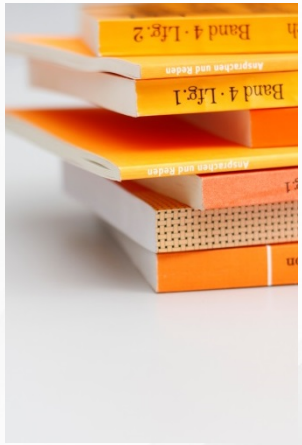
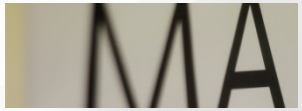
INSTITUT FÜR DEUTSCHE SPRACHE



Sandra Hansen-Morath
Sascha Wolfer



STATISTIK MIT R

Inferenzstatistik

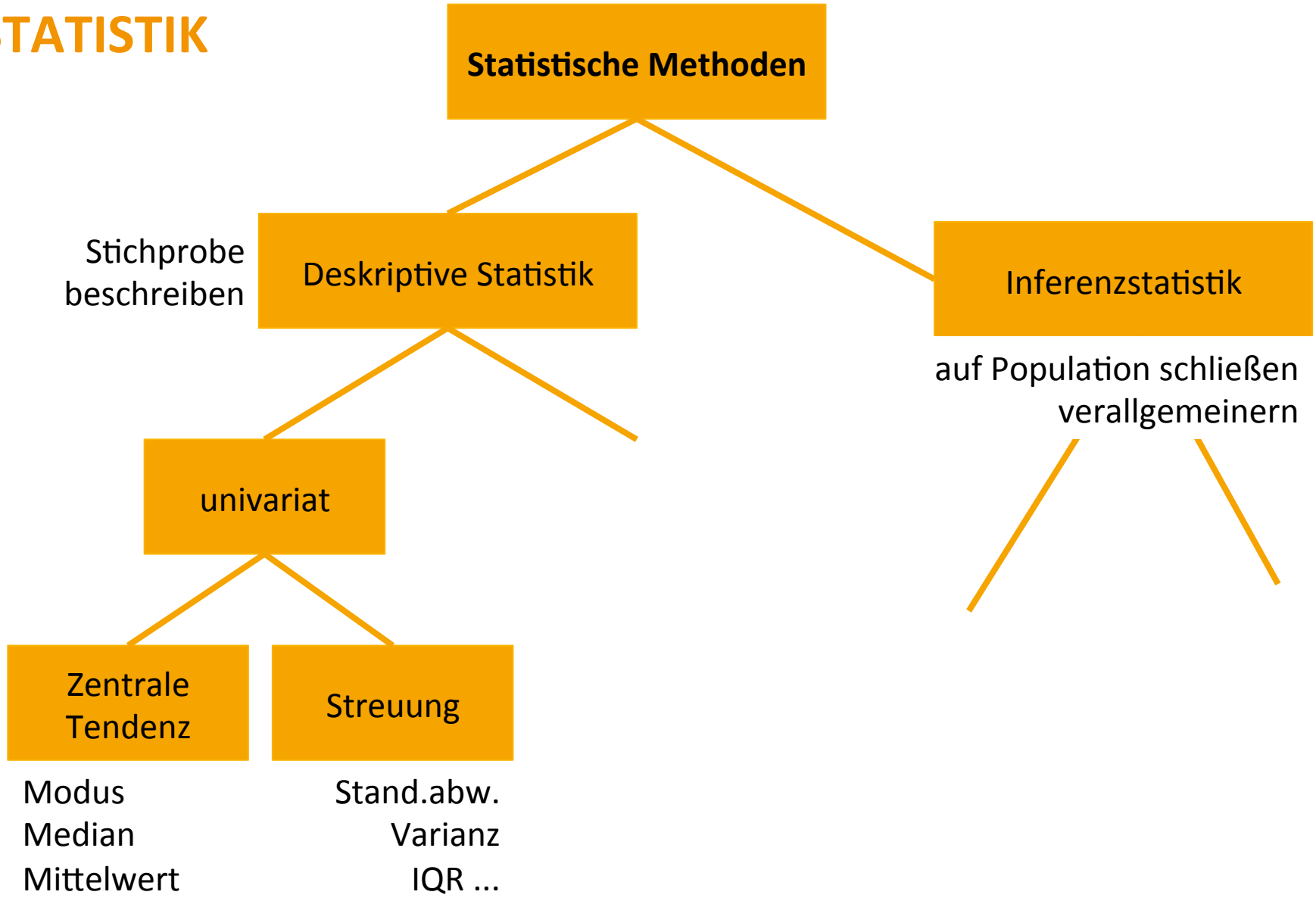
Mitglied der Leibniz-Gemeinschaft

BISHER

Deskriptive Statistik → Quantitative Aussagen über Stichproben:

- Häufigkeiten
- Maße der zentralen Tendenz:
  - Modus
  - Median
  - Mittelwert
- Dispersionsmaße:
  - Range
  - Varianz
  - Standardabweichung

STATISTIK


Statistische Methoden
Stichprobe beschreiben
Deskriptive Statistik
Inferenzstatistik
auf Population schließen verallgemeinern
univariat
Zentrale Tendenz
Streuung
Modus
Median
Mittelwert
Stand.abw.
Varianz
IQR ...

INFERENZSTATISTIK

Inferenzstatistik → Hypothesen sollen für alle Elemente (z.B. Menschen) einer Grundgesamtheit (Population) gelten

INFERENZSTATISTIK

Inferenzstatistik → Hypothesen sollen für alle Elemente (z.B. Menschen) einer Grundgesamtheit (Population) gelten

→ Die Untersuchungen, die zur Überprüfung dieser Hypothesen durchgeführt werden, beziehen sich jedoch nur auf eine Stichprobe.

INFERENZSTATISTIK

