100 HDMI-Kabeln sind 15 defekte. Eines wird zufällig herausgegriffen und getestet. Wie wahrscheinlich ist es defekt? $\frac{15}{100}$

wichtig

Beispiel

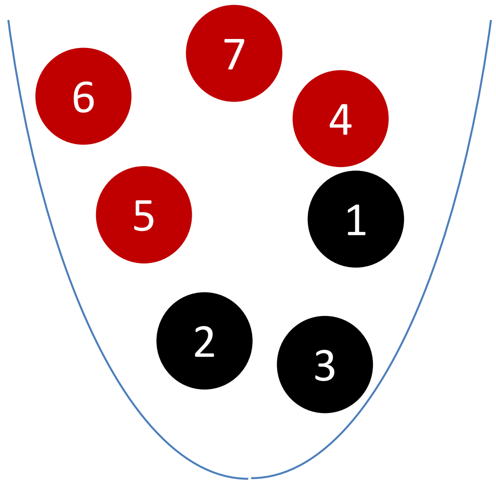
Eine Website hat 90 000 kostenlos registrierte und 10 000 zahlende Nutzer.

Ein Besucher startet eine neue Session, und füllt gerade die Login-Maske aus. Noch wissen wir also nicht, wer es ist. Wie wahrscheinlich ist es, dass es sich um einen kostenlos registrierten Nutzer handelt? (und es deshalb Sinn macht, ihm Werbung für ein Upgrade einzublenden)

Beispiel

Die Kugeln in einer Urne sind von 1 bis 7 durchnummeriert.
Kugeln **1-3** sind schwarz, **4-7** sind rot.
Wie wahrscheinlich ist es, dass eine zufällig gezogene Kugel

a) rot ist? (Ereignis R)
b) geradzahlig ist? (Ereignis G)
c) rot *und* geradzahlig ist?
d) rot *oder* geradzahlig ist?



Lösung

Alle Kugeln werden mit gleicher Wahrscheinlichkeit gezogen, also gilt:

a) $P(R) = 4/7$, da 4 von 7 gleich wahrscheinlichen Kugeln rot sind
b) $P(G) = 3/7$, " 3 von 7 " " " geradzahlig sind.
c) $P(R \cap G) = 2/7$, " 2 von 7 " " " rot und geradz.
d) $P(R \cup G) = 5/7$, " 5 von 7 " " " rot und geradz.
oder: $P(R \cup G) = P(R) + P(G) - P(R \cap G) = (4 + 3 - 2)/7 = 5/7$

Wie bestimmt man Wahrscheinlichkeiten? (3)

-> Aus anderen bereits bekannten Wahrscheinlichkeiten

Aus den Kolmogorov-Axiomen folgt neben den uns bereits bekannten Formeln:

Satz 26.11

a) **Additionsregel für vereinbare Ereignisse**:
Für beliebige (also nicht notwendigerweise unvereinbare) Ereignisse *A* und *B* gilt:

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

b) **Additionsregel für mehr als zwei unvereinbare Ereignisse**:
Für mehr als zwei Ereignisse $A_1 \ldots, A_n$, die paarweise unvereinbar sind, gilt:

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n)$$

„*Paarweise unvereinbar*" bedeutet, dass für jedes Paar von Ereignissen A_j, A_k $(j \neq k)$ gilt, dass A_j mit A_k unvereinbar ist. Insgesamt kann also ***höchstens eines*** der Ereignisse $A_1 \ldots A_n$ eintreten.

c) Die Wahrscheinlichkeit ist monoton, d.h. es gilt

$P(A) \leq P(B)$, wenn $A \subseteq B$, d.h. *A* ein Spezialfall von *B* ist.

Die Kolmogorov-Axiome und obige Rechengesetze gelten analog auch für relative Häufigkeiten!

wichtig

Beispiel

Hans hat einen ungünstigen Geburtstagstermin. Über die Jahre beträgt die Wahrscheinlichkeit, dass es an seinem Geburtstag

- regnet (Ereignis R): $P(R) = 60\%$
- kalt ist (Ereignis K): $P(K) = 70\%$
- regnet und kalt ist: $P(R \cap K) = 50\%$

Wie wahrscheinlich ist es, dass es an seinem nächsten Geburtstag

a) nicht regnet?
b) regnet oder kalt ist?
c) nicht regnet, und kalt ist?

Beispiel

Lösung

a) $P(\overline{R}) = 1 - P(R) = 1 - 0.6 = 0.4 = 40\%$

b) $P(R \cup K) = P(R) + P(K) - P(R \cap K) = 0.6 + 0.7 - 0.5 = 0.8 = 80\%$

c) Es gilt:

$$0.7 = P(K) = P(\ (R \cap K) \cup (\overline{R} \cap K)\ ) =^{*} P(R \cap K) + P(\overline{R} \cap K) =$$
$$= \quad 0.5 \quad + P(\overline{R} \cap K)$$

Offensichtlich folgt: $P(\overline{R} \cap K) = 0.7 - 0.5 =$ **<u>20%</u>**

(Bei * wurde benutzt, dass $(R \cap K)$ mit $(\overline{R} \cap K)$ *unvereinbar* ist.)

Visualisierung mit Kreuztabelle:

	R	$\overline{R}$	**Σ**
K	50%	20%	**70%**
$\overline{K}$	10%	20%	**30%**
Σ	**60%**	**40%**	**100%**

$P(R \cap K)$ $P(\overline{R} \cap K)$ **P(K)** **P($\overline{K}$)** $P(R \cap \overline{K})$ $P(\overline{R} \cap \overline{K})$ **P(R)** **P($\overline{R}$)**

Beispiel

Für die Projekte eines *Softwarebüros* betrachten wir die Ereignisse

R: Das Projekt wird **rechtzeitig** fertig

V: Das Projekt wird moderat **verspätet** fertig (Dauer ***bis 20%*** länger als geplant)

S: Das Projekt wird **stark verspätet** fertig (Dauer ***über 20%*** länger als geplant)

K: Die geplanten **Kosten** werden überschritten

Dabei sei P(*R*) = 1/6, P(*V*) = 1/3 , P(*K*) = 70% .

Bestimmen Sie, falls ohne weitere Informationen möglich:

a) $\mathrm{P}(\overline{R})$ b) P(*R* ∪ *V*) c) P(S) d) P(*S* ∪ *K*)

Lösung

a) $\mathrm{P}(\overline{R}) = 1 - P(R) = 1 - \frac{1}{6} = \frac{5}{6} \approx$ **<u>83 %</u>**

b) *R* und *V* sind unvereinbar, also gilt:

$P(R \cup V) = P(R) + P(V) = 1/6 + 1/3 =$ **<u>50%</u>**

c) (*R* ∪ *V*) ist das Gegenereignis zu *S*

Also gilt $\mathrm{P}(S) = 1 - P(R \cup V) = 1 - 3/6 = 3/6 =$ **<u>50%</u>** nach Satz 26.11 a

d) S und K sind vereinbar, P(*S* ∪ *K*) ist deshalb ohne Information über P(S∩K) **<u>nicht ermittelbar</u>**. (P(S∩K) lässt sich nicht berechnen, da nicht sicher ist, dass S und K voneinander stochastisch unabhängig sind.)

Beispiel

Bei einem manipulierten Würfel wurde empirisch ermittelt:

$$P(1)=\frac{1}{12},\quad P(6)=\frac{1}{4} \text{ und } P(2)=P(3)=P(4)=P(5)=\frac{1}{6}$$

(Die „1“ in „P(1)“ bezeichne dabei das Ereignis, dass eine Eins gewürfelt wird, „P(2),…,P(6) analog.)

a) Bestimmen Sie P(„gerade Augenzahl“)

b) Bestimmen Sie P(„ungerade Augenzahl“)

Da die Ergebnisse „1“, „2“, …, „6“ jeweils miteinander unvereinbar sind, kann man die Wahrscheinlichkeiten einfach addieren:

$$\text{a) } P(\{2,4,6\}) \overset{\text{nach 26.10}}{=} P(2)+P(4)+P(6) = \frac{1}{6}+\frac{1}{6}+\frac{1}{4}=\frac{7}{12} \approx \underline{58\,\%}$$

$$\text{b) } P(\{1,3,5\}) = P(\overline{\{2,4,6\}}) \underset{\text{nach 26.11a}}{=} 1 - P(\{2,4,6\}) = 1-\underset{\text{nach (a)}}{\frac{7}{12}} \approx \underline{42\,\%}$$

Wahrscheinlichkeit (1d): ***Subjektiver*** W-Begriff

Wie bestimmt man Wahrscheinlichkeiten (4) -> **durch Schätzen**

„Wie wahrscheinlich ist es, dass der Flughafen BER 2017 in Betrieb geht?“

Kein wiederholbares Zufallsexperiment; statistischer Wahrscheinlichkeitsbegriff nicht anwendbar.

