

# Stochastik und Statistik

Vorlesung Sommersemester 2016  
Dienstag 12 – 14 Uhr, M114 (HGB)  
Donnerstag 12 – 14 Uhr, A021 (HGB)

Vorlesung: Fabian Scheipl, Bernd Bischl  
Übung: Giuseppe Casalicchio

**Kontakt:** <vorname.nachname@stat.uni-muenchen.de>

**Webseite:** <https://www.elab.moodle.elearning.lmu.de>,  
Kursbereich Statistik

**Schein:** Erwerb durch Klausur

# Übungen

- Erste Übung am 19./20.04:  
Dienstag 14 - 16 Uhr  
Mittwoch 14 - 16 Uhr  
Mittwoch 16 - 18 Uhr **s.t.**
- Ort: HGB M209
- MELDEN SIE SICH BITTE ÜBER DIE WEBSEITE AN:
- Moodle-Einschreibeschlüssel: **Sto&Sta16**

# Literatur

- Held (2008): "Methoden der statistischen Inferenz. Likelihood und Bayes", Spektrum, 29,95 EUR (304 Seiten)
- Dümbgen (2003): "Stochastik für Informatiker", Springer Verlag, 30,79 EUR (268 Seiten)
- Fahrmeir, Künstler, Pigeot, Tutz (2010): "Statistik: Der Weg zur Datenanalyse", 7. Auflage, Springer Verlag, 30,79 EUR (610 Seiten)
- Georgii (2009): "Stochastik", 4. Auflage, deGruyter, 29,95 EUR (404 Seiten)
- Grimmett, Stirzaker (2001): "Probability and Random Processes", 3rd Edition, Oxford University Press, ca. 40 EUR (608 Seiten)
- Ligges (2008): "Programmieren mit **R**", 3. Auflage, Springer Verlag, 33,87 EUR (251 Seiten)
- Skript zur Vorlesung auf der Webseite

Die Graphiken und Beispiele im Skript und in der Vorlesung wurden mit **R** erstellt. Das Analysesystem **R** ist für alle gängigen Betriebssysteme frei erhältlich unter

<http://www.R-Project.org>

Eine geeignete deutschsprachige Einführung ist das Buch von Uwe Ligges “Programmieren mit **R**” (siehe Literaturliste). Auf der obigen Webseite sind elektronische Handbücher und diverse Einführungen in diese Programmiersprache erhältlich. Am Institut für Statistik werden Vorlesungen und Kurse zu **R** angeboten. Eine sehr gute IDE für **R** ist das frei erhältliche **RStudio**.

# Inhaltsverzeichnis

- 1 Einleitung
- 2 Laplace-Wahrscheinlichkeiten und diskrete Modelle
- 3 Bedingte Wahrscheinlichkeiten, stoch. Unabhängigkeit
- 4 Diskrete Zufallsvariablen
- 5 Erwartungswerte, Varianzen und Kovarianzen
- 6 Statistische Inferenz
- 7 Markov-Ketten
- 8 Stetige Zufallsvariablen
- 9 Statistische Inferenz II
- 10 Lineare Regression

# 1. Einleitung

$$\text{Stochastik} = \begin{cases} \text{Wahrscheinlichkeitstheorie} \\ \text{Statistik} \end{cases}$$

- Wahrscheinlichkeitstheorie: Mathematische Beschreibung von zufälligen Phänomenen
- Statistik:
  - ▶ Erhebung, Auswertung und Interpretation von Daten
  - ▶ Quantifizierung von Unsicherheit

# Ist Stochastik für die (Bio)Informatik wichtig?

- Simulation von zufälligen Prozessen am Computer  
Bsp.: Verbreitung von Epidemien
- Analyse von statistischen/randomisierten Algorithmen  
Bsp.: Quicksort,
- Statistische Analyse von Daten aus der Biologie und Genetik  
Bsp.: Genexpression und Überlebenszeit
- Stochastische Modelle für das Auftreten von Daten  
Bsp.: Hardy-Weinberg-Gesetz

vgl. Journal *Bioinformatics*, Table of Contents

# Wahrscheinlichkeit

Grundbegriff der Stochastik: Wahrscheinlichkeit  $P(A)$  für das Auftreten eines bestimmten Ereignisses  $A$

$P(A) = 1$  :  $A$  tritt mit Sicherheit ein

$P(A) = 0$  :  $A$  tritt mit Sicherheit *nicht* ein

$P(A) = p \in (0, 1)$  : Ereignis  $A$  tritt mit Wahrscheinlichkeit  $p$  ein

Interpretation?



