MaTeGnu

Mathematik mit **Te**chnologie an **G**rundvorstellungen orientiert **n**achhaltig **u**nterrichten

Modul 4: Stochastik

Teil 2

Jürgen Roth & Susanne Digel

10.02.2025

Didaktik der Mathematik Sekundarstufen

RPTU Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

MaTeGnu

Modul 4: Stochastik

Teil 1

1. Stochastik in den Bildungsstandards
2. Aspekte der Stochastik in der Sek. I
3. Wahrscheinlichkeitsrechnung

Teil 2

4. Beurteilende Statistik

GeoGebra-Buch
„Didaktik der Stochastik"
https://roth.tel/stochastik

RPTU

4 Beurteilende Statistik

RPTU
Rheinland-Pfälzische
Technische Universität
Kaiserslautern
Landau

MaTeGnu

Kapitel 4:
Beurteilende Statistik

GeoGebra-Buch
„Didaktik der Stochastik“
https://roth.tel/stochastik

dms.nuw.rptu.de/mategnu

RPTU

Lehrplanalternativen RLP

Grundkurs
(ca. 42 Stunden)

Stochastik 1

wahlweise

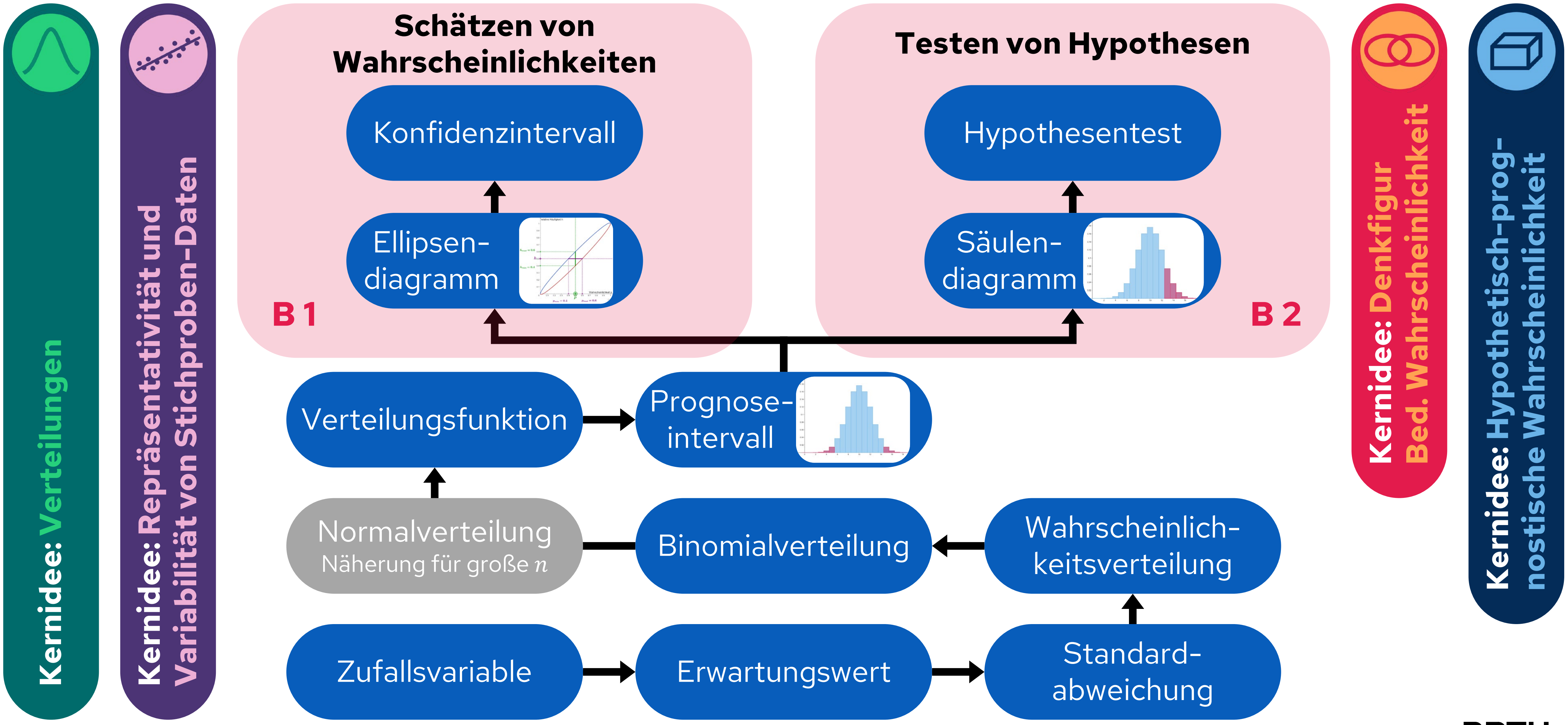
- Stochastik 2 – B1: Schätzen von Wahrscheinlichkeiten
- Stochastik 2 – B2: Testen von Hypothesen

Leistungskurs
(ca. 70 Stunden)

Stochastik

MBWWK (2015). Lehrplan Mathematik – Grund- und Leistungsfach in der gymnasialen Oberstufe (Mainzer Studienstufe)

Aufbau der beurteilenden Statistik
Grundkurs



Aufbau der beurteilenden Statistik
Leistungskurs

Kernidee: Verteilungen

Kernidee: Repräsentativität und Variabilität von Stichproben-Daten

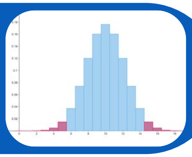
Vernetzen über Prognoseintervall

Hypothesentest

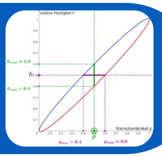
Konfidenzintervall

Verteilungsfunktion

Prognose-intervall



Ellipsen-diagramm



Normalverteilung
Näherung für große n

Binomialverteilung

Wahrscheinlich-keitsverteilung

Zufallsvariable

Erwartungswert

Standard-abweichung

Kernidee: Denkfigur Bed. Wahrscheinlichkeit

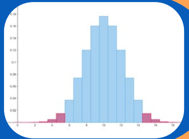
Kernidee: Hypothetisch-prog-nostische Wahrscheinlichkeit

Stochastik-Curriculum
Sekundarstufen I und II

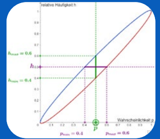
RPTU
Didaktik der Mathematik Sekundarstufen

Beschreibende Statistik
- Erhebung – Daten sammeln
- Graphische Darstellungsformen
- Mittelwerte & Streuungsmaße

Beurteilende Statistik
- Hypothesentest
- Konfidenzintervall
- Verteilungsfunktion
- Prognose-intervall



- Ellipsen-diagramm



- Normalverteilung
  Näherung für große n
- Binomialverteilung
- Wahrscheinlich-keitsverteilung
- Zufallsvariable
- Erwartungswert
- Standard-abweichung

Wahrscheinlich-keitsrechnung
- (Mehrstufige) Zufallsexperimente
- Hypothetisch-prognostische Wahrscheinlichkeit
- Stoch. (Un-)Abhängigkeit
- Bedingte Wahrscheinlichkeit

Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

MaTeGnu

Kapitel 4: Beurteilende Statistik

GeoGebra-Buch „Didaktik der Stochastik“ https://roth.tel/stochastik

Kernideen im Fokus
bei Zufallsvariable, Erwartungswert und Standardabweichung

RPTU
Didaktik der Mathematik
Sekundarstufen

Kernidee
Denkfigur: Bedingte Wahrscheinlichkeit

- Konditionales Denken
  (Wenn …, dann …)
- Schnitt-Wahrscheinlichkeit
  ↔ totale Wahrscheinlichkeit

Kernidee
hypothetisch-prognostische Wahrscheinlichkeit

- Konzept des Bezweifelns
- Modell (Wahrscheinlichkeit)
  ↔ Realität (rel. Häufigkeit)
- aus Erfahrung wird Erwartung

Kernidee
Repräsentativität und Variabilität von Stichproben-Daten

- Repräsentative Stichprobe
- Wachsende Stichprobe
  → reduziert Variabilität

Kernidee
Verteilungen

- Struktur erfassen
  (Form, Streuung, Maximum, …)
- Gesetz der großen Zahlen
- Sichtweise: lokal ↔ global

Lehrplan RLP: Leistungskurs

	Ziele / Inhalte (Sach- & Methodenkompetenz)	**Hinweise zur Unterrichtsgestaltung und Methodenkompetenz**
6.	Die Begriffe „Zufallsgröße” und „Wahrscheinlichkeitsverteilung“ kennen und an Beispielen erläutern (4.12g)	
7.	Die Begriffe „Erwartungswert“, „Varianz“ und „Standardabweichung“ einer diskreten Zufallsgröße kennen und anwenden (2.06g, 2.07g, 5.01g, 5.04g)	Bei der Anwendung in Sachaufgaben kommt es vor allem darauf an, dass die SuS verstehen, welche Folgerungen man aus den Kennwerten für das Sachproblem ziehen kann. Hier bietet es sich an, exemplarisch eine statistische Erhebung zu planen und zu beurteilen.

Lehrplan RLP: Grundkurs – Stochastik 1

	Ziele / Inhalte (Sach- & Methodenkompetenz)	**Hinweise zur Unterrichtsgestaltung und Methodenkompetenz**
8.	Erwartungswert und Standardabweichung für Binomialverteilungen berechnen und anwenden (2.06g, 2.07g, 5.01g, 5.04g)	Die Formeln sollen anhand von Plausibilitätsbetrachtungen gewonnen werden. Bei der Anwendung in Sachaufgaben kommt es vor allem darauf an, dass die Schülerinnen und Schüler verstehen, welche Folgerungen man aus den Kennwerten für das Sachproblem ziehen kann. Hier bietet es sich an, exemplarisch eine statistische Erhebung zu planen und zu beurteilen.

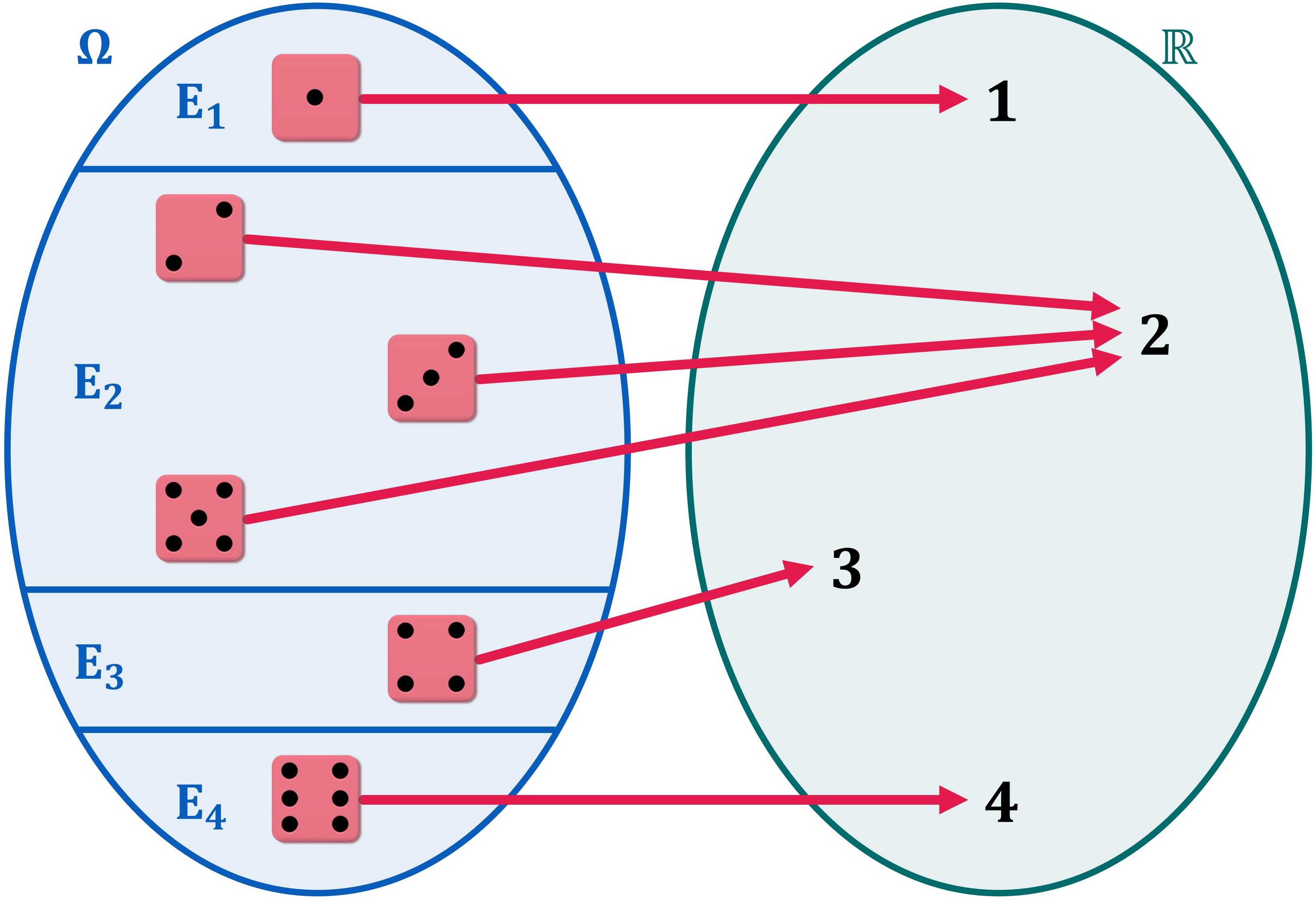
Beispiel: Würfeln mit einem Würfel

Beispiel:
Würfeln mit einem Würfel

Die Zufallsvariable X ordnet jedem Ergebnis ω die Anzahl x der Teiler der geworfenen Augenzahl zu:

$X : \Omega \to \mathbb{R}, \omega \mapsto x$

Bemerkung
Obige Zufallsvariable X ist nur eine von sehr vielen möglichen Zufallsvariablen zum Zufallsexperiment Würfeln mit einem Würfel. Man könnte z.B. auch die Zufallsvariable Y definieren, die der Augenzahl 6 den Wert 4 € und allen anderen Augenzahlen den Wert -1 € (als Gewinnausschüttung) zuordnet.



Zufallsvariable

Achtung: Es gibt diskrete und stetige Zufallsvariablen!
Wir betrachten zunächst nur diskrete Zufallsvariablen.

Definition Zufallsvariable

Eine Funktion X, die jedem Ergebnis $\omega \in \Omega$ eine reelle Zahl $x \in \mathbb{R}$ zuordnet, heißt **Zufallsvariable** (auch **Zufallsgröße**).

$$X: \Omega \to \mathbb{R}, \omega \mapsto x$$

Bemerkungen

- Eine Zufallsvariable X ordnet einem Ergebnis ω jeweils eine reelle Zahlen x zu ($X(\omega) = x$), wobei verschiedenen Ergebnissen durchaus dieselbe Zahl zugeordnet werden kann.
- Ein Ereignis E, d.h. eine Teilmenge des Ergebnisraums Ω ($E \subseteq \Omega$), die alle Ergebnisse $\omega \in \Omega$ enthält, denen ein bestimmter Wert x der Zufallsvariable X zugeordnet wird, bezeichnet man mit: $E = \{X = x\} := \{\omega \in \Omega \mid X(\omega) = x\}$

Beispiel: Würfeln mit einem Würfel

Eine Ergebnismenge Ω ist $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Die Zufallsvariable X ordnet jedem Ergebnis ω die Anzahl der Teiler der geworfenen Augenzahl zu:
$X: \Omega \to \mathbb{R}, \omega \mapsto x :=$ Anzahl der Teiler von ω

Ereignisse

- Augenzahl hat einen Teiler ($1 = 1 \cdot 1$):
  $E_1 = \{X = 1\} := \{\omega \in \Omega \mid X(\omega) = 1\} = \{1\}$
- Augenzahl hat zwei Teiler ($2 = 1 \cdot 2;\ 3 = 1 \cdot 3;\ 5 = 1 \cdot 5$):
  $E_2 = \{X = 2\} := \{\omega \in \Omega \mid X(\omega) = 2\} = \{2; 3; 5\}$
- Die Augenzahl hat drei Teiler ($4 = 1 \cdot 4 = 2 \cdot 2$):
  $E_3 = \{X = 3\} := \{\omega \in \Omega \mid X(\omega) = 3\} = \{4\}$
- Die Augenzahl hat vier Teiler ($6 = 1 \cdot 6 = 2 \cdot 3$):
  $E_4 = \{X = 4\} = \{\omega \in \Omega \mid X(\omega) = 4\} = \{6\}$

Weitere sinnvolle Notationen:

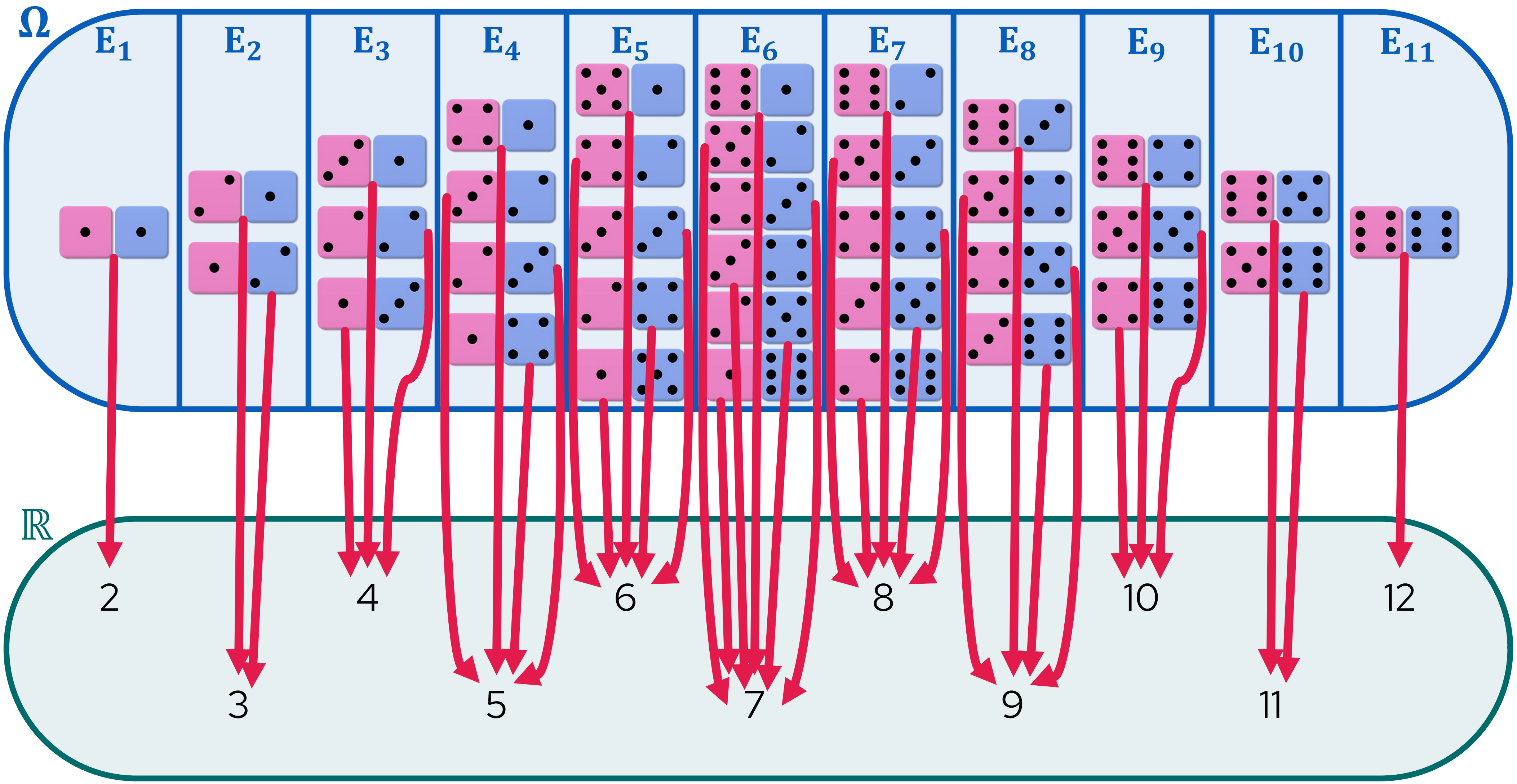
- $\{X > 1\} = \{X = 2\} \cup \{X = 3\} \cup \{X = 4\}$
- $\{1 < X < 4\} = \{X = 2\} \cup \{X = 3\}$

Beispiel: Würfeln mit zwei Würfeln

Beispiel: Würfeln mit zwei Würfeln

Die Zufalls-variable X ordnet jedem Ergebnis ω die Augen-summe x der geworfenen Augenzahl zu:

$X: \Omega \to \mathbb{R},$
$\omega \mapsto x$



Beispiel: $E_3 := \{\omega \in \Omega | X(\omega) = 4\} = \{(1;3);(2;2);(3;1)\}$

Beispiel: Würfeln mit zwei Würfeln

Beispiel: Würfeln mit zwei Würfeln

Die Zufallsvariable X ordnet jedem Ergebnis ω die Augensumme x der geworfenen Augenzahl zu:

$$X: \Omega \to \mathbb{R}, \omega \mapsto x$$

Die Zufallsvariable Y ordnet jedem Ergebnis ω die Augenzahl x des ersten Würfels zu:

$$Y: \Omega \to \mathbb{R}, \omega \mapsto y$$

Die Zufallsvariable Z ordnet jedem Ergebnis ω die Augenzahl y des zweiten Würfels zu:

$$Z: \Omega \to \mathbb{R}, \omega \mapsto z$$

Bemerkungen

- Zu einem Zufallsexperiment können sehr verschiedene Zufallsvariablen definiert werden. Links stehen drei Beispiele zum Zufallsexperiment „Würfeln mit zwei Würfeln".
- Mit Zufallsvariablen kann gerechnet werden. So gilt z.B. für die Zufalls-variablen X, Y und Z und alle $\omega \in \Omega$ im Beispiel links:
  $X(\omega) = (Y + Z)(\omega) := Y(\omega) + Z(\omega)$

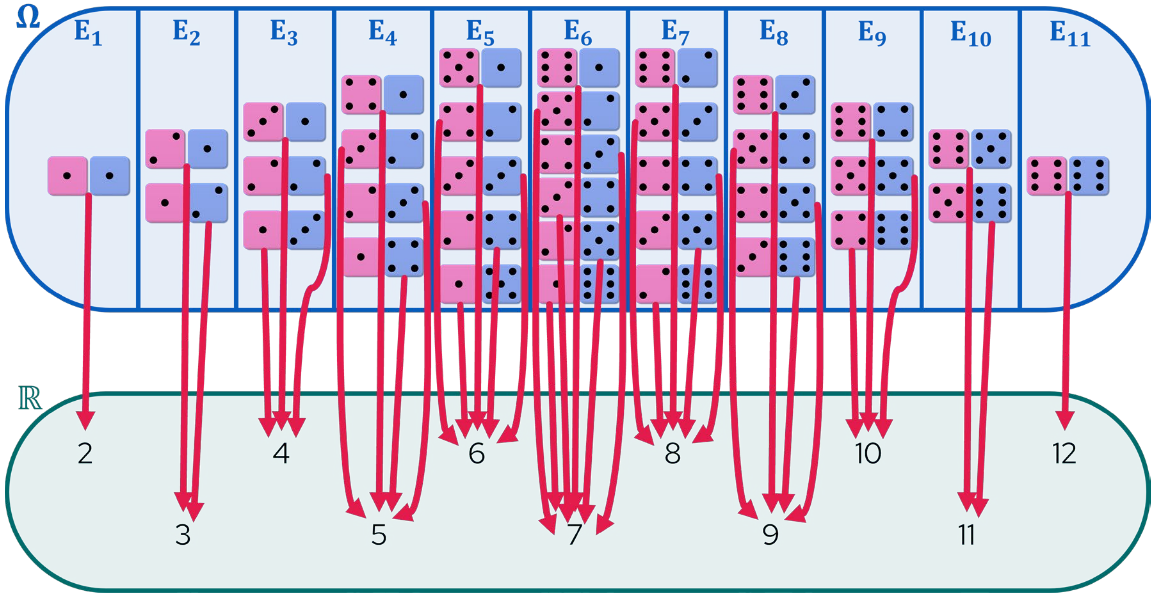
Präsenzaufgabe: Welche Wertebereiche besitzen folgende Zufallsvariablen im obigen Beispiel des zweifachen Würfelwurfs?

a) $Y - Z$ **b)** $Y \cdot Z$ **c)** $Y - 2Z$

Beispiel: Würfeln mit zwei Würfeln

Beispiel: Würfeln mit zwei Würfeln
Die Zufallsvariable X ordnet jedem Ergebnis ω die Augensumme x der geworfenen Augenzahl zu:
$X: \Omega \to \mathbb{R}, \omega \mapsto x$

Aufgabe: Welche mittlere Augensumme $\mu = E(X)$ ist beim Würfeln mit zwei Würfeln zu erwarten?



Idee: Die Augensummen treten mit unterschiedlichen Wahrscheinlichkeiten auf, sie sollten also gewichtet mit der jeweiligen Wahrscheinlichkeit in den im Mittel erwarteten Wert (**Erwartungswert**) eingehen.

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

$$\mu = E(X) = 2\cdot\frac{1}{36} + 3\cdot\frac{2}{36} + 4\cdot\frac{3}{36} + 5\cdot\frac{4}{36} + 6\cdot\frac{5}{36} + 7\cdot\frac{6}{36} + 8\cdot\frac{5}{36} + 9\cdot\frac{4}{36} + 10\cdot\frac{3}{36} + 11\cdot\frac{2}{36} + 12\cdot\frac{1}{36} = \frac{252}{36} = 7$$

Zusammenhang zwischen arithmetischem Mittel und Erwartungswert

Didaktik der Mathematik Sekundarstufen

Zufallsexperiment: Es wird 20-mal mit einem Spielwürfel gewürfelt. Die Zufallsvariable X ordnet jedem Ergebnis ω die geworfene Augenzahl x zu: $X: \Omega \to \{1,2,3,4,5,6\}, \omega \mapsto x$

Ergebnis einer Versuchsreihe von $n = 20$ Würfelwürfen: $6, 3, 3, 3, 5, 4, 3, 1, 6, 2, 3, 3, 3, 6, 5, 4, 6, 1, 5, 5$

Realität: Arithmetisches Mittel

$$\bar{x} = \frac{x_1+x_2+\cdots+x_n}{n} = \frac{6+3+3+3+5+4+3+1+6+2+3+3+3+6+5+4+6+1+5+5}{20}$$
$$= \frac{(1+1)+(2)+(3+3+3+3+3+3+3)+(4+4)+(5+5+5+5)+(6+6+6+6)}{20} = \frac{1\cdot 2+2\cdot 1+3\cdot 7+4\cdot 2+5\cdot 4+6\cdot 4}{20}$$

Mit $H_n(x)$ = „Absolute Häufigkeit des Auftretens der Augenzahl x bei n Würfelwürfen“ ergibt sich:

$$\bar{x} = \frac{1\cdot H_n(1)+2\cdot H_n(2)+3\cdot H_n(3)+4\cdot H_n(4)+5\cdot H_n(5)+6\cdot H_n(6)}{n}$$
$$= 1\cdot\frac{H_n(1)}{n} + 2\cdot\frac{H_n(2)}{n} + 3\cdot\frac{H_n(3)}{n} + 4\cdot\frac{H_n(4)}{n} + 5\cdot\frac{H_n(5)}{n} + 6\cdot\frac{H_n(6)}{n}$$

Mit $h_n(x) = \frac{H_n(x)}{n}$ = „Relative Häufigkeit des Auftretens der Augenzahl x bei n Würfelwürfen“ folgt:

$$\bar{x} = x_1 \cdot h_n(x_1) + \cdots + x_k \cdot h_n(x_k)$$

Modell: Erwartungswert

$$\mu = E(X) := x_1 \cdot P(X = x_1) + \cdots + x_k \cdot P(X = x_k)$$

Erwartungswert

Didaktik der Mathematik Sekundarstufen

Definition Erwartungswert
Für eine Zufallsvariable $X : \Omega \to \mathbb{R}$ mit der Wertemenge $\{x_1, \dots, x_k\}$ heißt

$$E(X) = x_1 \cdot P(X = x_1) + \cdots + x_k \cdot P(X = x_k)$$

der **Erwartungswert** von X.

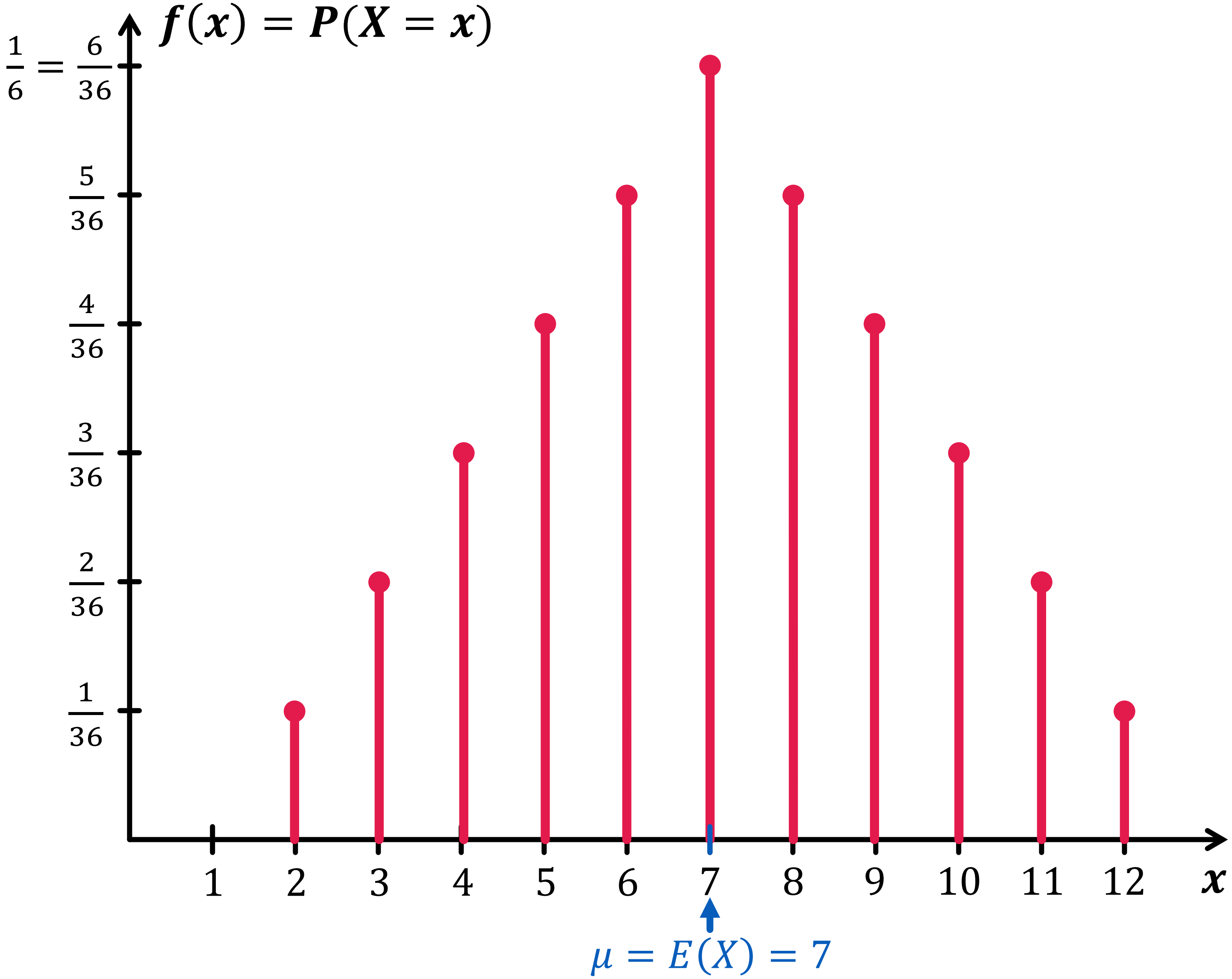
Ausgedrückt mit dem Summenzeichen:

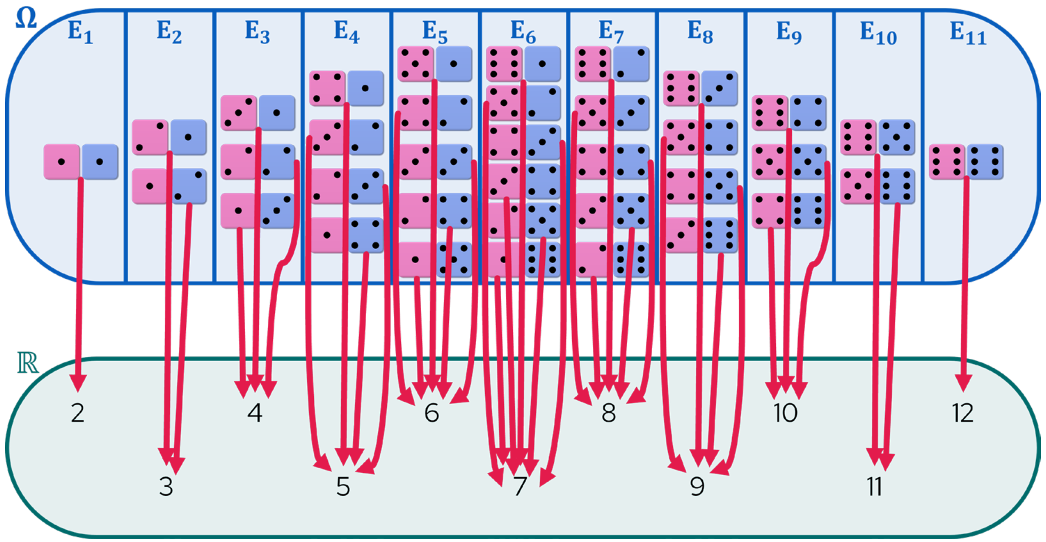
$$E(X) = \sum_{i=1}^{k} x_i \cdot P(X = x_i)$$

Bemerkungen

- Statt $E(X)$ schreibt man auch $\mu(X)$ bzw. nur μ, wenn man nur eine einzige Zufallsvariable betrachtet.
- Bei mehreren Zufallsvariablen X_i ist die Bezeichnung $E(X_i) = \mu_i$ für die Erwartungswerte üblich.
- Interpretiert man die Wahrscheinlichkeitsverteilung als Masseverteilung mit den $P(X = x_i)$ als punktförmigen Massen mit der Gesamtmasse 1, dann lässt sich $\mu = E(X)$ als die Stelle deuten, an der man unterstützen muss, damit die „Waage“ im Gleichgewicht bleibt.

