

Statistik in Maturaarbeiten

Patric Müller <patric.mueller@kfr.ch>

34. Schweizerischer Tag über Mathematik und Unterricht

11.09.2024



Kantonsschule Freudenberg Zürich
Liceo Artistico
Schweizerisch-Italienisches Kunstgymnasium

ETH zürich

- 1 Einführung
- 2 Datenerhebung
- 3 Modellwahl
- 4 Datenanalyse
- 5 Referenzen
- 6 *Vertiefung

- 1 Einführung
- 2 Datenerhebung
- 3 Modellwahl
- 4 Datenanalyse
- 5 Referenzen
- 6 *Vertiefung

Einführung

- In vielen Maturaarbeiten werden Umfragen und/oder Feldstudien durchgeführt. Diese muss(t)en die **wissenschaftlichen Kriterien** erfüllen.
- Oft reichen die Statistikenntnisse der SuS nicht aus, um ihre Datensätze vollständig zu analysieren.
- In diesem Seminar werden wichtige Aspekte einer **statistischen Beratung für Gymnasiast:innen**, die die Maturaarbeit schreiben behandelt.

Alle im Vortrag erwähnten Unterlagen sind hier zur Verfügung gestellt:
<https://polybox.ethz.ch/index.php/s/F7saNMOFRtmju3D>

Phasen eines Feldexperiments

Ein Feldexperiment oder die Durchführung einer Umfrage besteht aus 3 Phasen.

- Die Datenerhebung
 - ▶ Fragestellung (was will ich erreichen/untersuchen).
 - ▶ Methodik (wie werde ich es tun).
 - ▶ Durchführung der Studie auf dem Feld.
- Modellwahl für die Daten
 - ▶ Ist ausschlaggebend für die Datenanalyse.
 - ▶ Wird oft unterschätzt.
- Datenanalyse und Darstellung der Ergebnisse
 - ▶ Wahl geeigneter statistischer Methoden.
 - ▶ Beschreibende/ beurteilende Statistik?
 - ▶ Inhaltliche und optische Darstellung der Ergebnisse.

1 Einführung

2 Datenerhebung

- Fragestellung
- Datenerhebung
- Externe Einflüsse und Kontrollgruppe
- Stichprobengröße und Randomisierung
- Latente Variablen
- Systematischer Fehler
- Repräsentative Stichprobe
- Tipps aus der Erfahrung

3 Modellwahl

4 Datenanalyse

5 TMU-Referenzen

Datenerhebung

- ✓ **Qualitativ gute Daten** zu erheben ist mindestens so wichtig, wie gute Analysen zu machen.
- ⚠ Der Computer akzeptiert ohne zu protestieren gute, wie auch schlechte Daten und 'spuckt' ein/irgendein Ergebnis aus.
- *Worauf soll man achten bei der Phase der Datenerhebung? (Planung und Durchführung eines Experiments/Fragebogens)*

Klare Fragestellung und Messgrößen

- **Klare Fragestellungen** definieren.
 - ✓ Alle Konzepte/Begriffe sind wissenschaftlich definiert und messbar.
 - ✗ Diese Arbeit soll zeigen, dass eine gesunde Ernährung, die Gesundheit verbessert.
 - ✓ [...] die tägliche Aufnahme von 100mg Vitamin C das Risiko senkt, an einem grippalen Infekt zu erkranken.
- Definition **präzise Messgrößen**.
 - ▶ Wenig, viel,... und weitere subjektive Werte sollten womöglich vermieden werden, zugunsten von genauen Werten. (Objektive, lineare Skala)
 - ✓ Bereits vorhandene Messmethoden verwenden (z.B. Dyslexie, Krankheiten).

Externe Einflüsse und Kontrollgruppe

- **Setting** des Experiments soll so gut, wie möglich, für alle Probanden gleich sein.
 - ✗ Kaltduschen 2 Wochen lang. (Temperatur? Wie lange? Wie oft? Wann?...)
 - ✓ 12 Muffins pro Rezept gebacken, bei genau den gleichen Konditionen (Temperatur, Zeit, Muffinsform, Platz im Ofen etc.). Gewichtsmessung nach 0, 10 und 20 Minuten. Schmid (2021)
- **Kontrollgruppe** bei Vergleichen!
 - ✗ kalt zu duschen stärkt das Immunsystem? → Alle Probanden duschen kalt.
 - ✓ → Die Hälfte der Probanden duschen kalt, die anderen normal.