Man kann subjektive Einschätzungen als ***Subjektive Wahrscheinlichkeiten*** ausdrücken. Deren Wert lässt sich nicht objektiv verifizieren, aber man kann damit basierend auf der eigenen Einschätzung Entscheidungen optimieren.

Umgekehrt kann man daraus, welche Wettquoten man annehmen würde und welche nicht, rückwärts ermitteln, wie man die Wahrscheinlichkeiten einschätzt.

Die subjektiven Wahrscheinlichkeiten, die die selbe Person verschiedenen Ereignissen beimisst, sollten den Kolmogorov-Axiomen genügen, ansonsten führt das zu widersprüchlichen Entscheidungen und ermöglicht Arbitrage-Wetten gegen die Person.

1.2 Was Sie gelernt haben sollten

- Wie kann man Wahrscheinlichkeiten bestimmen?
- Wann reicht einfaches Zählen?
- Wie berechnet man die Wahrscheinlichkeit von zusammengesetzten Ereignissen („***oder***“, „***nicht***“, ...)?
- Solche Berechnungen korrekt zu Papier bringen können.

1.A.6, 1.A.7, 1.B.5, 1.B.1

Fahrplan der Statistik-Vorlesung

1. Wahrscheinlichkeitsrechnung von Ereignissen

- *Zufall, Wahrscheinlichkeit, Zufallsereignis*
- Wie bestimmt man Wahrscheinlichkeiten
- *bedingte* Wahrscheinlichkeit, stochastische *Unabhängigkeit*

2. Beschreibende Statistik von ***quantitativen Merkmalen*** und Wahrscheinlichkeitsrechnung von ***Zufallsvariablen***

- *Mittelwert/Erwartungswert, Quantile, Varianz, Standardabweichung*
- Häufigkeitsverteilung / *Wahrscheinlichkeitsverteilung*
- Wichtige Verteilungstypen (*Binomial-, Poisson-, Exponential-, Normalverteilung*)
- Multivariate Statistik: *Kovarianz* und *Korrelation*
- lineare Regression

3. Schließende Statistik

- Lernen von Modellparametern aus Beispielen (z.B. Maximum Likelihood Ansatz)
- Anzahl erforderlicher Beispiele abschätzen

1.3 Bedingte Wahrscheinlichkeiten

- 1.1 Grundbegriffe
  - Zufall
  - relative/absolute Häufigkeiten
  - Wahrscheinlichkeiten

- 1.2 Wie bestimmt man Wahrscheinlichkeiten?
  - Durch Zählen: Laplace-Experimente
  - Aus bereits bekannten Wahrscheinlichkeiten: Rechengesetze für Wahrscheinlichkeiten

- 1.3 Bedingte Wahrscheinlichkeiten
  - Umgang mit „Teilwissen“

Diskussion

Aus langjähriger Erfahrung sei bekannt, dass es am 17.11. eines Jahres mit

- 60% Wahrscheinlichkeit regnet,
- 70% Wahrscheinlichkeit kalt ist

Wie wahrscheinlich ist es, dass es am 17.11.2016 regnet und kalt ist?

Einführendes Beispiel

Die Wahrscheinlichkeit, dass es an einem bestimmten Tag im Jahr regnet (Ereignis R) sei 60%, dass es kalt ist (Ereignis K) 70%, **und dass beides zusammen eintritt 50%.**

Sie sehen morgens aus dem Fenster: Es regnet.
Wie wahrscheinlich ist es, dass es kalt ist?

	R	$\overline{R}$	
K	$P(R \cap K)$ 50%	$P(\overline{R} \cap K)$ 20%	**P(K) 70%**
$\overline{K}$	$P(R \cap \overline{K})$ 10%	$P(\overline{R} \cap \overline{K})$ 20%	**P($\overline{K}$) 30%**
	P(R) 60%	**P($\overline{R}$) 40%**	**100%**

Lösung:

Die Wahrscheinlichkeit, dass es regnet und kalt ist, beträgt 50%. Diese Wahrscheinlichkeit gilt, wenn noch nichts über den Tag bekannt ist. Hier ist aber schon klar, dass es regnet. **Unter dieser Bedingung** ist die Wahrscheinlichkeit höher:

$$P(K \mid R) = \frac{P(K \cap R)}{P(R)} = \frac{0.5}{0.6} \approx 83\%$$

Wenn man schon weiß, dass es regnet, wie wahrscheinlich ist es kalt?
Schätzwert:
An wie viel Prozent ***der Regentage*** ist es kalt

Wenn man noch nichts weiß,
wie wahrscheinlich ist es kalt und regnet?
Schätzwert:
An wie viel Prozent ***aller*** Tage regnet es und ist kalt?

Die Wahrscheinlichkeit, dass es kalt ist, ist also davon abhängig, ob es regnet oder nicht. (Bei Regen ist sie höher.) Das nennt man ***stochastische Abhängigkeit*** zwischen K und R.

Bedingte Wahrscheinlichkeit

Ziel: Mit ***teilweise bekanntem Ergebnis*** eines Zufallsexperimentes umgehen.

Definition 26.15 (***bedingte Wahrscheinlichkeit***)
Seien A, B Ereignisse, und von einer Durchführung des Experimentes sei bekannt, dass *A* eingetreten ist.

Wahrscheinlichkeit von ***B* unter der Bedingung *A* :** $P(B|A) := \dfrac{P(A \cap B)}{P(A)}$

Damit gleichbedeutend ist offensichtlich:

Satz 26.17 (**Multiplikationssatz**)
Gegeben sind die Ereignisse *A* und *B* mit Wahrscheinlichkeiten ungleich Null. Dann gilt:

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

Ein Prozentwert, der sich **nicht auf alle** sondern **nur auf bestimmte Fälle bezieht**, beschreibt eine ***bedingte*** Wahrscheinlichkeit. (oder bedingte Häufigkeit)

Beispiele bedingter Wahrscheinlichkeiten:

1. ***Wenn*** eine Festplatte 3 Jahre ohne Ausfall funktioniert hat, ***wie wahrscheinlich ist dann***, dass Sie auch im 4. Jahr nicht ausfällt?
2. Wie wahrscheinlich klickt im Web ein Besucher, ***der ein Tablet benutzt***, auf die Anzeige?

wichtig

Gültigkeit von Wahrscheinlichkeiten

Seien A, B, C Ereignisse zu einem Zufallsexperiment.

P(A) ist zu verwenden,
wenn über eine Durchführung eines Zufallsexperimentes **nichts bekannt** ist.

P(A | B) ist zu verwenden,
wenn bekannt ist, dass **B eingetreten** ist, und **sonst nichts bekannt** ist.

P(A | B, C) = $P(A \mid B \cap C)$ ist zu verwenden,
wenn bekannt ist, dass **B und C eingetreten** ist, und **sonst nichts bekannt** ist.

Berechnete Wahrscheinlichkeiten werden mit neuen Informationen ungültig.

