9. Vorlesung Statistik I
Parameterschätzung: Unterschiede zwischen Populationen

Unterschiede zwischen Populationen

Unterschiede zwischen Populationen

- Die bisher besprochenen inferenzstatistischen Verfahren erlauben uns, aus einer einfachen Zufallsstichprobe auf eine deskriptivstatistische Maßzahl in **einer** Population zu schließen:

    - Falls wir uns für den Mittelwert einer stetigen Variable in dieser Population interessieren, ziehen wir eine einfache Zufallsstichprobe und berechnen ein Konfidenzintervall für μ.

    - Falls wir uns für die relative Häufigkeit einer Messwertausprägung einer diskreten Variable in dieser Population interessieren, ziehen wir eine einfache Zufallsstichprobe und berechnen ein Konfidenzintervall für π.

- Sehr häufig interessiert man sich in der Psychologie jedoch auch für **Unterschiede zwischen mehreren Populationen**.

Beispiele

- Beispiel 1: Uns interessiert der Unterschied zwischen Psychologie- und BWL-Student*innen in der durchschnittlichen (stetigen) Leistungsmotivation:
  - Population 1: Psychologie-Student*innen
  - Population 2: BWL-Student*innen
  - Interessierender Unterschied: $\bar{x}_{Psychologie} - \bar{x}_{BWL}$

- Beispiel 2: Uns interessiert der Unterschied in der durchschnittlichen (stetigen) Depressionsschwere zwischen depressiven Personen, die eine Psychotherapie erhalten haben, und solchen, die keine Therapie erhalten haben:
  - Population 1: Personen, die eine Therapie erhalten haben
  - Population 2: Personen, die keine Therapie erhalten haben
  - Interessierender Unterschied: $\bar{x}^*_{Therapie} - \bar{x}^*_{keineTherapie}$

Bemerkung:
Wofür der Stern * steht, besprechen wir später

- Beispiel 3: Uns interessiert der Unterschied in der durchschnittlichen (stetigen) Matheleistung von Personen vor und nach einer Nachhilfestunde:
  - Population 1: Personen vor der Nachhilfe
  - Population 2: Die gleichen Personen nach der Nachhilfe
  - Interessierender Unterschied: $\bar{x}_{vorher} - \bar{x}_{nachher}$

Arten der Stichprobenziehung

- Die drei Beispiele unterscheiden sich in der Art der Stichprobenziehung:
  - Beispiel 1: Wir ziehen eine einfache Zufallsstichprobe aus der Population der Psycholog*innen und eine einfache Zufallsstichprobe aus der Population der BWLer*innen.
  - Beispiel 2: Wir ziehen zunächst eine einfache Zufallsstichprobe aus der Gesamtpopulation der depressiven Personen. Ein zufällig ausgewählter Teil der Personen in der Stichprobe erhält eine Psychotherapie, der Rest nicht.
  - Beispiel 3: Wir ziehen eine einfache Zufallsstichprobe aus der Gesamtpopulation der Schüler*innen. Alle Personen in dieser Stichprobe erhalten Nachhilfe.
- Wir werden uns nun anschauen, wie wir diese drei Arten der Stichprobenziehung im Rahmen der Wahrscheinlichkeitstheorie formalisieren können.

Zufallsvariablen in Beispiel 1 I

- In Beispiel 1 haben wir zwei einfache Zufallsstichproben und daher die gleiche Situation wie im Fall von einer einzigen Population, nur eben zweimal:

- Wir haben eine Stichprobe von Psycholog*innen mit Stichprobenumfang n_1 und Messwerte x_{11}, x_{12}, …, x_{1n_1}, die als Realisationen der Zufallsvariablen X_{11}, X_{12}, …, X_{1n_1} aufgefasst werden können. x_{11} ist die Leistungsmotivation der ersten Person in der Stichprobe der Psycholog*innen, x_{12} ist die Leistungsmotivation der zweiten Person in der Stichprobe der Psycholog*innen, …, x_{1n_1} ist die Leistungsmotivation der n_1 - ten Person in der Stichprobe der Psycholog*innen.

- Wir haben eine Stichprobe von BWLer*innen mit Stichprobenumfang n_2 und Messwerte x_{21}, x_{22}, …, x_{2n_2}, die als Realisationen der Zufallsvariablen X_{21}, X_{22}, …, X_{2n_2} aufgefasst werden können. x_{21} ist die Leistungsmotivation der ersten Person in der Stichprobe der BWLer*innen, x_{22} ist die Leistungsmotivation der zweiten Person in der Stichprobe der BWLer*innen, …, x_{2n_2} ist die Leistungsmotivation der n_2 - ten Person in der Stichprobe der BWLer*innen.

Zufallsvariablen in Beispiel 1 II

- Falls das Histogramm der Leistungsmotivation in beiden Populationen durch die Wahrscheinlichkeitsdichtefunktion einer Normalverteilung approximiert werden kann, gilt für die Zufallsvariablen:
  - Stichprobe der Psycholog*innen : $X_{1i} \overset{\text{iid}}{\sim} N(\mu_1, \sigma_1^2)$
  - Stichprobe der BWLer*innen : $X_{2i} \overset{\text{iid}}{\sim} N(\mu_2, \sigma_2^2)$
  - Die Zufallsvariablen aus der Stichprobe der Psycholog*innen sind unabhängig von den Zufallsvariablen aus der Stichprobe der BWLer*innen.
- μ_1 entspricht dem Mittelwert $\bar{x}_{Psychologie}$ der Leistungsmotivation in der **Population** der Psycholog*innen und σ_1^2 entspricht der empirischen Varianz $s^2_{emp\ Psychologie}$ der Leistungsmotivation in der **Population** der Psycholog*innen.
- μ_2 entspricht dem Mittelwert $\bar{x}_{BWL}$ der Leistungsmotivation in der **Population** der BWLer*innen und σ_2^2 entspricht der empirischen Varianz $s^2_{emp\ BWL}$ der Leistungsmotivation in der **Population** der BWLer*innen.

Zufallsvariablen in Beispiel 1 III

- Wir nehmen zusätzlich an, dass die empirischen Varianzen in der **Population** der BWLer*innen und Psycholog*innen gleich sind, also $s^2_{emp\ Psychologie} = s^2_{emp\ BWL}$ gilt.

- Hieraus folgt, dass auch die Varianzen der Zufallsvariablen gleich sind, also $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

- Die Plausibilität dieser Annahme werden wir – genau wie die der Normalverteilungsannahme – später überprüfen.

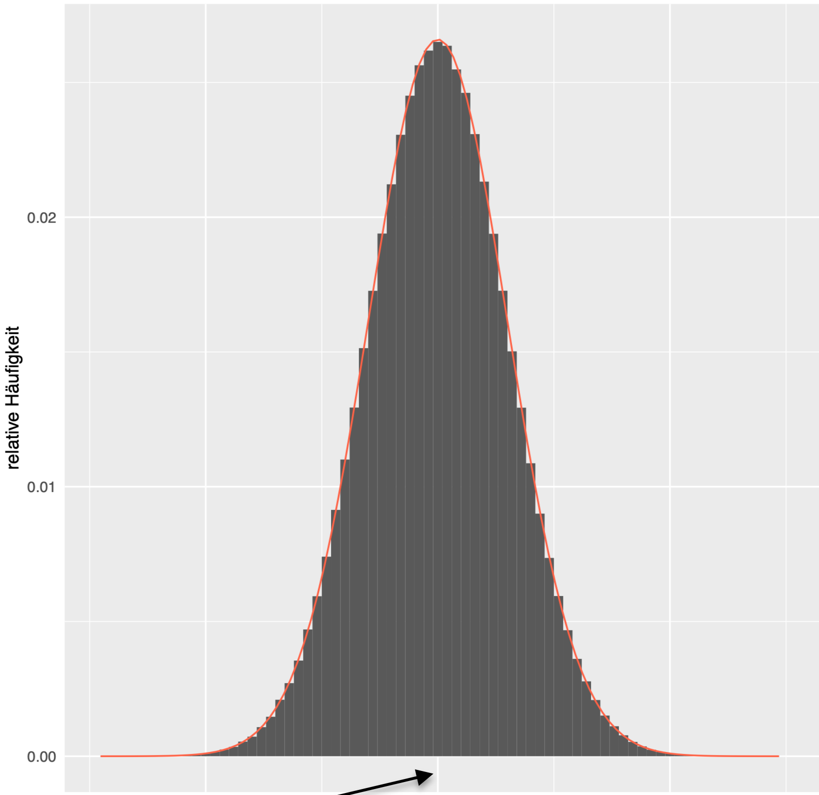
Zufallsvariablen in Beispiel 1 IV

- Zusammengefast ist die Situation also wie folgt:
  - Stichprobe der Psycholog*innen : $X_{1i} \overset{\text{iid}}{\sim} N(\mu_1, \sigma^2)$
  - Stichprobe der BWLer*innen : $X_{2i} \overset{\text{iid}}{\sim} N(\mu_2, \sigma^2)$
  - Die Zufallsvariablen aus der Stichprobe der Psycholog*innen sind unabhängig von den Zufallsvariablen aus der Stichprobe der BWLer*innen.
- μ_1 entspricht dem Mittelwert $\bar{x}_{Psychologie}$ der Leistungsmotivation in der Population der Psycholog*innen und σ^2 entspricht der empirischen Varianz $s^2_{emp\ Psychologie}$ der Leistungsmotivation in der Population der Psycholog*innen.
- μ_2 entspricht dem Mittelwert $\bar{x}_{BWL}$ der Leistungsmotivation in der Population der BWLer*innen und σ^2 entspricht der empirischen Varianz $s^2_{emp\ BWL}$ der Leistungsmotivation in der Population der BWLer*innen.

