Forward probability

Question 1

The probability of rain on Saturday and Sunday are 50% and 20%, respectively. Any rain will be light with 90%, and heavy with 10% probability on Saturday. Any rain on Sunday will be light.
Q_1: What is the probability of light rain on both days?
Q_2: What is the probability of raining during the weekend?

Question 2

One bag of candy contains 3 pieces of taffy and 4 pieces of caramel, while the other contains 1 piece of taffy and 5 pieces of caramel. We draw one piece of candy from each bag.
Q_1: With what probability are the two drawn pieces of candy different?
Q_2: What if we draw them from the same (but randomly chosen) bag?

Question 3

In a simple game, we flip a coin 3-times. The reward we get depends on the outcome in the following way: 3 heads → win $100, 2 heads → win $40, 1 head → nothing, 0 heads → lose $200.
Q: What is the expectation value and the standard deviation of the reward?

Question 4

10 selected Hogwarts students are randomly lined up for questioning.
Q_1: What is the probability that Potter, Granger and Weasley are standing next to each other?
Q_2: What if they are standing in a circle?

Question 5

5 male dancers, a, b, c, d, e and 5 female dancers $\alpha, \beta, \gamma, \delta, \epsilon$ form dancing couples randomly.
Q: What is the probability that c dances with γ?

Question 6

The 21 Insight Fellows are grouped into 3 equal-size groups randomly.
Q: What is the probability that Derrick and Gaurav end up in the same group?

Question 7

The Krampus distributes 10 pieces of candy randomly between the stockings of 4 kids, A, B, C, D.
Q: What is the probability that A doesn't get any?
(The answer depends on the details of the randomization procedure.)

Questions 8

We keep throwing a fair, ten-sided die, with numbers $1, 2, 3, \ldots 10$ on its sides, and count the number of times "1" comes up.
Q_1: What is the probability that we get two "1"s in the first twenty throws?
Q_2: What is the probability that we get the first "1" at the tenth throw?
Q_3: What is the probability that we get the third "1" at the thirtieth throw?

Probabilistic inference

Problem

- **Model comparison**: Given observations (aka data) $D = (d_1, d_2, \dots d_N)$, how plausible is each of the proposed models $\{M_1, M_2, \dots\}$?
- **Prediction**: Given the above calculated plausibility of each model, what is the prediction for the next observation, d_{N+1}?

Solution

1. **Prior probability**: Think hard about the details of each model M_i, irrespective of the data D, and come up with prior probabilities, $\boxed{\mathcal{P}(M_1), \mathcal{P}(M_2)\dots}$. (Here $P(M_i)$ is the probability with which we *think* model i is true, disregarding the observations D.)

2. **Forward probability**: Think hard about the details of each model M_i, and calculate the probability of observing the data D, assuming that model i is true, $\boxed{\mathcal{P}(D|M_i),\ i = 1, 2, \dots}$.
(A common assumption is that D is composed of independent, identically distributed (iid) observations d_j, in which case we need to calculate $\boxed{\mathcal{P}(d|M_i)}$ for all d individual observations, and combine them to get $\mathcal{P}(D|M_i) = \prod_{j=1}^{N} \mathcal{P}(d_j|M_i)$)

3. **Posterior probability**: Bayes theorem allows relating the probability of a model M_i (given observations D) to the probabilities of observations D (given models $\{M_k\},\ k = 1, 2, \dots$),

$$\boxed{\mathcal{P}(M_i|D)} = \frac{\mathcal{P}(D|M_i)\mathcal{P}(M_i)}{\mathcal{P}(D)} = \frac{\mathcal{P}(D|M_i)\mathcal{P}(M_i)}{\sum_k \mathcal{P}(D|M_k)\mathcal{P}(M_k)} \propto \mathcal{P}(D|M_i)\mathcal{P}(M_i).$$

   There are different strategies making sense of this result:

