

Wahrscheinlichkeitstheorie anhand der Vorlesung  
von Dr. Arleta Szkola,  
Universität Leipzig

Phil Trommer

16. Januar 2019

# Inhaltsverzeichnis

<b>I</b>	<b>Einleitung</b>	<b>4</b>
<b>II</b>	<b>Thematischer Vorlesungsbeginn</b>	<b>6</b>
<b>1</b>	<b>Laplace - Modell</b>	<b>6</b>
1.1	Definition . . . . .	6
<b>2</b>	<b>Kombinatorik</b>	<b>8</b>
2.1	Definition Karthesisches Produkt . . . . .	9
2.2	Definition Endlichkeit . . . . .	10
2.3	Definition Mächtigkeit . . . . .	10
2.4	Lemma Mächtigkeit des karthesischen Produktes . . . . .	11
2.5	Satz Binomialkoeffizient . . . . .	11
2.6	Binomischer Lehrsatz . . . . .	11
<b>3</b>	<b>Diskrete Wahrscheinlichkeitsmodelle</b>	<b>12</b>
3.1	Definition Abzählbar & Unendlich . . . . .	12
3.2	Definition Diskreter W-Raum . . . . .	12
3.3	Definition Ereignisraum . . . . .	12
3.4	Satz (Begründung für die Interpretation) . . . . .	13
3.5	Beispiel Urne . . . . .	13
3.6	Definition Mengen in Funktionen . . . . .	14
<b>4</b>	<b>Zufallsvariablen</b>	<b>15</b>
4.1	Definition Zufallsvariable . . . . .	15
4.2	Theorem . . . . .	15
4.2.1	Beispiel Augensumme . . . . .	16
4.3	Definition Bernoulli-Verteilung . . . . .	16
4.4	Definition Bernoulli-Experiment . . . . .	17
4.5	Definition Indikatorfunktion . . . . .	17
4.5.1	Beispiel . . . . .	17
4.5.2	Definition . . . . .	18
4.6	Definition Gleichverteilte Zufallsvariablen . . . . .	18
4.7	Definition Binomialverteilung . . . . .	18
4.8	Definition . . . . .	19
4.8.1	Hypergeometrisch verteilte Zufallsvariablen . . . . .	19
4.9	Definition Zähldichte . . . . .	20
4.10	Definition Geometrische Verteilung . . . . .	20
4.11	Definition Poisson-Verteilung . . . . .	21

<b>5</b>	<b>Erwartungswerte und Varianz</b>	<b>22</b>
5.1	Definition Reelwertigen Zufallsvariable . . . . .	22
5.2	Definition Erwartungswert . . . . .	22
5.3	Theorem . . . . .	24
5.4	Definition . . . . .	24
5.5	Lemma . . . . .	25
5.6	Theorem . . . . .	25
5.7	Definition Varianz . . . . .	26
5.7.1	Endliche Zufallsvariable . . . . .	26
5.7.2	Abzählbare Zufallsvariable . . . . .	26
5.8	Definition Standardabweichung . . . . .	26
5.9	Definition Zweiter Moment . . . . .	27
5.10	Theorem . . . . .	28
5.10.1	Beispiele . . . . .	28
<b>6</b>	<b>Stochastische Unabhängigkeit</b>	<b>29</b>
6.1	Definition . . . . .	29
6.2	Definition Unabhängige Zufallsvariablen . . . . .	30
6.3	Lemma Unabhängigkeit . . . . .	31
6.3.1	Beispiel 1 . . . . .	31
6.3.2	Beispiel 2 . . . . .	31
6.4	Definition Familie von Ereignissen . . . . .	32
6.5	Definition Unabhängige Familie . . . . .	32
<b>7</b>	<b>Bedingte Wahrscheinlichkeit</b>	<b>32</b>
7.1	Definition Bedingte Wahrscheinlichkeit . . . . .	34
7.2	Satz von der totalen Wahrscheinlichkeit . . . . .	36
7.3	Satz Formel von Bayer . . . . .	36
7.4	Satz Existenz und Eindeutigkeit bedingter W-Maße . . . . .	38
<b>8</b>	<b>Gemeinsame Verteilung</b>	<b>40</b>
8.1	Definition Gemeinsame Verteilung . . . . .	41
8.2	Definition Produktmaß . . . . .	41
8.3	Satz Unabhängige Familie . . . . .	42
<b>9</b>	<b>Das schwache Gesetz der großen Zahlen (GGZ)</b>	<b>43</b>
9.1	Satz Bernoullis GGZ . . . . .	44
9.2	Lemma Rechenregeln für die Varianz . . . . .	44
9.3	Lemma Tschebyscheff-Menge . . . . .	45

# Teil I

# Einleitung

## Inhalte

- Wahrscheinlichkeitstheorie
- mathematische Statistik
- beide unter dem Begriff *Stochastik* zusammengefasst

## Literaturempfehlungen

- Hans-Otto-Georgii : “Stochastik”

## Ausgangsfrage:

Ist Zufall etwas fundamentales?

## Zentrale Frage:

Was ist die **WS** ( Wahrscheinlichkeit) eines zufälligen Ereignisses?  
Ursprünglich existierten folgenden 2 Definitionen.

### (1) Frequentistische Definition:

Die WS eines zufälligen Ereignisses ist der Grenzwert der relativen Häufigkeit des Eintretens dieses Ereignisses bei vielen Wiederholungen.

### (2) Bayes'sche Definition:

Die WS eines zufälligen Ereignisses ist ein Maß dafür, wie stark man vom Eintreten dieses Ereignisses überzeugt ist.

## Moderner Zugang

- Der moderne Zugang geht auf Alexander Kolmogorov zurück  
     $\hookrightarrow$  *Kolmogorov'schen Axiome*

- A1 Die WS eines Ereignisses ist eine reelle Zahl  $x$  mit  $0 \leq x \leq 1$
- A2 Das sichere Ereignis hat die WS  $= 1$
- A3 Eine abzählbare Vereinigung sich gegenseitig ausschließender Ereignisse hat die WS gleich der Summe der einzelnen WS.
  - Diese Axiome bilden die Grundlage der modernen Formulierung der **WT** (Wahrscheinlichkeitstheorie)  
“moderne Formulierung”  
 $\hookrightarrow$  basiert auf dem Konzept des *W-Raumes* (Wahrscheinlichkeitsraum)
- Sei  $\Omega$  (Omega) eine Menge (die Elemente heißen *Elementarereignisse*)
  - Sei  $\mathcal{A}$  ein geeignetes System von Teilmengen von  $\Omega$  ( diese Mengen heißen *Ereignisse*)
  - Und sei  $P: \mathcal{A} \rightarrow [0,1]$  eine Abbildung  
 $\hookrightarrow$  *W-Maß* (Wahrscheinlichkeitsmaß)
  - Die Abbildung  $P$  erfüllt dabei diese Rechenregeln:
    - (i)  $P(\Omega) = 1$
    - (ii)  $P(A^c) = 1 - P(A)$  , mit  $A^c = \Omega \setminus A$  und  $\forall A \in \mathcal{A}$
    - (iii)  $P\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} P(A_n)$ ,  $A_n \in \mathcal{A}$
- Damit heißt  $(\Omega, \mathcal{A}, P)$  ein W-Raum.

## Teil II

# Thematischer Vorlesungsbeginn

## 1 Laplace - Modell

↪ einfachstes Modell für ein Zufallsexperiment

Sei  $\Omega$  eine Menge

↪  $\Omega$  ist Ergebnismenge eines Zufallsexperimentes, wenn  $\Omega$  alle möglichen Ausgänge des Experiments erfasst

↪ jedes  $\omega \in \Omega$  ist ein mögliches Ergebnis/Ausgang welches *Elementarereignis* genannt wird.

Ist  $\Omega$  eine höchstens abzählbare Menge, dann heißt jede Teilmenge  $E$  von  $\Omega$  ein Ereignis. ( $\forall E \in \mathcal{P}(\Omega)$  ist ein Ereignis)

### 1.1 Definition

Ein Zufallsexperiment heißt *Laplace-Experiment* der Ordnung  $N \in \mathbb{N}$ , wenn die Ergebnismenge  $\Omega$  endlich ist mit  $|\Omega| = N$  und die Elementarereignisse alle gleichwahrscheinlich sind, d.h. die WS für das Eintreten des Elementarereignisses  $\omega \in \Omega$  ist  $\frac{1}{N}$ . (Formal:  $P(\omega) = \frac{1}{N}$ ,  $\forall \omega \in \Omega$ )

### Beispiel “Spielwürfel”

↪ Augenzahlen 1 - 6 entsprechen den möglichen Ausgängen des Experiments

↪  $\Omega = \{1,2,3,4,5,6\}$  Ergebnismenge

$N = |\Omega| = 6$  ist die Ordnung

Im Laplace-Modell ist die WS für das Eintreten eines Ereignisses  $E \in \mathcal{P}(\Omega)$  gegeben durch

$$P(E) = \frac{|E|}{N}$$

Denn gemäß den A1 - A3 gilt

$$P(E) =_{A3} \sum_{\omega \in E} P(\omega) = \sum_{\omega \in E} \frac{1}{N} = \frac{|E|}{N}$$

Beispiel

E: es fällt eine ungerade Augenzahl,  $E = \{1,3,5\} = \{1\} \cup \{3\} \cup \{5\}$

$$\hookrightarrow P(E) = P(\{1,3,5\}) = P(\{1\}) + P(\{3\}) + P(\{5\})$$

$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

Beispiel:

A ist ein Alphabet mit 5 Buchstaben, d.h.  $|A| = 5$

Frage:

Wie groß ist die WS, das ein zufällig gewähltes Wort der Länge 3 genau 2 verschiedene Buchstaben enthält?

Antwort:

$$\Omega = A \times A \times A$$

$$|\Omega| = |A^3| = 125$$

Ereignis E: Wort der Länge 3 enthält genau 2 verschiedene Buchstaben

$$|E| = ? \quad |E| = 5 * 4 * 3 = 60$$

$\hookrightarrow$  Möglichkeiten, um das doppelt vorkommende Symbol zu wählen

Daraus folgt

$$P(E) = \frac{60}{125}$$

Beispiel: "Wiederholtes Werfen eines Spielwürfels"

(fair, 6-seitig)

Der Würfel wird n -mal geworfen und jedes mal die Augenzahl notiert.

$\hookrightarrow$  Es handelt sich um ein Laplace-Experiment der Ordnung  $N = ?$

$$\hookrightarrow \Omega = \{1, \dots, 6\}^n = \{1, \dots, 6\} \times \{1, \dots, 6\} \times \dots$$

$$\hookrightarrow N = |\Omega| = 6^n$$

Für jedes  $\omega = (\omega_1, \dots, \omega_n) \in \Omega$  gilt  $P(\{\omega\}) = \frac{1}{6^n} = 6^{-n}$  Wir betrachten

- a)  $E_i$  : beim i-ten Wurf fällt die "6"
- b)  $\tilde{E}_i$  : nur beim i-ten Wurf fällt die 6
- c)  $E$  : die 6 fällt genau einmal

Zu a)

$$E_i = \{\omega = (\omega_1, \dots, \omega_n) \in \Omega | \omega_i = 6\}$$

$$\Rightarrow |E_i| = 6^{n-i} \Rightarrow P(E_i) = \frac{6^{n-1}}{6^n} = \frac{1}{6}$$

Zu b)

$$\tilde{E}_i = \{\omega = (\omega_1, \dots, \omega_n) \in \Omega | \omega_i = 6, \omega_k \neq 6, \forall k \neq i\}$$

$$\Rightarrow |\tilde{E}_i| = 5^{n-i} * 1 \Rightarrow P(\tilde{E}_i) = \frac{5^{n-1}}{6^n} = \frac{1}{6} * \left(\frac{5}{6}\right)^{n-1}$$

Zu c)

$$E = \bigcup_{i=1}^n \tilde{E}_i = \tilde{E}_1 \cup \tilde{E}_2 \cup \dots \cup \tilde{E}_n$$

$$\Rightarrow |E| = \sum_{i=1}^n |\tilde{E}_i| = n * 5^{n-1} \Rightarrow P(E) = \frac{n}{6} * \left(\frac{5}{6}\right)^{n-1}$$

## 2 Kombinatorik

Unser Kontext: Laplace-Experiment

$\hookrightarrow$  Formel für WS eines Ereignisses

$$E \subset \Omega \quad P(E) = \frac{|E|}{|\Omega|}$$

Das Urnenmodell:

- Urne enthält endlich viele gleichartige (in Größe, Gewicht, etc.) Kugeln
- Die Urne ist formal eine Menge  
 $\hookrightarrow$  wir denken uns die Kugeln durchnummeriert
- Die Kugeln werden nacheinander *blind* der Urne entnommen

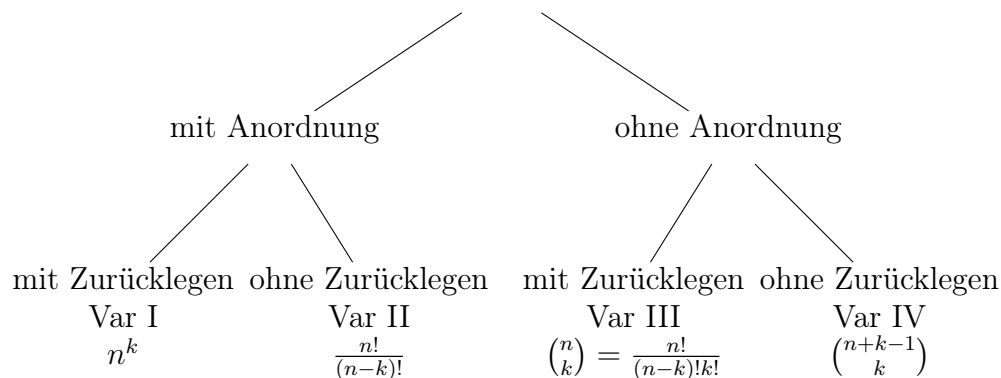
Man unterscheidet die Varianten:

- Wiederholtes Ziehen mit Zurücklegen
- Wiederholtes Ziehen ohne Zurücklegen

Sowie

- mit Anordnung (unter Beachtung der Reihenfolge der Ausgänge der Einzelbeziehungen)
- ohne Anordnung

$\hookrightarrow$  insgesamt 4 Varianten





Es gilt  $|U| = n$ ,  $k$  ist stetig

**Zu Var I :**

“Anzahl der 7 stelligen Telefonnummern ( $10^7$ )”

**Zu Var II :**

“Platzierung der ersten 3 Gewinner eines Wettbewerbes mit  $n$  Teilnehmern”

$$= n * (n - 1)(n - 2)$$

**Zu Var IV :**

“Wahlergebnisse einer Wahl in Parteien &  $k$  Wähler (mit je einer Stimme)”

$$\Omega = \{(a_1, \dots, a_k), a_i \in \bigcup, \forall i = 1, \dots, k$$

$$\text{und } a_1 \leq a_2 \leq a_3 \leq \dots \leq a_k\}$$

$\hookrightarrow$  hiermit wird die Vorgabe “ohne Anordnung” formuliert

Betrachte auf  $\Omega$  die Abbildung  $f$ , die durch folgende Vorschrift definiert ist.

$$(a_1, \dots, a_k) \rightarrow (a_1 + 0, a_2 + 1, a_3 + 2, \dots, a_k + k - 1)$$

Die Abbildung  $f$  bildet  $\Omega$  bijektiv auf  $B = f(\Omega)$  ab, wobei

$$B = \left\{ (b_1, \dots, b_k) \mid b_1 < b_2 < \dots < b_k, \quad b_i \in \bigcup_{n+k-1}, \forall i = 1, \dots, k \right\}$$

Damit ist  $|\Omega| = |B|$  ( $\Omega$  und  $B$  sind gleichmächtig).  $B$  ist aber die Ergebnismenge des Zufallsexperimentes der **Var III** im Urnenmodell.

**Wir beobachten:** Die Ergebnismengen von Zufallsexperimenten die jeweils eine der Varianten im Urnenmodell entsprechen, lassen sich in der Form eines karthesischen Produktes von (endlichen) Mengen beziehungsweise einer Teilmenge davon, darstellen.

## 2.1 Definition Karthesisches Produkt

Seien  $A, B$  zwei beliebige Mengen. Dann heißt

$$A \times B := \{(a, b) : a \in A, b \in B\}$$

karthesisches Produkt von  $A$  und  $B$ . Ist  $k > 2$  und  $A_1, \dots, A_k$  Mengen so definieren wir entsprechend

$$A_1 \times \dots \times A_k := \{(a_1, \dots, a_k) : a_i \in A_i, i = 1, \dots, k\}$$

Sind  $A_i = A, i = 1, \dots, k$ , so schreiben wir kurz  $A^k$  für  $A_1 \times \dots \times A_k$  Bemerkung:

$(a, b) \in A \times B$  ist ein *geordnetes Paar*.

$$\hookrightarrow (a, b) \neq (b, a)$$

Wichtig ist die Unterscheidung zwischen  $(a, b)$  und  $\{a, b\}$  da hier  $\{a, b\} = \{b, a\}$  gilt.

## 2.2 Definition Endlichkeit

Eine Menge  $A$  die nur endlich viele Elemente enthält, heißt *endlich*. Sonst heißt sie *unendliche Menge*. Die Anzahl der Elemente einer endlichen Menge  $A$  wird *Mächtigkeit* von  $A$  genannt und mit  $|A|$  bezeichnet.

## 2.3 Definition Mächtigkeit

Zwei Mengen  $A, B$  heißen *gleichmächtig*, wenn eine bijektive Abbildung

$$f : A \rightarrow B$$

existiert.

**Beispiel:**  $\mathbb{N}$  und  $\mathbb{Z}$  sind gleichmächtig, da folgende Abbildung bijektiv ist

$$f : \mathbb{N} \rightarrow \mathbb{Z}, n \rightarrow f(n) := \begin{cases} \frac{n}{2}, & n \text{ gerade} \\ \frac{1-n}{2}, & n \text{ ungerade} \end{cases}$$

Explizite Darstellung der Ergebnismengen der Variante I-IV

- Var I:

$$\{(a_1, \dots, a_k) : a_i \in \{1, \dots, k\}, \forall i = 1, \dots, k\} = \{1, \dots, n\}^k =: \Omega_I$$

- Var II:

$$\{(a_1, \dots, a_k) \in \{1, \dots, n\}^k : a_i \neq a_j \forall i, j = 1, \dots, k \text{ mit } i \neq j\} =: \Omega_{II}$$

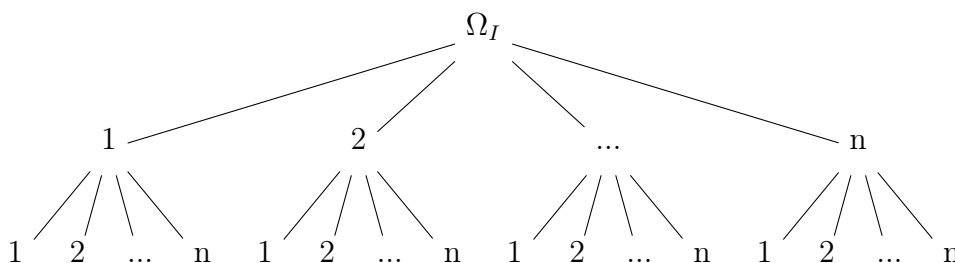
- Var III:

$$\{(a_1, \dots, a_k) \in \{1, \dots, n\}^k : a_1 < a_2 < \dots < a_k\} =: \Omega_{III}$$

- Var IV:

$$\{(a_1, \dots, a_k) \in \{1, \dots, n\}^k : a_1 \leq a_2 \leq \dots \leq a_k\} =: \Omega_{IV}$$

- Baumdiagramm:



- Jeder der Pfade stellt einen möglichen Ausgang des Zufallsexperimentes der Variante I im Urnenmodell dar.
- Für Varianten II - IV müssten bestimmte Pfade weggelassen bzw. miteinander identifiziert werden (z.B. kein doppeltes Vorkommen).

## 2.4 Lemma Mächtigkeit des kartesischen Produktes

Für zwei endliche Mengen  $A, B$  gilt

$$|A \times B| = |A| * |B|$$

## 2.5 Satz Binomialkoeffizient

$$\binom{n}{k} = \binom{n}{n-k}$$

Beweis:

- 1)  $\binom{n}{k}$  = Anzahl der  $k$ -elementigen Teilmengen von  $\{1, \dots, n\}$
- 2) Mit jeder  $k$ -elementigen Teilmenge wird auch ihr Komplement  $T^C$  festgelegt. Es handelt sich dann um eine  $(n-k)$ -elementige Teilmenge von  $\{1, \dots, n\}$

$$\hookrightarrow |\{T \subset \{1, \dots, n\} : |T| = k\}| = |\{T^C : T \subset \{1, \dots, n\} \text{ und } |T^C| = k\}|$$

1) & 2) zusammen ergeben die Aussage ■

## 2.6 Binomischer Lehrsatz

Für  $n \in \mathbb{N}$  und  $a, b \in \mathbb{R}$  gilt:

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k * b^{n-k}$$

Beweis:

$$(a+b)^n = (a+b)(a+b) * \dots (a+b)$$

Durch ausmultiplizieren erhalten wir eine Summe, bei der jeder Term ein Produkt von  $n$  Faktoren ist. Jeder der Faktoren kann nur die Werte "a" oder "b" annehmen.

$\hookrightarrow$  Die Terme sind von der Gestalt  $a^k b^{n-k}$ , wobei  $k = 0, \dots, n$ . Für jedes feste  $k$  gibt es genau  $\binom{n}{k}$  Möglichkeiten ■

### 3 Diskrete Wahrscheinlichkeitsmodelle

Als Verallgemeinerung des Laplace-Modells im folgendem Sinn:

- endliche oder abzählbar unendliche Ergebnismenge
- die Elementarereignisse (typischerweise) nicht mehr alle gleichwahrscheinlich

#### 3.1 Definition Abzählbar & Unendlich

Eine Menge  $\Omega$  heißt abzählbar unendlich, wenn sie mit  $\mathbb{N}$  gleichmächtig ist.

#### 3.2 Definition Diskreter W-Raum

Sei  $\Omega$  eine höchstens abzählbare nichtleere Menge und

$$\rho : \Omega \rightarrow [0, 1], \omega \mapsto \rho(\omega)$$

eine Funktion mit

$$\sum_{\omega \in \Omega} \rho(\omega) = 1$$

Dann heißt  $\rho$  eine (*Zähl*)-*Dichte*, W-Vektor oder Gewichtsfunktion auf  $\Omega$ . Die Abbildung

$$P : \mathcal{P}(\Omega) \rightarrow [0, 1], E \mapsto P(E) := \sum_{\omega \in E} \rho(\omega)$$

wird als W-Maß auf  $\mathcal{P}(\Omega)$  genannt.  $(\Omega, \mathcal{P}(\Omega), P)$  wird als *diskreter W-Raum* bezeichnet.

#### 3.3 Definition Ereignisraum

Sei  $\Omega$  eine höchstens abzählbare Ergebnismenge. Dann versteht man unter einem *Ereignisraum* das Paar  $(\Omega, \mathcal{P}(\Omega))$ .

**Frage:** Warum wird  $P$  als W-Maß bezeichnet?

- Der Wert  $P(E)$ , den  $P$  dem Ereignis  $E \in \mathcal{P}(\Omega)$  zuordnet, wird als Wahrscheinlichkeit des Ereignisses  $E$  interpretiert
- Speziell für Elementarereignisse  $\{\omega\} \in \mathcal{P}(\Omega)$  erhalten wir gemäß der Definition von  $P$  die Relation

$$P(\{\omega\}) = \rho(\omega)$$

### 3.4 Satz (Begründung für die Interpretation)

Ist  $\rho : \Omega \rightarrow [0, 1]$  eine Gewichtungsfunktion, d.h.

$$\sum_{\omega \in \Omega} \rho(\omega) = 1,$$

dann erfüllen die Funktionswerte  $P(E), E \in \mathcal{P}(\Omega)$  des zugehörigen W-Maßes

$$P : \mathcal{P}(\Omega) \rightarrow [0, 1], E \mapsto P(E) = \sum_{\omega \in E} \rho(\omega)$$

die *kolmogorowschen Axiome* A1 bis A3 (siehe Einleitung).

