

# Algorithms and Probability

Week 7

G09 - mkilic

2.IV.2026

# Overview

---

1. Zufallsvariablen
2. Erwartungswert
3. Indikatorvariablen
4. Randomisiertes Quicksort
5. Varianz
6. Wichtige Diskrete Verteilungen
7. Bedingte Zufallsvariablen
8. Mehrere Zufallsvariablen

# Roadmap

---

## 1. Graphentheorie

- Zusammenhang
- Kreise
- Matchings
- Färbungen

## 2. W'keitstheorie

- Bedingte W'keit
- Unabhängigkeit
- (mehrere) Zufallsvariablen
- Diskrete Verteilungen
- Abschätzen von W'keiten
- Randomisierte Algorithmen

## 3. Algorithmen

- Lange-Bunte Pfade
- MaxFlow
- MinCut
- Kleinster umschliessender Kreis
- Konvexe Hülle

# Zufallsvariablen

---

## Zufallsvariable

Eine Zufallsvariable ist eine Abbildung  $X : \Omega \rightarrow \mathbb{R}$ , wobei  $\Omega$  die Ergebnismenge eines Wahrscheinlichkeitsraumes ist.

Wir bezeichnen mit  $W_X$  den Wertebereich einer Zufallsvariable:

$$W_X := X(\Omega) = \{x \in \mathbb{R} \mid \exists \omega \in \Omega \text{ mit } X(\omega) = x\}$$

---

<sup>0</sup>Engl. Zufallsvariable : Random Variable

# Dichte- und Verteilungsfunktion

---

**Dichtefunktion:**

$$f_X : \mathbb{R} \rightarrow [0, 1], x \mapsto \Pr[X = x]$$

**Verteilungsfunktion**

$$F_X : \mathbb{R} \rightarrow [0, 1], x \mapsto \Pr[X \leq x]$$

wobei wir haben

$$\Pr[X \leq y] = \sum_{x \in W_x: x \leq y} \Pr[X = x] = \Pr[\{\omega \in \Omega \mid X(\omega) \leq y\}]$$

# Erwartungswert

---

## Erwartungswert

Zu einer Zufallsvariablen  $X$  definieren wir den *Erwartungswert*  $\mathbb{E}[X]$  durch

$$\mathbb{E}[X] := \sum_{x \in W_X} x \cdot \Pr[X = x],$$

sofern die Summe absolut konvergiert. Ansonsten sagen wir, dass der Erwartungswert undefiniert ist.

## Erwartungswert-2

Ist  $X$  eine Zufallsvariable, so gilt:

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega) \cdot Pr[\omega]$$

# Erwartungswert

## Erwartungswert-3

Ist  $X$  eine Zufallsvariable mit  $W_X \subseteq \mathbb{N}_0$ . Dann gilt:

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \Pr[X \geq i]$$

*Beweis.* Nach Definition gilt

$$\begin{aligned} \mathbb{E}[X] &= \sum_{i=0}^{\infty} i \cdot \Pr[X = i] = \sum_{i=0}^{\infty} \sum_{j=1}^i \Pr[X = i] \\ &= \sum_{j=1}^{\infty} \sum_{i=j}^{\infty} \Pr[X = i] = \sum_{j=1}^{\infty} \Pr[X \geq j]. \end{aligned}$$

# Linearität des Erwartungswerts

---

*Der Erwartungswert einer Summe von Zufallsvariablen ist die Summe der Erwartungswerte der Zufallsvariablen.*

# Linearität des Erwartungswerts

---

## Linearität des Erwartungswerts

**Satz 2.33.** (*Linearität des Erwartungswerts*) Für Zufallsvariablen  $X_1, \dots, X_n$  und  $X := a_1X_1 + \dots + a_nX_n + b$  mit  $a_1, \dots, a_n, b \in \mathbb{R}$  gilt

$$\mathbb{E}[X] = a_1\mathbb{E}[X_1] + \dots + a_n\mathbb{E}[X_n] + b.$$

# Indikatorvariablen

---

## Indikatorvariablen

Für ein Ereignis  $A \subseteq \Omega$  ist die zugehörige Indikatorvariable  $X_A$  definiert durch:

$$X_A(\omega) := \begin{cases} 1, & \text{falls } \omega \in A \\ 0, & \text{sonst} \end{cases}$$

