COS 445 - PSet 3, Problem 3

Anakin Skywalker

March 31, 2021

Problem 3: Noisy Cascades

Part a

We wish to find the probability as a function of α that exactly the sequence S guesses will be observed, conditioned on the bag containing two red balls and one blue ball.

It is given that the bag contains two red balls and one blue ball. Since balls are randomly drawn from the bag each time, a given ball is red with probability $\frac{2}{3}$ and blue with probability $\frac{1}{3}$. However, because of the noise value α, it is possible for a ball's color to be announced differently. To compute the final probability, we need to multiply the probability that R balls are reported red by the probability that B balls are reported blue.

The first probability can be found by adding the probability that a red ball is actually reported as red with the probability that a blue ball is reported as red:

$$
\begin{aligned}
&= \frac{2}{3} \cdot (1 - \frac{\alpha}{2}) + \frac{1}{3} \cdot \frac{\alpha}{2} \\
&= \frac{2}{3} - \frac{2\alpha}{6} + \frac{\alpha}{6} \\
&= \frac{4 - \alpha}{6}
\end{aligned}
\tag{1}
$$

The second probability can be computed in a similar manner:

$$
\begin{aligned}
&= \frac{1}{3} \cdot (1 - \frac{\alpha}{2}) + \frac{2}{3} \cdot \frac{\alpha}{2} \\
&= \frac{1}{3} - \frac{\alpha}{6} + \frac{2\alpha}{6} \\
&= \frac{2 - \alpha}{6}
\end{aligned}
\tag{2}
$$

Therefore, the final probability that the sequence is observed conditioned on there being two red balls and one blue ball is:

$$
(\frac{4 - \alpha}{6})^R \cdot (\frac{2 - \alpha}{6})^B
\tag{3}
$$

Part b

We wish to find the probability that the bag contains two red balls and one blue ball, conditioned on having seen exactly the sequence S of revealed guesses and thereafter drawing a red ball from the bag.

We apply Bayes' rule in order to rewrite the probability as follows:

$$\frac{\left(\frac{4-\alpha}{6}\right)^{R+1} \cdot \left(\frac{2-\alpha}{6}\right)^{B} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \left(\frac{4-\alpha}{6}\right)^{R+1} \cdot \left(\frac{2-\alpha}{6}\right)^{B} + \frac{1}{2} \cdot \left(\frac{4-\alpha}{6}\right)^{B} \cdot \left(\frac{2-\alpha}{6}\right)^{R+1}} \tag{4}$$

The first part of the numerator is the probability S is observed conditioned on the bag containing two red balls and one blue ball, which was calculated in the previous part. This is multiplied by the probability that the bag contains two red balls and one blue ball, which is given as $\frac{1}{2}$. The denominator is simply the sum of the probability that S is observed conditioned on the bag containing two red balls and one blue ball, and the probability that S is observed when the bag contains one red ball and two blue balls. This denominator then gives the total probability that the exact sequence S is observed.

Part c

We wish to find how many more red revealed guesses than blue revealed guesses (or vice versa) would be needed before one's own draw is ignored.

To do so, we can compare the probability that there are two red balls and one blue ball conditioned on S with the probability that there are two blue balls and one red ball conditioned on S. We already know the first probability as it was calculated in the previous part, with the required change being to decrease the exponent for the probability an observed ball is red from $R+1$ to R (since we are not assuming that we draw a red ball anymore). The second probability is also easily determined following from Bayes' rule:

$$\frac{\left(\frac{4-\alpha}{6}\right)^{B} \cdot \left(\frac{2-\alpha}{6}\right)^{R} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \left(\frac{4-\alpha}{6}\right)^{R} \cdot \left(\frac{2-\alpha}{6}\right)^{B} + \frac{1}{2} \cdot \left(\frac{4-\alpha}{6}\right)^{B} \cdot \left(\frac{2-\alpha}{6}\right)^{R}} \tag{5}$$

