CS 520: CoinBot

16:198:520

Instructor: Wes Cowan

The purpose of this assignment is to get you thinking probabilistically, and refresh you on basic probability mechanics and uses. You should be familiar with Bernoulli random variables, independence, Bayes' theorem, conditional probability, and the law of total probability, at the very least.

1 Preliminaries

1) Argue that if a coin has a probability of heads p, then out of n independent flips, if X is the number of heads, then

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}. \tag{1}$$

This is a classic result in probability theory, but ***why*** *is it true? Consider marginalizing, or expressing the simple event in terms of more complicated events.*

2) Someone gives you a coin that is either Coin A, or Coin B. Without prior knowledge, is there any reason to think one option is more plausible or likely than the other? What does this say about $\mathbb{P}(\text{Coin A})$ and $\mathbb{P}(\text{Coin B})$?

2 The Coin Bot

You are given a mysterious coin. It is either Coin A, which has a probability of 0.4 of giving heads and 0.6 of giving tails, or it is Coin B, which has a probability of 0.7 of giving heads and 0.3 of giving tails. Initially, you have no reason to think that either coin is more likely than the other. However, by flipping the coin multiple times, you plan on using the resulting flips (your data) to reason about which coin it is more likely to be, A or B.

Let $F_1, F_2, F_3, \ldots, F_n$ be the sequence of flips collected through time n (i.e., $F_t = 0$ means the tth flip was tails, $F_t = 1$ means the tth flip was heads). We may quantify our belief in either coin at any given point in time with the two functions

$$\begin{aligned} p_A(n) &= \mathbb{P}(\text{Coin} = A | F_1, F_2, \ldots, F_n) \\ p_B(n) &= \mathbb{P}(\text{Coin} = B | F_1, F_2, \ldots, F_n). \end{aligned} \tag{2}$$

3) Prove that, for any n, $p_A(n) + p_B(n) = 1$.

4) Prove that for any n, $p_A(n)$ and $p_B(n)$ depend only on *the number of heads and tails recorded* rather than the specific order of flips collected. Give an explicit formula for both in terms of the respective coin probabilities, and the total number of heads recorded in n flips, Heads_n.

Given a set of data $F_1, \ldots, F_n$, we can construct the following decision rule for guessing what the coin is: let $\text{Guess}_n = A$ if $p_A(n) \geq p_B(n)$, or take $\text{Guess}_n = B$ if $p_B(n) > p_A(n)$. *Is it actually possible for $p_A(n) = p_B(n)$?*

5) Simplify the test $p_B(n) > p_A(n)$ as much as possible to a simple test on the value of Heads_n.

6) What is $\mathbb{P}(\text{Guess}_n = B | \text{Coin} = A)$? That is, if the coin were *actually* A, what is the probability that after n flips, we would guess that the coin is B? What is $\mathbb{P}(\text{Guess}_n = A | \text{Coin} = B)$? *Hint: What can / should you marginalize on?* Give answers for $n = 5, 10, 100$.

7) Combining the results of the previous question, what is

$$\mathbb{P}\left(\text{Guess}_n \neq \text{Coin}\right)? \tag{3}$$

What should you marginalize on? What are your priors? Give answers for $n = 5, 10, 100$.

8) What is the smallest n such that $\mathbb{P}\left(\text{Guess}_n \neq \text{Coin}\right) \leq 0.1$? How does this inform how many times you ought to experiment by flipping?

7) The answer to the previous question can inform how many experiments you ought to do. But *once you have performed those experiments*, it is reasonable to ask - what is the probability that you guessed wrong? i.e., what is

$$\mathbb{P}\left(\text{Guess}_n \neq \text{Coin}|F_1, F_2, \ldots, F_n\right) \text{ or equivalently } \mathbb{P}\left(\text{Guess}_n \neq \text{Coin}|\text{Heads}_n\right)? \tag{4}$$

What should you marginalize on? Given the data, is Guess_n *random?*

9) What value of Heads_n maximizes $\mathbb{P}\left(\text{Guess}_n \neq \text{Coin}|\text{Heads}_n\right)$ as in the previous question? Initially, what is the probability of this value of Heads_n occurring? *(Marginalize!)*

10) Freeform question: Suppose that you know that the coin is Coin A, but after some (unknown) number of flips, someone is going to swap it for Coin B without telling you. Design a test CoinBot can use to try to determine when / if the coin has been swapped. What can you say about the probability of CoinBot declaring the coin has been swapped before it was actually swapped (false positives)? What can you say about how many flips it takes for CoinBot to realize the coin has been swapped?