**Beweis:**

zu A1) zu zeigen ist  $0 \leq P(E) \leq 1, \forall E \in \mathcal{P}(\Omega)$

$$P(E) \stackrel{\text{def}}{=} \sum_{\omega \in E} \rho(\omega) \geq \sum_{\omega \in E} 0$$

$$\hookrightarrow \rho(\omega) \geq 0, \forall \omega \in \Omega$$

$$P(E) \stackrel{\text{def}}{=} \sum_{\omega \in E} \rho(\omega) \leq \sum_{\omega \in E \cup E^c} \rho(\omega) = \sum_{\omega \in \Omega} \rho(\omega) = 1$$

$$\hookrightarrow \text{da } \rho(\omega) \geq 0, \forall \omega \in \Omega$$

zu A2)

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

zu A3) Seien  $E_1, \dots, E_n$  paarweise disjunkte Ereignisse

$$P\left(\bigcup_{i=1}^n E_i\right) = \sum_{\omega \in \bigcup_{i=1}^n (E_i)} \rho(\omega) = \sum_{i=1}^n P(E_i)$$

### 3.5 Beispiel Urne

Sei  $U$  eine Urne mit  $n$  Kugeln, d.h.  $|U| = n$ . Die Kugeln in  $U$  besitzen ein weiteres Unterscheidungsmerkmal, hier die Farbe.

- Sei  $S \subset U$  die Teilmenge der schwarzen Kugeln
- Sei  $W \subset U$  die Teilmenge der weißen Kugeln
- Gelte  $S \cup W = U$

Betrachte das *Zufallsexperiment*  $\mathcal{E}$ :

Es wird eine Kugel aus  $M$  entnommen und die Farbe notiert.

$\hookrightarrow$  Ergebnismenge von  $\mathcal{E} : \Omega = \{s, w\}$

$\rightarrow$  **Frage:** Wie wahrscheinlich ist es, das ich eine schwarze Kugel ziehe?

**Antwort:**  $P(E) = \frac{|S|}{M} = \tilde{P}(S)$

**Begründung:** Dem Experiment  $\mathcal{E}$  liegt ein Laplace-Experiment  $\tilde{\mathcal{E}}$  zugrunde.

$\tilde{\mathcal{E}}$ : Es wird eine Kugel aus  $M$  gezogen mit dem zugehörigen  $W$ -Raum

$$(\tilde{\Omega} = U, \mathcal{P}(U), \tilde{P}) \quad \text{wobei} \quad \tilde{P}(E) = \frac{|E|}{n}$$

Betrachten wir jetzt die Abbildung

$$f(\tilde{\omega}) = \begin{cases} s & , \text{ falls } \tilde{\omega} \in S \\ w & , \text{ sonst } \hat{=} \tilde{\omega} \in W \end{cases}$$

Dann erhalten wir die Ergebnismenge  $\Omega$  von  $\mathcal{E}$  als Bildmenge von  $\tilde{\Omega} = U$  unter der Abbildung  $f$ , das heißt

$$\Omega = f(\tilde{\Omega})$$

und die Gewichtsfunktion:

$$\rho : \Omega \rightarrow [0, 1]$$

ergibt sich aus

$$\rho(\omega) = P(\{\omega\}) = \tilde{P}(f^{-1}(\{\omega\})) = \frac{|f^{-1}(\{\omega\})|}{n}$$

### 3.6 Definition Mengen in Funktionen

Seien  $A, B$  nichtleere Mengen und

$$f : A \rightarrow B$$

eine Abbildung mit dem Definitionsbereich  $A$  und Bildbereich  $B$ . Für jedes  $M \subset A$  heißt die Menge

$$f(M) = \{f(a) \in B : a \in M\} \subset B$$

Die *Urbildmenge* von  $S \subset B$  und  $f$  ist gegeben durch

$$f^{-1}(S) := \{a \in A : f(a) \in S\}$$

**Beispiel:**

$$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto e^x, \quad f^{-1}([0, 1]) = \mathbb{R}$$

## 4 Zufallsvariablen

### 4.1 Definition Zufallsvariable

- $(\Omega, \mathcal{P}(\Omega'), P)$  diskreter W-Raum
- $\Omega$  höchstens abzählbar  
 $\hookrightarrow X : \Omega' \rightarrow \Omega$  heißen Zufallsvariablen mit Werten in  $\Omega$  und Verteilung

$$P_X(E) := P(X^{-1}(E)), \forall E \subset \Omega$$

- Mit dem Konzept einer Zufallsvariable werden in der Wahrscheinlichkeitstheorie Zufallsexperimente beschrieben.
- Eine Zufallsvariable mit Werten in  $\Omega$  beschreibt ein Zufallsexperiment, dessen Ergebnismenge gleich  $\Omega$  oder einer Teilmenge von  $\Omega$  ist. Genauer gibt  $X(\Omega')$  die Ergebnismenge des Zufallsexperimentes an.
- Die Verteilung  $P_X$  der Zufallsvariable  $X$  ist formal ein W-Maß. Damit stellt das Tripel  $(\Omega, \mathcal{P}\Omega, P_X)$  einen diskreten W-Raum dar.

### 4.2 Theorem

Seien  $(\Omega', \mathcal{P}(\Omega'), P)$  ein diskreter W-Raum und  $X$  eine Abbildung auf  $\Omega'$ , dann ist durch

$$P_X : \mathcal{P}(\Omega) \rightarrow [0, 1]$$

$$E \mapsto P(X^{-1}(E))$$

wobei  $\Omega = X(\Omega')$  ein W-Maß auf  $\mathcal{P}(\Omega)$  definiert.

**Beweis:** Wir konstruieren eine Gewichtsfunktion

$$\rho : \Omega \rightarrow [0, 1]$$

so dass  $\forall E \in \mathcal{P}(\Omega)$  gilt:

$$P_X(E) = \sum_{\omega \in E} \rho(\omega)$$

**Unser Ansatz:**

$$\forall \omega \in \Omega$$

$$\rho(\omega) = P(X^{-1}(\{\omega\})) \quad \text{mit} \quad X^{-1}(\{\omega\}) \subset \Omega'$$

Da  $P$  nach Voraussetzung ein W-Maß ist, gilt dann  $\rho(\omega) \in [0, 1]$ .  
 Zum anderen erhalten wir

$$\sum_{\omega \in \Omega} \rho(\omega) \stackrel{\text{def}}{=} \sum_{\omega \in \Omega} P(X^{-1}(\{\omega\})) \stackrel{i}{=} P\left(\bigcup_{\omega \in \Omega} X^{-1}(\{\omega\})\right) = P(\Omega') = 1 \quad \blacksquare$$

#### 4.2.1 Beispiel Augensumme

Zufallsexperiment  $\mathcal{E}$ : Es werden 2 Spielwürfel nacheinander geworfen und die Augensumme notiert.

$\hookrightarrow$  **Frage:** Was ist das geeignete W-Modell für  $\mathcal{E}$ ?

**alternativ:** Wie sieht die Zufallsvariable  $X$ : “Augensumme von zwei 6-seitigen Würfeln” formal aus?

**Antwort:** Die Ergebnismenge von  $\mathcal{E}$  ist  $\Omega = \{2, \dots, 12\} \Rightarrow |\Omega| = 11$

**Beobachtung:** Das Laplace-Modell scheint ungeeignet, da es mit dem Zufallsexperiment  $\mathcal{E}$ : “Ein Spielwürfel wird 2 mal nacheinander geworfen und die Augenzahl notiert” nicht konsistent ist.

Wir formulieren eine geeignete Abbildung:

$$\begin{aligned} X : \Omega' &\rightarrow \Omega \quad \text{wobei} \quad \Omega' = \{1, \dots, 6\}^2, |\Omega'| = 36 \\ &\hookrightarrow (a, b) \rightarrow a + b \\ \Rightarrow \rho(\omega) &= P(\{\omega\}) := P'(X^{-1}(\{\omega\})) = \frac{|X^{-1}(\{\omega\})|}{36} \\ &= \frac{|\{(a, b) : a, b \in \{1, \dots, 6\}, a + b = \omega\}|}{36} \\ &\hookrightarrow \rho(\omega) = \frac{1}{36} * (6 - |\omega - 7|), \omega \in \{1, \dots, 12\} = \Omega \end{aligned}$$

Die Verteilung  $P_X$  ist dann das W-Maß zur Gewichtsfunktion  $\rho$ , die in der vorherigen Gleichung festgelegt ist. Es gibt mehrere wichtige Verteilungen.

#### 4.3 Definition Bernoulli-Verteilung

Sei  $p \in [0, 1]$ . Die Verteilung auf  $\mathcal{P}(\Omega)$ , wobei  $\Omega = \{0, 1\}$ , zur Gewichtsfunktion  $\rho(1) = p$  und  $\rho(0) = 1 - p$  heißt *Bernoulli-Verteilung zum Parameter  $p$*  und wird mit  $B_p$  bezeichnet.

---

<sup>i</sup>Siehe 3.4



## 4.4 Definition Bernoulli-Experiment

Ein Zufallsexperiment heißt *Bernoulli-Experiment*, wenn es durch eine Bernoulli -verteilte Zufallsvariable beschrieben wird.

Bernoulli-Experimente haben folgende Eigenschaften:

- Ergebnismenge =  $\{0, 1\}$
- Modellieren zum Beispiel das Werfen einer Münze, wobei “Zahl” und “Kopf” durch die Werte “0” bzw. “1” kodiert werden
- Allgemein interpretiert man

$$\begin{aligned} 1 &\hat{=} \text{Treffer/Erfolg} \\ 0 &\hat{=} \text{kein Treffer/Misserfolg} \end{aligned}$$

$\hookrightarrow B_p$  ist die Bernoulli-Verteilung mit Erfolgswahrscheinlichkeit  $p$

- Häufig steht “1” sinngemäß für “Ja” und “0” für “Nein”, wenn die betreffende Zufallsvariable das Eintreten bzw. Nicht-Eintreten eines Ereignisses beschreibt.

## 4.5 Definition Indikatorfunktion

Seien  $\Omega$  eine Menge und  $M \subseteq \Omega$ . Die auf  $\Omega$  definierte Funktion

$$1_M(\omega) := \begin{cases} 1 & , \omega \in M \\ 0 & , \text{sonst} \end{cases}$$

heißt *Indikatorfunktion* von  $M$  auf  $\Omega$ .

### 4.5.1 Beispiel

**Beispiel 1** Ein 6-seitiger Würfel wird geworfen und die Augenzahl “6” als Treffer gewertet.

$\hookrightarrow$  Zufallsvariable: Es wird ein Treffer erzielt ist Bernoulli-verteilt zum Parameter  $p = \frac{1}{6}$   
 $\hookrightarrow$  formal:

$$\begin{aligned} X : \{1, \dots, 6\} &\rightarrow \{0, 1\} \\ \omega &\rightarrow 1_{\{6\}}(\omega) \end{aligned}$$

**Beispiel 2** Ein 6-seitiger Würfel wird dreimal nacheinander geworfen und jeweils die Augenzahl notiert. Ab einer Augensumme =15 wird ein Gewinn ausgezahlt.

$\hookrightarrow$  Zufallsvariable  $Z$ : Gewinnauszahlung  
 $\hookrightarrow$  formall:

$$Z = 1_{\{X \geq 15\}} \circ X^{\text{ii}}$$

---

<sup>ii</sup>Zufallsvariable: Augensumme nach 3 Würfeln

somit  $Z : \{1, \dots, 6\}^3 \rightarrow \{0, 1\}$

#### 4.5.2 Definition

Seien  $f : A \rightarrow B$  und  $g : B \rightarrow C$  zwei Abbildungen. Dann ist die Verkettung  $g \circ f$  definiert als die Abbildung

$$g \circ f : A \rightarrow C, \quad a \mapsto g \circ f(a) := g(f(a))$$

#### 4.6 Definition Gleichverteilte Zufallsvariablen

Sei  $\Omega$  eine endliche Menge. Das W-Maß auf  $\Omega$  mit der (konstanten) Zähldichte

$$\rho(\omega) = \frac{1}{|\Omega|}, \quad \omega \in \Omega$$

heißt *Gleichverteilung auf  $\Omega$*  und wird mit  $\mathcal{U}_\Omega$  bezeichnet.

#### 4.7 Definition Binomialverteilung

Seien  $n \in \mathbb{N}$  und  $p \in (0, 1)$ . Das W-Maß auf  $\{0, 1, \dots, n\}$  mit der Zähldichte

$$\rho(k) := \binom{n}{k} p^k (1-p)^{n-k}, \quad \forall k = 0, 1, \dots, n$$

heißt *Binomialverteilung zu den Parametern  $n$  und  $p$* . (Notation =  $B_{n,p}$ )

**Beispiel:** Beim  $n$ -maligen Werfen eines 6-seitigen Würfels wird jedes Mal, wenn die Augenzahl “6” fällt, ein Treffer gezählt.

