Excellent health statistics - smokers are less likely to die of age related illnesses.'

statistics (def.):

(1) a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data,

(2) the only science in which two recognized experts, using exactly the same set of data, may come to completely opposite conclusions.

Course introduction

Erik Štrumbelj

2019

Why Bayesian?

- Computationally intensive but conceptually simple.
- Less prone to misinterpretation.
- Straightforward decision-making.
- Likely the future of how statistics is taught and applied. And definitely should be part of every statistician's toolbox.

Course objectives

Main: To apply Bayesian statistics in practice.

- Learn about the Bayesian view on probability and inference.
- Start using tools for Bayesian statistics.
- Understand the principles of Markov Chain Monte Carlo and their implications for Bayesian computation.

Illustrative examples

Course information

Prerequisites: Basic probability theory, basic statistics, R programming.

Organization: 6 lectures + 2 hands-on sessions.

Requirements:

- Take-home problem set (after first 4 lectures),
- Final project (deadline: before end of the school year).

Materials:

- `https://github.com/bstatcomp/hse_bayesian`
- All the slides, code and data will be added before each lecture.
- Further reading references will be included at end of each lecture.

Final project information

Goal: Demonstrate your ability to perform a Bayesian analysis from start to finish (data, modelling, computation and interpretation/decision-making). You are encouraged to propose your own topic. If you don't have one, let me know and I will recommend one. You may use any tools for Bayesian statistics.

Timeline:

- Submit topic proposal (as soon as possible; May 15 at the latest).
- Submit first draft (at least 3-4 weeks before final version is due, to allow for comments and revision).
- Submit final version (before end of school year).

Final submission must be a dynamic report (pdf or html; Sweave, Markdown or Notebook), including all the data required to reproduce the results. The report must clearly and concisely describe all the steps of your analysis.

Contact: `erik.strumbelj@fri.uni-lj.si`

You *may* be allowed to work in groups on the same topic as long as each member must can work on a distinct part of the problem and will submit their own final report that focuses on their individual contributions.

The tools that you will see me use

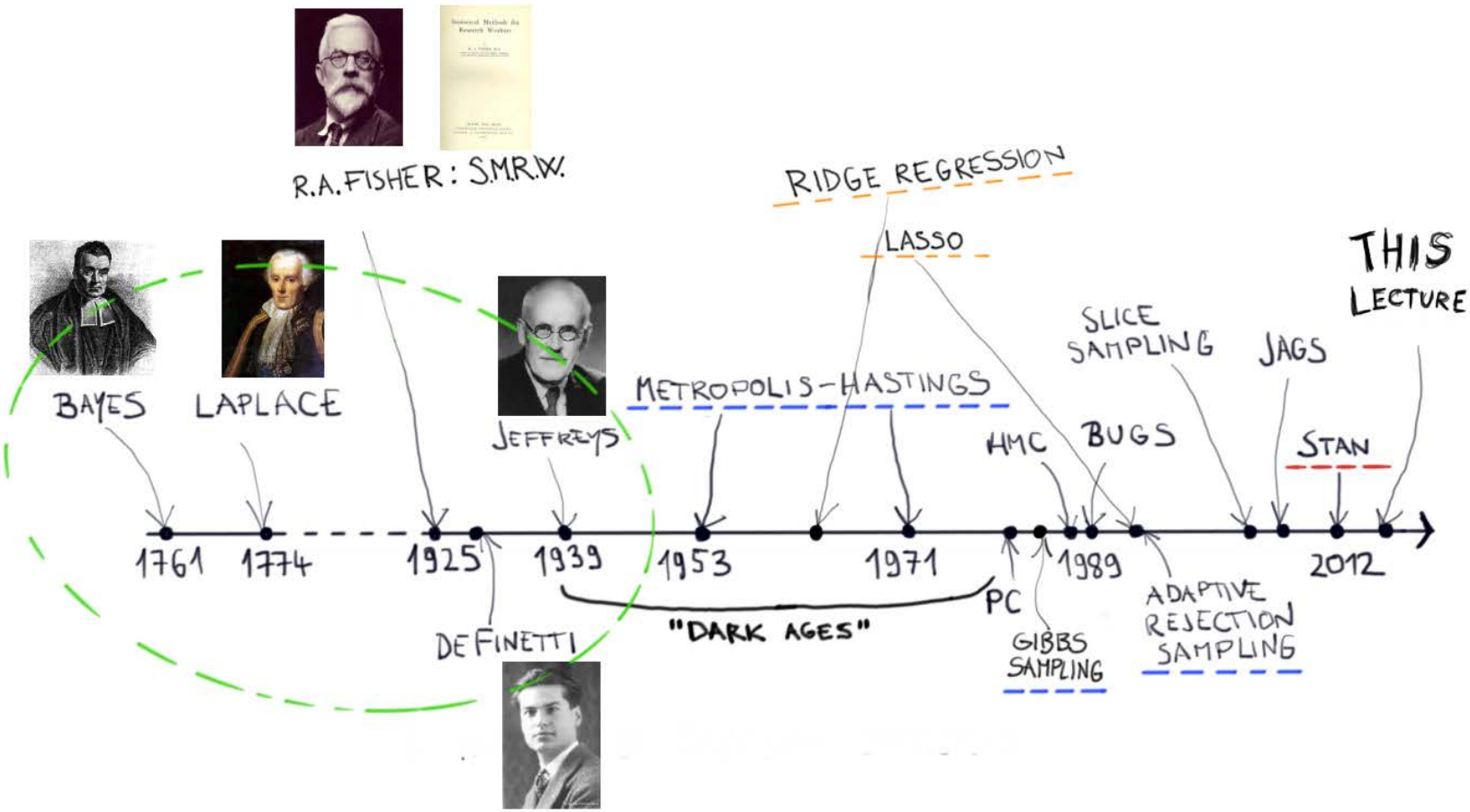
Software:

- **R** + **RStudio**,
- **ggplot2** package for visualization,
- **Stan** for Bayesian inference.

Reporting:

- **LaTeX** + **Texmaker**,
- **RStudio** + dynamic reports (**sweave**, **R markdown**, **R notebook**).

The Bayesian statistics timeline



Lectures outline

1. **Probabilistic thinking**
2. Principles of Bayesian inference
3. Probabilistic programming with Stan
4. Estimation, group comparison and linear regression

——————— break ———————

5. A gentle introduction to Markov Chain Monte Carlo
6. Hands-on session 1
7. Hierarchical modelling
8. Hands-on session 2 & Where to go from here