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

Zwei Seiten einer Medaille: Realität, Modell, Modelle

Realität

Stichprobe

x, h

$\longleftrightarrow$

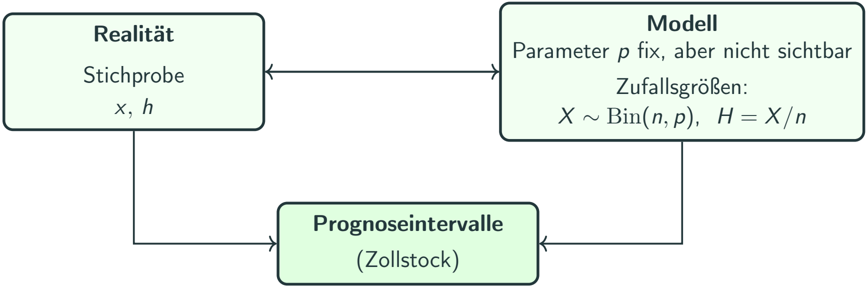
Modell

Parameter p fix, aber nicht sichtbar

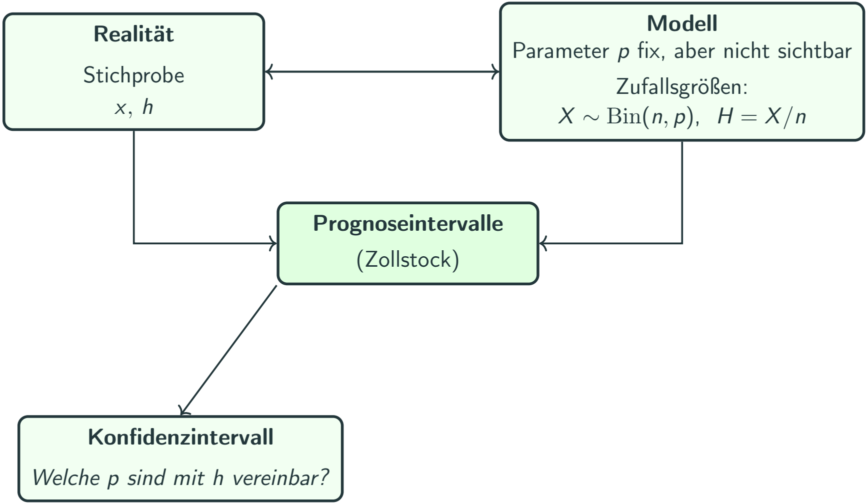
Zufallsgrößen:

$X \sim \mathrm{Bin}(n, p), \quad H = X/n$

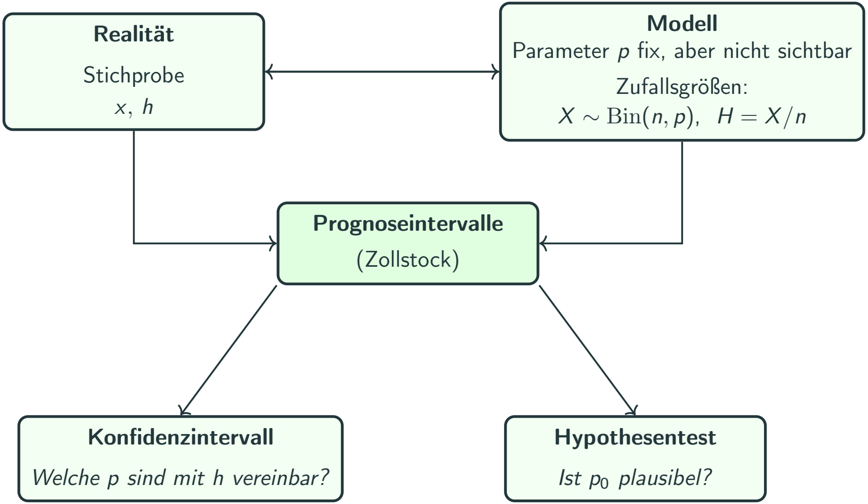
Zwei Seiten einer Medaille: Realität, Modell, Modelle



Zwei Seiten einer Medaille: Realität, Modell, Modelle



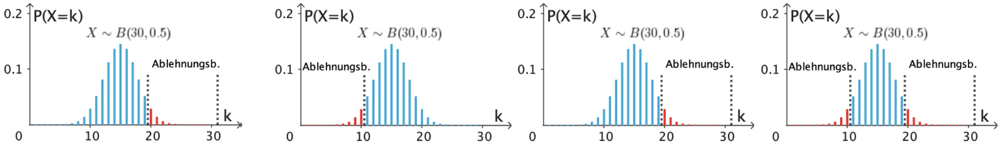
Zwei Seiten einer Medaille: Realität, Modell, Modelle



Hypothesentests - eine Übersicht

- **Alternativtest**
- **Signifikanztest**
  - einseitiger Test
    - linksseitiger Test
    - rechtsseitiger Test
  - zweiseitiger Test

Alternativtest		linksseitiger Test		rechtsseitiger Test		zweiseitiger Test	
H_a	H_b	H_0	H_1	H_0	H_1	H_0	H_1
$p = p_0$	$p = p_1$	$p \geq p_0$	$p < p_0$	$p \leq p_0$	$p > p_0$	$p = p_0$	$p \neq p_0$



Testen bewertet die Daten – nicht die Wahrheit

Kernidee: Der Test bewertet nicht die Wahrheit einer Hypothese, sondern nur, wie gut die beobachteten Daten zu dieser Hypothese passen. Die Hypothese selbst ist vor der Ziehung entweder wahr oder falsch; der Test kann das nicht ändern. Formal geschieht dies durch das Gedankenexperiment:

„Wie ungewöhnlich wären solche Daten, *wenn* H_0 gegolten hätte?“

Der Test beurteilt also *nur die Daten* im Licht des Modells, nicht die Hypothese selbst.

Testen bewertet die Daten – nicht die Wahrheit

Kernidee: Der Test bewertet nicht die Wahrheit einer Hypothese, sondern nur, wie gut die beobachteten Daten zu dieser Hypothese passen. Die Hypothese selbst ist vor der Ziehung entweder wahr oder falsch; der Test kann das nicht ändern. Formal geschieht dies durch das Gedankenexperiment:

„Wie ungewöhnlich wären solche Daten, *wenn* H_0 gegolten hätte?“

Der Test beurteilt also *nur die Daten* im Licht des Modells, nicht die Hypothese selbst.

Kurz und prägnant: Der Test urteilt nicht über die Wahrheit einer Hypothese, sondern über die *Kompatibilität der Daten* mit dieser Hypothese. Die Hypothese ist vor der Stichprobe entweder wahr oder falsch; der Test bewertet nur, ob die beobachteten Daten unter H_0 typisch oder ungewöhnlich wären.

Testen bewertet die Daten – nicht die Wahrheit

Kernidee: Der Test bewertet nicht die Wahrheit einer Hypothese, sondern nur, wie gut die beobachteten Daten zu dieser Hypothese passen. Die Hypothese selbst ist vor der Ziehung entweder wahr oder falsch; der Test kann das nicht ändern.
Formal geschieht dies durch das Gedankenexperiment:

„Wie ungewöhnlich wären solche Daten, *wenn* H_0 gegolten hätte?“

Der Test beurteilt also *nur die Daten* im Licht des Modells, nicht die Hypothese selbst.

Kurz und prägnant: Der Test urteilt nicht über die Wahrheit einer Hypothese, sondern über die *Kompatibilität der Daten* mit dieser Hypothese. Die Hypothese ist vor der Stichprobe entweder wahr oder falsch; der Test bewertet nur, ob die beobachteten Daten unter H_0 typisch oder ungewöhnlich wären.

Zeitliche Ordnung

1. Vor der Stichprobenziehung ist H_0 entweder wahr oder falsch.
2. Danach wird eine Stichprobe gezogen.
3. Der Test berechnet im Gedankenexperiment, wie häufig Daten wie die beobachteten unter der Annahme H_0 auftreten würden.

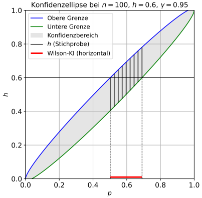
Der Test trifft keine Aussage über die Wahrheit von H_0, sondern bewertet ausschließlich die Daten anhand eines vorher festgelegten Regelwerks.

Konfidenzintervalle - WILSON, WALD und Clopper-Pearson

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

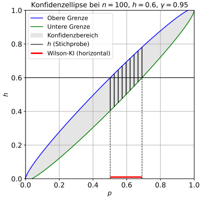
Konfidenzintervalle - WILSON, WALD und Clopper-Pearson

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



Konfidenzintervalle - WILSON, WALD und Clopper-Pearson

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

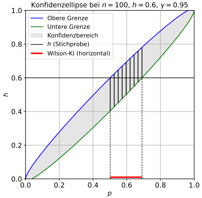


WALD: Approximation der Lösungen von WILSON; in GeoGebra und Rechnern implementiert

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

Konfidenzintervalle - WILSON, WALD und Clopper-Pearson

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



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 $\mathbb{P}_p(p \in \mathrm{KI}) \geq 1-\alpha$ für alle p.
Grenzen über F-Verteilung

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

Zufallsgröße vs. Realisation

Die zentrale Unterscheidung:

- Zufallsgröße $\Rightarrow$ Gedankenexperiment *vor der Ziehung*.
- Realisation $\Rightarrow$ konkretes Ergebnis *nach der Ziehung*.

Zufallsgröße vs. Realisation

Die zentrale Unterscheidung:

- Zufallsgröße $\Rightarrow$ Gedankenexperiment *vor der Ziehung*.
- Realisation $\Rightarrow$ konkretes Ergebnis *nach der Ziehung*.

Zufallsgröße

- Gedankenexperiment vor der Stichprobe
- beschreibt alle möglichen Ausgänge
- Beispiel: $H = X/n$
- „Alle denkbaren Werte von H“

Zufallsgröße vs. Realisation

Die zentrale Unterscheidung:

- Zufallsgröße $\Rightarrow$ Gedankenexperiment *vor der Ziehung*.
- Realisation $\Rightarrow$ konkretes Ergebnis *nach der Ziehung*.

Zufallsgröße

- Gedankenexperiment vor der Stichprobe
- beschreibt alle möglichen Ausgänge
- Beispiel: $H = X/n$
- „Alle denkbaren Werte von H"

Realisation

- konkreter Wert nach der Ziehung
- einer der möglichen Ausgänge
- Beispiel: $h = x/n$
- „Der tatsächlich beobachtete Wert"

Zufallsgröße vs. Realisation

Die zentrale Unterscheidung:

- Zufallsgröße $\Rightarrow$ Gedankenexperiment *vor der Ziehung*.
- Realisation $\Rightarrow$ konkretes Ergebnis *nach der Ziehung*.

Zufallsgröße

- Gedankenexperiment vor der Stichprobe
- beschreibt alle möglichen Ausgänge
- Beispiel: $H = X/n$
- „Alle denkbaren Werte von H“

Realisation

- konkreter Wert nach der Ziehung
- einer der möglichen Ausgänge
- Beispiel: $h = x/n$
- „Der tatsächlich beobachtete Wert“

Merksatz: Zufallsgröße $\Rightarrow$ Modell, Realisation $\Rightarrow$ Daten.

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 und Prognoseintervalle bereit
- lebt *vor* der Stichprobenziehung

Trenne Modell und Realität messerscharf!

Modellwelt

- beschreibt Zufallsgrößen (z. B. $H = X/n$)
- stellt Wahrscheinlichkeiten 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

Leitprinzip der gesamten Inferenz

Trenne Modell und Realität messerscharf!

Modellwelt

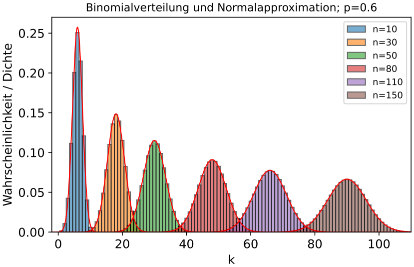
- beschreibt Zufallsgrößen (z. B. $H = X/n$)
- stellt Wahrscheinlichkeiten und Prognoseintervalle bereit
- lebt *vor* der Stichprobenziehung

Datenwelt

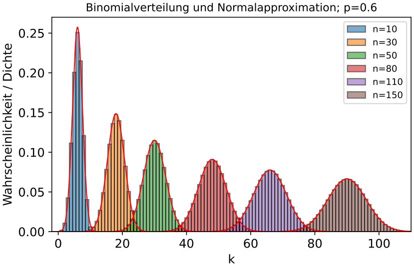
- enthält die beobachtete Realisation (z. B. $h = x/n$)
- ist das, was tatsächlich passiert ist
- beginnt *nach* der Stichprobenziehung

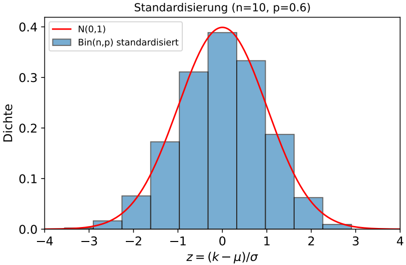
Statistik = präzises Wechseln zwischen beiden Welten.

