Computational Modeling - Assignment 2 - Part 2

Focusing on Predictions – Updates to Teachers Cognitive Science Knowledge

Caroline Casey

02/03/2018

Link to Github: https://github.com/Carolinecasey17/ComputationalModels.Sem4.Portfolio2.Part2.git

1) Write a paragraph discussing how assessment of prediction performance is different in Bayesian vs. frequentist models

In frequentist models, we calculate a specific value such as RMSE which we use as measure of prediction error. However, the RMSE value tells us nothing about the uncertainty inherent in the value, so there is no easy way to tell how probable the value is. In Bayesian assessment of prediction performance however, the output is a distribution of likely errors. It is much easier to see the uncertainty inherent in the model. In frequentist models, we make predictions of the model using only data from the current experiment, where in Bayesian models, we encode past knowledge of similar experiments as a prior, which we combine with the current experiment data to make a conclusion. So, in Bayesian statistics, we acknowledge that there is a prior probability.

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

First, we create a function to calculate posterior given number of hits, n of possibilities and size of grid to see prediction errors. We define the posteriors from our old data (previous portfolio) from all teachers to create a new model with the old data as our new posterior.

RESULTS

RICCARDO:

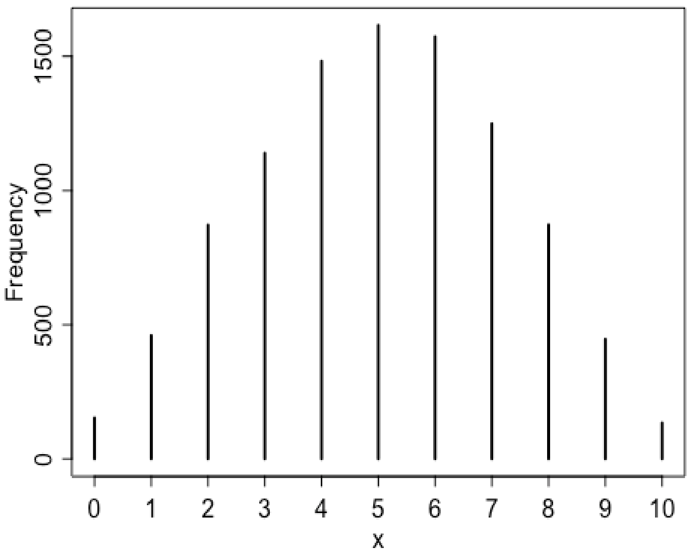

Figure 1 – New Riccardo Posterior Probability

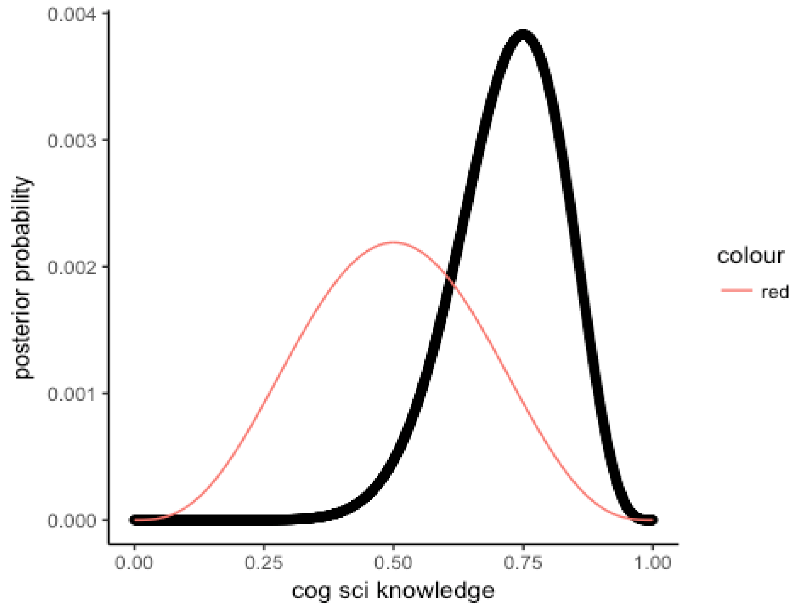

Figure 2 – The old Riccardo Posterior Probability (red line) vs the New Posterior Probability (black line)

Figure 1 illustrates Riccardo's predicted performance with the old Riccardo posterior used as prior. When checking the probability mass needed to include the result given, we needed PI = 0.9, which is a very large spread of probability intervals. This in turn indicates that the model did not predict performance very accurately even though the posterior was used as prior to update predictions on this data. Figure 3 illustrates the comparison of the old posterior probability and the new posterior probability of Riccardo.