   (a) **Maximum likelihood estimate**: Pick the model that has the highest probability of producing the data (disregarding of its prior probability)

$$\hat{M}_{\mathrm{MLE}} = \arg\max_M \Big[\mathcal{P}(D|M)\Big].$$

   (b) **Maximum a posteriori estimate**: Pick the model with the highest posterior probability

$$\hat{M}_{\mathrm{MAP}} = \arg\max_M \Big[\mathcal{P}(M|D)\Big] = \arg\max_M \Big[\mathcal{P}(D|M)\mathcal{P}(M)\Big].$$

   (c) **Full distribution**: Report the full set of posterior probabilities,

$$\Big\{\mathcal{P}(M_1|D),\, \mathcal{P}(M_2|D),\, \dots\Big\}.$$

   Or report summary statistics of this distribution, if it makes sense.

4. **Prediction**: Averaging the forward probabilities of all models weighted by their posterior probabilities give the probability of the next observation $d_{N+1} = d$,

$$\boxed{\mathcal{P}(d|D)} = \sum_i \mathcal{P}(d, M_i|D) = \sum_i \mathcal{P}(d|D, M_i)P(M_i|D) = \sum_i \mathcal{P}(d|M_i)\mathcal{P}(M_i|D).$$

   This is often approximated by taking only the most plausible model into account (either $\hat{M}_{\mathrm{MLE}}$ or $\hat{M}_{\mathrm{MAP}}$) to get $\mathcal{P}(d|D) \approx \mathcal{P}(d|\hat{M})$, resulting in a fast but often less robust result.

Examples

- Two discrete models:
  $D = \{$series of daily maximum temperatures$\}$, $M_1 =$ "It's December.", $M_2 =$ "It's January.".
- Many discrete models:
  $D = \{$series of drawn cards at a blackjack table$\}$, $M_k =$ "The foot contains k deck.", $k = 1, 2, 3, \ldots$
- One set of continuously many models:
  $D = \{$series of coin toss outcomes$\}$, $M_p =$ "A coin toss gives heads with probability p.", $p \in [0, 1]$
- Continuously many models against one discrete:
  $D = \{$series of coin toss outcomes$\}$, $M_{\text{null}} =$ "Coin is fair." $M_p =$ (see above)

Model comparison

Question 9

A blood test for an disease gives positive result on your sample.
Q: What is the probability that the sample actually has the disease? (Let's assume that the prevalence of disease is one in 10,000, and the test has a false positive rate of 1%, and no false negative rate.)

Question 10

You have a new burglar alarm installed in your house. It is advertised to be able to detect 99.9% of all burglaries. From the user manual you also learn that earthquakes have a tendency to set it off with 20% probability.
Q_1: One day, while you are at work, you get an automated message from your new burglar alarm saying it went off. What is the probability that a break-in happened? (Let's assume a break-in rate of one in five years, and an earthquake rate of two per year.)
Q_2: A minute after this, you learn that there was an earthquake near your house. Now, what is the probability that a break-in happened?

Prediction

Question 11

We pick one of the following two coins at random, and toss it five times. Coin 1 is a fair coin with tail and head sides, but coin 2 is a trick coin, with heads on both sides.
Q_1: For each coin, what is the probability of tossing five heads?
Q_2: Given that five heads came up, what is the probability that we chose the trick coin?
Q_3: Given that five heads came up, what is the probability of getting a head for the sixth toss?

Solutions

Solution 1

From the text directly:

$$\mathcal{P}(\text{rain Sat}) = 0.5, \qquad \mathcal{P}(\text{light}|\text{rain Sat}) = 0.9, \qquad \mathcal{P}(\text{heavy}|\text{rain Sat}) = 0.1$$
$$\mathcal{P}(\text{rain Sun}) = 0.2, \qquad \mathcal{P}(\text{light}|\text{rain Sun}) = 1$$

