Statistik
04. z-Test und t-Test

Roland Schäfer

Institut für Germanistische Sprachwissenschaft
Friedrich-Schiller-Universität Jena

stets aktuelle Fassungen: `https://github.com/rsling/VL-Deutsche-Syntax`

Inhalt

Übersicht

Übersicht

- Wann sind Unterschiede zwischen Mittelwerten signifikant?

Übersicht

- Wann sind Unterschiede zwischen Mittelwerten signifikant?
- Mittelwerte in Grundgesamtheiten und Stichproben

Literatur

- Gravetter & Wallnau (2007)
- Bortz & Schuster (2010)

- oder eben gleich Fisher (1935)

Wiederholungen

Tests in Fishers Philosophie

1. Nullhypothese (H0) festlegen: Der theoretisch angenommene Effekt existiert nicht (z. B.: Die Versuchsperson [VP] kann nicht erkennen, ob Tee oder Milch zuerst in der Tasse war).
2. Stichprobengröße und Versuchsaufbau festlegen (z. B. acht Tassen mit vier *Tee zuerst*-Tassen; VP kennt das Verhältnis)
3. *sig*-Niveau festlegen: Wie unwahrscheinlich darf das Ergebnis unter Annahme der H0 sein, damit wir die H0 zurückweisen.
4. Experiment durchführen, Ergebnis messen.
5. p-Wert berechnen: Wie wahrscheinlich war es, dieses Ergebnis oder ein extremeres Ergebnis zu erreichen, wenn die H0 die Welt korrekt beschreibt.
6. Wenn $p \leq sig$, dann H0 zurückweisen: Entweder der Effekt existiert (z. B. die VP kann die Reihenfolge des Einschenkens erkennen) oder ein seltenes Ereignis ist eingetreten.

Einschränkungen und Probleme bei der Interpretation

- Voraussetzung: echte Zufallsstichprobe
- Ergebnis: kein Beweis
- keine Auskunft darüber, wie „wahrscheinlich“ der Effekt ist
- keine Auskunft darüber, wie stark wir von der Existenz des Effekts überzeugt sein sollten (= *inverse probability*)
- jede H0-Zurückweisung: nur ein kleinteiliger Hinweis auf einen Effekt
- substantielle theoretische Hypothese oft und hart testen!

- Sensitivity: keine Auskunft über die Stärke des Effekts
  - große Stichprobe → hohe Sensitivität
  - kleine Strichprobe → niedrige Sensitivität
  - je sensitiver desto leichter werden schwache Effekte signifikant
  - Abhilfe bei Neyman-Pearson: Power (Teststärke) vor dem Experiment
  - quasi-kompatibel zu Fisher: Effektstärke nach dem Experiment

Und beim Konfidenzintervall?

Am Beispiel des 95%-Konfidenzintervalls (KI)

- Falsch: Wir können zu 95% sicher sein, dass der wahre Wert im KI liegt.
- Falsch: Der wahre Wert liegt mit 95% Wahrscheinlichkeit im KI.

- Warum? Wenn der wahre Wert nicht im geschätzten KI liegt, ist die Wahrscheinlichkeit 1, dass er nicht im KI liegt.
- Fakten haben die Wahrscheinlichkeit 1.

- Richtig: Entweder liegt der wahre Wert im KI oder ein seltenes Ereignis ist eingetreten
- „selten“ heißt: nur in 5 von 100 Fällen (im Grenzwert)

Exakter vs. asymptotischer Test

- exakter Test:
  - Die Wahrscheinlichkeitverteilung ist bekannt und wird direkt zugrunde gelegt (= Berechnung der exakten Wahrhscheinlichkeit).
  - Fisher-Test, Binomialtest
  - hohe Sensitivität
  - geeignet für kleine Stichproben
  - oft rechenintensiv
- approximativer oder asymptotischer Test:
  - Die Wahrscheinlichkeitsverteilung ist nicht bekannt (oder kann mathematisch nicht effizient zugrundegelegt werden) und es wird ein Differenzwert berechnet, der asymptotisch eine bekannte Verteilung hat.
  - χ^2-Test, t-Test, ANOVA
  - oft wird Normalverteilung approximiert
  - wegen asymptotischer Natur weniger sensitiv (= größere Stichprobe)

