# Vorlesungsfolien Stochastik 1 für Studierende der Informatik

Hendrik Baumann

Universität Hamburg

2016

## Kapitel 1: Einleitung

# Der Begriff Stochastik

Die Stochastik beschäftigt sich mit

- der Beschreibung und Analyse zufälliger Ereignisse und Prozesse
- und ihrer Einbettung in mathematische Modelle.

Etymologisch kommt der Begriff Stochastik aus dem Griechischen und bedeutet etwa

- „die Kunst des geschickten Vermutens“.

## 1.1 Historische Anfänge

# Erste Schritte

- 1654: Der Chevalier de Méré stellt Blaise Pascal (1623-1662) Fragen bezüglich Glücksspielen.
- Gesucht: Gewinnchancen, faire Einsätze, Ruinwahrscheinlichkeiten.
- Pascal beginnt daraufhin einen Briefwechsel mit Pierre de Fermat (1601-1665).
- Fermat findet Antworten.
- Der Begriff „Wahrscheinlichkeit“ tauchte in dieser Phase noch nicht auf.

# Glücksspiel: Ein Beispiel (1)

Eine der Aufgaben des Chevalier de Méré:

- Zwei Spieler (A und B) werfen neunmal eine faire Münze.
- Fällt Zahl, so gewinnt A eine Runde, fällt Bild, so gewinnt B eine Runde.
- Wer zuerst fünf Runden für sich entschieden hat, bekommt den kompletten Einsatz.
- Durch höhere Gewalt muss das Spiel nach sieben Runden beim Stand von 4:3 für A abgebrochen werden.
- Wie soll der Einsatz aufgeteilt werden?

# Glücksspiel: Ein Beispiel (2)

Naheliegende Ideen waren für damalige Spieler

- die Aufteilung im Verhältnis der gewonnenen Runden, also 4:3,
- oder die Aufteilung umgekehrt zum Verhältnis der noch zum Gesamtsieg benötigten gewonnenen Runden, also 2:1.

Pascal und Fermat:

- Aufteilung im Verhältnis der Gewinnchancen.

# Glücksspiel: Ein Beispiel (3)

- B kann nur gewinnen, wenn in den letzten beiden Runden jeweils Bild fällt.
- Wird eine Münze sehr oft geworfen, so fällt in etwa der Hälfte der Fälle Bild.
- Wird sehr oft eine Münze zweimal hintereinander geworfen, so fällt in etwa einem Viertel der Fälle zweimal Bild.
- Vom Spielstand 4:3 aus startend wird bei sehr häufiger Wiederholung der letzten zwei Würfe Spieler B noch in etwa einem Viertel aller Fälle gewinnen, Spieler A in etwa drei Viertel der Fälle.
- Die Aufteilung sollte also im Verhältnis 3:1 erfolgen.

# Zufallsexperimente und Ereignisse

Pascal und Fermat gingen davon aus, dass die Glücksspielsituation

- zwar unterschiedliche Ausgänge haben kann,
- aber im Prinzip beliebig oft wiederholbar ist.

Durch diese Eigenschaften werden *Zufallsexperimente* charakterisiert. Die Ausgänge eines Zufallsexperiment werden *Ereignisse* genannt.

# Relative Häufigkeiten und Wahrscheinlichkeiten

- Wird ein Zufallsexperiment $n$ mal ausgeführt, lässt sich für ein fixiertes Ereignis die *relative Häufigkeit* bestimmen.
- Dazu wird die Anzahl der Experimente, in denen das Ereignis eingetreten ist, durch $n$ geteilt.
- Die Erfahrung zeigte Pascal und Fermat, dass die relative Häufigkeit für $n \to \infty$ konvergiert.
- Dieser Grenzwert wird seit Jacob Bernoulli (1654-1705) und Abraham de Moivre (1667-1754) als *Wahrscheinlichkeit* bezeichnet.

# Axiomatisierung der Wahrscheinlichkeitstheorie

- 1655-1931: Fortschritte in der Wahrscheinlichkeitsrechnung (Fermat, Pascal, Huygens, Bernoulli, de Moivre, Bayes, Euler, Lagrange, Laplace, Legrende, Gauß, Poisson, Tschebyscheff, Lyapunov, Markov)
- Wahrscheinlichkeitsrechnung aber noch sehr umstritten: Mathematik unterscheidet zwischen wahren und falschen Aussagen. Wie passt das zu zufälligen Ereignissen?
- 1931-1933: Axiomatisierung der Wahrscheinlichkeitstheorie durch Andrej Nikolajewitsch Kolmogorov (1903-1987), Einbettung in Analysis und Maßtheorie
- Grundidee: Wahrscheinlichkeiten folgen Axiomen, diese sind wichtigen Eigenschaften relativer Häufigkeiten nachempfunden.
- Seitdem: Echte mathematische Disziplin, rasante Entwicklung

## 1.2 Zufall

Zusätzlich zu unklaren Begrifflichkeiten birgt die Stochastik für viele Menschen ein weiteres Problem: Für den menschlichen Beobachter wirken zum Beispiel

- das Ergebnis des Wurfes mit einer fairen Münze,
- die Wartezeit an einer Ampel,
- der Zeitpunkt des Zerfalls eines Atomkerns und
- der Börsenkurs einer Aktie

dem „Zufall“ unterworfen.

**Aber was ist eigentlich Zufall?**

# Wirklich Zufall? (1)

Das Ergebnis des Münzwurfes hängt kausal (deterministisch)

- von Wurfkraft, Effet, Raumklima, ... ab,
- die physikalischen Gesetze hierfür sind prinzipiell bekannt.

Die Wartezeit an einer Ampel hängt kausal (deterministisch)

- von dem Verkehr und der Steuerung der Ampel ab (Länge der Grün- und Rotphasen, Kontaktschaltungen, ... ),
- diese Abhängigkeiten sind vergleichsweise einfach.

Trotzdem wird in beiden Fällen von Glück und Pech, also Synonymen für Zufall gesprochen. Warum?

# Wirklich Zufall? (2)

- Die Physik eines Münz- oder Würfelwurfs ist sehr komplex.
- Die Einflussfaktoren sind für einen Spieler nicht hinreichend genau messbar.
- Um die Wartezeit an einer Ampel zu bestimmen, fehlen Verkehrsteilnehmern Informationen (andere Verkehrsteilnehmer und die Ampelsteuerung betreffend).

# Wirklich Zufall? (3)

Der Begriff Zufall wird umgangssprachlich fast ausschließlich verwendet, wenn ein Ereignis zwar kausale Ursachen hat,

- aber diese Kausalität hochkomplex ist und Informationen über die Ursachen nicht oder nicht genügend genau messbar sind,
- die Ursachen nicht beobachtbar sind oder
- die genaue Form der Kausalität bisher nicht erfassbar ist.