Using $\mathcal{P}(A \text{ AND } B) = \mathcal{P}(A, B) = \mathcal{P}(A|B) \times \mathcal{P}(B)$, we can write the light rain probabilities for each day separately:

$$\begin{aligned}
\mathcal{P}(\text{light}, \text{rain Sat}) &= \mathcal{P}(\text{light}|\text{rain Sat}) \times \mathcal{P}(\text{rain Sat}) = 0.9 \times 0.5 = 0.45 \\
\mathcal{P}(\text{light}, \text{rain Sun}) &= \mathcal{P}(\text{light}|\text{rain Sun}) \times \mathcal{P}(\text{rain Sun}) = 1 \times 0.2 = 0.2
\end{aligned}$$

These being independent events, we can use $\mathcal{P}(A, B) = \mathcal{P}(A) \times \mathcal{P}(B)$, and write the probability of light rain on both days:

$$\mathcal{P}(\text{light}, \text{rain Sat}) \times \mathcal{P}(\text{light}, \text{rain Sun}) = 0.45 \times 0.2 = 0.09 = 9\% \quad (\mathrm{A}_1)$$

Using $\mathcal{P}(\text{NOT } A) = 1 - \mathcal{P}(A)$, we can write

$$\begin{aligned}
\mathcal{P}(\text{rain during weekend}) &= 1 - \mathcal{P}(\text{no rain at all}) \\
&= 1 - \mathcal{P}(\text{no rain Sat}) \times \mathcal{P}(\text{no rain Sun}) \\
&= 1 - (1 - 0.5) \times (1 - 0.2) = 1 - 0.5 \times 0.8 = 1 - 0.4 = 0.6 = 60\% \quad (\mathrm{A}_2)
\end{aligned}$$

Solution 2

Let's label the bags A (containing $3t, 4c$) and B (containing $1t, 5c$), and label the draws $A1, B1$. The probability of getting different pieces of candy is

$$\begin{aligned}
\mathcal{P}(\text{diff}) &= \mathcal{P}([A1 = t, B1 = c] \text{ OR } [A1 = c, B1 = t]) \\
&= \mathcal{P}(A1 = t)\mathcal{P}(B1 = c) + \mathcal{P}(A1 = c)\mathcal{P}(B1 = t) \\
&= \frac{3}{7} \times \frac{5}{6} + \frac{4}{7} \times \frac{1}{6} = \frac{19}{42} \quad (\mathrm{A}_1)
\end{aligned}$$

For the second question, let's distinguish the two possible cases: bag A and bag B. If bag A is chosen, the probability of different candies is

$$\begin{aligned}
\mathcal{P}(\text{diff}|A) &= \mathcal{P}(\text{A2 = t, A1 = c}) + \mathcal{P}(\text{A2 = c, A1 = t}) \\
&= \mathcal{P}(\text{A2 = t}|\text{A1 = c})\mathcal{P}(\text{A1 = c}) + \mathcal{P}(\text{A2 = c}|\text{A1 = t})\mathcal{P}(\text{A1 = t}) \\
&= \frac{3}{6} \times \frac{4}{7} + \frac{4}{6} \times \frac{3}{7} = \frac{4}{7}
\end{aligned}$$

And if bag B is chosen, then

$$\begin{aligned}
\mathcal{P}(\text{diff}|B) &= \mathcal{P}(\text{B2 = t, B1 = c}) + \mathcal{P}(\text{B2 = c, B1 = t}) \\
&= \mathcal{P}(\text{B2 = t}|\text{B1 = c})\mathcal{P}(\text{B1 = c}) + \mathcal{P}(\text{B2 = c}|\text{B1 = t})\mathcal{P}(\text{B1 = t}) \\
&= \frac{1}{5} \times \frac{5}{6} + \frac{5}{5} \times \frac{1}{6} = \frac{1}{3}.
\end{aligned}$$

