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 $\rightarrow$ win \$100, 2 heads $\rightarrow$ win \$40, 1 head $\rightarrow$ nothing, 0 heads $\rightarrow$ 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)

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

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

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

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

Model comparison

Question 9

A blood test for an disease gives positive result on a sample.
Q: What is the probability that the sample actually has the disease? (Let's assume that the prevalence of the 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?