



Einführung in die Statistik

Praktische Übung – Jürgen Hermes – IDH – SoSe 2023

Programm heute

- Besprechung Testat-relevanter Themen
 - Einführung in R - Grundlagen, die im ersten Testat schon verwendet werden mussten
 - Visualisierung: Histogramme, Punktplots, Boxplots
 - Deskriptive Statistik:
 - Präsentationen 9 und 10 (Einführung, Mittelwerte, Dispersionsmaße, Korrelation)
 - Präsentation 11 (Regression) bis Folie 20 (Hausaufgabe)
 - Präsentation 12 (Inferenzstatistik) bis Folie 6 (Signifikanzniveau)
 - Testat: Insgesamt 120 Punkte - 48 Punkte sind ausreichend für die Studienleistung
 - 48 Punkte im Teil-Testat 1 (25.5.)
 - 72 Punkte im Teil-Testat 2 (12.7.)
 - davon 48 Punkte im R-Programmierteil
 - weitere 24 Punkte in ILIAS-Testat (Multiple Choice und Lückentext)
 - Fragen?
- Kurze Einführung Inferenzstatistik



Diese und die folgenden Folien sind erstellt worden von Sascha Wolfer für seinen Kurs "Statistik mit R" an der Uni Basel. Ich nutze sie mit seiner freundlichen Genehmigung. DOI für die Materialien ist

[10.5281/zenodo.7431504](https://doi.org/10.5281/zenodo.7431504)

Inferenzstatistik

Inferenzstatistik

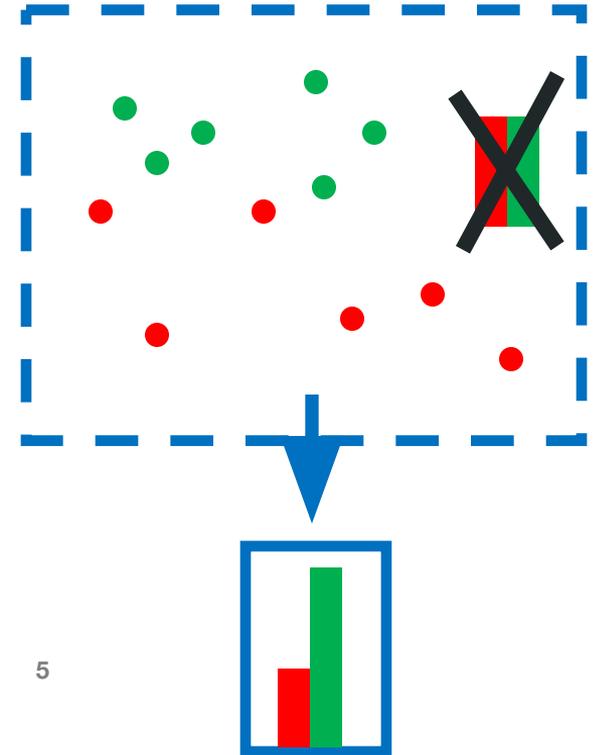
Man nennt dieses Vorgehen auch *null-hypothesis significance testing (NHST)*.

- Mit inferenzstatistischen Verfahren wollen wir von unserer **Stichprobe** auf die zugrundeliegende **Grundgesamtheit (Population)** verallgemeinern.
- Wir testen anhand unserer Stichprobe eine **Hypothese**, von der wir wissen möchten, ob sie in der Population gilt.
- Die Hypothese, die getestet wird, ist immer die **Nullhypothese**.
 - H_0 : Es gibt *keinen* Zusammenhang / Unterschied (= Effekt).
- Unsere Forschungshypothese ist stets die **Alternativhypothese**.
 - H_1 : Es gibt einen Zusammenhang / Unterschied.
- Wir wollen herausfinden, ob wir die Nullhypothese mit ausreichend großer Wahrscheinlichkeit **ablehnen** können.