Stichprobegrösse, Randomisierung

● Stichprobegrösse und Randomisierung.

- ✓ Faustregel: $n \geq 10p$, das heisst pro geschätztem Parameter sollte man mindestens 10 Beobachtungen haben. (Harrell (2015)).
- ✓ Faustregel: Bei Fragebögen mindestens 30 – 50 Kandidaten.
- ✓ Werden Probanden in Kategorien unterteilt, dann sollte jede Kategorie etwa die gleiche Anzahl an Probanden haben.
 - ✗ 20 Schwarze Katzen, 20 Braune K., 20 S. Hunde, 1 B. H. (**Leverage Point**)
- ✓ Womöglich zufällige Gruppeneinteilung der Probanden.

Latente Variablen

- **Latente Variablen** = nicht gemessene Faktoren, die das Resultat beeinflussen (können).
 - ✓ Vor der Durchführung überlegen!
 - ✓ Nicht alle Faktoren können bedacht/kontrolliert werden. Dank der Randomisierung, bleiben die wissenschaftlichen Schlussfolgerungen valide, auch wenn latente Variablen vorhanden sind, solange diese keine systematischen Fehler erzeugen.
- Beispiele:
 - ▶ Untersuchung: Geschlechtsabhängige Lohnunterschiede. Mögliche latente Variablen: Alter, Erfahrung, Qualifikation, ...
 - ▶ Lebenserwartung in Abhängigkeit der Anzahl Fast Food. Latente Variable: Wohlstand(?)
 - ▶ Nicht alle Mitglieder der gleichen Familie in dieselbe Gruppe einteilen. (mögliche Korrelation)

Systematischer Fehler

- **Systematische Fehler** können in der Regel nicht 'mathematisch' erkannt werden.
- ⚠ Sie machen den ganzen Datensatz und die ganze Analyse nichtig.
- Bsp:
 - ▶ Wie viel verdienst du? (Systematischer Fehler, wegen unklarer Frage. Besser: Wie viel ist dein Bruttolohn pro Stunde?)
 - ▶ Hast du heute schon... (Uhrzeit der Frage beeinflusst die Antwort).
 - ▶ Messung der Grösse einer Person mit Schuhen.

Räpresentative Stichprobe

- Die Stichprobe soll räpresentativ für die ganze untersuchte Bevölkerung sein.
 - ✗ Fragebogen in der Vorlesung: 'Wie oft besuche ich die Vorlesung?'
 - ✓ Als Probanden, N SuS pro Klasse zufällig auswählen.

Weitere Tipps aus der Erfahrung

- ✓ Es empfiehlt sich **VOR** der Durchführung eines Experiments, die Methodik einer Expertin/einem Experten zur Überprüfung zu zeigen!
- ✓ Oft kann man mit kleinen Anpassungen die Datenerhebung massiv verbessern.
- ✓ Schon vor der Durchführung sollen die SuS klare und präzise **Fragestellungen formulieren** und die **statistisch relevanten Messgrößen** erkennen.
- ✓ Qualitativ gute Datensätze sind wichtiger als z.B. eine strenge Überprüfung aller Modellannahmen. (Systematische Fehler müssen aber vermieden werden).
- ✓ Daten und Ergebnisse sind immer mit kritischem Blick zu hinterfragen!

1 Einführung

2 Datenerhebung

3 **Modellwahl**

- Beispiel
- Wahl passender Modellen

4 Datenanalyse

5 Referenzen

6 *Vertiefung

Datensatz Beispiel

- Man betrachte folgenden Datensatz

Nr.	1	2	3	4	5	6	7	8	9	10
Wert	0	2	1	2	0	0	2	1	2	0

Welche Analyse eignet sich zum Datensatz?