# Subjektivistische Interpretation

## Wahrscheinlichkeit aus Wetteinsatz

Frage : "Wie sicher bist **Du**, dass das Ereignis  $A$  eintreten wird?"

: "Wie viel Einsatz  $E$  wirst **Du** maximal setzen, wenn beim Eintreten von  $A$  ein Gewinn  $G$  ausbezahlt wird?"

$$\rightsquigarrow P(A) = \frac{E}{G}$$

Wahrscheinlichkeit als Maß für **Deine** Unsicherheit.

## Beispiel: Spiel mit drei Bechern

Unter einen von drei gleichartigen Bechern wird eine weiche Kugel gelegt. Nun beginnt der Anbieter, die Becher vor den Augen des Spielers zu vertauschen.

Der Spieler muss nach einer gewissen Zeit sagen, unter welchem Becher die Kugel liegt. Wenn er die Kugel findet, gewinnt er den doppelten Einsatz.

Ereignis  $A$ : "Kugel gefunden"

- Spieler glaubt:  $P(A) > 1/2$ , möglichst nahe bei Eins
- Anbieter glaubt:  $P(A) < 1/2$ , im Idealfall  $P(A) = 1/3$

# Frequentistische Interpretation

“Häufigkeitsinterpretation”

Angenommen das Zufallsexperiment kann beliebig oft wiederholt werden, dann konvergiert die **relative Häufigkeit** des Eintretens des Ereignisses  $A$  gegen die Wahrscheinlichkeit  $P(A)$ .

Klassisches Beispiel: Wiederholtes Werfen eines “fairen” Würfels; Simulation durch Funktion `sample()` in **R**; Berechnung der **kumulierten relativen Häufigkeiten** von interessierenden Ereignissen.

## 2. Laplace-Wahrscheinlichkeiten und diskrete Modelle

Prinzip von Laplace (Pierre-Simon de Laplace [1749-1827]):

*“Wenn nichts dagegen spricht, gehen wir davon aus, dass alle Elementarereignisse gleichwahrscheinlich sind.”*

Frage: Was sind “Elementarereignisse”?

Ein Element  $\omega$  einer **Grundgesamtheit**  $\Omega$  nennt man **Elementarereignis**.

## 2.1 Laplace Wahrscheinlichkeiten

Betrachte die endliche Grundgesamtheit von Elementarereignissen

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

Für ein Ereignis  $A \subseteq \Omega$  definiert man die **Laplace-Wahrscheinlichkeit** als die Zahl

$$P(A) := \frac{|A|}{|\Omega|} = \frac{|A|}{n}$$

wobei  $|A|$  die Anzahl der Elemente in  $A$  ist.

# Folgerungen und Erweiterungen

- Jedes Elementarereignis  $\omega_i, i = 1, \dots, n$  hat also die Wahrscheinlichkeit  $P(\{\omega_i\}) = \frac{1}{n}$ .
- Die Wahrscheinlichkeit von  $\Omega$  ist  $P(\Omega) = 1$ .
- Die entsprechende **Abbildung**  $\mathcal{P}(\Omega) \rightarrow [0, 1]$  nennt man auch **diskrete Gleichverteilung** auf  $\Omega$ .
- Hierbei nennt man  $\mathcal{P}(\Omega)$  die **Potenzmenge** (Menge aller Teilmengen) von  $\Omega$  (nicht zu verwechseln mit  $P(\Omega)$ !).

