Kapitel 5

Stochastik

5.1 Ereignisse und ihre Wahrscheinlichkeiten

Folie 261

Definition 5.1 (Wahrscheinlichkeitsraum)**.**
Ein *endlicher Wahrscheinlichkeitsraum* (Ω, P) besteht aus einer endlichen, nicht-leeren Menge Ω von *Ergebnissen* bzw. *Elementarereignissen*, denen *Wahrscheinlichkeiten* $\mathrm{P}(\omega) = p_\omega \in \mathbb{R}$ für jedes $\omega \in \Omega$ zugeordnet sind, so dass gilt:

$$0 \ \leqslant \ p_\omega \ \leqslant \ 1, \quad \text{für jedes } \omega \in \Omega, \qquad \text{und} \qquad \sum_{\omega \in \Omega} p_\omega \ = \ 1.$$

Die Menge Ω fassen wir hierbei als die Menge aller möglichen Ergebnisse eines Zufallsexperiments auf. Für jedes $\omega \in \Omega$ gibt $\mathrm{P}(\omega)$ die Wahrscheinlichkeit an, dass eine einmalige Durchführung des Zufallsexperiments das Ergebnis ω liefert.

Folie 262

Beispiel 5.2 (2-maliger Münzwurf)**.**
Wir werfen zwei mal hintereinander eine „faire" Münze, d.h. eine Münze, die jeweils mit Wahrscheinlichkeit $\frac{1}{2}$ auf „Kopf" bzw. auf „Zahl" landet.[1] Als Wahrscheinlichkeitsraum betrachten wir dazu die Menge

$$\Omega \ = \ \{KK,\ KZ,\ ZK,\ ZZ\},$$

wobei z.B. KK für das Elementarergeignis steht, dass die Münze bei beiden Würfen auf „Kopf" landet, und ZK für das Elementarereignis steht, dass

[1] Wir gehen davon aus, dass das Landen auf „Kopf" bzw. „Zahl" die einzigen möglichen Ergebnisse sind — d.h. es kann nie passieren, dass die Münze „auf dem Rand stehenbleibt".

die Münze beim ersten Wurf auf „Zahl“ und beim zweiten Wurf auf „Kopf“ landet. Jedes der 4 möglichen Elementarereignisse hat hier dieselbe Wahrscheinlichkeit, d.h. für jedes $\omega \in \Omega$ ist hier $\mathrm{P}(\omega) = \frac{1}{4}$.

Folie 263

Definition 5.3 (Ereignisse)**.**
Ein *Ereignis* ist eine Menge von Ergebnissen, d.h. eine Teilmenge von Ω.
Die *Wahrscheinlichkeit* eines Ereignisses $A \subseteq \Omega$ ist definiert als

$$\mathrm{P}(A) \quad := \quad \sum_{\omega \in A} \mathrm{P}(\omega)\,.$$

Wir schreiben $\bar{A}$ um das *Komplement* von A zu bezeichnen, d.h. $\bar{A} := \Omega \setminus A$.

Anschaulich bedeutet die Aussage „Ereignis A tritt ein“, dass wir als Ergebnis eines Zufallsexperiments ein Elementarereignis $\omega \in A$ erhalten. Dies geschieht mit Wahrscheinlichkeit $\mathrm{P}(A)$.

Die Aussage „Ereignis A tritt nicht ein“ entspricht gerade der Situation, in der wir als Ergebnis eines Zufallsexperiments ein Elementarereignis $\omega \in \bar{A}$ erhalten.

Folie 264

Beispiel 5.4. Für den in Beispiel 5.2 betrachteten Wahrscheinlichkeitsraum Ω ist z.B.

$$A \quad := \quad \{KK,\ KZ,\ ZK\}$$

das Ereignis, bei dem bei mindestens einem der beiden Münzwürfe die Münze auf „Kopf“ landet. Die Wahrscheinlichkeit, dass dieses Ereignis eintritt, ist $\mathrm{P}(A) = \frac{3}{4}$.
Entsprechend ist $\bar{A} = \{ZZ\}$ das Ereignis, bei dem bei keinem der beiden Münzwürfe die Münze auf „Kopf“ landet. Dieses Ereignis tritt mit Wahrscheinlichkeit $\mathrm{P}(\bar{A}) = \frac{1}{4}$ ein.

Folie 265

Bemerkung 5.5 (Regeln zum Rechnen mit Wahrscheinlichkeiten)**.**
Man kann sich leicht davon überzeugen, dass für jeden endlichen Wahrscheinlichkeitsraum (Ω, P) und alle Ereignisse A und B gilt:

$$\mathrm{P}(\bar{A}) \quad = \quad 1 \;-\; \mathrm{P}(A)\,, \tag{5.1}$$

$$P(A \cup B) \quad = \quad P(A) \;+\; P(B) \;-\; P(A \cap B)\,. \tag{5.2}$$

Für $s \in \mathbb{N}_{\geqslant 1}$ und für beliebige Ereignisse $A_1, \ldots, A_s \subseteq \Omega$ gilt:

$$P(A_1 \cup \cdots \cup A_s) \quad \leqslant \quad \sum_{i=1}^{s} P(A_i)\,. \tag{5.3}$$

Hier ist $P(A_1 \cup \cdots \cup A_s)$ gerade die Wahrscheinlichkeit dafür, dass mindestens eins der Ereignisse $A_1, \ldots, A_s$ eintritt, d.h. die Wahrscheinlichkeit dafür, dass wir als Ergebnis des Zufallsexperiments ein Elementarereignis $\omega \in A_1 \cup \cdots \cup A_s$ erhalten.
Die Abschätzung (5.3) ist auch unter der Bezeichnung *Union Bound* bekannt.
Sind die Ereignisse $A_1, \ldots, A_s$ paarweise disjunkt, d.h. ist $A_i \cap A_j = \emptyset$ für alle $i, j \in \{1, \ldots, s\}$ mit $i \neq j$, so gilt sogar

$$P(A_1 \cup \cdots \cup A_s) \quad = \quad \sum_{i=1}^{s} P(A_i)\,. \tag{5.4}$$

Beweis. Da (Ω, P) ein endlicher Wahrscheinlichkeitsraum ist, gilt insbes.

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

Beweis von (5.1): Für ein beliebiges $A \subseteq \Omega$ gilt: $A \cap \bar{A} = \emptyset$ und $A \cup \bar{A} = \Omega$. Somit gilt:

$$1 \;=\; \sum_{\omega \in \Omega} P(\omega) \quad = \quad \sum_{\omega \in A} P(\omega) + \sum_{\omega \in \bar{A}} P(\omega) \quad = \quad P(A) + P(\bar{A}).$$

Somit ist $P(\bar{A}) = 1 - P(A)$.

Beweis von (5.2): Betrachte nun beliebige $A, B \subseteq \Omega$. Es gilt:
$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$, und die drei Mengen $(A \setminus B)$, $(B \setminus A)$ und $(A \cap B)$ sind paarweise disjunkt. Außerdem gilt: $A = (A \setminus B) \,\dot\cup\, (A \cap B)$ und $B = (B \setminus A) \,\dot\cup\, (A \cap B)$. Somit gilt:

$$\begin{aligned}
P(A) + P(B) \quad &= \quad \sum_{\omega \in A} P(\omega) \;+\; \sum_{\omega \in B} P(\omega) \\
&= \quad \sum_{\omega \in A \setminus B} P(\omega) \;+\; \sum_{\omega \in A \cap B} P(\omega) \;+\; \sum_{\omega \in B \setminus A} P(\omega) \;+\; \sum_{\omega \in A \cap B} P(\omega) \\
&= \quad \sum_{\omega \in A \cup B} P(\omega) \;+\; \sum_{\omega \in A \cap B} P(\omega) \\
&= \quad P(A \cup B) + P(A \cap B).
\end{aligned}$$

Somit gilt also: $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$.

Beweis von (5.3): Sei nun $s \in \mathbb{N}_{\geqslant 1}$ und seien $A_1, \ldots, A_s \subseteq \Omega$ und sei $V = A_1 \cup \cdots V_s$. Für jedes $\omega \in V$ gibt es mindestens ein $i \in [s]$ so dass $\omega \in A_i$ ist — aber es könnte auch mehrere verschiedene $i_1, i_2, \ldots$ geben, so dass $\omega \in A_{i_1}$ und $\omega \in A_{i_2}$ etc. ist. Auf jeden Fall gilt:

$$\sum_{i=1}^{s} \mathrm{P}(A_i) \;=\; \sum_{i=1}^{s} \sum_{\omega \in A_i} \mathrm{P}(\omega) \;\geqslant\; \sum_{\omega \in V} \mathrm{P}(\omega) \;=\; \mathrm{P}(V)\,.$$

Somit gilt also: $\mathrm{P}(A_1 \cup \cdots \cup A_s) \leqslant \sum_{i=1}^{s} \mathrm{P}(A_i)$.

Beweis von (5.4): Seien nun $A_1, \ldots, A_s \subseteq \Omega$ paarweise disjunkt und sei $V := A_1 \cup \cdots \cup A_s$. Dann gilt für jedes $\omega \in V$: Es gibt genau ein $i \in [s]$ so dass $\omega \in A_i$ ist. Daher gilt:

$$\mathrm{P}(V) \;=\; \sum_{\omega \in V} \mathrm{P}(\omega) \;=\; \sum_{i=1}^{s} \sum_{\omega \in A_i} \mathrm{P}(\omega) \;=\; \sum_{i=1}^{s} \mathrm{P}(A_i)\,.$$

Somit gilt für paarweise disjunkte Mengen $A_1, \ldots, A_s \subseteq \Omega$:

$$\mathrm{P}(A_1 \cup \cdots \cup A_s) \;=\; \mathrm{P}(A_1) + \cdots + \mathrm{P}(A_s)\,.$$

$\square$

Folie 266

Aus den in Bemerkung 5.5 bewiesenen Rechenregeln ergibt sich insbesondere, dass die Abbildung P eines endlichen Wahrscheinlichkeitsraums (Ω, P) eine *Wahrscheinlichkeitsverteilung* im Folgenden Sinne ergibt (wobei wir beachten, dass $\mathrm{P}(\{\omega\}) = \mathrm{P}(\omega)$ ist).

Definition 5.6. Sei Ω eine nicht-leere Menge. Eine *Wahrscheinlichkeitsverteilung* (auch: ein *Wahrscheinlichkeitsmaß*) für Ω ist eine Abbildung[2] $\mathrm{P} : \mathcal{P}(\Omega) \to [0,1]$, die die folgenden Eigenschaften besitzt:

1. $\mathrm{P}(\Omega) = 1$.
2. Für alle $A \subseteq \Omega$ gilt: $\mathrm{P}(A) \geqslant 0$.
3. Für alle $A, B \subseteq \Omega$ mit $A \cap B = \emptyset$ gilt: $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$.

Folie 267

[2] Wir schreiben $[0,1]$ um die Menge aller reellen Zahlen r mit $0 \leqslant r \leqslant 1$ zu bezeichnen.

Gleichverteilung und Laplacescher Wahrscheinlichkeitsraum

Definition 5.7. Sei Ω eine endliche nicht-leere Menge.
Die *Gleichverteilung* auf Ω ist der endliche Wahrscheinlichkeitsraum (Ω, P), bei dem P wie folgt definiert ist: Für jedes $\omega \in \Omega$ ist $\mathrm{P}(\omega) = \frac{1}{|\Omega|}$.
D.h.: Jedes Elementarereignis tritt mit derselben Wahrscheinlichkeit auf.

