CS-E4820 Machine Learning: Advanced Probabilistic Methods (spring 2021)

Pekka Marttinen, Santosh Hiremath, Tianyu Cui, Yogesh Kumar, Zheyang Shen, Alexander Aushev, Khaoula El Mekkaoui, Shaoxiong Ji, Alexander Nikitin, Sebastiaan De Peuter, Joakim Järvinen.

Exercise 1, due on Tuesday 26th January at 23:00.

Contents

Problem 1: Coins

There are two bent coins (c_1 and c_2) with different properties, and your objective is to guess which coin was used (i.e. the value of random variable $C \in \{c_1, c_2\}$), after learning whether the result of the coin toss (i.e. the random variable $X \in \{\mathrm{h}, \mathrm{t}\}$) was heads or tails.

As prior knowledge, we know the probability of each coin resulting in tails: $p(X = \mathrm{t} \mid C = c_1) = \theta_1$ and $p(X = \mathrm{t} \mid C = c_2) = \theta_2$. In addition, the prior probability for using coin c_1 is known: $p(C = c_1) = \pi_1$.

Give the posterior probability of coin c_1 being used for the toss, $p(C = c_1 \mid X)$, in terms of θ_1, θ_2 and π_1, for both $X = \mathrm{t}$ and $X = \mathrm{h}$.

Furthermore, plot the posterior probability of coin c_1, $p(C = c_1 \mid X = \mathrm{t})$, as a function of θ_1, when we have $\theta_2 = 0.5$ and $\pi_1 = 0.5$.

Write your solutions in LateX or attach a picture in the answer cell provided below. You can add a picture using the command `!(imagename_in_the_folder.jpg)`. Latex in here works similarly as you would write it normally! You can use some of the definitions from the exercise description as a reference. The list of valid Latex commands in Jypyter notebook can be found here: http://www.onemathematicalcat.org/MathJaxDocumentation/TeXSyntax.htm

Solution

The posterior probability follows from Bayes' theorem,

$$p(A \mid B) = \frac{p(B \mid A)p(A)}{p(B)},$$

and the principle of marginalization,

$$p(A) = \sum_b p(A, B = b) = \sum_b p(A \mid B = b)p(B = b).$$

Applying Bayes' theorem in (1.1) and marginalization in (1.2), we get the following solution for $p(C = c_1 \mid X = \mathrm{t})$:

$$
\begin{align}
p(C = c_1 \mid X = \mathrm{t}) &= \frac{p(X = \mathrm{t} \mid C = c_1) p(C = c_1)}{p(X = \mathrm{t})} \tag{1.1} \\
&= \frac{p(X = \mathrm{t} \mid C = c_1) p(C = c_1)}{\sum_{i=1}^{2} p(X = \mathrm{t} \mid C = c_i) p(C = c_i)} \tag{1.2} \\
&= \frac{\theta_1 \pi_1}{\theta_1 \pi_1 + \theta_2 (1 - \pi_1)}.
\end{align}
$$

Similarly, for the $X = \mathrm{h}$ case:

$$
p(C = c_1 \mid X = \mathrm{h}) = \frac{(1 - \theta_1)\pi_1}{(1 - \theta_1)\pi_1 + (1 - \theta_2)(1 - \pi_1)}.
$$

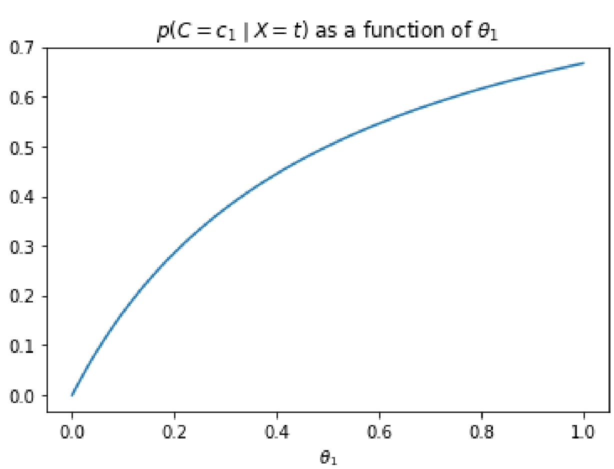
[3]:

```python
import matplotlib.pyplot as plt
import numpy as np

# define some variables:
theta_2 = 0.5
pi_1 = 0.5

# compute posterior probability of c_1
#theta_1 = ... # choose a reasonable range
#post_c1 = ... # posterior
### BEGIN SOLUTION
theta_1 = np.linspace(0, 1, 100)
post_c1 = theta_1 * pi_1 / (theta_1 * pi_1 + theta_2 * (1 - pi_1))
### END SOLUTION

# plot the result
plt.plot(theta_1, post_c1)
plt.title('$p(C = c_1 \\mid X = t)$ as a function of $\\theta_1$')
plt.xlabel('$\\theta_1$')
plt.show()
```



