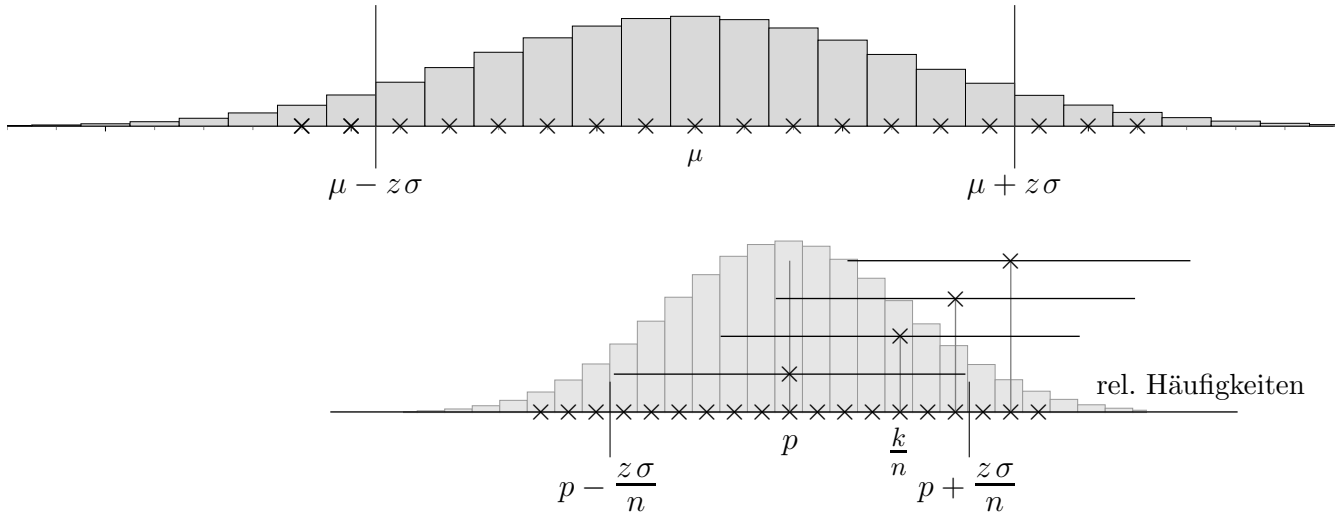


Überdeckungswahrscheinlichkeit

Überdeckt das Konfidenzintervall mit 95%iger Wahrscheinlichkeit p ?

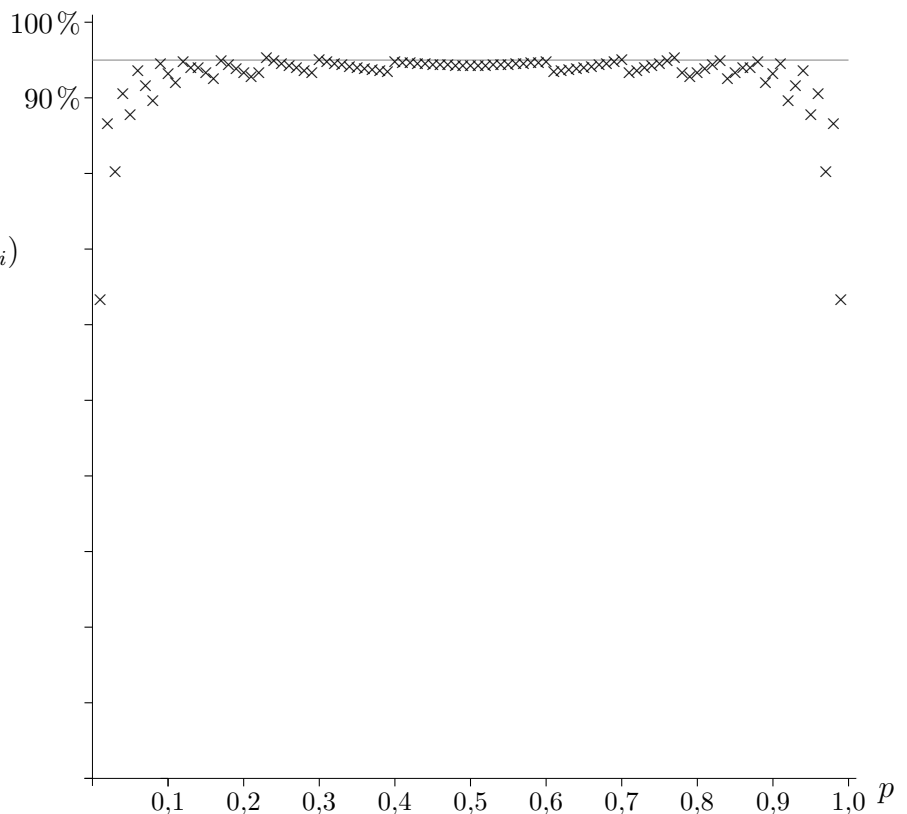


Wir betrachten eine binomialverteilte Zufallsvariable X , $n = 100$,
und zu jedem Stichprobenergebnis $X = k_i$, $i = 0 \dots n$, das genäherte Wald-Konfidenzintervall

