Datenanalyse und Statistik
4. Wahrscheinlichkeitsrechnung

Marcus Hudec

It is remarkable that a science which began with the consideration of games of chance should have become the most important object of human knowledge ...

The most important questions of life are, for the most part, really only problems of probability.

Pierre Simon de Laplace
Theorie Analytique des Probabilities, 1812

Zufall und Erkenntnis

- Klassischer wissenschaftstheoretischer Fortschrittsglauben:
  Unsicherheit und Unschärfe (Zufall) können durch fortschreitende wissenschaftliche Erkenntnis reduziert werden
- Gegensätzliche Standpunkte:
  - Demokrit:
    Die Natur ist in ihrer Grundlage streng determiniert
    Zufälliges entspricht dem *Nichterkannten*
  - Epikur:
    Der Zufall ist immanenter Bestandteil der Natur und der Erscheinungen unserer Welt

Zweifel am deterministischen Weltbild

Das Gewebe dieser Welt ist aus Notwendigkeit **und** Zufall gebildet; die Vernunft des Menschen stellt sich zwischen beide und weiß sie zu beherrschen; sie behandelt das Notwendige als den Grund ihres Daseins; das Zufällige weiß sie zu lenken, zu leiten und zu nutzen, ...

Johann Wolfgang von Goethe

Dissertation von Karl Marx (1841) beschäftigte sich über den Unterschied in der Naturphilosophie Demokrits und Epikurs

Paradigmenwechsel im 20.Jahrhundert

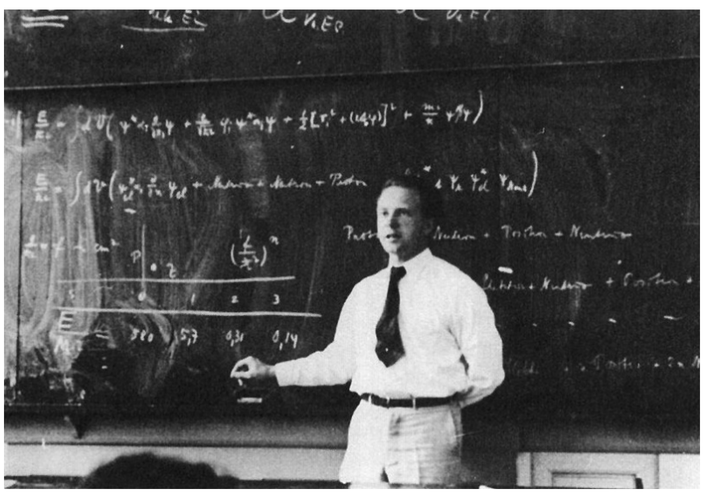
- Erkenntnisse der theoretischen Physik
  "Existenz der Wahrscheinlichkeit in der Natur"
- Thermodynamik (Boltzmann)
  - Bewegungen von Molekülen werden nicht durch deterministische Gesetze der Newtonschen Mechanik sondern durch **Wahrscheinlichkeitsgesetze** gesteuert
- Quantenmechanik (Heisenberg)
  - Radioaktivität: freie Neutronen zerfallen **zufällig**
  - Ihre Anzahl gehorcht jedoch einem bestimmten **Gesetz**
- Kalkül des Zufalls ~ Wahrscheinlichkeitsrechnung
- Einstein: "Gott würfelt nicht"

Unschärferelation der Quantenmechanik

- Grundpostulat der modernen Physik:

  Die Elementarvorgänge im materiellen Geschehen entziehen sich grundsätzlich einer exakten raum-zeitlichen Darstellung. Voraussagen über Ort und Geschwindigkeit der kleinsten Partikeln haben immer nur den Charakter von Wahrscheinlichkeitsaussagen.

  Werner Heisenberg



Wahrscheinlichkeit und Biologie

- Theorie der Evolution: umweltbedingt kommt es zu zufälligen Änderungen des Genotyps
  - Crossing Over
  - Mutation
- Evolution ist keine Entwicklung von primitiven zu komplexen Lebensformen sondern eine Entwicklung von weniger angepassten Arten zu besser angepassten
- Die Rolle des Zufalls für moderne Biologie und Genetik
  ==> M. Eigen "Das Spiel“
- Die Rolle des Zufalls für moderne Physik und Biologie
  ==> L. Tarassow „Wie der Zufall will? - Vom Wesen der Wahrscheinlichkeit“

Bildungsauftrag: Wahrscheinlichkeiten verstehen

Unsere Gesellschaft muss stärker lernen, Risiken zu bewerten, ganz generell gesprochen. Das Leben mit der Chance und dem Risiko ist ein wichtiges gesellschaftliches Problem.

Ich finde es in einer komplexer werdenden Welt auch wichtig, Kinder bereits frühzeitig an solche Abwägungen heranzuführen, die sie später immer wieder vornehmen müssen.

Im Kindergarten und in der Schule können Kinder spielerisch lernen, was Wahrscheinlichkeit und Risiko bedeuten.

Angela Merkel in einem Interview 12/2006

Definitionen - 1

- Kalkulatorische Erfassung des Phänomen Zufalls
- Zufallsvorgang

  Vorgang mit ungewissem Ausgang

  Es gibt mehrere mögliche Ergebnisse des Vorgangs

  Das Ergebnis bei einer Durchführung ist nicht mit Sicherheit vorhersehbar
- Zufallsexperiment

  Der Vorgang ist unter gleichen Randbedingungen beliebig oft wiederholbar

Bestimmung der Wahrscheinlichkeit

- Das Prinzip "günstige Fälle" dividiert durch alle "möglichen Fälle" im Rahmen eines Modells gleicher Wahrscheinlichkeiten für die Einzelereignisse
  Laplace klassischer Wahrscheinlichkeitsbegriff
  einfache ideale Glücksspiele
- Statistische Wahrscheinlichkeit für das Eintreten eines Ereignisses ist jener Wert, bei dem sich die relative Häufigkeit bei einer wachsenden Zahl von Versuchswiederholungen stabilisiert
  Mises frequentistischer Wahrscheinlichkeitsbegriff
  alle Zufallsexperimente
- Das Konzept subjektiver Wahrscheinlichkeiten
  de Finetti subjektiver Wahrscheinlichkeitsbegriff
  alle Zufallsvorgänge

Beispiele:

- Laplace: Teilnahme an einer Lotterie
  Anzahl der Möglichkeiten (verschiedene Lose) sei 100.000
  Anzahl der „günstigen“ Fälle (gewinnende Lose) sei 5.000
  Ich besitze ein Los ➔ Wahrscheinlichkeit = 5%, dass ich einen Gewinn erhalte
- Voraussetzung: Echte Zufallsauswahl – alle Lose haben die gleiche Chance gezogen zu werden.

Beispiele:

- Mises

  Wiederholte Durchführung eines Zufallsexperiments (z.B. Würfelwurf) und Ermittlung der relativen Häufigkeit, wie oft ein bestimmtes Ereignis auftritt.
  Verwende diese relative Häufigkeit, als Maß für die Wahrscheinlichkeit
- Voraussetzung: Wiederholbarkeit muss möglich sein
- De Finetti

  Subjektive Einschätzung kann Zweifel an Zufälligkeit berücksichtigen
  Bei Lotterie wenig relevant aber z.B. bei Sportwetten

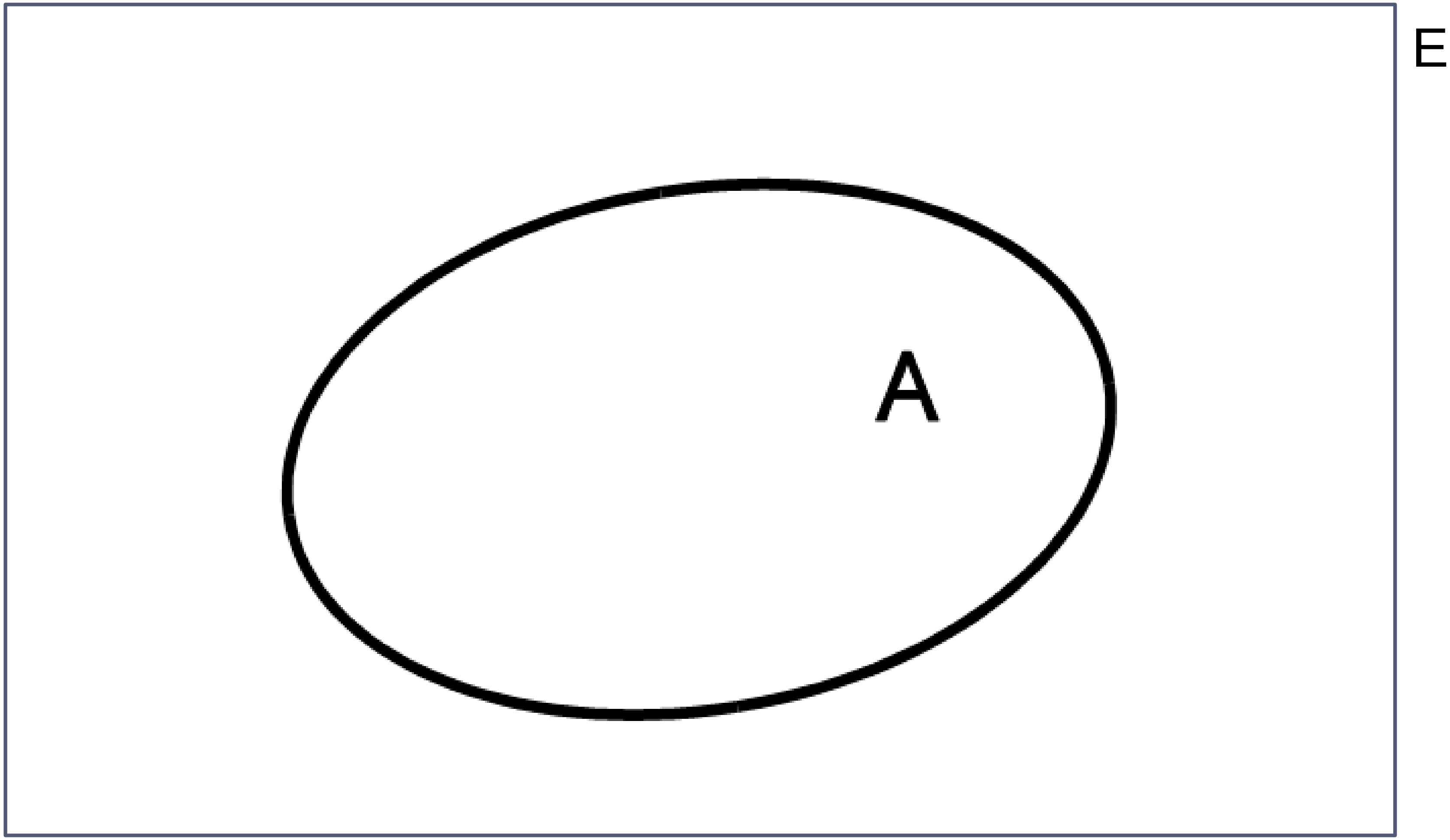
Was ist Wahrscheinlichkeit?

- Wahrscheinlichkeit ist der Grad der Gewissheit für das Eintreten von Ereignissen
- Je nach Sichtweise können Wahrscheinlichkeiten hergeleitet werden, als
  - ein Maß für die Unsicherheit konkreter Ereignisse aufgrund der möglichen Alternativen (Laplace: „Günstige durch Mögliche")
  - ein Maß für die relative Häufigkeit des Auftretens von Ereignissen
  - ein Maß für den Grad an persönlicher Überzeugung

Definitionen

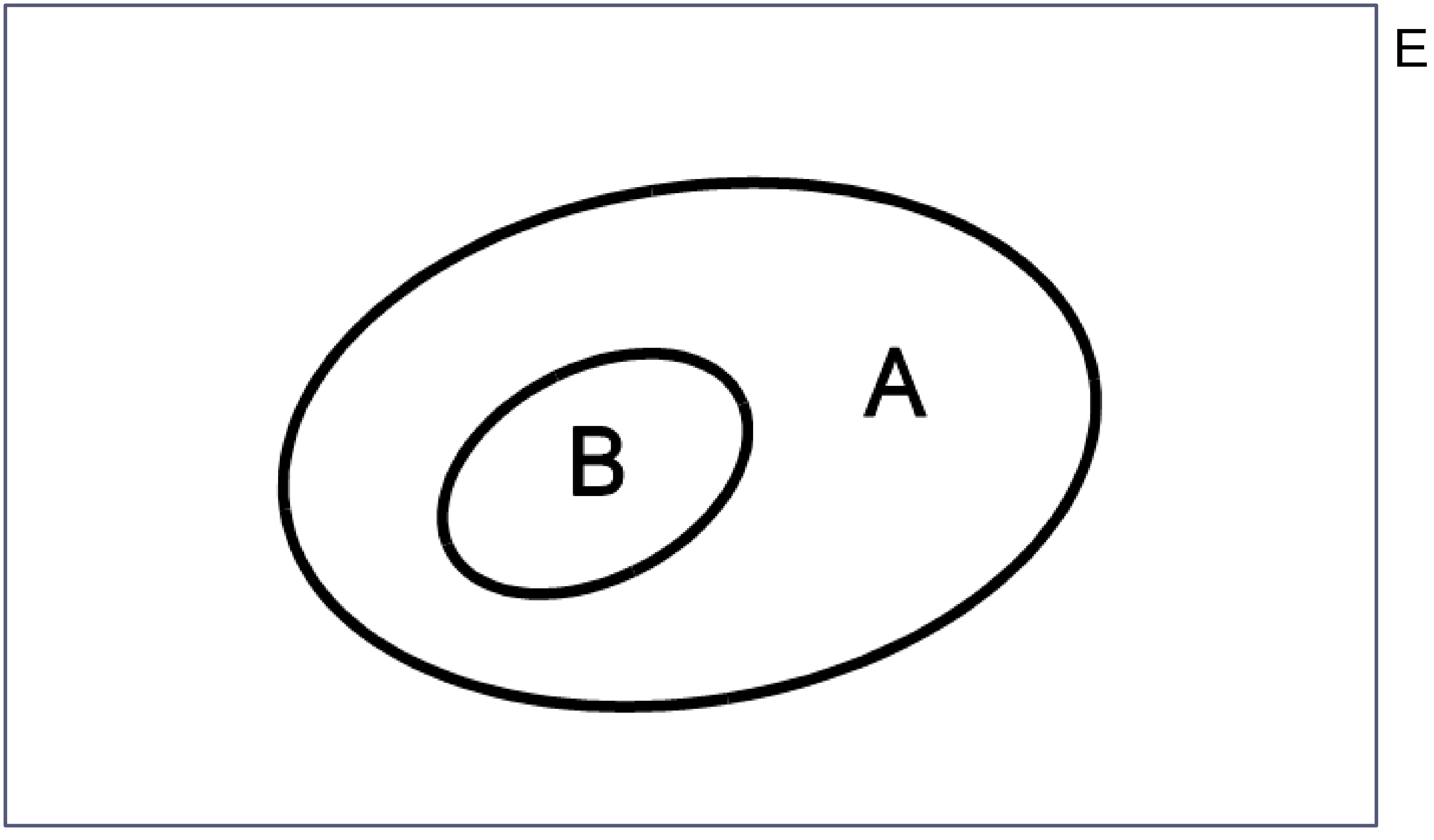
- Ein möglicher Ausgang (mögliches Ergebnis) eines Zufallsexperiments wird als Elementarereignis bezeichnet.
  $e1, e2, \ldots$
- Die Menge aller Elementarereignisse eines Zufallsexperiments wird als Ergebnismenge oder Stichprobenraum bezeichnet.
  $E=\{e1, e2, \ldots\}$
- Eine Teilmenge A der Ergebnismenge heißt (zusammengesetztes) Ereignis. Ein Ereignis A tritt ein, wenn ein Ergebnis beobachtet wird, das zu A gehört.
  $A \subset E$