## Beispiel: Augensumme von zwei Würfeln

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

$$|\Omega| = n = 6^2 = 36$$

Sei  $A_k$  das Ereignis "Augensumme ist  $k$ ".

Dann gilt:

$$P(A_k) = \frac{6 - |k - 7|}{36} \quad \text{für } k = 2, \dots, 12$$

## Beispiel: Skatspiel

Beim Skatspiel werden 32 verschiedene Karten, darunter 4 Buben an 3 Spieler verteilt. Jeder Spieler erhält 10 Karten. 2 Karten kommen in den Skat. Wie groß ist nun die Wahrscheinlichkeit der Ereignisse:

$A_1$  := "Spieler 1 erhält alle Buben"

$A_2$  := "Jeder Spieler erhält genau einen Buben"



# Laplace-Verteilungen sind zu speziell:

Beispiele:

- Unfairer Würfel
- Wahrscheinlichkeit für Knabengeburt
- Seltene Ereignisse, z.B. hard disk failure, Mutationen, ...

↪ Elementarereignisse hier nicht gleichwahrscheinlich!

Ein weiteres Problem ist, dass manchmal  $|\Omega|$  unendlich ist.

## Beispiel für unendliche Grundgesamtheiten

Man interessiere sich für die Anzahl der Würfe einer fairen Münze bis zum ersten Mal Zahl eintritt.

$$\Omega \text{ unendlich: } \Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \dots\} = \{1, 2, 3, 4, \dots\} = \mathbb{N}$$

Allgemein mit  $\omega_i = i$ :

$$P(\{\omega_i\}) = \frac{1}{2^i} \quad i = 1, 2, 3, \dots$$

$$\sum_{i=1}^{\infty} P(\{\omega_i\}) = \sum_{i=1}^{\infty} \frac{1}{2^i} = 1 \quad (\text{geom. Reihe}) \text{ vgl. Analysis}$$

## 2.2 Diskrete Wahrscheinlichkeitsräume

Ein diskreter Wahrscheinlichkeitsraum ist ein Paar  $(\Omega, P)$  wobei

- $\Omega$  eine abzählbare Grundgesamtheit ist und
- $P$  ein diskretes Wahrscheinlichkeitsmaß, das jeder Teilmenge  $A \subset \Omega$  eine Wahrscheinlichkeit  $P(A)$  zuordnet.

Diese definiert man wieder über die Wahrscheinlichkeiten  $P(\{\omega\})$  der Elementarereignisse  $\omega \in A$ :

$$P(A) = \sum_{\omega \in A} P(\{\omega\})$$

# Diskrete Wahrscheinlichkeitsräume II

wobei für  $P(\{\omega\})$  gelten muss:

$$0 \leq P(\{\omega\}) \leq 1 \quad \text{für alle } \omega$$

$$\text{und } \sum_{\omega \in \Omega} P(\{\omega\}) = 1.$$

Beispiel für unendlichen Wahrscheinlichkeitsraum:  $\Omega = \mathbb{N}$ ;  $\omega_i = i$

Betrachte z.B.  $P(\{\omega_i\}) = 1/(i(i+1))$  oder  $P(\{\omega_i\}) = 1/2^i$

## 2.3 Axiome von Kolmogorov [1903-1987]

Wir betrachten eine beliebige (abzählbare) Grundgesamtheit  $\Omega$  und eine Funktion  $P$  auf der Potenzmenge  $\mathcal{P}(\Omega)$ , die jedem Ereignis  $A \subseteq \Omega$  eine Wahrscheinlichkeit zuordnet.