Dieser Wahrscheinlichkeitsraum (Ω, P) wird auch *Laplacescher Wahrscheinlichkeitsraum* genannt.

Ist (Ω, P) der Laplacesche Wahrscheinlichkeitsraum, so gilt für jedes Ereignis $A \subseteq \Omega$: $\mathrm{P}(A) = \frac{|A|}{|\Omega|}$.
Beispiel: Der in Beispiel 5.2 betrachtete Wahrscheinlichkeitsraum ist ein Laplacescher Wahrscheinlichkeitsraum.

Folie 268

Vorsicht beim "intuitiven" Argumentieren mit Wahrscheinlichkeiten

Bei der Analyse eines Zufallsexperiments sollte man sich nicht nur auf die eigene Intuition verlassen, sondern — um sicherzugehen, dass bei der Analyse kein Unsinn verzapft wird — auch die mathematische Definition der Wahrscheinlichkeiten benutzen. Oft ist die folgende *3-Schritt-Methode* hilfreich:

1. *Finde den richtigen Wahrscheinlichkeitsraum:*

   Bestimme alle möglichen Ergebnisse des Zufallsexperiments und deren Wahrscheinlichkeiten. D.h.: Bestimme die Menge Ω der Elementarereignisse und die Wahrscheinlichkeit $\mathrm{P}(\omega)$ für jedes $\omega \in \Omega$.

2. *Bestimme die relevanten Ereignisse:*

   Für welche konkreten Ereignisse $A \subseteq \Omega$ wollen wir wissen, mit welcher Wahrscheinlichkeit sie auftreten?

3. *Berechne die Wahrscheinlichkeiten der relevanten Ereignisse:*

   Für jedes Ereignis $A \subseteq \Omega$ gilt: $\mathrm{P}(A) = \sum_{\omega \in A} \mathrm{P}(\omega)$.

Beispiel: Diese Methode haben wir in Beispiel 5.4 angewendet um die Wahrscheinlichkeit dafür auszurechnen, dass bei zwei hintereinander ausgeführten Münzwürfen mindestens einmal die Münze auf "Kopf" landet.

Folie 269

Beispiel 5.8. Wir nutzen einen “fairen” Würfel, d.h. bei jedem Wurf des Würfels erhalten wir eine der Zahlen $1, 2, 3, 4, 5, 6$ als Ergebnis; und zwar jede dieser Zahlen mit der gleichen Wahrscheinlichkeit $\frac{1}{6}$.

Wir führen folgendes Zufallsexperiment aus: Wir würfeln zwei mal hintereinander.

Fragen:

(a) Wie groß ist die Wahrscheinlichkeit dafür, dass die Summe der in beiden Würfen erzielten Ergebnisse größer als 10 ist?

(b) Wie groß ist die Wahrscheinlichkeit dafür, dass die zweite gewürfelte Zahl größer als die erste ist?

Um diese beiden Fragen zu beantworten, nutzen wir die oben beschriebene 3-Schritt-Methode.

Folie 270

Schritt 1: Finde den richtigen Wahrscheinlichkeitsraum:

Jedes Tupel $(i, j) \in [6] \times [6]$ repräsentiert einen Ausgang des Zufallsexperiments — nämlich dass der erste Wurf die Zahl i und der zweite Wurf die Zahl j ergeben hat. Daher wählen wir die folgende Menge Ω von Elementarereignissen: $\Omega := [6] \times [6] = \{(i, j) \ : \ i \in [6], \ j \in [6]\}$.

Da jedes dieser Elementarereignisse mit derselben Wahrscheinlichkeit $\frac{1}{36}$ auftritt, wählen wir $\mathrm{P}(\omega) = \frac{1}{36}$ für jedes $\omega \in \Omega$. Somit haben wir einen Laplaceschen Wahrscheinlichkeitsraum, d.h. eine Gleichverteilung auf Ω, mit $|\Omega| = 36$.

Folie 271

Schritt 2: Bestimme die relevanten Ereignisse:

Bei Frage (a) wollen wir wissen: Wie groß ist die Wahrscheinlichkeit dafür, dass die Summe der in beiden Würfen erzielten Ergebnisse größer als 10 ist?

D.h. wir interessieren uns für das Ereignis A, das aus allen Elementarereignissen $(i, j) \in \Omega$ besteht, für die gilt: $i + j > 10$.
Sei also $A := \{(i, j) \in \Omega \ : \ i + j > 10\}$. Offensichtlicherweise gilt:
$A = \{ \, (5, 6), \ (6, 5), \ (6, 6) \, \}$.

Bei Frage (b) wollen wir wissen: Wie groß ist die Wahrscheinlichkeit dafür, dass die zweite gewürfelte Zahl größer als die erste ist?

D.h. wir interessieren uns für das Ereignis B, das aus allen Elementarereignissen $(i, j) \in \Omega$ besteht, für die gilt: $i < j$.

Sei also $B := \{(i,j) \in \Omega \ : \ i < j\}$.

Folie 272

Schritt 3: Berechne die Wahrscheinlichkeiten der Ereignisse A und B:
Da (Ω, P) ein Laplacescher Wahrscheinlichkeitsraum mit $|\Omega| = 36$ ist, gilt für jedes Ereignis $C \subseteq \Omega$: $\mathrm{P}(C) = \frac{|C|}{36}$.

Insbesondere gilt für $A = \{\,(5,6),\ (6,5),\ (6,6)\,\}$, dass $|A| = 3$ ist. Somit ist $\mathrm{P}(A) = \frac{3}{36} = \frac{1}{12}$. Die Antwort auf Frage (a) lautet also: "Die Wahrscheinlichkeit dafür, dass die Summe der in beiden Würfen erzielten Augenzahlen größer als 10 ist, ist genau $\frac{1}{12}$."

Folie 273

Für $B = \{(i,j) \in \Omega \ : \ i < j\}$ gilt: $\mathrm{P}(B) = \frac{|B|}{36}$. Aber wie können wir $|B|$ ausrechnen, ohne langwierig alle Tupel in B hinschreiben zu müssen? Durch geschicktes Abzählen!

Für jedes $i \in [6]$ sei X_i die Menge aller $j \in [6]$ mit $j > i$, d.h. $j \geqslant i{+}1$. Offensichtlicherweise ist $|X_i| = 6{-}i$ (von den 6 Elementen in $[6]$ gehören die ersten i nicht zu X_i, während die restlichen $6{-}i$ Elemente zu X_i dazugehören).
B ist dann die Vereinigung der Mengen $B_i := \{i\} \times X_i$ für jedes $i \in [6]$. Da diese Mengen paarweise disjunkt sind gilt:[3]

$$|B| \ = \ \sum_{i=1}^{6} |\{i\}{\times}X_i| \ = \ \sum_{i=1}^{6} |X_i| \ = \ \sum_{i=1}^{6} (6{-}i) \ = \ \sum_{k=0}^{5} k \ = \ \tfrac{5\cdot(5+1)}{2} \ = \ \tfrac{30}{2} \ = \ 15\,.$$

Somit ist $\mathrm{P}(B) = \frac{|B|}{36} = \frac{15}{36} = \frac{5}{12}$. Die Antwort auf Frage (b) lautet also: "Die Wahrscheinlichkeit dafür, dass die zweite gewürfelte Zahl größer als die erste ist, ist genau $\frac{5}{12}$."

Folie 274

Das Geburtstagsproblem

Beispiel 5.9. Beim *Geburtstagsproblem* (auch bekannt unter dem Begriff *Geburtstagsparadoxon*) geht es um die folgende Situation.

In einem Raum befinden sich k Personen.

Frage: Wie groß ist die Wahrscheinlichkeit dafür, dass es darunter zwei Personen gibt, die am gleichen Tag Geburtstag[4] haben?

[3] Wir wenden hier die Gaußsche Summenformel an: $\sum_{k=0}^{n} = \frac{n(n+1)}{2}$, siehe Satz 2.44.
[4] ohne Berücksichtigung des Geburtsjahrs

Verblüffende Antwort: Schon bei nur $k = 25$ Personen ist die Wahrscheinlichkeit, dass zwei davon am gleichen Tag Geburtstag haben, größer als $\frac{1}{2}$. Und bei $k = 50$ Personen ist sie größer als $0,96$.

Dies lässt sich wie folgt errechnen. Wir wenden die 3-Schritt-Methode an.

Erfahrungsbericht: Am 23.01.2023 wurde dieses Zufallsexperiment unter den Teilnehmer:innen der Veranstaltung „Diskrete Strukturen“ durchgeführt. Am Experiment nahmen insgesamt 58 Personen teil.
Ergebnis: Es gab einen Tag, an dem 3 Personen Geburtstag haben und zwei weitere Tage, an denen je 2 Personen Geburtstag haben.

Folie 275

Schritt 1: Finde den richtigen Wahrscheinlichkeitsraum

Wir machen die folgenden vereinfachenden Annahmen (die nicht unbedingt exakt der Realität entsprechen):

- Niemand ist in einem Schaltjahr geboren; somit gibt es genau 365 Möglicheiten dafür, an welchem Tag jemand Geburtstag hat.
- Geburtstage sind über die 365 Tage des Jahres gleichverteilt.

Wir nummerieren die k Personen durch mit den Zahlen $1, \dots, k$, und wir nummerieren die 365 denkbaren Geburtstage durch mit den Zahlen $1, \dots, 365$. Als Wahrscheinlichkeitsraum wählen wir

$$\Omega \ := \ \{(a_1, \dots, a_k) \ : \ a_1, \dots, a_k \in \{1, \dots, 365\}\} \ = \ \{1, \dots, 365\}^k$$

Hierbei steht “a_i” dafür, dass Person i am Tag a_i Geburtstag hat.

Entsprechend der oben beschriebenen Annahmen gehen wir davon aus, dass wir eine Gleichverteilung auf Ω haben,d.h. jedes Elementarereignis $\omega \in \Omega$ hat die Wahrscheinlichkeit

$$\mathrm{P}(\omega) \ = \ \frac{1}{|\Omega|} \ = \ \frac{1}{365^k}\,.$$

Folie 276

Schritt 2: Finde die relevanten Ereignisse

Wir interessieren uns für das Ereignis, bei dem mindestens zwei der k Personen am gleichen Tag Geburtstag haben, d.h., für das Ereignis

$$A \ := \ \{(a_1, \dots, a_k) \in \Omega \ : \ \text{es gibt } i, j \in [k] \text{ mit } i \neq j \text{ und } a_i = a_j\}$$

Schritt 3: Berechne die Wahrscheinlichkeit des Ereignisses A

Wir wissen: $\mathrm{P}(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{365^k}$

Aber wie können wir $|A|$ ausrechnen?
Idee: Wir berechnen stattdessen $\mathrm{P}(\bar{A})$ und nutzen die Gleichung $\mathrm{P}(A) = 1 - \mathrm{P}(\bar{A})$.