Ereignis A als Teilmenge von E: $A \subset E$



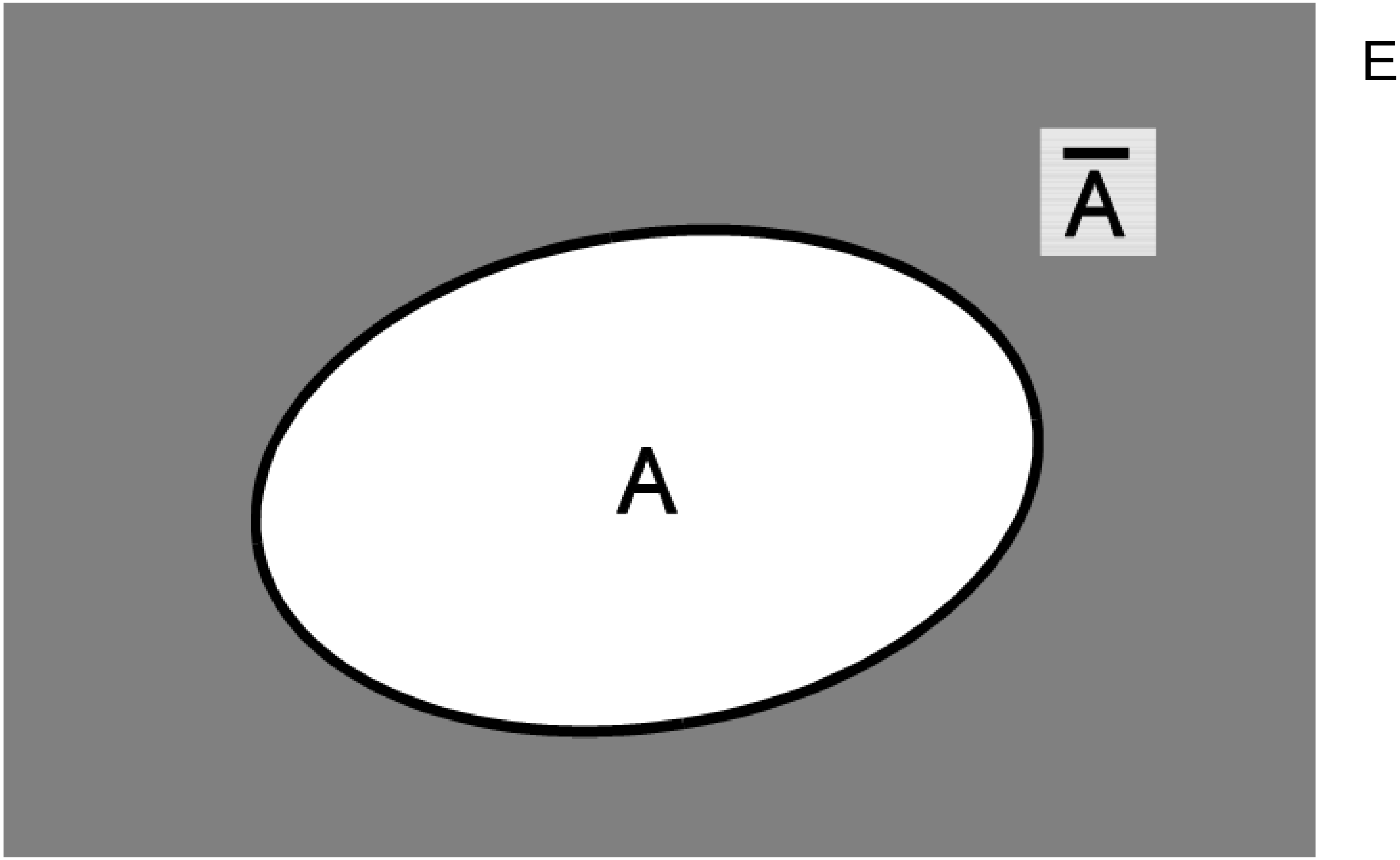
A ist Teil von E

B ist ein Teilereignis von A : $B \subset A \subset E$



B ist Teil von A

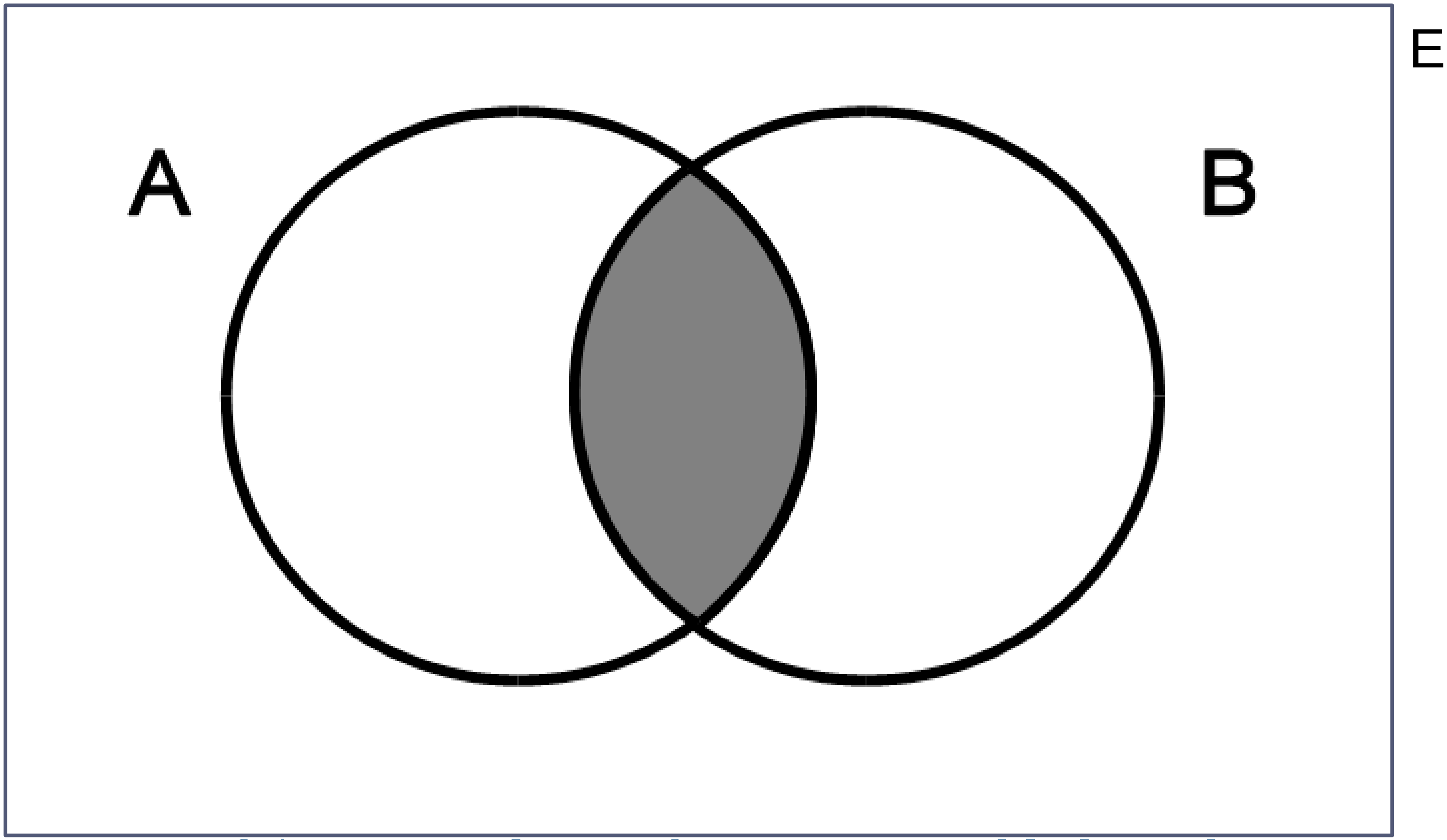
Komplementärereignis (Inversion)



