⟨ S t a t W t h 1 7 ⟩

2. Wahrscheinlichkeit

Werner Gurker

Institut für Stochastik und Wirtschaftsmathematik

Technische Universität Wien

Empirisches Gesetz der großen Zahlen (eGGZ)

Ein Zufallsexperiment wird n Mal wiederholt und $H_n(A)$ sei die absolute Häufigkeit eines bestimmten Ereignisses A.

Häufig hat man den Eindruck, dass sich mit wachsendem n die relative Häufigkeit $h_n(A) := H_n(A)/n$ von A einem Grenzwert nähert:

$$\lim_{n\to\infty} h_n(A) \ \rightsquigarrow \ P(A)$$

Es liegt nahe, den Grenzwert $P(A)$ als Wahrscheinlichkeit (des Eintritts) von A zu bezeichnen.

Beispiel zum empirischen GGZ

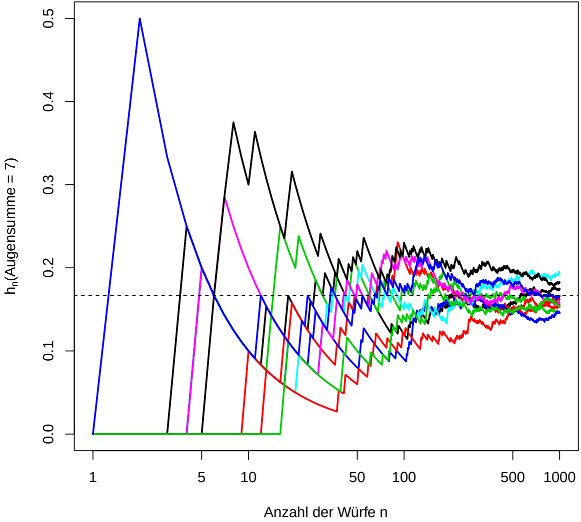
Zwei Würfel werden geworfen und A sei das Ereignis, dass die Augensumme gleich sieben ist:

$$A = \{\text{Augensumme} = 7\}$$

Wir betrachten 10 simulierte Wurffolgen zu je 1000 Würfen.

Am Anfang ist die Fluktuation von $h_n(A)$ noch hoch, wird aber mit wachsendem n geringer und $h_n(A)$ scheint sich einem Grenzwert zu nähern.

Relative Häufigkeit der Augensumme 7



$$\lim_{n\to\infty} h_n(A) \rightsquigarrow \frac{1}{6}$$

Subjektive Wahrscheinlichkeiten

Der frequentistische (oder objektivistische) Wahrscheinlichkeitsbegriff beruht auf der Annahme, dass ein statistisches Experiment unter (mehr oder weniger) identischen Bedingungen beliebig oft wiederholbar ist. Das ist aber nicht immer der Fall.

Bsp: Ein Wettermoderator behauptet, dass die Wahrscheinlichkeit für Regen während eines Freiluftkonzerts 30% beträgt.

Diese Form von subjektiver Wahrscheinlichkeit lässt sich nicht als Grenzwert von relativen Häufigkeiten interpretieren. Vielmehr handelt es sich um einen – auf persönlicher Expertise basierenden – Grad des Vertrauens in die Behauptung.

Merkmalraum

Die möglichen Ergebnisse von statistischen Experimenten lassen sich in einer Grundmenge zusammenfassen. Die Menge aller möglichen Versuchsausgänge nennt man den Merkmalraum:

$$\Omega = \Big\{\omega \,\Big|\, \omega \text{ möglicher Versuchsausgang}\Big\}$$

Ein Merkmalraum kann endlich, unendlich – abzählbar unendlich oder überabzählbar, ein– oder mehrdimensional sein.

Beispiel 1

Wirft man zwei Münzen, wobei eine die „erste“ und die andere die „zweite“ ist, besteht der Merkmalraum aus vier Elementen:

$$\Omega = \big\{(H, H), (H, T), (T, H), (T, T)\big\}$$

Dabei bedeutet z. B. (T, H), dass die erste Münze auf „Zahl“ (*tail*) und die zweite Münze auf „Kopf“ (*head*) fällt.

Sind die Münzen symmetrisch, ist es naheliegend, den einzelnen Versuchsausgängen die gleiche Wahrscheinlichkeit $1/4$ zuzuordnen.

Beispiel 2

Wirft man zwei Würfel, wobei einer der „erste“ und der andere der „zweite“ ist, und beobachtet die geworfenen Augenzahlen, kann man als Merkmalraum die Menge aller 2–Tupel nehmen:

$$\Omega_1 = \big\{(i,j)\,\big|\, i,j = 1,2,3,4,5,6\big\}, \quad |\Omega_1| = 36$$

Dabei bedeutet das Paar (i,j), dass der erste Würfel auf i und der zweite auf j fällt.

Sind die Würfel symmetrisch, wird man den einzelnen Versuchsausgängen die gleiche Wahrscheinlichkeit $1/36$ zuordnen.

Beispiel 2 (Forts.)

Interessiert nur die Augensumme, kann man auch den folgenden Merkmalraum nehmen:

$$\Omega_2 = \{2, 3, \dots, 12\}, \quad |\Omega_2| = 11$$

Ein Nachteil dieses Merkmalraums besteht darin, dass die einzelnen Elemente $\omega \in \Omega_2$ nicht „gleichwahrscheinlich" sind, die des zuerst betrachteten Raumes Ω_1 aber schon.

Beispiel 3

Besteht das statistische Experiment im Messen der Lebensdauer (in Betriebsstunden) einer elektronischen Komponente, nimmt man als Merkmalraum ein halbunendliches Intervall:

$$\Omega = \left\{x \,\middle|\, 0 \leq x < \infty\right\} = [0, \infty)$$

Dabei handelt es sich um ein stetiges Merkmal, da – prinzipiell – jedes Element von $[0, \infty)$ als Versuchsausgang in Frage kommt.

Da viele Komponenten gleich bei der ersten Inbetriebnahme ausfallen können, sollte 0 ein Element des Merkmalraums sein.

Beispiel 4

Das statistische Experiment bestehe in der zufälligen Auswahl von k Objekten aus einer Menge von n (unterscheidbaren) Objekten, bezeichnet mit $M = \{1, 2, \ldots, n\}$ $(k \leq n)$.

Die Auswahl erfolge in der Weise, dass ein einmal gewähltes Objekt nicht noch einmal gewählt werden kann. Für die Wahl eines passenden Merkmalraums hat man zwei Möglichkeiten.

- Spielt die Reihenfolge der Auswahl keine Rolle, nimmt man die Menge aller k–elementigen Teilmengen:

$$\Omega_1 = \left\{ B \,\middle|\, B \subseteq M,\ |B| = k \right\}, \quad |\Omega_1| = \binom{n}{k}$$

Beispiel 4 (Forts.)