4

Signifikanzniveau p

- Wir nehmen *hypothetisch* an, dass in der Grundgesamtheit die Nullhypothese gilt.
- Wir ziehen zufällig eine Stichprobe aus der Grundgesamtheit.
- p gibt uns an, wie wahrscheinlich die Ergebnisse in der Stichprobe sind, wenn in der Grundgesamtheit tatsächlich die Nullhypothese gelten sollte.
- Wenn p **klein genug** ist, lehnen wir die Nullhypothese ab und nehmen stattdessen die Alternativhypothese an.



Signifikanzniveau p

Permutationstests erlauben es, Effekte mit weniger Annahmen zu überprüfen. Siehe u.a. hier: <https://statologie.de/permutationstest-r/>

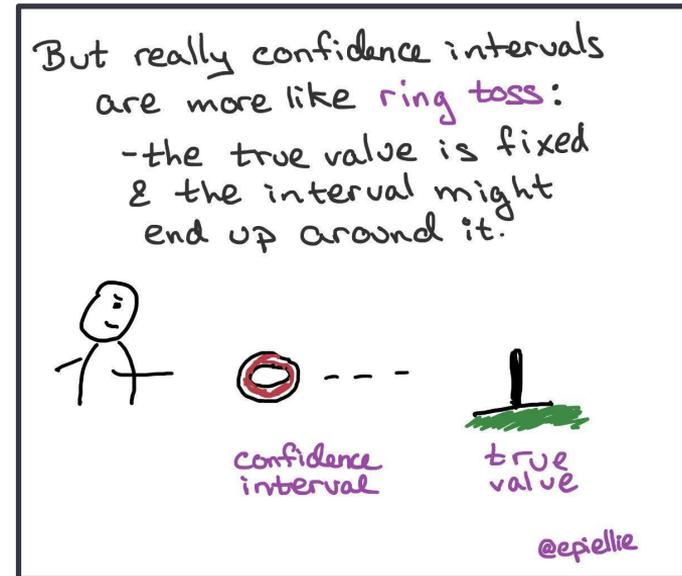
- "Klein genug": kleiner 5% ($p < 0,05$)
 - Andere Schwellenwerte: 0,01 (1%); 0,001 (0,1%)
- Unbedingte Voraussetzung: Zufällige Stichprobenziehung, ansonsten *sampling bias*
- p kann uns **nicht** dabei helfen, die Nullhypothese zu **bestätigen**.
 - Nicht-Effekte können also nicht interpretiert werden! Es ist somit in diesem Paradigma äußerst schwierig, Gleichheit oder Nicht-Zusammenhänge nachzuweisen.
- Alternativhypothese wird nie "bewiesen", wir können lediglich mit hoher Wahrscheinlichkeit die Nullhypothese ablehnen.

p -Werte in R & Effektgrößen

- Viele statistische Methoden geben p -Werte aus.
 - `summary()` von linearen Modellen
 - `cor.test()` statt `cor()`
- Statistische Signifikanz ist aber nicht alles. Auch (sehr) kleine Effekte können signifikant sein.
 - z. B. sehr kleine Unterschiede zwischen Gruppen, schwache Zusammenhänge
- Auch sehr große Effekte sind manchmal *nicht* signifikant ...
 - ... z. B., weil das assoziierte Rauschen sehr groß ist.
- 5% ist eine **arbiträre Grenze!**
 - Ist $p = 0,049$ wirklich veröffentlichungswert und $p = 0,051$ nicht?

