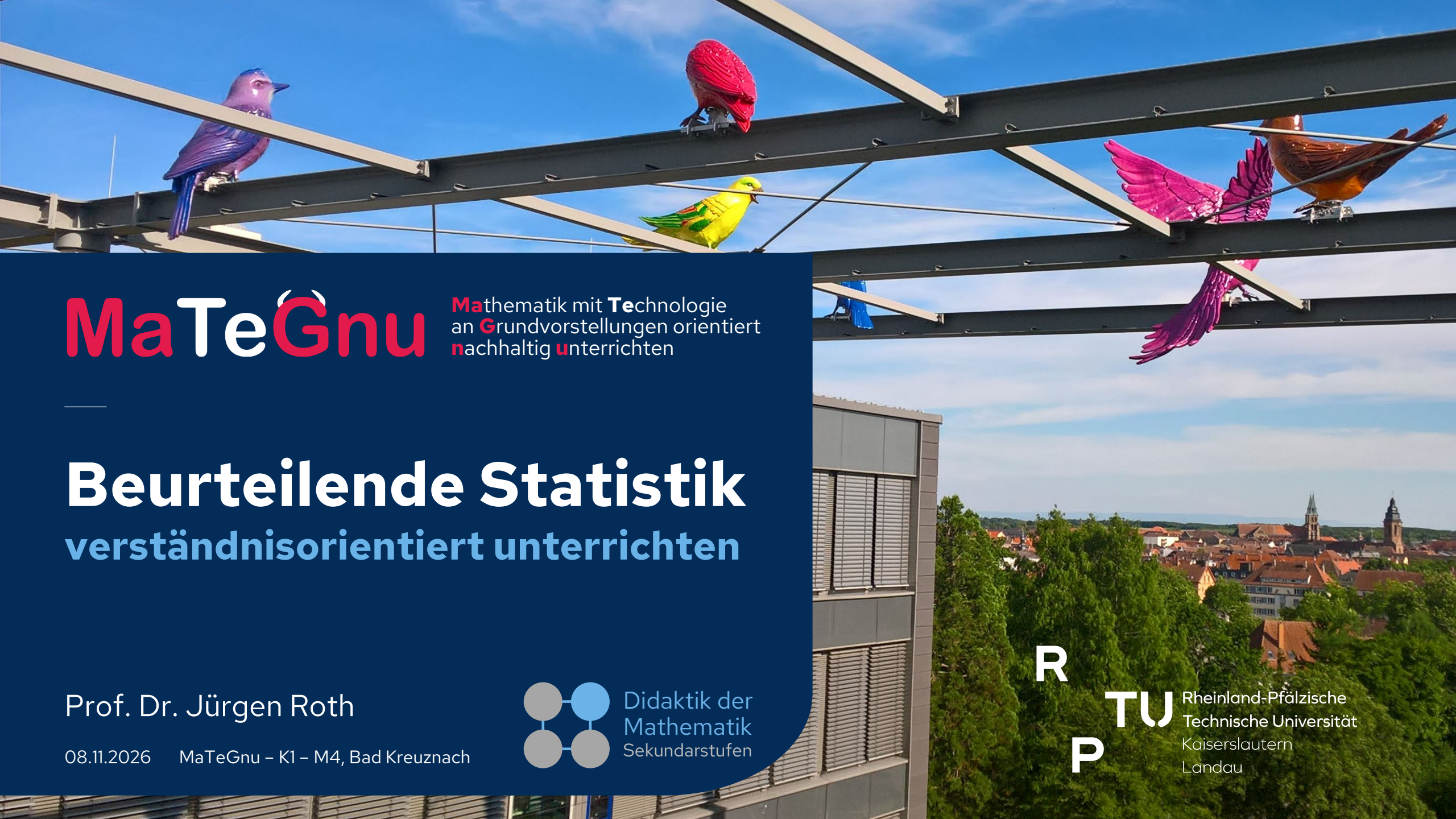


MaTeGnu

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

# Beurteilende Statistik

## verständnisorientiert unterrichten

Prof. Dr. Jürgen Roth

Didaktik der Mathematik Sekundarstufen

08.11.2026 MaTeGnu – K1 – M4, Bad Kreuznach

RPTU Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

# Kernideen der Stochastik

Didaktik der Mathematik Sekundarstufen



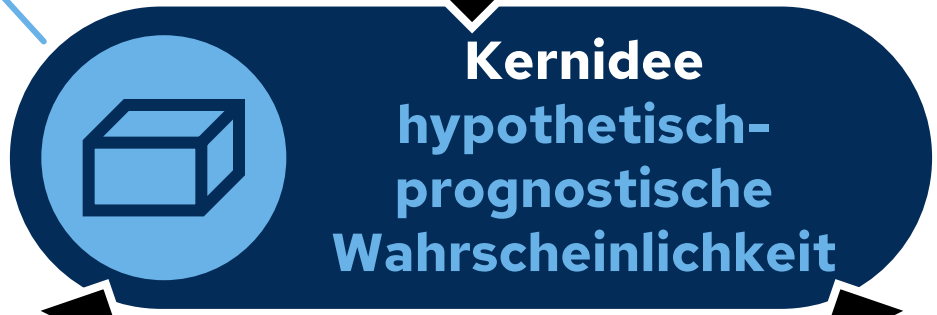
**Kernidee**
**Konditionale Denkfigur**

**„Wenn … , dann …"-Strukturen**

**Beispiel:**
Wenn man weiß, dass das Ereignis A eingetreten ist, wie groß ist dann die Wahrscheinlichkeit für das Eintreten von Ereignis B?

**Wahrscheinlichkeiten sind mathematische Modelle**, die …

- relative Häufigkeiten in der Realität vorhersagen und
- bei schlechter Passung zur Realität modifiziert werden.



**Kernidee**
**hypothetisch-prognostische Wahrscheinlichkeit**

**Globale Muster in Daten** …

- als Modell-Elemente zur Beschreibung der Realität
- als Vergleichselemente für die Struktur der Stichproben-Daten

**„Gute" Zufallsstichproben**

- Beschaffenheit und Generierung



**Kernidee**
**Repräsentativität und Variabilität von Stichproben-Daten**



**Kernidee**
**Verteilungen**



Ossadnik, H., Roth, J. & Engel, J. (2025). Kernideen zu Hypothesentests entwickeln und validieren. In L. Schick, M. Platz & A. Lambert (Hrsg.). Beiträge zum Mathematikunterricht 2025 (S. 1062-1065). WTM-Verlag.

# Kernidee: Hypothetisch-prognostische Wahrscheinlichkeit

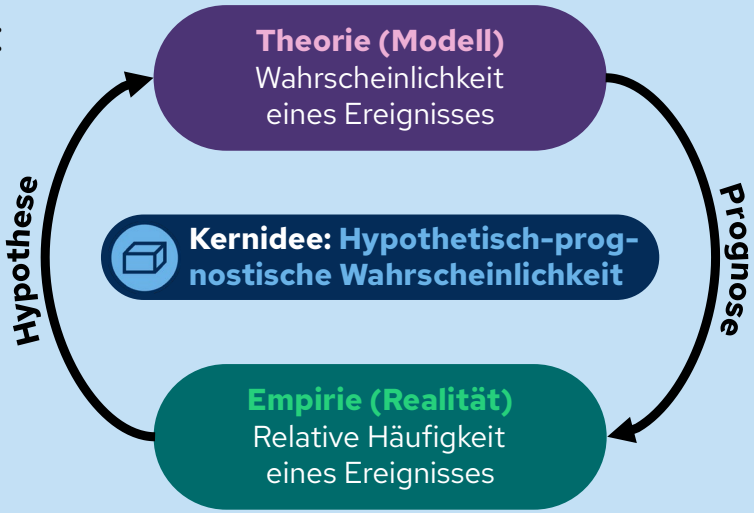


**Kernidee**
**Hypothetisch-prognostische Wahrscheinlichkeit**

**Wahrscheinlichkeiten sind mathematische Modelle**, die
- … relative Häufigkeiten in der Realität vorhersagen und
- … bei schlechter Passung zur Realität modifiziert werden.

## Hypothetisch-prognostische Wahrscheinlichkeit

- Diese Kernidee setzt Wahrscheinlichkeiten (Modelle) und relativen Häufigkeiten (Realität) durch zwei Perspektiven in Beziehung zueinander:
  - **Prognostisch:**
    Wahrscheinlichkeiten sind mathematische Modelle, die vorhersagen, um welchen Wert relative Häufigkeiten bei mehrfacher Versuchswiederholung in der Realität zufallsabhängig schwanken werden.
  - **Hypothetisch:**
    Modelle können angezweifelt und modifiziert werden, wenn sie die Realität bei mehrfacher Versuchswiederholung nicht gut abbilden.
- Diese Kernidee betont damit den Modellcharakter von Wahrscheinlichkeiten und die Tatsache, dass Modelle sich an der Realität messen lassen müssen.






Ossadnik, H., Roth, J. & Engel, J. (2025). Kernideen zu Hypothesentests entwickeln und validieren. In L. Schick, M. Platz & A. Lambert (Hrsg.). Beiträge zum Mathematikunterricht 2025 (S. 1062-1065). WTM-Verlag.

# Kernidee: Konditionale Denkfigur



**Kernidee**
**Konditionale Denkfigur**

## „Wenn … , dann …"-Strukturen

**Beispiel:** Wenn man weiß, dass das Ereignis A eingetreten ist, wie groß ist dann die Wahrscheinlichkeit für das Eintreten von Ereignis B?

## Wahrscheinlichkeit → Anteil (mit natürlichen Häufigkeiten)

- Übergehen von Wahrscheinlichkeiten zu Anteilen und klären, was dabei das Ganze (die Bezugsgröße) und was der Teil ist.
- So kann z. B. zwischen
  - totalen Wahrscheinlichkeiten (das Ganze ist die Grundgesamtheit) und
  - bedingten Wahrscheinlichkeiten (das Ganze ist die Teilmenge der Grundgesamtheit, die die Bedingung charakterisiert)

  unterschieden werden.
- Eine bedingte Wahrscheinlichkeit wird hier im Sinne des Anteils der Teilgruppe A an der durch B eingeschränkten Bezugsgruppe aufgefasst.

**Beispiel:**

**Das Ganze:** Mögl. Würfel-Ergebnisse

| | |
|---|---|
| 2 | 4 |
| 3 | 6 |
| 5 | 1 |

**Der Teil:** Primzahl

| | |
|---|---|
| 2 | 4 |
| 3 | 6 |
| 5 | 1 |

**Das Ganze:** Primzahl

| | |
|---|---|
| 2 | 4 |
| 3 | 6 |
| 5 | 1 |

**Der Teil:** Gerade Zahl



Post, M. & Prediger, S. (2020). „Nur das ‚und' ist anders, sonst ist es gleich". Mit bedingten Wahrscheinlichkeiten sprachsensibel umgehen. Mathematik 5-10, 14(53), 38–41.
Roth, J. (2026). Wahrscheinlichkeiten verstehen: Kernideen und geeignete Baumdiagramme nutzen. Preprint.

# Kernidee: Verteilungen



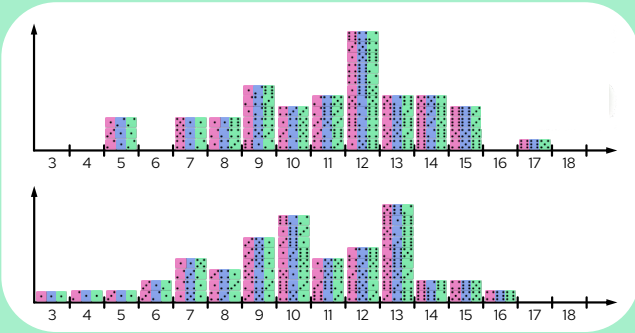
**Kernidee**
**Verteilungen**

## Globale Muster in Daten …

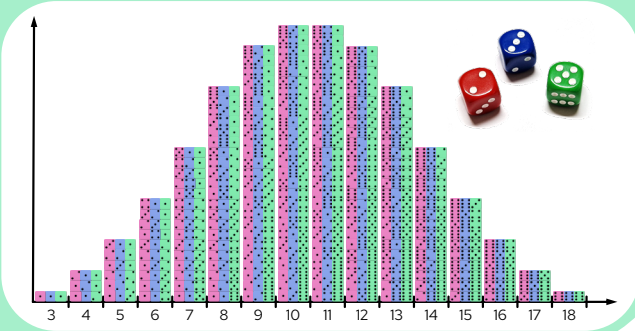
- als **Modell**-Elemente zur Beschreibung der Realität
- als **Vergleich**selemente für die Struktur der Daten in Stichproben

## Empirische Verteilung ↔ theoretische Verteilung

- **Empirische Verteilungen …**
  - basieren auf beobachteten Daten und beschreiben deren Variabilität
  - sind Verteilungen von Häufigkeiten bzw. relativen Häufigkeiten
  - dienen zum Vergleich der Struktur der Daten in verschiedenen Stichproben




- **Theoretische Verteilungen …**
  - modellieren und quantifizieren Variabilität
  - sind Verteilungen von Wahrscheinlichkeiten
  - sind Modelle zur Beschreibung der Realität






Ossadnik, H., Roth, J. & Engel, J. (2025). Kernideen zu Hypothesentests entwickeln und validieren. In L. Schick, M. Platz & A. Lambert (Hrsg.). Beiträge zum Mathematikunterricht 2025 (S. 1062-1065). WTM-Verlag.

# Kernidee: Repräsentativität und Variabilität von Stichproben-Daten



**Kernidee**
**Repräsentativität und Variabilität von Stichproben-Daten**

**Beschaffenheit und Generierung von „guten" Zufallsstichproben**

## Repräsentativität und Variabilität von Stichproben-Daten

- **Repräsentativität einer Stichprobe**
  - … bezüglich der Population reflektieren.
  - Entscheidet über die Generalisierbarkeit der Ergebnisse.
- **Akzeptanz der Variabilität von Stichprobendaten**
  - Verständnis dafür entwickeln, dass eine Stichprobe ein Element einer Menge von potenziellen anderen Stichproben ist.
  - Verständnis für das Gesetz der großen Zahlen anbahnen.





Ossadnik, H., Roth, J. & Engel, J. (2025). Kernideen zu Hypothesentests entwickeln und validieren. In L. Schick, M. Platz & A. Lambert (Hrsg.). Beiträge zum Mathematikunterricht 2025 (S. 1062-1065). WTM-Verlag.

# Lehrplanalternativen RLP

Didaktik der Mathematik Sekundarstufen

## Grundkurs
(ca. 42 Stunden)

**Stochastik 1**

**+ wahlweise**

**Stochastik 2 – B1**
Schätzen von Wahrscheinlichkeiten

**Stochastik 2 – B2**
Testen von Hypothesen

## Leistungskurs
(ca. 70 Stunden)

**Stochastik**

Online Workshop mit Jürgen Roth:
**Hypothesentest verständnisorientiert erarbeiten**
14.04.2027, 16:00-18:00, geogebra-rlp.de/zoom


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

# Aufbau der beurteilenden Statistik
## Grundkurs



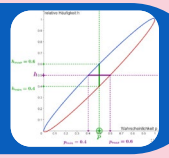
**Kernidee: Verteilungen**

**Kernidee: Repräsentativität und Variabilität von Stichproben-Daten**

**Schätzen von Wahrscheinlichkeiten**

Konfidenzintervall

Ellipsen-diagramm

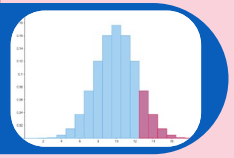


B 1

**Testen von Hypothesen**

Hypothesentest

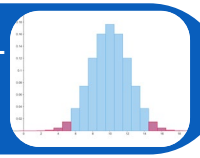
Histo-gramm



B 2

Verteilungsfunktion

Prognose-intervall



Normalverteilung
Näherung für große $n$

Binomialverteilung

Wahrscheinlich-keitsverteilung

Zufallsvariable

Erwartungswert

Standard-abweichung

**Kernidee: Konditionale Denkfigur**

**Kernidee: Hypothetisch-prog-nostische Wahrscheinlichkeit**

# Aufbau der beurteilenden Statistik
## Leistungskurs

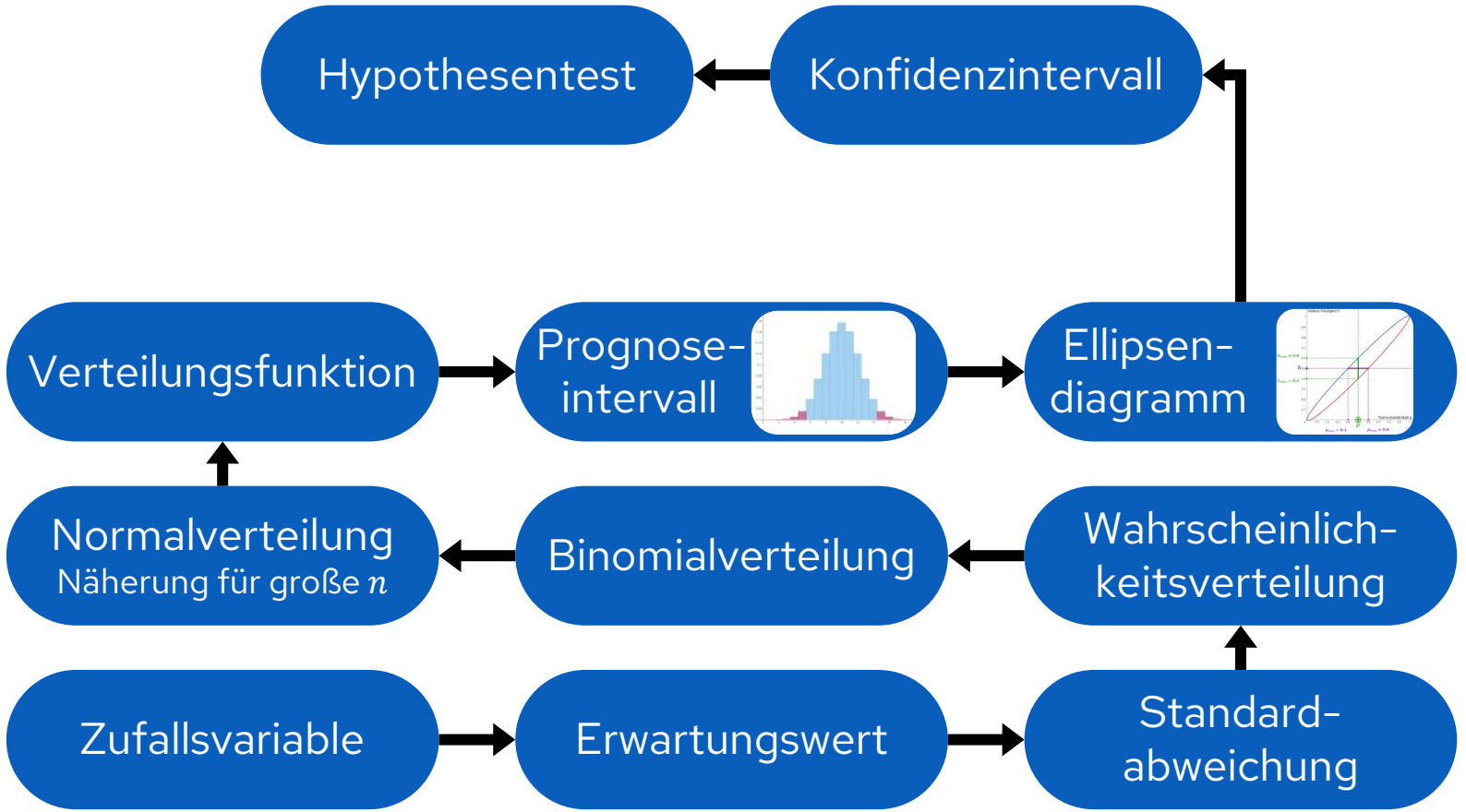


Kernidee: Verteilungen

Kernidee: Repräsentativität und Variabilität von Stichproben-Daten

Kernidee: Konditionale Denkfigur

Kernidee: Hypothetisch-prognostische Wahrscheinlichkeit



# Stochastik-Curriculum

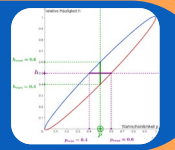
## Sekundarstufen I und II

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
- Stoch. (Un-)Abhängigkeit
- Hypothetisch-prognostische Wahrscheinlichkeit
- Bedingte Wahrscheinlichkeit



RPTU



Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

MaTeGnu

# Beurteilende Statistik

1. Hypothetisch-prognostische Wahrscheinlichkeit
2. Mehrstufige Zufallsexperimente: Baumdiagramme und Pfadregeln
3. Bedingte Wahrscheinlichkeit
4. Verteilungen: Binomialverteilung & Co
5. **Beurteilende Statistik**



Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

MaTeGnu

# 5. Beurteilende Statistik


5.1 Prognoseintervall

5.2 Konfidenzintervall

5.3 Hypothesentest

MaTeGnu

# 5.1 Prognoseintervall

# 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) | |

# Lehrplan RLP: Grundkurs

## Wahlpflichtgebiet B1: Schätzen von Wahrscheinlichkeiten

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 $\Phi$ (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.<br>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.