$$C(k_i) = \left[h_i - 1,96 \cdot \sqrt{\frac{h_i \cdot (1 - h_i)}{n}} \mid h_i + 1,96 \cdot \sqrt{\frac{h_i \cdot (1 - h_i)}{n}} \right], \quad h_i = \frac{k_i}{n}$$

Erläutere die Funktion
und ihren Graphen.

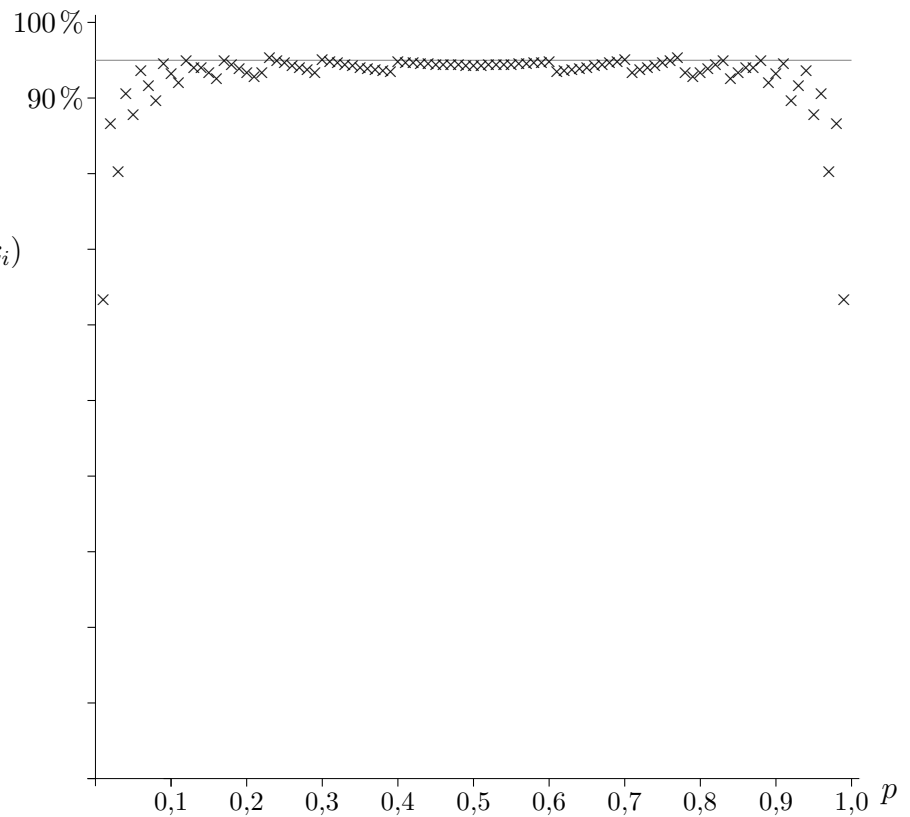
$$p \longrightarrow \sum_{\substack{i=0 \dots n \\ p \in C(k_i)}} P_p(X = k_i)$$



Überdeckungswahrscheinlichkeit

Erläutere die Funktion
und ihren Graphen.

$$p \longrightarrow \sum_{\substack{i=0 \dots n \\ p \in C(k_i)}} P_p(X = k_i)$$



Die Funktion klärt die Frage, mit welcher Wahrscheinlichkeit ein Konfidenzintervall die zugrunde liegende (unbekannte) Wahrscheinlichkeit überdeckt. Sei p gegeben. Wir schauen hinter den Vorhang. Zu jedem Stichprobenergebnis k_i gehört ein Konfidenzintervall $C(k_i)$, das mit der Wahrscheinlichkeit $P_p(X = k_i)$ gebildet wird und p entweder überdeckt oder nicht. Betrachten wir nun alle Konfidenzintervalle $C(k_i)$, die p überdecken, und zählen die zugehörigen Wahrscheinlichkeiten $P_p(X = k_i)$ zusammen. Mit dieser Wahrscheinlichkeit wird ein Konfidenzintervall, das mit einem zufälligen Stichprobenergebnis gebildet wird, p enthalten. Der Graph beinhaltet diese Wahrscheinlichkeiten für mögliche Werte p , Schrittweite 0,01.