$A \cup A' = E$

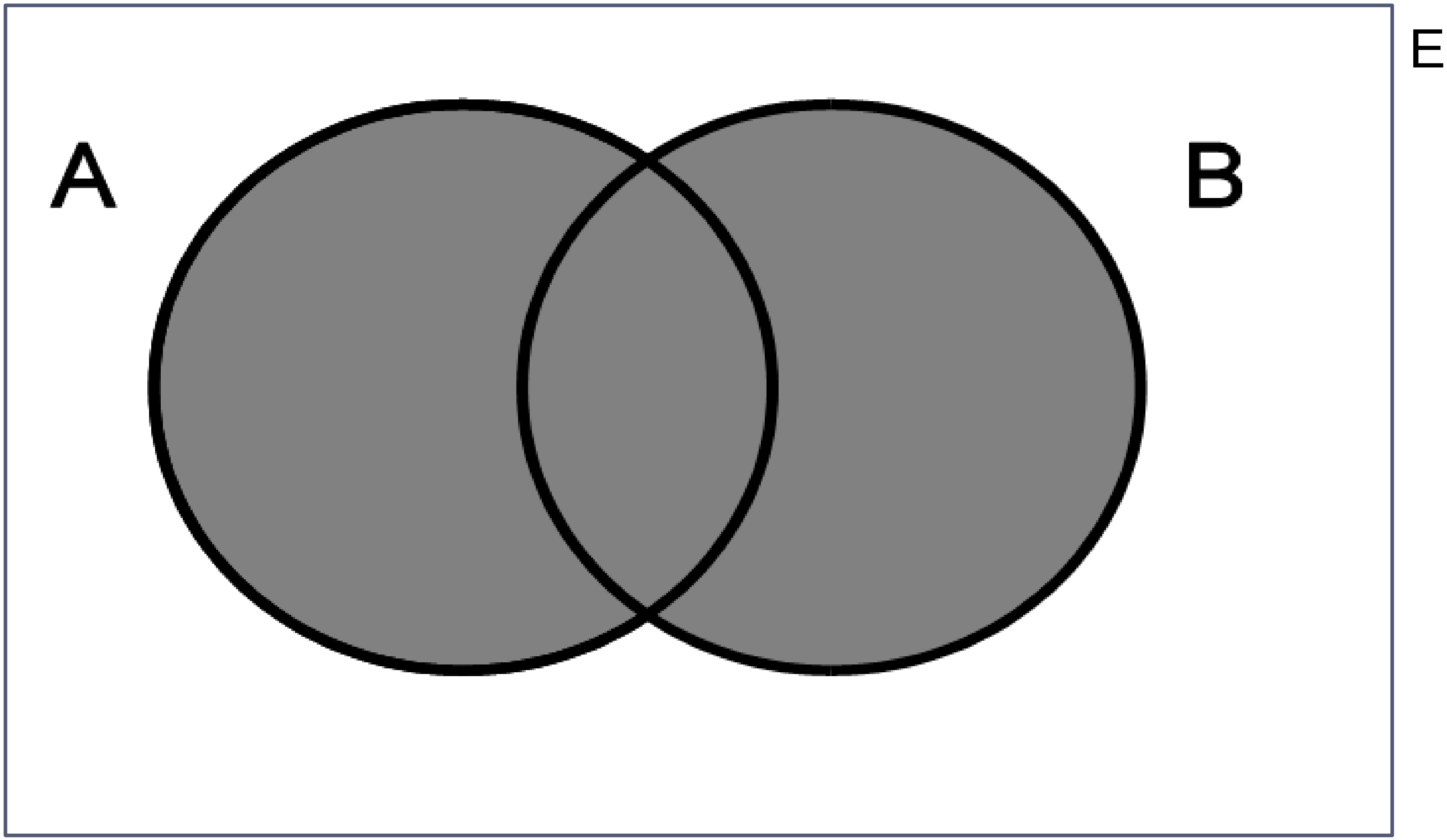
Nicht A

Durchschnitt von A und B „UND“



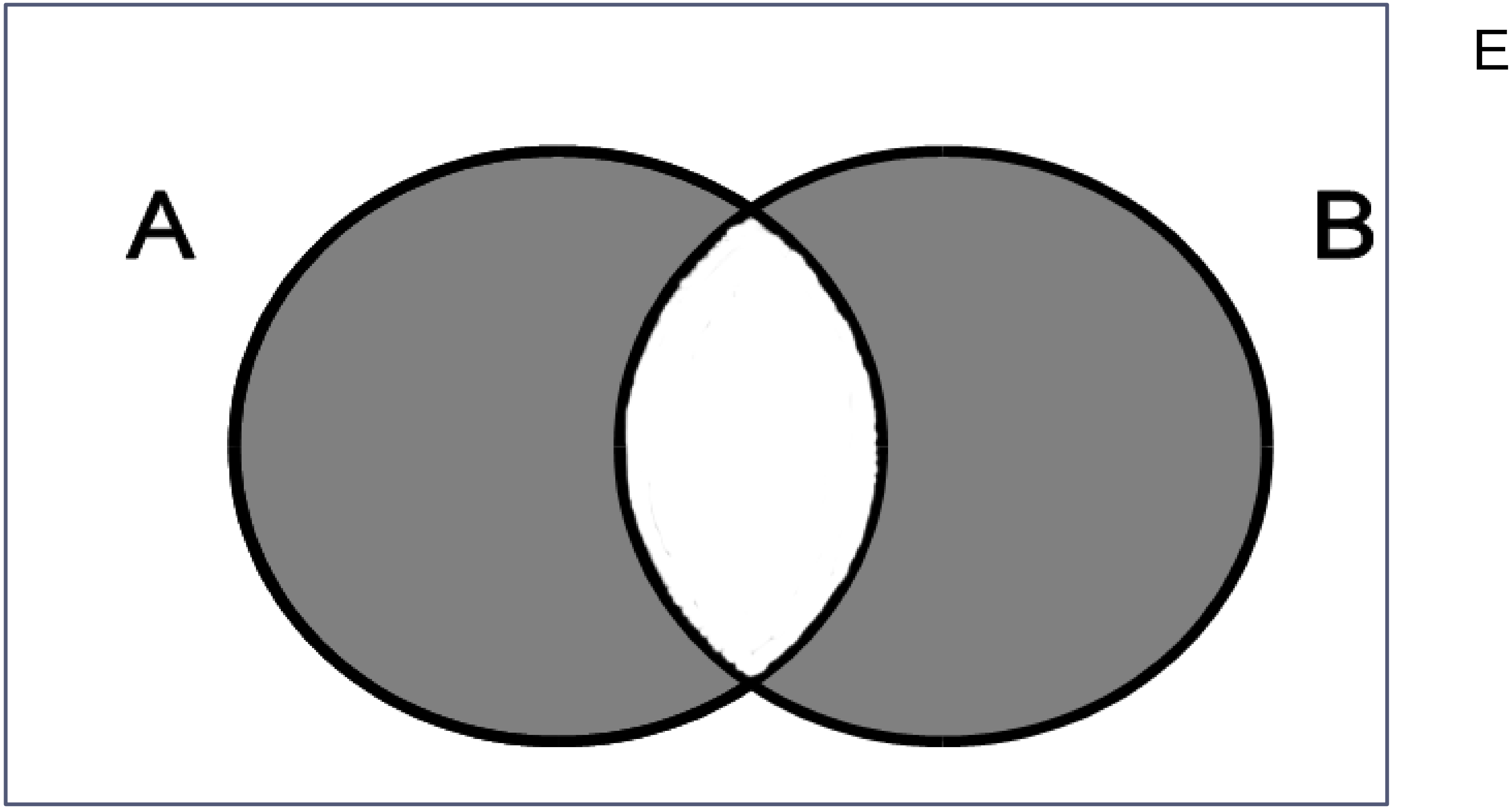
$A \cap B = \{e | \; e \in A \text{ und } e \in B\}$ sowohl als auch

Vereinigung von A und B „ODER“



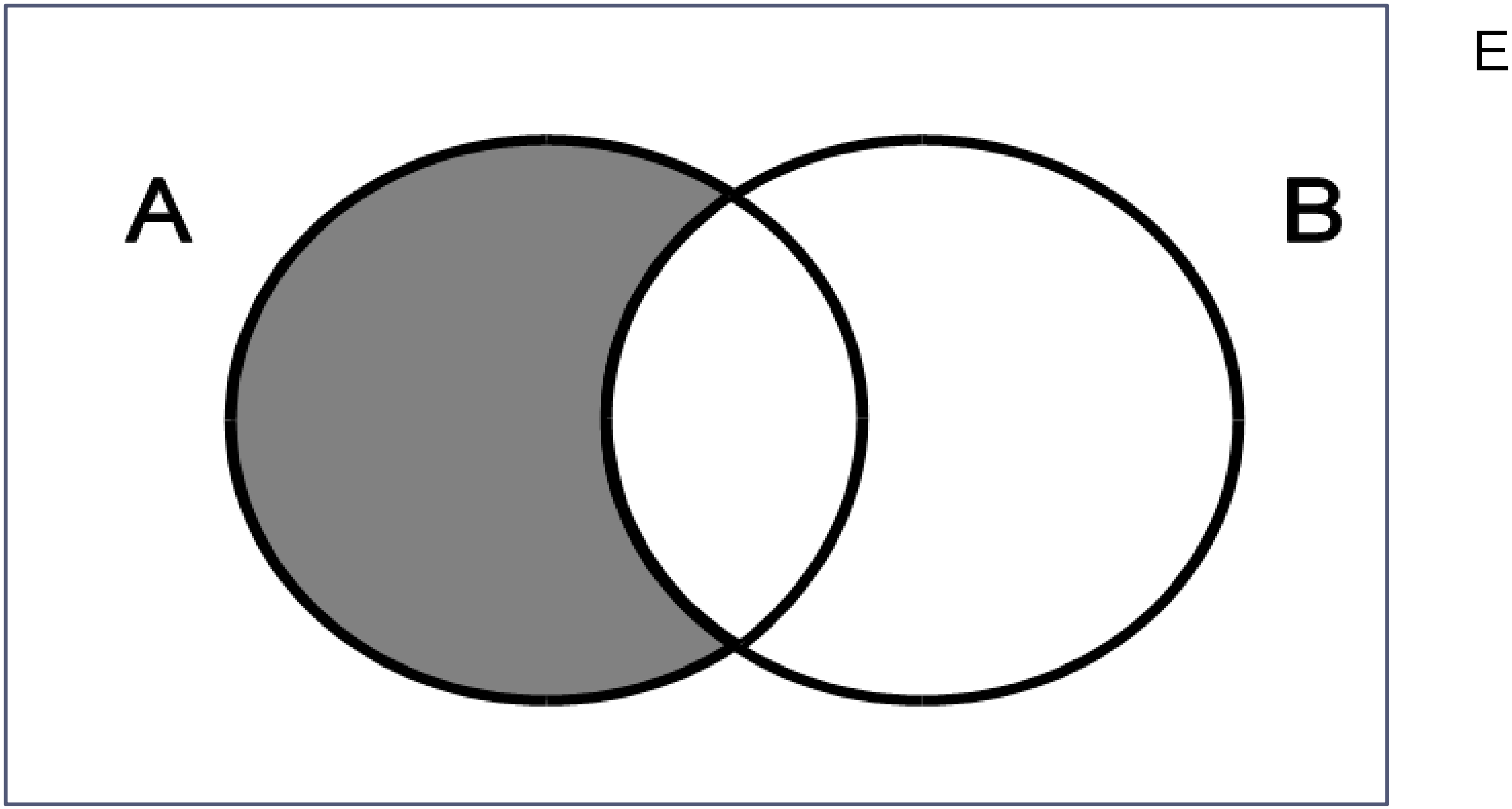
$A \cup B = \{e \mid e \in A$ oder $e \in B\}$ oder (aber nicht exklusiv)

Exklusives Oder



$(A \cap B') \cup (A' \cap B) = (A \cup B) \setminus (A \cap B)$ entweder oder

Differenzmenge: A ohne B



$A \setminus B = \{e | \, e \in A$ und nicht $e \in B\}$ ohne

Beachte: $E \setminus A = A'$

Wahrscheinlichkeitsmaß

- Ein Wahrscheinlichkeitsmaß P ist eine Abbildung, die allen möglichen Elementarereignissen eines Zufallsexperiments eine Zahl zuordnet und dabei den Axiomen von Kolmogorov genügt.
- Die Wahrscheinlichkeit eines beliebigen Ereignisses A ergibt sich durch die Summe der Wahrscheinlichkeiten jener Elementarereignisse, die in A enthalten sind.

• Positivität	$0 \leq P(A) \leq 1$
• Normierung	$P(E) = 1$
• Additivität	$P(A_1 \cup A_2) = P(A_1) + P(A_2)$ falls $A_1 \cap A_2 = \emptyset$

Die Axiome in Worten

- Für jedes Ereignis ist die Wahrscheinlichkeit eine reelle Zahl zwischen 0 und 1.
- Das sichere Ereignis hat die Wahrscheinlichkeit 1.
- Das unmögliche Ereignis hat die Wahrscheinlichkeit null.
- Die Wahrscheinlichkeit einer Vereinigung abzählbar vieler unterschiedlicher Elementarereignisse entspricht der Summe der Wahrscheinlichkeiten der einzelnen Elementarereignisse.

- Die Summe der Wahrscheinlichkeit von 2 komplementären Ereignissen ist 1.
  $P(A \cup A')=P(E)=1 \Rightarrow \quad P(A)+P(A')=1 \Rightarrow P(A')=1-P(A)$

Aus den Axiomen lassen sich Theoreme ableiten

- Die Summe der Wahrscheinlichkeit von 2 komplementären Ereignissen ist 1.
  <u>Beweis</u>:
  $A \cup A' = E$ $\quad P(E)=1 \quad P(A \cup A')=1$
  $P(A)+P(A')=1 \quad P(A')=1-P(A)$
- Die Wahrscheinlichkeit der Vereinigung von 2 beliebigen Ereignissen A und B lässt sich wie folgt schreiben:
  $P(A \cup B) = P(A)+P(B)-P(A \cap B)$
  <u>Beweis</u>:
  $A \cup B = (A \cap B') \cup (A \cap B) \cup (A' \cap B)$
  $P(A \cup B) = P(A \cap B') + P(A \cap B) + P(A' \cap B)$
  $P(A \cup B) + P(A \cap B) = P(A \cap B') + P(A \cap B) + P(A' \cap B) + P(A \cap B)$
  $P(A \cup B) + P(A \cap B) = P(A) + P(B)$
  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ q.e.d.

Einfache Beispiele

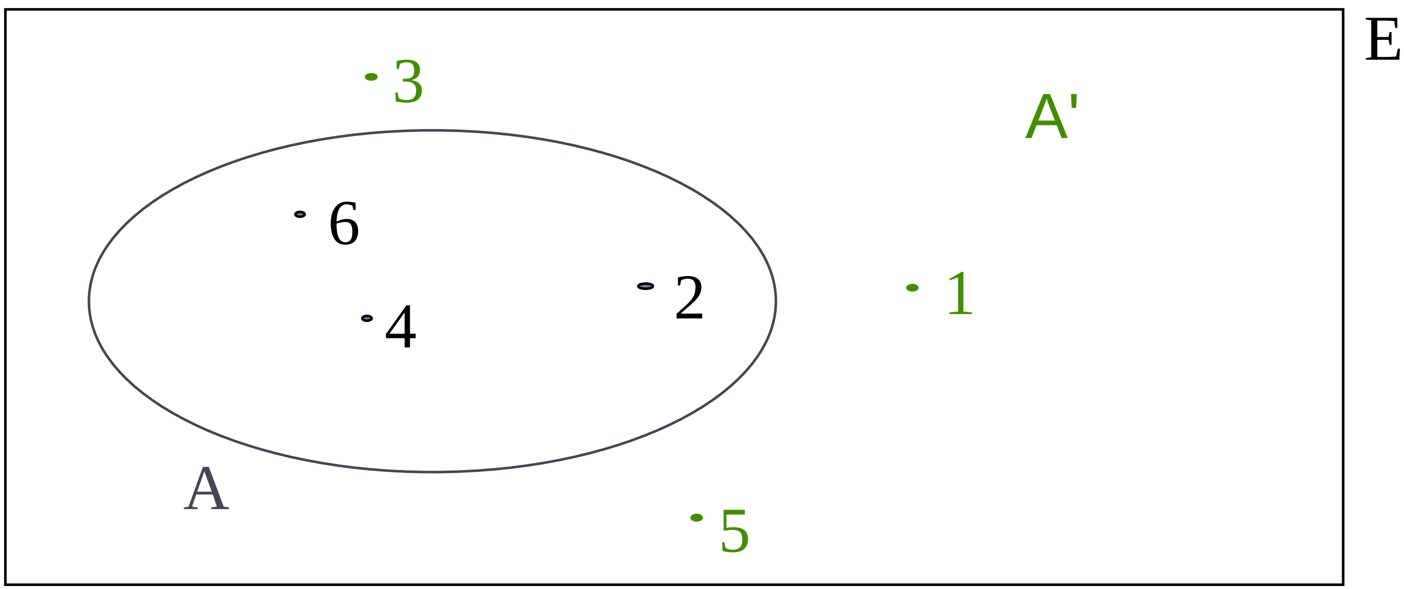
- Eine Münze wird einmal geworfen
  - Elementarereignisse: Kopf, Adler
  - $P(K) = 1/2$
  - $P(A) = 1/2$
  - $E = K \cup A$ $P(K \cup A) = 1/2 + 1/2 = 1$
- Eine Münze wird zweimal geworfen
  - 4 Elementarereignisse: KK, KA, AK, AA
  - $P(KK) = P(KA) = P(AK) = P(AA) = 1/4$
  - Zusammengesetztes Ereignis:
  - A = "zwei gleiche Ergebnisse"
  - $A = \{KK, AA\} = P(A) = 1/4 + 1/4 = 1/2$

