

# Mehr als ein Testergebnis: Konfidenzintervalle im Unterricht erlebbar machen

Von der Ja/Nein-Entscheidung des Hypothesentests zur Intervallschätzung – praxisnah mit Zollstock, Ellipse, GeoGebra und Python

---

Reimund Vehling

20.02.2026

# Leitgedanken

---

## Zentrale Leitfrage

Was bedeutet es, wenn wir aus einer Stichprobe auf die Grundgesamtheit schließen?

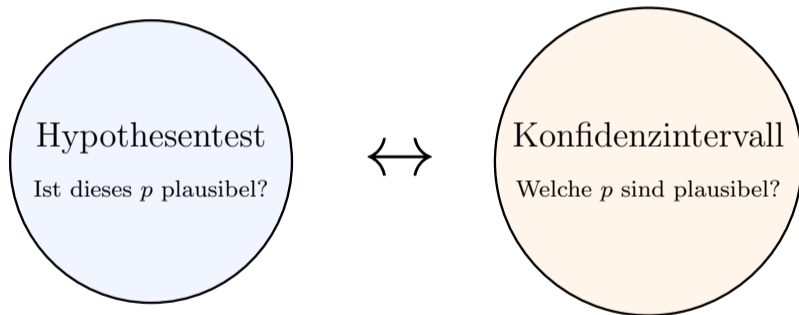
# Worum es mir heute wirklich geht

## Zentrale Leitfrage

Was bedeutet es, wenn wir aus einer Stichprobe auf die Grundgesamtheit schließen?

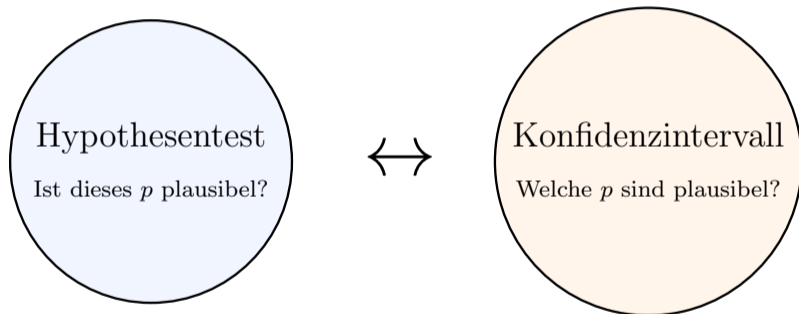
## Vier Eckpfeiler meines Ansatzes

- Unterscheide streng: **Modell** vs. **Realisation**.
- Unterscheide streng: **Zufallsgröße  $X$  (Modell)** vs. **Realisation  $x$  (Realität)**.
- Nutze: **Handlungen, Bilder und Gedankenexperimente**, um Unsicherheit verständlich zu machen - vor dem Erlernen formaler Verfahren.
- Suche: **Anschlussfähigkeit zur Universität**, ohne die formale Bildungsbarriere unnötig zu erhöhen - Strukturen sichtbar machen, nicht die technische Oberfläche.



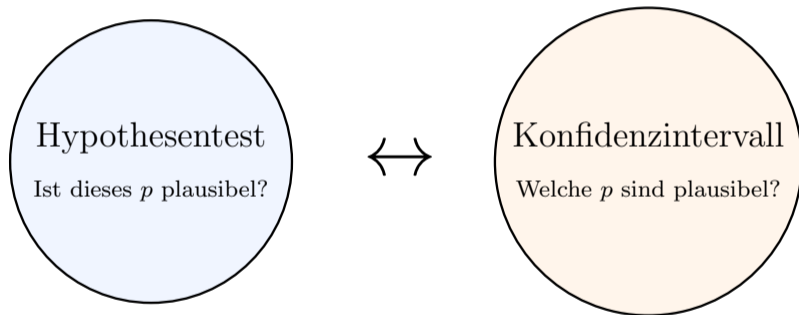
## Zwei Seiten einer Medaille

*Was beide verbindet: dieselbe Stichprobe, dieselben Daten – nur die andere Frage.*



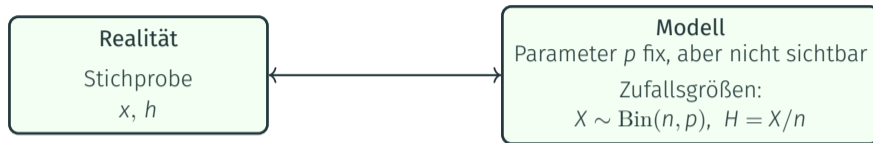
## Zwei Seiten einer Medaille

*Was beide verbindet: dieselbe Stichprobe, dieselben Daten – nur die andere Frage.*

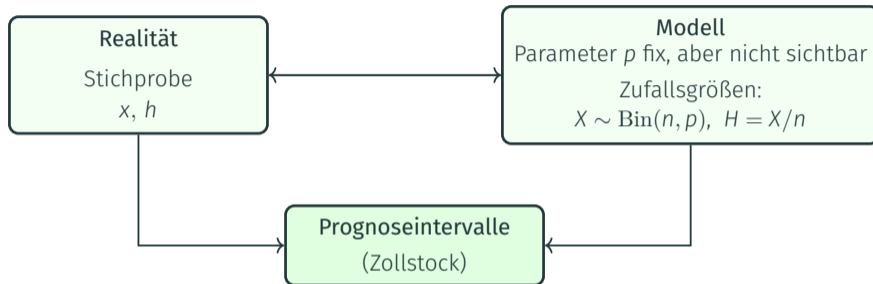


*Der Test auf Niveau  $\alpha$  verwirft  $H_0 \iff p_0 \notin KI$*

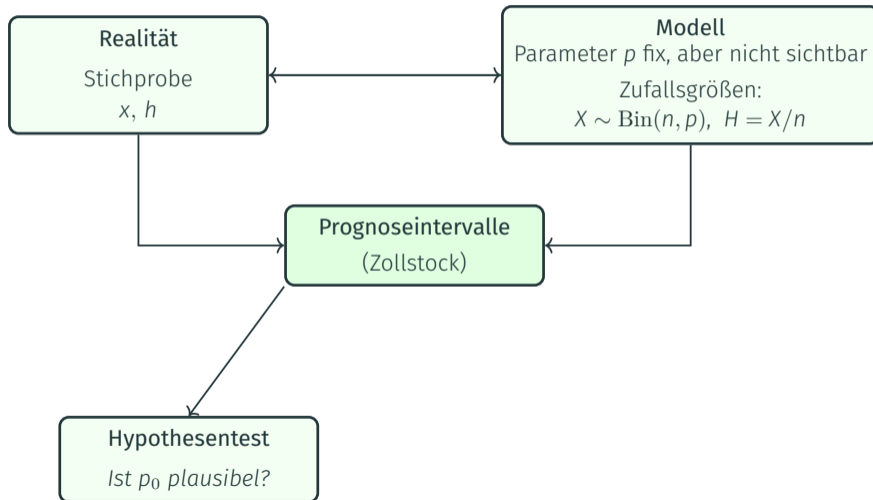
## Zwei Seiten einer Medaille: Realität, Modell, Modelle



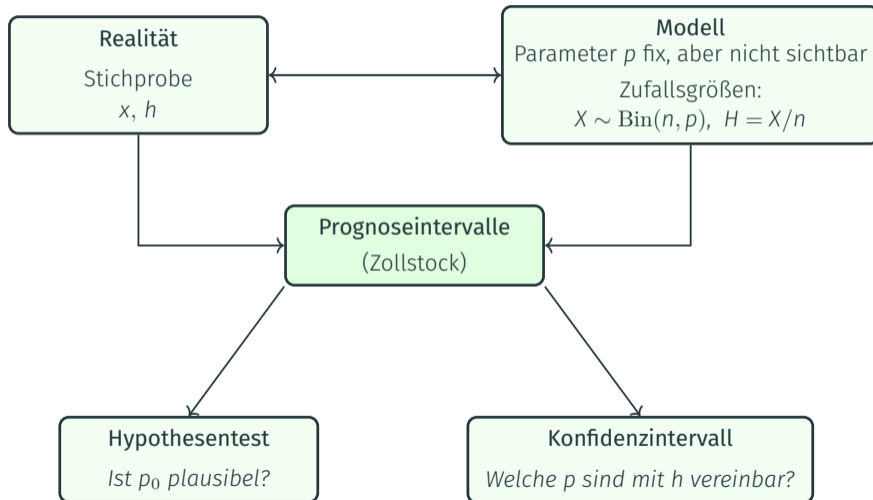
## Zwei Seiten einer Medaille: Realität, Modell, Modelle



# Zwei Seiten einer Medaille: Realität, Modell, Modelle



# Zwei Seiten einer Medaille: Realität, Modell, Modelle



# Leitprinzip der gesamten Inferenz

Trenne Modell und Realität messerscharf!

# Leitprinzip der gesamten Inferenz

Trenne Modell und Realität messerscharf!

## Modellwelt

- beschreibt Zufallsgrößen (z. B.  $H = X/n$ )
- stellt Wahrscheinlichkeiten, Verteilungen und Prognoseintervalle bereit
- lebt *vor* der Stichprobenziehung

# Leitprinzip der gesamten Inferenz

## Trenne Modell und Realität messerscharf!

### Modellwelt

- beschreibt Zufallsgrößen (z. B.  $H = X/n$ )
- stellt Wahrscheinlichkeiten, Verteilungen und Prognoseintervalle bereit
- lebt *vor* der Stichprobenziehung

### Datenwelt

- enthält die beobachtete Realisation (z. B.  $h = x/n$ )
- ist das, was tatsächlich passiert ist
- beginnt *nach* der Stichprobenziehung

Ein Konfidenzintervall ist nicht erfahrungsbasiert, es ist

- *kein Erfahrungsurteil.*
- *kein Glaubensmaß*
- *keine subjektive Unsicherheit*

Ein Konfidenzintervall ist nicht erfahrungsbasiert, es ist

- *kein Erfahrungsurteil.*
- *kein Glaubensmaß*
- *keine subjektive Unsicherheit*

**... sondern ...**

eine Eigenschaft eines Zufallsexperiments unter einem Modell.

Ein Konfidenzintervall ist nicht erfahrungsbasiert, es ist

- *kein Erfahrungsurteil.*
- *kein Glaubensmaß*
- *keine subjektive Unsicherheit*

**... sondern ...**

eine Eigenschaft eines Zufallsexperiments unter einem Modell.

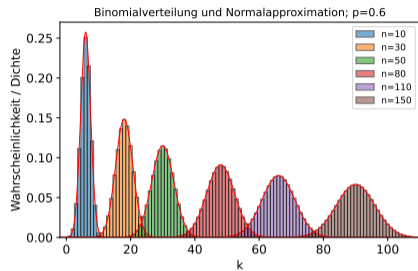
Erfahrung motiviert Modelle.

Konfidenzintervalle sind Aussagen über Verfahren innerhalb eines Modells.

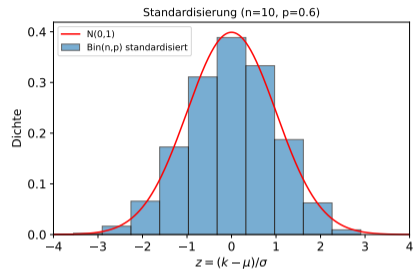
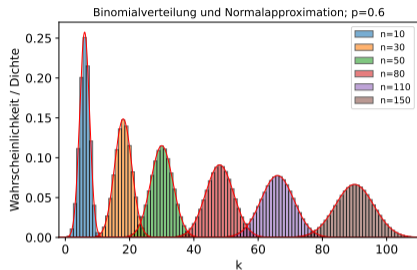
# Prognoseintervalle

---

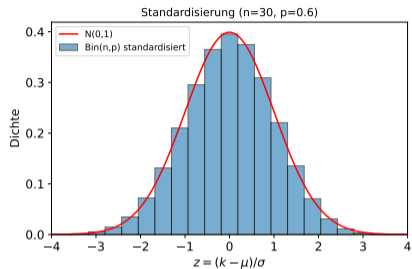
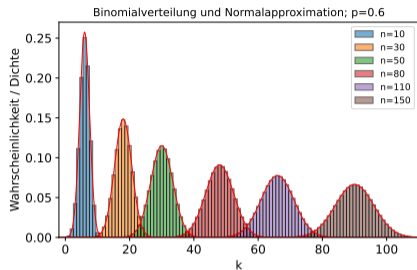
# Der Zollstock als rücktransformierte Normalapproximation



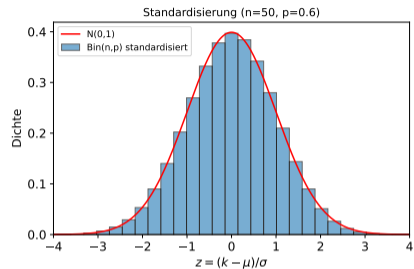
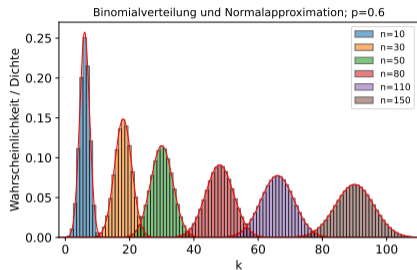
# Der Zollstock als rücktransformierte Normalapproximation



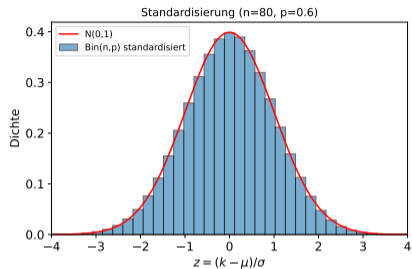
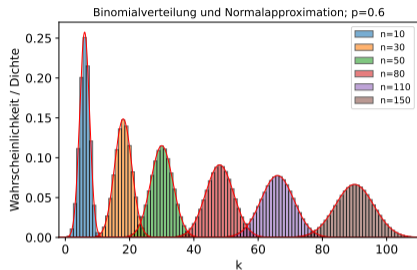
# Der Zollstock als rücktransformierte Normalapproximation



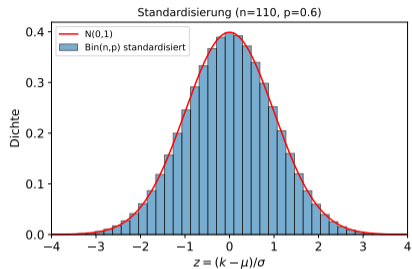
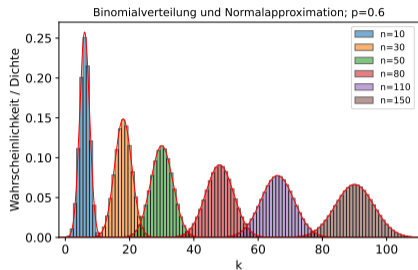
# Der Zollstock als rücktransformierte Normalapproximation



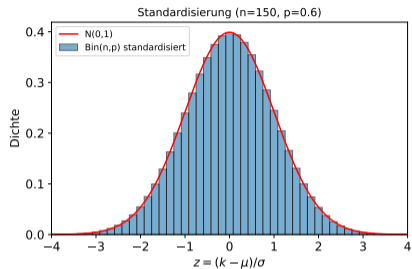
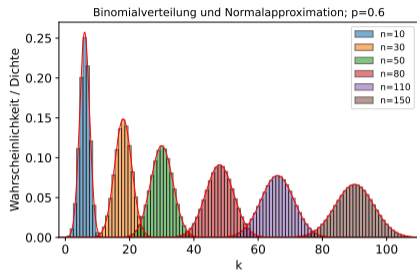
# Der Zollstock als rücktransformierte Normalapproximation



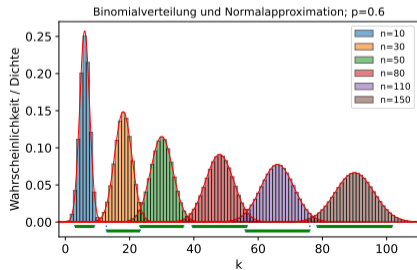
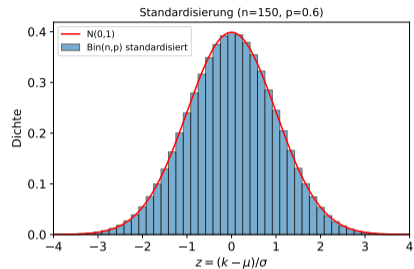
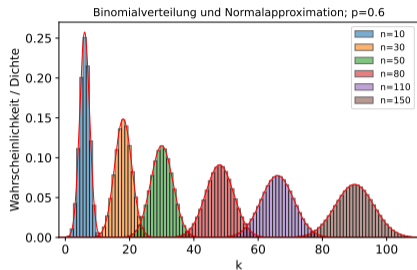
# Der Zollstock als rücktransformierte Normalapproximation



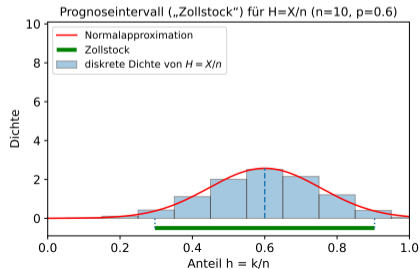
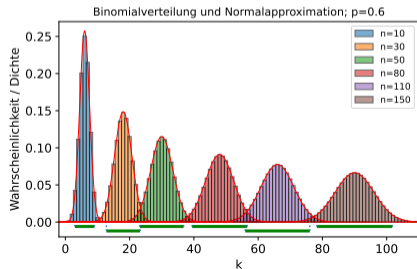
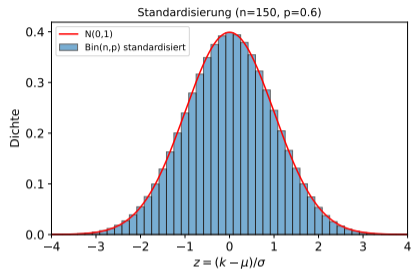
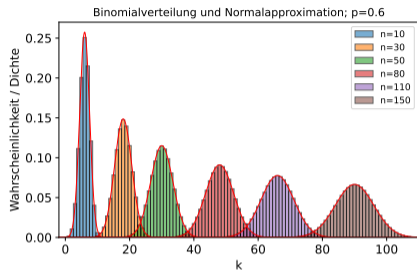
# Der Zollstock als rücktransformierte Normalapproximation



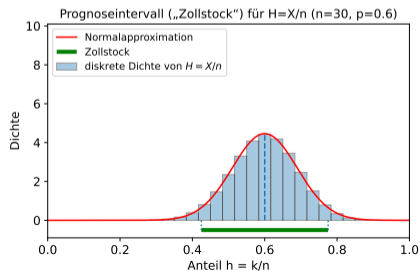
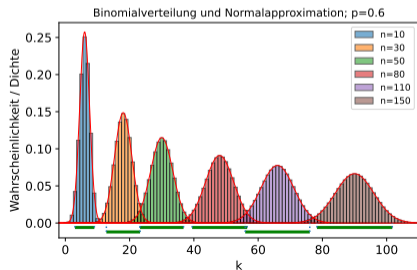
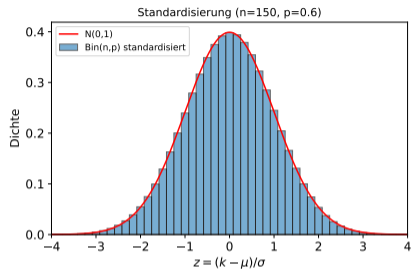
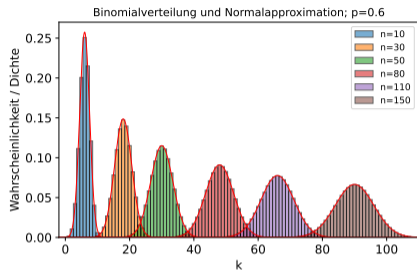
# Der Zollstock als rücktransformierte Normalapproximation



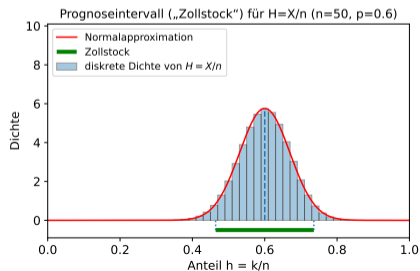
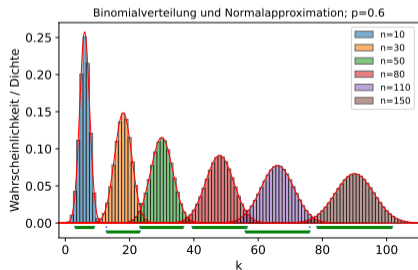
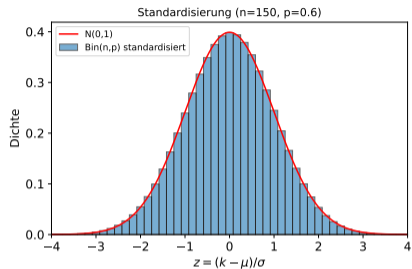
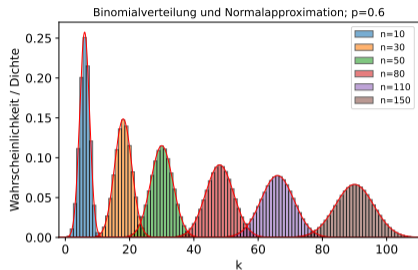
# Der Zollstock als rücktransformierte Normalapproximation



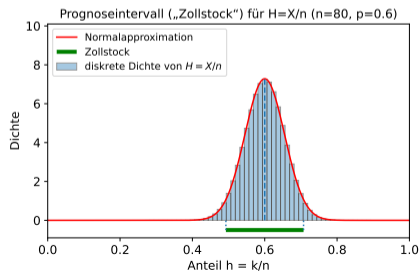
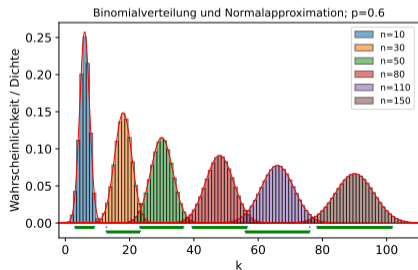
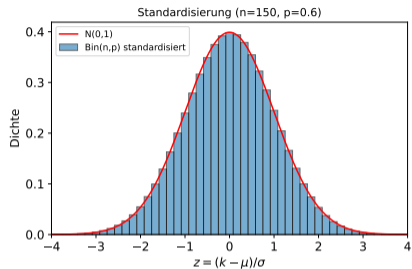
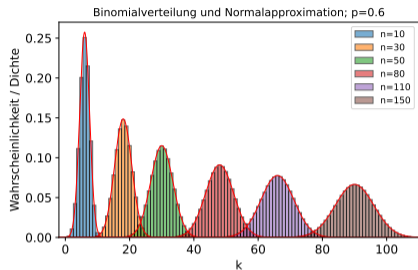
# Der Zollstock als rücktransformierte Normalapproximation



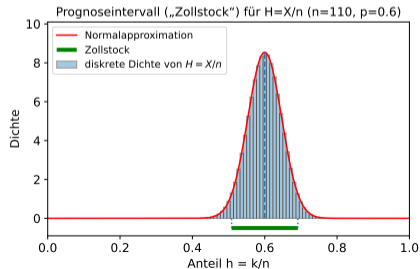
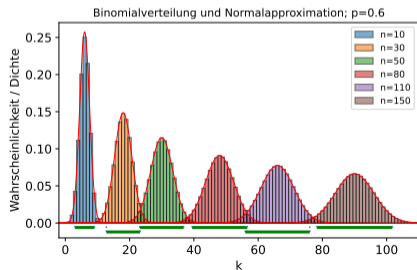
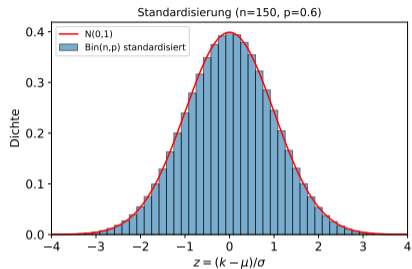
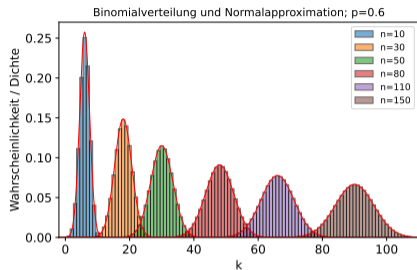
# Der Zollstock als rücktransformierte Normalapproximation



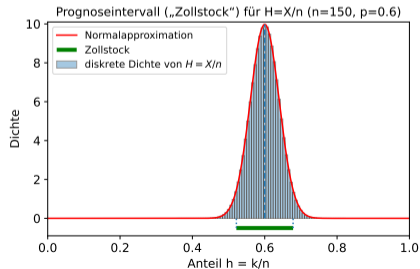
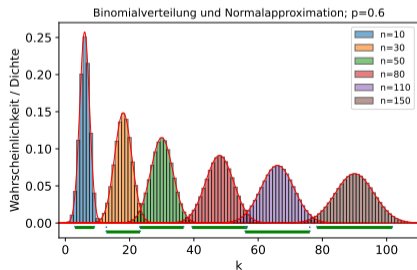
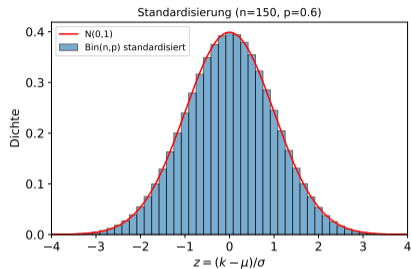
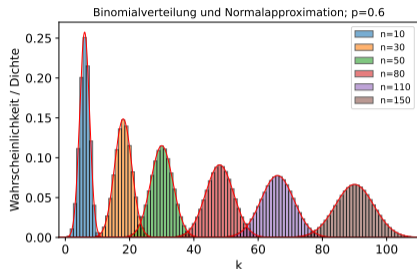
# Der Zollstock als rücktransformierte Normalapproximation



# Der Zollstock als rücktransformierte Normalapproximation



# Der Zollstock als rücktransformierte Normalapproximation



# Prognoseintervall - der Zollstock

Ein 95%-**Prognoseintervall** zu einer Trefferwahrscheinlichkeit  $p$  ist ein Intervall, in das die *zufällige* Trefferanzahl  $X$  bzw. die *zufällige* Trefferhäufigkeit  $H = X/n$  bei einem *bevorstehenden* Bernoulliexperiment gegebener Länge  $n$  mit einer Mindestwahrscheinlichkeit 95% fällt.

# Prognoseintervall - der Zollstock

Ein 95%-**Prognoseintervall** zu einer Trefferwahrscheinlichkeit  $p$  ist ein Intervall, in das die *zufällige* Trefferanzahl  $X$  bzw. die *zufällige* Trefferhäufigkeit  $H = X/n$  bei einem *bevorstehenden* Bernoulliexperiment gegebener Länge  $n$  mit einer Mindestwahrscheinlichkeit 95% fällt.

Prognoseintervalle schauen in die Zukunft. Man schließt mit ihnen vom Modell (der Wahrscheinlichkeit) auf die Realität (relative Häufigkeit).

Bei gegebener Trefferwahrscheinlichkeit  $p$  liegt die *zufällige* Trefferhäufigkeit  $H = X/n$  mit ca. 95% Sicherheit im Prognoseintervall

$$\left[ p - 1,96\sqrt{\frac{p(1-p)}{n}} ; p + 1,96\sqrt{\frac{p(1-p)}{n}} \right].$$

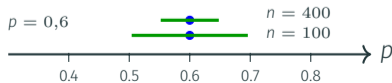
# Prognoseintervall - der Zollstock

Ein 95%-**Prognoseintervall** zu einer Trefferwahrscheinlichkeit  $p$  ist ein Intervall, in das die *zufällige* Trefferanzahl  $X$  bzw. die *zufällige* Trefferhäufigkeit  $H = X/n$  bei einem *bevorstehenden* Bernoulliexperiment gegebener Länge  $n$  mit einer Mindestwahrscheinlichkeit 95% fällt.

Prognoseintervalle schauen in die Zukunft. Man schließt mit ihnen vom Modell (der Wahrscheinlichkeit) auf die Realität (relative Häufigkeit).

Bei gegebener Trefferwahrscheinlichkeit  $p$  liegt die *zufällige* Trefferhäufigkeit  $H = X/n$  mit ca. 95% Sicherheit im Prognoseintervall

$$\left[ p - 1,96\sqrt{\frac{p(1-p)}{n}} ; p + 1,96\sqrt{\frac{p(1-p)}{n}} \right].$$



# Prognoseintervalle - Denken in Verteilungen

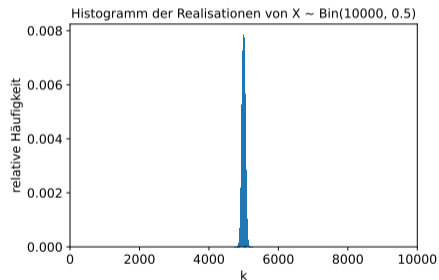
Für eine binomialverteilte Zufallsgröße  $X$  gilt:  $n = 10\,000$ ;  $p = 0,5$ .

Finden Sie Aufgaben der Form  $P(X = k)$ ;  $P(X \leq k)$ ;  $P(X \geq k)$ ;  $P(a \leq X \leq b), \dots$  .

# Prognoseintervalle - Denken in Verteilungen

Für eine binomialverteilte Zufallsgröße  $X$  gilt:  $n = 10\,000$ ;  $p = 0,5$ .

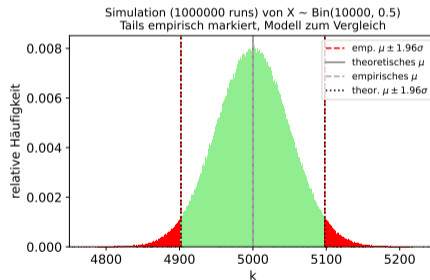
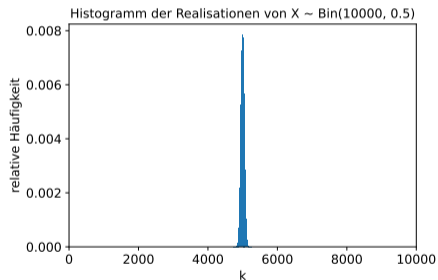
Finden Sie Aufgaben der Form  $P(X = k)$ ;  $P(X \leq k)$ ;  $P(X \geq k)$ ;  $P(a \leq X \leq b)$ , ... .



# Prognoseintervalle - Denken in Verteilungen

Für eine binomialverteilte Zufallsgröße  $X$  gilt:  $n = 10\,000$ ;  $p = 0,5$ .

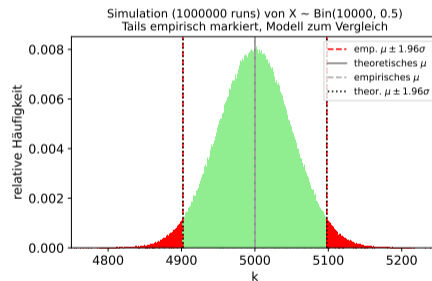
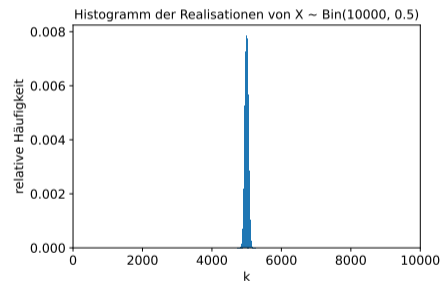
Finden Sie Aufgaben der Form  $P(X = k)$ ;  $P(X \leq k)$ ;  $P(X \geq k)$ ;  $P(a \leq X \leq b)$ , ... .



# Prognoseintervalle - Denken in Verteilungen

Für eine binomialverteilte Zufallsgröße  $X$  gilt:  $n = 10\,000$ ;  $p = 0,5$ .

Finden Sie Aufgaben der Form  $P(X = k)$ ;  $P(X \leq k)$ ;  $P(X \geq k)$ ;  $P(a \leq X \leq b)$ , ... .



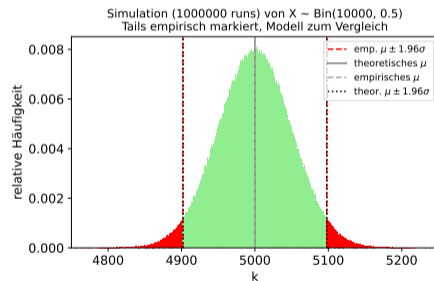
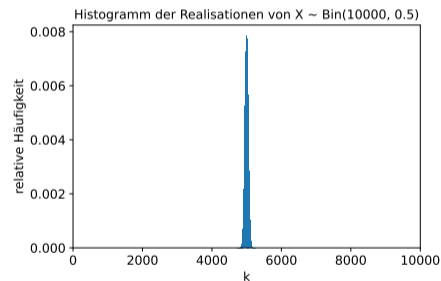
Idee (ohne Rechner):  $\mu = n \cdot p = 5000$ ;  $\sigma = \sqrt{10000 \cdot 0,5 \cdot 0,5} = 50$ ;  $2\sigma = 100$

Folgerung: In ca. 95% aller zukünftigen Fälle gilt:  $4900 \leq X \leq 5100$ .