Problem 2: False positive paradox

Consider a hypothetical lie detector that is ''fairly reliable'', in the sense that it will correctly detect 98% of all lies, and also classify as true 98% of all statements that are actually true. This lie detector is being used in an attempt to detect academic dishonesty, by asking''did you cheat?'' from all students participating in an exam of a machine learning course. (This example is still hypothetical.)

For the purposes of this question, assume as prior knowledge that there are 300 students taking the exam, and a single student has chosen to cheat. We will further assume that all students deny having cheated. If the detector now flags a particular student X as a cheater, how likely is it that X has, in fact, cheated in the exam?

Write your solutions in LateX or attach a picture in the answer cell provided below.

Solution

Let us denote by X whether the student cheated or not ($X \in \{\text{cheater}, \text{honest}\}$), and by D the lie detector report ($D \in \{\text{lie}, \text{truth}\}$).

We have the following prior knowledge:

$$\begin{aligned}
p(X = \text{cheater}) &= \frac{1}{300}\ , \\
p(X = \text{honest}) &= \frac{299}{300}\ , \\
p(D = \text{lie} \mid X = \text{cheater}) &= \frac{98}{100}\ , \\
p(D = \text{lie} \mid X = \text{honest}) &= 1 - p(D = \text{truth} \mid X = \text{honest}) \\
&= 1 - \frac{98}{100} = \frac{2}{100}\ , \\
p(D = \text{lie}) &= \sum_X p(D = \text{lie} \mid X) p(X)\ , \\
&= \frac{98}{100}\frac{1}{300} + \frac{2}{100}\frac{299}{300} = \frac{696}{30000}\ .
\end{aligned}$$

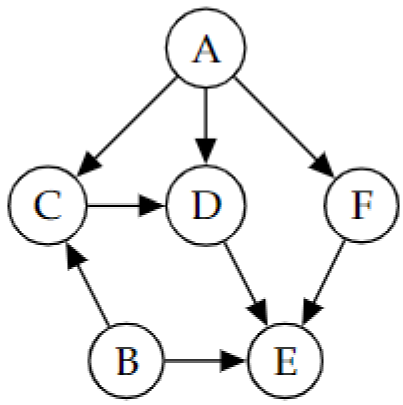
Applying Bayes' theorem again,

$$\begin{aligned}
p(X = \text{cheater} \mid D = \text{lie}) &= \frac{p(D = \text{lie} \mid X = \text{cheater}) p(X = \text{cheater})}{p(D = \text{lie})} \\
&= \frac{\frac{98}{100}\frac{1}{300}}{\frac{696}{30000}} = \frac{98}{696} \approx 0.14.
\end{aligned}$$

In other words, despite having a ''98% reliable'' detector, if cheating is sufficiently rare, it is far more likely that a student flagged as a liar by the detector is telling the truth than lying.

Problem 3: Markov blanket definition

Consider the Bayesian network in below. What is the Markov blanket of each variable? (see Barber: Bayesian Reasoning and Machine Learning, ch. 2.1, Definition 2.5)



Define Markov blanket for each variable A, B, C, D, E, F. You answer should list the nodes that form the Markov blanket for each node. For example, for node A, your answer should look like so $A = \{B, D, E\}$

Solution

$A = \{B, C, D, F\}$
$B = \{A, C, D, E, F\}$
$C = \{A, B, D\}$
$D = \{A, B, C, E, F\}$
$E = \{B, D, F\}$
$F = \{A, B, D, E\}$