Beispiel: Würfelwurf

- mögliche Elementarereignisse 1, 2, 3, 4, 5, 6
- $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$
- $P(E) = 1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 1$
- (Zusammengesetzte) Ereignisse
- A ... gerade Augenzahl $A=\{2, 4, 6\}$
- $P(A) = 1/6 + 1/6 + 1/6 = 3/6$
- B ... Augenzahl < 3 $B=\{1, 2\}$
- $P(B) = 1/6 + 1/6 = 2/6$

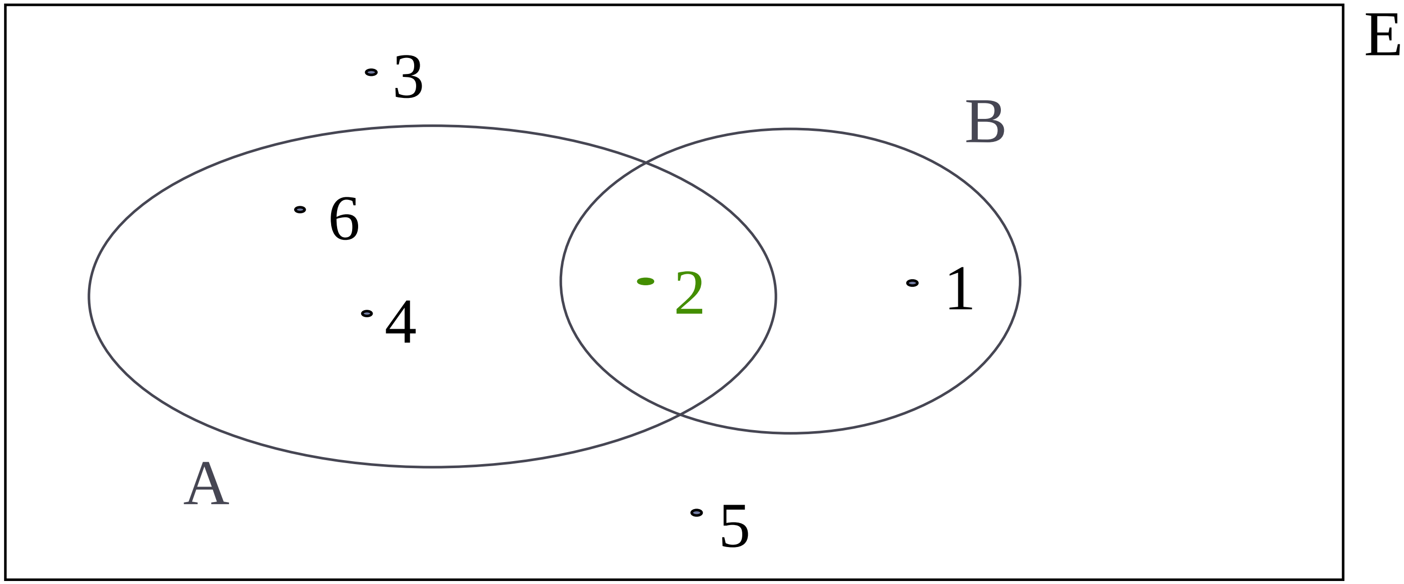
Komplementbildung

- A' ... Nicht A ➔ ungerade Augenzahl
- A' = {1, 3, 5}
- A' = E \ A
- P(A') = P(E) - P(A) = 1 - P(A) = 1 - 3/6 = 3/6



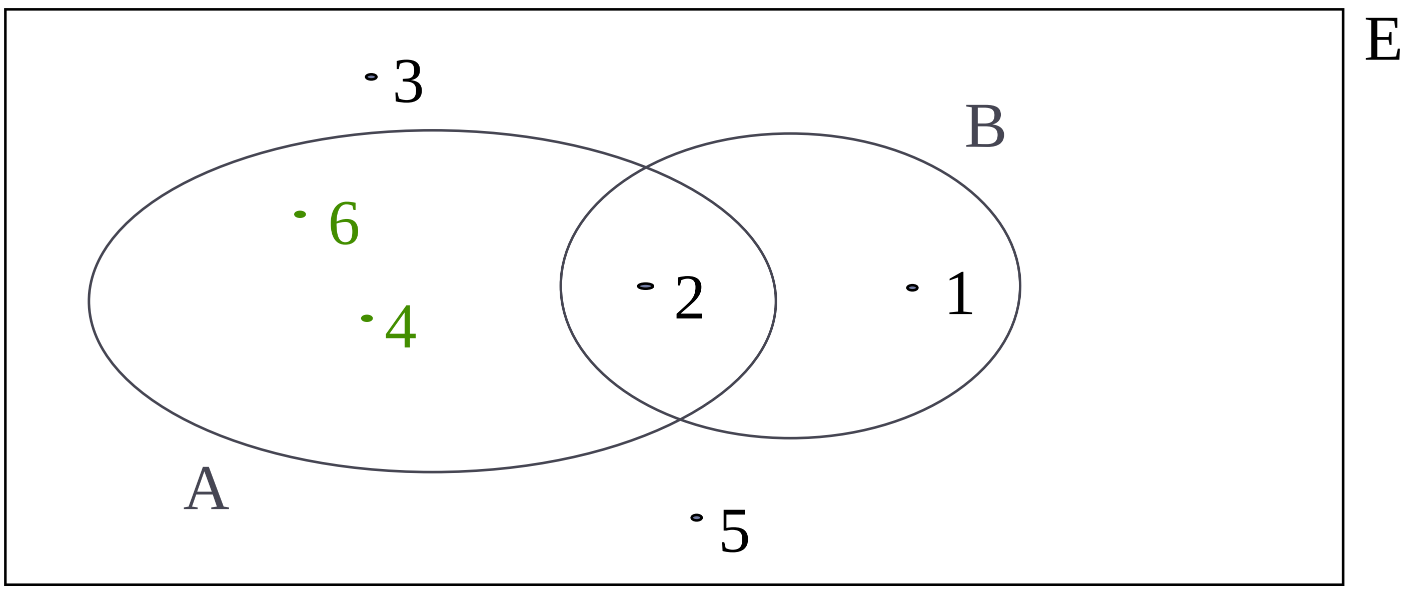
Durchschnittsbildung

- $A \cap B$ ... gerade Augenzahl <u>und</u> Augenzahl kleiner 3
- $A \cap B = \{2\}$
- $P(A \cap B) = 1/6$



Differenzmenge

- $A \setminus B$ ...   gerade Augenzahl ohne den Augenzahlen kleiner 3
- $A \setminus B = \{4, 6\}$
- $P(A \setminus B) = P(A) - P(A \cap B) = 2/6$



Beispiel zur Ereignisalgebra

		Alter		
		<=25	>25	Gesamt
Sex	männlich	20	30	50
	weiblich	40	10	50
	Gesamt	60	40	100

2 Merkmale: Geschlecht, Alter

A...(Geschlecht = männlich) .... n(A)=50 ... h(A)=0,50

A‘...(Geschlecht = nicht männlich)
(Geschlecht = weiblich) .... n(A‘)=50 ... h(A‘)=0,50

B...(Alter <= 25) .... n(B)=60 ... h(B)=0,60

B‘...(Alter > 25) .... n(B‘)=40 ... h(B‘)=0,40

Beispiel zur Ereignisalgebra

		Alter		
		<=25	>25	Gesamt
Sex	männlich	**20**	30	50
	weiblich	40	10	50
	Gesamt	60	40	100

Durchschnitt

$A \cap B$ ~ (Geschlecht = männlich) und (Alter <= 25)

$n(A \cap B) = 20$ ... $h(A \cap B) = 0{,}20$

Man beachte $P(A \cap B)$ ist offensichtlich nicht immer gleich dem Produkt $P(A)$ mal $P(B)$! (Details im nächsten Kapitel)

Beispiel zur Ereignisalgebra

		Alter		
		<=25	>25	Gesamt
Sex	männlich	**20**	**30**	50
	weiblich	**40**	10	50
	Gesamt	60	40	100

<u>Vereinigung</u>

$A \cup B$ ~ (Geschlecht = männlich) oder (Alter <= 25)

$n(A \cup B) = 90$ (ergibt sich aus 20+30+40)

$h(A \cup B) = 0{,}90$

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

$0{,}90 = 0{,}50 + 0{,}60 - 0{,}20$

Beispiel zur Ereignisalgebra

		Alter		
		<=25	>25	Gesamt
Sex	männlich	20	**30**	50
	weiblich	40	10	50
	Gesamt	60	40	100

<u>Differenzmenge</u>

$A \setminus B$ ~ (Geschlecht = männlich) ohne (Alter <= 25)

$n(A \setminus B) = 30 \quad ... h(A \setminus B) = 0{,}30$

Man beachte: $P(A \setminus B) = P(A) - P(A \cap B)$

$0{,}30 \quad = 0{,}50 - 0{,}20$

ZUSAMMENFASSUNG

Manchmal gilt: $P(A \cap B) = P(A) \,.\, P(B)$
Manchmal gilt: $P(A \cap B) \neq P(A) \,.\, P(B)$
➔ Aufklärung darüber Unabhängigkeit (nächste Lektion)

Immer gilt:

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

Die Wahrscheinlichkeit eines durch Vereinigung definierten Ereignisses ist die Summe der Wahrscheinlichkeiten der beiden Ereignisse minus der Wahrscheinlichkeit des Durchschnitts (Doppelzählung!)

Immer gilt:

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

Einschub
Grundlagen der Kombinatorik

Univ.Prof. Dr. Marcus Hudec

Kombinatorik

- Wie komme ich zu den relevanten Anzahlen von möglichen und günstigen Fällen?
  - Vollständige Enumeration aller Fälle
  - Mathematische Berechnung
  - Kombinatorisches Kalkül
- **Kombinatorik** ist jenes Teilgebiet der Mathematik, das sich mit der Bestimmung der Zahl möglicher Anordnungen oder Auswahlen von
  - unterscheidbaren oder nicht unterscheidbaren Objekten
  - mit oder ohne Beachtung der Reihenfolge
- beschäftigt. Sie bildet ein wichtiges Hilfsmittel der Wahrscheinlichkeitsrechnung

Überblick über 4 Stichprobenmodelle

	Mit Zurücklegen	Ohne Zurücklegen
Mit Berücksichtigung der Anordnung (Variationen)	$N^* = N^n$ I	$N^* = N\,(N-1)\ \ldots\ (N-(n-1))$ II falls n=N: N! (Permutation)
Ohne Berücksichtigung der Anordnung (Kombinationen)	$N^* = \binom{N+n-1}{n}$ III	$N^* = N\,(N-1)\ \ldots\ (N-(n-1))/n!$ $N^* = \binom{N}{n}$ IV

N*…Anzahl der unterscheidbaren Stichproben
N … Umfang der Grundgesamtheit
n … Anzahl der Elemente in der Stichprobe

Modelle mit Berücksichtigung der Anordnung

- **Ziehen mit Zurücklegen** (Situation I)
  - n Ziehungen aus einer Gesamtheit vom Umfang N
  - Jede Ziehung erfolgt von der ursprünglichen Grundgesamtheit;
  - das Ergebnis eines Ziehungsschrittes beeinflusst nicht das Ergebnis des nächsten Schrittes;
  - In der Stichprobe können Objekte mehrfach vorkommen

Anzahl der möglichen Stichproben:

$$N^* = N^n$$

Beispiele für Variation mit Zurücklegen (I)

- Wie viele verschiedene Zustände kann man mit 8 binären Informationseinheiten (0/1) [mit 8 Bit] darstellen?
  $N=2$ $n=8$
  Ziehen mit Zurücklegen
  mit Berücksichtigung der Anordnung
  $2^8 = 256$

Beispiele für Variationen mit Zurücklegen (I)

- Wie viele verschiedene Wörter bestehend aus 4-Buchstaben kann man bilden?
  $N=26 \qquad n=4$
  Ziehen mit Zurücklegen
  mit Berücksichtigung der Anordnung
  $26^4 = 456.976$
- Wie viele verschiedene sechsstellige Zahlen kann man aus den Ziffern 0 bis 9 bilden? (vgl. Joker beim Lotto)
  $N=10 \qquad n=6$
  Ziehen mit Zurücklegen
  mit Berücksichtigung der Anordnung
  $10^6 = 1.000.000$

Modelle mit Berücksichtigung der Anordnung

- **Ziehen ohne Zurücklegen (Situation II)**
  - Jede einzelne Ziehung erfolgt von einer verringerten Grundgesamtheit;
  - das Ergebnis eines Ziehungsschrittes beeinflusst das Ergebnis des nächsten Schrittes;
  - In der Stichprobe kann ein Objekt nur maximal einmal vorkommen

Anzahl der möglichen Stichproben:

$$N^* = N\,(N-1)\,(N-2)\,\ldots\,(N-(n-1))$$

$$N^* = \frac{N!}{(N-n)!}$$

Beispiele für Variation ohne Zurücklegen (II)

- Wie viele Möglichkeiten gibt es 3 Kandidaten aus einer Liste von 5 auszuwählen und sie dabei zu reihen?
  $N=5 \qquad n=3$
  Ziehen ohne Zurücklegen
  mit Berücksichtigung der Anordnung
  Situation II
  5*4*3 = 60

Permutationen

- Spezialfall: $n=N$
- d.h. wir wählen alle Elemente aus
- N^*~ Anzahl von Anordnungen von N unterscheidbaren Objekten
- $N^*=N(N-1)(N-2)...(N-(n-1))=N(N-1)...1=N!$
- N-Fakultät
- Beachte: $0!=1$

Beispiele:

- Wie viele Möglichkeiten gibt es 3 Kandidaten zu reihen?
  N=3 n=3
  Ziehen ohne Zurücklegen
  mit Berücksichtigung der Anordnung
  $3! = 6$
- Permutation
- abc acb
- bac bca
- cab cba

