

### 3 Konfidenzintervalle

- Konfidenzintervalle sind das Ergebnis von Intervallschätzungen.
- Sicheres Wissen über Grundgesamtheiten kann man anhand von Stichproben nicht gewinnen.
- Aber mit Hilfe der Statistik können Intervalle angegeben werden, innerhalb derer sich die Parameter der Grundgesamtheit wahrscheinlich bewegen. Diese Bandbreiten nennt man Konfidenzintervalle.

#### *Beispiel*

- Mit einer Wahrscheinlichkeit von 95 % wird der Stimmenanteil der Partei LILA zwischen 35 % und 41 % liegen.
- Mit einer Wahrscheinlichkeit von 99 % liegt der Ertrag mit dem Dünger XYZ zwischen 8,7 und 9,3 dz/ha.

Wie kommt man aber zu diesen Aussagen?

## 4 Auswahl eines Konfidenzintervalls

- Um ein Intervall anzugeben, in dem der geschätzte Wert für die Grundgesamtheit „wahrscheinlich“ liegt, benötigt man eine Irrtumswahrscheinlichkeit.
- Ein  $\alpha$  von 5 % bedeutet, dass diese Intervall den gesuchten Wert der Grundgesamtheit mit 95 % Wahrscheinlichkeit enthält.
- Eine (gesetzte) Irrtumswahrscheinlichkeit von 5 % bedeutet, dass das Intervall die mittleren 95% der Kennwerteverteilung überdeckt, Beziehungsweise jeweils 2,5 % auf der linken und der rechten Seite der Kennwerteverteilung werden nicht berücksichtigt.

## ***Grenzen des Vertrauensintervalls***

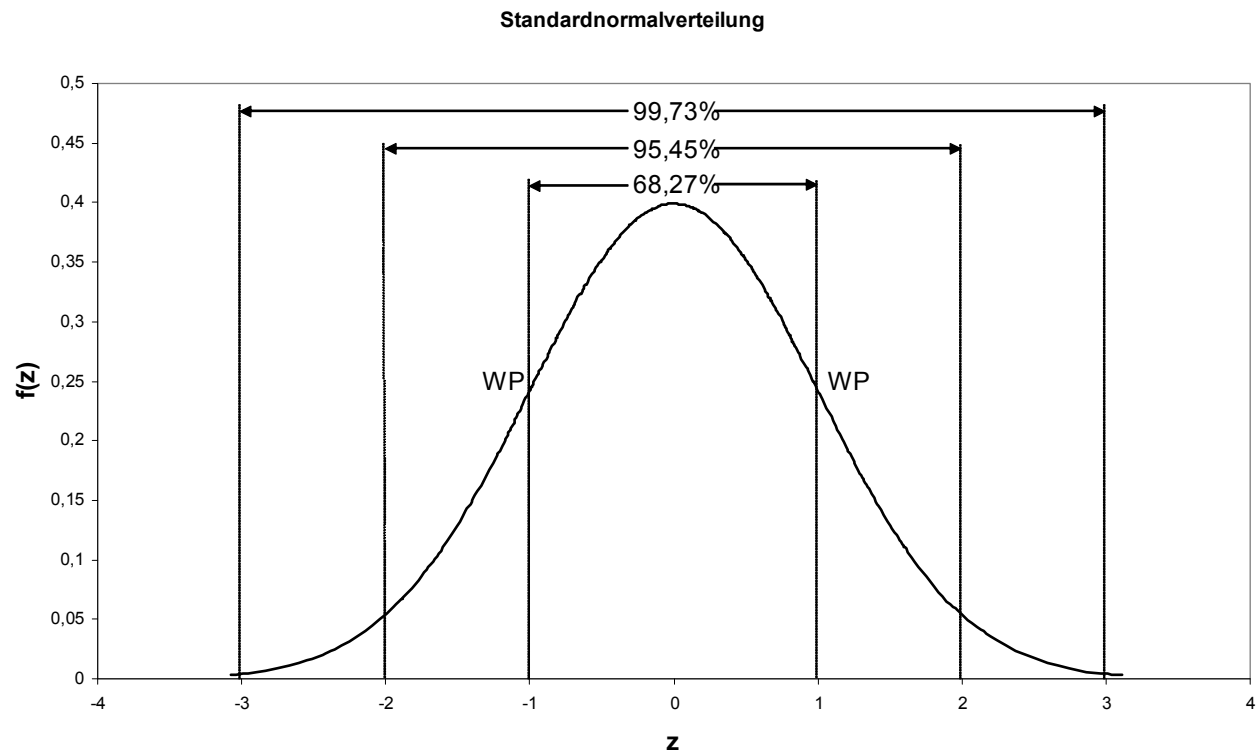
Allgemein:

- Die linke (untere) Grenze des Konfidenz-intervalls liegt bei:
  - **Schätzwert–Standardfehler \* (1- $\alpha$ /2)-Quantilwert**
- Die rechte (obere) Grenze des Konfidenz-intervalls liegt bei
  - **Schätzwert+Standardfehler \* (1- $\alpha$ /2)-Quantilwert**

Quantile sind Punkte einer nach Rang oder Größe der Einzelwerte sortierten statistischen Verteilung.

## Vorteil der Standardnormalverteilung

z	Fläche	z	Fläche
-2,576	0,005	0,675	0,75
-2,325	0,01	1,282	0,90
<b>-1,96</b>	<b>0,025</b>	<b>1,645</b>	<b>0,95</b>
<b>-1,645</b>	<b>0,05</b>	<b>1,96</b>	<b>0,975</b>
-1,282	0,10	2,325	0,99
-0,675	0,25	2,576	0,995
0	0,5		



- Etwa 68 % der Stichprobenergebnisse liegen im Bereich von  $\pm 1$  Standardfehler um den wahren Wert (in der Grundgesamtheit)
- Etwa 95 % der Stichprobenergebnisse liegen im Bereich von  $\pm 2$  Standardfehlern. Oder genauer exakt 95 % liegen im Bereich  $\pm 1,96$  Standardfehlern um den wahren Wert
- Etwa 99 % der Stichprobenergebnisse liegen im Bereich von  $\pm 2,5$  Standardfehlern. Oder genauer exakt 99 % liegen im Bereich  $\pm 2,58$  Standardfehlern um den wahren Wert

#### Hinweis

- Die Verteilungsfunktionen der wichtigsten Zufallsverteilungen sind in allen brauchbaren Statistiklehrbüchern zumindest auszugsweise tabelliert.

## ***Zentrale Bedeutung des Standardfehlers***

- Für die Berechnung des Konfidenzintervalls benötigen wir den Standardfehler (die Standardabweichung der Stichprobenverteilung)

$$SE = \hat{\delta}_{\bar{x}} = \frac{\hat{\delta}}{\sqrt{n}} \text{ (SE, standard error)}$$

Wovon hängt der Stichprobenfehler ab?

Für die Größe des Standardfehlers sind 2 Aspekte ausschlaggebend

- Stichprobengröße (je größer  $n$ , umso kleiner SE)
- Streuung der Werte in der Grundgesamtheit (Je größer die Streuung, umso größer der SE)

→ Bei Daten in großen Stichproben kann man die bekannten t-Werte nehmen:

### **Beispiel Alter**

- Stichprobe von 100 Personen aus Dresden
- Mittelwert ist 42
- Standardabweichung ist 11,1
- Der Vertrauensbereich soll 95 % betragen ( $\alpha=5\%$ ), der zugehörige t-Wert ist 1,96

- Untere Grenze  $42 - 1,96 \cdot \frac{11,1}{\sqrt{100}} = 39,8$

- Obere Grenze  $42 + 1,96 \cdot \frac{11,1}{\sqrt{100}} = 44,2$

Mit 95 % Wahrscheinlichkeit, sind die Dresdner zwischen 39,8 und 44,2 Jahren alt.



***Beispiel: Konfidenzintervall für Mittelwerte (kleine Stichprobe)***

- Eine zufällig ausgewählte Gruppe von 11 Studierenden der Soziologie hat an einem IQ-Test teilgenommen. Die resultierenden Werte sind 203, 195, 193, 193, 193, 188, 185, 184, 172, 170 und 162. Wir möchten nun das arithmetische Mittel der IQ-Werte für die entsprechende Population schätzen und als Indikator der Genauigkeit der Schätzung ein 90 % Konfidenzintervall bestimmen.

