MATHEMATISCHE GRUNDLAGEN

STOCHASTIK

a.) Ereignisse und Wahrscheinlichkeiten

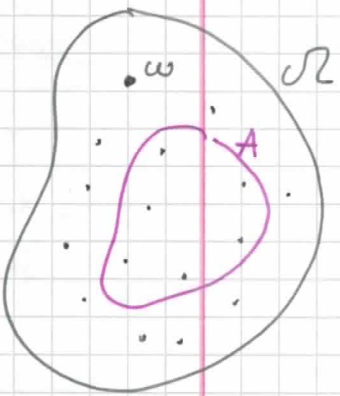
Def: Sei Ω eine endliche oder abzählbare Menge (der Ergebnisraum). Eine Wahrscheinlichkeitsverteilung über Ω ist eine Funktion

→ was wird hier verteilt?

$$p : \Omega \to [0,1], \text{ so dass } \sum_{\omega \in \Omega}' p(\omega) = 1$$

ein $\omega \in \Omega$ heißt Elementarereignis

$A \subseteq \Omega$ heißt Ereignis $\quad p(A) = \sum_{\omega \in A} p(\omega)$



- die ω sind häufig die Ausgänge eines sog. Zufallsexperiments, deshalb heißt Ω Ergebnisraum
- die Wahrscheinlichkeiten von Ereignissen, die $p(A)$ verhalten sich wie die Größen von Flächen → man darf an $p(\;)$ als an Fläche(größe) denken, wobei die Gesamtfläche 1 ist.

Beispiel 1. Wir spielen Roulette (eine Runde), die Elementarereignisse seien die Farben

$p(\text{schwarz}) = \frac{18}{37}$

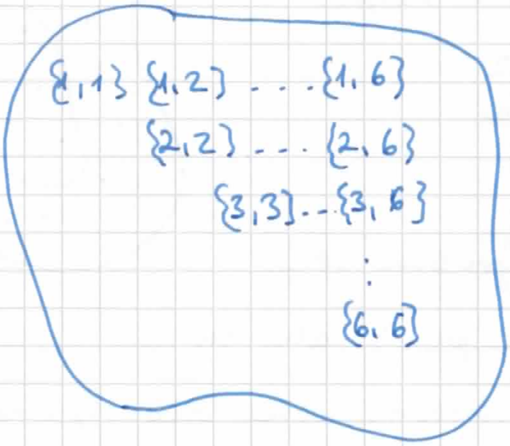
$p(\text{rot}) = \frac{18}{37}$

$p(\text{grün}) = \frac{1}{37}$

- Beachte, dass die Elementarereignisse nicht unbedingt gleichwahrscheinlich sind! Falls sie gleichwahrscheinlich sind, haben wir eine <u>Gleichverteilung</u> oder <u>uniforme</u> Verteilung.

- Zum selben Experiment kann man verschiedenste Ereignisräume definieren. Bei einer Runde Roulette, könnten die Elementarereignisse z.B. auch die Zahlen 0, 1, 2, ... , 36 sein, dann hätten wir sogar (in unserem idealisierten Modell) eine Gleichverteilung.

<u>Beispiel 2</u>: Wir würfeln mit zwei Würfeln (einem Roten und einem Blauen). Die Elementarereignisse seien diesmal die <u>ungeordnete</u> Zahlenpaare.



$$P(\{i,i\}) = \frac{1}{36}$$

$$\underset{i\neq j}{P(\{i,j\})} = \frac{2}{36}$$

Welche der Beiden ist <u>kein</u> Ereignis über diesem Ereignisraum?

A: die Summe der gewürfelten Zahlen ist 10

$\left(P(A) = \frac{3}{36}\right)$

B: die mit dem blauen Würfel gewürfelte Zahl ist 4

keine Teilmenge von Ω ⟶ kein Ereignis

b.) Unabhängige Ereignisse

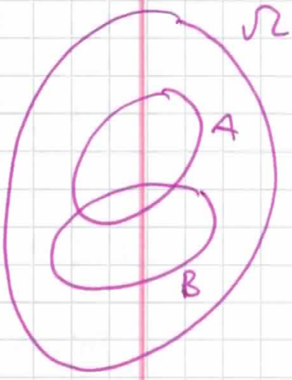
Seien A und B zwei Ereignisse $A, B \subset \Omega$

Was für Ereignisse sind $A \cup B$ und $A \cap B$, und was wissen wir über ihre Wahrscheinlichkeiten

$$P(A \cup B) \text{ und } P(A \cap B)?$$

$\|$

Andere Notation: $P(A \cdot B)$



Wann gilt $P(A \cup B) = P(A) + P(B)$?

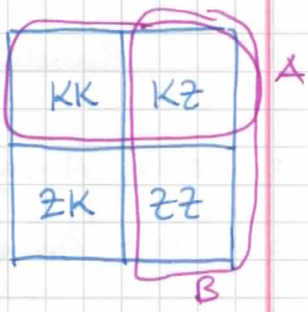
Nur wenn A und B sich ausschließende Ereignisse sind.

Beispiele

1. Wir werfen eine Münze 2mal

A: der erste ist Kopf

B: der zweite ist Zahl



$$P(A \cup B) = \frac{3}{4}$$

$A \cup B$: der erste ist Kopf oder der zweite ist Zahl

$A \cdot B$: der erste ist Kopf und der zweite ist Zahl

$$P(A \cdot B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$$

Es gilt immer: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
(wie bei Flächen oder Mengen)

Gilt es immer, dass $P(A \cap B) = P(A) \cdot P(B)$? NEIN!

zB. $P(A \cap \bar{A}) = 0$

Nur wenn A und B sog. unabhängige Ereignisse sind (zB. die Ausgänge von unabhängigen Experimenten sind unabhängige Ereignisse.) Aber wie können wir Unabhängigkeit mathematisch definieren?