# Verständnisanker: Gläserne Urne

Didaktik der Mathematik Sekundarstufen

## Prognoseintervall

**Gläserne Urne**



**Modell ($p$)**

**Bekannt:**
**Trefferwahrscheinlichkeit $p$**
in der
**Grundgesamtheit**

schließen auf

**Wirklichkeit ($h$)**

?

**Gesucht:**
**Relative Häufigkeiten $h$**
mit denen in einer
**Stichprobe**
zu rechnen ist.

**Leitfrage zur gläsernen Urne**

Wie lassen sich Ergebnisse einer Stichprobenziehung vorhersagen (Prognose)?

| | Prognoseintervall |
|---|---|
| **Gesamtheit** | bekannt |
| **Modell** | durchsichtige **gläserne Urne** |
| **Stichprobe** | unbekannt |
| **Schluss von ... auf ...** | Gesamtheit → Stichprobe |
| **Fragestellung** | Wie viele Erfolge $k$ wird es bei $n$ Versuchen geben, wenn ich die Erfolgswahrscheinlichkeit kenne? |
| $n$: Anzahl der Versuche<br>$k$: Anzahl der Erfolge<br>$p$: Wahrscheinlichkeit<br>$h$: relative Häufigkeit | Das Prognoseintervall enthält alle relativen Häufigkeiten $h$ einer Stichprobe, die mit der bekannten Wahrscheinlichkeit $p$ verträglich sind. |



Ossadnik, H. & Engelhardt, A. (2024). Konfidenzintervalle verstehensorientiert unterrichten – Das Urnenmodell als Verständnisanker in einer digital angereicherten Lernumgebung. Mathematik lehren, 247, 38-44.

# Prognose aufstellen





KI-generiert

## Situation

- Eine (durchsichtige) **gläserne Urne** wird mit **15 blauen** und **25 grauen Kugeln** befüllt.
- Aus dieser Urne wird (ohne hinzusehen) $n = 50$-mal jeweils eine Kugel gezogen, festgehalten, ob sie blau ist, die Kugel in die Urne zurückgelegt und gemischt, bevor eine neue Kugel gezogen wird.
- Das Ergebnis ist eine Stichprobe aus $50$ Kugeln.
- **Prognose:**
  Mit welchen Anzahlen $k$ bzw. welchen relativen Häufigkeiten $h = \frac{k}{n}$ der gezogenen blauen Kugeln ist in der Stichprobe aus $n = 50$ Ziehungen zu rechnen?
- Kann man das vor dem Ziehen vorhersagen?
  Wie sicher sind wir uns gegebenenfalls?

# Idee: Stichprobe















# Verständnisanker: Gläserne Urne

Didaktik der Mathematik Sekundarstufen

**Situation**

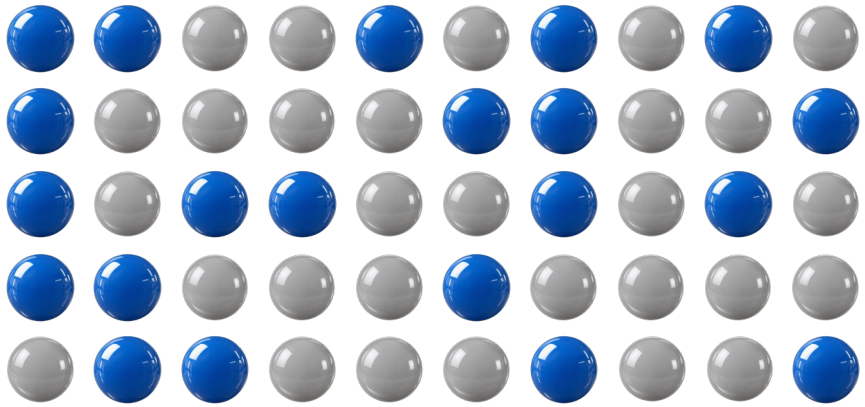
- Eine Urne enthält **15 blaue** und **25 graue** Kugeln.
- $n = 50$-mal eine Kugel mit Zurücklegen ziehen.

**Treffer**

- **Blaue Kugel** gezogen.
- Anzahl der Treffer: $k$



$k = 21$



**Prognoseintervall**

$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 95{,}4\%$

$[\,\mu - 2\sigma; \mu + 2\sigma\,] = [11{,}91; 25{,}59]$

**Voraussetzung:** Binomialverteilung?

Ja (Zurücklegen und Mischen)

$p = \frac{15}{40} = 0{,}375$ und $n = 50$

$P(X = k) = \binom{50}{k} \cdot 0{,}375^k \cdot (1 - 0{,}375)^{50-k}$

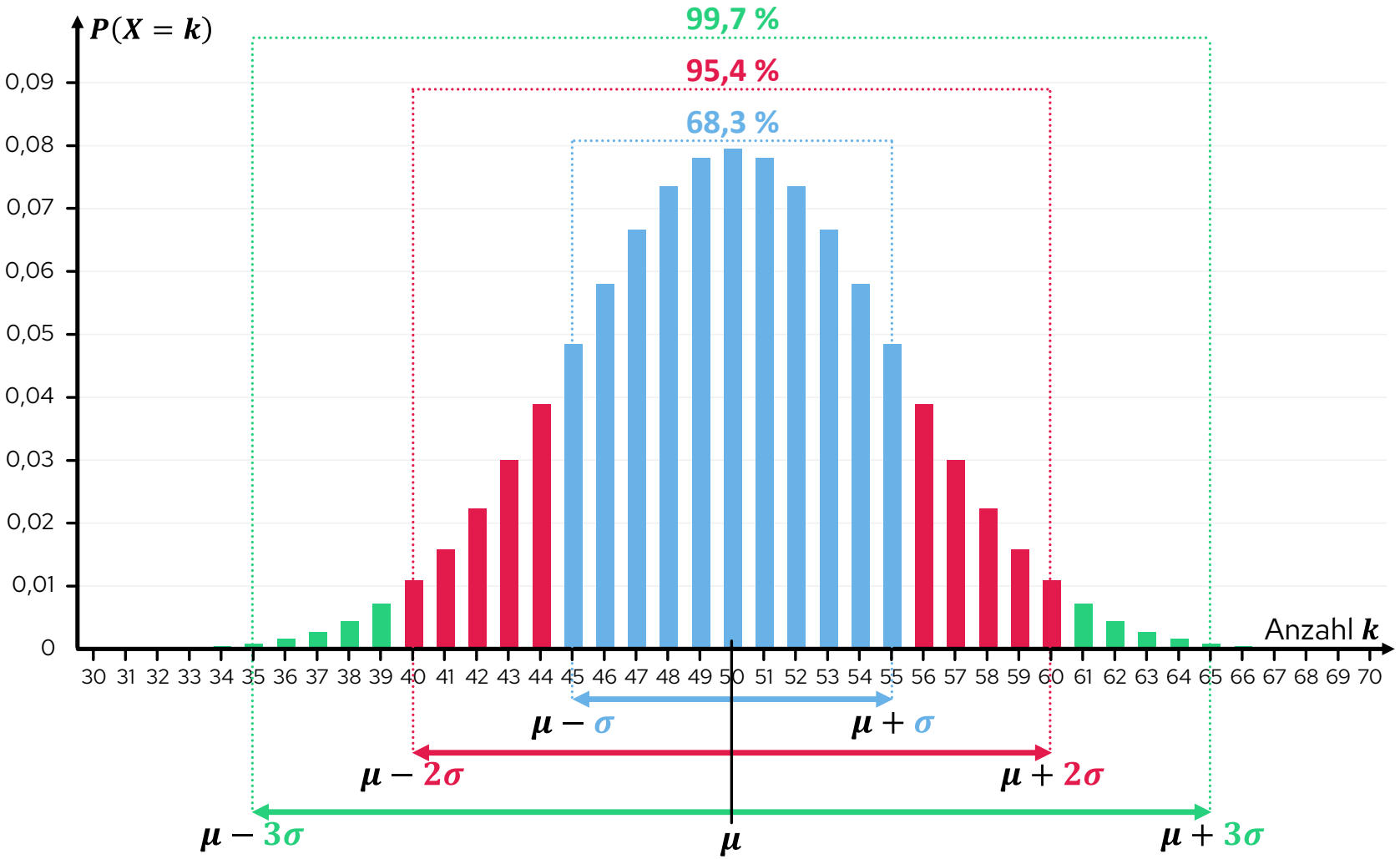
$\mu = n \cdot p = 18{,}75 \approx 19$

$\sigma = \sqrt{n \cdot p \cdot (1 - p)} \approx 3{,}42$

**Gewählt:** 95,4% Sicherheitswahrscheinl.

# Sigma-Regeln und Binomialverteilung





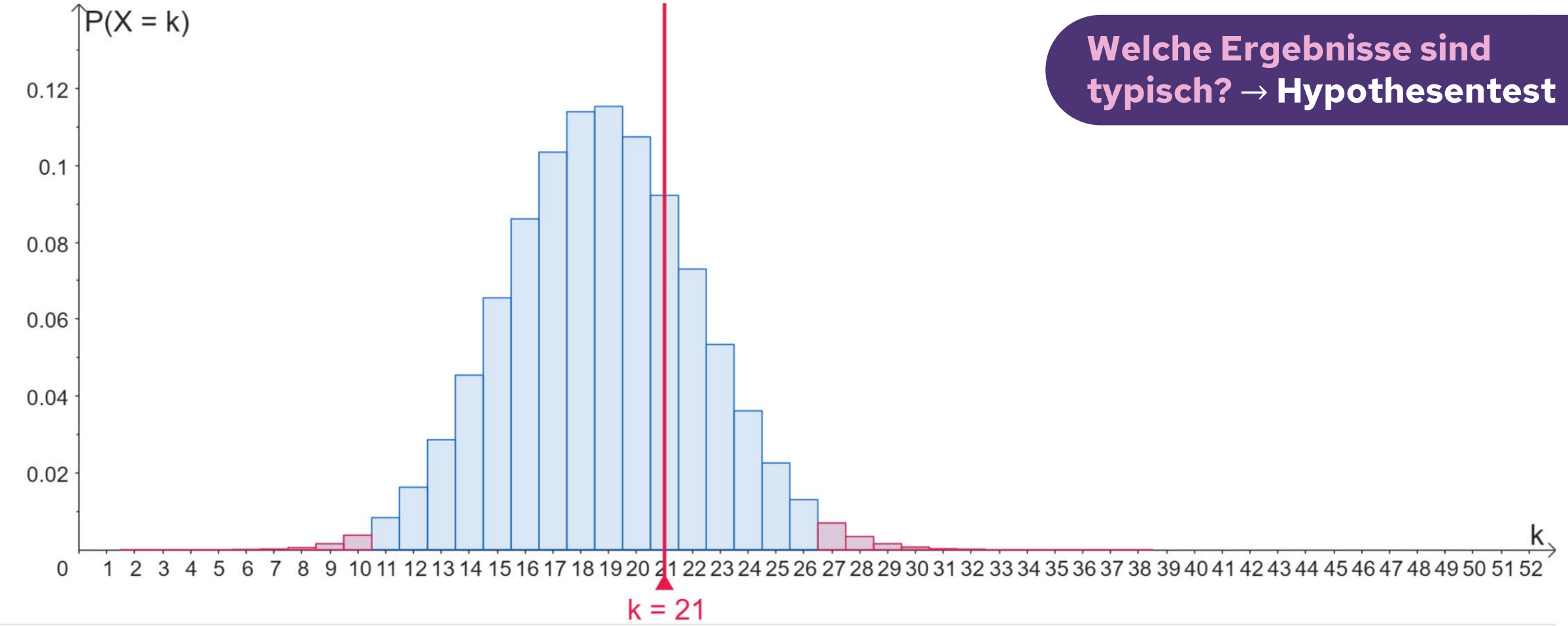

## Sigma-Regeln

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

- Die Sigma-Regeln gelten für Normalverteilungen.
- Bei Binomialverteilungen gelten sie näherungsweise, für $\sigma = \sqrt{n \cdot p \cdot (1-p)} > 3$.

# Ziehen mit Zurücklegen






**Welche Ergebnisse sind typisch? → Hypothesentest**

n = 50 $p_{blau}$ = 0.375 k = 21 c = 2

Sicherheitswahrscheinlichkeit: 95.4 %

# Daumenformeln für Prognoseintervalle





## $\sqrt{n}$-Daumenformel

- Die Länge des Prognoseintervalls für absolute Häufigkeiten wächst bei einem Stichprobenumfang $n$ proportional zu $\sqrt{n}$.
- Sie verdoppelt sich, wenn man den Stichprobenumfang vervierfacht.
- $[\, n \cdot p - \sqrt{n}; n \cdot p + \sqrt{n}\,]$

## $\frac{1}{\sqrt{n}}$-Daumenformel

- Die Länge des Prognoseintervalls für relative Häufigkeiten schrumpft bei einem Stichprobenumfang $n$ proportional zu $\frac{1}{\sqrt{n}}$.
- Sie halbiert sich, wenn man den Stichprobenumfang vervierfacht.
- $\left[\, p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}}\,\right]$

dividieren durch $n$

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

**Bemerkung:** Ersetzt man in den Termen der Intervallgrenzen **2** durch $c$, so lassen sich – nach den Sigma-Regeln – auch andere Sicherheitswahrscheinlichkeiten realisieren.



## 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. 95,4% Sicherheit im $2\sigma$-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. 95,4% 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.



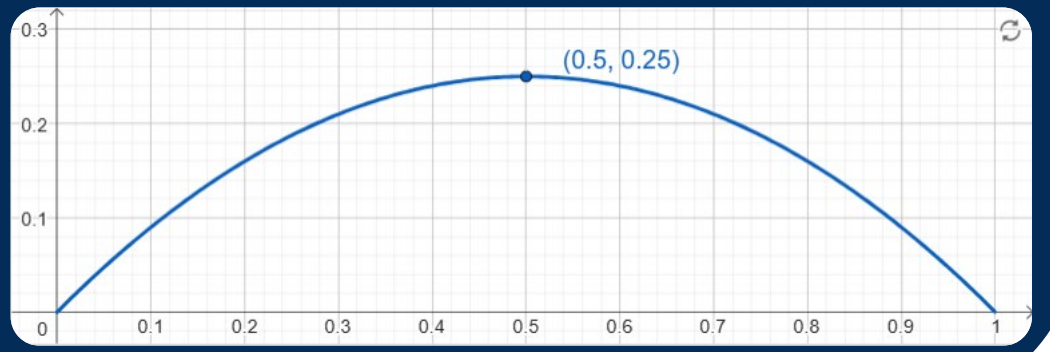
* **Bemerkung:** Die Teilmengeneigenschaften in beiden Sätzen ergeben sich aus folgender Ungleichung: $p \cdot (1-p) \leq \frac{1}{4}$

# 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:

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

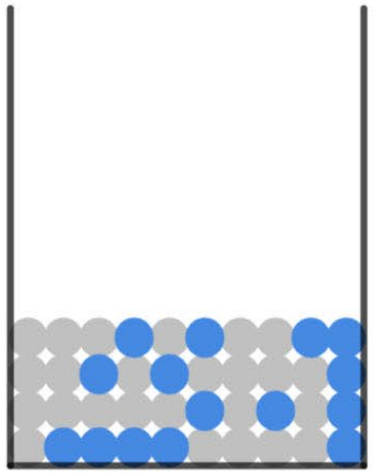
# Simulieren: Viele Versuchsreihen

Workshop

Zurücksetzen

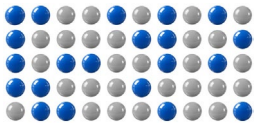
25 15

Urne befüllen



Eine Versuchsreihe

Zurück



Einstellungen zur Simulation

Simulieren

$\text{Anzahl}_{\text{Versuchsreihen}} = 400$

$\text{Länge}_{\text{Versuchsreihe}} = 50$

Einstellungen zur Simulation

☑ Grenzen fixieren

Anteil der Ergebnisse, die **im** Intervall liegen, an der Anzahl der Versuchsreihen:

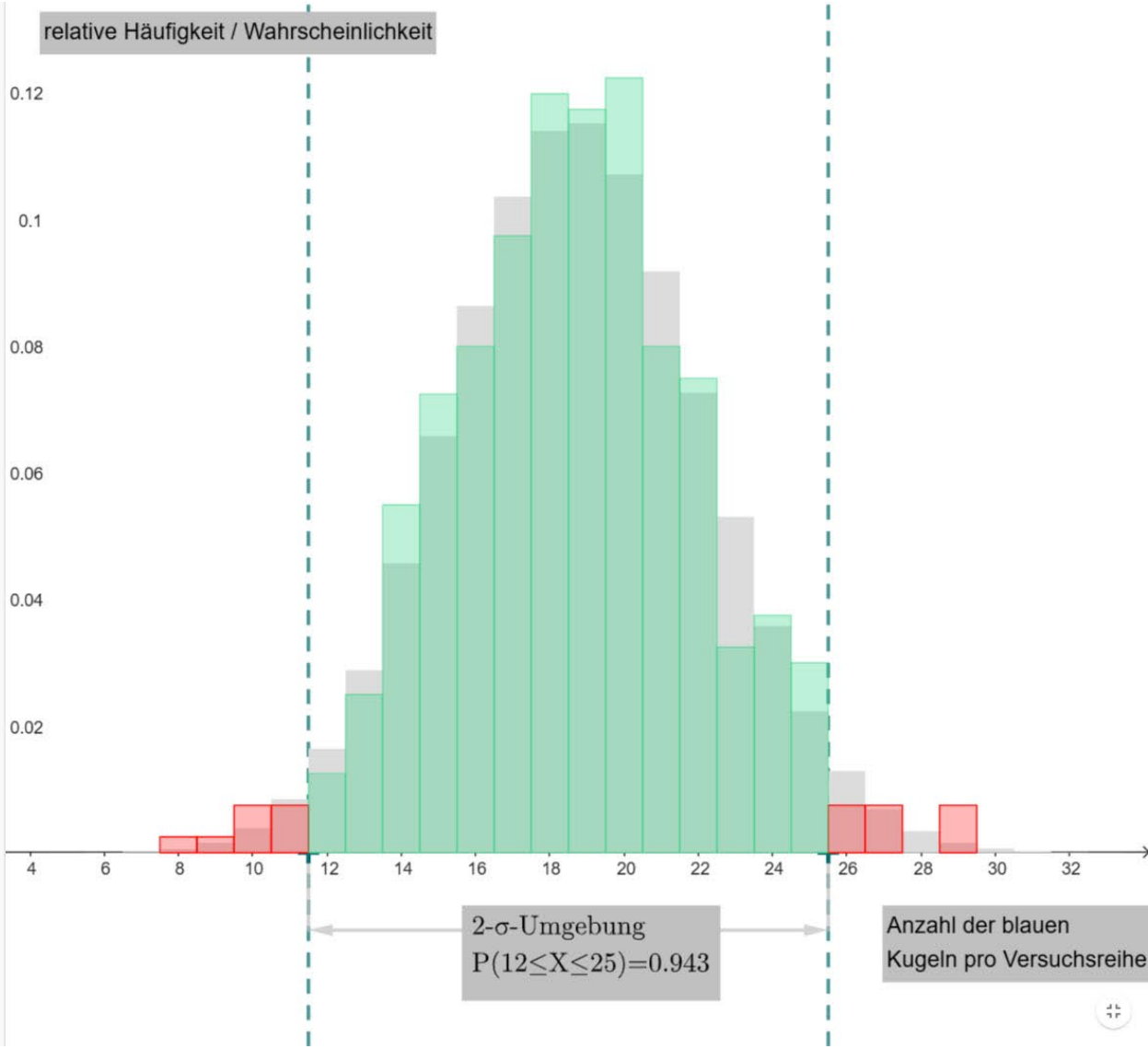
$$\frac{\text{Anzahl der Ergebnisse im Intervall}}{\text{Anzahl der Versuchsreihen}} = \frac{383}{400}$$

$= 0.958 = 95.75\%$

☐ Punktediagramm

☐ gestapeltesPunktediagramm

☑ Säulendiagramm ☑ Binomialverteilung






https://www.geogebra.org/m/wbmhsnba#material/yduyfnkn

Vielen Dank an Henrik Ossadnik für die Vorlage!