Es gilt:

$$\begin{array}{rcll} \bar{A} & = & \Omega \setminus A & \\ & = & \{(a_1, \dots, a_k) \in \Omega & : \ a_1, \dots, a_k \text{ sind paarweise verschieden}\} \\ & = & \{(a_1, \dots, a_k) \in [n]^k & : \ a_1, \dots, a_k \text{ sind paarweise verschieden}\} \end{array}$$

für $n := 365$. Aus Kapitel 4, Gleichung (4.2) wissen wir, dass Folgendes gilt:

$$|\bar{A}| \;=\; (n)_k \;=\; n \cdot (n{-}1) \cdot (n{-}2) \cdots (n{-}k{+}1)$$

Also gilt:

$$\mathrm{P}(\bar{A}) \;=\; \frac{|\bar{A}|}{|\Omega|} \;=\; \frac{(n)_k}{n^k} \;=\; \prod_{i=0}^{k-1} \frac{n-i}{n} \;=\; \prod_{i=1}^{k-1} \Big(1 - \frac{i}{n}\Big)$$

Um dies zu vereinfachen, nutzen wir die Abschätzung (4.8), die besagt:

$$\text{Für alle } x \in \mathbb{R} \text{ ist } \ 1 + x \leqslant e^x\,.$$

Für $x := -\frac{i}{n}$ liefert dies: $(1 - \frac{i}{n}) \leqslant e^{-\frac{i}{n}}$. Somit ist

$$\mathrm{P}(\bar{A}) \;\leqslant\; \prod_{i=1}^{k-1} e^{-\frac{i}{n}} \;=\; e^{\sum_{i=1}^{k-1} -\frac{i}{n}}\,.$$

Unter Verwendung der Gaußschen Summenformel erhalten wir:

$$\sum_{i=1}^{k-1} -\frac{i}{n} \;=\; -\frac{1}{n} \cdot \sum_{i=1}^{k-1} i \;=\; -\frac{1}{n} \cdot \frac{(k-1) \cdot k}{2} \;=\; -\frac{(k-1) \cdot k}{2n}\,.$$

Also ist

$$\mathrm{P}(\bar{A}) \;\leqslant\; e^{-\frac{(k-1)\cdot k}{2n}} \qquad \text{und} \qquad \mathrm{P}(A) \;=\; 1 - \mathrm{P}(\bar{A}) \;\geqslant\; 1 - e^{-\frac{(k-1)\cdot k}{2n}}\,. \tag{5.5}$$

Durch Einsetzen von $n = 365$ erhalten wir:

$$\mathrm{P}(A) \;\geqslant\; 1 - e^{-\frac{(k-1)\cdot k}{730}}\,.$$

Ausrechnen mit dem Taschenrechner liefert für $k = 25$ Personen:

$$\mathrm{P}(A) \quad \geqslant \quad 1 - e^{\frac{-600}{730}} \quad \geqslant \quad 1 - 0,43958 \quad \geqslant \quad 0,56\ .$$

Für $n = 30$ erhalten wir:

$$\mathrm{P}(A) \quad \geqslant \quad 1 - e^{\frac{-870}{730}} \quad \geqslant \quad 1 - 0,30368 \quad \geqslant \quad 0,7\ .$$

Und für $n = 50$ erhalten wir:

$$\mathrm{P}(A) \quad \geqslant \quad 1 - e^{\frac{-2450}{730}} \quad \geqslant \quad 1 - 0,03486 \quad \geqslant \quad 0,96\ .$$

Folie 277

Bezug zur Informatik: Kollisionen beim Hashing

Um schnellen Zugriff auf einzelne Daten zu haben, verwendet man beim Programmieren gerne *Hashtabellen* bzw. *Hashfunktionen*.
Dabei werden einzelne Datensätze, die aus einer großen Menge D stammen können, auf "Schlüssel" aus einer (i.d.R. kleineren) Menge $S = \{1, \ldots, n\}$ (mit $n \in \mathbb{N}_{\geqslant 1}$) abgebildet.
Eine Hashfunktion ist eine Funktion $h : D \to S$, so dass man bei Eingabe eines Datenelements $d \in D$ den "Hashwert" $h(d)$ schnell berechnen kann.
Informationen über die verfügbaren Datenwerte werden in einem Array $A[1 \ldots n]$ abgespeichert — und zwar für den Datenwert d im Array-Eintrag $A(h(d))$.
Wenn man auf die für d gespeicherten Informationen zugreifen will, berechnet man einfach schnell den Hashwert $i := h(d)$ und schaut dann im Array-Eintrag $A(i)$ nach. Problematisch dabei sind *Kollisionen* — also wenn mehrere verschiedene Datenwerte auf denselben Hashwert i abgebildet werden. Dann speichert man im Array-Eintrag i.d.R. eine Liste aller entsprechenden Datenwerte ab, die man dann wieder mühsam durchsuchen muss.
Das "Geburtstagsproblem" zeigt, dass Kollisionen mit recht großer Wahrscheinlichkeit auftreten können. Für die Situation, in der wir n Schlüssel zur Verfügung haben und insgesamt k Datenwerte gespeichert werden, gibt uns Gleichung (5.5) die Wahrscheinlichkeit dafür an, dass es Kollisionen gibt.
Beachte: Schon bei $k \approx \sqrt{n}$ ist diese Wahrscheinlichkeit recht groß.

5.2 Zufallsvariablen

Folie 278

Definition 5.10 (Zufallsvariable)**.**
Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum.
Eine *Zufallsvariable* ist eine Funktion

$$X : \Omega \to \mathbb{R}\,.$$

Für jede Zahl $r \in \mathbb{R}$ definieren wir die Wahrscheinlichkeit dafür, dass die Zufallsvariable X den Wert r annimmt, durch

$$\mathrm{P}(X = r) \ := \ \mathrm{P}(A) \quad \text{für das Ereignis} \quad A := \{\omega \in \Omega \ : \ X(\omega) = r\}.$$

Folie 279

Beispiel 5.11 (Münzwürfe und Zufallsvariablen)**.**
Für eine beliebige, feste Zahl $p \in [0,1]$ haben wir eine Münze zur Verfügung, die mit Wahrscheinlichkeit p auf „Kopf“ und mit Wahrscheinlichkeit $1-p$ auf „Zahl“ fällt. Für eine feste Zahl $s \in \mathbb{N}_{\geqslant 1}$ werfen wir s-mal hintereinander die Münze.

Formal gesehen, betrachten also den Wahrscheinlichkeitsraum (Ω, P), dessen Elementarereignisse Worte der Länge s über dem Alphabet $\{K, Z\}$ sind, d.h. $\Omega = \{K, Z\}^s$.
Jedes Elementarereignis $\omega = \omega_1 \cdots \omega_s \in \Omega$ tritt mit Wahrscheinlichkeit $\mathrm{P}(\omega) := p^k(1-p)^{s-k}$ ein, wobei k die Anzahl der Vorkommen des Buchstabens K im Wort ω ist.

Folie 280

Für jeden Münzwurf $i \in \{1, \ldots, s\}$ betrachten wir die Zufallsvariable X_i, der wir den Wert 1 zuordnen, falls die Münze im i-ten Wurf auf „Kopf“ gelandet ist, und der wir den Wert 0 zuordnen, falls die Münze im i-ten Wurf auf „Zahl“ gelandet ist. Wir schreiben dafür kurz

$$X_i \ := \ \begin{cases} 1 & \text{falls die Münze im } i\text{-ten Wurf auf „Kopf“ landet} \\ 0 & \text{sonst} \end{cases}$$

und meinen damit die Zufallsvariable $X_i : \Omega \to \mathbb{R}$, bei der für jedes Elementarereignis $\omega \in \Omega$ gilt:

$$X_i(\omega) \ = \ \begin{cases} 1 & \text{falls an der } i\text{-ten Position des Worts } \omega \text{ der Buchstabe } K \text{ steht} \\ 0 & \text{sonst.} \end{cases}$$

Offensichtlicherweise gilt für jedes $i \in \{1, \ldots, s\}$, dass

$$P(X_i = 1) \;=\; p \qquad \text{und} \qquad P(X_i = 0) \;=\; 1 - p\,.$$

Folie 281

Zusätzlich zu den Zufallsvariablen $X_1, \ldots, X_s$ betrachten wir noch eine weitere Zufallsvariable X, die angeben soll, bei wie vielen der s Münzwürfe die Münze auf „Kopf“ gelandet ist. Wir schreiben dazu kurz

$$X \;\;:=\;\; \sum_{i=1}^{s} X_i$$

und meinen damit die Zufallsvariable $X : \Omega \to \mathbb{R}$, bei der für jedes Elementarereignis $\omega \in \Omega$ gilt:[5]

$$X(\omega) \quad = \quad \sum_{i=1}^{s} X_i(\omega).$$

Man kann sich leicht davon überzeugen, dass für jedes $k \in \{0, \ldots, s\}$ gilt:

$$P(X = k) \quad = \quad \binom{s}{k}\; p^k \;(1-p)^{s-k}.$$

Dieser Wert gibt die Wahrscheinlichkeit dafür an, dass bei den s Münzwürfen die Münze insgesamt genau k-mal auf „Kopf“ landet.

5.3 Erwartungswert

Folie 282

Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum und sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable. Da Ω endlich ist, ist auch das *Bild* von Ω unter X, d.h. die Menge $\mathit{Bild}(X) := \{X(\omega) \;:\; \omega \in \Omega\}$ endlich.

Definition 5.12 (Erwartungswert einer Zufallsvariablen)**.**
Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum und sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable. Der *Erwartungswert* der Zufallsvariablen X ist definiert als

$$E(X) \;\;:=\;\; \sum_{r \in \mathit{Bild}(X)} r \cdot P(X = r).$$

[5] *Zur Erinnerung:* Der Binomialkoeffizient $\binom{n}{k}$ gibt an, wie viele verschiedene k-elementige Teilmengen eine n-elementige Menge besitzt. Es gilt: $\binom{n}{k} = \frac{n!}{k!\cdot(n-k)!}$.

Anschaulich ist der Erwartungswert der „zu erwartende“ Wert, den wir erhalten, wenn wir das Zufallsexperiment sehr oft wiederholen und den Durchschnitt der dabei für X erhaltenen Werte berechnen.

Man kann sich leicht davon überzeugen, dass Folgendes gilt:

$$\mathrm{E}(X) \quad = \quad \sum_{\omega\in\Omega} X(\omega)\cdot \mathrm{P}(\omega)\,. \tag{5.6}$$

Beweis. Es gilt:

$$\begin{aligned}
\mathrm{E}(X) \quad &= \quad \sum_{r\in \mathit{Bild}(X)} r\cdot \mathrm{P}(X=r)\\
&= \quad \sum_{r\in \mathit{Bild}(X)} r\cdot \mathrm{P}(\{\omega\in\Omega \ :\ X(\omega)=r\})\\
&= \quad \sum_{r\in \mathit{Bild}(X)} r\cdot \Big(\sum_{\substack{\omega\in\Omega:\\ X(\omega)=r}} \mathrm{P}(\omega)\Big)\\
&= \quad \sum_{r\in \mathit{Bild}(X)} \ \sum_{\substack{\omega\in\Omega:\\ X(\omega)=r}} X(\omega)\cdot \mathrm{P}(\omega)\\
&= \quad \sum_{\omega\in\Omega} X(\omega)\cdot \mathrm{P}(\omega)\,.
\end{aligned}$$

□

Folie 283

Beispiel 5.13 (Erwartungswert beim Würfeln)**.**
Wir werfen einen herkömmlichen Würfel, d.h. einen Würfel, der 6 Seiten besitzt, die mit den Augenzahlen 1 bis 6 durchnummeriert sind, und bei dem jede Augenzahl mit der selben Wahrscheinlichkeit gewürfelt wird. Sei X die Zufallsvariable, die die gewürfelte Augenzahl beschreibt. Der Erwartungswert von X ist

$$\mathrm{E}(X) \quad = \quad \sum_{i=1}^{6} i\cdot \mathrm{P}(X=i) \quad = \quad \sum_{i=1}^{6} i\cdot\tfrac{1}{6} \quad = \quad \tfrac{1}{6}\cdot\tfrac{6\cdot 7}{2} \quad = \quad \tfrac{7}{2} \quad = \quad 3,5\,.$$

Folie 284

Man beachte, dass die Zufallsvariable X^2 das Quadrat der gewürfelten Augenzahl beschreibt. Der Erwartungswert von X^2 ist

$$\mathrm{E}(X^2) \;=\; \sum_{i=1}^{6} i^2 \cdot \mathrm{P}(X^2 = i^2) \;=\; \sum_{i=1}^{6} i^2 \cdot \mathrm{P}(X = i) \;=\; \tfrac{1}{6} \cdot \sum_{i=1}^{6} i^2 \;=\; \tfrac{91}{6} \;=\; 15,16666\cdots.$$

Insbesondere ist hier

$$\mathrm{E}(X^2) \;\neq\; \mathrm{E}(X)^2,$$

da $\mathrm{E}(X)^2 = (3,5)^2 = 12,25$ ist.