Für den Erwartungswert von  $X_A$  gilt:  $\mathbb{E}[X_A] = Pr[A]$

## Indikatorvariablen - II

---

**Indikatorvariable**  $I_A(\omega) := \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$

- Schnitt von zwei Indikatorvariablen  $I_A, I_B$ :

$$I_{A \cap B} = I_A \cdot I_B$$

*Das Produkt von Indikatorvariablen ist ebenfalls eine Indikatorvariable, nämlich der Indikator für den Schnitt der beteiligten Ereignisse.*

## Indikatorvariablen - II

---

**Indikatorvariable**  $I_A(\omega) := \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$

- Schnitt von zwei Indikatorvariablen  $I_A, I_B$ :

$$I_{A \cap B} = I_A \cdot I_B$$

*Das Produkt von Indikatorvariablen ist ebenfalls eine Indikatorvariable, nämlich der Indikator für den Schnitt der beteiligten Ereignisse.*

- Komplement von einer Indikatorvariable  $I_A$ :

$$I_{\bar{A}} = 1 - I_A$$

## Indikatorvariablen - II

---

**Indikatorvariable**  $I_A(\omega) := \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$

- Schnitt von zwei Indikatorvariablen  $I_A, I_B$ :

$$I_{A \cap B} = I_A \cdot I_B$$

*Das Produkt von Indikatorvariablen ist ebenfalls eine Indikatorvariable, nämlich der Indikator für den Schnitt der beteiligten Ereignisse.*

- Komplement von einer Indikatorvariable  $I_A$ :

$$I_{\bar{A}} = 1 - I_A$$

- Vereinigung von zwei Indikatorvariablen  $I_A$  und  $I_B$ :

$$I_{A \cup B} \rightarrow \text{Siebformel}$$

# Indikatorvariablen - Beweis der Siebformel

## Siebformel: Beweis

Wir wollen die W'keit von Vereinigung allgemeiner Ereignissen berechnen. Sei

$$B := A_1 \cup \dots \cup A_n$$

Weiter seien  $I_i = I_{A_i}$  und  $I_B$  die zugehörigen Indikatorvariablen. Beobachte

$$\begin{aligned}\bar{B} &= \overline{A_1 \cup \dots \cup A_n} = \bar{A}_1 \cap \dots \cap \bar{A}_n \\ I_{\bar{B}} &= \prod_{i=1}^n I_{\bar{A}_i} = \prod_{i=1}^n (1 - I_{A_i}) \\ &= 1 - \sum_i I_{A_i} + \sum_{i < j} I_{A_i} I_{A_j} - \sum_{i < j < k} I_{A_i} I_{A_j} I_{A_k} + \dots + (-1)^n I_{A_1} I_{A_2} \dots I_{A_n}\end{aligned}$$

# Indikatorvariablen - Beweis der Siebformel - II

## Siebformel: Beweis

Linearität des Erwartungswertes und  $\mathbb{E}[I_A] = \Pr[A]$  besagt:

$$\mathbb{E}[I_{\bar{B}}] = 1 - \sum_i \mathbb{E}[I_{A_i}] + \sum_{i < j} \mathbb{E}[I_{A_i} I_{A_j}] - \sum_{i < j < k} \mathbb{E}[I_{A_i} I_{A_j} I_{A_k}] + \cdots + (-1)^n \mathbb{E}[I_{A_1} I_{A_2} \cdots I_{A_n}]$$

also gilt

$$1 - \Pr[B] =$$

$$1 - \sum_i \Pr[A_i] + \sum_{i < j} \Pr[A_i \cap A_j] - \sum_{i < j < k} \Pr[A_i \cap A_j \cap A_k] + \cdots + (-1)^n \Pr\left[\bigcap_{i=1}^n A_i\right]$$

Das ist genau die Siebformel!

# Randomisiertes Quicksort

---

```
1  if < r then
2      p ← Uniform({l, l + 1, ..., r})
3      t ← PARTITION(A, l, r, p)
4      QUICKSORT(A, l, t - 1)          \\Block I
5      QUICKSORT(A, t + 1, r)        \\Block II
6
```

Quicksort

Wir sind interessiert für Laufzeit. Sie ist von der Anzahl der Vergleiche bestimmt.

