Method 2
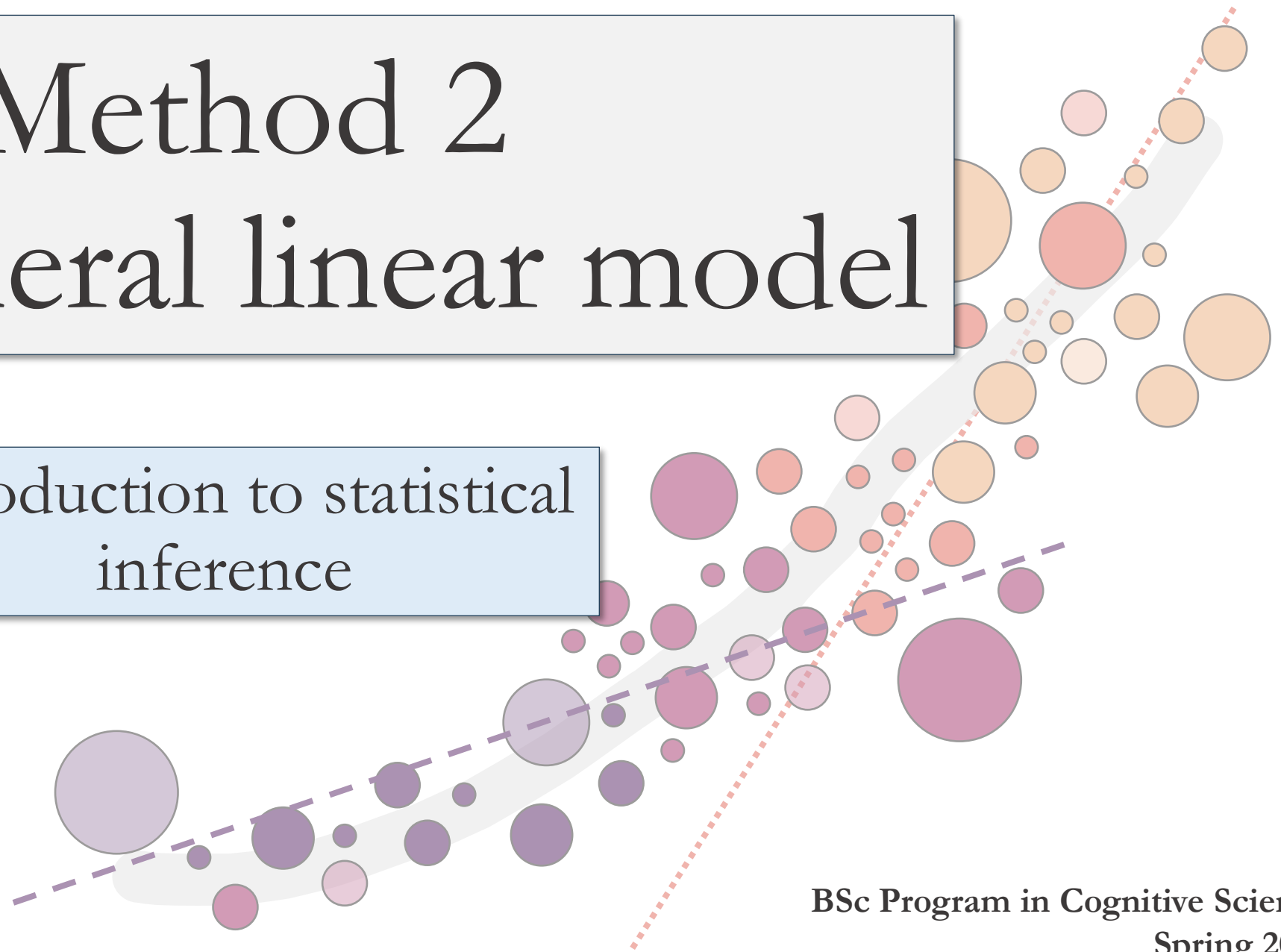# The general linear model

Introduction to statistical inference

BSc Program in Cognitive Science

Spring 2024

Nicolas Legrand – Researcher– ILAB (Interacting Minds Centre)

AARHUS UNIVERSITY

Who am I?