# Wirklich Zufall? (4)

- Zerfall von Atomkernen: Es gibt keine lokale Ursachen für die konkreten Zeitpunkte.
- Echter (objektiver) Zufall in der Quantenphysik? Physik-philosophische Diskussion...
- Börsenkurs: Abhängig von Entscheidung anderer Marktteilnehmer.
- Freier Wille als echter Zufall? Philosophische, theologische und psychologische Diskussion...

## 1.3 Stochastische Modellierung

# Stochastische Modellierung

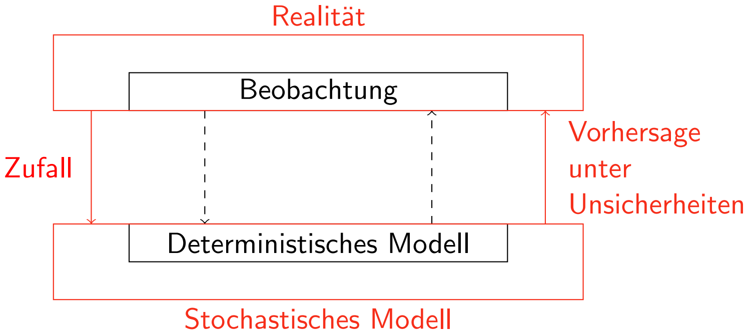
- Mathematische Modellierung: Eine oder mehrere Zielgrößen werden als Funktionen von Einflussfaktoren charakterisiert.
- Deterministischer Fall: Gleiche Werte der Einflussfaktoren liefern gleiche Werte der Zielgrößen.
- Stochastischer Fall: Bei gleichen Werten der Einflussfaktoren sind verschiedene Werte möglich.
- Zumeist: Zielgröße hängt von weiteren, nicht beobachtbaren oder nicht messbaren Ursachen ab, diese werden als „Zufallseinfluss“ modelliert.

# Verteilungen

- Zur Modellierung von Zufallseinflüssen nötig: Quantifizierung des Zufalls, verschiedene *Verteilungstypen* und Parameter.
- Beispiel 1: Beschrifte fairen Würfel mit Zahlen 0 bis 5, betrachte Ergebnis eines einfachen Wurfs.
- Beispiel 2: Werfe faire Münze fünfmal, zähle, wie häufig Zahl erscheint.
- In Beispiel 1: Jede Zahl erscheint mit gleicher Wahrscheinlichkeit, also $\frac{1}{6}$.
- In Beispiel 2: Anzahlen unterschiedlich wahrscheinlich, die Wahrscheinlichkeit für fünfmal Zahl ist etwa $\left(\frac{1}{2}\right)^5 = \frac{1}{32}$.

# Deterministische Modellierung...

# ... und stochastische Modellierung



# Aufgaben der Stochastik

Wesentliche Aufgabe der Stochastik ist die *Quantifizierung von Unsicherheiten*

- bei der Modellierung der Zufallseinflüsse und
- bei der Vorhersage.

Durch Vergleich (testweise) vorhergesagter Daten und realer Daten ist darüberhinaus die

- Validierung stochastischer Modelle

eine wichtige Aufgabe.

# Bereiche der Stochastik

- Diese Veranstaltung: Basiswissen, Verteilungstypen, Rechnen mit Wahrscheinlichkeiten, einfache Grenzwertsätze.
- Stochastische Prozesse: Zeitliche geordnete Prozesse mit Zufallseinflüssen (Markovketten, Brownsche Bewegung, Warteschlangentheorie,...).
- Stochastische Simulation: Nachbildung komplexer stochastischer Situationen auf dem Rechner.
- Statistik: Parameterschätzung, Modellvalidierung, Auswertung stochastischer Simulationen,...

# Einsatzgebiete

- Glücksspiele: Faire Einsätze, Gewinnchancen, ...
- Finanzmärkte: Optionspreisbewertungen,...
- Physik, Chemie: Kinetik, Teilchenbewegungen, molekulare Reaktionen, Versuchsplanung,...
- Biologie, Medizin, Pharmazie: Populationsentwicklungen, Ausbreitung von Epidemien, Medikamententests, Therapiebewertungen,...
- Informatik: Analyse von Algorithmen (zufällige Daten), randomisierte Algorithmen, Monte-Carlo-Methoden, Leistungsbewertung von Servern und Rechnernetzen,...
- Industrie und Dienstleistung: Verkehrsinfrastruktur, Logistik, Lagerhaltung, Produktionsprozesse, Callcenter,...
- ...

# Beispiel: Gestaltung von Bahnfahrplänen

- Zufallseinflüsse verursachen Verspätungen einzelner Züge.
- Aufgrund des dicht befahrenen Schienennetzes und Anschlussmöglichkeiten pflanzen sich diese Verspätungen fort und verstärken sich teilweise.
- Aufgaben:
  - Quantifizierung der Quellverspätungen,
  - Quantifizierung der Verspätungsfortplanzung,
  - Gestaltung eines möglichst robusten Fahrplans.

# Beispiel: Auslegung von Service-Callcentern

- Kunden rufen zu zufälligen Zeitpunkten an, Gesprächsdauern sind ebenfalls zufällig.
- Sind keine Service-Mitarbeiter verfügbar, werden die Kunden unzufrieden.
- Betreiber sind an möglichst hoher Auslastung der Mitarbeiter interessiert.
- Aufgaben:
  - Quantifizierung der Zufallseinflüsse,
  - Bestimmen eines sinnvoll gewählten LoS (Level of Service), etwa: Mit welcher Wahrscheinlichkeit wird ein anrufender Kunde innerhalb einer Minute mit einem Mitarbeiter verbunden?
  - Festlegen einer Anzahl von Mitarbeitern, um dieses Ziel zu erreichen.

# Beispiel: Produktionsplanung

- Produktionsprozesse unterliegen vielfältigen stochastischen Einflüssen (Qualität der Grundwerkstoffe, Fehler durch Mensch und Maschinen).
- Dadurch: Zeiten in Produktionsabschnitten zufällig (Nachbearbeitung,...), Warteschlangen bauen sich auf.
- Notwendigkeit von Warteplatz, halbfertige Produkte sind totes Kapital, zeitliche Unregelmäßigkeiten der Produktion.
- Aufgaben:
  - Quantifizierung und Fortpflanzung der Zufallseinflüsse,
  - Gestaltung eines Produktionsprozesses, bei dem möglichst wenig Warteplatz benötigt wird; dieser muss von Beginn an eingeplant werden.

# Problem: Vermittlung stochastischer Resultate

- Häufig: Anwender wollen konkrete Vorhersagen für einzelne Ereignisse.
- Bei Zufallseinflüssen ist dies natürlich nicht möglich.
- Stochastische Resultate haben für Anwendungen nur bei häufiger Wiederholung oder langer Laufzeit Sinn.
- Weiteres Problem: Kontraintuitivität stochastischer Resultate.