- Soll die Reihenfolge der Auswahl berücksichtigt werden, nimmt man:

$$\Omega_2 = \left\{(x_1, \dots, x_k) \,\middle|\, x_i \in M;\; x_i \neq x_j \text{ für } i \neq j\right\}, \quad |\Omega_2| = \frac{n!}{(n-k)!}$$

Da in beiden Fällen die Merkmalräume aus „gleichwahrscheinlichen“ Elementen bestehen, können Wahrscheinlichkeiten auf einfache Weise zugeordnet werden.

Ereignisse

Allgemein nennt man Teilmengen $A \subseteq \Omega$ eines Merkmalraumes Ω Ereignisse.

Gilt für einen Versuchsausgang $\omega \in \Omega$, dass $\omega \in A$, sagt man, dass A eingetreten ist.

Die Ereignisse werden in einem Ereignissystem $\mathcal{A}$ zusammengefasst:

$$\mathcal{A} = \left\{ A \,\middle|\, A \subseteq \Omega \text{ ist ein Ereignis} \right\}$$

Zusammengesetzte Ereignisse

Um neben einfachen Ereignissen $A, B, \ldots$ auch zusammengesetzte Ereignisse wie beispielsweise:

A und B treten ein

A oder B treten ein

A tritt nicht ein

etc.

betrachten zu können, benötigt das Ereignissystem eine entsprechende algebraische Struktur.

Ereignissysteme sind σ–Algebren

Ein System $\mathcal{A}$ von Teilmengen des Merkmalraums Ω nennt man eine σ–Algebra über Ω, wenn gilt:

(1) $\Omega \in \mathcal{A}$ (d. h., der Merkmalraum selbst ist ein Ereignis)

(2) Für $A \in \mathcal{A}$ gilt $A^c \in \mathcal{A}$ (d. h., $\mathcal{A}$ ist abgeschlossen unter Komplementbildung)

Bem: Für das Komplement A^c von A schreibt man auch $\overline{A}$.

(3) Für eine Folge $A_1, A_2, \ldots$ aus $\mathcal{A}$ gilt $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ (d. h., $\mathcal{A}$ ist abgeschlossen unter abzählbaren Vereinigungen)

Messraum

Ist Ω ein Merkmalraum und $\mathcal{A}$ eine σ–Algebra über Ω, nennt man das Paar $(\Omega, \mathcal{A})$ einen Messraum.

Die Elemente von $\mathcal{A}$ nennt man messbare Mengen.

De Morgan'sche Regeln

Für Teilmengen $A_1, A_2, \ldots$ von Ω gilt:

$$\left(\bigcap_{i=1}^{\infty} A_i\right)^c = \bigcup_{i=1}^{\infty} A_i^c \quad \text{und} \quad \left(\bigcup_{i=1}^{\infty} A_i\right)^c = \bigcap_{i=1}^{\infty} A_i^c$$

Folgerungen

(1) Das sichere und das unmögliche Ereignis sind Elemente jedes Ereignissystems $\mathcal{A}$, d. h. $\Omega \in \mathcal{A}$ und $\emptyset \in \mathcal{A}$.

(2) Ereignissysteme sind abgeschlossen gegenüber abzählbaren Durchschnitten:

$$A_1, A_2, \ldots \in \mathcal{A} \quad \Longrightarrow \quad \bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$$

(3) Die Abgeschlossenheit gegenüber Vereinigungen und Durchschnitten gilt auch für endlich viele Ereignisse:

$$A_1, A_2, \ldots, A_n \in \mathcal{A} \quad \Longrightarrow \quad \bigcup_{i=1}^{n} A_i \in \mathcal{A}, \quad \bigcap_{i=1}^{n} A_i \in \mathcal{A}$$

Folgerungen (Forts.)

(4) Einzelne Versuchsausgänge $\omega \in \Omega$ sind *keine* Ereignisse! Das entsprechende Ereignis lautet korrekt $\{\omega\} \subset \Omega$.

Bem: Einelementige Ereignisse $\{\omega\}$ sind nicht automatisch auch Elemente des Ereignissystems! In der Praxis geht man aber davon aus, dass $\{\omega\} \in \mathcal{A}$ für alle $\omega \in \Omega$.

(5) Die kleinste σ–Algebra über Ω besteht nur aus dem unmöglichen und dem sicheren Ereignis: $\mathcal{A} = \{\emptyset, \Omega\}$.

(6) Die größte σ–Algebra über Ω ist die Potenzmenge, d. h. die Menge aller Teilmengen des Merkmalraums:

$$\mathcal{A} = \{A \,|\, A \subseteq \Omega\} = \mathcal{P}(\Omega)$$

Festlegung 1

Ist der Merkmalraum Ω endlich oder abzählbar unendlich, d. h., ist er höchstens abzählbar, wählt man stets die größte σ–Algebra, d. h. die Potenzmenge von Ω, als Ereignissystem.

Der Messraum ist in diesem Fall gegeben durch:

$$\big(\Omega, \mathcal{A} = \mathcal{P}(\Omega)\big)$$

Beispiel 1

Das statistische Experiment bestehe im Werfen eines Würfels und im Bestimmen der geworfene Augenzahl.

Der Merkmalraum ist gegeben durch $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Das zugehörige Ereignissystem ist nach Festlegung 1 die Potenzmenge von Ω:

$$\mathcal{A} = \mathcal{P}(\Omega) = \{\emptyset, \{1\}, \{2\}, \ldots, \{6\}, \{1, 2\}, \{1, 3\}, \ldots, \Omega\}$$

Das Ereignissystem besteht aus 64 Elementen (Teilmengen von Ω):

$$\left|\mathcal{A}\right| = \left|\mathcal{P}(\Omega)\right| = 2^6 = 64$$

Beispiel 1 (Forts.)

Beispielsweise lässt sich das Ereignis, dass die geworfene Augenzahl eine gerade Zahl ist, wie folgt formulieren:

$$A = \{2, 4, 6\} \quad \widehat{=} \quad \text{Augenzahl gerade}$$

Das Ereignis, dass die Augenzahl ungerade ist, lässt sich durch das Komplement von A ausdrücken:

$$A^c = \{1, 3, 5\} \quad \widehat{=} \quad \text{Augenzahl ungerade}$$

Beispiel 2

Ein Würfel werde solange geworfen, bis zum ersten Mal ein „Sechser“ kommt.

In diesem Fall ist $\Omega = \{1, 2, \dots\} = \mathbb{N}$ der Merkmalraum.

Als Ereignissystem wählt man nach Festlegung 1 die Potenzmenge von Ω:

$$\mathcal{A} = \mathcal{P}(\Omega)$$

$\mathcal{A}$ ist nicht mehr abzählbar.

