NCRM Intro to BDA 4 Bayes rule (contd)

Now, our choice of non informative priors might not be the most appropriate one in this case. Let's say we are doing this analysis in July and we know that it is very, very unlikely to be cold outside, we can adjust our priors to be slightly more informative, we can say that the prior probability of it being hot outside is 90%. The probability of it being warm outside is 9%. And the probability of it being cold outside is only 1%. Please pause the video now, and try calculating the posterior distribution for the parameter values using these new priors and our old likelihood function.

Here are the results of these calculations. To obtain this table, I multiplied the prior probability of it being cold by the likelihood associated with being cold and put that number down in the third row of this table. I multiplied the prior probablility of it being warm outside, by the likelihood associated with it being a warm outside and wrote down that number in this third cell in the third row, I repeated the same operation with the values in the last column of this table. Once obtained these products, I normalised them to add up to one, and that gave me the posterior probabilities that you can see in the bottom row of this table.

Let's try plotting this information. So here on the left hand side, you can see the prior probabilities and the posterior probabilities that we obtained in the first analysis. On the right hand side, you can see the priors and the posteriors that we have obtained just now. So in the first case, here, you can see that the posterior reflects the relative values of the likelihoods, we can see that our in our posterior beliefs, the probability of it being cold is significantly larger than the probability of it being a warm, which is also larger than the probability of it being hot.

When we use an informative prior, then the posterior distribution is in some way, a compromise between that informative prior and the likelihood. In our curriculum in these new results, the most likely category, the most likely value of the weather parameter is warm. The second most likely value of the parameter, the weather parameter is hot, and the third most likely, value of this parameter is cold.

Now, let's take this example even further. Let's change the data that we have at hand. Instead of observing four persons wearing coats, and one person wearing sandals, let's say, we observed two persons were wearing coats, and five persons wearing sandals. Please pause the video now, and try calculating the posterior probabilities for different data types using these data and the non informative priors.

As before, the first step is to extract the variable values from the description of what we have observed through the window. The value of y the number of people wearing coats is two, and the total number of people observed is seven. Once I apply the probability mass function for the binomial distribution, with the parameters specified in our likelihood function, and the data, I'm obtaining the following values for the likelihood of being cold being warm and being hot.

Here are the results of computing the posterior in this case. If I, if you plot the results of these analysis, side by side, for two different data sets, we can see that the direction the the information contained in the data, does in fact change the posterior beliefs about the weather. And that shouldn't be of surprise, we hit roughly the same number of observations in dataset one and in dataset two. We observed five people in the first case and seven people in the second case.

Let's try looking at the situation when our data sets have the same share. So the people are wearing coats and sandals, but differ in the amount of observations they have. We should expect that when we have more observations, the evidence that we have in this data set should have a stronger effect on our assir than in the case when our data set is smaller, and contains less evidence for making the conclusions. And this is exactly what what's happening here. So here on the chart, you can see for analysis, the top row shows the analysis that we performed with non informative priors. And the bottom row shows the analysis that we performed with informative prayers, the ones that we described earlier. The left panels contain the analysis that we performed with the data suggests, according to which, we had four people wearing coats and one person wearing sandals. The analysis, shown in the right column, are based on the data according to which 40 people were wearing coats, and 10 people or were wearing sandals. So the data used to do it from analysis in the show in the right column provides a much stronger, much larger amount of evidence for making conclusions about the weather outside. And as you can see here, in both cases, it did update our beliefs about the weather, much stronger than the data with fewer observations. And in both cases, we see a roughly the same posterior distribution, which should tell you that when the evidence is strong, the initial prior is less likely significantly less likely to affect your conclusions than when you have a much lower amount of evidence.