Wir nennen  $P$  eine **Wahrscheinlichkeitsverteilung** auf  $\Omega$ , wenn sie folgende Eigenschaften erfüllt:

$$\text{A1) } P(A) \geq 0 \quad \text{für beliebige } A \subseteq \Omega$$

$$\text{A2) } P(\Omega) = 1$$

$$\text{A3) } P(A \cup B) = P(A) + P(B) \quad \text{für disjunkte Ereignisse} \\ A, B \subseteq \Omega$$

# Folgerungen

- $P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$   
für **paarweise disjunkte** Ereignisse  $A_1, A_2, \dots, A_n \subset \Omega$
- $P(A) \leq P(B)$  falls  $A \subset B$
- Definiere das **Komplement** von  $A$ :  $\bar{A} = \Omega \setminus A$ .  
Dann gilt  $P(\bar{A}) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$  für beliebige  $A, B \subset \Omega$

↪ Darstellung im Venn-Diagramm (John Venn [1834-1923])

# Die Siebformel von Sylvester-Poincaré

James Sylvester [1814-1897]

Jules Henri Poincaré [1854-1912]

Für beliebiges  $n \in \mathbb{N}$  und  $A_1, A_2, \dots, A_n \subset \Omega$  gilt:

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n) &= \sum_i P(A_i) - \sum_{i < j} P(A_i \cap A_j) \\ &\quad + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) \\ &\quad \pm \dots + (-1)^{n+1} \cdot P(A_1 \cap A_2 \cap \dots \cap A_n) \end{aligned}$$

Daher:  $P(A \cup B \cup C) = ???$

# Bonferroni Ungleichungen

→ Abschätzungen von  $P(A_1 \cup A_2 \cup \dots \cup A_n)$

Für beliebige Ereignisse  $A_1, A_2, \dots, A_n$  ist

$$P(\cup_{i=1}^n A_i) = \begin{cases} \leq \sum_i P(A_i) \\ \geq \sum_i P(A_i) - \sum_{i < j} P(A_i \cap A_j) \end{cases}$$



## 3.1 Bedingte Wahrscheinlichkeiten

Für Ereignisse  $A, B \subset \Omega$  mit  $P(B) > 0$  definiert man die bedingte Wahrscheinlichkeit von  $A$  gegeben  $B$  als die Zahl

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

↪ Darstellung im Venn-Diagramm

## Beispiel: Spiel mit drei Bechern

Spieler macht einen Trick und markiert einen der drei Becher.

Sei

$A :=$  "Spieler findet den richtigen Becher"

$B :=$  "Spielanbieter legt die Kugel unter den markierten Becher"

Dann:

$$P(A|B) = 1$$

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

# Eigenschaften von bedingten Wahrscheinlichkeiten

$$P(B|B) = 1$$

$$P(\bar{B}|B) = 0$$

$$P(A|B) \geq 0 \text{ für beliebige } A \subset \Omega$$

$$P((A_1 \cup A_2)|B) = P(A_1|B) + P(A_2|B) \\ \text{für } A_1 \text{ und } A_2 \text{ disjunkt}$$

Daher:

Als Funktion von  $A \subset \Omega$  ist  $P(A|B)$  (bei festem  $B$ !) eine Wahrscheinlichkeitsverteilung

## Beispiel: Skat

Sei

$A :=$  "Mindestens eine der acht Karokarten liegt im Skat"

$B :=$  "Spieler 1 erhält beim Austeilen keine der acht Karokarten"

Berechne  $P(A)$  und  $P(A|B)$  und vergleiche diese.

# Multiplikationssatz

Für beliebige Ereignisse  $A_1, A_2, \dots, A_n$  mit  $P(A_1 \cap A_2 \cap \dots \cap A_n) > 0$  gilt:

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

wobei man die rechte Seite offensichtlich auch in jeder anderen möglichen Reihenfolge faktorisieren kann.

Wir schreiben im Folgenden auch gerne  $P(A_1, A_2) := P(A_1 \cap A_2)$  etc.

Insbesondere gilt also

$$\begin{aligned} P(A_1, A_2) &= P(A_1) \cdot P(A_2|A_1) \\ P(A_1, A_2, A_3) &= P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1, A_2) \end{aligned}$$

# Satz der totalen Wahrscheinlichkeit

Sei  $B_1, B_2, \dots, B_n$  eine **disjunkte Zerlegung** von  $\Omega$ :

- 1  $B_1, B_2, \dots, B_n$  paarweise disjunkt:  $B_i \cap B_j = \emptyset \quad \forall i \neq j$
- 2  $B_1 \cup B_2 \cup \dots \cup B_n = \Omega$

Falls zusätzlich  $P(B_i) > 0$  für  $i = 1, \dots, n$  so gilt für jedes  $A \subset \Omega$ :

$$P(A) = \sum_{i=1}^n P(A|B_i) \cdot P(B_i)$$

→ Illustration im Venn-Diagramm

# Ein wichtiger Spezialfall

Insbesondere gilt

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

da  $B, \bar{B}$  eine disjunkte Zerlegung von  $\Omega$  ist.

