**ECE 3077, Summer 2014**

**Homework #1**

**Due Thursday May 22, in class**

**Reading: B&T 1.1–1.5**

1. Using you class notes, prepare a 1-2 paragraph summary of what we talked about in class in the last week. I do not want just a bulleted list of topics, I want you to use complete sentences and establish context (Why is what we have learned relevant? How does it connect with other things you have learned here or in other classes?). The more insight you give, the better.

2. The theory of probability relies only on Kolmogorov's three axioms: a probability law $P$ on a sample space $\Omega$ must obey

   (i) $P(A) \geq 0$ for any event $A$ (i.e., for any subset $A$ of the possible outcomes $\Omega$).

   (ii) If $A$ and $B$ are disjoint events, meaning $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.

   (iii) $P(\Omega) = 1$.

   From these, we can derive any number of properties of $P$. Here are four examples (some of which are also given in the notes):

   (iv) If $A \subset B$, then $P(A) \leq P(B)$.

   (v) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ for all events $A$ and $B$.

   (vi) $P(A \cup B) \leq P(A) + P(B)$ for all events $A$ and $B$.

   (vii) $P(A^c) = 1 - P(A)$.

   For example, to establish (iv) we can write $B$ as[1]

   $$B = (B \cap A) \cup (B \cap A^c) = A \cup (B \cap A^c)$$

   where the last equality follows from the fact that $A \subset B$. Then using the axioms

   $$\begin{aligned} P(B) = P(A \cup (B \cap A^c)) &= P(A) + P(B \cap A^c) \qquad &&\text{(since } A \text{ and } B \cap A^c \text{ are disjoint)} \\ &\geq P(A) &&\text{(since } P(B \cap A^c) \geq 0\text{).} \end{aligned}$$

   (a) Derive (v), (vi), and (vii) from Kolmogorov's axioms.

   (b) Let $A$ and $B$ be arbitrary events. Suppose we know that $P(A \cup B) = 2/3$, $P(A \cap B) = 1/12$, and $P(B \cap A^c) = 1/3$. What is $P(A \cap B^c)$?

   (c) Out of the students in this class, 60% can attend my office hours if I hold them on Monday, 75% can attend them on Tuesday, and 40% can attend either day. What is the probability that a randomly selected will not be able to attend my office hours?

[1] If you are having trouble visualizing this, draw a Venn diagram with $A$ as a subset of $B$, then clearly label the region corresponding to $A$ and $B \cap A^c$.

(d) There is a 50% chance it rains on Saturday, a 60% chance it rains on Sunday, and a 70% chance is rains on Saturday OR Sunday. What is the percent chance it rains on Saturday AND Sunday?

(e) True or False: the following statements are consistent:

- the chance it rains on Saturday is 50%
- the chance it rains on Sunday is 70%
- the chance it rains on both Saturday AND Sunday is 15%

Justify your answer.

3. The pieces of candy you are eating come in three flavors: orange, grape, and cherry. When you reach into the bag, the probability you pull out an orange or grape piece is 0.7, while the probability you pull out a grape or cherry piece is 0.9. What is the probability that you pull out an orange or cherry piece?

4. Suppose you are competing in a strange tournament in which you will play three games against three different opponents. You win a prize only if you win two consecutive games. You calculate your chance of beating opponent $A$ as 50%, opponent $B$ as 75%, and opponent $C$ as 10%. If you could pick the order in which you play the three opponents, what order would you choose? What is your probability of winning the prize in this case?

5. Two presidential candidates, generically labeled $D$ (for Democrat) and $R$ (for Republican), are in a heated election battle. A U.S. presidential election can be broken down into 50 separate subelections, one for each state, and the winner of each subelection is awarded a certain number of points (i.e. electoral votes, the number of which can vary from contest to contest). There are 538 total points available, and so accumulating 270 or more will win the election. We will assume[2] that the outcome of a subselection in one state does not affect the outcome of the others (i.e., the subelections in each state are independent).

   Suppose that 46 of the 50 contests have already been more or less decided, and the score is $D$ 247, $R$ 227. Four contests remain for the final 64 points. Here is a breakdown which includes how many points each of the remaining contests is worth, along with the probability that each player wins that contest[3]:

| state | points | $P(D \text{ wins})$ | $P(R \text{ wins})$ |
|---|---|---|---|
| FL | 29 | 0.45 | 0.55 |
| OH | 18 | 0.66 | 0.34 |
| WI | 10 | 0.77 | 0.23 |
| CO | 9 | 0.62 | 0.38 |

   (a) There are $2^4 = 16$ different win/loss combinations for these last 4 contests. Which ones result in $R$ begin elected?

[2]This assumption generally does not hold in practice, but nonetheless, these results can be instructive.
[3]This is actually not too far from what the 2012 election looked like in August 2012.

(b) What is the probability that $R$ is elected?

6. An actual article in the *Los Angeles Times* from August 24, 1987, discussing the statistical risks of an AIDS infection, observed that the transmission rates for AIDS were actually estimated to be surprisingly low. The reporter quotes a study listing the risk of contracting the virus after a single act of intercourse with an infected partner being only about one in 500. The reporter then went on to reason:

   > Statistically, 500 acts of intercourse with one infected partner or 100 acts with five partners lead to a 100% probability of infection (statistically, not necessarily in reality).

   Following this line of reasoning, 1000 acts of intercourse would lead to a 200% probability of infection ("statistically").

   (a) Explain the flaw in the reporter's reasoning.

   (b) Assuming the risk is correctly estimated at one in 500, what would be the actual probability of infection for a person who had 500 acts of intercourse with an infected partner?

7. In MATLAB, the command `rand` generates a single random number on $[0,1]$ according to the uniform law. Try it. The command `rand(m,1)` generates a column vector of length $m$ where each entry is generated independently according to the uniform law. Try it with $m = 37$.

   If we set `x= (rand < .32)`, the variable $x$ will be either 0 (if the random number is greater than 0.32) or 1 (if the random number is less than .32). It shouldn't be hard to see that $x = 1$ with probability 0.32 and $x = 0$ with probability 0.68. Similarly, `x=(rand(37,1) > .32)` will create a vector of 37 variables, each of which is 1 or 0 with probability 0.32 or 0.68, respectively. It should be clear how to generalize this to vector lengths other than 37 and probabilities other than 0.32

   So if I tell you that basketball player Dwight is a 45% free throw shooter, you can simulate a sequence of 100 free throws with

```
FT = (rand(100,1) < .45);
```

   The entry `FT(k)` will be 1 if Dwight made free throw number $k$, and 0 if he missed. The result of `sum(FT)` will tell you the total number of free throws out of the 100 that Dwight made.

   (a) Write a MATLAB function `free_throws.m` that takes a probability $p$ and a positive integer $N$ and returns a simulated sequence of free throws of length $N$, with a probability of success $p$.

   (b) Use your function to estimate the probability that Dwight makes at least 5 of his next 10 free throws. This will require you running it many times (say at least $10,000$) and collecting the results in a smart way.

(c) Use your script to estimate the probability that Dwight makes at least 50 of his next 100 free throws.

(d) Use your script to estimate the probability that Dwight makes at least 500 of his next 1000 free throws.

Turn in all of your code, along with your answers to (b)–(d).