Der Zollstock als rücktransformierte Normalapproximation

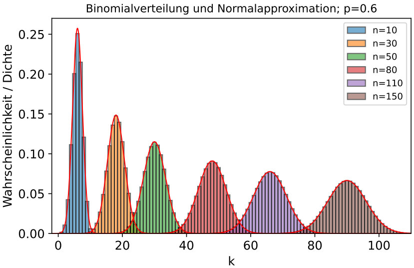


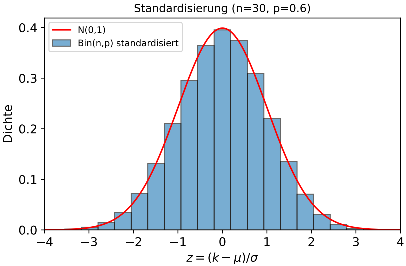
Der Zollstock als rücktransformierte Normalapproximation



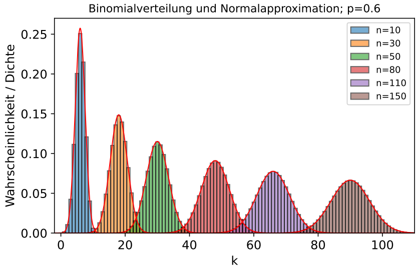


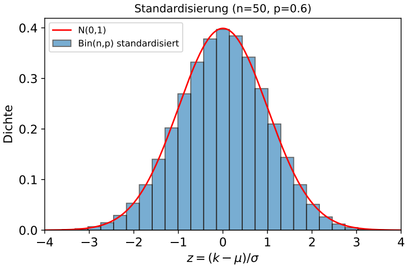
Der Zollstock als rücktransformierte Normalapproximation



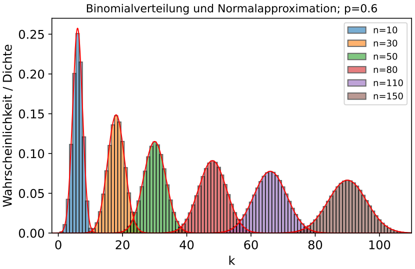


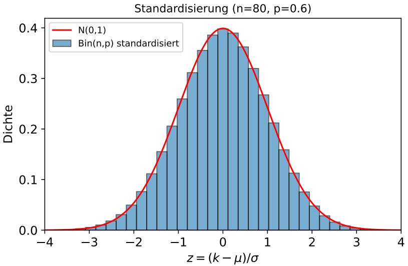
Der Zollstock als rücktransformierte Normalapproximation



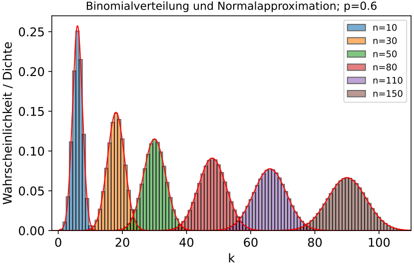


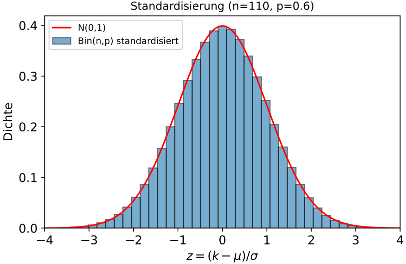
Der Zollstock als rücktransformierte Normalapproximation



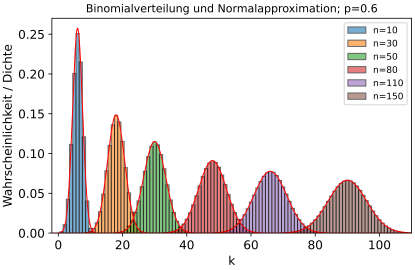


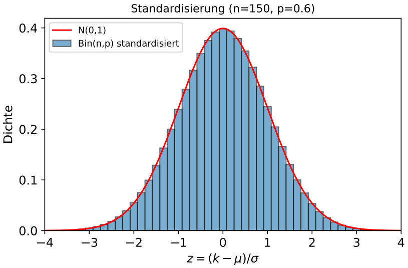
Der Zollstock als rücktransformierte Normalapproximation



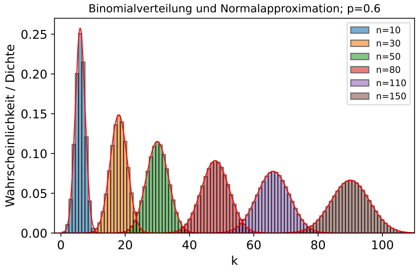


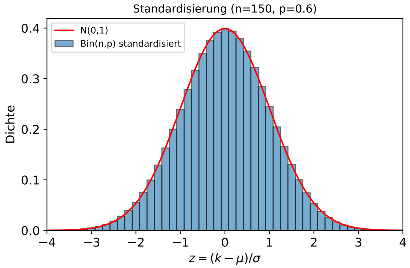
Der Zollstock als rücktransformierte Normalapproximation

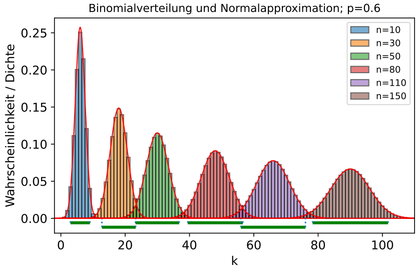




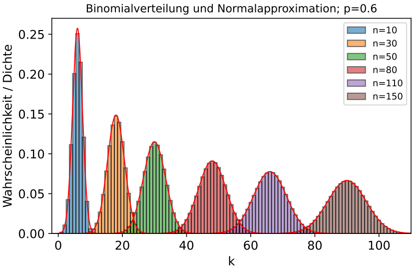
Der Zollstock als rücktransformierte Normalapproximation

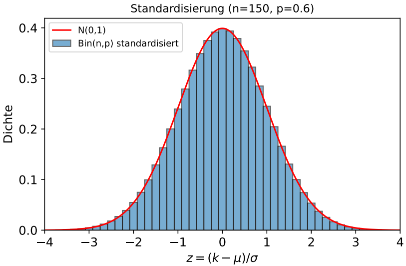


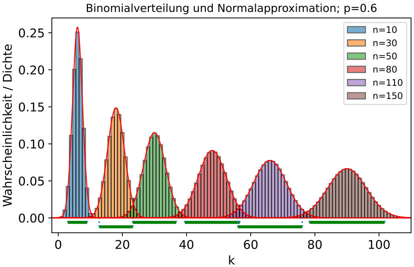


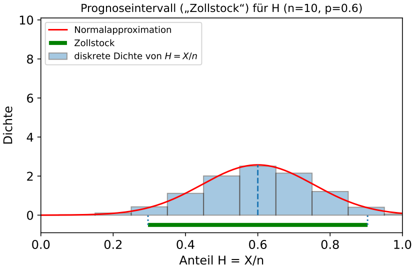


Der Zollstock als rücktransformierte Normalapproximation

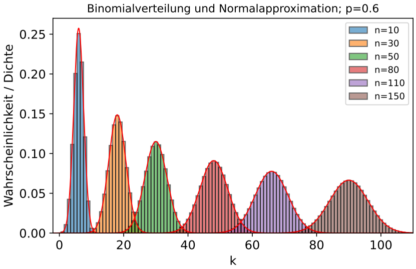


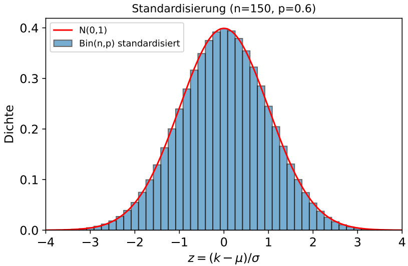


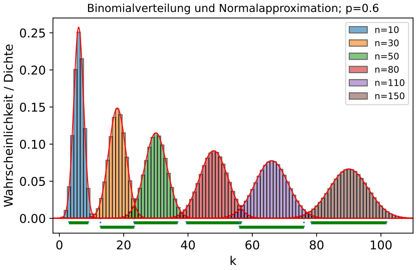


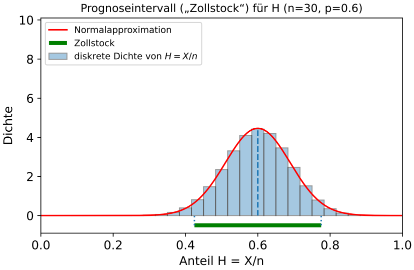


Der Zollstock als rücktransformierte Normalapproximation

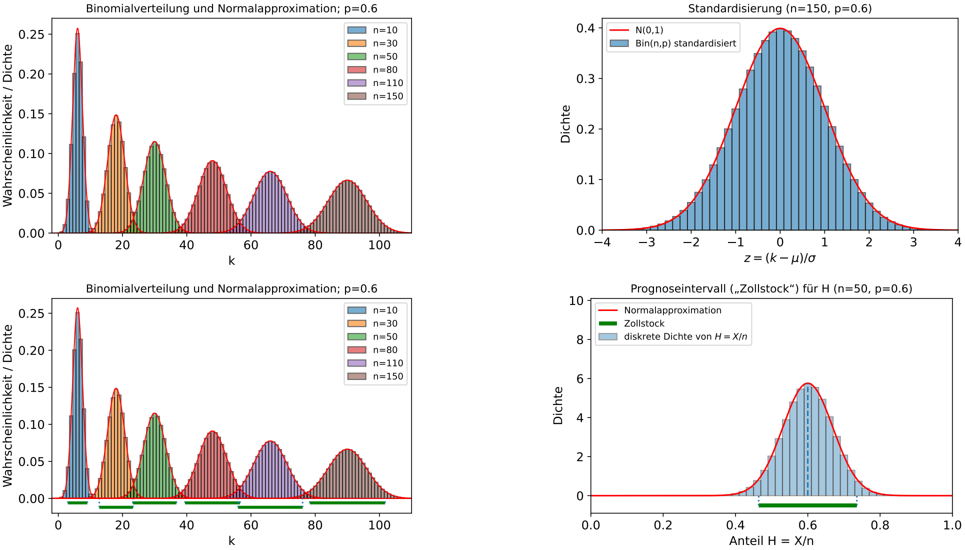




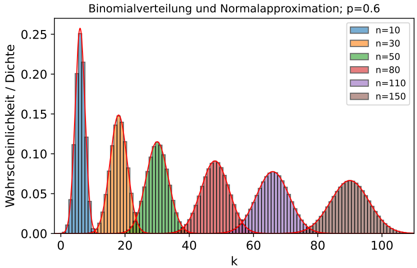


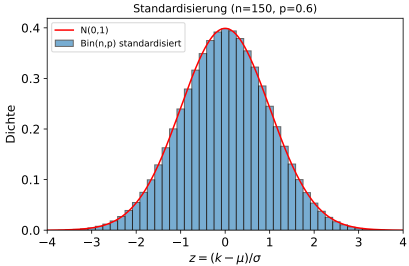


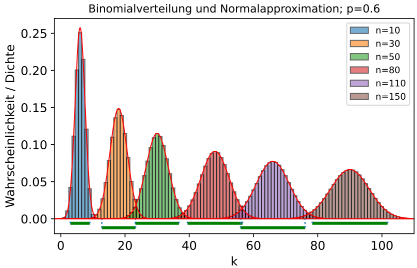
Der Zollstock als rücktransformierte Normalapproximation

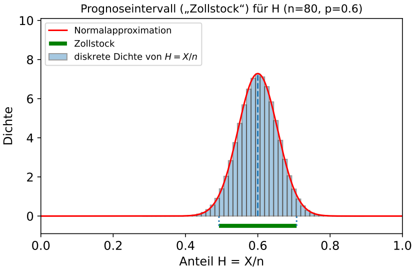


Der Zollstock als rücktransformierte Normalapproximation

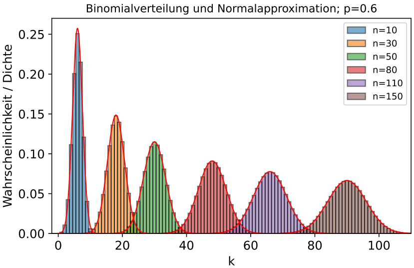


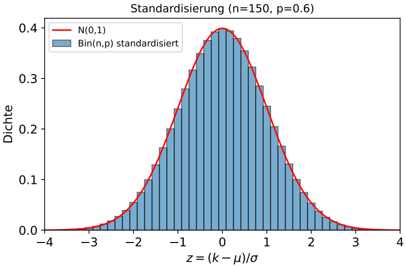


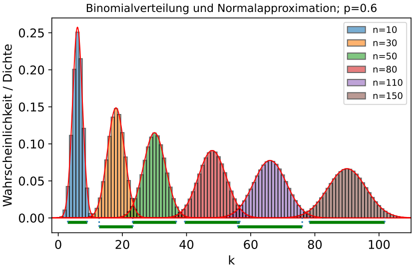


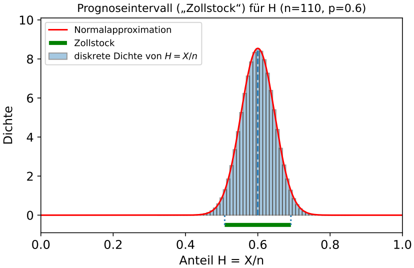


Der Zollstock als rücktransformierte Normalapproximation

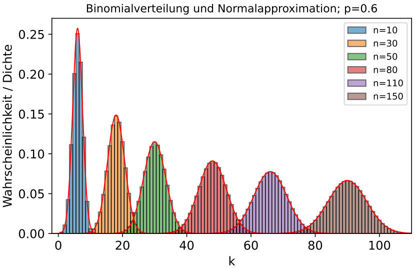


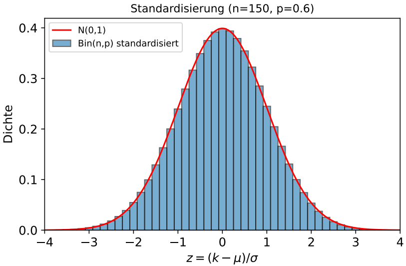


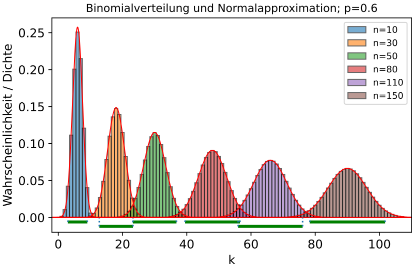


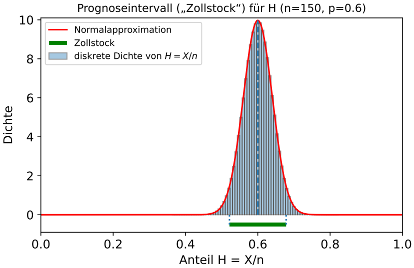


Der Zollstock als rücktransformierte Normalapproximation









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

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

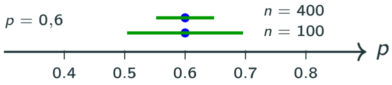
Prognoseintervall - der Zollstock

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

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



Planung des Stichprobenumfangs – mit Zollstöcken

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

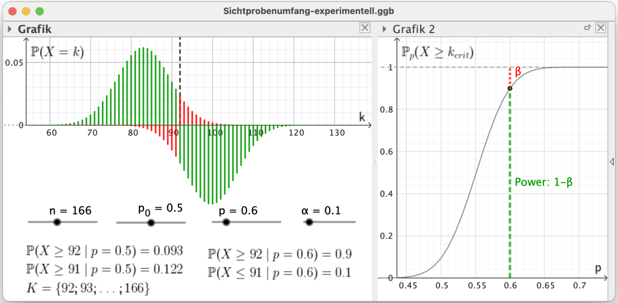
Ausgangsfrage: Wie groß muss der Stichprobenumfang n mindestens sein, damit ein Test auf Niveau $\alpha = 0{,}1$ für $H_0 : p \leq 0{,}5$ gegen $H_1 : p > 0{,}5$ mit Wahrscheinlichkeit $\gamma = 0{,}9$ die richtige Entscheidung (H_1 trifft zu) gibt, falls $p = 0{,}6$ gilt?