# Prognoseintervalle - Denken in Verteilungen

Für eine binomialverteilte Zufallsgröße  $X$  gilt:  $n = 10\,000$ ;  $p = 0,5$ .

Finden Sie Aufgaben der Form  $P(X = k)$ ;  $P(X \leq k)$ ;  $P(X \geq k)$ ;  $P(a \leq X \leq b)$ , ... .



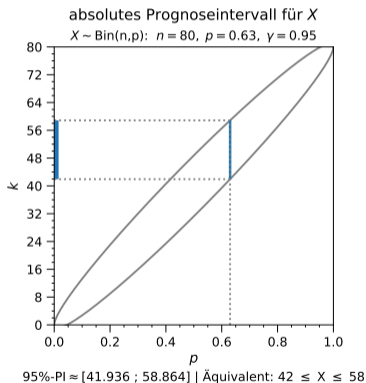
Idee (ohne Rechner):  $\mu = n \cdot p = 5000$ ;  $\sigma = \sqrt{10000 \cdot 0,5 \cdot 0,5} = 50$ ;  $2\sigma = 100$

Folgerung: In ca. 95% aller zukünftigen Fälle gilt:  $4900 \leq X \leq 5100$ . ( $0,49 \leq \frac{X}{n} \leq 0,51$ )

# Prognoseintervalle: absolut und relativ - Normierung liefert Einsichten

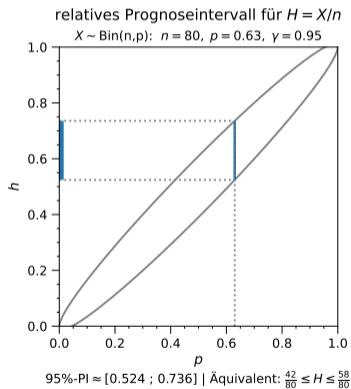
## Absolutes Prognoseintervall für $X$

$$\left[ n \cdot p - z \sqrt{np(1-p)}, n \cdot p + z \sqrt{np(1-p)} \right]$$



## Relatives Prognoseintervall für $H = X/n$

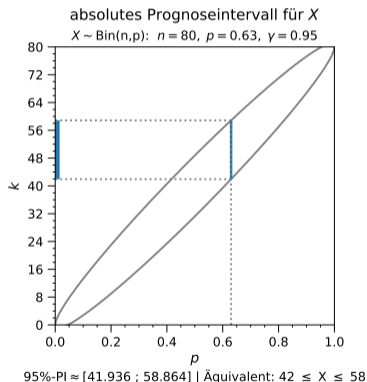
$$\left[ p - z \sqrt{p(1-p)/n}, p + z \sqrt{p(1-p)/n} \right]$$



# Prognoseintervalle: absolut und relativ - Normierung liefert Einsichten

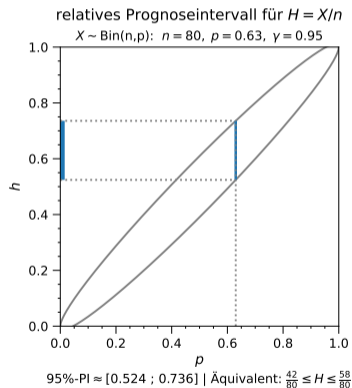
## Absolutes Prognoseintervall für $X$

$$\left[ n \cdot p - z \sqrt{np(1-p)}, n \cdot p + z \sqrt{np(1-p)} \right]$$



## Relatives Prognoseintervall für $H = X/n$

$$\left[ p - z \sqrt{p(1-p)/n}, p + z \sqrt{p(1-p)/n} \right]$$



Lineare Transformation  $H = \frac{X}{n}$ . Skalierung ändert die Darstellung, nicht die Struktur.  
Normierung ist eine lineare Transformation - das  $1/\sqrt{n}$ -Gesetz wird sichtbar.

# Hypothesentest

---

Ein Hypothesentest entscheidet, ob beobachtete Daten unter Annahme von  $H_0$  ungewöhnlich sind.

Die Entscheidungsregel wird *vor* den Daten festgelegt.

## Zeitliche Ordnung eines Tests

1. **Vor der Stichprobe:** Modell, Signifikanzniveau und Entscheidungsregel werden festgelegt.
2. **Dann:** Die Stichprobe wird erhoben.
3. **Danach:** Die festgelegte Regel wird auf die Daten angewendet.

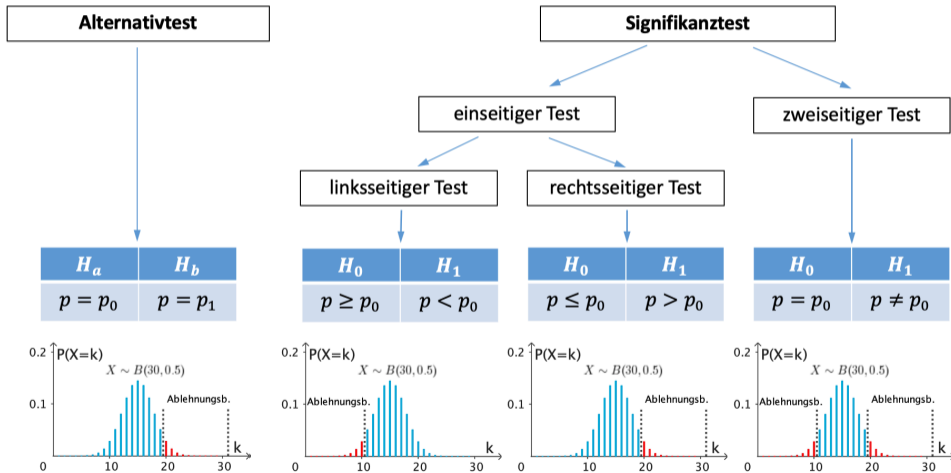
## Gedankenexperiment eines Tests:

*Wie ungewöhnlich wären diese Daten,  
wenn  $H_0$  gegolten hätte?*

Testen bewertet Daten – nicht die Wahrheit.

# Hypothesentests - ganz schön verwirrend, trotz Diagramm

## Hypothesentest – eine erste Übersicht



## Hypothesentest – eine erste Übersicht

Signifikanztest

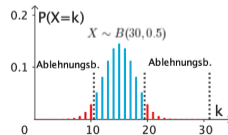
zweiseitiger Test

$H_0$

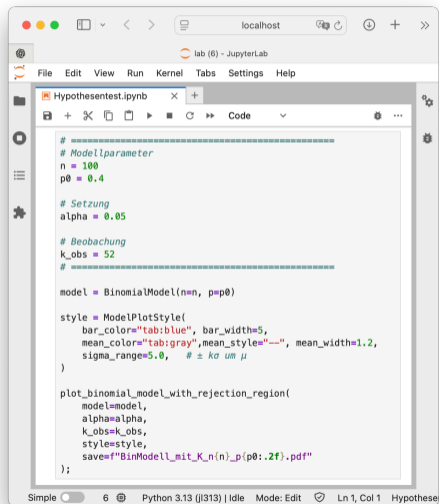
$H_1$

$p = p_0$

$p \neq p_0$



# Modell - Setzung - Style - Ausgabe - Interpretation



The screenshot shows a JupyterLab window titled "lab (6) - JupyterLab" with a browser address bar at "localhost". The notebook file is named "Hypothesentest.ipynb". The code is written in Python and includes comments in German. The code defines model parameters, sets the significance level, and generates a plot with a rejection region. The status bar at the bottom indicates "Simple" mode, Python 3.13 (j1313) kernel, and the current position is "Ln 1, Col 1".

```
# =====  
# Modellparameter  
n = 100  
p0 = 0.4  
  
# Setzung  
alpha = 0.05  
  
# Beobachtung  
k_obs = 52  
# =====  
  
model = BinomialModel(n=n, p=p0)  
  
style = ModelPlotStyle(  
    bar_color="tab:blue", bar_width=5,  
    mean_color="tab:gray", mean_style="--", mean_width=1.2,  
    sigma_range=5.0, # ± kσ um μ  
)  
  
plot_binomial_model_with_rejection_region(  
    model=model,  
    alpha=alpha,  
    k_obs=k_obs,  
    style=style,  
    save=f"BinModell_mit_K_n{n}_p{p0:.2f}.pdf"  
);
```

# Modell - Setzung - Style - Ausgabe - Interpretation

```
lab (6) - JupyterLab
File Edit View Run Kernel Tabs Settings Help
Hypothesentest.ipynb
# -----
# Modellparameter
n = 100
p0 = 0.4

# Setzung
alpha = 0.05

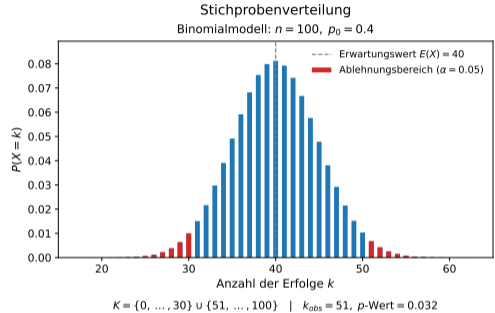
# Beobachtung
k_obs = 52
# -----

model = BinomialModel(n=n, p=p0)

style = ModelPlotStyle(
    bar_color="tab:blue", bar_width=5,
    mean_color="tab:gray", mean_style="---", mean_width=1.2,
    sigma_range=5.0, # ± kσ um μ
)

plot_binomial_model_with_rejection_region(
    model=model,
    alpha=alpha,
    k_obs=k_obs,
    style=style,
    save=f"BinModell_mit_K_n{n}_p{p0:.2f}.pdf"
);

Simple 6 Python 3.13 (j1313) | Idle Mode: Edit Ln 1, Col 1 Hypothese
```



# Modell - Setzung - Style - Ausgabe - Interpretation

```
localhost
lab (6) - JupyterLab
File Edit View Run Kernel Tabs Settings Help
Hypothesentest.ipynb
# -----
# Modellparameter
n = 100
p0 = 0.4

# Setzung
alpha = 0.05

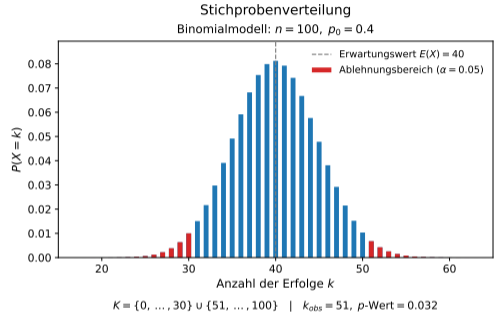
# Beobachtung
k_obs = 52
# -----

model = BinomialModel(n=n, p=p0)

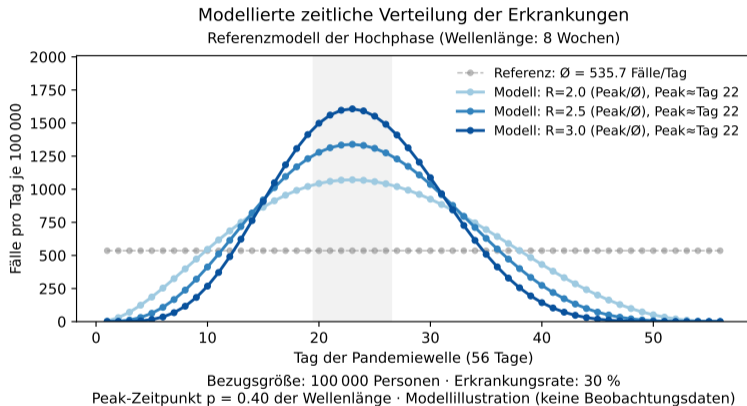
style = ModelPlotStyle(
    bar_color="tab:blue", bar_width=5,
    mean_color="tab:gray", mean_style="---", mean_width=1.2,
    sigma_range=5.0, # ± kσ um μ
)

plot_binomial_model_with_rejection_region(
    model=model,
    alpha=alpha,
    k_obs=k_obs,
    style=style,
    save=f"BinModell_mit_K_n{n}_p{p0:.2f}.pdf"
);

Simple 6 Python 3.13 (j1313) | Idle Mode: Edit Ln 1, Col 1 Hypothese
```



# Trennung von Modell und Realität - außerhalb der Schule



# Die Power eines Tests - Grundwissen (wäre schön)

## Die Power eines Testes

- Die Power beschreibt die Leistungsfähigkeit eines Tests – nicht die Relevanz eines Effekts.
- Relevanz ist keine stochastische Kategorie - sie gehört zur Fragestellung, nicht zum Test.
- Ob ein Effekt relevant ist, entscheidet nicht der Test, sondern der Kontext der Fragestellung.
- Mit genügend großem Stichprobenumfang kann jeder noch so kleine Unterschied statistisch auffällig werden.
- Ein großes  $n$  macht Tests empfindlich, nicht unbedingt wichtiger.

## Die Power eines Testes

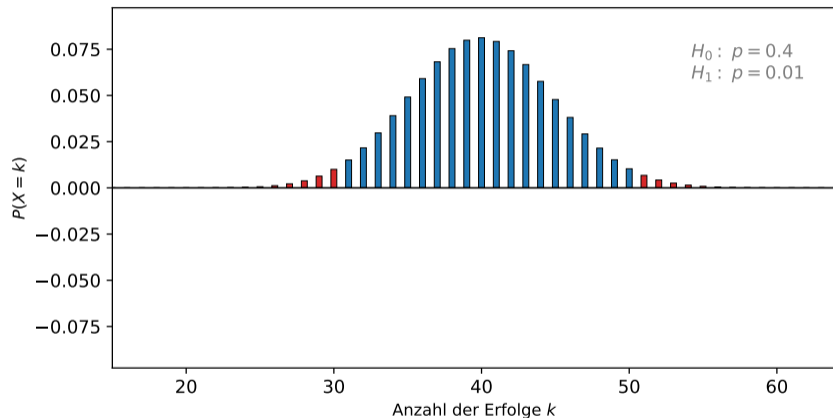
- Die Power beschreibt die Leistungsfähigkeit eines Tests – nicht die Relevanz eines Effekts.
- Relevanz ist keine stochastische Kategorie - sie gehört zur Fragestellung, nicht zum Test.
- Ob ein Effekt relevant ist, entscheidet nicht der Test, sondern der Kontext der Fragestellung.
- Mit genügend großem Stichprobenumfang kann jeder noch so kleine Unterschied statistisch auffällig werden.
- Ein großes  $n$  macht Tests empfindlich, nicht unbedingt wichtiger.

Die Power (Teststärke) eines zweiseitigen Binomialtests beschreibt die Wahrscheinlichkeit, die Nullhypothese  $H_0 : p = p_0$  korrekt abzulehnen, wenn eine bestimmte Alternativhypothese  $H_1 : p = p_1$ ; ( $p_1 \neq p_0$ ) wahr ist.

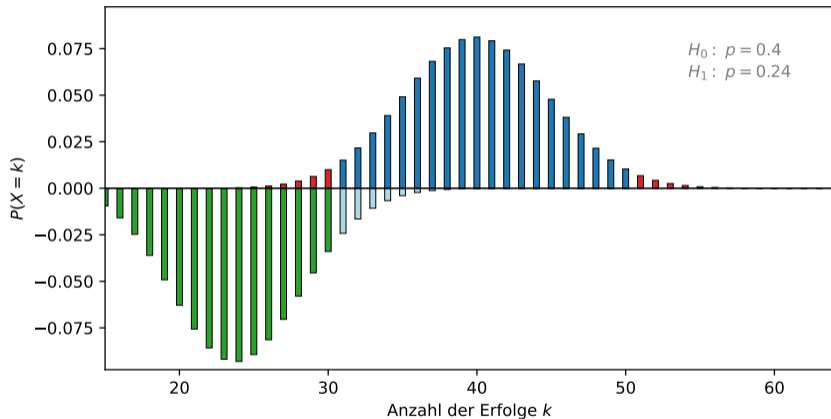
$$\text{Power} = \mathbb{P}(X \leq k_u | p_1) + \mathbb{P}(X \geq k_o | p_1), \quad K = \{0, \dots, k_u\} \cup \{k_o, \dots, n\}$$

# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen

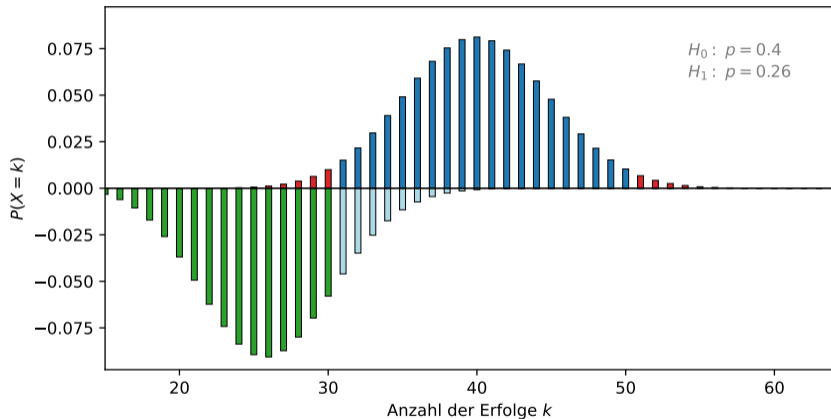
# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



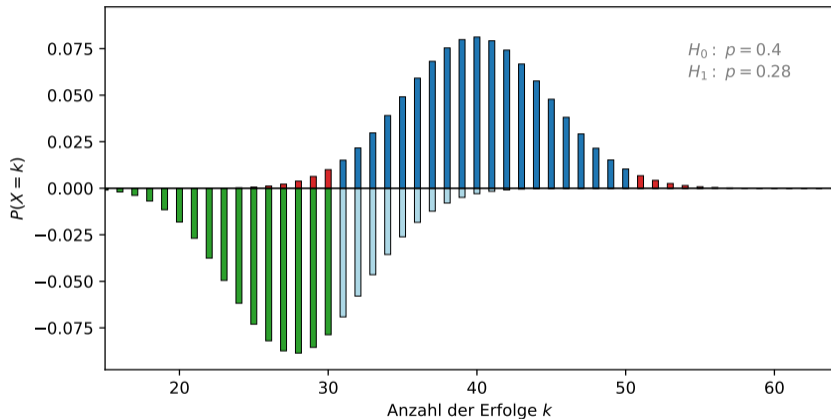
# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



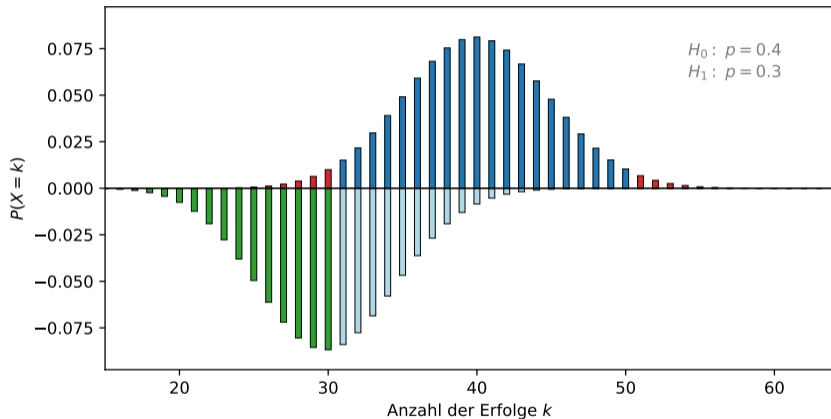
# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



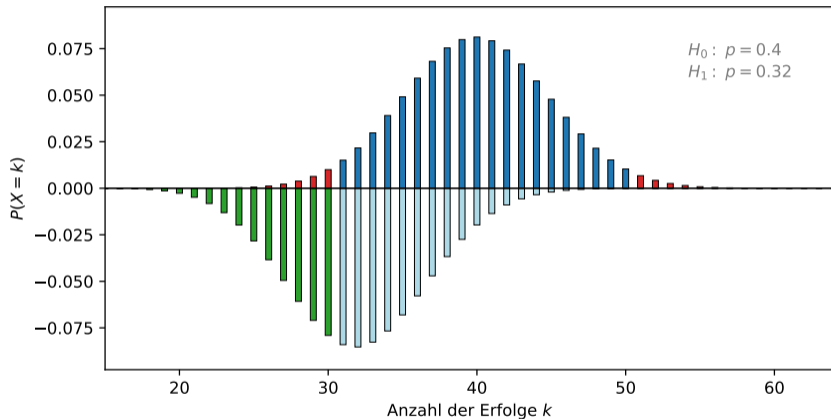
# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



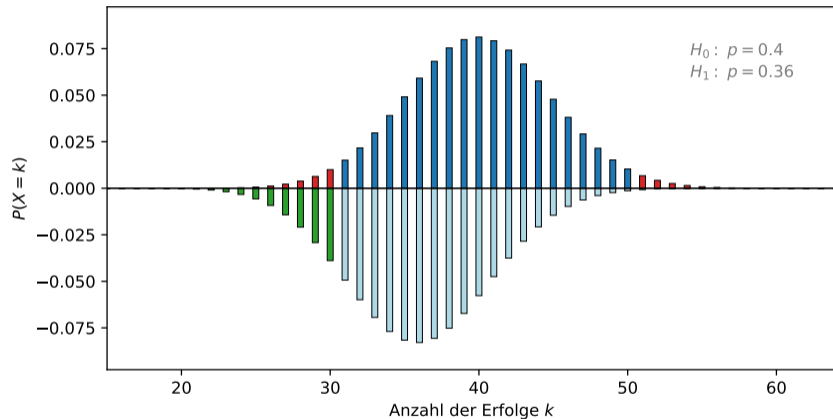
# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



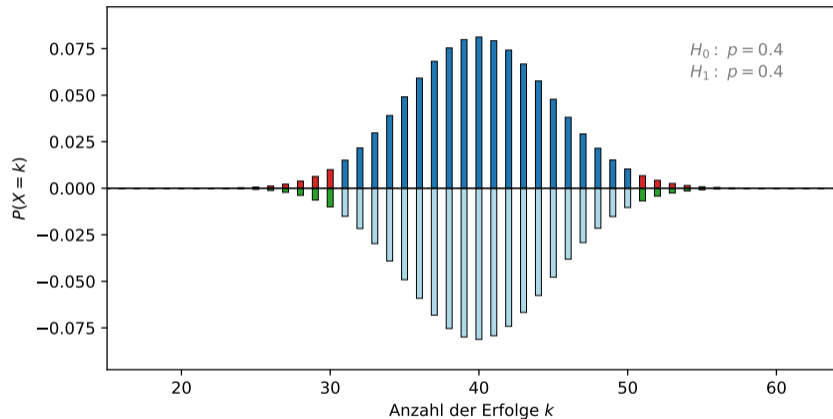
# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



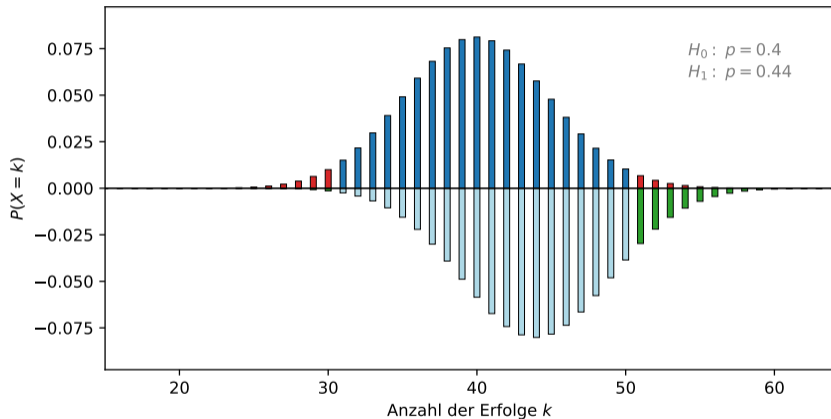
# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



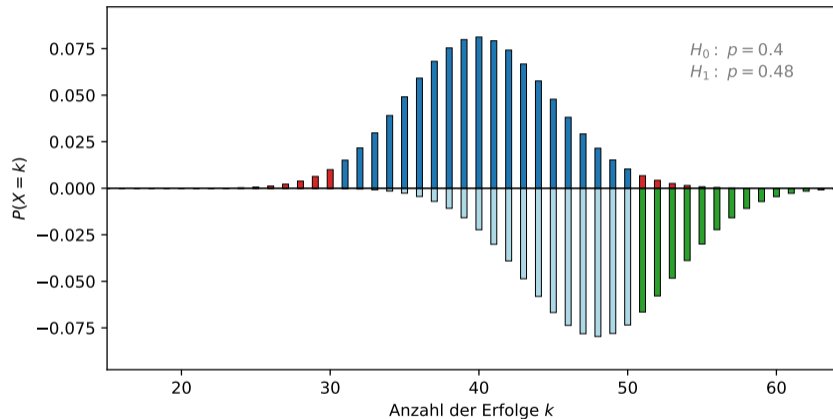
# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



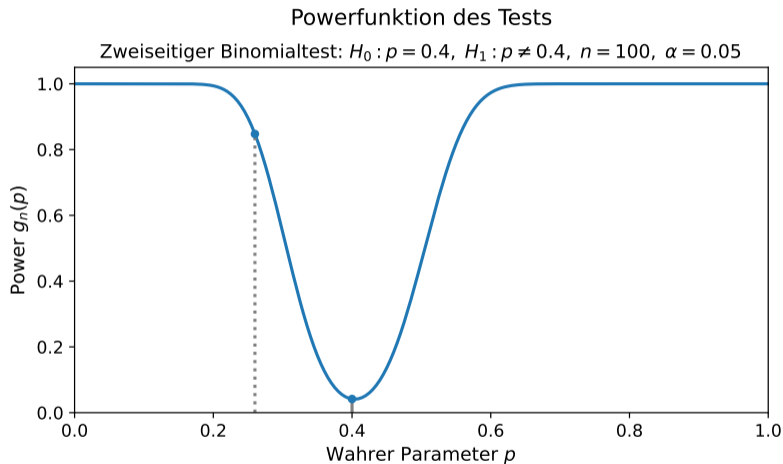
# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



# Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen

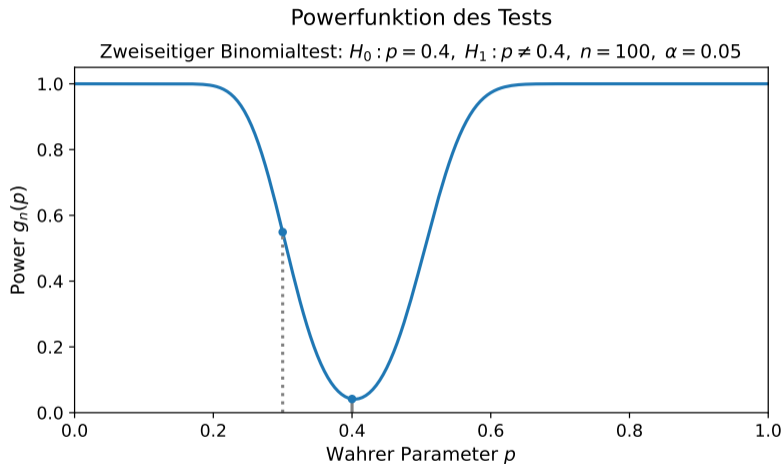


# Powerfunktion $g_n(p) = \mathbb{P}_p(X \in K)$ - Wahrscheinlichkeitsmasse von $p$ in $K$



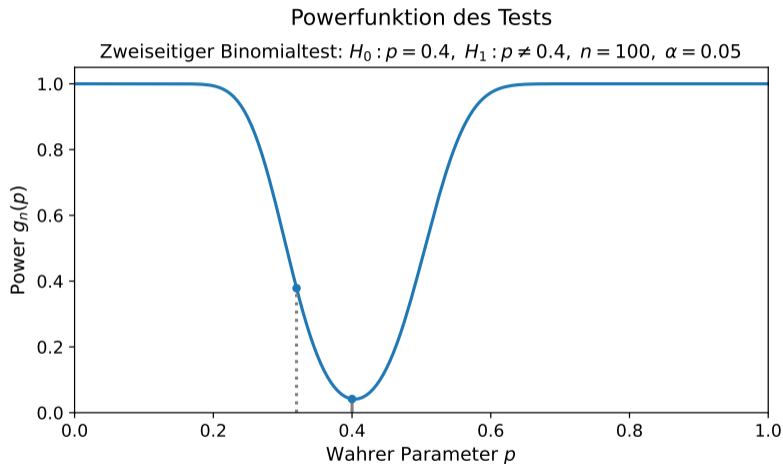
$K = \{0, \dots, 30\} \cup \{51, \dots, 100\}$  | Beispiel:  $p = 0.26 \Rightarrow g_{100}(0.26) = 0.847$

# Powerfunktion $g_n(p) = \mathbb{P}_p(X \in K)$ - Wahrscheinlichkeitsmasse von $p$ in $K$



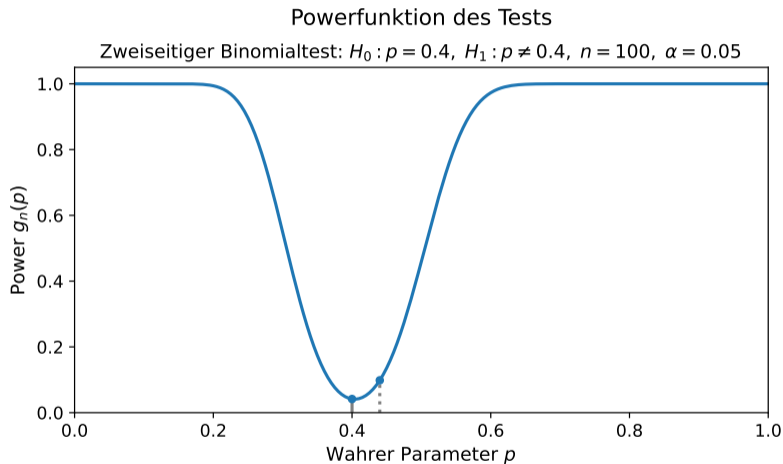
$K = \{0, \dots, 30\} \cup \{51, \dots, 100\}$  | Beispiel:  $p = 0.3 \Rightarrow g_{100}(0.3) = 0.549$

# Powerfunktion $g_n(p) = \mathbb{P}_p(X \in K)$ - Wahrscheinlichkeitsmasse von $p$ in $K$



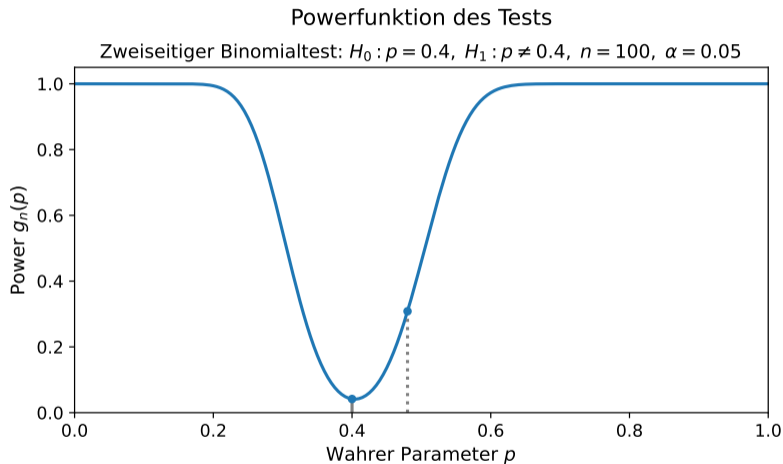
$K = \{0, \dots, 30\} \cup \{51, \dots, 100\}$  | Beispiel:  $p = 0.32 \Rightarrow g_{100}(0.32) = 0.378$

# Powerfunktion $g_n(p) = \mathbb{P}_p(X \in K)$ - Wahrscheinlichkeitsmasse von $p$ in $K$



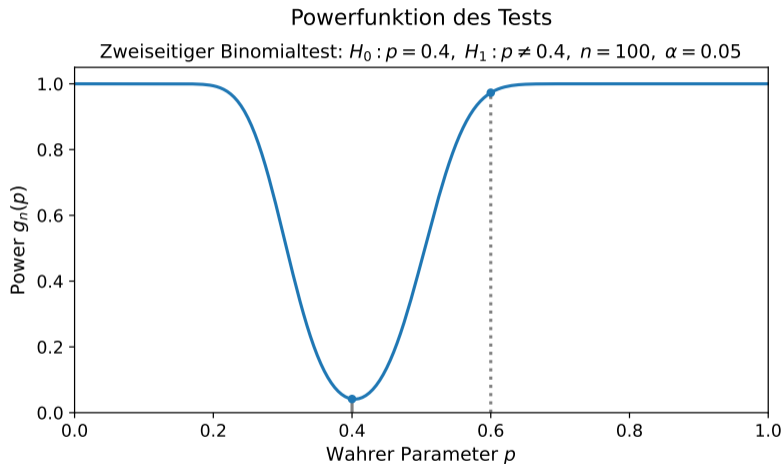
$K = \{0, \dots, 30\} \cup \{51, \dots, 100\}$  | Beispiel:  $p = 0.44 \Rightarrow g_{100}(0.44) = 0.098$

# Powerfunktion $g_n(p) = \mathbb{P}_p(X \in K)$ - Wahrscheinlichkeitsmasse von $p$ in $K$



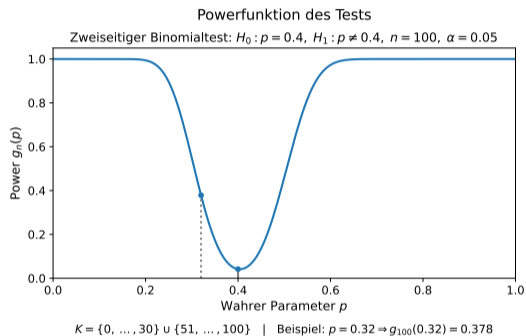
$K = \{0, \dots, 30\} \cup \{51, \dots, 100\}$  | Beispiel:  $p = 0.48 \Rightarrow g_{100}(0.48) = 0.308$

# Powerfunktion $g_n(p) = \mathbb{P}_p(X \in K)$ - Wahrscheinlichkeitsmasse von $p$ in $K$

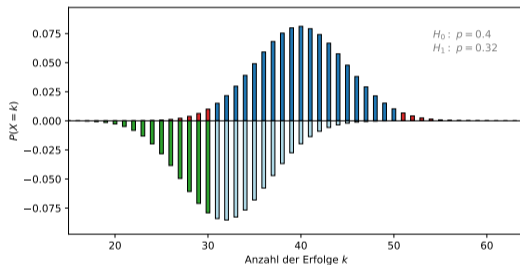
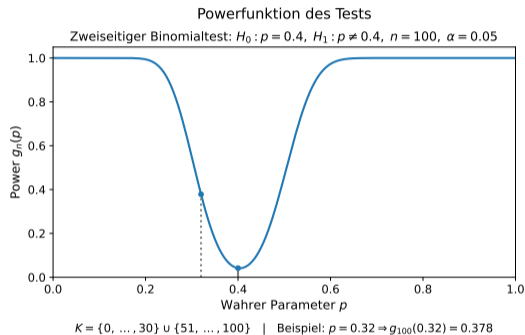


$K = \{0, \dots, 30\} \cup \{51, \dots, 100\}$  | Beispiel:  $p = 0.6 \Rightarrow g_{100}(0.6) = 0.973$

# Powerfunktion und Spiegelung



# Powerfunktion und Spiegelung



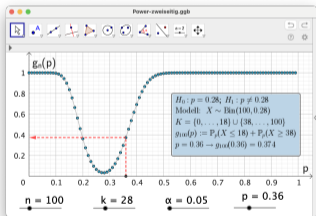
# Ein Modell – zwei digitale Perspektiven

## GeoGebra & Python – zwei Werkzeuge, ein Modell

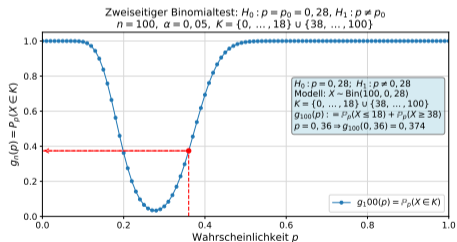
**GeoGebra** Schulnah, interaktiv und zeilenorientiert, sofort sichtbare Objekte

**Python** Präzise, reproduzierbar, ideal für Simulation und professionelles Arbeiten

### GeoGebra: Diskrete Powerfunktion



### Python: Diskrete Powerfunktion



*Didaktische Leitidee: GeoGebra zum Verstehen – Python zum Produzieren.*

## Planung des Stichprobenumfangs – sinnvoll probieren

**Ausgangsfrage:** Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit ein Test auf Niveau  $\alpha = 0,1$  für  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit Wahrscheinlichkeit  $\gamma = 0,9$  die richtige Entscheidung ( $H_1$  trifft zu) gibt, falls  $p = 0,6$  gilt?