Wahrscheinlichkeitsverteilung





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

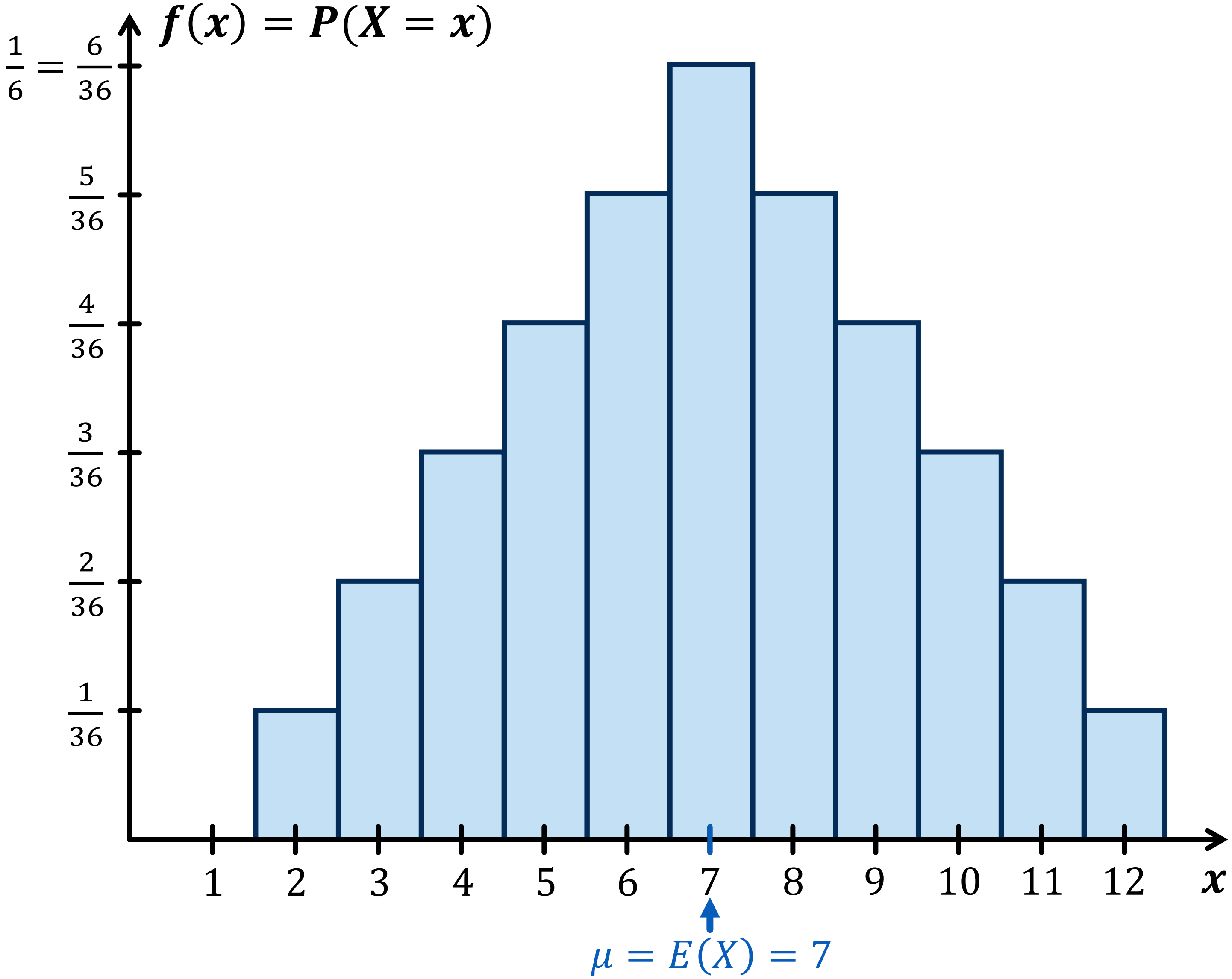
Beispiel: Würfeln mit zwei Würfeln

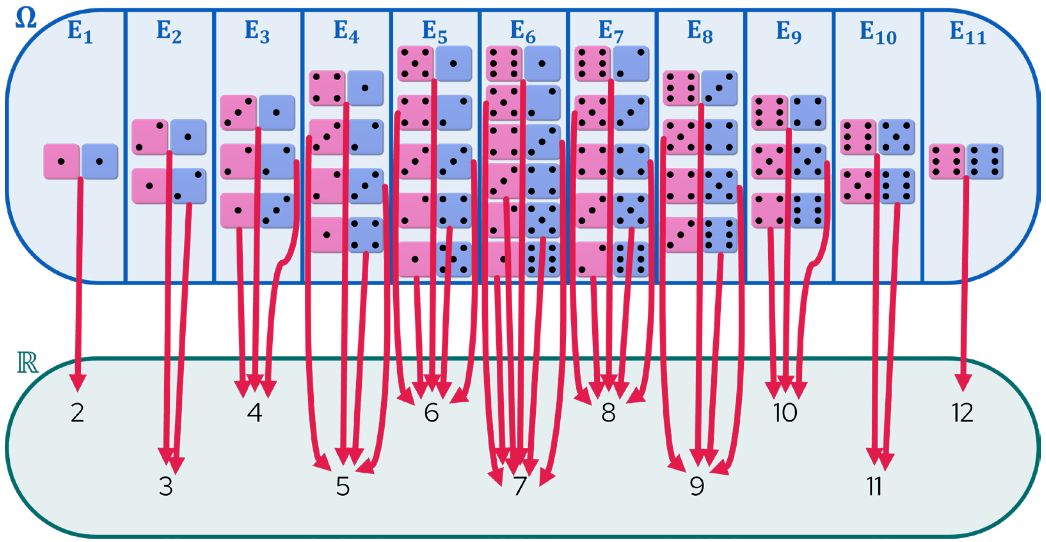
Die Zufallsvariable X ordnet jedem Ergebnis ω die Augensumme x der geworfenen Augenzahl zu:

$$X: \Omega \to \mathbb{R}, \omega \mapsto x$$

Wahrscheinlichkeitsverteilung

Didaktik der Mathematik Sekundarstufen





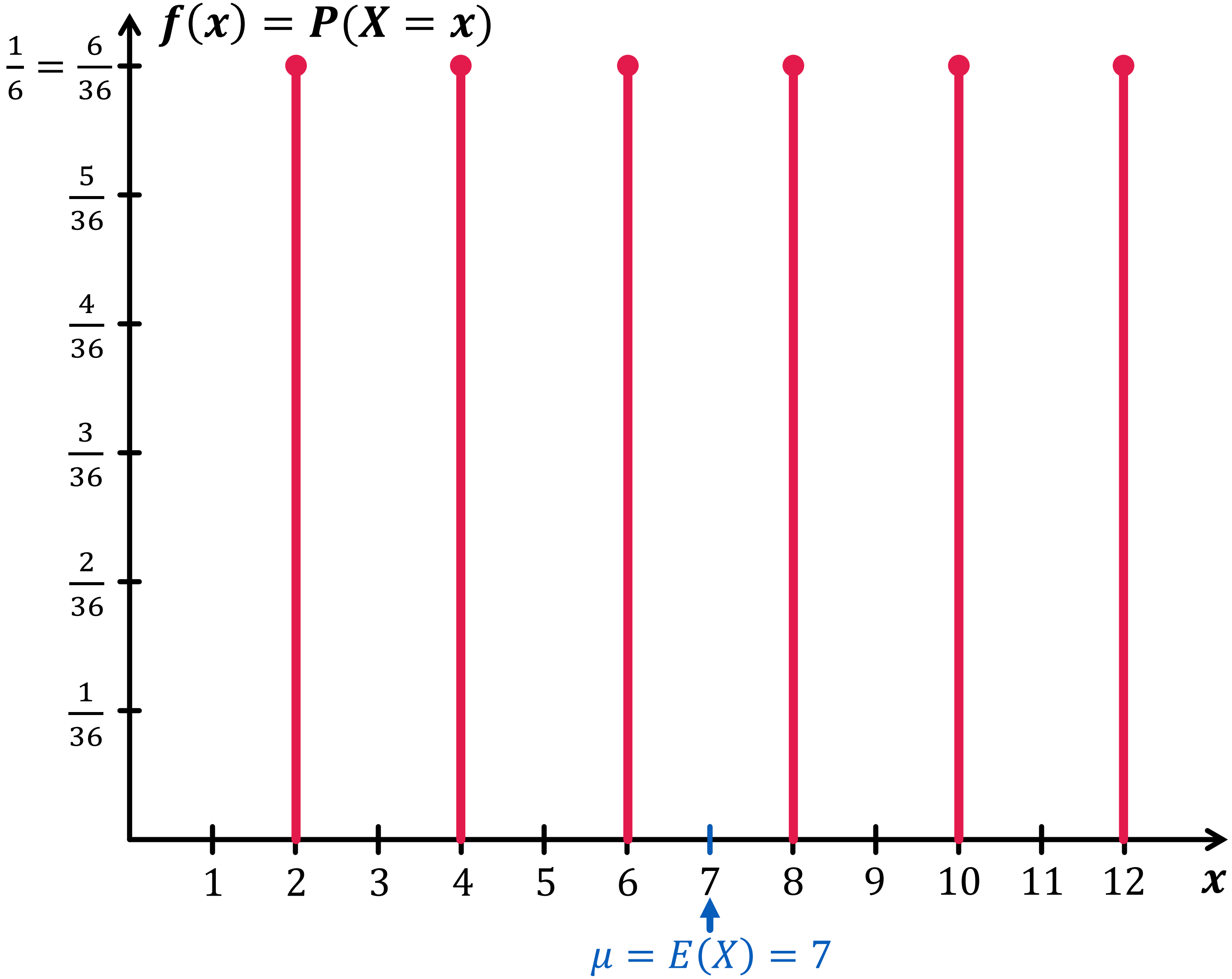
x	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

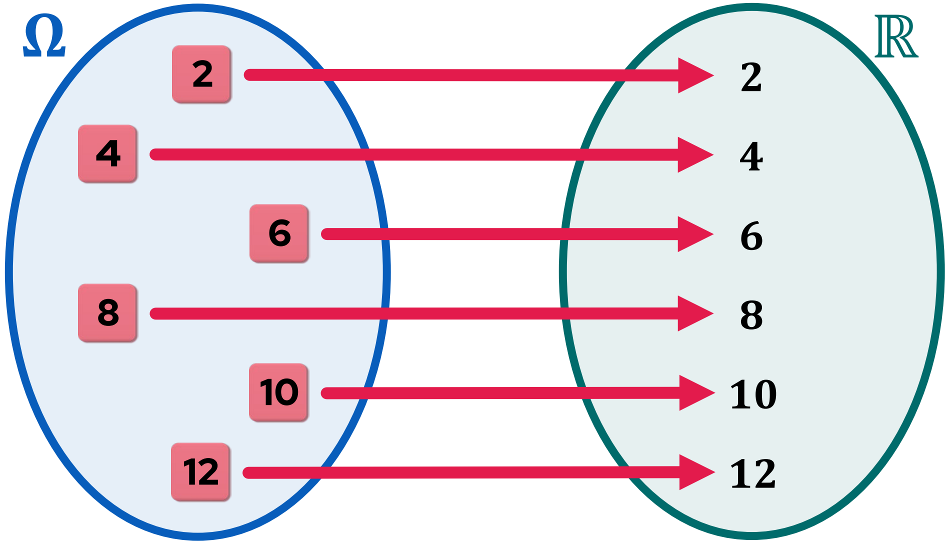
Beispiel: Würfeln mit zwei Würfeln
Die Zufallsvariable X ordnet jedem Ergebnis ω die Augensumme x der geworfenen Augenzahl zu:

$$X: \Omega \to \mathbb{R}, \omega \mapsto x$$

RPTU

Wahrscheinlichkeitsverteilung





x	2	4	6	8	10	12
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Beispiel: Würfeln mit einem Würfel mit den Augenzahlen 2, 4, 6, 8, 10, 12
Die Zufallsvariable X ordnet jedem Ergebnis ω die geworfene Augenzahl x zu: $X: \Omega \to \mathbb{R}, \omega \mapsto x$

Warum Streuungsmaße?

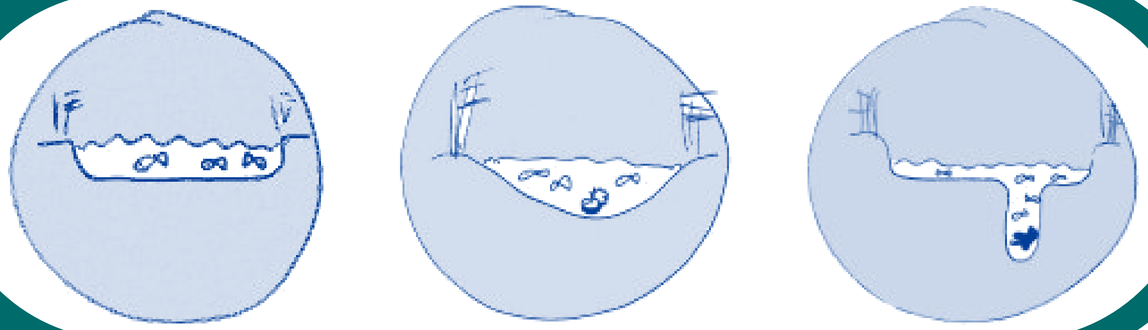
Didaktik der Mathematik Sekundarstufen

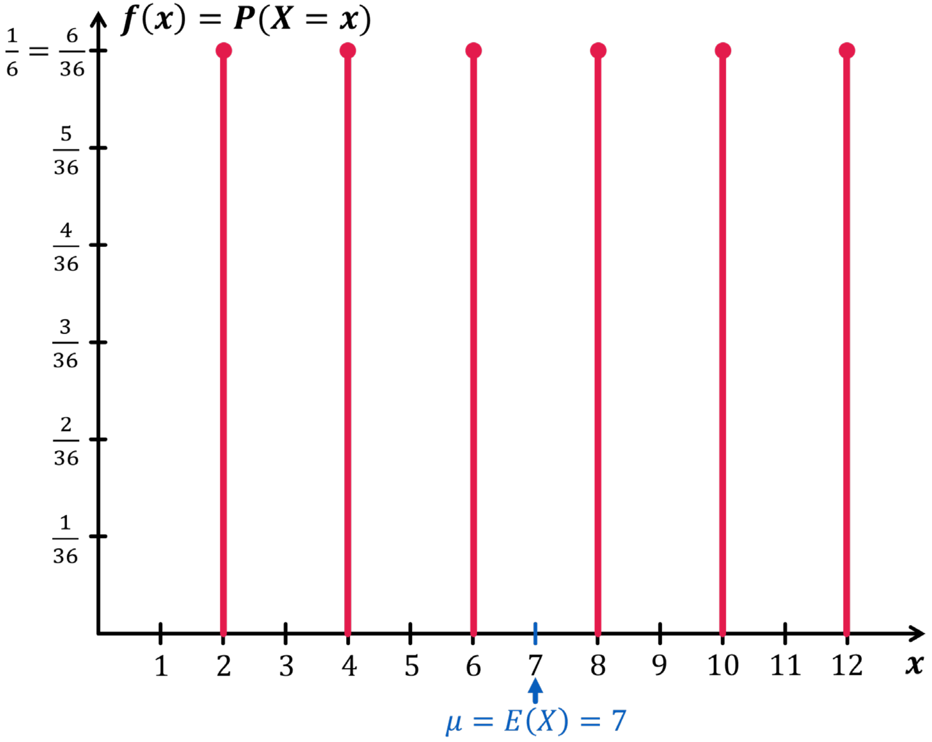
Durchschnitte (Erwartungswerte) allein machen nicht glücklich!

Krämer, W. (20068). So lügt man mit Statistik. Piper

Beispiel: Flussquerschnitte gleicher mittlerer Tiefe

madin

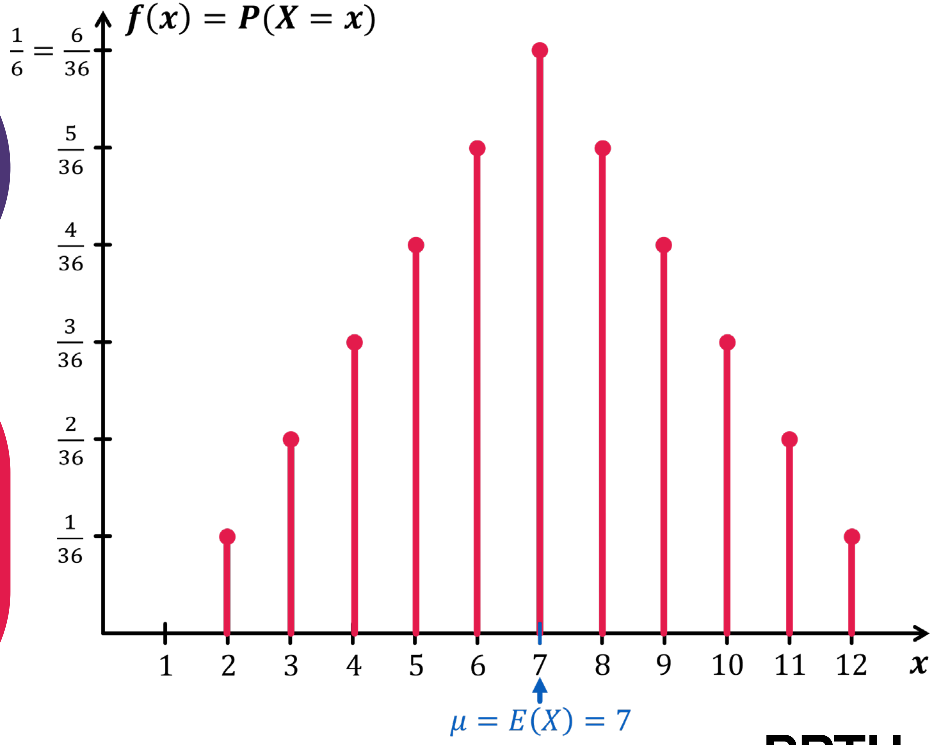




https://www.geogebra.org/m/wzm6mwpt

Wahrscheinlichkeits-verteilungen mit demselben Erwartungswert $E(X)$.

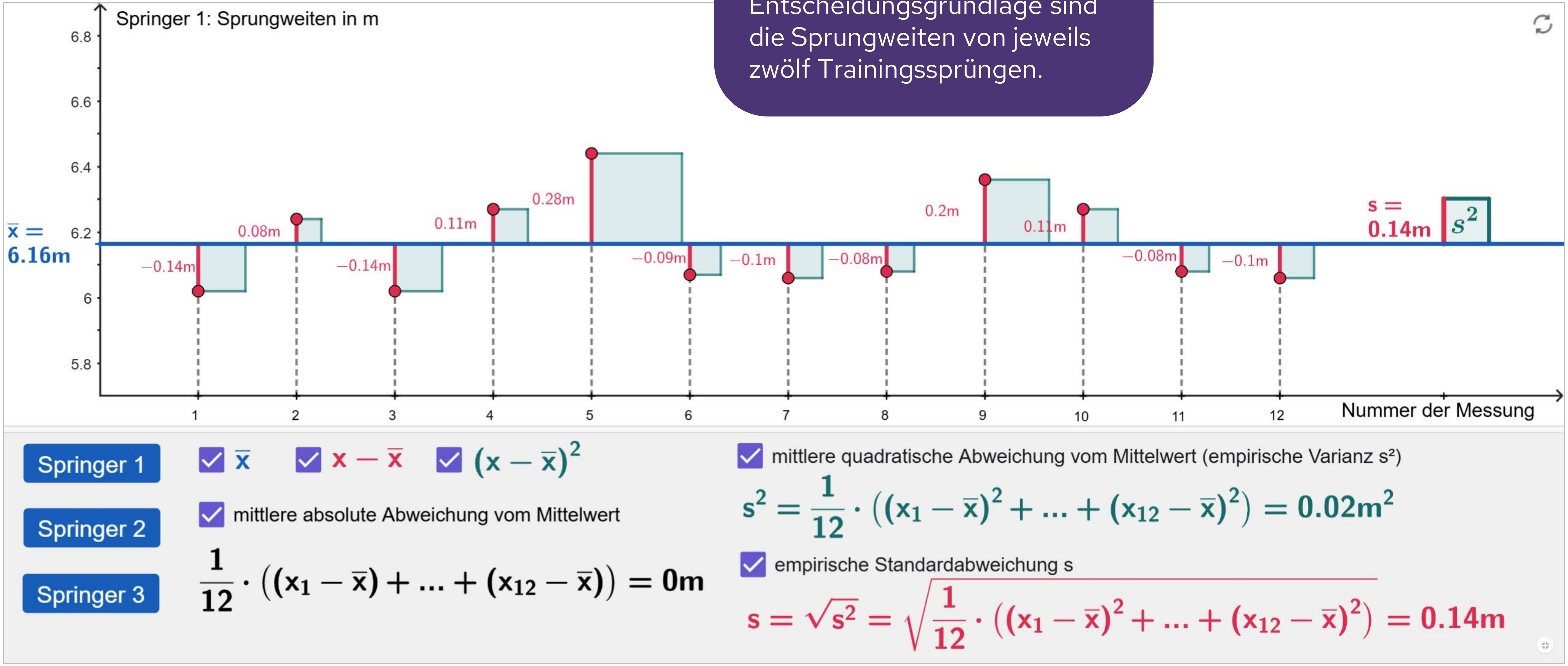
Gesucht:
Mittlere Abweichung der Ausprägungen der Zufallsvariable vom Erwartungswert als Maß für die Streuung.



https://www.geogebra.org/m/zahmzfqx

Warum Streuungsmaße?

Ein **Weitsprung-Trainer** muss entscheiden, welchen der von ihm betreuten drei Springer er zum Wettkampf schicken soll. Entscheidungsgrundlage sind die Sprungweiten von jeweils zwölf Trainingssprüngen.



 https://www.geogebra.org/m/ttaugsmv

Zusammenhang: Empirische Standardabweichung s und Standardabweichung σ

RPTU Didaktik der Mathematik Sekundarstufen

Zufallsexperiment: Es wird 20-mal mit einem Spielwürfel gewürfelt. Die Zufallsvariable X ordnet jedem Ergebnis ω die geworfene Augenzahl x zu: $X: \Omega \to \{1,2,3,4,5,6\}, \omega \mapsto x$

Ergebnis einer Versuchsreihe von $n = 20$ Würfelwürfen: 6, 3, 3, 3, 5, 4, 3, 1, 6, 2, 3, 3, 3, 6, 5, 4, 6, 1, 5, 5

Realität: Arithm. Mittel s^2 der quadratischen Abweichungen vom arithm. Mittel $\overline{x} = 3{,}5$

$$s^2 = \frac{(x_1-\overline{x})^2+(x_2-\overline{x})^2+\cdots+(x_n-\overline{x})^2}{n} = \frac{(1-3{,}5)^2\cdot 2+(2-3{,}5)^2\cdot 1+(3-3{,}5)^2\cdot 7+(4-3{,}5)^2\cdot 2+(5-3{,}5)^2\cdot 4+(6-3{,}5)^2\cdot 4}{20}$$

Mit $H_n(x) =$ „Absolute Häufigkeit des Auftretens der Augenzahl x bei n Würfelwürfen" ergibt sich:

$$s^2 = \frac{(1-\overline{x})^2\cdot H_n(1)+(2-\overline{x})^2\cdot H_n(2)+(3-\overline{x})^2\cdot H_n(3)+(4-\overline{x})^2\cdot H_n(4)+(5-\overline{x})^2\cdot H_n(5)+(6-\overline{x})^2\cdot H_n(6)}{n}$$

$$= (1-\overline{x})^2\cdot\frac{H_n(1)}{n} + (2-\overline{x})^2\cdot\frac{H_n(2)}{n} + (3-\overline{x})^2\cdot\frac{H_n(3)}{n} + (4-\overline{x})^2\cdot\frac{H_n(4)}{n} + (5-\overline{x})^2\cdot\frac{H_n(5)}{n} + (6-\overline{x})^2\cdot\frac{H_n(6)}{n}$$

Mit $h_n(x) = \frac{H_n(x)}{n} =$ „Relative Häufigkeit des Auftretens der Augenzahl x bei n Würfelwürfen" folgt:

$$s^2 = (x_1-\overline{x})^2\cdot h_n(x_1) + \cdots + (x_k-\overline{x})^2\cdot h_n(x_k) \quad\Rightarrow\quad s = \sqrt{(x_1-\overline{x})^2\cdot h_n(x_1) + \cdots + (x_k-\overline{x})^2\cdot h_n(x_k)}$$

Modell: **Varianz σ^2:** $\sigma^2 := (x_1-\mu)^2\cdot P(X=x_1) + \cdots + (x_k-\mu)^2\cdot P(X=x_k)$

Standardabweichung σ:

$$\sigma = \sqrt{\sigma^2} := \sqrt{(x_1-\mu)^2\cdot P(X=x_1) + \cdots + (x_k-\mu)^2\cdot P(X=x_k)}$$

Streuungsmaße
Varianz und Standardabweichung

Bemerkung: Würden zur Messung der mittleren Abweichung vom Erwartungswert nur die Differenzen $x_i - \mu$ gebildet, dann würden sich Abweichungen nach unten und oben über die unterschiedlichen Vorzeichen gegenseitig aufheben.

Durch das Quadrieren der Terme $(x_i - \mu)^2$ entfällt dieser Effekt.

Definition Varianz

Für eine Zufallsvariable $X\colon \Omega \to \mathbb{R}$ mit der Wertemenge $\{x_1, \dots, x_k\}$ und dem Erwartungswert $E(X) = \mu$ heißt

$$\sigma^2 = \mathrm{Var}(X) = E\left((X-\mu)^2\right)$$

$$= (x_1 - \mu)^2 \cdot P(X = x_1) + \cdots + (x_k - \mu)^2 \cdot P(X = x_k)$$

die **Varianz** von X.

Bemerkung: Die Maßeinheit der Varianz $\mathrm{Var}(X)$ ist das Quadrat der Einheit, in der die Zufallsvariable X gemessen wird.

Zur leichteren Interpretierbarkeit wurde deshalb mit der Standardabweichung $\sigma(X)$ eine Maßzahl konstruiert, die dieselbe Einheit wie X hat.

Für die Varianz schreibt man auch σ^2.

Definition Standardabweichung

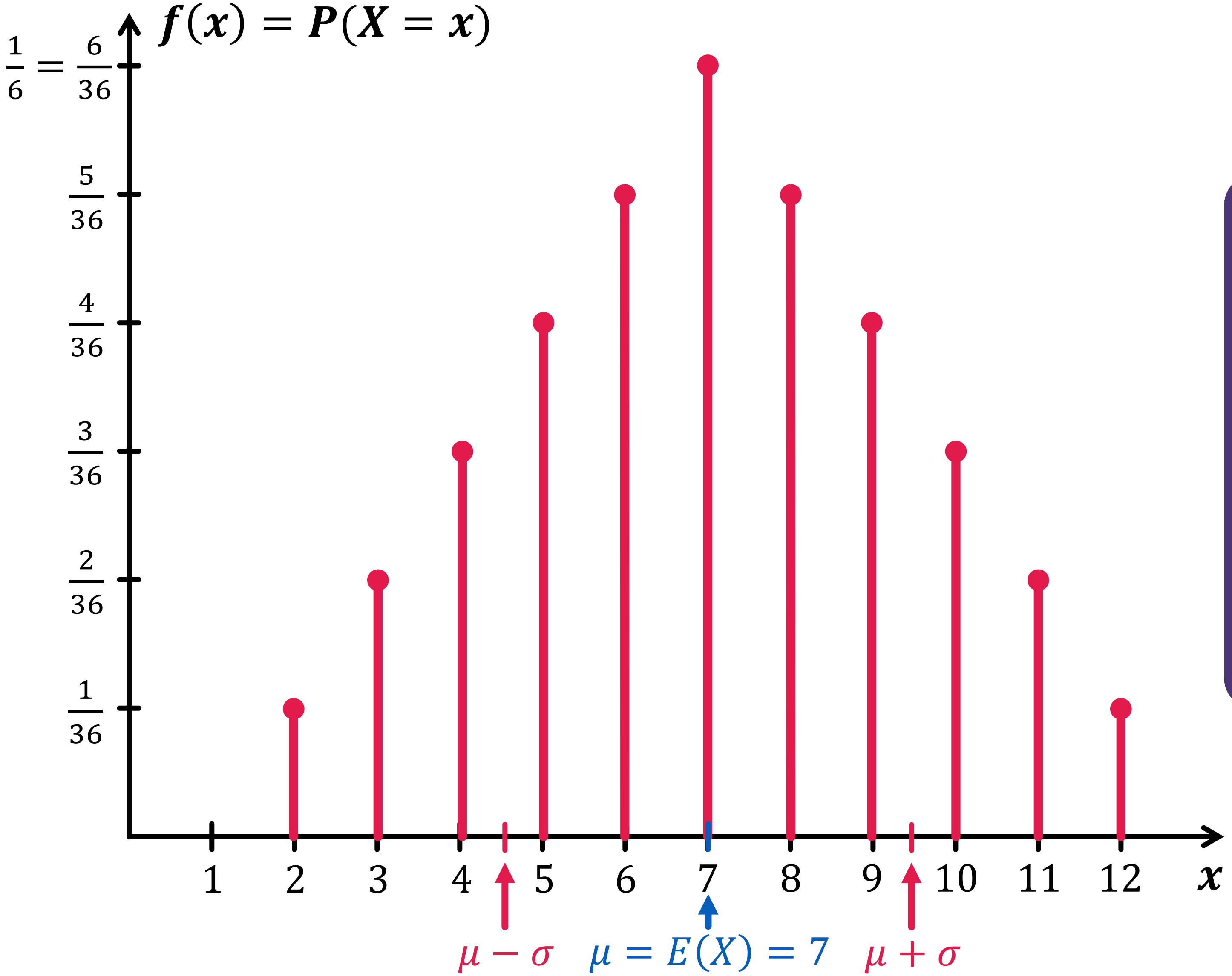
Für eine Zufallsvariable $X\colon \Omega \to \mathbb{R}$ mit der Wertemenge $\{x_1, \dots, x_k\}$ und dem Erwartungswert $E(X) = \mu$ heißt

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

$$= \sqrt{(x_1 - \mu)^2 \cdot P(X = x_1) + \cdots + (x_k - \mu)^2 \cdot P(X = x_k)}$$

die **Standardabweichung** von X.

