Final Reflection Assignment for Stat 495 **Due: Wednesday, Dec. 20th by 11:59 pm**

This activity will allow you to reflect on your experiences in our class and how they have impacted your journey in statistics. It will allow you to revisit your portfolio and related reflection materials from earlier this semester.

To successfully complete this assignment, you will need to:

- Put a copy of your Homework 5 write-up and your final paper in your Portfolio folder in your Github repo (to round out your portfolio)
- Review the following activities from earlier this semester:
    - The goals you set for the course in your Readmes in your repos.
    - Portfolio Review for Stat 495 (submitted on Gradescope if you need to find it fast)
    - Portfolio InClassActivity (you had as a handout which may have your comments on it, can reference electronic copy in the course repo)
    - The reflection incorporated with Homework 1 (which is in your Portfolio folder in your Github repo)
    - The Hmk 3 problem where you choose a skill to practice
    - Your Hmk 5 and final paper submissions
- Fill out the worksheet below in relation to the Hmk 3 problem, Hmk 5, and final project assignments.
- Write a letter to me with your final responses and insights as directed below (still in this document, so it can all be submitted to Gradescope). The worksheet should help with part of the letter.
- Submit the worksheet portion and letter to Gradescope in the Final Reflection assignment, and put a copy in your repo's Portfolio folder.

You can delete this page and the letter instructions that are below from your final submission if you like, but it's not necessary. You'll just skip over them when assigning Gradescope pages anyway.

Don't worry about the spacing or making the worksheet sections look "pretty". I suggest bolding the skills to select the ones you want to keep or just deleting the ones that don't apply to the respective pieces.

Worksheet

Consider your Hmk 3 problem practicing a skill as a portfolio piece (or at least, an activity where you demonstrated statistical knowledge). Review it briefly / skim over it.

1. Which of the following statistical skills are you practicing, applying, and/or demonstrating in this portfolio piece?

Descriptive Statistics
Exploratory Data Analysis
Design
Model Assessment (Comparison)
Estimation
Other Statistical Methods
Appropriate Graphical Displays
Regression (of any kind)
Resampling Methods (Bootstrap, Randomization Tests)
Hypothesis Testing (and Sampling Distributions)
Understanding Variability
Understanding Probability
Working with Statistical Notation / Writing Models

2. Which of the following coding-related skills are you practicing and/or demonstrating in this portfolio piece?

Data wrangling
Good programming practices (comments, organization)
Reproducible workflow
Understanding code syntax

3. Now let's consider communication skills in relation to this portfolio piece. Which of the following are you practicing and/or demonstrating?

Using writing to demonstrate and help develop my understanding of new concepts
Use writing to communicate statistical knowledge and understanding to:
- **A teacher**
- **My classmates / peers**
- A non-expert audience

The ability to make a clear, convincing, coherent, well-organized argument
Present ideas in a logical order with an appropriate narrative thread
Provide sufficient interpretation/explanation of figures/graphics
Provide an appropriate amount of supporting evidence without too much extraneous information
Provide a conclusion that summarizes your findings, their importance and implications, and sets forth proposals for future work

4. What do you consider as the strengths of this portfolio piece?

For Additional Problem 3 in Hmk 3, I did a thorough job in performing EDA for all of the relevant variables. I also gave useful interpretations of the graphical output, trying to glean as much information from the preliminary analysis of the raw variables as I could.

5. What do you consider as the weaknesses of this portfolio piece?

In answering this exploratory prompt, I lacked depth. Beyond observing how all of the variables behave on the surface level, I believe there were more techniques I could have used to do effective EDA. Also, I could have mentioned what my next steps for analysis might be, including any ideas of potential relationships or models that I want to explore.

Consider your Hmk 5 submission as a portfolio piece. Review it briefly / skim over it.

1. Which of the following statistical skills are you practicing, applying, and/or demonstrating in this portfolio piece?

Descriptive Statistics
Exploratory Data Analysis
Design
Model Assessment (Comparison)
Estimation
Other Statistical Methods
Appropriate Graphical Displays
Regression (of any kind)
Resampling Methods (Bootstrap, Randomization Tests)
Hypothesis Testing (and Sampling Distributions)
Understanding Variability
Understanding Probability
Working with Statistical Notation / Writing Models

2. Which of the following coding-related skills are you practicing and/or demonstrating in this portfolio piece?

Data wrangling
Reproducible workflow
Good programming practices (comments, organization)
Understanding code syntax

3. Now let's consider communication skills in relation to this portfolio piece. Which of the following are you practicing and/or demonstrating?