# Kapitel 2: Diskrete Wahrscheinlichkeitsräume – Grundbegriffe

## 2.1 Ergebnisse und Ereignisse

# Einfache stochastische Modelle (1)

- Beispiel vom Chevalier de Méré: Münzwurfspiel wird abgebrochen, zwei Runden stehen noch aus.
- Spieler B gewinnt genau, wenn bei den letzten beiden Würfen Bild erscheint.
- Mögliche stochastische Modellierung: Zufallsexperiment mit zwei Ausgängen, A gewinnt oder B gewinnt, Wahrscheinlichkeiten $\frac{3}{4}$ bzw. $\frac{1}{4}$.
- Alternative: Vier Ausgänge (Zahl,Zahl), (Zahl,Bild), (Bild,Zahl), (Bild,Bild), Wahrscheinlichkeiten jeweils $\frac{1}{4}$.
- Bei Alternative: Die ersten drei Ausgänge sind günstig für A, der letzte für B.
- Welches ist die richtige Modellierung?

# Einfache stochastische Modelle (2)

- Wurf mit zwei fairen Würfeln, von Interesse ist die Augensumme.
- Modellierung 1: Zufallsexperiment mit 11 Ausgängen, den Zahlen $2, 3, \ldots, 12$.
- Modellierung 2: Zufallsexperiment mit 21 Ausgängen, den Paaren $(i, j)$ mit $1 \leq i \leq j \leq 6$ (die Würfelergebnisse der Größe nach aufsteigend geordnet).
- Modellierung 3: Zufallsexperiment mit 36 Ausgängen, den Paaren $(i, j) \in \{1, \ldots, 6\}^2$ (gedanklich: ein roter und ein grüner Würfel, die Augenzahl des roten Würfels wird zuerst aufgeführt).
- Bei Modellierung 2 und 3: Ereignis „7 gewürfelt“ beinhaltet mehrere Ausgänge des Zufallsexperiments.
- Wie sehen die Wahrscheinlichkeiten für die Ausgänge des Zufallsexperiments jeweils aus? Was ist die richtige Modellierung?

# Einfache stochastische Modelle (3)

- Alle obigen Modellierungen sind in Ordnung.
- In Hinblick auf die Berechnung von Wahrscheinlichkeiten bietet es allerdings Vorteile,
  - das Experiment möglichst detailliert zu modellieren,
  - und interessante Ereignisse aus mehreren möglichen Ausgängen des Zufallsexperiments zusammenzusetzen.
- In den Beispielen also sinnvoll: Ausgänge $\{\text{Zahl}, \text{Bild}\}^2$ bzw. $\{1, \ldots, 6\}^2$.
- Detaillierte Modelle können darüberhinaus für ähnliche Fragestellungen wiederverwendet werden.

# Ergebnisse und Ergebnismengen

**Definition 2.1.1**
Die möglichen Ausgänge eines Zufallsexperiments heißen *Ergebnisse* oder *Elementarereignisse*. Die Menge aller Ergebnisse heißt *Ergebnismenge* oder *Ergebnisraum*.

Besonders gebräuchlich ist die Bezeichnung $\Omega$ für Ergebnismengen.

- Wurf mit einem fairen Würfel: $\Omega = \{1, \ldots, 6\}$.
- Wurf mit zwei fairen Würfeln: $\Omega = \{1, \ldots, 6\}^2$.
- Anzahl wartender Kunden an einem Schalter: $\Omega = \mathbb{N}_0$.
- Wartezeit an einer Ampel: $\Omega = [0, \infty)$.

# Diskrete Mengen

Die Mächtigkeit der Ergebnismengen in den einfachen Beispielen unterscheidet sich wesentlich. Es wird sich herausstellen, dass Ergebnismengen mit höchstens abzählbar unendlich vielen Elementen einfacher zu behandeln sind.

**Definition 2.1.2**
In der Stochastik heißt eine Menge mit endlich vielen oder abzählbar unendlich vielen Elementen *diskret*.

*Achtung:* Der Begriff „diskret“ ist in der Mathematik nicht eindeutig.

# Ereignisse

- Wurf mit zwei fairen Würfeln, $\Omega = \{1, \ldots, 6\}^2$.
- „7 gewürfelt“ umfasst die Ergebnisse $(1,6), (2,5), \ldots, (6,1)$.
- Idee: Ereignisse sind *Mengen von Ergebnissen*.

**Definition 2.1.3**
Ist $\Omega \neq \emptyset$ eine diskrete Ergebnismenge, so heißen die Teilmengen von $\Omega$ *Ereignisse*. Die Menge aller Ereignisse, also die Potenzmenge, wird mit $2^\Omega$ bezeichnet.

# Ereignisse – einfache Beispiele

- Wurf mit einem fairen Würfel, $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Ereignis „gerade Augenzahl“: $\{2, 4, 6\}$.
- Wurf mit zwei fairen Würfeln, $\Omega = \{1, 2, \dots, 6\}^2$.
- Ereignis „gerade Augensumme“:

$$\{(i,j) \in \Omega : \ 2|i+j\}.$$

- Ereignis „Augensumme höchstens 5“:

$$\{(i,j) \in \Omega : \ i+j \leq 5\}.$$

- Ereignis „Pasch gewürfelt“:

$$\{(i,i) : \ i \in \{1, \dots, 6\}\}.$$

# Warum Mengenschreibweise?

- Wurf mit zwei fairen Würfeln, $\Omega = \{1, 2, \ldots, 6\}^2$.
- $A \subset \Omega$ sei Ereignis „gerade Augensumme“.
- $B \subset \Omega$ sei Ereignis „Augensumme höchstens 5“.
- Das Ereignis „gerade Augensumme $\leq 5$“ wird einfach durch $A \cap B$ dargestellt.

Die Mengenoperation $\cap$ hat also eine anschauliche Interpretation. Dies gilt auch für alle anderen Mengenoperationen.

# Interpretation der Mengenoperationen

Es sei $\Omega \neq \emptyset$ eine Ergebnismenge und $A, B, A_1, A_2, \ldots \subset \Omega$ seien Ereignisse. Dann sind alle üblichen Mengenoperationen interpretierbar:

| | |
|---|---|
| $A \cap B$ | Ereignisse $A$ und $B$ treten gleichzeitig ein. |
| $\bigcap_{n=1}^{\infty} A_n$ | Alle Ereignisse $A_1, A_2, \ldots$ treten gleichzeitig ein. |
| $A \cup B$ | Mindestens eines der Ereignisse $A$ oder $B$ tritt ein. |
| $\bigcup_{n=1}^{\infty} A_n$ | Mindestens eines der Ereignisse $A_1, A_2, \ldots$ tritt ein. |
| $A^C = \Omega \setminus A$ | Ereignis $A$ tritt nicht ein. |
| $A \setminus B$ | Ereignis $A$ tritt ein, $B$ aber nicht. |

*Achtung:* Symbolik für Komplementbildung in der Mathematik nicht einheitlich (häufig auch $\overline{A}$), hier immer $A^C$.