ST 4.

Definition: Die Ereignisse A und B heißen sind genau dann **unabhängig**, wenn $P(A \cdot B) = P(A) \cdot P(B)$.

2\. Wir werfen eine Münze 2mal

A: die beiden sind gleich

B: der zweite ist eine Zahl

$P(A \cup B) = \frac{3}{4}$

Sind A und B unabhängig?

$P(A) = \frac{1}{2}$ $P(B) = \frac{1}{2}$ $P(A \cdot B) = \frac{1}{4}$ → JA!

3\. eine Runde Roulette

R: der Ausgang ist rot

U: der Ausgang ist ungerade

Sind R und U unabhängig?

$P(R) = \frac{18}{37}$ $P(U) = \frac{18}{37}$ $P(R \cdot U) = \frac{9}{37}$

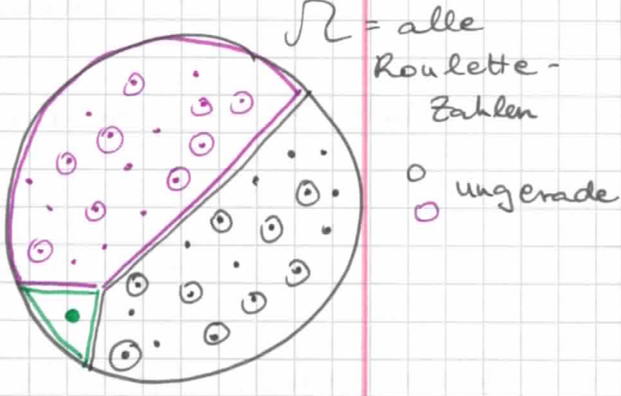
↓ weniger als die Hälfte von allen ist ungerade

↓ die Hälfte der roten ist ungerade

$P(R) \cdot P(U) < P(R \cdot U)$ → NEIN!

c.) Bedingte Wahrscheinlichkeit

1. eine Runde Roulette



Wir legen auf Ungerade nach dem Drehen sehe ich ohne Brille nur dass es rot ist. Haben sich meine Chancen ein wenig verbessert?

JA! von $\frac{18}{37} < \frac{1}{2}$ auf $\frac{9}{18} = \frac{1}{2}$

In dem Moment wenn ich sehe dass der Ausgang rot ist, hat sich für mich der Ereignisraum auf den roten Teil beschränkt.

Die Ereignisse haben für mich jetzt neue Wahrscheinlichkeiten: früher $P(U) = \frac{18}{37}$; jetzt $P(U|R) = \frac{9}{18}$

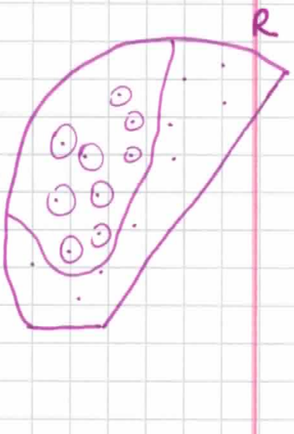
$P(U|R)$ = die Wahrscheinlichkeit von Ungeradem Ausgang angenommen (unter der Bedingung) dass der Ausgang rot ist.

Definition: $P(U|R)$ heißt die bedingte Wahrscheinlichkeit von Ereignis U bezüglich Ereignis R

Wie berechnet man die bedingte Wahrscheinlichkeit?

mein neuer Ereignisraum ist die Menge (das Ereignis) R

es muss zB. $P(R|R) = 1$ gelten



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

ST 6.

2. Wir werfen eine Münze 2mal.

Wie hoch ist die bedingte Wahrscheinlichkeit

$$P(\text{zweite ist Kopf} \mid \text{erste ist Zahl})\,?$$

Intuitiv: $= P(\text{zweite ist Kopf}) = \frac{1}{2}$, weil die beiden Würfe „nichts miteinander zu tun haben", sie sind unabhängig.

Behauptung: $P(A|B) = P(A)$ genau dann, wenn A und B unabhängige Ereignisse sind.

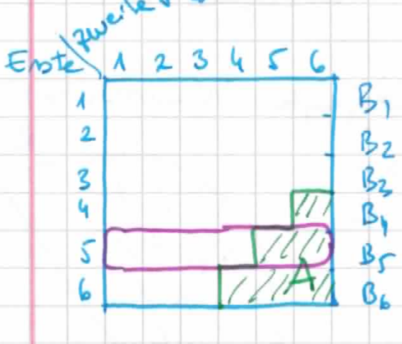
Warum? $$P(A|B) \underset{\substack{\downarrow \\ \text{Def.}}}{=} \frac{P(A \cdot B)}{P(B)} \underset{\substack{\Updownarrow \\ \text{unabhängig}}}{=} \frac{P(A)\cdot P(B)}{P(B)} = P(A)$$

3. Wir würfeln mit einem Würfel 2mal hintereinander. Wir gewinnen wenn die Summe der Würfe mindestens 10 ist.

$$P(\geqslant 10) = \frac{6}{36} = \frac{1}{6} \qquad (\text{warum?})$$

Der erste Wurf wird eine 5. Haben sich nach diesem Wurf unsere Chancen verbessert?