- Wir benötigen also zuerst den Stichprobenmittelwert und die geschätzte Populationsvarianz:

$$\bar{x} = \frac{203 + 195 \dots + 162}{11} = 185,27$$

$$\begin{aligned}\hat{\delta} &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \sqrt{\frac{1}{10} [(203 - 185,27)^2 + (195 - 185,27)^2 + \dots + (162 - 185,27)^2]} \\ &= \sqrt{154,818} = 12,443\end{aligned}$$

$$SE = \hat{\delta}_{\bar{x}} = \frac{\hat{\delta}}{\sqrt{n}} = \frac{12,443}{\sqrt{11}} = 3,752$$

- Jetzt benötigen wir noch die t-Werte für  $n-1 = 10$  Freiheitsgrade und eine Konfidenz von 90 %. Diese findet man in den Tabellenanhängen für  $t_{10, Konf=90\%} = 1,813$ . Nun können die Grenzen des Konfidenzintervalls bestimmt werden:

$$\text{Untere Grenze } \bar{x} - \hat{\delta}_{\bar{x}} \cdot t_{df, Konf} = 185,27 - 3,752 \cdot 1,813 = 178,47$$

$$\text{Obere Grenze } \bar{x} + \hat{\delta}_{\bar{x}} \cdot t_{df, Konf} = 185,27 + 3,752 \cdot 1,813 = 192,07$$

Wir können also mit 90 % Sicherheit sagen, dass das Intervall zwischen 178 und 192 den wahren IQ-Wert der Soziologie-Studierenden enthält.

***Was beeinflusst die Größe eines Konfidenzintervalls?***

- Stichprobengröße
- Streuung der Werte
- Größe des Effekts

**Tab.1** 95%-Konfidenzintervalle für die Differenz von Mittelwerten zweier Gruppen bei hypothetischen Daten (systolischer Blutdruck in mm Hg) mit variierender Differenz, Stichprobengröße und Variabilität.

Nr.	Stichprobenumfänge	Medikament Mittelwert (SD)	Placebo Mittelwert (SD)	Differenz der Mittelwerte	95%-Konfidenz- intervall	P-Wert
1	$n_1 = n_2 = 10$	160 (22)	180 (22)	20	[-0,7; 40,7]	0,057
2	$n_1 = n_2 = 10$	160 (15)	180 (15)	20	[5,9; 34,1]	0,008
3	$n_1 = n_2 = 20$	160 (22)	180 (22)	20	[5,9; 34,1]	0,007
4	$n_1 = n_2 = 50$	170 (22)	180 (22)	10	[1,3; 18,7]	0,025
5	$n_1 = n_2 = 5$	140 (50)	180 (50)	40	[-32,9; 112,9]	0,242
6	$n_1 = n_2 = 1000$	178 (12)	180 (12)	2	[0,9; 3,1]	< 0,001

SD = Standardabweichung

### ***Vorteil von Konfidenzintervallen***

- Liefern im Gegensatz zu den p-Werten Informationen in der gleichen Skala wie die untersuchte Variable.
- Hier lässt sich also auch etwas über die Unsicherheit zur Stärke des Effekts sagen.
- Man kann die Signifikanz direkt ablesen.

## 5 Veröffentlichung von Ergebnissen

### *Von Signifikanztests*

- keine Sternchen berichten
- Exakte p-Werte (gerundet) angeben, z.B.  $p = 0,02$ , oder bei sehr niedrigen p-Werten das niedrigste unterschrittene Signifikanzniveau
- Z.B. Kühne, M.; Böhme, R. (2006): Effekte von Informationsstand, Wissen und Einstellungsstärke von Befragten auf die Antwortstabilität in Online-Befragungen mit Selbstrekrutierung. ZUMA-Nachrichten. 59, S. 42-71.

erzeugen. Die Zentralität einer Einstellung identifiziert etwas seltener Personen mit starken Einstellungen zu einem Thema. Dieser Unterschied findet sich auch in den Mittelwerten dieser beiden Indikatoren wieder ( $M_{\text{Intensität}} = 3,82 > M_{\text{Zentralität}} = 3,28$ ;  $t(1.340) = -9,7$ ;  $p < 0,001$ ).

...

Einstellung zum Thema Studiengebühren in der Stichprobe vertreten sind. Die hier verwendeten Indikatoren der Einstellungsstärke zeigen den vermuteten Zusammenhang. Zentralität und Intensität korrelieren für sozialwissenschaftliche Fragestellungen relativ stark miteinander ( $r = 0,50$ ;  $p < 0,001$ ).

## ***Von Konfidenzintervallen***

- Die Vertrauensbereich, der Kennwert und die Grenzen sollten berichtet werden.
- Z.B. (Mittelwert[M]; 0,10; 95 % CI; 0.01 – 0.70)
- ([M]; 1,35; 95 % CI; 1,23 – 1,56)