Surprise - They're Different!
Comparing Frequentist and Bayesian Approaches in Public Policy

Stefani Langehennig, PhD[a], Zach del Rosario, PhD[b], Mine Dogucu, PhD[c]

[a] *University of Denver,*

[b] *Olin College of Engineering,*

[c] *University of California Irvine,*

Abstract

Implementing quantitative methodological techniques is a crucial piece of understanding public policy. While quasi-experimental, spatial (diffusion), and mixed methods are most commonly used when teaching policy studies, little research exists on using Bayesian approaches for policy learning, or how the outcomes from traditional quantitative approaches differ from a Bayesian approach. We propose an applied learning activity for students of public policy that exposes them to Bayesian methods and explores the differences between this statistical paradigm and more commonly used approaches. We do this using a structured interrogation of the The Climate and Economic Justice Screening Tool (CEJST) and the epistemological framings in the 5E model [1]. The activity illustrates the importance of statistical assumptions, and by extension, the impact that different quantitative methods have on understanding public policy. The goal of the study is to introduce students, instructors, and practitioners of public policy to a new way of using statistics, equipping them with the tool set and logical processes necessary to apply either approach as they see fit when studying public policy.

Keywords: Active Learning, Public Policy, Bayesian Statistics, Undergraduate Education

1. Introduction

Social scientists routinely make use of quantitative methods to understand the complex world around them. Approaches employed range from quasi-experimental, spatial (diffusion), and econometric techniques, to methods that are more qualitative in nature. While Bayesian approaches are not completely missing from public policy and related fields (see [2]; [3]), they are underutilized for policy learning among academics and policy specialists.

In this paper, we propose an applied learning activity for students of public policy - though its applications extend beyond this discipline - that exposes them to Bayesian methods. The

---

*Corresponding author

Email addresses: `stefani.langehennig@du.edu` (Stefani Langehennig, PhD), `zdelrosario@olin.edu` (Zach del Rosario, PhD), `mdogucu@uci.edu` (Mine Dogucu, PhD)

activity explores the differences between Bayesian approaches and more commonly used statistical techniques to introduce students, instructors, and practitioners to new and different ways of using statistics to investigate real-world problems. Through the application of this activity, our hope is that using these methods will equip them with the tool set and logical processes necessary to apply different quantitative approaches as they see fit when studying public policy.

What follows is an in-depth examination of how teaching different quantitative methods results in a more robust understanding of public policy. We start by reviewing the research to date on cross-disciplinary approaches to using quantitative methods, as well as how this influences students' assumptions about their own learning using the epistemological framings in the 5E Model [1]. Next, we introduce our applied activity and describe its implementation in the classroom. We then discuss our findings from implementing the activity, followed by our conclusions from the study.

2. Background

Literature review here. High-level framing:

- What's being/has been done in public policy education
- Cross-disciplinary approaches to quantitative methodology education
  - Focus on social sciences
  - Why it's important/works
- Epistemological framings/5E Model

Overview of model when ready:

- *Engage*: Get students interested
- *Explore*: Students do self-directed inquiry
- *Explain*: Give students conceptual tools
- *Elaborate*: Let students work with the tools
- *Evaluate*: Assess the learning outcomes

3. Applied Activity: Comparing Frequentist & Bayesian Approaches

We designed, implemented, and evaluated an activity that highlights the differences between Frequentist and Bayesian statistics. Our primary work included designing an activity that takes students through a structured interrogation of a dataset, with the pilot activity launching during spring of 2024. The activity is designed to take place during one or two class sessions. We evaluated this activity's impact on students' statistical knowledge and epistemological framings using the 5E Model [1], discussed more below.

The activity is focused on a real dataset, discussed more below, to which groups of students are given guided statistical analysis. Students follow a structured process to analyze the dataset and interpret their results. However, different groups receive different versions of the activity: some receive a Frequentist approach, while the others receive a Bayesian approach. By carefully crafting the analyses to reach different conclusions, we aim to surprise students with diverging conclusions. The activity concludes with a final full-group discussion,

where we highlight the importance of statistical assumptions, completing the comparison of Frequentist and Bayesian approaches.

The activity learning objectives are three-fold. First, students should be able to evaluate multiple hypotheses using inferential statistics; second, students should be able to connect their evaluation of hypotheses to real-world factors; and third, students should be able to state the primary statistical assumptions for Frequentist and Bayesian inference, and understand how they can lead to different conclusions. These learning objectives stem from our overall learning goal of engineering a "classroom controversy" to motivate students to find their own understanding of how Frequentist and Bayesian assumptions can lead to different conclusions (and by extension, real-world decisionmaking).