Ein Bild kann helfen...

Planung des Stichprobenumfangs – mit Zollstöcken

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?

Ein Bild kann helfen...



Planung des Stichprobenumfangs – mit Zollstöcken

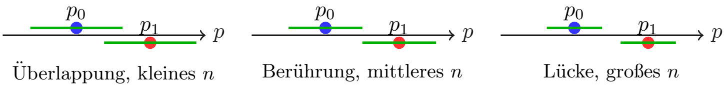
Test $H_0 : p \leq 0{,}5$ gegen $H_1 : p > 0{,}5$ mit $\alpha = 0{,}1$, $\gamma = 0{,}9$, $p_1 = 0{,}6$.

Zwei Zollstöcke aber auch...

Planung des Stichprobenumfangs – mit Zollstöcken

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

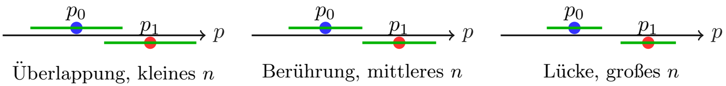
Zwei Zollstöcke aber auch...



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: $g(n) = p_1 - \Phi^{-1}(\gamma)\sqrt{p_1(1-p_1)/n}$

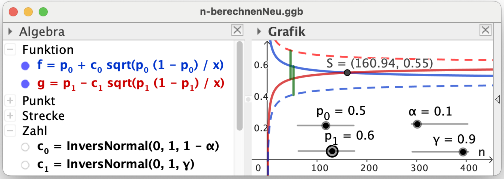
Planung des Stichprobenumfangs – mit Zollstöcken

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...

Überlappung, kleines n — Berührung, mittleres n — Lücke, großes n

- 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: $g(n) = p_1 - \Phi^{-1}(\gamma)\sqrt{p_1(1-p_1)/n}$

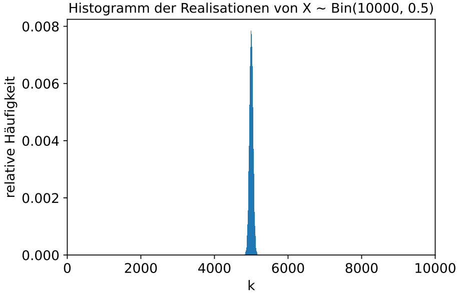


Prognoseintervalle - Denken in Verteilungen

Für eine binomialverteilte Zufallsgröße X gilt: $n = 10000; p = 0,5$.
Finden Sie Aufgaben der Form $P(X = k); P(X \leq k); P(X \geq k); P(a \leq X \leq b), \ldots$.

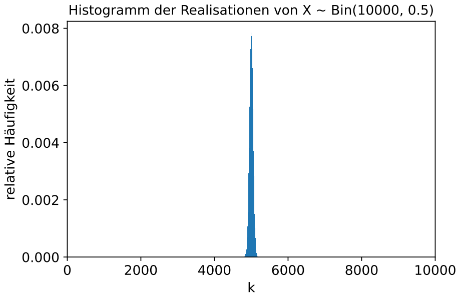
Prognoseintervalle - Denken in Verteilungen

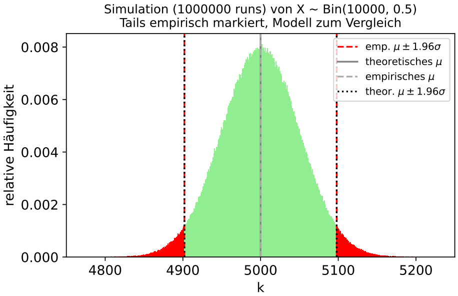
Für eine binomialverteilte Zufallsgröße X gilt: $n = 10000; p = 0,5$.
Finden Sie Aufgaben der Form $P(X = k); P(X \leq k); P(X \geq k); P(a \leq X \leq b), \dots$.



Prognoseintervalle - Denken in Verteilungen

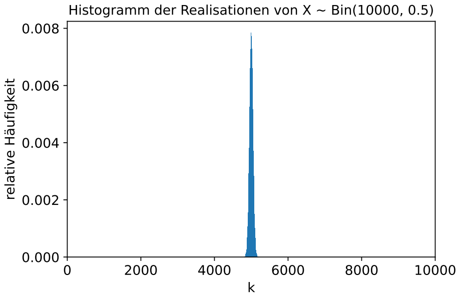
Für eine binomialverteilte Zufallsgröße X gilt: $n = 10000; p = 0,5$.
Finden Sie Aufgaben der Form $P(X = k); P(X \leq k); P(X \geq k); P(a \leq X \leq b), \ldots$.

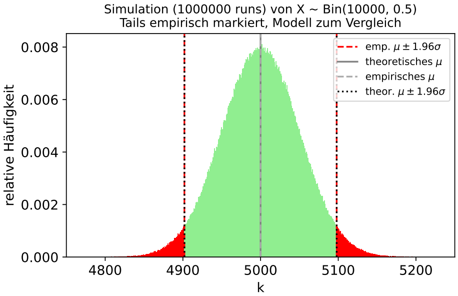




Prognoseintervalle - Denken in Verteilungen

Für eine binomialverteilte Zufallsgröße X gilt: $n = 10000; p = 0,5$.
Finden Sie Aufgaben der Form $P(X = k); P(X \leq k); P(X \geq k); P(a \leq X \leq b), \ldots$.



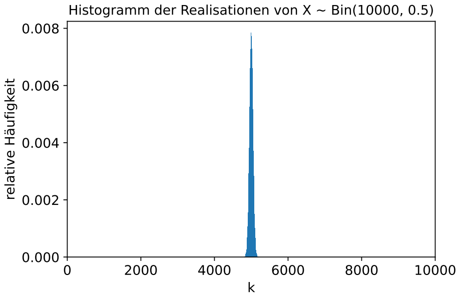


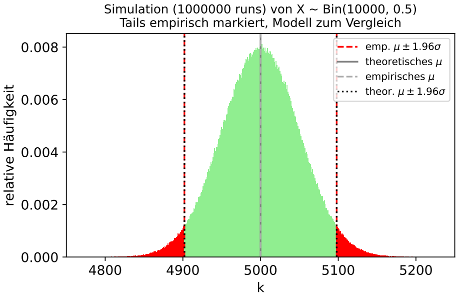
Idee (ohne Rechner): $\mu = n \cdot p = 5000; \quad \sigma = \sqrt{10000 \cdot 0{,}5 \cdot 0{,}5} = 50; \; 2\sigma = 100$

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

Prognoseintervalle - Denken in Verteilungen

Für eine binomialverteilte Zufallsgröße X gilt: $n = 10000; p = 0,5$.
Finden Sie Aufgaben der Form $P(X = k); P(X \leq k); P(X \geq k); P(a \leq X \leq b), \dots$.

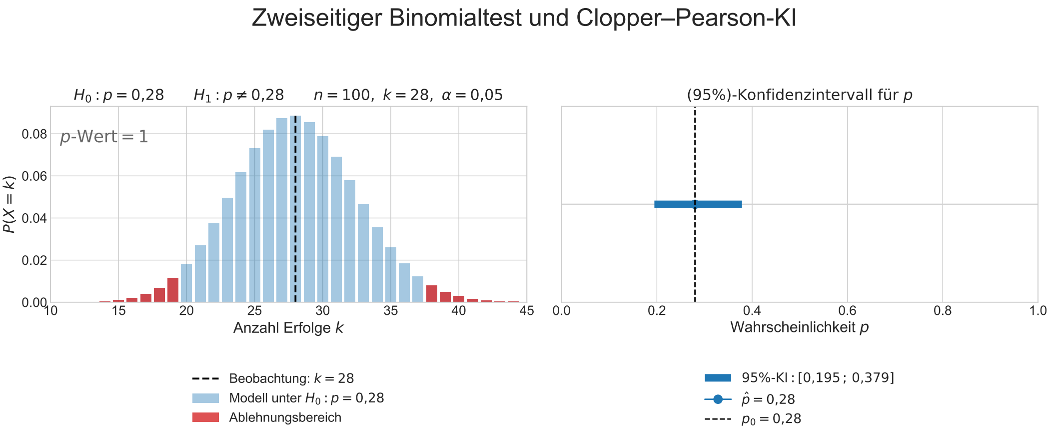




Idee (ohne Rechner): $\mu = n \cdot p = 5000; \quad \sigma = \sqrt{10000 \cdot 0{,}5 \cdot 0{,}5} = 50; \; 2\sigma = 100$

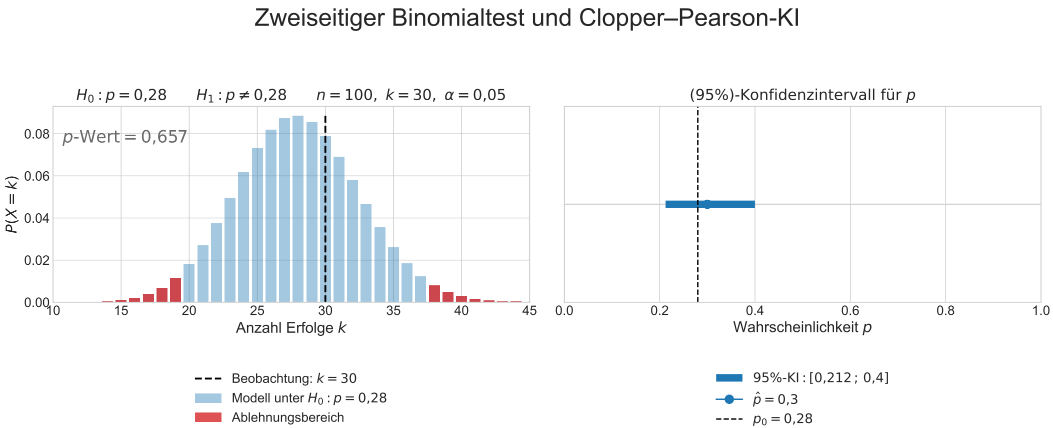
Folgerung: In ca. 95% aller *zukünftigen* Fälle gilt: $4900 \leq X \leq 5100$. $(0{,}49 \leq \frac{X}{n} \leq 0{,}51)$

Zweiseitiger Test und Konfidenzintervall – 11 Situationen

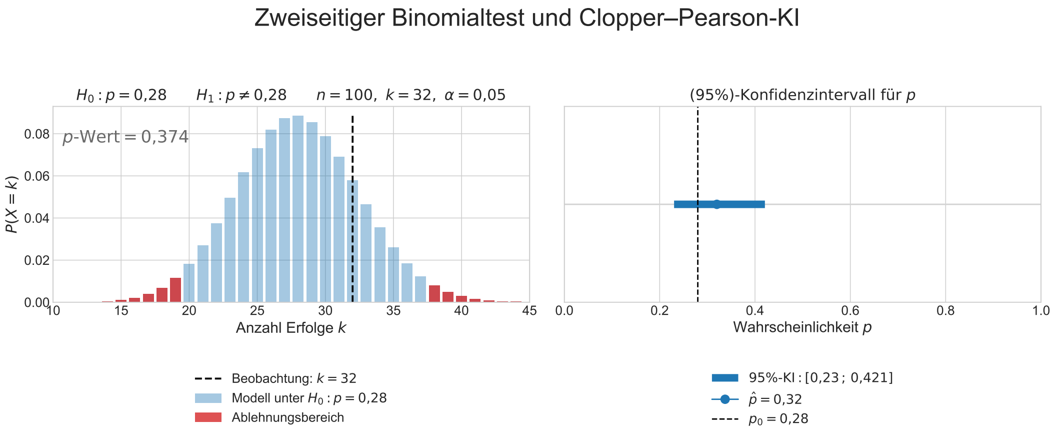


Beobachtung: $k = 28$

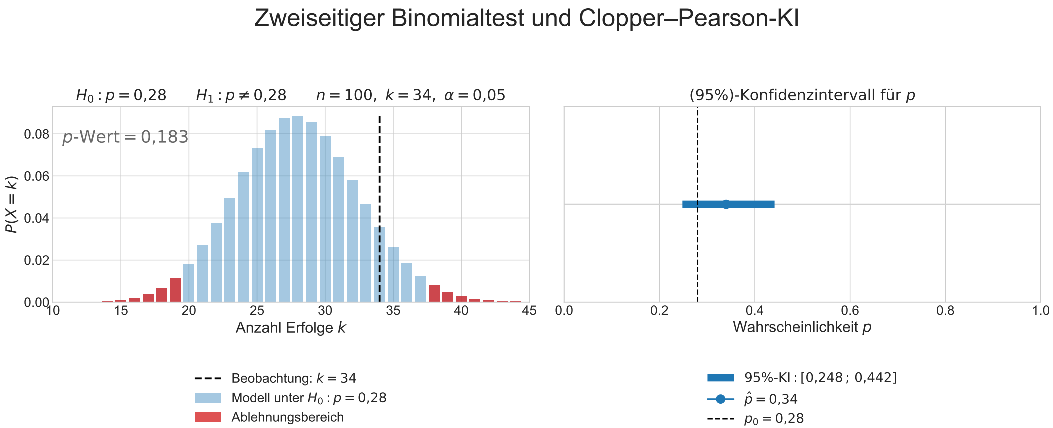
Zweiseitiger Test und Konfidenzintervall – 11 Situationen