Modelle ohne Berücksichtigung der Anordnung

- Ziehen ohne Zurücklegen (Situation IV)
  - Jede einzelne Ziehung erfolgt von einer verringerten Grundgesamtheit;
  - das Ergebnis eines Ziehungsschrittes beeinflusst das Ergebnis des nächsten Schrittes;
  - In der Stichprobe kann ein Objekt nur maximal einmal vorkommen
  - **Vorgang entspricht der Teilmengenbildung**

Anzahl der möglichen Stichproben:

$$N^* = N\,(N-1)\,(N-2)\,\ldots\,(N-(n-1))/n!$$

Vereinfachte Schreibweise

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} = \binom{N}{N-n} = \frac{N \cdot (N-1) \cdot \ldots \cdot (N-n+1)}{n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1}$$

Binomialkoeffizient

$$\binom{N}{0} = 1 \qquad \binom{N}{1} = N \qquad \binom{N}{N-1} = N$$

$$\sum_{i=0}^{N} \binom{N}{i} = 2^n$$

Excel-Funktion: KOMBINATIONEN(n; k)
R-Function: choose(n, k)

Beispiele:

- Wie viele Möglichkeiten gibt es 3 Kandidaten aus einer Liste von 5 (a,b,c,d,e) auszuwählen, wobei Ihre Reihung irrelevant ist?
  N=5 n=3
  Ziehen ohne Zurücklegen
  ohne Berücksichtigung der Anordnung

abc abd abe
acd ace
ade
bcd bce
bde
cde

$$\binom{5}{3} = \frac{5!}{2!3!} = \frac{1\cdot 2\cdot 3\cdot 4\cdot 5}{1\cdot 2\cdot 1\cdot 2\cdot 3} = \frac{5\cdot 4\cdot 3}{1\cdot 2\cdot 3} = 10$$

$$\binom{5}{3} = \binom{5}{2} = \frac{5\cdot 4}{1\cdot 2}$$

Enumeration:10 Möglichkeiten

Kombinatorik:10 Möglichkeiten

Modelle ohne Berücksichtigung der Anordnung

- Ziehen mit Zurücklegen (Situation III)
  - Jede Ziehung erfolgt von der ursprünglichen Grundgesamtheit;
  - das Ergebnis eines Ziehungsschrittes beeinflusst nicht das Ergebnis des nächsten Schrittes;
  - In der Stichprobe können Objekte mehrfach vorkommen

Anzahl der möglichen Stichproben:

$$N^* = \binom{N+n-1}{n} = \frac{(N+n-1)\cdot(N+n-2)\cdot\ldots\cdot(N+1)\cdot N}{n\cdot(n-1)\cdot\ldots\cdot 2\cdot 1}$$

Beispiele für Kombinationen mit Zurücklegen (III)

- In einer Woche ereigneten sich 10 tödliche Verkehrsunfälle. Wie viele unterschiedliche Häufigkeitsverteilungen der Zahl der Unfälle klassifiziert nach dem Wochentag gibt es ?
  N=7 n=10
  Ziehen mit Zurücklegen
  ohne Berücksichtigung der Anordnung
  ➔Sit. III

$$\binom{7+10-1}{10} = \binom{16}{10} = \binom{16}{6} = \frac{16\cdot 15\cdot 14\cdot 13\cdot 12\cdot 11}{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6} = 8008$$

- entspricht der allgemeinen Frage: Wie viele mögliche Verteilungen von n Kugeln auf N Fächer mit beliebiger Kapazität gibt es?

Überblick über Stichprobenmodelle

	Mit Zurücklegen	Ohne Zurücklegen
Mit Berücksichtigung der Anordnung (Variationen)	$N^* = N^n$ I	$N^* = N\,(N-1)\,...\,(N-(n-1))$ II falls $n=N$: $N!$ (Permutation)
Ohne Berücksichtigung der Anordnung (Kombinationen)	$N^* = \binom{N+n-1}{n}$ III	$N^* = N\,(N-1)\,...\,(N-(n-1))/n!$ $N^* = \binom{N}{n}$ IV

Diskussion der Formeln

- N* aus Sit. II (Variation ohne Zurücklegen)ist gleich N* aus Sit. IV (Kombination ohne Zurücklegen) mal n!
- Sit. II unterscheidet sich von Sit. IV eben nur dadurch, dass die Reihenfolge relevant ist.
- Die Formel aus III (Kombination mit Zurücklegen) entspricht, der aus IV (Kombination ohne Zurücklegen), nur ist die Zahl der Elemente der Menge aus der wir die Stichprobe ziehen konzeptionell größer: N+(n-1) statt N
  - Idee: Ab dem zweiten der n Ziehungsschritte vergrößern wir jeweils komparativ die Grundmenge um ein Element

Kombinatorik in R (1)

```
# Kombinatorische Funktionen in R
# ================================

# Faktorielle
cbind(0:10, factorial(0:10))

# Anzahl der Kombinationen
choose(8, 3)
choose(8, 5)

# Generieren aller Kombinationen
combn(letters[1:6], 2)
t(combn(letters[1:5], 3))

# Zufälliges Ziehen einer Kombination ohne Zurücklegen
sample(letters[1:5], 3)
# Zufälliges Ziehen einer Kombination mit Zurücklegen
sample(letters[1:8], 16, rep=T)
```

R Console

```
> # Kombinatorische Funktionen in R
> # ================================
>
> # Faktorielle
> cbind(0:10, factorial(0:10))
      [,1]    [,2]
 [1,]    0       1
 [2,]    1       1
 [3,]    2       2
 [4,]    3       6
 [5,]    4      24
 [6,]    5     120
 [7,]    6     720
 [8,]    7    5040
 [9,]    8   40320
[10,]    9  362880
[11,]   10 3628800
>
```

Kombinatorik in R (2)

```
> # Anzahl der Kombinationen
> choose(8, 3)
[1] 56
> choose(8, 5)
[1] 56
>
> # Generieren aller Kombinationen
> combn(letters[1:6], 2)
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14] [,15]
[1,] "a"  "a"  "a"  "a"  "a"  "b"  "b"  "b"  "b"  "c"   "c"   "c"   "d"   "d"   "e"
[2,] "b"  "c"  "d"  "e"  "f"  "c"  "d"  "e"  "f"  "d"   "e"   "f"   "e"   "f"   "f"
> t(combn(letters[1:5], 3))
      [,1] [,2] [,3]
 [1,] "a"  "b"  "c"
 [2,] "a"  "b"  "d"
 [3,] "a"  "b"  "e"
 [4,] "a"  "c"  "d"
 [5,] "a"  "c"  "e"
 [6,] "a"  "d"  "e"
 [7,] "b"  "c"  "d"
 [8,] "b"  "c"  "e"
 [9,] "b"  "d"  "e"
[10,] "c"  "d"  "e"
>
> # Zufälliges Ziehen einer Kombination ohne Zurücklegen
> sample(letters[1:5], 3)
[1] "d" "e" "c"
> # Zufälliges Ziehen einer Kombination mit Zurücklegen
> sample(letters[1:8], 16, rep=T)
 [1] "a" "b" "a" "f" "d" "h" "g" "g" "h" "a" "e" "d" "d" "e" "d" "b"
>
```

Beispiel: Geburtstagsproblem

- Wie groß ist die Wahrscheinlichkeit, dass unter n Personen zumindest zwei am selben Kalendertag Geburtstag haben (natürlich unabhängig vom Alter/Geburtsjahr)?
- Wie viele Personen müssen anwesend sein, dass die Wahrscheinlichkeit zumindest ein Personenpaar mit gleichem Geburtstag zu erhalten, größer als 50% ist?

Beispiel: Geburtstagsproblem

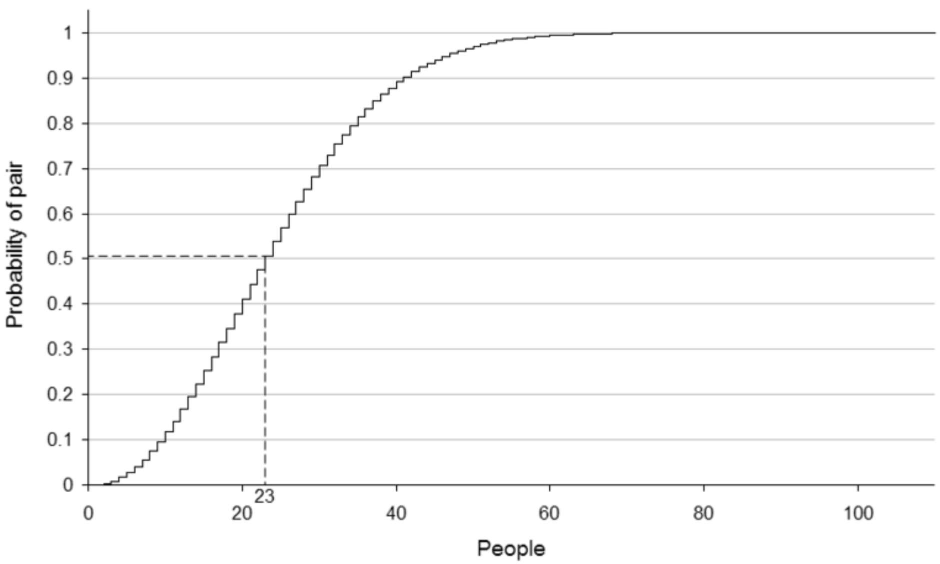
- Wie groß ist die Wahrscheinlichkeit, dass unter n Personen zumindest 2 am selben Kalendertag Geburtstag haben (unabhängig vom Alter) ?
- X sei das Ereignis 2 oder mehr haben am selben Tag Geburtstag $P(X) = 1 - P(X')$
- X' ist das Ereignis, dass alle Personen an verschiedenen Tagen Geburtstag haben
  - 2 Personen: $P(X') = 364/365$
  - 3 Personen: $P(X') = (364/365)*(363/365)$ usw.
- n=30 $P(X') = 0{,}294$ $P(X) = 0{,}706$
- n=50 $P(X') = 0{,}030$ $P(X) = 0{,}970$
- ab n =23 $P(X) > 50\%$

Beispiele

- Wie groß ist die Wahrscheinlichkeit, bei 6 Würfelwürfen alle 6 Augenzahlen genau einmal zu würfeln?
- Mögliche Ergebnisse: n^n $6^6=46.656$
- Günstige Ergebnisse: $n!$ $6!=720$
- Wahrscheinlichkeit: $n! / n^n$
- Für $n=6$: 0,01543

Formale Lösung

- $P(X')=N*(N-1)*...*(N-(n-1))/N^n$
- $N=365$
- n=Anzahl der Personen
- $P(X)=1 - N*(N-1)*...*(N-(n-1))/N^n$



n	$p(n)$
10	12.0%
20	41.0%
23	50.7%
30	70.0%
50	97.0%
100	99.99996%

Modifiziertes Beispiel

- Wie groß ist die Wahrscheinlichkeit, dass unter n Personen zumindest einer am selben Kalendertag Geburtstag hat, wie der Vortragende? Ab wie vielen Personen ist die Chance größer als 50%?
  - P(X) = 1- P(X')
- X' ist das Ereignis, dass alle Personen nicht am Stichtag Geburtstag haben
  - 1 Personen: P(X') = 364/365
  - 2 Personen: P(X') = (364/365)*(364/365) usw.
- n=30 P(X') = 0,921 P(X) = 0,079
- n=50 P(X') = 0,872 P(X) = 0,128
- ab n =253 P(X) > 50%

Bedingte Wahrscheinlichkeit

- Das Konzept bedingter Wahrscheinlichkeit erlaubt zu untersuchen, inwieweit sich die Wahrscheinlichkeiten für das Eintreten von Ereignissen durch das Eintreten anderer Ereignisse ändern.
- Entwicklung anhand eines empirischen Beispiels mit 2 Merkmalen und einer sog. 4-Feldertafel
- Merkmal: Gesundheitszustand mit den Ausprägungen krank (disease) (D+) oder gesund(D-)
- Merkmal: Testergebnis mit den Ausprägungen Test positiv oder negativ (T+ bzw. T-)
- Von Interesse sind hier die Wahrscheinlichkeiten krank zu sein P(D+) oder einen positiven Test zu haben (P(T+) sondern insbesondere die Wahrscheinlichkeit krank zu sein, wenn ein positiver Test vorliegt: P(D+|T+) – aber auch P(T+|D+)

Beispiel zur bedingten Wahrscheinlichkeit

- Anhand eines Labortests (Digitalis-Konzentration im Blut) kann das Vorliegen einer bestimmten Herz-Krankheit diagnostiziert werden. 1975 wurde dazu folgende Statistik veröffentlicht:
- T+…positiver Test T- negativer Test
- D+…Krankheit D- gesund

	D+	D-	Total
T+	25	14	39
T-	18	78	96
Total	43	92	135

Randverteilungen

	D+	D-	Total
T+	0,185	0,104	0,289
T-	0,133	0,578	0,711
Total	0,318	0,682	1,000

Randverteilung (marginale Verteilung):

P(D+) = 0,318 P(D-) = 0,682

P(T+) = 0,289 P (T-) = 0,711

Die Randverteilung eines Merkmals ergibt sich jeweils durch Summation über alle Ausprägungen des anderen Merkmals.

Bedingte Verteilungen

	D+	D-	Total
T+	25	14	39
T-	18	78	96
Total	43	92	135

	D+	D-	Total
T+	0,185	0,104	0,289
T-	0,133	0,578	0,711
Total	0,318	0,682	1,000

Wir interessieren uns nun für die Krankheits-wahrscheinlichkeit, gegeben der Test ist positiv

Bedingte Verteilung:

$P(D+|T+) = 25/39 = 0{,}64$

$P(D+|T+) = P(D+ \cap T+)/P(T+) = 0{,}185/0{,}289 = 0{,}64$

$P(D-|T+) = 14/39 = 1 - P(D+|T+) = 0{,}36$

$P(D-|T+) = P(D- \cap T+)/P(T+) = 0{,}104/0{,}289 = 0{,}36$

Berechnung bedingter Wahrscheinlichkeiten

Bedingte Verteilung gegeben ein negativer Test liegt vor:

P(D-|T-) = 78/96 = 0,578 / 0,711= 0,813

P(D+|T-) = 18/96 = 0,133 / 0,711= 0,187

	D+	D-	Total
T+	0,64	0,36	0,289
T-	0,187	0,813	0,711
Total	0,318	0,682	1,000

In obiger Tabelle sind in den Zeilen die bedingten Verteilungen des Gesundheitszustandes bei Kenntnis des Testergebnisses ausgewiesen (Zeilenprozent).