Folie 285

Linearität des Erwartungswerts

Bemerkung 5.14 (*Linearität des Erwartungswerts*)**.**
Man kann sich leicht davon überzeugen, dass für jeden endlichen Wahrscheinlichkeitsraum (Ω, P), für alle Zufallsvariablen X und Y und für alle Zahlen $a \in \mathbb{R}$ gilt:

$$\mathrm{E}(aX) \;=\; a \cdot \mathrm{E}(X) \qquad \text{und} \qquad \mathrm{E}(X+Y) \;=\; \mathrm{E}(X) + \mathrm{E}(Y).$$

Daraus folgt, dass für alle $s \in \mathbb{N}_{\geqslant 1}$, alle Zufallsvariablen $X_1, \ldots, X_s$ und alle Zahlen $a_1, \ldots, a_s \in \mathbb{R}$ gilt:

$$\mathrm{E}\Big(\sum_{i=1}^{s} a_i X_i\Big) \;=\; \sum_{i=1}^{s} a_i \cdot \mathrm{E}(X_i).$$

Beweis. Wir nutzen (5.6). Daraus folgt:

$$\begin{aligned} a \cdot \mathrm{E}(X) &= a \cdot \sum_{\omega \in \Omega} X(\omega) \cdot \mathrm{P}(\omega) \\ &= \sum_{\omega \in \Omega} a \cdot X(\omega) \cdot \mathrm{P}(\omega) \\ &= \sum_{\omega \in \Omega} (aX)(\omega) \cdot \mathrm{P}(\omega) \\ &= \mathrm{E}(aX). \end{aligned}$$

Des Weiteren gilt:

$$\begin{aligned}
\mathrm{E}(X)+\mathrm{E}(Y) &= \Big(\sum_{\omega\in\Omega} X(\omega)\cdot \mathrm{P}(\omega)\Big) + \Big(\sum_{\omega\in\Omega} Y(\omega)\cdot \mathrm{P}(\omega)\Big) \\
&= \sum_{\omega\in\Omega} \big(X(\omega)+Y(\omega)\big)\cdot \mathrm{P}(\omega) \\
&= \sum_{\omega\in\Omega} (X+Y)(\omega)\cdot \mathrm{P}(\omega) \\
&= \mathrm{E}(X+Y)\,.
\end{aligned}$$

Es folgt:

$$\mathrm{E}(\sum_{i=1}^{s} a_i X_i) \;=\; \sum_{i=1}^{s} \mathrm{E}(a_i X_i) \;=\; \sum_{i=1}^{s} a_i \mathrm{E}(X_i)\,.$$

□

Folie 286

Beispiel 5.15 (Erwartungswert bei Münzwürfen)**.**
Wir betrachten wieder das Szenario aus Beispiel 5.11, bei dem für jedes $i \in \{1,\ldots,s\}$ die Zufallsvariable X_i angibt, ob die Münze beim i-ten Münzwurf auf „Kopf“ gelandet ist, und bei dem die Zufallsvariable $X := \sum_{i=1}^{s} X_i$ angibt, bei wie vielen der s Münzwürfe die Münze auf „Kopf“ gelandet ist. Für jedes $i \in \{1,\ldots,s\}$ hat X_i den Erwartungswert

$$\mathrm{E}(X_i) \;=\; 0\cdot \mathrm{P}(X_i=0) \;+\; 1\cdot \mathrm{P}(X_i=1) \;=\; \mathrm{P}(X_i=1) \;=\; p.$$

Gemäß der Linearität des Erwartungswerts ist der Erwartungswert der Zufallsvariablen $X := \sum_{i=1}^{s} X_i$ daher der Wert

$$\mathrm{E}(X) \;=\; \sum_{i=1}^{s} \mathrm{E}(X_i) \;=\; p\cdot s.$$

Folie 287

Bemerkung 5.16. Die in Bemerkung 5.14 formulierte Linearität des Erwartungswerts besagt, dass Summen und konstante Vielfache aus der Bildung des Erwartungswerts herausgezogen werden können, so dass

$$\mathrm{E}\Big(\sum_{i=1}^{s} a_i X_i\Big) \;=\; \sum_{i=1}^{s} a_i\cdot \mathrm{E}(X_i).$$

Im Gegensatz dazu dürfen wir Produkte i.d.R. nicht einfach aus der Bildung des Erwartungswerts herausziehen. Beispielsweise haben wir in Beispiel 5.13 eine Situation kennen gelernt, in der gilt:

$$\mathrm{E}(X\cdot X) \;\neq\; \mathrm{E}(X)\cdot \mathrm{E}(X)\,.$$

5.4 Varianz

Folie 288

Die *Varianz* ist ein Maß dafür, wie weit die tatsächlichen Werte einer Zufallsvariablen X vom Erwartungswert $\mathrm{E}(X)$ abweichen.

Definition 5.17 (Varianz einer Zufallsvariablen)**.**
Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum und sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable. Die *Varianz* der Zufallsvariablen X ist definiert als

$$\mathrm{Var}(X) \quad := \quad \mathrm{E}\Big(\big(X - \mathrm{E}(X)\big)^2\Big).$$

Somit ist die Varianz einer Zufallsvariablen X definiert als der Erwartungswert des Quadrats der Abweichung von X zum Erwartungswert von X. Anschaulich ist die Varianz von X also der „zu erwartende“ Wert, den wir erhalten, wenn wir das Zufallsexperiment sehr oft wiederholen, jeweils die Abweichung des für X erzielten Werts vom Erwartungswert $\mathrm{E}(X)$ berechnen, und den Durchschnitt der Quadrate dieser Abweichungen berechnen.

Folie 289

Regeln zum Rechnen mit Varianzen

Bemerkung 5.18 (Regeln zum Rechnen mit Varianzen)**.**
Für jede Zufallsvariable X über einem endlichen Wahrscheinlichkeitsraum (Ω, P) und für alle Zahlen $a, b \in \mathbb{R}$ gilt:

$$\mathrm{Var}(X) \;=\; \mathrm{E}(X^2) - \mathrm{E}(X)^2 \qquad \text{und} \qquad \mathrm{Var}(aX+b) \;=\; a^2 \cdot \mathrm{Var}(X).$$

Beweis. Es gilt:

$$\begin{aligned}
\mathrm{Var}(X) &\overset{\text{Def.}}{=} \mathrm{E}\Big((X - \mathrm{E}(X))^2\Big) \\
&= \mathrm{E}\Big(X^2 - 2{\cdot}\mathrm{E}(X){\cdot}X + \mathrm{E}(X)^2\Big) \\
&\overset{\text{Linearität}}{=} \mathrm{E}(X^2) - 2\mathrm{E}(X)\cdot\mathrm{E}(X) + \mathrm{E}(X)^2 \cdot \mathrm{E}(1) \\
&\overset{\mathrm{E}(1)=1}{=} \mathrm{E}(X^2) - \mathrm{E}(X)^2.
\end{aligned}$$

Unter Verwendung dieser Gleichung erhalten wir:

$$\begin{aligned}
\mathrm{Var}(aX) &= \mathrm{E}((aX)^2) - \mathrm{E}(aX)^2 \\
&\overset{\text{Linearität}}{=} \mathrm{E}(a^2 \cdot X^2) - \big(a \cdot \mathrm{E}(X)\big)^2 \\
&\overset{\text{Linearität}}{=} a^2 \cdot \mathrm{E}(X^2) - a^2 \cdot \mathrm{E}(X)^2 \\
&= a^2 \cdot \Big(\mathrm{E}(X^2) - \mathrm{E}(X)^2\Big) \\
&= a^2 \cdot \mathrm{Var}(X).
\end{aligned}$$

Außerdem gilt für jede Zufallsvariable Y:

$$\begin{aligned}
\mathrm{Var}(Y+b) &= \mathrm{E}\Big(\big(Y + b - \underbrace{\mathrm{E}(Y+b)}_{\overset{\text{Linearität}}{=}\ \mathrm{E}(Y)+b}\ \big)^2\Big) \\
&= \mathrm{E}\Big(\big(Y - \mathrm{E}(Y)\big)^2\Big) \\
&= \mathrm{Var}(Y).
\end{aligned}$$

Insgesamt erhalten wir daher für die Zufallsvariable $Y := aX$, dass

$$\mathrm{Var}(aX+b) \quad = \quad \mathrm{Var}(aX) \quad = \quad a^2 \cdot \mathrm{Var}(X).$$

□

Folie 290

Beispiel 5.19 (Varianz beim Würfeln)**.**
Wir betrachten das Szenario aus Beispiel 5.13, bei dem die Zufallsvariable X die beim 1-maligen Werfen eines 6-seitigen Würfels erzielte Augenzahl angibt. Wir wissen bereits, dass der Erwartungswert $\mathrm{E}(X) = 3,5 = \frac{7}{2}$ ist, und dass der Erwartungswert $\mathrm{E}(X^2) = \frac{91}{6}$ ist.
Die Varianz von X ist der Wert

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2) - \mathrm{E}(X)^2 &&= \tfrac{91}{6} - \left(\tfrac{7}{2}\right)^2 &&= \tfrac{91}{6} - \tfrac{49}{4} \\
&= \tfrac{182-147}{12} &&= \tfrac{35}{12} &&= 2,916666\cdots.
\end{aligned}$$

Folie 291

Bemerkung 5.20 (Nicht-Linearität der Varianz)**.**
Anhand der Regel

$$\mathrm{Var}(aX+b) \quad = \quad a^2 \cdot \mathrm{Var}(X)$$

lässt sich leicht sehen, dass *die Varianz i.d.R. nicht linear ist*:
Beispielsweise erhalten wir für $a = 2$ und $b = 0$, dass

$$\mathrm{Var}(X + X) \quad = \quad 4 \cdot \mathrm{Var}(X) \quad \neq \quad \mathrm{Var}(X) + \mathrm{Var}(X),$$

falls $\mathrm{Var}(X) \neq 0$.

5.5 Schranken

Folie 292

In diesem Abschnitt stellen wir zwei Werkzeuge bereit, mit deren Hilfe wir abschätzen können, mit welcher Wahrscheinlichkeit, der Wert einer Zufallsvariablen X „weit" vom Erwartungswert $\mathrm{E}(X)$ abweicht.