Beispiel:

Eine KFZ-Versicherung will die Schadenswahrscheinlichkeit eines Versicherten für das nächste Jahr schätzen:

P(Schaden) = **10 %** (ohne Informationen über den Versicherten)

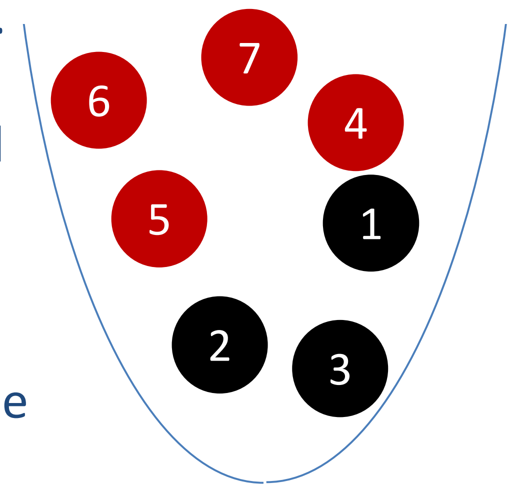
P(Schaden | Führerschein seit 1 Jahr) = **30 %**

P(Schaden| Führerschein seit 1 Jahr, 40000km unfallfrei) = **7 %**

(einfaches) Beispiel

Die Kugeln in einer Urne sind von 1 bis 7 durchnummeriert.
Kugeln **1-3** sind schwarz, **4-7** sind rot.
Wie wahrscheinlich ist es, dass eine zufällig gezogene Kugel

a) … rot ist? (Ereignis *R*)
b) … geradzahlig ist? (Ereignis *G*)
c) … rot *und* geradzahlig ist?
d) … geradzahlig ist, wenn sie schon gesehen haben, dass sie rot ist, die Nummer aber nicht erkennen konnten?



Lösung

a) 4/7 , da 4 von 7 gleich wahrscheinlichen Kugeln rot sind
b) 3/7 , “ 3 von 7 “ “ “ geradzahlig sind.
c) 2/7, “ 2 von 7 “ “ “ rot und geradzahlig sind.
d) Gesucht: Wahrscheinlichkeit von ***G unter der Bedingung R***:

$$P(G\,|\,R) = \frac{2}{4} = \frac{P(G \cap R)}{P(R)} = \frac{2/7}{4/7}$$

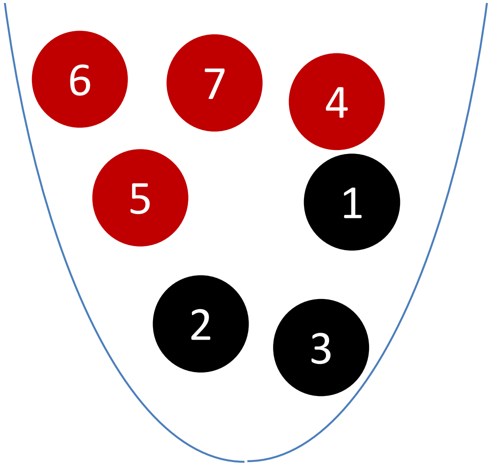
wenn man schon weiß, dass die Kugel rot ist, wie wahrscheinlich ist sie geradzahlig?

wenn man nichts weiß, wie wahrscheinlich ist die Kugel rot und geradzahlig?

Beispiel (Rechenregeln, bedingte W.)

Aus einer Urne mit vier roten und drei schwarzen Kugeln ziehen wir ohne Zurückzulegen nacheinander zwei Kugeln. Verwenden Sie die Ereignisbezeichner

A := „erstes Kugel ist rot“,
B := „zweite Kugel ist rot“



Formalisieren Sie die gesuchten Größen und bestimmen Sie:

a) Wie wahrscheinlich ist es, *zwei schwarze* Kugeln zu ziehen?

b) Wie wahrscheinlich ist es, *zuerst eine rote, dann eine schwarze* zu ziehen?

c) Wie wahrscheinlich ist es, *eine rote und eine schwarze* zu ziehen? (egal in welcher Reihenfolge)

d) Wie wahrscheinlich ist es, dass die erste *oder* zweite Kugel rot ist?

Lösung

a) Gesucht ist $P(\bar{A} \cap \bar{B})$. Es ist $P(\bar{A}\,) = \frac{3}{7}$ und $P(\bar{B}|\bar{A}) = \frac{2}{6}$, da nach dem Eintreten von $\bar{A}$ nur noch 6 Kugeln übrig sind, davon 2 schwarze. Daher ist

$$P(\bar{A} \cap \bar{B}) = P(\bar{A}) \cdot P(\bar{B}|\bar{A}) = \frac{3}{7} \cdot \frac{2}{6} \approx \underline{\mathbf{14.3\ \%}}$$

b) Die Wahrscheinlichkeit, zuerst eine rote und dann eine schwarze zu ziehen ist

$$P\left(A \cap \overline{B}\right) = P(A) \cdot P\left(\overline{B}\middle|A\right) = \frac{4}{7} \cdot \frac{3}{6} \approx \underline{\mathbf{28.6\ \%}}$$

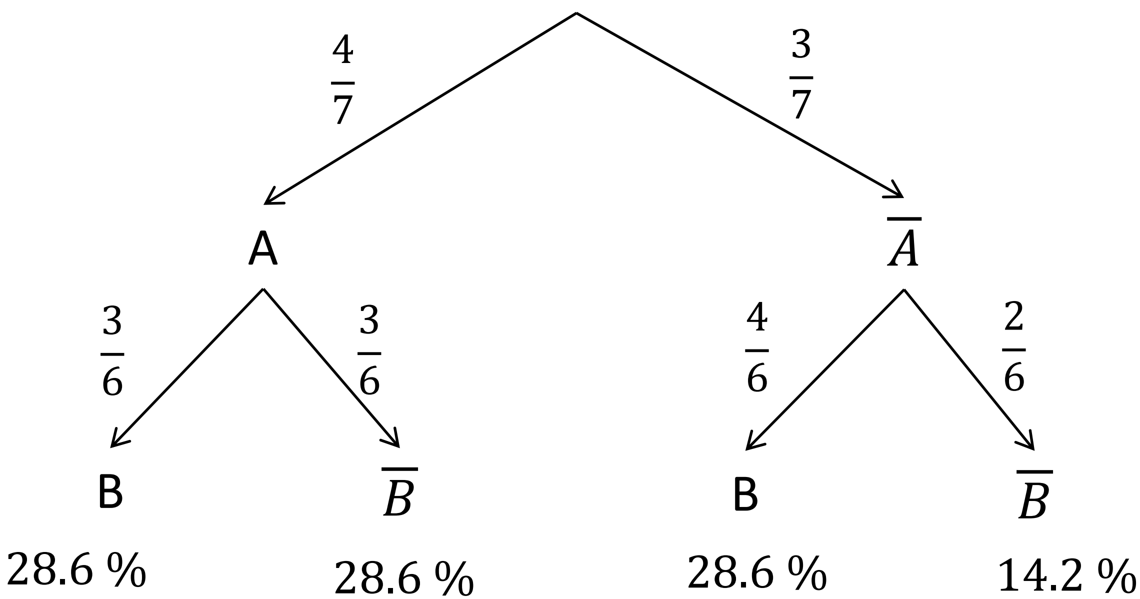
denn zuerst sind es 4 rote unter 7, danach 3 schwarze unter 6 Kabeln.

c) Das Ereignis tritt ein, wenn entweder die erste Kugel rot ist und die zweite schwarz oder umgekehrt die erste schwarz und die zweite rot ist. Da beide Varianten **unvereinbar** sind, gilt:

P(***genau eine*** rote) =

$$= P\left(\overline{A} \cap B\right) + P\left(A \cap \overline{B}\right) = P\left(\overline{A}\right) \cdot P\left(B\middle|\overline{A}\right) + P(A) \cdot P\left(\overline{B}\middle|A\right) = \frac{3}{7} \cdot \frac{4}{6} + \frac{4}{7} \cdot \frac{3}{6} \approx \underline{\mathbf{57.1\ \%}}$$

Visualisierung der Lösung des Beispiels mit Wahrscheinlichkeitsbaum



A: erste Kugel rot
B: zweite Kugel rot

a) $P(\overline{A}\cap\overline{B}) = \frac{3}{7}\cdot\frac{2}{6} \approx 14.2\%$

b) $P(\overline{A}\cap B) = \frac{3}{7}\cdot\frac{4}{6} \approx 28.6\%$

c) $P\left((A\cap\overline{B})\cup(\overline{A}\cap B)\right) = \frac{4}{7}\cdot\frac{3}{6}+\frac{3}{7}\cdot\frac{4}{6} \approx 57.1\%$