RPTU

# Prognoseintervall rechnerisch bestimmen



**Aufgabe:** Es werden 120 Würfe mit einem fairen Spielwürfel durchgeführt. Bestimmen Sie ein 95,4%-Prognoseintervall für die Anzahl $X$ (absolute Häufigkeit) der dabei geworfenen 6er.

| Bearbeitungsschritte | Ergebnisse |
|---|---|
| (1) Liegt eine Binomialverteilung vor? | Ja, mit $\boldsymbol{n} = 120$ und $\boldsymbol{p} = \frac{1}{6}$. |
| (2) Erwartungswert bestimmen | $\boldsymbol{\mu} = \boldsymbol{n} \cdot \boldsymbol{p} = 120 \cdot \frac{1}{6} = 20$ |
| (3) Standardabweichung bestimmen | $\boldsymbol{\sigma} = \sqrt{\boldsymbol{n} \cdot \boldsymbol{p} \cdot (1 - \boldsymbol{p})} = \sqrt{120 \cdot \frac{1}{6} \cdot \frac{5}{6}} \approx 4{,}1$ |
| (4) Untere Grenze Prognoseintervall für absH | $\boldsymbol{\mu} - 2\boldsymbol{\sigma} = 20 - 2 \cdot 4{,}1 = 11{,}8$ |
| (5) Obere Grenze Prognoseintervall für absH | $\boldsymbol{\mu} + 2\boldsymbol{\sigma} = 20 + 2 \cdot 4{,}1 = 28{,}2$ |
| (6) Prognoseintervall für absH unter Beachtung ganzzahliger Grenzen angeben („richtig runden“) | $[12; 28]$ |
| (7) Probe durchführen: Qualität beurteilen | $\boldsymbol{X} \sim \boldsymbol{B}_{120;\frac{1}{6}}$, mit $P(12 \leq \boldsymbol{X} \leq 28) = 0{,}96373$ |

# Prognoseintervall rechnerisch bestimmen

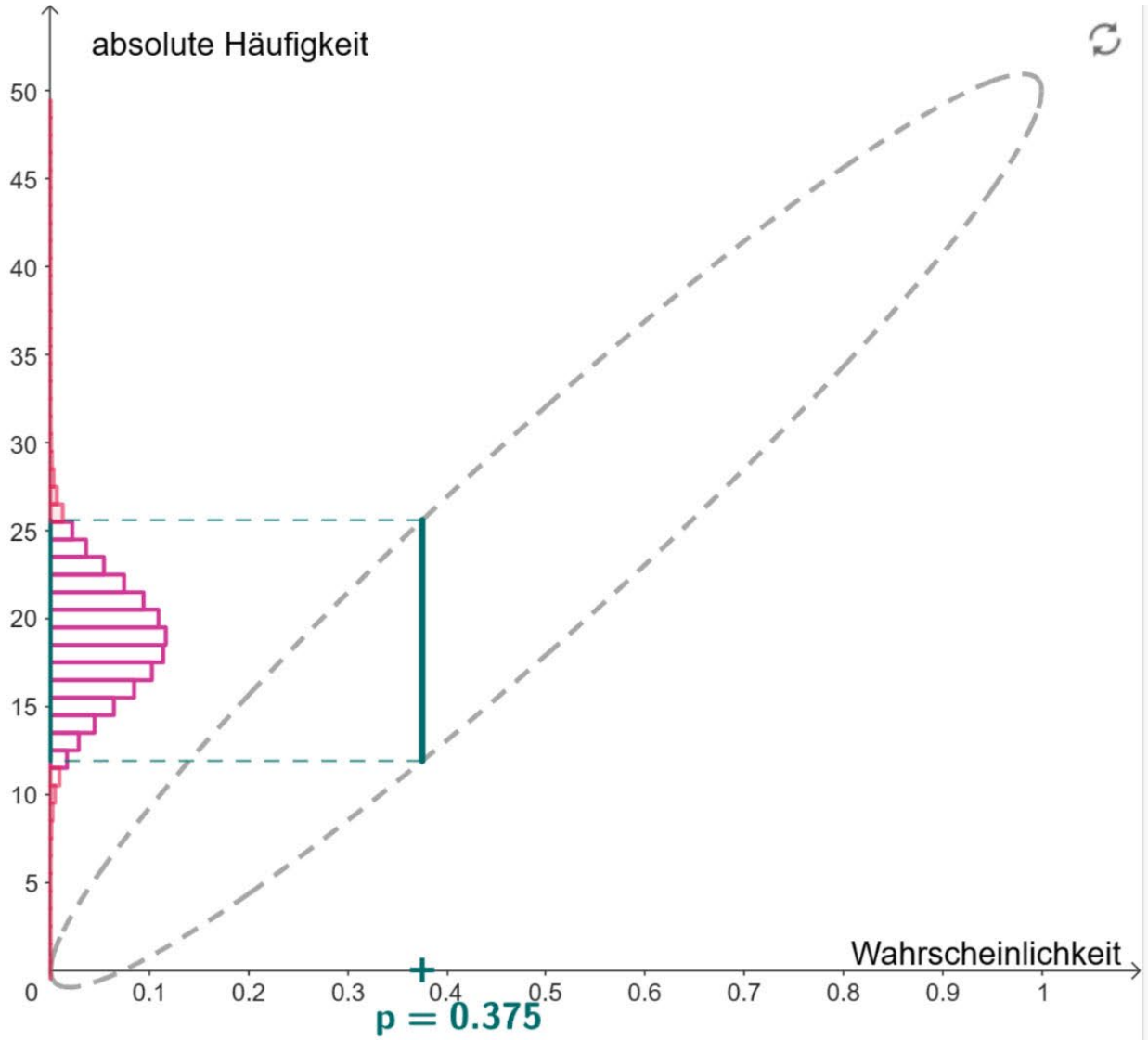


**Aufgabe:** Es werden 120 Würfe mit einem fairen Spielwürfel durchgeführt. Bestimmen Sie ein 95,4%-Prognoseintervall für die relative Häufigkeit $\boldsymbol{h}$ der dabei geworfenen 6er.

| Bearbeitungsschritte | Ergebnisse |
|---|---|
| (1) Liegt eine Binomialverteilung vor? | Ja, mit $\boldsymbol{n} = 120$ und $\boldsymbol{p} = \frac{1}{6}$. |
| (2) Untere Grenze 95,4%-Prognoseintervall für relative Häufigkeiten | $\boldsymbol{p} - 2 \cdot \frac{\sqrt{\boldsymbol{p} \cdot (1-\boldsymbol{p})}}{\sqrt{\boldsymbol{n}}} = \frac{1}{6} - 2 \cdot \frac{\sqrt{\frac{1}{6} \cdot \left(1 - \frac{1}{6}\right)}}{\sqrt{120}} \approx 0{,}099$ |
| (3) Untere Grenze 95,4%-Prognoseintervall für relative Häufigkeiten | $\boldsymbol{p} + 2 \cdot \frac{\sqrt{\boldsymbol{p} \cdot (1-\boldsymbol{p})}}{\sqrt{\boldsymbol{n}}} = \frac{1}{6} + 2 \cdot \frac{\sqrt{\frac{1}{6} \cdot \left(1 - \frac{1}{6}\right)}}{\sqrt{120}} \approx 0{,}235$ |
| (4) 95,4%-Prognoseintervall für rel. Häufigk. | $[0{,}099; 0{,}235]$ |
| (5) Probe durchführen: Qualität beurteilen | $\boldsymbol{X} \sim \boldsymbol{B}_{120;\frac{1}{6}}$, mit $P(0{,}099 \cdot 120 \leq \boldsymbol{X} \leq 0{,}235 \cdot 120) = 0{,}9637$ |

# Ellipsendiagramm verstehen

**Wozu dient das Ellipsendiagramm?**




Stichprobenumfang n = 50

p = 0.375

c = 2

**Sicherheitswahrscheinlichkeit:** 95.45 %

**Prognoseintervall zu absoluten Häufigkeiten**

$[n \cdot p - c \cdot \sqrt{n \cdot p \cdot (1-p)};\ n \cdot p + c \cdot \sqrt{n \cdot p \cdot (1-p)}]$

$\approx [11, 26]$

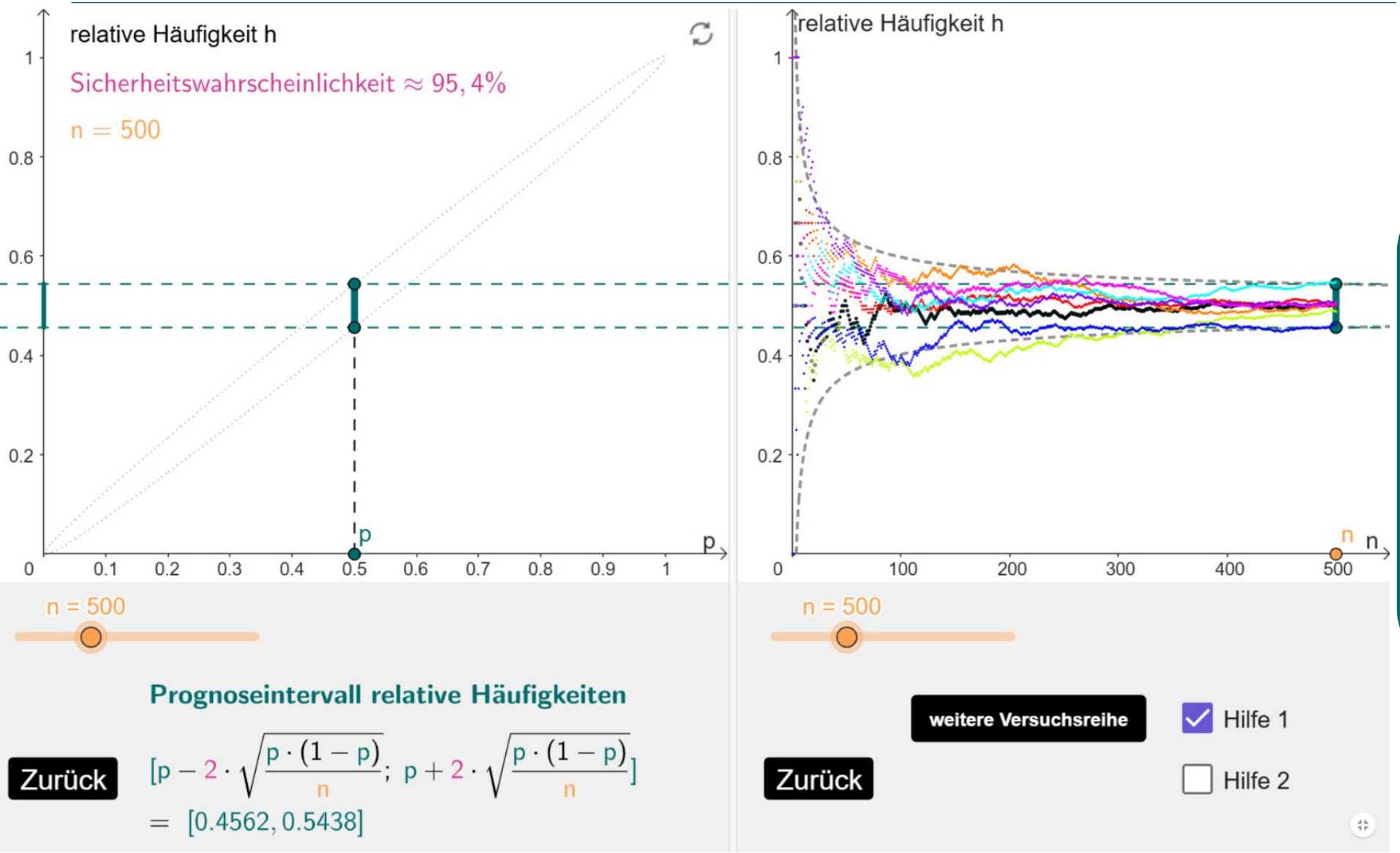
Weiter Zurücksetzen


https://www.geogebra.org/m/wbmhsnba#material/a8mag2hu

Vielen Dank an Henrik Ossadnik für die Vorlage!

# Wurzeltrichterdiagramm verstehen






## 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/wbmhsnba#material/spnvx8gc

Vielen Dank an Henrik Ossadnik für die Vorlage!

# Ellipsen- und Wurzeltrichterdiagramm

Didaktik der Mathematik Sekundarstufen

| Diagrammtyp | | Ellipsendiagramm | Wurzeltrichterdiagramm |
|---|---|---|---|
| **Was wird dargestellt?** | x-Achse | Wahrscheinlichkeit $p$ | Stichprobenumfang $n$ |
| | y-Achse | relative Häufigkeit $h$ | relative Häufigkeit $h$ |
| **Was kann abgelesen werden?** | | **Prognoseintervalle** in Abhängigkeit von der Wahrscheinlichkeit $p$ bei festem Versuchsumfang $n$<br>↑ **vertikales Ablesen** | **Prognoseintervalle** in Abhängigkeit vom Versuchsumfang $n$ bei fester Wahrscheinlichkeit $p$ |
| | | **Konfidenzintervalle**<br>→ **horizontales Ablesen** | --- |
| **Welche Funktion hat die Darstellung?** | | ■ Zusammenhang von Prognose- und Konfidenzintervall darstellen<br>■ Ein Konfidenzintervall enthält alle Wahrscheinlichkeiten $p$, in deren Prognoseintervall die beobachtete relative Häufigkeit $h$ liegt. | ■ Länge des Prognoseintervalls in Abhängigkeit von $n$ visualisieren<br>■ Hinführung zum $\frac{1}{\sqrt{n}}$-Gesetz |



Ossadnik, H. & Engelhardt, A. (2024). Konfidenzintervalle verstehensorientiert unterrichten – Das Urnenmodell als Verständnisanker in einer digital angereicherten Lernumgebung. Mathematik lehren, 247, 38-44.

MaTeGnu

# 5.2 Konfidenzintervall

# Prognoseintervall ↔ Konfidenzintervall

Lateinisch: confidere = vertrauen

Didaktik der Mathematik Sekundarstufen

## Prognoseintervall

**Gläserne Urne**



**Bekannt:**
**Trefferwahr-scheinlichkeit $p$**
in der
**Grundgesamtheit**

**Modell ($p$)**

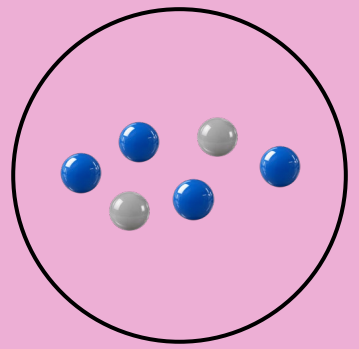
schließen auf




**Gesucht:**
**Relative Häufigkeiten $h$**
mit denen in einer
**Stichprobe**
zu rechnen ist.

**Wirklichkeit ($h$)**

## Konfidenzintervall



**Bekannt:**
**Relative Häufigkeit $h$**
in der
**Stichprobe**

**Wirklichkeit ($h$)**

schließen auf

**Schwarze Urne**




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

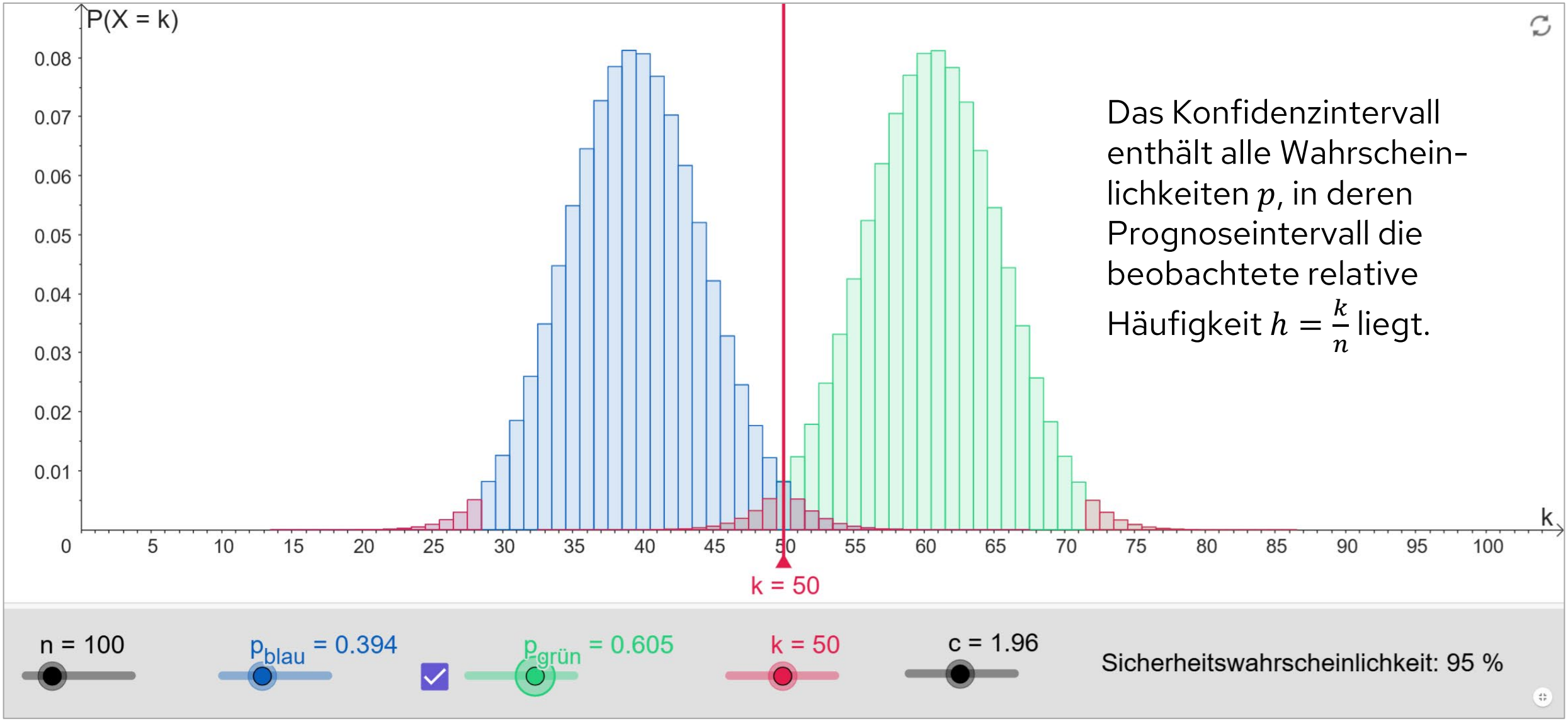
**Modell ($p$)**



Ossadnik, H. & Engelhardt, A. (2024). Konfidenzintervalle verstehensorientiert unterrichten – Das Urnenmodell als Verständnisanker in einer digital angereicherten Lernumgebung. Mathematik lehren, 247, 38-44.

# Konfidenzintervall






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

# Ellipsendiagramm – Prognoseintervall

**Bemerkung:** Ersetzt man in den Termen zur Berechnung der Intervallgrenzen **2** durch $c$, so lassen sich – nach den Sigma-Regeln – auch andere Sicherheitswahrscheinlichkeiten realisieren.