Beobachtung: $k = 30$

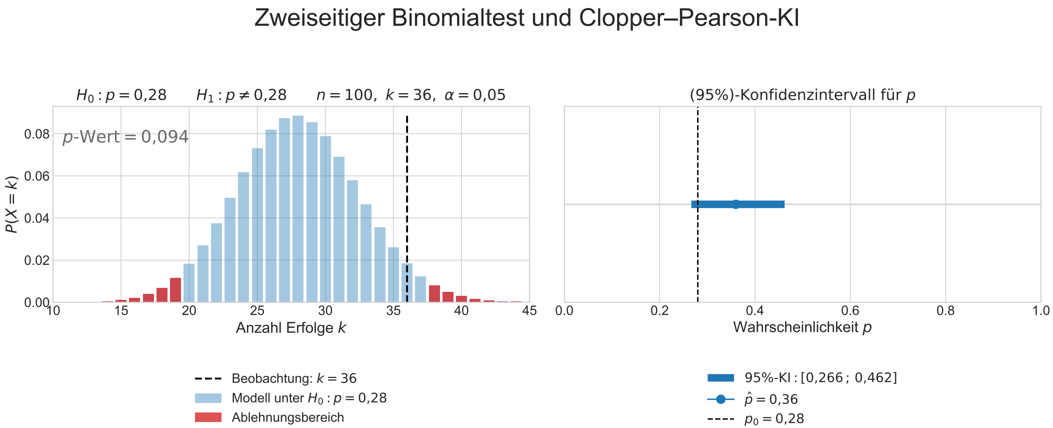


Beobachtung: $k = 32$

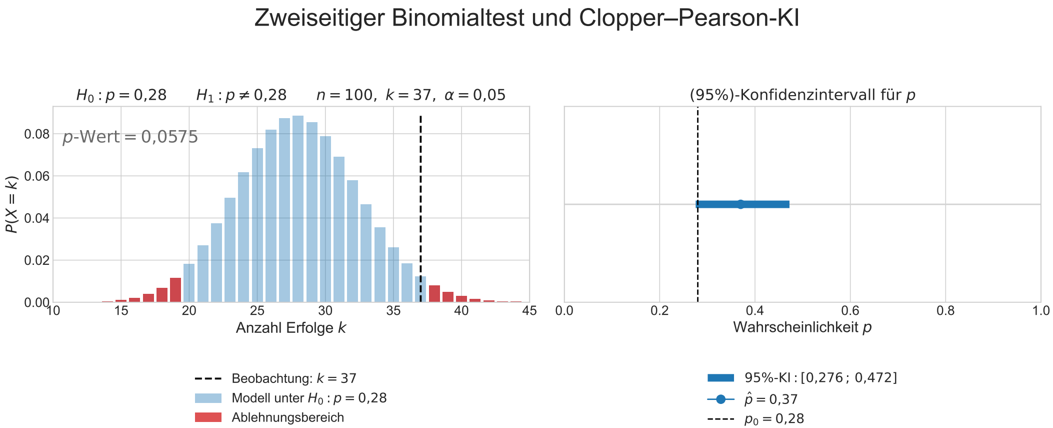


Beobachtung: $k = 34$

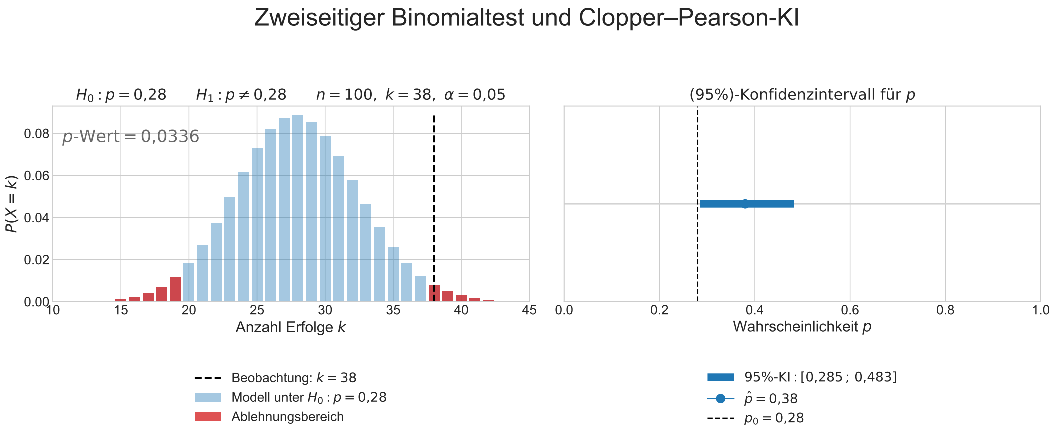
Zweiseitiger Test und Konfidenzintervall – 11 Situationen



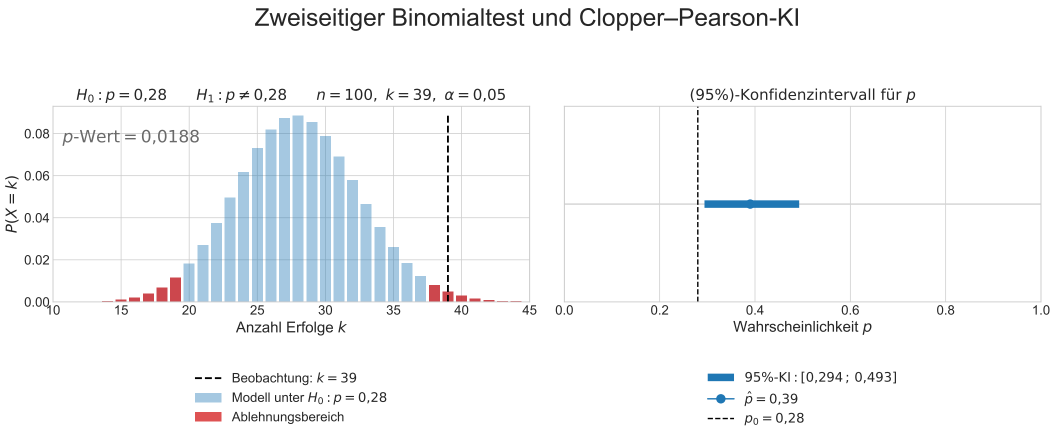
Beobachtung: $k = 36$



Beobachtung: $k = 37$

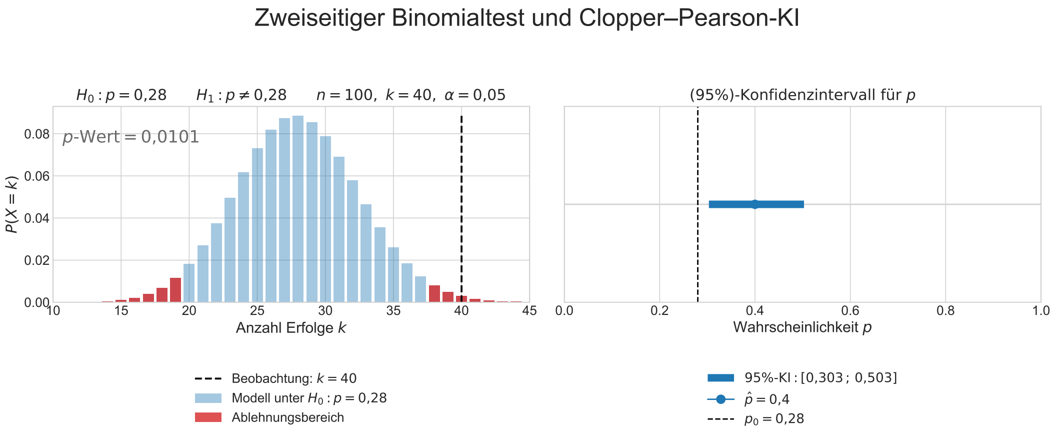


Beobachtung: $k = 38$



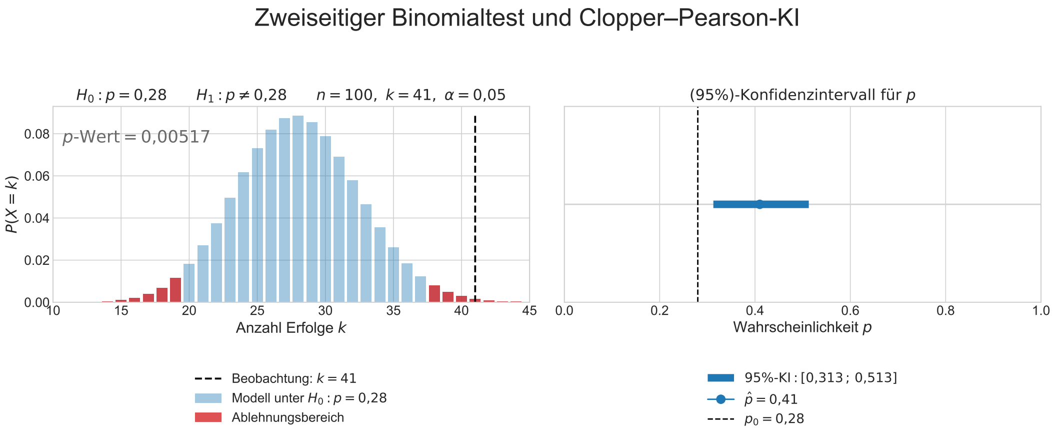
Beobachtung: $k = 39$

Zweiseitiger Test und Konfidenzintervall – 11 Situationen



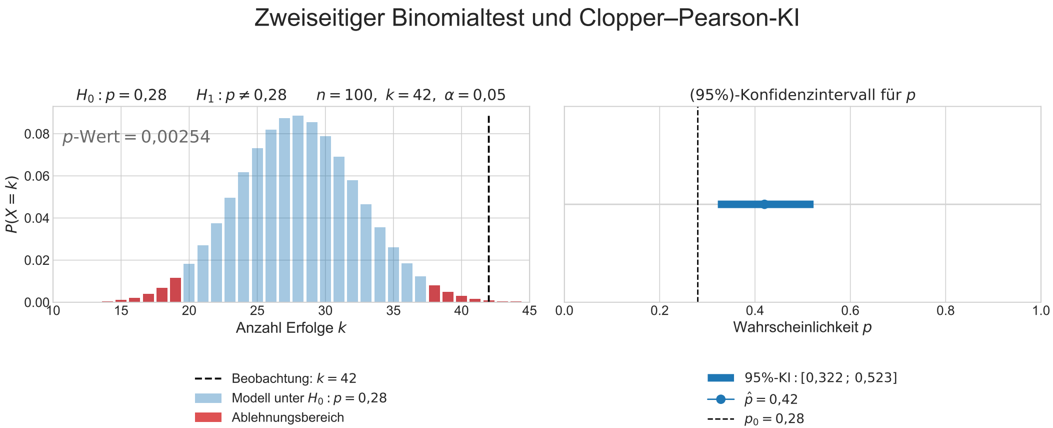
Beobachtung: $k = 40$

Zweiseitiger Test und Konfidenzintervall – 11 Situationen



Beobachtung: $k = 41$

Zweiseitiger Test und Konfidenzintervall – 11 Situationen



Beobachtung: $k = 42$

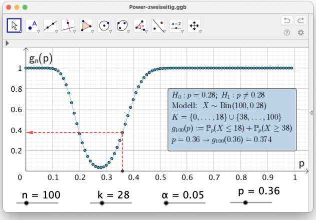
Ein Modell – zwei digitale Perspektiven

GeoGebra & Python – zwei Werkzeuge, ein Modell

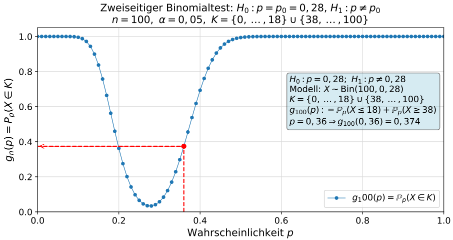
GeoGebra Schulnah, interaktiv und zeilenorientiert, sofort sichtbare Objekte

Python Präzise, reproduzierbar, ideal für Simulation und professionelles Arbeiten

