Machine Learning Assignment 27

Maia Dimas

Problem 27-3

Suppose you toss a coin 10 times and get the result **HHHHT HHHHH**. From this result, you estimate that the coin is biased and generally lands on heads 90% of the time. But how sure can you be? Let's quantify it.

A: Compute the likelihood $P(\text{HHHHT HHHHH} \mid k)$ where $P(H) = k$. Remember that the likelihood is just the probability of getting the result HHHHT HHHHH under the assumption that $P(H) = k$. Your answer should be expressed in terms of k

$$
\begin{aligned}
P(\text{HHHHT HHHHH} \,|\, k) &= k \cdot k \cdot k \cdot k \cdot (1-k) \cdot k \cdot k \cdot k \cdot k \cdot k \\
&= k^9 \cdot (1-k) \\
&= k^9 - k^{10}
\end{aligned}
$$

B: Create a probability distribution $P(k|\text{ HHHHT HHHHH})$ that is proportional to the likelihood $P(\text{HHHHT HHHHH }|k)$.

In order to find the constant that the probability distribution is multiplied by, you must find the integral of the likelihood. Thus, the constant will be the inverse of that answer so that the integral of the probability distribution will work out to one.

$$\begin{aligned}
\int_0^1 c \cdot P(\text{HHHHT HHHHH} \,|\, k)\,\mathrm{d}k &= \int_0^1 c \cdot (k^9 - k^{10})\,\mathrm{d}k \\
&= c \cdot \left[\frac{k^{10}}{10} - \frac{k^{11}}{11}\right]_0^1 \quad \text{(Integrate)} \\
&= c \cdot \left(\frac{1}{10} - \frac{1}{11} + 0\right) \\
&= c \cdot \frac{1}{110}
\end{aligned}$$

$$\begin{aligned}
\int_0^1 c \cdot P(\text{HHHHT HHHHH} \,|\, k)\,\mathrm{d}k &= 1 \\
c \cdot \frac{1}{110} &= 1 \\
c &= 110
\end{aligned}$$

As $P(k \,|\, \text{HHHHT HHHHH}) = c \cdot P(\text{HHHHT HHHHH} \,|\, k)$, you know the integral of $P(k)$ will equal $c \cdot \frac{1}{110}$ and as it is known that $c = 110$, you can replace the c to be...

$$110 \cdot \frac{1}{110} = 1$$

Thus making $P(k)$ a proper probability distribution.

C: Using the prior distribution $\boldsymbol{P(k)} \sim \boldsymbol{U}[0,1]$, what was the prior probability that the coin was biased towards heads? In other words, what was $\boldsymbol{P(k > 0.5)}$?

With the prior distribution, there's about a 50% chance of the coin being biased towards heads.

$$
\begin{aligned}
P(k > 0.5) &= \int_{0.5}^{1} p(k)\,\mathrm{d}k \\
&= \int_{0.5}^{1} 1\,\mathrm{d}k \\
&= [k]_{0.5}^{1} \\
&= 1 - 0.5 \\
&= 0.5
\end{aligned}
$$

D:Using the posterior distribution $\boldsymbol{P(k|\text{HHHHT HHHHH})}$, what was the posterior probability that the coin was biased towards heads? In other words, what is $\boldsymbol{P(k{>}0.5|\text{HHHHT HHHHH})}$.

There's about a 99% chance that the coin was biased towards heads

$$
\begin{aligned}
P(k > 0.5|\text{HHHHT HHHHH}) &= \int_{0.5}^{1} 110 \cdot k^9 - k^{10}\,\mathrm{d}k \\
&= 110 \cdot \left[\frac{k^{10}}{10} - \frac{k^{11}}{11}\right]_{0.5}^{1} \qquad \text{(Intigrate)} \\
&= 110 \cdot \left(\frac{1}{10} - \frac{1}{11} - \left(\frac{(0.5)^{10}}{10} - \frac{(0.5)^{11}}{11}\right)\right) \\
&= 110 \cdot (0.009037642045) \\
&= 0.994140625
\end{aligned}
$$

E: Compare your answers in parts (c) and (d). Did the probability that the coin was biased towards heads increase or decrease, after observing the sequence of flips? Why does this make intuitive sense?

The probability increased. This just makes sense as the prior distribution has no data to go off of and was just a constant until data was added. After adding in all the heads, of course the chance that the coin will be heads would go up.

F: Using the posterior distribution, what is the most probable value of k? In other words, what is value of k at which $P(k|\text{HHHHT HHHHH})$ reaches a maximum? Show your work using the first or second derivative test.

The most likely value of k is 0.9

$$
\begin{aligned}
P(k|\text{HHHHT HHHHH}) &= 110 \cdot (k^9 - k^{10}) \\
P'(k|\text{HHHHT HHHHH}) &= 110 \cdot (9k^8 - 10k^9) \\
0 &= 110 \cdot (9k^8 - 10k^9) \\
0 &= 9k^8 - 10k^9 \\
k &= 0.9
\end{aligned}
$$

G: Why does your answer to (f) make sense? What's the intuition here?

It makes sense because the math says its true. It makes intuitional sense as the coin was flipped 10 times and was only tails once. That's 90% of the time

H: What is the probability that the bias k lies within 0.05 of your answer to part (g)? In other words, what is the probability that 0.85 < k <0.95?

Around a 41% chance that k lays between 0.85 and 0.95

$$
\begin{aligned}
P(0.85 < k < 0.95|\text{HHHHT HHHHH}) &= \int_{0.85}^{0.95} 110 \cdot k^9 - k^{10} \, \mathrm{d}k \\
&= 110 \cdot \left[\frac{k^{10}}{10} - \frac{k^{11}}{11} \right]_{0.85}^{0.95} \quad \text{(Integrate)} \\
&= 110 \cdot \left(\frac{0.95^{10}}{10} - \frac{0.95^{11}}{11} - \left(\frac{0.85^{10}}{10} - \frac{0.85^{11}}{11} \right) \right)
\end{aligned}
$$

I: Fill in the blank: you can be 99% sure that P(H) is at least ___.

Around 0.53

$$\begin{aligned}
0.99 &= \int_v^1 110 \cdot k^9 - k^{10} \,\mathrm{d}k \\
&= 110 \cdot \left[\frac{k^{10}}{10} - \frac{k^{11}}{11}\right]_v^1 \quad \text{(Integrate)} \\
&= 110 \cdot \left(\frac{1}{10} - \frac{1}{11} - \left(\frac{v^{10}}{10} - \frac{v^{11}}{11}\right)\right) \\
&= 10v^{11} - 11v^{10} + 1 \\
v &= 0.530184 \quad \text{(Wolfram Alpha)}
\end{aligned}$$