Aggregating the two gives

$$\begin{aligned}
\mathcal{P}(\text{diff}) &= \mathcal{P}(\text{diff}|A)\mathcal{P}(A) + \mathcal{P}(\text{diff}|B)\mathcal{P}(B) \\
&= \frac{4}{7} \times 0.5 + \frac{1}{3} \times 0.5 = \frac{19}{42} \quad (\mathrm{A}_2)
\end{aligned}$$

Solution 3

From the text directly, the reward for each outcome is

$$R(3h) = 200, \quad R(2h) = 40, \quad R(1h) = 0, \quad R(0h) = -200.$$

The probabilities of the outcomes are

$$\begin{aligned}
\mathcal{P}(3h) &= \mathcal{P}(hhh) = [\mathcal{P}(h)]^3 = (0.5)^3 = 0.125, \\
\mathcal{P}(2h) &= \mathcal{P}(hht) + \mathcal{P}(hth) + \mathcal{P}(thh) = 3[\mathcal{P}(h)]^2\mathcal{P}(t) = 3 \times (0.5)^3 = 0.375, \\
\mathcal{P}(1h) &= \mathcal{P}(htt) + \mathcal{P}(tht) + \mathcal{P}(tth) = 3\mathcal{P}(t)[\mathcal{P}(t)]^2 = 3 \times (0.5)^3 = 0.375, \\
\mathcal{P}(0h) &= \mathcal{P}(ttt) = [\mathcal{P}(t)]^3 = (0.5)^3 = 0.125.
\end{aligned}$$

The expectation value of the reward is

$$\begin{aligned}
\langle R \rangle = \mathbb{E}(R) &= \sum_{\text{out in outcomes}} R(\text{out}) \times \mathcal{P}(\text{out}) \\
&= 100 \times 0.125 + 40 \times 0.375 + 0 - 200 \times 0.125 = (\$)2.5 \quad \text{(A)}
\end{aligned}$$

Solution 4

Let's label the positions in the line with 1, 2, 3 . . . 9, 10.
Each student gets randomly assigned to a position. The first student has 10 choices, the second has 9, the third has 8 and so on. The total number of possible line-ups is then

$$10 \times 9 \times 8 \ldots \times 2 \times 1 = 10!$$

When counting the line-ups where P, W and G are standing next to each other, we need to reserve a 3-size block for them. There's a total of 8 such blocks:

$$[1, 2, 3] \text{ or } [2, 3, 4] \text{ or } [3, 4, 5] \ldots \text{ or } [7, 8, 9] \text{ or } [8, 9, 10].$$

For each choice of the above blocks the remaining 7 students can arrange themselves on the remaining 7 places in 7! different ways.
And in the same time, P,W,G can arrange themselves in 3! different ways withing their block. The total number of such line-ups is $8 \times 7! \times 3!$, and, then, the probability that they are standing next to each other is

$$\mathcal{P} = \frac{8 \times 7! \times 3!}{10!} = \frac{8! \times 6}{10 \times 9 \times 8!} = \frac{1}{15}. \quad (\text{A}_1)$$

If they are standing in a circle, that's equivalent to the line whose front and rear are made neighboring. This way there are two more ways to reserve a 3-size block:

$$[9, 10, 1] \text{ or } [10, 1, 2]$$

which gets added to the previous 8 possibilites. The total number of arrangements where P,W,G stand next to each other is then $(8 + 2) \times 7! \times 3!$, and the probability is

$$\mathcal{P} = \frac{10 \times 7! \times 3!}{10!} = \frac{6 \times 7!}{9 \times 8 \times 7!} = \frac{1}{12}. \quad (\text{A}_2)$$

Solution 5

Let's assume the female dancers pick their partners. α can choose from 5 different choices, β can choose from the 4 remaining, γ can choose from the 3 remaining, ... The total number of pairings is

$$5 \times 4 \times \ldots 2 \times 1 = 5!$$

If we lock γ's choice to be c. Then the remaining 4 females can choose only from the remaining 4 males. Similarly to the 5-5 case, now, this can happen 4! different ways. The probability that c dances with γ is

$$\mathcal{P} = \frac{4!}{5!} = \frac{4!}{5 \times 4!} = \frac{1}{5}. \quad \text{(A)}$$