$\hookrightarrow$  Zu  $X$ : Anzahl der Treffer bei  $n$  Versuchen

$X$  ist *binomialverteilt*.

**Herleitung:** Das zugrunde liegende Zufallsexperiment ist

$\mathcal{E}'$ : ein fairer 6-seitiger Würfel wird  $n$ -mal geworfen und jeweils die Augenzahl notiert.  $\mathcal{E}'$  stellt ein Laplace-Experiment der Ordnung  $N = 6^n$  dar. Bezeichne  $\Omega'$  die Ergebnismenge, d.h.  $\Omega' = \{1, \dots, 6\}^n$ , dann ist  $(\Omega', \mathcal{P}(\Omega'), \mathcal{U}_{\Omega'})$  der zugehörige Laplace-Raum.

$$X : \Omega' \rightarrow \{0, 1, \dots, n\}$$

$$\omega = \{\omega_1, \dots, \omega_n\} \mapsto \sum_{i=1}^n 1_{\{6\}}(\omega_i)$$

Diese Abbildung stellt die betreffende Zufallsvariable  $X$  dar!

$$\begin{aligned}\hookrightarrow P_X(\{k\}) &= \mathcal{U}_{\Omega'}(X^{-1}(\{k\})) \\ &= \frac{|X^{-1}(\{k\})|}{6^n} \\ &= \frac{\binom{n}{k} * 5^{n-k}}{6^{n-k+k}} = \binom{n}{k} \left(\frac{1}{6}\right) * \left(\frac{5}{6}\right)^{n-k}\end{aligned}$$

## 4.8 Definition

Seien  $P$  ein W-Maß und  $X$  eine Zufallsvariable. Wir schreiben  $X \sim P$ , wenn die Verteilung  $P_X$  von  $X$  gleich  $P$  ist.

### 4.8.1 Hypergeometrisch verteilte Zufallsvariablen

**Beispiel** In einer Urne befinden sich  $N$  Kugeln, davon sind  $k$  rot und die anderen  $N - k$  Kugeln sind weiß. Es werden  $m$  Kugeln blind gezogen.

$\hookrightarrow$  Frage: Wie ist folgende Zufallsvariable  $X$  verteilt?

$X$ : Anzahl der roten Kugeln in der Stichprobe (vom Umfang  $m$ )

Wir stellen  $X$  als eine Abbildung auf dem Laplace-Raum  $(\Omega', \mathcal{P}(\Omega'), P)$  dar, der zum Laplace-Experiment  $\mathcal{E}'$  gehört.

$\hookrightarrow \mathcal{E}'$ : Es werden  $m$  Kugeln blind und gleichzeitig aus der Urne gezogen

$\hookrightarrow$  Ergebnismenge  $\Omega'$  zu  $\mathcal{E}'$ :

$$\begin{aligned}\Omega' &= \{X \subseteq \{1, \dots, N\} : |X| = m\} \\ &= \{(\omega_1, \dots, \omega_m) : \forall i = 1, \dots, m : \omega_i \in \{1, \dots, N\}, \\ &\quad \text{und } \omega_1 < \omega_2 < \dots < \omega_m\}\end{aligned}$$

$$\hookrightarrow |\Omega'| = \binom{N}{m}$$

Bezeichne jetzt  $K$  die Teilmenge der roten Kugeln. Die Zufallsvariable  $X$  wird als Abbildung auf  $(\Omega', \mathcal{P}(\Omega'), \mathcal{U}_{\Omega'})$  dargestellt.

$$X : \omega \in \Omega' \mapsto \sum_{i=1}^m 1_K(\omega_i)$$

$\hookrightarrow$  Verteilung  $P_X$  von  $X$  ergibt sich wie folgt:

$$\begin{aligned}P_X(\{l\}) &= \mathcal{U}_{\Omega'}(X^{-1}(\{l\})), \forall l = 0, \dots, \min\{m, k\} \\ &= \frac{\binom{k}{l} * \binom{N-k}{m-l}}{\binom{N}{m}}\end{aligned}$$

In folgenden Fällen ist  $X^{-1}(\{l\}) = \emptyset$ :

- $l > k$
- $N - k < m - l$

$$\hookrightarrow X(\Omega') = \{l * \max\{0, \min(N - k)\} \leq l \leq \min\{m, k\}\}$$

**Bemerkung:** Seien  $n \in \mathbb{N}$  und  $k \in \mathbb{Z}$ . Dann gilt

$$\binom{n}{k} = \begin{cases} 0 & , k < 0 \\ 0 & , k > n \\ \frac{n(n-1)\dots(n-k+1)}{k!} & , \text{sonst} \end{cases}$$

#### 4.9 Definition Zählerdichte

Seien  $N, m, k \in \mathbb{N}$  mit  $K < N$  und  $m \leq N$ . Das W-Maß mit der *Zählerdichte*

$$\rho : \{0, 1, \dots, m\} \rightarrow [0, 1], \quad l \mapsto \frac{\binom{k}{l} \binom{n-k}{m-l}}{\binom{N}{m}}$$

Beispiele für Zufallsvariablen, die abzählbar unendlich viele Werte annehmen.

- *Geometrische Verteilung*

#### 4.10 Definition Geometrische Verteilung

Sei  $p \in (0, 1)$ . Das W-Maß mit der Zählerdichte

$$k \in \mathbb{N}_0 \rightarrow \rho(k) := (1 - p)^k * p$$

heißt *geometrische Verteilung* zum Parameter  $p$  und wird mit  $G_p$  bezeichnet.

**Bemerkung:**

$$\sum_{k=0}^{\infty} \rho(k) = \sum_{k=0}^{\infty} (1 - p)^k p \quad \text{iii}$$

**Beispiel:** Eine Münze wird so lange geworfen bis  $Z$  fällt.

$\hookrightarrow$  Zufallsvariable:  $X$  Anzahl der Versuche, bis  $Z$  fällt ist geometrisch verteilt zum Parameter  $p = 0,5$ .

---

iii geometrische Reihe

**Diskussion:** Da es beliebig viele Versuche sein können bis  $Z(=1)$  fällt, scheint folgende Menge geeignet, in  $X$  auf  $\Omega$  zu modellieren.

$$\Omega = \{(x_1, x_2, x_3, \dots) : x_i \in \{0, 1\} : \forall i \in \mathbb{N}\}$$

Modifizierte Version: Wir brechen nach  $n$  Versuchen ab, unabhängig davon ob ein Treffer erzielt wurde oder nicht.

$$\tilde{\Omega} = \{(x_1, \dots, x_n) : x_i \in \{0, 1\} \forall i = 1, \dots, n\}$$

$X$  ist Anzahl der Versuche bis eine 1 fällt. Dies wird dargestellt als Abbildung auf den Laplace-Raum  $(\tilde{\Omega}, \mathcal{P}(\tilde{\Omega}), P)$

**Konkret:**

$$X : (x_1, \dots, x_n) \in \tilde{\Omega} \rightarrow \begin{cases} k \in \{0, \dots, n-1\} & , \text{so dass } x = 1 \text{ und } x_j = 0 \\ n & , \text{sonst} \end{cases}$$

$$\hookrightarrow \rho(k) = P_X(\{k\}) = \frac{|X^{-1}(\{k\})|}{2^n} = \frac{2^{n-k-1}}{2^n} = \left(\frac{1}{2}\right)^k * \frac{1}{2} \quad \forall k = 0, \dots, n-1$$

Das nennt man auch *Poisson-Verteilung*

## 4.11 Definition Poisson-Verteilung

Sei  $d > 0$ . Das W-Maß mit der Zähldichte

$$\rho : \mathbb{N}_0 \rightarrow [0, 1], k \mapsto e^{-\lambda} \frac{\lambda^k}{k!}$$

heißt *Poisson-Verteilung* zum Parameter  $d$  und wird mit  $P_\lambda$  bezeichnet.

**Bemerkung:**

- $\lambda$  wird als mittlere Rate interpretiert, mit der ein Ereignis in einem vorgegebenen Zeitfenster  $J$  beobachtet wird
- $\mathcal{P}_\lambda$  ist die Verteilung der Zufallsvariablen “Anzahl der Ergebnisse in  $J$ ”

**Beispiele:**

- Anzahl der ankommenden E-Mails/Tag
- Anzahl der Versicherungsfälle pro Jahr
- Anzahl der Kunden pro Stunde

**Bemerkungen:**

$$\rho(k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, \dots$$

$$\hookrightarrow \text{Dichte } e^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

## 5 Erwartungswerte und Varianz

### 5.1 Definition Reelwertigen Zufallsvariable

Eine Zufallsvariable, die Werte in  $\mathbb{R}$  annimmt heißt *reelle Zufallsvariable*. Alle bisherigen Zufallsvariablen sind Beispiele reeller Zufallsvariablen. Genauer handelt es sich bei den Zufallsvariablen um *diskret verteilte* Zufallsvariablen. Das für diese ist, dass die Werte in einer höchstens abzählbaren Menge  $\Omega \subset \mathbb{R}$  annehmen:

- $\Omega = \{0, 1\}$  bei  $X \sim \mathcal{B}_p$  ( $p \in [0, 1]$ )
- $\Omega = \{0, 1, \dots, n\}$  bei  $X \sim \mathcal{B}_{n,p}$
- $\Omega = \{0, 1, \dots, m\}$  bei  $X \sim \mathcal{H}_{m,k,N-k}$
- $\Omega = \mathbb{N}$  bei  $X \sim \mathcal{P}_\lambda$  sowie  $X \sim \mathcal{G}_p$

Jede auf einem diskreten W-Raum  $(\Omega', \mathcal{P}(\Omega'), P)$  definierte Zufallsvariable  $X$  gehört zu den diskreten Zufallsvariablen, denn  $X(\Omega')$  ist eine höchstens abzählbare Menge.

### 5.2 Definition Erwartungswert

Sei  $X$  eine reelle Zufallsvariable, die Werte in einer endlichen Menge  $\Omega \subset \mathbb{R}$  annimmt, dann heißt der Wert

$$\mathbb{E}(X) := \sum_{x \in \Omega} x P(X = x) = \sum_{x \in \Omega} x P_X(\{x\})$$

*Erwartungswert von  $X$ .*

Sei  $X$  eine Zufallsvariable, die abzählbar unendlich viele Werte  $x_i, i \in \mathbb{N}$  annehmen kann, jeweils mit der Wahrscheinlichkeit  $p_i$ , so heißt

$$\mathbb{E}(X) := \sum_{i=1}^{\infty} x_i p_i$$

*Erwartungswert (EW) von  $X$ , wenn die eben aufgeführte Reihe absolut konvergent ist.*

$$\hookrightarrow \sum_{i=1}^{\infty} x_i \text{ absolut konvergent} \Leftrightarrow \sum_{i=1}^{\infty} |x_i| < \infty$$

Der Erwartungswert ist eine reelle Zahl.

### Interpretation:

- Diese Zahl gibt den Wert an, den die betreffende Zufallsvariable im Mittel annimmt. Dabei gewichten wir die einzelnen Werte  $x$ , die  $X$  annehmen kann, entsprechend der Wahrscheinlichkeit  $p_x$  des Eintretens des Elementarereignisses  $X = x$ .
- Wenn wir das Experiment  $X$  unter identischen Bedingungen und ohne gegenseitige Beeinflussung oft wiederholen, erwarten wir, dass das arithmetische Mittel der einzelnen Ergebnisse für ein geeignetes W-Modell nah bei  $\mathbb{E}(X)$  liegt

### Beispiele

a) Erwartungswert von  $X \sim \mathcal{B}_p$

$$\mathbb{E}(X) = 0(1-p) + 1 \cdot p = p$$

b)  $\Omega = \{x_1, \dots, x_n\} \subset \mathbb{R}$  und  $X \sim \mathcal{U}_\Omega$  :

$$\mathbb{E}(X) = \sum_{i=1}^n x_i \frac{1}{n} = \frac{1}{n} \sum_{i=1}^n x_i \text{ entspricht Mittel der Zahlen } x_1, \dots, x_n$$

c) Erwartungswert von  $X \sim \mathcal{B}_{n,p}$

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^n k \mathcal{B}_{n,p}(\{k\}) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n k \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n np \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \\ &= np \sum_{l=0}^{n-1} \binom{n-1}{l} p^l (1-p)^{n-1-l} = np \end{aligned}$$

d) Erwartungswert von  $X \sim \mathcal{P}_\lambda$ :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^{\infty} k \mathcal{P}_\lambda(\{k\}) \\ &= \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda \end{aligned}$$

## Zusammenhang zwischen Poisson-und Binomialverteilung

### Beispiel:

- Sei  $\lambda > 0$  die mittlere Anzahl von Kunden, die ein Blumengeschäft im Zeitintervall  $I$  (z.B.  $I = 10$  Stunden) betreffen.
  - Wir zerlegen das Zeitintervall  $I$  in  $n$  gleich große Teilintervalle
  - Mit wachsendem  $n$  wird bei festen  $I$  die Länge der Teilintervalle<sup>iv</sup> immer kürzer
- ⇨ Ansatz für die Wahrscheinlichkeit  $p_n$ , dass ein Teilintervall der Länge  $\frac{I}{n}$  ein Kunde den Laden betritt

$$p_n = \frac{\lambda}{n}$$

**Annahme:** Ob ein Kunde im gegebenen Teilintervall den Laden betritt ist unabhängig davon, ob im anderen Teilintervall das Ereignis eintritt.

