Leila Peitsch
6/3/2023
Mentor: Erin Lipman

Bayesian Perspectives on Probability and Statistics: Final Write-up

In this Directed Reading Program (DRP), I was able to explore a Bayesian perspective of statistics. In many statistics courses, students are often taught under the Frequentist philosophy - that is using confidence intervals, null hypothesis significance testing, p-value analysis - which is often hard to interpret and can lead to illogical conclusions. The Bayesian perspective, however, relies on prior information along with the new data to form updated posterior information. Within this program, I was able to learn as well as appreciate the usefulness of Bayesian statistics and build an understanding of how to utilize Bayesian models in order to learn about the world around us.

Within the early weeks of the DRP program, I read from the textbook *Bayes Rules! An Introduction to Applied Bayesian Modeling* which provided a paced, novice-friendly introduction to the world of the Bayesian perspective. Erin helped me through the first section of the textbook that covered "Bayesian Foundations" where we covered the differences between Bayesian and Frequentist perspectives and the different ways to use Bayesian statistics within applied modeling - such as Bayesian modeling for discrete and continuous probability functions, and the different conjugate families: Normal-Normal, Beta-Binomial, and Gamma-Poisson. I also learned about some of the nuances behind Bayesian inference, such as the balance between how much influence the prior and new data has on the posterior often depends on how Bayesian statisticians form a prior and how much data is being considered when forming the posterior.

With the presentation in mind, within the latter half of the quarter, Erin and I worked on a project in which I would be able to use what I learned and conduct a Bayesian analysis of some dataset. I learned that Bayesian statistics can be very useful in removing noise from data, especially on a county level (using data from very populated counties to estimate understanding of less populated counties) - and with my interest in the mental health crisis, Erin found a dataset that collected data on how many mental health providers there are within each county in Washington state. Since we were observing the ratio of mental health providers per 100,000 people, we used a Gamma-Poisson model where the variables λ_i represented the rate of mental

health providers per 100k people and x_i represented the number of mental health providers in some *i*-th county. Using the dataset from County Health Rankings & Roadmaps about *Frequent Mental Distress*, in *R*, we were able to model the rates from the given dataset and then the updated rates from our Bayesian analysis that provided a clearer understanding of how many mental health providers there probably are given the number of mental health providers and the current population within each county.

Overall, this DRP was an informative and fun experience that I am so grateful to have been able to participate in. During my STAT 311 class, I was admittedly a little confused by the short "introduction" to Bayesian statistics we had in class, as much of the class was taught from the Frequentist perspective - but due to Erin's expertise in the field and her ability to explain the ins and outs of Bayesian statistics on a beginner level (as well as kindly clarify any misunderstandings I had throughout the experience), I was able to take away information I wouldn't have gotten to learn otherwise as well as an unforgettable experience within my undergraduate career so far. Looking ahead, I hope to find a place within statistics academia, and within there I hope I am able to further learn about Bayesian perspectives - and maybe even promote Bayesian statistics to my doubtful peers.