Beweis?

## 3.2 Der Satz von Bayes I

Thomas Bayes [1701-1761]

Dieser Satz beruht auf der Asymmetrie der Definition von bedingten Wahrscheinlichkeiten:

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

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



## Der Satz von Bayes II

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

totale W'keit

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

Allgemeiner gilt für eine disjunkte Zerlegung  $B_1, \dots, B_n$

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^n P(A|B_i)P(B_i)}$$

# Interpretation

$P(B_i)$  “*a-priori*-Wahrscheinlichkeiten”

$P(B_i|A)$  “*a-posteriori*-Wahrscheinlichkeiten”

Nach Beobachtung von  $A$  ändert sich die Wahrscheinlichkeit von  $B_i$  von  $P(B_i)$  zu  $P(B_i|A)$

## Beispiel: Diagnostischer Test

$K$  := “Person ist krank”

$T$  := “Test auf Krankheit ist positiv”

Üblicherweise kennt man die:

$$\text{Sensitivität } P(T|K) \Rightarrow P(\bar{T}|K) = 1 - P(T|K)$$

$$\text{Spezifität } P(\bar{T}|\bar{K}) \Rightarrow P(T|\bar{K}) = 1 - P(\bar{T}|\bar{K})$$

$P(K)$  heißt **Prävalenz**

## Beispiel: Diagnostischer Test II

Zahlenbeispiel:

$$P(T|K) = 0.222 \quad P(\bar{T}|\bar{K}) = 0.993 \quad P(K) = 0.0264$$

Es ergibt sich

$$P(T) = P(T|K)P(K) + (1 - P(T|\bar{K}))(1 - P(K)) \approx 0.012676$$

$$P(K|T) = P(T|K)P(K)/P(T) \approx 0.462$$

$$P(\bar{K}|\bar{T}) \approx 0.979$$

## Beispiel: Creutzfeldt-Jakob

	CJD		
14-3-3	+	-	Total
+	126	7	133
-	8	97	105
Total	134	104	238

Die Sensitivität ist  $P(14-3-3 = + | \text{CJD} = +) = 126/134 = 0.94$  und die Spezifität ist  $P(14-3-3 = - | \text{CJD} = -) = 97/104 = 0.93$ .

# Chancen

Oft ist es einfacher, W'keiten als Chancen, im Sinne von Wettchancen (etwa 1:10), engl. "odds" aufzufassen. Für eine W'keit  $\pi$  besteht zur Chance  $\gamma$  der Zusammenhang

$$\gamma = \frac{\pi}{1 - \pi}$$

$$\pi = \frac{\gamma}{1 + \gamma}$$

# Eine Version vom Satz von Bayes

Eine elementare Umformung liefert

$$\frac{P(K|T)}{P(\bar{K}|T)} = \frac{P(K)}{P(\bar{K})} \cdot \frac{P(T|K)}{P(T|\bar{K})}$$

$$\textit{Posterior Odds} = \textit{Prior Odds} \cdot \textit{Likelihood Ratio}$$

$$\textit{Posteriori Chance} = \textit{Priori Chance} \cdot \textit{Likelihood Quotient}$$

Im Beispiel ergibt sich:

Priori Chance: 0.027

Likelihood Quotient: 31.7

↪ Posteriori Chance: 0.858

## Beispiel: O.J. Simpson Prozess

O.J. Simpson gab zu, seine Frau geschlagen zu haben.

Er gab aber nie zu, seine Frau ermordet zu haben.