Folie 293

Die Markov-Ungleichung

Satz 5.21 (Markov-Ungleichung)**.**
Sei (Ω, P) *ein endlicher Wahrscheinlichkeitsraum und sei* $X : \Omega \to \mathbb{R}_{\geqslant 0}$ *eine Zufallsvariable, die nur Werte* $\geqslant 0$ *annehmen kann. Für jede reelle Zahl* $a > 0$ *gilt:*

$$\mathrm{P}(X \geqslant a) \quad \leqslant \quad \frac{\mathrm{E}(X)}{a}.$$

Beweis. Es gilt:

$$\begin{aligned} \mathrm{E}(X) \quad &= \sum_{x \in \mathit{Bild}(X)} x \cdot \mathrm{P}(X = x) \\ &\geqslant \sum_{\substack{x \in \mathit{Bild}(X) \\ \mathrm{mit} x \geqslant a}} a \cdot \mathrm{P}(X = x) \quad = \quad a \cdot \mathrm{P}(X \geqslant a). \end{aligned}$$

□

Bemerkung 5.22. Aus der Markov-Ungleichung folgt direkt, dass

$$\mathrm{P}(X \geqslant c{\cdot}\mathrm{E}(X)) \quad \leqslant \quad \tfrac{1}{c}$$

für jede Zahl $c > 0$ gilt und für jede Zufallsvariable $X : \Omega \to \mathbb{R}_{\geqslant 0}$, die nur Werte $\geqslant 0$ annehmen kann und deren Erwartungswert > 0 ist.

Folie 294

Ein Beispiel zur Anwendung der Markov-Ungleichung

Bemerkung 5.23. *Randomisiertes Quicksort* ist die Variante des Sortieralgorithmus *Quicksort*, bei der in jedem einzelnen Schritt das Pivotelement *zufällig* und gleichverteilt aus den prinzipiell denkbaren Pivotelementen gezogen wird.

Eine detaillierte Laufzeitanalyse zeigt, dass die *erwartete Laufzeit* von randomisiertem Quicksort bei Eingabe einer zu sortierenden Liste der Länge n von der Form $c \cdot n \cdot \log_2(n)$ ist, für eine geeignete Konstante c.

Aus der Markov-Ungleichung folgt, dass die Wahrscheinlichkeit, das Pech zu haben, bei einem konkreten Lauf von randomisiertem Quicksort eine Laufzeit von $\geqslant 10 \cdot c \cdot n \cdot \log_2(n)$ (also dem 10-fachen der erwarteten Laufzeit) zu erwischen, höchstens $\frac{1}{10}$ ist.

Folie 295

Die Tschebyscheff-Ungleichung

Satz 5.24 (Tschebyscheff-Ungleichung)**.**
Für jeden endlichen Wahrscheinlichkeitsraum (Ω, P)*, für jede Zufallsvariable* $X : \Omega \to \mathbb{R}$ *und für jede reelle Zahl* $a > 0$ *gilt:*

$$\mathrm{P}\Big(|X - \mathrm{E}(X)| \;\geqslant\; a\Big) \quad \leqslant \quad \frac{\mathrm{Var}(X)}{a^2}.$$

Beweis. Wir betrachten die Zufallsvariable $Y := |X - \mathrm{E}(X)|$ und wenden die Markov-Ungleichung auf die Zufallsvariable Y^2 an. Wir erhalten:

$$\mathrm{P}(|X - \mathrm{E}(X)| \geqslant a) \quad = \quad \mathrm{P}(Y \geqslant a) \quad = \quad \mathrm{P}(Y^2 \geqslant a^2) \quad \overset{\text{Markov}}{\leqslant} \quad \frac{\mathrm{E}(Y^2)}{a^2}.$$

Hierbei gilt:

$$\mathrm{E}(Y^2) \quad = \quad \mathrm{E}((X - \mathrm{E}(X))^2) \quad \overset{\text{Def.}}{=} \quad \mathrm{Var}(X).$$

□

Ein Beipiel zur Anwendung der Tschebyscheff-Ungleichung behandeln wir in Kapitel 5.7.

5.6 Paarweise Unabhängigkeit

Folie 296

Definition 5.25. Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum.

(a) Zwei Ereignisse A und B heißen (stochastisch) *unabhängig*, wenn gilt:

$$\mathrm{P}(A \cap B) \quad = \quad \mathrm{P}(A) \cdot \mathrm{P}(B).$$

Für jede Zahl $s \in \mathbb{N}_{\geqslant 1}$ heißen die Ereignisse $A_1, \ldots, A_s$ *paarweise unabhängig*, wenn für alle $i, j \in \{1, \ldots, s\}$ mit $i \neq j$ die Ereignisse A_i und A_j unabhängig voneinander sind.

(b) Zwei Zufallsvariablen X und Y heißen *unabhängig*, wenn für alle Werte $r, s \in \mathbb{R}$ gilt:

$$\mathrm{P}(X = r \text{ und } Y = s) \quad = \quad \mathrm{P}(X = r) \cdot \mathrm{P}(Y = s).$$

Für jede Zahl $s \in \mathbb{N}_{\geqslant 1}$ heißen die Zufallsvariablen $X_1, \ldots, X_s$ *paarweise unabhängig* (kurz: *pw. unabh.*), wenn für alle $i, j \in \{1, \ldots, s\}$ mit $i \neq j$ die Zufallsvariablen X_i und X_j unabhängig voneinander sind.

Folie 297

Vorsicht beim Begriff der (stochastischen) Unabhängigkeit

Bemerkung 5.26. Per Definition sind 2 Ereignisse A und B genau dann (stochastisch) unabhängig, wenn gilt: $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B)$.

Falls (Ω, P) ein Laplacescher Wahrscheinlichkeitsraum ist, so wissen wir, dass Folgendes für alle Ereignisse A und B gilt:

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

Also sind die Ereignisse A und B genau dann (stochastisch) unabhängig, wenn gilt:

$$|A \cap B| \quad = \quad \frac{|A| \cdot |B|}{|\Omega|}.$$

Stochastische Unabhängigkeit hat *nichts* damit zu tun, ob die Ereignisse disjunkt sind oder nicht.
Beispiel: Ist $\mathrm{P}(A) > 0$ und $\mathrm{P}(B) > 0$ und $A \cap B = \emptyset$, so ist $\mathrm{P}(A \cap B) = 0 \neq \mathrm{P}(A) \cdot \mathrm{P}(B)$. Somit sind die Ereignisse A, B *nicht* unabhängig!

Folie 298

Beispiel 5.27.

(a) Wir betrachten das Szenario aus Beispiel 5.2, bei dem eine faire Münze zwei mal hintereinander geworfen wird.

Sei A_1 das Ereignis, dass die Münze beim ersten Wurf auf „Zahl" landet, sei A_2 das Ereignis, dass die Münze beim zweiten Wurf auf „Kopf" landet, und sei A_3 das Ereignis, dass die Münze bei beiden Münzwürfen auf der gleichen Seite landet. Also ist

$$A_1 = \{ZK,\ ZZ\}\ , \quad A_2 = \{KK,\ ZK\}\ , \quad A_3 = \{KK,\ ZZ\}\ ,$$

und es gilt

$$\mathrm{P}(A_1) = \tfrac{1}{2}\ , \qquad \mathrm{P}(A_2) = \tfrac{1}{2}\ , \qquad \mathrm{P}(A_3) = \tfrac{1}{2}\ .$$

Außerdem gilt

$$\begin{array}{rcccccl} \mathrm{P}(A_1 \cap A_2) & = & \mathrm{P}(\{ZK\}) & = & \frac{1}{4} & = & \mathrm{P}(A_1) \cdot \mathrm{P}(A_2). \\ \mathrm{P}(A_1 \cap A_3) & = & \mathrm{P}(\{ZZ\}) & = & \frac{1}{4} & = & \mathrm{P}(A_1) \cdot \mathrm{P}(A_3). \\ \mathrm{P}(A_2 \cap A_3) & = & \mathrm{P}(\{KK\}) & = & \frac{1}{4} & = & \mathrm{P}(A_2) \cdot \mathrm{P}(A_3). \end{array}$$

Daher sind die drei Ereignisse A_1, A_2, A_3 paarweise unabhängig.

Man beachte, dass trotz der paarweisen Unabhängigkeit die Ereignisse hier *nicht* paarweise disjunkt sind. Außerdem gilt hier

$$\mathrm{P}(A_1 \cap A_2 \cap A_3) \quad \neq \quad \mathrm{P}(A_1) \cdot \mathrm{P}(A_2) \cdot \mathrm{P}(A_3),$$

denn $A_1 \cap A_2 \cap A_3 = \emptyset$, d.h. die drei Ereignisse können nicht alle gleichermaßen eintreten, und daher ist $\mathrm{P}(A_1 \cap A_2 \cap A_3) = 0$, wohingegen $\mathrm{P}(A_1) \cdot \mathrm{P}(A_2) \cdot \mathrm{P}(A_3) = \frac{1}{8}$ ist.

(b) Entsprechend dem in Teil (a) betrachteten Szenario definieren wir die Zufallsvariablen Y_1, Y_2, Y_3 wie folgt:

$$Y_1 \ := \ \begin{cases} 1 & \text{falls die Münze beim 1. Wurf auf „Zahl" landet} \\ 0 & \text{sonst} \end{cases}$$

$$Y_2 \ := \ \begin{cases} 1 & \text{falls die Münze beim 2. Wurf auf „Kopf" landet} \\ 0 & \text{sonst} \end{cases}$$

$$Y_3 \ := \ \begin{cases} 1 & \text{falls die Münze bei beiden Würfen auf der gleichen Seite landet} \\ 0 & \text{sonst} \end{cases}$$

Somit gilt für jedes $i \in \{1,2,3\}$

$$P(Y_i = 1) \quad = \quad P(A_i) \quad = \quad \tfrac{1}{2} \qquad \text{und} \qquad P(Y_i = 0) \quad = \quad \tfrac{1}{2}.$$

Man kann sich leicht davon überzeugen, dass die drei Zufallsvariablen Y_1, Y_2, Y_3 paarweise unabhängig sind.

Man beachte, dass trotz der paarweisen Unabhängigkeit die drei Zufallsvariablen nicht „vollständig unabhängig" voneinander sind. Beispielsweise gilt

$$P(Y_1 = 1 \text{ und } Y_2 = 1 \text{ und } Y_3 = 0) \quad \neq \quad P(Y_1 = 1) \cdot P(Y_2 = 1) \cdot P(Y_3 = 0),$$

da $P(Y_1 = 1 \text{ und } Y_2 = 1 \text{ und } Y_3 = 0) = P(\{ZK\}) = \frac{1}{4}$ ist, wohingegen $P(Y_1 = 1) \cdot P(Y_2 = 1) \cdot P(Y_3 = 0) = \frac{1}{8}$ ist.

Folie 299

Für *paarweise unabhängige* Zufallsvariablen gelten für Erwartungswert und Varianz noch die folgenden sehr nützlichen Regeln.

