

# Bayesian data analysis (Aalto fall 2024)

## Practical matters

- Book: Gelman, Carlin, Stern, Dunson, Vehtari & Rubin: Bayesian Data Analysis, Third Edition. (online PDF available)
- The course website has more detailed information  
[https://avehtari.github.io/BDA\\_course\\_Aalto/Aalto2024.html](https://avehtari.github.io/BDA_course_Aalto/Aalto2024.html)
- Timetable: see the course website
- TAs: David Kohns, Mélanie Guhl, Noa Kallioinen, Anna Riha, Varun Shanmugam, Maksim Sinelnikov, Teemu Säilynoja



# Bayesian data analysis

## Schedule

- No lecture on 9th September as I and most TAs are in StanCon conference
- Yes lecture on 14th October although it is evaluation week
- No lecture on 24th October as I'm away
- Project presentations on week starting 9th December

# Bayesian data analysis

## prerequisites

- Basic terms of probability theory
  - probability, probability density, distribution
  - sum, product rule, and Bayes' rule
  - expectation, mean, variance, median
- Some algebra and calculus
- Basic visualization techniques (R or Python)
  - histogram, density plot, scatter plot

These will be tested with the first assignment round

# Bayesian data analysis

## prerequisites

- What to do if the course seems to be too difficult
  - refresh your memory on prerequisites (see the course web site for some links)
  - ask for help
  - consider reading Regression and Other Stories  
<https://avehtari.github.io/ROS-Examples/>
  - consider reading Statistical rethinking + watching videos  
<https://xcelab.net/rm/statistical-rethinking/>

# Bayesian data analysis

## Course contents

- Background (Ch 1)
- Model, likelihood, prior, posterior, integration (Ch 2)
- Integration in multiparameter models (Ch 3)
- Basic integration methods (Ch 10)
- Markov chain Monte Carlo integration (Ch 11–12)
- Stan and probabilistic programming
- Hierarchical models (Ch 5)
- Model checking (Ch 6)
- Evaluating and comparing models (Ch 7 + extra material)
- Decision analysis (Ch 9)
- Large sample properties and Laplace approximation (Ch 4)
- Bayesian workflow (project)

# Bayesian data analysis

## Different learning styles

- Reading
- Listening lectures
- Solving problems
  - mathematical derivations
  - programming

# Bayesian data analysis

## Assessment

- Assignments 40%, e-exam 10%, and project work and presentation 40%
  - Minimum of 50% of points must be obtained from each

# Bayesian data analysis

- Lectures describe basics and give broader overview (recorded and made available)
  - written material has all the details and self-study is possible
- Supporting material and assignments in [https://avehtari.github.io/BDA\\_course\\_Aalto/Aalto2024.html](https://avehtari.github.io/BDA_course_Aalto/Aalto2024.html)
  - reading instructions and chapter notes
  - demos (very useful for assignments)
  - slides (not very useful without the lectures)
  - video clips
  - links to additional material
- R demos [https://avehtari.github.io/BDA\\_course\\_Aalto/demos.html#BDA\\_R\\_demos](https://avehtari.github.io/BDA_course_Aalto/demos.html#BDA_R_demos)
- (Python demos [https://avehtari.github.io/BDA\\_course\\_Aalto/demos.html#BDA\\_Python\\_demos](https://avehtari.github.io/BDA_course_Aalto/demos.html#BDA_Python_demos))
- Aalto Zulip chat instance (link in MyCourses)



# Bayesian data analysis

## Assignments

- Weekly assignments (some have two weeks time)
  - R (Python) simulation assignments
  - Stan probabilistic programming assignments (via R (Python))
- Related R (Python) demos available (see the course web site)
- TAs available: the web page for TA session times
- Assignment deadlines on Sunday (see detailed info in the course web page)
  - we recommend to submit before Friday 3pm as TAs are not available during the weekend
  - we allow the late submission on Sunday as some students are working on weekdays
- Submit assignment answers via MyCourses Quizzes
- For some later assignments submit also a report in FeedbackFruits

# Bayesian data analysis

R

- We use R in the course as there are more packages for Stan and statistical analysis in general in R
  - this will make specifically the later part of the course easier

# Bayesian data analysis

## Assignments

- Assignments are available in the course website
- Assignments are returned by answering MyCourses Quizzes and in some later rounds also by submitting to FeedbackFruits

# Assignments

## MyCourses Quizzes

- Assignment answerd submitted via MyCourses Quizzes
  - multiple choice questions
  - multiresponse questions
  - numerical value questions (with tolerance)

# Assignments

## FeedbackFruits

- Peergrading used in BDA course since 2016
  - peergrade.io used 2016–2023, but it was shut down in 2023
  - FeedbackFruits is missing some features, and that's why we use it only for some assignments and for the project report
- Submit report as PDF
- Each student grades 3 assignment reports (randomly distributed)
- Detailed grading instructions – rubric (available also on the course website)
- Graders provide also text feedback
- Quality of grading is evaluated, too
- See more at [https://avehtari.github.io/BDA\\_course\\_Aalto/assignments.html](https://avehtari.github.io/BDA_course_Aalto/assignments.html)
  - FeedbackFruits instructions still in progress

# Assignments

## Feedbackfruits

- Login with Aalto account
- Combined score: 90% submission performance, 10% feedback performance

# Assignments

## Plagiarism and empty reports

- It's OK to discuss assignments with others
- It's OK to use code from the demos (mention the source)
- It's OK to use AI, but need to mention when and how used
  - Warning: I have tested these and they can provide very vague or completely wrong results for the course contents
  - Might be most useful for getting ideas for code and markdown syntax
- Don't copy reports from others or from internet
- Don't submit empty, almost empty or nonsense report
  - these will be problematic for other students
  - if you see such, send TAs a message and mark it as problematic in Peergrade and get another one for grading

# E-exam

- All exam questions picked from the assignment questions
- No programming tasks
- Choose your own time to take the e-exam from 18 November to 13 December



# Project work

- Project work in groups of 1–3
  - combines all the pieces learned in one project work
  - R notebook / Quarto report
  - project report peer graded (30% of the project score)
  - oral presentation graded by me and TAs (70% of the project score)
- More about projects later

# Zulip chat

[bda2024.zulip.cs.aalto.fi](https://bda2024.zulip.cs.aalto.fi)

- Aalto login, hosted by Aalto IT, deleted after one year
- The web interface is better, but the mobile app has gained push notifications, too
- Different streams for announcements, general, assignments, etc.

# RStudio, Quarto, R markdown

- RStudio is a great IDE for R
- Quarto is a markdown language for making reports mixing text, code, equations, tables, etc
  - *Quarto is the next iteration of R Markdown, and allows you can create dynamic content with Python, R, Julia, and Observable, author documents as plain text markdown or Jupyter notebooks, and output to multiple format types.*
- RStudio has also visual editor for Quarto (and R markdown) making it easy for new users
- RStudio is also installed in Aalto JupyterHub

- No need to install anything locally, everything can be done in Aalto JupyterHub
  - The **notebooks** folder is the only persistent folder (stays there if you sign out) so save everything to that folder!
  - You may access JupyterHub folder as a network drive by SMB mounting it on your own computer
  - See more in the course FAQ
- There is some support for local installations (see FAQ in the course web page)

# FAQ

- [https://avehtari.github.io/BDA\\_course\\_Aalto/FAQ.html](https://avehtari.github.io/BDA_course_Aalto/FAQ.html)
- For example,
  - R packages used in demos
  - Installing aaltobda package
  - Installation problems
  - Remote access
  - Tidyverse and pipes
  - What if I missed some deadline or wasn't able to do some part of the course