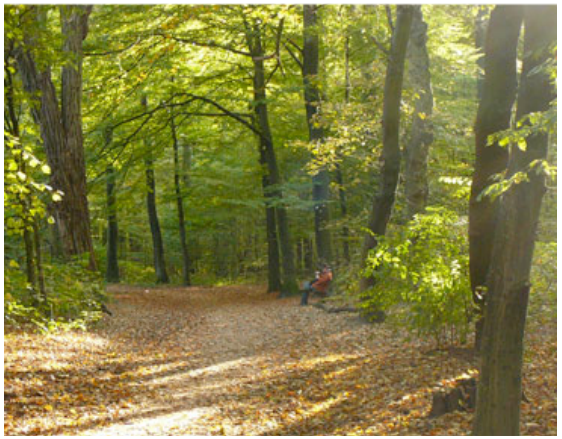
Inferenzstatistik → Hypothesen sollen für alle Elemente (z.B. Menschen) einer Grundgesamtheit (Population) gelten

→ Die Untersuchungen, die zur Überprüfung dieser Hypothesen durchgeführt werden, beziehen sich jedoch nur auf eine Stichprobe.

→ **Sind die Untersuchungsergebnisse von einer Stichprobe auf die Grundgesamtheit übertragbar bzw. verallgemeinerbar???**

INFERENZSTATISTIK

- Inferenzstatistische Verfahren geben in der Regel die Wahrscheinlichkeit dafür an, ob ein Zusammenhang oder Unterschied (ein Effekt) auf bloßen Zufall zurückzuführen ist.
- Konvention: Die Annahme einer zufälligen Zufallsschwankung in der Stichprobe wird verworfen wenn diese Wahrscheinlichkeit **unter 5%** beträgt **($p < .05$)**.



INFERENZSTATISTIK / STICHPROBENFEHLER

- Jedes Ergebnis, das aus einer Stichprobe gewonnen wird, ist bis zu einem gewissen Grad vom Zufall abhängig ...
- ... und dadurch mit einem „Fehler“ behaftet

INFERENZSTATISTIK / STICHPROBENFEHLER

- Jedes Ergebnis, das aus einer Stichprobe gewonnen wird, ist bis zu einem gewissen Grad vom Zufall abhängig ...
- ... und dadurch mit einem „Fehler“ behaftet.