# Randomisiertes Quicksort

---

Definiere:  $X_{i,j} :=$  Indikatorvariable für " $a_i$  wird mit  $a_j$  verglichen"

- Pivotelement wird mit allen anderen Elementen im selben Block verglichen. Sonst gibt es keine Vergleiche und das Pivotelement ist danach in keinem Block mehr.
- Damit das Paar  $(a_i, a_j)$  verglichen werden können, muss einer der beiden als Pivotelement gewählt werden. Jedes Paar wird maximal einmal verglichen.

# Randomisiertes Quicksort

---

Definiere:  $X_{i,j} :=$  Indikatorvariable für "  $a_i$  wird mit  $a_j$  verglichen "

- Pivotelement wird mit allen anderen Elementen im selben Block verglichen. Sonst gibt es keine Vergleiche und das Pivotelement ist danach in keinem Block mehr.
- Damit das Paar  $(a_i, a_j)$  verglichen werden können, muss einer der beiden als Pivotelement gewählt werden. Jedes Paar wird maximal einmal verglichen.

$$X = \sum_{1 \leq i < j \leq n} X_{i,j}$$

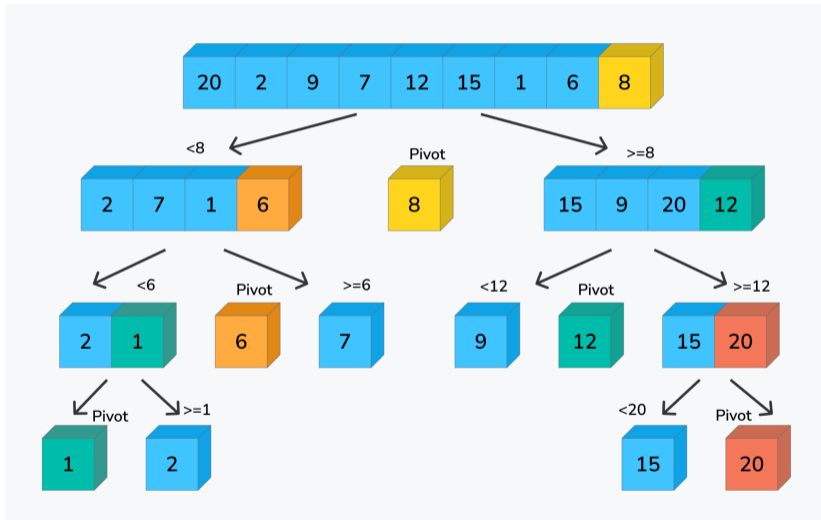
Für ein Paar  $(a_i, a_j)$  gibt es zwei Möglichkeiten:

1.  $a_i$  oder  $a_j$  wird als Pivot gewählt:  $X_{i,j} = 1$  mit W'keit  $\frac{2}{j-i+1}$
2. Einer von  $a_{i+1}, \dots, a_{j-1}$  wird gewählt:  $X_{i,j} = 0$

Also haben wir:  $\mathbb{E}[X_{i,j}] = Pr[X_{i,j} = 1] = \frac{2}{j-i+1}$

$$\mathbb{E}[X] \stackrel{\text{lin.}}{=} \sum_{1 \leq i < j \leq n} \mathbb{E}[X_{i,j}] = \sum_{1 \leq i < j \leq n} \frac{2}{j-i+1} \dots \leq 2n \ln(n)$$

# Randomisiertes Quicksort



# Varianz

---

## Varianz

Für eine Zufallsvariable  $X$  mit  $\mu = \mathbb{E}[X]$  definieren wir die *Varianz*  $\text{Var}[X]$  durch

$$\text{Var}[X] := \mathbb{E}[(X - \mu)^2] = \sum_{x \in W_x} (x - \mu)^2 \cdot \text{Pr}[X = x]$$

Die Grösse  $\sigma := \sqrt{\text{Var}[X]}$  heisst *Standardabweichung* von  $X$

Wir sind eigentlich interessiert für die erwartete Abweichung vom Erwartungswert.