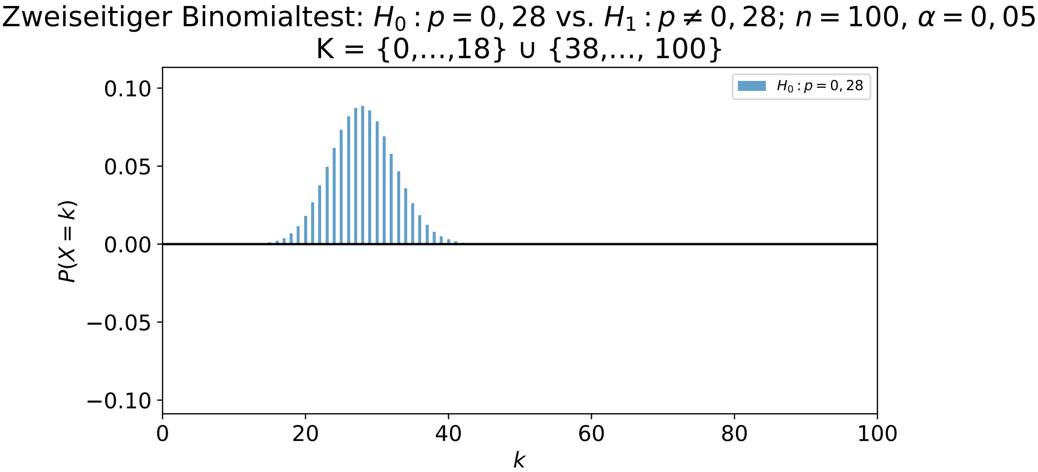
GeoGebra: Diskrete Powerfunktion



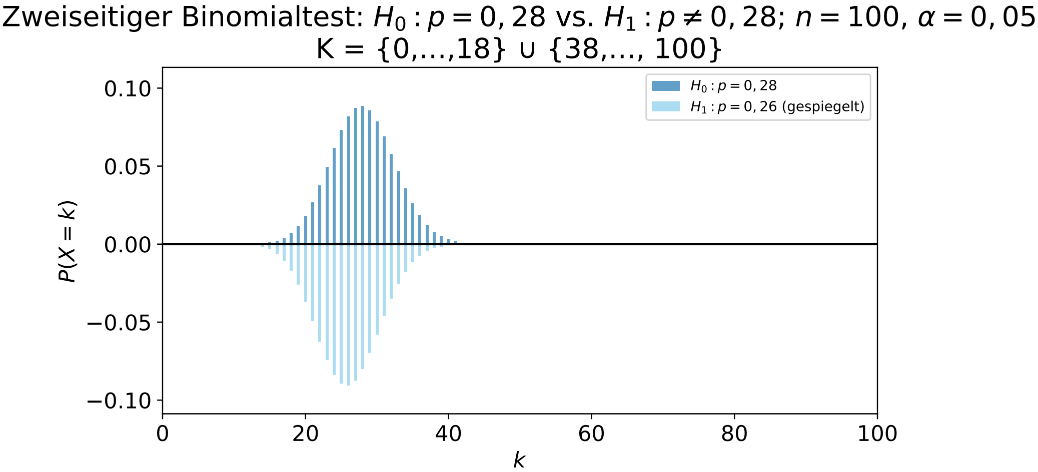
Python: Diskrete Powerfunktion

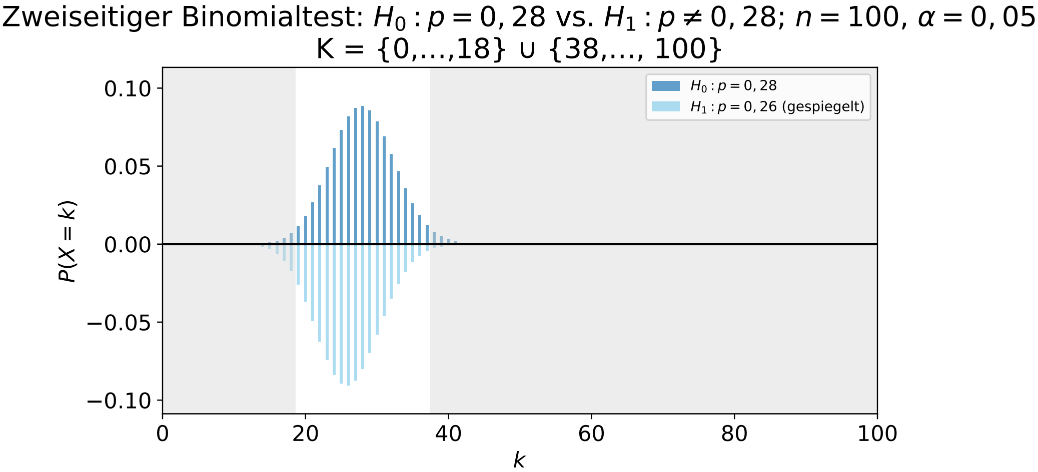


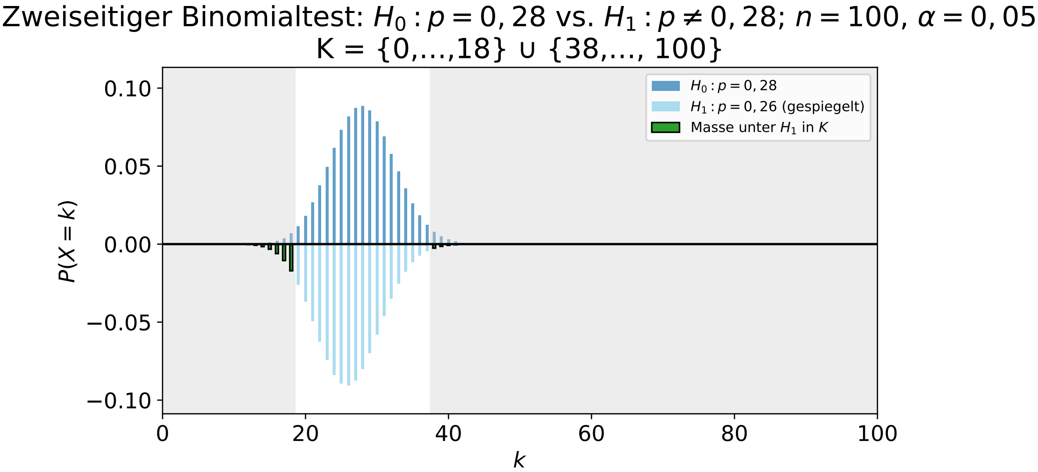
Didaktische Leitidee: GeoGebra zum Verstehen – Python zum Produzieren.


Zweiseitiger Binomialtest: $H_0: p = 0,28$ vs. $H_1: p \neq 0,28$; $n = 100$, $\alpha = 0,05$
K = {0,...,18} ∪ {38,..., 100}
$H_0: p = 0,28$
0.10
0.05
0.00
−0.05
−0.10
$P(X = k)$
0
20
40
60
80
100
k

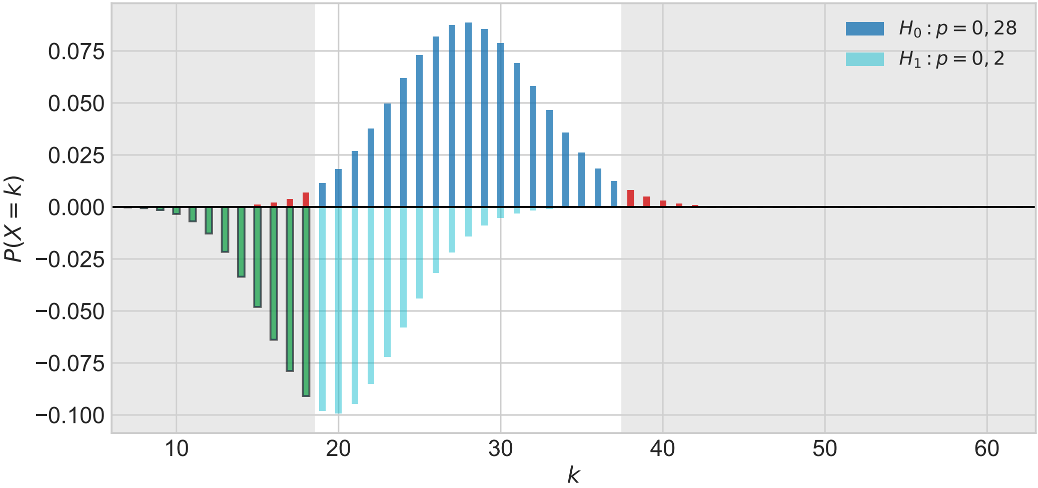
Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



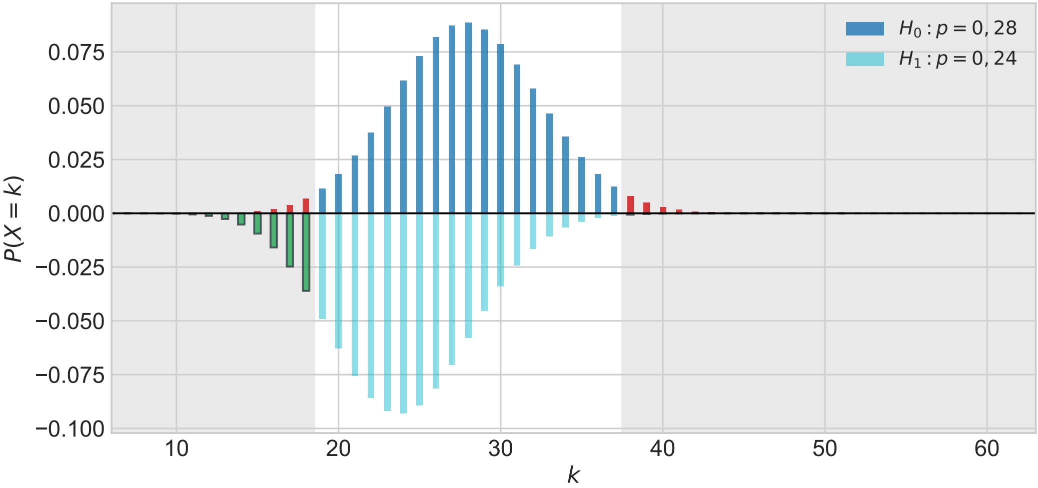

Zweiseitiger Binomialtest: $H_0: p = 0,28$ vs. $H_1: p \neq 0,28$; $n = 100$, $\alpha = 0,05$
K = {0,...,18} ∪ {38,..., 100}
$H_0: p = 0,28$
$H_1: p = 0,26$ (gespiegelt)
0.10
0.05
0.00
−0.05
−0.10
$P(X = k)$
0
20
40
60
80
100
k


Zweiseitiger Binomialtest: $H_0: p = 0,28$ vs. $H_1: p \neq 0,28$; $n = 100$, $\alpha = 0,05$
K = {0,...,18} ∪ {38,..., 100}
$H_0: p = 0,28$
$H_1: p = 0,26$ (gespiegelt)
Masse unter H_1 in K
$P(X = k)$
k
0.10
0.05
0.00
−0.05
−0.10
0
20
40
60
80
100

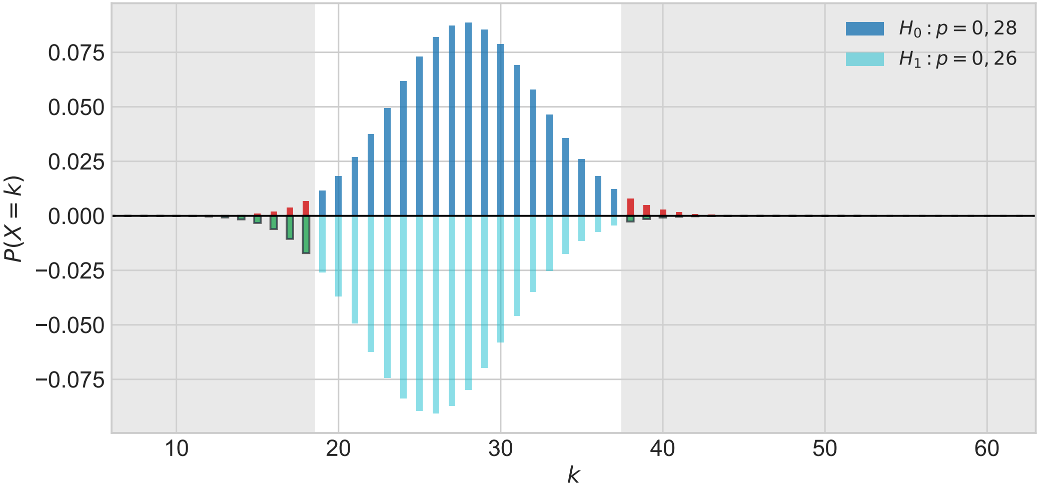
Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



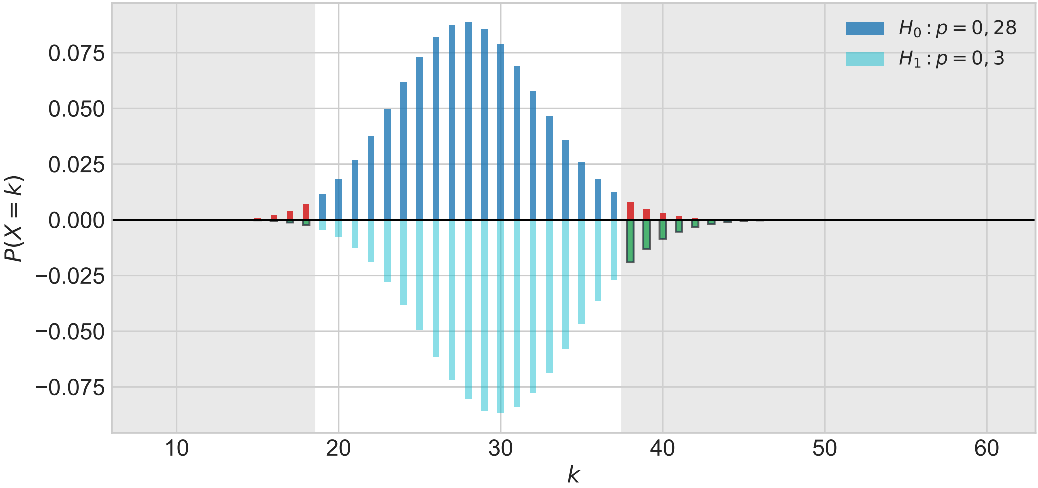
Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



