Solutions Chapter 2 - The subjective worlds of Frequentist and Bayesian statistics

Alexandra Hotti

September 2019

1 Assignment 2.1 - The deterministic nature of random coin throwing

Suppose that, in an idealised world, the ultimate fate of a thrown coin heads or tails is deterministically given by the angle at which you throw the coin and its height above a table. Also in this ideal world, the heights and angles are discrete. However, the system is chaotic (highly sensitive to initial conditions), and the results of throwing a coin at a given angle and height are shown in Table 2.1.

Table 1: The results of a coin throw from a given angle and height above a table.

Angle (degrees)	0.2	0.4	0.6	0.8	1
0	T	H	T	T	H
45	H	T	T	T	T
90	H	H	T	T	H
135	H	H	T	H	T
180	H	H	T	H	H
225	H	T	H	T	T
270	H	T	T	T	H
315	T	H	H	T	T

1.1 Assignment 2.1.1

Suppose that all combinations of angles and heights are equally likely to be chosen. What is the probability that the coin lands heads up?

Answer: To calculate the probability of head we simply divide the number of times the coin will land on heads with the total number of coin configurations.

$$Pr(H) = \frac{\#heads}{\#outcomes} = \frac{19}{40} = 0.475$$

1.2 Assignment 2.1.2

Now suppose that some combinations of angles and heights are more likely to be chosen than others, with the probabilities shown in Table 2. What are the new probabilities that the coin lands heads up?

Table 2: The probability that a given person throws a coin at a particular angle, and at a certain height above a table.

Angle (degrees)	0.2	0.4	0.6	0.8	1
0	0.05	0.03	0.02	0.04	0.04
45	0.03	0.02	0.01	0.05	0.02
90	0.05	0.03	0.01	0.03	0.02
135	0.02	0.03	0.04	0.00	0.04
180	0.03	0.02	0.02	0.00	0.03
225	0.00	0.01	0.04	0.03	0.02
270	0.03	0.00	0.03	0.01	0.04
315	0.02	0.03	0.03	0.02	0.01

Answer: Now we weight the outcomes in Table 1 with the probabilities in Table 2. From the calculations below we find that: Pr(H) = 0.5

$$Pr(H) = \frac{\sum_i^{noHeads} prob_i}{\sum_j^{noOutcomes} prob_j} =$$

$$\frac{0.03+0.04+0.03+0.05+0.03+0.02+0.02+0.03+0+0.03+0.02+0+0.03+0+0.04+0.03+0+0.04+0.03+0.03}{1} =$$

$$= \frac{0.5}{1} = 0.5$$

1.3 Assignment 2.1.3

We force the coin-thrower to throw the coin at an angle of 45 degrees. What is the probability that the coin lands heads up?

Answer: Now we are constrained to have an angle of 45 degrees. Thus, we only weight the head coin throws with an angle by 45 degrees by the the probabilities in Table 2.

$$Pr(H) = \frac{0.03}{0.03+0.02+0.01+0.05+0.02} = 0.23076923076 \approx 0.23$$

1.4 Assignment 2.1.4

We force the coin-thrower to throw the coin at a height of 0.2m. What is the probability that the coin lands heads up?

Answer: Now we are constrained in a similar manner as in the previous question.

$$Pr(H) = 0.69565217391 \approx 0.7$$

1.5 Assignment 2.1.5

If we constrained the angle and height to be fixed, what would happen in repetitions of the same experiment?

Answer: There is only one possible outcome, so the coin will always land with the same side up.

2 Assignment 2.3 - Model Choice

Suppose that you have been given the data contained in *subjective_overfitShort.csv* and are asked to find a 'good' statistical model to fit the (x, y) data.

2.1 Assignment 2.3.1

Fit a linear regression model using least squares. How reasonable is the fit?

Answer: The results can be found in Figure 1. The linear regression model is the blue line. Evidently a line is not a good choice as a model for these data-points. However, the situation could change given that we acquire more points.

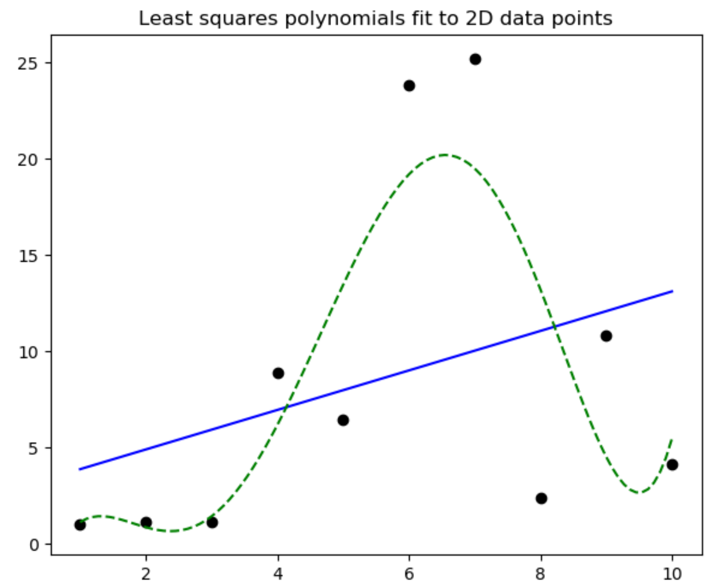

Figure 1: Two least square polynomials fit to a data set using a least square error of the residuals.

2.2 Assignment 2.3.2

. Fit a quintic (powers up to the fifth) model to the data. How does its fit compare to that of the linear model?

Answer: The results can be found in Figure 1. The quintic regression model is the green dotted line. Now the model fits the data very closely. Since we have so few data points it could be that the model would we very different given additional samples.

2.3 Assignment 2.3.3 2.3.4

. You are now given new data contained within *subjective_overfitLong.csv*. This contains data on 1000 replications of the same experiment, where the x values are held fixed. Using the least squares fits from the first part of this question, compare the performance of the linear regression model with that of the quintic model. Which of the two models do you prefer, and why?

Answer: The mean squared error of the residuals of the linear regression model fitted to 100 data points was approximately 12. The same value for the quintic model fitted to 100 data points was approximately 13.5. Thereby, the simpler model is less overfitted to the data and thus becomes better at generalizing and classifying new data points. Therefore, the linear regression model is the better classifier.