Datensatz Beispiel

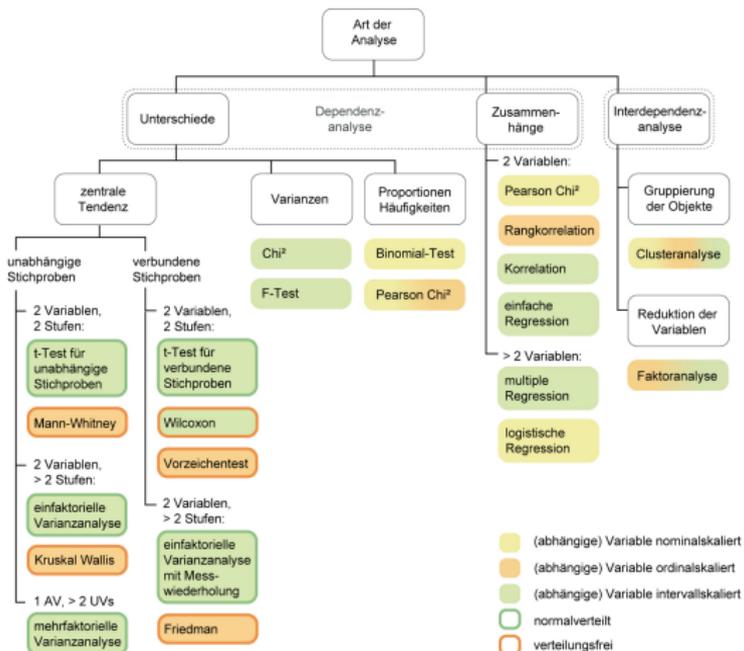
Nr.	1	2	3	4	5	6	7	8	9	10
Wert	0	2	1	2	0	0	2	1	2	0

- Fall 1: Werte 1,2 sind zwei Dünger, 0 ist die KG (kein Dünger).
 - ▶ Daten sind kategoriell (Faktoren)
 - ✗ Die Berechnung des Mittelwerts ist sinnlos.
- Fall 2: Antwort zur Frage: 'Hast du schon einen Kreuzbandriss gehabt?', 0=Nie, 1=Einmal, 2=Mehrmals.
 - ▶ Daten sind ordinalskaliert, aber nicht metrisch.
- Fall 3: Wie oft hast du auf dein Handy in den letzten 5' geschaut?
 - ▶ Daten sind metrisch
 - ✓ Die Berechnung des Mittelwerts ist sinnvoll.

Wahl passender Modellen

- Die Wahl des Modells beeinflusst die darauffolgende Datenanalyse.
- Es ist sehr wichtig, das passende Modell zu wählen.
 - ▶ Tools wie <https://datatab.de/> oder <https://www.statstest.com/> oder https://www.methodenberatung.uzh.ch/de/datenanalyse_spss.html helfen bei der Modellwahl und der Analyse.
 - ▶ Die Modellwahl beinhaltet eine mathematische und eine interpretative Komponente.
 - ▶ Computer führen Analyse und Berechnungen auch bei falscher Modellwahl aus.
- Modellannahmen können /sollen mathematisch überprüft werden. (siehe später).

Beispiel Entscheidungsbaum Modell und Analyse



Quelle: https://www.methodenberatung.uzh.ch/de/datenanalyse_sps.html

1 Einführung

2 Datenerhebung

3 Modellwahl

4 Datenanalyse

- Vorbemerkung - Wichtige Hinweise
- Paradebeispiel t-Test
- Paradebeispiel lineare Regression
- Paradebeispiel ANOVA