- ## PhD in Psychology
  2014 - 2018 - Université de Caen (France)
  - *The consolidation and suppression of individual and collective memory.*
  EEG, machine learning, ECG, Python, memory suppression, fMRI, collective memory, dreams, memory schemas

- ## Master in Neurobiology and behaviors
  2013 - 2014 - Université de Caen (France)
  - *Sleep-dependent prospective memory consolidation and dreaming*

- ## Master in Cognitive Science
  2011 - 2013 - Université Lumière Lyon 2 (France)
  - *What is phenomenal consciousness?*
  - *The neuropsychoanalysis of dreaming.*

- ## Bachelor in Mathematics and computer sciences
  2009 - 2011 - Université de Grenoble 2 (France)

- ## Bachelor in Philosophy
 **2008 - 2011** - Université de Grenoble 2 (France)

Dec. 2022 – Curr.

Researcher at IMC (Ilab, sup. Chris Mathys) in computational psychiatry. Using Bayesian nonparametric methods to create models of delusions. Developing a new neural network library for predictive coding (pyhgf).

Jul. 2019 – Dec. 2022

Postdoctoral fellow in computational psychiatry at CFIN (Embodied Computation Group, sup. Micah Allen). Creating new methods to measure cardiac interoception. Developing Python toolboxes for signal processing (Systole), psychophysic tasks (Cardioception), and metacognition modeling (metadpy).
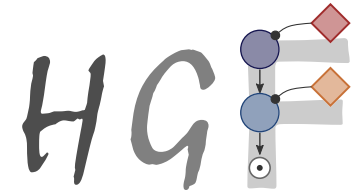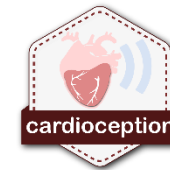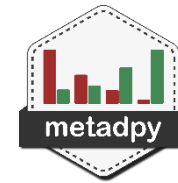
Python is my main programming language. I also have several years of experience using and teaching R and Matlab.

Overview

Topics

- The General Linear Model

- Regression modelling

- Mathematical foundations

- Linear algebra (vectors, matrices, determinants, eigen-analysis,...)

- Calculus (infinite series, derivatives, integrals,...)

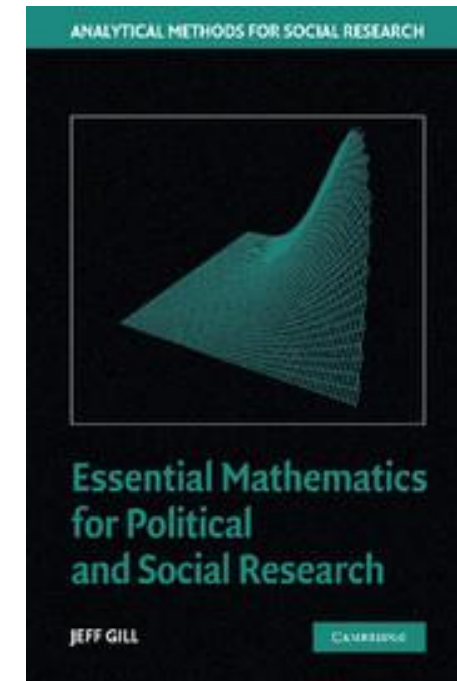- Generalized Linear Models (e.g., logistic regression)

