Portfolio 2, part 1

Josephine Hillebrand Hansen

Script: https://github.com/JosephineHH/Assignment-2-2018/blob/master/Assignment2.Rmd

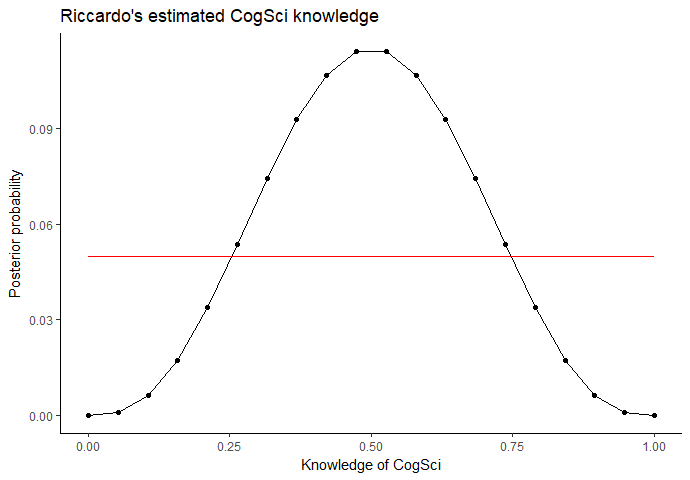
# Question 1) What is Riccardo’s estimated knowledge of CogSci?

We assume that we know nothing about Riccardo’s knowledge of CogSci. Therefore, we assume a flat prior, all outcomes are all equally likely.

Riccardo has a 50 % probability of knowing more than chance.

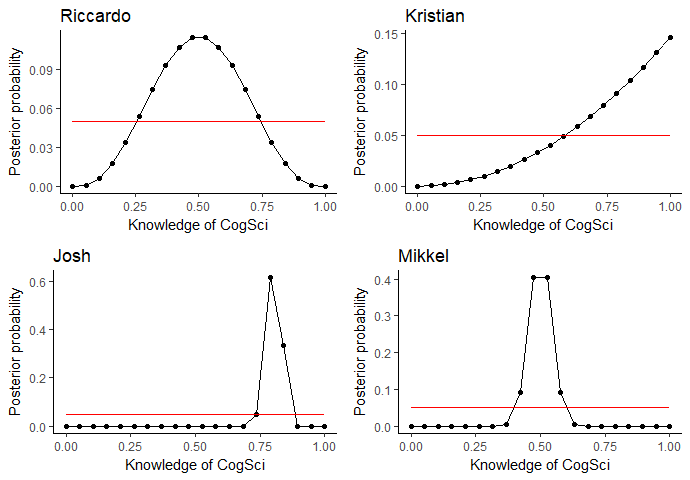
Using quadratic approximation, it was found that Riccardo’s posterior has a mean value of 0.5 with a standard deviation of 0.2 and the 89 % percentile interval is between 0.17 and 0.83.

Riccardo’s estimated knowledge of CogSci is plotted below. The black line shows the posterior probability and the red line shows the prior probability:



# Question 2) Estimate all the teachers’ knowledge of CogSci

The teachers knowledge of CogSci is plotted below:

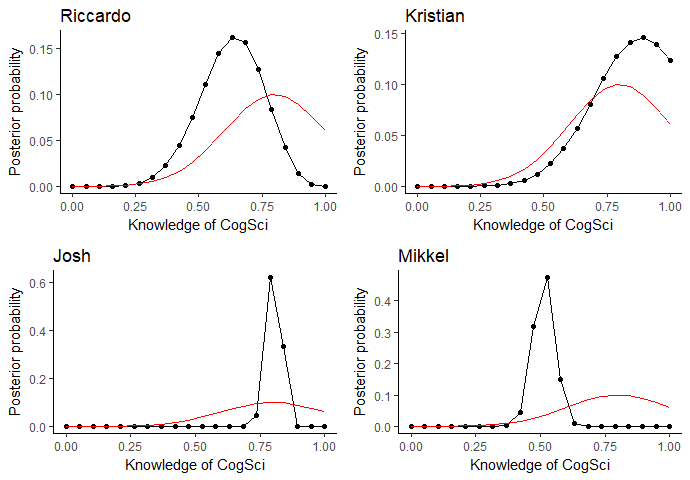


The following can be seen about the teacher’s CogSci knowledge from the plots:

Riccardo and Mikkel are both centered around 50 and both have a 50 % probability of performing above chance level. However, the distribution of Riccardo’s possible knowledge is much broader, so he still has the possibility of being worse or better, whereas we can be more sure that Mikkel is around his peak distribution. For Kristian and Josh. Josh is for sure good, whereas Kristian has the potential to know everything, but this is still unsure. Thus, Josh would be a safe choice.

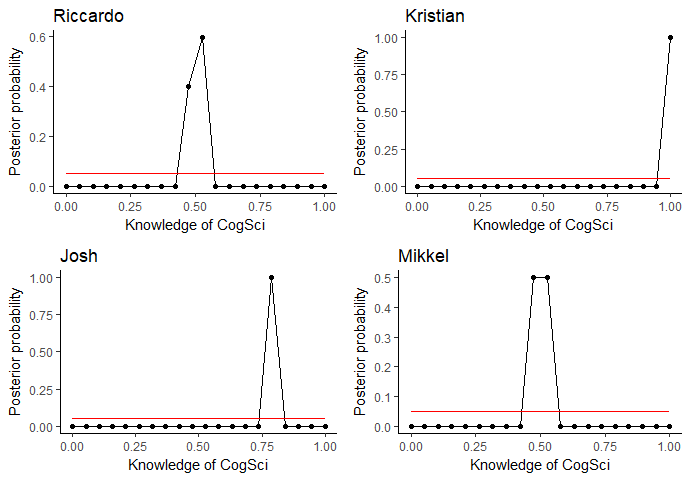
# Question 3) Change the prior to a normal distribution with a mean of 0.8 and a standard deviation of 0.2. How does this change the distribution?