Nach der allgemeinen Kontinuumshypothese gilt:

$$\left|\mathcal{A}\right| = \left|\mathcal{P}(\mathbb{N})\right| = 2^{|\mathbb{N}|} = |\mathbb{R}|$$

Erzeugte σ–Algebra

Ist Ω der Merkmalraum eines statistischen Experiments, so versteht man unter der von einem Mengensystem:

$$\mathcal{G} = \{A : A \subseteq \Omega\}$$

erzeugten σ–Algebra die kleinste σ–Algebra, die die Elemente von $\mathcal{G}$ umfasst.

Beispiel

Wirft man einen Würfel und beobachtet die Augenzahl, so lautet der Merkmalraum $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Besteht das Mengensystem $\mathcal{G}$ beispielsweise nur aus einer Menge:

$$\mathcal{G} = \big\{\{1, 2\}\big\}$$

so ist die von $\mathcal{G}$ erzeugte σ–Algebra gegeben durch:

$$\mathcal{A} = \big\{\emptyset, \{1, 2\}, \{3, 4, 5, 6\}, \Omega\big\}$$

Borelmengen

Der Merkmalraum sei ganz $\mathbb{R}$, d. h. $\Omega = \mathbb{R}$.

Als Mengensystem $\mathcal{G}$ nehmen wir das System der links offenen und rechts abgeschlossenen (endlichen) Intervalle:

$$\mathcal{G} = \big\{(a, b] \,\big|\, a, b \in \mathbb{R} \text{ mit } a \leq b\big\} \qquad \longleftarrow \text{ keine } \sigma\text{–Algebra!}$$

Die kleinste von $\mathcal{G}$ erzeugte σ–Algebra $\mathcal{B}$, die alle Elemente von $\mathcal{G}$ umfasst, nennt man die Borel σ–Algebra oder die Borelmengen.

$\mathcal{B}$ ist eine echte Teilmenge der Potenzmenge von $\mathbb{R}$:

$$\mathcal{B} \subset \mathcal{P}(\mathbb{R})$$

Welche Teilmengen von $\mathbb{R}$ sind Borelmengen?

Allgemein gilt, dass alle Teilmengen von $\mathbb{R}$, die man sich vorstellen oder zeichnen kann, Borelmengen sind.

Speziell sind alle Arten von Intervallen Borelmengen:

$$(a, b], \quad (a, b), \quad [a, b], \quad (-\infty, b], \quad (a, \infty), \quad \ldots$$

Auch alle einpunktigen Mengen:

$$\{a\}, \quad a \in \mathbb{R}$$

sind Borelmegen.

Gibt es überhaupt Teilmengen von $\mathbb{R}$, die keine Borelmengen sind?

Wie bereits erwähnt, gilt $\mathcal{B} \subset \mathcal{P}(\mathbb{R})$.

Es gibt also Teilmengen von $\mathbb{R}$, die keine Borelmengen sind.

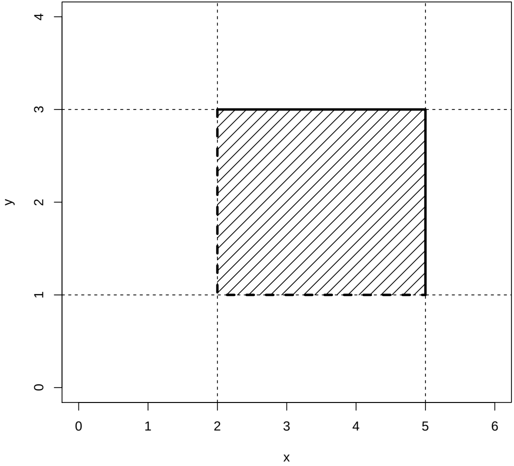
Derartige Mengen entziehen sich allerdings weitgehend der Vorstellung.

Darüber hinaus gilt:

$$\left|\mathcal{B}\right| = \left|\mathbb{R}\right|, \quad \left|\mathcal{P}(\mathbb{R})\right| \succ \left|\mathbb{R}\right|$$

D. h. aber, dass es *sehr viel mehr* Teilmengen von $\mathbb{R}$ gibt, die keine Borelmengen sind, als es Borelmengen gibt!

Beispiel für einen halboffenen Quader



Mehrdimensionale Borelmengen

Analog zum 1–dimensionalen Fall lässt sich die Borel σ–Algebra über $\mathbb{R}^k$ (für $k \geq 2$) definieren.

Beispielsweise besteht für $k = 2$ das erzeugende Mengensystem aus den halboffenen Quadern:

$$\mathcal{G}_2 = \left\{(a, b] \times (c, d] \,\middle|\, a, b, c, d \in \mathbb{R} \text{ mit } a \leq b,\ c \leq d\right\}$$

Die 2–dimensionalen Borelmengen $\mathcal{B}_2$ sind dann definiert als die kleinste σ–Algebra, die alle Mengen aus $\mathcal{G}_2$ umfasst.

Analog definiert man allgemein die k–dimensionalen Borelmengen $\mathcal{B}_k$.

Festlegung 2

Ist der Merkmalraum Ω eine überabzählbare Teilmenge von $\mathbb{R}^k$ $(k \geq 1)$, z. B. ein ein- oder mehrdimensionales Intervall, wählt man als Ereignissystem stets die entsprechende Borel σ–Algebra.

Der Messraum lautet in diesem Fall:

$$\big(\Omega, \mathcal{A} = \mathcal{B}(\Omega)\big)$$

$$\mathcal{B}(\Omega) \ = \ \text{Borel } \sigma\text{–Algebra über } \Omega$$

Alle praktisch relevanten Ereignisse werden dadurch erfasst.

Wahrscheinlichkeitsmaß

Gegeben sei ein Messraum $(\Omega, \mathcal{A})$.

Eine Abbildung $P : \mathcal{A} \longrightarrow \mathbb{R}$ heißt ein Wahrscheinlichkeitsmaß (kurz: W–Maß) auf $(\Omega, \mathcal{A})$, wenn sie die folgenden Eigenschaften erfüllt:

(1) $P(A) \geq 0$ für alle $A \in \mathcal{A}$

(2) $P(\Omega) = 1$

(3) Für eine Folge $A_1, A_2, \ldots$ von (paarweise) disjunkten Ereignissen (d. h. $A_i \cap A_j = \emptyset$ für $i \neq j$) gilt die σ–Additivität:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Wahrscheinlichkeitsraum

Durch ein W–Maß P lässt sich ein Messraum $(\Omega, \mathcal{A})$ zum Wahrscheinlichkeitsraum (kurz: W–Raum) erweitern:

$$\text{W–Raum:} \quad \big(\Omega, \mathcal{A}, P\big)$$

Behauptung 1

$$P(\emptyset) = 0$$

Beweis: Schreibe Ω wie folgt:

$$\Omega = \underbrace{\Omega \cup \emptyset \cup \emptyset \cup \cdots}_{\text{paarweise disjunkt}}$$

Nach (2) und (3) gilt:

$$\underbrace{P(\Omega)}_{1} = \underbrace{P(\Omega)}_{1} + P(\emptyset) + P(\emptyset) + \cdots \quad \Longrightarrow \quad P(\emptyset) = 0$$

Behauptung 2

Für eine endliche Folge $A_1, A_2, \ldots, A_n$ von (paarweise) disjunkten Ereignissen gilt:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

Beweis: Folgt aus (3) und Behauptung 1: Setze $A_i = \emptyset$, $i > n$.

