Intervallschätzung

Konfidenzintervalle

$\alpha \in (0,1)$, Signifikanzniveau/Fehler

Def: Ein $(1-\alpha)$-Konfidenzintervall $C_\vartheta(\vec{X})$, basierend auf einem Punktschätzer $\hat{\vartheta}$, enthält den echten (unbekannten) Parameter ϑ in (mindestens) $1-\alpha$ [$\cdot 100\%$] aller möglichen Stichproben.

$$\mathbb{P}(\vartheta \in C_\vartheta(\vec{X})) \geq 1-\alpha$$

Anmerkung: Ein Konfidenzintervall wird „besser", je

- kleiner/ kürzer/ genauer es ist.
- kleiner die Irrtumswahrscheinlichkeit α ist
- größer der Stichprobenumfang n ist

$n = \text{const.}$: Länge $\downarrow$ $\Leftrightarrow$ $\alpha \uparrow$

$(1-\alpha)$-Konfidenzintervalle für μ (Normalverteilung oder $n \geq 30$)

$$\frac{\overline{X}-\mathbb{E}(\overline{X})}{\sqrt{\mathbb{V}ar(\overline{X})}} = \frac{\overline{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\overline{X}-\mu}{\sigma} \cdot \sqrt{n} \approx \mathcal{N}(0,1)$$

(1) σ bekannt

$C_\mu(\vec{x}) = \left(\overline{X} - \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}} \, ; \, \overline{X} + \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}\right)$ zweiseitig

$C_\mu(\vec{x}) = \left(-\infty, \overline{X} + \frac{\sigma}{\sqrt{n}} \cdot z_{1-\alpha}\right)$ einseitig nach oben beschränkt

$C_\mu(\vec{x}) = \left(\overline{X} - \frac{\sigma}{\sqrt{n}} \cdot z_{1-\alpha}, \infty\right)$ einseitig nach unten beschränkt

(2) σ unbekannt (geschätzt durch s) $\sqrt{\frac{1}{n-1} \cdot \left(\overline{x^2} - \overline{x}^2\right)}$

$C_\mu(\vec{x}) = \left(\overline{X} - \frac{s}{\sqrt{n}} \cdot t_{n-1,1-\frac{\alpha}{2}} \, , \, \overline{X} + \frac{s}{\sqrt{n}} \cdot t_{n-1,1-\frac{\alpha}{2}}\right)$ zweiseitig

$C_\mu(\vec{x}) = \left(-\infty, \overline{X} + \frac{s}{\sqrt{n}} \cdot t_{n-1,1-\alpha}\right)$ einseitig nach oben beschränkt

$C_\mu(\vec{x}) = \left(\overline{X} - \frac{s}{\sqrt{n}} \cdot t_{n-1,1-\alpha}, \infty\right)$ einseitig nach unten beschränkt

__Anmerkung:__ $\frac{\overline{X}-\mu}{s} \sqrt{n} \sim t(n-1)$ „Student t-Verteilung"

$X_1, \ldots, X_n$ i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$

Beweis: zweiseitige Intervalle aus (1). Rest analog.

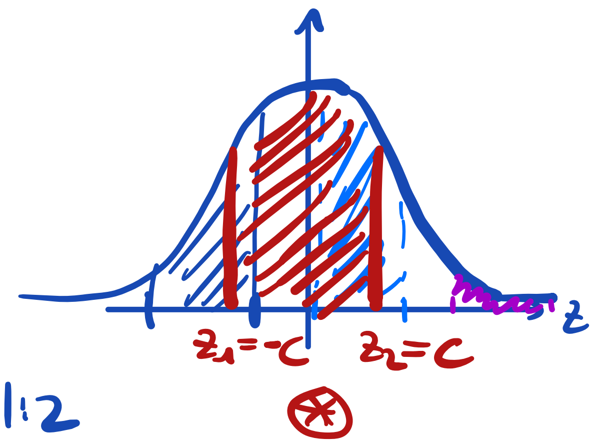
$$\mathbb{P}(a \le \bar{X} \le b) = \mathbb{P}\left(\underbrace{\frac{a - \mathbb{E}(\bar{X})}{\sqrt{Var(\bar{X})}}}_{z_1} \le \underbrace{\frac{\bar{X} - \mathbb{E}(\bar{X})}{\sqrt{Var(\bar{X})}}}_{\frac{\bar{X}-\mu}{\sigma}\cdot\sqrt{n}} \le \underbrace{\frac{b - \mathbb{E}(\bar{X})}{\sqrt{Var(\bar{X})}}}_{z_2} \right) = \mathbb{P}\left(z_1 \le \frac{\bar{X}-\mu}{\sigma}\sqrt{n} \le z_2\right)$$