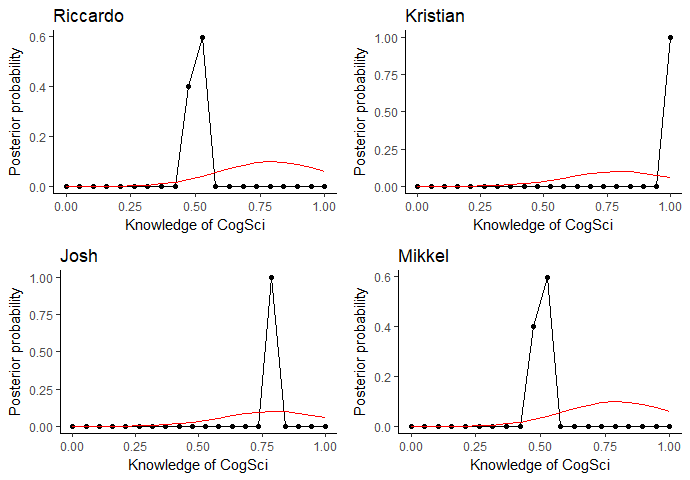
The prior was changed to a normal distribution with a mean of 0.8 and a standard deviation of 0.2. The plots for the 4 teachers can be seen below:



From the plots can be seen that the teachers who answered many questions (Josh and Mikkel) are not affected that much by the prior. However, the teachers with little data, Kristian and Riccardo has their posterior probabilities shifted a lot more.

# Question 4) Let us go and collect more data. Do we still see a difference between the results?





The prior hardly makes any difference when we collect this much data. This intuitively also makes sense, as there is a more data collected will give more weight to the data, than to the prior. The more data, the more we can rely on the data. The only difference is that Mikkel’s posterior distribution is shifted slightly towards the right.

# Question 5) Imagine that you are a skeptic and think that your teachers do not know anything about CogSci. How would you operationalize that belief?

Given that 0 is negative knowledge of CogSci, 0.5 is random chance and 1 is awesome cognitive science super powers and someone performing at chance level will therefore perform at random chance. Thus, we would not expect them to perform below 0.5, which would give them negative knowledge. Therefore, I would operationalize the belief by implementing a prior that is normally distributed with a mean of 0.5. If we expect them to know nothing, we expect them to perform at chance level, which is modelled by the normal distributed prior. We also expect very little chance of performing above and below, which would be normally distributed. #write about the SD

Portfolio 2, part 2

# Question 1) How is assessment of prediction performance different in Bayesian versus frequentist models

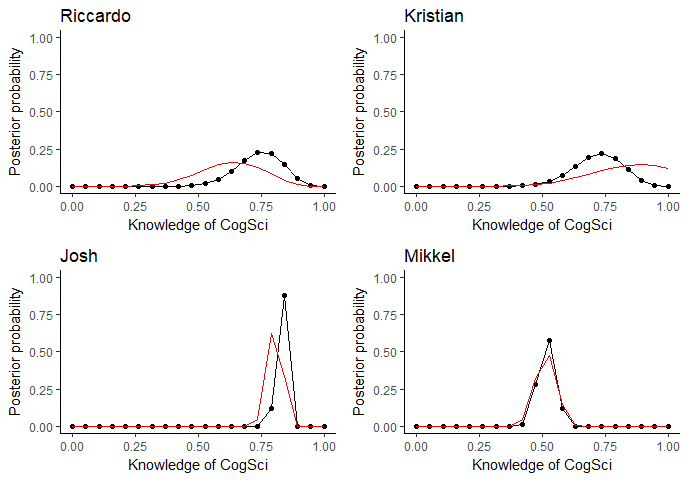
In frequentist models performance is predicted by taking the model and from the model predict values of new data. The prediction performance is then assessed by comparing the predicted values to actual values to evaluate how well the model performed.

In Bayesian models prediction performance is evaluated by putting the performance for last time as the prior and then the performance this time as the posterior and then evaluating how much the distribution changes. We therefore evaluate how well the model performs by looking at how much the distribution changes from the prior distribution to the posterior distribution.

In frequentist statistics we get a single predicted value for each data point, that we can evaluate against the measurements. In Baysian statistics we get a distribution of likelihood of different outcomes. This does not mean that we cannot evaluate the results, but we need to do it in a different way. One way would be to measure the difference in distributions.

# Question 2) Provide at least one plot and one written line discussing prediction errors for each of the teachers

I believe (or want to believe (-; ) that the university would not hire stupid teachers, and therefore I will use the posterior created last time using the optimal prior as our new prior.



It seems that Riccardo gets more likely to be better, Kristian more likely to be worse (actually Kristian and Riccardo and Kristian seem to be equally good now). Josh performs slightly worse and Mikkel performs about the same. The new data has the biggest impact on Riccardo’s and Kristian’s data as their prior was more uncertain (last time they answered fewer questions).