5 Referenzen

6 *Vertiefung

Wichtige Hinweise

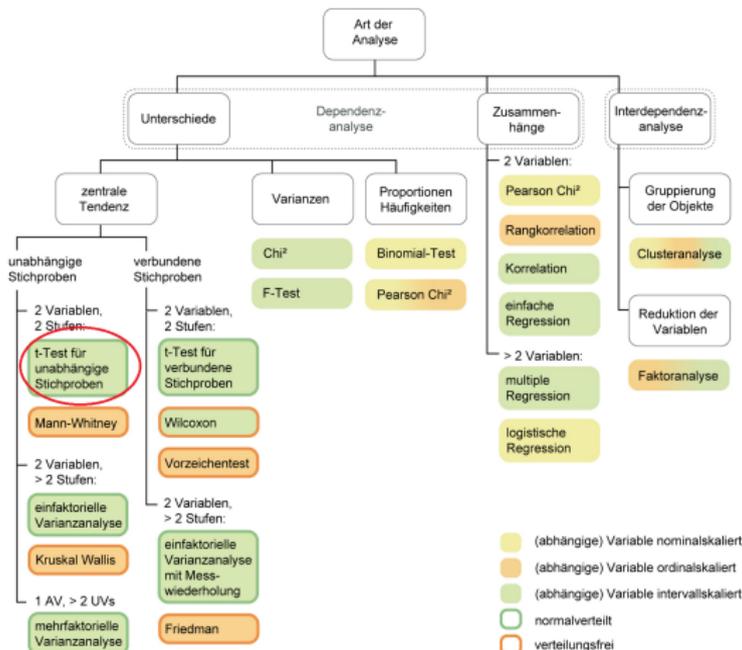
- Zuerst mit den 'einfachen Sachen' beginnen. (Deskriptive Statistik: Darstellung und Merkmale der Daten).
 - ✓ SuS bekommen ein erstes Erfolgserlebnis und motivieren sich dann für komplexere Analysen.
 - ✓ Sind SuS mit 'einfachen' Analysen schon überfordert, dann rate ich davon ab, beurteilende Statistik zu machen. Die Zeit ist viel besser in einer klaren und qualitativ hochwertigen Darstellung der Daten investiert.
 - ✓ SuS bekommen ein 'Gefühl' für den Datensatz. Sie merken oft selber, welche Fragen am interessantesten sind.
- Erst danach eine beschränkte Auswahl an Fragen mittels beurteilender Statistik untersuchen.

Beispiel t-Test

- Arbeit von Inaan Shah (2021), Körperbild und Körperzufriedenheit.
- Gemessene Werte: Körperzufriedenheit und Geschlecht von Jugendlichen.

```
frau <- c(33,45,88,78,...  
mann <- c(48,62,92,83,...
```
- Fragestellung: Ist die Körperzufriedenheit der Männer unterschiedlich als die der Frauen?
- Modellwahl: 2 unabhängige Stichproben, metrischer Daten, (normalverteilt).
- Art der Analyse: Unterschied der Erwartungswerte (zentrale Tendenz).

Beispiel t-Test



Beispiel t-Test mit R

- Analyse mit R, Methode: t-test.

```
> t.test(mann,frau,alternative = "two.sided",var.equal = TRUE)
```

Two Sample t-test

data: mann and frau

t = 1.585, df = 136, p-value = 0.1153

- Ergebnisse:

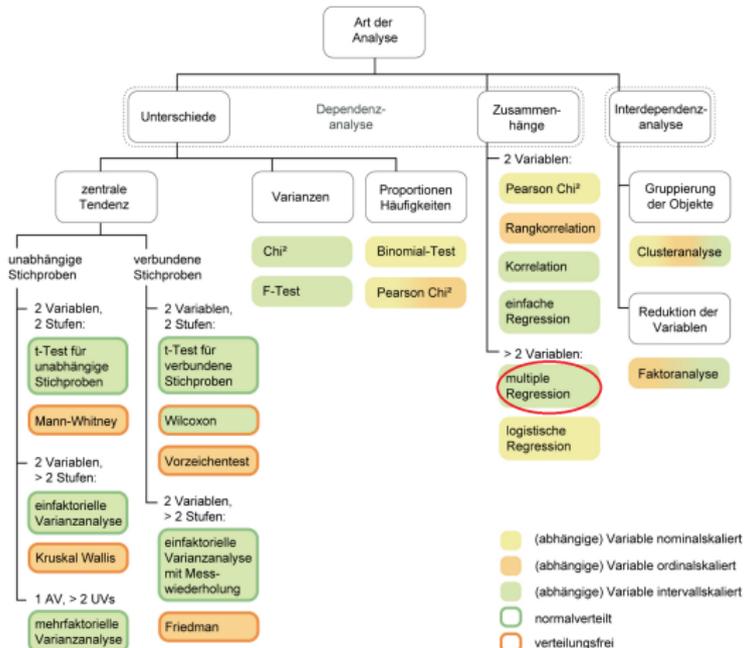
- ✓ Es wurde kein signifikanter Unterschied ($p > 5\%$) der Zufriedenheit zwischen Frauen und Männer festgestellt. (Das heisst nicht, dass es keinen gibt!)