Lösung (Fortsetzung)

d) Ein „oder“ ist in mathematischem Kontext immer als nicht exklusives „oder“ gemeint.
Lösungsweg über Gegenereignis:

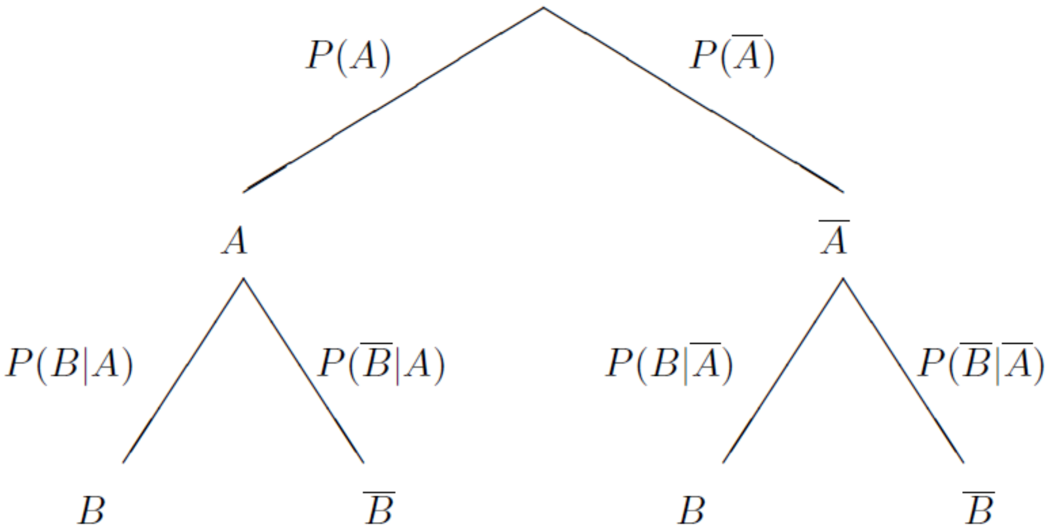
$$P(A \cup B) = 1 - P\left(\overline{A \cup B}\right) = 1 - P\left(\overbrace{\overline{A} \cap \overline{B}}^{\text{aus (a) bekannt}}\right) = 1 - \frac{3}{7} \cdot \frac{2}{6} \approx \underline{\mathbf{85.7\%}}$$

nach Satz 26.11a

Mengenlehre:
$\left(\overline{A \cup B}\right) = \left(\overline{A} \cap \overline{B}\right)$

Das Gegenteil von „beide nicht rot“

Wahrscheinlichkeitsbäume



	A	$\overline{\text{A}}$	Sum
B	$P(A \cap B)$	$P(\overline{A} \cap B)$	$\boldsymbol{P(B)}$
$\overline{\text{B}}$	$P(A \cap \overline{B})$	$P(\overline{A} \cap \overline{B})$	$\boldsymbol{P(\overline{B})}$
Sum	$\boldsymbol{P(A)}$	$\boldsymbol{P(\overline{A})}$	

Bedingte Wahrscheinlichkeiten lassen sich als ***Wahrscheinlichkeitsbaum*** visualisieren.

Voraussetzung: Es muss sichergestellt sein, dass **an jeder Verzweigung immer genau einer der Fälle** eintritt.

Durch **Multiplizieren** der (bedingten) Wahrscheinlichkeiten **entlang eines Pfades** erhält man die zugehörige totale Wahrscheinlichkeit.

Beispiel: $P\left(A \cap \overline{B}\right) = P(A) \cdot P(\overline{B}|A)$

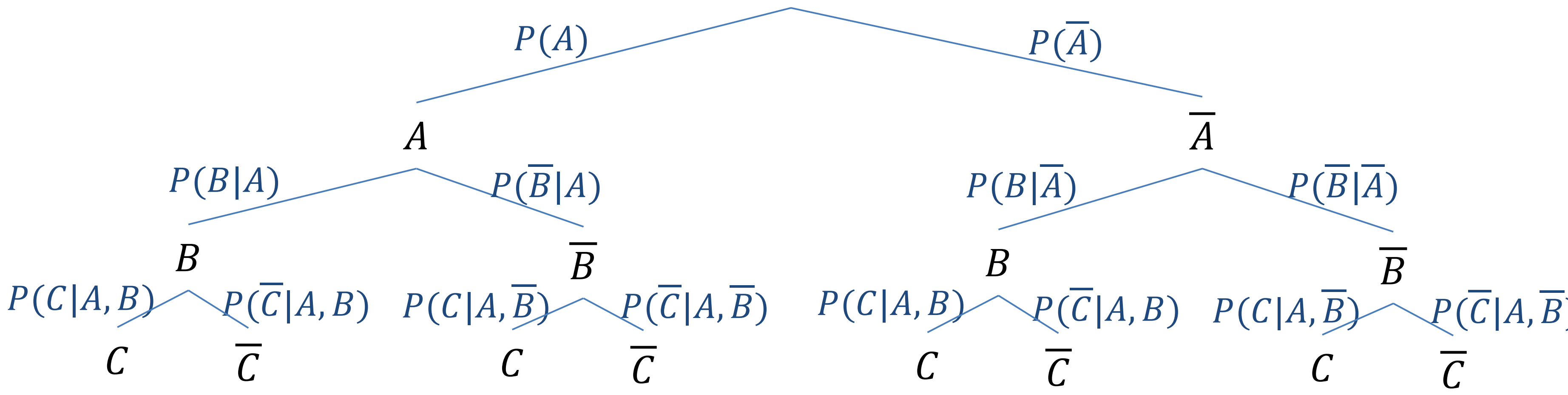
Der selbe Sachverhalt lässt sich auch ganz anders, z.B. als **Kreuztabelle der totalen Wahrscheinlichkeiten** angegeben.

Voraussetzung: Es muss sichergestellt sein, dass sich jedes Ergebnis **genau einer der Spalten** sowie **genau einer der Zeilen** (und damit genau einer Zelle) zuordnen lässt.

wichtig

Wahrscheinlichkeitsbäume

Wahrscheinlichkeitsbäume und der Multiplikationssatz lassen sich auch auf mehr als 2 Ereignisse verallgemeinern:



Aus dem Baum kann man z.B. ablesen:

$$P(A \cap B \cap C) = P(A) \cdot P(B \mid A) \cdot P(C \mid A, B)$$

$$P(\overline{A} \cap B \cap \overline{C}) = P(\overline{A}) \cdot P(B \mid \overline{A}) \cdot P(\overline{C} \mid \overline{A}, B)$$

Dabei wurde „ *A, B*“ als Kurzschreibweise für „ A∩B “ verwendet.
Im Multiplikationssatz (Satz 26.17) ist *A* durch „$A \cap B$“ sowie *B* durch *C* zu substituieren, dann ergibt sich:

$$P(A \cap B \cap C) = P(A \cap B) \cdot P(\mathrm{C} \mid \mathrm{A} \cap \mathrm{B})$$

Stochastische Unabhängigkeit, Multiplikationssatz für unabhängige Ereignisse

Definition 26.20
Zwei Ereignisse *A* und *B* (P(*A*)>0, P(B)>0) sind voneinander ***stochastisch unabhängig***, wenn eine (und damit alle) der folgenden äquivalenten Bedingungen erfüllt sind:

$$P(B \mid A) = P(B)\ , \qquad P(B \mid A) = P(B \mid \bar{A})\ ,$$
$$P(A \mid B) = P(A)\ , \qquad P(A \mid B) = P(A \mid \bar{B})\ ,$$
$$P(A \cap B) = P(A) * P(B)$$

ACHTUNG: ***Logische Unabhängigkeit***, d.h. dass alle Kombinationen von A bzw. $\overline{A}$ mit B bzw. $\overline{B}$ *möglich* sind, **genügt nicht** für stochastische Unabhängigkeit: Auch die ***Wahrscheinlichkeit***, dass B eintritt, darf nicht davon abhängen, ob A eintritt. Wenn wir kurz von ***Unabhängigkeit*** sprechen ist **immer stochastische Unabhängigkeit** gemeint.