Einer seiner Verteidiger sagte, die Wkeit, dass jemand, der seine Frau schlägt, auch ermordet, ist  $1/1000$ .

Aber: Das Gericht ist mehr interessiert an der Wkeit, dass O.J. Simpson der Mörder seiner Frau ist, gegeben er hat sie geschlagen **und sie ist ermordet worden**.



## Beispiel: O.J. Simpson Prozess II

Zwei Ansätze zur Lösung! Sei

$S$  : “Ein Mann hat seine Frau geschlagen”

$M$  : “Die Frau wurde ermordet”

$G$  : “Der Mann ist des Mordes schuldig”

$$\frac{P(G|S, M)}{P(\bar{G}|S, M)} = \frac{P(G|M)}{P(\bar{G}|M)} \cdot \frac{P(S|G, M)}{P(S|\bar{G}, M)} \quad (1)$$

$$\text{aber auch} = \frac{P(G|S)}{P(\bar{G}|S)} \cdot \frac{P(M|G, S)}{P(M|\bar{G}, S)} \quad (2)$$

## Beispiel: O.J. Simpson Prozess III

Ansatz (1) liefert (unter geeigneter Vorinformation)

$$\frac{P(G|S, M)}{P(\bar{G}|S, M)} = 4.08$$

Ansatz (2) liefert (unter anderer Vorinformation)

$$\frac{P(G|S, M)}{P(\bar{G}|S, M)} = 1.0$$

## Beachte:

Alle Aussagen für Wahrscheinlichkeiten gelten auch für bedingte Wahrscheinlichkeiten (bei fester Bedingung  $C$  mit  $P(C) > 0$ ).

Daher z.B.

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

und eben auch:

$$\begin{aligned} \frac{P(G | S, M)}{P(\bar{G} | S, M)} &= \frac{P(G | M)}{P(\bar{G} | M)} \cdot \frac{P(S | G, M)}{P(S | \bar{G}, M)} \\ \text{und} \quad \frac{P(G | S, M)}{P(\bar{G} | S, M)} &= \frac{P(G | S)}{P(\bar{G} | S)} \cdot \frac{P(M | G, S)}{P(M | \bar{G}, S)} \end{aligned}$$

## 3.3 Stochastische Unabhängigkeit

Frage: Wann sind 2 Ereignisse  $A, B$  unabhängig?

→ Illustration im Venn-Diagramm

Motivation über bedingte Wahrscheinlichkeiten:

Zwei Ereignisse  $A, B$  sind unabhängig, wenn

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

bzw.

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

# Definition und Folgerungen

Zwei Ereignisse  $A, B$  sind unabhängig, wenn

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

gilt.

Beachte: Voraussetzung  $P(B) > 0$  und  $P(A) > 0$  hier nicht nötig.

Folgerungen:

Sind  $A$  und  $B$  unabhängig, dann sind auch  $\bar{A}$  und  $B$ ,  $A$  und  $\bar{B}$  und auch  $\bar{A}$  und  $\bar{B}$  unabhängig. Beweis?

## Beispiel: Zweimaliges Würfeln

Ein fairer Würfel wird zweimal hintereinander geworfen. Sei

$A$ : “Beim 1. Würfelwurf eine Sechs”

$B$ : “Beim 2. Würfelwurf eine Sechs”

Bei jedem Würfelwurf ist die Grundgesamtheit  $\Omega = \{1, 2, 3, 4, 5, 6\}$ .

Nach Laplace gilt  $P(A) = P(B) = 1/6$ .

Bei “unabhängigem” Werfen gilt somit

$$P(A \cap B) = P(A) \cdot P(B) = 1/36$$

## Beispiel: Zweimaliges Würfeln mit Tricks

Angenommen der Würfelwerfer ist besonders am Werfen von einem Pasch interessiert. Er kann den zweiten Wurf ein wenig steuern und würfelt mit W'keit  $1/2$  das gleiche wie beim ersten Wurf. Die anderen Ergebnisse seien dann gleichverteilt mit W'keit  $0.1$  .