Interpretation bedingter Wahrscheinlichkeiten

Summary

Offensichtlich verändert die Kenntnis des Testergebnisses meine Krankheitswahrscheinlichkeiten:

- P(D+) = 0,318
- Bei einem positiven Test gilt P(D+|T+) = 25/39 = 0,64
- Bei einem negativen Test gilt P(D+|T-) = 18/96 = 0,187

d.h. der Test ist informativ für das Merkmal Gesundheitszustand; allerdings ist der Test weit davon entfernt perfekt zu sein; er verändert nur den Grad der Sicherheit

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

<u>Lesehinweis:</u>

P(A|B) … Wahrscheinlichkeit, dass ein Ereignis A eintritt, gegeben [oder unter der Bedingung], dass das Ereignis B eingetreten ist

Maßzahlen für die Güte von diagnostischen Tests

	D+	D-	Total
T+	0,185	0,104	0,289
T-	0,133	0,578	0,711
Total	0,318	0,682	1,000

$P(D+) = 0{,}318$ $\qquad$ $P(T+) = 0{,}289$

- Sensitivität des Tests

$P(T+|D+) = 25/43 = 0{,}185/0{,}318 = 0{,}581$

- Spezifität des Tests

$P(T-|D-) = 78/92 = 0{,}578 \,/\, 0{,}682 = 0{,}848$

Sensitivität

- Von einem guten diagnostischen Test wünschen wir uns, dass er möglichst viele Kranke erkennt, das heißt, diese durch ein positives Ergebnis anzeigt. Der Anteil unter allen Kranken, die positiv getestet werden, heißt *Sensitivität*, da er angibt, wie sensibel der Test auf das Vorliegen der Krankheit reagiert.
- Sensitivität: $P(T+|D+)$ … Wahrscheinlichkeit eines positiven Testergebnisses gegeben der Proband ist krank

Spezifität

- Weiters wünschen wir uns, dass der Test möglichst spezifisch ist, also nur auf das Vorliegen der Krankheit anspricht. Jeder nicht Erkrankte, der trotzdem positiv getestet wird, deutet auf einen Mangel an Spezifität [~ P(T+|D-)] hin.
- Als *Spezifität* des Tests bezeichnen wir deshalb den Anteil der korrekt negativ Getesteten unter den nicht Erkrankten.
- Spezifität: P(T-|D-) …Wahrscheinlichkeit eines negativen Testergebnisses gegeben der Proband ist gesund

Verallgemeinerung

- Die Begrifflichkeiten „krank"- „gesund" lassen sich natürlich beliebig auf andere binäre Merkmale übertragen
- Ebenso können wir die Begrifflichkeit des diagnostischen Tests auf das eines bedingenden Merkmals verallgemeinern
- Praxisbeispiel:
  Nutzer eines Webshops tätigt eine Kauf „JA –NEIN"
  Anzahl Besuche im Shop in der letzten Woche ≥ 3 „JA –NEIN"
- Sensitivität des Merkmals #-Besuche für Kauf ist die Wahrscheinlichkeit eines Kaufs, gegeben der User hat den Shop in der letzten Woche schon 3-mal oder öfter besucht
- Spezifität des Merkmals Verweildauer für Kauf ist die Wahrscheinlichkeit keinen Kauf zu tätigen, gegeben der User hat den Shop in der letzten Woche weniger als 3-mal besucht

Statistische Qualität

- Durch die beiden Kriterien Spezifität und Sensitivität kann die statistische Qualität eines diagnostischen Tests charakterisiert werden.
- Wünschenswert ist es, dass ein Test in beiden Kriterien möglich nahe an 100% herankommt.
- Leider wird dieses Idealziel in der Praxis nicht erreicht. Sowohl Kranke als auch Gesunde können positiv oder negativ getestet werden. Deshalb kann aus dem Testergebnis nicht sicher, sondern nur mit einer bestimmten Wahrscheinlichkeit auf das Vorliegen der Krankheit geschlossen werden.

Prädiktiver Wert

- Von Interesse sind in der Praxis folgende bedingten Wahrscheinlichkeiten:
- Der positive prädiktive Wert oder auch Voraussagewert eines positiven Testergebnisses, gibt die Wahrscheinlichkeit an, krank zu sein, wenn ein positiver Test vorliegt
  P(D+|T+)
- Der negative prädiktive Wert oder auch Voraussagewert eines negativen Testergebnisses, gibt die Wahrscheinlichkeit an, gesund zu sein, wenn ein negativer Test vorliegt
  P(D-|T-)

Allgemeine Fragestellung

- Die Anwendung dieser Überlegungen gehen wie schon aufgezeigt weit über diagnostische Tests in der Medizin hinaus
  - Beispiele:
  - Alkomat … Test auf Alkoholisierung
  - Lügendetektor … Test auf Wahrheitsgehalt von Aussagen
- Letztlich von Relevanz bei jeder binären Entscheidung unter Unsicherheit auf der Basis empirischer Evidenz
  Dabei wird man bei vielen Fragestellungen nicht ein bedingdes Merkmal sondern die Kombination mehrer Merkmale verwenden
  - Automatische Erkennung von Falschgeld
  - Warenkorb-Analyse
  - Tests auf Kreditwürdigkeit

Bedingte Wahrscheinlichkeit

Die bedingte Wahrscheinlichkeit eines Ereignisses A unter der Bedingung, dass das Ereignis B eingetreten ist (wobei P(B)>0 sein muss), ist wie folgt definiert:

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

Multiplikationssatz für zwei Ereignisse

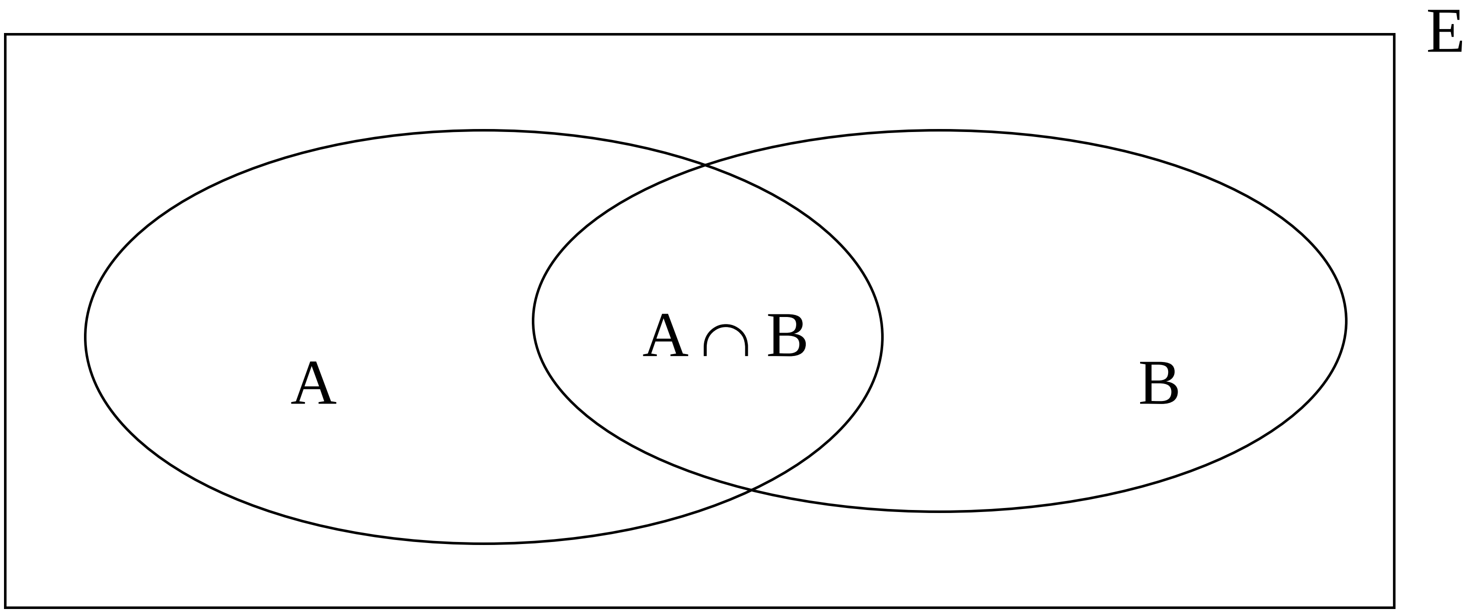
Falls P(A|B)=P(A) (d.h. A ist unabhängig von B) genau dann gilt:

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

Ansonsten bei Abhängigkeit des Ereignisses A von B gilt:

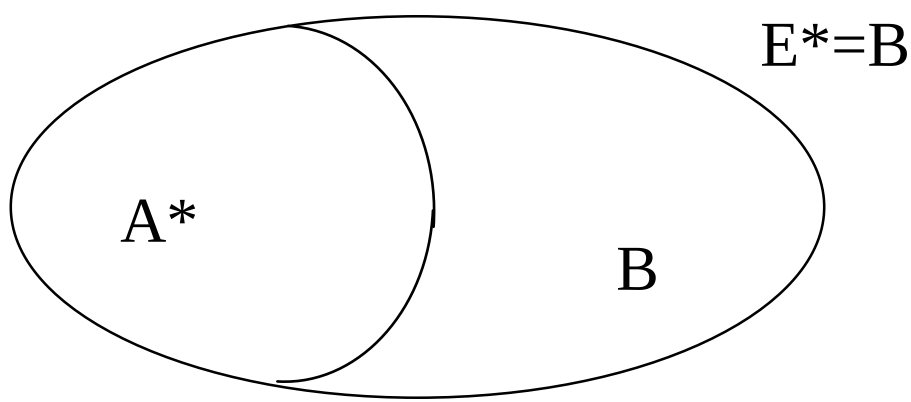
$$P(A \cap B) \neq P(A) \,.\, P(B)$$

Visualisierung des Prinzips der bedingten Wahrscheinlichkeiten



$P(A|B) \sim P(A^*)$

Durch die Bedingung kommt es zu einer Einschränkung des Ereignisraumes



Beispiel

- Für einen männlichen Österreicher gelten folgende Wahrscheinlichkeiten (Sterbetafel 1980/81):
  - P(Alter $\geq$ 70) = 0,59
  - P(Alter $\geq$ 80) = 0,28
- Wie groß ist die Wahrscheinlichkeit, dass ein Mann, der den 70. Geburtstag feiert, auch den 80. Geburtstag feiern kann ?

  P(Alter $\geq$ 80| Alter $\geq$ 70) =
  P(Alter $\geq$ 80 $\cap$ Alter $\geq$ 70) / P(Alter $\geq$ 70) =

  P(Alter $\geq$ 80) / P(Alter $\geq$ 70) = 0,28 / 0,59 = 0,47

  - Es ist evident, dass Berechnungen über Prämien von Lebensversicherungen oder Rentensystemen auf bedingten Wahrscheinlichkeiten basieren müssen!

Berechnung marginaler Wahrscheinlichkeiten

	D+	D-	Total
T+	0,185	0,104	0,289
T-	0,133	0,578	0,711
Total	0,318	0,682	1,000

Durch die Summation der gemeinsamen Wahrscheinlichkeiten

$$P(T+) = P(T+ \cap D+) + P(T+ \cap D-) = 0{,}185 + 0{,}104 = 0{,}289$$

	D+	D-	Total
T+	0,581	0,152	0,289
T-	0,419	0,848	0,711
Total	0,318	0,682	1,000

Durch diegewichtete Summation der bedingten Wahrscheinlichkeiten

$$P(T+) = P(T+|D+) \cdot P(D+) + P(T+|D-) \cdot P(D-) =$$
$$= 0{,}581 * 0{,}318 + 0{,}152 * 0{,}682 = 0{,}289$$

Totale Wahrscheinlichkeit

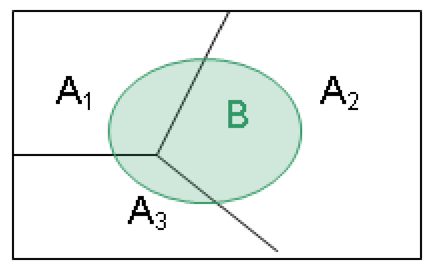
$A=(A \cap B) \cup (A \cap B')$ ... Partition von A auf Basis von B

$P(A) = P(A \cap B) + P(A \cap B') =$ Axiom 3

$P(A|B)\cdot P(B) + P(A|B')\cdot P(B')$ Multiplikationstheorem

Beantwortung von Wahrscheinlichkeitsaussagen unter Berücksichtigung verschiedener Szenarien

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



Beispiel zur bedingten Wahrscheinlichkeit

- Wir verändern die Zahlen des vorigen Beispiels:
- T+...positiver Test T- negativer Test
- D+...Krankheit D- gesund

	D+	D-	Total
T+	12	8	20
T-	48	32	80
Total	60	40	100

$P(T+) = 0{,}2$ $P(D+)=0{,}6$

$P(D+|T+) = 12/20 = 0{,}6$ $P(D+|T-) = 48/80 =0{,}6$

$P(D+ \cap T+) = 12/100 = 0{,}12 = P(D+)\cdot P(T+) = 0{,}2*0{,}6$

STOCHASTISCHE UNABHÄNGIGKEIT

Stochastische Unabhängigkeit (Beispiel)

Nach der Veränderung der Wahrscheinlichkeiten verändert die Kenntnis des Testergebnisses meine Krankheitswahrscheinlichkeiten nicht mehr:

P(D+) = 0,60

Bei einem positiven Test gilt P(D+|T+) = 12/20 = 0,60

Bei einem negativen Test gilt P(D+|T-) = 48/80 = 0,60

	D+	D-	Total
T+	0,60	0,40	0,20
T-	0,60	0,40	0,80
Total	0,60	0,40	1,00

Ein Test, der solche Ergebnisse liefert, ist nicht informativ für das Merkmal Gesundheitszustand. Die bedingten Wahrscheinlichkeiten und die marginale Wahrscheinlichkeit sind gleich.