Oft nimmt man Unabhängigkeit aufgrund des Versuchsaufbaus als offensichtlich gegeben an und nutzt das, um $P(A \cap B)$ aus $P(A)$ und $P(B)$ zu berechnen:

Beispiel: Die Wahrscheinlichkeit, beim ersten *und* zweiten Wurf eines Würfels Sechs zu werfen (Ereignis S_1 bzw. S_2), ist $\mathrm{P}(S_1 \cap S_2) = P(S_1) \cdot P(S_2 \mid S_1) = P(S_1) \cdot P(S_2) = \frac{1}{6} \cdot \frac{1}{6}$, da sich die Würfe offensichtlich nicht gegenseitig beeinflussen, also S_2 von S_1 stochastisch unabhängig ist.

wichtig

Beweis der Äquivalenzen aus Def 26.20

Für zwei Ereignisse A, B eines Zufallsexperimentes mit P(A)>0 und P(B)>0 sind folgende Bedingungen äquivalent:

$$P(B \mid A) = P(B) \qquad (1)$$

$$P(A \cap B) = P(A) \quad P(B) \qquad (2)$$

$$P(A \mid B) = P(A) \qquad (3)$$

Beweis:

Nach dem Multiplikationssatz gilt:

$P(A|B)P(B) = P(A \cap B)$ **(4a)** und $P(B|A)P(A) = P(A \cap B)$ **(4b)**

Wir zeigen (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1) :

(1) $\Rightarrow$ (2) : Mit (1) wird aus (4b):

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

Daraus folgt unmittelbar (2).

(2) $\Rightarrow$ (3) : (4a) in Verbindung mit (2) ergibt:

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

Daraus folgt unmittelbar (3) (da P(B)>0 angenommen wurde)

(3) $\Rightarrow$ (1): (4a) und (4b) zusammengesetzt ergeben:

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

Mit (3) folgt: $P(A)P(B) = P(B|A)P(A)$

Daraus folgt unmittelbar (1) (da P(A)>0 angenommen wurde)

Beweis der Äquivalenzen aus Def 26.20

Für zwei Ereignisse A, B eines Zufallsexperimentes mit P(A)>0 und P(B)>0 sind folgende Bedingungen äquivalent:

$P(B \mid A) = P(B)$ (1)

$P(B \mid A) = P(B \mid \bar{A})$ (2)

Beweis:

Wir zeigen (2) $\Rightarrow$(1):

Nach dem Satz von der totalen W. gilt:

$$P(B) = P(A \cap B) + P(\bar{A} \cap B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A}) = \quad \text{(mit (2))}$$

$$= P(B|A)P(A) + P(B|A)P(\bar{A}) = P(B|A)\big(P(A) + P(\bar{A})\big) = P(B|A) \cdot 1$$

Also gilt (1)

Wir zeigen (1) $\Rightarrow$(2):

Oben wurde bereits gezeigt, dass $P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A})$

Mit (1) erhält man: $P(B|A) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A})$

Umformen ergibt: $P(B|A) - P(B|A)P(A) = P(B|\bar{A})P(\bar{A})$

$$\Leftrightarrow \quad P(B|A)\big(1 - P(A)\big) = P(B|\bar{A})P(\bar{A})$$

$$\Leftrightarrow \quad P(B|A)P(\bar{A}) = P(B|\bar{A})P(\bar{A})$$

$$\Leftrightarrow \quad P(B|A) = P(B|\bar{A})$$

Stochastische Unabhängigkeit

Satz

Für zwei Ereignisse *A, B* sind folgende Aussagen äquivalent:

- A und B …
- A und $\bar{B}$ …
- $\bar{A}$ und B …
- $\bar{A}$ und $\bar{B}$ …

… sind voneinander stochastisch unabhängig

Beispiel:

Wenn das Ereignis *„es regnet"* unabhängig vom Ereignis *„es ist windig"* ist, dann ist auch *„es regnet nicht"* stochastisch unabhängig von *„es ist nicht windig"*.

wichtig

Bedingte Wahrscheinlichkeit (Bsp.)

Beispiel (zum Nachdenken):

Zwei faire Würfel werden im Würfelbecher geworfen und Sie sollen auf die Augensumme 2 wetten.

a) Angenommen, Sie haben mitbekommen, dass mindestens eine Eins dabei ist:
Geben Sie die Wahrscheinlichkeit von Augensumme 2 unter der Bedingung „mindestens eine Eins“ an.

Für welche der folgenden Präzisierungen von „Sie haben mitbekommen“ gilt Ihre Ergebnis aus (a)?

b) Die Würfel sind rot und blau, und Sie haben gesehen, dass der rote Würfel Eins ist; vom blauen haben Sie nichts gesehen.

c) Sie haben nichts gesehen, aber ein Schiedsrichter schaut beide Würfel an und sagt Ihnen dann, ob mindestens eine Eins dabei ist oder nicht.

Lösung zu Beispiel:

Wir verwenden folgende Ereignisse:

M1 :=„mindestens einer der Würfel ist Eins“

A2 :=„die Augensumme ist Zwei“

Wenn man die Würfel unterscheidet, dann gibt es 36 mögliche Ergebnisse, die alle jeweils eine Wahrscheinlichkeit von 1/36 haben. 11 dieser Ergebnisse enthalten *mindestens eine Eins* und nur *Eines Augensumme 2*.

a) $$P(A2|M1) = \frac{P(M1 \cap A2)}{P(M1)} = \frac{P(Zwei\ Einser)}{P(Mindestens\ eine\ Eins)} = \frac{1/36}{11/36} = \frac{1}{11}$$

Naheliegender Fehlschluss: Der eine Würfel ist Eins, über den anderen nichts bekannt, also ist die Wahrscheinlichkeit, dass er Eins ist 1/6.
Was ist daran falsch? Die Entscheidung, welcher der Würfel „der Eine“ und welcher „der Andere“ ist, erfolgt abhängig vom Wert. Dadurch ist die Wahrscheinlichkeit, dass der andere NICHT Eins ist, höher als 5/6.

b) Jetzt ist etwas anderes bekannt als in (a), nämlich dass der rote Würfel Eins ist und über den blauen nichts bekannt ist. Bei 6 der 36 Ergebnisse ist der rote Würfel Eins, nur bei einem einzigen ist der rote Würfel Eins und die Augensumme Zwei:

$$P(A2 \mid der\ rote\ Würfel\ ist\ Eins) = \frac{P(zwei\ Einser)}{P(der\ rote\ Würfel\ ist\ Eins)} = \frac{1/36}{6/36} = \frac{1}{6}$$

c) siehe (a)

Man muss also wirklich das reproduzierbare Zufallsexperiment zu den vorliegenden Ereignissen kennen, sprich zwischen (b) und (c) unterscheiden können, um eindeutig bestimmen zu können, mit welcher Wahrscheinlichkeit die Augensumme gleich 2 ist. Die Formulierung „unter der Bedingung, dass mindestens eine Eins dabei ist“ aus (a) hat aber eine eindeutige Bedeutung, nämlich im Sinne von (c).

Bedingte Wahrscheinlichkeit (Bsp.)

Beispiel (zur Unterstützung des vorhergehenden Beispiels):

In einer Trommel sind 10 000 Lose, davon 100 Gewinnlose. Zwei Lose werden gezogen und ungeöffnet nebeneinander hingelegt. Wir interessieren uns für folgende Ereignisse:

L: Das linke Los ist ein Gewinn.
R: Das rechte Los ist ein Gewinn.

Welche Variante ist Ihnen lieber:
i) Sie bekommen das linke, ich das rechte Los. (noch ungeöffnet)
ii) Ich darf mir beide Lose anschauen, und dann wählen, welches Sie und welches ich bekomme.

Lösung:
Dadurch dass ich bei ii das Los wählen darf, sinkt die Wahrscheinlichkeit dafür, dass das übrigbleibende Los ein Gewinnlos ist. Nur wenn beide Lose Gewinnlose sind, was sehr unwahrscheinlich ist, gewinnen Sie.