Parametrische und nichtparametrische Tests

- parametrischer Test:
  - Messung eines Parameters/mehrerer Parameter der Grundgesamtheit
  - (Parameter entsprechen in der Messung einer Variable)
  - zum Beispiel Mittelwert oder Varianz
  - Voraussetzung: bekannte Wahrscheinlichkeitsverteilung der Variable
  - z. B. t-Test (mittel), ANOVA (Varianz)
- nichtparametrischer Test:
  - keine direkte Messung eines zufallsverteilten Parameters
  - zum Beispiel Ränge oder Zähldaten
  - keine Verteilungsannahmen (auch: *verteilungsfreier Test*)
  - z. B. χ^2, Binomialtest, H-Test, U-Test

t-Test

Fragestellung beim z-Test und beim Einstichproben-t-Test

- Mittel μ über X in der Grundgesamtheit bekannt
  (z. B. mittlere Satzlänge im Korpus).

Fragestellung beim z-Test und beim Einstichproben-t-Test

- Mittel μ über X in der Grundgesamtheit bekannt
  (z. B. mittlere Satzlänge im Korpus).
- Stichprobe (z. B. der Grundriss von PE) zeigt gemessenes Mittel $\bar{x}$.

Fragestellung beim z-Test und beim Einstichproben-t-Test

- Mittel μ über X in der Grundgesamtheit bekannt (z. B. mittlere Satzlänge im Korpus).
- Stichprobe (z. B. der Grundriss von PE) zeigt gemessenes Mittel $\bar{x}$.
- Ist die Abweichung signifikant?

Fragestellung beim z-Test und beim Einstichproben-t-Test

- Mittel μ über X in der Grundgesamtheit bekannt (z. B. mittlere Satzlänge im Korpus).
- Stichprobe (z. B. der Grundriss von PE) zeigt gemessenes Mittel $\bar{x}$.
- Ist die Abweichung signifikant?
- H0: $\bar{x} = \mu$

z-Test

Wäre die Varianz der GG als $s^2(X)$ bekannt:

- SF(X) bei Stichprobengröße n ausrechnen, und...

Aufgabe: Mit Ihrer Stichprobe aus NaB und $\mu = 6.8$ sowie $s^2(X) = 10.8$ z-Test rechnen. (Bzw. erstmal die nötigen Werte ausrechnen. Wir besprechen dann die Interpretation als Test.)

z-Test

Wäre die Varianz der GG als $s^2(X)$ bekannt:

- SF(X) bei Stichprobengröße n ausrechnen, und…
- mit $z = \frac{\bar{x}-\mu}{SF(X)}$ einen Signifikanztest über Normalverteilung rechnen

Aufgabe: Mit Ihrer Stichprobe aus NaB und $\mu = 6.8$ sowie $s^2(X) = 10.8$ z-Test rechnen. (Bzw. erstmal die nötigen Werte ausrechnen. Wir besprechen dann die Interpretation als Test.)

z-Test

Wäre die Varianz der GG als $s^2(X)$ bekannt:

- SF(X) bei Stichprobengröße n ausrechnen, und...
- mit $z = \frac{\bar{x}-\mu}{SF(X)}$ einen Signifikanztest über Normalverteilung rechnen
- Problem aber leider: $SF(X) = \frac{s(X)}{\sqrt{n}}$

Aufgabe: Mit Ihrer Stichprobe aus NaB und $\mu = 6.8$ sowie $s^2(X) = 10.8$ z-Test rechnen. (Bzw. erstmal die nötigen Werte ausrechnen. Wir besprechen dann die Interpretation als Test.)

z-Test

Wäre die Varianz der GG als $s^2(X)$ bekannt:

- SF(X) bei Stichprobengröße n ausrechnen, und…
- mit $z = \frac{\bar{x}-\mu}{SF(X)}$ einen Signifikanztest über Normalverteilung rechnen

- Problem aber leider: $SF(X) = \frac{s(X)}{\sqrt{n}}$
- und $s^2(X)$ meist nicht bekannt!