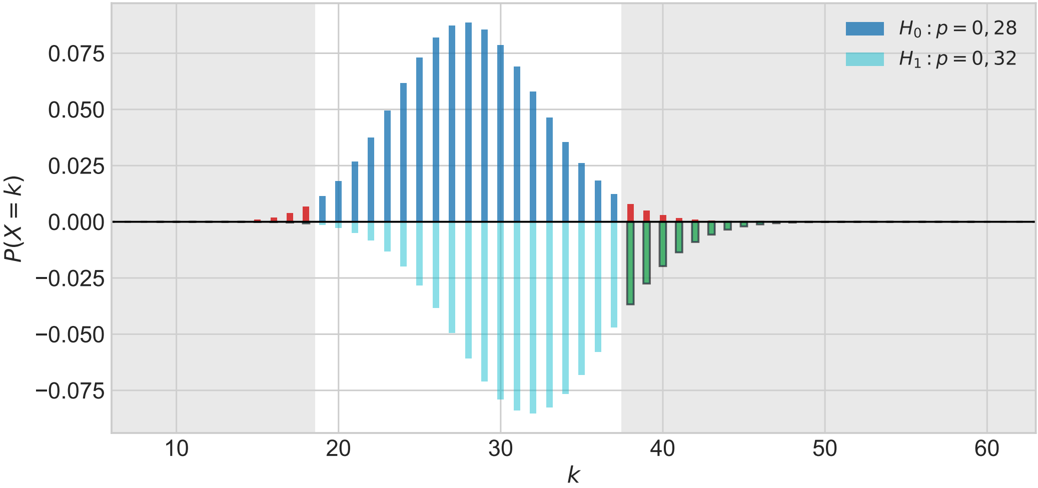
Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



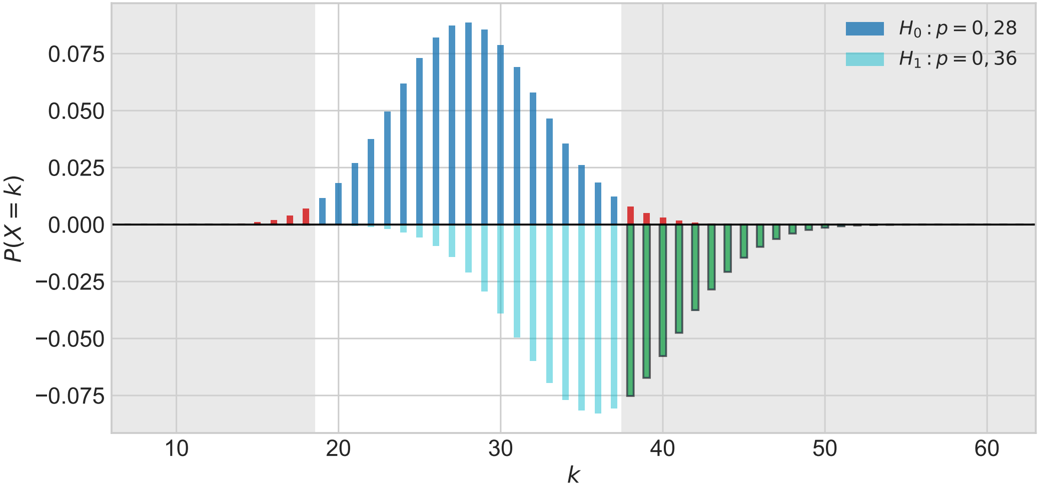
Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



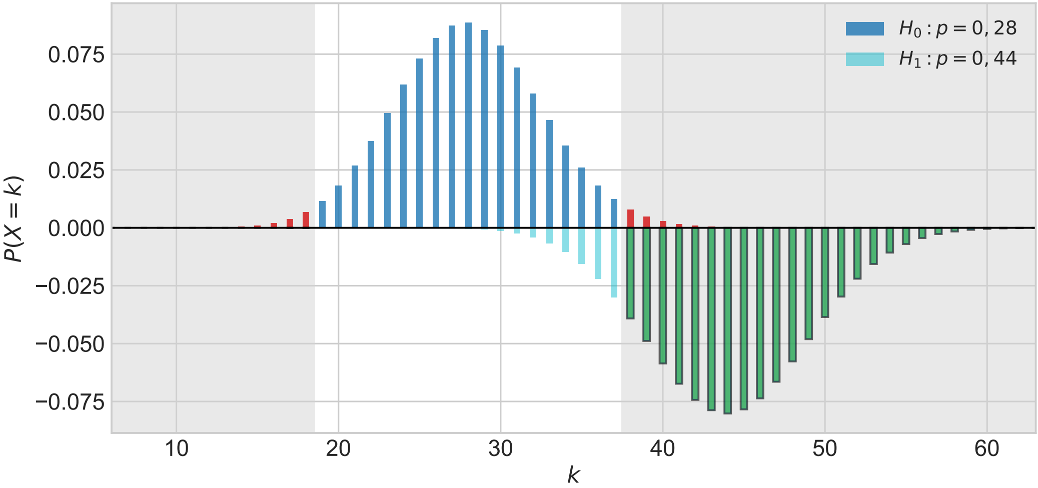
Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen

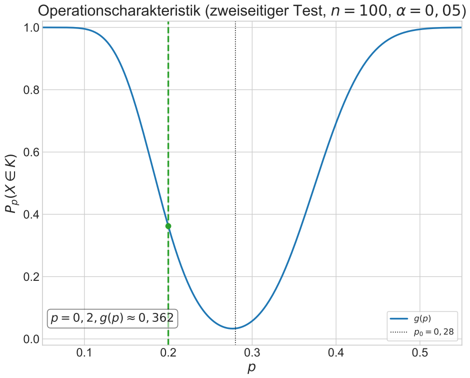


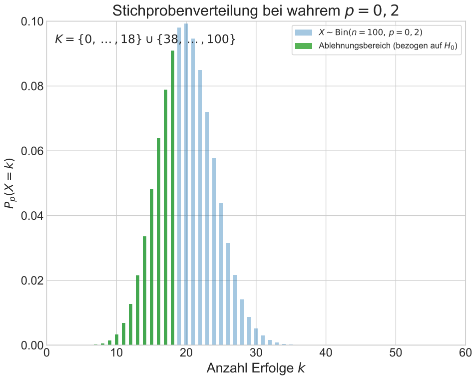
Die Power ohne Formeln – mit Spiegelung und Denken in Verteilungen



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

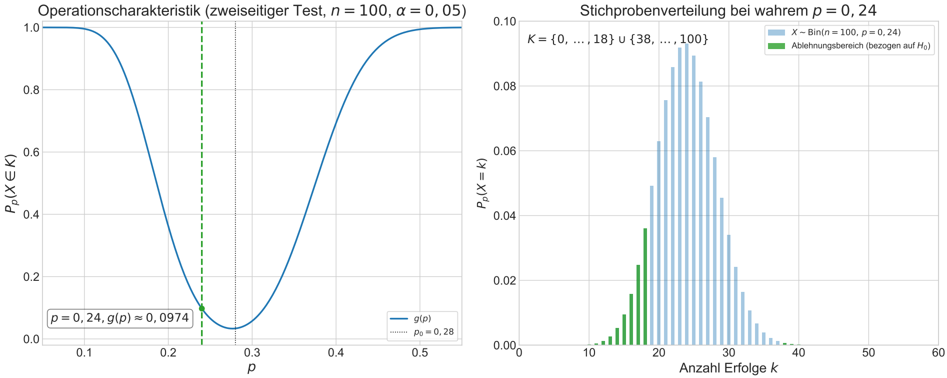
zweiseitiger Hypothesentest: $H_0: p_0 = 0{,}28; \quad H_1: p \neq 0{,}28$





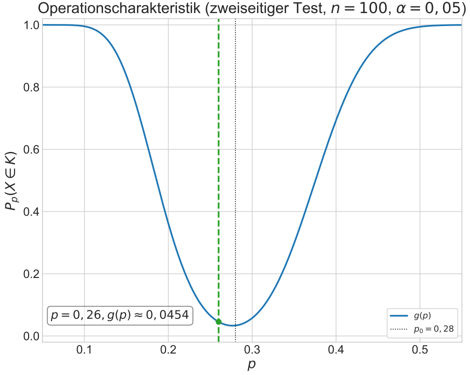
Powerfunktion $g_n(p) = \mathbb{P}_p(X \in K)$ - Wahrscheinlichkeitsmasse von p in K

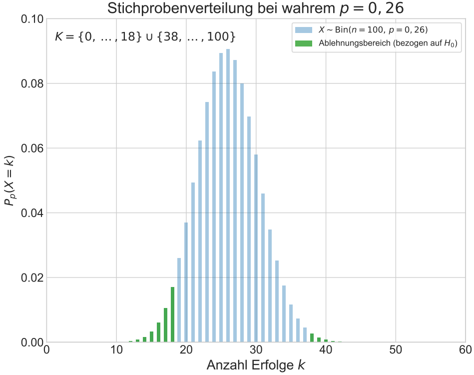
zweiseitiger Hypothesentest: $H_0 : p_0 = 0{,}28; \quad H_1 : p \neq 0{,}28$



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

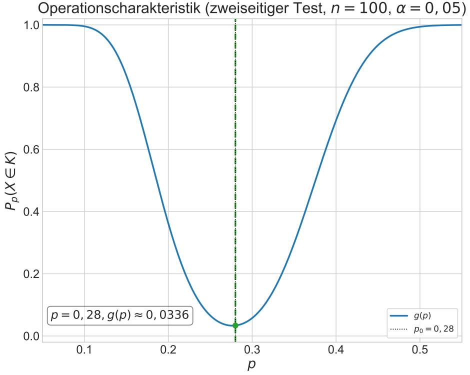
zweiseitiger Hypothesentest: $H_0 : p_0 = 0{,}28; \quad H_1 : p \neq 0{,}28$

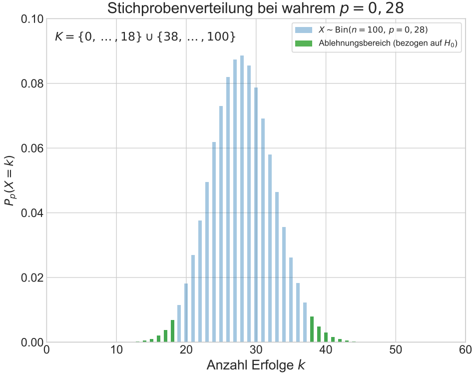




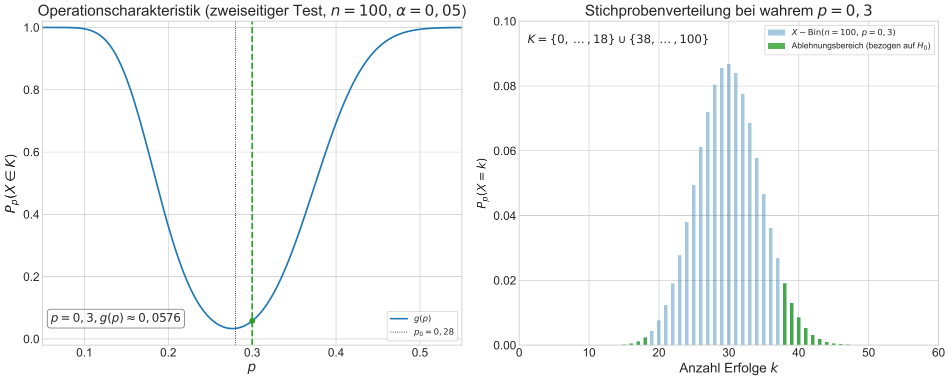
Powerfunktion $g_n(p) = \mathbb{P}_p(X \in K)$ - Wahrscheinlichkeitsmasse von p in K

zweiseitiger Hypothesentest: $H_0: p_0 = 0{,}28; \quad H_1: p \neq 0{,}28$

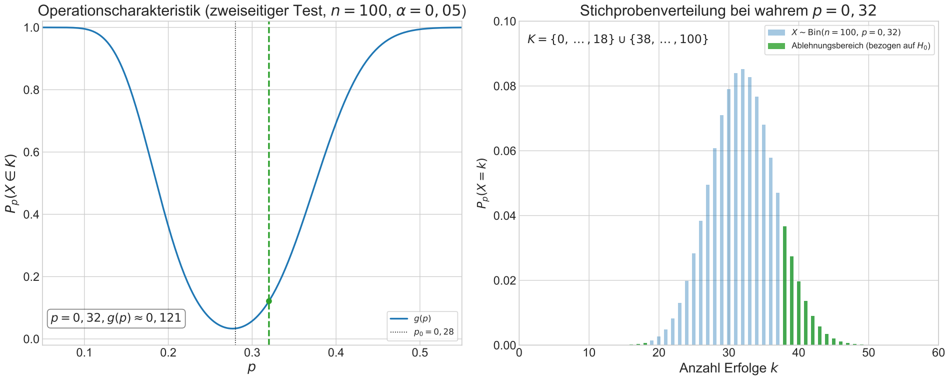




zweiseitiger Hypothesentest: $H_0 : p_0 = 0{,}28; \quad H_1 : p \neq 0{,}28$

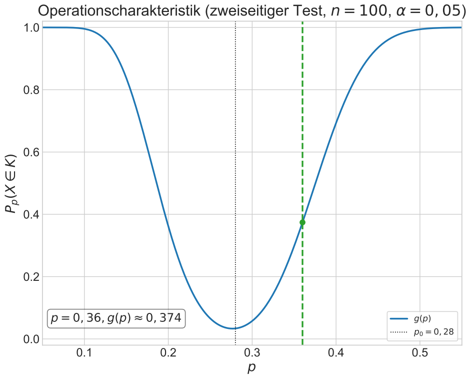


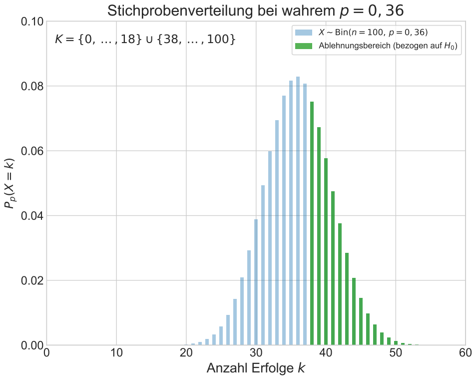
zweiseitiger Hypothesentest: $H_0 : p_0 = 0{,}28; \quad H_1 : p \neq 0{,}28$



