6. Parameterschätzung, Konfidenzintervalle

Lernziele:

- Sie können Konfidenzintervalle für einen unbekannten Erwartungswert berechnen.
- Sie sehen den Zusammenhang zwischen der Konstruktion eines Konfidenzintervalls und dem ZGWS.
- Sie verstehen, dass Konfidenzintervalle zufällig sind.
- Sie verstehen die Zusammenhänge zwischen Stichprobenumfang, Konfidenzniveau und Konfidenzintervall und können dieses Wissen auf verschiedene Aufgabentypen anwenden.

Literatur:

- Teschl Band 2, Kap. 30.1 - 30.3
- Zucchini, Kap. 7.6.2 + 7.6.3
- Arens et al., Kap 40.4

Situation:

Basierend auf i.i.d. Zufallsvariablen $X_1, \ldots, X_n$ (Messungen) soll ein unbekannter Parameter der Verteilung der Grundgesamtheit, z. B. der Erwartungswert μ, geschätzt werden.

Voraussetzung: ZGWS ist anwendbar!

Bei kleinem Stichprobenumfang ($n < 30$) ist die Grundgesamtheit näherungsweise normalverteilt bzw. der Stichprobenumfang ist hinreichend groß ($n \geq 30$), so dass die Summe bzw. der Mittelwert der X_i nach dem zentralen Grenzwertsatz näherungsweise normalverteilt ist.

Punktschätzer

- Für Erwartungswert: Stichprobenmittel

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

- Für Varianz: Stichprobenvarianz

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Eigenschaften:

- Schätzwert für wahren Parameter, aber keine Aussage über Unsicherheit der Schätzung
- Geringe Sicherheit, dass wahrer Parameterwert getroffen wird

Intervallschätzer

$$\bar{X} \sim N_{\mu, \frac{\sigma^2}{n}}$$

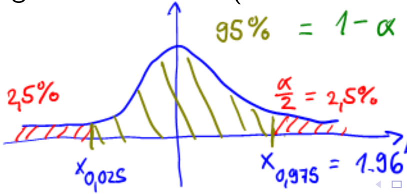
Mit Hilfe des ZGWS wird ein **Konfidenzintervall** konstruiert, das den wahren Parameter mit einer vorgegebenen Sicherheit, dem sog. **Konfidenzniveau** $1-\alpha$ überdeckt.

(α: Irrtumswahrscheinlichkeit)

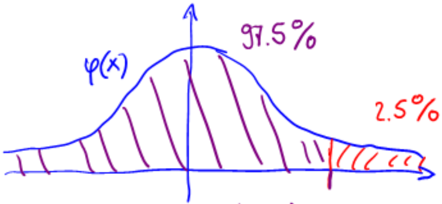
Eigenschaften:

- Intervall für wahren Parameter, aber mit vorgegebener Sicherheit liegt dieser im Intervall
- Vorgabe einer großen Sicherheit (95% bzw. 99%) evtl. auch 90%

$95\% = 1-\alpha \Rightarrow \alpha = 5\%$



$x_{0{,}975} = 1.96 = \text{qnorm}(0.975) = \Phi^{-1}(0.975)$



$\Phi^{-1}(0.975) = x_{0.975}$

Φ: Verteilungsfkt. der Standardnormalvert.

Φ^{-1}: Quantilsfkt. der Standardnormalvert.

$$P(-a \leq \bar{x} \leq a) \geq 0.95$$

$$P\Big(\underbrace{-x_{0.975}}_{-1.96} \leq \underbrace{\frac{\bar{x} - \overset{\text{unbekannt}}{\mu}}{\sigma}\sqrt{n}}_{\sim N_{0,1}} \leq \underbrace{x_{0.975}}_{1.96}\Big) \geq 0.95$$

bekannt

Daraus lässt sich jetzt ein Intervall für μ mit der vorgeg. Sicherheit 95% bestimmen:

$$P\left(-1.96 \cdot \frac{\sigma}{\sqrt{n}} \leq \bar{x} - \mu \leq 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right) \geq 0.95$$

$$P\left(\underbrace{-\bar{x}}_{\cdot(-1)} - 1.96 \cdot \frac{\sigma}{\sqrt{n}} \leq \underbrace{-\mu}_{\cdot(-1)} \leq -\bar{x} + \underbrace{1.96}_{\cdot(-1)} \cdot \frac{\sigma}{\sqrt{n}}\right) \geq 0.95$$

$$P\left(\bar{x} + 1.96 \cdot \frac{\sigma}{\sqrt{n}} \geq \mu \geq \bar{x} - 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right) \geq 0.95$$

$$P\left(\underbrace{\bar{x} - 1.96 \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96 \cdot \frac{\sigma}{\sqrt{n}}}_{\text{95\%-Konfidenzintervall für } \mu}\right) \geq 0.95$$

