Progress Report: Week 4

Reading:

I read chapter 9 of *Think Bayes*, which contained a method for extending Bayesian reasoning into more than one dimension. I thought that this explanation for how this work is pretty reasonable, and am starting to wonder if there is a problem that couldn't be solved better by using Bayesian methods.

I also read Downey's "Bayesian election forecasting" which was entirely about how to predict elections using Bayesian methods in response to Silver's post explaining how his model works. I thought that this blog post didn't really provide me a bunch of new information, but rather that the most helpful part of the blog post was that it was a good example for what the finished product of a case study might look like.

I read chapter 10 of *Think Bayes*, which was a chapter about variable hypotheses and in the context of heights of men and women. I didn't really like this chapter as much as some of the other chapters in the book, mostly because I felt like there was a lot of vocabulary sort of thrown at me, and I wasn't expecting that.

I also read Davidson-Pilon's "The Multi-Armed Bandit Problem". This contained a summary of a method for deciding which slot machine had the highest payout rate. I thought that as far as interesting applications of Bayesian methods go, that this was a pretty intuitively explained problem, and I really liked that he called the slot machines "Multi-Armed Bandidts".

Exercises:

This week, I went back and did a few of the exercises from Chapter 7 of *Think Bayes*. I did the light bulb problem, and my likelihood function looked like this

```
    def Likelihood(self, data, hypo):
        """Computes the likelihood of the data under the hypothesis.

        hypo: The rate at which bulbs break per month
        data: the number of broken bulbs
        """
        lam = hypo
        k = data
```

```
        like = thinkbayes2.EvalPoissonPmf(lam,k)
        return like
```

Which gave me the output that 6 more light bulbs in April as the most likely answer. I don't think that I did this problem right, because it seems a little weird that I'd get exactly 3 as my mean light bulb breaking rate per month if before my mean was 50 (assuming that all light bulb breaking rates were equally likely in the beginning.

Case Study:

This week, Claire and I continued work on the case study we thought that we were doing last week: our predictor of guilt in a crime TV show. We made a simple version of our likelihood function, and realized some potential flaws in our case study, or perhaps just the way that we have implemented it. Our likelihood function is

```
    def Likelihood(self, data, hypo):
        """Computes the likelihood of the data under the hypothesis.

        hypo: that a certain character commited the crime
        data: time, characters implicated, how strongly
        """
        #Assuming that for this case, the hypothesis is correct
        time = data[0]
        characterTuples = data [1:]
        print (characterTuples)

        for charTup in characterTuples:
            if hypo == charTup[0]:
                  like = charTup[1] * time/10
                return like
            else:
                like=0.5*time/10
        return like
```

Here what we've done is we've said that our data will take the form `(time in the show, (character1, how strongly the evidence implicates them or how strongly it exhonerates), (characte4, how strongly the evidence is for or against them),...)` Where the data is assessed from a scale of 0 (innocent) to 1(guilty) and 0.5 means that the data doesn't sway the opinion of who did it.

Something that we've noticed with this potential case study is that this particular implementation really just tells us exactly what we think is happening in the show, and while it actively depicts who might be the mostly likely suspect at any point in the show , this is not something that we don't know already. At this point, Claire and I aren't quite sure what we're going to do with this case study. Whether we try a different method of approaching the likelihood function or whether we scrap it and try to find a different case study. We're thinking that we'll model the data as an instrument with "error" that is dependent on when the data appears in over the course of the television show.

Claire and I had a pretty long chat about this, and we think that we're going to switch our case study to something completely different. Now, we're thinking that we're going to do something similar to the GPS problem, but with removing noise from sensor data. We've both had a bit of problems dealing with noisy sensor data, and after the GPS problem that we did in class on Tuesday, we were wondering whether we could apply the same idea to our own noisy sensor data to potentially get a more accurate reading on where a LIDAR, for example, is telling us we are. We're planning on gathering the LIDAR data from the Neatos used for Comp Robo (Claire is taking this class). For this case study, we'll use readings from a LIDAR and a graph similar to the graph we generated for the GPS problem, assuming we start this knowing that we are in an x by y rectangular room. If time allows, we will also use this Bayesian smoothing method to implement a more accurate mapping of a room.

Reflection:

This week, I feel like I did a lot of trying to figure out what to do for a case study, once I realized that my plan wasn't really going to work out. That being said, I think that, in the end, Claire and I came up with a pretty interesting one, and I'm glad that I don't really have to try to force a case study that I'm not sure is all that insightful. I also think that this case study is more applicable than the guilt predictor that we were originally trying, and I am looking forward to working on this case study. I think that as this class is getting closer and closer to over, I think

I've learned a lot more in this class than I think I have, which is always a little bit of a weird feeling for me, especially when I feel a little lost almost every class. This week, I also didn't do quite as much work as I should have, or rather I spent a lot of time thinking about a possible case study, and not a ton of time doing exercises. This in combination with not doing as much work as I should have for this class meant that I don't really feel that I've produced as much as I feel like I should have this week.