⇨ Die Wahrscheinlichkeit dafür, dass im Intervall  $I$  der Blumenladen von  $k$  Kunden besucht wird, wird entsprechend durch die Binomialverteilung bestimmt:

$$\mathcal{B}_{n, p_n = \frac{\lambda}{n}}(\{k\}) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

Dann lässt sich zeigen:

$$\lim_{n \rightarrow \infty} \mathcal{B}_{n, \frac{\lambda}{n}}(\{k\}) = \mathcal{P}_\lambda(\{k\})$$

### 5.3 Theorem

Seien  $\lambda > 0$  und  $p_n \in (0, 1)$ ,  $n \in \mathbb{N}$  eine Folge, so dass  $\lim_{n \rightarrow \infty} n p_n = \lambda$ . Dann gilt für jedes  $k \in \mathbb{N}$ :

$$\lim_{n \rightarrow \infty} \mathcal{B}_{n, p_n}(\{k\}) = \mathcal{P}_\lambda(\{k\})$$

### 5.4 Definition

Seien  $f : \mathbb{N} \rightarrow \mathbb{R}$  und  $g : \mathbb{N} \rightarrow \mathbb{R}$  zwei Folgen

$$f(k) \sim g(k) \text{ für } k \rightarrow \infty \Leftrightarrow \frac{f(k)}{g(k)} = 1 \text{ für } k \rightarrow \infty$$

---

<sup>iv</sup> =  $\frac{I}{n}$



## 5.5 Lemma

Sei  $k \in \mathbb{N}$ , dann gilt im Grenzwert  $n \rightarrow \infty$

$$\binom{n}{k} \sim \frac{n^k}{k!}$$

### Beweis zu Lemma 5.4

$$\begin{aligned} \binom{n}{k} &= \frac{n!}{(n-k)!k!} = \frac{n(n-1)\dots(n-k+1)}{k!} \\ &= \frac{n^k}{k!} \frac{(n-1)(n-2)\dots(n-k+1)}{n^{k-1}} \\ &= \frac{n^k}{k!} \frac{n(1-\frac{1}{n})}{n} \frac{n(1-\frac{2}{n})}{n} \dots \frac{n(1-\frac{k-1}{n})}{n} \\ &\sim \frac{n^k}{k!} \end{aligned}$$

## 5.6 Theorem

Sei  $\Omega', \mathcal{P}(\Omega'), P$  ein diskreter W-Raum und  $X, Y$  zwei Zufallsvariablen mit  $\mathbb{E}(X)$  und  $\mathbb{E}(Y)$  definiert. Dann gilt:

- a)  $\mathbb{E}(cX) = c\mathbb{E}(X) \quad \forall c \in \mathbb{R}$
- b)  $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$
- c)  $X \leq Y \Rightarrow \mathbb{E}(X) \leq \mathbb{E}(Y)$

**Beweis:** a)

$$\mathbb{E}(cX) \stackrel{1}{=} \sum_{z \in \Omega} z P_{cX}(\{z\}) = \sum_{z \in c\Omega} z P_X\left(\left\{\frac{z}{c}\right\}\right) = \sum_{x \in \Omega} cx P_X(\{x\}) = c\mathbb{E}(X)$$

1.  $cX$  ist eine Zufallsvariable mit Werten in  $c\Omega =: \{cx \in \mathbb{R} : x \in \Omega\}$  eine Verteilung  $P_{cX}$
- 2.

$cX : \Omega \mapsto c\Omega$  ist eine Zufallsvariable über  
 $x \mapsto cX \quad (\Omega, \mathcal{P}(\Omega), P_X)$ , wobei

$$P_X = P \circ X^{-1}$$

$$\hookrightarrow P_{cX} = P_X \circ (cX)^{-1} \Rightarrow P_{cX}(\{z\}) = P_X(cX = z) = P_X\left(\left\{\frac{z}{c}\right\}\right)$$

## 5.7 Definition Varianz

### 5.7.1 Endliche Zufallsvariable

Sei  $X$  eine endlich Zufallsvariable, mit Werten  $x \in \Omega$ , mit Wahrscheinlichkeit  $p_x$ . Dann heißt:

$$\sum_{x \in \Omega} (x - \mathbb{E}(X))^2 p_x =: \text{Var}(X)$$

die *Varianz*<sup>v</sup> von  $X$ .

### 5.7.2 Abzählbare Zufallsvariable

Sei  $X$  eine abzählbar, unendliche Zufallsvariable, mit den Werten  $x \in \Omega$ ,  $i \in \mathbb{N}$ , mit einem abzählbaren  $\Omega$  mit Wahrscheinlichkeit  $p_i$  und der Erwartungswert  $\mathbb{E}(X)$  ist konvergent. Dann heißt:

$$\sum_{i=1}^{\infty} (x_i - \mathbb{E}(X))^2 p_i =: \text{Var}(X)$$

die Varianz von  $X$ .

**Bemerkung:**

$$\text{Var}(X) \geq 0$$

$$\text{Var}(X) = \infty \text{ sind möglich}$$

$$\forall X \exists v \in \mathbb{R} \cup \{+\infty\} \quad \text{Var}(X) = v$$

## 5.8 Definition Standardabweichung

Sei  $X$  eine Zufallsvariable mit  $\text{Var}(X) \in \mathbb{R}$ , dann wird

$$\sigma(X) = \sqrt{\text{Var}(X)}$$

*Standardabweichung* genannt.

**Interpretation:** Die Varianz gibt den Mittelwert der Abweichung zum Quadrat der Zufallsvariable  $X$  gegenüber des Erwartungswertes  $\mathbb{E}(X)$  an.

---

<sup>v</sup>Auch *Streuung* genannt.

a)  $X \sim \mathcal{B}_p$ , mit  $\mathbb{E}(X) = p$

$$\begin{aligned}
 \text{Var}(X) &= (1-p)^2 p + (0-p)^2 (1-p) \\
 &= (1-p) (p(1-p) + p^2) \\
 &= (1-p) (p - p^2 + p^2) \\
 &= (1-p) p \\
 \text{und } \sigma(X) &= \sqrt{(1-p)p}
 \end{aligned} \tag{1}$$

b)  $X \sim \mathcal{U}_{\Omega_n}$  mit  $\Omega_n = \{x_1, \dots, x_n\}$ ,  $x_i \in \mathbb{R}$

$$\begin{aligned}
 \text{Dann ist } \mathbb{E}(X) &= \frac{1}{n} \sum_{x \in \Omega_n} x =: \bar{x} \\
 \hookrightarrow \text{Var}(X) &= \sum_{i=1}^n (x_i - \bar{x})^2 \frac{1}{n} \\
 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \\
 &= \frac{1}{n} \sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\
 &= \frac{1}{n} \sum_{i=1}^n x_i^2 - \frac{2\bar{x}}{n} \sum_{i=1}^n x_i + \bar{x}^2 \\
 &= \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 \\
 &= \mathbb{E}(X^2) - \mathbb{E}(X)^2
 \end{aligned}$$

## 5.9 Definition Zweiter Moment

Sei  $X$  eine Zufallsvariable mit Werten  $x \in \Omega$  mit Wahrscheinlichkeiten  $p_x$ , wenn

$$\sum_{x \in \Omega} x^2 p_x$$

konvergent, dann nennt man diesen Wert *das zweite Moment von  $X$* .

$$\mathbb{E}(X^2) = \sum_{x \in \Omega} x^2 p_x$$

## 5.10 Theorem

Sei  $X$  eine Zufallsvariable mit reellen, diskreten Werten und dem Erwartungswert  $\mathbb{E}(X) = m$ . Wenn  $\mathbb{E}(X^2) \in \mathbb{R}$ , dann gilt:

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

.

### 5.10.1 Beispiele

**Beispiel 1:**  $X \sim \mathcal{B}_{n,p}$ ,  $\mathbb{E}(X) = np$ ,  $\text{Var}(X) = ?$  bzw.  $\sigma(X) = ?$

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}(X^2) - (np)^2 \\ &= \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} - (np)^2 \\ &= \sum_{k=1}^n npk \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} \\ &= np \sum_{l=0}^{n-1} (l+1) \binom{n-1}{l} p^l (1-p)^{(n-1)-l}, \quad \tilde{X} \sim \mathcal{B}_{n-1,p} \\ &= np \left( \mathbb{E}(\tilde{X}) + 1 \right) - (np)^2 = np((n-1)p + 1) - (np)^2 = np(1-p) \end{aligned}$$

und  $\sigma(X) = \sqrt{\text{Var}(X)} = \sqrt{np(1-p)}$

**Beispiel 2**  $X$  ist eine Zufallsvariable mit Werten in  $\{-1, 1\}$  und zugehöriger Zähldichte

$$\begin{aligned} \rho(-1) &= \frac{1}{2} \text{ und } \rho(1) = \frac{1}{2} \\ \hookrightarrow \mathbb{E}(X) &= (-1) \frac{1}{2} + 1 \cdot \frac{1}{2} = 0 \\ \text{Var}(X) &= (-1)^2 \frac{1}{2} + 1^2 \cdot \frac{1}{2} = \frac{1}{2} + \frac{1}{2} = 1 \\ \hookrightarrow \sigma(X) &= \sqrt{1} = 1 \end{aligned}$$

### Beispiel 3

$$X \sim P_\lambda \quad (\lambda > 0), \quad \mathbb{E}(X) = \lambda$$

Wir brauchen das zweite Moment von  $X$ :

$$\begin{aligned}
 \mathbb{E}(X^2) &= \sum_{k=0}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} \\
 &= \lambda \cdot \sum_{k=1}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} \\
 &= \lambda \cdot \sum_{l=0}^{\infty} (l+1) e^{-\lambda} \frac{\lambda^l}{l!}, \quad l := k-1 \\
 &= \lambda \cdot (\lambda + 1) = \lambda^2 + \lambda
 \end{aligned}$$

## 6 Stochastische Unabhängigkeit

### 6.1 Definition

Sei  $(\Omega, \mathcal{P}(\Omega), P)$  ein diskreter W-Raum. Zwei Ereignisse  $E, F \in \mathcal{P}(\Omega)$  heißen *stochastisch Unabhängig* bezüglich  $P$ , wenn

$$P(E \cap F) = P(E) \cdot P(F)$$

gilt.

