Sheet 3 solutions

June 5, 2020

Exercise 1: Bayes Rule

Suppose you are a witness to a nighttime hit-and-run accident involving a taxi in Athens. All taxi cars in Athens are blue or green. You swear under oath that the taxi was blue. Extensive testing shows that, under the dim lighting conditions, discrimination between blue and green is 75% reliable.

(a) *Given your statement as a witness and given that 9 out of 10 Athenian taxis are green, what is the probability of the taxi being blue?*

Let us define two random variables x and y; let the first denote the actual color of the taxi, while let the second denote the color observed by the witness. Both random variables can have two possible values, respectively, g for green and b for blue.

From the text *"under the dim lighting conditions, discrimination between blue and green is 75% reliable."* we can infer the following probabilities:

$$\begin{aligned} p(y = g \mid x = g) &= 0.75 \\ \Rightarrow p(y = b \mid x = g) &= 0.25 \\ p(y = b \mid x = b) &= 0.75 \\ \Rightarrow p(y = g \mid x = b) &= 0.25 \end{aligned}$$

While, since we know that *"9 out of 10 Athenian taxis are green"*, we have as a prior:

$$\begin{aligned} p(x = g) &= 0.9 \\ p(x = b) &= 0.1 \end{aligned}$$

What we're interested in is to compute the probability that the taxi is blue given the prior distribution on taxis and the statement of the witness, namely $p(x = b|y = b)$. Using the conditional probability (see "Probabilistic Robotics" slide 9):

$$p(x = b, y = b) = p(x = b \mid y = b)p(y = b) \tag{1}$$

We can compute it as follows:

$$
\begin{aligned}
p(x = b \mid y = b) &= \frac{p(x = b, y = b)}{p(y = b)} \\
&= \frac{p(y = b \mid x = b)p(x = b)}{p(y = b)} \\
&= \frac{p(y = b \mid x = b)p(x = b)}{p(y = b, x = g) + p(y = b, x = b)} \\
&= \frac{p(y = b \mid x = b)p(x = b)}{p(y = b \mid x = g)p(x = g) + p(y = b \mid x = b)p(x = b)} \qquad (2)
\end{aligned}
$$

Notice that the resulting formula is, indeed, exactly Bayes' rule.

Since we know all of the probabilities in (2), we can substitute the numerical values and find:

$$p(x = b \mid y = b) = \frac{0.75 \cdot 0.1}{0.25 \cdot 0.9 + 0.75 \cdot 0.1} = 0.25$$

Thus, given the prior chance of the taxi being blue and the observation of the witness, the most likely conclusion is that the witness is in fact wrong.

(b) *Is there a significant change if 7 out of 10 Athenian taxis are green?*

This exercise is exactly the same as the previous one, with a slight change on the prior distribution on taxi colors, i.e., this time we have:

$$
\begin{aligned}
p(x = g) = 0.7 \\
p(x = b) = 0.3
\end{aligned}
$$

If we plug into (2) the new values, this time we find:

$$p(x = b \mid y = b) = \frac{0.75 \cdot 0.3}{0.25 \cdot 0.7 + 0.75 \cdot 0.3} = 0.5626$$

Contrary to the previous exercise, this time the most likely conclusion is that the witness is right.

(c) *Suppose now that there is a second witness who swears that the taxi is green. Unfortunately he is color blind, so he has only a 50% chance of being right. How would this change the estimate from (b)?*

Since the second witness has only a 50% chance of being right we are not expecting to gain any additional information on the color of the taxi. Intuitively, this is because we can regard him as nothing more than a fair coin toss, hence we are not expecting any change in estimate from (b).

Nevertheless, we can prove this formally as follows. Let z denote the random variable that describes the color observed by the second witness, with possible values g and b. The probability we seek to compute is thus $p(x = b \mid y = b, z = g)$.

Do note that we assume the two witnesses to be independent given the state of the hidden variable x, i.e. the true color of the car. This assumption is typically used in the Bayes filter – not only in this particular setup:

$$p(y = b, z = g|x = b) = p(y = b|x = b)p(z = g|x = b) \tag{3}$$

We will use it in the following derivation:

$$\begin{aligned} p(x = b \mid y = b, z = g) &= \frac{p(x = b, y = b, z = g)}{p(y = b, z = g)} \\ &= \frac{p(y = b, z = g \mid x = b)p(x = b)}{p(y = b, z = g)} \\ &= \frac{p(y = b \mid x = b)p(z = g \mid x = b)p(x = b)}{p(y = b, z = g)} \end{aligned}$$

