ENGG 107: Bayesian Statistical Modeling and Computation

Setting up computation
(Meeting #5)

Motivating questions for today:

1. How to install and use R and R-Studio?
2. How to learn R syntax?
3. How to document your analysis?
4. How to make your analysis method FAIR?

Current Schedule

Week starting	Class 1	Class 2	Class 3
January 2, 2023	No Class	Introduction to the class	Bayesian Approach I
January 9, 2023	Bayesian Approach II	Designing a project	Setting up Computation
January 16, 2023	No class (MLK day)	Analytical Solutions / Intuition	Bayesian Workflow
January 23, 2023	Precalibration /(Bayes) Monte Carlo	MCMC Part 1	MCMC Part 2
January 30, 2023	Students pitch project ideas	Writing a method section	Catching up / review
February 6, 2023	Convergence diagnostics	Checking Assumptions	Deep Uncertainty
February 13, 2023	Getting Solid Priors	Links to Decision-Making	Links to Decision-Making
February 20, 2023	Model Choice	Emulation	Sensitivity analysis
February 27, 2023	Communication of results	Student Presentations	Student Presentations
March 6, 2023	Class Debriefing / Resarch links	No Class	No Class

Assignments Review

- Please use the links (R and RStudio) to download and install the latest versions of R and RStudio for your preferred operating system.
- Bring your laptop with the install (If you cannot use R locally, R is available on the Thayer Linux systems located in MacLean 210 and Cummings G13.
- Please review Lab#0 in Applegate and Keller (2016) (and ignore the "SWIRL" part)
- Please review: "An Introduction to R"

  => http://cran.r-project.org/doc/contrib/Lam-IntroductionToR_LHL.pdf
- Please review simpleR - Using R for Introductory Statistics"

  => http://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf

Questions for today

1. How to install and use R and R-Studio?

2. How to learn R syntax?

3. How to document your analysis?

4. How to make your analysis method FAIR?

Questions for today

1. How to install and use R and R-Studio?

2. How to learn R syntax?

3. How to document your analysis?

4. How to make your analysis method FAIR?

How to install R and R-studio?

https://rstudio-education.github.io/hopr/starting.html

Questions for today

1. How to install and use R and R-Studio?

2. How to learn R syntax?

3. How to document your analysis?

4. How to make your analysis method FAIR?

158

How to learn R syntax?

- Review Lab#0 in Applegate and Keller (2016) (and ignore the "SWIRL" part)
- Review: "An Introduction to R"

  => http://cran.r-project.org/doc/contrib/Lam-IntroductionToR_LHL.pdf
- Review simpleR - Using R for Introductory Statistics"

  => http://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf

Questions for today

1. How to install and use R and R-Studio?

2. How to learn R syntax?

3. How to document your analysis?

4. How to make your analysis method FAIR?

160

How to document your analysis?

- Decide on GUI vs workbook vs **script**

- Use of **repositories**

- Document your code with comments

Questions for today

1. How to install and use R and R-Studio?

2. How to learn R syntax?

3. How to document your analysis?

4. How to make your analysis method FAIR?

What is a FAIR analysis?

Findable, **A**ccessible, **I**nterpretable, **R**eproducible

- Chue Hong, N. P., Katz, D. S., Barker, M., Lamprecht, A.-L., Martinez, C., Psomopoulos, F. E., et al. (2022). *FAIR Principles for Research Software (FAIR4RS Principles)*. https://doi.org/10.15497/RDA00068
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Du Sert, N. P., et al. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9. Retrieved from https://www.nature.com/articles/s41562-016-0021%C3%82%C2%A0
- Stark, P. B. (2018). Before reproducibility must come preproducibility. *Nature*, *557*(7707), 613. https://doi.org/10.1038/d41586-018-05256-0
- Wicherts, J. M., Bakker, M., & Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PloS One*, *6*(11), e26828. https://doi.org/10.1371/journal.pone.0026828
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18

Where to go for resources?

- [CRAN Task View Bayesian Inference](#)

- [LaplacesDemon: Complete Environment for Bayesian Inference](#)

- [abc: Tools for Approximate Bayesian Computation (ABC)](#)

- Applegate, P. J., & Keller, K. (Eds.). (2016). [Risk analysis in the Earth Sciences: A Lab manual](#). 2nd edition. Leanpub.
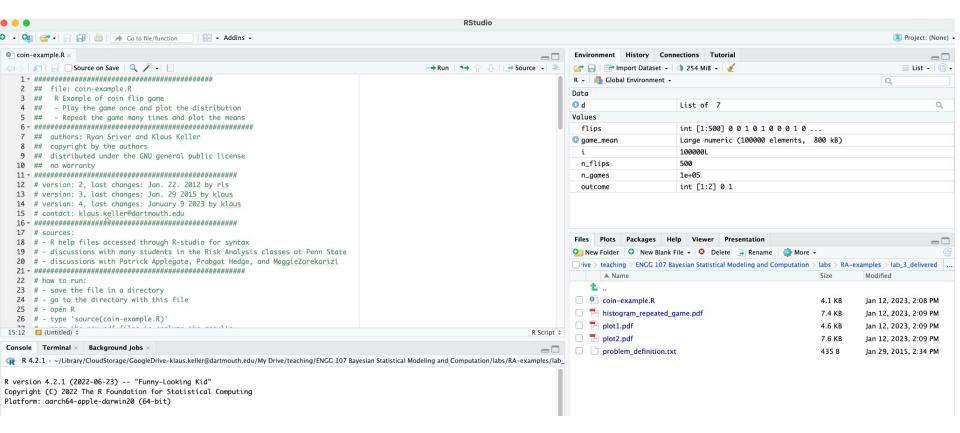
- [LearnBayes: Functions for Learning Bayesian Inference](#)

Let's dive into a simple example analysis

```
############################################################
# problem definition:
# You flip a coin 500 times
# Heads=1, Tails=0
# What do you expect as a mean outcome?
# What is a 95% confidence interval for this mean outcome?
############################################################


##########
Approaches:
- Make a flow diagram of the method
- sketch the figures you want to produce
- review the function "sample"
- recycle the old script components
- hack away and test
- document…
```

Let's dive into a code example

Reading Assignments for Next Class

- Core:
    - D'Agostini, G. (2003). *Bayesian reasoning in data analysis: A critical introduction*. Singapore: World Scientific Publishing. (Chapter 6, an easy and elegant introduction).

- Supplementary (if you are interested in background)
    - Pei, Y., Biswas, S., Fussell, D. S., & Pingali, K. (2019). An elementary introduction to Kalman filtering. *Communications of the ACM*, *62*(11), 122–133. https://doi.org/10.1145/3363294
    - Kalman Filter. (n.d.). Retrieved November 29, 2022, from https://en.wikipedia.org/wiki/Kalman_filter
    - Chen, Z., & Others. (2003). Bayesian filtering: From Kalman filters to particle filters, and beyond. *Statistics*, *182*(1), 1–69.

Review

Do you know

1. How to install R and R-Studio?
2. How to learn R syntax?
3. What is a FAIR analysis?
4. Why you may want to opt for a FAIR analysis?
5. How to run, read, and reproduce an example analysis script?