Satz 5.28 (Erwartungswert & Varianz von pw. unabh. Zufallsvariablen)**.** *Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum, sei $s \in \mathbb{N}_{\geqslant 1}$ und seien $X_1, \ldots, X_s$ paarweise unabhängige Zufallsvariablen. Dann gilt:*

$$E(X_i \cdot X_j) \quad = \quad E(X_i) \cdot E(X_j)$$

für alle $i, j \in \{1, \ldots, s\}$ mit $i \neq j$, und es gilt

$$\operatorname{Var}\Big(\sum_{i=1}^{s} X_i\Big) \quad = \quad \sum_{i=1}^{s} \operatorname{Var}(X_i).$$

Beweis. Seien $i, j \in \{1, \ldots, s\}$ mit $i \neq j$. Wir setzen $Y := X_i$ und $Z := X_j$. Es gilt:

$$\begin{aligned} E(Y) \cdot E(Z) \quad &= \quad \Big(\sum_{y \in \mathit{Bild}(Y)} y \cdot P(Y = y)\Big) \cdot \Big(\sum_{z \in \mathit{Bild}(Z)} z \cdot P(Z = z)\Big) \\ &= \quad \sum_{\substack{y \in \mathit{Bild}(Y), \\ z \in \mathit{Bild}(Z)}} y \cdot z \cdot P(Y = y) \cdot P(Z = z) \end{aligned}$$

Laut Voraussetzung sind Y und Z unabhängig, d.h. es gilt für alle Werte $y, z \in \mathbb{R}$, dass

$$P(Y = y \text{ und } Z = z) \quad = \quad P(Y = y) \cdot P(Z = z).$$

Somit ist

$$\begin{aligned}
\mathrm{E}(Y)\cdot\mathrm{E}(Z) &= \sum_{\substack{y\in \mathit{Bild}(Y),\\ z\in \mathit{Bild}(Z)}} y\cdot z\cdot \mathrm{P}(Y=y \text{ und } Z=z)\\
&= \sum_{x\in \mathit{Bild}(YZ)} x\cdot\Big(\sum_{\substack{y\in \mathit{Bild}(Y),\\ z\in \mathit{Bild}(Z)\\ \text{mit } yz=x}} \mathrm{P}(Y=y \text{ und } Z=z)\Big)\\
&= \sum_{x\in \mathit{Bild}(YZ)} x\cdot \mathrm{P}(YZ=x)\\
&= \mathrm{E}(YZ).
\end{aligned}$$

Wir haben also gezeigt, dass für alle $i,j\in\{1,\dots,s\}$ mit $i\neq j$ gilt:

$$\mathrm{E}(X_i\cdot X_j) \quad=\quad \mathrm{E}(X_i)\cdot\mathrm{E}(X_j).$$

Um zu zeigen, dass

$$\mathrm{Var}(\sum_{i=1}^{s} X_i) \quad=\quad \sum_{i=1}^{s}\mathrm{Var}(X_i)$$

ist, beachten wir, dass $\mathrm{Var}(Z)=\mathrm{E}(Z^2)-\mathrm{E}(Z)^2$ für jede Zufallsvariable Z gilt. Speziell für

$$Z \quad:=\quad \sum_{i=1}^{s} X_i$$

ist

$$Z^2 \quad=\quad \sum_{i,j=1}^{s} X_i{\cdot}X_j \quad=\quad \sum_{i=1}^{s}(X_i)^2 \;+\; \sum_{\substack{i,j\in\{1,\dots,s\}\\ \text{mit } i\neq j}} X_i{\cdot}X_j.$$

Auf Grund der Linearität des Erwartungswerts gilt daher:

$$\mathrm{E}(Z^2) \quad=\quad \sum_{i=1}^{s}\mathrm{E}({X_i}^2) \;+\; \sum_{\substack{i,j\in\{1,\dots,s\}\\ \text{mit } i\neq j}} \mathrm{E}(X_i{\cdot}X_j). \tag{5.7}$$

Andererseits gilt auf Grund der Linearität des Erwartungswerts auch, dass $\mathrm{E}(Z)=\sum_{i=1}^{s}\mathrm{E}(X_i)$. Somit ist

$$\begin{aligned}
\mathrm{E}(Z)^2 &= \sum_{i,j=1}^{s}\mathrm{E}(X_i){\cdot}\mathrm{E}(X_j)\\
&= \sum_{i=1}^{s}\mathrm{E}(X_i)^2 \;+\; \sum_{\substack{i,j\in\{1,\dots,s\}\\ \text{mit } i\neq j}} \mathrm{E}(X_i){\cdot}\mathrm{E}(X_j).
\end{aligned}$$

Die paarweise Unabhängigkeit der Zufallsvariablen $X_1, \ldots, X_s$ liefert uns, dass $\mathrm{E}(X_i \cdot X_j) = \mathrm{E}(X_i) \cdot \mathrm{E}(X_j)$ für alle $i, j \in \{1, \ldots, s\}$ mit $i \neq j$ gilt. Somit ist

$$\mathrm{E}(Z)^2 \quad = \quad \sum_{i=1}^{s} \mathrm{E}(X_i)^2 \quad + \sum_{\substack{i,j \in \{1,\ldots,s\} \\ \text{mit } i \neq j}} \mathrm{E}(X_i \cdot X_j) \tag{5.8}$$

Aus den Gleichungen (5.7) und (5.8) erhalten wir, dass

$$\begin{aligned} \mathrm{Var}(Z) &= \mathrm{E}(Z^2) - \mathrm{E}(Z)^2 \\ &= \sum_{i=1}^{s} \mathrm{E}({X_i}^2) \quad - \quad \sum_{i=1}^{s} \mathrm{E}(X_i)^2 \\ &= \sum_{i=1}^{s} \Big(\mathrm{E}({X_i}^2) - \mathrm{E}(X_i)^2 \Big) \\ &= \sum_{i=1}^{s} \mathrm{Var}(X_i). \end{aligned}$$

□

5.7 Ein Beispiel zur Verwendung der Tschebycheff-Ungleichung

Folie 300

Der Zufallssurfer

Suchmaschinen im Internet beruhen u.a. darauf, die von ihnen gefundene Liste der Ergebnisse auf eine eingegebene Suchanfrage möglichst gut nach ihrer "Relevanz" zu sortieren, so dass die am besten geeigneten Treffer möglichst weit oben in der Liste der Suchergebnisse angezeigt werden.

Eine Grundlage des von Suchmaschinen verwendeten Verfahrens beruht auf dem Modell des *Zufallssurfers* (engl.: *Random Surfer*). Hierbei stellt man sich vor, dass die einzelnen Webseiten Knoten eines gerichteten Graphen sind, und dass ein Link, der von einer Webseite auf eine andere Webseite führt, einer gerichteten Kante entspricht.

Ein "zufällig vorgehender" Web-Surfer beginnt auf einer beliebigen Webseite und klickt sich nach und nach durchs Internet, indem er jeweils einem zufällig ausgewählten Link folgt. Dies entspricht dem in der Stochastik seit vielen Jahrzehnten untersuchten Konzept der *Markov-Ketten*.

Als Maß für das "Renommee" einer einzelnen Webseite nimmt man dann die Wahrscheinlichkeit, mit der der Zufallssurfer nach langem hin- und her-klicken auf dieser Webseite landet. Und dieses “Renommee” ist einer der Bestandteile, die in die Sortierung der Suchergebnisse eingehen, um möglichst relevante Treffer möglichst weit oben in der Liste der Suchergebnisse anzuzeigen.

Die Analyse der Bewegungen des Zufallssurfers ist für die Veranstaltung „Diskrete Strukturen“ zu anspruchsvoll. Im folgenden Beispiel betrachten wir ein ähnliches, aber deutlich einfacheres Szenario.

Folie 301

Beispiel 5.29 (Der Frosch)**.**
In einem Teich befinden sich Steine in einer langen Reihe. Die Steine sind durchnummeriert mit $\dots, -2, -1, 0, 1, 2, \dots$
Ein Frosch sitzt anfänglich auf Stein 0. Dann beginnt er mit gleicher Wahrscheinlichkeit $\frac{1}{2}$ immer wieder entweder nach rechts oder nach links zu springen.

Frage: Mit welcher Wahrscheinlichkeit wird der Frosch nach n Sprüngen vom Anfangsstein 0 mindestens a Steine weit entfernt sein (für eine gegebene Zahl a)?

Für jedes $i \in [n]$ sei $X_i = +1$, falls der Frosch im i-ten Schritt nach rechts springt, und $X_i = -1$, falls er im i-ten Schritt nach links springt.
Sei $X := \sum_{i=1}^{n} X_i$. D.h. X gibt an, auf welchem Stein er sich nach n Sprüngen befindet.

Für jedes $i \in [n]$ gilt: $\mathrm{E}(X_i) = (-1) \cdot \frac{1}{2} + (+1) \cdot \frac{1}{2} = 0$.
Also gilt:

$$\mathrm{E}(X) \;=\; \mathrm{E}\Big(\sum_{i=1}^{n} X_i\Big) \;=\; \sum_{i=1}^{n} \mathrm{E}(X_i) \;=\; \sum_{i=1}^{n} 0 \;=\; 0\,.$$

D.h.: Nach n Sprüngen sitzt der Frosch “im Erwartungswert” wieder auf seinem ursprünglichen Ausgangspunkt.

Aber mit welcher Wahrscheinlichkeit hat er dennoch eine Entfernung $\geqslant a$ vom Startpunkt (für eine gegebene Zahl a)?

Die Zufallsvariable $|X - 0| = |X|$ gibt die Entfernung vom Startpunkt nach n Sprüngen an. Die Tschebycheff-Ungleichung liefert (beachte dazu, dass $\mathrm{E}(X) = 0$ ist):

$$\mathrm{P}(|X| \geqslant a) \;=\; \mathrm{P}(|X - \mathrm{E}(X)| \geqslant a) \;\leqslant\; \frac{\mathrm{Var}(X)}{a^2}\,. \tag{5.9}$$

Wir nutzen Satz 5.28, um die Varianz $\mathrm{Var}(X)$ auszurechnen: Die Zufallsvariablen $X_1, \ldots, X_n$ sind paarweise unabhängig voneinander. Daher gilt:

$$\mathrm{Var}(X) \quad = \quad \mathrm{Var}(\sum_{i=1}^{n} X_i) \quad = \quad \sum_{i=1}^{n} \mathrm{Var}(X_i)\,.$$

Für jedes einzelne $i \in [n]$ gilt:
$\mathrm{E}((X_i)^2) = (-1)^2 \cdot \frac{1}{2} + (+1)^2 \cdot \frac{1}{2} = 1$.
Also ist

$$\mathrm{Var}(X_i) \quad = \quad \mathrm{E}((X_i)^2) - \mathrm{E}(X_i) \quad = \quad 1 - 0 \quad = \quad 1\,.$$

Somit gilt:

$$\mathrm{Var}(X) \quad = \quad \sum_{i=1}^{n} \mathrm{Var}(X_i) \quad = \quad n\,.$$

Wir setzen dies in (5.9) ein und erhalten:

$$\mathrm{P}(|X| \geqslant a) \quad \leqslant \quad \frac{\mathrm{Var}(X)}{a^2} \quad = \quad \frac{n}{a^2}\,.$$

Somit gilt für jedes $a > 0$: Die Wahrscheinlichkeit, dass der Frosch nach n Sprüngen einen Abstand $\geqslant a$ vom Startpunkt hat, ist $\leqslant \frac{n}{a^2}$.
Insbes. gilt z.B. bei $n = 100$ Sprüngen: Die Wahrscheinlichkeit, dass er am Ende einen Abstand von $\geqslant 20$ vom Startpunkt hat, ist $\leqslant \frac{100}{20^2} = \frac{100}{400} = \frac{1}{4}$.