Ergebnisraum:



A: sei das Ereignis, dass die Summe ≥ 10 ist

nach dem ersten Wurf hat sich der Ereignisraum auf die Zeile 5 eingeschränkt

$$P(\geq 10 \mid \text{Erster} = 5) = \frac{1}{3} = \left(\frac{P(\geq 10 \text{ und Erster} = 5)}{P(\text{Erster} = 5)} = \frac{\frac{2}{36}}{\frac{6}{36}} \right)$$

$$P(\geq 10 \mid \text{Erster} = 6) = \frac{1}{2}$$

$$P(\geq 10 \mid \text{Erster} = 4) = \frac{1}{6}$$

$$P(\geq 10 \mid \text{Erster} = 3) = 0$$

---

d.) Die Formel für die totale Wahrscheinlichkeit

Wir hätten $P(\geq 10)$ auch so berechnen können:

$$\begin{aligned} P(\geq 10) = {} & P(\text{Erster} = 4) \cdot P(\geq 10 \mid \text{Erster} = 4) + \\ & + P(\text{Erster} = 5) \cdot P(\geq 10 \mid \text{Erster} = 5) + \\ & + P(\text{Erster} = 6) \cdot P(\geq 10 \mid \text{Erster} = 6) \\ & \left(+ P(\text{Erster} = 3) \cdot \underset{0}{\underset{\parallel}{P(\geq 10 \mid \text{Erster} = 3)}} + \underset{0}{\underset{\parallel}{\ldots}} \right) \end{aligned}$$

$$= \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{2}{6} + \frac{1}{6} \cdot \frac{3}{6} = \frac{1}{6}$$

Warum stimmt das, und wie kann man diese Regel allgemein formulieren?

Sei B_i : das Ereignis, dass Erster $= i$

A : die Summe ≥ 10

$B_1 \; B_2 \; B_3 \; B_4 \; B_5 \; B_6$ bilden ein sog.

vollständiges Ereignissystem, weil $\bigcup B_i = \Omega$ und

$B_i \cap B_j = \emptyset \quad \forall i \neq j$

ST8 .

Es gilt $A = (A\cdot B_1)\cup(A\cdot B_2)\cup(A\cdot B_3)\cup\underbrace{(A\cdot B_4)}_{\text{Erster}=4 \text{ und Summe} \geq 10}\cup(A\cdot B_5)\cup(A\cdot B_6)$

$$P(A) = P(A\cdot B_1)+P(A\cdot B_2)+P(A\cdot B_3)+P(A\cdot B_4)+P(A\cdot B_5)+$$
$$+P(A\cdot B_6)$$

Hier dürfen wir die Wahrscheinlichkeiten summieren, weil wir sich ausschließende Ereignisse haben. (der Erste Wurf kann nicht gleichzeitig 4 und 5 sein)

Aber immer gilt, dass

$$P(A B_i) = P(B_i)\cdot P(A|B_i)$$

deshalb:

$$\boxed{P(A) = P(B_1)\cdot P(A|B_1) + \ldots + P(B_4)\cdot P(A|B_4) + P(B_5)\cdot P(A|B_5) + P(B_6)\cdot P(A|B_6)}$$

Diese war unsere Ausgangsformel. Man nennt sie die Formel für die totale Wahrscheinlichkeit.

(Wir haben diese Formel schon beim Sekretärproblem verwendet:

$$p = \sum_{i=r+1}^{n} \underbrace{\frac{1}{n}}_{P(b=i)} \cdot \underbrace{\frac{r}{i-1}}_{P(\text{Erfolg}\mid b=i)}$$

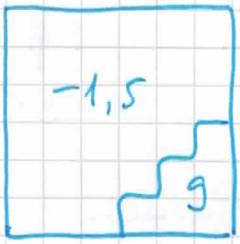
für $i = 1, 2, \ldots n$ ein vollständiges Ereignissystem)

e.) Zufallsvariablen

Nehmen wir an, dass wir 9 € gewinnen, falls die Summe der Würfe ≥ 10, aber 1,5 € bezahlen sollen, falls die Summe < 10. Sollen wir an so einem Spiel teilnehmen?

Was uns jetzt interessiert ist zB. ~~der~~ unser erwarteter Gewinn ((besonders wenn das Spiel oft wiederholt wird)) Wie modellieren wir diese Situation?

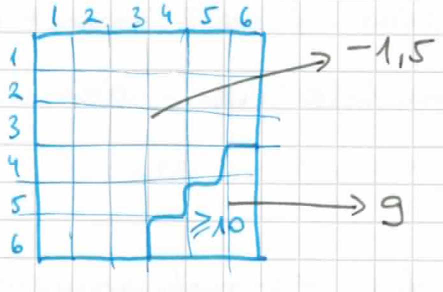
(Wir könnten uns einen neuen Ereignisraum ausdenken mit zwei Elementarereignissen



$P(9) = \frac{1}{6}$ $P(-1{,}5) = \frac{5}{6}$

und dann den Erwartungswert (definieren und) berechnen.

Was wir formal tun, ist fast dasselbe.)



Wir weisen jedem Elementarereignis den entsprechenden Gewinn zu, somit also haben wir jedem Elementarereignis eine reelle Zahl zugewiesen. Gemäß der W.-Verteilung auf Ω, diese Zahl ist 9 mit Wahrscheinlichkeit $\frac{1}{6}$, und $-1{,}5$ mit Wahrscheinlichkeit $\frac{5}{6}$, diese Zahl (diese Zuweisung) ist eine Zufallsvariable.

Man kann zu demselben Ereignisraum verschiedenste Zufallsvariablen definieren.

Definition: Eine **Zufallsvariable** ist eine Funktion

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

- wir haben natürlicherweise eine Wahrscheinlichkeitsverteilung über die reellen Werte von X erhalten (aus den Wahrscheinlichkeiten $p(\omega)$)
- die möglichen Werte von X, (und die zugehörigen Wahrscheinlichkeiten) kann man auf der Zahlengerade (im Koordinatensystem) darstellen.

