

Statistik und Wahrscheinlichkeitsrechnung

- Wintersemester 2020/21 -

Kapitel 03: Wahrscheinlichkeitsrechnung

Prof. Dr. Adrian Ulges

Angewandte Informatik (B.Sc.) / Informatik - Technische Systeme (B.Sc.) / Wirtschaftsinformatik (B.Sc.)

> Fachbereich DCSM Hochschule RheinMain

Outline



- 1. Grundlegende Definitionen
- 2. Der Wahrscheinlichkeitsbegriff
- 3. Der allgemeine Additionssatz
- 4. Bedingte Wahrscheinlichkeiten
- Der Multiplikationssatz
- 6. Unabhängigkeit von Ereignissen
- 7. Der Satz der totalen Wahrscheinlichkeit
- 8. Die Formel von Bayes

Definition: Zufallsexperiment Bilder: [6] [1] [5]



Ein Zufallsexperiment ist ein Vorgang ...

- ... der sich unter den gleichen Bedingungen beliebig oft wiederholen lässt
- ... und dessen Ergebnis sich nicht sicher vorhersagen lässt.

Die uns aus der deskriptiven Statistik bekannten Stichproben $x_1, ..., x_n$ entstehen durch die wiederholte Durchführung von Zufallsexperimenten.







Definition: Ergebnismenge Bilder: [7] [3] [8]



Wir nennen die Menge aller möglichen Resultate des Experiments die Ergebnismenge Ω .

Beispiele

- ► Würfeln eines Würfels:
 - $\Omega = \{1, 2, 3, 4, 5, 6\}$ (endlich)
- ► Anzahl der Tage bis zum Defekt einer Festplatte:

$$\Omega = \{0, 1, 2, ...\} = \mathbb{N}$$
 (unendlich)

Körpergröße einer zufällig ausgewählten Person (in cm): $\Omega = [0,400]$ (überabzählbar)







Definition: Ereignis Bild: [2]



Definition (Ereignis)

Ein Ereignis ist eine Teilmenge $A \subseteq \Omega$. Gilt für das Ergebnis eines Zufallsexperiments ω dass $\omega \in A$, sagen wir: "Das Ereignis A ist eingetreten."

Beispiel: Würfeln eines Würfels

- $A = \text{``Augenzahl ist} \geq 5\text{''} = \{5,6\}$
- ightharpoonup B = "Augenzahl ist gerade" = $\{2,4,6\}$
- Wir würfeln eine 5
 - \rightarrow A ist eingetreten, B nicht.



Besondere Ereignisse

- ► Elementarereignisse sind einelementig (z.B. {1}, {2}, ..., {6}).
- $ightharpoonup \Omega$ nennen wir das sichere Ereignis, \varnothing das unmögliche Ereignis.

Verknüpfung von Ereignissen [13]



Da wir Ereignisse als Mengen definiert haben, können wir sie mittels Mengenoperationen verknüpfen. Seien A, B Ereignisse. Dann stellen die folgenden Verknüpfungen ebenfalls Ereignisse dar:

$\mathbf{A} \cup \mathbf{B}$: A oder B tritt ein (oder beide).	Ω Ω
A ∩ B (oder kurz: A , B): A und B treten beide ein.	Δ A B D
$ar{f A}=\Omegaackslash A$: A tritt nicht ein.	A A Q
Falls $A \cap B = \emptyset$, können A und B nicht gleichzeitig auftreten. Die Ereignisse "schließen sich aus".	A B Ω

Do-Ereignisse-Yourself Bild: [2]

Wir werfen zwei vierseitige Würfel.

- \blacktriangleright Wie lautet die Ergebnismenge Ω?
- ▶ Berechnen Sie A := "Der 2. Würfel zeigt ein höheres Ergebnis als der erste Würfel"
- ▶ Berechnen Sie *B* := "Die Augensumme ist durch 3 teilbar"
- ▶ Berechnen Sie *C* := "A tritt ein, aber B nicht"
- ► Schließen *A* und *B* einander aus?



Do-Ereignisse-Yourself Bild: [2]



Outline



- 1. Grundlegende Definitionen
- 2. Der Wahrscheinlichkeitsbegriff
- 3. Der allgemeine Additionssatz
- 4. Bedingte Wahrscheinlichkeiten
- 5. Der Multiplikationssatz
- 6. Unabhängigkeit von Ereignissen
- 7. Der Satz der totalen Wahrscheinlichkeit
- 8. Die Formel von Bayes

Der Begriff "Wahrscheinlichkeit" Bilder: [9] [10] [11]





Pierre-Simon Laplace



Richard von Mises



Andrei Kolmogorov

"Wahrscheinlichkeit": Laplace'scher Begriff



Definition (Laplace-Experiment)

Wir bezeichnen ein Zufallsexperiment als Laplace-Experiment, falls zwei Bedingungen erfüllt sind:

- 1. Die Ergebnismenge Ω ist endlich
- 2. Alle Elementarereignisse sind gleich wahrscheinlich.