Graphische Darstellung der Populationen

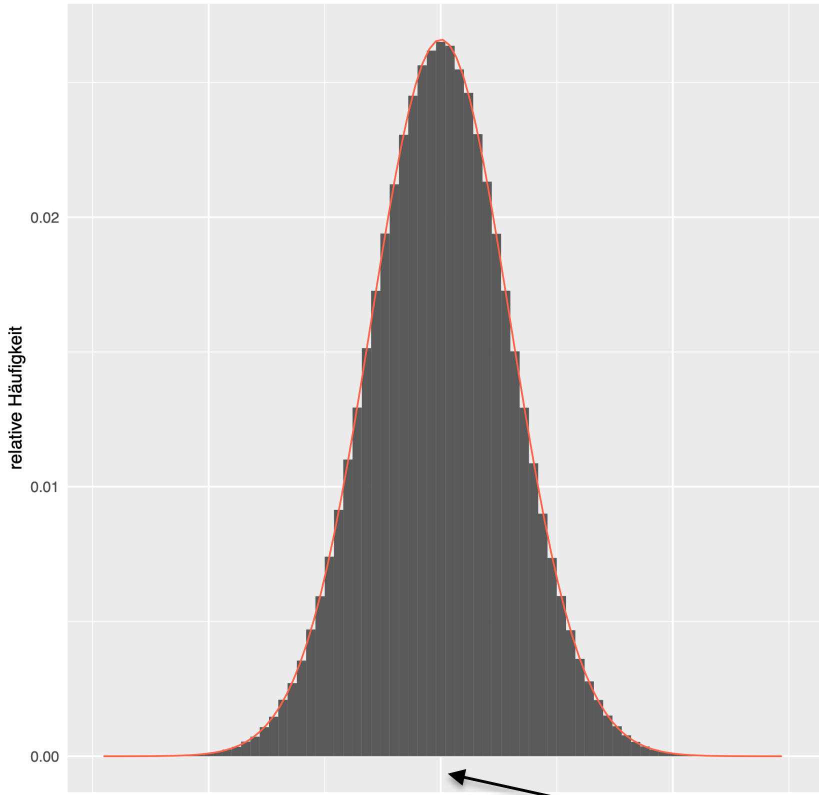
- Graphische Darstellung:

Population der Psycholog*innen



$\bar{x}_{Psychologie}$ Leistungsmotivation

Population der BWLer*innen



Leistungsmotivation $\bar{x}_{BWL}$

Parameterdifferenz

- Der Unterschied $\bar{x}_{Psychologie} - \bar{x}_{BWL}$ zwischen den Populationen der Psycholog*innen und BWLer*innen in der durchschnittlichen Leistungsmotivation entspricht somit der **Parameterdifferenz** $\mu_1 - \mu_2$.

- Falls wir Aussagen über $\bar{x}_{Psychologie} - \bar{x}_{BWL}$ treffen wollen, müssen wir im Rahmen inferenzstatistischer Verfahren zu Aussagen über die Parameterdifferenz $\mu_1 - \mu_2$ gelangen.

- Wir werden in der heutigen Vorlesung besprechen, wie wir Schätzfunktionen und Konfidenzintervalle für $\mu_1 - \mu_2$ konstruieren können.

Zufallsvariablen in Beispiel 2 I

- In Beispiel 2 haben wir eine einfache Zufallsstichprobe, die wir nach der Ziehung zufällig in eine Experimentalgruppe und eine Kontrollgruppe aufteilen:

- Wir haben eine Stichprobe von Personen, die eine Psychotherapie erhalten haben, mit Stichprobenumfang n_1 und Messwerte x_{11}, x_{12}, …, x_{1n_1}, die als Realisationen der Zufallsvariablen X_{11}, X_{12}, …, X_{1n_1} aufgefasst werden können. x_{11} ist die Depressionsschwere der ersten Person in dieser Stichprobe, x_{12} ist die Depressionsschwere der zweiten Person in dieser Stichprobe, …, x_{1n_1} ist die Depressionsschwere der n_1 - ten Person in dieser Stichprobe.

- Wir haben eine Stichprobe von Personen, die keine Psychotherapie erhalten haben, mit Stichprobenumfang n_2 und Messwerte x_{21}, x_{22}, …, x_{2n_2}, die als Realisationen der Zufallsvariablen X_{21}, X_{22}, …, X_{2n_2} aufgefasst werden können. x_{21} ist die Depressionsschwere der ersten Person in dieser Stichprobe, x_{22} ist die Depressionsschwere der zweiten Person in dieser Stichprobe, …, x_{2n_2} ist die Depressionsschwere der n_2 - ten Person in dieser Stichprobe.

Zufallsvariablen in Beispiel 2 II

- Falls das Histogramm der Depressionsschwere in der Population mit Therapie und in der Population ohne Therapie jeweils durch die Wahrscheinlichkeitsdichtefunktion einer Normalverteilung approximiert werden kann, gilt für die Zufallsvariablen:
  - Stichprobe mit Psychotherapie: $X_{1i} \overset{\text{iid}}{\sim} N(\mu_1, \sigma_1^2)$
  - Stichprobe ohne Psychotherapie: $X_{2i} \overset{\text{iid}}{\sim} N(\mu_2, \sigma_2^2)$
  - Die Zufallsvariablen aus der Stichprobe der Personen mit Psychotherapie sind unabhängig von den Zufallsvariablen aus der Stichprobe der Personen ohne Psychotherapie.

- μ_1 entspricht dem Mittelwert $\bar{x}^*_{Therapie}$ der Depressionsschwere in der Population der Personen mit Therapie und σ_1^2 entspricht der empirischen Varianz $s^2_{emp\ Therapie}$ der Depressionsschwere in dieser Population.

- μ_2 entspricht dem Mittelwert $\bar{x}^*_{keineTherapie}$ der Depressionsschwere in der Population der Personen ohne Therapie und σ_2^2 entspricht der empirischen Varianz $s^2_{emp\ keineTherapie}$ der Depressionsschwere in dieser Population.

Zufallsvariablen in Beispiel 2 III

- Wie in Beispiel 1 nehmen wir zusätzlich an, dass die empirischen Varianzen in beiden Populationen gleich sind, also $s^2_{emp\ Therapie} = s^2_{emp\ keineTherapie}$ gilt.

- Hieraus folgt, dass auch die Varianzen der Zufallsvariablen gleich sind, also $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

- Die Plausibilität dieser Annahme werden wir – genau wie die der Normalverteilungsannahme – später überprüfen.

Zufallsvariablen in Beispiel 2 IV

- Wir haben also – trotz leicht unterschiedlicher Stichprobenziehung – die gleiche Situation wie in Beispiel 1:
  - Stichprobe mit Therapie: $X_{1i} \overset{\text{iid}}{\sim} N(\mu_1, \sigma^2)$
  - Stichprobe ohne Therapie: $X_{2i} \overset{\text{iid}}{\sim} N(\mu_2, \sigma^2)$
  - Die Zufallsvariablen aus der Stichprobe mit Therapie sind unabhängig von den Zufallsvariablen aus der Stichprobe ohne Therapie.