Streuungsmaß: Standardabweichung



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

Varianz σ^2

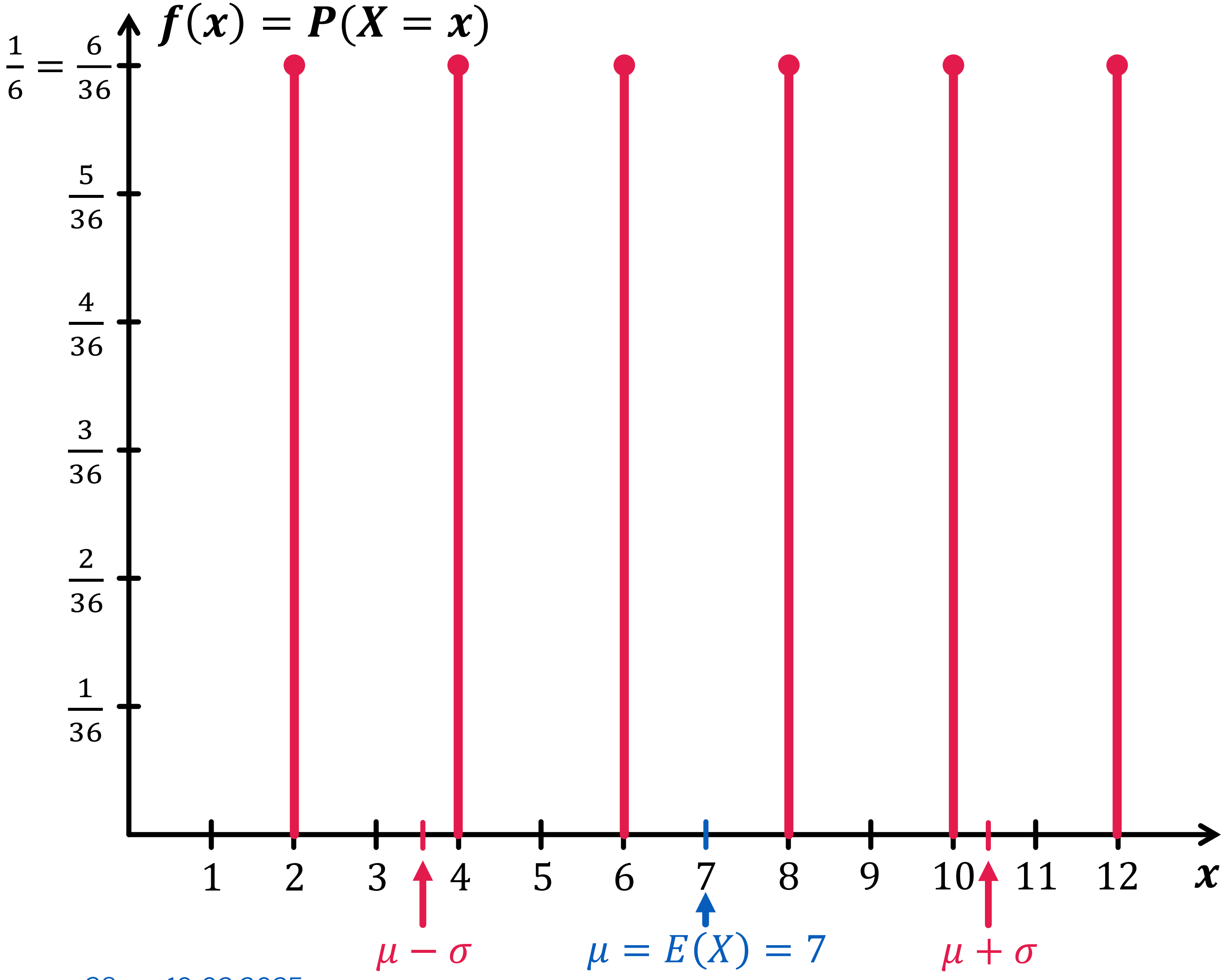
$$\begin{aligned}\sigma^2 &= \mathrm{Var}(X)\\ &= (2-7)^2\cdot\frac{1}{36} + (3-7)^2\cdot\frac{2}{36} + (4-7)^2\cdot\frac{3}{36}\\ &\quad + (5-7)^2\cdot\frac{4}{36} + (6-7)^2\cdot\frac{5}{36} + (7-7)^2\cdot\frac{6}{36}\\ &\quad + (8-7)^2\cdot\frac{5}{36} + (9-7)^2\cdot\frac{4}{36} + (10-7)^2\cdot\frac{3}{36}\\ &\quad + (11-7)^2\cdot\frac{2}{36} + (12-7)^2\cdot\frac{1}{36}\\ &= \frac{210}{36} \approx 5{,}8333\end{aligned}$$

Standardabweichung σ

$$\sigma = \sqrt{\mathrm{Var}(X)} = \sqrt{\frac{210}{36}} \approx 2{,}4152$$

https://www.geogebra.org/m/zahmzfqx

Streuungsmaß: Standardabweichung



x	2	4	6	8	10	12
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Varianz σ^2

$$\begin{aligned}\sigma^2 &= \mathrm{Var}(X)\\ &= (2-7)^2 \cdot \frac{1}{6} + (4-7)^2 \cdot \frac{1}{6}\\ &\quad + (6-7)^2 \cdot \frac{1}{6} + (8-7)^2 \cdot \frac{1}{6}\\ &\quad + (10-7)^2 \cdot \frac{1}{6} + (12-7)^2 \cdot \frac{1}{6}\\ &= \frac{70}{6} \approx 11{,}6667\end{aligned}$$

Standardabweichung σ

$$\sigma = \sqrt{\mathrm{Var}(X)} = \sqrt{\frac{70}{6}} \approx 3{,}4157$$

https://www.geogebra.org/m/zahmzfqx

Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

MaTeGnu

Kapitel 4: Beurteilende Statistik

GeoGebra-Buch „Didaktik der Stochastik“ https://roth.tel/stochastik

Kernideen im Fokus bei Wahrscheinlichkeits-verteilung und Dichte- bzw. Verteilungsfunktion

RPTU Didaktik der Mathematik Sekundarstufen

Kernidee
Denkfigur: Bedingte Wahrscheinlichkeit

- Konditionales Denken
  (Wenn ..., dann ...)
- Schnitt-Wahrscheinlichkeit
  ↔ totale Wahrscheinlichkeit

Kernidee
hypothetisch-prognostische Wahrscheinlichkeit

- Konzept des Bezweifelns
- Modell (Wahrscheinlichkeit)
  ↔ Realität (rel. Häufigkeit)
- aus Erfahrung wird Erwartung

Kernidee
Repräsentativität und Variabilität von Stichproben-Daten

- Repräsentative Stichprobe
- Wachsende Stichprobe
  → reduziert Variabilität

Kernidee
Verteilungen

- Struktur erfassen
  (Form, Streuung, Maximum, ...)
- Gesetz der großen Zahlen
- Sichtweise: lokal ↔ global

Lehrplan RLP: Leistungskurs

	Ziele / Inhalte (Sach- & Methodenkompetenz)	**Hinweise zur Unterrichtsgestaltung und Methodenkompetenz**
8.	Die Begriffe „Bernoullikette“, „Binomialverteilung“ verstehen und die Formel zur Berechnung der Werte einer Binomialverteilung herleiten (5.04g)	Die explizite Berechnung von Werten der Binomialverteilung soll nur exemplarisch mit wenigen Stufen durchgeführt werden. Beim Lösen von Anwendungsaufgaben werden i.d.R. Rechner benutzt. Anknüpfend an vorangegangene Erfahrungen bietet es sich auch an, Bernoulliketten zu simulieren und Werte der Binomialverteilung auf diese Weise zu bestimmen.
9.	Die Formeln für Erwartungswert und Standardabweichung einer Binomialverteilung kennen und anwenden (5.04g)	Die Formeln können anhand von Plausibilitätsbetrachtungen gewonnen werden; ein Beweis ist nicht gefordert.
10.	Eigenschaften der Binomialverteilung kennen, begründen und anwenden (5.04g)	Zur Veranschaulichung charakteristischer Eigenschaften ist die Darstellung von Binomialverteilungen durch Histogramme hilfreich; der Einsatz geeigneter Computerprogramme wird empfohlen.

Lehrplan RLP: Leistungskurs

RPTU Didaktik der Mathematik Sekundarstufen

	Ziele / Inhalte (Sach- & Methodenkompetenz)	**Hinweise zur Unterrichtsgestaltung und Methodenkompetenz**
11.	Sachaufgaben zur Binomialverteilung lösen (5.04g, 4.12g)	Die SuS sollen erfahren, dass viele Zufallsexperimente im täglichen Leben durch eine Binomialverteilung ausreichend gut modelliert werden können.
12.	Verstehen, wie man Wahrscheinlichkeiten einer binomialverteilten Zufallsgröße näherungsweise mit Hilfe der Gaußschen Integralfunktion Φ (Standard-Normalverteilung) bestimmt (5.09e, 5.10e)	Die Möglichkeit der Approximation soll anschaulich, z.B. anhand von Histogrammen, einsichtig gemacht werden. Hierfür empfiehlt sich der Einsatz eines geeigneten Computerprogramms. Die Bestimmung der Näherungswerte erfolgt mit Hilfe von Tabellen oder Rechnern. Im Rahmen des pädagogischen Freiraums können darauf aufbauend die Normalverteilung definiert und Anwendungsbeispiele behandelt werden.
13.	Funktionsterm, Graph und Eigenschaften der Gaußfunktion Φ kennen (5.09e, 5.10e)	Zur Veranschaulichung charakteristischer Eigenschaften ist die Darstellung von Binomialverteilungen durch Histogramme hilfreich; der Einsatz geeigneter Computerprogramme wird empfohlen.

Lehrplan RLP: Grundkurs
Stochastik 1

	Ziele / Inhalte (Sach- & Methodenkompetenz)	Hinweise zur Unterrichtsgestaltung und Methodenkompetenz
6.	Die Begriffe „Bernoullikette“ und „Binomialverteilung“ verstehen und wissen, wie man die Werte einer Binomialverteilung bestimmen kann (5.04g)	Eine explizite Berechnung der Werte der Binomialverteilung soll nur exemplarisch mit wenigen Stufen durchgeführt werden. Die Herleitung der entsprechenden Formel ist nicht gefordert. Beim Lösen von Anwendungsaufgaben werden Tabellen für die Binomialverteilung benutzt. Binomialverteilungen sollen auch grafisch dargestellt werden (Histogramme).
7.	Sachaufgaben zur Binomial-verteilung lösen (5.04g)	Die Formeln sollen anhand von Plausibilitätsbetrachtungen gewonnen werden. Bei der Anwendung in Sachaufgaben kommt es vor allem darauf an, dass die SuS verstehen, welche Folgerungen man aus den Kennwerten für das Sachproblem ziehen kann. Hier bietet es sich an, exemplarisch eine statistische Erhebung zu planen und zu beurteilen.

 MBWWK (2015). Lehrplan Mathematik – Grund- und Leistungsfach in der gymnasialen Oberstufe (Mainzer Studienstufe), S. 34

Wahrscheinlichkeitsverteilung und Verteilungsfunktion

Bemerkung

- $P(X = x)$ ist die Wahrscheinlichkeit des Ereignisses, das zum Wert x der Zufallsvariable X gehört.
- Sprechweise: „$P(X = x)$ ist die Wahrscheinlichkeit, dass die Zufallsvariable X den Wert x annimmt."
- Notation
  $P(X = x) := P(\{X = x\}) = P(\{\omega \in \Omega \mid X(\omega) = x\})$

Definition Wahrscheinlichkeitsverteilung

Ist X eine Zufallsvariable, dann heißt $f\colon \mathbb{R} \to [0;1], x \mapsto f(x)$ mit

$$\begin{aligned} f(x) &:= P(X = x) \\ &= P(\{X = x\}) = P(\{\omega \in \Omega \mid X(\omega) = x\}) \end{aligned}$$

Wahrscheinlichkeitsverteilung der Zufallsvariable X.

Definition Verteilungsfunktion

Ist X eine Zufallsvariable, dann heißt $F\colon \mathbb{R} \to [0;1], x \mapsto F(x)$ mit

$$\begin{aligned} F(x) &:= P(X \leq x) \\ &= P(\{X \leq x\}) = P(\{\omega \in \Omega \mid X(\omega) \leq x\}) \end{aligned}$$

Verteilungsfunktion der Zufallsvariable X.

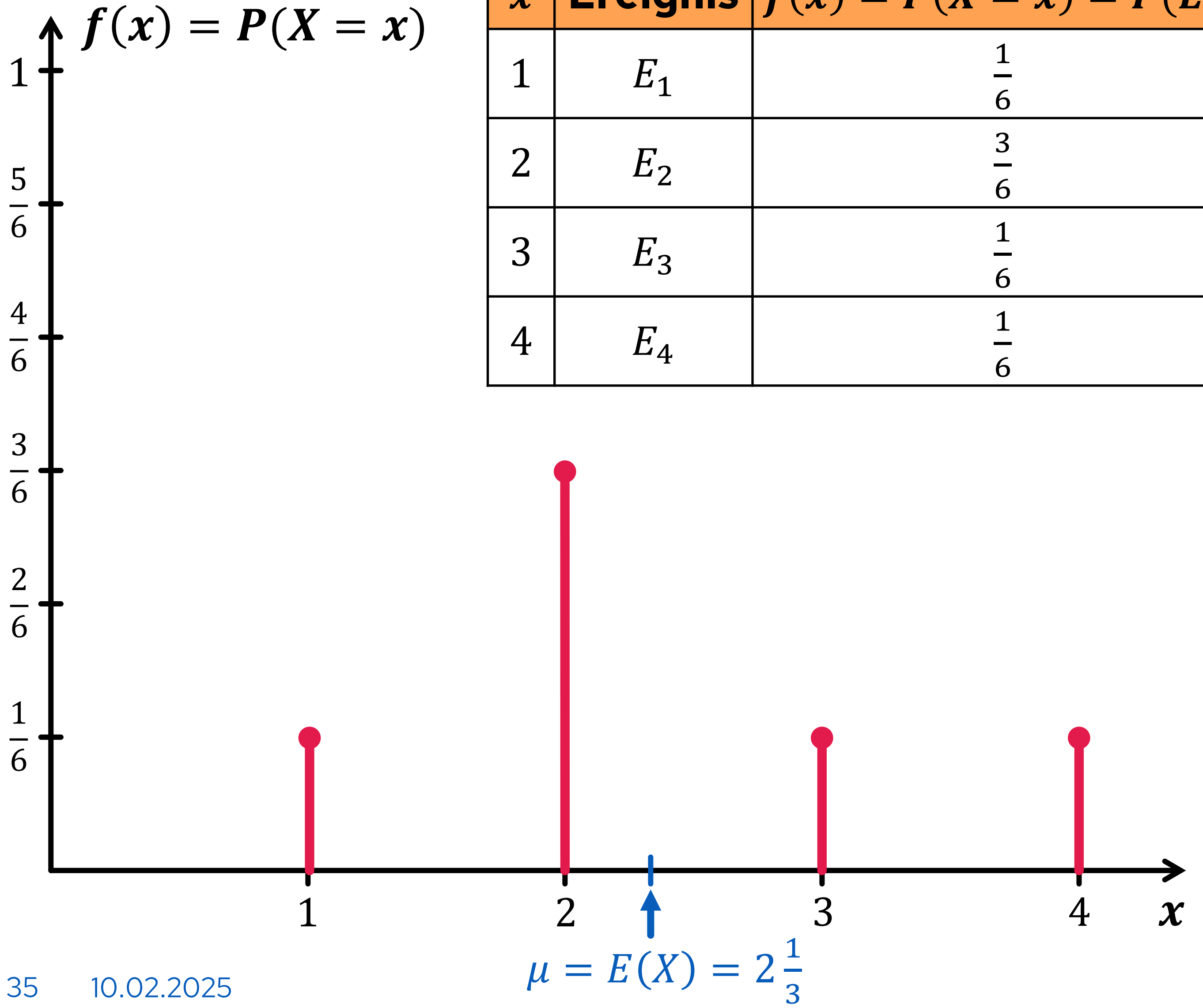
Bemerkung

Konkret werden durch eine Wahrscheinlichkeitsverteilung die Wahrscheinlichkeiten so auf disjunkte Ereignisse E_1 bis E_n verteilt, dass gilt:

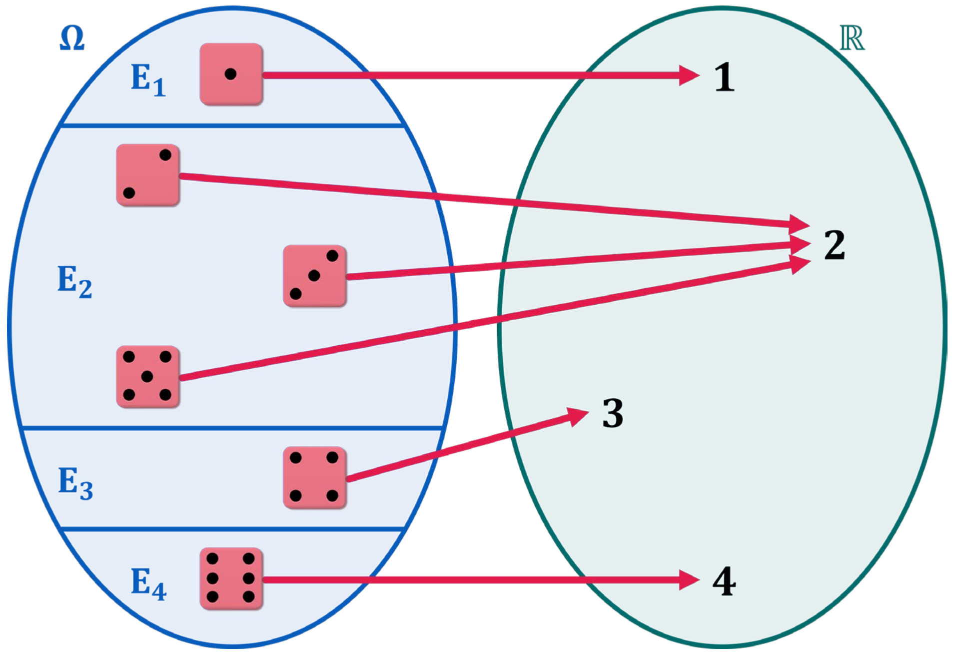
(1) $E_1 \cup \dots \cup E_n = \Omega$

(2) $P(E_1) + \dots + P(E_n) = 1$

Wahrscheinlichkeitsverteilung

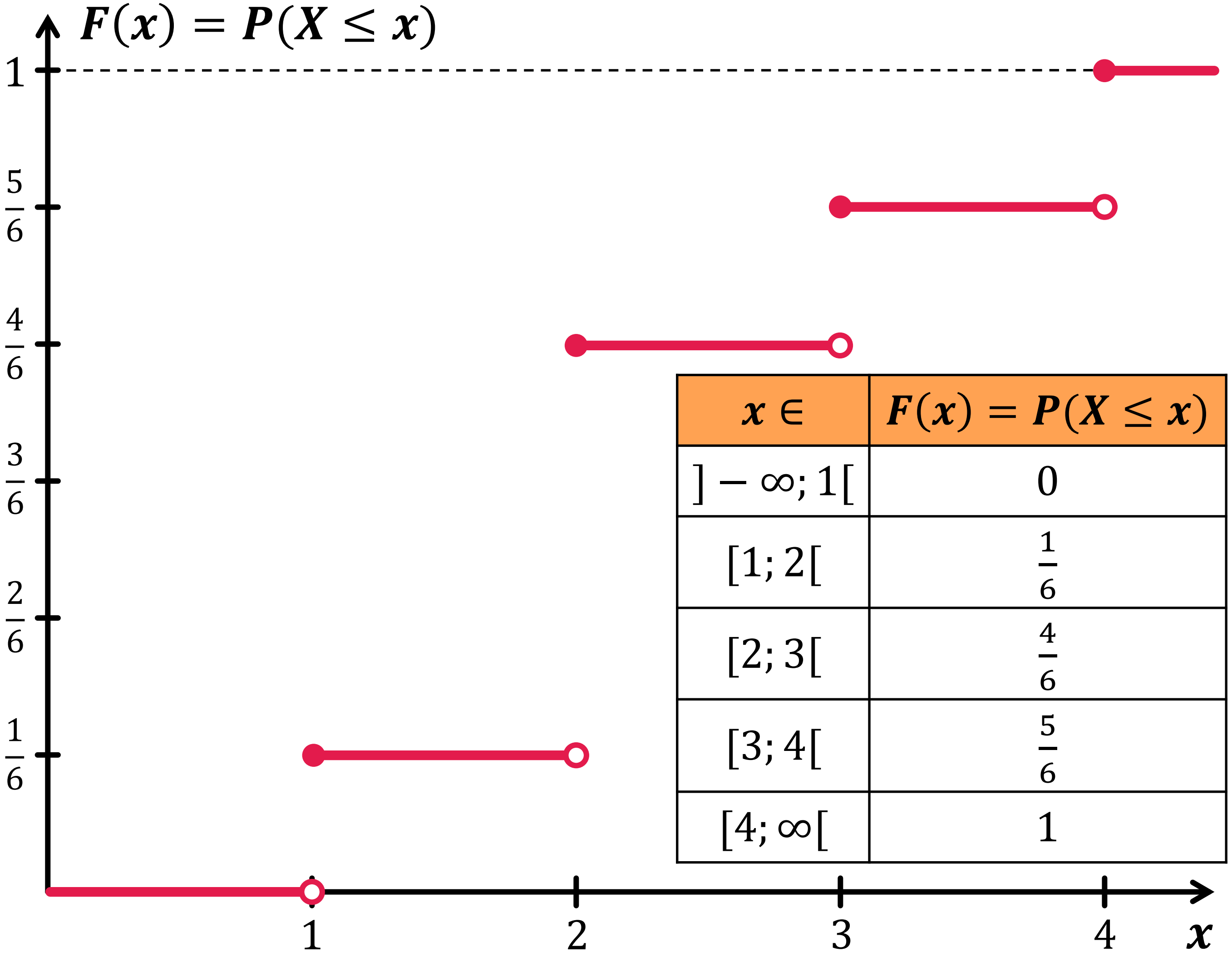


x	Ereignis	$f(x) = P(X = x) = P(E_x)$
1	E_1	$\frac{1}{6}$
2	E_2	$\frac{3}{6}$
3	E_3	$\frac{1}{6}$
4	E_4	$\frac{1}{6}$

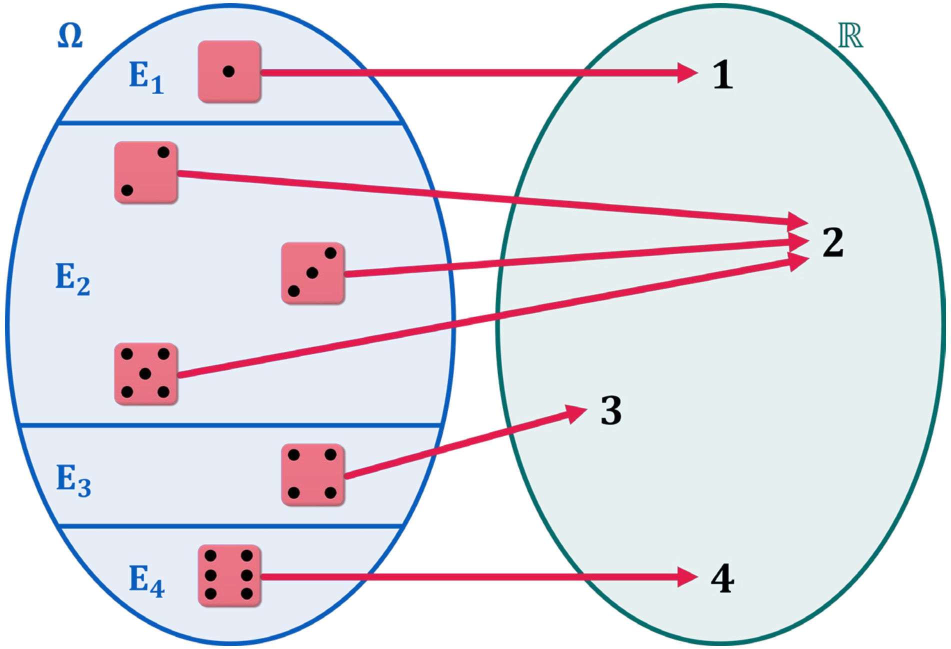


Beispiel: Würfeln mit einem Würfel
Die diskrete Zufallsvariable X ordnet jedem Ergebnis ω die Anzahl x der Teiler der geworfenen Augenzahl zu: $X: \Omega \to \mathbb{R}, \omega \mapsto x$

Verteilungsfunktion

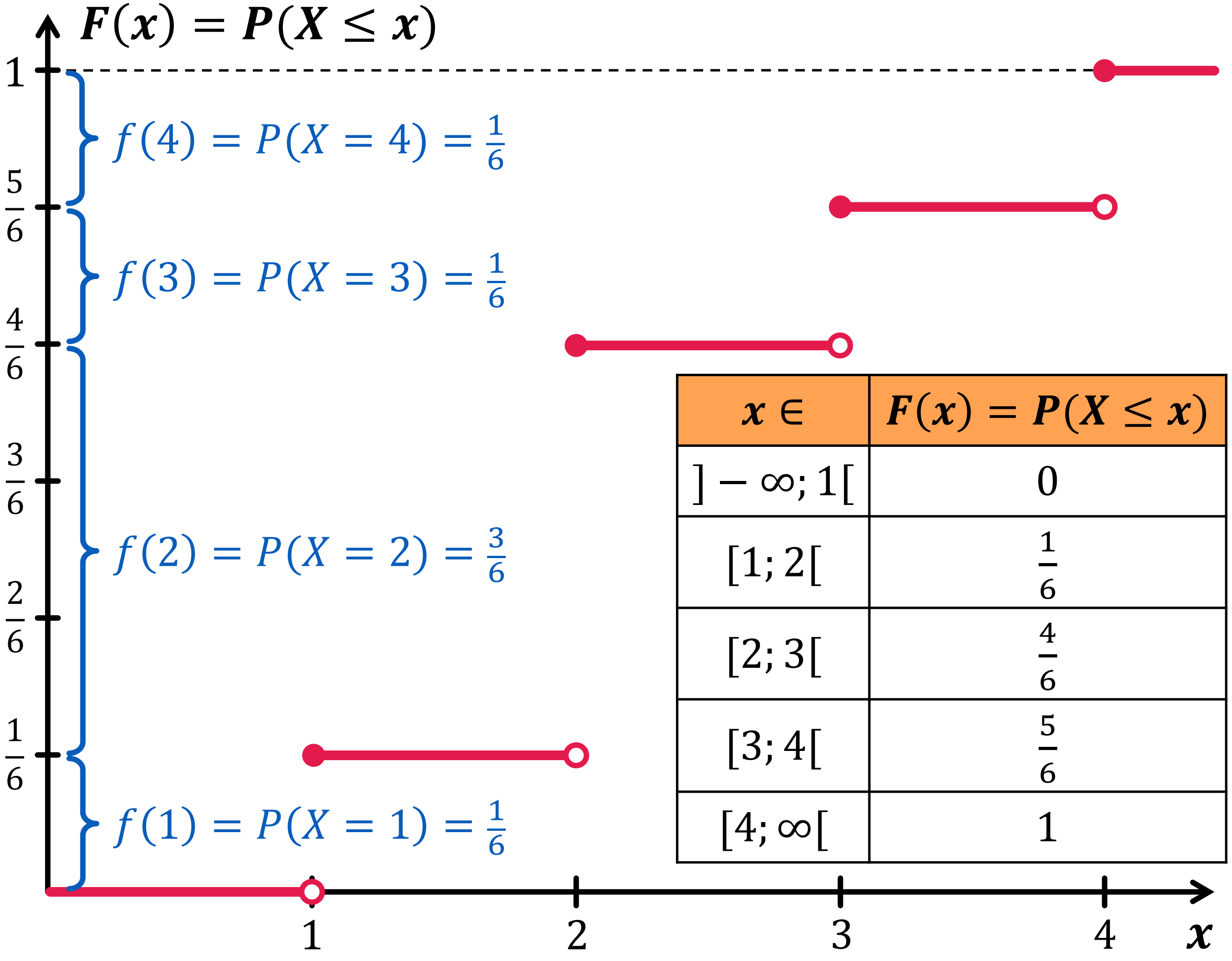


$x \in$	$F(x) = P(X \leq x)$
$]-\infty; 1[$	0
$[1; 2[$	$\frac{1}{6}$
$[2; 3[$	$\frac{4}{6}$
$[3; 4[$	$\frac{5}{6}$
$[4; \infty[$	1

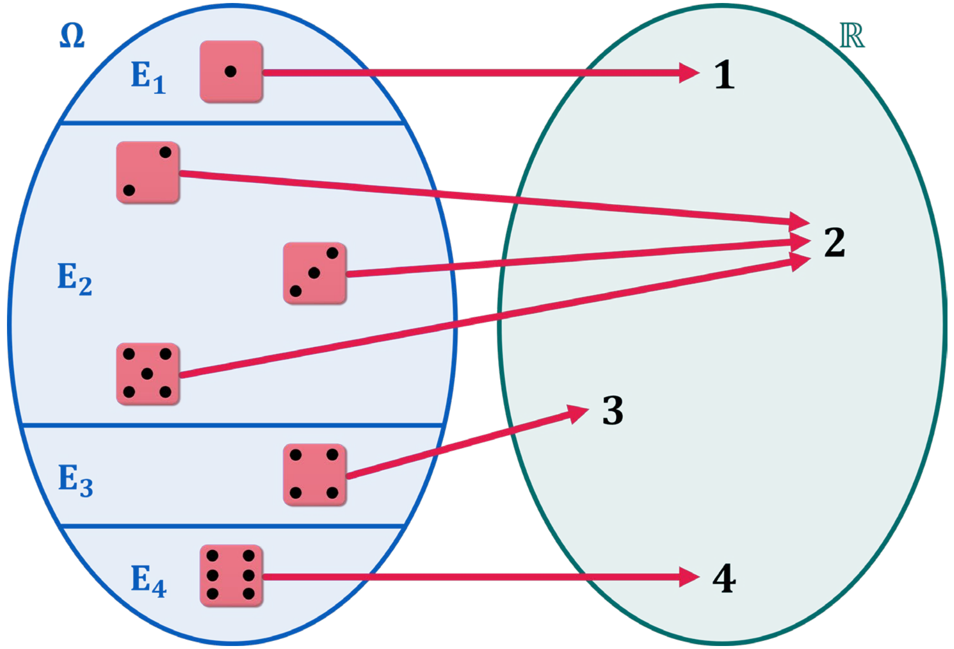


Beispiel: Würfeln mit einem Würfel
Die diskrete Zufallsvariable X ordnet jedem Ergebnis ω die Anzahl x der Teiler der geworfenen Augenzahl zu: $X: \Omega \to \mathbb{R}, \omega \mapsto x$

Verteilungsfunktion



$x \in$	$F(x) = P(X \leq x)$
$]-\infty; 1[$	0
$[1; 2[$	$\frac{1}{6}$
$[2; 3[$	$\frac{4}{6}$
$[3; 4[$	$\frac{5}{6}$
$[4; \infty[$	1



Eigenschaften von Verteilungsfunktionen

- $P(X > a) = 1 - F(a)$
- $P(a < X \leq b) = F(b) - F(a)$
- Bei diskreten Zufallsvariablen $P(X = x_{i+1}) = F(x_{i+1}) - F(x_i)$

RPTU

Binomialverteilung: Voraussetzungen

Bernoulli-Experiment [1]

- Ein Zufallsexperiment bei dem nur interessiert, ob ein bestimmter Ausgang vorliegt oder nicht, man also nur die beiden Ergebnisse „Treffer" oder „Niete" hat, nennt man **Bernoulli-Experiment**.
- Ein passender Ergebnisraum ist $\Omega = \{0,1\}$ mit der Trefferwahrscheinlichkeit $P(\{1\}) = p$ und der Nietenwahrscheinlichkeit $P(\{0\}) = 1 - p = q$.

Bernoulli-Kette: Ein n-stufiges Zufallsexperiment mit Ergebnisraum $\Omega = \{0,1\}^n$, bei dem (für $i = 1, \dots, n$) die Ereignisse

E_i: „Menge aller n-Tupel mit der 1 an der i-ten Stelle" („Treffer in der i-ten Stufe des Experiments")

- unabhängig sind und
- die gleiche Wahrscheinlichkeit $P(E_i) = p$ haben,

heißt **Bernoulli-Kette** mit den Parametern n und p. Der Parameter n heißt Länge der Bernoulli-Kette.