In diesem Fall berechnen wir die Wahrscheinlichkeit von A, notiert als P(A), als:

$$P(A) = \frac{\#A}{\#\Omega}$$

Beispiel: Werfen eines fairen Würfels

- ► Es gilt: $P(\{1\}) = P(\{2\}) = ... = P(\{6\}) = \frac{1}{6}$
- ► Beispiele komplizierterer Ereignisse:

$$P(\{2,4,6\}) = \frac{\#\{2,4,6\}}{\#\Omega} = \frac{3}{6} = \frac{1}{2}$$
 $P(\{5,6\}) = \frac{\#\{5,6\}}{\#\Omega} = \frac{2}{6} = \frac{1}{3}$

Sind dies Laplace-Experimente?



► Festplatten-Defekt

 $lackbox{ }\Omega=\mathbb{N}$ (Anzahl der Tage bis zum Defekt)

Black Jack

- Ziehen von 2 Karten aus einem Stapel von 52
- $\Omega = \{ A \subset \{1, ..., 52\} \mid \#A = 2 \}$

Werfen von Reißzwecken

- Elementarereignisse O ("Spitze zeigt nach oben") und U ("Spitze zeigt nach unten")

"Wahrscheinlichkeit": Laplace'scher Begriff



Kritik am Laplaceschen Wahrscheinlichkeitsbegriff

In der Praxis sind die Elementarereignisse oft *nicht* gleich wahrscheinlich. *Wie verfahren wir in komplizierteren Fällen?*

Beispiele

- Werfen eines gezinkten Würfels: $\Omega = \{1, ..., 6\}$ mit $P(\{6\}) > P(\{1\})$
- Werfen von zwei Würfeln und Addieren der Augen: $\Omega = \{2, 3, ..., 7, ..., 11, 12\}$ mit $P(\{7\}) > P(\{2\})$
- ► Zugriffshäufigkeiten eines Webservers pro Wochentag: $\Omega = \{MO, DI, MI, DO, FR, SA, SO\}$ mit $P(\{MI\}) > P(\{SO\})$



Relative Häufigkeiten und Wahrscheinlichkeiten



- Im Allgemeinen müssen Elementarereignisse nicht gleich wahrscheinlich sein.
- ▶ Dies kann z.B. physikalisch bedingt sein (gezinkter Würfel).
- ► Wie definieren wir nun Wahrscheinlichkeiten?

Ansatz: Relative Häufigkeiten → Wahrscheinlichkeiten

- ▶ Wir wiederholen das Zufallsexperiment *n*-mal.
- ► Ein beliebiges Ereignis A besitze nach n Wiederholungen die relative Häufigkeit $h_n(A)$.
- ▶ Diese Häufigkeit konvergiert falls wir das Zufallsexperiment beliebig oft unter gleichen Bedingungen wiederholen – gegen die Wahrscheinlichkeit P(A).
- ▶ Dies ist von Mises' sogenannte Limes-Definition.

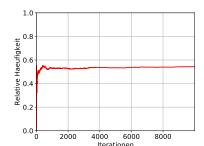
Beispiel: Ist ein Würfel gezinkt?



- ► Wir würfeln 3, 5, 6, 5, 2, 1, 4, 4, 2, ...
- \blacktriangleright Wir betrachten das Ereignis A = "Augenzahl gerade".
- ▶ Wir messen A's relative Häufigkeit nach n=1, 2, ... Versuchen:

n	1	2	3	4	5	6	7	8	9	
ω_n	3	5	6	5	2	1	4	4	2	
$h_n(A)$	0	0	$\frac{1}{3}$	$\frac{1}{4}$	<u>2</u> 5	$\frac{2}{6}$	$\frac{3}{7}$	<u>4</u> 8	<u>5</u>	

▶ Beobachtung: $h_n(A)$ konvergiert gegen 54% (Würfel gezinkt).





Kolmogorov verallgemeinert Laplace' Definition. Er definiert "Wahrscheinlichkeiten" als Werte, die gewisse Bedingungen (oder *Axiome*) erfüllen:

- ightharpoonup Gegeben sei ein Zufallsexperiment mit Ergebnismenge Ω.
- ► Wir definieren ein sog. Wahrscheinlichkeitsmaß *P*, das jedem Ereignis eine Wahrscheinlichkeit zwischen 0 und 1 zuordnet.