Solution 6

Let's label the groups with A, B, C. The total number of ways 21 people can be groupped into these 7-people groups is $21!/(7!)^3$. To see this, we imagine the following procedure:

1. Choose 7 people out of 21, and assign them to group A
2. Choose 7 people out of the remaining 14, and assign them to group B
3. Assign the remaining 7 people to group C

The first step can be done in $\binom{21}{7}$ ways, the second step in $\binom{14}{7}$ ways, while the third step is deterministic. The total possible ways is

$$\binom{21}{7}\binom{14}{7} = \frac{21!}{7! \times 14!}\frac{14!}{7! \times 7!} = \frac{21!}{(7!)^3}.$$

Now, when counting the number of grouppings where D and G are in the same group, we can investigate the 3 possible cases separately.

First, if D and G are in group A. This limits the free positions to 5 in group A, therefore the steps are the following:

1. Choose 5 people out of 19, and assign them to group A
2. Choose 7 people out of the remaining 14, and assign them to group B
3. Assign the remaining 7 people to group C

The number of ways is

$$\binom{19}{5}\binom{14}{7} = \frac{19!}{5! \times 14!}\frac{14!}{7! \times 7!} = \frac{19!}{5!(7!)^2}.$$

Similarly, if D and G are in group B, the number of ways is

$$\binom{19}{7}\binom{12}{5} = \frac{19!}{7! \times 12!}\frac{12!}{5! \times 7!} = \frac{19!}{5!(7!)^2}.$$

And if they are in group C, the number of ways is

$$\binom{19}{7}\binom{12}{7} = \frac{19!}{7! \times 12!}\frac{12!}{7! \times 5!} = \frac{19!}{5!(7!)^2}.$$

The total number of ways when they are in the same group is then $3 \times 19!/(5! \times (7!)^2)$, and the probability of this is

$$\mathcal{P} = 3 \times \frac{19!}{5! \times (7!)^2}\Bigg/\frac{21!}{(7!)^3} = \frac{3 \times 7 \times 6}{21 \times 20} = \frac{3}{10}. \quad \text{(A)}$$

Solution 7

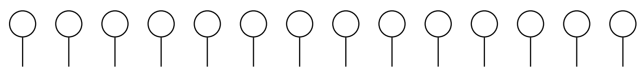
Depending on the exact procedure the Krampus uses to randomly distribute the candy, more than one solution is possible.

If the Krampus takes each candy, one after the other, out of his bag, and decides randomly to whom to give it, then the probability of A not getting any candy is

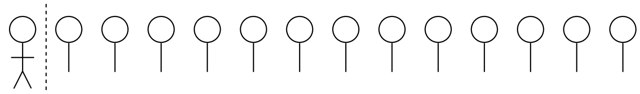
$$\mathcal{P}(\text{no candy}) = [\mathcal{P}(\text{not this candy})]^{10} = \left(\frac{3}{4}\right)^{10}. \quad \text{(A)}$$

However, if the Krampus simply decides how many pieces of candy to give to each kid, in which case a candy distribution of [10, 0, 0, 0] is as probable as [3, 2, 2, 3], (and any other combination of four integers between 0 and 10, which add up to 10) then the result is different.
Let's calculate the total number of different ways the Krampus can distribut the candy using this procedure.