Analog dazu sinkt beim vorhergehenden Beispiel die Wahrscheinlichkeit, dass der übrigbleibende Würfel Eins ist, wenn wir zuvor nach Inspektion beider Würfel einen Einserwürfel (so vorhanden) entfernen.

Zusammenfassung der Rechengesetze aus Kapitel 1

1. $P(\{\}) = 0$ (Wahrscheinlichkeit des unmöglichen Ereignisses)
2. $P(\Omega) = 1$ (Wahrscheinlichkeit des sicheren Ereignisses)
3. $P(\bar{A}) = 1 - P(A)$ (Wahrscheinlichkeit des Gegenereignisses)

4. Ist Ereignis *A* mit *B* ***unvereinbar***, so gilt:
   $P(A \cup B) = P(A) + P(B)$
5. Für beliebige Ereignisse *A, B* gilt:
   $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

„ODER“

6. Für stochastisch ***unabhängige*** Ereignisse *A, B* gilt:
   $P(A \cap B) = P(A) \cdot P(B)$
7. Für beliebige Ereignisse *A, B* gilt:
   $P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$

„UND“

Alle diese Regeln gelten völlig analog auch für relative Häufigkeiten.

Übungsaufgabe:
Formulierungen für bedingte Wahrscheinlichkeiten

Ein Spam-Filter prüft ankommende Emails
und löscht diejenigen, die er für Spam hält.
Formalisieren Sie unter Verwendung der Ereignisse

S= „Email ist (wirklich) Spam“

L=„Email wird vom Spam-Filter gelöscht“

	S	$\overline{S}$
L	470	1
$\overline{L}$	10	19

Wie wahrscheinlich ist es, dass

a) eine Mail, die Spam ist, trotzdem nicht gelöscht wird?

b) eine ankommende Spam-Mail nicht gelöscht wird?

c) eine ankommende Mail Spam ist, der nicht gelöscht wird?

d) eine ankommende Mail Spam ist und trotzdem nicht gelöscht wird?

e) eine Mail, die nicht gelöscht wird, in Wahrheit Spam ist?

f) eine Mail nicht gelöscht wird, obwohl sie Spam ist?

g) Welcher Anteil der ankommenden Mails sind Spam-Mails, die nicht gelöscht werden?

h) Welcher Anteil der ankommenden Mails, die nicht gelöscht werden, sind Spam-Mails?

Beispiel (totale W., Bayes)

Beispiel

Ein Laptophersteller bezieht 20% seiner Displays vom Lieferanten L1, 30% von L2 und 50% von L3. Dabei weisen 6% der Displays von L1 Fehler auf, aber nur 4% der Displays von L2 und 3% derjenigen von L3:

	L1	L2	L3
Anteil	20 %	30 %	50 %
Fehlerquote	6 %	4 %	3 %

a) Welcher Anteil aller Laptops hat (langfristig) ein fehlerhaftes Display?
(Bzw: Mit welcher Wahrscheinlichkeit hat ein wahllos herausgegriffener Laptop ein fehlerhaftes Display?)

Tipp: Berechnen Sie zunächst $P(L_i \cap \text{Fehler})$ (für i=1,2,3)

b) Welcher Anteil aller fehlerhaften Displays stammt (langfristig) von L3?
(Bzw: Wenn ein wahllos herausgegriffener Laptop ein fehlerhaftes Display hat, wie wahrscheinlich ist es dann, dass es von L3 stammt.)

Beispiel (Lösung)

a)

Für einen zufällig herausgegriffenen Laptop gilt laut Angabe:

$P(L1) = 0.2$; $P(L2)=0.3$; $P(L3)=0.5$; $P(F \mid L1)=0.06$; $P(F \mid L2)=0.04$; $P(F \mid L3)=0.03$

Also ist die Wahrscheinlichkeit, dass er von *L1 und fehlerhaft* ist,

$P(L1 \cap F) = P(L1) \cdot P(F \mid L1) = 0.2 \cdot 0.06 = 1.2\ \%$ (Multiplikationssatz).

Ebenso gilt:

$P(L2 \cap F) = P(L2) \cdot P(F \mid L2) = 0.3 \cdot 0.04 = 1.2\ \%$

$P(L3 \cap F) = P(L3) \cdot P(F \mid L3) = 0.5 \cdot 0.03 = 1.5\ \%$

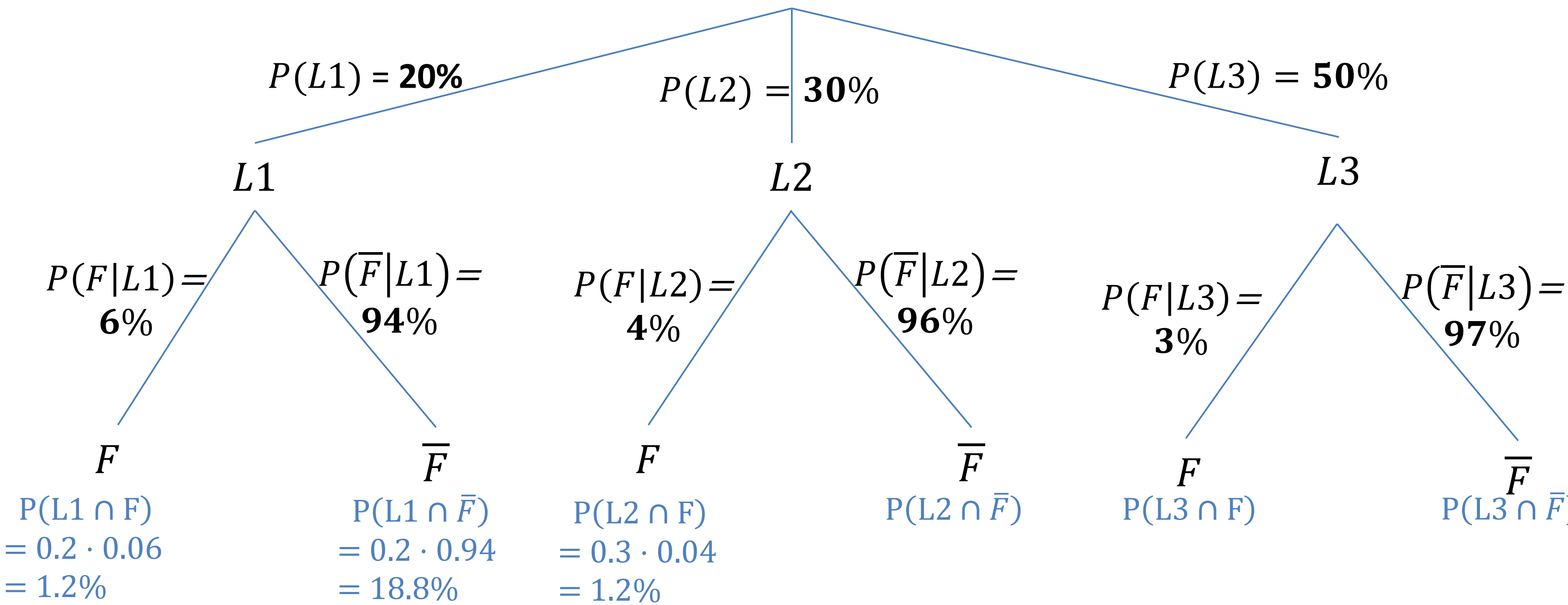
Insgesamt ist die Wahrscheinlichkeit, ein defektes Display zu erwischen also

$P(F) = P(L1 \cap F) + P(L2 \cap F) + P(L3 \cap F) = 1.2\% + 1.2\% + 1.5\% = \underline{\mathbf{3.9\%}}$

Beispiel (Lösung)

Die Informationen lassen sich als Wahrscheinlichkeitsbaum visualisieren:

F: *Display ist fehlerhaft*



Wahrscheinlichkeitsbäume eignen sich vor allem zur Visualisierung von **mehrstufigen Abläufen** und zugehörigen Berechnungen mit dem **Multiplikationssatz**.

Bei der oberen Verzweigung gibt es hier ***drei*** Äste.