$$= \mathbb{P}\left(z_1 \cdot \frac{\sigma}{\sqrt{n}} \le \bar{X} - \mu \le z_2 \cdot \frac{\sigma}{\sqrt{n}}\right) = \mathbb{P}\left(\bar{X} - z_2 \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} - z_1 \cdot \frac{\sigma}{\sqrt{n}}\right) \ge 1 - \alpha$$

$$\overset{(*)}{=} \mathbb{P}\left(\boxed{\bar{X} - c \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + c \cdot \frac{\sigma}{\sqrt{n}}} \right) \ge 1 - \alpha$$

→ Konfidenzintervall

$$\Phi\left(\frac{\bar{X} + c \cdot \frac{\sigma}{\sqrt{n}} - \bar{X}}{\frac{\sigma}{\sqrt{n}}}\right) - \Phi\left(\frac{\bar{X} - c \cdot \frac{\sigma}{\sqrt{n}} - \bar{X}}{\frac{\sigma}{\sqrt{n}}}\right)$$



$$= \Phi(c) - \Phi(-c) = 2 \cdot \Phi(c) - 1 \ge 1 - \alpha \quad | +1 \quad | :2$$

$$\Phi(c) \ge 1 - \frac{\alpha}{2} \quad | \; \Phi^{-1}(\ldots)$$

$$\underline{\underline{c \ge}} \; \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) =: \underline{\underline{z_{1-\frac{\alpha}{2}}}}$$

$(1-\alpha)$-Konfidenzintervalle für σ^2 (Normalverteilung)

(1) μ bekannt

$$C_{\sigma^2}(\vec{X}) = \left(\frac{n \cdot S_*^2}{\chi^2_{n,1-\frac{\alpha}{2}}} , \frac{n \cdot S_*^2}{\chi^2_{n,\frac{\alpha}{2}}} \right) \quad \text{mit} \quad S_*^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \mu)^2$$ zweiseitig

$$C_{\sigma^2}(\vec{X}) = \left(0 , \frac{n \cdot S_*^2}{\chi^2_{n,\alpha}} \right)$$ einseitig nach oben beschränkt

$$C_{\sigma^2}(\vec{X}) = \left(\frac{n \cdot S_*^2}{\chi^2_{n,1-\alpha}} , \infty \right)$$ einseitig nach unten beschränkt

(2) μ unbekannt

$$C_{\sigma^2}(\vec{X}) = \left(\frac{(n-1) \cdot S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}} , \frac{(n-1) \cdot S^2}{\chi^2_{n-1,\frac{\alpha}{2}}} \right) \quad \text{mit} \quad S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2$$ zweiseitig

$$C_{\sigma^2}(\vec{X}) = \left(0 , \frac{(n-1) \cdot S^2}{\chi^2_{n-1,\alpha}} \right)$$ einseitig nach oben beschränkt

$$C_{\sigma^2}(\vec{X}) = \left(\frac{(n-1) \cdot S^2}{\chi^2_{n-1,1-\alpha}} , \infty \right)$$ einseitig nach unten beschränkt