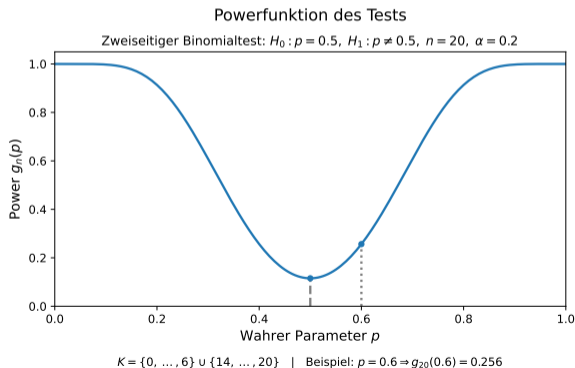
## Planung des Stichprobenumfangs – sinnvoll probieren

**Ausgangsfrage:** Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit ein Test auf Niveau  $\alpha = 0,1$  für  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit Wahrscheinlichkeit  $\gamma = 0,9$  die richtige Entscheidung ( $H_1$  trifft zu) gibt, falls  $p = 0,6$  gilt?

Grafiken können helfen...

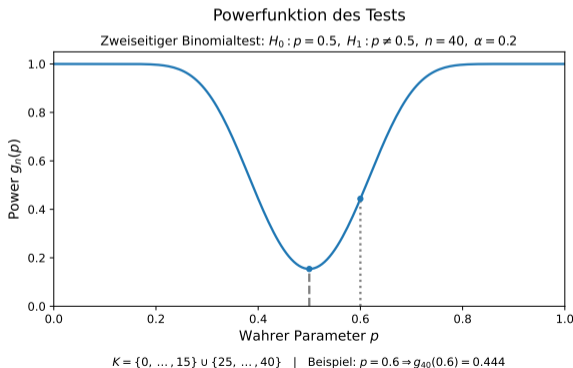
# Planung des Stichprobenumfangs – sinnvoll probieren

**Ausgangsfrage:** Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit ein Test auf Niveau  $\alpha = 0,1$  für  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit Wahrscheinlichkeit  $\gamma = 0,9$  die richtige Entscheidung ( $H_1$  trifft zu) gibt, falls  $p = 0,6$  gilt?



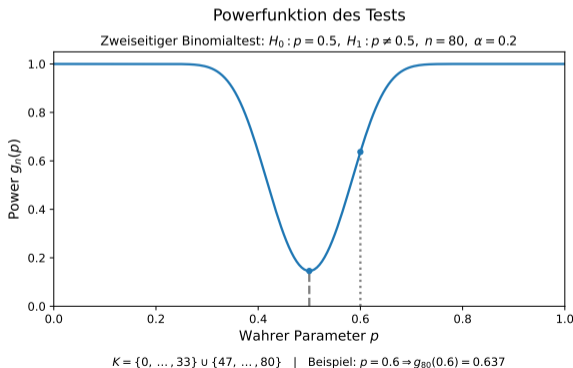
# Planung des Stichprobenumfangs – sinnvoll probieren

**Ausgangsfrage:** Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit ein Test auf Niveau  $\alpha = 0,1$  für  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit Wahrscheinlichkeit  $\gamma = 0,9$  die richtige Entscheidung ( $H_1$  trifft zu) gibt, falls  $p = 0,6$  gilt?



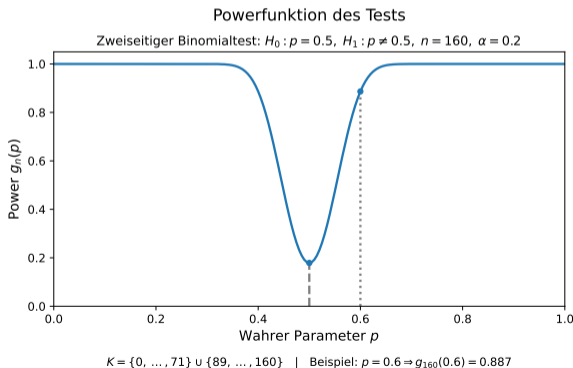
# Planung des Stichprobenumfangs – sinnvoll probieren

**Ausgangsfrage:** Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit ein Test auf Niveau  $\alpha = 0,1$  für  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit Wahrscheinlichkeit  $\gamma = 0,9$  die richtige Entscheidung ( $H_1$  trifft zu) gibt, falls  $p = 0,6$  gilt?



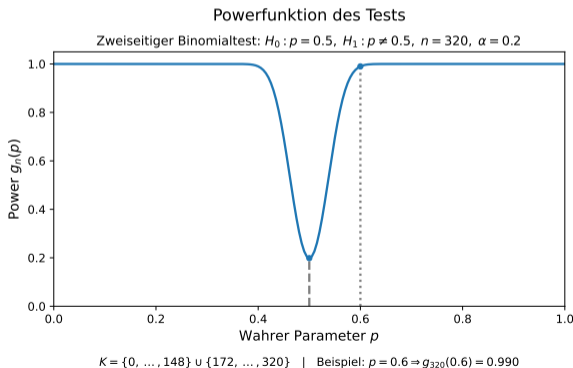
# Planung des Stichprobenumfangs – sinnvoll probieren

**Ausgangsfrage:** Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit ein Test auf Niveau  $\alpha = 0,1$  für  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit Wahrscheinlichkeit  $\gamma = 0,9$  die richtige Entscheidung ( $H_1$  trifft zu) gibt, falls  $p = 0,6$  gilt?



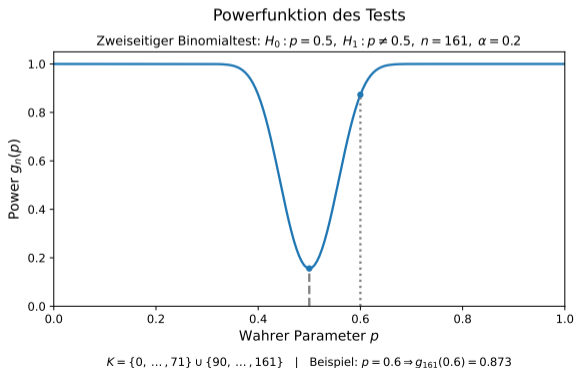
# Planung des Stichprobenumfangs – sinnvoll probieren

**Ausgangsfrage:** Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit ein Test auf Niveau  $\alpha = 0,1$  für  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit Wahrscheinlichkeit  $\gamma = 0,9$  die richtige Entscheidung ( $H_1$  trifft zu) gibt, falls  $p = 0,6$  gilt?



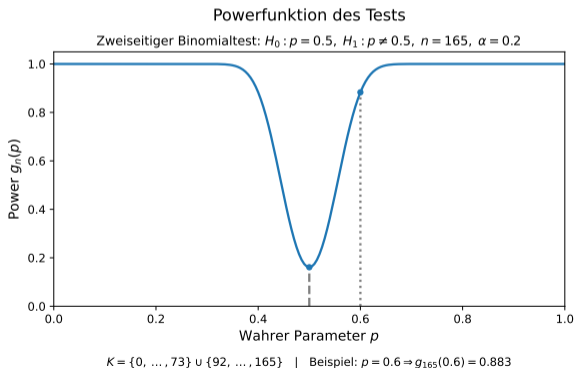
# Planung des Stichprobenumfangs – sinnvoll probieren

**Ausgangsfrage:** Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit ein Test auf Niveau  $\alpha = 0,1$  für  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit Wahrscheinlichkeit  $\gamma = 0,9$  die richtige Entscheidung ( $H_1$  trifft zu) gibt, falls  $p = 0,6$  gilt?



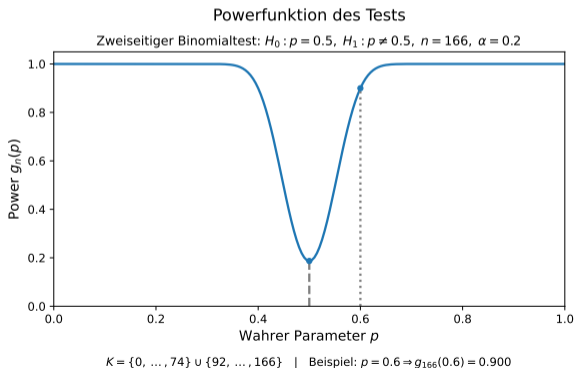
# Planung des Stichprobenumfangs – sinnvoll probieren

**Ausgangsfrage:** Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit ein Test auf Niveau  $\alpha = 0,1$  für  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit Wahrscheinlichkeit  $\gamma = 0,9$  die richtige Entscheidung ( $H_1$  trifft zu) gibt, falls  $p = 0,6$  gilt?



# Planung des Stichprobenumfangs – sinnvoll probieren

**Ausgangsfrage:** Wie groß muss der Stichprobenumfang  $n$  mindestens sein, damit ein Test auf Niveau  $\alpha = 0,1$  für  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit Wahrscheinlichkeit  $\gamma = 0,9$  die richtige Entscheidung ( $H_1$  trifft zu) gibt, falls  $p = 0,6$  gilt?



## Planung des Stichprobenumfangs – mit Zollstöcken

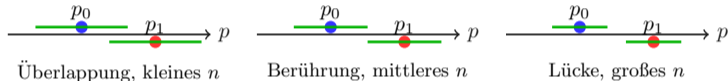
Test  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit  $\alpha = 0,1$ ;  $\gamma = 0,9$ ;  $p_1 = 0,6$ .

Zwei Zollstöcke aber auch...

# Planung des Stichprobenumfangs – mit Zollstöcken

Test  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit  $\alpha = 0,1$ ;  $\gamma = 0,9$ ;  $p_1 = 0,6$ .

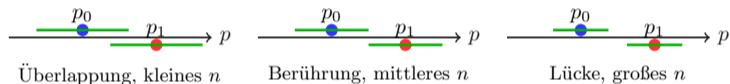
Zwei Zollstöcke aber auch...



# Planung des Stichprobenumfangs – mit Zollstöcken

Test  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit  $\alpha = 0,1$ ;  $\gamma = 0,9$ ;  $p_1 = 0,6$ .

Zwei Zollstöcke aber auch...

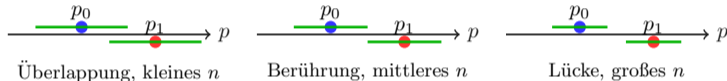


- rechte Grenze des 80%-PI unter  $p_0$ :  $f(n) = p_0 + \Phi^{-1}(1-\alpha)\sqrt{p_0(1-p_0)/n}$
- linke Grenze des 90%-PI unter  $p_1$ :  $g(n) = p_1 - \Phi^{-1}(\gamma)\sqrt{p_1(1-p_1)/n}$

# Planung des Stichprobenumfangs – mit Zollstöcken

Test  $H_0 : p \leq 0,5$  gegen  $H_1 : p > 0,5$  mit  $\alpha = 0,1$ ;  $\gamma = 0,9$ ;  $p_1 = 0,6$ .

Zwei Zollstöcke aber auch...



- rechte Grenze des 80%-PI unter  $p_0$ :  $f(n) = p_0 + \Phi^{-1}(1-\alpha)\sqrt{p_0(1-p_0)/n}$
- linke Grenze des 90%-PI unter  $p_1$ :  $g(n) = p_1 - \Phi^{-1}(\gamma)\sqrt{p_1(1-p_1)/n}$

GeoGebra Classic

Löse  $\left( p_0 + \text{InversNormal}(0, 1, 1-\alpha)\sqrt{p_0 \cdot \frac{1-p_0}{n}} = p_1 - \text{InversNormal}(0, 1, \gamma)\sqrt{p_1 \cdot \frac{1-p_1}{n}} \right)$

$\approx \{n = 160.94\}$

# IQB-Aufgabe 2022 - n gesucht mit Grafik

Gemeinsame Abituraufgabenpools der Länder

**Pool für das Jahr 2022**

Aufgaben für das Fach Mathematik

Kurzbeschreibung

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	WTR

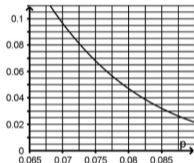
2 Ein Händler vermutet, dass die Fitnessarmbänder eines bestimmten Herstellers besonders häufig Fehler aufweisen. Um einen Anhaltspunkt für den Anteil der fehlerhaften Armbänder unter allen Fitnessarmbändern dieses Herstellers zu gewinnen, führt er einen Signifikanztest mit der Nullhypothese „Der Anteil der fehlerhaften Armbänder beträgt mindestens 7 %.“ durch.

Für diesen Test gilt:

- ♦ Die Nullhypothese wird abgelehnt, wenn höchstens vier Armbänder fehlerhaft sind.
- ♦ Der Abbildung kann die Wahrscheinlichkeit für den Fehler erster Art in Abhängigkeit vom Anteil  $p$  fehlerhafter Armbänder entnommen werden.

a Ermitteln Sie den Umfang der für den Test verwendeten Stichprobe.

b Geben Sie an, welche Überlegung den Händler dazu veranlasst haben könnte, die gewählte Nullhypothese der Alternative „Der Anteil der fehlerhaften Armbänder beträgt höchstens 7 %.“ vorzuziehen. Begründen Sie Ihre Angabe.



5

3

# IQB-Aufgabe 2022 - n gesucht mit Grafik

Gemeinsame Abituraufgabenpools der Länder

**Pool für das Jahr 2022**

Aufgaben für das Fach Mathematik

Kurzbeschreibung

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	WTR

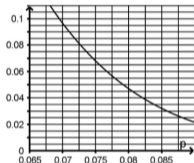
2 Ein Händler vermutet, dass die Fitnessarmbänder eines bestimmten Herstellers besonders häufig Fehler aufweisen. Um einen Anhaltspunkt für den Anteil der fehlerhaften Armbänder unter allen Fitnessarmbändern dieses Herstellers zu gewinnen, führt er einen Signifikanztest mit der Nullhypothese „Der Anteil der fehlerhaften Armbänder beträgt mindestens 7 %.“ durch.

Für diesen Test gilt:

- ♦ Die Nullhypothese wird abgelehnt, wenn höchstens vier Armbänder fehlerhaft sind.
- ♦ Der Abbildung kann die Wahrscheinlichkeit für den Fehler erster Art in Abhängigkeit vom Anteil  $p$  fehlerhafter Armbänder entnommen werden.

a Ermitteln Sie den Umfang der für den Test verwendeten Stichprobe.

b Geben Sie an, welche Überlegung den Händler dazu veranlasst haben könnte, die gewählte Nullhypothese der Alternative „Der Anteil der fehlerhaften Armbänder beträgt höchstens 7 %.“ vorzuziehen. Begründen Sie Ihre Angabe.



5

3

2 a Die Abbildung liefert für  $p = 0,07$  einen Wert zwischen 0,095 und 0,1.

5

Y: Anzahl der fehlerhaften Armbänder

$$P_{0,07}^{112}(Y \leq 4) \approx 0,101, P_{0,07}^{113}(Y \leq 4) \approx 0,097, P_{0,07}^{114}(Y \leq 4) \approx 0,093$$

Die Stichprobe hatte also einen Umfang von 113 Armbändern.

b Der Händler möchte vermeiden, irrtümlich von einem zu geringen Anteil fehlerhafter Armbänder auszugehen, da dies beispielsweise zu unerwartet vielen Reklamationen durch Kunden führen könnte. Bei der gewählten Nullhypothese beträgt das Risiko für diesen Irrtum höchstens 9,7 % und ist damit gering. Die Wahrscheinlichkeit dafür, irrtümlich von einem zu großen Anteil auszugehen, kann dagegen wesentlich größer sein.

3

# IQB-Aufgabe 2022 - n gesucht mit Grafik

Gemeinsame Abituraufgabepools der Länder

**Pool für das Jahr 2022**

Aufgaben für das Fach Mathematik

Kurzbeschreibung

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	WTR

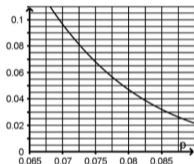
2 Ein Händler vermutet, dass die Fitnessarmbänder eines bestimmten Herstellers besonders häufig Fehler aufweisen. Um einen Anhaltspunkt für den Anteil der fehlerhaften Aribänder unter allen Fitnessarmbändern dieses Herstellers zu gewinnen, führt er einen Signifikanztest mit der Nullhypothese „Der Anteil der fehlerhaften Aribänder beträgt mindestens 7 %.“ durch.

Für diesen Test gilt:

- Die Nullhypothese wird abgelehnt, wenn höchstens vier Aribänder fehlerhaft sind.
- Der Abbildung kann die Wahrscheinlichkeit für den Fehler erster Art in Abhängigkeit vom Anteil  $p$  fehlerhafter Aribänder entnommen werden.

a Ermitteln Sie den Umfang der für den Test verwendeten Stichprobe.

b Geben Sie an, welche Überlegung den Händler dazu veranlasst haben könnte, die gewählte Nullhypothese der Alternative „Der Anteil der fehlerhaften Aribänder beträgt höchstens 7 %.“ vorzuziehen. Begründen Sie Ihre Angabe.



5

3

2 a Die Abbildung liefert für  $p = 0,07$  einen Wert zwischen 0,095 und 0,1.

5

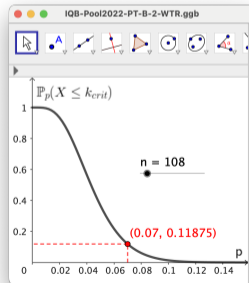
Y: Anzahl der fehlerhaften Aribänder

$$P_{0,07}^{112}(Y \leq 4) \approx 0,101, P_{0,07}^{113}(Y \leq 4) \approx 0,097, P_{0,07}^{114}(Y \leq 4) \approx 0,093$$

Die Stichprobe hatte also einen Umfang von 113 Aribändern.

b Der Händler möchte vermeiden, irrtümlich von einem zu geringen Anteil fehlerhafter Aribänder auszugehen, da dies beispielsweise zu unerwartet vielen Reklamationen durch Kunden führen könnte. Bei der gewählten Nullhypothese beträgt das Risiko für diesen Irrtum höchstens 9,7 % und ist damit gering. Die Wahrscheinlichkeit dafür, irrtümlich von einem zu großen Anteil auszugehen, kann dagegen wesentlich größer sein.

3



# IQB-Aufgabe 2022 - n gesucht mit Grafik

Gemeinsame Abituraufgabepools der Länder

**Pool für das Jahr 2022**

Aufgaben für das Fach Mathematik

Kurzbeschreibung

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	WTR

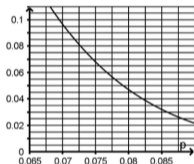
2 Ein Händler vermutet, dass die Fitnessarmbänder eines bestimmten Herstellers besonders häufig Fehler aufweisen. Um einen Anhaltspunkt für den Anteil der fehlerhaften Aribänder unter allen Fitnessarmbändern dieses Herstellers zu gewinnen, führt er einen Signifikanztest mit der Nullhypothese „Der Anteil der fehlerhaften Aribänder beträgt mindestens 7 %.“ durch.

Für diesen Test gilt:

- Die Nullhypothese wird abgelehnt, wenn höchstens vier Aribänder fehlerhaft sind.
- Der Abbildung kann die Wahrscheinlichkeit für den Fehler erster Art in Abhängigkeit vom Anteil  $p$  fehlerhafter Aribänder entnommen werden.

a Ermitteln Sie den Umfang der für den Test verwendeten Stichprobe.

b Geben Sie an, welche Überlegung den Händler dazu veranlasst haben könnte, die gewählte Nullhypothese der Alternative „Der Anteil der fehlerhaften Aribänder beträgt höchstens 7 %.“ vorzuziehen. Begründen Sie Ihre Angabe.



5

3

2 a Die Abbildung liefert für  $p = 0,07$  einen Wert zwischen 0,095 und 0,1.

5

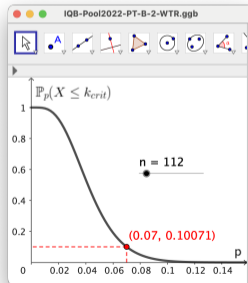
Y: Anzahl der fehlerhaften Aribänder

$$P_{0,07}^{112}(Y \leq 4) \approx 0,101, P_{0,07}^{113}(Y \leq 4) \approx 0,097, P_{0,07}^{114}(Y \leq 4) \approx 0,093$$

Die Stichprobe hatte also einen Umfang von 113 Aribändern.

b Der Händler möchte vermeiden, irrtümlich von einem zu geringen Anteil fehlerhafter Aribänder auszugehen, da dies beispielsweise zu unerwartet vielen Reklamationen durch Kunden führen könnte. Bei der gewählten Nullhypothese beträgt das Risiko für diesen Irrtum höchstens 9,7 % und ist damit gering. Die Wahrscheinlichkeit dafür, irrtümlich von einem zu großen Anteil auszugehen, kann dagegen wesentlich größer sein.

3



# IQB-Aufgabe 2022 - n gesucht mit Grafik

Gemeinsame Abituraufgabenpools der Länder

**Pool für das Jahr 2022**

Aufgaben für das Fach Mathematik

Kurzbeschreibung

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	WTR

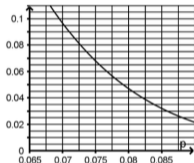
2 Ein Händler vermutet, dass die Fitnessarmbänder eines bestimmten Herstellers besonders häufig Fehler aufweisen. Um einen Anhaltspunkt für den Anteil der fehlerhaften Armbänder unter allen Fitnessarmbändern dieses Herstellers zu gewinnen, führt er einen Signifikanztest mit der Nullhypothese „Der Anteil der fehlerhaften Armbänder beträgt mindestens 7 %.“ durch.

Für diesen Test gilt:

- Die Nullhypothese wird abgelehnt, wenn höchstens vier Armbänder fehlerhaft sind.
- Der Abbildung kann die Wahrscheinlichkeit für den Fehler erster Art in Abhängigkeit vom Anteil  $p$  fehlerhafter Armbänder entnommen werden.

a Ermitteln Sie den Umfang der für den Test verwendeten Stichprobe.

b Geben Sie an, welche Überlegung den Händler dazu veranlasst haben könnte, die gewählte Nullhypothese der Alternative „Der Anteil der fehlerhaften Armbänder beträgt höchstens 7 %.“ vorzuziehen. Begründen Sie Ihre Angabe.



5

3

2 a Die Abbildung liefert für  $p = 0,07$  einen Wert zwischen 0,095 und 0,1.

5

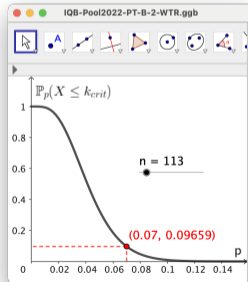
Y: Anzahl der fehlerhaften Armbänder

$$P_{0,07}^{112}(Y \leq 4) \approx 0,101, P_{0,07}^{113}(Y \leq 4) \approx 0,097, P_{0,07}^{114}(Y \leq 4) \approx 0,093$$

Die Stichprobe hatte also einen Umfang von 113 Armbändern.

b Der Händler möchte vermeiden, irrtümlich von einem zu geringen Anteil fehlerhafter Armbänder auszugehen, da dies beispielsweise zu unerwartet vielen Reklamationen durch Kunden führen könnte. Bei der gewählten Nullhypothese beträgt das Risiko für diesen Irrtum höchstens 9,7 % und ist damit gering. Die Wahrscheinlichkeit dafür, irrtümlich von einem zu großen Anteil auszugehen, kann dagegen wesentlich größer sein.

3



# IQB-Aufgabe 2022 - n gesucht mit Grafik

Gemeinsame Abituraufgabepools der Länder

**Pool für das Jahr 2022**

Aufgaben für das Fach Mathematik

Kurzbeschreibung

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	WTR

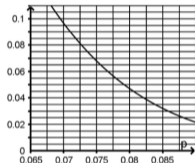
2 Ein Händler vermutet, dass die Fitnessarmbänder eines bestimmten Herstellers besonders häufig Fehler aufweisen. Um einen Anhaltspunkt für den Anteil der fehlerhaften Armbänder unter allen Fitnessarmbändern dieses Herstellers zu gewinnen, führt er einen Signifikanztest mit der Nullhypothese „Der Anteil der fehlerhaften Armbänder beträgt mindestens 7 %.“ durch.

Für diesen Test gilt:

- Die Nullhypothese wird abgelehnt, wenn höchstens vier Armbänder fehlerhaft sind.
- Der Abbildung kann die Wahrscheinlichkeit für den Fehler erster Art in Abhängigkeit vom Anteil  $p$  fehlerhafter Armbänder entnommen werden.

a Ermitteln Sie den Umfang der für den Test verwendeten Stichprobe.

b Geben Sie an, welche Überlegung den Händler dazu veranlasst haben könnte, die gewählte Nullhypothese der Alternative „Der Anteil der fehlerhaften Armbänder beträgt höchstens 7 %.“ vorzuziehen. Begründen Sie Ihre Angabe.



5

3

2 a Die Abbildung liefert für  $p = 0,07$  einen Wert zwischen 0,095 und 0,1.

5

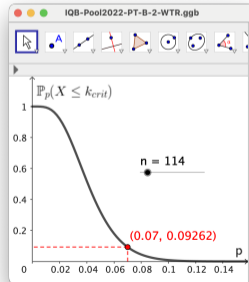
Y: Anzahl der fehlerhaften Armbänder

$$P_{0,07}^{112}(Y \leq 4) \approx 0,101, P_{0,07}^{113}(Y \leq 4) \approx 0,097, P_{0,07}^{114}(Y \leq 4) \approx 0,093$$

Die Stichprobe hatte also einen Umfang von 113 Armbändern.

b Der Händler möchte vermeiden, irrtümlich von einem zu geringen Anteil fehlerhafter Armbänder auszugehen, da dies beispielsweise zu unerwartet vielen Reklamationen durch Kunden führen könnte. Bei der gewählten Nullhypothese beträgt das Risiko für diesen Irrtum höchstens 9,7 % und ist damit gering. Die Wahrscheinlichkeit dafür, irrtümlich von einem zu großen Anteil auszugehen, kann dagegen wesentlich größer sein.

3



# IQB-Aufgabe 2022 - n gesucht mit Grafik

Gemeinsame Abituraufgabepools der Länder

**Pool für das Jahr 2022**

Aufgaben für das Fach Mathematik

Kurzbeschreibung

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	WTR

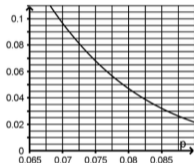
2 Ein Händler vermutet, dass die Fitnessarmbänder eines bestimmten Herstellers besonders häufig Fehler aufweisen. Um einen Anhaltspunkt für den Anteil der fehlerhaften Armbänder unter allen Fitnessarmbändern dieses Herstellers zu gewinnen, führt er einen Signifikanztest mit der Nullhypothese „Der Anteil der fehlerhaften Armbänder beträgt mindestens 7 %.“ durch.

Für diesen Test gilt:

- Die Nullhypothese wird abgelehnt, wenn höchstens vier Armbänder fehlerhaft sind.
- Der Abbildung kann die Wahrscheinlichkeit für den Fehler erster Art in Abhängigkeit vom Anteil  $p$  fehlerhafter Armbänder entnommen werden.

a Ermitteln Sie den Umfang der für den Test verwendeten Stichprobe.

b Geben Sie an, welche Überlegung den Händler dazu veranlasst haben könnte, die gewählte Nullhypothese der Alternative „Der Anteil der fehlerhaften Armbänder beträgt höchstens 7 %.“ vorzuziehen. Begründen Sie Ihre Angabe.



5

3

2 a Die Abbildung liefert für  $p = 0,07$  einen Wert zwischen 0,095 und 0,1.

5

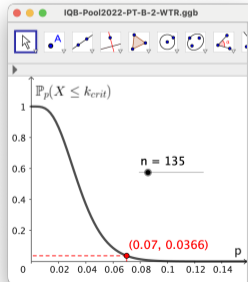
Y: Anzahl der fehlerhaften Armbänder

$$P_{0,07}^{112}(Y \leq 4) \approx 0,101, P_{0,07}^{113}(Y \leq 4) \approx 0,097, P_{0,07}^{114}(Y \leq 4) \approx 0,093$$

Die Stichprobe hatte also einen Umfang von 113 Armbändern.

b Der Händler möchte vermeiden, irrtümlich von einem zu geringen Anteil fehlerhafter Armbänder auszugehen, da dies beispielsweise zu unerwartet vielen Reklamationen durch Kunden führen könnte. Bei der gewählten Nullhypothese beträgt das Risiko für diesen Irrtum höchstens 9,7 % und ist damit gering. Die Wahrscheinlichkeit dafür, irrtümlich von einem zu großen Anteil auszugehen, kann dagegen wesentlich größer sein.

3



# IQB-Aufgabe 2022 - n gesucht mit Grafik

Gemeinsame Abituraufgabenpools der Länder

**Pool für das Jahr 2022**

Aufgaben für das Fach Mathematik

Kurzbeschreibung

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	WTR

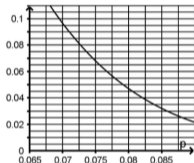
2 Ein Händler vermutet, dass die Fitnessarmbänder eines bestimmten Herstellers besonders häufig Fehler aufweisen. Um einen Anhaltspunkt für den Anteil der fehlerhaften Aribänder unter allen Fitnessarmbändern dieses Herstellers zu gewinnen, führt er einen Signifikanztest mit der Nullhypothese „Der Anteil der fehlerhaften Aribänder beträgt mindestens 7 %.“ durch.

