# Question 1:

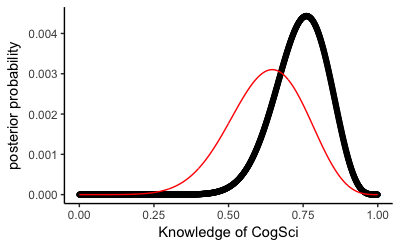
## Differences in assessment of prediction performance in Bayesian and Frequentist statistics.

The main difference is that, where we earlier have predicted a single outcome, we now have a whole distribution for the prediction (our predictive posterior). Before, we would have only predicted the most likely value, now we sample many values from a distribution. This means that we for now use more qualitative methods to decide when there is a difference or not, and whether the predictions was correct or not. So we see if the observed value seems to fit in the prediction distribution. (I can imagine there are ways to quantify this as well – for example something like mean difference between observed value and each of the predictions, or something like that). It also means that a measure of uncertainty takes a very different shape. We can’t no longer just look at how far from the prediction our result was, or how far they generally are (if we have more datapoints than 1). We also have to take into account the uncertainty in the distribution, and whether this is the reason that the observed value isn’t the most likely one – or if it is because the model predicted poorly (i.e. that the underlying distribution was different).

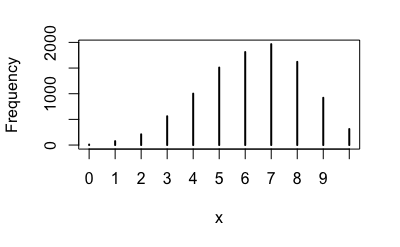
# Question 2:

I have for each teacher used the old posterior as predictor (using the optimistic prior as original prior), and used grid approximation to calculate a new posterior, using the new data as likelihood. I then qualitatively assess whether there has been a large change to the knowledge estimates.  
I have also made a predictive posterior and assess whether the observed amounts of correct answers seem realistic.   
Lastly, I have plotted the difference between the predictions and the observed value to look for a skew in error. The output has been standardized by dividing with the amount of questions asked.

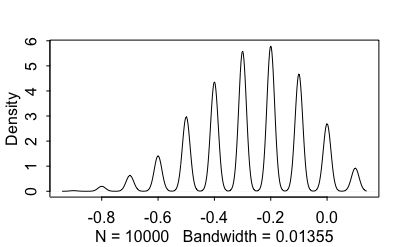
## Riccardo:



It can be seen that the parameter estimate increases distinctly, and also grows more certain, with the inclusion of the new data.

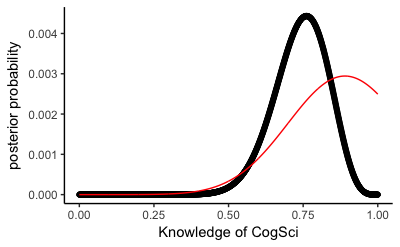


Riccardo got 9 correct answers. This is a possible but improbable result, positively skewed.

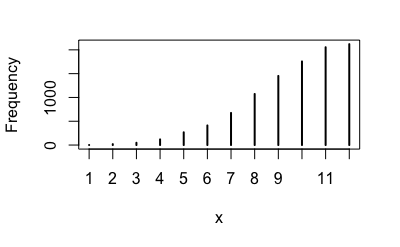


Here it can be seen that most predicted results are below the observed, indicating that Riccardo’s knowledge was underestimated by the model.

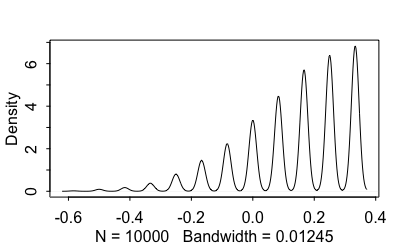
## Kristian:



For Kristian, the estimate gets lower and much more certain with the inclusion of the new data.

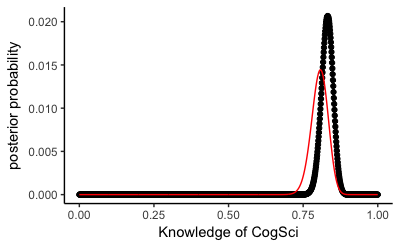


Kristian got 8 correct answers. This seems to be a slightly unlikely result.

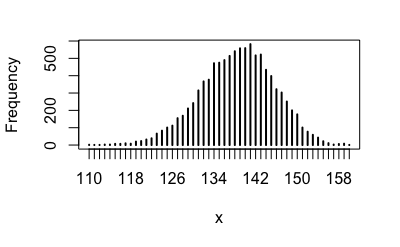


Here it can be seen that most predictions are above Kristian’s score, indicating that he has been overestimated slightly.

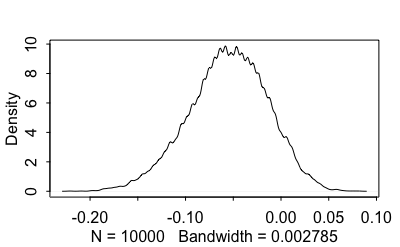
## Josh:



It can be seen that the parameter estimate increases distinctly, and also grows more certain, with the inclusion of the new data. Slightly less here than in Riccardo’s example.

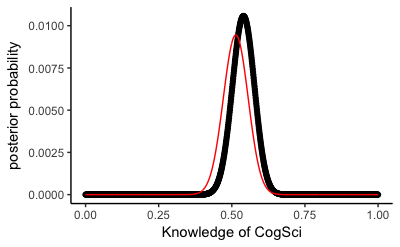


Josh got 148 correct answers. This is a possible but improbable result, again positively skewed.

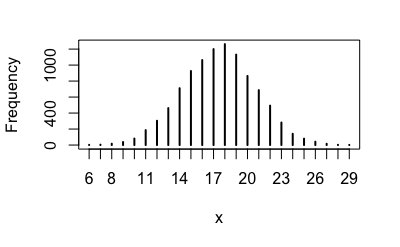


Here it can be seen that most predicted results are below the observed, indicating that Josh’s knowledge was underestimated by the model, here even more strongly than in Riccardo’s instance.

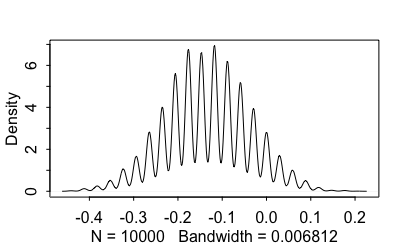
## Mikkel:



It can be seen that the parameter estimate increases distinctly, and also grows more certain, with the inclusion of the new data. Slightly less here than in Riccardo’s example.



Mikkel got 22 correct answers. This is a possible but improbable result, again positively skewed.



Here it can be seen that most predicted results are below the observed, indicating that Mikkel’s knowledge was underestimated by the model, here even more strongly than in Riccardo’s instance.

## Summary

The inclusion of the new data in all cases reduces uncertainty. In all cases except Kristian, the model has underestimated the knowledge rates of the teachers. For Kristian it is opposite, possibly because his prior was not normally distributed, and only grew higher in the extreme case of 100% knowledge.