**Ellipsendiagramm**

☑ oberer blauer Teil

$$f_{oben}(p) = p + c \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}}$$

☑ unterer roter Teil

$$f_{unten}(p) = p - c \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}}$$

**Prognoseintervall**

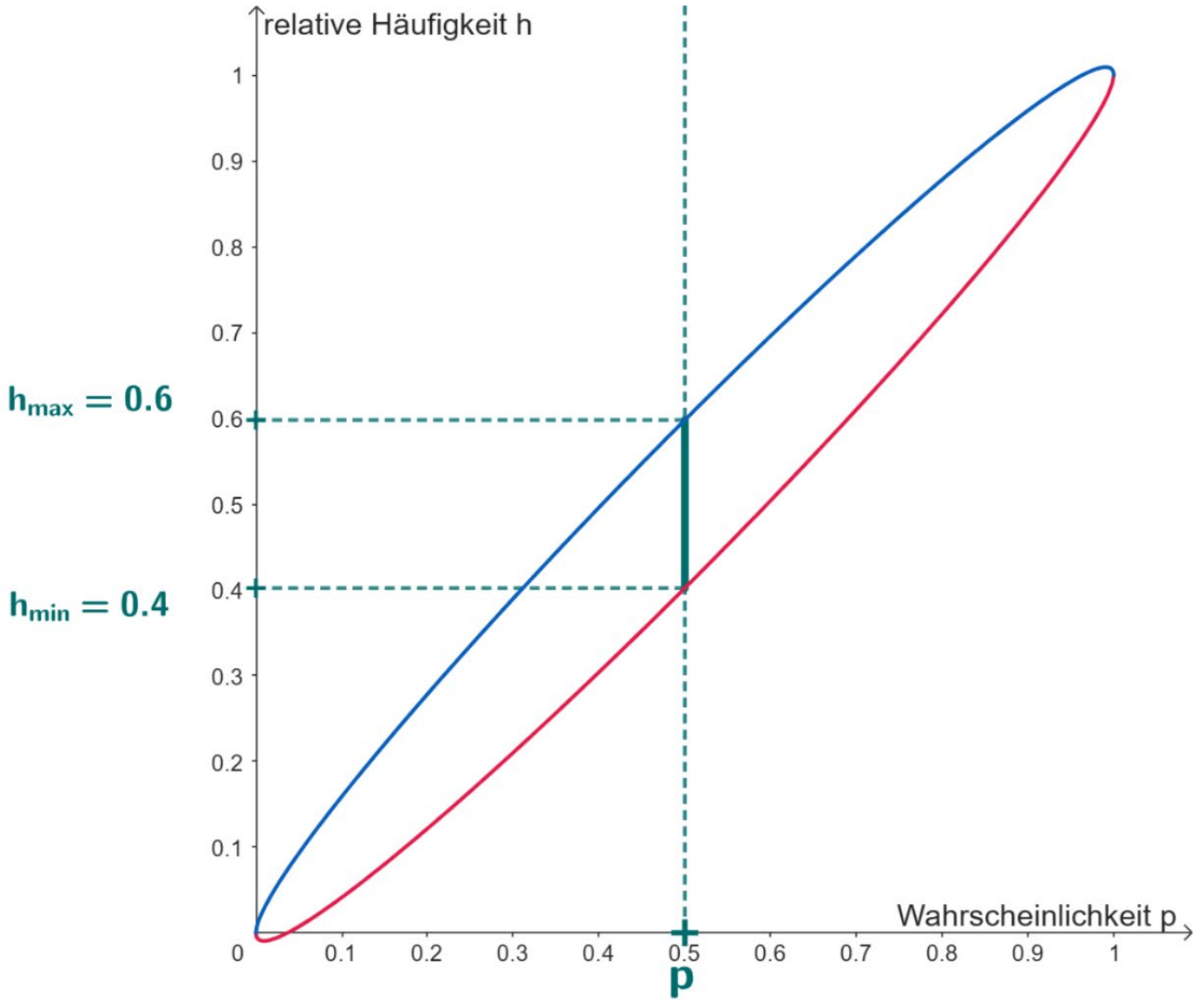
$$\left[ p - 2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} ; p + 2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} \right]$$

☑ Prognoseintervall

☐ Konfidenzintervall

c = 1.96

n = 100



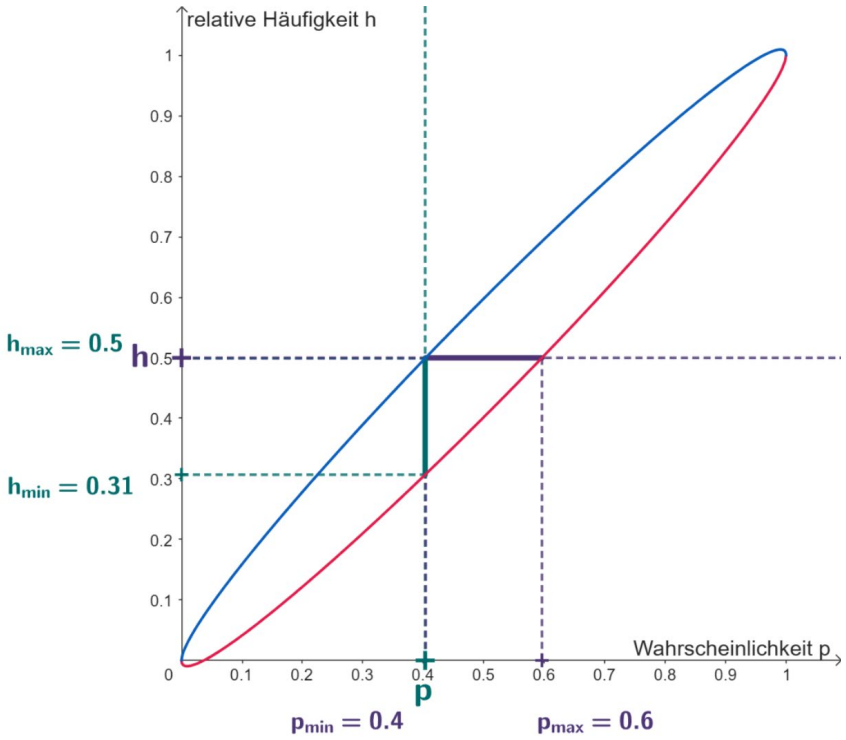

 https://www.geogebra.org/m/wbmhsnba#material/ypgqsvcd

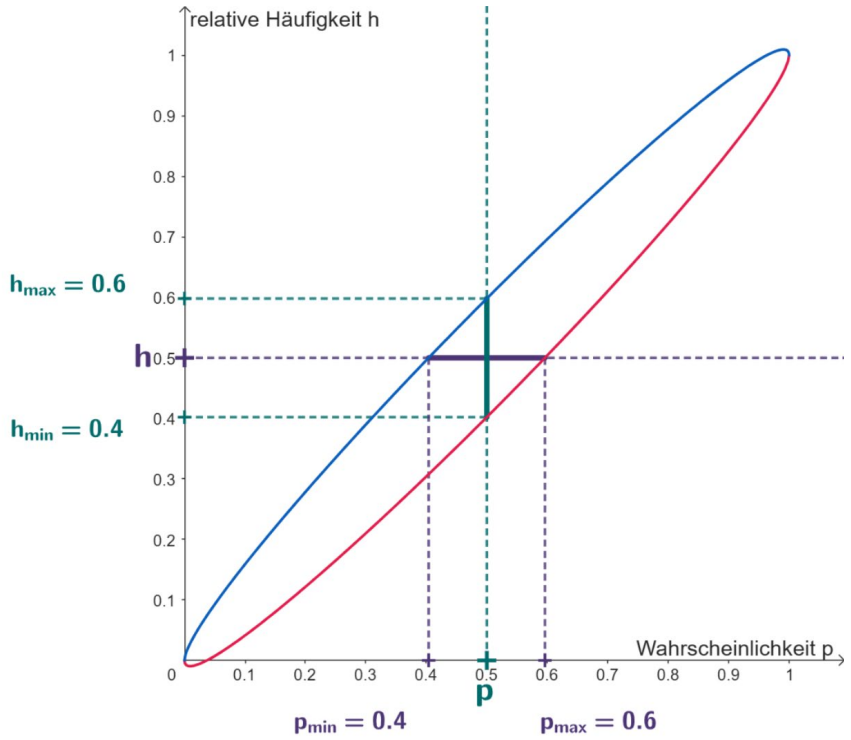
RPTU

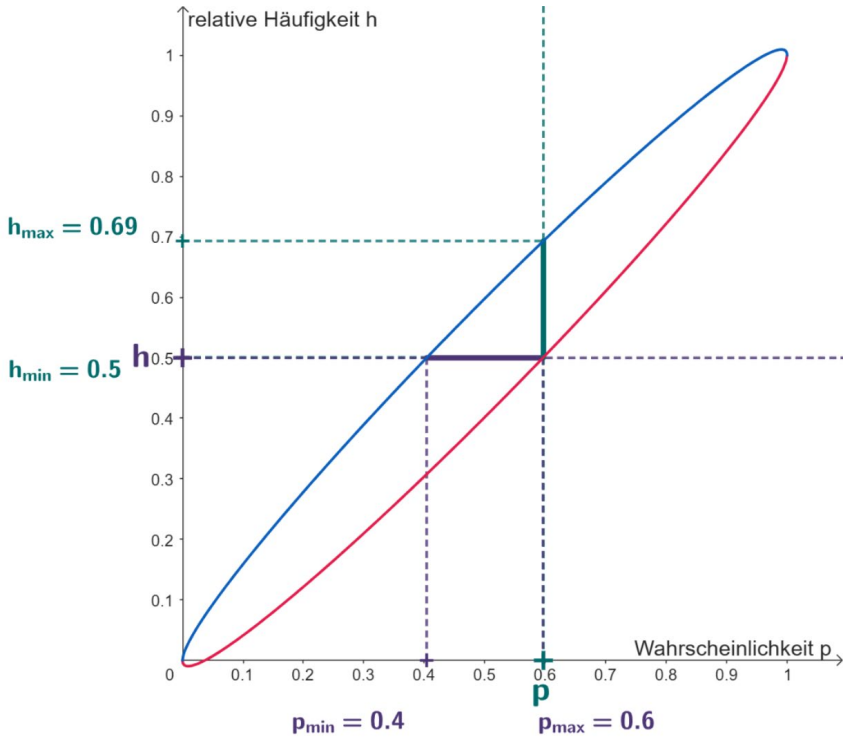
# Ellipsendiagramm – Konfidenzintervall[1]



**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.

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

[1] Lateinisch: confidere = vertrauen

# Verständnisanker: Schwarze Urne

Didaktik der Mathematik Sekundarstufen

Eine undurchsichtige **schwarze Urne** wurde mit **40 Kugeln** befüllt, von denen eine unbekannte Anzahl blau ist.

Es wird nun n – Mal hintereinander jeweils eine Kugel gezogen, deren Farbe bestimmt, die Kugel wieder in die Urne zurückgelegt und gemischt, bevor eine neue Kugel gezogen wird.

Die relative Häufigkeit **h** der dabei gezogenen blauen Kugeln wird notiert.

Gesucht ist ein Intervall in dem die Wahrscheinlichkeit **p** dafür liegt, dass beim einmaligen Ziehen eine blaue Kugel gezogen wird.

n = 50

Ziehung aus der Urne

☐

Weiter



**Workshop**



https://www.geogebra.org/m/wbmhsnba#material/tkxngsec

Vielen Dank an Henrik Ossadnik für die Vorlage!

# Ellipsendiagramm – Konfidenzintervall

Lateinisch: confidere = vertrauen

**Bemerkung:** Ersetzt man in den Termen zur Berechnung der Intervallgrenzen **2** durch $c$, so lassen sich – nach den Sigma-Regeln – auch andere Konfidenzniveaus realisieren.

## Grenzen des Konfidenzintervalls bestimmen

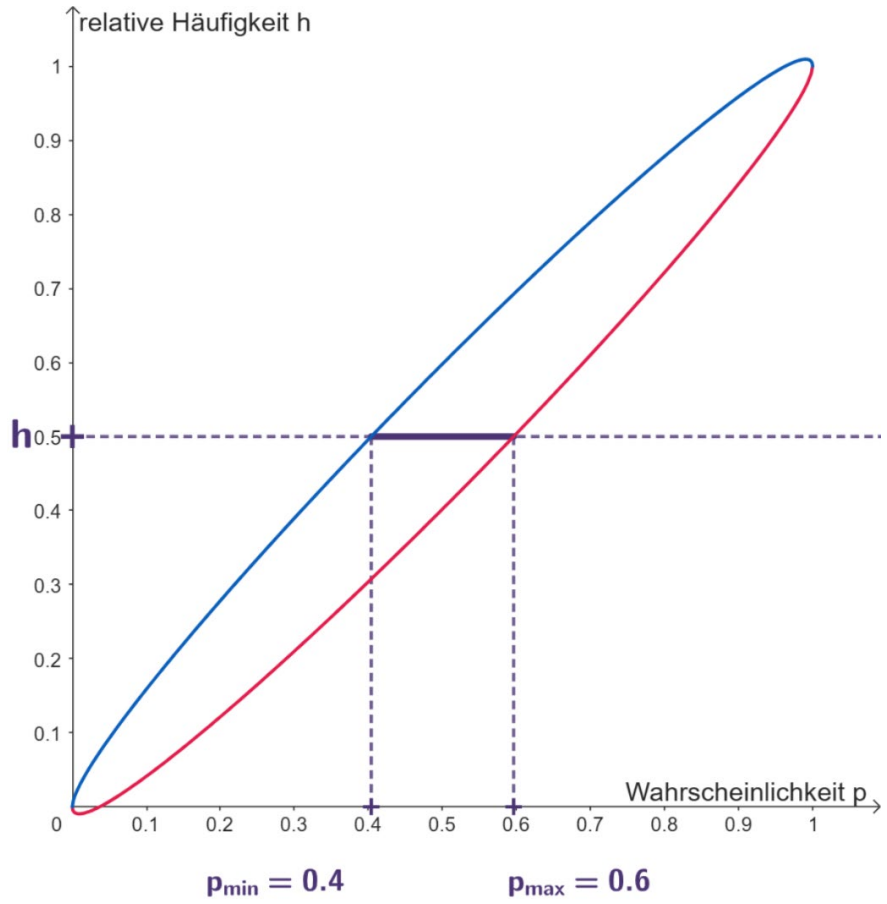
- Ablesen am Ellipsendiagramm oder
- (numerisches) Lösen folgender Gleichungen (*) und (**)
  (Die rechten Seiten sind die Terme für die Grenzen des jeweiligen Prognoseintervalls.)

**Linke Grenze** des Konfidenzintervalls:

$$h = p + 2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} \qquad (*)$$

**Rechte Grenze** des Konfidenzintervalls:

$$h = p - 2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} \qquad (**)$$




**Hinweis:** Die Lösungen der Gleichungen (*) und (**) können mit GeoGebra CAS bestimmt werden.

## (Wald-)Konfidenzintervall (Näherung für große $n$)

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

Daumenformel für Überschläge

**Bemerkung:**
Die Teilmengeneigenschaft $*$ ergibt sich aus der Ungleichung: $h \cdot (1-h) \leq \frac{1}{4}$.



RPTU

# Ellipsendiagramm – Konfidenzintervall

Lateinisch: confidere = vertrauen

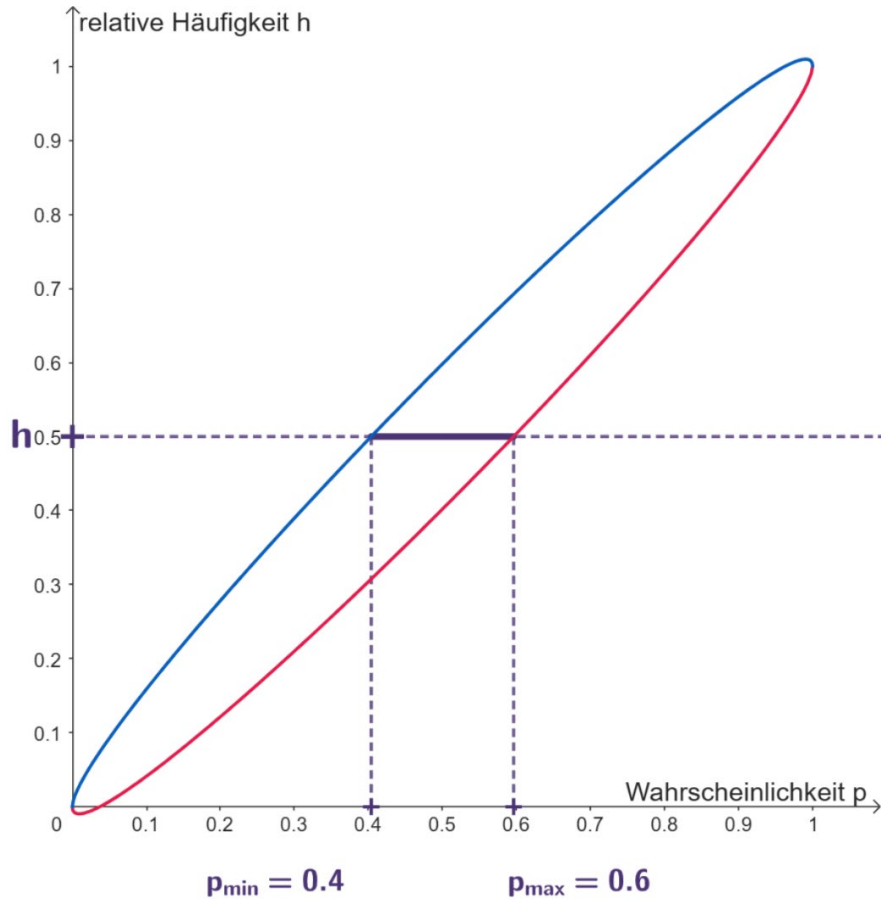
**Bemerkung:** Ersetzt man in den Termen zur Berechnung der Intervallgrenzen **2** durch $c$, so lassen sich – nach den Sigma-Regeln – auch andere Konfidenzniveaus realisieren.

## Grenzen des Konfidenzintervalls bestimmen

- Ablesen am Ellipsendiagramm oder
- (numerisches) Lösen folgender Gleichungen (*) und (**)
  (Die rechten Seiten sind die Terme für die Grenzen des jeweiligen Prognoseintervalls.)

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

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




## 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. mit CAS)

## (Wald-)Konfidenzintervall (Näherung für große $n$)

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

Daumenformel für Überschläge

**Bemerkung:**
Die Teilmengeneigenschaft $*$ ergibt sich aus der Ungleichung: $h \cdot (1-h) \leq \frac{1}{4}$.



RPTU

# Eigenschaften von Prognose- und Konfidenzintervallen



## 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 $\boldsymbol{\gamma = 95\%}$.
- Man nennt $\boldsymbol{\gamma = 95\%}$ Sicherheitswahrscheinlichkeit.
- Die Länge von Prognoseintervallen halbiert sich bei Vervierfachung des Versuchsumfangs.

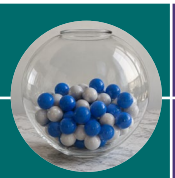
## Konfidenzintervall

- 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 $\boldsymbol{\gamma = 95\%}$.
- Man nennt $\boldsymbol{\gamma = 95\%}$ Konfidenzniveau (oder Vertrauenswahrscheinlichkeit).
- Bei Vervierfachung des Versuchsumfangs halbiert sich die Länge von Wilson-Konfidenzintervallen annähernd, von Wald-Konfidenzintervallen genau.


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

# Prognoseintervall ↔ Konfidenzintervall

* **Bemerkung:** Ersetzt man in den Termen zur Berechnung der Intervallgrenzen **2** durch $c$, so lassen sich – nach den Sigma-Regeln – auch andere $\gamma$-Werte realisieren (hier: $\gamma = 95{,}4\%$).

| | **Prognoseintervall** | **(Wald-)Konfidenzintervall** |
|---|---|---|
| **Gesamtheit** | bekannt  | unbekannt  |
| **Modell** | durchsichtige **gläserne Urne** | undurchsichtige **schwarze Urne** |
| **Stichprobe** | unbekannt | bekannt |
| **Schluss von … auf …** | Gesamtheit → Stichprobe | Stichprobe → Gesamtheit |
| **Fragestellung** | Wie viele Erfolge $\boldsymbol{k}$ wird es bei $\boldsymbol{n}$ Versuchen geben, wenn ich die Erfolgswahrscheinlichkeit kenne? | Welche Wahrscheinlichkeit hat ein Ereignis, das in $\boldsymbol{n}$ Versuchen $\boldsymbol{k}$-mal aufgetreten ist? |
| **95,4%-Intervall*** | $\left[ p - 2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} ; p + 2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} \right]$ | $\left[ h - 2 \cdot \frac{\sqrt{h \cdot (1-h)}}{\sqrt{n}} ; h + 2 \cdot \frac{\sqrt{h \cdot (1-h)}}{\sqrt{n}} \right]$ |
| **Ablesen im Ellipsendiagramm** | vertikal | horizontal |
| $\boldsymbol{n}$: Anzahl der Versuche<br>$\boldsymbol{k}$: Anzahl der Erfolge<br>$\boldsymbol{p}$: Wahrscheinlichkeit<br>$\boldsymbol{h}$: relative Häufigkeit | Das Prognoseintervall enthält alle relativen Häufigkeiten $\boldsymbol{h}$ einer Stichprobe, die mit der bekannten Wahrscheinlichkeit $\boldsymbol{p}$ verträglich sind. | Das Konfidenzintervall enthält alle Wahrscheinlichkeiten $\boldsymbol{p}$, in deren Prognoseintervall die beobachtete relative Häufigkeit $\boldsymbol{h}$ liegt. |

MaTeGnu

# 5.3 Hypothesentest

# Hypothesentest: Kernideen im Fokus



**Kernidee**
**Konditionale Denkfigur**

**„Wenn … , dann …"-Strukturen**

**Beispiel:**
Wenn man weiß, dass das Ereignis A eingetreten ist, wie groß ist dann die Wahrscheinlichkeit für das Eintreten von Ereignis B?

