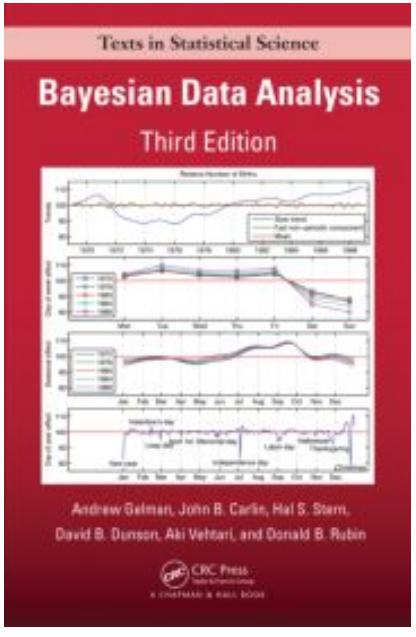
Bayesian data analysis (Aalto fall 2020)

- Book: Gelman, Carlin, Stern, Dunson, Vehtari & Rubin: Bayesian Data Analysis, Third Edition.
- The course website has more detailed information than these slides

  https://avehtari.github.io/BDA_course_Aalto/Aalto2020.html
- Timetable: see the course website
  - TAs: Alejandro Catalina, Akash Dhaka, Kunal Ghosh, Noa Kallioinen, Antgon Mallastom Topi Paananen, Teemu Säilynoja



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

These will be tested with the first assignment round

- Background (Ch 1)
- Single-parameter models (Ch 2)
- Multiparameter models (Ch 3)
- Computational methods (Ch 10)
- Markov chain Monte Carlo (Ch 11–12)
- Stan and probabilistic programming
- Hierarchical models (Ch 5)
- Model checking (Ch 6)
- Evaluating and comparing models (Ch 7)
- Decision analysis (Ch 9)
- Large sample properties and Laplace approximation (Ch 4)
- In addition you learn workflow for Bayesian data analysis

- Treatment/control
  - randomize patients to treatment or control
  - is the treatment effective?

- Treatment/control
  - randomize patients to treatment or control
  - is the treatment effective?
- Continuous valued treatment
  - randomize patients with different dosages
  - which dosage is sufficient without too many side effects?

- Treatment/control
  - randomize patients to treatment or control
  - is the treatment effective?
- Continuous valued treatment
  - randomize patients with different dosages
  - which dosage is sufficient without too many side effects?
- Different effects for different patients?
  - Is the treatment effect different for male/female, child/adult, light/heavy, ...

- Basic visualisation techniques
- Binomial distribution – Algae
- Normal distribution – Windshield
- Difference between binomials – Treatment/control
- Difference between normals – Windshield
- Generalized linear model (GLM) + importance sampling – Bioassay
- GLM + Metropolis + convergence diagnostics – Bioassay
- GLM + Bioassay + Stan
- Linear model + Stan
- Hierarchical model + Stan
- Model seletion + Stan

- Exercises 2/3, and project work and presentation 1/3
  - Minimum of 50% of points must be obtained from both the project work and the exercises.

Bayesian data analysis

- Pre-recorded lectures describe basics and give broader overview
  - written material has all the details and self-study is possible
- Supporting material and assignments in https://avehtari.github.io/BDA_course_Aalto/Aalto2020.html
  - reading instructions and chapter notes
  - demos
  - slides (not very useful without the videos)
  - 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 chat instance

- Weekly exercises (some have two week time)
  - R (Python) simulation exercises
  - Stan probabilistic programming exercises (via R (Python))
- Related R (Python) demos available
- TAs available: see Oodi for exercise sessions
- Exercise deadlines on Sunday (see detailed info in the course web page)
- After exercise deadline grading period Monday–Tuesday
- Students grade 3 other exercises using peergrade.io

- 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

- Exercises are given on PeerGrade (also available in the course website)
- Exercises are returned and graded on Peergrade

Exercises
peergrade.io

- Used in BDA course since 2016
- Each student grades 3 exercises (randomly distributed)
- Detailed grading instructions – rubric (available also on the course website)
- Also text feedback
- Possible to flag inappropriate grading
- TAs check flagged gradings
- Possible to give thumb up for great feedback
  - those who give good feedback will get bonus points

Exercises
peergrade.io

- Combined score: 70% submission performance, 30% feedback performance

- Combined score: 70% submission performance, 30% feedback performance
- Hand-in score:
  - averaging the scores from peers
  - after flagging teacher may overrule the score
  - different exercises have different weight

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

- Combined score: 70% submission performance, 30% feedback performance
- Hand-in score:
  - averaging the scores from peers
  - after flagging teacher may overrule the score
  - different exercises have different weight

  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.

Peergrade.io
Registration

- Go to peergrade.io/join
- Use class code: (see MyCourses announcements)
- Use your Aalto email or we can't match you to your student id

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
  - oral presentation graded by me and TAs