Querschnittsbereich 1:
Epidemiologie, Medizinische Biometrie und Medizinische Informatik

- Übungsmaterial -

Erstellt von Mitarbeitern des IMISE und des ZKS Leipzig

4. Übung

Konfidenzintervalle für Anteile und Mittelwerte

WS 2018/19

Konfidenzintervalle für Anteile und Mittelwerte

Beispiel für diagnostische Tests (Übung 2)

Güte des genetischen Kriteriums

Pathologie	**Genetisches Kriterium** ≥1 ∊4 Allel	kein ∊4 Allel	
Alzheimer +	1142	628	1770
Alzheimer –	133	285	418
	1275	913	2188

Sensitivität = …………   Spezifität = ……..………

Konfidenzintervalle für Anteile und Mittelwerte

- **Punktschätzer**
  Schätzer, die einen Wert angeben, nennt man auch Punktschätzer.
- **Intervallschätzer**
  Punktschätzer sagen nichts über die Genauigkeit der Schätzung aus. Deshalb konstruiert man um den Punktschätzer ein Intervall (Konfidenzintervall oder kurz KI).

Konfidenzintervalle für Anteile und Mittelwerte

1. Konfidenzintervall für Anteile

(Synonym: relative Häufigkeiten)

Konfidenzintervall für einen Anteil

Aufgabe 1

→ **Berechnen Sie das 95%-Konfidenzintervall für die Sensitivität von** $\frac{1142}{1770} = 64{,}5\%$ **.**

Konfidenzintervall für einen Anteil

Anteil in Stichprobe: $\hat{p} = \frac{k}{n}$

$k\ldots$ Ereignisse
$n\ldots$ Stichprobenumfang

Anteil in Grundgesamtheit: p (wahrer Wert)

Voraussetzung: $n \cdot \hat{p} \cdot (1-\hat{p}) \geq 9$

Formel: $$\underbrace{\hat{p} - z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}\cdot(1-\hat{p})}{n}}}_{\text{untere Grenze}} \leq p \leq \underbrace{\hat{p} + z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}\cdot(1-\hat{p})}{n}}}_{\text{obere Grenze}}$$

95%-Konfidenzintervall:

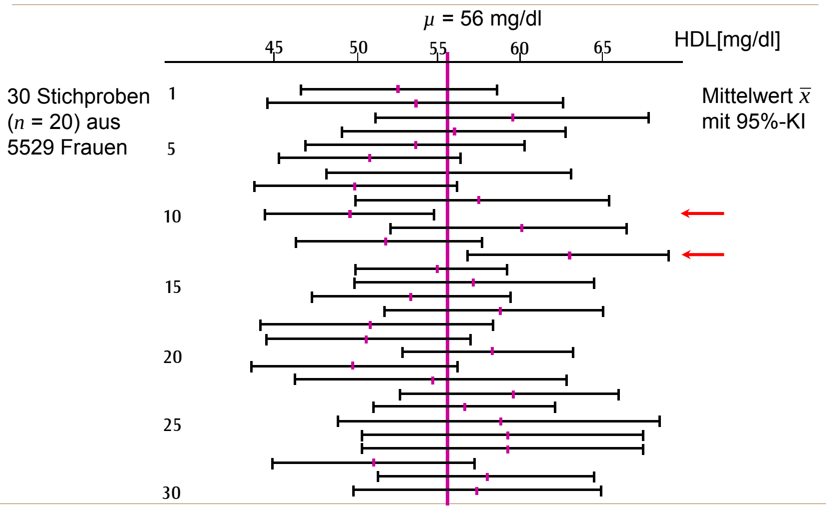
$$\hat{p} - z_{1-0{,}05/2} \cdot \sqrt{\frac{\hat{p}\cdot(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{1-0{,}05/2} \cdot \sqrt{\frac{\hat{p}\cdot(1-\hat{p})}{n}}$$

Bedeutung des Konfidenzintervalls

- Die Grenzen des 95%-Konfidenzintervalls sind so gewählt, dass das zufällige Intervall mit ca. 95%-iger Wahrscheinlichkeit den wahren Wert überdeckt.
- Bei wiederholter Anwendung werden also ca. 95% der so berechneten (festen) Intervalle den wahren Wert enthalten.