# Varianz

---

## Varianz und Erwartungswert

Für eine beliebige Zufallsvariable  $X$  gilt

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

**Beweis:**

$$\begin{aligned}\mathbb{E}[(X - \mu)^2] &= \mathbb{E}[X^2 - 2\mu X + \mu^2] && \text{(Linearität des } \mathbb{E} \text{)} \\ &= \mathbb{E}[X^2] - 2\mu\mathbb{E}[X] + \mu^2 && (\mu = \mathbb{E}[X]) \\ &= \mathbb{E}[X^2] - \mathbb{E}[X]^2\end{aligned}$$



## Varianz II

---

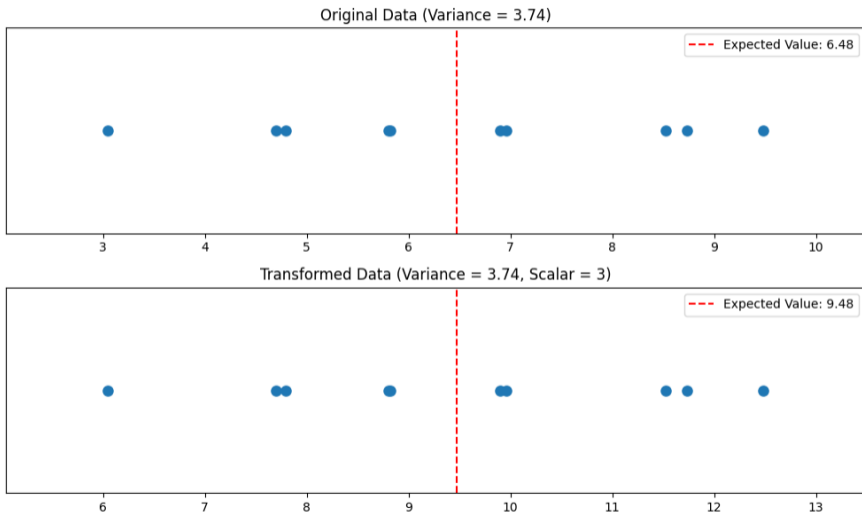
Satz 2.41.

Für eine beliebige Zufallsvariable  $X$  und  $a, b \in \mathbb{R}$  gilt

$$\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X]$$

Folgt aus beiden vorherigen Folien

# Invariance of the variance under scalar addition



# Variance Basics

---

- Variance of a random variable  $X$ :

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

- For independent random variables  $X$  and  $Y$ :
  - $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$
  - $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$  (Das gilt auch für unabhängige ZV )
  - $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$
  - $\text{Var}[X \cdot Y] = \mathbb{E}[X^2]\mathbb{E}[Y^2] - (\mathbb{E}[X])^2(\mathbb{E}[Y])^2$
- If  $\mathbb{E}[Y] = 0$ , then:

$$\text{Var}[X \cdot Y] = \mathbb{E}[X^2] \cdot \text{Var}[Y]$$

## Exercise: Intransitive Dice

---

Alice, Bob, and Clara just found three fair dice. Instead of the numbers 1 to 6, the six sides of the dice show the following numbers:

**Die 1:** 2, 2, 4, 4, 9, 9

**Die 2:** 1, 1, 6, 6, 8, 8

**Die 3:** 3, 3, 5, 5, 7, 7

Bob and Clara want to play a game: Bob starts by choosing a die. Next, Clara chooses a die. Then, both of them roll the dice and the person whose die shows the higher number wins.

Alice computes the expected value of the outcome of each die. She claims that their choices do not matter, because all dice have the same expected value.

Bob additionally computes the variance for the outcome of each die. He claims that the first two dice have a higher variance and are thus better than the third die. Thus, he chooses die 1.

Clara chooses die 3 and claims that her probability of winning was  $\frac{5}{9}$  all along — no matter which die Bob chose.

Compute the expected value and variance for each die. What do you think about the claims of Alice, Bob, and Clara?

# Bernoulli-Verteilung

---

## Bernoulli-Verteilte Zufallsvariablen

Eine Zufallsvariable  $X$  mit  $W_X = \{0, 1\}$  und der Dichte

$$f_X(x) = \begin{cases} p & \text{für } x = 1, \\ 1 - p & \text{für } x = 0, \\ 0 & \text{sonst} \end{cases}$$

heißt bernoulli-verteilt wobei  $p =$  Erfolgswahrscheinlichkeit.

