The Journal of Open Source Education

Productive Controversy: A Classroom Activity to Contrast Frequentist and Bayesian Paradigms

Zachary del Rosario [1], **Stefani Langehennig** [2], **and Mine Dogucu** [3]

1 Olin College of Engineering **2** University of Denver **3** University of California, Irvine

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Submitted: 12 July 2024
Published: unpublished

Statement of Need

As a discipline, statistics is quite special. While most quantitative sciences operate under a single *paradigm* (a set of common, unchallenged assumptions) (Kuhn, 1996), statistics has two: frequentist and Bayesian. That statisticians operate under two paradigms *that sometimes give different results* is disturbing to some practicing scientists (Efron, 2005). Many students do not even get a chance to observe this controversy: A recent survey of high-ranking institutions found that only ~30% offer *any* Bayesian statistics course (Dogucu & Hu, 2022).

Far from being an academic distinction, modern students must have a minimal comprehension of the importance of paradigmatic differences. For students to become practicing statisticians, they must develop a skeptical disposition, particularly towards statistical assumptions (Wild & Pfannkuch, 1999). For students to be productive and active citizens, they must have foundational statistical literacy (Engel, 2017). A foundational component of statistical literacy is the basic understanding that quantitative results are based on analytic assumptions. Our proposed activity is a constructivist approach to generate a *productive controversy* that highlights this fundamental insight.

Our Story

This activity was developed under the NSF-funded project *Bayes BATS* [^1]. The third author (Dogucu), as co-PI on this NSF grant, recruited the first and second authors (del Rosario and Langehennig) to participate in a multi-day Bayesian statistics faculty boot camp aimed at developing and disseminating Bayesian statistics educational materials. Through discussions at the workshop, del Rosario and Langehennig came up with the idea to generate a "classroom controversy" to highlight the differences between frequentist and Bayesian statistics. With continued mentorship from Dogucu, the three developed the proposed activity.

Learning Module

In this section we briefly describe the learning module. Overall, the module is a one-class activity (~100 minutes) that explores the Climate and Economic Justice Screening Tool (CEJST) dataset using in-class (R) Quarto notebooks, small-group discussions, and full-class discussions. For instructors and learners in less-resourced environments, it is possible to run this activity in a "pure paper" form (see Practical Tips below).

In addition to this paper, we have recorded a video overview of the activity and written a run-of-show document

JOSE
The Journal of Open Source Education

Instructional Design

The activity is designed using the 5E Instructional Model, consisting of *engage*, *explore*, *explain*, *elaborate*, and *evaluate* (Duran & Duran, 2004). The 5E model follows a constructivist theory of learning (Saunders & Wong, 2020), in part by spending minimal time in passive lecture (for us, the *explain* phase) and a majority of instructional time with students actively engaged in learning activities (for us, the *explore* and *elaborate* phases).

The activity actually runs through *two* 5E "loops": The first 5E loop has students use a statistical analysis on the CEJST dataset to answer a question. The activity has been designed to have different groups come to *opposite* conclusions, motivating the second (faster) 5E loop where students investigate how they came to different conclusions.

The activity comes in both frequentist and Bayesian form, which *crucially* is not made apparent to students at the beginning of the activity. This is intended to give students a sense of ownership over their results (first 5E loop), motivating them to dig deeper when they discover that other groups came to opposite conclusions (second 5E loop).

Learning Objectives

The activity learning goals are stated directly for students in the 01-introduction activity file. However, the *overall* goal of this activity—hidden at first from students—is to illustrate how different assumptions strongly impact analysis results.

Contents

The activity is divided into several Quarto notebooks, written in the R programming language:

(First 5E Loop) - 00-context: Introduces the CEJST dataset - 01-introduction: Introduces the ideas of statistical paradigms, inference, and answering research questions with statistics - `02-activity` is the bulk of the activity. It comes in two version: - 02a-activity-freq: Frequentist version - 02b-activity-bayes: Bayesian version - 03-simplified: A short reference guide to differences between frequentist and Bayesian approaches (Second 5E Loop) - 04-conclusion: The moment the "productive controversy" is revealed and discussed - 05-survey: An optional survey to measure student learning

Practical Tips

- The activity is intended to be run in a single class session (~100 minutes) with students in groups of ~4.
  - Since the activity *strongly relies* on there being two versions (frequentist and Bayesian) of the activity "in play", it is wise to pre-assign students to groups and distribute the two versions to roughly half the class (either via your LMS, or with paper handouts).
- The Quarto notebooks are quite detailed, which should allow the instructor to focus on classroom observation and management. Make sure to review all notebooks before running the activity to "get a sense" of the activity.
- The repository includes a `Makefile` that automates the "assembly" of the learning materials:
  - Code approach: Run `make code` from the root directory to assemble all of the notebooks in the `freq/` and `bayes/` folders.
  - "Pure paper" approach: Run `make pdfs` from the root directory to create print-ready PDFs for the `freq/` and `bayes/` versions. Note that this requires the Quarto CLI tool.
    - For the `bayes/` version of the activity, this will also produce "posterior pamphlets" {`99-postMA.pdf`, `99-postMN.pdf`, `99-postNH.pdf`} to enable students

JOSE
The Journal of Open Source Education

in the "pure paper" approach to see the result of their choice of prior.