Für diesen Test gilt:

- Die Nullhypothese wird abgelehnt, wenn höchstens vier Aribänder fehlerhaft sind.
- Der Abbildung kann die Wahrscheinlichkeit für den Fehler erster Art in Abhängigkeit vom Anteil  $p$  fehlerhafter Aribänder entnommen werden.

a Ermitteln Sie den Umfang der für den Test verwendeten Stichprobe.

b Geben Sie an, welche Überlegung den Händler dazu veranlasst haben könnte, die gewählte Nullhypothese der Alternative „Der Anteil der fehlerhaften Aribänder beträgt höchstens 7 %.“ vorzuziehen. Begründen Sie Ihre Angabe.



5

3

2 a Die Abbildung liefert für  $p = 0,07$  einen Wert zwischen 0,095 und 0,1.

5

Y: Anzahl der fehlerhaften Aribänder

$$P_{0,07}^{112}(Y \leq 4) \approx 0,101, P_{0,07}^{113}(Y \leq 4) \approx 0,097, P_{0,07}^{114}(Y \leq 4) \approx 0,093$$

Die Stichprobe hatte also einen Umfang von 113 Aribändern.

b Der Händler möchte vermeiden, irrtümlich von einem zu geringen Anteil fehlerhafter Aribänder auszugehen, da dies beispielsweise zu unerwartet vielen Reklamationen durch Kunden führen könnte. Bei der gewählten Nullhypothese beträgt das Risiko für diesen Irrtum höchstens 9,7 % und ist damit gering. Die Wahrscheinlichkeit dafür, irrtümlich von einem zu großen Anteil auszugehen, kann dagegen wesentlich größer sein.

3

$$K = [0, \dots, 4]$$

$$0.0965$$



$$p = 0.07$$

$$\alpha = 0.193$$

$$n = 135$$

# IQB-Aufgabe 2022 - n gesucht mit Grafik

Gemeinsame Abituraufgabenpools der Länder

**Pool für das Jahr 2022**

Aufgaben für das Fach Mathematik

Kurzbeschreibung

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	WTR

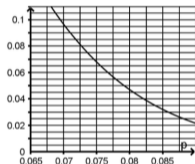
2 Ein Händler vermutet, dass die Fitnessarmbänder eines bestimmten Herstellers besonders häufig Fehler aufweisen. Um einen Anhaltspunkt für den Anteil der fehlerhaften Armbänder unter allen Fitnessarmbändern dieses Herstellers zu gewinnen, führt er einen Signifikanztest mit der Nullhypothese „Der Anteil der fehlerhaften Armbänder beträgt mindestens 7 %.“ durch.

Für diesen Test gilt:

- Die Nullhypothese wird abgelehnt, wenn höchstens vier Armbänder fehlerhaft sind.
- Der Abbildung kann die Wahrscheinlichkeit für den Fehler erster Art in Abhängigkeit vom Anteil  $p$  fehlerhafter Armbänder entnommen werden.

a Ermitteln Sie den Umfang der für den Test verwendeten Stichprobe.

b Geben Sie an, welche Überlegung den Händler dazu veranlasst haben könnte, die gewählte Nullhypothese der Alternative „Der Anteil der fehlerhaften Armbänder beträgt höchstens 7 %.“ vorzuziehen. Begründen Sie Ihre Angabe.



5

3

2 a Die Abbildung liefert für  $p = 0,07$  einen Wert zwischen 0,095 und 0,1.

5

Y: Anzahl der fehlerhaften Armbänder

$$P_{0,07}^{112}(Y \leq 4) \approx 0,101, P_{0,07}^{113}(Y \leq 4) \approx 0,097, P_{0,07}^{114}(Y \leq 4) \approx 0,093$$

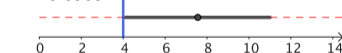
Die Stichprobe hatte also einen Umfang von 113 Armbändern.

b Der Händler möchte vermeiden, irrtümlich von einem zu geringen Anteil fehlerhafter Armbänder auszugehen, da dies beispielsweise zu unerwartet vielen Reklamationen durch Kunden führen könnte. Bei der gewählten Nullhypothese beträgt das Risiko für diesen Irrtum höchstens 9,7 % und ist damit gering. Die Wahrscheinlichkeit dafür, irrtümlich von einem zu großen Anteil auszugehen, kann dagegen wesentlich größer sein.

3

$$K = [0, \dots, 4]$$

$$0.0965$$



$$p = 0.07$$

$$\alpha = 0.193$$

$$n = 108$$

# IQB-Aufgabe 2023 - argumentieren mit Grafik

Gemeinsame Abituraufgabenpools der Länder **Kurzbeschreibung**

## Pool für das Jahr 2023

Aufgaben für das Fach Mathematik

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	MMS/WTR

**b** Es soll davon ausgegangen werden, dass 80 000 Personen an dem Spiel teilnehmen werden. Der Erwartungswert der Anzahl der Personen mit zwei Strandkörben wird mit  $\mu$  bezeichnet. Ermitteln Sie den kleinsten möglichen ganzzahligen Wert von  $c$ , für den die Anzahl der Personen mit zwei Strandkörben mit einer Wahrscheinlichkeit von mindestens 80 % im Intervall  $[\mu - c; \mu + c]$  liegt.

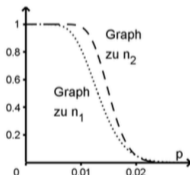
4

**c** Die Wahrscheinlichkeit dafür, dass eine zufällig ausgewählte Person nach der Teilnahme am Gewinnspiel eine Reise bei dem Reiseunternehmen bucht, wird mit  $p$  bezeichnet. Für das Unternehmen wäre eine Verlängerung des Gewinnspiels für  $p \geq 2\%$  mit Vorteilen verbunden, für  $p < 2\%$  dagegen mit finanziellen Verlusten. Die Nullhypothese „ $p$  beträgt mindestens 2 %.“ soll auf einem Signifikanzniveau von 5 % getestet werden. Je größer der Umfang der Stichprobe gewählt wird, desto teurer ist die Durchführung des Tests.

5

Das Gewinnspiel soll nur dann nicht verlängert werden, wenn die Nullhypothese aufgrund des Testergebnisses abgelehnt werden müsste.

Die Abbildung stellt für zwei mögliche Stichprobenumfänge  $n_1$  und  $n_2$  mit  $n_1 < n_2$  in Abhängigkeit von  $p$  die Wahrscheinlichkeit dafür dar, dass das Ergebnis des Tests im Ablehnungsbereich liegt. Geben Sie an, unter welcher Bedingung sich der größere Stichprobenumfang für das Unternehmen lohnen könnte, und begründen Sie Ihre Angabe unter Verwendung der Abbildung sowie des Fehlers zweiter Art.



# IQB-Aufgabe 2023 - argumentieren mit Grafik

Gemeinsame Abituraufgabenpools der Länder **Kurzbeschreibung**

**Pool für das Jahr 2023**

Aufgaben für das Fach Mathematik

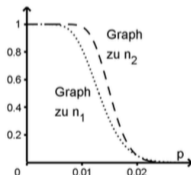
Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	MMS/WTR

**b** Es soll davon ausgegangen werden, dass 80 000 Personen an dem Spiel teilnehmen werden. Der Erwartungswert der Anzahl der Personen mit zwei Strandkörben wird mit  $\mu$  bezeichnet. Ermitteln Sie den kleinsten möglichen ganzzahligen Wert von  $c$ , für den die Anzahl der Personen mit zwei Strandkörben mit einer Wahrscheinlichkeit von mindestens 80 % im Intervall  $[\mu - c; \mu + c]$  liegt.

**c** Die Wahrscheinlichkeit dafür, dass eine zufällig ausgewählte Person nach der Teilnahme am Gewinnspiel eine Reise bei dem Reiseunternehmen bucht, wird mit  $p$  bezeichnet. Für das Unternehmen wäre eine Verlängerung des Gewinnspiels für  $p \geq 2\%$  mit Vorteilen verbunden, für  $p < 2\%$  dagegen mit finanziellen Verlusten. Die Nullhypothese „ $p$  beträgt mindestens 2 %.“ soll auf einem Signifikanzniveau von 5 % getestet werden. Je größer der Umfang der Stichprobe gewählt wird, desto teurer ist die Durchführung des Tests.

Das Gewinnspiel soll nur dann nicht verlängert werden, wenn die Nullhypothese aufgrund des Testergebnisses abgelehnt werden müsste.

Die Abbildung stellt für zwei mögliche Stichprobenumfänge  $n_1$  und  $n_2$  mit  $n_1 < n_2$  in Abhängigkeit von  $p$  die Wahrscheinlichkeit dafür dar, dass das Ergebnis des Tests im Ablehnungsbereich liegt. Geben Sie an, unter welcher Bedingung sich der größere Stichprobenumfang für das Unternehmen lohnen könnte, und begründen Sie Ihre Angabe unter Verwendung der Abbildung sowie des Fehlers zweiter Art.



**c** Der Fehler zweiter Art kann auftreten, wenn  $p < 0,02$  gilt, beispielsweise für  $p = 0,014$ . Für diesen Wert von  $p$  liegt der Graph zu  $n_2$  oberhalb des Graphen zu  $n_1$ . Damit ist die Wahrscheinlichkeit für den Fehler zweiter Art, d. h. die Wahrscheinlichkeit dafür, das Gewinnspiel irrtümlich zu verlängern, für  $n_2$  kleiner als für  $n_1$ . Bei größerem Stichprobenumfang ist für das Unternehmen also das Risiko, durch eine Verlängerung des Gewinnspiels finanzielle Verluste zu erleiden, geringer. Gleichzeitig sind aber die Kosten für die Durchführung des Tests höher. Damit könnte sich der größere Stichprobenumfang für das Unternehmen lohnen, wenn die zusätzlichen Kosten für den größeren Stichprobenumfang verhältnismäßig gering wären.

5

4

5

# IQB-Aufgabe 2023 - argumentieren mit Grafik

Gemeinsame Abituraufgabenpools der Länder **Kurzbeschreibung**

**Pool für das Jahr 2023**

Aufgaben für das Fach Mathematik

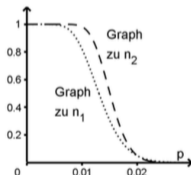
Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	MMS/WTR

**b** Es soll davon ausgegangen werden, dass 80 000 Personen an dem Spiel teilnehmen werden. Der Erwartungswert der Anzahl der Personen mit zwei Strandkörben wird mit  $\mu$  bezeichnet. Ermitteln Sie den kleinsten möglichen ganzzahligen Wert von  $c$ , für den die Anzahl der Personen mit zwei Strandkörben mit einer Wahrscheinlichkeit von mindestens 80 % im Intervall  $[\mu - c; \mu + c]$  liegt.

**c** Die Wahrscheinlichkeit dafür, dass eine zufällig ausgewählte Person nach der Teilnahme am Gewinnspiel eine Reise bei dem Reiseunternehmen bucht, wird mit  $p$  bezeichnet. Für das Unternehmen wäre eine Verlängerung des Gewinnspiels für  $p \geq 2\%$  mit Vorteilen verbunden, für  $p < 2\%$  dagegen mit finanziellen Verlusten. Die Nullhypothese „ $p$  beträgt mindestens 2 %.“ soll auf einem Signifikanzniveau von 5 % getestet werden. Je größer der Umfang der Stichprobe gewählt wird,

desto teurer ist die Durchführung des Tests. Das Gewinnspiel soll nur dann nicht verlängert werden, wenn die Nullhypothese aufgrund des Testergebnisses abgelehnt werden müsste.

Die Abbildung stellt für zwei mögliche Stichprobenumfänge  $n_1$  und  $n_2$  mit  $n_1 < n_2$  in Abhängigkeit von  $p$  die Wahrscheinlichkeit dafür dar, dass das Ergebnis des Tests im Ablehnungsbereich liegt. Geben Sie an, unter welcher Bedingung sich der größere Stichprobenumfang für das Unternehmen lohnen könnte, und begründen Sie Ihre Angabe unter Verwendung der Abbildung sowie des Fehlers zweiter Art.

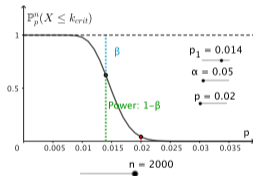


4

5

**c** Der Fehler zweiter Art kann auftreten, wenn  $p < 0,02$  gilt, beispielsweise für  $p = 0,014$ . Für diesen Wert von  $p$  liegt der Graph zu  $n_2$  oberhalb des Graphen zu  $n_1$ . Damit ist die Wahrscheinlichkeit für den Fehler zweiter Art, d. h. die Wahrscheinlichkeit dafür, das Gewinnspiel irrtümlich zu verlängern, für  $n_2$  kleiner als für  $n_1$ . Bei größerem Stichprobenumfang ist für das Unternehmen also das Risiko, durch eine Verlängerung des Gewinnspiels finanzielle Verluste zu erleiden, geringer. Gleichzeitig sind aber die Kosten für die Durchführung des Tests höher. Damit könnte sich der größere Stichprobenumfang für das Unternehmen lohnen, wenn die zusätzlichen Kosten für den größeren Stichprobenumfang verhältnismäßig gering wären.

5



# IQB-Aufgabe 2023 - argumentieren mit Grafik

Gemeinsame Abituraufgabenpools der Länder **Kurzbeschreibung**

**Pool für das Jahr 2023**

Aufgaben für das Fach Mathematik

Anforderungsniveau	Prüfungsteil	Sachgebiet <sup>1</sup>	digitales Hilfsmittel
erhöht	B	Stochastik	MMS/WTR

**b** Es soll davon ausgegangen werden, dass 80 000 Personen an dem Spiel teilnehmen werden. Der Erwartungswert der Anzahl der Personen mit zwei Strandkörben wird mit  $\mu$  bezeichnet. Ermitteln Sie den kleinsten möglichen ganzzahligen Wert von  $c$ , für den die Anzahl der Personen mit zwei Strandkörben mit einer Wahrscheinlichkeit von mindestens 80 % im Intervall  $[\mu - c; \mu + c]$  liegt.

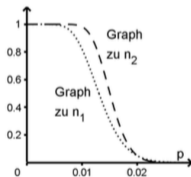
4

**c** Die Wahrscheinlichkeit dafür, dass eine zufällig ausgewählte Person nach der Teilnahme am Gewinnspiel eine Reise bei dem Reiseunternehmen bucht, wird mit  $p$  bezeichnet. Für das Unternehmen wäre eine Verlängerung des Gewinnspiels für  $p \geq 2\%$  mit Vorteilen verbunden, für  $p < 2\%$  dagegen mit finanziellen Verlusten. Die Nullhypothese „ $p$  beträgt mindestens 2 %.“ soll auf einem Signifikanzniveau von 5 % getestet werden. Je größer der Umfang der Stichprobe gewählt wird, desto teurer ist die Durchführung des Tests.

5

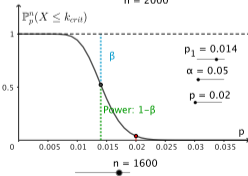
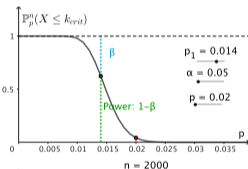
Das Gewinnspiel soll nur dann nicht verlängert werden, wenn die Nullhypothese aufgrund des Testergebnisses abgelehnt werden müsste.

Die Abbildung stellt für zwei mögliche Stichprobenumfänge  $n_1$  und  $n_2$  mit  $n_1 < n_2$  in Abhängigkeit von  $p$  die Wahrscheinlichkeit dafür dar, dass das Ergebnis des Tests im Ablehnungsbereich liegt. Geben Sie an, unter welcher Bedingung sich der größere Stichprobenumfang für das Unternehmen lohnen könnte, und begründen Sie Ihre Angabe unter Verwendung der Abbildung sowie des Fehlers zweiter Art.



**c** Der Fehler zweiter Art kann auftreten, wenn  $p < 0,02$  gilt, beispielsweise für  $p = 0,014$ . Für diesen Wert von  $p$  liegt der Graph zu  $n_2$  oberhalb des Graphen zu  $n_1$ . Damit ist die Wahrscheinlichkeit für den Fehler zweiter Art, d. h. die Wahrscheinlichkeit dafür, das Gewinnspiel irrtümlich zu verlängern, für  $n_2$  kleiner als für  $n_1$ . Bei größerem Stichprobenumfang ist für das Unternehmen also das Risiko, durch eine Verlängerung des Gewinnspiels finanzielle Verluste zu erleiden, geringer. Gleichzeitig sind aber die Kosten für die Durchführung des Tests höher. Damit könnte sich der größere Stichprobenumfang für das Unternehmen lohnen, wenn die zusätzlichen Kosten für den größeren Stichprobenumfang verhältnismäßig gering wären.

5



Konfidenzintervalle - wie werden sie  
berechnet und wo steckt der Zufall?

---

Ergebnis der Fortbildung

## Definition

Ein **Konfidenzintervall** zu einer beobachteten relativen Häufigkeit  $h$  besteht aus allen denjenigen Wahrscheinlichkeiten ( $p$ -Werten), in deren Prognoseintervall die beobachtete relative Häufigkeit  $h$  liegt.

# Konfidenzintervall - der Versuch einer Definition

## Ergebnis der Fortbildung

### Definition

Ein **Konfidenzintervall** zu einer beobachteten relativen Häufigkeit  $h$  besteht aus allen denjenigen Wahrscheinlichkeiten ( $p$ -Werten), in deren Prognoseintervall die beobachtete relative Häufigkeit  $h$  liegt.

## Ergebnis am Schreibtisch

### Definition

Ein **Konfidenzintervall** zu einer beobachteten relativen Häufigkeit  $h$  ist die Menge aller Wahrscheinlichkeiten  $p$  im Binomialmodell  $\text{Bin}(n, p)$ , in deren Prognoseintervall die beobachtete relative Häufigkeit  $h$  liegt.

## Definition

Ein Konfidenzintervall zu einer beobachteten relativen Häufigkeit  $h$  ist die Menge aller Wahrscheinlichkeiten  $p$  im Binomialmodell  $\text{Bin}(n, p)$ , in deren Prognoseintervall die beobachtete relative Häufigkeit  $h$  liegt.

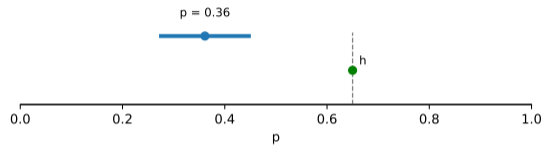
## Definition

Ein Konfidenzintervall zu einer beobachteten relativen Häufigkeit  $h$  ist die Menge aller Wahrscheinlichkeiten  $p$  im Binomialmodell  $\text{Bin}(n, p)$ , in deren Prognoseintervall die beobachtete relative Häufigkeit  $h$  liegt.

Wo ist hier der Zufall?

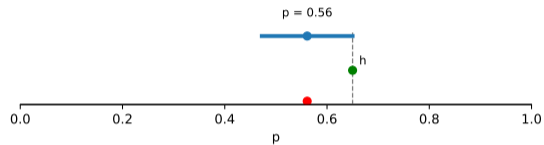
# Konfidenzintervalle freirubbeln - und berechnen

Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$



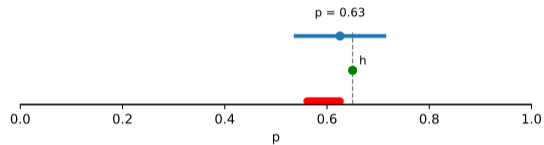
# Konfidenzintervalle freirubbeln - und berechnen

Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$



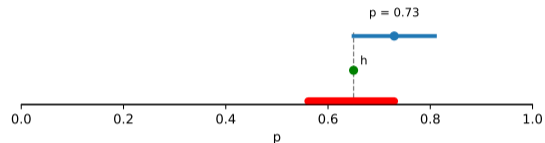
# Konfidenzintervalle freirubbeln - und berechnen

Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$



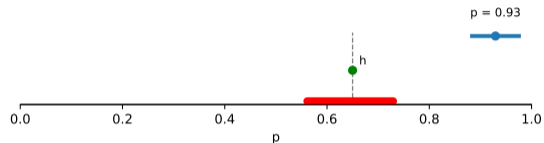
# Konfidenzintervalle freirubbeln - und berechnen

Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$



# Konfidenzintervalle freirubbeln - und berechnen

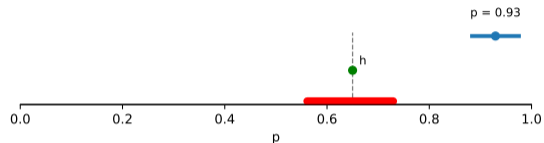
Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$



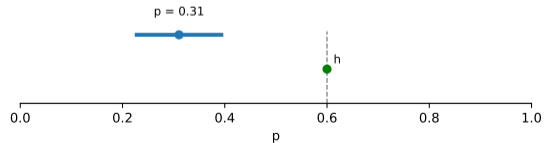
Stichprobenziehung:  $n = 120$ ;  $x = 72$ ;  $h = 0,60$

# Konfidenzintervalle freirubbeln - und berechnen

Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$

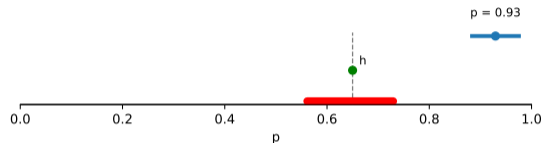


Stichprobenziehung:  $n = 120$ ;  $x = 72$ ;  $h = 0,60$

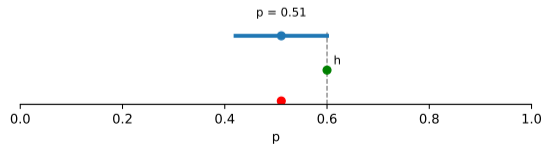


# Konfidenzintervalle freirubbeln - und berechnen

Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$

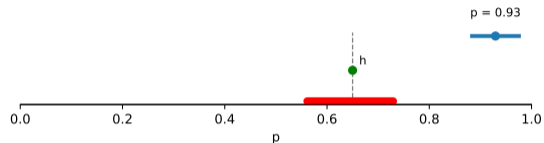


Stichprobenziehung:  $n = 120$ ;  $x = 72$ ;  $h = 0,60$

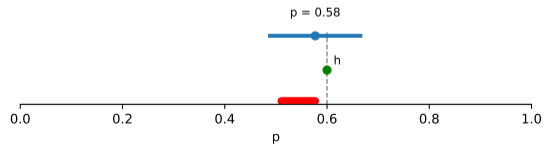


# Konfidenzintervalle freirubbeln - und berechnen

Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$

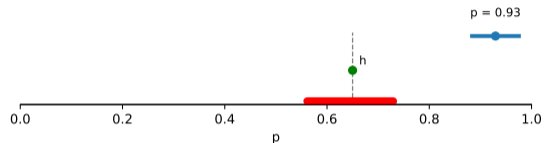


Stichprobenziehung:  $n = 120$ ;  $x = 72$ ;  $h = 0,60$

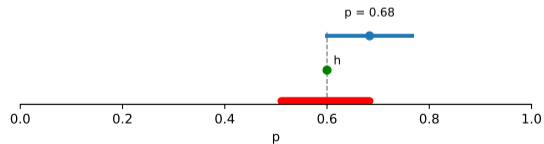


# Konfidenzintervalle freirubbeln - und berechnen

Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$

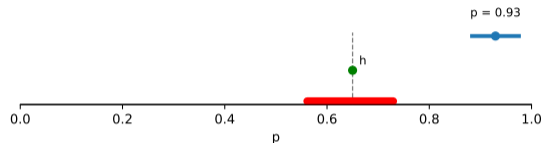


Stichprobenziehung:  $n = 120$ ;  $x = 72$ ;  $h = 0,60$

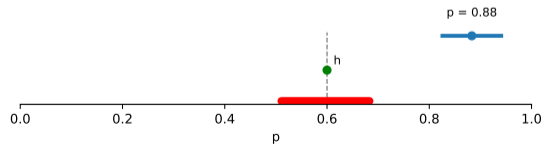


# Konfidenzintervalle freirubbeln - und berechnen

Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$

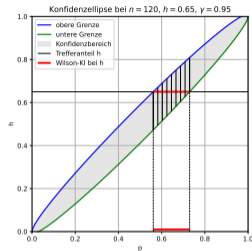
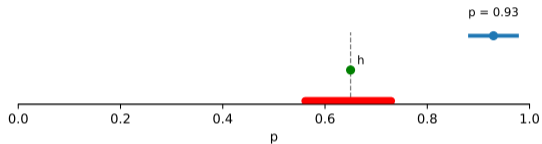


Stichprobenziehung:  $n = 120$ ;  $x = 72$ ;  $h = 0,60$

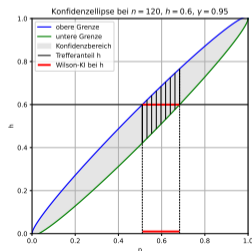
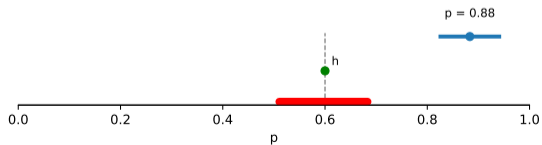


# Konfidenzintervalle freirubbeln - und berechnen

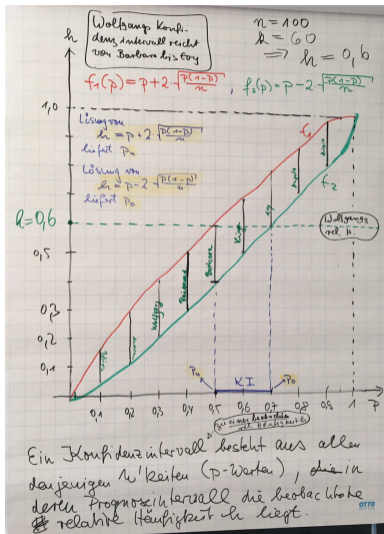
Stichprobenziehung:  $n = 120$ ;  $x = 78$ ;  $h = 0,65$



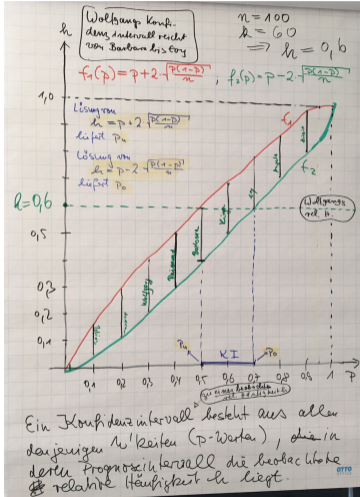
Stichprobenziehung:  $n = 120$ ;  $x = 72$ ;  $h = 0,60$



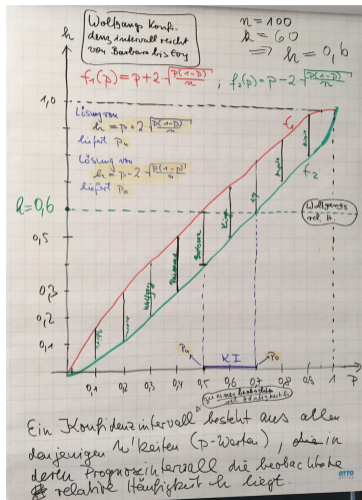
# Konfidenzellipse (do it yourself) - während einer denkwürdigen Fortbildung



# Die Konfidenzellipse (do it yourself) - während einer denkwürdigen Fortbildung



# Die Konfidenzellipse als Denkraum



## Mehr als eine Grafik

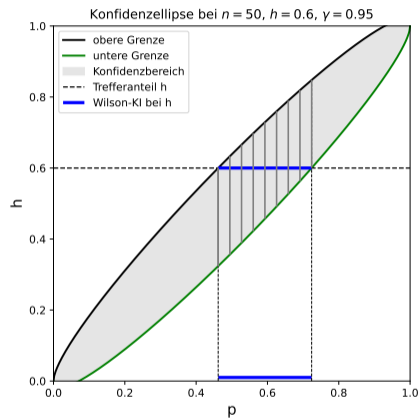
Die Konfidenzellipse ist kein Poster und kein Rechenhilfsmittel, sondern ein gemeinsamer **Denkraum**.

- Sie macht Abhängigkeiten sichtbar.
- Sie erlaubt Perspektivwechsel (Schnitte).
- Sie verbindet Geometrie und Stochastik.
- Sie trägt Prognose- und Konfidenzintervalle.

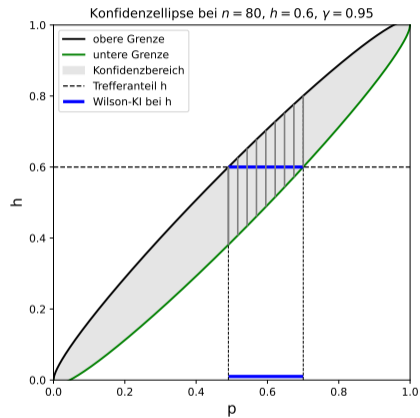
*Nicht anschauen – benutzen.*

*Immer wieder – im Klassenraum – gemeinsam.*

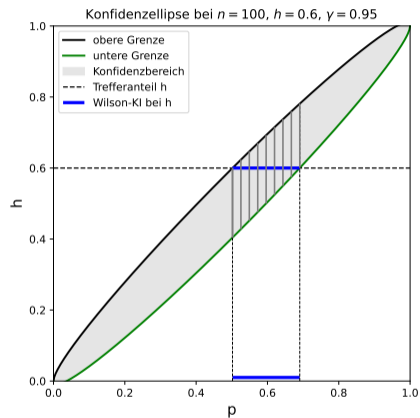
# Konfidenzellipsen – ein Überblick, falls $n$ verändert wird



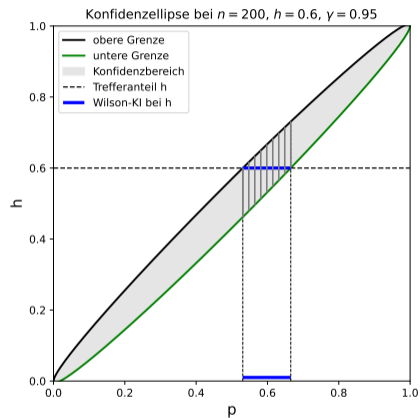
# Konfidenzellipsen – ein Überblick, falls $n$ verändert wird



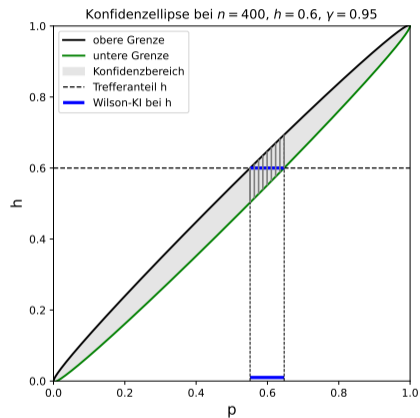
# Konfidenzellipsen – ein Überblick, falls $n$ verändert wird



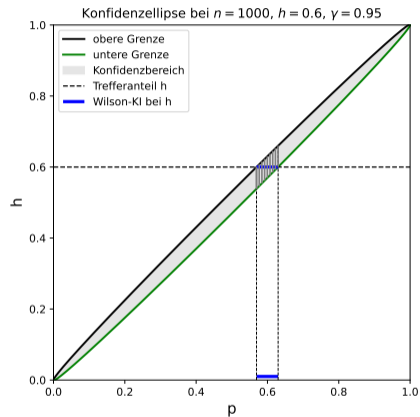
# Konfidenzellipsen – ein Überblick, falls $n$ verändert wird



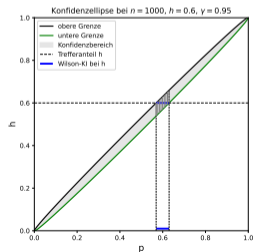
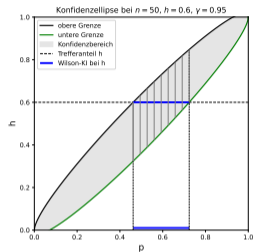
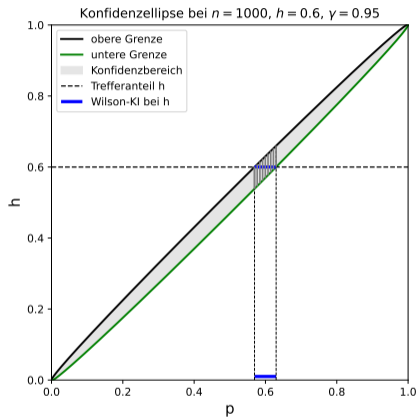
# Konfidenzellipsen – ein Überblick, falls $n$ verändert wird



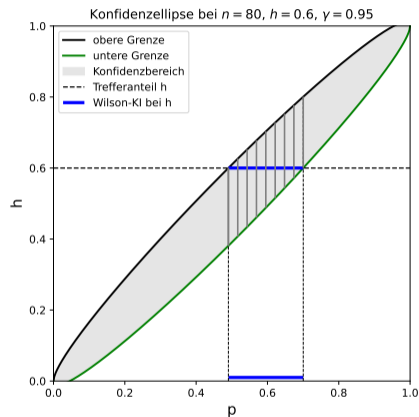
# Konfidenzellipsen – ein Überblick, falls $n$ verändert wird



# Konfidenzellipsen – ein Überblick, falls $n$ verändert wird



# Konfidenzellipsen – ein Überblick, falls $n$ verändert wird



Schnitt des oberen Astes der Ellipse mit  $h = 0,6$  liefert die linke Grenze  $p_u$  des Konfidenzintervalls:

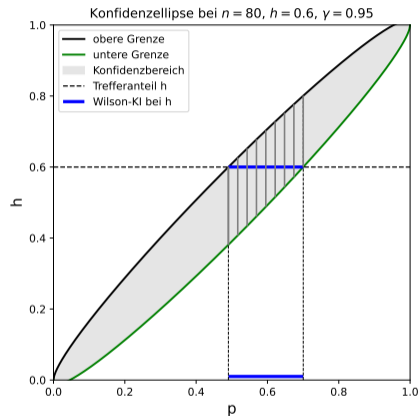
$$0,6 = p_u + 1,96\sqrt{\frac{p_u(1-p_u)}{n}}$$

Schnitt des unteren Astes der Ellipse

mit  $h = 0,6$  liefert die rechte Grenze  $p_o$  des Konfidenzintervalls:

$$0,6 = p_o - 1,96\sqrt{\frac{p_o(1-p_o)}{n}}$$

# Konfidenzellipsen – ein Überblick, falls $n$ verändert wird

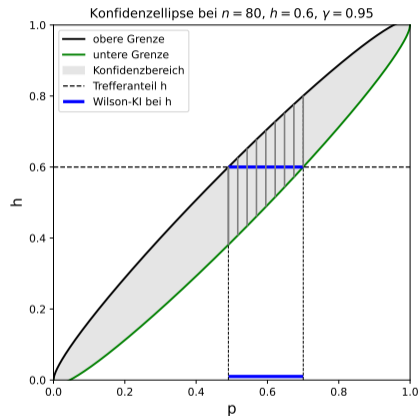


Die Konfidenzellipse hilft ...