Wir haben aus den Elementarereignissen **Zahlen** gemacht. Jetzt spielen die Werte von X die Rolle der Elementarereignisse. (Wenn die Elementarereignisse selber reelle Zahlen sind, das definiert „automatisch" eine Zufallsvariable, aber wir können beliebige andere X über diesen Ereignisraum definieren.)

Unser **erwarteter Gewinn** ist der Erwartungswert dieser Zufallsvariable X:

$$E[X] = \frac{1}{6}\cdot 9 + \frac{5}{6}(-1,5) = \frac{3}{2} - \frac{5}{6}\cdot\frac{3}{2} = \frac{1}{6}\cdot\frac{3}{2} = \frac{1}{4} > 0$$

⇒ Zumindest wenn das Spiel mehrmals wiederholt wird, spielen wir mit. In welchem Sinne modelliert $E[X]$ einen durchschnittlichen Wert?

f.) Der Erwartungswert

Definition: Der Erwartungswert der Zufallsvariable X ist

$$E[X] = \sum_{r\in\mathbb{R}} r\cdot P(X=r) = \sum_{\omega\in\Omega} X(\omega)\cdot p(\omega)$$

→ da Ω (endlich oder) abzählbar ist, hat die rechte Summe abzählbar viele Summanden wie $\sum_{i=1}^{n}$ $\sum_{i=1}^{\infty}$

die linke Summe hat abzählbar viele r Werte, so dass $P(X=r)>0$ (nur die $X(\omega)$ Werte).

→ Vorsicht! X ist kein Ereignis, $P(\cancel{X})$ hat keinen Sinn!

Ereignisse haben jetzt z.B. diese Form:

$P(X=r)$
$P(X\geq a)$
$P(a\leq X\leq b)$
$P(X=3$ oder $X>7)$ usw.

→ Nur Zufallsvariablen haben Erwartungswert (nur Zahlen) Ereignisse wie "rot", "König" bzw. deren Wahrscheinlichkeitsverteilungen, natürlich nicht.

Kann man aus dem Erwartungswert irgend etwas über die Wahrscheinlichkeit eines einzigen Wertes folgern?

Beispiel: Ein randomisierter Algorithmus hat für eine fixierte Eingabe w die erwartete Laufzeit „100 Schritte". $E[L] = 100$

Wie hoch kann die Wahrscheinlichkeit höchstens sein, dass der Algorithmus (eine Berechnung) mindestens 1000 Schritte braucht?

$$P(L \geq 1000)\ ?$$

{ Hier sind die möglichen Bitfolgen 0110010... die der Algorithmus erhält, die Elementarereignisse (wir können gleichlange Bitfolgen betrachten – wenn es eine Maximumlänge gibt –, oder jede Bitfolge bis zum letzten angeforderten Bit bevor der Algorithmus / die Berechnung hält); die Laufzeit/Schrittzahl ist eine Zufallsvariable L mit den möglichen Werten 0, 1, 2, 3, ..., n, ... Für eine Antwort auf die Frage ist entscheidend, dass alle möglichen Werte der Laufzeit, nichtnegativ sind!]

→ die gefragte Wahrscheinlichkeit

$P(L \geq 100)$

$P(L \geq 1000) = \text{Prob}(\text{Alg braucht} \geq 1000 \text{ Schritte})$ ist höchstens $\frac{1}{10}$

$P(L \geq 1000)$

Warum? Nehmen wir an z.B. $\text{Prob}(\geq 1000 \text{ Schritte}) = \frac{1}{5}$

Dann hätten wir erwartete Laufzeit mindestens

$$\geq 1000 \cdot \frac{1}{5} = 200 > 100 \ ↯. \text{ Widerspruch!}$$

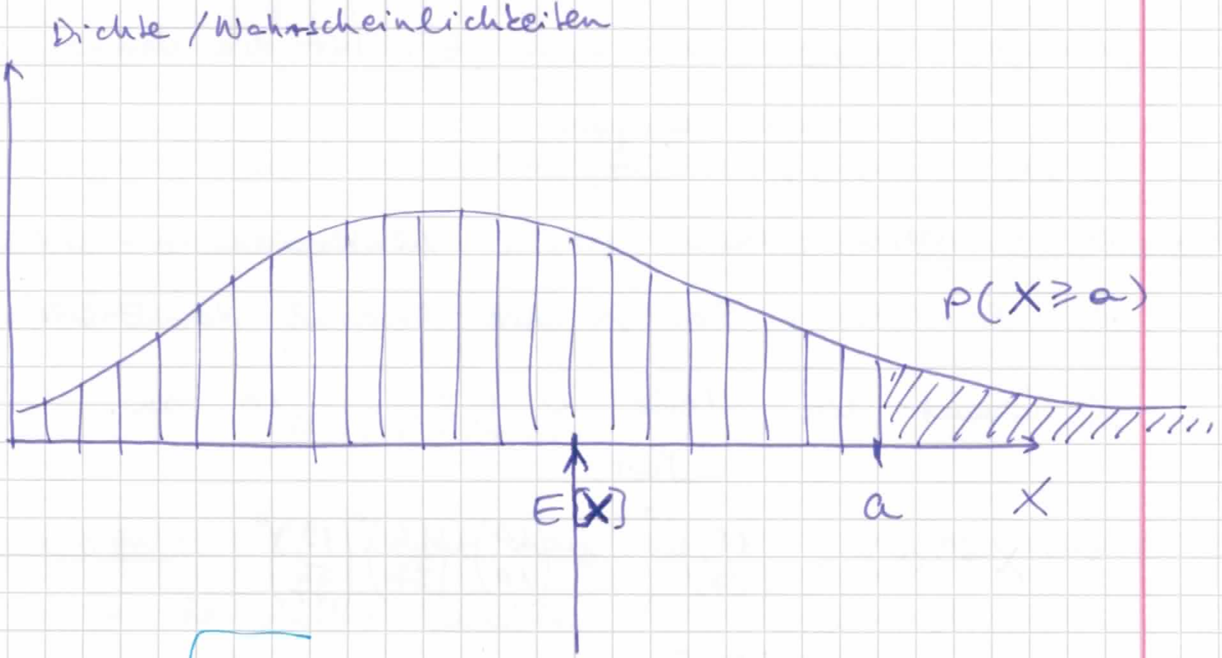
(beachte, dass wir mit negativen Werten die 200 nicht „wiedergutmachen" können, da die Laufzeit L nichtnegativ ist)