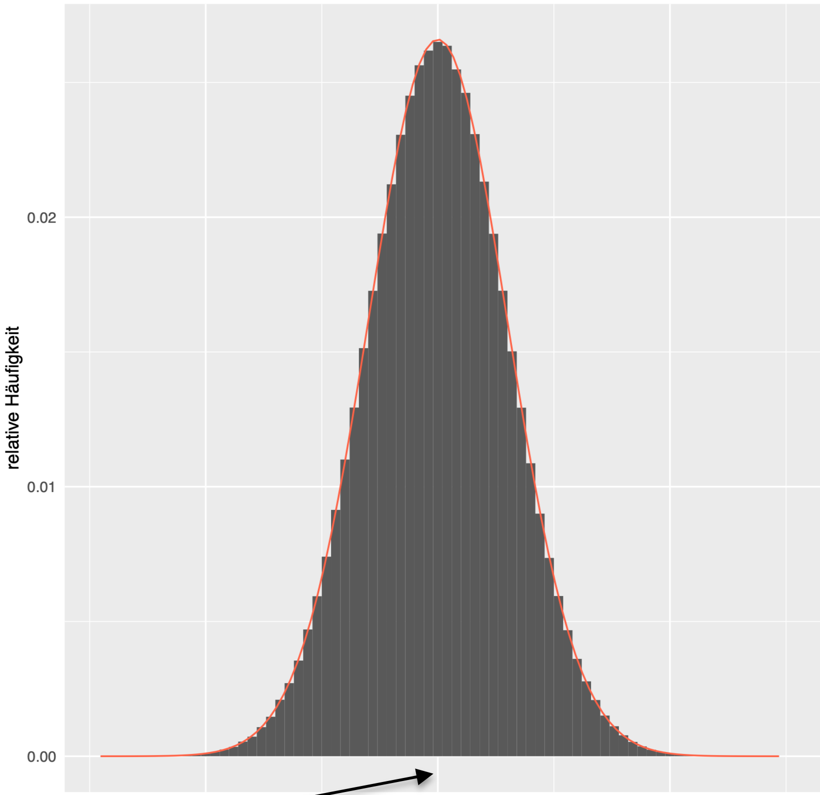
- μ_1 entspricht dem Mittelwert $\bar{x}^*_{Therapie}$ der Depressionsschwere in der Population der Personen mit Therapie und σ^2 entspricht der empirischen Varianz $s^2_{emp\ Therapie}$ der Depressionsschwere in dieser Population.

- μ_2 entspricht dem Mittelwert $\bar{x}^*_{keineTherapie}$ der Depressionsschwere in der Population der Personen ohne Therapie und σ^2 entspricht der empirischen Varianz $s^2_{emp\ keineTherapie}$ der Depressionsschwere in dieser Population.

Graphische Darstellung der Populationen

- Graphische Darstellung:

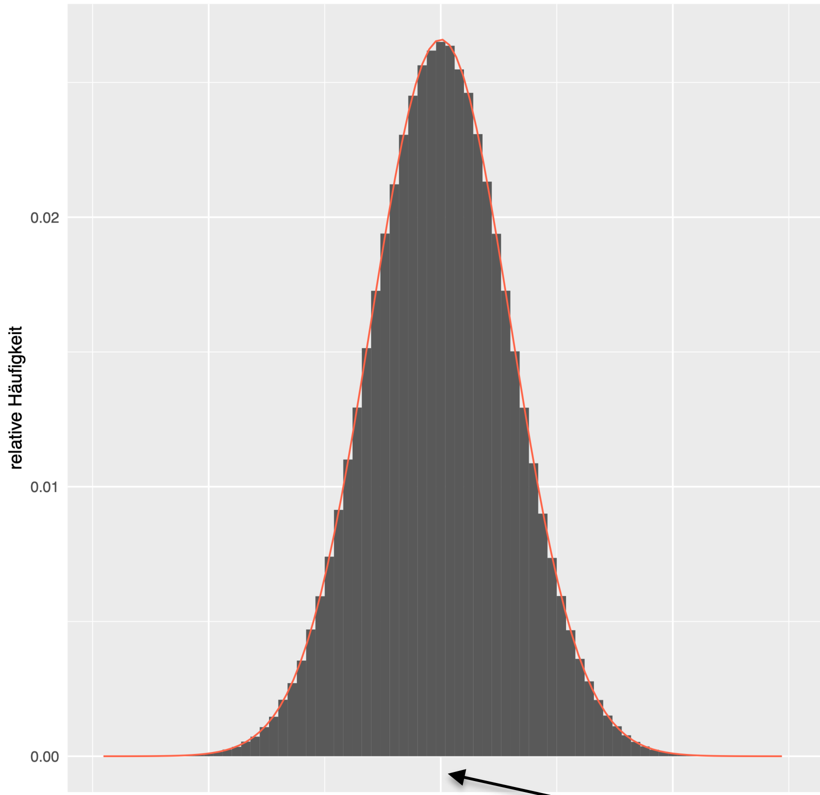
Population mit Therapie



$\bar{x}^*_{Therapie}$

Depressionsschwere

Population ohne Therapie



Depressionsschwere

$\bar{x}^*_{keineTherapie}$

Parameterdifferenz in Beispiel 2

- Der Unterschied $\bar{x}^*_{Therapie} - \bar{x}^*_{keineTherapie}$ zwischen den Populationen mit und ohne Therapie in der durchschnittlichen Depressionsschwere entspricht somit der Parameterdifferenz $\mu_1 - \mu_2$.

- Falls wir Aussagen über $\bar{x}^*_{Therapie} - \bar{x}^*_{keineTherapie}$ treffen wollen, müssen wir im Rahmen inferenzstatistischer Verfahren zu Aussagen über die **Parameterdifferenz** $\mu_1 - \mu_2$ gelangen.

- Wir werden in der heutigen Vorlesung besprechen, wie wir Schätzfunktionen und Konfidenzintervalle für $\mu_1 - \mu_2$ konstruieren können.

Zufallsvariablen in Beispiel 3 I

- In Beispiel 3 haben wir eine einfache Zufallsstichprobe von Schüler*innen mit Stichprobenumfang n, bei denen wir zweimal – einmal vor der Nachhilfe und einmal nach der Nachhilfe – die Matheleistung erfassen.

- Wir haben Messwerte x_{11}, x_{12}, …, x_{1n} aller Schüler*innen vor der Nachhilfe, die als Realisationen der Zufallsvariablen X_{11}, X_{12}, …, X_{1n} aufgefasst werden können. x_{11} ist die Matheleistung der ersten Schüler*in vor der Nachhilfe, x_{12} ist die Matheleistung der zweiten Schüler*in vor der Nachhilfe, …, x_{1n} ist die Matheleistung der n - ten Schüler*in vor der Nachhilfe.

- Wir haben Messwerte x_{21}, x_{22}, …, x_{2n} **der gleichen Schüler*innen** nach der Nachhilfe, die als Realisationen der Zufallsvariablen X_{21}, X_{22}, …, X_{2n} aufgefasst werden können. x_{21} ist die Matheleistung der ersten Schüler*in nach der Nachhilfe, x_{22} ist die Matheleistung der zweiten Schüler*in nach der Nachhilfe, …, x_{2n} ist die Matheleistung der n - ten Schüler*in nach der Nachhilfe.

Zufallsvariablen in Beispiel 3 II

- Falls das Histogramm der Matheleistung in beiden Populationen (Schüler*innen vor und nach der Nachhilfe) durch die Wahrscheinlichkeitsdichtefunktion einer Normalverteilung approximiert werden kann, gilt für die Zufallsvariablen:
  - Vor der Nachhilfe: $X_{1i} \overset{\text{iid}}{\sim} N(\mu_1, \sigma_1^2)$
  - Nach der Nachhilfe: $X_{2i} \overset{\text{iid}}{\sim} N(\mu_2, \sigma_2^2)$
  - Die Zufallsvariablen vor der Nachhilfe sind **nicht unabhängig** von den Zufallsvariablen nach der Nachhilfe, da es sich bei ihren Realisationen um **Messwerte der gleichen Personen** handelt.

- μ_1 entspricht dem Mittelwert $\bar{x}_{vorher}$ der Matheleistung in der Population vor der Nachhilfe und σ_1^2 entspricht der empirischen Varianz $s^2_{emp\ vorher}$ der Matheleistung in dieser Population.

- μ_2 entspricht dem Mittelwert $\bar{x}_{nachher}$ der Matheleistung in der Population nach der Nachhilfe und σ_2^2 entspricht der empirischen Varianz $s^2_{emp\ nachher}$ der Matheleistung in dieser Population.

Zufallsvariablen in Beispiel 3 III

- Obwohl die beiden Zufallsvariablen X_{1i} und X_{2i} für jede Schüler*in i abhängig sind, sind die Differenzen

$$X_{i\,Diff} = X_{1i} - X_{2i}$$

für jede Schüler*in i immer noch unabhängig, da wir die Schüler*innen im Rahmen einer einfachen Zufallsstichprobe gezogen haben.

Zufallsvariablen in Beispiel 3 IV

- Die Realisationen $x_{i\,Diff} = x_{1i} - x_{2i}$ der Zufallsvariablen $X_{i\,Diff} = X_{1i} - X_{2i}$ entsprechen für jede Schüler*in *i* der Differenz zwischen dem Messwert x_{1i} vor der Nachhilfe und dem Messwert x_{2i} nach der Nachhilfe.

- Beispieldaten:

$\boldsymbol{x_{1i}}$	$\boldsymbol{x_{2i}}$	$\boldsymbol{x_{i\,Diff} = x_{1i} - x_{2i}}$
5	1	4
4	5	-1
8	7	1
5	7	-2

Zufallsvariablen in Beispiel 3 V

- Für die Zufallsvariablen $X_{i\ Diff}$ gilt:

$$X_{i\ Diff} \overset{\text{iid}}{\sim} N\left(\mu_1 - \mu_2, \sigma^2_{Diff}\right)$$

  wobei σ^2_{Diff} der empirischen Varianz der **Differenz** der Matheleistung vor und nach der Nachhilfe in der Population entspricht. σ^2_{Diff} entspricht also **nicht** der empirischen Varianz der Matheleistung innerhalb der beiden Populationen.