We note that the denominator is identical, because it is still the probability that S is observed. Additionally, since it is equally likely that the bag contains two red balls and one blue ball or two blue balls and one red ball the numerator still contains $\frac{1}{2}$. While there are many ways to compare these two probabilities, the simplest is to divide one by the other. Because we want to know how many more red revealed guesses are required, we divide the probability that there are two red balls and one blue ball conditioned on S with the probability that there are two blue balls and one red ball conditioned on S:

$$
\begin{aligned}
&= \frac{\left(\frac{4-\alpha}{6}\right)^{R} \cdot \left(\frac{2-\alpha}{6}\right)^{B} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \left(\frac{4-\alpha}{6}\right)^{R} \cdot \left(\frac{2-\alpha}{6}\right)^{B} + \frac{1}{2} \cdot \left(\frac{4-\alpha}{6}\right)^{B} \cdot \left(\frac{2-\alpha}{6}\right)^{R}} \cdot \frac{\frac{1}{2} \cdot \left(\frac{4-\alpha}{6}\right)^{R} \cdot \left(\frac{2-\alpha}{6}\right)^{B} + \frac{1}{2} \cdot \left(\frac{4-\alpha}{6}\right)^{B} \cdot \left(\frac{2-\alpha}{6}\right)^{R}}{\left(\frac{4-\alpha}{6}\right)^{B} \cdot \left(\frac{2-\alpha}{6}\right)^{R} \cdot \frac{1}{2}} \\
&= \frac{1}{2} \cdot \frac{\left(\frac{4-\alpha}{6}\right)^{R} \cdot \left(\frac{2-\alpha}{6}\right)^{B}}{\left(\frac{4-\alpha}{6}\right)^{B} \cdot \left(\frac{2-\alpha}{6}\right)^{R}} \\
&= \frac{1}{2} \cdot \left(\frac{4-\alpha}{2+\alpha}\right)^{R-B}
\end{aligned}
\tag{6}
$$

Therefore, a player will ignore their own draw when $\frac{1}{2} \cdot \left(\frac{4-\alpha}{2+\alpha}\right)^{R-B} > 1$. We also wish to prove that when $\alpha = 1$ a cascade will never happen. We first rewrite the inequality as follows:

$$
\begin{aligned}
\frac{1}{2} \cdot \left(\frac{4-\alpha}{2+\alpha}\right)^{R-B} &> 1 \\
\left(\frac{4-\alpha}{2+\alpha}\right)^{R-B} &> 2
\end{aligned}
\tag{7}
$$

If we then set $\alpha = 1$, the inequality becomes the following:

$$
\begin{aligned}
\left(\frac{4-1}{2+1}\right)^{R-B} &> 2 \\
\left(\frac{3}{3}\right)^{R-B} &> 2 \\
1^{R-B} &> 2 \\
1 &> 2
\end{aligned}
\tag{8}
$$

This is a contradiction, meaning that a cascade will never happen when $\alpha = 1$. To prove that a cascade will eventually happen when $\alpha < 1$, we again rewrite the inequality:

$$
\begin{aligned}
\left(\frac{4-\alpha}{2+\alpha}\right)^{R-B} &> 2 \\
(R-B) \cdot \log_2\left(\frac{4-\alpha}{2+\alpha}\right) &> \log_2(2) \\
(R-B) \cdot \log_2\left(\frac{4-\alpha}{2+\alpha}\right) &> 1 \\
R - B &> \frac{1}{\log_2\left(\frac{4-\alpha}{2+\alpha}\right)}
\end{aligned}
\tag{9}
$$

Intuitively, when $\alpha < 1$ we observe that $R - B$ will be defined, meaning that the ultimate probability a sequence will cause a cascade is nonzero.

Part d

I don't know.

Part e

We wish to find the probability that the eventual cascade is correct when $\alpha < 1$.

Let $P(c)$ be the probability of a correct cascade and $P(i)$ be the probability of an incorrect cascade. We already know that $P(c) + P(i) = 1$. Furthermore, we can equate $P(c)$ to the following expression:

$$
\begin{aligned}
P(c) &= \sum_c P(s) \\
&= k \cdot P(i)
\end{aligned}
\tag{10}
$$

Where k is the revealed guess at step k in the sequence S. Therefore, the probability that the eventual cascade is correct when $\alpha < 1$ is equivalent to $k \cdot P(i)$.