Regression for Linguists

WiSe23/24

Daniela Palleschi

Lecture 07/09/23

Table of contents

Course overview

Moodle

- lecture materials

Kursübersicht

- Leistungspunkte
  - 3LP
    * 1LP:
    * 1LP:
    * 1LP:
- Office hours: Wednesdays, 15.00-16.00 (by appointment)

Course aims

By the end of this course, you will

- blah blah

What will you learn?

- linear regression
- multiple regression
- logistic regression
- mixed models
- using the `lme4` package
- how to apply these models appropriately to a variety of data types

What will you *not* learn?

- stuff

Part I

Overview

Syllabus

```
Wrote 0 references to './references.bib'
```

```
v Reading from "Regression for Linguists WiSe23/24".
```

```
v Range 'Sheet1'.
```

Meeting	Lecture	Topic	Vorbereitung
2023-10-10	1	Equation of a line	Winter (2019): Ch. 1-3
2023-10-11	2	Linear regression	Winter (2019): Ch. 4 Winter (2013)
2023-10-12	3	Continuous predictors	Winter (2019): Ch. 5 Winter (2013)
2023-10-10	4	Multiple linear regression	Winter (2019): Ch. 6 Winter (2013)
2023-10-11	5	Categorical predictors	Winter (2019): Ch. 7 Winter (2013)
2023-10-12	6	Model assumptions	
2023-10-10	7	Logistic regression	Winter (2019): Ch. 12
2023-10-11	8	Log odds, logits, and odds ratio	
2023-10-12	9	Foundational Ideas	Vasishth & Nicenboim (2016)
2024-01-12	10	Linear mixed models	Winter (2019): Ch. 14 Winter & Grice (2021); unt
2024-01-12	11	Linear mixed models	
2024-01-26	12		
2024-01-26	13		
2024-02-09	14		
2024-02-09	15		

Resources and Set-up

Error: not found: Winter_2013, Winter_2014, sondregger_regression_2023, baayen_2008, jaeger_1

Resources

This course is mainly based on (**winter_statistics_2019?**), which is an excellent introduction into regression for linguists. For even more introductory tutorials, I recommend going through (**Winter_2013?**) and (**Winter_2014?**). For a more intermediate textbook, I'd recommend (**sondregger_regression_2023?**).

If you're interested in the foundational writings on the topic of linear mixed models in (psycho)linguistic research, I'd recommend reading (**baayen_2008?**); (**jaeger_2008?**); (**barr_2013?**); (**matschucek_2017?**).

Assumptions about you

For this course, I assume that you are familiar with more classical statistical tests, such as the t-test, Chi-square test, etc. I also assume you are familiar with measures of central tendency (mean, median, mode) measures dispersion/spread (standard deviation), and with the concept of a normal distribution. Lacking this knowledge will not impeded your progress in the course, but is an important foundation on which we'll be building. We can review these concepts in-class as needed.

Software

- R: a statistical programming language (the underlying language)
- RStudio: an program that facilitates working with R; our preferred IDE integrated development environment
- LaTeX: a typesetting system that generates documents in PDF format
- why R?
  - R and RStudio are open-source and free software
  - they are widely used in science and business

Install R

- we need the free and open source statistical software R to analyze our data
- download and install R: https://www.r-project.org

Install RStudio

- we need RStudio to work with R more easily
- Download and install RStudio: https://rstudio.com
- it can be helpful to keep English as language in RStudio
  - we will find more helpful information if we search error messages in English on the internet
- If you have problems installing R or RStudio, check out this help page (in German): http://methods-berlin.com/wp-content/uploads/Installation.html

Install LaTeX

- we will not work with LaTeX directly, but it is needed in the background
- Download and install LaTeX: https://www.latex-project.org/get/

resources

- many aspects of this course are inspired by (**nordmann_applied_2022?**) and (**wickham_r_nodate?**)
  - both freely available online (in English)
- for German-language resources, visit the website of Methodengruppe Berlin

Troubleshooting (EN: Troubleshooting)

- Error messages are very common in programming, at all levels.
- How to find solutions for these error messages is an art in itself
- Google is your friend! If possible, google in English to get more information

Literaturverzeichnis

Part II

Part I: Foundations

Part III

Part II: Mixed models

Part IV

Reports

References