# Interpretation von Mengenbeziehungen

| | |
|---|---|
| $\Omega$ | Sicheres Ereignis, tritt immer ein. |
| $\emptyset$ | Unmögliches Ereignis, tritt nie ein. |
| $A \subset B$ | Tritt Ereignis $A$ ein, so auch $B$. |
| $A \cap B = \emptyset$ | Ereignisse $A$ und $B$ treten niemals gleichzeitig ein. |

**Definition 2.1.4**
Im Fall $A \cap B = \emptyset$ heißen die Ereignisse $A$ und $B$ auch *disjunkt* oder *unvereinbar*. Gilt $A_i \cap A_j = \emptyset$ für alle $i, j \in \mathbb{N}$ mit $j \neq i$, so wird $(A_n)_{n \in \mathbb{N}}$ als Folge *paarweise disjunkter* Ereignisse bezeichnet.

*Vereinbarung:* Wird in dieser Veranstaltung $A \subset B$ geschrieben, bedeutet dies nur $x \in A \Rightarrow x \in B$; die Mengen können auch gleich sein. Ist $A$ eine echte Teilmenge von $B$, und ist dies wichtig, so wird dies über die Symbolik $\subsetneq$ betont.

# Rechnen mit Ereignissen

Da Ereignisse Mengen sind, gelten alle gewöhnlichen Eigenschaften der Mengenoperationen.

- Kommutativität: $A \cap B = B \cap A$, $A \cup B = B \cup A$.
- Assoziativität: $A \cap (B \cap C) = (A \cap B) \cap C$, $A \cup (B \cup C) = (A \cup B) \cup C$.
- Distributivität: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
- Komplementbildung: $\overline{\overline{A}} = A$, $\overline{\Omega} = \emptyset$, $\overline{\emptyset} = \Omega$.
- Reziprozität: $\overline{A \cap B} = \overline{A} \cup \overline{B}$, $\overline{A \cup B} = \overline{A} \cap \overline{B}$.
- Regeln von de Morgan: $\overline{\bigcap_{n=1}^{\infty} A_n} = \bigcup_{n=1}^{\infty} \overline{A}_n$, $\overline{\bigcup_{n=1}^{\infty} A_n} = \bigcap_{n=1}^{\infty} \overline{A}_n$.
- Reziproke Mengeninklusion: $A \subset B \;\Leftrightarrow\; \overline{B} \subset \overline{A}$.

# Ereignisse als Mengen – einige Bemerkungen

- Logische Verknüpfungen und Mengenverknüpfungen hängen eng zusammen, etwa in

$$A \cap B = \{x : \; x \in A \wedge x \in B\}.$$

- Logische Verknüpfungen beziehen sich aber auf Aussagen, Mengenverknüpfungen auf Mengen. Nicht verwechseln!
- Der Umgang mit Ereignissen erfordert einen sorgfältigen Umgang mit Mengen. Im Zweifelsfall lieber eine Klammer zu viel setzen! Was bedeutet

$$A \cup B \cap C?$$

- Der Begriff „Elementarereignis“ ist in der Literatur nicht eindeutig.
  - Hier: Identifikation mit Ergebnissen $\omega \in \Omega$.
  - Gelegentlich auch: Identifikation mit einelementigen Teilmengen $\{\omega\}$ für $\omega \in \Omega$.

# Unendlich große Ergebnismengen

- In den einfachen Würfelbeispielen war die Ergebnismenge immer endlich.
- Auch die Ergebnismengen $\mathbb{N} = \{1, 2, 3, \ldots\}$ oder $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ kommen in vielen sinnvollen stochastischen Modellen vor:
  - Anzahl wartender Kunden an einem Bedienschalter,
  - Anzahl zerfallender Atomkerne in einer vorgegebenen Zeitspanne,
  - Anzahl ein Service-Callcenter anrufender Kunden in einer vorgegebenen Zeitspanne.
- Hat die Ergebnismenge abzählbar unendlich viele Elemente, so gibt es überabzählbar viele Ereignisse.
- Begründung: Die Potenzmenge kann bijektiv auf die Menge aller unendlich langen 0-1-Folgen abgebildet werden.

## Zusammenfassung 2.1

- Die möglichen Ausgänge eines Zufallsexperiments heißen Ergebnisse.
- Interessierende Ereignisse können aus mehreren Ergebnissen zusammengesetzt werden.
- Ereignisse sind daher Teilmengen der der Ergebnismenge.
- Vorteil: Benutzen des Mengen-Formalismus.

## 2.2 Wahrscheinlichkeitsmaße

# Relative Häufigkeiten

Historisch gesehen entstanden Wahrscheinlichkeiten als Abstraktion relativer Häufigkeiten.

**Definition 2.2.1**

Ein Zufallsexperiment wird $n$ mal wiederholt. Tritt das Ereignis $A$ dabei $H_n(A)$ mal auf, so ist die *relative Häufigkeit* durch

$$r_n(A) = \frac{H_n(A)}{n}$$

gegeben.

Für eine feste Anzahl von Wiederholungen ist $r_n(\cdot)$ eine Funktion

$$r_n : 2^{\Omega} \to \mathbb{R}.$$

# Eigenschaften relativer Häufigkeiten

Betrachte eine diskrete Ergebnismenge $\Omega \neq \emptyset$; $r_n(\cdot)$ bezeichne die relative Häufigkeit bei $n$ Wiederholungen des Zufallsexperiments. Einfache Beobachtungen sind:

- Die relativen Häufigkeiten haben Werte in $[0, 1]$.
- Es gilt $r_n(\emptyset) = 0$ und $r_n(\Omega) = 1$.
- Für beliebige Ereignisse $A$ gilt $r_n\left(A^C\right) = 1 - r_n(A)$.
- Für Ereignisse $A \subset B$ gilt $r_n(A) \leq r_n(B)$.
- Sind $A$ und $B$ disjunkte Ereignisse, so gilt $r_n(A \cup B) = r_n(A) + r_n(B)$.
- Für eine Folge $(A_i)_{i \in \mathbb{N}}$ paarweise disjunkter Ereignisse gilt

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

# Der Übergang $n \to \infty$

- Historisch lange als Erfahrungswert bekannt: Relative Häufigkeiten stabilisieren sich, d.h. $r_n(A)$ konvergiert für $n \to \infty$.
- Dieser Grenzwert soll die Wahrscheinlichkeit eines Ereignisses werden, $P(A) = \lim_{n \to \infty} r_n(A)$.
- Dies als Definition zu verwenden, führt auf keine befriedigende Theorie.
- Kolmogorovs Idee: Eigenschaften relativer Häufigkeiten übertragen sich auf die Wahrscheinlichkeiten.
- Definiere *Wahrscheinlichkeitsmaße* axiomatisch anhand möglichst weniger solcher Eigenschaften.

# Die Kolmogorov-Axiome

**Definition 2.2.2**

Es sei $\Omega \neq \emptyset$ eine diskrete Ergebnismenge. Ist eine Funktion $P : 2^\Omega \to \mathbb{R}$

- nichtnegativ, d.h. $P(A) \geq 0$ für alle $A \subset \Omega$,
- normiert, d.h. $P(\Omega) = 1$,
- $\sigma$-additiv, d.h. gilt für jede Folge $(A_n)_{n\in\mathbb{N}}$ paarweise disjunkter Ereignisse
$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n),$$

so heißt $P$ *diskretes Wahrscheinlichkeitsmaß* auf $2^\Omega$ (auch: auf $\Omega$). Das Tupel $(\Omega, 2^\Omega, P)$ oder kurz $(\Omega, P)$ heißt dann *diskreter Wahrscheinlichkeitsraum*.