Dann ist zwar  $P(A) = 1/6$  und auch  $P(B) = 1/6$ , aber

$$P(A \cap B) = 1/12 > 1/36$$

Die Ereignisse  $A$  und  $B$  sind also abhängig, da

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

# Unabhängigkeit von mehr als zwei Ereignissen

Allgemeiner lässt sich die Unabhängigkeit von mehr als zwei Ereignissen definieren:

Die Ereignisse  $A_1, A_2, \dots, A_n$  sind (stochastisch) unabhängig, wenn für alle Teilmengen  $I \subset \{1, 2, \dots, n\}$  mit  $I = \{i_1, i_2, \dots, i_k\}$  gilt:

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k})$$

*Bemerkung:* Aus der paarweisen Unabhängigkeit folgt **nicht** die Unabhängigkeit von mehr als zwei Ereignissen.



## Beispiel zur paarweisen Unabhängigkeit

$$\begin{aligned}\Omega &= \{0, 1, 2, 3\} && \text{Laplace-Wahrscheinlichkeitsraum} \\ A_i &= \{0\} \cup \{i\} && \text{mit } i = 1, 2, 3\end{aligned}$$

(etwa einmaliges Ziehen aus einer Urne). Dann gilt:

$$P(A_i) = \frac{1}{2} \text{ und } P(A_i \cap A_j) = \frac{1}{4} = P(A_i) \cdot P(A_j) \text{ für alle } i \neq j.$$

Aber:

$$\begin{aligned}P(A_1 \cap A_2 \cap A_3) &= \frac{1}{4} \\ &\text{und} \\ P(A_1) \cdot P(A_2) \cdot P(A_3) &= \frac{1}{8}\end{aligned}$$

$A_1, A_2, A_3$  sind also **nicht** unabhängig.

# Bedingte Unabhängigkeit

Sei  $C$  ein beliebiges Ereignis mit  $P(C) > 0$ . Zwei Ereignisse  $A$  und  $B$  nennt man **bedingt unabhängig gegeben**  $C$ , wenn

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

gilt.

Man überzeugt sich leicht, dass weder aus unbedingter Unabhängigkeit bedingte Unabhängigkeit (bzgl. einem Ereignis  $C$ ), noch aus bedingter Unabhängigkeit bzgl. einem Ereignis  $C$  unbedingte Unabhängigkeit folgt.

## 3.4 Das Hardy-Weinberg-Gesetz

Population von diploiden Organismen; zwei Allele  $a$  und  $b$

Drei Genotypen  $aa$ ,  $ab$  und  $bb$ , Wahrscheinlichkeitsverteilung  $p_{aa}$ ,  $p_{ab}$ ,  $p_{bb}$  sei **unabhängig** vom Geschlecht

Unter bestimmten weiteren Voraussetzungen (“**zufällige**” Paarung von Individuen, keine Selektion, keine Mutation etc.) gilt, dass die Wahrscheinlichkeitsverteilung der drei Genotypen über Generationen hinweg konstant bleibt.

# Das Hardy-Weinberg-Gleichgewicht

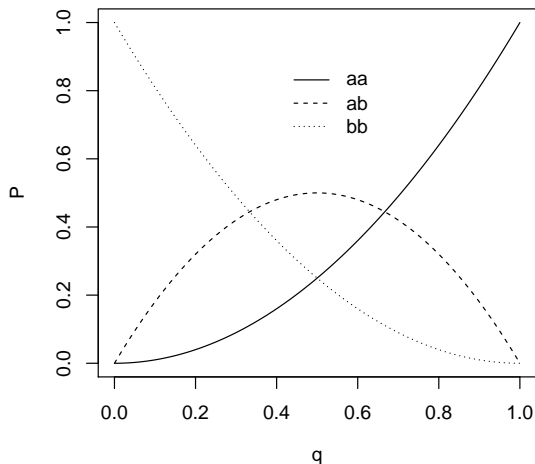
Sei  $x \in \{aa, ab, bb\}$  der Genotyp eines zufällig ausgewählten Individuums. Ist die Population im H-W-Gleichgewicht, so gilt:

$$p_x = \begin{cases} q^2 & \text{für } x = aa \\ 2q(1 - q) & \text{für } x = ab \\ (1 - q)^2 & \text{für } x = bb \end{cases}$$

mit  $q \in [0, 1]$ . Der Parameter  $q$  ist die Häufigkeit des Allels  $a$ .