Konfidenzintervall für unbekannten Erwartungswert μ bei bekannter Varianz σ^2

Konstruktion des Konfidenzintervalls:

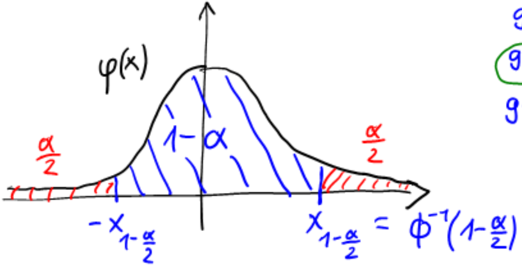
- $(n > 30)$
  n Wiederholungen eines Zufallsexperiments (z. B. Messungen): $X_1, \ldots, X_n$ sind i.i.d. ZV mit $E[X_i] = \mu$ und $Var[X_i] = \sigma^2$
  unbekannt bekannt
- Nach ZGWS gilt näherungsweise: $\frac{\bar{X}-\mu}{\sigma}\sqrt{n} \sim N_{0,1}$
- Also liegt $\frac{\bar{X}-\mu}{\sigma}\sqrt{n}$ mit einer Wahscheinlichkeit von 95% zwischen dem 2.5%- und dem 97.5%-Quantil der Standardnormalverteilung Φ, d. h. zwischen $\Phi^{-1}(0.025) \approx -1.96$ und $\Phi^{-1}(0.975) \approx 1.96$.
- Dann liegt μ mit einer Wahscheinlichkeit von 95% zwischen $\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}$ und $\bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$
- 95%-Konfidenzintervall für μ: $I = \left]\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right[$,
  d. h. $P(\mu \in I) \geq 0.95$
  abhängig von $\bar{X}, \sigma, n, \alpha$

Allgemein:

Das Konfidenzintervall für den Erwartungswert μ bei bekannter Varianz σ^2 zum Konfidenzniveau $1-\alpha$ ist (α: Irrtumswahrscheinlichkeit)

$$I = \left]\bar{X} - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}, \bar{X} + \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}\right[$$

Konfidenzniveau $1-\alpha$	$\frac{\alpha}{2}$	$\Phi^{-1}\left(1-\frac{\alpha}{2}\right)$
90%	5%	$\Phi^{-1}(0.95) \approx 1.645$
95%	2.5%	$\Phi^{-1}(0.975) \approx 1.96$
99%	0.5%	$\Phi^{-1}(0.995) \approx 2.576$



Beispiel 6.1:

Ein Signal mit Erwartungswert μ wird von A nach B übertragen. An B wird ein verrauschtes Signal $\mu + R$ empfangen mit $R \sim N_{0,4}$. Bei 9-maliger Messung des Signals erhält man:

$4 = \sigma^2$

! Voraussetzung für Anwendung des ZGWS

$x_1 = 5, x_2 = 8.5, x_3 = 12, x_4 = 15, x_5 = 7, x_6 = 9, x_7 = 7.5, x_8 = 6.5, x_9 = 10.$

$\bar{x} = \frac{161}{18} \approx 9$

Bestimmen Sie das Konfidenzintervall für μ zum Konfidenzniveau 95%.

$= 1-\alpha$

$\Rightarrow \Phi^{-1}(0.975) = 1.96$

$$I = \left]\bar{x} - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{\sigma}{\sqrt{n}} \,;\, \bar{x} + \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{\sigma}{\sqrt{n}}\right[= \left]9 - 1.96\cdot\frac{2}{3} \,;\, 9 + 1.96\cdot\frac{2}{3}\right[$$

$$= \left]7.69 \,;\, 10.31\right[$$

Konfidenzintervall für unbekannten Erwartungswert μ bei unbekannter Varianz

Konstruktion des Konfidenzintervalls:

unbekannt

- Verwendung der Stichprobenvarianz S^2 als Schätzer für σ^2
- Nach ZGWS gilt näherungsweise: $\frac{\bar{X}-\mu}{S}\sqrt{n} \sim t_{n-1}$
- Wegen der Symmetrie der t-Verteilung ist das Konfidenzintervall für den Erwartungswert μ bei unbekannter Varianz geschätzt durch S^2 zum Konfidenzniveau $1-\alpha$

$$I = \left]\bar{X} - t_{n-1}^{-1}\left(1-\frac{\alpha}{2}\right)\frac{S}{\sqrt{n}}, \bar{X} + t_{n-1}^{-1}\left(1-\frac{\alpha}{2}\right)\frac{S}{\sqrt{n}}\right[$$

qt$(1-\frac{\alpha}{2}, n-1)$

Übung:

qt$(0.975, 8) \approx 2.306$

Wie ändert sich für Beispiel 6.1 das Konfidenzintervall bei unbekannter Varianz?