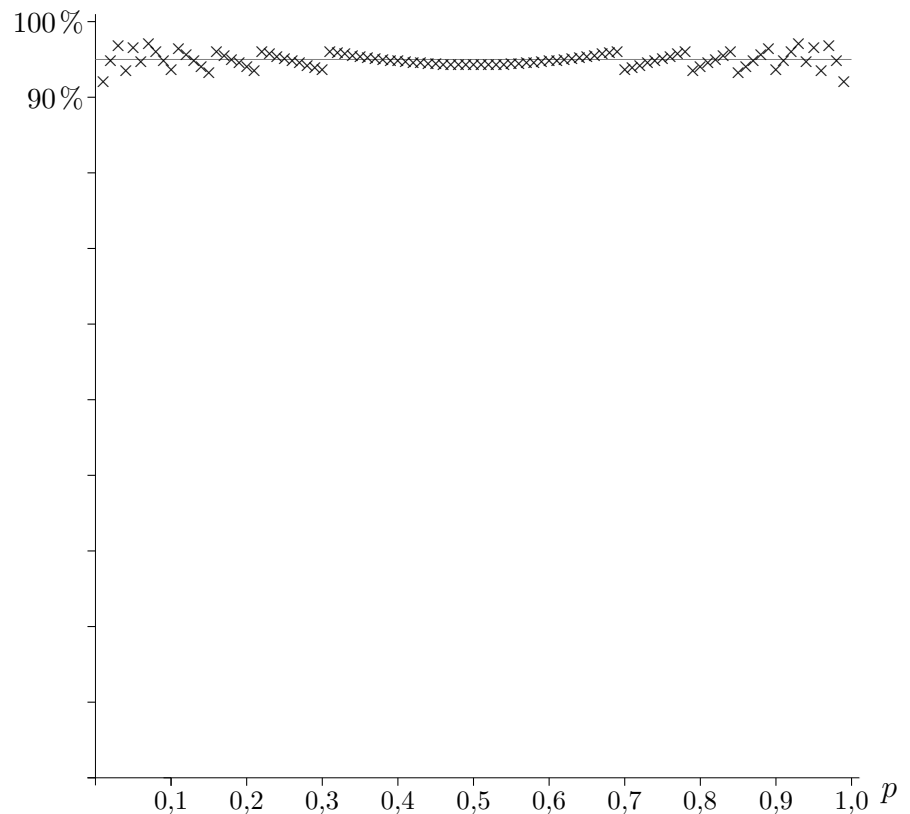
Überdeckungswahrscheinlichkeit

Eine Erhöhung der Überdeckungswahrscheinlichkeit lässt sich durch eine genauere Berechnung der Konfidenzintervalle (Wilson) erzielen:

$$C(k_i) = [a_i, b_i]$$

a_i ist Lösung der Gleichung $k_i = \mu + 1,96\sigma$ (Variable p)