**Wahrscheinlichkeiten sind mathematische Modelle**, die …

- relative Häufigkeiten in der Realität vorhersagen und
- bei schlechter Passung zur Realität modifiziert werden.

**Kernidee**
**hypothetisch-prognostische Wahrscheinlichkeit**

**Globale Muster in Daten** …

- als Modell-Elemente zur Beschreibung der Realität
- als Vergleichselemente für die Struktur der Stichproben-Daten

**„Gute" Zufallsstichproben**

- Beschaffenheit und Generierung

**Kernidee**
**Repräsentativität und Variabilität von Stichproben-Daten**

**Kernidee**
**Verteilungen**



Ossadnik, H., Roth, J. & Engel, J. (2025). Kernideen zu Hypothesentests entwickeln und validieren. In L. Schick, M. Platz & A. Lambert (Hrsg.). Beiträge zum Mathematikunterricht 2025 (S. 1062-1065). WTM-Verlag.

# Lehrplan RLP: Leistungskurs



| | **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.<br><br>Besondere Bedeutung kommt der Interpretation des Ergebnisses eines Hypothesentests zu. Dabei sollen die Schülerinnen und Schüler auch die Grenzen des Verfahrens erkennen.<br><br>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

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.<br>Besondere Bedeutung kommt der Interpretation des Ergebnisses eines Hypothesentests zu. Dabei sollen die Schüler/innen auch die Grenzen des Verfahrens erkennen.<br>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.

# Warum Hypothesentest im MU?

Didaktik der Mathematik Sekundarstufen

| | |
|---|---|
| **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

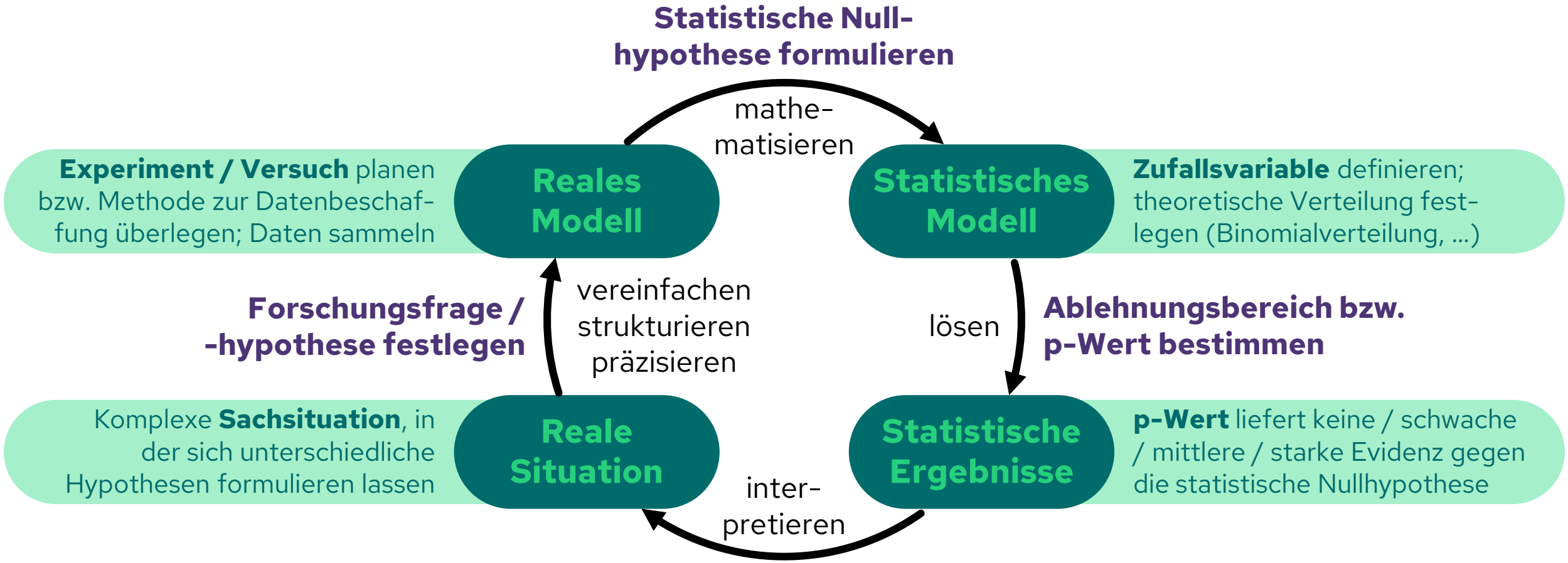
- Begriffe der Wahrscheinlichkeitsrechnung werden produktiv zusammengeführt.
  - Wiederholen, in Beziehung setzen und anwenden
  - oder aus Beurteilungssituationen herleiten
- Arbeitsweise von empirisch arbeitenden Wissenschaftlern exemplarisch erfahrbar:
  - Aufstellen von Hypothesen
  - Erproben im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

**Ziel:** Tragfähige Grundvorstellungen zum Hypothesentest aufbauen



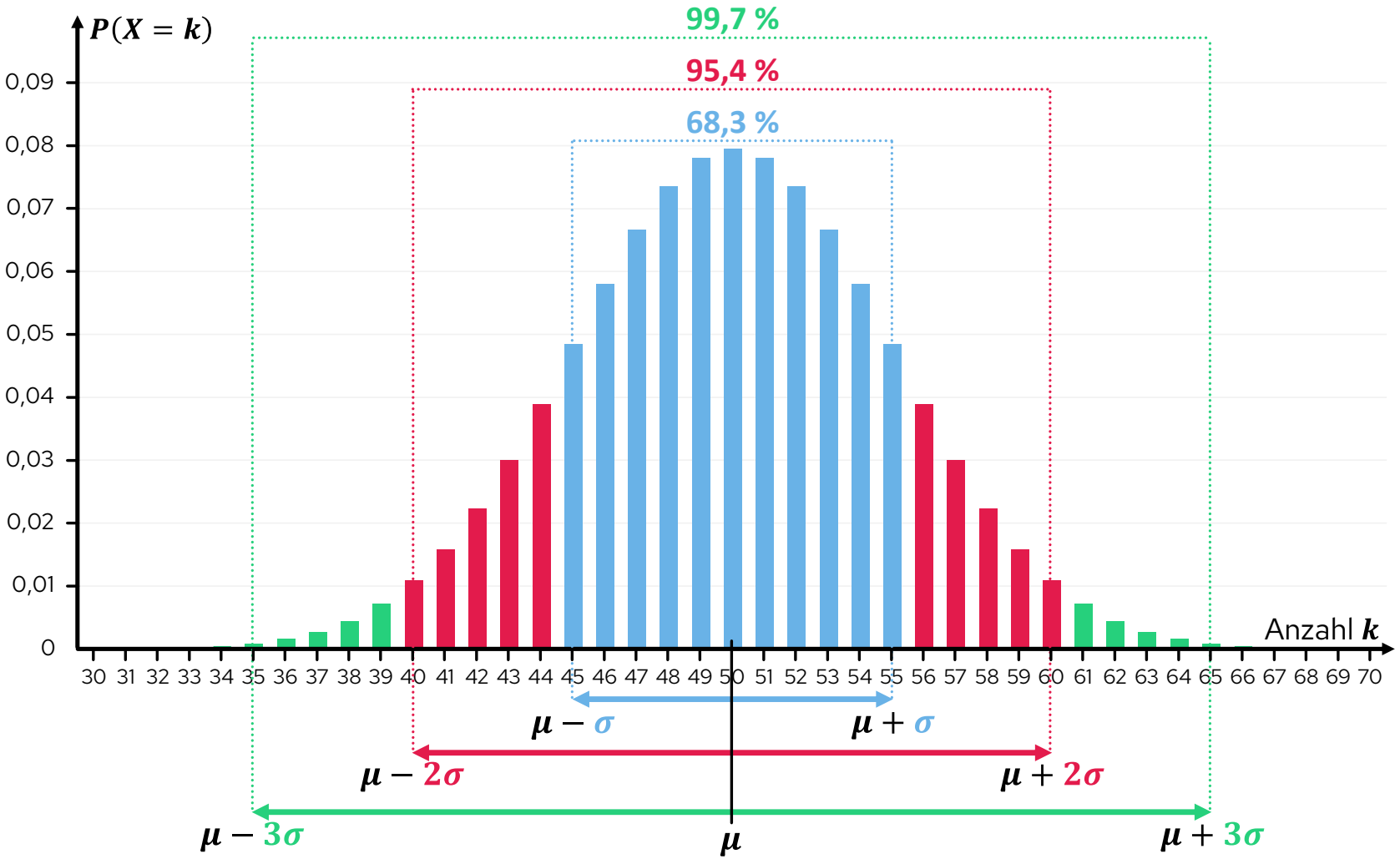
# Hypothesentest: Modellierungsschritte





# Idee: Abgleich der Daten mit der theoretischen Verteilung






## Sigma-Regeln

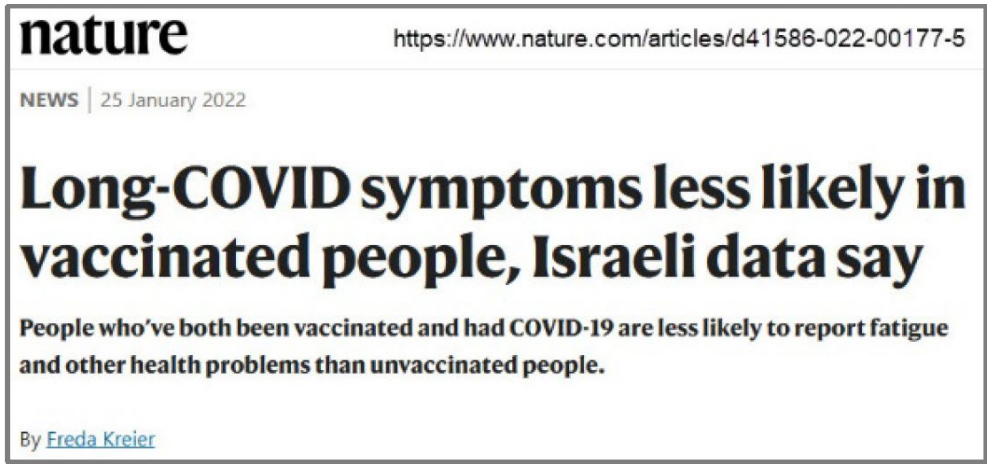
| $c$ | $P(\mu - c \cdot \sigma \leq X \leq \mu + c \cdot \sigma)$ |
|---|---|
| **1** | $\approx 68{,}3\,\%$ |
| 1,64 | $\approx \mathbf{90{,}0}\,\%$ |
| 1,96 | $\approx \mathbf{95{,}0}\,\%$ |
| **2** | $\approx 95{,}4\,\%$ |
| 2,58 | $\approx \mathbf{99{,}0}\,\%$ |
| **3** | $\approx 99{,}7\,\%$ |

- Die Sigma-Regeln gelten für Normalverteilungen.
- Bei Binomialverteilungen gelten sie näherungsweise, für $\sigma = \sqrt{n \cdot p \cdot (1-p)} > 3$.

# Long-COVID-Studie



**Long-COVID-Studie** (Kuodi et al. 2022)
In einer israelischen Studie wurden zwischen März 2020 und November 2021 Probanden, die mit einem PCR-Test positiv auf Corona getestet wurden, befragt, inwiefern bei ihnen nach der Infektion auch Symptome von Long-Covid auftraten.



nature https://www.nature.com/articles/d41586-022-00177-5

NEWS | 25 January 2022

**Long-COVID symptoms less likely in vaccinated people, Israeli data say**

People who've both been vaccinated and had COVID-19 are less likely to report fatigue and other health problems than unvaccinated people.

By Freda Kreier

**Tabelle:** Daten aus der Studie von Kuodi et al. (2022)

| Impfstatus | Probanden | |
|---|---|---|
| | insgesamt | mit Kopfschmerzen |
| geimpft (2 Dosen) | 294 | 41 (14%) |
| geimpft (1 Dosis) | 340 | 80 (24%) |
| ungeimpft | 317 | 69 (22%) |
| **insgesamt** | **951** | **190 (20%)** |

## Daten aus der Tabelle

- In der Gruppe der doppelt-geimpften Personen hatte ein geringerer Anteil Kopfschmerzen.
- Handelt es sich hier um ein signifikantes Ergebnis?
- Wären dieses und ähnliche Ergebnisse nicht auch aufgrund von Zufallsschwankungen zu erwarten, wenn in Wirklichkeit das Risiko für Kopfschmerzen für alle Gruppen gleich ist?

 https://doi.org/10.1038/s41541-022-00526-5

# Hypothesentest: Modellierungsschritte

Didaktik der Mathematik Sekundarstufen

**$H_0$** Die Wahrscheinlichkeit für Kopfschmerzsymptome ist unabhängig vom Impfstatus und liegt bei $p = 20\%$.

**Statistische Nullhypothese formulieren**

mathematisieren

**Datenerhebung:** Long-COVID-Kopfschmerzen ↔ Impfung

**Reales Modell**

**Statistisches Modell**

**Binomialverteilung** mit $n = 294$ (Anzahl der Doppelgeimpften) und $p = 0,2$

Haben geimpfte Personen weniger Long-COVID-Symptome?

**Forschungsfrage / -hypothese festlegen**

vereinfachen
strukturieren
präzisieren

lösen

**Ablehnungsbereich bzw. p-Wert bestimmen**

Prognoseintervall zur Sicherheitswahrscheinlichkeit 95%

**Long-COVID-Symptome** bei geimpften Personen

**Reale Situation**

**Statistische Ergebnisse**

**p-Wert** liefert keine / schwache / mittlere / starke Evidenz gegen die statistische Nullhypothese

interpretieren

**Kritisch reflektieren**

- Was bedeutet das statistische Resultat für die Sachsituation?
- Welchen Beitrag leistet der Signifikanztest zur Klärung der Forschungsfrage?
- Sind weitere Untersuchungen notwendig?



RPTU

# Long-COVID-Studie
## Modellierungsschritte



**Aufgabe 1:** Geben Sie für diese Situation eine passende Wahrscheinlichkeitsverteilung an und bestimmen Sie das zugehörige Prognoseintervall zur Sicherheitswahrscheinlichkeit von 95%.

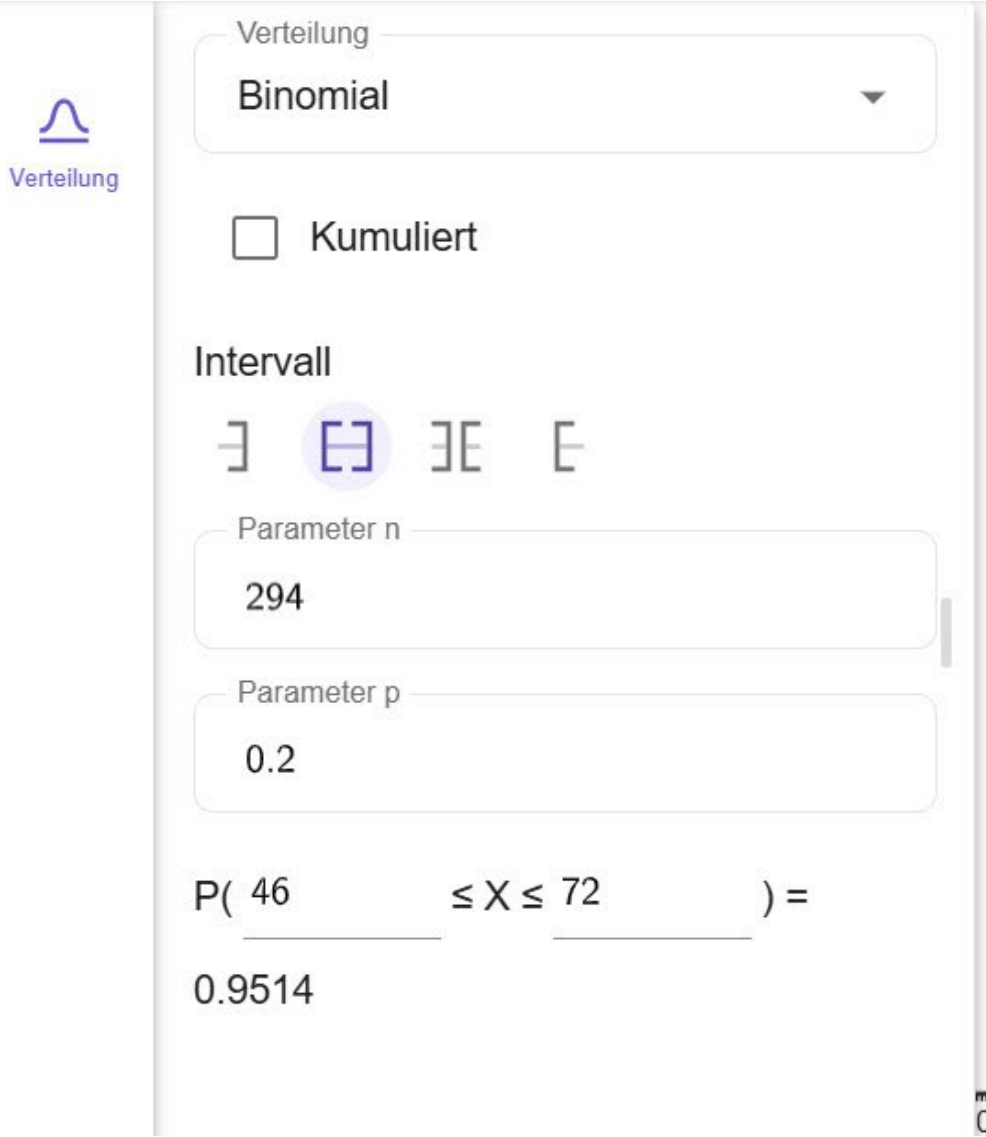
- **Modellannahme:** Kopfschmerzen nach einer Corona-Infektion treten zufällig & unabhängig vom Impfstatus auf.
- **Nullhypothese:** Das Kopfschmerz-Risiko für zweifach-geimpfte Personen ist mit 20% genauso hoch wie für alle.
- **Theoretische Verteilung:** Die Zufallsgröße $X$ der „Kopf-schmerz-Anzahl bei zweifach-geimpften Personen“ ist binomialverteilt mit den Parametern $n = 294$ (Anzahl der untersuchten zweifach-geimpfte Patienten) und $p_0 = 0{,}20$ (Wahrscheinlichkeit für das Auftreten von Kopfschmerzen).
- Es gilt also:
$$P(X = k) = B(294; 0{,}20; k) = \binom{294}{k} \cdot 0{,}20^k \cdot (1 - 0{,}20)^{294-k}$$