- *Vertiefung: Modellannahmen können mittels eines QQ-Plot und levene-Test überprüft werden.*

Beispiel lineare Regression

- Arbeit von Géraldine Schmid (2021), Vegane Ei-Alternativen beim Backen.
- In einer Umfrage wurde gefragt, wie oft pro Woche die Versuchsperson vegane Mahlzeiten isst. (Abhängige Variable Y)
- Fragestellung: Haben Alter, Geschlecht, Einkommen und Ausbildung (unabhängige Variablen) einen Einfluss auf die Anzahl vegane Mahlzeiten? Welcher?
- Modellwahl: Alle Variablen sind metrisch.
- Art der Analyse: Es wird einen (linearen) Zusammenhang vermutet. → Multiple lineare Regression.
- $Y_i = \beta_0 + \beta_1 x_i^{(A)} + \beta_2 x_i^{(G)} + \beta_3 x_i^{(E)} + \beta_4 x_i^{(T)} + E_i$
- *Vertiefung: Modellannahme mit Residuenanalyse (d.h. Tukey-Anscombe und QQ-Plots) überprüfen.*

Beispiel lineare Regression



Beispiel lineare Regression mit R

- Multiple lineare Regression mit R:

```
> library(readxl)
> da <- read_excel("./Daten/Datensatz_Vegan.xlsx")
> fit <- lm(ProWoche~Alter + Geschlecht + Einkommen + TeuBil, data=da)
> coef(summary(fit))
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	5.00323	0.42359	11.812	5.08e-25
Alter	-0.03228	0.00728	-4.433	1.50e-05
Geschlecht	0.12609	0.12894	0.978	3.29e-01
Einkommen	0.00453	0.00196	2.308	2.20e-02
TeuBil	-1.27428	0.44475	-2.865	4.59e-03

✓ Alter, Bildung (und Einkommen) sind signifikant.

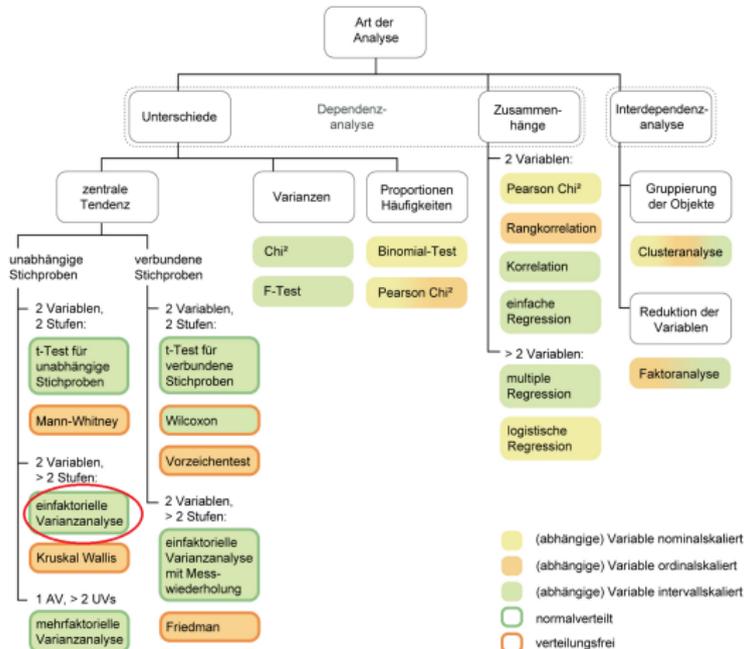
- Gefittetes Modell:

$$Y = 5 - 0.032x^{(A)} + 0.13x^{(G)} + 0.0045x^{(E)} - 1.27x^{(T)}$$

Beispiel ANOVA

- Arbeit von Géraldine Schmid (2021), Vegane Ei-Alternativen beim Backen. (Fortsetzung)
- Es wurden Muffins nach 5 unterschiedlichen Rezepten (1x mit Ei und 4x mit verschiedenen veganen Ei-Alternativen) gekocht.
- Backzeit, Temperatur, andere Zutaten, usw bleiben konstant während des ganzen Experiments.
- Abhängige Variable: Gewichtsverlust der Muffins in den ersten 20 Minuten nach dem Backen.
- Fragestellung: Ist der Gewichtsverlust vom Rezept abhängig?
- Modellwahl: Abhängige Variable ist metrisch. Unabhängige Variable (UV) ist **nominalskaliert** (5 Stufen).
- Art der Analyse: Unterschied des Erwartungswerts. UV mit mehreren Stufen. → Es eignet sich eine ANOVA.

Beispiel ANOVA



Beispiel ANOVA mit R