# Einfache Folgerungen aus den Axiomen

Kolmogorovs Axiome für Wahrscheinlichkeitsmaße umfassen nicht alle oben festgestellten Eigenschaften relativer Häufigkeiten. Nicht berücksichtigte Eigenschaften folgen allerdings direkt aus den Axiomen.

**Satz 2.2.1**

*Es sei $(\Omega, P)$ ein diskreter Wahrscheinlichkeitsraum und es seien $A, B \subset \Omega$. Dann gilt:*

(a) $P(\emptyset) = 0$.

(b) $P(A \cup B) = P(A) + P(B)$ *für disjunkte* $A, B$ *(Additivität).*

(c) $P\left(A^C\right) = 1 - P(A)$.

(d) $P(A) \leq P(B)$ *für* $A \subset B$ *(Monotonie).*

(e) $P(A) \leq 1$.

# Beweis Satz 2.2.1

(a) Für beliebiges $A$ sind die Mengen $A, \emptyset, \emptyset, \ldots$ paarweise disjunkt, $\sigma$-Additivität liefert

$$P(A) = P(A \cup \emptyset \cup \emptyset \cup \ldots) = P(A) + P(\emptyset) + P(\emptyset) + \ldots,$$

also $P(\emptyset) = 0$.

(b) Sind $A$ und $B$ disjunkt, so sind $A, B, \emptyset, \emptyset, \ldots$ paarweise disjunkt, $\sigma$-Additivität liefert

$$\begin{aligned} P(A \cup B) &= P(A \cup B \cup \emptyset \cup \emptyset \cup \ldots) \\ &= P(A) + P(B) + P(\emptyset) + P(\emptyset) + \ldots = P(A) + P(B). \end{aligned}$$

# Beweis Satz 2.2.1 (Fortsetzung)

(c) $A$ und $A^C$ sind disjunkt mit $A \cup A^C = \Omega$. Additivität liefert also

$$P(A) + P\left(A^C\right) = P\left(A \cup A^C\right) = P(\Omega) = 1.$$

(d) Für $A \subset B$ gilt $B = A \cup (B \setminus A)$, wobei $A$ und $B \setminus A$ disjunkt sind. Additivität und Nichtnegativität liefern

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

(e) $P(A) \leq 1$ folgt aus $A \subset \Omega$, $P(\Omega) = 1$ und der Monotonie.

# Ein einfaches Beispiel

- Wurf mit zwei fairen Würfeln, $\Omega = \{1, \ldots, 6\}^2$.
- Sinnvolle Annahme: $P(\{(i,j)\}) = \frac{1}{36}$ für alle $(i,j) \in \Omega$.
- $A$ und $B$ seien „Augensumme 6“ bzw. „Augensumme 7“, also

$$A = \{(1,5), (2,4), \ldots, (5,1)\}, \quad B = \{(1,6), (2,5), \ldots, (6,1)\}.$$

- Additivität liefert $P(A) = \frac{5}{36}$ und $P(B) = \frac{6}{36} = \frac{1}{6}$.
- Da $A$ und $B$ disjunkt sind, gilt auch $P(A \cup B) = P(A) + P(B) = \frac{11}{36}$.

# Nicht-disjunkte Vereinigungen

- Ersetze $B$ nun durch „Pasch gewürfelt“, also

$$B = \{(1,1), \ldots, (6,6)\}.$$

- $A$ bleibt unverändert „6 gewürfelt“.
- Dann gilt $P(B) = \frac{6}{36}$ und natürlich immer noch $P(A) = \frac{5}{36}$.
- Um $P(A \cup B)$ zu bestimmen, muss zunächst die komplette Menge $A \cup B$ bestimmt werden. Mit etwas Aufwand: Diese enthält 10 Ergebnisse.
- Es gilt also

$$P(A \cup B) = \frac{10}{36} \neq \frac{5}{36} + \frac{6}{36} = P(A) + P(B).$$

- Woran liegt das?

# Nicht-disjunkte Vereinigungen

Der Grund ist das „doppelte Zählen" des Ergebnisses $(3, 3)$, das sowohl in $A$ als auch $B$ liegt. Für nicht-disjunkte Vereinigungen ist das Berechnen von Wahrscheinlichkeiten etwas komplizierter.

**Satz 2.2.2 (Additionssatz)**

*Es sei $(\Omega, P)$ ein diskreter Wahrscheinlichkeitsraum, $A, B \subset \Omega$. Dann gilt*

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

# Beweis Satz 2.2.2

Schreibe $A = (A \setminus B) \cup (A \cap B)$ und $B = (B \setminus A) \cup (A \cap B)$ sowie $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$, in jedem Fall handelt es sich um disjunkte Vereinigungen. Additivität liefert

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

# Anwendung auf das Würfelbeispiel

- Im Würfelbeispiel mit den Ereignissen $A$ „6 gewürfelt" und $B$ „Pasch gewürfelt" war $P(A) = \frac{5}{36}$ und $P(B) = \frac{6}{36}$.
- Es gilt $A \cap B = \{(3,3)\}$, also $P(A \cap B) = \frac{1}{36}$. Der Additionssatz liefert also

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{10}{36}.$$

- Die Bestimmung von $A \cap B$ ist im Beispiel deutlich einfacher als die von $A \cup B$.
- Auch sonst wird häufig die Bestimmung der Wahrscheinlichkeit für den Schnitt zweier Ereignisse deutlich einfacher sein als für ihre Vereinigung.

# Die Siebformel von Sylvester-Poincaré

Als Verallgemeinerung des Additionssatzes gilt:

**Satz 2.2.3 (Siebformel von Sylvester-Poincaré)**

*Es sei $(\Omega, P)$ ein diskreter Wahrscheinlichkeitsraum, $A_1, \ldots, A_n \subset \Omega$. Dann gilt*

$$\begin{aligned} P\left(\bigcup_{k=1}^{n} A_k\right) &= \sum_{k=1}^{n} P(A_k) - \sum_{\substack{k_1,k_2=1\\k_1<k_2}}^{n} P\left(A_{k_1} \cap A_{k_2}\right) \\ &\quad + \sum_{\substack{k_1,k_2,k_3=1\\k_1<k_2<k_3}}^{n} P\left(A_{k_1} \cap A_{k_2} \cap A_{k_3}\right) \\ &\quad + \cdots + (-1)^{n-1} \cdot P(A_1 \cap \cdots \cap A_n). \end{aligned}$$

# Die Siebformel für drei Ereignisse

In den meisten Anwendungen wird die Siebformel nur auf zwei oder drei Ereignisse angewendet. Für zwei Ereignisse handelt es sich um den Additionssatz, für drei Ereignisse ergibt sich