<u>Die Güte einer Stichprobe hängt ab von:</u>

1) Größe der Stichprobe

→ je größer, desto besser

2) Streuung in der Stichprobe

→ je kleiner, desto besser

STICHPROBENFEHLER

- Um empirische Ergebnisse von der Stichprobe auf die Grundgesamtheit verallgemeinern zu können, muss man die **Größe des Stichprobenfehlers** kennen.
- Dieser ergibt sich aus der Differenz des Mittelwerts der Grundgesamtheit und dem Mittelwert der Stichprobe:

$$e_i = \mu - \bar{x}_i$$

STICHPROBENFEHLER

- Um empirische Ergebnisse von der Stichprobe auf die Grundgesamtheit verallgemeinern zu können, muss man die **Größe des Stichprobenfehlers** kennen.
- Dieser ergibt sich aus der Differenz des Mittelwerts der Grundgesamtheit und dem Mittelwert der Stichprobe:

$$e_i = \boxed{\mu} - \bar{x}_i$$

→ wir brauchen eine zuverlässige Schätzung für den Mittelwert der Grundgesamtheit

STICHPROBENFEHLER

Ein Gedankenspiel:

- Wir entnehmen einer Grundgesamtheit *k* Zufallsstichproben gleicher Größe *n.*
- Für jede dieser Stichproben wird der Mittelwert berechnet.
- All diese Stichprobenmittelwerte sind Schätzwerte für den unbekannten Parameter μ der Grundgesamtheit.
- Jeder einzelne Stichprobenmittelwert hat einen eigenen Fehler:

$$e_1 = \mu - \bar{x}_1$$

$$e_2 = \mu - \bar{x}_2$$

$$\ldots$$

$$e_k = \mu - \bar{x}_k$$

STICHPROBENFEHLER

- Die ermittelten Stichprobenwerte bilden wiederum selbst eine Verteilung: Die Stichprobenverteilung des arithmetischen Mittels.

→ der Mittelwert der Stichprobenverteilung ist der beste Schätzwert für den Mittelwert der Grundgesamtheit

→ „X quer quer“ – der Mittelwert der Mittelwerte

- Die Schätzung ist umso genauer,
  - je kleiner die Streuung der Stichprobenverteilungen
  - je mehr ausreichend große Stichproben

Es gilt: $\bar{\bar{x}} = \mu \ \ für \ \ k \rightarrow \infty$

STANDARDFEHLER

- Der Standardfehler gibt an, wie nah ein so erhobener Mittelwert am wahren Mittelwert liegt
- Die Standardabweichung des Mittelwerts (= **Standardfehler**) ist die Standardabweichung in der Grundgesamtheit geteilt durch die Wurzel aus n

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

- Je größer die Variabilität eines Merkmals in der Grundgesamtheit, desto größer der Standardfehler des Mittelwerts
- Je größer die Stichproben, desto kleiner dieser Fehler.

KURZE ZUSAMMENFASSUNG

- Eine Grundgesamtheit hat einen bestimmten Mittelwert, den wir nicht kennen.
- Entnimmt man dieser Grundgesamtheit immer wieder Stichproben, würde man jedes Mal einen leicht anderen Mittelwert erhalten.
- Die Mittelwerte aus den verschiedenen Stichproben verteilen sich normal. Der Mittelwert der Mittelwerte („x quer quer“) entspricht dem tatsächlichen Mittelwert der Grundgesamtheit.
- Der Standardfehler gibt an, wie nah ein so erhobener Mittelwert am wahren Mittelwert liegt.

