Homework 1 8/29 - 9/2

STA 335 Due 11:59pm Monday, September 5 Name:

Instructions: Write-up complete solutions to the following problems and submit answers on Gradescope. Your solutions should be neatly-written, show all work and computations, include figures or graphs where appropriate, and include some written explanation of your method or process (enough that I can understand your reasoning without having to guess or make assumptions). A rubric for homework problems appears on the final page of this assignment.

- Unless otherwise noted, problem numbers are taken from the 2nd edition of Blitzstein and Hwang's *Intro to Probability*.

Monday 8/29

Chapter 1

26, 27, 38

Wednesday 8/31

Chapter 1

45, 60

Additional Problems

AP1. Suppose Oliver has a belief system assigning a number $P_o(A)$ between 0 and 1 to every event $A \subset S$ for some sample space S. This represents Oliver's degree of belief about how likely A is to occur. For every event A, Oliver is willing to pay $P_o(A)$ dollars to buy from you a certificate that says

> "The owner of this certificate can redeem it from the seller for \$1 if A occurs, and for \$0 if A does not occur."

Likewise, Oliver is willing to sell such a certificate at the same price. In fact, he will buy/sell any number of certificates for this price.

However, as a four year old, Oliver stubbornly refused to accept the axioms of probability. In particular, there are a particular pair of disjoint events A and B with

$$P_o(A \cup B) \neq P_o(A) + P_o(B).$$

Suppose Oliver was gifted \$10 on his birthday. Show that there is a sequence of transactions that Oliver is willing to make that will **guarantee** he will lose all \$10. (You can assume that after all certificates have been bought and sold, we know for certain whether event A and whether event B occurred).

Friday 9/2

Additional Problems

AP2. We will revisit the following famous problem due to de Montmort several times throughout the term:

> Consider a well-shuffled deck of n cards, labeled 1 through n. You flip over the cards one by one, saying the numbers 1 through n as you do so. You win the game if, at some point, the number you say aloud is the same as the number on the card being flipped over (for example, if the 7th card in the deck has the label 7). For large n, the probability of winning is approximately $1 - e^{-1}$.

(a) For each of $n = 5, 10, 50, 100$, write a program in R that uses the `sample` function to play one iteration of the game and output whether the game results in a win or a loss. Then use the `replicate` function to simulate a large number of trials of the game in order to approximate the probability of winning. Verify numerically that the probability of winning indeed approaches $1 - e^{-1}$ as n gets larger.

(b) Modify the program you designed in the previous part to instead output the number of matches in one iteration of the game (again, for each of $n = 5, 10, 50, 100$).

(c) Recall that the **mean** of a data set is the arithmetic average of the values in the set, while the **standard deviation** measures the variability of values in the set. Use `replicate` to simulate a large number of games using the program from the previous part. Then use the `mean` and `sd` functions to compute the mean and standard deviation of the number of matches in a large number of trials. What do you observe about the mean and standard deviation as n gets larger?

AP3. Suppose each of 10 balls is independently placed into one of 10 boxes, with all boxes equally likely. Write a program in R to estimate the probability that **exactly** one box is empty.

Hint: it may be useful to use the `unique` function, which returns the subset of unique entries of a vector, along with the `length` function, which returns the length of a vector.

General Rubric

Points	Criteria
5	The solution is correct **and** well-written. The author leaves no doubt as to why the solution is valid.
4.5	The solution is well-written, and is correct except for some minor arithmetic or calculation mistake.
4	The solution is technically correct, but author has omitted some key justification for why the solution is valid. Alternatively, the solution is well-written, but is missing a small, but essential component.
3	The solution is well-written, but either overlooks a significant component of the problem or makes a significant mistake. Alternatively, in a multi-part problem, a majority of the solutions are correct and well-written, but one part is missing or is significantly incorrect
2	The solution is either correct but not adequately written, or it is adequately written but overlooks a significant component of the problem or makes a significant mistake.
1	The solution is rudimentary, but contains some relevant ideas. Alternatively, the solution briefly indicates the correct answer, but provides no further justification
0	Either the solution is missing entirely, or the author makes no non-trivial progress toward a solution (i.e. just writes the statement of the problem and/or restates given information)
Notes:	For problems with multiple parts, the score represents a holistic review of the entire problem. Additionally, half-points may be used if the solution falls between two point values above.