**Folgerung 2.2.1**

*Es sei $(\Omega, P)$ ein diskreter Wahrscheinlichkeitsraum, $A, B, C \subset \Omega$. Dann gilt*

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

# Warum Ergebnisse und Ereignisse? (1)

- Im Würfelbeispiel: $P(\{\omega\}) = \frac{1}{36}$ für alle $\omega \in \Omega = \{1, \ldots, 6\}^2$.
- Für die Bestimmung der Wahrscheinlichkeit beliebiger Ereignisse wurde die Anzahl günstiger Ergebnisse gezählt und mit $\frac{1}{36}$ multipliziert.
- Kann $P(\{\omega\})$ verschiedene Werte annehmen, so reicht Zählen nicht aus. Die $\sigma$-Additivität liefert aber immer noch

$$P(A) = P\left(\bigcup_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in A} P(\{\omega\}).$$

- Somit lässt sich die Wahrscheinlichkeit beliebiger Ereignisse auf „Einzelwahrscheinlichkeiten“ zurückführen.
- Dies wird nun durch Wahrscheinlichkeitsvektoren formalisiert.

**Definition 2.2.3**
Es sei $\Omega \neq \emptyset$ eine diskrete Ergebnismenge. Ein Vektor $p = (p_\omega)_{\omega\in\Omega}$ mit

- $\forall \omega \in \Omega : \; p_\omega \geq 0$ (Nichtnegativität) und
- $\sum\limits_{\omega\in\Omega} p_\omega = 1$ (Normiertheit)

heißt *Wahrscheinlichkeitsvektor* oder *Zähldichte*.

Hat $\Omega$ (abzählbar) unendlich viele Elemente, so ist $p$ ein unendlich langer Vektor. Da $p_\omega$ nichtnegativ ist, ist die Reihe $\sum p_\omega$ stets absolut konvergent, insbesondere auch unbedingt konvergent.

**Satz 2.2.4**

*Es sei $\Omega \neq \emptyset$ eine diskrete Ergebnismenge.*

(a) *Ist $P : 2^\Omega \to \mathbb{R}$ ein Wahrscheinlichkeitsmaß, so wird durch $p_\omega = P(\{\omega\})$ ein Wahrscheinlichkeitsvektor $p = (p_\omega)_{\omega\in\Omega}$ definiert.*

(b) *Ist $p = (p_\omega)_{\omega\in\Omega}$ ein Wahrscheinlichkeitsvektor, so wird durch*

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

*ein Wahrscheinlichkeitsmaß definiert.*

# Beweis von Satz 2.2.4

(a) Da $P$ ein Wahrscheinlichkeitsmaß ist, gilt insbesondere $p_\omega = P(\{\omega\}) \geq 0$, außerdem ist

$$\sum_{\omega \in \Omega} p_\omega = \sum_{\omega \in \Omega} P(\{\omega\}) = P\left(\bigcup_{\omega \in \Omega} \{\omega\}\right) = P(\Omega) = 1.$$

(b) Es müssen die drei Eigenschaften von Wahrscheinlichkeitsmaßen nachgeprüft werden:

- Nichtnegativität: $P(A) = \sum_{\omega \in A} p_\omega \geq 0$.
- Normiertheit: $P(\Omega) = \sum_{\omega \in \Omega} p_\omega = 1$.
- $\sigma$-Additivität: Es seien $A_1, A_2, \ldots$ paarweise disjunkt. Unbedingte Konvergenz liefert

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{\omega \in \bigcup_{n=1}^{\infty} A_n} p_\omega = \sum_{n=1}^{\infty} \sum_{\omega \in A_n} p_\omega = \sum_{n=1}^{\infty} P(A_n).$$

# Warum Ergebnisse und Ereignisse? (2)

- Im Würfelbeispiel: $P(\{\omega\}) = \frac{1}{36}$ für alle $\omega \in \Omega = \{1, \dots, 6\}^2$.
- Diese Annahme erscheint physikalisch sinnvoll.
- Noch einfacher: Wurf eines fairen Würfels, $\Omega = \{1, \dots, 6\}$, physikalisch plausible Annahme $P(\{\omega\}) = \frac{1}{6}$ für alle $\omega = 1, \dots, 6$.
- Ähnlich: Wurf einer fairen Münze, $\Omega = \{\text{Zahl}, \text{Bild}\}$, physikalisch plausibel: $P(\{\omega\}) = \frac{1}{2}$ für alle $\omega \in \Omega$.
- Darauf aufbauend: $n$-facher Wurf einer fairen Münze, $\Omega = \{\text{Zahl}, \text{Bild}\}^n$, $P(\{\omega\}) = \frac{1}{2^n}$ für alle $\omega \in \Omega$.
- $n$-facher Wurf eines fairen Würfels, $\Omega = \{1, \dots, 6\}^n$, $P(\{\omega\}) = \frac{1}{6^n}$.

In allen Beispielen ist der Detailgrad der Ergebnisse so hoch, dass es einfach ist, für $P(\{\omega\})$ sinnvolle Annahmen aufzuschreiben. Kompliziertere Ereignisse können dann zusammengesetzt werden.

# Die Laplace-Annahme

In vielen Fällen erscheint es pausibel, dass alle Ergebnisse eines Zufallsexperimentes gleichwahrscheinlich sind.

**Definition 2.2.4**

Es sei $\Omega \neq \emptyset$ die Ergebnismenge eines Zufallsexperiments mit $|\Omega| < \infty$. Die Annahme $p_\omega = P(\{\omega\}) = \frac{1}{|\Omega|}$ heißt *Laplace-Annahme*. Das entstehende Wahrscheinlichkeitsmaß

$$P : 2^\Omega \to \mathbb{R},\; P(A) = \frac{|A|}{|\Omega|}$$

heißt *diskrete Gleichverteilung* oder *Laplacesche Wahrscheinlichkeitsverteilung*. Der zugehörige Wahrscheinlichkeitsraum $(\Omega, P)$ heißt *Laplacescher Wahrscheinlichkeitsraum*.

## Zusammenfassung 2.2

- Kolmogorov-Axiome für Wahrscheinlichkeiten sind Eigenschaften relativer Häufigkeiten nachempfunden.
- Berechnung von Wahrscheinlichkeiten „komplizierter“ Ereignisse lässt sich auf Wahrscheinlichkeiten „einfacher“ Ereignisse zurückführen.
- Insbesondere: Zusammensetzung von Wahrscheinlichkeiten beliebiger Ereignisse aus Wahrscheinlichkeitsvektoren.
- In vielen Fällen lässt sich die Ergebnismenge so wählen, dass eine Laplace-Annahme gerechtfertigt ist. Dann lassen sich Wahrscheinlichkeiten besonders einfach bestimmen (Zählen).