- Prognoseintervalle von Konfidenzintervallen zu unterscheiden,
- Berechnungen zu verstehen,
- Abhängigkeiten zu entdecken und zu begründen.

Sie ist keine Darstellung der Verfahrenswahrscheinlichkeit.

# Konfidenzellipsen – ein Überblick, falls $n$ verändert wird



Die Konfidenzellipse hilft ...

- Prognoseintervalle von Konfidenzintervallen zu unterscheiden,
- Berechnungen zu verstehen,
- Abhängigkeiten zu entdecken und zu begründen.

Sie ist keine Darstellung der Verfahrenswahrscheinlichkeit.

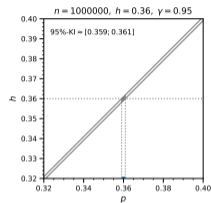
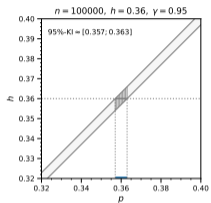
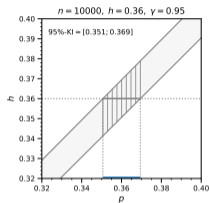
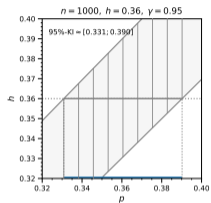
Überdeckungswahrscheinlichkeit: Zahlenwert

Sicherheitswahrscheinlichkeit: Wird oft benutzt, umgangssprachlich besetzt, 95% sicher - was soll ich mir darunter vorstellen?

Verfahrenswahrscheinlichkeit: Die Wahrscheinlichkeit gehört dem Verfahren, nicht dem Ergebnis; der Perspektivwechsel wird deutlich.

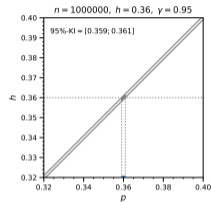
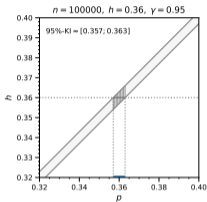
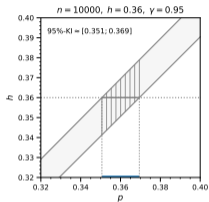
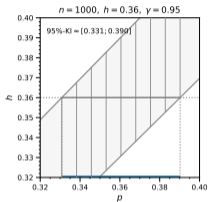
# Konfidenzintervall - Abhängigkeit vom $n$ und $\gamma$

$n$  verschieden,  $\gamma = 0,95$

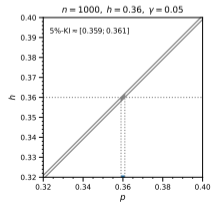
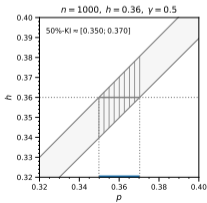
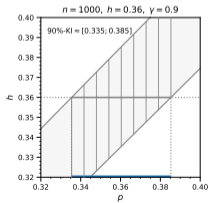
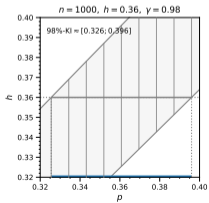


# Konfidenzintervall - Abhängigkeit vom $n$ und $\gamma$

$n$  verschieden,  $\gamma = 0,95$



$n = 1000$ ,  $\gamma$  verschieden



# Konfidenzellipse – mehr als ein Programm

**Konfidenzintervalle mit der Konfidenzellipse**

Die Konfidenzellipse erlaubt die geometrische Darstellung von Konfidenzintervallen, deren untere und obere Grenzen durch zwei Funktionen  $f$  und  $g$  gegeben sind durch

$$f(p) = p + z \cdot \sqrt{\frac{p(1-p)}{n}} \quad \text{und} \quad g(p) = p - z \cdot \sqrt{\frac{p(1-p)}{n}},$$

mit  $z = \Phi^{-1}\left(\frac{1+\gamma}{2}\right)$ .

Ausgewertet werden diese Konfidenzintervalle für einen **unbekannten** Parameterwert  $p$  durch Schnitte mit der Parallelen zur  $x$ -Achse mit Abstand  $\hat{h}$ .

Als geometrisches Hilfsmittel macht die Konfidenzellipse Abhängigkeiten von den Modellparametern  $n$ ,  $\gamma$  und dem Stichprobenergebnis  $\hat{h}$  sichtbar.

Der Unterschied zwischen Prognose- und Konfidenzintervallen wird erfahrbar:

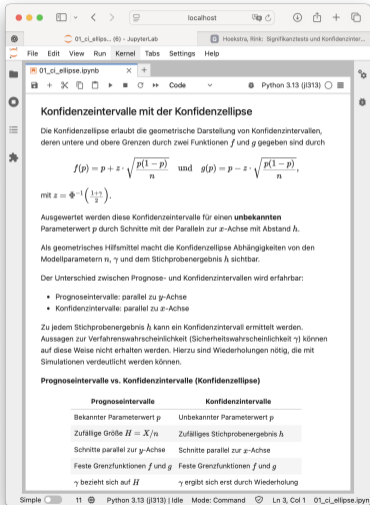
- Prognoseintervalle: parallel zu  $y$ -Achse
- Konfidenzintervalle: parallel zu  $x$ -Achse

Zu jedem Stichprobenergebnis  $\hat{h}$  kann ein Konfidenzintervall ermittelt werden. Aussagen zur Verfahrenswahrscheinlichkeit (Sicherheit/wahrscheinlichkeit  $\gamma$ ) können auf diese Weise nicht erhalten werden. Hierzu sind Wiederholungen nötig, die mit Simulationen verdeutlicht werden können.

**Prognoseintervalle vs. Konfidenzintervalle (Konfidenzellipse)**

Prognoseintervalle	Konfidenzintervalle
Bekannter Parameterwert $p$	Unbekannter Parameterwert $p$
Zufällige Größe $H = X/n$	Zufälliges Stichprobenergebnis $\hat{h}$
Schnitte parallel zur $y$ -Achse	Schnitte parallel zur $x$ -Achse
Feste Grenzfunktionen $f$ und $g$	Feste Grenzfunktionen $f$ und $g$
$\gamma$ bezieht sich auf $H$	$\gamma$ ergibt sich erst durch Wiederholung

# Konfidenzellipse – mehr als ein Programm



**Konfidenzintervalle mit der Konfidenzellipse**

Die Konfidenzellipse erlaubt die geometrische Darstellung von Konfidenzintervallen, deren untere und obere Grenzen durch zwei Funktionen  $f$  und  $g$  gegeben sind durch

$$f(p) = p + z \cdot \sqrt{\frac{p(1-p)}{n}} \quad \text{und} \quad g(p) = p - z \cdot \sqrt{\frac{p(1-p)}{n}},$$

mit  $z = \Phi^{-1}\left(\frac{1+\gamma}{2}\right)$ .

Ausgewertet werden diese Konfidenzintervalle für einen **unbekannten** Parameterwert  $p$  durch Schnitte mit der Parabeln zur  $x$ -Achse mit Abstand  $h$ .

Als geometrisches Hilfsmittel macht die Konfidenzellipse Abhängigkeiten von den Modellparametern  $n$ ,  $\gamma$  und dem Stichprobenergebnis  $h$  sichtbar.

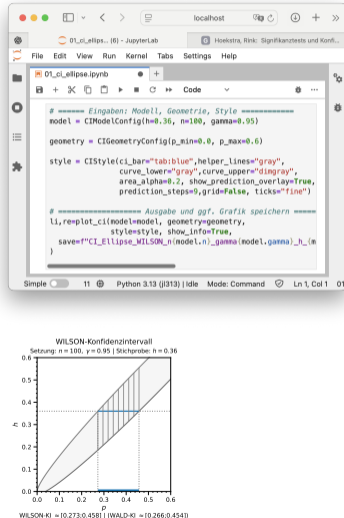
Der Unterschied zwischen Prognose- und Konfidenzintervallen wird erfahrbar:

- Prognoseintervalle: parallel zu  $y$ -Achse
- Konfidenzintervalle: parallel zu  $x$ -Achse

Zu jedem Stichprobenergebnis  $h$  kann ein Konfidenzintervall ermittelt werden. Aussagen zur Verfahrenswahrscheinlichkeit (Sicherheit/wahrscheinlichkeit  $\gamma$ ) können auf diese Weise nicht erhalten werden. Hierzu sind Wiederholungen nötig, die mit Simulationen verdeutlicht werden können.

**Prognoseintervalle vs. Konfidenzintervalle (Konfidenzellipse)**

Prognoseintervalle	Konfidenzintervalle
Bekannter Parameterwert $p$	Unbekannter Parameterwert $p$
Zufällige Größe $H = X/n$	Zufälliges Stichprobenergebnis $h$
Schnitte parallel zur $y$ -Achse	Schnitte parallel zur $x$ -Achse
Feste Grenzfunktionen $f$ und $g$	Feste Grenzfunktionen $f$ und $g$
$\gamma$ bezieht sich auf $H$	$\gamma$ ergibt sich erst durch Wiederholung



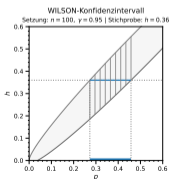
```
# ===== Eingaben: Modell, Geometrie, Style =====
model = CIModelConfig(h=0.36, n=100, gamma=0.95)

geometry = CIGeometryConfig(p_min=0.0, p_max=0.6)

style = CIStyle(ci_bar="tab:blue", helper_lines="gray",
               curve_lower="gray", curve_upper="dimgrey",
               area_alpha=0.2, show_prediction_overlay=True,
               prediction_steps=9, grid=False, ticks="fine")

# ===== Ausgabe und ggf. Grafik speichern =====
li, re = plot_ci(model=model, geometry=geometry,
                style=style, show_info=True,
                save="CI_Ellipse_WILSON_n(model.n)_gamma(model.gamma)_h_(m)")
```

**WILSON-Konfidenzintervall**  
Setzung:  $n = 100$ ,  $\gamma = 0.95$  | Stichprobe:  $h = 0.36$



WILSON-KI = [0.273; 0.458] | WALD-KI = [0.266; 0.454]

# Konfidenzellipse – mehr als ein Programm

**Konfidenzintervalle mit der Konfidenzellipse**

Die Konfidenzellipse erlaubt die geometrische Darstellung von Konfidenzintervallen, deren untere und obere Grenzen durch zwei Funktionen  $f$  und  $g$  gegeben sind durch

$$f(p) = p + z \cdot \sqrt{\frac{p(1-p)}{n}} \quad \text{und} \quad g(p) = p - z \cdot \sqrt{\frac{p(1-p)}{n}},$$

mit  $z = \Phi^{-1}\left(\frac{1+\gamma}{2}\right)$ .

Ausgewertet werden diese Konfidenzintervalle für einen **unbekannten** Parameterwert  $p$  durch Schnitte mit der Parabeln zur  $x$ -Achse mit Abstand  $h$ .

Als geometrisches Hilfsmittel macht die Konfidenzellipse Abhängigkeiten von den Modellparametern  $n$ ,  $\gamma$  und dem Stichprobenergebnis  $h$  sichtbar.

Der Unterschied zwischen Prognose- und Konfidenzintervallen wird erfahrbar:

- Prognoseintervalle: parallel zu  $y$ -Achse
- Konfidenzintervalle: parallel zu  $x$ -Achse

Zu jedem Stichprobenergebnis  $h$  kann ein Konfidenzintervall ermittelt werden. Aussagen zur Verfahrenswahrscheinlichkeit (Sicherheit/wahrscheinlichkeit  $\gamma$ ) können auf diese Weise nicht erhalten werden. Hierzu sind Wiederholungen nötig, die mit Simulationen verdeutlicht werden können.

**Prognoseintervalle vs. Konfidenzintervalle (Konfidenzellipse)**

Prognoseintervalle	Konfidenzintervalle
Bekannter Parameterwert $p$	Unbekannter Parameterwert $p$
Zufällige Größe $H = X/n$	Zufälliges Stichprobenergebnis $h$
Schnitte parallel zur $y$ -Achse	Schnitte parallel zur $x$ -Achse
Feste Grenzfunktionen $f$ und $g$	Feste Grenzfunktionen $f$ und $g$
$\gamma$ bezieht sich auf $H$	$\gamma$ ergibt sich erst durch Wiederholung

```
# ===== Eingaben: Modell, Geometrie, Style =====
model = CIModelConfig(h=0.36, n=100, gamma=0.95)
geometry = CIGeometryConfig(p_min=0.0, p_max=0.6)
style = CIStyle(ci_bar="tab:blue", helper_lines="gray",
               curve_lower="gray", curve_upper="dimgrey",
               area_alpha=0.2, show_prediction_overlay=True,
               prediction_steps=9, grid=False, ticks="fine")

# ===== Ausgabe und ggf. Grafik speichern =====
li, re = plot_ci(model=model, geometry=geometry,
                style=style, show_info=True,
                save=f"CI_Ellipse_WILSON_n({model.n})_gamma({model.gamma})_h({m})")
```

**WILSON-Konfidenzintervall**  
Setzung:  $n = 100$ ,  $\gamma = 0.95$  | Stichprobe:  $h = 0.36$

WILSON-KI = [0.273; 0.458] | WALD-KI = [0.266; 0.454]



## Konfidenzintervalle - WILSON, WALD und Clopper-Pearson

**WILSON:** Normalapproximation, Grenzen durch Lösen der Gleichung  $(h - p)^2 = \frac{z^2}{n} p(1 - p)$

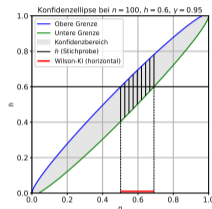
# Konfidenzintervalle - WILSON, WALD und Clopper-Pearson

**WILSON:** Normalapproximation, Grenzen durch Lösen der Gleichung  $(h - p)^2 = \frac{z^2}{n} p(1 - p)$

$$h = p_u - z\sqrt{p_u(1 - p_u)/n}$$

$$h = p_o + z\sqrt{p_o(1 - p_o)/n}$$

$$p_{u,o} = (1 - a) \cdot \frac{1}{2} + a \cdot h \mp z \cdot \sqrt{\frac{1}{4a} \cdot \frac{1-a}{a} + \frac{h(1-h)}{n}}; \quad a := \frac{1}{1+z^2/n}$$



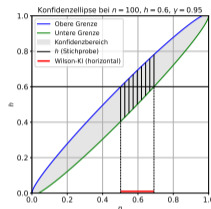
# Konfidenzintervalle - WILSON, WALD und Clopper-Pearson

**WILSON:** Normalapproximation, Grenzen durch Lösen der Gleichung  $(h - p)^2 = \frac{z^2}{n} p(1 - p)$

$$h = p_u - z\sqrt{p_u(1 - p_u)/n}$$

$$h = p_o + z\sqrt{p_o(1 - p_o)/n}$$

$$p_{u,o} = (1 - a) \cdot \frac{1}{2} + a \cdot h \mp z \cdot \sqrt{\frac{1}{4a} \cdot \frac{1-a}{a} + \frac{h(1-h)}{n}}; \quad a := \frac{1}{1+z^2/n}$$



**WALD:** Approximation der Lösungen von WILSON; in GeoGebra und Rechnern implementiert

$$\left[ h - z_{1-\alpha/2} \sqrt{h(1-h)/n}, h + z_{1-\alpha/2} \sqrt{h(1-h)/n} \right]$$

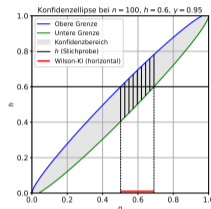
# Konfidenzintervalle - WILSON, WALD und Clopper-Pearson

**WILSON:** Normalapproximation, Grenzen durch Lösen der Gleichung  $(h - p)^2 = \frac{z^2}{n} p(1 - p)$

$$h = p_u - z\sqrt{p_u(1 - p_u)/n}$$

$$h = p_o + z\sqrt{p_o(1 - p_o)/n}$$

$$p_{u,o} = (1 - a) \cdot \frac{1}{2} + a \cdot h \mp z \cdot \sqrt{\frac{1}{4a} \cdot \frac{1-a}{a} + \frac{h(1-h)}{n}}; \quad a := \frac{1}{1+z^2/n}$$



**WALD:** Approximation der Lösungen von WILSON; in GeoGebra und Rechnern implementiert

$$\left[ h - z_{1-\alpha/2} \sqrt{h(1-h)/n}, h + z_{1-\alpha/2} \sqrt{h(1-h)/n} \right]$$

**Clopper-Pearson (exakt):** Inversion der Binomialverteilung; garantiert  $\mathbb{P}_p(\text{KI} \ni p) \geq 1 - \alpha$  für alle  $p$ .  
Grenzen über F-Verteilung

$$p_{\text{low}} = \frac{k}{k + (n - k + 1) F_{1-\alpha/2; 2(n-k+1), 2k}}, \quad p_{\text{high}} = \frac{(k + 1) F_{1-\alpha/2; 2(k+1), 2(n-k)}}{(n - k) + (k + 1) F_{1-\alpha/2; 2(k+1), 2(n-k)}}.$$

# Vergleich von Konfidenzintervallen

**Tabelle 1:**  $n = 100$ ;  $\gamma = 0,95$

k	$\hat{p}$	Wald	Wilson	Clopper–Pearson
30	0,3	[0,2102 ; 0,3898]	[0,2189 ; 0,3958]	[0,2124 ; 0,3998]
32	0,32	[0,2286 ; 0,4114]	[0,2367 ; 0,4166]	[0,2302 ; 0,4208]
34	0,34	[0,2472 ; 0,4328]	[0,2546 ; 0,4372]	[0,2482 ; 0,4415]
36	0,36	[0,2659 ; 0,4541]	[0,2727 ; 0,4576]	[0,2664 ; 0,4621]
37	0,37	[0,2754 ; 0,4646]	[0,2818 ; 0,4678]	[0,2756 ; 0,4724]
38	0,38	[0,2849 ; 0,4751]	[0,291 ; 0,4779]	[0,2848 ; 0,4825]
39	0,39	[0,2944 ; 0,4856]	[0,3002 ; 0,488]	[0,294 ; 0,4927]

Die drei Verfahren liefern leicht unterschiedliche, aber sehr ähnliche Intervalle. Wichtig ist nicht die Formel, sondern der gemeinsame Grundgedanke: Ein KI gibt an, welche  $p$ -Werte mit den Daten verträglich sind.

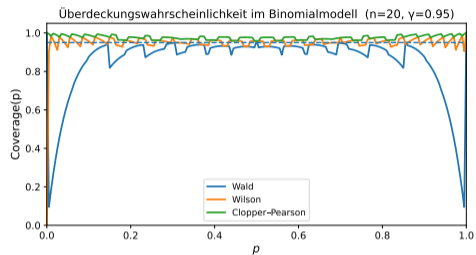
# Vergleich von Konfidenzintervallen

Tabelle 2:  $n = 1000$ ;  $\gamma = 0,95$

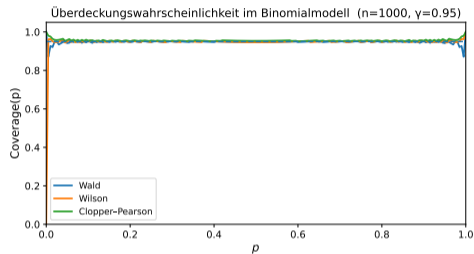
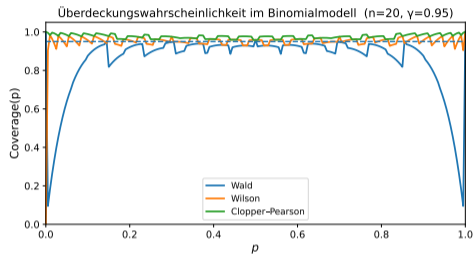
k	$\hat{p}$	Wald	Wilson	Clopper-Pearson
300	0,3	[0,2716 ; 0,3284]	[0,2724 ; 0,3291]	[0,2717 ; 0,3295]
320	0,32	[0,2911 ; 0,3489]	[0,2918 ; 0,3496]	[0,2912 ; 0,3499]
340	0,34	[0,3106 ; 0,3694]	[0,3113 ; 0,3699]	[0,3106 ; 0,3703]
360	0,36	[0,3302 ; 0,3898]	[0,3308 ; 0,3902]	[0,3302 ; 0,3906]
370	0,37	[0,3401 ; 0,3999]	[0,3406 ; 0,4004]	[0,34 ; 0,4008]
380	0,38	[0,3499 ; 0,4101]	[0,3504 ; 0,4105]	[0,3498 ; 0,4109]
390	0,39	[0,3598 ; 0,4202]	[0,3602 ; 0,4206]	[0,3596 ; 0,421]

Die drei Verfahren liefern leicht unterschiedliche, aber sehr ähnliche Intervalle. Wichtig ist nicht die Formel, sondern der gemeinsame Grundgedanke: Ein KI gibt an, welche  $p$ -Werte mit den Daten verträglich sind.

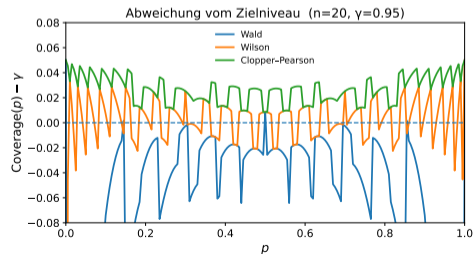
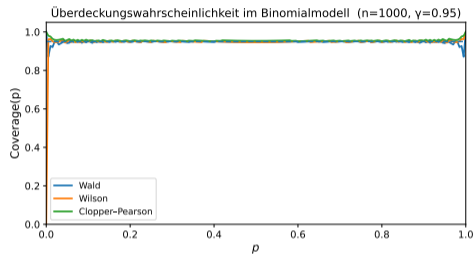
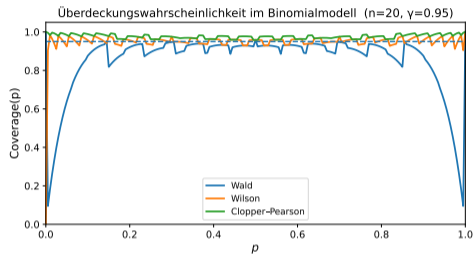
# Überdeckungswahrscheinlichkeiten mit Python



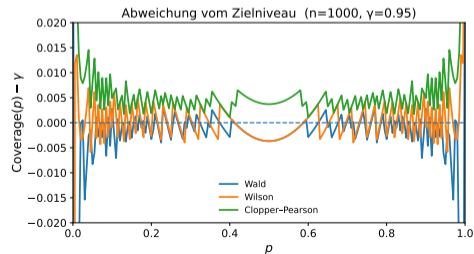
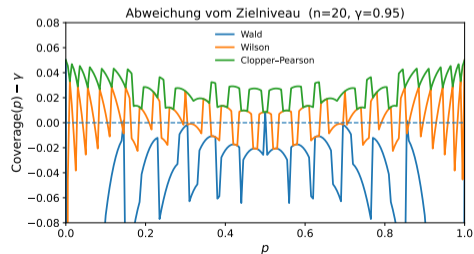
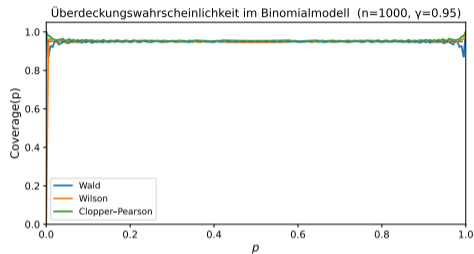
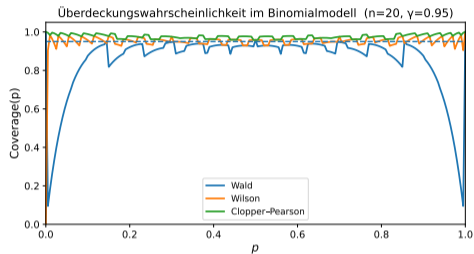
# Überdeckungswahrscheinlichkeiten mit Python



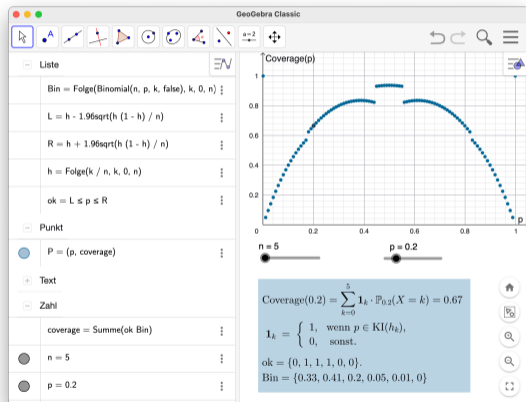
# Überdeckungswahrscheinlichkeiten mit Python



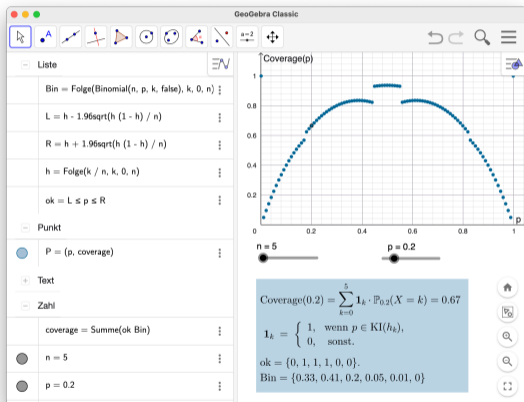
# Überdeckungswahrscheinlichkeiten mit Python



# Überdeckungswahrscheinlichkeiten mit GeoGebra



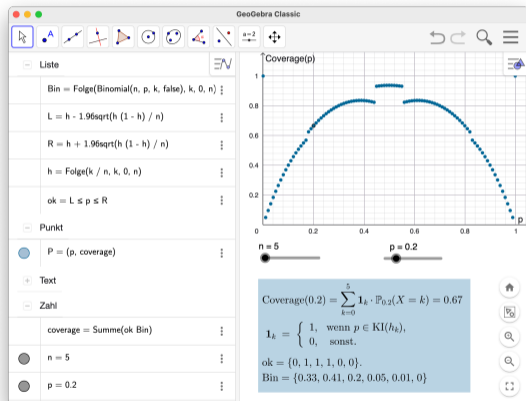
# Überdeckungswahrscheinlichkeiten mit GeoGebra - Trennung sichtbar machen



- **Gleichzeitigkeit mehrerer Ebenen**  
Modell – Realisation – Entscheidung – Wahrscheinlichkeit
- **Reduktion ohne Vereinfachung**  
Jede Zeile hat eine klar benennbare Funktion
- **Datentypen als Denkinstrumente**  
Listen – logische Ausdrücke – Funktionen – Punkte

*Kein neues Konzept – sondern Sichtbarmachung dessen, was im Modell ohnehin gedacht werden muss.*

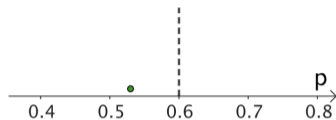
# Überdeckungswahrscheinlichkeiten mit GeoGebra - mehrere Ebenen, klar getrennt



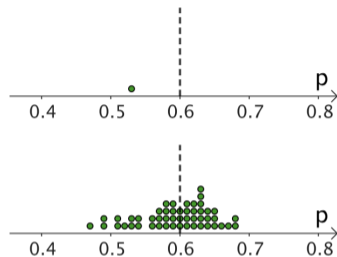
- **Modell**  
Alle möglichen Stichprobenergebnisse im Binomialmodell  
(Liste der Wahrscheinlichkeiten: Bin = Folge(...))
- **Entscheidung**  
Für jedes Ergebnis:  
Liegt der Parameter im zugehörigen Konfidenzintervall?  
(Boolesche Bewertung: ok = L ≤ p ≤ R)
- **Aggregation**  
Summe der Modellwahrscheinlichkeiten der „guten“ Fälle  
(Überdeckungswahrscheinlichkeit: Summe(ok Bin) )

Keine Simulation – sondern exakte Modellrechnung.  
Kein neues Objekt – sondern eine neue Sicht.

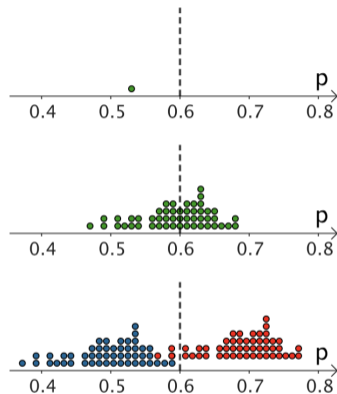
# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



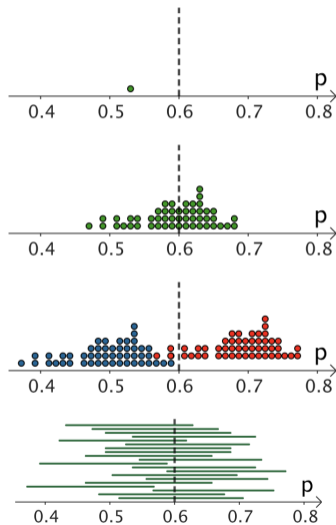
# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



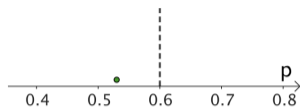
# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



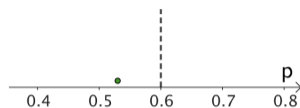
# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



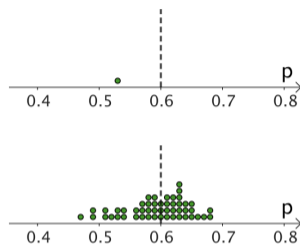
# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



## Eine Stichprobe – ein Ergebnis in der Realität

In der Realität beobachten wir genau einen Wert  $h$ . Der wahre Parameter liegt irgendwo in  $[0, 1]$  – mehr wissen wir nicht.

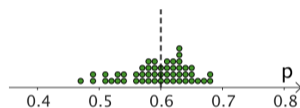
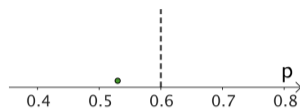
# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



Eine Stichprobe – ein Ergebnis in der Realität

In der Realität beobachten wir genau einen Wert  $h$ . Der wahre Parameter liegt irgendwo in  $[0, 1]$  – mehr wissen wir nicht.

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



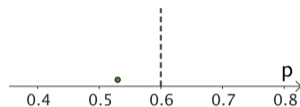
## Eine Stichprobe – ein Ergebnis in der Realität

In der Realität beobachten wir genau einen Wert  $h$ . Der wahre Parameter liegt irgendwo in  $[0, 1]$  – mehr wissen wir nicht.

## Gedankenexperiment im Modell

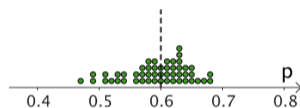
Im Modell denken wir viele mögliche Stichproben. So entsteht eine Verteilung möglicher Ergebnisse in unserem Kopf.

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



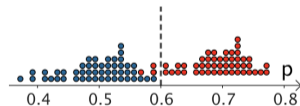
## Eine Stichprobe – ein Ergebnis in der Realität

In der Realität beobachten wir genau einen Wert  $h$ . Der wahre Parameter liegt irgendwo in  $[0, 1]$  – mehr wissen wir nicht.