Voraussetzung für die Verwendung eines Wahrscheinlichkeitsbaums: Es muss sichergestellt sein, dass **an jeder Verzweigung immer genau einer der Fälle** eintritt.

wichtig

Beispiel (Lösung)

F: *Display ist fehlerhaft*

Die selbe Information lässt sich auch als **Kreuztabelle** der totalen Wahrscheinlichkeiten von *Lieferant* und *F* visualisieren:

$P(L1 \cap F)$

	L1	L2	L3	Σ
F	**1.2%** $= 0.2 \cdot 0.06$	**1.2%** $= 0.3 \cdot 0.04$	**1.5%** $= 0.5 \cdot 0.03$	3.9%
$\overline{F}$	**18.8%** $= 0.2 \cdot 0.94$	**28.8%** $= 0.3 \cdot 0.96$	**48.5%** $= 0.5 \cdot 0.97$	96.1%
Σ	20%	30%	50%	

$P(F)$

$P(L1)$

Diese Art von Tabelle eignet sich vor allem zur Visualisierung von Berechnungen mit der Additionsregel für unvereinbare Ereignisse.

Voraussetzung dafür, dass die Zeilen- und Spaltensummen als Wahrscheinlichkeiten interpretierbar sind: Es muss sichergestellt sein, dass sich jedes Ergebnis **immer genau einer der Zeilen und genau einer der Spalten zuordnen lässt.** (D.h. sowohl Spalten als auch Zeilen definieren eine *Partitionierung* des Ergebnisraumes.)

wichtig

Beispiel (Lösung)

b)

Gesucht ist P(L3 | F).

Nach Definition der bedingten Wahrscheinlichkeit gilt:

$$P(L3\,|F) \;=\; \frac{P(L3 \cap F)}{P(F)} \;=\; \frac{0.015}{0.039} \approx \underline{\mathbf{38\ \%}}$$

Dabei wurden die Ergebnisse aus (a) verwendet.

Anschaulich: von 1000 zufällig herausgegriffenen Laptops sind in etwa 12 von L1 und fehlerhaft, in etwa 12 von L2 und fehlerhaft und in etwa 15 von L3 und fehlerhaft. Von insgesamt 39 fehlerhaften Displays sind also 15 von L3, also ca. 38%.

Zum Vergleich:

$$P(L1\,|F) \;= P(L2\,|F) \;= \frac{0.012}{0.039} \approx \mathbf{31\ \%}$$

Obwohl die Fehlerquote bei L3 am niedrigsten ist, stammen mehr fehlerhafte Displays von L3 als von L1 bzw. L2 – einfach weil L3 mehr Displays liefert als die anderen.

Satz von der Totalen Wahrscheinlichkeit

Mit diesem Satz kann man die totale (=nicht bedingte) Wahrscheinlichkeit von *A* aus bedingten Wahrscheinlichkeiten von *A* berechnen. Im vorhergehende Beispiel entsprach dabei E_k = „Display *ist von Lieferant k*" und *A* = „Display ist *defekt*"

Satz 26.23 (totale Wahrscheinlichkeit)

Wenn zu einem Zufallsexperiment die Ereignisse $E_1, E_2, \ldots, E_n$ eine Partitionierung von Ω bilden,

das heißt paarweise *unvereinbar* sind, und zusammen alle möglichen Ergebnisse umfassen, also $\Omega = E_1 \cup E_2 \cup \ldots \cup E_n$,
(also immer ***genau eines*** der Ereignisse eintreten muss)

dann

kann man die Berechnung der Wahrscheinlichkeit eines beliebigen Ereignis A aufspalten:

$$P(A) = \sum_{k=1}^{n} P(A \cap E_k) = \sum_{k=1}^{n} P(A|E_k) \cdot P(E_k)$$

Beweis: Unmittelbar aus der Additionsregel für unvereinbare Ereignisse
Bemerkung: Der Satz gilt analog auch für relative Häufigkeiten.

wichtig

Satz von der Totalen Wahrscheinlichkeit (Einfacher Spezialfall)

Jedes Ereignis E bildet zusammen mit dem zugehörigen Gegenereignis $\overline{E}$ eine Partitionierung. Der Satz von der totalen Wahrscheinlichkeit vereinfacht sich in diesem Spezialfall zu:

$$P(A) = P(A \cap E) + P\left(A \cap \overline{E}\right) =$$

$$= P(E) \cdot P(A|E) + P(\bar{E}) \cdot P(A|\bar{E})$$

Satz von Bayes

Mit dem Satz von Bayes kann man aus einer bedingten Wahrscheinlichkeit P(A|E) die ***umgekehrte bedingte Wahrscheinlichkeit*** P(E|A) berechnen:

Satz 26.34 (Satz von Bayes)

Für zwei beliebige Ereignisse A und E eines Zufallsexperimentes gilt:

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

Beweis:

Nach dem Multiplikationssatz (Satz 26.17) gilt:

$$P(A \cap E) = P(E) \cdot P(A\,|E) \;=\; P(A) \cdot P(E \mid A)$$

Auflösen nach P(E|A) liefert die gewünschte Behauptung.

wichtig

Satz von Bayes

Bemerkung:

Häufig ist bei Verwendung des Satz von Bayes $P(A)$ nicht bekannt, und muss mit dem Satz der totalen W. berechnet werden:

$$P(E\,|A) \overset{\text{Satz v. Bayes}}{=} \frac{P(E)\cdot P(A\,|E)}{P(A)} \overset{\text{Satz v. der totalen W.}}{=} \frac{P(E)\cdot P(A\,|E)}{P(E)\cdot P(A\,|E) + P(\overline{E})\cdot P(A\,|\,\overline{E}\,)}$$

oder im Fall einer beliebigen Partitionierung $E_1, E_2, \ldots, E_n$:

$$P(E_i\,|A) = \frac{P(E_i)\cdot P(A\,|E_i)}{P(A)} = \frac{P(E_i)\cdot P(A\,|E_i)}{\sum_{k=1}^{n} P(E_k)\cdot P(A\,|E_k)}$$

Dabei ist E_i eines der Ereignisse $E_1, E_2, \ldots, E_n$.

wichtig

Beispiel (Totale W. + Bayes)

Beispiel (analog 26.25)

Ein medizinischer Test erkennt Personen, die mit TBC infiziert sind, mit 98% Wahrscheinlichkeit (korrekterweise) als infiziert, und nicht infizierte Personen mit 99% Wahrscheinlichkeit (korrekterweise) als nicht infiziert. Weiterhin ist bekannt, dass von allen getesteten Personen 0.1% infiziert sind.

a) Wie wahrscheinlich ist es, dass der Test bei einer zufällig ausgewählten Person positiv ausfällt (d.h. TBC diagnostiziert wird)?

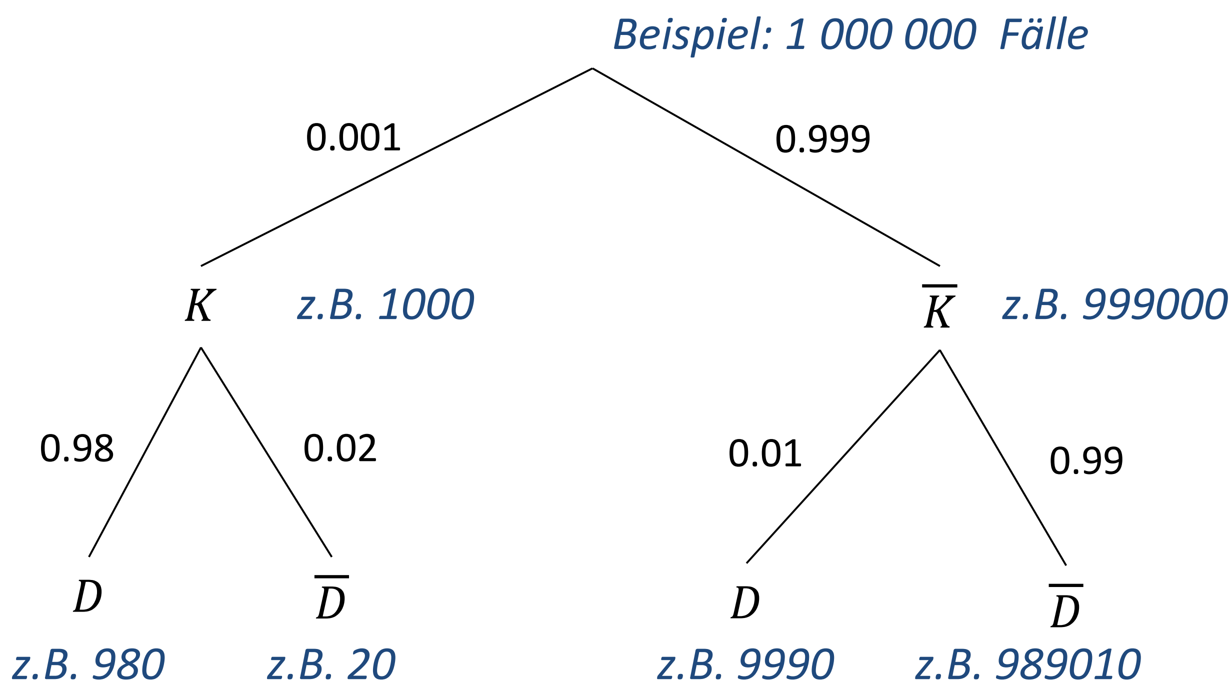
b) Eine Person lässt sich testen, der Test fällt positiv aus. Wie groß ist die Wahrscheinlichkeit, dass sie tatsächlich TBC hat?