Anmerkung: $$\frac{n-1}{\sigma^2} \cdot S^2 \sim \chi^2(n-1)$$

$n\bar{x}(1-\bar{x}) > 9$

$(1-\alpha)$-**Konfidenzintervall für p** (Binomialverteilung und n groß) ZGWS

$$C_p(\vec{X}) = \left(\bar{X} - \sqrt{\frac{\bar{X}\cdot(1-\bar{X})}{n}} \cdot z_{1-\frac{\alpha}{2}} \,,\ \bar{X} + \sqrt{\frac{\bar{X}\cdot(1-\bar{X})}{n}} \cdot z_{1-\frac{\alpha}{2}} \right)$$ zweiseitig

$$C_p(\vec{X}) = \left(0 \,,\ \bar{X} + \sqrt{\frac{\bar{X}\cdot(1-\bar{X})}{n}} \cdot z_{1-\alpha} \right)$$ einseitig nach oben beschränkt

$$C_p(\vec{X}) = \left(\bar{X} - \sqrt{\frac{\bar{X}\cdot(1-\bar{X})}{n}} \cdot z_{1-\alpha} \,,\ 1 \right)$$ einseitig nach unten beschränkt

Anmerkung: $X_1, \ldots, X_n$ i.i.d., $X_i \sim B(1,p)$

$$Var(\bar{X}) = \frac{Var(X_i)}{n} = \frac{p\cdot(1-p)}{n} \approx \frac{\bar{X}\cdot(1-\bar{X})}{n}$$

für „große" n (ZGWS)

→ für μ

Aufgabe: Im Auftrag einer Winzergenossenschaft soll für die durchschnittliche Abfüllmenge einer Flaschenabfüllanlage, mit der 700 ml-Weinflaschen gefüllt werden, ein 99%-Schätzintervall (zweiseitig) bestimmt werden. Die Abfüllmenge X wird dabei als normalverteilt mit einer Standardabweichung von 10 ml angesehen. Es werden $n = 10$ auf dieser Anlage abgefüllte Flaschen ausgewählt und die Füllmenge kontrolliert. Die Stichprobe ergab folgende Werte (Angabe in ml): $\sigma = 10$

700	706	698	695	695	705	698	710	705	700

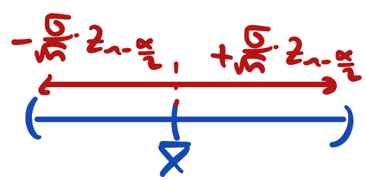
(a) Berechnen Sie das gesuchte Konfidenzintervall. Wie würde das Konfidenzintervall lauten, wenn die Standardabweichung nicht bekannt wäre?

(b) Wie groß muss der Stichprobenumfang mindestens sein, damit die Länge des 99%-Konfidenzintervalls höchstens 1 ml beträgt?

(c) Wie groß muss das Signifikanzniveau $\rightarrow \alpha$ gewählt werden, damit mit nur 40 Messungen für die obige Flaschenabfüllanlage erreicht werden kann, dass das zugehörige Konfidenzintervall höchstens 1 ml breit ist? Länge ≤ 1 mL

(d) Bestimmen Sie aus der Stichprobe auch ein zweiseitiges Konfidenzintervall für σ^2 zum Signifikanzniveau von 5%.

(a) $C_\mu(\vec{x}) = \left(\bar{X} - \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}} \,;\, \bar{X} + \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}\right) = (693{,}0545 \,;\, 709{,}3455)$



$n = 10$

$\sigma = 10$

$\bar{X} = \frac{1}{10} \cdot \sum_{i \in J} x_i = \frac{1}{10}(700 + 706 + \ldots + 700) = 701{,}2$

$\alpha = 0{,}01 \Rightarrow z_{1-\frac{\alpha}{2}} = z_{0{,}995} \approx 2{,}57582930$

(b) $C_\mu(\vec{x}) = \left(\bar{X} - \frac{S}{\sqrt{n}} \cdot t_{n-1,1-\frac{\alpha}{2}} ; \bar{X} + \frac{S}{\sqrt{n}} \cdot t_{n-1,1-\frac{\alpha}{2}}\right) = (696{,}0090 ; 706{,}3910)$

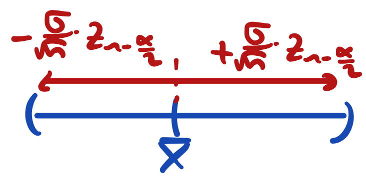
$$\bar{x} = \dots = 701{,}2$$

$$\left.\begin{array}{l} n = 10 \\ \alpha = 0{,}01 \end{array}\right\} t_{9;0{,}995} = 3{,}250$$

$$s^2 = \frac{n}{n-1}\left(\overline{x^2} - \bar{x}^2\right) = \frac{10}{9}\cdot\left(491704{,}4 - 701{,}2^2\right) = \frac{1148}{45}$$

$$\overline{x^2} = \frac{1}{10}\cdot\left(700^2 + 706^2 + \dots + 700^2\right) = 491704{,}4$$