Recruitment

The activity is designed for use in a classroom for non-statistics majors. To pilot the activity among students before disseminating widely, we recruited from two distinct populations at our respective institutions. Students recruited from the University of Denver Daniels College of Business have various majors ranging from Business Information and Analytics to Marketing, while students recruited from Olin College of Engineering have focus on a variety of engineering-related topics.

For both institutions, participants had to be at least 18 years of age or older and must have completed at least one entry-level statistics course. At the University of Denver, all students enrolled in a business school major must complete three courses that are part of their statistics sequence (INFO 1010, 1020, and 2020). Students who were currently enrolled in the Winter 2024 quarter of INFO 1020 received a Canvas announcement about the opportunity to participate in a research study involving an applied activity about statistical inference outside of normal class hours. XX students volunteered to participate after receiving the Canvas announcement. These students met with Dr. Langehennig and completed the activity over the course of approximately 120 minutes and received pizza at the end of the activity.

[PARAGRAPH HERE ON OLIN RECRUITMENT]

Activity Materials

All materials were created using the programming language `R` and can be rendered in .html or .pdf format for use. The materials are openly available for instructors on our GitHub repository. Important starter documents include the run of show, which outlines at a high level the different steps of the activity, as well as the artifacts used in the activity. Further, the learning objectives and details of the activity are fleshed out in the introduction document. Finally, instructors can watch a video overview of the activity on YouTube as well.

3.1. Activty Approach

To implement the activity, there are four steps, each discussed at length below:

1. Setting the context for the real world problem the class is exploring
2. Introducing the motivation for the activity (statistical approaches) given the context
3. Doing the applied activity
4. Closing out the activity

The activity is built around the aforementioned 5E Model Approach, where two loops of the model are working concurrently. The first loop is focused on the applied learning aspects of the activity, or the application of a statistical approach focused on current issues. For example, for *engage* the goal is to motivate students with current issues around climate and equity. The *explore* stage is the opportunity in which students get to do self-directed inquiry. For this activity, that means investigating the real-world dataset provided to them in small groups. *Explain* gives the students the conceptual tools they need to understand the different statistical approaches. Students will learn the basics of assessing and interpreting a fitted statistical model with the instructor. For *elaborate*, students get to work with the tools, meaning they apply the conceptual tools they learned to the real-world dataset. Finally, *evaluate* involves giving students the opportunity to reflect on their understanding of the concepts and application they just did through an instructor-facilitated class discussion.

The second loop is focused on the conceptual learning aspects of the activity, or the ownership of the results that students discovered. *Engage* is focused on the different questions around the data (context) that should motivate their search for an explanation for outcomes using statistical approaches. *Explore* and *explain* capture the introduction to statistical inference broadly defined, the learning objectives for the activity, and the high-level critical differences in Frequentist and Bayesian approaches. For *elaborate*, students articulate the basic concepts of assessing and interpreting a fitted statistical model and come to conclusions related to the research question and hypotheses. Finally, *evaluate* uses the concepts and what students learned in the application to articulate and refine their understanding of the differences between Frequentist and Bayesian approaches.

3.1.1. Problem Context

Given our interest in teaching students new approaches to examining real-world public policy problems, we start our activity by introducing The Climate and Economic Justice Screening Tool (CEJST). The CEJST is the result of President Biden's Executive Order issued in January 2021. The tool is used to identify and subsequently help communities disadvantaged by the burdens stemming from climate change in government social programs. While the data covers a number of burdens (health, transportation, and workforce development, for example), we focus on the sustainability aspects of the tool for our activity, including climate change, energy, and legacy pollution burdens on communities.

We begin with a straightforward explanation of the dataaset, situating it in the contemporary dialogue around climate change. Specifically, we theorize there may be a relationship between climate change burdens and minorities residing in Census tracts. By providing them with this context, we get students to think about possible questions - and by extension, hypotheses - they may be able to explore using statistical inference.