KRISTIAN:

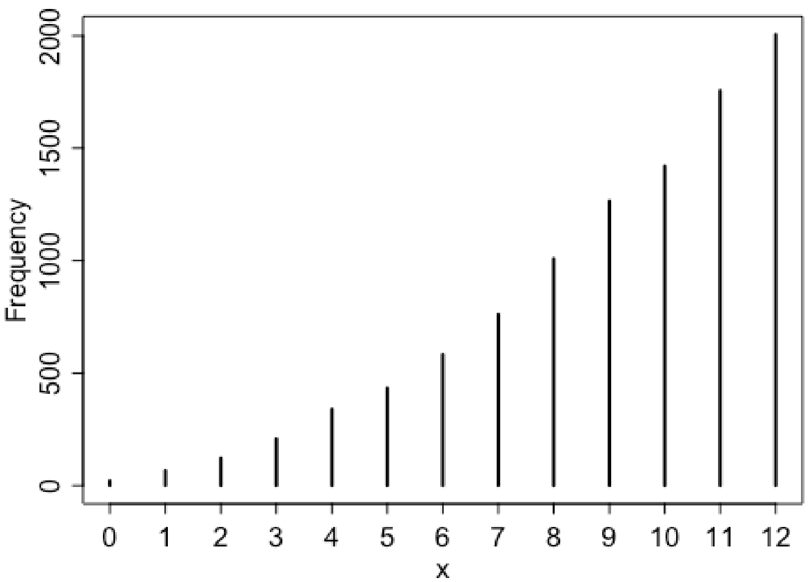

Figure 3 - New Kristian Posterior Probability

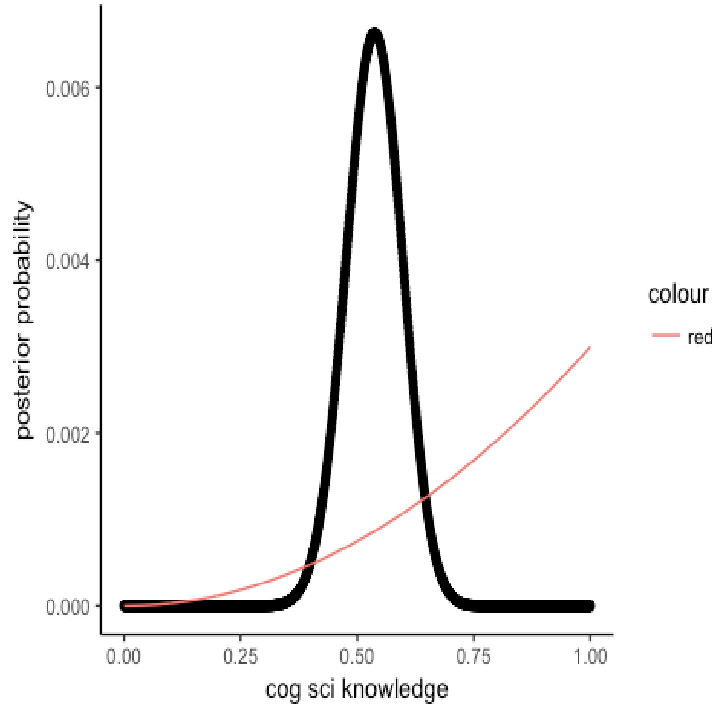

Figure 4 - The old Riccardo Posterior Probability (red line) vs the New Posterior Probability (black line)

Figure 3 illustrates Kristian's predicted performance with the old Kristian posterior used as prior. When checking the probability mass needed to include the result given, we needed HPDI = 0.7, which is a very large spread of probability intervals. This implies that the model did not predict performance very accurately even though the posterior was used as prior to update predictions on this data. Figure 4 illustrates the comparison of the old posterior probability and the new posterior probability of Kristian.