Beachte: Im Allgemeinen beschreiben **zwei** Parameter eine diskrete Wahrscheinlichkeitsverteilung mit **drei** möglichen Ausprägungen, das H-W-Gleichgewicht wird jedoch nur von **einem** Parameter beschrieben.

# Das Hardy-Weinberg-Gleichgewicht



Hardy-Weinberg-Gleichgewicht für die Genotypen  $aa$ ,  $bb$  und  $ab$ .

# Beweisskizze

Starte mit beliebigen Wahrscheinlichkeiten  $p_{aa}$ ,  $p_{ab}$  und  $p_{bb}$   
(mit  $p_{aa} + p_{ab} + p_{bb} = 1$ ).

Berechne die bedingten Wahrscheinlichkeiten für das Auftreten eines bestimmten Genotyps eines Nachkommen in Abhängigkeit vom Genotyp der Eltern.

Satz der totalen Wahrscheinlichkeit liefert H-W-Gleichgewicht mit  $q = p_{aa} + p_{ab}/2$ .

Neustart im H-W-Gleichgewicht liefert wieder H-W-Gleichgewicht!

## Anwendungsbeispiel

Genotyp  $bb$  verursacht bestimmte Krankheit,  $aa$  und  $ab$  sind aber äußerlich nicht erkennbar.

Aus Kenntnis der relativen Häufigkeit/Wahrscheinlichkeit  $p_{bb}$  lassen sich (unter der Annahme, dass die Population im H-W-Gleichgewicht ist) die anderen Wahrscheinlichkeiten berechnen:

$$\begin{aligned}p_{aa} &= (1 - \sqrt{p_{bb}})^2 \\p_{ab} &= 2(1 - \sqrt{p_{bb}})\sqrt{p_{bb}}\end{aligned}$$

Beispiel:  $p_{bb} = 0.0001 \rightsquigarrow p_{aa} = 0.9801$  und  $p_{ab} = 0.0198$

# Das Hardy-Weinberg-Ungleichgewicht

Allgemeinere Formulierung mit zwei Parametern:

$$p_x = \begin{cases} q^2 + d & \text{für } x = aa \\ 2q(1 - q) - 2d & \text{für } x = ab \\ (1 - q)^2 + d & \text{für } x = bb \end{cases}$$

mit Ungleichgewichtskoeffizienten  $d$ .

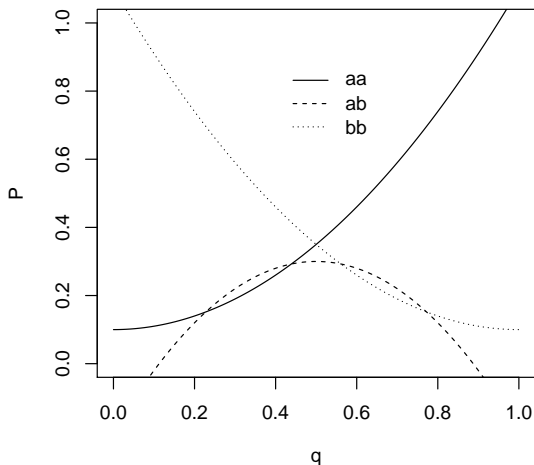
Problem: Beschränkter Wertebereich für  $q$  und  $d$

Später:

- Statistische Schätzung von  $q$  und  $d$
- Statistischer Test auf "Nullhypothese"  $H_0 : d = 0$



# Das Hardy-Weinberg-Ungleichgewicht ( $d = 0.1$ )



Hardy-Weinberg-Ungleichgewicht mit  $d = 0.1$  für die Genotypen  $aa$ ,  $bb$  und  $ab$ .