Beispiel

- Münzwurf: $\Omega = \{Z, K\}$
- ► Ereignisse: ∅, {*Z*}, {*K*}, {*Z*, *K*}
- ► Wahrscheinlichkeitsmaß P, zum Beispiel
 - $P(\emptyset) = 0$
 - $P({Z}) = 0.4$
 - $P(\{K\}) = 0.6$
 - $P(\{Z,K\}) = 1$



- ▶ Wir fassen sämtliche möglichen Ereignisse $A \subseteq \Omega$, denen wir Wahrscheinlichkeiten zuordnen wollen, in einer sogenannten Sigma-Algebra Σ zusammen.
- \triangleright Wir nennen Σ auch Ereignismenge des Zufallsexperiments.
- ► Beispiel: Sechsseitiger Würfel

$$\begin{split} \Sigma &= \Big\{\{1\},\{2\},...,\{6\},\\ &\{1,2\},\{1,3\},\{1,4\},...,\{5,6\},\\ &\{1,2,3\},\{1,2,4\},\{1,2,5\},...,\{4,5,6\},\\ &...\\ &\{1,2,3,4,5,6\}\Big\} \end{split}$$

Anmerkungen

- ▶ Wir stellen uns unter Σ die Potenzmenge $\mathcal{P}(\Omega)$ vor.
- In Wirklichkeit ist es etwas komplizierter, falls Ω überabzählbar ist.



Definition (Wahrscheinlichkeitsmaß)

Gegeben sei ein Zufallsexperiment mit Ereignismenge Σ . Dann nennen wir $P:\Sigma \to [0,1]$ ein Wahrscheinlichkeitsmaß, falls P folgende Eigenschaften erfüllt:



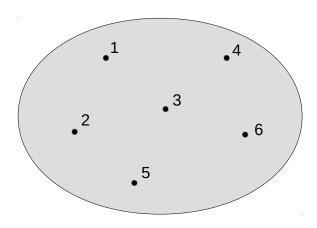
Anmerkungen

- \blacktriangleright Wir nennen (Ω, Σ, P) einen Wahrscheinlichkeitsraum.
- ▶ Die Wahrscheinlichkeitsdefinition nach Laplace erfüllt ebenfalls die Kolmogorov-Axiome. Hier liegt kein Widerspruch vor – Kolmogorovs Definition ist nur allgemeiner.
- ▶ Die Axiome werden uns im Folgenden erlauben, Regeln für das Rechnen mit Wahrscheinlichkeiten einzuführen.

Do-it-yourself: Gezinkter Würfel



Wir wissen: $P(\{1\})=P(\{2\})=P(\{3\})=P(\{4\})=\frac{1}{10}$ und $P(\{1,3,6\})=\frac{7}{10}$.



Berechne $P(\{6\})$ und $P(\{5\})$. Ist P nun für <u>alle</u> $A \in \Sigma$ definiert?









Outline



- 1. Grundlegende Definitionen
- 2. Der Wahrscheinlichkeitsbegriff
- 3. Der allgemeine Additionssatz
- 4. Bedingte Wahrscheinlichkeiten
- 5. Der Multiplikationssatz
- 6. Unabhängigkeit von Ereignissen
- 7. Der Satz der totalen Wahrscheinlichkeit
- 8. Die Formel von Bayes

Allgemeiner Additionssatz



Bisher wissen wir, dass für disjunkte Ereignisse A, B gilt:

$$P(A \cup B) = P(A) + P(B).$$

Für allgemeine Ereignisse A, B gilt verallgemeinern wir:

Satz (Allgemeiner Additionssatz)

Es sei (Ω, Σ, P) ein Wahrscheinlichkeitsraum und $A, B \in \Sigma$ zwei Ereignisse. Dann gilt:

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

Anmerkungen

Dies schließt den speziellen Additionssatz mit ein: Sind A und B disjunkt, so erhalten wir $P(A \cap B) = 0$ und somit $P(A \cup B) = P(A) + P(B)$.

Allgemeiner Additionssatz: Beweis



Outline



- 1. Grundlegende Definitionen
- 2. Der Wahrscheinlichkeitsbegriff
- 3. Der allgemeine Additionssatz
- 4. Bedingte Wahrscheinlichkeiten
- Der Multiplikationssatz
- 6. Unabhängigkeit von Ereignissen
- 7. Der Satz der totalen Wahrscheinlichkeit
- 8. Die Formel von Bayes





- ▶ Wir wissen jetzt wie wir $P(A \cup B)$ berechnen.
- ► Komplizierter wird es, wenn wir $P(A \cap B)$ (kurz: P(A, B)) berechnen wollen, d.h. die Wahrscheinlichkeit für ein Auftreten von A und B.