(c) $C_\mu(\vec{x}) = \Big(\underbrace{\bar{X} - \frac{\sigma}{\sqrt{n}}\cdot z_{1-\frac{\alpha}{2}}}_{C_u} ; \underbrace{\bar{X} + \frac{\sigma}{\sqrt{n}}\cdot z_{1-\frac{\alpha}{2}}}_{C_o}\Big)$



$n = 40$
$\sigma = 10$

$$\text{Länge} = C_o - C_u = \boxed{2\cdot\frac{\sigma}{\sqrt{n}}\cdot z_{1-\frac{\alpha}{2}} \leq 1} \quad \Big| \cdot \frac{\sqrt{n}}{20}$$

$$z_{1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}}{20} = \frac{\sqrt{40}}{20} = \frac{\sqrt{10}}{10} \approx 0{,}3162 \approx 0{,}32 \quad \big| \Phi(\dots)$$

am Rand

$$1 - \frac{\alpha}{2} \leq 0{,}625516$$

in der Tabelle

$$\Rightarrow \underline{\underline{\alpha \geq 0{,}7490}}$$

(d) $C_{\sigma^2}(\vec{X}) = \left(\frac{(n-1)\cdot s^2}{\chi^2_{n-1;1-\frac{\alpha}{2}}} ; \frac{(n-1)\cdot s^2}{\chi^2_{n-1;\frac{\alpha}{2}}} \right) = (12{,}0842 \; ; \; 85{,}0374)$

$s^2 = \frac{1148}{45}$

$\left.\begin{array}{l} n=10 \\ \alpha=0{,}05 \end{array}\right\} \begin{array}{l} \chi^2_{9;0.975} = 19 \\ \chi^2_{9;0{,}025} = 2{,}7 \end{array}$

Aufgabe: Von den für eine Wahlhochrechnung zufällig befragten Bürgern hatten 10% für den Bund der Unzufriedenen (BDU) gestimmt. Geben Sie ein zweiseitiges Konfidenzintervall zum Niveau 0.99 für das tatsächliche Wahlergebnis p des BDU an, falls die Zahl der Befragten (a) $n = 50$ und (b) $n = 1000$ betragen hatte.

$$C_p(\vec{X}) = \left(\bar{X} - \sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \cdot z_{1-\frac{\alpha}{2}} \;;\; \bar{X} + \sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \cdot z_{1-\frac{\alpha}{2}} \right)$$

$$\bar{X} = 0{,}1$$

$$\alpha = 0{,}01 \Rightarrow z_{1-\frac{0{,}01}{2}} = z_{0{,}995} \approx 2{,}57582930$$

(a) $n=50$: $\underline{\underline{C_p(\vec{X}) = (0 \;;\; 0{,}2093)}}$

(b) $n=1000$: $\underline{\underline{C_p(\vec{X}) = (0{,}0756 \;;\; 0{,}1244)}}$

Aufgabe: Der Intelligenzquotient von Mathematikern kann als eine normalverteilte Zufallsgröße X angenommen werden, d.h. $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$. Für die 101 Testpersonen ergaben sich der Stichprobenmittelwert $\overline{x} = 120$ und die Stichprobenvarianz $s_X^2 = 228$.

(a) Bestimmen Sie das einseitig nach unten beschränkte 95%-Konfidenzintervall für den wahren Mittelwert des Intelligenzquotienten von Mathematikern.

(b) Wie groß muss der Stichprobenumfang mindestens sein, damit das zweiseitige 95%-Konfidenzintervall für den wahren Mittelwert des Intelligenzquotienten von Mathematikern bei bekannter Varianz $\sigma_X^2 = 225$ höchstens die Länge 4 hat?