BEISPIEL: MESSUNG VON KÖRPERGRÖßEN

- Feststellen der mittleren Größe aller in Freiburg lebenden Menschen (Grundgesamtheit = 229.144 Menschen, Stand Dez. 2011)
  - → Alle diese Menschen versammeln sich auf dem neuen Messplatz für unsere Messung.
  - → Wir nehmen zufällig 10 Leute aus der Masse heraus und messen sie. Wir notieren uns den Mittelwert der Messung und schicken die Leute zurück in die Masse.
  - → Das ganze machen wir 50 mal; am Ende haben wir 50 Mittelwerte.
  - → Aus diesen 50 Mittelwerten wird ein Mittelwert gebildet: „x quer quer“.
- Das ist unsere beste Schätzung für die mittlere Größe der Menschen in Freiburg = μ.
- Je mehr bzw. je größere Stichproben wir ziehen, desto besser wird diese Schätzung

INFERENZSTATISTIK

- Grundidee: Wir schätzen Populationsparameter (Mittelwert und Standardabweichung) durch die Kennwerte von Stichproben.

	Stichprobe	Population
Arithmetisches Mittel	$\overline{x}$	μ
Varianz	s^2	σ^2

HYPOTHESEN TESTEN

- In der Inferenzstatistik werden Hypothesen getestet

→ Es wird getestet, ob wir uns sicher genug sein können, die Nullhypothese abzulehnen und damit die Alternativhypothese anzunehmen

Nullhypothese (H0)	Alternativhypothese (H1)
Es gibt keinen Unterschied. Es gibt keinen Zusammenhang. Das Merkmal ist gleich verteilt.	Es gibt einen Unterschied. Es gibt einen Zusammenhang. Das Merkmal ist nicht gleichverteilt

ALTERNATIVHYPOTHESE

Die H1 kann gerichtet und/oder spezifisch sein:

- **ungerichtet, unspezifisch**:

  Äpfel haben ein anderes Gewicht als Birnen.
- **gerichtet, unspezifisch**:

  Äpfel sind schwerer als Birnen.
- **gerichtet, spezifisch**:

  Äpfel sind im Schnitt um 20 Gramm schwerer als Birnen.
- **ungerichtet, spezifisch**:

  meist nicht sinnvoll

HYPOTHESEN TESTEN

- Ziel ist es, die Nullhypothese mit großer Sicherheit abzulehnen und die Alternativhypothese anzunehmen.
- Hypothesen werden nicht bewiesen, sondern mit großer Sicherheit angenommen/abgelehnt.
- Die Irrtumswahrscheinlichkeit *p* gibt an, wie unsicher wir uns sind.

FEHLERTYPEN

In der Inferenzstatistik spricht man von verschiedenen Fehlertypen

In der GG gilt:

Wir postulieren:

	H0	**H1**
H0	korrekt, langweilig	β-Fehler
H1	α-Fehler	korrekt

***p* = Wahrscheinlichkeit für den Alphafehler**

DAS ALPHANIVEAU

- Unterschiede in der Statistik werden als signifikant bezeichnet, wenn die Wahrscheinlichkeit, dass sie durch Zufall zustande gekommen sind, nicht über einer gewissen Schwelle liegt.
- Irrtumswahrscheinlichkeit: Mit der Wahrscheinlichkeit *p* wird die H0 abgelehnt, obwohl sie richtig ist.
- Wenn $p < \alpha$, wird die H0 verworfen (= die H1 angenommen).

- Konventionen:
  - 5 % ($p < .05$) → signifikant
  - 1 % ($p < .01$) → hoch signifikant
  - 0,1% ($p < .001$) → höchst signifikant

KONFIDENZINTERVALL

- Konfidenzintervalle geben einen Bereich an, bei dem wir uns zu x-prozentiger Wahrscheinlichkeit sicher sind, dass ein bestimmter Wert (z.B. Mittelwert) darin liegt.
- Je sicherer wir uns sein wollen, desto größer muss das Konfidenzintervall sein.
- Je mehr Rauschen in unserer Stichprobe, desto größer muss das Konfidenzintervall sein.
- Das Konfidenzintervall für den Mittelwert errechnet sich aus:

$$\mathrm{CI}_{\mathrm{lower/upper}} = \bar{x} \pm u_\alpha \cdot s_{\bar{x}}$$

KONFIDENZINTERVALL

Standardfehler des Mittelwerts

Mittelwert der Stichprobe

$$\mathrm{CI}_{\mathrm{lower/upper}} = \bar{x} \pm u_\alpha \cdot s_{\bar{x}}$$

Ergibt zwei Werte: untere und obere Grenze, immer gleich weit entfernt vom Mittelwert