Beispiel (Karten)

- Aus einem Stapel von Spielkarten ziehen wir zwei Karten ohne Zurücklegen.
- Wir betrachten die Ereignisse
 - ► *A* = "Die erste Karte ist eine Kreuz-Karte"
 - \triangleright B = "Die zweite Karte ist eine Kreuz-Karte"

$$P(A) = \frac{\text{\# Kreuz-Karten}}{\text{\# alle Karten}} = \frac{8}{32} = 0.25$$

$$P(B) = ?$$



Definition (Bedingte Wahrscheinlichkeit)

Es sei (Ω, Σ, P) ein Wahrscheinlichkeitsraum mit Ereignissen $A, B \in \Sigma$, sowie $P(A) \neq 0$. Wir definieren die Wahrscheinlichkeit des Ereignisses $B \in \Sigma$ unter der Annahme dass $A \in \Sigma$ eintritt als:

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

Wir nennen P(B|A) eine bedingte Wahrscheinlichkeit.

Anmerkungen

- ► Auf der rechten Seite des Trennstrichs (A) steht unser "Wissen": Ereignisse, die wir als gegeben annehmen können.
- ▶ Der Unterschied zwischen P(B) und P(B|A) besteht in dieser Zusatzinformation A.



Anmerkungen (cont'd)

Für bedingte Wahrscheinlichkeiten gelten dieselben Rechenregeln wie für "normale" Wahrscheinlichkeiten, zum Beispiel:

- Gegenereignis: $P(\overline{B}|A) = 1 P(B|A)$
- Additionssatz:

$$P(B_1 \cup B_2 | A) = P(B_1 | A) + P(B_2 | A) - P(B_1 \cap B_2 | A)$$

- ▶ Sicheres Ereignis: $P(\Omega|A) = 1$
- ▶ ..
- ► Achtung: Es gilt nicht $P(B|A) + P(B|\overline{A}) = 1!$

Bedingte Wahrscheinlichkeiten: Do-it-yourself



- Würfeln mit zwei Würfeln, Ergebnis = Summe der Augen $(\Omega = \{2, ..., 12\}).$
- ► Elementarereignisse mit Wahrscheinlichkeiten

- Ereignisse
 - ightharpoonup A = "Das Ergebnis ist ≥ 7 "
 - \triangleright B = "Das Ergebnis ist eine gerade Zahl"

Berechnen Sie P(B|A).

Bedingte Wahrscheinlichkeiten: Do-it-yourself



Wahrscheinlichkeiten notieren



Als die Titanic sank, waren 35% aller Passagiere Frauen. Drei von vier weiblichen Passagieren überlebten das Unglück. 10% der männlichen Passagiere waren Jungen unter 18. Von diesen überlebten 34%. Bei den restlichen Männern betrug die Überlebensquote 19%.



Outline



- 1. Grundlegende Definitionen
- 2. Der Wahrscheinlichkeitsbegriff
- 3. Der allgemeine Additionssatz
- 4. Bedingte Wahrscheinlichkeiten
- 5. Der Multiplikationssatz
- 6. Unabhängigkeit von Ereignissen
- 7. Der Satz der totalen Wahrscheinlichkeit
- 8. Die Formel von Bayes

Der Multiplikationssatz



Aus der Definition der bedingten Wahrscheinlichkeit folgt unmittelbar der sogenannte Multiplikationssatz:

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

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

Satz (Multiplikationsssatz)

Es sei (Ω, Σ, P) ein Wahrscheinlichkeitsraum und $A, B \in \Sigma$ zwei Ereignisse mit $P(A) \neq 0$ und $P(B) \neq 0$. Dann gilt:

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

Remarks

Wir können nun Wahrscheinlichkeiten für das Auftreten mehrerer Ereignisse (d.h. $A \cap B$ oder A, B) zu berechnen.