Wir formulieren diese Beobachtung allgemein:

Theorem: Sei $X \geq 0$ eine <u>nichtnegative</u> Zufallsvariable und $a \geq E[X]$ eine Zahl. Es gilt:

$$P(X \geq a) \leq \frac{E[X]}{a} \qquad P(X \geq k \cdot E[X]) \leq \frac{1}{k}$$

das ist die <u>Markoff-Ungleichung</u>.



Beweis:

$$\boxed{E[X]} = \sum_r r \cdot P(X=r) = \sum_{r<a} r \cdot P(X=r) + \sum_{r \geq a} r \cdot P(X=r) \geq$$

$$\geq \sum_{r \geq a} r \cdot P(X=r) \geq \sum_{r \geq a} a \cdot P(X=r) = a \cdot \sum_{r \geq a} P(X=r) = \boxed{a \cdot P(X \geq a)}$$

Oft ist es nützlich wenn man Zufallsvariablen addieren, multiplizieren oder in beliebige Funktionen einsetzen kann:

Definition: Seien X und Y Zufallsvariablen über dem selben Ereignisraum. Dann sind $X+Y$, $X \cdot Y$ oder beliebige $f(X)$ oder $f(X, Y)$ auch Zufallsvariablen.

Beispiel 1. Wir würfeln 2mal hintereinander.

Seien X, bzw. Y die Werte des ersten und des zweiten Wurfes. $X+Y$ bedeutet die Summe der zwei Würfe: eine Zufallsvariable mit möglichen Werten $2, 3, \ldots, 11, 12$

Die Verteilung von X und Y ist jeweils uniform (Gleichverteilung) über $1, \ldots, 6$. Ist die Verteilung von $X+Y$ auch uniform?

Beispiel 2. Wir setzen 10 mal hintereinander auf Rot in Roulette. Sei X die Anzahl der Erfolge. X nimmt die Werte $0, 1, 2, \ldots, 10$ an.

$$P(X=10) = \left(\overset{P(\text{Rot})}{\frac{18}{37}}\right)^{10} = \binom{10}{10} \cdot \left(\frac{18}{37}\right)^{10} \left(\frac{19}{37}\right)^{0}$$

$$P(X=0) = \left(\underset{P(\text{schwarz/grün})}{\frac{19}{37}}\right)^{10}$$

$$P(X=3) = \binom{10}{3} \cdot \left(\frac{18}{37}\right)^{3} \cdot \left(\frac{19}{37}\right)^{7}$$

Warum?

$\binom{10}{3}$ Möglichkeiten für die Position der Erfolge

● ● ● ○ ○ ○ ○ ○ ○ ○

● ● ○ ● ○ ○ ○ ○ ○ ○

● ● ○ ○ ● ○ ○ ○ ○ ○

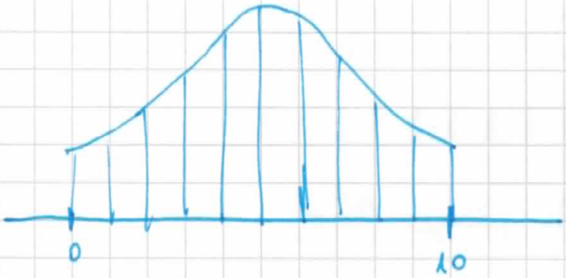
⋮

○ ○ ○ ○ ○ ○ ○ ● ● ●

Die $\binom{10}{3}$ Möglichkeiten schließen sich gegenseitig aus und alle haben $\text{Prob} = \left(\frac{18}{37}\right)^{3} \cdot \left(\frac{19}{37}\right)^{7}$

d.h. die Verteilung von X ist eine Binomialverteilung mit den Parametern $p = \frac{18}{37}$ $n = 10$

definiert durch $P(X = k) = \binom{10}{k} \cdot \left(\frac{18}{37}\right)^k \cdot \left(\frac{19}{37}\right)^{10-k}$

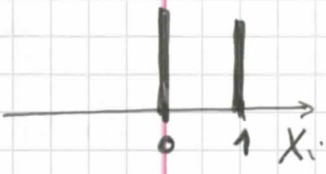


Sei jetzt für das selbe „Zufallsexperiment von 10 Runden" die Variable $X_i = 1$ falls das i-te Spiel rot ergibt, und $X_i = 0$ sonst. Dann gilt

$$X = X_1 + X_2 + X_3 + \dots + X_{10}$$

Definition: X_i nennt man auch Indikatorvariable oder Bernoulli-Variable (sie nimmt die Werte 0 oder 1 an, und zeigt („indiziert") damit ob ein bestimmtes Ereignis eingetreten ist).

$$\boxed{E[X_i] = 1 \cdot \frac{18}{37} + 0 \cdot \frac{19}{37} = P(X_i = 1)}$$



Theorem: Es gilt stets: (nicht nur für unabhängige Variablen)

$$E[X + Y] = E[X] + E[Y]$$

$$E[\lambda \cdot X] = \lambda \cdot E[X]$$

$$E[X + \lambda] = E[X] + \lambda \quad \text{für beliebige } \lambda \in \mathbb{R}$$

Was ergibt dieses Theorem für unsere Beispiele?

