4 Introduction to regression: Assignment

Instructions

Submit your answers to the 6 first exercises through Aula Global.

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)

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.

Use the remainder of your time to start working on your final report.

References

Franke, Michael, and Timo Benjamin Roettger. 2019. "Bayesian Regression Modeling (for Factorial Designs): A Tutorial," July. Center for Open Science. 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.