Aufgabe: Ist die Voraussetzung für eine Bernoulli-Kette jeweils erfüllt?

a) Eine Person, die bisher nicht Basketball gespielt hat, wirft zehnmal hintereinander auf einen Basketballkorb.

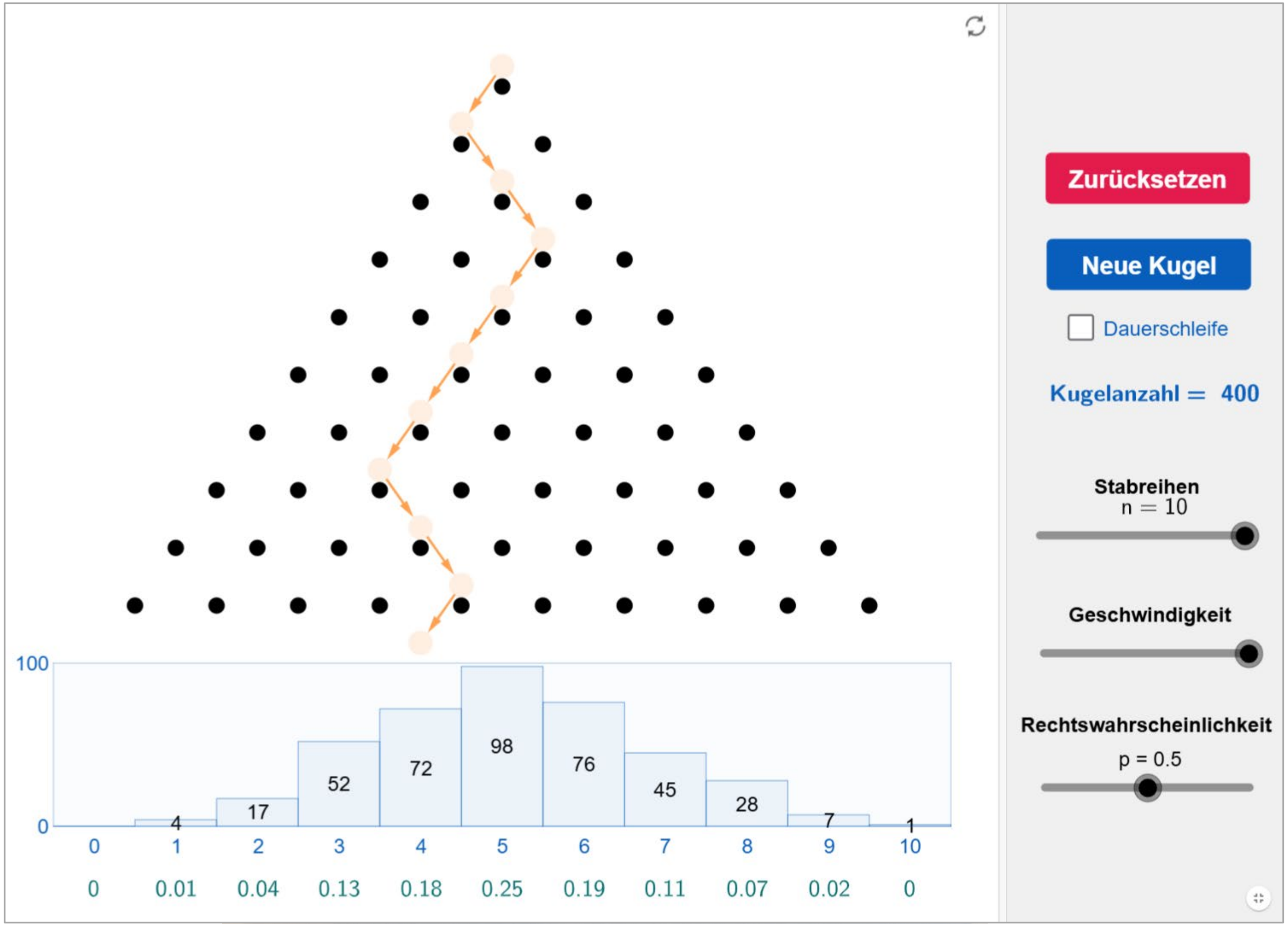
b) Zehn zufällig gewählte Personen werfen je einmal auf einen Basketballkorb.

c) Eine Person wirft während zehn Wochen immer Montags in der Pause einmal auf einen Basketballkorb.

1 Benannt nach Jacob Bernoulli (*06.01.1655, †16.08.1705) • https://de.wikipedia.org/wiki/Jakob_I_Bernoulli

Galton-Brett

https://phet.colorado.edu/sims/html/plinko-probability/latest/plinko-probability_all.html?locale=de



 https://www.geogebra.org/m/wn7enhda

Herleitung: Wahrscheinlichkeit $B(n;p;k)$ für k Treffer in einer Bernoulli-Kette der Länge n

Didaktik der Mathematik Sekundarstufen

$$B(n;p;k) = \underbrace{\binom{n}{k}}_{} \cdot \underbrace{p^k \cdot (1-p)^{n-k}}_{}$$

Summenregel: Anzahl aller Pfade mit k Treffern

Produktregel: Wahrscheinlichkeit eines Pfades der Länge n mit k Treffern

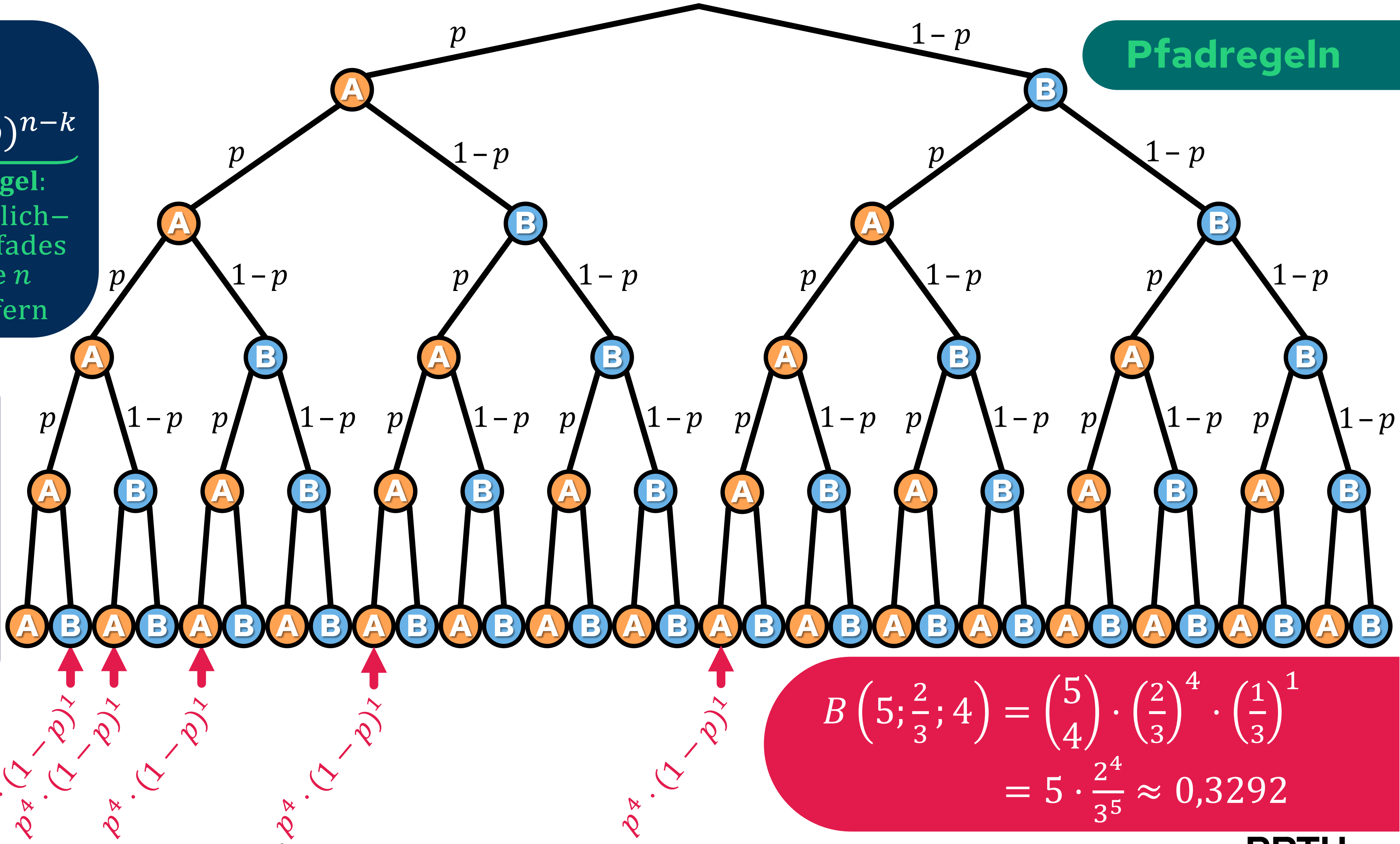
Länge der Kette: $n = 5$

Treffer: A

Trefferanzahl: $k = 4$

Trefferwahrscheinlichkeit: $p = \frac{2}{3}$

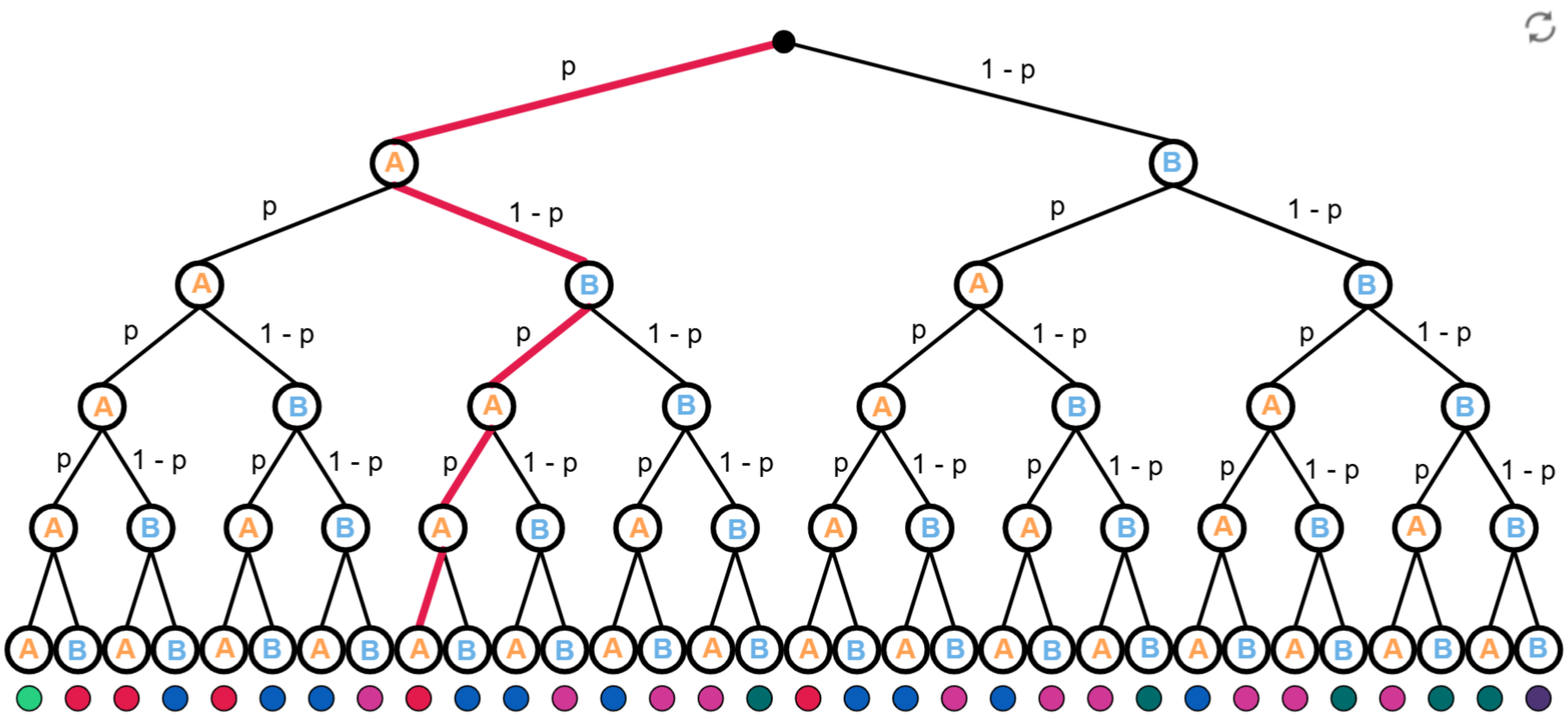
Pfadregeln

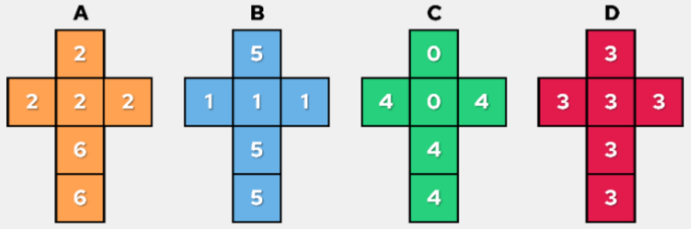


$$B\left(5;\frac{2}{3};4\right) = \binom{5}{4} \cdot \left(\frac{2}{3}\right)^4 \cdot \left(\frac{1}{3}\right)^1 = 5 \cdot \frac{2^4}{3^5} \approx 0{,}3292$$

Herleitung: Wahrscheinlichkeit $B(n; p; k)$ für k Treffer in einer Bernoulli-Kette der Länge n

RPTU Didaktik der Mathematik Sekundarstufen





$p^5 \cdot (1-p)^0 \approx 0.132$

$p^4 \cdot (1-p)^1 \approx 0.066$

$p^3 \cdot (1-p)^2 \approx 0.033$

$p^2 \cdot (1-p)^3 \approx 0.016$

$p^1 \cdot (1-p)^4 \approx 0.008$

$p^0 \cdot (1-p)^5 \approx 0.004$

$$p = \frac{2}{3}$$

Würfelauswahl

Reset

A und B | A und C | A und D

B und C | B und D | C und D

Anzahl der Stufen = 5

Pfad

 https://www.geogebra.org/m/m6zg8yaw

Binomialverteilung

Satz: Trefferwahrscheinlichkeit in einer Bernoulli-Kette [1]

Für ein Bernoulli-Experiment mit der Trefferwahrscheinlichkeit $p \in \,]0;1[$ ist die Wahrscheinlichkeit $B(n;p;k)$ für k Treffer in einer Bernoulli-Kette der Länge n gegeben durch ($k \in \{0,1,\dots n\}$):

$$B(n;p;k) = P(X=k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Beweisidee
Vgl. Folie 40

Definition: Binomialverteilung

- Die Wahrscheinlichkeitsverteilung

$$B(n;p)\colon k \mapsto B(n;p;k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

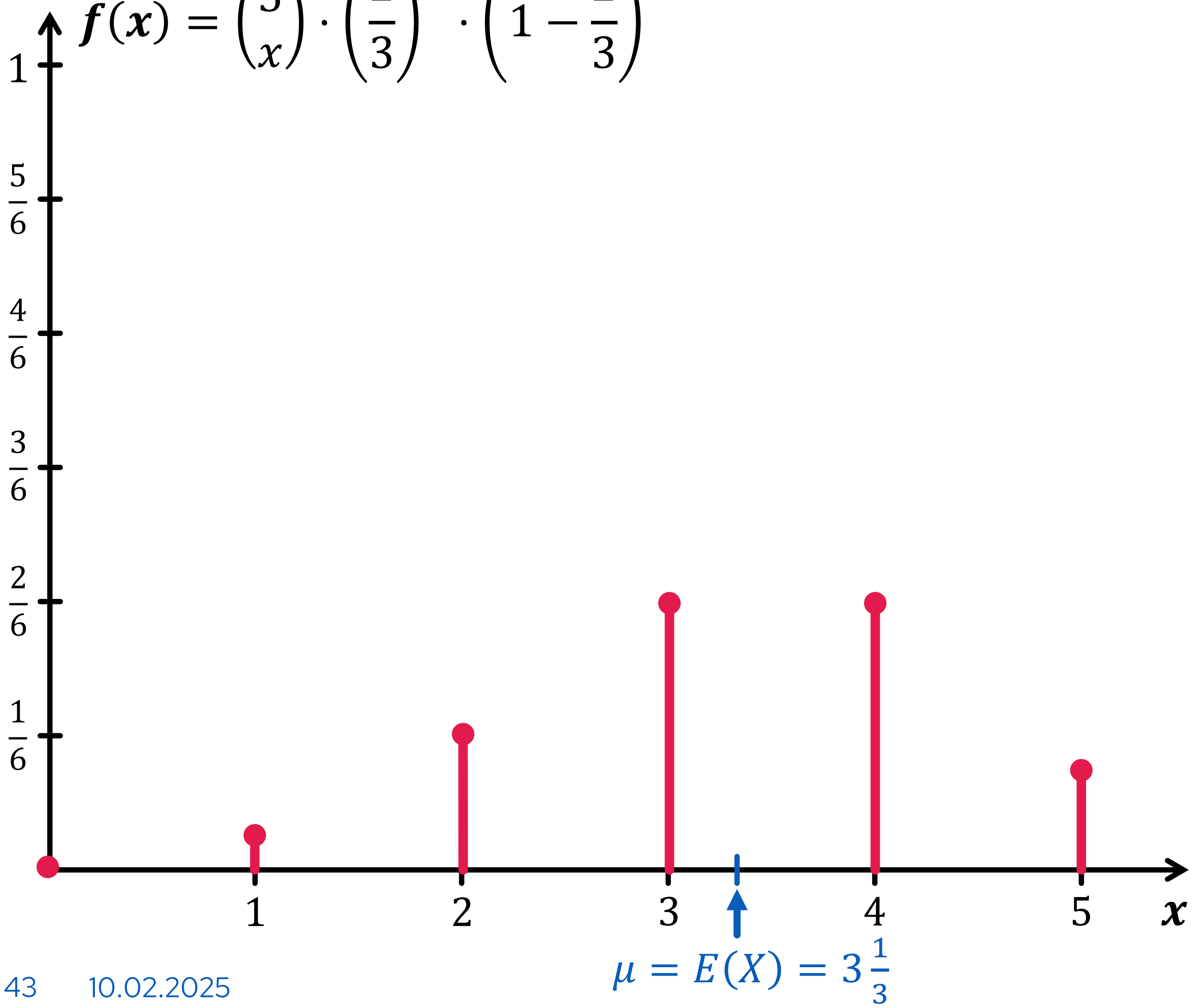
  mit $k \in \{0,1,\dots n\}$ und $p \in \,]0;1[$ heißt **Binomialverteilung** mit den Parametern n und p.
- Eine Zufallsgröße X mit $P(X=k) = B(n;p;k)$ heißt binomial verteilt nach der Binomialverteilung $B(n;p)$.

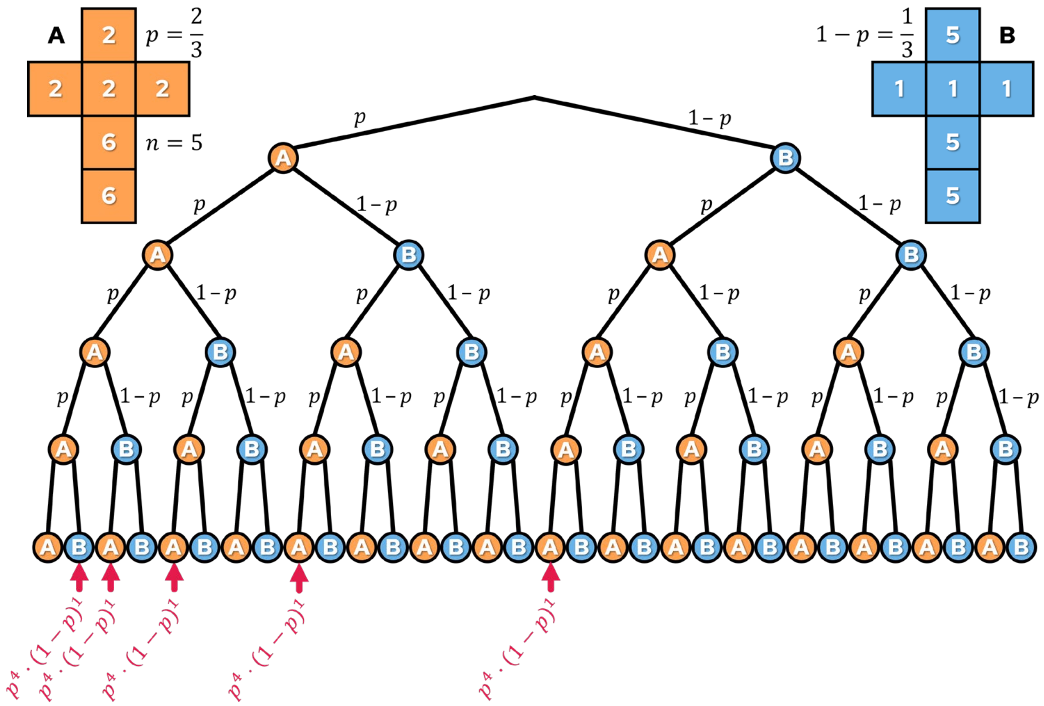
[1] Benannt nach Jacob Bernoulli (*06.01.1655, †16.08.1705) • https://de.wikipedia.org/wiki/Jakob_I_Bernoulli

Wahrscheinlichkeitsverteilung
Beispiel: Bernoulli-Kette → Binomialverteilung

Didaktik der Mathematik Sekundarstufen

$$f(x) = \binom{5}{x} \cdot \left(\frac{2}{3}\right)^{x} \cdot \left(1 - \frac{2}{3}\right)^{n-x}$$





Beispiel: Fünfmaliges Würfeln mit zwei chinesischen Würfeln

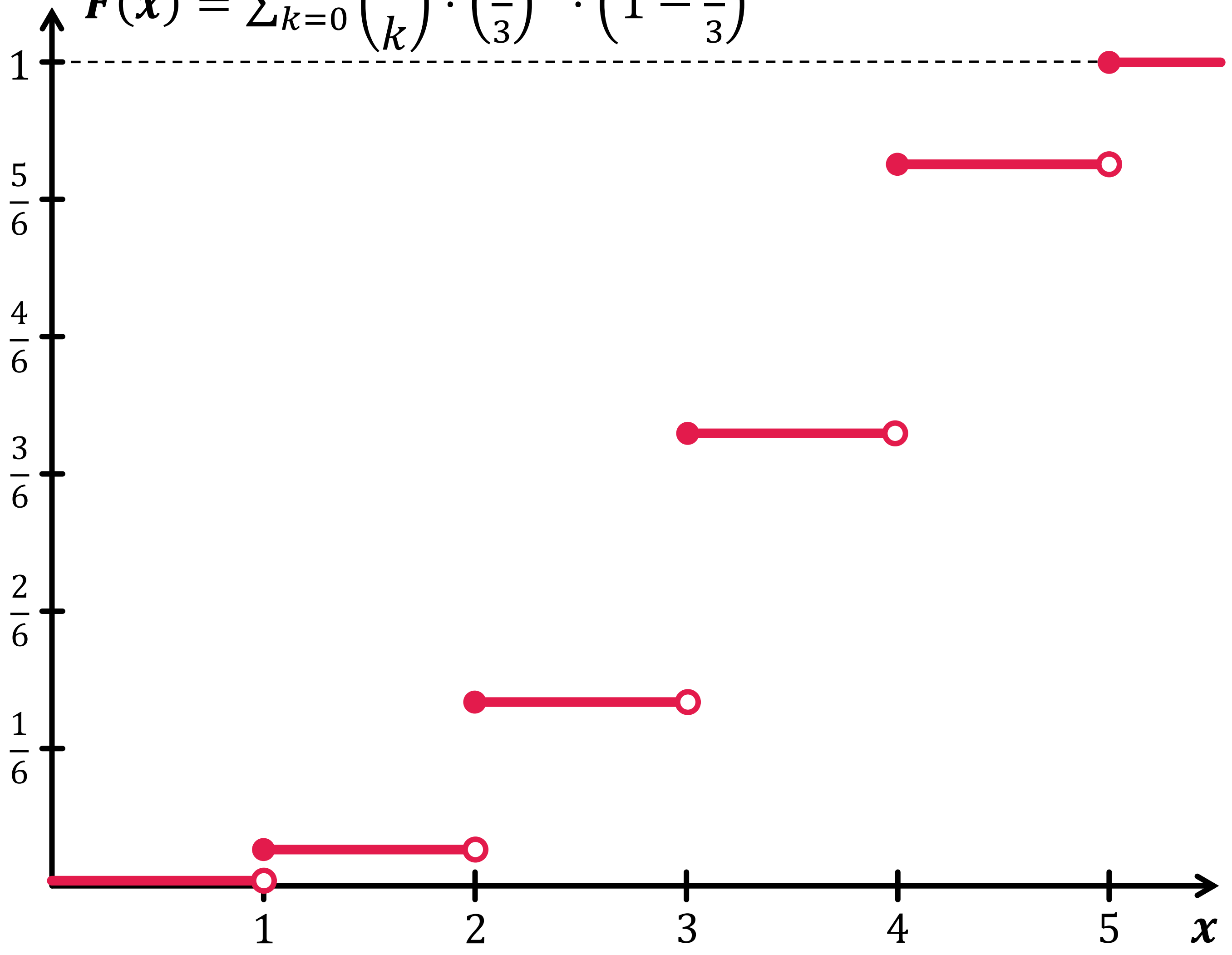
Die diskrete Zufallsvariable X ordnet jedem Ergebnis ω des fünfmaligen Würfelns ($n = 5$) die Anzahl x der Gewinne des orangen Würfels $\left(\text{Gewinnwahrscheinlichkeit } p = \frac{2}{3}\right)$ zu:

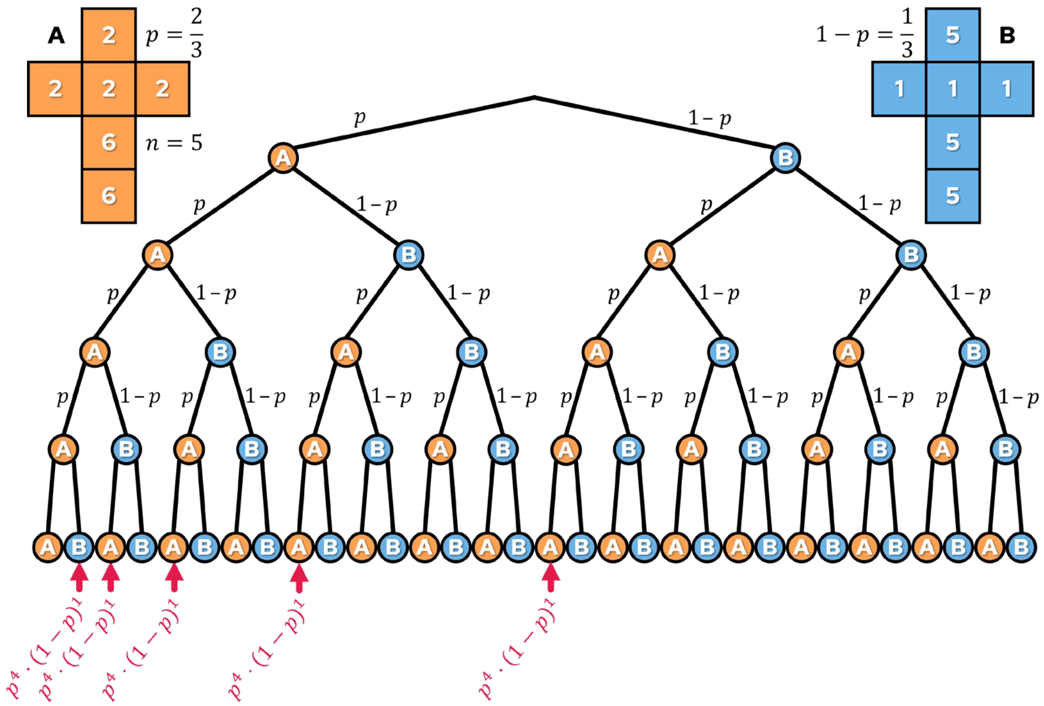
$$X\colon \Omega \to \mathbb{R}, \omega \mapsto x$$

Verteilungsfunktion

Beispiel: Bernoulli-Kette → Binomialverteilung

$$F(x) = \sum_{k=0}^{x} \binom{5}{k} \cdot \left(\frac{2}{3}\right)^k \cdot \left(1-\frac{2}{3}\right)^{n-k}$$





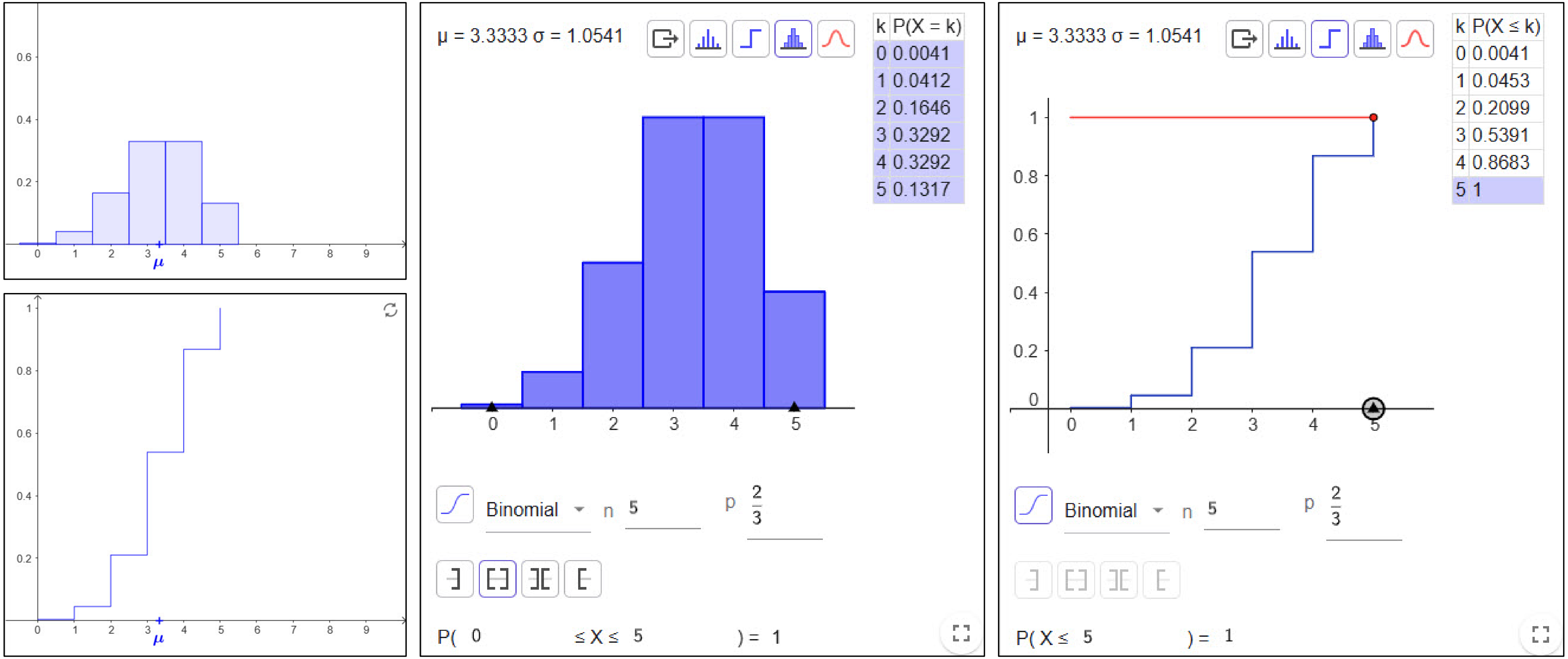
Beispiel: Fünfmaliges Würfeln mit zwei chinesischen Würfeln

Die diskrete Zufallsvariable X ordnet jedem Ergebnis ω des fünfmaligen Würfelns ($n = 5$) die Anzahl x der Gewinne des orangen Würfels $\left(\text{Gewinnwahrscheinlichkeit } p = \frac{2}{3}\right)$ zu:

$$X: \Omega \to \mathbb{R}, \omega \mapsto x$$

Binomialverteilung mit GeoGebra

Didaktik der Mathematik Sekundarstufen



Aufgabe: Erkunden Sie Eigenschaften der Binomialverteilung anhand der GeoGebra-Datei.