1.) Die erwartete Summe von zwei Würfen

$$E[X+Y] = E[X] + E[Y] = 3.5 + 3.5 = 7$$

2. Die erwartete Anzahl der Ausgänge Rot

$$E[X] = E[X_1] + E[X_2] + \ldots + E[X_{10}] = 10 \cdot \frac{18}{37} = n \cdot p(\text{rot})$$

$\longrightarrow$ $n \cdot p$ ist der Erwartungswert der Binomialverteilung

---

<u>Bemerkung:</u> In beiden Beispielen waren X und Y bzw. $X_1 \, X_2 \ldots X_n$ sog. <u>unabhängige Zufallsvariablen</u>

<u>Definition:</u> X und Y sind <u>unabhängige Zufallsvariablen</u>, wenn

$$P(X=a,\ Y=b) = P(X=a) \cdot P(Y=b)$$ für beliebige a und b gilt.

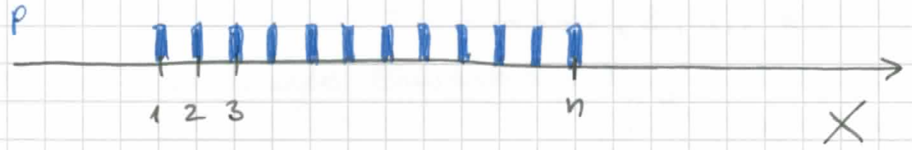
$X_1 \, X_2 \ldots X_n$ sind unabhängig, wenn für jede $a_1 a_2 \ldots a_n$

$$P(X_1=a_1,\ X_2=a_2, \ldots, X_n=a_n) = P(X_1=a_1) \cdot P(X_2=a_2) \ldots P(X_n=a_n)$$

(Die Beispiele können wir auch als zwei bzw. 10 unabhängige Zufallsexperimente betrachten. Wiederholte Experimente entsprechen immer unabhängigen Zufallsvariablen.)

g.) Wichtige Verteilungen

→ uniforme, oder Gleichverteilung

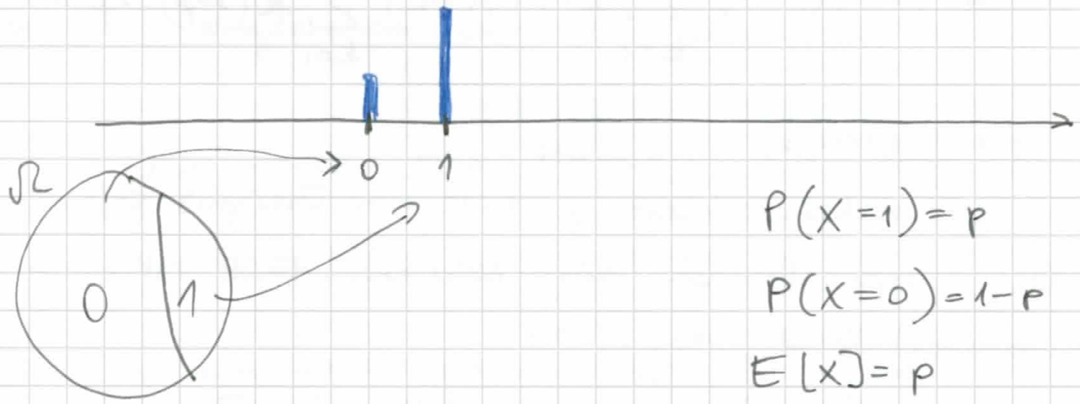


$$P(X=i) = \frac{1}{n} \quad \text{für } 1 \leq i \leq n$$

$$E[X] = \frac{n+1}{2}$$

(die Werte von X dürfen andere sein)

→ binäre oder Bernoulli-Variable



$$P(X=1) = p$$

$$P(X=0) = 1-p$$

$$E[X] = p$$

→ Binomialverteilung mit Parametern n und p

$$P(X=k) = \binom{n}{k} \cdot p^{k} \cdot (1-p)^{n-k} \qquad \text{für } k = 0, 1, 2, \dots, n$$

$$E[X] = n \cdot p$$

→ Geometrische Verteilung mit Parameter p

Wir führen ein Experiment so oft aus, bis Erfolg (ein Ereignis mit Wahrscheinlichkeit p) eintritt.

X sei die Anzahl der ausgeführten Experimente.

$$X = 1, 2, 3, \dots$$

$$P(X = k) = (1-p)^{k-1} \cdot p$$

Beispiel: Wir würfeln so lange bis wir eine 3 würfeln. Wie oft müssen wir würfeln?

$$P\ (X = 9) = \left(\frac{5}{6}\right)^{8} \cdot \frac{1}{6}$$

Wir wissen: $E[X] = \frac{1}{p}$ z.B. hier $E[X] = 6$

Es gilt also: $E[X] = \sum_{k=1}^{\infty} k \cdot P(X = k) = \sum_{k=1}^{\infty} k\,(1-p)^{k-1} \cdot p = \frac{1}{p}$

Warum?