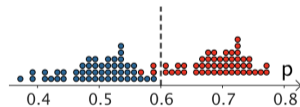
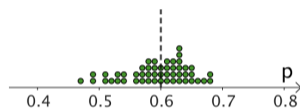
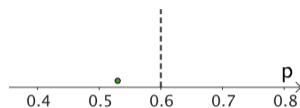


## Gedankenexperiment im Modell

Im Modell denken wir viele mögliche Stichproben. So entsteht eine Verteilung möglicher Ergebnisse in unserem Kopf.



# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



## Eine Stichprobe – ein Ergebnis in der Realität

In der Realität beobachten wir genau einen Wert  $h$ . Der wahre Parameter liegt irgendwo in  $[0, 1]$  – mehr wissen wir nicht.

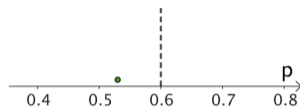
## Gedankenexperiment im Modell

Im Modell denken wir viele mögliche Stichproben. So entsteht eine Verteilung möglicher Ergebnisse in unserem Kopf.

## Grenzen durch gedachte Extremfälle im Modell

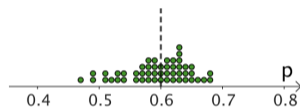
Um den unbekannt Parameter einzugrenzen, betrachten wir zu jeder möglichen Stichprobe die kleinste und größte noch plausible Lage.

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



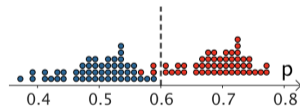
## Eine Stichprobe – ein Ergebnis in der Realität

In der Realität beobachten wir genau einen Wert  $h$ . Der wahre Parameter liegt irgendwo in  $[0, 1]$  – mehr wissen wir nicht.



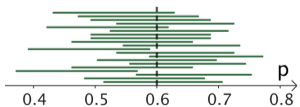
## Gedankenexperiment im Modell

Im Modell denken wir viele mögliche Stichproben. So entsteht eine Verteilung möglicher Ergebnisse in unserem Kopf.

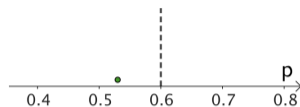


## Grenzen durch gedachte Extremfälle im Modell

Um den unbekannt Parameter einzugrenzen, betrachten wir zu jeder möglichen Stichprobe die kleinste und größte noch plausible Lage.

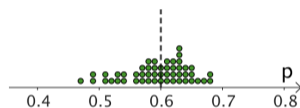


# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



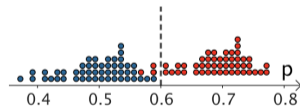
## Eine Stichprobe – ein Ergebnis in der Realität

In der Realität beobachten wir genau einen Wert  $h$ . Der wahre Parameter liegt irgendwo in  $[0, 1]$  – mehr wissen wir nicht.



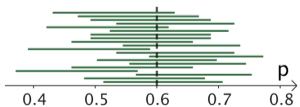
## Gedankenexperiment im Modell

Im Modell denken wir viele mögliche Stichproben. So entsteht eine Verteilung möglicher Ergebnisse in unserem Kopf.



## Grenzen durch gedachte Extremfälle im Modell

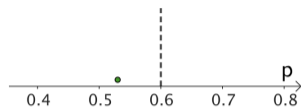
Um den unbekannt Parameter einzugrenzen, betrachten wir zu jeder möglichen Stichprobe die kleinste und größte noch plausible Lage.



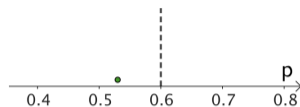
## Das Intervall als Verfahren im Modell

Wiederholen wir dieses Vorgehen gedanklich, so überdecken die entstehenden Intervalle den wahren Parameter in einem festen Anteil der Fälle.

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



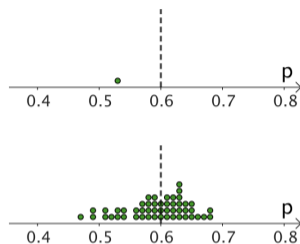
# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



## Realisation

Ein konkretes Intervall zu einem beobachteten Stichprobenergebnis  $h$ .

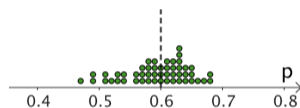
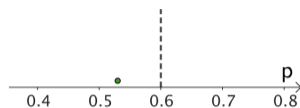
# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



## Realisation

Ein konkretes Intervall zu einem beobachteten Stichprobenergebnis  $h$ .

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



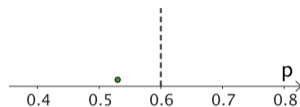
## Realisation

Ein konkretes Intervall zu einem beobachteten Stichprobenergebnis  $h$ .

## Zufallsgröße

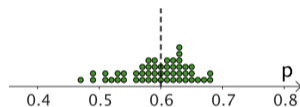
Das Stichprobenergebnis wird als Zufallsgröße  $H = X/n$  modelliert.

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



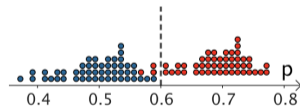
## Realisation

Ein konkretes Intervall zu einem beobachteten Stichprobenergebnis  $h$ .

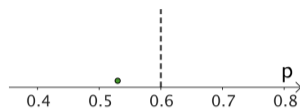


## Zufallsgröße

Das Stichprobenergebnis wird als Zufallsgröße  $H = X/n$  modelliert.

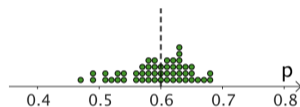


# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



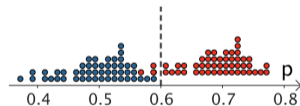
## Realisation

Ein konkretes Intervall zu einem beobachteten Stichprobenergebnis  $h$ .



## Zufallsgröße

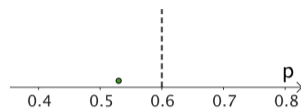
Das Stichprobenergebnis wird als Zufallsgröße  $H = X/n$  modelliert.



## Zufällige Intervallgrenzen

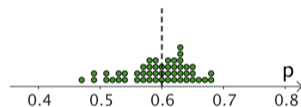
Die linke und rechte Grenze  $L(H)$  und  $R(H)$  sind selbst Zufallsgrößen.

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



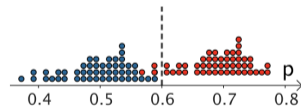
## Realisation

Ein konkretes Intervall zu einem beobachteten Stichprobenergebnis  $h$ .



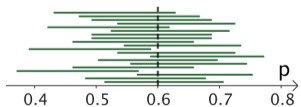
## Zufallsgröße

Das Stichprobenergebnis wird als Zufallsgröße  $H = X/n$  modelliert.

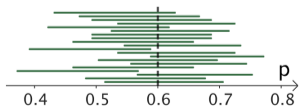
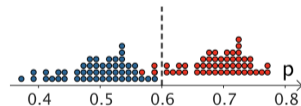
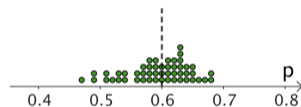
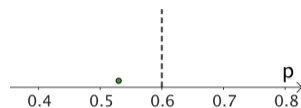


## Zufällige Intervallgrenzen

Die linke und rechte Grenze  $L(H)$  und  $R(H)$  sind selbst Zufallsgrößen.



# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



## Realisation

Ein konkretes Intervall zu einem beobachteten Stichprobenergebnis  $h$ .

## Zufallsgröße

Das Stichprobenergebnis wird als Zufallsgröße  $H = X/n$  modelliert.

## Zufällige Intervallgrenzen

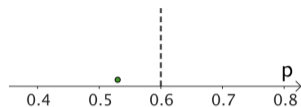
Die linke und rechte Grenze  $L(H)$  und  $R(H)$  sind selbst Zufallsgrößen.

## Konfidenzverfahren

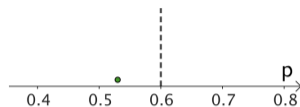
Das Verfahren ist so konstruiert, dass gilt

$$\mathbb{P}_p([L(H), R(H)] \ni p) = 0,95.$$

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



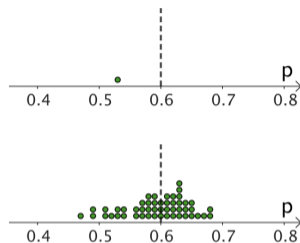
# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



1. WALD-Intervall als Realisation  $\rightarrow$  Realität

$$\left[ h \mp 1,96 \sqrt{\frac{h(1-h)}{n}} \right]$$

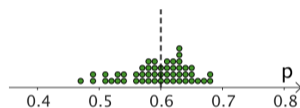
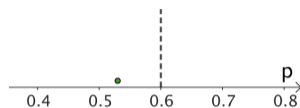
# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



1. WALD-Intervall als Realisation  $\rightarrow$  Realität

$$\left[ h \mp 1,96 \sqrt{\frac{h(1-h)}{n}} \right]$$

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



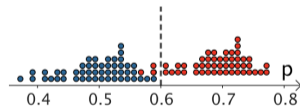
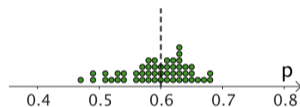
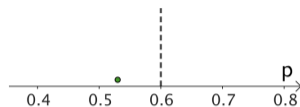
1. WALD-Intervall als Realisation  $\rightarrow$  Realität

$$\left[ h \mp 1,96 \sqrt{\frac{h(1-h)}{n}} \right]$$

2. Viele mögliche  $h$ -Werte  $\rightarrow$  Modell

$$H = \frac{X}{n}$$

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



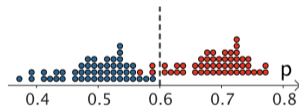
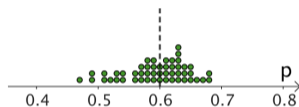
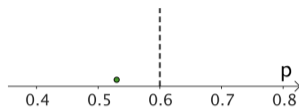
1. WALD-Intervall als Realisation  $\rightarrow$  Realität

$$\left[ h \mp 1,96 \sqrt{\frac{h(1-h)}{n}} \right]$$

2. Viele mögliche  $h$ -Werte  $\rightarrow$  Modell

$$H = \frac{X}{n}$$

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



1. WALD-Intervall als Realisation  $\rightarrow$  Realität

$$\left[ h \mp 1,96 \sqrt{\frac{h(1-h)}{n}} \right]$$

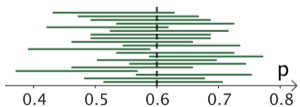
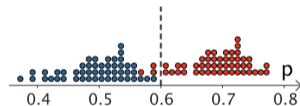
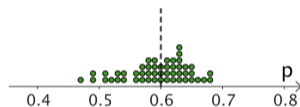
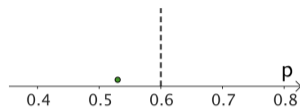
2. Viele mögliche  $h$ -Werte  $\rightarrow$  Modell

$$H = \frac{X}{n}$$

3. Grenzen als Zufallsgrößen  $L(H)$  und  $R(H)$   $\rightarrow$  Modelle

$$\left[ L(H), R(H) \right] = \left[ H \mp 1,96 \sqrt{\frac{H(1-H)}{n}} \right]$$

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



1. WALD-Intervall als Realisation  $\rightarrow$  Realität

$$\left[ h \mp 1,96 \sqrt{\frac{h(1-h)}{n}} \right]$$

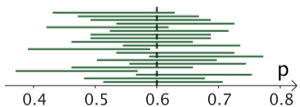
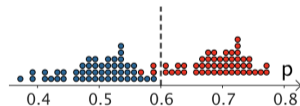
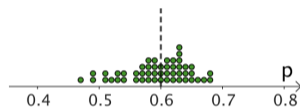
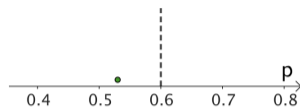
2. Viele mögliche  $h$ -Werte  $\rightarrow$  Modell

$$H = \frac{X}{n}$$

3. Grenzen als Zufallsgrößen  $L(H)$  und  $R(H)$   $\rightarrow$  Modelle

$$\left[ L(H), R(H) \right] = \left[ H \mp 1,96 \sqrt{\frac{H(1-H)}{n}} \right]$$

# Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren



1. WALD-Intervall als Realisation  $\rightarrow$  Realität

$$\left[ h \mp 1,96 \sqrt{\frac{h(1-h)}{n}} \right]$$

2. Viele mögliche  $h$ -Werte  $\rightarrow$  Modell

$$H = \frac{X}{n}$$

3. Grenzen als Zufallsgrößen  $L(H)$  und  $R(H)$   $\rightarrow$  Modelle

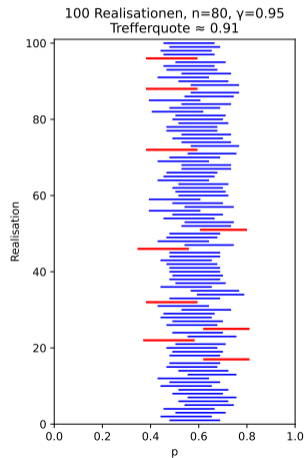
$$\left[ L(H), R(H) \right] = \left[ H \mp 1,96 \sqrt{\frac{H(1-H)}{n}} \right]$$

4. KI als Verfahren (viele Intervalle)  $\rightarrow$  Modelle

$$\mathbb{P}_p \left( \left[ L(H), R(H) \right] \ni p \right) = 0,95$$

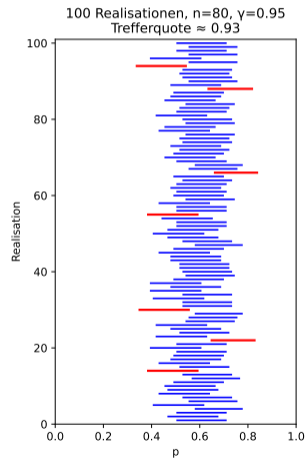
# Simulationen - der Weg zum Verständnis

Anteil der KI, die den wahren Parameter enthalten



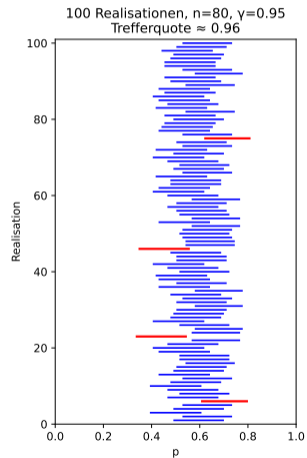
# Simulationen - der Weg zum Verständnis

Anteil der KI, die den wahren Parameter enthalten



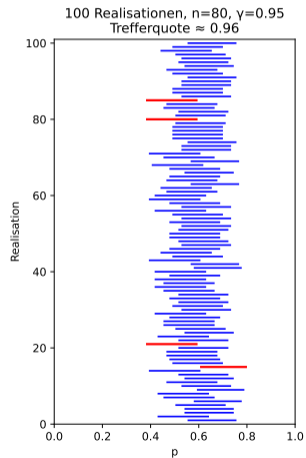
# Simulationen - der Weg zum Verständnis

Anteil der KI, die den wahren Parameter enthalten



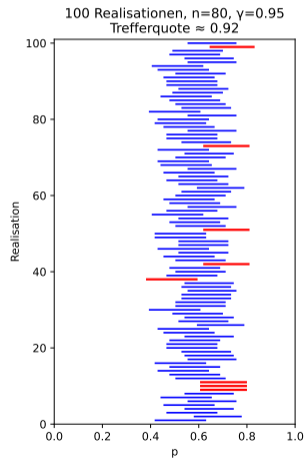
# Simulationen - der Weg zum Verständnis

Anteil der KI, die den wahren Parameter enthalten



# Simulationen - der Weg zum Verständnis

Anteil der KI, die den wahren Parameter enthalten

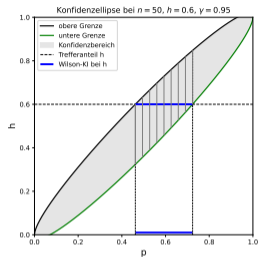


## Wo steckt hier eigentlich der Zufall?

- Der Zufall steckt **nicht** im Parameter.
- Er steckt **nicht** im Vertrauen.
- Er steckt **nicht** im konkreten Intervall.

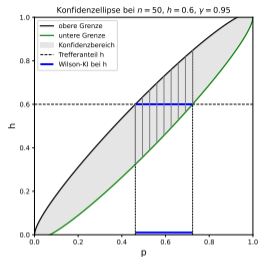
## Wo steckt hier eigentlich der Zufall?

- Der Zufall steckt **nicht** im Parameter.
- Er steckt **nicht** im Vertrauen.
- Er steckt **nicht** im konkreten Intervall.



## Wo steckt hier eigentlich der Zufall?

- Der Zufall steckt **nicht** im Parameter.
- Er steckt **nicht** im Vertrauen.
- Er steckt **nicht** im konkreten Intervall.

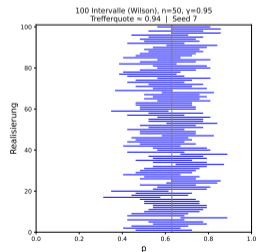
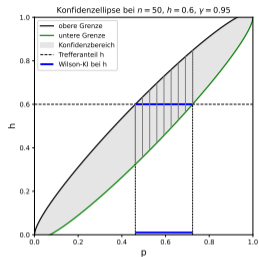


Er steckt im Gedanken der Wiederholung.

## Wo steckt hier eigentlich der Zufall?

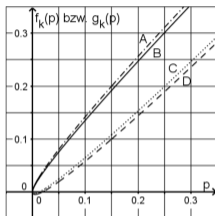
- Der Zufall steckt **nicht** im Parameter.
- Er steckt **nicht** im Vertrauen.
- Er steckt **nicht** im konkreten Intervall.

Er steckt im Gedanken der Wiederholung.



# IQB-Pool für das Jahr 2018 - PT B - WTR - Aufgabe 1e), 1f)

Auf der Grundlage der Stichprobe können mithilfe der abgebildeten Graphen A, B, C und D für den zu ermittelnden Schätzwert Konfidenzintervalle zu den Sicherheitswahrscheinlichkeiten 90 % und 95 % bestimmt werden.



Jeder der Graphen lässt sich durch eine der folgenden Funktionen beschreiben:

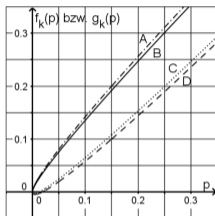
$$f_k : p \mapsto p - k \cdot \sqrt{\frac{p(1-p)}{n}} \quad g_k : p \mapsto p + k \cdot \sqrt{\frac{p(1-p)}{n}}$$

Dabei ist  $p \in [0;1]$ ,  $k \in \mathbb{R}^+$  und  $n$  der Umfang der Stichprobe.

- f** Bestimmen Sie ein Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. Entscheiden Sie, ob die Aussage „Im Mittel ist einer von fünf Bildschirmen fehlerhaft.“ durch das bestimmte Konfidenzintervall gestützt werden kann. 3
- g** Der Mitarbeiter zieht eine zweite Stichprobe vom Umfang  $2n$ . Auch in dieser Stichprobe sind 15 % der Bildschirme fehlerhaft. Begründen Sie, dass die Länge des zugehörigen Konfidenzintervalls im Vergleich zur ersten Stichprobe – bei gleicher Sicherheitswahrscheinlichkeit – geringer, aber nicht halb so groß ist. 3

# IQB-Pool für das Jahr 2018 - PT B - WTR - Aufgabe 1e), 1f)

Auf der Grundlage der Stichprobe können mithilfe der abgebildeten Graphen A, B, C und D für den zu ermittelnden Schätzwert Konfidenzintervalle zu den Sicherheitswahrscheinlichkeiten 90 % und 95 % bestimmt werden.



Jeder der Graphen lässt sich durch eine der folgenden Funktionen beschreiben:

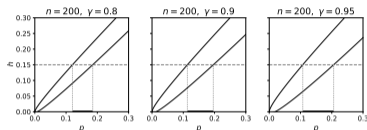
$$f_k : p \mapsto p - k \cdot \sqrt{\frac{p(1-p)}{n}} \quad g_k : p \mapsto p + k \cdot \sqrt{\frac{p(1-p)}{n}}$$

Dabei ist  $p \in [0;1]$ ,  $k \in \mathbb{R}^+$  und  $n$  der Umfang der Stichprobe.

- f** Bestimmen Sie ein Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. Entscheiden Sie, ob die Aussage „Im Mittel ist einer von fünf Bildschirmen fehlerhaft.“ durch das bestimmte Konfidenzintervall gestützt werden kann. 3
- g** Der Mitarbeiter zieht eine zweite Stichprobe vom Umfang  $2n$ . Auch in dieser Stichprobe sind 15 % der Bildschirme fehlerhaft. Begründen Sie, dass die Länge des zugehörigen Konfidenzintervalls im Vergleich zur ersten Stichprobe – bei gleicher Sicherheitswahrscheinlichkeit – geringer, aber nicht halb so groß ist. 3

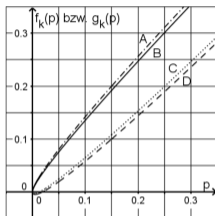
Lösungshinweise mit Variationen zu f)

Zeichnen: gleiches  $n$ , verschiedenes  $\gamma$



# IQB-Pool für das Jahr 2018 - PT B - WTR - Aufgabe 1e), 1f)

Auf der Grundlage der Stichprobe können mithilfe der abgebildeten Graphen A, B, C und D für den zu ermittelnden Schätzwert Konfidenzintervalle zu den Sicherheitswahrscheinlichkeiten 90 % und 95 % bestimmt werden.



Jeder der Graphen lässt sich durch eine der folgenden Funktionen beschreiben:

$$f_k : p \mapsto p - k \cdot \sqrt{\frac{p(1-p)}{n}} \quad g_k : p \mapsto p + k \cdot \sqrt{\frac{p(1-p)}{n}}$$

Dabei ist  $p \in [0;1]$ ,  $k \in \mathbb{R}^+$  und  $n$  der Umfang der Stichprobe.

**f** Bestimmen Sie ein Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. Entscheiden Sie, ob die Aussage „Im Mittel ist einer von fünf Bildschirmen fehlerhaft.“ durch das bestimmte Konfidenzintervall gestützt werden kann.

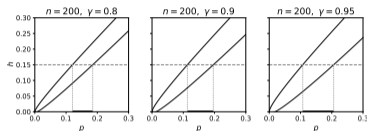
**g** Der Mitarbeiter zieht eine zweite Stichprobe vom Umfang  $2n$ . Auch in dieser Stichprobe sind 15 % der Bildschirme fehlerhaft. Begründen Sie, dass die Länge des zugehörigen Konfidenzintervalls im Vergleich zur ersten Stichprobe – bei gleicher Sicherheitswahrscheinlichkeit – geringer, aber nicht halb so groß ist.

3

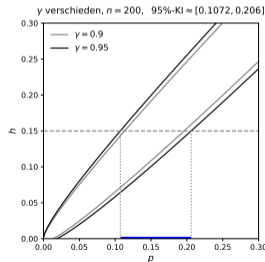
3

Lösungshinweise mit Variationen zu f)

Zeichnen: gleiches  $n$ , verschiedene  $\gamma$

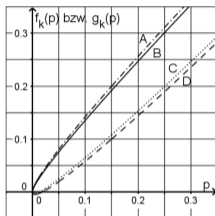


Zeichnen: Konfidenzellipsen



# IQB-Pool für das Jahr 2018 - PT B - WTR - Aufgabe 1e), 1f)

Auf der Grundlage der Stichprobe können mithilfe der abgebildeten Graphen A, B, C und D für den zu ermittelnden Schätzwert Konfidenzintervalle zu den Sicherheitswahrscheinlichkeiten 90 % und 95 % bestimmt werden.



Jeder der Graphen lässt sich durch eine der folgenden Funktionen beschreiben:

$$f_k : p \mapsto p - k \cdot \sqrt{\frac{p(1-p)}{n}} \quad g_k : p \mapsto p + k \cdot \sqrt{\frac{p(1-p)}{n}}$$

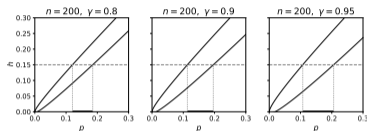
Dabei ist  $p \in [0;1]$ ,  $k \in \mathbb{R}^+$  und  $n$  der Umfang der Stichprobe.

**f** Bestimmen Sie ein Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. Entscheiden Sie, ob die Aussage „Im Mittel ist einer von fünf Bildschirmen fehlerhaft.“ durch das bestimmte Konfidenzintervall gestützt werden kann. 3

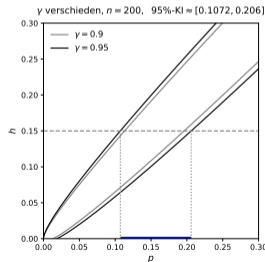
**g** Der Mitarbeiter zieht eine zweite Stichprobe vom Umfang  $2n$ . Auch in dieser Stichprobe sind 15 % der Bildschirme fehlerhaft. Begründen Sie, dass die Länge des zugehörigen Konfidenzintervalls im Vergleich zur ersten Stichprobe – bei gleicher Sicherheitswahrscheinlichkeit – geringer, aber nicht halb so groß ist. 3

Lösungshinweise mit Variationen zu f)

Zeichnen: gleiches  $n$ , verschiedene  $\gamma$



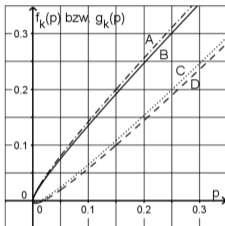
Zeichnen: Konfidenzellipsen



Erläutern:  $0,2 \in 95\text{-KI} \rightarrow$  Aussage kann gestützt werden

# IQB-Pool für das Jahr 2018 - PT B - WTR - Aufgabe 1f), 1g)

Auf der Grundlage der Stichprobe können mithilfe der abgebildeten Graphen A, B, C und D für den zu ermittelnden Schätzwert Konfidenzintervalle zu den Sicherheitswahrscheinlichkeiten 90 % und 95 % bestimmt werden.



Jeder der Graphen lässt sich durch eine der folgenden Funktionen beschreiben:

$$f_k : p \mapsto p - k \cdot \sqrt{\frac{p(1-p)}{n}} \quad g_k : p \mapsto p + k \cdot \sqrt{\frac{p(1-p)}{n}}$$

Dabei ist  $p \in [0; 1]$ ,  $k \in \mathbb{R}^+$  und  $n$  der Umfang der Stichprobe.

**f** Bestimmen Sie ein Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. Entscheiden Sie, ob die Aussage „Im Mittel ist einer von fünf Bildschirmen fehlerhaft.“ durch das bestimmte Konfidenzintervall gestützt werden kann.

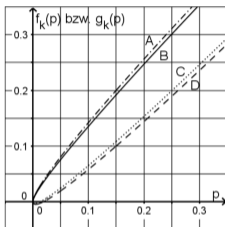
3

**g** Der Mitarbeiter zieht eine zweite Stichprobe vom Umfang  $2n$ . Auch in dieser Stichprobe sind 15 % der Bildschirme fehlerhaft. Begründen Sie, dass die Länge des zugehörigen Konfidenzintervalls im Vergleich zur ersten Stichprobe – bei gleicher Sicherheitswahrscheinlichkeit – geringer, aber nicht halb so groß ist.

3

# IQB-Pool für das Jahr 2018 - PT B - WTR - Aufgabe 1f), 1g)

Auf der Grundlage der Stichprobe können mithilfe der abgebildeten Graphen A, B, C und D für den zu ermittelnden Schätzwert Konfidenzintervalle zu den Sicherheitswahrscheinlichkeiten 90 % und 95 % bestimmt werden.



Jeder der Graphen lässt sich durch eine der folgenden Funktionen beschreiben:

$$f_k : p \mapsto p - k \cdot \sqrt{\frac{p(1-p)}{n}} \quad g_k : p \mapsto p + k \cdot \sqrt{\frac{p(1-p)}{n}}$$

Dabei ist  $p \in [0;1]$ ,  $k \in \mathbb{R}^+$  und  $n$  der Umfang der Stichprobe.

**f** Bestimmen Sie ein Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. Entscheiden Sie, ob die Aussage „Im Mittel ist einer von fünf Bildschirmen fehlerhaft.“ durch das bestimmte Konfidenzintervall gestützt werden kann.

3

**g** Der Mitarbeiter zieht eine zweite Stichprobe vom Umfang  $2n$ . Auch in dieser Stichprobe sind 15 % der Bildschirme fehlerhaft. Begründen Sie, dass die Länge des zugehörigen Konfidenzintervalls im Vergleich zur ersten Stichprobe – bei gleicher Sicherheitswahrscheinlichkeit – geringer, aber nicht halb so groß ist.

3

Lösungshinweise zu g)

Erwartungshorizon (IQB):

**g** Für die Länge des Konfidenzintervalls ergibt sich:

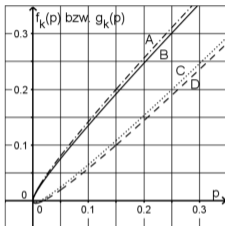
$$2k \cdot \sqrt{\frac{0,15 \cdot 0,85}{2n}} = 2k \cdot \sqrt{\frac{0,15 \cdot 0,85}{n}} \cdot \frac{1}{\sqrt{2}}$$

$$\text{Es gilt: } \frac{1}{\sqrt{2}} < 1, \frac{1}{\sqrt{2}} \neq \frac{1}{2}$$

3

# IQB-Pool für das Jahr 2018 - PT B - WTR - Aufgabe 1f), 1g)

Auf der Grundlage der Stichprobe können mithilfe der abgebildeten Graphen A, B, C und D für den zu ermittelnden Schätzwert Konfidenzintervalle zu den Sicherheitswahrscheinlichkeiten 90 % und 95 % bestimmt werden.



Jeder der Graphen lässt sich durch eine der folgenden Funktionen beschreiben:

$$f_k : p \mapsto p - k \cdot \sqrt{\frac{p(1-p)}{n}} \quad g_k : p \mapsto p + k \cdot \sqrt{\frac{p(1-p)}{n}}$$

Dabei ist  $p \in [0;1]$ ,  $k \in \mathbb{R}^+$  und  $n$  der Umfang der Stichprobe.

**f** Bestimmen Sie ein Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. Entscheiden Sie, ob die Aussage „Im Mittel ist einer von fünf Bildschirmen fehlerhaft.“ durch das bestimmte Konfidenzintervall gestützt werden kann.

3

**g** Der Mitarbeiter zieht eine zweite Stichprobe vom Umfang  $2n$ . Auch in dieser Stichprobe sind 15 % der Bildschirme fehlerhaft. Begründen Sie, dass die Länge des zugehörigen Konfidenzintervalls im Vergleich zur ersten Stichprobe – bei gleicher Sicherheitswahrscheinlichkeit – geringer, aber nicht halb so groß ist.

3

Lösungshinweise zu g)

Erwartungshorizon (IQB):

**g** Für die Länge des Konfidenzintervalls ergibt sich:

$$2k \cdot \sqrt{\frac{0,15 \cdot 0,85}{2n}} = 2k \cdot \sqrt{\frac{0,15 \cdot 0,85}{n}} \cdot \frac{1}{\sqrt{2}}$$

Es gilt:  $\frac{1}{\sqrt{2}} < 1$ ,  $\frac{1}{\sqrt{2}} \neq \frac{1}{2}$

3

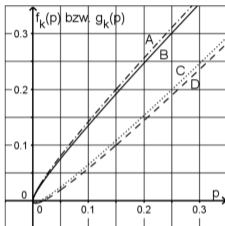
Vorsicht: Das gilt nur für WALD-Konfidenzintervalle.

$$\text{WALD-KI: } \left[ h - k \cdot \sqrt{\frac{h(1-h)}{n}}, h + k \cdot \sqrt{\frac{h(1-h)}{n}} \right]$$

$$\rightarrow \text{Länge des WALD-KI: } 2k \cdot \sqrt{\frac{h(1-h)}{n}}$$

# IQB-Pool für das Jahr 2018 - PT B - WTR - Aufgabe 1f), 1g)

Auf der Grundlage der Stichprobe können mithilfe der abgebildeten Graphen A, B, C und D für den zu ermittelnden Schätzwert Konfidenzintervalle zu den Sicherheitswahrscheinlichkeiten 90 % und 95 % bestimmt werden.



Jeder der Graphen lässt sich durch eine der folgenden Funktionen beschreiben:

$$f_k : p \mapsto p - k \cdot \sqrt{\frac{p(1-p)}{n}} \quad g_k : p \mapsto p + k \cdot \sqrt{\frac{p(1-p)}{n}}$$

Dabei ist  $p \in [0;1]$ ,  $k \in \mathbb{R}^+$  und  $n$  der Umfang der Stichprobe.

**f** Bestimmen Sie ein Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. Entscheiden Sie, ob die Aussage „Im Mittel ist einer von fünf Bildschirmen fehlerhaft.“ durch das bestimmte Konfidenzintervall gestützt werden kann.