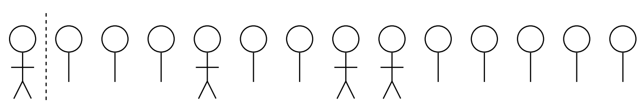
1. Let's draw 4 + 10 lollipops:



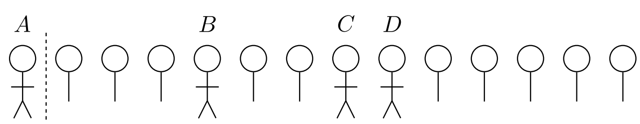
2. Turn the first one into a kid:



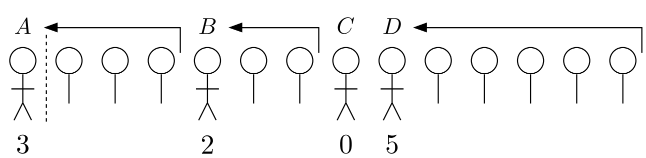
3. Choose 3 out of the 13 remaining lollipops randomly, and turn them into kids too:



4. Starting from left, label the kids with their names, A, B, C, D, in order:



5. Assign the lollipops to the kid on their left, and count how many each kid got.



In the above procedure only step 3 is random, and it can be done in $\binom{13}{3}$ different ways.

The cases where A doesn't get any candy correspond to deterministically choosing the second lollipop (out of 14) to be a kid. Then, in step 3 the two remaining kids needs to be chosen out of the 12 remaining lollipops, which can be done in $\binom{12}{2}$ different ways.

The probability of A getting no candy is

$$\mathcal{P} = \binom{12}{2} \Big/ \binom{13}{3} = \frac{12!}{2! \times 10!} \frac{3! \times 10!}{13!} = \frac{3}{13}. \quad \text{(A}'\text{)}$$

Solution 8

The number of "1"s, n_1 in $n_{\mathrm{tot}} = 20$ throws is a binomial variable. The probability of throwing two "1"s is

$$\mathcal{P}(n_1 = 2|n_{\mathrm{tot}} = 20, p = 0.1) = \binom{n_{\mathrm{tot}}}{n_1} p^{n_1}(1-p)^{n_{\mathrm{tot}}-n_1} = \binom{20}{2}(0.1)^2(0.9)^{18} = 0.2852. \quad (\mathrm{A}_1)$$

In order to get the first "1" at the tenth throw, we need to (1) not get "1" in 9 throws AND (2) get "1" in the tenth throw. The probability of this happening is

$$\mathcal{P}(\text{no "1" in 9 throws"}) \times \mathcal{P}(\text{throw "1"}) = (1-p)^9 \times p = (0.9)^9(0.1) = 0.0387. \quad (\mathrm{A}_2)$$

In order to get the third "1" at the thirtieth throw, we need to (1) get exactly 2 "1"s in 29 throws AND (2) get a "1" in the thirtieth one. The probability of this happening is

$$\mathcal{P}(n_1 = 2|n_{\mathrm{tot}} = 29, p = 0.1) \times p = \binom{29}{2}(0.1)^2(0.9)^{27} \times 0.1 = 0.0236. \quad (\mathrm{A}_3)$$

Solution 9

Let's denote the different event with the following abbreviations. "+/−" means the test was positive/negative, "ill" means the sample has the disease, "heathy" means the sample doesn't have the disease.

This is a probabilistic inference problem with a single data point $D =$ "$+''$, and two competing models $M_1 =$ "healthy", $M_2 =$ "ill". Let's follow the steps outlined for probabilistic inference:

1. From the known prevalence of the disease we deduce the priors:

$$\mathcal{P}(\text{ill}) = 0.0001, \qquad \mathcal{P}(\text{healthy}) = 0.9999$$

2. From the false rates of the test, we can deduce the forward probabilities:

$$\begin{array}{ll} \mathcal{P}(+|\text{healty}) = 0.01, & \mathcal{P}(-|\text{healty}) = 0.99 \\ \mathcal{P}(+|\text{ill}) = 1, & \mathcal{P}(-|\text{ill}) = 0 \end{array}$$