Aufgabe: Mit Ihrer Stichprobe aus NaB und $\mu = 6.8$ sowie $s^2(X) = 10.8$ z-Test rechnen. (Bzw. erstmal die nötigen Werte ausrechnen. Wir besprechen dann die Interpretation als Test.)

Annahme beim t-Test mit einer Stichprobe

- Wir kennen μ oder haben eine Hypothese (z. B. $\mu = 0.5$).

Annahme beim t-Test mit einer Stichprobe

- Wir kennen μ oder haben eine Hypothese (z. B. $\mu = 0.5$).
- Wir haben eine Stichprobe x mit n und bekannten $\bar{x}$ und $s^2(x)$.

Annahme beim t-Test mit einer Stichprobe

- Wir kennen μ oder haben eine Hypothese (z. B. $\mu = 0.5$).
- Wir haben eine Stichprobe x mit n und bekannten $\bar{x}$ und $s^2(x)$.

- anders als bei z-Test: Wir schätzen $SF(X) \approx SF(x)$!

t-Formel

$$t = \frac{\bar{x} - \mu}{SF(x)}$$

t-Formel

$$t = \frac{\bar{x} - \mu}{SF(x)}$$

Bitte rechnen für Satzlängen (in Wörtern):

$$\mu = 7.3$$

$$x = [6, 3, 12, 16, 8, 15, 9, 9, 2, 11]$$

Lösung

1. $\bar{x} = 9.1$

Lösung

1. $\bar{x} = 9.1$
2. $s^2(x) = 21.43$

Lösung

1. $\bar{x} = 9.1$
2. $s^2(x) = 21.43$
3. $s(x) = 4.63$

Lösung

1. $\bar{x} = 9.1$
2. $s^2(x) = 21.43$
3. $s(x) = 4.63$
4. $SF(x) = \frac{4.63}{\sqrt{10}} = 1.464$

Lösung

1. $\bar{x} = 9.1$
2. $s^2(x) = 21.43$
3. $s(x) = 4.63$
4. $SF(x) = \frac{4.63}{\sqrt{10}} = 1.464$
5. $t = \frac{9.1-7.3}{1.464} = 1.229$

Lösung

1. $\bar{x} = 9.1$
2. $s^2(x) = 21.43$
3. $s(x) = 4.63$
4. $SF(x) = \frac{4.63}{\sqrt{10}} = 1.464$
5. $t = \frac{9.1-7.3}{1.464} = 1.229$

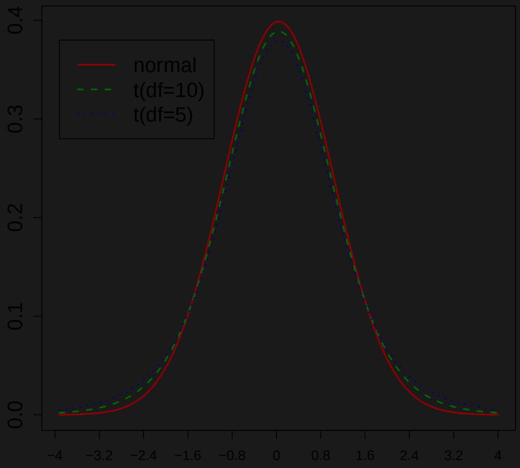
Lösung

1. $\bar{x} = 9.1$
2. $s^2(x) = 21.43$
3. $s(x) = 4.63$
4. $SF(x) = \frac{4.63}{\sqrt{10}} = 1.464$
5. $t = \frac{9.1-7.3}{1.464} = 1.229$

Und was sagt uns $t = 1.229$?

t-Verteilung

Während die z-Werte normalverteilt sind,
flacht die Verteilung der t-Werte durch die Schätzung
je nach df verglichen mit der Normalverteilung ab.



df und Signifikanz

- $df = n - 1$ ($\bar{x}$ muss für $s^s(x)$ bekannt sein)

df und Signifikanz

- $df = n - 1$ ($\bar{x}$ muss für $s^s(x)$ bekannt sein)
- Welche t-Werte machen $1 - \alpha$ der Werte aus?