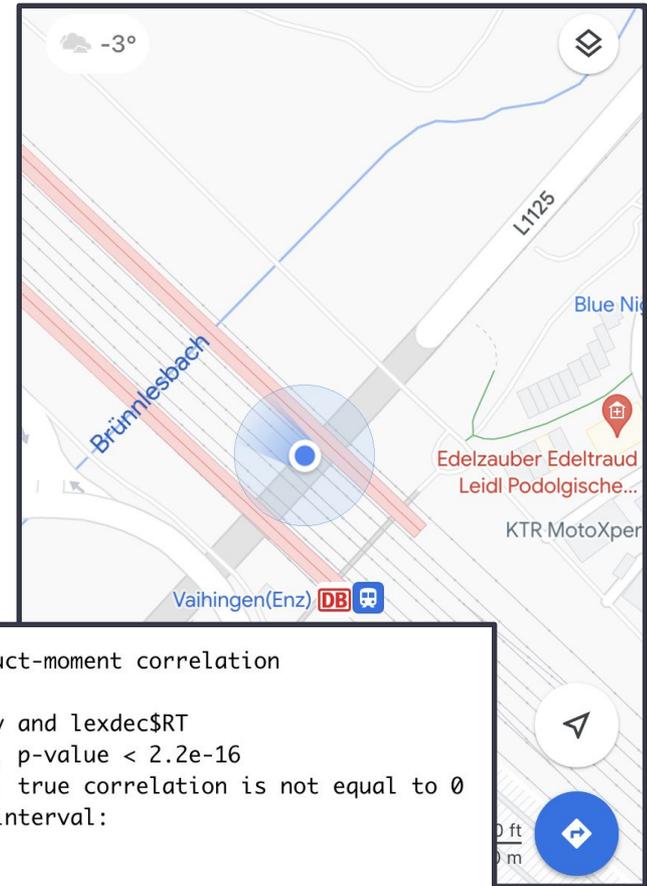
Konfidenzintervalle

- Wichtiges Konzept in der frequentistischen Inferenzstatistik
- Schätzen wir einen Parameter der Population mit einem bestimmten Verfahren, enthält ein bestimmter Anteil der Intervalle den wahren Wert.
 - 95%, 99%, 99,9%, ...
- Werfen wir wiederholt Ringe, um den echten Wert in der Population zu treffen, liegen 95%, 99%, ... der Ringe um den Wert.



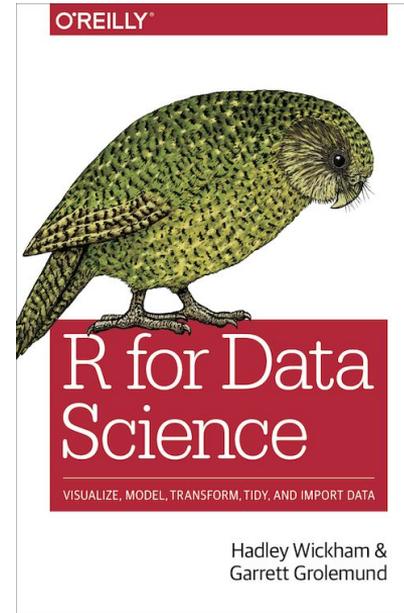
Konfidenzintervalle

- Konfidenzintervalle werden in R für eine Reihe von Tests ausgegeben.
- `confint(<Modell>)` gibt Konfidenzintervalle für die Effektschätzer aus.
 - `level = .95`

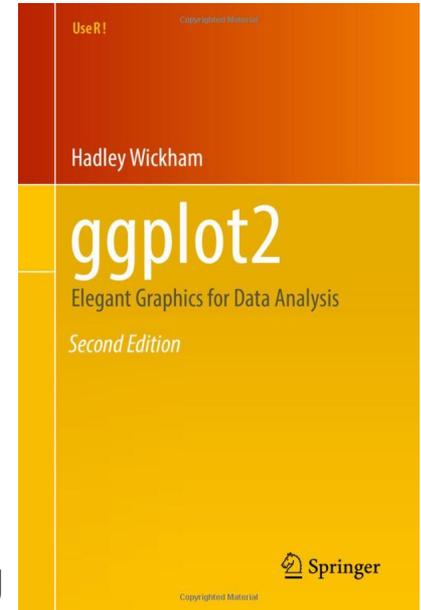


Hier geht's weiter...

- Fortgeschrittene Regressionsverfahren: Gemischte Modelle
- Explorative Verfahren: Clusteranalysen, Hauptkomponentenanalyse, Multidimensionale Skalierung
- Visualisierungen mit dem Paket `{ggplot2}`
- Datenmanipulation mit dem Paket `{dplyr}`
- ...



<https://r4ds.had.co.nz>



<https://ggplot2-book.org>