```
> da <- read_excel("./Daten/Daten_Feuchtigkeit.xlsx")
> da$Rezept <- as.factor(da$Rezept) # Wichtig!
> fit <- aov(Feuchtigkeitsverlust~Rezept, data=da)
> summary(fit)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Rezept	4	41.6	10.4	51.5	<2e-16 ***
Residuals	48	9.7	0.2		

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- ✓ Der Faktor Rezept ist signifikant, da $p < 2e - 16 \ll 0.05$.
- ✓ Zur Darstellung nominalskalierten Variablen eignet sich oft ein Boxplot. (s. Unterlagen).
- *Vertiefung: Um zu wissen, welche Rezepte sich voneinander unterscheiden, kann man paarweise t-Tests durchführen (s. Unterlagen).*
 - ⚠ *Dabei muss auf das "Multiple testing Problem" geachtet werden. (s. Unterlagen)*

Take home message

- Planung und Durchführung eines Experiments (Datenerhebung) müssen strenge wissenschaftliche Kriterien und Qualitätsansprüche erfüllen.
- Die Modellwahl ist z.T. sehr schwierig, aber sehr wichtig. Sie bestimmt die Art der Analyse. Online Tools können bei der Modellwahl unterstützen.
- Zuerst mit der beschreibenden Statistik anfangen.
- Mit dem Statistikprogramm R lässt sich jede Art von Datensätze analysieren.
- T-tests, Regressionen und ANOVAs gehören zu den häufigsten Modellen für Datensätze in Maturaarbeiten.

- 1 Einführung
- 2 Datenerhebung
- 3 Modellwahl
- 4 Datenanalyse
- 5 Referenzen**
 - Unterlagen
 - Nützliche Links
 - Literatur
- 6 *Vertiefung

Unterlagen

Unter folgendem Link finden Sie alle Unterlagen.

- ✓ Vortrag
- ✓ **R-Skript** mit Datensätzen
- ✓ **Merkblatt für SuS** "Schritt für Schritt durch die Maturaarbeit: Von der Fragestellung zur Statistik"
- ✓ Erklärung der Residuenanalyse.



<https://polybox.ethz.ch/index.php/s/F7saNM0FRtmju3D>

Links

- R und RStudio herunterladen
<https://posit.co/download/rstudio-desktop/>
- Einführungen in R
<https://www.cyclismo.org/tutorial/R/>
Sauer (2019)
- Wahl des Modells und der statistischen Analyse
<https://datatab.de/>
https://www.methodenberatung.uzh.ch/de/datenanalyse_spss.html
<https://www.statstest.com/>

Literatur I

Frank Harrell. *Regression Modeling Strategies*. Springer, 2015.

Jonas Himmelberger. *Handys und Hirne*. 2023.

Lukas Meier. *Wahrscheinlichkeitsrechnung und Statistik: Eine Einführung für Verständnis, Intuition und Überblick*. Springer Berlin Heidelberg, 2020. ISBN 9783662614877. URL

<https://people.math.ethz.ch/~meier/teaching/book-intro/>.

Lukas Meier. *ANOVA and Mixed Models: A Short Introduction Using R*. Chapman & Hall/CRC The R Series. CRC Press, 2022. ISBN 9781000776775. URL

<https://people.math.ethz.ch/~meier/teaching/anova/>.

Patric Müller. *Lineare Regression, WBL-Notizen*. 2023. URL

<https://polybox.ethz.ch/index.php/s/1aG9NyPTAVb5JA6>.

Myra L Samuels, Jeffrey A Witmer, and Andrew Schaffner. *Statistics for the life sciences*. Pearson Education, 2012.