Man notiert:  $X \sim \text{Bernoulli}(p)$

- $\mathbb{E}[X] = p$
- $\text{Var}[X] = p(1 - p)$

# Binomialverteilung

---

## Binomialverteilung

Eine Zufallsvariable  $X$  mit  $W_X = \{0, 1, \dots, n\}$  und der Dichte

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x \in \{0, 1, \dots, n\} \\ 0, & \text{sonst.} \end{cases}$$

heißt binomialverteilt. Man schreibt  $X \sim \text{Bin}(n, p)$

- $\mathbb{E}[X] = np$
- $\text{Var}[X] = np(1-p)$

*mehrere Bernoulli-verteilte Variablen*

# Geometrische Verteilung

*so lange wiederholen bis man erfolgreich ist*

## Geometrische Verteilung

Wenn ein einzelner Versuch mit Wahrscheinlichkeit  $p$  gelingt, so ist die Anzahl der Versuche bis zum Erfolg geometrisch verteilt:  $X \sim \text{Geo}(p)$  Für die Dichte und Verteilung einer geometrisch verteilten Zufallsvariable gilt:

$$f_X(i) = \begin{cases} p(1-p)^{i-1} & \text{für } i \in \mathbb{N}, \\ 0 & \text{sonst.} \end{cases}$$

$$\text{(Vert)} \quad F_X(n) = \Pr[X \leq n] = \sum_{i=1}^n \Pr[X = i] = \sum_{i=1}^n p(1-p)^{i-1} = 1 - (1-p)^n.$$

- $\mathbb{E}[X] = \frac{1}{p}$
- $\text{Var}[X] = \frac{1-p}{p^2}$

# Gedächtnislosigkeit

---

” Die Wahrscheinlichkeit, dass wir gleich im ersten Versuch ‘Kopf’ sehen, ist identisch zu der Wahrscheinlichkeit, dass wir nach 1000 Fehlversuchen beim 1001ten Wurf ‘Kopf’ erhalten. Die Tatsache, dass wir die Münze schon lange (aber erfolglos) geworfen haben, hat also keinerlei Einfluss auf den Erfolg im nächsten Wurf.”

# Gedächtnislosigkeit - Formal

---

## Gedächtnislosigkeit

Ist  $X \sim \text{Geo}(p)$ , so gilt für alle  $s, t \in \mathbb{N}$ :

$$\Pr[X \geq s + t \mid X > s] = \Pr[X \geq t]$$

**Beweis:**  $F_X(n) = 1 - (1 - p)^n, n \in \mathbb{N}$ . Dann gilt  $\Pr[X \geq n] = (1 - p)^n$ . Dann mithilfe der Def. von bedingten W'keit folgt:

$$\Pr[X \geq s + t \mid X > s] = \frac{\Pr[X \geq s + t]}{\Pr[X > s]} = \frac{(1 - p)^{s+t-1}}{(1 - p)^s} = (1 - p)^{t-1} = \Pr[X \geq t],$$

---

<sup>0</sup>Engl. *Memorylessness*

## Warten auf den $n$ -ten Erfolg

---

$Z$  bezeichnet die Anzahl der Versuche bis zum  $n$ -ten erfolgreichen Experiment. Wenn  $Z = z$  ist, so wurden also genau  $n$  erfolgreiche und  $z - n$  nicht erfolgreiche Experimente durchgeführt. Da nach Definition von  $Z$  das letzte Experiment erfolgreich sein muss, steht der Zeitpunkt des  $n$ -ten erfolgreichen Experimentes fest. Die übrigen  $n - 1$  erfolgreichen Experimente können beliebig auf die restlichen  $z - 1$  Experimente verteilt werden.

# Warten auf den n-ten Erfolg

---

## Warten auf den n-ten Erfolg

Falls  $Z$  **negativ binomialverteilt** für die Dichte gilt:

$$f_Z(z) = \binom{z-1}{n-1} \cdot p^n (1-p)^{z-n}.$$

- $\mathbb{E}[Z] = \frac{n}{p}$  (*Linearität des Erwartungswertes*)

# Coupon-Collector-Problem: Bio-Twist

---

## Problem: Coupon Collector - Bio Twist

There is a new virus that gets a random mutation every time a new person is infected. We know that there exists only  $n$  many possible mutations. How many patients (including patient 0) must get infected on average such that we can see all the mutations and develop medication for them?