Intuitiv ein Beweis:

mit Wahrscheinlichkeit p tritt das Ereignis sofort beim 1sten Versuch ein; in diesem Fall ist der (erwartete) Wert von X = 1

mit Wahrscheinlichkeit 1-p tritt das Ereignis beim 1sten Versuch nicht ein: dann ist der Erwartungswert 1+ E[X]

Deshalb $E[X] = 1 \cdot p + (1 + E[X]) \cdot (1-p)$

$$p \cdot E[X] = 1$$

$$E[X] = \frac{1}{p}$$

(Offiziell haben wir eben den Satz vom totalen Erwartungswert benutzt:

Der Satz vom totalen Erwartungswert

$$E[X] = E[X|B_1]\cdot P(B_1) + E[X|B_2]\cdot P(B_2)$$

wobei hier B_1 und B_2 das folgende vollständige Ereignissystem bilden:

B_1 : Erfolg beim ersten Experiment

B_2 : kein Erfolg beim ersten Experiment

und der bedingte Erwartungswert wird so definiert:

$$E[X|B] = \sum_{r\in\mathbb{R}} r\cdot P(X=r|B)$$

mit den bedingten Wahrscheinlichkeiten $P(X=r|B)$

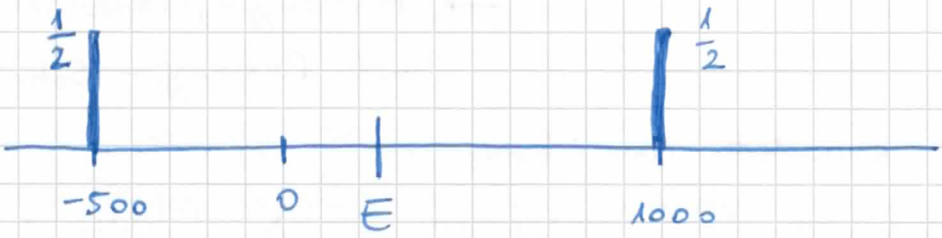
statt $P(X=r)$)]

h.) Die Varianz

Beispiel: Vergleichen wir die folgenden zwei Angebote:

→ wir werfen eine Münze; falls Kopf, bekommen wir 1000 €, falls Zahl, bezahlen wir 500 €.

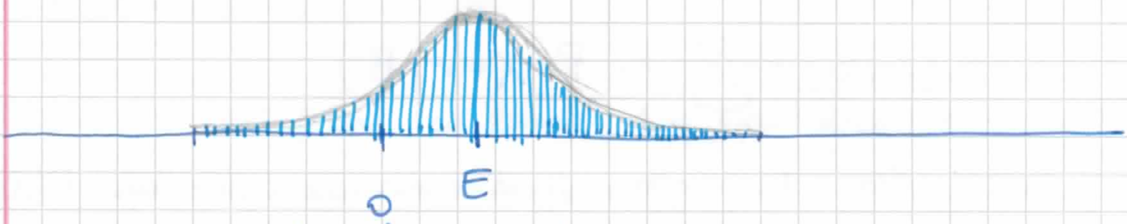
Der erwartete Gewinn ist $E = \frac{1}{2}\cdot 1000 + \frac{1}{2}(-500) = 250$ €



→ Wir werfen eine Münze 1000 mal; jedes mal wenn Kopf, bekommen wir 1 €, und jedes mal wenn Zahl, bezahlen wir 50 Cent.

Der erwartete Gewinn bei einem Wurf ist 25 Cent.

Der erwartete Gesamtgewinn ist 250 €, wie oben.



die Wahrscheinlichkeitsverteilung sieht jetzt so aus

Angenommen wir hätten viel Geduld (oder einen automatisierten Prozess), würden wir wahrscheinlich das 2. Angebot annehmen (wenn wir nicht riskieren möchten). Was ist hier anders?

Was uns interessiert, ist

→ die „durchschnittliche", also erwartete Abweichung vom Erwartungswert → die Varianz

→ die Wahrscheinlichkeit, dass der Gewinn vom Erwartungswert viel zu weit fällt abzuschätzen

→ Markoff-, Tschebischeff-, und Chernoff-Schranken

(Die erwartete Abweichung von E könnte die erwartete Distanz von $E[X]$ sein, also $E[|X - E[X]|]$ stattdessen betrachtet man häufiger den Erwartungswert von $(X - E[X])^2$, die Varianz:)

Definition: Die Varianz einer Zufallsvariable X ist

$$V[X] = E[(X - E[X])^2]$$

eine Maßzahl für die Streuung um den Erwartungswert.

[Sei: $E = E[X]$ E ist also eine Konstante, während X die Zufallsvariable ist

$$V[X] = E[(X-E)^2] = E[X^2 - 2\cdot X\cdot E + E^2] = E[X^2] - 2E\cdot E[X] + E^2 =$$

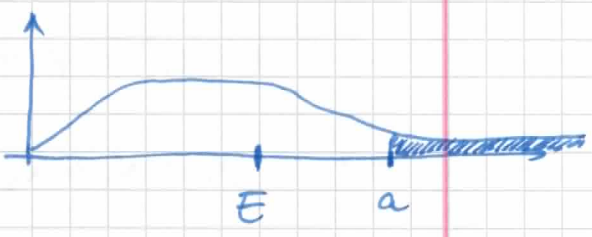
$$= E[X^2] - 2E^2 + E^2 = E[X^2] - (E[X])^2 \quad]$$

Wir sind nicht nur den unbequemen Absolutwert | | losgeworden. Bei $(X-E)^2$ fallen die große Abweichungen $(X-E)$ mehr ins Gewicht. (werden mehr „bestraft" durch einen hohe V[X]]