3. The posterior can be calculated by putting these together to get

$$\begin{aligned} \mathcal{P}(\text{ill}|+) &= \frac{\mathcal{P}(+|\text{ill})\mathcal{P}(\text{ill})}{\sum_{i=1,2}\mathcal{P}(+|M_i)P(M_i)} = \frac{\mathcal{P}(+|\text{ill})\mathcal{P}(\text{ill})}{\mathcal{P}(+|\text{ill})\mathcal{P}(\text{ill}) + \mathcal{P}(+|\text{healthy})\mathcal{P}(\text{healthy})} = \\ &= \frac{1\times10^{-4}}{1\times10^{-4} + 0.01\times0.9999} = \frac{1}{100.99} \approx 1\%. \quad (\mathrm{A}) \end{aligned}$$

Solution 10

Let's use the following notation. B and no$B$: (no) break-in, Q and no$Q$: (no) earthquake, A: alarm set off. We are going to use probabilistic inference, where the single data point is the alarm going off, $D = A$, and disjunct models are $M_1 = (B \text{ and } Q)$, $M_2 = (\text{no}B \text{ and } Q)$, $M_3 = (B \text{ and no}Q)$, and $M_4 = (\text{no}B \text{ and no}Q)$.

1. Using the base rates of B and Q, we can write

$$\mathcal{P}(B) = 1/(5 \times 365.25) = 5.5 \times 10^{-4}, \quad \mathcal{P}(Q) = 2/365.25 = 5.5 \times 10^{-3}$$

   assuming that break-ins are independent of earthquakes, we can write

$$\begin{aligned}
&\mathcal{P}(B \text{ and } Q) = \mathcal{P}(B)\mathcal{P}(Q) = 3 \times 10^{-6} \\
&\mathcal{P}(B \text{ and no}Q) = \mathcal{P}(B)(1 - \mathcal{P}(Q)) = 5.5 \times 10^{-4} \\
&\mathcal{P}(\text{no}B \text{ and } Q) = (1 - \mathcal{P}(B))\mathcal{P}(Q) = 5.5 \times 10^{-3} \\
&\mathcal{P}(\text{no}B \text{ and no}Q) = (1 - \mathcal{P}(B))(1 - \mathcal{P}(Q)) = 0.9940
\end{aligned}$$

2. Two of the forward probabilities can be readily deduced from the accuracy and the earthquake sensitivity of the alarm.

$$\mathcal{P}(A|B \text{ and no}Q) = 0.999, \qquad \mathcal{P}(A|\text{no}B \text{ and } Q) = 0.2$$

   We are not given explicit information about the forward probabilities of the other two models. Let's use the following realistic assumptions.

$$\mathcal{P}(A|B \text{ and } Q) = 1, \qquad \mathcal{P}(A|\text{no}B \text{ and no}Q) = 0$$

3. The posterior probability of each model $\mathcal{P}(M_i|D)$ can be calculated the following way. First, let's calculate the *non-normalized* weights $(\mathcal{P}(D|M_i)\mathcal{P}(M_i))$ of each model:

$$\begin{aligned}
\mathcal{P}(A|B,Q)\mathcal{P}(B,Q) &= 1 \times 3 \times 10^{-6} = 3 \times 10^{-6} \\
\mathcal{P}(A|B,\text{no}Q)\mathcal{P}(B,\text{no}Q) &= 0.999 \times 5.5 \times 10^{-4} = 5.5 \times 10^{-4} \\
\mathcal{P}(A|\text{no}B,Q)\mathcal{P}(\text{no}B,Q) &= 0.2 \times 5.5 \times 10^{-3} = 1.1 \times 10^{-3} \\
\mathcal{P}(A|\text{no}B,\text{no}Q)\mathcal{P}(\text{no}B,\text{no}Q) &= 0 \times 0.9940 = 0
\end{aligned}$$

   Then, normalize each of them with the total sum $Z = \sum_i \mathcal{P}(D|M_i)\mathcal{P}(M_i) = 0.00165$. This gives us

$$\begin{aligned}
\mathcal{P}(B,Q|A) &= 3 \times 10^{-6}/0.00165 = 0.0018 \approx 0.2\% \\
\mathcal{P}(B,\text{no}Q|A) &= 5.5 \times 10^{-4}/0.00165 = 0.333 \approx 33\% \\
\mathcal{P}(\text{no}B,Q|A) &= 1.1 \times 10^{-3}/0.00165 = 0.665 \approx 67\% \\
\mathcal{P}(\text{no}B,\text{no}Q|A) &= 0
\end{aligned}$$

   Now we know everything we need to answer the two questions.