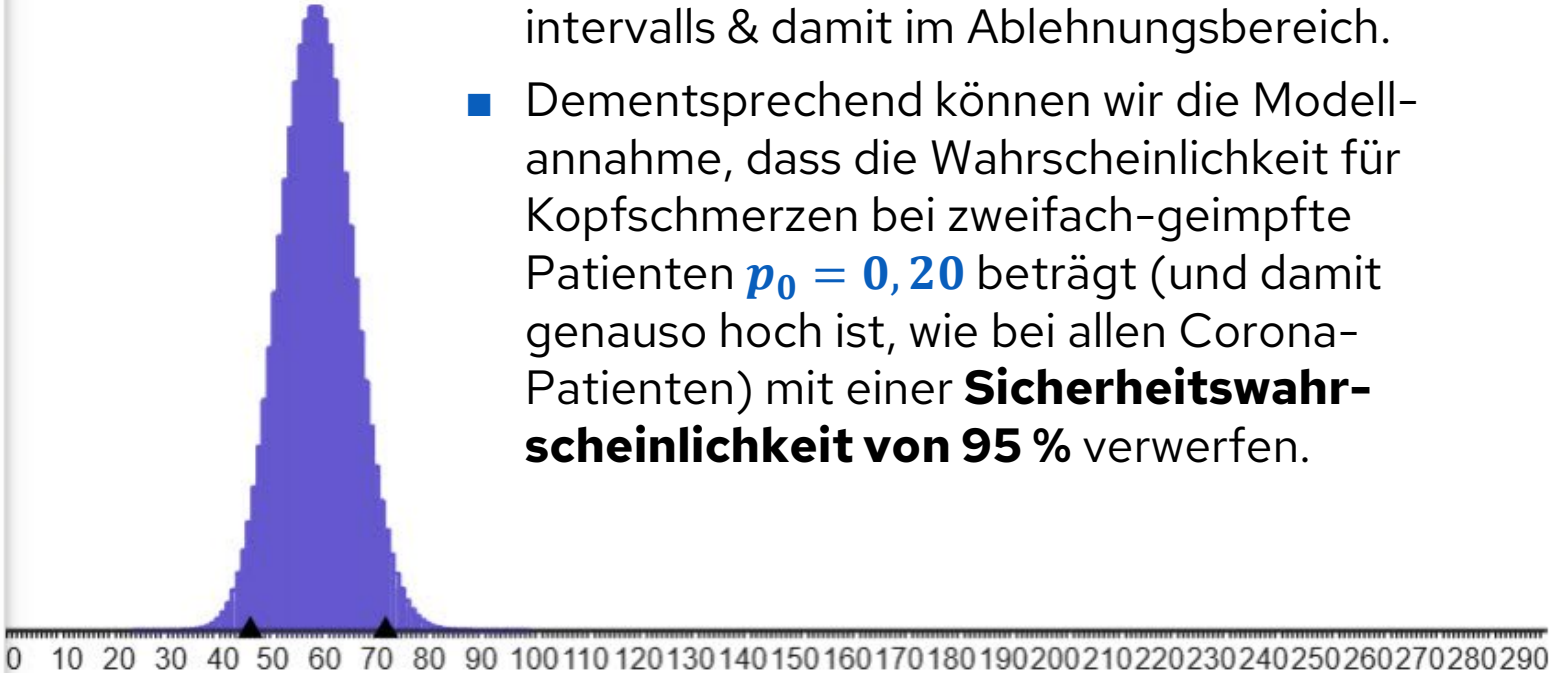
- **Prognoseintervall bestimmen:**
  - Welche Sicherheitswahrscheinlichkeit soll genutzt werden? D.h., ab welcher Abweichung vom Erwartungswert $\boldsymbol{\mu = n \cdot p_0 = 294 \cdot 0{,}20 = 58{,}8}$ möchte man die Modellannahme $p_0 = 0{,}20$ verwerfen?
  - Hier ist die Sicherheitswahrscheinlichkeit 95% vorgegeben.
  - Dies bedeutet, dass man bei großen Abweichungen der Testgröße $X$ (Anzahl der berichteten Kopfschmerzen) vom Erwartungswert, die nur mit einer Wahrscheinlichkeit von maximal $\boldsymbol{\alpha = 1 - 0{,}95 = 0{,}05 = 5\%}$ (Signifikanzniveau) auftreten, die **Modellannahme** verwerfen wird.
  - Das lässt sich mit dem Wahrscheinlich-keitsrechner von GeoGebra bestimmen.

# Long-COVID-Studie
## Modellierungsschritte

**Tabelle:** Daten aus der Studie von Kuodi et al. (2022)

| Impfstatus | Probanden | |
|---|---|---|
| | insgesamt | mit Kopfschmerzen |
| geimpft (2 Dosen) | 294 | 41 (14%) |
| geimpft (1 Dosis) | 340 | 80 (24%) |
| ungeimpft | 317 | 69 (22%) |
| **insgesamt** | **951** | **190 (20%)** |




### Wo liegt der Wert der Testgröße $X$?

- Der Testgröße $X$, die Anzahl berichteter Kopfschmerzen, hat hier den Wert **41**.
- Dieser liegt außerhalb des Prognose-intervalls & damit im Ablehnungsbereich.
- Dementsprechend können wir die Modell-annahme, dass die Wahrscheinlichkeit für Kopfschmerzen bei zweifach-geimpfte Patienten $p_0 = 0,20$ beträgt (und damit genauso hoch ist, wie bei allen Corona-Patienten) mit einer **Sicherheitswahr-scheinlichkeit von 95 %** verwerfen.

# Long-COVID-Studie

## Modellierungsschritte

**Tabelle:** Daten aus der Studie von Kuodi et al. (2022)

| Impfstatus | Probanden | |
|---|---|---|
| | insgesamt | mit Kopfschmerzen |
| geimpft (2 Dosen) | 294 | 41 (14%) |
| geimpft (1 Dosis) | 340 | 80 (24%) |
| ungeimpft | 317 | 69 (22%) |
| **insgesamt** | **951** | **190 (20%)** |

Didaktik der Mathematik Sekundarstufen

Verteilung

Verteilung

Binomial

Kumuliert

Intervall

Parameter n

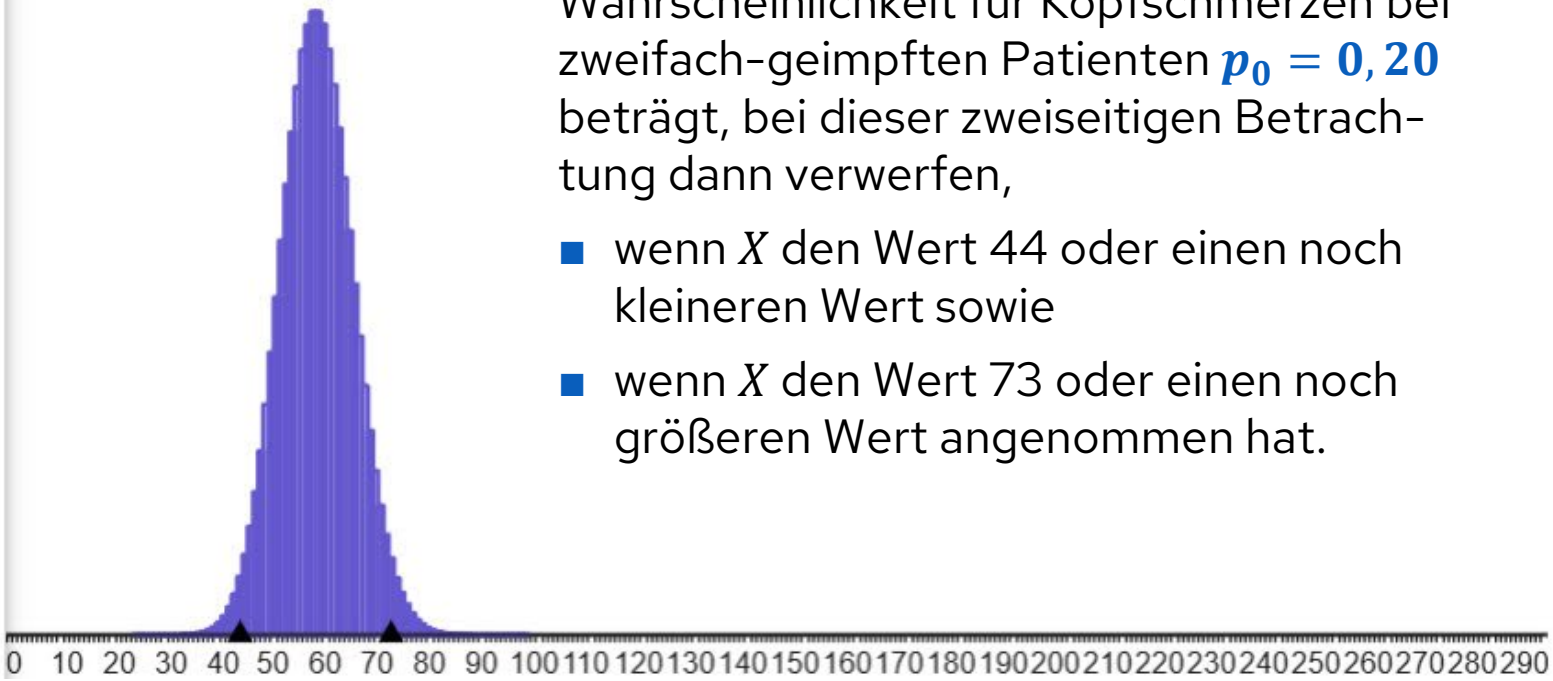
294

Parameter p

0.2

P(X ≤ 44 ) + P(X ≥ 73 ) =

0.0161 + 0.0251 = 0.0412




### Bemerkung

Wir können die Modellannahme, dass die Wahrscheinlichkeit für Kopfschmerzen bei zweifach-geimpften Patienten $p_0 = 0,20$ beträgt, bei dieser zweiseitigen Betrachtung dann verwerfen,

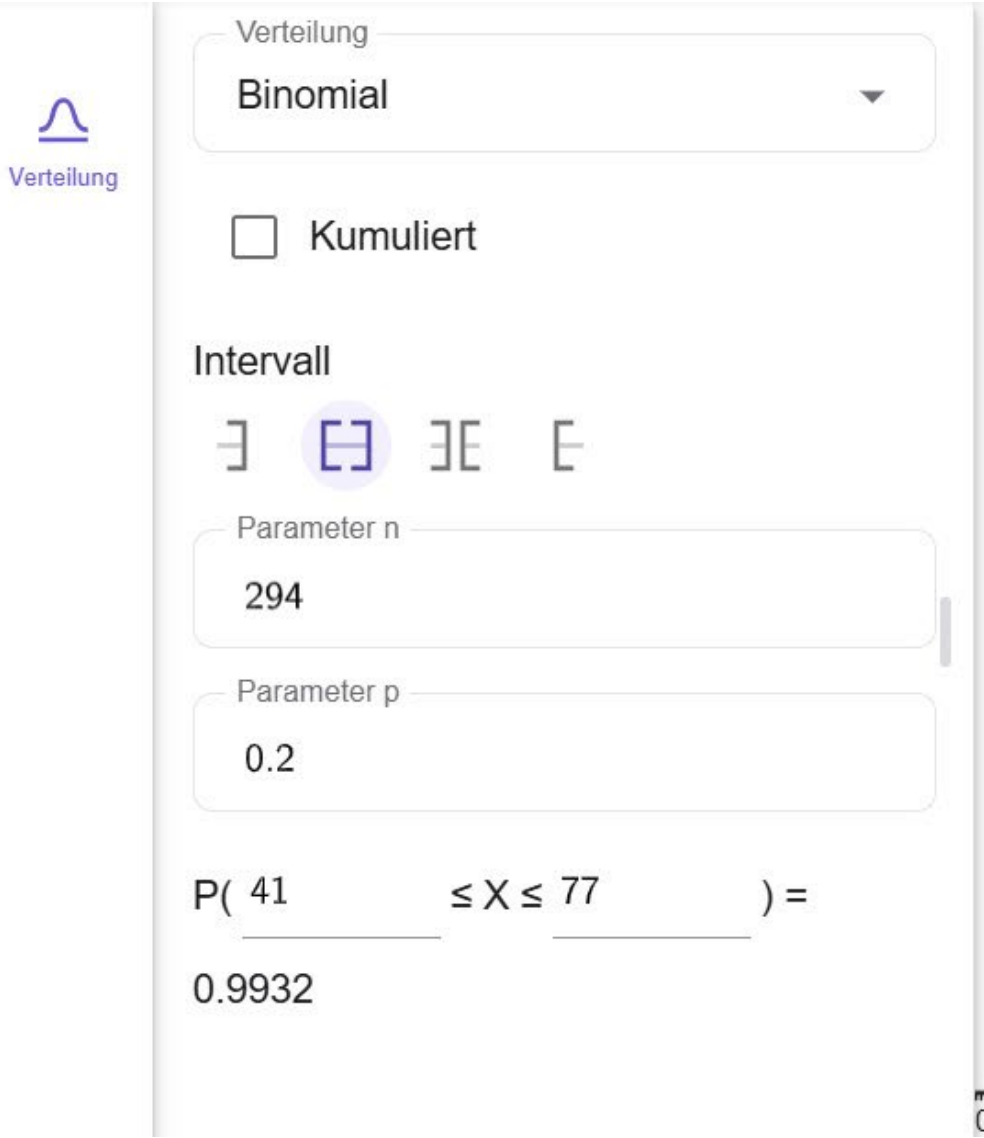
- wenn $X$ den Wert 44 oder einen noch kleineren Wert sowie
- wenn $X$ den Wert 73 oder einen noch größeren Wert angenommen hat.

# Long-COVID-Studie
## Modellierungsschritte

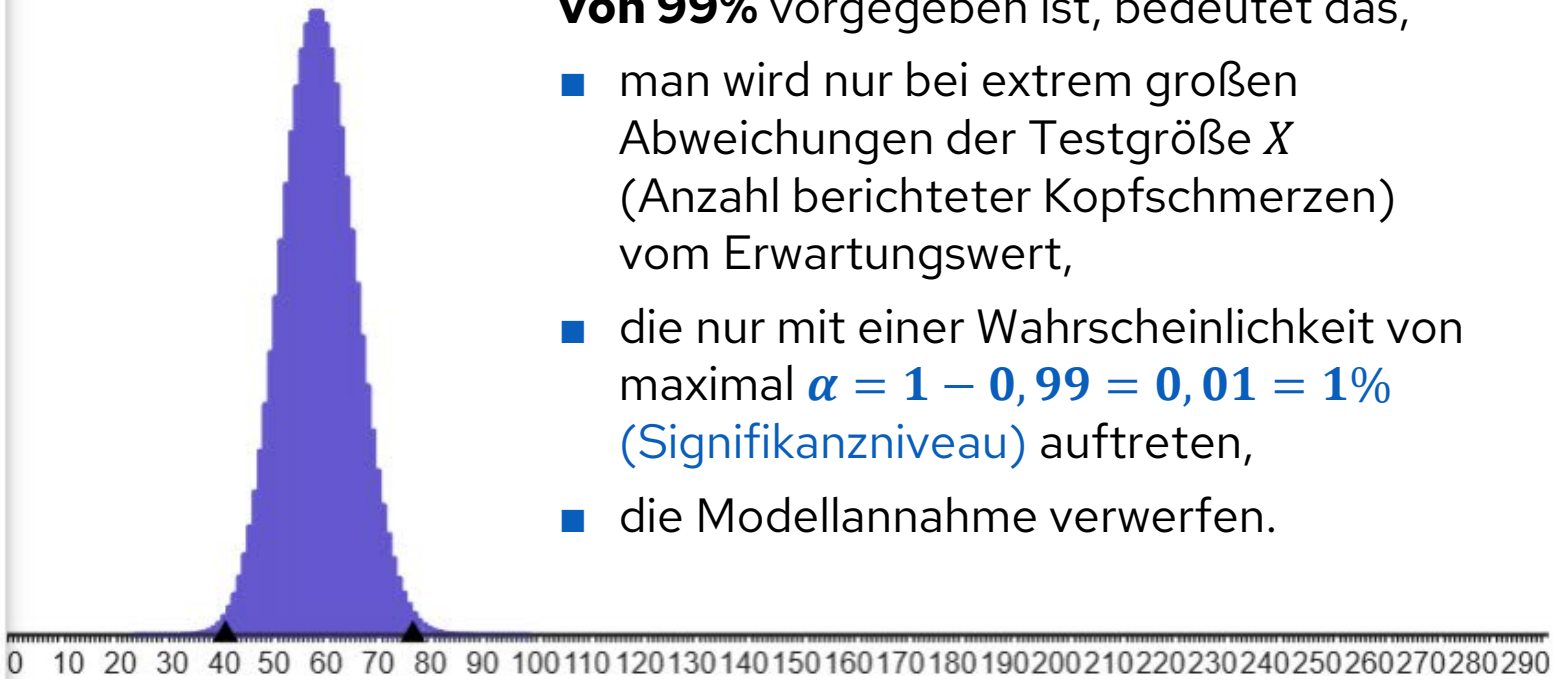
**Tabelle:** Daten aus der Studie von Kuodi et al. (2022)

| Impfstatus | Probanden | |
|---|---|---|
| | insgesamt | mit Kopfschmerzen |
| geimpft (2 Dosen) | 294 | 41 (14%) |
| geimpft (1 Dosis) | 340 | 80 (24%) |
| ungeimpft | 317 | 69 (22%) |
| **insgesamt** | **951** | **190 (20%)** |






**Aufgabe 2:** Was hätte sich verändert, wenn das Prognoseintervall zur Sicherheitswahrscheinlichkeit 99% bestimmt worden wäre?




Da eine **Sicherheitswahrscheinlichkeit von 99%** vorgegeben ist, bedeutet das,

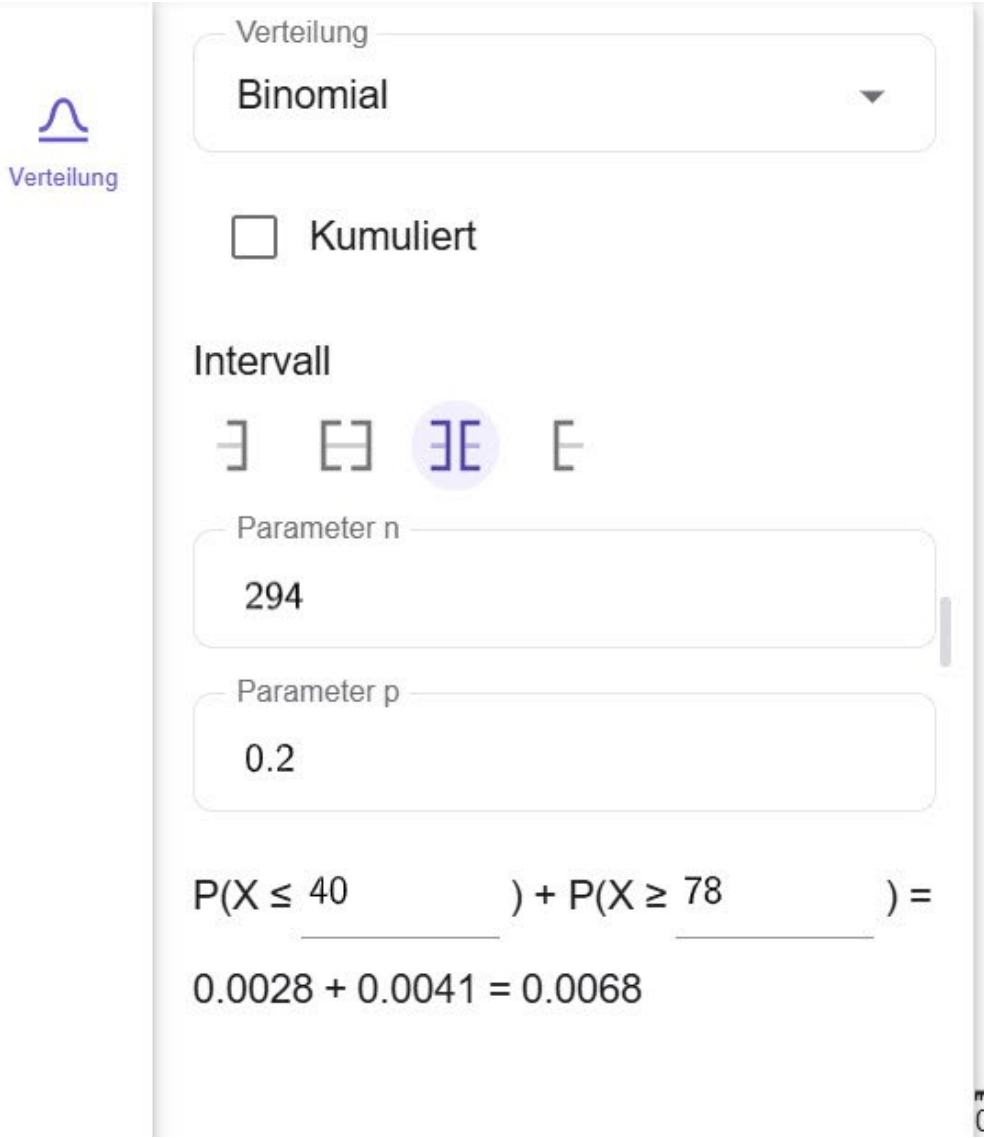
- man wird nur bei extrem großen Abweichungen der Testgröße $X$ (Anzahl berichteter Kopfschmerzen) vom Erwartungswert,
- die nur mit einer Wahrscheinlichkeit von maximal $\alpha = 1 - 0,99 = 0,01 = 1\%$ (Signifikanzniveau) auftreten,
- die Modellannahme verwerfen.

