6CCS3AIN & 7CCSMAIN, 2018, Tutorial 02 <small>(Version 1.1)</small>

1. From the joint distribution on page 38 of the slides, compute:

    (a) $P(cavity)$
    (b) $\mathbf{P}(Toothache)$.
    (c) $\mathbf{P}(Toothache|cavity)$.
    (d) $P(catch \vee cavity)$
    (e) $\mathbf{P}(Cavity|toothache \vee catch)$

    Recall what the difference is between P and $\mathbf{P}$ (see page 31 of the slides).

2. You take a test T to tell whether you have a disease D. The test comes back positive. You know that test is 95% accurate (the probability of testing positive when you do have the disease is 0.95, and the probability of testing negative when you don't have the disease is also 0.95). You also know that the disease is rare, only 1 person is 10,000 gets the disease.

    What is the probability that you have the disease?

    How would this change if the disease was more common, say affecting 1 person in 100?

3. Consider two tests, A and B, for a virus.

    Test A is 95% effective at recognizing the virus when it is present (that is 95% of the time that the virus is present, the test detects it), but has a 10% false positive rate (that is, 10% of the time it indicates the virus is present when it is not).

    Test B is 90% effective at recognizing the virus, but has a 5% false positive rate.

    The two test use different, independent, methods of indentifying the virus.

    1% of all people have the virus.

    Joe tests positive for the virus using test A. Bob tests positive using test B. Who is more likely to have the virus?

4. Look at the meningitis example on page 71 of the notes. Carry out the calculation without using the value for $P(s)$.

    This uses the trick of considering the denominator to be a normalization constant α.

    You'll need a value of $P(s|\neg m)$ to do this — use $P(s|\neg m) = 0.1$.

5. Now the optional computational part of the tutorial.

    On KEATS you can find the file `underwear.py`. This is the "wrong underwear" example from the notes coded using a Python package called `pomegranate`:

    https://pomegranate.readthedocs.io/en/latest/index.html

    `pomegranate` provides tools for probabilistic inference. You can install it using:

    conda install pomegranate

    if you have Anaconda installed[1]. Otherwise, see the webpage.

    With `pomegranate` installed, you can run the underwear example with:

    python underwear.py

    from the command line. As set up, this gives you the result for the first observation of the wrong underwear. (You can check this against the value in the slides.) To run the calculation for the second observation of the wrong underwear, you need to change the value of `prior_cheat`, the prior probability of cheating.

    What do you change it to? For the second observation, the prior is the value you just obtained for the first observation. Do this, and verify that you get the same result as in the slides.

    Now try building your own probability model for the Meningitis example.

---

[1]If you don't have Anaconda installed, you can install that from https://www.anaconda.com/download. If you are planning to do much with Python, Anaconda is pretty handy to have since it makes it easy to install new packages.