- In diesem Fall hängen die inferenzstatistischen Verfahren nur von dem Parameter σ^2_{Diff} ab und wir müssen keine Annahme bezüglich der Gleichheit von σ^2_1 und σ^2_2 treffen.

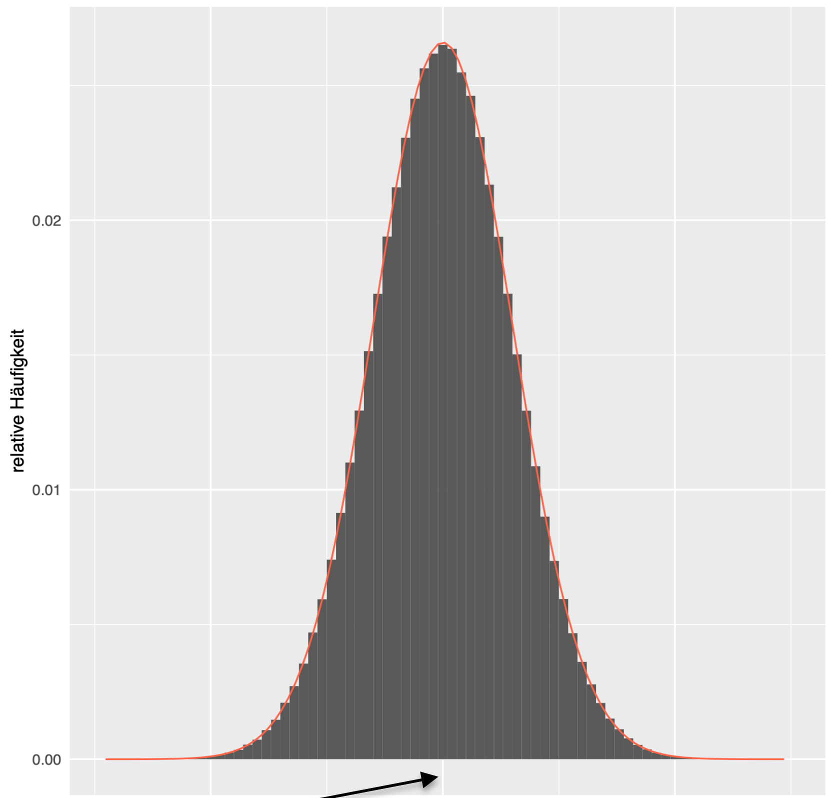
Zufallsvariablen in Beispiel 3 VI

- Zusammengefasst ist die Situation also wie folgt:
  - $X_{i\,Diff} \overset{\text{iid}}{\sim} N(\mu_1 - \mu_2, \sigma^2_{Diff})$
- μ_1 entspricht dem Mittelwert $\bar{x}_{vorher}$ der Matheleistung in der Population vor der Nachhilfe.
- μ_2 entspricht dem Mittelwert $\bar{x}_{nachher}$ der Matheleistung in der Population nach der Nachhilfe.
- σ^2_{Diff} entspricht der empirischen Varianz der Differenz der Matheleistung vor und nach der Nachhilfe in der Population.

Graphische Darstellung der Populationen

- Graphische Darstellung:

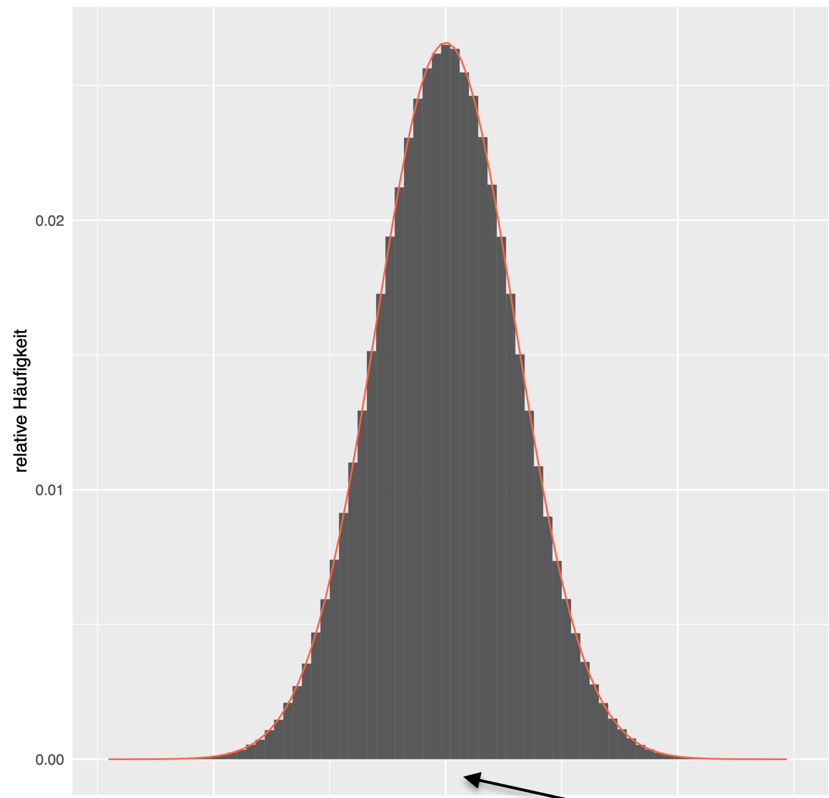
Population vorher



$\bar{x}_{vorher}$

Matheleistung

Population nachher



Matheleistung

$\bar{x}_{nachher}$

Parameterdifferenz in Beispiel 3

- Der Unterschied $\bar{x}_{vorher} - \bar{x}_{nachher}$ zwischen den Populationen der Schüler*innen vor und nach der Nachhilfe in der durchschnittlichen Matheleistung entspricht somit der **Parameterdifferenz** $\mu_1 - \mu_2$.

- Falls wir Aussagen über $\bar{x}_{vorher} - \bar{x}_{nachher}$ treffen wollen, müssen wir im Rahmen inferenzstatistischer Verfahren zu Aussagen über die Parameterdifferenz $\mu_1 - \mu_2$ gelangen.

- Wir werden in der heutigen Vorlesung besprechen, wie wir Schätzfunktionen und Konfidenzintervalle für $\mu_1 - \mu_2$ konstruieren können.

Abhängige und unabhängige Stichproben

- In der Situation von Beispiel 1 und 2 liegen aufgrund der unterschiedlichen Personen in den beiden Stichproben zwei **unabhängige Stichproben** vor.

- In der Situation von Beispiel 3 liegen aufgrund der gleichen Personen in den beiden Stichproben zwei **abhängige Stichproben** vor.

- In beiden Situationen interessieren wir uns für die Parameterdifferenz $\mu_1 - \mu_2$.

- Wir müssen jedoch je nach Situation aufgrund der unterschiedlichen Ausgangslage auf unterschiedliche inferenzstatistische Verfahren zurückgreifen.

Bemerkung

- Abhängige Stichproben entstehen in den meisten Fällen - wie in Beispiel 3 - durch wiederholte Messung bei den gleichen Personen.

- Es können aber auch Stichproben mit unterschiedlichen Personen abhängig sein, falls zum Beispiel Paare von Personen (z.B. Zwillingspaare) zufällig gezogen werden und jeweils eine Person des Personenpaars der Stichprobe 1 und die andere Person der Stichprobe 2 zugeordnet wird.

- Auch in diesem Fall würde man inferenzstatistische Verfahren für abhängige Stichproben verwenden, weil zwei AV-Werte die zum gleichen Paar gehören „ähnlicher" sein sollten als zwei AV-Werte von unterschiedlichen Paaren (genau wie im Fall mit Messwiederholung).

Zwischengliederung

- Bislang:
  - Unterschiede zwischen Populationen

- Jetzt:
  - Punktschätzung von Parameterdifferenzen

Punktschätzung von Parameterdifferenzen

Unabhängige Stichproben

Unabhängige Stichproben I

- Ausgangssituation:
  - $X_{11}, X_{12}, \ldots, X_{1n_1}$ mit $X_{1i} \overset{\text{iid}}{\sim} N(\mu_1, \sigma^2)$
  - $X_{21}, X_{22}, \ldots, X_{2n_2}$ mit $X_{2i} \overset{\text{iid}}{\sim} N(\mu_2, \sigma^2)$
  - alle diese Zufallsvariablen sind unabhängig.
- Ziel: Erwartungstreue, effiziente und konsistente Schätzfunktion für die Parameterdifferenz $\mu_1 - \mu_2$

Unabhängige Stichproben II

- Wir wissen, dass wir μ_1 und μ_2 jeweils durch die Mittelwerte in den beiden Stichproben schätzen können:

$$\hat{\mu}_1 = \bar{X}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i}$$

$$\hat{\mu}_2 = \bar{X}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i}$$

- Idee: Wir verwenden die Differenz

$$\bar{X}_{Diff} = \bar{X}_1 - \bar{X}_2$$

dieser beiden Schätzfunktionen als Schätzfunktion für $\mu_1 - \mu_2$.