Multiplikationssatz: Allgemeine Form



Der Multiplikationssatz gilt nicht nur für <u>zwei</u> Ereignisse *A*, *B*, sondern für ganze *Ereignisketten*:

Satz (Multiplikationsssatz für > 2 Ereignisse)

Es sei (Ω, Σ, P) ein Wahrscheinlichkeitsraum, und es seien $A_1, A_2, ..., A_n \in \Sigma$ Ereignisse mit $P(A_1), P(A_2), ..., P(A_{n-1}) \neq 0$. Dann gilt:

$$P(A_1, A_2, ..., A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1, A_2) \cdot ...$$

 $\cdot P(A_n|A_1, ..., A_{n-1})$

Anmerkungen

Ahnlich wie beim vorherigen Satz können wir auf der rechten Seite die Reihenfolge der Ereignisse $A_1, ..., A_n$ vertauschen.

Outline



- 1. Grundlegende Definitionen
- 2. Der Wahrscheinlichkeitsbegriff
- 3. Der allgemeine Additionssatz
- 4. Bedingte Wahrscheinlichkeiten
- Der Multiplikationssatz
- 6. Unabhängigkeit von Ereignissen
- 7. Der Satz der totalen Wahrscheinlichkeit
- 8. Die Formel von Bayes

Unabhängigkeit von Ereignissen



Oft nehmen wir an, dass Ereignisse einander nicht beeinflussen:

Beispiel 1: Zweifacher Würfelwurf

- \triangleright A = "Wir würfeln eine 6 im ersten Wurf."
- \triangleright B = "Wir würfeln eine 6 im zweiten Wurf."

Beispiel 2: Befragung von Wählern

- ightharpoonup A = "Die erste Person wählt die SPD."
- \triangleright B = "Die zweite Person wählt die SPD."

Beispiel 3: Ziehe eine Karte

- ightharpoonup A = "Die Karte ist eine Kreuzkarte."
- \triangleright B = "Die Karte ist eine Bildkarte."

In diesen Fällen können wir annehmen, dass die Wahrscheinlichkeit von B durch das Eintreten von A nicht beeinflusst wird. Es gilt also: P(B|A) = P(B). Wir nennen A und B unabhängig.

Unabhängigkeit von Ereignissen



Satz (Unabhängigkeit von Ereignissen)

Es sei (Ω, Σ, P) ein Wahrscheinlichkeitsraum und $A, B \in \Sigma$.

Dann nennen wir A und B genau dann unabhängig, wenn eine der folgenden Eigenschaften erfüllt ist:

- 1. P(B|A) = P(B) (falls P(A) > 0)
- 2. P(A|B) = P(A) (falls P(B) > 0)
- 3. $P(A, B) = P(A) \cdot P(B)$

Anmerkungen

- ▶ Falls P(A) > 0 und P(B) > 0, gilt: $(1.) \Leftrightarrow (2.) \Leftrightarrow (3.)$ (gilt also eine der drei Eigenschaften, gelten automatisch alle).
- Wir nennen (3.) den Multiplikationssatz für unabhängige Ereignisse.

Unabhängigkeit von Ereignissen: Beweis



Beispiel 1: Ziehen von zwei Karten



Wir ziehen zwei Karten aus einem 32-Blatt-Kartenspiel (geordnet, ohne Zurücklegen). Es ergibt sich ein Laplace-Experiment mit $\#\Omega=32\times31$:

$$\Omega = \{(K7, K8), ..., (K7, HA), (K8, K7), ..., (K8, HA)..., (HA, HK)\}$$

Wir untersuchen auf Unabhängigkeit:

- ► *A* = "Erste Karte ist Kreuz"
- ightharpoonup B = "Zweite Karte ist Kreuz"

Beispiel 2: Ziehen von zwei Karten



Wir untersuchen auf Unabhängigkeit:

- ► *A* = "Erste Karte ist Kreuz"
- \triangleright B = "Zweite Karte ist Bild"

Rechenregeln: Zusammenfassung



$$P(A \cap B) = \begin{cases} P(A) \cdot P(B) & \text{falls } A, B \text{ unabhängig} \\ P(A) \cdot P(B|A) & \text{sonst} \end{cases}$$

$$P(A \cup B) = \begin{cases} P(A) + P(B) & \text{falls } A, B \text{ disjunkt} \\ P(A) + P(B) & \text{sonst} \\ -P(A \cap B) \end{cases}$$

Outline



- 1. Grundlegende Definitionen
- 2. Der Wahrscheinlichkeitsbegriff
- 3. Der allgemeine Additionssatz
- 4. Bedingte Wahrscheinlichkeiten
- 5. Der Multiplikationssatz
- 6. Unabhängigkeit von Ereignissen
- 7. Der Satz der totalen Wahrscheinlichkeit
- 8. Die Formel von Bayes

Der Satz der totalen Wahrscheinlichkeit



In der Praxis treten die gesuchten Ereignisse A oft unter verschiedenen "Teilbedingungen" $E_1, ..., E_n$ auf.