For computational reproducibility, session info from a successful build of the activity materials is provided in this file.

Our Experiences

del Rosario: I ran this with the "pure paper" approach in a colleague's Data Science class at Olin College of Engineering. In this approach, I printed two copies of the activity packet for each 4-person group: This required pairs to work together to make sense of the activity. I also printed several copies of each "posterior pamphlet" to enable each group in the Bayesian arm of the activity to make their choice of prior.

Going into the activity, I prepared a short (~5 minute) lecture on the CEJST dataset, including a sketch of the EDA graph from the 00-context notebook. My aim with this lecture was to use provocative questions to help students draw connections between the data and the context ("What does energy burden have to do with sustainability?"), basic comprehension of the statistical questions ("What does a positive association between energy burden and percent Black mean?"), and the heart of statistical inference ("Should we blindly trust this curve?").

I used the following timings to facilitate the activity, calling out instructions such as "You should be finishing up 01 Introduction and moving on to the 02 Activity."

	Phase	Time
	Lecture	5 min
00	Context	5 min
01	Introduction	10 min
02	CEJST Activity	40 min
03	Simplified 1-pager	
04	Conclusion: Facilitated discussion	10 min
	Conclusion: Small group discussions	15 min

The "key points" that I added (beyond the activity notebooks) were the introductory lecture and facilitated discussion. Using my observations during the CEJST activity, I was able to find groups that had opposite conclusions. I asked for their permission to call on them during the facilitated discussion, which allowed me to unambiguously highlight the "productive controversy". I found this to be a highly effective way to get students engaged: Students talked animatedly in the small group discussion about their different approaches.

I noticed that the groups in the Bayesian arm tended to progress a fair bit slower; this is likely because the Bayesian analysis involved distributions, while the frequentist analysis presented confidence intervals and point estimates (simpler mathematical objects). As a facilitator, I had to invent additional tasks for the frequentist groups.

Langehennig: Similar to del Rosario, I ran the activity using the pure paper approach at the University of Denver in an introductory undergraduate business school statistics class. I used the order and timings outlined above during a normally scheduled lecture. The students were separated into two groups using a random number generator, resulting in four 3-person Frequentist groups and four 3-person Bayesian groups. The students were given their respective packets and worked together on the activity, with Frequentists on one side of the room and Bayesians on the other side. The Bayesians had the additional printed artifact with the various posteriors for their activity that they needed to come collect once they arrived at the "02 Activity" phase.

To start the activity, we had a class-wide discussion on the dataset being used by pulling from the 00-context and 01-introduction documents. This not only introduced the real-

world context in which we would be working, but also got the students to think more critically about the connection between choosing and using statistical methods to answer practical questions. I gave them the opportunity to ask questions about both documents before turning to the applied activity.

Once this initial class-wide discussion was complete, the students turned to the CEJST activity and worked together while I circulated the room, answering any lingering clarifying or topical questions related to the activity. On the Bayesian side, groups debated the differences in outcomes for the analysis conditional on the posterior pamphlet they chose. The groups noticed that other groups were coming to different conclusions around them, so I encouraged them to think about why that was the case. As del Rosario experienced, students on the Bayesian side took longer to complete the activity relative to their Frequentist peers. Since the Frequentists finished sooner, I had them collect their conclusions, write them down, and distill three primary takeaways from their analysis to share with the class during concluding facilitated discussion.

During both the facilitated and small-group discussions, the students not only had spirited debates about whether there is a "right" or "correct" statistical technique to use for the problem, but also had a number of questions for me about the best way to justify their choices when using statistical inference.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No.s #2215879, #2215920, and #2215709.

References

Dogucu, M., & Hu, J. (2022). The Current State of Undergraduate Bayesian Education and Recommendations for the Future. *The American Statistician*, *76*(4), 405–413. https://doi.org/10.1080/00031305.2022.2089232

Duran, L. B., & Duran, E. (2004). *The 5E Instructional Model: A Learning Cycle Approach for Inquiry-Based Science Teaching.*

Efron, B. (2005). Bayesians, Frequentists, and Scientists. *Journal of the American Statistical Association*, *100*(469), 1–5. https://doi.org/10.1198/016214505000000033

Engel, J. (2017). STATISTICAL LITERACY FOR ACTIVE CITIZENSHIP: A CALL FOR DATA SCIENCE EDUCATION. *STATISTICS EDUCATION RESEARCH JOURNAL*, *16*(1), 44–49.

Kuhn, T. S. (1996). *The structure of scientific revolutions* (3rd ed). University of Chicago Press. ISBN: 978-0-226-45808-3

Saunders, L., & Wong, M. A. (2020). *Instruction in Libraries and Information Centers.* Windsor & Downs Press. https://doi.org/10.21900/wd.12

Wild, C. J., & Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review*, *67*(3), 223–248. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x