Using writing to demonstrate and help develop my understanding of new concepts
Use writing to communicate statistical knowledge and understanding to:
- **A teacher**
- **My classmates / peers**
- A non-expert audience

The ability to make a clear, convincing, coherent, well-organized argument
Present ideas in a logical order with an appropriate narrative thread
Provide sufficient interpretation/explanation of figures/graphics
Provide an appropriate amount of supporting evidence without too much extraneous information
Provide a conclusion that summarizes your findings, their importance and implications, and sets forth proposals for future work

4. What do you consider as the strengths of this portfolio piece?

In general, I like how I organized the report and allowed a thread of logic within the sections. I also tried a lot of different models and compared them via their performances on training/test sets for the first time.

5. What do you consider as the weaknesses of this portfolio piece?

I was not communicating well to a non-expert audience when I wrote about things like LASSO without explicitly explaining the origin of the name and concept. I also failed to include a lot of steps in logic that I took to decide how to proceed in my analysis. I used and interpreted some methods wrong, and I didn't include the final variables that I decided to include in my predictive model, which was the entire point of the piece.

Consider your final report submission as a portfolio piece. Review it briefly / skim over it.

1. Which of the following statistical skills are you practicing, applying, and/or demonstrating in this portfolio piece?

Descriptive Statistics	**Appropriate Graphical Displays**
Exploratory Data Analysis	Regression (of any kind)
Design	Resampling Methods (Bootstrap, Randomization Tests)
Model Assessment (Comparison)	Hypothesis Testing (and Sampling Distributions)
Estimation	Understanding Variability
Other Statistical Methods	**Understanding Probability**

Working with Statistical Notation / Writing Models

2. Which of the following coding-related skills are you practicing and/or demonstrating in this portfolio piece?

Data wrangling	**Good programming practices (comments, organization)**
Reproducible workflow	**Understanding code syntax**

3. Now let's consider communication skills in relation to this portfolio piece. Which of the following are you practicing and/or demonstrating?

Using writing to demonstrate and help develop my understanding of new concepts
Use writing to communicate statistical knowledge and understanding to:
- **A teacher**
- **My classmates / peers**
- A non-expert audience

The ability to make a clear, convincing, coherent, well-organized argument
Present ideas in a logical order with an appropriate narrative thread
Provide sufficient interpretation/explanation of figures/graphics
Provide an appropriate amount of supporting evidence without too much extraneous information

Provide a conclusion that summarizes your findings, their importance and implications, and sets forth proposals for future work

4. What do you consider as the strengths of this portfolio piece?

Drawing comparisons to concepts/methods that STAT 495 students are familiar with allowed me to construct an argument for using GMMs rather than just introducing the topic. I also ended up doing a toy example, a simulation, and an application, all on small scales. I hope this helped to show the versatility of the concept. I also worked to do citations properly.

5. What do you consider as the weaknesses of this portfolio piece?

The flow of the paper could have been adjusted slightly, and the sections that delve into the fundamental underpinnings of GMMs could use a bit more depth and more examples.

Review your earlier portfolio work (the prior worksheet, in-class activity, and Hmk1 portfolio reflection) and your stated goals for the course (Readme and reflection) and the responses to the new portfolio pieces above.

1. What skills have you developed further or feel you demonstrated well this semester? Do these skills align with your stated goals from earlier this semester?

I stated in my portfolio reflection that I had trouble carrying out simulation-related tasks. In general, I'd also wanted to delve deeper into statistical concepts by looking at applications of the theoretical material and utilizing analytical tools. I feel that I've mostly been able to achieve these goals, as I constructed a simulation of the GMM in my final report and found multiple packages/functions to demonstrate the implementation of the concept.

2. Reflect on your journey through the course. What did you learn this semester (through these assignments and our coursework) about statistics, statistical thinking, and writing in statistics?

Something that I wasn't prepared for was the theoretical component of GMMs being hard to decipher and then communicate. I learned, though, that it was necessary to give the concept enough time to sink in for myself before I could reconstruct it for others. Of course, I've learned massive amounts about new statistical concepts and methods, especially dipping into machine learning, which I'd always been afraid of. But more importantly, I've learned that the most effort goes into figuring out how to structure the flow and depth of whatever I am trying to communicate.

3. Describe a moment during the semester where you felt an idea or concept "clicked" for you. What was the idea/concept (statistical or writing related) and what transpired in the moment it "clicked"?

The moment when I understood the LASSO visual was when I realized how truly effective a statistical visual could be. In a way, I needed the background on LASSO to understand what was being described, but the learning flowed in the other direction as well, where the visual allowed me to absorb the concept and further explain it to someone else.

4. We engaged in several activities to support your writing for the report, ranging from class discussions of statistical papers and discussions with a Writing Associate to trying out SMART goals in class. How did these activities (broadly speaking) inform your writing this semester, particularly on the final paper?