df und Signifikanz

- $df = n - 1$ ($\bar{x}$ muss für $s^s(x)$ bekannt sein)
- Welche t-Werte machen $1 - \alpha$ der Werte aus?
- `> qt(c(0+0.05/2, 1-0.05/2), df=9)`
  $\Rightarrow$ 2.262157.. - 2.262157

df und Signifikanz

- $df = n - 1$ ($\bar{x}$ muss für $s^s(x)$ bekannt sein)
- Welche t-Werte machen $1 - \alpha$ der Werte aus?
- `> qt(c(0+0.05/2, 1-0.05/2), df=9)`
  $\Rightarrow$ 2.262157.. - 2.262157
- Der errechnete t-Wert ist nicht signifikant.

df und Signifikanz

- $df = n - 1$ ($\bar{x}$ muss für $s^s(x)$ bekannt sein)
- Welche t-Werte machen $1 - \alpha$ der Werte aus?
- `> qt(c(0+0.05/2, 1-0.05/2), df=9)`
  $\Rightarrow$ 2.262157.. – 2.262157
- Der errechnete t-Wert ist nicht signifikant.
- H0: $\mu = \bar{x}$ nicht zurückgewiesen.

Effektstärke

- Signifikanz ≠ starker Effekt

Effektstärke

- Signifikanz ≠ starker Effekt
- Effektstärke beim t-Test für Stichprobe x:

Effektstärke

- Signifikanz $\neq$ starker Effekt
- Effektstärke beim t-Test für Stichprobe x:

Effektstärke

- Signifikanz ≠ starker Effekt
- Effektstärke beim t-Test für Stichprobe x:

$$\text{Cohens } d = \frac{\bar{x}-\mu}{s(x)}$$

Effektstärke

- Signifikanz ≠ starker Effekt
- Effektstärke beim t-Test für Stichprobe x:

$$\text{Cohens } d = \frac{\bar{x}-\mu}{s(x)}$$

- Herleitung/Erklärung: Gravetter & Wallnau, Kap. 9

Erklärung der Varianz

- ähnlich der Effektstärke:
  Welcher Anteil der Varianz in den Daten
  wird durch die Unabhängige erklärt?

Erklärung der Varianz

- ähnlich der Effektstärke:
  Welcher Anteil der Varianz in den Daten
  wird durch die Unabhängige erklärt?

$$\text{Cohens } r^2 = \frac{t^2}{t^2 + df}$$

Erklärung der Varianz

- ähnlich der Effektstärke:
  Welcher Anteil der Varianz in den Daten
  wird durch die Unabhängige erklärt?

$$\text{Cohens } r^2 = \frac{t^2}{t^2+df}$$

- Herleitung/Erklärung: Gravetter & Wallnau, Kap. 9

Zwei-Stichproben t-Test

- zwei Grundgesamtheiten (z. B. dt. Sätze im 19. und im 20. Jh.)

Zwei-Stichproben t-Test

- zwei Grundgesamtheiten (z. B. dt. Sätze im 19. und im 20. Jh.)
- dazu: zwei Stichproben (je eine) mit einem Mittelwert (z. B. Länge)

Zwei-Stichproben t-Test

- zwei Grundgesamtheiten (z. B. dt. Sätze im 19. und im 20. Jh.)
- dazu: zwei Stichproben (je eine) mit einem Mittelwert (z. B. Länge)
- Interesse: anhand der zwei Stichproben zeigen,
  dass sie (sehr wahrscheinlich) aus zwei Grundgesamtheiten kommen

Zwei-Stichproben t-Test

- zwei Grundgesamtheiten (z. B. dt. Sätze im 19. und im 20. Jh.)
- dazu: zwei Stichproben (je eine) mit einem Mittelwert (z. B. Länge)
- Interesse: anhand der zwei Stichproben zeigen,
  dass sie (sehr wahrscheinlich) aus zwei Grundgesamtheiten kommen
- H0: $\mu_1 - \mu_0 = 0$