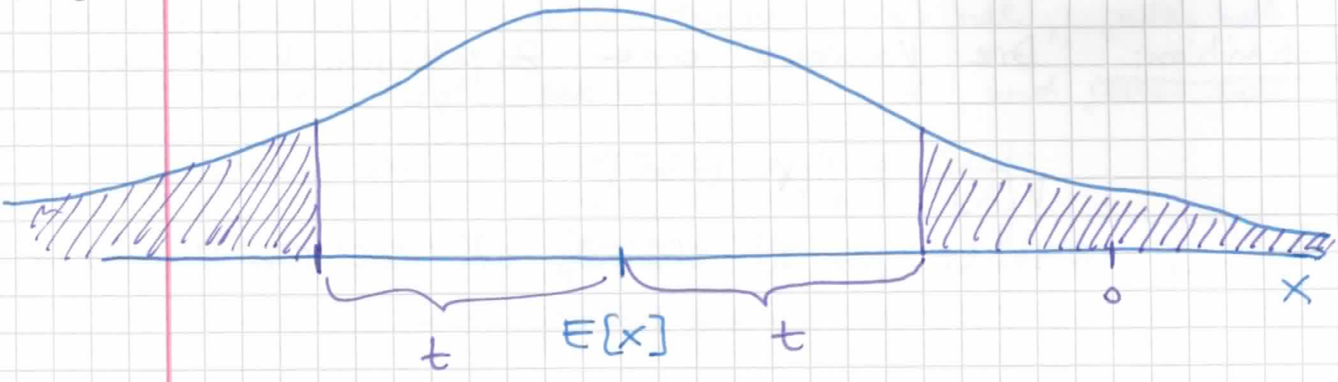
i.) Abweichungen vom Erwartungswert

① Die Markoff-Ungleichung gab eine obere Schranke für die Wahrscheinlichkeit von zu großen Werten von X für nichtnegative Zufallsvariablen.

$$P(X \geq a) \leq \frac{E[X]}{a}$$



② Für beliebige Zufallsvariablen erhalten wir jetzt anhand der Varianz eine obere Schranke für die Wahrscheinlichkeit, dass X zu weit wegfällt von E[X]



die Tschebyscheff-Ungleichung

$$P(|X-E[x]|\geq t)\leq \frac{V[x]}{t^2}$$

(Für den Beweis verwendet man die Markoff-Ungleichung:

$$P(|X-E|\geq t) = P((X-E)^2\geq t^2) \underset{\substack{\downarrow \\ \text{Markoff}}}{\leq} \frac{E[(X-E)^2]}{t^2} = \frac{V[x]}{t^2}$$
)

③ die Chernoff-Ungleichungen

Seien $X_1, X_2 \dots X_k$ unabhängige binäre Zufallsvariablen mit den $P(X_i=1)=p_i$ Erfolgswahrscheinlichkeiten, und sei X die Gesamtanzahl von Erfolgen. Also gilt:

→ $E[X_i] = p_i\cdot 1 + (1-p_i)\cdot 0 = p_i$

→ $X = X_1 + X_2 + \dots + X_k = \sum_{i=1}^{k} X_i$

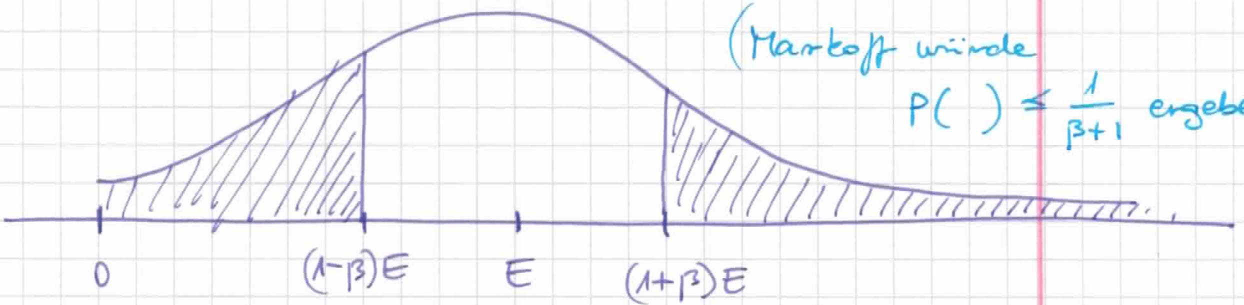
→ $E[x] = \sum_{i=1}^{k} E[X_i] = \sum_{i=1}^{k} p_i$

In diesem Spezialfall gibt es schärfere Abschätzungen:

Sei $E = E\{X]$

die <u>Chernoff-Schranken</u> sind

a.) $$P(X \geq (1+\beta)\cdot E) \leq \left(\frac{e^{\beta}}{(1+\beta)^{1+\beta}}\right)^{E} \leq e^{-\frac{E\beta^2}{3}}$$ für jedes $\beta > 0$



(Markoff würde $P(\ ) \leq \frac{1}{\beta+1}$ ergeben)

b.) $$P(X \leq (1-\beta)\cdot E) \leq \left(\frac{e^{-\beta}}{(1-\beta)^{1-\beta}}\right)^{E} \leq e^{-E\cdot\frac{\beta^2}{2}}$$ für $0<\beta<1$

(Im allgemeinen erhält man eine Chernoff-Schranke, wenn man die Markoff-Ungleichung für die Zufallsvariable $e^{\alpha\cdot X}$ mit einem geeigneten α anwendet.

Was bedeutet dies in unserem Fall?

Sei $t = (1+\beta)E$

$$P(X \geq t) = P\left(e^{\alpha\cdot X} \geq e^{\alpha\cdot t}\right) \leq \frac{E\left[e^{\alpha\cdot X}\right]}{e^{\alpha\cdot t}}$$

Wir nutzen die Unabhängigkeit der X_i aus

$$e^{\alpha\cdot X} = \prod_{i=1}^{k} e^{\alpha\cdot X_i}$$

$$E\left\{\prod_{i=1}^{k} e^{\alpha\cdot X_i}\right] \underset{\text{Unabhängigkeit}}{=} \prod_{i=1}^{k} E\left\{e^{\alpha\cdot X_i}\right]$$

Man erhält die Ungleichung a.) mit

$t = (1+\beta)\cdot E$ und $\alpha = \ln(1+\beta)$)