

Übung 04

Hypothesen testen

Korrelation, Wilcoxon-test und χ^2 in R

INFI-IS

5xHWII

Albert Greinöcker

November 15, 2021



CCA - COMPETENCE CENTRE

HTL Anichstraße

1 Hypothesen testen

So wie im Student Performance - Datensatz besprochen soll nun ein etwas größerer Fragebogen kennengelernt und teilweise ausgeweitet werden. Ein Download des Datensatzes plus einem PDF in dem genaue Beschreibungen zu den Variablen zu finden sind, befindet sich im Moodle unter: european social survey.

1.1 Importieren des Datensatzes

Die entsprechende Datei entpacken und die Datei ESS8e02.1_F1.csv so wie im gemeinsamen Skript zum European Social Survey beschrieben, importieren.

```
1 d <- read.csv("ESS8e02.1_F1.csv", sep=";", encoding="UTF-8")
```

Im Datensatz sind die einzelnen Variablen in Zahlen kodiert (so wie es sich gehört). Will man diesen Zahlen Beschriftungen zuordnen, macht man das auf diese Art:

```
1 d$gndr <- factor(d$gndr, levels=c(1,2,9), labels = c("Male", "Female", "No Answer"))
2 d$vote <- factor(d$vote, levels=c(1,2,3,7,8,9), labels = c("Yes", "No", "Not eligible to vote", "Refusal", "Don't know", "No answer"))
```

Listing 1: Hier wird z.B. der Variable gndr (Geschlecht) die Werte "Male", "Female", "No Answer" für die Zahlen 1, 2, 9 zugeordnet.

In einzelnen Auswertungen wird der Vergleich zwischen 2 Ländern oder die Situation in nur einem Land untersucht. Aus diesem Grund müssen Subsets mit nur diesen Ländern gebildet werden. Das geht z.B. so:

```
1 d_at <- d[d$cntry == 'AT',]
2 d_at_it <- d[d$cntry %in% c('AT', 'IT'),]
```

Listing 2: Hier wird 1x ein Datensatz aller Österreicher und 1x ein Datensatz mit Österreichern und Italienern erzeugt.

Leider kann es dann sein, wenn - wie im Fall der Länder - schon Beschriftungen automatisch zugeordnet wurden, dass unbenutzte Labels (hier: die Länderkürzel für nicht mehr vorhandene Länder) in der Auswertung auftauchen. Diese kann man so löschen:

```
1 d_at_it$cntry <- droplevels(d_at_it$cntry)
```

1.2 Hypothesen Testen

In den eigentlichen Aufgabenstellungen sollen bestimmte Hypothesen getestet werden. Es soll bei jeder Hypothese...

- a. eine Aufbereitung der Daten gemacht werden,
- b. eine grafische Veranschaulichung (sofern möglich) gemacht werden,
- c. ein passender Hypothesentest ausgewählt und ausgeführt werden und
- d. eine Interpretation der Ergebnisse erstellt werden.

Hier die Hypothesen (in Klammern die betroffenen Variablen):

- a. Männer haben mehr Glauben in die Polizei (`rstplc`, `gender`).
- b. Es besteht ein negativer Zusammenhang bei "mehr Strom aus nuklearer Energie" und "mehr Strom aus Solarenergie" (`elgnuc`, `elgsun`).
- c. In Österreich ist der Eindruck, dass sich der Klimawandel schlecht auf die Menschen auswirkt, stärker als in Ungarn (`ccgdbd`).
- d. Frauen stimmen einem bedingungslosen Grundeinkommen eher zu (`basinc`).

Bitte noch 2 Hypothesen selbst wählen und überprüfen!