3

**g** Der Mitarbeiter zieht eine zweite Stichprobe vom Umfang  $2n$ . Auch in dieser Stichprobe sind 15 % der Bildschirme fehlerhaft. Begründen Sie, dass die Länge des zugehörigen Konfidenzintervalls im Vergleich zur ersten Stichprobe – bei gleicher Sicherheitswahrscheinlichkeit – geringer, aber nicht halb so groß ist.

3

Lösungshinweise zu g)

Erwartungshorizon (IQB):

**g** Für die Länge des Konfidenzintervalls ergibt sich:

$$2k \cdot \sqrt{\frac{0,15 \cdot 0,85}{2n}} = 2k \cdot \sqrt{\frac{0,15 \cdot 0,85}{n}} \cdot \frac{1}{\sqrt{2}}$$

$$\text{Es gilt: } \frac{1}{\sqrt{2}} < 1, \frac{1}{\sqrt{2}} \neq \frac{1}{2}$$

3

Vorsicht: Das gilt nur für WALD-Konfidenzintervalle.

$$\text{WALD-KI: } \left[ h - k \cdot \sqrt{\frac{h(1-h)}{n}}, h + k \cdot \sqrt{\frac{h(1-h)}{n}} \right]$$

$$\rightarrow \text{Länge des WALD-KI: } 2k \cdot \sqrt{\frac{h(1-h)}{n}}$$

WILSON-KI:

Grenzen dieses Intervalls sind Lösungen von

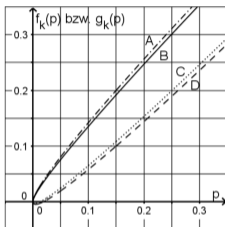
$$h = p_u + k \cdot \sqrt{\frac{p_u(1-p_u)}{n}}; \quad h = p_o - k \cdot \sqrt{\frac{p_o(1-p_o)}{n}}$$

$\rightarrow$  Länge des WILSON-KI:

$$\frac{n}{k^2 + n} \cdot 2k \cdot \sqrt{\left(\frac{k}{2n}\right)^2 + \frac{h(1-h)}{n}}$$

# IQB-Pool für das Jahr 2018 - PT B - WTR - Aufgabe 1f), 1g)

Auf der Grundlage der Stichprobe können mithilfe der abgebildeten Graphen A, B, C und D für den zu ermittelnden Schätzwert Konfidenzintervalle zu den Sicherheitswahrscheinlichkeiten 90 % und 95 % bestimmt werden.



Jeder der Graphen lässt sich durch eine der folgenden Funktionen beschreiben:

$$f_k : p \mapsto p - k \cdot \sqrt{\frac{p(1-p)}{n}} \quad g_k : p \mapsto p + k \cdot \sqrt{\frac{p(1-p)}{n}}$$

Dabei ist  $p \in [0; 1]$ ,  $k \in \mathbb{R}^+$  und  $n$  der Umfang der Stichprobe.

**f** Bestimmen Sie ein Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. Entscheiden Sie, ob die Aussage „Im Mittel ist einer von fünf Bildschirmen fehlerhaft.“ durch das bestimmte Konfidenzintervall gestützt werden kann.

**g** Der Mitarbeiter zieht eine zweite Stichprobe vom Umfang  $2n$ . Auch in dieser Stichprobe sind 15 % der Bildschirme fehlerhaft. Begründen Sie, dass die Länge des zugehörigen Konfidenzintervalls im Vergleich zur ersten Stichprobe – bei gleicher Sicherheitswahrscheinlichkeit – geringer, aber nicht halb so groß ist.

Lösungshinweise zu g)

Erwartungshorizon (IQB):

<b>g</b> Für die Länge des Konfidenzintervalls ergibt sich:	3
$2k \cdot \sqrt{\frac{0,15 \cdot 0,85}{2n}} = 2k \cdot \sqrt{\frac{0,15 \cdot 0,85}{n}} \cdot \frac{1}{\sqrt{2}}$	
Es gilt: $\frac{1}{\sqrt{2}} < 1$ , $\frac{1}{\sqrt{2}} \neq \frac{1}{2}$	

Vorsicht: Das gilt nur für WALD-Konfidenzintervalle.

$$\text{WALD-KI: } \left[ h - k \cdot \sqrt{\frac{h(1-h)}{n}}, h + k \cdot \sqrt{\frac{h(1-h)}{n}} \right]$$

$$\rightarrow \text{Länge des WALD-KI: } 2k \cdot \sqrt{\frac{h(1-h)}{n}}$$

**WILSON-KI:**

Grenzen dieses Intervalls sind Lösungen von

$$h = p_u + k \cdot \sqrt{\frac{p_u(1-p_u)}{n}}; \quad h = p_o - k \cdot \sqrt{\frac{p_o(1-p_o)}{n}}$$

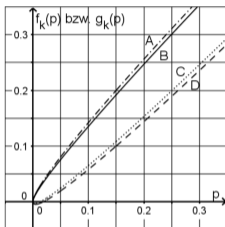
$\rightarrow$  Länge des WILSON-KI:

$$\frac{n}{k^2 + n} \cdot 2k \cdot \sqrt{\left(\frac{k}{2n}\right)^2 + \frac{h(1-h)}{n}}$$

$$\text{sehr gute Näherung } (n \gg k): 2k \cdot \sqrt{\frac{h(1-h)}{n}}$$

# IQB-Pool für das Jahr 2018 - PT B - WTR - Aufgabe 1f), 1g)

Auf der Grundlage der Stichprobe können mithilfe der abgebildeten Graphen A, B, C und D für den zu ermittelnden Schätzwert Konfidenzintervalle zu den Sicherheitswahrscheinlichkeiten 90 % und 95 % bestimmt werden.



Jeder der Graphen lässt sich durch eine der folgenden Funktionen beschreiben:

$$f_k : p \mapsto p - k \cdot \sqrt{\frac{p(1-p)}{n}} \quad g_k : p \mapsto p + k \cdot \sqrt{\frac{p(1-p)}{n}}$$

Dabei ist  $p \in [0; 1]$ ,  $k \in \mathbb{R}^+$  und  $n$  der Umfang der Stichprobe.

**f** Bestimmen Sie ein Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. Entscheiden Sie, ob die Aussage „Im Mittel ist einer von fünf Bildschirmen fehlerhaft.“ durch das bestimmte Konfidenzintervall gestützt werden kann.

**g** Der Mitarbeiter zieht eine zweite Stichprobe vom Umfang  $2n$ . Auch in dieser Stichprobe sind 15 % der Bildschirme fehlerhaft. Begründen Sie, dass die Länge des zugehörigen Konfidenzintervalls im Vergleich zur ersten Stichprobe – bei gleicher Sicherheitswahrscheinlichkeit – geringer, aber nicht halb so groß ist.

Lösungshinweise zu g)

Erwartungshorizon (IQB):

<b>g</b> Für die Länge des Konfidenzintervalls ergibt sich:	3
$2k \cdot \sqrt{\frac{0,15 \cdot 0,85}{2n}} = 2k \cdot \sqrt{\frac{0,15 \cdot 0,85}{n}} \cdot \frac{1}{\sqrt{2}}$	
Es gilt: $\frac{1}{\sqrt{2}} < 1$ , $\frac{1}{\sqrt{2}} \neq \frac{1}{2}$	

Vorsicht: Das gilt nur für WALD-Konfidenzintervalle.

$$\text{WALD-KI: } \left[ h - k \cdot \sqrt{\frac{h(1-h)}{n}}, h + k \cdot \sqrt{\frac{h(1-h)}{n}} \right]$$

$$\rightarrow \text{Länge des WALD-KI: } 2k \cdot \sqrt{\frac{h(1-h)}{n}}$$

**WILSON-KI:**

Grenzen dieses Intervalls sind Lösungen von

$$h = p_u + k \cdot \sqrt{\frac{p_u(1-p_u)}{n}}; \quad h = p_o - k \cdot \sqrt{\frac{p_o(1-p_o)}{n}}$$

$\rightarrow$  Länge des WILSON-KI:

$$\frac{n}{k^2 + n} \cdot 2k \cdot \sqrt{\left(\frac{k}{2n}\right)^2 + \frac{h(1-h)}{n}}$$

sehr gute Näherung ( $n \gg k$ ):  $2k \cdot \sqrt{\frac{h(1-h)}{n}}$

**WALD-KI:** Näherung von WILSON-KI

**WILSON-KI:**  $\text{Bin}(n, p) \sim N(\mu, \sigma^2)$

# IQB-Pool für das Jahr 2019 - PT B - WTR - Aufgabe 1d), 1e)

Kurz nach einer Änderung im Herstellungsverfahren stellt das Unternehmen den Anteil fehlerhafter Geräte von 25 % infrage. Um bei einer Sicherheitswahrscheinlichkeit von 95 % einen Schätzwert für den Anteil fehlerhafter Geräte zu ermitteln, wird eine Stichprobe von 100 Geräten betrachtet. Abbildung 1 zeigt die Graphen der folgenden für  $p \in [0;1]$  definierten Funktionen:

$$f: p \mapsto p - 1,96 \cdot \sqrt{\frac{p(1-p)}{100}} \quad g: p \mapsto p + 1,96 \cdot \sqrt{\frac{p(1-p)}{100}}$$

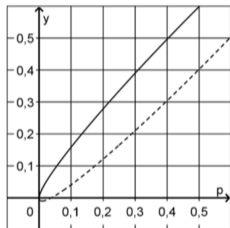


Abb. 1

- d Bestimmen Sie grafisch alle möglichen Anzahlen fehlerhafter Geräte in der Stichprobe, für die bei einer Sicherheitswahrscheinlichkeit von 95 % jeweils Anlass dazu bestehen würde, die Korrektheit des gegebenen Anteils fehlerhafter Geräte infrage zu stellen. 3
- e Die betrachtete Stichprobe enthält 19 fehlerhafte Geräte. Bestimmen Sie grafisch das zu dieser Anzahl gehörende Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. 3

# IQB-Pool für das Jahr 2019 - PT B - WTR - Aufgabe 1d), 1e)

Kurz nach einer Änderung im Herstellungsverfahren stellt das Unternehmen den Anteil fehlerhafter Geräte von 25 % infrage. Um bei einer Sicherheitswahrscheinlichkeit von 95 % einen Schätzwert für den Anteil fehlerhafter Geräte zu ermitteln, wird eine Stichprobe von 100 Geräten betrachtet. Abbildung 1 zeigt die Graphen der folgenden für  $p \in [0,1]$  definierten Funktionen:

$$f: p \mapsto p - 1,96 \cdot \sqrt{\frac{p(1-p)}{100}} \quad g: p \mapsto p + 1,96 \cdot \sqrt{\frac{p(1-p)}{100}}$$

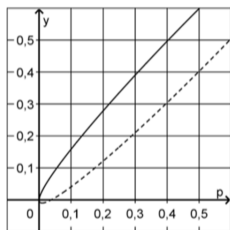


Abb. 1

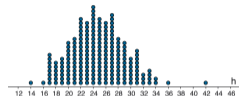
- d Bestimmen Sie grafisch alle möglichen Anzahlen fehlerhafter Geräte in der Stichprobe, für die bei einer Sicherheitswahrscheinlichkeit von 95 % jeweils Anlass dazu bestehen würde, die Korrektheit des gegebenen Anteils fehlerhafter Geräte infrage zu stellen.
- e Die betrachtete Stichprobe enthält 19 fehlerhafte Geräte. Bestimmen Sie grafisch das zu dieser Anzahl gehörende Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %.

3

3

Lösungshinweise mit Variationen zu d)

Simulieren: 100 Simulationen mit  $p = 0,25$  und  $n = 100$



# IQB-Pool für das Jahr 2019 - PT B - WTR - Aufgabe 1d), 1e)

Kurz nach einer Änderung im Herstellungsverfahren stellt das Unternehmen den Anteil fehlerhafter Geräte von 25 % infrage. Um bei einer Sicherheitswahrscheinlichkeit von 95 % einen Schätzwert für den Anteil fehlerhafter Geräte zu ermitteln, wird eine Stichprobe von 100 Geräten betrachtet. Abbildung 1 zeigt die Graphen der folgenden für  $p \in [0,1]$  definierten Funktionen:

$$f: p \mapsto p - 1,96 \cdot \sqrt{\frac{p(1-p)}{100}} \quad g: p \mapsto p + 1,96 \cdot \sqrt{\frac{p(1-p)}{100}}$$

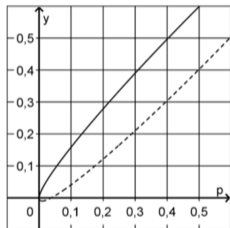


Abb. 1

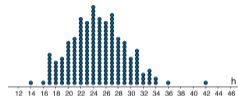
- d Bestimmen Sie grafisch alle möglichen Anzahlen fehlerhafter Geräte in der Stichprobe, für die bei einer Sicherheitswahrscheinlichkeit von 95 % jeweils Anlass dazu bestehen würde, die Korrektheit des gegebenen Anteils fehlerhafter Geräte infrage zu stellen.
- e Die betrachtete Stichprobe enthält 19 fehlerhafte Geräte. Bestimmen Sie grafisch das zu dieser Anzahl gehörende Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %.

3

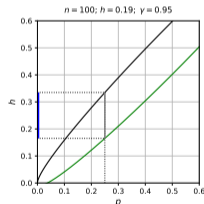
3

Lösungshinweise mit Variationen zu d)

Simulieren: 100 Simulationen mit  $p = 0,25$  und  $n = 100$



Zeichnen: Konfidenzellipse



# IQB-Pool für das Jahr 2019 - PT B - WTR - Aufgabe 1d), 1e)

Kurz nach einer Änderung im Herstellungsverfahren stellt das Unternehmen den Anteil fehlerhafter Geräte von 25 % infrage. Um bei einer Sicherheitswahrscheinlichkeit von 95 % einen Schätzwert für den Anteil fehlerhafter Geräte zu ermitteln, wird eine Stichprobe von 100 Geräten betrachtet. Abbildung 1 zeigt die Graphen der folgenden für  $p \in [0,1]$  definierten Funktionen:

$$f: p \mapsto p - 1,96 \cdot \sqrt{\frac{p(1-p)}{100}} \quad g: p \mapsto p + 1,96 \cdot \sqrt{\frac{p(1-p)}{100}}$$

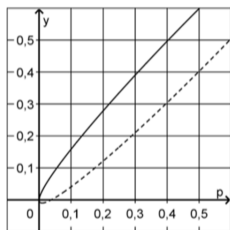


Abb. 1

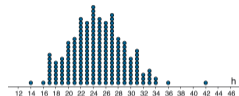
- d Bestimmen Sie grafisch alle möglichen Anzahlen fehlerhafter Geräte in der Stichprobe, für die bei einer Sicherheitswahrscheinlichkeit von 95 % jeweils Anlass dazu bestehen würde, die Korrektheit des gegebenen Anteils fehlerhafter Geräte infrage zu stellen.
- e Die betrachtete Stichprobe enthält 19 fehlerhafte Geräte. Bestimmen Sie grafisch das zu dieser Anzahl gehörende Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %.

3

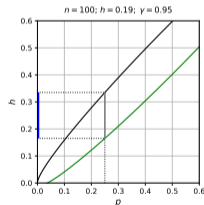
3

Lösungshinweise mit Variationen zu d)

Simulieren: 100 Simulationen mit  $p = 0,25$  und  $n = 100$



Zeichnen: Konfidenzellipse



Rechnen: 95 %-Prognoseintervall:

$$\left[ np - 1,96 \sqrt{np(1-p)}; np + 1,96 \sqrt{np(1-p)} \right]$$

$$[16,51; 33,49] \rightarrow [17; 33]$$

# IQB-Pool für das Jahr 2019 - PT B - WTR - Aufgabe 1d), 1e)

Kurz nach einer Änderung im Herstellungsverfahren stellt das Unternehmen den Anteil fehlerhafter Geräte von 25 % infrage. Um bei einer Sicherheitswahrscheinlichkeit von 95 % einen Schätzwert für den Anteil fehlerhafter Geräte zu ermitteln, wird eine Stichprobe von 100 Geräten betrachtet. Abbildung 1 zeigt die Graphen der folgenden für  $p \in [0,1]$  definierten Funktionen:

$$f: p \mapsto p - 1,96 \cdot \sqrt{\frac{p(1-p)}{100}} \quad g: p \mapsto p + 1,96 \cdot \sqrt{\frac{p(1-p)}{100}}$$

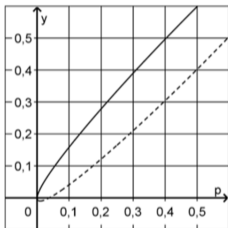


Abb. 1

- d Bestimmen Sie grafisch alle möglichen Anzahlen fehlerhafter Geräte in der Stichprobe, für die bei einer Sicherheitswahrscheinlichkeit von 95 % jeweils Anlass dazu bestehen würde, die Korrektheit des gegebenen Anteils fehlerhafter Geräte infrage zu stellen. 3
- e Die betrachtete Stichprobe enthält 19 fehlerhafte Geräte. Bestimmen Sie grafisch das zu dieser Anzahl gehörende Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. 3

# IQB-Pool für das Jahr 2019 - PT B - WTR - Aufgabe 1d), 1e)

Kurz nach einer Änderung im Herstellungsverfahren stellt das Unternehmen den Anteil fehlerhafter Geräte von 25 % infrage. Um bei einer Sicherheitswahrscheinlichkeit von 95 % einen Schätzwert für den Anteil fehlerhafter Geräte zu ermitteln, wird eine Stichprobe von 100 Geräten betrachtet. Abbildung 1 zeigt die Graphen der folgenden für  $p \in [0,1]$  definierten Funktionen:

$$f: p \mapsto p - 1,96 \cdot \sqrt{\frac{p(1-p)}{100}} \quad g: p \mapsto p + 1,96 \cdot \sqrt{\frac{p(1-p)}{100}}$$

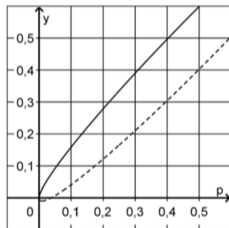
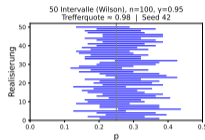


Abb. 1

- d Bestimmen Sie grafisch alle möglichen Anzahlen fehlerhafter Geräte in der Stichprobe, für die bei einer Sicherheitswahrscheinlichkeit von 95 % jeweils Anlass dazu bestehen würde, die Korrektheit des gegebenen Anteils fehlerhafter Geräte infrage zu stellen. 3
- e Die betrachtete Stichprobe enthält 19 fehlerhafte Geräte. Bestimmen Sie grafisch das zu dieser Anzahl gehörende Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %. 3

Lösungshinweise mit Variationen zu e)

Simulieren: 50 Simulationen mit  $p = 0,25$ ;  $n=100$ ;  $\gamma=0,95$



# IQB-Pool für das Jahr 2019 - PT B - WTR - Aufgabe 1d), 1e)

Kurz nach einer Änderung im Herstellungsverfahren stellt das Unternehmen den Anteil fehlerhafter Geräte von 25 % infrage. Um bei einer Sicherheitswahrscheinlichkeit von 95 % einen Schätzwert für den Anteil fehlerhafter Geräte zu ermitteln, wird eine Stichprobe von 100 Geräten betrachtet. Abbildung 1 zeigt die Graphen der folgenden für  $p \in [0,1]$  definierten Funktionen:

$$f: p \mapsto p - 1,96 \cdot \sqrt{\frac{p(1-p)}{100}} \quad g: p \mapsto p + 1,96 \cdot \sqrt{\frac{p(1-p)}{100}}$$

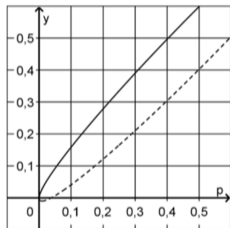


Abb. 1

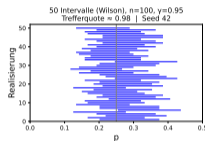
- d Bestimmen Sie grafisch alle möglichen Anzahlen fehlerhafter Geräte in der Stichprobe, für die bei einer Sicherheitswahrscheinlichkeit von 95 % jeweils Anlass dazu bestehen würde, die Korrektheit des gegebenen Anteils fehlerhafter Geräte infrage zu stellen.
- e Die betrachtete Stichprobe enthält 19 fehlerhafte Geräte. Bestimmen Sie grafisch das zu dieser Anzahl gehörende Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %.

3

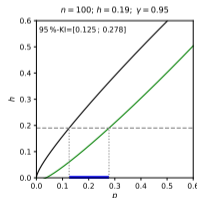
3

Lösungshinweise mit Variationen zu e)

Simulieren: 50 Simulationen mit  $p = 0,25$ ;  $n=100$ ;  $\gamma=0,95$



Zeichnen: Konfidenzellipse



# IQB-Pool für das Jahr 2019 - PT B - WTR - Aufgabe 1d), 1e)

Kurz nach einer Änderung im Herstellungsverfahren stellt das Unternehmen den Anteil fehlerhafter Geräte von 25 % infrage. Um bei einer Sicherheitswahrscheinlichkeit von 95 % einen Schätzwert für den Anteil fehlerhafter Geräte zu ermitteln, wird eine Stichprobe von 100 Geräten betrachtet. Abbildung 1 zeigt die Graphen der folgenden für  $p \in [0,1]$  definierten Funktionen:

$$f: p \mapsto p - 1,96 \cdot \sqrt{\frac{p(1-p)}{100}} \quad g: p \mapsto p + 1,96 \cdot \sqrt{\frac{p(1-p)}{100}}$$

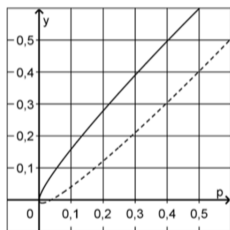


Abb. 1

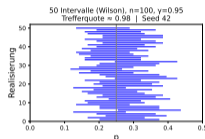
- d Bestimmen Sie grafisch alle möglichen Anzahlen fehlerhafter Geräte in der Stichprobe, für die bei einer Sicherheitswahrscheinlichkeit von 95 % jeweils Anlass dazu bestehen würde, die Korrektheit des gegebenen Anteils fehlerhafter Geräte infrage zu stellen.
- e Die betrachtete Stichprobe enthält 19 fehlerhafte Geräte. Bestimmen Sie grafisch das zu dieser Anzahl gehörende Konfidenzintervall zur Sicherheitswahrscheinlichkeit 95 %.

3

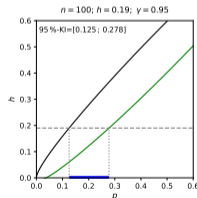
3

Lösungshinweise mit Variationen zu e)

Simulieren: 50 Simulationen mit  $p = 0,25$ ;  $n=100$ ;  $\gamma=0,95$



Zeichnen: Konfidenzellipse



Rechnen: 95 %-WILSON-Konfidenzintervall:

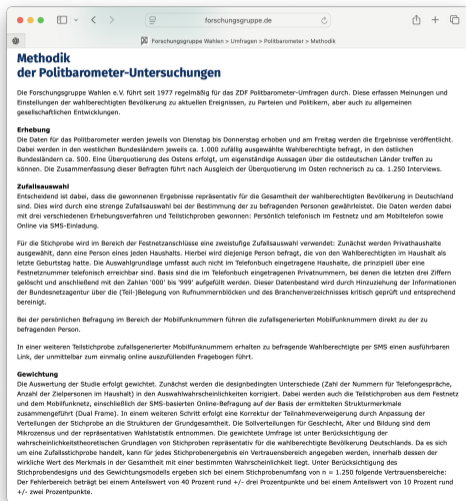
$$h = p + 1,96 \sqrt{\frac{p(1-p)}{n}}; p_l \approx 0,125$$

$$h = p - 1,96 \sqrt{\frac{p(1-p)}{n}}; p_r \approx 0,278$$

## Anwendungen - eine kleine Auswahl

---

# Methodik-Politbarometer, Forschungsgruppe Wahlen



The screenshot shows a web browser window with the URL 'forschungsgruppe.de'. The page title is 'Methodik der Politbarometer-Untersuchungen'. The content is organized into several sections:

- Methodik der Politbarometer-Untersuchungen**

Die Forschungsgruppe Wahlen e.V. führt seit 1977 regelmäßig für das ZDF Politbarometer-Umfragen durch. Diese erfassen Meinungen und Einstellungen der wahlberechtigten Bevölkerung zu aktuellen Ereignissen, zu Parteien und Politikern, aber auch zu allgemeinen gesellschaftlichen Entwicklungen.
- Erhebung**

Die Daten für das Politbarometer werden jeweils von Dienstag bis Donnerstag erhoben und am Freitag werden die Ergebnisse veröffentlicht. Dabei werden in den westlichen Bundesländern jeweils ca. 1.000 zufällig ausgewählte Wahlberechtigte befragt, in den östlichen Bundesländern ca. 500. Eine Überquotierung des Ostens erfolgt, um eigenständige Aussagen über die ostdeutschen Länder treffen zu können. Die Zusammenfassung dieser Befragten führt nach Ausgleich der Überquotierung im Osten rechnerisch zu ca. 1.250 Interviews.
- Zufallsauswahl**

Entscheidend ist dabei, dass die gewonnenen Ergebnisse repräsentativ für die Gesamtheit der wahlberechtigten Bevölkerung in Deutschland sind. Dies wird durch eine strenge Zufallsauswahl bei der Bestimmung der zu befragenden Personen gewährleistet. Die Daten werden dabei mit drei verschiedenen Erhebungsverfahren und Teilstichproben gewonnen: Persönlich telefonisch im Festnetz und am Mobiltelefon sowie Online via SMS-Einladung.

Für die Stichprobe wird im Bereich der Festnetzanschlüsse eine zweistufige Zufallsauswahl verwendet: Zunächst werden Privathaushalte ausgewählt, dann eine Person eines Jeden Haushalts. Hierbei wird diejenige Person befragt, die von den Wahlberechtigten im Haushalt als letzte Geburtstag hatte. Die Auswahlgrundlage umfasst auch nicht im Telefonbuch eingetragene Haushalte, die prinzipiell über eine Festnetznummer telefonisch erreichbar sind. Basis sind die im Telefonbuch eingetragenen Privatnummern, bei denen die letzten drei Ziffern gelöscht und anschließend mit den Zahlen '000' bis '999' aufgefüllt werden. Dieser Datenbestand wird durch Hinzuziehung der Informationen der Bundeszentralagentur über die (Teil-)Belegung von Rufnummernblöcken und des Branchenverzeichnisses kritisch geprüft und entsprechend bereinigt.

Bei der persönlichen Befragung im Bereich der Mobilfunknummern führen die zufallsgenerierten Mobilfunknummern direkt zu der zu befragenden Person.

In einer weiteren Teilstichprobe zufallsgenerierter Mobilfunknummern erhalten zu befragende Wahlberechtigte per SMS einen ausführbaren Link, der unmittelbar zum einmalig online auszufüllenden Fragebogen führt.
- Gewichtung**

Die Auswertung der Studie erfolgt gewichtet. Zunächst werden die designbedingten Unterschiede (Zahl der Nummern für Telefongespräche, Anzahl der Zielpersonen im Haushalt) in den Auswahlwahrscheinlichkeiten korrigiert. Dabei werden auch die Teilstichproben aus dem Festnetz und dem Mobilfunknetz, einschließlich der SMS-basierten Online-Befragung auf der Basis der ermittelten Strukturmerkmale zusammengeführt (Dual Frame). In einem weiteren Schritt erfolgt eine Korrektur der Teilnahmeverweigerung durch Anpassung der Verteilungen der Stichprobe an die Strukturen der Grundgesamtheit. Die Sollverteilungen für Geschlecht, Alter und Bildung sind dem Mikrozensus und der repräsentativen Wahlstatistik entnommen. Die gewichtete Umfrage ist unter Berücksichtigung der wahrscheinlichkeitstheoretischen Grundlagen von Stichproben repräsentativ für die wahlberechtigte Bevölkerung Deutschlands. Da es sich um eine Zufallsstichprobe handelt, kann für jedes Stichprobenergebnis ein Vertrauensbereich angegeben werden, innerhalb dessen der wirkliche Wert des Merkmals in der Gesamtheit mit einer bestimmten Wahrscheinlichkeit liegt. Unter Berücksichtigung des Stichprobendesigns und des Gewichtungmodells ergeben sich bei einem Stichprobenumfang von  $n = 1.250$  folgende Vertrauensbereiche: Der Fehlerbereich beträgt bei einem Anteilswert von 40 Prozent rund +/- drei Prozentpunkte und bei einem Anteilswert von 10 Prozent rund +/- zwei Prozentpunkte.

# Forschungsgruppe Wahlen: Vertrauensbereiche in der Praxis

- Dual-Frame-Stichprobe (Festnetz, Mobilfunk, SMS-Online): unterschiedliche Auswahlwahrscheinlichkeiten.
- **Gewichtung** korrigiert
  - designbedingte Unterschiede (z. B. Haushaltsgröße, Zahl der Rufnummern),
  - und **Nonresponse** durch Anpassung an Strukturen des Mikrozensus (Geschlecht, Alter, Bildung).
- Ergebnis: Die gewichtete Stichprobe ist – *im Modell* – **repräsentativ** für die Grundmenge.
- Da eine modellierte Zufallsstichprobe zugrunde liegt, werden **Vertrauensbereiche** angegeben:
  - bei 40 % etwa  $\pm 3$  Prozentpunkte, bei 10 %: etwa  $\pm 2$  Prozentpunkte.

## Designfaktor: Modell $\rightarrow$ Realität

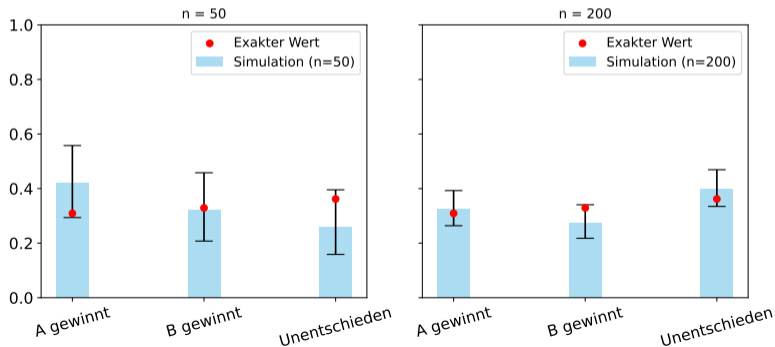
Gewichtung bläht die Varianz auf. Praktisch arbeitet man mit einem *effektiven Stichprobenumfang*  
 $n_{\text{eff}} = n / \text{Designfaktor}$ .

# Simulationen ohne Konfidenzintervalle - wertlos

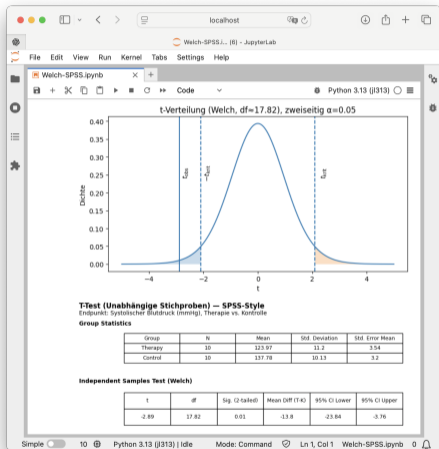
Das Spiel *Differenz trifft*: Die Proportionalstrategie ist nicht die „optimale“ Strategie

$$A = (3, 5, 4, 3, 2, 1); \quad B = (3, 7, 4, 3, 1, 0); \quad p = \left(\frac{3}{18}, \frac{5}{18}, \frac{4}{18}, \frac{3}{18}, \frac{2}{18}, \frac{1}{18}\right)$$

Simulation vs. exakte Werte für verschiedene Stichprobenumfänge

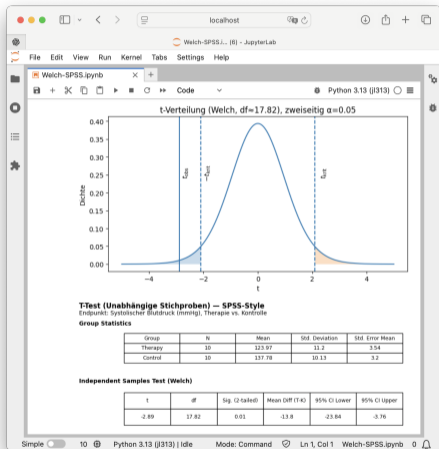


# Anwendung in der Medizin: Welch-t-Test



Teststatistik, Verteilung, kritischer Bereich.