JOSH:

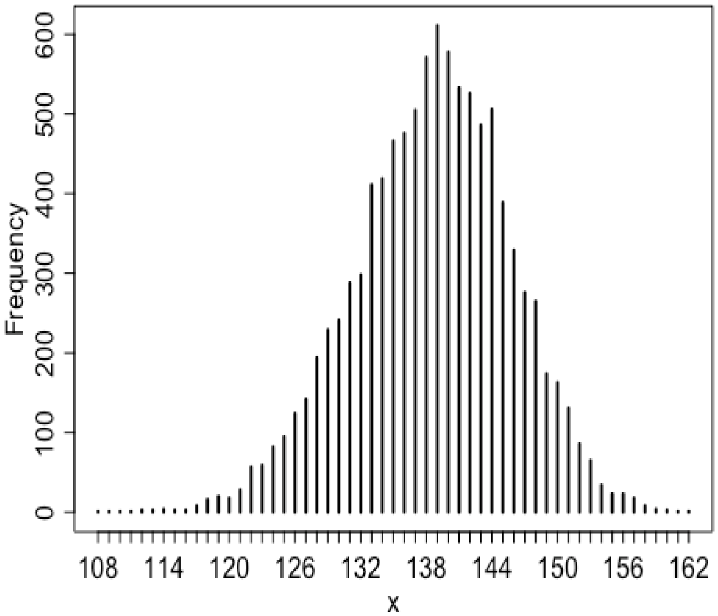

Figure 5 - New Josh Posterior Probability

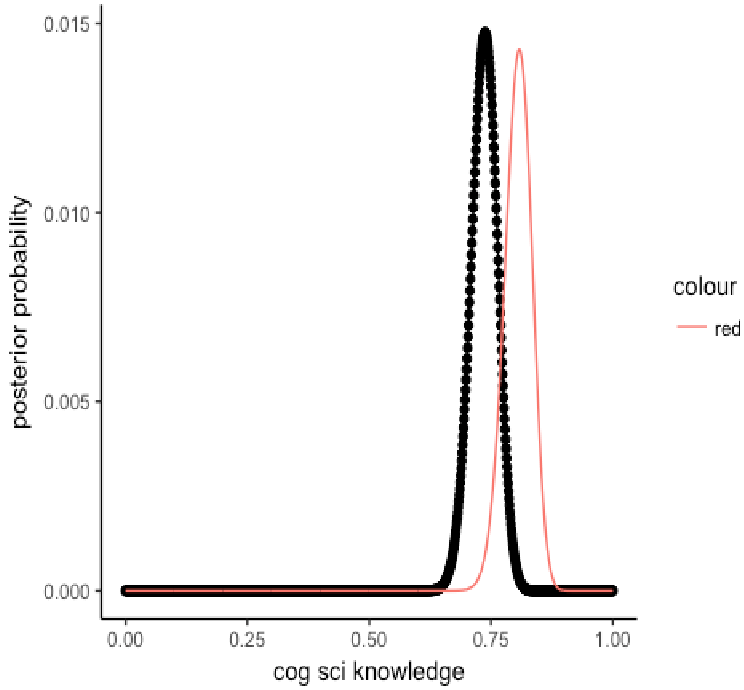

Figure 6 - The old Riccardo Posterior Probability (red line) vs the New Posterior Probability (black line)

In Figure 5, we see Josh' predicted performance with the old Josh posterior used as prior. When checking the probability mass needed to include the result given, we needed HPDI = 0.9, which is a very large spread of probability intervals and therefore indicates that the model did predict performance very poorly even though the posterior was used as prior to update predictions on this data. Figure 6 illustrates the comparison of the old posterior probability and the new posterior probability of Josh.