# Long-COVID-Studie
## Modellierungsschritte

**Tabelle:** Daten aus der Studie von Kuodi et al. (2022)

| Impfstatus | Probanden | |
|---|---|---|
| | insgesamt | mit Kopfschmerzen |
| geimpft (2 Dosen) | 294 | 41 (14%) |
| geimpft (1 Dosis) | 340 | 80 (24%) |
| ungeimpft | 317 | 69 (22%) |
| **insgesamt** | **951** | **190 (20%)** |

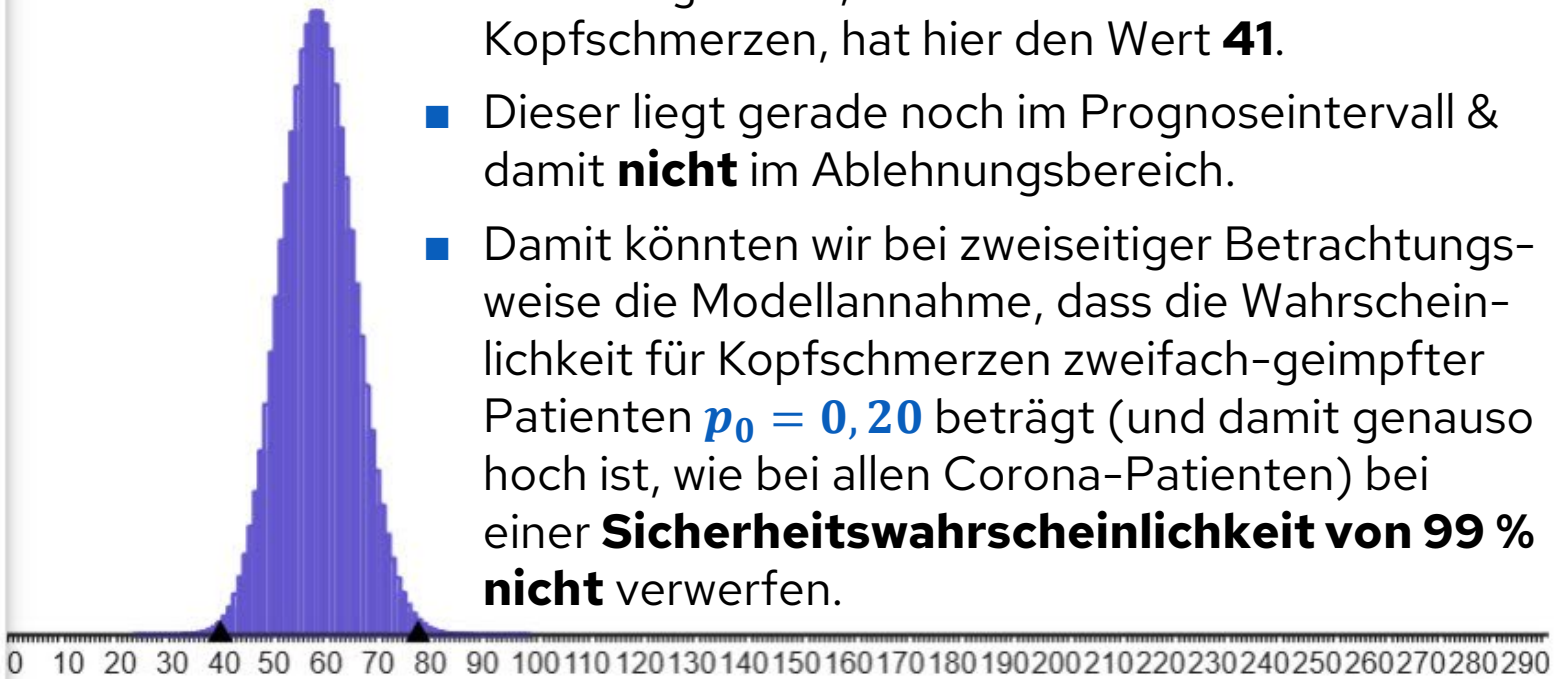
Didaktik der Mathematik Sekundarstufen




**Aufgabe 2:** Was hätte sich verändert, wenn das Prognoseintervall zur Sicherheitswahrscheinlichkeit 99% bestimmt worden wäre?

### Wo liegt der Wert der Testgröße $X$?

- Der Testgröße $X$, die Anzahl berichteter Kopfschmerzen, hat hier den Wert **41**.
- Dieser liegt gerade noch im Prognoseintervall & damit **nicht** im Ablehnungsbereich.
- Damit könnten wir bei zweiseitiger Betrachtungsweise die Modellannahme, dass die Wahrscheinlichkeit für Kopfschmerzen zweifach-geimpfter Patienten $p_0 = 0,20$ beträgt (und damit genauso hoch ist, wie bei allen Corona-Patienten) bei einer **Sicherheitswahrscheinlichkeit von 99 %** **nicht** verwerfen.



# Long-COVID-Studie
## Modellierungsschritte



**Aufgabe 3:** Formulieren Sie vor dem Hintergrund Ihrer Analyse des Prognoseintervalls für diese Situation ein passendes Modell sowie eine geeignete Null- und Alternativhypothese.

- **Modellannahme:** Aufgrund der Daten besteht die gerichtete Vermutung, dass zweifach Geimpfte ein Kopfschmerz-Risiko von weniger als 20% haben.
- **Nullhypothese:** Das Kopfschmerz-Risiko für zweifach-geimpfte Personen beträgt 20%.
- **Alternativhypothese:** Das Kopfschmerz-Risiko für zweifach-geimpfte Personen ist kleiner als 20%.
- **Theoretische Verteilung:** Die Zufallsgröße (und Testgröße) $X$ der „Kopfschmerz-Anzahl bei zweifach-geimpften Personen" ist binomialverteilt mit den Parametern $n = 294$ und $p_0 = 0{,}20$. Es gilt also:

$$P(X = k) = B(294; 0{,}20; k) = \binom{294}{k} \cdot 0{,}20^k \cdot (1 - 0{,}20)^{294-k}$$

**Aufgabe 4:** Nutzen Sie den Wahrscheinlichkeitsrechner in GeoGebra zum Testen Ihrer Nullhypothese mittels Konstruktion einer Entscheidungsregel und des zugehörigen Ablehnungsbereichs zu einem Signifikanzniveau von 1 %.

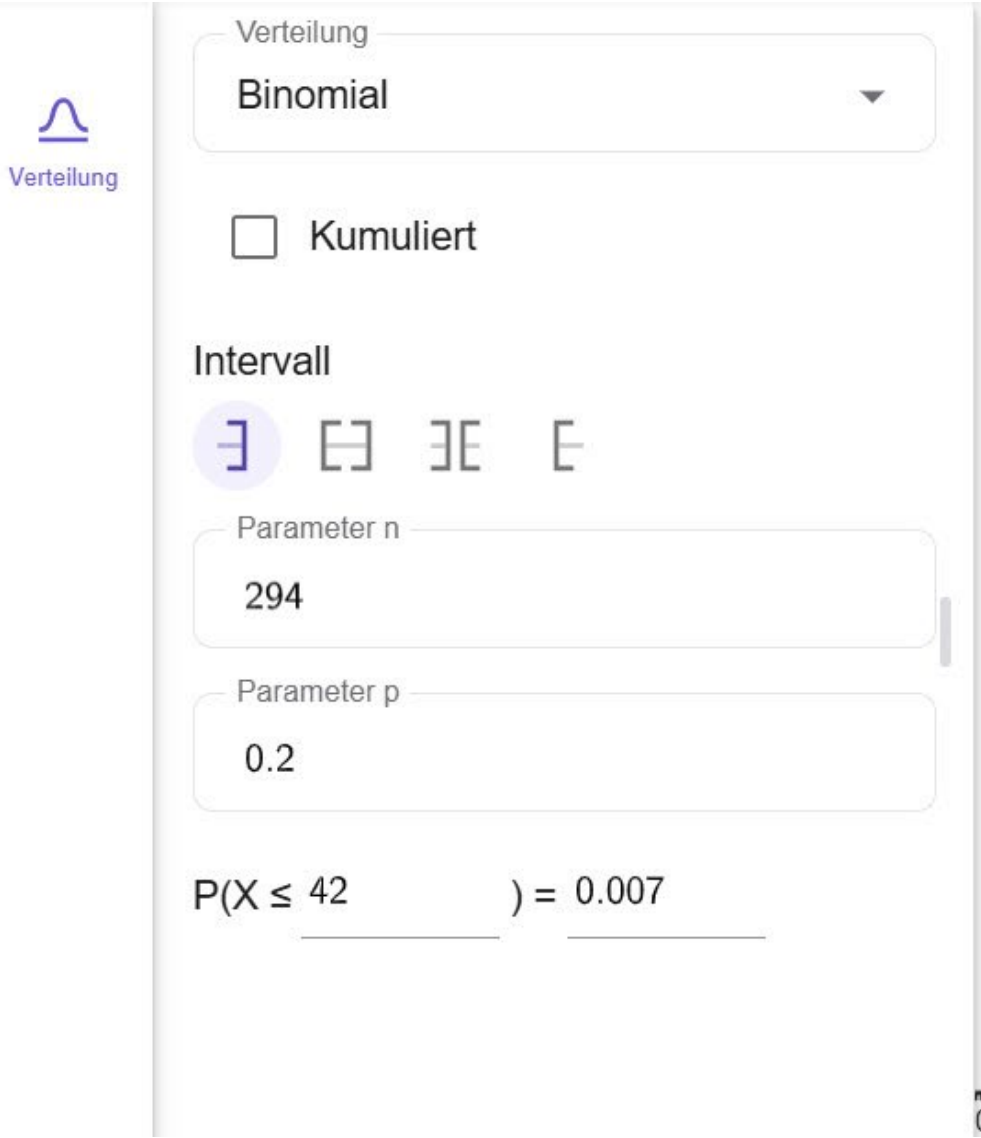
Prüfen Sie mit dieser das beobachtete Ergebnis.

- **Entscheidungsregel:** Liegt das Stichprobenergebnis von 41 berichteten Kopfschmerzen im Ablehnungsbereich $\{\mathbf{0}, \ldots, \boldsymbol{a}\}$, der eine Eintrittswahrscheinlichkeit von $\boldsymbol{P(X \leq a) \leq 0.01}$ haben sollte, so ist die Nullhypothese anzuzweifeln.
- **Bestimmung des maximalen Ablehnungsbereichs mit GeoGebra:** Mithilfe von GeoGebra erhält man $P(X \leq 43) = 0.0108$ und $P(X \leq 42) \leq 0.007$.

  $\Rightarrow$ Ablehnungsbereich: $\{0, \ldots, 42\}$

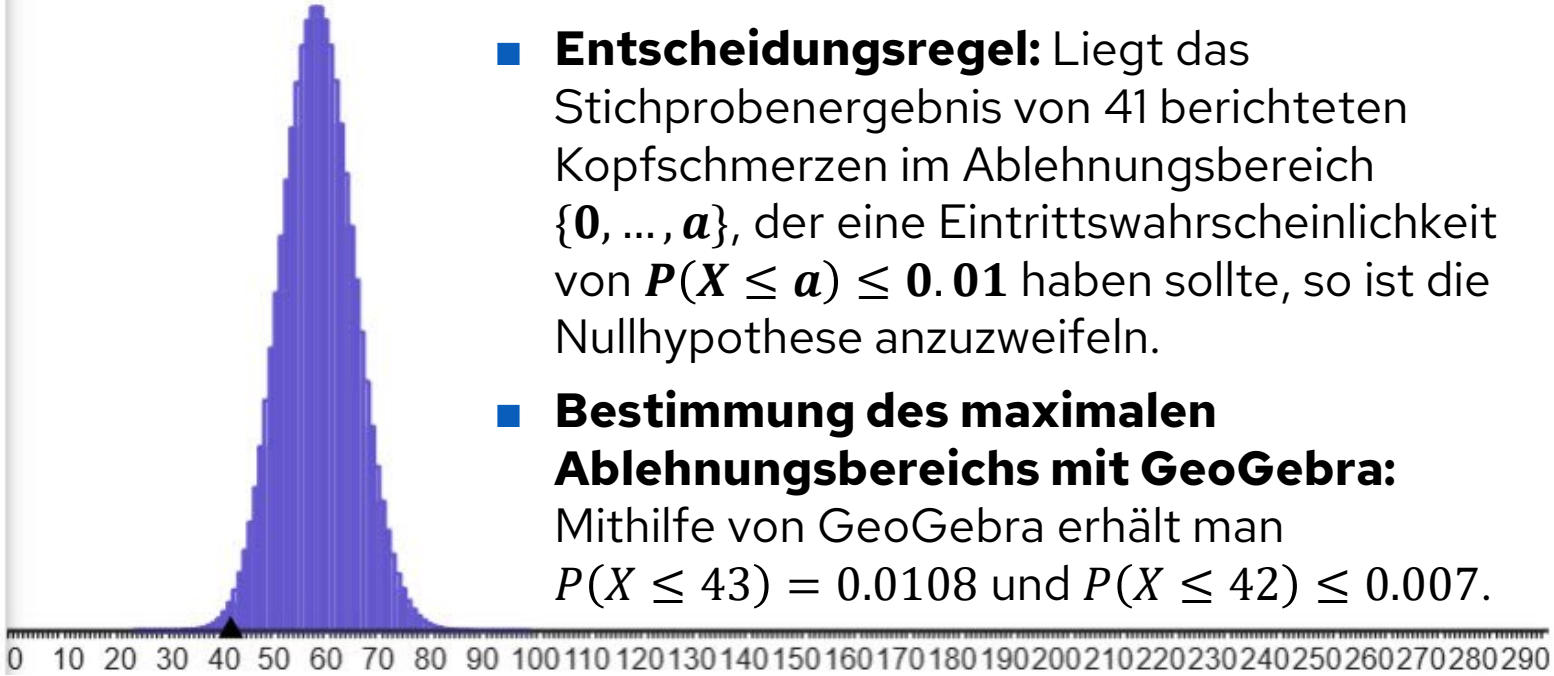
# Long-COVID-Studie
## Modellierungsschritte






**Aufgabe 4:** Nutzen Sie den Wahrscheinlichkeitsrechner in GeoGebra zum Testen Ihrer Nullhypothese mittels Konstruktion einer Entscheidungsregel und des zugehörigen Ablehnungsbereichs zu einem Signifikanzniveau von 1 %.

Prüfen Sie mit dieser das beobachtete Ergebnis.




- **Entscheidungsregel:** Liegt das Stichprobenergebnis von 41 berichteten Kopfschmerzen im Ablehnungsbereich $\{0, \ldots, a\}$, der eine Eintrittswahrscheinlichkeit von $P(X \leq a) \leq 0.01$ haben sollte, so ist die Nullhypothese anzuzweifeln.
- **Bestimmung des maximalen Ablehnungsbereichs mit GeoGebra:** Mithilfe von GeoGebra erhält man $P(X \leq 43) = 0.0108$ und $P(X \leq 42) \leq 0.007$.

$\Rightarrow$ Ablehnungsbereich: $\{0, \ldots, 42\}$

# Long-COVID-Studie
## Modellierungsschritte



### Beobachtung und Schlussfolgerung

- Da 41 im Ablehnungsbereich liegt, ist die Nullhypothese, ein Kopfschmerz-Risiko von 20% bei Zweifachgeimpften, anzuzweifeln.
- Dies ist kohärent zur Vermutung, dass das Risiko geringer als 20% sein könnte.
- Wegen der Signifikanz des Ergebnisses (sowie der beobachteten Ergebnisse & Hintergrundtheorien zum Einfluss von Impfungen) gehen wir daher von einem geringeren Risiko als 20% aus.
- Dies tun wir, obwohl wir wissen, dass auch ein Fehler 1. Art vorliegen könnte, dass wir die Nullhypothese verwerfen, obwohl sie in Wirklichkeit richtig ist.
- Die Wahrscheinlichkeit für das Auftreten dieses Fehlers beträgt aber nur 1%, entspricht also dem gewählten Signifikanzniveau $\alpha = 1\%$.

**Aufgabe 4:** Nutzen Sie den Wahrscheinlichkeitsrechner in GeoGebra zum Testen Ihrer Nullhypothese mittels Konstruktion einer Entscheidungsregel und des zugehörigen Ablehnungsbereichs zu einem Signifikanzniveau von 1 %.

Prüfen Sie mit dieser das beobachtete Ergebnis.

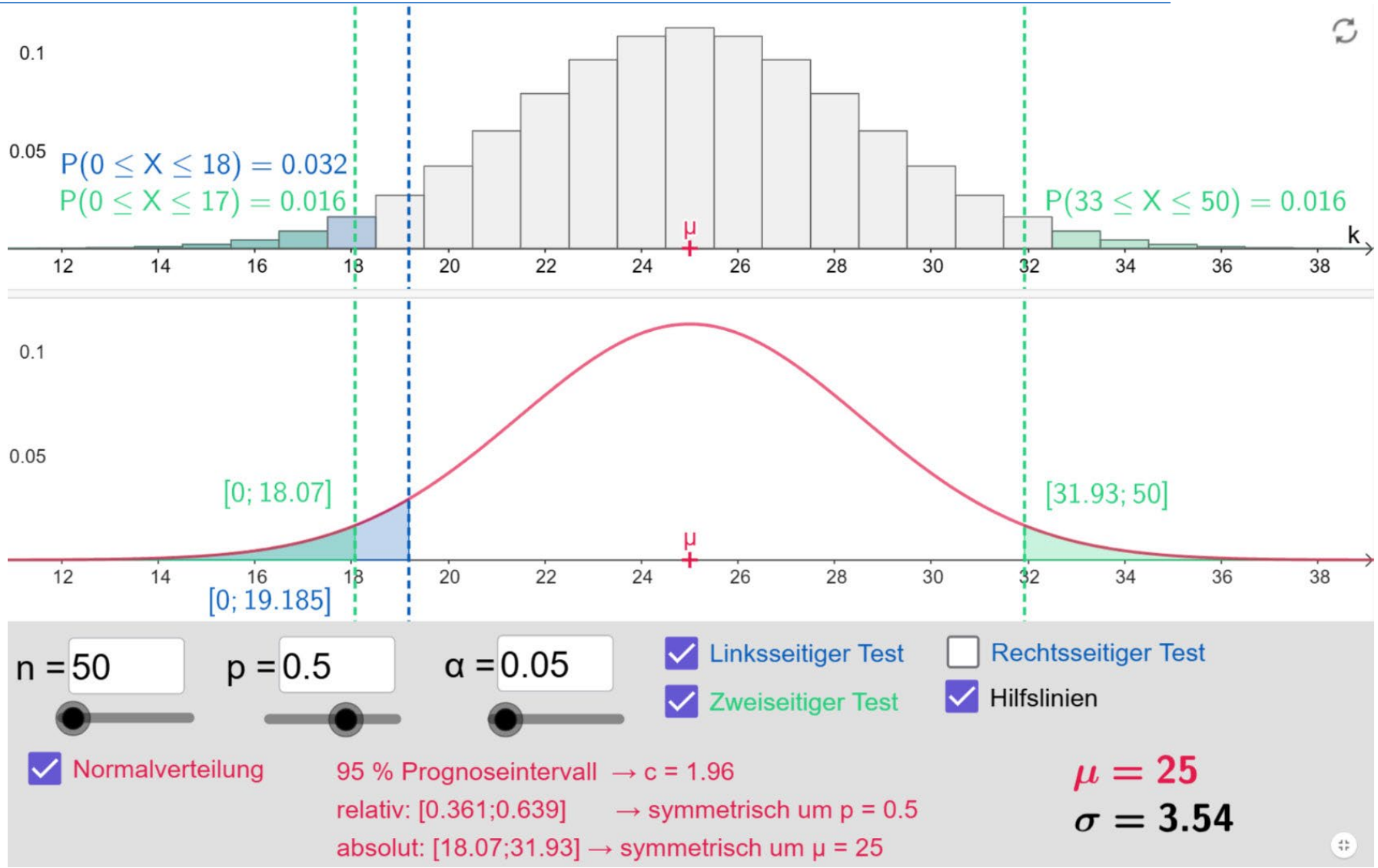
### Bemerkung

Ein Vergleich mit dem Ergebnis aus **Aufgabe 2** zeigt, dass sich der Ablehnungsbereich (hier 1%) bei einseitiger Betrachtungsweise, nicht auf zwei Seiten „verteilt“ sondern nur auf eine Seite der Verteilung liegt und damit die Grenze auf dieser Seite näher an den Erwartungswert heranrückt.