# Coupon-Collector-Problem: Bio-Twist - II

## Coupon Collector - Bio Twist

Teile den ganzen Prozess in Phasen. In der Phase  $i$ , hatte das Virus schon  $i - 1$  Mutationen.

- $X :=$  Anzahl Infektionen insgesamt
- Sei  $X_i :=$  Anzahl Infektionen in der Phase  $i$

Also gilt  $X = \sum_{i=1}^n X_i$  Wir wissen auch, dass  $X_i \sim \text{Geo}(p_i)$  Wir wollen einer der  $n - i + 1$  Mutationen sehen, die noch nicht aufgetaucht sind. Somit erhalten wir

$$p_i = \frac{n - i + 1}{n} \text{ und } \mathbb{E}[X_i] = \frac{n}{n - i + 1}$$

$$\mathbb{E}[X] \stackrel{\text{Lin. des E}}{=} \sum_{i=1}^n \mathbb{E}[X_i] = \sum_{i=1}^n \frac{n}{n - i + 1} = n \cdot \sum_{i=1}^n \frac{1}{i} = n \cdot H_n,$$

## Coupon-Collector-Problem: Bio-Twist - III

---

Wir wissen  $H_n := \sum_{i=1}^n \frac{1}{i} = \ln n + O(1)$  also folgt

$$\mathbb{E}[X] = n \ln n + O(n)$$

**Fazit:**  $n \ln n + O(n)$  Menschen im Durchschnitt müssen infiziert werden damit wir alle Mutationen beobachten können.

# Poisson-Verteilung

---

## Poisson-Verteilung

Poisson-Verteilung mit Parameter  $\lambda$  ist definiert durch die Dichtefunktion

$$f_X(i) = \begin{cases} \frac{e^{-\lambda} \lambda^i}{i!} & \text{für } i \in \mathbb{N}_0 \\ 0 & \text{sonst.} \end{cases}$$

Es gilt:

$$\mathbb{E}[X] = \text{Var}[X] = \lambda$$

Man kann die Poisson-Verteilung als Grenzwert einer Binomialverteilung  $\text{Bin}(n, \frac{\lambda}{n})$  erhalten.

# Unabhängigkeit von Zufallsvariablen

---

## Unabhängigkeit von Zufallsvariablen

Die Zufallsvariablen  $X_1, \dots, X_n$  heißen unabhängig, genau dann wenn für alle  $(x_1, \dots, x_n) \in W_{X_1} \times \dots \times W_{X_n}$  wir haben:

$$\Pr[X_1 = x_1, \dots, X_n = x_n] = \Pr[X_1 = x_1] \cdots \Pr[X_n = x_n].$$

**Alternativ:**

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$$

für alle  $(x_1, \dots, x_n) \in W_{X_1} \times \dots \times W_{X_n}$

# Bedingte Zufallsvariable

## Bedingte Zufallsvariable

Sei  $X$  Zufallsvariable auf  $W$ -raum  $\Omega$ ,  $A \subseteq \Omega$  Ereignis mit  $\Pr[A] > 0$ . Die bedingte Zufallsvariable  $X|A$  ist dieselbe Funktion wie  $X$ , aber der Definitionsbereich ist auf die Menge  $A$  eingeschränkt.

$$f_{X|A} : \mathbb{R} \rightarrow [0, 1], \quad x \mapsto \Pr[X = x | A]$$

$X$  ist unabhängig von  $A$ , falls  $f_{X|A} = f_X$ .

$$F_{X|A} : \mathbb{R} \rightarrow [0, 1], \quad x \mapsto \Pr[X \leq x | A]$$

$$\mathbb{E}[X | A] := \sum_{x \in W_X} x \cdot \Pr[X = x | A] = \frac{1}{\Pr[A]} \sum_{\omega \in A} X(\omega) \cdot \Pr[\omega]$$

**Erinnerung:**  $X = x$  ist auch ein Ereignis, also " $X = x \subseteq \Omega$ "

# Gemeinsame Dichten und Randdichten

---

## Gemeinsame Dichte

$$f_{X,Y}(x, y) = \Pr[X = x, Y = y]$$

## "Randdichte von X" (= Dichte von X)

$$f_X(x) = \sum_{y \in W_Y} f_{X,Y}(x, y) \quad (\text{nach dem Additionssatz})$$

The End