Bedeutung des Konfidenzintervalls

Beispiel: HDL-Werte (Frauen)



Konfidenzintervall für einen Anteil
Aufgabe 2

Angenommen es wäre in dem Beispiel aus Aufgabe 1 nur jeder 10. Patient untersucht worden.

→ **Berechnen Sie das 95%-Konfidenzintervall für die Sensitivität von** $\frac{114}{177} = 64{,}4\%$ **.**

→ **Vergleichen Sie das Ergebnis aus Aufgabe 1 und 2.**

Konfidenzintervall für einen Anteil
Aufgabe 3a

- Seit dem Giftgasunfall in Soveso hatten die der Dioxin-Belastung ausgesetzten Männer 88 Jungen und 103 Mädchen gezeugt.
- Die Geschlechterverteilung in der Normalbevölkerung ist wie 106 Jungen zu 100 Mädchen.

→ **Berechnen Sie sowohl für Seveso als auch für die Normalbevölkerung den Anteil der Jungengeburten und vergleichen Sie beide Anteile!**

Konfidenzintervall für einen Anteil

Aufgabe 3b

Eine Schlagzeile in der Presse behauptete, dass Dioxin die Zeugung von Jungen verringert.

Ist diese Schlagzeile berechtigt?

Konfidenzintervalle für Anteile und Mittelwerte

2. Konfidenzintervall für einen Mittelwert

Konfidenzintervall für einen Mittelwert

(1) Standardabweichung σ in Grundgesamtheit <u>bekannt</u>:

$$\bar{x} - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

95%-Konfidenzintervall: $\bar{x} - z_{1-0,05/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{1-0,05/2} \frac{\sigma}{\sqrt{n}}$

(2) Standardabweichung σ in Grundgesamtheit <u>unbekannt</u>:

$$\bar{x} - t_{f;1-\alpha/2} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{f;1-\alpha/2} \frac{s}{\sqrt{n}}$$

95%-Konfidenzintervall: $\bar{x} - t_{f;1-0,05/2} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{f;1-0,05/2} \frac{s}{\sqrt{n}}$

s … Standardabweichung (aus Stichprobe geschätzt)
Freiheitsgrad $f = n - 1$
für $n > 100$ setze $t_{f;1-0,05/2} = z_{1-0,05/2} = 1{,}96$

Konfidenzintervall für einen Mittelwert

Aufgabe 4

- Dr. X hatte bei 25 Frauen Brustkrebs diagnostiziert. Die Frauen waren im Mittel $\bar{x}$ = 55,0 Jahre alt. Die berechnete Standardabweichung s ist 10,0. Aus der Literatur ist ihm bekannt, dass das Erkrankungsalter bei früheren Untersuchungen im Populationsmittel bei 62,0 Jahren liegt.
- Dr. X fragt sich, ob seine Stichprobe mit der Patientengruppe der Grundgesamtheit in Bezug auf das Alter vergleichbar ist?

→ **Berechnen Sie zur Beantwortung dieser Frage das 95%-Konfidenzintervall.**

Konfidenzintervall für einen Mittelwert
Aufgabe 5

Berechnen Sie für beide Therapiearme das 95%-KI für den „Body-mass index“:

TABLE 1. BASE-LINE CHARACTERISTICS OF THE PATIENTS.*

CHARACTERISTIC	CONVENTIONAL TREATMENT (N=783)	INTENSIVE TREATMENT (N=765)
Male sex — no. (%)	557 (71)	544 (71)
Age — yr	62.2±13.9	63.4±13.6
Body-mass index†	25.8±4.7	26.2±4.4

*Plus–minus values are means ±SD.

†The body-mass index is the weight in kilograms divided by the square of the height in meters.

(Van den Berghe et al., 2001)

Konfidenzintervalle für Anteile und Mittelwerte

Hausaufgabe

Berechnen Sie das 95%-KI für beide Arme für „Septicemia during intensive care“.
(siehe Van den Berghe et al. 2001 – table 4.)

Berechnen Sie das 95%-KI für beide Arme für „Age“.
(siehe Van den Berghe et al. 2001 – table 1.)

Beurteilen Sie mithilfe des berechneten 95%-KI, ob sich die Behandlungsgruppen signifikant unterscheiden.