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

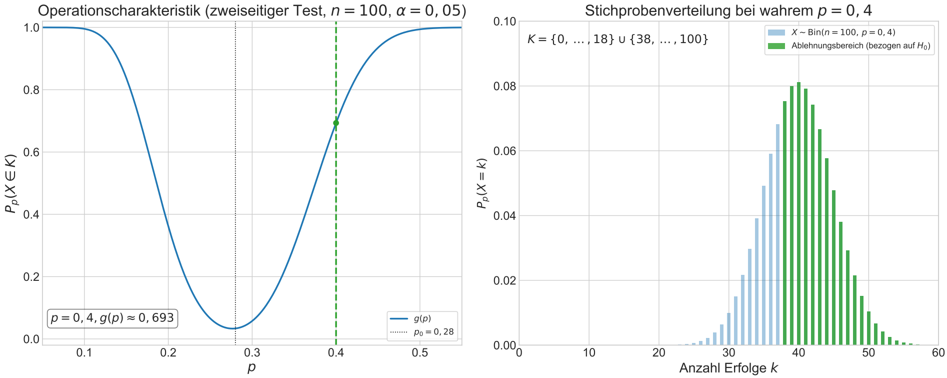
zweiseitiger Hypothesentest: $H_0 : p_0 = 0{,}28; \quad H_1 : p \neq 0{,}28$





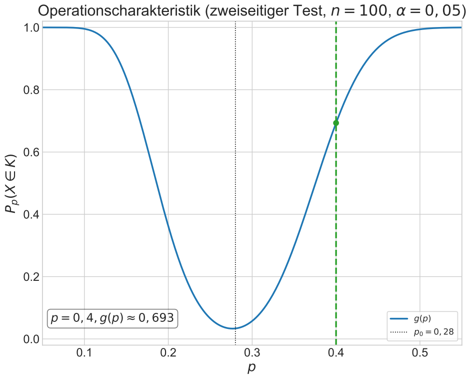
Powerfunktion $g_n(p) = \mathbb{P}_p(X \in K)$ - Wahrscheinlichkeitsmasse von p in K

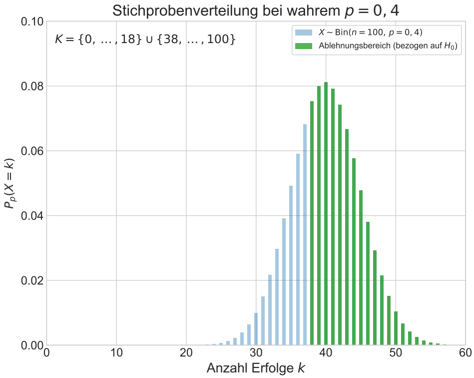
zweiseitiger Hypothesentest: $H_0: p_0 = 0{,}28; \quad H_1: p \neq 0{,}28$



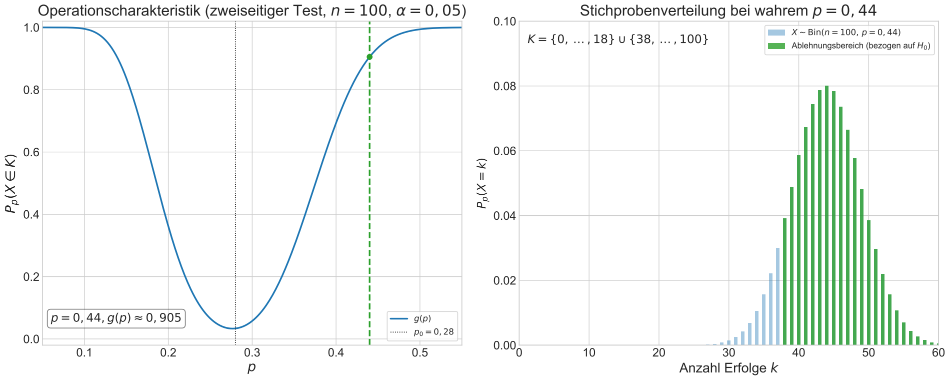
Powerfunktion $g_n(p) = \mathbb{P}_p(X \in K)$ - Wahrscheinlichkeitsmasse von p in K

zweiseitiger Hypothesentest: $H_0 : p_0 = 0{,}28; \quad H_1 : p \neq 0{,}28$



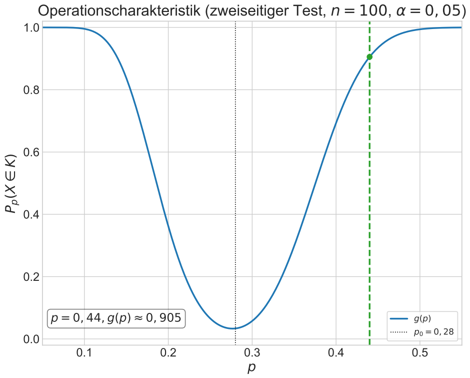


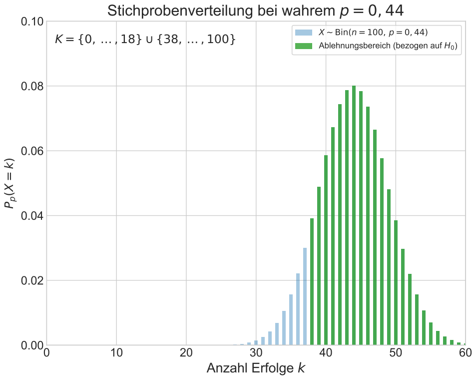
zweiseitiger Hypothesentest: $H_0: p_0 = 0{,}28; \quad H_1: p \neq 0{,}28$



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

zweiseitiger Hypothesentest: $H_0: p_0 = 0{,}28; \quad H_1: p \neq 0{,}28$

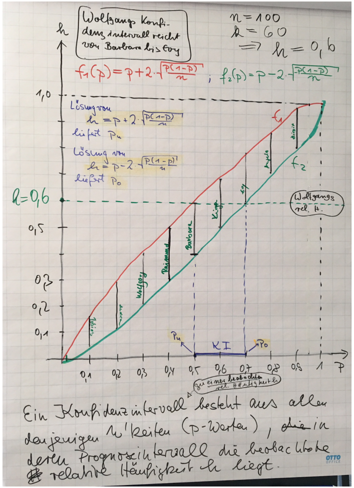




95%-PI von $p_0 = 0{,}28$: $[0{,}192\,;\,0{,}368]$ 95%-PI von $\mu = 28$: $[20\,;\,36]$

Konfidenzellipse - selbst erstellt

Ein Produkt während einer denkwürdigen Fortbildung im Mellingen.



Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren

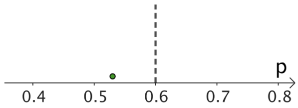
1. WALD-Intervall als Realisation $\rightarrow$ Realität

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

Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren

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

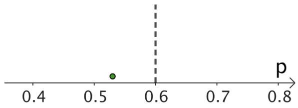
$$\left[\, h \mp 1{,}96\sqrt{\frac{h(1-h)}{n}}\,\right]$$



Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren

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

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



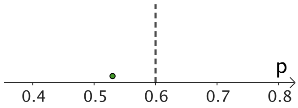
2. Viele mögliche *h*-Werte $\rightarrow$ Modell

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

Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren

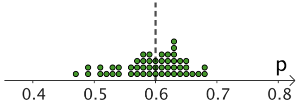
1. **WALD-Intervall als Realisation** → Realität

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



2. **Viele mögliche *h*-Werte** → Modell

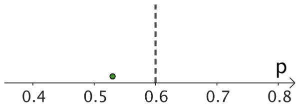
$$H = \frac{X}{n}$$



Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren

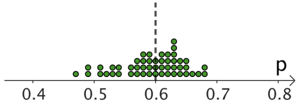
1. WALD-Intervall als Realisation $\to$ Realität

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



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

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



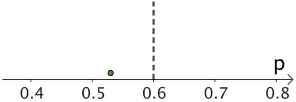
3. Grenzen als Zufallsgrößen $L(H)$ und $R(H)$ $\to$ Modelle

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

Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren

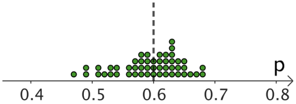
1. WALD-Intervall als Realisation $\rightarrow$ Realität

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



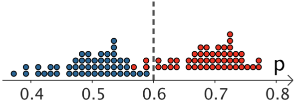
2. Viele mögliche *h*-Werte $\rightarrow$ Modell

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



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

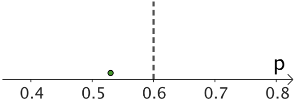
$$\big[L(H), R(H)\big] = \left[\, H \mp 1{,}96\sqrt{\tfrac{H(1-H)}{n}}\,\right]$$



Vom Zahlenintervall zum Zufallsintervall – der Weg zum Konfidenzverfahren

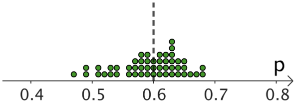
1. WALD-Intervall als Realisation $\rightarrow$ Realität

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



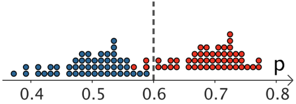
2. Viele mögliche h-Werte $\rightarrow$ Modell

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



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

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

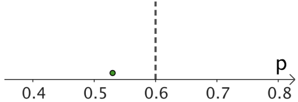


4. KI als Verfahren (viele Intervalle) $\rightarrow$ Modelle

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

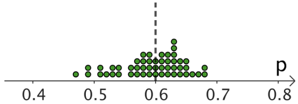
1. **WALD-Intervall als Realisation** $\rightarrow$ Realität

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



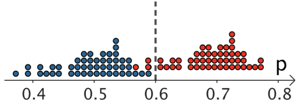
2. **Viele mögliche h-Werte** $\rightarrow$ Modell

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



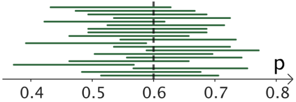
3. **Grenzen als Zufallsgrößen $L(H)$ und $R(H)$** $\rightarrow$ Modelle

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

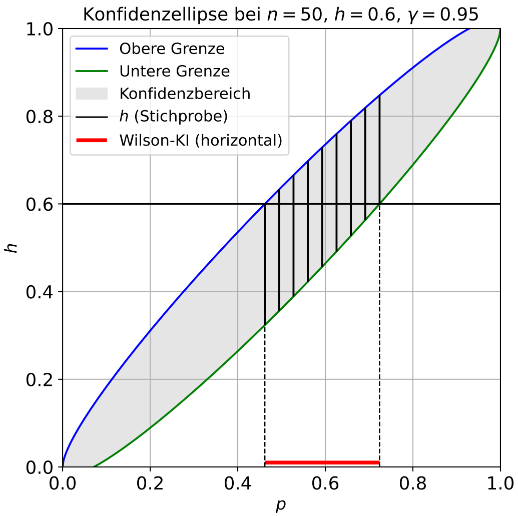


4. **KI als Verfahren (viele Intervalle)** $\rightarrow$ Modelle

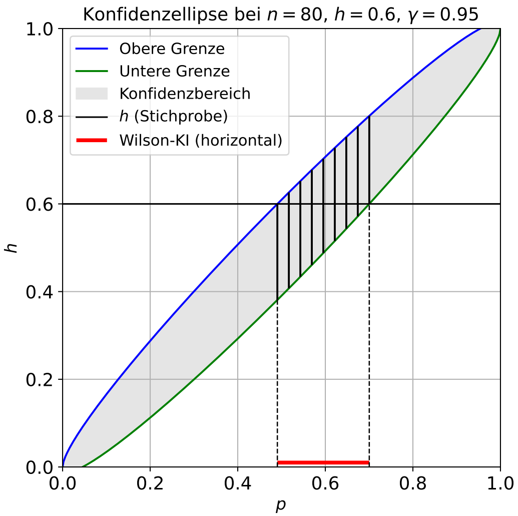
$$\mathbb{P}_p\big([L(H), R(H)] \ni p\big) \;=\; 0{,}95$$



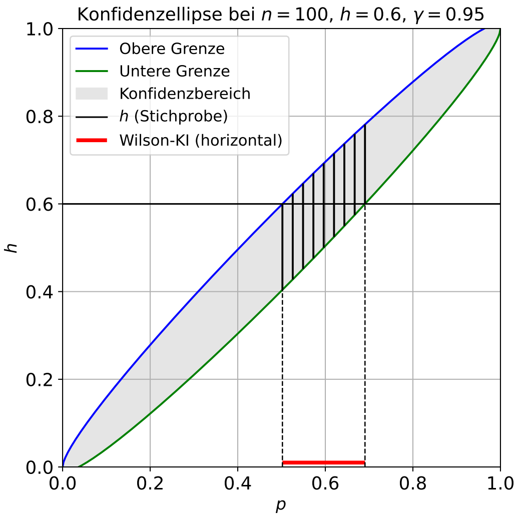
Konfidenzellipsen – ein Überblick, falls n verändert wird