Binomialverteilung: Erwartungswert

Aufgabe: Bestimmen Sie den Erwartungswert für die Anzahl (absolute Häufigkeit) k der Treffer bei einer Bernoulli-Kette der Länge n und der Trefferwahrscheinlichkeit p.

- Bei einer Bernoulli-Kette der Länge n ordnet die Zufallsvariable X_i dem i-ten Bernoulli-Experiment der Kette den Wert 0 für Niete und den Wert 1 für Treffer mit folgender Wahrscheinlichkeitsverteilung zu:

x_i	0	1
$P(X_i = x_i)$	$1-p$	p

- Die Zufallsgröße $X = X_1 + X_2 + \cdots + X_n$ beschreibt also die Anzahl (absolute Häufigkeit) der Treffer in einer Bernoulli-Kette mit den Parametern n und p, wobei die X_i unabhängig sind.

- Da für die Erwartungswerte $E(X_i)$ gilt

$$\begin{aligned} \boldsymbol{E(X_i)} &= 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1) \\ &= 0 \cdot (1-p) + 1 \cdot p = \boldsymbol{p} \end{aligned}$$

folgt für den **Erwartungswert $E(X)$**:

$$\begin{aligned} \boldsymbol{E(X)} &= E(X_1 + X_2 + \cdots + X_n) \\ &\overset{*}{=} E(X_1) + E(X_2) + \cdots + E(X_n) \\ &= \underbrace{p + p + \cdots + p}_{n \text{ Summanden}} = \boldsymbol{n \cdot p} \end{aligned}$$

Satz: Erwartungswert der binomialverteilten Zufallsvariable „absolute Trefferhäufigkeit“ X

Für den **Erwartungswert $E(X)$** einer $B(n;p)$-verteilten Zufallsvariable „absolute Trefferhäufigkeit“ X gilt: $\boldsymbol{E(X) = n \cdot p}$

Satz: Erwartungswert der binomialverteilten Zufallsvariable „relative Trefferhäufigkeit“ h

Für den **Erwartungswert $E(h)$** einer $B(n;p)$-verteilten Zufallsvariable „relative Trefferhäufigkeit“ h gilt: $\boldsymbol{E(h) = \frac{E(X)}{n} = \frac{n \cdot p}{n} = p}$

* Hinweis: Hier wird die Additivität des Erwartungswerts, also $E(X+Y) = E(X) + E(Y)$, benutzt. Das wäre eigentlich noch zu zeigen.

Binomialverteilung: Standardabweichung

Didaktik der Mathematik Sekundarstufen

Aufgabe: Bestimmen Sie die Standardabweichung für die Anzahl k der Treffer bei einer Bernoulli-Kette der Länge n und der Trefferwahrscheinlichkeit p.

- Die Varianz der Zufallsgröße $X = X_1 + X_2 + \cdots + X_n$ lässt sich ebenfalls über die Varianzwerte der unabhängigen Zufallsgrößen X_i bestimmen:
- Mit $\mu_i = E(X_i) = p$ gilt für die Varianzen $\mathrm{Var}(X_i)$:

$$
\begin{aligned}
\mathbf{Var}(\boldsymbol{X_i}) &= (0-\mu_i)^2 \cdot P(X_i = 0) + (1-\mu_i)^2 \cdot P(X_i = 1) \\
&= (0-p)^2 \cdot (1-p) + (1-p)^2 \cdot p \\
&= p^2 \cdot (1-p) + (1-p)^2 \cdot p \\
&= p \cdot (1-p) \cdot \underbrace{[p + (1-p)]}_{=1} \\
&= \boldsymbol{p \cdot (1-p)}
\end{aligned}
$$

Daraus folgt für die **Varianz $\mathrm{Var}(X)$**:

$$
\begin{aligned}
\mathbf{Var}(\boldsymbol{X}) &= \mathrm{Var}(X_1 + X_2 + \cdots + X_n) \\
&\overset{*}{=} \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n) \\
&= \underbrace{p \cdot (1-p) + \cdots + p \cdot (1-p)}_{n \text{ Summanden}} \\
&= \boldsymbol{n \cdot p \cdot (1-p)}
\end{aligned}
$$

$$\Rightarrow \boldsymbol{\sigma(X) = \sqrt{n \cdot p \cdot (1-p)}}$$

Satz: Standardabweichung einer binomialverteilten Zufallsvariable X

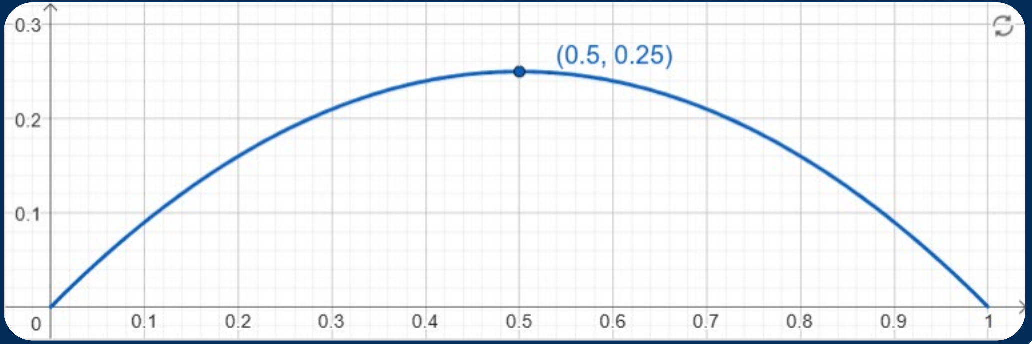
Für die Standardabweichung $\sigma(X)$ einer $B(n;p)$-verteilten Zufallsvariable X gilt: $\boldsymbol{\sigma(X) = \sqrt{n \cdot p \cdot (1-p)}}$

* Hinweis: Hier wird die Additivität der Varianz, also $Var(X+Y) = Var(X) + Var(Y)$, benutzt. Das wäre eigentlich noch zu zeigen.

Binomialverteilung: Standardabweichung

Bemerkung: Für die Standardabweichung $\boldsymbol{\sigma(X) = \sqrt{n \cdot p \cdot (1-p)}}$ einer Binomialverteilung lässt sich eine obere Abschätzung angeben, die unabhängig von p ist:

- Der Term $p \cdot (1-p)$ unter der Wurzel kann wegen $0 \leq p \leq 1$ nicht größer als $\frac{1}{4}$ werden. Man sieht das, wenn man die Funktion $p \mapsto p \cdot (1-p) = -p^2 + p$ im Definitionsbereich $0 \leq p \leq 1$ betrachtet.
- Der Graph dieser quadratischen Funktion ist eine nach unten geöffnete Parabel mit dem Scheitel $\left(\frac{1}{2}, \frac{1}{4}\right)$.
- Damit folgt:
  $\boldsymbol{\sigma(X) = \sqrt{n \cdot p \cdot (1-p)}} \leq \sqrt{n \cdot \frac{1}{4}} = \frac{\sqrt{n}}{2}$

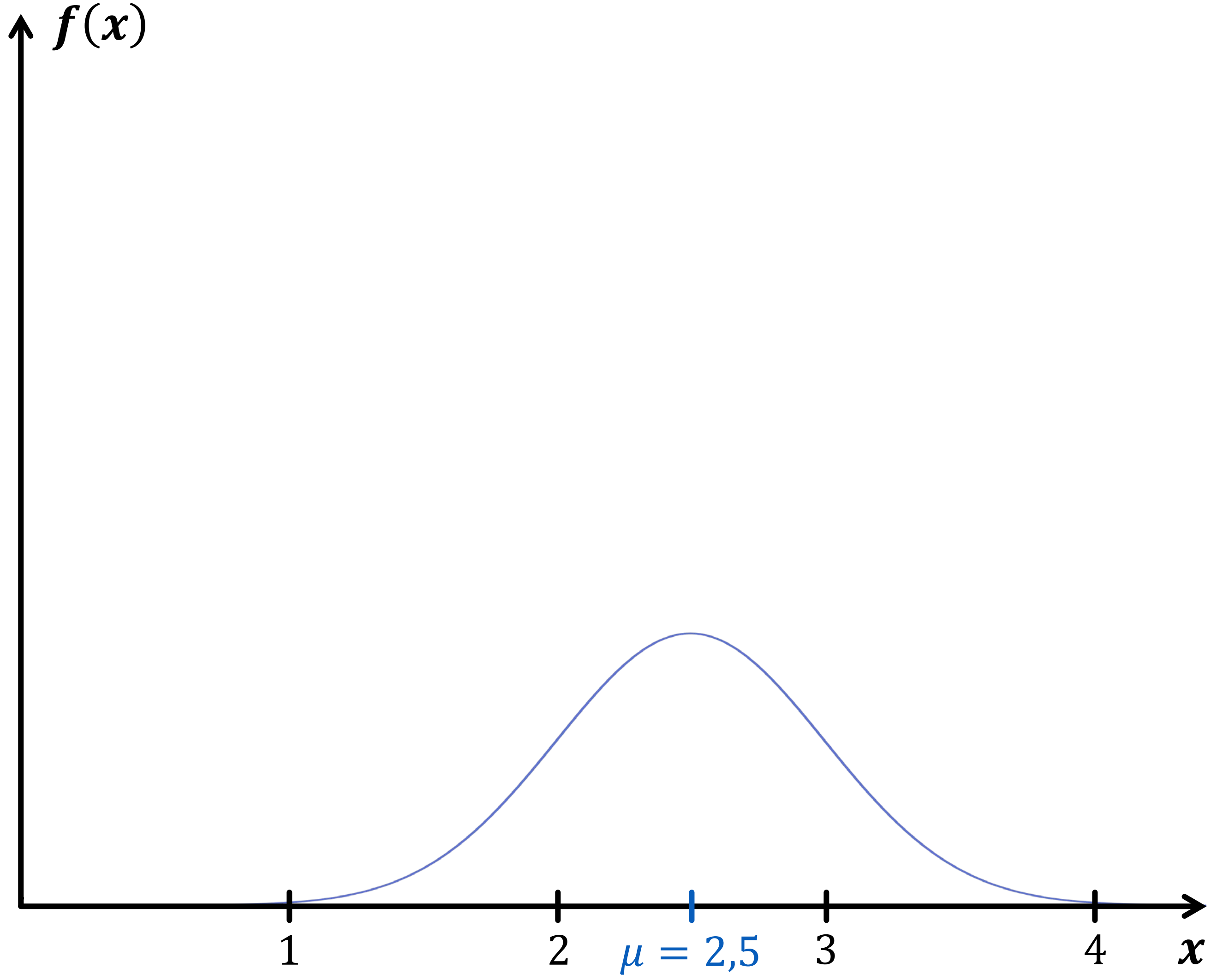


Satz: Standardabweichung einer binomialverteilten Zufallsvariable X

Für die Standardabweichung $\sigma(X)$ einer $B(n; p)$-verteilten Zufallsvariable X gilt:

$$\boldsymbol{\sigma(X) = \sqrt{n \cdot p \cdot (1-p)} \leq \frac{\sqrt{n}}{2}}$$

Normalverteilung: Dichtefunktion



https://youtu.be/cTyPuZ9-JZ0
https://www.geogebra.org/m/vvb5qz4t

Bemerkung

- Hier ist der Graph der Dichtefunktion f einer stetigen Wahrscheinlichkeitsverteilung, konkret einer Normalverteilung
$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$
mit $\mu = 2{,}5$ und $\sigma = 0{,}5$ dargestellt.
- Für eine stetige Wahrscheinlichkeitsverteilung f gilt:
$$\int_{-\infty}^{\infty} f(x)dx = 1$$

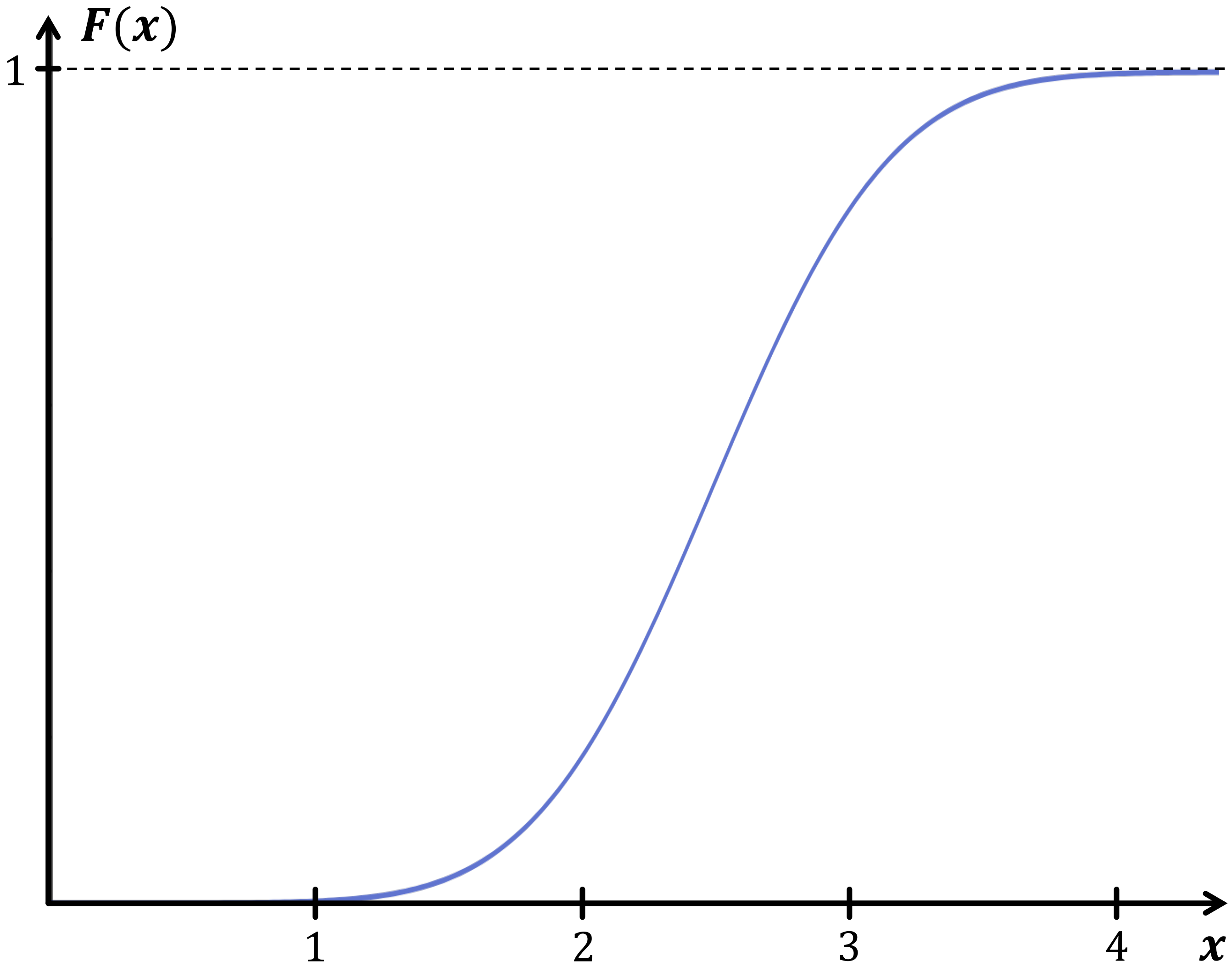
Bemerkung

Eine Normalverteilung mit den Parametern $\mu = 0$ und $\sigma = 1$ heißt auch φ.

Unnötig!

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$

Normalverteilung: Verteilungsfunktion



Bemerkung

- Hier ist der Graph der Verteilungsfunktion F einer Normalverteilung
$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$
mit $\mu = 2{,}5$ und $\sigma = 0{,}5$ dargestellt.
- Für solche Verteilungsfunktionen F gilt:
$$F(x) = \int_{-\infty}^{x} f(t)dt$$

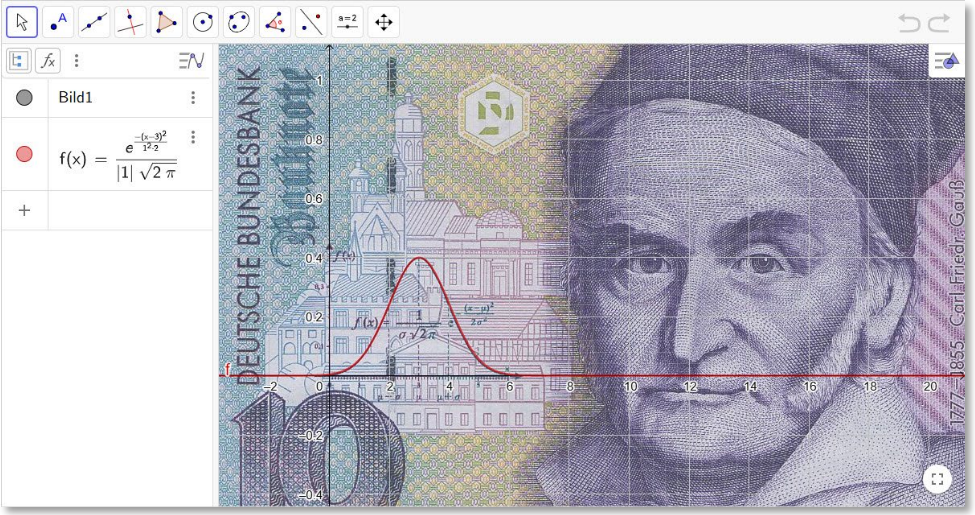
Bemerkung

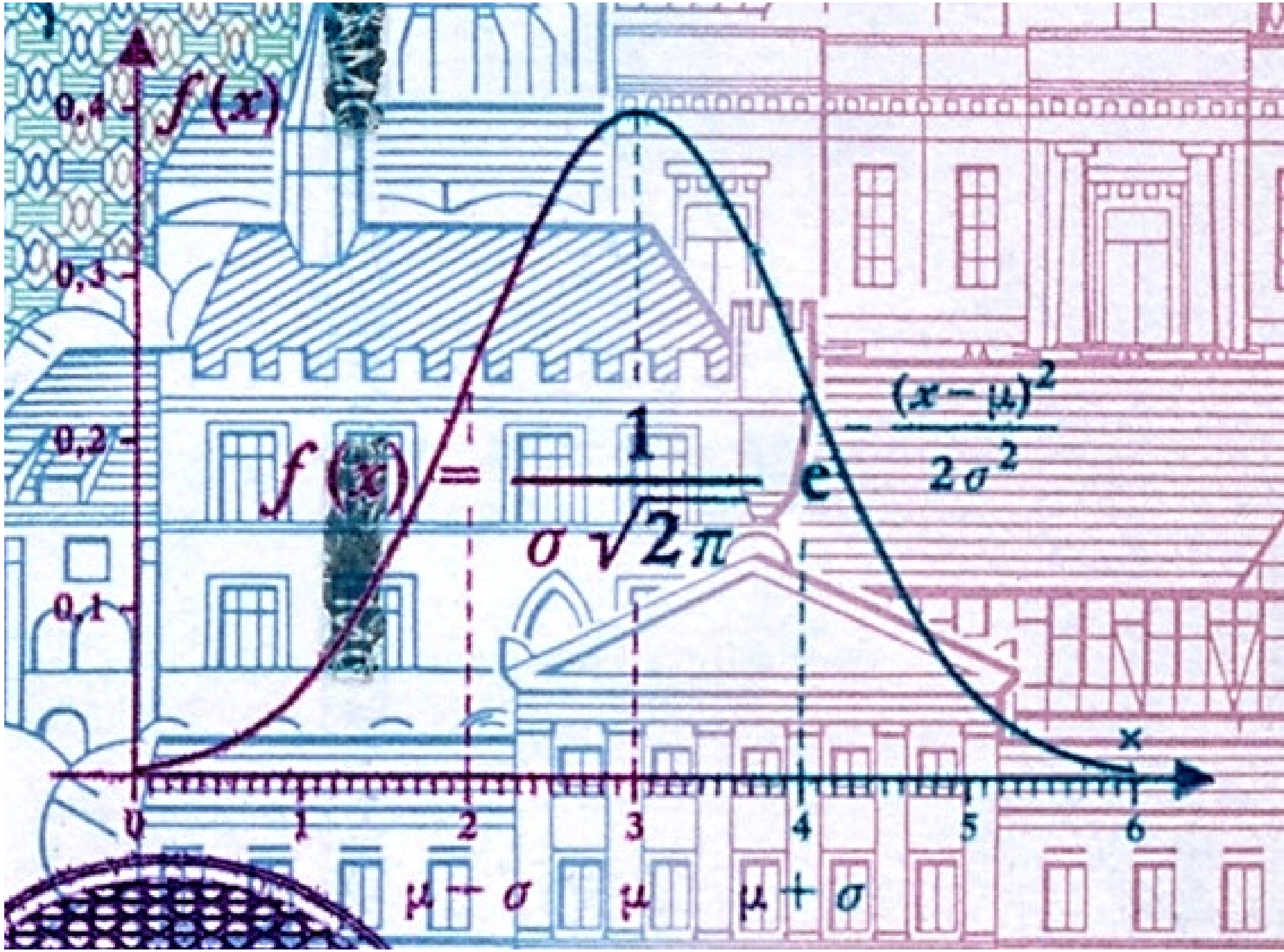
Die Verteilungsfunktion zur Normalverteilung φ mit den Parametern $\mu = 0$ und $\sigma = 1$ heißt auch Φ.

Unnötig!

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)dt = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

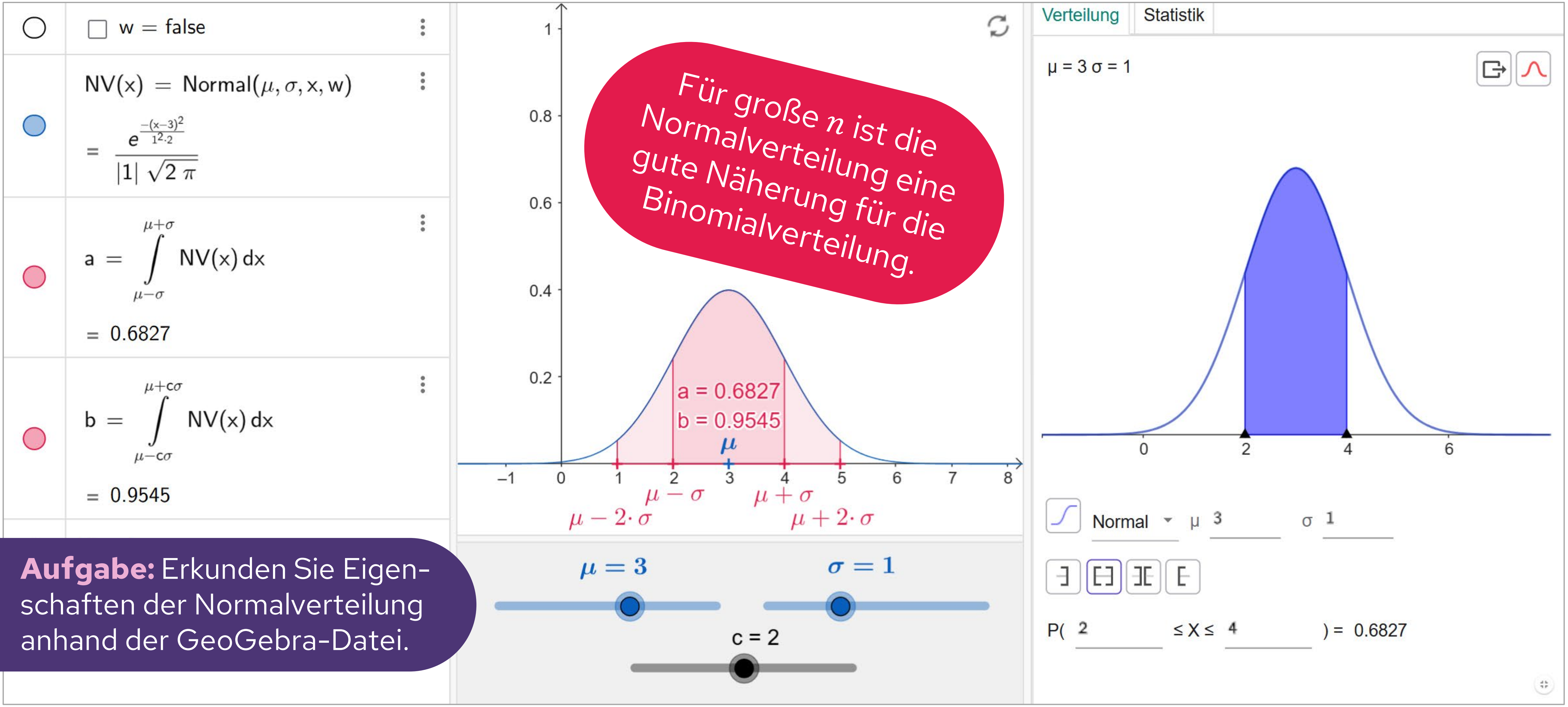
Normalverteilung mit GeoGebra





https://www.geogebra.org/m/s6xmfncc
https://www.didaktik.mathematik.uni-wuerzburg.de/history/ausstell/gauss/geldschein.html

Normalverteilung mit GeoGebra



Für große n ist die Normalverteilung eine gute Näherung für die Binomialverteilung.

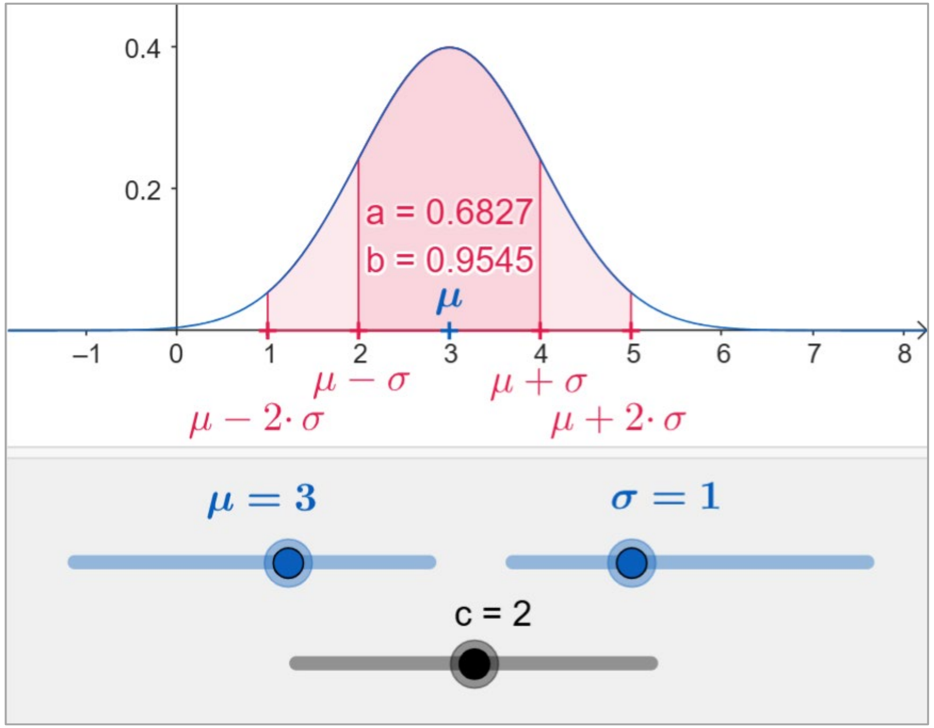
Aufgabe: Erkunden Sie Eigenschaften der Normalverteilung anhand der GeoGebra-Datei.

Sigma-Regeln (empirisch bestimmt)

Hier wird folgende Wahrscheinlichkeit in Abhängigkeit von $\boldsymbol{c}$ untersucht:

$$P(\boldsymbol{\mu} - \boldsymbol{c} \cdot \boldsymbol{\sigma} \leq \boldsymbol{X} \leq \boldsymbol{\mu} + \boldsymbol{c} \cdot \boldsymbol{\sigma})$$

$$= \int_{\boldsymbol{\mu}-\boldsymbol{c}\cdot\boldsymbol{\sigma}}^{\boldsymbol{\mu}+\boldsymbol{c}\cdot\boldsymbol{\sigma}} \frac{1}{\boldsymbol{\sigma} \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\boldsymbol{\mu})^2}{2\cdot\boldsymbol{\sigma}^2}} dx$$



Sigma-Regeln: Ist $\boldsymbol{X}$ eine normalverteilte Zufallsgröße, mit dem Erwartungswert $\boldsymbol{\mu}$ und der Standardabweichung $\boldsymbol{\sigma}$ so gilt:

- $P(\boldsymbol{\mu} - \boldsymbol{\sigma} \leq \boldsymbol{X} \leq \boldsymbol{\mu} + \boldsymbol{\sigma}) \approx 68{,}3\ \%$
- $P(\boldsymbol{\mu} - 1{,}64 \cdot \boldsymbol{\sigma} \leq \boldsymbol{X} \leq \boldsymbol{\mu} + 1{,}64 \cdot \boldsymbol{\sigma}) \approx 90{,}0\ \%$
- $P(\boldsymbol{\mu} - 1{,}96 \cdot \boldsymbol{\sigma} \leq \boldsymbol{X} \leq \boldsymbol{\mu} + 1{,}96 \cdot \boldsymbol{\sigma}) \approx 95{,}0\ \%$
- $P(\boldsymbol{\mu} - 2 \cdot \boldsymbol{\sigma} \leq \boldsymbol{X} \leq \boldsymbol{\mu} + 2 \cdot \boldsymbol{\sigma}) \approx 95{,}4\ \%$
- $P(\boldsymbol{\mu} - 2{,}58 \cdot \boldsymbol{\sigma} \leq \boldsymbol{X} \leq \boldsymbol{\mu} + 2{,}58 \cdot \boldsymbol{\sigma}) \approx 99{,}0\ \%$
- $P(\boldsymbol{\mu} - 3 \cdot \boldsymbol{\sigma} \leq \boldsymbol{X} \leq \boldsymbol{\mu} + 3 \cdot \boldsymbol{\sigma}) \approx 99{,}7\ \%$

Bemerkung

Die Laplace-Bedingung $\sigma = \sqrt{n \cdot p \cdot (1-p)} > 3$ muss erfüllt sein, damit man die Binomialverteilung durch die Normalverteilung annähern kann. In diesem Fall gelten die Sigma-Regeln näherungsweise auch für die Binomialverteilung.

 https://www.geogebra.org/m/bkaxwfme

Rheinland-Pfälzische
Technische Universität
Kaiserslautern
Landau

MaTeGnu

Kapitel 4:
Beurteilende Statistik

GeoGebra-Buch
„Didaktik der Stochastik“
https://roth.tel/stochastik

dms.nuw.rptu.de/mategnu

RPTU

Kernideen im Fokus
bei Prognose- und Konfidenzintervall

Didaktik der Mathematik
Sekundarstufen

Kernidee
Denkfigur: Bedingte Wahrscheinlichkeit

- Konditionales Denken
  (Wenn …, dann …)
- Schnitt-Wahrscheinlichkeit
  ↔ totale Wahrscheinlichkeit

Kernidee
hypothetisch-prognostische Wahrscheinlichkeit

- Konzept des Bezweifelns
- Modell (Wahrscheinlichkeit)
  ↔ Realität (rel. Häufigkeit)
- aus Erfahrung wird Erwartung

Kernidee
Repräsentativität und Variabilität von Stichproben-Daten

- Repräsentative Stichprobe
- Wachsende Stichprobe
  → reduziert Variabilität

Kernidee
Verteilungen

- Struktur erfassen
  (Form, Streuung, Maximum, …)
- Gesetz der großen Zahlen
- Sichtweise: lokal ↔ global

Lehrplan RLP: Leistungskurs

Didaktik der Mathematik Sekundarstufen

	Ziele / Inhalte (Sach- & Methodenkompetenz)	**Hinweise zur Unterrichtsgestaltung und Methodenkompetenz**
14.	Den Begriff "Konfidenzintervall" und das Verfahren zur Bestimmung eines Konfidenzintervalls für eine unbekannte Wahrscheinlichkeit verstehen (5.07e)	
15.	Den Zusammenhang zwischen dem Stichprobenumfang und der Länge des Konfidenzintervalls verstehen (5.06g, 5.07e)	
16.	Sachaufgaben zu Konfidenzintervallen lösen und die Ergebnisse interpretieren (5.07e)	

MBWWK (2015). Lehrplan Mathematik – Grund- und Leistungsfach in der gymnasialen Oberstufe (Mainzer Studienstufe), S. 55

Lehrplan RLP: Grundkurs

Wahlpflichtgebiet B1: Schätzen von Wahrscheinlichkeiten

RPTU Didaktik der Mathematik Sekundarstufen

	Ziele / Inhalte (Sach- & Methodenkompetenz)	Hinweise zur Unterrichtsgestaltung und Methodenkompetenz
1.	Verstehen, wie man Wahrscheinlichkeiten einer binomialverteilten Zufallsgröße näherungsweise mit Hilfe der Gaußschen Integralfunktion Φ (Standard-Normalverteilung) bestimmt (5.09e, 5.10e)	Die Möglichkeit der Approximation soll anschaulich, z.B. anhand von Histogrammen, einsichtig gemacht werden. Hierfür empfiehlt sich der Einsatz eines geeigneten Computerprogramms. Die Bestimmung der Näherungswerte erfolgt mit Hilfe von Tabellen oder Rechnern.
2.	Den Begriff "Konfidenzintervall" verstehen und wissen, wie man ein Konfidenzintervall für eine unbekannte Wahrscheinlichkeit bestimmt (5.07e).08e)	
3.	Den Zusammenhang zwischen dem Stichprobenumfang und der Länge des Konfidenzintervalls verstehen (5.06g, 5.07e)	
4.	Sachaufgaben zu Konfidenzintervallen lösen und die Ergebnisse interpretieren (5.06g, 5.07e)	

Zentrales Anliegen des Wahlpflichtgebiets "Schätzen von Wahrscheinlichkeiten" ist es, dass die SuS das Verfahren zur Bestimmung von Konfidenzintervallen verstehen und zum Lösen von Sachproblemen aus unterschiedlichen Bereichen anwenden. Dabei erfährt der Wahrscheinlichkeitsbegriff eine Erweiterung und Vertiefung.
Als Voraussetzung wird eine Möglichkeit, Wahrscheinlichkeiten einer binomialverteilten Zufallsgröße näherungsweise zu ermitteln, bereitgestellt.