AARHUS UNIVERSITY

Resources
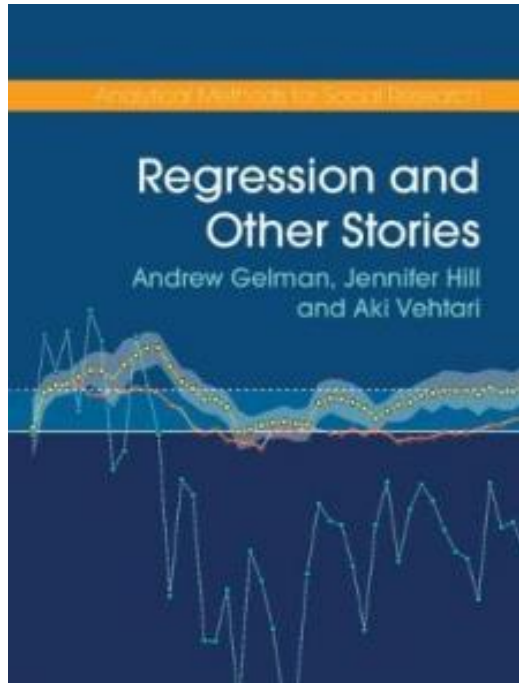
Textbook:
Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and Other Stories* (Analytical
Methods for Social Research). Cambridge: Cambridge University
Press. doi:10.1017/9781139161879
Please get a copy!
Free PDF: https://avehtari.github.io/ROS-Examples/

Textbook:
Gill, J. (2006). Essential Mathematics for Political and Social
Research (Analytical
Methods for Social Research). Cambridge: Cambridge University
Press. doi:10.1017/CBO9780511606656
No need to buy it! You have access to PDFs of the chapters via
the Royal Library.

Resources

Code:
• This course's repository: https://github.com/methods-2-f24/

• All code and data in the book: https://github.com/avehtari/ROS-Examples

• **Please get a free GitHub account**

Videos:
• This course is on YouTube!
https://www.youtube.com/playlist?list=PLvJwKACYy5_MTdnrzxx_1sN389dS9OB3S

• Order slightly different: we'll do GHV chapters 1-5 in the first 3 weeks, but after that, you can watch the videos in the order of the playlist

Schedule

Course week	Week of the year	Topics and readings
1	7	Regression and the GLM: overview, data and measurement, (GHV[1]1,2)
2	8	Basic methods, statistical inference (GHV 3,4)
3	9	Statistical inference (continued), simulation (GHV 4,5)
4	10	Math basics: functions, equations, polynomials, logarithms (Gill[2]1)
5	11	Linear algebra basics: vectors, matrices, norms, transposition (Gill 3)
6	12	More linear algebra: geometry, determinants, rank, inversion, eigenvectors (Gill 4)
7	15	Scalar calculus: derivatives, integrals, fundamental theorem (Gill 5)
8	16	More calculus: root finding, extrema, Lagrange multipliers, vector calculus (Gill 6)
9	17	Conceptual foundations and history of the GLM, model fitting (GHV 6,7,8)
10	18	Fitting GLMs: prediction, Bayesian inference (GHV 9)
11	19	Multiple predictors, interactions (GHV 10)
12	20	Model comparison, assumptions and diagnostics (GHV 11)
13	21	Transformations, predictive simulations (GHV 12) [no class, just lecture]

Introduction to statistical inference

▪ **Portfolio 1**

Mathematical foundations

▪ **Portfolio 2**

Generalized Linear Models (GLM)

▪ **Portfolio 3**

1 - Gelman, A., Hill, J., & Vehtari, A. (2020). Regression and Other Stories (Analytical Methods for Social Research). Cambridge: Cambridge University Press. doi:10.1017/9781139161879

2 - Gill, J. (2006). Essential Mathematics for Political and Social Research (Analytical Methods for Social Research). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511606656

Exam

- Portfolio consisting of 3 assignments

- Each assignment will require you to create an R Markdown notebook consisting of a mix of text and code.

- Due:
    1. End of week 9 (Sunday 3 March, 23:59)
    2. End of week 16 (Sunday 21 April, 23:59)
    3. End of week 17 (Sunday 26 May, 23:59)

You will receive a (short) feedback message from us on your portfolio assignments that you can use for improvements before finalizing your hand-ins.

https://kursuskatalog.au.dk/en/course/115680/Methods-2-The-General-Linear-Model

Ordinary examination and re-examination:
The exam consists of a portfolio containing some assignments. The total length of the portfolio is 3-7 assignments.

Their form and length will be announced on Blackboard by the teacher at the start of the semester. The portfolio may include products. Depending on their length, and subject to the teacher's approval, these products can replace some of the standard pages in the portfolio.