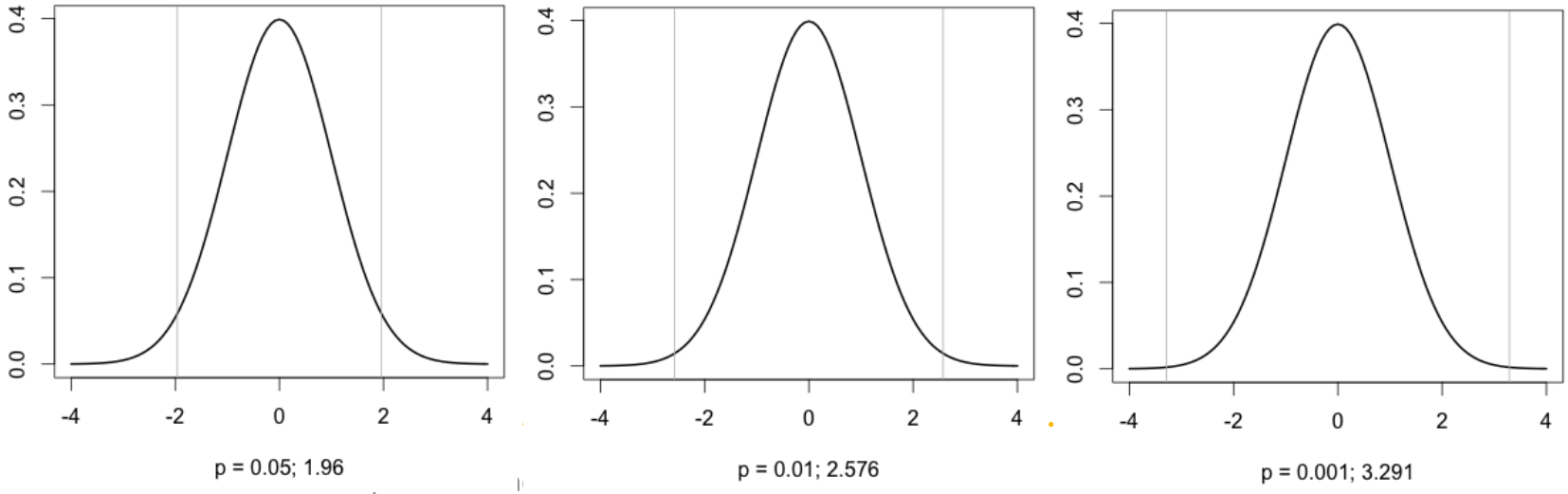
„u von alpha“: Mit Irrtumswahrscheinlichkeit assoziierter Wert

„U VON ALPHA“

Mit jedem α-Niveau ist ein bestimmter u-Wert verbunden.