- Der Schätzwert für $\mu_1 - \mu_2$ wäre dann

$$\bar{x}_{Diff} = \bar{x}_1 - \bar{x}_2$$

also einfach die Differenz der Mittelwerte der beiden Stichproben.

Unabhängige Stichproben III

- Erwartungswert von $\bar{X}_{Diff}$:

$$E\left(\bar{X}_{Diff}\right) = E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2$$

- Varianz von $\bar{X}_{Diff}$:

$$Var\left(\bar{X}_{Diff}\right) = Var(\bar{X}_1 - \bar{X}_2) = Var(\bar{X}_1) + Var(\bar{X}_2) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}$$

- Standardfehler von $\bar{X}_{Diff}$:

$$SE\left(\bar{X}_{Diff}\right) = SD\left(\bar{X}_{Diff}\right) = \sqrt{Var\left(\bar{X}_{Diff}\right)} = \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}$$

Unabhängige Stichproben IV

- Gütekriterien:

  ✓ $\bar{X}_{Diff}$ ist erwartungstreu

  ✓ $\bar{X}_{Diff}$ ist effizient (der Beweis ist mal wieder sehr schwierig)

  ✓ $\bar{X}_{Diff}$ ist konsistent für $n_1 \to \infty$ und $n_2 \to \infty$

Unabhängige Stichproben V

- Obwohl wir uns primär für $\mu_1 - \mu_2$ interessieren, werden wir im Folgenden auch eine Schätzfunktion für σ^2 benötigen.

- Da σ^2 die Varianz in beiden Stichproben ist, kämen als Schätzfunktion sowohl

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2$$

aus der ersten Stichprobe als auch

$$S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2$$

aus der zweiten Stichprobe in Betracht.

Unabhängige Stichproben VI

- Eine Schätzfunktion mit noch geringerem Standardfehler erhält man jedoch, indem man einen nach den Stichprobengrößen n_1 und n_2 gewichteten Mittelwert aus diesen beiden Schätzfunktionen bildet:

$$S_{pool}^2 = \frac{(n_1 - 1) \cdot S_1^2 + (n_2 - 1) \cdot S_2^2}{n_1 + n_2 - 2}$$

- Diese Schätzfunktion nennt man auch **gepoolte Varianz**.

- Als Schätzwert für σ^2 ergibt sich damit:

$$s_{pool}^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}$$

✓ Man kann zeigen, dass die Schätzfunktion S_{pool}^2 erwartungstreu, effizient und konsistent für σ^2 ist.

Unabhängige Stichproben VII

- Bemerkung: Für gleiche Stichprobengrößen $n_1 = n_2$ ergibt sich einfach der ungewichtete Mittelwert von S_1^2 und S_2^2 :

$$S_{pool}^2 = \frac{S_1^2 + S_2^2}{2}$$

Beispiel I

- Beispiel 1: Uns interessiert der Unterschied zwischen Psycholog*innen und BWLer*innen in der durchschnittlichen Leistungsmotivation.

- Wir ziehen zwei unabhängige einfache Zufallsstichproben:
  - Stichprobe 1 aus der Population der Psycholog*innen mit Umfang $n_1 = 101$.
  - Stichprobe 2 aus der Population der BWLer*innen $n_2 = 51$.

- In diesem Fall entspricht die uns interessierende Differenz $\bar{x}_{Psychologie} - \bar{x}_{BWL}$ der Parameterdifferenz $\mu_1 - \mu_2$.

Beispiel II

- In der Stichprobe der Psycholog*innen erhalten wir eine durchschnittliche Leistungsmotivation von $\bar{x}_1 = 165$ und einen Schätzwert der Varianz von $s_1^2 = 81$.

- In der Stichprobe der BWLer*innen erhalten wir eine durchschnittliche Leistungsmotivation von $\bar{x}_2 = 170$ und einen Schätzwert der Varianz von $s_2^2 = 100$.

- Als Schätzwert für $\mu_1 - \mu_2$, also für den Unterschied zwischen Psycholog*innen und BWLer*innen in der durchschnittlichen Leistungsmotivation, ergibt sich damit

$$\bar{x}_{Diff} = \bar{x}_1 - \bar{x}_2 = 165 - 170 = -5$$

- Als Schätzwert für σ^2 ergibt sich:

$$s_{pool}^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2} = \frac{(101 - 1) \cdot 81 + (51 - 1) \cdot 100}{101 + 51 - 2} = 87.33$$

Abhängige Stichproben

Abhängige Stichproben I

- Ausgangssituation:
  - $X_{i\,Diff} = X_{i1} - X_{i2}$
  - $X_{i\,Diff} \overset{\text{iid}}{\sim} N\left(\mu_1 - \mu_2, \sigma^2_{Diff}\right)$

- Ziel: Erwartungstreue, effiziente und konsistente Schätzfunktion für die Parameterdifferenz $\mu_1 - \mu_2$

Abhängige Stichproben II

- Auch für abhängige Stichproben können wir als Schätzfunktion für $\mu_1 - \mu_2$ wieder

$$\bar{X}_{Diff} = \bar{X}_1 - \bar{X}_2$$

verwenden.

- Der Schätzwert für $\mu_1 - \mu_2$ ist damit auch hier einfach die Differenz

$$\bar{x}_{Diff} = \bar{x}_1 - \bar{x}_2$$

der Mittelwerte $\bar{x}_1$ und $\bar{x}_2$ in den beiden Stichproben.

Abhängige Stichproben III

- Erwartungswert von $\bar{X}_{Diff}$:

$$E(\bar{X}_{Diff}) = E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2$$

- Varianz von $\bar{X}_{Diff}$:

$$Var(\bar{X}_{Diff}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_{i\,Diff}\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_{i\,Diff}\right) = \frac{n}{n^2} Var(X_{i\,Diff}) = \frac{\sigma_{Diff}^2}{n}$$

- Standardfehler von $\bar{X}_{Diff}$:

$$SE(\bar{X}_{Diff}) = SD(\bar{X}_{Diff}) = \sqrt{Var(\bar{X}_{Diff})} = \sqrt{\frac{\sigma_{Diff}^2}{n}}$$

Abhängige Stichproben IV

- Gütekriterien:

  ✓ $\bar{X}_{Diff}$ ist erwartungstreu

  ✓ $\bar{X}_{Diff}$ ist effizient

  ✓ $\bar{X}_{Diff}$ ist konsistent für $n \rightarrow \infty$

Abhängige Stichproben V

- Obwohl wir uns primär für $\mu_1 - \mu_2$ interessieren, werden wir im Folgenden auch eine Schätzfunktion für σ^2_{Diff} benötigen.

- Eine erwartungstreue, effiziente und konsistente Schätzfunktion für σ^2_{Diff} ist:

$$S^2_{Diff} = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_{i\,Diff} - \bar{X}_{Diff}\right)^2$$

- Die Realisation dieser Schätzfunktion, also der Schätzwert für σ^2_{Diff} ist somit:

$$s^2_{Diff} = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_{i\,Diff} - \bar{x}_{Diff}\right)^2$$

Abhängige Stichproben VI

- Beispieldaten:

x_{1i}	x_{2i}	$x_{i\,Diff} = x_{1i} - x_{2i}$
5	1	4
4	5	-1
8	7	1
5	7	-2

$\bar{x}_1 = 5.5$ $\bar{x}_2 = 5$ $\bar{x}_{Diff} = 0.5$

- Der Schätzwert für σ^2_{Diff} wäre in diesem Fall:

$$s^2_{Diff} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{i\,Diff} - \bar{x}_{Diff}\right)^2 = \frac{1}{4-1}\left[(4-0.5)^2 + (-1-0.5)^2 + (1-0.5)^2 + (-2-0.5)^2\right] = 7$$

Beispiel I

- Beispiel 3: Uns interessiert der Unterschied in der durchschnittlichen Matheleistung von Schüler*innen vor und nach einer Nachhilfestunde.

- Wir ziehen eine einfache Zufallsstichprobe von $n = 100$ Schüler*innen.
  - Wir erfassen die Matheleistung aller Schüler*innen in unserer Stichprobe mithilfe eines Mathetests (höhere Werte im Test entsprechen einer höheren Matheleistung).
  - Alle Schüler*innen in unserer Stichprobe erhalten Nachhilfeunterricht.
  - Wir erfassen die Matheleistung aller Schüler*innen in unserer Stichprobe nochmals mithilfe des Mathetests.

