Note: This homework assignment covers Chapter 2.

1. A local television weatherman announces, "There is a 30 percent chance of rain tomorrow." What is your interpretation of this statement?

(i) It will rain tomorrow for 30 percent of the time. That is, for 7.2 hours tomorrow, it will be raining. For the remaining 16.8 hours, it will not be raining.
(ii) It will rain tomorrow in 30 percent of the region covered by the local television station. It will not rain in the other 70 percent of the region.
(iii) Among all local meteorologists, 30 percent of them think that it will rain tomorrow. The remaining 70 percent of the meteorologists think that it will not rain tomorrow.
(iv) Thirty percent of all inhabitants of the region covered by this local television station will see rain at least once during their day tomorrow; the remaining 70 percent will not see rain during their day.
(v) It will rain on 30 percent of the days in which this same forecast is made.

I think (v) is the "best" interpretation, but in reality all 5 interpretations are valid depending on how you conceptualize the underlying sample space for such a probability assignment; i.e., $P(A) = 0.30$, where A is the event it rains. For each possible interpretation, describe a sample space that would make each interpretation valid.

2. Suppose that S is a sample space and that A and B are events in S with $P(A) = 0.6$, $P(B) = 0.5$, and $P(A \cap B) = 0.2$. Compute each of the following probabilities. In each part, state clearly what probability law(s) you are using.
(a) $P(\overline{A})$
(b) $P(A \cup B)$
(c) $P(\overline{A} \cap B)$ *Hint:* Note that $B = (A \cap B) \cup (\overline{A} \cap B)$. Draw Venn Diagrams to convince yourself this is true.
(d) $P(\overline{A} \cup \overline{B})$
(e) $P(B|A)$
(f) Are A and B independent? Why or why not?

Remark: The following are called *DeMorgan's Laws:*

$$\begin{aligned} P(\overline{A} \cup \overline{B}) &= P(\overline{A \cap B}) = 1 - P(A \cap B) \\ P(\overline{A} \cap \overline{B}) &= P(\overline{A \cup B}) = 1 - P(A \cup B). \end{aligned}$$

Draw Venn Diagrams to convince yourself that each law is correct.

3. The use of plant appearance in prospecting for ore deposits is called geobotanical prospecting. One indicator of copper is a small mint with a mauve-colored flower. Suppose that, for a certain region, there is a 30 percent chance that the soil has a high copper content and a 23 percent chance that the mint will be present there. In addition, we know that if the copper content is high, there is a 70 percent chance that the mint will be present. Let A denote the event that a soil sample has high copper content, and let B denote the event that the mint is present.
(a) Find the probability that the mint will not be present.
(b) Find the probability that the copper content will be high and the mint will be present.
(c) Find the probability that the copper content will be high given that the mint is present.
(d) Are the events A and B independent? Explain.

4. A plane is missing and is presumed to have crashed in one of two regions: R_1 or R_2 (no other regions are possible). Field experts from the NTSB have projected that

- there is a 30 percent chance that the plane crashed in Region 1
- there is a 70 percent chance that the plane crashed in Region 2.

A search party will be successful at finding the crashed plane with probability 0.8 for Region 1 and 0.4 for Region 2.
(a) What is the probability that the crashed plane will be found in Region 2?
(b) What is the probability that the crashed plane will not be found?
(c) What is the probability that the plane crashed in Region 1 given that the search was successful?

5. Reliability engineers often work with systems having components connected in parallel. In this problem, we will interpret the phrase "in parallel" as follows: The system is reliable (i.e., it is functioning) if **at least one** of the components is functioning. As a frame of reference, consider a two-engine aircraft.

- If both engines are functioning, the aircraft is functioning (at least in lieu of non-engine related problems).
- If only one engine is functioning, the aircraft still functions (although losing one engine would warrant an immediate landing).
- If both engines are not functioning, the aircraft is not functioning.

In this problem, we will denote by n the number of components in a parallel system.