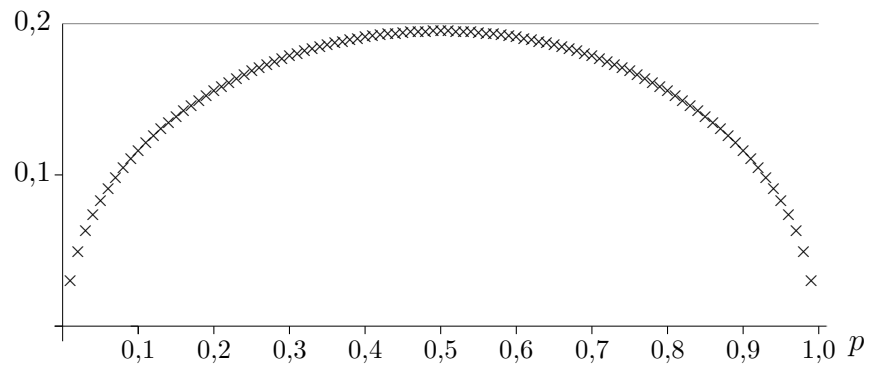
b_i löst $k_i = \mu - 1,96\sigma$



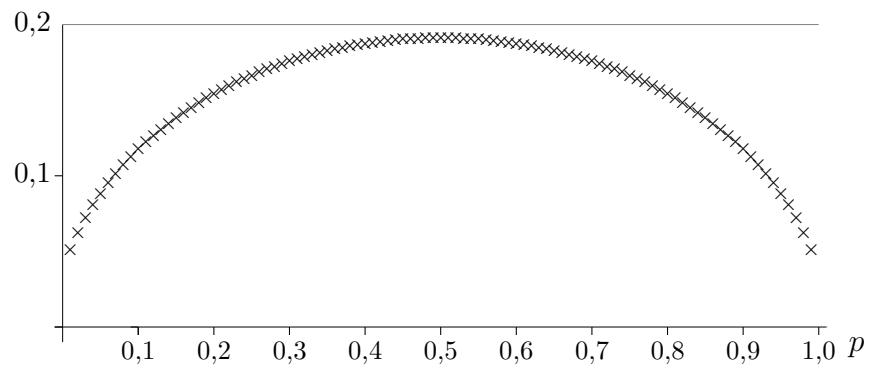
Da die Normalverteilung die Binomialverteilung für eine geringe Anzahl n von Versuchen und für sehr kleine und sehr große Trefferwahrscheinlichkeiten p nur unzulänglich approximiert, betragen nicht alle Überdeckungswahrscheinlichkeiten mindestens 95%.

Mittlere Konfidenzintervalllänge

Wald-Konfidenzintervalle



Wilson-Konfidenzintervalle



Überdeckungswahrscheinlichkeit Clopper-Pearson

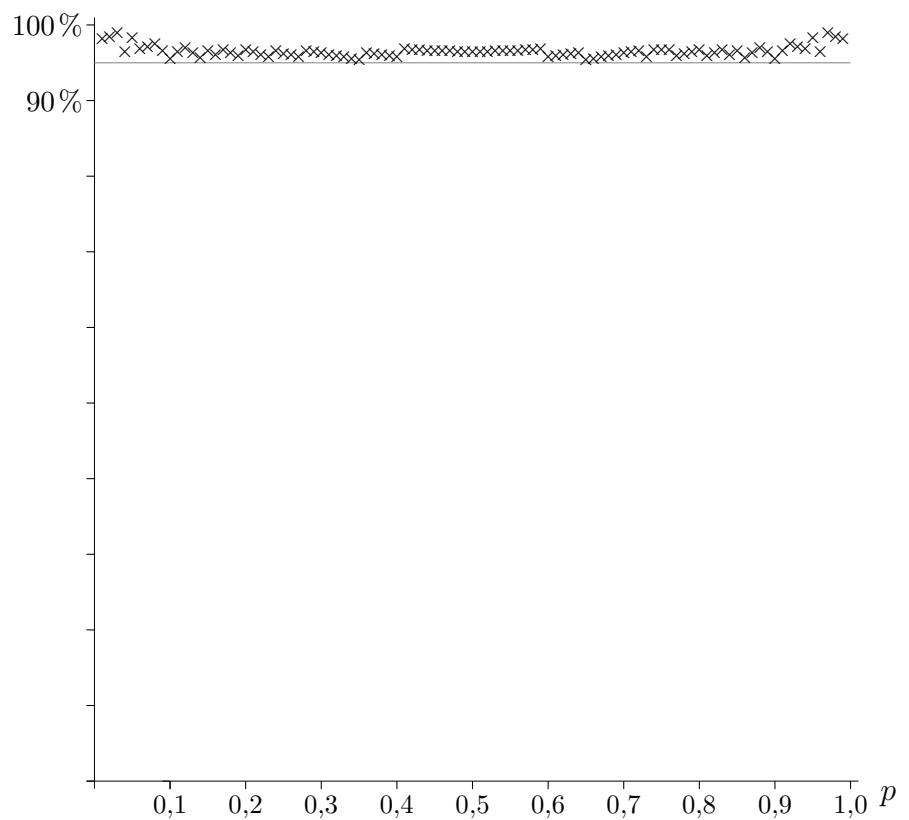
Für die exakten Konfidenzintervalle gilt:

$$C(k_i) = [a_i, b_i]$$

$$a_i = 1 - \text{BetaInv}\left(1 - \frac{1-\alpha}{2}, n - k + 1, k\right), \quad a_i = 0 \text{ für } k = 0$$

$$b_i = 1 - \text{BetaInv}\left(\frac{1-\alpha}{2}, n - k, k + 1\right), \quad b_i = 1 \text{ für } k = n$$

$$P_p^n(X \leq k) = \text{Beta}(1 - p, n - k, k + 1)$$



Konfidenzintervalle