- In diesem Fall entspricht die uns interessierende Differenz $\bar{x}_{vorher} - \bar{x}_{nachher}$ der Parameterdifferenz $\mu_1 - \mu_2$.

Beispiel II

- Vor der Nachhilfe erhalten wir in unserer Stichprobe eine durchschnittliche Matheleistung von $\bar{x}_1 = 110$.

- Nach der Nachhilfe erhalten wir in unserer Stichprobe eine durchschnittliche Matheleistung von $\bar{x}_2 = 100$.

- Als Schätzwert für $\mu_1 - \mu_2$, also für den Unterschied in der durchschnittlichen Matheleistung vor und nach der Nachhilfe, ergibt sich damit

$$\bar{x}_{Diff} = \bar{x}_1 - \bar{x}_2 = 110 - 100 = 10$$

- Als Schätzwert für σ^2_{Diff} erhalten wir:

$$s^2_{Diff} = 99$$

Zwischengliederung

- Bislang:
    - Unterschiede zwischen Populationen
    - Punktschätzung von Parameterdifferenzen

- Jetzt:
    - Intervallschätzung von Parameterdifferenzen

Intervallschätzung von Parameterdifferenzen

Unabhängige Stichproben

Unabhängige Stichproben I

- Das Vorgehen bei der Konstruktion eines Konfidenzintervalls für die Parameterdifferenz $\mu_1 - \mu_2$ bei unabhängigen Stichproben ist analog zur Konstruktion des Parameters μ im Fall von einer einzigen Population.

- Wir gehen von der Schätzfunktion $\bar{X}_{Diff} = \bar{X}_1 - \bar{X}_2$ aus.

- Wir z-standardisieren diese Schätzfunktion:

$$Z = \frac{\bar{X}_{Diff} - E\left(\bar{X}_{Diff}\right)}{\sqrt{Var\left(\bar{X}_{Diff}\right)}} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}$$

- Wir ersetzen alle unbekannten Parameter im Nenner durch die entsprechende Schätzfunktion und erhalten so eine Zufallsvariable *T*:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_{pool}^2}{n_1} + \frac{S_{pool}^2}{n_2}}}$$

Unabhängige Stichproben II

- Die Wahrscheinlichkeitsverteilung P_T dieser Zufallsvariable *T* ist genau bestimmbar:

$$T \sim t(n_1 + n_2 - 2)$$

- Wir können dann auf Basis dieser t-Verteilung die Quantile $t_{\frac{\alpha}{2}}$ und $t_{1-\frac{\alpha}{2}}$ berechnen, so dass

$$P\left(t_{\frac{\alpha}{2}} \leq T \leq t_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

- Dann setzen wir *T* ein und lösen die Ungleichungen nach $\mu_1 - \mu_2$ auf:

$$P\left((\bar{X}_1 - \bar{X}_2) - t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S_{pool}^2}{n_1} + \frac{S_{pool}^2}{n_2}} \leq \mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S_{pool}^2}{n_1} + \frac{S_{pool}^2}{n_2}} \right) = 1 - \alpha$$

Unabhängige Stichproben III

- Dies ergibt das zufällige Konfidenzintervall

$$I(X_1, \dots, X_n) = \left[(\bar{X}_1 - \bar{X}_2) - t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S_{pool}^2}{n_1} + \frac{S_{pool}^2}{n_2}}, (\bar{X}_1 - \bar{X}_2) + t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S_{pool}^2}{n_1} + \frac{S_{pool}^2}{n_2}}\right]$$

und das konkrete in unserer Stichprobe berechenbare Konfidenzintervall

$$I(x_1, \dots, x_n) = \left[(\bar{x}_1 - \bar{x}_2) - t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s_{pool}^2}{n_1} + \frac{s_{pool}^2}{n_2}}, (\bar{x}_1 - \bar{x}_2) + t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s_{pool}^2}{n_1} + \frac{s_{pool}^2}{n_2}}\right]$$

Beispiel I

- Beispiel 1: Uns interessiert der Unterschied zwischen Psycholog*innen und BWLer*innen in der durchschnittlichen Leistungsmotivation.

- Wir ziehen zwei unabhängige einfache Zufallsstichproben:
  - Stichprobe 1 aus der Population der Psycholog*innen mit Umfang $n_1 = 101$.
  - Stichprobe 2 aus der Population der BWLer*innen $n_2 = 51$.

- In diesem Fall entspricht die uns interessierende Differenz $\bar{x}_{Psychologie} - \bar{x}_{BWL}$ der Parameterdifferenz $\mu_1 - \mu_2$.

Beispiel II

- In der Stichprobe der Psycholog*innen erhalten wir eine durchschnittliche Leistungsmotivation von $\bar{x}_1 = 165$ und einen Schätzwert der Varianz von $s_1^2 = 81$.

- In der Stichprobe der BWLer*innen erhalten wir eine durchschnittliche Leistungsmotivation von $\bar{x}_2 = 170$ und einen Schätzwert der Varianz von $s_2^2 = 100$.

- Als Schätzwert für $\mu_1 - \mu_2$, also für den Unterschied zwischen Psycholog*innen und BWLer*innen in der durchschnittlichen Leistungsmotivation, ergibt sich damit

$$\bar{x}_{Diff} = \bar{x}_1 - \bar{x}_2 = 165 - 170 = -5$$

- Als Schätzwert für σ^2 ergibt sich:

$$s_{pool}^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2} = \frac{(101 - 1) \cdot 81 + (51 - 1) \cdot 100}{101 + 51 - 2} = 87.33$$

Beispiel III

- Wir wählen ein Konfidenzniveau von $1 - \alpha = 0.95$. Wir suchen also $t_{1-\frac{\alpha}{2}} = t_{0.975}$.
- $\nu = n_1 + n_2 - 2 = 101 + 51 - 2 = 150$
- Damit können wir $t_{0.975}$ in R berechnen:

```
> qt(0.975, 150)
[1] 1.975905
```

- Also ist $t_{0.975} \approx 1.98$

Beispiel IV

- Als konkretes Konfidenzintervall ergibt sich damit:

$$I(x_1, \ldots, x_n) = \left[(\bar{x}_1 - \bar{x}_2) - t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s_{pool}^2}{n_1} + \frac{s_{pool}^2}{n_2}}, (\bar{x}_1 - \bar{x}_2) + t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s_{pool}^2}{n_1} + \frac{s_{pool}^2}{n_2}}\right]$$

$$= \left[-5 - 1.98 \cdot \sqrt{\frac{87.33}{101} + \frac{87.33}{51}}, -5 + 1.98 \cdot \sqrt{\frac{87.33}{101} + \frac{87.33}{51}}\right]$$

$$= [-8.18, -1.82]$$

- Die plausiblen Werte für $\mu_1 - \mu_2$ und somit für den Unterschied zwischen Psycholog*innen und BWLer*innen in der durchschnittlichen Leistungsmotivation liegen also zwischen -8.18 und -1.82.

Abhängige Stichproben

Abhängige Stichproben I

- Auch bei abhängigen Stichproben ist das Vorgehen bei der Konstruktion eines Konfidenzintervalls für die Parameterdifferenz $\mu_1 - \mu_2$ analog zur Konstruktion des Parameters μ im Fall einer einzigen Population.

- Wir gehen von der Schätzfunktion $\bar{X}_{Diff} = \bar{X}_1 - \bar{X}_2$ aus.

- Wir z-standardisieren diese Schätzfunktion:

$$Z = \frac{\bar{X}_{Diff} - E\left(\bar{X}_{Diff}\right)}{\sqrt{Var\left(\bar{X}_{Diff}\right)}} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2_{Diff}}{n}}}$$

- Wir ersetzen alle unbekannten Parameter im Nenner durch die entsprechende Schätzfunktion und erhalten so eine Zufallsvariable *T*:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S^2_{Diff}}{n}}}$$

Abhängige Stichproben II

Die Wahrscheinlichkeitsverteilung P_T dieser Zufallsvariable *T* ist genau bestimmbar:

$$T \sim t(n-1)$$

- Wir können dann auf Basis dieser t-Verteilung die Quantile $t_{\frac{\alpha}{2}}$ und $t_{1-\frac{\alpha}{2}}$ berechnen, so dass

$$P\left(t_{\frac{\alpha}{2}} \leq T \leq t_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

- Dann setzen wir *T* ein und lösen die Ungleichungen nach $\mu_1 - \mu_2$ auf:

$$P\left((\bar{X}_1 - \bar{X}_2) - t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S_{Diff}^2}{n}} \leq \mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S_{Diff}^2}{n}}\right) = 1-\alpha$$

Abhängige Stichproben III

- Dies ergibt das zufällige Konfidenzintervall

$$I(X_1, \dots, X_n) = \left[(\bar{X}_1 - \bar{X}_2) - t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S^2_{Diff}}{n}}, (\bar{X}_1 - \bar{X}_2) + t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S^2_{Diff}}{n}} \right]$$

und das konkrete in unserer Stichprobe berechenbare Konfidenzintervall

$$I(x_1, \dots, x_n) = \left[(\bar{x}_1 - \bar{x}_2) - t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s^2_{Diff}}{n}}, (\bar{x}_1 - \bar{x}_2) + t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s^2_{Diff}}{n}} \right]$$

Beispiel I

- Beispiel 3: Uns interessiert der Unterschied in der durchschnittlichen Matheleistung von Schüler*innen vor und nach einer Nachhilfestunde.