Behauptung 3

$$P(A^c) = 1 - P(A)$$

Beweis: A und A^c sind disjunkt und $A \cup A^c = \Omega$; nach (2) und (3) (oder Behauptung 2) gilt:

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c)$$

Behauptung 4

$$A \subseteq B \quad \Longrightarrow \quad P(A) \leq P(B)$$

Beweis:

$$A \subseteq B \quad \Longrightarrow \quad B = \underbrace{A \cup (A^c \cap B)}_{\text{disjunkt}}$$

Mit Behauptung 2 folgt:

$$P(B) = P(A) + \underbrace{P(A^c \cap B)}_{\geq 0} \quad \Longrightarrow \quad P(B) \geq P(A)$$

Behauptung 5

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

Beweis:

$$A \cup B = \underbrace{A \cup (A^c \cap B)}_{\text{disjunkt}} \quad \Longrightarrow \quad P(A \cup B) = P(A) + P(A^c \cap B) \;\; (*)$$

$$B = \underbrace{(A \cap B) \cup (A^c \cap B)}_{\text{disjunkt}} \quad \Longrightarrow \quad P(A^c \cap B) = P(B) - P(A \cap B) \;\; (**)$$

$$(*) \;\wedge\; (**) \quad \Longrightarrow \quad \text{Beh.}$$

Chancen (Odds)

Die Chancen (*engl.* Odds) $o(A)$ für den Eintritt des Ereignisses A sind definiert als der Quotient aus Wahrscheinlichkeit $P(A)$ und Gegenwahrscheinlichkeit $P(A^c) = 1 - P(A)$:

$$o(A) = \frac{P(A)}{1 - P(A)}$$

Umgekehrt:

$$P(A) = \frac{o(A)}{1 + o(A)}$$

Chancenverhältnis (Odds–Ratio)

Die Chancen $o(A)$ und $o(B)$ von zwei Ereignissen A und B werden häufig durch das Chancenverhältnis (*engl.* Odds–Ratio) miteinander verglichen:

$$r(A, B) = \frac{o(A)}{o(B)} = \frac{P(A) \big/ \big[1 - P(A)\big]}{P(B) \big/ \big[1 - P(B)\big]}$$

Beispiel

Die Gewinnwahrscheinlichkeit bei einem Spiel sei 0.75. Die Chancen, das Spiel zu gewinnen, stehen also 75 zu 25 oder 3 zu 1 oder $o = 3$. D. h., zu gewinnen ist dreimal wahrscheinlicher als zu verlieren.

Bei einem anderen Spiel beträgt die Gewinnwahrscheinlichkeit 0.9. Die Chancen, dieses Spiel zu gewinnen, stehen also 9 zu 1 oder $o = 9$. D. h., zu gewinnen ist neunmal wahrscheinlicher als zu verlieren.

Die Odds-Ratio beträgt $r = 9/3 = 3$. D. h., die Gewinnchancen sind beim zweiten Spiel um den Faktor 3 günstiger.

Endliche W-Räume

Ein W-Raum $(\Omega, \mathcal{P}(\Omega), P)$ heißt endlich, wenn der Merkmalraum Ω eine endliche Menge ist.

Für endliche W–Räume ist das W–Maß P durch die Wahrscheinlichkeiten der Elementarereignisse $\{\omega\}$, $\omega \in \Omega$, eindeutig bestimmt:

$$P\big(\{\omega\}\big) \geq 0 \quad \forall \omega \in \Omega \qquad \text{und} \qquad \sum_{\omega \in \Omega} P\big(\{\omega\}\big) = 1$$

Wahrscheinlichkeit eines beliebigen Ereignisses $A \in \mathcal{P}(\Omega)$:

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

Laplace–Raum

Sind für einen endlichen W–Raum $(\Omega, \mathcal{P}(\Omega), P)$ die Elementarereignisse $\{\omega\}$ gleichwahrscheinlich, nennt man den W–Raum einen Laplace–Raum.

In diesem Fall gilt:

$$P\big(\{\omega\}\big) = \frac{1}{|\Omega|} \quad \text{für alle} \quad \omega \in \Omega$$

Wahrscheinlichkeit eines beliebigen Ereignisses $A \in \mathcal{P}(\Omega)$:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der Elemente von } A}{\text{Anzahl der Elemente von } \Omega}$$

Klassische W–Definition

Die übliche Sprechweise im Laplace–Raum lautet:

Die Wahrscheinlichkeit für ein Ereignis A ist der Quotient aus der Zahl der für (den Eintritt von) A günstigen Fälle und der Zahl der möglichen Fälle:

$$P(A) = \frac{\text{Zahl der für } A \text{ günstigen Fälle}}{\text{Zahl der möglichen Fälle}} = \frac{g}{m}$$

Das nennt man auch die klassische W–Definition.

Achtung: Diese W–Definition gilt nur im Laplace–Raum !

Abzählende Kombinatorik

Bei Anwendungen der klassischen W–Definition spielt naturgemäß das Zählen eine große Rolle:

Wieviele Elemente hat der Merkmalraum?

Wieviele Elemente hat ein bestimmtes Ereignis?

In einfachen Fällen können die Elemente direkt abgezählt werden, in den meisten Fällen muss man aber auf Methoden der abzählenden Kombinatorik zurückgreifen:

$\longrightarrow$ Permutationen, Kombinationen, Variationen, ...

(Vgl. Skriptum/Buch S. 107f für eine kurze Zusammenfassung.)

Beispiel

Ein Würfel wird dreimal hintereinander geworfen. Mit welcher Wahrscheinlichkeit bilden die Augenzahlen eine monoton wachsende Folge?

Merkmalraum:

$$\Omega = \{(x_1, x_2, x_3) : x_i = 1, \ldots, 6\}, \quad |\Omega| = 6^3 = 216$$

$(\Omega, \mathcal{P}(\Omega), P)$ ist ein Laplace–Raum:

$$P(\{(x_1, x_2, x_3)\}) = \frac{1}{|\Omega|} = \frac{1}{216}$$

Beispiel (Forts.)