The class discussions of statistical papers were helpful and implicitly guided me as I wrote my paper, but I feel that it happened rather early on, and I didn't consciously have the ideas we talked about in mind as I should have. The SMART goals were helpful in making me understand what I needed to do, but also what I actually could do. The peer review was biggest boost in showing me what was missing from my paper, because I had seen the material for several weeks by that point, and the peer review revealed the pieces that were natural in my mind that I needed to communicate.

Letter Instructions

Purpose:

In this letter, I hope to learn about your journey through the course and progress made towards your goals for the course that you set way back when you created your repos and the goals made in your Hmk 1 reflection.

I also hope you'll share examples of what you learned about statistics / statistical writing that went well (allowing you to demonstrate your knowledge) as well as a challenge you encountered working on the final paper and how you got through it.

Finally, I hope to learn about your perspective of your portfolio, what you are proud of in it, and whether it is an accurate reflection of your progress as a statistics major at Amherst.

Specifics:

Kristen from the writing center suggested a letter format would work well for this, so let's try it out! Remember to start with the date and a greeting, try to show your ideas with sufficient detail that I can really understand, perhaps even picture what you are saying, and don't forget a closing!

Within the letter to me, be sure to craft a narrative that allows you to:

- Re-express the goals you set at the start of the semester (in your repo Readme and reflection for Hmk 1).
- Describe the progress you made towards achieving those goals over the course of the semester.
- Describe two things you learned about writing and illustrate them with examples from our assignments.
- Describe a challenge you encountered while writing for the final paper and how you got through it.
- Reflect on the portfolio now as a whole and comment on how it shows your development as a statistics major at Amherst. Is it complete? Is there anything else you'd like to add to it before you graduate? (Some items may not be written yet.)
- Be sure to include one thing you are proud of when considering your overall portfolio. This could be anything ranging from demonstrated improvement in writing, to having multiple examples to showcase your abilities, to getting through an exposition on a new topic.
- Before closing, what are your next adventures in the realm of statistics? Are you taking any statistics courses next semester? Going into a data analytics position? Etc.

Craft a response, rather than responding to the bullet points.

Assessment:

The worksheet is graded on completion and earnest effort.

The letter is graded similarly, with responses to the above bullets in narrative form and required letter structure (date, greeting, closing) as the parts that need to be completed.

I hope the letter provides a good opportunity for reflection on everything you've learned this semester!

Dec 19, 2023

Dear Professor Wagaman,

I hope you've had a wonderful semester! I myself started it intimidated by the "capstone" course of the major, and just having gotten a little wrecked in STAT 370, I wasn't sure if I was ready for the next and final Statistics course. But right off the bat, I knew I was going to enjoy the class because of its strong structure and its application-driven nature. My goals at the start of the semester were to become more succinct, as I felt that I frequently spent too much space on extraneous information. I also wanted to be more conscious about my audience and to understand how that would translate to how I explain concepts and conclusions. I found difficulty in carrying out more probability and simulation-related tasks, so I wanted to work to strengthen my simulation skills. In general, I wanted to delve deeper into statistical concepts by looking at applications of the material and utilizing analytical tools, and I looked forward to producing a thorough, large-scale project from beginning to end.

Through the multiple writing practices we had throughout the semester, I learned the holes that I frequently leave in my communication of concepts and more often my code. In my final paper on GMMs, I happened to construct a simulation of the concept and found multiple packages/functions to demonstrate the implementation of the concept. I only meant to do an application, but in peer reviewing my report, Mahathi helped me realize that I'd done a toy example, a serious of simulations, and an application (none on a very large scale though). It was also fun to sift through a number of packages for GMMs. I wanted to work on my long-term time management at the beginning of the semester. It was definitely frustrating to pump out major sections of the report weeks in advance, but the incremented spurts of stressed work allowed me to do specific steps thoroughly. This added up and having the work spread out was truly so beneficial to the entire process and my final product.

Something that I wasn't prepared for was the theoretical component of GMMs being hard to decipher and then communicate. I learned, though, that it was necessary to give the concept enough time to sink in for myself before I could reconstruct it for others. Of course, I've learned massive amounts about new statistical concepts and methods, especially dipping into machine learning, which I'd always been afraid of. But more importantly, I've learned that the most effort goes into figuring out how to structure the flow and depth of whatever I am trying to communicate. I believe my portfolio as a whole demonstrates progress both conceptually and in terms of delivery. The statistical methods that I've learned over these few years have built upon each other, and in diving deeper into these concepts, I've also had to learn to maintain an effective level of communication.

For the future, I had planned to apply for jobs over Thanksgiving break but got caught up in a few things and am now currently in the process (I will probably be emailing to ask for a recommendation letter soon!). Next semester, though, I will be taking Data Structures and perhaps another class at UMass, if they can make my enrollment go through.

Thank you for being a phenomenal professor, from Data Science to now! I've had a great semester and learned massively, and I'm so glad that you were there to lead us through it.

Happy Holidays!

Cassie Jin