- Wir ziehen eine einfache Zufallsstichprobe von $n = 100$ Schüler*innen.
  - Wir erfassen die Matheleistung aller Schüler in unserer Stichprobe mithilfe eines Mathetests (höhere Werte im Test entsprechen einer höheren Matheleistung).
  - Alle Schüler*innen in unserer Stichprobe erhalten Nachhilfeunterricht.
  - Wir erfassen die Matheleistung aller Schüler*innen in unserer Stichprobe nochmals mithilfe des Mathetests.

- In diesem Fall entspricht die uns interessierende Differenz $\bar{x}_{vorher} - \bar{x}_{nachher}$ der Parameterdifferenz $\mu_1 - \mu_2$.

Beispiel II

- Vor der Nachhilfe erhalten wir in unserer Stichprobe eine durchschnittliche Matheleistung von $\bar{x}_1 = 110$.

- Nach der Nachhilfe erhalten wir in unserer Stichprobe eine durchschnittliche Matheleistung von $\bar{x}_2 = 100$.

- Als Schätzwert für $\mu_1 - \mu_2$, also für den Unterschied in der durchschnittlichen Matheleistung vor und nach der Nachhilfe, ergibt sich damit

$$\bar{x}_{Diff} = \bar{x}_1 - \bar{x}_2 = 110 - 100 = 10$$

- Als Schätzwert für σ^2_{Diff} erhalten wir:

$$s^2_{Diff} = 99$$

Beispiel III

- Wir wählen ein Konfidenzniveau von $1 - \alpha = 0.95$. Wir suchen also $t_{1-\frac{\alpha}{2}} = t_{0.975}$.
- $\nu = n - 1 = 100 - 1 = 99$
- Damit können wir $t_{0.975}$ in R berechnen:

```
> qt(0.975, 99)
[1] 1.984217
```

- Also ist $t_{0.975} \approx 1.98$

Beispiel IV

- Als konkretes Konfidenzintervall ergibt sich damit:

$$I(x_1, \dots, x_n) = I(x_1, \dots, x_n) = \left[(\bar{x}_1 - \bar{x}_2) - t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s^2_{Diff}}{n}}, (\bar{x}_1 - \bar{x}_2) + t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s^2_{Diff}}{n}}\right]$$

$$= \left[10 - 1.98 \cdot \sqrt{\frac{99}{100}}, 10 + 1.98 \cdot \sqrt{\frac{99}{100}}\right]$$

$$= [8.03, 11.97]$$

- Die plausiblen Werte für $\mu_1 - \mu_2$ und somit für den Unterschied in der Population zwischen Schüler*innen vor und nach der Nachhilfe in der durchschnittlichen Matheleistung liegen also zwischen 8.03 und 11.97.

Kausale Interpretation von Parameterunterschieden

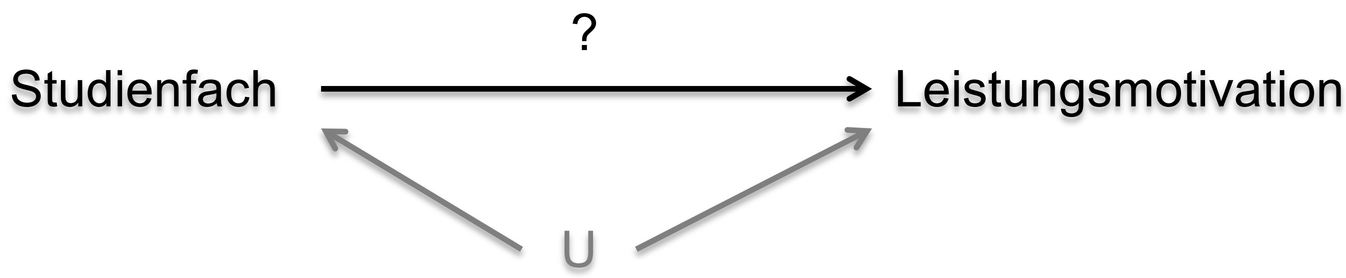
Kausale Interpretation in den 3 Beispielen

- In welchen der 3 Beispielen können wir die Schätzwerte und Konfidenzintervalle kausal interpretieren?
- Zur Erinnerung: Wenn es einen kausalen Einfluss der UV auf die AV gibt, sollte eine Manipulation der UV zu einer Veränderung in der AV führen (aber nicht umgekehrt).

- Beispiel 1: Schätzen $\bar{x}_{Diff}$ und das entsprechende KI den kausalen Effekt des Studienfachs (BWL vs. Psychologie) auf die Leistungsmotivation?
- Beispiel 2: Schätzen $\bar{x}_{Diff}$ und das entsprechende KI den kausalen Effekt der Psychotherapie (Therapie vs. keine Therapie) auf die Depressionsschwere?
- Beispiel 3: Schätzen $\bar{x}_{Diff}$ und das entsprechende KI den kausalen Effekt der Nachhilfestunde (vorher vs. nachher) auf die Matheleistung?

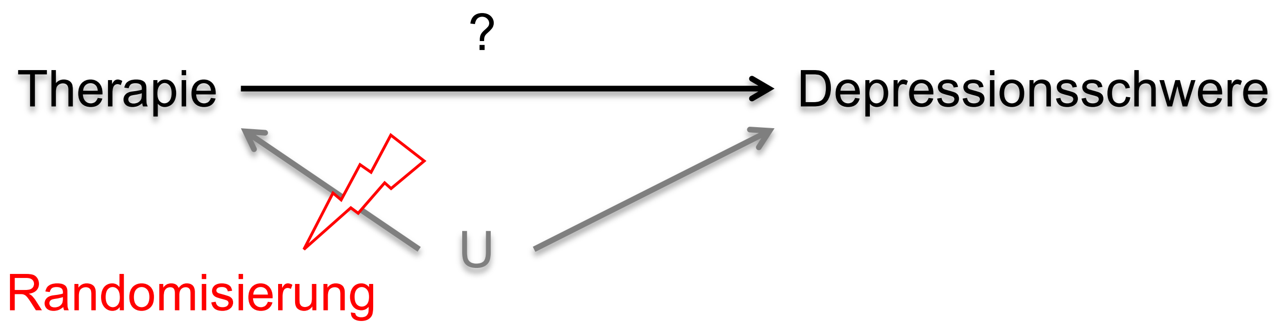
Kausale Interpretation Beispiel 1

- Beispiel 1: Schätzen $\bar{x}_{Diff}$ und das entsprechende KI den kausalen Effekt des Studienfachs (BWL vs. Psychologie) auf die Leistungsmotivation?
- Antwort: Nein
- Das Studienfach ist nicht zufällig zugeteilt (sondern von den Personen selbst gewählt). Es wird viele Drittvariablen (U) geben, die sich kausal sowohl auf die Wahl des Studienfachs, als auch auf die Leistungsmotivation auswirken.
- Berücksichtigt man U in der statistischen Analyse (z.B. im Rahmen einer Multiplen Linearen Regression; siehe Statistik II) ist eine kausale Interpretation mit zusätzlichen Annahmen eventuell doch möglich.



Kausale Interpretation Beispiel 2

- Beispiel 2: Schätzen $\bar{x}_{Diff}$ und das entsprechende KI den kausalen Effekt der Psychotherapie (Therapie vs. keine Therapie) auf die Depressionsschwere?
- Antwort: Ja
- Weil die Personen zufällig zur Therapiegruppe zugeteilt wurden, ist damit ausgeschlossen, dass Drittvariablen (U) vorliegen, die sich kausal sowohl auf die Gabe der Therapie, als auch auf die Depressionsschwere auswirken können.
- Das hier verwendete Studiendesign ist der Prototyp für kontrollierte randomisierte Studien (RCTs) in der Psychologie, Medizin und anderen empirischen Wissenschaften.



Kausale Interpretation Beispiel 3

- Beispiel 3: Schätzen $\bar{x}_{Diff}$ und das entsprechende KI den kausalen Effekt der Nachhilfestunde (vorher vs. nachher) auf die Matheleistung?
- Antwort: Nein
- Es wurden nur Personen betrachtet, die eine Nachhilfestunde bekommen haben. Wir können zwar feststellen, inwiefern die durchschnittliche Matheleistung nach der Nachhilfestunde höher ist als vorher. Ob das wirklich (nur) an der Nachhilfestunde liegt oder (auch) an anderen Dingen, die in der Zwischenzeit passiert sind, wissen wir aber nicht. Dafür bräuchten wir eine Kontrollgruppe, die keine Nachhilfestunde bekommt, aber deren Matheleistung trotzdem zweimal gemessen wird.