(a) $C_\mu(\vec{x}) = \left(\bar{x} - \frac{s}{\sqrt{n}} \cdot t_{n-1,1-\alpha} \,;\, \infty\right) = \underline{\underline{(117{,}5059 \,,\, \infty)}}$

$\bar{x} = 120$

$s^2 = 228$

$\left.\begin{array}{l} n = 101 \\ \alpha = 0{,}05 \end{array}\right\} t_{100,\,0{,}95} = 1{,}66$

(b) $C_\mu(\vec{x}) = \left(\bar{x} - \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}} \,;\, \bar{x} + \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}\right)$

$\sigma^2 = 225, \quad \alpha = 0{,}05 \Rightarrow z_{0{,}975} \approx 1{,}96$

$\text{Länge} = 2 \cdot \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}} \leq 4 \iff n \geq \left(\frac{1}{2} \cdot \sigma \cdot z_{1-\frac{\alpha}{2}}\right)^2 = 216{,}09 \rightarrow \underline{\underline{n \geq 217}}$

Aufgabe: Ein eventuell manipulierter Würfel soll bezüglich der Wahrscheinlichkeit des Auftretens einer „6“ beurteilt werden. Der Anteil der „6“-en bei einer Stichprobe mit 1000 Würfen beträgt 200.

(a) Bestimmen Sie ein möglichst kleines, zweiseitige Konfidenzintervall zum Konfidenzniveau 0,95. $1-\alpha$ Verwenden Sie dabei eine geeignete Approximation der Verteilung des Schätzers für p. Ist davon auszugehen, dass der Würfel manipuliert ist?

(b) Wieviele Würfe sind notwendig, damit die Länge des Konfidenzintervalls aus Aufgabenteil (a) höchstens 0,02 beträgt?

(c) Wie groß muss die Irrtumswahrscheinlichkeit gewählt werden, damit mit nur 1000 Würfen erreicht werden kann, dass das zugehörige Konfidenzintervall höchstens 0,02 breit ist? Ist dieses Konfidenzniveau für statistische Untersuchungen geeignet?

(a) $$C_p(\vec{x}) = \left(\bar{x} - \sqrt{\frac{\bar{x}\cdot(1-\bar{x})}{n}}\cdot z_{1-\frac{\alpha}{2}} \; ; \; \bar{x} + \sqrt{\frac{\bar{x}\cdot(1-\bar{x})}{n}}\cdot z_{1-\frac{\alpha}{2}}\right) = \underline{\underline{(0{,}1752 \; ; \; 0{,}2248)}}$$

$$n = 1000$$

$$p_0 = \frac{1}{6} \approx 0{,}1667 \notin C_p(\vec{x})$$

$\Rightarrow$ signifikante Abweichung

$$\bar{x} = \frac{1}{n}\cdot\sum_{i=1}^{n} x_i = \frac{1}{1000}\cdot 200 = \frac{1}{5}$$

am Rand

$$\alpha = 0{,}05 \Rightarrow z_{\boxed{0{,}975}} \approx 1{,}96$$

in der Tabelle

(b) $$\text{Länge} = 2\cdot\sqrt{\frac{\bar{x}(1-\bar{x})}{n}}\cdot z_{1-\frac{\alpha}{2}} \leq 0{,}02 \Rightarrow n \geq \left(100\cdot\sqrt{\bar{x}(1-\bar{x})}\cdot z_{1-\frac{\alpha}{2}}\right)^2 \approx 6146{,}56$$

$$\Rightarrow \underline{\underline{n \geq 6147}}$$

(c) $\text{Länge} = 2 \cdot \sqrt{\frac{\bar{x}(1-\bar{x})}{n}} \cdot z_{1-\frac{\alpha}{2}} \leq 0{,}02$

am Rand

$$\Rightarrow z_{1-\frac{\alpha}{2}} \leq 0{,}01 \cdot \sqrt{\frac{n}{\bar{x} \cdot (1-\bar{x})}} = 0{,}01 \cdot \sqrt{\frac{1000}{92 \cdot 0{,}8}} \approx 0{,}79 \quad \Big| \, \Phi(\ldots)$$

in die Tabelle schauen

$$1 - \frac{\alpha}{2} \leq 0{,}785236$$

in der Tabelle

$\Rightarrow \underline{\underline{\alpha \geq 0{,}4295}}$ (sehr hoher Fehler)