S. Sauer. *Moderne Datenanalyse mit R: Daten einlesen, aufbereiten, visualisieren, modellieren und kommunizieren*. FOM-Edition. Springer Fachmedien Wiesbaden, 2019. ISBN 9783658215873.

Literatur II

Géraldine Schmid. *Vegane Ei-Alternativen beim Backen*. 2021.

Inaan Shah. *Körperbild und Körperzufriedenheit. Eine Untersuchung bei Jugendlichen an der Kantonsschule Freudenberg*. 2021.

Werner Alfred Stahel. *Statistische Datenanalyse*. Vieweg und Sohn, Braunschweig, 2002.

Danke für die Aufmerksamkeit

1 Einführung

2 Datenerhebung

3 Modellwahl

4 Datenanalyse

5 Referenzen

6 *Vertiefung

- Interpretationsfehler
- Modellannahmen überprüfen
- R und Rstudio
- Multiple testing problem
- Paradebeispiel mixed ANOVA

Klassische Interpretationsfehler, die es zu vermeiden gilt

- ⚠ Deskriptive Statistik erlaubt keine Schlussfolgerungen.
 - ✗ Falls Medi A im Schnitt (Mittelwert der Stichprobe) 10 Stunden wirkt und Medi B nur 8, kann man nicht sagen "Medi A wirkt länger als Medi B".
 - ✓ Eine solche Aussage ist nur möglich nach einem entsprechenden signifikanten Testergebnis.
- ⚠ Ein negativer Test lässt keine Schlüsse zu. Keine Signifikanz bedeutet nicht, dass H_0 korrekt ist.
- ⚠ Keine Daten manipulieren, wie z.B. unerwünschte Messungen willkürlich und unbegründet löschen.
- ⚠ Resultate nicht ausser Kontext anwenden.
 - ▶ Aus der Studie: mit 100mg Vitamin C erkrankt man 10% weniger, mit 50mg 5% weniger (linearer Zusammenhang).
 - ✓ Zulässige Aussage: mit 70mg Vitamin C erkrankt man 7% weniger.
 - ✗ Unzulässige Aussage: mit 1000mg erkrankt man 100% weniger. Vorhersage ist zu weit weg von den gemessenen Daten!

Modellannahmen überprüfen

- Viele Modellannahmen werden mit Hilfe von statistischen Tests nach ihrer Plausibilität untersucht.
- Modellannahmen können deswegen nie bewiesen werden, sondern nur widerlegt!
- Nicht widerlegte Modellannahmen gelten als vertretbar.
- Beispiele von typischen Modellannahmen und dazu passende Arten sie zu überprüfen.
 - ▶ Annahme: Die Varianz der zwei Stichproben ist die gleiche → **levene Test**
 - ▶ Annahme: Fehler sind i.i.d. (identical independent distributed). → Residuenanalyse (Tukey-Anscombe plot). Siehe Müller (2023) und Meier (2020)
 - ▶ Annahme: Fehler sind normalverteilt → qqPlot
 - ▶ Annahme: Keine systematische Fehler → kann kaum überprüft werden.

R, Rstudio und Alternativen

- R ('sprache') und RStudio (Programm) herunterladen und installieren:
<https://posit.co/download/rstudio-desktop/>
- ✓ Das 'coden' in RStudio kann schnell und sektoriell gelernt werden. Programmierkenntnisse sind nicht notwendig (können aber der Lernprozess beschleunigen).
- ✓ Daten aus Excel (.xlsx / .csv) können schnell und mühelos importiert werden.
- ✓ Basiskenntnisse in R können im Studium/Arbeitswelt wertvoll sein.
 - Excel deckt folgende Bereiche der Statistik: Beschreibende Statistik, t-tests, 1,2-way ANOVA, einfache lineare Regression.
 - SPSS ist eine (kostenpflichtige) Alternative zu R.