(a) Suppose $n = 2$. If the two components are functioning **independently**, each with probability p, show that the system reliability r_2 is given by

$$r_2 = 1 - (1-p)^2.$$

(b) Generalize the result in part (a) to consider a parallel system with n components (each functioning independently with probability p). That is, show that the system reliability is

$$r_n = 1 - (1-p)^n.$$

(c) I have a parallel system with $n = 4$ components (each functioning independently with probability p). How unreliable can the individual components be and still have a system with reliability $r_4 = 0.999$?

(d) So far in this problem, we have made two critical assumptions:

(A1) the components function independently

(A2) the components each function with the same probability p.

Give a real-life example where Assumption (A1) is likely violated. Give a real-life example where Assumption (A2) is likely violated. Do not use the plane example I used at the outset. Explain your examples sufficiently so I can understand.

6. *Streptococcal pharyngitis*, which is more commonly known as "strep throat," is a contagious disease caused by Group A streptococcal infection. Most of you have likely been tested for strep throat at some point in your lives (usually the doctor will swab your throat with a Q-tip). The most common type of test to detect strep throat is a "rapid test." The rapid strep test works by taking your Q-tip sample and detecting in it the presence of a carbohydrate antigen unique to group A *streptococcus*. This is a fairly simple procedure, and, in fact, one advantage of this particular test is that the results are usually available in minutes. Therefore, you can get your diagnosis from your doctor while you wait in his/her office. Although simple and fast, rapid tests are not perfect. A particular rapid test has sensitivity 0.95 and specificity 0.97. That is, among all patients who are strep-positive, the rapid test will test positively 95 percent of time (i.e., 5 percent will be "false negatives"). Similarly, among all patients who are strep-negative, the rapid test will test negatively 97 percent of time (i.e., 3 percent will be "false positives").

(a) Suppose the probability an individual is strep-positive is p (the population-level prevalence of the disease). If an individual's test result is positive, use Bayes' Rule to show that the probability the individual is strep-positive is equal to

$$\text{PPV}(p) = \frac{0.95p}{0.95p + 0.03(1-p)}.$$

Similarly, if an individual's test result is negative, show that the probability the individual is strep-negative is equal to

$$\text{NPV}(p) = \frac{0.97(1-p)}{0.97(1-p) + 0.05p}.$$

Note: The symbols PPV and NPV stand for "positive predictive value" and "negative predictive value," respectively. These are commonly-used terms in medical epidemiology.
(b) Graph both PPV(p) and NPV(p) as functions of p over $0 < p < 1$ (use R or Maple or whatever). You should see that PPV(p) is an increasing function of p while NPV(p) is a decreasing function of p. Explain why this is true.
(c) Suppose that $p = 0.10$. Calculate PPV(p) and NPV(p) for this value of p and interpret (in words) what these probabilities mean. If you tested positively, would you feel "convinced" that you were actually strep-positive?

After you leave the doctor's office and while you are on your way to the pharmacist to get antibiotics (if your rapid test was positive), your doctor might direct your Q-tip swab to be subjected to a more powerful testing procedure, known as a throat-culture test. For this test, your Q-tip sample is put in a culture that allows infections to grow. If the infection grows, the culture is positive. If no infection grows, your culture is negative. Culture tests can be much more accurate than rapid tests. In fact, most throat-culture tests have sensitivity and specificity exceeding 0.99. The problem is they can take much longer to perform (1-2 days while waiting for growth). This is obviously why rapid tests are given first; you are not going to wait in the doctor's office for 2 days to get your results.

In parts (d) and (e), continue to assume that $p = 0.10$. Suppose a throat-culture test has sensitivity and specificity both equal to 0.999.

(d) Suppose that your rapid test was positive, so that your (updated) probability of being strep-positive is PPV(0.10). Now, if your follow-up throat-culture test is also positive, give me a new update on the probability that you are strep-positive.
(e) Suppose that your rapid test was negative, so that your (updated) probability of being strep-negative is NPV(0.10). Now, if your follow-up throat-culture test is also negative, give me a new update on the probability that you are strep-negative.