# Bayesian data analysis (Aalto fall 2023)

- Book: Gelman, Carlin, Stern, Dunson, Vehtari & Rubin: Bayesian Data Analysis, Third Edition. (online PDF available)
- The course website has more detailed information than these slides
  https://avehtari.github.io/BDA_course_Aalto/Aalto2023.html
- Timetable: see the course website
- TAs: David Kohns, Noa Kallioinen, Andrew Johnson, Leevi Lindgren, Anna Riha, Niko Siccha, Maksim Sinelnikov, Teemu Säilynoja
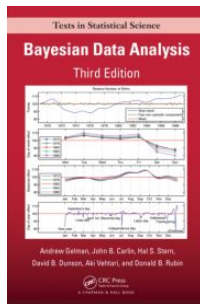
Bayesian data analysis

Pre-requisites

- 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

- What to do if the course seems to be too difficult
  - refresh your memory on pre-requisites (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

- Assignments 60%, and project work and presentation 40%
  - Minimum of 50% of points must be obtained from both the project work and the assignments.

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/Aalto2023.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
- (Python demos https://avehtari.github.io/BDA_course_Aalto/demos.html#BDA_Python_demos)
- Aalto Zulip chat instance (link in MyCourses)

Bayesian data analysis

- 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
- After the assignment deadline, the grading period Monday–Tuesday
- Students grade 3 other assignments using peergrade.io

Bayesian data analysis
R vs Python

- We strongly recommend using R in the course as there are more packages for Stan and statistical analysis in general in R
- If you are already fluent in Python, but not in R, then using Python may be easier, but it can still be more useful to learn also R

Bayesian data analysis

Assignments

- Assignments are available in the course website
- Assignments are returned and graded in Peergrade

Assignments
peergrade.io

- Peergrading used in BDA course since 2016
- Each student grades 3 assignments (randomly distributed)
- Detailed grading instructions – rubric (available also on the course website)
- Also text feedback
- Possible to flag inappropriate grading (please, be kind!)
- TAs check flagged gradings
- Possible to give thumb up for great feedback
  - those who give good feedback will get bonus points
- See more at https://avehtari.github.io/BDA_course_Aalto/assignments.html

Assignments

- Combined score: 80% submission performance, 20% feedback performance

Assignments

- Combined score: 80% submission performance, 20% feedback performance
- Hand-in score:
  - averaging the scores from peers
  - after flagging, teacher may overrule the score
  - different assignments have different weights

  See details at
  http://help.peergrade.io/interfaces-and-features/
  grading-and-scores/the-hand-in-score

Assignments

- Combined score: 80% submission performance, 20% feedback performance
- Hand-in score:
  - averaging the scores from peers
  - after flagging, teacher may overrule the score
  - different assignments have different weights

  See details at
  http://help.peergrade.io/interfaces-and-features/
  grading-and-scores/the-hand-in-score
- Feedback score:
  - When students receive a review, they are asked to react to it using a scale ranging from "Not useful at all" to "Extremely useful".
  - These ratings each correspond to a score between 0% and 100%.
  - The feedback score is the average of the reaction scores.
  - "Somewhat useful. Could be more elaborate." is the baseline reaction.

Peergrade.io

Registration

- Go to BDA MyCourses page
- Click Peergrade and login with Aalto account

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

Project work

- Project work in groups of 1–3
  - combines all the pieces learned in one project work
  - R or Python notebook report
  - project report peer graded (40% of the project score)
  - oral presentation graded by me and TAs (60% of the project score)
- More about projects later

Zulip chat
bda2023.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 new 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

jupyter.cs.aalto.fi

- No need to install anything locally, everything can be done in Aalto JupyterHub
- There is some support for local installations (see FAQ in the course web page)

FAQ

- 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
  - I missed some deadline or wasn't able to do some part of the course