Folge strikt monoton: z. B. $(2, 5, 6)$

$$A_1 = \{(x_1, x_2, x_3) : x_1 < x_2 < x_3\}, \quad |A_1| = \binom{6}{3} = 20$$

$$P(A_1) = \frac{|A_1|}{|\Omega|} = \frac{20}{216} \doteq 0.0926$$

Folge monoton: z. B. $(2, 3, 3)$

$$A_2 = \{(x_1, x_2, x_3) : x_1 \leq x_2 \leq x_3\}, \quad |A_2| = \binom{6+3-1}{3} = 56$$

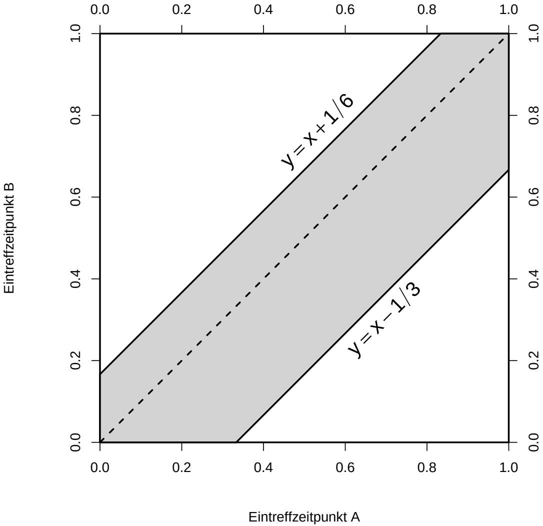
$$P(A_2) = \frac{|A_2|}{|\Omega|} = \frac{56}{216} \doteq 0.2593$$

Geometrische Wahrscheinlichkeiten

Die geometrische Wahrscheinlichkeitsdefinition lässt sich anwenden, wenn der Merkmalraum als geometrisches Objekt (z. B. Flächenstück) und Ereignisse als Teilbereiche interpretiert werden können, deren Wahrscheinlichkeit proportional zur Größe (z. B. Fläche) des Teilbereichs ist.

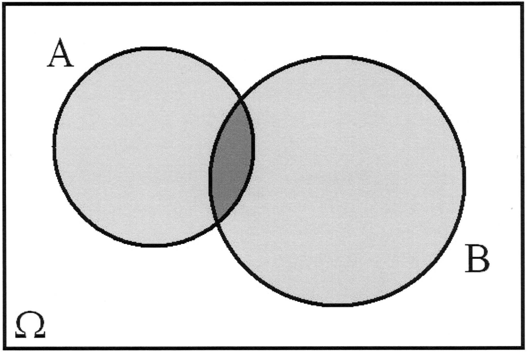
$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Größe von } A}{\text{Größe von } \Omega}$$

Rendevousproblem



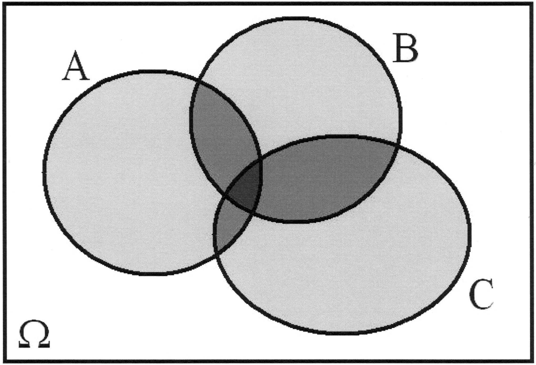
$$P(\text{Begegnung}) = \frac{|\text{graue Fläche}|}{|\text{gesamte Fläche}|}$$

Additionstheorem für zwei Ereignisse



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

Additionstheorem für drei Ereignisse



$$
\begin{aligned}
P(A \cup B \cup C) = & \; P(A) + P(B) + P(C) - \\
& - P(A \cap B) - P(A \cap C) - P(B \cap C) + \\
& + P(A \cap B \cap C)
\end{aligned}
$$

Allgemeines Additionstheorem

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{i_1 < i_2} P(A_{i_1} A_{i_2}) + \cdots$$

$$+ (-1)^{r+1} \sum_{i_1 < i_2 < \cdots < i_r} P(A_{i_1} A_{i_2} \cdots A_{i_r})$$

$$+ \cdots + (-1)^{n+1} P(A_1 A_2 \cdots A_n)$$

Boole'sche Ungleichung

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$$

Beweis:

$$\bigcup_{i=1}^{\infty} A_i = \underbrace{A_1}_{B_1} \cup \underbrace{(A_2 A_1^c)}_{B_2} \cup \underbrace{(A_3 A_1^c A_2^c)}_{B_3} \cup \cdots = \bigcup_{i=1}^{\infty} B_i$$

$$B_i \cap B_j = \emptyset,\ i \neq j \quad \text{und} \quad P(B_i) \leq P(A_i)$$

$$\sigma\text{–Add.} \Longrightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) \leq \sum_{i=1}^{\infty} P(A_i)$$

Bedingte Wahrscheinlichkeit

Gegeben sei ein W–Raum $(\Omega, \mathcal{A}, P)$ und $A, B \in \mathcal{A}$ seien zwei Ereignisse, wobei $P(B) > 0$.

Die bedingte Wahrscheinlichkeit von A gegeben B ist definiert durch:

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

Laplace–Raum: $P(A|B)$ der Anteil der für das Ereignis $A \cap B$ günstigen Fälle, bezogen auf die möglichen Fälle, die dem Ereignis B entsprechen:

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

$P(\cdot\,|B)$ ist ein W–Maß

Bedingte Wahrscheinlichkeiten haben alle Eigenschaften eines W–Maßes.

Für ein festes Ereignis B mit $P(B) > 0$ gilt:

(1) $P(A|B) \geq 0$ für alle $A \in \mathcal{A}$

(2) $P(\Omega|B) = 1$

(3) Für eine Folge $A_1, A_2, \ldots$ von (paarweise) disjunkten Ereignissen gilt:

$$P\left(\bigcup_{i=1}^{\infty} A_i \,\middle|\, B\right) = \sum_{i=1}^{\infty} P(A_i|B)$$

Multiplikationstheorem für zwei Ereignisse

Für $P(B) > 0$ gilt:

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

Für $P(A) > 0$ gilt:

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

Allgemeines Multiplikationstheorem

Für $A_1, A_2, \ldots, A_n \in \mathcal{A}$ mit $P(A_1 A_2 \cdots A_n) > 0$ gilt:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1})$$

Beweis: Anwendung der Definition der bedingten Wahrscheinlichkeit:

$$P(A_1)\,\frac{P(A_1A_2)}{P(A_1)}\,\frac{P(A_1A_2A_3)}{P(A_1A_2)}\,\cdots\,\frac{P(A_1A_2\cdots A_n)}{P(A_1A_2\cdots A_{n-1})} = P(A_1A_2\cdots A_n)$$

Geburtstagsproblem

Wenn sich in einem Raum n Personen befinden, mit welcher Wahrscheinlichkeit haben alle verschiedene Geburtstage?