It now has the form from slide 20, lecture 5:

$$p(x \mid \{z_1..z_n\}) = \eta_{\{1..n\}} \left[\prod_{i=1..n} p(z_i \mid x) \right] p(x).$$

In order to use the result from (b), we rewrite the equation in a recursive form:

$$\begin{aligned} &p(x = b \mid y = b, z = g) \\ \stackrel{BayesRule}{=} &\frac{p(x = b \mid y = b)p(z = g \mid x = b)p(y = b)}{p(y = b, z = g)} \\ = &\frac{p(z = g \mid x = b)p(x = b \mid y = b)}{p(z = g \mid y = b)} \end{aligned}$$

This has the form

probability of z given x** times **belief before** z times **normalizer

and is still general given the assumption (3). We can further expand the normalizer with the law of total probability to achieve the form of slide 21, lecture 5:

$$\begin{aligned} &p(x = b \mid y = b, z = g) \\ = &\frac{p(z = g \mid x = b)p(x = b \mid y = b)}{p(z = g, x = b \mid y = b) + p(z = g, x = g \mid y = b)} \\ = &\frac{p(z = g \mid x = b)p(x = b \mid y = b)}{p(z = g \mid x = b, y = b)p(x = b \mid y = b) + p(z = g \mid x = g, y = b)p(x = g \mid y = b)} \\ \stackrel{(3)}{=} &\frac{p(z = g \mid x = b)p(x = b \mid y = b)}{p(z = g \mid x = b)p(x = b \mid y = b) + p(z = g \mid x = g)p(x = g \mid y = b)} \end{aligned}$$

For the special case of a blind person (whose observation does not depend on the state x), we have

$$p(z = g \mid x = b) = p(z = g \mid x = g)$$

and thus

$$
\begin{aligned}
&p(x = b \mid y = b, z = g) \\
=&\frac{p(z = g \mid x = b)p(x = b \mid y = b)}{p(z = g \mid x = b)p(x = b \mid y = b) + p(z = g \mid x = b)p(x = g \mid y = b)} \\
=&\frac{p(x = b \mid y = b)}{p(x = b \mid y = b) + p(x = g \mid y = b)} \\
=&p(x = b \mid y = b)
\end{aligned}
$$

Which is consistent with the initial intuition of non-informativeness of the second measurement z.

Exercise 2: Bayes Filter

A vacuum cleaning robot is equipped with a cleaning unit to clean the floor. Furthermore, the robot has a sensor to detect whether the floor is clean or dirty. Neither the cleaning unit nor the sensor are perfect.

From previous experience you know that the robot succeeds in cleaning a dirty floor with a probability of

$$p(x_{t+1} = \textit{clean} \mid x_t = \textit{dirty}, u_{t+1} = \textit{vacuum-clean}) = 0.7,$$

where x_{t+1} is the state of the floor after having vacuum-cleaned, u_{t+1} is the control command, and x_t is the state of the floor before performing the action.

The probability that the sensor indicates that the floor is clean although it is dirty is given by $p(z = \textit{clean} \mid x = \textit{dirty}) = 0.3$, and the probability that the sensor correctly detects a clean floor is given by $p(z = \textit{clean} \mid x = \textit{clean}) = 0.9$.

Unfortunately, you have no knowledge about the current state of the floor. However, after cleaning the floor the sensor of the robot indicates that the floor is clean.

(a) *Compute the probability that the floor is still dirty after the robot has vacuum-cleaned it. Use an appropriate prior distribution and justify your choice.*

Let us first list all of the variables that are considered and their possible values:

- Let x_0 be the state of the floor prior to the cleaning action. x_0 can have as values either c (clean) or d (dirty).
- Let x_1 be the state of the floor after to the cleaning action. x_1 can have as values either c (clean) or d (dirty).
- Let z be the measurement of the robot after the cleaning action. z can have as values either c (clean) or d (dirty).
- Let u be the command for the execution of the cleaning action. Here we will only consider vc (vacuum clean) as a possible action.