**Beispiel:** Es wird ein blauer und ein roter Spielwürfel geworfen. Im entsprechenden Laplace-Modell mit dem Ergebnisraum  $\Omega = \{1, \dots, 6\}^2$  sind die Ereignisse:

- a)  $B$ : der blaue Würfel zeigt die Augenzahl “2”  
 $R$ : der rote Würfel zeigt die Augenzahl “4”  
 stochastisch unabhängig, denn

$$|B| = |\{(b, r) \in \Omega : b = 2\}| = 6 \Rightarrow P(B) = \frac{6}{36} = \frac{1}{6}$$

$$|R| = |\{(b, r) \in \Omega : r = 4\}| = 6 \Rightarrow P(R) = \frac{1}{6}$$

$$|B \cap R| = |\{(2, 4)\}| = 1 \Rightarrow P(B \cap R) = \frac{1}{36}$$

$$\text{d.h. } P(B \cap R) = P(B) \cdot P(R)$$

- b)  $B$  und  $R_{>2}$ : der rote Würfel zeigt die Augenzahl  $> 2$  stochastisch unabhängig, denn

$$|R_{>2}| = |\{(b, r) \in \Omega : r > 2\}| = 6 \cdot 4 = 24 \Rightarrow P(R_{>2}) = \frac{2}{3}$$

$$|B \cap R_{>2}| = \frac{1}{9} = \frac{2}{3} \cdot \frac{1}{6} = P(B) \cdot P(R_{>2})$$

- c)  $B$  und  $E_g$ : beide Augenzahlen sind gerade  
stochastisch *abhängig*, denn

$$\begin{aligned} |E_g| &= |\{(b, r) \in \Omega : b, r \in \{2, 4, 6\}\}| = 9 \Rightarrow P(E_g) = \frac{1}{4} \\ |E_g \cap B| &= |\{(b, r) \in \Omega : b = 2, r \in \{2, 4, 6\}\}| = 3 \\ \Rightarrow P(E_g \cap B) &= \frac{1}{12} \neq \frac{1}{4} \cdot \frac{1}{6} = P(E_g) \cdot P(B) \end{aligned}$$

- d)  $B$  und  $E$ : beide Augenzahlen sind gerade oder beide sind ungerade

$$E = E_g \cup E_u \Rightarrow P(E) = \frac{18}{36} = \frac{1}{2}$$

so dass  $B$  und  $E$  stochastisch unabhängig sind bezüglich  $P = \mathcal{U}_\Omega$

**Bemerkung:** Sei  $(\Omega, \mathcal{P}(\Omega))$  ein diskreter Ereignisraum. Dann sind je zwei disjunkte Ereignisse  $E, F \in \mathcal{P}(\Omega)$  mit  $E, F \neq \emptyset$  stochastisch abhängig bezüglich jedem W-Maß auf  $\mathcal{P}(\Omega)$ , dessen zugehörige Zähldichte  $\rho$  auf  $\Omega$  strikt positiv ist ( $\rho(\omega) > 0 \forall \omega \in \Omega$ ), denn

$$P(E \cap F) = P(\emptyset) = 0 \neq \underbrace{P(E)}_{=\sum_{\omega \in E} \rho(\omega) > 0} \cdot P(F) > 0$$

## 6.2 Definition Unabhängige Zufallsvariablen

Seien  $(\Omega, \mathcal{P}(\Omega), P)$  ein diskreter W-Raum und  $X, Y$  zwei Zufallsvariablen auf  $(\Omega, \mathcal{P}(\Omega), P)$  jeweils mit Werten in  $\Omega_X$  beziehungsweise  $\Omega_Y$ . Dann heißen  $X$  und  $Y$  unabhängige Zufallsvariablen, wenn  $\{X \in E\}$  und  $\{Y \in F\}$  bezüglich  $P$  stochastisch unabhängig sind für jede Wahl von Ereignissen  $E \subset \Omega_X$  und  $F \subset \Omega_Y$ .

**Beachte:**

$$\{X \in E\} = X^{-1}(\underbrace{E}_{\in \Omega_X}) = \{\omega \in \Omega : X(\omega) \in E\}$$

Wir können die Unabhängigkeit von  $X$  &  $Y$  wie folgt definieren(\*):  
 $X, Y$  sind unabhängige Zufallsvariablen

$$:\Leftrightarrow P(X \in E, Y \in F) = P(X \in E) \cdot P(Y \in F) \quad \forall E \subseteq \Omega_X, \forall F \subseteq \Omega_Y$$

Aus dieser Relation folgt(\*\*):

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y) \quad \forall x \in \Omega_X \text{ und } \forall y \in \Omega_Y$$

Umgekehrt impliziert (\*\*) die Relation (\*), d.h. dass  $X$  und  $Y$  unabhängig sind. Wie sieht man das?

- $x \in \Omega \mapsto \rho_X(x) := P(X = x)$  ist die Gewichtsfunktion der Verteilung  $P_X$  von  $X$
- $y \in \Omega_Y \mapsto \rho_Y(y) := P(Y = y)$  ist die Gewichtsfunktion der Verteilung  $P_Y$  von  $Y$
- $(x, y) \in \Omega_X \times \Omega_Y \mapsto P(X = x, Y = y)$  ist die Gewichtsfunktion der Verteilung einer Zufallsvariable mit Werten aus  $\Omega_X \times \Omega_Y$

### 6.3 Lemma Unabhängigkeit

Unter der Voraussetzung aus der Definition Unabhängige Zufallsvariablen (siehe 6.2) sind zwei Zufallsvariablen  $X$  und  $Y$  auf  $(\Omega, \mathcal{P}(\Omega), P)$  unabhängig genau dann, wenn für jedes  $x \in \Omega_X$  und jedes  $y \in \Omega_Y$  gilt

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$

#### 6.3.1 Beispiel 1

$\mathcal{E}$ : n-maliges Werfen eines Spielwürfels

$\hookrightarrow$  Der W-Raum zu diesem Zufallsexperiment ist:

$$(\Omega, \mathcal{P}(\Omega), P = \mathcal{U}_\Omega) \text{ mit } \Omega := \{1, \dots, 6\}^n$$

Seien  $1 \leq k, l \leq n$  mit  $k \neq l$ . Dann sind

- $X_k$ : Augenzahl bei Wurf  $k$ , und
- $X_l$ : Augenzahl bei Wurf  $l$

Denn für jedes  $x \in \Omega_{X_k} := X_k(\Omega)$  und jedes  $y \in \Omega_{X_l} := X_l(\Omega)$  erhalten wir

$$\begin{aligned} P(X_k = x, X_l = y) &= P(\{\omega \in \Omega : \omega_k = x, \omega_l = y\}) \\ &= \frac{6^{n-2}}{6^n} = \frac{1}{6} \cdot \frac{1}{6} = P(X_k = x) \cdot P(X_l = y) \end{aligned}$$

Nach Lemma 6.3 sind somit  $X_k$  und  $X_l$  zwei unabhängige Variablen.

#### 6.3.2 Beispiel 2

**Gegeben:** eine Urne mit 10 gleichartigen, aber nummerierten Kugeln.

**Zufallsexperiment:** Stichprobe ohne Zurücklegen: es werden 2 Kugeln nacheinander gezogen

- $X$ : Nummer der ersten gezogenen Kugel
- $Y$ : Nummer der zweiten gezogenen Kugel

## 6.4 Definition Familie von Ereignissen

Seien  $(\Omega, \mathcal{P}(\Omega), P)$  ein diskreter W-Raum und  $I$  eine beliebige, nicht leere Indexmenge. Eine Familie von Ereignissen  $E_i \in \mathcal{P}(\Omega) \mid i \in I$ , heißt *stochastisch unabhängig*, wenn für jede endliche Teilmenge  $J \in I$

$$P\left(\bigcap_{j \in J} E_j\right) = \prod_{j \in J} P(E_j)$$

gilt.

**Beispiel:** (Kontext: Beispiele zur Definition 6.1)

Die Familie der Ereignisse  $B, E$  und  $R$  ist nicht stochastisch unabhängig, obwohl die Ereignisse *paarweise* stochastisch unabhängig sind (Wie wir bereits nachgerechnet haben), denn

$$B \cap R \cap E = \emptyset \Rightarrow P(\emptyset) = 0 \neq P(E) \cdot P(R) \cdot P(B)$$

**Bemerkung:** Auch die Familie  $B, E$  und  $R_4$ : **roter Würfel zeigt "4"** ist stochastisch *nicht* unabhängig.

## 6.5 Definition Unabhängige Familie

Seien  $(\Omega, \mathcal{P}(\Omega), P)$  ein diskreter W-Raum,  $I$  eine beliebige, nicht-leere Indexmenge und  $X_i, i \in I$  eine Familie von Zufallsvariablen auf  $(\Omega, \mathcal{P}(\Omega), P)$  jeweils mit Werten in einer höchstens abzählbaren  $\Omega_{X_i}$ . Dann heißt die Familie  $X_i, i \in I$  unabhängig, wenn für jede Wahl von Ereignissen  $E_i \subset \Omega_i, i \in I$  eine stochastisch unabhängige Familie ist.

**Standardbeispiel:** Seien  $n \in \mathbb{N}$ ,  $\Omega = \{1, \dots, 6\}^n$  und

$$\underbrace{(\Omega, \mathcal{P}(\Omega), P = \mathcal{U}_\Omega)}_{\text{Experiment } \mathcal{E}: n\text{-maliges würfeln eines Würfels}}$$

Wir betrachten:

$$X_i : (\omega_1, \dots, \omega_n) \in \Omega \mapsto \omega_i$$

$$P(X_{i_1} = x_{i_1}, \dots, X_{i_m} = x_{i_m}) = \frac{6^{n-m}}{6^n} = \frac{1}{6^m} = \prod_{l=1}^m P(X_{i_l} = x_{i_l})$$

## 7 Bedingte Wahrscheinlichkeit

**Formaler Rahmen:**  $(\Omega, \mathcal{P}(\Omega), P)$  diskretes W-Raum

Gegeben:  $A, B \subseteq \Omega$  zwei Ereignisse

Annahme:  $P(A) > 0$



**Szenario:** Wir beobachten, dass  $A$  eingetreten ist.

$\hookrightarrow$  Frage: Wie ändert sich unter der obigen Annahme unsere Wahrscheinlichkeitsbewertung für  $B$ ?

$\hookrightarrow$  Fallunterscheidung:

- Sind  $A$  und  $B$  stochastisch unabhängig  
 $\hookrightarrow$  kein Anlass für eine Neubewertung
- $A$  &  $B$  sind nicht stochastisch unabhängig  
 $\hookrightarrow$  Wie genau soll die Neubewertung ausfallen?

**Beispiel:** Gegeben ist eine Urne mit  $N$  Kugeln, davon genau  $s$  schwarze und  $w$  weiße. D.h.  $N = s + w$ .  $\hookrightarrow$  Modell:  $U = \underbrace{\{1, \dots, w\}}_{\text{Teilmenge der weißen Kugeln}} \cup \underbrace{\{w+1, \dots, s+w\}}_{\text{Teilmenge der schwarzen Kugeln}}$

**Zufallsexperiment:** Es werden zwei Kugeln gezogen, ohne Zurücklegen.

$\hookrightarrow$  Ergebnismenge :  $\Omega = \{(k, l) : k, l \in U, k \neq l\}$

$\hookrightarrow |\Omega| = N \cdot (N - 1) = (s + w) \cdot (s + w - 1)$

$P = \mathcal{U}_\Omega$  Gleichverteilt auf  $\Omega$

Wir betrachten die Ereignisse:

$A$  : 1. gezogene Kugel ist weiß

$B$  : 2. gezogene Kugel ist weiß

$A$  und  $B$  sind nicht stochastisch unabhängig bezüglich  $P = \mathcal{U}_\Omega$ , wobei  $\Omega = \{(k, l) : k, l \in U, k \neq l\}$ , denn

- $P(A) = \frac{|A|}{(w+s)(w+s-1)} = \frac{w(w+s-1)}{(w+s)(w+s-1)} = \frac{w}{w+s}$
- $P(B) = P(A)$ , da  $|A| = |B|$  ( $f : A \rightarrow B, (k, l) \rightarrow (l, k)$  ist eine Bijektion)
- $P(A \cap B) = \frac{|A \cap B|}{(w+s)(w+s-1)} = \frac{w(w-1)}{(w+s)(w+s-1)} \neq \left(\frac{w}{w+s}\right)^2 = P(A) \cdot P(B)$   
 $\hookrightarrow$  Bedingung für stochastische Unabhängigkeit nicht erfüllt  $\Rightarrow A$  &  $B$  stochastisch unabhängig bezüglich  $P = \mathcal{U}_\Omega$

**Annahme:**  $A$  ist eingetreten (1. gezogene Kugel ist weiß)

$\hookrightarrow$  **Frage:** Wie groß ist die Wahrscheinlichkeit für  $B$  unter dieser Annahme?

**Antwort:**

$$\frac{w-1}{(w+s-1)} \neq \frac{w}{s+w} = P(B)$$

Hier gilt:

$$\frac{w-1}{w+s-1} = \frac{P(A \cap B)}{P(A)} = \frac{w(w-1)}{(w+s)(w+s-1)} \cdot \frac{w+s}{w}$$

## 7.1 Definition Bedingte Wahrscheinlichkeit

Sei  $(\Omega, \mathcal{P}(\Omega), P)$  ein diskreter W-Raum und  $A \subseteq \Omega$  mit  $P(A) > 0$ . Dann heißt für jedes  $B \subseteq \Omega$

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

*bedingte Wahrscheinlichkeit* für  $B$  gegen  $A$ .

