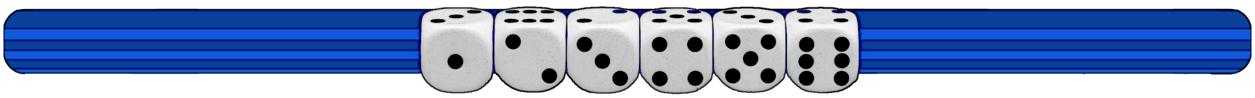
Exercise sheet 1: Hidden Markov models

Exercise 1 - Two kinds of dice



A casino uses two kinds of dice: 98% of dice are fair and 2% are loaded. The loaded die has a probability of 0.5 to show number six and 0.1 for the numbers one to five.

Question 1A When we pick up a die from a table at random, what is the probability of rolling a six?

Hint 1 : Formulae

$$L = \text{Loaded} \quad F = \text{Fair} \quad \mathcal{O} = \text{Observation} P(\mathcal{O}) = P(F) \times P(\mathcal{O}|F) + P(L) \times P(\mathcal{O}|L)$$

Hint 2 : Calculation Method

$$P(6) = 0.98 \times \frac{1}{6} + 0.02 \times \frac{1}{2}$$

Solution

$$P(6) = 0.173\overline{3}$$

Question 1B We pick up a die from a table at random and roll [6 6 6]. What is the probability, that the die is loaded.

Hint 1 : Formulae

$$P(L|\mathcal{O}) = \frac{P(L,\mathcal{O})}{P(\mathcal{O})} P(L,\mathcal{O}) = P(\mathcal{O}|L) \times P(L)$$

Hint 2 : Calculation Method

$$\begin{aligned} P(L|\mathcal{O}) &= \frac{P(\mathcal{O}|L) \times P(L)}{P(\mathcal{O}|L) \times P(L) + P(\mathcal{O}|F) \times P(F)} \\ &= \frac{(\frac{1}{2})^3 \times 0.02}{(\frac{1}{2})^3 \times 0.02 + (\frac{1}{6})^3 \times 0.98} \end{aligned}$$

Solution

$$P(L|\mathcal{O}) = 35.53\%$$

Question 1C How many sixes in a row would we need to roll to be at least 90% sure that the die is loaded?

Hint 1 : Formulae

$$P(L|\mathcal{O}) = \frac{P(\mathcal{O}|L) \times P(L)}{P(\mathcal{O}|L) \times P(L) + P(\mathcal{O}|F) \times P(F)}$$

Hint 2 : Calculation Method

$$\begin{aligned}
P(L|\mathcal{O}) = \frac{\frac{2}{100} \times (\frac{1}{2})^n}{\frac{2}{100} \times (\frac{1}{2})^n + \frac{98}{100} \times (\frac{1}{6})^n} & \geq 0.9 && |\text{ split } (\frac{1}{6})^n && (1)\\
\iff \frac{\frac{2}{100} \times (\frac{1}{2})^n}{\frac{2}{100} \times (\frac{1}{2})^n + \frac{98}{100} \times (\frac{1}{2})^n \times (\frac{1}{3})^n} & \geq \frac{9}{10} && |\text{ factorize} && (2)\\
\iff \frac{\frac{2}{100} \times (\frac{1}{2})^n}{\frac{2}{100} \times (\frac{1}{2})^n \times (1 + 49 \times (\frac{1}{3})^n)} & \geq \frac{9}{10} && |\text{ simplify, given } n > 0 && (3)\\
\iff \frac{1}{1 + 49 \times (\frac{1}{3})^n} & \geq \frac{9}{10} && |\text{ cross-multiply, given } n > 0 && (4)\\
\iff \frac{9}{10}(1 + 49 \times (\frac{1}{3})^n) & \leq 1 && |\text{ rewrite} && (5)\\
\iff (\frac{1}{3})^n & \leq \frac{1}{441} && |\ \ln() && (6)\\
\iff n \times ln(\frac{1}{3}) & \leq ln(\frac{1}{441}) && |\times \frac{1}{ln(\frac{1}{3})} && (7)\\
\iff n & \geq \frac{ln(\frac{1}{441})}{ln(\frac{1}{3})} && && (8)\\
\iff n & \geq 5.542487... && && (9)
\end{aligned}$$