It must be possible to carry out an individual assessment. So if some parts of the portfolio have been produced by a group, it must be stated clearly which parts each student is responsible for, and which parts the group as a whole is responsible for.

The complete portfolio must be submitted for assessment in the Digital Exam system.
Each student submits a portfolio."

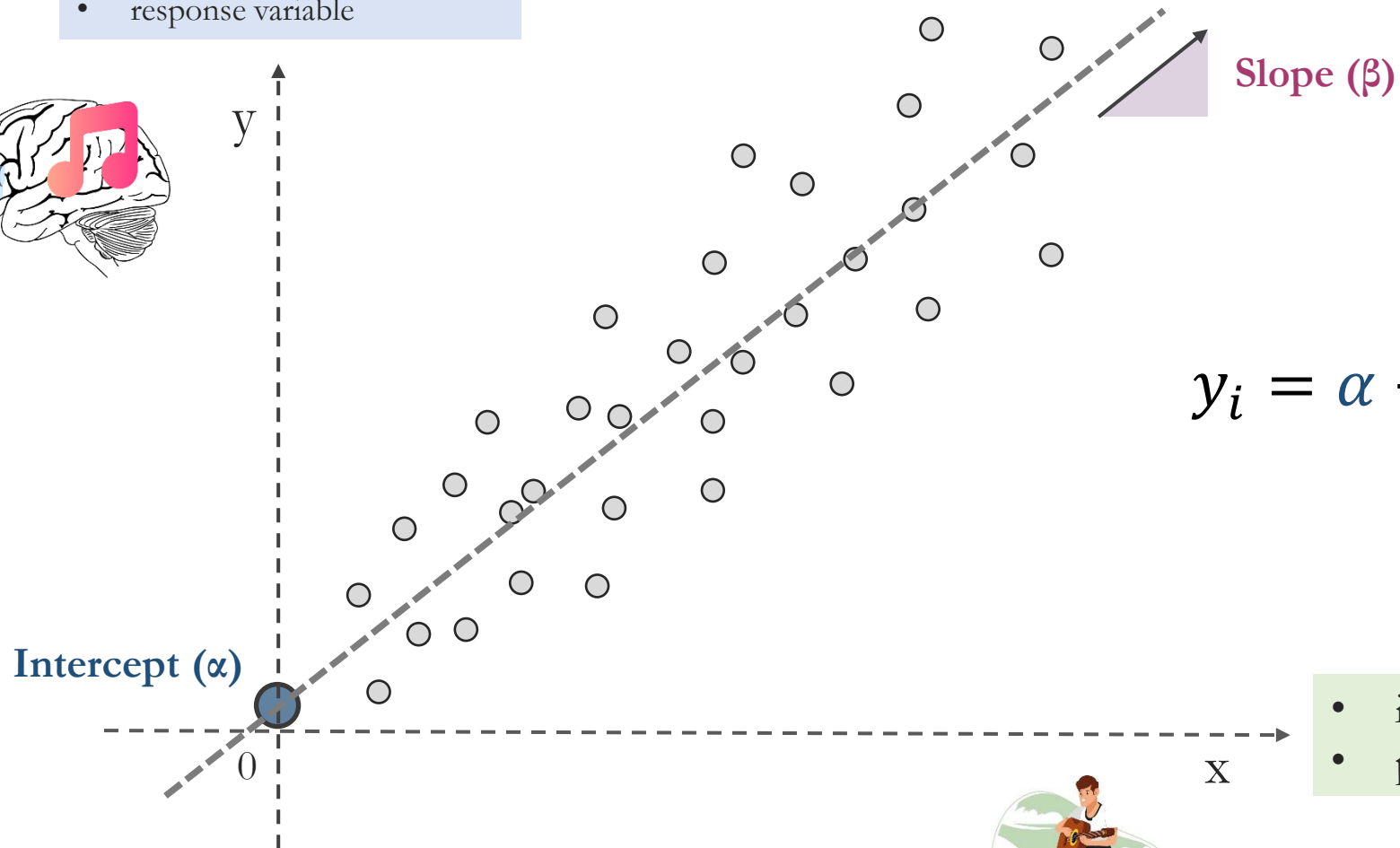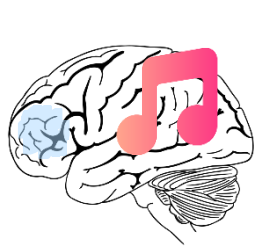# Regression and the GLM:
overview, data and measurement