**Beachte:** Im Spezialfall, dass  $A$  &  $B$  stochastisch unabhängig bezüglich  $P$  sind, gilt

$$P(B|A) = P(B) \quad \left( \text{denn } P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A) \cdot P(B)}{P(A)} \right)$$

**Bemerkung:** Bedingen macht keine Aussagen über Kausalitäten!

$\hookrightarrow$  Beispiel:

- 1. gezogene Kugel bleibt verdeckt
- 2. gezogene Kugel ist weiß (d.h.  $B$  tritt ein)

$\hookrightarrow$  **Frage:** Wahrscheinlichkeit dafür, dass  $A$  eingetreten ist, d.h. die verdeckte Kugel weiß ist?

$\hookrightarrow$  **Antwort:**  $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{w-1}{w+s-1} \neq P(A)$

$\hookrightarrow$  Das Ereignis  $B$  hat keinen Einfluss auf das Eintreten von  $A$ . Trotzdem bewirkt  $B$  eine Neubewertung der Wahrscheinlichkeit für  $A$ . Das liegt daran dass  $B$  Informationen aus  $A$  enthält.  $B$  wird nämlich durch  $A$  beeinflusst.

**Beispiel:** Ziehen einer Kugel aus einer von 2 Urnen.

Urne 1: enthält 7 rote und 3 schwarze Kugeln

Urne 2: enthält 9 rote und 1 schwarze Kugel

Es wird zufällig entschieden aus welcher Urne gezogen wird. (z.B. auf der Grundlage des Ergebnisses eines Münzwurfes)

**Ergebnismenge des Zufallsexperimentes:**

$$\Omega = \{1, 2\} \times \{1, \dots, 10\} = \{(a, b) : a \in \{1, 2\}, b \in \{1, \dots, 10\}\}$$

Wir betrachten die Ereignisse:

$$A = \{1\} \times \{1, \dots, 10\} = \{(a, b) \in \Omega : a = 1\} \text{ (es wird aus Urne 1 gezogen)}$$

$$B = \{2\} \times \{1, \dots, 10\} = \{(a, b) \in \Omega : a = 2\} \text{ (es wird aus Urne 2 gezogen)}$$

$$R = \{(a, b) \in \Omega, a = 1, b \leq 7\} \cup \{(a, b) \in \Omega : a = 2, b \leq 9\}$$

$$S = \{(a, b) \in \Omega, a = 1, b > 7\} \cup \{(a, b) \in \Omega : a = 2, b = 10\}$$

$\hookrightarrow$  **Frage:**

- Wahrscheinlichkeit, dass eine rote Kugel gezogen wird, wenn die Wahl für Urne 1 fällt. D.h. wir fragen nach der bedingten Wahrscheinlichkeit:

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

- Wahrscheinlichkeit dafür, dass eine schwarze Kugel gezogen wird, wenn die Wahl auf Urne 1 fällt:

$$\begin{aligned} P(S|A) &= \frac{P(S \cap A)}{P(A)} \\ &= \frac{|S \cap A|}{|\Omega|} \cdot \frac{|\Omega|}{|A|} \\ &= \frac{|S \cap A|}{|A|} \\ &= \frac{3}{10} \end{aligned}$$

- Wie groß ist die Wahrscheinlichkeit, eine rote Kugel zu ziehen? D.h.  $P(R) = ?$

$$P(R) = \underbrace{\frac{1}{2} \cdot \frac{7}{10} + \frac{1}{2} \cdot \frac{9}{10}}_{\text{Wofür stehen diese Zahlen genau}}$$

$$\begin{aligned}
P(R) &= P(A) \cdot P(R|A) + P(B) \cdot P(R|B) \\
&= P(A) \cdot \frac{P(R \cap A)}{P(A)} + P(B) \cdot \frac{P(R \cap B)}{P(B)} \\
&\text{(da gilt: } A \cap B = \emptyset \Rightarrow (A \cap R) \text{ und } (B \cap R) \text{ sind disjunkt)} \\
&= P((R \cap A) \cup (R \cap B)) \\
&= P\left(R \cap \underbrace{(A \cup B)}_{=\Omega}\right) \\
&= P(R)
\end{aligned}$$

## 7.2 Satz von der totalen Wahrscheinlichkeit

Sei  $(\Omega, \mathcal{P}(\Omega), P)$  ein diskreter W-Raum und  $n \in \mathbb{N}$ . die Ereignisse  $H_1, \dots, H_n \subset \Omega$  seien eine Zerlegung von  $\Omega$ , d.h.

$$H_i \cap H_j = \emptyset \quad \forall i, j \in \{1, \dots, n\}, i \neq j \text{ und } \bigcup_{i=1}^n H_i = \Omega$$

Gilt  $P(H_j) > 0 \quad \forall j = 1, \dots, n$ , dann gilt für jedes Ereignis  $A \subset \Omega$

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

Eine weitere mögliche Frage: Wie groß ist die Wahrscheinlichkeit, dass die gezogene rote Kugel aus Urne 1 stammt. (siehe 7.1)

$\hookrightarrow$  Frage:  $P(A|R) = ?$

$\hookrightarrow$  "Problem": Aus der Aufgabenstellung ist statt  $P(A|R)$  die bedingte Wahrscheinlichkeit  $P(R|A)$  leicht herzuleiten

$\hookrightarrow$  Ansatz siehe "Formel von Bayer" (siehe 7.3)

## 7.3 Satz Formel von Bayer

Für alle Ereignisse  $A, B \subseteq \Omega$  mit  $P(A) > 0$  gilt

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

**Beweis:**

$$\begin{aligned}\frac{P(A|B) \cdot P(B)}{P(A)} &= \frac{\frac{P(A \cap B)}{P(B)} \cdot P(B)}{P(A)} \\ &= \frac{P(A \cap B)}{P(A)} \\ &= P(B|A)\end{aligned}$$

**Beispiel:** Ziehen einer Kugel aus einer von zwei Urnen:

Urne 1: genau 7 rote Kugeln und 3 schwarze Kugeln  $\Rightarrow P(R|A) = \frac{7}{10}$

Urne 2: genau 1 rote und 9 schwarze Kugeln  $\Rightarrow P(R|B) = \frac{1}{10}$

enthält.

$$\hookrightarrow P(R) = P(R|A) \cdot P(A) + P(R|B) \cdot P(B) = \frac{2}{5}$$

$\hookrightarrow$  **Frage:**  $P(A|R) = ?$  ( $\hat{=}$  Wahrscheinlichkeit dafür, dass die gezogene Kugel aus Urne 1 stammt, wenn die Kugel rot ist.)

**Antwort:**

$$P(A|R) = \frac{P(R|A) \cdot P(A)}{P(R)} = \frac{\frac{7}{10} \cdot \frac{1}{2}}{\frac{2}{5}} = \frac{7}{8}$$

**Zusammenfassung:** Durch das Eintreten eines Ereignisses  $A$  muss die Wahrscheinlichkeit von allen anderen Ereignissen angepasst werden.

$\hookrightarrow$  Übergang zu einem neuen/modifizierten W-Maß  $P_A$ , welches die Information über das Eintreten von  $A$  berücksichtigt.<sup>vi</sup>

Wir fordern:

F1)  $P_A(A) = 1$ , **Motivation:**  $A$  stellt ein sicheres Ereignis dar

F2) Es gibt ein  $k_A > 0$ , so dass für jedes  $C \subset \mathcal{A}$  gilt:

$$P_A(C) = k_A \cdot P(C)$$

$\hookrightarrow$  **Motivation/Beispiel:**

$$\text{WS für } = \begin{cases} \text{Schnee} & = \frac{1}{3} \\ \text{Regen} & = \frac{1}{3} \\ \text{niederschlagsfrei} & = \frac{1}{3} \end{cases}$$

---

<sup>vi</sup>  $P \rightsquigarrow^A P_A$

## 7.4 Satz Existenz und Eindeutigkeit bedingter W-Maße

Sei  $(\Omega, \mathcal{P}(\Omega), P)$  ein diskreter W-Raum und  $A \subset \Omega$  mit  $P(A) > 0$ . Dann existiert genau ein W-Maß

$$P_A : \mathcal{P}(\Omega) \rightarrow [0, 1]$$

mit den Eigenschaften F1) und F2), nämlich

$$(P) \ni B \mapsto P_A(B) := P(B|A) \quad (2)$$

**Beweis:**

1) Wir verifizieren, dass Gleichung (2) F1) & F2 erfüllt:

F1)

$$P_A(A) = P(A|A) = \frac{P(A \cap A)}{P(A)} = 1$$

F2)

$$C \subset A \Rightarrow P_A(C) = P(C|A) = \frac{P(C \cap A)}{P(A)} = \frac{1}{P(A)} \cdot P(C)$$

$\hookrightarrow$  F2) ist erfüllt mit  $k_A = \frac{1}{P(A)}$

2) Sei  $P_A: \mathcal{P}(\Omega) \rightarrow [0, 1]$  ein W-Maß mit F1) & F2). Dann gilt:

$$P_A \text{ F1) erfüllt} \Rightarrow P_A(A^C) = 1 - P_A(A) = 0 \quad (3)$$

nun ist  $(A, A^C)$  eine Zerlegung von  $\Omega$

$\hookrightarrow$  Nun betrachten wir für die beliebige  $B \subset \Omega$  die Zerlegung

$$B = (B \cap A) \dot{\cup} (B \cap A^C)$$

$$\hookrightarrow P_A(B) = P_A(\underbrace{B \cap A}_{\subset A}) + P(\underbrace{B \cap A^C}_{\subset A^C})$$

$= 0, \text{ wegen (3)}$

$$\hookrightarrow P_A(B) = P_A(B \cap A) = k_A \cdot P(B \cap A) \quad (4)$$

Setze in Gleichung (4) :  $B = A$

$$\hookrightarrow 1 = P_A(B = A) = k_A \cdot P(A \cap A) = k_A \cdot P(A)$$

$$\Leftrightarrow k_A = \frac{1}{P(A)}$$

**Beispiel:** Ein Beutel mit 3 Münzen:

- eine gewöhnliche Münze: **KZ**

- eine Münze: **KK**
- eine Münze: **ZZ**

Eine Münze wird zufällig gezogen und geworfen:

**Annahme:** Ergebnis ist “Zahl” (oben zu sehen)

**Frage:** Was ist die Wahrscheinlichkeit dafür, dass auf der Rückseite (unten) “Kopf” liegt?

Wir formulieren die relevanten Ereignisse.

A: gezogene Münze ist ZZ

B: gezogene Münze ist KK

C: gezogene Münze ist KZ

D: geworfene Münze zeigt nach oben “Z” (Zahl)

E: geworfene Münze zeigt nach unten “K” (Kopf)

Mit dieser Bezeichnung lautet die gestellte Frage:  $P(E|D) = ?$

Definitionsgemäß gilt:

$$P(E|D) = \frac{P(E \cap D)}{P(D)}$$

$$\begin{aligned} \hookrightarrow P(D) &= P(D|A) \cdot P(A) + P(D|B) \cdot P(B) + P(D|C) \cdot P(C) \\ &= 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} \\ &= \frac{1}{2} \end{aligned}$$

$$\begin{aligned} P(D \cap E) &= P(D \cap E|A) \cdot P(A) + \dots \\ &= 0 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} \\ &= \frac{1}{6} \end{aligned}$$

$$\hookrightarrow P(E|D) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

## 8 Gemeinsame Verteilung

**Ausgangspunkt:** Es existiert ein diskreter W-Raum  $(\Omega, \mathcal{P}(\Omega), P)$ .

Gegeben seien zwei Abbildungen  $X$  und  $Y$  auf  $(\Omega, \mathcal{P}(\Omega), P)$ .

$\hookrightarrow$  Sie stellen Zufallsvariablen dar mit jeweils Werten in  $X(\Omega) := \Omega_X$  und  $Y(\Omega) := \Omega_Y$ , d.h.

$$X : \Omega \rightarrow \Omega_X \text{ und } Y : \Omega \rightarrow \Omega_Y$$

$X$  und  $Y$  heißen unabhängige (siehe 6.2), wenn gilt:

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B), \forall A \subset \Omega_X, \forall B \subset \Omega_Y$$

Rechte Seite:  $A \subset \Omega_X \rightarrow P(X \in A) \hat{=} \text{Verteilung von } X$

$B \subset \Omega_Y \rightarrow P(Y \in B) \hat{=} \text{Verteilung von } Y$

$$\text{Linke Seite: } (A \in \Omega_X, B \in \Omega_Y) \rightarrow P(X \in A, Y \in B) \quad (5)$$

**Frage:** Stellt die Abbildung (siehe Gleichung(5)) ebenfalls eine Verteilung/W-Maß dar?

**Antwort:** Das ist nicht der Fall, da der Definitionsbereich der Abbildung (Gleichung 5) zu klein ist. Er lässt sich nur mit einer echten Teilmenge der Potenzmenge des kartesischen Produktes  $\Omega_X \times \Omega_Y$  identifizieren.

