3.18

Intro

In this exercise we will perform Bayesian **hypothesis testing** in a coin tossing problem. The details of Bayes factor as a technique for model selection are descbribed in page 165, in the chapter of Bayesian statistics. Also, there is an similar example showing how to use Bayes factor for the coin problem under the same hypothesis in the same page. The only difference is that it works with the sequence of results of the coin tosses (i.e. $D = \{head, head, tail, head, ...\}$) instead of the summary statistics N, N_1. Since Bayes factor uses the marginal likelihood, the choice of the dataset (original or summary statistics) gives different results, as we discussed in the previous exercise.

For those that never heard about hyposthesis testing, here is a quick overview. When we want to test an intersting hypothesis about our dataset (alternative hypothesis), it is common to compare it with a more conservative one (null hypothesis). The comparison is made using the data and there is more than one way to do it. The most popular way is using p-value, which is a technique from frequentist statistics. In Bayesian statistics, we can use the Bayes factor to make our selection and this will be done in this exercise. Since both results ($N = 10, N_1 = 9$ and $N = 100, N_1 = 90$) have a high proportion of heads, we might expect that Bayes factor will favour the alternative hypothesis (the coin can have any bias).

Solution

The Bayes factor is defined by the following equation:

$$BF_{1,0} \triangleq \frac{p(D|M_1)}{p(D|M_0)} \tag{1}$$

where D is the data and M_0, M_1 are the null and alternative hypothesis, respectively. From the last exercise, we know that $p(D|M_1) = \frac{1}{N+1}$. So, now we need to calculate the marginal likelihood for the null hypothesis.

$$p(D|M_0) = \int_0^1 p(D|\theta)p(\theta|M_0)d\theta = p(D|\theta = 0,5) = \binom{N}{N_1} 0.5^{N_1} 0.5^{N-N_1} = \binom{N}{N_1} 0.5^N \tag{2}$$

In Equation 2, we used the fair coin hypothesis to simplify the integral ($p(\theta|M_0) = \delta(\theta - 0.5)$). So the final expression for our Bayes factor is:

$$BF_{1,0} = \frac{1}{N+1} \frac{1}{\binom{N}{N_1} 0.5^N} = \frac{2^N}{(N+1)\binom{N}{N_1}} \tag{3}$$

Now, let's perform the actual calculations. For the first dataset, we have $N = 10, N_1 = 9$. Therefore:

$$BF_{1,0} = \frac{2^{10}}{(10+1)\binom{10}{9}} = 9.31 \tag{4}$$

From Jeffreys scale of evidence (page 165), we have moderate evidence to prefer M_1 over M_0.

For the second dataset, we have $N = 100, N_1 = 90$. Therefore:

$$BF_{1,0} = \frac{2^{100}}{(100+1)\binom{100}{90}} = 7.25 \times 10^{14} \tag{5}$$

From Jeffreys scale of evidence, we have decisive evidence to prefer M_1 over M_0.

Conclusion

In this question we performed hypothesis testing using bayesian model selection and the Bayes factor. We used the result of exercise 3.17 to help us to calculate the marginal likelihood ratio. As we suspected in the beginning, the high proportion of heads in the results favored the alternative hypothesis over the null. In another words, our data support the hypothesis that the coin can have any bias much more than support the hypothesis of a fair coin.

Another important point is that despite having the same proportions ($\frac{N_1}{N} = \frac{9}{10} = \frac{90}{100} = 90\%$) the second dataset presented a much more stronger evidence towards the alternative hypothesis than the first. This make us suspect that having more datapoints following a given rule help us to pinpoint the best model, which is very reasonable with all other results of this chapter.