Zwei-Stichproben t-Test

- zwei Grundgesamtheiten (z. B. dt. Sätze im 19. und im 20. Jh.)
- dazu: zwei Stichproben (je eine) mit einem Mittelwert (z. B. Länge)
- Interesse: anhand der zwei Stichproben zeigen,
  dass sie (sehr wahrscheinlich) aus zwei Grundgesamtheiten kommen
- H0: $\mu_1 - \mu_0 = 0$
- hier also: eine unabhängige Variable (Jahrhundert)
  und eine abhängige Variable (Satzlänge), gemessen als Mittel

Genereller Ansatz

Allgemein funktioniert der t-Test immer so:

$$t = \frac{Stichprobenwert - Grundgesamtheitswert}{Standardfehler}$$

Genereller Ansatz

Allgemein funktioniert der t-Test immer so:

$$t = \frac{Stichprobenwert-Grundgesamtheitswert}{Standardfehler}$$

Jetzt geht man per Hypothese von zwei GG und zwei Stichproben aus, also:

Genereller Ansatz

Allgemein funktioniert der t-Test immer so:

$$t = \frac{Stichprobenwert - Grundgesamtheitswert}{Standardfehler}$$

Jetzt geht man per Hypothese von zwei GG und zwei Stichproben aus, also:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SF(x_1 - x_2)}$$

Genereller Ansatz

Allgemein funktioniert der t-Test immer so:

$$t = \frac{Stichprobenwert - Grundgesamtheitswert}{Standardfehler}$$

Jetzt geht man per Hypothese von zwei GG und zwei Stichproben aus, also:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SF(x_1 - x_2)}$$

- Wir testen also auf die Differenz der Unterschiede.

Genereller Ansatz

Allgemein funktioniert der t-Test immer so:

$$t = \frac{Stichprobenwert{-}Grundgesamtheitswert}{Standardfehler}$$

Jetzt geht man per Hypothese von zwei GG und zwei Stichproben aus, also:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SF(x_1 - x_2)}$$

- Wir testen also auf die Differenz der Unterschiede.
- Per H0 wird gesetzt: $\mu_1 - \mu_2 = 0$

Schätzung des Standardfehlers

Für gleichgroße Stichproben:

$$SF(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s^2(x_1)}{n_1} + \frac{s^2(x_2)}{n_2}}$$

Schätzung des Standardfehlers

Für gleichgroße Stichproben:

$$SF(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s^2(x_1)}{n_1} + \frac{s^2(x_2)}{n_2}}$$

- Problem: Beitrag zum SF von beiden Stichproben gleich.

Schätzung des Standardfehlers

Für gleichgroße Stichproben:

$$SF(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s^2(x_1)}{n_1} + \frac{s^2(x_2)}{n_2}}$$

- Problem: Beitrag zum SF von beiden Stichproben gleich.
- Besser: zusammengefasste Varianz, und daraus dann SF.

Schätzung des Standardfehlers

Für gleichgroße Stichproben:

$$SF(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s^2(x_1)}{n_1} + \frac{s^2(x_2)}{n_2}}$$

- Problem: Beitrag zum SF von beiden Stichproben gleich.
- Besser: zusammengefasste Varianz, und daraus dann SF.

Schätzung des Standardfehlers

Für gleichgroße Stichproben:

$$SF(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s^2(x_1)}{n_1} + \frac{s^2(x_2)}{n_2}}$$

- Problem: Beitrag zum SF von beiden Stichproben gleich.
- Besser: zusammengefasste Varianz, und daraus dann SF.

$$s_p^2(x_1, x_2) = \frac{(\sum_{i=1}^{n}(x_{1,i}-\bar{x}_1)^2)+(\sum_{i=1}^{n}(x_{2,i}-\bar{x}_2)^2)}{(n_1-1)+(n_2-1)} = \frac{SQ(x_1)+SQ(x_2)}{(n_1-1)+(n_2-1)}$$

Schätzung des Standardfehlers

Für gleichgroße Stichproben:

$$SF(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s^2(x_1)}{n_1} + \frac{s^2(x_2)}{n_2}}$$

- Problem: Beitrag zum SF von beiden Stichproben gleich.
- Besser: zusammengefasste Varianz, und daraus dann SF.