- α = 95% → u = 1,96
- α = 99% → u = 2,58
- α = 99,9% → u = 3,29

Das heißt: Je sicherer wir uns sein wollen, desto größer wird das Intervall



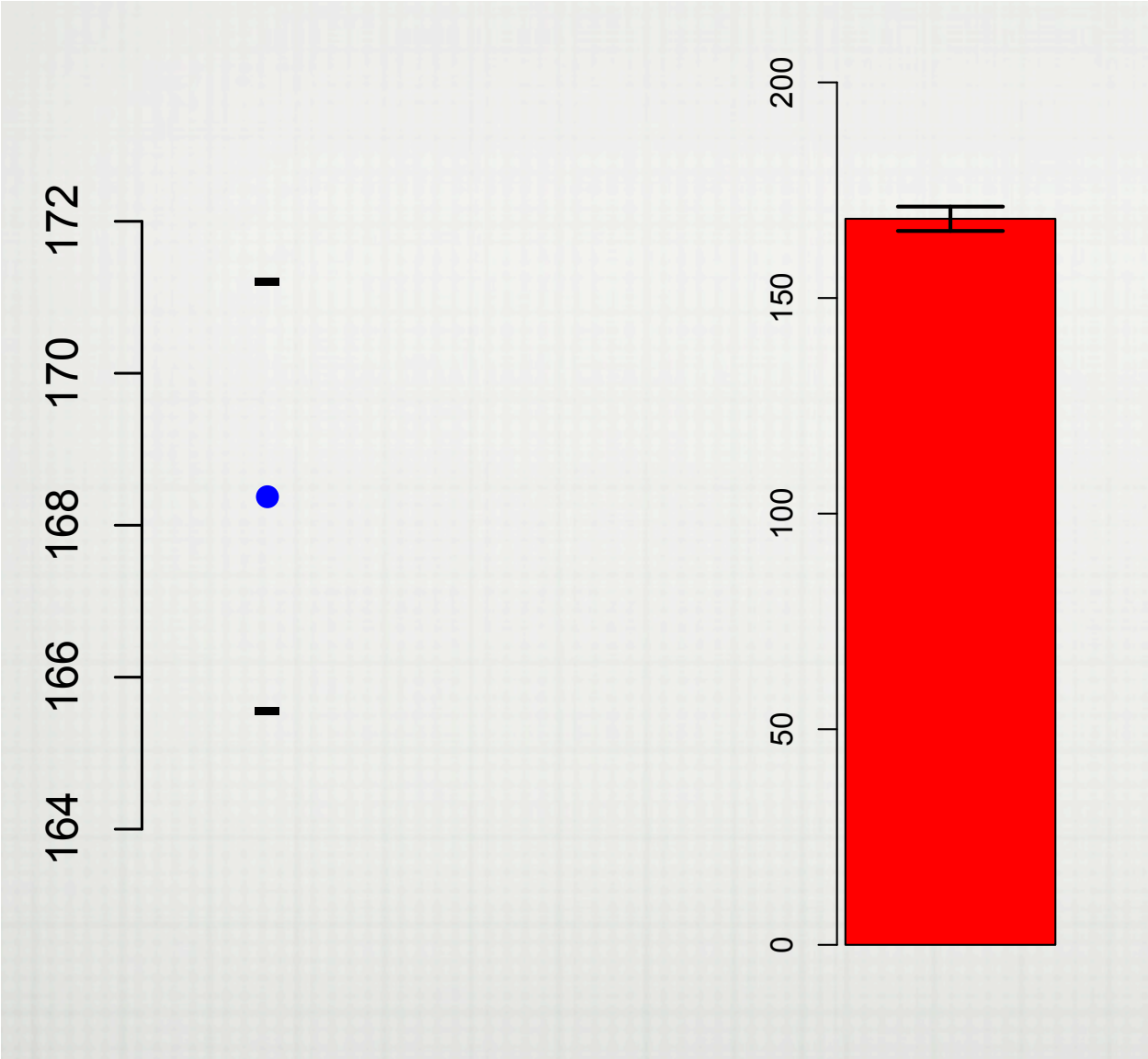
BEISPIEL

- Wir schätzen aufgrund von Körpergrößen die mittlere Körpergröße der Population:
  - Mittelwert: 168,375
  - s = 7,06
  - n = 24
  - α = 95%, u = 1,96

- Standardfehler = 7,06 / Wurzel (24) = 1,44

- CI = 168,375 +/- 1,96 * 1,44 = 168,375 +/- 2,82

- Aufgrund unserer Stichprobe nehmen wir mit 95%iger Sicherheit an, dass sich der Mittelwert der Population zwischen 165,6 und 171,2 cm befindet.


164
166
168
170
172
0
50
100
150
200

BEGRIFFE

- Inferenzstatistik
- Population / Grundgesamtheit
- Effekt / Signal
- Rauschen
- Stichprobe
- Stichprobenfehler
- Standardfehler
- Nullhypothese H0
- Alternativhypothese H1
- gerichtete / spezifische H.
- Irrtumswahrscheinlichkeit
- α-Fehler, β-Fehler
- Signifikanz
- Konfidenzintervall