To dig into the context of our real-world problem further, we also provide embedded code snippets and output from `R` of some high-level exploratory data analysis (EDA) for the students to review and discuss. We begin by focusing our attention on a few variables of interest for EDA. We start with the *energy burden percentile*, which captures the percentile of energy cost as well as energy-related pollution within a census tract, as well as the *percent of African-American or Black alone*, which captures the percent of African-American or Black

individuals in a census tract.[1] The instructor walks the students through basic data wrangling, providing prompt questions to get them thinking about the substantive implications of descriptive statistics and visualizations. Below is an example of the output, which contains the code used to create the figure, as well as a short description and a prompt question for the students.

```
# Plot relationship
p <- df_small %>%
  ggplot(aes(
    percent_black_or_african_american_alone,
    energy_burden_percentile
  )) +
  geom_hex() +
  xlab("% Black or African-American in Census Tract") +
  ylab("Energy Burden Percentile in Census Tract") +
  geom_smooth()

# Remove legend
p + theme(legend.position = "none")
```
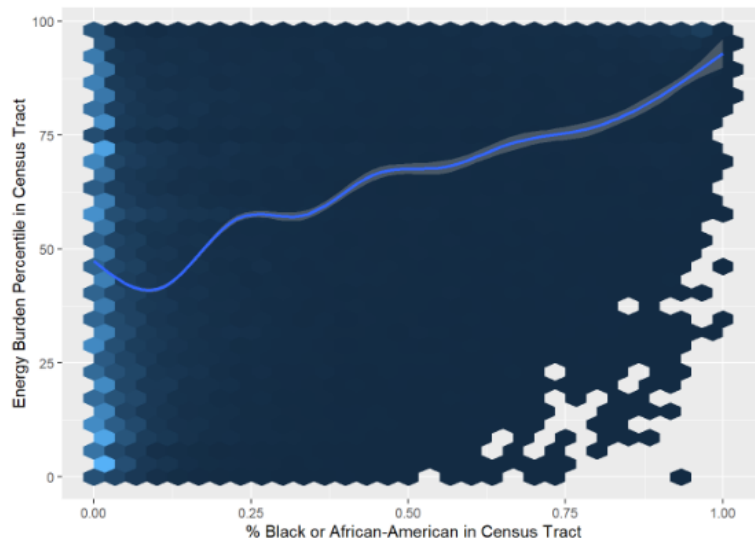

As the figure shows, there is a positive relationship between the percent of Black or African-Americans living in a census tract and the energy burden percentile in the respective census tract.

> **Visualizing Variables of Interest**
> What other ways might we choose to visualize our variables of interest? Is there anything else concerning or interesting based on our visualizations?

After students have investigated the dataset and have a more thorough understanding of the problem at hand, the instructor turns their attention to the introduction document, discussed next.

---

[1]While we guide our students to the variables we want to explore for this implementation of the activity, instructors can modify the activity such that students explore the dataset and identify variables of interest on their own. If they know how to use statistical software, they can also conduct the EDA and analyses in R on their own, rather than giving them the completed version as we do here.

3.1.2. Activity Introduction

This portion of the activity is instructor led, with them walking students through the learning objectives and warmup questions, which in turn initiates a group-wide discussion on statistical inference. The instructor also discusses inference at a high level, offering more pointed discussion around crafting a research question and hypotheses. At this point, the instructor turns students to the simplified critical differences one-pager that introduces them to the primary differences between the Frequentist and Bayesian paradigms. To keep the exercise manageable, we focus students' attention on *general inference* and *model summaries.*[2] Table 1 and Table 2 showcase the differences outlined in the critical differences one-pager between general inference and model summaries:

Table 1: General Inference

Frequentist	Bayesian
Deduction from $\Pr(\text{data} \mid \text{H0})$, by setting α in advance	Induction from $\Pr(\theta \mid \text{data})$, starting with $\Pr(\theta)$
Accept H1 if $\Pr(\text{data} \mid \text{H0}) < \alpha$	$1-\alpha\%$ of most likely parameter values fall within a $1-\alpha$ HPD
Accept H0 if $\Pr(\text{data} \mid \text{H0})\ \alpha$	

Table 2: Model Summaries

Frequentist	Bayesian
Point estimates and standard errors	Descriptions of the posterior distribution such as means and quantiles
95% confidence intervals indicating that 19/20 times the interval covers the true parameter value	Highest posterior density intervals indicating region of highest posterior probability
	$1-\alpha\%$ of most likely parameter values fall within a $1-\alpha$ HPD

The introductory discussion of the activity wraps up with the instructor introducing the research question and associated hypothesis the class will test with their respective statistical approach. Specifically, students with will be assessing whether Black Americans experience a disproportionate level of energy expenditure using inferential statistics and the CEJST dataset.[3]

---

[2]You can access the full one-pager that is instructor-facing on the associated GitHub repository. In addition to comparing general inference and model summaries, it also includes comparisons between fixed variables, interpreting probabilities, and model inference.

