13.8 Given the full joint distribution shown in Figure 13.3, calculate the following:
a. **P**(toothache) .
b. **P**(Cavity) .
c. **P**(Toothache | cavity) .
d. **P**(Cavity | toothache ∨ catch) .
Answer:
The main point of this exercise is to understand the various notations of bold versus non-bold P, and uppercase versus lowercase variable names. The rest is easy, involving a small matter of addition.
a. This asks for the probability that *Toothache* is true.
P (toothache) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2

b. This asks for the vector of probability values for the random variable *Cavity*. It has two values, which we list in the order ⟨true, false⟩.
First add up 0.108 + 0.012 + 0.072 + 0.008 = 0.2.
Then we have
P (Cavity) = ⟨0.2,0.8⟩ .

c. This asks for the vector of probability values for *Toothache*, given that *Cavity* is true.
P (Toothache |cavity) = ⟨(.108 + .012)/0.2, (0.072 + 0.008)/0.2⟩ = ⟨0.6, 0.4⟩

d. This asks for the vector of probability values for *Cavity*, given that either *Toothache* or *Catch* is true. First compute
P(toothache ∨catch) = 0.108+0.012+0.016+0.064+ 0.072 + 0.144 = 0.416.
Then
P(Cavity|toothache ∨ catch) =⟨(0.108 + 0.012 + 0.072)/0.416, (0.016 + 0.064 + 0.144)/0.416⟩ = ⟨0.4615, 0.5384⟩

13.13 Consider two medical tests, A and B, for a virus. Test A is 95% effective at recognizing the virus when it is present, but has a 10% false positive rate (indicating that 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 tests use independent methods of identifying the virus. The virus is carried by 1% of all people. Say that a person is tested for the virus using only one of the tests, and that test comes back positive for carrying the virus. Which test returning positive is more indicative of someone really carrying the virus? Justify your answer mathematically.
Answer:
Let V be the statement that the patient has the virus, and A and B the statements that the medical tests A and B returned positive, respectively. The problem statement gives:
P(V) = 0.01
P(A|V) = 0.95
P(A|¬V) = 0.10
P(B|V) = 0.90
P(B|¬V) = 0.05
The test whose positive result is more indicative of the virus being present is the one whose posterior probability, P(V |A) or P(V |B) is largest. One can compute these probabilities directly

from the information given, finding that P(V |A) = 0.0876 and P(V |B) = 0.1538, so B is more indicative.

Equivalently, the questions is asking which test has the highest posterior odds ratio P(V |A)/P(¬V |A). From the odd form of Bayes theorem:

$$\frac{P(V|A)}{P(\neg V|A)} = \frac{P(A|V)}{P(A|\neg V)}\frac{P(V)}{P(\neg V)}$$

we see that the ordering is independent of the probability of V , and that we just need to compare the likelihood ratios P(A|V)/P(A|¬V) = 9.5 and P(B|V)/P(V |¬V) = 18 to find the answer.

13.15 After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease and that the test is 99% accurate (i.e., the probability of testing positive when you do have the disease is 0.99, as is the probability of testing negative when you don't have the disease). The good news is that this is a rare disease, striking only 1 in 10,000 people of your age. Why is it good news that the disease is rare?
What are the chances that you actually have the disease?
Answer:
We are given the following information:
P(test|disease) = 0.99
P(¬test|¬disease) = 0.99
P(disease) = 0.0001
and the observation test. What the patient is concerned about is P(disease|test). Roughly speaking, the reason it is a good thing that the disease is rare is that P(disease|test) is proportional to P(disease), so a lower prior for disease will mean a lower value for P(disease|test). Roughly speaking, if 10,000 people take the test, we expect 1 to actually have the disease, and most likely test positive, while the rest do not have the disease, but 1% of them (about 100 people) will test positive anyway, so P(disease|test) will be about 1 in 100. More precisely, using the normalization equation from page 480:

$$\begin{aligned} &P(disease|test) \\ &= \frac{P(test|disease)P(disease)}{P(test|disease)P(disease)+P(test|\neg disease)P(\neg disease)} \\ &= \frac{0.99\times 0.0001}{0.99\times 0.0001+0.01\times 0.9999} \\ &= .009804 \end{aligned}$$

The moral is that when the disease is much rarer than the test accuracy, a positive test result does not mean the disease is likely. A false positive reading remains much more likely.
Here is an alternative exercise along the same lines: A doctor says that an infant who predominantly turns the head to the right while lying on the back will be right-handed and one who turns to the left will be left-handed. Isabella predominantly turned her head to the left. Given that 90% of the population is right handed, what is Isabella's probability of being right-handed if the test is 90% accurate? If it is 80% accurate? The reasoning is the same, and the answer is 50% right-handed if the test is 90% accurate,
69% right-handed if the test is 80% accurate.

Additional Exercise:

Consider again the example application. Suppose the doctor decides to order a second laboratory test for the same patient, and suppose the second test returns a positive result as well. What are the probabilities of cancer and ~cancer following these two tests. Assume the two tests are independent.

• hMAP ? Based on P(canner|++) or P(¬ canner|++)

Answer:

Now we want to calculate the probabililty of cancer given two positive test results "++". We essentially perform the same calculation as before. The only difference is that we substitute the posterior which we calculated in problem one for the prior probabilities we used.
That is, we substitute P(cancer|+) and P(~cancer|+) for P(cancer) and P(~cancer) respectively.
This gives us:
P(cancer|++) = P(+|cancer) x P(cancer|+) /(P(+|cancer) x P(cancer|+) + P(+|~cancer) x P(~cancer|+))

We previously found that P(cancer|+) = .21. We also know that P(~cancer|+) = 1 - P(cancer|+) = .79. This gives us

P(cancer|++) = .98 x .21 / (.98 x .21 + .03 x .79)
= .21 / (.21 + .02)
= .91

So unfortunately, after a second positive test, we are much more certain that the patient has cancer. Sad news for the patient but good news for Bayes rule which is helping us to be sure are the patients really affected by cancer.That kind of information can help us to detect the problem and try to get treatment rather than staying in dark.