4 Introduction to regression: Assignment

Instructions

Submit your answers to the 6 first exercises through Aula Global (quiz section **in your seminar**). Upload the remaining exercise concerning peer review through Aula Global as well (upload task **in grup gran**).

Regression

Consider the model `pitch ~ 1 + gender`, fit on the data from Franke & Roettger (2019).

```
df <- read.csv("https://tinyurl.com/polite-data") #download data

model <- lm(formula = pitch ~ 1 + gender, #fit model
            data = df)

#N.B.: this is equivalent to:
#model <- lm (df$pitch ~ 1 + df$gender)
```

1. Is this model equivalent to `pitch ~ gender` or `pitch ~ gender + context`?
2. Is `pitch` the outcome, a predictor, or neither?
3. Is `gender` the outcome, a predictor, or neither?
4. What is the predicted `pitch` of a `M`-gender speaker? Round your answer to the first decimal after the comma (if there are decimals)
5. What is the predicted `pitch` of a `F`-gender speaker? Round your answer to the first decimal after the comma (if there are decimals)
6. What is the estimated difference between `M`- and `F`-gendered speakers? Round your answer to the first decimal after the comma (if there are decimals)

Peer review (part I)

Submit answers to these questions concerning your analysis report as a **single page** PDF-file through Aula Global. This submission makes up 50% of your peer-review grade. Remember that **your submission must be anonymous**. That is: do not put any information that can identify you in the PDF nor in the name of the file (like your name).

1. What is your general research question?
2. Why do you think this question is interesting? What does an answer to it tell us?
3. What is your specific research question?
4. What kind of data would you use to address Question 2 if you had unlimited resources?
5. What kind of data are you planning to use to address Question 2 within the scope of this class?
    - How will you obtain it?
    - How much will you collect?
    - Do you think that is enough data to address Question 2? Why (not)?

Elaborate with as much detail as you (currently) have. It is perfectly acceptable if you answer in bullet points and single sentences. A randomly selected peer will give you feedback on them

Self-study

Read Winter (2013), pages 1-10: https://arxiv.org/ftp/arxiv/papers/1308/1308.5499.pdf. If you find this interesting, read on until page 21. Pages 10-21 cover topics we will briefly touch on but will not be able to discuss in detail. If you're still interested, read the entire tutorial. The second part of the tutorial touches on a very powerful and useful extension of regression models that we won't cover in this class.

References

Franke, Michael, and Timo Benjamin Roettger. 2019. "Bayesian Regression Modeling (for Factorial Designs): A Tutorial," July. https://doi.org/10.31234/osf.io/cdxv3.

Winter, Bodo. 2013. "Linear Models and Linear Mixed Effects Models in R with Linguistic Applications." *CoRR* abs/1308.5499. http://arxiv.org/abs/1308.5499.