Vgl. dazu auch das Applet M4.V.15b App Signifikanztest & Prognoseintervall.

# Signifikanztest: Einseitig ↔ Zweiseitig



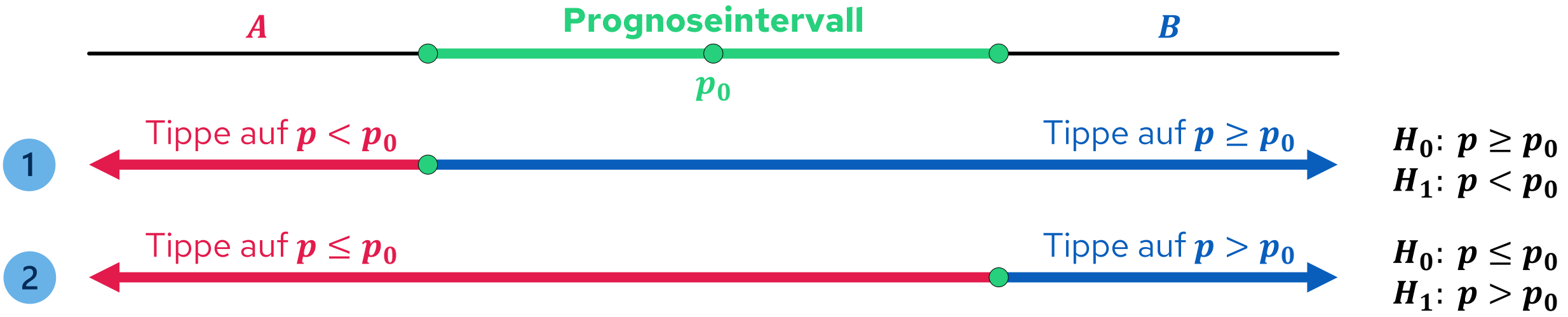



https://www.geogebra.org/m/wbmhsnba#material/fscea3s8

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

# Einseitige Hypothesentests:
## Welche Seite soll es sein?

Didaktik der Mathematik Sekundarstufen




**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:
## 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.



### 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

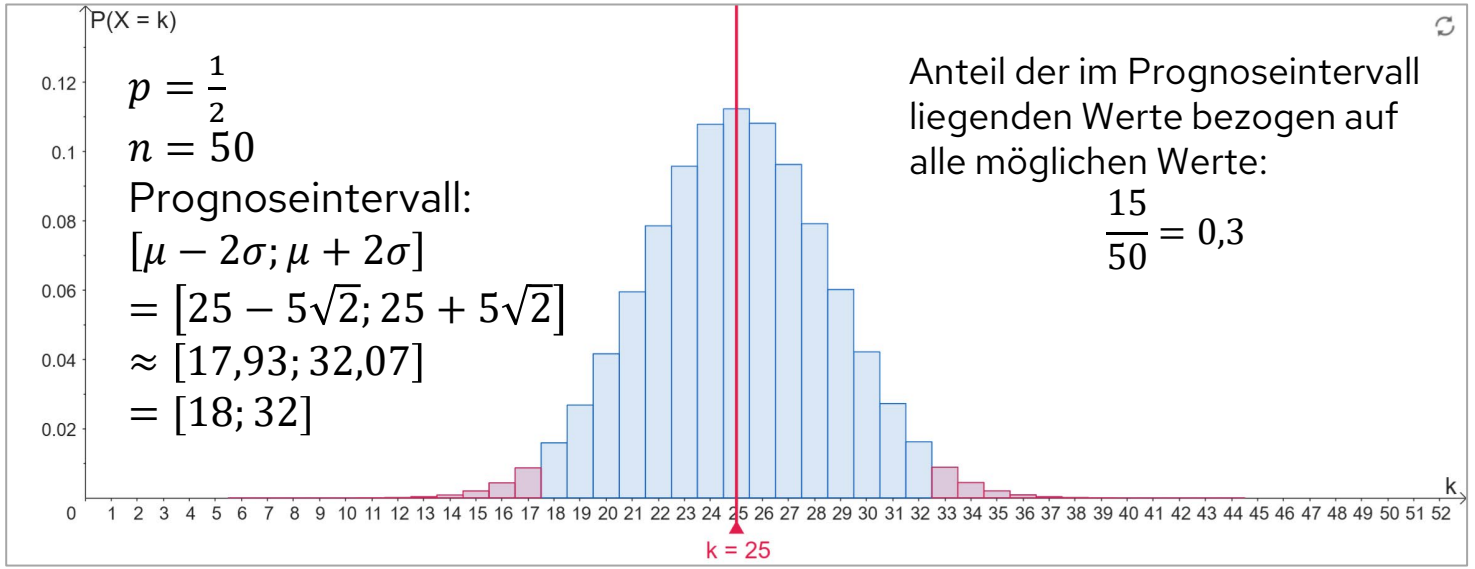
## Der Einfluss des Stichprobenumfangs $n$

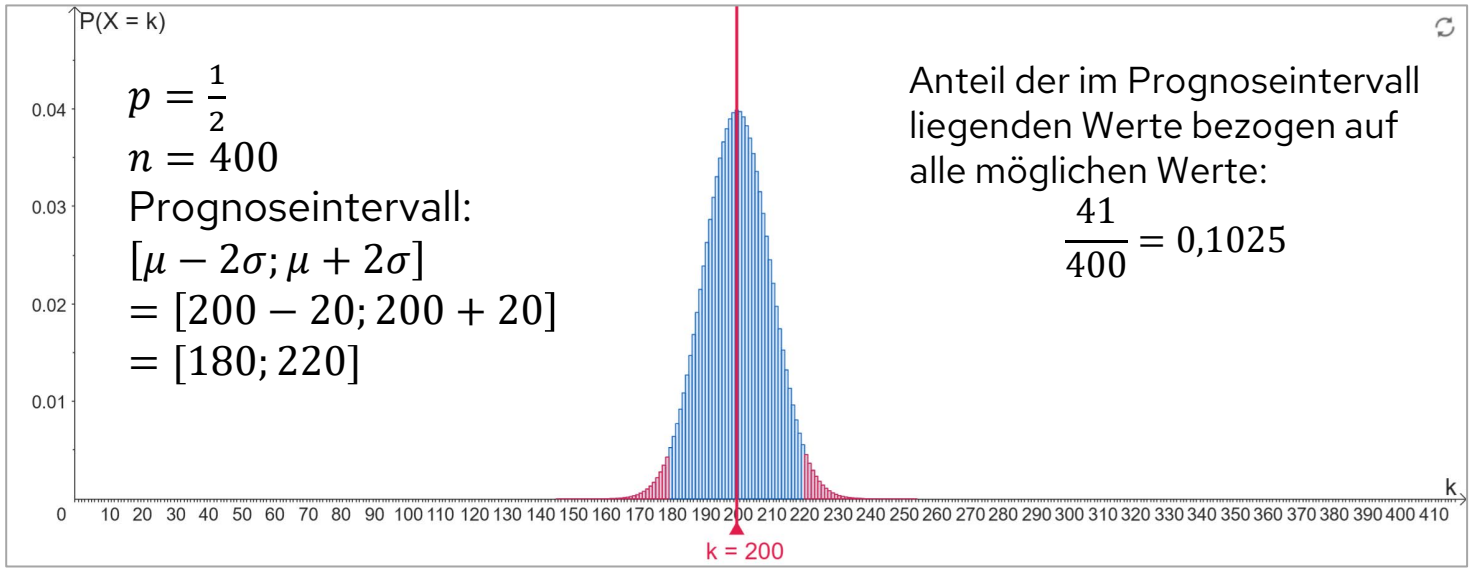


**Bemerkungen:** Bei einseitigen Tests ist zur Entschlüsselung von Standpunkten und Interessenlagen der zweiseitige Blick sehr hilfreich.

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 etwas 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

# Fehler 1. Art und Fehler 2. Art

Didaktik der Mathematik Sekundarstufen

| | | Realität | |
|---|---|---|---|
| | | $H_0$ richtig | $H_0$ falsch |
| Test | $H_0$ richtig | ✓ | ✗ **Fehler 2. Art** ($\beta$-Fehler) |
| | $H_0$ falsch | ✗ **Fehler 1. Art** ($\alpha$-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



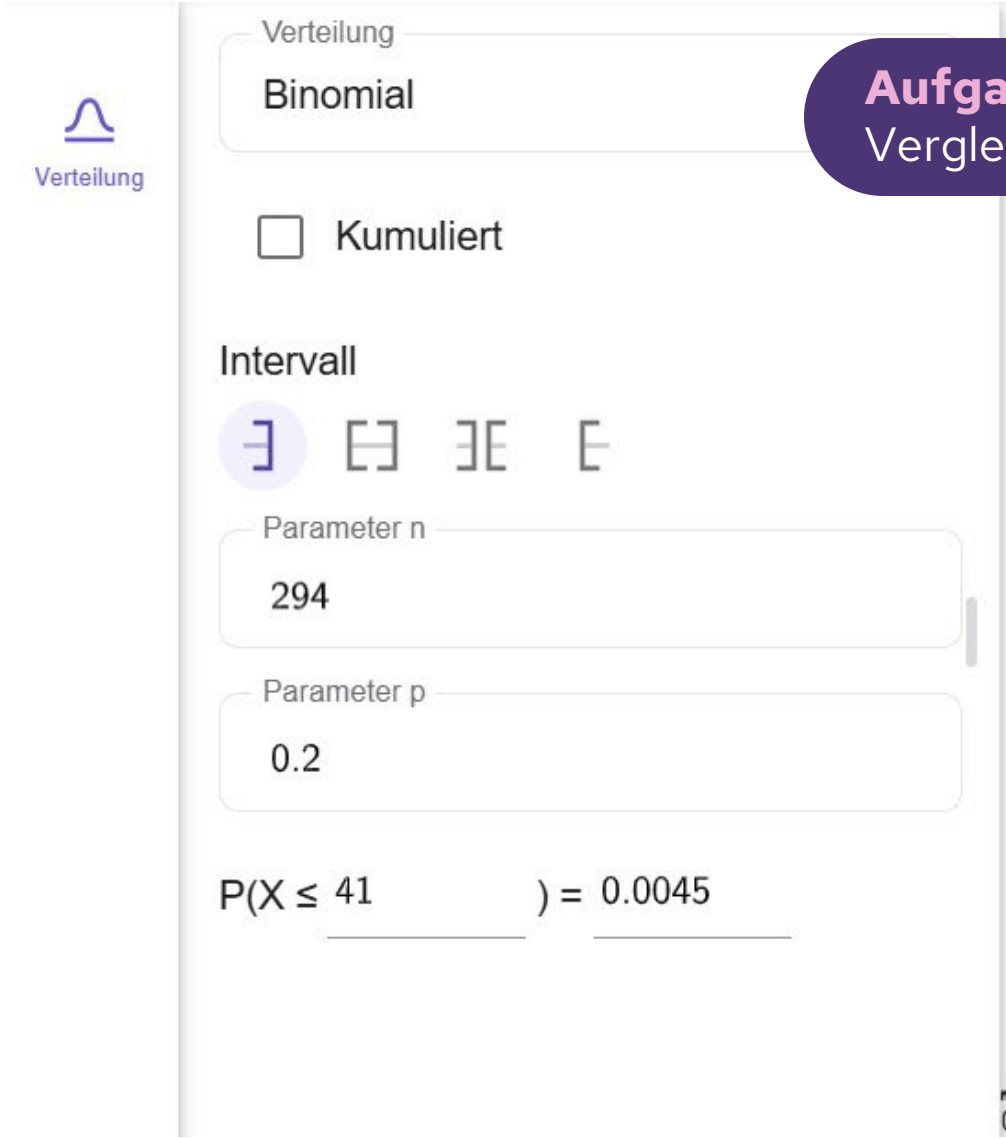
RPTU

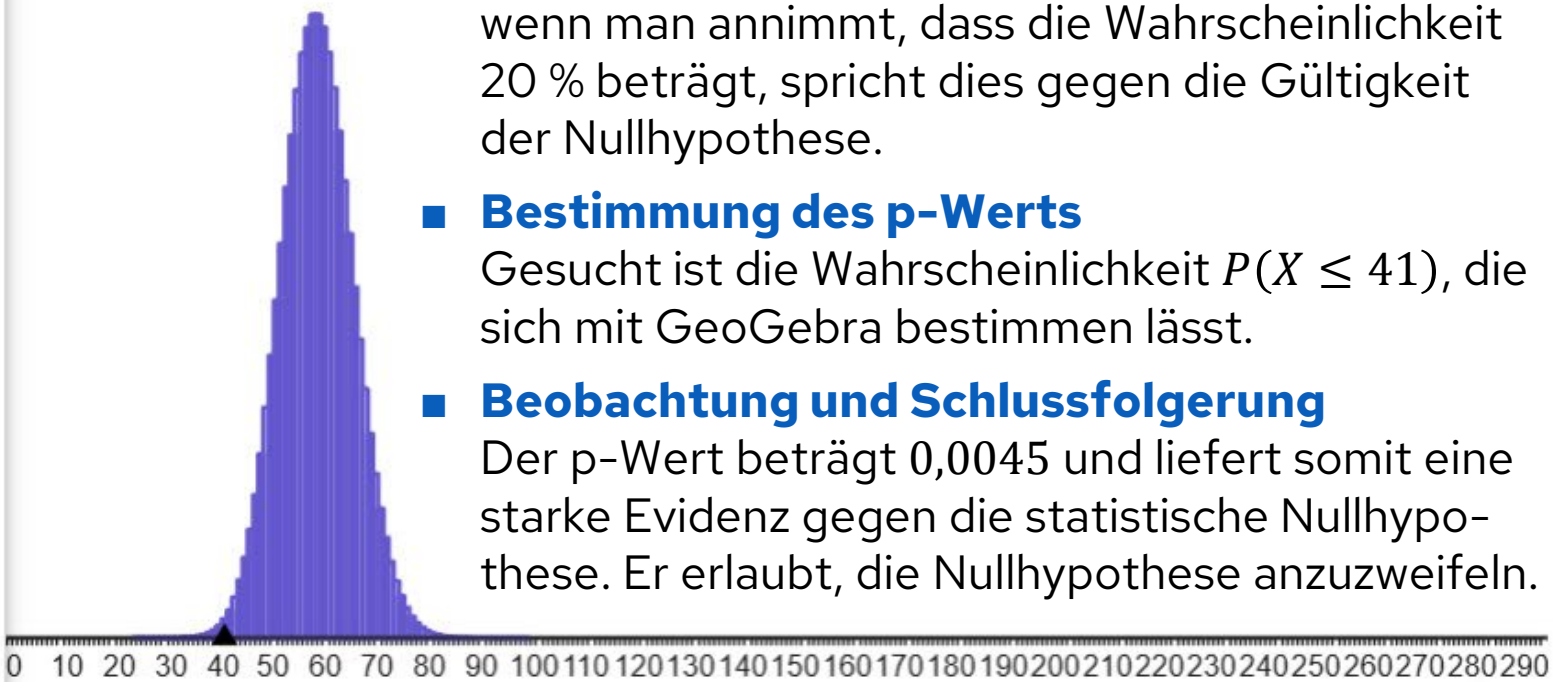
# Long-COVID-Studie
## Modellierungsschritte

| Impfstatus | Probanden | |
|---|---|---|
| | insgesamt | mit Kopfschmerzen |
| geimpft (2 Dosen) | 294 | 41 (14%) |
| geimpft (1 Dosis) | 340 | 80 (24%) |
| ungeimpft | 317 | 69 (22%) |
| **insgesamt** | **951** | **190 (20%)** |

Didaktik der Mathematik Sekundarstufen

**Aufgabe 5:** Wiederholen Sie den Hypothesentest mit der p-Wert-Methode. Vergleichen Sie das Vorgehen und ihr Ergebnis mit dem aus **Aufgabe 4**.




- **Entscheidungsregel**
  Kann man mit dem Signifikanztest zeigen, dass die aufgetretenen und extremere Daten mit einer geringen Wahrscheinlichkeit auftreten würden, wenn man annimmt, dass die Wahrscheinlichkeit 20 % beträgt, spricht dies gegen die Gültigkeit der Nullhypothese.
- **Bestimmung des p-Werts**
  Gesucht ist die Wahrscheinlichkeit $P(X \leq 41)$, die sich mit GeoGebra bestimmen lässt.
- **Beobachtung und Schlussfolgerung**
  Der p-Wert beträgt $0{,}0045$ und liefert somit eine starke Evidenz gegen die statistische Nullhypothese. Er erlaubt, die Nullhypothese anzuzweifeln.

# Hypothesentest: p-Wert-Methode



## Übersicht über die p-Wert-Methode nach Fisher

Man bestimmt die kumulierte Wahrscheinlichkeit für den aufgetretenen Wert und extremere Werte, unter der Annahme, dass die Nullhypothese gilt, den sogenannten **p-Wert**.

**(1)** Fischers Idee war es den exakten p-Wert zu veröffentlichen.

**(2)** Alternativ ist der Vergleich des p-Werts mit einem Signifikanzniveau möglich:

→ Ist der p-Wert kleiner als das Signifikanzniveau $\alpha$, so spricht dies gegen die Gültigkeit der Nullhypothese.

| Größe des p-Werts | Sprechweise | |
|---|---|---|
| p-Wert $\leq 1\%$ | statistisch hochsignifikante Abweichung | starke Evidenz gegen die statistische Nullhypothese |
| $1\% <$ p-Wert $\leq 5\%$ | statistisch signifikante Abweichung | mittlere Evidenz gegen die statistische Nullhypothese |
| $5\% <$ p-Wert $\leq 10\%$ | | schwache Evidenz gegen die statistische Nullhypothese |
| $10\% <$ p-Wert | keine statistisch signifikante Abweichung | keine Evidenz gegen die statistische Nullhypothese |

# Was bedeutet „signifikant“?



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

## Was bedeutet ein signifikantes Testergebnis?

- Der Begriff „Signifikanz“ bezieht sich auf „eine Welt“, in der die Nullhypothese $H_0$ gültig ist: Man betrachtet die Wahrscheinlichkeit für das Eintreten des Stichprobenergebnisses oder noch extremerer Ergebnisse unter der Annahme, die Nullhypothese $H_0$ sei richtig.
- Diese Wahrscheinlichkeit wird **$p$-Wert** genannt.
- **Ein Testergebnis ist auf dem $1\%$-Niveau signifikant, wenn der** zum Testergebnis gehörige **$p$-Wert kleiner oder gleich $1\%$ ist** $(p \leq 0,01)$.

## Was bedeutet ein signifikantes Testergebnis nicht?

- Der Signifikanztest ist kein Beweis für die Gültigkeit der statistischen Nullhypothese, wenn diese nicht verworfen werden kann.
- Der Signifikanztest ist kein Beweis für die Alternativhypothese, wenn die Nullhypothese verworfen werden kann. Auch wenn ein signifikantes Ergebnis dafür spricht, dass die statistische Nullhypothese falsch ist, kann sie dennoch richtig sein (Fehler 1. Art).
- Der Signifikanztest liefert keine Aussage über die Wahrscheinlichkeit für die Gültigkeit der Null- oder der Alternativhypothese.



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

# Kontakt




**Prof. Dr. Jürgen Roth**

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

j.roth@rptu.de

juergen-roth.de
dms.math.rptu.de
mategnu.de


RPTU