$I =]6.63; 11.37[$

Zusammenhänge

- Falls I das $(1-\alpha)$–Konfidenzintervall für μ, dann: $P(\mu \in I) \geq 1-\alpha$.
  Welche Größe ist hier zufällig? $\bar{X}$, S ← Länge des Intervalls

MP des Intervalls

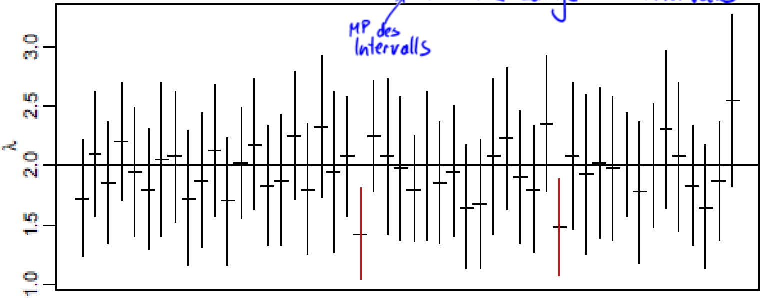


Abbildung: 95%–Konfidenzintervalle zu 50 Stichproben vom Umfang $n = 40$

- Wie verändert sich das $(1-\alpha)$–Konfidenzintervall, wenn man
  n vergrößert $\Rightarrow$ I wird kürzer
  bzw. α verkleinert? $\Rightarrow$ I wird länger

Aufgabentypen

(1) **Geg.:** Stichprobe vom Umfang n, Konfidenzniveau $1-\alpha$
Ges.: Konfidenzintervall I

(2) **Geg.:** $\bar{X}$, σ, Konfidenzniveau $1-\alpha$, Konfidenzintervall I der Länge
geg. $L = 2\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}$
Ges.: Stichprobenumfang n ges.

$$\sqrt{n} > 2\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\frac{\sigma}{L}$$

(3) **Geg.:** Stichprobe vom Umfang n, Konfidenzintervall I der Länge L
Ges.: Konfidenzniveau $1-\alpha$ $\quad L = 2\cdot\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{\sigma}{\sqrt{n}}$

$$\Leftrightarrow \frac{L\sqrt{n}}{2\sigma} = \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \;\Big|\; \Phi(\,) \Rightarrow 1-\frac{\alpha}{2} = \underbrace{\Phi\left(\frac{L}{2}\frac{\sqrt{n}}{\sigma}\right)}_{\text{pnorm}\left(\frac{L\sqrt{n}}{2\sigma}\right)}$$

Beispiel 6.2:

Bei 10 Messungen X_i der Elastizitätsgrenze von Baustahl ergibt sich ein Mittelwert $\bar{X} = 347.5$ [N/mm^2]. Die Varianz der X_i sei $\sigma^2 = 105$. $\Rightarrow \sigma = \sqrt{105}$

Wie viele Messungen $\overset{\text{ges.}}{n}$ müssen durchgeführt werden, damit die Länge des Konfidenzintervalls zum Konfidenzniveau $\underbrace{95\%}_{1-\alpha}$ höchstens $\underbrace{8}_{=L}$ [N/mm^2] beträgt?

$$\sqrt{n} > 2 \cdot 1.96 \cdot \frac{\sqrt{105}}{8} \quad \Longrightarrow \quad n \geq 26$$

Beispiel 6.3:

Bei der Befragung von 124 Studierenden nach ihrer Körpergröße ergibt sich ein Mittelwert $\bar{X} = 1.814$ [m] und eine Stichprobenstandardabweichung $S = 0.074$ [m]. Bestimmen Sie das Konfidenzniveau zu dem das Konfidenzintervall $I =]1.800, 1.828[$ für die durchschnittlich zu erwartende Körpergröße der Studentengruppe angegeben werden kann.

$n =$

ges.

$$1 - \frac{\alpha}{2} = \underbrace{pt\left(\frac{L \cdot \sqrt{124}}{2\,S}, 123\right)}_{0.9814} = pt\left(2.1067, 123\right)$$

$$L = 1.828 - 1.800 = 0.028$$

$$\Rightarrow \quad \alpha = 2 \cdot (1 - 0.9814) \approx 3.7\,\%$$

d.h. Konfidenzniveau 96.3 %