## 2.3 Diskrete Zufallsvariablen

# Einfache Beispiele

- Betrachte wieder Wurf mit zwei fairen Würfeln.
- Gute Wahl $\Omega = \{1, \dots, 6\}^2$, da dann Laplace-Annahme sinnvoll ist.
- Häufig aber nur Augensumme von Interesse, daher läge $\Omega' = \{2, \dots, 12\}$ als Ergebnismenge näher.
- Betrachte $n$-maliges Werfen einer fairen Münze.
- Laplace-Annahme für $\Omega = \{0, 1\}^n$, wobei 1 für „Zahl geworfen" und 0 für „Bild geworfen" stehen möge.
- Häufig nur Anzahl „Zahl gefallen" von Interesse, daher naheliegender: $\Omega' = \{0, 1, \dots, n\}$.

# Weitere Beispiele

- Beispiel Produktionslinie: Detaillierte Modellierung aller Produktionsstationen mit Zufallseinflüssen führt auf große, hochdimensionale Ergebnismenge.
- Häufig nur von Interesse: Gesamtdurchlaufzeit eines Produktes (Wert aus $[0, \infty)$).
- Beispiel Verspätungsfortpflanzung in Bahnnetzen: Detaillierte Modellierung aller Linien, Bahnhöfe, usw. führt auf große, hochdimensionale Ergebnismenge.
- Häufig nur von Interesse: Anzahl von Reisenden, die einen Zug verpassen oder mehr als eine Stunde später als geplant ihr Ziel erreichen (zwei Werte aus $\mathbb{N}_0$).

# Der Begriff Merkmal

- In diesen Situationen wird der Begriff *Merkmal* verwendet.
- In diesem Sinne ist ein Merkmal eine beobachtbare Größe in einem Zufallsexperiment.
- Durch Betrachtung eines einzelnen Merkmals oder weniger Merkmale anstatt des gesamten Zufallsexperimentes können Details verloren gehen.
- Wird nur das Merkmal „Augensumme“ beim Würfeln zweier Würfel beobachtet, sind zumeist keine Rückschlüsse auf die einzelnen Würfelergebnisse möglich.

# Merkmale als Funktionen

- Betrachte beim Wurf zweier fairer Würfel das Merkmal „Augensumme“ und die möglichen Ergebnismengen $\Omega = \{1, \ldots, 6\}^2$ und $\Omega' = \{2, \ldots, 12\}$.
- Beobachtung: Die Augensumme ist eine Abbildung $\Omega \ni (i, j) \mapsto i + j \in \Omega'$.
- $n$-faches Werfen einer fairen Münze, Merkmal „Anzahl Zahl“, mögliche Ergebnismengen $\Omega = \{0, 1\}^n$ und $\Omega' = \{0, \ldots, n\}$.
- Beobachtung: Die Anzahl geworfener Zahlen ist eine Abbildung $\Omega \ni (a_1, \ldots, a_n) \mapsto a_1 + \ldots + a_n \in \Omega'$.
- Dies gilt auch für die anderen Beispiele (mit komplizierteren Funktionen).

# Diskrete Zufallsvariablen

**Definition 2.3.1**
Es sei $(\Omega, P)$ ein diskreter Wahrscheinlichkeitsraum und $\Omega' \neq \emptyset$ eine diskrete Menge. Dann heißt jede Abbildung $X : \Omega \to \Omega'$ *Zufallsvariable*.

Zufallsvariablen formalisieren also die Reduktion eines Zufallsexperiments auf ein Merkmal. *Achtung:* Trotz des (historisch gewachsenen) Begriffes „Zufallsvariable“ handelt es sich um Funktionen.

# Wahrscheinlichkeiten im Bildbereich

- Bisher: Zufallsvariablen bilden Ergebnisse aus einer Ergebnismenge $\Omega$ auf Ergebnisse aus einer anderen Ergebnismenge $\Omega'$ ab.
- Gewünscht: Wahrscheinlichkeiten im Bildbereich.
- Etwa: Wie wahrscheinlich ist Augensumme $\in \{6, 7\}$ beim Wurf mit zwei fairen Würfeln?
- Hier einfache Idee: Welche Ergebnisse aus $\Omega$ führen auf die Ergebnisse $6, 7 \in \Omega'$?
- Dies sind $(1, 5), (2, 4), \dots, (5, 1), (1, 6), (2, 5), \dots, (6, 1)$.
- Das zugehörige Ereignis $\{(1, 5), \dots, (6, 1)\}$ hat im Originalraum $(\Omega, P)$ die Wahrscheinlichkeit $\frac{11}{36}$.
- Daher nimmt das Merkmal Augensumme mit Wahrscheinlichkeit $\frac{11}{36}$ einen der Werte aus $\{6, 7\}$ an.

# Formalisierung

- Formal ist ein Wahrscheinlichkeitsmaß $P'$ auf $2^{\Omega'}$ gesucht.
- Eben verwendet: $P'(\{6,7\}) = P(\{\omega \in \Omega :\ X(\omega) \in \{6,7\})$.
- Allgemein: $P'(A') = P(\{\omega \in \Omega :\ X(\omega) \in A'\}) = P\left(X^{-1}(A')\right)$ für Ereignisse $A' \subset \Omega'$.

**Erinnerung 2.3.1**

Es sei $f : D \to W$ eine Funktion und $B \subset W$. Dann heißt

$$f^{-1}(B) := \{x \in D :\ f(x) \in B\} \subset D$$

das *Urbild* von $B$ unter $f$. Für das Urbild gilt

$$f^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right) = \bigcup_{n=1}^{\infty} f^{-1}(B_n).$$

# Wirklich ein Wahrscheinlichkeitsmaß?

Es stellt sich noch die Frage, ob durch $P'(A') = P\left(X^{-1}(A')\right)$ tatsächlich ein Wahrscheinlichkeitsmaß definiert wird. Dazu werden die drei Axiome überprüft:

- Nichtnegativität ist klar.
- Normiertheit: $P'(\Omega') = P\left(X^{-1}(\Omega')\right) = P(\Omega) = 1.$
- $\sigma$-Additivität: Sind $A'_1, A'_2, \ldots \subset \Omega'$ paarweise disjunkt, so auch $X^{-1}(A'_1), X^{-1}(A'_2), \ldots \subset \Omega$. Es folgt

$$P'\left(\bigcup_{n=1}^{\infty} A'_n\right) = P\left(X^{-1}\left(\bigcup_{n=1}^{\infty} A'_n\right)\right) = P\left(\bigcup_{n=1}^{\infty} X^{-1}(A'_n)\right)$$

$$= \sum_{n=1}^{\infty} P\left(X^{-1}(A'_n)\right) = \sum_{n=1}^{\infty} P'(A'_n).$$

# Bildmaß und Verteilung

**Satz 2.3.1**
*Es sei* $(\Omega, P)$ *ein Wahrscheinlichkeitsraum,* $\Omega' \neq \emptyset$ *diskret,* $X : \Omega \to \Omega'$ *eine Zufallsvariable und* $P_X : 2^{\Omega'} \to \mathbb{R}$ *durch*

$$P_X(A') = P\left(X^{-1}(A')\right) = P(\{\omega \in \Omega : \ X(\omega) \in A'\}), \ A' \subset \Omega'$$

