Introduction to Bayesian Statistics: Practicum 1

In 2012, the prevalence of HIV among the general U.S. population was estimated to be 0.38% (www.cdc.gov). Rapid diagnostic tests have been developed to test for the presence of HIV infection in as little as 10 minutes. Suppose one such diagnostic test was evaluated in a case-control study and it was observed that out of 1000 subjects with HIV, 967 obtained positive test results and out of 1000 subjects without HIV, 985 obtained negative test results.

1. From the information given above write down, using probability notation, each of the following items. Once you have written down the probability, calculate the value of each probability below. None of these requires much calculation:

1. The probability of having the disease in the general US population.
2. The probability of NOT having the disease in the general US population.
3. The sensitivity of the rapid test.
4. The Probability of a positive test result given the subject dos NOT have the disease.
5. The specificity of the rapid test.
6. The probability of a negative test result given the subject has the disease.

2. Draw a Venn diagram that illustrates the following two events:

1. Disease Status in the general US population (has two levels)
2. Test result of the rapid diagnostic test in the general US population (has two levels)

3. Although it is not directly given, let’s try to find the probability of a positive test result (i.e. P(T+)).

1. First, notice from the Venn diagram in question 2 that the event of a positive test can be split into two mutually exclusive pieces; 1. one that is the intersection of disease positive and test positive and 2. one that is the intersection of disease negative and test positive. (Hint: the probability of a positive test will then be the sum of these two parts.)
2. Next simply write down the two intersections above in probability notation.
3. Recall the formula for the intersection and write what the above intersections are equal to in terms of quantities given in question 1.
4. Finally, compute these quantities and sum them to get the probability of a positive test.

BONUS!!! 4. Now that you’ve done the hard part, let’s see if we can put this together with what we’ve learned about conditional probabilities to compute the positive predictive value of the rapid test in this population (i.e. P(D+|T+)). Can you also compute the negative predictive value of the rapid test (i.e. P(D-|T-))?