Solution

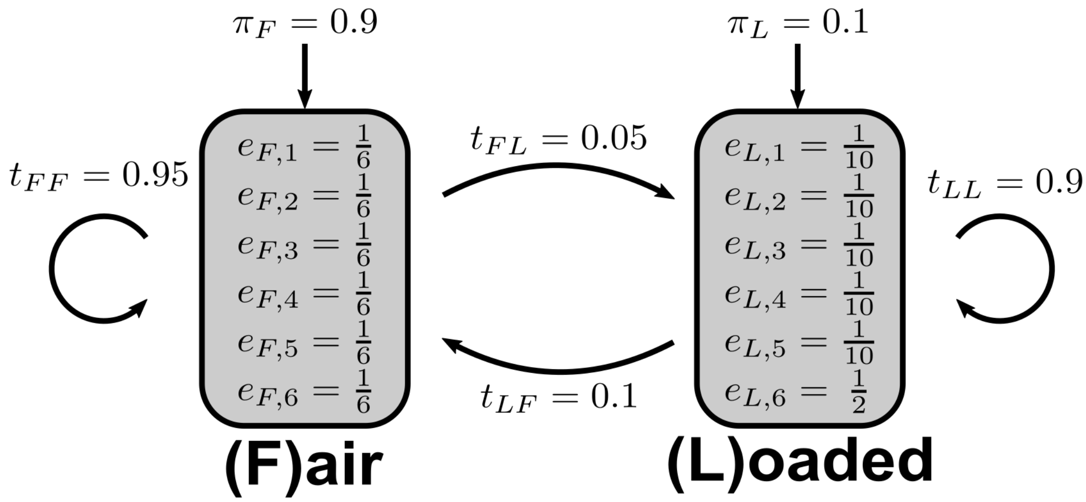
$n = 6$, as only Integers make sense here (just trying would also work)

Exercise 2 - The occasionally cheating casino

In a casino they use a fair die most of the time, but occasionally they switch to a loaded die. The loaded die has a probability 0.5 to show number six and probability 0.1 for the numbers one to five. Assume that the casino switches from a fair to a loaded die with probability 0.05 before each roll, and that the probability of switching back is 0.1. The probability to start a game with the fair die is 0.9.

Question 2A Draw a graphical representation of the described Hidden Markov model.

Solution



Question 2B Given an observed sequence of outcomes $\mathcal{O} = 3661634$ and two possible state sequences $s_1 = LLLFFFF$ and $s_2 = FFFFFFF$ (where F = Fair and L = Loaded), what are the joint probabilities $P(\mathcal{O}, p_1)$ and $P(\mathcal{O}, p_2)$ in the HMM described above?

Hint 1 : Formulae

$$P(\mathcal{O}, p_x) = P(\mathcal{O}|p_x) \times P(p_x)$$

Hint 2 : Calculation Method

$$\begin{aligned}
P(s_1) =& \pi_L \times t_{LL}^2 \times t_{LF} \times t_{FF}^3 = 0.1 \times 0.9^2 \times 0.1 \times (0.95)^3 = 0.0069 \\
P(s_2) =& \pi_F \times t_{FF}^6 = 0.9 \times 0.95^6 = 0.6616 \\
P(\mathcal{O}|s_1) =& t_{L,3} \times t_{L,6}^2 \times t_{F,1} \times t_{F,6} \times t_{F,3} \times t_{F,4} = 0.1 \times 0.5^2 \times (\frac{1}{6})^4 = 1.9 \times 10^{-5} \\
P(\mathcal{O}|s_2) =& t_{F,3} \times t_{F,6}^2 \times t_{F,1} \times t_{F,6} \times t_{F,3} \times t_{F,4} = (\frac{1}{6})^7 = 3.57 \times 10^{-6} \\
P(\mathcal{O}, s_1) =& P(\mathcal{O}|s_1) \times P(s_1) = 1.9 \times 10^{-5} \times 0.0069 \\
P(\mathcal{O}, s_2) =& P(\mathcal{O}|s_2) \times P(s_2) = 3.57 \times 10^{-6} \times 0.6616
\end{aligned}$$

Solution

$$\begin{aligned}
P(\mathcal{O}, s_1) &= 1.34 \times 10^{-7} \\
P(\mathcal{O}, s_2) &= 2.36 \times 10^{-6}
\end{aligned}$$

Question 2C Give an observation $\mathcal{O} = 1662$, how many possible state sequences exist in the described HMM?

Hint 1 The actual observation does not matter in this case because all emission probabilities are > 0. This there are 2^4 possible state sequences.

Solution There are 16 possible state sequences.

Exercise 3 - Programming assignment

Programming assignments are available via Github Classroom and contain automatic tests.

We recommend doing these assignments as they will help you further your understanding of this topic.

Access the Github Classroom link: Programming Assignment: Sheet 01.