Furthermore, we know that:

$$\begin{aligned}
p(x_1 = c \mid x_0 = d, u = vc) &= 0.7 \\
p(x_1 = d \mid x_0 = d, u = vc) &= 0.3 \\
p(z = c \mid x_1 = d) &= 0.3 \\
p(z = d \mid x_1 = d) &= 0.7 \\
p(z = c \mid x_1 = c) &= 0.9 \\
p(z = d \mid x_1 = c) &= 0.1
\end{aligned}$$

We can safely assume that if the floor is initially clean it will stay so even after a cleaning action, since no further details are given in this respect. Namely:

$$\begin{aligned}
p(x_1 = c \mid x_0 = c, u = vc) &= 1 \\
p(x_1 = d \mid x_0 = c, u = vc) &= 0
\end{aligned}$$

We are not given an initial belief on the state of the floor, therefore the most reasonable prior on the state of the floor would be the least informative one. This corresponds to setting:

$$p(x_0 = c) = 0.5 \qquad p(x_0 = d) = 0.5$$

Nevertheless, for the sake of generality, in this solution we will assume an arbitrary prior on the initial state of the floor:

$$p(x_0 = c) = q \qquad p(x_0 = d) = 1 - q$$

Where q can be chosen as any value such that $q \in [0, 1]$.

We're now aiming to compute the probability of the floor being dirty after executing the cleaning action and detecting the floor as clean. In other words we want to compute $p(x_1 = d \mid z = c, u = vc)$. In order to do so we can use the Bayes filter formula:

$$\begin{aligned}
p(x_1 = d \mid z = c, u = vc) = \eta\, p(z = c \mid x_1 = d) \sum_{x_0} p(x_1 = d | x_0, u = vc) p(x_0) &= \\
= \eta\, p(z = c \mid x_1 = d) \big(p(x_1 = d | x_0 = c, u = vc) p(x_0 = c) + & \\
p(x_1 = d | x_0 = d, u = vc) p(x_0 = d) \big) &= \\
= \eta\, 0.3 \left(0 \cdot q + 0.3 \cdot (1 - q) \right) = (0.09 - 0.09q)\, \eta &
\end{aligned}$$

We are missing the value of the normalizer η in order to obtain the probability value. We can compute it by taking into account the fact that our belief over x_1 needs to be normalized to 1, in other words, the following needs to hold:

$$p(x_1 = d \mid z = c, u = vc) + p(x_1 = c \mid z = c, u = vc) = 1 \tag{4}$$

We can thus apply again the Bayes filter formula to compute the value of the term we are missing as a function of η, thus:

$$\begin{aligned}
p(x_1 = c \mid z = c, u = vc) &= \eta\, p(z = c \mid x_1 = c) \sum_{x_0} p(x_1 = c | x_0, u = vc) p(x_0) = \\
&= \eta\, p(z = c \mid x_1 = c) \big(p(x_1 = c | x_0 = c, u = vc) p(x_0 = c) + \\
&\qquad\qquad p(x_1 = c | x_0 = d, u = vc) p(x_0 = d) \big) = \\
&= \eta\, 0.9 \left(1 \cdot q + 0.7 \cdot (1 - q) \right) = (0.63 + 0.27q)\, \eta
\end{aligned}$$

We can now plug the values we found in (4) and solve for η:

$$\begin{gathered}
(0.09 - 0.09q)\, \eta + (0.63 + 0.27q)\, \eta = 1 \\
\Rightarrow \eta = \frac{1}{0.72 + 0.18q}
\end{gathered}$$

This finally gives us:

$$p(x_1 = d \mid z = c, u = vc) = (0.09 - 0.09q)\, \eta = \frac{0.09 - 0.09q}{0.72 + 0.18q} \tag{5}$$

We can then substitute q with particular values for our prior and obtain a value for the resulting probability, for example:

- For the non informative prior of $q = 0.5$: $p(x_1 = d \mid z = c, u = vc) \approx 0.0556$
- If we assume the floor to be deterministically dirty at the beginning ($q = 0$): $p(x_1 = d \mid z = c, u = vc) = 0.125$

Note that we could have also computed the same value for η by realizing that:

$$\eta = \frac{1}{p(z = c)} = \frac{1}{p(z = c | x_1 = c) p(x_1 = c) + p(z = c | x_1 = d) p(x_1 = d)}$$

Where:

$$\begin{aligned}
p(x_1 = c) &= \sum_{x_0} p(x_1 = c | x_0, u = vc) p(x_0) \\
p(x_1 = d) &= \sum_{x_0} p(x_1 = d | x_0, u = vc) p(x_0)
\end{aligned}$$

(b) *Which prior gives you a lower bound for that probability?*

Intuitively, if the prior is a deterministically clean floor we expect the floor to be clean even after the vacuum cleaning action.

We can realize mathematically that this is indeed the case by noting that the minimum of (5) is achieved by $q = 1$. In other words if we take as prior:

$$p(x_0 = c) = 1 \qquad p(x_0 = d) = 0$$

we obtain the minimum possible value of $p(x_1 = d \mid z = c, u = vc) = 0$