Nicolas Legrand – Researcher– ILAB (Interacting Minds Centre)

AARHUS UNIVERSITY

Linear regressions

- dependent variable
- outcome variable
- response variable

Slope (β)

$$y_i = \alpha + \beta x_i + \epsilon_i$$

Intercept (α)

y

0

x

- independent variable
- predictor

Nicolas Legrand – Researcher– ILAB (Interacting Minds Centre)
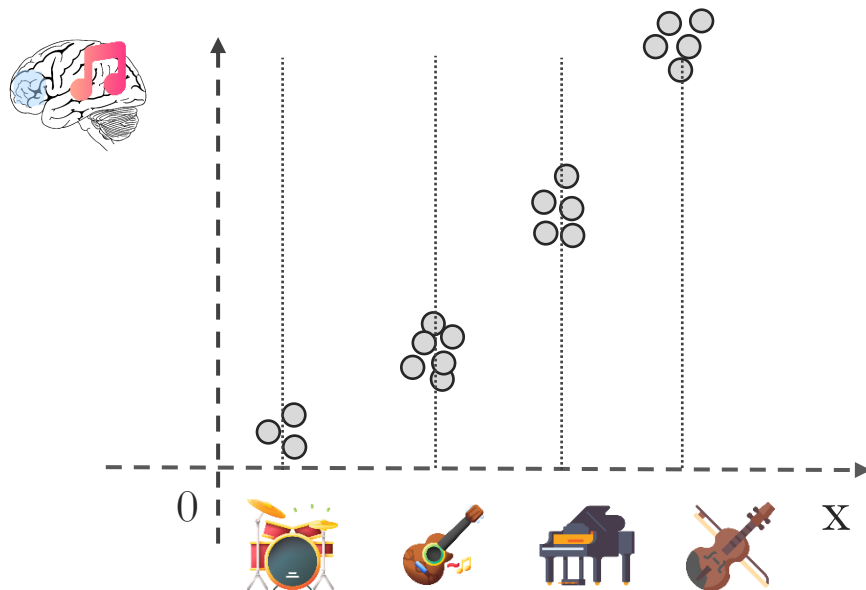
AARHUS UNIVERSITY

Generalizations of linear models
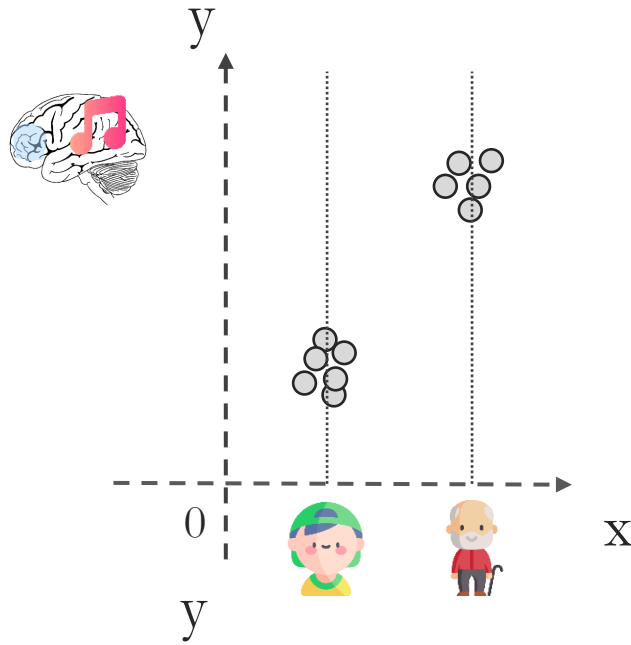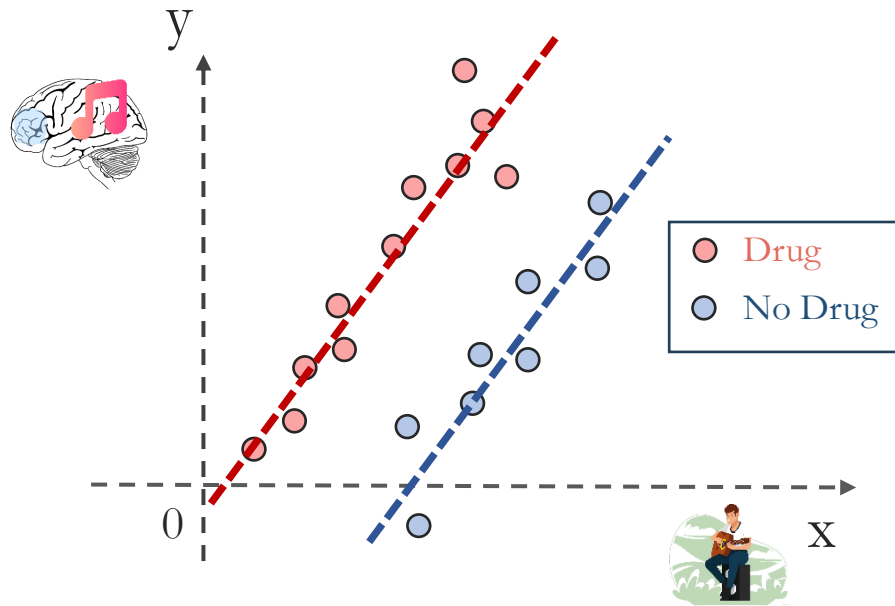