In the first questions we only know that the alarm went off, but have no idea whether there was an earthquake or a break-in. To find the probability of a break-in, we need to marginalize over the two possible earthquake outcomes. This gives us

$$\mathcal{P}(B|A) = \mathcal{P}(B,Q|A) + \mathcal{P}(B,\text{no}Q|A) = 0.2\% + 33\% \approx 33\%. \quad (\text{A}_1)$$

In the second question, we know for sure that an earthquake has happened. This means the probability that there was a break-in is

$$\mathcal{P}(B|Q,A) = \frac{\mathcal{P}(B,Q|A)}{\mathcal{P}(Q|A)} = \frac{\mathcal{P}(B,Q|A)}{\mathcal{P}(B,Q|A) + \mathcal{P}(\text{no}B,Q|A)} = \frac{0.0018}{0.0018 + 0.665} = 0.0027 \approx 0.3\%. \quad (\text{A}_2)$$

Note: The knowledge about the occurrence of an earthquake lowers the probability of a break-in. This phenomenon is called "explaining away".

Solution 11

Let us denote the two hypothesis with M_0 = "fair coin" and M_1 = "trick coin". We are going to use probabilistic inference to determine the probability of each hypothesis after the data D = "five heads in a row"(hhhhh) has been observed.

1. The prior probabilities for the two models are hinted to be equal,

$$\mathcal{P}(\text{fair}) = 0.5, \qquad \mathcal{P}(\text{trick}) = 0.5$$

2. We deduce the forward probabilities:

$$\mathcal{P}(\text{h}|\text{fair}) = 1/2, \qquad \mathcal{P}(\text{h}|\text{trick}) = 1.$$

Using only these, we can answer the first question:

$$\mathcal{P}(D|\text{fair}) = \Big[\mathcal{P}(\text{h}|\text{fair})\Big]^5 = \frac{1}{32}, \qquad \mathcal{P}(D|\text{trick}) = \Big[\mathcal{P}(\text{h}|\text{trick})\Big]^5 = 1. \quad (\text{A}_1)$$

3. The posterior probability of the "trick" model is

$$\mathcal{P}(\text{trick}|D) = \frac{\mathcal{P}(D|\text{trick})\mathcal{P}(\text{trick})}{\mathcal{P}(D|\text{fair})\mathcal{P}(\text{fair}) + \mathcal{P}(D|\text{trick})\mathcal{P}(\text{trick})} = \frac{1 \times 0.5}{\frac{1}{32} \times 0.5 + 1 \times 0.5} = \frac{32}{33}. \quad (\text{A}_2)$$

while the probability of "fair" is $\mathcal{P}(\text{fair}|D) = 1 - \mathcal{P}(\text{trick}|D) = 1/33$.

4. Weighting the forward probabilities of the two models with their posterior probabilitites gives us the prediction for the 6th toss:

$$\begin{aligned}
\mathcal{P}(6^{\text{th}}\text{ toss} \to \text{h} \mid \text{hhhhh}) = \mathcal{P}(\text{h}|D) &= \sum_i \mathcal{P}(\text{h}|M_i)\mathcal{P}(M_i|D) \\
&= \mathcal{P}(\text{h}|\text{fair})\mathcal{P}(\text{fair}|D) + \mathcal{P}(\text{h}|\text{trick})\mathcal{P}(\text{trick}|D) \\
&= \frac{1}{2} \times \frac{1}{33} + 1 \times \frac{32}{33} = \frac{65}{66} \approx 98.5\%. \quad (\text{A}_3)
\end{aligned}$$