Prognose- und Konfidenzintervalle

Prognoseintervall

- Ein 95%-Prognoseintervall zu einer Trefferwahrscheinlichkeit p ist ein Intervall, in das die zufällige Trefferanzahl X bzw. die zufällige relative Trefferhäufigkeit $h = \frac{X}{n}$ bei einer Bernoulli-Kette der Länge n mit der Wahrscheinlichkeit 95% fällt.
- Mit einem Prognoseintervall schließt man vom Modell (der Wahrscheinlichkeit p) auf die Wirklichkeit (die relative Häufigkeit h).

Modell (p) → Wirklichkeit (h)

Konfidenzintervall

- Das Konfidenzintervall zu einer relativen Häufigkeit h enthält alle Wahrscheinlichkeiten p, die nach Beobachtung von h nicht bezweifelt werden müssen.
- Das Konfidenzintervall enthält also alle Wahrscheinlichkeiten p, in deren Prognoseintervall die beobachtete relative Häufigkeit h liegt.
- Mit einem Konfidenzintervall schließt man von der Wirklichkeit (der relativen Häufigkeit h) auf das Modell (die Wahrscheinlichkeit p).

Wirklichkeit (h) → Modell (p)

Riemer, W. & Vehling, R. (2020). Prognose- und Konfidenzintervalle: beurteilende Statistik mit Sinn und Verstand. *MU, 66*(4), 11-25

$\sqrt{n}$-Gesetz und $\frac{1}{\sqrt{n}}$-Gesetz

Bemerkung: Ersetzt man in den Termen der Intervallgrenzen $\mathbf{2}$ durch $\boldsymbol{c}$, so lassen sich – entsprechend Folie 53 – auch andere Sicherheitswahrscheinlichkeiten realisieren.

RPTU Didaktik der Mathematik Sekundarstufen

Satz: $\sqrt{n}$-Gesetz der absoluten Häufigkeit

Prognoseintervall einer binomialverteilten Zufallsvariable „absolute Trefferhäufigkeit" X

Bei einer Binomialverteilung mit $\mu = n \cdot p$ und $\sigma = \sqrt{n \cdot p \cdot (1-p)}$ liegt die Trefferanzahl X mit ca. 95% Sicherheit im 2σ-Prognoseintervall.

$$[\,\mu - 2\sigma, \mu + 2\sigma\,] = \left[n \cdot p - 2\sqrt{n \cdot p \cdot (1-p)}; n \cdot p + 2\sqrt{n \cdot p \cdot (1-p)}\right] \overset{*}{\subseteq} [\,n \cdot p - \sqrt{n}; n \cdot p + \sqrt{n}\,]$$

dividieren durch n

Satz: $\frac{1}{\sqrt{n}}$-Gesetz der relativen Häufigkeit

Prognoseintervall einer binomialverteilten Zufallsvariable „relative Trefferhäufigkeit" $h = \frac{X}{n}$

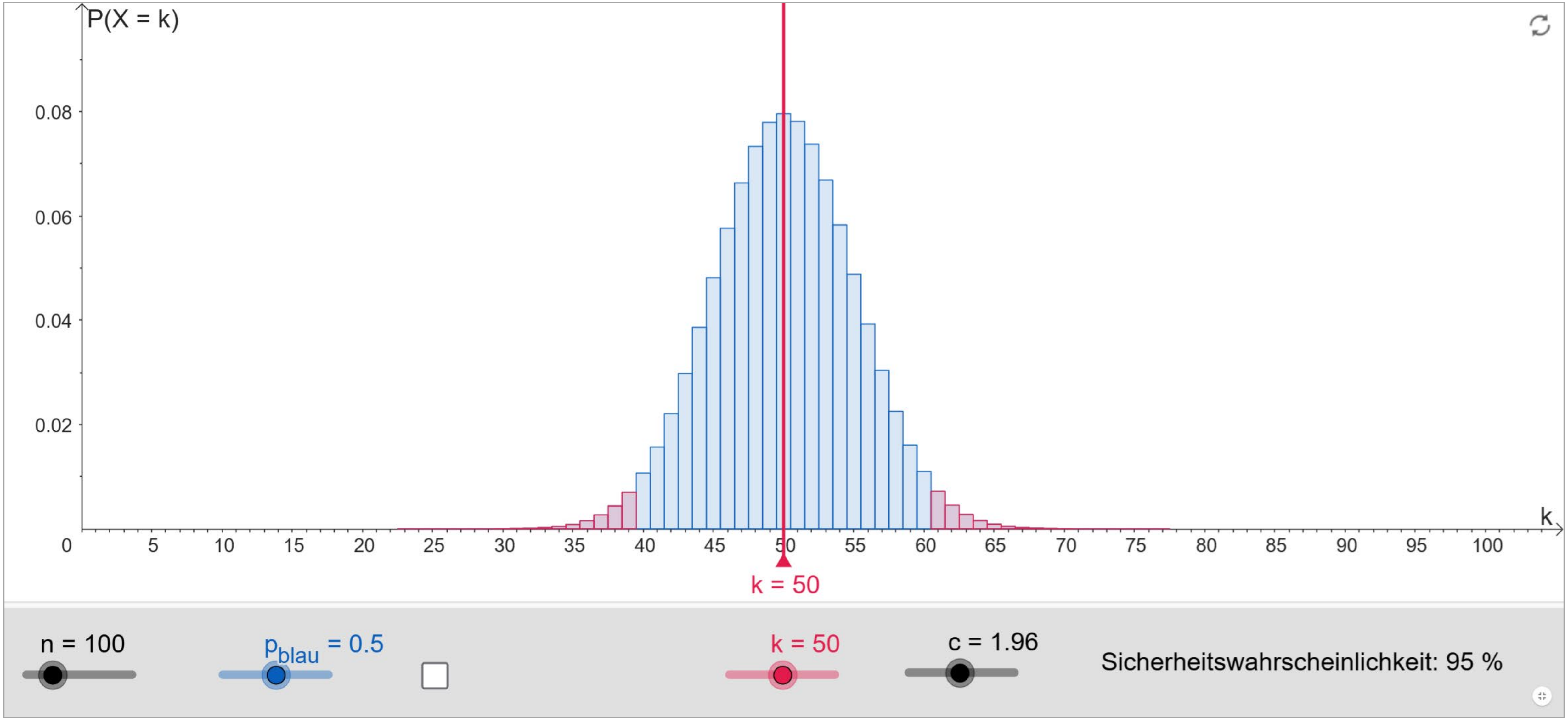
Bei gegebener Trefferwahrscheinlichkeit p einer Binomialverteilung liegt die relative Trefferhäufigkeit $h = \frac{X}{n}$ mit ca. 95% Sicherheit im Prognoseintervall.

$$\left[\frac{\mu}{n} - 2\frac{\sigma}{n}, \frac{\mu}{n} + 2\frac{\sigma}{n}\right] = \left[p - 2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}}; p + 2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}}\right] \overset{*}{\subseteq} \left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}}\right]$$

Die Länge des Prognoseintervalls nimmt umgekehrt proportional mit der Wurzel aus der Länge n der Bernoulli-Kette ab. Sie halbiert sich, wenn man n vervierfacht.

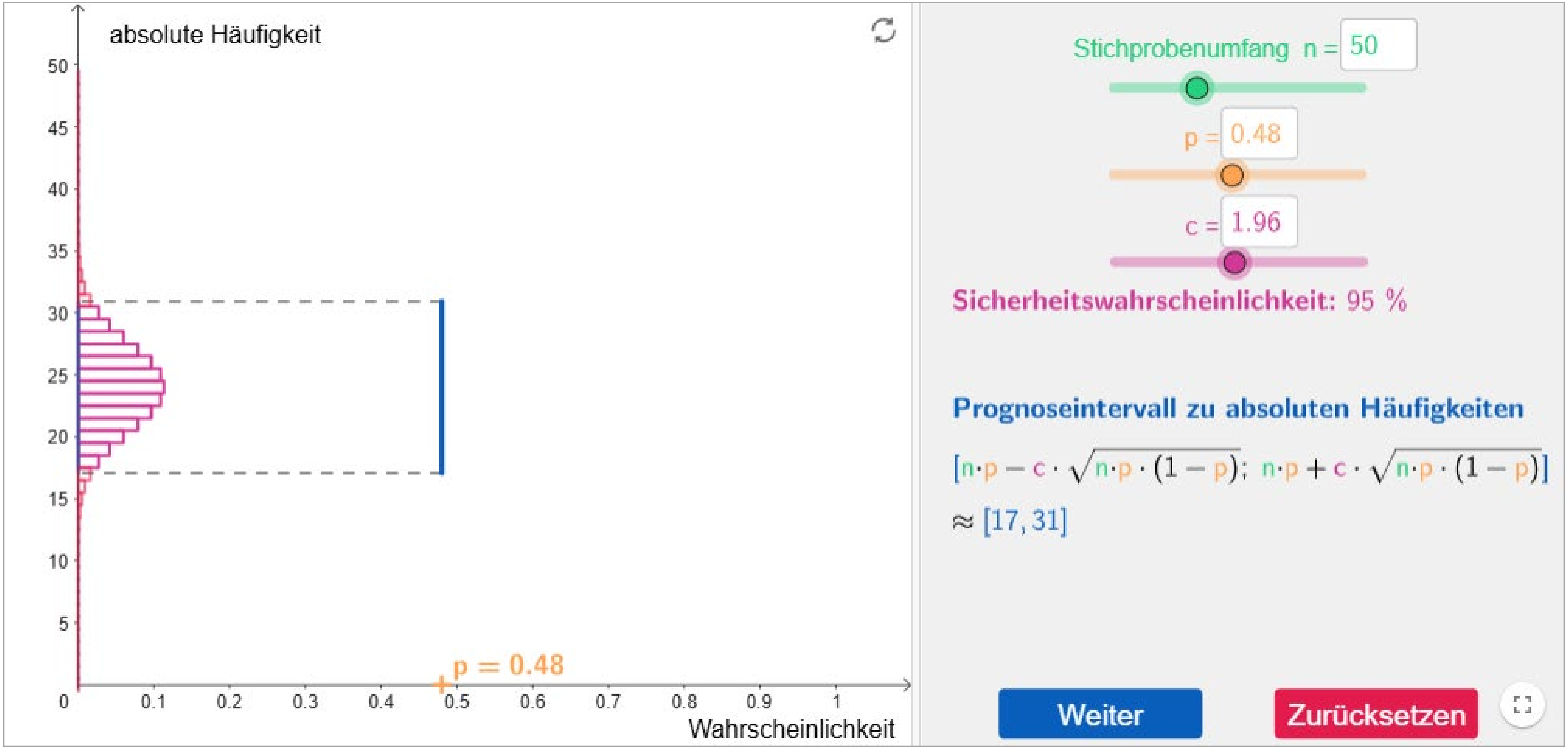
* **Bemerkung:** Die Teilmengeneigenschaften in beiden Sätzen ergeben sich aus folgender Ungleichung (vgl. Folie 48): $p \cdot (1-p) \leq \frac{1}{4}$

Prognoseintervall



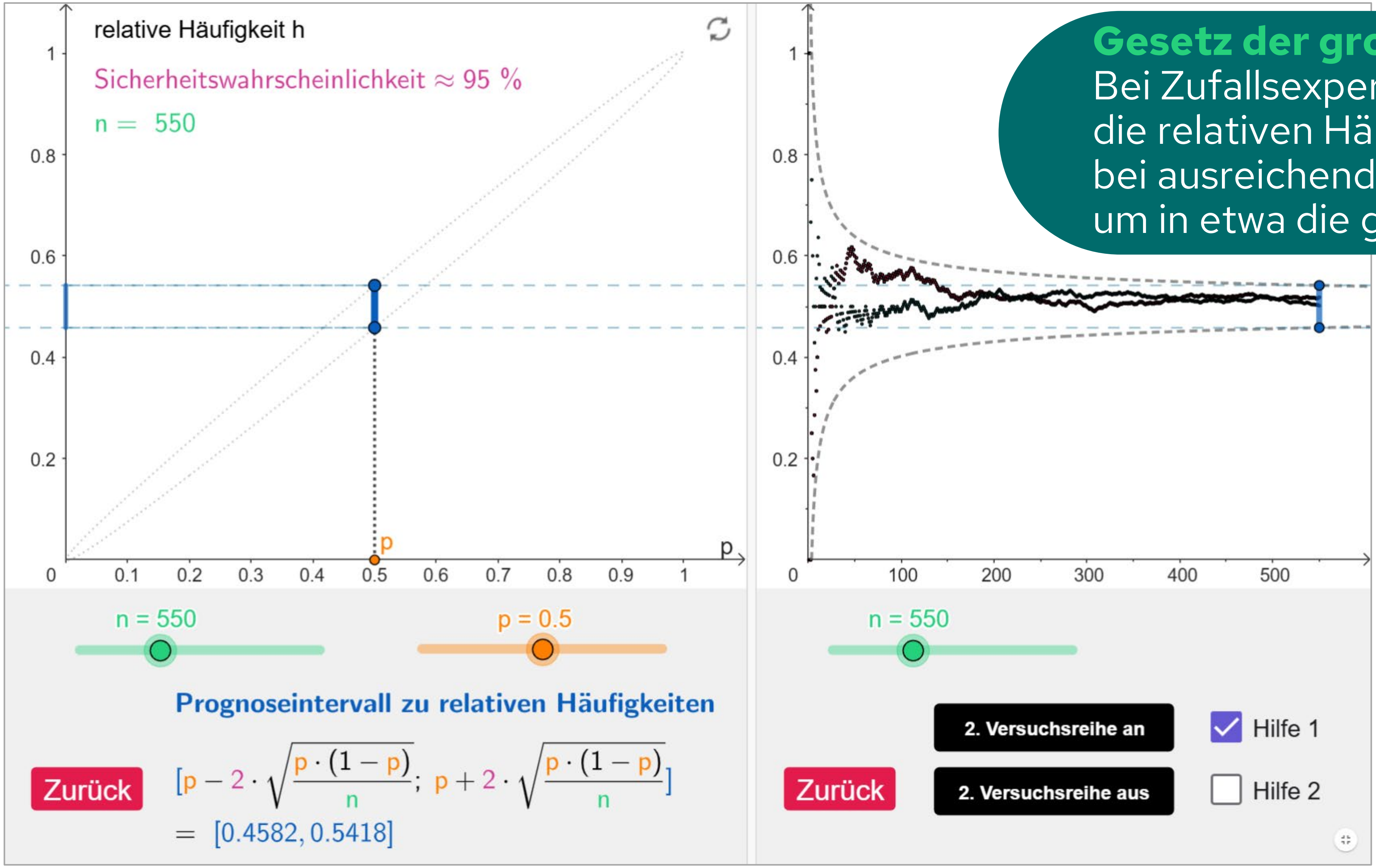
 https://www.geogebra.org/m/tecchxzt

Ellipsendiagramm verstehen



 https://www.geogebra.org/m/vp6thpjs Vielen Dank an Henrik Ossadnik für die Vorlage!

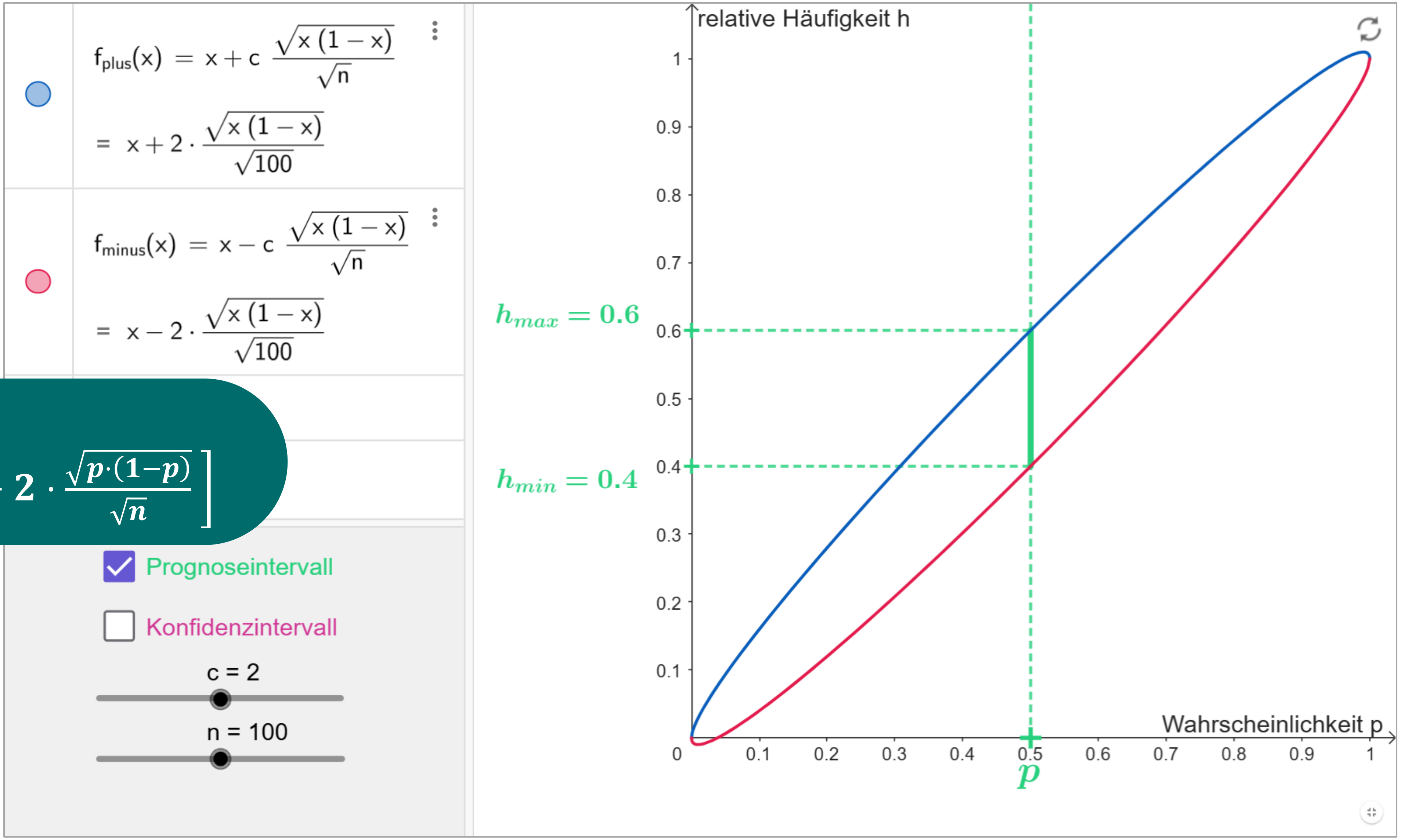
Ellipsendiagramm – Prognoseintervall



Gesetz der großen Zahlen
Bei Zufallsexperimenten pendeln sich die relativen Häufigkeiten für Ereignisse bei ausreichend großem n immer wieder um in etwa die gleiche Zahl ein.

 https://www.geogebra.org/m/vtm99k25 Vielen Dank an Henrik Ossadnik für die Vorlage!

Ellipsendiagramm – Prognoseintervall



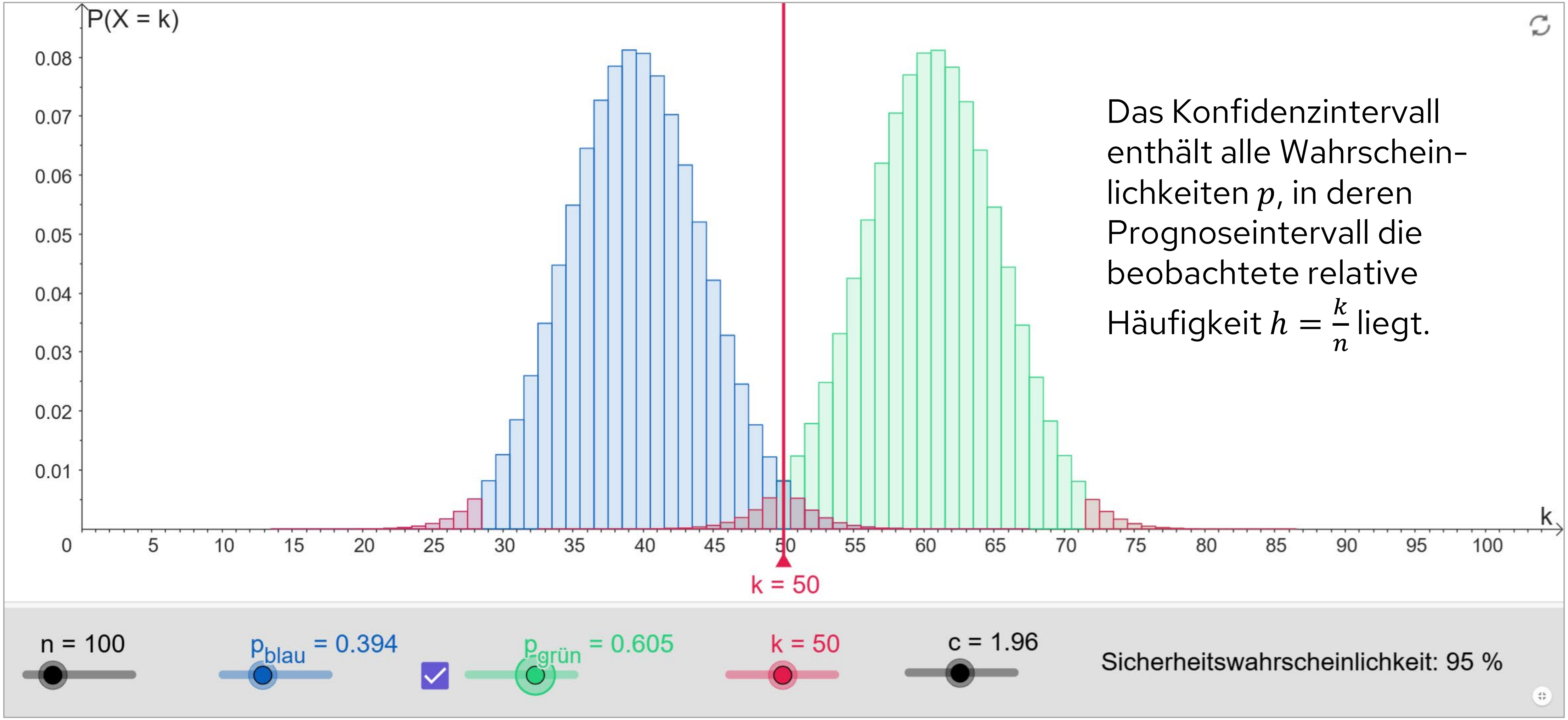
Prognoseintervall

$$\left[p - 2 \cdot \frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\,;\, p + 2 \cdot \frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}\right]$$

Bemerkung: Ersetzt man in den Termen der Intervallgrenzen 2 durch c, so lassen sich – entsprechend Folie 53 – auch andere Sicherheitswahrscheinlichkeiten realisieren.

 https://www.geogebra.org/m/ms82gcy6

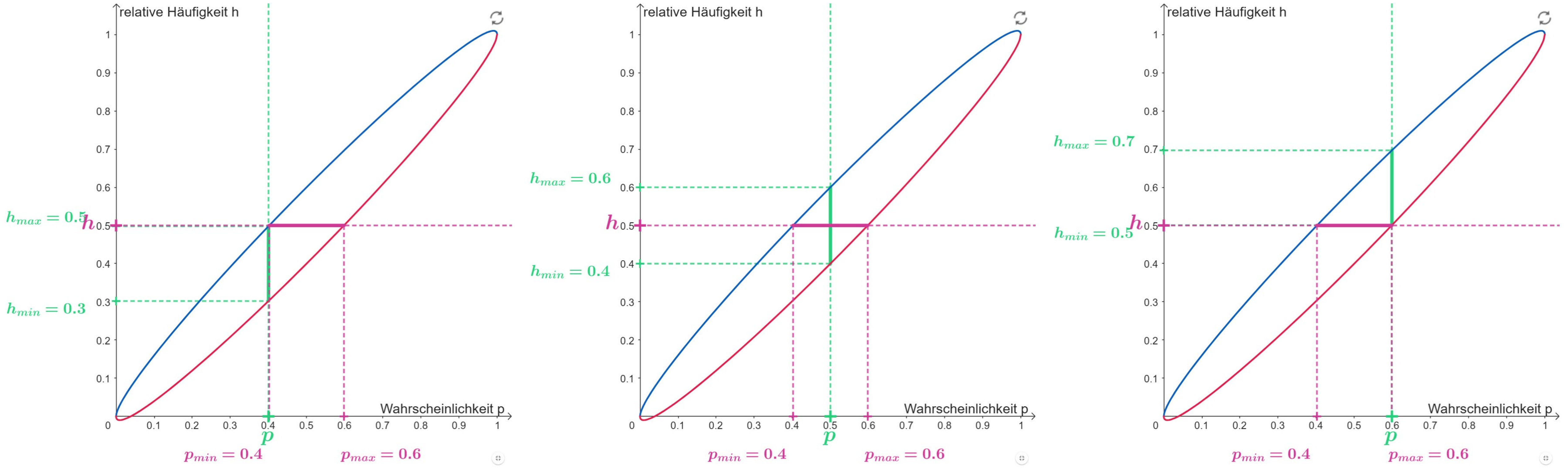
Konfidenzintervall



Das Konfidenzintervall enthält alle Wahrscheinlichkeiten p, in deren Prognoseintervall die beobachtete relative Häufigkeit $h = \frac{k}{n}$ liegt.

Ellipsendiagramm – Konfidenzintervall[1]

Aufgabe: Welche Wahrscheinlichkeiten p müssen bei Beobachtung einer relativen Häufigkeit von $h = 0{,}5$ bei einer Sicherheitswahrscheinlichkeit von 95% bezweifelt werden?



Ergebnis: Es sollten alle Wahrscheinlichkeiten p bezweifelt werden, in deren Prognoseintervall die beobachtete relative Häufigkeit h nicht liegt. → Das Konfidenzintervall enthält alle Wahrscheinlichkeiten p, in deren Prognoseintervall die beobachtete relative Häufigkeit h liegt.

[1] Lateinisch: confidere = vertrauen

Ellipsendiagramm – Konfidenzintervall[1]

Didaktik der Mathematik Sekundarstufen

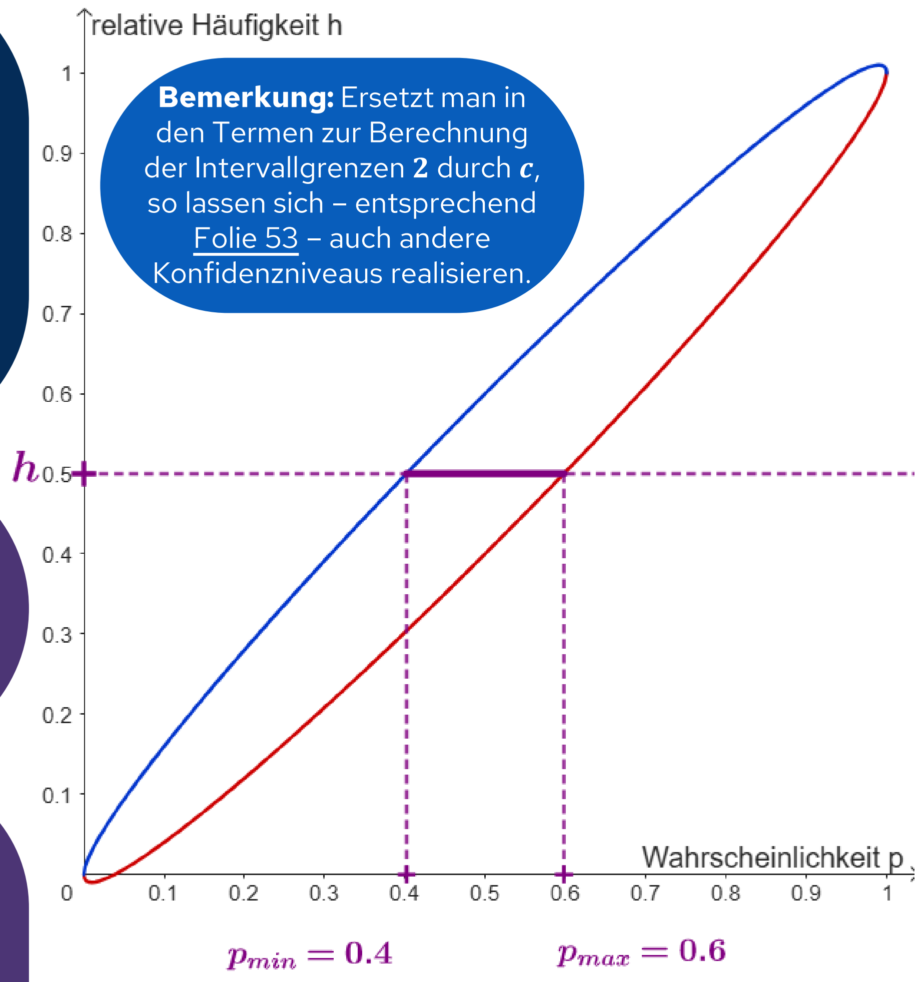
Grenzen des Konfidenzintervalls bestimmen

- Ablesen am Ellipsendiagramm (vgl. Folie 65) oder
- (numerisches) Lösen folgender Gleichungen (*) und (**)

(Die rechten Seiten sind die Terme für die Grenzen des Prognoseintervalls.)

- Linke Grenze: $h = p + 2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}}$ (*)
- Rechte Grenze: $h = p - 2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}}$ (**)

Bemerkung: Ersetzt man in den Termen zur Berechnung der Intervallgrenzen $\mathbf{2}$ durch $\boldsymbol{c}$, so lassen sich – entsprechend Folie 53 – auch andere Konfidenzniveaus realisieren.