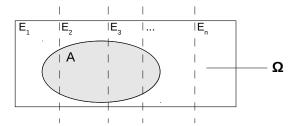
Beispiel

- ► *A* = "Studi besteht Klausur"
- $ightharpoonup E_1, ..., E_n =$ "Studi studiert AI, ITS, WI, AI-dual, ..."

Mengentheoretische Sichtweise

Berechne P(A), wenn A sich aus Teilmengen

 $A \cap E_1$, $A \cap E_2$, ..., $A \cap E_n$ zusammensetzt.



Der Satz der totalen Wahrscheinlichkeit



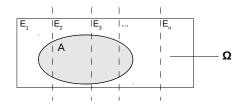
Satz (Totale Wahrscheinlichkeit)

Es sei (Ω, Σ, P) ein Ereignisraum, und die Ereignisse $E_1, ..., E_n \in \Sigma$ bilden eine Partitionierung von Ω , d.h.:

$$E_1 \cup ... \cup E_n = \Omega \quad \text{und} \quad E_j \cap E_k = \emptyset \quad \text{für } 1 \leq j < k \leq n$$

Dann gilt für jedes Ereignis $A \in \Sigma$:

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



Totale Wahrscheinlichkeit: Beispiel



Ein Computerhersteller bezieht Festplatten mit unterschiedlichen Ausschussraten von drei verschiedenen Lieferanten:

	Lieferant 1	Lieferant 2	Lieferant 3
Anteil	40%	25%	35%
Ausschuss	2%	1%	3%

Wir bestimmen die Wahrscheinlichkeit, dass eine beliebige gelieferte Festplatte einen Defekt aufweist:

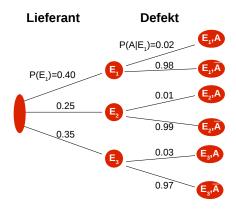
- ► *A* = "Festplatte weist einen Defekt auf."
- $ightharpoonup E_1, E_2, E_3 =$ "Festplatte stammt von Lieferant 1,2,3."





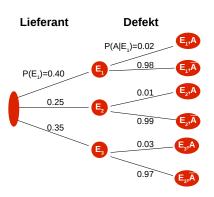
Wir können uns das Beispiel auch als ein grafisches Modell – einen sogenannten Ereignisbaum – vorstellen:

	Lieferant 1	Lieferant 2	Lieferant 3
Anteil	40%	25%	35%
Ausschuss	2%	1%	3%



Totale Wahrscheinlichkeit: Ereignisbäume



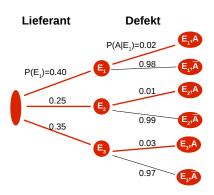


Wir können $P(E_k, A)$ berechnen, indem wir die Wahrscheinlichkeiten entlang der Pfade aufmultiplizieren:

- $P(E_1, A) = P(E_1) \cdot P(A|E_1) = 0.4 \cdot 0.02 = 0.008$
- $P(E_2, \overline{A}) = P(E_2) \cdot P(\overline{A}|E_2) = 0.25 \cdot 0.99 = 0.2475$
- **.**..

Totale Wahrscheinlichkeit: Ereignisbäume





Gemäß dem Satz der totalen Wahrscheinlichkeit erhalten wir P(A), indem wir die Wahrscheinlichkeiten aller Pfade, die A umfasst, aufsummieren!

$$P(A) = \underbrace{0.40 \cdot 0.02}_{\text{Pfad 1}} + \underbrace{0.25 \cdot 0.01}_{\text{Pfad 2}} + \underbrace{0.35 \cdot 0.03}_{\text{Pfad 3}} = 0.0201$$





Wir können das Prinzip des grafischen Modells auf weitere Ebenen erweitern. Zum Beispiel mit 3 Ebenen:

Satz (Satz der totalen Wahrscheinlichkeit (3 Ebenen))

Sei (Ω, Σ, P) ein Ereignisraum, $A \in \Sigma$, und $B_1, ..., B_n \in \Sigma$ und $C_1, ... C_m \in \Sigma$ Partitionierungen von Ω . Dann gilt:

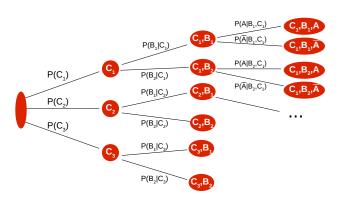
$$P(A) = \sum_{i,j} P(A, B_i, C_j)$$

$$= \sum_{i,j} P(B_i, C_j) \cdot P(A|B_i, C_j)$$

$$= \sum_{i,j} P(C_i) \cdot P(B_i|C_j) \cdot P(A|B_i, C_j).$$

Totale Wahrscheinlichkeit: Erweiterung





Um P(A) zu berechnen...