# Anwendung in der Medizin: Welch-t-Test



Der Welch-t-Test (zwei unabhängige Gruppen) - mit SPSS-Ausgabe

Wir simulieren systolische Blutdruckdaten (mmHg) für zwei unabhängige Gruppen (Therapie vs. Kontrolle) und testen die Nullhypothese

$$H_0: \mu_{\text{Therapie}} - \mu_{\text{Kontrolle}} = 0$$

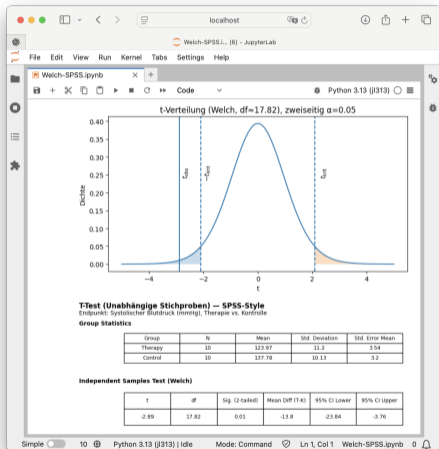
mit dem Welch-t-Test, der keine Varianzgleichheit voraussetzt. Die Teststatistik

$$t = \frac{\bar{x}_T - \bar{x}_K}{\sqrt{\frac{s_T^2}{n_T} + \frac{s_K^2}{n_K}}}$$

misst den beobachteten Mittelwertunterschied relativ zu seiner zufälligen Streuung. Die (nicht-ganzzahligen) Freiheitsgrade werden über die Welch-Satterthwaite-Approximation bestimmt; Entscheidungsstruktur und Interpretation bleiben identisch zum klassischen t-Test.

Teststatistik, Verteilung, kritischer Bereich.

# Anwendung in der Medizin: Welch-t-Test



Teststatistik, Verteilung, kritischer Bereich.

**Der Welch-t-Test (zwei unabhängige Gruppen) - mit SPSS-Ausgabe**

Wir simulieren systolische Blutdruckdaten (mmHg) für zwei unabhängige Gruppen (Therapie vs. Kontrolle) und testen die Nullhypothese

$$H_0: \mu_{\text{Therapie}} - \mu_{\text{Kontrolle}} = 0$$

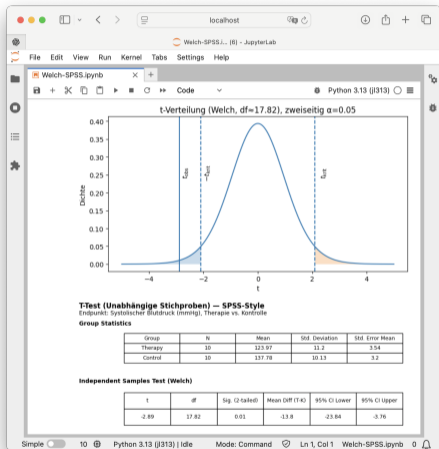
mit dem **Welch-t-Test**, der keine Varianzgleichheit voraussetzt. Die Teststatistik

$$t = \frac{\bar{x}_T - \bar{x}_K}{\sqrt{\frac{s_T^2}{n_T} + \frac{s_K^2}{n_K}}}$$

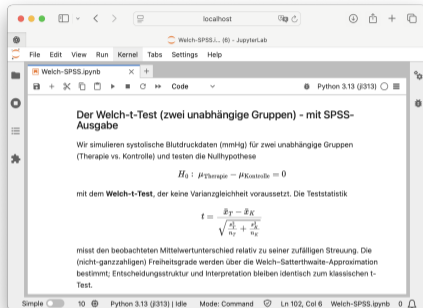
misst den beobachteten Mittelwertunterschied relativ zu seiner zufälligen Streuung. Die (nicht-ganzzahligen) Freiheitsgrade werden über die Welch-Satterthwaite-Approximation bestimmt; Entscheidungsstruktur und Interpretation bleiben identisch zum klassischen t-Test.

Gleiche Struktur wie beim Hypothesentest:  
Teststatistik, Referenzverteilung, Entscheidungsregel.

# Anwendung in der Medizin: Welch-t-Test



Teststatistik, Verteilung, kritischer Bereich.



Gleiche Struktur wie beim Hypothesentest:

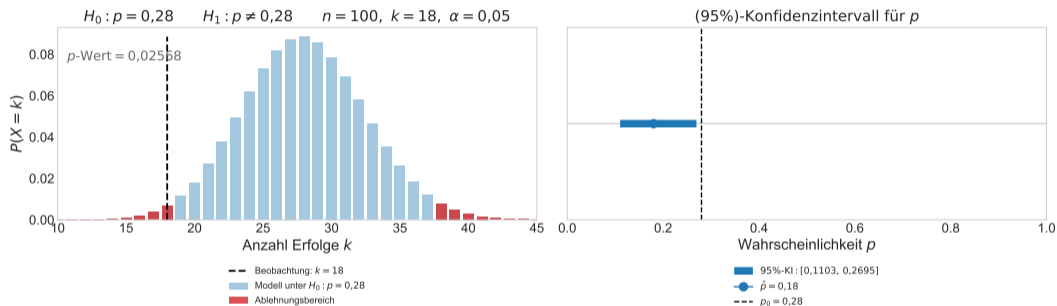
- Teststatistik, Referenzverteilung, Entscheidungsregel.
- Kein neues Prinzip – nur andere Sprache (SPSS). Test und Konfidenzintervall sind zwei Seiten derselben Struktur.

# Hypothesentest und Konfidenzintervalle - zwei Seiten einer Medaille

---

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

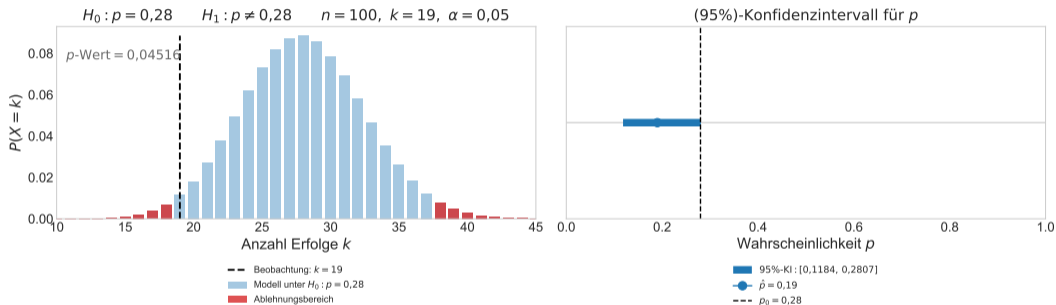


Beobachtung:  $k = 18$

Der Test auf Niveau  $\alpha$  verwirft  $H_0 \iff p_0 \notin KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

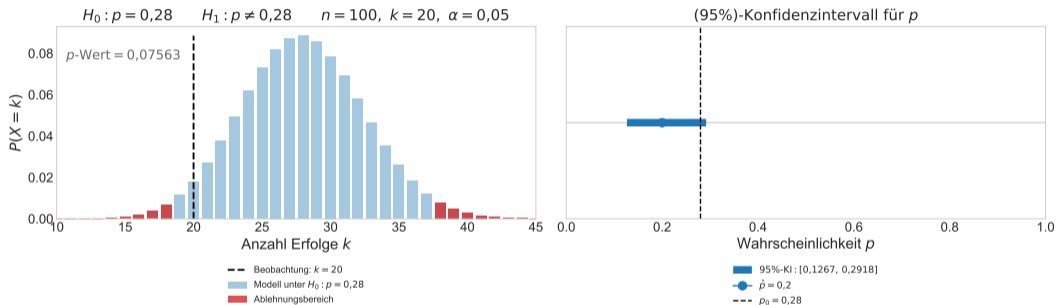


Beobachtung:  $k = 19$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

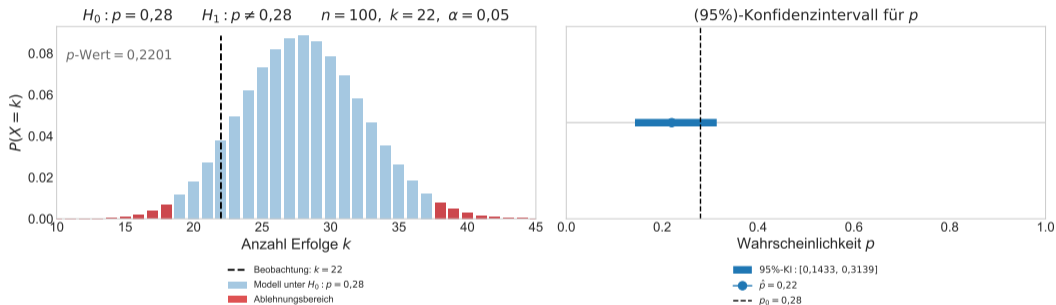


Beobachtung:  $k = 20$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

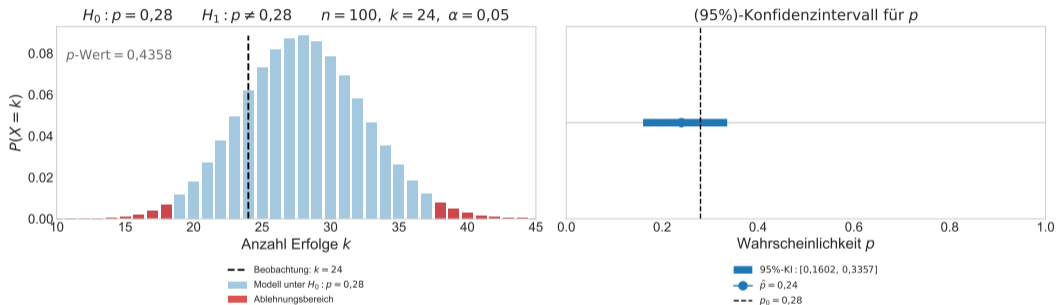


Beobachtung:  $k = 22$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

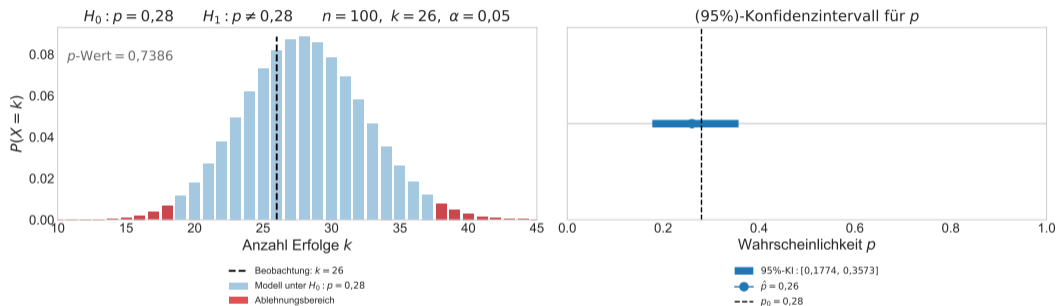


Beobachtung:  $k = 24$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

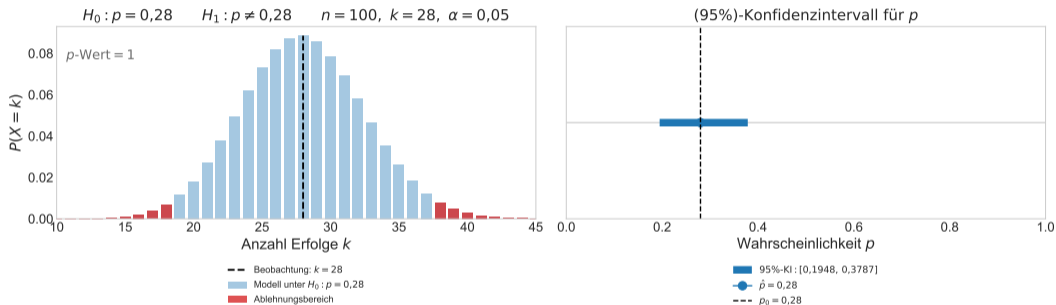


Beobachtung:  $k = 26$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

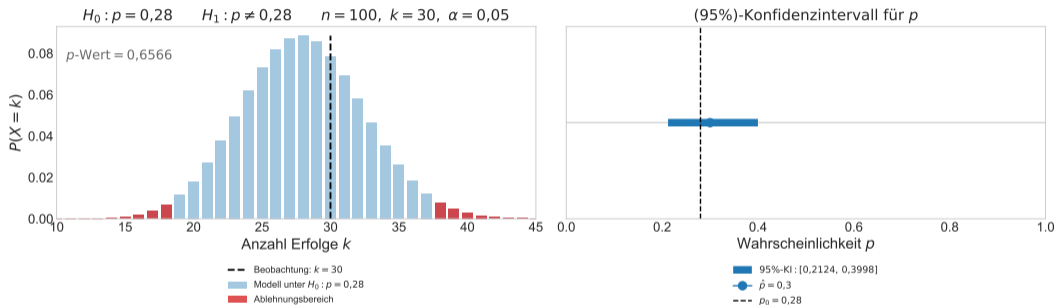


Beobachtung:  $k = 28$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

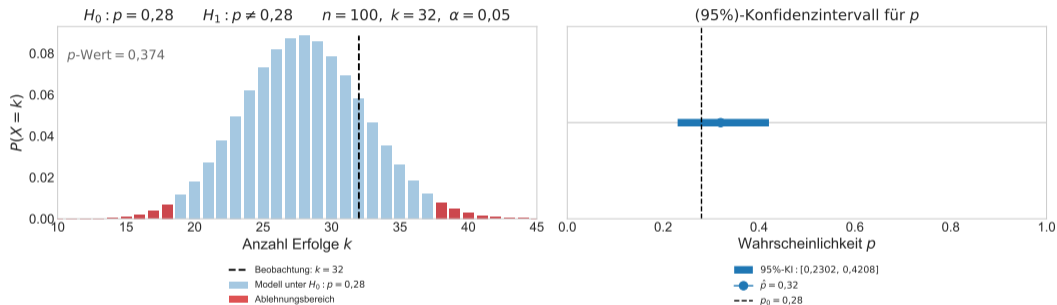


Beobachtung:  $k = 30$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

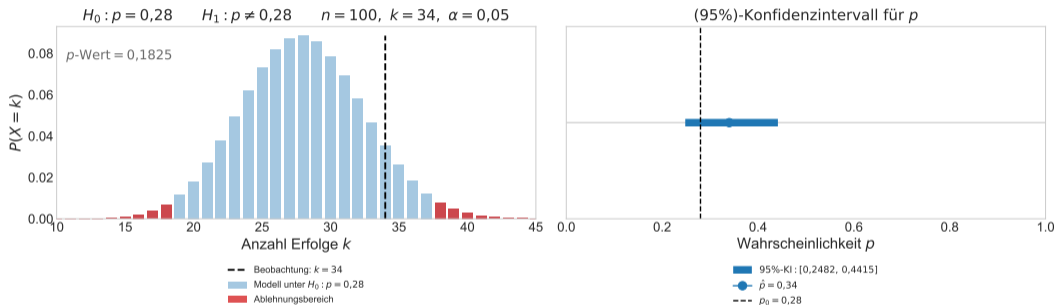


Beobachtung:  $k = 32$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

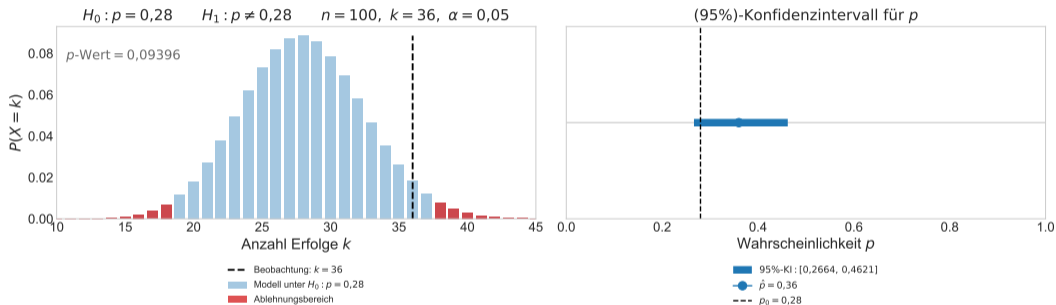


Beobachtung:  $k = 34$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

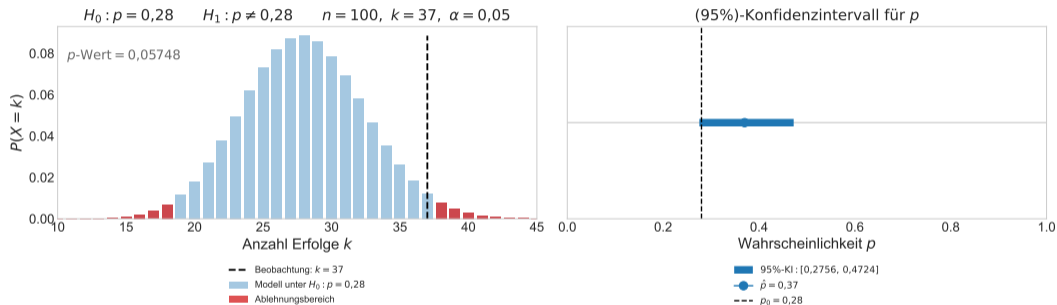


Beobachtung:  $k = 36$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

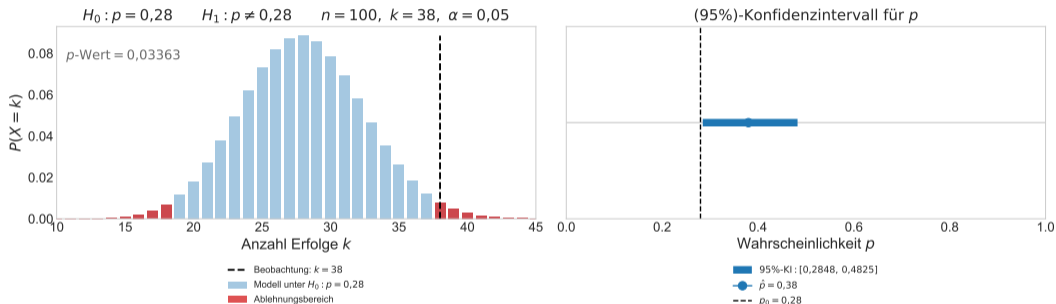


Beobachtung:  $k = 37$

Der Test auf Niveau  $\alpha$  verwirft nicht  $H_0 \iff p_0 \in KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

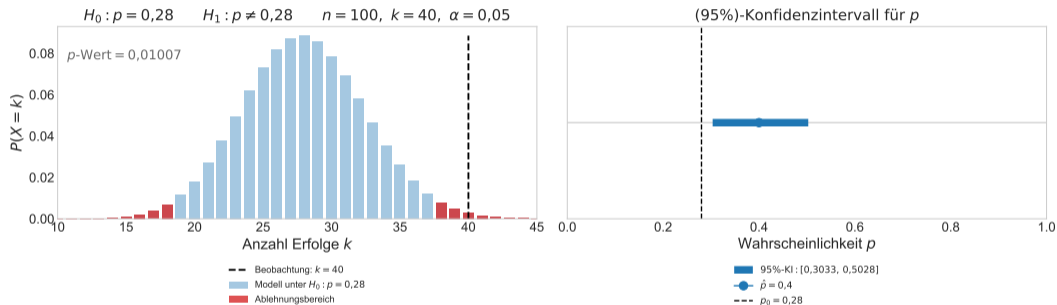


Beobachtung:  $k = 38$

Der Test auf Niveau  $\alpha$  verwirft  $H_0 \iff p_0 \notin KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI

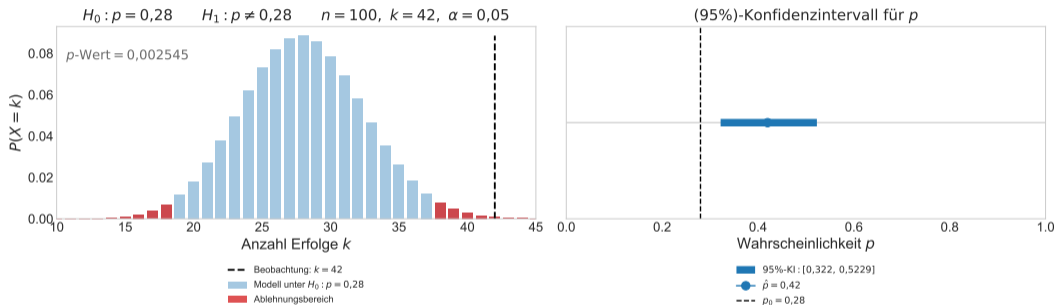


Beobachtung:  $k = 40$

Der Test auf Niveau  $\alpha$  verwirft  $H_0 \iff p_0 \notin KI$

# Zweiseitiger Test und Konfidenzintervall – 15 Situationen

## Zweiseitiger Binomialtest und Clopper–Pearson-KI



Beobachtung:  $k = 42$

Der Test auf Niveau  $\alpha$  verwirft  $H_0 \iff p_0 \notin KI$

Ein zweiseitiger Test zum Niveau  $\alpha$  verwirft  $H_0: p = p_0$



$p_0 \notin (1 - \alpha)$ -Konfidenzintervall.

## **Hypothesentest**

Beim Hypothesentest liegt der Zollstock fest –  
und die Stichprobe bewegt sich.

## **Konfidenzintervall**

Beim Konfidenzintervall liegt die Stichprobe fest –  
und wir verschieben den Zollstock, um ein Intervall zu konstruieren,  
dass sich bewegt.

In der Realität bleibt die Stichprobe gleich.

Die Bewegung bezieht sich auf gedankliche Wiederholungen.

Sicht von oben.

Sicht von oben.

**Konfidenzintervalle** beschreiben Verfahrenseigenschaften bei Wiederholung, nicht die Unsicherheit eines konkreten Parameters.

Sicht von oben.

**Konfidenzintervalle** beschreiben Verfahrenseigenschaften bei Wiederholung, nicht die Unsicherheit eines konkreten Parameters.

**Hypothesentests** beschreiben Verfahrenseigenschaften bei Wiederholung, nicht die Wahrscheinlichkeit einer konkreten Hypothese.

Sicht von oben.

**Konfidenzintervalle** beschreiben Verfahrenseigenschaften bei Wiederholung, nicht die Unsicherheit eines konkreten Parameters.

**Hypothesentests** beschreiben Verfahrenseigenschaften bei Wiederholung, nicht die Wahrscheinlichkeit einer konkreten Hypothese.

Etwas konkreter.

Sicht von oben.

**Konfidenzintervalle** beschreiben Verfahrenseigenschaften bei Wiederholung, nicht die Unsicherheit eines konkreten Parameters.

**Hypothesentests** beschreiben Verfahrenseigenschaften bei Wiederholung, nicht die Wahrscheinlichkeit einer konkreten Hypothese.

Etwas konkreter.

Ein **Konfidenzintervall** beschreibt die Menge aller Parameterwerte, die mit den beobachteten Daten vereinbar sind, bei einem gegebenen Sicherheitsniveau.

Sicht von oben.

**Konfidenzintervalle** beschreiben Verfahrenseigenschaften bei Wiederholung, nicht die Unsicherheit eines konkreten Parameters.

**Hypothesentests** beschreiben Verfahrenseigenschaften bei Wiederholung, nicht die Wahrscheinlichkeit einer konkreten Hypothese.

Etwas konkreter.

Ein **Konfidenzintervall** beschreibt die Menge aller Parameterwerte, die mit den beobachteten Daten vereinbar sind, bei einem gegebenen Sicherheitsniveau.

Ein **Hypothesentest** ist ein Schnitt durch dieses Intervall.

## Rückblick und Ausblick

---

## Worum es mir heute wirklich geht

### Zentrale Leitfrage

Was bedeutet es, wenn wir aus einer Stichprobe auf die Grundgesamtheit schließen?

# Worum es mir heute wirklich geht

## Zentrale Leitfrage

Was bedeutet es, wenn wir aus einer Stichprobe auf die Grundgesamtheit schließen?

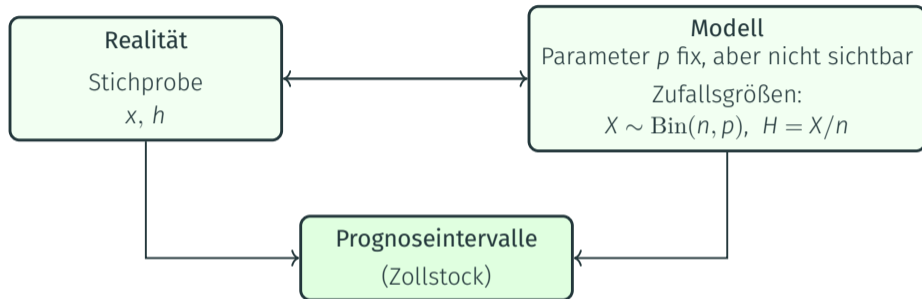
## Vier Eckpfeiler meines Ansatzes

- Unterscheide streng: **Modell** vs. **Realisation**.
- Unterscheide streng: **Zufallsgröße  $X$  (Modell)** vs. **Realisation  $x$  (Realität)**.
- Nutze: **Handlungen, Bilder und Gedankenexperimente**, um Unsicherheit verständlich zu machen - vor dem Erlernen formaler Verfahren.
- Suche: **Anschlussfähigkeit zur Universität**, ohne die formale Bildungsbarriere unnötig zu erhöhen - Strukturen sichtbar machen, nicht die technische Oberfläche.

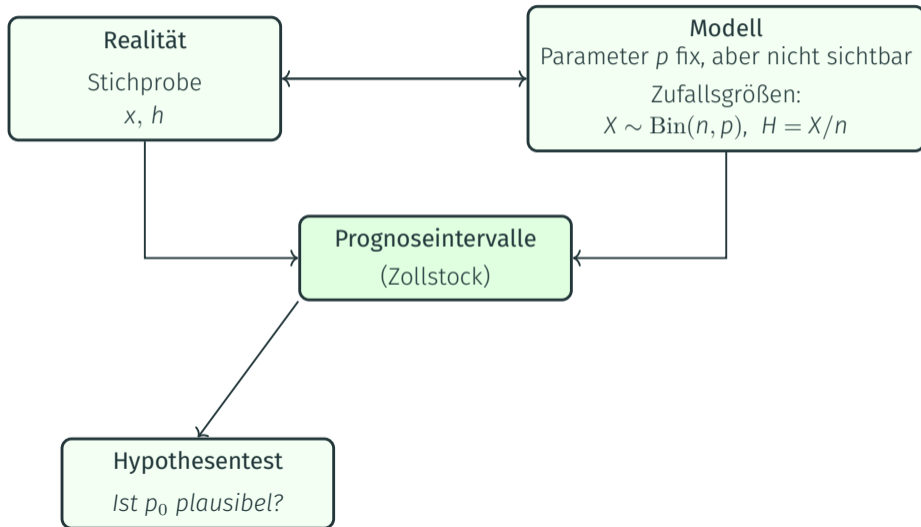
## Zwei Seiten einer Medaille: Realität, Modell, Modelle



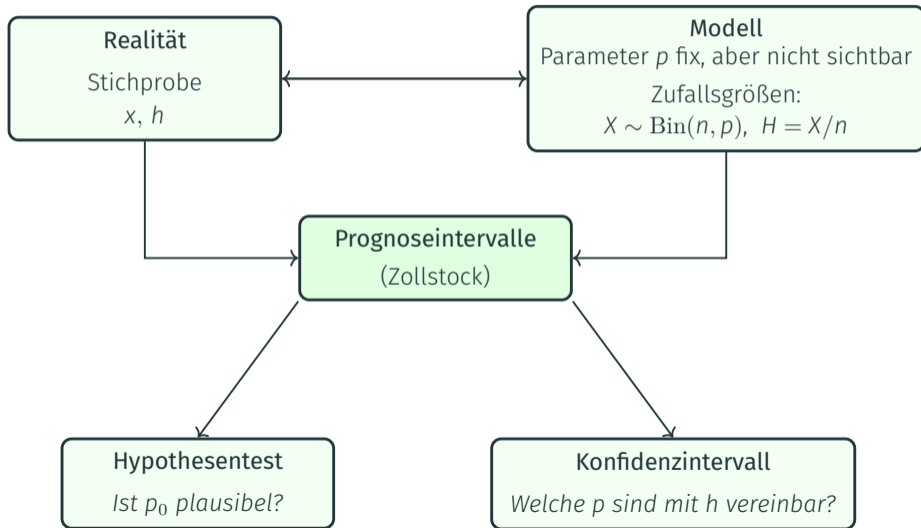
## Zwei Seiten einer Medaille: Realität, Modell, Modelle



## Zwei Seiten einer Medaille: Realität, Modell, Modelle



# Zwei Seiten einer Medaille: Realität, Modell, Modelle



Das Fundament in drei Sätzen.

## Das Fundament in drei Sätzen.

- Modelle sind formale Ersatzwelten.
- Tests und Konfidenzintervalle bewerten Daten, nicht Wahrheit.
- Zufallsvariablen leben im Gedankenexperiment, Daten in der Realität.

## Das Fundament in drei Sätzen.

- Modelle sind formale Ersatzwelten.
- Tests und Konfidenzintervalle bewerten Daten, nicht Wahrheit.
- Zufallsvariablen leben im Gedankenexperiment, Daten in der Realität.

## Beispiel: Münzwurf

## Das Fundament in drei Sätzen.

- Modelle sind formale Ersatzwelten.
- Tests und Konfidenzintervalle bewerten Daten, nicht Wahrheit.
- Zufallsvariablen leben im Gedankenexperiment, Daten in der Realität.

## Beispiel: Münzwurf

- Wir beobachten die Realität - hier sind es die Häufigkeiten
- Wir legen ein Modell fest - und damit Wahrscheinlichkeiten.
- Wir rechnen strikt im Modell.
- Die Realität ist unser Prüfstein, nicht unser Rechenraum.
- Kurz und knackig: Häufigkeiten beobachten wir, Wahrscheinlichkeiten legen wir fest.

## Das Fundament in drei Sätzen.

- Modelle sind formale Ersatzwelten.
- Tests und Konfidenzintervalle bewerten Daten, nicht Wahrheit.
- Zufallsvariablen leben im Gedankenexperiment, Daten in der Realität.

## Beispiel: Münzwurf

- Wir beobachten die Realität - hier sind es die Häufigkeiten
- Wir legen ein Modell fest - und damit Wahrscheinlichkeiten.
- Wir rechnen strikt im Modell.
- Die Realität ist unser Prüfstein, nicht unser Rechenraum.
- Kurz und knackig: Häufigkeiten beobachten wir, Wahrscheinlichkeiten legen wir fest.

*Daten können ein Modell stützen oder infrage stellen – sie können es nicht wahr machen.  
Modelle können sich bewähren – sie werden dadurch aber nicht wahr.*

## Zwei Seiten einer Medaille



## Zwei Seiten einer Medaille

*Was beide verbindet: dieselbe Stichprobe, dieselben Daten – nur die andere Frage.*



# Zwei Seiten einer Medaille

*Was beide verbindet: dieselbe Stichprobe, dieselben Daten – nur die andere Frage.*



*Der Test auf Niveau  $\alpha$  verwirft  $H_0 \iff p_0 \notin KI$*

These: Anschlussfähigkeit entsteht durch Ehrlichkeit.

**These: Anschlussfähigkeit entsteht durch Ehrlichkeit.**

Universitäre Anschlussfähigkeit heißt nicht:

- mehr Formeln
- mehr Beweise
- mehr Symbole

**These: Anschlussfähigkeit entsteht durch Ehrlichkeit.**

Universitäre Anschlussfähigkeit heißt nicht:

- mehr Formeln
- mehr Beweise
- mehr Symbole

Sondern hier:

- saubere Begriffe
- korrekte Objekt-Ebenen
- keine stillschweigenden Bedeutungswechsel

**These: Anschlussfähigkeit entsteht durch Ehrlichkeit.**

Universitäre Anschlussfähigkeit heißt nicht:

- mehr Formeln
- mehr Beweise
- mehr Symbole

Sondern hier:

- saubere Begriffe
- korrekte Objekt-Ebenen
- keine stillschweigenden Bedeutungswechsel

Konkret:

- $p$  bleibt Parameter
- Intervalle sind Zufallsobjekte
- Wahrscheinlichkeit bezieht sich auf Verfahren
- Bilder und Code widersprechen der Theorie nicht

These: Stoff wird nicht einfacher, wenn man ihn weichzeichnet.

**These: Stoff wird nicht einfacher, wenn man ihn weichzeichnet.**

Vermeide:

- „Stell dir das so vor ...“
- „Im Grunde heißt das ...“
- „Das ist jetzt nicht ganz korrekt, aber ...“

**These: Stoff wird nicht einfacher, wenn man ihn weichzeichnet.**

Vermeide:

- „Stell dir das so vor ...“
- „Im Grunde heißt das ...“
- „Das ist jetzt nicht ganz korrekt, aber ...“

Stattdessen:

- „Wir betrachten im Modell ...“
- „Wir konstruieren folgendes Objekt ...“
- „Im Gedankenexperiment ...“
- „Diese Grafik repräsentiert das Modell, nicht die Realität““

**These: Stoff wird nicht einfacher, wenn man ihn weichzeichnet.**

Vermeide:

- „Stell dir das so vor ...“
- „Im Grunde heißt das ...“
- „Das ist jetzt nicht ganz korrekt, aber ...“

Stattdessen:

- „Wir betrachten im Modell ...“
- „Wir konstruieren folgendes Objekt ...“
- „Im Gedankenexperiment ...“
- „Diese Grafik repräsentiert das Modell, nicht die Realität““

*Das entlastet kognitiv, selbst wenn der Anspruch höher ist.  
Das ist nicht schwerer – nur ehrlicher.*