$$s_p^2(x_1, x_2) = \frac{(\sum_{i=1}^{n}(x_{1,i}-\bar{x}_1)^2)+(\sum_{i=1}^{n}(x_{2,i}-\bar{x}_2)^2)}{(n_1-1)+(n_2-1)} = \frac{SQ(x_1)+SQ(x_2)}{(n_1-1)+(n_2-1)}$$

$$SF(x_1 - x_2) = \sqrt{\frac{s_p^2(x_1,x_2)}{n_1} + \frac{s_p^2(x_1,x_2)}{n_2}}$$

Schätzung des Standardfehlers

Für gleichgroße Stichproben:

$$SF(\bar{x_1} - \bar{x_2}) = \sqrt{\frac{s^2(x_1)}{n_1} + \frac{s^2(x_2)}{n_2}}$$

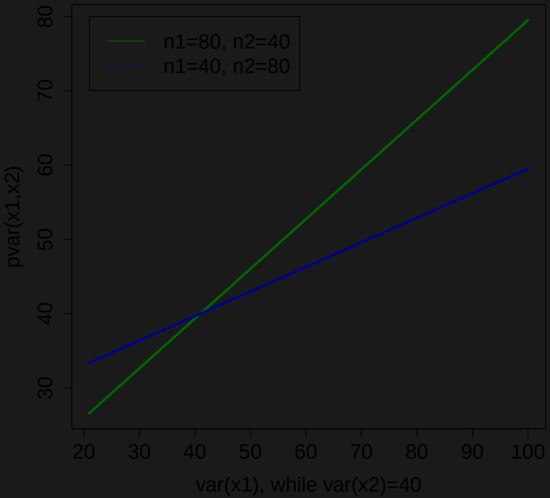
- Problem: Beitrag zum SF von beiden Stichproben gleich.
- Besser: zusammengefasste Varianz, und daraus dann SF.

$$s_p^2(x_1, x_2) = \frac{(\sum_{i=1}^{n}(x_{1,i} - \bar{x_1})^2) + (\sum_{i=1}^{n}(x_{2,i} - \bar{x_2})^2)}{(n_1 - 1) + (n_2 - 1)} = \frac{SQ(x_1) + SQ(x_2)}{(n_1 - 1) + (n_2 - 1)}$$

$$SF(x_1 - x_2) = \sqrt{\frac{s_p^2(x_1, x_2)}{n_1} + \frac{s_p^2(x_1, x_2)}{n_2}}$$

Mehr: Gravetter & Wallnau, Kap. 10

Illustration der zusammengefassten Varianz



Rechnen des Tests

t-Wert

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SF(x_1 - x_2)} = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{SF(x_1 - x_2)} = \frac{\bar{x}_1 - \bar{x}_2}{SF(x_1 - x_2)}$$

Rechnen des Tests

t-Wert

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SF(x_1 - x_2)} = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{SF(x_1 - x_2)} = \frac{\bar{x}_1 - \bar{x}_2}{SF(x_1 - x_2)}$$

Freiheitsgrade

$$df = df(x_1) + df(x_2) = (n_1 - 1) + (n_2 - 1)$$

Rechnen des Tests

t-Wert

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SF(x_1 - x_2)} = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{SF(x_1 - x_2)} = \frac{\bar{x}_1 - \bar{x}_2}{SF(x_1 - x_2)}$$

Freiheitsgrade

$$df = df(x_1) + df(x_2) = (n_1 - 1) + (n_2 - 1)$$

Effektstärke

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2}}$$

t-Wert

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SF(x_1 - x_2)} = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{SF(x_1 - x_2)} = \frac{\bar{x}_1 - \bar{x}_2}{SF(x_1 - x_2)}$$

Freiheitsgrade

$$df = df(x_1) + df(x_2) = (n_1 - 1) + (n_2 - 1)$$

Effektstärke

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2}}$$

Erklärung der Varianz

$$r^2 = \frac{t^2}{t^2 + df}$$

Übung

Bitte „von Hand in R“ t-Test für folgende zwei Stichproben bei $\alpha = 0.05$ rechnen:

$$x_1 = [11, 11, 8, 8, 11, 9, 8, 11, 9, 8]$$
$$x_2 = [10, 14, 14, 13, 11, 14, 10, 14, 12, 10]$$