5.8 Bedingte Wahrscheinlichkeiten

Folie 302

Definition 5.30. Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum. Seien A, B zwei Ereignisse.
Die *bedingte Wahrscheinlichkeit* $\mathrm{P}(A|B)$ für das Ereignis A unter der Bedingung B ist definiert durch

$$\mathrm{P}(A|B) \quad := \quad \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}\,.$$

Die Wahrscheinlichkeit $\mathrm{P}(A|B)$ bezeichnet man auch als *a-posteriori Wahrscheinlichkeit* von A bzgl. B.

Interpretation: Wir stellen uns vor, dass wir ein Zufallsexperiment durchgeführt haben und bereits wissen, dass Ereignis B eingetreten ist. Der Wert $\mathrm{P}(A|B)$ gibt an, mit welcher Wahrscheinlichkeit hierbei auch das Ereignis A eingetreten ist.

Folie 303

Beispiel 5.31. Wir führen unabhängig voneinander drei Würfe mit einer fairen Münze durch. Der Wahrscheinlichkeitsraum ist also (Ω, P) mit $\Omega = \{K, Z\}^3$ und $\mathrm{P}(\omega) = \frac{1}{|\Omega|} = \frac{1}{8}$ für jedes $\omega \in \Omega$.

Sei A das Ereignis, dass mind. 2 der 3 Münzwürfe das Ergebnis "K" liefern. Dann ist $\mathrm{P}(A) = \frac{4}{8} = \frac{1}{2}$.

Nehmen wir mal an, dass wir den ersten Münzwurf gesehen haben und wissen, dass dessen Ergebnis "Z" war. Nun wollen wir wissen, mit welcher Wahrscheinlichkeit unter den insgesamt 3 Münzwürfen dennoch das Ereignis A eingetreten ist.

Sei B das Ereignis, dass beim ersten Münzwurf das Ergebnis "Z" eingetreten ist. Dann ist $\mathrm{P}(B) = \frac{1}{2}$.

Außerdem ist $\mathrm{P}(A \cap B) = \frac{1}{8}$, denn $A \cap B = \{ZKK\}$.

Somit ist

$$\mathrm{P}(A|B) \;=\; \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \;=\; \tfrac{1}{8} \cdot \tfrac{2}{1} \;=\; \tfrac{1}{4}\,.$$

Folie 304

Bemerkung 5.32. Zwei Ereignisse A, B sind genau dann (stochastisch) unabhängig voneinander, wenn gilt: $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B)$.
Somit gilt für beliebige Ereignisse A, B:

$$\mathrm{P}(A|B) = \mathrm{P}(A) \quad \Longleftrightarrow \quad A \text{ und } B \text{ sind (stochastisch) unabhängig.}$$

Folie 305

Rechenregeln

Bemerkung 5.33. Für jeden Wahrscheinlichkeitsraum (Ω, P) und alle Ereignisse A und B gilt:

(a) *Multiplikationssatz für Wahrscheinlichkeiten*:

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

(b) *Satz von Bayes*: Ist $\mathrm{P}(A) \neq 0$ und $\mathrm{P}(B) \neq 0$, so gilt:

$$\mathrm{P}(A|B) \;=\; \frac{\mathrm{P}(A)}{\mathrm{P}(B)} \cdot \mathrm{P}(B|A)\,.$$

(c) *Satz der totalen Wahrscheinlichkeit*: Sei $\ell \in \mathbb{N}_{\geqslant 1}$ und seien $B_1, \ldots, B_\ell \subseteq \Omega$, so dass gilt: $\Omega = \bigcup_{i \in [\ell]} B_i$ und die Mengen $B_1, \ldots, B_\ell$ sind paarweise disjunkt. Dann gilt:

$$\mathrm{P}(A) \quad = \quad \sum_{i=1}^{\ell} \mathrm{P}(A \cap B_i) \quad = \quad \sum_{i=1}^{\ell} \mathrm{P}(B_i) \cdot \mathrm{P}(A|B_i)\,.$$

Beweis. (a): Dies folgt direkt aus Definition 5.30.

(b): Gemäß Definition 5.30 gilt: $\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$ und $\mathrm{P}(B|A) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)}$. Somit gilt: $\mathrm{P}(A \cap B) = \mathrm{P}(A|B) \cdot \mathrm{P}(B)$ und $\mathrm{P}(A \cap B) = \mathrm{P}(B|A) \cdot \mathrm{P}(A)$. Also gilt: $\mathrm{P}(A|B) \cdot \mathrm{P}(B) = \mathrm{P}(B|A) \cdot \mathrm{P}(A)$. Teilen durch $\mathrm{P}(B)$ liefert: $\mathrm{P}(A|B) = \frac{\mathrm{P}(A)}{\mathrm{P}(B)} \cdot \mathrm{P}(B|A)$.

(c): Es gilt: $A = \bigcup_{i=1}^{\ell} (A \cap B_i)$, und die Ereignisse $(A \cap B_i)$ für $i \in \{1, \ldots, \ell\}$ sind paarweise disjunkt. Gemäß (5.4) gilt daher:

$$\mathrm{P}(A) \quad = \quad \sum_{i=1}^{\ell} \mathrm{P}(A \cap B_i)\,.$$

Außerdem gilt gemäß Definition 5.30 für jedes $i \in [\ell]$:

$$\mathrm{P}(A|B_i) \quad = \quad \frac{\mathrm{P}(A \cap B_i)}{\mathrm{P}(B_i)}\,,$$

d.h.: $\mathrm{P}(A \cap B) = \mathrm{P}(B_i) \cdot \mathrm{P}(A|B_i)$. Also gilt:

$$\mathrm{P}(A) \quad = \quad \sum_{i=1}^{\ell} \mathrm{P}(A \cap B_i) \quad = \quad \sum_{i=1}^{\ell} \mathrm{P}(B_i) \cdot \mathrm{P}(A|B_i)\,.$$

$\square$

Folie 306

Beispiel 5.34. Eine Studie[6] zur Zuverlässigkeit von Antigen-Schnelltests auf SARS-CoV-2 kam zu folgendem Ergebnis innerhalb einer Gruppe von Personen, die Symptome einer SARS-CoV-2 Erkrankung hatten und einen Antigen-Schnelltest durchführten:

[6] https://www.cochrane.de/news/aktueller-cochrane-review-wie-zuverlaessig-sind-corona-sc

(a) 5% der Personen war tatsächlich an SARS-CoV-2 erkrankt;
95% der Personen war gesund;

(b) $4,5\%$ der Personen hatte ein positives Testergebnis;
$95,5\%$ der Personen hatte ein negatives Testergebnis;

(c) $99,47\%$ der gesunden Personen hatte ein negatives Testergebnis;
$0,53\%$ der gesunden Personen hatte ein positives Testergebnis (d.h. der Test war "falsch positiv").

(d) 80% der kranken Personen hatte ein positives Testergebnis;
20% der kranken Personen hatte ein negatives Testergebnis (d.h. der Test war "falsch negativ");

Folie 307

Wir wollen nun die beiden folgenden Fragen beantworten:

1. Wenn ich ein negatives Testergebnis habe: Wie groß ist die Wahrscheinlichkeit, dass ich tatsächlich gesund bin?
2. Wenn ich ein positives Testergebnis habe: Wie groß ist die Wahrscheinlichkeit, dass ich tatsächlich krank bin?

Um diese Fragen zu beantworten, betrachten wir die folgenden Ereignisse:

- N: "das Testergebnis ist negativ" (dann ist $\overline{N}$ das Ereignis "das Testergebnis ist positiv")
- G: "die gestestete Person ist gesund" (dann ist $\overline{G}$ das Ereignis "die getestete Person ist krank")

Wir sehen das Testergebnis und wissen daher, welches der beiden Ereignisse N bzw. $\overline{N}$ eingetreten ist.
Bei Frage 1 wollen wir wissen, wie groß die Wahrscheinlichkeit $\mathrm{P}(G|N)$ ist.
Bei Frage 2 wollen wir wissen, wie groß die Wahrscheinlichkeit $\mathrm{P}(\overline{G}|\overline{N})$ ist.

Folie 308

Wir wissen:

- Gemäß (a) ist $\mathrm{P}(G) = 0,95$ und $\mathrm{P}(\overline{G}) = 0,05$.
- Gemäß (b) ist $\mathrm{P}(N) = 0,955$ und $\mathrm{P}(\overline{N}) = 0,045$.
- Gemäß (c) ist $\mathrm{P}(N|G) = 0,9947$ und $\mathrm{P}(\overline{N}|G) = 0,0053$.

- Gemäß (d) ist $P(\overline{N}|\overline{G}) = 0,8$ und $P(N|\overline{G}) = 0,2$.

Um die Fragen 1 und 2 zu beantworten, können wir den Satz von Bayes benutzen. Für Frage 1 erhalten wir:

$$P(G|N) \quad = \quad \frac{P(G)}{P(N)} \cdot P(N|G) \quad = \quad \frac{0,95}{0,955} \cdot 0,9947 \quad = \quad 0,98949\cdots .$$

D.h.: Wenn ich ein negatives Testergebnis habe, ist (gemäß der zitierten Studie) die Wahrscheinlichkeit, dass ich tatsächlich gesund bin, etwa $98,95\%$.
Für Frage 2 erhalten wir:

$$P(\overline{G}|\overline{N}) \quad = \quad \frac{P(\overline{G})}{P(\overline{N})} \cdot P(\overline{N}|\overline{G}) \quad = \quad \frac{0,05}{0,045} \cdot 0,8 \quad = \quad 0,\overline{8} .$$

D.h.: Wenn ich ein positives Testergebnis habe, ist (gemäß der zitierten Studie) die Wahrscheinlichkeit, dass ich tatsächlich krank bin, $88,\overline{8}\%$. Es besteht also immerhin noch eine Wahrscheinlichkeit von $11,\overline{1}\%$ dass ich gesund bin.

5.9 Einige wichtige Wahrscheinlichkeitsräume

Folie 309

Einfache Urnenmodelle

Beispiel 5.35. In einer Urne befinden sich n Kugeln, die mit den Zahlen $1, \ldots, n$ durchnummeriert sind. Wir ziehen zufällig nacheinander k Kugeln. Wir nutzen im Folgenden unsere Erkenntnisse aus Kapitel 4.2.

Variante 1: *Mit Zurücklegen und mit Berücksichtigung der Reihenfolge*

Dies entspricht dem Wahrscheinlichkeitsraum $(\Omega_{\text{mm}}, P_{\text{mm}})$ mit

$$\Omega_{\text{mm}} \ := \ \{(a_1, \ldots, a_k) \ : \ a_1, \ldots, a_k \in [n]\}$$

mit $|\Omega_{\text{mm}}| = n^k$ und $P_{\text{mm}}(\omega) = \frac{1}{|\Omega_{\text{mm}}|}$ für alle $\omega \in \Omega_{\text{mm}}$.

Variante 2: *Ohne Zurücklegen und mit Berücksichtigung der Reihenfolge*

Dies entspricht dem Wahrscheinlichkeitsraum $(\Omega_{\text{om}}, P_{\text{om}})$ mit

$$\Omega_{\text{om}} \ := \ \{(a_1, \ldots, a_k) \in [n]^k \ : \ a_1, \ldots, a_k \text{ sind paarweise verschieden}\}$$

mit $|\Omega_{\text{om}}| = (n)_k$ und $P_{\text{om}}(\omega) = \frac{1}{|\Omega_{\text{om}}|}$ für alle $\omega \in \Omega_{\text{om}}$.