Two sample t-test

$$y_i = \alpha + \beta x_i + X_i^T \cdot \epsilon$$

Factorial ANOVA

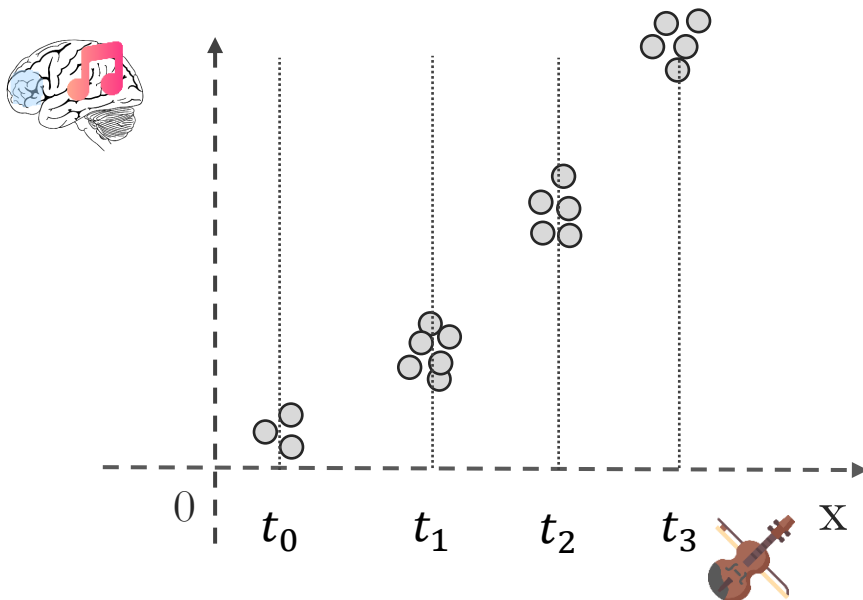$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + X_i^T \epsilon$$

AARHUS UNIVERSITY

Generalizations of linear models

ANCOVA

$$y_{ij} = \mu + \alpha_i + \beta(x_{ij} - \overline{x}_j) + \epsilon_{ij}$$

Repeated-measure ANOVA

$$y_{ij} = \alpha + \beta_j + \epsilon_i$$

The generalized linear model (GLM)

The residuals can come from any distribution from the exponential family

The link between the predictor variables and the outcomes can be any function

$$\mu_i = \mathbf{X}_i^T \cdot \beta \;\rightarrow\; g(\mu_i) = \mathbf{X}_i^T \beta$$