**Beispiel:**  $\underbrace{\{1, 2\}}_{=\Omega_M} \times \underbrace{\{1, \dots, 6\}}_{=\Omega_W} \hat{=} \text{Ergebnismenge zum Zufallsexperiment, bei dem eine Münze und ein Würfel geworfen werden. Betrachte das Ereignis:}$

$$E := \{(1, 2), (2, 1)\} \subseteq \Omega_M \times \Omega_W$$

$\hookrightarrow$  Offensichtlich existiert keine Teilmenge  $A \subset \Omega_M$  und  $B \subset \Omega_W$ , so dass

$$E = A \times B = \{(x, y) : x \in A, y \in B\}$$

**Diskussion:** Wir stellen eine minimale Forderung an die Erweiterung des Definitionsbereiches (aus Gleichung 5).

$\hookrightarrow$  abstrakte Formulierungen: Sei  $\Omega$  eine Menge und  $\mathcal{A}$  ein System von Ereignissen bzw. Teilmengen von  $\Omega$ . Wir fordern:

1.  $\Omega \in \mathcal{A}$
2.  $E \in \mathcal{A} \Rightarrow E^c = \Omega \setminus E \in \mathcal{A}$



3.  $I$  eine abzählbare Indexmenge und  $E_i \in \mathcal{A}_i, i \in I$

$$\Rightarrow \bigcup_{i \in I} E_i \in \mathcal{A}_i$$

Im obigen Beispiel:

$$\begin{aligned} E &= \{(1, 2), (2, 1)\} = \{(1, 2)\} \cup \{(2, 1)\} \\ &= \underbrace{\{(1)\}}_{=\Omega_M} \times \underbrace{\{(2)\}}_{=\Omega_W} \end{aligned}$$

Forderung (1)-(3) auf dem fiktivem System von Ereignissen in  $\Omega_X \times \Omega_Y$ , welches mit der Potenzmenge  $\mathcal{P}(\Omega_X \times \Omega_Y)$  übereinstimmt.

## 8.1 Definition Gemeinsame Verteilung

Seien  $X$  und  $Y$  zwei Zufallsvariablen auf einem gemeinsamen diskreten W-Raum  $(\Omega, \mathcal{P}(\Omega), P)$  mit Werten in  $\Omega_X$  und  $\Omega_Y$ . Die *gemeinsame Verteilung* von  $X$  und  $Y$  ist definiert als das W-Maß auf  $\mathcal{P}(\Omega_X \times \Omega_Y)$  welches durch

$$E \in \mathcal{P}(\Omega_X \times \Omega_Y) \mapsto P((X, Y) \in E)$$

gegeben ist. Wir bezeichnen es mit  $P_{(X,Y)}$ . Beachte:

$$P_{(X,Y)}(E) = P((X, Y) \in E) = P(\{\omega \in \Omega : (X(\omega), Y(\omega)) \in E\})$$

**Bemerkung:** Man kann zeigen, dass es eine eindeutige Fortsetzung der Abbildung (Gleichung 5) zu einem W-Maß auf  $\mathcal{P}(\Omega_X \times \Omega_Y)$  gibt und diese dann mit der gemeinsamen Verteilung von  $X$  und  $Y$  übereinstimmt.

## 8.2 Definition Produktmaß

Seien  $Q_1$  und  $Q_2$  zwei diskrete W-Maße jeweils auf der Potenzmenge  $\mathcal{P}(\Omega_1)$  und  $\mathcal{P}(\Omega_2)$ . Dann heißt das W-Maß  $P$  auf  $\mathcal{P}(\Omega_1 \times \Omega_2)$  mit der Eigenschaft:

$$P(E_1 \times E_2) = Q_1(E_1) \cdot Q_2(E_2), \forall E_i \in \mathcal{P}(\Omega_i), i = 1, 2$$

das *Produktmaß* von  $Q_1$  und  $Q_2$  und wird mit  $Q_1 \otimes Q_2$  bezeichnet.

### 8.3 Satz Unabhängige Familie

Seien  $n \in \mathbb{N}$ ,  $n \geq 2$ . Die gemeinsame Verteilung einer Familie  $X_i$ ,  $i = 1, \dots, n$ , von Zufallsvariablen jeweils mit der Verteilung  $P_X$  ist gleich dem Produktmaß

$$\bigotimes_{i=1}^n P_{X_i}$$

genau dann wenn die Familie *unabhängig* ist.

Die Gewichtsfunktion  $\rho$  der gemeinsamen Verteilung  $P_{X_1, \dots, X_n}$  einer unabhängigen Familie ist gleich dem Produkt der Gewichtsfunktion  $\rho_i$  von  $X_i$ ,  $i = 1, \dots, n$ , d.h.

$$\rho(x_1, \dots, x_n) = \prod_{i=1}^n \rho_i(x_i) \quad , \quad \text{für jedes } (x_1, \dots, x_n) \in \Omega_1 \times \dots \times \Omega_n$$

**Beispiel:**

$$\rho(x_1, \dots, x_n) = \frac{1}{6^n} = \frac{1}{6} \cdot \dots \cdot \frac{1}{6} = \rho(x_1) \cdot \dots \cdot \rho(x_n)$$

als Gewichtsfunktion der Verteilung  $P_{(X,Y)}$  wobei

$X$  : Augenzahl bei  $i$  Versuchen

**Annahme:**  $X$  und  $Y$  sind **nicht** unabhängig!

$\hookrightarrow$  Wir benutzen das Konzept der bedingten Wahrscheinlichkeit/W-Maße:

Für jedes  $A \in \mathcal{P}(\Omega_X)$  mit  $P_X(A) > 0$  gilt

$$P_{(X,Y)}(A \times B) = \underbrace{P_X(A)}_{=P_X(A)} \cdot P(Y \in B | X \in A)$$

**“Beweis”:**

$$\text{Linke Seite} = P(X^{-1}(A) \cap Y^{-1}(B))$$

$$\text{Rechte Seite} = P(X^{-1}(A)) \cdot \frac{P(Y^{-1}(B) \cap X^{-1}(A))}{P(X^{-1}(A))}$$

## 9 Das schwache Gesetz der großen Zahlen (GGZ)

**Ziel:** Das Langzeitverhalten von Mittelwerten erfassen

### Bernoulli's GGZ

$\hookrightarrow$  Es beschreibt das Verhalten der Langzeitmittelwerte von Beobachtungswerten, die man bei unabhängig durchgeführten Bernoulli-Experimenten erhält.

**Genauer:** Seien  $p \in (0, 1)$  und  $X \sim B_p$ , d.h.  $X$  ist Bernoulli-verteilt mit Erfolgswahrscheinlichkeit  $p$

$\hookrightarrow$  Beispiel Münzwurf  $\rightarrow$  bei fairer Münze ist  $p = \frac{1}{2}$

Sei  $n \in \mathbb{N}$ . Es werden unabhängig voneinander  $n$  Bernoulli-Experimente durchgeführt, die jeweils die Zufallsvariablen

$$X_i \sim B_p, \quad i = 1, \dots, n$$

beschrieben werden.

$\hookrightarrow$  Wir interessieren uns für die

- $\underbrace{\text{absolute Häufigkeit}}_{\hat{=} \text{ zufällige Anzahl der Treffer}}$  der  $\underbrace{\text{Treffer}}_{\hat{=} X_i=1}$  bei  $n$  Einzelexperimenten.

$$\hookrightarrow S_n = X_1 + \dots + X_n$$

- beziehungsweise die relative Häufigkeit der Treffer  $\hat{=} \frac{1}{n} S_n =: \bar{S}_n$

Wir erinnern uns:  $S_n \sim B_{n,p}$ , d.h.  $S_n$  ist binomialverteilt zu den Parametern  $n$  und  $p$ .

**Ziel:** Das Verhalten der relativen Häufigkeit der Treffer mit wachsender Anzahl  $n$  der Versuche zu erfassen.

$\hookrightarrow$  **Frage:**  $n = 10000$  Versuche mit einer fairen Münze

$\hookrightarrow$  Wie viele Treffer?

$\rightarrow$  **Antwort:** ungefähr  $5000 \hat{=} \frac{1}{2}$  als relative Häufigkeit der Treffer.

$\hookrightarrow$  Variante: mit einer unfairen Münze mit Erfolgswahrscheinlichkeit  $p (\neq \frac{1}{2})$

$\hookrightarrow$  **Frage:** Wie hoch ist die relative Häufigkeit der Treffer?

**Antwort:** ungefähr  $p$

Es ist noch nicht klar, wie wir unsere Antwort formal darstellen können. Beispielsweise führt folgender naheliegender Ansatz nicht zum Ziel.

$$\begin{aligned} P \left( \left\{ \frac{1}{2n} \sum_{i=1}^{2n} X_i = \frac{1}{2} \right\} \right) &= B_{2n, \frac{1}{2}}(\{n\}) \\ &= \binom{2n}{n} \cdot \frac{1}{2^{2n}} \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

## 9.1 Satz Bernoullis GGZ

Sei  $p \in (0, 1)$ . Seien  $X_i \sim B_p$ ,  $i \in \mathbb{N}$  und

$$\bar{S}_n := \frac{1}{n} \sum_{i=1}^n X_i, \quad n \in \mathbb{N}$$

Dann gilt für ein beliebiges  $\epsilon > 0$ :

$$\lim_{n \rightarrow \infty} P(|\bar{S}_n - p| \geq \epsilon) = 0$$

Folgende Resultate gehen in den Beweis von Satz 9.1 ein.

1.

$$\mathbb{E}(\bar{S}_n) = \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \cdot np = p$$

2.

$$\text{Var}(\bar{S}_n) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{p(1-p)}{n}$$

3. Tschebyscheff-Ungleichung

## 9.2 Lemma Rechenregeln für die Varianz

Sei  $X$  eine reelle, diskrete Zufallsvariable mit  $\mathbb{E}(X) = m$  und  $\text{Var}(X) = v$ . Dann gilt für alle  $a, b \in \mathbb{R}$ :

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

**Beweis:**

$$\text{Var}(aX) = \mathbb{E}((aX)^2) - (\mathbb{E}(aX))^2$$

$$\begin{aligned} \text{Linierität des Erwartungswertes (siehe 5.6)} &= a^2 \mathbb{E}(X^2) - (a \mathbb{E}(X))^2 \\ &= a^2 \cdot (\mathbb{E}(X^2) - (\mathbb{E}(X))^2) = a^2 \cdot \text{Var}(X) \end{aligned}$$

zu zeigen:  $\text{Var}(x + b) = \text{Var}(X)$

$$\begin{aligned} \text{Var}(x + b) &= \mathbb{E}(\underbrace{(X + b)^2}_{X^2 + 2bX + b^2}) - \underbrace{(\mathbb{E}(X + b))^2}_{(\mathbb{E}(X) + b)^2} \\ &= \mathbb{E}(X^2) + 2b\mathbb{E}(X) + b^2 - ((\mathbb{E}(X))^2 + 2b\mathbb{E}(X) + b^2) \end{aligned}$$

### 9.3 Lemma Tschebyscheff-Menge

Sei  $X$  eine reelle, diskrete Zufallsvariable mit  $\mathbb{E}(X) = m$ , und der Varianz  $\text{Var}(X) = v$ . Dann gilt für jedes  $\epsilon > 0$ :

$$P(|X - m| \geq \epsilon) \leq \frac{\text{Var}(X)}{\epsilon^2}$$

**Beweis:** Es ist zu zeigen  $\text{Var}(X) \geq \epsilon^2 \cdot P(|X - m| \geq \epsilon)$ . Sei  $\tilde{X} := (X - m)^2$  wobei zu beachten ist, dass  $\tilde{X} \geq 0$

$$\text{Var}(X) = \mathbb{E}(\tilde{X}) = \sum_{\tilde{x} \in \tilde{\Omega}} \tilde{x} \cdot P_{\tilde{X}}(\{\tilde{x}\}) \geq \sum_{\tilde{x} \in M} \tilde{x} P_{\tilde{X}}(\{\tilde{x}\})$$

wobei  $M \subseteq \tilde{\Omega}$ , speziell  $M := \{\tilde{X} \geq \epsilon^2\}$