Exkurs: Average Treatment Effect (ATE)

Was bedeutet der kausale Effekt in Beispiel 2 genau?

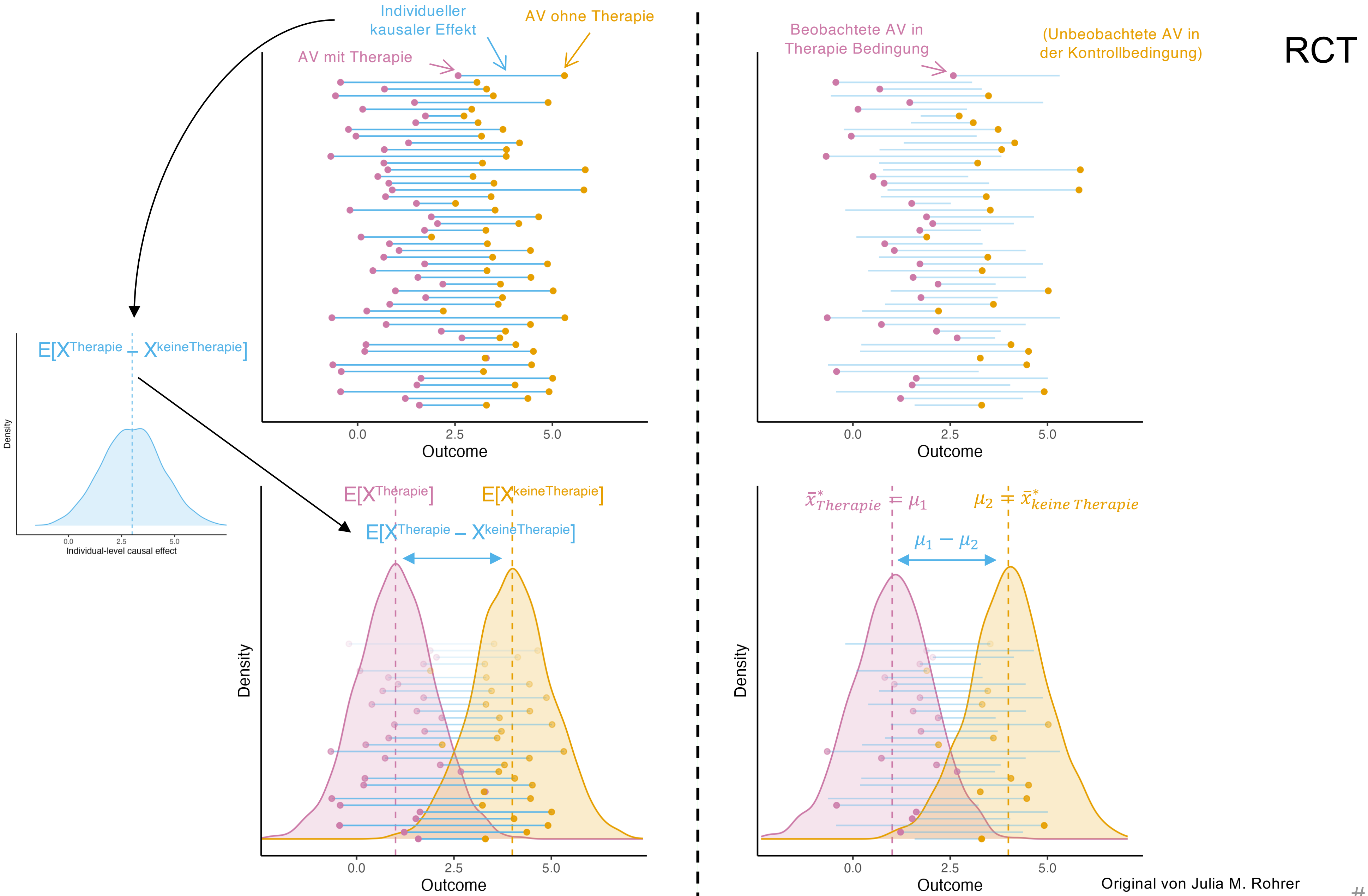
* Annahme: RCT, d.h. zufällige Zuteilung zur Experimental- oder Kontrollgruppe

- $\bar{x}^*_{Therapie} - \bar{x}^*_{keineTherapie} \stackrel{*}{=} E\left(X_i^{Therapie} - X_i^{keineTherapie}\right)$
- $X_i^{Therapie}$: Die Depressionsschwere der Person i, unter der Annahme dass sie eine Psychotherapie bekommt.
- $X_i^{keineTherapie}$: Die Depressionsschwere der gleichen Person i, unter der Annahme dass sie keine Psychotherapie bekommt.
- $X_i^{Therapie} - X_i^{keineTherapie}$: Der individuelle kausale Effekt der Person i. Gibt für jede Person in der Population an, wie diese von der Therapie profitieren würde. Kann nie beobachtet werden, weil in der Realität jede Person immer entweder Therapie bekommt oder nicht.
- $E\left(X_i^{Therapie} - X_i^{keineTherapie}\right)$: Der „durchschnittliche" kausale Effekt, gemittelt über alle Personen in der Population. Diese Größe wird in der Literatur als **Average Treatment Effect (ATE)** bezeichnet. Im Gegensatz zu den individuellen Effekten, ist der ATE die Art von kausalem Effekt, die in randomisierten kontrollierten Studien (RCTs) geschätzt werden kann.

Lehrstuhl für Psychologische Methodenlehre und Diagnostik der Ludwig-Maximilians-Universität München

Exkurs: Grafische Veranschaulichung des ATE

Vorlesung Statistische Methoden I WS 25/26



Original von Julia M. Rohrer (https://juliarohrer.com/resources/)

Exkurs: Was genau ist $\bar{x}^*_{Therapie} - \bar{x}^*_{keineTherapie}$?

- Die Größe $\bar{x}^*_{Therapie} - \bar{x}^*_{keineTherapie}$, die wir in Beispiel 2 geschätzt haben, bezieht sich also auf eine hypothetische Population, in der alle Personen zufällig entweder Therapie bekommen haben oder nicht. Diese Größe schätzen wir, indem wir eine einfache Zufallsstichprobe aus allen Personen ziehen und diese Personen dann zufällig entweder der Experimental- oder der Kontrollgruppe zuweisen.

- $\bar{x}^*_{Therapie} - \bar{x}^*_{keineTherapie}$ bezieht sich *nicht* auf die tatsächlichen Populationen von Personen, die in der Realität Therapie absolviert haben oder nicht. In der Realität werden Personen nicht zufällig zu einer Therapie zugeteilt, sondern entscheiden selbst, ob Sie Psychotherapie machen oder nicht.

- Zum Beispiel gibt es in Deutschland vermutlich sehr viel mehr Personen ohne Psychotherapie also solche mit. Außerdem entscheiden sich vermutlich mehr depressive Personen für eine Psychotherapie, sodass in der Realität die mittlere Depressionsschwere von Personen nach Psychotherapie vermutlich höher ist als bei zufälliger Zuteilung. Den tatsächlichen Mittelwertsunterschied könnten wir schätzen, wenn wir eine einfache Zufallsstichprobe aus den Personen mit Therapie ziehen und eine aus den Personen ohne. Die damit geschätzte deskriptivstatistische Größe wäre allerdings theoretisch viel weniger interessant, weil Sie uns nicht die kausale Fragestellung beantwortet, wie gut die Therapie wirkt.

Zusammenfassung

- Für Forschungsfragen sind häufig nicht nur Schätzungen einzelner Parameter interessant, sondern auch Schätzungen von **Parameterdifferenzen**.
- Falls Parameterdifferenzen $\mu_1 - \mu_2$ geschätzt werden sollen, muss man unterscheiden, ob es sich um **abhängige** oder **unabhängige** Stichproben handelt.
- Als **Punktschätzung** kann in beiden Fällen die erwartungstreue, effiziente und konsistente Schätzfunktion $\bar{X}_1 - \bar{X}_2$ verwendet werden.
- Für eine **Intervallschätzung** können für beide Fälle jeweils Konfidenzintervalle berechnet werden, die nach der gleichen Logik des KIs für μ aufgebaut sind.
- Wird zuerst eine einfache Zufallsstichprobe aus der Population gezogen und dann jede gezogene Person zufällig einer der beiden Gruppen der UV zugeteilt, kann $\mu_1 - \mu_2$ als kausaler Effekt der diskreten UV auf die stetige AV interpretiert werden.
- (Bemerkung: Eine Schätzung von Parameterdifferenzen $\pi_1 - \pi_2$ für diskrete Merkmale ist ebenfalls möglich. Entsprechende Konfidenzintervalle haben wir jedoch aus Zeitgründen nicht besprochen.)