Folie 310

Variante 3: *Ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge*

Dies entspricht dem Wahrscheinlichkeitsraum $(\Omega_{\text{oo}}, \mathrm{P}_{\text{oo}})$ mit

$$\Omega_{\text{oo}} \ := \ \{X \subseteq [n] \ : \ |X| = k\}$$

mit $|\Omega_{\text{oo}}| = \binom{n}{k}$ und $\mathrm{P}_{\text{oo}}(\omega) = \frac{1}{|\Omega_{\text{oo}}|}$ für alle $\omega \in \Omega_{\text{oo}}$.

Variante 4: *Mit Zurücklegen und ohne Berücksichtigung der Reihenfolge*

Dies entspricht dem Wahrscheinlichkeitsraum $(\Omega_{\text{mo}}, \mathrm{P}_{\text{mo}})$ mit

$$\Omega_{\text{mo}} \ := \ \{f : [n] \to \mathbb{N} \ : \ \sum_{x=1}^{n} f(x) = k\}$$

mit $|\Omega_{\text{mo}}| = \binom{k+n-1}{n-1}$ und $\mathrm{P}_{\text{mo}}(\omega) = \frac{1}{|\Omega_{\text{mo}}|}$ für alle $\omega \in \Omega_{\text{mo}}$.

Folie 311

Produktexperimente

Wir stellen uns vor, dass wir nacheinander n Zufallsexperimente unabhängig voneinander durchführen und deren Ergebnisse in einem n-Tupel zusammenfassen. Beispielsweise könnten wir dreimal hintereinander eine Münze werfen, danach zweimal würfeln und dann noch ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge vier Kugeln aus einer Urne mit insgesamt 15 Kugeln ziehen.

Definition 5.36. Sei $n \in \mathbb{N}_{\geqslant 1}$, und für jedes $i \in [n]$ sei (Ω_i, P_i) ein endlicher Wahrscheinlichkeitsraum.
Das *Produkt der Wahrscheinlichkeitsräume* (Ω_i, P_i) für $i \in [n]$ ist (Ω, P) mit

$$\Omega \ := \ \Omega_1 \times \cdots \times \Omega_n$$

und, für jedes $\omega = (\omega_1, \ldots, \omega_n) \in \Omega$,

$$\mathrm{P}(\omega) \ := \ \prod_{i=1}^{n} \mathrm{P}_i(\omega_i)\,.$$

Man kann sich leicht davon überzeugen, dass (Ω, P) ein endlicher Wahrscheinlichkeitsraum ist (Details: Übungsaufgabe).

Folie 312

Die Bernoulli-Verteilung

Ein *Bernoulli-Experiment* ist ein Zufallsexperiment, bei dem es nur 2 verschiedene Ausgänge gibt. Das Paradebeispiel dafür ist der Wurf einer Münze. Wir sprechen von einem "Erfolg", wenn die Münze auf "Kopf" landet, und von einem "Misserfolg", wenn sie auf "Zahl" landet.
Es sei p (mit $0 \leqslant p \leqslant 1$) die "Erfolgswahrscheinlichkeit", d.h. die Wahrscheinlichkeit, dass die Münze auf "Kopf" landet.
Als Wahrscheinlichkeitsraum nutzen wir (Ω, P) mit $\Omega = \{K, Z\}$ und $\mathrm{P}(K) = p$ und $\mathrm{P}(Z) = 1{-}p$. Die Verteilung P wird auch die *Bernoulli-Verteilung* genannt.
Die Zufallsvariable X mit $X(K) = 1$ und $X(Z) = 0$ wird auch *Indikatorvariable* genannt. Sie hat Erwartungswert

$$\mathrm{E}(X) \quad = \quad 1 \cdot \mathrm{P}(X=1) \ + \ 0 \cdot \mathrm{P}(X=0) \quad = \quad p$$

und Varianz

$$\mathrm{Var}(X) \quad = \quad \mathrm{E}(X^2) - \mathrm{E}(X)^2 \quad = \quad p - p^2 \quad = \quad p \cdot (1{-}p)\,.$$

Folie 313

Die *Binomialverteilung* $B(n,p)$

Wir führen unabhängig voneinander n Bernoulli-Experimente durch, bei denen jedes dieselbe Erfolgswahrscheinlichkeit p besitzt.
Für jedes $i \in [n]$ sei X_i die Indikatorvariable, die angibt, ob das i-te Experiment erfolgreich war oder nicht. D.h.: $\mathrm{P}(X_i = 1) = p$, $\mathrm{P}(X_i = 0) = 1 - p$, $\mathrm{E}(X_i) = p$ und $\mathrm{Var}(X_i) = p \cdot (1{-}p)$.
Wir betrachten die Zufallsvariable $S := \sum_{i=1}^n X_i$. D.h. S gibt an, wie viele der n Experimente erfolgreich waren.
Für jedes $k \in \{0, \ldots, n\}$ gilt:

$$\mathrm{P}(S = k) \quad = \quad \tbinom{n}{k} \cdot p^k \cdot (1{-}p)^{n-k}\,.$$

Es gilt gemäß der Linearität des Erwartungswerts:

$$\mathrm{E}(S) \quad = \quad \sum_{i=1}^{n} \mathrm{E}(X_i) \quad = \quad n \cdot p\,.$$

Da die Zufallsvariablen $X_1, \ldots, X_n$ paarweise unabhängig sind, gilt gemäß Satz 5.28:

$$\mathrm{Var}(S) \quad = \quad \sum_{i=1}^{n} \mathrm{Var}(X_i) \quad = \quad n \cdot p \cdot (1{-}p)\,.$$

Folie 314

Die *geometrische Verteilung*

Wir führen wieder unabhängig voneinander n Bernoulli-Experimente durch, bei denen jedes dieselbe Erfolgswahrscheinlichkeit p besitzt. Es sei $0 < p < 1$.
Für jedes $i \in [n]$ sei X_i die Indikatorvariable, die angibt, ob das i-te Experiment erfolgreich war oder nicht.
Wir betrachten die Zufallsvariable $Y := \min\{i \in [n] \ : \ X_i = 1\}$, d.h.: $Y = k$ bedeutet, dass die ersten $k{-}1$ Durchführungen des Zufallsexperiments zu einem Misserfolg führten und die k-te Durchführung erfolgreich war. Es gilt:

$$\mathrm{P}(Y = k) \quad = \quad (1{-}p)^{k-1} \cdot p\,.$$

Wenn wir jetzt kein festes n mehr vorgeben, sondern einfach das Bernoulli-Experiment immer weiter hintereinander durchführen, bis wir den ersten "Erfolg" erzielt haben, betrachten wir den *diskreten Wahrscheinlichkeitsraum* (Ω, P) mit $\Omega = \mathbb{N}_{\geqslant 1}$ und $\mathrm{P}(k) = (1{-}p)^{k-1} \cdot p$ für alle $k \in \Omega$.
Man kann nachrechnen (hier ohne Beweis), dass P eine Wahrscheinlichkeitsverteilung auf Ω ist und dass die Zufallsvariable Y den Erwartungswert $\mathrm{E}(Y) = \frac{1}{p}$ und die Varianz $\mathrm{Var}(Y) = \frac{1-p}{p^2}$ hat.

5.10 Ein weiteres Beispiel: Das Monty Hall Problem

Folie 315

Beispiel 5.37 (Das Monty Hall Problem)**.** In der TV-Spielshow „Geh aufs Ganze!" stehen drei "Tore" namens "Tor 1", "Tor 2", "Tor 3" zur Verfügung. Hinter einem davon verbirgt sich der Hauptpreis, hinter einem weiteren ein Trostpreis, und hinter dem verbliebenen eine Niete, die "der Zonk" genannt wird.

Das Spiel läuft zwischen dem Showmaster und einem Kandidaten ab, und zwar wie folgt:

1. Der Kandidat sucht sich eins der drei Tore aus.

2. Der Showmaster öffnet eines der beiden verbliebenen Tore — und zwar eins, hinter dem sich entweder der Zonk oder der Trostpreis befinden.

3. Der Kandidat kann entscheiden, ob er bei seinem in Punkt 1 gewählten Tor bleibt, oder ob er zu dem anderen noch geschlossenen Tor wechselt.

Danach ist das Spiel beendet und der Kandidat gewinnt das, was sich hinter dem Tor verbirgt, für das er sich in Punkt 3 entschieden hat.

Frage: Was sollte der Kandidat in Schritt 3 tun, um seine Chancen auf den Hauptgewinn zu maximieren?

Dieses Problem ist auch unter den Begriffen *Monty Hall Problem* bzw. *Ziegenproblem* bekannt.

Folie 316

Zur Beantwortung der Frage wenden wir wieder die 3-Schritt-Methode an.

Schritt 1: Finde den richtigen Wahrscheinlichkeitsraum

Wir wählen $\Omega := \{(i,j) \ : \ i,j \in \{1,2,3\}\}$. Hierbei steht $\omega = (i,j)$ für das Elementarereignis, bei dem der Hauptgewinn sich hinter "Tor i" verbirgt und der Kandidat in Punkt 1 "Tor j" wählt

Wir gehen davon aus, dass der Showmaster die Preise bzw. Nieten zufällig auf die 3 Tore verteilt hat und dass der Kandidat in Punkt 1 zufällig eins der drei Tore wählt (jedes mit derselben Wahrscheinlichkeit). Somit gilt für jedes $\omega \in \Omega$, dass $\mathrm{P}(\omega) = \frac{1}{|\Omega|} = \frac{1}{9}$ ist.

Folie 317

Schritt 2: Finde die relevanten Ereignisse

Wir untersuchen 2 verschiedene Strategien, die der Kandidat verfolgen könnte.

Strategie 1: Er bleibt bei dem in Punkt 1 gewählten Tor.

Strategie 2: Er wechselt in Punkt 3 zu dem (einzigen noch verbliebenen) anderen Tor.

Dass er bei Strategie 1 den Hauptgewinn bekommt, entspricht dem Ereignis $A := \{(i,i) \ : \ i \in [3]\}$.

Dass er bei Strategie 2 den Hauptgewinnn bekommt, entspricht dem Ereignis $B := \{(i,j) \ : \ i,j \in [3], \ i \neq j\} = \bar{A}$.

Schritt 3: Berechne die Wahrscheinlichkeiten der Ereignisse A und B

$\mathrm{P}(A) = \frac{|A|}{|\Omega|} = \frac{3}{9} = \frac{1}{3}$.

$\mathrm{P}(B) = \mathrm{P}(\bar{A}) = 1 - \mathrm{P}(A) = 1 - \frac{1}{3} = \frac{2}{3}$.

D.h.: Um eine möglichst große Gewinnwahrscheinlichkeit zu haben, sollte der Kandidat in Punkt 3 das Tor wechseln. Dann wird er mit Wahrscheinlichkeit $\frac{2}{3}$ den Hauptgewinn erhalten.

Intuitive Erläuterung: Wenn der Showmaster in Schritt 2 nicht ein Tor öffnen, sondern einfach fragen würde, ob der Kandidat lieber darauf wetten will, dass der Hauptgewinn sich hinter einem der beiden anderen (nicht von ihm in Schritt 1 gewählten) Tore befindet, könnte man leicht einsehen, dass "Wechseln" eine Erfolgswahrscheinlichkeit von $\frac{2}{3}$ hat.

5.11 Literaturhinweise

Als vertiefende Lektüre seien die Kapitel 11 und 12 von [Juk08] sowie das Lehrbuch [Kre91] empfohlen.

Quellennachweis: Teile dieses Kapitels basieren auf Teilen der Kapitel 11 und 12 von [Juk08].