Konfidenzintervall (exaktes Wilson-Intervall)

$$\left[\frac{h + \frac{2}{n} - 2 \cdot \sqrt{\frac{1}{n^2} + \frac{h \cdot (1-h)}{n}}}{\frac{4}{n} + 1} ; \frac{h + \frac{2}{n} + 2 \cdot \sqrt{\frac{1}{n^2} + \frac{h \cdot (1-h)}{n}}}{\frac{4}{n} + 1} \right]$$

Erhältlich nach Umformen von (*) und (**) sowie Lösen quadratischer Gleichungen (ggf. CAS)

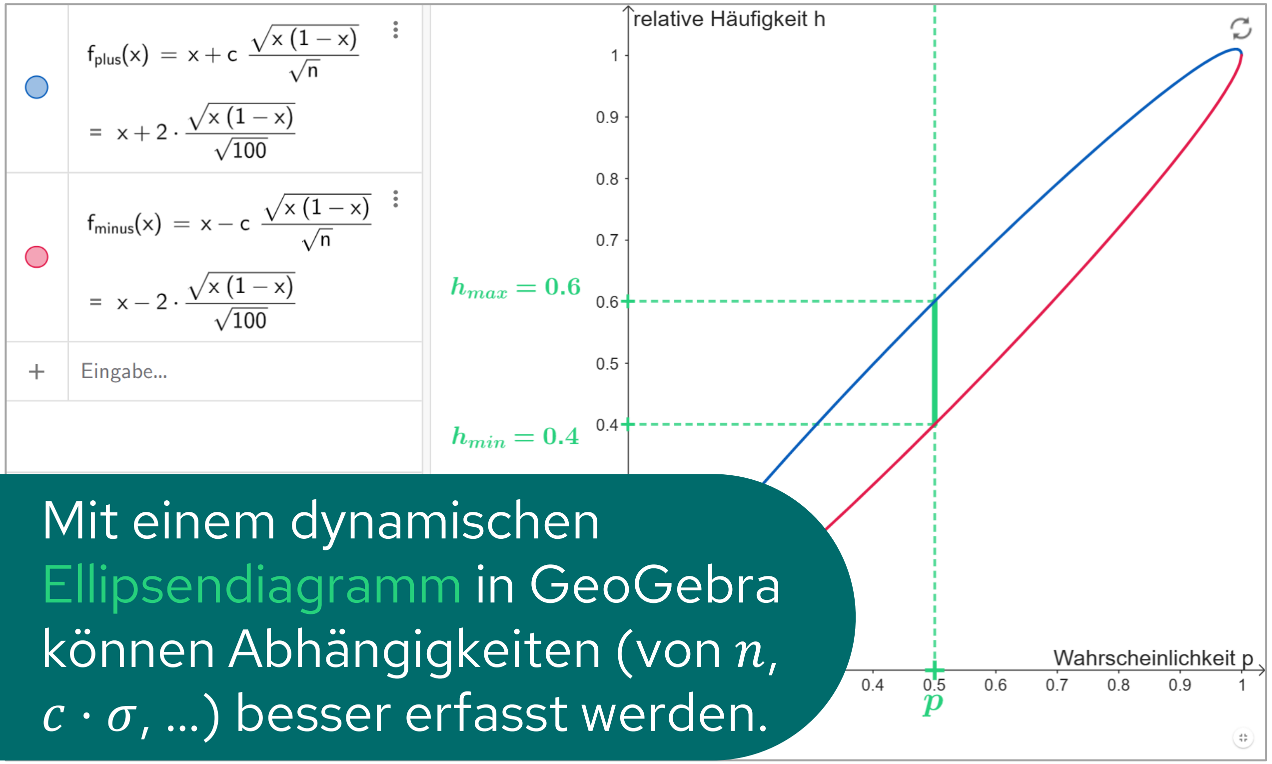
Konfidenzintervall (Näherung für große n)

$$\underbrace{\left[h - 2 \cdot \frac{\sqrt{h \cdot (1-h)}}{\sqrt{n}} ; h + 2 \cdot \frac{\sqrt{h \cdot (1-h)}}{\sqrt{n}} \right]}_{\text{Näherung: Wald-Intervall}} \subseteq \underbrace{\left[h - \frac{1}{\sqrt{n}} ; h + \frac{1}{\sqrt{n}} \right]}_{\text{Faustformel für Überschläge}}$$

[1] Lateinisch: confidere = vertrauen

Prognoseintervall
Ellipsendiagramm ↔ Tabellendarstellung

Didaktik der Mathematik Sekundarstufen



Mit einem dynamischen Ellipsendiagramm in GeoGebra können Abhängigkeiten (von n, $c \cdot \sigma$, ...) besser erfasst werden.

k↓ p→	0,15	0,16	0,17	0,18	0,19	0,2	0,21	0,22	0,23	0,24	0,25	0,26
0	0,0003	0,0002	9E-05	5E-05	3E-05	1E-05	8E-06	4E-06	2E-06	1E-06	6E-07	3E-07
1	0,0029	0,0017	0,001	0,0006	0,0003	0,0002	0,0001	6E-05	3E-05	2E-05	1E-05	5E-06
2	0,0142	0,009	0,0056	0,0035	0,0021	0,0013	0,0008	0,0005	0,0003	0,0002	9E-05	5E-05
3	0,046	0,0312	0,0208	0,0137	0,0088	0,0057	0,0036	0,0022	0,0014	0,0008	0,0005	0,0003
4	0,1121	0,0808	0,0572	0,0399	0,0274	0,0185	0,0123	0,0081	0,0052	0,0033	0,0021	0,0013
5	0,2194	0,1677	0,1259	0,0929	0,0673	0,048	0,0337	0,0233	0,0159	0,0107	0,007	0,0046
6	0,3613	0,2919	0,2314	0,1801	0,1376	0,1034	0,0764	0,0555	0,0397	0,028	0,0194	0,0132
7	0,5188	0,4406	0,3672	0,3004	0,2413	0,1904	0,1477	0,1126	0,0844	0,0623	0,0453	0,0324
8	0,6681	0,5929	0,5168	0,4424	0,372	0,3073	0,2495	0,1991	0,1563	0,1206	0,0916	0,0684
9	0,7911	0,7282	0,6597	0,5878	0,5151	0,4437	0,3759	0,313	0,2564	0,2066	0,1637	0,1276
10	0,8801	0,8339	0,7797	0,7187	0,6527	0,5836	0,5136	0,4448	0,379	0,3178	0,2622	0,2129
11	0,9372	0,9071	0,8691	0,8232	0,77	0,7107	0,6467	0,5799	0,5122	0,4456	0,3816	0,3218
12	0,9699	0,9525	0,9286	0,8978	0,8595	0,8139	0,7617	0,7037	0,6415	0,5767	0,511	0,4461
13	0,9868	0,9777	0,9643	0,9456	0,9209	0,8894	0,851	0,8058	0,7544	0,6977	0,637	0,5739
14	0,9947	0,9904	0,9836	0,9734	0,9589	0,9393	0,9138	0,8819	0,8435	0,7987	0,7481	0,6925
15	0,9981	0,9962	0,993	0,988	0,9803	0,9692	0,9538	0,9335	0,9074	0,8753	0,8369	0,7925
16	0,9993	0,9986	0,9973	0,995	0,9913	0,9856	0,9771	0,9653	0,9492	0,9282	0,9017	0,8694
17	1	1	0,999	0,9981	0,9964	0,9937	0,9895	0,9832	0,9741	0,9616	0,9449	0,9234
18	1	1	1	0,9993	0,9987	0,9975	0,9955	0,9925	0,9878	0,9809	0,9713	0,9582
19	1	1	1	1	1	0,9991	0,9982	0,9969	0,9946	0,9912	0,9861	0,9788

Verknüpfung zur algebraischen Darstellung der Intervalle beim Ellipsendiagramm deutlicher.

Ablesen von Tabellen ist auch unreflektiert und unverstanden möglich. → Eher problematisch.

Aktuelle Werkzeugkompetenz ist der Umgang mit digitalen Werkzeugen wie z. B. GeoGebra.

Papier-Tabellen ablesen keine aktuell relevante Werkzeug-Kompetenz mehr (nur noch historisch).

 Die Abbildung der Excel-Tabelle stammt von Christian Fahse.

Eigenschaften von Prognose- und Konfidenzintervallen

Bemerkung: Durch Ersetzen von $\mathbf{2}$ durch $\boldsymbol{c}$ in den Berechnungstermen der Intervallgrenzen (vgl. Folie 53) lassen sich auch andere γ-Werte realisieren.

Didaktik der Mathematik
Sekundarstufen

Prognoseintervall

- Prognoseintervalle enthalten relative Häufigkeiten.
- Prognoseintervalle sind symmetrisch zur Wahrscheinlichkeit $\boldsymbol{p}$.
- Im 95%-Prognoseintervall zu einer bekannten Wahrscheinlichkeit $\boldsymbol{p}$ liegt die relative Häufigkeit $\boldsymbol{h}$ mit der Wahrscheinlichkeit $\boldsymbol{\gamma} = \mathbf{95\%}$.
- Man nennt $\boldsymbol{\gamma} = \mathbf{95\%}$ Sicherheitswahrscheinlichkeit.
- Die Länge von Prognoseintervallen halbiert sich bei Vervierfachung des Versuchsumfangs.

Konfidenzintervall

- Konfidenzintervalle enthalten Wahrscheinlichkeiten.
- Wilson-Konfidenzintervalle sind nur für $\boldsymbol{h} = 0{,}5$ symmetrisch zur relativen Häufigkeit $\boldsymbol{h}$. Wald-Konfidenzintervalle sind stets symmetrische zu $\boldsymbol{h}$.
- Das (zufallsabhängige) 95%-Konfidenzintervall zu einer beobachteten relativen Häufigkeit $\boldsymbol{h}$ überdeckt die unbekannte (feste) Wahrscheinlichkeit $\boldsymbol{p}$ mit der Wahrscheinlichkeit $\boldsymbol{\gamma} = \mathbf{95\%}$.
- Man nennt $\boldsymbol{\gamma} = \mathbf{95\%}$ Konfidenzniveau (oder Vertrauenswahrscheinlichkeit).
- Bei Vervierfachung des Versuchsumfangs halbieren sich Wilson-Konfidenzintervalle annähernd, Wald-Konfidenzintervalle genau.

Riemer, W. & Vehling, R. (2020). Prognose- und Konfidenzintervalle: beurteilende Statistik mit Sinn und Verstand. MU, 66(4), 11-25

Prognoseintervall ↔ Konfidenzintervall

RPTU | Didaktik der Mathematik Sekundarstufen

Prognoseintervall

Situation
Aufgrund empirischer Untersuchungen oder theoretischer Überlegungen bekannte Trefferwahrscheinlichkeit p in der Grundgesamtheit.

Bekannt:
Trefferwahrscheinlichkeit p
in der
Grundgesamtheit

Modell (p)

schließen auf

Ziel
Aussage darüber, welche absoluten Trefferanzahlen X bzw. relativen Trefferhäufigkeit h in einer Stichprobe mit der Trefferwahrscheinlichkeit p „verträglich" sind.

Gesucht:
Prognoseintervall
Trefferanzahlen X / –häufigkeiten h
mit denen in einer
Stichprobe
zu rechnen ist.

Wirklichkeit (h)

Konfidenzintervall

Situation
In einer Stichprobe mit dem Umfang n tritt die Merkmalsausprägung mit der relativen Häufigkeit h auf.

Bekannt:
Relative Häufigkeit h
in der
Stichprobe

Wirklichkeit (h)

schließen auf

Ziel
Aussage über die unbekannte Wahrscheinlichkeit p für das Eintreffen des Merkmals in der Grundgesamtheit.

Gesucht:
Konfidenzintervall
Wahrscheinlichkeiten p
in der
Grundgesamtheit
die zu h passen.

Modell (p)

Lateinisch: confidere = vertrauen

Stichprobe und Konfidenzintervall

Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

MaTeGnu

Kapitel 4: Beurteilende Statistik

Kernideen im Fokus
beim Hypothesentest

Didaktik der Mathematik Sekundarstufen

Kernidee Denkfigur: Bedingte Wahrscheinlichkeit

- Konditionales Denken (Wenn …, dann …)
- Schnitt-Wahrscheinlichkeit ↔ totale Wahrscheinlichkeit

Kernidee hypothetisch-prognostische Wahrscheinlichkeit

- Konzept des Bezweifelns
- Modell (Wahrscheinlichkeit) ↔ Realität (rel. Häufigkeit)
- aus Erfahrung wird Erwartung

Kernidee Repräsentativität und Variabilität von Stichproben-Daten

- Repräsentative Stichprobe
- Wachsende Stichprobe → reduziert Variabilität

Kernidee Verteilungen

- Struktur erfassen (Form, Streuung, Maximum, …)
- Gesetz der großen Zahlen
- Sichtweise: lokal ↔ global

Lehrplan RLP: Leistungskurs

Didaktik der Mathematik Sekundarstufen

	Ziele / Inhalte (Sach- & Methodenkompetenz)	**Hinweise zur Unterrichtsgestaltung und Methodenkompetenz**
17.	Die Struktur des Hypothesentests verstehen (5.08e)	
18.	Sachaufgaben zum Testen von Hypothesen lösen und die Ergebnisse interpretieren (5.08e)	Die Sachprobleme werden so vorgegeben, dass sie durch Binomialverteilungen modelliert werden können. Gegebenenfalls werden die Binomialverteilungen durch die Normalverteilung approximiert. Besondere Bedeutung kommt der Interpretation des Ergebnisses eines Hypothesentests zu. Dabei sollen die Schülerinnen und Schüler auch die Grenzen des Verfahrens erkennen. Zumindest einmal sollen die Schülerinnen und Schüler zu einem offen formulierten Sachproblem einen Hypothesentest entwerfen, gesuchte Größen berechnen und die Konsequenzen der Ergebnisse für den Sachverhalt erörtern. Im Leistungskurs soll dies weitgehend selbstständig in Gruppen- oder Partnerarbeit erfolgen.

Lehrplan RLP: Grundkurs
Wahlpflichtgebiet B2: Testen von Hypothesen

RPTU Didaktik der Mathematik Sekundarstufen

	Ziele / Inhalte (Sach- & Methodenkompetenz)	**Hinweise zur Unterrichtsgestaltung und Methodenkompetenz**
1.	Das Vorgehen beim Testen von Hypothesen verstehen (5.08e)	
2.	Verstehen, welche Fehlentscheidungen beim Hypothesentest auftreten können & wissen, wie man die Wahrscheinlichkeiten dafür ermittelt (5.08e)	
3.	Sachaufgaben zum Testen von Hypothesen lösen und die Ergebnisse interpretieren (5.08e)	Die Sachprobleme werden so vorgegeben, dass sie durch Binomialverteilungen modelliert werden können. Besondere Bedeutung kommt der Interpretation des Ergebnisses eines Hypothesentests zu. Dabei sollen die Schüler/innen auch die Grenzen des Verfahrens erkennen. Zumindest einmal sollen die Schüler/innen zu einem offen formulierten Sachproblem einen Hypothesentest entwerfen, gesuchte Größen berechnen und die Konsequenzen der Ergebnisse für den Sachverhalt erörtern. Die im Grundkurs hierfür angemessene Unterrichtsform ist in der Regel das Unterrichtsgespräch oder die angeleitete Gruppenarbeit.

Zentrales Anliegen des Wahlpflichtgebiets “Testen von Hypothesen” ist es, dass die Schüler/innen das Verfahren verstehen & zum Lösen von Sachproblemen aus unterschiedlichen Bereichen anwenden. Im Zusammenhang mit der Diskussion von Fehlerwahrscheinlichkeiten erfährt der Wahrscheinlichkeitsbegriff eine Erweiterung & Vertiefung.

 MBWWK (2015). Lehrplan Mathematik – Grund- und Leistungsfach in der gymnasialen Oberstufe (Mainzer Studienstufe), S. 36

Warum Hypothesentest im MU?

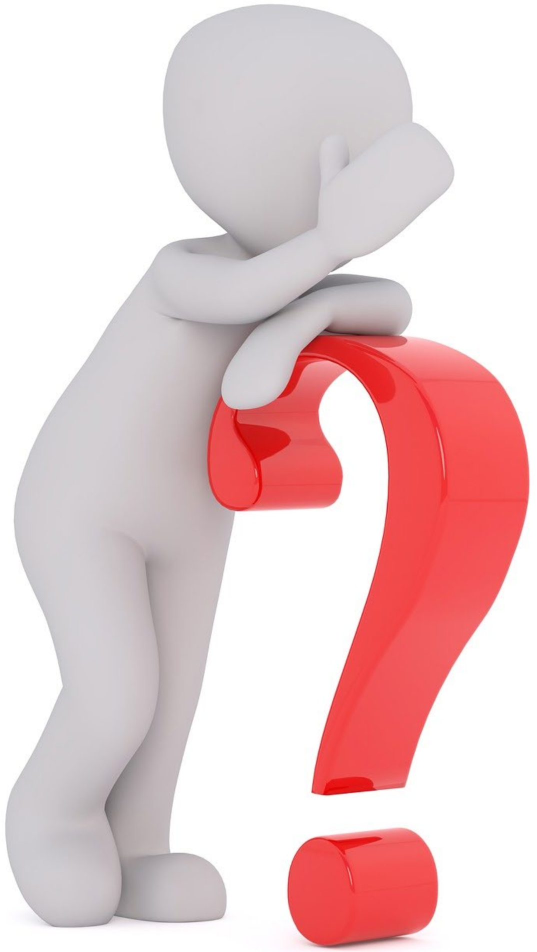
Klassischer Hypothesentest	Typische Methode der beurteilenden Statistik
Ziel mathematischer Bildung	Verständigen und reflektierten Umgang mit gängigen Methoden fördern
Problem	Testen von Hypothesen wird häufig als unverstandenes Schema abgearbeitet
Gegenmaßnahme	Stärkung des inhaltlich-anschaulichen Verstehens und Argumentierens

Leuders, T. (2005). Darf das denn wahr sein? – Eine schüleraktive Entdeckung der Grundidee des Hypothesentestens durch Simulation mit Tabellenkalkulation. *PM, 47*(4), 8-16

Hypothesentest: Probleme

Probleme beim Hypothesentest

- Das Konzept ist anspruchsvoll
- Es werden schwierige Begriffe benötigt.
  - bedingte Wahrscheinlichkeit
  - Fehler 1. Art und 2. Art
  - Binomialverteilung, -koeffizient
  - …
- Angemessene Interpretation der Ablehnungs- bzw. Entscheidungsbereiche selbst für Experten problematisch
- Realitätsbezug macht die Problemsituation eher komplexer also durchsichtiger
- Algebraisch auf Schulniveau berechenbare Fälle sind sehr beschränkt
  → Binomialverteilung oder Black-Box-Formeln



Hypothesentest: Chancen

Chancen beim Hypothesentest

- Begriffe der Wahrscheinlichkeitsrechnung werden produktiv zusammengeführt.
  - Wiederholen, in Beziehung setzen und anwenden
  - oder aus Beurteilungssituationen herleiten
- Arbeitsweise von empirisch arbeitenden Wissenschaftlern exemplarisch erfahrbar:
  - Aufstellen von Hypothesen
  - Erproben im Experiment
  - Verwerfen bzw. vorläufiges Akzeptieren nach rationalen Kriterien

Ziel: Tragfähige Grundvorstellungen zum Hypothesentest aufbauen

Hypothesentest: Erarbeitungsprinzipien

Didaktik der Mathematik
Sekundarstufen

Aufbau tragfähiger, inhaltlicher Begriffsvorstellungen

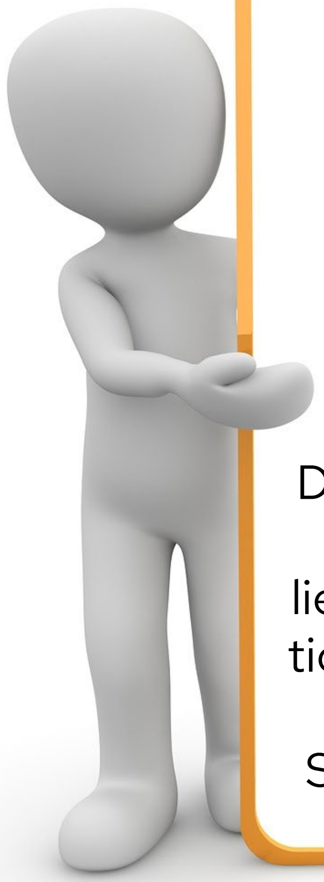
- Entschleunigung der Begriffsbildung
- Arbeiten mit vorläufigen Begriffen
- hohe Schüleraktivität
- Kein rezeptartiges Arbeiten!

Schwerpunkt: Inhaltlich-anschauliches Arbeiten

Formale, symbolische und algebraischer Elemente sehr dosiert verwenden!

Computerwerkzeuge

- Heuristisch und explorativ nutzen
- Umsetzung obiger Prinzipien



Genetische Erarbeitung

Das Verfahren des Hypothesentestens wird aus einer Problemsituation von Lernenden aktiv entwickelt.

Die dem Hypothesen-testen zugrunde liegenden Argumenta-tionsfiguren bilden den roten Faden der Schülererkundungen.

Hypothesentest: Erarbeitungsprinzipien

RPTU | Didaktik der Mathematik Sekundarstufen

Schritt	**Fachliche Beschreibung** (Rückschauperspektive der „fertigen Mathematik“)	**Inhaltlich-anschauliches Argumentieren** aus Sicht der Schüler	**Erwartete Schülertätigkeit** bei der Erarbeitung eines Verfahrens
(1)	Hypothese **H** aufstellen und Versuchsausgang **A** beobachten (H oft Gegenteil der Vermutung)	Du behauptest der Würfel sei in Ordnung? Ich habe 6-mal die 6 gewürfelt.	Zu untersuchende Situation & konkurrierende Annahmen festlegen. Ausgang bestimmen.
(2)	Man fragt: Wenn **H** richtig ist, wie wahrscheinlich ist dann **A**?	Es wäre zwar möglich, aber extrem unwahrscheinlich, dass der Würfel in Ordnung ist.	Berechnung oder Simulation einer erwarteten Verteilung, Beurteilung der möglichen Ausgänge nach der Wahrscheinlichkeit
(3)	Bei unwahrscheinlichen Ausgängen wird die Hypothese ***H*** verworfen.	Also glaube ich dir nicht! Natürlich könnte ich mich irren und du hast doch Recht. Das ist aber extrem unwahrscheinlich.	Kriterien für eine Ablehnungsentscheidung bestimmen; Beurteilung der Situation aus Sicht der getroffenen Annahmen

Leuders, T. (2005). Darf das denn wahr sein? – Eine schüleraktive Entdeckung der Grundidee des Hypothesentestens durch Simulation mit Tabellenkalkulation. *PM*, *47*(4), 8–16

Nullhypothese: Einseitig ↔ Zweiseitig

Bemerkung: Ziele der Beteiligten sind nicht klar: Wollen z. B. Jugendschützer beruhigt schlafen, wünschen sie sich niedrige Messwerte unterhalb des Prognoseintervalls zu $p_0 = 0{,}05$, wollen sie erfolgreich klagen, wünschen sie sich Messwerte oberhalb des Prognoseintervalls.

RPTU Didaktik der Mathematik Sekundarstufen

Aufgabe

Eine Fabrik erzeugt Getränke mit einem Alkoholgehalt von 5 %. Dabei treten beim Alkoholgehalt Schwankungen auf. Die Hypothese H_0, dass der Alkoholgehalt gleich dem Sollwert $p_0 = 0{,}05$ ist, soll anhand einer Stichprobe überprüft werden.

Aufgrund der Interessenlage der Personen, die die Untersuchung vornehmen, sind drei Fälle zu unterscheiden:

Die Überprüfung wird durchgeführt durch eine

a) **Eichkommission**, die an einer Abweichung vom Sollwert $p_0 = 0{,}05$ sowohl nach unten als auch nach oben interessiert ist,

b) **Verbraucherorganisation**, die daran interessiert ist, dass tatsächlich die 5 % Alkohol in den Flaschen sind. Sie stellt misstrauisch die Frage, ob der wahre Alkoholgehalt kleiner als der Sollwert ist,

c) **Jugendschutzkommission**, die befürchtet, dass zu viel Alkohol in den Flaschen ist, um die Konsumenten möglichst schnell alkoholabhängig zu machen. Sie stellt ebenfalls misstrauisch die Frage, ob der wahre Alkoholgehalt größer als der Sollwert ist.

Die Interessenslagen führen zu folgenden Testverfahren:

a) $H_0\colon p = p_0$ gegen $H_1\colon p \neq p_0$ — zweiseitig

b) $H_0\colon p \leq p_0$ gegen $H_1\colon p > p_0$ — einseitig

c) $H_0\colon p \geq p_0$ gegen $H_1\colon p < p_0$ — einseitig

Bemerkung: H_0 nennt man Nullhypothese, H_1 nennt man Alternativhypothese.

Stoyan, D. (2011). Statistische Tests in Gymnasiallehrbüchern. *Stochastik in der Schule, 31*(1), 28-32
Riemer, W. (2023). Statistik unterrichten – Eine handlungsorientierte Didaktik der Stochastik. Kallmeyer, 111f

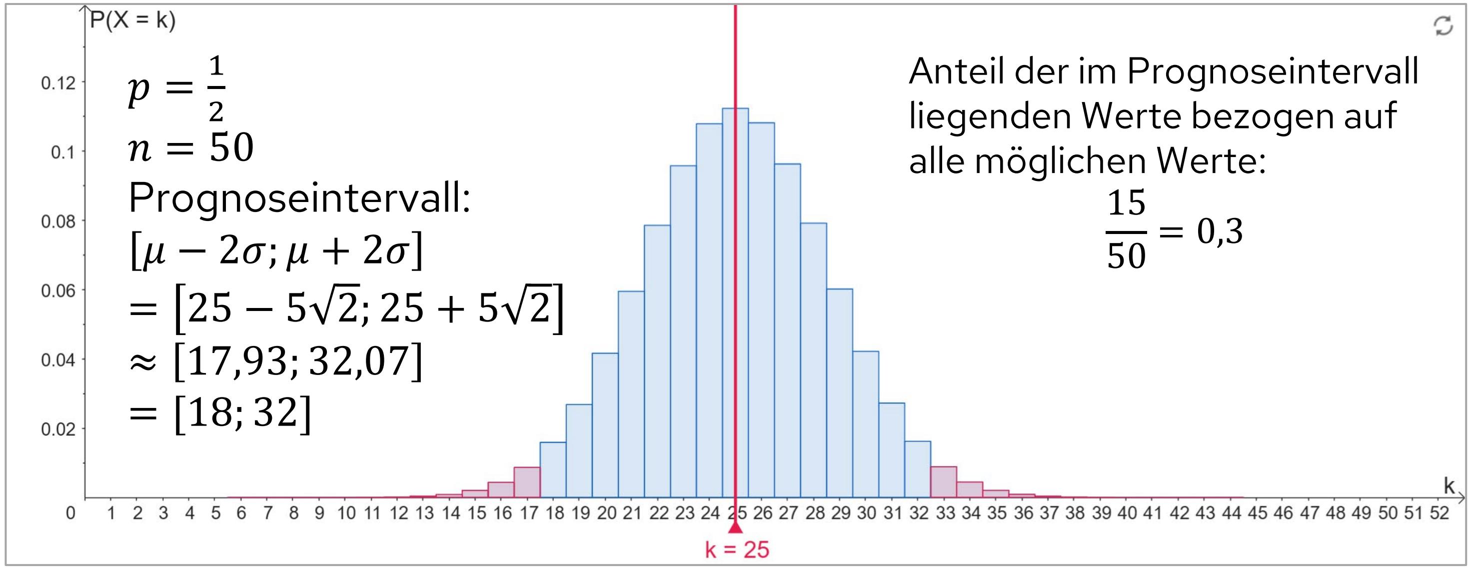
Signifikanztest: Einseitig ↔ Zweiseitig
Der Einfluss des Stichprobenumfangs n

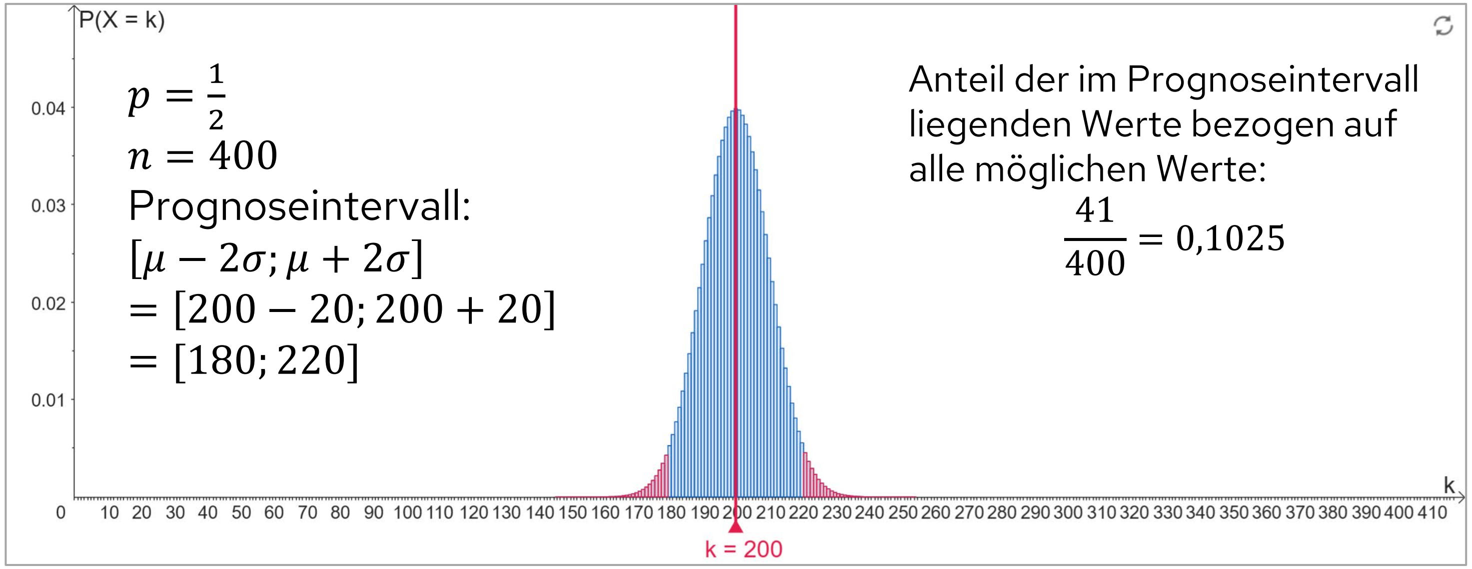
R TU P Didaktik der Mathematik Sekundarstufen

Bemerkungen: Bei einseitigen Tests ist zur Entschlüsselung von Standpunkten und Interessenlagen der zweiseitige Blick sehr hilfreich.

Der zweiseitige Blick auf einseitige Signifikanztests lässt den Einfluss des Stichprobenumfangs n hervortreten:

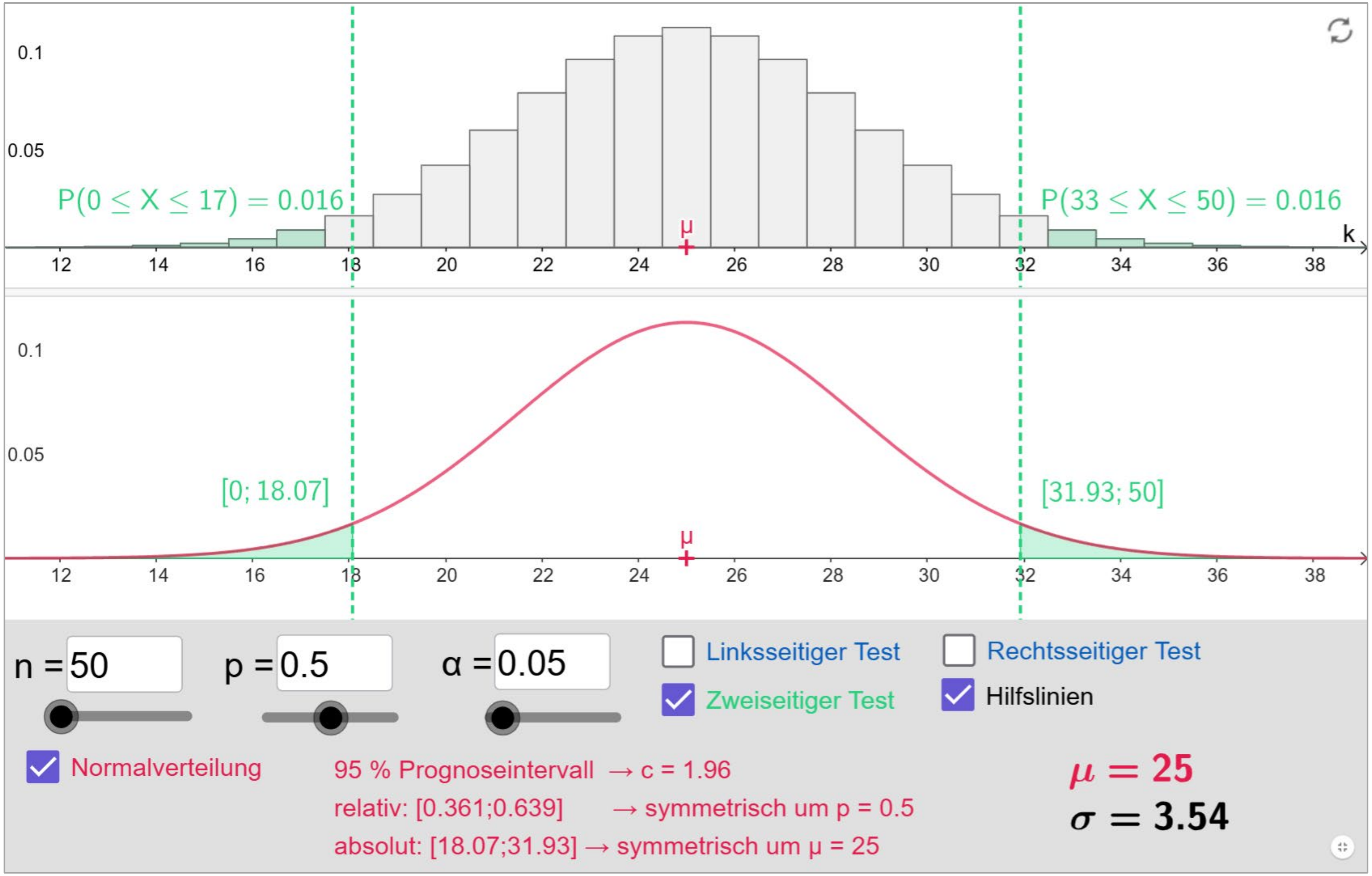
- Wenn n sehr groß ist, wird das Prognoseintervall sehr klein, d.h. der Anteil der im Prognoseintervall liegenden Werte bezogen auf alle möglichen Werte wird sehr klein.
- Wenn dann die Beobachtung ein wenig außerhalb des Prognoseintervalls liegt, sind die Abweichungen trotz statistischer Signifikanz inhaltlich in der Regel irrelevant.