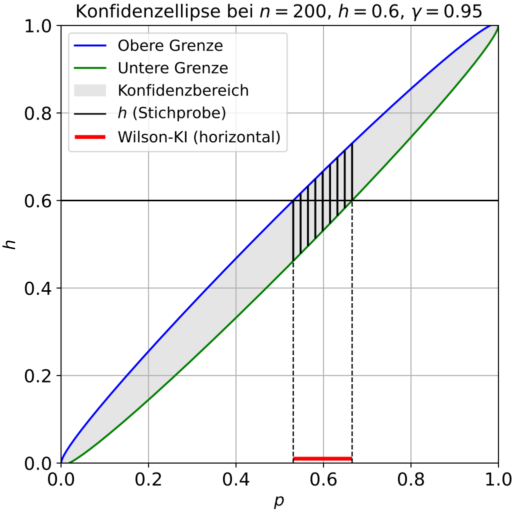
Konfidenzellipsen – ein Überblick, falls n verändert wird



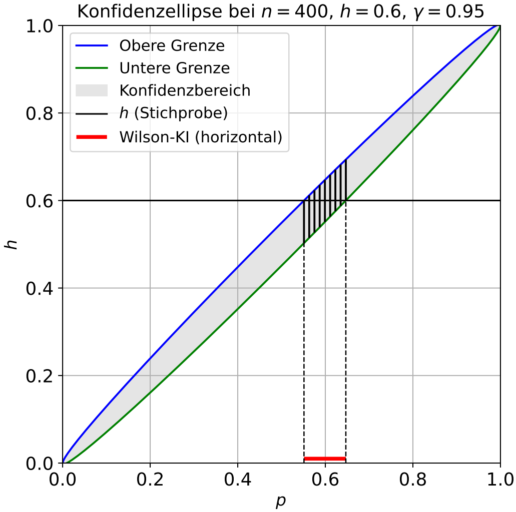
Konfidenzellipsen – ein Überblick, falls n verändert wird



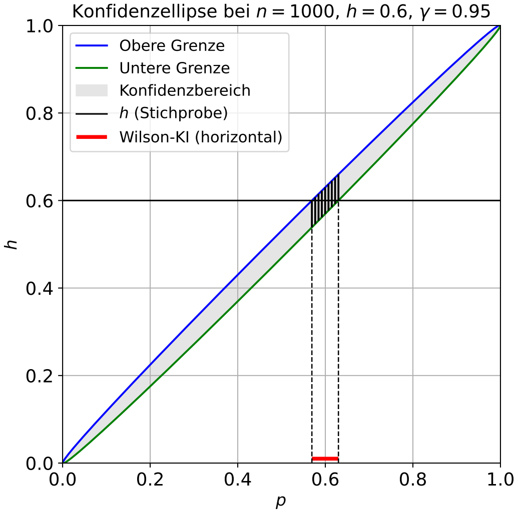
Konfidenzellipsen – ein Überblick, falls n verändert wird



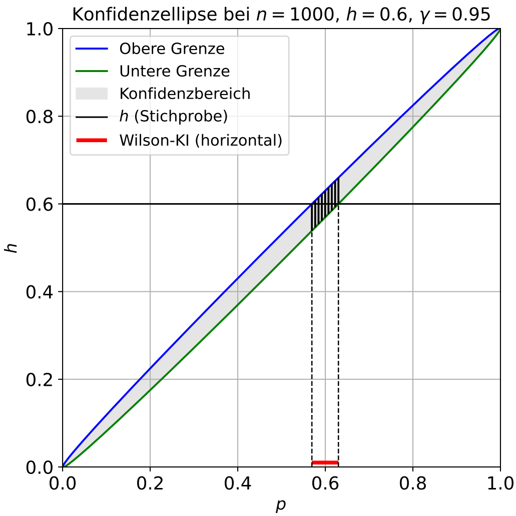
Konfidenzellipsen – ein Überblick, falls n verändert wird

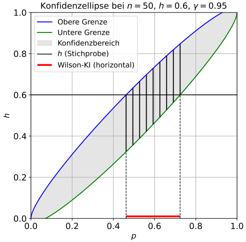


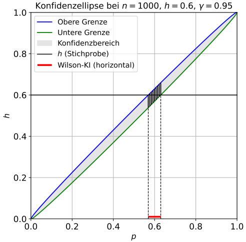
Konfidenzellipsen – ein Überblick, falls n verändert wird



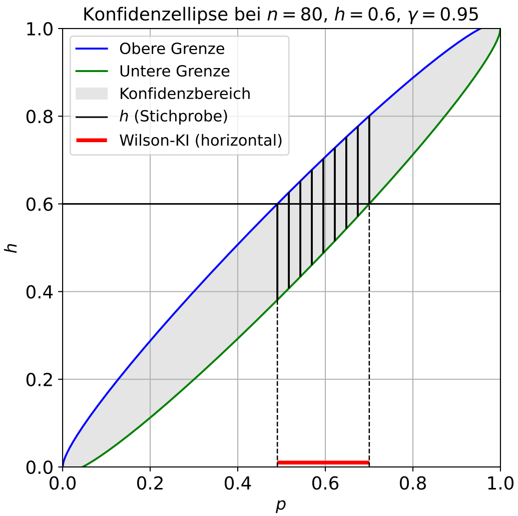
Konfidenzellipsen – ein Überblick, falls n verändert wird







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:

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

Vergleich von Konfidenzintervallen

k	$\hat{p}$	Wald	Wilson	Clopper–Pearson
28	0,28	[0,192 ; 0,368]	[0,2014 ; 0,3749]	[0,1948 ; 0,3787]
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]
40	0,4	[0,304 ; 0,496]	[0,3094 ; 0,498]	[0,3033 ; 0,5028]
41	0,41	[0,3136 ; 0,5064]	[0,3187 ; 0,508]	[0,3126 ; 0,5129]
42	0,42	[0,3233 ; 0,5167]	[0,328 ; 0,5179]	[0,322 ; 0,5229]

Vergleich von Konfidenzintervallen

k	$\hat{p}$	Wald	Wilson	Clopper–Pearson
280	0,28	[0,2522 ; 0,3078]	[0,2531 ; 0,3086]	[0,2524 ; 0,3089]
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]
400	0,4	[0,3696 ; 0,4304]	[0,3701 ; 0,4307]	[0,3695 ; 0,4311]
410	0,41	[0,3795 ; 0,4405]	[0,3799 ; 0,4408]	[0,3793 ; 0,4412]
420	0,42	[0,3894 ; 0,4506]	[0,3898 ; 0,4508]	[0,3892 ; 0,4513]

Vom WALD-Intervall zur Konfidenzmethode

1. WALD-KI als *Realisation:* $[h \pm 1{,}96\sqrt{h(1-h)/n}]$

2. Wo steckt der Zufall? $h = \frac{x}{n} \quad \Rightarrow \quad H = \frac{X}{n}$

3. Grenzen werden zu Zufallsgrößen

$$L(H) = H - 1{,}96\sqrt{H(1-H)/n}, \qquad R(H) = H + 1{,}96\sqrt{H(1-H)/n}$$

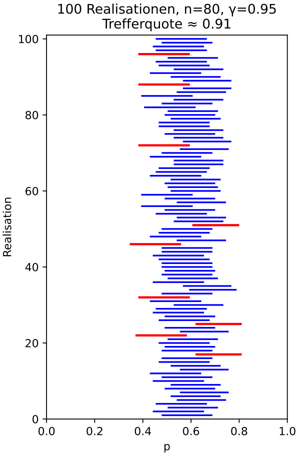
4. Das eigentliche Konfidenzintervall ist ein *Verfahren*

$$P\big([L(H), R(H)] \ni p\big) = 0.95$$

(Notation: $A \ni x$ bedeutet „A enthält x“.)

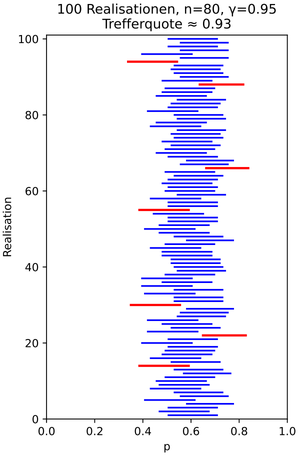
Simulationen - der Weg zum Verständnis

Anteil der KI, die den wahren Parameter enthalten

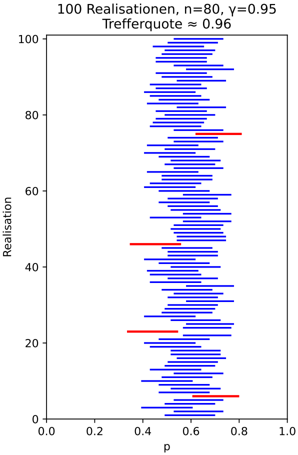


Simulationen - der Weg zum Verständnis

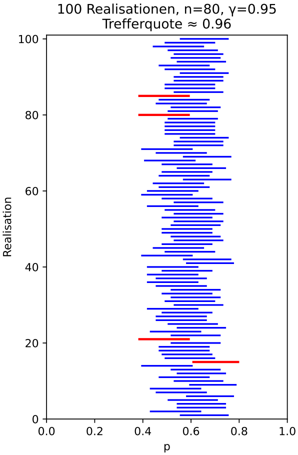
Anteil der KI, die den wahren Parameter enthalten



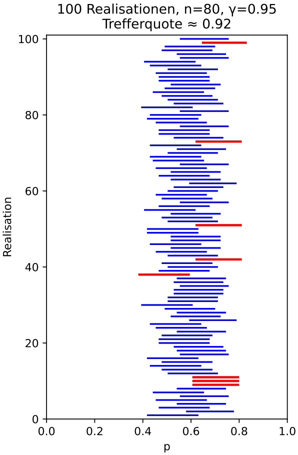
Anteil der KI, die den wahren Parameter enthalten



Anteil der KI, die den wahren Parameter enthalten



Anteil der KI, die den wahren Parameter enthalten



Die Logik des Konfidenzintervalls — in 5 Sätzen

1. Zufall steckt nur in Zufallsgrößen, nicht in Beobachtungen.

Die Logik des Konfidenzintervalls — in 5 Sätzen

1. Zufall steckt nur in Zufallsgrößen, nicht in Beobachtungen.
2. Vor der Ziehung ist $H = X/n$ eine Zufallsgröße – und alles, was davon abhängt, ist ebenfalls zufällig.

Die Logik des Konfidenzintervalls — in 5 Sätzen

1. Zufall steckt nur in Zufallsgrößen, nicht in Beobachtungen.
2. Vor der Ziehung ist $H = X/n$ eine Zufallsgröße – und alles, was davon abhängt, ist ebenfalls zufällig.
3. Ein KI entsteht aus einer Regel: $H \mapsto [L(H), R(H)]$.

Die Logik des Konfidenzintervalls — in 5 Sätzen

1. Zufall steckt nur in Zufallsgrößen, nicht in Beobachtungen.
2. Vor der Ziehung ist $H = X/n$ eine Zufallsgröße – und alles, was davon abhängt, ist ebenfalls zufällig.
3. Ein KI entsteht aus einer Regel: $H \mapsto [L(H), R(H)]$.
4. Diese Regel hat eine garantierte Trefferwahrscheinlichkeit:

$$\mathbb{P}([L(H), R(H)] \ni p) = 1 - \alpha.$$

(Notation: $A \ni x$ bedeutet „A enthält x“.)

Die Logik des Konfidenzintervalls — in 5 Sätzen

1. Zufall steckt nur in Zufallsgrößen, nicht in Beobachtungen.
2. Vor der Ziehung ist $H = X/n$ eine Zufallsgröße – und alles, was davon abhängt, ist ebenfalls zufällig.
3. Ein KI entsteht aus einer Regel: $H \mapsto [L(H), R(H)]$.
4. Diese Regel hat eine garantierte Trefferwahrscheinlichkeit:

$$\mathbb{P}([L(H), R(H)] \ni p) = 1 - \alpha.$$

   (Notation: $A \ni x$ bedeutet „A enthält x“.)

5. Nach der Ziehung wird aus dem zufälligen Intervall ein konkretes Intervall $[L(h), R(h)]$ – ohne Zufall.

Was entspricht was? - Hypothesentest vs. Konfidenzintervall

Hypothesentest

- **Gegenstand:** Punkt-Hypothese p_0
- **Zollstock:** fest über p_0
- **Zufallsgröße:** $H = X/n$
- **Realisation:** $h = x/n$
- **Regel:** Verwerfe H_0 falls $h \notin PI(p_0)$
- **Zeitlicher Ablauf:**
  Regel $\rightarrow$ Ziehung $\rightarrow$ Entscheidung

Konfidenzintervall

- **Gegenstand:** Menge möglicher p-Werte
- **Zollstock:** wandert über alle p
- **Zufallsgröße:** gesamte Intervallfamilie $PI(p)$
- **Realisation:** Menge aller p mit $h \in PI(p)$
- **Regel:** $KI = \{p : h \in PI(p)\}$
- **Zeitlicher Ablauf:**
  Ziehung $\rightarrow$ KI-Konstruktion

Zentrale Beziehung

$$p_0 \notin KI \quad \Longleftrightarrow \quad \text{Der Test auf Niveau } \alpha \text{ verwirft } H_0.$$