Stochastische Unabhängigkeit (Beispiel)

Man beachte im Beispiel:

$$P(D+ \cap T+) = 12/100 = 0{,}12$$

$$P(D+ \cap T+) = P(D+) \cdot P(T+|D+) =$$

In diesem Fall

$$P(D+) \cdot P(T+) = 0{,}2 * 0{,}6 = 0{,}12$$

Die gemeinsame Wahrscheinlichkeit ergibt sich im Fall stochastischer Unabhängigkeit einfach aus dem Produkt der marginalen Wahrscheinlichkeiten.

Die gemeinsame absolute Häufigkeit ergibt sich im Fall stochastischer Unabhängigkeit aus dem Produkt der marginalen absoluten Häufigkeiten durch die Gesamtzahl der Beobachtungen.

Stochastische Unabhängigkeit (Theorie)

- Zwei Ereignisse A und B heißen stochastisch unabhängig, wenn gilt:

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

- Korollar:
  Wenn zwei Ereignisse unabhängig sind gilt:
  $P(A|B) = P(A)$ bzw. $P(B|A) = P(B)$.
  d.h. die bedingten Wahrscheinlichkeiten sind gleich den marginalen Wahrscheinlichkeiten
- Bei Unabhängigkeit enthält die isolierte Beobachtung jedes Merkmals für sich bereits die selbe Information, wie die gemeinsame Beobachtung ihres Vorkommens

Beispiel

- Aus der Statistik einer Versicherung ist bekannt, dass 10% aller Personen in einem Jahr einen Unfall erleiden.
- Diskutiere die Wahrscheinlichkeit, dass eine Person in einem Intervall von 2 Jahren unfallfrei ist!

	Jahr2		
Jahr1	Unfall	kein Unfall	Summe
Unfall	1	9	10
kein Unfall	9	81	90
Summe	10	90	100

Unabhängigkeit
81%

	Jahr2		
Jahr1	Unfall	kein Unfall	Summe
Unfall	10	0	10
kein Unfall	0	90	90
Summe	10	90	100

Pechvogel
90%

	Jahr2		
Jahr1	Unfall	kein Unfall	Summe
Unfall	0	10	10
kein Unfall	10	80	90
Summe	10	90	100

Aus Schaden klug
80%

Beispiel: Unabhängige Ereignisse

- In einem Flugzeug gibt es 2 von einander unabhängige automatische Navigationssysteme A und B. Die Verfügbarkeit für das System A sei 0,99 und für B 0,96.
- Wie groß ist die Wahrscheinlichkeit, dass der Pilot zu einer manuellen Navigation greifen muss?
- A ... System A funktioniert $\quad P(A) = 0{,}99$
- B ... System B funktioniert $\quad P(B) = 0{,}96$
- $P(A \text{ ist defekt}) = P(A') = 1 - 0{,}99 = 0{,}01$
- $P(B \text{ ist defekt}) = P(B') = 1 - 0{,}96 = 0{,}04$
- $P(\text{beide Systeme defekt}) = P(A' \cap B') =$
  $= 0{,}01 \times 0{,}04 = 0{,}0004$

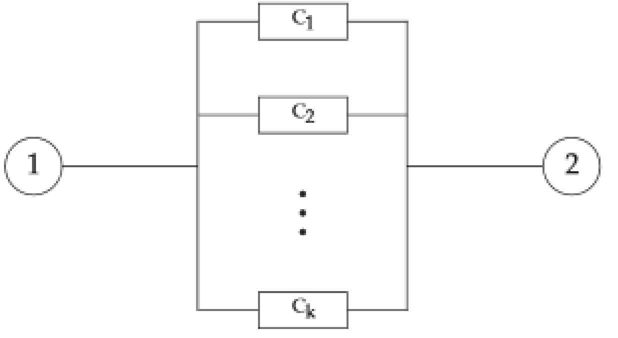
Beispiel: Unabhängige Ereignisse

- Eine Expertenkommission besteht aus 3 Experten A, B, C. Jeder Experte hat eine individuelle Irrtumswahrschein-lichkeit, die wie folgt gegeben ist:
  - P(A irrt) = P(A) = 0,10
  - P(B irrt) = P(B) = 0,15
  - P(C irrt) = P(C) = 0,12
- Wie groß ist die Wahrscheinlichkeit, dass die Meinung der Mehrheit korrekt ist, wenn die 3 Experten voneinander unabhängig urteilen?
- $P(\text{Mehrheit irrt nicht}) = P(A' \cap B' \cap C) + P(A' \cap B \cap C') + P(A \cap B' \cap C') + P(A' \cap B' \cap C') = 0{,}9 \times 0{,}85 \times 0{,}12 + 0{,}9 \times 0{,}15 \times 0{,}88 + 0{,}1 \times 0{,}85 \times 0{,}88 + 0{,}9 \times 0{,}85 \times 0{,}88 = 0{,}0918 + 0{,}1188 + 0{,}0748 + 0{,}6732 = 0{,}9586$

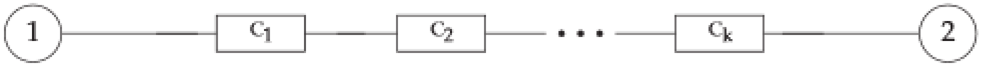
Anwendung in der Systemtheorie

On Systems in Series, Systems in Parallel, and Reliability

- A parallel system consists of k components $c_1, \ldots, c_k$ arranged in such a way that the system works if and only if at least one of the k components functions properly.



- A series system consists of k components $c_1, \ldots, c_k$ arranged in such a way that the system works if an only if all of the components function properly.

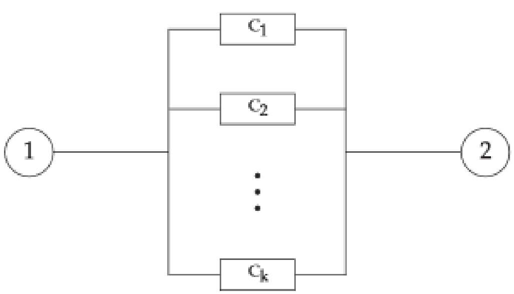


- The reliability of a system is the probability that the system works.

Reliabilität eines parallelen Systems

Example: Parallel system with k mutually independent components

Let $c_1, \ldots, c_k$ denote the k components in a parallel system. Assume the k components operate independently, and $P(c_j \text{ works }) = p_j$. What is the reliability of the system?



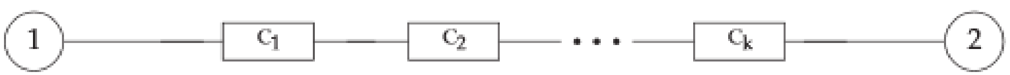
$$
\begin{aligned}
P(\text{ system works}) &= P(\text{ at least one component works}) \\
&= 1 - P(\text{ all components fail }) \\
&= 1 - P(c_1 \text{ fails and } c_2 \text{ fails } \ldots \text{ and } c_k \text{ fails }) \\
&= 1 - \prod_{j=1}^{k} (1 - p_j).
\end{aligned}
$$

Reliabilität eines seriellen Systems

Example: System in series with k mutually independent components

Let $c_1, \ldots, c_k$ denote the k components in a system. Assume the k components are connected in series, operate independently, and $P(c_j \text{ works }) = p_j$. What is the reliability of the system?

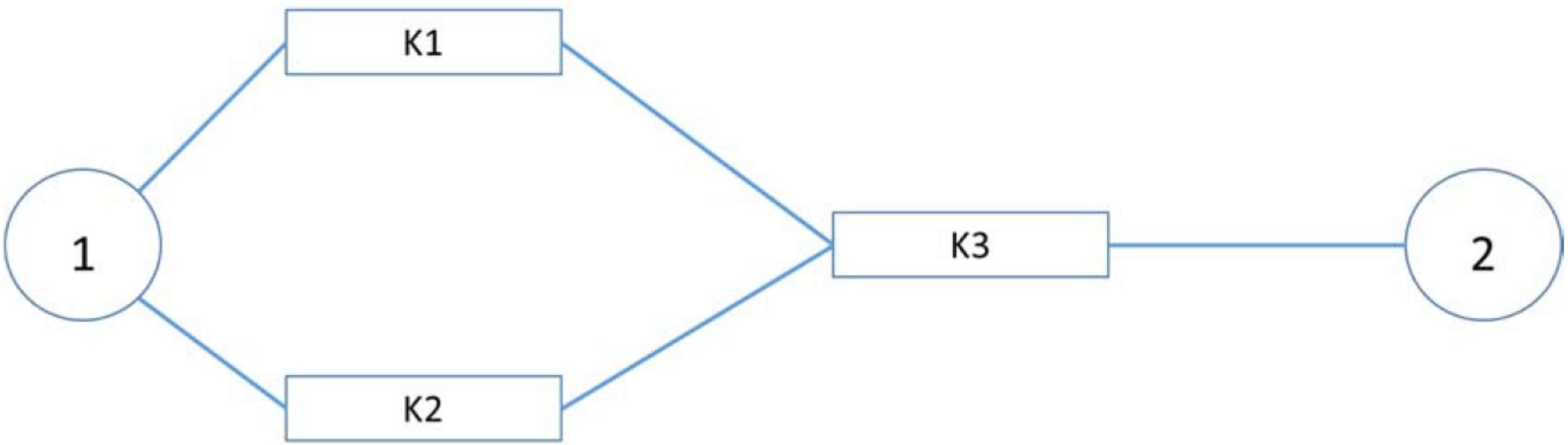
$$\begin{aligned} P(\text{ system works}) &= P(\text{ all components work}) \\ &= \prod_{j=1}^{k} p_j. \end{aligned}$$



(2) Bestimmen Sie für das nachfolgend dargestellte System bestehend aus 3 unabhängig arbeitenden Komponenten K1, K2, K3 die Reliabilität, d.h. die Wahrscheinlichkeit des Funktionierens des Systems.

Es gelten folgende Wahrscheinlichkeiten:

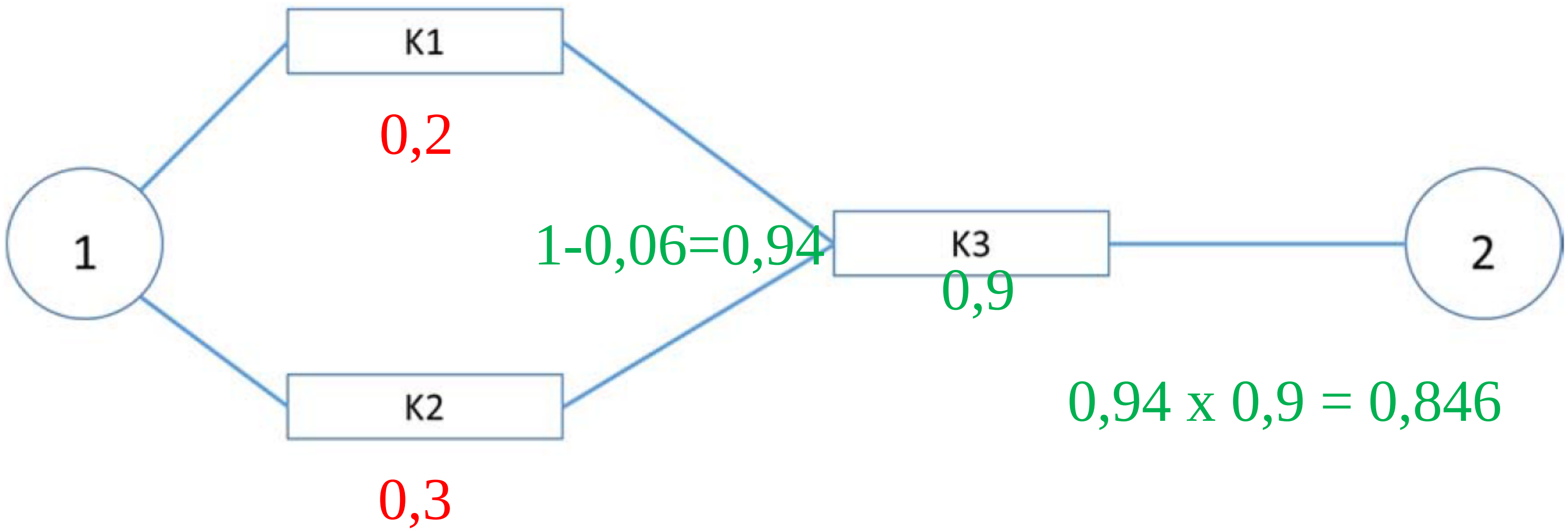
Prob(K1=defekt) =0,2 Prob(K2=defekt) =0,3 Prob(K3=funktioniert) =0,9



(2) Bestimmen Sie für das nachfolgend dargestellte System bestehend aus 3 unabhängig arbeitenden Komponenten K1, K2, K3 die Reliabilität, d.h. die Wahrscheinlichkeit des Funktionierens des Systems.

Es gelten folgende Wahrscheinlichkeiten:

Prob(K1=defekt) =0,2 Prob(K2=defekt) =0,3 Prob(K3=funktioniert) =0,9



Beispiel: Totale Wahrscheinlichkeit

- In einer Population mit gleichen Anteilen von Männern und Frauen wurde festgestellt, dass 5% der Männer und 1% der Frauen farbenblind sind.
- Wie groß ist die Wahrscheinlichkeit, dass eine beliebig (zufällig) herausgegriffene Person farbenblind ist?
- Notation:
  - F...farbenblind N...normalsichtig
  - M...männlich W...weiblich

$$P(F) = P(F|M).P(M) + P(F|W).P(W) =$$
$$= 0{,}05 * 0{,}5 + 0{,}01 * 0{,}5 = 0{,}03$$

Anwendung des Satzes der Totalen Wahrscheinlichkeit

Im Prinzip: Gewogenes Arithmetisches Mittel

Beispiel: Theorem von Bayes

- Daten zur Farbenblindheit

	a priori Wahrschein-lichkeiten
M	0,5
W	0,5
Gesamt	1

- Wie verändern sich diese Wahrscheinlichkeiten, gegeben die Person ist farbenblind ?