Annahme: Jede Person hat mit gleicher Wahrscheinlichkeit an einem der 365 möglichen Tage Geburtstag.

Ereignis $A =$ Alle Geburtstage verschieden

$$P(A) = \underbrace{\left(\frac{365}{365}\right)}_{\text{1. Person}} \underbrace{\left(\frac{364}{365}\right)}_{\text{2. Person}} \cdots \underbrace{\left(\frac{365-n+1}{365}\right)}_{\text{n. Person}}$$

Dabei wird das Multiplikationstheorem verwendet.

Geburtstagsproblem (Forts.)

Schon für $n \geq 23$ ist $P(A) < \frac{1}{2}$.

Gibt es 23 oder mehr Personen, ist die Wahrscheinlichkeit dafür, dass zumindest zwei am gleichen Tag geboren sind, größer als 50% .

Das nennt man das „Geburtstagsparadoxon“.

Gibt es 41 oder mehr Personen, ist die Wahrscheinlichkeit dafür, dass zumindest zwei am gleichen Tag geboren sind, größer als 90% .

Partition des Merkmalraums

Ereignisse $C_1, C_2, \ldots \in \mathcal{A}$ bilden eine Partition des Merkmalraums Ω, wenn sie:

(1) paarweise disjunkt sind:

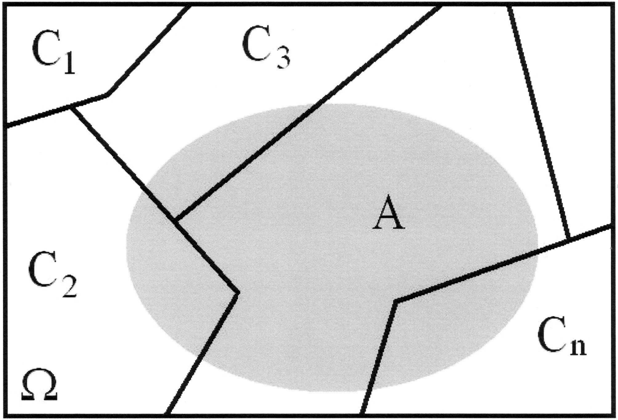
$$C_i \cap C_j = \emptyset \quad \text{für } i \neq j$$

(2) zusammen den Merkmalraum ergeben:

$$\bigcup_{i=1}^{\infty} C_i = \Omega$$

Partitionen $\{C_i\}$ können endlich oder abzählbar unendlich sein.

Venn–Diagramm zur vollständigen Wahrscheinlichkeit



Satz von der vollständigen Wahrscheinlichkeit

Ist $\{C_i\}$ eine Partition von Ω, lässt sich die Wahrscheinlichkeit für ein Ereignis A wie folgt berechnen:

$$P(A) = \sum_{i=1}^{\infty} P(A|C_i)P(C_i)$$

Beweis:

$$A = \bigcup_{i=1}^{\infty} \underbrace{(A \cap C_i)}_{\text{disjunkt}} \implies P(A) = \sum_{i=1}^{\infty} P(A \cap C_i) \quad (\sigma\text{–Add.})$$

$$\text{Multiplikationsth.} \quad \implies \quad P(A) = \sum_{i=1}^{\infty} P(A|C_i)P(C_i)$$

Beispiel

In einem Behälter befinden sich – gut gemischt – drei Typen von Batterien im Verhältnis 20 : 30 : 50.

$$\left\{\begin{array}{c}\text{Typ 1}\\ \text{Typ 2}\\ \text{Typ 3}\end{array}\right\} \text{ arbeitet mit W. } \left\{\begin{array}{c}0.7\\ 0.4\\ 0.3\end{array}\right\} \text{ länger als 100 Stunden}$$

Eine Batterie wird zufällig entnommen.

Mit welcher Wahrscheinlichkeit arbeitet diese Batterie länger als 100 Stunden?

Beispiel (Forts.)

A = Ausgewählte Batterie arbeitet länger als 100 Stunden

B_i = Batterie vom Typ i wird ausgewählt $\quad (i = 1, 2, 3)$

Satz v. d. vollst. Wahrscheinlichkeit:

$$\begin{aligned} P(A) &= \sum_{i=1}^{3} P(A|B_i)P(B_i) \\ &= (0.7)(0.2) + (0.4)(0.3) + (0.3)(0.5) \\ &= 0.41 \end{aligned}$$

Antwort: Die Wahrscheinlichkeit beträgt 41%.

Bayes'sche Formel

$C_1, C_2, \ldots \in \mathcal{A}$ sei eine Partition (d. h. eine disjunkte Zerlegung) des Merkmalraums Ω und $P(C_i) > 0$ für alle $i = 1, 2, \ldots$.

Dann gilt für ein Ereignis A mit $P(A) > 0$:

$$P(C_i|A) = \frac{P(A|C_i)P(C_i)}{\sum_{j=1}^{\infty} P(A|C_j)P(C_j)}$$

Beweis: Multiplikationsth. & Satz v. d. vollst. Wahrsch. $\Longrightarrow$

$$P(C_i|A) = \frac{P(C_i \cap A)}{P(A)} = \frac{P(A|C_i)P(C_i)}{P(A)} = \frac{P(A|C_i)P(C_i)}{\sum_{j=1}^{\infty} P(A|C_j)P(C_j)}$$

Sprechweise

$P(C_i)$ nennt man die A-priori–Wahrscheinlichkeit von C_i.

D. h., die W. von C_i vor dem Eintritt von A.

$P(C_i|A)$ nennt man die A-posteriori–Wahrscheinlichkeit von C_i.

D. h., die W. von C_i nach dem Eintritt von A.

Die Ereignisse C_i aus der Partition von Ω nennt man auch Hypothesen.

Beispiel

Die Kommissarin ist zu 60% (= A-priori–W.) davon überzeugt, dass der Verdächtige auch der Täter ist.

Neues Beweismittel: Der Täter hat zu 90% eine bestimmte Eigenart, die in der Bevölkerung zu 20% vorkommt.

Der Verdächtige hat diese Eigenart.

Wie ändert sich dadurch die Einschätzung der Kommissarin?

M. a. W., wie groß ist die A-posteriori–W. ?

Beispiel (Forts.)

T = Verdächtiger ist der Täter

E = Verdächtiger hat die Eigenart

Bayes'sche Formel:

$$
\begin{aligned}
P(T|E) &= \frac{P(E|T)P(T)}{P(E|T)P(T) + P(E|T^c)P(T^c)} \\
&= \frac{(0.9)(0.6)}{(0.9)(0.6) + (0.2)(0.4)} \\
&= \frac{54}{62} = \frac{27}{31} \doteq 0.871
\end{aligned}
$$

Antwort: Der Verdächtige ist a-posteriori zu 87% auch der Täter.