Übung

Bitte „von Hand in R“ t-Test für folgende zwei Stichproben bei $\alpha = 0.05$ rechnen:

$$x_1 = [11, 11, 8, 8, 11, 9, 8, 11, 9, 8]$$
$$x_2 = [10, 14, 14, 13, 11, 14, 10, 14, 12, 10]$$

Und überprüfen mit:

```
> t.test(x1, x2)
```

Voraussetzungen prüfen I

Die GGs müssen normalerverteilt sein:

```
shapiro.test(x)
```

Wenn $p \leq 0.05$ wird die Nullhypothese des Shapiro-Wilk-Tests verworfen – H0: Die Werte stammen aus einer normalverteilten GG.

Voraussetzungen prüfen I

Die GGs müssen normalerverteilt sein:

```
shapiro.test(x)
```

Wenn $p \leq 0.05$ wird die Nullhypothese des Shapiro-Wilk-Tests verworfen – H0: Die Werte stammen aus einer normalverteilten GG.

Die Varianzen müssen homogen sein:

```
var.test(x1, x2)
```

Auch hier: $p \leq 0.05$ weist die H0 zurück (sehr informell) – H0: Die Varianzen von x1 und x2 sind homogen.

Voraussetzungen prüfen I

Die GGs müssen normalerverteilt sein:

```
shapiro.test(x)
```

Wenn $p \leq 0.05$ wird die Nullhypothese des Shapiro-Wilk-Tests verworfen –
H0: Die Werte stammen aus einer normalverteilten GG.

Die Varianzen müssen homogen sein:

```
var.test(x1, x2)
```

Auch hier: $p \leq 0.05$ weist die H0 zurück (sehr informell) –
H0: Die Varianzen von x1 und x2 sind homogen.

Solche Tests sind umstritten, weil sie i. d. R. viel zu empfindlich reagieren.
Zuur u. a. (2009) empfehlen z. B. grafische Methoden (bei linearen Modellen).

Voraussetzungen prüfen II

Wenn Voraussetzungen nicht erfüllt sind:

- steigt das Risiko für Typ 1-Fehler

Voraussetzungen prüfen II

Wenn Voraussetzungen nicht erfüllt sind:

- steigt das Risiko für Typ 1-Fehler
- nicht-parametrische Alternative nehmen

Voraussetzungen prüfen II

Wenn Voraussetzungen nicht erfüllt sind:

- steigt das Risiko für Typ 1-Fehler
- nicht-parametrische Alternative nehmen
- Daten transformieren

Voraussetzungen prüfen II

Wenn Voraussetzungen nicht erfüllt sind:

- steigt das Risiko für Typ 1-Fehler
- nicht-parametrische Alternative nehmen
- Daten transformieren
- sich über Robustheit des Test ggü. verletzten Annahmen informieren (oft schwer zugängliche und kontroverse Spezialliteratur)

Nächste Woche | Überblick

Einzelthemen

1. Inferenz
2. Deskriptive Statistik
3. Nichtparametrische Verfahren
4. z-Test und t-Test
5. ANOVA
6. Freiheitsgrade und Effektstärken
7. Power und Severity
8. Lineare Modelle
9. Generalisierte Lineare Modelle
10. Gemischte Modelle

Literatur I

Bortz, Jürgen & Christof Schuster. 2010. *Statistik für Human- und Sozialwissenschaftler*. 7. Aufl. Berlin: Springer.

Fisher, Ronald A. 1935. *The Design of Experiments*. London: Macmillan.

Gravetter, Frederick J. & Larry B. Wallnau. 2007. *Statistics for the Behavioral Sciences*. 7. Aufl. Belmont: Thomson.

Zuur, Alain F., Elena N. Ieno, Neil Walker, Anatoly A. Saveliev & Graham M. Smith. 2009. *Mixed effects models and extensions in ecology with R*. Berlin etc.: Springer.

Autor

Kontakt

Prof. Dr. Roland Schäfer
Institut für Germanistische Sprachwissenschaft
Friedrich-Schiller-Universität Jena
Fürstengraben 30
07743 Jena

https://rolandschaefer.net
roland.schaefer@uni-jena.de

Lizenz

Creative Commons BY-SA-3.0-DE