[3]An extension of the activity could have the students develop the research questions and hypotheses on their own. Given the activity we implemented is meant to span one class period, we provide those for the students.

3.1.3. Activity Application

The activity application step is the heart of the exercise. Students are assigned a random number generated by a random number generator and put into groups based on their number to go through an applied statistical analysis. There are two versions that are circulated: the Frequentist analysis and the Bayesian analysis. The analyses that are given to the students are completed - they only receive the output of the analysis with associated questions to help them think through the different parts of the analysis before they come to any conclusions.

It is important to note that the data used for each analysis is the exact same for both of the groups, as is the hypothesis that the students are testing. Additionally, the students are asked to assess the same conceptual things, regardless of which activity they receive. They will use the *general inference* and *model summaries* comparison discussed in the Section 3.1.2 to diagnose the outputs of the models from the analyses.

Frequentist Analysis

[TO DO]
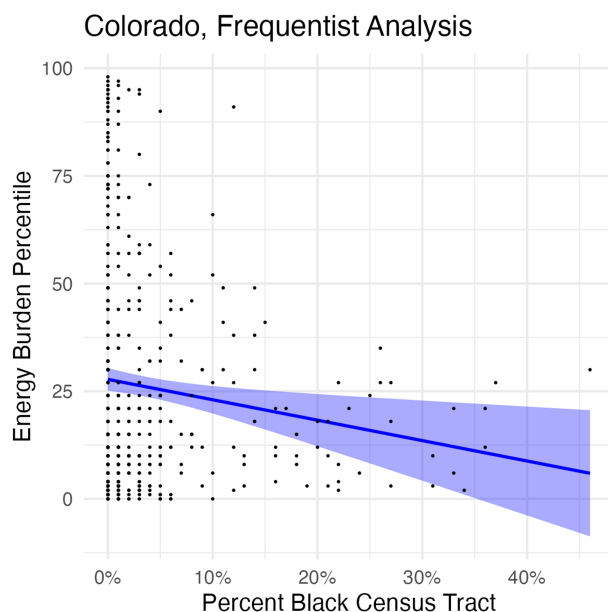
Bayesian Analysis

[TO DO]

3.1.4. Activity Closing

After students have had the opportunity to work through their respective applied analyses and discuss the questions in the associated document, the groups come back together for a full-class discussion. In addition to students jotting down any remaining questions they may have about each of the following questions, they are posed to the whole class for discussion and are tied to the learning objectives in Section 3:

- What can we say about our hypothesis?

Colorado, MA-based Prior
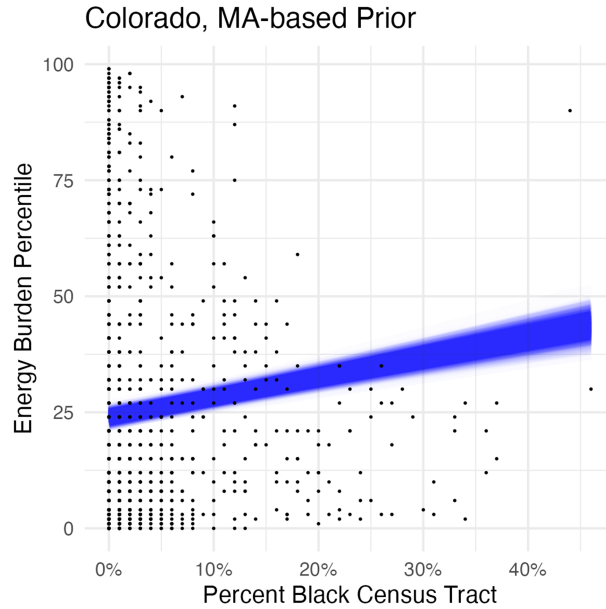
- How would you answer our research question now that we have analyzed the data?
- What can we conclude about the relationship between sustainability and disadvantaged communities? What might you recommend from a policy-making perspective?

We use these questions for a few reasons. The first is so the entire class can hear the impressions of both groups regarding the statistical approach used in their analyses. The second is to play into the "controversy" or differences between the approaches to further engage students on the importance of assumptions for statistical conclusions. The instructor facilitates a debate between the two groups using the critical differences one-pager discussed in Section 3.1.2. Each group likely thinks their conclusions are "correct" based on their analyses, however it is important to point out that the controversy cannot be resolved; rather the results from our analyses are conditional on the assumptions chosen for the analysis, implying that choosing appropriate assumptions is critical to a sound analysis.