Generalized Linear Model (GLM)[1]

$$\mathbf{Y} = \mathbf{XB} + \mathrm{E}$$

Multiple predictors and outcome variables. Include ANOVA, t-test, ANCOVA…

General Linear Models

Multiple Linear Regressions

Linear Regressions

$$y = \mathbf{X}\beta + \epsilon$$

One outcome, multiple predictors variables
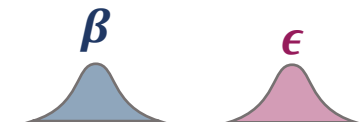
$$y_i = \alpha + \beta x_i + \epsilon_i$$

One predictor, one outcome variable
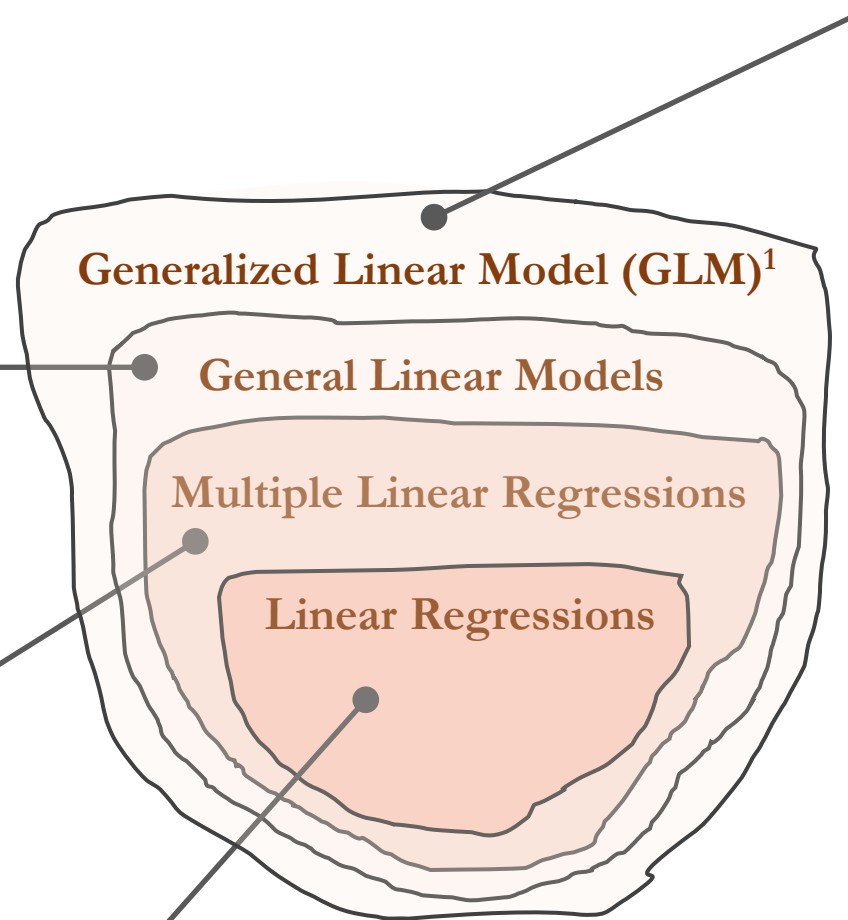
Statistical inference

Three challenges of statistics:
1. Generalizing from sample to population
2. Generalizing from treatment to control group
3. Generalizing from observed measurements to underlying constructs of interest

$\boldsymbol{\beta}$ ϵ

1 - Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized Linear Models. In Journal of the Royal Statistical Society. Series A (General) (Vol. 135, Issue 3, p. 370). JSTOR. https://doi.org/10.2307/2344614

Nicolas Legrand – Researcher– ILAB (Interacting Minds Centre)

AARHUS UNIVERSITY

Statistical inference

$$y_i = \alpha + \beta x_i + \epsilon_i$$

Probability Theory

Bayesian inference

$$P(\alpha|y) = \frac{P(y|\alpha)P(\alpha)}{P(y)}$$

α β ϵ

Calculus

Nicolas Legrand – Researcher– ILAB (Interacting Minds Centre)

AARHUS UNIVERSITY