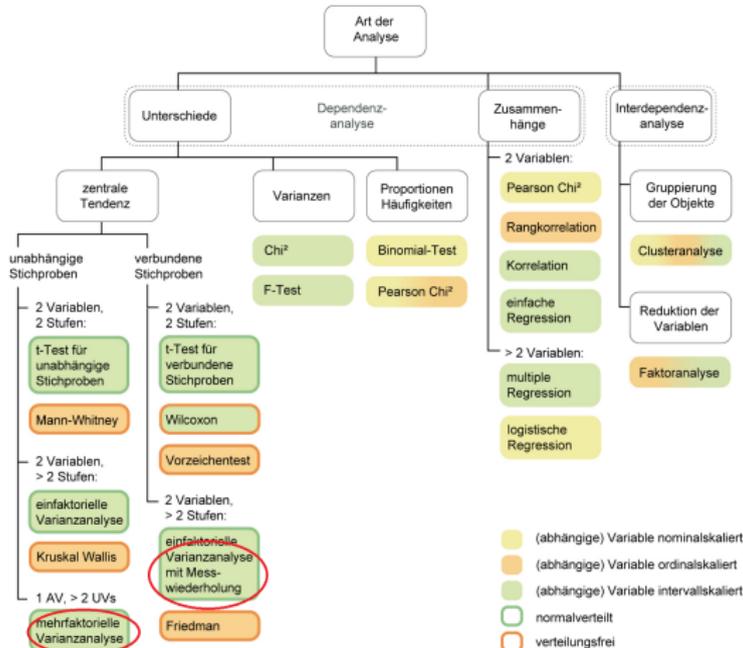
multiple testing problem

- Das Führen von mehreren Tests ist mathematisch problematisch...
- ... Per Definition hat ein Test eine Wahrscheinlichkeit $\alpha = 5\%$ eine nicht existierende Signifikanz zu finden (**False Positive**). Anders gesagt: Ist H_0 korrekt, hat man mit Wahrscheinlichkeit $\alpha = 5\%$ 'eine schlechte Stichprobe' und somit ein positives Testergebnis (H_A).
- Werden mehrere Tests durchgeführt, erhöht sich die Chance mindestens einen False Positive zu bekommen.
- Das Problem wird umgegangen in dem das Signifikanzniveau (korrigiert) verkleinert wird.
 - ▶ Bonferroni-Korrektur: Niveau für einen einzelnen Test wird durch N (Anzahl Tests) dividiert. Signifikant sind nur die Tests mit $p\text{-Wert} < 0.05/N$.

mixed ANOVA

- Arbeit von Jonas Himmelberger (2023), Handys und Hirne.
 - Jeder Proband muss 20 Objekte merken. 20 Minuten Pause. Proband muss die Objekte auflisten.
 - Der Test hat drei Varianten, einmal herrscht in der Pause Ruhe, einmal werden Medien passiv konsumiert und einmal aktiv.
 - Jeder Proband führt alle 3 Varianten durch. (An drei unterschiedlichen Tagen).
 - Unabhängige Variablen sind: Alter (jung/erwachsen), Medienkonsum (wenig/viel), soz. Medienkonsum (2 Niveaus) und die Testvariante.
 - Die abhängige Variable sind die Testergebnisse (metrisch).
- ⚠ Modellwahl: Hier haben wir wiederholte Messungen, da Probanden mehrmals getestet werden.
- Art der Analyse: mixed ANOVA.

Beispiel mixed ANOVA



mixed ANOVA mit R

- UV Test ist '**within**', denn bei der selben Person variiert der Test.
- Alter, Medien und soz. Medien sind '**between**' UV, denn sie bleiben während der ganzen Studie für die Probanden konstant.

```
> fit1 <- ezANOVA(data=da, dv=Resultat1, wid = Person, within = Test, between = c(A
> summary(fit1$aov)
```

```
Error: Person
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Alter	1	32.5	32.46	1.451	0.234
Medienkonsum	1	0.4	0.36	0.016	0.900
socmed	1	0.3	0.27	0.012	0.913
Alter:Medienkonsum	1	1.9	1.95	0.087	0.769
Alter:socmed	1	44.0	43.96	1.966	0.167
Medienkonsum:socmed	1	0.6	0.64	0.029	0.866
Alter:Medienkonsum:socmed	1	9.4	9.45	0.423	0.519
Residuals	47	1051.0	22.36		

```
Error: Person:Test
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Test	2	79.7	39.86	9.433	0.000185 ***
Alter:Test	2	23.3	11.65	2.756	0.068671 .
[...]					

- ✓ Faktor Test (Pausenaktivität) ist signifikant.