4. Evaluation

In addition to exposing students to different statistical paradigms, we are interested in students' awareness of Bayesian methods and their *epistemological framings* — their assumptions about the nature and accessibility of "truth" [1]. Using pre- and post-pre-activity surveys, we measure students' self-reported familiarity of fundamental Bayesian ideas. Specifically, we measure their attitudes before the activity, as well as their change in attitudes after the activity, with respect to ideas around statistical inference. The reason for using a post-pre design after the activity is straightforward. We want to capture changes in self-perceived attitudes about a topic by asking participants to consider where they think their beliefs were before the activity, followed by where they think they are now (see [4]). The participants give themselves two ratings to capture this before and after reflection, as shown below.

To what degree do you (dis) agree with the following statement: There is no uncertainty in the results of a statistical analysis.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
Before activity	○	○	○	○	○
After activity	○	○	○	○	○

In 1-2 sentences, describe your reasoning for your answers to Q1 (just above).

We used a similar format for each of the learning objectives outlined in Section 3. The survey includes the following questions with Likert responses ranging from "Strongly disagree" to "Strongly agree":

1. *To what degree do you (dis) agree with the following statement: There is no uncertainty in the results of a statistical analysis.*
2. *To what degree do you (dis) agree with the following statement: The results of a statistical analysis should not depend on the analyst's assumptions.*

Each question that is associated with a learning objectives also gives students an open-ended opportunity to elaborate on their Likert responses. The survey ends by asking students the following open-ended question:

From the activity, what did you learn about the differences between Frequentist and Bayesian statistics? Please provide as much information as you can.

The post-pre survey is implemented after the activity closing, when students have had the opportunity to talk through the outcomes of each approach with their peers and the instructor. The results of the pilot implementation at the University of Denver and Olin College of Engineering are discussed next.

4.1. Results

Results from surveys here once activity has been implemented.

Implications of the initial results.

5. Conclusion

[overview of findings from evaluation (surveys) here]

This activity bridges the gap between the common Frequentist approach oft taught in both undergraduate and graduate statistics classes and the Bayesian paradigm, to which

many students have not been exposed. We do this using an applied approach, with an eye towards answering education research questions. While many teaching methods highlight the theoretical similarities and differences between frequentists and Bayesians, our activity moves beyond by grounding the comparison in a real-data application, as well as measuring the impact of applying both frameworks in the classroom. In doing so, we hope to introduce students to a new way of using statistics, equipping them with the tool set and logical processes necessary to apply either the frequentist or Bayesian (or both) approaches as they see fit.

Our activity uses an active learning approach, rather than passive lecture. Active learning has been shown to result in superior learning outcomes for students, particularly those from underrepresented groups [5]. In this way, our proposed activity will promote broader impacts of Bayesian thinking.

Our more speculative research goal—to promote more nuanced epistemological framings among students—has further potential impacts. [1] argue that a "sophisticated" personal epistemology is actually achieved when one has access to multiple epistemological framings and can choose to switch between them based on what is productive for the context at hand. Students who can recognize and critique the assumptions underpinning their analyses (treating them as tentative), but carry out their analyses respecting those analyses (treating them as true) will likely be more effective as practicing statisticians. Getting students to recognize the importance of assumptions—and to practice adopting different assumptions— will be a critical first step in developing these multiple epistemological framings.

References

[1] A. Elby, D. Hammer, Epistemological resources and framing: A cognitive framework for helping teachers interpret and respond to their students' epistemologies, Personal epistemology in the classroom: Theory, research, and implications for practice 4 (1) (2010) 409–434.

[2] J. Gill, C. Witko, Bayesian analytical methods: A methodological prescription for public administration, Journal of Public Administration Research and Theory 23 (2) (2013) 457–494.

[3] S. E. Fienberg, Bayesian models and methods in public policy and government settings (2011).

[4] B. Hiebert, K. Magnusson, The power of evidence: Demonstrating the value of career development services, Career Development Practice in Canada: Perspectives, Principles, and Professionalism (2014) 489–530.

[5] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, M. P. Wenderoth, Active learning increases student performance in science, engineering, and mathematics, Proceedings of the national academy of sciences 111 (23) (2014) 8410–8415.