

# Introduction to Generalized Linear Models

General course information

Filippo Gambarota

University of Padova

2022/2023

# About me

- I'm a post-doc researcher at the Department of Developmental Psychology and Socialization
- I did a PhD in Experimental Psychology studying the unconscious working memory processing
- I work with Professor Gianmarco Altoè on data analysis in Psychology, especially meta-analysis

# Office hours

I do not have official office hours but we can schedule an appointment.  
you can write me at **[filippo.gambarota@unipd.it](mailto:filippo.gambarota@unipd.it)**:

- my office is the 027, first floor Psico 1 building (pink building)
- we can also schedule on Zoom



We are going to spend the following hours together:

Day	Date	Time	Room
Wednesday	26/04/2023	10:30-12:30	3L
Thursday	27/04/2023	10:30-12:30	3L
Tuesday	02/05/2023	10:30-12:30	3L
Wednesday	03/05/2023	10:30-12:30	3L
Thursday	04/05/2023	10:30-12:30	3L
Tuesday	09/05/2023	10:30-12:30	3L
Wednesday	10/05/2023	10:30-12:30	3L
Thursday	11/05/2023	10:30-12:30	3L
Tuesday	16/05/2023	10:30-12:30	3L
Wednesday	17/05/2023	10:30-12:30	3L
Thursday	18/05/2023	10:30-12:30	3L

# Materials

The slides will be structured intermixing:

- R code
- Theory and Formulas (not a lot :) )
- Plots (a lot!)
- Examples and exercises

All the slides and extra materials are available on Moodle. During the course and for the exercises we will use some **custom functions** that I wrote for the course. To download and use the function you can download the `utils-glm.R` file from Moodle or from here

**<https://stat-teaching.github.io/SMDA-2023/R/utils-glm.R>**

# Materials

- slides with the `#extra` tag are very specific but useful topics that will be eventually covered but are not part of the core course/exam

The course materials are created using R (4.3.0). To organize the materials I used **R Projects**, a feature available with R Studio. I highly suggest you to use R Projects that significantly improves your workflow. A brief tutorial is available here <https://r4ds.had.co.nz/workflow-projects.html>.



# R code

I mainly use R for my daily work but a deep understanding of R is not necessary.

Slides are created with R Markdown (`.rmd` files) and distributed in pdf and all source scripts are available.

I wrote several custom functions that are used in the slide and maybe during the exercises.

# R Packages

We will use, directly or indirectly (for custom functions), several R packages. You can easily install all packages using:

```
pkgs <- c("effects", "tidyverse", "MASS",  
          "broom", "reshape2", "ggeffects",  
          "performance", "see", "car", "devtools")  
install.packages(pkgs)
```

# Pipes

Sometimes in my code you will see a symbol like this `|>`, this is called **pipe**. Sometimes it is also written as `%>%` that is a different pipe coming from the `magrittr` package. The pipe is a very simple way to write R code when you need to apply multiple functions in succession. Practically, the pipe apply a function to an element, for example:

```
# these two are the same  
mean(x)
```

```
# to x apply the mean function  
x |> mean()
```

# Pipes

Beyond this silly example, when we need to use multiple nested functions the pipe makes the code more readable:

```
x <- runif(10)
```

```
# without pipe
```

```
exp(min(round(x, 2)))
```

```
## [1] 1.105171
```

```
# with pipe
```

```
x |>
```

```
  round(2) |>
```

```
  min() |>
```

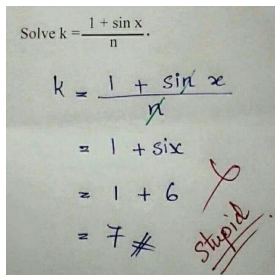
```
  exp()
```

```
## [1] 1.105171
```

# Theory and Formulas

I tried to reduce the amount of formulas. I prefer to make practical examples and showing the R code.

The probability of making an error or typo is close to 1 (Shepard, 2023), if you find something strange raise the hand or write me an email :)



A photograph of a piece of paper with handwritten mathematical work. At the top, it says 'Solve  $k = \frac{1 + \sin x}{n}$ .' Below this, the student has written several lines of work, each with a correction or cancellation. The first line is  $k = \frac{1 + \sin x}{n}$ , where the  $n$  in the denominator is crossed out with a green checkmark. The second line is  $= 1 + \sin x$ . The third line is  $= 1 + 6$ . The fourth line is  $= 7$ , followed by a crossed-out  $7$  and the word 'Stupid' written in red ink and underlined twice.

$$\begin{aligned}\text{Solve } k &= \frac{1 + \sin x}{n} \\ k &= \frac{1 + \sin x}{\cancel{n}} \\ &= 1 + \sin x \\ &= 1 + 6 \\ &= 7 \quad \cancel{7} \quad \text{Stupid}\end{aligned}$$

# Examples and exercises

I tried to make practical examples whenever possible and relevant.  
Furthermore we will see some exercises and case studies.

# Rules of the game

1. **Participate!** If you have questions, doubts, comments, etc. please ask
2. If something is **not clear or is discordant** with other information from previous courses, tell me.
3. **Participate!**
4. Try do to **exercises and case studies**
5. **Participate!**
6. If you can, **bring your laptop with R**

# Prepare the course folder #optional


- Install the packages of slide *R Packages*
- Download and extract the **template course folder** from here
- Open the SMDA-2023.Rproj file
- Now the command `devtools::load_all()` will load each function within the R/ folder. You can consider this folder as an R Package



# Final note

This is my first teaching experience with this course. Suggestions and critique are welcome.



 [filippo.gambarota@unipd.it](mailto:filippo.gambarota@unipd.it)

 [github.com/filippogambarota](https://github.com/filippogambarota)