Formalisieren Sie dazu zunächst die Angaben mit geeigneten Bezeichnungen und zeichnen Sie einen Wahrscheinlichkeitsbaum. TIPP: Wählen Sie Ihre Bezeichner so, dass sich sowohl die gegebenen als auch die gesuchten Größen leicht ausdrücken lassen.

Beispiel (Totale W. + Bayes)

K: *wirklich krank*
D: *Diagnose „krank“*

$$P(K) = 0.001$$
$$P(D \mid K) = 0.98$$
$$P(\overline{D} \mid \overline{K}) = 0.99$$
$$\Rightarrow$$
$$P(\overline{K}) = 0.999$$
$$P(\overline{D} \mid K) = 0.02$$
$$P(D \mid \overline{K}) = 0.01$$



Satz v. d. totalen W.

a) $\mathrm{P}(D) = P(D \cap K) + P(D \cap \overline{K}) = P(K) \cdot P(D \mid K) + P(\overline{K}) \cdot P(D \mid \overline{K}) =$
$= 0.001 \cdot 0.98 + 0.999 \cdot 0.01 = 0.00098 + 0.00999 = \underline{\mathbf{1.097\ \%}}$

Satz v. Bayes

b) $$\mathrm{P}(K \mid D) = \frac{P(K) \cdot P(D \mid K)}{P(D)} =$$

$$= \frac{P(K) \cdot P(D \mid K)}{P(K) \cdot P(D \mid K) + P(\overline{K}) \cdot P(D \mid \overline{K})} = \frac{0.001 \cdot 0.98}{0.001 \cdot 0.98 + 0.999 \cdot 0.01} = \frac{0.00098}{0.01097} = \underline{\mathbf{8.9\ \%}}$$

1.3 Was Sie gelernt haben sollten

- Bedeutung von ***bedingter Wahrscheinlichkeit*** und ***Unabhängigkeit***.
- Bedingte Wahrscheinlichkeiten interpretieren und in **Textaufgaben** als solche erkennen können. (Merke: Wenn es nicht darum geht, bei welchem Anteil von *allen*, sondern bei welchem Anteil einer bestimmten *Teilmenge* der Versuche ein Ereignis eintritt, ist es eine *bedingte* W.)
- Die **Notation** für bedingte Wahrscheinlichkeiten beherrschen.
- Die **Wahrscheinlichkeit von „*A* und *B*"** ausrechnen können, wenn A und *B* unabhängig sind, aber auch, wenn sie stochastisch abhängig sind.
- Zusammenhang zwischen ***Wahrscheinlichkeitsbaum***, bedingter W., und **totaler W**. und **Kreuztabellen** der totalen W. herstellen.
- Bedingte Wahrscheinlichkeiten aus anderen Wahrscheinlichkeiten ausrechnen
- Aus einer beliebigen Mischung von angegebenen bedingten und totalen Wahrscheinlichkeiten jede gesuchte bedingte oder totale Wahrscheinlichkeit berechnen. (z.B. auch mit Satz v. Bayes, Satz von der totalen W.)
- Kreuztabellen der totalen Wahrscheinlichkeiten bzw. Häufigkeiten nutzen, um derartige Berechnungen durchzuführen.

Übungsaufgaben: Alle verbleibenden vom Übungsblatt zu Kapitel 1

Bedingte Wahrscheinlichkeiten (Vertiefung)

Zwischen Wahrscheinlichkeiten unter der selben Bedingung gelten die gleichen Gesetze wie zwischen totalen (=„normalen“) Wahrscheinlichkeiten. D.h. wenn man in einem Rechengesetz bei allen Wahrscheinlichkeiten die selbe Bedingung hinzufügt, bleibt es gültig.

Beispiele:

$$P(\bar{A} \mid C) = 1 - P(A \mid C)$$
$$P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$$
$$P(A \cap B \mid C) = P(A \mid C) \cdot P(B \mid A \cap C)$$

Definition (bedingte Unabhängigkeit) Zwei Ereignisse A und B heißen ***stochastisch unabhängig unter der Bedingung C***, wenn eine (und damit alle drei) der folgenden äquivalenten Bedingungen erfüllt sind:

$$P(A \cap B \mid C) = P(A \mid C) \cdot P(B \mid C)$$
$$P(B \mid A, C) = P(B \mid C)$$
$$P(A \mid B, C) = P(A \mid C)$$

Dabei steht das Komma für „und“, also z.B. $P(A \mid B, C) := P(A \mid B \cap C)$

Beispiel: Unter der Bedingung, dass jemand krank ist, ist es unabhängig, ob dies von Testverfahren 1 bzw. Testverfahren 2 korrekt erkannt wird.

wichtig

Beispiel (bedingte Unabhängigkeit)

Sie haben einen ***gezinkten*** Würfel, bei dem die Sechs mit Wahrscheinlichkeit 1/2 geworfen wird,

sowie einen gleich aussehenden ***normalen*** Würfel.

Sie nehmen wahllos einen davon und würfeln zwei Mal nacheinander.

a) Wie wahrscheinlich ist es, dass Sie zwei Mal Sechs werfen?
TIPP: Benutzen Sie folgende Zwischenergebnisse:

   i. Wie wahrscheinlich ist es, dass Sie zwei Mal Sechs werfen, falls Sie den *normalen* Würfel erwischt haben?

   ii. Wie wahrscheinlich ist es, dass Sie zwei Mal Sechs werfen, falls Sie den *gezinkten* Würfel erwischt haben?

b) Angenommen, beim ersten Wurf kam eine Sechs. Wie wahrscheinlich ist es dann, dass auch beim zweiten Wurf eine Sechs kommt?

1. Was Sie gelernt haben sollten

- Was versteht man unter „Zufall“ bzw. Zufallsereignis?
- Grundbegriffe „**Wahrscheinlichkeit**“, „**relative Häufigkeit**“ und „**absolute Häufigkeit**“ sicher anwenden.
- Ereignisse mit den Operatoren „**und**“, „**oder**“, „**nicht**“ zu zusammengesetzten Ereignissen kombinieren.
- Wahrscheinlichkeiten von Ereignissen bestimmen durch
  - Zählen (bei **Laplace-Experiment**). Auch beurteilen, ob ein Laplace-Experiment vorliegt.
  - Berechnen aus den bekannten Wahrscheinlichkeiten anderer Ereignisse. (**Rechengesetze, Satz v.d. totalen W., Satz v. Bayes**)
- Stochastische **Unabhängigkeit** und **Unvereinbarkeit** von Ereignissen auseinanderhalten und erkennen.
- Begriff und Konzept der **bedingten Wahrscheinlichkeit** in Anwendungssituationen richtig umsetzen.
- Notation beherrschen. Rechenwege korrekt aufschreiben und angegebene Informationen sowohl in Formelnotation als auch in Textform korrekt interpretieren.