Odds–Form der Bayes'schen Formel

Betrachten wir in der obigen Sprechweise nur eine Hypothese H und ihre Gegenhypothese H^c, so lässt sich die Bayes'sche Formel auf Basis der Odds auch wie folgt schreiben:

$$\underbrace{\frac{P(H|A)}{P(H^c|A)}}_{\text{A-posteriori–Odds}} = \underbrace{\frac{P(H)}{P(H^c)}}_{\text{A-priori–Odds}} \times \underbrace{\frac{P(A|H)}{P(A|H^c)}}_{\text{Likelihood-Quotient}}$$

Beispiel (Forts.)

$$\underbrace{\frac{27}{4}}_{\text{A-posteriori–Odds}} = \underbrace{\frac{6}{4}}_{\text{A-priori–Odds}} \times \underbrace{\frac{9}{2}}_{\text{Likelihood-Quotient}}$$

D. h., die A-posteriori–Odds, dass der Verdächtige auch der Täter ist, stehen 27:4 .

$$\Longrightarrow \quad P(T|E) = \frac{27}{27+4} = \frac{27}{31} \doteq 0.871 \quad (87\%)$$

Diagnostischer Test

Ein Bluttest reagiert zu 95% positiv, wenn eine Krankheit vorliegt:

$$P(T+|D+) = 0.95 \quad (\text{Sensitivität})$$

Reagiert aber auch zu 1% „falsch positiv“ :

$$P(T+|D-) = 0.01, \qquad P(T-|D-) = 0.99 \quad (\text{Spezifität})$$

Man nimmt an, dass 0.5% der Bevölkerung erkrankt sind:

$$P(D+) = 0.005 \quad (\text{Prävalenz})$$

Diagnostischer Test (Forts.)

Mit welcher Wahrscheinlichkeit ist eine zufällig ausgewählte Person, deren Bluttest positiv ist, tatsächlich erkrankt?

Bayes'sche Formel:

$$\begin{aligned} P(D+|T+) &= \frac{P(T+|D+)P(D+)}{P(T+|D+)P(D+) + P(T+|D-)P(D-)} \\ &= \frac{(0.95)(0.005)}{(0.95)(0.005) + (0.01)(0.995)} \\ &= \frac{95}{294} = 0.323 \end{aligned}$$

Diagnostischer Test (Forts.)

Odds-Form der Bayes'schen Formel:

$$\underbrace{\frac{95}{199}}_{\text{A-posteriori–Odds}} = \underbrace{\frac{5}{995}}_{\text{A-priori–Odds}} \times \underbrace{\frac{95}{1}}_{\text{Likelihood-Quotient}}$$

Die A-posteriori–Odds sind – bei positivem Test – um den Faktor 95 höher als die A-priori–Odds.

Unabhängige Ereignisse

I. A. ist die bedingte Wahrscheinlichkeit von A gegeben B nicht gleich der unbedingten Wahrscheinlichkeit von A:

$$P(A|B) \neq P(A)$$

$$P(A|B) \geq P(A) \quad \text{oder} \quad P(A|B) \leq P(A)$$

$$P(A|B) = P(A) \quad \longrightarrow \quad A \text{ und } B \text{ sind „unabhängig"}$$

$$P(A|B) = P(A) \quad \Longrightarrow \quad P(A \cap B) = P(A)P(B)$$

Unabhängigkeit von zwei Ereignissen

Zwei Ereignisse $A, B \in \mathcal{A}$ nennt man (stochastisch) unabhängig (kurz: A, B ua.; man schreibt manchmal auch $A \perp B$), wenn:

$$P(AB) = P(A)P(B)$$

Andernfalls sagt man, dass die Ereignisse A und B abhängig sind.

$$A,\ B \text{ ua.} \quad \Longleftrightarrow \quad B,\ A \text{ ua.}$$

Beispiel

Augenzahl bei einem Würfel: $\Omega = \{1, 2, 3, 4, 5, 6\}$

$$A = \{1, 2\}, \quad B = \{2, 3, 4\}$$

$$P(AB) = P\big(\{2\}\big) = \frac{1}{6} = \underbrace{P(A)}_{\frac{1}{3}} \times \underbrace{P(B)}_{\frac{1}{2}}$$

$\Longrightarrow \quad A,\ B$ unabhängig

Beispiel (Forts.)

$$A = \{1,2,3\}, \quad B = \{2,3,4\}$$

$$P(AB) = P\big(\{2,3\}\big) = \frac{1}{3} \neq \underbrace{P(A)}_{\frac{1}{2}} \times \underbrace{P(B)}_{\frac{1}{2}}$$

$\Longrightarrow \quad A,\ B$ nicht unabhängig

Sind disjunkte Ereignisse unabhängig?

Augenzahl bei einem Würfel: $\Omega = \{1, 2, 3, 4, 5, 6\}$

$$A = \{1, 2\},\ B = \{3, 4\} \quad \Longrightarrow \quad A \cap B = \emptyset \quad \text{disjunkt}$$

$$P(AB) = P(\emptyset) = 0 \neq \underbrace{P(A)}_{\frac{1}{3}} \times \underbrace{P(B)}_{\frac{1}{3}}$$

$$\Longrightarrow \quad A,\ B \text{ nicht unabhängig}$$

Achtung: Disjunkte Ereignisse A und B mit $P(A) > 0$ und $P(B) > 0$ sind nicht unabhängig!

Sind unabhängige Ereignisse disjunkt?

Augenzahl bei einem Würfel: $\Omega = \{1,2,3,4,5,6\}$

$$A = \{1,2,3\},\ B = \{3,4\} \quad \Longrightarrow \quad A \cap B = \{3\} \quad \text{nicht disjunkt}$$

$$P(AB) = P\big(\{3\}\big) = \frac{1}{6} = \underbrace{P(A)}_{\frac{1}{2}} \times \underbrace{P(B)}_{\frac{1}{3}}$$

$$\Longrightarrow \quad A,\ B \text{ unabhängig}$$

Achtung: Unabhängige Ereignisse A und B mit $P(A) > 0$ und $P(B) > 0$ sind nicht disjunkt !

Behauptung 1

$$A, B \text{ ua.} \quad \Longrightarrow \quad (1)\ A, B^c \text{ ua.} \quad (2)\ A^c, B \text{ ua.} \quad (3)\ A^c, B^c \text{ ua.}$$

Beweis für (1):

$$A = \underbrace{AB \cup AB^c}_{\text{disjunkt}}$$

$$\Longrightarrow \quad P(A) = P(AB) + P(AB^c) = \underbrace{P(A)P(B)}_{A, B \text{ ua.}} + P(AB^c)$$

$$\Longrightarrow \quad P(AB^c) = P(A) - P(A)P(B) = P(A)\big[1 - P(B)\big] = P(A)P(B^c)$$

Unabhängigkeit von drei Ereignissen

Drei Ereignisse $A, B, C \in \mathcal{A}$ sind (stochastisch) unabhängig, wenn:

$$
\begin{aligned}
(1)\ P(ABC) &= P(A)P(B)P(C)\\
(2)\ P(AB) &= P(A)P(B)\\
(3)\ P(AC) &= P(A)P(C)\\
(4)\ P(CB) &= P(C)P(B)
\end{aligned}
$$

Gilt nur (2), (3), (4), nennt man die Ereignisse paarweise unabhängig.