MIKKEL:

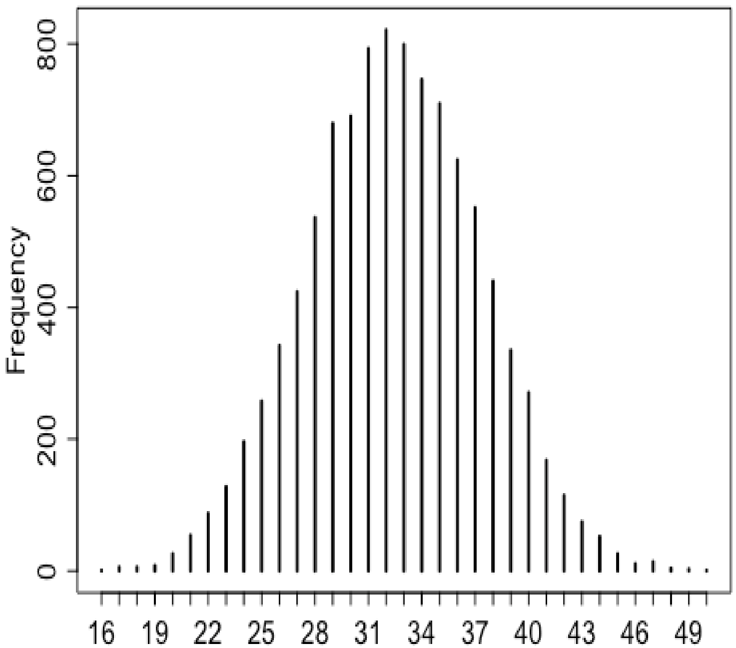

Figure 7 - New Mikkel Posterior Probability

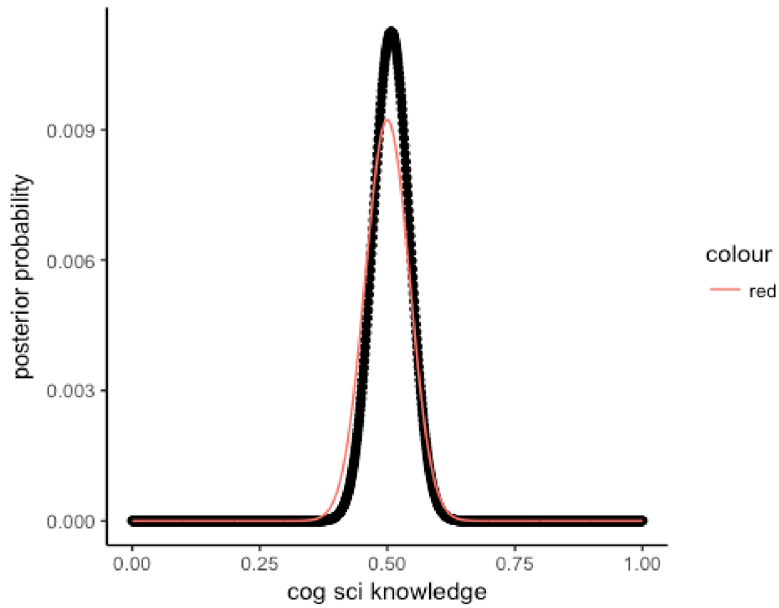

Figure 8 - The old Riccardo Posterior Probability (red line) vs the New Posterior Probability (black line)

In Figure 7, Mikkel's predicted performance with the old Josh posterior used as prior, is illustrated. When checking the probability mass needed to include the result given, we needed PI = 0.2, which is a very small spread of probability intervals and therefore indicates that the model did predict performance very suitable when the posterior was used as prior to update predictions on this data. Figure 8 illustrates the comparison of the old posterior probability and the new posterior probability of Mikkel.

CONCLUSION:

In conclusion, Mikkel's data was predicted most accurately (0.2), then Kristian's data (0.7) and then Josh' and Riccardo's as the worst predicted (0.9). Mikkel, Kristian and

Riccardo's results makes great sense if you compare to the percentage of their correct results from the old data to the new. However, Josh' data seems strange, since the comparison of the % of correct results of old and new data are close, but still the prediction is poor. This could be due to that Josh's data contains a relative big sample size, and therefore is rather precise in predictions. Thus, just a small change in accuracy will affect its precision. The overlapping of the distributions is not that big, even though the means are close. When we have a big sample size, we have a more accurate estimate of the model. We get a big error because the posteriors are quite narrow, and therefore the predictions we get are quite confident, which leads to poor predictions of new data.