Riemer, W. (2023). Statistik unterrichten – Eine handlungsorientierte Didaktik der Stochastik. Kallmeyer, 111ff • https://geogebra.org/m/tecchxzt

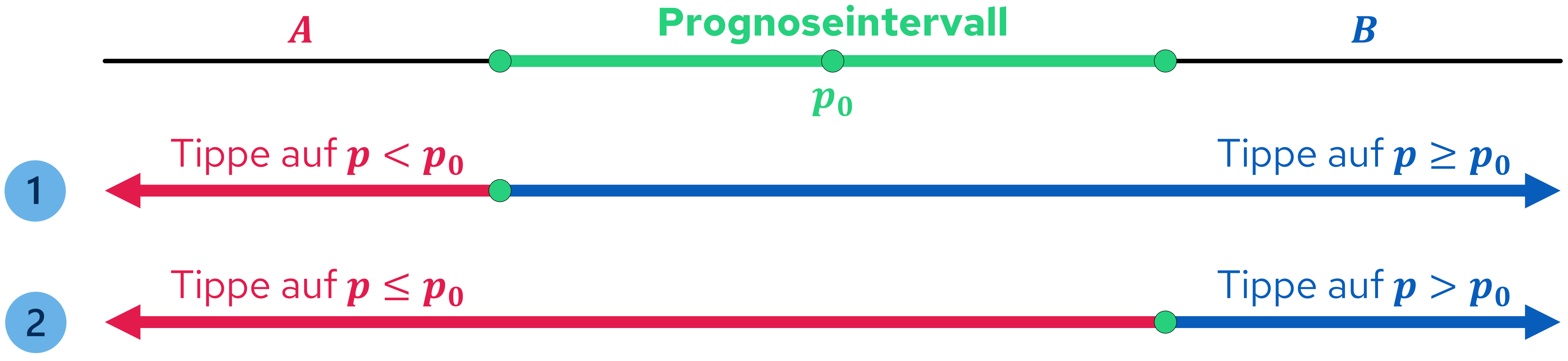
Signifikanztest: Einseitig ↔ Zweiseitig



https://geogebra.org/m/fpduh95y

 Riemer, W. (2023). Statistik unterrichten – Eine handlungsorientierte Didaktik der Stochastik. Kallmeyer, 111ff

Einseitige Hypothesentests:
Welche Seite soll es sein?



H_0: $p \geq p_0$
H_1: $p < p_0$

H_0: $p \leq p_0$
H_1: $p > p_0$

Obere Zeile 1: Sollte ein falscher Tipp auf A (also das fälschliche Verwerfen von B) mehr Verlust bringen (negativere Konsequenzen haben) als ein falscher Tipp auf B, wird man mit dem Tippen auf A vorsichtig sein und im Zweifel bei B: $p \geq p_0$ bleiben, also nicht nur, wenn h oberhalb, sondern auch, wenn h im Prognose-intervall von p_0 liegt. Statistisch gesprochen wird B dann zur Nullhypothese H_0 und das Verfahren zum linksseitigen Test H_0: $p \geq p_0$ gegen die Alternativhypothese H_1: $p < p_0$.

Untere Zeile 2: Wenn der falsche Tipp auf B (also das fälschliche Verwerfen von A) negativere Konsequenzen hat als der falsche Tipp auf A, bleibt man, wenn h im Prognoseintervall von p_0 liegt, bei A: $p \leq p_0$. A wird zur Nullhypothese H_0. Man testet rechtsseitig H_0: $p \leq p_0$ gegen H_1: $p > p_0$.

 Riemer, W. (2023). Statistik unterrichten – Eine handlungsorientierte Didaktik der Stochastik. Kallmeyer, 115

Nullhypothese
„Die verlorenen Mädchen von Gorleben“

Biostatistik-Studie

In einer Studie beschreibt der Biostatistiker Hagen Scherb eine statistisch signifikante Abweichung im Geschlechterverhältnis Neugeborener rund um das Atommüll-Zwischenlager in Gorleben.

Er untersuchte den Anteil der Jungen bzw. Mädchengeburten nach dem Beginn der Einlagerung von Atommüll im Jahr 1996 in das Zwischenlager, das 2 km südlich von Gorleben liegt.

Scherb und seine Kollegen werten „die verlorenen Mädchen von Gorleben“ als Nachweis dafür, dass eine Strahlenbelastung stattgefunden hat. (Scherb et al. 2012)

Aufgabe

Untersuchen Sie, ob das Geschlechterverhältnis rund um Gorleben im Vergleich zu Deutschland tatsächlich statistisch signifikant verschoben ist.

Nutzen Sie dazu folgende im Jahr 2011 vom Niedersächsischen Landesgesundheitsamt (NLGA) veröffentlichten Daten zu den Geburten in sämtlichen Gemeinden im 35 km-Umkreis um das Atommüll-Zwischenlager in Gorleben:

	1991-1995	1996-2009	Summe
Jungen	3477	10040	13517
Mädchen	3421	9169	12590
Summe	6898	19209	26107

Schäfer, A. (2016). Weniger Mädchen in Gorleben?! – Statistische Signifikanz verstehen. *Mathematik lehren, 197*, 40-44
Schäfer, A. (2017). Das Formulieren der Nullhypothese beim Signifikanztest mit Blick auf eine authentische Anwendung. *Stochastik in der Schule, 37*(3), 18-24

Nullhypothese
„Die verlorenen Mädchen von Gorleben“

RPTU Didaktik der Mathematik Sekundarstufen

Passende Nullhypothese

Die Wahrscheinlichkeit für eine Jungengeburt im 35 km-Umkreis um das Atommüll-Zwischenlager Gorleben ist nach der Einlagerung des Atommülls gleich der Wahrscheinlichkeit für eine Jungengeburt in Deutschland im selben Zeitraum:

$$H_0 : p = P(\text{Jungengeburt}) = 0{,}513$$

Bemerkung

Die Zahl 0,513 ergibt sich als arithmetisches Mittel der Anteile für eine Jungengeburt in Deutschland in den Jahren 1996 bis 2009.

Ergebnis

Der p-Wert – also die Wahrscheinlichkeit dafür, dass unter den 19209 Geburten im Zeitraum von 1996 bis 2009 **10040** oder mehr Jungen geboren wurden, unter der Annahme, dass die Nullhypothese H_0 gilt – lässt sich mit Hilfe der Binomialverteilung $B(n; p) = B(19209; 0{,}513)$ für die Zufallsvariable $\boldsymbol{X}$ = „Anzahl der geborenen Jungen“ wie folgt berechnen:

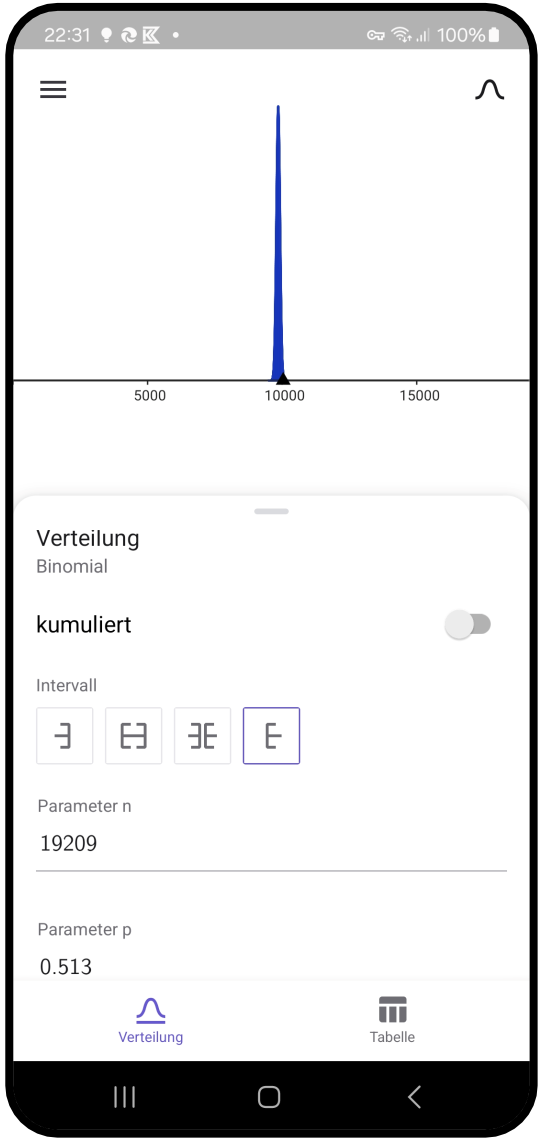
$$P(\boldsymbol{X} \geq \mathbf{10040} \text{ unter der Annahme, dass } H_0 \text{ gilt})$$

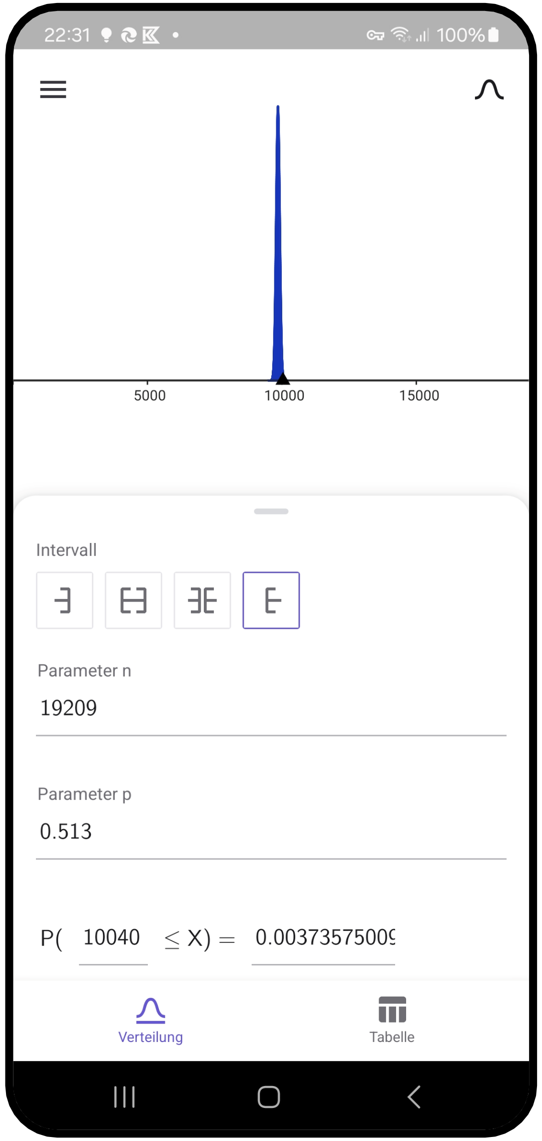
$$= \sum_{k=10040}^{19209} \binom{19209}{k} \cdot 0{,}513^k \cdot (1 - 0{,}513)^{19209-k}$$

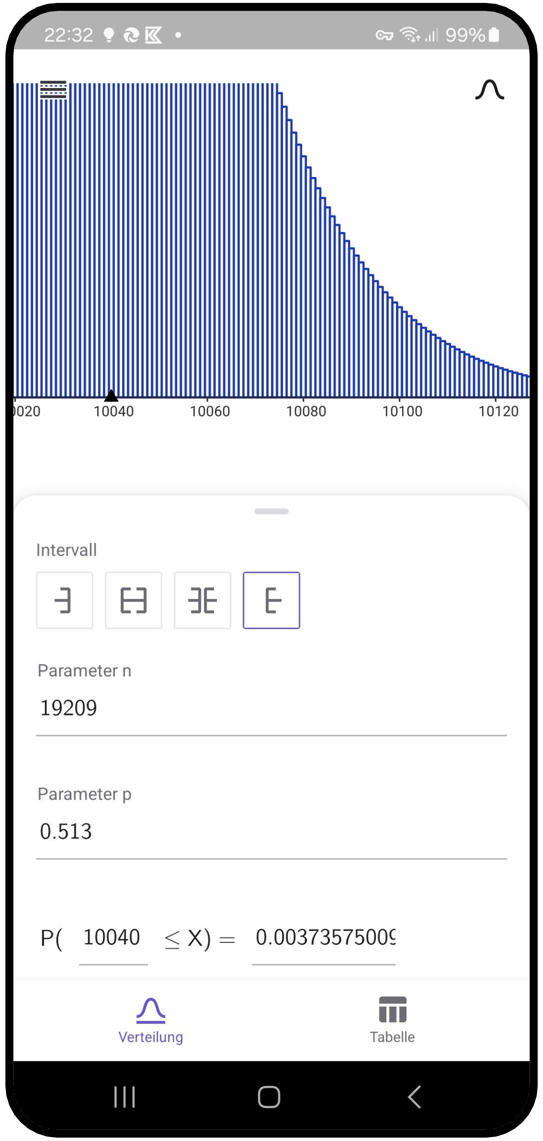
$$\approx 0{,}0037$$

Schäfer, A. (2017). Das Formulieren der Nullhypothese beim Signifikanztest mit Blick auf eine authentische Anwendung. *Stochastik in der Schule, 37*(3), 18-24

Nullhypothese
„Die verlorenen Mädchen von Gorleben“







 Schäfer, A. (2017). Das Formulieren der Nullhypothese beim Signifikanztest mit Blick auf eine authentische Anwendung. *Stochastik in der Schule, 37*(3), 18-24

Nullhypothese
„Die verlorenen Mädchen von Gorleben“

RPTU Didaktik der Mathematik Sekundarstufen

Kriterien zum Aufstellen von Nullhypothesen

- Es lohnt sich, die Nullhypothese zu untersuchen.
- Die Nullhypothese kann mit den vorliegenden Daten untersucht werden.
- Die Intention ist klar: Es geht um die Bewertung von Daten bzw. Stichprobenergebnissen.
- Die interessierende Population wird benannt.
- Die interessierenden Merkmale/Zufallsgrößen werden benannt und es sind entsprechende Daten verfügbar.

Statistische Nullhypothese

Die Wahrscheinlichkeit für eine Jungengeburt ist nach der Einlagerung des Atommülls in Gorleben gleich der Wahrscheinlichkeit für eine Jungengeburt in Deutschland im selben Zeitraum.

Was kann man statistisch untersuchen?

Mathematik

Welt

Substanzielle Nullhypothese

Der höhere Anteil an Jungengeburten hängt mit der Nähe zum Atommüll-Zwischenlager zusammen.

Schäfer, A. (2017). Das Formulieren der Nullhypothese beim Signifikanztest mit Blick auf eine authentische Anwendung. *Stochastik in der Schule, 37*(3), 18-24

Nullhypothese
„Die verlorenen Mädchen von Gorleben“

„Nullhypothesen“ von Studierenden

(1) „Der höhere Anteil an Jungengeburten hängt nicht mit der Nähe zum Atommüllzwischenlager zusammen.“

(2) „Es werden nicht signifikant mehr Mädchen als Jungen geboren.“

(3) „Die niedrige Mädchengeburtenrate liegt nicht an der Strahlung des radioaktiven Materials, sondern ist Zufall.“

(4) „Es werden mehr Mädchen als Jungen geboren.“

(5) „In Gorleben werden weniger Mädchen geboren.“

(6) „Im Umkreis von Gorleben werden im Vergleich zu Deutschland weniger Mädchen geboren.“

(7) „Es werden genauso viele Jungen geboren wie Mädchen.“

Bemerkung

Die Studierenden-Formulierungen der Nullhypothesen erfüllen die Kriterien zum Aufstellen von Nullhypothesen auf der vorherigen Folie oft nicht:

- Die Hypothesen **(1)** und **(3)** kann man mit den vorliegenden Daten nicht untersuchen.
- Bei den Hypothesen **(4)** bis **(7)** werden die interessierenden Merkmale falsch benannt.
- Nur bei Hypothese **(6)** wird die interessierende Population explizit angegeben …

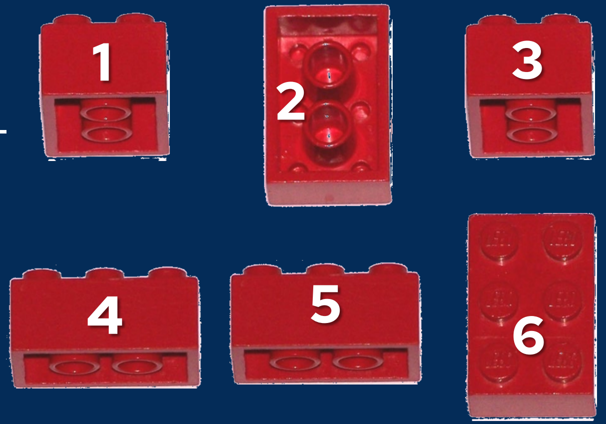
Schäfer, A. (2017). Das Formulieren der Nullhypothese beim Signifikanztest mit Blick auf eine authentische Anwendung. *Stochastik in der Schule, 37*(3), 18-24

Hypothesentest: Würfeln mit Lego-Stein oder Klemmbaustein

RPTU Didaktik der Mathematik Sekundarstufen

Situation: Frau Müller ist Mathematiklehrerin und möchte die Würfelversuche mit dem Lego-Baustein in Ihrem aktuellen Mathematik-Grundkurs durchführen. Ihre bisherigen Erfahrungen mit echten roten Lego-Steinen zeigen, dass die Häufigkeitsverteilungen für die verschiedenen Seiten bei allen näherungsweise gleich sind. Ihr aktueller Grundkurs ist sehr groß und die bisher verwendeten Legosteine reichen nicht für alle Schülerinnen und Schüler. Zum Glück hat sie in Größe, Gewicht & Form den Lego-Steinen vergleichbare andere gelbe Klemmbausteine zuhause. Sie fragt sich, ob sie diese einfach neben den Lego-Steinen verwenden kann, oder ob sie andere Wahrscheinlichkeiten für das Auftreten der Augenzahlen aufweisen.

LEGO-Stein
Aus langen Versuchsreihen geschätzte Wahrscheinlichkeitsverteilung für die roten Lego-Steine:



Augenzahl x	1	2	3	4	5	6
$P(X = x)$	0,007	0,500	0,007	0,086	0,086	0,314

Aufgabe: Entwerfen Sie eine Zufallsstichprobe und einen zugehörigen Hypothesentest, mit dem Sie die Frage von Frau Müller klären können und führen Sie den Hypothesentest anschließend durch.

Hypothesentest: Würfeln mit Lego-Stein  oder Klemmbaustein 

RPTU Didaktik der Mathematik Sekundarstufen

Schritt	**Fachliche Beschreibung** (Rückschauperspektive der „fertigen Mathematik“)	**Umsetzung am Beispiel** Würfeln mit Lego-Stein oder Klemmbaustein
(1)	Hypothese H aufstellen	**Mögliche Nullhypothese:** Das Auftreten der Augenzahl 2 ist beim gelben Klemmbaustein genauso wahrscheinlich wie beim roten Lego-Stein.
(2)	Hypothetische Verteilung realisieren	**Hypothetische Verteilung:** Passt die Verteilung des Lego-Würfels für den Klemmbaustein?
(3)	Art des Hypothesentests festlegen (zweiseitig, einseitig (welche Seite))	**Art des Hypothesentests:** Getestet wird zweiseitig, die Nullhypothese $H_0\colon P_{gelb}(X=2) = P_{rot}(X=2)$ gegen $H_1\colon P_{gelb}(X=2) \neq P_{rot}(X=2)$.
(4)	Signifikanzniveau festlegen	**Signifikanzniveau $(0,1\%)$:** Frau Müller will die Nullhypothese nur verwerfen, wenn die Stichprobe deutlich vom Erwartungswert abweicht.
(5)	Zufallsstichprobe festlegen (n) und durchführen	**Zufallsstichprobe:** 200-mal mit dem gelben Klemmbaustein würfeln und die Anzahl der dabei auftretenden Augenzahlen 2 festhalten.
(6)	Bei unwahrscheinlichen Ausgängen wird die Hypothese H verworfen.	Ergebnisse, die nicht im Prognoseintervall zur Sicherheitswahrscheinlichkeit 99,9% liegen, führen zum **Verwerfen der Nullhypothese**.

Hypothesentest: Würfeln mit Lego-Stein oder Klemmbaustein

Situation: Frau Müller ist Mathematiklehrerin und möchte die Würfelversuche mit dem Lego-Baustein in Ihrem aktuellen Mathematik-Grundkurs durchführen. Ihre bisherigen Erfahrungen mit echten roten Lego-Steinen zeigen, dass die Häufigkeitsverteilungen für die verschiedenen Seiten bei allen näherungsweise gleich sind. Ihr aktueller Grundkurs ist sehr groß und die bisher verwendeten Legosteine reichen nicht für alle Schülerinnen und Schüler. Zum Glück hat sie in Größe, Gewicht & Form den Lego-Steinen vergleichbare andere gelbe Klemmbausteine zuhause. Sie fragt sich, ob sie diese einfach neben den Lego-Steinen verwenden kann, oder ob sie andere Wahrscheinlichkeiten für das Auftreten der Augenzahlen aufweisen.

LEGO-Stein

Aus langen Versuchsreihen geschätzte Wahrscheinlichkeitsverteilung für den Lego-Stein:

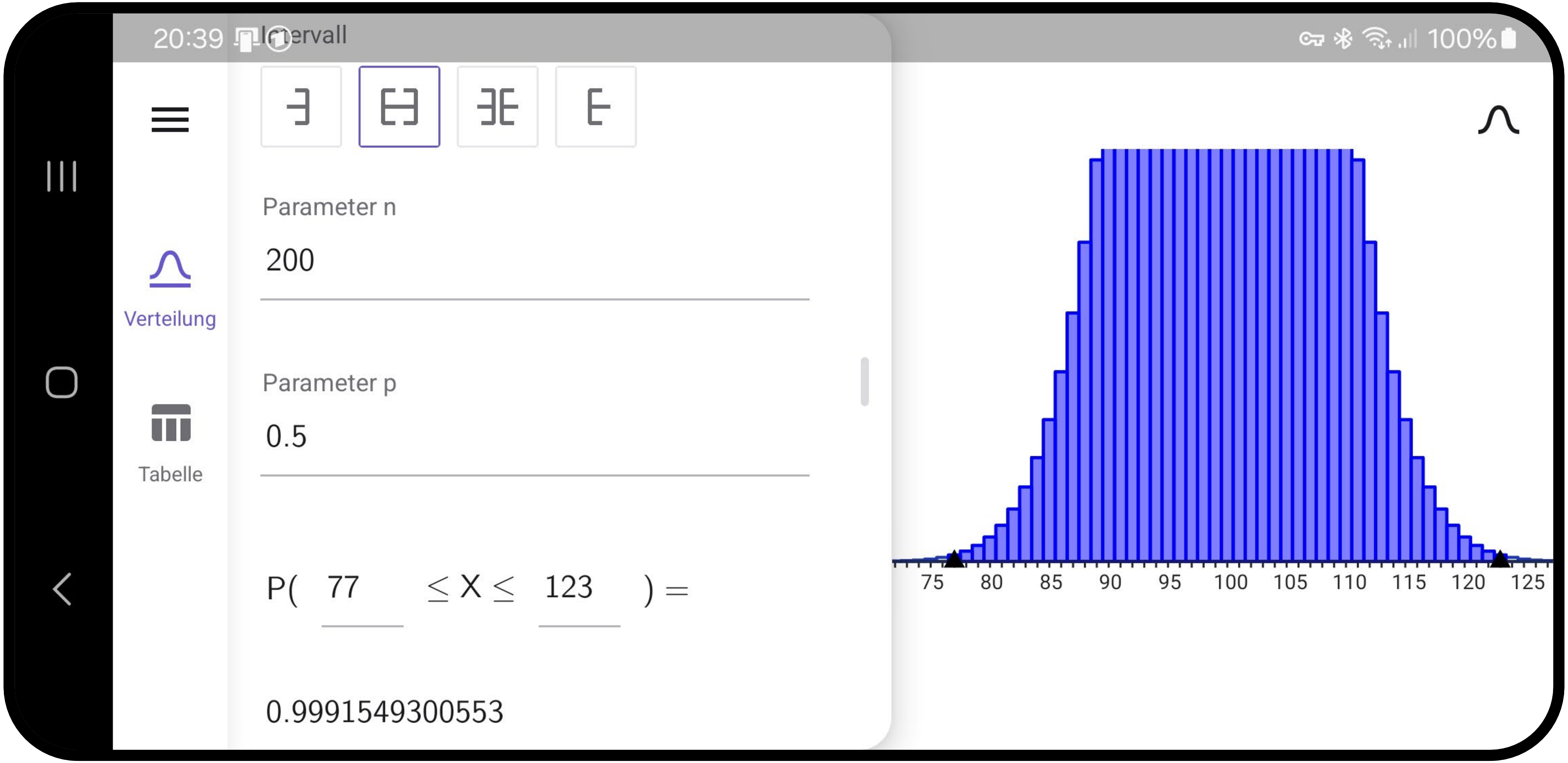
Augenzahl x	1	2	3	4	5	6
$P(X = x)$	0,007	0,500	0,007	0,086	0,086	0,314

Ergebnis der Zufallsstichprobe

Häufigkeitsverteilung nach $n = 200$ Würfen mit dem Klemmbaustein:

Augenzahl x	1	2	3	4	5	6
$H(X = x)$	21	72	12	33	34	28

Hypothesentest: Würfeln mit Lego-Stein oder Klemmbaustein

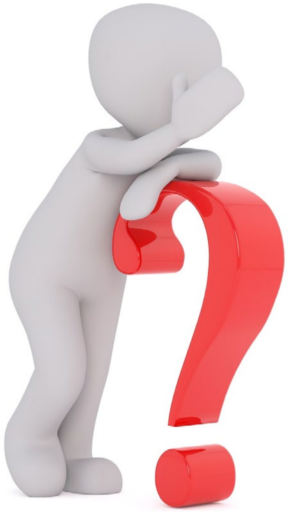


Exkurs: Fehler 1. Art und Fehler 2. Art

		Realität	
		H_0 richtig	H_0 falsch
Test	H_0 richtig	✓	✗ **Fehler 2. Art** (β-Fehler)
	H_0 falsch	✗ **Fehler 1. Art** (α-Fehler)	✓

Wird die Nullhypothese irrtümlich abgelehnt, so bezeichnet man dies als **Fehler erster Art**.

Das **Signifikanzniveau** ist der Wert, den die Wahrscheinlichkeit für den Fehler erster Art nicht überschreiten soll.



Wird die Nullhypothese irrtümlich nicht abgelehnt, so bezeichnet man dies als **Fehler zweiter Art**.

Aufgaben zur Erarbeitung der Ideen zu den Fehlern 1. und 2. Art: Ω-mathe

Was bedeutet „signifikant“?

https://de.wikipedia.org/wiki/P-Wert

Aufgabe
Mit einem Testverfahren wird überprüft, eine Stichprobe signifikant von einem Modell abweicht.

Es stellt sich heraus, dass die Abweichung auf dem 1%-Niveau signifikant ist.

Welche der folgenden Aussagen lassen sich aus dieser Tatsache folgern? (Es können dabei einzelne, alle oder auch keine der Aussagen zutreffen.)

(1) Es ist eindeutig bewiesen, dass die Nullhypothese falsch ist.
(2) Die Wahrscheinlichkeit des Zutreffens der Nullhypothese ist gefunden worden.
(3) Es ist eindeutig bewiesen, dass die Alternativhypothese wahr ist.
(4) Man kann nun die Wahrscheinlichkeit ableiten, dass die Alternativhypothese richtig ist.
(5) Entscheidet man sich nun, die Nullhypothese zu verwerfen, dann kennt man jetzt die Wahrscheinlichkeit dafür, dass diese Entscheidung falsch ist.
(6) Wenn man das Experiment sehr oft wiederholen würde, würde man in 99% der Fälle ein signifikantes Ergebnis bekommen.

Alle diese Aussagen sind falsch!

Der Begriff „Signifikanz“ bezieht sich auf „eine Welt“, in der H_0 gültig ist:

Man betrachtet die **Wahrscheinlichkeit des Stichprobenergebnisses oder noch extremerer Ergebnisse unter der Annahme, H_0 sei richtig**.

Diese „bedingte“ Wahrscheinlichkeit wird **p-Wert** genannt.

Ein Testergebnis ist genau dann auf dem 1%-Niveau signifikant. wenn der zum Testergebnis gehörige p-Wert kleiner als 1% ist.

 Krauss, S. & Wassner, C. (2001). Wie man das Testen von Hypothesen einführen sollte. Stochastik in der Schule, 21(1), 29-34

Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

MaTeGnu

Kapitel 4: Beurteilende Statistik

4.1 Zufallsvariable, Erwartungswert und Standardabweichung

4.2 Wahrscheinlichkeitsverteilung und Verteilungsfunktion

4.3 Prognose- und Konfidenzintervall

4.4 Hypothesentest

4.5 Ausgewählte Unterrichtsvorschläge

GeoGebra-Buch „Didaktik der Stochastik“ https://roth.tel/stochastik

dms.nuw.rptu.de/mategnu

RPTU

Unterrichtsvorschläge

Didaktik der Mathematik
Sekundarstufen

Geschmackstests

Riemer, W. & Petzolt, W. (1997). Geschmackstests: Spannende und verbindende Experimente. Mathematik lehren, 85, 16-19

Glücksrad auf schiefer Ebene

http://www.riemer-koeln.de/cmbasic/?gluecksrad Riemer, W. & Vehling, R. (2021). Stochastik erkunden – Ideenreiche Arbeitsblätter mit GeoGebra. Friedrich Verlag, 80-85

Bleistiftrollen – Beurteilende Statistik im Federmäppchen

http://www.riemer-koeln.de/cmbasic/?stifte
Riemer, W. & Seebach, G. (2011). Bleistiftrollen – Beurteilende Statistik im Federmäppchen. In R. Kaenders, R. Schmidt (Hrsg.). Mit GeoGebra mehr Mathematik verstehen (S. 69-83). Wiesbaden: Vieweg+Teubner

Zufällige Schwankungen untersuchen

Biehler, R., Griese, B., Niesporek, R. & Prömmel, A. (2022). Zufällige Schwankungen verstehen. *MatheWelt, 232*.

Zufallssinn

Leuders, T. (2005). Darf das denn wahr sein? – Eine schüleraktive Entdeckung der Grundidee des Hypothesentestens durch Simulation mit Tabellenkalkulation. *PM, 47*(4), 8-16

Viele weitere ausgearbeitete Unterrichtsbeispiele finden sich bei Wolfgang Riemer:
http://www.riemer-koeln.de/cmbasic/?kallmeyer-verlag

Kontakt

Prof. Dr. Jürgen Roth & Dr. Susanne Digel

RPTU
Rheinland-Pfälzische Technische Universität
Kaiserslautern-Landau
Didaktik der Mathematik (Sekundarstufen)
Fortstraße 7, 76829 Landau

j.roth@rptu.de; s.digel@rptu.de

juergen-roth.de
dms.nuw.rptu.de/mategnu

RPTU