- ▶ Berechne für die Pfade die Wahrscheinlichkeiten $P(A, B_i, C_j)$ mittels Aufmultiplizieren von der Wurzel bis zum jeweiligen Blatt.
- Summiere Wahrscheinlichkeiten der Pfade zu A auf.

Totale Wahrscheinlichkeiten: Beispiel Bild: [4]

In einer Show ziehen drei Spieler nacheinander je einen von vier Umschlägen. Drei Umschläge enthalten Nieten, einer den Hauptpreis (ein Auto).

Wir definieren $A_1/A_2/A_3 :=$ "Spieler 1/2/3 zieht das Auto".

- ► Skizzieren Sie den Ereignisbaum.
- ▶ Mit welcher Wahrscheinlichkeit gewinnt Spieler 3 das Auto?

Totale Wahrscheinlichkeiten: Beispiel (cont'd)



Outline



- 1. Grundlegende Definitionen
- 2. Der Wahrscheinlichkeitsbegriff
- 3. Der allgemeine Additionssatz
- 4. Bedingte Wahrscheinlichkeiten
- Der Multiplikationssatz
- 6. Unabhängigkeit von Ereignissen
- 7. Der Satz der totalen Wahrscheinlichkeit
- 8. Die Formel von Bayes

Die Formel von Bayes Bild: [17]



Satz (Die Formel von Bayes)

Sei (Ω, Σ, P) ein Ereignisraum, $A \in \Sigma$ mit $P(A) \neq 0$, und $B_1, ..., B_n \in \Sigma$ eine Partitionierung von Ω . Dann gilt für alle i = 1, ..., n (falls $P(B_i) \neq 0$):

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

Warum ist die Formel von Bayes (extrem) nützlich?

Oft kennen wir die Verteilung "in eine Richtung" $P(A|B_i)$ und wollen die Verteilung in die andere Richtung $P(B_i|A)$ ermitteln. Dies ist mit der Formel von Bayes' direkt möglich.



Die Formel von Bayes: Beweis



Die Formel von Bayes



Anmerkungen

Laut der obigen Form der Bayes'schen Formel benötigen wir noch P(A). Diesen Term können wir mittels der totalen Wahrscheinlichkeit ermitteln $\left(P(A) = \sum_{j=1}^{n} P(B_j) \cdot P(A|B_j)\right)$ und erhalten:

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

▶ Ein Spezialfall entsteht, falls die Partitionierung nur in zwei Teile B, \overline{B} vorgenommen wird:

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

Beispiel: Medizinischer Test [15]



Ein medizinischer Test soll entscheiden, ob eine Person mit Tuberkulose infiziert ist. Wir wissen, dass der Test in 99% der Fälle die korrekte Antwort liefert, sowohl für gesunde als auch für kranke Patienten. Man weiß außerdem, dass 0.1% der Bevölkerung infiziert sind. Eine Person lässt sich testen, der Test fällt positiv aus. Wie groß ist die Wahrscheinlichkeit, dass die Person in der Tat mit Tuberkulose infiziert ist?

► Ereignis *K* := "Person ist krank" (=mit Tuberkulose infiziert)

► Ereignis T := "Der Test schlägt an"

Beispiel: Medizinischer Test



Beispiel 2: Spam-Filter



Ist eine E-Mail Spam oder "echt" ("ham")?

► Wir definieren das Ereignis

$$S :=$$
 "Die Mail ist Spam"

 Idee: Das Auftreten bestimmter Schlüsselworter (wie z.B. viagra) dient als Indikator für Spam-Mails.
 Wir definieren also das Ereignis

▶ 19% aller Spam-Mails (aber nur 1% der Ham-Mails) enthalten das Wort "viagra":

$$P(V|S) = 0.19 \text{ und } P(V|\bar{S}) = 0.01$$

▶ Wir wissen außerdem, dass 25% aller Mails SPAM sind:

$$P(S) = 0.25$$

Beispiel 2: Spam-Filter



Bauen Sie mit Hilfe der Bayes'schen Formel einen (einfachen) Spam-Filter:

- ▶ Berechnen Sie P(S|V)!
- ▶ Berechnen Sie $P(S|\bar{V})!$

Beispiel 2: Spam-Filter



Spam-Filter: Ausblick***



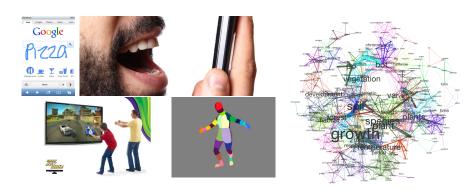
- Kommt "viagra" in der Mail vor, würde sich unser (sehr simpler) Filter also für "Spam" entscheiden, ansonsten nicht.
- ► In der Praxis stützen wir unsere Entscheidung nicht auf einen Begriff (viagra), sondern auf viele!
- ► Hierzu definieren wir einen Vokabular von Worten (oder Termen) w₁,..., w_m (z.B. viagra, nigeria, dollar, Papa, Geburtstag, ...). Potenziell kann dieser Grundvokabular groß sein und sich dynamisch ändern.



Bayes im Machine Learning Bilder: [18] [12] [14] [16]



- ▶ Die Bayes'sche Regel bildet im maschinellen Lernen die Grundlage vieler Modelle.
- Spam-Filter sind nur ein Beispiel: Es gibt zahlreiche andere Anwendungen.



References I



- A surface weather analysis for the United States on October 21, 2006. https://en.wikipedia.org/wiki/Weather_map (retrieved: Oct 2016).
- [2] Andreas Hopf: Fair Dice, View 2 (Cast galvanised iron fair dice, numbers engraved inlaid black). https://flic.kr/p/5SunDf (retrieved: Nov 2016).
- [3] DieWespe: HDR-HIT-ACTUAL. https://flic.kr/p/qBV7S (retrieved: Nov 2016).
- [4] kfant: Zonk (Ich wollte früher auch immer so einen haben :D). https://flic.kr/p/yn1se (retrieved: Nov 2016).
- [5] Lany Lane's Photo Stream: Envie de plein de chose. https://flic.kr/p/bx7psH (retrieved: Nov 2016).
- [6] Olli Henze: 4 Asse. http://www.ohenze.de/ (retrieved: Nov 2016).
- [7] Ulrica Törning: Yatzy. https://flic.kr/p/84JVjL (retrieved: Nov 2016).
- [8] Maurice de Bévère (Morris). Die Dalton-Brüder. In <u>Lucky Luke</u>, 1949.
- [9] Sophie Feytaud. Pierre Simon Laplace.

https://commons.wikimedia.org/w/index.php?curid=20879 (retrieved: Nov 2016). This image appears identical t the cover image used by Gillispie et al. They cite the portrait as an 1842 posthumous portrait by Madame Feytaud, courtesy of the Académie des Sciences, Paris.

References II



[10] Erlangen Konrad Jacobs.

Richard von Mises.

https://commons.wikimedia.org/w/index.php?curid=6995231 (retrieved: Nov 2016), http://owpdb.mfo.de/detail?photo_id=2896. CC BY-SA 2.0 de.

[11] Erlangen Konrad Jacobs.

Richard von Mises.

https://commons.wikimedia.org/w/index.php?curid=11829175 (retrieved: Nov 2016), http://owpdb.mfo.de/detail?photoID=7493. CC BY-SA 2.0 de.

[12] Microsoft.

Kinect Joy Ride Xbox 360 Game for Kinect.

https://www.microsoftstore.com/store/msusa/en_US/pdp/ Kinect-Joy-Ride-Xbox-360-Game-for-Kinect/productID.253660600 (retrieved: Nov 2016).

[13] L. Papula.

Mathematik für Ingenieure und Naturwissenschaftler, volume 3.

Friedrich Vieweg & Sohn Verlagsgesellschaft mbH. Braunschweig/Wiesbaden, 4 edition, 2001.

- [14] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake. Real-time Human Pose Recognition in Parts from Single Depth Images. In Proc. CVPR, pages 1297–1304, 2011.
- [15] G. Teschl and S. Teschl. Mathematik Für Informatiker, volume 2 of eXamen.press. Springer-Verlag, 2 edition, 2007.
- [16] Tethne Tutorials.

Generating and Visualizing Topic Models with Tethne and MALLET. http://diging.github.io/tethne/api/tutorial.mallet.html (retrieved: Nov 2016).

References III



[17] Unbekannt.

Thomas Bayes (vermutlich nicht authentisch). https://commons.wikimedia.org/w/index.php?curid=14532025 (retrieved: Nov 2016). Gemeinfrei.

[18] VeryPDF.

What is OCR (Optical Character Recognition)?

http://www.verypdf.com/pdf2txt/what-is-ocr2.htm (retrieved: Nov 2016).