Achtung: Aus der paarweisen Unabhängigkeit folgt i. A. nicht die (vollständige) Unabhängigkeit.

Achtung: Aus (1) folgt i. A. weder (2) noch (3) noch (4).

Behauptung 2

Sind A, B, C ua., dann ist A (oder A^c) auch unabhängig von jedem Ereignis, das sich aus B und C bilden lässt.

Beispiele:

A und $B \cup C$ sind ua.

A und $B \cap C^c$ sind ua.

A^c und $B^c \cup C$ sind ua.

etc.

Analoges gilt klarerweise auch für B und C.

Unabhängigkeit von n Ereignissen

Die n Ereignisse $A_1, A_2, \ldots, A_n \in \mathcal{A}$ sind (stochastisch) unabhängig, wenn für jede Teilmenge $\{i_1, i_2, \ldots, i_r\}$ von $\{1, 2, \ldots, n\}$ gilt:

$$P(A_{i_1} A_{i_2} \cdots A_{i_r}) = P(A_{i_1}) P(A_{i_2}) \cdots P(A_{i_r})$$

Unendlich viele Ereignisse sind (stochastisch) unabhängig, wenn jede endliche Teilmenge (stochastisch) unabhängig ist.

Behauptung 3

Sind $A_1, A_2, \ldots, A_n \in \mathcal{A}$ unabhängig, dann sind auch $B_1, B_2, \ldots, B_k$, $k \leq n$, unabhängig, wobei jedes B_i entweder A_i oder A_i^c ist.

Beispiel:

$$A, B, C, D, E \text{ ua.} \quad \Longrightarrow \quad A, B^c, C^c, D, E^c \text{ ua.}$$

Unabhängigkeit in der Praxis

- Zur Feststellung der Unabhängigkeit von n Ereignissen sind insgesamt $2^n - n - 1$ Bedingungen zu überprüfen.

  Bsp: Für 10 Ereignisse müssen 1013 Bedingungen überprüft werden!

- Vielfach lässt sich aus der Art eines (statistischen) Experiments auf Unabhängigkeit schließen.

  Bsp: Wird eine Münze wiederholt geworfen, so ist die Annahme nicht unplausibel, dass die (Ergebnisse der) einzelnen Würfe voneinander unabhängig sind.

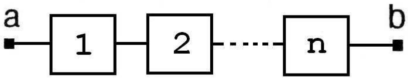
Unabhängigkeit in der Praxis (Forts.)

- In vielen Fällen ist Unabhängigkeit eine – oft nicht überprüfbare – Voraussetzung für weitere Berechnungen.

  Bsp: Bei komplexen Systemen aus vielen Einzelkomponenten lässt sich die Abhängigkeit zwischen den Komponenten meist nicht einfach beschreiben, sodass man in erster Näherung von deren Unabhängigkeit hinsichtlich des Ausfallverhaltens ausgeht.

Seriensystem

Wenn ein System, bestehend aus n Komponenten, nur dann funktioniert, wenn alle Komponenten funktionieren, spricht man von einem Seriensystem.



Seriensystem (Forts.)

Wenn jede Komponente unabhängig von den anderen Komponenten, mit der Wahrscheinlichkeit p_i, für $i = 1, 2, \ldots, n$, funktioniert, mit welcher Wahrscheinlichkeit funktioniert dann das System?

C_i = Ereignis, dass Komponente i funktioniert

C = Ereignis, dass das System funktioniert

$$C = \bigcap_{i=1}^{n} C_i \quad \Longrightarrow \quad P(C) = \prod_{i=1}^{n} P(C_i) = \prod_{i=1}^{n} p_i$$

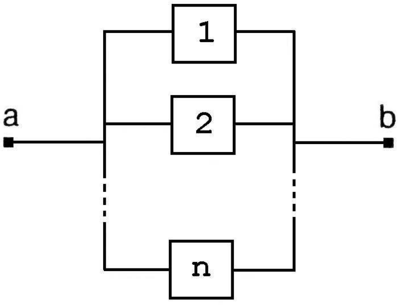
Spezialfall:

$$p_1 = p_2 = \cdots = p_n = p \quad \Longrightarrow \quad P(C) = p^n$$

Parallelsystem

Wenn ein System, bestehend aus n Komponenten, solange funktioniert, solange zumindest eine Komponente funktioniert, spricht man von einem Parallel-system.

Systeme dieser Art nennt man auch (vollständig) redundante Systeme.



Parallelsystem (Forts.)

Wenn jede Komponente unabhängig von den anderen Komponenten, mit der Wahrscheinlichkeit p_i, für $i = 1, 2, \ldots, n$, funktioniert, mit welcher Wahrscheinlichkeit funktioniert dann das System?

C_i = Ereignis, dass Komponente i funktioniert

C = Ereignis, dass das System funktioniert

$$C = \bigcup_{i=1}^{n} C_i \qquad C^c = \left(\bigcup_{i=1}^{n} C_i\right)^c = \bigcap_{i=1}^{n} C_i^c$$

Parallelsystem (Forts.)

$$P(C) = 1 - P(C^c) = 1 - P\left(\bigcap_{i=1}^{n} C_i^c\right)$$

$$= 1 - \prod_{i=1}^{n} P(C_i^c) = 1 - \prod_{i=1}^{n} \left[1 - P(C_i)\right]$$

$$= 1 - \prod_{i=1}^{n} (1 - p_i)$$

Spezialfall:

$$p_1 = p_2 = \cdots = p_n = p \quad \Longrightarrow \quad P(C) = 1 - (1 - p)^n$$

Mehrstufige Experimente

Viele Experimente laufen in mehreren Stufen ab, wobei abhängig von den Ergebnissen einer Stufe verschiedene Ergebnisse auf der folgenden Stufe möglich sind.

$\Omega_i =$ Merkmalraum der i–ten Stufe, $i = 1, 2, \ldots, n$

$\Omega =$ Merkmalraum des n–stufigen Experiments

$$\Longrightarrow \quad \text{Produktraum:} \quad \Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$$

Bei der Analyse mehrstufiger Experimente kommen auf natürliche Weise bedingte Wahrscheinlichkeiten ins Spiel.

2–stufiges Experiment: W–Baum

Zuerst wird ein regelmäßiger Tetraeder geworfen.

Ist die Augenzahl k (= 1, 2, 3, 4) werden anschließend k Münzen geworfen.

Der Versuchsausgang ist die Zahl der erzielten "Köpfe".