Theorem von Bayes

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

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

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

Satz von der totalen Wahrscheinlichkeit

Beispiel: Theorem von Bayes

```
M … männlich
W … weiblich
F … farbenblind
N … normalsichtig
```

- P(M|F) = P(F|M) x P(M) / P(F)
- P(F) = P(F|M) x P(M) + P(F|W) x P(W) =
  = 0,05 x 0,5 + 0,01 x 0,5 = 0,03
  Satz von der totalen Wahrscheinlichkeit
- P(M|F) = 0,05 x 0,5 / 0,03 = 5/6 = 0,833
- P(W|F) = 1 - P(M|F) = 1 - 5/6 = 1/6 = 0,167

- P(M|N) = P(N|M) x P(M) / P(N)
  = 0,95 x 0,5 / 0,97 = 0,49
- P(W|N) = 1- P(M|N) = 0,51
- Man beachte den unterschiedlichen Informationsgehalt von „farbenblind" bzw. „nicht farbenblind" in Bezug auf die veränderten Wahrscheinlichkeiten des Geschlechts!

Beispiel: Theorem von Bayes

- Zusammenfassung der Daten zur Farbenblindheit

	a priori Wahrschein-lichkeiten	Posterior gegeben Farbenblind	Posterior gegeben Normalsichtig
M	0,5	0,833	0,49
W	0,5	0,167	0,51
	1	1	1

Beispiel: Theorem von Bayes

- P(D+|T+) = P(T+|D+) x P(D+) / P(T+)
- P(T+) = 39/135 = 0,289
- P(T+|D+) = 25/43 =0,581
- P(D+) = 43/135 = 0,318
- P(D+|T+) = 0,581 x 0,318 / 0,289 = 0,64

Daten:

	D+	D-	Total
T+	25	14	39
T-	18	78	96
Total	43	92	135

Beispiel: Theorem von Bayes

- Zusammenfassung der Daten

	a priori Wahrschein-lichkeiten	Posterior gegeben positiver Test	Posterior gegeben negativer Test
D+	0,318	0,640	0,187
D-	0,682	0,360	0,848
	1	1	1

Beispiel: Mammographie

- Die Daten stammen aus Kerlinowske et al. 1996, JAMA „Likelihood Ratios for Modern Screening Mammography - Risk of Breast Cancer Based on Age and Mammographic Interpretation“
- Die Wahrscheinlichkeit, dass eine symptomfreie Frau im Alter von 55 Jahren Brustkrebs hat, beträgt 0,6% (d.h. die Prävalenz = $P(D+) = 0{,}006$)

Beispiel

- Wenn eine symptomfreie Frau im Alter von 55 Jahren Brustkrebs hat, dann beträgt die Wahrscheinlichkeit, dass sie einen positiven Mammografie-Befund P(T+) erhält, 94 Prozent.
- Sensitivität des Tests = 0,94
- Wenn eine dieser Frauen jedoch keinen Brustkrebs (D-) hat, dann beträgt die Wahrscheinlichkeit, dass sie dennoch einen positiven Mammografie-Befund erhält nur 7 Prozent.
- Spezifität des Tests = 0,93

Zentrale Frage

Eine 55-jährige Frau, ohne einschlägige Symptome, ist dem Rat ihres Arztes gefolgt, im Rahmen der Brustkrebsfrüherkennung jedes Jahr eine Mammografie durchführen zu lassen. Bei einer solchen Untersuchung erhält sie einen positiven Befund. Schockiert über das Ergebnis, fragt sie ihren Arzt:

«Heißt das, ich habe Brustkrebs?»

«Nein, das kann man noch nicht sicher sagen.»

«Wie hoch ist die Wahrscheinlichkeit, dass ich tatsächlich Brustkrebs habe?»

Beispiel

Ihre Schätzung für die korrekte Wahrscheinlichkeit, dass die Patientin mit einer positiven Mammographie tatsächlich Brustkrebs hat, lautet

_ _ , _ %

Schema der Diagnostik

A priori Wahrscheinlichkeit einer Erkrankung (Prävalenz)

↓

Diagnostischer Test

P(T-|D-) Spezifität

P(T+|D+) Sensitivität

↓

Posteriore Wahrscheinlichkeit einer Erkrankung

Falls Test positiv ist	Falls Test negativ ist
P(D+|T+) = ???	P(D-|T-) = ???

- Welche korrekte statistische Angabe kann der Arzt der Patientin geben?
- Prävalenz = $P(D+) = 0,006$
- Sensitivität des Tests
  $P(T+|D+) = 0,94$
- Spezifität des Tests bzw. 1-Spez
  $P(T-|D-) = 0,93$ bzw. $P(T+|D-) = 0,07$
- $P(D+|T+) = ???$

Theorem von Bayes

a priori Wahrschein-lichkeit → Theorem von Bayes → posteriore Wahrschein-lichkeit

$$P(D+\,|\,T+) = \frac{P(T+\,|\,D+)\cdot P(D+)}{P(T+)}$$

$$P(T+) = P(T+\,|\,D+)\cdot P(D+) + P(T+\,|\,D-)\cdot P(D-)$$

$$P(D+\,|\,T+) = \frac{Sens\cdot \mathrm{Pr}\ddot{a}v}{Sens\cdot \mathrm{Pr}\ddot{a}v + (1-Spez)\cdot(1-\mathrm{Pr}\ddot{a}v)}$$

Anwendung des Bayes Theorem

- $P(T+) = P(T+|D+)*P(D+) + P(T+|D-)*P(D-)=$

$= 0,94*0,006 + 0,07*0,994 =0,07522$

- $P(D+|T+) = P(T+|D+)*P(D+)/P(T+)=$

$= 0,94*0,006/0,07522=0,07498$

- Die Wahrscheinlichkeit, dass die Patientin mit einer positiven Mammographie tatsächlich Brustkrebs hat, beträgt 7,5%
- In einer amerikanischen Studie lagen 95 von 100 befragten Ärzten in ihrer Schätzung zwischen 70% und 80%.

Formulierung in absoluten Zahlen

	Brustkrebs (D+)	Gesund (D-)	Summe
Test +			
Test -			
Summe			100.000

	Brustkrebs (D+)	Gesund (D-)	Summe
Test +			
Test -			
Summe	600	99.400	100.000

0,6% <=== Prävalenz

	Brustkrebs (D+)	Gesund (D-)	Summe
Test +	564		
Test -			
Summe	600	99.400	100.000

94,0% <=== Sensitivität

7,0% <=== 1 minus Spezifität

	Brustkrebs (D+)	Gesund (D-)	Summe
Test +	564	6.958	
Test -	36		
Summe	600	99.400	100.000

Formulierung in absoluten Zahlen

	Brustkrebs (D+)	Gesund (D-)	Summe
Test +	564	6.958	
Test -	36	92.442	
Summe	600	99.400	100.000

93,0% <=== Spezifität

	Brustkrebs (D+)	Gesund (D-)	Summe
Test +	564	6.958	7.522
Test -	36	92.442	92.478
Summe	600	99.400	100.000

	Brustkrebs (D+)	Gesund (D-)	Summe
Test +	564	6.958	7.522
Test -	36	92.442	92.478
Summe	600	99.400	100.000

564/7.522= 7,5%

Prävalenzabhängigkeit von Tests

- Ein ELISA zum Test auf HIV-Antikörper besitzt 99.99% Sensitivität und 98% Spezifität.
- Wir setzen diesen Test nun in zwei Situationen ein. In Population A ("Normalpopulation") liege die Prävalenz bei 0.01%.
- Population B („Risiko-Population") habe eine Prävalenz von 5%.
- In beiden Fällen wollen wir wissen, wie sicher wir bei einem positiven Test sein können, dass der Proband tatsächlich HIV-positiv ist.

- Ergebnis bei A: P(D+|T+) = 0,5%
- Ergebnis bei B: P(D+|T+) = 72%

Aussagekraft des Tests hängt stark von der Prävalenz ab!

Lügendetektoren und das Theorem von Bayes

Gastwirth(1978):

+ ... Test ergibt Person lügt - ... Test zeigt an Person lügt nicht

L ... Person lügt in Wirklichkeit W ... Person spricht die Wahrheit

P(+|L) = 0,88 P(-|L) = 0,12

P(-|W) = 0,86 P(+|W) = 0,14

a) Routinetest bei Personalselektion

P(W) = 0,99 P(L) = 0,01

P(+) = P(+|W)P(W) + P(+|L)P(L) =0,14*0,99+0,88*0,01=0,1474

P(L|+) = P(+|L)P(L)/P(+) = 0,88*0,01/0,1474=0,0597

P(W|+) = P(+|W)*P(W)/P(+) = 0,14*0,99/0,1474=0,9403

P(W|-) = P(-|W)*P(W)/P(-) = 0,86*0,99/0,853=0,998

P(L|-) = 1 - P(W|-) = 0,002

Lügendetektoren und das Theorem von Bayes

b) Verändern der subjektiven Wahrscheinlichkeit

b1) $P(W) = 0{,}50$ $\quad$ $P(L) = 0{,}50$

$P(+) = P(+|W)P(W) + P(+|L)P(L) = 0{,}14*0{,}5+0{,}88*0{,}5=0{,}51$

$P(L|+) = P(+|L)P(L)/P(+) = 0{,}88*0{,}5/0{,}51 = 0{,}863$

$P(W|+) = P(+|W)*P(W)/P(+) = 0{,}14*0{,}5/0{,}51 = 0{,}137$

$P(L|-) = P(-|L)P(L)/P(-) = 0{,}12*0{,}5/0{,}49 = 0{,}122$

$P(W|-)= 1- P(L|-) = 0{,}878$

b2) $P(W) = 0{,}20$ $\quad$ $P(L) = 0{,}80$

$P(+) = P(+|W)P(W) + P(+|L)P(L) = 0{,}14*0{,}5+0{,}88*0{,}5=0{,}732$

$P(L|+) = P(+|L)P(L)/P(+) = 0{,}88*0{,}8/0{,}732 = 0{,}96$

$P(W|+) = P(+|W)*P(W)/P(+) = 0{,}14*0{,}2/0{,}732 = 0{,}04$

$P(L|-) = P(-|L)P(L)/P(-) = 0{,}12*0{,}8/0{,}268 = 0{,}36$

$P(W|-)= 1- P(L|-) = 0{,}64$

Übungsbeispiele

- Eine Password-Policy schreibt als Minimalanforderung vor, dass ein Password aus mindestens 6 Zeichen bestehen muss, wobei das erste Zeichen ein Buchstabe sein muss und die restlichen Zeichen Buchstaben oder Ziffern sein können.
- Es wird nicht nach Groß- und Kleinbuchstaben unterschieden und wir setzen voraus, dass es 26 Buchstaben gibt.

a) Wie viele unterschiedliche Passwörter können mit den angegebenen Mindest-Kriterien generiert werden?

b) Wie viele unterschiedliche Passwörter können mit den angegebenen Minimalkriterien generiert werden, wenn als zusätzliche Anforderung formuliert wird, dass kein Zeichen öfter als einmal verwendet werden darf?

Übungsbeispiele

- Wie groß ist die Wahrscheinlichkeit bei Poker mit 52 Karten (4 Farben mit 13 Kartenwerten) in einer Hand bestehend aus 5 Karten 4 Karten mit gleichen Wert zu haben (also einen Poker zu haben)?

R Console

```
> # Mögliche Fälle
> m <- choose(52, 5)
> m
[1] 2598960
> # Günstige Fälle
> g <- 13 * choose(4, 4) * choose(48, 1)
> g
[1] 624
> # Probability
> g/m
[1] 0.000240096
>
```

Übungsbeispiele

- Wie groß ist die Wahrscheinlichkeit bei Poker mit 52 Karten (4 Farben mit 13 Kartenwerten) in einer Hand bestehend aus 5 Karten ein Full (1 Paar und 1 Drilling) zu haben?

R Console

```
> # Mögliche Fälle
> m <- choose(52, 5)
> m
[1] 2598960
> # Günstige Fälle
> g <- 13 * choose(4, 2) * 12 * choose(4, 3)
> g
[1] 3744
> # Probability
> g/m
[1] 0.001440576
> 
```

Übungsbeispiele

- 2 Basketballspieler A und B versuchen den Ball im Korb zu versenken.
  A trifft mit Prob(A)=1/3 für B gilt Prob(B) = 1/4.

a) Jeder wirft solange bis er einmal getroffen hat. Wie groß ist die Wahrscheinlichkeit, dass A schon beim zweiten Versuch einen Korb erzielt?

b) Wie groß ist die Wahrscheinlichkeit, dass B spätestens nach 6 Würfen seine Serie beendet hat?

c) Die Spieler werfen abwechselnd A beginnt, wie groß ist die Wahrscheinlichkeit, dass A vor B trifft?

Übungsbeispiele

Ein elektronischer Bauteil wird von 3 großen Herstellern produziert.

Hersteller	Marktanteil	durchschnittliche Ausfallsrate
A	50%	10%
B	40%	15%
C	10%	20%

a) Wie groß ist die durchschnittliche Ausfallsrate für Bauteile dieses Typs?

b) In einem System ist ein eingebautes Bauteil dieses Typs defekt. Sie wissen nicht von welchem Hersteller das eingebaute Bauteil stammt.

Bestimmen Sie die Wahrscheinlichkeiten dafür, dass es sich bei dem ausgefallenen Bauteil um ein Produkt des Herstellers A, B bzw. C handelt!