*definiert. Dann ist* $P_X$ *ein Wahrscheinlichkeitsmaß auf* $2^{\Omega'}$*.*

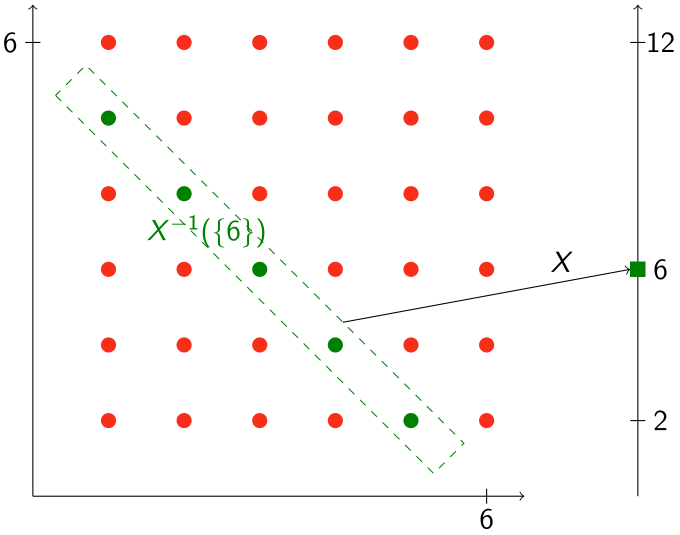
**Definition 2.3.2**
$(\Omega, P)$ sei ein diskreter Wahrscheinlichkeitsraum, $\Omega' \neq \emptyset$ eine diskrete Menge und $X : \Omega \to \Omega'$ eine Zufallsvariable. Dann heißt das in Satz 2.3.1 definierte Wahrscheinlichkeitsmaß $P_X$ auf $2^{\Omega'}$ *Bildmaß* von $P$ unter $X$ oder *Verteilung* von $X$ unter $P$.

# Verteilung der Augensumme

- Betrachte wieder die Augensumme beim Wurf zweier fairer Würfel.
- Zufallsvariable $X : \Omega \to \Omega'$ mit $\Omega = \{1, \dots, 6\}^2$ und $\Omega' = \{2, \dots, 12\}$.
- Laplace-Annahme für $P$. $P_X$ ist durch Angabe von $P_X(\{x\})$ für $x = 2, 3, \dots, 12$ eindeutig charakterisiert.
- Abzählen der Möglichkeiten führt auf

| $x$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $P_X(\{x\})$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $\frac{3}{36}$ | $\frac{2}{36}$ | $\frac{1}{36}$ |

# Urbild der Augensumme 6



# Bildmaß oder Verteilung?

- Die Begriffe Bildmaß und Verteilung wurden als (fast) synonym eingeführt.
- Der Unterschied lag in „…von $P$ unter $X$“ gegenüber „…von $X$ unter $P$“.
- Der Begriff Bildmaß betont, dass ein Wahrscheinlichkeitsmaß in ein anderes überführt wird.
- Der Begriff Verteilung stellt das zufällige Verhalten der Werte im Bildbereich, also der Werte des durch die Zufallsvariable beschriebenen Merkmals in den Vordergrund.
- Dabei tritt das Grundexperiment in den Hintergrund, es liefert „nur noch“ Möglichkeiten zur Berechnung der Wahrscheinlichkeiten.

# Übliche Notationen (1)

- Die Interpretation „$X$ ist der Wert eines Merkmals“ für eine Zufallsvariable legt Schreibweisen wie $P(X \in A')$ für „Das durch $X$ beschriebene Merkmal nimmt einen Wert aus $A'$ an“ nahe.
- Mathematisch gesehen: $X$ ist Funktion mit Werten in $\Omega'$, kein Element von $A'$.
- Aus historischen Gründen: Schreibweise hat sich durchgesetzt, sie ist als

$$P(X \in A') := P_X(A') = P\left(X^{-1}(A')\right) = P(\{\omega : \ X(\omega) \in A'\})$$

zu verstehen.

# Übliche Notationen (2)

Entsprechend auch

$$
\begin{aligned}
P(X = x) &= P_X(\{x\}) = P\left(X^{-1}(\{x\})\right) = P(\{\omega : X(\omega) = x\}),\\
P(X \leq x) &= P(\{\omega : X(\omega) \leq x\}),\\
P(X < x) &= P(\{\omega : X(\omega) < x\}),\\
&\vdots
\end{aligned}
$$

Im Würfelbeispiel wird daher auch geschrieben:

| $x$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $P(X = x)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $\frac{3}{36}$ | $\frac{2}{36}$ | $\frac{1}{36}$ |

# Ein weiteres Beispiel

- Eine faire Münze wird dreimal geworfen, $\Omega = \{0,1\}^3$, $P(\omega) = \frac{1}{8}$ für alle $\omega \in \Omega$ (Laplace-Annahme).
- Betrachtetes Merkmal: Anzahl Zahl (Anzahl Einsen).
- Zufallsvariable: $X : \Omega \to \Omega' = \{0,1,2,3\}$ mit $X((a_1,a_2,a_3)) = a_1 + a_2 + a_3$.
- Es ergeben sich die Urbilder

$$
\begin{aligned}
X^{-1}(\{0\}) &= \{(0,0,0)\}, \\
X^{-1}(\{1\}) &= \{(0,0,1),(0,1,0),(1,0,0)\}, \\
X^{-1}(\{2\}) &= \{(0,1,1),(1,0,1),(1,1,0)\}, \\
X^{-1}(\{3\}) &= \{(1,1,1)\}.
\end{aligned}
$$

- Damit ergibt sich

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ |

# Zufallsvariablen – einige Bemerkungen

- Bisher vorgestellt: Einfache Verteilungsannahme $P$ auf Grundraum $\Omega$, Zufallsvariable $X$ liefert neue Verteilung im Bildbereich.
- In vielen praktischen Situationen: Grundraum $\Omega$ und Wahrscheinlichkeitsmaß $P$ unbekannt, Annahmen über Verteilung einer Zufallsvariablen $X$ aufgrund von Erfahrungswerten.
- Dafür nötig: Grundwissen, welche Verteilung in welcher Situation plausibel ist.
- Laplace-Annahmen beruhen auf „physikalischer Austauschbarkeit“ der Ergebnisse (Indifferenzprinzip).
- Andere Verteilungsannahmen beruhen eher auf mathematischen Überlegungen.
- Grundgedanken dazu im Laufe dieses Semesters.

## Zusammenfassung 2.3

- Merkmale sind interessierende Größen in Zufallsexperimenten.
- Formal werden Merkmale durch Abbildungen zwischen Ergebnisräumen beschrieben, den Zufallsvariablen.
- Durch das zugrunde liegende Wahrscheinlichkeitsmaß und Zufallsvariablen wird ein Wahrscheinlichkeitsmaß im Bildbereich induziert.
- Dieses wird als Verteilung der Zufallsvariablen bezeichnet.